

**การเปรียบเทียบเทคนิค Naïve Bayes และ Binary Logistic Regression  
เพื่อการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการตัดสินใจศึกษาต่อของนักเรียน U.N.N. Academy**

**Comparison of Naïve Bayes and Binary Logistic Regression Techniques to Analyze  
the Factors Influencing Students' Decision to Enroll in U.N.N. Academy Language-School**

วรรณวิษา คุณย์มงคล\* และ สมชาย เล็กเจริญ

สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยรังสิต ถนนพหลโยธิน ตำบลหลักหก อำเภอเมือง  
จังหวัดปทุมธานี 12000 E-mail: signrsu\_wt@hotmail.com

**บทคัดย่อ**

การวิจัยในครั้งนี้ มีวัตถุประสงค์เพื่อการศึกษาปัจจัยที่มีอิทธิพลต่อการตัดสินใจศึกษาต่อในสถาบันสอนภาษา U.N.N. Academy โดยใช้เทคนิค Naïve Bayes เปรียบเทียบกับเทคนิค Binary Logistic Regression ซึ่งใช้แบบสอบถามเป็นเครื่องมือในการรวบรวมข้อมูลของนักเรียน จำนวน 150 คน สถิติที่ใช้ในการวิเคราะห์ข้อมูล คือ ค่าร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน เทคนิคการวิเคราะห์ความถดถอยโลจิสติกแบบไบนารี(Binary Logistic Regression analysis) เพื่อหาความสัมพันธ์ของข้อมูลและรูปแบบที่เหมาะสม โดยมีตัวแปรตาม 1 ตัว คือ การตัดสินใจศึกษาต่อของนักเรียน ที่มีค่าเพียง 2 ค่า (Dichotomous variable) คือ 0 และไม่เรียน/เรียน) และมีตัวแปรอิสระทั้งหมด 10 ตัวแปร ได้แก่ เพศ อายุ เกรดเฉลี่ย รายได้ครอบครัวต่อเดือน อาชีพบิดา อาชีพมารดา การศึกษาของบิดา การศึกษาของมารดา ประสบการณ์การศึกษาต่อ และหลักสูตรที่เคยเรียน ซึ่งผลการวิจัยพบว่า ปัจจัยที่มีอิทธิพลต่อการตัดสินใจศึกษาต่อ ได้แก่ เกรดเฉลี่ย รายได้ครอบครัวต่อเดือน ประสบการณ์การศึกษาต่อ และหลักสูตรที่เคยเรียน ด้วยค่าความถูกต้องที่ 94.7% และงานวิจัยนี้ยังนำการวิเคราะห์ข้อมูลโดยใช้เทคนิค Naïve Bayes มาเปรียบเทียบกับ Binary Logistic Regression ซึ่งได้ผลลัพธ์ที่สอดคล้องกัน

**คำสำคัญ:** ทฤษฎีเบย์ การวิเคราะห์ความถดถอยโลจิสติกแบบไบนารี

**Abstract**

The purpose of this study was to compare the Naïve Bayes and Binary Logistic Regression techniques to determine factors that influence students' decision to enroll at the U.N.N. Academy language school. The research was conducted using questionnaires to collect data from 150 students. The statistics used for the data analysis were calculated in percentage (%), ratio, standard deviation and Binary Logistic Regression. The Binary Logistic Regression model was used to analyze key factors that influence students' decision to enroll in accordance to data

relation. From 150 students, the student's decision was used as a dependent variable and 10 independent variables which were sex, age, GPA, household income, father's occupation, mother's occupation, father's education, mother's education, enrollment history and courses studied. For the design of the Binary Logistic Regression model, two variables in dichotomous variable technique were changed. The results showed that the four factors that influenced the students' decision to enroll were GPA, household income, enrollment history and courses taken. The prediction accuracy for the students was 94.7%. Furthermore, it was found that the analysis between the Naïve Bayes and Logistic Regression techniques were consistent.

