การทำนายผลผลิตข้าวโดยวิธีซัพพอร์ตเวกเตอร์แมชีน

รัชภูมิ ใจกล้า[1] พันธุ์ทิพย์ นนทรีย์[1] และ อรรถชัย จินตะเวช[1]

บทคัดย่อ

ประเด็น (Background):_การทำนายผลผลิตพืชจำเป็นต้องใช้แบบจำลองพืชและการทดลองจริง เพื่อการทดสอบความแม่นยำของแบบจำลอง ซึ่งในการทดลองในภาคสนามต้องใช้เวลานานและมีค่าใช้จ่ายสูง (สุนทร และ เมธี, 2535) เพื่อเป็นการประหยัดเวลาและค่าใช้จ่ายในการทำแบบจำลองพืช จึงได้เสนอวิธีการทำนายผลผลิตข้าวโดยวิธีซัพพอร์ทเวกเตอร์แมชีน

จุดประสงค์: เพื่อพัฒนาแบบจำลองการทำนายผลผลิตข้าว โดยใช้ข้อมูลนำเข้าของ DSSAT 4.0 (Minimum Data Set) ในการเรียนรู้และการทดสอบแบบจำลองการทำนายผลผลิตข้าวโดยวิธีซัพพอร์ทเวกเตอร์แมชีน

วิธีการศึกษา: พัฒนาแบบจำลองโดยใช้ข้อมูลที่ได้จากการคาดการณ์ผลผลิตข้าวจากการใช้โปรแกรม DSSAT 4.0 เป็นชุดข้อมูลสำหรับการเรียนรู้ (Learning) และสำหรับการทดสอบ โดยใช้โปรแกรม MATLAB ในการพัฒนาแบบจำลองการทำนายผลผลิตข้าวสำหรับการทำนายผลผลิตข้าวพันธุ์ กข23 ชุดดิน สุพรรณบุรี

ผลการศึกษา: จากการทำนายผลผลิตข้าวโดยวิธีซัพพอร์ทเวกเตอร์แมชีน พบว่าแบบจำลองสามารถทำนายได้มีความใกล้เคียงกับผลที่ได้จากการคาดการณ์โดยโปรแกรม DSSAT 4.0

สรุป: แบบจำลองทำนายผลผลิตข้าวนี้สามารถทำนายได้สำหรับข้าวพันธุ์ กข23 ชุดดินสุพรรณบุรีเท่านั้น ซึ่งในอนาคตจะสามารถทำได้หลายพันธุ์และหลายชุดดิน โดยจะต้องมีการเรียนรู้และทดสอบสำหรับข้าวพันธุ์และชุดดินอื่นๆ อีกต่อไป

คำสำคัญ: การทำนายผลผลิตข้าว, ซัพพอร์ตเวกเตอร์แมชีน

บทนำ

ในเอกสารฉบับนี้ได้เสนอวิธีการทำนายผลโดยใช้วิธีการทำนายโดยการใช้ซัพพอร์ทเวกเตอร์แมชีน สำหรับวิเคราะห์การถดถอย (Support Vector Regression: SVR) ซึ่งได้นำเอาผลที่ได้จากโปรแกรม DSSAT 4.0 ใช้เป็นชุดข้อมูลสำหรับการทดสอบการทำนายนี้ ในที่นี้ได้ทำการทดลองการทำนายผลผลิตข้าวพันธุ์ กข23 ชุดดิน สุพรรณบุรี ในปี 1963-1978 และข้อมูลอากาศของ สถานีทดลองเกษตรเขตชลประทาน นั่นก็หมายความว่าการทำนายผลผลิตข้าวนั้น จะแยกแบบจำลองออกไปตาม พันธุ์ข้าวและชุดดิน ในเอกสารฉบับนี้ได้ทำการทดลอง 1 แบบจำลอง คือ ข้าวพันธุ์ กข23 ชุดดิน สุพรรณบุรี

