

# การคาดการณ์ภาษีมูลค่าเพิ่มด้วยเทคนิคของเหมืองข้อมูล

## Forecasting VAT with Data Mining Techniques

วิระ จิริกิจอนุสรณ์\* และ ชุติรัตน์ จรัสกุลชัย\*

### บทคัดย่อ

ภาษีมูลค่าเพิ่มเป็นหนึ่งในแหล่งรายได้ของภาครัฐที่นำมาจัดทำงบประมาณในการพัฒนาประเทศ ประสิทธิภาพของการพยากรณ์รายได้จากภาษีมูลค่าเพิ่มจะช่วยให้ภาครัฐสามารถวางแผนงบประมาณอย่างมีประสิทธิภาพด้วย เทคนิคของเหมืองข้อมูลได้ถูกนำมาใช้ในงานประยุกต์หลายแขนง รวมทั้งการศึกษาวิจัยคาดการณ์ภาษีมูลค่าเพิ่ม อย่างไรก็ตาม ปัจจัยหรือตัวแปรที่ใช้ในการศึกษายังไม่ครอบคลุมทั้งหมด งานวิจัยฉบับนี้ศึกษาเพิ่มเติมปัจจัยอื่นๆ ที่อาจส่งผลต่อการคาดการณ์ และงานวิจัยนี้สร้างแบบจำลองด้วยเงื่อนไขที่หลากหลาย จากการศึกษาวิจัยพบว่าแบบจำลอง Multilayer perceptron ที่ใช้ตัวแปร ค่าดัชนีราคาผู้บริโภค ดัชนีสินค้านำเข้า ดัชนีการอุปโภคบริโภคภาคเอกชน ดัชนีการลงทุนภาคเอกชน รายจ่ายภาครัฐ จะให้ค่าการคาดการณ์ผลการจัดเก็บภาษีมูลค่าเพิ่ม ที่มีค่าผิดพลาดในการคาดการณ์เฉลี่ยน้อยที่สุด และมีการคาดการณ์ถูกต้องเฉลี่ยคิดเป็น 98 %

**คำสำคัญ:** Linear regression, Data Mining, Multilayer Perceptron

### Abstract

Value Added Tax revenue is one of resources for government budgeting. The effectiveness of forecasting VAT revenue will reflect the government budgeting. Data Mining techniques has been applied in several application areas, Included research forecasting VAT. However, factors, or variables used in previous study do not consider other variables. This research focuses on other factors that may affect the

\* คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

forecasting. Various models have been built on many conditions. The study found that the Multilayer perceptron model which used the variable value to the consumer price index, the import products index, the private consumption index, the private investment index and the public expenditure would get value the VAT collect estimations, that were the mistake value from the estimations at the minimum average and 98% for the correct average.

**Keyword:** Linear regression, Data Mining, Multilayer, Perceptron

### 1. บทนำ

ภาษีมูลค่าเพิ่มเป็นภาษีชนิดหนึ่งที่กรมสรรพากรมีหน้าที่จัดเก็บทุกเดือน และเป็นภาษีที่มีมูลค่าการจัดเก็บที่สูง โดยในปี พ.ศ. 2552 กรมสรรพากร มียอดผลการจัดเก็บภาษีทั้งหมดประมาณ 1,180,000 ล้านบาท ซึ่งภาษีมูลค่าเพิ่มมี ผลการจัดเก็บประมาณ 23% ของยอดรวมการจัดเก็บภาษีทั้งหมด จึงมีการนำค่าผลการจัดเก็บภาษีมูลค่าเพิ่มไปใช้ในการประมาณการรายได้ของรัฐเพื่อกำหนดโครงการต่างๆ ดังนั้นจึงเห็นได้ว่าการคาดการณ์ผลการจัดเก็บภาษีมูลค่าเพิ่มจึงมีประโยชน์ที่ใช้ในการบริหารงานของรัฐบาล ปัจจุบันอัลกอริทึมในการคาดการณ์ มีการศึกษาค้นคว้าเพิ่มเติม โดยเฉพาะในศาสตร์ของเหมืองข้อมูล (Data Mining) เช่นงานวิจัยของสนิทวงศ์ [1] ซึ่งได้ทำการวิจัยการคาดการณ์ผลการจัดเก็บภาษีมูลค่าเพิ่มโดยใช้กรรมวิธี neuron network ซึ่งทำการนำข้อมูลผลการจัดเก็บภาษีมูลค่าเพิ่มในอดีตมาใช้ในการวิเคราะห์ผลการจัดเก็บในเดือนต่อไป ซึ่งข้อจำกัดของงานวิจัยนี้ คือยังไม่ได้มีการนำปัจจัยภายนอกเข้ามาใช้

