# บทที่ 1

# บทนำ

## ความเป็นมาและความสำคัญของปัญหา

การพยากรณ์ (Forecasting) หมายถึง การคาดคะเน หรือการทำนายการเกิดเหตุการณ์ หรือสภาพการณ์ต่าง ๆ ในอนาคต โดยอาศัยข้อมูล ประสบการณ์ ความรู้ความสามารถของผู้พยากรณ์ที่เกิดขึ้นในอดีต มาทำการศึกษาแนวโน้มหรือรูปแบบของการเกิดเหตุการณ์ในอนาคต ซึ่งต้องอาศัยการพิจารณาอย่างละเอียดถี่ถ้วน การพยากรณ์จึงมีบทบาท และเป็นองค์ประกอบที่สำคัญของการตัดสินใจเกี่ยวกับการจัดการเพื่อให้บรรลุผลสำเร็จ ทั้งในเรื่องของวัตถุประสงค์และเป้าหมายของการปฏิบัติงาน ซึ่งวัตถุประสงค์หลักของการพยากรณ์คือ ต้องการนำข้อมูลเพื่อนำมาใช้ในวางแผนและตัดสินใจเกี่ยวกับการดำเนินงาน (ทรงศิริ แต้สมบัติ, 2539) ทั้งนี้เพราะการวางแผนและการตัดสินใจต่างก็เกี่ยวข้องกับเหตุการณ์ในอนาคต ซึ่งโดยทั่วไปเหตุการณ์ในอนาคตเป็นสิ่งที่ควบคุมไม่ได้ เพราะฉะนั้นการพยากรณ์เหตุการณ์ในอนาคตจึงมีความจำเป็นอย่างมากที่ผู้บริหารระดับต่างๆ จะนำมาใช้เป็นเครื่องมือในการวางแผนและการตัดสินใจเช่น ในระบบการค้าปลีก การพยากรณ์ยอดขายสินค้าจะช่วยในการวางแผนส่งเสริมการจำหน่ายสินค้าให้กับลูกค้าได้อย่างมีประสิทธิภาพและสอดคล้องกับสถานการณ์ที่คาดว่าจะเกิดขึ้นในอนาคต ซึ่งนอกจากจะพยากรณ์ยอดขายสินค้าแล้ว ยังเป็นเครื่องมือที่ช่วยในการควบคุมและรักษาส่วนแบ่งตามตลาดให้มีความต่อเนื่องอีกด้วย (สมเกียรติ เกตุเอี่ยม, 2546) Wheelwriht และ Makridakis (1980) ได้จำแนกหน้าที่ทางด้านบริหารที่จำเป็นต้องใช้การพยากรณ์ดังนี้ การวางแผนทางด้านการผลิต การวางแผนการจัดการด้านสินค้าคงคลัง การจัดซื้อ การตลาด การเงินและบัญชี บุคลากร และการวางนโยบาย นั่นแสดงว่า การพยากรณ์มีบทบาทอย่างมากในสังคมและหากการพยากรณ์มีความใกล้เคียงกับความเป็นจริงที่จะเกิดขึ้นในอนาคต ย่อมส่งผลทำให้การวางแผนและการตัดสินใจเป็นไปอย่างมีประสิทธิภาพ การพัฒนาก็จะเป็นไปด้วยความก้าวหน้าและมั่นคงต่อไป

การพยากรณ์แบ่งออกเป็น 2 ประเภท ประเภทที่ 1 คือ การพยากรณ์เชิงคุณภาพ (Qualitative forecasting or Judgmental forecasting) เป็นการพยากรณ์ที่ต้องอาศัยวิจารณญาณ ประสบการณ์ ความรู้ความสามารถ หรือวิสัยทัศน์ของผู้พยากรณ์ หรือของผู้เชี่ยวชาญที่จะพยากรณ์เรื่องใดเรื่องหนึ่ง โดยเฉพาะการพยากรณ์เชิงคุณภาพ ส่วนใหญ่จึงไม่มีรูปแบบ กฎเกณฑ์และสูตรที่แน่นอนในการ

