

บทคัดย่อ

การบาดเจ็บที่เกิดจากอุบัติเหตุจราจรทางบกเป็นปัญหาที่สำคัญอย่างยิ่งของประเทศไทย เพราะการเกิดอุบัติเหตุ这不仅นอกจากจะส่งผลกระทบต่อการสูญเสียบุคคลที่เป็นกำลังสำคัญของชาติ แล้วยังส่งผลกระทบต่อคุณภาพชีวิต ครอบครัวและสังคมอีกด้วย โดยเฉพาะอย่างยิ่งในกรณีที่เกิดความพิการขึ้นเพราะความพิการดังกล่าวจะส่งผลกระทบต่อสังคมและเศรษฐกิจโดยตรง ไม่ว่าจะเป็นส่วนบุคคล หรือด้านครอบครัว และส่วนรวมทั้งทางตรงและทางอ้อม เนื่องจากการทุพพลภาพหรือพิการทำให้บุคคลไม่สามารถช่วยเหลือตนเอง และประกอบกิจวัตรประจำวันได้ และไม่สามารถประกอบอาชีพของตนได้ ทำให้บุคคลตกเป็นภาระต่อพ่อ แม่ พี่น้อง หรือคนอื่น ๆ ในครอบครัว รวมไปถึงประเทศชาติ

ในงานวิจัยนี้ได้นำข้อมูลหัตถิยภูมิซึ่งจัดเก็บโดยศูนย์สิรินธรเพื่อการฟื้นฟูสมรรถภาพทางการแพทย์แห่งชาติ (ดารณี และคณะ, 2551) มาวิเคราะห์ทางสถิติเพื่อหาปัจจัยที่ก่อให้เกิดความพิการได้แก่ ข้อมูลอายุ เพศ ช่วงเวลาเกิดอุบัติเหตุ การใช้แอลกอฮอล์หรือยาระหว่างการขับขี่ การใช้หมวกนิรภัยและเข็มขัดนิรภัย เป็นต้น โดยการนำปัจจัยเหล่านี้มาพัฒนาตัวแบบในการพยากรณ์ระดับความรุนแรงของการพิการโดยใช้เทคนิคโครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอน ผลที่ได้จากการวิจัยพบว่าตัวแบบในการพยากรณ์ที่ได้มีความแม่นยำถึงร้อยละ 80.60 ซึ่งเป็นผลลัพธ์ที่อยู่ในระดับยอมรับได้ในกรณีที่ระดับความรุนแรงของการพิการเป็นแบบเข้าเกณฑ์ จากนั้นได้นำผลลัพธ์ที่ได้ไปพัฒนาตัวโมเดลด้วยภาษา EDEN ซึ่งเป็นการพัฒนาโปรแกรมแนวใหม่ที่เน้นการใช้โปรแกรมเป็นเครื่องมือเพื่อหาองค์ความรู้เชิงลึกของสิ่งที่เราต้องการค้นหา

Abstract

This research project introduces a computer-based model for predicting the severity of injuries in road traffic accidents. Using accident data from surveys at hospitals in Thailand, standard data mining techniques were applied to train and test a multilayer perceptron neural network. The resulting neural network specification was loaded into an interactive environment called EDEN that enables further exploration of the computer-based model. Although the model can be used for the classification of accident data in terms of injury severity (in a similar way to other data mining tools), the EDEN tool enables deeper exploration of the underlying factors that might affect injury severity in road traffic accidents. The aim of this paper is to describe the development of the computer-based model and to demonstrate the potential of EDEN as an interactive tool for knowledge discovery. The prediction accuracy of the developed model is about 80.60% which is acceptable to use.

Key Words: Disability, Artificial Neural Network, Interactive Environments, Data Mining.

Executive Summary

การบาดเจ็บที่เกิดจากอุบัติเหตุจราจรทางบกเป็นปัญหาที่สำคัญอย่างยิ่งของประเทศไทย เพราะการเกิดอุบัติเหตุนอกจากกระทบต่อการสูญเสียบุคคลที่เป็นกำลังสำคัญของชาติ ยังส่งผลกระทบต่อคุณภาพชีวิต ครอบครัวและสังคมอีกด้วย โดยเฉพาะอย่างยิ่งเมื่อเกิดความพิการดังกล่าวยังส่งผลกระทบต่อสังคมและเศรษฐกิจโดยตรง ไม่ว่าจะเป็นด้านส่วนบุคคล ครอบครัว และส่วนรวม กล่าวคือการทุพพลภาพหรือพิการทำให้บุคคลไม่สามารถช่วยเหลือตนเองและประกอบกิจวัตรประจำวัน และไม่สามารถประกอบอาชีพของตนได้ ทำให้บุคคลตกเป็นภาระต่อพ่อ แม่ พี่น้อง ลูกหลาน ญาติ สังคม และเป็นเทศชาติ

ในงานวิจัยนี้ได้นำข้อมูลทุติยภูมิซึ่งจัดเก็บโดยศูนย์สรีรศาสตร์เพื่อการฟื้นฟูสมรรถภาพทางการแพทย์แห่งชาติ (ดารณี และคณะ, 2551) มาวิเคราะห์เพื่อหาปัจจัยด้วยเทคนิคทางสถิติ และได้ปัจจัยที่สัมพันธ์ต่ออุบัติเหตุที่ก่อให้เกิดความพิการได้แก่ ข้อมูลอายุ เพศ ช่วงเวลาเกิดอุบัติเหตุ แอลกอฮอล์ ยา การใช้หมวกนิรภัยและเข็มขัดนิรภัย โดยได้นำปัจจัยเหล่านี้มาพัฒนาตัวแบบในการพยากรณ์ระดับความรุนแรงของการพิการด้วยเทคนิคโครงข่ายประสาทเทียม โดยใช้โปรแกรม Weka ซึ่งเป็นเครื่องมือที่นิยมใช้ในการสร้างตัวพยากรณ์เบื้องต้น โดยตัวพยากรณ์ที่สร้างได้มีความแม่นยำร้อยละ 80.60 จากนั้นได้นำผลลัพธ์ที่ได้ไปพัฒนาตัวโมเดลด้วยภาษา Eden ซึ่งเป็นการพัฒนาโปรแกรมแนวใหม่ที่เน้นการใช้โปรแกรมเป็นเครื่องมือเพื่อหาองค์ความรู้เชิงลึกของสิ่งที่เราต้องการค้นหา ซึ่งหมายถึงกลไกการทำงานของตัวพยากรณ์ด้วยเทคนิคโครงข่ายประสาทเทียม เพื่อหาปัจจัย และความสัมพันธ์ภายในโหนดซ่อนของโครงข่าย

อย่างไรก็ตามตัวพยากรณ์ที่ได้แม้จะมีความแม่นยำที่ร้อยละ 80.60 ซึ่งมีค่าอยู่ในระดับที่ยอมรับได้ แต่ผลลัพธ์การพยากรณ์สำหรับกรณีที่ระดับความรุนแรงของการพิการเป็นแบบไม่เข้าเกณฑ์ยังไม่เป็นที่น่าพอใจ จึงอาจจะต้องปรับโครงข่ายประสาทเทียมใหม่ รวมไปถึงการค้นหาปัจจัยที่อาจจะมีผลความรุนแรงของการพิการ ตลอดจนการประยุกต์ใช้เทคนิคเหมือนข้อมูลตัวอื่น ๆ เช่น เจเนติกอัลกอริทึม การใช้ข่ายงานแบบเบย์ เป็นต้น

บทที่ 1

บทนำ

รายงานนี้เป็นส่วนหนึ่งของโครงการวิจัยภายใต้การสนับสนุนของเงินงบประมาณแผ่นดิน ประจำปีงบประมาณ 2554 มหาวิทยาลัยนเรศวร โดยคณะผู้วิจัยได้รับการอนุมัติให้ทำงานวิจัยนี้ โดยมีชื่อโครงการและรายละเอียดเกี่ยวกับโครงการวิจัยดังต่อไปนี้

ชื่อโครงการ

(ภาษาไทย)	การวิเคราะห์ปัจจัยที่มีผลต่อความพิการจากอุบัติเหตุจราจรทางบก ด้วยเทคนิคเหมืองข้อมูล
(ภาษาอังกฤษ)	An analysis of disability causes from road traffic injury using data mining techniques

คณะผู้วิจัย(ระบุสังกัดภาควิชา) และสัดส่วนที่ทำงานวิจัย (%)

หัวหน้าโครงการวิจัย

ผู้ช่วยศาสตราจารย์ ดร. จรัสศรี รุ่งรัตนอุบล (40%)
อาจารย์สังกัดภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ
มหาวิทยาลัยนเรศวร

ผู้วิจัยร่วม

ผู้ช่วยศาสตราจารย์ ดร. อนามัย นาอุดม (40%)
อาจารย์สังกัดภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์
มหาวิทยาลัยนเรศวร

แพทย์หญิง ดารณี สุวพันธ์ (20%)
สังกัดศูนย์สิรินธรเพื่อการฟื้นฟูสมรรถภาพทางการแพทย์แห่งชาติ
กรมการแพทย์ กระทรวงสาธารณสุข

สถานที่จัดทำโครงการวิจัย

ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ และภาควิชาคณิตศาสตร์ คณะ
วิทยาศาสตร์ มหาวิทยาลัยนเรศวร

ความสำคัญและที่มาของปัญหาที่ทำการวิจัย

การบาดเจ็บที่เกิดจากอุบัติเหตุจราจรทางบกถือเป็นปัญหาที่สำคัญอย่างยิ่งของประเทศไทย ซึ่งนับวันจะมีความรุนแรงและก่อให้เกิดความเสียหายต่อประเทศเป็นอย่างมาก นอกจากนี้ยังเป็นปัญหาที่ติดอันดับหนึ่งในสามของปัญหาด้านสาธารณสุขมาโดยตลอด และก่อให้เกิดความสูญเสียทางเศรษฐกิจและสังคมเป็นอย่างสูง ภายหลังการบาดเจ็บอาจก่อให้เกิดความพิการตามมา ซึ่งองค์การอนามัยโลกคาดการณ์ไว้มีประมาณ 2% ของผู้ป่วยที่บาดเจ็บจะก่อให้เกิดความพิการ (ดารณี และคณะ, 2551) และความพิการดังกล่าวส่งผลกระทบต่อสังคมและเศรษฐกิจ ส่วนในด้านตัวบุคคลทำให้เกิดการสูญเสียและไม่สามารถช่วยเหลือตนเองในการประกอบกิจวัตรประจำวันได้ รวมไปถึงการประกอบอาชีพ ทำให้ตกเป็นภาระต่อญาติที่จะต้องดูแล จากรายงานการเฝ้าระวังการบาดเจ็บระดับชาติจากโรงพยาบาลเครือข่าย 19 แห่ง พบว่าสถิติการบาดเจ็บจากอุบัติเหตุจราจรทางบกในช่วงปี พ.ศ. 2546-2547 มีจำนวน 43,816 คน และมีจำนวนผู้บาดเจ็บถึง 44,232 คน ซึ่งจากจำนวนดังกล่าว จะเห็นว่ายังไม่มีการศึกษาถึงผลของการบาดเจ็บว่ามีปัจจัยใดบ้าง ที่ก่อให้เกิดความพิการ ซึ่งข้อมูลเหล่านี้จะเป็นสารประโยชน์ที่สำคัญในการวางแผนป้องกัน และการพัฒนางานฟื้นฟูสมรรถภาพของคนพิการในอนาคต

ศูนย์สิรินธรเพื่อการฟื้นฟูสมรรถภาพทางการแพทย์แห่งชาติ ได้ทำการวิจัยเรื่อง “อุบัติการณ์ความพิการและผลกระทบที่เกิดขึ้นจากสาเหตุอุบัติเหตุจราจรทางบก พ.ศ. 2549” ซึ่งเป็นการศึกษาเชิงสำรวจโดยการเก็บข้อมูลจากแผนกฉุกเฉินในโรงพยาบาลขนาดใหญ่ทั้งหมด 8 แห่งจากภาคต่าง ๆ ในประเทศไทยโดยมีวัตถุประสงค์เพื่อศึกษาอุบัติการณ์ของความพิการถาวรที่เกิดขึ้นภายหลังการบาดเจ็บรุนแรงจากอุบัติเหตุจราจรทางบก รวมไปถึงการศึกษาประเภทและระดับของความพิการ ตลอดจนศึกษาความสามารถในการทำกิจกรรมของคนพิการ นอกจากนี้ยังมีการศึกษาความสูญเสียเชิงเศรษฐกิจ และศึกษาเปรียบเทียบสภาวะสุขภาพของคนพิการก่อนและหลังได้รับบาดเจ็บ ผลที่ได้จากการวิจัยพบว่า จำนวนผู้บาดเจ็บไม่รุนแรง (not admitted) มีจำนวน 14,698 ราย และบาดเจ็บรุนแรง (admitted) จำนวน 9,737 ราย ทั้งนี้ไม่พบว่ามีความพิการถาวรในกลุ่มผู้บาดเจ็บไม่รุนแรง

ในขณะที่กลุ่มผู้บาดเจ็บรุนแรงพบว่ามีความพิการร้อยละ 4.6 เสียชีวิตในโรงพยาบาลร้อยละ 5.6 ญาติขอรับกลับไปเสียชีวิตที่บ้านร้อยละ 1.0 และความพิการที่พบมากที่สุดคือความพิการทางการเคลื่อนไหว สูงถึงร้อยละ 75.6 รองลงมาคือความพิการทางการมองเห็นและความพิการทางจิตใจหรือพฤติกรรมคิดเป็นร้อยละ 7.4 ส่วนความพิการทางการได้ยินหรือสื่อความหมายร้อยละ 5.0 และความพิการทางสติปัญญาหรือการเรียนรู้ร้อยละ 4.6

จากผลการวิจัยดังกล่าวข้างต้นจะเห็นได้ว่าการเก็บรวบรวมข้อมูลในงานวิจัยดังกล่าวได้นำเสนอและสรุปผลในแง่ของการวิเคราะห์ข้อมูลรวมเบื้องต้นที่เป็นความถี่และสถิติเชิงพรรณนาเท่านั้น เพื่อเป็นการนำข้อมูลไปใช้ให้เกิดประโยชน์สูงสุดรวมไปถึงการขยายขอบเขตของการสรุปผลที่ได้จากการวิจัยให้กว้างขึ้น ผู้วิจัยจึงมีแนวคิดที่จะประยุกต์เทคนิคเหมืองข้อมูลแบบต่าง ๆ เช่น Classification rule and regression tree (CART) และ Multivariate logistic regression analysis และการสร้างตัวแบบพยากรณ์รูปแบบต่างๆ เข้ามาเพื่อใช้ในการวิเคราะห์ข้อมูลและหาความสัมพันธ์ของตัวแปรต่าง ๆ ที่มีความซับซ้อนมากขึ้น และ

ศึกษารายละเอียดเชิงลึกของความสัมพันธ์ของตัวแปรต่าง ๆ ที่มีอิทธิพลต่อความพิการภายหลังการบาดเจ็บรุนแรงจากอุบัติเหตุจราจรทางบก รวมไปถึงการระบุปัจจัยและน้ำหนักของอิทธิพลสำหรับแต่ละตัวแปรที่มีผลต่อความพิการหลังการบาดเจ็บรุนแรง ซึ่งจะส่งผลให้สามารถเรียงลำดับความสำคัญของตัวแปรแต่ละตัวได้นอกจากนี้ผู้วิจัยคาดหวังว่าจะสามารถสร้างตัวแบบทางสถิติเพื่อพยากรณ์โอกาสที่จะเกิดความพิการถาวรเมื่อทราบสาเหตุหรือปัจจัยที่เกี่ยวข้องต่าง ๆ จากการเกิดอุบัติเหตุ ซึ่งผลที่ได้จากการวิจัยสามารถนำไปใช้ในการกำหนดทิศทางของมาตรการเฝ้าระวังและการป้องกันความรุนแรงจากอุบัติเหตุจราจรทางบก

นอกจากนี้เป็นที่คาดหวังไว้ว่าผลที่ได้จากการวิจัยนี้จะเป็แนวทางที่ดีสำหรับทีมรักษาหรือฟื้นฟูว่าจะมีแนวทางในการรักษาควบคุมการฟื้นฟูสมรรถภาพผู้ป่วยในโรงพยาบาลอย่างเป็นระบบเพื่อป้องกันการเกิดภาวะแทรกซ้อนอันจะก่อให้เกิดความพิการได้ ซึ่งกระทรวงสาธารณสุขสามารถนำผลสรุปที่ได้ไปกำหนดนโยบายในการลดความรุนแรงจากการบาดเจ็บได้อย่างมีประสิทธิภาพ

วัตถุประสงค์ของโครงการวิจัย

1. เพื่อศึกษาความสัมพันธ์ของปัจจัยต่าง ๆ ที่เกี่ยวข้องกับการเกิดอุบัติเหตุจราจรทางบกกับความพิการถาวรที่เกิดขึ้นภายหลังการบาดเจ็บ
2. เพื่อจัดลำดับอิทธิพลของปัจจัยต่าง ๆ ว่ามีอิทธิพลต่อการเกิดความพิการถาวรหลังการบาดเจ็บรุนแรงจากสาเหตุอุบัติเหตุจราจรมากทางบก
3. เพื่อสร้างตัวแบบในการพยากรณ์ความพิการถาวรภายหลังการบาดเจ็บรุนแรงเมื่อทราบสาเหตุต่าง ๆ จากอุบัติเหตุจราจรทางบก
4. เพื่อสร้างนักวิจัยรุ่นใหม่ที่มีผลงานเป็นที่ยอมรับในระดับนานาชาติ

โดยจะเป็นการวิจัยแบบ Retrospective study โดยใช้ข้อมูลที่เก็บรวบรวมไว้แล้วโดยศูนย์สิรินธรเพื่อการฟื้นฟูสมรรถภาพทางการแพทย์ซึ่งทำการสำรวจไว้ในปี 2549 จำนวนทั้งสิ้น 24,435 รายการ จากนั้นทำการคัดกรองข้อมูล วิเคราะห์ความสัมพันธ์ของข้อมูลและรวมไปถึงการพัฒนาตัวแบบเพื่อการพยากรณ์โดยใช้เทคนิค Artificial neural networks โดยใช้โปรแกรมทางสถิติ R ซึ่งเป็นโปรแกรมที่ใช้งานได้ฟรีและได้รับความนิยมและถูกนำมาใช้อย่างแพร่หลายในงานวิจัยด้านต่างๆ และใช้เครื่องมือเพื่อพัฒนาโมเดล Weka ซึ่งเป็นโปรแกรมที่พัฒนาขึ้นด้วยภาษาจาวา ใช้งานได้ฟรีและเป็นที่ยอมรับในการใช้งานวิเคราะห์และสร้างตัวแบบในการทำเหมืองข้อมูล