## ในการจัดทำารการคาดการณ์

ในงานวิจัยชิ้นนี้ จะนำเสนอการจัดทำแบบจำลองการคาดการณ์ผลการจัดเก็บมูลค่าเพิ่ม โดยพิจารณาเพิ่มเติมตัวแปรภายนอก เช่นปัจจัยของผลกระทบของภาวะเศรษฐกิจที่จะส่งต่อการจัดเก็บภาษีมูลค่าเพิ่ม เช่น ตัวแปร ดัชนีราคาผู้บริโภค ดัชนีสินค้านำเข้า ดัชนีการอุปโภคบริโภคภาคเอกชน ดัชนีการลงทุนภาคเอกชน และรายจ่ายภาครัฐในส่วนของการบริโภคที่ที่จะนำมาใช้ในการศึกษาคือ Linear Regression และ Multilayer Perceptron

## 2. วรรณกรรมที่เกี่ยวข้อง

### 2.1 การจัดทำแบบจำลองทางเศรษฐศาสตร์

การคาดการณ์ผลการจัดเก็บภาษีส่วนใหญ่จัดทำเป็นแบบจำลองทางเศรษฐศาสตร์ และส่วนใหญ่สมการถดถอย (Regression) เป็นที่นิยมใช้ในประเทศ

งานวิจัยที่ทำการศึกษถึงการคาดการณ์ การจัดเก็บภาษีมูลค่าเพิ่มในประเทศอิสราเอล Slobodnisky และ Drucker [2] มีการใช้ตัวแปรสภาพเศรษฐกิจ ที่จะทำการเพิ่มหรือลดอัตราการจัดเก็บภาษีมูลค่าเพิ่ม ซึ่งปัจจัยจะมาจากทางเศรษฐศาสตร์มหภาคเป็นหลัก ได้แก่ การผลิต การบริโภค การลงทุน การส่งออก การนำเข้า และในงานวิจัยได้กล่าวถึงแบบจำลองการคาดการณ์การจัดเก็บภาษีมูลค่าเพิ่มโดยจะนำตัวแปร อัตราการจัดเก็บภาษีมูลค่าเพิ่มการบริโภคภาคเอกชน รายจ่ายภาครัฐ ดัชนีราคาผู้บริโภคมาใช้ในการคาดการณ์ผลการจัดเก็บ ซึ่งในงานวิจัยให้ความสำคัญกับอัตราการจัดเก็บภาษีมูลค่าเพิ่มกับรายจ่ายภาครัฐจะมีความสัมพันธ์ในการจัดเก็บภาษีมูลค่าเพิ่มค่อนข้างมาก

นอกจากนี้ในงานวิจัย ที่ได้ทำการศึกษถึงการจัดเก็บภาษีของประเทศอินโดนีเซียของ Ikhsan และคณะ [3] โดยพิจารณาถึงภาษีบุคคลธรรมดา ภาษีนิติบุคคล ภาษีมูลค่าเพิ่ม โดยได้กล่าวถึงปัญหาในการจัดเก็บภาษีที่จัดเก็บได้น้อยที่มีผลจากความคลุมเครือในระบบการจัดเก็บ และในงานวิจัยจะกล่าวถึงปัจจัยการลงทุนภายใน ประเทศ ที่ส่งต่อการขยายตัวของเศรษฐกิจและทำให้การจัดเก็บภาษีจัดเก็บได้เพิ่มขึ้น และได้กล่าวถึงการคำนวณผลการจัดเก็บภาษีมูลค่าเพิ่มโดยงานวิจัยนำข้อมูล การบริโภค การลงทุน อัตราการเติบโตทางเศรษฐกิจ การนำเข้า มาใช้ในการพิจารณา

## 2.2 การถดถอยเชิงเส้น (Linear regression)

Linear regression เป็นสมการที่นิยมใช้ในงานวิจัยการจัดทำแบบจำลองทางเศรษฐศาสตร์ เนื่องจากแบบจำลองทางเศรษฐศาสตร์ จะมีแนวโน้มลักษณะความสัมพันธ์ของข้อมูลเป็นเชิงเส้น โดยลักษณะของสมการ linear regression จะเป็นการหาค่าความสัมพันธ์ของตัวแปรตามกับตัวแปรอิสระในลักษณะข้อมูลเชิงปริมาณ โดยจะมีลักษณะของ สมการคือ

$$Y = a + bx_1 + cx_2 + \dots \quad (1)$$

ซึ่งการหาค่าสัมประสิทธิ์ a b c จะมีสมการที่ใช้ดังนี้

$$na + b \sum_{i=1}^n x_{i1} + c \sum_{i=1}^n x_{i2} = \sum_{i=1}^n Y_i \quad (2)$$

$$a \sum_{i=1}^n x_{i1} + b \sum_{i=1}^n x_{i1}^2 + c \sum_{i=1}^n x_{i1}x_{i2} = \sum_{i=1}^n x_{i1}Y_i \quad (3)$$

