

บทที่ 3

วิธีการดำเนินการวิจัย

ในบทนี้จะกล่าวถึงการพิจารณาคัดเลือกองค์ประกอบของเว็บไซต์เพื่อนำมาใช้ในการประเมินระดับความน่าเชื่อถือของเว็บไซต์ และวิธีการวิจัย ซึ่งประกอบด้วย รูปแบบและวิธีการวิจัย ประชากรและกลุ่มตัวอย่าง เครื่องมือในการวิจัย การเก็บรวบรวมข้อมูล การเตรียมข้อมูล และการวิเคราะห์ข้อมูล เพื่อให้การดำเนินการทำวิจัยเป็นไปได้ตามที่ตั้งวัตถุประสงค์ไว้

3.1 การพิจารณาคัดเลือกองค์ประกอบในการประเมินระดับความน่าเชื่อถือของเว็บไซต์

จากงานวิจัยเกี่ยวกับความน่าเชื่อถือในระบบพาณิชย์อิเล็กทรอนิกส์ในหัวข้อที่ 2.2 สามารถนำมาสร้างเป็นตารางเปรียบเทียบองค์ประกอบที่สำคัญที่มีผลกระทบต่อความน่าเชื่อถือของเว็บไซต์ โดยพิจารณาเฉพาะองค์ประกอบที่ได้รับความน่าเชื่อถือสูง ได้ดังตารางที่ 3.1 เมื่อนำพิจารณาหาความเป็นไปได้ ที่จะนำองค์ประกอบเหล่านี้มาเพื่อทำการศึกษผลกระทบต่อความน่าเชื่อถือของเว็บไซต์ จึงได้คัดเลือกองค์ประกอบที่จะนำมาใช้ในงานวิจัย ดังนี้ คือ

1.การรักษาความปลอดภัยข้อมูลบัตรเครดิต เนื่องจากเป็นองค์ประกอบที่มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์ จากงานวิจัยของ Mahmood (2006) Lumsden (2009) และศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (2552)

2.การรักษาข้อมูลส่วนบุคคล เนื่องจากเป็นองค์ประกอบที่มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์ จากงานวิจัยของ Mahmood (2006) Lumsden (2009) และวรรณิดา กลิ่นดี (2551)

3.ข้อมูลบริษัท เนื่องจากเป็นองค์ประกอบที่มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์ จากงานวิจัยของ Lumsden (2009)

4.ข้อมูลในการติดต่อ เนื่องจากเป็นองค์ประกอบที่มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์ จากงานวิจัยของ Lumsden (2009) และวรรณิดา กลิ่นดี (2551)

5.การส่งสินค้าและคืนสินค้า เนื่องจากเป็นองค์ประกอบที่มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์ จากงานวิจัยของ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (2552) และวรรณิดา กลิ่นดี (2551)

6.การแสดงราคาสินค้าและค่าขนส่ง เนื่องจากเป็นองค์ประกอบที่มีความสำคัญต่อความน่าเชื่อถือของ เว็บไซต์ จากงานวิจัยของวรรณิดา กลิ่นดี (2551)

7.รูปลักษณ์ของเว็บไซต์ เนื่องจากเป็นองค์ประกอบที่มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์ จากงานวิจัยของ Mahmood (2006) Fogg (2003) Lumsden (2009) วงศ์วี ธิวิบูลย์สุข (2545) และวรรณิดา กลิ่นดี (2551)

8.ปริมาณรูปภาพต่อหน้าเว็บเพจ เนื่องจากการแสดงรูปภาพตัวอย่างสินค้าเป็นองค์ประกอบที่มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์ จากงานวิจัยของวรรณิดา กลิ่นดี (2551)

9.ปริมาณข้อมูลต่อหน้าเว็บเพจ เนื่องจากข้อมูลเป็นองค์ประกอบที่มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์ จากงานวิจัยของ Mahmood (2006) และ Fogg (2003)

10.ขนาดและชนิดของตัวอักษร เนื่องจากเป็นองค์ประกอบที่มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์ จากงานวิจัยของ วงศ์วี ธิวิบูลย์สุข (2545) และ ชัชพงศ์ ตั้งมณี และ วรวิทย์ อชิรานนท์ (2551)

ตารางที่ 3.1

องค์ประกอบของเว็บไซต์ที่มีผลกระทบต่อความน่าเชื่อถือของเว็บไซต์ที่รวบรวมมาจากงานวิจัยที่เกี่ยวข้อง เพื่อนำมาพิจารณาคัดเลือกเป็นองค์ประกอบของเว็บไซต์สำหรับใช้ในงานวิจัยนี้

Mahmood (2006)	เทคโนโลยีทางด้านความปลอดภัย ความน่าเชื่อถือในบุคคลที่สาม ความน่าเชื่อถือในชื่อเสียง ความน่าเชื่อถือจากการอ้างอิง รูปลักษณ์ของเว็บไซต์ แหล่งข้อมูล ความถูกต้อง ความคิดเห็นจากผู้มีชื่อเสียง การรักษาข้อมูลส่วนบุคคล
Fogg (2003)	รูปลักษณ์ของเว็บไซต์ โครงสร้างการวางข้อความ เน้นการให้ข้อมูลเกี่ยวกับบริษัท ข้อความมีประโยชน์ ความถูกต้องของข้อมูล ความมีชื่อเสียง โฆษณา
Lumsden (2009)	สัญลักษณ์ของแบรนด์เวอร์ชัน สัญลักษณ์ที่รับรองการใช้ระบบเข้ารหัส รับรองข้อมูลลูกค้าได้รับการควบคุมดูแล ข้อมูลเกี่ยวกับบริษัท ความคิดเห็นจากลูกค้าอื่น การออกแบบเว็บไซต์ ข้อมูลสำหรับติดต่อหาง่าย มีข้อในข้อมูลการติดต่อและมีผู้ช่วยเหลือ

ตารางที่ 3.1 (ต่อ)

องค์ประกอบของเว็บไซต์ที่มีผลกระทบต่อความน่าเชื่อถือของเว็บไซต์ที่รวบรวมมาจากการวิจัยที่เกี่ยวข้อง เพื่อนำมาพิจารณาคัดเลือกเป็นองค์ประกอบของเว็บไซต์สำหรับใช้ในงานวิจัยนี้

ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (2552)	ไม่ไว้วางใจผู้ขาย ไม่เห็นสินค้า ไม่มั่นใจในระบบการชำระเงิน การสั่งซื้อยุ่งยาก ไม่ต้องการส่งข้อมูลบัตรเครดิต กลัวเปลี่ยนสินค้าไม่ได้ ไม่ต้องการรอสินค้า
วงศ์รวี วิบูลย์สุข (2545)	ตำแหน่งการจัดวางขององค์ประกอบสอดคล้องกัน องค์ประกอบแต่ละประเภทมีขนาดเท่ากันทุกหน้า สีของพื้นหลังเป็นสีอ่อน ใช้ตัวอักษร Tahoma ขนาด 13 พิกเซล
รัชพงศ์ ตั้งมณี และ วรวิทย์ อชินานนท์ (2551)	การแสดงตรารับรองของกระทรวงพาณิชย์ การใช้แบบอักษรที่เป็นทางการ
วรรณิดา กลิ่นดี (2551)	การแสดงเครื่องหมายรับประกันการรักษาข้อมูลส่วนบุคคลของผู้ใช้งาน การแสดงชื่อบริษัท ที่อยู่ เบอร์โทรศัพท์ติดต่อ การแสดงเครื่องหมายรับประกันความปลอดภัยในการใช้งาน การแสดงรูปภาพตัวอย่างสินค้า การแสดงราคาสินค้าและค่าขนส่ง การแสดงนโยบายการซื้อและคืนสินค้า การแสดงวิธีการชำระเงิน การแสดงระบบตรวจสอบการสั่งซื้อสินค้า การจัดการข้อผิดพลาดที่เกิดขึ้นบนอินเทอร์เน็ต

3.2 วิธีการวิจัย (Methodology)

ในหัวข้อนี้จะเป็นการอธิบายถึงรูปแบบและวิธีการวิจัย ประชากรและกลุ่มตัวอย่าง เครื่องมือในการวิจัย การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูลสำหรับงานวิจัย

3.2.1 รูปแบบและวิธีการวิจัย

งานวิจัยนี้เป็นงานวิจัยเชิงสำรวจ โดยเก็บรวบรวมข้อมูลด้วยการสร้างเว็บไซต์ที่มีแบบสอบถามออนไลน์เพื่อประเมินระดับความน่าเชื่อถือของเว็บไซต์พาณิชย์อิเล็กทรอนิกส์ โดยข้อมูลจากการประเมินระดับความน่าเชื่อถือของเว็บไซต์ที่ได้จากการตอบแบบสอบถามจะนำมาเป็นเกณฑ์ในการเปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้ของเครื่อง 4 เทคนิค คือ ต้นไม้ตัดสินใจ โครงข่ายประสาทเทียม นาอูฟเบย์ และตรรกศาสตร์คลุมเครือ ในการประเมินระดับความน่าเชื่อถือของเว็บไซต์ เพื่อให้บรรลุวัตถุประสงค์ของงานวิจัยที่ต้องการศึกษาว่าเทคนิคการเรียนรู้ของเครื่องทั้ง 4 เทคนิค มีความสามารถในการประเมินระดับความน่าเชื่อถือของเว็บไซต์ได้แตกต่างจากที่คนประเมินหรือไม่ งานวิจัยนี้จึงมีการกำหนดตัวแปรต้น ตัวแปรตาม และสมมติฐานของงานวิจัยเพื่อให้นำไปสู่การทดสอบนัยสำคัญทางสถิติ และข้อสรุปของงานวิจัย ดังนี้

ตัวแปรต้น คือ เทคนิคการเรียนรู้ของเครื่องที่นำมาใช้เพื่อประเมินระดับความน่าเชื่อถือของเว็บไซต์ โดยมี 4 เทคนิค ได้แก่

1. เทคนิคต้นไม้ตัดสินใจ
2. เทคนิคโครงข่ายประสาทเทียม
3. เทคนิคนาอูฟเบย์
4. เทคนิคตรรกศาสตร์คลุมเครือ

ตัวแปรตาม คือ ระดับความน่าเชื่อถือของเว็บไซต์ที่ได้มาจากการประเมินด้วยเทคนิคแต่ละเทคนิค

สมมติฐานของงานวิจัย

จากวัตถุประสงค์ของงานวิจัยที่ต้องการศึกษาว่าเทคนิคการเรียนรู้ของเครื่องทั้ง 4 เทคนิค มีความสามารถในการประเมินระดับความน่าเชื่อถือของเว็บไซต์ได้แตกต่างจากที่คน

