

LAB 4 Linear models

ชื่อ-สกุล ตอน

1. โหลดชุดข้อมูล Auto-MPG (mile-per-gallon) จาก URL ในไฟล์ LAB
2. เขียนฟังก์ชัน `clean_data` เพื่อกำจัด rows ที่มี missing value ปรากฏ
3. จงอธิบายวิธีการจัดการข้อมูลใน feature “car name”
.....
.....
4. แบ่งชุดข้อมูลนี้ออกเป็น 80% และ 20% สำหรับ train และ test โมเดลตามลำดับ
5. เขียนสมการ linear regression ที่สร้างถูกสร้างโดยใช้ชุดข้อมูลฝึกฝนในข้อ 4 เพื่อทำนายค่า mile-per-gallon (mpg) ดังนี้

mpg =
.....
.....
6. ค่าความคาดเคลื่อนรวมเฉลี่ย (mean square error) บนชุดข้อมูล test และค่าคาดเคลื่อนรวมเฉลี่ย (mean absolute error) บนชุดข้อมูล test พร้อมให้เหตุผลถึงความแตกต่างระหว่างตัวชี้วัดทั้งสองนี้
.....
.....
7. โหลดชุดข้อมูล ‘transaction.csv’ ใน google classroom ซึ่งประกอบไปด้วยรายการยอดขายรายวันของสินค้า P1, P2 และ P3 ตามลำดับ เริ่มตั้งแต่วันที่ 1-1-2021 ถึง 31-12-2023 (จำนวน 730 วัน)
8. ค่าเฉลี่ยของยอดขายรายวันของ P1 = และ P3 =
9. สร้างโมเดล linear regression เพื่อทำนายยอดขายวันถัดไป $s(t+1)$ จากวันปัจจุบัน $s(t)$ ของสินค้า P1 โดยใช้ข้อมูล 700 วันแรก
10. ค่าความคาดเคลื่อนรวมเฉลี่ย MAE ของวันที่ 1-12-2023 ถึง 31-12-2023 เท่ากับ
11. ปรับปรุงโมเดลในข้อที่ 8 เพื่อทำนายยอดขายในสัปดาห์ถัดไป $w(t+1)$ จากยอดขายสัปดาห์ปัจจุบัน $w(t)$ และยอดขายสัปดาห์ก่อนหน้า $w(t-1)$ ของสินค้า P2
12. ความแม่นยำของโมเดลข้อ 10 โดยใช้ข้อมูล 4 สัปดาห์สุดท้าย

LAB 4 Linear models

ชื่อ-สกุล ตอน

13. เขียนสมการ linear regression ของโมเดลในข้อ 11

.....

.....