**Keywords:** Naive Bayes, Binary Regression Logistic

## 1. บทนำ

ธุรกิจสถาบันสอนภาษามีการเปิดตัวขึ้นมากมาย บวกกับการแข่งขันทางการตลาดที่สูงขึ้น และจากการศึกษาระบบของสถาบันสอนภาษาอื่นอื่น อะเคเดมี พบว่าเป็นสถาบันกวดวิชาเพื่อสอบเข้าในระดับประถมศึกษา มัธยมศึกษาตอนต้น มัธยมศึกษาตอนปลาย และมหาวิทยาลัย อีกทั้งยังเปิดสอนภาษาอังกฤษ ภาษาจีน ภาษาญี่ปุ่น รวมถึงดนตรีอีกด้วย โดยระบบบริหารจัดการสถาบันสอนภาษาแห่งนี้ ปัจจุบันยังคงใช้เอกสารเป็นส่วนใหญ่ และด้วยการดำเนินงานอย่างต่อเนื่องจนถึงปัจจุบัน ทำให้มีจำนวนเอกสารเพิ่มมากขึ้น ส่งผลให้เกิดการสูญหายของเอกสาร ข้อมูลในเอกสารไม่เป็นปัจจุบัน การที่สถาบันสอนภาษาฯ ขาดระบบข้อมูลที่ดีทำให้การสืบค้นข้อมูลต่างๆ ทำได้ลำบากอย่างมาก อีกทั้งทางสถาบันฯ ขาดระบบในการบริหารจัดการและช่วยสนับสนุนการตัดสินใจ จึงส่งผลให้ไม่สามารถวางแผนการเปิดหลักสูตรการเรียนการสอนและสร้างตารางการสอนของครูผู้สอนได้ดีพอ ทำให้เกิดเหตุการณ์ครูผู้สอนไม่เพียงพอหรือมากเกินไปจนไปอยู่เป็นไปประจำ

จากปัญหาที่กล่าวมาข้างต้น ผู้วิจัยจึงมีความคิดที่จะทำการพัฒนาระบบเพื่อช่วยในการวิเคราะห์และสนับสนุนการตัดสินใจเข้าศึกษาต่อ

สถาบันสอนภาษา U.N.N. Academy โดยนำเทคนิค Naïve Bayes มาทำการเปรียบเทียบกับเทคนิค Binary Logistic Regression เพื่อหาปัจจัยที่มีอิทธิพลต่อการตัดสินใจเข้าเรียนของนักเรียนและเพื่อทราบคุณลักษณะเบื้องต้นของนักเรียนที่มีแนวโน้มจะศึกษาต่อในหลักสูตรต่อไป นอกจากนั้นยังพัฒนาโปรแกรมขึ้นเพื่อช่วยในการจัดการข้อมูลและทดแทนระบบการทำงานเดิมของสถาบันฯ

จิราพร ยิ่งกว่าชาติ (2549) ได้กล่าวถึง Naïve Bayes ว่าเป็นเทคนิคที่ใช้ทฤษฎี Bayes Theorem ในการคำนวณความน่าจะเป็น ซึ่งถูกใช้ในการทำนายผลเมื่อทำการวิเคราะห์กรณีใหม่ การทำนายผลทำได้โดยการรวมผลของตัวแปรอิสระ (Independent variable) ที่มีต่อตัวแปรตาม (Dependent variable) ซึ่งสามารถทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระแต่ละตัวกับตัวแปรตามเพื่อใช้ในการสร้างเงื่อนไขความน่าจะเป็นสำหรับแต่ละความสัมพันธ์

อุไรวรรณ อมรมณี (2546) ได้นำเสนอ Binary Logistic Regression ในตัวแบบที่เรียกว่า Binary Logistic Regression Model จะอยู่ในรูปแบบของสมการเส้นถดถอย มีตัวแปรตามที่มีค่าเพียง 2 ค่า คือ 0 และ 1 ส่วนตัวแปรอิสระอาจมีเพียงตัวเดียวหรือหลายตัวก็ได้ (Dichotomous variable) หรือตัวแปรอิสระอาจเป็น

ถูกผสมที่แต่ละตัวลักษณะแตกต่างกันออกไป และจากการที่ Logistic Regression Analysis เป็นการทำนายค่าความน่าจะเป็นของการเกิดหรือไม่เกิดเหตุการณ์ที่สนใจ ซึ่งมีข้อมูลตัวแปรตามเป็น Dichotomous จึงมีความเกี่ยวข้องกับทฤษฎีความน่าจะเป็นแบบทวินาม ดังนั้นเทคนิคนี้อาจถูกเรียกว่า Binomial Logistic Regression

