บทที่ 4

ผลการวิจัย

ในบทที่ 3 ได้กล่าวถึงวิธีดำเนินงานวิจัยของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนด้วยเทคนิคเหมืองข้อมูล โดยใช้ข้อมูลภูมิหลัง และข้อมูลผลการเรียนในช่วงชั้นปีที่ 1 และ 2 ของนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศ และในบทนี้จะกล่าวถึงการผลการทดลองของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนด้วยเทคนิคเหมืองข้อมูล โดยแบ่งผลการทดลองออกเป็น 2 ผลการทดลอง ตามจุดมุ่งหมายของการวิจัย ได้แก่ ผลการทดลองที่ 1 เพื่อทราบประสิทธิภาพของการคัดเลือกตัวแปรด้วยเทคนิคการคัดเลือกชุดตัวแปรแบบอัตราส่วนเกนร่วมกับวิธีการค้นหาแบบจัดลำดับ ว่ามีผลต่อการคัดเลือกตัวแปรเท่าใด และมีอิทธิพลต่อค่าความถูกต้องในการทำนายเท่าใด และ ผลการทดลองที่ 2 เพื่อทราบประสิทธิภาพของการคัดเลือกตัวแปรด้วยสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันและสัมประสิทธิ์อีต้าว่ามีผลต่อการคัดเลือกตัวแปรเท่าใด และมีอิทธิพลต่อค่าความถูกต้องในการทำนายเท่าใด

ผลการทดลองที่ 1

การคัดเลือกตัวแปรด้วยเทคนิคการคัดเลือกชุดตัวแปรแบบอัตราส่วนเกนร่วมกับวิธีการค้นหาแบบจัดลำดับ

1. เทคนิคเคเนียเรสเนเบอร์

ข้อมูลชุดฝึกสอน

การทดสอบประสิทธิภาพการจัดหมวดหมู่ด้วยเทคนิคเคเนียร์เรสเนเบอร์ได้ทำการกำหนดค่า K จาก 1 ไปจนถึง 15 เพื่อหาจำนวน K ที่เหมาะสมที่สุด ผลการทดลองพบว่าค่าความถูกต้องลดลงและเพิ่มขึ้น ดังตาราง 19

ตาราง 19 แสดสงผลทดลองของโมเดลด้วยเทคนิคเคเนียร์เรสเนเบอร์

จำนวน K	CS 22 ตัวแปร	IT 22 ตัวแปร	CS 11 ตังแปร	IT 8 ตัวแปร	CS 4 ตัวแปร	IT 4 ตัวแปร	CS+IT 4 ตัวแปร
	ค่าความถูกต้อง (%)		ค่าความถูกต้อง(%)		ค่าความถูกต้อง (%)		ค่าความถูกต้อง (%)
1	65.289	**67.826**	**72.727**	65.217	76.033	67.826	77.119
2	**66.116**	60.000	72.727	69.565	**83.471**	70.435	78.390
3	65.289	70.435	71.901	**76.522**	82.645	79.130	79.661
4	66.116	66.087	76.033	75.652	80.165	79.130	80.932
5	63.636	66.087	75.207	80.870	82.645	**81.739**	**81.356**
6	62.810	64.348	75.207	79.130	83.471	80.000	80.509
7	63.636	66.087	78.512	80.870	83.471	80.870	82.203
8	66.942	60.870	79.339	81.739	81.818	80.000	80.085
9	66.116	66.957	79.339	81.739	81.818	85.217	80.932
10	65.289	67.826	79.339	78.261	84.298	83.478	81.780
11	63.636	69.565	78.512	80.870	81.818	83.478	83.051
12	61.984	70.435	80.165	80.000	85.124	82.609	82.203
13	61.984	67.826	78.512	80.000	83.471	84.348	83.051
14	63.636	70.435	79.339	78.261	85.950	81.739	83.051
15	61.984	68.696	76.033	78.261	81.818	84.348	83.475

จากตาราง 19 จะเห็นได้ว่า จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 66.116% และ 67.826% จำนวน 11 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ค่าความถูกต้องอยู่ที่ 72.727% จำนวน 8 ตัวแปรต้นของสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 76.522% จำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 83.471% และ 81.739% และเมื่อนำจำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศมารวมกันค่าความถูกต้องอยู่ที่ 81.356% จะเห็นได้ว่าค่าความถูกต้องหลังจากการนำ 2 สาขาวิชามารวมกัน ค่าความถูกต้องที่ได้ก็ยังค่อนข้างสูง

2. เทคนิคเนอีฟเบย์

ข้อมูลชุดฝึกสอน

การทดสอบประสิทธิภาพการจัดหมวดหมู่ด้วยเทคนิคเนอีฟเบย์ได้ทำการหาค่าความถูกต้อง ค่าความแม่นยำ และค่าความระลึก ดังตาราง 20

ตาราง 20 แสดงผลทดลองของโมเดลด้วยเทคนิคเนอีฟเบย์

	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
CS (22 attributes)	79.339	0.796	0.793
IT (22 attributes)	80.000	0.795	0.800
CS (11 attributes)	84.298	0.844	0.843
IT (8 attributes)	84.348	0.842	0.843
CS (4 attributes)	85.950	0.863	0.860
IT (4 attributes)	84.348	0.843	0.843
CS+IT (4 attributes)	84.322	0.843	0.843

จากตาราง 20 จะเห็นได้ว่า จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 79.339% และ 80.000% จำนวน 11 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ค่าความถูกต้องอยู่ที่ 84.298% จำนวน 8 ตัวแปรต้นของสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 84.348 % จำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 85.950% และ 84.348% และเมื่อนำจำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศมารวมกันค่าความถูกต้องอยู่ที่ 84.322% จะเห็นได้ว่าค่าความถูกต้องหลังจากการนำ 2 สาขาวิชามารวมกัน ค่าความถูกต้องที่ได้ก็ยังค่อนข้างสูง

สำนักงานคณะกรรมการวิจัยแห่งชาติ ห้องสมุดงานวิจัย

3. เทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ

ข้อมูลชุดฝึกสอน

การทดสอบประสิทธิภาพการจัดหมวดหมู่ด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับได้ทำการกำหนดอัตราการเรียนรู้จาก 0.1 ถึง 0.3 และปรับจำนวนโหนดของชั้นซ่อนเพื่อเปรียบเทียบค่าความถูกต้องของตัวแบบ ดังตาราง 21 ถึง ตาราง 27

ตาราง 21 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ สาขาวิชาวิทยาการคอมพิวเตอร์ กำหนดอัตราการเรียนรู้ = 0.1

	อัตราการเรียนรู้ =0.1			
	Model	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
	22:10:3	**80.992**	0.810	0.810
	22:11:3	79.339	0.793	0.793
CS (22 attributes)	22:12:3	79.339	0.793	0.793
	22:13:3	79.339	0.793	0.793
	22:14:3	78.512	0.785	0.785
	11:5:3	**90.909**	0.911	0.909
	11:6:3	90.083	0.903	0.901
CS (11 attributes)	11:7:3	90.083	0.903	0.901
	11:8:3	90.083	0.903	0.901
	11:9:3	90.083	0.903	0.901
	4:1:3	74.380	0.619	0.744
	4:2:3	**87.603**	0.878	0.876
CS (4 attributes)	4:3:3	87.603	0.878	0.876
	4:4:3	87.603	0.878	0.876
	4:5:3	87.603	0.878	0.876

จากตาราง 21 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.1 จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ ค่าความถูกต้องสูงสุดอยู่ที่ 80.992% (22:10:3) จำนวน 11 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 90.909% (11:5:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 87.603% (4:2:3)

ตาราง 22 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ สาขาวิชาวิทยาการคอมพิวเตอร์ กำหนดอัตราการเรียนรู้ = 0.2

	อัตราการเรียนรู้ =0.2			
	Model	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
CS (22 attributes)	22:10:3	**79.339**	0.793	0.793
	22:11:3	79.339	0.793	0.793
	22:12:3	79.339	0.793	0.793
	22:13:3	79.339	0.793	0.793
	22:14:3	78.512	0.786	0.785
CS (11 attributes)	11:5:3	**90.909**	0.911	0.909
	11:6:3	90.909	0.911	0.909
	11:7:3	90.909	0.911	0.909
	11:8:3	90.909	0.913	0.909
	11:9:3	90.909	0.913	0.909
CS (4 attributes)	4:1:3	74.380	0.619	0.744
	4:2:3	88.430	0.888	0.884
	4:3:3	88.430	0.888	0.884
	4:4:3	**89.256**	0.895	0.893
	4:5:3	88.430	0.888	0.884

จากตาราง 22 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.2 จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ ค่าความถูกต้องสูงสุดอยู่ที่ 79.339% (22:10:3) จำนวน 11 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 90.909% (11:5:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 89.256% (4:4:3)

