

# Young Developer Power - AI/ML Track Mini Project

## Instruction / Guidelines

- พยายามทำให้เต็มที่ สามารถส่งเท่าที่ทำได้/ทำเสร็จมาได้
- เป็นงานเดี่ยว ทำด้วยตัวเอง หาข้อมูลเพิ่มเติมออนไลน์หรือใช้ AI Tools ได้ (แต่ต้องเข้าใจนะ)

## Task Requirements

น้องเป็น AI/ML Developer ในทีมที่กำลังจะทำโมเดลเกี่ยวกับการสูญเสียลูกค้าของธนาคาร (ลูกค้ากำลังจะเลิกใช้ธนาคาร) โดยน้องได้รับมอบหมายให้

- ทำ EDA (Exploratory Data Analysis) เช่น
  - ทำ Data Visualization ของ Customer Profile เช่น % แต่ละประเทศ
  - การวิเคราะห์เชิงสถิติ เช่น หา Min, Max, Mean
- ทำ Data Preparation & Preprocessing
  - ใช้ข้อมูลที่ได้จาก EDA เพื่อตัดสินใจว่าต้องใช้ Column ไหนบ้าง ต้อง Clean Data ใหม่ (ข้อมูลว่าง ข้อมูลผิด)
- ทำโมเดล Classification เพื่อทำนายว่าลูกค้าจะใช้งานต่อหรือไม่ (Yes/No)
- (Optional) ทำ API สำหรับการใช้งานโมเดล

โดยใช้ [Dataset](#) นี้ในการสร้างโมเดล

## สิ่งที่ต้องส่ง

- Source Code ผ่าน GitHub หรือ GitLab
- เอกสารอื่น ๆ ตามความเหมาะสม
- Presentation สำหรับพรีเซนต์งาน

## อธิบาย Dataset

- RowNumber: ใอดีของ Row นั้น ๆ ไม่มีผลต่อข้อมูล
- CustomerId: ใอดีเฉพาะตัว (ไม่ซ้ำ) ของลูกค้า ไม่มีผลต่อการใช้งานของลูกค้า
- Surname: นามสกุล ไม่มีผลต่อการใช้งานของลูกค้า
- CreditScore: คะแนนความน่าเชื่อถือทางการเงิน (ยิ่งไปแล้วจะคืนไหม) คนที่มี CreditScore สูงมักจะไม่เลิกใช้งาน
- Geography: ที่อยู่ของลูกค้า อาจมีผลต่อการใช้งานของลูกค้า
- Gender: เพศ อาจมีผลต่อการใช้งานของลูกค้า
- Age: อายุ คนที่มีอายุมากกว่ามักจะไม่เลิกใช้งาน
- Tenure: จำนวนปีที่ลูกค้าใช้งานมาแล้ว หากใช้งานมานาน ก็มักจะเสี่ยงเลิกใช้งานน้อยกว่า
- Balance: เงินในบัญชี ลูกค้าที่มีเงินในบัญชีสูงมักจะไม่เลิกใช้งาน
- NumOfProducts: จำนวนผลิตภัณฑ์ที่ลูกค้าซื้อจากธนาคาร
- HasCrCard: ลูกค้ามีบัตรเครดิตหรือไม่ คนที่มีมักจะไม่เลิกใช้งาน
- IsActiveMember: ใช้งานเป็นประจำหรือไม่
- EstimatedSalary: เงินเดือนโดยประมาณ หากเงินเดือนสูงอาจจะไม่เลิกใช้งาน
- Exited: ลูกค้าได้เลิกใช้งานไปหรือยัง
- Complain: ลูกค้ามีเรื่องร้องเรียนหรือไม่
- Satisfaction Score: คะแนนความพึงพอใจจากลูกค้าที่มีต่อการจัดการเรื่องร้องเรียน
- Card Type: ประเภทการ์ดที่ลูกค้าใช้
- Points Earned: จำนวนคะแนนแต้มสะสมที่ลูกค้าได้รับจากการใช้ Credit การ์ด

Criteria

Criteria	Weights
Data Understanding	25%
มีการวิเคราะห์ทำความเข้าใจ Data ได้เหมาะสม	10%
สรุป Insights จาก EDA ได้	15%
Data Preparation & Preprocessing	20%
ตรวจสอบจัดการ Data ได้อย่างถูกต้อง	10%
เลือก Column มาใช้งานได้อย่างเหมาะสม	10%
Model Building & Evaluation	30%
การเลือก Model และปรับ Parameter เหมาะสม	15%
แสดงผลการเทรนและมีการวิเคราะห์ผลลัพธ์	15%
Code Quality & Documentation	15%
โค้ดอ่านง่าย มี Structure เป็นระบบ	5%
มีการทำ Documentation, วิธีรัน, Version Control	10%
Presentation	10%
พรีเซนตีสื่อสารได้ชัดเจน	5%
มีความเข้าใจ ตอบคำถามได้	5%