

บทที่ 1

บทนำ

โครงการเรื่อง การวิเคราะห์ข้อมูลผู้บาดเจ็บและเสียชีวิตในช่วงเทศกาลปีใหม่ประจำปี 2551 – 2558 ในบทนี้จะเป็นการกล่าวถึงความสำคัญ และที่มาของปัญหา วัตถุประสงค์ ประโยชน์ที่จะได้รับ ขอบเขตงานของคณะผู้จัดทำและผู้ใช้งานทั่วไปบนเว็บไซต์ เครื่องมือที่ใช้ในการพัฒนาเว็บไซต์ สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล และระยะเวลาในการดำเนินการ รวมถึงนิยามศัพท์เฉพาะของโครงการ ดังนี้

1.1 ความสำคัญและที่มาของปัญหา

ในยุคปัจจุบันเทคโนโลยีดิจิทัลเข้ามาเป็นส่วนหนึ่งในการดำรงชีวิตประจำวันของคนเราเป็นอย่างมาก หน่วยงานต่าง ๆ ไม่ว่าจะเป็นภาครัฐกิจ องค์กรต่าง ๆ รวมถึงหน่วยงานภาครัฐเอง ต่างก็มีการนำเทคโนโลยีต่าง ๆ เข้ามาช่วยเพื่อปรับปรุงประสิทธิภาพการดำเนินงาน และยกระดับการให้บริการไปสู่ประชาชนให้สามารถเข้าถึงบริการต่าง ๆ ได้สะดวก และรวดเร็วยิ่งขึ้น จึงทำให้เกิดข้อมูลสารสนเทศขึ้นอย่างมากมาขนาดมหึมา โดยเฉพาะอย่างยิ่งข้อมูลของภาครัฐที่จัดเก็บในรูปแบบของฐานข้อมูลนับเป็นเวลาหลายสิบปี และกำลังเพิ่มขึ้นอีกในปัจจุบัน และอนาคต ข้อมูลเหล่านี้สามารถนำมาใช้ประโยชน์ได้ทั้งสำหรับภาครัฐ ภาคเอกชน และประชาชนซึ่งทำให้เกิดเป็นแนวคิด วิธีการ รูปแบบทางธุรกิจ รวมถึงวิถีชีวิตใหม่ ๆ การแบ่งปันข้อมูลเพื่อนำมาใช้ประโยชน์ และสร้างนวัตกรรมนั้น ซึ่งถือเป็นสัญญาณเริ่มต้นของการขับเคลื่อนเศรษฐกิจด้วยข้อมูล อันจะช่วยให้เศรษฐกิจของประเทศเติบโตขึ้น และเพิ่มขีดความสามารถในการแข่งขันในเวทีโลกได้ ที่ผ่านมามากหลาย ๆ ประเทศมีการผลักดัน และดำเนินการเกี่ยวกับการเปิดเผยข้อมูลสู่สาธารณะกันอย่างจริงจัง เช่นเดียวกันกับในประเทศไทย หลายๆ หน่วยงานก็เริ่มเห็นถึงความสำคัญ และประโยชน์ของการเปิดเผยข้อมูลกันมากขึ้น ดังนั้นสำนักงานรัฐบาลอิเล็กทรอนิกส์ (องค์การมหาชน) (สรอ.) ซึ่งมีภารกิจหลักในการพัฒนา รัฐบาลอิเล็กทรอนิกส์ของประเทศให้มีคุณภาพ จึงได้พัฒนาระบบศูนย์กลางข้อมูลภาครัฐ หรือ data.go.th เพื่อเป็นการส่งเสริมและผลักดันให้เกิดการเปิดเผยข้อมูลภาครัฐสู่สาธารณะอย่างเป็นรูปธรรม เป็นช่องทางให้ผู้ให้บริการทั้งภาคประชาชน ภาคธุรกิจเอกชน รวมถึงหน่วยงานของรัฐ สามารถค้นหาและเข้าถึงข้อมูลที่มีคุณภาพของภาครัฐได้ง่าย โดยมีเป้าหมายสูงสุดเพื่อส่งเสริมให้ เกิดธรรมาภิบาล ส่งเสริมให้เกิดความร่วมมือของภาครัฐ และสร้างการมีส่วนร่วม