## 2. วัตถุประสงค์

การวิจัยในครั้งนี้ มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อการตัดสินใจศึกษาต่อในสถาบันฯ โดยใช้เทคนิค Naïve Bayes และ Binary Logistic Regression มาเปรียบเทียบกัน อีกทั้งยังพัฒนาโปรแกรมเพื่อช่วยบริหารจัดการข้อมูลและสนับสนุนการตัดสินใจของผู้บริหาร เพื่อทดแทนและปรับปรุงระบบการทำงานเดิม

## 3. อุปกรณ์และวิธีการ

### 3.1 ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการศึกษา คือ นักเรียนที่เข้ามาติดต่อสอบถามหรือสมัครเพื่อลงทะเบียนเรียนในสถาบันสอนภาษา U.N.N. Academy โดยผู้วิจัยใช้การสุ่มตัวอย่างแบบเจาะจง (Purposive sampling) ซึ่งเป็นการเลือกกลุ่มตัวอย่างโดยพิจารณาจากการตัดสินใจของผู้วิจัยเอง ลักษณะของกลุ่มที่เลือก คือ เป็นไปตามวัตถุประสงค์ของการวิจัย โดยวิธีการกำหนดขนาดของกลุ่มตัวอย่าง ผู้วิจัยได้กำหนดกลุ่มตัวอย่างที่ต้องการจะศึกษาวิจัยจำนวนทั้งสิ้น 150 คน

### 3.2 เครื่องมือที่ใช้ในการวิจัย

แบบสอบถามที่ผู้วิจัยสร้างขึ้น ลักษณะของคำถามเป็นแบบปลายปิด โดยแบ่งออกเป็น 3 ส่วน ได้แก่ ส่วนที่ 1 ข้อมูลส่วนตัวของผู้ตอบแบบสอบถาม ส่วนที่ 2 ข้อมูลการเข้าเรียนในสถาบันสอนภาษา และ

ส่วนที่ 3 ข้อมูลเกี่ยวกับปัจจัยที่มีผลต่อการตัดสินใจเข้าเรียนในสถาบันสอนภาษาของผู้เรียน แบ่งเป็น 6 ด้าน คือ ด้านสภาพแวดล้อมของสถาบัน ด้านคุณภาพการจัดการเรียนการสอน เหตุผลที่เข้าเรียนกวดวิชา/เสริมภาษา ด้านแหล่งการรับรู้ข้อมูลข่าวสาร ด้านผู้สอน และความคาดหวังของผู้เรียน

### 3.3 วิธีการเก็บรวบรวมข้อมูล

การวิจัยครั้งนี้ได้กำหนดวิธีการทดสอบเครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลดังนี้

3.3.1 นำแบบสอบถามเสนออาจารย์ที่ปรึกษาเพื่อตรวจทาน และนำมาปรับปรุงแก้ไขให้สมบูรณ์

3.3.2 นำแบบสอบถามที่ได้ปรับปรุงแก้ไขแล้วเสนอผู้ทรงคุณวุฒิหรือผู้เชี่ยวชาญ เพื่อพิจารณาตรวจสอบความเที่ยงตรงและความเหมาะสมของเนื้อหา โดยมีผู้เชี่ยวชาญจากคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยรังสิต จำนวน 3 ท่าน เป็นผู้ตรวจสอบ

3.3.3 นำผลคะแนนความเห็นจากผู้เชี่ยวชาญมาวิเคราะห์เพื่อหาดัชนีความสอดคล้อง (Index of Item Objective Congruence: IOC) (เรื่องอุไร ศรีนิลทา, 2535) และปรับปรุงแก้ไขจนทุกข้อคำถามมีค่า IOC ไม่ต่ำกว่า 0.5 ซึ่งถือว่าเป็นแบบสอบถามที่มีความตรงเชิงเนื้อหา

3.3.4 นำแบบสอบถามที่ผ่านการทดสอบแล้ว ไปใช้ในการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างต่อไป

### 3.4 การวิเคราะห์ข้อมูล

3.4.1 สถิติเชิงพรรณนา ได้แก่ ค่าความถี่และค่าร้อยละ ค่าเฉลี่ยตัวอย่าง ค่าเบี่ยงเบนมาตรฐาน S.D.

