

การทำนายสุขภาพดินของพืชโดยใช้โครงข่ายประสาทเทียมเชิงลึก

Prediction of soil health for plants using deep neural networks

ธวัชชัย พรหมรัตน์¹ และ ดุษฎี คำบุญเรือง^{2*}
Thawatchai Promrat¹ and Dusadee Khamboonrueng^{2*}

Received: 1 December 2023 ; Revised: 15 January 2024 ; Accepted: 9 February 2024

บทคัดย่อ

วัตถุประสงค์ของงานวิจัยนี้เพื่อสร้างโมเดลทำนายสุขภาพดินที่ ณ ปัจจุบัน โดยการจำลองสภาพแวดล้อมในการปลูกต้นถั่วเหลือง 4 สภาพแวดล้อมสภาพแวดล้อมละ 3 กระถางคือ สภาพแวดล้อมที่หนึ่งรดน้ำและใส่ปุ๋ยเคมี สภาพแวดล้อมที่สองรดน้ำและใส่ปุ๋ยอินทรีย์ สภาพแวดล้อมที่สามรดน้ำเพียงอย่างเดียว สภาพแวดล้อมที่สี่ไม่รดน้ำและไม่ใส่ปุ๋ย และได้พัฒนาระบบอินเทอร์เน็ตของสรรพสิ่งเพื่อใช้ในการเก็บข้อมูลสุขภาพดินเป็นระยะเวลา 2 เดือนทำให้ได้ข้อมูลทั้งหมด 475 เดต้าพอยต์ จากนั้นนำข้อมูลมาจัดกลุ่มเพื่อหาจำนวนกลุ่มที่เหมาะสมที่สุดด้วยอัลกอริทึม k-means ทำให้ได้จำนวนกลุ่มที่เหมาะสมกับข้อมูลสุขภาพดิน 3 กลุ่ม และให้ความหมายสุขภาพดินแต่ละกลุ่มคือ เหมาะสม ปานกลาง ปรับปรุง และระบุเลเบลให้กับข้อมูลสุขภาพดินแต่ละเดต้าพอยต์ที่เหมาะสมกับกลุ่ม เพื่อนำข้อมูลมาสร้างโมเดลทำนายสุขภาพดิน ณ ปัจจุบันโดยใช้อัลกอริทึมโครงข่ายประสาทเทียมเชิงลึกจากการทดลองทำให้ได้ความถูกต้องในการทำนายสุขภาพดิน 98 เปอร์เซนต์

คำสำคัญ: การทำนาย, โครงข่ายประสาทเทียมเชิงลึก, การจัดกลุ่ม, อินเทอร์เน็ตของสรรพสิ่ง, ดิน

Abstract

The objective of this research is to create a soil health prediction model at present by simulating the environmental conditions for growing yellow beans in four different environments. These environments involve three pots each, representing the following conditions: environment 1: watering and chemical fertilizer application, environment 2: watering and organic fertilizer application, environment 3: watering only, and environment 4: no watering or fertilization. Additionally, an Internet of Things (IoT) system was developed to collect soil health data over 2 months, resulting in a total of 475 data points. The collected data was then grouped to assign the most suitable number of clusters using the K-means algorithm. This process yielded three clusters that best fit the soil health data, each corresponding to different levels of soil health: Suitable, Moderate, and Improvement. Labels were assigned to each data point within the clusters, indicating the corresponding soil health level. Subsequently, the data was utilized to create a soil health prediction model using deep-learning neural network algorithms. Through experimentation, the model achieved 98% accuracy in predicting soil health.

Keywords: Predicting, Deep neural network, Clustering, Internet of things, Soil

¹ สาขาวิชาเทคโนโลยีสารสนเทศ ภาควิชาวิทยาศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏพระนครศรีอยุธยา 13000

² สาขาวิชาฟิสิกส์และวิทยาศาสตร์ทั่วไป ภาควิชาวิทยาศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครสวรรค์ 60000

¹ Major of Information Technology, Department of Science, Faculty of Science and Technology, Phranakhon Si Ayutthaya Rajabhat University, 13000 Thailand

² Major of Physics and General Sciences, Department of Science, Faculty of Science and Technology, Nakhon Sawan Rajabhat University, Nakhon Sawan, 60000 Thailand

