PORTFOLIO

แฟ้มสะสมผลงานประกอบการสัมภาษณ์



นายภูวดล วรบุตร

วิทยาศาสตรบัณฑิต มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ



ข้อมูลส่วนตัว

้ชื่อ นายภูวดล วรบุตร

ชื่อเล่น อั๊ต อายุ 24 ปี

วัน/เดือน/ปีเกิด 22 ตุลาคม 2542

กรุ๊ปเลือด A ศาสนา พุทธ

เชื้อชาติ ไทย สัญชาติ ไทย

ประวัติการศึกษา

2563 - 2566

มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

คณะ วิทยาศาสตร์และเทคโนโลยี

สาขา วิทยาการการจัดการข้อมูล

GPA 3.26

2559 - 2562

โรงเรียนพระแม่สกลสงเคราะห์ แผนการเรียนสาย วิทย์ - คณิต

ประวัติส่วนตัว

จุดมุ่งหมายในการทำงาน

ผมเป็นนักศึกษาจบใหม่มีความสนใจในสายงาน Data เนื่องจาก ตอนทำโปรเจคมีโอกาสได้ทำโปรเจคเกี่ยวกับ Machine Learning ได้ออกแบบโมเดลทำยังไงถึงจะได้ โมเดลที่ดีที่สุดและพอได้ศึกษาจริงๆเลยมีความชอบใน Machine Learning เพราะมีการทำอะไรสนุกๆ ได้หลาย อย่างและตอนฝึกงานได้เข้าไปอยู่ในทีม Data และมี โอกาสได้ทำงานไม่ว่าจะเป็นการเขียนสคริปต์สำหรับแยก ไฟล์รูปแบบ JSON ด้วยภาษา Python การแกะหาที่มา ของข้อมูลจากไฟล์Macro และการทำ ETL โดยใช้ภาษา C# แต่ตอนนั้นไม่มีความรู้และไม่เคยเรียนมาก่อนและมี เวลาจำกัด มันเป็นอะไรที่ท้าทายดีถึงแม้จะปวดหัวไปบ้าง แต่ก็ชอบในสายงานนี้รู้สีกทุกอย่างที่ทำในสายงานนี้เป็น อะไรที่ต้องหาความรู้อยู่ตลอดเวลาและดูสนุกไปหมด

Skills

Programming Language

- Python (Pandas , Numpy , Scikit-learn ,TensorFlow , Matplotlib) (Good)
- C# (Good)
- Golang (Fair)

Database

- SQL (Good)
- PL/SQL (Fair)

Web Developer

- HTML & CSS (Fair)
- JavaScript (Fair)

Tool

- Microsoft Office
- Rapidminer
- Visual Studio Code
- Visual Studio 2022
- Google Colab
- Jupyter Notebook

