

ระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน A system predicts severity of an accident on the road

สุกัญญา บุญชินวุฒิกุล และ วดี ร่มโพธิ์ชี และ สุภาพร บรรดาศักดิ์

ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ คณะวิทยาศาสตร์ ศรีราชา มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตศรีราชา
Emails: sukanya.booncu@ku.th, romphochee.ku73@gmail.com, jumbundasak@hotmail.com

บทคัดย่อ

ระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน A system predicts severity of an accident on the road จัดทำขึ้นในรูปแบบเว็บไซต์และนำเทคนิคเหมืองข้อมูล (Data Mining) โดยเลือกใช้เทคนิค decision tree Algorithm มาประยุกต์ใช้ร่วมด้วย จากข้อมูลการเกิดอุบัติเหตุของโรงพยาบาลแหลมฉบัง จังหวัดชลบุรี ในปีพ.ศ.2557-2559จำนวน 1,621 รายการ ให้ความถูกต้อง 81.16 เปอร์เซ็นต์ เพื่อช่วยในการคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนนให้แก่บุคคลทั่วไปทราบถึงความรุนแรงก่อนที่จะเริ่มเดินทางไม่ว่าจะใกล้หรือไกล และสามารถดูแผนที่เฝ้าระวังที่เคยเกิดอุบัติเหตุพร้อมระดับความรุนแรงในแต่ละพื้นที่ เพื่อความปลอดภัยทั้งร่างกายและทรัพย์สิน

ABSTRACT

A system predicts severity of an accident on the road in the form of site preparation and data mining techniques by using techniques decision tree Algorithm applied in. From the data of accidents of the hospital, province in the year. 2557-2559, number of 1,621 items, ensure the accuracy 81.16 percent. together with the application. To help predict the severity of an accident on the road to the public about the severity before starting the journey, whether near or far. You can see the surveillance had been an accident and the severity of each area. Safety and physical assets.

คำสำคัญ— เทคนิคเหมืองข้อมูล(Data Mining); อุบัติเหตุบนท้องถนน; decision tree

1. บทนำ

การเกิดอุบัติเหตุบนท้องถนนเป็นปัญหาร้ายแรงอันดับต้น ๆ ของประเทศไทยในยุคปัจจุบัน หน่วยงานต่าง ๆ จึงช่วยกันรณรงค์เพื่อควบคุมอุบัติเหตุ เนื่องจากอุบัติเหตุส่งผลให้มีผู้ได้รับบาดเจ็บ พิการ หรือร้ายแรงที่สุดคือเสียชีวิต อุบัติเหตุสามารถเกิดขึ้นได้ทุกเมื่อโดยที่เราไม่คาดคิด คนส่วนมากมักจะขาดความระมัดระวังและประมาทเลินเล่อต่อสถานการณ์ที่ยังไม่มาถึง จนทำให้เกิดความเสียหายทั้งร่างกายและทรัพย์สิน ผู้จัดทำจึงนำข้อมูลการเกิดอุบัติเหตุของโรงพยาบาลแหลมฉบัง จังหวัดชลบุรี ในปีพ.ศ.2557-2559 ข้อมูลมีจำนวนทั้งหมด 1,621 รายการ โดยจำแนกตามตำบลต่าง ๆ ในเขตจังหวัดชลบุรี และระดับความรุนแรงของผู้ประสบอุบัติเหตุ สามารถแบ่งออกเป็น 4 ประเภท ได้แก่ ประเภทที่ 1 บาดเจ็บวิกฤต ประเภทที่ 2 บาดเจ็บรุนแรง ประเภทที่ 3 บาดเจ็บปานกลาง ประเภทที่ 4 บาดเจ็บเล็กน้อย ทั้งนี้การบาดเจ็บตามระดับต่าง ๆ ขึ้นอยู่กับลักษณะการเกิดอุบัติเหตุในรูปแบบต่าง ๆ ที่แตกต่างกันไป

จากนั้นนำข้อมูลเหล่านี้มาวิเคราะห์เพื่อหาระดับความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน และผลที่ได้จากการวิเคราะห์ สามารถนำไปปรับใช้ในการกำหนดทิศทางของมาตรการเฝ้าระวังและคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนนเบื้องต้น พร้อมทั้งจัดทำระบบแผนที่เฝ้าระวังและแสดงผลการวิเคราะห์ให้บุคคลทั่วไปสามารถเข้าดูผลการคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน ในขณะเดียวกันผู้ดูแลระบบสามารถจัดการข้อมูลต่าง ๆ เกี่ยวกับฐานข้อมูล เรียกดูแผนที่เฝ้าระวังและสามารถคาดการณ์ได้เช่นเดียวกัน

2. วัตถุประสงค์

จัดทำระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนนและจัดทำแผนที่พร้อม ระดับความรุนแรงในแต่ละจุดจากชุดข้อมูล

3. วรรณกรรมที่เกี่ยวข้อง

ระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนนได้นำทฤษฎีต่าง ๆ มาใช้ดังนี้

3.1. เหมืองข้อมูล

เหมืองข้อมูล (data mining) [1] ตามศัพท์ที่ราชบัณฑิตสถานกำหนดไว้ หมายถึง การสกัดหรือการวิเคราะห์ ค้นหาข้อมูลที่ต้องการจากข้อมูลจำนวนมาก หรือการวิเคราะห์ข้อมูลที่ถูกออกแบบมาเพื่อระบบสนับสนุนความต้องการของผู้ใช้ในการค้นหาข้อมูลที่ต้องการจากข้อมูลจำนวนมากได้

เหมืองข้อมูล คือ กระบวนการที่ใช้หลักการทางคณิตศาสตร์ สถิติ เพื่อทำให้คอมพิวเตอร์สามารถเรียนรู้และจัดจำข้อมูลต่าง ๆ เพื่อใช้ในการค้นหารูปแบบ (Pattern) กฎเกณฑ์ (Rule) ของความสัมพันธ์ในชุดข้อมูล (Dataset) เพื่อสกัดข้อมูลซึ่งอยู่ในรูปของสารสนเทศ (Information) เพื่อใช้ในการกระบวนการสร้างองค์ความรู้และค้นหาองค์ความรู้ที่อยู่ในข้อมูล (Knowledge Discovery in Database: KDD) เพื่อนำผลการวิเคราะห์ข้อมูลมาใช้ในการตัดสินใจโดยมุ่งเน้นการหาค่าที่จะเกิดขึ้นในอนาคต (Prospective)