ประเมินหรือไม่ และเทคนิคการเรียนรู้ของเครื่องเทคนิคใดที่มีประสิทธิภาพมากที่สุดในการประเมินระดับความน่าเชื่อถือของเว็บไซต์เมื่อใช้การประเมินด้วยคนเป็นเกณฑ์ในการเปรียบเทียบ งานวิจัยนี้จึงตั้งสมมติฐานของการวิจัย 5 ข้อ ดังนี้คือ

สมมติฐานที่ 1: เทคนิคโครงข่ายประสาทเทียมเป็นเทคนิคที่มีประสิทธิภาพมากที่สุดในการประเมินระดับความน่าเชื่อถือของเว็บไซต์

การตั้งสมมติฐานโดยให้โครงข่ายเทียมเป็นเทคนิคที่มีประสิทธิภาพมากที่สุดในการประเมินระดับความน่าเชื่อถือของเว็บไซต์เนื่องจากโครงข่ายประสาทเทียมมีอัลกอริทึมแบ็กพรอพาเกชัน (Backpropagation Algorithm) ที่มีประสิทธิภาพในการเรียนรู้และสามารถปรับค่าน้ำหนักเพื่อให้สามารถทำนายกลุ่มได้อย่างถูกต้อง และการปรับค่าน้ำหนักสามารถปรับได้ด้วยพารามิเตอร์หลายค่า โครงข่ายประสาทเทียมจึงน่าจะเป็นเทคนิคที่มีประสิทธิภาพมากที่สุดในการประเมินระดับความน่าเชื่อถือของเว็บไซต์

สมมติฐานที่ 2: เทคนิคต้นไม้ตัดสินใจ มีความสามารถในการประเมินระดับความน่าเชื่อถือของเว็บไซต์ได้ไม่แตกต่างจากที่คนประเมิน

สมมติฐานที่ 3: เทคนิคโครงข่ายประสาทเทียม มีความสามารถในการประเมินระดับความน่าเชื่อถือของเว็บไซต์ได้ไม่แตกต่างจากที่คนประเมิน

สมมติฐานที่ 4: เทคนิคนาอ็พเบย์ มีความสามารถในการประเมินระดับความน่าเชื่อถือของเว็บไซต์ได้ไม่แตกต่างจากที่คนประเมิน

สมมติฐานที่ 5: เทคนิคตรรกศาสตร์คลุมเครือ มีความสามารถในการประเมินระดับความน่าเชื่อถือของเว็บไซต์ได้ไม่แตกต่างจากที่คนประเมิน

3.2.2 ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในงานวิจัยนี้ ได้แก่ ผู้ที่มีประสบการณ์ในการใช้อินเทอร์เน็ต มีการศึกษาตั้งแต่ชั้นมัธยมปลายเป็นต้นไป เพื่อให้เข้าใจข้อความในเว็บไซต์ขายเสื้อผ้าภาษาอังกฤษ

ที่นำมาใช้ในงานวิจัย ไม่จำกัดเพศ ไม่จำกัดประสบการณ์ในการซื้อสินค้าบนระบบพาณิชย์อิเล็กทรอนิกส์ และเป็นคนไทยเนื่องจากข้อความในแบบสอบถามเป็นภาษาไทย การที่งานวิจัยนี้เลือกเว็บไซต์ขายเสื้อผ้าสำเร็จรูปมาเป็นเครื่องมือในการทำวิจัย เนื่องจากรายงานผลการสำรวจกลุ่มผู้ใช้อินเทอร์เน็ตในประเทศไทย โดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (2552) พบว่าเสื้อผ้าสำเร็จรูปเป็นสินค้าที่มีคนนิยมซื้อผ่านทางอินเทอร์เน็ตเป็นอันดับสาม รองลงมาจากหนังสือ และการส่งของบริการ และจากข้อมูลผลการวิจัยของ วรณิดา กลิ่นดี (2551) พบว่าเสื้อผ้าสำเร็จรูปเป็นสินค้าที่มีคนนิยมซื้อผ่านทางอินเทอร์เน็ตเป็นอันดับสาม รองลงมาจากหนังสือ และซีดี หรือดีวีดี เพลง ภาพยนต์ แต่เนื่องจากยังไม่มีงานวิจัยใดที่ประเมินระดับความน่าเชื่อถือของเว็บไซต์ที่ขายเสื้อผ้าสำเร็จรูปผ่านทางระบบอิเล็กทรอนิกส์ อีกทั้งเสื้อผ้าสำเร็จรูปมีมูลค่าของสินค้าสูงกว่าหนังสือและซีดี หรือดีวีดี เพลง ภาพยนต์ การซื้อเสื้อผ้าสำเร็จรูปผ่านทางอินเทอร์เน็ตจึงมีความเสี่ยงมากกว่า งานวิจัยนี้จึงเลือกเว็บไซต์ขายเสื้อผ้าสำเร็จรูปผ่านทางระบบอิเล็กทรอนิกส์มาเป็นเครื่องมือในการทำวิจัย สำหรับการเลือกเว็บไซต์ที่เป็นภาษาอังกฤษเนื่องจากทำให้เว็บไซต์ที่นำมาใช้ในงานวิจัยมีความหลากหลายและทำให้สามารถคัดเลือกเว็บไซต์ได้โดยไม่มีข้อจำกัดในเรื่องประเทศ นอกจากนี้เว็บไซต์ขายเสื้อผ้าสำเร็จรูปที่เป็นภาษาอังกฤษจะมีองค์ประกอบของเว็บไซต์ เช่น การรักษาความปลอดภัยข้อมูล บัตรเครดิต และการรักษาข้อมูลส่วนบุคคล ที่เป็นองค์ประกอบของเว็บไซต์ที่สำคัญต่อความน่าเชื่อถือของเว็บไซต์ตามข้อมูลที่ได้ศึกษาจากงานวิจัยเกี่ยวข้อง งานวิจัยนี้จึงเลือกเว็บไซต์ที่เป็นภาษาอังกฤษมาใช้ในการวิจัย สำหรับการที่ผู้ตอบแบบสอบถามเป็นคนไทยแต่ให้ตอบแบบสอบถามเกี่ยวกับเว็บไซต์ที่เป็นภาษาอังกฤษเนื่องจากเว็บไซต์ขายเสื้อผ้าสำเร็จรูปที่เป็นภาษาอังกฤษมีองค์ประกอบของเว็บไซต์ที่มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์ เช่น การรักษาความปลอดภัยข้อมูล บัตรเครดิต และการรักษาข้อมูลส่วนบุคคล มากกว่าเว็บไซต์ขายเสื้อผ้าสำเร็จรูปที่เป็นภาษาไทย จึงเลือกใช้เว็บไซต์เสื้อผ้าสำเร็จรูปที่เป็นภาษาอังกฤษ ทั้งนี้เพื่อนำความคิดเห็นของคนไทยจากงานวิจัยนี้มาเป็นแนวทางสำหรับการพัฒนาปรับปรุงเว็บไซต์ขายเสื้อผ้าสำเร็จรูปของคนไทยให้มีความน่าเชื่อถือมากยิ่งขึ้น

งานวิจัยนี้ทำการเก็บข้อมูลจากแบบสอบถามระหว่างวันที่ 1-30 เมษายน 2553 จากรายงานผลการสำรวจกลุ่มผู้ใช้อินเทอร์เน็ตในประเทศไทย ปี 2552 โดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ พบว่าในปี 2551 ประเทศไทยมีจำนวนผู้ใช้อินเทอร์เน็ตจำนวน 16 ล้านคน และจากการสำรวจการมีการใช้เครื่องมือ/อุปกรณ์เทคโนโลยีสารสนเทศ ปี 2546 โดยสำนักงานสถิติแห่งชาติ พบว่าสัดส่วนของผู้ใช้อินเทอร์เน็ตที่มีระดับการศึกษาตั้งแต่

มัธยมศึกษาตอนปลายขึ้นไปมีจำนวนร้อยละ 76 ของจำนวนผู้ใช้อินเทอร์เน็ตทั้งหมด เมื่อนำมาคำนวณจำนวนประชากรผู้ใช้อินเทอร์เน็ตที่มีระดับการศึกษาตั้งแต่มัธยมศึกษาตอนปลายขึ้นไปจะได้ 12 ล้านคน และตามวิธีการคำนวณขนาดกลุ่มตัวอย่างของ Taro Yamane ที่ความคลาดเคลื่อนของการสุ่มตัวอย่างร้อยละ 5 ขนาดของกลุ่มตัวอย่างคือจำนวน 400 คน (1973 อ้างถึงใน บุญธรรม กิจปรีดาบริสุทธิ์, 2546) แต่เนื่องจากข้อจำกัดที่เว็บไซต์ขายเสื้อผ้าที่นำมาใช้ในงานวิจัยเป็นเว็บไซต์ภาษาอังกฤษ แบบสอบถามจึงได้รับการตอบกลับในจำนวนที่จำกัด โดยงานวิจัยนี้มีผู้ตอบแบบสอบถามออนไลน์จำนวน 103 คน ในจำนวนนี้มีแบบสอบถามที่ข้อมูลครบถ้วนสมบูรณ์และสามารถนำมาใช้เพื่อการวิจัยจำนวน 90 ชุด และมีแบบสอบถามที่ข้อมูลไม่สมบูรณ์และไม่สามารถนำมาใช้เพื่อการวิจัยจำนวน 13 ชุด ซึ่งเมื่อนำแบบสอบถามที่ครบถ้วนสมบูรณ์จำนวน 90 ชุด มาคำนวณหาความคลาดเคลื่อนของการสุ่มตัวอย่างโดยเปรียบเทียบกับจำนวนประชากร 12 ล้านคน พบว่าค่าความคลาดเคลื่อนของการสุ่มตัวอย่างของงานวิจัยนี้คิดเป็นร้อยละ 10.5

3.2.3 เครื่องมือในการวิจัย

ในหัวข้อนี้จะกล่าวถึงเครื่องมือในการวิจัย ซึ่งประกอบด้วย 2 ส่วน โดยส่วนแรก คือ เว็บไซต์พาณิชย์อิเล็กทรอนิกส์ที่คัดเลือกมาสำหรับใช้ในการประเมินความน่าเชื่อถือของเว็บไซต์ จะอธิบายถึงวิธีการคัดเลือกเว็บไซต์ และส่วนที่สอง คือแบบสอบถามออนไลน์ ที่มีรายละเอียดของแบบสอบถาม และวิธีการเก็บคะแนนความน่าเชื่อถือของเว็บไซต์