พยากรณ์ จึงไม่ต้องอาศัยตัวแบบทางคณิตศาสตร์ในการพยากรณ์ การพยากรณ์เชิงคุณภาพจัดว่าเป็นการพยากรณ์อย่างเป็นทางการ เนื่องจากต้องอาศัยการคิดอย่างเป็นระบบของผู้พยากรณ์ หรือของผู้เชี่ยวชาญ สำหรับประเภทที่ 2 ของการพยากรณ์คือ การพยากรณ์เชิงปริมาณ (Quantitative forecasting) เป็นการพยากรณ์ที่ต้องอาศัยความรู้ทางด้านคณิตศาสตร์และสถิติไปสร้างรูปแบบหรือสมการพยากรณ์ เพื่อจะพยากรณ์ข้อมูลหรือเหตุการณ์ในอนาคต ดังนั้น การพยากรณ์แบบนี้จะต้องมีข้อมูลในอดีตที่ผ่านมา ซึ่งข้อมูลในอดีตจะต้องอยู่ในรูปของตัวเลข หรือสามารถแปลงให้เป็นตัวเลขได้ และจะต้องมีปริมาณมากพอสมควร ถึงจะทำให้การพยากรณ์แบบนี้มีความน่าเชื่อถือ ซึ่งการพยากรณ์เชิงปริมาณเป็นที่นิยมมากกว่าการพยากรณ์เชิงคุณภาพ

การพยากรณ์เชิงปริมาณถูกแบ่งออกเป็น 2 ประเภท ตามปรัชญาของการพยากรณ์ดังนี้ *ประเภทที่ 1* เป็นวิธีการที่มีแนวคิดว่าพฤติกรรมในอดีตของสิ่งที่จะพยากรณ์ ควรจะเพียงพอที่จะพยากรณ์พฤติกรรมในอนาคตของมันเองได้ (วิชิต หล่อจีระชุณห์กุล, 2539) วิธีการพยากรณ์ในประเภทนี้ได้แก่ วิธีการทำให้เรียบ (Smoothing Method) วิธีเฉลี่ยเคลื่อนที่ (Moving Average Method) วิธีแยกส่วนประกอบ (Decomposition Method) การวิเคราะห์แนวโน้ม (Trend Analysis) และวิธีของบ๊อกซ์และเจ็นกิ้นส์ (Box-Jenkins Method) ส่วนการพยากรณ์เชิงปริมาณ *ประเภทที่ 2* เป็นวิธีการที่มีแนวคิดว่า พฤติกรรมของสิ่งที่จะพยากรณ์ถูกกำหนดขึ้นโดยสิ่งที่อยู่โดยรอบ ซึ่งมีความสัมพันธ์บางอย่างกับสิ่งที่จะพยากรณ์ วิธีการพยากรณ์ในประเภทนี้ได้แก่ การวิเคราะห์การถดถอย (Regression Analysis) การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) และการวิเคราะห์ทางเศรษฐมิติ (Econometric forecasting) การพยากรณ์เชิงปริมาณทั้งสองประเภทมีความขัดแย้งในแนวคิดพื้นฐาน และไม่สามารถสรุปได้ว่าประเภทใดดีกว่า การพยากรณ์เชิงปริมาณทั้งสองประเภทเป็นที่ยอมรับในวงการศึกษาและธุรกิจอุตสาหกรรม การพยากรณ์เชิงปริมาณประเภทที่ 1 จะได้เปรียบประเภทที่ 2 ในด้านข้อมูล การพยากรณ์เชิงปริมาณประเภทที่ 1 ใช้ข้อมูลในอดีตของสิ่งที่พยากรณ์เท่านั้น ไม่ได้ใช้ข้อมูลอย่างอื่น ส่วนการพยากรณ์ประเภทที่ 2 จะต้องใช้ข้อมูลของสิ่งที่อยู่โดยรอบซึ่งคาดว่ามีผลกระทบต่อสิ่งที่จะพยากรณ์ที่นอกเหนือไปจากข้อมูลในอดีตของสิ่งที่จะพยากรณ์ ดังนั้น จำนวนข้อมูลที่ใช้ในการพยากรณ์เชิงปริมาณประเภทที่ 1 จึงน้อยกว่าและเป็นข้อจำกัดในการเลือกระเบียบวิธีการพยากรณ์ ในการตัดสินใจที่จะลือกใช้วิธีพยากรณ์ใดนั้น ยังมีสิ่งที่ต้องพิจารณานอกเหนือจากประเภทของข้อมูล เนื่องจากระเบียบวิธีการแต่ละวิธีมีขีดความสามารถที่จำกัด อาจไม่เหมาะสมที่จะนำไปใช้ในสถานการณ์บางอย่างเช่น การพิจารณาของขอบเวลา ลักษณะของข้อมูล ระดับความรู้ของผู้พยากรณ์ค่าใช้จ่ายและความถูกต้อง (วิชิต หล่อจีระชุณห์กุล, 2546)