ของภาคประชาชน รวมถึงแก้ปัญหาการทุจริตคอร์รัปชัน การเมือง และเศรษฐกิจอย่างบูรณาการร่วมกันได้ การวิเคราะห์ข้อมูลผู้บาดเจ็บและเสียชีวิตในช่วงเทศกาลปีใหม่ คือการนำข้อมูลที่ได้ถูกรวบรวมไว้และถูกเผยแพร่ผ่านเว็บไซต์ Open Government Data of Thailand ซึ่งเป็นเว็บไซต์ที่เก็บรวบรวมชุดข้อมูลต่าง ๆ ของประเทศไทยซึ่งเป็นข้อมูลที่เปิดเผยได้ เพื่อให้ผู้ใช้บริการสามารถนำชุดข้อมูลไปศึกษาหรือวิเคราะห์ให้เกิดประโยชน์ต่อไปได้ การวิเคราะห์ข้อมูลปริมาณมหาศาลแสดงให้เห็นประโยชน์ต่าง ๆ ก็คือ ความรวดเร็วและประสิทธิภาพ ขณะที่หลายปีก่อนหน้านั้นหลายธุรกิจต้องรวบรวมข้อมูล ดำเนินการวิเคราะห์ และค้นหาข้อมูลที่สามารถนำมาใช้ประกอบการตัดสินใจในอนาคตได้ ส่วนในทุกวันนี้ธุรกิจสามารถระบุข้อมูลเชิงลึกสำหรับการกระทำที่ต้องการการตัดสินใจอย่างทันท่วงที ความสามารถในการทำงานได้เร็วขึ้น และยังคงไว้ซึ่งความคล่องตัว ช่วยให้หลายองค์กรสร้างความได้เปรียบด้านการแข่งขันอย่างที่ไม่เคยมีมาก่อน ซึ่งการนำข้อมูลผู้บาดเจ็บและเสียชีวิตในช่วงเทศกาลปีใหม่ประจำปี 2551 – 2558 มาทำการวิเคราะห์นั้น เนื่องจากต้องการทราบถึงข้อมูลในเชิงลึกของสาเหตุการบาดเจ็บและเสียชีวิต เช่น ปัจจัยที่นำมาซึ่งการเกิดอุบัติเหตุ อาทิเช่น การประพาดตัวของผู้ประสบอุบัติเหตุก่อนเกิดอุบัติเหตุ และปัจจัยอื่น ๆ ที่นำไปสู่การบาดเจ็บและเสียชีวิต เพื่อหาทางป้องกันและลดอัตราการบาดเจ็บและเสียชีวิต ซึ่งการวิเคราะห์ปัจจัยต่าง ๆ จากข้อมูลที่ได้มานั้นถือเป็นเรื่องสำคัญต่อการป้องกันภัยจากอุบัติเหตุต่าง ๆ ในช่วงเทศกาล แต่ด้วยจำนวนข้อมูลที่มีปริมาณมหาศาลทำให้ยากต่อการนำเสนอให้เข้าใจได้ง่าย ซึ่งถือเป็นข้อจำกัดหากต้องการนำข้อมูลชุดนี้ไปใช้เพื่อเผยแพร่ให้แก่บุคคลภายนอกหรือกลุ่มผู้ใช้ข้อมูลได้รับรู้

จากปัญหาข้างต้นดังนั้นคณะผู้จัดทำจึงได้นำชุดข้อมูลที่ได้มานั้นเข้าสู่กระบวนการจำแนกตามรายการชุดข้อมูล โดยการใช้ขีดความสามารถในการประมวลผลของโปรแกรมคอมพิวเตอร์เพื่อเพิ่มความน่าเชื่อถือในการวิเคราะห์ และประมวลผลข้อมูลให้เหมาะสมตรงกับความต้องการ เป็นการแยกแยะสิ่งที่จะพิจารณาออกเป็นส่วนย่อยที่มีความสัมพันธ์กัน เพื่อทำความเข้าใจแต่ละส่วนให้ชัดเจน รวมทั้งการสืบค้นความสัมพันธ์ของส่วนต่าง ๆ เพื่อดูว่าส่วนประกอบปลีกย่อยนั้นสามารถเข้ากันได้หรือไม่ มีความสัมพันธ์เกี่ยวเนื่องกันอย่างไร โดยคณะผู้จัดทำได้ทำการจำแนกข้อมูลออกเป็นประเภทต่าง ๆ ด้วยเทคนิคการทำเหมือนข้อมูลแบบ Classification เป็นกระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้ โดยอาศัยการเรียนรู้ข้อมูลเก่าในช่วงเวลาที่ผ่านมาในรูปแบบ Descriptive Analytics คือข้อมูลเหตุการณ์ที่เกิดขึ้นแล้ว เป็นการวิเคราะห์ข้อมูลแบบพื้นฐาน โดยจะเน้นไปที่ความสามารถในการอธิบายว่าเกิดอะไรขึ้นกับเหตุการณ์ และสามารถอธิบายได้ว่าทำไมถึงเกิดขึ้น โดย Classification สร้างกฎเพื่อช่วยในการตัดสินใจจากข้อมูลที่มีอยู่ เพื่อใช้ทำนายแนวโน้มการเกิดขึ้นของปัจจัยที่ก่อให้เกิด