3.4.2 สถิติเชิงวิเคราะห์ ได้แก่ การวิเคราะห์ข้อมูลด้วย Binary Logistic Regression และการวิเคราะห์ข้อมูลด้วย Naïve Bayes

#### 4. ผลการวิจัย

##### 4.1 การวิเคราะห์ความถดถอยโลจิสติกแบบไบนารี (Binary Logistic Regression Analysis)

จะทำการตรวจสอบหาความสัมพันธ์ระหว่างตัวแปรอิสระแต่ละตัว กับตัวแปรตาม ด้วยค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย (Simple Correlation Analysis) หรือ  $r$

ตารางที่ 1 แสดงความสัมพันธ์ระหว่างตัวแปรตาม และตัวแปรอิสระ

| Model Summary |                     |                      |                     |
|---------------|---------------------|----------------------|---------------------|
| Step          | -2 Log likelihood   | Cox & Snell R Square | Nagelkerke R Square |
| 1             | 36.165 <sup>a</sup> | .549                 | .850                |

จากตารางที่ 1 เป็นการวิเคราะห์เพื่อหาระดับความสัมพันธ์ของการตัดสินใจศึกษาต่อ กับตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมากและเป็นไปในทิศทางเดียวกันทั้งหมดซึ่งพบว่ามีค่าสัมประสิทธิ์สหสัมพันธ์ ( $r$ ) เท่ากับ 0.850 ซึ่งหมายถึงมีความสัมพันธ์กันมากและเป็นไปในทิศทางเดียวกัน

ตารางที่ 2 แสดงตาราง Classification Table

| Classification Table |                             |                             |     |      |      |
|----------------------|-----------------------------|-----------------------------|-----|------|------|
|                      |                             | Predicted                   |     |      |      |
|                      |                             | การตัดสินใจเข้าเรียนกวดวิชา | No  | Yes  |      |
| Step                 | การตัดสินใจเข้าเรียนกวดวิชา | No                          | 31  | 1    | 96.9 |
|                      | Yes                         | 7                           | 111 | 94.1 |      |
| 1                    | Overall Percentage          |                             |     |      | 94.7 |

จากตารางที่ 2 แสดงค่าร้อยละของการพยากรณ์ หรือค่าความถูกต้องของการทำนาย และการคำนวณพบว่า ตัวแบบสามารถทำนายนักเรียนที่ตัดสินใจไม่ศึกษาต่อ ได้ถูกต้องถึง 96.9% และทำนายนักเรียนที่ตัดสินใจศึกษาต่อ ได้ถูกต้อง 94.1% ในขณะที่

ภาพรวมจะเห็นได้ว่าสามารถทำนายได้อย่างถูกต้องถึง 94.7%

ตารางที่ 3 แสดงตาราง Variables in the Equation

| Variables in the Equation |         |           |       |       |
|---------------------------|---------|-----------|-------|-------|
|                           | B       | S.E.      | Wald  | Sig.  |
| GPA_1.00to2.00(1)         | -.429   | 15718.634 | .000  | 1.000 |
| GPA_2.01to2.50(1)         | -16.941 | 6756.588  | .000  | .998  |
| GPA_2.51to3.00(1)         | 2.068   | 8773.145  | .000  | 1.000 |
| GPA_3.01to3.50(1)         | 3.321   | 1.403     | 5.602 | .018  |
| IN_Lessthan10000(1)       | 41.540  | 19469.567 | .000  | .998  |
| IN_10001to20000(1)        | 21.645  | 15566.496 | .000  | .999  |
| IN_20001to30000(1)        | 19.463  | 15566.496 | .000  | .999  |
| IN_30001to40000(1)        | 19.079  | 12858.769 | .000  | .999  |
| IN_40001to50000(1)        | .985    | 14645.262 | .000  | 1.000 |
| STUDIED(1)                | -19.246 | 4036.512  | .000  | .996  |
| COURSE                    |         |           | .000  | .997  |
| COURSE(1)                 | -14.916 | 4036.512  | .000  | .997  |
| Constant                  | -50.494 | 64322.956 | .000  | .999  |