ปัจจุบันองค์กรได้นำเอาเทคโนโลยีสารสนเทศมาช่วยในการดำเนินงานขององค์กรอย่างแพร่หลาย องค์กรบางแห่งมีการจัดเก็บข้อมูลในการดำเนินงานของตน บางแห่งมีการจัดเก็บข้อมูลสำหรับการให้บริการแก่องค์กรอื่นๆ นอกจากนี้บางองค์กรยังมีการทำวิจัยเพื่อการศึกษาหาความรู้ใหม่ๆ ดังนั้นปัญหาที่พบของผู้ที่ทำงานคือ การหาทางที่ประมวลผลข้อมูลที่มีอยู่โดยใช้วิธีการเชิงวิทยาศาสตร์หรือวิธีการที่มีแบบแผนให้ได้สารสนเทศที่สามารถนำไปใช้ในรูปแบบที่เหมาะสมและมีคุณค่าต่อการนำไปเพื่อสนับสนุนในการดำเนินงาน การบริหาร และการตัดสินใจ ตลอดจนการนำเสนอในเชิงวิชาการที่ให้ความรู้แก่บุคคลทั่วไป และวิธีการหนึ่งในการวิเคราะห์ข้อมูลคือ การใช้เทคนิควิเคราะห์ทางสถิติเป็นส่วนหนึ่งในการแก้ปัญหาข้างต้น แต่จากการพัฒนาความก้าวหน้าทางเทคโนโลยีสารสนเทศและการจัดเก็บข้อมูลที่มีปริมาณมากทั้งจำนวนตัวแปรและจำนวนรายการของข้อมูล ประกอบกับข้อจำกัดบางประการของการวิเคราะห์เชิงสถิติซึ่งเป็นการวิเคราะห์ข้อมูลแบบ Model-Driven ทำให้การวิเคราะห์ข้อมูลแนวใหม่ที่เรียกว่า การทำเหมืองข้อมูล (Data Mining) ซึ่งเป็นการวิเคราะห์ข้อมูลแบบ Data-Driven โดยที่การทำเหมืองข้อมูลเป็นกระบวนการค้นหารูปแบบ และความสัมพันธ์ที่มีอยู่ในข้อมูลโดยอาศัยขั้นตอนวิธีทางคอมพิวเตอร์ผสมผสานกับวิธีการทางสถิติ ข้อดีของเทคนิคต่างๆในการทำเหมืองข้อมูลคือ ไม่มีข้อตกลงหรือเงื่อนไขเกี่ยวกับข้อมูล และสามารถค้นหารูปแบบและความสัมพันธ์ที่ไม่ใช่เชิงเส้น โดยผู้วิเคราะห์ ไม่จำเป็นต้องระบุตัวแบบด้วยตนเอง ปัจจุบันการทำเหมืองข้อมูลได้ถูกนำมาใช้ในการวิเคราะห์ข้อมูลทั้งทางวิทยาศาสตร์และสังคมศาสตร์โดยประยุกต์ใช้งานต่างๆ เช่น การจัดการความสัมพันธ์กับลูกค้า การพยากรณ์ทางการเงินและการตลาด การวิเคราะห์ข้อมูลในชีวสารสนเทศ เป็นต้น การวิเคราะห์เพื่อหารูปแบบที่เหมาะสมจะใช้โปรแกรม Clementine ซึ่งเป็นโปรแกรมสำเร็จรูปที่สามารถวิเคราะห์ โดยแบ่งเป็นกลุ่มการวิเคราะห์เป็น 5 กลุ่ม และค่าที่ใช้ในการวิเคราะห์แต่ละกลุ่มดังนี้ (1) การทำนายหรือพยากรณ์ (Prediction) ได้แก่ Neural net, C5.0, C&RT, Regresstion, Logistic regression (2) การจำแนกกลุ่มหรือจัดกลุ่ม (Classification) ได้แก่ Neural net, C5.0, C&RT (3) การแบ่งเป็นส่วน(Segmentation) ได้แก่ Kohonen, K-Mean, C5.0 (4) ความเกี่ยวข้องสัมพันธ์กัน (Association) ได้แก่ Apriori, GRI และ (5) การจัดลำดับ (Sequence ) ได้แก่ CARMA, Caprl