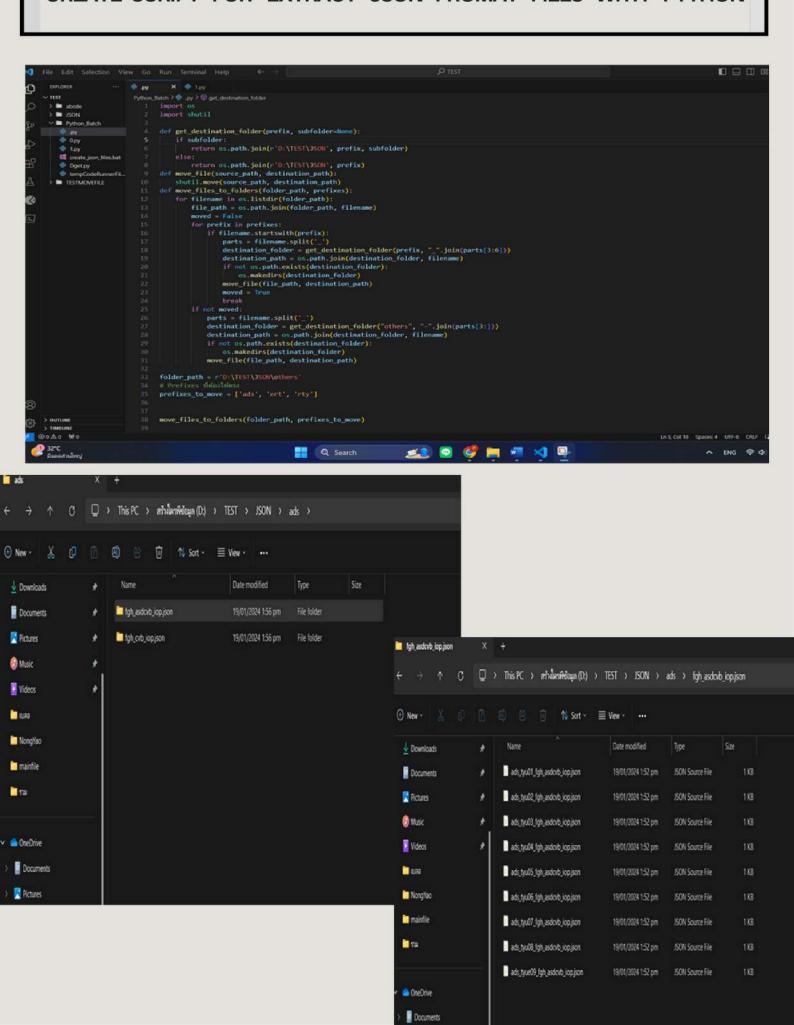
Database

- Microsoft SQL Server
- MySQL
- SQL Developer

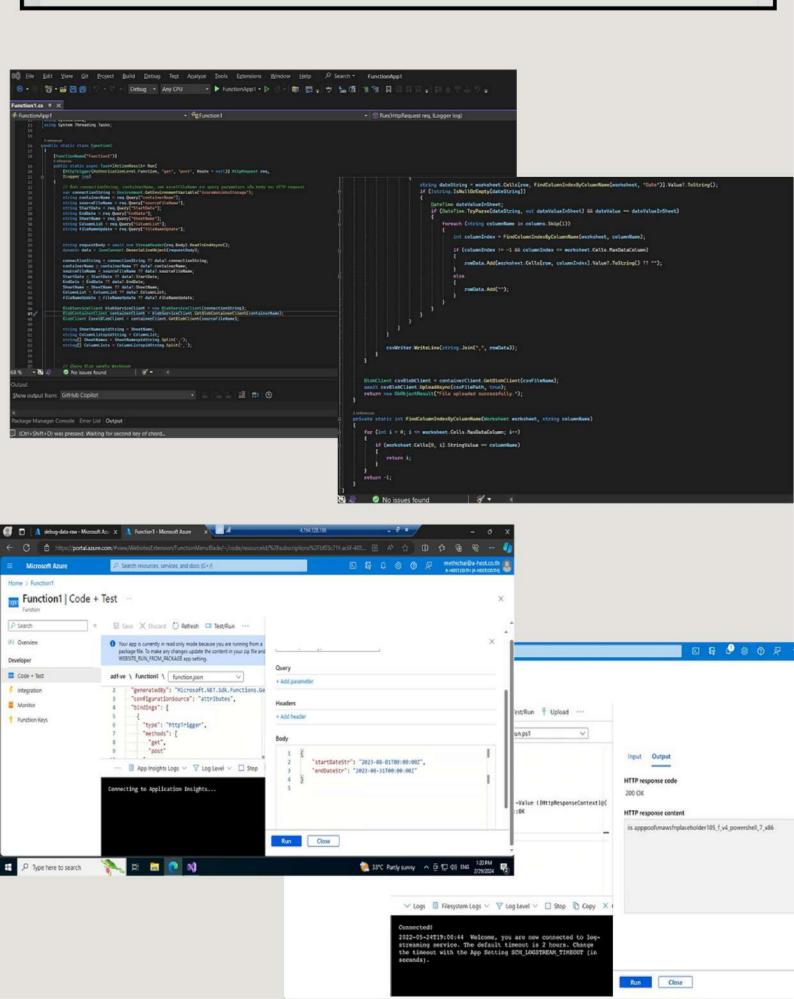
Data Visualization

- Power BI (DAX)
- Looker Studio

CREATE SCRIPT FOR EXTRACT JSON FROMAT FILES WITH PYTHON



Create C# Script on Azure Functions for transform data to csv format files



Project

การทำนายราคาทองคำด้วยการเรียนรู้ของเครื่องบนอนุกรมเวลาแบบแบน

วัตถุประสงค์

เพื่อออกแบบ พัฒนา และทดสอบแบบ จำลองราคาทองคำรายวัน และวิเคราะห์ ประเมินประสิทธิภาพของแบบจำลอง ตลอดจนเปรียบเทียบประสิทธิภาพของ แบบจำลองการทำนาย ราคาทองคำแบบ ต่างๆ โดยใช้ข้อมูลย้อนหลัง 10 ปี จาก เว็บทองคำ.com หรือ สมาคมค้าทองคำ และจะแบ่งข้อมูลเป็น 3 ส่วน เป็น 7 วัน 14 วัน และ 30 วันย้อนหลัง

อัลกอลิธึมที่ใช้

- Decision Tree Regression
- Random Forest Regression
- Artificial Neural Network
- Linear Regression
- Lasso Regression

เครื่องมือในการดำเนินโครงงาน

Microsoft Excel, Python, Google Colab

มีการใช้ไลบรารี่จาก Scikit – Learn และ TensorFlow

ทดสอบโมเดลด้วย

MAE, MSE, RMSE และโมเดล Baseline และสร้างแบบจำลองการซื้อ ขายเพื่อดูกำไรว่าโมเดลที่ได้ค่า MAE, MSE, RMSE น้อยที่สุดหรือดีที่สุดจะได้ กำไรหรือขาดทุน

```
df = pd.read_excel('/content/drive/MyDrive/Project/pit.xlsx')
   x123 = df.iloc[:, 1]
   x123_normalized = (x123 - x123.min()) / (x123.max() - x123.min())
   USD = df.iloc[:, 2]
   USD\_normalized = (USD - USD.min()) / (USD.max() - USD.min())
   plt.scatter(x123 normalized, USD normalized, c=USD normalized, label='USD')
   plt.xlabel('Goldspot Data (Normalized)')
   plt.ylabel("USD Data (Normalized)")
   plt.title("Scatter Plot of Normalized Goldspot vs Normalized USD")
   plt.colorbar(label='USD (Normalized)')
   plt.legend()
   plt.show()
   GOLDSPOT = df.lloc[:, 1]
   DATE = df.iloc[:, 0]
   plt.figure(figsize=(20, 6))
   plt.plot(DATE, GOLDSPOT)
   plt.grid(True)
   plt.xlabel('year')
   plt.ylabel('gold spot')
   plt.show()
   # sns.pairplot(df)
   data_train = pd.read_excel('/content/drive/MyDrive/Project/PDATA_Flantten_Time Series.xlsx', '7DAY')
   data_test = pd.read_excel('/content/drive/MyDrive/Project/PDATA_Flantten_Time Series.xlsx', 'DATA TEST 7DAY')
   train_target = data_train.iloc[:, 1] # Y
   train_factor = data_train.iloc[:, 4:] # x
   test_target = data_test.iloc[:, 1] # Y
   test_factor = data_test.iloc[:, 4:] # x
   2000
   1800
gold spod
   1400
   1200
                                                                                                                                                                                                                           2023
                 2013
                                     2014
                                                         2015
                                                                             2016
                                                                                                  2017
                                                                                                                      2018
                                                                                                                                          2019