ผู้บาดเจ็บและเสียชีวิตจากเทคนิคการจำแนกประเภทข้อมูล ด้วยการวิเคราะห์ในรูปแบบของแผนภูมิต้นไม้ หรือที่เรียกว่า Decision tree เป็นโครงสร้างที่ใช้แสดงกฎที่ได้จากเทคนิคการจำแนกประเภทข้อมูล โดย Decision tree แต่ละโหนดแสดงคุณลักษณะ (attribute) ส่วนโหนดลูกแสดงเงื่อนไขในการทดสอบ และโหนดปลาย (leaf node) แสดงกลุ่มที่กำหนดไว้ ซึ่งการจำแนกข้อมูลออกเป็นประเภทต่าง ๆ นี้จะช่วยให้เกิดความเข้าใจต่อข้อมูลและสิ่งที่ต้องการทราบอย่างแท้จริง และทางคณะผู้จัดทำยังได้ทำการนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพด้วยโปรแกรม Tableau Public และเผยแพร่ข้อมูลสารสนเทศนี้บน Web browser ที่เป็นที่ยอมรับในยูทิลิตี้เน็ตคือการเผยแพร่ทางสื่อออนไลน์ โดยใช้ภาษา HTML ในการพัฒนาหน้าเว็บไซต์สำหรับการเปิดเผยข้อมูลภาครัฐที่ได้มาจาก data.go.th หรือใน Open Government Data of Thailand เพื่อให้ผู้ใช้งานสามารถเข้าถึงข้อมูลได้อย่างเหมาะสม ซึ่งประโยชน์ที่ผู้ใช้งานทั่วไปจะได้รับจะเป็นในเรื่องของความเข้าใจ การรับรู้ และการตระหนักถึงความปลอดภัยในเหตุการณ์ต่าง ๆ สำหรับผู้ที่ทำการค้นหาและศึกษาข้อมูลก็สามารถดำเนินการผ่าน Web browser ที่ได้เผยแพร่ข้อมูลสารสนเทศของผู้บาดเจ็บและเสียชีวิตในเทศกาลปีใหม่ประจำปี 2551 – 2558 ที่ผ่านการสรุปจำแนกข้อมูลตามประเภทต่าง ๆ ได้สะดวกรวดเร็วยิ่งขึ้น รวมถึงมีความถูกต้อง สะดวก ลดความซ้ำซ้อน เพิ่มประสิทธิภาพให้กับการศึกษาค้นคว้า และทันต่อเวลา

1.2 วัตถุประสงค์

เพื่อวิเคราะห์ข้อมูลผู้บาดเจ็บและเสียชีวิตในเทศกาลปีใหม่ประจำปี 2551 – 2558 สำหรับเผยแพร่บนเว็บไซต์

1.3 ประโยชน์ที่จะได้รับ

ได้ข้อมูลสารสนเทศของผู้บาดเจ็บและเสียชีวิตในเทศกาลปีใหม่ประจำปี 2551 – 2558 สำหรับเผยแพร่บนเว็บไซต์

1.4 ขอบเขต

1.4.1 ขอบเขตคณะผู้จัดทำ

กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM

1.4.1.1 Business Understanding คณะผู้จัดทำทำความเข้าใจกับปัญหาให้อยู่ในรูป

ของการวิเคราะห์ข้อมูลทาง Data Mining โดยการวิเคราะห์ข้อมูลในประเด็นนี้ คือ ข้อมูลผู้บาดเจ็บและเสียชีวิตในช่วงเทศกาลปีใหม่ประจำปี 2551 – 2558 ซึ่งมีจำนวนข้อมูลมหาศาล

ทำให้ไม่สามารถทำความเข้าใจกับข้อมูลที่มีความซับซ้อนได้อย่างรวดเร็ว เช่น ต้องการทราบว่า ช่วงอายุใดมีอัตราการเกิดอุบัติเหตุสูงสุด

1.4.1.2 Data Understanding คณะผู้จัดทำทำการรวบรวมข้อมูล เพื่อตรวจสอบ รายละเอียด ปริมาณ และความน่าเชื่อถือของข้อมูลผู้บาดเจ็บและเสียชีวิตในช่วงเทศกาลปีใหม่ ประจำปี 2551 – 2558 ที่ได้จากเว็บไซต์ data.go.th