ในทางการศึกษา เกรดมีความสำคัญและมีความหมายต่อผู้เรียนทุกคน ผู้เรียนแต่ละคนต่างมุ่งหวังและพยายามที่จะเรียนให้ได้เกรดดีที่สุด เกรดที่ดีในแต่ละรายวิชารวมเป็นเกรดเฉลี่ยสะสม (GPA) เปรียบเสมือนใบเบิกทางที่ผู้เรียนใช้ในการศึกษาต่อสถาบันการศึกษาที่มีชื่อเสียง และมีโอกาสในการหางานที่ดีมีเกียรติยศ เกรดในแต่ละรายวิชาจึงเป็นตัวกำหนดชะตาชีวิตของผู้เรียน ครูผู้สอนแต่ละรายวิชามีส่วนในการกำหนดชะตาชีวิตผู้เรียนแต่ละคน (นงลักษณ์ วิรัชชัย, 2546)

ปัจจุบัน เกรดเฉลี่ยจึงได้ถูกนำมาเป็นเกณฑ์ในการพิจารณา เช่น การกำหนดคุณสมบัติของผู้สมัครเข้าศึกษาต่อระดับบัณฑิตศึกษาสาขาวิชาชีวสถิติ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น ปีการศึกษา 2546 คือ (1) เป็นผู้สำเร็จการศึกษาขั้นปริญญาตรีจากสถาบันที่ทบวงมหาวิทยาลัยรับรอง สาขาแพทยศาสตร์ ทันตแพทยศาสตร์ สาธารณสุขศาสตร์ เทคนิคการแพทย์ เภสัชศาสตร์ พยาบาลศาสตร์ หรือวิทยาศาสตร์สุขภาพสาขาอื่น ๆ สถิติ คณิตศาสตร์ หรือวิทยาศาสตร์สาขาอื่นๆ ที่เกี่ยวข้อง ซึ่งคณะกรรมการบัณฑิตศึกษาประจำคณะสาธารณสุขศาสตร์พิจารณาว่าเหมาะสม (2) เคยเรียนวิชาคณิตศาสตร์และสถิติ หรือคอมพิวเตอร์มาแล้วอย่างน้อย 6 หน่วยกิต (3) ได้คะแนนเฉลี่ยสะสมในระดับปริญญาตรีตลอดหลักสูตรไม่ต่ำกว่า 2.50 (4) เป็นผู้มีสุขภาพดีทั้งร่างกายและจิตใจ (บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น, 2546) นั่นคือ หากผู้สมัครมีเกรดเฉลี่ยต่ำกว่า 2.50 จะไม่มีสิทธิ์ในการสมัครในปีการศึกษานั้นเลย ซึ่งนอกจากนี้ ยังมีการนำเกรดมาพิจารณาเป็นคุณสมบัติในการสมัครเข้าทำงานเช่น การประปาส่วนภูมิภาครับสมัครบุคคลเพื่อสอบคัดเลือกขึ้นบัญชีรอการบรรจุแต่งตั้งเป็นพนักงานการประปาส่วนภูมิภาคในตำแหน่งที่ใช้วุฒิไม่ต่ำกว่าระดับปริญญาตรี