1. เว็บไซต์ที่คัดเลือกมาสำหรับใช้ในการประเมินความน่าเชื่อถือของเว็บไซต์

งานวิจัยนี้ใช้เว็บไซต์พาณิชย์อิเล็กทรอนิกส์จำนวน 50 เว็บไซต์ ที่ได้คัดเลือกจากเว็บไซต์ที่มีอยู่จริงบนอินเทอร์เน็ตเป็นเครื่องมือในการทำวิจัย โดยเลือกเว็บไซต์พาณิชย์อิเล็กทรอนิกส์ที่ขายเสื้อผ้าสำเร็จรูปที่ข้อมูลในเว็บไซต์เป็นภาษาอังกฤษ และแสดงข้อมูลตามความเป็นจริงที่บริษัทได้นำเสนอมาเพื่อขายสินค้าบนระบบพาณิชย์อิเล็กทรอนิกส์ วิธีคัดเลือกเว็บไซต์ในงานวิจัยนี้ทำโดยการคัดเลือกคำหลัก (keyword) ที่เกี่ยวข้องกับเว็บไซต์ขายเสื้อผ้าสำเร็จรูปมากที่สุดจำนวน 5 คำหลัก ในการค้นหาเว็บไซต์จากเครื่องมือค้นหา (search engine) โดยคำหลักที่ใช้ ได้แก่ buy clothing, clothing online, shopping clothing, online clothing

stores และ clothes shop และจะเลือกเฉพาะเว็บไซต์ 100 อันดับแรกจากการใช้คำหลักในการค้นหาแต่ละครั้งเนื่องจากอันดับที่มากกว่านี้จะตรงกับคำที่ค้นหาน้อยลง จากนั้นเว็บไซต์ที่เลือกมาทั้งหมดจะถูกนำมาคัดเลือกอีกครั้งว่าเป็นเว็บไซต์ที่เกี่ยวข้องกับการขายเสื้อผ้าสำเร็จรูป สามารถใช้งานได้จริง และไม่ซ้ำกับเว็บไซต์ที่พบในคำค้นหาก่อนหน้านี้ ซึ่งจากการคัดเลือกทำให้ได้เว็บไซต์สำหรับใช้ในงานวิจัยทั้งหมด 50 เว็บไซต์ โดยทำการเก็บลิงค์ของเว็บไซต์ไว้ในฐานข้อมูล ซึ่งโปรแกรมฐานข้อมูลที่ใช้สำหรับงานวิจัยนี้คือ MySQL เวอร์ชัน 5.0 และโปรแกรมที่ใช้สำหรับการจัดการฐานข้อมูลคือ phpMyAdmin เวอร์ชัน 3.2.5

2. แบบสอบถามออนไลน์

งานวิจัยนี้ใช้แบบสอบถามออนไลน์เป็นเครื่องมือในการเก็บรวบรวมข้อมูลสำหรับการวิเคราะห์ระดับความน่าเชื่อถือของเว็บไซต์พาณิชย์อิเล็กทรอนิกส์ที่เกี่ยวข้องกับการขายเสื้อผ้าสำเร็จรูป โดยข้อความในแบบสอบถามแบ่งออกเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 เป็นแบบสอบถามเกี่ยวกับข้อมูลส่วนบุคคลของผู้ตอบแบบสอบถาม มีจำนวน 6 ข้อ คือ

1. อายุของคุณอยู่ในช่วง
2. เพศ
3. ระดับการศึกษา
4. คุณมีประสบการณ์ในการใช้อินเทอร์เน็ตเป็นระยะเวลา
5. คุณใช้อินเทอร์เน็ตสัปดาห์ละกี่ชั่วโมง
6. คุณเคยใช้บริการหรือซื้อสินค้าผ่านทางเว็บไซต์หรือไม่

ตัวอย่างของหน้าจอแบบสอบถามส่วนที่ 1 ดังแสดงในภาพที่ 3.1

ส่วนที่ 2 เป็นการแสดงรายการองค์ประกอบหลัก 10 องค์ประกอบ ที่ได้คัดเลือกมาโดยการศึกษางานวิจัยที่เกี่ยวข้อง ได้แก่

- 1.การรักษาความปลอดภัยข้อมูลบัตรเครดิต
- 2.การรักษาข้อมูลส่วนบุคคล
- 3.ข้อมูลบริษัท
- 4.ข้อมูลในการติดต่อ
- 5.การส่งสินค้าและคืนสินค้า

- 6.การแสดงราคาสินค้าและค่าขนส่ง
- 7.รูปลักษณะของเว็บไซต์
- 8.ปริมาณรูปภาพต่อหนึ่งหน้าเว็บเพจ
- 9.ปริมาณข้อมูลต่อหนึ่งหน้าเว็บเพจ
- 10.ขนาดและชนิดของตัวอักษร

โดยให้ผู้ตอบแบบสอบถามให้คะแนนความน่าเชื่อถือของแต่ละองค์ประกอบของเว็บไซต์ และให้คะแนนว่าแต่ละองค์ประกอบของเว็บไซต์มีความสำคัญต่อความน่าเชื่อถือของเว็บไซต์มากน้อยเพียงไร คะแนนความน่าเชื่อถือขององค์ประกอบของเว็บไซต์มีคะแนน 0 – 20 คะแนน โดยที่ 0 คะแนน หมายถึง องค์ประกอบของเว็บไซต์ไม่มีความน่าเชื่อถือ และ 20 คะแนน หมายถึง องค์ประกอบของเว็บไซต์มีความน่าเชื่อถือมากที่สุด คะแนนความสำคัญขององค์ประกอบของเว็บไซต์มีคะแนน 0 – 10 คะแนน โดยที่ 0 หมายถึง องค์ประกอบของเว็บไซต์ไม่มีความสำคัญต่อการวัดระดับความน่าเชื่อถือของเว็บไซต์ และ 10 หมายถึง องค์ประกอบของเว็บไซต์มีความสำคัญต่อการวัดระดับความน่าเชื่อถือของเว็บไซต์มากที่สุด การกำหนดให้ช่วงคะแนนของความน่าเชื่อถือขององค์ประกอบของเว็บไซต์แตกต่างจากช่วงคะแนนความสำคัญขององค์ประกอบเนื่องจากคะแนนความน่าเชื่อถือขององค์ประกอบที่ได้จากการตอบแบบสอบถามจะถูกนำมาใช้เปรียบเทียบกับคะแนนความน่าเชื่อถือขององค์ประกอบที่ทำนายได้จากเทคนิคการเรียนรู้ของเครื่องจึงต้องการช่วงคะแนนที่มีความละเอียดเพื่อให้การเปรียบเทียบมีความแม่นยำมากยิ่งขึ้น

ในการตอบแบบสอบถามแต่ละครั้ง เว็บไซต์จำนวน 2 เว็บไซต์ จาก 50 เว็บไซต์ ที่ได้คัดเลือกไว้และเก็บอยู่ในฐานข้อมูล จะถูกสุ่มเลือกขึ้นมาโดยระบบ เพื่อให้ผู้ตอบแบบสอบถามประเมินระดับความน่าเชื่อถือของเว็บไซต์ ตัวอย่างของหน้าจอแบบสอบถามดังแสดงในภาพที่ 3.1 และ 3.2

ภาพที่ 3.1

หน้าจอของแบบสอบถามออนไลน์ ส่วนที่ 1

ส่วนที่1 ข้อมูลเกี่ยวกับผู้ตอบแบบสอบถาม

กรุณาตอบคำถามต่อไปนี้ โดยเลือกเพียงคำตอบเดียวสำหรับคำถามในแต่ละข้อ

1. อายุของคุณอยู่ในช่วง
☐ ต่ำกว่า 20 ปี ☐ 20 - 24 ปี ☐ 25 - 34 ปี ☐ 35 - 44 ปี ☐ 45 ปีขึ้นไป
2. เพศ
☐ หญิง ☐ ชาย
3. ระดับการศึกษา
☐ มัธยมศึกษาตอนปลาย
☐ กำลังศึกษาระดับปริญญาตรี
☐ จบการศึกษาระดับปริญญาตรี
☐ จบการศึกษาระดับปริญญาโท
4. คุณมีประสบการณ์ในการใช้อินเทอร์เน็ตเป็นระยะเวลา
☐ ต่ำกว่า 1 ปี ☐ 1 - 2 ปี ☐ 3 - 4 ปี ☐ 5 ปีขึ้นไป
5. คุณใช้อินเทอร์เน็ตสัปดาห์ละกี่ชั่วโมง
☐ น้อยกว่า 1 ชั่วโมง ☐ 1 - 5 ชั่วโมง ☐ 6 - 10 ชั่วโมง
☐ 11 - 15 ชั่วโมง ☐ 16 - 20 ชั่วโมง ☐ มากกว่า 20 ชั่วโมง
6. คุณเคยใช้บริการหรือซื้อสินค้าผ่านทางเว็บไซต์หรือไม่
☐ เคย ☐ ไม่เคย

ส่งข้อมูล

ล้างข้อมูล

ภาพที่ 3.2

หน้าจอของแบบสอบถามออนไลน์ ส่วนที่ 2

ส่วนที่2 แบบสอบถามระดับความน่าเชื่อถือของเว็บไซต์

กรุณาให้คะแนนความน่าเชื่อถือของเว็บไซต์ 2 เว็บไซต์ โดยพิจารณาความน่าเชื่อถือของเว็บไซต์ จากความน่าเชื่อถือขององค์ประกอบของเว็บไซต์ ความน่าเชื่อถือขององค์ประกอบของเว็บไซต์จะมีคะแนน อยู่ระหว่าง 0 - 20 คะแนน โดยที่

0 คะแนน หมายถึง องค์ประกอบนี้ของเว็บไซต์นี้ไม่มีความน่าเชื่อถือ
20 คะแนน หมายถึง องค์ประกอบนี้ของเว็บไซต์นี้มีความน่าเชื่อถือมากที่สุด

กรุณาให้ค่าความสำคัญขององค์ประกอบที่มีผลต่อการวัดระดับความน่าเชื่อถือของเว็บไซต์ โดยค่าความสำคัญขององค์ประกอบจะอยู่ระหว่าง 0-10 โดยที่

0 หมายถึง องค์ประกอบนี้ไม่มีความสำคัญต่อการวัดระดับความน่าเชื่อถือของเว็บไซต์
10 หมายถึง องค์ประกอบนี้มีความสำคัญต่อการวัดระดับความน่าเชื่อถือของเว็บไซต์มากที่สุด

เว็บไซต์ที่1: [onepost](#)

เว็บไซต์ที่2: [boredof](#)

องค์ประกอบของความน่าเชื่อถือของเว็บไซต์	เว็บไซต์ที่1	เว็บไซต์ที่2	ค่าความสำคัญ
1. การรักษาความปลอดภัยของข้อมูลบัตรเครดิต			
2. การรักษาข้อมูลส่วนบุคคล ได้แก่ ชื่อ ที่อยู่			
3. การแสดงข้อมูลบริษัท			
4. การแสดงข้อมูลการติดต่อ			
5. การส่งสินค้าและการคืนสินค้า			
6. การแสดงราคาสินค้าและค่าขนส่ง			
7. รูปแบบของเว็บไซต์			
8. ปริมาณรูปภาพต่อหนึ่งหน้าเว็บเพจ			
9. ปริมาณข้อมูลต่อหนึ่งหน้าเว็บเพจ			
10. ขนาดและชนิดของตัวอักษร			