ขอบเขตของงานวิจัยเป็นการวิจัยแบบ Retrospective study โดยใช้ข้อมูลที่เก็บรวบรวมไว้แล้วโดยศูนย์สิรินธรเพื่อการฟื้นฟูสมรรถภาพทางการแพทย์ซึ่งทำการสำรวจไว้ในปี 2549 จำนวนทั้งสิ้น 24,435 รายการ จากนั้นทำการคัดกรองข้อมูล และวิเคราะห์ความสัมพันธ์ของข้อมูลรวมไปถึงการพัฒนาตัวแบบเพื่อการพยากรณ์โดยใช้โปรแกรม R และโปรแกรมทางสถิติ SPSS ซึ่งเป็นโปรแกรมที่ได้รับความนิยมและถูกนำมาใช้อย่างแพร่หลายในงานวิจัยด้านต่าง ๆ จากนั้นนำผลที่ได้ มาวิเคราะห์เพิ่มเติมด้วยเทคนิคเหมืองข้อมูล (data mining techniques) แบบต่างๆ เพื่อพัฒนาเป็นโปรแกรมจำแนก เพื่อพยากรณ์ความรุนแรงของการบาดเจ็บ จากการวิจัยมาเผยแพร่ที่ศูนย์สิรินธรเพื่อการฟื้นฟูสมรรถภาพทางการแพทย์ เพื่อให้เกิดการทำงาน

ร่วมกันของทีมปฏิบัติงานของศูนย์ฯ และทีมนักวิจัยในการวิจัยและวิเคราะห์ข้อมูลขั้นสูง และทำการตีพิมพ์เผยแพร่ผลงานวิจัยในวารสารระดับนานาชาติจำนวน 1 เรื่อง

คำจำกัดความที่ใช้ในงานวิจัย

- 1. อุบัติเหตุจราจรทางบก** หมายถึง เหตุการณ์ที่เกิดขึ้นโดยไม่คาดคิดที่เกิดระหว่างการสัญจรไปมาบนทางบกของผู้ขับขี่ คนเดินเท้า หรือคนที่จูง ชี่ หรือไล่ต้อนสัตว์อันเกี่ยวเนื่องกับการคมนาคม การขนส่ง การเดินเท้า และยานพาหนะทุกชนิดที่ใช้บนถนนทั้งประเภทใช้เครื่องยนต์ หรือไม่ใช่เครื่องยนต์ที่ขับเคลื่อนรวมถึงล้อเลื่อนชนิดต่างๆ
- 2. บาดเจ็บไม่รุนแรง** หมายถึง กรณีบาดเจ็บจากอุบัติเหตุจราจรทางบกที่ได้รับการรักษาที่แผนกผู้ป่วยนอก และแผนกฉุกเฉินภายในระยะเวลาไม่เกิน 7 วัน ภายหลังเกิดอุบัติเหตุ และ แพทย์พิจารณาให้กลับบ้านได้ ทั้งนี้เจ้าหน้าที่จะติดตามอาการหลังจากกลับบ้านไป 3 สัปดาห์ เพื่อประเมินอาการและภาวะเสี่ยงของความพิการ
- 3. บาดเจ็บรุนแรง** หมายถึง กรณีผู้ได้รับบาดเจ็บจากอุบัติเหตุจราจรทางบก ที่เข้ารับการรักษาที่แผนกผู้ป่วยนอกและแผนกอุบัติเหตุและฉุกเฉินระยะเวลาไม่เกิน 7 วัน หลังประสบอุบัติเหตุ และแพทย์พิจารณาให้เข้าพักรักษาในโรงพยาบาล (Admit)
- 4. BR (Body Region)** หมายถึง ส่วนของร่างกายในแต่ละหมวดอวัยวะซึ่งจัดแบ่งตามระบบ IIS (Injury Severity Score) IIS ได้แก่
 - BR 1 การบาดเจ็บที่ศีรษะและลำคอ (รวมชั้นกลางและชั้นใน) : Head / Neck (Include middle and inner ear)
 - BR 2 การบาดเจ็บที่ใบหน้า (รวมกระบอกตา) : Face (Include eyeballs)
 - BR 3 การบาดเจ็บที่บริเวณทรวงอก : Thorax (Chest)
 - BR 4 การบาดเจ็บในช่องท้องและอวัยวะอุ้งเชิงกราน : Abdomen and pelvic contents
 - BR 5 การบาดเจ็บที่แขนขาและเชิงกราน : Extremities and pelvic girdle
 - BR 6 การบาดเจ็บภายนอกและผิวหนัง : External and body surface
- 5. AIS (Abbreviated Injury Scale)** คือ ระบบการจัดระดับความรุนแรงของการบาดเจ็บในแต่ละส่วนของร่างกายโดยการให้คะแนนตั้งแต่ต่ำไปมาก (จาก 1 ถึง 6) ใช้ตามระบบ AIS 85 โดยระบบการให้คะแนนนี้จะแบ่งเป็นระบบย่อยตามลักษณะของสิ่งที่ทำให้บาดเจ็บได้อีก 2 ลักษณะคือ Blunt (ทุ/ไม่มีคม) Penetrating (แหลม/มีคม)

รหัสคะแนน AIS 85 มีระดับความรุนแรง

1 Minor	เล็กน้อย
2 Moderate	ปานกลาง
3 Serious : not life threatening	มากแต่ไม่คุกคามต่อชีวิต
4 Severe : life threatening	มากและคุกคามต่อชีวิต
5 Critical : survival uncertain	วิกฤต, ไม่แน่ใจในโอกาสรอดชีวิต
6 Maximum injury	รุนแรงที่สุดส่วนใหญ่ไม่รอดชีวิต
9	ไม่ทราบว่ามีบาดเจ็บหรือไม่

6. **ความพิการ** หมายถึง ความผิดปกติหรือความบกพร่องทางร่างกาย ทางสติปัญญาหรือทางจิตใจที่ส่งผลให้บุคคลสูญเสียสมรรถภาพและเกิดความเสียเปรียบ ไม่สามารถปฏิบัติกิจกรรมประจำวันหรือทำสิ่งใดในการดำรงชีวิตได้ทัดเทียม หรือเท่ากับคนทั่วไป
7. **คนพิการ** หมายถึง คนที่มีความผิดปกติหรือบกพร่องทางร่างกาย ทางสติปัญญา หรือทางจิตใจ ตามประเภทและหลักเกณฑ์ที่กำหนดในกฎกระทรวงที่ 2 ออกตามความในพระราชบัญญัติการฟื้นฟูสมรรถภาพคนพิการ พ.ศ.2534 โดยประเมินหลังประสบอุบัติเหตุไปแล้ว 6 เดือน ซึ่งแบ่งออกเป็น 5 ประเภท
- 7.1.**คนพิการทางการมองเห็น** หมายถึง คนที่มีสายตาข้างที่ตีกว่าเมื่อใช้แว่นสายตาธรรมดาแล้วมองเห็นน้อยกว่า 6/18 หรือ 20/70 ลงไปจนมองไม่เห็นแม้แต่แสงสว่าง หรือ คนที่มีลานสายตาแคบกว่า 30 องศา สามารถแบ่งได้ 5 ระดับ

ตารางที่ 1 แสดงระดับความพิการของคนพิการทางการมองเห็น

ระดับที่	พิจารณาที่สายตา	พิจารณาที่ลานตา	ลักษณะความพิการ
1	6/18 ลงไปถึง 6/60 (สายตาเลือนราง)	แคบกว่า 30 องศา จนถึง 10 องศา	สายตาพิการ
2	น้อยกว่า 6/60 ลงไปถึง 3/60 (สายตาเลือนราง)	แคบกว่า 30 องศา จนถึง 10 องศา	สายตาพิการ
3	น้อยกว่า 3/60 ลงไปถึง 1/60	แคบกว่า 10 องศา จนถึง 5 องศา	ตาบอดขั้นหนึ่ง
4	น้อยกว่า 1/60 ลงไปถึงเห็น เพียงแสงสว่าง	แคบกว่า 5 องศาขึ้นไป	ตาบอดขั้นสอง
5	มองไม่เห็นแม้แต่แสงสว่าง	แคบกว่า 5 องศาขึ้นไป	ตาบอดขั้นสาม

7.2.คนพิการทางการได้ยินหรือสื่อความหมาย หมายถึง

- ก. คนพิการทางการได้ยิน หมายถึง คนที่ได้ยินเสียงที่มีความถี่ 500, 1,000, 2,000 เฮิรตซ์ในหูข้างที่ตีกว่า ที่มีความดังเฉลี่ยตามระดับต่อไปนี้
- ระดับที่ 1 26 – 40 เดซิเบล (หูตึงน้อย)
- ระดับที่ 2 41 – 55 เดซิเบล (หูตึงปานกลาง)
- ระดับที่ 3 56 – 70 เดซิเบล (หูตึงมาก)
- ระดับที่ 4 71 – 90 เดซิเบล (หูตึงอย่างรุนแรง)
- ระดับที่ 5 มากกว่า 90 เดซิเบล (หูหนวก)
- ข. คนพิการทางการสื่อความหมาย ได้แก่ คนที่มีความผิดปกติหรือความบกพร่องในการเข้าใจหรือการใช้ภาษาพูดจนไม่สามารถสื่อความหมายกับคนอื่นได้ ความผิดปกติทางการสื่อความหมายแบ่งเป็น 5 ระดับดังนี้
- ระดับที่ 1 สื่อความหมายในเนื้อหาที่มากกว่าการใช้ในกิจวัตรประจำวันหลังได้บั้งและมีปัญหาในการสื่อความคิดที่ซับซ้อน
- ระดับที่ 2 สื่อความหมายได้เฉพาะที่ใช้ในกิจวัตรประจำวันหลัก

- ระดับที่ 3 สื่อความหมายที่ใช้ในกิจวัตรประจำวันหลักได้รู้เรื่องบ้างไม่รู้เรื่องบ้าง
- ระดับที่ 4 สื่อความหมายได้เพียงการตอบรับหรือปฏิเสธ
- ระดับที่ 5 สื่อความหมายไม่ได้เลย

7.3.คนพิการทางกายหรือการเคลื่อนไหว หมายถึง คนที่มีความผิดปกติหรือความบกพร่องของร่างกายที่เห็นได้อย่างชัดเจน และไม่สามารถประกอบกิจวัตรหลักในชีวิตประจำวันได้หรือ คนที่มีการสูญเสียความสามารถในการเคลื่อนไหวมือ แขน ขา หรือลำตัว อันเนื่องมาจากแขนหรือขาขาด อัมพาตหรืออ่อนแรง โรคข้อหรืออาการปวดเรื้อรัง รวมทั้งโรคเรื้อรังของระบบการทำงานของร่างกายอื่นๆ ที่ทำให้ไม่สามารถประกอบกิจวัตรหลักในชีวิตประจำวันหรือดำรงชีวิตในสังคมอย่างคนปกติได้ ซึ่งสามารถแบ่งระดับความผิดปกติได้ดังนี้

ก. ความดังเฉลี่ยตามระดับต่อไปนี้

- ระดับที่ 1 ความผิดปกติหรือความบกพร่องของร่างกายที่ปรากฏให้เห็นชัดเจน แต่ยังสามารถประกอบกิจวัตรประจำวันได้
- ระดับที่ 2 ความผิดปกติหรือความบกพร่องในการเคลื่อนไหว ลำตัว มือ แขน หรือขา แต่ยัง สามารถประกอบกิจวัตรหลักในชีวิตประจำวันได้
- ระดับที่ 3 การสูญเสียความสามารถในการเคลื่อนไหว มือ แขน ลำตัว ซึ่งจำเป็นในการประกอบกิจวัตรหลักในชีวิตประจำวัน เพียง 1 รายการ
- ระดับที่ 4 การสูญเสียความสามารถในการเคลื่อนไหว มือ แขน ลำตัว ซึ่งจำเป็นในการประกอบกิจวัตรหลักในชีวิตประจำวัน เพียง 2 รายการ
- ระดับที่ 5 การสูญเสียความสามารถในการเคลื่อนไหว มือ แขน ลำตัว ซึ่งจำเป็นในการประกอบกิจวัตรหลักในชีวิตประจำวันมากกว่าครึ่งตัวหรือ 2 รายการขึ้นไป

7.4.คนพิการทางจิตใจหรือพฤติกรรม หมายถึง คนที่มีความผิดปกติหรือความบกพร่องทางจิตใจ หรือสมองในส่วนของการรับรู้ อารมณ์ ความคิดจนไม่สามารถควบคุมพฤติกรรมที่จำเป็นในการดูแลตนเองหรืออยู่กับผู้อื่น สามารถแบ่งระดับของความผิดปกติได้ดังนี้

- ระดับที่ 1 ผู้ที่ช่วยเหลือตัวเองไม่ได้ และรบกวนความสงบผู้อื่น
- ระดับที่ 2 ผู้ที่อยู่ตามลำพังไม่ยุ่งกับใคร และช่วยเหลือตัวเองไม่ได้
- ระดับที่ 3 ผู้ที่ช่วยเหลือตนเองได้บ้าง และอยู่ร่วมกับผู้อื่นได้บ้าง
- ระดับที่ 4 ผู้ที่ไม่ต้องพึ่งผู้อื่น ประกอบอาชีพเลี้ยงตัวได้
- ระดับที่ 5 ผู้ที่เลี้ยงตัวเองได้ และช่วยเหลือผู้อื่นได้

7.5.คนพิการทางสติปัญญาหรือการเรียนรู้ หมายถึง คนที่มีความผิดปกติหรือความบกพร่องทางสติปัญญาหรือสมองจนไม่สามารถเรียนรู้ด้วยวิธีการศึกษาปกติได้ ระดับความผิดปกติตามการทดสอบเชาว์ปัญญา สามารถแบ่งเป็น 5 ระดับดังนี้

- ระดับที่ 1 I.Q. = 80 - 90
- ระดับที่ 2 I.Q. = 70 - 79
- ระดับที่ 3 I.Q. = 50 - 69
- ระดับที่ 4 I.Q. = 35 - 49
- ระดับที่ 5 I.Q. น้อยกว่า 35

คนพิการทั้งหมดสามารถแบ่งเป็น 2 กลุ่มคือ กลุ่มที่เข้าเกณฑ์ที่สามารถจดทะเบียนเป็นคนพิการตามกฎหมายได้ และกลุ่มที่ไม่เข้าเกณฑ์ไม่สามารถจดทะเบียนเป็นคนพิการได้ ซึ่งมีหลักเกณฑ์แบ่งตามตารางที่ 2

ตารางที่ 2 แสดงเกณฑ์การประเมินคนพิการที่เข้าเกณฑ์การจดทะเบียนตาม พ.ร.บ.การฟื้นฟูสมรรถภาพ พ.ศ. 2534

ประเภทความพิการ	พิการเข้าเกณฑ์ตาม พ.ร.บ.	พิการไม่เข้าเกณฑ์ตาม พ.ร.บ.
1. คนพิการทางการมองเห็น	ระดับ 1 – 5	ตาบอดข้างเดียว ฯลฯ
2.		
2.1. คนพิการทางการได้ยิน	อายุน้อยกว่าหรือเท่ากับ 7 ปี ระดับ 2 – 5	ระดับ 1
	อายุมากกว่า 7 ปี ระดับ 3 – 5	ระดับ 1 – 2
2.2. คนพิการทางการสื่อ ความหมาย	ระดับ 3 – 5	ระดับ 1 – 2
3. คนพิการทางการกายหรือการ เคลื่อนไหว	ระดับ 3 – 5	ระดับ 1 – 2
4. คนพิการทางจิตใจหรือ พฤติกรรม	ระดับ 1 – 2	ระดับ 3 – 5
5. คนพิการทางสติปัญญาหรือ การเรียนรู้	ระดับ 3 – 5	ระดับ 1 – 2

8. ความสามารถในการทำกิจวัตรประจำวัน หมายถึง ความช่วยเหลือตัวเองของคนพิการในการทำกิจกรรมขั้นพื้นฐานต่างๆ ในชีวิตประจำวันที่คนทั่วไปทำได้ เช่น การรับประทานอาหาร การอาบน้ำ การเข้าห้องน้ำ การรับประทานอาหาร การดื่มน้ำ การจ่ายตลาด เป็นต้น โดยแบ่งเป็น 5 ระดับดังนี้
- ระดับที่ 1 ช่วยเหลือตนเองในการประกอบกิจวัตรหลักในชีวิตไม่ได้เลย หรือได้น้อยมาก และต้องการความช่วยเหลืออย่างมาก
- ระดับที่ 2 ช่วยเหลือตนเองในการประกอบกิจวัตรหลักในชีวิตประจำวันได้บ้างและต้องการการช่วยเหลือค่อนข้างมาก
- ระดับที่ 3 ช่วยเหลือตนเองในการประกอบกิจวัตรหลักในชีวิตประจำวันได้มาก อาจต้องการคนคอยแนะนำ หรือคอยระวังอยู่ด้านข้างหรือใช้อุปกรณ์เครื่องช่วยคนพิการหรือการปรับปรุงสิ่งของเครื่องใช้ประจำตัวแต่ไม่สามารถออกนอกบ้าน เพื่อศึกษาเล่าเรียนประกอบอาชีพหรือเข้าสังคมได้ด้วยตนเองแม้ในสิ่งแวดล้อมที่เอื้ออำนวยต่อคนพิการ
- ระดับที่ 4 ช่วยเหลือตนเองในการประกอบกิจวัตรหลักในชีวิตประจำวันได้เองอาจใช้อุปกรณ์เครื่องช่วยคนพิการหรือการปรับปรุงสิ่งของเครื่องใช้ประจำตัว สามารถออกนอกบ้าน เพื่อศึกษาเล่าเรียน ประกอบอาชีพ หรือเข้าสังคมได้ด้วยตนเองในสิ่งแวดล้อมที่เอื้ออำนวยต่อคนพิการ

ระดับที่ 5 ช่วยเหลือตัวเองในการประกอบกิจวัตรหลักในชีวิตประจำวันได้เองอาจใช้อุปกรณ์เครื่องช่วยคนพิการ หรือการปรับปรุงสิ่งของเครื่องใช้ประจำตัว สามารถออกนอกบ้าน เพื่อศึกษาเล่าเรียน ประกอบอาชีพ หรือเข้าสังคมได้เยี่ยมคนปรกติ

9. **ความสูญเสียอันเนื่องมาจากอุบัติเหตุจราจรทางบก** หมายถึง ความสูญเสียทรัพย์สิน ค่าใช้จ่ายในการจัดการอุบัติเหตุจราจรทางบก ความสูญเสียเชิงเศรษฐกิจเนื่องจากการขาดงาน ค่าใช้จ่ายด้านการรักษาพยาบาล ความสูญเสียซึ่งอันสะท้อนให้เห็นถึงความเจ็บปวดและความทุกข์ทรมานจากอุบัติเหตุของผู้ประสบอุบัติเหตุหรือครอบครัว