ตรวจเอกสาร

เนื่องจากงานวิจัยนี้เป็นการประยุกต์ใช้ซัพพอร์ตเวกเตอร์แมชีน เพื่อการประมาณค่าฟังก์ชันหรือเรียกอีกอย่างว่าการวิเคราะห์การถดถอย (Regression) เนื้อหาหลักของซัพพอร์ตเวกเตอร์แมชีนจึงเกี่ยวกับการวิเคราะห์การ

[1] ศูนย์วิจัยเพื่อเพิ่มผลผลิตทางเกษตร คณะเกษตรศาสตร์ มหาวิทยาลัยเชียงใหม่

ถดถอย ส่วนทฤษฎีของซัพพอร์ตเวกเตอร์แมชีนสำหรับการแบ่งกลุ่มจะกล่าวถึงให้รู้หลักการเบื้องต้นพอสังเขป ดังนั้นเนื้อหาในบทนี้จึงประกอบด้วย พื้นฐานที่เกี่ยวกับซัพพอร์ตเวกเตอร์แมชีน ซัพพอร์ตเวกเตอร์แมชีนสำหรับการแบ่งกลุ่ม และในส่วนสุดท้ายจะกล่าวถึงซัพพอร์ตเวกเตอร์แมชีนสำหรับการถดถอย

พื้นฐานที่เกี่ยวกับซัพพอร์ตเวกเตอร์แมชีน

ซัพพอร์ตเวกเตอร์แมชีนเป็นเครื่องช่วยเรียนรู้ (Machine Learning Tasks) ที่มีพื้นฐานมาจากทฤษฎีการเรียนรู้จากสถิติ (Statistical Learning Theory) (T. Xu, X.N. Dong) เหมือนกับโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) แต่โครงสร้างแตกต่างกัน นั่นคือ ซัพพอร์ตเวกเตอร์แมชีนใช้หลักการลดความเสี่ยงเชิงโครงสร้างให้ต่ำสุด (Structural Risk Minimization: SRM) ในขณะที่โครงข่ายประสาทเทียมใช้หลักการลดความเสี่ยงเชิงทดลองให้ต่ำสุด (Empirical Risk Minimization: ERM) ซัพพอร์ตเวกเตอร์แมชีนสามารถประยุกต์ใช้กับงาน 2 รูปแบบคือ การแบ่งกลุ่ม (Classification) และการวิเคราะห์การถดถอย (Regression) หรือการประมาณค่าฟังก์ชันสำหรับการทำนายผลผลิตข้าวเป็นการประยุกต์ใช้เพื่อประมาณค่าฟังก์ชัน

ซัพพอร์ตเวกเตอร์แมชีนสำหรับการถดถอย (Support Vector Regression: SVR)

โครงสร้างของซัพพอร์ตเวกเตอร์แมชีนสำหรับการถดถอยแตกต่างจากโครงสร้างของซัพพอร์ตเวกเตอร์แมชีนสำหรับการแบ่งกลุ่ม แต่อยู่บนหลักการของการใช้ซัพพอร์ตเวกเตอร์เพื่อเป็นตัวแทนของกลุ่มข้อมูลที่ใช้สำหรับสอนเช่นเดียวกัน

