



ใบรับรองวิทยานิพนธ์  
บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตร์มหาบัณฑิต (วิทยาการคอมพิวเตอร์)

ปริญญา

วิทยาการคอมพิวเตอร์

วิทยาการคอมพิวเตอร์

สาขา

ภาควิชา

เรื่อง แบบจำลองเพื่อใช้สำหรับการเฝ้าระวังโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีกของไทย  
โดยเทคนิคการเหมืองข้อมูล

Model for Avian Influenza Surveillance of Poultry Husbandary in Thailand Using Data Mining

นามผู้วิจัย นายสิรพจน์ สันตติวุฒิ

ได้พิจารณาเห็นชอบโดย

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

( รองศาสตราจารย์อนงค์นาฏ ศรีวิหค, Ph.D. )

หัวหน้าภาควิชา

( ผู้ช่วยศาสตราจารย์ศิริกร จันทร์นวล, M.S. )

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

( รองศาสตราจารย์กัญญา ชีระกุล, D.Agr. )

คณบดีบัณฑิตวิทยาลัย

วันที่ ..... เดือน ..... พ.ศ. ....

วิทยานิพนธ์

เรื่อง

แบบจำลองเพื่อใช้ในการเฝ้าระวังโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีกของไทย โดยเทคนิคการ  
เหมืองข้อมูล

Model for Avian Influenza Surveillance of Poultry Husbandary in Thailand Using Data Mining

โดย

นายสิรพจน์ สันตติวุฒิ

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์  
เพื่อความสมบูรณ์แห่งปริญญาวิทยาศาสตรมหาบัณฑิต(วิทยาการคอมพิวเตอร์)

พ.ศ. 2554

สิริพจน์ สันตจิตฺติ 2554: แบบจำลองเพื่อใช้สำหรับการเฝ้าระวังโรคไข้หวัดนกในพื้นที่  
เลี้ยงสัตว์ปีกของไทยโดยเทคนิคการทำเหมืองข้อมูล ปรินญาวิทยาสตรมหาบัณฑิต  
(วิทยาการคอมพิวเตอร์) สาขาวิทยาการคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์ อาจารย์  
ที่ปรึกษาวิทยานิพนธ์หลัก: รองศาสตราจารย์อนงค์นาถ ศรีวิหค, Ph.D. 218 หน้า

ในปัจจุบันโรคไข้หวัดนกเป็นปัญหาที่ส่งผลกระทบเป็นวงกว้างในหลายด้านได้แก่คุณภาพชีวิต ด้าน  
การปศุสัตว์ และ ด้านเศรษฐกิจ ด้วยผลกระทบที่เกิดขึ้นในหลายด้าน ทำให้นักวิจัยมากมายได้คิดค้นวิธีการ หรือ  
เครื่องมือ ที่ทำให้ทราบถึงปัจจัยหรือตัวกระตุ้นที่ทำให้เกิดการระบาดของโรคไข้หวัดนก ซึ่งทำให้สามารถ  
ป้องกันหรือลดความรุนแรงของการระบาดของโรคไข้หวัดนกในประเทศไทยลงได้ ดังนั้นงานวิจัยนี้มี  
วัตถุประสงค์ที่จะนำเสนอแบบจำลองของการทำนายลักษณะพื้นที่การเลี้ยงสัตว์ปีกของไทยที่มีโอกาสที่จะเกิด  
การระบาดของโรคไข้หวัดนก โดยใช้วิธีการทำเหมืองข้อมูล และ วัดประสิทธิภาพของแบบจำลองที่สร้างขึ้น  
ด้วยเทคนิคการทำเหมืองข้อมูลที่แตกต่างกัน เพื่อใช้เป็นข้อมูลในการเลือกแบบจำลองที่เหมาะสมและมี  
ประสิทธิภาพดีที่สุดมาใช้ในการเฝ้าระวังโรคไข้หวัดในพื้นที่เลี้ยงสัตว์ปีกของไทย

ข้อมูลที่ใช้ในงานวิจัยนี้ได้แก่ รายงานสัตว์ปีกป่วยตายตั้งแต่ 2547-2550 จำนวนข้อมูลทั้งสิ้น 8,152  
ระเบียน ประกอบด้วย พื้นที่ ที่มีสัตว์ปีกป่วยตายด้วยโรคไข้หวัดนกจำนวน 1,351 ระเบียน และพื้นที่ ที่มีสัตว์  
ปีกป่วยตายแต่ไม่มีสาเหตุจากโรคไข้หวัดนก จำนวน 6,801 ระเบียน โดยขั้นตอนในการวิจัยเริ่มจาก การนำ  
ข้อมูลจากรายงานสัตว์ปีกป่วยตาย มาผ่านขั้นตอนแบ่งช่วงข้อมูล (Discretize) คัดเลือกคุณลักษณะ (Feature  
Selection) และ ทำการถ่วงน้ำหนักให้กับชุดข้อมูล (Feature Weighting) โดยใช้วิธี (1) DECISION MAKING  
TRIAL And EVALUATION LABORATORY (DEMATEL) และ (2) Emerging Pattern เพื่อศึกษา  
ประสิทธิภาพในการทำนายข้อมูล ของแบบจำลองที่สร้างขึ้น โดยใช้เทคนิค 3 แบบ คือ 1) Decision Tree (J48)  
2) Decision Tree(Naive Bayes Tree) และ 3) ข่ายงานแบบเบย์ส์ (Bayesian Network) ต่อมาดำเนินการวัดผล  
และ ทำการคัดเลือก แบบจำลองที่มีประสิทธิภาพด้านการแนกข้อมูลที่ดีที่สุด จากผลการวิจัยพบว่าแบบจำลอง  
ที่มีประสิทธิภาพดีที่สุดคือ แบบจำลองที่สร้างจาก Bayesian Network ซึ่งสร้างจากชุดข้อมูลที่ ดำเนินการแบ่ง  
ช่วงข้อมูล (Discretize) ที่เป็นตัวเลข ดำเนินการคัดเลือก คุณลักษณะ และ ดำเนินการถ่วงน้ำหนักข้อมูลด้วย  
น้ำหนักที่หามาจากวิธีการ Emerging Pattern

ซึ่งจากแบบจำลองทำให้ทราบว่าคุณลักษณะ ที่ส่งผลต่อการระบาดของโรคไข้หวัดนกประกอบด้วย  
เดือนที่สัตว์ปีกป่วยตาย (month) ความชื้นของอากาศ (humidity) จำนวนชั่วโมงที่มีแสงแดด (day light hours)  
อุณหภูมิ (tempurature) ความสูงจากระดับน้ำทะเลของพื้นที่เลี้ยงสัตว์ปีก (height) ลักษณะของพื้นที่เลี้ยงสัตว์ปีก  
(land scape) จำนวนสัตว์ปีกที่ตาย (death) และ จำนวนสัตว์ปีกที่ถูกทำลาย (destruction) สามารถนำความรู้ที่ได้  
จากการศึกษาครั้งนี้ไปสร้างกฎเพื่อใช้ในการเฝ้าระวังโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีกของไทยต่อไป

Sirapote Santatiwut 2011: Model for Avian Influenza Surveillance of Poultry Husbandary in Thailand Using Data Mining. Master of Science (Computer Science), Major Field: Computer Science, Department of Computer Science. Thesis Advisor: Associate Professor Anongnart Srivihok, Ph.D. 218 pages.

Avian Influenza namely, Bird Flu, is a virulent virus which destroys poultries severely and some harmful to human. It has threatened Thailand economics since there are many Thai people who are live stock farmers and earn their living by poultry productions. In the past researchers proposed the models, procedures and techniques to predict and prevent Avian Influenza epidemic. This present study proposes a classification model for predicting the characteristics of poultry farms in Thailand that having high risks of Avian Influenza epidemic. This proposed model is built by Data Mining Techniques.

In this research, data were undertaken from reports of death and ill poultries from the Department of Livestocks, Ministry of Agriculture in period 2004 - 2007 and Weather Information Metheology Department of the same periods. There are totally 8,152 records of death and ill poultries. Data set included 1) Death and ill poultry from Avian Influenza was 1,351 records and 2) death and ill poultry from other reasons was 6,801 records. The experiments included three steps. First step, data from the above two sources were combined and preprocessed. Then, discretization, attribute selection and weighting data by two different methods were applied. The first weighting method was Decision Making Trial and Evaluation Laboratory (DEMATEL) and the second method, instance weights were computed from Emerging Pattern (EP) algorithm. Second step, classification models were generated by using Data Mining techniques: Bayesian Network, Decision Tree and Naive Bayes Tree. Classification Performances of each Model were evaluated by using accuracy rate, precision value recall value and F-measure value.

Results of this study showed that the model built from Bayesian Network using discretized data set, selected attributes, and weighting data computed from Emerging Pattern had the highest performance with 95.98 % accuracy, and precision (0.943) and Recall (0.939). Further, factors which contribute to the epidermic of Bird Flu in poultry farms were proposed. They included (1) month of death, (2) humidity, (3) day light hours, (4) Temperature, (5) height, (6) landscape, (7) number of dead poultries and (8) number of destroyed poultry. Lastly, the production rules which concerning the Bird Flu epidermics were presented.

---

Student's signature

---

Thesis Advisor's signature

## กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จลุล่วงได้ด้วยความสามารถ และ ความช่วยเหลืออย่างดียิ่งจาก รองศาสตราจารย์ ดร.อนงค์นาฏ ศรีวิหค ที่ได้ชี้แนะแนวทางรวมทั้งให้คำแนะนำ วิธีการและ ขั้นตอนในการดำเนินงาน และแนวทางในการแก้ไขปัญหา รวมทั้งตรวจสอบแก้ไขข้อบกพร่องที่เกิดขึ้นในงานวิจัยชิ้นนี้มาโดยตลอด และ ขอกราบขอบพระคุณคณะกรรมการในการสอบ ที่ให้คำแนะนำและแนวทางในการทำงานวิจัยนี้ให้มีความสมบูรณ์และถูกต้องมากยิ่งขึ้น และคณาจารย์ทุกท่านที่กรุณาให้คำแนะนำ และข้อเสนอแนะในการดำเนินการจัดทำวิทยานิพนธ์นี้ ผู้จัดทำรู้สึกซาบซึ้งและขอขอบพระคุณเป็นอย่างสูง

ขอกราบขอบพระคุณ คุณแม่ คุณพ่อ ที่ได้ให้กำลังใจ และ แนะนำแนวทางในการทำวิจัย และ ขอขอบคุณเพื่อนๆ พี่ๆ ที่อยู่ บริษัท เอเชียน พร็อพเพอร์ตี้ และ เพื่อนๆ พี่ๆ ที่ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเกษตรศาสตร์ ที่ได้ให้ความช่วยเหลือ คำแนะนำต่างๆ และ กำลังใจต่างๆ ที่ช่วยให้มาถึงตรงจุดนี้ได้

ขอขอบพระคุณ กรมปศุสัตว์ ที่ช่วยแนะนำแหล่งข้อมูลที่น่ามาใช้ในการวิจัยในครั้งนี้ และ ขอขอบพระคุณกรมอุตุนิยมวิทยา ที่ได้แนะนำแหล่งข้อมูลสภาพอากาศที่น่ามาใช้ในการวิจัยนี้

ขอขอบพระคุณ โครงการปริญญาโทภาคปกติ ที่ให้การสนับสนุนอุปกรณ์และสถานที่ ตลอดระยะเวลาในการทำวิทยานิพนธ์ คุณค่าและประโยชน์อันพึงมีจากวิทยานิพนธ์นี้ ขอมอบแด่ผู้มีพระคุณทุกท่าน ด้วยความเคารพยิ่ง

สิรพจน์ สันตติวุฒิ

มิถุนายน 2554

## สารบัญ

## หน้า

สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(12)
คำอธิบายสัญลักษณ์ และ คำย่อ	(14)
คำนำ	1
วัตถุประสงค์	3
การตรวจเอกสาร	4
อุปกรณ์และวิธีการ	44
อุปกรณ์	44
วิธีการ	50
ผลและวิจารณ์	74
ผล	74
วิจารณ์	186
สรุปและข้อเสนอแนะ	189
สรุป	189
ข้อเสนอแนะ	191
เอกสารและสิ่งอ้างอิง	192
ภาคผนวก	195
ประวัติการศึกษา และการทำงาน	218

## สารบัญตาราง

ตารางที่		หน้า
1	จำนวนข้อมูลที่อยู่ภายใต้กิ่งของต้นไม้ที่จำแนกด้วย Naive Bayes แยกตามสมาชิกใน Attribute Humidity และ ผลเฉลยของข้อมูลว่าพื้นที่เลี้ยงสัตว์ปีกดังกล่าวเกิดโรคไข้หวัดนกหรือไม่	18
2	จำนวนข้อมูลที่อยู่ภายใต้กิ่งของต้นไม้ที่จำแนกด้วย Naive Bayes แยกตามสมาชิกใน Attribute Landscape และ ผลเฉลยของข้อมูลว่าพื้นที่เลี้ยงสัตว์ปีกดังกล่าวเกิดโรคไข้หวัดนกหรือไม่	18
3	จำนวนข้อมูลที่อยู่ภายใต้กิ่งของต้นไม้ที่จำแนกด้วย Naive Bayes แยกตามสมาชิกใน Attribute Height และ ผลเฉลยของข้อมูลว่าพื้นที่เลี้ยงสัตว์ปีกดังกล่าวเกิดโรคไข้หวัดนก	19
4	จำนวนข้อมูลที่อยู่ภายใต้กิ่งของต้นไม้ที่จำแนกด้วย Naive Bayes แยกตามสมาชิกใน Attribute Temperature และ ผลเฉลยของข้อมูลว่าพื้นที่เลี้ยงสัตว์ปีกดังกล่าวเกิดโรคไข้หวัดนก	19
5	จำนวนข้อมูลที่อยู่ภายใต้กิ่งของต้นไม้ที่จำแนกด้วย Naive Bayes แยกตามสมาชิกใน Attribute Death และ ผลเฉลยของข้อมูลว่าพื้นที่เลี้ยงสัตว์ปีกดังกล่าวเกิดโรคไข้หวัดนก	20
6	เปรียบเทียบงานวิจัยที่เกี่ยวกับการสร้างแบบจำลองของโรคไข้หวัดใหญ่ ปอดอักเสบ และ โรคไข้หวัดนก	39
7	ตัวอย่างข้อมูลที่ได้จากรายงานสัตว์ปีก ป่วยตาย จากกรมปศุสัตว์	45
8	รายละเอียดของตัวแปรที่ใช้ในการวิจัย	49
9	ตัวอย่างข้อมูลที่ใช้ในการทดลองที่รวบรวมจากข้อมูลสัตว์ปีกป่วยตาย ข้อมูลสภาพภูมิอากาศ และ ข้อมูลลักษณะพื้นที่	55
10	คุณลักษณะของข้อมูลที่ถูกลำมาแบ่งช่วงข้อมูลเนื่องจากเป็นคุณลักษณะที่มีชนิดข้อมูลเป็นตัวเลข	56
11	ความหมายของสมาชิกที่อยู่ในคุณลักษณะ ภูมิประเทศ (Land Scape)	57
12	ลักษณะข้อมูลที่ผ่านขั้นตอนการแบ่งช่วงข้อมูล (Discretization)	58

## สารบัญตาราง (ต่อ)

ตารางที่		หน้า
13	คุณลักษณะที่ผ่านการคัดเลือก คุณลักษณะข้อมูล โดยวิธี ClassifierSubsetEvaluator	59
14	ตัวอย่างชุดข้อมูลที่ผ่านวิธีการคัดเลือกคุณลักษณะ	60
15	ลักษณะข้อมูลที่ผ่านการถ่วงน้ำหนักด้วยน้ำหนักที่หามาด้วยวิธีการ DEMATEL	63
16	ตัวอย่าง Item set ที่นำไปใช้ในการคำนวณน้ำหนักของข้อมูล	64
17	ลักษณะของชุดข้อมูลที่นำมาใช้ในงานวิจัย	65
18	ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบ ประสิทธิภาพของแบบจำลองระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่ได้ ผ่านขั้นตอนการเตรียมข้อมูล กับแบบจำลองที่สร้างจากชุดข้อมูลที่ผ่านการ คัดเลือกคุณลักษณะ (Attribute Selection)	66
19	ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบ ประสิทธิภาพของแบบจำลองระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่ได้ ผ่านขั้นตอนการเตรียมข้อมูล กับแบบจำลองที่สร้างจากชุดข้อมูลที่ผ่านการ แบ่งช่วงข้อมูล (Discretization)	67
20	ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบ ประสิทธิภาพของแบบจำลองระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ผ่าน ขั้นตอนการคัดเลือกคุณลักษณะ และ ไม่แบ่งช่วงข้อมูล และไม่ ถ่วงน้ำหนัก ด้วยน้ำหนักที่หามาจาก DEMATEL กับแบบจำลองที่สร้างจากชุดข้อมูลที่ผ่าน ขั้นตอนการคัดเลือกคุณลักษณะ กับขั้นตอนการถ่วงน้ำหนักด้วยน้ำหนักที่หา มาจาก DEMATEL	68
21	ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบ ประสิทธิภาพของแบบจำลองระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่ได้ ผ่านการแบ่งช่วง ไม่ได้ผ่านการคัดเลือกคุณลักษณะ และ ไม่ได้ทำการถ่วง น้ำหนัก กับแบบจำลองที่ผ่านขั้นตอนการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจาก DEMATEL	69

## สารบัญตาราง (ต่อ)

ตารางที่		หน้า
22	ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองที่สร้างจากชุดข้อมูลที่ผ่านขั้นตอนการแบ่งช่วงข้อมูล และไม่ผ่านการถ่วงน้ำหนัก กับแบบจำลองที่ผ่านขั้นตอนการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจาก DEMATEL กับ แบบจำลองที่ผ่านขั้นตอนการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจาก Emerging Pattern	70
23	ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ผ่านขั้นตอนการแบ่งช่วงข้อมูล และคัดเลือกคุณลักษณะ และคำนวณน้ำหนักโดยใช้วิธี DEMATEL และ Emerging Pattern	71
24	การแบ่งข้อมูลใช้ในการเรียนรู้ และ ทดสอบอัลกอริทึม แบบวิธี 10 fold Cross Validation	73
25	น้ำหนักของแต่ละ Attribute ที่ได้มาจากวิธีการ DEMATEL โดยไม่ผ่านการคัดเลือก คุณลักษณะ	74
26	น้ำหนักของแต่ละ Item set ที่มาจากกลุ่มผลเฉลยที่เป็น Positive ด้วยวิธี Emerging Pattern	75
27	น้ำหนักของแต่ละ Item set ที่มาจากกลุ่มผลเฉลยที่เป็น Negative ด้วยวิธี Emerging Pattern	76
28	ประสิทธิภาพของแบบจำลองที่สร้างด้วยข่ายงานแบบ เบย์ส์ (แสดงประสิทธิภาพการจำแนกข้อมูล (Accuracy))	78
29	แสดง Precision , Recall และ F_measure ของแบบจำลองที่สร้างจาก ข่ายงานแบบเบย์ส์	79
30	ความน่าจะเป็นแบบมีเงื่อนไข ของ Attribute Month ที่สัมพันธ์กับ ผลเฉลยของข้อมูลการระบาดของโรคไข้หวัดนก	82
31	ความน่าจะเป็นแบบมีเงื่อนไข ของ Attribute Height ที่มีความสัมพันธ์กับผลเฉลยของข้อมูล ในช่วงเวลา 12 เดือน สำหรับกรณี Positive Class	84

### สารบัญตาราง (ต่อ)

ตารางที่		หน้า
32	ความน่าจะเป็นแบบมีเงื่อนไข ของ Attribute Height ที่มีความสัมพันธ์กับผล เฉลี่ยของข้อมูล ในช่วงเวลา 12 เดือน สำหรับกรณี Negative Class	85
33	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับความ สูงจากระดับน้ำทะเลในช่วงเวลาเดือน มกราคม ถึง เดือน กุมภาพันธ์ สำหรับ กรณี Positive Class	86
34	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความ สูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือนมีนาคม จนถึงเดือน เมษายน สำหรับกรณี Positive Class	87
35	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความ สูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน พฤษภาคม จนถึงเดือน มิถุนายน สำหรับกรณี Positive Class	88
36	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความ สูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน กรกฎาคม จนถึงเดือน สิงหาคม สำหรับกรณี Positive Class	89
37	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความ สูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน กันยายน จนถึงเดือน ตุลาคม สำหรับกรณี Positive Class	90
38	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน พฤศจิกายน จนถึงเดือน ธันวาคม สำหรับกรณี Positive Class	91
39	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือนมกราคม จนถึงเดือน กุมภาพันธ์ สำหรับกรณี Negative Class	93
40	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือนมีนาคม จนถึงเดือน เมษายน สำหรับกรณี Negative Class	94

## สารบัญตาราง (ต่อ)

ตารางที่		หน้า
41	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือนพฤษภาคม จนถึงเดือนมิถุนายน สำหรับกรณี Negative Class	95
42	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน กรกฎาคม จนถึงเดือนสิงหาคม สำหรับกรณี Negative Class	96
43	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน กันยายน จนถึงเดือนตุลาคม สำหรับกรณี Negative Class	97
44	ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน พฤศจิกายน จนถึงเดือนธันวาคม สำหรับกรณี Negative Class	98
45	ความน่าจะเป็นแบบมีเงื่อนไขของ ลักษณะของพื้นที่ (Land Shape) ที่สัมพันธ์กับความสูงจากระดับน้ำทะเล (Height) สำหรับกรณี Positive Class	100
46	ความน่าจะเป็นแบบมีเงื่อนไขของ ลักษณะของพื้นที่ (Land Shape) ที่สัมพันธ์กับความสูงจากระดับน้ำทะเล (Height) สำหรับกรณี Negative Class	101
47	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์กับความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงระดับความสูง น้อยกว่า 3.5 เมตร (H1) หรือ ช่วงระดับความสูงมากกว่า 3.5 ถึง 6.5 เมตร (H2) และลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Shape) สำหรับ กรณี Positive Class	102
48	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงมากกว่า 6.5 เมตร ถึง 12 เมตร (H3) หรือ ช่วงระดับความสูงมากกว่า 12 เมตร ถึง 31 เมตร (H4) และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Shape) สำหรับ กรณี Positive Class	103

## สารบัญตาราง (ต่อ)

ตารางที่		หน้า
49	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงมากกว่า 31 เมตร ถึง 34.5 เมตร (H5) หรือ ช่วงระดับความสูงมากกว่า 34.5 เมตร ถึง 97 เมตร (H6) และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class	104
50	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงมากกว่า 97 เมตร ถึง 124.5 เมตร (H7) หรือ ช่วงระดับความสูงมากกว่า 124.5 เมตร ถึง 382.5 เมตร (H8) และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class	105
51	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงมากกว่า 382.5 เมตร ถึง 388.5 เมตร (H9) หรือ ระดับความสูงที่อยู่ในช่วง มากกว่า 388.5 เมตร (H10) และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class	106
52	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงระดับความสูงน้อยกว่า 3.5 เมตร (H1) หรือ ช่วงระดับความสูงมากกว่า 3.5 ถึง 6.5 เมตร (H2) และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class	108
53	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงมากกว่า 6.5 เมตร ถึง 12 เมตร (H3) หรือ ช่วงระดับความสูงมากกว่า 12 เมตร ถึง 31 เมตร (H4) และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class	109

## สารบัญตาราง (ต่อ)

ตารางที่		หน้า
54	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงระดับความสูงมากกว่า 31 เมตร ถึง 34.5 เมตร (H5) หรือ ช่วงระดับความสูง มากกว่า 34.5 เมตร ถึง 97 เมตร (H6) และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class	110
55	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงระดับความสูงมากกว่า 97 เมตร ถึง 124.5 เมตร (H7) หรือ ช่วงระดับความสูง มากกว่า 124.5 เมตร ถึง 382.5 เมตร (H8) และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class	111
56	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงระดับความสูงมากกว่า 382.5 เมตร ถึง 388.5 เมตร (H9) หรือ ช่วงระดับความสูง มากกว่า 388.5 เมตร (H10) และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class	112
57	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยที่ถูกทำลาย (Destroy) ที่สัมพันธ์ กับ จำนวนสัตว์ปีกที่ตาย (Death) ที่อยู่ในช่วงน้อย 5 ตัว (D1) หรือ อยู่ในช่วง มากกว่า 5 ตัว ถึง 18 ตัว และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class	114
58	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยที่ถูกทำลาย (Destroy) ที่สัมพันธ์ กับ จำนวนสัตว์ปีกที่ตาย (Death) อยู่ในช่วง มากกว่า 18 ตัว ถึง 98 ตัว หรือ อยู่ในช่วงมากกว่า 98 ตัว และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class	115

## สารบัญตาราง (ต่อ)

ตารางที่		หน้า
59	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยที่ถูกทำลาย (Destroy) ที่สัมพันธ์ กับ จำนวนสัตว์ปีกที่ตาย (Death) ในช่วงน้อย 5 ตัว (D1) หรือ อยู่ในช่วง มากกว่า 5 ตัว ถึง 18 ตัว และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class	117
60	ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยที่ถูกทำลาย (Destroy) ที่สัมพันธ์ กับ จำนวนสัตว์ปีกที่ตาย (Death) อยู่ในช่วง มากกว่า 18 ตัว ถึง 98 ตัว หรือ อยู่ในช่วงมากกว่า 98 ตัว และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class	118
61	ตัวอย่างความน่าจะเป็นแบบมีเงื่อนไขของ ความชื้น (Humidity) ที่สัมพันธ์ กับ ช่วงเดือนธันวาคม (Month) สภาพพื้นที่ (Landscape) อุณหภูมิ (Temperature) และ ความสูงของพื้นที่ (Height) ในกรณีที่มีผลเฉลยเป็น Positive Class	120
62	ตัวอย่างความน่าจะเป็นแบบมีเงื่อนไขของ ความชื้น (Humidity) ที่สัมพันธ์ กับ ช่วงเดือนธันวาคม (Month) สภาพพื้นที่ (Landscape) อุณหภูมิ (Temperature) และ ความสูงของพื้นที่ (Height) ในกรณีที่มีผลเฉลยเป็น Negative Class	122
63	ตัวอย่างความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนชั่วโมงที่มีแสงแดด (Day Light Hour) ที่สัมพันธ์ กับ ช่วงเดือนที่มีการระบาด (Month) สภาพพื้นที่ (Landscape) และ ความสูงของพื้นที่ (Height) ในกรณีที่มีผลเฉลยเป็น Positive Class	124
64	ตัวอย่างความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนชั่วโมงที่มีแสงแดด (Day Light Hour) ที่สัมพันธ์ กับ ช่วงเดือนที่มีการระบาด (Month) สภาพพื้นที่ (Landscape) และ ความสูงของพื้นที่ (Height) ในกรณีที่มีผลเฉลยเป็น Negative Class	125

## สารบัญตาราง (ต่อ)

ตารางที่		หน้า
65	ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วง ระยะเวลาเดือน มกราคม	128
66	ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วง ระยะเวลาเดือน กรกฎาคม	130
67	ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วง ระยะเวลาเดือน สิงหาคม	134
68	ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วง ระยะเวลาเดือน กันยายน	137
69	ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วง ระยะเวลาเดือน ตุลาคม	140
70	ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วง ระยะเวลาเดือน พฤศจิกายน	152
71	ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วง ระยะเวลาเดือน ธันวาคม	159
72	ประสิทธิภาพของแบบจำลองที่สร้างด้วยต้นไม้ตัดสินใจ (J48) โดยวัดจาก ความถูกต้องของแบบจำลองในการจำแนกข้อมูล	162
73	ค่า Precision Recall และ F-measure ที่แสดงประสิทธิภาพของแบบจำลองที่ สร้างจากต้นไม้ตัดสินใจ (J48) โดยใช้ชุดข้อมูลที่มีวิธีการเตรียมที่แตกต่างกัน	163
74	กฎที่ได้มาจากต้นไม้ตัดสินใจ (J48) จากแบบจำลองในภาพที่ 8	166
75	กฎที่ได้จากแบบจำลองที่สร้างจากต้นไม้ตัดสินใจ (J48) ตามภาพที่ 9	168
76	กฎที่ได้จากแบบจำลองที่สร้างจากต้นไม้ตัดสินใจตามภาพที่ 10	170
77	ประสิทธิภาพการจำแนกข้อมูล (Accuracy) ของแบบจำลองที่สร้างด้วย วิธีการ Naïve Bayes Tree (NB Tree)	171
78	ค่า Precision Recall และ F-measure ในการจำแนกข้อมูลของแบบจำลองที่ สร้างจากวิธีการ Naïve Bayes Tree (NB Tree)	172

## สารบัญตาราง (ต่อ)

ตารางที่		หน้า
79	ร้อยละของโอกาสที่ละทิ้งของ NB Tree จะมีผลเฉลยเป็น Positive เมื่อความชื้นสัมพัทธ์อยู่ในช่วง น้อยกว่า 60.5% (Hu1)	179
80	ร้อยละของโอกาสที่ละทิ้งของ Naïve Bayes Tree จะมีผลเฉลยเป็น Positive เมื่อ ความชื้นสัมพัทธ์อยู่ในช่วง 60.5 % ถึง 75.09% (Hu2)	180
81	ร้อยละของโอกาสที่ละทิ้งของ NB Tree จะมีผลเฉลยเป็น Positive เมื่อความชื้นสัมพัทธ์ อยู่ในช่วง 60.5% - 75.09% (Hu3)	182
82	เปรียบเทียบประสิทธิภาพของความถูกต้องในการจำแนกข้อมูล (accuracy) ของแบบจำลองที่สร้างขึ้นด้วยอัลกอริทึมการจำแนกข้อมูลทั้ง 3 แบบ	183
83	ค่า Precision Recall และ F-measure ของแต่ละแบบจำลองที่แบบจำลองที่สร้างขึ้นด้วยอัลกอริทึมการจำแนกข้อมูลทั้ง 3 แบบ	184

## สารบัญภาพ

ภาพที่		หน้า
1	แสดงลักษณะของ Bayesian Network	11
2	ตัวอย่างแบบจำลองที่สร้างด้วย Bayesian Network	13
3	ตัวอย่างข้อมูลความชื้น และ ชั่วโมงที่มีแสงแดด จากกรมอุตุนิยมวิทยา	46
4	ตัวอย่างข้อมูลอุณหภูมิของแต่ละจังหวัดในประเทศไทย	47
5	ตัวอย่างข้อมูลระดับความสูงจากระดับน้ำทะเลของแต่ละพื้นที่ในประเทศไทย จาก กรมอุตุนิยมวิทยา	48
6	ขั้นตอนในการทดลอง	52
7	แสดงลักษณะแบบจำลอง Bayesian Network ที่สร้างจากชุดข้อมูลที่ผ่าน กระบวนการถ่วงน้ำหนักและการคัดเลือกคุณลักษณะ ( ชุดข้อมูล H)	80
8	แบบจำลองที่สร้างด้วย อัลกอริทึม Decision Tree (J48) ที่สร้างจากชุดข้อมูลที่ผ่าน การ Discretize การคัดเลือก Attribute และ การถ่วงน้ำหนัก (ในกรณีที่มีจำนวน สัตว์ที่ถูกทำลาย อยู่ในช่วง 18 – 86 ตัว	165
9	แบบจำลองที่สร้างด้วย อัลกอริทึม Decision Tree (J48) ที่สร้างจากชุดข้อมูลที่ผ่าน การ Discretize การคัดเลือก Attribute และ การถ่วงน้ำหนัก (ในกรณีที่มีจำนวน สัตว์ปีกที่ถูกทำลาย อยู่ในช่วง 1 -18 ตัว	167
10	แบบจำลองที่สร้างด้วย อัลกอริทึม Decision Tree (J48) ที่สร้างจากชุดข้อมูลที่ผ่าน การ Discretize การคัดเลือก Attribute และ การถ่วงน้ำหนัก (ในกรณีที่มีจำนวนสัตว์ปีกที่ถูกทำลาย อยู่ในช่วงน้อยกว่า 1 ตัว, ช่วง 88 – 969 ตัว และ ช่วง มากกว่า 969 ตัว	169
11	ลักษณะแบบจำลองของ โรคใช้หัดคนกที่สร้างด้วย Naive Bayes Tree เมื่อ ความชื้นสัมพัทธ์ในอากาศ อยู่ในช่วง 60.5% ถึง 75.09% และ อยู่ในลักษณะภูมิ ประเทศแบบ คาบสมุทรมหาไต้ชายฝั่งตะวันตก คาบสมุทรมหาไต้ชายฝั่ง ตะวันออก ชายฝั่งภาคตะวันออกเฉียงใต้ ทิวเขาและหุบเขาภาคตะวันตก และ แอ่ง สกลนคร	174
12	แบบจำลองที่สร้างด้วย Naive Bayes Tree เมื่อ ความชื้นสัมพัทธ์ในอากาศ อยู่ ในช่วง 60.5% ถึง 75.09% และ อยู่ในพื้นที่ที่มีลักษณะ ที่ราบลุ่มน้ำตอนบน	175

### สารบัญภาพ (ต่อ)

ภาพที่		หน้า
13	แบบจำลองใช้หัดคนกที่สร้างด้วย Naive Bayes Tree เมื่อ ความชื้นสัมพัทธ์ในอากาศ อยู่ในช่วง 60.5% ถึง 75.09% และมีลักษณะภูมิประเทศเป็นแบบ บริเวณขอบที่ราบภาคกลาง, บริเวณแอ่งโคราช และ ทิวเขาและหุบเขาภาคเหนือ	176
14	ลักษณะของแบบจำลองใช้หัดคนกที่สร้างด้วย Naive bayes Tree เมื่อ ความชื้นสัมพัทธ์ในอากาศ อยู่ในช่วงมากกว่า 75.09% (Hu3)	177
15	ลักษณะของแบบจำลองใช้หัดคนกที่สร้างด้วย Naive Bayes Tree เมื่อ ความชื้นสัมพัทธ์ในอากาศ อยู่ในช่วง น้อยกว่า 60.5% (Hu1) หรือ ความชื้นสัมพัทธ์ในอากาศอยู่ในช่วง มากกว่า 75.09 (Hu3)	178

### คำอธิบายสัญลักษณ์และคำย่อ

Bayes Net	=	Bayesian Network
NBTree	=	Naive Bayes Tree
EP	=	Emerging Patthern
HW	=	การคำนวณคุณลักษณะด้วยวิธีการของ Hall
CSE	=	Classification Subset Evaluator
BF	=	Best First search
DEMATEL	=	Decision Making Trial And Evaluation Laboratory

## แบบจำลองเพื่อใช้สำหรับการเฝ้าระวังโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีกของไทย โดย เทคนิคการเหมืองข้อมูล

### Model for Avian Influenza Surveillance of Poultry Husbandary in Thailand Using Data Mining

#### คำนำ

เนื่องด้วยโรคไข้หวัดนกเป็นโรคที่ส่งผลกระทบต่อเป็นวงกว้างทั้งในด้านคุณภาพชีวิต ด้านสาธารณสุข ด้านเกษตรกรรมและปศุสัตว์ และ ด้านเศรษฐกิจ ด้วยเหตุนี้ จึงทำให้มีนักวิจัยพยายามที่จะค้นหาวิธี และ เครื่องมือที่ใช้เฝ้าระวัง ควบคุม และ ยับยั้ง การระบาดของโรคไข้หวัดนก เช่น การคิดค้นเครื่องมือที่ช่วยให้การวินิจฉัยโรคไข้หวัดนก ที่มีความรวดเร็ว และ แม่นยำ หรือ การสร้างเครื่องมือเพื่อจำลองจำนวนผู้ป่วยที่จะเกิดขึ้น โดยสมมุติเหตุการณ์ว่ามีการระบาดของโรคไข้หวัดนกโดยใช้ปัจจัยด้านสภาพแวดล้อม และ ช่วงระยะเวลาที่มีการระบาด หรือ การค้นหาเส้นทางการระบาดของโรคไข้หวัดนก เป็นต้น ซึ่งการคิดค้นวิธีการ และ เครื่องมือ ในการเฝ้าระวังโรคนั้นต้องอาศัยข้อมูลจำนวนมาก โดยเฉพาะการสร้างแบบจำลอง ซึ่งต้องทำการค้นหาลักษณะสำคัญของข้อมูล ที่สามารถบ่งชี้ได้ว่า ถ้าข้อมูลที่มีอยู่ มีลักษณะตามแบบจำลองที่สร้างขึ้น สามารถทำนายผลของข้อมูลดังกล่าว ว่ามีร้อยละความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกเป็นเท่าไร

การค้นหาลักษณะที่สำคัญของข้อมูลที่อยู่ในชุดข้อมูลนั้น เรียกว่าการทำ เหมืองข้อมูล ซึ่งการทำเหมืองข้อมูลมีหลากหลายวิธี ขึ้นอยู่กับรูปแบบผลลัพธ์ที่ต้องการ นอกจากนั้น สามารถใช้วิธีการทางด้านการทำเหมืองข้อมูลเพื่อค้นหา คุณลักษณะที่มีอยู่ในชุดข้อมูลมีความสัมพันธ์กับการเกิดการระบาดของโรคไข้หวัดนกอย่างไร เพื่อที่จะได้นำเอาความสัมพันธ์เหล่านั้นมาสร้างเป็นแบบจำลองเพื่อใช้ในการเฝ้าระวังการระบาดของโรคไข้หวัดนก

นอกจาก ใช้วิธีการทำเหมืองข้อมูลมาหาความสัมพันธ์ระหว่างคุณลักษณะกับการระบาดของโรคไข้หวัดนกเพื่อสร้างเป็นแบบจำลองแล้ว สามารถนำเอาเทคนิคการทำเหมืองข้อมูลมาใช้ในการหาระดับความสำคัญของแต่ละคุณลักษณะที่ส่งผลต่อการระบาดของโรคไข้หวัดนกเพื่อนำมาใช้ร่วมกับการสร้างแบบจำลอง เพื่อให้แบบจำลองที่มีความแม่นยำและความถูกต้องในการทำนายเพิ่มขึ้น

งานวิจัยนี้เป็นการสร้างแบบจำลองเพื่อใช้ในการเฝ้าระวังการระบาดของโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีกของไทย โดยใช้วิธีการทางด้านการทำเหมืองข้อมูลมาใช้ในการสร้างแบบจำลอง ซึ่งวิธีการที่นำมาใช้ในการสร้างแบบจำลองคือการจำแนกข้อมูล (Classification) ซึ่งเป็นวิธีการหนึ่งในวิธีการทำเหมืองข้อมูล โดยเทคนิคที่นำมาใช้ในการสร้างแบบจำลองของงานวิจัยนี้ ประกอบด้วย Decision Tree (J48), Naïve Bayes Tree และ Bayesian Network

ขั้นตอนในการสร้างแบบจำลองนั้นเริ่มจาก การนำเอาข้อมูลที่เกี่ยวข้องกับการระบาดของโรคไข้หวัดนก มาทำการคัดเลือกคุณลักษณะที่มีความสัมพันธ์ต่อการระบาดของโรคไข้หวัดนก ซึ่งคุณลักษณะที่ได้รับการคัดเลือกจะถูกนำไปใช้ในการสร้างแบบจำลอง และ ทำการถ่วงน้ำหนักให้กับข้อมูล โดยวิธีการที่นำมาใช้ในการถ่วงน้ำหนักมี 2 วิธีประกอบด้วย วิธี DEMATEL และ Emerging Pattern ซึ่งสาเหตุที่ใช้การถ่วงน้ำหนัก 2 วิธีเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองที่มีการถ่วงน้ำหนักที่แตกต่างกัน เมื่อทำการถ่วงน้ำหนักให้กับข้อมูลเรียบร้อยแล้วจึงสร้างแบบจำลองด้วยเทคนิคการทำเหมืองข้อมูล ทำการวัดผลประสิทธิภาพแบบจำลอง และ ทำการเลือกแบบจำลองที่มีประสิทธิภาพดีที่สุดไปใช้ในการสร้างกฎ เพื่อนำมาใช้ในการเฝ้าระวังโรคไข้หวัดนก

## วัตถุประสงค์

1. เพื่อศึกษาเทคนิคทางการทำเหมืองข้อมูลด้านการจำแนกข้อมูลและนำไปใช้สร้างแบบจำลองที่จะช่วยให้ทราบถึงลักษณะของพื้นที่เสี่ยงสัตว์ปีกของไทยที่มีความเสี่ยงจะเกิดการระบาดของโรคไข้หวัดนก
2. เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองที่สร้างจากอัลกอริทึมทางการทำเหมืองข้อมูลที่แตกต่างกัน
3. เพื่อให้ได้มาซึ่งแบบจำลองที่สามารถนำมาใช้ในการทำนายลักษณะของพื้นที่เสี่ยงสัตว์ปีกของไทยที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนก

## ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถนำเทคนิคทางการทำเหมืองข้อมูลในด้านการจำแนกข้อมูลมาใช้ในการสร้างแบบจำลองเพื่อใช้สำหรับการเฝ้าระวังโรคไข้หวัดนกในพื้นที่เสี่ยงสัตว์ปีกของไทย
2. ได้แบบจำลองเพื่อใช้สำหรับการเฝ้าระวังโรคไข้หวัดนกในพื้นที่เสี่ยงสัตว์ปีกของไทยที่มีประสิทธิภาพในการทำนายที่ดี

## ขอบเขตของงานวิจัย

1. ข้อมูลที่ใช้คือข้อมูลจากรายงานจำนวนสัตว์ปีกป่วยตาย ของกรมปศุสัตว์ตั้งแต่ปี พ.ศ. 2547-2550
2. ข้อมูลสภาพภูมิอากาศ จากกรมอุตุนิยมวิทยาตั้งแต่ปี พ.ศ. 2547-2550
3. ข้อมูลสภาพภูมิประเทศของพื้นที่ที่พบสัตว์ปีกป่วยตายประกอบด้วย ความสูง ลักษณะภูมิประเทศ จากกรมอุตุนิยมวิทยาตั้งแต่ปี พ.ศ. 2547 -2550

## การตรวจเอกสาร

### ทฤษฎีที่เกี่ยวข้อง

งานวิจัยครั้งนี้เป็นการนำเทคนิคทางด้านการทำเหมืองข้อมูลมาประยุกต์ใช้ร่วมกับข้อมูลรายงานสัตว์ปีกป่วยตาย ตั้งแต่ปี 2547-2550 ข้อมูลภูมิอากาศ และข้อมูลภูมิประเทศของพื้นที่พบสัตว์ปีกป่วยตาย เพื่อสร้างแบบจำลองที่ใช้ในการเฝ้าระวังโรคไข้หวัดนกในพื้นที่ที่มีการเลี้ยงสัตว์ปีกของประเทศไทย โดยทฤษฎีและความรู้ที่เกี่ยวข้องกับการวิจัยในครั้งนี้ ได้แก่

1. ความรู้เกี่ยวกับโรคไข้หวัดนก
2. ทฤษฎีด้านการทำเหมืองข้อมูล
3. ทฤษฎีด้านการคัดเลือกคุณลักษณะของข้อมูล
4. วิธีการคำนวณน้ำหนักให้กับข้อมูล
5. ทฤษฎีที่การวัดประสิทธิภาพของแบบจำลอง

#### 1. ความรู้เกี่ยวกับโรคไข้หวัดนก

โรคไข้หวัดนก (อรุณี, 2002) เป็นโรคไข้หวัดใหญ่ชนิดหนึ่งที่เกิดจากเชื้อไวรัสไข้หวัดใหญ่ ซึ่งเชื้อไข้หวัดใหญ่ แบ่ง ออกเป็น 3 กลุ่ม ได้แก่ ชนิด A, B และ C โดยเชื้อไข้หวัดนกอยู่ในกลุ่ม A ซึ่งเป็นกลุ่มเดียวกับ เชื้อไข้หวัดใหญ่ในคน

เชื้อไวรัสไข้หวัดนกอยู่ในตระกูล Orthomyxoviridae ซึ่งเป็น RNA ของไวรัสชนิดมีเปลือกหุ้ม ซึ่งไวรัสที่อยู่ในกลุ่มนี้ มี Surface antigens ที่สำคัญ ได้แก่ Haemagglutinin (H) ซึ่งมี 15 ชนิด และ Neuraminidase (N) มี 9 ชนิด ซึ่งเชื้อไวรัสไข้หวัดนกนั้น มี H ชนิดที่ 5 และ N ชนิดที่ 1 จึงทำให้เชื้อไข้หวัดนกมีชื่อเรียกตามโครงสร้างว่า H5N1

เชื้อไวรัสไข้หวัดนก ไม่สามารถทนทานต่ออากาศร้อนได้ เช่น ถ้าอยู่ในอุณหภูมิ 50 องศาเซลเซียส จะมีชีวิตอยู่ได้นานเพียง 3 ชั่วโมง และ ไม่สามารถทนทานต่อสารเคมีที่มีคุณสมบัติในการละลายไขมัน เช่น สารเคมีประเภท ฟอรั่มาลีน หรือ ประเภทสาร Oxidizing Agents และเชื้อสามารถอยู่ได้นานจากสิ่งที่ยับถ่ายหรือสารคัดหลั่งที่ออกมาจากร่างกายของสัตว์ปีกที่ป่วย เช่น น้ำมูก น้ำตา น้ำลาย เสมหะ และ อุจจาระ

การแพร่กระจายของเชื้อ ไข้หวัดนกจากสัตว์ที่ติดเชื้อมายังสัตว์ปีกที่ยังไม่ได้ติดเชื้อส่วนใหญ่มักจะติดต่อผ่านทางสิ่งยับถ่ายของสัตว์ปีกที่เป็นโรค หรือ การติดเชื้อจากแหล่งน้ำที่มีการเลี้ยงสัตว์ปีกที่เป็นโรคแต่ไม่แสดงอาการให้เห็น เช่น นกเป็ดน้ำ เป็นต้น

ลักษณะอาการของโรคไข้หวัดนกในสัตว์ปีก ขึ้นอยู่กับ หลายปัจจัย เช่น ชนิดสัตว์ปีก, อายุ, สถานะความเครียด และ โรคแทรกซ้อน เป็นต้น ลักษณะอาการที่พบโดยทั่วไปได้แก่

- ชูบพอม ซึม ไม่กินอาหาร ขนยุ่ง จำนวนไข่ไถ่ลดลง
- ไอ จาม หายใจลำบาก น้ำตาไหล หน้าบวม และ หงอนมีสีคล้ำ
- อาจมีอาการของระบบประสาทและท้องเสีย

ซึ่งในสัตว์ปีกบางชนิดที่มีอาการรุนแรงจะตายกระทันหัน โดยไม่แสดงอาการ ซึ่งขึ้นอยู่กับชนิดของเชื้อ ชนิดสัตว์ปีก เป็นต้น

ในปัจจุบันสถานการณ์การระบาดของโรคไข้หวัดนกในประเทศไทย ยังคงมีการระบาดอยู่แต่ไม่รุนแรงเท่ากับในอดีตที่ผ่านมาอันเนื่องมาจากในปัจจุบันมีมาตรการในการป้องกัน การควบคุม รวมถึงมีการเฝ้าระวังที่ดีเช่น การออกมาตรการขนย้ายสัตว์ปีกข้ามจังหวัดที่ต้องการแจ้งขออนุญาต การเลี้ยงไก่ในระบบปิด การออกกฎควบคุมการเลี้ยงเป็ดไล่ทุ่ง ซึ่งมาตรการเหล่านี้เป็นมาตรการในการป้องกันการแพร่ระบาดของโรคไข้หวัดนกไม่ให้เกิดการระบาดไปยังพื้นที่ใกล้เคียง และ ป้องกันการระบาดของโรคไข้หวัดนก ฆ่า ในพื้นที่เลี้ยงสัตว์ปีก

## 2. ทฤษฎีการทำเหมืองข้อมูล

การทำเหมืองข้อมูล เป็นกระบวนการในการค้นหารูปแบบ ความสัมพันธ์ และ แนวทางของข้อมูลที่สำคัญที่ซ่อนอยู่ในข้อมูลที่มีปริมาณจำนวนมาก โดยอาศัยหลักการของทางสถิติ และ วิธีการเรียนรู้ของเครื่องจักร (Machine Learning) ความรู้ที่ได้จากการทำเหมืองข้อมูลมีอยู่ 4 รูปแบบ ประกอบด้วย

- กฎความสัมพันธ์
- การจำแนกข้อมูล
- การแบ่งกลุ่มข้อมูล
- การทำนายข้อมูล

2.1 กฎความสัมพันธ์ (Association Rule) เป็นการแสดงความสัมพันธ์ของ เหตุการณ์ ที่เกิดขึ้นพร้อมกัน อัลกอริทึมสำหรับการค้นหาเหตุการณ์ (Item set) ที่นำมาใช้ในการสร้างกฎนั้น มีอยู่มากมาย แต่ที่ได้รับความนิยม และเป็นที่ยูจคือ **Apiori Algorithm** ซึ่งเป็นอัลกอริทึมที่ใช้ในการหาความสัมพันธ์ระหว่าง Attribute ที่มีอยู่ในชุดข้อมูล ซึ่งมีการนำเสนอครั้งแรกในงานวิจัยของ Rakesh Agrawal และ Ramakrishnan Srikant ในปี 1994 (Agrawal and Srikant, 1994) ขั้นตอนการหาความสัมพันธ์ของ APRIORI Algorithm มีดังต่อไปนี้

ขั้นตอนที่ 1 คัดเลือก ค่าของข้อมูลที่อยู่ในแต่ละ Attribute ที่อยู่ภายใต้ชุดข้อมูลเดียวกัน (Item set) ที่จะนำมาใช้ในการสร้างกฎ โดยคัดเลือกจาก ค่าของข้อมูล ที่มี ค่า support สูงกว่าค่าที่กำหนดไว้ ค่า support เป็น ค่าที่แสดงถึงสัดส่วนของ Transaction ที่มีค่าของข้อมูลที่ต้องการอยู่ใน Transaction นั้น เทียบกับจำนวน Transaction ทั้งหมด ซึ่งเขียนเป็นสมการทางคณิตศาสตร์ได้ดังต่อไปนี้

$$Supp(X \cup Y) = \frac{n(X \cup Y)}{n(\text{transaction})} \quad (1)$$

โดยที่  $n(X \cup Y)$  คือ จำนวน Transaction ที่มีทั้ง X และ Y อยู่ภายใต้ Transaction เดียวกัน

$n(\text{Transaction})$  คือ จำนวน Transaction ทั้งหมดที่อยู่ใน Data set เดียวกัน

X และ Y คือ สมาชิกใน Item set ที่แต่ละตัวมีค่า Support มากกว่าค่าที่กำหนด

ขั้นตอนที่ 2 เมื่อเลือก Item set ได้แล้ว ทำการผสมสมาชิกใน Item Set เหล่านั้นทั้งหมด เพื่อสร้างสมาชิกของ Candidate Set แล้วนำสมาชิกใน Candidate Set ไปวัดค่า Support ของแต่ละสมาชิกที่ทำการผสมขึ้นมานั้น และนำสมาชิกของ Candidate Set ที่มีค่า Support ไปเปรียบเทียบกับค่า minimum Support ที่กำหนดไว้ ถ้าสมาชิกของ Candidate Set มีค่า Support น้อยกว่า ค่า minimum support ที่กำหนดไว้ ให้ทำการตัดทิ้ง Item ที่ผสมขึ้นมานั้น ทิ้งไป แต่ ถ้ามากกว่าค่า minimum Support ที่กำหนดไว้ ไม่ต้องทำการตัด สมาชิก ที่ผสม ขึ้นมานั้นออกจาก สมาชิกของ Candidate Set เมื่อทำการเปรียบเทียบครบทุกสมาชิกใน Candidate Set จึงทำการผสมสมาชิกที่เหลืออยู่ใน Candidate set ใหม่อีกครั้ง แล้วทำการเปรียบเทียบ จนกว่าจะไม่สามารถตัดทิ้ง สมาชิก ที่อยู่ใน candidate set ได้อีก

ขั้นตอนที่ 3 เมื่อได้ สมาชิก ที่อยู่ใน Candidate set ที่ไม่สามารถผสม ได้อีก จึงนำเอา สมาชิกที่อยู่ใน Candidate set มาสร้างเป็นกฎ แล้วนำไปวัดค่าความเชื่อมั่น (Confidence) เพื่อเลือกกฎที่มีค่าความเชื่อมั่น สูงกว่าที่ค่า Minimum Confidence ที่ได้กำหนดไว้ การหาค่าความเชื่อมั่น (Confidence) มีสมการทางคณิตศาสตร์ดังต่อไปนี้

$$CONF(X \rightarrow Y) = \frac{SUPP(X \cup Y)}{SUPP(X)} \quad (2)$$

โดยที่ SUPP (XUY) คือ ค่าSupport ของ Item set ที่มี X และ Y เป็นส่วนประกอบ

SUPP (X) คือ ค่าSupport ของ Item set ที่มี X เป็นส่วนประกอบ

ค่าความเชื่อมั่น (Confidence) เป็นค่าที่แสดงถึงความน่าเชื่อถือของกฎที่ได้มา ซึ่งเป็นการวัดประสิทธิภาพของกฎที่ได้จากการทำ Association Rule ที่นิยมนำมาใช้ มีอยู่ 3 ประเภทได้แก่

- 1) ค่าความเชื่อมั่น (Confidence) เป็นการวัดค่าความน่าเชื่อถือของกฎ
- 2) ค่า LIFT เป็นการวัดค่าความน่าสนใจของกฎ โดยเป็นการวัดความถี่ของการเกิด

Transaction ที่มีลักษณะเป็นไปตามกฎที่ได้มาจากการหากฎโดยใช้ Apriori ลักษณะของสูตรที่ใช้ในการคำนวณมีดังต่อไปนี้

$$LIFT(X \rightarrow Y) = \frac{CONF(X \rightarrow Y)}{SUPP(Y)} \quad (3)$$

โดยที่  $CONF(X \rightarrow Y)$  คือ ค่าความเชื่อมั่นของกฎที่มีส่วนประกอบของ X และ Y

3) ค่า Leverage เป็นการวัดการขึ้นต่อกันของ Item Set ที่นำมาใช้สร้างกฎโดยมีสูตรการคำนวณดังต่อไปนี้

$$Leverage(X \rightarrow Y) = P(X \cap Y) - P(X)P(Y) \quad (4)$$

โดยที่  $P(X \text{ and } Y)$  คือ ความน่าจะเป็นของ Item Set X และ Y

$P(x)$  คือ ความน่าจะเป็นของ Item set X

$P(Y)$  คือ ความน่าจะเป็นของ Item set Y

2.2 การจำแนกข้อมูล (Classification Data) เป็นการนำข้อมูลมาทำการจัดกลุ่มตามเงื่อนไขของกลุ่มที่ได้กำหนดไว้เช่น การจำแนกข้อมูลโรคไข้หวัดนก ตามผลการตรวจซึ่งประกอบด้วย ผลการตรวจว่าเป็นโรคไข้หวัดนก กับ ไม่เป็นโรคไข้หวัดนก ขั้นตอนวิธี ที่นำมาใช้ในการจำแนกข้อมูลประกอบด้วย ต้นไม้ตัดสินใจ (Decision Tree) การจำแนกข้อมูลด้วยวิธีการ Naïve Bayes การจำแนกข้อมูลโดยใช้ข่ายงานแบบเบย์ (Bayesian Network) เป็นต้น ซึ่งการจำแนกข้อมูลมักจะนำไปใช้กับการหาลักษณะของข้อมูล เพื่อระบุประเภทของข้อมูล เช่นการค้นหาคุณลักษณะของพื้นที่ที่เกิดการระบาดของโรคไข้หวัดนก

2.2.1 ต้นไม้ตัดสินใจ (Decision Tree) เป็นวิธีการเรียนรู้ของเครื่องจักรที่นำมาใช้ในการจำแนกข้อมูลที่มีผลเฉลยที่มีลักษณะเป็นข้อมูลที่มีค่าไม่ต่อเนื่อง (Discrete Value) ซึ่งขั้นตอนการจำแนกข้อมูลแสดงออกมาเป็นในรูปของ ต้นไม้ ที่ประกอบด้วย โหนดที่เป็นจุดเริ่มต้นของต้นไม้ ที่เรียกว่า โหนดราก (Root Node) โหนดที่เป็นตัวจำแนกข้อมูล ที่ไม่ใช่โหนดราก หรือที่เรียกว่า โหนดพ่อแม่ (Parent Node) และ โหนดที่เป็นผลลัพธ์ของการจำแนก ที่เรียกว่า โหนดลูก (Leave Node) การแสดงความสัมพันธ์ ระหว่างโหนด ในแผนภาพต้นไม้ จะแสดงเป็นเส้นเชื่อมระหว่าง โหนด แต่ละ โหนด (Mitchell, 1998) ซึ่ง สามารถแสดงผลในอีกรูปแบบได้คือ รูปแบบของเงื่อนไข (If-Then) ซึ่งเป็นรูปแบบที่ง่ายต่อความเข้าใจ และ การนำไปใช้ในการจำแนกข้อมูล ตัวอย่างของอัลกอริทึม ที่ใช้ในการสร้างต้นไม้ตัดสินใจ มีดังต่อไปนี้

ID3 เป็นอัลกอริทึมที่ใช้สำหรับการสร้าง ต้นไม้ตัดสินใจซึ่งคิดขึ้นโดย Ross Quinlan ซึ่งมีหลักการและขั้นตอนการสร้างดังต่อไปนี้ (Mitchell, 1998)

ขั้นตอนที่ 1 การคัดเลือกตัวแปรที่จะนำมาสร้างเป็น โหนด ราก โดยการคัดเลือกจากค่าความสามารถในการจำแนกข้อมูลตามผลเฉลยที่กำหนดไว้หรือที่เรียกว่า Information Gain โดยการคำนวณค่า Information Gain นั้นจะคำนวณจากค่าความปะปนการของข้อมูล หรือที่เรียกว่า ค่า Entropy ซึ่งสูตรในการคำนวณค่า Entropy มีดังนี้คือ

$$Entropy(S) = -P_{(+)} \log_2 P_{(+)} - P_{(-)} \log_2 P_{(-)} \quad (5)$$

โดยที่  $P_{(+)}$  คือ สัดส่วนจำนวนข้อมูลที่มีผลเป็นบวกต่อจำนวนข้อมูลทั้งหมด  
 $P_{(-)}$  คือ สัดส่วนจำนวนข้อมูลที่มีผลเป็นลบต่อจำนวนข้อมูลทั้งหมด  
 $S$  คือ ตัวแปรที่นำมาวัดค่า Entropy

เมื่อหา ค่า Entropy ได้แล้วจึงมาหาค่า Information Gain ซึ่งสูตรในการหาค่า information Gain คือ

$$Gain(S, A) \equiv Entropy(S) - \sum_{v \in value(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (7)$$

โดยที่  $S$  แทนจำนวนข้อมูลทั้งหมด  
 $A$  แทน Attribute ที่ต้องการนำมาหาค่า Information Gain  
 $v$  แทน ค่าของ Attribute A  
 $S_v$  แทน จำนวนข้อมูลที่มีค่าเท่ากับ  $v$

ขั้นตอนที่ 2 เลือก Attribute ที่มีค่า Information Gain มากที่สุดนำไปสร้างเป็น โหนดราก ส่วนเส้นที่เชื่อมระหว่าง โหนด (Node) แทนความหมายถึง ค่าที่อยู่ใน Attribute ที่ถูกนำมาเป็น โหนด ในต้นไม้ตัดสินใจ และ ในแต่ละเส้นเชื่อมจะมีจำนวนของข้อมูลที่มีค่าใน Attribute ที่นำมาสร้างเป็น โหนด (Node) ตรงกับ ค่าที่อยู่ในแต่ละเส้นเชื่อม ซึ่งจำนวนข้อมูลดังกล่าวไปถูกใช้ในการนำไปหา ค่า Information Gain ของแต่ละ Attribute เพื่อเลือก Attribute ที่เหลือ มาเป็น โหนด (Node) ในต้นไม้ตัดสินใจ

J48 เป็นอัลกอริทึมสำหรับสร้าง ต้นไม้ตัดสินใจ ที่พัฒนาขึ้นต่อจาก ID3 (Mitchell ,1998) เพื่อแก้ปัญหาการทำงานของ ID3 Algorithm ซึ่งมีดังต่อไปนี้

Overfitting เกิดจาก ต้นไม้ตัดสินใจที่สร้างขึ้นมีขนาดของต้นไม้ที่ใหญ่เกินไป และ ผลของการจำแนกข้อมูลในข้อมูลที่ใช้ในการสร้างต้นไม้ตัดสินใจ สามารถจำแนกได้ 100% แต่เมื่อไปทดสอบกับข้อมูลทดสอบที่แบ่งไว้ปรากฏว่า ผลการจำแนกนั้นมี Error ที่สูง ซึ่งใน J48 ได้ทำการแก้ไขโดยการเพิ่มการ pruning ซึ่งเป็นการตัดกิ่งของ Decision Tree ที่ไม่จำเป็น ออกไป ซึ่งการทำ Pruning มี 2 รูปแบบคือ แบบ Pre Pruning (Johannes and Gerhard, 1994) หมายถึง ระหว่างการเรียนรู้ เพื่อสร้าง Decision Tree กับชุดข้อมูล Training Set จะทำเรียนรู้ไม่ครบทุกข้อมูลที่อยู่ใน Training Set เมื่อนำแบบจำลองไปทดสอบ ทำให้ได้ Decision Tree ที่ไม่สามารถจำแนกข้อมูลทั้งหมดที่อยู่ในชุด Training Set ได้อย่างถูกต้อง 100%

ตัวอย่างอัลกอริทึมของ Decision Tree ที่ใช้วิธีการ Pruning แบบ Pre Pruning ได้แก่ FOIL และ Mfoil แบบ Post Pruning หมายถึง ในช่วงของการเรียนรู้เพื่อสร้าง Decision Tree จากชุดข้อมูล ที่เรียกว่า Growing set ที่ถูกแบ่งมาจากชุดข้อมูล Training set จะไม่มีการ Pruning ในระหว่างการสร้าง Decision Tree แต่จะมีการ Pruning หลังจากที่ได้ Decision Tree เรียบร้อยแล้ว โดยที่การ Pruning จะใช้ชุดข้อมูล Pruning Set ที่ได้จากการแบ่งข้อมูลจาก training set เช่นเดียวกับ Growing set มาทำการตัดกิ่งของ Decision Tree จนกว่าการ Pruning จะทำให้ความแม่นยำในการจำแนกข้อมูล ของ Decision Tree ลดลง

อัลกอริทึม Decision Tree (ID3) นั้นไม่รองรับข้อมูลที่เป็นตัวเลข แต่ใน J48 นั้นรองรับในเรื่องของการจำแนกข้อมูลที่มี Attribute เป็นแบบตัวเลข ในอัลกอริทึม ID3 การค้นหา Attribute เพื่อนำไปสร้าง โหนด ราก และ โหนด พ่อแม่ จะวัดจากค่า information Gain ของแต่ละ Attribute ซึ่งค่า Information Gain ที่เหมาะสมสำหรับแต่ละ Attribute อาจจะมี Bias เกิดขึ้นได้ถ้า Attribute ที่เลือกมานั้นสามารถจำแนกข้อมูลที่ใช้ในการสร้าง Decision Tree ได้ 100% ผลเสียคือเมื่อนำ Decision Tree ไปทดสอบกับชุดข้อมูลทดสอบ จะทำให้ได้ความแม่นยำและความถูกต้องในการจำแนกไม่ดี ดังนั้นใน J48 Algorithm จึงได้เปลี่ยนการเลือก Attribute โดยใช้ค่า Information Gain เป็นใช้ค่า Gain Ratio ซึ่งเป็นนำ Information Gain ทำการถ่วงน้ำหนัก ด้วยค่า Split information ซึ่งเป็นค่าที่แสดงถึงค่า Entropy ที่แท้จริงของชุดข้อมูลที่มีต่อค่าของ Attribute ซึ่งสูตรในการคำนวณหาค่า Split Information คือ

$$SplitInformation(S, A) \equiv - \sum_{i=1}^c \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|} \quad (8)$$

โดยที่  $S_i$  แทนจำนวนข้อมูลที่แบ่งตามจำนวนสมาชิกที่อยู่ใน Attribute A  
 $I$  แทนจำนวนสมาชิกที่อยู่ใน Attribute A  
 $S$  แทนจำนวนข้อมูลทั้งหมด

เมื่อหาค่า SplitInformation ได้แล้ว จึงนำมาหาค่า Gain Ratio จากสูตร

$$GainRatio(S, A) \equiv \frac{Gain(S, A)}{SplitInformation(S, A)} \quad (9)$$

เมื่อทำการหาค่า Gain Ratio ของแต่ละ Attribute ได้แล้วจึงทำการเลือก Attribute ที่มีค่า Gain Ratio มากที่สุดนำไปสร้างเป็น โหนดราก และทำการคำนวณหาค่า Gain Ratio ของ แต่