CRISP-DM ย่อมาจาก Cross-Industry Standard Process for Data Mining เป็น Workflow มาตรฐานในการทำ data mining ประกอบไปด้วย 6 ขั้นตอนต่อไปนี้ 1.Business Understanding เป็นขั้นตอนแรกของ CRISP-DM ทำความเข้าใจกับปัญหาและระบุ output หรือเป้าหมายที่ต้องการได้จากการวิเคราะห์ข้อมูลด้วย data mining 2.Data Understanding เป็นจุดเริ่มต้นของการเก็บรวบรวมข้อมูล ระบุคุณภาพของ data ข้อมูลถูกต้องน่าเชื่อถือ ข้อมูลที่ได้มามีปริมาณมากพอหรือยัง ข้อมูลที่ได้มีความเหมาะสม มีรายละเอียดเพียงพอต่อการนำไปใช้ในการวิเคราะห์ 3.Data Preparation ขั้นตอนการเตรียมข้อมูลเป็นขั้นตอนที่ใช้เวลานานที่สุด เนื่องจากโมเดลที่ได้จากการทำ data mining จะให้ผลลัพธ์ที่ถูกต้องหรือไม่ขึ้นอยู่กับคุณภาพของข้อมูลที่ใช้ แบ่งออกเป็น 3 ขั้นตอนย่อย คือ 3.1.ทำการคัดเลือกข้อมูล (Data Selection) กำหนดเป้าหมายก่อนว่าจะทำการวิเคราะห์อะไร เลือกเฉพาะข้อมูลที่เกี่ยวข้องกับสิ่งที่ทำการวิเคราะห์ 3.2.ทำการกลั่นกรองข้อมูล (Data Cleaning) ลบข้อมูลที่มีความซ้ำซ้อน แก้ไขข้อมูลที่ผิดพลาด (ข้อมูลที่ผิดรูปแบบ, ข้อมูลที่หายไป, ข้อมูลที่แปลกแยกจากข้อมูลอื่น) 3.3.แปลงรูปแบบของข้อมูล (Data transformation) เป็นขั้นตอนการเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมในการใช้ในการวิเคราะห์ตามอัลกอริทึมของ data mining ที่เลือกใช้ 4.Modeling เลือกเทคนิคเพื่อใช้ในการวิเคราะห์ข้อมูล และสร้างแบบจำลอง ด้วยเทคนิค data mining

5.Evaluation ประเมินหรือวัดประสิทธิภาพของโมเดลจากการวิเคราะห์ข้อมูลในขั้นตอนของ Modeling เพื่อให้โมเดลที่ได้มีคุณภาพสูงสุด เป็นการประเมินโมเดลก่อนนำไปใช้งานจริง ถ้าโมเดลที่ได้ไม่มีคุณภาพมากพอ จะต้องเลือกเทคนิคใหม่ เพื่อให้เหมาะสมกับข้อมูลที่สุด เมื่อจบ process นี้ จะได้ผลลัพธ์จากการทำเหมืองข้อมูล 6.Deployment เป็นกระบวนการนำโมเดลหรือผลจากการวิเคราะห์ที่ได้ไปใช้งานจริง

3.2. ต้นไม้การตัดสินใจ

ต้นไม้การตัดสินใจ (Decision Tree) [2] เป็นการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปแบบโครงสร้างต้นไม้และมีการทำงานแบบ Supervised Learning (คือการเรียนรู้แบบมีครูสอน) สามารถสร้างแบบจำลองการจัดหมวดหมู่ได้จากกลุ่มตัวอย่างข้อมูลที่กำหนดไว้ก่อนล่วงหน้าและพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดหมวดหมู่ได้ด้วยรูปแบบของต้นไม้ (Tree) โครงสร้างประกอบด้วย Root Node, Child และ Leaf Node ในที่นี้เลือกใช้ J48 Algorithm ในการสร้างโมเดล

การสร้างโมเดล decision tree จะทำการคัดเลือกแอตทริบิวต์ที่มีความสัมพันธ์กับคลาสมากที่สุดขึ้นมาเป็นโหนดบนสุดของ tree (root node) หลังจากนั้นก็จะหาแอตทริบิวต์ถัดไปเรื่อย ๆ ในการหาความสัมพันธ์ของแอตทริบิวต์นี้จะใช้ตัววัด ที่เรียกว่า Information Gain (IG) คำนวณได้จากสมการดังนี้

$$IG(\text{parent, child}) = \text{entropy}(\text{parent}) - [p(c_1) \times \text{entropy}(c_1) + p(c_2) \times \text{entropy}(c_2) + \dots] \quad (1)$$

โดยการคำนวณค่า $\text{entropy}(c_i)$ คำนวณได้จากสมการดังนี้

$$\text{entropy}(c_i) = -p(c_i) \log p(c_i) \quad (2)$$

ต่อไปจะลองคำนวณค่าแต่ละแอตทริบิวต์เทียบกับคลาสเพื่อหาแอตทริบิวต์ที่มีค่า IG มากที่สุดมาเป็น root ของ decision tree โดย target class ของงานนี้ คือ ระดับความรุนแรงที่แบ่งออกเป็น 4 ประเภท 1 บาดเจ็บบริเวณ 2 บาดเจ็บรุนแรง 3 บาดเจ็บปานกลาง 4 บาดเจ็บเล็กน้อย

3.3. ทฤษฎีเกี่ยวกับอุบัติเหตุ

อุบัติเหตุ (Incidence) [8] หมายถึง เหตุการณ์ที่เกิดขึ้น อาจเป็นเหตุการณ์ที่ร้ายหรือเหตุการณ์ที่ดีก็ได้ แต่ถ้าเป็นอุบัติเหตุแล้วมักจะนึกถึงแต่เหตุร้ายไม่คิดว่าจะเป็นเรื่องดี จึงตรงกับคำว่า

