

บทที่ 5

สรุปผล อภิปราย และข้อเสนอแนะ

ระบบสนับสนุนการตัดสินใจในการเลือกใช้แพคเกจอินเทอร์เน็ตมือถือโดยใช้ต้นไม้ตัดสินใจ เป็นระบบที่พัฒนาขึ้นโดยใช้หลักการเหมืองข้อมูล (Data Mining) และเทคนิคต้นไม้ตัดสินใจ เพื่อช่วยในการตัดสินใจเลือกใช้แพคเกจอินเทอร์เน็ตมือถือ จากการวิจัยสามารถสรุปผลการวิจัย อภิปรายผลการวิจัยและนำมาซึ่งข้อเสนอแนะได้ดังนี้

5.1 สรุปผลการวิจัย

ในการจัดทำวิจัยครั้งนี้มีวัตถุประสงค์เพื่อสร้างโมเดลที่มีความเหมาะสมและถูกต้อง สำหรับการเลือกใช้แพคเกจอินเทอร์เน็ตมือถือ โดยใช้เทคนิคต้นไม้ตัดสินใจ ซึ่งใช้วิธีการหาอัลกอริทึมที่เหมาะสมโดยการเปรียบเทียบระหว่างอัลกอริทึมของต้นไม้ตัดสินใจ 2 อัลกอริทึม คือ ID3 และ C4.5 (J48) เนื่องจากเป็นอัลกอริทึมที่นิยมนำมาใช้ในการจัดหมวดหมู่ของข้อมูลที่ทราบผลลัพธ์แน่นอน

โดยการนำอัลกอริทึมดังกล่าวมาทดสอบกับกลุ่มข้อมูลสำหรับการเรียนรู้ (Training Data) ที่ได้เตรียมไว้จำนวน 1,000 ชุด เพื่อสร้างโมเดลในโปรแกรม Weka 3.6.2 และนำโมเดลที่ให้ผลค่าความถูกต้องมากที่สุดไปใช้ในการทดสอบกับชุดข้อมูลทดสอบ (Testing Data) จำนวน 500 ชุด ได้ผลการสร้างโมเดลและค่าความถูกต้องของโมเดลดังนี้

5.1.1 อัลกอริทึม ID3 โมเดลที่ได้มีโหนดราก (Root Node) คือ อายุระหว่าง 26-25 ปี มีค่าความถูกต้อง (Correctly Classified Instances) จากข้อมูล 1000 Instances โมเดลสามารถทำนายข้อมูลถูกต้อง 923 Instances หรือคิดเป็น 92.3% ของทั้งหมด มีค่าความไม่ถูกต้อง (Incorrectly Classified Instances) จากข้อมูล 1000 Instances โมเดลทำนายข้อมูลไม่ถูกต้อง 77 Instances หรือคิดเป็น 7.7% ของทั้งหมด และ มีค่าความคลาดเคลื่อนระหว่างค่าจริงและค่าที่พยากรณ์ได้ (Root Mean Squared Error : RMSE) เท่ากับ 0.1225

5.1.2 อัลกอริทึม C4.5 (J48) โมเดลที่ได้มีโหนดราก (Root Node) คือ ปริมาณการใช้งาน Youtube น้อยกว่า 3 ชั่วโมงต่อวัน มีค่าความถูกต้อง (Correctly Classified Instances) จากข้อมูล 1000 Instances โมเดลสามารถทำนายข้อมูลถูกต้อง 906 Instances หรือคิดเป็น 90.6% ของทั้งหมด มีค่าความไม่ถูกต้อง (Incorrectly Classified Instances) จากข้อมูล 1000 Instances โมเดลทำนายข้อมูลไม่ถูกต้อง 94 Instances หรือคิดเป็น 9.4% ของทั้งหมด และ มีค่าความคลาดเคลื่อนระหว่างค่าจริงและค่าที่พยากรณ์ได้ (Root Mean Squared Error : RMSE) เท่ากับ 0.1424

ส่วนค่า Confusion Matrix พบว่าผลจากการใช้อัลกอริทึมทั้ง 2 วิธี มีรูปแบบการกระจายตัวของข้อมูลที่มีลักษณะคล้ายกัน และอัลกอริทึม ID3 มีจำนวนข้อมูลค่าจริง กับจำนวนข้อมูลจากการทำนายมีผลลัพธ์ตรงกันมากกว่าผลที่ได้จากอัลกอริทึม C4.5 (J48)

5.1.3 เนื่องจากอัลกอริทึม ID3 มีค่าความถูกต้องและ Confusion Matrix มีค่าจำนวนข้อมูลค่าจริง กับจำนวนข้อมูลจากการทำนายมีผลลัพธ์ตรงกันมากกว่า อัลกอริทึม C4.5 (J48) ผู้วิจัยจึงนำอัลกอริทึม ID3 มาทดสอบกับชุดข้อมูลทดสอบ (Testing Data) จำนวน 500 ชุด ผลที่ได้ คือ มีค่าข้อมูลที่ทำนายถูกต้อง (Correctly Classified Instances) เท่ากับ 92.2 % มีค่าข้อมูลที่ทำนายไม่ถูกต้อง (Incorrectly Classified Instances) เท่ากับ 7.8% และเมื่อพิจารณา Confusion Matrix มีค่าของข้อมูลค่าจริง กับจำนวนข้อมูลจากการทำนาย แบ่งตาม Package และนำมาหาค่าเฉลี่ยรวมของทุก Package ได้ค่าเฉลี่ยรวมเท่ากับ 83.06%