2.2.2 Bayesian Network เป็นกราฟชนิดหนึ่ง que สร้างมาจากทฤษฎีของเบย์ส์ซึ่ง เป็นทฤษฎีที่อธิบายถึงการขึ้นต่อกันของตัวแปร โดยใช้ความน่าจะเป็นแบบมีเงื่อนไขในการอธิบาย ความสัมพันธ์ระหว่างตัวแปร โดยแสดง ออกมาเป็นในรูปแบบ Direct Acyclic Graph (DAG) ซึ่ง รูปแบบในการแสดงว่า คุณสมบัติที่อยู่ในชุดข้อมูลมีการขึ้นต่อกัน โดยแสดงออกมาเป็นสัญลักษณ์ ประกอบด้วย โหนด แทน คุณสมบัติของชุดข้อมูล (Attribute) และ เส้นเชื่อมแทนทิศทางการขึ้นต่อกัน (Mitchell, 1998)



ภาพที่ 1 ลักษณะของ Bayesian Network

จากภาพเราสามารถแปลความหมายได้ว่า ตัวแปรสุ่ม B ขึ้นกับตัวแปรสุ่ม A ซึ่งเมื่อนำมาเขียนในรูปแบบความน่าจะเป็นแบบมีเงื่อนไขคือ  $P(B|A)$  ซึ่งนำมาใช้ในการคำนวณความน่าจะเป็นแบบมีเงื่อนไข ของ A ที่ขึ้นกับ B ซึ่งเขียนเป็นรูปสมการทางคณิตศาสตร์ ตามทฤษฎีของเบย์ส์จะได้ว่า

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (10)$$

โดยที่  $P(A|B)$  แทนความน่าจะเป็นแบบมีเงื่อนไขของ A ขึ้นกับ B  
 $P(B|A)$  แทนความน่าจะเป็นแบบมีเงื่อนไขของ B ขึ้นกับ A  
 $P(A)$  แทนความน่าจะเป็นของ A  
 $P(B)$  แทนความน่าจะเป็นของ B

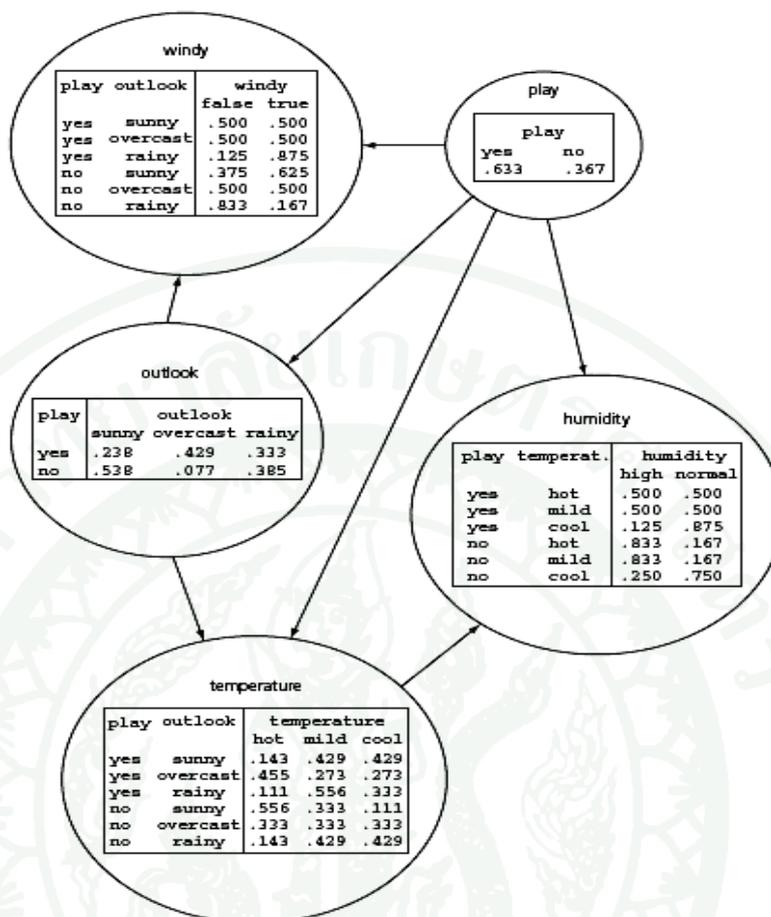
หากแบบจำลองมีมากกว่า 2 ตัวแปร สามารถเขียนในรูปแบบสมการเชิงคณิตศาสตร์ได้เป็น

$$P(X_1, X_2, X_3, \dots, X_n) = \prod_{i=1}^n (P(X_i | \text{Parent}(X_i))) \quad (11)$$

โดยที่  $X_1, X_2, \dots, X_n$  แทน ตัวแปรที่อยู่ในชุดข้อมูล  
 $\text{Parent}(X_i)$  แทน ตัวแปรที่เป็นเงื่อนไขของตัวแปร  $X_i$

เมื่อได้แบบจำลองที่สร้างจากอัลกอริทึม Bayesian Network สิ่งที่ได้ออกมาพร้อมกันคือ ตารางความน่าจะเป็นแบบมีเงื่อนไขของแต่ละคุณลักษณะที่ปรากฏใน แบบจำลองที่สร้างด้วย Bayesian Network

ดังนั้นในการคำนวณหาความน่าจะเป็นของทั้งหมดตามความสัมพันธ์ของแบบจำลองจึงต้องใช้ตารางความน่าจะเป็นที่ถูกสร้างขึ้นมากพร้อมกับแบบจำลองมาในการคำนวณ ซึ่งตัวอย่างของการหาความน่าจะเป็น และ ตัวอย่างของแบบจำลองที่สร้างด้วย Bayesian Network มีดังต่อไปนี้ (Witten and Frank, 2005)



ภาพที่ 2 ตัวอย่างแบบจำลองที่สร้างด้วย Bayesian Network

ที่มา: หนังสือ Data Mining Practical Machine Learning Tools and Technique. Witten and Frank (2005)

ตัวอย่างแบบจำลองในภาพเป็นเรื่องเกี่ยวกับสภาพภูมิอากาศที่ส่งผลต่อการออกไปเล่นกีฬากลางแจ้ง ซึ่งในชุดข้อมูลมี Attribute ดังต่อไปนี้ Outlook, Humidity, Temperature และ windy มี Class ของข้อมูลคือ Play โดยจากรูปภาพจะเห็นตารางความน่าจะเป็นของแต่ละ Attribute ซึ่ง ด้านฝั่งซ้ายของตาราง แทน ค่าของ Attribute ที่ ถูก Attribute ที่เราสนใจอ้างอิง และ ด้านฝั่งขวาแต่ละคอลัมน์ แทน ค่าของ Attribute ที่สนใจ

ตัวอย่างเช่น ต้องการหาค่าความน่าจะเป็นของ Attribute Play มีค่า เท่ากับ Yes โดยที่ Outlook = rainy, Temperature = Cool, Humidity = high และ Windy = true ซึ่งจากภาพที่ 2 เห็นได้ว่า  $P(\text{Attribute Play} = \text{Yes})$  มีค่าความน่าจะเป็นเท่ากับ 0.633 และ  $P(\text{Outlook} = \text{Rainy} | \text{Attribute Play} = \text{Yes})$  มีค่าความน่าจะเป็นเท่ากับ 0.333 และ  $P(\text{Attribute Temperature} = \text{Cool} | \text{Attribute Play} = \text{yes}, \text{Attribute Outlook} = \text{Rainy})$  มีค่าความน่าจะเป็นเท่ากับ 0.333 และ  $P(\text{Attribute Humidity} = \text{High} | \text{Attribute Play} = \text{Yes}, \text{Attribute Temperature} = \text{Cool})$  มีค่าความน่าจะเป็นเท่ากับ 0.125 และ  $P(\text{Attribute Windy} = \text{True} | \text{Attribute Play} = \text{Yes}, \text{Attribute Outlook} = \text{Rainy})$  มีค่าความน่าจะเป็นเท่ากับ 0.875 เนื่องจากต้องการหาค่าความน่าจะเป็นของ Attribute Play = Yes เมื่อ Attribute Outlook = rainy, Temperature = Cool, Humidity = Height และ Windy = True ดังนั้น เราจึงนำความน่าจะเป็นที่ได้กล่าวมาข้างต้น ทำการคูณกัน จะได้ ความน่าจะเป็นของ Attribute Play = Yes โดยที่ Outlook = rainy, Temperature = cool, Humidity = High และ Windy = true มีค่าความน่าจะเป็นเท่ากับ 0.0077 ซึ่งถ้าเราหาค่าความน่าจะเป็นของ Attribute Play = no โดยที่ Attribute Outlook = Rainy, Temperature = Cool, Humidity = High, Windy = true มีค่าความน่าจะเป็นเท่ากับ 0.0075 ซึ่งในหลักการของค่าความน่าจะเป็น เมื่อนำเอาค่าความน่าจะเป็นของ Play = yes โดยที่ Outlook = rainy, Temperature = Cool, Humidity = High, Windy = true มารวมกับ ค่าความน่าจะเป็นของ Play = No โดยที่ Outlook = rainy, Temperature = Cool, Humidity = High, Windy = true ต้อง เท่ากับ 1 แต่จากค่าความน่าจะเป็นที่หามาแล้ว เมื่อนำมารวมกันได้เท่ากับ 0.0152 เพื่อให้ได้ ความน่าจะเป็นที่ถูกต้อง จึงต้องนำเอาความน่าจะเป็นที่หามาได้ ของ กรณีที่ Attribute Play = Yes ซึ่งเท่ากับ 0.0077 มาหารด้วยผลรวมของค่าความน่าจะเป็นของทั้งสองกรณี ซึ่งเท่ากับ 0.0152 ซึ่งจะได้ว่า ความน่าจะเป็นที่แท้จริงของ Attribute Play = Yes ในกรณีที่ Outlook = rainy, temperature = cool, Humidity = High และ Windy = true มีค่าเท่ากับ 0.5065

ในการสร้างแบบจำลองด้วย Bayesian Network ในอัลกอริทึม จะมีการทำงานสองส่วน ประกอบด้วย ส่วนที่ใช้ในการสร้างโครงสร้างของแบบจำลองที่สร้างด้วย เทคนิค Bayesian Network กับส่วนที่ใช้การหาตารางความน่าจะเป็นแบบมีเงื่อนไขของแต่ละ Attribute ที่อยู่ในแบบจำลอง ในส่วนของการเรียนรู้เพื่อสร้างแบบจำลอง ด้วย Bayesian Network มีอัลกอริทึมที่ใช้ดังต่อไปนี้

Hill Climbing เป็นการใช้เทคนิคในการค้นหาเส้นเชื่อมที่เหมาะสมในการลบ หรือ เพิ่มระหว่าง Attribute โดยที่ไม่ได้ยึดตามลำดับของ Attribute (Bouckart, 2008)

K2 เป็นการใช้เทคนิคการค้นหาแบบ Hill Climbing ในการเพิ่มเส้นเชื่อมที่เหมาะสมระหว่าง Attribute ในชุดข้อมูลในการสร้างแบบจำลอง Bayesian Network โดยที่ยึดตามลำดับของ Attribute ในข้อมูล (Bouckart, 2008)

TAN (Tree Augmented Naïve Bayes) เป็นการนำเอาชุดข้อมูลมาทำการสร้างต้นไม้ตัดสินใจและ ใช้ Mutual Information Test Conditioned กับ Class ของข้อมูล และ นำเอา Class ขข้อมูล มาสร้าง เป็นคลาสพ่อแม่ของ Tree และ ทำการเพิ่มเส้นเชื่อมระหว่าง Class กับ Attribute และ จึงทำการคำนวณหาค่า ความน่าจะเป็นแบบมีเงื่อนไข ระหว่าง Attribute (Bouckart, 2008)

ในส่วนอัลกอริทึมที่ใช้สำหรับการหาตารางความน่าจะเป็นมีหลายอัลกอริทึม ประกอบด้วย SimpleEstimator เป็นวิธีการในการสร้างตารางความน่าจะเป็นแบบมีเงื่อนไขโดยที่สูตรในการค้นหาคือ

$$P(x_i = k | pa(x_i) = j) = \frac{N_{ijk} + N'_{ijk}}{N_{ij} + N'_{ij}} \quad (12)$$

โดยที่	$Pa(x_i)$ แทน Parent Node ของ Attribute ที่ i
	J แทนค่าของ Parent Node ของ Attribute ที่ i
	$x_i$ แทน Attribute ที่ i
	K แทนค่าของ Attribute x
	$N_{ij}$ แทนจำนวนข้อมูลทั้งหมดที่มีค่าของ Parent Node ของ Attribute ที่ i เท่ากับ j
	$N_{ijk}$ แทนจำนวนข้อมูลในชุดข้อมูลที่มีค่า Attribute $x_i$ มีค่าเท่ากับ K
	$N'_{ijk}$ แทนค่า Alpha ซึ่งปกติกำหนดไว้ที่ 0.5

2.2.3 Naïve Bayes Tree (NB Tree) เป็นเทคนิคในการจำแนกวิธีหนึ่งที่เป็นที่ผสมกันระหว่าง วิธีการจำแนกแบบต้นไม้ตัดสินใจ และ วิธีการจำแนกข้อมูลแบบ Naïve Bayes โดยผลลัพธ์ที่ได้จากการใช้เทคนิคนี้ในการจำแนกข้อมูล ทำให้ได้แบบจำลองที่รองรับตัวแปรที่ไม่มี ความสัมพันธ์กัน และสามารถรองรับกรณีที่มีการจำแนกข้อมูลนั้นต้องใช้หลายๆตัวแปรช่วยกัน จำแนก ซึ่งคุณสมบัติเหล่านี้เป็นคุณสมบัติของ Naïve bayes และ มีความรวดเร็วในการจำแนก ข้อมูล และสามารถนำไปใช้จำแนกข้อมูลที่มีจำนวนมากได้ (Kahovi, 1996) ขั้นตอนในการสร้าง แบบจำลอง Naïve Bayes Tree มีดังต่อไปนี้

ขั้นตอนที่ 1 ทำการคำนวณหาค่า Utility ของแต่ละ Attribute ที่เป็นตัว Split ข้อมูล ซึ่งใน ข้อมูลที่เป็นตัวเลข ค่า Threshold ที่ใช้ในการแบ่งข้อมูลจะถูกคำนวณขึ้นมาโดยใช้วิธีคำนวณค่า Entropy ของ Attribute ซึ่งเป็นวิธีที่ใช้ในการสร้าง Decision Tree

ขั้นตอนที่ 2 เลือก Attribute ที่มีค่า Utility มากที่สุดมาก มาเป็น Node ในการแบ่งข้อมูล

ขั้นตอนที่ 3 ทำการทดสอบ Attribute ที่จะถูกนำมาสร้าง เป็น Node ว่าถ้า Attribute ที่นำมา สร้างเป็น Node นั้นมีค่า Utility ของ Node น้อยกว่า ค่า Utility ของ Node ปัจจุบัน แล้ว จึงทำการ สร้างตัวจำแนก Naïve Bayes ที่ โหนด นั้น

ขั้นตอนที่ 4 ทำการแบ่งข้อมูลตามค่าของข้อมูลที่อยู่ใน Attribute ที่นำมาสร้าง Node ใน แบบจำลอง ถ้า Attribute ที่นำมาสร้างเป็น Node เป็นแบบที่มีค่าที่เป็นตัวเลข ให้ทำการแบ่งข้อมูล ตามค่า Threshold ที่คำนวณออกมา สำหรับแต่ละ โหนดถูกทำตามขั้นตอนที่ 1 อีกครั้ง

ในการสร้าง NB Tree พบว่าจะมีการใช้ ค่า Utility ของ ตัวแปรในการนำมาใช้ในการ คัดเลือก Attribute ซึ่งมีค่าดังต่อไปนี้

Utility of node เป็นการคำนวณโดยการแบ่งข้อมูล และ ทำการคำนวณความแม่นยำ โดย วัดจากการนำ Node ไปทำการจำแนกข้อมูลโดยใช้ Naïve Bayes

Utility of Split เป็นการคำนวณโดยการถ่วงน้ำหนักผลรวมของ Utility of Node โดยที่ น้ำหนักที่ทำการถ่วงนั้นคือสัดส่วนของจำนวนข้อมูลที่ถูกแบ่งโดย node ต่อจำนวนข้อมูลทั้งหมด

การแบ่งข้อมูลจะถือว่ามีความสำคัญ ต่อเมื่อ การแบ่งข้อมูลนั้นทำให้ค่าความผิดพลาดของการทำนาย ลดลง มากกว่า 5 % และมีจำนวนข้อมูลที่อยู่ภายใต้ node ที่ สร้างขึ้นมามากกว่า 30 instance (Kahovi, 1996)

2.2.4 Naïve Bayes Classifier เป็นวิธีการจำแนกข้อมูล โดยใช้ทฤษฎีของเบย์ส์มาใช้ในการจำแนกข้อมูล โดยนำเอา ค่าของข้อมูลที่อยู่ใน แต่ละ Attribute มาทำการหาความน่าจะเป็นแบบมีเงื่อนไขที่มีต่อ Class ของข้อมูล ซึ่งสามารถเขียนเป็นสมการทางคณิตศาสตร์ได้ดังต่อไปนี้

$$V_{MAP} = \arg \max_{v_j \in V} P(v_j | a_1, a_2, \dots, a_n)$$

$$V_{MAP} = \arg \max_{v_j \in V} \frac{P(a_1, a_2, \dots, a_n | v_j) P(v_j)}{P(a_1, a_2, \dots, a_n)}$$

$$V_{MAP} = \arg \max_{v_j \in V} P(a_1, a_2, a_3, \dots, a_n | v_j) P(v_j)$$

$$V_{MAP} = \arg \max_{v_j \in V} P(v_j) \prod_i P(a_i | v_j)$$

โดยที่	$V_{MAP}$	แทน ค่าความน่าจะเป็นแบบมีเงื่อนไขที่มากที่สุดที่ส่งผลต่อผลเฉลย
	$a_1, \dots, a_n$	แทน ค่าของข้อมูลที่อยู่ในแต่ละ Attribute
	$V_j$	แทน ผลเฉลยของข้อมูล
	$P(V_j)$	แทน ความน่าจะเป็นของผลเฉลยของข้อมูล

ตัวอย่างแสดงการคำนวณ หาค่าความน่าจะเป็นแบบมีเงื่อนไขของ Naïve Bayes Tree โดยสมมุติข้อมูลตามตารางที่ 1 ถึง 5

**ตารางที่ 1** จำนวนข้อมูลที่อยู่ภายใต้กิ่งของต้นไม้ที่จำแนกด้วย Naive Bayes แยกตาม สมาชิกใน Attribute Humidity และ ผลเฉลยของข้อมูลว่าพื้นที่เลี้ยงสัตว์ปีกดังกล่าวเกิดโรคไข้หวัดนก(Positive)หรือไม่ (Negative)

Attribute Humidity	Positive(0.9)	Negative(0.1)
Hu1	1.000	1.000
Hu2	11.477	1.290
Hu3	1.000	1.000
Total	13.477	3.290

**ตารางที่ 2** จำนวนข้อมูลที่อยู่ภายใต้กิ่งของต้นไม้ที่จำแนกด้วย Naive Bayes แยกตาม สมาชิกใน Attribute Landscape และ ผลเฉลยของข้อมูลว่าพื้นที่เลี้ยงสัตว์ปีกดังกล่าวเกิดโรคไข้หวัดนก (Positive)หรือไม่ (Negative)

Attribute (Land Scape)	Positive (0.9)	Negative (0.1)
S1	1.000	1.000
S2	1.000	1.000
S3	1.000	1.000
S4	1.000	1.000
S5	1.000	1.000
S6	11.477	1.290
S7	1.000	1.000
S8	1.000	1.000
S9	1.000	1.000
S10	1.000	1.000
Total	20.477	10.290

ตารางที่ 3 จำนวนข้อมูลที่อยู่ภายใต้กิ่งของต้นไม้ที่จำแนกด้วย Naive Bayes แยกตาม สมาชิกใน Attribute Height และ ผลเฉลยของข้อมูลว่าพื้นที่เลี้ยงสัตว์ปีกดังกล่าวเกิดโรคไข้หวัดนก (Positive) หรือ ไม่ (Negative)

Attribute (Height)	Positive(0.9)	Negative(0.1)
H1	11.477	1.290
H2	1.000	1.000
H3	1.000	1.000
H4	1.000	1.000
H5	1.000	1.000
H6	1.000	1.000
H7	1.000	1.000
H8	1.000	1.000
H9	1.000	1.000
H10	1.000	1.000
Total	20.477	10.290

ตารางที่ 4 จำนวนข้อมูลที่อยู่ภายใต้กิ่งของต้นไม้ที่จำแนกด้วย Naive Bayes แยกตาม สมาชิกใน Attribute Temperature และ ผลเฉลยของข้อมูลว่าพื้นที่เลี้ยงสัตว์ปีกดังกล่าวเกิดโรคไข้หวัดนก (Positive) หรือ ไม่ (Negative)

Attribute (Temperature)	Positive (0.9)	Negative (0.1)
T1	1.000	1.000
T2	1.000	1.000
T3	11.47	1.290
T4	1.000	1.000
Total	14.47	4.290

ตารางที่ 5 จำนวนข้อมูลที่อยู่ภายใต้กิ่งของต้นไม้ที่จำแนกด้วย Naive Bayes แยกตาม สมาชิกใน Attribute Death และ ผลเฉลยของข้อมูลว่าพื้นที่เลี้ยงสัตว์ปีกดังกล่าวเกิดโรคไข้หวัดนก (Positive Class) หรือ ไม่ (Negative Class)

Attribute (Death)	Positive (0.9)	Negative (0.1)
D1	1.000	1.000
D2	1.000	1.000
D3	1.000	1.000
D4	9.550	1.290
Total	12.550	4.290

จากตารางที่ 1 ถึง ตารางที่ 5 เป็นผลที่ได้จากการสร้างแบบจำลองของต้นไม้ตัดสินใจ ด้วย Naïve Bayes โดยที่ในตารางที่ 1 ถึงตารางที่ 5 ตรงแถวแรกของตารางที่อยู่ในวงเล็บ แสดงค่าความน่าจะเป็นพื้นฐาน ของผลเฉลยของข้อมูลที่เป็นสมาชิกภายใต้แต่ละกิ่งของต้นไม้ตัดสินใจซึ่งได้จากการที่โปรแกรมทำการคำนวณให้ ส่วนข้อมูลภายในตารางแสดงให้เห็นถึงจำนวนข้อมูลที่เป็นสมาชิกของกิ่งต้นไม้ตัดสินใจโดยจะแยกข้อมูลตาม สมาชิกที่อยู่ในแต่ละ Attribute กับผลเฉลยของข้อมูล

สมมติว่า จากการสร้างแบบจำลองด้วย Naïve Bayes Tree ได้กิ่งของต้นไม้ที่ประกอบด้วย attribute ดังต่อไปนี้คือ Humidity = Hu 2, Landscape = S6, Height = H1, Temperature = T3, Death = D4 จากสูตรการคำนวณของ Naïve Bayes

$$V_{MAP} = \arg \max_{v_j \in V} P(v_j) \prod_i P(a_i | v_j) \quad (13)$$

โดย i แทนจำนวน Attribute ซึ่งเท่ากับ 5

j แทนจำนวนผลเฉลยของข้อมูล ซึ่งเท่ากับ 2 ซึ่งประกอบด้วย  $V_1 = \text{Positive}$

$V_2 = \text{Negative}$

ดังนั้น สามารถเขียนได้ว่า

$$V_{MAP} = \arg \max_{v_1 \in V} P(v_1) \prod_{i=5} P(a_i | v_1)$$

$$V_{Map} = \arg \max_{v_1 \in V} P(v_1) P(a_1 = Hu2 | v_1) P(a_2 = S6 | v_1)$$

$$P(a_3 = H1 | v_1) P(a_4 = T3 | v_1) P(a_5 = D4 | v_1)$$

แทนค่าลงไปในสูตร

$$V_{v_1} = 0.9 * 0.851 * 0.56 * 0.56 * 0.793 * 0.76$$

$$V_{v_1} = 0.144755443$$

$$V_{v_2} = 0.1 * 0.392 * 0.125 * 0.125 * 0.3 * 0.3$$

$$V_{v_2} = 0.000055125$$

จากสูตร

$$V_{MAP} = \arg \max_{v_j \in V} P(v_j) \prod_i P(a_i | v_j)$$

ดังนั้น

$$V_{MAP} = 0.144755443$$

คำนวณหาความน่าจะเป็นโดยที่

$$P(V_{Map}) = \frac{0.144755443}{0.144755443 + 0.000055125}$$

$$P(V_{Map}) = 0.99961933$$

เพราะฉะนั้นสรุปได้ว่า เมื่อ กิ่งของต้นไม้ที่สร้างจาก Naïve Bayes Tree ประกอบด้วย Humidity = Hu2, Landscape = S6, Height = H1, Temperature = T3, Death = D4 จะมีความน่าจะเป็นเกิดโรคไข้หวัดนกเท่ากับ 0.9996

2.3 การแบ่งกลุ่มข้อมูล (Clustering Data) เป็นการนำข้อมูลมาจัดกลุ่มซึ่งคล้ายกับการจำแนกข้อมูล เพียงแต่จะไม่มีข้อกำหนดข้อมูลตั้งต้นในการแบ่งกลุ่ม (Unsupervised Learning) ไว้ก่อน และใช้ความเหมือนกันของข้อมูลเป็นตัวแบ่งว่าข้อมูลนี้ควรอยู่ในกลุ่มไหน ตัวอย่างอัลกอริทึมที่ใช้ในการแบ่งกลุ่มข้อมูล เช่น K-Mean algorithm และ Fuzzy C-Mean Algorithm เป็นต้น

2.4 การทำนายข้อมูล (Prediction) เป็นการนำเอาข้อมูลมาสร้างฟังก์ชันเพื่อเป็นแบบจำลองที่ใช้ในการทำนายที่มีความผิดพลาดของการทำนายน้อยที่สุด อัลกอริทึม ในประเภทนี้ได้แก่การสร้างแบบจำลองด้วยวิธีการ Regression

### 3. ทฤษฎีการเลือกคุณลักษณะของข้อมูล

การคัดเลือกคุณลักษณะของข้อมูล (Attribute selection) เป็นเทคนิคพื้นฐานของการเตรียมข้อมูลสำหรับการทำเหมืองข้อมูล ในการคัดเลือกคุณลักษณะข้อมูลที่มีความสัมพันธ์กัน และมีอิทธิพลต่อผลเฉลย เพื่อนำมาสร้างแบบจำลองที่มีประสิทธิภาพในการจำแนกข้อมูล หรือ การทำนายข้อมูล โดยทั่วไปเราสามารถจำแนกเทคนิคการคัดเลือกคุณลักษณะของข้อมูล ออกเป็น 2 ประเภทได้แก่ การจัดอันดับคุณลักษณะ (Feature Ranking) และ การคัดเลือกกลุ่มคุณลักษณะของข้อมูล (Subset Selection)

3.1 การจัดอันดับคุณลักษณะ (Feature Ranking) เป็นการนำเอาแต่ละคุณลักษณะที่มีในข้อมูลมาทำการจัดอันดับ ซึ่งในการจัดอันดับใช้คะแนนของแต่ละคุณลักษณะในการจัดอันดับคุณลักษณะ แล้วเลือกเอาคุณลักษณะที่มีอันดับที่ดีไปใช้ในการทำเหมืองข้อมูล เช่นนำไปสร้างแบบจำลองเพื่อใช้ในการจำแนกข้อมูล เป็นต้น (Witten and Frank, 2005) ขั้นตอนในการจัดอันดับคุณลักษณะมีขั้นตอนดังต่อไปนี้

ขั้นตอนที่ 1 คำนวณหาคะแนนของแต่ละคุณลักษณะที่มีอยู่ในข้อมูล

ขั้นตอนที่ 2 ทำการเรียงอันดับคุณลักษณะของข้อมูลโดยเรียงตามลำดับคะแนนจากมากไปหาน้อย

วิธีการในการคำนวณที่ใช้ในการหาคะแนนมีดังต่อไปนี้

3.1.1 Pearson Correlation Coefficient เป็นการวัดความสัมพันธ์ร่วมกันระหว่างคุณลักษณะกับผลเฉลยของข้อมูล โดยมักจะใช้กับข้อมูลที่มีลักษณะเป็นตัวเลข ซึ่งสูตรที่ใช้ในการคำนวณ คือ

$$R(i) = \frac{\sum_{k=1}^m (X_{k,i} - \bar{X})(Y - \bar{Y})}{\sqrt{\sum_{k=1}^m (X_{k,i} - \bar{X})^2 \sum_{k=1}^m (Y - \bar{Y})^2}} \quad (14)$$

โดยที่	i	แทน index ของแต่ละคุณลักษณะของข้อมูล
	k	แทน index ของข้อมูลที่อยู่ภายใต้ชุดข้อมูล
	$X_{k,i}$	แทน ค่าของสมาชิกที่อยู่ภายใต้คุณลักษณะของข้อมูลที่ถูกนำมาหา Correlation
	$\bar{X}$	แทน ค่าเฉลี่ยของสมาชิกที่อยู่ภายใต้คุณลักษณะของข้อมูลที่ถูกนำมาหา Correlation
	Y	แทน ผลเฉลยของข้อมูล
	$\bar{Y}$	แทน ค่าเฉลี่ยของผลเฉลยของข้อมูล

ค่า R(i) จะมีค่าอยู่ที่ -1 ถึง 1 ซึ่งสามารถแปลความหมายได้ถึงว่า คุณลักษณะที่หามานี้มี ความสัมพันธ์ร่วมกับคลาสที่เป็น Negative และ Positive มากน้อยเพียงใด

3.1.2 Information theoretic Ranking Criteria เป็นการจัดอันดับของคุณลักษณะของข้อมูลโดยการวัดค่าของการขึ้นต่อกัน ระหว่างคุณลักษณะกับผลเฉลยของข้อมูล ซึ่งมีสูตรในการคำนวณดังต่อไปนี้ (Witten and Frank, 1996)

$$I(i) = \sum_{X_j} \sum_Y P(X = x_i, Y = y) \log \frac{P(X = x_i, Y = y)}{P(X = x_i)P(Y = y)} \quad (15)$$

โดยที่  $P(X=x_i, Y=y)$  แทนความน่าจะเป็นร่วมเมื่อ  $X=x_i$  และ  $Y=y$

$P(X=x_i)$  แทนความน่าจะเป็นเมื่อ  $X = x_i$

$P(Y=y)$  แทนความน่าจะเป็นเมื่อ  $Y=y$

$Y$  แทนผลเฉลยของชุดข้อมูล

$X$  แทนคุณลักษณะของข้อมูล

$x_i$  แทนสมาชิกที่อยู่ภายใต้คุณลักษณะของข้อมูล

การหาคะแนนของแต่ละคุณลักษณะด้วยวิธีนี้ไม่เหมาะกับชุดข้อมูลที่มีลักษณะเป็นตัวเลข ซึ่งถ้าจะทำการวัดด้วยวิธีนี้ต้องทำการแบ่งช่วงข้อมูล ก่อนที่จะมาเข้าสู่ตรรกะคำนวณดังกล่าว

3.2 การเลือกกลุ่มคุณลักษณะของข้อมูล (Subset Selection) เป็นการประเมินกลุ่มของคุณลักษณะของข้อมูล กลุ่มใดที่เหมาะสมต่อการนำไปใช้ในการสร้างแบบจำลอง อัลกอริทึมของการเลือกกลุ่มของคุณลักษณะสามารถแบ่งได้เป็น 3 ประเภท คือ Wrapper, Filter และ Embed (Witten and Frank, 1996)

3.2.1 Wrapper เป็นวิธีที่นำการเรียนรู้ของเครื่องจักรมาใช้ในการให้คะแนนตามความสามารถในการทำนายแต่ละกลุ่มย่อยของคุณลักษณะของข้อมูล

3.2.2 Filter เป็นวิธีที่ทำการคัดเลือกกลุ่มย่อยของคุณลักษณะในช่วงที่ทำการ Preprocess Data ซึ่งกลุ่มของคุณลักษณะจะเป็นอิสระต่ออัลกอริทึมที่จะนำมาใช้ในการจำแนก

3.2.3 Embed เป็นวิธีการคัดเลือกคุณลักษณะของข้อมูลที่เกิดขึ้นในช่วงของการนำเอาอัลกอริทึมมาฝึกกับชุดข้อมูลสำหรับ Training ซึ่งคุณลักษณะที่เกิดขึ้นมานั้นจะขึ้นอยู่กับอัลกอริทึมที่นำมาใช้ในการ Training ของอัลกอริทึม

#### 4. การคำนวณน้ำหนักให้กับข้อมูล

การปรับปรุงประสิทธิภาพ ของแบบจำลองจากที่ได้ค้นคว้ามีหลายวิธี แต่วิธีที่น่าสนใจ และทำการศึกษามาไว้ในงานวิจัยนี้มีวิธีการถ่วงน้ำหนักเพื่อปรับปรุงประสิทธิภาพจำนวน 3 วิธีซึ่งประกอบด้วย

4.1 Decision making Trial and Evaluation Laboratory (DEMATEL) และ Analytic Network Processing (ANP) เป็นวิธีการสำหรับการสร้างแบบจำลอง และ วิเคราะห์ความสัมพันธ์องค์ประกอบของปัญหาที่เกิดขึ้น ผู้ที่นำเสนอคือ Gabus และ Fontela ในปี 1973 (Wu, 2008) ส่วน Analytic Network Processing ผู้ที่นำเสนอคือ Thomus L Saaty ในปี 1996 การทำงานของ DEMATEL และ ANP มีดังต่อไปนี้

ขั้นตอนที่ 1 Generating the direct –relation matrix ทำการวัดระดับความสัมพันธ์ระหว่างองค์ประกอบของปัญหา โดยทำการให้คะแนนในแต่ละองค์ประกอบที่มีความสัมพันธ์กัน โดยถ้ามีความสัมพันธ์ต่อกันมาก ให้ 3 คะแนน แต่ถ้าไม่มีความสัมพันธ์จะให้ 0 คะแนน

ขั้นตอนที่ 2 Normalizing the direct-relation matrix หลังจากผ่านขั้นตอนที่ 1 จะได้ matrix direct-relation สมมุติให้เป็น Matrix A นำ matrix A ไปทำ Normalize ด้วยสมการ

$$X = K . A$$

$$K = \frac{1}{\max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij}}, i, j = 1, 2, 3, \dots, n \quad (16)$$

โดยที่ A คือ Direct Relation matrix

X คือ Normalizing Direct Relation matrix

$a_{i,j}$  คือ สมาชิกใน Direct Relation Matrix

K คือ ค่าคงที่สำหรับการ Normalized direct – relation Matrix

ขั้นตอนที่ 3 Attaining the total-relation matrix ทำการสร้าง the total-Relation matrix (T) โดยสมการ

$$T = X(1 - X)^{-1} \quad (17)$$

โดยที่

T คือ Total-relation matrix

X คือ Normalizing Direct Relation Matrix

ขั้นตอนที่ 4 Producing a Causal Diagram ทำการรวมค่าระหว่างแถว กับ คอลัมน์ n ของ Matrix T เพื่อทำการสร้าง Causal Diagram โดยให้ผลรวมตามแนว Row เป็น ค่า D และ ค่าตามแนว คอลัมน์ n เป็นค่า R ถ้าความสัมพันธ์ในปัจจัยที่มีค่าผลต่างระหว่างค่า D และ ค่า R เป็นค่าบวก แสดงว่า ปัจจัยนั้นเป็นเหตุ แต่ถ้าปัจจัยไหนที่มีค่าผลต่างระหว่าง ค่า D และ R มีค่าน้อยกว่าศูนย์ แสดงว่าปัจจัยที่กำลังพิจารณานี้เป็นผลที่เกิดขึ้น โดยที่ค่า D และ R มีสูตรในการคำนวณดังต่อไปนี้

$$T = [t_{i,j}]_{n \times n} \quad i, j = 1, 2, 3, \dots, n \quad (18)$$

$$D = \left[ \sum_{j=1}^n t_{i,j} \right]_{n \times 1} = [t_i]_{n \times 1} \quad (19)$$

$$R = \left[ \sum_{i=1}^n t_{i,j} \right]_{1 \times n} = [t_j]_{n \times 1} \quad (20)$$

โดยที่ T แทน Total –Relation Matrix

$T_{ij}$  แทน สมาชิกใน Total – Relation Matrix

I แทนของสมาชิกของ Total Relation Matrix ตามแนว คอลัมน์

J แทนลำดับของสมาชิกของ Total Relation Matrix ตามแนว แถว

ขั้นตอนที่ 5 Inner Dependence Matrix ในขั้นตอนนี้นำเอาค่าที่มากที่สุด ใน คอลัมน์ total-relation matrix มาหารกับทุกค่าที่อยู่ภายใต้ คอลัมน์ เดียวกัน ซึ่ง จะได้ ผลรวมตามแนว คอลัมน์ ของ total-relation matrix เท่ากับ 1 ดังนั้นจะได้ inner dependence matrix ออกมา

ขั้นตอนที่ 6 สร้าง Unweighted Supermatrix นำ Inner dependence Matrix ของ แต่ละกลุ่ม ของปัจจัยที่ได้ นำมาสร้างเป็น Unweighted Supermatrix

ขั้นตอนที่ 7 สร้าง Limited Super Matrix นำ Unweighted Super Matrix ไปสร้างเป็น Limit Supermatrix ซึ่งเป็น Matrix ที่แสดงถึงน้ำหนักของแต่ละองค์ประกอบของปัญหา

4.2 วิธีการคำนวณน้ำหนักโดยใช้ต้นไม้ตัดสินใจ เป็นวิธีที่นำเอาต้นไม้ตัดสินใจ และวิธีการ Bagging มาใช้ในการหาน้ำหนักของคุณลักษณะที่อยู่ในชุดข้อมูลซึ่งมีขั้นตอนดังต่อไปนี้ (Hall, 2006)

ขั้นตอนที่ 1 กำหนดจำนวนรอบของการสร้างต้นไม้ตัดสินใจเพื่อใช้ในการคำนวณหาน้ำหนักให้กับคุณลักษณะของข้อมูล

ขั้นตอนที่ 2 กำหนดจำนวนข้อมูลที่ใช้ในการสร้างต้นไม้ตัดสินใจ โดยกำหนดเป็น % จากจำนวนข้อมูลที่อยู่ในชุดฝึก โดยที่ข้อมูลที่ถูกนำมาใช้นั้นจะถูกสุ่มมาจากชุดข้อมูลที่ใช้สำหรับฝึก และ ทำการสุ่มทุกครั้งในแต่ละรอบที่ทำการสร้างต้นไม้ตัดสินใจ

ขั้นตอนที่ 3 ทำการสร้างต้นไม้ตัดสินใจโดยที่ต้นไม้ที่สร้างมานั้นยังไม่ได้ทำการ ตัดแต่ง

ขั้นตอนที่ 4 เมื่อสร้างต้นไม้ตัดสินใจได้แล้ว ตรวจสอบว่า มีคุณลักษณะใดที่ไม่ปรากฏในต้นไม้ตัดสินใจ ถ้าพบ ให้นำน้ำหนักใน คุณลักษณะนั้น เท่ากับ 0

ขั้นตอนที่ 5 ในกรณีที่ในต้นไม้ตัดสินใจนั้นมี คุณลักษณะปรากฏให้ทำการวัดค่าความ ความลึก (d) น้อยสุดจาก โหนด ราก จนถึง โหนดที่มี คุณลักษณะดังกล่าวปรากฏอยู่ และเมื่อหาได้ แล้วให้นำไปเข้าสู่ตรรกการคำนวณน้ำหนัก

$$Weight = \frac{1}{\sqrt{d}} \quad (21)$$

โดยที่  $d$  คือ ความสูงน้อยสุดจาก โหนดราก จนถึง โหนดที่มีคุณลักษณะที่กำลังหาน้ำหนักปรากฏอยู่

ขั้นตอนที่ 6 คำนวณน้ำหนักของคุณลักษณะที่มีปรากฏอยู่ใน ต้นไม้ตัดสินใจที่สร้างจนครบทุกคุณลักษณะ

ขั้นตอนที่ 7 ทำการสร้างต้นไม้ตัดสินใจ และ ทำการคำนวณน้ำหนักของแต่ละคุณลักษณะของต้นไม้ตัดสินใจที่สร้างขึ้นในแต่ละต้นจนกว่าจะครบรอบของการสร้างต้นไม้ตัดสินใจที่กำหนดไว้ในขั้นตอนที่ 1

ขั้นตอนที่ 8 หาค่าเฉลี่ยของน้ำหนักในแต่ละคุณลักษณะที่ได้มาจากต้นไม้ตัดสินใจที่สร้างขึ้น โดยที่จำนวนต้นไม้เท่ากับจำนวนรอบที่กำหนดไว้ในขั้นตอนที่ 1

ขั้นตอนที่ 9 สรุปว่าแต่ละคุณลักษณะมีน้ำหนักเท่าไร แล้วนำไปใช้ในการปรับปรุงประสิทธิภาพของแบบจำลอง

4.3 การคำนวณหาน้ำหนักของคุณลักษณะที่อยู่ในข้อมูลโดยใช้วิธี Emerging Pattern เป็นการนำเอา Item set ที่อยู่ในชุดข้อมูลที่มีค่าความน่าจะเป็นอยู่ใน คลาสใดคลาสหนึ่ง มากกว่าอีกคลาสหนึ่งอย่างมีนัยสำคัญ นำมาคำนวณหาน้ำหนักให้กับข้อมูลแต่ละ Instance โดยมีวิธีการคำนวณดังต่อไปนี้ (Alhammady, 2006)

ขั้นตอนที่ 1 แบ่งข้อมูลออกเป็นสองชุดตามผลเฉลยของข้อมูล

ขั้นตอนที่ 2 ทำการค้นหา Item set ของข้อมูลที่อยู่ในแต่ละชุดข้อมูล และ ค้นหา ค่า Support ของแต่ละ Item set จากสูตร

$$S_{D_i}(X) = \frac{\text{Count}_{D_i}(X)}{|D_i|} \quad (22)$$

โดยที่  $S_{D_i}(X)$  แทน ค่า Support ของ Itemset ที่มี ต่อชุด  $D_i$   
 $\text{Count}_{D_i}(X)$  แทน จำนวนข้อมูลที่อยู่ในชุด  $D_i$  ที่มี Item set  $X$  ปรากฏอยู่  
 $|D_i|$  แทน จำนวนข้อมูลที่อยู่ในชุด  $D_i$  โดยที่  $i = 1,2,3,4\dots n$

ขั้นตอนที่ 3 หาค่า Growth Rate ให้กับแต่ละ Item set ที่ได้มาจากแต่ละชุดข้อมูล  
 ค่า Growth Rate คือสัดส่วนของค่า Support ของแต่ละ item set ที่มีต่อ ผลเฉลยของข้อมูล  
 ซึ่งแสดงถึง อำนาจในการจำแนกข้อมูล ของ Item set โดยสูตรการคำนวณคือ

$$Gr_{D_i \rightarrow D_j}(X) = \begin{cases} 0, & \text{if } (S_i(X) = 0 \cap S_j(X) = 0) \\ \infty, & \text{if } (S_i(X) = 0 \cap S_j(X) \neq 0) \\ \frac{S_j}{S_i}, & \text{Otherwise} \end{cases} \quad (23)$$

โดยที่  $Gr_{D_i \rightarrow D_j}(X)$  แทนค่า Growth Rate ระหว่าง ชุดข้อมูลที่แบ่งจากขั้นตอนที่ 1  
 $S_i(X)$  แทนค่า Support ของ Item set ที่คำนวณมาจากชุดข้อมูล  $I$  ที่แบ่งเมื่อขั้นตอนที่ 1  
 $S_j(X)$  แทนค่า Support ของ Item set ที่คำนวณมาจากชุดข้อมูล  $J$  ที่แบ่งเมื่อขั้นตอนที่ 1  
 เมื่อหาค่า Growth Rate ได้แล้วทำการคัดเลือก Item set โดยทำการเลือก Item set ที่มีค่า  
 $Gr > 1$  ซึ่ง Item set ที่ผ่านการคัดเลือกจะเรียกว่า Emerging Pattern เรียกว่า  $e$

ขั้นตอนที่ 4 หาค่า Strength ของ Emerging Pattern โดยมีสูตรในการคำนวณคือ

$$\alpha_{C_j}(e) = \frac{gr_{C_j}(e)}{1 + gr_{C_j}(e)} * S_{C_j}(e) \quad (24)$$

โดยที่  $C_j$  แทนกลุ่มผลเฉลยของข้อมูล

$gr_{C_j}(e)$  แทน ค่า Growth Rate ของ Emerging Pattern ที่ปรากฏในชุดข้อมูลที่มีผลเฉลย อยู่ใน Class  $C_j$

$S_{C_j}(e)$  แทนค่า Support ของ Emerging Pattern ที่ปรากฏอยู่ในชุดข้อมูลที่มีผลเฉลย อยู่ใน Class  $C_j$

ขั้นตอนที่ 5 นำ ค่า Strength ของแต่ละ Emerging pattern ที่หามาใส่ลงไปในแต่ละ Instance ที่มี Emerging Pattern นั้น ปรากฏอยู่ และหาผลรวมของ ค่า Strength ของ Emerging Pattern ทั้งหมดที่ปรากฏอยู่ใน แต่ละ Instances โดยสูตรการคำนวณคือ

$$\beta_{C_j}(i) = \sum_{e \subseteq i, e \in E_{C_j}} \alpha_{C_j}(e) \quad (25)$$

โดยที่  $E_{C_j}$  แทน set ของ Emerging Pattern ที่เป็นของชุดข้อมูลที่มีผลเฉลยอยู่ใน คลาส  $C_j$

$I$  แทน instance ที่มีผลเฉลยอยู่ในคลาส  $C_j$

ขั้นตอนที่ 6 นำผลรวมของ Emerging Pattern ที่หามาจากขั้นตอนที่ 5 ของแต่ละ Instance มาหาค่าน้ำหนักของแต่ละ Instance โดยหาจากสูตร

$$w_{C_j}(i) = \frac{\beta_{C_j}(i)}{\text{median } C_j} \quad (26)$$

โดยที่  $\beta_{C_j}(i)$  แทน ผลรวมค่า Strength ของ Emerging Pattern ที่พบอยู่ใน Instance  $i$

Median  $C_j$  แทน ค่ามัธยฐานของผลรวมค่า Strength ของ Emerging Pattern

ขั้นตอนที่ 7 นำน้ำหนักไปใช้ในการสร้างแบบจำลอง

## 5. ทฤษฎีการวัดประสิทธิภาพของแบบจำลอง

5.1 Precision เป็นการวัดในเรื่องความแม่นยำในการจำแนกข้อมูลว่าข้อมูลที่ถูกละกอริทึมจำแนกว่าอยู่ในคลาสมีสัดส่วนเป็นเท่าไรจากจำนวนข้อมูลที่มีผลเฉลยจริง อยู่ในคลาส A

$$precision = \frac{tp}{tp + fp} \quad (27)$$

โดยที่ tp แทนค่า True Positive ซึ่งหมายถึงจำนวนข้อมูลทั้งหมดที่จำแนก คลาสถูก  
fp แทนค่า False Positive ซึ่งหมายถึง จำนวนข้อมูลทั้งหมดที่ถูกจำแนก ผิดคลา  
แต่ผลเฉลยของข้อมูลอยู่ในคลาสเดียวกัน

5.2 Recall เป็นการวัดสัดส่วนความถูกต้องระหว่างข้อมูลที่จำแนกและตรงกับผลเฉลยว่าอยู่ในคลาส A กับ ข้อมูลทั้งหมดที่ได้รับการจำแนกว่าอยู่ในคลาส A

$$recall = \frac{tp}{tp + fn} \quad (28)$$

โดยที่ tp แทนค่า true positive ซึ่งหมายถึง จำนวนข้อมูลทั้งหมดที่จำแนกว่าอยู่ ในคลาส A และผลเฉลยของข้อมูลก็อยู่ในคลาส A  
fn แทนค่า False Negative ซึ่งหมายถึง จำนวนข้อมูลทั้งหมดที่ถูกจำแนกว่าอยู่ในคลาส A แต่ผลเฉลยอยู่ในคลาส B

5.3 F-measure เป็นค่ากลางที่ใช้ในการวัดประสิทธิภาพที่ใช้เป็นตัวแทนในการเปรียบเทียบประสิทธิภาพ กับ อัลกอริทึมตัวอื่น แทนการใช้ค่า Precision และ Recall เพราะถ้าใช้ค่า Precision และ Recall จะไม่สามารถสรุปได้ว่า อัลกอริทึมไหนมีประสิทธิภาพดีกว่ากัน สูตรในการคำนวณ ค่า F – measure มีดังต่อไปนี้

$$F - measure = \frac{2 * (precision * recall)}{(precision + recall)} \quad (29)$$

## งานวิจัยที่เกี่ยวข้องกับวิทยานิพนธ์

จากการศึกษางานวิจัยในอดีตที่ผ่านมาที่เกี่ยวข้องกับวิทยานิพนธ์ฉบับนี้สามารถจัดกลุ่มการศึกษางานวิจัยในอดีตที่ได้ไปศึกษามาเป็น 3 กลุ่มได้แก่ งานวิจัยที่เกี่ยวกับลักษณะและ การระบาดของโรคไข้หวัดใหญ่ ปอดอักเสบ และ โรคไข้หวัดนก งานวิจัยที่เกี่ยวข้องกับการสร้างแบบจำลองการระบาด ของโรคต่างๆที่เกิดขึ้น และ งานวิจัยที่เกี่ยวข้องกับการหาหน้าหนัก และการถ่วงน้ำหนักของข้อมูล

### 1. งานวิจัยที่เกี่ยวกับลักษณะและ การระบาดของโรคไข้หวัดใหญ่ ปอดอักเสบ และโรคไข้หวัดนก

จากการศึกษางานวิจัยในอดีตที่ผ่านมาพบว่า มีงานวิจัยจำนวนมากที่ได้ทำการศึกษาเกี่ยวกับลักษณะและการระบาดของโรคไข้หวัดนก ไข้หวัดใหญ่ และปอดอักเสบ เช่น งานวิจัยที่ศึกษาเกี่ยวกับ การระบาดของโรคไข้หวัดใหญ่ และ โรคปอดอักเสบในพื้นที่ของประเทศสหรัฐอเมริกา ฝรั่งเศส และ ออสเตรเลีย (Vibound, 2002) โดยศึกษาถึงความสัมพันธ์ของพื้นที่ และ ช่วงเวลา ระหว่างประเทศที่เป็นเป้าหมายทั้งสามประเทศ ซึ่งผลจากการศึกษาคือ มี 2 ประเทศที่มีลักษณะการระบาดคล้ายคลึงกัน ทั้งพื้นที่ และช่วงเวลาที่มีการระบาดมากที่สุดได้แก่ ประเทศอเมริกา และฝรั่งเศส ซึ่งแบบจำลองที่นำมาใช้ในงานวิจัยเรื่องนี้คือ Regression Model

สำหรับในประเทศไทยงานวิจัยที่ศึกษาเกี่ยวกับการศึกษาลักษณะของการระบาดของโรคไข้หวัดนกชนิดรุนแรงในปี พ.ศ. 2547 – พ.ศ. 2549 (วีรพงษ์ และ ถนอม, 2550) เพื่อแสดงให้เห็นถึงความสัมพันธ์ของปัจจัยที่ส่งผลต่อการระบาดของโรคไข้หวัดนก โดยนำ รายงานสัตว์ปีกป่วยตาย มาทำการวิเคราะห์

ซึ่งผลสรุปจากการศึกษานี้คือปัจจัยที่มีความเสี่ยงสูงต่อการเกิดโรคไข้หวัดนกประกอบด้วย ลักษณะพื้นที่ภาคกลาง ช่วงฤดูฝน และ ฤดูหนาว ชนิดสัตว์ที่เลี้ยงได้แก่ ไก่พื้นเมือง และเป็ด

## 2. งานวิจัยที่เกี่ยวข้องกับการสร้างแบบจำลองการระบาด ของโรคต่างๆที่เกิดขึ้น

จากการศึกษางานวิจัยในอดีตที่ผ่านมาพบว่า มีงานวิจัยจำนวนมากที่ได้สร้างแบบจำลองการระบาดของโรคต่างๆที่เกิดขึ้น เช่น การสร้างแบบจำลอง Bayesian Network เพื่อใช้ในการเฝ้าระวังการระบาดของโรค Anthrax ที่เกิดขึ้นจากการก่อการร้าย (Cooper *et al.*, 2004) ซึ่ง Bayesian Network ที่นำเสนอมานั้น ประกอบด้วย Sub Bayesian Network 2 ส่วน ได้แก่ Bayesian Network ส่วนที่สร้างจากข้อมูลผู้ป่วย และ ส่วนที่เป็น คุณลักษณะที่ส่งผลโดยตรงต่อการติดเชื้อ Antrax ของผู้ป่วย และ คุณลักษณะที่ส่งผลต่อประชากรทุกคน เมื่อได้แบบจำลองแล้วจึงนำไปสร้างเป็น โปรแกรมที่ใช้ในการตรวจจับการปล่อยเชื้อ Antrax โดยวัดประสิทธิภาพจากความเร็วและความถูกต้องในการตรวจจับ การปล่อยเชื้อ Antrax ซึ่งเครื่องมือที่ใช้วัดประสิทธิภาพของโปรแกรม คือ AMOC Curve ซึ่งในการวัดประสิทธิภาพของ โปรแกรมนี้ จะวัดในสถานะที่มีปริมาณเชื้อ Antrax ในอากาศ แตกต่างกัน ซึ่งพบว่า ประสิทธิภาพในการ ตรวจจับการปล่อยเชื้อ Antrax ขึ้นอยู่กับ ปริมาณเชื้อในอากาศ และ ต่อมามีการปรับปรุงประสิทธิภาพของแบบจำลอง โดยเพิ่มเรื่องทิศทางลมของการปล่อยเชื้อ Antrax ทำให้ประสิทธิภาพการตรวจจับของ โปรแกรมที่สร้างจากแบบจำลองดีขึ้น, การศึกษาเกี่ยวกับการหาลักษณะประชากรที่มีความเสี่ยงที่จะเป็นโรคไข้หวัดใหญ่ และ ปอดอักเสบ (Abbas *et al.*, 2004) ซึ่งวิธีการศึกษาคือ การนำข้อมูลลักษณะของประชากร มาสร้างเป็นแบบจำลองด้วยวิธีข่ายงานแบบเบย์ ซึ่งผลที่ได้จากการศึกษาคือ สามารถนำเอา Bayesian Network มาใช้ในการระบุลักษณะของประชากรที่มีความเสี่ยงสูงที่จะเป็นโรคไข้หวัดใหญ่ และ ปอดอักเสบ ซึ่งประโยชน์ทางด้านสาธารณสุข คือ สามารถจัดการทรัพยากรทางด้านสาธารณสุขที่มีอยู่อย่างจำกัดได้อย่างมีประสิทธิภาพ และ ลักษณะประชากรที่รายได้ต่ำ มีลักษณะเชื้อชาติผิวดำ, เพศ หญิง และ อยู่ในวัยเด็ก มีความเสี่ยงสูงที่จะเป็นโรคไข้หวัดใหญ่ และ ปอด ซึ่งต่อมาได้มีนำแบบจำลองที่ได้มาจากงานวิจัยข้างต้นมาศึกษาต่อเพื่อใช้วิเคราะห์ประสิทธิภาพของการควบคุมโรคไข้หวัดใหญ่ และ โรคปอดอักเสบ ของวัคซีน ว่ามีประสิทธิภาพต่อการควบคุมโรคอย่างไร (Abbas *et al.*, 2005) โดยทำการวัดกับประชากรที่มีความเสี่ยงที่จะเป็นโรคไข้หวัดใหญ่ และ ปอดอักเสบ ซึ่งผลที่ได้คือ การใช้ มีผลต่อการลดลงของกลุ่มประชากรที่มีความเสี่ยงที่จะเป็นโรคไข้หวัดใหญ่ และ ประสิทธิภาพของวัคซีนมีผลต่อจำนวนผู้ป่วยที่เป็นโรคไข้หวัดใหญ่และ ปอดอักเสบ

ต่อมาได้มีงานวิจัยที่นำเสนอได้นำเสนอแบบจำลอง Dynamic Bayesian Network ที่สร้างจากชุดข้อมูลที่ใช้ในการเฝ้าระวังโรคไข้หวัดใหญ่ (Sebastiani *et al.*, 2005) ที่มีลักษณะ Time Series จำนวน 4 ข้อมูล ได้แก่ ข้อมูลจำนวนผู้ป่วยเด็ก และ ผู้ใหญ่ ที่ป่วยเป็นระบบทางเดินหายใจที่ถูกส่งตัวไปที่แผนกฉุกเฉินทางระบบทางเดินหายใจเป็นรายสัปดาห์

ข้อมูลจำนวนผู้ป่วยโรคไข้หวัดใหญ่เป็นรายสัปดาห์ และ ข้อมูลจำนวนผู้เสียชีวิตด้วยโรคไข้หวัดใหญ่ และ ปอดอักเสบ เพื่อนำแบบจำลองที่ได้ไปใช้ในการค้นหาสัญญาณเบื้องต้นของกระบวนของโรคไข้หวัดใหญ่ ซึ่งในการสร้างแบบจำลองจะเริ่มต้นจากการหา Regression model ให้กับข้อมูล และทำการเลือก regression model ของชุดข้อมูลทั้งสิ้นชุด มาสร้างเป็นแบบจำลอง Dynamic Bayesian Network ผลที่ได้คือ ทำให้ทราบว่า จำนวนผู้ป่วยเด็กที่ป่วยเป็นโรกระบบทางเดินหายใจ ที่ถูกส่งไปแผนก ฉุกเฉิน เป็นตัวชี้วัดเบื้องต้นการระบาดของโรคไข้หวัดใหญ่ และเป็นการแสดงให้เห็นว่า แบบจำลอง Bayesian Network สามารถนำมาใช้ในการเฝ้าระวังโรคไข้หวัดใหญ่ได้

ต่อมาได้มีงานวิจัยที่นำเสนอแบบจำลองต้นไม้ตัดสินใจ (Arithmetic Decision tree) เพื่อใช้ประเมิน ค่าใช้จ่ายในการรักษาของผู้ป่วยเมื่อเกิดการระบาดของโรคไข้หวัดใหญ่ (Hak *et al.* , 2006) โดยใช้ข้อมูลค่าใช้จ่ายในการรักษาโรคไข้หวัดใหญ่ ที่มีการระบาดในอดีต ซึ่งผลที่ได้พบว่า ค่าใช้จ่ายที่เกิดขึ้นในการรักษาโรคไข้หวัดใหญ่ ในแต่ละช่วงอายุ ซึ่งช่วงอายุที่เสียค่าใช้จ่ายในการรักษาโรคไข้หวัดใหญ่สูงที่สุดคือ ช่วงอายุ 20-64 ปี

จากงานวิจัยที่ได้ศึกษามาข้างต้นสามารถสังเกตได้ว่า แบบจำลองนั้นถูกสร้างจากการนำเอาอัลกอริทึม Bayesian Network, Decision Tree มาใช้ในการสร้างแบบจำลอง แต่จากการศึกษา งานวิจัยที่ผ่านมาพบว่าการนำเสนอแบบจำลองที่ไม่ได้สร้างจาก อัลกอริทึม Bayesian Network, Decision Tree แต่นำเอาปัจจัยทางด้านสภาพแวดล้อมมาสร้างเป็นสมการความน่าจะเป็นทางคณิตศาสตร์ เพื่อนำมาใช้ในการคำนวณหาค่าความน่าจะเป็นของพื้นที่ที่มีโอกาสเกิดการระบาดของโรคไข้หวัดนก และ นำความน่าจะเป็นที่ได้ไปใช้ในการทำนายจำนวนพื้นที่ที่เกิดการระบาดของโรคไข้หวัดนก ซึ่งงานวิจัยแรกที่ได้ค้นคว้าและมีการนำเสนอแบบจำลองที่ไม่ได้สร้างจากอัลกอริทึม Bayesian Network และ Decision Tree ได้แก่งานวิจัยที่นำเสนอแบบจำลองการระบาดของโรคไข้หวัดนกโดยใช้ตัวแปรทางด้านสภาพของพื้นที่ ของประเทศสาธารณรัฐประชาชนจีน และ ระยะเวลาที่มีการระบาด (Li *et al.*, 2004) เพื่อใช้ทำนายว่า ในช่วงระยะเวลาที่มีการระบาดของโรค

ไข้หวัดนกเกิดขึ้นในประเทศจีน จำนวนพื้นที่ที่เกิดการระบาดของโรคไข้หวัดนกเป็นเท่าไร ซึ่งผลที่ได้จากแบบจำลองการระบาดของโรคไข้หวัดนกคือ จำนวนพื้นที่ที่พบการระบาดของโรคไข้หวัดนก และ ความน่าจะเป็นของการระบาดของโรคไข้หวัดนกขึ้นอยู่กับ ระยะทางห่างกับจุดที่มีการระบาดของโรคไข้หวัดนก และ ระยะห่างจากเส้นทางคมนาคม ประโยชน์ แบบจำลองนี้คือ สามารถนำไปใช้ในการกำหนดนโยบายในการเฝ้าระวังและควบคุมโรคไข้หวัดนกที่เหมาะสม ได้

ต่อมาได้มีการนำเอาแบบจำลองที่ไม่ได้สร้างจากอัลกอริทึม Bayesian Network และ Decision Tree ไปทำการศึกษาและพัฒนาต่อเช่น สร้างแบบจำลองการแพร่ระบาดของโรคไข้หวัดนกซึ่งแสดงให้เห็นถึงความสัมพันธ์ระหว่าง เวลาที่มีการระบาด และ ลักษณะพื้นที่ โดยสร้างเป็นสมการความน่าจะเป็น เพื่อคำนวณถึงโอกาสที่จะเกิดการระบาดของโรคไข้หวัดนกในพื้นที่ๆทำการประเมิน ซึ่งผลของการทำวิจัยนี้ทำให้ทราบว่า ความน่าจะเป็นของการระบาดของโรคไข้หวัดนกแปรผกผันกับระยะทางและช่วงเวลาของการระบาด ซึ่งเป็นผลมาจากช่วงชีวิตของเชื้อไวรัส และ อุณหภูมิ (Upadhyay *et al.*, 2007) และ การนำเสนอ Transmission Model ที่ใช้ในการทำนายการระบาดของโรคไข้หวัดนกในคน (Prapaphan and Anongnart, 2007) โดยตัวแปรที่ใช้ในการสร้างแบบจำลองนี้ได้แก่ อุณหภูมิ ระยะเวลาที่มีการระบาด จำนวนประชากรที่มีความเสี่ยงที่จะเป็น โรคไข้หวัดนก ระดับความรุนแรงของการระบาดของโรคไข้หวัดนก และ ประสิทธิภาพของวัคซีน โดยผลที่ได้จากการนำเอาแบบจำลองนี้ไปใช้คือ ความน่าจะเป็นที่จะเกิดการระบาดของโรคไข้หวัดนกในคน ผลกระทบที่เกิดขึ้นในกรณีที่มีการใช้วัคซีน กับ ไม่มีการใช้วัคซีน ซึ่งผลคือ ในกรณีที่มีการใช้วัคซีนจะทำให้ความน่าจะเป็นที่จะทำให้เกิดการระบาดของโรคไข้หวัดนกลดลง

### 3. งานวิจัยที่เกี่ยวข้องกับการหาน้ำหนัก และการถ่วงน้ำหนักของข้อมูล

จากการศึกษางานวิจัยในอดีตในเรื่องการหาน้ำหนักของข้อมูล และ การถ่วงน้ำหนักของข้อมูล พบว่ามีวิธีการในถ่วงน้ำหนัก และ นาน้ำหนักให้กับข้อมูลหลายวิธี แต่ที่ได้ทำการศึกษาครั้งนี้ การศึกษาถึงวิธีการที่เหมาะสมในการค้นหา ยุทธศาสตร์ในการจัดทำ การจัดการความรู้ ที่เหมาะสม โดยใช้วิธีการ Analytic Network Process (ANP) (Wu, 2008) ซึ่งเป็นวิธีที่ใช้ในการค้นหา ยุทธศาสตร์ในการจัดทำ การจัดการความรู้ แต่เนื่องด้วย ANP มีข้อเสียในเรื่องของไม่สามารถรองรับเรื่องการขึ้นต่อกัน ของ องค์ประกอบ ในการเลือก ยุทธศาสตร์ในการจัดทำ การจัดทำความรู้ จึงต้องนำเอา Decision Making Trial and Evaluation Laboratory (DEMATEL) ซึ่งเป็นวิธีการในการศึกษาสร้าง และ วิเคราะห์แบบจำลองที่แสดงความสัมพันธ์ที่เป็นเหตุเป็นผลกัน