ตาราง 23 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ สาขาวิชาวิทยาการคอมพิวเตอร์ กำหนดอัตราการเรียนรู้ = 0.3

	อัตราการเรียนรู้ = 0.3			
	Model	**ค่าความถูกต้อง (%)**	**ค่าความแม่นยำ**	**ค่าความระลึก**
CS (22 attributes)	22:10:3	77.686	0.777	0.777
	22:11:3	**80.165**	0.802	0.802
	22:12:3	79.339	0.793	0.793
	22:13:3	79.339	0.793	0.793
	22:14:3	79.339	0.793	0.793
CS (11 attributes)	11:5:3	**90.909**	0.913	0.909
	11:6:3	90.083	0.904	0.901
	11:7:3	90.909	0.913	0.909
	11:8:3	90.909	0.913	0.909
	11:9:3	90.909	0.913	0.909
CS (4 attributes)	4:1:3	73.554	0.614	0.736
	4:2:3	**89.256**	0.895	0.893
	4:3:3	88.430	0.888	0.884
	4:4:3	87.603	0.88	0.876
	4:5:3	86.777	0.87	0.868

จากตาราง 23 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.3 จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ ค่าความถูกต้องสูงสุดอยู่ที่ 80.165% (22:11:3) จำนวน 11 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 90.909% (11:5:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 89.256% (4:2:3)

ตาราง 24 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดสาขาวิชาเทคโนโลยีสารสนเทศ อัตราการเรียนรู้ = 0.1

	อัตราการเรียนรู้ =0.1			
	Model	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
IT (22 attributes)	22:10:3	**80.000**	0.806	0.8
	22:11:3	78.261	0.79	0.783
	22:12:3	77.391	0.779	0.774
	22:13:3	78.261	0.79	0.783
	22:14:3	79.130	0.798	0.791
IT (8 attributes)	8:3:3	77.391	0.772	0.774
	8:4:3	77.391	0.772	0.774
	8:5:3	78.261	0.783	0.783
	8:6:3	77.391	0.772	0.774
	8:7:3	**79.130**	0.791	0.791
IT (4 attributes)	4:1:3	75.652	0.658	0.757
	4:2:3	**80.870**	0.806	0.809
	4:3:3	80.870	0.806	0.809
	4:4:3	80.870	0.806	0.809
	4:5:3	80.870	0.806	0.809

จากตาราง 24 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.1 จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ ค่าความถูกต้องสูงสุดอยู่ที่ 80.000% (22:10:3) จำนวน 11 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 79.130% (8:7:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 80.870% (4:2:3)

ตาราง 25 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดสาขาวิชาเทคโนโลยีสารสนเทศ อัตราการเรียนรู้ = 0.2

	อัตราการเรียนรู้ =0.2			
	Model	**ค่าความถูกต้อง (%)**	**ค่าความแม่นยำ**	**ค่าความระลึก**
	22:10:3	79.130	0.796	0.791
	22:11:3	77.391	0.777	0.774
IT (22 attributes)	22:12:3	78.261	0.785	0.783
	22:13:3	**80.000**	0.804	0.8
	22:14:3	79.130	0.796	0.791
	8:3:3	77.391	0.772	0.774
	8:4:3	76.522	0.766	0.765
IT (8 attributes)	8:5:3	78.261	0.783	0.783
	8:6:3	77.391	0.768	0.774
	8:7:3	**80.870**	0.809	0.809
	4:1:3	73.913	0.701	0.739
	4:2:3	80.870	0.806	0.809
IT (4 attributes)	4:3:3	80.870	0.809	0.809
	4:4:3	**81.739**	0.817	0.817
	4:5:3	81.739	0.817	0.817

จากตาราง 25 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.2 จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ ค่าความถูกต้องสูงสุดอยู่ที่ 80.000% (22:13:3) จำนวน 11 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 80.870% (8:7:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 81.739% (4:4:3)

ตาราง 26 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดสาขาวิชาเทคโนโลยีสารสนเทศ อัตราการเรียนรู้ = 0.3

	อัตราการเรียนรู้ =0.3			
	Model	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
	22:10:3	79.130	0.796	0.791
	22:11:3	77.391	0.777	0.774
IT (22 attributes)	22:12:3	77.391	0.781	0.774
	22:13:3	**80.000**	0.804	0.8
	22:14:3	78.261	0.787	0.783
	8:3:3	78.261	0.779	0.783
	8:4:3	75.652	0.755	0.757
IT (8 attributes)	8:5:3	**80.000**	0.802	0.8
	8:6:3	79.130	0.791	0.791
	8:7:3	80.000	0.795	0.8
	4:1:3	73.913	0.701	0.739
	4:2:3	80.870	0.806	0.809
IT (4 attributes)	4:3:3	80.870	0.809	0.809
	4:4:3	**81.739**	0.817	0.817
	4:5:3	80.000	0.800	0.800

จากตาราง 26 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.3 จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ ค่าความถูกต้องสูงสุดอยู่ที่ 80.000% (22:13:3) จำนวน 11 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 80.000% (8:5:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 81.739% (4:4:3)

สาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 4 ตัวแปรต้น

การทดสอบประสิทธิภาพการจัดหมวดหมู่ด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับได้ทำการทำการกำหนดอัตราการเรียนรู้จาก 0.1 ถึง 0.3 และปรับจำนวนโหนดของชั้นซ่อน เพื่อเปรียบเทียบค่าความถูกต้องของตัวแบบ ด้วยจำนวนตัวแปร 4 ตัวแปรต้น และ 1 ตัวแปรตาม ดังตาราง 27

ตาราง 27 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับของสองสาขาวิชา จำนวน 4 ตัวแปรต้น

Model	อัตราการเรียนรู้	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
4:1:3	0.1	81.356	0.821	0.814
4:1:3	0.2	81.356	0.820	0.814
4:1:3	0.3	81.356	0.820	0.814
4:2:3	0.1	81.356	0.817	0.814
4:2:3	0.2	81.780	0.822	0.818
4:2:3	0.3	81.356	0.817	0.814
4:3:3	0.1	81.780	0.821	0.818
4:3:3	0.2	83.051	0.833	0.831
4:3:3	0.3	**84.322**	0.845	0.843
4:4:3	0.1	83.051	0.832	0.831
4:4:3	0.2	82.203	0.823	0.822
4:4:3	0.3	83.051	0.831	0.831
4:5:3	0.1	83.051	0.832	0.831
4:5:3	0.2	83.898	0.84	0.839
4:5:3	0.3	83.898	0.84	0.839

จากตาราง 26 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ 0.1- 0.3 จำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศ ค่าความถูกต้องสูงสุดอยู่ที่ 84.322% (4:3:3)

4. เทคนิคต้นไม้ตัดสินใจ

ข้อมูลชุดฝึกสอน

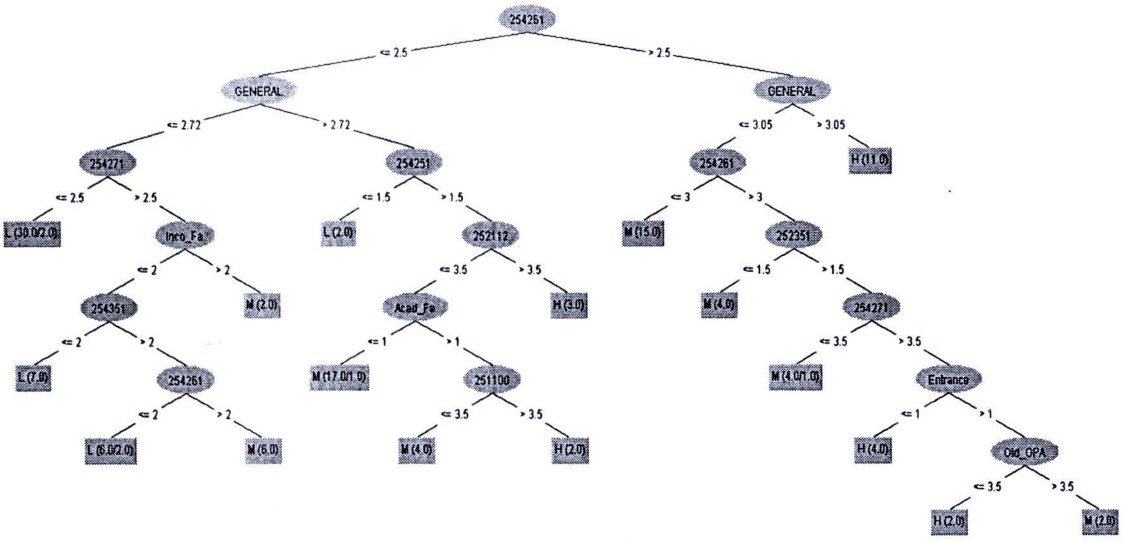
การทดสอบประสิทธิภาพการจัดหมวดหมู่ด้วยเทคนิคต้นไม้ตัดสินใจได้ทำการหาค่าความถูกต้อง ค่าความแม่นยำ และค่าความระลึก ดังตาราง 28