จากตารางที่ 3 แสดงให้เห็นค่า  $p$  (Sig.) ของตัวแปรแต่ละตัว โดยกำหนดให้ตัวแปรที่มีนัยสำคัญ  $p < 0.05$  ดังนั้นสามารถสรุปได้ว่า มีตัวแปรอิสระ 1 ตัว คือ เกรดเฉลี่ย 3.01-3.50 (GPA\_3.01 to 3.50) เป็นตัวแปรที่ส่งผลให้นักเรียนตัดสินใจศึกษาต่ออย่างมีนัยสำคัญ และจากการวิเคราะห์ความถดถอยโลจิสติกแบบไบนารี จะได้รูปแบบสมการการพยากรณ์ความเป็นได้ในการตัดสินใจศึกษาต่อตามรูปที่ 1 ดังนี้

$$Z = -50.494 - 0.429(\text{GPA}_{1.00\text{to}2.00}) - 16.941(\text{GPA}_{2.01\text{to}2.50}) + 2.068(\text{GPA}_{2.51\text{to}3.00}) + 3.321(\text{GPA}_{3.01\text{to}3.50}) + 41.540(\text{IN}_{\text{Lessthan}10000}) + 21.645(\text{IN}_{10001\text{to}20000}) + 19.463(\text{IN}_{20001\text{to}30000}) + 19.079(\text{IN}_{30001\text{to}40000}) + 0.985(\text{IN}_{40001\text{to}50000}) - 19.246(\text{Studied}) - 14.916(\text{Course})$$

รูปที่ 1 สมการการพยากรณ์ความเป็นได้ในการตัดสินใจศึกษาต่อ

##### 4.2 การวิเคราะห์หาปัจจัยที่มีอิทธิพลในการตัดสินใจ (Naïve Bayes)

ตารางที่ 4 แสดงผลการวิเคราะห์หาปัจจัยที่มีอิทธิพลในการตัดสินใจ โดยใช้สถิติ Naïve Bayes

**Key Influencers Report for 'DECISION'**

Key Influencers and their impact over the values of 'DECISION'

Filter by 'Column' or 'Favors' to see how various columns influence 'DECISION'

| Column           | Value | Favors | Relative Impact |
|------------------|-------|--------|-----------------|
| STUDIED          | 1     | 1      |                 |
| COURSE           | 2     | 1      |                 |
| GPA_2_01to2_50   | 1     | 1      |                 |
| GPA_2_51to3_00   | 1     | 1      |                 |
| IN_40001to50000  | 1     | 1      |                 |
| IN_30001to40000  | 1     | 1      |                 |
| MOCC_EMP         | 1     | 1      |                 |
| COURSE           | 9     | 0      |                 |
| GPA_3_01to3_50   | 1     | 0      |                 |
| IN_Lessthan10000 | 1     | 0      |                 |
| IN_10001to20000  | 1     | 0      |                 |
| GPA_3_51to4_00   | 1     | 0      |                 |
| FOCC_SHK         | 1     | 0      |                 |

จากตารางที่ 4 จะพบว่าตัวแปรที่มี Relative impact มากที่สุด ได้แก่ ประสบการณ์การศึกษาต่อ (Studied) หลักสูตรที่เคยศึกษาต่อ (Course) เกรดเฉลี่ย (GPA) และรายได้ครอบครัว (IN) ซึ่งถือได้ว่าปัจจัยเหล่านี้มีอิทธิพลในการตัดสินใจศึกษาต่อ โดยผลการวิเคราะห์ดังกล่าวสามารถสรุปคุณลักษณะของนักเรียนที่จะตัดสินใจศึกษาต่อมากที่สุด ได้ดังนี้ 1) นักเรียนมีประสบการณ์การเรียนกวดวิชามาก่อน 2) นักเรียนเคยเรียนกวดวิชาในหลักสูตร Group course 3) นักเรียนมีเกรดเฉลี่ย 2.01-3.00 และ 4) ครอบครัวของนักเรียนมีรายได้ต่อเดือน 30,001-50,000 บาท

## 5. การอภิปรายผล

จากผลการวิเคราะห์ความสัมพันธ์และวิเคราะห์หาปัจจัยที่มีอิทธิพลในการตัดสินใจเข้าเรียนพบว่า ปัจจัยที่มีความสัมพันธ์กับการตัดสินใจอย่างมีนัยสำคัญทางสถิติคือ เกรดเฉลี่ย หลักสูตรที่เคยเข้าเรียน ประสบการณ์การเข้าเรียน รายได้ครอบครัว และอาชีพมารดา โดยสามารถอภิปรายผลได้ดังนี้

