

บทที่ 5

สรุปการวิจัย และข้อเสนอแนะ

การศึกษาวิจัยเรื่อง การประยุกต์ระบบ GPS และการทำเหมืองข้อมูลเพื่อทำนาย และป้องกันการเกิดอุบัติเหตุของรถโดยสารสาธารณะบริการระหว่างจังหวัด ผู้วิจัยได้พัฒนากล้องข้อมูลระบบ GPS โดยใช้หลักการ Big Data รองรับข้อมูลขนาดใหญ่ เพิ่มประสิทธิภาพการสืบค้น การวิเคราะห์ข้อมูลที่รวดเร็ว และการทำเหมืองข้อมูลการเกิดอุบัติเหตุรถโดยสารสาธารณะเพื่อสร้างแบบจำลองทำนาย ได้ข้อสรุปผลการวิจัยการอภิปรายผล และข้อเสนอแนะดังนี้

1 สรุปผลการวิจัย

1.1 สร้างระบบจัดการข้อมูล GPS

ผู้วิจัยนำข้อมูลติดตามรถโดยสารสาธารณะระบบ GPS มาสร้างคลังข้อมูล โดยใช้โปรแกรม MONGODB จัดการฐานข้อมูลให้สามารถรองรับกับข้อมูลขนาดใหญ่ และผ่านกระบวนการอีทีแอล (EXTRACT TRANSFORM AND LOAD : ETL) ซึ่งผู้วิจัยได้ออกแบบส่วนของการนำเข้าข้อมูลอยู่ในรูปแบบของเท็กซ์ไฟล์ (CSV) ทำการคัดเลือก คัดแยกข้อมูลที่ไม่เกี่ยวข้อง ให้ได้รูปแบบข้อมูลถูกต้อง แล้วใช้คำสั่ง COMMAND LINE ของโปรแกรม MongoDB เป็นเครื่องมือนำเข้าคลังข้อมูล GPS

1.2 สร้างแบบจำลองการทำนาย

ข้อมูลที่ใช้ในการทำเหมืองข้อมูลได้มาจากคลังข้อมูลระบบ GPS และคลังข้อมูลการเกิดอุบัติเหตุรถโดยสารสาธารณะที่สร้างขึ้นมา ซึ่งใช้เทคนิคการทำเหมืองข้อมูลมาสร้างระบบการวิเคราะห์ใช้อัตราความเร็วของรถโดยสารสาธารณะ และทำนายการเกิดอุบัติเหตุรถโดยสารสาธารณะ ส่วนขั้นตอนการทำเหมืองข้อมูลใช้รูปแบบจำลองของ CRISP - DM Model โดยใช้ข้อมูลติดตามรถโดยสารสาธารณะ ข้อมูลสถานะของรถโดยสารสาธารณะ ข้อมูลประวัติรถโดยสาร ข้อมูลสถานะของรถโดยสาร ข้อมูลการจัดกลุ่มรถโดยสาร ข้อมูลการประเภทรถโดยสาร และข้อมูลตำแหน่งที่ตั้งต่างๆกำหนดความสัมพันธ์ระหว่างข้อมูลเพื่อนำเข้าคลังข้อมูลระบบ GPS และข้อมูลการเกิดอุบัติเหตุรถโดยสารสาธารณะระหว่างพ.ศ. 2556-2559 จำนวน 406 ระเบียบ ประกอบด้วย

ข้อมูลการเกิดอุบัติเหตุโดยสาธารณะ ข้อมูลความเสียหายของรถยนต์คู่กรณี ข้อมูลรายละเอียดรถคู่กรณี ข้อมูลประวัติพนักงาน ข้อมูลความเสียหายของบริษัท ข้อมูลขั้นตอนทำงาน และรายจ่าย ประวัติดูโดยสาธารณะ ข้อมูลประเภทรถโดยสาธารณะ กำหนดความสัมพันธ์ระหว่างข้อมูลเพื่อนำเข้าคลังข้อมูลการเกิดอุบัติเหตุโดยสาธารณะ