* Corresponding author: Dusadee Khamboonrueng E-mail: Dusadee.k@nsru.ac.th

บทนำ

สุขภาพดินเป็นปัจจัยหลักต่อปริมาณและคุณภาพของผลผลิตทางการเกษตรซึ่งค่าอุณหภูมิ (Temperature) ค่าความชื้น (Humid) ค่าความเป็นกรด-ด่าง (pH) ค่าความเค็ม (Salinity) ค่าการนำไฟฟ้า (Electrical Conductivity : EC) และค่าธาตุอาหารหลักในดิน (NPK) ได้แก่ ไนโตรเจน (N) ฟอสฟอรัส (P) โพแทสเซียม (K) ค่าเหล่านี้มีความสำคัญต่อการเจริญเติบโตของพืช (กรมวิทย์ฯ, 2561; สุภาวดี หนูสิน, 2564) อย่างไรก็ตามเมื่อสุขภาพดิน ณ ปัจจุบันไม่เหมาะสมกับพืชที่ปลูก ทำให้พืชมีการเจริญเติบโตช้า ใบไหม้ หรือ ใบหนา การรับรู้ข้อมูลสุขภาพดินข้างต้นจึงเป็นสิ่งจำเป็นเพื่อการรักษาสมาดุลในดินของพืชที่ปลูก งานวิจัยได้สังเกตเห็นความสำคัญการนำข้อมูลสุขภาพดิน อย่างไรก็ตามการตรวจสอบสุขภาพดินนั้นสามารถทำได้โดยการนำอุปกรณ์ตรวจสอบสุขภาพดินไปปักลงบนดินที่ปลูกต้นพืชแล้วทำการตรวจวัดค่าต่างๆ ข้างต้น แต่วิธีการนี้ต้องใช้แรงงานคนเป็นหลักหากมีการทำการเกษตรที่มีพื้นที่ขนาดใหญ่ เช่น การทำการเกษตรเป็นรูปแบบฟาร์ม เป็นต้น ทำให้เกิดความล่าช้าและใช้เวลานานกว่าจะตรวจวัดค่าต่างๆ ของพืชในฟาร์มได้ งานวิจัยนี้จึงได้พัฒนาระบบอินเทอร์เน็ตของสรรพสิ่ง (Wang *et al.*, 2020, Yubolsai *et al.*, 2022; ภาณุ เดชทิพย์อักษร และคณะ, 2022) เพื่อเก็บและรายงานข้อมูลสุขภาพดินที่ทดลองปลูกกับต้นถั่วเหลือง โดยได้จำลองสภาพแวดล้อมออกเป็น 4 ประเภทคือ สภาพแวดล้อมที่หนึ่ง รดน้ำและใส่ปุ๋ยเคมี สภาพแวดล้อมที่สองรดน้ำและใส่ปุ๋ยอินทรีย์ สภาพแวดล้อมที่สามรดน้ำเพียงอย่างเดียว และสภาพแวดล้อมที่สี่ไม่รดน้ำและไม่ใส่ปุ๋ย อย่างไรก็ตามระบบอินเทอร์เน็ตของสรรพสิ่งสามารถเก็บและรายงานผลข้อมูลสุขภาพดินที่เป็นค่าข้อมูลดิบบนสมาร์ตโฟนได้เพียงอย่างเดียว อีกทั้งข้อมูลที่ได้จากเซนเซอร์บางครั้งมีความคลาดเคลื่อนส่งผลให้ค่าที่ได้บางครั้งไม่ถูกต้อง จากปัญหาดังกล่าวข้างต้น ดังนั้นงานวิจัยนี้จะทำการทำนายสุขภาพดินด้วยข้อมูลมัลติพีเจอร์ เพื่อลดความผิดพลาดของข้อมูลสุขภาพที่ได้จากเซนเซอร์ด้วยเทคนิคการเรียนรู้ของเครื่องและปัญญาประดิษฐ์ ซึ่งปัจจุบันได้รับความนิยมอย่างกว้างขวางโดยเฉพาะอัลกอริทึมโครงข่ายประสาทเทียมถูกนำมาประยุกต์ใช้กับงานด้านต่างๆ เช่นอุตสาหกรรมการพยากรณ์ และการท่องเที่ยว เป็นต้น (Liu *et al.*, 2016; Phyo and Jeenanunta., 2019) ดังนั้นงานวิจัยนี้นำข้อมูลสุขภาพดินมาสร้างโมเดลทำนายสุขภาพดิน ณ ปัจจุบัน โดยใช้โครงข่ายประสาทเทียมเรียนรู้เชิงลึก (Deep Neural Network : DNN) (Hamena and Meshoul., 2018) ทำให้สามารถปรับปรุงดินให้เหมาะสมกับพืชที่ปลูกได้อย่างเหมาะสมทันทั่วทั้ง

การทดลอง

1. อุปกรณ์การทดลอง

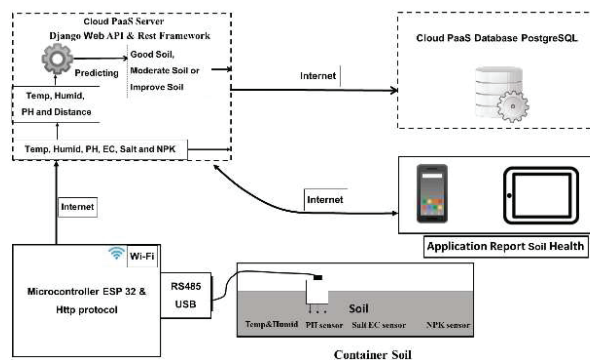


Figure 1 System architecture of the soil health

จาก Figure 1 สถาปัตยกรรมระบบอินเทอร์เน็ตของสรรพสิ่งสำหรับเก็บและรายงานข้อมูลสุขภาพดินสามารถแบ่งออกเป็น 4 องค์ประกอบด้วยกันดังต่อไปนี้

องค์ประกอบที่หนึ่งประกอบด้วยไมโครคอนโทรลเลอร์ (Microcontroller) เซนเซอร์ตรวจจับสุขภาพดิน (Soil Sensors) และ RS 485 สำหรับรับส่งข้อมูลระหว่างกัน มีหน้าที่ตรวจจับและส่งข้อมูลสุขภาพดิน โดยไมโครคอนโทรลเลอร์จะส่งข้อมูลที่ได้จากเซนเซอร์ไปยังเว็บแอปพลิเคชันโดยใช้โปรโตคอลเอชทีทีพี (HTTP) ผ่านเครือข่ายไวไฟ (พินิตา พงษ์ไพบูลย์ และ คณะ, 2020) และก่อนนำเซนเซอร์ไปใช้งานได้ทำการสอบเทียบผลการวัดคุณภาพดินที่ได้จากเซนเซอร์กับกรมพัฒนาที่ดิน

องค์ประกอบที่สองคอมพิวเตอร์แม่ข่ายงานวิจัยนี้ได้ประยุกต์ใช้เทคโนโลยีคลาวด์แบบแพลตฟอร์มแอสอะเซอร์วิส (Platform as a Service : PaaS) (ประจวบ เพิ่มสุวรรณ และ มณฑล นุ่นละอง, 2556) เพื่อใช้สำหรับการติดตั้งเว็บส่วนต่อประสานโปรแกรมประยุกต์ (Application Program Interface : API) (ชูพันธุ์ รัตนโกคา และคณะ, 2565) ซึ่งถูกพัฒนาด้วยภาษาไพทอนจิงโกเว็บเฟรมเวิร์ค (Django Web Framework) ทำหน้าที่เป็นตัวกลางในการรับและส่งข้อมูลระหว่างไมโครคอนโทรลเลอร์ฐานข้อมูล และโมบายแอปพลิเคชัน นอกจากนี้ยังทำหน้าที่ในการสกัดข้อมูลค่าระยะห่างระหว่างข้อมูล (Distance) เพื่อนำข้อมูลไปใช้สำหรับทำนายข้อมูลสุขภาพดิน