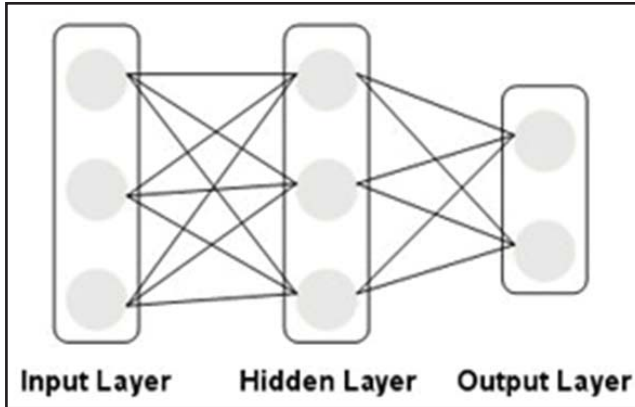
$$a \sum_{i=1}^n x_{i2} + b \sum_{i=1}^n x_{i1}x_{i2} + c \sum_{i=1}^n x_{i2}^2 = \sum_{i=1}^n x_{i2}Y_i \quad (4)$$

มีงานวิจัยทางเศรษฐศาสตร์ที่มีการนำ Linear regression มาใช้ในการจัดทำแบบจำลอง ได้แก่งานวิจัยการคาดการณ์ราคาทองคำ Ismail และคณะ [4] โดยนำตัวแปรทางเศรษฐศาสตร์เข้ามาใช้ได้แก่ ดัชนีสินค้านำเข้า อัตราการแลกเปลี่ยนค่าเงินยูโรกับค่าเงินดอลลาร์ อัตราเงินเฟ้อ ปริมาณเงินในระบบ ค่าตลาดหลักทรัพย์ New York ข้อมูลการจัดอันดับความน่าเชื่อถือ ค่าพันธบัตรรัฐบาล และดัชนีค่าเงินดอลลาร์ ซึ่งแบบจำลองที่สร้างมีความถูกต้องในการคาดการณ์ 85.2% และในงานวิจัยได้มีการหาค่า Correlation ระหว่างตัวแปรต่างๆ เพื่อนำมาใช้ในการเลือกตัวแปรที่มีความสัมพันธ์กันมาสร้างเป็นแบบจำลอง โดยงานวิจัยได้เลือกดัชนีสินค้านำเข้า อัตราการแลกเปลี่ยน อัตราเงินเฟ้อ ปริมาณเงินในระบบ มาสร้างเป็นแบบจำลองซึ่งแบบจำลองมีความถูกต้องในการคาดการณ์ 70%

## 2.3 Multilayer Perceptron

Multilayer Perceptron เป็นอัลกอริทึมรูปแบบหนึ่งของกรรมวิธี neuron network ที่มีการนำไปใช้ในงานวิจัยด้าน Data mining ที่ต้องการคาดการณ์ข้อมูลที่เป็นลักษณะเชิงปริมาณ โดยการทำงานเป็นการประมวลผล โดยใช้การคำนวณการให้น้ำหนักของการเชื่อมต่อในแต่ละโหนด ลักษณะการเชื่อมต่อเป็นแบบเครือข่ายและมีการปรับค่าถ่วงน้ำหนัก

ของการส่งผ่านข้อมูลไปในแต่ละโหนด ซึ่งจะมีการส่งค่าย้อนกลับ เพื่อปรับค่าน้ำหนักจนให้ได้ผลลัพธ์ที่ยอมรับได้ที่จะนำไปใช้ในการทำแบบจำลองต่อไป ในการสร้างเครือข่ายจะมีระดับโหนดข้อมูลเข้า (input node) และระดับโหนดปิดซ่อน (hidden node) จำนวนชั้นปิดซ่อนอาจมีได้หลายชั้น ดังภาพที่ 1



