**เฉลยแบบประเมิน Lab 2**

**แถบสีเหลือง คือ คำตอบที่ถูกต้อง, ไอคอนรูปคน คือ คำตอบที่คุณเลือก**

1. หากต้องการสร้าง training pipeline สำหรับ regression model เมื่อพิจารณา dataset พบว่าเป็นข้อมูลชนิดตัวเลขจำนวนหลายคอลัมน์โดยแต่ละคอลัมน์แสดงค่าด้วยหน่วยที่แตกต่างกัน ดังนั้น จึงต้อง transform ให้ค่าเหล่านี้อยู่ในสเกลเดียวกัน ข้อใดคือ module ที่ควรใช้ใน pipeline?

Select Columns in a Dataset

Normalize Data

คำตอบที่คุณเลือก

Clean Missing Data

2. จากแบบฝึกหัดทำนายราคารถ ข้อใดกล่าวถูกต้องเกี่ยวกับ ขั้นตอนการ Split data

Splitting mode = Split Rows, Fraction of rows in the first output dataset = 0.7

คำตอบที่คุณเลือก

Splitting mode = Split Rows, Fraction of rows in the first output dataset = 0.3

Splitting mode = Split Columns, Fraction of rows in the first output dataset = 0.7

Splitting mode = Split Columns, Fraction of rows in the first output dataset = 0.3

3. ผล Evaluation result จากแบบฝึกหัดทำนายราคารถยนต์ข้อใดถูกต้อง (ค่าโดยประมาณ)

Mean\_Absolute\_Error = 0.085

Relative\_Squared\_Error = 1955.6

Relative\_Absolute\_Error = 0.2456

คำตอบที่คุณเลือก

Coefficient\_of\_Determination = 2892.3

4. ข้อใดคือสาเหตุของการแบ่งข้อมูลออกเป็น training set และ validation sets

Data is split into two sets in order to create two models, one model with the training set and a different model with the validation set.

Splitting data into two sets enables you to compare the labels that the model predicts with the actual known labels in the original dataset.

คำตอบที่คุณเลือก

Only split data when you use the Azure Machine Learning Designer, not in other machine learning scenarios.

5. ข้อใดกล่าวถูกต้องเกี่ยวกับ การนำคอลัมน์ของข้อมูลที่ไม่ต้องการนำไป Train ออก

เป็นขั้นตอนหนึ่งของ Data Transformations

ใช้ module Select Columns in Dataset

Add ทุกคอลัมน์ ยกเว้น normalized-losses ของ Dataset Automobile price data

ถูกทุกข้อ

คำตอบที่คุณเลือก