                                                                                                                                                              2020
                                                                                                                                                                                  2021
                                                                                                                                                                                                      2022
                                                                                                                          lasso_model = Lasso(positive=True)
lasso_model.fit(train_factor, train_target)
predictions = lasso_model.predict(test_factor)
                                                                                                                          predictions = np.round(predictions, 2)
                                                                                                                          date = data_test.floc[:, 0]
                                                                                                                          for idx, (d, target, pred) in enumerate(zip(date, test_target, predictions), start=1):
        plt.figure(figsize=(15, 6))
                                                                                                                            date_str = d.strftime("%Y-%m-%d")
# print(f"Row (idx): Date:(date_str), Actual: (target), Predicted: (pred)")
        plt.plot(test_target.values.flatten(), label='test_target')
        plt.plot(predictions.flatten(), label='predictions')
                                                                                                                          mse = mean_squared_error(test_target, predictions)
print(f"MSE:", mse)
        plt.legend()
        plt.show()
                                                                                                                          mae = mean_absolute_error(test_target, predictions)
                                                                                                                          print(f"MAE: {mae}")
                                                                                                                          rmse = mean_squared_error(test_target,predictions, squared=False)
        RES=test_target - predictions
                                                                                                                          print(f"RMSE: (rmse)")
        sns.displot(RES,bins=30, kde=True)
                                                                                                                          r_squared = r2_score(test_target, predictions)
                                                                                                                          print(f"R-squared: (r_squared)")
                                                                                                                          MSE: 223.2899517915311
MAE: 11.005667752443001
RMSE: 14.942889673404242
```