Accidents ซึ่งนิยมใช้กันมาจนเป็นที่ยอมรับแล้ว ส่วนอุบัติเหตุ (Accident) คือ เหตุการณ์ที่เกิดขึ้นโดยไม่ตั้งใจ ไม่คาดฝันมาก่อน เป็นผลให้เกิดความเสียหายต่อทรัพย์สิน ร่างกาย และจิตใจ รวมทั้งอาจเป็นอันตราย ถึงแก่เสียชีวิตได้ทั้งกับตนเองและผู้อื่น

ความปลอดภัยเป็นสิ่งสำคัญต่อการใช้รถใช้ถนน เนื่องจากเป็นผลต่อทั้งตัวผู้ขับขี่และเพื่อนร่วมทางรอบข้าง แต่ทำไมสถิติการเกิดอุบัติเหตุทางรถยนต์จึงเพิ่มมากขึ้นทุกขณะ ซึ่งเมื่อตัดปัจจัยของการเพิ่มปริมาณรถยนต์บนท้องถนนที่มีโอกาสทำให้ตัวเลขผู้ประสบอุบัติเหตุเพิ่มมากขึ้นแล้ว “ความประมาท” และ “การขาดความรู้เรื่องการใช้รถใช้ถนน” น่าจะเป็นสาเหตุหลักของอุบัติเหตุที่น่าสลดใจเหล่านี้

4. วิจัยที่เกี่ยวข้อง

ผกามาศ ไพโรจน์และวิภาดา เวทย์ประสิทธิ์ [3] หลักสูตรการจัดการเทคโนโลยีสารสนเทศ คณะวิศวกรรมศาสตร์ ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ ได้ทำการศึกษารื่อง การจำแนกข้อมูลระบบเฝ้าระวังการบาดเจ็บจากอุบัติเหตุทางถนน โดยใช้เทคนิคเหมืองข้อมูล กรณีศึกษาโรงพยาบาลหาดใหญ่ จังหวัดสงขลา จุดประสงค์ในการทำคือ นำเสนอปัจจัยที่มีความสัมพันธ์ส่งผลต่อสถานะ การบาดเจ็บของผู้ประสบอุบัติเหตุ ที่เข้ารักษา ณ ห้องฉุกเฉิน ของโรงพยาบาลหาดใหญ่ งานวิจัยนี้ใช้ข้อมูลจากระบบเฝ้าระวังการบาดเจ็บของโรงพยาบาลหาดใหญ่ ที่ได้เก็บบันทึกไว้ตั้งแต่ ปี พ.ศ. 2547 ถึง พ.ศ. 2550 มีข้อมูลเฉพาะอุบัติเหตุเท่านั้น จำนวน 28,970 รายการโดยเริ่มต้นจากการทำความสะอาดข้อมูลและคัดเลือกตัวแปรโดยเทคนิค chi-square test แล้วทำการวิเคราะห์ ด้วย ต้นไม้การตัดสินใจ Naïve bayes และโครงข่ายประสาทเทียม ผลการทดลองพบว่า เทคนิค ต้นไม้การตัดสินใจให้ผลการทดลองที่มีความถูกต้องมากกว่า Naïve bayes และโครงข่ายประสาทเทียม และสรุปการเกิดอุบัติเหตุจากผลการวิเคราะห์พบว่า พาหนะที่ใช้และการดื่มแอลกอฮอล์มีผลอย่างมากต่อการเกิดอุบัติเหตุ โดยผู้ที่ดื่มแอลกอฮอล์มีโอกาสเกิดอุบัติเหตุสูงถึง 96.66 เปอร์เซ็นต์ และยัง มีพฤติกรรมอื่น ๆ อีกที่ส่งผลต่อสถานะของผู้บาดเจ็บ เช่น การใช้ยา การสวมหมวกกันน็อคและการใช้เข็มขัดนิรภัย ข้อเสนอแนะของงานวิจัยนี้ คือ ข้อมูลในอดีตที่มีจำนวนมาก ๆ อาจจะมีข้อมูลที่เสียหายหรือไม่ครบถ้วน การบันทึกข้อมูลลงไปยังเครื่องที่จะกรอกให้ครบถ้วนเพื่อความสมบูรณ์ของข้อมูลและเพื่อประโยชน์การใช้ข้อมูลต่อไป

อมรภัทร์ หาญโคกกรวด ,จारी ทองคำ, ฉัตรตระกูลสมบัติธีระ และ ชัยนันท์ สมพงษ์ [4] ได้ทำการศึกษารื่องแบบจำลองทำนายความเสี่ยงการเกิดอุบัติเหตุบนท้องถนน โดยเทคนิคถ่วงน้ำหนัก โดยมีวัตถุประสงค์ในการทำแบบจำลองทำนาย

ความเสี่ยงการเกิดอุบัติเหตุบนท้องถนน โดยรวบรวมข้อมูลมาจากสถานีภูธรจังหวัดสกลนครตั้งแต่ปี พ.ศ. 2552 ถึง พ.ศ. 2556 จำนวน 699 รายการแบ่งข้อมูลออกเป็น 3 กลุ่ม คือ เสียชีวิต บาดเจ็บเล็กน้อย และบาดเจ็บสาหัส ทำการวิเคราะห์ข้อมูลโดยใช้เทคนิค ถ่วงน้ำหนัก (Bagging : B) ,จำแนกเบย์ (Naïve bayes : NB) และเครื่องมือสนับสนุนเส้นสมมติ (Support vector machine : SVM) ผลการวิจัยพบว่า สาเหตุสำคัญของการเกิดอุบัติเหตุบนท้องถนนส่วนใหญ่เกิดจากมนุษย์ จากการวิเคราะห์เทคนิค ถ่วงน้ำหนักให้ความถูกต้องสูงสุด 67.67 เปอร์เซ็นต์ และแบบจำลองจำแนกเบย์ให้ผลลัพธ์ 71.30 เปอร์เซ็นต์ ซึ่งให้ประสิทธิภาพสูงที่สุดแตเมื่อนำเทคนิคถ่วงน้ำหนักเข้ามาผสมรวมเพื่อช่วยให้มีประสิทธิภาพมากขึ้น ทำให้แบบจำลองถ่วงน้ำหนัก ร่วมกับแบบจำแนกเบย์ทำให้ผลลัพธ์ 71.60 เปอร์เซ็นต์ ซึ่งเพิ่มจากเดิม 0.30 เปอร์เซ็นต์ ซึ่งจะเห็นว่าเทคนิคถ่วงน้ำหนักช่วยให้แบบจำลองมีประสิทธิภาพมากขึ้น เพื่อเพิ่มความถูกต้องในการทำนายของแบบจำลอง