10. **สภาวะสุขภาพ** หมายถึงว่า การแสดงออกของบุคคลในลักษณะของความรู้ ความเข้าใจ ความเชื่อ ความรู้สึกนึกคิด ค่านิยมและการกระทำหรือไม่กระทำที่เกี่ยวข้องกับสุขภาพหรือมีผลต่อสุขภาพของตนเอง รวมถึงความสุขสมบูรณ์ของทรัพยากรบุคคลด้านกาย จิตใจและสังคม และสภาวะสุขภาพของบุคคลจะเป็นอย่างไรล้วนเกี่ยวข้องกับพฤติกรรมสุขภาพของบุคคล พฤติกรรมของกลุ่ม และสังคม การมีสภาวะดีเป็นสิ่งที่ทุกคนสามารถกระทำได้ในการดำเนินวิถีชีวิตของบุคคล ครอบครัว ชุมชน สังคม ภายใต้กรอบบุคลิกภาพของปัจเจกบุคคล สังคม และวัฒนธรรม

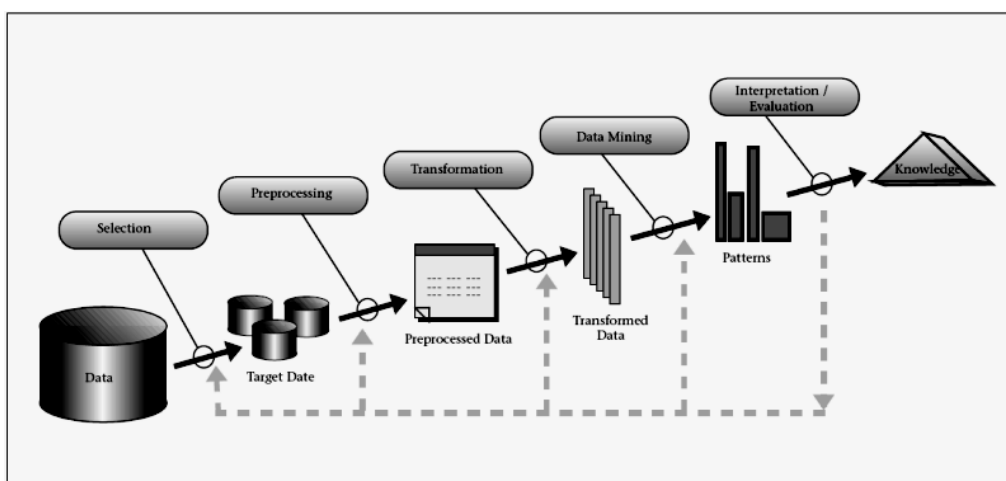
บทที่ 2

กรอบแนวคิดทฤษฎี และงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงรายละเอียดเกี่ยวกับความสำคัญและที่มาของปัญหาวิจัย รวมไปถึงการทบทวนวรรณกรรมของการศึกษาที่เกี่ยวข้อง และรายละเอียดโดยรวมของปัจจัยที่มีอิทธิพลต่อระดับความรุนแรงของการพิการจาก อุบัติเหตุจากรถทางบก ตลอดจนกระบวนการพัฒนาตัวแบบเพื่อการพยากรณ์โดยใช้เทคนิคแบบต่างๆ เช่น ต้นไม้ตัดสินใจ (decision tree) นาอิวเบย์ (naïve bayes) โครงข่ายประสาทเทียม เป็นต้น ซึ่งรายละเอียดเป็นดังนี้

วิธีการดำเนินการเพื่อค้นหาค้นหาองค์ความรู้ (Knowledge Discovery in Database)

การค้นหาค้นหาองค์ความรู้ เป็นกระบวนการในการค้นหาลักษณะแฝงของข้อมูลที่อยู่ในกลุ่มข้อมูลจำนวนมาก ซึ่งมีขั้นตอนการทำ Data Mining เป็นกระบวนการที่สำคัญในการค้นหาลักษณะที่น่าสนใจของข้อมูลเหล่านี้ เช่น รูปแบบ ความสัมพันธ์ การเปลี่ยนแปลง โครงสร้างที่เด่นชัด หรือ ลักษณะที่ผิดปกติของข้อมูลจากข้อมูลจำนวนมากๆ ที่เก็บอยู่ในฐานข้อมูล หรือแหล่งที่เก็บข้อมูลอื่นๆ ซึ่งวิธีการต่างๆ ที่นำมาใช้ในการทำ mining นี้ก็มีวัตถุประสงค์ต่างกันขึ้นอยู่กับผลลัพธ์ของ กระบวนการโดยรวมที่ต้องการ ดังนั้นจึงควรมีการนำเสนอวิธีการที่หลากหลายสำหรับเป้าหมายที่แตกต่างกัน เพื่อให้ได้ผลลัพธ์ ที่เหมาะสมตามที่ต้องการ หลังจากนำไปใช้งานแล้ว และเนื่องจากความแพร่หลายของการจัดเก็บข้อมูลในลักษณะที่เป็น รูปแบบทางอิเล็กทรอนิกส์ และความต้องการในการเปลี่ยนข้อมูลเหล่านั้นให้เป็นข้อมูลที่มีประโยชน์ต่อการนำไปประยุกต์ ใช้ในงานด้านต่างๆ เช่น การวิเคราะห์ด้านการตลาด การบริหารธุรกิจ รวมถึงระบบที่ช่วยสนับสนุนการตัดสินใจ เป็นต้น ดังนั้นจึงทำให้การนำเหมืองข้อมูล (data mining) มาใช้ โดยการทำงานแบ่งออกเป็น 5 ขั้นตอนดังนี้ (KDD2011)



ภาพที่ 1 แสดงกระบวนการของ KDD

1. การคัดเลือกข้อมูล (Data Selection) เป็นการระบุถึงแหล่งข้อมูลที่จะนำมาใช้ในการทำ mining รวมถึงการนำข้อมูลที่ต้องการออกมาจากฐานข้อมูลเพื่อทำการพิจารณาในเบื้องต้นต่อไป
2. การกรองข้อมูล (Data Cleaning) เป็นกระบวนการที่ทำให้เกิดความมั่นใจในคุณภาพของข้อมูลที่จะนำมาใช้ วิเคราะห์ ว่าถูกต้อง โดยการนำข้อมูลที่ไม่ถูกต้องออก
3. การแปลงรูปแบบข้อมูล (Data Transformation) เป็นการแปลงข้อมูลที่เลือกมาให้อยู่ในรูปแบบที่เหมาะสม สำหรับการนำไปใช้วิเคราะห์ตามอัลกอริทึม (Algorithm) และแบบจำลองที่ใช้ในการทำ data mining ต่อไป
4. การทำ Mining ข้อมูล (Data Mining) การใช้เทคนิคภายใน Data Mining เพื่อทำการ Mine ข้อมูล โดยทั่วไป ประเภทของงานตามลักษณะของแบบจำลองที่ใช้ในการทำ Data Mining นั้นสามารถแบ่งกลุ่มได้เป็น 2 ประเภทใหญ่ๆ คือ
 - o **การสร้างแบบจำลองในการทำนาย (predictive modeling, supervised modeling)** เป็นการคาดคะเนลักษณะหรือประมาณค่าที่ชัดเจนของข้อมูลที่จะเกิดขึ้นโดยใช้พื้นฐานจากข้อมูลที่ผ่านมาในอดีต ในที่นี้ ทุกข้อมูลจะมีคุณสมบัติหนึ่งเรียกว่าคำตอบ (output) ซึ่งค่าของคุณสมบัตินี้จะเป็นค่าที่ใช้ในการทำนายผลของข้อมูล อัลกอริทึมประเภทนี้จะมุ่งเน้นในการแบ่งแยกข้อมูลออกเป็นกลุ่มตามค่าคุณสมบัติของคำตอบ ซึ่งถ้าค่าคุณสมบัติของคำตอบมีค่าไม่ต่อเนื่อง จะเรียกกระบวนการที่ใช้แบ่งแยกว่า การจำแนก (classification) ถ้าค่าคุณสมบัติของคำตอบมีค่าต่อเนื่องจะเรียกกระบวนการที่ใช้แบ่งแยกว่า การถดถอย (regression)
 - o **การสร้างแบบจำลองในการบรรยาย (descriptive modeling, unsupervised modeling)** เป็นการหาแบบจำลองเพื่ออธิบายลักษณะบางอย่างของข้อมูลที่มีอยู่ ซึ่งโดยส่วนมากจะเป็นลักษณะการแบ่งกลุ่มให้กับข้อมูล ในที่นี้ อาจเป็นการหาความสัมพันธ์ต่างๆ (association) หรือหาการจัดกลุ่มข้อมูล (clustering) ซึ่งไม่ได้มีจุดมุ่งหมายเพื่อการทำนาย (บุญเสริม กิจศิริกุล, 2545 ; ปาลจิตต์ พันทวาที, 2552)
5. การวิเคราะห์และประเมินผลลัพธ์ที่ได้ (Result Analysis and Evaluation) เป็นขั้นตอนการแปลความหมาย และการประเมินผลลัพธ์ที่ได้ว่ามีความเหมาะสมหรือตรงกับวัตถุประสงค์ที่ต้องการหรือไม่ โดยทั่วไปควรมีการแสดงผลในรูปแบบ ที่สามารถเข้าใจได้โดยง่าย

การคัดเลือกข้อมูล

ข้อมูลที่จะนำมาใช้เป็นข้อมูลทุติยภูมิที่ได้มาจากรายงาน เรื่อง “อุบัติการณ์ความพิการและผลกระทบที่เกิดขึ้นจากสาเหตุอุบัติเหตุจราจรทางบก พ.ศ. 2549” โดยการจัดเก็บข้อมูลเป็นไปตามระเบียบวิธีที่เสนอ โดย แพทย์หญิง ดารณี สุวพันธ์ ที่เสนอไว้ในรายงาน

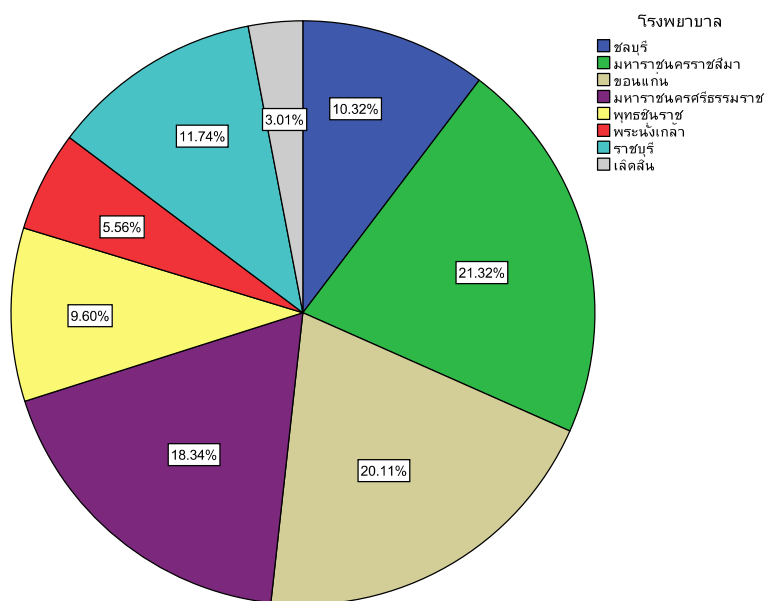
การกรองข้อมูล

ข้อมูลที่ได้รับมาเป็นข้อมูลอุบัติเหตุจราจรทางบกที่ผู้ป่วยมีอาการรุนแรง (กล่าวคือต้อง admit) มีทั้งหมด 9,738 ระเบียบ โดยจัดเก็บจากโรงพยาบาลทั้งหมด 8 แห่ง ดังแสดงด้วยแผนที่ ดังภาพที่ 2



ภาพที่ 2 แผนที่แสดงตำแหน่งของโรงพยาบาลทั้ง 8 แห่ง

ตามจังหวัดต่างๆ ในประเทศไทย ซึ่งช่วงเดือนของการเกิดอุบัติเหตุสามารถจำแนกตามโรงพยาบาลได้ดังภาพที่ 3 และตามช่วงเดือนที่เกิดอุบัติเหตุได้ดังตารางที่ 3

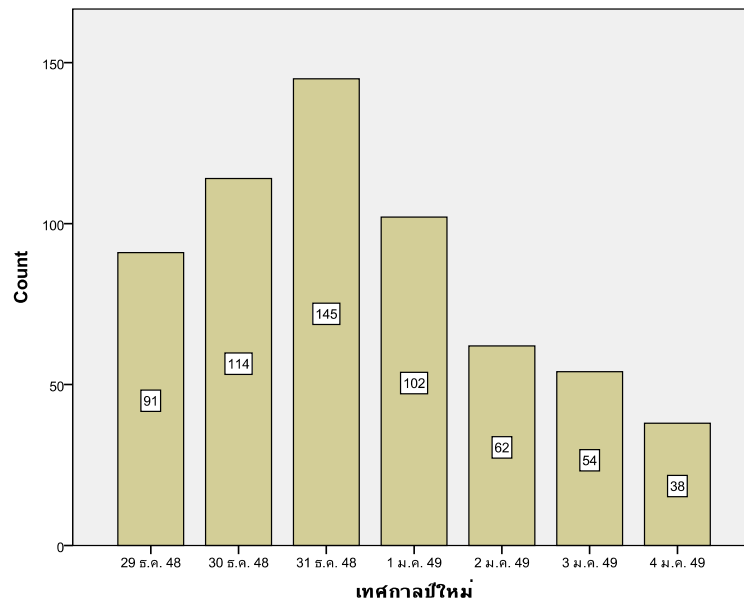


ภาพที่ 3 แสดงสัดส่วนการเกิดอุบัติเหตุในแต่ละโรงพยาบาล

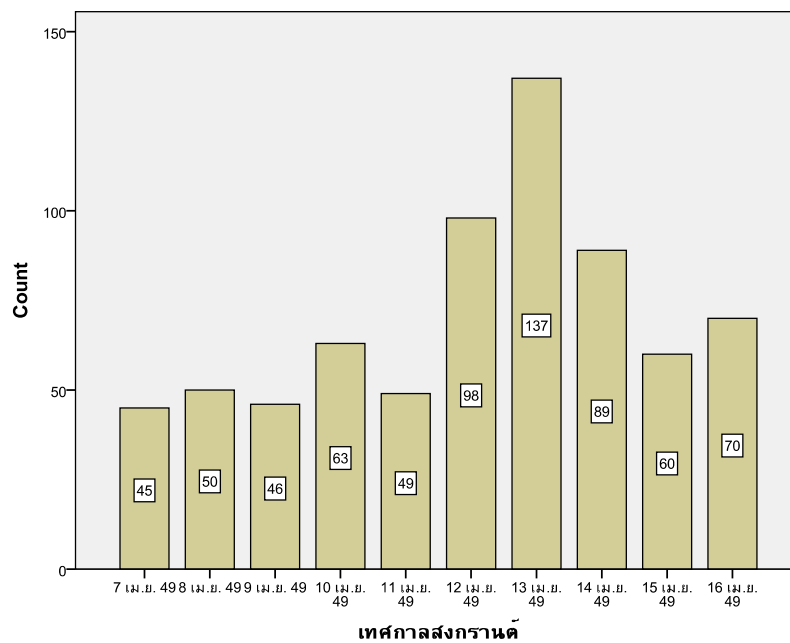
ตารางที่ 3 แสดงจำนวนอุบัติเหตุจำแนกตามโรงพยาบาล และเดือนที่เกิดเหตุ

		โรงพยาบาล								Total
		ชลบุรี	มหาสารคาม นครราชสีมา	ขอนแก่น	มหาสารคาม ศรีธรรมราช	พุทธ ชินราช	พระนั่ง เกล้า	ราชบุรี	เลิดสิน	
ช่วงเดือนที่เกิดเหตุ	มกราคม	289	395	317	273	79	71	225	48	1697
	กุมภาพันธ์	161	251	291	289	144	65	175	44	1420
	มีนาคม	173	324	390	306	194	116	238	52	1793
	เมษายน	148	380	307	324	202	99	197	50	1707
	พฤษภาคม	123	286	292	266	147	86	172	51	1423
	1 - 25 มิถุนายน	50	278	246	240	94	76	106	33	1123
	26 - 31 ธันวาคม	57	161	115	87	74	28	29	15	566
Total		1001	2075	1958	1785	934	541	1142	293	9726

โดยจำนวนอุบัติเหตุรุนแรงที่เกิดในช่วง เทศกาลสงกรานต์ และปีใหม่ มีจำนวนทั้งหมด 606 และ 707 ตามลำดับ ซึ่งรวมทั้งเป็น 1,313 คิดเป็น 13.5% ของอุบัติเหตุทั้งหมด โดยมีความถี่ของการเกิดแฉกแฉงตามวันได้ดัง ภาพที่ 4 และ 5

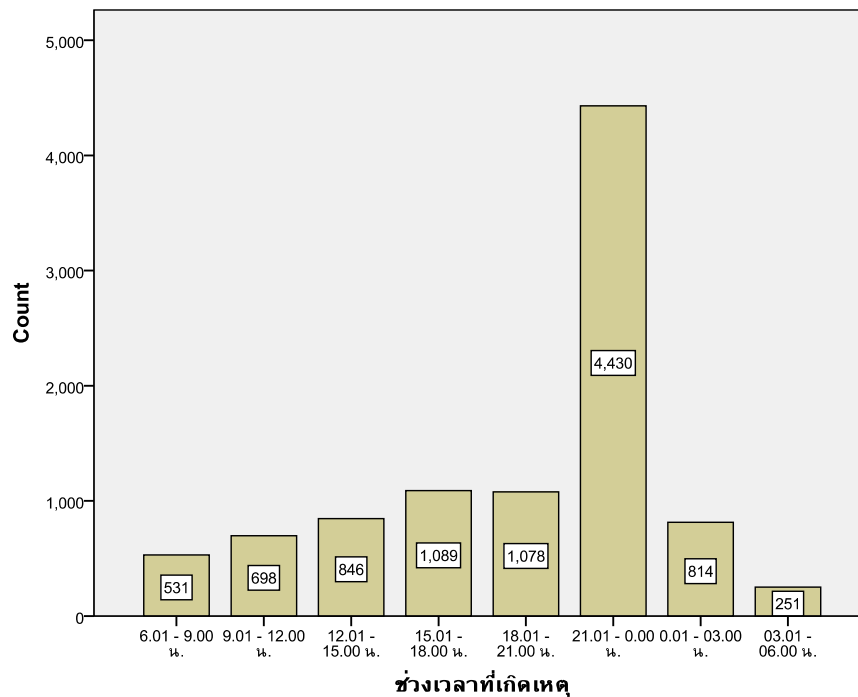


ภาพที่ 4 กราฟแท่งแสดงการเกิดอุบัติเหตุ ช่วงเทศกาลปีใหม่



ภาพที่ 5 กราฟแท่งแสดงการเกิดอุบัติเหตุ ช่วงเทศกาลสงกรานต์

และเมื่อจำแนกตามช่วงเวลาการเกิดเหตุจะได้ดังกราฟแท่งภาพที่ 6 ช่วงเวลา 3 ชั่วโมงถึงเที่ยงคืนเป็นช่วงที่เกิดอุบัติเหตุมากที่สุด



