บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

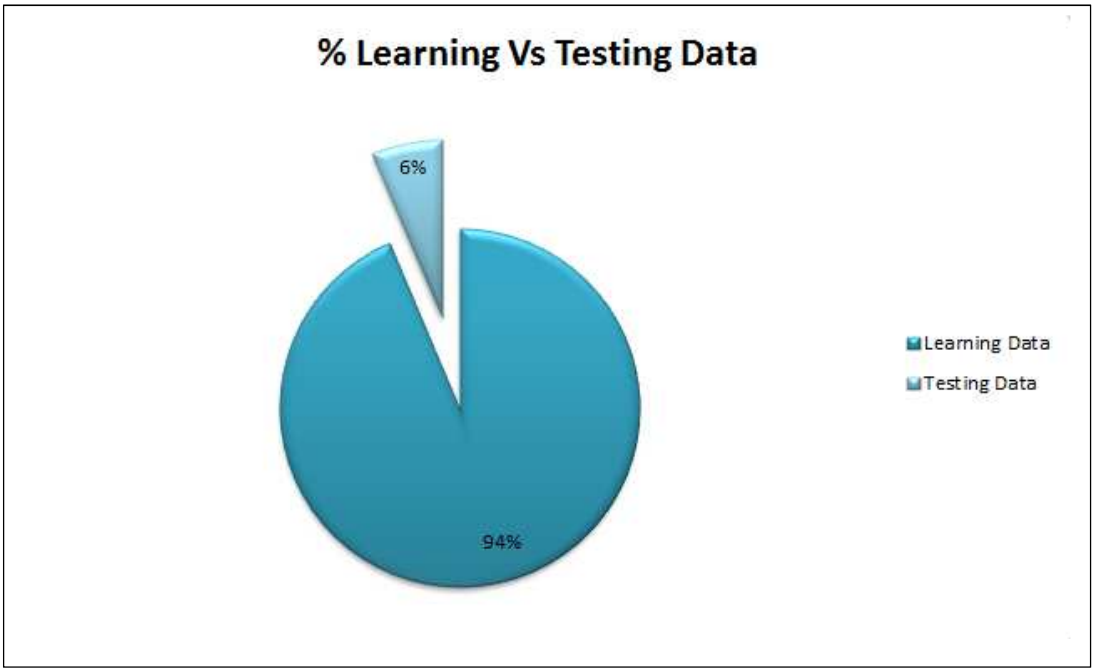
5.1 สรุปผลการวิจัย

จากปัญหาการขาดการเชื่อมโยงของการปฏิบัติงานเกี่ยวกับงานทบทวนวงเงิน ระหว่างผู้ปฏิบัติงาน 2 ฝ่าย คือ เจ้าหน้าที่สินเชื่อ ซึ่งมีบทบาทหน้าที่ในปฏิบัติงานเป็นด่านหน้าของหน่วยธุรกิจของธนาคาร คือ มีหน้าที่รับผิดชอบในการให้บริการและติดตามดูแลลูกค้าอย่างใกล้ชิด กับอีกฝ่ายหนึ่ง คือ เจ้าหน้าที่ทบทวนวงเงิน ซึ่งมีบทบาทหน้าที่เป็นผู้ปฏิบัติงานเบื้องหลัง คือ เมื่อลูกค้าได้รับอนุมัติสินเชื่อเงินกู้เบิกเกินบัญชีเรียบร้อยแล้ว เจ้าหน้าที่ทบทวนวงเงิน มีหน้าที่รับผิดชอบในการพิจารณาทบทวนวงเงินของบัญชีเงินกู้เบิกเกินบัญชีที่ใกล้ถึงวันครบกำหนดอายุสัญญา หากพิจารณาแล้วเห็นว่าเป็นลูกค้าดีและอยู่ในเกณฑ์ที่กำหนด ก็จะทำการอนุมัติให้ขยายวันครบกำหนดอายุสัญญาออกไป ซึ่งลูกค้าจะสามารถใช้วงเงินนั้นต่อไปได้ โดยไม่จำเป็นต้องยื่นขอสินเชื่อใหม่อีกครั้ง