Tibebe Beshah และ Shawndra Hill [5], Department of Information Science Operations and Information Management ,Department Addis Ababa University, Ethiopia University of Pennsylvania, Philadelphia, PA ได้ทำการศึกษารื่อง Mining Road Traffic Accident Data to Improve Safety: Role of Road-related Factors on Accident Severity in Ethiopia การทำเหมืองข้อมูลอุบัติเหตุบนท้องถนนเพื่อลดความรุนแรง จากปัจจัยด้านความรุนแรงบนท้องถนนในประเทศเอธิโอเปีย ซึ่งมีความต้องการในการหาวิธีในการลดความรุนแรงของการเกิดอุบัติเหตุจากหน่วยงานจราจรและประชาชนทั่วไป โดยมีข้อมูลทั้งหมด 18,288 รายการและแบ่งความรุนแรงของอุบัติเหตุเป็น 4 ชุด คือ Fatal, Severe injury, Slight injury, Property loss จุดประสงค์ของงานนี้คือ 1. การสำรวจข้อมูลพื้นฐานที่มีความเกี่ยวข้องกับ การเกิดอุบัติเหตุบนท้องถนนที่ส่งผลกระทบต่อยานพาหนะและความรุนแรงของการเกิดอุบัติเหตุ 2. การคาดการณ์ความรุนแรงของการเกิดอุบัติเหตุโดยใช้เทคนิคการทำเหมืองข้อมูลที่แตกต่างกันและการเปรียบเทียบรูปแบบการจัดหมวดหมู่มาตรฐานสำหรับงานนี้ ส่วนใหญ่มุ่งเน้นไปที่การคาดการณ์หรือการจัดตั้งปัจจัยสำคัญที่มีอิทธิพลต่อการบาดเจ็บรุนแรง ในงานวิจัยนี้ได้ทำการวิเคราะห์ข้อมูลโดยใช้เทคนิค Decision Tree (J48), Naive Bayes และ K-Nearest Neighbors โดย J48 ได้ความถูกต้อง 80.22 เปอร์เซ็นต์ Naive Bayes ได้ความถูกต้อง 79.99 เปอร์เซ็นต์ และ K-Nearest Neighbors ได้ความถูกต้อง 80.8281 เปอร์เซ็นต์ ซึ่งไม่แตกต่างกันมาก และผลการวิเคราะห์คือ สูญเสียทรัพย์สินจากอุบัติเหตุประมาณ 75 เปอร์เซ็นต์ ได้รับบาดเจ็บเล็กน้อยประมาณ 10 เปอร์เซ็นต์ บาดเจ็บอย่างรุนแรง

ประมาณ 8 เปอร์เซ็นต์ ผลจากการศึกษานี้สามารถนำมาใช้ประโยชน์เพื่อส่งเสริมความปลอดภัยทางถนน และผู้จัดทำได้คิดริเริ่มที่จะสร้างแบบจำลองที่จะรวมปัจจัยบนท้องถนนที่เกี่ยวข้องพัฒนาเครื่องมือสนับสนุนการตัดสินใจสำหรับการจราจรในประเทศเอโอเปียต่อไป

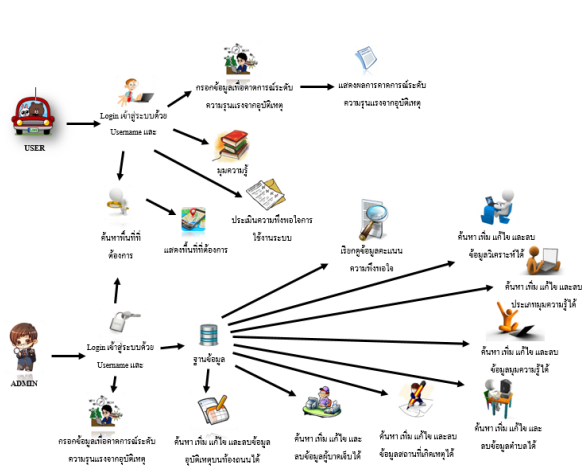
สุนิดา ศรีสุริยะชัย และ รังสิพรรณ มฤคทัต [6] คณะวิศวกรรมศาสตร์ สาขาวิชาเทคโนโลยีการจัดการสารสนเทศ มหาวิทยาลัยมหิดล ได้ทำการศึกษาเรื่อง การหารูปแบบของปรากฏการณ์โดยใช้ต้นไม้การตัดสินใจ: กรณีศึกษา อุบัติเหตุการจราจรในจังหวัดนครปฐม จุดประสงค์ของงานนี้ คือ สร้างเป็นโพรไฟล์ของอุบัติเหตุการจราจรที่เกิดขึ้นในจังหวัดนครปฐม สำหรับการวิเคราะห์ปัญหา ตลอดจนการหามาตรการแก้ไขต่อไป และข้อมูลอุบัติเหตุการจราจรที่ใช้ในงานวิจัยนี้ได้รับความอนุเคราะห์จากศูนย์เรนทร กระทรวงสาธารณสุข และจากสถานีตำรวจภูธรในจังหวัดนครปฐม ข้อมูลแบ่งออกเป็น 3 ชุดคือ 1. ข้อมูลการปฏิบัติงานประจำวันของหน่วยบริการการแพทย์ฉุกเฉิน จากศูนย์เรนทร จำนวน 1208 รายการ 2. ข้อมูลผู้ประสบอุบัติเหตุการจราจร ช่วงเทศกาลปีใหม่ ระหว่างวันที่ 27 ธันวาคม 2547 ถึง 5 มกราคม 2548 จากศูนย์เรนทร จำนวน 736 รายการ 3. ข้อมูลคดีอุบัติเหตุการจราจรที่เกิดขึ้นในจังหวัดนครปฐม จำนวน 1103 คดี กระบวนการในการใช้โมเดล J48 อัลกอริทึม C4.5 ผลการวิจัยคือ ต้นไม้การตัดสินใจสะท้อนภาพของปรากฏการณ์ที่เกิดขึ้นในแต่ละกลุ่มได้พอสมควร แต่การใช้ต้นไม้การตัดสินใจยังมีข้อจำกัดอยู่ ดังนั้น การสร้างโพรไฟล์ของอุบัติเหตุการจราจรยังต้องอาศัยการทำเหมืองข้อมูลหลาย ๆ แบบประกอบกัน เช่น Naïve Bayes มาวิเคราะห์เปรียบเทียบด้วย