องค์ประกอบที่สามฐานข้อมูล (Database) ที่ถูกติดตั้งไว้บนคลาวด์รูปแบบ PaaS เช่นเดียวกันทำหน้าที่จัดเก็บข้อมูลดิบและผลการทำนายสุขภาพดิน

องค์ประกอบที่สี่โมบายแอปพลิเคชันทำหน้าที่รายงานข้อมูลดิบและผลการทำนายสุขภาพดิน ณ ปัจจุบัน

2. การเก็บข้อมูล

งานวิจัยได้ออกแบบวิธีการเก็บตัวอย่างข้อมูลสุขภาพดิน ที่ปลูกถั่วเหลืองเพื่อนำข้อมูลมาใช้ในการสร้างโมเดลทำนายข้อมูลสุขภาพดิน ซึ่งข้อมูลที่เก็บประกอบด้วย ข้อมูลค่าอุณหภูมิ ค่าความชื้น ค่าความเป็นกรด-ด่าง ค่าการนำไฟฟ้า ค่าความเค็ม ค่าไนโตรเจน ค่าฟอสฟอรัส และค่าโพแทสเซียม งานวิจัยได้เตรียมดินใช้สำหรับปลูกต้นถั่วเหลืองทั้งหมด 12 กระถาง โดยขนาดของกระถาง 17 นิ้วและแบ่งสภาพแวดล้อมการดูแลต้นถั่วเหลืองออกเป็น 4 สภาพแวดล้อมๆ ละ 3 กระถาง คือ สภาพแวดล้อมที่หนึ่งรดน้ำและใส่ปุ๋ยเคมี สภาพแวดล้อมที่สองรดน้ำและใส่ปุ๋ยอินทรีย์ สภาพแวดล้อมที่สามรดน้ำเพียงอย่างเดียว สภาพแวดล้อมที่สี่ไม่รดน้ำและไม่ใส่ปุ๋ย ทุกสภาพแวดล้อมจะเป็นสภาพแวดล้อมแบบเปิด ช่วงเดือนแรกของการปลูกต้นถั่วเหลืองทุกกระถางจะได้รับการดูแลเหมือนกันเพื่อให้ต้นถั่วเหลืองมีการเจริญเติบโตใกล้เคียงกันมากที่สุด จากนั้นเมื่อต้นถั่วเหลืองมีอายุครบ 1 เดือนได้ทำการปรับสภาพแวดล้อมของต้นถั่วเหลืองทั้งสี่สภาพแวดล้อมดังที่ได้กล่าวข้างต้น และนำระบบอินเทอร์ของสรรพสิ่งมาใช้เก็บข้อมูลสุขภาพดินวันละหนึ่งครั้งเป็นระยะเวลา 2 เดือนทำให้ได้ข้อมูลทั้งหมด 475 เดต้าพอยต์ (ธวัชชัย พรหมรัตน์ และ ดุษฎี คำบุญเรือง, 2566)

3. การสกัดคุณลักษณะข้อมูลสุขภาพดิน

การสกัดคุณลักษณะข้อมูลงานวิจัยนี้ได้นำข้อมูลสุขภาพดินมาตรวจสอบค่าสัมประสิทธิ์สหสัมพันธ์ (Pearson Correlation)(Chanrashekar and Sahin, 2013) ดังสมการที่ (1) โดยพิจารณาหาค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่าเท่ากับ 1 มาสกัดฟีเจอร์ด้วยการหาระยะห่างระหว่างข้อมูล (Euclidean Distance) (Sukreep *et al.*, 2015) ดังสมการที่ (2) เพื่อลดความ

ซ้ำซ้อนของข้อมูลในการนำข้อมูลไปจัดกลุ่มและการสร้างโมเดลทำนายสุขภาพดิน

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (1)$$

r ค่าสัมประสิทธิ์สหสัมพันธ์

x_i ค่าตัวแปร x ณ ชุดข้อมูลที่ i

$\bar{x}$ ค่าเฉลี่ยของตัวแปร x

y_i ค่าตัวแปร y ณ ชุดข้อมูลที่ i

$\bar{y}$ ค่าเฉลี่ยของตัวแปร y

$$d = \sqrt{(\Delta E)^2 + (\Delta S)^2 + (\Delta N)^2 + (\Delta P)^2 + (\Delta K)^2} \quad (2)$$

$$\Delta E = E_i - E_{i-1}$$

$$\Delta S = S_i - S_{i-1}$$

$$\Delta N = N_i - N_{i-1}$$

$$\Delta P = P_i - P_{i-1}$$

$$\Delta K = K_i - K_{i-1}$$

จากสมการที่ (2) d เป็นผลรวมของค่าหาระยะห่างของข้อมูล ΔE เป็นค่าการนำไฟฟ้า ΔS เป็นค่าความเค็ม ΔN คือค่าไนโตรเจน ΔP เป็นค่าฟอสฟอรัส และ ΔK เป็นค่าโพแทสเซียม i เป็นค่าปัจจุบัน และ $i-1$ ค่าก่อนหน้า ดังนั้นงานวิจัยนี้ใช้ข้อมูลในการจัดกลุ่มและการสร้างโมเดลทำนายสุขภาพดิน 4 ฟีเจอร์ ประกอบด้วยค่าอุณหภูมิ ค่าความชื้น ค่าความเป็นกรด-ด่าง และ ค่าผลรวมระยะห่างระหว่างข้อมูล ดังตัวอย่างข้อมูลใน Table 1