หลักการหาระนาบเชิงเส้น

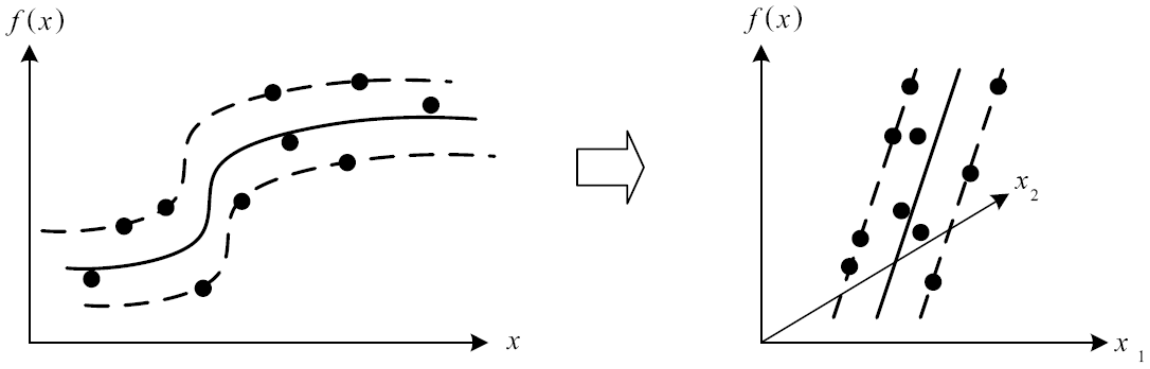
หลักการหาระนาบเชิงเส้นของซัพพอร์ตเวกเตอร์แมชีนสำหรับการถดถอย เริ่มต้นที่การสอนซัพพอร์ตเวกเตอร์แมชีนด้วยเซตข้อมูล $\{x_i, y_i\}_{i=1}^{\ell}, x_i \in \Re^n, y_i \in \Re$ เมื่อ x_i คือเวกเตอร์ข้อมูลอินพุต y_i คือข้อมูลเอาต์พุต เพื่อหาฟังก์ชันเชิงเส้น $f(x)$ ที่ใช้แทนสมการระนาบเกิน (Hyperplane) ของกลุ่มข้อมูลที่นำมาฝึกสอนทั้งหมดดังสมการ (1) เมื่อ w คือเวกเตอร์น้ำหนักและ b คือค่าไบอัส ดังนั้นเมื่อมีเวกเตอร์ x ใดๆ ก็จะสามารถประมาณค่า y ได้โดยใช้สมการ (1) นั่นคือ $f(x) = y'$ นั่นเอง

$$f(x) = \langle w \bullet x \rangle + b \tag{1}$$

ก่อนการสร้างระนาบเกิน สถาปัตยกรรมของซัพพอร์ตเวกเตอร์แมชีนได้มีการออกแบบให้มีการแมปข้อมูลจากปริภูมิอินพุตไปอยู่ในปริภูมิที่สูงขึ้นเรียกว่าปริภูมิลักษณะเด่นด้วยฟังก์ชันเคอร์เนล เพื่อให้สามารถหาฟังก์ชันเชิงเส้นที่เหมาะสมสำหรับใช้แทนสมการระนาบเกินได้ ดังในภาพที่ 1 ซึ่งแสดงถึงความแตกต่างในการวิเคราะห์การถดถอยแบบไม่เชิงเส้นในปริภูมิอินพุตและการวิเคราะห์การถดถอยเชิงเส้นในปริภูมิลักษณะเด่น ทำให้เวกเตอร์ข้อมูลอินพุตเปลี่ยนไปเป็นข้อมูลที่มีมิติสูงขึ้นในปริภูมิลักษณะเด่น การแมปข้อมูลด้วยฟังก์ชันเคอร์เนลแสดงดังสมการ (2) เมื่อ Φ คือฟังก์ชันการแมป (Mapping Function) นั่นคือ $x_i \Rightarrow K(x_i, x_j)$ ทำให้สมการระนาบเกินที่ต้องการหาเปลี่ยนเป็นสมการ (3)