การสร้างตัวแบบสำหรับการวิเคราะห์ใช้อัตราความเร็วของรถโดยสาธารณะ และตัวแบบจำลองทำนายการเกิดอุบัติเหตุโดยสาธารณะ อธิบายได้ดังนี้

1.2.1 แบบจำลองสำหรับการวิเคราะห์ใช้อัตราความเร็วของรถโดยสาธารณะ

สามารถออกแบบจำลองการวิเคราะห์อัตราความเร็วรถโดยสาธารณะได้ 2 แบบ ได้แก่ การใช้ความเร็วรถของกลุ่มรถยนต์ และการใช้ความเร็วรถของตำแหน่งที่ตั้ง ซึ่งใช้เทคนิคเรียนรู้ต้นไม้ตัดสินใจ Decision Tree Learner ด้วยการกำหนดเป้าหมาย (TARGET) ของฟิลด์หมายเลขรถ (TRUCK_NAME) และระบุเงื่อนไขที่เพื่อการวิเคราะห์ใน INCLUDE BLOCK ของ Decision Tree Learner Node เช่น ฟิลด์ความเร็ว (R_SPEED) เพื่อทำนายแบบจำลองการวิเคราะห์ความเร็วรถของกลุ่มรถยนต์ หรือฟิลด์ความเร็ว (R_SPEED) กับฟิลด์ตำแหน่งที่ตั้ง (TAM_TNAME) เพื่อทำนายแบบจำลองวิเคราะห์การใช้ความเร็วรถของตำแหน่งที่ตั้ง และกำหนดคุณลักษณะการวัดค่าด้วย Quality Measure (QM) ในรูป Gain / Gain Ratio ผลการวิจัยพบว่าสามารถจำแนกกลุ่มรถโดยสาธารณะ และกลุ่มตำแหน่งที่ตั้งในรูปแบบภาพต้นไม้ตัดสินใจ ซึ่งสามารถอธิบายผลลัพธ์ที่ได้ เช่น การเปรียบเทียบร้อยละ กับจำนวนครั้งทั้งหมด เช่น ข้อมูลหมายเลขรถ 42 (158) ใช้ความเร็วมากกว่าหรือเท่ากับ 90 กม.ต่อชั่วโมง มีจำนวน 999 ครั้ง คิดเป็นร้อยละ 10.7 ของจำนวนครั้งรวม (TOTAL) 9379 ครั้ง และอีกหนึ่งตัวอย่าง เช่น หมายเลข 36 (36) ใช้อัตราความเร็วรถผ่านสถานที่บางไฟจำนวน 5 ครั้งคิดเป็นร้อยละ 50 ของจำนวนทั้งหมด 10 ครั้ง ที่ความเร็วมากกว่าหรือเท่ากับ 93 กม.ต่อชั่วโมง

1.2.2 แบบจำลองทำนายการเกิดอุบัติเหตุโดยสาธารณะ ใช้เทคนิคการจัด

กลุ่ม (Clustering) ภูมิความสัมพันธ์ (Association) และ การจำแนก (Classification) โดยใช้ข้อมูลการเกิดอุบัติเหตุโดยสาธารณะ ซึ่งกระบวนการออกแบบจำลองเพื่อการทำนายต้องการจำแนกกลุ่มออกเป็นสองกลุ่ม คือ กลุ่มเสี่ยง (เป็นบุคคลที่ถูกจัดให้มีอัตราการก่อให้เกิดอุบัติเหตุการขับรถโดยสาธารณะมากที่สุด) และกลุ่มเฝ้าระวัง (เป็นบุคคลที่ถูกจัดว่าอาจจะก่อให้เกิดอุบัติเหตุการขับรถโดยสาธารณะ) แบบจำลองเพื่อการทำนายของกลุ่มเสี่ยง มี 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การแบ่งกลุ่มเสี่ยง โดยใช้ข้อมูลในส่วนของคุณสมบัติบุคคล ได้แก่ อายุ การศึกษา ภูมิลำเนา น้ำหนักตัว ส่วนสูง ประสบการณ์ทำงาน รายได้ และแบ่งข้อมูลด้วยเทคนิค Split Test ได้แก่ Training Data และ Testing Data ทำการทดสอบแบบจำลองด้วยอัลกอริทึม Simple