ภาพที่ 1 ลักษณะการเชื่อมต่อของ node

มีงานวิจัยทางเศรษฐศาสตร์ ที่มีการใช้ neuron network ในการคาดการณ์ ได้แก่ งานวิจัยการคาดการณ์ราคาหุ้นของตลาดหลักทรัพย์ Tehran ในประเทศอิหร่าน Ahangar และคณะ [5] ซึ่งในงานวิจัย มีการเปรียบเทียบการคาดการณ์โดยใช้ linear regression เทียบกับ neural network ซึ่งแบบจำลอง linear regression ได้มีการนำตัวแปร อัตราการเติบโตของสินค้าอุตสาหกรรม อัตราเงินเฟ้อ อัตราหนี้สินรวมของผู้ถือหุ้น ระดับของบริษัทส่วนแบ่งผลกำไรมาใช้ในการสร้างแบบจำลอง ส่วน neural network มีการใช้ตัวแปรส่วนแบ่งผลกำไร ขนาดของหุ้น อัตราหนี้สินรวมของผู้ถือหุ้น ผลรวมรายได้จากการขายอัตรา เงินเฟ้อ ปริมาณเงินในระบบ อัตราการเติบโตของสินค้าอุตสาหกรรม มาใช้ในการสร้างแบบจำลอง ซึ่งในงานวิจัยได้สรุปว่าแบบจำลอง neural network มีการคาดการณ์ราคาหุ้นที่ให้ความแม่นยำมากกว่าแบบจำลอง linear regression

### 3. วิธีการดำเนินการวิจัย

ในการศึกษาวิจัยงานต่างประเทศ [2], [3] มีการใช้ตัวแปรปัจจัยอื่นๆ เช่น ค่าดัชนีราคาผู้บริโภค ดัชนีสินค้านำเข้า ดัชนีการอุปโภคบริโภคภาคเอกชน ดัชนีการลงทุนภาคเอกชน รายจ่ายภาครัฐและผลการจัดเก็บภาษีมูลค่าเพิ่มเข้ามาใช้ในการจัดทำแบบจำลอง โดยงานที่จัดทำเป็นการ

นำเสนอ การทดลองสร้างแบบจำลองในลักษณะใหม่ ซึ่งข้อมูลต่างๆ ที่นำมาใช้งาน ได้นำมาจากการให้บริการข้อมูลผ่านทาง website ของสำนักดัชนีเศรษฐกิจการค้าและ website ของธนาคารแห่งประเทศไทย โดยลักษณะข้อมูลเป็นแบบรายเดือน

#### 3.1 การเตรียมข้อมูล

ข้อมูลที่ใช้ในงานวิจัยชิ้นนี้จัดเก็บเป็นรายเดือนตั้งแต่ปี พ.ศ. 2544 ถึงปี พ.ศ. 2552 จัดเก็บในรูปแบบไฟล์ csv

ตัวแปรในงานวิจัยนี้เป็นข้อมูลที่มาจากหลายแหล่งข้อมูล เช่น ดัชนีราคาผู้บริโภคจาก website ของสำนักดัชนีเศรษฐกิจการค้า ดัชนีสินค้านำเข้า ดัชนีการอุปโภคบริโภคภาคเอกชน ดัชนีการลงทุนภาคเอกชน รายจ่ายภาครัฐจาก website ของธนาคารแห่งประเทศไทย และข้อมูลผลการจัดเก็บภาษีมูลค่าเพิ่ม จากกรมสรรพากร โดยจะทำการแบ่งข้อมูล 70% มาทำการสอนระบบเพื่อสร้างแบบจำลองคือ ข้อมูลตั้งแต่เดือน ม.ค. ปี พ.ศ. 2544 ถึง เดือน ธ.ค. ปี พ.ศ. 2549 และแบ่งข้อมูล 30% มาทำการทดสอบการคาดการณ์ผลการจัดเก็บ คือข้อมูลตั้งแต่เดือน ม.ค. ปี พ.ศ. 2550 ถึง เดือน มี.ค. ปี พ.ศ. 2552 เพื่อตรวจสอบค่าผิดพลาดในการคาดการณ์

##### 3.1.1 การทำ Normalization

ดำเนินการจัดเตรียมข้อมูลโดยการนำข้อมูลมาจัดทำ Normalization เพื่อปรับสเกลของตัวแปรต่างๆ ให้มีความใกล้เคียงกัน โดยใช้ min-max normalization ตามสูตร

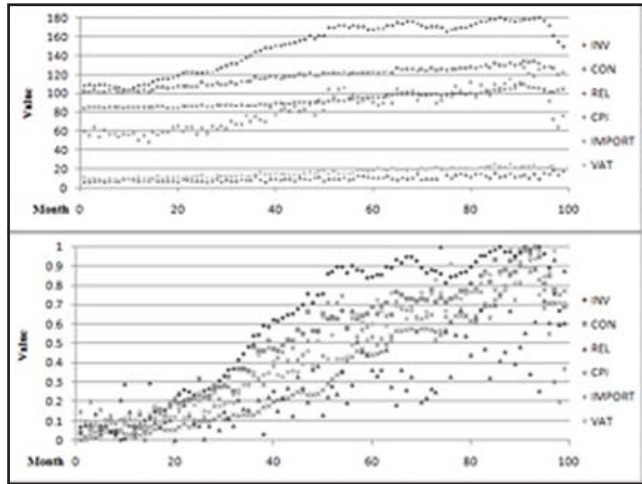
$$v' = \frac{v - \min}{\max - \min} \quad (5)$$