S. Krishnaveni ,Ph. D (CS) [7] Research Scholar,Karpagam University Dr. M. Hemalatha,Asst. Professor & Head, Dept of Software Systems Karpagam University ได้ทำการศึกษาเรื่อง A Perspective Analysis of Traffic Accident using Data Mining Techniques การวิเคราะห์ข้อมูลอุบัติเหตุบนการจราจร โดยใช้เหมืองข้อมูล การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อการพัฒนาแบบจำลองโดยการทำการเหมืองข้อมูล เพื่อสนับสนุนการวิเคราะห์ความรุนแรงของอุบัติเหตุจราจรทางถนนเพื่อประโยชน์ในการป้องกันและควบคุมอุบัติเหตุบนท้องถนน งานนี้เลือกใช้เทคนิค Naïve Bayes, AdaBoostM1, PART, J48 และ Random Forest ชุดข้อมูลสำหรับการศึกษามีข้อมูลอุบัติเหตุจราจรของปี 2008 รวมจำนวน 34,575 รายการ ข้อมูลชุดนี้มีแต่ข้อมูลคนขับรถเท่านั้นและไม่มีรวมถึงผู้โดยสาร การเตรียมข้อมูลแบ่งออกเป็น 3 ประเภท 1. ข้อมูลที่เกี่ยวข้องกับอุบัติเหตุ 14,576 รายการ 2. ข้อมูลจากผู้ขับและยานพาหนะ 9,628 รายการ 3. ข้อมูลที่เกี่ยวข้องกับการ

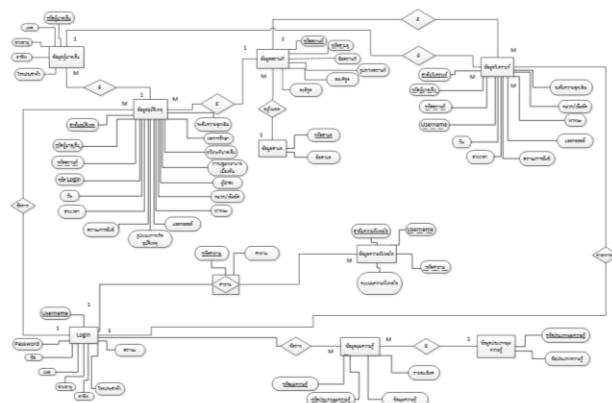
บาดเจ็บ 10,371 รายการ ผลของงานวิจัยนี้ คือ แบบจำลองได้ทำการวิเคราะห์ข้อมูลทั้งหมด 3 ประเภท โดยข้อมูลชุดที่ 1 ข้อมูลที่เกี่ยวข้องกับอุบัติเหตุ Naïve Bayes ให้ความถูกต้อง 68.47 เปอร์เซ็นต์ J48 ให้ความถูกต้อง 71.34 เปอร์เซ็นต์ AdaBoostM1 ให้ความถูกต้อง 67.07 เปอร์เซ็นต์ PART ให้ความถูกต้อง 72.03 เปอร์เซ็นต์ และ Random Forest 75.61 เปอร์เซ็นต์ ข้อมูลชุดที่ 2 ข้อมูลจากผู้ขับและยานพาหนะ Naïve Bayes ให้ความถูกต้อง 51.05 เปอร์เซ็นต์ J48 ให้ความถูกต้อง 51.75 เปอร์เซ็นต์ AdaBoostM1 ให้ความถูกต้อง 49.64 เปอร์เซ็นต์ PART ให้ความถูกต้อง 52.33 เปอร์เซ็นต์ และ Random Forest 53.84 เปอร์เซ็นต์ ข้อมูลชุดที่ 3 ข้อมูลที่เกี่ยวข้องกับการบาดเจ็บ Naïve Bayes ให้ความถูกต้อง 33.05 เปอร์เซ็นต์ J48 ให้ความถูกต้อง 34.23 เปอร์เซ็นต์ AdaBoostM1 ให้ความถูกต้อง 27.60 เปอร์เซ็นต์ PART ให้ความถูกต้อง 35.38 เปอร์เซ็นต์ และ Random Forest 36.49 เปอร์เซ็นต์ สรุปคือเทคนิค Random Forest สามารถให้ความถูกต้องของข้อมูลทั้ง 3 ประเภทได้มากกว่าเทคนิคอื่น ๆ

5. วิธีดำเนินการ

ระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน ผู้ใช้งานระบบนี้จะแบ่งเป็น 2 ส่วน คือ บุคคลทั่วไป และผู้ดูแลระบบ ซึ่งการทำงานของระบบสำหรับผู้ใช้งานระบบนั้นมีลักษณะดังนี้ คือ ผู้ใช้งานระบบทั้ง 2 ประเภท ต้องเข้าสู่ระบบโดยการใส่ชื่อผู้ใช้และรหัสผ่านก่อนเข้าสู่ระบบทุกครั้ง แต่สำหรับผู้ที่ไม่เคยใช้ระบบนี้ก็สามารถลงทะเบียนและสามารถเข้าใช้ระบบได้เช่นกัน ในส่วนของบุคคลทั่วไป สามารถค้นหาพื้นที่ที่ต้องการเดินทางไปเพื่อดูประวัติการเกิดอุบัติเหตุ สามารถคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุอุบัติเหตุ สามารถดูข้อมูลเกี่ยวกับความรู้ต่าง ๆ ได้เพื่อปรับใช้ในชีวิตรประจำวัน สามารถประเมินความพึงพอใจหลังจากใช้งานระบบนี้ และในส่วนของผู้ดูแลระบบ สามารถค้นหา เพิ่ม แก้ไข และลบข้อมูลต่าง ๆ ในฐานข้อมูลประกอบไปด้วย ข้อมูลอุบัติเหตุบนท้องถนน ข้อมูลผู้บาดเจ็บ ข้อมูลสถานที่เกิดเหตุ ข้อมูลตำบล ข้อมูลมุมมอง ความรู้ ประเภทมุมมอง ความรู้ ข้อมูลวิเคราะห์ สามารถเรียกดูข้อมูลคะแนนความพึงพอใจเพื่อดูผลการประเมิน และสามารถดูแผนที่และคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุได้เช่นเดียวกับบุคคลทั่วไป