Table 1 Examples raw data of soil health

Temp	Humidity	PH	Distance
34.0	6.0	8.1	34
35.0	26.0	7.5	35
27.0	32.12	7.2	27

4. การจัดกลุ่ม

สำหรับการจัดกลุ่มข้อมูลงานวิจัยนี้ได้ใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) (วิทยา ปัญญา และคณะ, 2565) แบบไม่กำหนดคำตอบ (Unsupervised Learning) (Allenbrand, 2024) เพื่อจัดกลุ่มข้อมูลที่มีความคล้ายกันให้อยู่ในกลุ่มเดียวกัน อย่างไรก็ตามด้วยความหลากหลายของอัลกอริทึมการจัดกลุ่มจึงได้ศึกษาและเปรียบเทียบเพื่อได้อัลกอริทึมที่มีความเหมาะสมในการจัดกลุ่มข้อมูล

พบว่าอัลกอริทึม K-means เป็นอัลกอริทึมที่มีการนำมาใช้สำหรับการจัดกลุ่มมากที่สุด (Ezugwu *et al.*, 2022; Li *et al.*, 2024) รวมถึงมีความเหมาะสมกับข้อมูลสุขภาพดินที่ได้จากทั้ง 4 สภาพแวดล้อม และได้ใช้วิธีวิเคราะห์จำนวนกลุ่มที่เหมาะสมจากจุดโค้งของกราฟ (Elbow Point) และให้อัลกอริทึม K-mean ระบุกลุ่มให้กับข้อมูลทั้ง 475 เดต้าพอยต์เพื่อนำข้อมูลไปใช้สำหรับสร้างโมเดลทำนายสุขภาพดิน

5. การสร้างโมเดลการทำนายสุขภาพดิน

ในขั้นตอนการสร้างโมเดลสำหรับทำนายข้อมูลสุขภาพดิน หลังจากที่ได้มีการจัดกลุ่มข้อมูลและกำหนดคำตอบให้กับข้อมูลสุขภาพดินทั้ง 475 เดต้าพอยต์เสร็จเรียบร้อยแล้ว จะนำข้อมูลสุขภาพดินมาสอนให้คอมพิวเตอร์เรียนรู้ข้อมูลด้วยอัลกอริทึมโครงข่ายประสาทเชิงลึก เพื่อสร้างโมเดลทำนายสุขภาพดินที่ดีที่สุด และนำไปใช้สำหรับทำนายสุขภาพของดิน ณ ปัจจุบัน โดยได้ทำการทดลองสร้างโมเดลและกำหนดค่าไฮเปอร์พารามิเตอร์ (Hanif *et al.*, 2023) ดัง Table 2

Table 2 Hyperparameter for creating model

Hidden Layer	Neural	Activation Function	Optimizer	Loss Function	Epochs	Cross-validation
2	16, 24, 32	Relu	Adam	sparse_categorical_crossentropy	50	StratifiedKFold 3,5,7
3	16, 24, 32	Relu	Adam	sparse_categorical_crossentropy	50	StratifiedKFold 3,5,7
4	16, 24, 32	Relu	Adam	sparse_categorical_crossentropy	50	StratifiedKFold 3,5,7

ผลการทดลองและอภิปรายผล

1. ระบบอินเทอร์เน็ทของสรรพสิ่ง

หลังจากที่พัฒนาระบบอินเทอร์เน็ทของสรรพสิ่ง ได้นำระบบมาทดลองเก็บและรายงานข้อมูลสุขภาพดินที่ใช้สำหรับการปลูกต้นถั่วเหลือง 12 กระถาง 4 สภาพแวดล้อมดัง Figure 2 และ Figure 3 ในส่วนของแอปพลิเคชันรายงานข้อมูลสุขภาพดินผ่านสมาร์ตโฟน ซึ่งได้ออกแบบหน้าจอตั้ง Figure 4 ภาพทางซ้ายเป็นหน้าจอหลักของแอปพลิเคชัน เมื่อกดปุ่มตรวจสอบสุขภาพดินด้านล่างจะแสดงรายงานค่าอุณหภูมิและความชื้น ซึ่งเป็นของเมนู T&M ถูกกำหนดให้เป็นหน้าจอเริ่มต้น นอกจากนี้ยังมีเมนู PH จะรายงานค่าความเป็นกรด-ด่างของดิน เมนู Salinity รายงานค่าความเค็มและค่าการนำไฟฟ้า และเมนูสูตรทำ NPK รายงานค่าแร่ธาตุหลักในดินคือ ไนโตรเจน ฟอสฟอรัส โพแทสเซียม ซึ่งจากผลการทดลองพบว่าระบบสามารถเก็บและรายงานข้อมูลสุขภาพดินได้ถูกต้อง



Figure 2 Example soybean plants 12 pots
(ที่มา: ธวัชชัย พรหมรัตน์ และ ดุษฎิ คำบุญเรือง, 2566)

ซึ่งมีความลึกของชั้นปกปิด (Hidden Layer) ที่ 2 ชั้น 3 ชั้น และ 4 ชั้น โดยมีการกำหนดจำนวนนิวรัลที่ 16 นิวรัล 24 นิวรัล และ 32 นิวรัล ของแต่ละชั้นปกปิดในส่วนของการกำหนดค่าสูตรที่ใช้สำหรับการปรับค่าน้ำหนักคือ อัดัม สำหรับสูตรที่ใช้ในการคำนวณค่ากำหนดค่าความผิดพลาดคือ sparse_categorical_crossentropy การกำหนดฟังก์ชันกระตุ้น (Activation Function) คือเรคติไฟด์ไลน์เนี่ยยูนิต (Rectified Linear Units, ReLu) และชั้นเอาต์พุต Softmax ส่วนการกำหนดจำนวนรอบในการสอนข้อมูล 50 รอบ



Figure 3 Collection of the soil health data for soybean
(ที่มา: โยธิน เกตุเมฆ และ ธวัชชัย พรหมรัตน์, 2567)

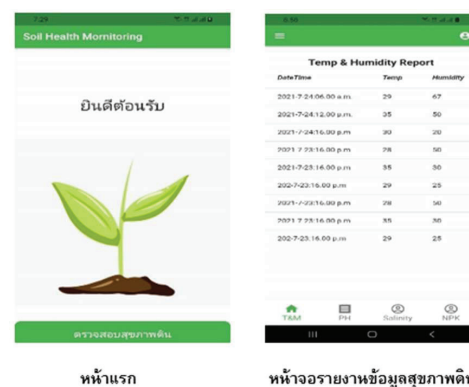


Figure 4 Screen of the mobile application



Figure 5 Characteristic soybean plant of each environment

จาก Figure 5 ลักษณะของต้นถั่วเหลืองแต่ละสภาพแวดล้อม ต้นถั่วเหลืองสภาพแวดล้อมที่หนึ่งรดน้ำและใส่ปุ๋ยเคมี และสภาพแวดล้อมที่สองรดน้ำและใส่ปุ๋ยอินทรีย์ มีลักษณะการเจริญเติบโตใกล้เคียงกันทั้งความสมบูรณ์ของลำต้นและใบ สภาพแวดล้อมที่สามรดน้ำเพียงอย่างเดียวจะมีลักษณะใบเหี่ยวเป็นบางใบและการเจริญเติบโตช้ากว่าสภาพแวดล้อมที่หนึ่งรดน้ำและใส่ปุ๋ยเคมี และสภาพแวดล้อมที่สองรดน้ำและใส่ปุ๋ยอินทรีย์ สภาพแวดล้อมที่สี่ไม่รดน้ำและไม่ใส่ปุ๋ย ต้นถั่วเหลืองลักษณะแตกต่างจากทั้ง 3 สภาพแวดล้อมก่อนหน้านี้อย่างชัดเจนกล่าวคือเมื่อเวลาผ่านไปต้นถั่วเหลืองจะค่อยๆ เหี่ยวและยืนต้นตาย

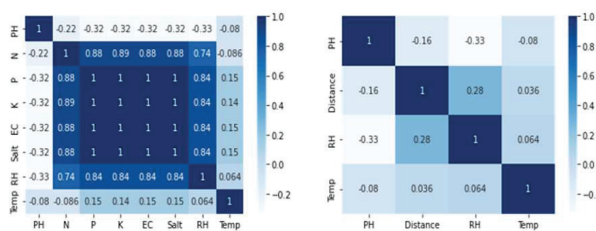


Figure 6 Value of the Correlation coefficient

2. การสกัดคุณลักษณะข้อมูลสุขภาพดิน

จาก Figure 6 ทางด้านซ้ายแสดงให้เห็นว่ามีค่าสัมประสิทธิ์สหสัมพันธ์ก่อนการสกัดคุณลักษณะของข้อมูล มีค่าเท่ากับ 1 อยู่ 5 ค่าคือ ค่าการนำไฟฟ้า ไนโตรเจน ฟอสฟอรัส โพแทสเซียม และความเค็ม ซึ่งค่าสัมประสิทธิ์สหสัมพันธ์มีค่าเท่ากับ 1 เมื่อนำข้อมูลมาพิจารณาทั้ง 475 เดต้าพอยต์แล้ว ข้อมูลที่ได้ไม่มีฟีเจอร์ใดที่มีความโดดเด่นหรือสามารถระบุความแตกต่างของข้อมูลได้ชัดเจน ดังนั้นจึงได้นำข้อมูลที่มีค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่าเท่ากับ 1 มาหาค่าระยะห่างระหว่างข้อมูลทำให้ได้ค่าสัมประสิทธิ์สหสัมพันธ์น้อยกว่า 0.5 ทุกค่าดัง Figure 6 ขวา

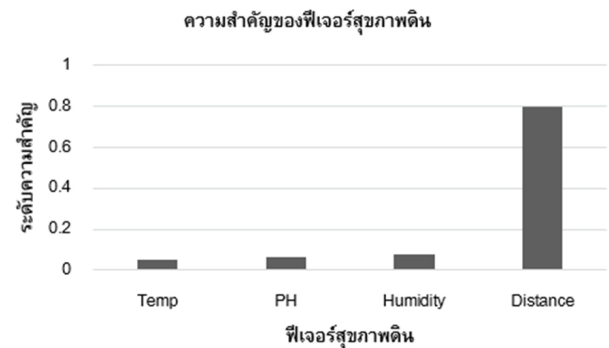


Figure 7 Comparison feature importance of soil health

หลังจากที่ได้ค่าระยะห่างระหว่างข้อมูลแล้วได้นำข้อมูลทั้ง 475 เดต้าพอยต์มาตรวจหาค่าความสำคัญของฟีเจอร์ (Feature importance) โดยใช้เทคนิคการเรียนรู้ของเครื่องด้วยอัลกอริทึม Extra Trees Classifier ทำให้ได้ค่าความสำคัญของฟีเจอร์ ดัง Figure 7 ซึ่งแสดงให้เห็นว่าค่าระยะห่างของข้อมูลมีความสำคัญสูงที่สุดในการทำนายสุขภาพดิน คือ มีค่าอยู่ที่ 0.8 ในขณะที่ค่าอุณหภูมิ ความเป็นกรด-ด่าง และความชื้นนั้นมีค่าน้อยกว่าโดยมีค่าดังนี้ 0.05 0.06 และ 0.07 ตามลำดับ ดังนั้นแสดงให้เห็นว่าค่าระยะห่างระหว่างข้อมูลมีความสำคัญต่อการนำข้อมูลสุขภาพดินไปใช้สำหรับการจัดกลุ่มและสร้างโมเดลทำนายสุขภาพดิน

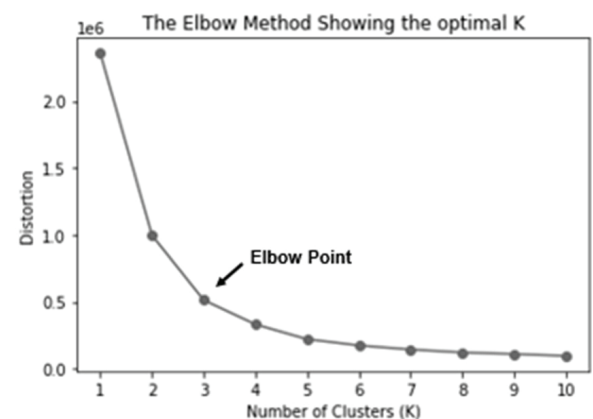


Figure 8 Elbow analysis for optimal k

3. การจัดกลุ่ม

สำหรับการจัดกลุ่มงานวิจัยนี้ได้ทำการทดลองหาจำนวนกลุ่มที่เหมาะสมที่สุดจากทั้งหมด 10 กลุ่มด้วยอัลกอริทึม K-means จากกราฟ Figure 8 แสดงให้เห็นว่าจำนวนกลุ่มที่เหมาะสมกับข้อมูลสุขภาพดินคือ 3 กลุ่ม (ธวัชชัย พรหมรัตน์ และ ดุชนฎิ์ คำบุญเรือง, 2566) ดังนั้นงานวิจัยนี้จึงได้ระบุชื่อกลุ่มของสุขภาพดินที่ได้ทั้ง 3 กลุ่มคือสุขภาพดินเหมาะสม ปานกลาง และต้องปรับปรุง และทำให้สามารถทราบจำนวนข้อมูลสุขภาพดินของแต่ละกลุ่มดังตัวอย่างการกระจายค่าระยะห่างระหว่าง

ข้อมูลของแต่ละเต้าพอยต์ดัง Figure 9 ซึ่งจากการจัดกลุ่มทำให้ได้ข้อมูลกลุ่มของสุขภาพที่ดินที่ต้องปรับปรุง 357 เต้าพอยต์ มีค่าระยะห่างระหว่างข้อมูล 0 - 63 สุขภาพดินปานกลาง 105 เต้าพอยต์มีค่าระยะห่างระหว่างข้อมูล 62 - 215 และพบว่าสุขภาพดินเหมาะสม 13 เต้าพอยต์มีค่าระยะห่างระหว่างข้อมูล 237 - 571

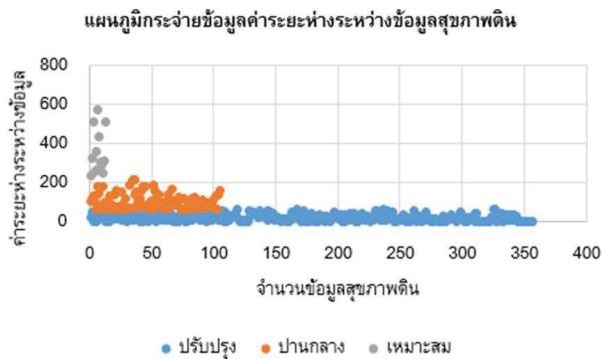


Figure 9 Distribution of the individual distance value all data points

Table 3 Result of confusion matrix from training data set using DNN

Soil health	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
improve	97	100	99	98
medium	100	100	100	
good	100	90	95	

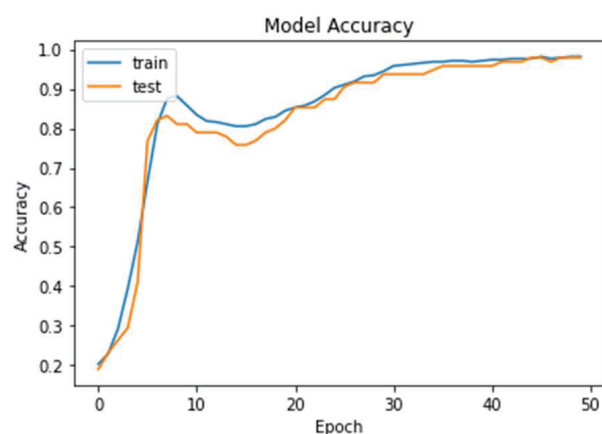


Figure 10 Interpretation of accuracy using DNN

กราฟสำหรับเปรียบเทียบความถูกต้องการเลิร์นและทดสอบข้อมูลสุขภาพดินจากภาพที่ 10 แสดงให้เห็นว่าสิ้นน้ำเงินเส้นโค้งของการเลิร์นและสีส้มเส้นโค้งการทดสอบนั้นอยู่ในระดับที่ใกล้เคียงกันแสดงให้เห็นว่าอัลกอริทึมสามารถเลิร์นข้อมูลได้ดีซึ่งขณะที่กราฟเปรียบเทียบค่าความผิดพลาดของโมเดลจาก Figure 11 มีโอเวอร์ฟิตติงเพียงเล็กน้อยเนื่องจากเส้นโค้งสีน้ำเงินอยู่ต่ำกว่าเส้นสีส้ม อย่างไรก็ตามหากมีการนำระบบไปใช้กับการเกษตรที่เป็นแบบฟาร์ม จะทำให้มีจำนวน

4. โมเดลทำนายข้อมูลสุขภาพดิน

สำหรับการสร้างโมเดลทำนายสุขภาพดินของต้นถั่วเหลือง ได้ทำการทดลองตามแผนการทดลองจาก Table 1 ทำให้ได้โมเดลที่มีถูกต้องมากที่สุด 98 เปอร์เซนต์ดังแสดงใน Table 3 และพบว่าการกำหนดค่าไฮเปอร์พารามิเตอร์ความลึกของชั้นปกปิดจำนวน 3 ชั้นและจำนวนนิวรัลเน็ตเวิร์ค 32 นิวรัล ในส่วนของการแบ่งข้อมูลสำหรับสอนและทดสอบข้อมูลสุขภาพดินแบบ StratifiedKfold เพื่อลดการเกิดโอเวอร์ฟิตติงที่เกิดจากจำนวนข้อมูลแต่ละกลุ่มไม่เท่ากัน จากการทดลองพบว่าจำนวน k เท่ากับ 5

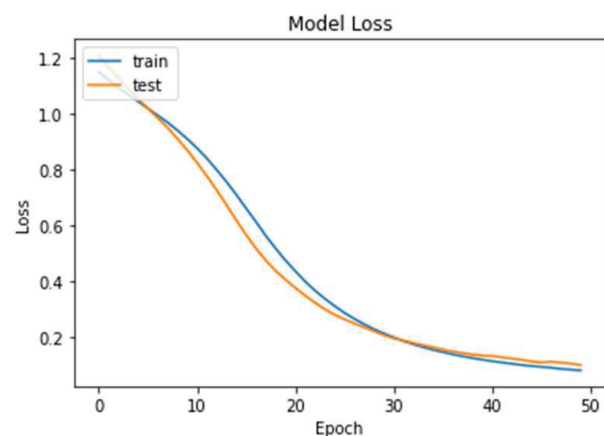


Figure 11 Interpretation of loss value using DNN

ข้อมูลสุขภาพดินเพิ่มมากขึ้นอย่างต่อเนื่องและสามารถนำไปปรับปรุงโมเดล ทำให้โมเดลมีประสิทธิภาพมากขึ้นรวมถึงสามารถลดค่าโอเวอร์ฟิตติงได้

สรุปผลการทดลองและข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างโมเดลทำนายสุขภาพดิน ณ ปัจจุบัน โดยได้ประยุกต์ใช้เทคนิคทางด้านปัญญาประดิษฐ์ อินเทอร์เน็ตของสรรพสิ่ง คลาวด์ และ แอปพลิเคชัน

บนสมาร์ตโฟน ได้นำพัฒนาระบบอินเทอร์เนตของสรรพสิ่งเพื่อเก็บข้อมูลสุขภาพดินจากทั้ง 4 สภาพแวดล้อมเป็นระยะเวลา 2 เดือนทำให้ได้ข้อมูลทั้งหมด 475 เดต้าพอยต์ จากสภาพแวดล้อมทั้ง 4 ดังนี้ สภาพแวดล้อมที่หนึ่งรดน้ำและใส่ปุ๋ยเคมี สภาพแวดล้อมที่สองรดน้ำและใส่ปุ๋ยอินทรีย์ สภาพแวดล้อมที่สามรดน้ำเพียงอย่างเดียว สภาพแวดล้อมที่สี่ไม่รดน้ำและไม่ใส่ปุ๋ย ซึ่งค่าสุขภาพดินที่เก็บคือ ค่าอุณหภูมิ (Temperature) ค่าความชื้น (Humid) ค่าความเป็นกรด-ด่าง (pH) ค่าความเค็ม (Salinity) ค่าการนำไฟฟ้า (Electrical Conductivity : EC) และค่าธาตุอาหารหลักในดิน (NPK) ได้แก่ ไนโตรเจน (N) ฟอสฟอรัส (P) โพแทสเซียม (K) นอกจากนี้ได้พัฒนาแอปพลิเคชันบนสมาร์ตโฟนเพื่อรายงานสุขภาพดินค่าต่างๆ ที่กล่าวมาข้างต้นทำให้สามารถรับรู้สุขภาพดิน ณ ปัจจุบัน ซึ่งเป็นข้อมูลดิบที่ได้จากเซอร์ แต่ข้อมูลอาจมีความคลาดเคลื่อนของค่าที่ได้จากเซนเซอร์ จึงนำไปสู่การทำนายสุขภาพดินของต้นถั่วเหลือง โดยได้ทดลองสกัดคุณลักษณะข้อมูลโดยใช้ระยะทางแบบยุคลิดเพื่อหาพีเจอร์ที่มีความสำคัญกับการทำนายสุขภาพดินของต้นถั่วเหลืองทำให้พบว่า พีเจอร์ที่ระยะห่างของข้อมูลมีความสำคัญมากที่สุดรองลงมาคือ อุณหภูมิ ความเป็นกรด-ด่าง และความชื้น ตามลำดับ จากนั้นได้นำข้อมูลทั้ง 4 ค่ามาจัดกลุ่มโดยใช้อัลกอริทึม k-means ทำให้ได้จำนวนกลุ่มที่เหมาะสมที่สุดคือ 3 กลุ่ม คือสุขภาพดิน เหมาะสม ปานกลาง และปรับปรุง และกำหนดเลเบลให้กับข้อมูลแต่ละเดต้าพอยต์เพื่อนำข้อมูลมาสร้างโมเดลทำนายสุขภาพดินโดยใช้อัลกอริทึมโครงข่ายประสาทเทียมเชิงลึกจากการทดลองสร้างโมเดลตามค่าไฮเปอร์พารามิเตอร์จาก Table 2 ทำให้ได้ผลความถูกต้อง 98 เปอร์เซ็นต์ ซึ่งสามารถนำไปใช้สำหรับทำนายสุขภาพดินและหากสุขภาพดิน และหากพบว่าสภาพดิน ณ ปัจจุบัน เป็นดินที่ไม่เหมาะสมกับพืชที่ปลูกจะทำให้สามารถปรับสภาพหรือบำรุงดินให้กับพืชได้อย่างทันทั่วทั้ง

ข้อเสนอแนะ

นอกจากข้อมูลสุขภาพดินที่เป็นปัจจัยสำคัญแล้วยังมีสภาพสภาพอากาศที่เป็นปัจจัยที่มีผลต่อการเจริญเติบโตของพืช หากจะให้พืชเจริญเติบโตได้เต็มที่ควรคำนึงถึงสภาพอากาศด้วย ในส่วนของการนำระบบไปใช้งานจริงในสมาร์ตฟาร์มที่มีขนาดใหญ่ จะต้องปรับสเกลของระบบให้มีความเหมาะสมกับขนาดของฟาร์ม นอกจากนี้ในอนาคตหากมีการนำระบบไปใช้กับฟาร์มและมีการเก็บข้อมูลอย่างต่อเนื่องทำให้ได้ข้อมูลเพิ่มมากขึ้นจะสามารถข้อมูลมาปรับปรุงให้โมเดลมีประสิทธิภาพเพิ่มขึ้น

เอกสารอ้างอิง

- กรมอุตุฯ. (2561). ระบบสังเกตการณ์ด้วยเสียงบนเทคโนโลยีสรรพสิ่งเพื่อประยุกต์ควบคุมมอเตอร์ในงานด้านเกษตรกรรม. มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ.
- ชนิดา ยุกตไสย์, วิสวะ กุลนะ, & คมกฤษณ์ ชูเรือง. (2022). ระบบการวัดและบันทึกค่า pH ของดิน โดยใช้ Internet of Things (IoT). วารสารวิจัย และพัฒนานวัตกรรมภาคลุ่มน้ำโขง, 1(1), 1-7.
- รัชชัย พรหมรัตน์ และ ดุษฎี คำบุญเรือง. (2566). การจัดกลุ่มข้อมูลคุณภาพดินที่ได้จากการปลูกต้นถั่วเหลืองโดยใช้เทคนิคการเรียนรู้ของเครื่อง. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ "ราชภัฏกรุงเก่า" ครั้งที่ 6 (2023 6th EEAAT & ARU CON.) ภาคบรรยาย (Oral Presentation) วันที่ 21 - 22 ธันวาคม พ.ศ. 2566 ณ มหาวิทยาลัยราชภัฏพระนครศรีอยุธยา, 751-758. <https://www.aru.ac.th/arucon/2023/EEAAT-ARUCON2023Oral-ProceedingsV1.pdf>
- ประจวบ เพิ่มสุวรรณ และ มณฑล นุ่นละออง. (2013). ปัจจัยที่มีผลต่อการยอมรับระบบคลาวด์คอมพิวเตอร์ของผู้ที่งานด้านเทคโนโลยีสารสนเทศ. *Rajabhat Chiang Mai Research Journal*, 14(2), 47-60.
- ภาณุ เดชทิพย์อักษร, อรรถพล วิเวก, & อนุพงศ์ ไพโรจน์. (2022). การประยุกต์ใช้งานแอปพลิเคชันปัญญาประดิษฐ์ร่วมกับอินเทอร์เนตของ สรรพสิ่งบนคลาวด์ในอาคารอัจฉริยะ. วารสารวิศวกรรมศาสตร์มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา, 7(1), 52-61.
- โยธิน เกตุเมฆ และ รัชชัย พรหมรัตน์. (2567). การตรวจสอบและรายงานคุณภาพดินของพืชโดยใช้อินเทอร์เนตของสรรพสิ่ง. รายงานสืบเนื่องจากการประชุมวิชาการระดับปริญญาตรีด้านคอมพิวเตอร์ภูมิภาคเอเชียครั้งที่ 12 (Oral Conference Proceeding) วันที่ 21-23 กุมภาพันธ์ 2567 ณ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม, 400-408.
- วิทยา ปัญญา และวณัช รมสายหยุด. (2023). วิธีการสร้างแบบจำลองเชิงทำนายพฤติกรรมการผลิต เงื่อนไขการปล่อยชั่วคราวของศาลจากชุดข้อมูลที่ไม่ สมดุลโดยใช้เทคนิคการเรียนรู้ของเครื่อง. วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม, 42(2), 7-17.
- สุภาวดี หนูสิน. (2021). ผลกระทบจากการใช้ปุ๋ยเคมี ทางภาคเกษตร. วารสารบริหารการพัฒนา นวัตกรรมเชิงบูรณาการ, 1(2), 34 - 43.

- Allenbrand, C. (2024). Supervised and unsupervised learning models for pharmaceutical drug rating and classification using consumer generated reviews. *Healthcare Analytics*, 5, 100288.
- Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1), 16-28.
- Ezugwu, A. E., Ikotun, A. M., Oyelade, O. O., Abualigah, L., Agushaka, J. O., Eke, C. I., & Akinyelu, A. A. (2022). A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence*, 110, 104743.
- Hamena, S., & Meshoul, S. (2018, October). Deep neural network for classification and prediction of oxygen binding proteins. *Proceedings of the 2018 2nd International Conference on Computational Biology and Bioinformatics*, 72-77.
- Hanifi, S., Cammarono, A., & Zare-Behtash, H. (2024). Advanced hyperparameter optimization of deep learning models for wind power prediction. *Renewable Energy*, 221, 119700.
- Li, J., Ma, R., Deng, M., Cao, X., Wang, X., & Wang, X. (2024). A comparative study of clustering algorithms for intermittent heating demand considering time series. *Applied Energy*, 353, 122046.
- Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, 234, 11-26.
- Phyo, P. P., & Jeenanunta, C. (2019). Electricity load forecasting using a deep neural network. *Engineering and Applied Science Research*, 46(1), 10-17.
- Pongpaibool, P., Niruntasukrat, A., & Meesublak, K. (2020). แพลตฟอร์มอินเทอร์เน็ตของสรรพสิ่ง ฐานรากสำคัญสู่การสร้างสรรค์นวัตกรรม ดิจิทัล. *NBTC Journal*, 4(4), 269-287.
- Rattanapoka, C., Jannet, P., & Moonpanyo, T. (2022). ระบบบริการการจำแนกเพศและประเมินอายุ บุคคลจากภาพใบหน้าด้วยวิธีการเรียนรู้เชิงลึก. *The Journal of Industrial Technology*, 18(2), 164-179. <https://ojs.kmutnb.ac.th/index.php/jointtech/article/view/5820>
- Sukreep, S., Mongkolnam, P., & Nukoolkit, C. (2015). Detect the daily activities and in-house locations using smartphone. In *Recent Advances in Information and Communication Technology 2015. Proceedings of the 11th International Conference on Computing and Information Technology (IC2IT)*. Springer International Publishing, 215-225. https://link.springer.com/chapter/10.1007/978-3-319-19024-2_22.
- Wang, J., Chang, L., Aggarwal, S., Abari, O., & Keshav, S. (2020, June). Soil moisture sensing with commodity RFID systems. In *Proceedings of the 18th International Conference on Mobile Systems, applications, and Services*, 273-285. <https://web.cs.ucla.edu/~omid/Papers/Mobisys20.pdf>.