โดยค่า min คือค่าข้อมูลที่มีค่าน้อยที่สุดในชุดข้อมูล ค่า max คือ ค่าข้อมูลที่มีค่ามากที่สุดในชุดข้อมูล ค่า v คือ ค่าปัจจุบันที่นำมาทำ Normalization ค่า v' คือ ค่าที่ได้จากการทำ Normalization และนำข้อมูลมาใช้ในการจัดทำแบบจำลองต่อไป

##### 3.1.2 การหาค่า Information gain และ Gain ratio

ในศาสตร์ของการเหมืองข้อมูลการหาความสัมพันธ์ของตัวแปรแต่ละตัวแปรที่มีผลต่อการพยากรณ์การจัดเก็บภาษีมูลค่าเพิ่ม โดยใช้ information gain และ gain ratio เพื่อจะคัดเลือกตัวแปรไปทดสอบแบบจำลอง โดยจะทำการแบ่งช่วงข้อมูลของแต่ละตัวแปร เป็นจำนวน 10 ช่วง แล้วนำเข้าคำนวณในโปรแกรม weka ซึ่งผลความสัมพันธ์ของ

ตัวแปรกับค่า VAT (ผลการจัดเก็บภาษีมูลค่าเพิ่ม) เรียงลำดับแสดงในตารางที่ 1



ภาพที่ 2 เปรียบเทียบลักษณะการกระจายของข้อมูลก่อนทำ normalization และข้อมูลหลังจากทำ normalization แล้ว

ตารางที่ 1 ความสัมพันธ์ของตัวแปรเทียบกับค่า VAT

| Order of relation | Gain ratio, Information gain |
|-------------------|------------------------------|
| 1                 | CON (ดัชนีการอุปโภคบริโภค)   |
| 2                 | CPI (ดัชนีราคาผู้บริโภค)     |
| 3                 | INV (ดัชนีการลงทุนภาคเอกชน)  |
| 4                 | IMPORT (ดัชนีสินค้านำเข้า)   |
| 5                 | REL (รายจ่ายภาครัฐ)          |

### 3.2 การสร้างแบบจำลอง

การทดลองจะทำการสร้างแบบจำลองการคาดการณ์รายเดือนออกเป็น 2 แบบจำลองหลัก คือ แบบจำลอง Linear regression และแบบจำลอง Multilayer Perceptron

#### 3.2.1 การสร้างแบบจำลองการคาดการณ์ linear regression

การสร้างแบบจำลองแบบ Linear regression จำแนกเป็น 4 แบบย่อยตามปัจจัยนำเข้าดังนี้ คือ

##### 3.2.1.1 แบบจำลอง linear regression แบบที่ 1

แบบจำลอง linear regression แบบที่ 1 ประกอบด้วยตัวแปร ค่าดัชนีราคาผู้บริโภค ดัชนีสินค้านำเข้าดัชนีการอุปโภคบริโภคภาคเอกชน ดัชนีการลงทุนภาคเอกชนรายจ่ายภาครัฐและผลการจัดเก็บภาษีมูลค่าเพิ่ม

##### 3.2.1.2 แบบจำลอง linear regression แบบที่ 2

แบบจำลอง linear regression แบบที่ 2 เลือกตัวแปร

ตามงานวิจัยของ Slobodnisky, Drucker [2] ประกอบด้วยตัวแปร ดัชนีราคาผู้บริโภค ดัชนีการอุปโภคบริโภคภาคเอกชน รายจ่ายภาครัฐและผลการจัดเก็บภาษีมูลค่าเพิ่ม

##### 3.2.1.3 แบบจำลอง linear regression แบบที่ 3

แบบจำลอง linear regression แบบที่ 3 เลือกตัวแปรตามงานวิจัยของ Ikhsan, Trialdi, และ Syahrial [3] ประกอบด้วยตัวแปร ดัชนีการลงทุนภาคเอกชน ดัชนีการอุปโภคบริโภคภาคเอกชน ดัชนีสินค้านำเข้าและผลการจัดเก็บภาษีมูลค่าเพิ่ม

##### 3.2.1.4 แบบจำลอง linear regression แบบที่ 4

แบบจำลอง linear regression แบบที่ 4 เลือกตัวแปรโดยพิจารณาจากค่าของ information gain และ gain ratio โดยเลือกเฉพาะที่มีความสัมพันธ์มากที่สุด 3 อันดับแรกคือ ดัชนีการอุปโภคบริโภค ดัชนีราคาผู้บริโภค และดัชนีการลงทุนภาคเอกชนและผลการจัดเก็บภาษี