โดยมีคุณสมบัติดังนี้ (1) เป็นผู้มีผลการเรียนได้เกรดเฉลี่ยไม่ต่ำกว่า 2.50 (2) อายุไม่เกิน 35 ปีบริบูรณ์ ณ วันปิดรับสมัคร (3) กรณีเพศชายต้องพ้นภาระการเกณฑ์ทหารแล้ว (4) มีความสามารถในการใช้คอมพิวเตอร์ (5) สามารถไปปฏิบัติงานประจำที่ส่วนภูมิภาคได้ (สมัครงาน ,ฉบับที่ 475 หน้า 43) จะพบได้ว่า เกรดเฉลี่ยมีส่วนในการกำหนดชะตาชีวิตของคนได้ หากได้เกรดเฉลี่ยถึงเกณฑ์ที่กำหนด นั่นหมายถึง การได้โอกาสในการแสดงความสามารถของตนเองได้มากขึ้น รวมทั้งการได้สิทธิเท่าคนอื่นและมากกว่าคนที่ไม่มีสิทธิ ถ้าเราสามารถพยากรณ์เกรดเฉลี่ยได้ล่วงหน้าจะสามารถทำให้เราได้เตรียมการและป้องกันการเกิดผลลัพธ์ที่เราไม่ต้องการได้ จากความมากมายของวิธีวิเคราะห์ของการพยากรณ์ และความก้าวหน้าทางเทคโนโลยีซึ่งถูกปรุงแต่งให้ทันสมัยอยู่เสมอเพื่อให้เหมาะกับความต้องการของมนุษย์ นั้น ปัญหาที่เราต้องคิดวิเคราะห์คือ การเลือกใช้ที่เหมาะสมและดีที่สุด ดังนั้นผู้วิจัยจึงมีความสนใจที่จะศึกษาการเปรียบเทียบผลของการพยากรณ์เกรดเฉลี่ยด้วยเทคนิคคลีเมนไทน์ ซึ่งเป็นโปรแกรมสำเร็จรูปแบบหนึ่งของ Spss ที่ได้มีการพัฒนาจนถึง version 8.5 กับการถดถอยพหุคูณ ซึ่งเป็นวิธีวิเคราะห์จะเป็นแบบดั้งเดิมที่นิยมใช้ทั่วไปในโปรแกรมสำเร็จรูป Spss pc+ ผลการวิเคราะห์ที่ได้จะเป็นประโยชน์ต่อการตัดสินใจในการเลือกใช้วิธีวิเคราะห์ที่เหมาะสมและดีที่สุดกับข้อมูล เพื่อให้ได้ผลการพยากรณ์ที่ใกล้เคียงความจริงมากที่สุด โดยข้อมูลที่นำมาวิเคราะห์เพื่อเปรียบเทียบผลการพยากรณ์ของแต่ละวิธีวิเคราะห์นั้นได้มาจากฐานข้อมูล ซึ่งได้มีการจัดเก็บไว้แล้วจัดเป็นข้อมูลแบบทุติยภูมิ ตัวอย่างข้อมูลเช่น เกรดเฉลี่ยขณะที่ศึกษาในมหาวิทยาลัยอุบลราชธานี เกรดเฉลี่ยสะสมเมื่อสำเร็จการศึกษาในระดับมัธยมศึกษาตอนปลาย เพศ อายุ คะแนนวัดความรู้