ภาพที่ 6 แสดงจำนวนอุบัติเหตุจำแนกตามเวลาที่เกิดเหตุ

โดยเมื่อจำแนกตามอาการบาดเจ็บ ในเบื้องต้นที่แบ่งเป็น 5 ระดับ คือไม่มีความพิการ เสียชีวิต พิการที่เข้าหลักเกณฑ์ พิการที่ยังไม่เข้าหลักเกณฑ์ และผิดปกติอาจจะพิการ โดยแยกตามช่วงเวลาซึ่งแสดงได้ดังตารางที่ 4

ตารางที่ 4 จำนวนผู้พิการจำแนกตามช่วงเวลาที่เกิดเหตุ

		ช่วงเวลาที่เกิดเหตุ								Total
		6.01 - 9.00 น.	9.01 - 12.00 น.	12.01 - 15.00 น.	15.01 - 18.00 น.	18.01 - 21.00 น.	21.01 - 0.00 น.	0.01 - 3.00 น.	3.01 - 6.00 น.	
สรุปการประเมิน	ไม่มีความพิการ	461	632	756	977	956	3945	691	221	8639
ความพิการ	เสียชีวิต	44	39	48	70	65	288	73	18	645
	พิการที่เข้าหลักเกณฑ์	20	16	27	25	42	125	28	8	291
	พิการที่ยังไม่เข้าหลักเกณฑ์	1	3	3	3	7	20	7	1	45
	ผิดปกติอาจจะพิการ	2	3	2	2	5	20	5	1	40
Total		528	693	836	1077	1075	4398	804	249	9660

โดยในจำนวนเมื่อสิ้นสุดการติดตามอาการในเวลา 3 เดือน จะมีเพียง 1,064 ระเบียบที่สุดท้ายแล้วเกิดการเสียชีวิต หรือพิการ ซึ่งคิดเป็นร้อยละ 11.1 ของอุบัติเหตุที่มีความรุนแรง และเมื่อจำแนกตามความพิการ และการเสียชีวิต จะได้ข้อมูลดังตารางที่ 5

ตารางที่ 5 แสดงจำนวนผู้พิการและเสียชีวิต ซึ่งได้ข้อมูลมากจากค่าตัวแปร pikan และ r_pikan

	ความถี่	ร้อยละ
1. ไม่มีความพิการ	8560	88.9
2. เสียชีวิต	641	6.7
3. พิการเข้าเกณฑ์	306	3.2
4. พิการไม่เข้าเกณฑ์	117	1.2
รวม	9624	

โดยในข้อมูลชุดนี้ได้จัดเก็บข้อมูลต่างๆ กว่า 50 ปัจจัยหรือตัวแปร ได้แก่

ชื่อตัวแปร	ความหมาย
prov	ชื่อโรงพยาบาล
sit	สิทธิการรักษา
adate	วันที่เกิดเหตุ
atime	เวลาเกิดเหตุ
age	อายุ
sex	เพศ
occu	อาชีพ
injp	ผู้บาดเจ็บเป็น
injt	พาหนะ
injer	การบาดเจ็บเกิดจาก
tinj	ลักษณะการบาดเจ็บ
risk1	แอลกอฮอล์
risk2	ยาเสพติด
risk3	เข็มฉีดยา
risk4	หมวกนิรภัย
airway	ดูแลการหายใจ
blood	ห้ามเลือด
splint	Splint/Slab
iv	IV Fluid
a2h	การมาโรงพยาบาล
a2h1	จากที่เกิดเหตุอื่น
a2h2	ผู้นำส่งคือ
h2h	มาโดย
h2h1	ใบส่งต่อ

ชื่อตัวแปร	ความหมาย
h2h2	Ambulance
pikan	ไม่พิการ ตาย หรือพิการ
r_pikan	คนพิการ
rr_diag	วิเคราะห์ความพิการเข้าเกณฑ์หรือไม่
t_comma	Hx. Consciousness
staer	ออกจาก ER โดย
disc	จำหน่ายจากหอผู้ป่วยโดย
a221	ลักษณะพิการก่อนเกิดเหตุ
a222	การจดทะเบียนตาม พรบ.
a31	ถนนที่เกิดเหตุ
a41	ทางลาดชัน
a421	ทางตรง
a43	ทางแยก
a44	ทางม้าลาย
a45	ทางโค้ง
a46	ทางมีสิ่งกีดขวาง
a51	ตำแหน่งการนั่งของผู้บาดเจ็บ
a6	ลักษณะการชน

เป็นต้น นอกจากนี้ยังมีตัวแปรอื่นๆ ที่ระบุค่าอื่นๆ เช่น ลักษณะการดูแลผู้พิการ ผู้ดูแลคนพิการ ที่อยู่ การวินิจฉัยอาการในขั้นตอนต่างๆ เป็นต้น ซึ่งในที่นี้ผู้วิจัยได้คัดเลือกมาเพียงบางส่วนเท่านั้น

การแปลงรูปแบบข้อมูล

ข้อมูลต่างๆ ได้ถูกปรับหน่วยให้สอดคล้องกับเทคนิคที่เลือกใช้ในการสร้างตัวแบบพยากรณ์ เช่น อายุ และ เวลาที่เกิดเหตุ ลักษณะยานพาหนะและ ถนน เป็นต้น โดยในที่นี้อายุจะปรับเป็นค่าระดับ 8 ระดับ

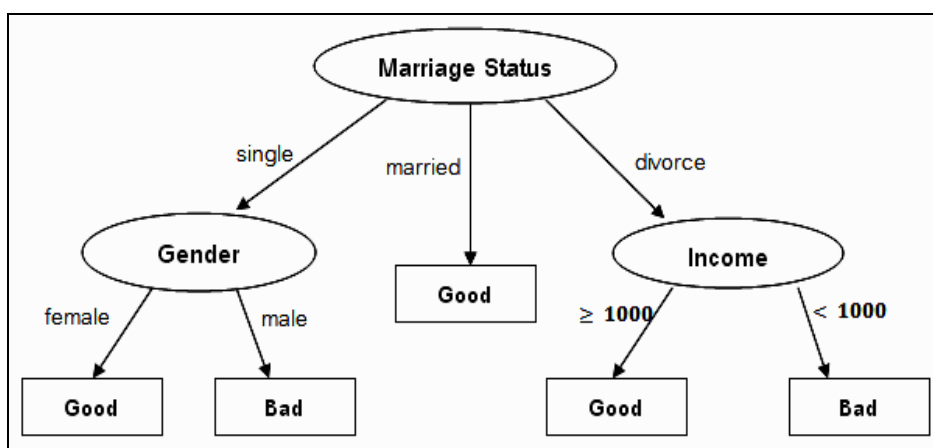
การทำ Mining ข้อมูล

เทคนิคต้นไม้ตัดสินใจ (Decision tree)

ต้นไม้ตัดสินใจ เป็นการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปของต้นไม้ตัดสินใจ ซึ่งต้นไม้ตัดสินใจนั้นมีการทำงานแบบการเรียนรู้แบบมีผู้สอน (Supervised Learning) คือ การสร้างแบบจำลองการจำแนกได้จากกลุ่มตัวอย่างของข้อมูลที่ได้กำหนดไว้ก่อนล่วงหน้า หรือที่เรียกว่า ชุดข้อมูลเรียนรู้ (Training Set) ได้อัตโนมัติ และสามารถพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจำแนกได้ด้วย

ต้นไม้ตัดสินใจเป็นวิธีที่ได้รับความนิยมเนื่องจากความไม่ซับซ้อนของอัลกอริทึม สามารถตีความและเข้าใจลักษณะของรูปแบบข้อมูลได้ง่าย

รูปแบบของต้นไม้ (Tree) จะประกอบด้วยโหนด (Node) โดยโหนดแรกสุดจะเรียกว่า โหนดราก (Root Node) จากโหนดรากก็จะแตกออกเป็นโหนดลูก และที่โหนดลูก ก็จะสามารถแตกลูกของตัวเอง ซึ่งโหนดที่ระดับสุดท้ายจะเรียกว่า โหนดใบ (Leaf Node) โดยโหนดใบจะเป็นคำตอบของต้นไม้ หลักการพื้นฐานของการสร้างต้นไม้ตัดสินใจ จะเป็นการสร้างในลักษณะจากบนลงล่าง (top – down) คือ จะเริ่มจากการสร้างรากของต้นไม้ก่อนแล้วจึงแตกกิ่งไปจนถึงใบ



ภาพที่ 7 แสดงตัวอย่างของต้นไม้ตัดสินใจ

จากภาพจะเห็นว่า ในโหนดราก จะได้ลักษณะประจำ (Attribute) คือ Marriage Status โดยมีการแตกกิ่งตามค่าของลักษณะประจำเป็น 3 แฉก คือ single, married และ divorce ถ้าลักษณะประจำ Marriage Status มีค่าเป็น married โหนดลูกที่ได้จะเป็นใบ หรือในที่นี้คือ Good แต่ถ้า ค่าของลักษณะเฉพาะเป็น single หรือ divorce โหนดลูกของคำตอบเหล่านี้จะเป็นลักษณะประจำอื่นที่สามารถแตกออกเป็นโหนดลูกได้อีก จนลำดับสุดท้ายของโหนดจะเป็นโหนดใบที่ให้ผลลัพธ์เป็น Good หรือ Bad ซึ่งเป็นคำตอบของการจำแนกนี้

จะเห็นว่า จากโหนดรากจนถึงโหนดใบจะมีเพียงเส้นทางเดียวเท่านั้น ซึ่งเส้นทางนี้จะอธิบายถึงกฎที่ใช้สำหรับการจำแนกของแต่ละกลุ่ม ซึ่งในแต่ละโหนดใบนั้นอาจเป็นกลุ่มเดียวกัน ซึ่งเกิดจากเหตุผลที่แตกต่างกันได้

ประสิทธิภาพของต้นไม้ตัดสินใจไม่ได้อยู่ที่การสร้างต้นไม้ตัดสินใจเพื่อให้สามารถจำแนกชุดข้อมูลเรียนรู้ได้อย่างถูกต้องเท่านั้น แต่ต้องสามารถจำแนกข้อมูลจากตัวอย่างใหม่ๆ ที่นอกเหนือจากนั้นได้อย่างถูกต้องด้วย ดังนั้นการสร้างต้นไม้ตัดสินใจจึงต้องมีชุดข้อมูลทดสอบ (test set) ที่จะใช้ตรวจสอบความถูกต้องของต้นไม้ตัดสินใจด้วย

ข้อจำกัดของต้นไม้ตัดสินใจ

- การจำแนกกลุ่มแบบต้นไม้ตัดสินใจในกรณีเป็นข้อมูลที่มีค่าต่อเนื่อง เช่น ข้อมูลรายได้ ข้อมูลราคา ต้องทำการแปลงให้อยู่ในช่วงหรือตัดเป็นกลุ่มก่อน

- เมื่ออัลกอริทึมเลือกที่จะใช้ค่าไหนเป็นตัวจำแนกแล้วก็จะไม่สนใจค่าอื่นที่อาจมีความสำคัญเช่นกัน
- การจัดการกับข้อมูลที่ไม่ทราบค่า อาจมีผลกระทบกับผลของต้นไม้ตัดสินใจ
- ต้นไม้ที่มีระดับชั้นมากเกินไป จะทำให้ข้อมูลที่ผ่านโหนดแตกออกเป็นชิ้นเล็กชิ้นน้อย ซึ่งข้อมูลเหล่านั้นจะไม่มีประโยชน์ในการทำมาใช้ในการวิเคราะห์
- ปัญหาเรื่อง ความเฉพาะเจาะจงเกินไป (Overfitting) หรือ การฝึกมากเกินไป (Overtraining) เกิดจากการที่แบบจำลองได้เรียนรู้เข้าไปถึงรายละเอียดของข้อมูลมากเกินไป จะทำให้เกิดโหนดที่เป็นส่วนเฉพาะเจาะจงกับกลุ่มข้อมูลที่ใช้ในการเรียนรู้ ซึ่งต้องหาวิธีการในการตัดกิ่งนี้ออกไป (เชษฐา จิรไพศาลกุล, 2550 ; นฤพนธ์ ว่องประชาณุกุล, 2548 ; Wikipedia, 2555)

เทคนิคนาอิวเบย์ (Naïve bayes)

ตัวจำแนกแบบนาอิวเบเซียน หรือ เบย์อย่างง่าย (Bayesian Classification) เป็นขั้นตอนหนึ่งที่มีประสิทธิภาพมากในการจำแนกข้อมูล โดยการเรียนรู้ปัญหาที่เกิดขึ้นเพื่อนำมาสร้างเงื่อนไขการจำแนกข้อมูลด้วยการใช้ทฤษฎีความน่าจะเป็นแบบมีเงื่อนไข

ทฤษฎีความน่าจะเป็นแบบมีเงื่อนไข ถ้าเหตุการณ์ A และ B เป็นเหตุการณ์ใดๆ ที่ไม่เป็นอิสระต่อกัน เราสามารถหาความน่าจะเป็นของเหตุการณ์ A เมื่อทราบเหตุการณ์ B ที่เกิดขึ้นแล้ว เรียกว่า ความน่าจะเป็นแบบมีเงื่อนไข เขียนแทนด้วย $P(A|B)$ ซึ่งเกิดจากความน่าจะเป็นของเหตุการณ์ A และเหตุการณ์ B ที่เกิดขึ้นด้วยกัน เขียนแทนด้วย $P(A,B)$ หาด้วยความน่าจะเป็นของเหตุการณ์ B เขียนแทนด้วย $P(B)$ ดังสมการ

$$P(A|B) = \frac{P(A,B)}{P(B)} \text{ เมื่อ } P(B) > 0 \quad (1)$$

การแบ่งกลุ่มแบบนาอิวเบเซียน (Naive Bayesian Classifier) เป็นการคำนวณหาความน่าจะเป็นของแต่ละกลุ่มข้อมูล ซึ่งในที่นี้เรียกว่า คลาส (Class) เมื่อกำหนดลักษณะประจำ (Attribute) และค่าของลักษณะประจำแต่ละตัวที่จะใช้ในการทำนาย ซึ่งเป็นการคำนวณหาความน่าจะเป็นของแต่ละคำตอบที่เป็นไปได้ในคลาสแล้วทำการเปรียบเทียบกัน ค่าความน่าจะเป็นที่สูงที่สุดของคลาสใดๆ จะเป็นผลของการทำนายเพียงค่าเดียว โดยถือว่าลักษณะประจำแต่ละตัวมีความเป็นอิสระต่อกัน (Class conditional independence) ซึ่งมีการทำงานดังนี้

1) ให้ X เป็นเซตของเหตุการณ์หนึ่งๆ จะได้ $X = \{x_1, x_2, \dots, x_n\}$ เมื่อกำหนดให้ x_1 คือค่าของ X ในลักษณะประจำที่ $A_1, A_2, \dots, A_n$ ตามลำดับ

2) สมมติให้มีคลาสที่เป็นไปได้จำนวน m คลาส คือ $C_1, C_2, \dots, C_m$ สำหรับเหตุการณ์ X โดยการทำนายจะเลือกค่าความน่าจะเป็นของคลาสที่สูงที่สุดที่ทำให้เกิดเหตุการณ์ X มาเป็นผลของการทำนาย

$$P(C_i|X) > P(C_j|X) \quad (2)$$

สำหรับ $1 \leq j \leq m, j \neq i$

จากทฤษฎีของเบย์

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \quad (3)$$

3) เมื่อ $P(X)$ เป็นค่าคงที่สำหรับทุกคลาส ดังนั้น การพิจารณาจึงเลือกจากค่า $P(X|C_i)P(C_i)$ ที่สูงที่สุด โดย $P(C_i) = s_i / S$ เมื่อ S_i คือจำนวนข้อมูลที่เป็นคลาส C_i และ S คือจำนวนข้อมูลทั้งหมด

4) เนื่องจากนาอ์ฟเบเซียนมีเงื่อนไขว่าลักษณะประจำแต่ละตัวเป็นอิสระต่อกัน (Independent) ดังนั้น

$$P(X|C_i) = \prod_{k=1}^n P(X_k | C_i) \quad (4)$$

ค่าความน่าจะเป็น $P(x_1|C_i), P(x_2|C_i), \dots, P(x_n|C_i)$ สามารถคำนวณได้จาก ลักษณะประจำของข้อมูลที่นำมาทำการเรียนรู้ โดยแบ่งเป็น 2 กรณี คือ

ถ้า A_k เป็นข้อมูลเชิงคุณภาพ $P(x_k|C_i) = \frac{S_{ik}}{S_i}$ เมื่อ S_{ik} คือจำนวนข้อมูลที่เป็นคลาส C_i ซึ่งมีค่า x_k สำหรับลักษณะประจำที่ A_k และ S_i คือจำนวนข้อมูลที่เป็นคลาส C_i

ถ้า A_k เป็นข้อมูลเชิงปริมาณ จะได้

$$P(X_k|C_i) = g(x_k; \mu_{C_i}, \sigma_{C_i}) \quad (5)$$

$$P(X_k|C_i) = \frac{1}{\sqrt{2\pi}\sigma_{C_i}} e^{-\frac{(x_k - \mu_{C_i})^2}{2\sigma_{C_i}^2}} \quad (6)$$

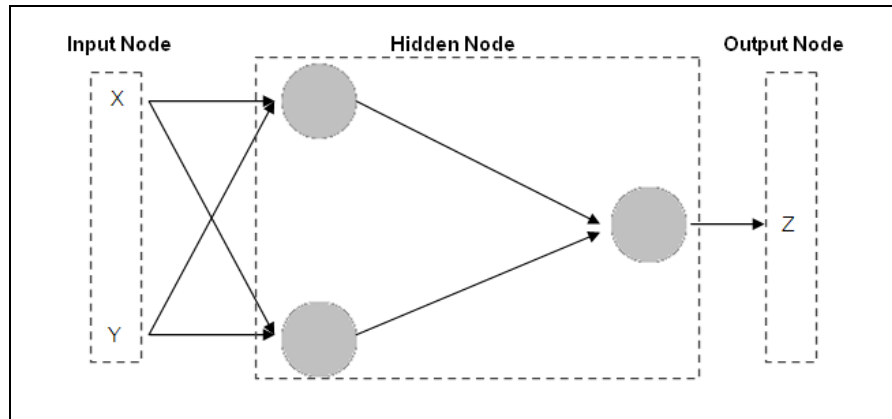
เมื่อ $g(x_k; \mu_{C_i}, \sigma_{C_i})$ คือ ฟังก์ชันเกาส์ (Gaussian Function) สำหรับลักษณะประจำ $A_k; \mu_{C_i}$ และ σ_{C_i} คือ ค่าเฉลี่ย และ ส่วนเบี่ยงเบนมาตรฐานของลักษณะประจำ A_k ที่เป็นคลาส C_i ตามลำดับ