เมื่อทำการเตรียมข้อมูลเรียบร้อยแล้วจะทำการนำข้อมูลทั้งหมดในแต่ละแบบจำลอง เข้าคำนวณในโปรแกรม Weka ในส่วนของฟังก์ชัน linear regression เพื่อหาแบบจำลองออกมา ซึ่งเมื่อได้แบบจำลองจะทำการทดสอบการคาดการณ์ผลการจัดเก็บตั้งแต่เดือน ม.ค. ปี พ.ศ. 2550 ถึง เดือน มี.ค. ปี พ.ศ. 2552

#### 3.2.2 การสร้างแบบจำลองการคาดการณ์ Multilayer Perceptron

แบบจำลอง Multilayer Perceptron เป็นหนึ่งในตัวแบบของโครงข่ายประสาทเทียม ในการสร้างแบบจำลองเพื่อใช้ในการพยากรณ์จะเป็นการทดลองแบบสุ่มและตรวจสอบ (Trial and Error) การกำหนดปัจจัยนำเข้า จำนวนโหนด (input node) เป็นส่วนแรกของการจัดเตรียมข้อมูลก่อนการประมวลผล (Preprocessing) ดังที่ได้กล่าวในส่วนก่อนหน้าข้อมูลต้องมีการปรับสเกล เพื่อไม่ให้ปัจจัยนำเข้าแต่ละตัวแปรมีค่าข้อมูลต่างกันมาก ในงานวิจัยนี้กำหนดให้โหนดข้อมูลนำเข้าใช้ค่าตัวแปร VAT จำแนกตามช่วงเวลาที่แตกต่างกัน โดย 2 ช่วงเวลาจะเป็นการนำค่า VAT ก่อนหน้าสองเดือนมาเป็นข้อมูลนำเข้า (Mul VAT 2) เพื่อพยากรณ์เดือนถัดไป ในงานวิจัยนี้จะกำหนดช่วงเวลาตั้งแต่ 2 จนถึง 11 ดังนั้น Mul VAT 11 จะมีข้อมูลนำเข้าค่า VAT 11 เดือนก่อนหน้ามาเป็นข้อมูลนำเข้าเพื่อพยากรณ์เดือนถัดไปและมีการทดสอบการกำหนดโหนดข้อมูลนำเข้าโดยใช้ตัวแปร

จากแบบจำลอง linear regression แบบที่ 1 คือนำค่าตัวแปร CPI CON INV REL IMPORT มาใช้ (Mul All 5) และจำแนกตัวแปรทุกตัวตามช่วงเวลาที่แตกต่างกัน โดยใช้ 3 ช่วงเวลาดูก่อนหน้ามาเป็นโหนดข้อมูลนำเข้า (Mul All 15) ดังแสดงในตารางที่ 2 โดยโหนดผลลัพธ์ คือ ค่าพยากรณ์ของการจัดเก็บภาษีของเดือนถัดไป ทั้งนี้การสร้างแบบจำลองดังกล่าวก็เพื่อใช้ในการพยากรณ์ของการจัดเก็บภาษีของเดือนถัดไป นอกจากนี้การใช้ Perceptron ดังกล่าว นักวิจัยยังต้องกำหนดค่าตัวแปรหลายตัวซึ่งนักวิจัยควรมีประสบการณ์ในการกำหนดตัวแปรเหล่านี้ประกอบด้วย ตัวแปรการเรียนรู้ (Learning rate), ตัวแปรกระตุ้นการปรับค่าน้ำหนัก (Momentum) จำนวน Hidden Layer ตัวแปรที่ใช้ในการตัดสินใจหยุดการสอบระบบเมื่อความผิดพลาดน้อยกว่าค่าที่กำหนด (validation set size) และท้ายสุดตัวแปรจำนวนรอบของการสอนระบบ Training Time/Epochs การกำหนดข้อมูลทดสอบแสดงในตารางที่ 2

ตารางที่ 2 การสร้างแบบจำลอง Multilayer Perceptron

| Model      | Input Nodes | Hidden Layer | Learning Rate | Momentum |
|------------|-------------|--------------|---------------|----------|
| Mul VAT 2  | 2           | 65           | 0.025         | 0.025    |
| Mul VAT 3  | 3           | 35           | 0.025         | 0.3      |
| Mul VAT 4  | 4           | 35           | 0.1           | 0.2      |
| Mul VAT 5  | 5           | 35           | 0.025         | 0.2      |
| Mul VAT 6  | 6           | 35           | 0.025         | 0.3      |
| Mul VAT 7  | 7           | 35           | 0.025         | 0.05     |
| Mul VAT 8  | 8           | 45           | 0.05          | 0.2      |
| Mul VAT 9  | 9           | 35           | 0.05          | 0.2      |
| Mul VAT 10 | 10          | 25           | 0.05          | 0.1      |
| Mul VAT 11 | 11          | 35           | 0.025         | 0.05     |
| Mul All 5  | 5           | 25           | 0.025         | 0.025    |
| Mul All 15 | 15          | 25           | 0.05          | 0.05     |