ระหว่าง องค์ประกอบที่เป็นส่วนในการเลือกกลยุทธ์ในการจัดการความรู้ และยังสามารถจัดการปัญหาการขึ้นต่อกันขององค์ประกอบในการเลือก ยุทธศาสตร์ในการจัดทำ การจัดการความรู้ซึ่งในการศึกษานี้ได้ใช้เหตุการณ์สมมุติในการศึกษา

ซึ่งผลการศึกษางานวิจัยนี้ได้สรุปว่า การค้นหา ยุทธศาสตร์ในการจัดการความรู้ ด้วย ANP และ DEMATEL นั้นสามารถแก้ปัญหการขึ้นต่อกันของปัจจัยที่ใช้ในการเลือก ยุทธศาสตร์ในการจัดการความรู้ ทำให้สามารถเลือก ยุทธศาสตร์ในการจัดการความรู้ ที่มีประสิทธิภาพอันส่งผลในเรื่องของ การ Implementing KM ให้ประสบความสำเร็จ

งานวิจัยที่ได้ทำการศึกษาดังกล่าวคือ การถ่วงน้ำหนักข้อมูลโดยใช้ต้นไม้ตัดสินใจ และวิธีการ Bagging (Hall, 2007) ในการหาค่าน้ำหนักสำหรับแต่ละคุณลักษณะที่อยู่ในชุดข้อมูล ซึ่งเมื่อได้กลุ่มต้นไม้ตัดสินใจแล้ว จะทำการวัดความลึกน้อยสุดของระหว่าง โหนดราก กับ โหนดของคุณลักษณะที่ต้องการหาค่าน้ำหนัก ในแต่ละต้นไม้ตัดสินใจที่ถูกสร้างขึ้น ต่อมาเมื่อได้ความลึกของคุณลักษณะของข้อมูล จึงเข้าสู่ตรรกการคำนวณเพื่อหาค่าถ่วงน้ำหนักของคุณลักษณะ และ เมื่อหาค่าน้ำหนักของคุณลักษณะจากต้นไม้ตัดสินใจได้ครบทุกต้นแล้ว จึงนำค่าน้ำหนักของคุณลักษณะที่หามาจากต้นไม้ตัดสินใจแต่ละต้นที่ได้สร้างขึ้น มาหาค่าเฉลี่ย ซึ่งจะได้เป็นค่าน้ำหนักของแต่ละคุณลักษณะ และ นำน้ำหนักที่ได้มานำไปใช้กับอัลกอริทึมสำหรับการจำแนกข้อมูลซึ่งในงานวิจัยนี้คือ อัลกอริทึม Naïve Bayes สำหรับ ข้อมูลที่ถูกนำมาใช้ในงานวิจัยดังกล่าวนี้

เป็นข้อมูลเฉพาะที่ใช้สำหรับการทดสอบอัลกอริทึมจาก UCI Respository มีจำนวนชุดข้อมูลทั้งสิ้น 28 ชุดข้อมูล ซึ่งผลการทดลองจากการนำเอา Naïve Bayes ที่ผ่านการถ่วงน้ำหนักพบว่ามีประสิทธิภาพในการจำแนกข้อมูลดีกว่าอัลกอริทึมดีว่าการจำแนกข้อมูล โดยใช้อัลกอริทึมอื่นๆ

งานวิจัยที่ได้ทำการศึกษาเกี่ยวกับการถ่วงน้ำหนักข้อมูล และการหาน้ำหนักข้อมูล งานสุดท้ายคืองานวิจัยที่นำเสนอวิธีการสร้างต้นไม้ตัดสินใจ โดยใช้ค่า Emerging Pattern (EP) ในการนำมาหาค่าน้ำหนักของข้อมูล (Alhammady, 2006) และนำมาใช้ในสูตรสำหรับการสร้างต้นไม้ตัดสินใจ และทำการเปรียบเทียบประสิทธิภาพของแบบจำลองที่สร้างด้วยอัลกอริทึมต้นไม้ตัดสินใจ แบบธรรมดา กับแบบที่นำเรื่องของ Emerging Pattern มาใช้สร้างต้นไม้ตัดสินใจ พบว่าแบบจำลองที่นำเอาเรื่อง Emerging Pattern มาใช้นั้นมีประสิทธิภาพในการจำแนกดีกว่า และสามารถทนต่อ Error ที่อยู่ในชุดข้อมูลที่นำมาใช้ในการสร้างแบบจำลองการจำแนกข้อมูล

จากงานวิจัยที่ได้นำเสนอจะพบว่า แต่ละงานวิจัยที่มีการศึกษาเกี่ยวกับโรคไข้หวัดนกนั้น มุ่งเน้นไปที่ เรื่องของ การสร้างแบบจำลองเพื่อทดสอบสมมุติฐาน , เพื่อหารูปแบบการกระจายของโรคไข้หวัดนก แต่ยังไม่มีการนำเอาข้อมูลโรคไข้หวัดนก มาหาปัจจัยและความสัมพันธ์เชิงพื้นที่ที่ทำให้เกิดการระบาดของโรคไข้หวัดนก แล้วนำไปสร้างแบบ ซึ่งจะมีประโยชน์ในด้านของการเฝ้าระวังโรคไข้หวัดนกในประเทศไทยได้

ตารางที่ 6 เปรียบเทียบงานวิจัยที่เกี่ยวกับการสร้างแบบจำลองของโรคไข้หวัดใหญ่ ปอดอักเสบ และ โรคไข้หวัดนก

ชื่อ	ผู้แต่ง/ปีที่พิมพ์	สิ่งที่ต้องการนำเสนอ	โครงสร้างแบบจำลอง	ผลลัพธ์
Computational Epidemiology: Bayesian Disease surveillance	Kaja Abbas <i>et. al</i> /2004	ต้องการหาปัจจัยทางลักษณะประชากรที่ส่งผลทำให้เกิดการระบาดของโรคปอดอักเสบ และ ไข้หวัดใหญ่ โดยใช้ Bayesian Net work ในการหาลักษณะของประชากร	แบบจำลอง Bayesian Network ที่ถูกสร้างขึ้น ประกอบด้วยคุณลักษณะ ดังต่อไปนี้ เชื้อชาติ, เพศ, อายุ, รายได้, อาการป่วยที่พบ (ไอ, มีไข้), โรคที่พบ	Bayesian network สามารถนำไปใช้ในการระบุจุดที่สำคัญในการควบคุม ป้องกัน และ เฝ้าระวังการติดเชื้อ ของโรคไข้หวัดใหญ่ และ ปอดอักเสบได้
Temporal Analysis of infectious Disease: Influenza	Kaja Abbas, <i>et. al</i> /2005	นำเอา BN และ SIR Epidemic Curve มาใช้ในการวิเคราะห์ถึงประสิทธิภาพของวัคซีนที่ส่งผลต่อการระบาดของโรคไข้หวัดใหญ่	แบบจำลอง Bayesian Network ที่ถูกสร้างขึ้น ประกอบด้วยคุณลักษณะ ดังต่อไปนี้ เชื้อชาติ, เพศ, อายุ, รายได้, อาการป่วยที่พบ (ไอ, มีไข้), โรคที่พบ และ ประสิทธิภาพของวัคซีน	ลักษณะของประชากรที่มีความเสี่ยงในการเกิดโรคไข้หวัดใหญ่คือ ประชากรที่เป็นเด็ก ,เชื้อชาติ สเปน และ โปรตุเกส และ Asian ซึ่งเมื่อให้ Vaccine ความเสี่ยงต่อการติดโรคลดลงอย่างมากตามประสิทธิภาพของ วัคซีน
Bayesian Biosurveillance of Disease Outbreaks	Gregory F Cooper, <i>et. al</i> / 2004	นำเสนอการวิจัยเกี่ยวกับการนำ Bayesian Network ไปใช้ในการสร้างแบบจำลองในเชิงระยะเวลา และ พื้นที่ ของ โรคไม่ติดต่อ โดยสร้างจากข้อมูลจำนวนประชากรซึ่งงานวิจัยนี้กล่าวถึงโรค Anthrax	แบบจำลอง Bayesian Network ที่สร้างขึ้น ประกอบด้วย 3 ส่วน ได้แก่ Bayesain Network ย่อย ที่สร้างจากข้อมูลผู้ป่วยโรค Anthrax, ส่วนที่ 2 คือ ส่วนที่แสดงถึงปัจจัยที่มีผลกระทบต่อคนในพื้นที่ที่ปล่อยเชื้อ และ ส่วนที่ 3 แสดงถึงผลกระทบ	ได้แบบจำลองที่สามารถนำไปพัฒนาเป็นระบบตรวจจับการปล่อยเชื้อ Anthrax ที่มีประสิทธิภาพ โดยวัดจากความเร็ว และ ความถูกต้องในการตรวจจับ และ ทำการปรับปรุงประสิทธิภาพให้ดีขึ้น โดยเพิ่มเรื่องทิศทางการปล่อยเชื้อ

ตารางที่ 6 (ต่อ)

ชื่อ	ผู้แต่ง/ปีที่พิมพ์	สิ่งที่ต้องการนำเสนอ	โครงสร้างแบบจำลอง	ผลลัพธ์
Modelling the health-economic impact of the next influenza pandemic in the netherlands	E. HAK, <i>et. al</i> /2006	พัฒนาแบบจำลองสำหรับการตัดสินใจ และ ประเมิน ภาระทางด้านสุขภาพ และ ต้นทุนทางการแพทย์ที่เกิดขึ้นเมื่อเกิดการระบาดของโรคไข้หวัดใหญ่ครั้งต่อไป ซึ่ง ประเมินจาก ภาระพื้นฐานในการดูแลสุขภาพในช่วงการระบาดของโรคไข้หวัดใหญ่ในอดีต	แบบจำลอง Arithmetic Decision tree ที่ถูกสร้างขึ้นประกอบด้วยคุณลักษณะดังต่อไปนี้ ช่วงอายุ, ความเสี่ยงที่จะเป็นโรคไข้หวัดใหญ่, วิธีการรักษาโรคที่เกี่ยวข้องกับระบบทางเดินหายใจ และ โรคระบบทางเดินหายใจที่ทำการรักษา	Arithmetic Decition Tree สามารถนำไปใช้ในการประมาณการค่าใช้จ่ายในการรักษาของประชากรได้ ซึ่งพบว่ากลุ่มประชากรที่มีค่ารักษาสูงสุดคือกลุ่มประชากรที่อยู่ในช่วงอายุ น้อยกว่า 65 ปี
Influenza Epidemics in the United States, France, and Australia, 1972 -1997	Cecile Viboud , <i>et. al</i> / 2002	ต้องการหาความเหมือนของการระบาดของไข้หวัดใหญ่ในสามประเทศได้แก่ ประเทศ อเมริกา, ฝรั่งเศส และ ออสเตรเลีย โดยวัดเรื่องอัตราการเสียชีวิตด้วยโรคไข้หวัดใหญ่ และ ช่วงเวลาที่มีการเสียชีวิตด้วยโรคไข้หวัดใหญ่สูงสุด	นำจำนวนผู้เสียชีวิตด้วยโรคปอดอักเสบ และ ไข้หวัดใหญ่ ใน แต่ละสัปดาห์ ตั้งแต่ปี 1972 – 1997 ของทั้งสามประเทศมาหา อัตราการเสียชีวิต ที่เพิ่มขึ้นโดยใช้ Regression Model มาช่วยในการหาอัตราการเสียชีวิต และ นำข้อมูลเดียวกันนี้ไปหา ช่วงเวลาที่เกิดการเสียชีวิตสูงที่สุด	ในมุมมองของ ความสัมพันธ์ของการระบาดทั้ง 3 ประเทศพบว่า ประเทศ อเมริกา และ ฝรั่งเศส มีความสัมพันธ์กัน ในเชิงของ จำนวนผู้เสียชีวิตที่เพิ่มขึ้น เป็นไปในทิศทางเดียวกัน และ ในมุมมองของช่วงเวลาที่มียาจำนวนผู้ป่วยที่เสียชีวิตด้วยโรคไข้หวัดใหญ่และ ปอดอักเสบ สูงที่สุดเหมือนกันคือ ประเทศ อเมริกา และ ฝรั่งเศส

## ตารางที่ 6 (ต่อ)

ชื่อ	ผู้แต่ง/ปีที่พิมพ์	สิ่งที่ต้องการนำเสนอ	โครงสร้างแบบจำลอง	ผลลัพธ์
Modeling the spread of bird flu and predicting outbreak diversity	Ranjit kumar Upadhyay, <i>et. al</i> /2007	นำเสนอแบบจำลอง Transmission model ในรูปแบบสมการความน่าจะเป็น เพื่อใช้ในการจำลองจำนวนพื้นที่ที่เกิดการระบาดของโรคไข้หวัดนก และ เพื่อให้เกิดความเข้าใจในเรื่องของการระบาดของโรคไข้หวัดนก	องค์ประกอบของแบบจำลองประกอบด้วย จำนวนพื้นที่ที่เกิดการระบาด, ช่วงชีวิตของไวรัสซึ่งสัมพันธ์กับอุณหภูมิในอากาศ, ระยะเวลาการระบาดของโรค และ ระยะทางระหว่างจุดที่ทำการประเมิน กับ จุดที่มีการระบาดอยู่ในขณะนี้	แบบจำลองที่ได้แสดงให้เห็นว่าความน่าจะเป็นของการเกิดไข้หวัดนกในแต่ละพื้นที่ จะแปรผกผัน กับระยะทาง และ ช่วงเวลาที่มีการระบาด
Adaptive Statistical Transmission Model of AVIAN Influenza for Outbrake Prediction in Human in thailand	Prapaphan Prawing, Anongnart Srivihok /2007	แบบจำลองการแพร่ระบาดของ โรคไข้หวัดนก ในคน โดยลักษณะของแบบจำลองเป็นแบบสมการความน่าจะเป็น เพื่อนำไปใช้ในการจำลองข้อมูลผู้ป่วยโรคไข้หวัดนกในคน และ เข้าในการแพร่ระบาดของโรคไข้หวัดนกในคน	แบบจำลอง การแพร่ระบาดของโรคไข้หวัดนกในคนประกอบด้วยคุณลักษณะดังต่อไปนี้ จำนวนการระบาดของ โคนกไข้หวัดนกในคน, ช่วงระยะเวลาที่มีการระบาด, ช่วงชีวิตของเชื้อไวรัสที่สัมพันธ์กับอุณหภูมิ, ความต้านทานโรคของประชากร และ ประสิทธิภาพของวัคซีน	แบบจำลองที่ได้ออกมาแสดงให้เห็นถึง ประสิทธิภาพของวัคซีน ที่แตกต่างกัน และ ช่วงเวลาของการระบาด มีผลต่อ ความน่าจะเป็นของการเกิดการระบาดของโรคไข้หวัดนก ในคนแบบผกผัน

ตารางที่ 6 (ต่อ)

ชื่อ	ผู้แต่ง/ปีที่พิมพ์	สิ่งที่ต้องการนำเสนอ	โครงสร้างแบบจำลอง	ผลลัพธ์
A Bayesian Dynamic Model for Influenza Surveillance	Paola sebastiani, et. al /2005	Dynamic Bayesian Network มาใช้หาความสัมพันธ์รวมกับของชุดข้อมูลที่ใช้ในการเฝ้าระวังโรคไข้หวัดนก จำนวน 4 ชุดข้อมูล เพื่อนำความสัมพันธ์ที่ได้มาวิเคราะห์หาตัวชี้วัดเบื้องต้นที่ทำให้ผู้เกี่ยวข้องกับโรคทางทราบถึงการเริ่มต้นการระบาดของโรคระบบทางเดินหายใจ	Dynamic Bayesian Network ถูกสร้างมาจาก Regression Model ของชุดข้อมูลที่ใช้ในการเฝ้าระวังโรคไข้หวัดใหญ่อันประกอบด้วย จำนวนเด็ก และ ผู้ใหญ่ ต่อสัปดาห์ที่ป่วยเป็นโรคทางเดินหายใจ และถูกส่งไปยังแผนกฉุกเฉินของ โรคระบบทางเดินหายใจ และ จำนวนผู้ป่วยที่มีอาการคล้ายโรคไข้หวัดใหญ่ และ จำนวนผู้ป่วยที่เสียชีวิตด้วยโรคไข้หวัดใหญ่และปอดอักเสบ	จากแบบจำลองที่ได้ทำให้ทราบว่า ตัวชี้วัดเริ่มแรกสุดที่ทำให้ทราบถึงการระบาดของโรคไข้หวัดใหญ่ได้แก่ จำนวนผู้ป่วยเด็กที่ป่วยเป็นโรคทางเดินหายใจที่ถูกส่งไปแผนกฉุกเฉินโรคระบบทางเดินหายใจ และ Dynamic Bayesian Network เป็นเครื่องมือที่เหมาะสมสำหรับการเฝ้าระวังโรคไข้หวัดใหญ่
Study on Transmission Model of Avian Influenza	Jingping Li, et. al /2004	แบบจำลองการแพร่ระบาดของโรคไข้หวัดนก ในประเทศจีน โดยลักษณะของแบบจำลองเป็นแบบสมการความน่าจะเป็น เพื่อนำไปใช้ในการจำลองข้อมูลพื้นที่การระบาดของโรคไข้หวัดนก และ อธิบายการแพร่ระบาดของโรคไข้หวัดนก	แบบจำลอง การแพร่ระบาดของโรคไข้หวัดนกในคนประกอบด้วยคุณลักษณะดังต่อไปนี้ จำนวนพื้นที่ที่เกิดการระบาดของโรคไข้หวัดนก จำนวนพื้นที่ที่ได้รับการควบคุมโรคไข้หวัดนก ช่วงชีวิตของเชื้อไวรัส ความต้านทานเชื้อไวรัส และ ระยะห่างของพื้นที่ที่เกิดการระบาด กับพื้นที่โดยรอบ	จำนวนพื้นที่ที่พบการระบาดของโรคไข้หวัดนก และ ความน่าจะเป็นของการระบาดของโรคไข้หวัดนกขึ้นอยู่กับ ระยะทางห่างกับจุดที่มีการระบาดของโรคไข้หวัดนก และ ระยะห่างจากเส้นทางคมนาคม ประโยชน์ แบบจำลองนี้คือสามารถนำไปใช้ในการกำหนดนโยบายในการเฝ้าระวังและควบคุมโรคไข้หวัดนกที่เหมาะสมได้

จากตารางที่ 6 ได้แสดงให้เห็นถึงการสร้างแบบจำลองที่แตกต่างกัน 3 ชนิด ได้แก่ แบบจำลองที่สร้างจาก Bayesian Network , Decision tree , Regression Model และ สมการความน่าจะเป็น ซึ่งในการนำเสนอแบบจำลองขึ้นอยู่กับลักษณะข้อมูลที่นำมาใช้ในการศึกษา ตัวอย่างเช่น ถ้าข้อมูลมีลักษณะที่อยู่รูปของอนุกรม เวลา (Time Series) มักสร้างแบบจำลองด้วยวิธีการ Regression Model ก่อน แล้วจึงนำไปวิเคราะห์

สำหรับข้อมูลที่นำมาใช้ในการสร้างแบบจำลองของโรคไข้หวัดใหญ่ ปอดอักเสบ และ โรคไข้หวัดนก มักใช้ข้อมูลจำนวนผู้ป่วย จำนวนผู้เสียชีวิต และ ลักษณะโดยทั่วไปของผู้ป่วย และ พื้นที่ เช่น ลักษณะภูมิประเทศ อุณหภูมิ จำนวนช่วงเวลาในการระบาด ช่วงชีวิตของเชื้อไวรัสที่เป็นต้นกำเนิดของโรค ที่นำมาสร้างเป็นแบบจำลอง เป็นต้น มาใช้ในการสร้างแบบจำลอง เพื่อนำแบบจำลองที่ได้ไปใช้ในการอธิบายความสัมพันธ์ของปัจจัยที่ทำให้เกิดการระบาด นำไปใช้ในการจำลอง สถานการณ์การระบาดในอนาคต และ นำไปใช้ในการพิสูจน์ถึงประสิทธิภาพของเครื่องมือที่ใช้ในการควบคุมการระบาดของโรคไข้หวัดใหญ่ ปอดอักเสบ และ ไข้หวัดนก

จากงานวิจัยที่ได้กล่าวมาทั้งหมดสามารถนำไปเป็นข้อมูลเพื่อประกอบการเลือกวิธีการสร้างแบบจำลองที่เหมาะสมต่อข้อมูลที่นำมาใช้ในการศึกษานี้ ซึ่งได้แก่ Bayesian Network, Decision Tree, Naïve Bayes Tree

## อุปกรณ์และวิธีการ

### อุปกรณ์

อุปกรณ์ที่นำมาใช้ในงานวิจัยประกอบด้วย

เครื่องคอมพิวเตอร์ Note Book RAM 2 GB Hardisk 120 GB ระบบปฏิบัติการ  
Window XP Service Pack 3  
Weka Version 3.7.0  
Eclipse IDE V4.3  
โปรแกรม สถิติ SPSS V16

### ข้อมูล

ข้อมูลที่นำมาใช้ในการทำวิจัยนี้ประกอบด้วย

1. ข้อมูลการระบาดของโรคไข้หวัดนกในประเทศไทย ที่เกิดขึ้นตั้งแต่ปี 2547-2550 โดยนำข้อมูลมาจาก กรมปศุสัตว์ จำนวน 8,152 ระเบียบ แบ่งเป็น 2 กลุ่มคือ

1.1 ข้อมูลที่ยืนยันว่าเป็นพื้นที่มีสัตว์ปีกป่วยตายที่เกิดการระบาดของโรคไข้หวัดนก จำนวน 1,351 ระเบียบ

1.2 ข้อมูลที่ยืนยันว่าเป็นพื้นที่มีสัตว์ปีกป่วยตาย แต่ไม่ได้เกิดจากเชื้อโรคไข้หวัดนก จำนวน 6,801 ระเบียบ

ในชุดข้อมูลดังกล่าวมี Attribute ได้แก่ วันที่รายงาน, จังหวัดที่พบ, อำเภอ, หมู่บ้าน, ชนิดสัตว์ปีกที่พบ, จำนวนสัตว์ปีกที่เลี้ยง, จำนวนสัตว์ปีกที่ป่วย, จำนวนสัตว์ปีกที่ตาย, จำนวนสัตว์ปีกที่ถูกทำลาย และ ผลการทดสอบว่าเป็นพื้นที่ระบาดหรือไม่ใช่พื้นที่ระบาด

ตารางที่ 7 ตัวอย่างข้อมูลที่ได้จากรายงานสัตว์ปีกป่วยตาย ตั้งแต่ปี 2004 – 2007

วันที่รายงาน	จังหวัด	อำเภอ	ตำบล	หมู่ที่	ชนิดสัตว์	จำนวนสัตว์(ตัว)				ผลตรวจ
						ทั้งหมด	ป่วย	ตาย	ทำลาย	
1/9/2005	สระบุรี	หนองแซง	หนองกบ	1	ไก่พื้นเมือง	15	15	12	3	+
5/9/2005	สระบุรี	หนองแค	โพนทอง	4	ไก่พื้นเมือง	45	10	10	35	+
5/9/2005	สระบุรี	หนองแค	โพนทอง	4	ไก่พื้นเมือง	48	12	12	36	+
6/9/2005	สระบุรี	หนองแค	โพนทอง	5	ไก่พื้นเมือง	75	55	55	20	+
12/9/2005	กำแพงเพชร	คลองขลุง	วังบัว	10	ไก่พื้นเมือง	68	15	15	53	+
12/9/2005	กำแพงเพชร	คลองขลุง	วังบัว	8	ไก่พื้นเมือง	80	2	1	79	+
12/9/2005	กำแพงเพชร	คลองขลุง	วังบัว	8	ไก่พื้นเมือง	91	10	10	81	+
17/9/2005	กำแพงเพชร	คลองขลุง	วังบัว	5	ไก่เนื้อ	19,600	2,602	2,102	17,498	+
28/9/2005	กำแพงเพชร	เมืองกำแพงเพชร	ไตรตรังษ์	11	ไก่พื้นเมือง	5	1	1	4	+
28/9/2005	กำแพงเพชร	เมืองกำแพงเพชร	ไตรตรังษ์	11	ไก่พื้นเมือง	100	83	83	17	+
28/9/2005	กำแพงเพชร	เมืองกำแพงเพชร	ไตรตรังษ์	14	ไก่พื้นเมือง	35	33	30	5	+

หมายเหตุ เครื่องหมายบวก หมายถึง พื้นที่ที่มีการระบาดของโรคไขหวัดนก (Positive Class)

เครื่องหมายลบ หมายถึง พื้นที่ที่ไม่มีการระบาดของโรคไขหวัดนก (Negative Class)

ที่มา : กรมปศุสัตว์ กระทรวงเกษตรและสหกรณ์ (พ.ศ. 2550)

2. ข้อมูลความชื้น และ ชั่วโมงที่มีแสงแดด ตั้งแต่ปี 2004-2007 รวบรวมจาก กรมอุตุนิยมวิทยา มี Attribute ได้แก่ ชื่อศูนย์ในการวัดสภาพภูมิอากาศ, วันเดือนปีที่บันทึก, จำนวน ชั่วโมงที่มีแสงแดดต่อวัน และ ความชื้น

stn_name	stncode	year	month	dday	sunshine	avgrh
Nakhon Sawan*	400201	2004	1	1	9.6	65
Nakhon Sawan*	400201	2004	1	2	9.6	65
Nakhon Sawan*	400201	2004	1	3	9.5	64
Nakhon Sawan*	400201	2004	1	4	9.5	65
Nakhon Sawan*	400201	2004	1	5	9.3	66
Nakhon Sawan*	400201	2004	1	6	9.2	67
Nakhon Sawan*	400201	2004	1	7	9.5	65
Nakhon Sawan*	400201	2004	1	8	9.6	61
Nakhon Sawan*	400201	2004	1	9	8.6	66
Nakhon Sawan*	400201	2004	1	10	4.6	70
Nakhon Sawan*	400201	2004	1	11	5.1	68
Nakhon Sawan*	400201	2004	1	12	1.6	72
Nakhon Sawan*	400201	2004	1	13	3.7	78
Nakhon Sawan*	400201	2004	1	14	9	69
Nakhon Sawan*	400201	2004	1	15	8.7	70
Nakhon Sawan*	400201	2004	1	16	9.2	68
Nakhon Sawan*	400201	2004	1	17	8.7	72
Nakhon Sawan*	400201	2004	1	18	7.7	71
Nakhon Sawan*	400201	2004	1	19	9.2	68
Nakhon Sawan*	400201	2004	1	20	9.7	67
Nakhon Sawan*	400201	2004	1	21	6.1	68
Nakhon Sawan*	400201	2004	1	22	2.4	61
Nakhon Sawan*	400201	2004	1	23	8.7	65
Nakhon Sawan*	400201	2004	1	24	8.8	58
Nakhon Sawan*	400201	2004	1	25	8.9	55
Nakhon Sawan*	400201	2004	1	26	10.2	61
Nakhon Sawan*	400201	2004	1	27	10.1	61
Nakhon Sawan*	400201	2004	1	28	10	60
Nakhon Sawan*	400201	2004	1	29	9.3	65
Nakhon Sawan*	400201	2004	1	30	9.3	66
Nakhon Sawan*	400201	2004	1	31	8.1	69

ภาพที่ 3 แสดงตัวอย่างข้อมูล ความชื้น และ ชั่วโมงที่มีแสงแดด

ที่มา: กรมอุตุนิยมวิทยา (พ.ศ. 2550)

3. ข้อมูลอุณหภูมิในแต่ละจังหวัดของไทยตั้งแต่ปี 2004 – 2007 มี Attribute ดังต่อไปนี้  
ศูนย์สถานีตรวจอากาศของกรมอุตุนิยมวิทยา, วัน เดือน ปี ที่รายงานอุณหภูมิ, อุณหภูมิสูงสุด  
(หน่วย เซลเซียส), อุณหภูมิต่ำสุด (หน่วย เซลเซียส), อุณหภูมิเฉลี่ย (หน่วย เซลเซียส)

stn_name	stncod	year	month	dday	maxtmp	mintmp	meanter
U Thong Agromet	425301	2004	1	1	31.1	14.9	23
U Thong Agromet	425301	2004	1	2	31.6	15	23.3
U Thong Agromet	425301	2004	1	3	31.2	15.7	23.4
U Thong Agromet	425301	2004	1	4	32.2	15.5	23.8
U Thong Agromet	425301	2004	1	5	32.1	18.6	25.3
U Thong Agromet	425301	2004	1	6	31.5	16.3	23.9
U Thong Agromet	425301	2004	1	7	31.4	15.4	23.4
U Thong Agromet	425301	2004	1	8	31.5	15	23.2
U Thong Agromet	425301	2004	1	9	31.7	18	24.8
U Thong Agromet	425301	2004	1	10	31	20.2	25.6
U Thong Agromet	425301	2004	1	11	32.7	21.5	27.1
U Thong Agromet	425301	2004	1	12	25.6	21.7	23.6
U Thong Agromet	425301	2004	1	13	29.5	19.5	24.5
U Thong Agromet	425301	2004	1	14	31.7	20.8	26.2
U Thong Agromet	425301	2004	1	15	32.7	20.4	26.5
U Thong Agromet	425301	2004	1	16	32.5	20.9	26.7
U Thong Agromet	425301	2004	1	17	32.3	20.7	26.5
U Thong Agromet	425301	2004	1	18	33.4	22.8	28.1
U Thong Agromet	425301	2004	1	19	32.7	20.5	26.6
U Thong Agromet	425301	2004	1	20	32.8	19.1	25.9
U Thong Agromet	425301	2004	1	21	32.2	20.4	26.3
U Thong Agromet	425301	2004	1	22	27.8	21.6	24.7
U Thong Agromet	425301	2004	1	23	30.5	21.2	25.8
U Thong Agromet	425301	2004	1	24	28.7	19.2	23.9
U Thong Agromet	425301	2004	1	25	36.1	17.1	26.6
U Thong Agromet	425301	2004	1	26	29	15	22

ภาพที่ 4 ตัวอย่างข้อมูลอุณหภูมิของจังหวัดในประเทศไทย

ที่มา: กรมอุตุนิยมวิทยา (พ.ศ. 2550)

4. ข้อมูลความสูงจากระดับน้ำทะเลของแต่ละพื้นที่ รวบรวมจากกรมอุตุนิยมวิทยามี Attribute ดังต่อไปนี้ สถานีตรวจอากาศ, ตำแหน่งละติจูด, ตำแหน่งลองจิจูด, ความสูงจากระดับน้ำทะเล (หน่วย เมตร)

Station	STN-CODE	LATITUDE	LONGITUDE	MSL(M.)
CHIANG RAI AGROMET	303301	19.52.15	99.46.58	416
MAE JO AGROMET	327301	18.55.00	99.00.00	316
LAMPANG AGROMET	328301	18.19.00	99.17.00	314
NAN AGROMET	331301	18.52.00	100.45.00	264
SI SAMRONG AGROMET	373301	17.10.00	99.52.00	53
DOI MUSIR AGROMET	376301	16.45.00	98.56.00	860
PHICHIT AGROMET	386301	16.26.10	100.17.20	35
LOEI AGROMET	353301	17.24.00	101.44.00	258
SAKON NAKHON AGROMET	356301	17.07.00	104.03.00	191
NAKHON PHANOM AGROMET	357301	17.26.00	104.47.00	144
THAPHRA AGROMET	381301	16.20.00	102.49.00	165
ROI ET AGROMET	405301	16.04.00	103.37.00	153
UBON RATCHATHANI AGROMET	407301	15.14.00	105.02.00	130
SRISAKET AGROMET	408301	15.00.00	104.15.00	105

ภาพที่ 5 ตัวอย่างข้อมูลความสูงจากระดับน้ำทะเลของแต่ละพื้นที่ในประเทศไทย

ที่มา: กรมอุตุนิยมวิทยา (พ.ศ. 2550)

## ตัวแปรที่ใช้ในงานวิจัย

ตัวแปรที่นำมาใช้ในงานวิจัยนี้มีจำนวนทั้งสิ้น 14 ตัวแปร ซึ่งปรากฏอยู่ในตารางด้านล่าง

ตารางที่ 8 รายละเอียดของตัวแปรที่ใช้ในการวิจัย

ชื่อตัวแปร	ประเภทข้อมูล	คำอธิบาย
Month	Nominal	เดือนที่มีการพบสัตว์ปีกป่วยตาย
Land Scape	Nominal	ลักษณะพื้นที่
Height	Numeric	ความสูงของพื้นที่
Region	Nominal	ภูมิภาคของประเทศไทย
Type	Nominal	ประเภทของสัตว์ปีก
Quantity	Numeric	จำนวนสัตว์ปีกที่มีการเลี้ยงไว้
Sick	Numeric	จำนวนสัตว์ปีกที่พบอาการป่วย
Death	Numeric	จำนวนสัตว์ปีกที่ตาย
Destruction	Numeric	จำนวนสัตว์ปีกที่ถูกทำลาย
Temperature	Numeric	อุณหภูมิ
Humidity	Numeric	ความชื้น
Daylight Hour	Numeric	จำนวนชั่วโมงที่มีแสงแดดต่อวัน
Season	Nominal	ฤดูกาล

ผลเฉลยของข้อมูลประกอบด้วย พื้นที่เลี้ยงสัตว์ปีกที่ทดสอบแล้วว่าการระบาดของโรคไข้หวัดนก และ พื้นที่เลี้ยงสัตว์ปีกที่ทดสอบแล้วว่าไม่เป็นพื้นที่เกิดการระบาดของโรคไข้หวัดนก

ในการเลือกตัวแปรมาใช้ในการสร้างแบบจำลอง จะที่จะนำมาใช้สร้างแบบจำลอง เลือกมาจากการศึกษางานวิจัยในอดีตที่ได้ นำเสนอปัจจัยที่ส่งผลต่อการระบาด และนอกจากที่ใช้วิธีที่ได้กล่าวมาข้างต้นในการเลือกคุณลักษณะแล้ว ในงานวิจัยนำเอาการคัดเลือกคุณลักษณะ (Feature Selection) มาใช้ในการคัดเลือกคุณลักษณะเพื่อนำไปสร้างแบบจำลอง

## วิธีการ

ในการศึกษานี้มีวัตถุประสงค์ที่จะนำเอาข้อมูลจากรายงานสัตว์ปีก ป่วยตาย ของกรมปศุสัตว์ ข้อมูลด้านสภาพภูมิอากาศ ซึ่งประกอบด้วย อุณหภูมิ, จำนวนชั่วโมงที่มีแสงแดด, ความชื้นสัมพัทธ์ และ ข้อมูลด้านสภาพภูมิประเทศ ประกอบด้วย ความสูงจากระดับน้ำทะเล, ลักษณะของพื้นที่ เช่น พื้นที่ราบสูง เป็นต้น มาใช้ในการสร้างแบบจำลอง โดยที่เทคนิคที่นำมาใช้นั้นคือเทคนิคการทำเหมืองข้อมูล ซึ่งขั้นตอนในการทำเหมืองข้อมูลมีดังต่อไปนี้

ขั้นตอนที่ 1 กำหนดเรื่องที่ต้องการจะทำเหมืองข้อมูล

ขั้นตอนที่ 2 เลือกข้อมูลที่เกี่ยวข้องกับเรื่องที่ต้องการทำเหมืองข้อมูล เพื่อที่จะนำข้อมูลดังกล่าวไปใช้ในการทำเหมืองข้อมูล

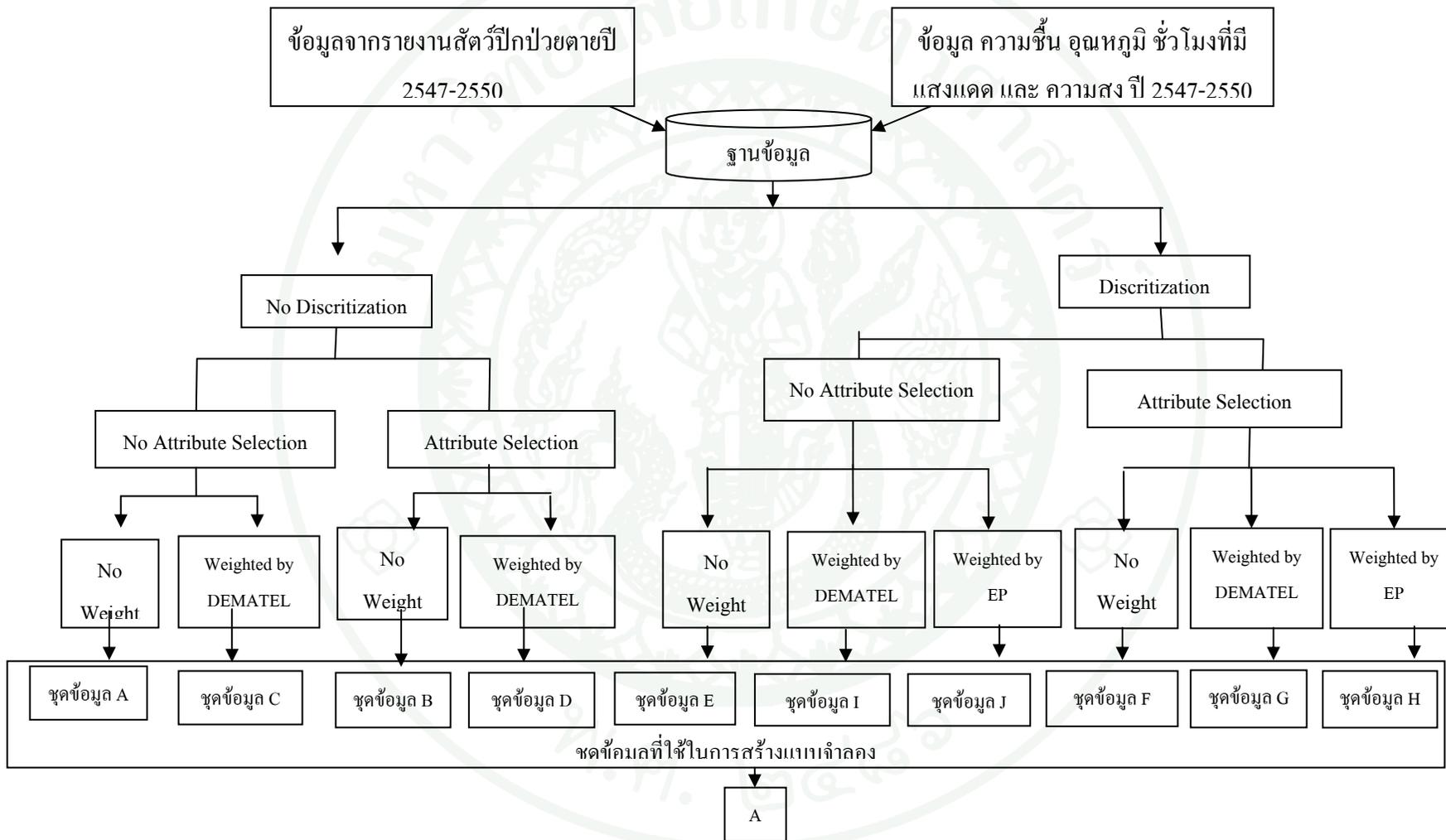
ขั้นตอนที่ 3 การเตรียมข้อมูล (Preprocess Data) คือการดำเนินการกับข้อมูลซึ่งประกอบด้วย (1) การคัดเลือกคุณลักษณะของข้อมูล (2) การแบ่งช่วงของข้อมูลที่เป็นแบบ Non Nominal Attribute (ข้อมูลที่มีคุณลักษณะต่อเนื่อง) ซึ่งเป็นการแบ่งข้อมูลที่มีลักษณะที่เป็นตัวเลขออกเป็นช่วงข้อมูล เพื่อที่จะนำเอา attribute ที่ถูกแบ่งช่วงข้อมูลไปใช้ในการทำเหมืองข้อมูลและ (3) การกำจัดข้อมูลที่มีค่าไม่เกาะกลุ่มกับค่าอื่นๆ ในชุดข้อมูล (Outlier)

ขั้นตอนที่ 4 นำเอาข้อมูลที่ผ่านการเตรียมข้อมูลเรียบร้อยแล้วเข้าสู่อัลกอริทึมที่ใช้ในการทำเหมืองข้อมูล ซึ่งในการนำข้อมูลไปเข้าสู่อัลกอริทึมที่ใช้ในการทำเหมืองข้อมูล นั้นต้องทำการแบ่งข้อมูลออกเป็น 2 ชุด ได้แก่ Training set และ Test Set โดยวิธีการแบ่งใช้วิธี 10 Fold Crossvalidation คือ แบ่งข้อมูล ออกเป็น 10 กลุ่ม แล้วผลัดกันเป็น Training Set และ Test Set จนครบ ทั้ง 10 กลุ่ม ซึ่งลักษณะของการสับเปลี่ยนกันเป็น Training Set และ Test Set โดยที่เริ่มต้นจะนำข้อมูลที่ อยู่ใน Training Set ไปทำการฝึกให้กับอัลกอริทึมที่ใช้ในการทำเหมืองข้อมูล ซึ่งเมื่อได้ผลลัพธ์ออกมาจะนำไปวัดประสิทธิภาพการทำนายของแบบจำลองด้วยชุดข้อมูล Test set

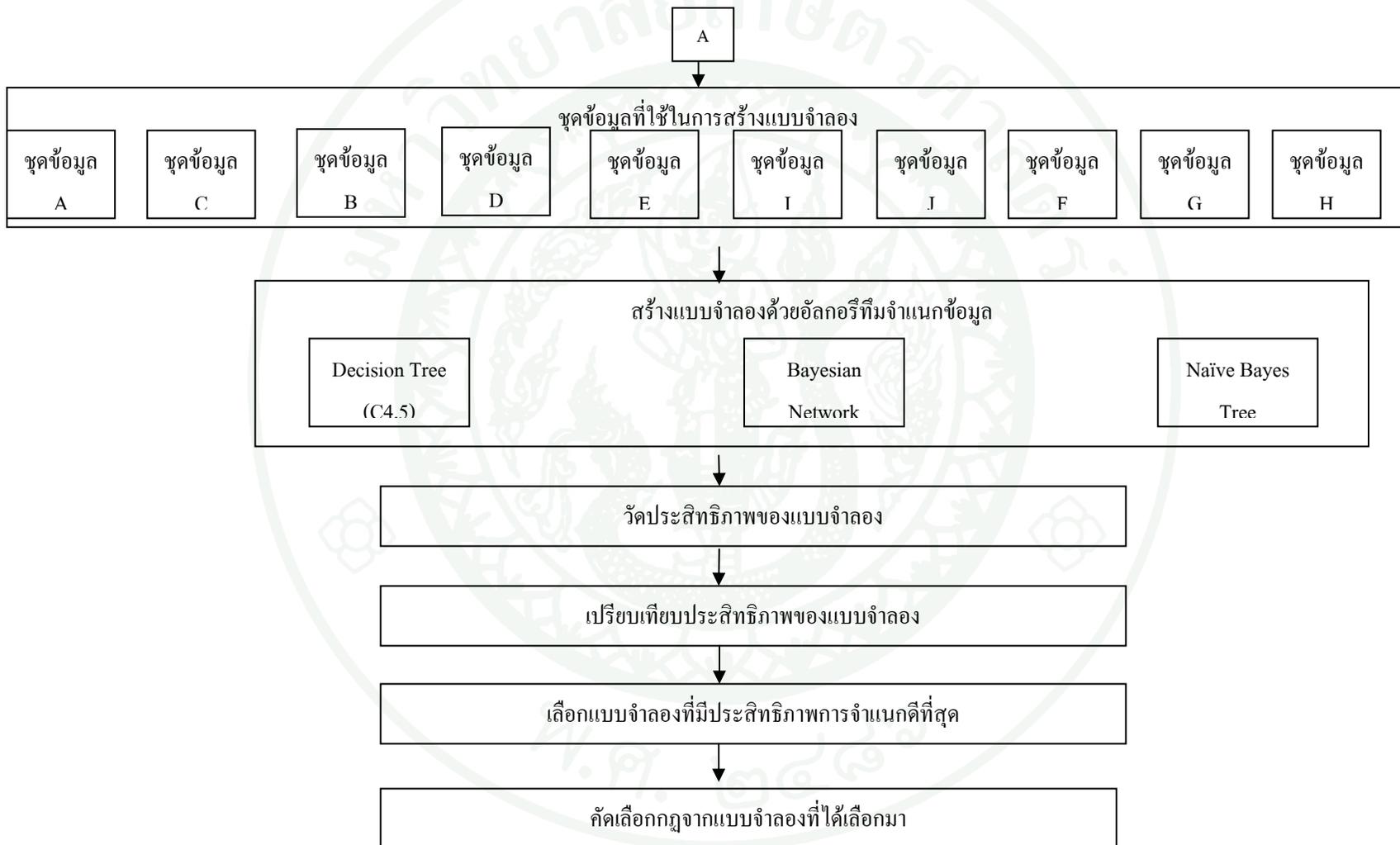
ขั้นตอนที่ 5 ทำการทดสอบประสิทธิภาพของแบบจำลองที่ได้จากอัลกอริทึมการทำเหมืองข้อมูลซึ่งได้แก่ ประสิทธิภาพการจำแนกข้อมูล เป็นต้น แล้ว จึงนำผลลัพธ์ที่ได้ไปประเมินความน่าสนใจ และ ความถูกต้องของผลลัพธ์

ขั้นตอนที่ 6 นำผลลัพธ์ที่ผ่านการวัดความน่าสนใจ และ ความถูกต้องเรียบร้อยแล้วไปใช้ประโยชน์ตามหัวข้อที่ได้กำหนดไว้





ภาพที่ 6 ขั้นตอนการทดลอง



ภาพที่ 6 (ต่อ)

ขั้นตอนที่ 1 รวบรวมข้อมูล นำข้อมูลรายงานสัตว์ปีกป่วยตายตั้งแต่ปี 2004 - 2007 ที่มีลักษณะตามภาพที่ 3 และ ข้อมูลอุณหภูมิ ความชื้นที่มีลักษณะตามภาพที่ 4 จำนวนชั่วโมงที่มีแสงแดด ที่มีลักษณะของข้อมูลดังภาพที่ 5 และความสูงของพื้นที่ จากกรมอุตุนิยมวิทยา ที่มีลักษณะของข้อมูลดังภาพที่ 6 มารวม เป็นชุดข้อมูล เพื่อใช้ในการทำวิจัย และ ทำการตัดข้อมูลที่ซ้ำออกไปจากชุดข้อมูล โดยพิจารณาจาก เลขที่ของข้อมูลจากรายงานสัตว์ปีกป่วยตายที่ซ้ำกัน หลังจากทำการกำจัดข้อมูลที่ซ้ำกันเรียบร้อยแล้วทำการบันทึกข้อมูลให้อยู่ในรูปแบบไฟล์นามสกุล .CSV ตัวอย่างข้อมูลที่ผ่านขั้นตอนที่ 1 มีลักษณะตามตารางที่ 9

ขั้นตอนที่ 2 การเตรียมข้อมูล ในการเตรียมข้อมูลเป็นขั้นตอนที่นำข้อมูลที่ผ่านขั้นตอนที่ 1 มาสร้างเป็นชุดข้อมูล เพื่อใช้ในการทำวิจัยซึ่งในแต่ละชุดข้อมูลมีวิธีการเตรียมข้อมูลดังต่อไปนี้  
วิธีการเตรียมข้อมูลที่ 1 การแบ่งช่วงข้อมูล (Discretization) เป็นการแบ่งช่วงข้อมูลที่มีลักษณะเป็นตัวเลข ซึ่งในชุดข้อมูลที่นำมาใช้ในการวิจัยนี้มี Attribute ที่มีข้อมูลเป็นลักษณะเป็นตัวเลขมีดังตามตารางที่ 10

ตารางที่ 9 ตัวอย่างข้อมูลที่ใช้ในการทดลองที่รวบรวมจากข้อมูลสัตว์ปีกป่วยตาย ข้อมูลสภาพอากาศ และข้อมูลลักษณะพื้นที่

Month	Height	Landscape	Region	Type	Quantity	sick	Death	Destruction	Temperature	humidity	Day Light		Result
											Hour	Season	
JUN	8	S1	South	C	30.00	2.00	5.00	0.00	25.80	77.51	4.56	Rainny	Negative
AUG	8	S1	South	C	29.00	5.00	0.00	0.00	25.80	75.67	4.57	Rainny	Negative
AUG	8	S1	South	D	60.00	0.00	4.00	0.00	25.80	75.67	4.57	Rainny	Negative
AUG	8	S1	South	C	15.00	2.00	2.00	0.00	25.80	75.67	4.57	Rainny	Negative
AUG	8	S1	South	D	60.00	4.00	5.00	0.00	25.80	75.67	4.57	Rainny	Negative
AUG	8	S1	South	C	50.00	5.00	15.00	18.00	25.80	75.67	4.57	Rainny	Negative
AUG	8	S1	South	C	10.00	10.00	10.00	0.00	25.80	75.67	4.57	Rainny	Negative
AUG	8	S1	South	C	16.00	0.00	3.00	11.00	25.80	75.67	4.57	Rainny	Negative
AUG	8	S1	South	C	21.00	0.00	15.00	0.00	25.80	75.67	4.57	Rainny	Negative
AUG	8	S1	South	C	18.00	0.00	8.00	0.00	25.80	75.67	4.57	Rainny	Negative
AUG	8	S1	South	C	3.00	0.00	1.00	0.00	25.80	75.67	4.57	Rainny	Negative
JUL	8	S1	South	C	25.00	0.00	5.00	0.00	26.10	93.00	4.35	Rainny	Negative
JUL	8	S1	South	D	35.00	10.00	15.00	0.00	25.90	88.00	4.35	Rainny	Negative

ตารางที่ 10 คุณลักษณะของข้อมูลที่ถูกนำมาแบ่งช่วงข้อมูลเนื่องจากเป็นคุณลักษณะที่มีชนิดข้อมูลเป็น ตัวเลข

ชื่อคุณลักษณะ	จำนวนช่วงของข้อมูล	ขอบเขตของข้อมูล	ตัวอักษรที่ใช้แทนขอบเขตข้อมูลในแต่ละคุณลักษณะ
Height	10	$\leq 3.5$ , (3.5-6.5] ,(6.5-12],[12-31],[31-34.5],[34.5-97],[97-124.5],[124.5-382.5],[382.5-388.5] , $> 388.5$	H1,H2,H3,H4,H5, H6,H7,H8,H9,H10
Quantity	4	$\leq 59.5$ , (59.5-120.5],[120.5-827] , $> 827$	Q1,Q2,Q3,Q4
Sick	4	$\leq 9.5$ ,(9.5-49.5],[49.5-784.5] , $> 784.5$	SI1,SI2,SI3,SI4
Death	4	$\leq 4.5$ ,(4.5-17.5],[17.5-97.5] , $> 97.5$	D1,D2,D3,D4
Destruction	5	$\leq 0.5$ ,(0.5-17.5],[17.5-87.5],[87.5-968.5] , $> 968.5$	De1,De2,De3,De4, De5
Temperature	4	$\leq 21.15$ , (21.15-27.85] , (27.85-32.05] , $> 32.05$	T1,T2,T3,T4
Humidity	3	$\leq 60.5$ , (60.5-75.09] , $> 75.09$	Hu1,Hu2,Hu3,
Day Light Hour	23	$\leq 2.360$ , (2.360-3.049] , (3.049-3.099] , (3.099-3.382] , (3.382-3.508] , (3.508-3.710] , (3.710-3.756] , (3.756-4.307] , (4.307-4.332] , (4.332-4.579] , (4.579-5.143] , (5.143-5.193] , (5.193-7.143] , (7.143-7.188] , (7.188-7.611] , (7.611-7.633] , (7.633-8.007] , (8.007-8.022] , (8.022-8.087] , (8.087-8.218] , (8.218-8.710] , (8.710-8.743] , $> 8.743876$	L1, L2, L3, L4, L5, L6, L7, L8, L9, L10, L11, L12, L13, L14, L15, L16, L17, L18 L19, L20, L21, L22, L23

จากตารางที่ 10 พบว่าในค่าของ Attribute ที่มีลักษณะเป็นตัวเลข หลังจากที่ได้แบ่งเป็นช่วงลักษณะการแสดงผลช่วงของข้อมูล จะแสดงออกมาในลักษณะเป็นช่วง กิ่งเปิด กิ่งปิด เขียนเป็นสัญลัษณ์ (x,y) หมายความว่า ค่าที่อยู่ในช่วงนี้ได้ต้องมีค่ามากกว่า x แต่มีค่าน้อยกว่าหรือเท่ากับ y ตัวอย่างเช่น ค่าของ Attribute Humidity ที่จะอยู่ในช่วง (60.5-75.09] ได้นั้น ต้องมีค่าที่มากกว่า 60.5 แต่น้อยกว่าหรือเท่ากับ 75.09

ตารางที่ 11 ความหมายของรายการในคุณลักษณะ Land Scape

คำภายในคุณลักษณะ Landscape	ชื่อสมาชิก
คาบสมุทรภาคใต้(ชายฝั่งตะวันตก)	S1
คาบสมุทรภาคใต้(ชายฝั่งตะวันออก)	S2
ชายฝั่งภาคตะวันออกเฉียงใต้	S3
ทิวเขาและหุบเขาภาคตะวันตก	S4
ทิวเขาและหุบเขาภาคเหนือ	S5
ที่ราบลุ่มดินดอนสามเหลี่ยมปากแม่น้ำ	S6
ที่ราบลุ่มน้ำดอนบน	S7
บริเวณขอบที่ราบ	S8
แอ่งโคราช	S9
แอ่งสกลนคร	S10

ตารางที่ 11 แสดงถึงการแทนค่าให้กับสมาชิกที่อยู่ใน คุณลักษณะ Landscape โดยแทนเป็นสัญลักษณ์ S1 จนถึง S10 สาเหตุที่ต้องแทนค่าด้วยสัญลักษณ์ดังกล่าวเพราะว่า สมาชิกที่อยู่ในคุณลักษณะ Landscape นั้นหมายถึง ลักษณะของภูมิประเทศของแต่ละพื้นที่ ซึ่งเป็น ภาษาไทย ตัวโปรแกรม WEKA เวอร์ชันที่ใช้อยู่ไม่รองรับภาษาไทย จึงจำเป็นต้องแทนค่าเป็นสัญลักษณ์ดังกล่าว

ตารางที่ 12 ลักษณะข้อมูลที่ผ่านมาขั้นตอนการแบ่งช่วงข้อมูล (Discretization)

Month	Height	Landscape	Region	Type	Quantity	sick	Death	Destruction	Temperature	humidity	Day Light Hour	Season	Result
JUN	'(6.5-12]'	S1	South	C	'(-inf-59.5]'	'(-inf-9.5]'	'(4.5-17.5]'	'(-inf-0.5]'	'(21.15-27.85]'	'(75.099638-inf)'	'(4.332413-4.579635]'	Rainny	Negative
AUG	'(6.5-12]'	S1	South	C	'(-inf-59.5]'	'(-inf-9.5]'	'(4.5-4.5]'	'(-inf-0.5]'	'(21.15-27.85]'	'(75.099638-inf)'	'(4.332413-4.579635]'	Rainny	Negative
AUG	'(6.5-12]'	S1	South	D	'(59.5-120.5]'	'(-inf-9.5]'	'(-inf-4.5]'	'(-inf-0.5]'	'(21.15-27.85]'	'(75.099638-inf)'	'(4.332413-4.579635]'	Rainny	Negative
AUG	'(6.5-12]'	S1	South	C	'(-inf-59.5]'	'(-inf-9.5]'	'(4.5-4.5]'	'(-inf-0.5]'	'(21.15-27.85]'	'(75.099638-inf)'	'(4.332413-4.579635]'	Rainny	Negative
AUG	'(6.5-12]'	S1	South	D	'(59.5-120.5]'	'(-inf-9.5]'	'(4.5-17.5]'	'(-inf-0.5]'	'(21.15-27.85]'	'(75.099638-inf)'	'(4.332413-4.579635]'	Rainny	Negative
AUG	'(6.5-12]'	S1	South	C	'(-inf-59.5]'	'(-inf-9.5]'	'(4.5-17.5]'	'(17.5-87.5]'	'(21.15-27.85]'	'(75.099638-inf)'	'(4.332413-4.579635]'	Rainny	Negative
AUG	'(6.5-12]'	S1	South	C	'(-inf-59.5]'	'(9.5-49.5]'	'(4.5-17.5]'	'(-inf-0.5]'	'(21.15-27.85]'	'(75.099638-inf)'	'(4.332413-4.579635]'	Rainny	Negative
AUG	'(6.5-12]'	S1	South	C	'(-inf-59.5]'	'(-inf-9.5]'	'(4.5-4.5]'	'(0.5-17.5]'	'(21.15-27.85]'	'(75.099638-inf)'	'(4.332413-4.579635]'	Rainny	Negative

วิธีการเตรียมข้อมูลที่ 2 การคัดเลือกคุณลักษณะ (Attribute Selection) นำชุดข้อมูลที่ผ่านการรวมข้อมูลตามขั้นตอนที่ 1 แล้วนำไปคัดเลือกคุณลักษณะ โดยในการคัดเลือกคุณลักษณะ ใช้วิธี ClassifierSubsetEvaluator เป็นการคัดเลือกคุณลักษณะ โดยใช้ อัลกอริทึมทางด้านการจำแนกข้อมูล ที่ประกอบด้วยอัลกอริทึม Decision Tree (J48) มาใช้ในการประเมินแต่ละ คุณลักษณะที่อยู่ในชุดข้อมูล และ เลือกคุณลักษณะที่ทำให้ผลการจำแนกข้อมูลมีความถูกต้องมากที่สุด

ตารางที่ 13 คุณลักษณะที่ผ่านการคัดเลือก คุณลักษณะข้อมูล โดยวิธี ClassifierSubsetEvaluator

ชื่อตัวแปร	ประเภทข้อมูล	คำอธิบาย
Month	Nominal	เดือนที่มีการพบสัตว์ปีกป่วยตาย
Land Scape	Nominal	ลักษณะพื้นที่
Height	Numeric	ความสูงของพื้นที่
Death	Numeric	จำนวนสัตว์ปีกที่ตาย
Destruction	Numeric	จำนวนสัตว์ปีกที่ถูกทำลาย
Temperature	Numeric	อุณหภูมิ
Humidity	Numeric	ความชื้น
Daylight Hours	Numeric	จำนวนชั่วโมงที่มีแสง

หลังจากการคัดเลือกคุณลักษณะข้อมูล จากจำนวนทั้งหมด 13 คุณลักษณะ ซึ่งประกอบด้วย Month, Height, Landscape, Region, Type, Quantity, Sick, Death, Destruction, Temperature, humidity, Daylight Hours, Season พบว่า คุณลักษณะ ที่ผ่านการคัดเลือกมีจำนวนทั้งสิ้น 8 คุณลักษณะ ซึ่งประกอบด้วย Month, Landscape, Height, Death, Destruction, Temperature, Humidity และ Daylight Hours

ตารางที่ 14 ตัวอย่างชุดข้อมูลที่ผ่านวิธีการคัดเลือกคุณลักษณะ (Attribute Selection)

Month	Height	Landscape	Death	Destructi		humid	Day	Result
				on	Temperature	ity	Light	
							Hour	
JUN	8	S1	5.00	0.00	25.80	77.51	4.56	Negative
AUG	8	S1	0.00	0.00	25.80	75.67	4.57	Negative
AUG	8	S1	4.00	0.00	25.80	75.67	4.57	Negative
AUG	8	S1	2.00	0.00	25.80	75.67	4.57	Negative
AUG	8	S1	5.00	0.00	25.80	75.67	4.57	Negative
AUG	8	S1	15.00	18.00	25.80	75.67	4.57	Negative
AUG	8	S1	10.00	0.00	25.80	75.67	4.57	Negative
AUG	8	S1	3.00	11.00	25.80	75.67	4.57	Negative
AUG	8	S1	15.00	0.00	25.80	75.67	4.57	Negative
AUG	8	S1	8.00	0.00	25.80	75.67	4.57	Negative
AUG	8	S1	1.00	0.00	25.80	75.67	4.57	Negative
JUL	8	S1	5.00	0.00	26.10	93.00	4.35	Negative
JUN	8	S1	5.00	0.00	25.80	77.51	4.56	Negative

วิธีการเตรียมข้อมูลที่ 3 ถ่วงน้ำหนักที่หามาด้วยวิธี DEMATEL(Weighted By DEMATEL Method) นำชุดข้อมูลที่ ผ่านขั้นตอนที่ 1 ไปทำการหาค่าน้ำหนักด้วยวิธี Decision Making Trial And Evaluation Laboratory DEMATEL) โดยมีขั้นตอนดังต่อไปนี้

ขั้นตอนที่ 1 นำชุดข้อมูล มาทำการหารระดับความสัมพันธ์ระหว่าง Attribute โดยหาจากค่าสัมประสิทธิ์สหสัมพันธ์ (Coefficient of Correlation) โดยโปรแกรม SPSS

ขั้นตอนที่ 2 หลังจากได้ค่าสัมประสิทธิ์สหสัมพันธ์ ของแต่ละตัวแปรแล้ว เปลี่ยนค่าสัมประสิทธิ์สหสัมพันธ์เป็นค่าบวก โดยใช้วิธีการหาค่าสัมบูรณ์ของ ค่าสัมประสิทธิ์สหสัมพันธ์

ขั้นตอนที่ 3 นำค่าสัมประสิทธิ์สหสัมพันธ์ของแต่ละตัวแปรที่หามาได้นำไปหาน้ำหนักด้วยวิธีการ DEMATEL ซึ่งผลลัพธ์ที่ได้คือ ค่าน้ำหนักของแต่ละ Attribute

ขั้นตอนที่ 4 เมื่อทำการหาน้ำหนักมาได้จึงทำการถ่วงน้ำหนักไปในชุดข้อมูลที่ผ่านขั้นตอนที่ 1 โดยที่ในการถ่วงน้ำหนักจะนำเอาน้ำหนักที่ได้ในแต่ละ Attribute กระจายลงไปในทุกค่าของ Attribute นั้น ตัวอย่าง Attribute : Month มีน้ำหนักเท่ากับ 0.04 และ เนื่องจาก Attribute Month มีสมาชิกประกอบด้วย JAN FEB MAR APRIL MAY JUN JULY AUG SEP OCT NOV DEC ด้วยเหตุนี้ สมาชิกทุกตัวของ Attribute Month จะมีน้ำหนักอยู่ที่ 0.04 เท่ากันเสมอ แต่ ถ้าใน Attribute ใดก็ตาม ที่มีสมาชิกของ Attribute เป็นค่าว่าง ดังนั้นจะมีน้ำหนักเท่ากับ 0 เสมอ

ขั้นตอนที่ 5 เมื่อถ่วงน้ำหนักครบทุก Attribute ที่อยู่ภายใต้ชุดข้อมูลดังกล่าวแล้ว ทำการหาผลรวมน้ำหนักที่เกิดจากการถ่วงน้ำหนักของแต่ละ Attribute ให้กลายเป็นน้ำหนักของแต่ละ Record

ขั้นตอนที่ 6 นำน้ำหนักของแต่ละ Record เก็บไว้ในเพิ่มข้อมูล โดยบันทึก ในรูปแบบของ CSV

ขั้นตอนที่ 7 นำน้ำหนักที่หามาไปทำการผนวกเข้ากับชุดข้อมูลที่ผ่านขั้นตอนที่ 1 โดยใช้โปรแกรม WEKA เพื่อให้โปรแกรมรับรู้ว่าเป็นน้ำหนักของชุดข้อมูล และทำการบันทึกชุดข้อมูลที่มีน้ำหนักดังกล่าวเป็นไฟล์นามสกุล .ARFF



ตารางที่ 15 ลักษณะข้อมูลที่ผ่านการถ่วงน้ำหนักด้วยน้ำหนักที่หามาด้วยวิธีการ DEMATEL