รูปที่ 1. รูปแบบการทำงานของระบบโดยรวม



รูปที่ 2. ER-diagram ของระบบคาดการณ์ความรุนแรงจากอุบัติเหตุบนท้องถนน

ตารางที่ 1. แอตทริบิวต์ข้อมูลที่จะนำไปคาดการณ์ความรุนแรง

ชื่อคอลัมน์	รายละเอียด
Date	วันที่เดินทางหรือต้องการคาดการณ์
Gender	เพศ
Age	ช่วงอายุ (ปี)
Job	อาชีพ
Condition not available	มีโรคประจำตัวหรือไม่
Scene	สถานที่ที่จะเดินทางไป
Time	เวลาที่เดินทาง
Status	สถานะการขับขี่
Vehicles	พาหนะที่ใช้
Drunk	ใช้แอลกอฮอล์หรือไม่
Safety	สวมหมวกหรือเข็มขัดนิรภัยหรือไม่
Severity	ระดับความรุนแรงโดยเป็น target class ของงานนี้จากการคาดการณ์ของคอมพิวเตอร์

จากตัวอย่างแอตทริบิวต์ข้อมูลที่จะนำมาใช้นั้นคือตัวอย่างข้อมูลที่จะให้ผู้ใช้งานกรอกข้อมูลนำไปวิเคราะห์ เพื่อนำไปคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุต่อไป

6. ผลการศึกษา/การทดลอง

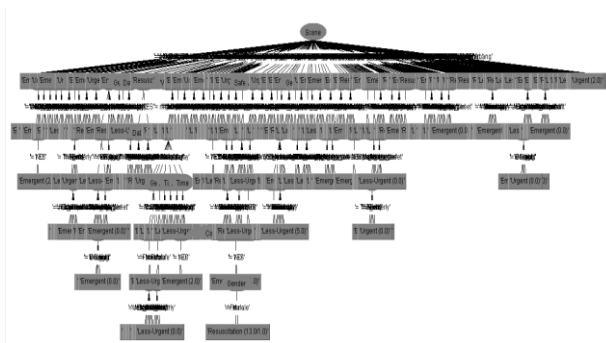
6.1 การวิเคราะห์และออกแบบระบบ

การสร้างเว็บไซต์ (Website) สำหรับบุคคลทั่วไป เรียกดูแผนที่เฝ้าระวัง คาดการณ์ความรุนแรงจากอุบัติเหตุบนท้องถนน ได้ทราบถึงมุมมองความรู้อย่างต่าง ๆ ที่เป็นประโยชน์ในการดำเนินชีวิต และประเมินผลความพึงพอใจหลังจากการใช้งานระบบนี้แล้ว สำหรับผู้ดูแลระบบ เรียกดูแผนที่เฝ้าระวัง คาดการณ์ความรุนแรงจากอุบัติเหตุบนท้องถนน จัดการข้อมูล เช่น ค้นหา เพิ่ม ลบ แก้ไขข้อมูลในฐานข้อมูล

การรวบรวมและจัดเก็บข้อมูล ได้รวบรวมข้อมูลทะเบียนบันทึกข้อมูลจราจร หน่วยงานอุบัติเหตุ - อุบัติเหตุ จากโรงพยาบาลแหลมฉบัง อ.ศรีราชา จ.ชลบุรี เป็นเวลา 2 ปี ตั้งแต่ปี พ.ศ. 2557 – 2559 จำนวนทั้งหมด 1,621 รายการ เพื่อนำมาวิเคราะห์ด้วย J48 Algorithm และจัดทำระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนนโดยมีตัวอย่างข้อมูลดังนี้

6.2. การคาดการณ์

จัดเตรียมข้อมูลอุบัติเหตุรายวันโดยจะนำไปวิเคราะห์เพื่อคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน โดย J48 Algorithm ทำการคำนวณค่า Information gain เพื่อใช้ในการวิเคราะห์และสร้างโมเดลเพื่อคาดการณ์ โดยใช้โปรแกรม WEKA version 3.6 จากข้อมูลห้องฉุกเฉินของโรงพยาบาลแหลมฉบัง จำนวน 1,621 รายการ ให้ความถูกต้อง 81.16 เปอร์เซ็นต์ โมเดลที่ได้แสดงดังรูปที่ 3



รูปที่ 3. โมเดลของระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน

เมื่อได้โมเดลก็จะนำข้อมูลที่ผู้ใช้กรอกเพื่อคาดการณ์มาวิเคราะห์กับโมเดลและทำการแสดงผลการคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุ โดยผลจะออกมาเป็นระดับความรุนแรงที่จะอาจจะได้รับเมื่อเกิดอุบัติเหตุขึ้นแบ่งเป็น 4 ระดับ คือ 1 บาดเจ็บวิกฤต 2 บาดเจ็บรุนแรง 3 บาดเจ็บปานกลาง และ 4 บาดเจ็บเล็กน้อย บนระบบเพื่อเตือนผู้ขับขี่หรือผู้ใช้รถ ใช้ถนนให้ขับขี่โดยปลอดภัย

7. สรุปผลและข้อเสนอแนะ

ระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน ได้พัฒนาขึ้นมาเพื่อคาดการณ์ความรุนแรงจากการวิเคราะห์ข้อมูลอุบัติเหตุที่เกิดขึ้นมาแล้ว และจัดทำแผนที่เฝ้าระวังพร้อมระบุความรุนแรงจากข้อมูลอุบัติเหตุ

ระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน จัดทำเป็นเว็บไซต์ (Website) พัฒนาโดย ASP.NET เป็นระบบเปิดและการใช้งานง่าย เหมาะสำหรับการเรียนรู้และการเปลี่ยนแปลงในอนาคต ในส่วนของฐานข้อมูลใช้ MS SQL Server 2012 มาใช้งาน