ปัจจัยด้าน เกรดเฉลี่ย ของ นักเรียน มีความสัมพันธ์กับการตัดสินใจเข้าเรียนกวดวิชามากที่สุด ด้วยค่าสัมประสิทธิ์สหสัมพันธ์ (r) เท่ากับ 0.640 ทั้งนี้จากการศึกษาพบว่า นักเรียนส่วนใหญ่ที่ตัดสินใจเข้า

เรียนมากที่สุด มีเกรดเฉลี่ยอยู่ในช่วง 2.01-2.50 รองลงมาคือ 2.51-3.00 ซึ่งถือว่ามีผลการเรียนอยู่ในระดับแย่งปานกลาง จึงเห็นได้ว่านักเรียนยังมีผลการเรียนแย่งเท่าไรก็มีแนวโน้มตัดสินใจเข้าเรียนมากขึ้นเท่านั้น ซึ่งสอดคล้องกับงานวิจัยของก้องเกียรติ บุญเสริม (2552) ที่พบว่า คะแนนเฉลี่ยสะสมของนักเรียนเป็นปัจจัยที่มีความสัมพันธ์กับการเรียนกวดวิชา โดยแสดงให้เห็นว่านักเรียนที่มีผลการเรียนแย่งจะแสวงหาความรู้จากแหล่งอื่นนอกเหนือจากการเรียนในโรงเรียนเพิ่มเติม รวมถึงนักเรียนที่มีผลการเรียนดีแล้วแต่มีความขยันใฝ่รู้ด้วยเช่นกัน

ปัจจัยที่มีความสัมพันธ์กันรองลงมาคือ ประสบการณ์การเข้าเรียนกวดวิชา และหลักสูตรที่เคยเรียน ซึ่งมีค่าสัมประสิทธิ์สหสัมพันธ์ (r) คือ 0.616 และ 0.635 ตามลำดับ ซึ่งเป็นไปได้อย่างยิ่งที่เมื่อนักเรียนตัดสินใจเข้าเรียนแล้วและมีผลการเรียนที่ดีขึ้น ทำให้ตัดสินใจเข้าเรียนต่อ ซึ่งสอดคล้องกับผลการวิเคราะห์หาปัจจัยที่มีอิทธิพลต่อการตัดสินใจเข้าเรียน ที่พบว่า ประสบการณ์การเข้าเรียนกวดวิชาเป็นปัจจัยที่มีอิทธิพลมากที่สุด รองลงมาคือหลักสูตรที่เคยเข้าเรียน ทั้งนี้เนื่องจากนักเรียนที่ตัดสินใจเข้าเรียนโดยส่วนใหญ่ จะเคยเรียนในหลักสูตร Group course มาก่อน ซึ่งเป็นการเรียนกวดวิชาแบบกลุ่ม นอกจากนักเรียนจะได้เพิ่มทักษะทางการเรียนแล้วนักเรียนยังได้พบกับสังคมใหม่ๆ ในการแลกเปลี่ยนความรู้ เสนอความคิดเห็น จึงเป็นไปได้อย่างยิ่งที่นักเรียนจะตัดสินใจเรียนในหลักสูตรต่อ ๆ ไป เพื่อเป็นการเพิ่มพัฒนาการและเสริมสร้างทักษะอย่างต่อเนื่อง ซึ่งสอดคล้องกับผลงานวิจัยของสาลินี คันธวะ (2550) ที่พบว่า ประสบการณ์เข้าเรียนเป็นปัจจัยที่มีอิทธิพลในการเรียนกวดวิชาของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ในเขตอำเภอเมือง จังหวัดลำปาง โดยผู้ที่เคยเรียนกวดวิชามาก่อนถึงร้อยละ 89 เลือกเรียนกวดวิชาต่อปัจจัยด้าน