การแบ่งหน้าที่ความรับผิดชอบที่แยกขาดจากกันเช่นนี้ ทำให้เกิดอุปสรรคในการทำงานระหว่างเจ้าหน้าที่สินเชื่อและเจ้าหน้าที่ทบทวนวงเงิน ดังนั้น เพื่อเป็นการแก้ไขปัญหาที่เกิดขึ้น ผู้วิจัยจึงแนวคิดที่จะนำความรู้ด้านการทำเหมืองข้อมูลมาประยุกต์ใช้กับงานทบทวนวงเงิน พร้อมกับทำการพัฒนาโปรแกรมพยากรณ์โอกาสที่บัญชีเงินกู้เบิกเกินบัญชีจะได้รับการทบทวนวงเงิน ที่มีลักษณะการทำงานแบบ Web Application โดยใช้ภาษา PHP และกำหนดให้เรียกใช้รูปแบบโมเดลการตัดสินใจโดยตรงจากโปรแกรม WEKA พร้อมกับมีการประเมินผลเชิงปริมาณจากการทดสอบข้อมูลผลการพิจารณาทบทวนวงเงินจริง มาเปรียบเทียบกับผลการพยากรณ์ที่ได้จากโปรแกรมพยากรณ์ฯ ว่ามีความถูกต้องน่าเชื่อถือ และสอดคล้องกับความเป็นจริงทางธุรกิจหรือไม่ เพียงใด และมีการประเมินผลเชิงคุณภาพจากความคิดเห็นของผู้ใช้โปรแกรมฯ ซึ่งเป็นผู้เชี่ยวชาญด้านสินเชื่อ

งานวิจัยนี้เป็นการประยุกต์ใช้โมเดลที่เกิดการเรียนรู้รูปแบบการตัดสินใจของโปรแกรม WEKA เวอร์ชั่น 3.6.2 โดยทำการเรียนรู้รูปแบบการตัดสินใจตามเงื่อนไขของหลักการของงานทบทวนวงเงิน หรือที่เรียกว่างาน Auto review ด้วยวิธีการจำแนกข้อมูล จากข้อมูลบัญชีเงินกู้เบิกเกินบัญชีที่เคยเข้าสู่กระบวนการทบทวนวงเงินล่วงหน้า ระหว่างเดือนมกราคม 2556 ถึงเดือน

กันยายน 2557 และเป็นข้อมูลของลูกค้าที่อยู่ในความดูแลของฝ่ายพิจารณาสินเชื่อธุรกิจ SME ของธนาคารกรุงศรีอยุธยา จำกัด (มหาชน) จำนวนทั้งสิ้น 13,765 เร็คคอร์ด สำหรับการดำเนินการวิจัยได้ทำการแบ่งข้อมูลออกเป็น 2 ชุด โดยใช้เกณฑ์เดือนของข้อมูลเป็นตัวแบ่งชุดข้อมูล ชุดแรกเป็นข้อมูลระหว่างเดือนมกราคม 2556 ถึงเดือนสิงหาคม 2557 จำนวน 12,893 เร็คคอร์ด ชุดนี้ใช้เพื่อการเรียนรู้และทดสอบโมเดลโดยโปรแกรม WEKA และชุดที่สองเป็นของเดือนกันยายน 2557 จำนวน 872 เร็คคอร์ด ชุดนี้ใช้เพื่อการประเมินประสิทธิภาพโปรแกรมพยากรณ์โอกาสที่บัญชีเงินกู้เบิกเกินบัญชีจะได้รับการทบทวนวงเงิน



ภาพที่ 5.1 สัดส่วนร้อยละของจำนวนข้อมูลที่ใช้เป็นชุดเรียนรู้และชุดทดสอบ

ผลการทดลองจากการประมวลผลข้อมูลด้วยเทคนิค J48 และเทคนิค ID3 พบว่า

1. เทคนิค J48 มีลำดับชั้นของการจำแนกข้อมูล 72 ชั้น (layers) และได้รูปแบบการตัดสินใจ 55 รูปแบบ