ส่งข้อมูล

ล้างข้อมูล

3.2.4 การเก็บรวบรวมข้อมูล

งานวิจัยนี้เก็บรวบรวมข้อมูลด้วยการสร้างเว็บไซต์ที่มีแบบสอบถามออนไลน์สำหรับใช้ในการประเมินระดับความน่าเชื่อถือของเว็บไซต์พาณิชย์อิเล็กทรอนิกส์ จากนั้นทำการประกาศข้อความในเว็บไซต์ www.pantip.com และ www.vcharkarn.com เพื่อให้ผู้ที่เห็นประกาศช่วยเข้ามาประเมินระดับความน่าเชื่อถือของเว็บไซต์พาณิชย์อิเล็กทรอนิกส์ในเว็บไซต์ที่ได้สร้างไว้ นอกจากนี้ ผู้วิจัยยังได้ทำการส่งอีเมลถึงนักศึกษาปริญญาโทภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยธรรมศาสตร์ ให้ช่วยเข้ามาประเมินระดับความน่าเชื่อถือของเว็บไซต์พาณิชย์อิเล็กทรอนิกส์ในเว็บไซต์ที่ได้สร้างไว้ และขอให้นักศึกษาปริญญาโทช่วยส่งอีเมลต่อถึงคนอื่นๆ ให้ช่วยเข้ามาประเมินระดับความน่าเชื่อถือของเว็บไซต์ด้วยเช่นกัน ข้อมูลจากการตอบแบบสอบถามแต่ละครั้งจะถูกเก็บไว้ในฐานข้อมูล และคะแนนความน่าเชื่อถือของแต่ละองค์ประกอบของเว็บไซต์ที่ได้จากการตอบแบบสอบถาม จะถูกนำมารวมกันและหาค่าเฉลี่ยเป็นคะแนนความน่าเชื่อถือของเว็บไซต์ที่มีคะแนนตั้งแต่ 0-20 คะแนน และนำคะแนนความน่าเชื่อถือของเว็บไซต์ มาจัดกลุ่มเป็น 5 ระดับ คือระดับ 0-4 โดยใช้การแจกแจงข้อมูลคะแนนความน่าเชื่อถือของเว็บไซต์ทุกเว็บไซต์ ด้วยค่าเฉลี่ยเลขคณิต ในกรณีที่ข้อมูลมีลักษณะสมมาตร หรือใช้การแจกแจงข้อมูลด้วยค่ามัธยฐาน ในกรณีที่ข้อมูลมีลักษณะไม่สมมาตร (กัลยา วานิชย์บัญชา, 2538)

เมื่อเก็บรวบรวมข้อมูลจากแบบสอบถามและนำคะแนนความน่าเชื่อถือของเว็บไซต์จากการตอบแบบสอบถามมาทำการจัดกลุ่มแล้ว ขั้นตอนต่อไปคือการดึงข้อมูลความน่าเชื่อถือขององค์ประกอบของเว็บไซต์ด้วยโปรแกรมเว็บครอว์เลอร์ โดยต้องสร้างเกณฑ์เพื่อให้ข้อมูลความน่าเชื่อถือของเว็บไซต์ที่เว็บครอว์เลอร์ดึงมาได้แบ่งออกเป็น 5 ระดับ ทั้งนี้เพื่อให้สามารถเปรียบเทียบกับระดับความน่าเชื่อถือของเว็บไซต์จากการตอบแบบสอบถามได้

1. การดึงข้อมูลองค์ประกอบของเว็บไซต์ด้วยโปรแกรมเว็บครอว์เลอร์

เว็บครอว์เลอร์ (Web Crawler) หรือที่เรียกว่า เว็บสไปเดอร์ (Web Spider) เป็นโปรแกรมที่ใช้สำหรับดาวน์โหลดข้อมูลจากอินเทอร์เน็ต โดยสามารถดึงข้อมูลที่มีอยู่ในเว็บไซต์มาเก็บไว้ที่ฮาร์ดไดรฟ์ของเครื่องคอมพิวเตอร์เพื่อนำข้อมูลมาดูหรือใช้งานในภายหลังได้ (Dustin Boswell, 2003) การดึงข้อมูลของเว็บครอว์เลอร์จะเริ่มต้นจากที่อยู่ของเว็บไซต์ (Uniform Resource Locator: URL) ของเว็บเพจแรก โดยจะทำการดึงข้อมูลที่มีอยู่ในเว็บเพจแรก และดึง

ข้อมูลที่อยู่ของเว็บไซต์ที่มีอยู่ในเว็บเพจแรกมาอยู่ในรายชื่อของเว็บไซต์ที่จะต้องทำการดึงข้อมูลในลำดับต่อไป เมื่อดึงข้อมูลและที่อยู่ข้อมูลของเว็บเพจแรกเสร็จแล้ว จึงทำการดึงข้อมูลและที่อยู่ของเว็บไซต์ตามรายชื่อของเว็บไซต์ที่จะต้องทำการดึงข้อมูล (Cho, Garcia-Molina & Page, 1998)

งานวิจัยนี้ใช้โปรแกรมเว็บครอว์เลอร์ ที่พัฒนาจากโปรแกรมไพธอน (Python) เวอร์ชัน 3.1.2 ในการดึงข้อมูลองค์ประกอบของเว็บไซต์ 50 เว็บไซต์ ที่ได้คัดเลือกไว้และเก็บอยู่ในฐานข้อมูล ตามเกณฑ์ที่กำหนด โดยโปรแกรมเว็บครอว์เลอร์จะทำการดึงข้อมูลองค์ประกอบของเว็บไซต์ที่ 1-5 ได้แก่ (1)การรักษาความปลอดภัยข้อมูลบัตรเครดิต (2)การรักษาข้อมูลส่วนบุคคล (3)ข้อมูลบริษัท (4)ข้อมูลในการติดต่อ และ (5)การส่ง/คืนสินค้า การที่กำหนดให้เว็บครอว์เลอร์ดึงข้อมูลเฉพาะองค์ประกอบที่ 1-5 เนื่องจากทั้ง 5 องค์ประกอบ เป็นองค์ประกอบที่สามารถสร้างเกณฑ์เพื่อแบ่งความน่าเชื่อถือขององค์ประกอบออกเป็น 5 ระดับได้ ยกตัวอย่างเช่น องค์ประกอบที่ (1)การรักษาความปลอดภัยข้อมูลบัตรเครดิต สามารถนำสัญลักษณ์การรักษาข้อมูลบัตรเครดิตของแบรนด์ที่แตกต่างกันที่พบในหน้าเว็บเพจมาเป็นเครื่องมือในการจำแนกระดับความน่าเชื่อถือของเว็บไซต์ได้ ทั้งนี้เนื่องจากแบรนด์ที่เป็นที่รู้จักและมีความน่าเชื่อถือสูงจะทำให้เว็บไซต์มีความน่าเชื่อถือสูงด้วย สำหรับองค์ประกอบที่ 6-10 ยังไม่มีเกณฑ์ที่เหมาะสมที่จะนำมาใช้ในการแบ่งระดับความน่าเชื่อถือขององค์ประกอบออกเป็น 5 ยกตัวอย่างเช่น องค์ประกอบที่ (7) ระบุลักษณะของเว็บไซต์ ยังไม่มีเกณฑ์ที่เหมาะสมที่จะจำแนกระดับความน่าเชื่อถือของรูปลักษณะของเว็บไซต์ออกเป็น 5 ระดับ

2. เกณฑ์ที่กำหนดให้เว็บครอว์เลอร์ในการดึงข้อมูลองค์ประกอบของเว็บไซต์

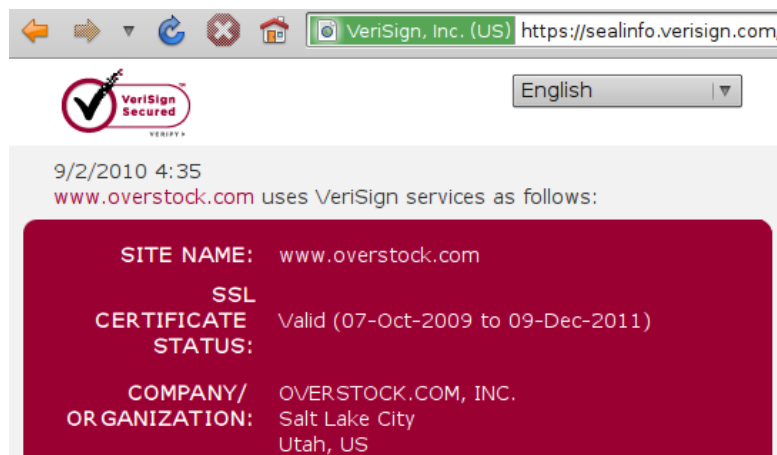
เพื่อให้เว็บครอว์เลอร์สามารถดึงข้อมูลระดับความน่าเชื่อถือขององค์ประกอบของเว็บไซต์มาเพื่อเปรียบเทียบกับระดับความน่าเชื่อถือของเว็บไซต์ที่ได้จากการตอบแบบสอบถาม จำเป็นต้องสร้างเกณฑ์ที่เหมาะสมสำหรับแต่ละองค์ประกอบของเว็บไซต์ โดยเกณฑ์ที่กำหนดขึ้นเพื่อให้เว็บครอว์เลอร์ใช้สำหรับดึงข้อมูลระดับความน่าเชื่อถือของเว็บไซต์จากองค์ประกอบของเว็บไซต์ที่ 1-5 มีดังต่อไปนี้

องค์ประกอบที่ 1: การรักษาความปลอดภัยข้อมูลบัตรเครดิต ใช้สัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตจากบริษัทที่เป็นบุคคลที่ 3 (third party) เป็นเกณฑ์ในการแบ่งระดับความน่าเชื่อถือของการรักษาความปลอดภัยของข้อมูลบัตรเครดิตออกเป็น 5 ระดับ โดยถ้า

ในเว็บไซค์มีสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิต เมื่อใช้เมาส์กดเลือกที่สัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิต แถบบอกที่อยู่ของเว็บเพจที่ปรากฏขึ้นจะนำหน้าด้วย https:// และในเว็บเพจจะมีรายละเอียดการจดทะเบียนการรักษาความปลอดภัยของเว็บไซค์ ตัวอย่างเว็บเพจของสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตดังแสดงในภาพที่ 3.3 งานวิจัยนี้จะให้สัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตของบริษัท 5 บริษัท ที่มีส่วนแบ่งตลาดรวมกันร้อยละ 90.69 เป็นแบรนด์หลัก โดยส่วนแบ่งตลาดของแต่ละบริษัทสำหรับสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิต ที่มาจากข้อมูลในเว็บเพจ <https://ssl.netcraft.com/ssl-sample-report//CMatch/certs>

ภาพที่ 3.3

ตัวอย่างเว็บเพจของสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิต



ตารางที่ 3.2

ส่วนแบ่งตลาดของสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิต
และจำนวนเว็บไซต์ที่ใช้สัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิต

แบรนด์	มีนาคม 2553 (ร้อยละ)	2549	2550	2551	2552
		มกราคม	มกราคม	มกราคม	มกราคม
VeriSign	47.52	222008	381757	443606	482037
Go Daddy	23.42	29209	58623	131495	237519
Comodo	15.44	56541	70249	122204	156602
Network Solutions	2.55	1768	7356	18448	25844
GlobalSign	1.76	0	2348	7297	17897
อื่นๆ	9.31	152851	47179	70958	94402

การใช้แบรนด์หลักของสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตเป็นเครื่องมือในการจำแนกระดับความน่าเชื่อถือของเว็บไซต์เนื่องจากแบรนด์หลักเป็นแบรนด์ที่ได้รับการยอมรับว่ามีความน่าเชื่อถือจึงทำให้เว็บไซต์ที่มีสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตของแบรนด์หลักมีความน่าเชื่อถือด้วย ตัวอย่างเกณฑ์ที่กำหนดให้เว็บเบราว์เซอร์ในการดึงข้อมูลการรักษาความปลอดภัยข้อมูลบัตรเครดิต ได้แก่ ถ้าพบสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตของแบรนด์หลักในหน้าแรกของเว็บไซต์ให้ระดับความน่าเชื่อถือของเว็บไซต์เป็นระดับ 4 และถ้าไม่พบสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตในเว็บไซต์ให้ระดับความน่าเชื่อถือของเว็บไซต์เป็นระดับ 0 เกณฑ์ที่กำหนดให้เว็บเบราว์เซอร์ในการดึงข้อมูลการรักษาความปลอดภัยข้อมูลบัตรเครดิตทั้ง 5 ระดับ ดังแสดงในตารางที่ 3.3 โดยคำหลักที่แสดงให้เว็บเบราว์เซอร์รู้ว่าในเว็บไซตมีสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิต คือ คำว่า https:// คำที่แสดงให้เว็บเบราว์เซอร์รู้ว่าเป็นแบรนด์หลัก ได้แก่ คำว่า verisign ที่แสดงว่าเป็นของแบรนด์ VeriSign ปรากฏอยู่หลัง https:// หรือ คำว่า godaddy ที่แสดงว่าเป็นของแบรนด์ Go Daddy ปรากฏอยู่หลัง https:// เป็นต้น

ตารางที่ 3.3

เกณฑ์ที่กำหนดให้เว็บครอว์เลอร์ในการดึงข้อมูลการรักษาความปลอดภัยข้อมูลบัตรเครดิต

ระดับ	องค์ประกอบที่ 1: การรักษาความปลอดภัยข้อมูลบัตรเครดิต
4	พบสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตของแบรนด์หลักหน้าแรก
3	พบสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตของแบรนด์หลักในหน้าอื่นที่ไม่ใช่หน้าแรก
2	พบสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตในหน้าแรก
1	พบสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตในหน้าอื่นที่ไม่ใช่หน้าแรก
0	ไม่พบสัญลักษณ์การรับรองการรักษาข้อมูลบัตรเครดิตในเว็บไซต์

องค์ประกอบที่ 2: การรักษาข้อมูลส่วนบุคคล ใช้ปริมาณข้อมูลที่พบในหน้าเว็บเพจของนโยบายการรักษาข้อมูลส่วนบุคคลเป็นเกณฑ์ในการแบ่งระดับความน่าเชื่อถือของการรักษาข้อมูลส่วนบุคคลออกเป็น 5 ระดับ เนื่องจากการแสดงนโยบายการรักษาข้อมูลส่วนบุคคลที่มีรายละเอียดมากจะทำให้เว็บไซต์ได้รับความน่าเชื่อถือสูง และการแสดงนโยบายการรักษาข้อมูลส่วนบุคคลที่มีรายละเอียดน้อยจะทำให้เว็บไซต์ได้รับความน่าเชื่อถือน้อย โดยถ้าปริมาณข้อมูลที่พบในหน้าเว็บเพจของนโยบายการรักษาข้อมูลส่วนบุคคลมีข้อมูลมากกว่าหนึ่งหน้ากระดาษ หรือ 2,000 ตัวอักษร โดยประมาณ จะกำหนดให้ระดับความน่าเชื่อถือของการรักษาข้อมูลส่วนบุคคลเป็นระดับ 4 เกณฑ์ที่กำหนดให้เว็บครอว์เลอร์ในการดึงข้อมูลการรักษาข้อมูลส่วนบุคคลทั้ง 5 ระดับ ดังแสดงในตารางที่ 3.4 โดยคำหลักที่แสดงให้เว็บครอว์เลอร์รู้ว่าเป็นเว็บเพจของการรักษาข้อมูลส่วนบุคคล ได้แก่ คำว่า Privacy Policy , Legal Policy หรือ Security Policy เป็นต้น

ตารางที่ 3.4

เกณฑ์ที่กำหนดให้เว็บครอว์เลอร์ในการดึงข้อมูลการรักษาข้อมูลส่วนบุคคล

ระดับ	องค์ประกอบที่ 2: การรักษาข้อมูลส่วนบุคคล
4	ในเพจของการรักษาข้อมูลส่วนบุคคลมีจำนวนตัวอักษร 2,000 อักขรขึ้นไป
3	ในเพจของการรักษาข้อมูลส่วนบุคคลมีจำนวนตัวอักษร 1,500-2,000 อักขร
2	ในเพจของการรักษาข้อมูลส่วนบุคคลมีจำนวนตัวอักษร 1,000-1,500 อักขร
1	ในเพจของการรักษาข้อมูลส่วนบุคคลมีจำนวนตัวอักษร 500-1,000 อักขร
0	ไม่พบการรักษาข้อมูลส่วนบุคคล

องค์ประกอบที่ 3: ข้อมูลบริษัท ใช้การแสดงข้อมูลของบริษัทในการแบ่งระดับความน่าเชื่อถือของบริษัทออกเป็น 5 ระดับ และใช้ปริมาณสาขาของบริษัทเป็นปัจจัยหนึ่งในการแบ่งระดับความน่าเชื่อถือของบริษัท เนื่องจากจำนวนสาขาที่มีมากขึ้นเป็นปัจจัยหนึ่งที่มีผลต่อยอดขายและความน่าเชื่อถือของบริษัท โดยบริษัทที่มียอดขายสูงมักจะมีสาขาเป็นจำนวนมากเช่นกัน ดังในรายงานส่วนแบ่งตลาดของเสื้อผ้าสำเร็จรูปของประเทศอังกฤษ จากเว็บเพจ www.tutor2u.net/business/marketing/market_analysis_marketshare_intro.asp ซึ่งแบรนด์ที่มียอดขายสูงจะมีสาขามากกว่า 20 สาขาขึ้นไป ดังนั้นจึงกำหนดให้บริษัทที่เป็นบริษัทมหาชนหรือมีสาขามากกว่า 20 สาขา มีระดับความน่าเชื่อถือของบริษัทเป็นระดับ 4 เกณฑ์ที่กำหนดให้เว็บครอว์เลอร์ในการดึงข้อมูลบริษัททั้ง 5 ระดับ ดังแสดงในตารางที่ 3.5 โดยคำหลักที่แสดงให้เว็บครอว์เลอร์รู้ว่าเป็นบริษัทมหาชน เช่น Public Limited Company หรือ PCL เป็นต้น คำหลักที่แสดงว่ามีรายงานประจำปี เช่น Annual Report หรือ Financial Statement เป็นต้น คำหลักที่แสดงว่ามีข้อมูลบริษัท เช่น Company Profile หรือ About Company เป็นต้น

ตารางที่ 3.5

เกณฑ์ที่กำหนดให้เว็บครอว์เลอร์ในการดึงข้อมูลบริษัท

ระดับ	องค์ประกอบที่ 3: ข้อมูลบริษัท
4	เป็นบริษัทมหาชน หรือ มีจำนวนสาขา 20 สาขาขึ้นไป
3	มีรายงานประจำปีหรือมีบริษัทผู้ตรวจสอบรับรองหรือมีสาขา 10-20 สาขา
2	แสดงข้อมูลผู้บริหาร แสดงข้อมูลบริษัท หรือมีสาขา 2-9 สาขา
1	แสดงข้อมูลบริษัท
0	ไม่แสดงข้อมูล

องค์ประกอบที่ 4: ข้อมูลในการติดต่อ ใช้การแสดงที่อยู่บริษัท เบอร์โทรศัพท์ สำหรับการติดต่อ อีเมลสำหรับการติดต่อ และการให้บริการตอบคำถามแบบออนไลน์ เป็นเกณฑ์ในการแบ่งระดับความน่าเชื่อถือของข้อมูลในการติดต่อออกเป็น 5 ระดับ โดยถ้าเว็บไซต์มีข้อมูลสำหรับการติดต่อครบทั้ง 4 ช่องทาง ระดับความน่าเชื่อถือของข้อมูลในการติดต่อจะอยู่ในระดับ 4 และถ้าเว็บไซต์มีข้อมูลสำหรับการติดต่อลดลงเหลือ 3 ช่องทาง ระดับความน่าเชื่อถือของข้อมูลในการติดต่อก็ลดลง 1 ระดับ เหลือเป็นระดับ 3 แต่ถ้าเว็บไซต์ไม่แสดงที่อยู่ของบริษัท ความน่าเชื่อถือของข้อมูลในการติดต่อจะเป็นระดับ 0 เกณฑ์ที่กำหนดให้เว็บครอว์เลอร์ในการดึงข้อมูลในการติดต่อทั้ง 5 ระดับ ดังแสดงในตารางที่ 3.6 โดยคำหลักที่แสดงให้เว็บครอว์เลอร์รู้ว่าเป็นที่อยู่ เช่น Address, Location หรือ Contact Information เป็นต้น

ตารางที่ 3.6

เกณฑ์ที่กำหนดให้เว็บครอว์เลอร์ในการดึงข้อมูลในการติดต่อ

ระดับ	องค์ประกอบที่ 4: ข้อมูลในการติดต่อ
4	แสดงที่อยู่บริษัท เบอร์โทรศัพท์ อีเมล และ ให้บริการตอบคำถามออนไลน์
3	แสดงที่อยู่บริษัท และ อีก 2 ช่องทางในการติดต่อ
2	แสดงที่อยู่บริษัท และ อีก 1 ช่องทางในการติดต่อ
1	แสดงที่อยู่บริษัท
0	ไม่แสดงที่อยู่บริษัท