ตาราง 28 แสดงการทดสอบประสิทธิภาพของโมเดลด้วยเทคนิคต้นไม้ตัดสินใจ

	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
CS (22 attributes)	65.289	0.658	0.653
IT (22 attributes)	77.391	0.772	0.774
CS (11 attributes)	70.248	0.713	0.702
IT (8 attributes)	83.478	0.833	0.835
CS (4 attributes)	75.207	0.755	0.752
IT (4 attributes)	78.261	0.785	0.783
CS+IT (4 attributes)	76.271	0.762	0.763

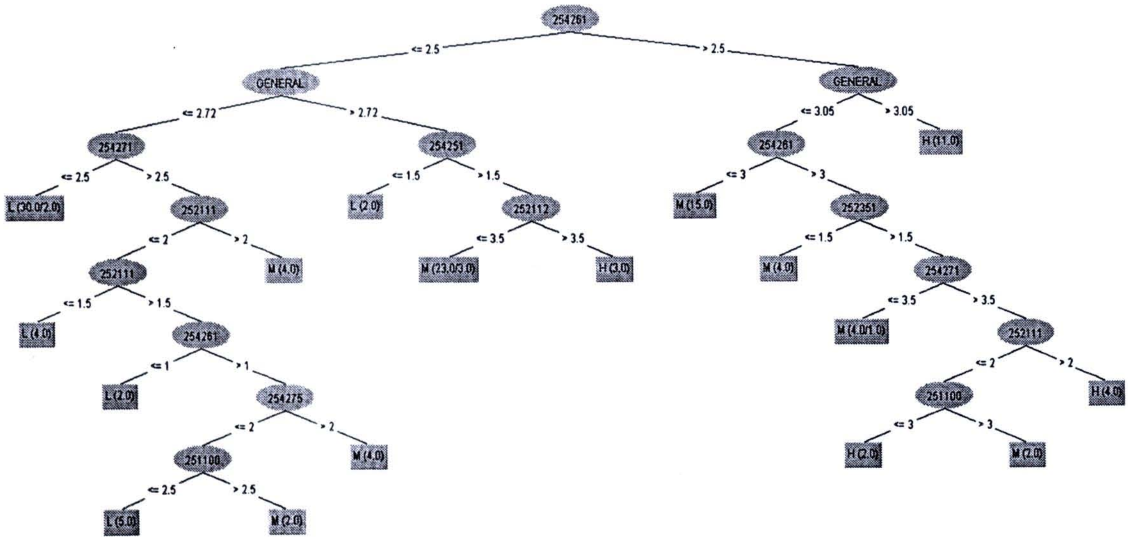
จากตาราง 28 จะเห็นได้ว่า จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 65.289% และ 77.391% จำนวน 11 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ค่าความถูกต้องอยู่ที่ 70.248% จำนวน 8 ตัวแปรต้นของสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 83.478% จำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 75.207% และ 78.261% และเมื่อนำจำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศมารวมกันค่าความถูกต้องอยู่ที่ 76.271% จะเห็นได้ว่าค่าความถูกต้องหลังจากการนำ 2 สาขาวิชามารวมกัน ค่าความถูกต้องที่ได้ก็ยังค่อนข้างสูง

ภาพ 29 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ จำนวน 22 ตัวแปรต้น



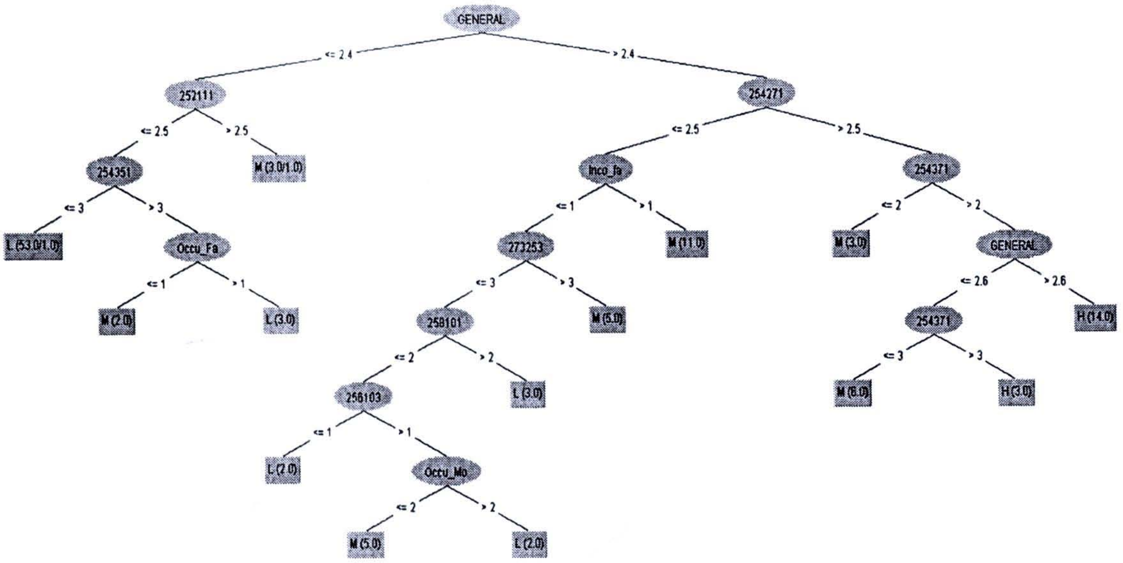
ภาพ 29 แสดงต้นไม้ตัดสินใจ จำนวน 22 ตัวแปรต้น สาขาวิชาวิทยาการคอมพิวเตอร์

ภาพ 30 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ จำนวน 11 ตัวแปรต้น



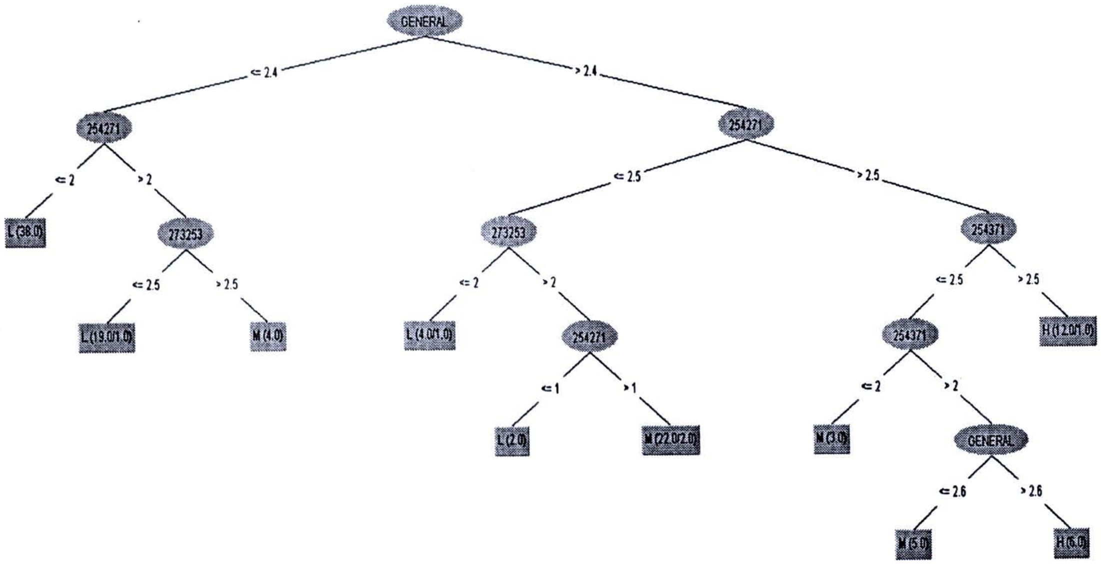
ภาพ 30 แสดงต้นไม้ตัดสินใจ จำนวน 11 ตัวแปรต้น สาขาวิชาวิทยาการคอมพิวเตอร์