รายได้ครอบครัว เป็นปัจจัยที่สำคัญอีกอย่างหนึ่งที่มีความสัมพันธ์กับตัวแปรการตัดสินใจเข้าเรียน โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ ( $r$ ) เท่ากับ 0.595 และเมื่อพิจารณาแล้วจะเห็นได้ว่าปัจจัยด้านรายได้จะมีความเกี่ยวข้องกับปัจจัยด้านอาชีพของผู้ปกครองด้วย โดยการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการตัดสินใจ แสดงให้เห็นว่า นักเรียนที่มีมารดาเป็นพนักงาน/ลูกจ้างเอกชน มีแนวโน้มตัดสินใจเข้าเรียน หมายความว่าปัจจัยด้านอาชีพของมารดามีอิทธิพลต่อการตัดสินใจเข้าเรียนด้วยเช่นกัน ซึ่งสอดคล้องกับผลงานวิจัยของก้องเกียรติ บุญเสริม (2552) ที่ได้กล่าวไว้ว่า ปัจจัยด้านรายได้รวมของครอบครัวเป็นปัจจัยที่แสดงถึงฐานะทางเศรษฐกิจและสังคมของผู้ปกครองหรือของครอบครัว หมายความว่านักเรียนที่มีความพร้อมและครอบครัวสามารถสนับสนุนในด้านค่าใช้จ่ายได้ จะมีแนวโน้มสูงในการตัดสินใจเรียนกวดวิชาด้วยการสนับสนุนดังกล่าว นอกจากนี้ผลงานวิจัยของศรารักษ์ ศิริพันธุ์ (2544) ยังพบว่า รายได้ครอบครัวเป็นปัจจัยที่มีอิทธิพลต่อการเลือกศึกษาต่อในระดับปริญญาโท ด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศของบัณฑิตด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ของสถาบันอุดมศึกษาของรัฐ ในสังกัดทบวงมหาวิทยาลัยในเขตกรุงเทพมหานครอีกด้วย

ในส่วนของปัจจัยที่มีผลต่อระดับการตัดสินใจเข้าเรียนสถาบันกวดวิชา พบว่า ด้านความคาดหวังของผู้เรียนมีระดับการตัดสินใจสูงสุด โดยมีค่าเฉลี่ยเท่ากับ 3.92 ซึ่งอยู่ในระดับมาก และปัจจัยที่มีอิทธิพลมากที่สุดในด้านความคาดหวังของผู้เรียนคือ สามารถเข้าใจในบทเรียนและมีผลการเรียนที่ดีขึ้น โดยมีค่าเฉลี่ยเท่ากับ 3.92 ซึ่งสอดคล้องกับผลงานวิจัยของเจริญพร แก้วละเอียด (2544) ที่พบว่า ปัจจัยด้านความคาดหวังของตัวผู้เรียนที่ต้องการเข้าศึกษาต่อในมหาวิทยาลัย เป็นปัจจัยที่มีผลต่อการเรียนกวดวิชาของนักเรียนชั้นมัธยมศึกษาตอนปลาย

## 6. สรุปผล

การเปรียบเทียบเทคนิค Naïve Bayes และ Binary Logistic Regression เพื่อการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการตัดสินใจศึกษาต่อของนักเรียน U.N.N. Academy โดยผู้วิจัยใช้แบบสอบถามเป็นเครื่องมือในการรวบรวมข้อมูลจำนวนทั้งสิ้น 150 กลุ่มตัวอย่าง และนำแบบสอบถามดังกล่าวมาประมวลผลตามระเบียบทางสถิติ โดยใช้โปรแกรมสำเร็จรูปทำการวิเคราะห์ข้อมูล ซึ่งมีขั้นตอนดังต่อไปนี้

1. ตรวจสอบความถูกต้อง สมบูรณ์ ครบถ้วนของแบบสอบถาม
2. บันทึกข้อมูลลงใน Worksheet และหาค่าความเชื่อมั่น
3. วิเคราะห์ข้อมูลส่วนบุคคลของกลุ่มตัวอย่าง โดยใช้สถิติพรรณนา ได้แก่ จำนวน ร้อยละ อัตราส่วน ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน รวมไปถึงค่าสัมประสิทธิ์ในการหาความสัมพันธ์ของตัวแปร และ Binary Logistic Regression เพื่อหาปัจจัยที่มีอิทธิพลต่อการตัดสินใจศึกษาต่อในสถาบันสอนภาษา U.N.N. Academy

นอกจากนั้น ผู้วิจัยยังได้นำข้อมูลจากแบบสอบถามไปวิเคราะห์ข้อมูล โดยใช้เทคนิค Naïve Bayes ด้วยโปรแกรม Microsoft SQL 2008 Data Mining Add-ins เพื่อหาปัจจัยที่มีอิทธิพลต่อการตัดสินใจศึกษาต่อในสถาบันฯ และทำการเปรียบเทียบผลลัพธ์ที่ได้จากการวิเคราะห์กับวิธี Binary Logistic Regression อีกด้วย