1.4.1.3 Data Preparation คณะผู้จัดทำทำการคัดเลือกข้อมูล และทำการ Data Cleaning ข้อมูลเพื่อให้ได้ข้อมูลที่สมบูรณ์ ด้วยการตัดข้อมูลที่ไม่น่าสนใจออกไป เช่น รหัสจังหวัด รหัสโรงพยาบาล เพื่อเหลือข้อมูลที่จำเป็นนำไปวิเคราะห์ข้อมูล

1.4.1.4 Modeling คณะผู้จัดทำวิเคราะห์ข้อมูลด้วยการทำเหมืองข้อมูลแบบ Data Classification เพื่อใช้ทำนายแนวโน้มการเกิดขึ้นของปัจจัยที่ก่อให้เกิดผู้บาดเจ็บและเสียชีวิตจาก เทคนิคการจำแนกประเภทข้อมูล ด้วยการสร้างโมเดล Decision Tree เพื่อจัดกลุ่มข้อมูลที่มีความสัมพันธ์กัน โดยใช้โปรแกรมที่ใช้ทำเหมืองข้อมูล

1.4.1.5 Evaluation คณะผู้จัดทำได้ข้อมูลที่ผ่านการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining จากการสร้างโมเดล Decision Tree และการสรุปผลข้อมูลเพื่อให้ตรงกับวัตถุประสงค์ที่ตั้งไว้

1.4.1.6 Deployment คณะผู้จัดทำนำผลข้อมูลที่ทำกรวิเคราะห์แล้ว แสดงผล ข้อมูลบน Web browser โดยใช้ชุดคำสั่ง HTML และ CSS3 ร่วมกับการนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม Tableau Public

1.4.2 ขอบเขตผู้ใช้งานทั่วไปบนเว็บเบราว์เซอร์

1.4.2.1 สามารถดูข้อมูลสารสนเทศของผู้บาดเจ็บและเสียชีวิตในช่วงเทศกาลปีใหม่ ประจำปี 2551 – 2558 ที่เผยแพร่ผ่านทางเว็บไซต์

1.4.2.2 สามารถดูข้อมูลสารสนเทศในลักษณะตารางสรุปข้อมูลสถิติผู้บาดเจ็บ และเสียชีวิต

1.4.2.3 สามารถดูข้อมูลสารสนเทศในลักษณะรูปแบบของแผนภูมิชนิดต่าง ๆ จาก โปรแกรม Tableau ได้

1.4.2.4 สามารถกรองข้อมูลสารสนเทศผ่านระบบของ Tableau บน Web browser ได้

1.5 เครื่องมือที่ใช้ในการพัฒนาโปรแกรม

1.5.1 Hardware

1.5.1.1 Notebook: ASUS K556U core i7 Ram 8 GB

1.5.1.2 Notebook: LENOVO G470 core i3 Ram 12 GB

1.5.2 Software

1.5.2.1 โปรแกรม Tableau Public ใช้ในการแสดงผลข้อมูลในรูปแบบของภาพ

1.5.2.2 โปรแกรม Weka 3.8.4 ใช้ในการวิเคราะห์ข้อมูลด้วยเทคนิคเหมืองข้อมูล

1.5.2.3 โปรแกรม RapidMiner Studio 9.5.1 ใช้ในการทำเทคนิคเหมืองข้อมูล

1.5.2.4 ชุดคำสั่งภาษา HTML4, HTML5 (Hyper Text Markup Language)

1.5.2.5 ชุดคำสั่งภาษา PHP (PHP Hypertext Preprocessor)

1.5.2.6 ชุดคำสั่ง CSS และ Bootstrap ที่ใช้ในการออกแบบเว็บ

1.5.2.7 โปรแกรม Xampp 5.5.38 ใช้ในการจำลอง web server

1.5.2.8 โปรแกรม FileZilla 3.40.0 โปรแกรมที่ใช้ติดต่อกับ FTP Server เพื่ออัปโหลด (Upload) ไฟล์ (File)

1.5.2.9 โปรแกรม Atom ใช้ในการเขียนคำสั่งและเขียนโปรแกรม

1.5.2.10 โปรแกรม Adobe XD ใช้ในการออกแบบหน้าเว็บไซต์

1.5.2.11 ระบบปฏิบัติการ : Window 10 ใช้ในการเปิดซอฟต์แวร์ต่าง ๆ

1.5.2.12 Microsoft office Word 2016 ใช้ในการทำเอกสารต่าง ๆ

1.5.2.13 Microsoft office Excel 2016 ใช้ในการ Clean Data กับชุดข้อมูล

1.6 สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล

1.6.1 สถานที่ใช้ในการดำเนินการศึกษา

มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพเชียงใหม่ 128 ถนนห้วยแก้ว ตำบลช้างเผือก อำเภอเมือง จังหวัดเชียงใหม่ 50300