K-Means เพื่อทำการแบ่งกลุ่มเสี่ยง ทำให้ได้ผลลัพธ์ของข้อมูลเหมือนกัน และได้กลุ่มเสี่ยงเป็นจำนวน 4 กลุ่ม ที่มีลักษณะเฉพาะกลุ่มเสี่ยงชัดเจน

ขั้นตอนที่ 2 นำกลุ่มเสี่ยงแต่ละกลุ่มมาทำการทดสอบร่วมกับข้อมูลองค์ประกอบของปัจจัยการเกิดอุบัติเหตุ ได้แก่ เส้นทางเดินรถ จำนวนชั่วโมงขับรถ ประเภทรถที่ใช้ และสถานที่เกิดเหตุ (บริษัท, สถานี, ภายนอก) ด้วยอัลกอริทึม Decision Tree Learner เพื่อหาองค์ประกอบปัจจัยการเกิดอุบัติเหตุในแต่ละกลุ่มเสี่ยงด้วย Decision Tree to Ruleset Node ให้ได้ผลลัพธ์การทำนายสามารถนำไปพัฒนาเป็นระบบการทำนายผลการเกิดอุบัติเหตุรถโดยสารสาธารณะ ดังตารางที่ 5.1

ตารางที่ 5.1 สรุปผลการทดสอบของกลุ่มเสี่ยง

กลุ่มเสี่ยง	ลักษณะเฉพาะกลุ่มบุคคล	การเกิดเหตุ	เกิดเหตุร้อยละ
กลุ่ม 1	ชาย -> กรุงเทพฯปริมณฑล ->	- ขับรถมากกว่า 8 ชั่วโมงต่อวัน	90.5
	อายุ 51-55 -> ประสบการณ์มากกว่า 7 ปี -	- ขับรถต่อวัน 7-8 ชั่วโมง รถตู้ V9905	50
	> การศึกษา(ประถม) -> รายได้เฉลี่ยต่อวัน 300-500 บาท	- ขับรถต่อวันน้อยกว่า 6 ชั่วโมง รถ BENZ	75
		เกิดเหตุที่สถานี	
กลุ่ม 2	ชาย -> ภาคอีสาน -> อายุ 30-35 ปี ->	- ขับรถ 7-8 ชั่วโมงต่อวัน	74-75
	ประสบการณ์ 1-2 ปี -> การศึกษา	- BENZ, GOLDEN DAGON, SUNLONG	73 - 83
	(ประถม) -> รายได้ต่อวัน 500-700 บาท	- สถานี, ภายนอก	74 - 75
กลุ่ม 3	ชาย -> ภาคกลาง -> อายุ 46--50 ปี ->	-เส้นทาง B389 เกิดเหตุภายในสถานี	100
	ประสบการณ์มากกว่า 7 ปี -> การศึกษา	- เส้นทาง B048 เกิดเหตุภายในสถานี	77.8
	ประถม -> รายได้ต่อวันมากกว่า 800 บาท	- ขับรถมากกว่า 8 ชั่วโมงต่อวันเส้นทาง B389	85
		- ขับรถมากกว่า 8 ชั่วโมงต่อวันเส้นทาง B048	58.5
กลุ่ม 4	ชาย -> ภาคตะวันออก -> อายุ 41-45 ->	- ขับรถน้อยกว่า 6 ชั่วโมงต่อวัน รถ BENZ	63.6
	ประสบการณ์น้อยกว่า 4 ปี -> การศึกษา	- รถ BENZ เกิดเหตุภายในสถานี	68.7
	(ประถม) -> รายได้เฉลี่ยต่อวันมากกว่า 800 บาท		

