

บทที่ 1

บทนำ

เอกสารฉบับนี้จัดทำเพื่อประกอบวิทยานิพนธ์เรื่อง การวิเคราะห์การสำเร็จการศึกษาของนักศึกษา ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ โดยระบบอนุมานนิรนัยวิธีแบบปรับตัวได้ โดยมีรายละเอียดของหลักการ ทฤษฎี เหตุผล และสมมุติฐาน วัตถุประสงค์ของงานวิจัย ประโยชน์ที่จะได้รับจากการศึกษาเชิงประยุกต์ ขอบเขตของโครงการ วิธีการวิจัย และสถานที่ที่ใช้ในการดำเนินงานวิจัย มีรายละเอียดดังต่อไปนี้

1.1 หลักการ ทฤษฎี เหตุผล และสมมุติฐาน

ปัญหาการฟื้นฟูสภาพหรือการสำเร็จการศึกษาช้ากว่าหลักสูตรของนักศึกษาก่อให้เกิดการสูญเสียต้นทุนด้านการเงินและเวลาของทั้งสถาบันศึกษาและนักศึกษา ทำให้ต้องมีการติดตามความก้าวหน้าและความเสี่ยงต่อปัญหาดังกล่าวด้วยเกรดเฉลี่ยสะสมที่กว้างแผนทางวิชาการใช้เป็นปัจจัยร่วม เพื่อประเมินผล และติดตามความก้าวหน้าของนักศึกษา [1][2] แต่เกรดเฉลี่ยสะสมเพียงปัจจัยเดียวไม่สามารถวิเคราะห์สถานภาพของนักศึกษาในอนาคตได้อย่างเหมาะสม เนื่องจากมีปัจจัยอื่นที่เกี่ยวข้อง เช่น เกรดเฉลี่ยสะสมของกลุ่มวิชาเอก กลุ่มวิชาปัญหาพิเศษ หน่วยกิตรวมของแต่ละกลุ่มวิชา เป็นต้น ซึ่งข้อมูลนักศึกษาในฐานะข้อมูลที่มีการเก็บอย่างต่อเนื่องเป็นจำนวนมากได้มีการนำมาวิเคราะห์ด้วยระบบคอมพิวเตอร์ เช่น การใช้เทคนิคทางสถิติแบบเดิมร่วมกับการจัดการด้านฐานข้อมูล แต่ไม่สามารถใช้กับชุดข้อมูลที่มีขนาดใหญ่ มีความซับซ้อน และข้อมูลที่ไม่ชัดเจน ทำให้เกิดการปรับปรุงเทคนิคและเครื่องมือสำหรับการวิเคราะห์ในชุดข้อมูลขนาดใหญ่ หนึ่งในเทคนิคนั้นคือ การทำเหมืองข้อมูล (Data mining) ที่ได้รับความสนใจจากนักวิจัยหลายแขนง เช่น การออกแบบฐานข้อมูล สถิติ การจดจำภาพแบบ กลไกการเรียนรู้ เป็นต้น [3] เมื่อนำมาประยุกต์ใช้กับการศึกษาเรียกว่า การทำเหมืองข้อมูลทางการศึกษา (Educational data mining) เพื่อให้ครอบคลุมปัญหาด้านการศึกษามากขึ้น และสามารถแก้ปัญหาที่มีความเฉพาะได้โดยตรง [4] การทำเหมืองข้อมูลทางการศึกษาจะเกี่ยวข้องกับการสกัดสิ่งที่เป็นประโยชน์หรือภาพแบบใหม่มาจากฐานข้อมูลด้านการศึกษานี้ เพื่อทำให้เกิดความเข้าใจ การปรับปรุงประสิทธิภาพการศึกษา และการประเมินกระบวนการเรียนรู้ของ