### 3.3 ทดสอบการคาดการณ์โดยใช้เทคนิค Moving average

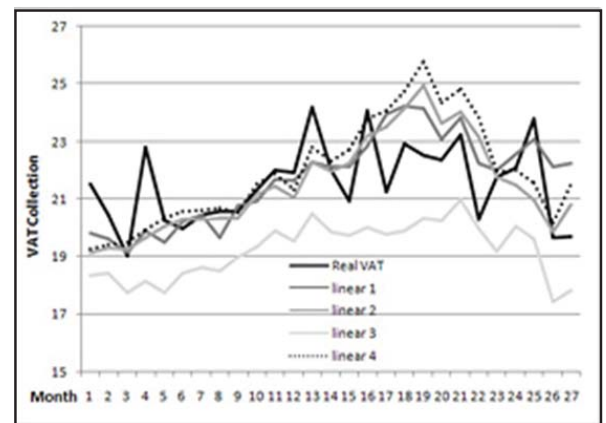
การใช้เทคนิค Moving average ทำโดยการนำค่าผลการจัดเก็บภาษีมูลค่าเพิ่ม เฉลี่ย 3 เดือนก่อนหน้ามาคาดการณ์ในเดือนถัดไป โดยจะทำการทดสอบการคาดการณ์ผลการจัดเก็บ ตั้งแต่เดือน ม.ค. ปี พ.ศ. 2550 ถึง เดือน มี.ค. ปี พ.ศ. 2552 เพื่อนำมาใช้ในการเปรียบเทียบผลการทดสอบการคาดการณ์ผลการจัดเก็บภาษีมูลค่าเพิ่มกับแบบจำลองแบบต่างๆ

## 4. ผลการศึกษา

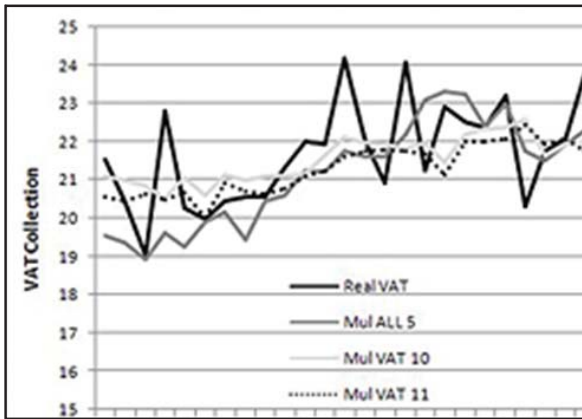
ผลการสร้างแบบจำลองการคาดการณ์ผลการจัดเก็บภาษีมูลค่าเพิ่มรายเดือน ในการทดสอบแบบจำลอง linear regression Multilayer Perceptron และการคาดการณ์โดยใช้เทคนิค Moving average ผลการทดสอบแสดงในตารางที่ 3

ตารางที่ 3 การคาดการณ์ของแบบจำลองที่ใช้ตัวแปรแบบต่างๆ โดยพิจารณา ค่าผิดพลาด Mean Absolute Deviation (MAD) หน่วย พันล้านบาท