5) ในการทำนายคลาสของเหตุการณ์ X จะทำการคำนวณ $P(X|C_i)P(C_i)$ สำหรับคลาสทุกคลาส และเหตุการณ์ X จะถูกจัดอยู่ในคลาสที่มีค่า $P(X|C_i)P(C_i)$ สูงสุด (พยุ่ง มีสัจ และ ญัญญา ห่อประชุม, 2548 ; สิทธิโชค มุกดาสกุลภิบาล, 2551)

เทคนิคโครงข่ายประสาทเทียม (Artificial neural networks)

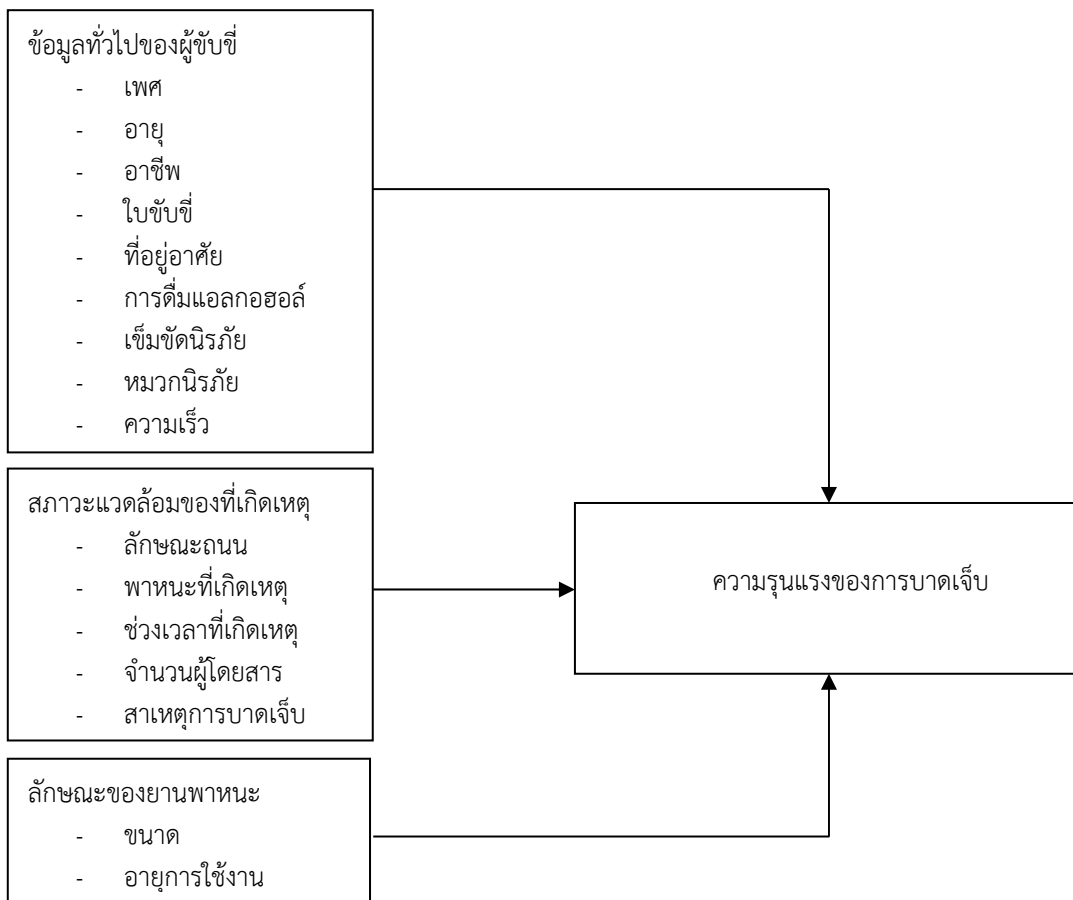
เทคนิคโครงข่ายประสาทเทียม (Artificial neural networks) ถูกนำมาใช้อย่างแพร่หลายในการศึกษารูปแบบความสัมพันธ์ของตัวแปรกลุ่มต่าง ๆ เมื่อรูปแบบความสัมพันธ์มีความซับซ้อน โดยเฉพาะอย่างยิ่งเมื่อชุดข้อมูลที่เก็บรวบรวมมาไม่เป็นไปตามข้อกำหนดเบื้องต้นของการวิเคราะห์ทางสถิติทั่วไป และจากรายงานการวิจัยในปัจจุบันพบว่าเทคนิคโครงข่ายประสาทเทียม สามารถใช้งานได้ดีในการสร้างตัวแบบเพื่อการพยากรณ์ และการศึกษาความสัมพันธ์ของกลุ่มตัวแปรต่าง ๆ ดังจะเห็นได้จากบทความวิจัยที่เผยแพร่อย่างต่อเนื่อง รวมทั้งมีการพัฒนารูปแบบต่าง ๆ ในขั้นตอนการทำงานของโครงข่ายประสาทเทียมเพื่อให้มี

ประสิทธิภาพมากยิ่งขึ้น การประยุกต์เทคนิคโครงข่ายประสาทเทียมเข้ามาวิเคราะห์ข้อมูล สามารถทำได้โดยการกำหนดรูปแบบโครงสร้างของโครงข่าย กล่าวคือจำนวนโหนดที่ใช้และจำนวนเลเยอร์ซ่อน ดังแสดงในภาพ 3 มีจำนวนข้อมูลเข้า 2 ค่าและข้อมูลออก 1 ค่า โดยในที่นี้ใช้โครงสร้างแบบ 1 เลเยอร์ซ่อนและ 2 โหนดซ่อน



ภาพที่ 8 แสดงโครงสร้างของชั้นเครือข่าย (Network Layer)

โดยกรอบและแนวคิดของงานวิจัยสามารถนำเสนอ ดังภาพที่ 3 โดยรวบรวมเกี่ยวกับอุบัติการณ์ความตาย ความพิการและผลกระทบที่เกิดขึ้นจากสาเหตุอุบัติเหตุจราจรทางบกเพื่อหารูปแบบความสัมพันธ์ของความรุนแรงของการพิการหรือตาย อันเนื่องมาจากปัจจัยต่าง ๆ ไม่ว่าจะเป็นปัจจัยด้านข้อมูลทั่วไปและพฤติกรรมการขับขี่ของผู้ขับขี่แต่ละคน ปัจจัยด้านสิ่งแวดล้อม เช่น สภาพของถนน และปัจจัยด้านสภาพของยานพาหนะที่ใช้ขับขี่



ภาพที่ 9 ตัวแปรในการวิจัยเรื่องการวิเคราะห์ปัจจัยที่มีผลต่อความพิการจากอุบัติเหตุจราจรทางบก

การทบทวนวรรณกรรม/สารสนเทศ (Information) ที่เกี่ยวข้อง

ในปัจจุบันองค์กรต่าง ๆ ทางด้านสาธารณสุขพยายามทำการศึกษาค้นคว้าวิจัยเพื่อลดความเสียหายอันเนื่องมาจากความพิการที่เกิดขึ้นหลังจากการเกิดอุบัติเหตุจราจรทางบก รวมไปถึงการกำหนดมาตรการต่าง ๆ เพื่อให้เกิดความปลอดภัยจากการจราจรทางบกมากขึ้น งานวิจัยส่วนใหญ่นำเสนอเกี่ยวกับปัจจัยที่มีผลต่อการเกิดอุบัติเหตุ การบาดเจ็บ และการเสียชีวิตจากอุบัติเหตุแต่ละครั้ง ซึ่งข้อมูลส่วนใหญ่ที่เก็บรวบรวมมามักจะมีขนาดใหญ่ จึงมีการประยุกต์ใช้วิธีวิเคราะห์ทางสถิติและเทคนิคเหมืองข้อมูลเข้ามาช่วยในการค้นหารูปแบบความสัมพันธ์ระหว่างปัจจัยต่าง ๆ จากข้อมูลที่เก็บรวบรวมมา วัตถุประสงค์ของการใช้เทคนิคเหมืองข้อมูลส่วนใหญ่จะเป็นลักษณะการจัดกลุ่มและการพยากรณ์ ซึ่งมีงานวิจัยหลากหลายประเภทที่ได้ทำการศึกษาและวิเคราะห์ความรุนแรงและการบาดเจ็บที่เกิดจากอุบัติเหตุจราจรทางบกงานวิจัยเหล่านี้มีวัตถุประสงค์เพื่อค้นหาและระบุปัจจัยที่มีผลกระทบต่อการบาดเจ็บที่เกิดขึ้น ดังจะเห็นจากงานวิจัยหลายชิ้นที่ศึกษาพฤติกรรมของผู้ขับขี่ว่ามีผลต่อความเสี่ยงต่อการเกิดอุบัติเหตุและการเสียชีวิตจากอุบัติเหตุ เช่น งานวิจัยของ Zajac and Ivan (2003) และ Keall et al. (2004) ได้รายงานไว้ว่าการดื่มแอลกอฮอล์มีผลในการเพิ่มความเสี่ยงในการเกิดอุบัติเหตุอย่างมีนัยสำคัญทางสถิติ นอกจากนี้ยังมีงานวิจัยอีกหลายชิ้นที่ทำการศึกษเกี่ยวกับประสิทธิภาพของอุปกรณ์ที่ใช้เพื่อความปลอดภัยขณะขับขี่ เช่น การใช้เข็มขัดนิรภัย หรือ

หมวกนิรภัย ที่มีต่อความรุนแรงของอาการบาดเจ็บที่เกิดจากอุบัติเหตุ (Bedard et al., 2002; Valent et al., 2002) ผลที่ได้จากการวิจัยชี้ให้เห็นว่าอุปกรณ์ต่าง ๆ เหล่านี้ช่วยลดระดับของอาการบาดเจ็บให้ลดลงอย่างมีนัยสำคัญทางสถิติ นอกจากนี้รายงานวิจัยอีกด้านหนึ่งทำการศึกษาเพื่อจัดกลุ่มของผู้ขับขี่ที่มีความเสี่ยงสูงในการเกิดอุบัติเหตุ หรือความรุนแรงของการบาดเจ็บ และการเสียชีวิตจากอุบัติเหตุ (Zhang et al., 2000; Valent et al., 2002) และมีงานวิจัยบางส่วนที่ทำการศึกษเกี่ยวกับอายุของผู้ขับขี่ พบว่าผู้ขับขี่สูงอายุ (มากกว่า 65 ปี) มีโอกาสที่จะเกิดความรุนแรงของการบาดเจ็บมากกว่ากลุ่มคนที่มีอายุน้อยกว่า เช่น ในงานวิจัยของ Zhang et al. (2000) ได้ทำการสำรวจพบว่าผู้ขับขี่ที่มีอายุตั้งแต่ 65 ปีขึ้นไปมีความเสี่ยงต่อการตายจากอุบัติเหตุสูงกว่าคนทั่วไป งานวิจัยบางชิ้นได้ทำการศึกษเกี่ยวกับความเสี่ยงและระดับอาการบาดเจ็บของผู้ใช้ถนนที่จะได้รับอันตรายจากอุบัติเหตุที่เกิดขึ้น (Zajac and Ivan, 2003)

ในส่วนของการประยุกต์ทฤษฎีทางสถิติและเทคนิคเหมืองข้อมูลเข้ามามีใช้ในการวิเคราะห์ข้อมูลเกี่ยวกับปัจจัยต่าง ๆ และผลกระทบที่เกิดจากอุบัติเหตุ นั้น ก็มีความหลากหลายเช่นกัน ในระยะแรก ๆ นั้น จะเริ่มจากทฤษฎีที่ง่ายและไม่ซับซ้อนมากนัก เช่น การใช้การทดสอบแบบที (t-test) เพื่อทดสอบความแตกต่างของค่าเฉลี่ย 2 กลุ่มใด ๆ ตลอดจนวิธีการทางสถิติที่มีความซับซ้อนมากขึ้น เช่นการวิเคราะห์การถดถอยเชิงเส้น (Linear regression analysis) นอกจากนี้ยังมีการใช้การทดสอบแบบไคสแควร์ (Chi-square test) เข้ามาช่วยในการทดสอบความสัมพันธ์ระหว่าง ระดับความรุนแรงของการบาดเจ็บกับเพศและอายุของผู้ขับขี่ (Ostrom and Eriksson, 2001) สำหรับปัญหาที่มีความซับซ้อนมากขึ้น หรือประกอบด้วยจำนวนตัวแปรมากขึ้น ผู้วิจัยมีการประยุกต์เทคนิคการวิเคราะห์ทางสถิติขั้นสูงขึ้นไปเข้ามามีใช้ เช่น Bedard et al. (2002) ประยุกต์ใช้การวิเคราะห์การถดถอยแบบโลจิสติกส์หลายตัวแปร (Multivariate logistic regression analysis) เข้ามาใช้ในการศึกษาอิทธิพลของตัวแปรต่าง ๆ เกี่ยวกับผู้พฤติกรรมของผู้ขับขี่ ลักษณะของยานยนต์ ที่มีผลต่อการเสียชีวิตจากอุบัติเหตุ ซึ่งผลที่ได้จากการวิจัยชี้ให้เห็นว่า ผู้ขับขี่สูงอายุ ผู้ขับขี่ที่เป็นเพศหญิง รวมไปถึงผู้ขับขี่ที่ดื่มแอลกอฮอล์ มีโอกาสสูงขึ้นไปที่จะเสียชีวิตจากอุบัติเหตุ Valent et al. (2002) ประยุกต์ใช้การถดถอยโลจิสติกส์ เข้ามาใช้ในการประเมินความสัมพันธ์ของลักษณะผู้ขับขี่และความรุนแรงของการบาดเจ็บจากอุบัติเหตุ โดยทำการประมาณตัวแบบแยกออกเป็นหลายตัวแบบตามประเภทของยานยนต์ เช่น รถยนต์ 4 ที่นั่ง รถโดยสาร รถบรรทุก และรถจักรยานยนต์ ตามลำดับ ต่อมา Yau (2004) ประยุกต์ใช้เทคนิคการวิเคราะห์การถดถอยแบบโลจิสติกส์เชิงขั้นตอน (Stepwise logistic regression analysis) ในการศึกษาอิทธิพลของภูมิสำเนา ธรรมชาติของผู้คนที่อยู่อาศัย ประเภทของยานยนต์ และปัจจัยทางสภาวะแวดล้อม ที่มีผลต่อความบาดเจ็บหลังการเกิดอุบัติเหตุ

นอกจากนี้ยังมีงานวิจัยที่มีลักษณะคล้ายคลึงกันที่เกี่ยวข้องกับระบบความปลอดภัย ซึ่งมีการวิจัยหลายชิ้นที่ทำการประยุกต์ใช้เทคนิคเหมืองข้อมูล เช่น ต้นไม้ช่วยการตัดสินใจ (Decision tree) เพื่อวิเคราะห์อัตราการเกิดอุบัติเหตุและระดับการบาดเจ็บที่เกิดจากอุบัติเหตุ (Kuhnert et al., 2000; Sohn and Shin, 2001) ในงานวิจัยของ Kuhnert et al. (2000) ได้ทำการเปรียบเทียบวิธีเหมืองข้อมูลที่แตกต่างกัน 3 วิธี ประกอบด้วย Logistic regression วิธี Classification and regression tree (CART) และวิธี Multivariate adaptive regression splines (MARS) เพื่อวิเคราะห์ข้อมูลเกี่ยวกับความบาดเจ็บที่เกิดจากอุบัติเหตุจราจรทางบก และผลจากการเปรียบเทียบพบว่าวิธี CART และ MARS สามารถสำรวจรูปแบบความสัมพันธ์ของตัวแปรได้ดีกว่าวิธี Logistic regression ต่อมา Sohn and Shin (2001) ทำการ

ประยุกต์ใช้ Classification tree และ Neural network และวิธี Logistic regression เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อการบาดเจ็บที่เกิดจากอุบัติเหตุ ผลการวิจัยระบุว่าอุปกรณ์ความปลอดภัยต่าง ๆ เช่น เข็มขัดนิรภัย และหมวกนิรภัย เป็นปัจจัยที่สำคัญที่สุดที่ลดอาการบาดเจ็บจากอุบัติเหตุที่เกิดขึ้น ทั้งนี้ตัวแปรอื่น ๆ ที่สำคัญเช่นเดียวกันประกอบด้วย ความเร็วของยานยนต์ก่อนเกิดอุบัติเหตุ การขับขี่ที่ผิดกฎจราจร ต่างมีผลต่อความรุนแรงของการบาดเจ็บเช่นกัน จะเห็นได้ว่างานวิจัยในลักษณะคล้ายคลึงกันถูกรายงานไว้ใน Chang and Yeh (2007) ที่ทำการศึกษาความสัมพันธ์ระหว่างตัวแปร เพศ อายุ และพฤติกรรมเสี่ยงในการขับขี่รถจักรยานยนต์ว่ามีผลต่อการเกิดอุบัติเหตุมากเพียงใดโดยใช้เทคนิคการวิเคราะห์ Cluster analysis เพื่อจัดกลุ่มของผู้ขับขี่จักรยานยนต์ออกเป็นกลุ่มที่แตกต่างกันตามระดับความเสี่ยงต่อการเกิดอุบัติเหตุ ผลที่ได้จากการวิจัยสรุปได้ว่าผู้ขับขี่เพศชายที่มีอายุน้อย ๆ มีอัตราการทำผิดกฎจราจรค่อนข้างสูงและมีความเสี่ยงในการเกิดอุบัติเหตุสูงเช่นเดียวกัน