1.2.3 แบบจำลองเพื่อการทำนายของกลุ่มเฝ้าระวัง เป็นการนำข้อมูลคุณสมบัติบุคคล ได้แก่ อายุ การศึกษา ภูมิลำเนา น้ำหนักตัว ส่วนสูง ประสบการณ์ทำงาน รายได้ มาทดสอบแบบจำลอง โดยใช้กฎความสัมพันธ์ (Association Rule) ด้วยวิธี Apriori เพื่อหาความสัมพันธ์ของคุณสมบัติบุคคล โดยกำหนดค่าความเชื่อมั่น Confidence ไม่ต่ำกว่าร้อยละ 60 และแสดงผลลัพธ์ของกฎ 50 จำนวน เพื่อให้ได้คุณลักษณะบุคคลที่แสดงค่าความเชื่อมั่น (Confidence) ของกฎความสัมพันธ์และค่าลิฟต์ (Lift) คือค่าสหสัมพันธ์ที่บ่งบอกถึงกฎความสัมพันธ์ที่มีค่ามากกว่า 1 เท่านั้น ทั้งนี้การกำหนดค่า Confidence ให้มีค่าระดับกลางๆ เพื่อให้ได้ผลลัพธ์ของข้อมูลจากกฎความสัมพันธ์อยู่ในระดับข้อมูลกลุ่มเฝ้าระวังและไม่มีข้อมูลถูกจัดไว้ในกลุ่มเสี่ยง ส่วนการกำหนดค่า Confidence ให้มีค่ามากหรือร้อยละ 80 ขึ้นไปผลลัพธ์ที่ได้มีจำนวนข้อมูลน้อย ซึ่งไม่ตรงกับความต้องการ จากผลการวิจัยพบว่า การหาคุณลักษณะบุคคลกลุ่มเฝ้าระวังด้วยกฎความสัมพันธ์ของอัลกอริทึม Apriori กำหนดค่าความเชื่อมั่น Confidence ไว้ที่ร้อยละ 60 และสามารถคัดเลือกค่าที่เหมาะสมมาใช้ในการพัฒนาโปรแกรมใช้งาน ดังตารางที่ 5.2

ตารางที่ 5.2 คุณลักษณะบุคคลกลุ่มเฝ้าระวัง

ลำดับ	ความสัมพันธ์นอกกลุ่มกับอุบัติเหตุ	Confidence	Lift
1	ภาคกลาง -> การศึกษาประถม	89	1.17
2	ประสบการณ์น้อย(1-2ปี) -> การศึกษาประถม	89	1.04
3	รายได้ต่อวัน(700-800บาท) ->การศึกษาประถม	78	1.03
4	อายุ 30-35 ปี -> น้ำหนักตัว 50 - 65 ก.ก.	72	1.15
5	ส่วนสูง 161-170 ซม. -> น้ำหนักตัว 50- 65 ก.ก.	72	1.15
6	ประสบการณ์ไม่ถึง 1 ปี -> ส่วนสูง161-170 ซม.	71	1.46
7	อายุ 45 - 50 ปี	66	1.05
8	การศึกษามัธยมต้น	66	1.04
9	รายได้เฉลี่ยต่อวัน 700 - 800บาท	66	1.04
10	ประสบการณ์น้อยมากไม่ถึง 1 ปี	66	1.04

1.3 สร้างโปรแกรมใช้งาน

การพัฒนาระบบทำนายการเกิดอุบัติเหตุของรถโดยสารสาธารณะ ผู้วิจัยใช้โปรแกรม Microsoft Visual FoxPro v9.0 ในการพัฒนาเพื่อความสามารถในสร้างและออกแบบการใช้งาน จากผลการประเมินความเชื่อมั่นของการทำนาย โดยผู้ใช้งานจำนวน 7 คน พบว่า ผลการประเมินความเชื่อมั่นโดยรวมมีค่าเฉลี่ยเท่ากับ 4.84 ถือว่าการทำนายการเกิดอุบัติเหตุของรถโดยสารสาธารณะ โดยรวมอยู่ในระดับดีมาก สามารถนำไปใช้เป็นเครื่องมือในการพิจารณาคัดเลือกพนักงานขับรถโดยสาร เพื่อลดความเสี่ยงต่อการเกิดอุบัติเหตุทางรถโดยสารสาธารณะ