Month	Height	Landscape	Region	Type	Quantity	sick	Death	Destruction	Temperature	humidity	Day		Result	Weight
											Light	Hour		
JUN	8.00	S1	South	C	30.00	2.00	5.00	0.00	25.80	77.51	4.56	Rainny	Negative	0.5
AUG	8.00	S1	South	C	29.00	5.00	0.00	0.00	25.80	75.67	4.57	Rainny	Negative	0.5
AUG	8.00	S1	South	D	60.00	0.00	4.00	0.00	25.80	75.67	4.57	Rainny	Negative	0.5
AUG	8.00	S1	South	C	15.00	2.00	2.00	0.00	25.80	75.67	4.57	Rainny	Negative	0.5
AUG	8.00	S1	South	D	60.00	4.00	5.00	0.00	25.80	75.67	4.57	Rainny	Negative	0.5
AUG	8.00	S1	South	C	50.00	5.00	15.00	18.00	25.80	75.67	4.57	Rainny	Negative	0.5
AUG	8.00	S1	South	C	10.00	10.00	10.00	0.00	25.80	75.67	4.57	Rainny	Negative	0.5
AUG	8.00	S1	South	C	16.00	0.00	3.00	11.00	25.80	75.67	4.57	Rainny	Negative	0.5
AUG	8.00	S1	South	C	21.00	0.00	15.00	0.00	25.80	75.67	4.57	Rainny	Negative	0.5
AUG	8.00	S1	South	C	18.00	0.00	8.00	0.00	25.80	75.67	4.57	Rainny	Negative	0.5
AUG	8.00	S1	South	C	3.00	0.00	1.00	0.00	25.80	75.67	4.57	Rainny	Negative	0.5
JUL	8.00	S1	South	C	25.00	0.00	5.00	0.00	26.10	93.00	4.35	Rainny	Negative	0.5
JUL	8.00	S1	South	D	35.00	10.00	15.00	0.00	25.90	88.00	4.35	Rainny	Negative	0.5

วิธีการสร้างที่ 4 ถ่วงน้ำหนักที่หามาด้วยวิธี Emerging Pattern (Weighted by EP Method) เป็นวิธีการเตรียมข้อมูลที่น่าชุดข้อมูลที่ผ่านวิธีการสร้างที่ 1 มาทำการหาค่าน้ำหนักด้วยวิธี Emerging pattern ซึ่งขั้นตอนการทำงานดังต่อไปนี้

ขั้นตอนที่ 1 นำชุดข้อมูลที่ 1 ทำการหา Item Set โดยใช้ อัลกอริทึม APRIORI ในการหา Item Set ซึ่ง Item set ที่หามานั้นมีลักษณะ เป็นกฎความสัมพันธ์ ระหว่างค่าของ Attribute ที่ปรากฏ ในชุดข้อมูล

ขั้นตอนที่ 2 นำ itemset ที่หามาได้ทำการหาค่าน้ำหนัก ด้วยวิธี Emerging pattern

ขั้นตอนที่ 3 นำน้ำหนักของแต่ละ Itemset ทำการถ่วงไปในแต่ละ Record ของชุดข้อมูล

ขั้นตอนที่ 4 หาผลรวมของน้ำหนักแต่ละ Itemset ที่ปรากฏอยู่ในแต่ละ Record กลายเป็น น้ำหนักของ Itemset ของแต่ละ Record

ขั้นตอนที่ 5 นำน้ำหนักของแต่ละ Record ไปทำการสร้างแฟ้มข้อมูลต่าง โดยบันทึกไว้ อยู่ในไฟล์ ในรูปแบบของ CSV

ขั้นตอนที่ 6 นำน้ำหนักที่หามาไปทำการผนวกเข้ากับชุดข้อมูลที่ 1 โดยผ่าน Code ของ โปรแกรม WEKA เพื่อให้โปรแกรมรับรู้ว่าเป็น น้ำหนักของชุดข้อมูล และทำการบันทึกชุดข้อมูล ที่มีน้ำหนักดังกล่าวเป็นไฟล์นามสกุล .ARFF

ตารางที่ 16 ตัวอย่าง ของ Item Set ที่นำไปใช้ในการคำนวณน้ำหนักของข้อมูล

หมายเลข	รายละเอียดของ Item set
I1	Month =Aug
I2	Month = Oct
I3	Month = Nov
I4	Temperature='(27.85-32.05]'
I5	Month=AUG Season=Rainny
I7	Landscap =G Region=Center Temperature='(27.85-32.05]'
I8	Heigt='(124.5-382.5]' Landscape=I Region=NorthEast

วิธีการสร้างทั้งหมดที่ได้กล่าวไปข้างต้นเป็นวิธีการดำเนินงานกับชุดข้อมูล โดยใช้ (1) การแบ่งช่วงข้อมูล (2) การคัดเลือกคุณลักษณะ (3) ถ่วงน้ำหนักที่หามาด้วยวิธี DEMATEL และ (4) ถ่วงน้ำหนักที่หามาด้วยวิธี Emerging Pattern ซึ่งในการวิจัยนี้นำเอาวิธีการดำเนินงานทั้ง 4 วิธีมาใช้ในการสร้างชุดข้อมูลที่ใช้ในการวิจัย ซึ่งได้ชุดข้อมูลสำหรับใช้ในการวิจัยทั้งสิ้น 10 ชุดข้อมูลตามตารางที่ 17

ตารางที่ 17 อธิบายลักษณะของชุดข้อมูลที่นำมาใช้ในงานวิจัย

Dataset	Discritization	Attribute Selection	Weight Method
A	N	N	No
B	N	Y	No
C	N	N	DEMATEL
D	N	Y	DEMATEL
E	Y	N	No
F	Y	Y	No
G	Y	Y	DEMATEL
H	Y	Y	Emerging Pattern
I	Y	N	Emerging Pattern
J	Y	N	DEMATEL

หมายเหตุ N = ไม่มีการดำเนินงาน Y = มีการดำเนินงาน

จาก ภาพที่ 6 และ ตารางที่ 17 พบว่าไม่มีการถ่วงน้ำหนักด้วยวิธี Emerging Pattern กับข้อมูลที่ไม่ได้ทำขั้นตอน Discretization อันเนื่องมาจาก ในการถ่วงน้ำหนักด้วยวิธี Emerging Pattern ต้องใช้ อัลกอริทึม ในค้นหา Item set ซึ่งอัลกอริทึม Apriori ไม่รองรับข้อมูลที่เป็นตัวเลขซึ่งยังไม่ได้ถูกแบ่งเป็นช่วง (Discritization) ได้ ดังนั้นข้อมูลที่ไม่ผ่านการแบ่งช่วงจึงไม่สามารถถ่วงน้ำหนักด้วย Emerging Pattern ได้

ขั้นตอนที่ 3 การสร้างแบบจำลองหลังจากที่ได้ชุดข้อมูลผ่านการเตรียมข้อมูลมาแล้ว  
ขั้นตอนต่อไปคือการสร้างแบบจำลอง โดยอัลกอริทึมที่นำมาใช้ในการสร้างแบบจำลอง  
ประกอบด้วย Bayesian Networks, Decision Tree(J48) และ Naïve Bayes Tree

**ตารางที่ 18** ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพ  
ของแบบจำลองระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่ได้ผ่านขั้นตอนการเตรียม  
ข้อมูล กับแบบจำลองที่สร้างจากชุดข้อมูลผ่านการ คัดเลือกคุณลักษณะ (Attribute  
Selection)

Algorithm	Data set	Discretization	Attribute Selection	Weight Method
Decision Tree	A	N	N	No
Bayes Net	A	N	N	No
NB Tree	A	N	N	No
Decision Tree	B	N	Y	No
Bayes Net	B	N	Y	No
NB Tree	B	N	Y	No

หมายเหตุ N = ไม่มีการดำเนินงาน Y = มีการดำเนินงาน

ตารางที่18 เป็นการแสดงความสัมพันธ์ของชุดข้อมูลที่ใช้ในการสร้างแบบจำลองที่มี  
ลักษณะแตกต่างกันสองชุด ได้แก่ชุดข้อมูลที่ไม่ผ่านการคัดเลือก คุณลักษณะ(ชุดข้อมูลแบบ A )  
และ ชุดข้อมูลที่ผ่านการคัดเลือกคุณลักษณะ (ชุดข้อมูลแบบ B) กับ อัลกอริทึมด้านการจำแนก  
ข้อมูลที่นำมาใช้ในการสร้างแบบจำลอง

เมื่อทำการสร้างแบบจำลองตามความสัมพันธ์ในตารางที่ 18 ทำให้ได้แบบจำลองที่นำมา  
เปรียบเทียบประสิทธิภาพระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่มีการคัดเลือกคุณลักษณะ  
กับแบบจำลองที่สร้างจากชุดข้อมูลที่มีการคัดเลือกคุณลักษณะ

**ตารางที่ 19** ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่ได้ผ่านขั้นตอนการเตรียมข้อมูล กับแบบจำลองที่สร้างจากชุดข้อมูลผ่านการแบ่งช่วงข้อมูล (Discretization)

Algorithm	Dataset	Discretization	Attribute Selection	Weight Method
Decision Tree	A	N	N	No
Bayes Net	A	N	N	No
NB Tree	A	N	N	No
Decision Tree	E	Y	N	No
Bayes Net	E	Y	N	No
NB Tree	E	Y	N	No

หมายเหตุ N = ไม่มีการดำเนินงาน Y = มีการดำเนินงาน

ตารางที่ 19 เป็นการแสดงความสัมพันธ์ของชุดข้อมูลที่ใช้ในการสร้างแบบจำลองที่มีลักษณะแตกต่างกันสองชุด ได้แก่ชุดข้อมูลที่ คุณลักษณะที่มีลักษณะข้อมูลเป็นตัวเลขไม่ผ่านการแบ่งช่วงข้อมูล (ชุดข้อมูลแบบ A ) และ ชุดข้อมูลที่คุณลักษณะที่มีลักษณะข้อมูลเป็นตัวเลขผ่านการแบ่งช่วงข้อมูล(ชุดข้อมูลแบบ E) กับ อัลกอริทึมด้านการจำแนกข้อมูลที่นำมาใช้ในการสร้างแบบจำลอง

เมื่อทำการสร้างแบบจำลองตามความสัมพันธ์ในตารางที่ 19 ทำให้ได้แบบจำลองที่นำมาเปรียบเทียบประสิทธิภาพระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่มีการแบ่งช่วงคุณลักษณะกับแบบจำลองที่สร้างจากชุดข้อมูลที่มีการแบ่งช่วงคุณลักษณะที่เป็นตัวเลข

**ตารางที่ 20** ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ผ่านขั้นตอนการคัดเลือกคุณลักษณะ และไม่แบ่งช่วงข้อมูล และไม่ ถ่วงน้ำหนักด้วยน้ำหนักที่หามาจาก DEMATEL กับแบบจำลองที่สร้างจากชุดข้อมูลที่ผ่านขั้นตอนการคัดเลือกคุณลักษณะกับขั้นตอนการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจาก DEMATEL

Algorithm	Data set	Discritization	Attribute Selection	Weight Method
Decision Tree	B	N	Y	No
Bayes Net	B	N	Y	No
NB Tree	B	N	Y	No
Decision Tree	D	N	Y	DEMATEL
Bayes Net	D	N	Y	DEMATEL
NB Tree	D	N	Y	DEMATEL

หมายเหตุ N = ไม่มีการดำเนินงาน Y = มีการดำเนินงาน

ตารางที่ 20 เป็นการแสดงความสัมพันธ์ของชุดข้อมูลที่ใช้ในการสร้างแบบจำลองที่มีลักษณะแตกต่างกันสองชุด ได้แก่ชุดข้อมูลที่มีการคัดเลือกคุณลักษณะ แต่ไม่มีการถ่วงน้ำหนักข้อมูล(ชุดข้อมูลแบบ B) และ ชุดข้อมูลที่มีการคัดเลือกคุณลักษณะและมีการถ่วงน้ำหนักข้อมูลด้วยวิธีการ DEMATEL (ชุดข้อมูลแบบ D) กับ อัลกอริทึมด้านการจำแนกข้อมูลที่นำมาใช้ในการสร้างแบบจำลอง

เมื่อทำการสร้างแบบจำลองตามความสัมพันธ์ในตารางที่ 20 ทำให้ได้แบบจำลองที่นำมาเปรียบเทียบประสิทธิภาพระหว่างแบบจำลองที่สร้างจากชุดข้อมูลมีการถ่วงน้ำหนัก กับแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่มีการถ่วงน้ำหนัก โดยที่ทั้งสองข้อมูลอยู่ภายใต้เงื่อนไขคือผ่านการคัดเลือกคุณลักษณะ

**ตารางที่ 21** ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่ได้ผ่านการแบ่งช่วง ไม่ได้ผ่านการคัดเลือกคุณลักษณะ และไม่ได้ทำการถ่วงน้ำหนัก กับแบบจำลองที่ผ่านขั้นตอนการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจาก DEMATEL

Algorithm	Dataset	Discritization	Attribute Selection	Weight Method
Decision Tree	A	N	N	No
Bayes Net	A	N	N	No
NB Tree	A	N	N	No
Decision Tree	C	N	N	DEMATEL
Bayes Net	C	N	N	DEMATEL
NB Tree	C	N	N	DEMATEL

หมายเหตุ N = ไม่มีการดำเนินงาน Y = มีการดำเนินงาน

ตารางที่ 21 เป็นการแสดงความสัมพันธ์ของชุดข้อมูลที่ใช้ในการสร้างแบบจำลองที่มีลักษณะแตกต่างกันสองชุด ได้แก่ชุดข้อมูลที่ไม่มีการถ่วงน้ำหนักข้อมูล (ชุดข้อมูลแบบ A) และชุดข้อมูลที่มีการถ่วงน้ำหนักข้อมูลด้วยวิธีการ DEMATEL (ชุดข้อมูลแบบ C) กับ อัลกอริทึมด้านการจำแนกข้อมูลที่นำมาใช้ในการสร้างแบบจำลอง

เมื่อทำการสร้างแบบจำลองตามความสัมพันธ์ในตารางที่ 21 ทำให้ได้แบบจำลองที่นำมาเปรียบเทียบประสิทธิภาพระหว่างแบบจำลองที่สร้างจากชุดข้อมูลมีการถ่วงน้ำหนัก กับแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่มีการถ่วงน้ำหนัก โดยที่ทั้งสองข้อมูลอยู่ภายใต้เงื่อนไขคือไม่ได้ผ่านการเตรียมข้อมูลเช่น การคัดเลือกคุณลักษณะ และ การแบ่งช่วงข้อมูล

**ตารางที่ 22** ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองที่สร้างจากชุดข้อมูลที่ผ่านขั้นตอนการแบ่งช่วงข้อมูล และไม่ผ่านการถ่วงน้ำหนัก กับแบบจำลองที่ผ่านขั้นตอนการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจาก DEMATEL กับ แบบจำลองที่ผ่านขั้นตอนการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจาก Emerging Pattern

Algorithm	Dataset	Discritization	Attribute Selection	Weight Method
Decision Tree	E	Y	N	No
Bayes Net	E	Y	N	No
NB Tree	E	Y	N	No
Decision Tree	I	Y	N	Emerging Pattern
Bayes Net	I	Y	N	Emerging Pattern
NB Tree	I	Y	N	Emerging Pattern
Decision Tree	J	Y	N	DEMATEL
Bayes Net	J	Y	N	DEMATEL
NB Tree	J	Y	N	DEMATEL

หมายเหตุ N = ไม่มีการดำเนินงาน Y = มีการดำเนินงาน

ตารางที่ 22 เป็นการแสดงความสัมพันธ์ของชุดข้อมูลที่ใช้ในการสร้างแบบจำลองที่มีลักษณะแตกต่างกันสองชุด ได้แก่ชุดข้อมูลที่ไม่มีการถ่วงน้ำหนักข้อมูล (ชุดข้อมูลแบบ E) และชุดข้อมูลที่มีการถ่วงน้ำหนักข้อมูลด้วยน้ำหนักที่หามาจากวิธีการ DEMATEL (ชุดข้อมูลแบบ I) และชุดข้อมูลที่มีการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ Emerging Pattern (ชุดข้อมูลแบบ J) กับ อัลกอริทึมด้านการจำแนกข้อมูลที่นำมาใช้ในการสร้างแบบจำลอง

เมื่อทำการสร้างแบบจำลองตามความสัมพันธ์ในตารางที่ 22 ทำให้ได้แบบจำลองที่นำมาเปรียบเทียบประสิทธิภาพระหว่างแบบจำลองที่สร้างจากชุดข้อมูล 3 ชุด ซึ่งประกอบด้วยแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่มีการถ่วงน้ำหนัก, แบบจำลองที่สร้างจากชุดข้อมูลที่มีการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ DEMATEL และ แบบจำลองที่สร้างจากชุดข้อมูลที่มีการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ Emerging Pattern โดยที่ทุกชุดข้อมูลอยู่ภายใต้เงื่อนไขคือผ่านการแบ่งช่วงข้อมูล

**ตารางที่ 23** ความสัมพันธ์ระหว่างชุดข้อมูลกับการสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองระหว่างแบบจำลองที่สร้างจากชุดข้อมูลที่ผ่านขั้นตอนการแบ่งช่วงข้อมูล และคัดเลือกคุณลักษณะ และคำนวณน้ำหนักโดยใช้วิธี DEMATEL และ Emerging Pattern

Algorithm	Dataset	Discritization	Attribute Selection	Weight Method
Decision Tree	F	Y	Y	No
Bayes Net	F	Y	Y	No
NB Tree	F	Y	Y	No
Decision Tree	G	Y	Y	DEMATEL
Bayes Net	G	Y	Y	DEMATEL
NB Tree	G	Y	Y	DEMATEL
Decision Tree	H	Y	Y	Emerging Pattern
Bayes Net	H	Y	Y	Emerging Pattern
NB Tree	H	Y	Y	Emerging Pattern

หมายเหตุ N = ไม่มีการดำเนินงาน Y = มีการดำเนินงาน

ตารางที่ 23 เป็นการแสดงความสัมพันธ์ของชุดข้อมูลที่ใช้ในการสร้างแบบจำลองที่มีลักษณะแตกต่างกันสองชุด ได้แก่ชุดข้อมูลที่ไม่มีการถ่วงน้ำหนักข้อมูล (ชุดข้อมูลแบบ F) และชุดข้อมูลที่มีการถ่วงน้ำหนักข้อมูลด้วยน้ำหนักที่หามาจากวิธีการ DEMATEL (ชุดข้อมูลแบบ G) และชุดข้อมูลที่มีการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ Emerging Pattern (ชุดข้อมูลแบบ H) กับ อัลกอริทึมด้านการจำแนกข้อมูลที่นำมาใช้ในการสร้างแบบจำลอง

เมื่อทำการสร้างแบบจำลองตามความสัมพันธ์ในตารางที่ 23 ทำให้ได้แบบจำลองที่นำมาเปรียบเทียบประสิทธิภาพระหว่างแบบจำลองที่สร้างจากชุดข้อมูล 3 ชุด ซึ่งประกอบด้วยแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่มีการถ่วงน้ำหนัก, แบบจำลองที่สร้างจากชุดข้อมูลที่มีการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ DEMATEL และ แบบจำลองที่สร้างจากชุดข้อมูลที่มีการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ Emerging Pattern โดยที่ทุกชุดข้อมูลอยู่ภายใต้เงื่อนไขคือผ่านการแบ่งช่วงข้อมูล และ การคัดเลือก คุณลักษณะ

เหตุผลที่ทำการเลือกอัลกอริทึม ข่ายงานแบบเบย์ส์ (Bayesian Network), ต้นไม้ตัดสินใจ (Decision Tree) และ Naïve bayes Tree มาใช้ในการศึกษาเพราะอัลกอริทึมทั้ง 3 ให้ผลลัพธ์ที่ได้คือแบบจำลองที่สามารถนำไปใช้ในการสร้างกฎเพื่อเฝ้าระวังโรคไข้หวัดนก และ มีงานวิจัยที่ผ่านมาในอดีตได้มีการใช้อัลกอริทึมเหล่านี้ในการศึกษามาก่อนแล้ว โดยเฉพาะ Bayesian Network และ ด้วยตัวข้อมูลที่นำมาใช้ในการศึกษาซึ่งเหมาะสมกับการใช้อัลกอริทึมดังกล่าวมาใช้ในการศึกษา

ขั้นตอนที่ 4 วัดประสิทธิภาพของแบบจำลอง ในการเรียนรู้ของแบบจำลอง และการทดสอบแบบจำลองเพื่อทำการเปรียบเทียบประสิทธิภาพ ต้องทำการแบ่งข้อมูลให้เหมาะสมถึงทำให้ได้แบบจำลองที่มีประสิทธิภาพ ซึ่งโดยทั่วไปการแบ่งข้อมูลเพื่อใช้ในการเรียนรู้ และ การทดสอบ จะแบ่งอัตราส่วนอยู่ที่ เรียนรู้ 70 % ข้อมูลทั้งหมด และ 30 % ของข้อมูลทั้งหมดคือทดสอบ ซึ่งวิธีนี้มีข้อเสียที่ทำให้การเรียนรู้ของแบบจำลองอาจจะไม่ได้เรียนรู้ครบทุกลักษณะข้อมูล จึงทำให้มีวิธีการที่ใช้ในการแบ่งข้อมูลเพื่อฝึกและเพื่อทดสอบขึ้นมา ซึ่งได้แก่ วิธี 10-fold Cross Validation

10-Fold Cross Validation เป็นการนำเอาชุดข้อมูลทั้งหมดมาทำการแบ่งออกเป็นชุดข้อมูลย่อย 10 ชุด โดยที่ ตั้งชุดที่ 1 ถึง ชุดที่ 9 นำไปใช้ในการเรียนรู้ และ ชุดที่ 10 ใช้ทดสอบ และ ทำการหมุนเวียนสับเปลี่ยนจนกว่า แบบจำลองที่ได้ทำการทดสอบครบทุกชุดข้อมูล ตัวอย่างการแบ่งข้อมูลเพื่อใช้ฝึก และ ทดสอบ เป็นไปตามตารางที่ 24

ตารางที่ 24 การแบ่งข้อมูลใช้ในการเรียนรู้ และ ทดสอบอัลกอริทึม แบบวิธี 10 fold Cross Validation

รอบที่	ชุดข้อมูลย่อยที่ใช้เรียนรู้	ชุดข้อมูลที่ใช้ทดสอบ
1	A1,A2,A3,A4,A5,A6,A7,A8,A9	A10
2	A1,A2,A3,A4,A5,A6,A7,A8,A10	A9
3	A1,A2,A3,A4,A5,A6,A7,A10,A9	A8
4	A1,A2,A3,A4,A5,A6,A10,A9,A8	A7
5	A1,A2,A3,A4,A5,A10,A9,A8,A7	A6
6	A1,A2,A3,A4,A10,A9,A8,A7,A6	A5
7	A1,A2,A3,A10,A9,A8,A7,A6,A5	A4
8	A1,A2,A10,A9,A8,A7,A6,A5,A4	A3
9	A1,A10,A9,A8,A7,A6,A5,A4,A3	A2
10	A10,A9,A8,A7,A6,A5,A4,A3,A2	A1

จากตารางที่ 24 แสดงให้เห็นว่า แบบจำลองได้เรียนรู้ครบทุกชุดข้อมูลย่อย ซึ่งหมายความว่าแบบจำลองได้เรียนรู้ครบทุกรูปแบบข้อมูลที่มีอยู่ในแต่ละชุดข้อมูลย่อย และได้ทดสอบแบบจำลองกับทุกรูปแบบข้อมูลที่อยู่ในชุดข้อมูลย่อย ในงานวิจัยนี้ใช้วิธีการแบ่งข้อมูลเพื่อเรียนรู้ และ ทดสอบประสิทธิภาพของแบบจำลองด้วยวิธี 10-fold Cross Validation ในการวัดประสิทธิภาพ ของแบบจำลองและ การเปรียบเทียบที่นำมาใช้ในงานวิจัยนี้ ใช้ค่าการจำแนกถูกต้อง, ค่า precision, ค่า Recall และ ค่า F-measure

ขั้นตอนที่ 5 เลือกแบบจำลองที่มีประสิทธิภาพการจำแนกที่ดีที่สุด ใช้ค่าประสิทธิภาพการจำแนกข้อมูล (accuracy) ค่า Precision และ Recall ในการนำมาเปรียบเทียบเพื่อเลือกแบบจำลองที่มีประสิทธิภาพ การจำแนกข้อมูลดีที่สุด เพื่อนำแบบจำลองไปทำการสกัดเป็นกฎเพื่อใช้งานในการเฝ้าระวังการระบาดของโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีก

ขั้นตอนที่ 6 กฎจากแบบจำลองที่ได้เลือกมา ทำการสกัดเป็นกฎเพื่อใช้งานในการเฝ้าระวังการระบาดของโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีก จากแบบจำลองที่มีประสิทธิภาพดีที่สุด

## ผลและวิจารณ์

### ผล

จากขั้นตอนการทดลอง งานวิจัยนี้ได้นำเอาข้อมูลรายงานสัตว์ปีก ป่วย ตาย , ข้อมูลสภาพภูมิอากาศ, ข้อมูลสภาพภูมิประเทศ มาเตรียมข้อมูลเพื่อนำไปใช้สร้างแบบจำลอง ด้วยวิธีการเตรียมข้อมูลซึ่งประกอบด้วย

1. การแบ่งช่วงข้อมูล (Discretize)
2. การคัดเลือกคุณลักษณะ (Attribute Selection)
3. การถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ DEMATEL

ซึ่งจากขั้นตอน การทดลองน้ำหนักของแต่ละคุณลักษณะที่หามาจากวิธีการ DEMATEL ได้ผลการทดลอง ดังต่อไปนี้

ตารางที่ 25 น้ำหนักของแต่ละ Attribute ที่ได้มาจากวิธีการ DEMATEL โดยไม่ผ่านการคัดเลือกคุณลักษณะ

Attribute	Weight
Month	0.040
Land Scape	0.040
Height	0.037
Region	0.040
Type	0.034
Quantity	0.040
Death	0.040
Destruction	0.034
Temp	0.035
Humidity	0.042
<b>Daylight Hour</b>	<b>0.043</b>

จากตารางน้ำหนักของคุณลักษณะที่หามาจากวิธีการของ DEMATEL พบว่า คุณลักษณะที่มีน้ำหนักมากที่สุดคือจำนวนชั่วโมงที่มีแสงแดด (Weight = 0.043) เป็นคุณลักษณะที่สำคัญมากที่สุดต่อการนำมาใช้ในการสร้างแบบจำลองเพื่อใช้เฝ้าระวังโรคไข้หวัดนก

4. ถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ Emerging Pattern ซึ่งได้ผลการทดลองดังต่อไปนี้

ตารางที่ 26 น้ำหนักของแต่ละ Item set ที่มาจากกลุ่มผลเฉลยที่เป็น Positive ด้วยวิธี Emerging Pattern

Item set	Weight
Month=OCT	0.255
Month=NOV	0.156
Height ='(6.5-12]'	0.348
Height ='(34.5-97]'	0.162
Landscape=S7	0.183
Landscape=S8	0.241
Death='(17.5-97.5]'	0.153
Destruction='(0.5-17.5]'	0.121
Destruction'(17.5-87.5]'	0.198
<b>humidity='(60.5-75.099638]'</b>	<b>0.430</b>
Daylight Hour='(8.743876-inf)'	0.136
Month=OCT humidity='(60.5-75.099638]'	0.182
Height ='(6.5-12]'	0.220
Height ='(6.5-12]'	0.274
Height ='(34.5-97]'	0.136
Landscape=S7 humidity='(60.5-75.099638]'	0.158
Death='(4.5-17.5]'	0.066
Destruction='(0.5-17.5]'	0.114
Temperature='(21.15-27.85]'	0.128

จากตารางที่ 26 แสดงให้เห็นถึง น้ำหนักของแต่ละ Item set ที่หามาจากวิธีการ Emerging Pattern ซึ่งพบว่า Item set ที่มีน้ำหนักมากที่สุดในกรณีเป็น Class Positive คือ Itemset ที่มีค่าของคุณลักษณะความชื้นสัมพัทธ์อยู่ที่ 60.5 % ถึง 75.09 % (Weight = 0.430)

ตารางที่ 27 น้ำหนักของแต่ละ Item set ที่มาจากกลุ่มผลเฉลยที่เป็น Negative ด้วยวิธี Emerging Pattern

Item set	Weight
Month=AUG	0.171
Height ='(124.5-382.5]'	0.324
Landscape=S9	0.200
Landscape=S5	0.176
Death='(-inf-4.5]'	0.315
Death='(4.5-17.5]'	0.196
<b>Destruction='(-inf-0.5]'</b>	<b>0.441</b>
Temperature='(21.15-27.85]'	0.255
humidity='(75.099638-inf)'	0.294
Daylight Hour='(5.193871-7.143635]'	0.141
Month=AUG Destruction='(-inf-0.5]'	0.120
Month=AUG humidity='(75.099638-inf)'	0.143
Month=OCT Temperature='(21.15-27.85]'	0.056
Height='(124.5-382.5] Landscape=S9	0.201
Height ='(124.5-382.5] Death='(-inf-4.5]'	0.153
Height ='(124.5-382.5] Death='(4.5-17.5]'	0.123
Height ='(124.5-382.5] humidity='(60.5-75.099638]'	0.152
Height ='(124.5-382.5] humidity='(75.099638-inf)'	0.123
Landscape=S9 Temperature='(21.15-27.85]'	0.107
Landscape=IS9 humidity='(60.5-75.099638]'	0.119

จากตารางที่ 27 แสดงให้เห็นถึง น้ำหนักของแต่ละ Item set ที่หามาจากวิธีการ Emerging Pattern ซึ่งพบว่า Item set ที่มีน้ำหนักมากที่สุดในกรณีที่เป็น Class Negative คือ Item set ที่มีค่าของคุณลักษณะจำนวนสัตว์ปีกที่ถูกทำลาย อยู่ในช่วง น้อยกว่า 1 ตัว (Weight = 0.44)

เมื่อนำชุดข้อมูลมาผ่านวิธีการเตรียมข้อมูล ผลที่ได้คือชุดข้อมูลที่ใช้ในการสร้างแบบจำลองจำนวน 10 ชุดข้อมูล ตามตารางที่ 18 ถึง 23 มาสร้างแบบจำลองโดยใช้อัลกอริทึมการจำแนกข้อมูล ซึ่งประกอบด้วย

1. ข่ายงานแบบเบย์ส์ (Bayesian Networks)
2. ต้นไม้ตัดสินใจ (Decision Tree (J48))
3. Naïve Bayes Tree

เมื่อได้แบบจำลอง จึงนำไปทำวัดประสิทธิภาพของแบบจำลองเพื่อเลือกแบบจำลองที่มีประสิทธิภาพดีที่สุดมาใช้ในการเฝ้าระวังโรคไข้หวัดนก ซึ่งผลของประสิทธิภาพของแบบจำลองมีดังต่อไปนี้

### 1. แบบจำลองที่สร้างจากข่ายงานแบบเบย์ส์

นำชุดข้อมูลที่เตรียมไว้ทั้งหมดสิบชุด ซึ่งประกอบด้วยชุดข้อมูล A B C D E F G H I และ J ที่แต่ละชุดผ่านการเตรียมข้อมูลด้วยวิธีการที่แตกต่างกัน เมื่อนำชุดข้อมูลทั้งสิบชุดมาทำสร้างแบบจำลองด้วยวิธี ข่ายงานแบบเบย์ส์ และ เมื่อสร้างแบบจำลองเสร็จนำมาทำการวัดประสิทธิภาพของแบบจำลอง ซึ่งเป็นไปตามตารางที่ 28 และ ตารางที่ 29

ตารางที่ 28 ประสิทธิภาพของแบบจำลองที่สร้างด้วยข่ายงานแบบ เบย์ส์ (แสดงประสิทธิภาพการจำแนกข้อมูล (Accuracy))

Algorithm	Data set	Correctly Classify(%)	Incorrect Classify(%)
Bayes Net	A	91.33	8.67
Bayes Net	B	91.05	8.95
Bayes Net	C	84.27	15.73
Bayes Net	D	84.27	15.73
Bayes Net	E	91.52	8.48
Bayes Net	F	91.41	8.59
Bayes Net	G	84.27	15.73
Bayes Net	H	95.98	4.02
<b>Bayes Net</b>	<b>I</b>	<b>96.47</b>	<b>3.53</b>
Bayes Net	J	85.21	14.79

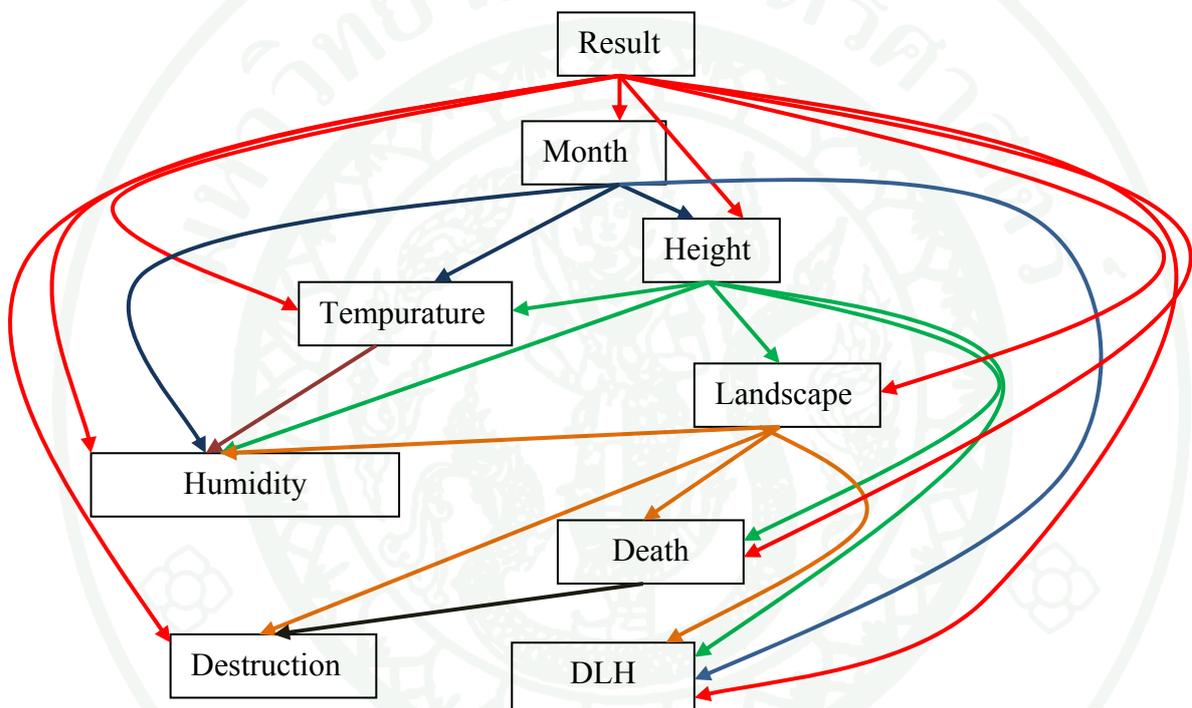
จากตารางที่ 28 พบว่า แบบจำลองที่สร้างด้วย Bayesian Network มีประสิทธิภาพการจำแนกถูกต้องมากที่สุดเมื่อแบบจำลองถูกสร้างจากชุดข้อมูล I ซึ่งเป็นชุดที่ผ่านการ Discretize ข้อมูล และมีการถ่วงน้ำหนักด้วยวิธี Emerging Pattern ซึ่งประสิทธิภาพการจำแนกถูกต้องอยู่ที่ 96.47 % และจำแนกผิด 3.53%

ตารางที่ 29 ค่า Precision, Recall และ F\_measure ของแบบจำลองที่สร้างจาก ข่ายงานแบบเบย์ส์

Algorithm	Data set	Precision	Recall	F-measure	Class
Bayes net	A	0.751	0.713	0.731	Positive
		0.946	0.853	0.49	Negative
Bayes net	B	0.751	0.688	0.718	Positive
		0.939	0.955	0.947	Negative
Bayes net	C	0.568	0.192	0.286	Positive
		0.859	0.971	0.912	Negative
Bayes net	D	0.568	0.192	0.286	Positive
		0.859	0.971	0.912	Negative
Bayes net	E	0.753	0.727	0.74	Positive
		0.946	0.953	0.949	Negative
Bayes net	F	0.759	0.707	0.732	Positive
		0.943	0.955	0.949	Negative
Bayes net	G	0.568	0.192	0.286	Positive
		0.859	0.971	0.912	Negative
Bayes net	H	<b>0.943</b>	<b>0.939</b>	<b>0.941</b>	<b>Positive</b>
		<b>0.968</b>	<b>0.971</b>	<b>0.969</b>	<b>Negative</b>
Bayes net	I	0.891	0.895	0.893	Positive
		0.979	0.979	0.979	Negative
Bayes net	J	0.561	0.333	0.423	Positive
		0.883	0.949	0.915	Negative

จากตารางที่ 29 แบบจำลอง ข่ายงานแบบเบย์ส์ มีค่า Precision, Recall และ F – measure สูงที่สุดเมื่อแบบจำลองดังกล่าวสร้างจากชุดข้อมูล H ซึ่งเป็นชุดข้อมูลที่ผ่านการ คัดเลือก Attribute มีการทำ Discretize และมีการถ่วงน้ำหนักให้กับข้อมูล ด้วยวิธี Emerging Pattern ซึ่งแบบจำลองที่สร้างจากชุดข้อมูลดังกล่าว ด้วยอัลกอริทึม Bayesian Network มีค่า Precision ในการจำแนกกว่าเป็น Class ที่มีค่าเป็นบวกที่สูง กว่าแบบจำลองข่ายงานแบบเบย์ส์ ที่สร้างจากชุดข้อมูลอื่นๆ

ในกรณีที่เปรียบเทียบแบบจำลองที่มีการถ่วงน้ำหนักให้กับข้อมูล โดยใช้วิธีการถ่วงที่แตกต่างกันคือ Emerging Pattern และ DEMATEL จากตารางการทดลองจะเห็นได้ว่า แบบจำลองที่ใช้การถ่วงน้ำหนักด้วยวิธี Emerging Pattern ให้ ความถูกต้องของการจำแนกข้อมูลที่สูงกว่าแบบจำลองที่ใช้วิธีการของ DEMATEL ในการคำนวณน้ำหนัก และ จากค่า Precision Recall และ F-measure จะพบว่าแบบจำลองข่ายงานแบบเบย์ส์ที่สร้างจากชุดข้อมูลที่มีการถ่วงน้ำหนักด้วย Emerging Pattern (H) จะให้ค่า Precision Recall และ F-measure สูงกว่า แบบจำลองข่ายงานแบบเบย์ส์ ที่มีการถ่วงน้ำหนักด้วย DEMATEL



ภาพที่ 7 แสดงลักษณะแบบจำลอง Bayesian Network ที่สร้างจากชุดข้อมูลที่ผ่านมากระบวนการถ่วงน้ำหนักและการคัดเลือกคุณลักษณะ ( ชุดข้อมูล H)

หมายเหตุ DLH = Day Light Hours

จากแบบจำลองในภาพที่ 7 แสดงให้เห็นว่า ทุกคุณลักษณะที่อยู่ในแบบจำลอง ซึ่งประกอบด้วย เดือนที่มีการรายงานสัตว์ปีกป่วยตาย (Month) ความสูงจากระดับน้ำทะเล (Height), อุณหภูมิ (Temperature) ลักษณะพื้นที่เลี้ยงสัตว์ปีก (Landscape) จำนวนสัตว์ปีกที่ตาย (Death), ความชื้นสัมพัทธ์ (Humidity) จำนวนชั่วโมงที่มีแสงแดด (Daylight hour) จำนวนสัตว์ปีกที่ถูกทำลาย (Destroy) มีผลต่อการคำนวณความน่าจะเป็นของการระบาดของโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีก และ ในแบบจำลองแสดงให้เห็นถึง ความสัมพันธ์ระหว่างคุณลักษณะดังต่อไปนี้

- 1.1 เดือนที่มีการรายงานสัตว์ปีกป่วยตาย, อุณหภูมิ จำนวนชั่วโมงที่มีแสงแดด ความสูงจากระดับน้ำทะเลของพื้นที่ และ ความชื้นของพื้นที่มีความสัมพันธ์ซึ่งกันและกัน
- 1.2 ความสูงจากระดับน้ำทะเล อุณหภูมิ จำนวนสัตว์ปีกที่ตาย ลักษณะพื้นที่เลี้ยงสัตว์ปีก, จำนวนชั่วโมงที่มีแสงแดด มีความสัมพันธ์ซึ่งกันและกัน
- 1.3 อุณหภูมิ ความชื้นของพื้นที่ มีความสัมพันธ์ซึ่งกันและกัน
- 1.4 ลักษณะพื้นที่เลี้ยงสัตว์ปีก จำนวนสัตว์ปีกที่ตาย จำนวนสัตว์ปีกที่ถูกทำลาย ความชื้นของพื้นที่ จำนวนสัตว์ปีกที่ถูกทำลาย (Destruction) มีความสัมพันธ์ซึ่งกันและกัน

จากภาพที่ 7 สามารถคำนวณหาค่าความน่าจะเป็นแบบมีเงื่อนไขระหว่างตัวแปรได้ดังตารางต่อไปนี้

ตารางที่ 30 ความน่าจะเป็นแบบมีเงื่อนไข ของ Attribute Month ที่สัมพันธ์กับ ผลเฉลี่ยของข้อมูลการระบาดของโรคไข้หวัดนก

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Pos	0.025	0	0.001	0	0	0	0.057	0.035	0.049	<b>0.443</b>	0.286	0.104
Neg	0.058	0.061	0.072	0.023	0.033	0.023	0.065	<b>0.287</b>	0.049	0.13	0.142	0.057

จากตารางที่ 30 ความน่าจะเป็นแบบมีเงื่อนไขระหว่าง Attribute Month กับ Attribute Result ที่มี  
ความน่าจะเป็นสูงที่สุด ในกรณีที่ Result มีค่าเท่ากับ Positive

$$\Pr(\text{Month} = \text{OCT} \mid \text{Result} = \text{Positive}) = 0.443$$

ซึ่งจากสมการความน่าจะเป็น สามารถแปลความหมายได้ว่า ความน่าจะเป็นของ Attribute  
month ที่มีค่าเท่ากับ OCT โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Positive มีค่าความน่าจะเป็น  
เท่ากับ 0.443

ดังนั้นเดือนตุลาคม เป็นเดือนที่มีความน่าจะเป็นสูงสุดของสัตว์ปีกที่ป่วยตายด้วยการ  
ระบาดของโรคไข้หวัดนก

ในกรณีที่ Result = Negative

$$\Pr(\text{Month} = \text{Aug} \mid \text{Result} = \text{Negative}) = 0.287$$

ซึ่งจากสมการความน่าจะเป็น สามารถแปลความหมายได้ว่า ความน่าจะเป็นของ Attribute  
month ที่มีค่าเท่ากับ Aug โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Negative มีค่าความน่าจะเป็น  
เท่ากับ 0.287

ดังนั้นในเดือนสิงหาคมเป็นเดือนที่มีโอกาสสูงสุดของสัตว์ปีกที่ป่วยตายด้วยโรคอื่นที่ไม่ใช่  
โรคไข้หวัดนก

ตารางที่ 31 ความน่าจะเป็นแบบมีเงื่อนไข ของ Attribute Height ที่มีความสัมพันธ์กับผลเฉลยของ ข้อมูล ในช่วงเวลา 12 เดือน สำหรับกรณี Positive class

Month	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10
JAN	0.038	0.012	0.111	0.012	0.031	0.653	0.012	0.010	0.012	0.012
FEB	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
MAR	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.301	0.078	0.078
APR	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
MAY	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
JUN	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
JUL	0.125	0.008	0.599	0.044	0.043	0.078	0.078	0.024	0.018	0.008
AUG	0.112	0.016	0.325	0.053	0.093	0.304	0.009	0.041	0.039	0.009
SEP	0.061	0.01	0.524	0.055	0.031	0.169	0.051	0.047	0.046	0.006
OCT	0.032	0.007	0.588	0.019	0.051	0.200	0.052	0.043	0.007	0.001
NOV	0.014	0.007	0.592	0.01	0.045	0.296	0.015	0.008	0.012	0.001
DEC	0.004	0.007	0.158	0.003	0.064	<b>0.749</b>	0.003	0.003	0.006	0.003

หมายเหตุ H1, H2, H3, ..., H10 แทนช่วงข้อมูลที่อยู่ใน Attribute Height ปรากฏอยู่ในตารางที่ 10

จากตารางที่ 31 ความน่าจะเป็นแบบมีเงื่อนไขระหว่าง Attribute Height กับ Attribute Result และ Attribute Month ที่มีความน่าจะเป็นสูงสุด ในกรณีที่ Result มีค่าเท่ากับ Positive

$$\Pr(\text{Height} = \text{H6} ((34.5-97)) \mid \text{Result} = \text{Positive}, \text{Month} = \text{DEC}) = 0.749$$

จากสมการความน่าจะเป็น สามารถแปลความหมายได้ว่า ความน่าจะเป็นที่ Attribute Height มีค่าเท่ากับ H6 (34.5 – 97) เมื่อ Attribute Result มีค่าเท่ากับ Positive และ Attribute Month มีค่าเท่ากับ DEC มีความน่าจะเป็นเท่ากับ 0.749 ดังนั้น การระบาดของโรคไข้หวัดนกมีโอกาสเกิดขึ้นสูงสุด(0.749)ในเงื่อนไขของความสูงของพื้นที่ ตั้งแต่ 34.5 - 97 เมตร และ ในเดือน ธันวาคม

ตารางที่ 32 ความน่าจะเป็นแบบมีเงื่อนไข ของ Attribute Height ที่มีความสัมพันธ์กับผลเฉลยของข้อมูล ในช่วงเวลา 12 เดือน สำหรับกรณี Negative class

Month	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10
JAN	0.081	0.069	0.068	0.087	0.007	0.065	0.022	0.509	0.003	0.089
FEB	0.057	0.047	0.065	0.028	0.011	0.099	0.02	0.572	0.004	0.095
MAR	0.015	0.030	0.04	0.018	0.006	0.044	0.017	<b>0.732</b>	0.002	0.097
APR	0.012	0.060	0.057	0.026	0.007	0.033	0.022	0.676	0.007	0.100
MAY	0.037	0.038	0.066	0.011	0.005	0.041	0.013	0.596	0.008	0.184
JUN	0.059	0.029	0.023	0.007	0.011	0.257	0.024	0.454	0.010	0.127
JUL	0.032	0.023	0.081	0.016	0.016	0.128	0.028	0.548	0.005	0.123
AUG	0.037	0.020	0.045	0.027	0.002	0.14	0.117	0.505	0.004	0.103
SEP	0.051	0.059	0.050	0.034	0.007	0.116	0.042	0.448	0.006	0.188
OCT	0.018	0.010	0.082	0.015	0.011	0.079	0.023	0.629	0.005	0.129
NOV	0.041	0.015	0.076	0.105	0.005	0.057	0.036	0.533	0.008	0.124
DEC	0.046	0.015	0.069	0.05	0.008	0.08	0.017	0.606	0.003	0.106

หมายเหตุ H1, H2, H3, ..., H10 แทนช่วงข้อมูลที่อยู่ใน Attribute Height ปรากฏอยู่ในตารางที่ 10

จากภาพที่ 7 แสดงให้เห็นว่า Node Height มี Node Result และ Node Month เป็น Parent Node เพราะฉะนั้น ความน่าจะเป็นของ Attribute Height สามารถเขียนออกมาในรูปแบบความน่าจะเป็นแบบมีเงื่อนไขได้เป็น  $\Pr(\text{Height} | \text{Result}, \text{Month})$

จากตารางที่ 32 แสดงถึงค่าความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Height ที่มี Attribute Month และ Attribute Result เป็นเงื่อนไขของความน่าจะเป็น ซึ่งจากตารางพบว่า ค่าความน่าจะเป็นที่มากที่สุดเมื่อ Attribute Result เท่ากับ Negative คือ

$$\Pr(\text{Height} = \text{H8}(124.5-382.5) | \text{Result} = \text{Negative}, \text{Month} = \text{MAR}) = 0.732$$

จากสมการความน่าจะเป็น สามารถแปลความหมายได้ว่า ความน่าจะเป็นที่ Attribute Height มีค่าเท่ากับ H8 (124.5-382.5] เมื่อ Attribute Result มีค่าเท่ากับ Negative และ Attribute Month มีค่าเท่ากับ MAR มีค่าความน่าจะเป็นเท่ากับ 0.732 ดังนั้น มีโอกาสสูงที่จะพบสัตว์ปีกป่วยตายด้วยโรคอื่นที่ไม่ใช่โรคไข้หวัดนก (0.732) ในเงื่อนไขความสูงของพื้นที่ คือ 124.5 – 382.5 เมตร และ ในเดือน มีนาคม

ตารางที่ 33 ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือนมกราคม จนถึงเดือน กุมภาพันธ์ สำหรับกรณี Positive class

Month	Height	T1	T2	T3	T4
JAN	H1	0.160	0.521	0.160	0.160
JAN	H2	0.250	0.250	0.250	0.250
JAN	H3	0.080	0.549	0.291	0.008
JAN	H4	0.250	0.250	0.250	0.250
JAN	H5	0.178	0.178	0.466	0.178
JAN	H6	0.017	0.949	0.017	0.017
JAN	H7	0.250	0.250	0.250	0.250
JAN	H8	0.082	0.665	0.171	0.082
JAN	H9	0.250	0.250	0.250	0.250
JAN	H10	0.250	0.250	0.250	0.250
FEB	H1	0.250	0.250	0.250	0.250
FEB	H2	0.250	0.250	0.250	0.250
FEB	H3	0.250	0.250	0.250	0.250
FEB	H4	0.250	0.250	0.250	0.250
FEB	H5	0.250	0.250	0.250	0.250
FEB	H6	0.250	0.250	0.250	0.250
FEB	H7	0.250	0.250	0.250	0.250

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

ตารางที่ 34 ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือนมีนาคม จนถึงเดือน เมษายน สำหรับกรณี

Positive class

Month	Height	T1	T2	T3	T4
MAR	H1	0.250	0.250	0.250	0.250
MAR	H2	0.250	0.250	0.250	0.250
MAR	H3	0.250	0.250	0.250	0.250
MAR	H4	0.250	0.250	0.250	0.250
MAR	H5	0.250	0.250	0.250	0.250
MAR	H6	0.250	0.250	0.250	0.250
MAR	H7	0.250	0.250	0.250	0.250
MAR	H8	0.145	0.564	0.145	0.145
MAR	H9	0.250	0.250	0.250	0.250
MAR	H10	0.250	0.250	0.250	0.250
APR	H1	0.250	0.250	0.250	0.250
APR	H2	0.250	0.250	0.250	0.250
APR	H3	0.250	0.250	0.250	0.250
APR	H4	0.250	0.250	0.250	0.250
APR	H5	0.250	0.250	0.250	0.250
APR	H6	0.250	0.250	0.250	0.250
APR	H7	0.250	0.250	0.250	0.250
APR	H8	0.250	0.250	0.250	0.250
APR	H9	0.250	0.250	0.250	0.250
APR	H10	0.250	0.250	0.250	0.250

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

ตารางที่ 35 ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความสูง  
ระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน พฤษภาคม จนถึงเดือน มิถุนายน สำหรับกรณี  
Positive class

Month	Height	T1	T2	T3	T4
MAY	H1	0.250	0.250	0.250	0.250
MAY	H2	0.250	0.250	0.250	0.250
MAY	H3	0.250	0.250	0.250	0.250
MAY	H4	0.250	0.250	0.250	0.250
MAY	H5	0.250	0.250	0.250	0.250
MAY	H6	0.250	0.250	0.250	0.250
MAY	H7	0.250	0.250	0.250	0.250
MAY	H8	0.250	0.250	0.250	0.250
MAY	H9	0.250	0.250	0.250	0.250
MAY	H10	0.250	0.250	0.250	0.250
JUN	H1	0.250	0.250	0.250	0.250
JUN	H2	0.250	0.250	0.250	0.250
JUN	H3	0.250	0.250	0.250	0.250
JUN	H4	0.250	0.250	0.250	0.250
JUN	H5	0.250	0.250	0.250	0.250
JUN	H6	0.250	0.250	0.250	0.250
JUN	H7	0.250	0.250	0.250	0.250
JUN	H8	0.250	0.250	0.250	0.250
JUN	H9	0.250	0.250	0.250	0.250
JUN	H10	0.250	0.250	0.250	0.250

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

**ตารางที่ 36** ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน กรกฎาคม จนถึงเดือน สิงหาคม สำหรับกรณี Positive class

Month	Height	T1	T2	T3	T4
JUL	H1	0.038	0.068	0.856	0.038
JUL	H2	0.219	0.342	0.219	0.219
JUL	H3	0.009	0.054	0.928	0.009
JUL	H4	0.089	0.493	0.329	0.089
JUL	H5	0.091	0.091	0.728	0.091
JUL	H6	0.038	0.210	0.714	0.038
JUL	H7	0.250	0.250	0.250	0.250
JUL	H8	0.133	0.336	0.397	0.133
JUL	H9	0.158	0.158	0.527	0.158
JUL	H10	0.221	0.338	0.221	0.221
AUG	H1	0.062	0.130	0.746	0.062
AUG	H2	0.207	0.378	0.207	0.207
AUG	H3	0.024	0.024	0.927	0.024
AUG	H4	0.109	0.193	0.588	0.109
AUG	H5	0.072	0.072	0.783	0.072
AUG	H6	0.026	0.149	0.799	0.026
AUG	H7	0.250	0.250	0.250	0.250
AUG	H8	0.128	0.172	0.573	0.128
AUG	H9	0.133	0.244	0.490	0.133
AUG	H10	0.250	0.250	0.250	0.250

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

ตารางที่ 37 ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน กันยายน จนถึงเดือน ตุลาคม สำหรับกรณี

Positive class

Month	Height	T1	T2	T3	T4
SEP	H1	0.079	0.079	0.764	0.079
SEP	H2	0.219	0.219	0.342	0.219
SEP	H3	0.012	0.048	0.929	0.012
SEP	H4	0.085	0.709	0.121	0.085
SEP	H5	0.126	0.126	0.622	0.126
SEP	H6	0.033	0.117	0.817	0.033
SEP	H7	0.09	0.09	0.731	0.09
SEP	H8	0.094	0.094	0.717	0.094
SEP	H9	0.097	0.097	0.709	0.097
SEP	H10	0.250	0.250	0.250	0.250
OCT	H1	0.021	0.021	0.936	0.021
OCT	H2	0.079	0.418	0.423	0.079
OCT	H3	0.001	0.133	0.864	0.001
OCT	H4	0.034	0.307	0.625	0.034
OCT	H5	0.014	0.014	0.959	0.014
OCT	H6	0.004	0.347	0.646	0.004
OCT	H7	0.013	0.812	0.162	0.013
OCT	H8	0.016	0.801	0.167	0.016
OCT	H9	0.083	0.262	0.573	0.083
OCT	H10	0.198	0.407	0.198	0.198

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

**ตารางที่ 38** ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ  
ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน พฤศจิกายน จนถึงเดือน ธันวาคม  
สำหรับ Positive class

Month	Height	T1	T2	T3	T4
NOV	H1	0.064	0.064	0.807	0.064
NOV	H2	0.109	0.205	0.577	0.109
NOV	H3	0.002	0.177	0.82	0.002
NOV	H4	0.085	0.394	0.437	0.085
NOV	H5	0.023	0.075	0.879	0.023
NOV	H6	0.004	0.617	0.376	0.004
NOV	H7	0.062	0.574	0.301	0.062
NOV	H8	0.095	0.714	0.095	0.095
NOV	H9	0.074	0.197	0.656	0.074
NOV	H10	0.250	0.250	0.250	0.250
DEC	H1	0.236	0.292	0.236	0.236
DEC	H2	0.184	0.448	0.184	0.184
DEC	H3	0.018	0.920	0.044	0.018
DEC	H4	0.25	0.250	0.250	0.250
DEC	H5	0.042	0.725	0.192	0.042
<b>DEC</b>	<b>H6</b>	<b>0.004</b>	<b>0.981</b>	<b>0.011</b>	<b>0.004</b>
DEC	H7	0.250	0.250	0.250	0.250
DEC	H8	0.250	0.250	0.250	0.250
DEC	H9	0.196	0.413	0.196	0.196
DEC	H10	0.250	0.250	0.250	0.250

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

จากตารางที่ 33 ถึง 38 ได้แสดงถึงค่าความน่าจะเป็นแบบมีเงื่อนไขของ Node Temperature ที่มี Parent Node คือ Node Result, Node Month, Node Height ซึ่งจากตารางความน่าจะเป็น แสดงให้ทราบว่า ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Temperature ที่มีค่าความน่าจะเป็นสูงสุด เมื่อ Result มีค่าเท่ากับ Positive ซึ่งอยู่ในตารางที่ 27.6 คือ

$$\Pr(\text{Temperature} = T2(21.15-27.85) | \text{Result} = \text{Positive}, \text{Month} = \text{DEC}, \text{Height} = H6(34.5-97)) = 0.981$$

จากสมการความน่าจะเป็นที่ได้กล่าวมาด้านบน สามารถแปลความหมายได้ว่า ความน่าจะเป็น ของ Attribute Temperature ที่มีค่าในช่วง 21.15 – 27.85 องศาเซลเซียส เมื่อ Attribute Result มีค่าเท่ากับ Positive และ Attribute Month มีค่าเท่ากับ DEC และ Attribute Height มีค่ามากกว่า 34.5-97 เมตร จะมีค่าความน่าจะเป็นเท่ากับ 0.981

ดังนั้น โอกาสที่มีสัตว์ปีกป่วยตายด้วยโรคไข้หวัดนกสูงสุด (0.981) ในเงื่อนไขของความสูงอยู่ในช่วง 34.5 – 97 เมตร และ ในเดือน ธันวาคม ที่อุณหภูมิในช่วง 21.15-27.85 องศาเซลเซียส

**ตารางที่ 39** ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ  
ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือนมกราคม จนถึงเดือน กุมภาพันธ์  
สำหรับกรณี Negative class

Month	Height	T1	T2	T3	T4
JAN	H1	0.032	0.784	0.153	0.032
JAN	H2	0.037	0.745	0.182	0.037
JAN	H3	0.037	0.765	0.161	0.037
JAN	H4	0.030	0.735	0.206	0.030
JAN	H5	0.183	0.369	0.265	0.183
JAN	H6	0.059	0.864	0.038	0.038
JAN	H7	0.278	0.534	0.094	0.094
JAN	H8	0.132	0.791	0.071	0.005
JAN	H9	0.243	0.243	0.27	0.243
JAN	H10	0.509	0.433	0.029	0.029
FEB	H1	0.042	0.635	0.282	0.042
FEB	H2	0.049	0.439	0.463	0.049
FEB	H3	0.062	0.297	0.604	0.037
FEB	H4	0.074	0.172	0.68	0.074
FEB	H5	0.138	0.138	0.585	0.138
FEB	H6	0.056	0.522	0.397	0.025
FEB	H7	0.096	0.503	0.306	0.096
FEB	H8	0.005	0.608	0.383	0.005
FEB	H9	0.216	0.216	0.352	0.216
FEB	H10	0.083	0.865	0.026	0.026

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

**ตารางที่ 40** ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ  
ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือนมีนาคม จนถึงเดือน เมษายน สำหรับ  
กรณี Negative class

Month	Height	T1	T2	T3	T4
MAR	H1	0.107	0.107	0.679	0.107
MAR	H2	0.062	0.171	0.704	0.062
MAR	H3	0.049	0.129	0.773	0.049
MAR	H4	0.092	0.188	0.579	0.141
MAR	H5	0.184	0.184	0.447	0.184
MAR	H6	0.045	0.400	0.509	0.045
MAR	H7	0.097	0.281	0.44	0.181
MAR	H8	0.014	0.407	0.548	0.031
MAR	H9	0.250	0.250	0.250	0.250
MAR	H10	0.022	0.920	0.036	0.022
APR	H1	0.210	0.210	0.371	0.210
APR	H2	0.087	0.305	0.522	0.087
APR	H3	0.090	0.194	0.579	0.137
APR	H4	0.149	0.352	0.284	0.214
APR	H5	0.250	0.250	0.250	0.250
APR	H6	0.128	0.128	0.615	0.128
APR	H7	0.164	0.164	0.436	0.236
APR	H8	0.010	0.437	0.443	0.109
APR	H9	0.250	0.250	0.250	0.250
APR	H10	0.058	0.364	0.521	0.058

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

**ตารางที่ 41** ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ  
ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือนพฤษภาคม จนถึงเดือน มิถุนายน  
สำหรับกรณี Negative Class

Month	Height	T1	T2	T3	T4
MAY	H1	0.093	0.515	0.299	0.093
MAY	H2	0.091	0.205	0.613	0.091
MAY	H3	0.060	0.414	0.466	0.060
MAY	H4	0.186	0.186	0.441	0.186
MAY	H5	0.250	0.250	0.250	0.250
MAY	H6	0.087	0.087	0.738	0.087
MAY	H7	0.173	0.375	0.278	0.173
MAY	H8	0.008	0.379	0.593	0.02
MAY	H9	0.214	0.359	0.214	0.214
MAY	H10	0.024	0.773	0.179	0.024
JUN	H1	0.086	0.583	0.245	0.086
JUN	H2	0.137	0.22	0.507	0.137
JUN	H3	0.155	0.336	0.353	0.155
JUN	H4	0.25	0.25	0.25	0.25
JUN	H5	0.217	0.217	0.348	0.217
JUN	H6	0.025	0.039	0.912	0.025
JUN	H7	0.154	0.154	0.539	0.154
JUN	H8	0.014	0.075	0.897	0.014
JUN	H9	0.227	0.227	0.32	0.227
JUN	H10	0.046	0.354	0.553	0.046

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

**ตารางที่ 42** ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ  
ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน กรกฎาคม จนถึงเดือน สิงหาคม  
สำหรับกรณี Negative Class

onth	Height	T1	T2	T3	T4
JUL	H1	0.063	0.255	0.619	0.063
JUL	H2	0.082	0.405	0.43	0.082
JUL	H3	0.028	0.183	0.761	0.028
JUL	H4	0.107	0.399	0.386	0.107
JUL	H5	0.108	0.108	0.676	0.108
JUL	H6	0.018	0.186	0.777	0.018
JUL	H7	0.071	0.412	0.446	0.071
JUL	H8	0.005	0.26	0.731	0.005
JUL	H9	0.197	0.197	0.409	0.197
JUL	H10	0.019	0.654	0.307	0.019
AUG	H1	0.015	0.408	0.562	0.015
AUG	H2	0.027	0.363	0.582	0.027
AUG	H3	0.012	0.373	0.602	0.012
AUG	H4	0.02	0.319	0.548	0.114
AUG	H5	0.166	0.166	0.502	0.166
AUG	H6	0.004	0.315	0.677	0.004
AUG	H7	0.005	0.578	0.413	0.005
AUG	H8	0.001	0.375	0.623	0.001
AUG	H9	0.109	0.369	0.412	0.109
AUG	H10	0.006	0.87	0.119	0.006

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

**ตารางที่ 43** ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ  
ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน กันยายน จนถึงเดือน ตุลาคม  
สำหรับกรณี Negative class

Month	Height	T1	T2	T3	T4
SEP	H1	0.055	0.243	0.648	0.055
SEP	H2	0.049	0.854	0.049	0.049
SEP	H3	0.055	0.292	0.597	0.055
SEP	H4	0.077	0.162	0.339	0.422
SEP	H5	0.195	0.195	0.416	0.195
SEP	H6	0.027	0.285	0.662	0.027
SEP	H7	0.065	0.218	0.653	0.065
SEP	H8	0.007	0.514	0.471	0.007
SEP	H9	0.208	0.208	0.376	0.208
SEP	H10	0.017	0.828	0.138	0.017
OCT	H1	0.058	0.241	0.643	0.058
OCT	H2	0.095	0.511	0.299	0.095
OCT	H3	0.015	0.622	0.348	0.015
OCT	H4	0.068	0.281	0.582	0.068
OCT	H5	0.088	0.088	0.735	0.088
OCT	H6	0.015	0.24	0.729	0.015
OCT	H7	0.047	0.751	0.155	0.047
OCT	H8	0.002	0.92	0.076	0.002
OCT	H9	0.146	0.243	0.464	0.146
OCT	H10	0.010	<b>0.971</b>	0.010	0.010

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

**ตารางที่ 44** ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute temperature ที่สัมพันธ์กับ  
ความสูงจากระดับน้ำทะเล ในช่วงเวลา ตั้งแต่เดือน พฤศจิกายน จนถึงเดือน ธันวาคม  
สำหรับกรณี Negative class

Month	Height	T1	T2	T3	T4
NOV	H1	0.026	0.358	0.589	0.026
NOV	H2	0.065	0.405	0.465	0.065
NOV	H3	0.015	0.298	0.672	0.015
NOV	H4	0.011	0.481	0.498	0.011
NOV	H5	0.137	0.215	0.51	0.137
NOV	H6	0.019	0.598	0.364	0.019
NOV	H7	0.029	0.611	0.33	0.029
NOV	H8	0.002	0.767	0.228	0.002
NOV	H9	0.106	0.136	0.651	0.106
NOV	H10	0.026	0.955	0.009	0.009
DEC	H1	0.053	0.781	0.113	0.053
DEC	H2	0.157	0.582	0.141	0.119
DEC	H3	0.037	0.853	0.072	0.037
DEC	H4	0.049	0.701	0.201	0.049
DEC	H5	0.169	0.314	0.349	0.169
DEC	H6	0.032	0.79	0.145	0.032
DEC	H7	0.112	0.489	0.288	0.112
DEC	H8	0.020	0.935	0.040	0.005
DEC	H9	0.250	0.250	0.250	0.250
DEC	H10	0.098	0.852	0.025	0.025