วิธีการโครงข่ายประสาทเทียม (Artificial neural network: ANN) ก็เป็นอีกหนึ่งวิธีที่ได้รับความนิยมและถูกนำมาประยุกต์ใช้สำหรับการพยากรณ์ข้อมูลเกี่ยวกับการเกิดอุบัติเหตุจราจรทางบกรวมไปถึงงานด้านอื่น ๆ เช่นในงานวิจัยของ Abdel-Aty and Abdelwahab (2004) ประยุกต์ใช้ ANN ในการวิเคราะห์การบาดเจ็บของผู้ขับขี่จากอุบัติเหตุ ผลที่ได้จากการวิจัยพบว่า เพศของผู้ขับขี่ ความเร็วขณะขับขี่ก่อนเกิดอุบัติเหตุ การใช้เข็มขัดนิรภัย และประเภทของรถยนต์มีผลต่อการบาดเจ็บของผู้ขับขี่ นอกจากนี้ยังมีการเปรียบเทียบความแม่นยำในการพยากรณ์ของ ANN และตัวแบบทางสถิติแบบอื่น ๆ เช่น Regression analysis พบว่า ANN มีความแม่นยำมากกว่า นอกจากนี้ยังมีงานวิจัยของ Chang and Wang (2006) ซึ่งใช้เทคนิค CART ในการวิเคราะห์ข้อมูลการเกิดอุบัติเหตุในกรุงเทพฯ ประเทศไทย ซึ่งมีการเก็บรวบรวมข้อมูลไว้ในปี ค.ศ. 2001 ซึ่งผลที่ได้พบว่าตัวแปรที่มีผลกระทบต่อความรุนแรงของการบาดเจ็บมากที่สุดคือประเภทของยานยนต์ที่ใช้ นอกจากนี้ยังพบว่า ผู้ใช้ถนน (คนเดินเท้า) รถจักรยานยนต์ และ คนขี่จักรยาน เป็นผู้ที่มีความเสี่ยงสูงที่จะได้รับบาดเจ็บมากกว่ายานยนต์ประเภทอื่น ๆ Chang (2005) ทำการศึกษาเปรียบเทียบประสิทธิภาพของ Negative binomial regression และ Artificial neural network สำหรับการวิเคราะห์ข้อมูลที่เก็บรวบรวมข้อมูลจากทางหลวงในประเทศไทย ระหว่างปี 1997-1998 และสรุปผลว่า ANN ให้ผลที่ดีกว่าทุกกรณีของการทดสอบ นอกจากนี้ Delen (2006) ทำการประยุกต์ใช้ ANN เพื่อพยากรณ์ระดับความรุนแรงของการพิการหลังได้รับอุบัติเหตุโดยศึกษาจากกลุ่มตัวแปรต่าง ๆ เช่น พฤติกรรมการขับขี่ สภาพยานพาหนะ และ สภาพถนน เป็นต้น และผลสรุปพบว่าระดับความรุนแรงของความพิการจะเปลี่ยนแปลงตามระดับค่าต่าง ๆ ของตัวแปรที่เกี่ยวข้อง

ตามที่ได้สรุปไว้ข้างต้น จะเห็นได้ว่าเทคนิคโครงข่ายประสาทเทียมถูกนำมาใช้อย่างแพร่หลาย และค่อนข้างจะประสบผลสำเร็จในการวิเคราะห์ข้อมูลทางด้านอุบัติเหตุจราจร ไม่ว่าจะเป็นเทคนิคสำหรับจำแนกกลุ่มหรือการพยากรณ์ ในงานวิจัยนี้ผู้วิจัยมีวัตถุประสงค์ที่จะประยุกต์โครงข่ายประสาทเทียมแบบ multi layer perceptron เข้ามาใช้ในการวิเคราะห์ข้อมูลที่เก็บรวบรวมไว้โดยศูนย์สิรินธรเพื่อการฟื้นฟูสมรรถภาพทางการแพทย์ ดังที่ได้รายงานไว้ในผลงานวิจัยเรื่อง “อุบัติเหตุการความพิการและผลกระทบที่เกิดขึ้นจากสาเหตุอุบัติเหตุจราจรทางบก พ.ศ. 2549” เพื่อทำให้สามารถระบุตัวแปรที่มีความสัมพันธ์ต่อระดับความรุนแรงของการบาดเจ็บ และสามารถแบ่งออกเป็นระดับย่อยที่แตกต่างกันได้ และนำผลสรุปที่ได้เปรียบเทียบกับตัวแบบทางสถิติที่ได้รับความนิยมอย่างแพร่หลายคือ Negative binomial model ทั้งนี้เพื่อเป็นแนวทาง

สำหรับหน่วยงานที่เกี่ยวข้องในการนำผลสรุปที่ได้ไปกำหนดนโยบายเพื่อลดโอกาสในการเกิดอุบัติเหตุรวมถึงลดความสูญเสียที่เกิดขึ้นหลังจากอุบัติเหตุจราจรทางบก

บทที่ 3

ผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ปัจจัยที่มีผลต่อการเกิดอุบัติเหตุจราจรทางบก โดยจากการศึกษาตัวแปรต่างๆ เพื่อใช้สร้างตัวแบบพยากรณ์ด้วยเน้นใช้เทคนิคเหมืองข้อมูล โดยได้แต่ประเด็นการสร้างตัวแบบ เป็น 4 รูปแบบขึ้นอยู่กับจำนวนตัวแปรนำเข้าที่ใช้และระดับข้อมูลออก

ตัวแบบพยากรณ์ที่ 1

โดยการใช้โครงข่ายประสาทเทียมและคัดเลือกตัวแปรเข้าด้วยการวิเคราะห์ค่าทางสถิติ ซึ่งได้ตัวแปรมาดังตารางที่ 6 และพิจารณาเพียงกรณีที่เกิดการพิการเท่านั้น ดังนั้นข้อมูลจะประกอบด้วยผู้พิการทั้งหมด 423 กรณี โดยในนี้มี 297 กรณีที่เกิดความพิการแบบเข้าเกณฑ์ และ 126 เกิดความพิการแบบไม่เข้าเกณฑ์

ตารางที่ 6 ผลการวิเคราะห์ตัวแปรด้วยวิธีการทางสถิติ

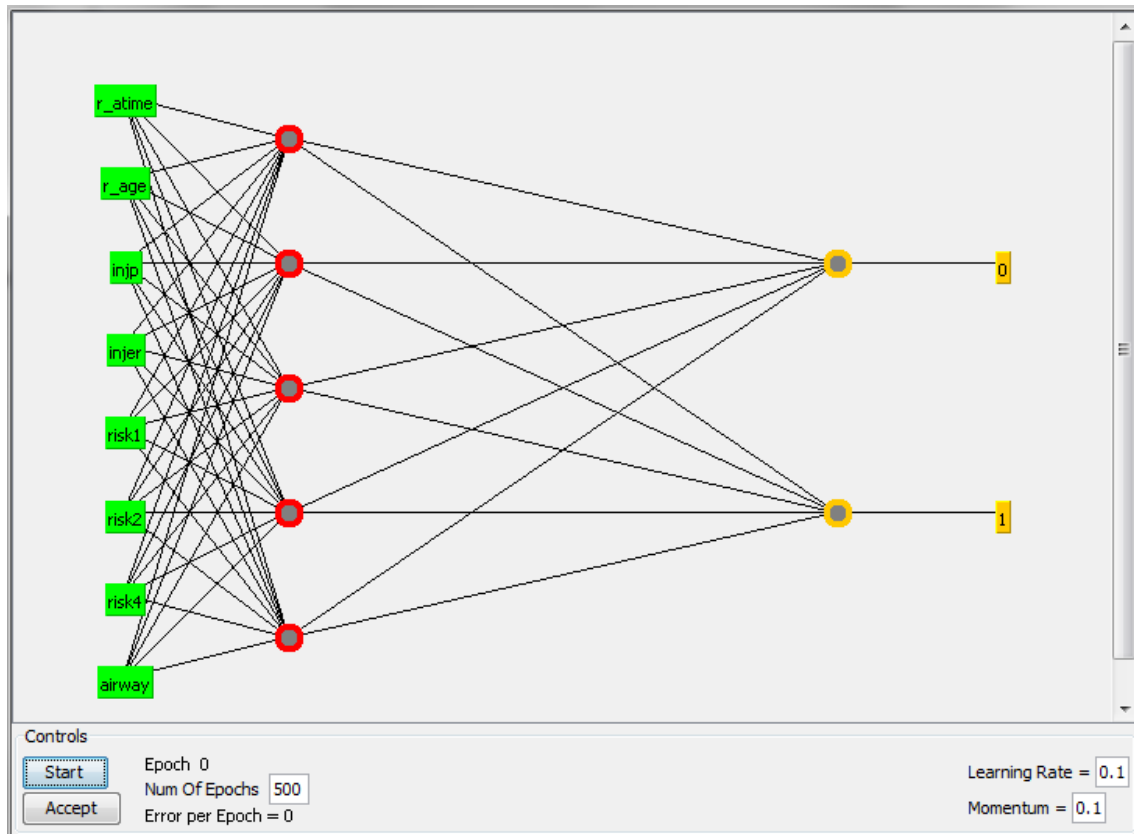
กลุ่มตัวแปร	ชื่อ/ลักษณะตัวแปร	r	p-value
ข้อมูลทั่วไปของผู้ขับขี่	age(อายุ)	-0.272	<0.001
	alcohol(สุรา)	0.112	0.021
	drug(ยาเสพติด)	0.144	0.003
	belt(เข็มขัดนิรภัย)	0.127	0.010
	position(ตำแหน่งของผู้บาดเจ็บ)	0.122	0.012
	cause(สาเหตุของการบาดเจ็บ)	0.170	<0.001
สภาวะแวดล้อมของที่เกิดเหตุ	time(เวลาในการเกิดเหตุ)	0.464	<0.001
การปฐมพยาบาลเบื้องต้น	airway(ลมหายใจ)	0.291	<0.001

โดยได้นำค่าเหล่านี้มาสร้างตัวพยากรณ์ด้วยวิธีโครงข่ายประสาทเทียมแบบ Multi-Layer perceptron ด้วยกลไก back-propagation ด้วยโปรแกรม Weka โดยใช้ค่า learning rate และ momentum เป็น 0.1 โดยเป็นโครงข่ายเป็นแบบ 1 ชั้นซ่อน ที่ประกอบด้วย 5 โหนดซ่อน โดยมี 8 ตัวแปรเข้า และ 2 ตัวแปรออก โดยค่าตัวแปรออก 2 ค่ามีความหมายดังนี้ คือ

0 หมายถึงเกิดการพิการแบบเข้าเกณฑ์

1 หมายถึงเกิดการพิการแบบไม่เข้าเกณฑ์

โดยจะได้ผลลัพธ์เป็นโครงข่ายดังภาพที่ 4

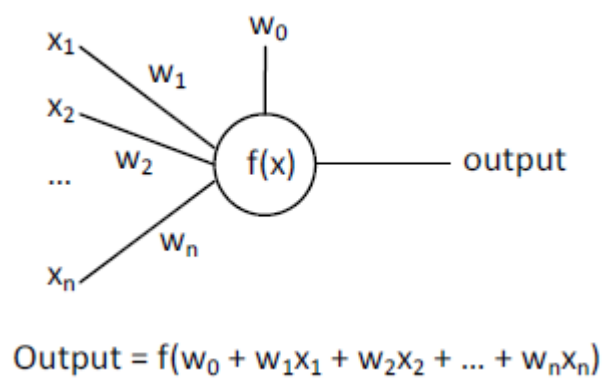


ภาพที่ 10 แสดงผลลัพธ์ที่ได้จากการสร้างโมเดลโครงข่ายประสาทเทียมด้วย Weka

โดยจะได้ค่าน้ำหนักของโหนดต่างดังเช่น

$$\text{node2} = \text{sigmoid}(-24.9 + 27.7 \cdot \text{atime} + (-17.8) \cdot \text{age} + (-25.5) \cdot \text{position} + 4.9 \cdot \text{cause} + 4.3 \cdot \text{drug} + 3.7 \cdot \text{belt} + 11.5 \cdot \text{airway})$$

โดย sigmoid เป็นฟังก์ชัน $f(x)$ ที่ใช้ในโครงข่ายประสาทเทียม ดังแสดงดังภาพที่ 11



ภาพที่ 11 แสดงกลไกการทำงานของโครงข่ายประสาทเทียม

ผลการประเมินตัวพยากรณ์

ตัวพยากรณ์ที่ได้มีความแม่นยำโดยรวมที่ 73.34% โดยมีตาราง Confusion matrix ดังต่อไปนี้

ค่าจริง \ ค่าพยากรณ์	0	1
0	230	43
1	60	54

จากตารางจะเห็นว่า พยากรณ์ 0 ผิดไป $43/273 = 15.8\%$

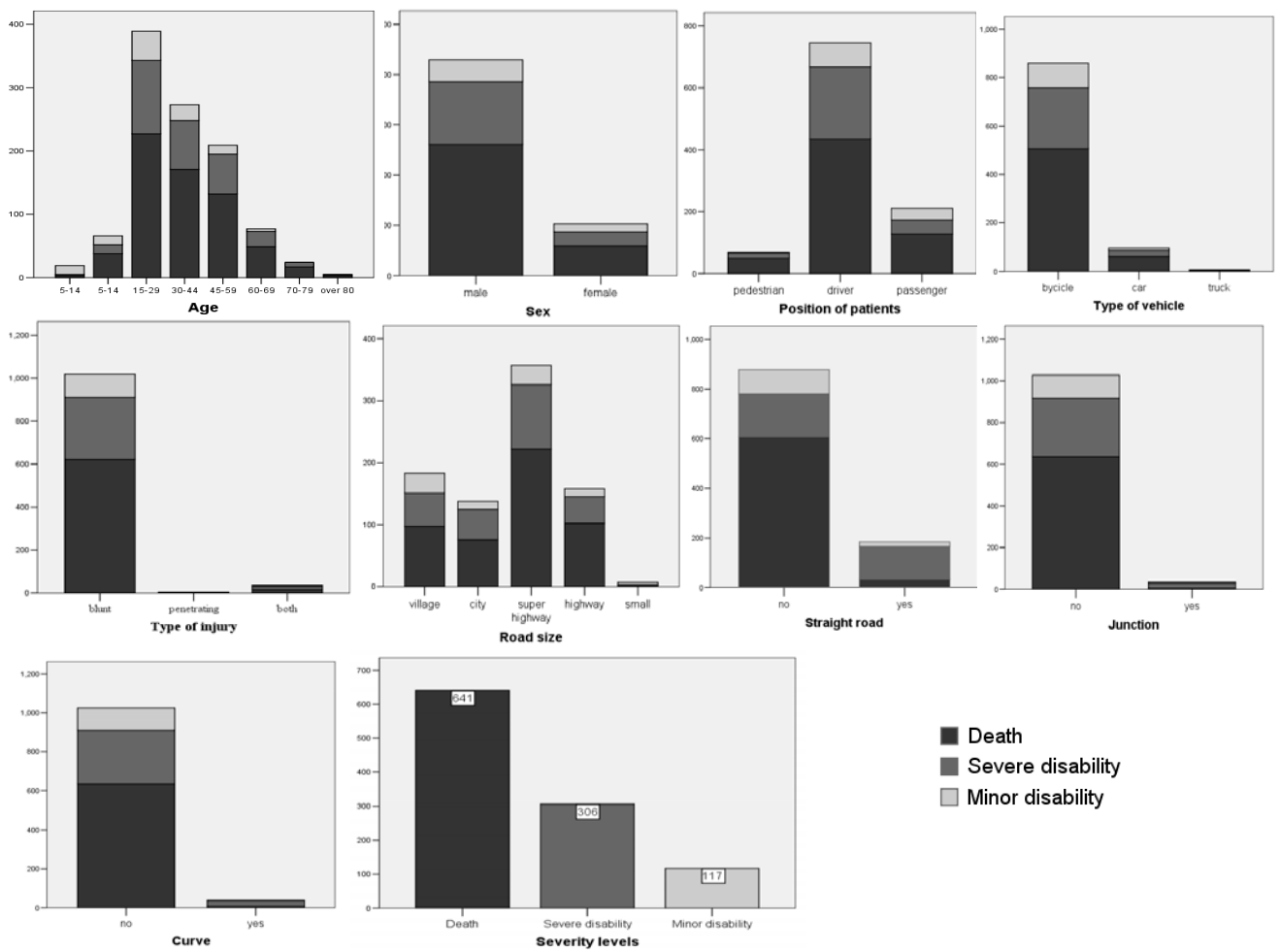
พยากรณ์ 1 ผิดไป $60/114 = 52.6\%$

จากผลการวิจัยจะเห็นว่าตัวพยากรณ์ที่ได้ในภาพรวมค่อนข้างดี แต่เมื่อมาพิจารณาในรายละเอียดจะเห็นว่าให้การทำนายที่แย่มากในกรณีที่คำตอบมีค่าเป็น 1 (กรณีพิการแบบไม่เข้าเกณฑ์)

ตัวแบบพยากรณ์ที่ 2

โดยการใช้โครงข่ายประสาทเทียมและคัดเลือกตัวแปรเข้าด้วยการวิเคราะห์ค่าทางสถิติ ซึ่งได้ตัวแปรมาดังตารางที่ 7 โดยพิจารณาเพิ่มกรณีที่เกิดการเสียชีวิตเข้าไป เพื่อให้ข้อมูลมีความสมดุลมากขึ้น ดังนั้นข้อมูลจะประกอบด้วยผู้พิการทั้งหมด 1,064 ระเบียบ โดยจะมี กรณีสิ้นชีวิต 641 ระเบียบ กรณีที่เกิดความพิการแบบเข้าเกณฑ์ 306 ระเบียบ และความพิการแบบไม่เข้าเกณฑ์ 117 ระเบียบ โดยเมื่อจำแนกตามปัจจัยจะได้การกระจายดังแสดงในภาพ 12 โดยจะเห็นได้จากแผนภูมิว่า

- ช่วงอายุ 15-29 ปีจะมีการเกิดอุบัติเหตุมากที่สุด และเพศชายเกิดอุบัติเหตุมากกว่าเพศหญิง
- ตำแหน่งคนขับได้รับบาดเจ็บมากที่สุด และพาหนะชนิด 2 ล้อเกิดอุบัติเหตุบ่อยสุด
- ถนนที่เกิดอุบัติเหตุมากที่สุดคือ super highway



ภาพที่ 12 แสดงการกระจายของความรุนแรงของการบาดเจ็บแยกตามปัจจัย

จากนั้นนำตัวแปรที่คัดเลือกมาเข้าสถิติเพื่อหาความสัมพันธ์ จะได้ตารางที่ 7

ตารางที่ 7 แสดงค่าความสัมพันธ์ทางสถิติของตัวแปรเข้า กับตัวแปรตาม

Variable Group	Variable name	r	P-value
Personal	Age (<i>age</i>)	-0.095	<0.001
	Sex (<i>sex</i>)	0.035	0.240
	Position of patient (<i>injp</i>)	0.050	0.089
	Type of vehicle (<i>injt</i>)	-0.038	0.215
Injury	Type of injury (<i>tinj</i>)	0.064	0.030
Situation	Road size (<i>a31</i>)	-0.074	0.015
	Straight road (<i>a421</i>)	0.323	<0.001
	Junction (<i>a43</i>)	0.164	<0.001
	Curved road (<i>a45</i>)	0.131	<0.001

โดยได้มีการกำหนดช่วงตัวแปรแต่ละตัวให้เป็นช่วงข้อมูลดังต่อไปนี้

- Age (0-4, 5-14, 15-29, 30-44, 45-59, 60-69, 70-79, over 80) แทนด้วยเลข {1, 2, 3,...,8}
- Sex ('male', 'female') แทนด้วยเลข {1, 2}
- Position of patient ('pedestrian', 'driver', 'passenger') แทนด้วยเลข {1, 2, 3}
- Type of vehicle ('bicycle/motorbike/three-wheel', 'car/pickup', 'truck/lorry/bus/minibus') แทนด้วยเลข {1, 2, 3}
- Type of injury ('blunt', 'penetrating', 'both') แทนด้วยเลข {1, 2, 3}
- Road size ('village road', 'city road', 'super highway', 'highway', 'small road') แทนด้วยเลข {1, 2, 3, 4, 5}
- Straight road ('no', 'yes') แทนด้วยเลข {0, 1}
- Junction ('no', 'yes') แทนด้วยเลข {0, 1}
- Curved road ('no', 'yes') แทนด้วยเลข {0, 1}