2 การอภิปรายผล

2.1 สร้างระบบจัดการข้อมูล GPS

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาล้างข้อมูลระบบ GPS ซึ่งในการพัฒนาล้างข้อมูลใช้ซอฟต์แวร์มอโกดีบี(MongoDB 2.6.7) มารองรับข้อมูลขนาดใหญ่ตามหลักการของ Big data จากแหล่งข้อมูลการติดตามรถโดยสารสาธารณะระบบ GPS เดิม ผู้วิจัยทำการรวบรวมข้อมูล ออกแบบคลังข้อมูล และสร้างกระบวนการ ETL ในการดึงข้อมูลเข้าสู่คลังข้อมูลที่พัฒนาขึ้นใหม่นั้น จำเป็นต้องใช้เครื่องมือติดต่อกับฐานข้อมูลเดิม และเปลี่ยนรูปแบบข้อมูลเป็นเท็กซ์ไฟล์ (CSV) ซึ่งในงานวิจัยใช้ด้วยโปรแกรม KNIME v2.12 เป็นตัวจัดการ ส่วนวิธีนำข้อมูลเข้าคลังข้อมูลที่พัฒนา จะใช้ฟังก์ชันของซอฟต์แวร์มอโกดีบี (MongoDB v2.6.7) ผ่านคำสั่ง Command Line ผลที่ได้จากการพัฒนาล้างข้อมูล GPS มีประสิทธิภาพด้านการประมวลผลข้อมูล การสืบค้นข้อมูลที่รวดเร็ว และสามารถพัฒนาระบบสนับสนุนด้านการประมวลผลเชิงวิเคราะห์ การทำเหมืองข้อมูล และสร้างรายงาน โปรแกรม KNIME v2.12 เป็นเครื่องมือที่ช่วยออกแบบสำหรับกลุ่มผู้ใช้งานในระดับผู้บริหารที่ต้องใช้ข้อมูลเพื่อประกอบการตัดสินใจ ซึ่งสอดคล้องกับงานวิจัยของ ประกายมาศ ศรีสุขทักษิณ และคณะ(2557) ที่ทำการทดสอบประสิทธิภาพของฐานข้อมูลระหว่างฐานข้อมูลเชิงสัมพันธ์ MySQL กับฐานข้อมูลไม่สัมพันธ์ MongoDB ซึ่งผลของการวิจัยพบว่าปริมาณข้อมูลที่นำมาทดสอบมีขนาดใหญ่ ฐานข้อมูลไม่สัมพันธ์ ของ MongoDB สามารถประมวลผลข้อมูล เช่น การค้นหา การอ่าน การบันทึก และการลบข้อมูลได้เร็วกว่าหลายเท่ามาก เมื่อเทียบกับฐานข้อมูลเชิงสัมพันธ์ MySQL ซึ่งสอดคล้องกับงานวิจัย