ภาพ 31 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 22 ตัวแปรต้น



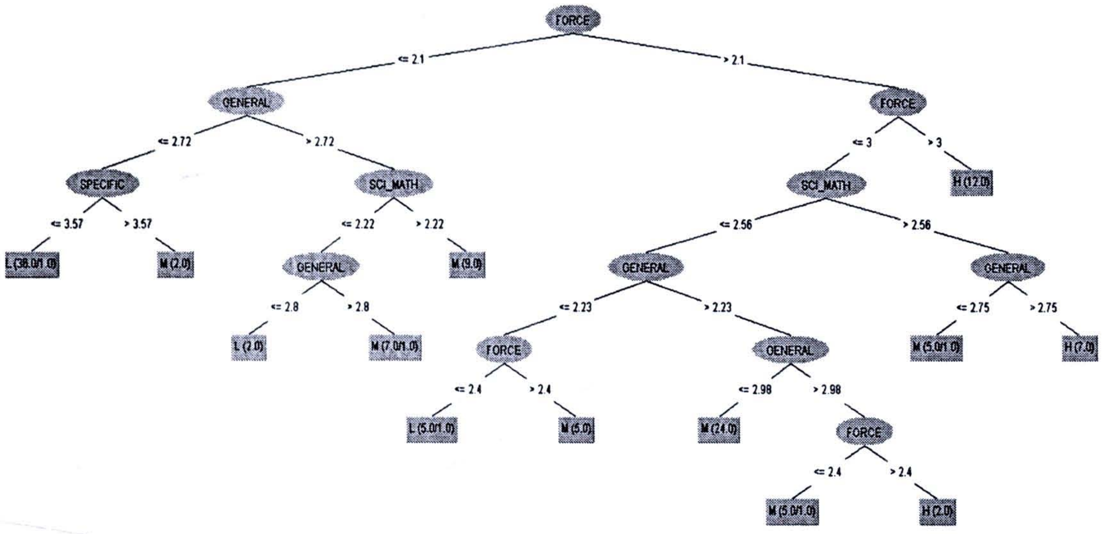
ภาพ 31 แสดงต้นไม้ตัดสินใจ จำนวน 22 ตัวแปรต้น สาขาวิชาเทคโนโลยีสารสนเทศ

ภาพ 32 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 8 ตัวแปรต้น



ภาพ 32 แสดงต้นไม้ตัดสินใจ จำนวน 8 ตัวแปรต้น สาขาวิชาเทคโนโลยีสารสนเทศ

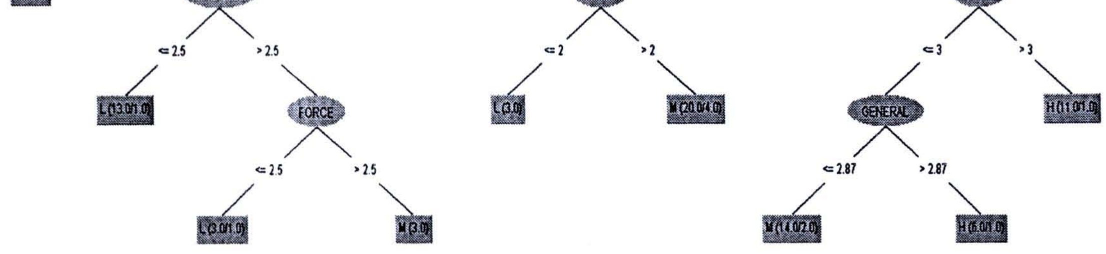
ภาพ 33 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ จำนวน 4 ตัวแปรต้น



ภาพ 33 แสดงต้นไม้ตัดสินใจ จำนวน 4 ตัวแปรต้น สาขาวิชาวิทยาการคอมพิวเตอร์

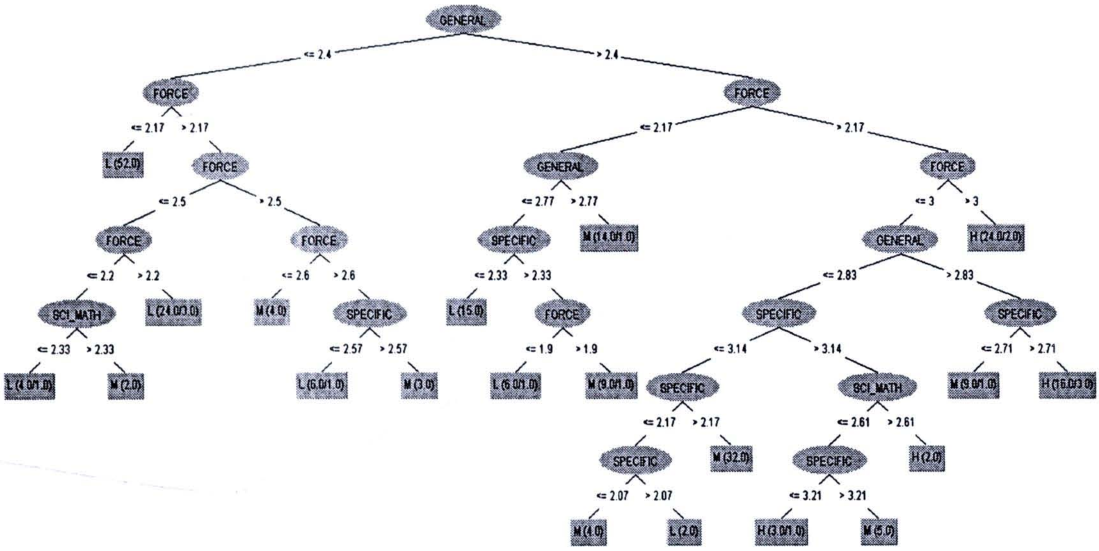
ภาพ 34 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 4 ตัวแปรต้น

สำนักงานคณะกรรมการวิจัยแห่งชาติ
ห้องสมุดงานวิจัย



ภาพ 34 แสดงต้นไม้ตัดสินใจ จำนวน 4 ตัวแปรต้น สาขาวิชาเทคโนโลยีสารสนเทศ

ภาพ 35 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 4 ตัวแปรต้น



ภาพ 35 แสดงต้นไม้ตัดสินใจของสองสาขาวิชา จำนวน 4 ตัวแปรต้น

ในตาราง 29 และ 30 เป็นการแสดงการคัดเลือกตัวแปรด้วยเทคนิคการคัดเลือกชุดตัวแปรแบบอัตราส่วนเกนร่วมกับวิธีการค้นหาแบบจัดลำดับ และทำการประเมินตัวแบบจากตัวแปรที่ได้จากการคัดเลือก จากข้อมูลชุดฝึกสอน

ตาราง 29 แสดงการทดสอบประสิทธิภาพของการจำแนกข้อมูลชุดฝึกสอน

		KNN	Naïve Bayes	BPNN	Decision Tree
ก่อนการคัดเลือกตัวแปร	CS (22 attributes)	66.116%	79.339%	80.992%	65.289%
	IT (22 attributes)	67.826%	80.000%	80.000%	77.391%
หลังการคัดเลือกตัวแปร	CS (11 attributes)	72.727%	84.298%	90.909%	70.248%
	IT (8 attributes)	76.522%	84.348%	80.870%	83.478%

ตาราง 30 แสดงการทดสอบประสิทธิภาพของการจำแนกข้อมูลชุดฝึกสอนที่รวมสาขา

		KNN	Naïve Bayes	BPNN	Decision Tree
แยกสาขาวิชา	CS (4 attributes)	83.471%	85.950%	89.256%	75.207%
	IT (4 attributes)	81.739%	84.348%	81.739%	78.261%
รวมสาขาวิชา	CS+IT (4 attributes)	81.356%	84.322%	84.322%	76.271%

นำตัวแปรที่ผ่านการคัดเลือกตัวแปรด้วยเทคนิคการคัดเลือกชุดตัวแปรแบบอัตราส่วนเกนร่วมกับวิธีการค้นหาแบบจัดลำดับ และนำมาหาค่าเฉลี่ยตามกลุ่มวิชา จำนวนตัวแปรต้น 4 ตัวแปร ได้แก่ หมวดวิชาศึกษาทั่วไป กลุ่มวิชาพื้นฐานวิทยาศาสตร์ และคณิตศาสตร์ กลุ่มวิชาพื้นฐานเฉพาะด้าน และกลุ่มวิชาเอกบังคับ และตัวแปรตาม 1 ตัวแปร คือ ผลสัมฤทธิ์ทางการเรียน มาใช้กับเทคนิคเหมืองข้อมูล 4 เทคนิค เพื่อนำมาเปรียบเทียบประสิทธิภาพการทำงาน ได้ค่าความถูกต้อง ค่าความระลึก และค่าความแม่นยำ ดังตาราง 31 ได้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของข้อมูลชุดฝึกสอน และข้อมูลชุดทดสอบ ดังตาราง 32

ผลการวิจัยพบว่า เทคนิคเนอีฟเบย์ และเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับให้ค่าความถูกต้องในการพยากรณ์ใกล้เคียงกัน และสูงกว่าเทคนิคเคเนียเรสเนเบอร์ และเทคนิคต้นไม้ตัดสินใจ ให้ค่าความถูกต้องต่ำสุด

จากค่าความถูกต้อง และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง แสดงให้เห็นว่าเทคนิคที่เหมาะสมที่จะนำมาสร้างตัวแบบ คือ เทคนิคเนอีฟเบย์ และเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับร่วมกัน เพราะให้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่ต่ำ และให้ค่าความถูกต้องค่อนข้างสูง มีความน่าเชื่อถือสูงเพราะนำเทคนิคที่มีความแม่นยำสูงทั้ง 2 เทคนิค มาร่วมกันพยากรณ์