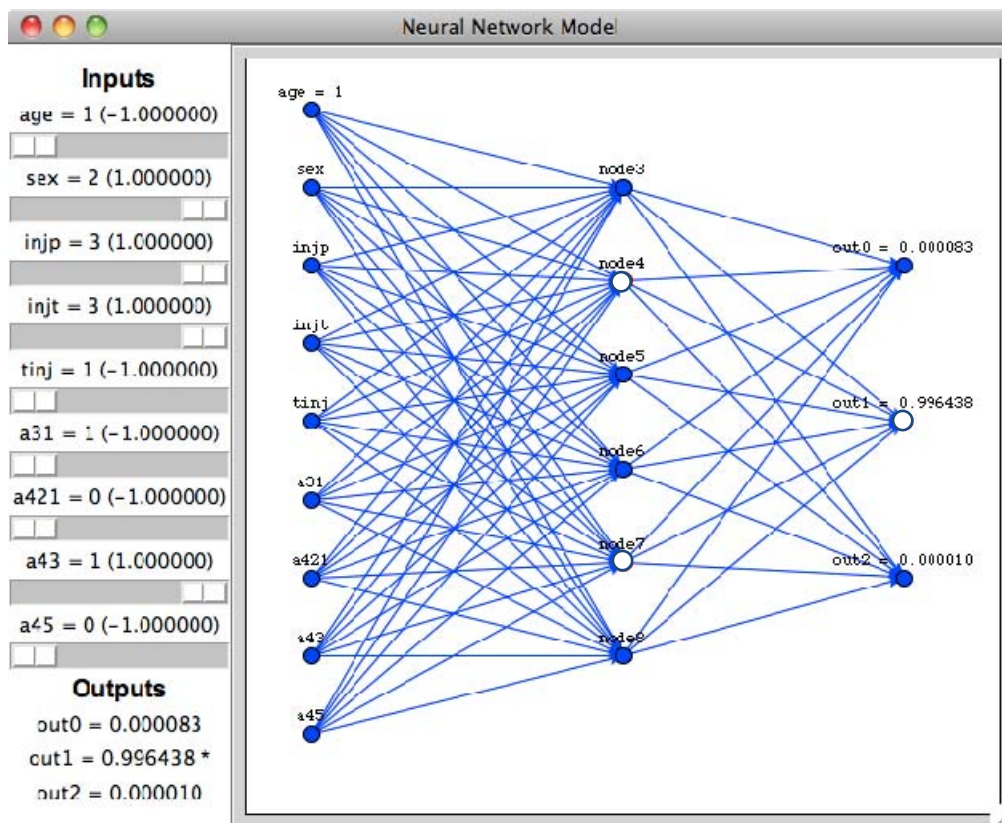
โดยคำตอบของการพยากรณ์จะเป็น 3 ระดับ คือ

1. Death
2. Disability inclusion criteria (Severe disability)
3. Disability non-inclusion criteria (Minor disability)

โดยเมื่อนำมาสร้างตัวแบบพยากรณ์ด้วยเทคนิคเหมืองข้อมูลแบบต่างๆ จะให้ค่าความแม่นยำดังต่อไปนี้

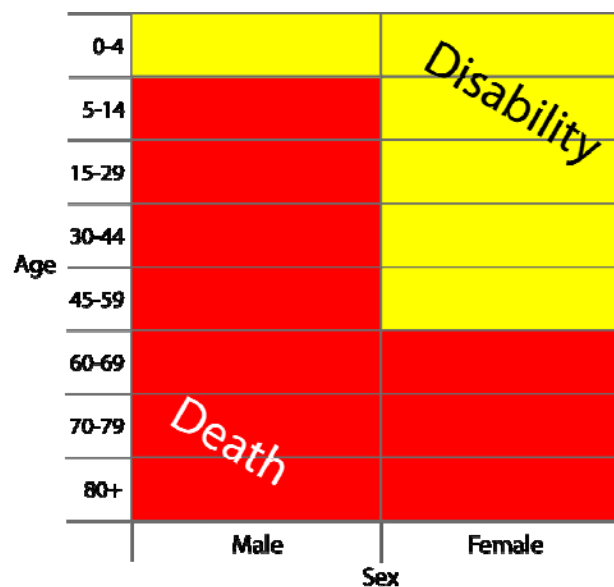
Classification technique:	Naïve Bayes	CART	Rand Tree	RBF	SMO (SVM)	MLP
Accuracy (Cross-validation)	75%	75%	71%	73%	75%	73%
Accuracy (full training)	74%	75%	80%	73%	75%	77%

โดยในที่นี้ได้เลือกเอาแบบโครงข่ายประสาทเทียมมาสร้างโปรแกรมพยากรณ์ดังแสดงในภาพที่ 13

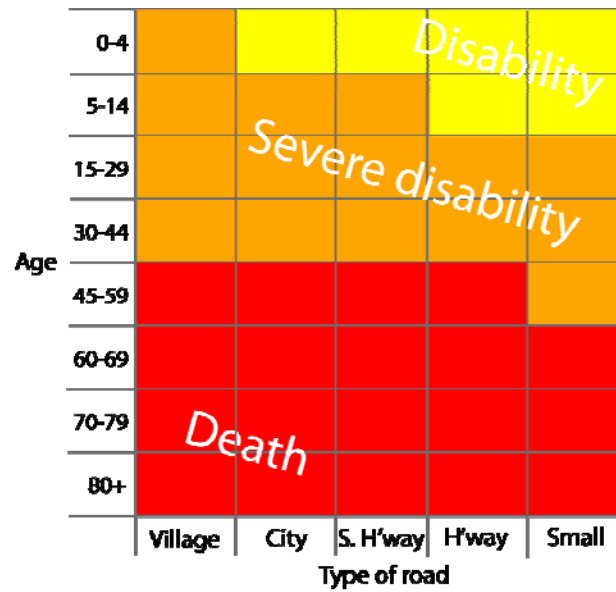


ภาพที่ 13 The neural network model in EDEN

และเมื่อวิเคราะห์จากโมเดลจะได้ว่า ความสัมพันธ์ระหว่างอายุ เพศ และความรุนแรงดังภาพที่ 14

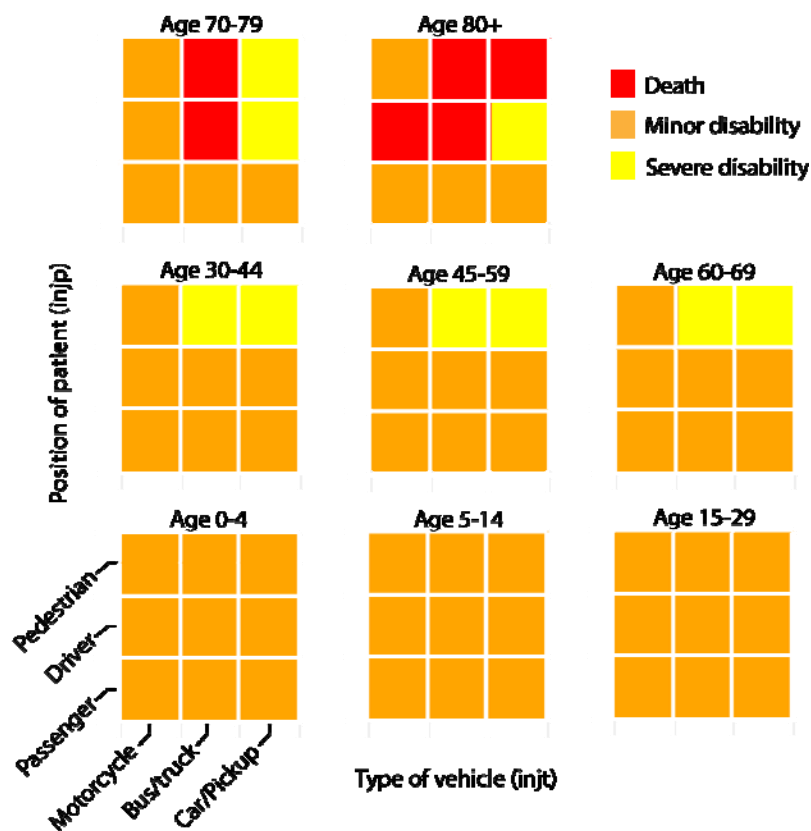


ภาพที่ 14 แสดงความสัมพันธ์ระหว่างอายุ เพศ และความรุนแรง



ภาพที่ 15 แสดงความสัมพันธ์ระหว่างอายุ ลักษณะถนน และความรุนแรง

และความสัมพันธ์ระหว่างอายุ ลักษณะถนน และความรุนแรงได้ดังภาพที่ 15 และจำแนกตามอายุ ชนิดของยานพาหนะ และความรุนแรงได้ดังภาพที่ 16

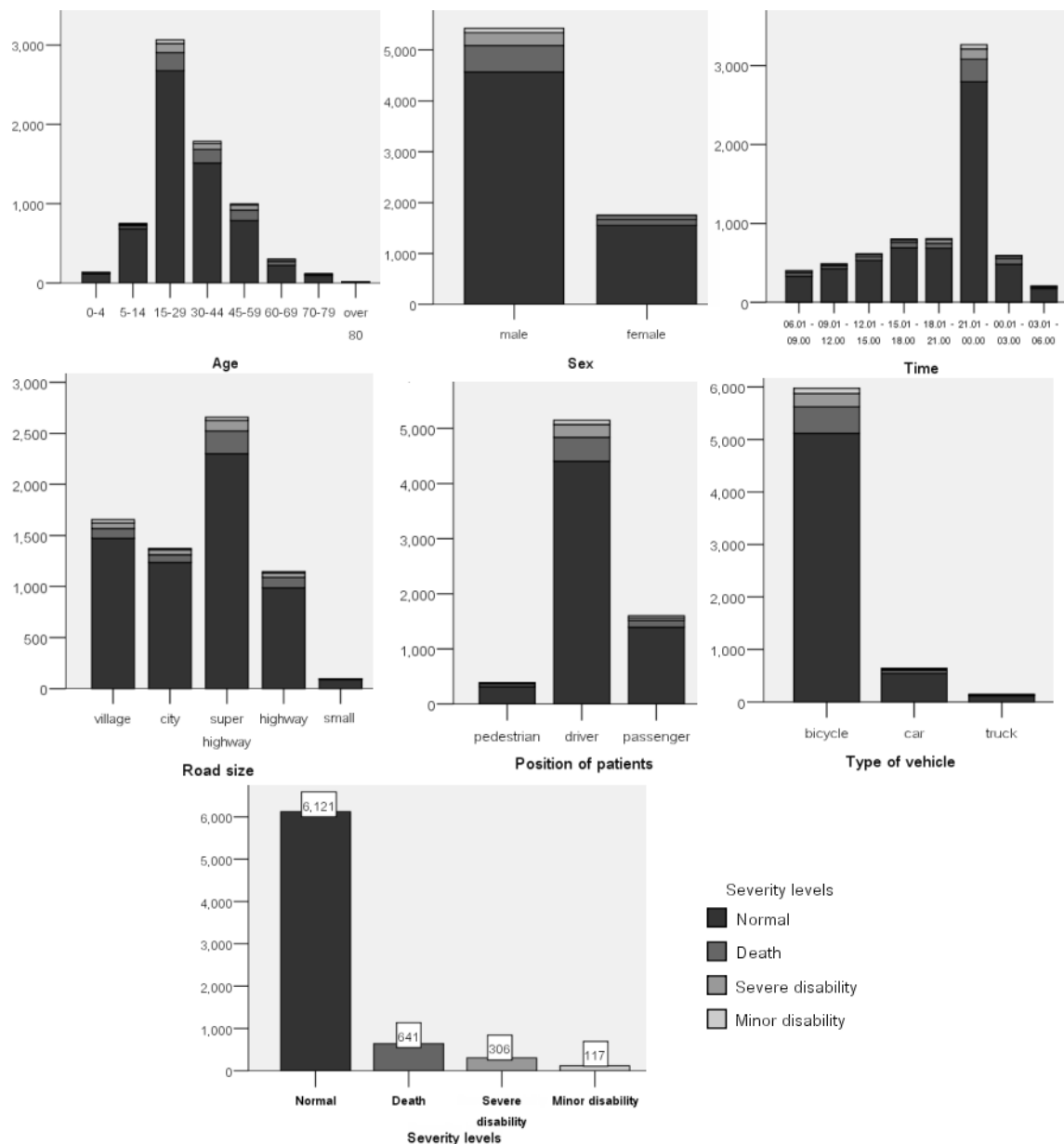


ภาพที่ 16 แสดงความสัมพันธ์ระหว่างอายุ ชนิดของยานพาหนะ และความรุนแรง

ซึ่งจากตัวแบบพยากรณ์ที่ 2 จะได้โปรแกรมเพื่อพยากรณ์ความรุนแรงของการเกิดอุบัติเหตุด้วย เทคนิค โคจรข่ายประสาทเทียม โดยมีความแม่นยำที่ร้อยละ 77 ทั้งนี้เพราะการลดจำนวนตัวแปรเข้ามามีผลให้หลาย ระเบียบมีค่าของตัวแปรเข้าที่เหมือนกัน แต่มีคำตอบที่แตกต่างกัน จึงทำให้ความแม่นยำที่ได้ค่อนข้างที่จะต่ำ

ตัวแบบพยากรณ์ที่ 3

เมื่อพิจารณาข้อมูลเพิ่มเติมโดยการนำข้อมูลในช่วงปกติ เข้ามาด้วยโดยข้อมูลจำแนกตามตัวแปร ต่อ ระดับความรุนแรงจะได้ดังภาพ 17



ภาพที่ 17 กราฟแท่งแสดงการจำแนกตาม 4 ระดับความรุนแรง

โดยเมื่อใช้การทดสอบทางสถิติจะเหลือตัวแปรต้นดังภาพ 18 และตัวแปรตามมี 4 ระดับ

Attribute	
age	0 - 4, 5 - 14, 15 - 29, 30 - 44, 45 - 59, 60 - 69, 70 - 79, and over 80
sex	'male' and 'female'
position of patient	'pedestrian', 'driver', or 'passenger'
type of vehicle	'bicycle/motorbike/three-wheel vehicle', 'car/pickup', or 'truck/lorry/bus/minibus'
type of injured	'Blunt', 'Penetrating', or 'Blunt & Penetrating'
road size	'village road', 'city road', 'super highway', 'highway', or 'small road'
straight road	'no' or 'yes'
junction	'no' or 'yes'
curved road	'no' or 'yes'
predicted	'normal', 'death', 'severe', or 'minor disability'

ภาพที่ 18 แสดงตัวแปรต้นที่ใช้ในการพัฒนาตัวแบบพยากรณ์

โดยที่เลือกใช้เทคนิค 3 เทคนิค โดยแต่ละเทคนิคให้ค่าความแม่นยำดังต่อไปนี้

	J48	RandomTree	Naive Bayes
Training Set	65.62%	70.95%	65.00%
Test Set	74.38%	72.21%	77.48%
Data Set	67.46%	71.15%	67.62%

และเมื่อนำไปพัฒนาโปรแกรมเพื่อใช้งานจะได้ดังภาพที่ 19

10.31.19.22/project/poster/main.php

ข้อมูลการบาดเจ็บจากอุบัติเหตุ

อายุ : 45 - 59
เพศ : ☒ ชาย ☐ หญิง
ผู้บาดเจ็บเป็น : คนขับขี
พาหนะของผู้บาดเจ็บ : จักรยานและสามล้อ
ลักษณะการบาดเจ็บ : Blunt

ข้อมูลที่เกิดเหตุ (ลักษณะเส้นทาง)
ถนนที่เกิดเหตุ : ถนนกรมทางหลวง
ทางตรง : ☐ ใช่ ☒ ไม่ใช่
ทางแยก : ☒ ใช่ ☐ ไม่ใช่
ทางโค้ง : ☐ ใช่ ☒ ไม่ใช่
ทำนายระดับความรุนแรง

ผลการทำนาย

Decision Tree

J48 : พิจารณาเข้าเกณฑ์

RandomTree : พิจารณาเข้าเกณฑ์

Bayes Classifier

Naive Bayes : พิจารณาเข้าเกณฑ์

Prediction

Prediction : พิจารณาเข้าเกณฑ์

Delete Value

ภาพที่ 19 โปรแกรมการพยากรณ์ความรุนแรงจากอุบัติเหตุการจราจรทางบก

เพราะคำตอบที่ระดับปกติมีจำนวนมากเกินไป จึงทำให้ความแม่นยำที่ได้มีค่าค่อนข้างต่ำ งานวิจัยนี้จึงได้ปรับข้อมูลใหม่เป็นรูปแบบพยากรณ์ที่ 4

ตัวแบบพยากรณ์ที่ 4

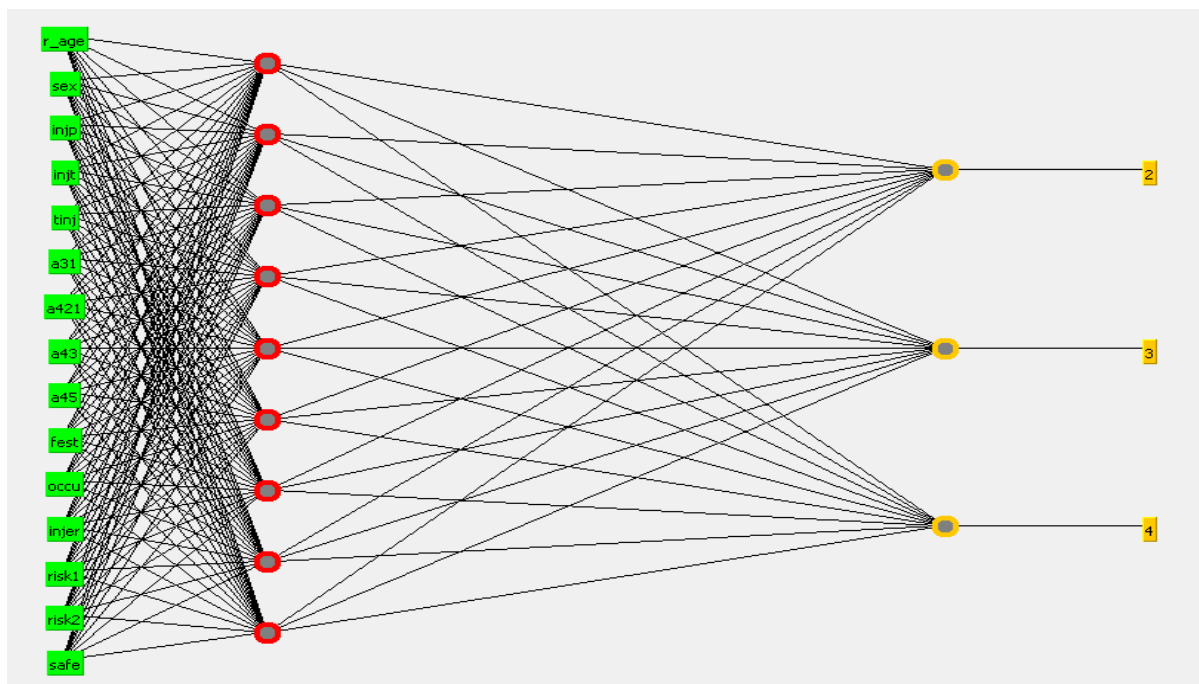
ตัวแบบพยากรณ์ที่ 4 มีการออกแบบใหม่โดยปรับปรุงจากตัวแบบที่ 1 และ 2 โดยได้เพิ่มตัวแปรเข้ามาโดยใช้สถิติ Pearson Chi-Square ได้ตัวแปรมาดังต่อไปนี้