$$K(x_i, x_j) = \langle \Phi(x_i) \bullet \Phi(x_j) \rangle \tag{2}$$

$$f(x) = \langle w \bullet K(x_i, x_j) \rangle + b \tag{3}$$

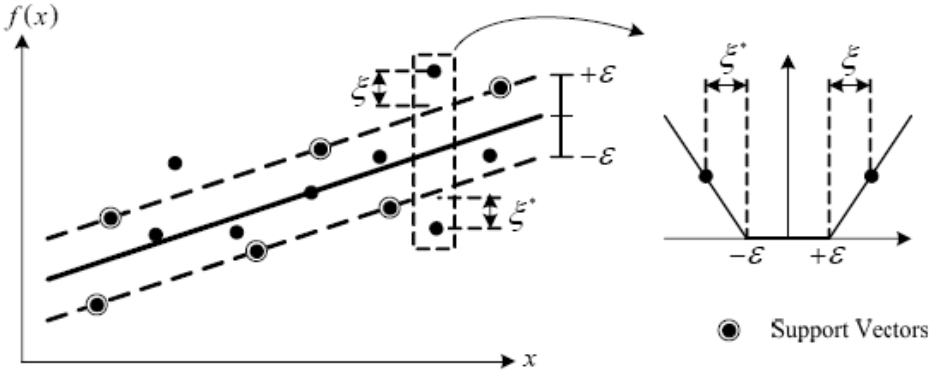


ปริภูมิอินพุต ปริภูมิลักษณะเด่น

ภาพที่ 1 การวิเคราะห์การถดถอยแบบไม่เชิงเส้นในปริภูมิอินพุต และการวิเคราะห์การถดถอยเชิงเส้นในปริภูมิลักษณะเด่น

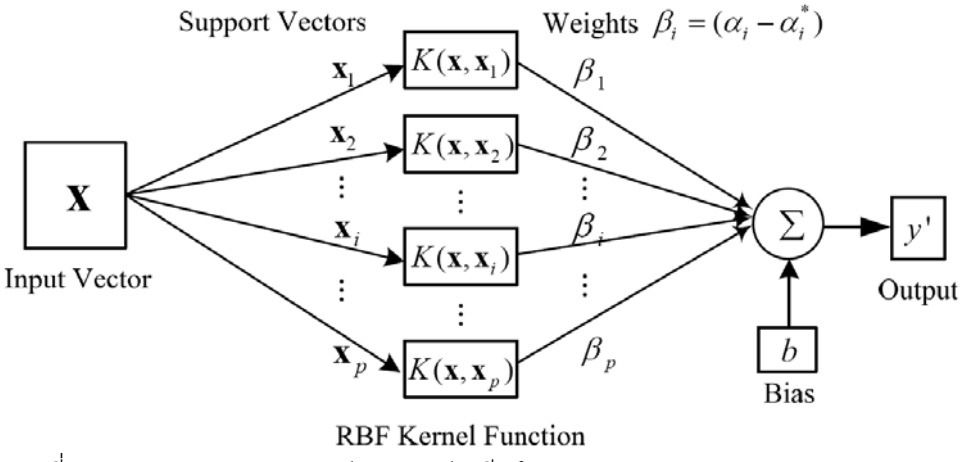
การสร้างระนาบเกิน $f(x)$ เพื่อประมาณค่าให้ใกล้เคียงกับ y ที่สุดจะต้องหาค่า w และ b ที่หมาะสมที่สุด การสร้างระนาบเกินสำหรับกรณีของซัพพอร์ตเวกเตอร์แมชีนสำหรับการถดถอย จะพิจารณาค่าความคลาดคลื่อนที่ยอมรับได้ ซึ่งอยู่ในรูปฟังก์ชันการสูญเสีย (Loss Function) และสำหรับกรณีการพยากรณ์ความต้องการไฟฟ้าเป็นการทำนายอนุกรมเวลา (Time Series Prediction) ซึ่งมีฟังก์ชันการสูญเสียให้เลือกใช้ 2 ชนิดคือ $\varepsilon - Insensitive$ และ Huber's Robust Loss Function (K.R. Muller) แต่ที่นิยมใช้มากกว่าคือ $\varepsilon - Insensitive$ ซึ่งมีความสัมพันธ์ดังสมการ (4) เมื่อ $L(y_i, f(x))$ คือ ความคลาดเคลื่อนที่เกิดขึ้น (ความแตกต่างระหว่าง y_i และ $f(x)$) โดยความหมายของสมการ(4) คือ ค่าความคลาดเคลื่อนที่น้อยกว่าหรือเท่ากับ ε ให้ถือว่าเป็นศูนย์และเป็นค่าความคลาดเคลื่อนที่ผู้ใช้งานยินยอมให้เกิดขึ้นได้ เช่นในกรณีการประยุกต์ใช้เพื่อทำนายผลผลิตข้าวกำหนดให้ ε = 0.001 ก็หมายความว่าผู้พยากรณ์ยอมรับผลการพยากรณ์ในสองหลักแรก และถือว่าการพยากรณ์ในหลักที่ 3 ไม่ค่อยมีความสำคัญและคลาดเคลื่อนได้เท่ากับ 0.001 นั่นเอง