หมายเหตุ T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ซึ่งปรากฏในตารางที่ 10

จากตารางที่ 39 จนถึง 44 ได้แสดงถึงค่าความน่าจะเป็นแบบมีเงื่อนไขของ Node Temperature ที่มี Parent Node คือ Node Result, Node Month, Node Height ซึ่งจากตารางความน่าจะเป็น แสดงให้เห็นว่า ความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Temperature ที่ค่าความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Temperature ที่มีค่าความน่าจะเป็นสูงที่สุดเมื่อ Result = Negative ซึ่งปรากฏในตารางที่ 43 คือ

$$\Pr(\text{Temperature} = T2(21.15 - 27.85) | \text{Result} = \text{Negative}, \text{Month} = \text{OCT}, \text{Height} = H10) = 0.971$$

จากสมการความน่าจะเป็นที่ได้กล่าวมาด้านบน สามารถแปลความหมายได้ว่า ความน่าจะเป็น ของ Attribute Temperature ที่มีค่าในช่วง 21.15 – 27.85 องศาเซลเซียส เมื่อ Attribute Result มีค่าเท่ากับ Negative และ Attribute Month มีค่าเท่ากับ OCT และ Attribute Height มีค่ามากกว่า 388.5 เมตร จะมีค่าความน่าจะเป็นเท่ากับ 0.971

ดังนั้น โอกาสที่มีสัตว์ปีกป่วยตายด้วยโรคอื่นที่ไม่ใช่โรคไข้หวัดนกสูงสุด (0.971) ในเงื่อนไข ความสูงมากกว่า 388.5 เมตร และในเดือนตุลาคม อุณหภูมิอยู่ในช่วง 21.15 ถึง 27.85 องศาเซลเซียส

ตารางที่ 45 ความน่าจะเป็นแบบมีเงื่อนไขของ ลักษณะของพื้นที่ (Land Scape) ที่สัมพันธ์กับ ความสูงจากระดับน้ำทะเล (Height) สำหรับกรณี Positive class

Height	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
H1	0.009	0.035	0.261	0.051	0.009	0.599	0.009	0.009	0.009	0.009
H2	0.037	0.131	0.306	0.247	0.037	0.095	0.037	0.037	0.037	0.037
H3	0.001	0.013	0.001	0.001	0.001	0.301	0.001	0.682	0.001	0.001
H4	0.016	0.016	0.016	0.246	0.016	0.15	0.016	0.493	0.016	0.016
<b>H5</b>	0.006	0.006	0.006	0.006	0.006	0.006	<b>0.944</b>	0.006	0.006	0.006
H6	0.001	0.012	0.068	0.001	0.025	0.001	0.888	0.001	0.001	0.001
H7	0.010	0.010	0.010	0.067	0.010	0.010	0.854	0.010	0.010	0.010
H8	0.010	0.010	0.010	0.010	0.021	0.010	0.010	0.010	0.801	0.108
H9	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.78	0.024	0.024
H10	0.086	0.086	0.086	0.086	0.224	0.086	0.086	0.086	0.086	0.086

หมายเหตุ S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

จากตารางที่ 45 แสดงถึงความน่าจะเป็นแบบมีเงื่อนไขของ Attribute landscape ที่มี Attribute Result กับ Attribute Height เป็น Parent Node ตามรูปที่ 8 ซึ่งพบว่าความน่าจะเป็นแบบมีเงื่อนไขที่มากที่สุดเมื่อ Result = Positive คือ

$$\Pr(\text{Landscape} = S7 \mid \text{Result} = \text{Positive}, \text{Height} = H5 (31-34.5)) = 0.9444$$

จากสมการความน่าจะเป็นที่ได้จากตารางที่ 45 แสดงได้ว่า ความน่าจะเป็นที่ Attribute Landscape มีลักษณะพื้นที่เป็นแบบ ที่ราบลุ่มน้ำตอนบนภาคกลาง โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Positive และ Attribute Height มีค่าอยู่ในช่วง 31-34.5 เมตร จะมีค่าเท่ากับ 0.9444

ดังนั้น โอกาสที่พบสัตว์ปีกป่วยตายด้วยโรคไข้หวัดนกสูงสุด (0.944) ในเงื่อนไข พื้นที่มีลักษณะภูมิประเทศเป็นแบบที่ราบลุ่มน้ำตื้นบนภาคกลาง และมีความสูงอยู่ในช่วง 31-34.5 เมตร

ตารางที่ 46 ความน่าจะเป็นแบบมีเงื่อนไขของ ลักษณะของพื้นที่ (Land Landscape) ที่สัมพันธ์กับ ความสูงจากระดับน้ำทะเล (Height) สำหรับกรณี Negative Class

Height	S1	S2	S3	S4	S5	S6	S7	S8	S99	S10
H1	0.022	0.3	0.319	0.174	0.004	0.163	0.004	0.004	0.004	0.004
H2	0.424	0.24	0.158	0.121	0.006	0.026	0.006	0.006	0.006	0.006
H3	0.06	0.091	0.003	0.003	0.003	0.153	0.003	0.679	0.003	0.003
H4	0.006	0.004	0.004	0.582	0.004	0.116	0.004	0.271	0.004	0.004
H5	0.027	0.027	0.027	0.027	0.027	0.027	0.761	0.027	0.027	0.027
H6	0.002	0.022	0.063	0.002	0.151	0.002	0.753	0.002	0.002	0.002
H7	0.003	0.003	0.003	0.103	0.003	0.003	0.871	0.003	0.003	0.003
H8	0.000	0.000	0.000	0.000	0.246	0.000	0.000	0.000	0.639	0.113
H9	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.668	0.037	0.037
H10	0.001	0.001	0.001	0.001	<b>0.987</b>	0.001	0.001	0.001	0.001	0.001

หมายเหตุ S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

H1, H2, H3, ..., H10 แทนช่วงข้อมูลที่อยู่ใน Attribute Height ปรากฏอยู่ในตารางที่ 10

จากตารางที่ 46 แสดงถึงความน่าจะเป็นแบบมีเงื่อนไขของ Attribute landscape ที่มี Attribute Result กับ Attribute Height เป็น Parent Node ตามรูปที่ 7 ซึ่งพบว่า ความน่าจะเป็นแบบมีเงื่อนไขที่มีค่ามากที่สุด เมื่อ Result = Negative คือ

$$Pr(\text{Landscape} = S5 \mid \text{Result} = \text{Negative}, \text{Height} = H10) = 0.987$$

จากสมการความน่าจะเป็นที่ได้มาจากตารางที่ 46 สามารถแสดงได้ว่า ความน่าจะเป็นที่ Attribute Landscape มีค่าเท่ากับ S5 (ทิวเขาและหุบเขาในภาคเหนือ) โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Negative และ Attribute Height มีค่าอยู่ในช่วง มากกว่า 388.5 เมตร มีค่าความน่าจะเป็นเท่ากับ 0.987

ดังนั้น โอกาสที่พบสัตว์ปีกป่วยตายด้วยโรคอื่นที่ไม่ใช่โรคไขหวัดนกสูงสุดในเงื่อนไข ลักษณะภูมิประเทศเป็นแบบทิวเขาและหุบเขาในภาคเหนือ และ มีความสูงมากกว่า 388.5 เมตร

**ตารางที่ 47** ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงระดับความสูง น้อยกว่า 3.5 เมตร (H1) หรือ ช่วงระดับความสูงมากกว่า 3.5 ถึง 6.5 เมตร (H2) และ ลักษณะของพื้นที่ เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive class

Height	Landscape	D1	D2	D3	D4
H1	S1	0.25	0.25	0.25	0.25
H1	S2	0.248	0.228	0.346	0.178
H1	S3	0.478	0.297	0.193	0.032
H1	S4	0.649	0.117	0.117	0.117
H1	S5	0.250	0.250	0.250	0.250
H1	S6	0.388	0.119	0.226	0.267
H1	S7	0.250	0.250	0.250	0.250
H1	S8	0.250	0.250	0.250	0.250
H1	S9	0.250	0.250	0.250	0.250
H1	S10	0.250	0.250	0.250	0.250
H2	S1	0.250	0.250	0.250	0.250
H2	S2	0.306	0.253	0.288	0.153
H2	S3	0.549	0.232	0.131	0.088
H2	S6	0.321	0.321	0.179	0.179
H2	S7	0.250	0.250	0.250	0.250
H2	S8	0.250	0.250	0.250	0.250
H2	S9	0.250	0.250	0.250	0.250
H2	S10	0.250	0.250	0.250	0.250

หมายเหตุ D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

**ตารางที่ 48** ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์  
กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงมากกว่า 6.5 เมตร ถึง 12 เมตร  
(H3) หรือ ช่วงระดับความสูงมากกว่า 12 เมตร ถึง 31 เมตร (H4) และ ลักษณะของ  
พื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class

Height	Landscape	D1	D2	D3	D4
H3	S1	0.250	0.250	0.250	0.250
H3	S2	0.332	0.339	0.264	0.066
H3	S3	0.250	0.250	0.250	0.250
H3	S4	0.250	0.250	0.250	0.250
H3	S5	0.250	0.250	0.250	0.250
H3	S6	0.303	0.200	0.29	0.207
H3	S7	0.250	0.250	0.250	0.250
H3	S8	0.267	0.143	0.232	0.358
H3	S9	0.250	0.250	0.250	0.250
H3	S10	0.250	0.250	0.250	0.250
H4	S1	0.250	0.250	0.250	0.250
H4	S2	0.250	0.250	0.250	0.250
H4	S3	0.250	0.250	0.250	0.250
H4	S4	0.439	0.219	0.287	0.054
H4	S5	0.250	0.250	0.250	0.250
H4	S6	0.259	0.08	0.08	0.581
H4	S7	0.250	0.250	0.250	0.250
H4	S8	<b>0.654</b>	0.099	0.09	0.158
H4	S9	0.250	0.250	0.250	0.250
H4	S10	0.250	0.250	0.250	0.250

หมายเหตุ D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11  
H1, H2, H3, ..., H10 แทนช่วงข้อมูลที่อยู่ใน Attribute Height ปรากฏอยู่ในตารางที่ 10

**ตารางที่ 49** ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์ กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงมากกว่า 31 เมตร ถึง 34.5 เมตร (H5) หรือ ช่วงระดับความสูงมากกว่า 34.5 เมตร ถึง 97 เมตร (H6) และ ลักษณะของ พื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class

Height	Landscape	D1	D2	D3	D4
H5	S1	0.250	0.250	0.250	0.250
H5	S2	0.250	0.250	0.250	0.250
H5	S3	0.250	0.250	0.250	0.250
H5	S4	0.250	0.250	0.250	0.250
H5	S5	0.250	0.250	0.250	0.250
H5	S6	0.250	0.250	0.250	0.250
H5	S7	0.358	0.105	0.355	0.183
H5	S8	0.250	0.250	0.250	0.250
H5	S9	0.250	0.250	0.250	0.250
H5	S10	0.250	0.250	0.250	0.250
H6	S1	0.250	0.250	0.250	0.250
H6	S2	0.452	0.336	0.141	0.071
H6	S3	0.377	0.311	0.245	0.067
H6	S4	0.250	0.250	0.250	0.250
H6	S5	0.373	0.161	0.282	0.184
H6	S6	0.250	0.250	0.250	0.250
H6	S7	0.376	0.251	0.292	0.081
H6	S8	0.250	0.250	0.250	0.250
H6	S9	0.250	0.250	0.250	0.250
H6	S10	0.250	0.250	0.250	0.250

หมายเหตุ D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

**ตารางที่ 50** ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์  
กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงมากกว่า 97 เมตร ถึง 124.5 เมตร  
(H7) และ ช่วงระดับความสูงมากกว่า 124.5 เมตร ถึง 382.5 เมตร (H8) และ ลักษณะ  
ของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class

Height	Landscape	D1	D2	D3	D4
H7	S1	0.250	0.250	0.250	0.250
H7	S2	0.250	0.250	0.250	0.250
H7	S3	0.250	0.250	0.250	0.250
H7	S4	0.298	0.496	0.103	0.103
H7	S5	0.250	0.250	0.250	0.250
H7	S6	0.250	0.250	0.250	0.250
H7	S7	0.397	0.251	0.303	0.049
H7	S8	0.250	0.250	0.250	0.250
H7	S9	0.250	0.250	0.250	0.250
H7	S10	0.250	0.250	0.250	0.250
H8	S1	0.250	0.250	0.250	0.250
H8	S2	0.250	0.250	0.250	0.250
H8	S3	0.250	0.250	0.250	0.250
H8	S4	0.250	0.250	0.250	0.250
H8	S5	0.195	0.195	0.414	0.195
H8	S6	0.250	0.250	0.250	0.250
H8	S7	0.250	0.250	0.250	0.250
H8	S8	0.250	0.250	0.250	0.250
H8	S9	0.547	0.209	0.23	0.015
H8	S10	0.239	0.072	0.346	0.343

หมายเหตุ D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

ตารางที่ 51 ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์  
กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในมากกว่า 382.5 เมตร ถึง 388.5 เมตร  
(H9) หรือ ช่วงระดับความสูงมากกว่า 388.5 เมตร (H10) และ ลักษณะของพื้นที่เลี้ยง  
สัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class

Height	Landscape	D1	D2	D3	D4
H9	S1	0.250	0.250	0.250	0.250
H9	S2	0.250	0.250	0.250	0.250
H9	S3	0.250	0.250	0.250	0.250
H9	S4	0.250	0.250	0.250	0.250
H9	S5	0.250	0.250	0.250	0.250
H9	S6	0.250	0.250	0.250	0.250
H9	S7	0.250	0.250	0.250	0.250
H9	S8	0.465	0.137	0.25	0.147
H9	S9	0.250	0.250	0.250	0.250
H9	S10	0.250	0.250	0.250	0.250
H10	S1	0.250	0.250	0.250	0.250
H10	S2	0.250	0.250	0.250	0.250
H10	S3	0.250	0.250	0.250	0.250
H10	S4	0.250	0.250	0.250	0.250
H10	S5	0.179	0.179	0.463	0.179
H10	S6	0.250	0.250	0.250	0.250
H10	S7	0.250	0.250	0.250	0.250
H10	S8	0.250	0.250	0.250	0.250
H10	S9	0.250	0.250	0.250	0.250
H10	S10	0.250	0.250	0.250	0.250

หมายเหตุ D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

จากตารางที่ 47 ถึง 51 ได้แสดงถึงความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Death ที่มี Attribute Result, Attribute height และ Attribute Landscape ซึ่งพบว่า ความน่าจะเป็นที่มีมากที่สุด เมื่อ Attribute Result = positive ซึ่งมีปรากฏในตารางที่ 48 คือ

$$\Pr(\text{Death} = D1 (< 4.5) | \text{Result} = \text{Positive}, \text{Height} = H4, \text{landscape} = S8) = 0.654$$

จากสมการความน่าจะเป็นข้างต้น สามารถแปลความหมายได้ว่า ความน่าจะเป็นของ Attribute Death มีค่าอยู่ในช่วงน้อยกว่า 5 ตัว โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Positive และ Attribute Height มีค่าอยู่ในช่วง 12 ถึง 31 เมตร และ Attribute Landscape มีค่าเท่ากับ บริเวณขอบที่ราบภาคกลาง จะมีความน่าจะเป็นเท่ากับ 0.654

ดังนั้น โอกาสที่มีสัตว์ปีกป่วย ตาย ด้วยโรคไข้หวัดนกสูงสุด(0.654) ในเงื่อนไขของความสูงของพื้นที่ อยู่ในช่วง 12 – 31 เมตร ในลักษณะภูมิประเทศ ขอบที่ราบภาคกลาง และมีจำนวนสัตว์ปีกป่วยตายน้อยกว่า 5 ตัว

**ตารางที่ 52** ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์  
กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงระดับความสูง น้อยกว่า 3.5 เมตร  
(H1) หรือ ช่วงระดับความสูงมากกว่า 3.5 ถึง 6.5 เมตร (H2) และ ลักษณะของพื้นที่  
เลี้ยงสัตว์ปีก (Landscape) สำหรับ กรณี Negative Class

Height	Landscape	D1	D2	D3	D4
H1	S1	0.363	0.319	0.125	0.192
H1	S2	0.445	0.454	0.038	0.063
H1	S3	0.727	0.19	0.059	0.025
H1	S4	0.661	0.228	0.087	0.023
H1	S5	0.250	0.250	0.250	0.250
H1	S6	0.603	0.217	0.103	0.076
H1	S7	0.250	0.250	0.250	0.250
H1	S8	0.250	0.250	0.250	0.250
H1	S9	0.250	0.250	0.250	0.250
H1	S10	0.250	0.250	0.250	0.250
H2	S1	0.563	0.354	0.062	0.021
H2	S2	0.446	0.297	0.123	0.134
H2	S3	0.403	0.326	0.212	0.058
H2	S4	0.536	0.284	0.136	0.045
H2	S5	0.250	0.250	0.250	0.250
H2	S6	0.451	0.231	0.141	0.176
H2	S7	0.250	0.250	0.250	0.250
H2	S8	0.250	0.250	0.250	0.250
H2	S9	0.250	0.250	0.250	0.250
H2	S10	0.250	0.250	0.250	0.250

หมายเหตุ D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

ตารางที่ 53 ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์  
กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงมากกว่า 6.5 เมตร ถึง 12 เมตร (H3)  
หรือ ช่วงระดับความสูงมากกว่า 12 เมตร ถึง 31 เมตร (H4) และ ลักษณะของพื้นที่เลี้ยง  
สัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class

Height	Landscape	D1	D2	D3	D4
H3	S1	0.44	0.48	0.04	0.04
H3	S2	0.639	0.174	0.137	0.05
H3	S3	0.250	0.250	0.250	0.250
H3	S4	0.250	0.250	0.250	0.250
H3	S5	0.250	0.250	0.250	0.250
H3	S6	0.393	0.369	0.167	0.071
H3	S7	0.250	0.250	0.250	0.250
H3	S8	0.559	0.314	0.084	0.044
H3	S9	0.250	0.250	0.250	0.250
H3	S10	0.250	0.250	0.250	0.250
H4	S1	0.231	0.231	0.307	0.231
H4	S2	0.250	0.250	0.250	0.250
H4	S3	0.250	0.250	0.250	0.250
H4	S4	0.744	0.22	0.027	0.009
H4	S5	0.250	0.250	0.250	0.250
H4	S6	0.422	0.433	0.085	0.06
H4	S7	0.250	0.250	0.250	0.250
H4	S8	<b>0.836</b>	0.129	0.017	0.019
H4	S9	0.250	0.250	0.250	0.250
H4	S10	0.250	0.250	0.250	0.250

หมายเหตุ D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

**ตารางที่ 54** ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์  
กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงระดับความสูงมากกว่า 31 เมตร  
ถึง 34.5 เมตร (H5) หรือ ช่วงระดับความสูง มากกว่า 34.5 เมตร ถึง 97 เมตร (H6) และ  
ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class

Height	Landscape	D1	D2	D3	D4
H5	S1	0.250	0.250	0.250	0.250
H5	S2	0.250	0.250	0.250	0.250
H5	S3	0.250	0.250	0.250	0.250
H5	S4	0.250	0.250	0.250	0.250
H5	S5	0.250	0.250	0.250	0.250
H5	S6	0.250	0.250	0.250	0.250
H5	S7	0.384	0.235	0.257	0.124
H5	S8	0.250	0.250	0.250	0.250
H5	S9	0.250	0.250	0.250	0.250
H5	S10	0.250	0.250	0.250	0.250
H6	S1	0.250	0.250	0.250	0.250
H6	S2	0.567	0.308	0.063	0.063
H6	S3	0.457	0.270	0.223	0.049
H6	S4	0.250	0.250	0.250	0.250
H6	S5	0.490	0.330	0.150	0.030
H6	S6	0.250	0.250	0.250	0.250
H6	S7	0.528	0.289	0.154	0.029
H6	S8	0.250	0.250	0.250	0.250
H6	S9	0.250	0.250	0.250	0.250
H6	S10	0.250	0.250	0.250	0.250

หมายเหตุ D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

ตารางที่ 55 ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์  
กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงระดับความสูงมากกว่า 97 เมตร ถึง  
124.5 เมตร (H7) หรือ ช่วงระดับความสูง มากกว่า 124.5 เมตร ถึง 382.5 เมตร (H8)  
และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class

Height	Landscape	D1	D2	D3	D4
H7	S1	0.250	0.250	0.250	0.250
H7	S2	0.250	0.250	0.250	0.250
H7	S3	0.250	0.250	0.250	0.250
H7	S4	0.618	0.32	0.033	0.029
H7	S5	0.250	0.250	0.250	0.250
H7	S6	0.250	0.250	0.250	0.250
H7	S7	0.52	0.362	0.113	0.005
H7	S8	0.250	0.250	0.250	0.250
H7	S9	0.250	0.250	0.250	0.250
H7	S10	0.250	0.250	0.250	0.250
H8	S1	0.250	0.250	0.250	0.250
H8	S2	0.250	0.250	0.250	0.250
H8	S3	0.250	0.250	0.250	0.250
H8	S4	0.250	0.250	0.250	0.250
H8	S5	0.534	0.385	0.071	0.01
H8	S6	0.250	0.250	0.250	0.250
H8	S7	0.250	0.250	0.250	0.250
H8	S8	0.250	0.250	0.250	0.250
H8	S9	0.599	0.316	0.075	0.01
H8	S10	0.473	0.433	0.084	0.01

หมายเหตุ D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

ตารางที่ 56 ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยตาย (Death) ที่สัมพันธ์  
กับ ความสูงจากระดับน้ำทะเล (Height) ที่อยู่ในช่วงระดับความสูงมากกว่า 382.5 เมตร  
ถึง 388.5 เมตร (H9) หรือ ช่วงระดับความสูง มากกว่า 388.5 เมตร (H10) และ ลักษณะ  
ของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class

Height	Landscape	D1	D2	D3	D4
H9	S1	0.250	0.250	0.250	0.250
H9	S2	0.250	0.250	0.250	0.250
H9	S3	0.250	0.250	0.250	0.250
H9	S4	0.250	0.250	0.250	0.250
H9	S5	0.250	0.250	0.250	0.250
H9	S6	0.250	0.250	0.250	0.250
H9	S7	0.250	0.250	0.250	0.250
H9	S8	0.567	0.273	0.099	0.062
H9	S9	0.250	0.250	0.250	0.250
H9	S10	0.250	0.250	0.250	0.250
H10	S1	0.250	0.250	0.250	0.250
H10	S2	0.250	0.250	0.250	0.250
H10	S3	0.250	0.250	0.250	0.250
H10	S4	0.250	0.250	0.250	0.250
H10	S5	0.643	0.311	0.039	0.008
H10	S6	0.250	0.250	0.250	0.250
H10	S7	0.250	0.250	0.250	0.250
H10	S8	0.250	0.250	0.250	0.250
H10	S9	0.250	0.250	0.250	0.250
H10	S10	0.250	0.250	0.250	0.250

หมายเหตุ D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

จากตารางที่ 52 ถึง 56 แสดงถึงความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Death ที่มี Attribute Result, Attribute height และ Attribute Landscape ซึ่งพบว่ากรณีที่มี Attribute Result = Negative ที่มีค่าความน่าจะเป็น มากที่สุด ซึ่งอยู่ในตารางที่ 53 คือ

$$\Pr(\text{Death} = D1 \mid \text{Result} = \text{Negative}, \text{Height} = H4, \text{Landscape} = S8) = 0.836$$

ซึ่งสามารถแปลความหมายได้ว่า ความน่าจะเป็นที่ Attribute Death มีค่าเท่ากับ D1 ซึ่งหมายถึงจำนวนสัตว์ปีกป่วยตายน้อยกว่า 1 ตัว โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Negative และ Attribute Height มีค่าเท่ากับ H4 ซึ่งหมายถึง มีความสูงจากระดับน้ำทะเลอยู่ในช่วง 12 เมตร ถึง 31 เมตร และ Attribute Landscape มีค่าเท่ากับ S8 ซึ่งหมายถึง บริเวณขอบที่ราบภาคกลาง มีความน่าจะเป็นเท่ากับ 0.836

ดังนั้น โอกาสที่สัตว์ปีกป่วยตายด้วยโรคอื่นที่ไม่ใช่โรคไข้หวัดนกสูงสุด (0.836) ในเงื่อนไขลักษณะภูมิประเทศเป็นแบบของที่ราบภาคกลาง และมีจำนวนสัตว์ปีกที่ตายน้อยกว่า 1 ตัว และมีระดับความสูงของพื้นที่อยู่ในช่วง 12-31 เมตร

ตารางที่ 57 ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยที่ถูกทำลาย (Destroy) ที่สัมพันธ์ กับ จำนวนสัตว์ปีกที่ตาย (Death) ที่อยู่ในช่วงน้อย 5 ตัว (D1) หรือ อยู่ใน ช่วง มากกว่า 5 ตัว ถึง 18 ตัว และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class

Death	Landscape	De1	De2	De3	De4	De5
D1	S1	0.200	0.200	0.200	0.200	0.200
D1	S2	0.136	0.383	0.308	0.052	0.121
D1	S3	0.109	0.481	0.286	0.074	0.051
D1	S4	0.045	0.209	0.263	0.045	0.438
D1	S5	0.092	0.248	0.072	0.24	0.347
D1	S6	0.078	0.250	0.331	0.152	0.190
D1	S7	0.032	0.340	0.377	0.133	0.118
D1	S8	0.083	0.301	0.162	0.095	0.359
D1	S9	0.055	0.394	0.451	0.080	0.020
D1	S10	0.137	0.137	0.318	0.271	0.137
D2	S1	0.200	0.200	0.200	0.200	0.200
D2	S2	0.277	0.108	0.499	0.058	0.058
D2	S3	0.076	0.516	0.257	0.122	0.029
D2	S4	0.235	0.407	0.214	0.072	0.072
D2	S5	0.121	0.121	0.450	0.121	0.188
D2	S6	0.040	0.350	0.482	0.033	0.096
D2	S7	0.034	0.391	0.455	0.104	0.017

หมายเหตุ De1, De2, De3, De4, De5 แทนช่วงข้อมูลที่อยู่ใน Attribute Destruction

ซึ่งปรากฏในตารางที่ 10

S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

ตารางที่ 58 ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยที่ถูกทำลาย (Destroy) ที่สัมพันธ์ กับ จำนวนสัตว์ปีกที่ตาย (Death) อยู่ในช่วง มากกว่า 18 ตัว ถึง 98 ตัว หรือ อยู่ในช่วงมากกว่า 98 ตัว และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Positive Class

Death	Landscape	De1	De2	De3	De4	De5
D3	S1	0.200	0.200	0.200	0.200	0.200
D3	S2	0.093	0.108	0.483	0.243	0.073
D3	S3	0.290	0.244	0.238	0.130	0.098
D3	S4	0.107	0.343	0.336	0.107	0.107
D3	S5	0.070	0.070	0.572	0.217	0.070
D3	S6	0.217	0.271	0.277	0.190	0.044
D3	S7	0.142	0.210	0.502	0.081	0.065
D3	S8	0.143	0.126	0.290	0.172	0.270
D3	S9	0.184	0.252	0.217	0.26	0.087
D3	S10	0.255	0.302	0.184	0.114	0.144
D4	S1	0.200	0.200	0.200	0.200	0.200
D4	S2	0.276	0.208	0.172	0.172	0.172
D4	S3	0.118	0.118	0.118	0.118	0.527
D4	S4	0.200	0.200	0.200	0.200	0.200
D4	S5	0.113	0.113	0.113	0.113	0.550
D4	S6	0.085	0.068	0.082	0.114	0.650
D4	S7	0.154	0.113	0.195	0.043	0.495
D4	S8	0.082	0.019	0.035	0.100	<b>0.764</b>

หมายเหตุ De1, De2, De3, De4, De5 แทนช่วงข้อมูลที่อยู่ใน Attribute Destruction

ซึ่งปรากฏในตารางที่ 10

S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

จากตารางที่ 57 ถึง ตารางที่ 58 ได้ อธิบายถึงความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Destruction ที่มี Attribute Result, Attribute Death, Attribute Landscape เป็นเงื่อนไข โดยที่ค่าความน่าจะเป็นที่มากที่สุดในกรณีที่ Result = Positive ซึ่งปรากฏในตารางที่ 58 คือ

$$\Pr(\text{Destruction} = \text{De5} | \text{Result} = \text{positive}, \text{Death} = \text{D4}, \text{Landscape} = \text{S8}) = 0.764$$

ซึ่งจากสมการสามารถอธิบายได้ว่า ความน่าจะเป็นที่ Attribute Destruction มีค่าเท่ากับ De5 ซึ่งหมายถึง จำนวนสัตว์ปีกที่ถูกทำลายอยู่ในช่วงมากกว่า 969 ตัว โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Positive และ Attribute Death มีค่าเท่ากับ D4 ซึ่งหมายถึง จำนวนสัตว์ปีกที่ตายอยู่ในช่วง มากกว่า 98 ตัว และ Attribute Landscape มีค่าเท่ากับ S8 ซึ่งหมายถึง ลักษณะภูมิประเทศเป็นแบบขอบที่ราบภาคกลาง มีความน่าจะเป็นเท่ากับ 0.764

ดังนั้น โอกาสที่พบสัตว์ปีกป่วยตายด้วยโรคไข้หวัดนกสูงสุด (0.764) ในเงื่อนไข มีจำนวนสัตว์ปีกที่ถูกทำลาย มากกว่า 969 ตัว และมีสัตว์ปีกที่ป่วยตายอยู่ในช่วง มากกว่า 98 ตัว และมีลักษณะภูมิประเทศแบบขอบที่ราบภาคกลาง

ตารางที่ 59 ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยที่ถูกทำลาย (Destroy) ที่สัมพันธ์ กับ จำนวนสัตว์ปีกที่ตาย (Death) ในช่วงน้อย 5 ตัว (D1) หรือ อยู่ในช่วงมากกว่า 5 ตัว ถึง 18 ตัว และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class

Death	Landscape	De1	De2	De3	De4	De5
D1	S1	<b>0.912</b>	0.034	0.018	0.018	0.018
D1	S2	0.79	0.155	0.032	0.012	0.012
D1	S3	0.712	0.193	0.072	0.012	0.012
D1	S4	0.899	0.053	0.034	0.008	0.006
D1	S5	0.869	0.107	0.021	0.001	0.001
D1	S6	0.719	0.213	0.037	0.016	0.016
D1	S7	0.643	0.212	0.122	0.016	0.006
D1	S8	0.685	0.162	0.136	0.008	0.01
D1	S9	0.735	0.197	0.066	0.001	0.001
D1	S10	0.775	0.151	0.058	0.01	0.005
D2	S1	0.866	0.043	0.043	0.024	0.024
D2	S2	0.881	0.067	0.018	0.018	0.018
D2	S3	0.762	0.105	0.060	0.045	0.027
D2	S4	0.834	0.075	0.057	0.017	0.017
D2	S5	0.798	0.151	0.043	0.006	0.002
D2	S6	0.682	0.177	0.098	0.022	0.022
D2	S7	0.683	0.163	0.14	0.01	0.004
D2	S10	0.655	0.192	0.13	0.017	0.006

หมายเหตุ De1, De2, De3, De4, De5 แทนช่วงข้อมูลที่อยู่ใน Attribute Destruction

ซึ่งปรากฏในตารางที่ 10

S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

**ตารางที่ 60** ความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนสัตว์ปีกป่วยที่ถูกทำลาย (Destroy) ที่สัมพันธ์กับ จำนวนสัตว์ปีกที่ตาย (Death) อยู่ในช่วง มากกว่า 18 ตัว ถึง 98 ตัว หรือ อยู่ในช่วงมากกว่า 98 ตัว และ ลักษณะของพื้นที่เลี้ยงสัตว์ปีก (Land Scape) สำหรับ กรณี Negative Class

Death	Landscape	De1	De2	De3	De4	De5
D3	S1	0.541	0.115	0.115	0.115	0.115
D3	S2	0.698	0.068	0.099	0.068	0.068
D3	S3	0.686	0.113	0.108	0.047	0.047
D3	S4	0.552	0.145	0.134	0.091	0.079
D3	S5	0.717	0.158	0.098	0.017	0.01
D3	S6	0.604	0.223	0.054	0.059	0.061
D3	S7	0.558	0.224	0.169	0.037	0.012
D3	S8	0.668	0.078	0.14	0.076	0.038
D3	S9	0.675	0.138	0.127	0.054	0.006
D3	S10	0.624	0.115	0.186	0.047	0.028
D4	S1	0.255	0.167	0.167	0.167	0.244
D4	S2	0.537	0.072	0.132	0.187	0.072
D4	S3	0.381	0.187	0.165	0.133	0.133
D4	S4	0.237	0.191	0.191	0.191	0.191
D4	S5	0.656	0.083	0.089	0.089	0.083
D4	S6	0.379	0.117	0.132	0.172	0.2
D4	S7	0.667	0.056	0.077	0.097	0.103
D4	S8	0.45	0.064	0.123	0.095	0.268
D4	S9	0.561	0.039	0.063	0.275	0.063
D4	S10	0.392	0.125	0.125	0.17	0.187

หมายเหตุ De1, De2, De3, De4, De5 แทนช่วงข้อมูลที่อยู่ใน Attribute Destruction

ซึ่งปรากฏในตารางที่ 10

S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

จากตารางที่ 59 ถึง ตารางที่ 60 เป็นตารางที่อธิบายถึงความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Destruction ที่มี Attribute Result, Attribute Death, Attribute Landscape เป็นเงื่อนไข โดยที่ค่าความน่าจะเป็นที่มากที่สุดในกรณีที่ Result = Negative ซึ่งอยู่ในตารางที่ 59 คือ

$$\Pr(\text{Destruction} = \text{De1} | \text{Result} = \text{Negative}, \text{Death} = \text{D1}, \text{landscape} = \text{S1}) = 0.912$$

ซึ่งจากสมการสามารถอธิบายได้ว่า ความน่าจะเป็นที่ Attribute Destruction มีค่าเท่ากับ De1 ซึ่งหมายถึง มีจำนวนสัตว์ปีกที่ถูกทำลาย น้อยกว่า 1 ตัว โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Negative และ Attribute Death มีค่าเท่ากับ D1 หมายถึง จำนวนสัตว์ปีกที่ตาย อยู่ในช่วง น้อยกว่า 5 ตัว และ Attribute Landscape มีค่าเท่ากับ S1 ซึ่งหมายถึง แบบคาบสมุทรภายใต้ชายฝั่งตะวันตก ได้ มีความน่าจะเป็นเท่ากับ 0.912

ดังนั้น โอกาสที่สัตว์ปีกป่วยตายด้วยโรคอื่นไม่ใช่โรคไข้หวัดนกสูงสุด (0.912) ในเงื่อนไข จำนวนสัตว์ปีกที่ถูกทำลายอยู่ในช่วง น้อยกว่า 1 ตัว และ จำนวนสัตว์ปีกที่ป่วยตายน้อยกว่า 5 ตัว และ ลักษณะภูมิประเทศเป็นแบบคาบสมุทรภายใต้ชายฝั่งตะวันตก

ตารางที่ 61 ตัวอย่างความน่าจะเป็นแบบมีเงื่อนไขของ ความชื้น (Humidity) ที่สัมพันธ์ กับ ช่วงเดือนธันวาคม (Month) สภาพพื้นที่ (Landscape) อุณหภูมิ(Temperature) และ ความสูงของพื้นที่ (Height) ในกรณีที่มีผลเฉลยเป็น Positive Class

Month	Landscape	Temperature	Height	Hu2	Hu3	Hu1
DEC	S7	T2	H1	0.333	0.333	0.333
DEC	S7	T2	H1	0.333	0.333	0.333
DEC	S7	T2	H10	0.333	0.333	0.333
DEC	S7	T2	H10	0.333	0.333	0.333
DEC	S7	T2	H2	0.333	0.333	0.333
DEC	S7	T2	H2	0.333	0.333	0.333
DEC	S7	T2	H3	0.333	0.333	0.333
DEC	S7	T2	H3	0.333	0.333	0.333
DEC	S7	T2	H4	0.333	0.333	0.333
DEC	S7	T2	H4	0.333	0.333	0.333
DEC	S7	T2	H5	0.897	0.051	0.051
DEC	S7	T2	H5	0.333	0.333	0.333
DEC	S7	T2	H6	<b>0.983</b>	0.004	0.013
DEC	S7	T2	H6	0.333	0.333	0.333
DEC	S7	T2	H7	0.333	0.333	0.333
DEC	S7	T2	H7	0.333	0.333	0.333
DEC	S7	T2	H8	0.333	0.333	0.333

หมายเหตุ Hu1, Hu2, Hu3 แทนช่วงข้อมูลที่อยู่ใน Attribute Humidity ซึ่งปรากฏในตารางที่ 10  
 S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11  
 H1, H2, H3, ..., H10 แทนช่วงข้อมูลที่อยู่ใน Attribute Height ปรากฏอยู่ในตารางที่ 10  
 T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ปรากฏอยู่ในตารางที่ 10

จากตารางที่ 61 เป็นตัวอย่างตารางที่แสดงความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Humidity ที่มี Attribute Result, Attribute Month Attribute Landscape Attribute Temperature และ Attribute Height เป็นเงื่อนไข โดยที่ค่าความน่าจะเป็นที่มากที่สุดในการกรณีที่ Result = Positive คือ

$$\Pr(\text{Humidity} = \text{Hu2} | \text{Result} = \text{Positive}, \text{Month} = \text{DEC}, \text{Landscape} = \text{S7}, \\ \text{Temperature} = \text{T2}, \text{Height} = \text{H6}) = 0.983$$

ซึ่งจากสมการความน่าจะเป็นสามารถแปลความหมายได้ว่า ความน่าจะเป็นที่ Attribute Humidity มีค่าเท่ากับ Hu2 ซึ่งหมายถึงมีความชื้นสัมพัทธ์อยู่ในช่วง 60.5 – 75.09 % โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Positive และ Attribute Month มีค่าเท่ากับ DEC และ Attribute Landscape มีค่าเท่ากับ S7 ซึ่งหมายถึง มีสภาพภูมิประเทศเป็นแบบ ที่ราบลุ่มน้ำภาคกลางตอนบน และ Attribute Temperature มีค่าเท่ากับ T2 ซึ่งหมายถึงอยู่ในอุณหภูมิช่วง 21.15-27.85 องศาเซลเซียส และ Attribute Height มีค่าเท่ากับ H6 ซึ่งหมายถึงมีความสูงจากระดับน้ำทะเลอยู่ในช่วง 34.5 – 97 เมตร มีค่าความน่าจะเป็นเท่ากับ 0.983

ดังนั้น โอกาสที่สัตว์ปีกป่วยตายด้วยโรคไข้หวัดนกสูงสุด (0.983) ในเงื่อนไข ความชื้นสัมพัทธ์อยู่ในช่วง 60.5 – 75.09 % และอยู่ในเดือน ธันวาคม และ อุณหภูมิอยู่ในช่วง 21.15-27.85 องศาเซลเซียส และ ลักษณะภูมิประเทศเป็นแบบ ที่ราบลุ่มน้ำภาคกลางตอนบน และ ความสูงของพื้นที่อยู่ในช่วง 34.5 – 97 เมตร

ตารางที่ 62 ตัวอย่างความน่าจะเป็นแบบมีเงื่อนไขของ ความชื้น (Humidity) ที่สัมพันธ์ กับ ช่วงเดือน ตุลาคม (Month) สภาพพื้นที่ (Landscape) อุณหภูมิ(Temperature) และ ความสูงของพื้นที่ (Height) ในกรณีที่มีผลเฉลยเป็น Negative Class

month	Landscape	Temperature	Height	Hu2	Hu3	Hu1
OCT	S10	T1	H1	0.333	0.333	0.333
OCT	S10	T1	H10	0.333	0.333	0.333
OCT	S10	T1	H2	0.333	0.333	0.333
OCT	S10	T1	H3	0.333	0.333	0.333
OCT	S10	T1	H4	0.333	0.333	0.333
OCT	S10	T1	H5	0.333	0.333	0.333
OCT	S10	T1	H6	0.333	0.333	0.333
OCT	S10	T1	H7	0.333	0.333	0.333
OCT	S10	T1	H8	0.333	0.333	0.333
OCT	S10	T1	H9	0.316	0.602	0.081
OCT	S10	T2	H1	0.333	0.333	0.333
OCT	S10	T2	H10	0.333	0.333	0.333
OCT	S10	T2	H2	0.333	0.333	0.333
OCT	S10	T2	H3	0.333	0.333	0.333
OCT	S10	T2	H4	0.333	0.333	0.333
OCT	S10	T2	H5	0.333	0.333	0.333
OCT	S10	T2	H6	<b>0.990</b>	0.008	0.002
OCT	S10	T2	H7	0.333	0.333	0.333
OCT	S10	T2	H8	0.333	0.333	0.333

หมายเหตุ Hu1, Hu2, Hu3 แทนช่วงข้อมูลที่อยู่ใน Attribute Humidity ซึ่งปรากฏในตารางที่ 10

S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

H1, H2, H3, ..., H10 แทนช่วงข้อมูลที่อยู่ใน Attribute Height ปรากฏอยู่ในตารางที่ 10

T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ปรากฏอยู่ในตารางที่ 10

จากตารางที่ 62 เป็นตัวอย่างตารางที่แสดงความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Humidity ที่มี Attribute Result, Attribute Month Attribute Landscape Attribute Temperature และ Attribute Height เป็นเงื่อนไข โดยที่ ค่าความน่าจะเป็นแบบมีเงื่อนไขที่มากที่สุดเมื่อ Result = Negative คือ

$$\Pr(\text{Humidity} = \text{Hu2} | \text{Result} = \text{Negative}, \text{Month} = \text{OCT}, \text{Landscape} = \text{S10}, \\ \text{Temperature} = \text{T2}, \text{Height} = \text{H6}) = 0.990$$

ซึ่งจากสมการความน่าจะเป็นสามารถแปลความหมายได้ว่า ความน่าจะเป็นที่ Attribute Humidity มีค่าเท่ากับ Hu2 ซึ่งหมายถึงมีความชื้นสัมพัทธ์อยู่ในช่วง 60.5 – 75.09 % โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Negative และ Attribute Month มีค่าเท่ากับ OCT และ Attribute Landscape มีค่าเท่ากับ S10 ซึ่งหมายถึง มีสภาพภูมิประเทศอยู่ในพื้นที่ที่เรียกว่า แอ่งสกลนคร และ Attribute Temperature มีค่าเท่ากับ T2 ซึ่งหมายถึงอยู่ในอุณหภูมิช่วงมากกว่า 21.15 ถึงน้อยกว่า 27.85 องศาเซลเซียส และ Attribute Height มีค่าเท่ากับ H6 ซึ่งหมายถึงมีความสูงจากระดับน้ำทะเลอยู่ในช่วง มากกว่า 34.5 เมตร ถึง น้อยกว่า 97 เมตร มีค่าความน่าจะเป็นเท่ากับ 0.990

ดังนั้น โอกาสที่สัตว์ปีกป่วยตายด้วยโรคอื่นที่ไม่ใช่โรคไข้หวัดนกสูงสุด (0.990) ในเงื่อนไข ความชื้นสัมพัทธ์อยู่ในช่วง 60.5 – 75.09 % และอยู่ในเดือน ตุลาคม อุณหภูมิอยู่ในช่วงมากกว่า 21.15 ถึง 27.85 องศาเซลเซียส ซึ่งมี ลักษณะภูมิประเทศเป็นแบบแอ่งสกลนคร และ ความสูงของพื้นที่อยู่ในช่วง มากกว่า 34.5 เมตร ถึง 97 เมตร

ตารางที่ 63 ตัวอย่างความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนชั่วโมงที่มีแสงแดด (Day Light Hour) ที่สัมพันธ์กับ ช่วงเดือนที่มีการระบาด (Month) สภาพพื้นที่ (Landscape) และ ความสูงของพื้นที่ (Height) ในกรณีที่มีผลเฉลยเป็น Positive Class

Month	Landscap	Height	Land																							
			L15	L13	L23	L21	L1	L20	L9	L18	L22	L3	L5	L14	L16	L10	L7	L8	L17	L2	L11	L12	L4	L19	L6	
NOV	S8	H1	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
NOV	S8	H4	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.18	0.04	
NOV	S8	H6	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
NOV	S8	H8	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
NOV	S8	H10	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
NOV	S8	H7	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
NOV	S8	H3	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0	0.01	0.06	0.02	0.01	0.03	0	0	<b>0.84</b>	0	
NOV	S8	H5	0	0	0	0	0	0	0	0	0	0	0	0	0	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
NOV	S8	H2	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
NOV	S8	H9	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.32	0.03	
JUL	S4	H1	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	

หมายเหตุ L1, L2, L3, ..., L23 แทนช่วงข้อมูลที่อยู่ใน Attribute Day light Hour ซึ่งปรากฏในตารางที่ 10  
 S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11  
 H1, H2, H3, ..., H10 แทนช่วงข้อมูลที่อยู่ใน Attribute Height ปรากฏอยู่ในตารางที่ 10

ตารางที่ 64 ตัวอย่างความน่าจะเป็นแบบมีเงื่อนไขของ จำนวนชั่วโมงที่มีแสงแดด (Day Light Hour) ที่สัมพันธ์กับ ช่วงเดือนที่มีการระบาด (Month) สภาพพื้นที่ (Landscape) และ ความสูงของพื้นที่ (Height) ในกรณีที่มีผลเฉลยเป็น Negative class

Month	Landscap	Height	Land																							
			L15	L13	L23	L21	L1	L20	L9	L18	L22	L3	L5	L14	L16	L10	L7	L8	L17	L2	L11	L12	L4	L19	L6	
AUG	S7	H1	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
AUG	S7	H4	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
AUG	S7	H6	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.01	0.2	0.01	0.03	0.13	0.01	0.01	0.01	0.01	0.01	0.18	
AUG	S7	H8	0.01	0.01	0.01	0.01	0.04	0.01	0.04	0.01	0.01	0.26	0.01	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
AUG	S7	H10	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
AUG	S7	H7	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0	0	0.01	0	0.03	0	0	0	0	0	0	
AUG	S7	H3	0	0	0	0	0.02	0	0	0.01	0	<b>0.85</b>	0	0	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
AUG	S7	H5	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.09	0.04	0.04	0.04	0.04	0.04	0.04	
AUG	S7	H2	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.06	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
AUG	S7	H9	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
AUG	S8	H1	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
AUG	S8	H4	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	

หมายเหตุ L1, L2, L3, ..., L23 แทนช่วงข้อมูลที่อยู่ใน Attribute Day light Hour ซึ่งปรากฏในตารางที่ 10

S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

H1, H2, H3, ..., H10 แทนช่วงข้อมูลที่อยู่ใน Attribute Height ปรากฏอยู่ในตารางที่ 10

จากตารางที่ 63 เป็นตารางที่แสดงตัวอย่างความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Daylight Hour ที่มี Attribute Result, Attribute Month Attribute Landscape และ Attribute Height เป็นเงื่อนไข โดยที่ค่าความน่าจะเป็นที่มากที่สุดในการกรณีที่ Result = Positive คือ

$$\Pr(\text{Day Light Hour} = L19 | \text{Result} = \text{Positive}, \text{Month} = \text{NOV}, \text{Landscape} = S8, \text{Height} = H3) = 0.84$$

จากสมการความน่าจะเป็น สามารถอธิบายได้ว่า ความน่าจะเป็นที่ Attribute Daylight hour มีค่าเท่ากับ L19 ซึ่งหมายถึงจำนวนชั่วโมงที่มีแสงแดดต่อวันอยู่ในช่วง 8.022 – 8.087 ชั่วโมง โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Positive และ Attribute Month มีค่าเท่ากับ NOV และ Attribute Landscape มีค่าเท่ากับ S8 ซึ่งหมายถึงมีลักษณะภูมิประเทศเป็นแบบ ขอบที่ราบภาคกลาง และ Attribute Height มีค่าเท่ากับ H3 ซึ่งหมายถึงมีระดับความสูงจากระดับน้ำทะเลอยู่ในช่วงมากกว่า 6.5 ถึงน้อยกว่า 12 เมตร มีค่าความน่าจะเป็นเท่ากับ 0.84

ดังนั้น โอกาสที่พบสัตว์ปีกป่วยตายด้วยโรคไข้หวัดนกสูงสุด(0.84) ในเงื่อนไข ช่วงเดือน พฤศจิกายน และมีลักษณะภูมิประเทศแบบ ขอบที่ราบภาคกลาง และมีระดับความสูงของพื้นที่อยู่ในช่วงมากกว่า 6.5 ถึง 12 เมตร

จากตารางที่ 64 เป็นตารางที่แสดงตัวอย่างความน่าจะเป็นแบบมีเงื่อนไขของ Attribute Daylight Hour ที่มี Attribute Result, Attribute Month, Attribute Landscape และ Attribute Height เป็นเงื่อนไข โดยที่ค่าความน่าจะเป็นแบบมีเงื่อนไขที่มากที่สุดเมื่อ Result = Negative คือ

$$\Pr(\text{Day Light Hour} = L3 | \text{Result} = \text{Negative}, \text{Month} = \text{AUG}, \text{Landscape} = S7, \text{Height} = H6) = 0.85$$

จากสมการความน่าจะเป็น สามารถอธิบายได้ว่า ความน่าจะเป็นที่ Attribute Daylight hour มีค่าเท่ากับ L3 ซึ่งหมายถึงจำนวนชั่วโมงที่มีแสงแดดต่อวันอยู่ในช่วง 3.049 – 3.099 ชั่วโมง โดยมีเงื่อนไขคือ Attribute Result มีค่าเท่ากับ Negative และ Attribute Month มีค่าเท่ากับ AUG และ Attribute Landscape มีค่าเท่ากับ S7 ซึ่งหมายถึงมีลักษณะภูมิประเทศเป็นแบบ ที่ราบภาคกลาง ตอนบน และ Attribute Height มีค่าเท่ากับ H6 มีระดับความสูงจากระดับน้ำทะเลอยู่ในช่วงมากกว่า 34.5 ถึง 97 เมตร มีค่าความน่าจะเป็นเท่ากับ 0.85

ดังนั้นโอกาสที่พบสัตว์ปีกป่วยตายด้วยโรคอื่นที่ไม่ใช่โรคไข้หวัดนกสูงสุด (0.85) ใน  
เดือนไซ จำนวนชั่วโมงที่มีแสงแดดต่อวันอยู่ในช่วง 3.049 – 3.099 ชั่วโมง และ ลักษณะภูมิประเทศ  
เป็นแบบที่ราบลุ่มแม่น้ำภาคกลางตอนบน และ มีความสูงของพื้นที่ ในช่วงมากกว่า 34.5 ถึง 97  
เมตร

จากตารางความน่าจะเป็นที่ได้กล่าวไปข้างต้น ยังไม่สามารถนำไปใช้งานได้ต้องผ่านการ  
นำความน่าจะเป็นของแต่ละ Attribute ที่ได้ไปหาความน่าจะเป็นรวมของทุก Attribute ตามสูตรการ  
คำนวณหาความน่าจะเป็นตามทฤษฎีของเบย์ส์ที่ว่า

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n (P(X_i | \text{Parent}(X_i)))$$

โดยที่

$X_1, X_2, \dots, X_n$

แทน Attribute ที่อยู่ชุดข้อมูลที่นำมาสร้างแบบจำลอง

Parent ( $X_i$ ) โดยที่  $i = 1, 2, 3, \dots, n$

แทน Node ที่เป็น Attribute ที่อยู่ใน Parent Node ของ Attribute  
ที่กำลังหาค่าความน่าจะเป็นแบบมีเงื่อนไข

หลังจากหาความน่าจะเป็นของแต่ละ Record ที่อยู่ในชุดข้อมูลจะทำให้ทราบได้ว่า Record  
ใดในชุดข้อมูลมีความน่าจะเป็นสูงที่สุด แล้วจึงนำแต่ละ Record ที่มีค่าความน่าจะเป็นสูงที่สุดไปใช้  
งานต่อไป

จากชุดข้อมูลที่เรานำมาใช้ในการสร้างแบบจำลองเพื่อเฝ้าระวังโรคไข้หวัดนก เมื่อทำการ  
หาความน่าจะเป็นเรียบร้อยแล้ว และเรียงลำดับข้อมูลที่มีความน่าจะเป็นสูงที่สุด ทำให้เราทราบว่า  
ลักษณะพื้นที่ และ สภาพภูมิอากาศแบบใดที่มีโอกาสที่จะเกิดการระบาดของโรคไข้หวัดนก ซึ่ง  
ตัวอย่างลักษณะพื้นที่ที่อยู่ในตารางที่ 65 ถึง ตารางที่ 71

ตารางที่ 65 ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วงระยะเวลา  
เดือน มกราคม

Month	Height	Landscape	Death	Destruction	Temperature	Humidity		
						ty	DLH	P*100
JAN	H6	S7	D3	De1	T2	Hu2	L11	50.958
JAN	H8	S10	D1	De4	T2	Hu2	L17	51.576
JAN	H8	S10	D4	De5	T2	Hu2	L21	53.850
JAN	H6	S7	D3	De1	T2	Hu2	L23	56.552
JAN	H1	S6	D2	De3	T3	Hu3	L1	57.310
JAN	H6	S7	D2	De2	T2	Hu1	L23	58.069
JAN	H6	S5	D3	De3	T2	Hu3	L15	59.722
JAN	H3	S6	D2	De2	T3	Hu2	L8	61.860
JAN	H8	S10	D3	De2	T2	Hu2	L17	61.914
JAN	H6	S7	D4	De1	T2	Hu2	L23	64.457
JAN	H9	S8	D2	De2	T3	Hu2	L17	65.976
JAN	H3	S8	D3	De3	T2	Hu3	L23	66.099
JAN	H6	S7	D3	De3	T2	Hu1	L23	68.546
JAN	H6	S7	D3	De2	T2	Hu2	L13	72.264
JAN	H3	S6	D1	De2	T3	Hu2	L8	75.310
JAN	H8	S10	D4	De3	T3	Hu2	L17	75.498
JAN	H6	S7	D1	De2	T2	Hu2	L23	82.281
JAN	H3	S8	D4	De4	T2	Hu3	L20	84.954
JAN	H6	S7	D2	De2	T2	Hu2	L23	87.387
JAN	H6	S7	D1	De3	T2	Hu2	L23	88.888
JAN	H6	S7	D2	De2	T2	Hu2	L15	90.328
JAN	H6	S7	D2	De3	T2	Hu2	L23	91.371
JAN	H8	S10	D3	De3	T3	Hu2	L17	91.474
JAN	H6	S7	D3	De3	T2	Hu2	L23	93.802
JAN	H3	S6	D3	De5	T2	Hu1	L21	93.984

ตารางที่ 65 (ต่อ)

Month	Heig	Lands	Death	Destruction	Temperature	Humidity	DLH	P*100
	ht	cape						
JAN	H6	S7	D2	De4	T2	Hu2	L20	95.763
JAN	H6	S7	D2	De2	T2	Hu2	L17	96.238
JAN	H3	S8	D4	De5	T3	Hu3	L11	96.283
JAN	H6	S7	D2	De4	T2	Hu2	L23	96.563
JAN	H6	S7	D3	De3	T2	Hu2	L21	97.539
JAN	H6	S7	D4	De3	T2	Hu2	L17	98.163
JAN	H6	S7	D3	De3	T2	Hu2	L17	98.242
JAN	H6	S7	D4	De5	T2	Hu2	L23	98.289
JAN	H6	S7	D4	De2	T2	Hu2	L17	98.706
JAN	H3	S6	D3	De5	T3	Hu2	L23	99.136
<b>JAN</b>	<b>H6</b>	<b>S7</b>	<b>D1</b>	<b>De5</b>	<b>T2</b>	<b>Hu2</b>	<b>L15</b>	<b>99.866</b>

หมายเหตุ DLH = Day Light Hour , P = Probability

จากตารางที่ 65 ลักษณะพื้นที่ เลี้ยงสัตว์ปีกที่มีโอกาสเกิดโรคไข้หวัดนกมากที่สุดคือ  
ลักษณะพื้นที่มีองค์ประกอบดังต่อไปนี้  
ในช่วงเดือนมกราคม

1. อยู่ในระดับความสูงจากระดับน้ำทะเล ระหว่าง 34.5 เมตร ถึง 97.5 เมตร
2. อุณหภูมิอยู่ในช่วง มากกว่า 21.15 องศาเซลเซียส แต่น้อยกว่า 27.85 องศาเซลเซียส
3. อยู่ในลักษณะพื้นที่ที่เรียกว่าพื้นที่ราบลุ่มน้ำภาคกลางตอนบน
4. จำนวนสัตว์ปีกที่ตายอยู่ในช่วง น้อยกว่า 5 ตัว
5. จำนวนสัตว์ปีกที่ถูกทำลาย มากกว่า 969 ตัว
6. ความชื้นสัมพัทธ์อยู่ในช่วง 60.5 -75.09 %
7. จำนวนชั่วโมงที่มีแสงแดด อยู่ในช่วง 7.18 – 7.66 ชั่วโมง/วัน

ตารางที่ 66 ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วงระยะเวลา  
เดือน กรกฎาคม

Month	Height	Landscape	Death	Destruction	Temperat		DLH	P*100
					ure	Humidity		
JUL	H3	S8	D1	De1	T3	Hu2	L11	50.113
JUL	H9	S8	D4	De3	T3	Hu3	L12	50.816
JUL	H2	S1	D4	De5	T2	Hu3	L13	52.896
JUL	H6	S7	D1	De2	T2	Hu2	L10	54.766
JUL	H6	S7	D1	De2	T3	Hu2	L10	55.362
JUL	H6	S7	D2	De3	T3	Hu2	L8	59.592
JUL	H8	S10	D3	De5	T3	Hu3	L8	60.486
JUL	H4	S4	D3	De2	T3	Hu2	L11	61.779
JUL	H6	S7	D3	De3	T3	Hu2	L8	63.304
JUL	H6	S5	D4	De5	T3	Hu3	L3	63.480
JUL	H3	S8	D4	De1	T2	Hu3	L11	67.974
JUL	H1	S6	D2	De2	T3	Hu3	L1	68.708
JUL	H1	S6	D4	De1	T3	Hu2	L11	75.725
JUL	H3	S8	D2	De2	T3	Hu3	L11	75.831
JUL	H6	S7	D3	De2	T3	Hu2	L23	77.877
JUL	H1	S6	D2	De3	T3	Hu3	L13	79.109
JUL	H3	S8	D2	De3	T2	Hu3	L11	82.164
JUL	H5	S7	D4	De5	T3	Hu2	L5	85.848
JUL	H3	S8	D1	De3	T3	Hu3	L11	85.922
JUL	H5	S7	D2	De2	T3	Hu2	L21	86.899
JUL	H8	S10	D4	De5	T2	Hu3	L1	87.090
JUL	H1	S6	D3	De1	T3	Hu2	L6	87.386
JUL	H3	S6	D4	De4	T3	Hu3	L21	88.117

## ตารางที่ 66 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
JUL	H6	S5	D4	De5	T2	Hu3	L3	88.369
JUL	H3	S6	D4	De4	T3	Hu3	L1	88.630
JUL	H3	S6	D3	De1	T3	Hu2	L23	88.754
JUL	H6	S7	D3	De3	T3	Hu2	L23	89.408
JUL	H6	S7	D4	De2	T3	Hu3	L3	90.341
JUL	H1	S6	D2	De2	T3	Hu2	L13	91.060
JUL	H6	S7	D4	De5	T2	Hu3	L1	91.616
JUL	H3	S8	D3	De1	T3	Hu2	L11	93.130
JUL	H6	S5	D3	De4	T3	Hu2	L3	93.179
JUL	H9	S8	D2	De3	T3	Hu2	L11	93.884
JUL	H9	S8	D3	De3	T3	Hu3	L12	94.114
JUL	H1	S6	D4	De4	T3	Hu1	L12	94.222
JUL	H3	S8	D3	De3	T3	Hu3	L11	94.882
JUL	H3	S8	D1	De2	T3	Hu2	L11	94.984
JUL	H6	S5	D4	De5	T3	Hu2	L3	95.894
JUL	H1	S6	D2	De3	T3	Hu3	L12	96.154
JUL	H3	S6	D1	De2	T2	Hu3	L1	96.548
JUL	H3	S8	D4	De4	T3	Hu2	L17	96.767
JUL	H3	S8	D4	De1	T3	Hu2	L11	97.160
JUL	H3	S6	D3	De2	T2	Hu3	L1	97.220
JUL	H1	S6	D1	De2	T3	Hu2	L13	97.237
JUL	H1	S6	D1	De2	T3	Hu2	L15	97.441
JUL	H3	S8	D3	De3	T2	Hu3	L11	97.465

## ตารางที่ 66 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature		DLH	P*100
JUL	H3	S8	D2	De3	T3	Hu2	L11	97.825
JUL	H1	S6	D3	De2	T3	Hu2	L23	98.076
JUL	H1	S6	D2	De3	T3	Hu2	L13	98.140
JUL	H1	S6	D4	De2	T3	Hu2	L13	98.463
JUL	H3	S6	D3	De2	T3	Hu2	L12	98.464
JUL	H6	S7	D3	De4	T3	Hu2	L3	98.574
JUL	H3	S6	D4	De5	T3	Hu2	L8	98.594
JUL	H3	S8	D1	De3	T3	Hu2	L11	98.595
JUL	H3	S8	D3	De2	T3	Hu2	L11	98.686
JUL	H1	S6	D3	De3	T3	Hu2	L17	98.771
JUL	H1	S6	D4	De2	T3	Hu2	L23	98.772
JUL	H6	S7	D3	De3	T3	Hu2	L3	98.907
JUL	H1	S6	D4	De3	T3	Hu2	L15	98.969
JUL	H3	S6	D1	De2	T3	Hu2	L1	99.005
JUL	H1	S6	D2	De3	T3	Hu2	L15	99.144
JUL	H1	S6	D2	De2	T3	Hu2	L15	99.213
JUL	H3	S6	D4	De5	T3	Hu2	L13	99.229
JUL	H1	S6	D4	De4	T3	Hu2	L12	99.274
JUL	H3	S6	D4	De5	T3	Hu2	L17	99.283
JUL	H3	S6	D3	De2	T3	Hu2	L1	99.338
JUL	H3	S8	D4	De2	T3	Hu2	L12	99.390
JUL	H1	S6	D4	De2	T3	Hu2	L12	99.482
JUL	H3	S8	D4	De5	T3	Hu3	L11	99.502

## ตารางที่ 66 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
JUL	H3	S8	D3	De3	T3	Hu2	L11	99.595
JUL	H3	S6	D4	De5	T3	Hu2	L21	99.596
JUL	H1	S6	D3	De4	T3	Hu2	L12	99.629
JUL	H9	S8	D1	De5	T3	Hu2	L12	99.656
JUL	H3	S6	D1	De3	T3	Hu2	L1	99.688
JUL	H3	S8	D3	De2	T3	Hu2	L12	99.786
JUL	H1	S6	D3	De4	T3	Hu2	L15	99.844
JUL	H3	S6	D3	De3	T3	Hu2	L1	99.855
JUL	H3	S8	D4	De4	T3	Hu2	L12	99.857
JUL	H1	S6	D4	De5	T3	Hu2	L12	99.883
JUL	H3	S8	D4	De5	T3	Hu2	L11	99.904
JUL	H3	S8	D3	De3	T3	Hu2	L12	99.906
<b>JUL</b>	<b>H3</b>	<b>S8</b>	<b>D4</b>	<b>De5</b>	<b>T3</b>	<b>Hu2</b>	<b>L12</b>	<b>99.984</b>

หมายเหตุ DLH = Day Light Hour P = Probability

จากตารางที่ 66 ลักษณะพื้นที่ เลี้ยงสัตว์ปีกที่มีโอกาสเกิดโรคไข้หวัดนกมากที่สุดคือ  
ลักษณะพื้นที่มีองค์ประกอบดังต่อไปนี้

1. ในช่วงเดือน กรกฎาคม
2. มีความสูงจากระดับน้ำทะเล ระหว่าง 6.5 เมตร ถึง 12 เมตร
3. อุณหภูมิอยู่ในช่วง 27.85 ถึง 32.05 องศาเซลเซียส
4. อยู่ในพื้นที่ที่มีลักษณะอยู่ในบริเวณขอบที่ราบภาคกลาง
5. มีจำนวนสัตว์ปีกที่ตายอยู่ในช่วงมากกว่า 98 ตัว
6. มีจำนวนสัตว์ปีกที่ถูกทำลายอยู่ในช่วงมากกว่า 969
7. มีความชื้นสัมพัทธ์ ในช่วง 60.5 -75.09 %
8. มีจำนวนชั่วโมงที่มีแสงแดด อยู่ในช่วง 5.14 – 5.19 ชั่วโมง/วัน

ตารางที่ 67 ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วงระยะเวลา  
เดือน สิงหาคม

Landsca								
Month	Height	pe	Death	Destruction	Temperature	Humidity	DLH	P*100
AUG	H9	S8	D2	De3	T3	Hu2	L8	51.354
AUG	H6	S7	D3	De3	T3	Hu3	L1	52.187
AUG	H1	S3	D2	De3	T3	Hu3	L11	54.045
AUG	H6	S7	D1	De2	T3	Hu2	L1	54.494
AUG	H6	S7	D4	De5	T3	Hu3	L5	57.972
AUG	H8	S10	D4	De5	T2	Hu3	L8	59.151
AUG	H6	S5	D1	De4	T3	Hu3	L5	60.366
AUG	H6	S7	D3	De3	T3	Hu2	L6	60.680
AUG	H9	S8	D3	De2	T3	Hu2	L8	63.472
AUG	H6	S7	D3	De2	T3	Hu2	L1	67.796
AUG	H6	S7	D1	De3	T3	Hu2	L1	68.492
AUG	H6	S7	D1	De4	T3	Hu3	L5	69.382
AUG	H5	S7	D4	De4	T3	Hu3	L13	69.390
AUG	H6	S7	D2	De3	T3	Hu2	L5	71.025
AUG	H1	S6	D3	De2	T3	Hu2	L1	71.567
AUG	H3	S6	D4	De3	T3	Hu2	L1	71.593
AUG	H3	S8	D4	De3	T3	Hu2	L4	72.069
AUG	H9	S8	D3	De3	T3	Hu2	L4	72.503
AUG	H6	S7	D2	De2	T3	Hu2	L1	72.809
AUG	H3	S6	D3	De4	T3	Hu2	L2	73.007
AUG	H3	S6	D2	De2	T3	Hu2	L1	77.079
AUG	H3	S8	D2	De2	T3	Hu2	L8	77.953
AUG	H1	S6	D1	De2	T3	Hu2	L1	78.057

## ตารางที่ 67 (ต่อ)

Month	Height	Landscape	Destruction		Temperature	Humidity	DLH	P*100
			Death	ction				
AUG	H6	S7	D2	De3	T3	Hu2	L1	81.192
AUG	H3	S8	D2	De3	T3	Hu3	L11	81.337
AUG	H4	S6	D1	De5	T3	Hu2	L1	82.955
AUG	H1	S3	D1	De3	T3	Hu2	L2	83.052
AUG	H6	S7	D2	De3	T3	Hu2	L23	84.527
AUG	H5	S7	D3	De2	T3	Hu2	L1	85.709
AUG	H3	S8	D3	De2	T3	Hu2	L8	86.873
AUG	H3	S6	D4	De3	T3	Hu2	L5	87.341
AUG	H6	S7	D1	De4	T3	Hu3	L1	87.431
AUG	H6	S7	D3	De3	T3	Hu2	L1	87.630
AUG	H5	S7	D1	De2	T3	Hu2	L1	88.302
AUG	H5	S7	D3	De2	T3	Hu2	L4	88.313
AUG	H6	S7	D3	De3	T3	Hu2	L15	88.440
AUG	H9	S8	D4	De5	T2	Hu3	L11	88.456
AUG	H5	S7	D4	De4	T3	Hu2	L23	88.862
AUG	H3	S8	D4	De5	T3	Hu3	L8	89.426
AUG	H4	S8	D1	De3	T2	Hu3	L11	89.604
AUG	H1	S6	D4	De4	T3	Hu2	L11	90.407
AUG	H6	S7	D1	De4	T3	Hu2	L6	91.380
AUG	H1	S6	D2	De4	T2	Hu3	L11	92.351
AUG	H5	S7	D1	De2	T3	Hu2	L4	92.867
AUG	H1	S6	D4	De5	T2	Hu3	L11	92.962
AUG	H1	S6	D2	De2	T3	Hu2	L11	93.093

ตารางที่ 67 (ต่อ)

Month	Height	Landscape	Death	Destructio				
				n	Temperature	Humidity	DLH	P*100
AUG	H1	S6	D4	De5	T3	Hu2	L13	96.579
AUG	H3	S8	D4	De2	T3	Hu3	L11	96.882
AUG	H3	S8	D3	De4	T3	Hu2	L13	96.888
AUG	H1	S6	D3	De3	T3	Hu3	L11	97.250
AUG	H3	S6	D4	De5	T3	Hu2	L1	97.527
AUG	H1	S6	D4	De5	T3	Hu2	L1	97.536
AUG	H5	S7	D4	De5	T3	Hu3	L8	97.700
AUG	H9	S8	D4	De5	T3	Hu3	L11	98.124
AUG	H5	S7	D3	De2	T3	Hu2	L23	98.177
AUG	H3	S6	D3	De2	T3	Hu2	L13	98.223
AUG	H3	S6	D3	De3	T3	Hu2	L13	98.333
AUG	H5	S7	D3	De3	T3	Hu2	L20	98.803
AUG	H4	S8	D4	De5	T3	Hu2	L11	99.058
AUG	H3	S8	D4	De4	T3	Hu2	L11	99.566
AUG	H3	S6	D4	De5	T3	Hu2	L13	99.693
<b>AUG</b>	<b>H3</b>	<b>S8</b>	<b>D4</b>	<b>De5</b>	<b>T3</b>	<b>Hu2</b>	<b>L11</b>	<b>99.886</b>

หมายเหตุ DLH = Day Light Hour, P = Probability

จากตารางที่ 67 ลักษณะพื้นที่ เลี้ยงสัตว์ปีกที่มีโอกาสเกิดโรคไข้หวัดนกมากที่สุดคือ  
ลักษณะพื้นที่มีองค์ประกอบดังต่อไปนี้

1. ในช่วงเดือน สิงหาคม
2. มีระดับความสูงจากระดับน้ำทะเลอยู่ระหว่าง 6.5 ถึง 12 เมตร
3. อุณหภูมิ อยู่ในช่วง 27.85 ถึง 32.05 องศาเซลเซียส
4. ลักษณะพื้นที่เป็นขอบที่ราบภาคกลาง
5. จำนวนสัตว์ปีกป่วยตายอยู่ในช่วง 98-969 ตัว

6. จำนวนสัตว์ปีกที่ถูกทำลาย อยู่ในช่วงมากกว่า 969 ตัว
7. ความชื้นสัมพัทธ์ อยู่ในช่วง 60.5 – 75.09 %
8. จำนวนชั่วโมงที่มีแสงแดดต่อวัน อยู่ในช่วง 4 – 5 ชั่วโมง

**ตารางที่ 68** ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไขหวัดนกในช่วงระยะเวลา  
เดือน กันยายน

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
SEP	H6	S3	D3	De5	T3	Hu2	L15	51.334
SEP	H6	S7	D2	De3	T2	Hu3	L1	52.154
SEP	H7	S7	D3	De3	T3	Hu3	L13	52.504
SEP	H6	S7	D3	De2	T3	Hu2	L8	55.436
SEP	H3	S8	D1	De2	T3	Hu2	L13	59.332
SEP	H6	S7	D3	De3	T2	Hu3	L1	61.60
SEP	H6	S3	D3	De3	T2	Hu3	L11	64.271
SEP	H6	S7	D4	De5	T3	Hu3	L11	65.766
SEP	H4	S8	D1	De4	T4	Hu2	L13	65.885
SEP	H1	S6	D2	De2	T3	Hu3	L9	69.801
SEP	H6	S7	D1	De2	T3	Hu3	L8	73.842
SEP	H1	S6	D4	De4	T2	Hu3	L1	75.602
SEP	H4	S4	D2	De2	T2	Hu3	L9	76.662
SEP	H6	S7	D3	De2	T3	Hu3	L1	78.985
SEP	H1	S6	D2	De2	T3	Hu2	L15	80.512
SEP	H6	S7	D2	De3	T3	Hu3	L8	82.303
SEP	H1	S6	D3	De2	T3	Hu3	L1	82.573
SEP	H6	S7	D2	De2	T3	Hu3	L1	84.131
SEP	H6	S7	D2	De2	T3	Hu2	L23	85.171
SEP	H1	S3	D3	De4	T3	Hu3	L15	86.006
SEP	H6	S7	D3	De3	T3	Hu3	L2	86.483
SEP	H6	S3	D2	De4	T2	Hu2	L1	88.649

## ตารางที่ 68 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
SEP	H3	S8	D2	De2	T3	Hu2	L13	89.786
SEP	H3	S8	D2	De4	T3	Hu2	L13	91.420
SEP	H6	S7	D1	De2	T3	Hu3	L7	91.922
SEP	H1	S6	D4	De3	T3	Hu3	L1	92.497
SEP	H6	S5	D1	De5	T3	Hu3	L8	93.043
SEP	H6	S7	D1	De4	T3	Hu3	L13	93.253
SEP	H7	S7	D3	De3	T3	Hu3	L21	94.317
SEP	H3	S8	D1	De3	T3	Hu2	L13	94.933
SEP	H6	S7	D3	De2	T3	Hu3	L7	95.095
SEP	H1	S6	D4	De2	T3	Hu3	L9	95.244
SEP	H6	S7	D4	De5	T3	Hu3	L20	96.380
SEP	H1	S6	D1	De2	T3	Hu2	L20	96.658
SEP	H1	S6	D3	De3	T3	Hu3	L9	96.775
SEP	H6	S7	D1	De3	T3	Hu3	L7	96.894
SEP	H1	S6	D2	De3	T3	Hu2	L20	96.977
SEP	H6	S7	D2	De3	T3	Hu3	L7	97.087
SEP	H1	S6	D3	De2	T3	Hu2	L20	97.458
SEP	H5	S7	D4	De5	T3	Hu3	L8	97.882
SEP	H9	S8	D2	De2	T3	Hu3	L8	97.991
SEP	H4	S6	D4	De3	T2	Hu2	L1	97.995
SEP	H3	S6	D2	De3	T3	Hu3	L9	98.018
SEP	H5	S7	D2	De3	T3	Hu3	L23	98.085
SEP	H9	S8	D2	De3	T3	Hu3	L8	98.353

## ตารางที่ 68 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
SEP	H5	S7	D4	De5	T3	Hu3	L23	99.135
SEP	H9	S8	D3	De3	T3	Hu3	L8	99.233
SEP	H1	S6	D4	De5	T3	Hu2	L9	99.322
SEP	H6	S7	D4	De5	T3	Hu3	L7	99.322
SEP	H4	S8	D4	De5	T3	Hu3	L9	99.326
SEP	H4	S8	D4	De4	T2	Hu2	L9	99.408
SEP	H3	S8	D2	De3	T3	Hu3	L9	99.478
SEP	H3	S6	D4	De5	T3	Hu3	L13	99.544
SEP	H9	S8	D4	De5	T3	Hu3	L9	99.571
SEP	H3	S8	D2	De3	T3	Hu2	L9	99.727
SEP	H3	S8	D3	De2	T3	Hu2	L9	99.836
SEP	H3	S8	D4	De3	T3	Hu2	L9	99.932
SEP	H3	S8	D4	De5	T3	Hu3	L9	99.937
SEP	H4	S8	D4	De5	T2	Hu2	L9	99.952
<b>SEP</b>	<b>H3</b>	<b>S8</b>	<b>D4</b>	<b>De5</b>	<b>T3</b>	<b>Hu2</b>	<b>L9</b>	<b>99.987</b>

หมายเหตุ DLH = Day Light Hour , P = Probability

จากตารางที่ 68 ลักษณะพื้นที่ เลี้ยงสัตว์ปีกที่มีโอกาสเกิดโรคไข้หวัดนกมากที่สุดคือ  
ลักษณะพื้นที่มีองค์ประกอบดังต่อไปนี้

1. ในช่วงเดือนกันยายน
2. มีระดับความสูงจากระดับน้ำทะเลอยู่ระหว่าง 6.5 ถึง 12 เมตร
3. อุณหภูมิ อยู่ระหว่าง 27.85 ถึง 32.05 องศาเซลเซียส
4. ลักษณะพื้นที่เป็นขอบที่ราบสูง
5. จำนวนสัตว์ปีกป่วยตายอยู่ระหว่าง 98 ถึง 969 ตัว
6. จำนวนสัตว์ปีกที่ถูกทำลาย อยู่ในช่วงมากกว่า 969 ตัว

7. ความชื้นสัมพัทธ์ อยู่ในช่วง 60.5 – 75.09 %
8. จำนวนชั่วโมงที่มีแสงแดด อยู่ในช่วง 4.3 – 4.33 ชั่วโมง/วัน

ตารางที่ 69 ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วงระยะเวลา  
เดือน ตุลาคม

Month	Lands			Death	Destruction	Temperature	Humidity	DLH	P*100
	Height	cape							
OCT	H8	S9	D3	De4	T2	Hu2	L13	50.507	
OCT	H6	S7	D1	De3	T3	Hu3	L11	50.837	
OCT	H6	S3	D1	De3	T2	Hu3	L15	51.318	
OCT	H8	S9	D1	De2	T3	Hu2	L13	51.743	
OCT	H6	S7	D3	De2	T3	Hu3	L17	52.236	
OCT	H3	S2	D2	De3	T2	Hu3	L11	54.085	
OCT	H8	S9	D2	De2	T3	Hu2	L21	54.355	
OCT	H8	S9	D3	De3	T2	Hu2	L21	56.370	
OCT	H8	S9	D3	De2	T3	Hu2	L21	56.388	
OCT	H2	S4	D2	De2	T3	Hu2	L15	56.929	
OCT	H8	S9	D1	De3	T3	Hu2	L21	57.832	
OCT	H4	S6	D2	De3	T3	Hu3	L13	57.887	
OCT	H3	S2	D3	De3	T2	Hu3	L11	60.968	
OCT	H6	S7	D4	De4	T3	Hu3	L15	62.470	
OCT	H2	S2	D2	De2	T3	Hu3	L11	65.154	
OCT	H1	S3	D1	De2	T3	Hu3	L15	67.277	
OCT	H1	S3	D2	De2	T3	Hu3	L17	67.771	
OCT	H3	S8	D4	De5	T2	Hu3	L1	68.210	
OCT	H3	S6	D4	De4	T3	Hu3	L10	68.520	
OCT	H4	S4	D1	De3	T2	Hu2	L18	68.601	
OCT	H6	S7	D2	De2	T3	Hu3	L15	68.837	
OCT	H1	S3	D2	De2	T3	Hu3	L15	69.395	

## ตารางที่ 69 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
OCT	H1	S6	D1	De2	T3	Hu1	L23	71.036
OCT	H5	S7	D2	De1	T3	Hu2	L23	71.627
OCT	H3	S2	D3	De4	T2	Hu3	L11	72.453
OCT	H6	S7	D4	De4	T3	Hu3	L14	77.356
OCT	H6	S7	D1	De3	T3	Hu3	L13	77.686
OCT	H1	S3	D2	De3	T3	Hu3	L15	77.764
OCT	H1	S3	D4	De2	T3	Hu3	L15	79.529
OCT	H4	S4	D2	De3	T3	Hu2	L18	79.812
OCT	H1	S3	D2	De3	T3	Hu3	L4	80.966
OCT	H3	S8	D4	De3	T2	Hu3	L13	81.430
OCT	H8	S9	D3	De3	T2	Hu3	L21	81.474
OCT	H6	S7	D1	De3	T3	Hu3	L15	81.572
OCT	H4	S6	D4	De4	T3	Hu2	L23	82.524
OCT	H4	S6	D4	De4	T3	Hu3	L13	83.006
OCT	H7	S7	D1	De3	T2	Hu2	L17	83.067
OCT	H2	S4	D1	De3	T2	Hu2	L13	83.355
OCT	H3	S8	D2	De2	T2	Hu2	L23	84.143
OCT	H2	S4	D1	De3	T2	Hu3	L13	84.562
OCT	H4	S4	D1	De3	T3	Hu2	L18	84.711
OCT	H6	S2	D1	De2	T2	Hu3	L1	85.628
OCT	H4	S4	D3	De4	T3	Hu3	L18	85.742
OCT	H1	S3	D3	De2	T3	Hu2	L13	85.878
OCT	H9	S8	D2	De2	T3	Hu3	L18	88.331

## ตารางที่ 69 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
OCT	H5	S7	D3	De2	T3	Hu3	L13	89.300
OCT	H9	S8	D2	De3	T3	Hu3	L18	89.891
OCT	H6	S7	D2	De2	T3	Hu3	L14	89.921
OCT	H2	S3	D4	De4	T3	Hu2	L15	90.350
OCT	H6	S7	D3	De2	T3	Hu3	L14	90.516
OCT	H3	S8	D1	De2	T3	Hu1	L18	90.785
OCT	H3	S8	D2	De3	T3	Hu1	L18	91.295
OCT	H3	S8	D1	De3	T3	Hu3	L15	91.558
OCT	H9	S8	D4	De3	T3	Hu3	L18	91.612
OCT	H3	S8	D3	De2	T2	Hu2	L23	91.759
OCT	H6	S3	D2	De4	T2	Hu3	L8	91.875
OCT	H4	S4	D3	De2	T3	Hu2	L18	91.976
OCT	H6	S5	D2	De3	T2	Hu2	L14	92.062
OCT	H7	S7	D1	De2	T2	Hu2	L23	92.120
OCT	H7	S7	D2	De2	T2	Hu2	L21	92.239
OCT	H7	S7	D2	De3	T2	Hu2	L13	92.584
OCT	H3	S8	D4	De3	T3	Hu3	L13	92.704
OCT	H7	S7	D2	De3	T2	Hu2	L23	92.918
OCT	H5	S7	D4	De3	T3	Hu3	L21	93.400
OCT	H5	S7	D1	De2	T3	Hu3	L13	93.502
OCT	H6	S2	D3	De4	T2	Hu3	L1	93.613
OCT	H5	S7	D3	De3	T3	Hu3	L13	93.708
OCT	H4	S6	D4	De2	T3	Hu3	L13	93.723

## ตารางที่ 69 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
OCT	H6	S7	D1	De3	T3	Hu3	L23	93.801
OCT	H9	S8	D2	De3	T2	Hu2	L18	94.038
OCT	H3	S8	D3	De2	T3	Hu3	L13	94.156
OCT	H6	S7	D3	De3	T3	Hu3	L23	94.626
OCT	H1	S6	D1	De2	T3	Hu2	L17	94.645
OCT	H6	S7	D1	De3	T3	Hu3	L14	94.678
OCT	H7	S7	D1	De2	T3	Hu2	L23	94.690
OCT	H7	S7	D2	De2	T3	Hu2	L21	94.714
OCT	H6	S7	D3	De4	T3	Hu3	L23	94.742
OCT	H3	S6	D2	De2	T3	Hu2	L20	94.954
OCT	H1	S6	D2	De2	T3	Hu1	L23	95.062
OCT	H3	S2	D2	De3	T3	Hu2	L11	95.076
OCT	H7	S7	D4	De3	T2	Hu2	L21	95.083
OCT	H6	S3	D3	De3	T2	Hu2	L2	95.158
OCT	H6	S7	D3	De3	T3	Hu3	L14	95.270
OCT	H1	S3	D1	De2	T3	Hu2	L23	95.552
OCT	H7	S7	D2	De2	T2	Hu2	L23	95.613
OCT	H7	S7	D3	De2	T2	Hu2	L23	95.752
OCT	H5	S7	D4	De4	T3	Hu3	L23	95.803
OCT	H6	S7	D2	De2	T2	Hu2	L20	96.006
OCT	H2	S6	D2	De3	T2	Hu3	L1	96.015
OCT	H7	S4	D3	De2	T3	Hu3	L14	96.037
OCT	H1	S3	D3	De2	T3	Hu2	L23	96.078

## ตารางที่ 69 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
OCT	H5	S7	D2	De3	T3	Hu3	L13	96.522
OCT	H6	S7	D4	De5	T3	Hu3	L13	96.557
OCT	H3	S8	D3	De5	T2	Hu3	L13	96.568
OCT	H7	S7	D3	De2	T3	Hu2	L23	96.583
OCT	H9	S8	D1	De3	T3	Hu3	L18	96.631
OCT	H9	S8	D3	De2	T3	Hu3	L18	96.643
OCT	H3	S6	D2	De3	T3	Hu2	L20	96.811
OCT	H2	S3	D3	De3	T3	Hu2	L15	96.853
OCT	H7	S7	D2	De3	T2	Hu2	L21	96.889
OCT	H6	S5	D3	De3	T2	Hu2	L14	96.901
OCT	H6	S3	D4	De5	T3	Hu2	L23	96.918
OCT	H6	S7	D1	De2	T2	Hu2	L23	96.974
OCT	H5	S7	D3	De3	T3	Hu3	L21	96.978
OCT	H7	S7	D1	De3	T2	Hu2	L13	97.042
OCT	H2	S4	D1	De2	T3	Hu2	L15	97.171
OCT	H6	S7	D2	De2	T3	Hu2	L13	97.191
OCT	H5	S7	D4	De2	T3	Hu3	L23	97.207
OCT	H5	S7	D3	De2	T3	Hu2	L21	97.236
OCT	H7	S7	D3	De2	T3	Hu2	L13	97.280
OCT	H6	S7	D3	De2	T3	Hu2	L13	97.298
OCT	H3	S8	D3	De2	T3	Hu1	L18	97.505
OCT	H6	S7	D2	De3	T2	Hu3	L14	97.535
OCT	H1	S6	D1	De2	T3	Hu2	L23	97.542

## ตารางที่ 69 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
OCT	H5	S7	D4	De3	T3	Hu3	L23	97.643
OCT	H3	S6	D1	De1	T3	Hu2	L18	97.724
OCT	H9	S8	D3	De3	T3	Hu3	L18	97.778
OCT	H1	S6	D3	De2	T3	Hu2	L4	97.853
OCT	H5	S7	D3	De2	T3	Hu3	L23	97.866
OCT	H5	S7	D2	De3	T3	Hu3	L23	97.886
OCT	H7	S7	D2	De3	T3	Hu2	L23	97.926
OCT	H5	S7	D2	De3	T3	Hu2	L21	98.120
OCT	H3	S8	D4	De5	T3	Hu3	L15	98.238
OCT	H6	S7	D1	De2	T3	Hu2	L23	98.251
OCT	H7	S7	D3	De3	T2	Hu2	L21	98.269
OCT	H7	S7	D3	De3	T2	Hu2	L13	98.356
OCT	H3	S8	D4	De4	T3	Hu1	L18	98.438
OCT	H3	S8	D4	De2	T3	Hu3	L18	98.446
OCT	H7	S7	D4	De4	T2	Hu2	L23	98.464
OCT	H3	S6	D3	De4	T3	Hu3	L13	98.468
OCT	H6	S7	D4	De5	T3	Hu3	L14	98.473
OCT	H2	S3	D4	De5	T3	Hu2	L15	98.487
OCT	H5	S7	D2	De2	T3	Hu2	L23	98.494
OCT	H1	S3	D1	De3	T3	Hu2	L23	98.536
OCT	H6	S7	D2	De2	T2	Hu2	L23	98.552
OCT	H4	S4	D3	De3	T3	Hu2	L18	98.578
OCT	H3	S8	D4	De5	T2	Hu2	L23	98.605

## ตารางที่ 69 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
OCT	H6	S7	D1	De2	T3	Hu2	L14	98.684
OCT	H3	S6	D2	De2	T2	Hu2	L23	98.701
OCT	H1	S6	D2	De3	T3	Hu2	L23	98.729
OCT	H3	S6	D2	De3	T2	Hu2	L23	98.736
OCT	H3	S6	D3	De4	T3	Hu2	L20	98.791
OCT	H5	S7	D3	De3	T3	Hu3	L23	98.812
OCT	H6	S7	D2	De2	T2	Hu2	L14	98.845
OCT	H7	S7	D3	De3	T2	Hu2	L23	98.850
OCT	H1	S6	D4	De2	T3	Hu2	L1	98.869
OCT	H6	S3	D2	De2	T3	Hu2	L23	98.886
OCT	H2	S3	D1	De3	T3	Hu2	L15	98.925
OCT	H3	S8	D3	De2	T3	Hu2	L21	98.947
OCT	H1	S3	D2	De2	T3	Hu2	L23	98.979
OCT	H6	S7	D2	De4	T2	Hu3	L14	98.984
OCT	H3	S6	D2	De5	T3	Hu3	L18	99.041
OCT	H6	S5	D4	De5	T3	Hu2	L14	99.042
OCT	H3	S8	D2	De5	T2	Hu2	L21	99.062
OCT	H3	S6	D2	De2	T3	Hu2	L21	99.071
OCT	H4	S8	D2	De2	T3	Hu3	L18	99.089
OCT	H3	S8	D4	De5	T3	Hu1	L23	99.108
OCT	H6	S7	D2	De4	T3	Hu2	L13	99.110
OCT	H6	S7	D2	De3	T2	Hu2	L23	99.141
OCT	H3	S6	D2	De3	T2	Hu2	L18	99.173

## ตารางที่ 69 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
OCT	H5	S7	D3	De2	T3	Hu2	L23	99.212
OCT	H6	S7	D3	De2	T3	Hu2	L23	99.218
OCT	H3	S6	D2	De4	T3	Hu2	L18	99.272
OCT	H1	S6	D2	De2	T3	Hu2	L18	99.305
OCT	H3	S6	D2	De3	T3	Hu3	L18	99.310
OCT	H5	S7	D1	De2	T3	Hu2	L23	99.322
OCT	H3	S8	D3	De4	T3	Hu1	L18	99.328
OCT	H3	S6	D2	De2	T3	Hu3	L18	99.337
OCT	H6	S7	D3	De3	T2	Hu2	L23	99.382
OCT	H5	S7	D2	De3	T3	Hu2	L23	99.392
OCT	H6	S7	D3	De3	T2	Hu2	L14	99.392
OCT	H3	S6	D4	De4	T3	Hu3	L18	99.419
OCT	H6	S3	D2	De3	T2	Hu2	L2	99.424
OCT	H3	S8	D4	De5	T2	Hu2	L21	99.428
OCT	H1	S3	D2	De4	T3	Hu2	L15	99.429
OCT	H6	S7	D4	De5	T2	Hu2	L20	99.432
OCT	H6	S3	D2	De3	T3	Hu2	L23	99.436
OCT	H3	S6	D1	De3	T2	Hu2	L18	99.448
OCT	H6	S3	D1	De3	T2	Hu2	L23	99.465
OCT	H6	S3	D2	De2	T2	Hu2	L23	99.469
OCT	H6	S7	D3	De5	T3	Hu3	L13	99.472
OCT	H3	S8	D1	De3	T3	Hu2	L18	99.487
OCT	H3	S8	D3	De3	T2	Hu2	L18	99.489

## ตารางที่ 69 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
OCT	H3	S6	D3	De1	T3	Hu2	L23	99.548
OCT	H3	S6	D3	De3	T2	Hu2	L18	99.567
OCT	H3	S6	D3	De4	T2	Hu2	L23	99.569
OCT	H3	S6	D2	De2	T3	Hu2	L13	99.570
OCT	H6	S7	D2	De3	T3	Hu2	L14	99.615
OCT	H3	S8	D3	De2	T3	Hu3	L18	99.630
OCT	H6	S7	D3	De3	T3	Hu2	L14	99.639
OCT	H3	S8	D1	De2	T3	Hu2	L18	99.661
OCT	H6	S7	D3	De3	T3	Hu2	L23	99.666
OCT	H1	S3	D2	De3	T3	Hu2	L23	99.668
OCT	H3	S6	D3	De3	T3	Hu3	L18	99.669
OCT	H3	S8	D3	De5	T2	Hu2	L23	99.682
OCT	H6	S7	D1	De5	T3	Hu3	L13	99.689
OCT	H3	S6	D2	De3	T3	Hu2	L21	99.700
OCT	H6	S7	D2	De4	T2	Hu2	L23	99.704
OCT	H5	S7	D3	De3	T3	Hu2	L23	99.711
OCT	H3	S8	D4	De4	T3	Hu3	L18	99.711
OCT	H1	S6	D4	De4	T3	Hu2	L18	99.715
OCT	H3	S8	D3	De4	T3	Hu2	L13	99.721
OCT	H3	S6	D4	De3	T3	Hu2	L21	99.727
OCT	H3	S8	D2	De5	T3	Hu1	L18	99.738
OCT	H3	S8	D4	De5	T3	Hu3	L13	99.744
OCT	H3	S8	D2	De2	T3	Hu2	L18	99.745

## ตารางที่ 69 (ต่อ)

month	Height	Landscape	Death	Destruction	Temperature	Humidity		
						ity	DLH	P*100
OCT	H3	S6	D3	De4	T2	Hu2	L18	99.788
OCT	H3	S6	D2	De2	T3	Hu2	L23	99.792
OCT	H5	S7	D2	De4	T3	Hu2	L23	99.812
OCT	H6	S7	D2	De4	T3	Hu2	L14	99.814
OCT	H3	S8	D4	De5	T3	Hu3	L23	99.825
OCT	H3	S8	D4	De5	T3	Hu1	L18	99.830
OCT	H5	S7	D4	De5	T3	Hu2	L23	99.839
OCT	H3	S6	D1	De2	T3	Hu2	L23	99.842
OCT	H3	S8	D3	De3	T3	Hu3	L18	99.848
OCT	H3	S6	D1	De3	T3	Hu2	L13	99.850
OCT	H3	S8	D2	De3	T3	Hu2	L18	99.852
OCT	H3	S6	D3	De4	T3	Hu2	L21	99.862
OCT	H3	S6	D3	De2	T3	Hu2	L23	99.871
OCT	H3	S6	D2	De3	T3	Hu2	L23	99.882
OCT	H3	S6	D2	De2	T3	Hu2	L18	99.884
OCT	H3	S6	D3	De2	T3	Hu2	L18	99.887
OCT	H6	S7	D4	De5	T3	Hu2	L23	99.891
OCT	H3	S6	D1	De4	T3	Hu2	L13	99.893
OCT	H3	S8	D4	De5	T2	Hu2	L18	99.898
OCT	H3	S6	D4	De5	T2	Hu2	L18	99.901
OCT	H3	S6	D4	De5	T3	Hu3	L18	99.902
OCT	H6	S7	D3	De5	T3	Hu2	L13	99.907
OCT	H6	S7	D4	De5	T3	Hu2	L14	99.909

## ตารางที่ 69 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
OCT	H3	S8	D4	De5	T3	Hu2	L13	99.921
OCT	H3	S8	D4	De5	T3	Hu2	L17	99.922
OCT	H3	S8	D3	De2	T3	Hu2	L18	99.924
OCT	H1	S6	D3	De4	T3	Hu2	L18	99.928
OCT	H3	S8	D4	De4	T3	Hu2	L18	99.937
OCT	H3	S6	D3	De4	T3	Hu2	L13	99.938
OCT	H7	S7	D1	De5	T2	Hu2	L23	99.939
OCT	H6	S7	D1	De4	T3	Hu2	L23	99.944
OCT	H6	S7	D3	De5	T2	Hu2	L23	99.945
OCT	H3	S6	D4	De2	T3	Hu2	L18	99.948
OCT	H3	S6	D3	De3	T3	Hu2	L23	99.949
OCT	H3	S8	D2	De4	T3	Hu2	L18	99.949
OCT	H3	S8	D4	De5	T3	Hu3	L18	99.949
OCT	H3	S8	D3	De5	T2	Hu2	L18	99.955
OCT	H3	S8	D2	De5	T3	Hu3	L18	99.959
OCT	H3	S6	D3	De4	T3	Hu2	L23	99.962
OCT	H3	S8	D3	De3	T3	Hu2	L18	99.964
OCT	H3	S8	D4	De5	T3	Hu2	L23	99.965
OCT	H3	S6	D3	De3	T3	Hu2	L18	99.965
OCT	H3	S6	D1	De3	T3	Hu2	L18	99.969
OCT	H3	S8	D3	De4	T3	Hu2	L18	99.970
OCT	H3	S6	D3	De4	T3	Hu2	L18	99.970
OCT	H3	S6	D4	De5	T3	Hu2	L23	99.971

## ตารางที่ 69 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
OCT	H1	S6	D4	De5	T3	Hu2	L18	99.975
OCT	H3	S8	D1	De5	T3	Hu3	L18	99.982
OCT	H3	S8	D3	De5	T3	Hu2	L18	99.983
OCT	H3	S6	D4	De5	T3	Hu2	L18	99.988
OCT	H3	S8	D3	De5	T3	Hu3	L18	99.988
OCT	H3	S8	D4	De5	T3	Hu2	L18	99.988
OCT	H3	S8	D2	De5	T3	Hu2	L18	99.989
OCT	H3	S6	D2	De5	T3	Hu2	L23	99.993
OCT	H3	S6	D2	De5	T3	Hu2	L18	99.994
OCT	H3	S6	D3	De5	T3	Hu2	L23	99.995
<b>OCT</b>	<b>H3</b>	<b>S6</b>	<b>D3</b>	<b>De5</b>	<b>T3</b>	<b>Hu2</b>	<b>L18</b>	<b>99.996</b>

หมายเหตุ DLH = Day Light Hour, P = Probability

จากตารางที่ 69 ลักษณะพื้นที่ เลี้ยงสัตว์ปีกที่มีโอกาสเกิดโรคไข้หวัดนกมากที่สุดคือ  
ลักษณะพื้นที่มีองค์ประกอบดังต่อไปนี้

1. ในช่วงเดือน ตุลาคม
2. มีความสูงจากระดับน้ำทะเล อยู่ระหว่าง 6.5 เมตร ถึง 12 เมตร
3. อุณหภูมิอยู่ในช่วง 27.85 ถึง 32.05 องศาเซลเซียส
4. อยู่ในพื้นที่ที่มีลักษณะอยู่ในบริเวณที่ราบลุ่มสามเหลี่ยมปากแม่น้ำ
5. มีจำนวนสัตว์ปีกที่ตายอยู่ระหว่าง 18 ถึง 98 ตัว
6. มีจำนวนสัตว์ปีกที่ถูกทำลายอยู่ในช่วงมากกว่า 969 ตัว
7. มีความชื้นสัมพัทธ์ ในช่วง 60.5 -75.09 %
8. มีจำนวนชั่วโมงที่มีแสงแดด อยู่ในช่วง 8 ชั่วโมง/วัน

ตารางที่ 70 ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วงระยะเวลา  
เดือน พฤศจิกายน

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
NOV	H7	S7	D2	De3	T2	Hu2	L23	50.238
NOV	H3	S8	D2	De1	T3	Hu2	L23	53.355
NOV	H6	S7	D1	De4	T3	Hu3	L13	53.802
NOV	H3	S6	D2	De1	T3	Hu2	L23	53.980
NOV	H4	S6	D4	De5	T2	Hu2	L23	58.990
NOV	H3	S8	D3	De3	T3	Hu3	L1	61.200
NOV	H7	S4	D2	De2	T3	Hu2	L16	62.144
NOV	H2	S3	D4	De1	T3	Hu2	L13	63.038
NOV	H3	S8	D3	De3	T2	Hu3	L21	63.250
NOV	H6	S3	D4	De4	T3	Hu2	L1	64.097
NOV	H6	S7	D2	De2	T3	Hu3	L16	65.468
NOV	H3	S6	D3	De1	T2	Hu2	L23	69.583
NOV	H3	S8	D3	De1	T2	Hu2	L23	69.714
NOV	H2	S4	D2	De2	T2	Hu2	L15	69.832
NOV	H3	S8	D1	De1	T3	Hu2	L23	70.026
NOV	H4	S6	D4	De5	T2	Hu3	L1	70.192
NOV	H6	S3	D1	De3	T2	Hu3	L17	71.316
NOV	H6	S3	D3	De2	T3	Hu3	L13	73.311
NOV	H6	S7	D1	De4	T2	Hu3	L13	73.764
NOV	H3	S8	D4	De5	T3	Hu3	L1	74.283
NOV	H3	S8	D4	De3	T3	Hu3	L23	75.054
NOV	H3	S8	D3	De1	T3	Hu1	L23	75.228
NOV	H1	S6	D1	De2	T3	Hu2	L23	76.766

## ตารางที่ 70 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
NOV	H4	S4	D2	De3	T3	Hu1	L22	79.197
NOV	H6	S7	D3	De3	T2	Hu3	L1	80.771
NOV	H2	S3	D1	De2	T3	Hu2	L13	82.622
NOV	H3	S6	D4	De1	T3	Hu2	L20	84.724
NOV	H4	S8	D3	De3	T3	Hu2	L22	85.886
NOV	H1	S6	D2	De2	T3	Hu2	L15	86.956
NOV	H6	S7	D2	De2	T2	Hu3	L23	87.137
NOV	H6	S7	D1	De2	T2	Hu1	L23	88.002
NOV	H6	S7	D3	De1	T2	Hu2	L20	88.227
NOV	H1	S6	D3	De2	T3	Hu2	L15	91.031
NOV	H6	S7	D1	De3	T2	Hu2	L15	91.062
NOV	H1	S6	D3	De2	T3	Hu2	L21	91.224
NOV	H1	S6	D3	De2	T3	Hu1	L15	91.297
NOV	H3	S6	D3	De2	T2	Hu2	L23	91.684
NOV	H3	S8	D2	De3	T2	Hu2	L23	92.364
NOV	H6	S7	D3	De1	T2	Hu2	L23	92.375
NOV	H4	S8	D3	De2	T3	Hu2	L23	92.421
NOV	H6	S7	D2	De2	T2	Hu1	L23	92.667
NOV	H6	S3	D4	De5	T3	Hu2	L1	93.054
NOV	H3	S8	D4	De4	T3	Hu3	L22	94.219
NOV	H1	S3	D2	De3	T3	Hu2	L23	94.406
NOV	H1	S6	D1	De2	T3	Hu2	L22	95.088
NOV	H2	S3	D4	De5	T3	Hu1	L20	95.110

## ตารางที่ 70 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
NOV	H6	S3	D3	De3	T3	Hu1	L23	96.015
NOV	H3	S6	D2	De2	T3	Hu2	L21	96.336
NOV	H3	S8	D3	De3	T2	Hu2	L23	96.384
NOV	H6	S7	D1	De3	T2	Hu1	L23	96.415
NOV	H3	S8	D1	De2	T3	Hu2	L23	96.642
NOV	H3	S6	D1	De2	T3	Hu2	L1	96.884
NOV	H6	S7	D1	De3	T2	Hu2	L1	96.908
NOV	H6	S5	D2	De5	T2	Hu3	L16	96.932
NOV	H6	S5	D3	De3	T3	Hu3	L16	97.050
NOV	H4	S4	D3	De5	T3	Hu2	L22	97.184
NOV	H6	S7	D1	De2	T2	Hu2	L23	97.192
NOV	H6	S7	D3	De3	T2	Hu1	L23	97.194
NOV	H3	S6	D3	De2	T3	Hu2	L21	97.216
NOV	H6	S7	D2	De3	T2	Hu2	L1	97.607
NOV	H6	S7	D3	De2	T2	Hu2	L23	97.890
NOV	H3	S6	D3	De2	T3	Hu2	L23	97.894
NOV	H6	S7	D2	De2	T2	Hu2	L23	97.970
NOV	H6	S7	D1	De2	T2	Hu2	L16	98.048
NOV	H9	S8	D4	De5	T2	Hu2	L22	98.079
NOV	H6	S7	D3	De3	T2	Hu2	L1	98.105
NOV	H5	S7	D4	De5	T2	Hu2	L23	98.186
NOV	H6	S7	D1	De3	T2	Hu2	L23	98.400
NOV	H3	S8	D3	De4	T2	Hu2	L23	98.448

## ตารางที่ 70 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
NOV	H6	S7	D2	De3	T2	Hu2	L23	98.576
NOV	H3	S6	D2	De3	T3	Hu2	L23	98.768
NOV	H6	S7	D3	De3	T2	Hu2	L20	98.820
NOV	H6	S7	D1	De2	T3	Hu2	L23	98.870
NOV	H6	S7	D3	De2	T2	Hu2	L16	98.893
NOV	H3	S8	D1	De2	T2	Hu2	L22	98.923
NOV	H6	S7	D2	De2	T2	Hu2	L16	98.939
NOV	H6	S7	D4	De3	T2	Hu2	L16	98.961
NOV	H6	S3	D2	De3	T3	Hu1	L23	98.963
NOV	H3	S6	D4	De5	T2	Hu1	L23	99.001
NOV	H3	S8	D4	De5	T2	Hu2	L23	99.005
NOV	H3	S6	D4	De4	T3	Hu2	L17	99.006
NOV	H3	S6	D3	De3	T3	Hu2	L21	99.038
NOV	H6	S7	D3	De3	T2	Hu2	L23	99.076
NOV	H3	S8	D4	De5	T3	Hu1	L23	99.182
NOV	H5	S7	D4	De5	T3	Hu2	L15	99.256
NOV	H6	S3	D2	De3	T2	Hu2	L23	99.268
NOV	H1	S6	D1	De3	T3	Hu1	L22	99.276
NOV	H3	S8	D4	De5	T3	Hu2	L13	99.296
NOV	H9	S8	D3	De2	T3	Hu2	L22	99.329
NOV	H6	S7	D2	De2	T3	Hu2	L23	99.346
NOV	H3	S8	D1	De3	T3	Hu2	L20	99.393
NOV	H3	S8	D3	De4	T2	Hu2	L20	99.410

## ตารางที่ 70 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
NOV	H1	S6	D3	De4	T3	Hu1	L22	99.486
NOV	H6	S7	D3	De2	T3	Hu2	L23	99.522
NOV	H6	S7	D1	De2	T3	Hu2	L16	99.533
NOV	H6	S7	D2	De4	T2	Hu2	L23	99.533
NOV	H3	S6	D3	De2	T2	Hu1	L22	99.537
NOV	H3	S6	D3	De2	T2	Hu2	L22	99.551
NOV	H3	S8	D2	De3	T2	Hu2	L22	99.561
NOV	H6	S7	D2	De3	T3	Hu2	L23	99.562
NOV	H9	S8	D3	De3	T3	Hu2	L22	99.634
NOV	H3	S8	D4	De3	T3	Hu2	L23	99.666
NOV	H3	S8	D3	De2	T2	Hu2	L22	99.688
NOV	H6	S7	D3	De3	T3	Hu2	L23	99.704
NOV	H3	S8	D4	De4	T3	Hu1	L22	99.705
NOV	H6	S7	D2	De2	T3	Hu2	L16	99.717
NOV	H1	S6	D1	De4	T3	Hu2	L22	99.722
NOV	H6	S3	D4	De5	T3	Hu2	L23	99.728
NOV	H3	S8	D3	De3	T3	Hu1	L22	99.777
NOV	H6	S7	D2	De3	T3	Hu2	L16	99.783
NOV	H3	S8	D4	De5	T2	Hu1	L22	99.813
NOV	H5	S7	D3	De5	T3	Hu2	L15	99.842
NOV	H3	S8	D1	De2	T3	Hu2	L22	99.844
NOV	H5	S7	D3	De3	T3	Hu2	L23	99.851
NOV	H3	S6	D1	De2	T3	Hu2	L22	99.852

ตารางที่ 70 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity		
						y	DLH	P*100
NOV	H3	S8	D2	De2	T3	Hu2	L22	99.855
NOV	H3	S8	D2	De3	T3	Hu2	L22	99.856
NOV	H3	S8	D3	De4	T3	Hu1	L22	99.860
NOV	H3	S6	D2	De2	T3	Hu2	L22	99.860
NOV	H6	S3	D4	De5	T3	Hu2	L20	99.866
NOV	H3	S8	D3	De3	T2	Hu2	L22	99.876
NOV	H3	S6	D4	De4	T3	Hu2	L22	99.893
NOV	H5	S7	D3	De4	T3	Hu2	L23	99.906
NOV	H3	S6	D3	De2	T3	Hu2	L22	99.908
NOV	H3	S6	D2	De5	T3	Hu2	L23	99.913
NOV	H3	S8	D3	De4	T2	Hu2	L22	99.914
NOV	H9	S8	D4	De5	T3	Hu2	L22	99.920
NOV	H3	S6	D4	De5	T2	Hu1	L22	99.920
NOV	H3	S8	D4	De3	T3	Hu2	L22	99.921
NOV	H5	S7	D4	De5	T3	Hu2	L23	99.923
NOV	H6	S7	D2	De5	T2	Hu2	L23	99.924
NOV	H6	S7	D4	De5	T3	Hu2	L23	99.926
NOV	H3	S6	D2	De3	T3	Hu2	L22	99.931
NOV	H3	S8	D4	De5	T3	Hu2	L21	99.949
NOV	H3	S8	D1	De3	T3	Hu2	L22	99.950
NOV	H3	S8	D2	De5	T3	Hu2	L22	99.952
NOV	H3	S8	D3	De2	T3	Hu2	L22	99.954
NOV	H3	S8	D4	De5	T3	Hu1	L22	99.956

## ตารางที่ 70 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
NOV	H3	S6	D4	De5	T2	Hu3	L22	99.971
NOV	H3	S8	D4	De4	T3	Hu2	L22	99.972
NOV	<b>H3</b>	<b>S8</b>	<b>D4</b>	<b>De5</b>	<b>T2</b>	<b>Hu2</b>	<b>L22</b>	<b>99.978</b>

หมายเหตุ DLH = Day Light Hour, P = Probability

จากตารางที่ 70 ลักษณะพื้นที่ เลี้ยงสัตว์ปีกที่มีโอกาสเกิดโรคไข้หวัดนกมากที่สุดคือ ลักษณะพื้นที่มีองค์ประกอบดังต่อไปนี้

1. ในช่วงเดือนพฤศจิกายน
2. มีระดับความสูงจากระดับน้ำทะเลอยู่ระหว่าง 6.5 - 12 เมตร
3. อุณหภูมิ อยู่ระหว่าง 21.15 – 27.85 องศาเซลเซียส
4. ลักษณะพื้นที่เป็นขอบที่ราบภาคกลาง
5. จำนวนสัตว์ปีกป่วยตายในช่วง มากกว่า 98 ตัว
6. จำนวนสัตว์ปีกที่ถูกทำลาย อยู่ในช่วงมากกว่า 968.5 ตัว
7. ความชื้นสัมพัทธ์ อยู่ในช่วง 60.5 – 75.09 %
8. จำนวนชั่วโมงที่มีแสงแดด อยู่ในช่วง 8 ชั่วโมง/วัน

ตารางที่ 71 ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกในช่วงระยะเวลา  
เดือน ธันวาคม

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
DEC	H6	S7	D2	De1	T2	Hu2	L21	53.665
DEC	H3	S6	D2	De1	T2	Hu2	L23	55.774
DEC	H6	S3	D1	De2	T3	Hu2	L13	60.146
DEC	H6	S7	D1	De3	T2	Hu2	L13	61.657
DEC	H6	S3	D4	De2	T3	Hu2	L13	62.033
DEC	H3	S2	D2	De2	T2	Hu3	L13	73.752
DEC	H6	S7	D1	De1	T2	Hu2	L23	76.911
DEC	H6	S3	D1	De3	T3	Hu2	L13	78.902
DEC	H2	S4	D2	De2	T3	Hu2	L13	84.395
DEC	H6	S7	D2	De1	T2	Hu2	L23	85.223
DEC	H6	S7	D2	De2	T2	Hu2	L20	92.342
DEC	H6	S7	D2	De2	T2	Hu1	L23	95.351
DEC	H6	S7	D3	De4	T2	Hu2	L15	96.158
DEC	H3	S6	D1	De3	T2	Hu2	L23	97.639
DEC	H6	S7	D2	De2	T2	Hu2	L21	97.802
DEC	H6	S7	D3	De2	T2	Hu2	L21	97.863
DEC	H6	S7	D4	De2	T2	Hu2	L21	97.875
DEC	H3	S6	D4	De5	T3	Hu2	L11	97.991
DEC	H6	S7	D1	De3	T2	Hu2	L21	98.254
DEC	H6	S7	D1	De2	T2	Hu2	L23	98.474
DEC	H6	S7	D2	De3	T2	Hu2	L21	98.745
DEC	H6	S7	D3	De3	T2	Hu2	L21	98.993
DEC	H6	S7	D3	De2	T2	Hu2	L23	99.090

## ตารางที่ 71 (ต่อ)

Month	Height	Landscape	Death	Destruction	Temperature	Humidity	DLH	P*100
DEC	H3	S8	D2	De3	T2	Hu2	L15	99.170
DEC	H3	S8	D2	De2	T2	Hu2	L22	99.321
DEC	H3	S8	D1	De3	T2	Hu2	L15	99.371
DEC	H6	S7	D2	De2	T2	Hu2	L23	99.375
DEC	H6	S7	D2	De3	T2	Hu2	L23	99.407
DEC	H9	S8	D2	De5	T2	Hu1	L22	99.474
DEC	H5	S7	D3	De2	T2	Hu2	L23	99.575
DEC	H3	S8	D2	De3	T2	Hu2	L22	99.584
DEC	H6	S7	D4	De2	T2	Hu2	L23	99.586
DEC	H3	S8	D4	De5	T3	Hu1	L22	99.594
DEC	H6	S7	D3	De4	T2	Hu2	L23	99.626
DEC	H6	S7	D4	De3	T2	Hu2	L23	99.695
DEC	H6	S7	D3	De3	T2	Hu2	L23	99.705
DEC	H9	S8	D3	De5	T2	Hu1	L22	99.707
DEC	H5	S7	D4	De5	T3	Hu2	L23	99.716
DEC	H5	S7	D2	De2	T2	Hu2	L23	99.732
DEC	H5	S7	D3	De3	T2	Hu2	L23	99.842
DEC	H6	S7	D2	De4	T2	Hu2	L23	99.851
DEC	H3	S8	D4	De4	T2	Hu2	L22	99.858
DEC	H6	S7	D4	De5	T2	Hu2	L23	99.868
DEC	H5	S7	D4	De5	T2	Hu2	L23	99.950
DEC	H6	S7	D1	De4	T2	Hu2	L23	99.950
<b>DEC</b>	<b>H6</b>	<b>S7</b>	<b>D3</b>	<b>De5</b>	<b>T2</b>	<b>Hu2</b>	<b>L23</b>	<b>99.977</b>

หมายเหตุ DLH = Day Light Hour, P = Probability

จากตารางที่ 71 ลักษณะพื้นที่เลี้ยงสัตว์ปีกที่มีโอกาสเกิดโรคไข้หวัดนกมากที่สุดคือ ลักษณะพื้นที่มีองค์ประกอบดังต่อไปนี้

1. ในช่วงเดือน ธันวาคม
2. มีระดับความสูงจากระดับน้ำทะเลอยู่ระหว่าง 34.5 ถึง 97 เมตร
3. อุณหภูมิ อยู่ระหว่าง 21.15 ถึง 27.85 องศาเซลเซียส
4. ลักษณะพื้นที่เป็นที่ราบลุ่มน้ำภาคกลางตอนบน
5. จำนวนสัตว์ปีกป่วยตายในช่วง 18 ถึง 98 ตัว
6. จำนวนสัตว์ปีกที่ถูกทำลาย อยู่ในช่วงมากกว่า 969 ตัว
7. ความชื้นสัมพัทธ์ อยู่ในช่วง 60.5 – 75.09 %
8. จำนวนชั่วโมงที่มีแสงแดด อยู่ในช่วง มากกว่า 8 ชั่วโมง/วัน

จากตารางที่ 65 ถึง 71 ได้อธิบายถึงลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดโรคไข้หวัดนกซึ่งลักษณะพื้นที่ ได้มาจากการสกัดมาจากแบบจำลองที่สร้างด้วย Bayesian Network แล้วเลือกเอาลักษณะพื้นที่ที่มีความน่าจะเป็นที่สูงมาสร้างเป็นกฎเพื่อใช้ในการเฝ้าระวังของโรคไข้หวัดนกซึ่งได้จำนวนกฎทั้งสิ้น 1313 กฎ ซึ่งทำให้ได้ทราบว่า ลักษณะพื้นที่ที่มีความเสี่ยงที่จะเกิดโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีก เช่น

ลักษณะพื้นที่ที่มีความสูงจากระดับน้ำทะเลอยู่ระหว่าง 6.5 ถึง 12 เมตร อุณหภูมิอยู่ในช่วง 21.15 องศาเซลเซียส ถึง 32.05 องศาเซลเซียส ลักษณะพื้นที่เป็นแบบ ที่ราบลุ่มน้ำภาคกลางตอนบน มีสัตว์ปีกที่ตายอยู่ระหว่าง 18 – 98 ตัว และมีสัตว์ปีกที่ถูกทำลายมากกว่า 969 ตัว และ อยู่ในช่วงเดือน ธันวาคม

## 2. แบบจำลองที่ด้วยวิธีต้นไม้ตัดสินใจ (Decision Tree J48)

นำชุดข้อมูลที่เตรียมไว้ทั้งหมดสิบชุด ซึ่งประกอบด้วยชุดข้อมูล A B C D E F G H I และ J ที่แต่ละชุดผ่านการเตรียมข้อมูลด้วยวิธีการที่แตกต่างกัน นำมาสร้างแบบจำลองด้วยวิธี ต้นไม้ตัดสินใจ (Decision Tree) และ นำแบบจำลองที่ได้ ทำการวัดประสิทธิภาพของแบบจำลอง ซึ่งประสิทธิภาพของแบบจำลองเป็นไปตามตารางที่ 72 และ ตารางที่ 73

ตารางที่ 72 ประสิทธิภาพของแบบจำลองที่สร้างด้วยต้นไม้ตัดสินใจ (J48) โดยวัดจากความถูกต้องของแบบจำลองในการจำแนกข้อมูล

Algorithm	Data set	Correctly Classify(%)	Incorrect Classify(%)
Decision Tree (J48)	A	89.25	10.75
Decision Tree (J48)	B	89.65	10.35
Decision Tree (J48)	C	89.35	10.65
Decision Tree (J48)	D	90.18	9.82
Decision Tree (J48)	E	89.76	10.24
Decision Tree (J48)	F	89.65	10.35
Decision Tree (J48)	G	83.52	16.48
Decision Tree (J48)	H	94.14	5.86
<b>Decision Tree (J48)</b>	<b>I</b>	<b>95.19</b>	<b>4.81</b>
Decision Tree (J48)	J	84.02	15.98

จากตารางที่ 72 แสดงให้เห็นถึงประสิทธิภาพการจำแนกข้อมูล (Accuracy) ของแบบจำลองที่สร้างด้วย Decision Tree (J48) ซึ่งจากตารางที่ 72 แบบจำลองที่สร้างด้วย Decision Tree ที่มีประสิทธิภาพการจำแนกข้อมูลที่ดีที่สุด (Accuracy) คือ แบบจำลองที่สร้างจากข้อมูล ที่ผ่านการแบ่งช่วงข้อมูล และ ทำการถ่วงน้ำหนักให้กับข้อมูล โดยน้ำหนักที่นำมาถ่วงให้กับชุดข้อมูล มาจากวิธีการหาน้ำหนักด้วยวิธี Emerging Pattern (ชุดข้อมูล I) โดยมีประสิทธิภาพการจำแนกข้อมูล (Accuracy) อยู่ที่ 95.19 %

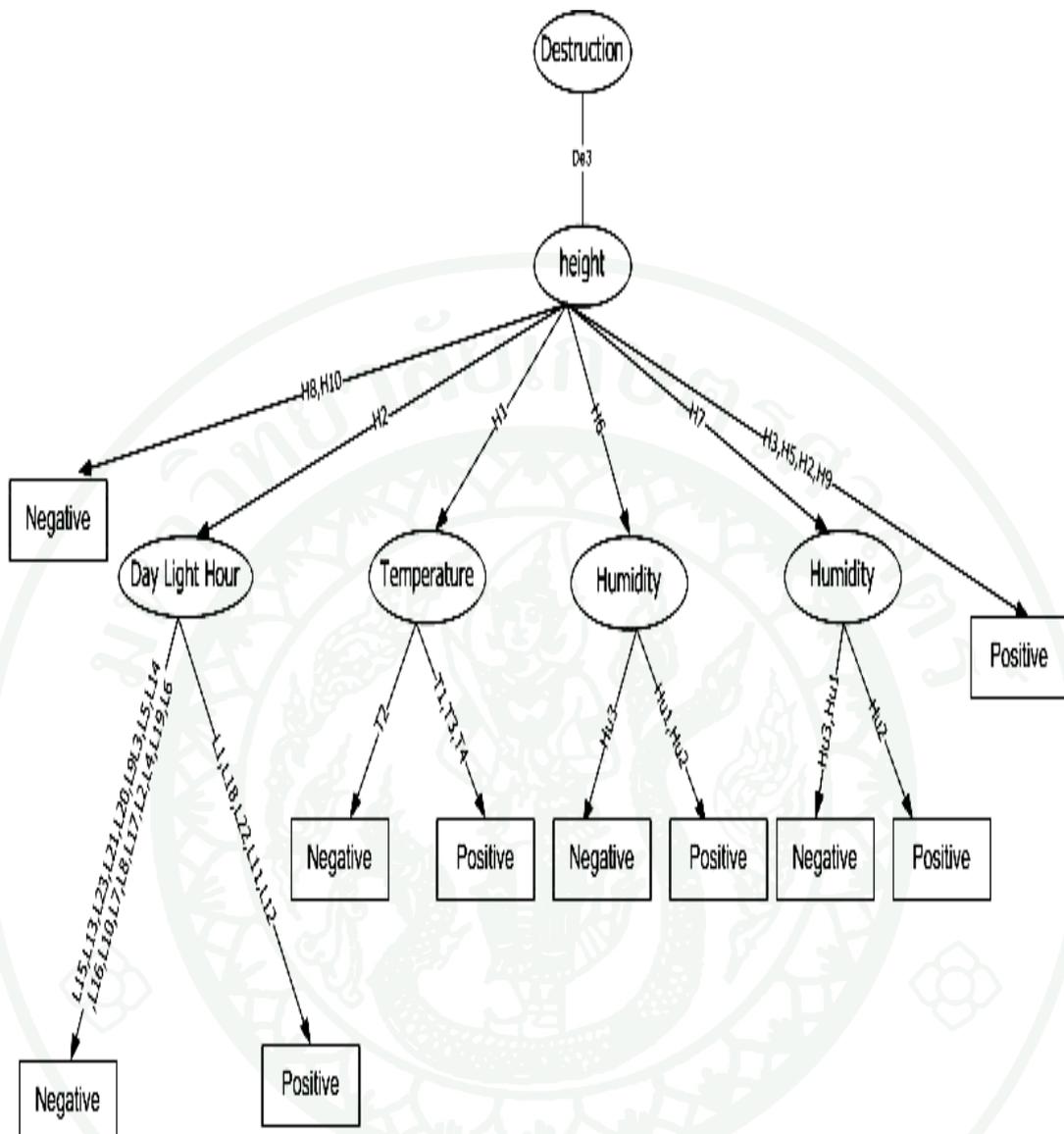
ตารางที่ 73 ค่า Precision Recall และ F-measure ที่แสดงประสิทธิภาพของแบบจำลองที่สร้างจากต้นไม้ตัดสินใจ (J48) โดยใช้ชุดข้อมูลที่มีวิธีการเตรียมที่แตกต่างกัน

Algorithm	Data set	Precision	Recall	F-measure	Class
Decision	A	0.724	0.568	0.637	Positive
Tree (J48)		0.918	0.957	0.937	Negative
Decision	B	0.808	0.609	0.695	Positive
Tree (J48)		0.926	0.971	0.956	Negative
Decision	C	0.724	0.539	0.618	Positive
Tree (J48)		0.916	0.961	0.938	Negative
Decision	D	0.767	0.58	0.943	Positive
Tree (J48)		0.921	0.965	0.943	Negative
Decision	E	0.778	0.535	0.634	Positive
Tree (J48)		0.913	0.970	0.941	Negative
Decision	F	0.801	0.499	0.615	Positive
Tree (J48)		0.907	0.975	0.94	Negative
Decision	G	0.000	0.000	0.000	Positive
Tree (J48)		0.835	1.000	0.91	Negative
Decision	<b>H</b>	<b>0.940</b>	<b>0.885</b>	<b>0.912</b>	<b>Positive</b>
Tree (J48)		<b>0.942</b>	<b>0.971</b>	<b>0.956</b>	<b>Negative</b>
Decision	I	0.904	0.791	0.884	Positive
Tree (J48)		0.96	0.983	0.972	Negative
Decision	J	0.000	0.000	0.000	Positive
Tree (J48)		0.840	1.000	0.913	Negative

จากตารางที่ 73 แสดงให้เห็นถึงความแม่นยำ ในการจำแนกข้อมูลของแบบจำลองที่สร้าง จาก Decision Tree (J48) โดยวัดจากค่า Precision Recall และ F-measure ซึ่งจะพบว่า แบบจำลองที่ สร้างจากชุดข้อมูลที่มีการ Discretize การคัดเลือก Attribute และ การถ่วงน้ำหนักข้อมูล ซึ่ง น้ำหนักที่ใช้ในการ ถ่วงได้จากการหาน้ำหนักด้วยวิธี Emerging Pattern (ชุดข้อมูล H) เป็น แบบจำลองที่มีค่า Precision Recall และ F-measure สูงที่สุด ทั้งกรณีที่มีผลเฉลยข้อมูลเป็น Positive และ Negative

จากตารางผลการทดลองที่ 72 และ 73 จะพบว่า แบบจำลองที่สร้างด้วย ต้นไม้ตัดสินใจ (J48) ที่มีความถูกต้องในการจำแนกข้อมูล (Accuracy) ดีที่สุดคือ แบบจำลองที่สร้างจากชุดข้อมูลที่ ผ่านการ Discretize และ ถ่วงน้ำหนัก แต่ไม่ได้ทำการคัดเลือก Attribute (ชุดข้อมูล I) และ แบบจำลองที่สร้างจาก ต้นไม้ตัดสินใจ (J48) ที่มีค่า Precision , Recall และ F-measure ดีที่สุดคือ แบบจำลองที่สร้างจากชุดข้อมูลที่ผ่านการ Discretize ,การคัดเลือก Attribute และ การถ่วงด้วย น้ำหนักที่หามาจาก วิธี Emerging Pattern (ชุดข้อมูล H)

ดังนั้นในการเลือก แบบจำลองที่สร้างจากต้นไม้ตัดสินใจ (J48) ที่จะนำไปเปรียบเทียบกับ แบบจำลองที่สร้างจากอัลกอริทึมชนิดอื่นๆ นี้ จึงเลือกแบบจำลองที่สร้างจากชุดข้อมูล H เพราะใน การเลือกเอาแบบจำลองไปใช้ในการเฝ้าระวังโรคไข้หวัดนก พิจารณาจากความแม่นยำในทำนายว่า เป็นข้อมูลที่มีผลเฉลยเป็นบวก ซึ่งวัดได้จากค่า Precision ของแบบจำลอง ซึ่งจากตารางที่ 73 แสดง ให้เห็นว่า แบบจำลองที่สร้างด้วย ต้นไม้ตัดสินใจ (J48) โดยสร้างจากชุดข้อมูล H เป็นแบบจำลองที่ มีค่า Precision สูงกว่าแบบจำลองที่สร้างด้วย ต้นไม้ตัดสินใจ (J48) โดยสร้างชุดข้อมูลชนิด อื่นๆ



ภาพที่ 8 แบบจำลองที่สร้างด้วย อัลกอริทึม Decision Tree (J48) ที่สร้างจากชุดข้อมูลที่ผ่านการ Discretize การคัดเลือก Attribute และ การถ่วงน้ำหนัก (ในกรณีที่จำนวนสัตว์ที่ถูกทำลายอยู่ในช่วง 18 – 88 ตัว)