องค์ประกอบที่ 5: การส่งหรือคืนสินค้า ใช้การแสดงค่าขนส่ง การคืนหรือเปลี่ยนสินค้าได้ การยกเลิกการสั่งซื้อสินค้าได้ และ ติดตามสินค้าได้ เป็นเกณฑ์ในการแบ่งระดับความน่าเชื่อถือของการส่งหรือคืนสินค้าออกเป็น 5 ระดับ โดยถ้าเว็บไซต์มีบริการสำหรับการส่งหรือคืนสินค้าครบทั้ง 4 บริการ ระดับความน่าเชื่อถือของการส่งหรือคืนสินค้าจะอยู่ในระดับ 4 ถ้าเว็บไซต์มีข้อมูลสำหรับการการส่งหรือคืนสินค้าออกลดลงเหลือ 3 บริการ ระดับความน่าเชื่อถือของการส่งหรือคืนสินค้าก็ลดลง 1 ระดับ เหลือเป็นระดับ 3 และถ้าเว็บไซต์มีข้อมูลสำหรับการการส่งหรือคืนสินค้าออกลดลงเหลือ 2 บริการ ระดับความน่าเชื่อถือของการส่งหรือคืนสินค้าก็ลดลงอีก 1 ระดับ เหลือเป็นระดับ 2 เกณฑ์ที่กำหนดให้เว็บครอว์เลอร์ในการดึงข้อมูลในการติดต่อทั้ง 5 ระดับ ดังแสดงในตารางที่ 3.7 โดยคำหลักที่แสดงให้เว็บครอว์เลอร์รู้ว่าเป็นการส่ง/คืนสินค้า เช่น Return Policy, Refund Policy หรือ Delivery Policy เป็นต้น

ตารางที่ 3.7

เกณฑ์ที่กำหนดให้เว็บครอว์เลอร์ในการดึงข้อมูลในการส่งหรือคืนสินค้า

ระดับ	องค์ประกอบที่ 5: การส่ง/คืนสินค้า
4	แสดงค่าขนส่ง คืน/เปลี่ยนสินค้าได้ ยกเลิกการสั่งซื้อได้ ติดตามสินค้าได้
3	มีบริการน้อยกว่าระดับ 4 จำนวน 1 บริการ
2	มีบริการน้อยกว่าระดับ 4 จำนวน 2 บริการ
1	มีบริการน้อยกว่าระดับ 4 จำนวน 3 บริการ
0	ไม่มีบริการ

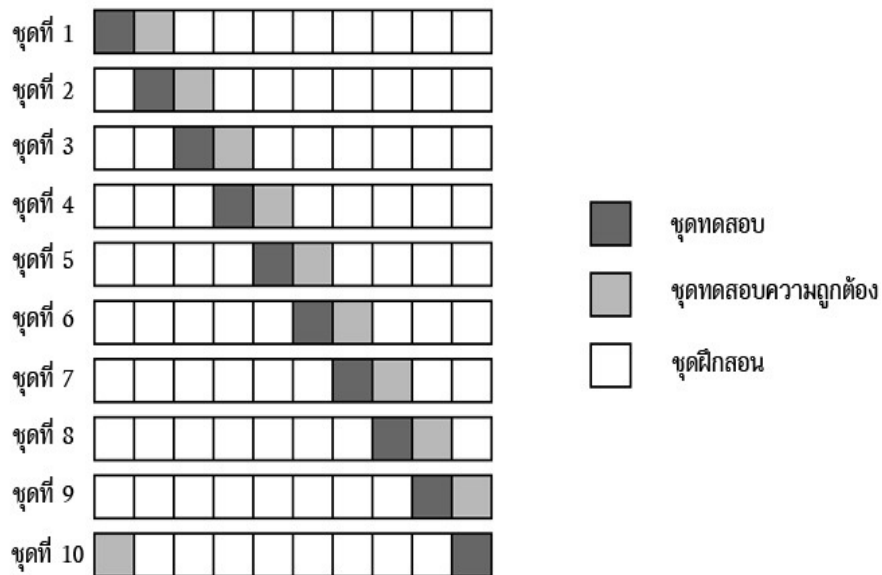
3.2.5 การเตรียมข้อมูล

สำหรับเทคนิคการเรียนรู้ของเครื่อง 4 เทคนิค ที่จะใช้ในการการประเมินระดับความน่าเชื่อถือของเว็บไซต์ จะเป็นการนำโมเดลที่มีอยู่ในโปรแกรมคอมพิวเตอร์ 2 โปรแกรม คือ โปรแกรมเวก้า (weka) เวอร์ชัน 3.7.2 และโปรแกรมแมทแล็บ (matlab) เวอร์ชัน 7.1 มาใช้งานวิจัย โดยเทคนิคต้นไม้ตัดสินใจ โครงข่ายประสาทเทียม และนาอีฟเบย์ จะใช้โมเดลที่มีอยู่ในโปรแกรมเวก้ามาดำเนินการทดลอง ส่วนตรรกศาสตร์คลุมเครือ จะใช้โมเดลที่มีอยู่ในโปรแกรมแมทแล็บมาดำเนินการทดลอง

การเตรียมข้อมูลเพื่อจะใช้กับเทคนิคการเรียนรู้ของเครื่อง 4 เทคนิค ทำโดยการแบ่งข้อมูลของเว็บไซต์ 50 เว็บไซต์ ที่ได้เลือกไว้ ออกเป็น 2 ส่วน ส่วนแรกเป็นข้อมูลจาก 40 เว็บไซต์ เพื่อเป็นตัวอย่างสำหรับให้ระบบเรียนรู้ โดยตัวแปรอินพุตของงานวิจัยนี้มี 5 ตัวแปร ได้แก่ ระดับคะแนนขององค์ประกอบของเว็บไซต์ที่ 1-5 ที่ดึงมาด้วยโปรแกรมเว็บครอว์เลอร์ ซึ่งแต่ละองค์ประกอบมีระดับคะแนน 0 – 4 สำหรับตัวแปรเอาต์พุตมี 1 ตัวแปร คือ คะแนนความน่าเชื่อถือของเว็บไซต์ที่ได้จากการตอบแบบสอบถาม ซึ่งได้จากการนำคะแนนขององค์ประกอบของเว็บไซต์ที่ 1 - 5 จากการตอบแบบสอบถามมารวมกันเพื่อเป็นคะแนนความน่าเชื่อถือโดยรวมของเว็บไซต์ที่มีคะแนนตั้งแต่ 0-20 คะแนน ในการการสอนให้ระบบเรียนรู้จะใช้การทดสอบแบบไขว้ข้ามสิบชุด (10 fold cross validation) เนื่องจากการทดสอบแบบไขว้ข้ามสิบชุดให้ค่าความแตกต่างระหว่างความคลาดเคลื่อนจริงกับความคลาดเคลื่อนที่คาดหวัง (biased error) น้อยกว่าการทดสอบแบบไขว้ข้ามสองชุด และประหยัดทรัพยากรทางคอมพิวเตอร์กว่าการทดสอบแบบไขว้ข้ามของทุกตัวอย่าง (Rodriguez, Perez & Lozano, 2010) นอกจากนี้การทดสอบแบบไขว้ข้ามสิบชุดเป็นหนึ่งในทางเลือกที่ดีที่สุดที่ประมาณค่าความคลาดเคลื่อน (error estimate) ได้แม่นยำที่สุด (Lazakidou, 2010) โดยวิธีการทดสอบชุดข้อมูลแบบไขว้ข้ามสิบชุด คือ ให้ 1 ชุดข้อมูลเป็นชุดที่ใช้ทดสอบความถูกต้อง (validation) อีก 1 ชุดข้อมูลเป็นชุดทดสอบ ส่วนที่เหลือ 8 ชุดเป็นชุดสำหรับเรียนรู้ โดยชุดข้อมูลทั้งหมดจะถูกสลับเรียงกัน (ปกรณ, 2550) ดังแสดงในภาพที่ 3.4 ข้อมูลส่วนที่สองสำหรับงานวิจัยนี้ คือข้อมูลสำหรับให้ระบบทำการประเมินระดับความน่าเชื่อถือของเว็บไซต์ โดยเป็นข้อมูลของเว็บไซต์จำนวน 10 เว็บไซต์ ที่เหลือหลังจากได้แบ่งข้อมูลส่วนแรกสำหรับให้ตัวอย่างเรียนรู้

ภาพที่ 3.4

การทดสอบชุดข้อมูลแบบไขว้ข้ามสิบชุด (10 fold cross validation)



แหล่งที่มา: ปกรณ์ จารุตระกูลชัย (2550)

1. การเตรียมข้อมูลของต้นไม้ตัดสินใจ

การเตรียมข้อมูลของต้นไม้ตัดสินใจสามารถทำได้โดยการปรับเปลี่ยนค่าพารามิเตอร์ที่เหมาะสม ดังนี้คือ

1. ปัจจัยความเชื่อมั่น (confidence factor) ซึ่งใช้สำหรับการตัดเล็ม (prune) ต้นไม้ตัดสินใจ โดยปัจจัยความเชื่อมั่นจะมีค่าอยู่ระหว่าง 0-1 การกำหนดให้ปัจจัยความเชื่อมั่นมีค่าน้อยจะทำให้มีจำนวนครั้งของการตัดเล็มต้นไม้มากขึ้น ในขณะที่การกำหนดให้ปัจจัยความเชื่อมั่นมีค่ามากจะทำให้มีจำนวนครั้งของการตัดเล็มต้นไม้ลดลง (Lazakidou, 2010) การตัดเล็มต้นไม้ตัดสินใจมีวัตถุประสงค์เพื่อเอากิ่งที่มีความน่าเชื่อถือน้อยที่สุดออกไป ซึ่งจะช่วยให้ต้นไม้สามารถจัดกลุ่มข้อมูลที่นำมาทดสอบได้ถูกต้องมากยิ่งขึ้นและทำให้การจัดกลุ่มมีความรวดเร็วขึ้น (Cai, 2006) งานวิจัยนี้ได้ทดลองปรับค่าปัจจัยความเชื่อมั่นโดยเริ่มจากค่าปัจจัยความเชื่อมั่นที่โปรแกรมเวก้ากำหนดไว้ (default) คือ 0.25 แล้วปรับลดและเพิ่มจากค่าที่โปรแกรมเวก้ากำหนดไว้ โดยพิจารณาว่าการปรับค่าปัจจัยความเชื่อมั่นให้ลดลงหรือเพิ่มขึ้นที่ทำให้ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ย (RMSE) มีค่าลดลง ถ้าการปรับลดค่าปัจจัยความเชื่อมั่นทำให้ค่ารากที่

สองของค่าความคลาดเคลื่อนเฉลี่ยมีค่าลดลง จะทำการปรับลดค่าปัจจัยความเชื่อมั่นให้ลดลงอีกในครั้งต่อไป จนกระทั่งค่าปัจจัยความเชื่อมั่นที่ทำให้ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยมีค่าน้อยที่สุด โดยงานวิจัยนี้ได้ทดลองปรับค่าปัจจัยความเชื่อมั่นที่ 0.05, 0.1, 0.25 และ 0.5