$$L(y_i, f(x)) = \begin{cases} 0 & for |y_i - f(x)| \le \varepsilon \\ |y_i - f(x)| - \varepsilon & othorwise \end{cases} \quad (4)$$



ภาพที่ 2 การหาระนาบเกินที่เหมาะสมที่สุดสำหรับใช้แทนกลุ่มข้อมูลและการกำหนดแนวขอบระนาบเกินด้วยฟังก์ชัน $\varepsilon - Insensitive$

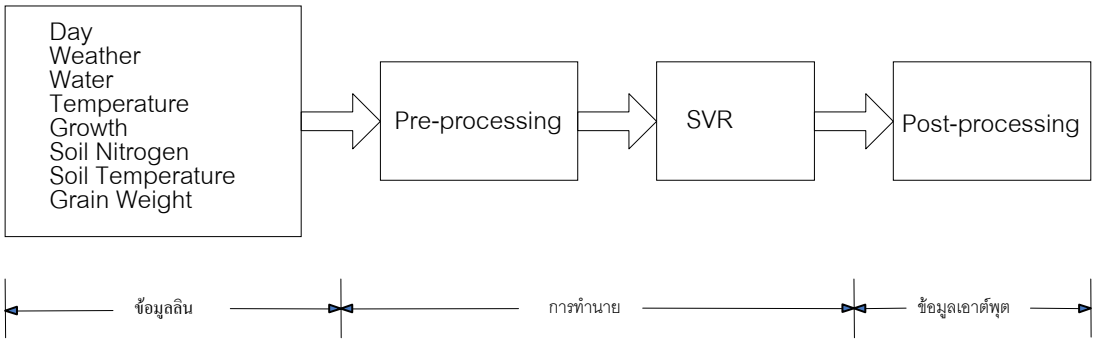
สถาปัตยกรรมของซัพพอร์ตเวกเตอร์แมชีนสำหรับการถดถอยในกรณีนี้แสดงดังภาพที่ 3 เมื่อ y' คือ ค่าประมาณของ y ที่มีเวกเตอร์ x เป็นลักษณะเด่น ซัพพอร์ตเวกเตอร์ x_i เวกเตอร์น้ำหนัก β_i และ ค่าไบอัส (b) ได้มาจากขั้นตอนการสอนซัพพอร์ตเวกเตอร์แมชีน



ภาพที่ 3 สถาปัตยกรรมของซัพพอร์ตเวกเตอร์แมชีนสำหรับการถดถอย

วิธีการ

ขั้นตอนในการทำการทำนายนี้ได้แบ่งออกเป็น 3 ส่วน คือ ส่วนของข้อมูลอินพุต (Input data) การทำนาย (Prediction) และส่วนของข้อมูลเอาต์พุต (Output data) ดังแสดงในภาพที่ 4



ภาพที่ 4 ขั้นตอนการทำนายผลผลิตของข้าว

ข้อมูลอินพุต คือ ข้อมูลผลผลิตของข้าวและปัจจัยในอดีตที่เกี่ยวข้อง โดยข้อมูลผลผลิตของข้าวในอดีตและปัจจุบันมีหน่วยกิโลกรัมต่อไร่ (kg/rai) ส่วนปัจจัยที่เกี่ยวข้องประกอบด้วย ข้อมูล ข้าวพันธุ์ กข. 23 ข้อมูลชุดดิน สุพรรณบุรี วันที่จะทำการทำนาย สภาพอากาศ ปริมาณน้ำในดิน อุณหภูมิสะสม ปริมาณไนโตรเจน อุณหภูมิในดิน และข้อมูลผลผลิตของข้าวในวันที่ผ่านมา