แบบจำลองการซื้อขาย

Action

Action:

Action:

Actions

Action:

Action:

Action:

Action:

Action:

Action:

Profit:

Profit:

Buy/Sell:

Buy/Sell:

Buy/Sell: 0,

Action:

เลียโอกาส ใม่ทำการชื่อขาย

da da

ใม่ฟาการซื้อขาย

бa

เสียโลกาส ไม่ฟาการซื้อขาย

do.

เสียโลทาส

เสียโลยาส

"ไม่ฟาการซื้อขาย

ไม่ฟาการซื้อขาย

```
date_str = d.strftime('%Y-%m-%d')
    if actual1 < predictions:
        buy sell indicator = 1
        profit = actual2 - actual1
        action_description = 12a
        if actual1 < actual2:
          buy_sell_indicator = 0
           profit = 0
           action_description = "เสียโอกาส"
          buy_sell_indicator = 0
           action_description = "ไม่ฟาการซื้อขาย"
    buy_sell.append(buy_sell_indicator)
    profits.append(profit)
    if action_description == "เสียโอกาส":
       missed_opportunities_count += 1
     elif action_description ==
       no trade count += 1
     elif action_description == "#a":
        buy_count += 1
        if profit > 0:
          profitable_buy_count += 1
        elif profit < 0:
           loss_buy_count += 1
        if profit == 0:
Row 290: Date: 2022-12-12,
                                                                    Actual 2: 1784.5,
                                                                                               Predictions tomorow:
                                                                                                                                           Buy/Sell:
                                                                                                                                                                       Profit:
                                                                                                                                                                                   -4.5.
 low 291: Date: 2022-12-13,
                                      Actual 1:
                                                                    Actual 2: 1810.5
                                                                                               Predictions tomorow:
                                                                                                                     [1784.31],
                                                                                                                                                                       Profit:
Row 292: Date: 2022-12-14,
                                      Actual 1:
                                                 1810.5,
                                                                    Actual 2: 1797.0,
                                                                                               Predictions tomorow:
                                                                                                                     [1808.72],
                                                                                                                                           Buy/Sell:
                                                                                                                                                                      Profit:
Row 293: Date: 2022-12-15,
                                                                    Actual 2: 1780.5,
Row 294: Date: 2022-12-16.
                                                                    Actual 2: 1793.5.
                                                                                                                                         Buw/Sell:
                                      Actual 1:
                                                 1780.5.
                                                                                               Predictions tomorow:
                                                                                                                     [1781.L
                                                                                                                                                                    Profit:
 ow 295: Date: 2022-12-17,
                                                                    Actual 2:
 low 296: Date: 2022-12-19.
                                      Actual 1:
                                                                    Actual 2: 1786.0.
                                                                                               Predictions tomorow:
                                                                                                                     [1794.55],
                                                                                                                                           Buy/Sell:
                                                                                                                                                                      Profit:
 low 297: Date: 2022-12-20.
                                      Actual 1:
 low 298: Date: 2022-12-21.
                                      Actual 1:
                                                 1816.0.
                                                                    Actual 2: 1819.0.
                                                                                               Predictions tomorow:
                                                                                                                     [1814.22].
                                                                                                                                                                       Profit:
 low 299: Date: 2022-12-22,
                                                                    Actual 2: 1794.5.
                                                                                                                     [1818.36].
                                      Actual 1:
                                                 1819.0,
                                                                                               Predictions tomorow:
                                                                                                                                           Buy/Sell:
                                                                                                                                                                       Profit:
 low 300: Date: 2022-12-23,
                                                                    Actual 2: 1799.0,
Row 301: Date: 2022-12-24.
                                      Actual 1:
                                                 1799.0.
                                                                    Actual 2: 1799.0.
                                                                                               Predictions tomorow:
                                                                                                                     [1798,41].
                                                                                                                                           Buy/Sell:
                                                                                                                                                                       Profit:
 ow 302: Date: 2022-12-26,
Row 303: Date: 2022-12-27.
                                      Actual 1:
                                                 1805.0.
                                                                    Actual 2: 1811.0.
                                                                                               Predictions tomorow:
                                                                                                                     [1804.25].
                                                                                                                                           Buy/Sell:
                                                                                                                                                                       Profit:
 low 304: Date: 2022-12-28,
```