2. เทคนิค ID3 มีลำดับชั้นของการจำแนกข้อมูล 181 ชั้น (layers) และได้รูปแบบการตัดสินใจ 132 รูปแบบ นอกจากนี้ ยังพบว่าเทคนิค ID3 มีรูปแบบการตัดสินใจที่ไม่สามารถพยากรณ์ข้อมูลได้ จำนวน 4 รูปแบบ คือ แสดงผลการพยากรณ์เป็น “null”

3. ค่า True Positive Rate (Average) หรือค่า Recall (Average) หรือค่าที่ใช้วัดผลการพยากรณ์ตามรูปแบบโมเดลที่ถูกค้นพบ แล้วให้คำตอบตรงกับข้อมูลจริง ซึ่งทั้งของเทคนิค J48 และ ID3 มีค่าเท่ากันคือ 0.969

4. ค่า False Positive Rate (Average) หรือค่าที่ใช้วัดผลการพยากรณ์ตามรูปแบบโมเดลที่ถูกค้นพบ แล้วให้คำตอบเป็นเท็จ หมายถึง การที่ข้อมูลจริงไม่ได้อยู่ในคลาสนั้นแต่คำตอบจากโมเดลบอกว่าอยู่ในคลาสนั้น ของเทคนิค J48 และ ID3 มีค่าเท่ากับ 0.274 และ 0.275 ตามลำดับ

5. ค่า Precision (Average) หรือค่าที่ใช้วัดความเชื่อมั่นของผลการพยากรณ์ ของเทคนิค J48 และ ID3 มีค่าเท่ากับ 0.968 และ 0.967 ตามลำดับ

จากการวิเคราะห์และเปรียบเทียบข้อดีข้อเสียของเทคนิค J48 และเทคนิค ID3 พบว่ารูปแบบโมเดลการตัดสินใจที่ได้จากเทคนิค J48 มีข้อดีมากกว่าเทคนิค ID3 ทั้งในแง่ของลำดับชั้นการตัดสินใจที่ไม่เยอะจนเกินไป ดังที่แสดงไว้ในบทที่ 3 และเทคนิค J48 มีค่าดัชนีที่ใช้วัดประสิทธิภาพของรูปแบบโมเดลการตัดสินใจ ที่แสดงถึงความมีประสิทธิภาพในการจำแนกข้อมูลหรือมีความสามารถในการพยากรณ์ข้อมูลได้ดีกว่าเทคนิค ID3

ส่วนประกอบของโปรแกรมพยากรณ์โอกาสที่บัญชีเงินกู้เบิกเกินบัญชีจะได้รับการทบทวนวงเงิน มี 3 ส่วน ดังนี้

1. หน้าจอหลัก ใช้สำหรับการพยากรณ์

2. หน้าจอแสดง Decision Tree สำหรับผู้จัดทำข้อมูลใช้เพื่อทำความเข้าใจรูปแบบการตัดสินใจของตัวโปรแกรมพยากรณ์

3. หน้าจอแสดง Decision Tree สำหรับผู้ใช้ทั่วไปใช้เพื่อทำความเข้าใจรูปแบบการตัดสินใจของตัวโปรแกรมพยากรณ์ เพื่อความเข้าใจง่าย ผู้วิจัยจึงได้แปลงข้อมูลทางเทคนิค ให้แสดงผลเป็นเปอร์เซนต์ (%) ของค่าความน่าจะเป็น