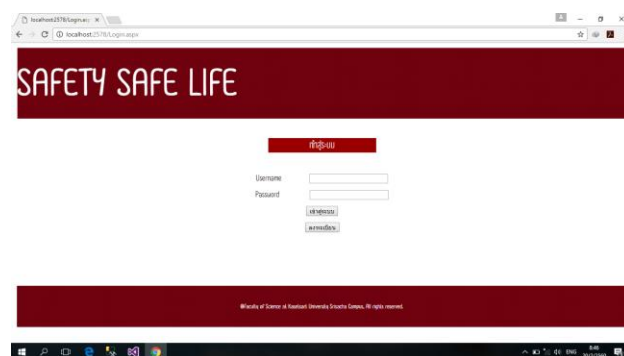
8. การอภิปรายผล

บุคคลทั่วไปในปัจจุบันไม่มีความระมัดระวังในการขับขี่หรือเดินทางไปไหนมาไหน ทำให้เกิดอุบัติเหตุอยู่บ่อยครั้ง ผลที่ตามมาคือการได้รับความบาดเจ็บทั้งร่างกายและรวมถึงการสูญเสียซึ่งทรัพย์สิน ทั้งในส่วนของผู้ขับขี่ ผู้โดยสารหรือรวมไปถึงผู้ร่วมใช้เส้นทางเดียวกันทั้งสิ้น

จึงได้พัฒนาระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน ให้แก่บุคคลทั่วไปได้เข้าใช้ระบบเพื่อคาดการณ์ความรุนแรงในจุดพื้นที่ที่จะเดินทางเพื่อรับรู้ความรุนแรงที่อาจจะเกิดแล้วเตรียมรับมือกับสถานการณ์นั้น ๆ และรับรู้ความรู้ต่าง ๆ เพื่อปรับใช้ในสถานการณ์ต่าง ๆ

ระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน เป็นเพียงการคาดการณ์ความรุนแรงเบื้องต้นเท่านั้น อาจจะเกิดหรือไม่ก็ได้ ไม่ได้มีเจตนาที่จะให้บุคคลที่เข้าใช้ระบบเกิดความตื่นตระหนกหรือวิตกกังวลมากเกินไปจนเกินเหตุ

ระบบคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนนมีการออกแบบหน้าจอเข้าระบบดังรูปที่ 4



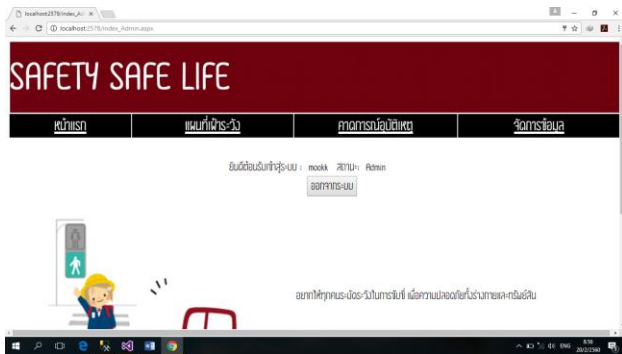
รูปที่ 4. หน้าเข้าสู่ระบบสำหรับผู้ใช้งานทุกประเภท

จากรูปที่ 4 จะเป็นการเข้าสู่ระบบสำหรับผู้ใช้งานทุกประเภท ที่มีชื่อผู้ใช้และรหัสผ่านแล้ว ถ้าผู้ไม่เคยใช้ระบบสามารถลงทะเบียนและใช้งานระบบนี้ได้เช่นกัน ผู้ใช้งานระบบทุกคนที่ต้องการเข้าใช้ระบบนี้ต้องเข้าสู่ระบบก่อนทุกครั้ง



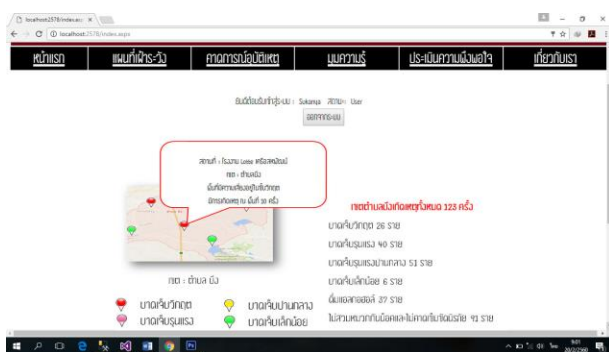
รูปที่ 5. หน้าหลักสำหรับบุคคลทั่วไป

จากรูปที่ 5 จะเป็นหน้าหลักสำหรับบุคคลทั่วไปที่เข้าสู่ระบบเรียบร้อยแล้ว หน้าจอก็จะแสดงให้เห็นว่า ใครเป็นผู้ใช้อยู่สถานะของผู้ใช้ และเมนูต่าง ๆ ที่สามารถใช้งานในระบบนี้ได้



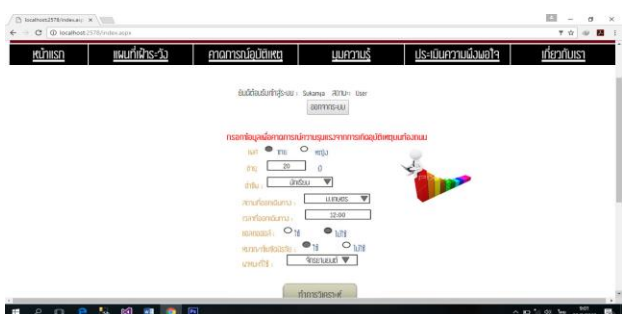
รูปที่ 6. หน้าหลักสำหรับผู้ดูแลระบบ

จากรูปที่ 6 จะเป็นหน้าหลักสำหรับผู้ดูแลระบบที่เข้าสู่ระบบเรียบร้อยแล้ว หน้าจอก็จะแสดงให้เห็นว่า ใครเป็นผู้ใช้อยู่สถานะของผู้ใช้ และเมนูต่าง ๆ ที่สามารถใช้งานในระบบนี้ได้