สถานศึกษาเดิม ภูมิลำเนาของนักศึกษา วุฒิการศึกษาของผู้ปกครองนักเรียน อาชีพของผู้ปกครองนักเรียน แหล่งอุดหนุนการศึกษา จำนวนพี่น้องในครอบครัว รายได้บิดามารดารวมกันต่อเดือน วิธีรับเข้าการสังกัดคณะ เป็นต้น ในการวิจัยครั้งนี้ ผู้วิจัยเลือกใช้ข้อมูลเกี่ยวกับทะเบียนประวัตินักศึกษาของงานทะเบียนประวัตินักศึกษา กองบริการการศึกษา มหาวิทยาลัยอุบลราชธานี ซึ่งจะเป็นประโยชน์ต่อการวางแผนและการบริหารต่อนักศึกษา และงานทะเบียนประวัติ เพราะงานทะเบียนเป็นหัวใจทางวิชาการของมหาวิทยาลัย การได้ทราบถึงรายงาน และข้อสรุปจะมีผลต่อการนำไปพัฒนามหาวิทยาลัยและนักศึกษาให้มีคุณภาพต่อไป

**คำถามวิจัย**

วิธีวิเคราะห์ใดระหว่างเทคนิคคลีเมนไทน์กับเทคนิคการถดถอยพหุคูณที่ให้ผลพยากรณ์เกรดเฉลี่ยมีความคลาดเคลื่อนน้อยที่สุด

**วัตถุประสงค์ของการวิจัย**

1. เพื่อประยุกต์ใช้วิธีวิเคราะห์ข้อมูลในการพยากรณ์เกรดเฉลี่ยโดยเทคนิคคลีเมนไทน์ และเทคนิคการถดถอยพหุคูณ

2. เพื่อตรวจสอบผลการพยากรณ์ด้วยเทคนิคคลีเมนไทน์ และเทคนิคการถดถอยพหุคูณ โดยใช้เกณฑ์ในการเปรียบเทียบจากการวัดความคลาดเคลื่อน 3 แบบ ได้แก่ รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error หรือ RMSE) ค่ามัธยฐานของค่าสัมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ (Median Absoluate Percentage Error หรือ MdAPE) และร้อยละที่ดีกว่า (Percent Better )

**ขอบเขตการวิจัย**

การวิจัยครั้งนี้กำหนดตัวแปรต้นและตัวแปรตามดังนี้

ตัวแปรต้น ได้แก่ เกรดเฉลี่ยสะสมเมื่อสำเร็จการศึกษาระดับมัธยมศึกษาตอนปลาย คะแนนสอบคัดเลือก เพศ ภูมิลำเนา การศึกษาของบิดา การศึกษาของมารดา อาชีพของบิดา อาชีพของมารดาวิธีรับเข้า จำนวนพี่น้องในครับครัว การสังกัดคณะ และรายได้ต่อเดือนของครอบครัว (บิดามารดารวมกัน)

ตัวแปรตาม คือ เกรดเฉลี่ยสะสมเมื่อสิ้นสุดการศึกษาชั้นปีที่ 1

**คำจำกัดความที่ใช้ในการวิจัย**

*นักศึกษา* หมายถึง ผู้ที่ได้ลงทะเบียนและได้ศึกษาในมหาวิทยาลัยอุบลราชธานี ในปีการศึกษา 2544-2546

*ภูมิหลังของนักศึกษา* หมายถึง ลักษณะที่เกี่ยวกับตัวนักศึกษาและครอบครัวนักศึกษา จำนวน 9 ตัวแปร คือ เพศ การสำเร็จมัธยมศึกษาตอนปลาย คะแนนเฉลี่ยสะสมสิ้นสุดระดับมัธยมศึกษาตอนปลาย ภูมิลำเนา การศึกษาของบิดา การศึกษาของมารดา อาชีพของบิดา อาชีพของมารดา และรายได้ต่อเดือนของครอบครัว (บิดามารดารวมกัน)

*ผลสัมฤทธิ์ทางการเรียนหรือเกรดเฉลี่ยสะสม (GPA)* หมายถึง ผลการประเมินความรู้ความสามารถในการเรียนทุกรายวิชาในรูปเกรดเฉลี่ย (Grade Point Average) เมื่อสิ้นสุดแต่ละภาค