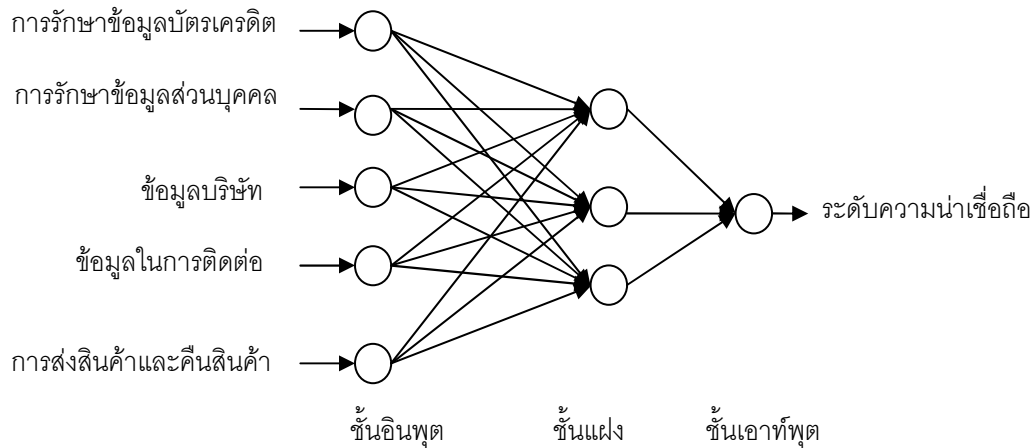
2. จำนวนตัวอย่างที่น้อยที่สุดต่อโหนดใบ (minimum number of instances per leaf) เป็นการกำหนดว่าโหนดใบที่จะสร้างขึ้นโหนดต่อไปควรมีจำนวนตัวอย่างที่น้อยที่สุดจำนวนเท่าใด หรืออีกนัยหนึ่งคือการบ่งบอกว่าต้นไม้ตัดสินใจมีความเป็นทั่วหรือมีความเฉพาะเจาะจงมากน้อยเพียงไร ถ้ากำหนดจำนวนตัวอย่างที่น้อยที่สุดต่อโหนดใบให้มีค่าสูงจะเป็นการสร้างต้นไม้ตัดสินใจที่มีความเป็นทั่วไป แต่ถ้ากำหนดจำนวนตัวอย่างที่น้อยที่สุดต่อโหนดใบให้มีค่าน้อยจะเป็นการสร้างต้นไม้ตัดสินใจที่มีความเฉพาะเจาะจงมากยิ่งขึ้น (Lazakidou, 2010) งานวิจัยนี้ได้ทดลองปรับจำนวนตัวอย่างที่น้อยที่สุดต่อโหนดใบโดยเริ่มจากจำนวนตัวอย่างที่น้อยที่สุดต่อโหนดใบที่โปรแกรมเวก้ากำหนดไว้ คือ จำนวน 2 ตัวอย่างต่อโหนดใบ แล้วจึงทำการปรับลดหรือเพิ่มจำนวนตัวอย่างต่อโหนดใบจนกระทั่งได้จำนวนตัวอย่างที่น้อยที่สุดต่อโหนดใบที่ทำให้ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยมีค่าน้อยที่สุด โดยงานวิจัยนี้ได้ทดลองปรับจำนวนตัวอย่างที่น้อยที่สุดต่อโหนดใบเป็น 2, 3 และ 5

2. การเตรียมข้อมูลสำหรับโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมสำหรับงานวิจัยนี้เป็นโครงข่ายประสาทเทียมหลายชั้นแบบไปข้างหน้า ดังแสดงในภาพที่ 3.5 โดยโครงข่ายประสาทเทียมประกอบด้วยชั้นอินพุตจำนวน 1 ชั้น ชั้นแฝงจำนวน 1 ชั้น และชั้นเอาต์พุตจำนวน 1 ชั้น สำหรับการเตรียมข้อมูลของโครงข่ายประสาทเทียม ได้กำหนดให้ชั้นอินพุตมีจำนวนโหนด 5 โหนด โดยแต่ละโหนดแทนองค์ประกอบของเว็บไซต์ที่เกี่ยวข้องกับระดับความน่าเชื่อถือของเว็บไซต์ และกำหนดให้ชั้นเอาต์พุตมีจำนวนโหนด 1 โหนด โดยโหนดเอาต์พุตแทนระดับความน่าเชื่อถือของเว็บไซต์ อัลกอริทึมที่ใช้ในการเรียนรู้ของโครงข่ายประสาทเทียมเป็นแบบแบ็กพรอพาเกชัน

ภาพที่ 3.5

การเตรียมข้อมูลของโครงข่ายประสาทเทียม



การปรับค่าพารามิเตอร์ของโครงข่ายประสาทเทียม

ในการเตรียมข้อมูลของโครงข่ายประสาทเทียมจะมีการปรับค่าพารามิเตอร์ที่เกี่ยวข้อง เพื่อหาค่าพารามิเตอร์ที่ทำให้โครงข่ายประสาทเทียมมีประสิทธิภาพมากที่สุดสำหรับงานวิจัย โดยการปรับค่าพารามิเตอร์สำหรับโครงข่ายประสาทเทียมจะเริ่มจากค่าที่โปรแกรมเมอร์กำหนดไว้ (default) คือ ค่าอัตราการเรียนรู้ 0.3 ค่าโมเมนตัม 0.2 จำนวนรอบในการเรียนรู้ 500 รอบ สำหรับจำนวนโหนดในชั้นแฝงโปรแกรมเมอร์ไม่ได้กำหนดค่าเริ่มแรกไว้จึงทดลองปรับโดยเริ่มจาก 2 โหนด เป็นต้นไป โดยการปรับค่าพารามิเตอร์สำหรับโครงข่ายประสาทเทียมจะทำการปรับค่าพารามิเตอร์ จนกระทั่งได้ค่าพารามิเตอร์ที่ทำให้ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยมีค่าน้อยที่สุด งานวิจัยนี้ได้ทดลองปรับค่าพารามิเตอร์สำหรับโครงข่ายประสาทเทียมดังนี้ คือ

1. ค่าอัตราการเรียนรู้ ทดลองปรับค่าเป็น 0.1, 0.2, และ 0.3
2. ค่าโมเมนตัม ทดลองปรับค่าเป็น 0.1, 0.2, และ 0.3
3. จำนวนรอบในการเรียนรู้ ทดลองปรับค่าเป็น 500, 1000, และ 1500 รอบ
4. จำนวนโหนดในชั้นแฝง ทดลองปรับค่าเป็น 2, 3 และ 4 โหนด

3. การเตรียมข้อมูลสำหรับนาอูฟเบย์

ในการเตรียมข้อมูลสำหรับนาอ์ฟเบย์สามารถทำได้โดยการปรับค่าพารามิเตอร์ของการแจกแจงความถี่ของคุณลักษณะ โดยในงานวิจัยนี้จะทดลองใช้โมเดลการแจกแจงแบบปกติ (Normal Distribution)

4. การเตรียมข้อมูลสำหรับตรรกศาสตร์คลุมเครือ

ตรรกศาสตร์คลุมเครือสำหรับงานวิจัยนี้เป็นโมเดลที่ประกอบด้วยตัวแปรอินพุต ตัวแปรเอาต์พุต กฎและเงื่อนไข ดังแสดงในภาพที่ 3.6 โดยขั้นตอนการเตรียมข้อมูลมีดังนี้

1. สร้างตัวแปรอินพุตที่เป็นสมการคลุมเครือ ซึ่งตัวแปรอินพุตสำหรับตรรกศาสตร์คลุมเครือในงานวิจัยนี้คือ องค์ประกอบของเว็บไซต์ทั้ง 5 องค์ประกอบ ได้แก่ (1)การรักษาความปลอดภัยข้อมูลบัตรเครดิต (2)การรักษาข้อมูลส่วนบุคคล (3)ข้อมูลบริษัท (4)ข้อมูลในการ และ (5)การส่ง/คืนสินค้า

2. สร้างตัวแปรเอาต์พุตที่เป็นสมการคลุมเครือ ซึ่งตัวแปรเอาต์พุตในงานวิจัยนี้คือ ระดับความน่าเชื่อถือของเว็บไซต์

3. สร้างฟังก์ชันการเป็นสมาชิกให้กับตัวแปรอินพุต และตัวแปรเอาต์พุต สำหรับงานวิจัยนี้เลือกใช้ฟังก์ชันการเป็นสมาชิกแบบกซ์เซียน เนื่องจากเป็นฟังก์ชันที่มีความสมมาตรและมีความเรียบ โดยกำหนดให้แต่ละตัวแปรอินพุต และเอาต์พุต มีฟังก์ชันการเป็นสมาชิก 2 ฟังก์ชัน การกำหนดให้มีฟังก์ชันการเป็นสมาชิก 2 ฟังก์ชัน เนื่องจากงานวิจัยนี้มีตัวแปรอินพุต 5 ตัวแปร ซึ่งคือองค์ประกอบของเว็บไซต์ 5 องค์ประกอบ ได้แก่ (1)การรักษาความปลอดภัยข้อมูลบัตรเครดิต (2)การรักษาข้อมูลส่วนบุคคล (3)ข้อมูลบริษัท (4)ข้อมูลในการติดต่อ และ (5)การส่ง/คืนสินค้า การให้แต่ละตัวแปรอินพุตมีค่าฟังก์ชันการเป็นสมาชิก 2 ฟังก์ชัน ทำให้สามารถสร้างกฎอนุมานคลุมเครือที่เป็นไปได้ทั้งหมดจำนวน 32 ข้อ โดยจำนวนกฎอนุมานคลุมเครือที่เป็นไปได้ทั้งหมดคำนวณได้จากจำนวนฟังก์ชันการเป็นสมาชิกยกกำลังด้วยจำนวนตัวแปรอินพุต ซึ่งสำหรับงานวิจัยนี้คือ 2^5 หรือ 32 ข้อ ถ้ากำหนดให้มีฟังก์ชันการเป็นสมาชิกเพียงฟังก์ชันเดียวจะสามารถสร้างกฎอนุมานคลุมเครือได้เพียง 1 ข้อ ซึ่งทำให้ไม่สามารถจัดกลุ่มระดับความน่าเชื่อถือขององค์ประกอบของเว็บไซต์ได้ และถ้ากำหนดให้มีฟังก์ชันการเป็นสมาชิก 3 ฟังก์ชัน จะสามารถสร้างกฎอนุมานคลุมเครือได้มากถึง 243 ข้อ ซึ่งเป็นการยากต่อการจัดการและมีความละเอียดเกินความจำเป็น สำหรับชื่อของฟังก์ชันการเป็นสมาชิกในงานวิจัยนี้จะให้ชื่อฟังก์ชันว่า ฟังก์ชันระดับต่ำ และฟังก์ชันระดับสูง ฟังก์ชันการเป็นสมาชิกแบบกซ์เซียนเป็นฟังก์ชันที่มีค่าพารามิเตอร์ คือ (σ, m) โดย σ หมายถึง ค่าความเบี่ยงเบน และ m หมายถึง ค่ากลางของฟังก์ชัน ซึ่งในงานวิจัยนี้จะเตรียมข้อมูล