missed_opportunities_count = 0 no_trade_count = 0 buy_count = 0 profitable buy count = 0

predictions = pd.DataFrame({'Predictions': predictions})

for idx, (d, actual1, actual2, predictions) in enumerate(zip(date, Actual1, Actual2, predictions), start=1):

predictions = predictions.shift(periods=-1) predictions = np.array(predictions)

loss_buy_count = 0 zero_profit_buy_count = 0 buy_sell_indicator_zero_count = 0

predictions Actual1 = test_target Actual2 = test_target.shift(-1)

buy_sell = [] profits = []

Row 305: Date: 2022-12-29.

ow 306: Date: 2022-12-30,

Actual 1: 1807.5.

1816.0.

ผลรวมของกำไร: 73.0 จำนวนการซื้อ: 96 จำนวนการซื้อกำไร: 47 จำนวนการซื้อขาดทุน: 46 จำนวนการซื้อแต่ไม่ได้กำไร 3 จำนวนครั้งที่ไม่ซื้อ: 211 จำนวนครั้งที่ ที่เสียโอกาส: 100

Predictions tomorow:

Predictions tomorow:

Predictions tomorow: [nan],

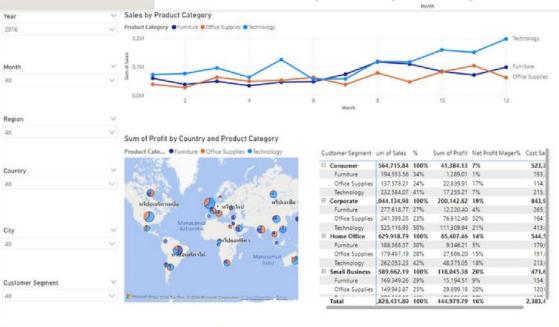
Actual 2: 1816.0.