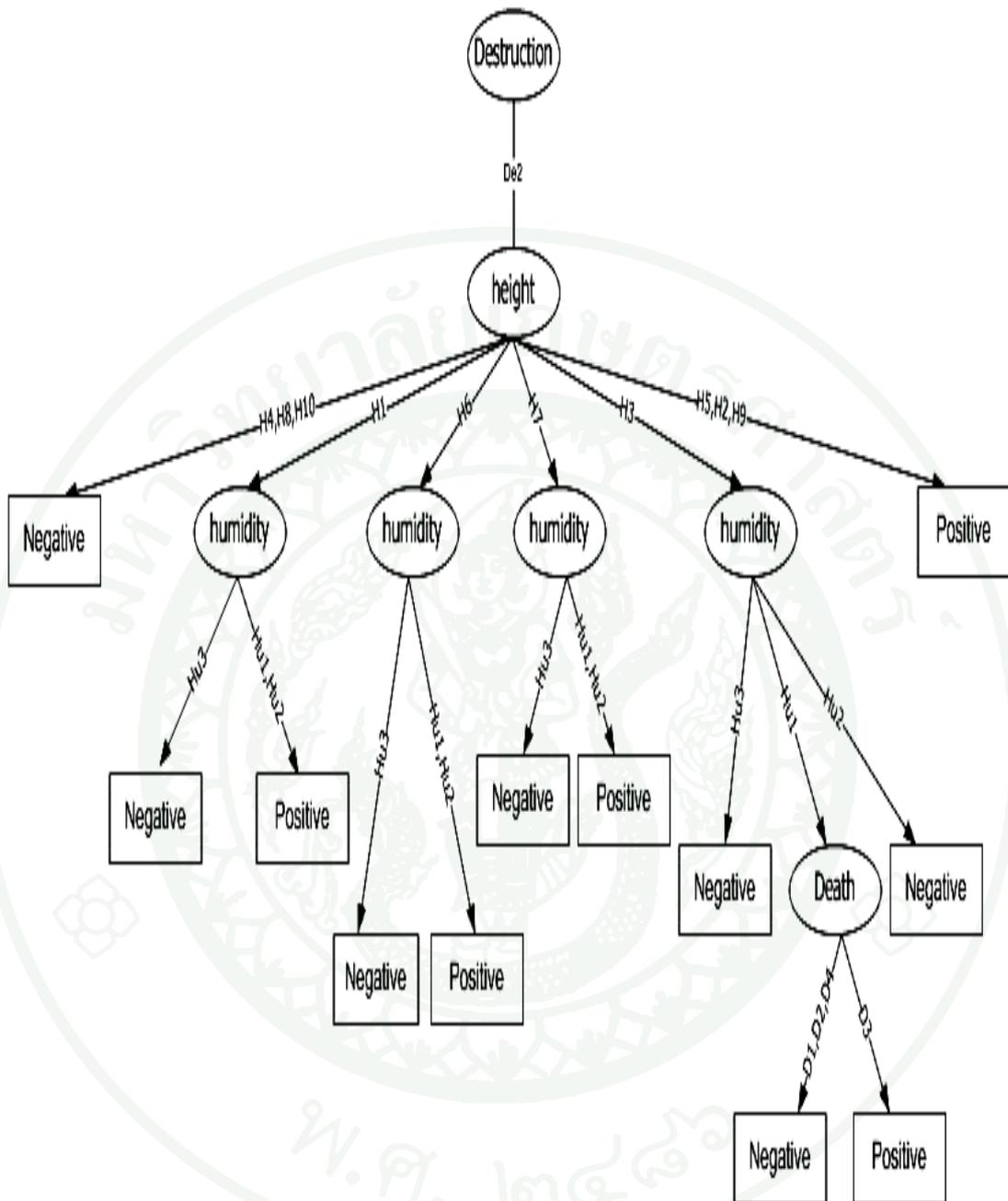
จากภาพที่ 8 แสดงให้เห็นถึงลักษณะของต้นไม้ตัดสินใจที่สร้างจากชุดข้อมูลที่ผ่านการ Discretize, การคัดเลือก Attribute และ การถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ Emerging pattern ซึ่งจากรูปสามารถนำเอาความสัมพันธ์ที่อยู่ในภาพที่ 8 นำไปสร้างเป็นกฎในรูปแบบของ Production Rule ดังตัวอย่างตามตารางด้านล่าง

ตารางที่ 74 กฎที่ได้มาจากต้นไม้ตัดสินใจ (J48) จากแบบจำลองในภาพที่ 8

กฎ	จำนวนข้อมูลที่อยู่ภายใต้กฎ
IF (Destruction = (17.5-87.5], Height = (34.5-97], Humidity = (60.5-75.099638] ) Then (Result = Positive)	183 Record
IF (Destruction = (17.5-87.5], Height = (97-124.5] Humidity = (60.5-75.099638] ) Then (Result = Positive)	23 Record
IF (Destruction = (17.5-87.5], Height = (6.5-12] Then (Result = Positive)	201 Record

ซึ่งจากตารางที่ 74 พบว่า กฎที่มีความสามารถในการจำแนกข้อมูลมากที่สุด คือ กฎ IF (Destruction = (17.5-87.5], Height = (6.5-12] Then (Result = Positive) สามารถอธิบายได้ว่า

ในพื้นที่ที่มีการทำลายสัตว์ปีกเป็นจำนวนในช่วง 18 ถึง 88 ตัว และ มีความสูงจากระดับน้ำทะเลอยู่ระหว่าง 6.5 ถึง 12 เมตร แล้ว พื้นที่ที่มีลักษณะดังกล่าวมีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนก



ภาพที่ 9 แบบจำลองที่สร้างด้วย อัลกอริทึม Decision Tree (J48) ที่สร้างจากชุดข้อมูลผ่านการ Discretize การคัดเลือก Attribute และ การถ่วงน้ำหนัก (ในกรณีที่มีจำนวนสัตัวฝึกที่ถูกทำลาย อยู่ในช่วง 1 -18 ตัว)

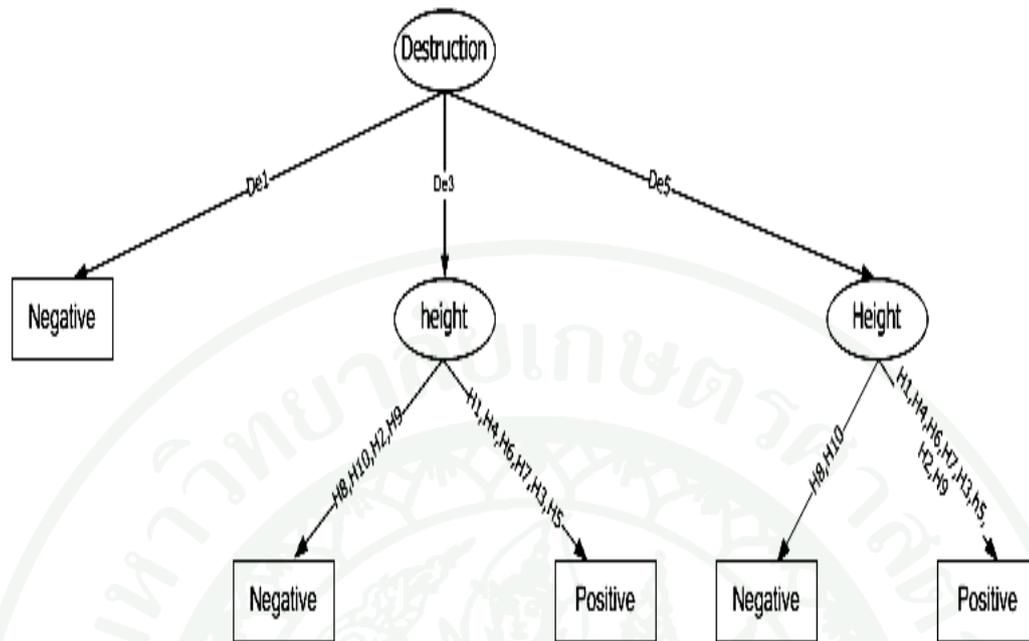
จากภาพที่ 9 แสดงให้เห็นถึงลักษณะของต้นไม้ตัดสินใจที่สร้างจากชุดข้อมูลที่ผ่านการ Discretize, การคัดเลือก Attribute และ การถ่วงน้ำหนักด้วยน้ำหนักที่มาจากวิธีการ Emerging pattern ซึ่งจากภาพที่ 9 สามารถ นำไปสร้างเป็นกฎในรูปแบบของ Production Rule ดังตัวอย่างตาม ตาราง ที่ 43

ตารางที่ 75 กฎที่ได้จากแบบจำลองที่สร้างจากต้นไม้ตัดสินใจตามภาพที่ 9

กฎ	จำนวนข้อมูล
If(Destruction = (0.5-17.5], Height = '(34.5-97], humidity = '(60.5-75.099638]) Then (Result = Positive)	138 Record
If(Destruction = (0.5-17.5], Height = (6.5-12] , humidity = '(60.5-75.099638]) Then (Result = Positive)	157 Record

ซึ่งจากตารางที่ 75 พบว่า กฎที่มีความสามารถในการจำแนกข้อมูลมากที่สุด คือ กฎ  
 If(Destruction = (0.5-17.5], Height = (6.5-12] , humidity = '(60.5-75.099638])  
 Then (Result = Positive)

อธิบายความหมายได้ว่า ในพื้นที่ที่มีการทำลายสัตว์ปีกในช่วง 1 ถึง 18 ตัว ในระดับความสูงจากระดับน้ำทะเล อยู่ระหว่าง 6.5 ถึง 12 เมตร ที่มีความชื้น อยู่ระหว่าง 60.5 ถึง 75.09 % พื้นที่นั้นมีความเสี่ยงที่จะเกิดโรคไข้หวัดนก



ภาพที่ 10 แบบจำลองที่สร้างขึ้นด้วย อัลกอริทึม Decision Tree (J48) ที่สร้างจากชุดข้อมูลที่ผ่านการ Discretize การคัดเลือก Attribute และ การถ่วงน้ำหนัก (ในกรณีที่ จำนวนสัตว์ปีกที่ถูกทำลาย อยู่ในช่วงน้อยกว่า 1 ตัว, ช่วง 88 – 969 ตัว และ ช่วง มากกว่า 969 ตัว

จากภาพที่ 10 แสดงให้เห็นถึงลักษณะของต้นไม้ตัดสินใจที่สร้างจากชุดข้อมูลที่ผ่านการ Discretize, การคัดเลือก Attribute และ การถ่วงน้ำหนักด้วยน้ำหนักที่มาจากวิธีการ Emerging pattern ซึ่งจากภาพที่ 10 นำไปสร้างเป็นกฎในรูปแบบของ Production Rule ดังตัวอย่างตามตารางที่

ตารางที่ 76 กฎที่ได้จากแบบจำลองที่สร้างจากต้นไม้ตัดสินใจตามภาพที่ 10

กฎ	จำนวนข้อมูล
IF(Destruction = '(87.5-968.5], Height = '(6.5-12]) Then (Result = Positive)	102.8 Record
IF(Destruction = '(968.5-inf), Height = '(6.5-12]) Then (Result = Positive)	327.78 Record

ซึ่งจากตารางที่ 76 พบว่า กฎที่มีความสามารถในการจำแนกข้อมูลมากที่สุด คือ กฎ IF(Destruction = '(968.5-inf), Height = '(6.5-12]) Then (Result = Positive) ซึ่งอธิบายความหมายได้ว่า ในพื้นที่ที่การทำลายสัตว์ปีกอยู่ในช่วง มากกว่า 968 ตัว ในความสูงจากระดับน้ำทะเล อยู่ระหว่าง 6.5 เมตร ถึง 12 เมตร พื้นที่นั้นมีความเสี่ยงที่จะเกิดโรคไข้หวัดนก

### 3. แบบจำลองที่สร้างด้วย Naïve Bayes Tree

นำชุดข้อมูลที่เตรียมไว้ทั้งหมดสิบชุด ซึ่งประกอบด้วยชุดข้อมูล A B C D E F G H I และ J ที่แต่ละชุดผ่านการเตรียมข้อมูลด้วยวิธีการที่แตกต่างกัน มาสร้างแบบจำลองด้วยวิธี Naïve Bayes Tree ซึ่งเป็นอัลกอริทึมด้านการจำแนกข้อมูลที่ทำงานร่วมกันระหว่าง อัลกอริทึมส์ ต้นไม้ตัดสินใจ และ Naïve Bayes หลังจากสร้างแบบจำลองแล้ว นำแบบจำลองที่ได้ไปวัดประสิทธิภาพซึ่งได้ผลตามตารางที่ 77 และ ตารางที่ 78

ตารางที่ 77 ประสิทธิภาพการจำแนกข้อมูล (Accuracy) ของแบบจำลองที่สร้างด้วย วิธีการ Naïve Bayes Tree (NB Tree)

Algorithm	Dataset	Correctly Classify(%)	Incorrect Classify(%)
NB Tree	A	88.50	11.50
NB Tree	B	88.50	11.50
NB Tree	C	90.39	9.61
NB Tree	D	89.95	10.05
NB Tree	E	88.79	11.21
NB Tree	F	89.14	10.86
NB Tree	G	84.07	15.93
NB Tree	H	94.96	5.04
<b>NB Tree</b>	<b>I</b>	<b>95.74</b>	<b>4.26</b>
NB Tree	J	85.45	14.55

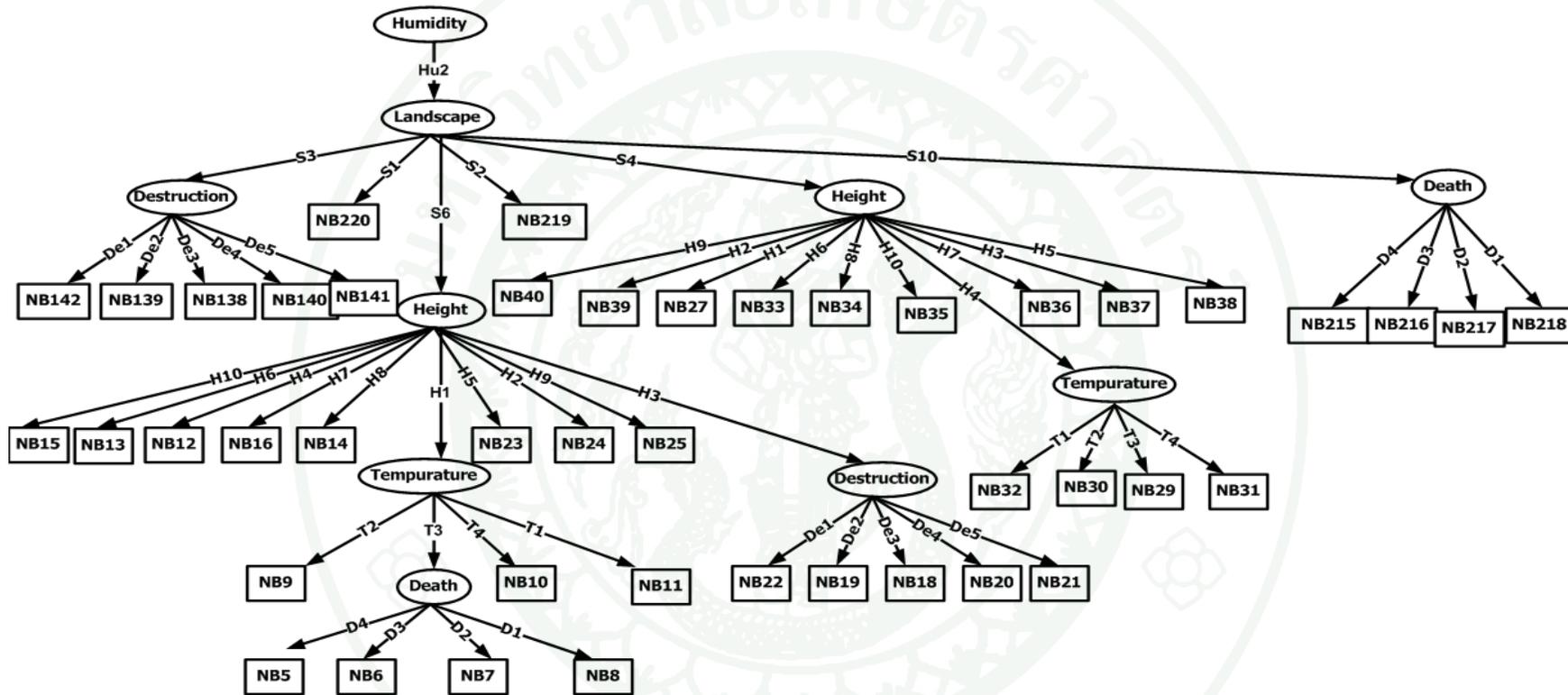
จากตารางที่ 77 แสดงให้เห็นถึงประสิทธิภาพการจำแนกข้อมูล (accuracy) ของแบบจำลองที่สร้างด้วย Naïve bayes Tree ซึ่งนำเอาอัลกอริทึมส์ Naïve bayes มารวมกับ อัลกอริทึมส์ Decision Tree เพื่อช่วยกันจำแนกข้อมูล ซึ่งแบบจำลองที่สร้างจากข้อมูลที่ผ่านการ Discretize, การถ่วงน้ำหนัก ซึ่งน้ำหนักคำนวณมาจากวิธี Emerging Pattern เป็นแบบจำลองที่มีประสิทธิภาพการจำแนกข้อมูลเท่ากับ 95.74% ซึ่งสูงกว่าแบบจำลองที่สร้างด้วย NB Tree จากชุดข้อมูลแบบอื่นๆ

ตารางที่ 78 ค่า Precision Recall และ F-measure ในการจำแนกข้อมูลของแบบจำลองที่สร้างจากวิธีการ Naïve Bayes Tree (NB Tree)

Algorithm	Dataset	Precision	Recall	F-measure	Class
NB Tree	A	0.679	0.581	0.626	positive
		0.919	0.945	0.932	Negative
NB Tree	B	0.673	0.597	0.633	positive
		0.922	0.942	0.932	Negative
NB Tree	C	0.76	0.582	0.659	Positive
		0.924	0.965	0.944	Negative
NB Tree	D	0.788	0.534	0.637	Positive
		0.914	0.972	0.942	Negative
NB Tree	E	0.685	0.6	0.639	Positive
		0.922	0.945	0.934	Negative
NB Tree	F	0.697	0.611	0.651	Positive
		0.925	0.947	0.936	Negative
NB Tree	G	0.544	0.205	0.298	Positive
		0.800	0.966	0.910	Negative
<b>NB Tree</b>	<b>H</b>	<b>0.911</b>	<b>0.945</b>	<b>0.928</b>	<b>Positive</b>
		<b>0.971</b>	<b>0.952</b>	<b>0.961</b>	<b>Negative</b>
NB Tree	I	0.863	0.88	0.872	Positive
		0.976	0.973	0.974	Negative
NB Tree	J	0.566	0.381	0.456	Positive
		0.889	0.945	0.916	Negative

จากตารางที่ 78 แสดงถึงค่าความแม่นยำของการจำแนกข้อมูลซึ่งวัดได้จากค่า Precision Recall และ F-measure ซึ่งจากตารางแสดงให้เห็นว่า แบบจำลองที่สร้างจากชุดข้อมูลที่ผ่านการ Discretize การคัดเลือก Attribute และการถ่วงน้ำหนัก ซึ่งน้ำหนักที่ใช้ในการถ่วงน้ำหนักมาจากวิธี Emerging Pattern มีค่า Precision Recall และ F-measure สูงกว่าแบบจำลอง NB Tree ที่สร้างจากชุดข้อมูลแบบอื่นๆ

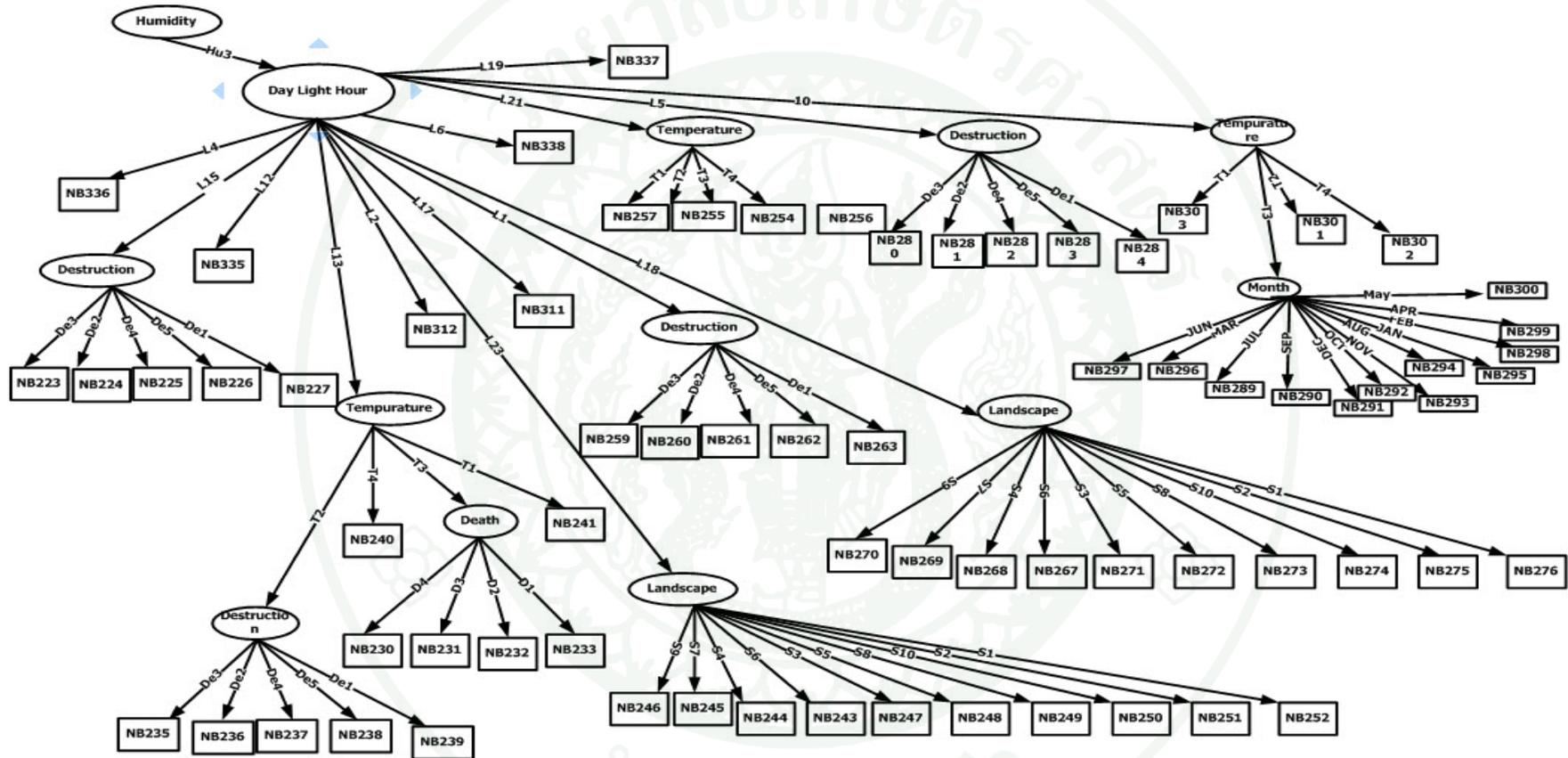
ลักษณะของแบบจำลองที่สร้างด้วย Naïve Bayes Tree ลักษณะโครงสร้างคล้ายกับแบบจำลองที่สร้างจาก Decision Tree แต่ในส่วนของ Leave Node จะเป็นตารางแสดงการจำแนกข้อมูลด้วย Naïve Bayes



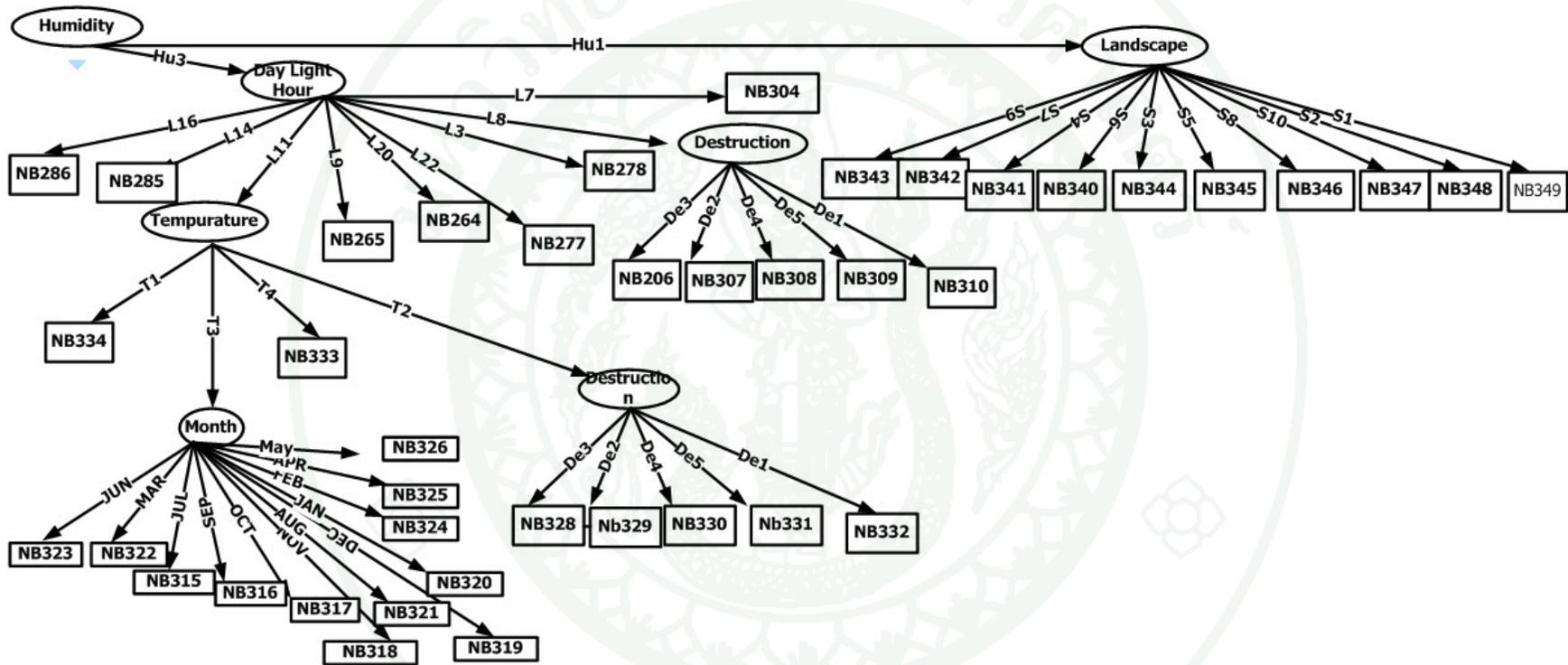
ภาพที่ 11 ลักษณะแบบจำลองของ โรคข้าวหัดนกที่สร้างด้วย Naive Bayes Tree เมื่อ ความชื้นสัมพัทธ์ในอากาศ อยู่ในช่วง 60.5% ถึง 75.09% และ อยู่ในลักษณะภูมิประเทศแบบ คาบสมุทรภาคใต้ชายฝั่งตะวันตก คาบสมุทรภาคใต้ชายฝั่งตะวันออก ชายฝั่งภาคตะวันออกเฉียงใต้ ทิวเขาและหุบเขาภาคตะวันตก และ แอ่งสกลนคร







ภาพที่ 14 ลักษณะของแบบจำลองใช้หัวคนทีสร้างด้วย Naive bayes Tree เมื่อ ความชื้นสัมพัทธ์ในอากาศ อยู่ในช่วงมากกว่า 75.09% (Hu3)



ภาพที่ 15 ลักษณะของแบบจำลองใช้หาคณกที่สร้างด้วย Naive Bayes Tree เมื่อ ความชื้นสัมพัทธ์ในอากาศ อยู่ในช่วง น้อยกว่า 60.5% (Hu1) หรือ ความชื้นสัมพัทธ์ในอากาศอยู่ในช่วง มากกว่า 75.09 (Hu3)

จากรูปภาพที่ 11,12,13,14 และ 15 แสดงให้เห็นถึงลักษณะของแบบจำลองที่สร้างด้วย Naïve Bayes Tree ซึ่งมีลักษณะในส่วนที่เป็น Parent node เหมือนกับ Decision Tree แต่ในส่วนของ Leave Node จะไม่แสดงส่วนของ Class ของข้อมูล แต่จะแสดงออกมาเป็นจำนวนข้อมูลที่ถูกรวบรวมด้วย Naïve Bayes ตัวอย่างเช่น ในภาพที่ 11 เห็นได้ว่า โหนดรากของแบบจำลองคือ Humidity ส่วนที่เป็น Parent Node ประกอบด้วย โหนด Day Light Hours, โหนด Destruction, โหนด Landscape, โหนด Temperature และ โหนด Month ส่วนที่เป็น Leave Node เป็นส่วนที่เป็น การจำแนกข้อมูลตาม Class ของข้อมูล ซึ่งจำแนกข้อมูลโดยใช้ Naïve Bayes Classifier ซึ่งให้ผลออกมาคือจำนวนข้อมูลแยกตาม Class ในภาพจะแทนด้วยสัญลักษณ์ NB ซึ่งหมายถึงการจำแนก ข้อมูลด้วย Naïve Bayes Classifier และใน 1 Leave Node มี Naïve Bayes Classifier ได้เพียงแค่ 1 Naïve Bayes Classifier เท่านั้น ซึ่งจากแบบจำลอง มี Leave Node ทั้งหมด 300 Leave node จะมี Naïve Bayes Classifier ทั้งหมด 300 ตัว

ดังนั้นเมื่อได้ Naïve Bayes Classifier จากแบบจำลอง Naïve bayes Tree แล้วจึงนำเอาแต่ละ Leave Node ของ Naïve Bayes Tree ที่มีทั้งหมด 300 Leave Node มาหาร้อยละของโอกาสที่แต่ละกิ่งมีผลเฉลยเป็น ค่า Positive ซึ่งในการคำนวณนั้นใช้วิธีการ ของ Naïve Bayes Classifier ในการหาร้อยละของโอกาส ซึ่งผลจากการหาร้อยละของ โอกาส เป็นดังตารางที่ 79, 80, 81

**ตารางที่ 79** ร้อยละของ โอกาสที่กิ่งของ NB Tree จะมีผลเฉลยเป็น Positive เมื่อความชื้นสัมพัทธ์ อยู่ในช่วง น้อยกว่า 60.5% (Hu1)

รายละเอียดของแต่ละกิ่งของ Naïve bayes Tree	ร้อยละของโอกาสที่มีผลเฉลยเป็น Positive
Hu1, S3	56.750
Hu1, S6	65.670
Hu1, S8	72.630

**หมายเหตุ** Hu1, Hu2, Hu3 แทนช่วงข้อมูลที่อยู่ใน Attribute Humidity ซึ่งปรากฏในตารางที่ 10 S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11

จากตารางที่ 79 กิ่งของ Naïve Bayes Tree ที่แสดงให้เห็นถึงลักษณะของพื้นที่ที่เลี้ยงสัตว์ ปีกของไทยที่ผลเฉลยของข้อมูลมีโอกาสที่เป็น Positive มากที่สุดคือ กิ่งใน Naïve Bayes Tree ที่ แสดงถึงลักษณะพื้นที่ ที่มี ความชื้นสัมพัทธ์ น้อยกว่า 60.5 % และ มีลักษณะพื้นที่ที่เป็นแบบขอบ ที่ราบทางภาคกลาง

ตารางที่ 80 ร้อยละของโอกาสที่กิ่งของ Naïve Bayes Tree จะมีผลเฉลยเป็น Positive เมื่อ ความชื้นสัมพัทธ์อยู่ในช่วง 60.5 % ถึง 75.09% (Hu2)

รายละเอียดของแต่ละกิ่งของ Naïve Bayes Tree	ร้อยละของโอกาสที่มีผลเฉลยเป็น Positive
Hu2, S6, H1, T3, D1	55.566
Hu2, S10, D4	67.080
Hu2, S8, H4	69.637
Hu2, S7, T3, H5, JUL	80.877
Hu2, S6, H1, T3, D2	87.693
Hu2, S7, T2, NOV, DE4	94.480
Hu2, S8, H9	94.522
Hu2, S7, T2, NOV, DE5	95.070
Hu2, S5, T3, H6,	95.362
Hu2, S7, T2, JAN, De5	96.233
Hu2, S3, de2	96.802
Hu2, S5, T2, H6	97.049
Hu2, S3, de5	97.734
Hu2, S3, de3	97.823
Hu2, S7, T3, H5, SEP	98.010
Hu2, S7, T2, DEC	98.661

## ตารางที่ 80 (ต่อ)

รายละเอียดของแต่ละกิ่งของ Naïve Bayes Tree	ร้อยละของโอกาสที่มีผลเฉลยเป็น Positive
Hu2, S7, T3, H5, De3	99.897
Hu2, S8, H3, T2, De5, OCT	99.906
Hu2, S7, T3, H5, OCT, De5	99.981
Hu2, S7, T3, H5, OCT, De2	99.984
Hu2, S6, H3, De5	99.985
Hu2, S8, H3, T3, D2, De2	99.988
Hu2, S7, T3, H5, OCT, De4	99.987
Hu2, S8, H3, T2, De5, DEC	99.988
Hu2, S6, H3, De3	99.988
Hu2, S6, H4	99.991
Hu2, S7, T3, H5, NOV	99.991
Hu2, S8, H3, T3, D2, De5	99.995
Hu2, S8, H3, T3, D2, De3	99.996
<b>Hu2, S8, H3, T3, D4</b>	<b>99.999</b>

หมายเหตุ Hu1, Hu2, Hu3 แทนช่วงข้อมูลที่อยู่ใน Attribute Humidity ซึ่งปรากฏในตารางที่ 10  
 S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11  
 H1, H2, H3, ..., H10 แทนช่วงข้อมูลที่อยู่ใน Attribute Height ปรากฏอยู่ในตารางที่ 10  
 T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ปรากฏอยู่ในตารางที่ 10  
 De1, De2, De3, De4, De5 แทนช่วงข้อมูลที่อยู่ใน Attribute Destruction ซึ่งปรากฏใน  
 ตารางที่ 10  
 D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
 L1, L2, L3, ..., L23 แทนช่วงข้อมูลที่อยู่ใน Attribute Day Light Hour ซึ่งปรากฏใน  
 ตารางที่ 10

และจากตารางที่ 80 ลักษณะพื้นที่เลี้ยงสัตว์ปีกที่มีโอกาสมากที่สุดที่จะเกิดโรคไข้หวัดนก คือลักษณะพื้นที่ที่มีความชื้นอยู่ที่ 60.5%-75.09% ลักษณะของภูมิประเทศ ตั้งอยู่บริเวณ ขอบที่ราบ ภาคกลาง โดยที่มีความสูงจากระดับน้ำทะเลอยู่ในช่วง 6.5 เมตร ถึง 12 เมตร และอุณหภูมิอยู่ที่ 27.85 ถึง 32.05 เซลเซียส และจำนวนสัตว์ปีกป่วยตาย มากกว่า 87 ตัว ซึ่งลักษณะพื้นที่ดังกล่าวมีโอกาสเกิดโรคไข้หวัดนกสูงถึง 99.99 %

ตารางที่ 81 ร้อยละของ โอกาสที่ละกิ่งของ NB Tree จะมีผลเฉลยเป็น Positiveเมื่อ ความชื้นสัมพัทธ์ อยู่ในช่วง 60.5% - 75.09% (Hu3)

รายละเอียดของแต่ละกิ่งของ Naïve Bayes Tree	ร้อยละของ โอกาสที่มีผลเฉลยเป็น Positive
Hu3, L11, T2, De5	83.960
Hu3, L8, De5	84.436
Hu3, L13, T2, De5	91.504
Hu3, L15, De5	93.436
Hu3, L10, T3, OCT	94.100
Hu3, L15, De4	94.189
Hu3, L5, De5	95.703
Hu3, L7	97.827
Hu3, L9	98.373
<b>Hu3, L1, De5</b>	<b>99.193</b>

หมายเหตุ Hu1, Hu2, Hu3 แทนช่วงข้อมูลที่อยู่ใน Attribute Humidity ซึ่งปรากฏในตารางที่ 10  
 S1, S2, S3, ..., S10 แทนค่าที่อยู่ใน Attribute Landscape ปรากฏอยู่ในตารางที่ 11  
 H1, H2, H3, ..., H10 แทนช่วงข้อมูลที่อยู่ใน Attribute Height ปรากฏอยู่ในตารางที่ 10  
 T1, T2, T3, T4 แทนช่วงข้อมูลที่อยู่ใน Attribute Temperature ปรากฏอยู่ในตารางที่ 10  
 De1, De2, De3, De4, De5 แทนช่วงข้อมูลที่อยู่ใน Attribute Destruction ซึ่งปรากฏใน ตารางที่ 10  
 D1, D2, D3, D4 แทนช่วงข้อมูลที่อยู่ใน Attribute Death ซึ่งปรากฏในตารางที่ 10  
 L1, L2, L3, ..., L23 แทนช่วงข้อมูลที่อยู่ใน Attribute Day Light Hour ซึ่งปรากฏใน ตารางที่ 10

จากตารางที่ 81 กิ่งของ Naïve Bayes Tree ที่แสดงให้เห็นถึงลักษณะของพื้นที่ที่เลี้ยงสัตว์ปีกของไทยที่ผลเฉลยของข้อมูลมีโอกาสที่เป็น Positive มากที่สุดคือ กิ่งใน Naïve Bayes Tree ที่แสดงถึงลักษณะพื้นที่ ที่มี ความชื้นสัมพัทธ์มากกว่า 75.09 % (Hu3) และ จำนวนชั่วโมงที่มีแสงแดดอยู่ในช่วงน้อยกว่า 2.36 ชั่วโมง และมีจำนวนสัตว์ปีกที่ถูกทำลาย มากกว่า 969 ตัว

ลักษณะต้นไม้ที่สร้างจาก Naïve Bayes Tree สามารถนำมาสร้างเป็น Production Rule ซึ่งแสดงให้เห็นถึงความเป็นของแต่ละกิ่งของต้นไม้ตัดสินใจได้ เนื่องจากในระดับ Leave Node ของต้นไม้ตัดสินใจที่สร้างจาก NBTree มีการจำแนกข้อมูลด้วย Naïve Bayes Classifier อยู่ซึ่งทำให้ทราบว่าแต่ละกิ่งของต้นไม้ตัดสินใจมีความน่าจะเป็นที่จะมีผลเฉลยเป็นบวก (Positive) หรือ เป็นลบ (Negative) เป็นเท่าไร

จากตารางผลการทดลองที่แสดงถึงประสิทธิภาพการจำแนกข้อมูลของแต่ละแบบจำลองที่ทำการสร้างขึ้น โดยใช้อัลกอริทึมทางการจำแนกข้อมูล เมื่อนำประสิทธิภาพของแต่ละแบบจำลองที่นำมาเปรียบเทียบประสิทธิภาพ ตามตารางที่ 82 ซึ่งแสดงเรื่องประสิทธิภาพของความถูกต้องในการจำแนกข้อมูล (accuracy) ของแต่ละแบบจำลองที่ผ่านการคัดเลือก

**ตารางที่ 82** เปรียบเทียบประสิทธิภาพของความถูกต้องในการจำแนกข้อมูล (accuracy) ของแบบจำลองที่สร้างขึ้นด้วยอัลกอริทึมการจำแนกข้อมูลทั้ง 3 แบบ

Algorithm	ชุดข้อมูล	Correctly Classify (%)	Incorrectly Classify (%)
Bayes Net	H	95.98	4.02
Decision Tree (J48)	H	94.14	5.86
NB Tree	H	94.96	5.04
Bayes Net	I	<b>96.47</b>	3.53
Decision Tree (J48)	I	95.19	4.81
NB Tree	I	95.74	4.26

จากตารางที่ 82 เห็นได้ว่า แบบจำลองที่สร้างจากอัลกอริทึมส์ ข่ายงานแบบเบย์ส์ (Bayesian Network) เป็นแบบจำลองที่มีประสิทธิภาพการจำแนกข้อมูล (accuracy) สูงกว่าแบบจำลองชนิดอื่นๆ โดยที่มีประสิทธิภาพการจำแนกข้อมูล (accuracy) เท่ากับ 96.47 % และ ชุดข้อมูลที่จะทำให้แบบจำลองที่สร้างขึ้นมีประสิทธิภาพในการจำแนกข้อมูลที่ดีนั้น คือชุดข้อมูลที่มีการ ทำ Discretize Data การคัดเลือก Attribute และ การถ่วงน้ำหนักที่หามาจากวิธี Emerging Pattern

ตารางที่ 83 ค่า Precision Recall และ F-measure ของแต่ละแบบจำลองที่แบบจำลองที่สร้างขึ้นด้วยอัลกอริทึมส์การจำแนกข้อมูลทั้ง 3 แบบ

อัลกอริทึม	ชุดข้อมูล	Precision	Recall	F-measure	Class
Bayes net	H	<b>0.943</b>	<b>0.939</b>	<b>0.941</b>	<b>Positive</b>
		<b>0.968</b>	<b>0.971</b>	<b>0.969</b>	<b>Negative</b>
Bayes net	I	0.891	0.895	0.893	Positive
		0.979	0.979	0.979	Negative
J48	H	0.94	0.885	0.912	Positive
		0.942	0.971	0.956	Negative
J48	I	0.904	0.791	0.884	Positive
		0.96	0.983	0.972	Negative
NB Tree	H	0.911	0.945	0.928	Positive
		0.971	0.952	0.961	Negative
NB Tree	I	0.863	0.88	0.872	Positive
		0.976	0.973	0.974	Negative

จากตารางที่ 83 พบว่า แบบจำลองที่มีค่า Precision Recall และ F-measure ของ Class ที่เป็นบวก มากที่สุดคือแบบจำลอง Bayesian Network โดยที่มีค่า Precision ของ Class ที่เป็นบวก (Positive) เท่ากับ 0.943 Recall เท่ากับ 0.939 และ ค่า F-measure เท่ากับ 0.941 และชุดข้อมูลที่ทำให้แบบจำลองมีค่า Precision, Recall และ F-measure ที่ดีคือ ชุดข้อมูลที่ผ่านการ Discretize, การคัดเลือก Attribute และ การถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ Emerging pattern

สาเหตุที่นำค่า Precision Recall และ F-measure ของ Class Positive เป็นเกณฑ์ในการคัดเลือกแบบจำลองเพราะว่าด้วยข้อมูลโรคไข้หวัดนกที่นำมาสร้างแบบจำลอง มีข้อมูลสมาชิกที่เป็น Class Negative (83.42%) มากกว่า Class Positive (16.38 %) เพราะฉะนั้น จึงทำให้ค่า Precision ของ Class Negative ดีกว่า class Positive ในทุกแบบจำลอง และด้วยเป้าหมายของงานวิจัยนี้ที่ต้องการนำ แบบจำลองไปเฝ้าระวังโรคไข้หวัดนก ดังนั้น แบบจำลองที่นำมาใช้ในการเฝ้าระวังจึงต้องมีความแม่นยำในการทำนายว่าพื้นที่ที่เป็นพื้นที่ที่เสี่ยงสัตว์ปีกเป็นพื้นที่ที่มีความเสี่ยงที่จะเกิดโรคไข้หวัดนก มากกว่า ความแม่นยำในการทำนายว่าพื้นที่ที่เป็นพื้นที่ที่เสี่ยงสัตว์ปีกเป็นพื้นที่ที่ไม่เกิดโรคไข้หวัดนก



## วิจารณ์

### 1. การสร้างแบบจำลองโดยวิธีการจำแนกข้อมูล

เทคนิคการทำเหมืองข้อมูล ด้านการจำแนกข้อมูลที่นำมาใช้ในการวิจัยนี้ประกอบด้วย Bayesian Network Decision Tree Naïve Bayes Tree ซึ่งผลลัพธ์ที่ได้จากการนำเทคนิคเหล่านี้มาใช้คือ แบบจำลองที่แสดงถึงความสัมพันธ์ระหว่าง Attribute ที่อยู่ในชุดข้อมูล ประกอบด้วย ลักษณะของพื้นที่ที่เลี้ยงสัตว์ปีกของไทย อันได้แก่ ความสูงของพื้นที่ ลักษณะของพื้นที่ อุณหภูมิ จำนวนชั่วโมงที่มีแสงแดด และ ความชื้น อันแสดงให้เห็นว่า ปัจจัยทางด้านสภาพแวดล้อมส่งผลต่อโอกาสที่จะเกิดการระบาดของไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีก ซึ่งคล้ายกับงานวิจัย (Li *et al.*, 2004) ที่นำลักษณะพื้นที่มาใช้ในการคำนวณหาความน่าจะเป็นของพื้นที่ที่จะเกิดการระบาดของโรคไข้หวัดนก แต่มีจุดที่แตกต่างตรง ลักษณะพื้นที่ที่นำมาใช้ในการคำนวณหาความน่าจะเป็น ได้แก่ อุณหภูมิ ระยะทางของจุดที่เกิดการระบาดกับพื้นที่ที่ต้องการหาความน่าจะเป็น ช่วงระยะเวลาที่เกิดการระบาด เป็นต้น แต่สำหรับแบบจำลองที่งานวิจัยนี้นำเสนอ ประกอบด้วย อุณหภูมิ ลักษณะของพื้นที่ ความสูงของพื้นที่ ความชื้น และ จำนวนชั่วโมงที่มีแสงแดด ซึ่งแตกต่างกันเพราะ สมมุติฐานที่นำมาใช้ในการสร้างแบบจำลองที่แตกต่างกัน

แบบจำลองที่งานวิจัยนี้แนะนำเสนอคือแบบจำลองที่สร้างจาก Bayesian Network ซึ่งคล้ายกับงานวิจัย (Cooper *et al.*, 2004) , งานวิจัย (Sebastiani *et al.*, 2005) ซึ่งแตกต่างกับงานวิจัย (Li *et al.*, 2004), งานวิจัย (Praphaphan and Anongnart, 2007), งานวิจัย (Upadhyay, 2007)

## 2. การทำนายการระบาดของโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีก

เทคนิคการทำเหมืองข้อมูลที่นำมาใช้ในการทำวิจัยในครั้งนี้ทุกตัวให้ผลลัพธ์คือแบบจำลองที่สามารถนำไปทำการสกัดเป็นกฎเพื่อใช้ในการทำนายลักษณะของพื้นที่ที่เลี้ยงสัตว์ปีกของไทยที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนก โดยจากผลการทดลองเมื่อเราทำการสร้างกฎขึ้นมาจากแบบจำลองที่ได้ทำให้เราทราบว่า ลักษณะพื้นที่ที่เลี้ยงสัตว์ปีกของไทยที่มีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนก เป็นพื้นที่ที่มีความสูงจากระดับน้ำทะเล 6.5 -12 เมตรในลักษณะพื้นที่ที่เป็นที่ราบลุ่มดินดอนสามเหลี่ยมปากแม่น้ำ ความชื้นอยู่ที่ 60.5% ถึง 75.09 % และอุณหภูมิที่เหมาะสมคือ อยู่ในช่วง 27.85 องศาเซลเซียส ถึง 32.05 องศาเซลเซียส ซึ่งสอดคล้องกับงานวิจัยของ วิระพงษ์ (2550) รายงานว่าลักษณะพื้นที่ที่มีความถี่ในการเกิดโรคไข้หวัดนกคือ

พื้นที่ภาคกลาง และช่วงฤดูที่มีความถี่ในการเกิดการระบาดสูงที่สุดคือ ฤดูฝน และ ฤดูหนาว ซึ่งทั้งสองฤดูมีอุณหภูมิอยู่ในช่วง 27 องศาเซลเซียส ถึง 30 องศาเซลเซียส

## 3. การเปรียบเทียบประสิทธิภาพของแบบจำลองที่สร้างจากอัลกอริทึมทางการทำเหมืองข้อมูลที่แตกต่างกัน

จากตารางผลการทดลองที่ 82 และ 83 จะเห็นได้ว่าทุกแบบจำลองจะมีประสิทธิภาพการจำแนกข้อมูล (accuracy) ที่ใกล้เคียงกัน ความแม่นยำในการจำแนก ข้อมูล ซึ่งสามารถวัดได้จาก ค่า Precision Recall และ F-measure มีความใกล้เคียงกัน จากผลการทดลองทำให้ทราบว่า การถ่วงน้ำหนักข้อมูลด้วยค่าน้ำหนัก ที่หามาด้วยวิธี Emerging Pattern เป็นวิธีที่ทำให้แบบจำลองมีประสิทธิภาพในการจำแนกข้อมูลดีขึ้น ซึ่งเป็นไปตาม งานวิจัย (Alhammady *et al.*, 2006) ซึ่งนำน้ำหนัก ไปใช้กับ Decision Tree แต่ในงานวิจัยนี้นำน้ำหนักไปใช้กับ Bayesian Network ซึ่งทำให้ Bayesian Network มีประสิทธิภาพในการจำแนกข้อมูลสูงกว่าแบบจำลองอื่นๆ

จากตารางผลการทดลอง แสดงให้เห็นความแตกต่างระหว่างประสิทธิภาพของแบบจำลองที่สร้างจากชุดข้อมูล ที่ผ่านการถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธี Emerging Pattern และ ถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธี DEMATEL พบว่า ประสิทธิภาพของแบบจำลองที่สร้างมาจากชุดข้อมูลที่ถ่วงน้ำหนักที่หามาจากวิธี Emerging Pattern ดีกว่า แบบจำลองที่สร้างมาจากชุดข้อมูลที่ถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธี DEMATEL ซึ่งสาเหตุเป็นเพราะว่า น้ำหนักที่หามาจากวิธีการหลังสุด เป็นน้ำหนักที่แทนระดับความสัมพันธ์ระหว่าง Attribute เท่านั้น ไม่ได้แทนค่าระดับความสัมพันธ์ในการจำแนกระหว่าง Attribute กับ Class ของข้อมูล การคำนวณน้ำหนักโดยวิธีนี้ ในการจำแนกข้อมูลจึงให้ผลในการเพิ่มประสิทธิภาพการจำแนกข้อมูลของแบบจำลองไม่ชัดเจน แต่ น้ำหนักที่คำนวณมาจากวิธีการ Emerging Pattern นั้น ได้มาจาก Item set ที่ปรากฏอยู่ในชุดข้อมูล และในขั้นตอนการหาน้ำหนักมีการนำ Item set ที่ได้ไปหาค่าความเป็นสมาชิกอยู่ใน คลาสของข้อมูล Positive หรือ Negative ดังนั้นถ้า Item set นั้นมีค่าความเป็นสมาชิก อยู่ใน คลาสที่เป็น Positive มากกว่า คลาสที่เป็น Negative หมายความว่า Item set ซึ่งประกอบด้วย ค่าของ Attribute ต่างๆที่อยู่ในชุดข้อมูลนั้นมีความสัมพันธ์กับผลเฉลยของข้อมูลที่เป็น Positive มากกว่า Negative

ดังนั้น ถ้าใน Instance ที่อยู่ในชุดข้อมูล ปรากฏ Itemset ที่มีค่าความเป็นสมาชิกอยู่ในคลาที่เป็น Positive ดังนั้น Instance ดังกล่าวจึงมีโอกาสที่เป็น Positive คลาส มากกว่า Negative คลาส ซึ่งเมื่อทำการจำแนกข้อมูลด้วยแบบจำลองด้านการจำแนกข้อมูลทำให้สามารถจำแนกได้อย่างถูกต้องว่า Instance ดังกล่าวอยู่ในคลาที่เป็น Positive

เมื่อนำไปเปรียบเทียบกับแบบจำลองอื่นๆที่ได้มีการสร้างขึ้นมาพบว่า แบบจำลองที่งานวิจัยนี้เป็นแบบจำลองที่นำเสนอให้เห็นถึงการทำนายลักษณะพื้นที่ที่มีความเสี่ยงที่เกิดการระบาดของโรคไข้หวัดนก คล้ายงานวิจัย (Cooper, 2004) ที่สร้างแบบจำลองเฝ้าระวังโรค Antrax ที่เกิดจากการก่อการร้าย โดยใช้ Bayesian Network จุดที่แตกต่างระหว่างแบบจำลองคือ การนำข้อมูลทิศทางลมมาใช้ในการเฝ้าระวังการเกิดโรค Antrax สาเหตุเพราะ ทิศทางลมมีผลต่อการพัดพาเชื้อ Antrax จากพื้นที่พบเชื้อไปยังพื้นที่อื่นๆที่อยู่ในแนวทิศทางลม ทำให้ ผู้คนที่อยู่ในพื้นที่เหล่านั้นเกิดความเสี่ยงที่จะเกิดการติดเชื้อ Antrax

## สรุปและข้อเสนอแนะ

### สรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองของการเฝ้าระวังการระบาดของโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีก ในประเทศไทย ที่สร้างจากวิธีการทำเหมืองข้อมูลโดยการเปรียบเทียบประสิทธิภาพในการจำแนกข้อมูล อัลกอริทึมที่นำมาใช้การสร้างแบบจำลอง ซึ่งประกอบด้วย ต้นไม้ตัดสินใจ (J48) ข่ายงานแบบเบย์ส์ และ Naïve Bayes Tree ส่วนข้อมูลที่นำมาใช้ในการสร้างแบบจำลองนั้น มาจากรายงานสัตว์ปีกป่วยตาย ของกรมปศุสัตว์, ข้อมูลสภาพภูมิอากาศอันได้แก่ ข้อมูลอุณหภูมิของแต่ละพื้นที่ ข้อมูลความชื้นสัมพัทธ์ และ ข้อมูลจำนวนชั่วโมงที่มีแสงแดด ข้อมูลสภาพภูมิประเทศได้แก่ ข้อมูลความสูงจากระดับน้ำทะเลของแต่ละพื้นที่เป็นต้น

ในขั้นตอนการสร้างแบบจำลอง นอกจากการสร้างแบบจำลองด้วยอัลกอริทึมที่แตกต่างกัน การเตรียมข้อมูลเพื่อที่จะนำมาใช้ในการสร้างแบบจำลองก็แตกต่างกัน โดยแบ่งเป็น ชุดข้อมูลที่มีการแบ่งช่วงข้อมูล กับข้อมูลที่ไม่มีการแบ่งช่วงข้อมูล ชุดข้อมูลที่มีการคัดเลือกคุณลักษณะกับ ชุดข้อมูลที่ไม่ผ่านการคัดเลือกคุณลักษณะ และ ชุดข้อมูลที่ไม่มีการถ่วงน้ำหนัก กับ ชุดข้อมูลที่มีการถ่วงน้ำหนักข้อมูล โดยน้ำหนักที่นำมาใช้ในการถ่วงนั้น คำนวณจากวิธีการที่แตกต่างกัน 2 วิธี ได้แก่วิธีการ Decision Making Trial and Evaluation Laboratory (DEMATEL) และ วิธีการ Emerging Pattern (EP) สาเหตุที่ต้องมีหลายชุดข้อมูลที่นำมาใช้ในการสร้างแบบจำลอง เนื่องจากต้องการเปรียบเทียบประสิทธิภาพของแบบจำลองที่สร้างจากชุดข้อมูลที่มีวิธีการเตรียมข้อมูลที่แตกต่างกัน

ในการวัดประสิทธิภาพของแบบจำลองที่สร้างจากวิธีการทำเหมืองข้อมูลโดยใช้อัลกอริทึมการจำแนกข้อมูลที่แตกต่างกัน ใช้การวัดจากเปอร์เซ็นต์การจำแนกข้อมูลที่ถูกต้อง (Accuracy) และค่า Precision, Recall และ F-measure เพื่อวัดความสามารถในการจำแนกข้อมูล

จากผลการทดลองแสดงว่าจากแบบจำลองทั้งหมดที่ได้สร้างมานั้น แบบจำลองที่มีประสิทธิภาพในการจำแนกข้อมูลที่ได้รับการคัดเลือก คือ แบบจำลอง Bayesian Network ที่สร้างจากชุดข้อมูลที่ผ่านการ Discretize การคัดเลือก Attribute และ การถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธี Emerging Pattern มาใช้ในการเฝ้าระวังโรคไข้หวัดนก เพราะถึงแม้ว่าแบบจำลองจะมีประสิทธิภาพการจำแนกข้อมูล (Accuracy) น้อยกว่าแบบจำลอง Bayesian Network ที่สร้างจากชุดข้อมูลที่ผ่านการ Discretize การถ่วงน้ำหนักด้วยน้ำหนักที่หามาจากวิธีการ Emerging Pattern และไม่ผ่านการคัดเลือก Attribute แต่ค่า Precision Recall และ F-measure ของการจำแนกว่าเป็นคลาสบวก (Class Positive) สูงสุด เทียบกับทุกแบบจำลอง ซึ่งประสิทธิภาพในการจำแนกว่าเป็นคลาสที่เป็นบวกมีความจำเป็นต่อการเลือกแบบจำลองเพราะว่า การทำนายว่าพื้นที่ที่เสี่ยงสัตว์ปีกแบบใดมีความเสี่ยงที่จะเกิดการระบาดของโรคไข้หวัดนกเป็นเรื่องที่ละเอียดอ่อน จึงต้องเลือกแบบจำลองที่มีค่า Precision และ ค่า F-measure ในการทำนายว่าเป็นคลาสบวกที่สูงมาใช้งานในการเฝ้าระวังโรคไข้หวัดนกเป็นกรณีหลักในการตัดสินใจ

จากผลการทดลองสามารถสรุปได้ว่า แบบจำลองส่วนมากที่สร้างจากชุดข้อมูลที่ไม่มีการถ่วงน้ำหนัก มีประสิทธิภาพในการจำแนกข้อมูลน้อยกว่าแบบจำลองที่ใช้ชุดข้อมูลที่มีการถ่วงน้ำหนัก ยกเว้นกรณีชุดข้อมูลที่ใช้น้ำหนักที่หามาจากวิธี DEMATEL ที่มีประสิทธิภาพที่แย่กว่าแบบจำลองที่สร้างจากชุดข้อมูลที่ไม่ได้ถ่วงน้ำหนัก

ด้วยแบบจำลองนี้ทำให้ทราบว่า คุณลักษณะที่ทำให้พื้นที่เสี่ยงสัตว์ปีกมีโอกาสเกิดโรคไข้หวัดนกสูงสุด เป็นพื้นที่ที่พบสัตว์ปีกตายอยู่ในช่วง 18-98 ตัว และ อุณหภูมิอยู่ในช่วง 27.85 - 32.05 เซลเซียส และ พื้นที่ที่มีความสูงจากระดับน้ำทะเลอยู่ระหว่าง 6.5 ถึง 12 เมตร เป็นพื้นที่ราบลุ่มดินดอนสามเหลี่ยมปากแม่น้ำ เป็นต้น

### ข้อเสนอแนะ

จากการทำงานวิจัยนี้ทำให้พบปัญหาขึ้นต่างมากมายซึ่งสามารถแบ่งได้เป็น 2 ประเภท ได้แก่ ปัญหาทางด้านเทคนิค และ ปัญหาทางการจัดการ ซึ่งสำหรับปัญหาทางด้านเทคนิค เช่น ปัญหาที่ WEKA ไม่สามารถรองรับข้อมูลจำนวนมหาศาลได้เพราะ HEAP Size ที่ใช้ในโปรแกรมไม่เพียงพอ ซึ่งทำการแก้ไขโดยการ ปรับขนาด Heap ให้มีขนาดที่ใหญ่ขึ้นเพียงพอต่อการทำงานของโปรแกรม ปัญหาด้านข้อมูลที่มีลักษณะไม่สมบูรณ์ และ มีความซ้ำซ้อนกัน ซึ่งได้ทำการแก้ไขโดยการ ตัดข้อมูลที่ซ้ำซ้อนกัน และ ทำการ Cleansing ข้อมูล ก่อนนำไปใช้ในงานวิจัย เป็นต้น

สำหรับปัญหาทางด้าน การจัดการของการทำวิจัยได้แก่ การบริหารเวลาในการทำวิจัย ซึ่งปัญหานี้ถือเป็นปัญหาที่สำคัญมากต่อความสำเร็จ ของการทำงานวิจัย ดังนั้นจึงขอเสนอให้ผู้ทำการวิจัยต่อจากนี้ควรใส่ใจต่อปัญหานี้ มากที่สุด

แนวทางในการพัฒนางานวิจัยต่อไปในอนาคต คือทำการปรับปรุงงานวิจัยนี้ให้สามารถนำผลที่ได้ไปใช้งานในการเฝ้าระวังไข้หวัดนกในพื้นที่จริงได้อย่างมีประสิทธิภาพ โดยที่เริ่มต้นจากการในการคัดเลือก คุณลักษณะ เปลี่ยนจากการใช้ อัลกอริทึมในการคัดเลือก คุณลักษณะ เปลี่ยนเป็น การคัดเลือกโดยใช้ ผู้เชี่ยวชาญในด้าน โรคไข้หวัดนก ผู้คัดเลือก ซึ่งน่าจะทำได้แบบจำลอง และ กฎที่สามารถนำไปใช้ในการเฝ้าระวังของโรคไข้หวัดนกได้อย่างดี หรือ นำแบบจำลองที่ได้แนะนำนี้ไปใช้ในการทำนายความรุนแรงของสถานการณ์การระบาดของไข้หวัดนกในอนาคตได้ หรือ นำไปพัฒนาแบบจำลองดังกล่าวเพื่อใช้ในการทำนายโอกาสที่พื้นที่ในประเทศไทยจะเกิดการระบาดของโรคไข้หวัดนกจากคนสู่คน หรือพัฒนาแบบจำลองนี้เพื่อใช้ในการทำนายพื้นที่ในประเทศไทยที่จะเกิดการระบาดของโรคไข้หวัดนกซ้ำซ้อนขึ้น หรือนำแนวทางและวิธีการในการสร้างแบบจำลองที่ปรากฏในงานวิจัยนี้ไปใช้ในการสร้างแบบจำลองของโรคระบาดอื่นๆที่เกิดขึ้นในประเทศไทย เช่น โรคไข้หวัดใหญ่ 2009 เป็นต้น

## เอกสารและสิ่งอ้างอิง

- สุคนธ์ทิพย์ วงษ์พันธ์. 2551. การเปรียบเทียบเทคนิคการคัดเลือกคุณลักษณะที่เหมาะสมและอัลกอริทึมเพื่อจำแนกพฤติกรรมการกระทำคามผิดของนักเรียนระดับอาชีวศึกษา. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.
- หทัยชนก กรชี. 2551. การคำนวณคุณลักษณะข้อมูลสำหรับตัวจำแนกในอีฟเบสต์ โดยวิธีการจัดกลุ่มข้อมูล และ ต้นไม้ตัดสินใจ. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.
- วีรพงษ์ ธนพงศ์ธรรม และ ธนอม น้อยหอม. 2550. ระบาดวิทยาของโรคไข้หวัดนกในประเทศไทยระหว่างปี 2547-2550. สำนักควบคุม ป้องกันและบำบัดโรคสัตว์ กรมปศุสัตว์ กรมปศุสัตว์, กรุงเทพฯ
- อรุณี ชัยสิงห์. 2002. โรคไข้หวัดนก (Birdflu). แหล่งที่มา:  
[http://www.dld.go.th/niah/Technicarticle/Tech\\_AI.htm](http://www.dld.go.th/niah/Technicarticle/Tech_AI.htm), 22 เมษายน 2552.
- Abbas, K, A.R. Mikler, A Ramezani and S. Menezes. 2004. Computational Epidemiology : Bayesian Disease Surveillance, pp 1 – 12. *In Proceeding of the International Conference on Bioinformatic and its Application.*
- Abbas, K, A.R.Mikler and R.Gatti. 2005. Temporal Analysis of Infectious Disease : Influenza, pp 267-271. *In the 21<sup>st</sup> Annual ACM Symposium on Applied Computing*
- Alhammady, H and K. Ramahanarao. 2006. Using Emerging Pattern to Construct Weighted Decision Tree. *In IEEE Transaction on Knowledge and Data Engineering (18):* 865-876.
- Agrawal, R. and R. Srikant 1994. Fast Algorithm for Mining Association Rules, pp 487-499. *In Proceeding of the 20<sup>th</sup> International Conference on Very Large Data Base.*

Bouckaert, R.R. 2008. **Bayesian network Classifier In Weka for Version 3-5-8.** 47.

Cohen, W.W. 1995. Fast Effective Rule Induction, pp115-123 *In Machine Learning : Proceeding of The twelfth international Conference.*

Cooper, G.F., D.H. Dash, J.D. Levender, W.K.Wong, W. R.Hogan and M. M Wagner. 2004. Bayesain Biosurveillance Of Disease Outbreaks, pp 94-103. *In The twenty-seventh Conference on Uncertainty in Artificial Intelligence*

Furnkranz, J. and G. Widmer. 1994. Incremental Reduced Error Pruning. **OEFAI-TR (9):** 1-15.

HAK, E, M.J Meijboom and E. Buskens. 2006. Modeling the health Economic Impact of the next influenza Pandemic. *In Vaccine (24):* 6756-6760.

Hall, M. 2007. A Decision Tree Based Attribute Weighting filter of Naive Bayes. **knowledge based System (20):** 120-126.