ตาราง 31 แสดงการเปรียบเทียบการทดสอบประสิทธิภาพของข้อมูลชุดฝึกสอนและข้อมูลชุดทดสอบ

	Training Data			Testing Data		
	Accuracy (%)	Precision	Recall	Accuracy (%)	Precision	Recall
K-NN	81.356%	0.814	0.814	73.016%	0.747	0.730
Naïve Bayes	**84.322%**	0.843	0.843	**73.810%**	0.738	0.738
BPNN	**84.322%**	0.845	0.843	**73.016%**	0.748	0.730
Decision Tree	76.271%	0.763	0.763	70.635%	0.706	0.706

จากตาราง 31 จะเห็นว่าเทคนิคเนอีฟเบย์ และ เทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ ให้ค่าความถูกต้องในการพยากรณ์ใกล้เคียงกัน และสูงกว่าเทคนิคเคเนียเรสเนเบอร์ และเทคนิคต้นไม้ตัดสินใจ ให้ค่าความถูกต้องต่ำสุด ทั้งข้อมูลชุดฝึกสอนและข้อมูลชุดทดสอบ

ตาราง 32 แสดงการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของข้อมูลชุดฝึกสอนและข้อมูลชุดทดสอบ

	RMSE (Training Data)	RMSE (Testing Data)
Naïve Bayes	0.150	0.188
BPNN	0.148	0.185
Combine 2 algorithms	**0.140**	**0.172**

จากตาราง 32 จะเห็นว่าค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของเทคนิคเนอีฟเบย์ และเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับแบบทำงานร่วมกัน ให้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่ต่ำ

นำตัวแบบที่ได้สร้างขึ้นมาใช้ประโยชน์จริงในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของนิสิต หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศ ในรุ่นต่อ ๆ ไป ดังตาราง 33

ตาราง 33 แสดงการเปรียบเทียบประสิทธิภาพของข้อมูลชุดตรวจสอบ

	Validation Data			
	Accuracy (%)	Precision	Recall	RMSE
Naïve Bayes	74.8031 %	0.794	0.772	0.424
BPNN	77.1654 %	0.757	0.748	0.415
Combine 2 algorithms	-	-	-	**0.414**

จากตาราง 33 แสดงให้เห็นว่าเทคนิคที่เหมาะสมที่จะนำมาสร้างตัวแบบ คือ เทคนิคเนอีฟเบย์ และเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับแบบทำงานร่วมกัน เพราะให้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่ต่ำ มีความน่าเชื่อถือสูงเพราะนำเทคนิคที่มีความแม่นยำสูงทั้ง 2 เทคนิค มาร่วมกันพยากรณ์

ผลการทดลองที่ 2

การคัดเลือกตัวแปรด้วยสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน และสัมประสิทธิ์อีต้า

1. เทคนิคเคเนียเรสเนเบอร์

ข้อมูลชุดฝึกสอน

การทดสอบประสิทธิภาพการจัดหมวดหมู่ด้วยเทคนิคเคเนียร์เรสเนเบอร์ได้ทำการกำหนดค่า K มีค่าเริ่มต้น 1 ไปจนถึง 15 เพื่อหาจำนวน K ที่เหมาะสมที่สุด ผลการทดลองพบว่าค่าความถูกต้องลดลงและเพิ่มขึ้น ดังตาราง 34

ตาราง 34 แสดงการทดสอบประสิทธิภาพของโมเดลด้วยเทคนิคเคเนียร์เรสเนเบอร์

จำนวน K	CS 22 ตัวแปร	IT 22 ตัวแปร	CS 13 ตัวแปร	IT 13 ตัวแปร	CS 4 ตัวแปร	IT 4 ตัวแปร	CS+IT 4 ตัวแปร
	ค่าความถูกต้อง (%)		ค่าความถูกต้อง (%)		ค่าความถูกต้อง (%)		ค่าความถูกต้อง (%)
1	65.289	**67.826**	**73.554**	73.044	78.512	74.783	**82.203**
2	66.116	60.000	71.901	73.913	79.339	75.652	77.119
3	65.289	70.435	71.901	**83.478**	**80.992**	**79.130**	83.898
4	**66.116**	66.087	76.860	75.652	79.339	77.391	85.170
5	63.636	66.087	75.207	80.870	79.339	75.652	82.627
6	62.810	64.348	80.165	78.261	82.645	77.391	83.051
7	63.636	66.087	76.860	79.130	81.818	76.522	81.3560
8	66.942	60.870	77.686	79.130	83.471	75.652	81.780
9	66.116	66.957	77.686	76.522	80.165	76.522	82.627
10	65.289	67.826	77.686	70.435	82.645	78.261	82.627
11	63.636	69.565	78.512	74.783	80.992	80.870	83.475
12	61.984	70.435	75.207	75.652	81.818	80.000	84.746
13	61.984	67.826	6.860	72.174	81.818	85.217	83.051
14	63.636	70.435	75.207	75.652	83.471	80.000	84.746
15	61.984	68.696	78.512	73.044	82.645	82.609	85.593

จากตาราง 34 จะเห็นได้ว่า จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 66.116% และ 67.826% จำนวน 13 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ค่าความถูกต้องอยู่ที่ 73.554% จำนวน 13 ตัวแปรต้นของสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 83.478% จำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 80.992% และ 79.130% และเมื่อนำจำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศมารวมกันค่าความถูกต้องอยู่ที่ 82.203% จะเห็นได้ว่าค่าความถูกต้องหลังจากการนำ 2 สาขาวิชามารวมกัน ค่าความถูกต้องที่ได้ก็ยังค่อนข้างสูง

2. เทคนิคเนอีฟเบย์

ข้อมูลชุดฝึกสอน

การทดสอบประสิทธิภาพการจัดหมวดหมู่ด้วยเทคนิคเนอีฟเบย์ได้ทำการหาค่าความถูกต้อง ค่าความแม่นยำ และค่าความระลึก ดังตาราง 35

ตาราง 35 แสดงการทดสอบประสิทธิภาพของโมเดลด้วยเทคนิคเนอีฟเบย์

	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
CS (22 attributes)	79.339	0.796	0.793
IT (22 attributes)	80.000	0.795	0.80
CS (13 attributes)	82.645	0.830	0.826
IT (13 attributes)	85.217	0.848	0.852
CS (4 attributes)	84.298	0.849	0.843
IT (4 attributes)	86.957	0.868	0.870
CS+IT (4 attributes)	86.441	0.864	0.864

จากตาราง 35 จะเห็นได้ว่า จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 79.339% และ 80.000% จำนวน 13 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ค่าความถูกต้องอยู่ที่ 82.645% จำนวน 13 ตัวแปรต้นของสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 85.217% จำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 84.298% และ 86.957% และเมื่อนำจำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศมารวมกันค่าความถูกต้องอยู่ที่ 86.441% จะเห็นได้ว่าค่าความถูกต้องหลังจากการนำ 2 สาขาวิชามารวมกัน ค่าความถูกต้องที่ได้ก็ยังค่อนข้างสูง

3. เทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ

ข้อมูลชุดฝึกสอน

การทดสอบประสิทธิภาพการจัดหมวดหมู่ด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับได้ทำการกำหนดอัตราการเรียนรู้จาก 0.1 ถึง 0.3 เพื่อเปรียบเทียบค่าความถูกต้องของตัวแบบ ดังตาราง 36 ถึง ตาราง 42

ตาราง 36 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.1 สาขาวิชาวิทยาการคอมพิวเตอร์

	อัตราการเรียนรู้=0.1			
	Model	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
CS (22 attributes)	22:10:3	**80.992**	0.810	0.810
	22:11:3	79.339	0.793	0.793
	22:12:3	79.339	0.793	0.793
	22:13:3	79.339	0.793	0.793
	22:14:3	78.512	0.785	0.785
CS (13 attributes)	13:6:3	87.603	0.881	0.876
	13:7:3	87.603	0.881	0.876
	13:8:3	**88.430**	0.888	0.884
	13:9:3	88.430	0.888	0.884
	13:10:3	88.430	0.889	0.884
CS (4 attributes)	4:1:3	74.380	0.614	0.744
	4:2:3	**89.256**	0.893	0.893
	4:3:3	89.256	0.893	0.893
	4:4:3	89.256	0.893	0.893
	4:5:3	89.256	0.893	0.893

จากตาราง 36 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.1จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ ค่าความถูกต้องสูงสุดอยู่ที่ 80.992% (22:10:3) จำนวน 13 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 88.430% (13:8:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 89.256% (4:2:3)