2.2 สร้างแบบจำลองการทำนาย

งานวิจัยนี้ นำเสนอการใช้เทคนิคเหมืองข้อมูลในการสร้างแบบจำลอง Connolly&E.Begg (2002: 1232) กล่าวถึง "กระบวนการที่นำข้อมูลที่มีอยู่ในฐานข้อมูลขนาดใหญ่ มาทำการศึกษา วิเคราะห์ ทำความเข้าใจ นำผลลัพธ์ที่ได้จากการศึกษามาใช้ในการตัดสินใจทาง ธุรกิจ และการนำเทคนิคเหมืองข้อมูล เพื่อออกแบบจำลองจากข้อมูล" ที่แนะนำถึงเทคนิคและ ขั้นตอนการทำเหมืองข้อมูล ซึ่งสอดคล้องต่อการสร้างแบบจำลองเพื่อทำนายการเกิดอุบัติเหตุ โดยสารสนเทศ ของกลุ่มเสี่ยง และกลุ่มเฝ้าระวัง ด้วยการค้นหาคุณลักษณะบุคคลและปัจจัยการ เกิดอุบัติเหตุโดยสารร่วมกับเทคนิคการทำเหมืองข้อมูล เพื่อใช้เป็นข้อมูลประกอบการตัดสินใจคัด กรองบุคคลเพื่อเลือกพนักงานขับรถ ซึ่งผลการวิจัยสอดคล้องกับงานวิจัยของ พล โกสุววัฒนะ (2556: 81-82) สร้างแบบจำลองการแบ่งกลุ่มนักศึกษา เพื่อเห็นการจัดข้อมูลที่มีลักษณะประจำ เช่น อายุ เพศ ภูมิฐานะ ที่กระจกระบายให้อยู่เป็นกลุ่ม พบว่าการจัดกลุ่มนักศึกษาใช้เทคนิคเหมือง ข้อมูลแบบคลัสเตอร์ (Clustering) ด้วยวิธี Simple K-Means สามารถหากลุ่มที่มีความแตกต่าง และ จุดเด่นแต่ละกลุ่มได้อย่างชัดเจน และเสกสรรค์ วิสัยลักษณ์ (2558: 78-89) การสร้างแบบพยากรณ์ ผลการเรียนรู้ระดับมัธยมศึกษา การจัดกลุ่มข้อมูลนักศึกษาโดยใช้คุณลักษณะของนักศึกษาที่เป็น ข้อมูลประวัติส่วนตัวมาจัดแบ่งกลุ่ม 2 แบบ คือ ไม่มีการคัดเลือกคุณลักษณะ และวิธีการคัดเลือก คุณลักษณะ ด้วยอัลกอริทึม K-Means ผลที่ได้พบว่าแบบไม่จัดกลุ่มนำมาคัดเลือกคุณลักษณะ ร่วมกับโครงข่ายประสาทเทียมแบบมัลติเลเยอร์เพอร์เซ็ปตรอน ให้ค่าความถูกต้องร้อยละ 94.48 ซึ่ง นักเรียนสามารถนำผลการพยากรณ์ไปประกอบการตัดสินใจและวางแผนการเรียน และไพฑูรย์ งามจันทร์ (2550: 36-45) การพัฒนาระบบสนับสนุนการตัดสินใจเลือกสาขาการเรียนของ นักศึกษาระดับปริญญาตรี ใช้เทคนิคต้นไม้ตัดสินใจด้วยอัลกอริทึม C5.0 ในการสร้างตัวแบบ พบว่าผลลัพธ์ผลลัพธ์แสดงภาพออกมาเป็นต้นไม้ตัดสินใจ และสามารถจำแนกกฎ (Rule set) มีค่า การทำนายถูกต้องเท่ากับร้อยละ 76.84 นำตัวแบบที่ได้ไปพัฒนาระบบเลือกแผนการเรียนระดับ มัธยมศึกษาตอนปลาย ใช้เป็นเครื่องมือช่วยตัดสินใจเลือกแผนการเรียนได้ และ พิจิตรา จอมศรี (2549: 39-41) การทำนายเนื้อหาของเว็บโดยใช้เหมืองข้อมูล ได้ใช้เทคนิคการค้นหากฎความสัมพันธ์ ด้วยวิธี Apriori พบว่าสามารถจำแนกกฎความสัมพันธ์เนื้อหาของเว็บ และแสดงผลในรูปการจำแนกกฎ ความสัมพันธ์ ระบุถึงค่าความเชื่อมั่นและค่าสนับสนุน และให้ค่าความถูกต้องของตัวแบบข้อมูลเท่ากับ ร้อยละ 66.67 ดังนั้นสามารถนำแบบจำลองที่ได้ไปใช้ในการทำนายเนื้อหาเว็บได้ ซึ่งสอดคล้องกับแนว ทางการวิจัยการหาคุณลักษณะบุคคลที่ไม่ได้จัดอยู่ในกลุ่ม

นอกจากงานวิจัยได้นำเทคนิคการทำเหมืองข้อมูลมาวิเคราะห์ข้อมูลการเกิดอุบัติเหตุ โดยสารของพนักงานขับรถ เพื่อผลลัพธ์การทำนายจากตัวแบบที่สร้างได้ให้ค่าความสัมพันธ์ของ