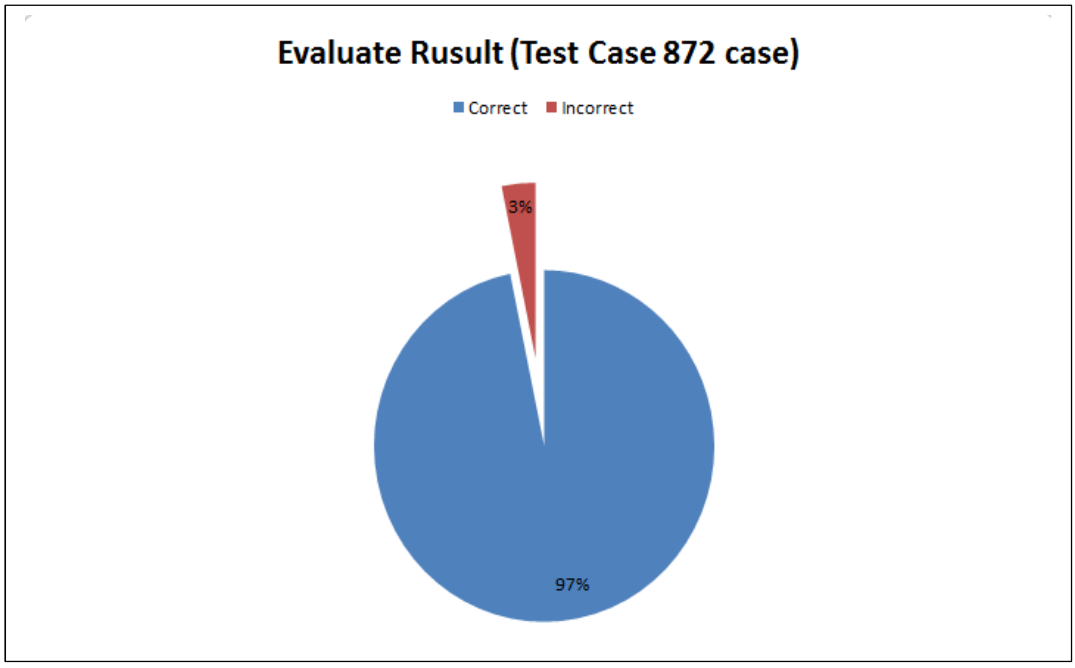
จากการนำข้อมูลชุดที่ 2 จำนวน 872 เร็คคอร์ด มาทำการทดสอบกับโปรแกรมพยากรณ์โอกาสที่บัญชีเงินกู้เบิกเกินบัญชีจะได้รับการทบทวนวงเงิน เพื่อประเมินประสิทธิภาพเชิงปริมาณของโปรแกรมฯ สรุปผลการทดสอบได้ดังนี้

1. ผลลัพธ์จากโปรแกรมพยากรณ์ กับผลการพิจารณาทบทวนวงเงินจากข้อมูลจริงมีความถูกต้องตรงกัน (Correctly Classified) จำนวน 844 เร็คคอร์ด คิดเป็นร้อยละ 96.79 หรือเรียกว่าค่า True Positive Rate หรือค่า Recall มีค่าเท่ากับ 0.9679

2. ผลลัพธ์จากโปรแกรมพยากรณ์ไม่ตรง กับผลการพิจารณาทบทวนวงเงินจากข้อมูลจริง (Incorrectly Classified) จำนวน 28 เร็คคอร์ด มีค่าเท่ากับ 0.0321 หรือคิดเป็นร้อยละ 3.21

3. รูปแบบโมเดลการติดสินใจที่ไม่ได้ทดสอบ เนื่องจากข้อมูลในชุดสอบไม่มีรูปแบบการตัดสินใจที่ตรงกับรูปแบบการตัดสินใจของโมเดล มีจำนวน 13 รูปแบบ

4. รูปแบบโมเดลการติดสินใจที่ผลการพยากรณ์ผิดพลาด จากผลการพิจารณาทบทวนวงเงินจากข้อมูลจริง ตั้งแต่ร้อยละ 20 ขึ้นไป มีจำนวน 11 รูปแบบ



ภาพที่ 5.2 สัดส่วนร้อยละของความถูกต้องของผลการพยากรณ์ (ชุดทดสอบ)