ตาราง 37 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.2 สาขาวิชาวิทยาการคอมพิวเตอร์

	อัตราการเรียนรู้=0.2			
	Model	**ค่าความถูกต้อง (%)**	**ค่าความแม่นยำ**	**ค่าความระลึก**
	22:10:3	**79.339**	0.793	0.793
	22:11:3	79.339	0.793	0.793
CS (22 attributes)	22:12:3	79.339	0.793	0.793
	22:13:3	79.339	0.793	0.793
	22:14:3	78.512	0.786	0.785
	13:6:3	87.603	0.881	0.876
	13:7:3	87.603	0.883	0.876
CS (13 attributes)	13:8:3	**88.430**	0.889	0.884
	13:9:3	88.430	0.889	0.884
	13:10:3	88.430	0.889	0.884
	4:1:3	74.380	0.614	0.744
	4:2:3	**88.430**	0.885	0.884
CS (4 attributes)	4:3:3	88.430	0.885	0.884
	4:4:3	87.603	0.877	0.876
	4:5:3	86.777	0.868	0.868

จากตาราง 37 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.2 จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ ค่าความถูกต้องสูงสุดอยู่ที่ 79.339% (22:10:3) จำนวน 13 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 88.430% (13:8:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 88.430% (4:2:3)

ตาราง 38 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.3 สาขาวิชาวิทยาการคอมพิวเตอร์

	อัตราการเรียนรู้=0.3			
	Model	**ค่าความถูกต้อง (%)**	**ค่าความแม่นยำ**	**ค่าความระลึก**
CS (22 attributes)	22:10:3	77.686	0.777	0.777
	22:11:3	**80.165**	0.802	0.802
	22:12:3	79.339	0.793	0.793
	22:13:3	79.339	0.793	0.793
	22:14:3	79.339	0.793	0.793
CS (13 attributes)	13:6:3	**88.430**	0.889	0.884
	13:7:3	87.603	0.883	0.876
	13:8:3	88.430	0.889	0.884
	13:9:3	88.430	0.889	0.884
	13:10:3	88.430	0.889	0.884
CS (4 attributes)	4:1:3	75.207	0.623	0.752
	4:2:3	**88.430**	0.885	0.884
	4:3:3	88.430	0.885	0.884
	4:4:3	87.603	0.877	0.876
	4:5:3	86.777	0.868	0.868

จากตาราง 38 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.3 จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ ค่าความถูกต้องสูงสุดอยู่ที่ 80.165% (22:11:3) จำนวน 13 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 88.430% (13:6:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 88.430% (4:2:3)

ตาราง 39 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.1 สาขาวิชาเทคโนโลยีสารสนเทศ

	อัตราการเรียนรู้=0.1			
	Model	**ค่าความถูกต้อง (%)**	**ค่าความแม่นยำ**	**ค่าความระลึก**
	22:10:3	**80.000**	0.806	0.8
	22:11:3	78.261	0.79	0.783
IT (22 attributes)	22:12:3	77.391	0.779	0.774
	22:13:3	78.261	0.79	0.783
	22:14:3	79.130	0.798	0.791
	13:6:3	**78.261**	0.788	0.783
	13:7:3	77.391	0.78	0.774
IT (13 attributes)	13:8:3	78.261	0.792	0.783
	13:9:3	78.261	0.792	0.783
	13:10:3	78.261	0.79	0.783
	4:1:3	78.261	0.688	0.783
	4:2:3	**86.087**	0.859	0.861
IT (4 attributes)	4:3:3	85.217	0.85	0.852
	4:4:3	85.217	0.85	0.852
	4:5:3	85.217	0.85	0.852

จากตาราง 39 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.1 จำนวน 22 ตัวแปรต้นของสาขาวิชาเทคโนโลยีสารสนเทศ ค่าความถูกต้องสูงสุดอยู่ที่ 80.000% (22:10:3) จำนวน 13 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 78.261% (13:6:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 86.087% (4:2:3)

สำนักงานคณะกรรมการวิจัยแห่งชาติ
ห้องสมุดงานวิจัย

ตาราง 40 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.2 สาขาวิชาเทคโนโลยีสารสนเทศ

	อัตราการเรียนรู้=0.2			
	Model	**ค่าความถูกต้อง (%)**	**ค่าความแม่นยำ**	**ค่าความระลึก**
IT (22 attributes)	22:10:3	79.130	0.796	0.791
	22:11:3	77.391	0.777	0.774
	22:12:3	78.261	0.785	0.783
	22:13:3	**80.000**	0.804	0.8
	22:14:3	79.130	0.796	0.791
IT (13 attributes)	13:6:3	76.522	0.768	0.765
	13:7:3	76.522	0.771	0.765
	13:8:3	79.130	0.797	0.791
	13:9:3	78.261	0.786	0.783
	13:10:3	**80.000**	0.802	0.8
IT (4 attributes)	4:1:3	78.261	0.688	0.783
	4:2:3	84.348	0.842	0.843
	4:3:3	**86.957**	0.87	0.87
	4:4:3	86.957	0.87	0.87
	4:5:3	86.957	0.87	0.87

จากตาราง 40 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.2 จำนวน 22 ตัวแปรต้นของสาขาวิชาเทคโนโลยีสารสนเทศ ค่าความถูกต้องสูงสุดอยู่ที่ 80.000% (22:13:3) จำนวน 13 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 80.000% (13:10:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 86.957% (4:3:3)

ตาราง 41 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.3 สาขาวิชาเทคโนโลยีสารสนเทศ

	อัตราการเรียนรู้=0.3			
	Model	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
	22:10:3	79.130	0.796	0.791
	22:11:3	77.391	0.777	0.774
IT (22 attributes)	22:12:3	77.391	0.781	0.774
	22:13:3	**80.000**	0.804	0.8
	22:14:3	78.261	0.787	0.783
	13:6:3	76.522	0.768	0.765
	13:7:3	76.522	0.774	0.765
IT (13 attributes)	13:8:3	76.522	0.766	0.765
	13:9:3	**77.391**	0.78	0.774
	13:10:3	77.391	0.777	0.774
	4:1:3	78.261	0.688	0.783
	4:2:3	84.348	0.842	0.843
IT (4 attributes)	4:3:3	85.217	0.853	0.852
	4:4:3	**86.087**	0.862	0.861
	4:5:3	86.087	0.862	0.861

จากตาราง 41 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ = 0.3 จำนวน 22 ตัวแปรต้นของสาขาวิชาเทคโนโลยีสารสนเทศ ค่าความถูกต้องสูงสุดอยู่ที่ 80.000% (22:13:3) จำนวน 13 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 77.391% (13:9:3) จำนวน 4 ตัวแปรต้นค่าความถูกต้องสูงสุดอยู่ที่ 86.087% (4:4:3)

สาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 4 ตัวแปรต้น

การทดสอบประสิทธิภาพการจัดหมวดหมู่ด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับได้ทำการทำการกำหนดอัตราการเรียนรู้จาก 0.1 ถึง 0.3 และปรับจำนวนโหนดของชั้นซ่อน เพื่อเปรียบเทียบค่าความถูกต้องของตัวแบบ ด้วยจำนวนตัวแปร 4 ตัวแปรต้น และ 1 ตัวแปรตาม ดังตาราง 42

ตาราง 42 แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ

Model	อัตราการเรียนรู้	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
4:1:3	0.1	74.576	0.798	0.746
4:1:3	0.2	75.000	0.795	0.75
4:1:3	0.3	76.271	0.802	0.763
4:2:3	0.1	86.864	0.869	0.869
4:2:3	0.2	86.864	0.869	0.869
4:2:3	0.3	86.017	0.86	0.86
4:3:3	0.1	86.864	0.869	0.869
4:3:3	0.2	86.864	0.869	0.869
4:3:3	0.3	86.441	0.864	0.864
4:4:3	0.1	86.864	0.869	0.869
4:4:3	0.2	86.864	0.868	0.869
4:4:3	0.3	89.831	0.900	0.898
4:5:3	0.1	86.441	0.864	0.864
4:5:3	0.2	88.136	0.881	0.881
4:5:3	0.3	88.983	0.89	0.89

จากตาราง 42 จะเห็นได้ว่า ถ้าทำการกำหนดอัตราการเรียนรู้ 0.1-0.3 จำนวน 4 ตัวแปรต้นของนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศ ค่าความถูกต้องสูงสุดอยู่ที่ 89.831% (4:4:3)