*เกรดเฉลี่ยสะสมเมื่อสำเร็จการศึกษามัธยมปลาย* หมายถึง ผลการประเมินความรู้ความสามารถทุกรายวิชาของตั้งแต่มัธยมศึกษาชั้นปีที่ 4-6 ในรูปเกรดเฉลี่ยเมื่อสิ้นสุดการศึกษามัธยมศึกษาชั้นปีที่ 6

*เกรดเฉลี่ยสะสมเมื่อศึกษาอยู่ชั้นปีที่ 1* หมายถึง ผลการประเมินความรู้ความสามารถทุกรายวิชาที่นักศึกษาลงทะเบียนเรียนในชั้นปีที่ 1 ในรูปเกรดเฉลี่ยเมื่อสิ้นสุดการศึกษาชั้นปีที่ 1

*คะแนนรวมสอบคัดเลือก* หมายถึง คะแนนที่มาจากผลการสอบวัดความรู้ในวิชาต่างๆที่คณะนั้นๆ กำหนดไว้เพื่อใช้ในการคัดเลือก

*การสอน (Train )* หมายถึง การทำงานของนิวรอลเน็ตเวิร์กเพื่อให้เกิดการเรียนรู้

*การเรียนรู้(Learning)* หมายถึง การปรับค่าให้เกิดความคลาดเคลื่อนน้อยที่สุดโดยการทำงานของนิวรอลเน็ตเวิร์ก

*จำนวนรอบการสอน* หมายถึง จำนวนครั้งที่ทำการสอนของนิวรอลเน็ตเวิร์กด้วยข้อมูลชุดเดิม การกำหนดให้จำนวนรอบการสอนมีค่าสูงเกินไปไม่ได้ทำให้นิวรอลเน็ตเวิร์กเรียนรู้ได้มากขึ้น ขึ้นอยู่กับลักษณะของรูปแบบนิวรอลเน็ตเวิร์ก

*ค่าโมเมนตัม* ($\alpha$) หมายถึง ค่าพารามิเตอร์ซึ่งเป็นค่าคงที่ที่ใช้ในการปรับค่าน้ำหนักมีค่าอยู่ระหว่าง 0 ถึง 1 ซึ่งค่าโมเมนตัมจะแสดงถึงสัดส่วนของน้ำหนักครั้งก่อนที่นำมาใช้ในการปรับน้ำหนักครั้งปัจจุบัน

*ค่าสัมประสิทธิ์การเรียนรู้* ($\eta$) เป็นค่าที่ควบคุมความคงตัว (Stability) ของค่าน้ำหนักและความเร็วในการปรับค่าน้ำหนักให้เข้าสู่ค่าที่สมดุล โดยค่าสัมประสิทธิ์การเรียนรู้มีค่าอยู่ระหว่าง 0 ถึง 1

*รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error หรือ RMSE)* หมายถึง ค่าวัดความคลาดเคลื่อนของการพยากรณ์ที่วัดจากค่าความคลาดเคลื่อนกำลังสองระหว่าง

ค่าพยากรณ์และค่าจริง ค่าวัดความคลาดเคลื่อนนี้จะไวต่อความคลาดเคลื่อนที่มีขนาดใหญ่ โดยมีเกณฑ์ที่ใช้ในการพิจารณาความเหมาะสม คือค่าวัดความคลาดเคลื่อนที่คำนวณได้ยิ่งต่ำ แสดงว่าวิธีการพยากรณ์นั้นมีความคลาดเคลื่อนน้อย ซึ่งมีสูตรในการคำนวณดังนี้

$$RMSE_m = \left( \frac{\sum_{s=1}^{s} \left( \hat{Y}_t - Y_t \right)^2}{S} \right)^{1/2}$$