5.2 อภิปรายผลการวิจัย

จากผลการวิจัยพบว่า ในการสร้างโมเดลเพื่อสนับสนุนการตัดสินใจในการเลือกใช้แพคเกจอินเทอร์เน็ตมือถือ โดยเปรียบเทียบการใช้อัลกอริทึมต้นไม้ตัดสินใจ คือ ID3 และ C4.5 (J48) โมเดลที่ถูกสร้างขึ้นจากการใช้อัลกอริทึม ID3 มีค่าความถูกต้องเท่ากับ 92.3% มากกว่าโมเดลที่ถูกสร้างขึ้นจากการใช้อัลกอริทึม C4.5 (J48) ซึ่งมีค่าความถูกต้องเท่ากับ 90.6% และโมเดลที่ถูกสร้างขึ้นจากการใช้อัลกอริทึม ID3 มีค่าการทำนายข้อมูลไม่ถูกต้องเท่ากับ 7.7% น้อยกว่าโมเดลที่ถูกสร้างขึ้นจากการใช้อัลกอริทึม C4.5 (J48) ซึ่งมีค่าการทำนายข้อมูลไม่ถูกต้องเท่ากับ 9.4% และโมเดลที่ถูกสร้างขึ้นจากการใช้อัลกอริทึม ID3 มีค่าความคลาดเคลื่อนเท่ากับ 0.1225 น้อยกว่าโมเดลที่ถูกสร้างขึ้นจากการใช้อัลกอริทึม C4.5 (J48) ซึ่งมีค่าความคลาดเคลื่อนเท่ากับ 0.1424

และเมื่อพิจารณาส่วนค่า Confusion Matrix พบว่าผลจากการใช้อัลกอริทึมทั้ง 2 วิธี มีรูปแบบการกระจายตัวของข้อมูลที่มีลักษณะคล้ายกัน และอัลกอริทึม ID3 มีจำนวนข้อมูลค่าจริง กับจำนวนข้อมูลจากการทำนายมีผลลัพธ์ตรงกันมากกว่าผลที่ได้จากอัลกอริทึม C4.5 (J48)

ผู้วิจัยจึงเลือกใช้โมเดลที่ถูกสร้างจากอัลกอริทึม ID3 ในการนำไปทดสอบกับชุดข้อมูลทดสอบ (Testing Data) จำนวน 500 ชุด ให้ผลการทดสอบโดยมีค่าความถูกต้อง (Correctly Classified Instances) เท่ากับ 92.2 % ซึ่งเป็นค่าความถูกต้องที่สูง ส่วนการวิเคราะห์ข้อมูลที่ทำนายไม่ถูกต้องมีค่าเท่ากับ 7.8% ซึ่งเป็นค่าที่ค่อนข้างต่ำ และจากค่า Confusion Matrix พบว่าผลของการทำนายจากโมเดลมีจำนวนข้อมูลค่าจริง กับจำนวนข้อมูลจากการทำนายมีผลลัพธ์ตรงกัน โดยแบ่งตาม Package และนำมาหาค่าเฉลี่ยรวมของทุก Package ได้ค่าเฉลี่ยรวมเท่ากับ 83.06% ซึ่งเป็น

ค่าเฉลี่ยของจำนวนข้อมูลค่าจริง กับจำนวนข้อมูลจากการทำนายมีผลลัพธ์ตรงกันอยู่ในระดับค่อนข้างสูง สามารถนำโมเดลไปใช้งานได้

ดังนั้นผู้วิจัยจึงเลือกใช้ โมเดลจากอัลกอริธึมต้นไม้ ID3 ในการนำไปพัฒนาระบบสนับสนุนการตัดสินใจในการเลือกใช้แพคเกจอินเทอร์เน็ตมือถือ เพราะมีค่าความถูกต้องของโมเดลและค่าเฉลี่ย Confusion Matrix จากทุก Package อยู่ในระดับค่อนข้างสูง

5.3 ข้อเสนอแนะ

เพื่อให้ระบบสนับสนุนการตัดสินใจในการเลือกใช้แพคเกจอินเทอร์เน็ตมือถือ โดยใช้ต้นไม้ตัดสินใจมีประสิทธิภาพมากขึ้น ทางผู้วิจัยเห็นว่าระบบควรได้รับการพัฒนาดังต่อไปนี้

5.3.1 ศึกษาและประยุกต์การใช้อัลกอริธึมอื่นๆที่เป็นการแบ่งกลุ่มเหมือนกันเพื่อช่วยสนับสนุนการตัดสินใจในการเลือกใช้แพคเกจอินเทอร์เน็ตมือถือ

5.3.2 ข้อมูลที่ใช้ในการวิเคราะห์มีจำนวนน้อยเกินไป ดังนั้น ในอนาคตควรมีการรวบรวมข้อมูลเพื่อทำการวิเคราะห์อีกครั้ง เพื่อให้ได้ผลของการวิเคราะห์ข้อมูลมีความสมบูรณ์มากขึ้น

5.3.3 เพื่อให้ได้โมเดลที่เหมาะสมที่สุดควรหาวิธีที่เหมาะสมสำหรับการตัดกิ่งต้นไม้ตัดสินใจในการสร้างโมเดล