การทำนาย งานวิจัยนี้ใช้ซัพพอร์ตเวกเตอร์แมชีนสำหรับการถดถอย (Support Vector Regression: SVR) ในการพยากรณ์ และทำการเปรียบเทียบกับการพยากรณ์โดยใช้โครงข่ายประสาท (Neural Network: NN) ซัพพอร์ตเวกเตอร์แมชีนสำหรับการถดถอยจะใช้ฟังก์ชันการสูญเสียแบบ $\varepsilon - Insensitive$ Loss Function และเลือกทดลองใช้ฟังก์ชันเคอร์เนล คือ ฟังก์ชันเรเดียลเบซิส (RBF Kernel Function) ในขั้นตอนการทำนายนี้จะมีการประมวลผลเบื้องต้น (Pre-Processing) เพื่อทำให้ข้อมูลสำหรับที่จะนำไปประมวลผลมีขนาดเล็กลง อยู่ในช่วงตั้งแต่

–1 ถึง +1 ทำให้เหมาะต่อการคำนวณ และเมื่อมีการประมวลผลเบื้องต้น จะต้องมีการประมวลผลตอนหลัง (Post-Processing) เพื่อแปลงข้อมูลที่มีขนาดน้อยๆ กลับมาเป็นข้อมูลจริง

ข้อมูลเอาต์พุต เป็นข้อมูลของผลผลิตที่ได้จากการทำนายในขั้นตอนนี้สิ่งสำคัญอีกประการคือ การหาค่าผิดพลาดจากการทำนาย งานวิจัยนี้ประเมินประสิทธิภาพอัลกอริทึมการพยากรณ์ คือ การหาค่าความผิดพลาดสมบูรณ์เฉลี่ย (Mean Absolute Error: MAE)

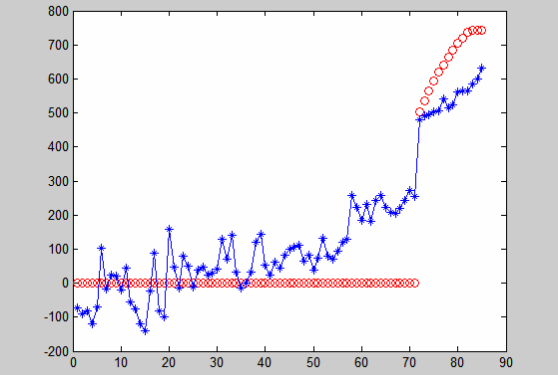
การออกแบบการทดลองและการแสดงประสิทธิภาพของการทำนาย

การออกแบบการทดลอง

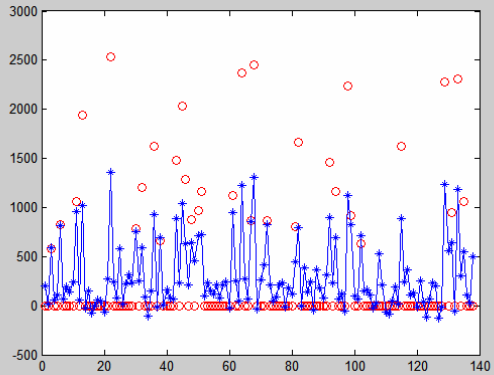
งานวิจัยนี้แบ่งข้อมูลออกเป็นสองส่วน คือ ชุดข้อมูลที่ใช้สำหรับการสอน(Training Set) หรือเรียนรู้ (Learning) และชุดข้อมูลที่ใช้สำหรับทดสอบประสิทธิภาพของระบบ (Test Set) การแบ่งข้อมูลออกเป็นสองส่วนนี้เพื่อจะหาแบบจำลองที่ดีที่สุดสำหรับการทำนายโดยข้อมูลสำหรับสอนเป็นข้อมูลตั้งแต่ปี 1963 ถึงปี 1977 ระยะเวลา 15 ปี จำนวน 10×1375 แซมเปิล (Samples) ส่วนข้อมูลสำหรับทดสอบเป็นข้อมูลปี 1978 จำนวน10×87 แซมเปิล