ตัวแปรอินพุตโดยทำการปรับค่า σ เป็น 0.3 และ 0.5 เนื่องจากค่า σ ระหว่าง 0.3 - 0.5 ทำให้ส่วนโค้งของฟังก์ชันมีลักษณะเป็นรูปประฆังที่ส่วนปลายทั้งสองครอบคลุมระดับความน่าเชื่อถือขององค์ประกอบทุกระดับตั้งแต่ระดับ 0 - 4 สำหรับฟังก์ชันระดับต่ำจะกำหนดให้ $m = 1$ และฟังก์ชันระดับสูงจะกำหนดให้ $m = 3$ เนื่องจากเป็นค่ากลางที่ทำให้ฟังก์ชันมีความสมมาตรและครอบคลุมระดับความน่าเชื่อถือของเว็บไซต์ทุกระดับตั้งแต่ระดับ 0-4 สำหรับตัวแปรเอาต์พุตจะทำการปรับค่า σ เป็น 1.5, 2 และ 2.5 ฟังก์ชันระดับต่ำให้ $m = 5$ และฟังก์ชันระดับสูงให้ $m = 15$ ตัวอย่างของฟังก์ชันการเป็นสมาชิกของตัวแปรอินพุต ดังแสดงในภาพที่ 3.7

4.สร้างกฎและเงื่อนไขเพื่อให้ระบบทำการแปลงข้อมูลออกมาเป็นเอาต์พุต โดยจำนวนกฎและเงื่อนไขที่มีจำนวนที่เหมาะสมจะช่วยลดความคลาดเคลื่อนในการจัดกลุ่มข้อมูลและง่ายต่อการจัดการ งานวิจัยนี้จะใช้กฎ 32 ข้อในการเตรียมข้อมูลสำหรับตรรกศาสตร์คลุมเครือ โดยตัวอย่างกฎที่ใช้ในงานวิจัยนี้ ได้แก่

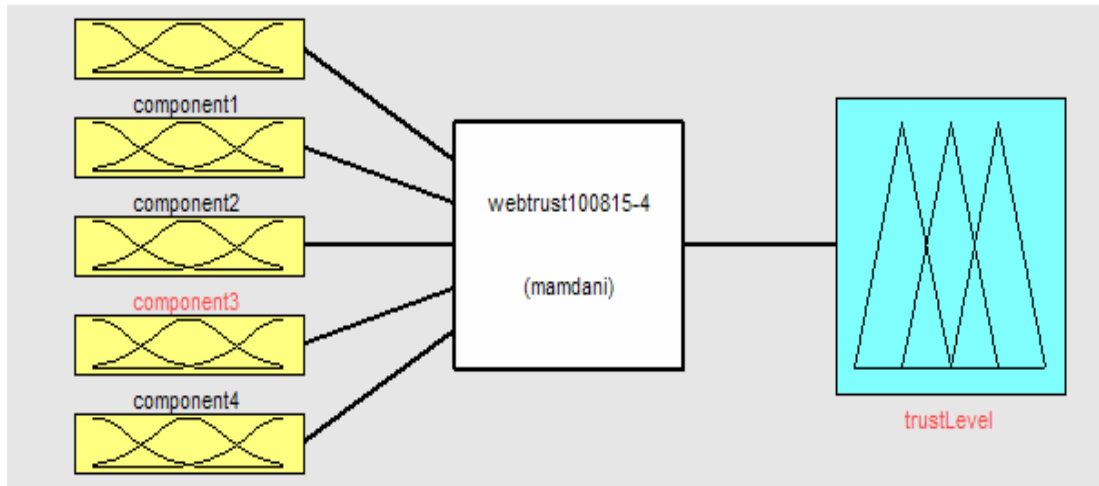
กฎข้อที่ 1 ถ้า การรักษาความปลอดภัยข้อมูลบัตรเครดิต มีระดับต่ำ และการรักษาข้อมูลส่วนบุคคล มีระดับต่ำ และข้อมูลบริษัท มีระดับต่ำ และข้อมูลในการติดต่อ มีระดับต่ำ และการส่ง/คืนสินค้า มีระดับต่ำ แล้ว ความน่าเชื่อถือของเว็บไซต์ มีระดับต่ำ

กฎข้อที่ 2 ถ้า การรักษาความปลอดภัยข้อมูลบัตรเครดิต มีระดับต่ำ และการรักษาข้อมูลส่วนบุคคล มีระดับสูง และข้อมูลบริษัท มีระดับต่ำ และข้อมูลในการติดต่อ มีระดับสูง และการส่ง/คืนสินค้า มีระดับสูง แล้ว ความน่าเชื่อถือของเว็บไซต์ มีระดับสูง

กฎข้อที่ 3 ถ้า การรักษาความปลอดภัยข้อมูลบัตรเครดิต มีระดับสูง และการรักษาข้อมูลส่วนบุคคล มีระดับสูง และข้อมูลบริษัท มีระดับสูง และข้อมูลในการติดต่อ มีระดับสูง และการส่ง/คืนสินค้า มีระดับสูง แล้ว ความน่าเชื่อถือของเว็บไซต์ มีระดับสูง

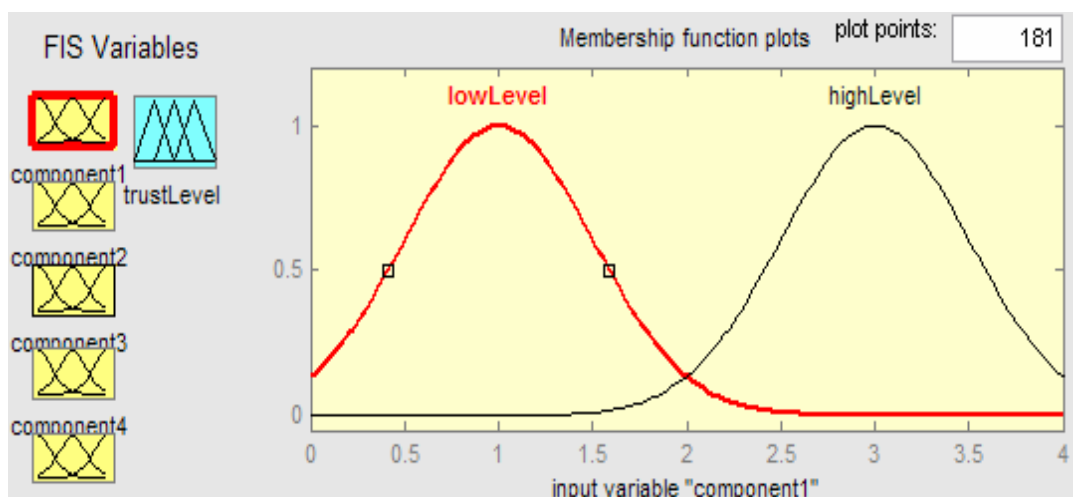
ภาพที่ 3.6

โมเดลของตรรกศาสตร์คลุมเครือที่ประกอบด้วยอินพุต 5 ตัวแปร และเอาต์พุต 1 ตัวแปร



ภาพที่ 3.7

ตัวอย่างฟังก์ชันการเป็นสมาชิกแบบกัซเซียนจำนวน 2 ฟังก์ชัน สำหรับตัวแปรอินพุต ที่มีค่าพารามิเตอร์ของฟังก์ชัน (0.5, 1) และ (0.5, 3)



3.2.6 การวิเคราะห์ข้อมูล (Data Analysis)

การวิเคราะห์ข้อมูลสำหรับงานวิจัยนี้ประกอบด้วย 5 ส่วน ดังนี้คือ

1.การวิเคราะห์ผลการทดลองจากแบบสอบถามในส่วนที่ 1 ซึ่งเป็นข้อมูลส่วนบุคคลของผู้เข้าร่วมการทดลอง ใช้การวิเคราะห์ข้อมูลทางสถิติ คือ ค่าความถี่ (Frequency) และร้อยละ (Percent)

2.การวิเคราะห์องค์ประกอบที่มีความสำคัญต่อระดับความน่าเชื่อถือของเว็บไซต์ โดยใช้การวิเคราะห์ค่าเฉลี่ยของคะแนนความสำคัญขององค์ประกอบของเว็บไซต์

3.การวิเคราะห์ว่าการประเมินระดับความน่าเชื่อถือของเว็บไซต์ด้วยเทคนิคการเรียนรู้ของเครื่องแต่ละเทคนิคมีความแตกต่างจากการประเมินระดับความน่าเชื่อถือของเว็บไซต์ด้วยการตอบแบบสอบถามอย่างมีนัยสำคัญทางสถิติหรือไม่ โดยใช้วิธีการหาค่าความแตกต่างระหว่างค่าเฉลี่ยของสองประชากรแบบจับคู่ (Paired Samples T Test) (กัลยา วานิชย์บัญชา, 2538) ถ้าผลการหาค่าความแตกต่างระหว่างค่าเฉลี่ยของสองประชากรแบบจับคู่พบว่าการประเมินระดับความน่าเชื่อถือของเว็บไซต์ด้วยเทคนิคแต่ละเทคนิคมีความแตกต่างจากการประเมินระดับความน่าเชื่อถือของเว็บไซต์ด้วยการตอบแบบสอบถามอย่างไม่มีนัยสำคัญทางสถิติ จึงทำการวิเคราะห์เปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้ของเครื่องแต่ละเทคนิคในการประเมินระดับความน่าเชื่อถือของเว็บไซต์

4.การวิเคราะห์หาค่าพารามิเตอร์ของแต่ละเทคนิคการเรียนรู้ของเครื่องที่ทำให้ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย (RMSE) (ปรีดา อนุสรณ์ธิรกุล และ พุทธิ ศิริแสงตระกูล, 2552) มีค่าน้อยที่สุด เพื่อนำผลการประเมินโดยค่าพารามิเตอร์ของแต่ละเทคนิคที่ให้ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุดมาเปรียบเทียบประสิทธิภาพในการประเมินระดับความน่าเชื่อถือของเว็บไซต์โดยใช้การประเมินด้วยคนเป็นเกณฑ์ในการเปรียบเทียบ

5.การใช้หลักการทางสถิติในการวิเคราะห์ความแปรปรวนแบบทางเดียว (One-Way ANOVA) ในเปรียบเทียบประสิทธิภาพในการประเมินระดับความน่าเชื่อถือของเว็บไซต์ของเทคนิคการเรียนรู้ของเครื่องโดยใช้การประเมินด้วยคนเป็นเกณฑ์ในการเปรียบเทียบ การใช้การวิเคราะห์ความแปรปรวนแบบทางเดียว เนื่องจากมีปัจจัยเดียวคือ ประสิทธิภาพในการประเมินระดับความน่าเชื่อถือของเว็บไซต์ของเทคนิคการเรียนรู้ของเครื่องที่ทำให้ระดับความน่าเชื่อถือของเว็บไซต์มีความแตกต่างกัน