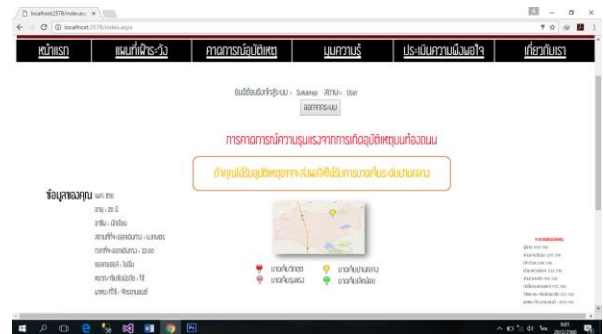
รูปที่ 7. หน้าแผนที่เฝ้าระวัง

จากรูปที่ 7 จะเป็นแผนที่ที่ระบุความรุนแรงในแต่ละจุดพื้นที่จากการเกิดอุบัติเหตุ แยกตามสีและระดับความรุนแรง 4 สีดังนี้ คือ สีแดง บาดเจ็บวิกฤต สีชมพู บาดเจ็บรุนแรง สีเหลือง บาดเจ็บปานกลาง และสีเขียว บาดเจ็บเล็กน้อย พร้อมระบุว่าข้อมูลโดยย่อในแต่ละพื้นที่



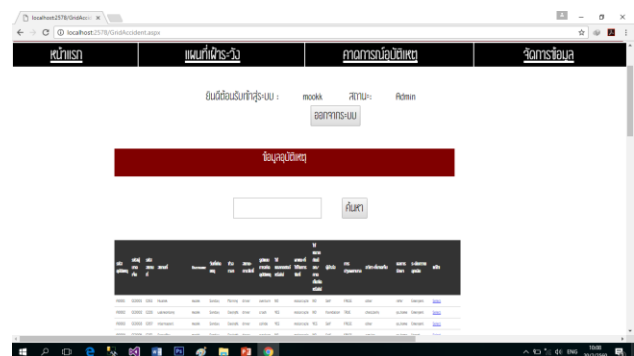
รูปที่ 8. หน้ากรอกข้อมูลเพื่อคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน

จากรูปที่ 8 เป็นการกรอกข้อมูลส่วนตัวเพื่อนำข้อมูลไปวิเคราะห์กับโมเดลที่สร้างก่อนหน้า และแสดงผลในรูปที่ 9



รูปที่ 9. หน้าแสดงผลการคาดการณ์ความรุนแรงจากการเกิดอุบัติเหตุบนท้องถนน

จากรูปที่ 9 จะแสดงข้อมูลที่ใช้กรอกมาเพื่อตรวจสอบว่าข้อมูลที่ผู้ใช้กรอกมาถูกต้อง และมีผลคาดการณ์ออกมาให้ทราบถึงระดับความรุนแรงที่อาจจะได้รับเมื่อเกิดอุบัติเหตุได้



รูปที่ 10. หน้าตัวอย่างการแสดงผลข้อมูลสำหรับผู้ดูแลระบบ



รูปที่ 11. การตัวอย่างการบันทึกข้อมูล

จากรูปที่ 10 และ 11 เป็นการยกตัวอย่างการแสดงผลข้อมูลและการบันทึกข้อมูลที่เข้ามาใหม่โดยผู้ดูแลระบบ โดยในนี้ยกตัวอย่างเป็นข้อมูลอุบัติเหตุทั้งหมด

9. เอกสารอ้างอิง

[1] ผู้ช่วยศาสตราจารย์ ดร. วีระยุทธ พิมพ์ภรณ์. เทคโนโลยีเหมืองข้อมูล (Data Mining Technologies) Conceptual /

Theoretical / Practical / Application. 199 ถนนสุขุมวิท ตำบลทุ่งสุขลา อำเภอศรีราชา ชลบุรี 20230: 2559

[2] Eakasit Pacharawongsakda, Ph.D. Practical Data Mining with RapidMiner Studio 7. บริษัท เอเชีย ดิจิตอล การพิมพ์ จำกัด ถนนงามวงศ์วาน แขวงลาดยาว เขตจตุจักร กรุงเทพฯ 10900: 2559

[3] ผกามาศ ไพโรจน์และวิภาดา เวทย์ประสิทธิ์. “การจำแนกข้อมูลระบบเผื่อระวังการบาดเจ็บจากอุบัติเหตุทางถนน โดยใช้เทคนิคเหมืองข้อมูล กรณีศึกษาโรงพยาบาลหาดใหญ่ จังหวัดสงขลา”. The 6TH National Conference on Computing and Information Technology, 2553, หน้า 219 – 224.

[4] อมรภัทร์ หาญโคกกรวด, จาริ ทองคำ, ฉัตรตระกูล สมบัติธีระและชัยนันท์ สมพงษ์. “แบบจำลองทำนายความเสี่ยงการเกิดอุบัติเหตุบนท้องถนน โดยเทคนิคถ่วงน้ำหนัก”. วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม การประชุมวิชาการมหาวิทยาลัยมหาสารคามวิจัย ครั้งที่ 9, 2556, หน้า 253 – 262.

[5] Tibebe Beshah and Shawndra Hill, “Mining Road Traffic Accident Data to Improve Safety: Role of Roadelated Factors on Accident Severity in Ethiopia ”, Ethiopia and Philadelphia: Addis Ababa University and University of Pennsylvania, pp 14-19, 2009.

[6] สุนิดา ศรีสุริยะชัย และรังสิพรรณ มฤคทัต. “การหารูปแบบของปรากฏการณ์โดยใช้ต้นไม้การตัดสินใจ: กรณีศึกษา อุบัติเหตุการจราจรในจังหวัดนครปฐม”. โครงการวิจัยเหมืองข้อมูลการจราจรสำหรับจังหวัดนครปฐม, 2548 .

[7] S.Krishnaveni and Dr.M.Hemalatha, “A Perspective Analysis of Traffic Accident using Data Mining Techniques” , International Journal of Computer Applications (0975 – 8887) ,Volume 23, No.7, pp. 40-48, June 2011.

[8] อุบัติเหตุคืออะไร. [ออนไลน์]. เข้าถึงได้จาก: <http://www.ipesp.ac.th/learning/supitcha/html/B1-1-1.html>. (วันที่ค้นข้อมูล : 2 พฤศจิกายน 2559)