การประยุกต์ใช้เทคนิค Naïve Bayes ทำให้ได้ผลการทำนายคุณลักษณะของนักเรียนที่จะตัดสินใจศึกษาต่อในสถาบันสอนภาษา U.N.N. Academy และปัจจัยที่มีอิทธิพลต่อการตัดสินใจศึกษาต่อนั้นได้แก่เกรดเฉลี่ย หลักสูตรที่เคยศึกษาต่อ ประสบการณ์ การศึกษาต่อและรายได้ครอบครัว ซึ่งรูปแบบในการ

ทำนายนี้เมื่อนำไปเปรียบเทียบกับผลการวิเคราะห์ความถดถอยโลจิสติกแบบไบนารี (Binary Logistic Regression Analysis) พบว่า ได้ผลการวิเคราะห์ที่มีความสอดคล้องกัน

## 7. กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณผู้บริหารและเจ้าหน้าที่สถาบันสอนภาษา U.N.N. Academy ที่ได้ให้ความอนุเคราะห์และความร่วมมือในการรวบรวมเอกสารข้อมูลต่าง ๆ ประกอบการวิจัย

## 8. เอกสารอ้างอิง

กัลวัฒน์ มัญชะสิงห์. (2554). การวิเคราะห์สหสัมพันธ์ (ออนไลน์). สืบค้นจาก: <http://kalawat.esu.ac.th> (15 พฤศจิกายน 2555).

ก้องเกียรติ บุญเสริม. (2552). ปัจจัยที่ส่งผลต่อการตัดสินใจเรียน กวดวิชา ของนักเรียนมัธยมศึกษาตอนต้น: การวิเคราะห์จำแนกกลุ่มพหุกับการวิเคราะห์ชีพพอร์ตเวกเตอร์แมชชีน. คุรุศาสตรมหาบัณฑิต. จุฬาลงกรณ์มหาวิทยาลัย

เจริญพร แก้วละเอียด. (2544). ค่านิยมที่มีผลต่อการเรียนกวดวิชาของนักเรียนชั้นมัธยมศึกษาตอนปลาย. ศึกษาศาสตรมหาบัณฑิต (การวัดผลและวิจัยการศึกษา. มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี.

จิราพร ยิ่งกว่าชาติ. (2549). การประยุกต์ใช้การเรียนรู้แบบเบย์ กับการสร้างแบบจำลองสำหรับทำนายผลสำเร็จการศึกษาของนักศึกษา. วิทยาศาสตร์มหาบัณฑิต, สาขาวิชาเทคโนโลยีสารสนเทศบัณฑิตวิทยาลัย มหาวิทยาลัยศรีปทุม.

เรืองอุไร ศรีนิลทา. (2535). ระเบียบวิธีวิจัย. กรุงเทพฯ : สำนักส่งเสริมและฝึกอบรม มหาวิทยาลัยเกษตรศาสตร์.

ศรารวณ ศิริพันธุ์. (2544). ปัจจัยที่มีอิทธิพลต่อการเลือกศึกษาต่อในระดับปริญญาโทด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ของบัณฑิตด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ สถาบันอุดมศึกษาของรัฐในสังกัดทบวงมหาวิทยาลัย ในเขตกรุงเทพฯ. คุรุศาสตร์อุตสาหกรรมมหาบัณฑิต สาขาวิชาคอมพิวเตอร์และเทคโนโลยีสารสนเทศ. มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี.

สวัสดิ์ชัย ศรีพนมธนากร. (2548). การวิเคราะห์การถดถอยโลจิสติก (ออนไลน์). สืบค้นจาก: <http://www.thairenu.com/logistic.html>. (15 สิงหาคม 2549).

สาลีณี คันธวะ. (2550). ปัจจัยที่มีอิทธิพลในการเรียนกวดวิชาของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ในเขตอำเภอเมือง จังหวัดลำปาง. คณะเศรษฐศาสตร์ มหาวิทยาลัยเชียงใหม่.

อุไรวรรณ อมรมิตรี. (2546). การวิเคราะห์ข้อมูลโดยใช้ Logistic Regression: ทางเลือกของการวิเคราะห์ความเสี่ยง. วารสารวิชาการมหาวิทยาลัยหอการค้าไทย: 21-35.