Li, J, Q. Ren, J. Yin and Xi Chen. 2004. Study On Transmission Model of Avian Influenza, pp 54-58. *In proceeding of 2004 International Conference on Information Acquisition*

Kahovi, R. 1996. Scaling up the Accuracy of naive Bayes Classifiers: A Decision Tree Hybrid, pp 202-207 *In Second International Conference on knowledge Discovery and Data mining.*

Mitchell, T.M. 1997. **Machine Learning.** 1 ed. McGraw-Hill Companies, Singapore

Prapaphan Prawing, Anongnart Srivihok. 2007. Adaptive Statistical transmission Model of Avian Influenza for Outbreak Prediction in human In Thailand, *In the 4th international Joint Conference on Computer Science and Software Engineering.*

- Sebastiani, P., K.D. Mandi, P. Szolovits, I.S. Konane and M.F. Ramoni. 2005. A Bayesian Dynamic Model for Influenza Surveillance. *In Journal of The American Statistical Association (100)* : 1-27.
- Upadhyay, R.K, N.Kumari and V.S.H Rao. 2007. Modeling the spread of bird flu and predicting outbreak diversity, pp 1638-1648 *In Press Non linear Annual Real world Application 2007*
- Vibound, C., P.Y. Boelle, K. Pakdaman, F. Carret, A.J. Vallerron and A. Flahault. 2002. Influenza Epidemic in the United States France and Australia 1972-1997. *In Emerging Infectious Disease 2004 (10)* : 32-39.
- Witten, I H and E. Frank. 2005. **Data Mining Practical Machine Learning Tools and Techniques.** 2 ed. Morgan kaufmann publisher, San Francisco.
- Wu, W.W. 2008. Choosing Knowledge Management Strategies by using a Combined ANP and DEMATEL Approach. *In Expert System With Application (35)*: 828-835.



ภาคผนวก



ผลงานวิจัยที่ได้นำเสนอ (Poster) ในที่ประชุมวิชาการ

Bankok international Conference on Avian Influenza 2008 ที่โรงแรมดุสิตธานี

วันที่ 23-25 มกราคม พ.ศ. 2551

PC-54

## Using Bayesian network for prediction model of avian influenza surveillance in Thai poultry farmland

Sirapote Santatiwut<sup>[1]</sup> Anongnart Srivihok<sup>[2]</sup>

Department of Computer science  
Faculty of Science Kasetsart University  
Jatujak , Bangkok 10900  
Email : g4964166@ku.ac.th<sup>[1]</sup> fsciagn@ku.ac.th<sup>[2]</sup>

This paper presents a prediction model for Avian Influenza in Thai Poultry Farmland. The aim of this study is to describe model and variables associated with Avian Influenza pandemic in Thailand corresponding to threat of an influenza pandemic. The variables of model include Site, Temperature, Season and Poultry Type. Data mining algorithm, Bayesian Network, was used for the prediction model. The measure of model prediction power is Accuracy Rate. This model will be useful for Avian Influenza Surveillance and strategic planning in control of Thai Poultry Farmland.



ผลงานทางวิชาการที่นำเสนอในที่ประชุมวิชาการ INCEB 2008  
ที่ โรงแรม แกรนด์ เมอร์เคียว ฟอรั่ม วันที่ 6-7 พฤศจิกายน พ.ศ. 2551

## Classification Algorithm for Avian Influenza Surveillance Model of Poultry Farms in Thailand

Sirapote Santatiwut<sup>[1]</sup>, Anongnart Srivihok<sup>[2]</sup>

Department of Computer Science, Faculty of Science, Kasetsart University, Thailand  
email: g4964166@ku.ac.th<sup>[1]</sup>, fsciang@ku.ac.th<sup>[2]</sup>

### Abstract

The pandemic of Avian Influenza (AI) in Thai poultry farms have threatened the growth of agriculture industry in Thailand. Since more than thousand poultries were dead and eradicated due to AI infections. The purpose of this study is to find the algorithm which can be used to build a classification model that predicts the site of poultry farms which are positive (infected) AI. Two classification algorithms: C4.5 and ID3 were tested on the prediction accuracy, precision and recall. The factors using for model generation including farm region, site characteristics, temperature, month, season, number of poultry, humidity, numbers daylight hours, numbers of sick poultries, numbers of dead and numbers of destroyed poultries.

Results showed that Decision Tree (C4.5) should be applied for Avian Influenza Surveillance (AIS) model. This proposed model is beneficial for predicting the positive site for AI epidemic. Thus the officials or farmers can use this model for AIS and protection of AI pandemic in the future.

**Keyword:** Classification, Decision Tree, Avian Influenza Surveillance

### 1. Introduction

At the present, the crisis of the bird Flu is critical as the important outbreak in the global level that leads to several disasters such as

economic, health, life quality and agriculture.

There are the serious pandemic of the Bird Flu, especially in Thailand. Journalists report Avian Influenza infections in poultry farms and some examples on human being in daily news.

The Avian Influenza (Bird Flu) is the name of virus which classified as the flu type H5N1, which is infected in poultry farm. It also can transfer to people through touching with sick poultry or infected animal corpses.

In Thailand, the serious of Bird Flu is considerable to be severe disease due to Thailand's economic is based on agriculture industry. In the past Bird Flu spreads rapidly and there were many infections in various zones. Researchers attempt to study on Bird Flu pandemic. This might provide information for efficiently protect and control the epidemic of disease, warning and watch out the danger of the Avian Influenza

In this paper, the classification model of Avian Influenza was proposed. The model was created by using a Decision Tree algorithm and the measurement of model efficiency by using accuracy rate, precision and recall.

### 2. Related Theory

#### 2.1 Decision Tree

Decision Tree is an algorithm which uses for the Data Mining and Classification by creating a prediction model showing the relationship between parameters and results. Decision Tree's structure consists of Leave Nodes expressed classification's

result and branch of Decision Tree that shown the linking of parameters classified for results of Leave Node.

There are many types of Decision Trees. In this study, two decision tree algorithms: ID3 and C4.5 were applied since they are most widely used in classification models.

ID3 [1] is an algorithm used for creating of a Decision Tree. This algorithm uses the concept of Information Entropy as:

$$I_E(i) = - \sum_{j=1}^m f(i, j) \log_2 f(i, j)$$

Similarly, C4.5 [8] is an algorithm used for creating of Decision Tree which was developed from ID3. The successor, C4.5 is using to predict classes in both numeric and non-numeric. However, C4.5 has more advantages as follows:

1. It can be used with both continuous and discrete parameters.
2. It complied with uncompleted parameters.
3. It is pruned after Decision Tree is generated.

### 2.2 Related Studies

Abbas and coworkers proposed the model of influenza in order to analyze time and patients who potentially have risk to be infected by flu in the spreading period. They tested hypothesis whether flu vaccine was capable to control flu well or not. Results showed that the vaccine performance was depending on its intense and efficiency [2].

Later, Abbas and coworkers used Bay Net to create a model to analyze how accurate the population was infected by the spread of flu [3]. Cooper and et al. [4] studied Causal Bayesian network and proposed a model: Spatial-temporal. This model was used to watch out and measure air-spreading epidemic disease which

can infect without contacting infected sources.

Li and his researchers [5] built a Transmission Model which used for explaining the spread of Bird Flu and predicting the future outbreak in mainland China. The outcome of this research is the model used for predicting infected areas in China.

Upadhyay and coworkers [6] focused on the spread of Avian Influenza in the terms of time and area in order to know how it was spread. This causes the model used for the efficient surveillance.

Sebastiani and group's research [7] presented the effects of flu by Dynamic Bayesian Network. This was used as a tool for the disease warning and the study of the express of patients who were sick in childhood with Severe Acute Respiratory Syndrome (SARS). These patients were treated in hospital's baby-care program for making the index of outbreak rates and death rates.

Hak and coworkers studied the model creation in order to monitor the side effects in Health – Economic area where flu spreads in the future. By using Decision Tree and basic data of the flu outbreak in the Netherlands, the model was proposed [9].

From the above studies, it is obviously seen that researchers use Bayesian Network to present the outbreak of flu and other diseases. However, the study of Avian Influenza in Thailand, which is one type to Flu, is never conducted by using data mining techniques. It is the purpose of this study to create a classification model of poultry farms to be infected by Avian Influenza.

### 3. Experiments

Data sets were obtained from the report of infected death of poultry since July 2004-2007 (the Bureau of Livestock) and weather reports of the same period from the Meteorological Department. There are totally about 8,501 records which classified as death proof from (1) the Bird Flu 1,647

records and (2) another reason, but not Bird Flu 6,854 records. Experimental Data were classified into 2 classes: (1) positive class which means dead poultry in farms infected by Bird flu and, (2) negative class which means dead poultry in farms not infected by Bird Flu.

There were 12 parameters in the model generating including (1) zone (2) month (3) season (4) Temperature (5) Humidity (6) Sun Light Day Hours (7) Number of poultry (8) number of sick poultry (9) number of death poultry (10) number of poultry was Destroyed (11) Type of poultry and (12) region as shown in Table 1.

To measure model's efficiency, RMSE (Root Mean Square Error), Recall, Precision and F-measure of the tested models were used for evaluations

Table 1 shows attributes of data in the model creation.

Name of variables	Attribute Type
Region	Nominal
Zone	Nominal
Month	Nominal
Season	Nominal
Temperature	Numeric
Humidity	Numeric
Sun light Day Hour	Numeric
Number Of Poultry	Numeric
Number of Death Poultry	Numeric
Number of Sick Poultry	Numeric
Number of poultry was Destroyed	Numeric
Type Poultry	Category

Three steps in the model creation are as follows:

#### Step 1. Data Preprocessing.

Data were collected from reports of infected death of poultry between July 2004-2007 (the Bureau of Livestock) and the weather reports of

the same periods from the Meteorological Department of Thailand. Uncompleted data and missing data were deleted. Then data files were formatted to CSV form for data mining software, WEKA version 3.4.13.

#### Step 2 Classification Models.

Create two classification models by using C4.5 and ID3 algorithms. Data set was obtained from step 1.

#### Step 3. Evaluation.

The measurements of model performances were conducted by using 10-fold cross validation. Then, precision, recall, F-Measure, accuracy rate and RMSE were used to test model efficiency. The comparisons of efficiency between two models (C4.5. and ID3) were conducted.

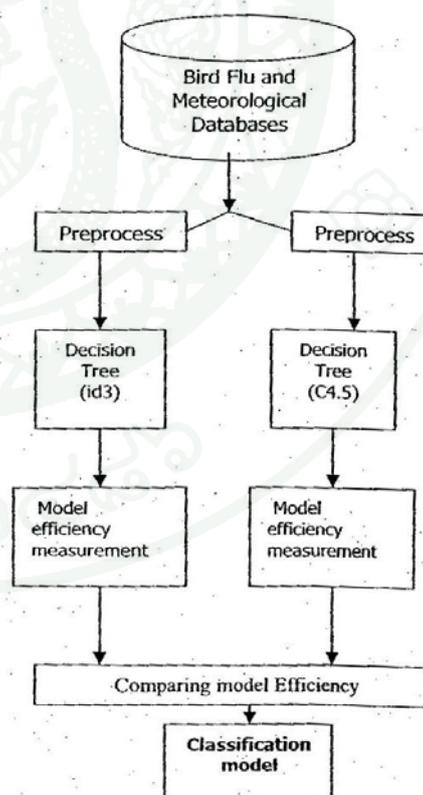


Figure 1. Framework of the study.

#### 4. Results

Results in table 2 showed that the prediction power of DT model (C4.5) is good with high correctly classified (91.11%) and low RMSE (0.2629). From Table 3, results of C4.5 predictions for Positive class, precision is 0.754, recall is 0.688 and F-measure is 0.72. Whereas, Negative class, precision is 0.939, recall is 0.955 and F-measure is 0.94 which is very good. However, Table 3 shows the prediction power of DT model (ID3) has lower accuracy rate (83.488%) than C4.5. Further ID3 provides the higher RMSE (0.3185) and lower precision and recall value than C4.5. Thus C4.5 algorithm has higher performances than ID3. Then, C4.5 was selected to generate the decision tree for AI model

Table 2 shows the comparisons of efficiency of C4.5 and ID3

Algorithm	Correctly Classified	Incorrectly classified	RMSE
C4.5	91.11	8.89	0.2629
ID3	83.4887	11.6045	0.3185

Table 3 shows the comparisons of Precision, Recall, F-measure of model created by Decision Tree in two classes

Algorithm	Precision	Recall	F-measure	Class
C4.5	0.754	0.688	0.72	Pos
	0.939	0.955	0.947	Neg
ID3	0.632	0.49	0.929	Pos
	0.911	0.948	0.552	Neg

The models of Bird Flu classification are generated by using C4.5 algorithm depicted in Figure 2, 3, 4 and 5 (see appendix).

Figure 2. shows the Bird Flu model created by Decision Tree algorithm. The left subtree of Bird Flu model is shown in Figure 3. This model of non-serious cases since no poultry was destroyed in the farms.

However, there were some positive cases indicated in this model.

Figure 4 showed the subtree(A) of this model which represents the conditions that less than 97 the number of poultry were destroyed in these areas. There were many positive cases occurred in rainy season. However, in summer, there was only the negative case.

Last, Figure 5 shows the sub tree represents numbers of poultry were destroyed more than 97 cases in those infected areas. In this sub tree (B), there are many positive cases of AI infections than negative cases.

From the classification model, it shows that most factors affecting the Bird Flu pandemic including poultry farm region is central part of Thailand and month is October and humidity is more than 60 % and temperature is more than 27.7 Celsius. No pandemic occurs in summer time.

Further another factors contribute to Bird Flu pandemic include the Winter Season (November and December) and Humidity  $\leq 75\%$  and poultry farms located in the central part of Thailand.

#### 5. Conclusions

From the above results, we conclude that Decision Tree (C4.5) algorithm is enable to generate a classification model. This model might be used to predict the infected site of bird flu in poultry farms of Thailand. Further, factors relate to Bird Flu pandemic in farms are proposed. These factors include central part of Thailand, winter month, humidity = 60-75%, temperature > 27.7 degree Celsius.

This model might be helpful in predicting the site of Bird Flu infection. The government, health related agency and poultry farms might use this model to prevent and watch for Bird Flu pandemic in Thailand.

### 6. Future Work

We might use Decision Tree model from this project to develop the surveillance warning system of the Bird Flu pandemic in the future. Moreover, other classification algorithms might be used to build models and evaluate their classification efficiency.

### 7. References

- [1] "Decision Tree Learning" Available at [http://en.wikipedia.org/wiki/Decision\\_Tree\\_Learning](http://en.wikipedia.org/wiki/Decision_Tree_Learning).
- [2] K.Abbas, A.R.Mikler, R.Gatti, "Temporal analysis of infectious disease : influenza". In *ACM Symposium on Applied Computing P267-271, 2005*.
- [3] K. Abbas , A.R. Mikler , Amir Ramezani, SHeena Menezes. "Computational Epidemiology: Bayesian Disease Surveillance". University of North Texas In *Proc. of the International Conference on Bioinformatics and its Applications, FL, USA, December 2004*.
- [4] G. F. Cooper , ET AL." Bayesian Biosurveillance of Disease Outbreaks." In *Conference on Uncertainty in Artificial Intelligence 2004*.
- [5] J. Li, Q. Ren ,Jianqin Yin and Xi Chen . "Study On Transmission Model of Avian Influenza" .*Proceeding of 2004 International Conference on Information Acquisition*
- [6] Ranjit Kumar Upadhyay , Nitu Kumari, V.Sree Hari Rao , "Modeling the spread of bird flu and predicting outbreak diversity". *In press Non linear Annual Real world Application 2007*
- [7] Paola Sebastiani , Kenneth D Mandl , Peter Szolovits, Issac s Konane , Marco F Ramoni , "A Bayesian Dynamic Model for Influenza Surveillance" . *Journal of the American Statistical Association 2005*
- [8] C4.5 . Available at [http://en.wikipedia.org/wiki/C4.5\\_Algorithm](http://en.wikipedia.org/wiki/C4.5_Algorithm)
- [9] E.Hak , ETAL . "Modeling the health-economic impact of the next influenza pandemic in the Netherlands" Available at [www.ScienceDirect.com](http://www.ScienceDirect.com)

Appendix

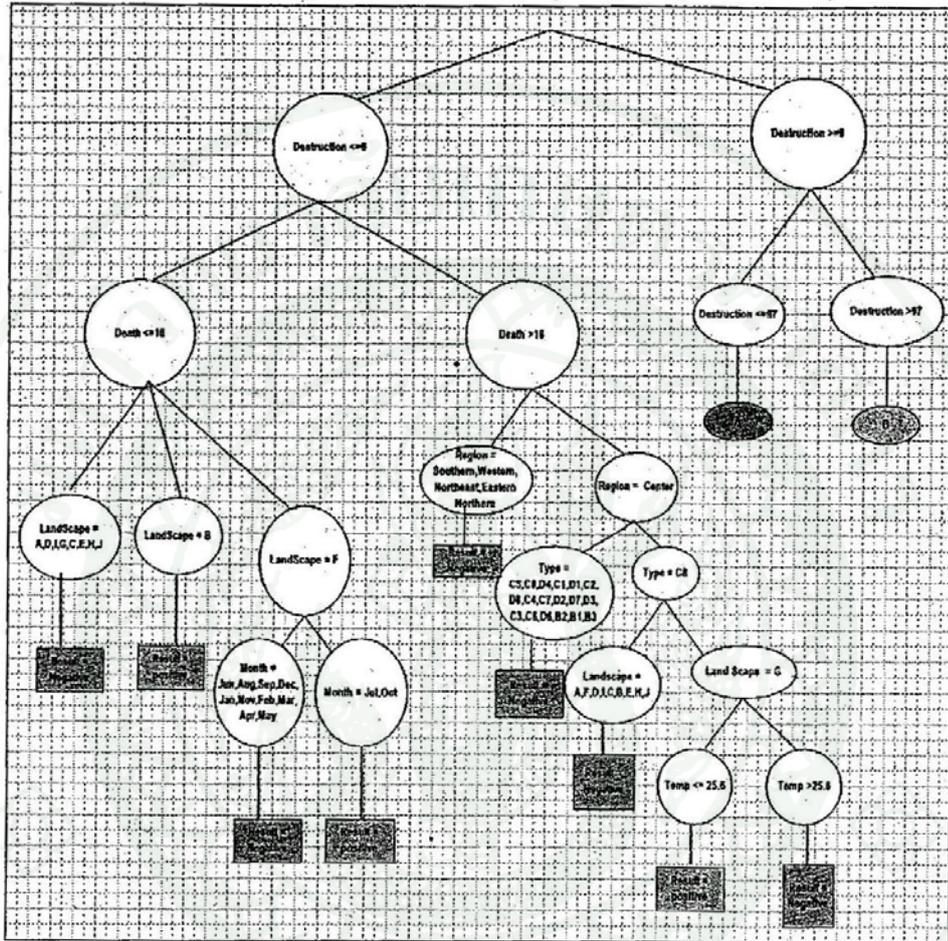


Figure 2. The final model created by using Decision Tree (C4.5)

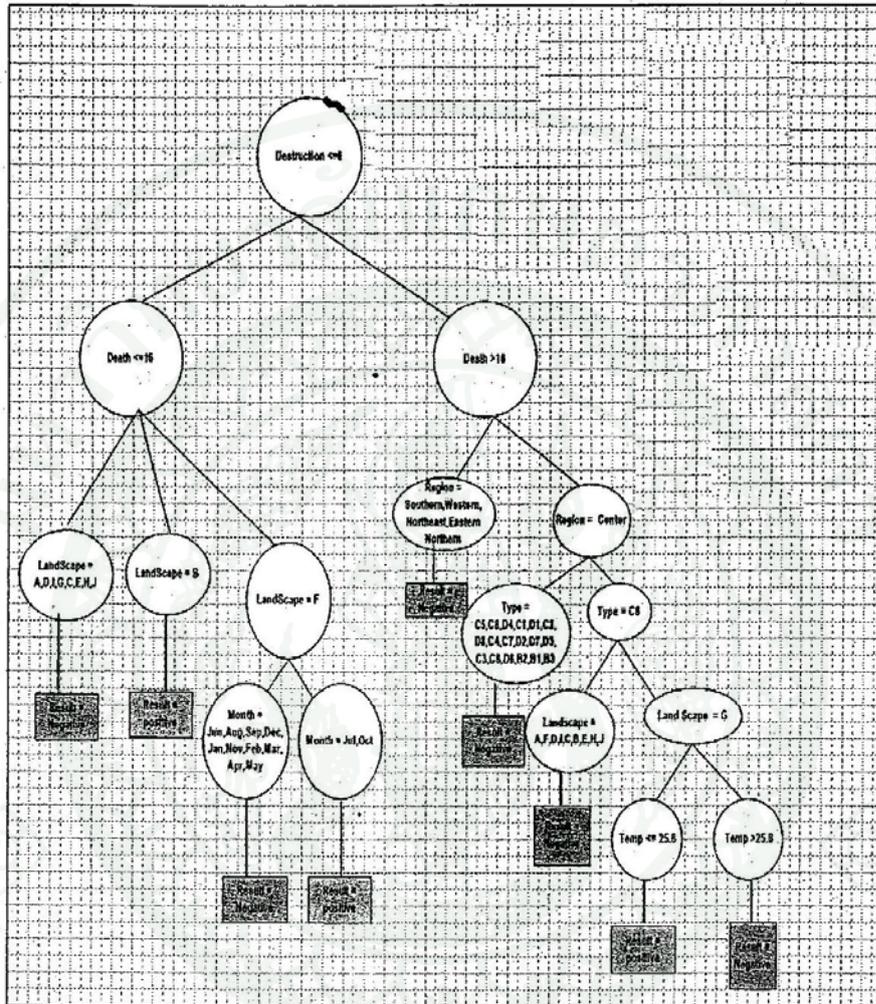


Figure 3. The model created by Decision Tree (left subtree of figure2: no poultries were destroyed)

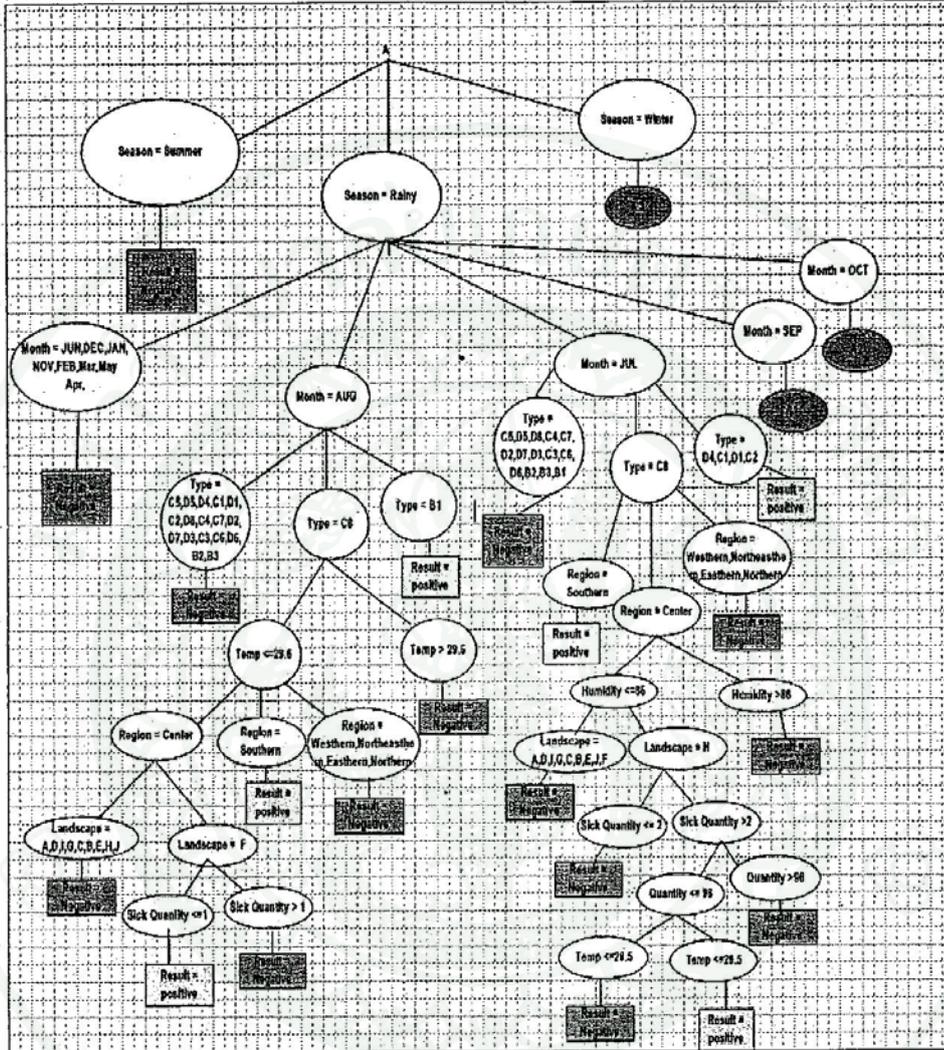


Figure 4. The sub model of Decision Tree ( subtree A of figure 2: Numbers of poultry were destroyed from 0 to 97)

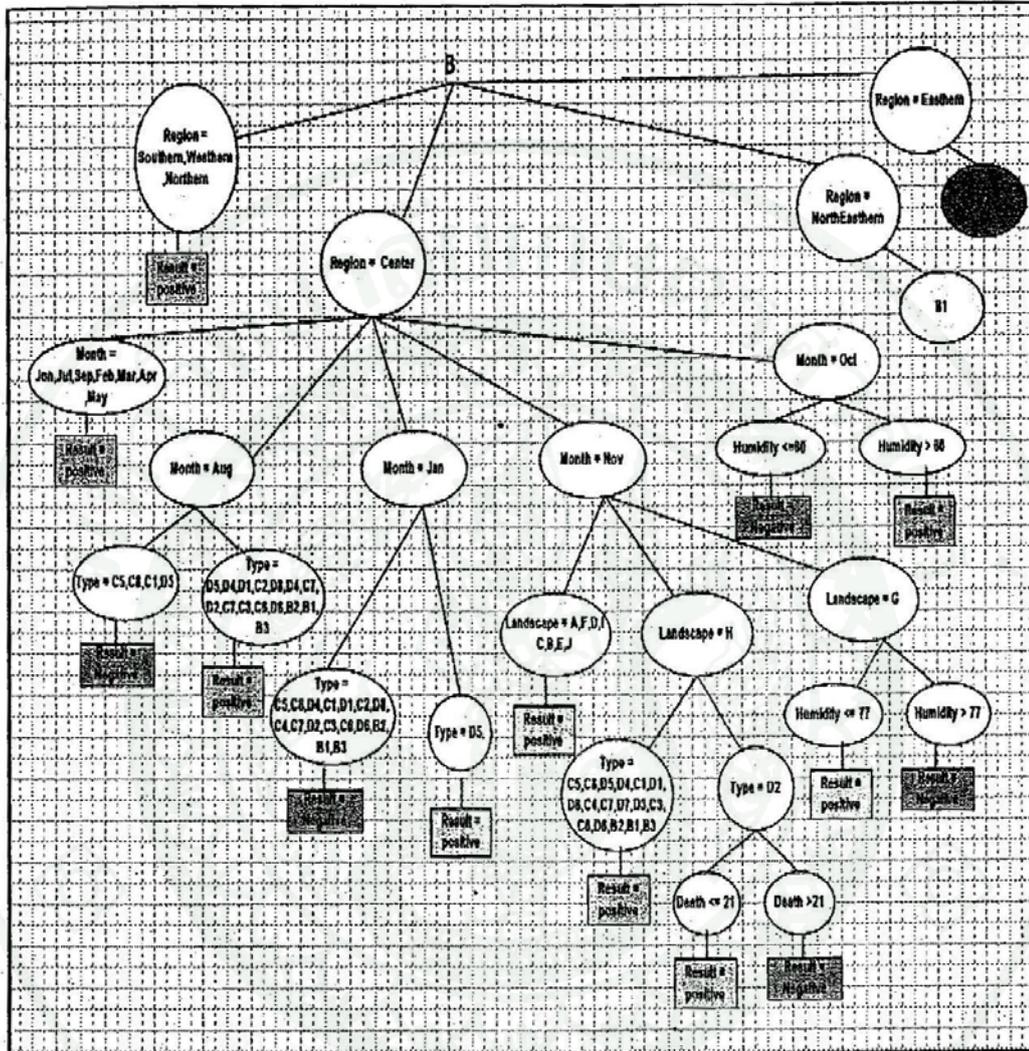


Figure 5. The sub model created by Decision Tree (subtree B of figure 2: Numbers of poultry were destroyed > 97)



ผลงานวิชาการที่นำเสนอในที่ประชุม JCSSE 2010  
ที่มหาวิทยาลัยรามคำแหง วันที่ 13-14 พฤศจิกายน พ.ศ. 2553

**แบบจำลองสำหรับการระบาดของโรคไข้หวัดนก  
สำหรับการเฝ้าระวังการระบาดของโรคไข้หวัดนกในพื้นที่เลี้ยงสัตว์ปีกของไทย  
โดยใช้เทคนิคเหมืองข้อมูล**

**Model for Avian Influenza Surveillance of Poultry Farms  
in Thailand by Data Mining Technique**

Sirapote Santatiwut<sup>1</sup> and Anongnart Srivihok<sup>2</sup>  
Computer Science, Kasetsart University, Bangkok, Thailand  
E-mail: [g4964166@ku.ac.th](mailto:g4964166@ku.ac.th)<sup>1</sup>, [fsciang@ku.ac.th](mailto:fsciang@ku.ac.th)<sup>2</sup>

**บทคัดย่อ**

ในงานวิจัยนี้นำเสนอแบบจำลองสำหรับการ  
จำแนกคุณลักษณะข้อมูลของโรคไข้หวัดนกเพื่อใช้ในการ  
เฝ้าระวังการระบาดในพื้นที่เลี้ยงสัตว์ปีกของไทย โดยที่  
แบบจำลองดังกล่าวสร้างโดยใช้อัลกอริทึมการจำแนก  
ข้อมูล คือ ต้นไม้ตัดสินใจ (Decision Tree) และข่ายงานแบบ  
เบย์ส์ (Bayesian Network) นอกจากนี้ได้มี การคัดเลือก  
คุณลักษณะ และ การถ่วงน้ำหนักของข้อมูลมาใช้ในการ  
ปรับปรุงประสิทธิภาพการทำงาน และทำการเปรียบเทียบ  
ประสิทธิภาพการทำงาน ระหว่าง แบบจำลองที่สร้างด้วย  
เทคนิคที่แตกต่างกัน เพื่อทดสอบว่าแบบจำลองใดให้ผลใน  
การทำงานได้ดีที่สุด ข้อมูลในการศึกษาได้จากข้อมูลสัตว์  
ปีกป่วยตาย ของ กรมปศุสัตว์ และ ข้อมูลสภาพภูมิประเทศ  
ของประเทศไทย ซึ่งการทดลองพบว่า แบบจำลองที่สร้าง  
จากเทคนิคข่ายงานแบบเบย์ส์ และการปรับปรุง  
ประสิทธิภาพการจำแนกข้อมูลด้วยการถ่วงน้ำหนัก  
ร่วมกับการคัดเลือกคุณลักษณะ ทำให้แบบจำลองมี  
ประสิทธิภาพในการทำงานได้ดีขึ้นกว่าแบบจำลองที่ไม่ได้  
มีการคัดเลือกคุณลักษณะ

**Abstract**

In this paper we propose a Classification Model  
for Avian Influenza Surveillance of Poultry Farms in  
Thailand . Bayesian Network and Decision Tree (C4.5)  
algorithms with attribute selection technique and  
Emerging Pattern Weighting Methods were used to  
create Classification Model for Avian Influenza. Data set  
were undertaken from both the Department of Livestock  
Ministry of Agriculture and Department of Meteorology .  
Results revealed that the performance of weighted  
Classification Model is better than non weighted .

**คำสำคัญ** ไข้หวัดนก การจำแนกข้อมูล การทำเหมือง  
ข้อมูล ข่ายงานแบบเบย์ส์ ต้นไม้ตัดสินใจ

บทนำ

โรคไข้หวัดนกเป็นโรคติดต่อชนิดหนึ่งที่สามารถติดต่อได้ทั้งในคนและสัตว์ ซึ่งผลกระทบจากการระบาดของโรคนั้นรุนแรงในหลายด้าน ไม่ว่าจะเป็นด้านเศรษฐกิจ สังคม คุณภาพชีวิต และ ด้านเกษตรกรรม จึงทำให้มีการกำหนดมาตรการ และ พัฒนาเครื่องมือสำหรับควบคุมการระบาด และ เฝ้าระวังการระบาดของโรคไข้หวัดนก เช่น มาตรการควบคุมการขนย้ายสัตว์ปีกข้ามพื้นที่ , มาตรการกำจัดสัตว์ปีกในพื้นที่ซึ่งเสี่ยงในรัศมี 5 กิโลเมตร โดยรอบจุดที่มีการระบาด เป็นต้น ในส่วนของงานวิจัย ก็มีงานวิจัยต่าง ๆ มากมายที่ค้นคว้าเพื่อให้ทราบลักษณะของเชื้อโรค เพื่อหาวิธีการควบคุม และ เฝ้าระวังที่มีประสิทธิภาพ เช่น การศึกษาลักษณะของเชื้อ ไข้หวัดนก การศึกษาลักษณะรวมชาติในการอยู่อาศัยของเชื้อ ไข้หวัดนก การคิดค้นเครื่องมือที่ช่วยในการวินิจฉัยโรค ไข้หวัดนกได้อย่างรวดเร็ว รวมถึง การคิดค้นเพื่อหาเครื่องมือที่ช่วยในการเฝ้าระวังจุดที่จะเกิดการระบาดของโรค ไข้หวัดนก เป็นต้น

วัตถุประสงค์ของงานวิจัยชิ้นนี้เพื่อนำเสนอแบบจำลองที่นำมาใช้ในการเฝ้าระวังการระบาดของโรค ไข้หวัดนกในพื้นที่เสี่ยงสัตว์ปีก โดยการเปรียบเทียบประสิทธิภาพของแบบจำลองที่สร้างด้วยเทคนิคการทำเหมืองข้อมูลคือ ข่ายงานแบบเบย์ส์ ค้นไม่คิดสนใจ

2. ทฤษฎี และ งานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีของเบย์

เป็นทฤษฎีที่ว่าด้วยเรื่องของความน่าจะเป็นแบบมีเงื่อนไขของตัวแปรสุ่ม ที่ขึ้นต่อกัน[1] และทฤษฎีนี้สามารถอธิบายการเปลี่ยนแปลงค่าความน่าจะเป็นเมื่อมี

การเปลี่ยนแปลงของเหตุการณ์ที่เป็นเงื่อนไข ซึ่งจะส่งผลให้ความน่าจะเป็นของผลสรุปเกิดการเปลี่ยนแปลงตามไปด้วย[8]

ทฤษฎีของเบย์ นำไปใช้ในการสร้างแบบจำลองเพื่อใช้ในการทำนายเหตุการณ์ที่มีความไม่แน่นอน ซึ่งแบบจำลองที่ได้สร้างขึ้นมานั้นจะมีลักษณะเป็นเหตุเป็นผลกันในเชิงความน่าจะเป็น แบบจำลองที่สร้างขึ้นมานี้เราเรียกว่า เครือข่ายงานแบบเบย์เชียน (Bayesian Network) และ วิธีการที่จะสรุปผลความน่าจะเป็นของเหตุการณ์ที่แบบจำลองทำนายนั้น เราเรียกว่า การอนุมานแบบเบย์เชียน (Bayesian Inference)

จากหลักการของทฤษฎีของเบย์ (Bayes' theorem) เมื่อนำมาเขียนเป็นสมการเชิงคณิตศาสตร์จะได้

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

P(A) คือ ค่าความน่าจะเป็นเบื้องต้นของ ตัวแปร A

P(B) คือค่าความน่าจะเป็นเบื้องต้นของ ตัวแปร B

P(B|A) คือความน่าจะเป็นแบบมีเงื่อนไขของตัวแปร B ขึ้นกับตัวแปร A

P(A|B) คือค่าความน่าจะเป็นของผลสรุปของ A ที่มีสาเหตุคือ B

2.2 ข่ายงานแบบเบย์

คือการนำความน่าจะเป็นแบบมีเงื่อนไขของตัวแปรสุ่มอิสระมาเขียนในลักษณะของสัญลักษณ์[1] ซึ่งเมื่อนำมาเขียนตาม สมการคณิตศาสตร์ ดังนี้

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | parent(X_i))$$

โดยที่

$X_i$  แทนตัวแปรที่อยู่ในชุดข้อมูลที่นำมาสร้างแบบจำลองด้วยข้อมูลแบบเบส

### 2.3 ต้นไม้ตัดสินใจ (Decision Tree)

Decision Tree เป็น อัลกอริทึมที่นำมาใช้ในการจำแนกข้อมูล (classification algorithm) ซึ่งเป็นเทคนิคหนึ่งของการทำเหมืองข้อมูลและนำมาใช้ในการสร้างแบบจำลองสำหรับการทำนาย ซึ่งแสดงให้เห็นถึงความสัมพันธ์ของตัวแปร กับ ผลลัพธ์ที่เชื่อมโยงกับตัวแปรนั้น [1] โดยในโครงสร้างของ Decision Tree จะประกอบด้วย ส่วนของใบ (leave node) ซึ่งแสดงผลลัพธ์ของการจำแนก และ ส่วนของกิ่งของ Decision Tree ซึ่งแสดงถึงการเชื่อมของตัวแปรที่นำมาใช้ในการจำแนกให้ได้ผลลัพธ์ใน leave node ออกมา

C4.5 เป็น algorithm ที่ใช้ในการสร้าง Decision Tree [1] ในการศึกษาครั้งนี้ใช้อัลกอริทึม C4.5 ในการสร้าง decision tree เนื่องจากเหตุผลดังต่อไปนี้ คือ (1) C4.5 ใช้กับข้อมูลที่เป็นแบบ Continuous และ Discrete (2) ใช้ชุดข้อมูลที่มีค่าไม่ครบถ้วนได้ และ (3) มีการทำ Prunning ต้นไม้หลังจากที่สร้าง Decision Tree แล้ว

### 2.4 การถ่วงน้ำหนักของข้อมูล (Weighting)

การถ่วงน้ำหนักของข้อมูลมีหลายวิธี แต่วิธีที่นำมาใช้คือวิธีการถ่วงน้ำหนักที่เรียกว่า Emerging Pattern [2] เป็นการนำเอา item set ที่อยู่ในชุดข้อมูลที่มีค่าความน่าจะเป็นอยู่ใน Class ใดคลาสหนึ่ง มากกว่า อีกคลาสหนึ่ง อย่างมีนัยสำคัญ นำมาคำนวณหาน้ำหนักให้กับข้อมูลแต่ละ instance โดยมีวิธีการคำนวณดังต่อไปนี้

ขั้นตอนที่ 1 แบ่งข้อมูลออกเป็นสองชุดตามผลเฉลยของข้อมูล

ขั้นตอนที่ 2 ทำการค้นหา item set ของข้อมูลที่อยู่ในแต่ละชุดข้อมูล และ ค้นหาค่า support ของแต่ละ Item set จากสูตร

$$S_{D_i}(X) = \frac{\text{Count}_{D_i}(X)}{|D_i|}$$

โดยที่

$S_{D_i}(X)$  คือ ค่า support ของ item set ที่มีค่าชุด ข้อมูล  $D_i$

โดยที่  $i = 1, 2, 3, \dots, n$

$\text{Count}_{D_i}(X)$  คือ จำนวนข้อมูลที่อยู่ในชุด  $D_i$  ที่มี item  $x$

$|D_i|$  คือ จำนวนข้อมูลที่อยู่ในชุด  $D_i$

ขั้นตอนที่ 3 หาค่า Growth Rate ให้กับแต่ละ Item set ที่ได้มาจากแต่ละชุดข้อมูล ค่า Growth Rate คือสัดส่วนของค่า support ของแต่ละ item set ที่มีต่อ ผลเฉลยของข้อมูล ซึ่งแสดงถึง อำนาจในการจำแนกข้อมูล ของ Item set โดยสูตรการคำนวณคือ

$$\text{Gr}_{D_i \rightarrow D_j}(X) = \begin{cases} 0 & \text{if } S_i(X) = 0 \text{ and } S_j(X) = 0 \\ \infty & \text{if } S_i(X) = 0 \text{ and } S_j(X) \neq 0 \\ \frac{S_j(X)}{S_i(X)} & \text{Other wise} \end{cases}$$

โดยที่

$\text{Gr}_{D_i \rightarrow D_j}(X)$  คือ ค่า Growth Rate ระหว่าง ชุดข้อมูลที่แบ่งจากขั้นตอนที่ 1

$S_i(X)$  คือ ค่า support ของ item set ที่คำนวณมาจากชุดข้อมูล  $I$  ที่แบ่งเมื่อขั้นตอนที่ 1

$S_j(X)$  คือ ค่า support ของ item set ที่คำนวณมาจากชุดข้อมูล  $J$  ที่แบ่งเมื่อขั้นตอนที่ 1

เมื่อหาค่า Growth Rate  $Gr(e)$  ได้แล้วทำการคัดเลือก item set โดยทำการเลือก item set ที่มีค่า  $Gr > 1$  ซึ่ง item set ที่ผ่านการคัดเลือกจะเรียกว่า Emerging Pattern เรียกว่า  $e$

ขั้นตอนที่ 4 หาค่า strength ของ Emerging Pattern โดยมีสูตรในการคำนวณคือ

$$\alpha_{C_j}(e) = \frac{gr_{C_j}(e)}{1 + gr_{C_j}(e)} * S_{C_j}(e)$$

โดยที่

$C_j$  คือชุดข้อมูลที่มีผลเฉลยของข้อมูลเป็นชุดเดียวกัน

$gr_{C_j}(e)$  คือค่า Growth Rate ของ Emerging Pattern ที่ปรากฏในชุดข้อมูลที่มีผลเฉลยอยู่ใน Class  $C_j$

$S_{C_j}(e)$  คือค่า support ของ Emerging Pattern ที่ปรากฏอยู่ในชุดข้อมูลที่มีผลเฉลยอยู่ใน Class  $C_j$

ขั้นตอนที่ 5 นำ Emerging pattern ที่หามาใส่ลงไปในแต่ละ Instance ที่มี Emerging Pattern ปรากฏอยู่ และหาผลรวมของค่า strength ของ Emerging Pattern ทั้งหมดที่ปรากฏอยู่ในแต่ละ instances โดยสูตรการคำนวณคือ

$$\beta_{C_j}(i) = \sum_{e \in E_{C_j}} \alpha_{C_j}(e)$$

โดยที่

$\beta_{C_j}(i)$  คือ ผลรวมของค่า strength ของ Emerging Pattern ที่ปรากฏในแต่ละ instance ของข้อมูล

$E_{C_j}$  คือ set ของ Emerging Pattern ที่เป็นของชุดข้อมูลที่มีผลเฉลยอยู่ใน คลาส  $C_j$

$i$  คือ instance ที่มีผลเฉลยอยู่ในคลาส  $C_j$

ขั้นตอนที่ 6 นำผลรวมของ Emerging Pattern ที่หามาจากขั้นตอนที่ 5 ของแต่ละ instance มาหาค่าน้ำหนักของแต่ละ instance โดยหาจากสูตร

$$w_{C_j}(i) = \frac{\beta_{C_j}(i)}{\text{median } C_j}$$

โดยที่

median  $C_j$  คือ ค่ามัธยฐานของ  $\beta_{C_j}(i)$  ของแต่ละ instance ที่อยู่ในชุดข้อมูลที่มีผลเฉลยอยู่ในกลุ่ม  $C_j$

### 2.5 การคัดเลือกคุณลักษณะของชุดข้อมูล (Selection Attribute)

ในการวิจัยนี้เราใช้การคัดเลือกคุณลักษณะของชุดข้อมูลมาใช้ในการปรับปรุงประสิทธิภาพการทำงานของแบบจำลอง โดยอัลกอริทึมที่นำมาใช้ในการคัดเลือกคุณลักษณะในงานวิจัยนี้คือ Classifier subset evaluator [3] ซึ่งเป็นการนำเอา อัลกอริทึมการจำแนกข้อมูลมาทำการประเมิน สัมผัสของ Attribute โดยทำการประเมินจากความแม่นยำของการจำแนกข้อมูลเมื่อใช้กับชุดของ Attribute ในการจำแนกข้อมูล

### 2.6 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการสร้างแบบจำลองของโรคไข้หวัดนก มีหลายแนวทาง เช่น Kaja Abbas และคณะ [4] ได้ทำการสร้างแบบจำลองของโรคไข้หวัดใหญ่ (Influenza Virus) เพื่อใช้ในการวิเคราะห์ช่วงเวลา และกลุ่มคนที่มีโอกาสเป็นโรคไข้หวัดใหญ่ ในช่วงเวลาที่มีการระบาด และ ทำการทดสอบสมมติฐานของแพทย์ที่กล่าวว่า วัคซีนของโรคไข้หวัดใหญ่ จะควบคุมการระบาดของโรคไข้หวัดใหญ่ได้ดีหรือไม่ขึ้นอยู่กับ ความเข้มข้น ของวัคซีน และ ประสิทธิภาพของวัคซีน

ต่อมา Abbas และคณะ [5] ได้สร้างแบบจำลองด้วย ข่ายงานแบบเบย์ เพื่อวิเคราะห์ผลลัพธ์ที่ตัวแปรทางด้านประชากรส่งผลต่อการระบาดของโรค ไข้หวัดใหญ่

Cooper และ คณะ [6] ศึกษาและนำ อัลกอริทึมเครือข่ายแบบเบย์ เรียกว่า Causal Bayesian network มาใช้สร้าง โมเดลที่มีลักษณะแบบ Spatial-Temporal เพื่อนำไปใช้สำหรับการตรวจจับโรคที่เกิดขึ้น โดยในงานวิจัยนี้มุ่งทำการศึกษาโรคที่มีลักษณะที่การแพร่กระจายทางอากาศซึ่งไม่ได้เกิดจากการสัมผัสโดยตรงกับเชื้อ

ในประเทศจีน Li และ คณะ [7] นำเสนอ Transmission model ที่ใช้ในการอธิบายการแพร่ระบาดของโรคไข้หวัดนกในสัตว์ปีก และ ทำนายสถานการณ์ในอนาคต ที่เกิดขึ้นในประเทศสาธารณรัฐประชาชนจีน ซึ่งสิ่งที่ได้คือ แบบจำลองที่ใช้ในการทำนายจำนวนพื้นที่ที่จะมีการระบาดของโรคไข้หวัดนก ในประเทศสาธารณรัฐประชาชนจีน

Upadhyay และ คณะ [8] เป็นงานวิจัยที่ต้องการศึกษาการแพร่ระบาดของโรคไข้หวัดนก (Avian Influenza) ทั้งในเชิงเวลา และ พื้นที่ เพื่อที่ได้เข้าใจกลไกการแพร่กระจายของเชื้อไข้หวัดนก ซึ่งผลที่ได้คือแบบจำลองที่สามารถนำมาใช้ในการเฝ้าระวังควบคุมโรคไข้หวัดนก ได้อย่างมีประสิทธิภาพ

Sebastiani และ คณะ [9] นำเสนอ ผลกระทบโรคไข้หวัดใหญ่ โดยใช้ Dynamic Bayesian Network เป็นเครื่องมือที่นำมาใช้ในการศึกษาการเฝ้าระวังโรค และ ศึกษาถึงลักษณะที่แสดงออกมาของผู้ป่วยที่ในวัยเด็กที่เป็นโรค Severe Acute Respiratory Syndrome (SARS) ที่ได้รับ

การดูแลทางด้าน กุมารเวช เพื่อนำมาใช้เป็นดัชนีของการเกิดโรคและ อัตราการตายที่จะเกิดขึ้น

จากงานวิจัยที่ได้กล่าวไปแล้วจะเห็นได้ว่ามีนักวิจัยหลายกลุ่มได้นำ อัลกอริทึมที่เกี่ยวข้องกับการจำแนกข้อมูล เช่น Bayesian Network มาใช้ในการทำนายการเกิดโรคไข้หวัดใหญ่ และ ใช้ในการดูความรุนแรงของการระบาดของโรคไข้หวัดใหญ่ แต่ไม่มีรายงานการใช้วิธีการคัดเลือกคุณลักษณะข้อมูล และ การถ่วงน้ำหนักมาเข้าร่วมกับ อัลกอริทึมจำแนกข้อมูลในการสร้างตัวแบบจำลองเพื่อใช้ในการทำนาย

ดังนั้นทางผู้วิจัยที่จะนำอัลกอริทึม Decision Tree และ Bayesian Network มาใช้ในการสร้างแบบจำลองเพื่อใช้ในการทำนายการเกิดโรคไข้หวัดนกโดย เริ่มจากการคัดเลือกคุณลักษณะ และ การถ่วงน้ำหนักข้อมูลเพื่อสร้างแบบจำลอง และ เปรียบเทียบประสิทธิภาพในการทำนายคลาสของข้อมูลของแบบจำลองทั้งสอง

### 3. ขั้นตอนการทดลอง

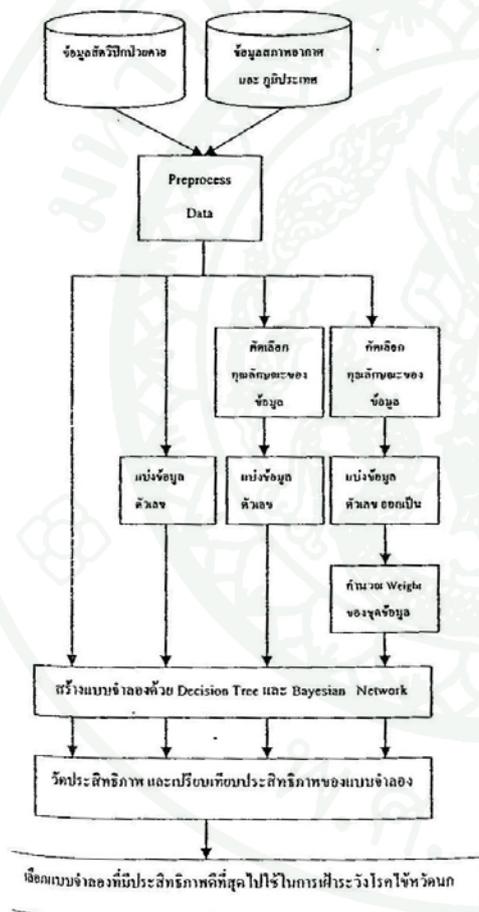
ในการสร้างแบบจำลองขึ้นโดยข้อมูลที่นำมาใช้ในการศึกษาครั้งนี้คือ รายงานสถานการณ์สัตว์ปีกป่วยตาย ตั้งแต่ เดือน กรกฎาคม 2547 – พฤศจิกายน 2550 ของกรมปศุสัตว์ โดยมีจำนวนข้อมูลทั้งสิ้น 8,152 ระเบียบ ซึ่งแบ่งเป็นกรณีที่มีการยืนยันแล้วว่าพื้นที่ที่พบการป่วยตายของสัตว์ปีกอันเกิดจากสาเหตุของโรคไข้หวัดนก (Positive) มีจำนวน 1,351 ระเบียบ และ กรณีที่มีการยืนยันว่าสัตว์ปีกที่ป่วยตาย ไม่ได้มาจากสาเหตุของเชื้อไข้หวัดนก มีจำนวนทั้งสิ้น 6,801 ระเบียบ (Negative) ตัวแปรที่นำมาใช้ในการสร้างแบบจำลอง ได้แก่ เดือน ลักษณะของพื้นที่ ความสูง ภูมิภาค ชนิดสัตว์ปีก จำนวนสัตว์ปีก จำนวนสัตว์

ปีกที่ป่วย จำนวนสัตว์ปีกที่ตาย จำนวนสัตว์ปีกที่ถูกทำลาย  
อุณหภูมิ ความชื้น ชั่วโมงที่แสงสว่าง และ ฤดูกาล

ในการวัดประสิทธิภาพของแบบจำลองใช้การ

วัดค่าความแม่นยำ จะดูจากค่า Recall ,ค่า Precision และ  
ค่า F-measure และ ความสามารถในการจำแนกที่ถูกต้อง  
(Accuracy) ของแบบจำลอง

ภาพที่ 1 แสดงแผนผังการทดลอง



ตารางที่ 1 การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูล  
(Accuracy) ระหว่าง แบบจำลอง ที่สร้างจาก Bayesian  
Network (BN) และ Decision Tree (DT)

Model	Discretize	Select Attribute	Weight	%Correct	%Incorrect
C4.5	N	N	N	89.25	10.75
BN	N	N	N	91.33	8.67
C4.5	Y	N	N	89.77	10.23
BN	Y	N	N	91.52	8.48
C4.5	N	Y	N	91.11	8.88
BN	N	Y	N	91.04	8.96
C4.5	Y	Y	N	89.65	10.35
BN	Y	Y	N	91.41	8.59
C4.5	Y	N	Y	95.19	4.81
BN	Y	N	Y	96.87	3.53
C4.5	Y	Y	Y	95.19	4.81
BN	Y	Y	Y	95.97	4.03

จากตารางที่ 1 พบว่า แบบจำลองที่ให้ประสิทธิภาพการ  
จำแนกข้อมูล (Accuracy) ดีที่สุดคือ แบบจำลองที่สร้างจาก  
Bayesian Network ซึ่งสร้างจากชุดข้อมูลที่มีการถ่วง  
น้ำหนัก โดยที่มีประสิทธิภาพในการจำแนกข้อมูล  
(Accuracy) อยู่ที่ 96.87 %

ตารางที่ 2 การเปรียบเทียบค่า Precision ,Recall และ F-measure ระหว่างแบบจำลอง ที่สร้างจาก Bayesian N Network (BN) และ Decision Tree (C4.5)

cl	Discreti ze	Select Attribut e	Weight	Precisio n	Recall	F- measure	Class
1	N	N	N	0.724	0.568	0.637	Pos
				0.918	0.957	0.937	Neg
2	N	N	N	0.751	0.713	0.731	Pos
				0.944	0.953	0.912	Neg
3	Y	N	N	0.778	0.535	0.634	Pos
				0.913	0.97	0.941	Neg
4	Y	N	N	0.753	0.727	0.74	Pos
				0.946	0.953	0.949	Neg
5	N	Y	N	0.808	0.609	0.695	Pos
				0.926	0.971	0.948	Neg
6	N	Y	N	0.751	0.688	0.718	Pos
				0.939	0.955	0.947	Neg
7	Y	Y	N	0.801	0.499	0.615	Pos
				0.907	0.975	0.94	Neg
8	Y	Y	N	0.759	0.707	0.732	Pos
				0.943	0.955	0.949	Neg
9	Y	N	Y	0.904	0.791	0.844	Pos
				0.96	0.983	0.972	Neg
10	Y	N	Y	0.891	0.895	0.893	Pos
				0.979	0.979	0.979	Neg

Model	Discreti ze	Select Attribut e	Weight	Precisio n	Recall	F- measure	Class
C4.5	Y	Y	Y	0.94	0.885	0.912	Pos
				0.942	0.971	0.956	Neg
BN	Y	Y	Y	0.943	0.939	0.941	Pos
				0.968	0.971	0.969	Neg

จากผลการทดลองจะเห็นได้ว่า ประสิทธิภาพการจำแนกข้อมูลของทั้งสองอัลกอริทึมมีความใกล้เคียงกัน คือ จำแนกถูกต้องมากกว่า 80% ส่วนค่า Precision ของทั้งสองแบบจำลองมากกว่า 0.7 ในกรณีที่มีข้อมูลที่นำมาใช้ไม่ได้มีการด่วงน้ำหนัก และ การคัดเลือกคุณลักษณะ แต่หลังจากคัดเลือกคุณลักษณะ และ ด่วงน้ำหนักข้อมูลแล้ว ทั้งสองแบบจำลองมีค่า Precision และ F-measure มากกว่า 0.9 และ จากตารางการทดลองจะเห็นได้ว่า การนำ Weight ไปใช้ร่วมกับอัลกอริทึมการจำแนกข้อมูลจะทำให้มี ประสิทธิภาพการจำแนกที่สูงกว่า การไม่นำ Weight ไปใช้ร่วมกับการจำแนกข้อมูล ทั้งกรณีของ Decision Tree และ Bayesian Network

5. สรุปผลการทดลอง

ในส่วนของการทดลองเห็นได้ว่า แบบจำลองที่สร้างด้วย Bayesian Network มีประสิทธิภาพในการจำแนกข้อมูลได้ถูกต้องดีกว่า แบบจำลองที่สร้างด้วย Decision Tree ยกเว้น กรณีที่มีข้อมูลมีการคัดเลือก Attribute แต่ไม่มีการ Discretize และ การด่วงน้ำหนัก เท่านั้นที่แบบจำลองที่สร้างด้วย Decision Tree มีประสิทธิภาพที่ดีกว่า Bayesian Network แต่ดีกว่ากันเพียงแค 0.07 ซึ่งเป็นผลมาจากการคัดเลือก Attribute ก่อนสร้างโมเดล และ เมื่อ

ดำเนินการถ่วงน้ำหนักของข้อมูล และการคัดเลือก Attribute เพื่อนำไปสร้างแบบจำลองพบว่า อัลกอริทึมส์ Bayesian Network มีประสิทธิภาพการทำนายที่ดีกว่า Decision Tree ตัวแบบที่สร้างขึ้นนี้ สามารถนำไปใช้ในการทำนายพื้นที่เสี่ยงในการติดเชื้อ หรือ ระบาดจากไข้หวัดนก โดยกำหนดค่าตัวแปรต่างๆ ของพื้นที่ และ นำผลการทำนายไปใช้ในการเฝ้าระวังเพื่อการเฝ้าระวังหรือป้องกันโรคในอนาคต

#### 6. แนวทางในอนาคต

วิธีการในการสร้างแบบจำลองเพื่อใช้ในการเฝ้าระวังโรคไข้หวัดนกในพื้นที่เสี่ยงตัวปีกลนั้น สามารถที่จะนำไปเป็นแนวทางในการพัฒนาระบบสารสนเทศเพื่อเฝ้าระวังโรคไข้หวัดนก โดยอาจจะใช้ข้อมูลสภาพแวดล้อม และ ข้อมูลอื่นที่เกี่ยวข้องนำมาสร้างแบบจำลองด้วยเทคนิคทางด้านการทำเหมืองข้อมูล ซึ่งผลลัพธ์ที่ได้คือแบบจำลองที่สามารถใช้ในการทำนายโอกาสที่จะเกิดการระบาดของโรคได้

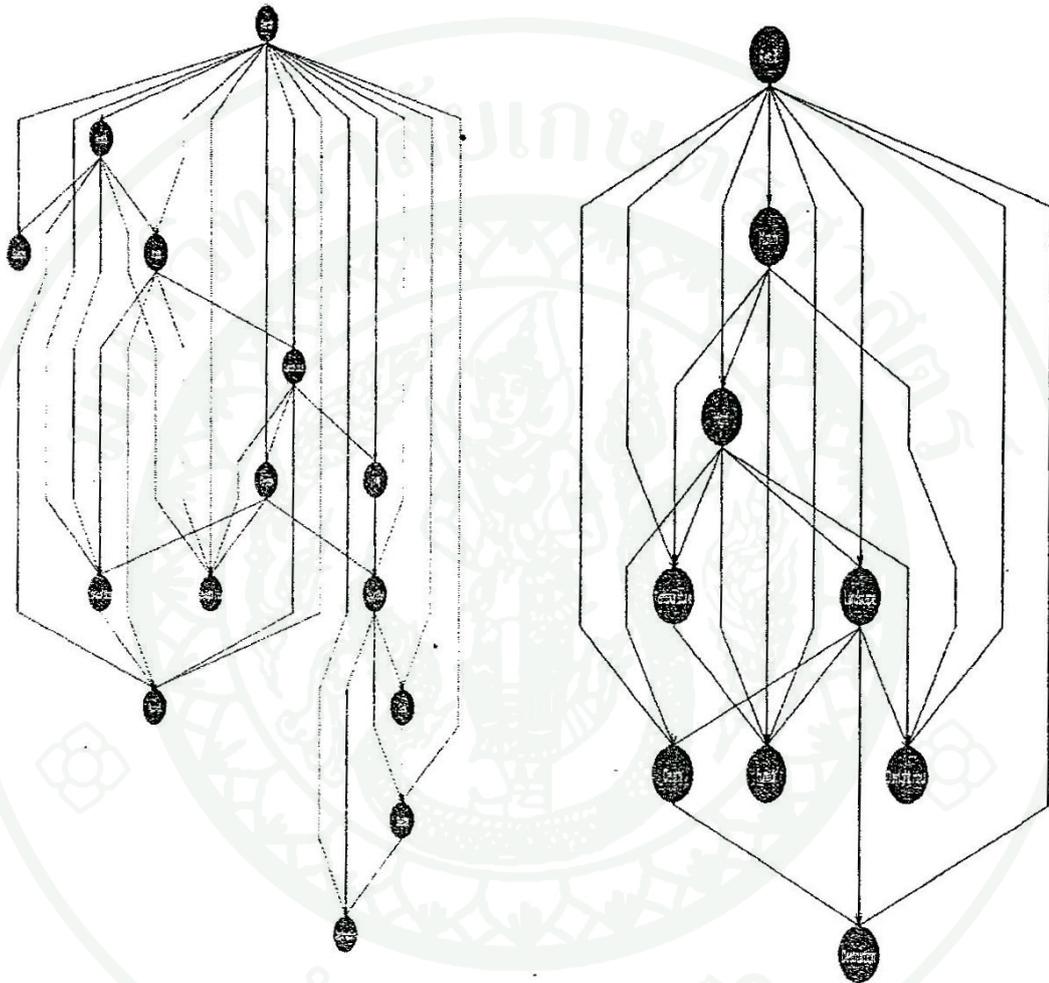
#### 7. เอกสารอ้างอิง

- [1] T.M. Mitchell. *Machine Learning*. Magraw-Hill .1997
- [2] H. Alhammady and K. hamahanarao . "Using Emerging Pattern to Construct Weight Decision Tree". IEEE Transaction on Knowledge and Data Engineering 2006
- [3] I. H. Witten and Eibe Frank. *Data Mining Practical Machine Learning Tools and Techniques*. Second Edition., Morgan Kaufmann, 2005
- [4] K. Abbas, A.R. Mikler, R. Gatti , "Temporal Analysis of Infectious Disease : Influenza". In *2005 ACM Symposium on Applied Computing*. P.267-271
- [5] K. Abbas , A.R. Mikler , Amir Ramezani, Sheena Menezes. "Computational Epidemiology: Bayesian Disease Surveillance". University of North Texas In *Proc. of the International Conference on Bioinformatics and its Applications*, FL, USA, December 2004.
- [6] G. F. Cooper , D.H. Dash, J. D. Levender , Weng Keen Wong ,W. R. Hogan ,M.M. Wagner. "Bayesian Biosurveillance of Disease Outbreaks." In *Conference on Uncertainty in Artificial Intelligence* 2004
- [7] J. Li, Q. Ren ,Jianqin Yin and Xi Chen . "Study On Transmission Model of Avian Influenza" .*Proceeding of 2004 International Conference on Information Acquisition*
- [8] Ranjit Kumar Upadhyay , Nitu Kumari, V. Sree Hari Rao , "Modeling the spread of bird flu and predicting outbreak diversity" . In *press Non linear Annual Real world Application* 2007
- [9] P. Sebastiani , K. D Mandl , P. Szolovits, I. s Konane , M. F Ramoni , "A Bayesian Dynamic Model for Influenza Surveillance" . *Journal of the American Statistical Association* 2005

## ภาพผนวก

ภาพที่ 2 แบบจำลองที่สร้างด้วย Bayesian Network ที่สร้างจากชุดข้อมูลผ่านการ Discretize และ ผ่านการถ่วงน้ำหนัก แต่ไม่ผ่านการ Select Attribute

ภาพที่ 3 แบบจำลองที่สร้างด้วย Bayesian Network ที่สร้างจากชุดข้อมูลผ่านการ Discretize และ ผ่านการถ่วงน้ำหนัก และ ผ่านการ Select Attribute



## ประวัติการศึกษา และการทำงาน

ชื่อ –นามสกุล	นายสิรพจน์ สันตติวุฒิ
วัน เดือน ปี ที่เกิด	วันที่ 16 มิถุนายน 2527
สถานที่เกิด	กรุงเทพ
ประวัติการศึกษา	วท.บ. (คณิตศาสตร์) มหาวิทยาลัยศรีนครินทรวิโรฒ
ตำแหน่งหน้าที่การงานปัจจุบัน	เจ้าหน้าที่คอมพิวเตอร์
สถานที่ทำงานปัจจุบัน	บมจ. เอเซียเน็ทเวิร์คส์ ดีเวลลอปเม้นท์
ผลงานดีเด่นและรางวัลทางวิชาการ	นำเสนอผลงานในที่ประชุมวิชาการ AI Conference 2008 ที่ โรงแรมดุสิตธานี กรุงเทพ นำเสนอผลงานวิชาการในงานประชุม วิชาการ INCEB 2008 ที่โรงแรม เมอร์เคียว กรุงเทพ นำเสนอผลงานวิชาการในงานประชุม JCSSE 2010 ที่ มหาวิทยาลัย รามคำแหง วันที่ 12-14 พฤษภาคม 2553 ทุนสนับสนุนการเสนอผลงานจากบัณฑิตวิทยาลัย มก.
ทุนการศึกษาที่ได้รับ	