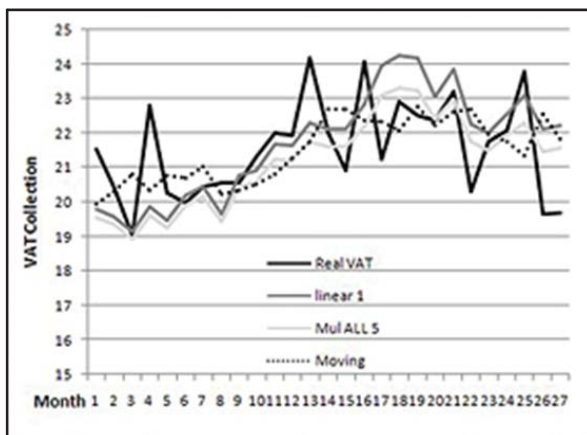
| Model      | CPI | INV | CON | RFL | IMPORT | VAT | MAD   |
|------------|-----|-----|-----|-----|--------|-----|-------|
| Linear 1   | 0   | 0   | 0   | 0   | 0      | -   | 1.057 |
| Linear 2   | 0   | -   | 0   | 0   | -      | -   | 1.086 |
| Linear 3   | -   | 0   | 0   | -   | 0      | -   | 2.318 |
| Linear 4   | 0   | 0   | 0   | -   | -      | -   | 1.197 |
| Mul VAT 2  | -   | -   | -   | -   | -      | 0   | 1.227 |
| Mul VAT 3  | -   | -   | -   | -   | -      | 0   | 1.068 |
| Mul VAT 4  | -   | -   | -   | -   | -      | 0   | 1.139 |
| Mul VAT 5  | -   | -   | -   | -   | -      | 0   | 1.146 |
| Mul VAT 6  | -   | -   | -   | -   | -      | 0   | 1.073 |
| Mul VAT 7  | -   | -   | -   | -   | -      | 0   | 1.053 |
| Mul VAT 8  | -   | -   | -   | -   | -      | 0   | 1.064 |
| Mul VAT 9  | -   | -   | -   | -   | -      | 0   | 1.047 |
| Mul VAT 10 | -   | -   | -   | -   | -      | 0   | 1.013 |
| Mul VAT 11 | -   | -   | -   | -   | -      | 0   | 1.015 |
| Mul All 5  | 0   | 0   | 0   | 0   | 0      | -   | 0.990 |
| Mul All 15 | 0   | 0   | 0   | 0   | 0      | -   | 1.327 |
| Moving     |     |     |     |     |        | 0   | 1.143 |



ภาพที่ 3 การจัดเก็บภาษีมูลค่าเพิ่มที่เก็บจริงเปรียบเทียบกับแบบจำลอง linear regression



ภาพที่ 4 การจัดเก็บภาษีมูลค่าเพิ่มที่เก็บจริงเปรียบเทียบกับแบบจำลอง Multilayer Perceptron



ภาพที่ 5 การจัดเก็บภาษีมูลค่าเพิ่มที่เก็บจริงเปรียบเทียบกับแบบจำลองแบบต่างๆ

## 5. บทสรุป

ผลการสร้างแบบจำลองแบบต่างๆ และการทดสอบการคาดการณ์ผลการจัดเก็บภาษีมูลค่าเพิ่ม จะเห็นว่า การคาดการณ์ผลการจัดเก็บรายเดือนโดยใช้แบบจำลอง Multilayer Perceptron โดยใช้ค่าตัวแปรทางเศรษฐศาสตร์ คือ ดัชนีราคาผู้บริโภค ดัชนีสินค้านำเข้า ดัชนีการอุปโภคบริโภคภาคเอกชน ดัชนีการลงทุนภาคเอกชน รายจ่ายภาครัฐ จำนวนตัวแปรละ 1 โหนดนำเข้า (Mul All 5) จะมีค่าผิดพลาดในการคาดการณ์น้อยที่สุด (ค่าผิดพลาด Mean Absolute

Deviation = 0.99 พันล้านบาท) ซึ่งตัวแปรเหล่านี้เป็นตัวแปรที่สามารถกระตุ้นได้ซึ่งจะทำให้ใช้ประโยชน์ในการคำนวณการวางแผนการจัดเก็บภาษีมูลค่าเพิ่มของทางภาครัฐได้ แต่ทั้งนี้ในการสร้างแบบจำลองเพื่อการใช้งานจริงต้องพิจารณาถึงวัตถุประสงค์ของตัวแปรที่ต้องการศึกษาถึงผลกระทบต่อผลการคาดการณ์ที่จะนำมาใช้ด้วย

## 6. เอกสารอ้างอิง

- [1] สนิทวงศ์ กมลภากรณ์ การประมาณผลการจัดเก็บภาษีมูลค่าเพิ่มโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพากชัน (Forecasting of VAT Collection by Backpropagation Neural Network). รายงานสาขาวิชาเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 1999.
- [2] Tatiana Slobodnitsky, "Lev Drucker VAT Revenue Forecasting in Israel," *Paper The Maurice Falk Institute for Economic Research in Israel*, pp. 1-19, 2007.
- [3] Mohamad Ikhsan, Ledi Trialidi, "Syarif Syahril Indonesia's new tax reform: Potential and direction," *Journal of Asian Economics* 16 2005, pp. 1029-1046, 2005.
- [4] Z.Ismail, A.Yahya, "A.Shabri Forecasting Gold Prices Using Multiple Linear Regression Method," *American Journal of Applied Sciences* 6, pp. 1509-1514, 2009.
- [5] Reza Gharoie Ahangar, "Mahmood Yahyazadehfard, Hassan Pournaghshband The Comparison of Methods Artificial Neural Network with Linear Regression Using Specific Variables for Prediction Stock Price in Tehran Stock Exchange, *International Journal of Computer Science and Information Security* Vol.7 No.2, pp. 38-46, 2010.