กลุ่มกับการเกิดอุบัติเหตุอยู่ในเกณฑ์สูง การประเมินผลจากผู้ใช้งานอยู่ในเกณฑ์ดีมาก การนำผลการทำนายไปใช้ประกอบการตัดสินใจคัดกรองบุคคลเพื่อเลือกพนักงานขับรถ

3 ปัญหาและอุปสรรค

3.1 ข้อมูลไม่ได้ถูกออกแบบเพื่อใช้ในการทำเหมืองข้อมูล ความเหมาะสมหรือความสมบูรณ์จึงไม่เพียงพอ ต้องมีการแทนค่า หรือปรับเปลี่ยนค่าต่างๆของข้อมูลให้เหมาะสม ซึ่งอาจมีผลต่อการวิเคราะห์ข้อมูลผิดพลาดได้

3.2 การนำข้อมูลจากโปรแกรมระบบ GPS ใช้เวลานาน เนื่องจากข้อมูลมีขนาดใหญ่ การจัดเก็บข้อมูลติดตามรถยนต์ถูกแบ่งตารางเป็นรายคัน ความซับซ้อนของข้อมูลจึงเพิ่มมากขึ้น

3.3 ข้อมูลบางส่วนไม่สามารถเปิดเผยได้ ซึ่งจะกระทบต่อธุรกิจของผู้ประกอบการเดินรถ จึงต้องใช้ข้อความอื่นแทนในการสื่อถึงความหมายที่ต้องการอ้างอิง

3.4 ข้อมูลบางส่วนไม่เพียงพอต่อความต้องการ และขาดหายไปจากการระบุข้อความของข้อมูลที่ไม่มีสมบูรณ์ หรือไม่ครบถ้วน เช่น ภูมิฐานะ อายุ ส่วนสูง รายได้ และน้ำหนัก เป็นต้น

3.5 ข้อมูลที่ใช้ในการทำตัวแบบทำนายการเกิดอุบัติเหตุโดยสาธารณะ การอ้างอิงข้อมูลมาจากผู้ประกอบการเดินรถแหล่งเดียว และข้อมูลมีจำนวน 406 ระเบียบ

3.6 ความร่วมมือจากหลายหน่วยงานในการใช้ข้อมูล มีข้อจำกัด กฎระเบียบ และไม่สามารถเผยแพร่ได้ซึ่งอาจจะทำให้แบบจำลองมีผลคลาดเคลื่อนกับความเป็นจริงได้

3.7 โปรแกรม KNIME เป็นเครื่องมือที่ใช้ในการวิเคราะห์ และประมวลผลข้อมูล มีความสามารถในการรองรับข้อมูลที่หลากหลายรูปแบบ ฟังก์ชันที่ถูกออกแบบช่วยในการวิเคราะห์ข้อมูลมีหลายชนิด และความหลากหลายของเครื่องอื่นๆ ผู้วิจัยมีเวลาจำกัดในการเรียนรู้ ซึ่งอาจทำให้เกิดข้อผิดพลาดได้

4 ข้อเสนอแนะ

ในการวิจัยนี้ได้พัฒนาค้างข้อมูลระบบ GPS โดยใช้โปรแกรม MongoDB v2.6.7 เพื่อรองรับข้อมูลติดตามรถโดยสารสาธารณะระบบ GPS ผู้สนใจสามารถนำหลักการพัฒนาค้างข้อมูลนี้ไปพัฒนาโดยใช้โปรแกรมอื่น เช่น Hadoop, CouchDB, HBase และ Cassandra เป็นต้นเพื่อเปรียบเทียบประสิทธิภาพและการประมวลผลเชิงวิเคราะห์จากข้อมูล Big Data

และการสร้างแบบจำลองทำนายการเกิดอุบัติเหตุรถโดยสารสาธารณะ ความสมบูรณ์ของแบบจำลองขึ้นอยู่กับข้อมูลที่นำมาใช้สร้างแบบจำลอง ควรมีข้อมูลเพิ่มเติมและเพิ่มข้อมูลการเกิดอุบัติเหตุรถโดยสารของผู้ประกอบการเดินรถเส้นทางอื่นๆ จะส่งผลให้เกิดความถูกต้องการทำนายดีมากยิ่งขึ้น