4. เทคนิคต้นไม้ตัดสินใจ

ข้อมูลชุดฝึกสอน

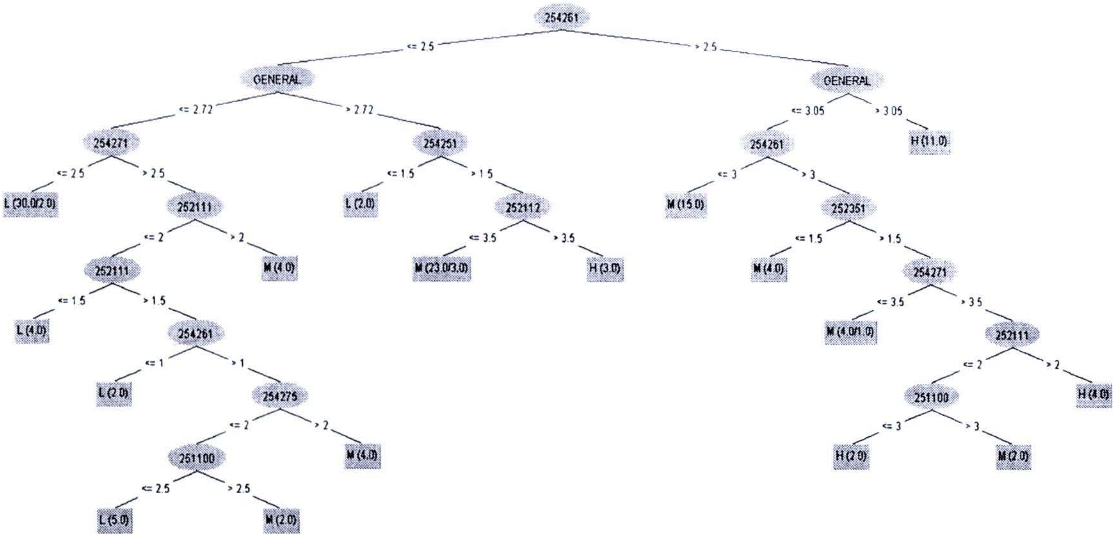
การทดสอบประสิทธิภาพการจัดหมวดหมู่ด้วยเทคนิคต้นไม้ตัดสินใจได้ทำการหาค่าความถูกต้อง ค่าความแม่นยำ และค่าความระลึก ดังตาราง 43

ตาราง 43 แสดงผลทดลองของโมเดลด้วยเทคนิคต้นไม้ตัดสินใจ

	ค่าความถูกต้อง (%)	ค่าความแม่นยำ	ค่าความระลึก
CS (22 attributes)	65.289	0.658	0.653
IT (22 attributes)	77.391	0.772	0.774
CS (13 attributes)	69.422	0.703	0.694
IT (13 attributes)	76.522	0.766	0.765
CS (4 attributes)	70.248	0.71	0.702
IT (4 attributes)	82.609	0.828	0.826
CS+IT (4 attributes)	79.237	0.798	0.792

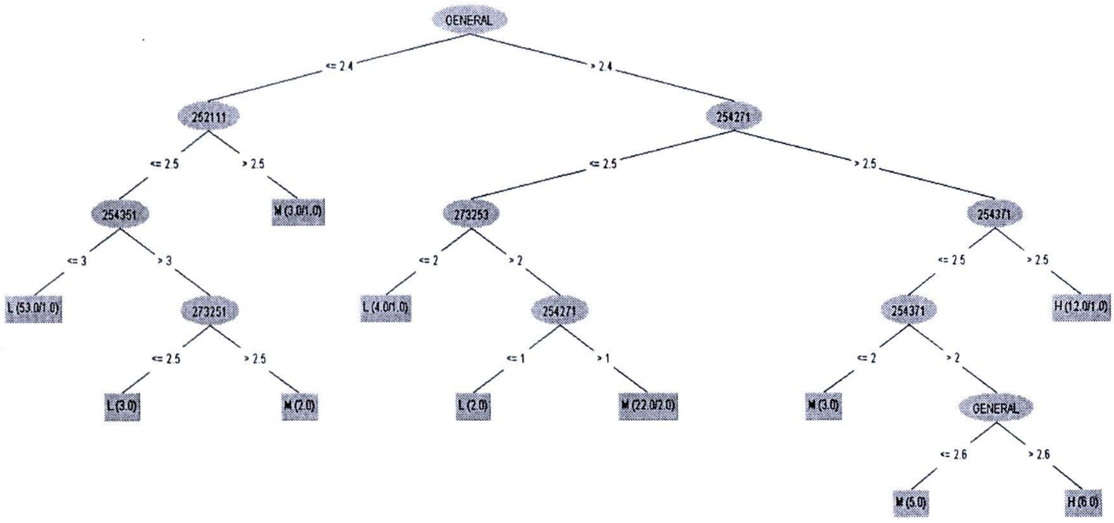
จากตาราง 43 จะเห็นได้ว่า จำนวน 22 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 65.289% และ 77.391% จำนวน 13 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 69.422% และ 76.522% จำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศค่าความถูกต้องอยู่ที่ 70.248% และ 82.609% และเมื่อนำจำนวน 4 ตัวแปรต้นของสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศมารวมกันค่าความถูกต้องอยู่ที่ 79.237% จะเห็นได้ว่าค่าความถูกต้องหลังจากการนำ 2 สาขาวิชามารวมกัน ค่าความถูกต้องที่ได้ก็ยังค่อนข้างสูง

ภาพ 36 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ จำนวน 13 ตัวแปรต้น



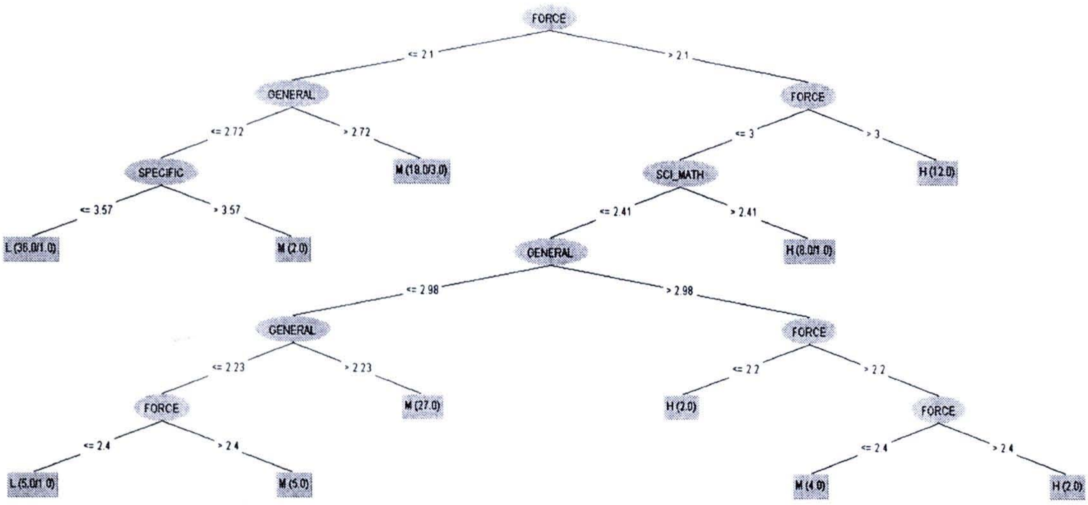
ภาพ 36 แสดงต้นไม้ตัดสินใจสาขาวิชาวิทยาการคอมพิวเตอร์ จำนวน 13 ตัวแปรต้น

ภาพ 37 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 13 ตัวแปรต้น



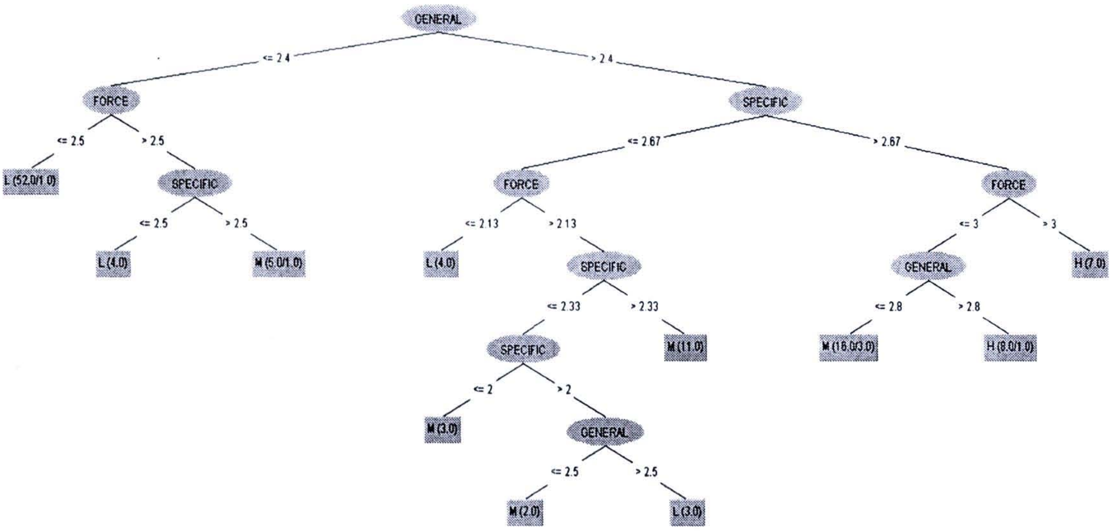
ภาพ 37 แสดงต้นไม้ตัดสินใจสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 13 ตัวแปรต้น