นอกจากนี้ ผู้วิจัยได้จัดทำแบบสอบถาม เพื่อประเมินประสิทธิภาพเชิงคุณภาพของโปรแกรมฯ จากผู้เชี่ยวชาญด้านสินเชื่อ จำนวน 33 คน แบ่งเป็นเพศชาย 18 คน และเพศหญิง 15 คน ส่วนใหญ่เป็นผู้มีตำแหน่งตั้งแต่ระดับ 7 หรือรองผู้จัดการฝ่ายขึ้นไป คิดเป็นร้อยละ 64 รองลงมาเป็นผู้มีตำแหน่งตั้งแต่ระดับ 5 - 6 หรือระดับหัวหน้างาน คิดเป็นร้อยละ 21 และตำแหน่งระดับ 1 – 4 หรือระดับเจ้าหน้าที่สินเชื่อ คิดเป็นร้อยละ 15 สามารถแปลผลการประเมินได้ดังนี้

1. เป็นผู้ที่ทราบเกี่ยวกับหลักเกณฑ์ของงานทบทวนวงเงิน ในระดับปานกลาง (จากคะแนนเต็ม 5 ได้คะแนนเฉลี่ย 3.27)

2. มีลูกค้าที่อยู่ในความดูแลที่เข้าร่วมโครงการของงานทบทวนวงเงิน ในระดับปานกลาง (จากคะแนนเต็ม 5 ได้คะแนนเฉลี่ย 2.70)

3. เห็นด้วยกับการนำ Utilization, LTV, Movement, Payment มาร่วมพิจารณาสำหรับการทบทวนวงเงินโอดี ในระดับมาก (จากคะแนนเต็ม 5 ได้คะแนนเฉลี่ย 3.88)

4. มีความคิดเห็นว่าการอธิบายรูปแบบการตัดสินใจด้วย Decision Tree ทำให้เข้าใจหลักเกณฑ์ของการทบทวนวงเงิน ในระดับมาก (จากคะแนนเต็ม 5 ได้คะแนนเฉลี่ย 3.52)

5. มีความคิดเห็นว่าการนำ Decision Tree ที่สร้างขึ้นจากโมเดลที่ได้จากการประมวลผลโดยโปรแกรม WEKA มีความสอดคล้องกับหลักความเป็นจริงทางธุรกิจ ในระดับปานกลาง (จากคะแนนเต็ม 5 ได้คะแนนเฉลี่ย 3.30)

6. มีความคิดเห็นว่าในส่วนผลลัพธ์ของการพยากรณ์ (Prediction) มีความน่าเชื่อถือและสอดคล้องกับความคาดหวัง ในระดับปานกลาง (จากคะแนนเต็ม 5 ได้คะแนนเฉลี่ย 3.24)

7. มีความคิดเห็นว่าโปรแกรมพยากรณ์ฯ ทำให้การสอบถาม หรือพยากรณ์โอกาสที่วงเงินเงินกู้เบิกเกินบัญชี (OD) จะได้รับการทบทวนวงเงิน ทำได้ง่ายขึ้น ในระดับมาก (จากคะแนนเต็ม 5 ได้คะแนนเฉลี่ย 3.55)

8. สรุปผลคะแนนเฉลี่ยในภาพรวมของการประเมินประสิทธิภาพเชิงคุณภาพของโปรแกรมพยากรณ์โอกาสที่บัญชีเงินกู้เบิกเกินบัญชีจะได้รับการทบทวนวงเงิน ในระดับปานกลาง (จากคะแนนเต็ม 5 ได้คะแนนเฉลี่ย 3.35)

5.2 ข้อเสนอแนะ

1. หากผู้สนใจต้องการนำแนวทางการพัฒนาโปรแกรมพยากรณ์โอกาสที่บัญชีเงินกู้เบิกเกินบัญชีจะได้รับการทบทวนวงเงินนี้ไปพัฒนาเพิ่มเติม อาจต้องทำการปรับเปลี่ยนเงื่อนไขสำหรับการจำแนกข้อมูลของแต่ละแอตทริบิวท์ ให้สอดคล้องกับนโยบายการทบทวนวงเงินของแต่ละองค์กร เนื่องจากแต่ละองค์กรอาจให้ความสำคัญกับแต่ละปัจจัยและมีมุมมองการจัดระดับความเสี่ยงที่แตกต่างกัน