นักศึกษาที่ดีขึ้น [5][6] หนึ่งในวิธีการสร้างโมเดลที่ได้รับความนิยมคือ โครงข่ายประสาทเทียม [7] ที่มีแบบขนานทำให้มีการประมวลผลที่เร็ว แต่ผลลัพธ์ที่ได้ไม่สามารถอธิบายในลักษณะของเหตุและผลได้ ส่วนระบบฟัซซี่สามารถอธิบายเหตุและผลที่สอดคล้องกับตรรกะความคิดของมนุษย์ เพราะมีการตีความในภาพถ้า-แล้ว (What-if analysis) แต่ไม่มีกระบวนการเรียนรู้ในการปรับแต่งโครงสร้างของกฎและตัวแปรของระบบ จึงต้องอาศัยผู้เชี่ยวชาญ เมื่อนำมาผสมผสานกันเป็นการเพิ่มข้อดีและลดข้อเสีย เรียกว่า ระบบนิวโรฟัซซี่ (Neuro-fuzzy) ระบบผสมนิวโรฟัซซี่ที่นิยมคือ ระบบอนุมานนิวโรฟัซซี่แบบปรับตัวได้ (Adaptive Neuro-Fuzzy Inference Systems, ANFIS) [8] ที่มีโครงสร้างแบบหลายอินพุตหนึ่งเอาต์พุต เทคนิคนี้จะใช้ในการทำนายเกรดเฉลี่ยสะสมของแต่ละหมวดหมู่หรือกลุ่มวิชาที่ได้มาจากการศึกษาเอกสารข้อบังคับมหาวิทยาลัยเชียงใหม่ ว่าด้วยการศึกษาระดับปริญญาตรี พ.ศ. 2550 และเอกสาร โครงสร้างหลักสูตรของหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ซึ่งตัวแปรอินพุต คือ เกรดเฉลี่ยสะสมของหมวดหมู่หรือกลุ่มวิชาในแต่ละภาคเรียนที่ไม่มีความแน่นอนหรือชัดเจนในการจัดกลุ่มข้อมูล ส่วนตัวแปรเอาต์พุต คือ เกรดเฉลี่ยสะสมของหมวดหมู่หรือกลุ่มวิชาในภาคเรียนที่ 11 จากนั้นจะนำเกรดเฉลี่ยที่ได้มาทำนายค่าความน่าจะเป็นของการสำเร็จการศึกษา

วัตถุประสงค์ของการศึกษานี้จึงมุ่งเน้นหาผลการสำเร็จการศึกษาด้วยระบบอนุมานนิวโรฟัซซี่แบบปรับตัวได้ และศึกษาเฉพาะ นักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ โดยใช้ข้อมูลลงทะเบียนเรียนของนักศึกษาที่เข้าศึกษาตั้งแต่ปีการศึกษา 2548 ถึง 2551 ในหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ปี 2548 ถึง 2551 ที่มีสถานภาพสำเร็จการศึกษาและพ้นสภาพเนื่องจากเกรดเฉลี่ยต่ำกว่าเกณฑ์ที่กำหนด

1.2 วัตถุประสงค์ของการวิจัย

การศึกษานี้มีวัตถุประสงค์เพื่อทำนายความน่าจะเป็นของการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ ด้วยระบบอนุมานนิวโรฟัซซี่แบบปรับตัวได้

1.3 ประโยชน์ที่จะได้รับจากการศึกษาเชิงประยุกต์

การศึกษานี้จะได้เกรดเฉลี่ยสะสมของแต่ละหมวดหมู่หรือกลุ่มวิชาในภาคเรียนที่ 11 หรือชั้นปีที่ 4 ภาคเรียนที่ 2 และ ค่าความน่าจะเป็นในแต่ละกลุ่มสถานภาพและชั้นปีที่สำเร็จการศึกษา ของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ สามารถใช้เป็นแนวทางในการแนะนำนักศึกษาได้

1.4 ขอบเขตของงานวิจัย

1.4.1 ขอบเขตข้อมูลที่ใช้ในงานวิจัย ข้อมูลที่นำมาใช้ในงานวิจัยเป็นข้อมูลนักศึกษาที่เข้าเรียนตั้งแต่ปีการศึกษา 2548 ถึง 2551 ในหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ เนื่องจากอยู่ในโครงสร้างหลักสูตรเดียวกัน ซึ่งมีสถานภาพสำเร็จการศึกษาหรือพ้นสภาพเนื่องจากผลการเรียนต่ำกว่าเกณฑ์เท่านั้น มีข้อมูลนักศึกษาทั้งหมด 323 คน ส่วนข้อมูลลงทะเบียนเรียนจะเลือกใช้ข้อมูลตั้งแต่ภาคเรียนที่ 1 จนถึงภาคเรียนที่ 4 หรือภาคเรียนที่ 5 และได้รับเกรด A, B+, B, C+, C, D+, D และ F เท่านั้น จากการศึกษาเอกสารข้อบังคับมหาวิทยาลัยเชียงใหม่ ว่าด้วยการศึกษาระดับปริญญาตรี เงื่อนไขที่เกี่ยวข้องกับการสำเร็จการศึกษานอกจากเกรดเฉลี่ยสะสมแล้วมีหน่วยกิต และเกรดเฉลี่ยสะสมของหมวดหมู่หรือกลุ่มวิชาย่อยตามโครงสร้างหลักสูตร ดังนั้นจึงใช้เป็นตัวแปรอินพุต ส่วนตัวแปรเอาต์พุตเป็นผลการสำเร็จการศึกษา แบ่งเป็น 6 แบบ ซึ่งแบ่งตามชั้นปีที่จบการศึกษา และตามสถานภาพ