ภาพ 38 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ จำนวน 4 ตัวแปรต้น



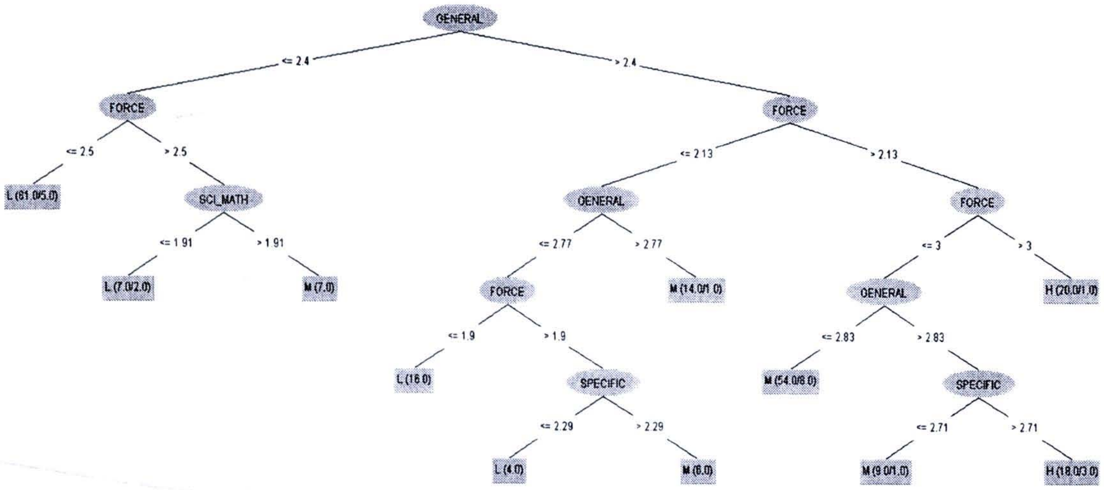
ภาพ 38 แสดงต้นไม้ตัดสินใจสาขาวิชาวิทยาการคอมพิวเตอร์ จำนวน 4 ตัวแปรต้น

ภาพ 39 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 4 ตัวแปรต้น



ภาพ 39 แสดงต้นไม้ตัดสินใจสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 4 ตัวแปรต้น

ภาพ 40 เป็นโครงสร้างต้นไม้ตัดสินใจของการพยากรณ์ผลสัมฤทธิ์ทางการเรียนใช้ข้อมูลนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศ จำนวน 4 ตัวแปรต้น



ภาพ 40 แสดงต้นไม้ตัดสินใจ จำนวน 4 ตัวแปรต้น ของสองสาขาวิชา

ในตาราง 44 และ 45 เป็นการแสดงการคัดเลือกตัวแปรด้วยค่าสัมประสิทธิ์สหสัมพันธ์และสัมประสิทธิ์อีต้า และทำการประเมินตัวแบบจากตัวแปรที่ได้จากการคัดเลือก จากข้อมูลชุดฝึกสอน

ตาราง 44 แสดงค่าความถูกต้องของการจำแนกข้อมูลชุดฝึกสอน

		KNN	Naïve Bayes	BPNN	Decision Tree
ก่อนการคัดเลือก	CS (22 attributes)	66.116%	79.339%	80.992%	65.289%
ตัวแปร	IT (22 attributes)	67.826%	80.000%	80.000%	77.391%
หลังการคัดเลือก	CS (13 attributes)	73.554%	82.645%	88.430%	69.422%
ตัวแปร	IT (13 attributes)	83.478%	85.217%	80.000%	76.522%

ตาราง 45 แสดงค่าความถูกต้องของการจำแนกข้อมูลชุดฝึกสอนที่รวมสาขา

		KNN	Naïve Bayes	BPNN	Decision Tree
แยกสาขาวิชา	CS (4 attributes)	80.992%	84.298%	89.256%	70.248%
	IT (4 attributes)	79.130%	86.957%	86.957%	82.609%
รวมสาขาวิชา	CS+IT (4 attributes)	82.203%	86.441%	89.831%	79.237%

นำตัวแปรที่ผ่านการคัดเลือกตัวแปรด้วยค่าสัมประสิทธิ์สหสัมพันธ์ และสัมประสิทธิ์อีต้า และนำมาหาค่าเฉลี่ยตามกลุ่มวิชา จำนวนตัวแปรต้น 4 ตัวแปร ได้แก่ หมวดวิชาศึกษาทั่วไป กลุ่มวิชาพื้นฐานวิทยาศาสตร์และคณิตศาสตร์ กลุ่มวิชาพื้นฐานเฉพาะด้าน และกลุ่มวิชาเอกบังคับ และตัวแปรตาม 1 ตัวแปร คือ ผลสัมฤทธิ์ทางการเรียน มาใช้กับเทคนิคเหมืองข้อมูล 4 เทคนิค เพื่อเปรียบเทียบประสิทธิภาพการทำงาน ได้ค่าความถูกต้อง ค่าความระลึก และค่าความแม่นยำ ดังตาราง 46 ได้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของข้อมูลชุดฝึกสอนและข้อมูลชุดทดสอบ ดังตาราง 47

ผลการวิจัยพบว่า เทคนิคเนอีฟเบย์ และ เทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับให้ค่าความถูกต้องในการพยากรณ์ใกล้เคียงกัน และสูงกว่าเทคนิคเคเนียเรสเนเบอร์ และเทคนิคต้นไม้ตัดสินใจ ให้ค่าความถูกต้องต่ำสุด

จากค่าความถูกต้อง และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง แสดงให้เห็นว่าเทคนิคที่เหมาะสมที่จะนำมาสร้างตัวแบบ คือ เทคนิคเนอีฟเบย์ และเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับร่วมกัน เพราะให้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่ต่ำ และให้ค่าความถูกต้องค่อนข้างสูง มีความน่าเชื่อถือสูงเพราะนำเทคนิคที่มีความแม่นยำสูงทั้ง 2 เทคนิค มาร่วมกันพยากรณ์

ตาราง 46 แสดงการเปรียบเทียบประสิทธิภาพของข้อมูลชุดฝึกสอน และข้อมูลชุดทดสอบ

	Training Data			Testing Data		
	Accuracy (%)	Precision	Recall	Accuracy (%)	Precision	Recall
K-NN	82.203%	0.821	0.822	75.397%	0.766	0.754
Naïve Bayes	**86.441%**	0.864	0.864	**78.571%**	0.795	0.786
BPNN	**89.831%**	0.900	0.898	**75.397%**	0.766	0.754
Decision Tree	79.237%	0.798	0.792	69.047%	0.709	0.690

จากตาราง 46 จะเห็นว่าเทคนิคเนอีฟเบย์ และเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ ให้ค่าความถูกต้องในการพยากรณ์ใกล้เคียงกัน และสูงกว่าเทคนิคเคเนียเรสเนเบอร์ และเทคนิคต้นไม้ตัดสินใจ ให้ค่าความถูกต้องต่ำสุด ทั้งข้อมูลชุดฝึกสอน และข้อมูลชุดทดสอบ

ตาราง 47 แสดงการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของข้อมูล

	RMSE (Training Data)	RMSE (Testing Data)
Naïve Bayes	0.236	0.162
BPNN	0.229	0.172
Combine 2 algorithms	**0.224**	**0.157**

จากตาราง 47 จะเห็นว่าค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของเทคนิคเนอีฟเบย์ และเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับแบบทำงานร่วมกัน ให้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่ต่ำ

นำตัวแบบที่ได้สร้างขึ้นมาใช้ประโยชน์จริงในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของนิสิต หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศ ในรุ่นต่อ ๆ ไป ดังตาราง 48

ตาราง 48 แสดงการเปรียบเทียบประสิทธิภาพของตัวแบบของข้อมูลชุดตรวจสอบ

	Validation Data			
	Accuracy (%)	Precision	Recall	RMSE
Naïve Bayes	78.740%	0.795	0.787	0.166
BPNN	76.378%	0.776	0.764	0.174
Combine 2 algorithms	-	-	-	**0.158**

จากตาราง 48 แสดงให้เห็นว่าเทคนิคที่เหมาะสมที่จะนำมาสร้างตัวแบบ คือ เทคนิคเนอีฟเบย์ และเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับแบบทำงานร่วมกัน เพราะให้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่ต่ำ มีความน่าเชื่อถือสูงเพราะนำเทคนิคที่มีความแม่นยำสูงทั้ง 2 เทคนิค มาร่วมกันพยากรณ์