2. LTV เป็นข้อมูลที่ว่าด้วยเรื่องมูลค่า หรือประเภทของหลักประกัน ที่ค้ำประกันต่อภาระหนี้ ซึ่งงานวิจัยนี้มีข้อมูล LTV ที่ระบุว่าเป็น CleanLoan หมายถึง ไม่มีมูลค่าหลักประกันที่ค้ำประกันวงเงินนั้นๆ ซึ่งในทางธุรกิจจะถือว่าวงเงินนั้นมีความเสี่ยงสูง และอาจจะพิจารณาไม่อนุมัติให้มีการทบทวนวงเงินในอนาคต แต่ผลการพยากรณ์ของโปรแกรมฯ ออกมาเป็น Pass ประเด็นนี้ผู้วิจัยได้ทำการตรวจสอบแล้ว พบว่า วงเงินเหล่านั้นมีการ Hold บัญชีเงินฝากออมทรัพย์ เพื่อหักชำระหนี้ แต่ในระบบข้อมูลหลักประกันไม่ได้บันทึกข้อมูลดังกล่าวไว้ และเมื่อข้อมูลดังกล่าวถูกตรวจพบโดยเจ้าหน้าที่ผู้เกี่ยวข้อง จึงทำการปรับเปลี่ยนผลการพิจารณาทบทวนวงเงินจาก No เป็น Pass เนื่องจากถือเสมือนว่าบัญชีเงินกู้เบิกเกินบัญชีนั้นๆ มีหลักประกันเป็นเงินฝาก

ดังนั้น เมื่อนำข้อมูลดังกล่าวมาเข้าสู่กระบวนการของการทำเหมืองข้อมูล จึงทำให้ผลการพยากรณ์ออกมาขัดแย้งกับความเป็นจริงในทางธุรกิจ แม้ว่าวงเงิน หรือข้อผิดพลาดเหล่านี้จะมีจำนวนน้อยมากก็ตาม แต่ถ้าหากนำโปรแกรมพยากรณ์โอกาสที่บัญชีเงินกู้เบิกเกินบัญชีจะได้รับการทบทวนออกใช้งานจริงในอนาคต อาจส่งผลต่อความน่าเชื่อถือของตัวโปรแกรมฯ ได้ ในการวิจัยครั้งต่อไป ต้องเพิ่มความระมัดระวังในการใช้คำที่อาจสื่อความหมายคลาดเคลื่อน

3. โปรแกรม WEKA ถูกพัฒนาโดยภาษา Java ผู้ใช้อาจพบปัญหาการใช้งาน อันเนื่องมาจากเวอร์ชั่นของ Java ที่ติดตั้งไว้ในฮาร์ดแวร์ของผู้ใช้อาจไม่สนับสนุนการทำงานของโปรแกรม ซึ่งผู้ใช้อาจต้องทำศึกษาค้นคว้าเพิ่มเติมเกี่ยวกับการกำหนดค่า (Configure) และเลือกใช้เวอร์ชั่นของ Java ให้เหมาะสม

4. เกร็ดความรู้เพิ่มเติมที่ผู้วิจัยค้นพบระหว่างการดำเนินการวิจัย และเพื่อเป็นวิทยาทานแก่ผู้ที่กำลังสนใจศึกษาเกี่ยวกับการทำเหมืองข้อมูลโดยใช้โปรแกรม WEKA ช่วยในการวิเคราะห์ข้อมูลมี 2 ประเด็น ดังนี้

4.1 หลักการตั้งชื่อแอตทริบิวท์ ห้ามมีช่องว่าง คั่นภายในชื่อของแอตทริบิวท์

4.2 ค่าข้อมูลในแต่ละแอตทริบิวท์ รวมถึงค่าข้อมูลที่เกิดจากการแปลงค่าข้อมูลจากตัวเลข เป็นรูปแบบนาม (Nominal) ห้ามมีช่องว่างคั่นระหว่างค่าข้อมูล หรือ Nominal นั้นๆ เพราะจะทำให้ผลการพยากรณ์ผิดพลาด โดยเฉพาะอย่างยิ่งในกรณีที่กำหนดให้มีการเรียกใช้โมเดลจากโปรแกรม WEKA โดยตรง