1.6.2 แหล่งรวบรวมข้อมูล

เว็บไซต์ Open Government Data of Thailand หรือ data.go.th

1.7 ระยะเวลาในการดำเนินการ

ตารางที่ 1.1 ระยะเวลาในการดำเนินการ

แผนการดำเนินการ	2563				
	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.
1. ศึกษาและกำหนดความต้องการ	←→				
2. ตรวจสอบความถูกต้องข้อมูล		←→			
3. วิเคราะห์ข้อมูลจำแนกตามกลุ่ม		←→			
4. ออกแบบระบบ			←→		
5. เขียนและทดสอบระบบ			←→	←→	
6. ตรวจสอบระบบโดยรวม				←→	
7. ประเมินการใช้งานระบบ				←→	←→
8. จัดทำเอกสารประกอบ โครงการ	←→				←→

1.8 นิยามศัพท์เฉพาะ

1.8.1 การวิเคราะห์ข้อมูล หมายความว่า เป็นการจัดระเบียบแยกแยะส่วนต่าง ๆ ของหลักฐาน หรือข้อมูลที่ได้ออกเป็นหมวดหมู่ เพื่อหาคำตอบตามความมุ่งหมาย และตามสมมติฐานที่ได้กำหนดไว้ การวิเคราะห์ข้อมูลนี้เป็นขั้นตอนการทำงานที่ต่อเนื่องมาจากการวัด การนับ และจัดเรียงลำดับข้อมูล ส่วนใหญ่เป็นเรื่องเกี่ยวกับการนำเอาวิธีการทางสถิติมาวิเคราะห์หาค่าตัวแปรหรือหาลักษณะของตัวแปร

1.8.2 เทศกาล หมายถึง เหตุการณ์ชนิดหนึ่ง ซึ่งตามปกติธรรมดาจัดตั้งขึ้นโดยชุมชนท้องถิ่น ที่มุ่งความสนใจและเฉลิมฉลองเอกลักษณ์บางอย่างของชุมชนนั้นและเทศกาลนั้น เทศกาลมักเกี่ยวข้องกับประเพณี ความเชื่อ หรือศาสนาของชุมชน เทศกาลส่วนใหญ่มักจัดขึ้นปีละครั้ง คือการถือเอาวันครบรอบปีของเทศกาลครั้งก่อนมาตั้งเป็นครั้งถัดไป

1.8.3 เว็บไซต์ หมายความว่า สื่อนำเสนอข้อมูลบนเครื่องคอมพิวเตอร์ หรือคือการรวบรวม หน้าเว็บเพจหลายหน้า ซึ่งเชื่อมโยงกันผ่านทางไฮเปอร์ลิงก์ ซึ่งต้องเปิดด้วยโปรแกรมเฉพาะทางที่เรียกว่า Web Browser โดยถูกจัดเก็บไว้ในเว็ลด์ไวด์เว็บ และเว็บไซต์นั้นถูกสร้างขึ้นด้วยภาษาทางคอมพิวเตอร์ที่เรียกว่า HTML (Hyper Text Markup Language)

1.9 บทสรุป

จากบทนำที่ได้กล่าวมาในข้างต้นทั้งหมดนั้น คณะผู้จัดทำได้เล็งเห็นความสำคัญของการนำเทคโนโลยีมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลผู้บาดเจ็บและเสียชีวิตในช่วงเทศกาลปีใหม่ ประจำปี 2551 – 2558 ด้วยกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM จากเทคนิคการทำเหมืองข้อมูลแบบ Classification ในรูปแบบของแผนภูมิต้นไม้ หรือที่เรียกว่า Decision tree และเผยแพร่ข้อมูลสารสนเทศบน web browser ที่เป็นที่ยอมรับในยุคอินเทอร์เน็ตคือการเผยแพร่ทางสื่อออนไลน์ โดยใช้ภาษา HTML ในการพัฒนาเว็บไซต์สำหรับการเปิดเผยข้อมูลภาครัฐที่ได้มาจาก data.go.th หรือใน Government Open Data of Thailand เนื่องจากเทคโนโลยีปัจจุบันเข้ามาช่วยในการจัดการข้อมูล หรือดำเนินงานต่าง ๆ ให้มีความสะดวกสบายเป็นอย่างมาก อีกทั้งผู้ใช้งานยังเข้าถึงข้อมูลได้ง่าย และมีประสิทธิภาพ