Variable Group	Variable Name	Pearson Chi-Square	p-value
Personal	Age {0-4, 5-14, 15-29, 30-44, 45-59, 60-69, 70-79, over 80}	188.79	<.001
	Gender {male, female}	21.25	<.001
	Occupation {public servant, police, state enterprise, employee, labor, dealer, farmer, student, other}	83.16	<.001
	Alcohol {yes, no, don't know}	1899.36	<.001
	Drug {yes, no, don't know}	1213.59	<.001
	Safety (Belt, Helmet) {no, yes, don't know}	549.23	<.001
Environment	Time {6.01 – 9.00, 9.01 – 12.00, 12.01 – 15.00, 15.01 – 18.00, 18.01 – 21.00, 21.01 – 0.00, 0.01 – 3.00, 3.01 – 6.00}	22.66	.363
	Festival {no, yes}	7.94	.047
Vehicle	Vehicle Type {motorbike, car, truck, other}	145.37	<.001
	Position in Vehicle {pedestrian, driver, passenger, don't know}	285.36	<.001
Injury	Cause of injury {fall, crash, other}	66.64	<.001
	Wounded type {blunt, penetrating, blunt and penetrating, other}	76.10	<.001
Road Condition	Road size {village, city, super highway, highway, small, don't know}	1302.43	<.001
	Straight road {no, yes}	496.62	<.001
	Junction {no, yes}	113.36	<.001
	Curve road {no, yes}	80.01	<.001

โดยในงานวิจัยนี้ได้ประยุกต์ใช้ตัวแบบพยากรณ์ทั้งหมด 3 แบบ โดยจะได้ค่าความแม่นยำดังแสดงในตารางต่อไปนี้

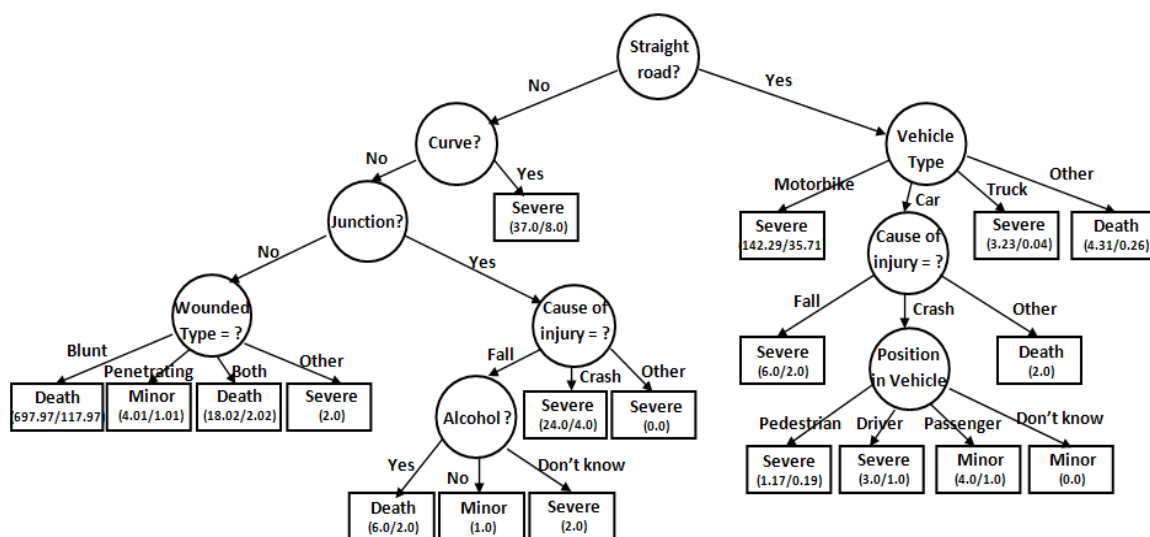
Data Technique	Training set		
	Sampling 90%	10-Folds	Modified 10-Folds
C4.5	74.08%	73.21%	74.21%
Naïve Bayes	73.63%	72.27%	74.43%
ANN L0.2/M0.1	81.01%	69.27%	81.39%
ANN L0.5/M0.2	80.31%	69.00.%	80.66%

ซึ่งจะเห็นว่าตัวแบบโครงข่ายประสาทเทียมจะให้ค่าความแม่นยำที่ดี กล่าวคือร้อยละ 80.66 โดยจะได้ตัวแบบโครงข่ายประสาทเทียมดังภาพ 20



ภาพที่ 20 แสดงตัวแบบโครงข่ายประสาทเทียมที่มี 1 ชั้นซ่อนที่มี 9 โหนดซ่อน

และตัวแบบที่ได้จากต้นไม้ตัดสินใจจะเป็นดังภาพ 21



ภาพที่ 21 แสดงตัวแบบพยากรณ์ที่ได้จากต้นไม้ตัดสินใจ

โดยจะเห็นว่าต้นไม้ตัดสินใจที่ได้จะสามารถสร้างเป็นกฎ ได้ 19 กฎ

ผลการวิจัย

จากข้อมูลทั้งหมด 9,738 ระเบียบ ซึ่งในนี้ไม่เกิดความพิการ 8,560 กรณี เสียชีวิตทั้งหมด 641 และ 297 กรณีที่เกิดความพิการแบบเข้าเกณฑ์ และ 126 เกิดความพิการแบบไม่เข้าเกณฑ์ และมีบางส่วนที่เป็นข้อมูลที่ไม่สมบูรณ์หรือขาดหายไป

โดยภาพรวมจังหวัดทางภาคตะวันออกเฉียงเหนือ (ขอนแก่น และ นครราชสีมา) จะเกิดอุบัติเหตุมากที่สุด และอุบัติเหตุจะเกิดมากในช่วงเดือนปลายธันวาคมถึงมกราคม เพราะเป็นช่วงเทศกาลสงกรานต์เก่าต้อนรับปีใหม่ โดยส่วนมากพาหนะของผู้บาดเจ็บคือ รถจักรยานยนต์ และผู้บาดเจ็บส่วนมากจะเป็นผู้ขับขี่

ตัวแบบพยากรณ์ที่ 1 แสดงให้เห็นว่าปัจจัยที่มีผลต่อความพิการจะได้แก่ อายุ สุรา ยาเสพติด เข็มขัดนิรภัย ตำแหน่งของผู้บาดเจ็บ สาเหตุของการบาดเจ็บ เวลาในการเกิดเหตุ และลมหายใจ ซึ่งเมื่อนำมาสร้างตัวพยากรณ์ด้วยโครงข่ายประสาทเทียมแบบหนึ่งชั้นซ่อนจะได้ความแม่นยำที่ 73.34% ในขณะที่ตัวแบบที่ 2 เพิ่มระดับความรุนแรง **เสียชีวิต** เข้าไปและได้ค่าความแม่นยำที่ 73% และ 77% สำหรับชุดข้อมูลฝึกสอนและข้อมูลทั้งหมด ตามลำดับ

ตัวแบบพยากรณ์ที่ 3 ได้เพิ่มระดับความรุนแรง **ปกติ** เข้าไป ทำให้ค่าความแม่นยำที่ค่อนข้างต่ำที่ 60% - 70% เพราะข้อมูลปกติมีจำนวนมากทำให้ตัวแบบเอนไปตามข้อมูลปกติ

ตัวแบบพยากรณ์ที่ 4 ปรับปรุงจากตัวแบบที่ 2 จนได้ค่าความแม่นยำเพิ่มขึ้นที่ 80.6%

โดยปัจจัยที่มีผลต่อความรุนแรงของอุบัติเหตุ ได้แก่ อายุ เพศ ตำแหน่งที่นั่ง ถนน ยานพาหนะ และถนนที่เกิดเหตุ

บทที่ 4

สรุปและข้อเสนอแนะ

จากผลการวิจัยที่ได้แสดงไปแล้วในบทที่ 3 จะเห็นได้ว่า การประยุกต์ใช้โครงข่ายประสาทเทียมแบบต่าง ๆ นั้นให้ผลสรุปดังต่อไปนี้

1. ปัจจัยที่มีผลต่อความพิการที่คัมมา เหลือเพียง 15 ตัว ดังตัวแบบพยากรณ์ที่ 4
2. ตัวพยากรณ์ด้วยโครงข่ายประสาทเทียมแบบหลายชั้นซ่อนให้ความแม่นยำสูงสุด โดยมีค่าความแม่นยำเท่ากับ 80.60% และให้ผลการทำนายที่ไม่ค่อยดีในกรณีที่ข้อมูลแต่ละกลุ่มมีจำนวนระเบียนค่อนข้างแตกต่างกัน กล่าวคือมีค่าตอบ กรณีที่เกิดความพิการแบบเข้าเกณฑ์ 297 และกรณีที่เกิดความพิการแบบไม่เข้าเกณฑ์เพียง 126 ซึ่งแตกต่างกันเกินสองเท่า
3. บทสรุปการเกิดอุบัติเหตุ โดยภาพรวมจังหวัดทางภาคตะวันออกเฉียงเหนือ (ขอนแก่น และ นครราชสีมา) จะเกิดอุบัติเหตุมากที่สุด และอุบัติเหตุจะเกิดมากในช่วงเดือนปลายธันวาคมถึง มกราคม เพราะเป็นช่วงเทศกาลส่งท้ายปีเก่าต้อนรับปีใหม่ โดยส่วนมากพาหนะของผู้บาดเจ็บ คือ รถจักรยานยนต์ และผู้บาดเจ็บส่วนมากจะเป็นผู้ขับขี่

ข้อเสนอแนะ

ข้อเสนอแนะที่ได้จากงานวิจัยนี้คือ การสร้างตัวพยากรณ์ด้วยการประยุกต์ใช้โครงข่ายประสาทเทียม ยังต้องมีการปรับปรุงในเรื่องของความแม่นยำ ปัจจัยหนึ่งที่อาจกระทบต่อผลการพยากรณ์คือจำนวนข้อมูลคำตอบที่มีสัดส่วนต่างกันมากเกินไป แต่อาจปรับปรุง โครงข่ายประสาทเทียมโดยเลือกแบบหลายชั้นซ่อน และจำนวนโหนดได้ นอกจากนี้อาจนำปัจจัยหรือตัวแปรตัวอื่นที่น่าจะเกี่ยวข้องกับการพิการเข้ามาในตัวแบบพยากรณ์ด้วย รวมไปถึงการใช้เทคนิคเหมืองข้อมูลแบบอื่น ๆ เช่น Naïve Bays เข้ามาประยุกต์ใช้กับข้อมูลดังกล่าวได้ด้วย และควรมีการศึกษาเพิ่มเติมกรณีการสุ่มตัวอย่างจากแต่ละกลุ่มเพื่อให้มีความสมดุลหรือใกล้เคียงกันมากที่สุดเพื่อเพิ่มความแม่นยำในการพยากรณ์ของตัวแบบ

- ดารณี สุวพันธ์ และคณะ. รายงานการวิจัยเรื่อง อุบัติการณ์ความพิการและผลกระทบที่เกิดขึ้นจากสาเหตุอุบัติเหตุจราจรทางบก พ.ศ. 2549. ศูนย์สิรินธรเพื่อการฟื้นฟูสมรรถภาพทางการแพทย์แห่งชาติ กรมการแพทย์ กระทรวงสาธารณสุข; 2551.
- Abdel-Aty, M.A., Abdelwahab, H.T., 2004. Predicting injury severity levels in traffic crashes: A modeling comparison. *Journal of transportation Engineering*, 130(2), 204-210.
- Bedard, M., Guyatt, G., Stones, M., Hirdes, J., 2002. The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. *Accident Analysis and Prevention*, 34(6), 717-727.
- Bozdogan, H. 2003, Statistical Data Mining and Knowledge discovery, Chapman & Hall/CRC, New York.
- Delen, D., Sharda, R., Bessonov, M., 2006. Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis and Prevention*, 38, 434-444.
- Dunham, N.H. 2003, Data mining introductory and advanced topics, Prentice Hall.
- Chang, L.Y., 2005. Analysis of freeway accident frequencies: Negative binomial regression versus artificial neural networks. *Safety Science*, 43, 541-557.
- Chang, L.Y., Chen, W.C., 2005. Data mining of tree-based models to analyze freeway accident frequency. *Journal of Safety Research*, 36, 365-375.
- Chang, L.Y., Wang, H. W., 2006. Analysis of traffic injury severity: An application of non-parametric classification tree techniques. *Accident Analysis and Prevention*. 38, 1019-1027.
- Chang, H.L., Yeh, T.H., 2007. Motorcyclist accident involvement by age, gender, and risky behaviors in Taipei, Taiwan. *Transportation Research Part F*, 10, 109-122.
- Kalyoncuglu, S. F., Tigdemir, M., 2004. An alternative approach for modeling and simulation of traffic data: artificial neural networks. *Simulation Modelling Practice and Theory*, 12, 351-362.
- Keall, M., Frith, W., Patterson, T., 2004. The influence of alcohol, age and number of passengers on the night-time risk of driver fatal injury in New Zealand. *Accident Analysis and Prevention*. 36(1), 49-61.
- Kdd, 2011 from website, http://www.kmitl.ac.th/~s9011142/hw2_kdd.html, retrieved on 16 August 2011.
- Kuhnert, P.M., Do, K., A., McClure, R., 2000. Combining non-parametric models with logistic regression: an application to motor vehicle injury data. *Computational Statistical Data Analysis*, 34(3). 371-386.

- Liao, C.W., Perng, Y. H., 2008. Data mining for occupational injuries in the Taiwan construction industry. *Safety Science*, 46, 1091-1102.
- Ostrom, M., Eriksson, A., 2001. Pedestrian fatalities and alcohol. *Accident Analysis and Prevention*. 33(2), 173-180.
- Scheetz, L.J., Zhang, J., Kolassa, J., 2009. Classification tree modeling to identify severe and moderate vehicular injuries in young and middle-aged adults. *Artificial Intelligence in Medicine*, 45, 1-10.
- Sohn, S.Y., Shin, H., 2001. Pattern recognition for road traffic accident severity in Korea. *Ergonomics*, 44(1), 107-117.
- Valent, F., Schiava, F., Savonitto, C., Gallo, T., Brusaferro, S., Barbone, F., 2002. Risk factors for fatal road accidents in Udine, Italy. *Accident Analysis and Prevention*. 34(1), 71-84.
- Yau, K., 2004. Risk factors affecting the severity of single vehicle traffic accident in Hong Kong. *Accident Analysis and Prevention*. 34(1), 71-84.
- Zajac, S., Ivan, J., 2003. Factors influencing injury severity of motor vehicle-crossing pedestrian crashes in rural Connecticut. *Accident Analysis and Prevention*. 35(3), 369-379.
- Zhang, J., Lindsay, J., Clarke, K., Robbins, G., Mao, Y., 2000. Factors affecting the severity of motor vehicle traffic crashes involving elderly drivers in Ontario. *Accident Analysis and Prevention*. 32(1), 117-125.

เอกสารการนำเสนอผลงาน

สารบัญ

บทคัดย่อ	I
Abstract	II
Executive Summary	III
บทที่ 1	1
บทนำ	1
ชื่อโครงการ	1
ความสำคัญและที่มาของปัญหาที่ทำการวิจัย	2
คำจำกัดความที่ใช้ในงานวิจัย	4
บทที่ 2	9
กรอบแนวคิดทฤษฎี และงานวิจัยที่เกี่ยวข้อง	9
วิธีการดำเนินการเพื่อค้นหาค้นหาองค์ความรู้ (Knowledge Discovery in Database)	9
การทำ Mining ข้อมูล	16
บทที่ 3	25
ผลการวิจัย	25
ตัวแบบพยากรณ์ที่ 1	25
ตัวแบบพยากรณ์ที่ 2	27
ตัวแบบพยากรณ์ที่ 3	32
ตัวแบบพยากรณ์ที่ 4	34
ผลการวิจัย	36
บทที่ 4	37
สรุปและข้อเสนอแนะ	37
ข้อเสนอแนะ	37
บรรณานุกรม	38
ภาคผนวก	40



รายงานวิจัยฉบับสมบูรณ์

โครงการ การวิเคราะห์ปัจจัยที่มีผลต่อความพิการจากอุบัติเหตุจราจรทางบก ด้วยเทคนิค
เหมืองข้อมูล

โดย ผศ. ดร. จรัสศรี รุ่งรัตนอุบล และคณะ

พฤษภาคม 2555

สัญญาเลขที่ R2554B070

รายงานวิจัยฉบับสมบูรณ์

โครงการ การวิเคราะห์ปัจจัยที่มีผลต่อความพิการจากอุบัติเหตุจราจรทางบก ด้วยเทคนิค
เหมืองข้อมูล

คณะผู้วิจัย

- | | |
|--------------------------------|---|
| 1. ผศ.ดร. จรัสศรี รุ่งรัตนอุบล | สังกัด ภาควิชาวิทยาการคอมพิวเตอร์และ
เทคโนโลยีสารสนเทศ
คณะวิทยาศาสตร์ ม.นเรศวร |
| 2. ผศ.ดร. อนามัย นาอุดม | สังกัด ภาควิชาคณิตศาสตร์
คณะวิทยาศาสตร์ ม.นเรศวร |
| 3. แพทย์หญิง ดารณี สุวพันธ์ | สังกัดศูนย์สิรินธรเพื่อการฟื้นฟูสมรรถภาพ
ทางการแพทย์แห่งชาติ
กรมการแพทย์ กระทรวงสาธารณสุข |

สนับสนุนโดยทุนวิจัยงบประมาณแผ่นดินประจำปี 2554
มหาวิทยาลัยนเรศวร



เอกสารงานวิจัยตามตัวชี้วัด

โครงการ การวิเคราะห์ปัจจัยที่มีผลต่อความพิการจากอุบัติเหตุจราจรทางบก ด้วยเทคนิค
เหมืองข้อมูล

โดย ผศ. ดร. จรัสศรี รุ่งรัตนอุบล และคณะ

พฤษภาคม 2555

สัญญาเลขที่ R2554B070