1.4.2 วิธีการวิจัย

- 1) ศึกษาเอกสารข้อบังคับมหาวิทยาลัยเชียงใหม่ ว่าด้วยการศึกษาระดับปริญญาตรี และเอกสารโครงสร้างหลักสูตรของหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ เพื่อหาเงื่อนไขการสำเร็จการศึกษา และสรุปเงื่อนไขการสำเร็จการศึกษา
- 2) เตรียมข้อมูลด้วยการคัดเลือกข้อมูลนักศึกษา คัดเลือกข้อมูลลงทะเบียนเรียนตัดค่าว่างเติมค่าว่างด้วยศูนย์ และจัดข้อมูลให้มีรูปแบบเดียวกัน
- 3) แบ่งชุดตัวแปรอินพุต เพื่อหาตัวแปรอินพุต
- 4) คัดเลือกตัวแปรอินพุตและตัวแปรเอาต์พุต
- 5) แบ่งข้อมูลเรียนรู้และข้อมูลทดสอบ
- 6) สร้างแบบจำลองเพื่อทำนายเกรดเฉลี่ยสะสมภาคเรียนที่ 11 หรือ ชั้นปี 4 ภาคเรียนที่ 2 ในแต่ละหมวดหมู่หรือกลุ่มวิชา ด้วยระบบอนุมานนิเวศวิธีแบบปรับตัวได้
 - 6.1) จัดกลุ่มข้อมูลเริ่มต้นด้วยเทคนิคการจัดกลุ่มแบบลอบออกและพีชชีมีน
 - 6.2) เรียนรู้แบบผสมผสานระหว่างแบบไปข้างหน้าด้วยวิธีประเมินกำลังสองน้อยสุด และแบบย้อนกลับด้วยการเคลื่อนลาดชันลงมา เพื่อปรับพารามิเตอร์
 - 6.3) ปรับพารามิเตอร์ของเทคนิคการจัดกลุ่มและพารามิเตอร์ของการเรียนรู้

- 7) วัดประสิทธิภาพแบบจำลองทำนายเกรดเฉลี่ยสะสมด้วยระบบอนุมานนิวโรฟัซซีแบบปรับตัวได้จากค่าความคลาดเคลื่อนกำลังสองน้อยสุด ถ้ามีค่าน้อยสุดจะเป็นแบบจำลองที่เหมาะสมที่สุด
- 8) นำเกรดเฉลี่ยสะสมของแต่ละหมวดหมู่หรือกลุ่มวิชาจากแบบจำลองที่สร้างจากระบบอนุมานนิวโรฟัซซีแบบปรับตัวได้ มาทำนายค่าความน่าจะเป็นของการสำเร็จการศึกษาในแต่ละกลุ่มเอทพุตที่กำหนดไว้ โดยมีเทคนิคดังนี้
 - 8.1) โครงข่ายการรู้จำแบบ เป็นโครงข่ายการเรียนรู้แบบไปข้างหน้า 2 ชั้นที่ใช้ฟังก์ชันการถ่ายโอนเป็นฟังก์ชันซอฟต์แมกซ์
 - 8.2) การเรียนรู้แบบเบย์อย่างง่ายจะใช้ค่าความน่าจะเป็นของข้อมูลก่อนหน้าโดยใช้กฎของเบย์ ทำนายค่าความน่าจะเป็นของแต่ละกลุ่ม
 - 8.3) ต้นไม้ตัดสินใจจะแบ่งข้อมูลของแต่ละตัวแปรอินพุตออกเป็นกลุ่มจากการคัดเลือกคุณสมบัติข้อมูล และคัดเลือกคุณสมบัติที่ดีที่สุดมาเป็นโหนดราก ต่อมาจะสร้างโหนดลูก และวนสร้างโหนดลูกและต้นไม้ย่อยของแต่ละกิ่งไปเรื่อยๆ
 - 8.4) ความถดถอยโลจิสติกพหุกลุ่ม เป็นเป็นการหาความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามด้วยสมการโลจิสติก
- 9) วัดประสิทธิภาพของแบบจำลองทำนายความน่าจะเป็นด้วยค่าความถูกต้อง ค่าความระลึก ค่าความแม่นยำ ค่าความเหวี่ยง และคะแนนไบรเออ
- 10) สรุปผลการทดลอง

1.4.3 เครื่องมือที่ใช้ในงานวิจัย

- 1) โปรแกรมจำลองทางคณิตศาสตร์ สำหรับสร้างแบบจำลอง
- 2) โปรแกรมไมโครซอฟท์แอ็กเซส เวอร์ชัน 2010 (Microsoft access 2010) สำหรับเก็บและเตรียมข้อมูล

1.4.4 สถานที่ที่ใช้ในการดำเนินงานวิจัย

- 1) ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่
- 2) หอสมุด มหาวิทยาลัยเชียงใหม่