ผลการศึกษา

ในการทดสอบการทำนายผลผลิตข้าวกับชุดข้อมูลในปี 1978 นั้น ผลผลิตข้าวหลังการเก็บเกี่ยวที่ได้จากการทำนายมีความใกล้เคียงกับผลผลิตข้าวหลังการเก็บเกี่ยวที่ได้ในปี 1978 คือ ผลที่ได้จากการทำนายมีปริมาณน้อยกว่าผลผลิตที่ได้ในปี 1978 ไป 109.12 (kg/rai) และค่า MAE เท่ากับ 651.76 ดังภาพที่ 5 ส่วนผลที่ได้จากการทดสอบแบบ 10% cross validation นั้นได้ได้ภาพที่ 6 ซึ่งค่า MAE เท่ากับ 240.24 ซึ่งค่าพารามิเตอร์ที่ใช้ในการเรียนรู้ คือ $\varepsilon = 0.82$, $C = 0.002$ และ $\sigma = 15$



ภาพที่ 5 ผลที่ได้จากการทดสอบกับข้อมูลปี 1978



ภาพที่ 6 ผลที่ได้จากการทดสอบแบบ 10% cross validation

สรุป

ผลการทำนายผลผลิตข้าวนั้นประกอบไปด้วยชุดข้อมูลสำหรับการเรียนรู้ ประกอบไปด้วย ข้อมูลวัน ข้อมูลอากาศ ข้อมูลน้ำในดิน ข้อมูลอุณหภูมิสะสม ข้อมูลอุณหภูมิในดิน ข้อมูลไนโตรเจนในดินและ ข้อมูลผลผลิตข้าว ซึ่ง

เป็นข้อมูลในปี 1963 – 1977 ผลที่ได้คือการหาค่า MAE โดยการทำ 10% cross validation ได้ 240.24 และผลการทดสอบกับข้อมูลที่ไม่เคยเห็นมาก่อน (Blind Test Set) เท่ากับ 651.76

ซึ่งในแบบจำลองนี้ สามารถทำนายได้เฉพาะข้าวพันธุ์ กข. 23 และชุดดินสุพรรณบุรี เท่านั้น เนื่องจากในการทำการเรียนรู้ (Learning) ให้กับแบบจำลองนั้นทำเฉพาะข้าวพันธุ์ กข.23 และชุดดินสุพรรณบุรีเท่านั้น ดังนั้นถ้าหากต้องการให้แบบจำลองนี้สามารถทำนายข้าวได้หลายพันธุ์และหลายชุดดินก็จำเป็นที่จะต้องมีชุดข้อมูลที่จะสอนให้กับแบบจำลองที่หลากหลายขึ้น

เอกสารอ้างอิง

สุนทร บูรณะวิริยะกูล และ เมธี เอกะสิงห์. 2535. การทดสอบแบบจำลอง CERES-Rice สำหรับผลผลิตข้าวในสภาพของเชียงใหม่. คณะเกษตรศาสตร์: มหาวิทยาลัยเชียงใหม่

T. Xu, R. He, P. Wang and D. Xu, “Input Dimension Reduction for Load Forecasting Based on Support Vector Machines,” *Proceedings of the 2004 IEEE International Conference on Electric Utility eregulation, Restructuring and Power Technologies*, pp.510-514, Hong Kong, Apr.2004.

X.N. Dong, W. Qiang and L. Jin-Chao, “Short Term Load Forecasting Model Using Support Vector Machine Based on Artificial Neural Network,” *Proceedings of Fourth International Conference on Machine Learning and Cybernetics,* pp.4260-4265, Guangzhou, Aug.2005.

K.R. Muller, A.J. Smola, G. Ratsch, B. Scholkopf, J. Kohlmorgen and V. Vapnik, “Predicting Time Series with Support Vector Machine,” *Proceedings of ICANN’97, Springer LNCS 1327*, pp.999-1004, 1997.