*ค่ามัธยฐานของค่าสัมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ (Median Absoluate Percentage Error หรือ MdAPE)* หมายถึง ค่าวัดความคลาดเคลื่อนของการพยากรณ์ที่วัดจากค่ามัธยฐานของความคลาดเคลื่อนของการพยากรณ์เทียบกับค่าจริง ค่าวัดความคลาเคลื่อนนี้เป็นค่าที่ไม่มีหน่วย โดยมีเกณฑ์ที่ใช้ในการพิจารณาความเหมาะสมคือ ค่าวัดความคลาดเคลื่อนที่คำนวณได้ยิ่งต่ำ แสดงว่าวิธีการพยากรณ์นั้นมีความคลาเคลื่อนน้อย ซึ่งมีสูตรในการคำนวณดังนี้

$$APE_m = \left| \frac{\hat{Y}_t - Y_t}{Y_t} \right| \times 100$$

เมื่อนำค่าสังเกต $APE_m$ มาเรียงลำดับจะได้

$MdAPE_m = (S+1)/2$ เมื่อ S เป็นจำนวนคี่

$MdAPE_m = (S/2)+1$ เมื่อ S เป็นจำนวนคู่

*ร้อยละที่ดีกว่า (Percent Better )* หมายถึง ค่าวัดความคลาดเคลื่อนของการพยากรณ์ที่วัดจากการหาผลต่างของความคลาดเคลื่อนจากการพยากรณ์กับการพยากรณ์เชิงสุ่ม โดยไม่คำนึงถึงทิศทางของความคลาดเคลื่อน โดยมีเกณฑ์ที่ใช้ในการพิจาณาความเหมาะสมคือค่าวัดความคลาดเคลื่อนที่คำนวณได้ยิ่งสูง แสดงว่าวิธีการพยากรณ์นั้นมีความคลาดเคลื่อนน้อย ซึ่งมีสูตรในการคำนวณดังนี้

$$Percent\ Better_m = \left( \frac{\sum_{s=1}^{s} J_s}{S} \right) \times 100$$

เมื่อ $J_s = 1$ ถ้า $\left| \hat{Y}_t - Y_t \right| < \left| F_{rw} - Y_t \right|$

$J_s = 0$ ในกรณีอื่น ๆ

**ประโยชน์ที่คาดว่าจะได้รับ**

ผลที่ได้จากการวิจัยครั้งนี้จะเป็นประโยชน์ทั้งในเชิงปฏิบัติและในเชิงวิชาการ ดังนี้

1. ประโยชน์ในเชิงปฏิบัติของการพยากรณ์เกรดเฉลี่ย สำหรับงานทะเบียนประวัติและประมวลผลจะเป็นประโยชน์ในด้านการวางแผนของหน่วยงาน เพื่อเป็นข้อมูลในการนำเสนอผู้บริหารเพื่อทราบสภาพและแนวโน้มทางวิชาการของนักศึกษา สำหรับนักศึกษานั้นจะทำให้สามารถได้ทราบถึงเกรดเฉลี่ยล่วงหน้า ซึ่งเมื่อทราบแล้วจะได้วางแผนและจัดการปรับปรุงตนเองให้เป็นไปในทิศทางที่ต้องการได้

2. ประโยชน์ในเชิงวิชาการของการพยากรณ์เกรดเฉลี่ยโดยเทคนิคคลีเมนไทน์ และเทคนิคการถดถอยพหุคูณจากข้อมูลทุติยภูมิ เป็นการนำวิธีวิทยาการวิจัยมาใช้กับงานวิจัยทางการศึกษาโดยการเปรียบเทียบผลการพยากรณ์ ของวิธีวิเคราะห์แบบดั้งเดิมกับสมัยใหม่ ผลการวิจัยในครั้งนี้จึงเป็นอีกทางหนึ่งของวิธีวิทยาการวิจัยที่จะเป็นประโยชน์ต่อการพยากรณ์ และเป็นประโยชน์ต่อการพัฒนาความก้าวหน้าทางด้านวิธีวิทยาการวิจัยในส่วนของวิธีวิเคราะห์ต่อไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย