**บทคัดย่อ**

รายงานฉบับนี้ได้ทำการศึกษาและวิเคราะห์ Prediction water data เพื่อศึกษาระดับน้ำและวิเคราะห์ระดับน้ำเพื่อให้นักศึกษานำข้อมูลไปต่อยอด นักศึกษาสาขาวิชาการวิเคราะห์และจัดการข้อมูลขนาดใหญ่ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

**กิตติกรรมประกาศ**

รายงาน Prediction water data สำเร็จได้ด้วยดีเนื่องจากได้รับความช่วยเหลือรายงานจาก อ.พิเชฐ คุณากรวงศ์ และ อ.ปองพล นิลพฤกษ์ ที่ได้ให้ข้อมูลแนะนำชี้แนะแนวทางในการศึกษาค้นคว้า ขั้นตอนวิธีการ จัดทำรายงานการจนสำเร็จตามวัตถุประสงค์ที่กำหนดไว้

คณะผู้จัดทำขอขอบพระคุณท่านที่ให้ความช่วยเหลือในเรื่องต่างๆและหวังเป็นอย่างยิ่งว่ารายงาน เรื่อง Prediction water data เรื่องนี้ จะเกิดประโยชน์ต่อผู้สนใจในการศึกษาต่อไป

คณะผู้จัดทำ

|  |  |
| --- | --- |
|  | หน้า |
| บทคัดย่อ | 1 |
| กิตติกรรมประกาศ | 2 |
| สารบัญ | 3 |
| บทที่ 1 บทนำ | 4 |
| 1.1 ที่มาและความสำคัญ | 4 |
| 1.2 วัตถุประสงค์ของรายงาน | 4 |
| 1.3 ขอบเขตของรายงาน | 4 |
| 1.4 ผลที่คาดว่าจะได้รับ | 5 |
| 1.5 ระยะเวลาการดำเนินงาน | 5 |
| บทที่ 2 เอกสารและทฤษฎีที่เกี่ยวข้อง | 6 |
| 2.1 Machine Learning | 6 |
| 2.2 Artificial Intelligence (AI) | 7 |
| 2.3 Orange Data Mining | 8-9 |
| 2.4 Excel | 10 |
| 2.5 การวัดระดับน้ำ | 11 |
| 2.6 ภาษา R | 12 |
| บทที่ 3 วิธีการดำเนินงาน | 13 |
| 3.1 แผนการดำเนินงาน | 13 |
| 3.2 ผู้รับผิดชอบโครงการ | 14 |
| บทที่ 4 ผลการดำเนินงาน | 15 |
| 4.1 วิธีการวัดผลทำนายระดับน้ำ | 15-17 |
| 4.2 วิธีการวัดผลทำนายค่าน้ำ | 18-19 |
| บทที่ 5 สรุปอภิปรายผลและข้อเสนอแนะ | 20 |
| 5.1 สรุปการประเมินผลการดำเนินโครงการ | 20 |
| 5.2 ปัญหาและอุปสรรค | 20 |

**สารบัญ**

**บทที่ 1**

**บทนำ**

**1.1 ที่มาและความสำคัญ** เป็นข้อมูลระดับน้ำของแม่น้ำในระบบชลประทานเพื่อให้นักศึกษาใช้กระบวนการการเรียนรู้ของเครื่อง (Machine Learning) ในการทำนายค่าระดับน้ำในช่วงเวลาต่าง ๆ

**1.2 วัตถุประสงค์ของรายงาน**

1**.**2.1 พัฒนาระบบสำหรับคาดการณ์ข้อมูลน้ำ

1.2.2 เพื่อคาดการณ์หรือทำนายระดับน้ำที่จะมาในอนาคต

1.2.3 เพื่อตอบสนองความต้องการของผู้ใช้งาน

1.2.4 เพื่อศึกษาความพึงพอใจของผู้ใช้งาน

**1.3 ขอบเขตการดำเนินงาน**

1.3.1 ระดับน้ำแม่น้ำในระบบชลประทาน

1.3.2 ขจัดข้อมูลผิดพลาด จัดรูปแบบข้อมูล

1.3.3 พัฒนาโมเดลการคาดการณ์ข้อมูลน้ำโดยใช้เทคนิคต่างๆ เช่น โมเดลทางสถิติ โมเดล Machine Learning

**1.4 ผลที่คาดว่าจะได้รับ**

1.4.1 โมเดลที่สามารถคาดการณ์ข้อมูลน้ำ เช่น ระดับน้ำ ปริมาณน้ำไหล คุณภาพน้ำ

1.4.2 โมเดลการคาดการณ์ข้อมูลน้ำที่มีความแม่นยำ

1.4.3 โมเดลที่มีความแม่นยำตามมาตรฐานที่กำหนด

1.4.4 รายงานสรุปผลการวิจัย

**1.5 ระยะเวลาการดำเนินงาน**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ลำดับ** | **รายการ** | 27-28  ก.พ.  2567 | 1-3  มี.ค.  2567 | 3-10  มี.ค.  2567 | 10-14  มี.ค.  2567 | 14-18  มี.ค.  2567 | 18-22  มี.ค.  2567 | 22-30  มี.ค.  2567 |
| **1** | ศึกษาและเก็บรวบรวมข้อมูล |  |  |  |  |  |  |  |
| **2** | วิเคราะห์และแก้ไขข้อมูล |  |  |  |  |  |  |  |
| **3** | ศึกษาโมเดลที่จะใช้ |  |  |  |  |  |  |  |
| **4** | ทดสอบและแก้ไขข้อผิดพลาด |  |  |  |  |  |  |  |
| **5** | ทดสอบโมเดลที่จะใช้ |  |  |  |  |  |  |  |
| **6** | นำมาแก้ไขและปรับปรุง |  |  |  |  |  |  |  |
| **7** | สรุปผลการทำงาน |  |  |  |  |  |  |  |

**บทที่ 2**

**เอกสารและทฤษฎีที่เกี่ยวข้อง**

**2.1 Machine Learning**

2.1.1Machine Learning คือ การทำให้คอมพิวเตอร์สามารถเรียนรู้ได้ด้วยตัวเอง โดยอาศัยข้อมูลคล้ายกับการเรียนรู้ของมนุษย์ Machine Learning จะเรียนรู้จากข้อมูลเหมือนกับที่เราเรียนรู้จากประสบการณ์ใช้ข้อมูลเพื่อสร้างโมเดล โมเดลคือสูตรหรือกฎที่คอมพิวเตอร์เรียนรู้ได้จากข้อมูลทำนายผลลัพธ์ของข้อมูลใหม่เมื่อเจอข้อมูลใหม่ โมเดลสามารถใช้ทำนายผลลัพธ์ได้ตัวอย่างการประยุกต์ใช้ Machine Learning: ระบบแนะนำสินค้า : เว็บไซต์ e-commerce ใช้ Machine Learning เพื่อแนะนำสินค้าที่ตรงกับความสนใจของลูกค้าการกรองอีเมลสแปม: ระบบอีเมลใช้ Machine Learning เพื่อกรองอีเมลที่ไม่ต้องการออกไปรถยนต์ไร้คนขับ: รถยนต์ไร้คนขับใช้ Machine Learning เพื่อรับรู้สภาพแวดล้อมและตัดสินใจในการขับ

2.1.2 Machine Learning มีกี่ประเภท? โดยทั่วไป Machine Learning แบ่งเป็น 3 ประเภทหลักๆการเรียนรู้แบบมีผู้สอน (Supervised Learning) ฝึกโมเดลจากข้อมูลที่มีคำตอบ เช่น จำแนกประเภทอีเมล (สแปมหรือไม่สแปม)

การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) หาโครงสร้างในข้อมูลโดยไม่ต้องใช้คำตอบ เช่น การจัดกลุ่มลูกค้าตามพฤติกรรมการซื้อสินค้า

การเรียนรู้แบบเสริมแรง (Reinforcement Learning) ฝึกโมเดลผ่านการลองผิดลองถูก เช่น การฝึกปัญญาประดิษฐ์ให้เล่นเกมความแตกต่างระหว่าง

2.1.3 Machine Learning กับ AI Machine Learning เป็นสาขาหนึ่งของปัญญาประดิษฐ์ (AI) ที่มุ่งเน้นไปที่การเรียนรู้จากข้อมูล AI เป็นแนวคิดกว้างกว่านั้น ครอบคลุมถึงความสามารถทางปัญญาต่างๆ ของมนุษย์ เช่น การเรียนรู้ การแก้ปัญหา การตัดสินใจ

**2.2 Artificial Intelligence (AI)**

2.2.1ปัญญาประดิษฐ์ (AI) คืออะไร?ปัญญาประดิษฐ์ (AI) หมายถึง ความสามารถทางปัญญาของเครื่องจักร ที่ถูกออกแบบมาให้เลียนแบบความสามารถทางปัญญาของมนุษย์

2.2.2 คุณสมบัติหลักของ (AI) การเรียนรู้ เรียนรู้จากข้อมูลและประสบการณ์การคิดวิเคราะห์ข้อมูลและหาความสัมพันธ์การแก้ปัญหา หาทางออกที่ดีที่สุดสำหรับปัญหาการตัดสินใจ ตัดสินใจอย่างมีเหตุผล การปรับตัวเข้ากับสถานการณ์ใหม่

ตัวอย่างการประยุกต์ใช้ AI ระบบแนะนำสินค้า แนะนำสินค้าที่ตรงกับความสนใจของลูกค้าการจดจำใบหน้า: ระบุบุคคลจากภาพถ่ายหรือวิดีโอการแปลภาษา แปลภาษาโดยอัตโนมัติ การขับขี่อัตโนมัติ ควบคุมรถโดยไม่ต้องมีคนขับการวินิจฉัยโรค วิเคราะห์ภาพถ่ายทางการแพทย์เพื่อวินิจฉัยโรคค

2.2.3 ประเภทของ AI Machine Learning เรียนรู้จากข้อมูลโดยไม่ต้องเขียนโปรแกรม Deep Learning ประเภทของ Machine Learning ที่ใช้โมเดลประสาทเทียม Natural Language Processing (NLP) เข้าใจและประมวลผลภาษาธรรมชาติ Computer Vision รับรู้และวิเคราะห์ภาพความแตกต่างระหว่าง

2.2.4 AI กับ Machine LearningAI แนวคิดกว้างกว่า ครอบคลุมถึงความสามารถทางปัญญาต่างๆMachine Learning สาขาหนึ่งของ AI ที่มุ่งเน้นไปที่การเรียนรู้จากข้อมูล

**2.3 Orange Data Mining**

2.3.1 Orange Data Mining เป็นซอฟต์แวร์โอเพนซอร์สสำหรับการวิเคราะห์ข้อมูล การเรียนรู้ของเครื่อง และการนำเสนอภาพข้อมูล

2.3.2 โมเดลของ Orange Data Mining ไม่ได้สร้างโมเดลของตัวเองโดยตรง แต่เป็นเครื่องมือที่ช่วยสร้างและวิเคราะห์โมเดล Machine Learning จากข้อมูลของผู้ใช้งาน Orange Data Mining มี Widget (ซึ่งเป็นเหมือนบล็อกตัวต่อ) สำหรับรองรับการทำงานกับโมเดล Machine Learning หลากหลายประเภท

* การเรียนรู้แบบมีผู้สอน (Supervised Learning) การจำแนกประเภท
* Classification แยกข้อมูลออกเป็นกลุ่ม เช่น จำแนกประเภทของสแปมอีเมลการถดถอย
* Regression ทำนายค่าต่อเนื่อง เช่น คาดการณ์ราคาบ้านการเรียนรู้แบบไม่มีผู้สอน
* Unsupervised Learning การจัดกลุ่มข้อมูล
* Clustering จัดกลุ่มข้อมูลที่มีลักษณะคล้ายคลึงกันการลดมิติ
* Dimensionality Reduction ลดจำนวนมิติของข้อมูลเพื่อการวิเคราะห์ที่ง่ายขึ้น

2.3.3 Orange Data Mining ช่วยอะไรคุณบ้าง

* โหลดข้อมูล โหลดข้อมูลจากไฟล์ หรือฐานข้อมูลต่างๆ
* เตรียมข้อมูล จัดการข้อมูลที่หายไป ปรับรูปแบบข้อมูล แปลงข้อมูล
* สร้างโมเดล เลือก Widget ที่เหมาะสมกับปัญหาของคุณ ปรับแต่งพารามิเตอร์ของโมเดล
* ฝึกโมเดล ฝึกโมเดลจากข้อมูลที่มีอยู่
* ประเมินผลโมเดล ประเมินประสิทธิภาพของโมเดล
* ใช้โมเดลทำนายผลลัพธ์ของข้อมูลใหม

2.3.4 การใช้ Widget สำหรับการสร้างโมเดลใน Orange Data Mining:

* Tree สร้างโมเดลต้นไม้ตัดสินใจ
* Linear Regression สร้างโมเดลการถดถอยเส้นตรง
* KNN สร้างโมเดล k-Nearest Neighbors
* Naive Bayes สร้างโมเดล Naive Bayes
* PCA สร้างโมเดลลดมิติด้วย Principal Component Analysis (PCA)

2.3.5 ข้อดีของการใช้ Orange Data Mining สร้างโมเดล:

* ใช้งานง่าย ใช้ Widget โดยไม่ต้องเขียนโปรแกรม
* รองรับโมเดลหลากหลาย รองรับโมเดล Machine Learning ที่หลากหลาย
* การนำเสนอภาพข้อมูล ช่วยในการสร้างกราฟและแสดงผลลัพธ์

**2.4 Excel**

2.4.1 Excel เป็นโปรแกรม spreadsheet ยอดนิยมจาก Microsoft ที่ใช้จัดการข้อมูล วิเคราะห์ข้อมูล และสร้างกราฟ

2.4.2 โครงสร้างของ Excel เซลล์: หน่วยพื้นฐานที่เก็บข้อมูลใน Excel แต่ละเซลล์มีพิกัดที่ระบุด้วยตัวอักษร (A, B, C ฯลฯ) และตัวเลข (1, 2, 3 ฯลฯ) เช่น A1, B2, C3 แถว: แนวตั้งของเซลล์คอลัมน์: แนวนอนของเซลล์ ชีต: ชุดของเซลล์ คล้ายกับหน้ากระดาษ แต่ละ Workbook ประกอบด้วยหลาย Sheet Workbook: ไฟล์ Excel ที่ประกอบด้วย Sheet หลายตัว

2.4.3 ประเภทของข้อมูล ตัวเลข เช่น 1, 2, 3, 10.5 ,

* ข้อความ ข้อความ เช่น "ชื่อ", "นามสกุล" "ประเทศไทย"วันที่
* ข้อมูลวันที่ เช่น 1/1/2024, 2/3/2023เวลา
* ข้อมูลเวลา เช่น 10:00, 13:30
* ตรรกะ ค่า TRUE หรือ FALSE

2.4.4 สูตรและฟังก์ชันสูตร การคำนวณหรืออ้างอิงข้อมูลในเซลล์อื่น ตัวอย่าง: =A1+B1

ฟังก์ชัน: สูตรสำเร็จรูปที่คำนวณค่าต่างๆ เช่น SUM(), AVERAGE(), COUNTIF()

2.4.5 การอ้างอิงเซลล์การอ้างอิงสัมพันธ์ อ้างอิงเซลล์โดยใช้ตัวอักษรและตัวเลข เช่น A1 B2 การอ้างอิงสัมบูรณ์: อ้างอิงเซลล์โดยใช้ $ ตัวอย่าง: $A$1การตั้งชื่อช่วง ตั้งชื่อให้กับกลุ่มเซลล์เพื่อการอ้างอิงที่ง่ายขึ้น

2.4.6 การจัดรูปแบบข้อมูล ตัวอักษร: เปลี่ยนขนาด สี รูปแบบตัวอักษรการจัดตำแหน่ง ปรับตำแหน่งของข้อมูลในเซลล์เส้นขอบ เพิ่มเส้นขอบให้กับเซลล์การแรเงา: เพิ่มสีพื้นหลังให้กับเซลล์

**2.5 การวัดระดับน้ำ**

2.5.1 การวัดระดับน้ำ หมายถึง การตรวจสอบและบันทึกระดับน้ำในแหล่งน้ำต่างๆ เช่น แม่น้ำ คลอง ทะเลสาบ เขื่อน ฯลฯ

2.5.2 หลักการวัดระดับน้ำ

* แรงดันไฮโดรสถิตวัดแรงดันที่เกิดจากน้ำเพื่อหาความสูงของน้ำ
* ลอยที่ลอยขึ้นหรือจมลงตามระดับน้ำคลื่นเสียง
* คลื่นเสียงที่ใช้วัดระยะทางไปยังผิวน้ำแม่เหล็กไฟฟ้า
* วัดการเปลี่ยนแปลงของสนามแม่เหล็กไฟฟ้าเมื่อระดับน้ำเปลี่ยนแปลง
* เลเซอร์วัดระยะทางไปยังผิวน้ำด้วยแสงเลเซอร์

2.5.3 วิธีการวัดระดับน้ำ

* เครื่องวัดระดับน้ำแบบลอยเชื่อมต่อกับตัวส่งสัญญาณ
* เครื่องวัดระดับน้ำแบบอัลตร้าโซนิค ส่งคลื่นเสียงไปยังผิวน้ำแล้ววัดเวลาสะท้อนกลับ
* เครื่องวัดระดับน้ำแบบแม่เหล็กไฟฟ้า วัดการเปลี่ยนแปลงของสนามแม่เหล็กไฟฟ้า
* เครื่องวัดระดับน้ำแบบเลเซอร์ ยิงแสงเลเซอร์ไปยังผิวน้ำแล้ววัดระยะทาง

2.5.3 ปัจจัยที่มีผลต่อการวัดระดับน้ำ:

* อุณหภูมิอุณหภูมิของน้ำส่งผลต่อความหนาแน่นของน้ำ
* ความเค็มความเค็มของน้ำส่งผลต่อความหนาแน่นของน้ำ
* คลื่นคลื่นในแหล่งน้ำส่งผลต่อความแม่นยำของการวัด
* สิ่งปนเปื้อนในน้ำส่งผลต่อความแม่นยำของการวัด

**2.6 ภาษา R**

2.6.1 โครงสร้างพื้นฐาน วัตถุ (Object) หน่วยพื้นฐานของข้อมูลใน R ประกอบด้วยชนิดข้อมูล (data type) และข้อมูล (data)ชนิดข้อมูล ประเภทของข้อมูล เช่น ตัวเลข (numeric), ข้อความ (character), ตรรกะ(logical), เวกเตอร์ (vector), เมทริกซ์ (matrix), ข้อมูลเชิงกลุ่ม (data frame), รายการ (list)โครงสร้างข้อมูล รูปแบบการจัดเก็บข้อมูล เช่น เวกเตอร์ เมทริกซ์ ข้อมูลเชิงกลุ่ม รายการฟังก์ชัน (Function) ชุดคำสั่งที่ทำงานเฉพาะอย่างตัวแปร (Variable) ตัวแทนสำหรับเก็บค่านิพจน์ (Expression) หน่วยคำนวณ ประกอบด้วยตัวแปร ฟังก์ชัน และสัญลักษณ์

2.6.2 การเขียนโปรแกรม ไวยากรณ์ (Syntax) กฎเกณฑ์ในการเขียนโปรแกรม บรรทัดคำสั่ง (Command line): การป้อนคำสั่งทีละบรรทัด สคริปต์ (Script) ชุดคำสั่งที่เก็บไว้ในไฟล์ แพ็กเกจ(Package): ชุดฟังก์ชันเพิ่มเติม การดีบัก (Debugging) การแก้ไขข้อผิดพลาด

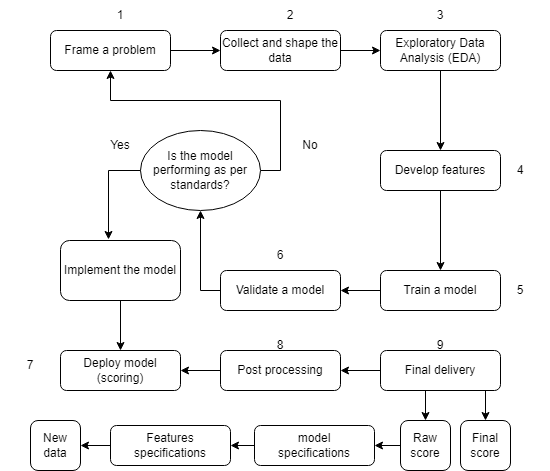
2.6.3 การวิเคราะห์ข้อมูลการอ่านข้อมูล นำเข้าข้อมูลจากไฟล์ แหล่งข้อมูลต่างๆ การจัดการข้อมูล: จัดเตรียมข้อมูลให้พร้อมสำหรับการวิเคราะห์ การวิเคราะห์ทางสถิติ วิเคราะห์ข้อมูลด้วยสถิติต่างๆ การสร้างกราฟ แสดงผลลัพธ์ข้อมูลในรูปแบบกราฟ การเรียนรู้ของเครื่อง พัฒนาระบบสำหรับทำนายผล

2.6.4 ทฤษฎีทางสถิติ การแจกแจงความน่าจะเป็น อธิบายความน่าจะเป็นของผลลัพธ์ การทดสอบสมมติฐาน ตรวจสอบสมมติฐานเกี่ยวกับข้อมูล การถดถอยเชิงเส้น อธิบายความสัมพันธ์ระหว่างตัวแปร การวิเคราะห์ความแปรปรวน อธิบายความแตกต่างระหว่างกลุ่ม การจำแนกประเภท แยกข้อมูลออกเป็นกลุ่ม

**บทที่ 3**

**วิธีการดำเนินงาน**

**3.1 แผนการดำเนินงาน**



**3.2 ผู้รับผิดชอบโครงงาน**

3.2.1 นายไทธานี รัตนพิเชษฐ 116510907029-2

3.2.2 นายธราเทพ ศิริมูน 116510907033-4

3.2.3 นายนครินทร์ ลังประเสริฐ 116510907036-7

**บทที่ 4**

**ผลการดำเนินงาน**

**4.1 วิธีการวัดผลทำนายระดับน้ำ**

Cross validation: วิธีการประเมินประสิทธิภาพโมเดลโดยแบ่งข้อมูลออกเป็นชุดย่อย

Number of folds: จำนวนชุดย่อยที่ใช้ในการประเมิน

Model: โมเดล Machine Learning ที่ใช้ในการทดสอบ

AUC: ค่า AUC (Area Under the ROC Curve) วัดประสิทธิภาพโมเดลในการจำแนก

ประเภท

CA: ค่า CA (Classification Accuracy) วัดความถูกต้องของโมเดลในการจำแนก ประเภท

F1: ค่า F1 Score วัดประสิทธิภาพโมเดลโดยรวม

Prec: ค่า Precision วัดความแม่นยำของโมเดล

Recall: ค่า Recall วัดความครอบคลุมของโมเดล

MCC: ค่า Matthews Correlation Coefficient วัดความสัมพันธ์ระหว่างการคาดการณ์ ของโมเดลกับค่าจริง

รูปภาพประกอบด้วย ข้อความ, ภาพหน้าจอ, ซอฟต์แวร์, ซอฟต์แวร์มัลติมีเดีย

คำอธิบายที่สร้างโดยอัตโนมัติ4.1.1 ใช้การ train model ด้วย Cross validation number of folds : 5 ครั้ง

เห็นได้ว่าการ

tree และ random Forest จะได้ค่าที่แม่นยำเทียบเท่ากัน

KNN และ Neural Network จะdiffกันค่อนข้างมาก

4.1.2 ใช้การ train model ด้วย Cross validation number of folds : 10 ครั้ง

รูปภาพประกอบด้วย ภาพหน้าจอ, ข้อความ, ซอฟต์แวร์, ซอฟต์แวร์มัลติมีเดีย

คำอธิบายที่สร้างโดยอัตโนมัติ

จากข้อมูลในภาพ มีโอกาสเกิด OverFitting กับโมเดล Neural Network โมเดล Neural Network มีคะแนน AUC, CA, F1, Precision, Recall และ MCC บนชุดข้อมูล Train สูงมาก (1.000)

โมเดล Tree มีคะแนน AUC, CA, F1, Precision, Recall และ MCC ใกล้เคียงกับ Random Forest (1) แต่ Random Forest (1) มีประสิทธิภาพการทำงานที่ดีกว่าเล็กน้อย

**4.2 วิธีการวัดผลทำนายค่าน้ำ**

รูปภาพประกอบด้วย ข้อความ, ภาพหน้าจอ, แสดง, ซอฟต์แวร์

คำอธิบายที่สร้างโดยอัตโนมัติ4.2.1 ใช้การ train model ด้วย Cross validation number of folds : 5 ครั้ง

เลือก Random Forest จะพบว่า การใช้ model Random Forest จะมี

มีค่าที่ Diff กันไม่มากเมื่อเทียบกับค่า kNN

Model Tree และ Linear Regression มีโอกาสที่เกิด Over Fiting มากที่สุด

โมเดล Linear Regression มีคะแนน R2 บนชุดข้อมูล Train สูงมาก (0.758)

4.2.2 ใช้การ train model ด้วย Cross validation number of folds : 10 ครั้ง เลือก Random Forest

รูปภาพประกอบด้วย ข้อความ, ภาพหน้าจอ, ซอฟต์แวร์, แสดง

คำอธิบายที่สร้างโดยอัตโนมัติ

โมเดล Random Forest มีประสิทธิภาพการทำงานที่ดีกว่าโมเดลอื่น ๆ

โมเดล Knn จะไม่ทนต่อค่า outlier และประสิทธิการทำงานปานกลางเมื่อเทียบกับ Random Forest

**บทที่ 5**

**สรุปอภิปรายผลและข้อเสนอแนะ**

**5.1 สรุปผลการรายงาน**

5.1.1 กลุ่มของเราเลือกใช้ Model Random Forest ด้วยเหตุผลคือมีคะแนน AUC, CA, F1, Precision, Recall, และ MCC สูงสุดสำหรับการทำนายระดับน้ำ และค่า MSE, RMSE, MAE, MAPE, และ R2 ดีที่สุดสำหรับการทำนายค่า ทั้งหมดถูกเทรนด้วยวิธี cross validation โดยใช้ Folds = 10 เพราะ Folds = 20 มีค่าที่แตกต่างน้อย

**5.2 ปัญหาและอุปสรรค**

5.2.1 ไม่ทราบช่วงเวลาของระดับที่แน่ชัด

5.2.1 โปรแกรม Orange canvas ไม่เอื้ออำนวยต่อการทำ Data Cleaning