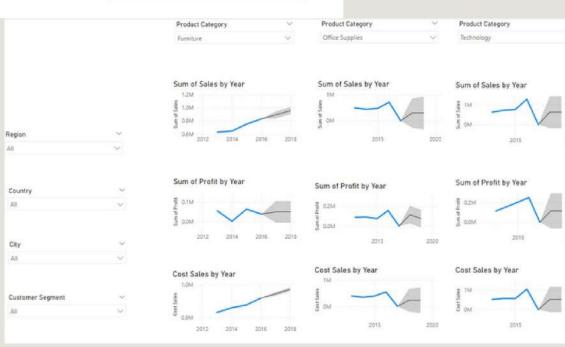
Actual 2: 1824.0,

จำนวนครั้งที่ ที่เสยโอกาส: 100 จำนวนครั้งที่ใม่ทำการซื้อขาย: 111

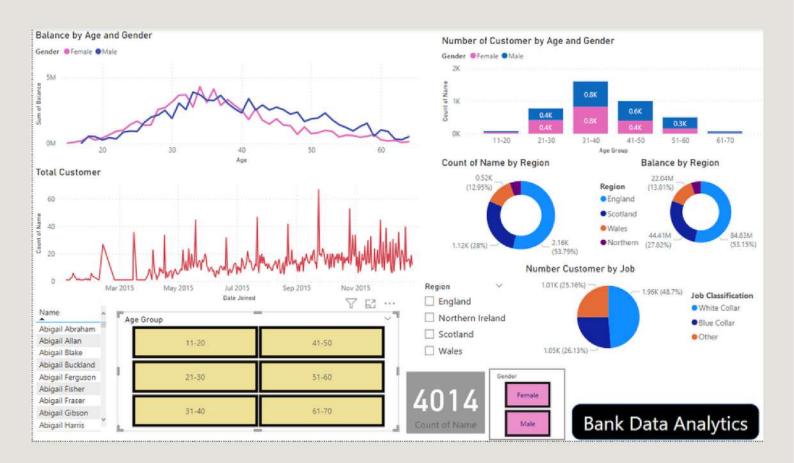
Data Visualization

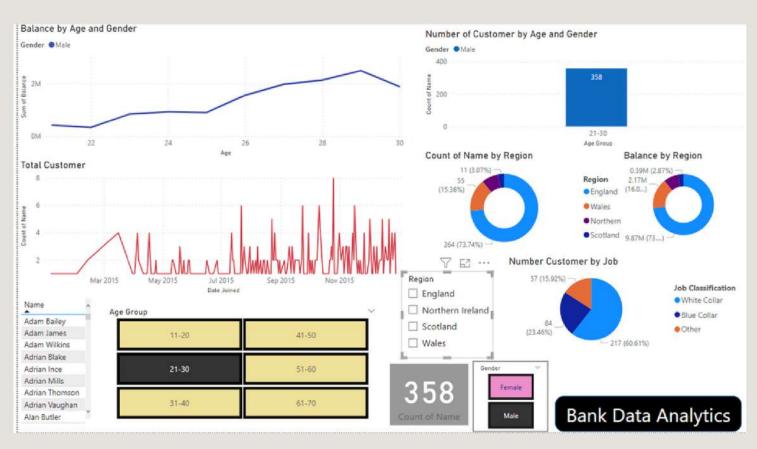






Data Visualization





SQL-ETL

```
SQLQuery1.sql - MI...r (MIND\athit (58))* # X
Object Explorer
                                            -USE Data Phuvadol:
Connect → + × = Y C -
                                            CREATE TABLE D_EMPLOYEE_ORA(

□ I MIND\SQLEXPRESS (SQL Server 16.)

                                                 EMPLOYEE CD INT.
                                                 EMPLOYEE_NAME NVARCHAR(200),
  Databases
                                                 EMPLOYEE_EMAIL NVARCHAR(50)

⊕ ■ System Databases

                                                 EMPLOYEE PHONE NUMBER NVARCHAR (200),

    ⊕ ■ Database Snapshots

                                                 HIRE_DATE DATE,

☐ ■ Data_Phuvadol

                                                 LAST DATE NVARCHAR(50),
       m Database Diagrams
                                                 JOB_TITLE NVARCHAR(35)

⊟ ■ Tables

                                                 EMPLOYEE_COMMISSION DECIMAL(5,2),
          MANAGER_CD INT,
          FileTables
                                                 MANAGER_NAME NVARCHAR(200),
          External Tables
                                                 ETL_DATE DATETIME2(6)
                                                 ETL_LAST_UPDATE DATETIME2(6)
          F Graph Tables
                                             );
          ⊞ dbo.D_DEPARTMENT_O
                                             USE Data_Phuvado

■ ■ dbo.D_EMPLOYEE_ORA

                                                                                                    SQLQuery1.sql - Ml...r (MIND\athit (58))* = X |
                                                                                                         USE Data_Phuvadol;
          CREATE TABLE D. Connect - # *# = Y C -*

■ ■ dbo.employees

                                                 DEPARTMENT_ | B MIND\SQLEXPRESS (SQL Server 16.)
          ⊞ III dbo.F_SALARY_ORA
                                                 DEPARTMENT I
                                                                □ ■ Databases
                                                                                                         CREATE OR ALTER VIEW D_EMPLOYEE_ORA_V AS

■ dbo.job_history

                                                 STREET_ADDRE
                                                                   E.EMPLOYEE_ID as EMPLOYEE_CD,

    ⊞ dbo.jobs

                                                 POSTAL CODE

    ■ Database Snapshots

                                                                                                         CONCAT(TRIM(E.FIRST_NAME), '', TRIM(E.LAST_NAME)) AS EMPLOYEE_NAME, CONCAT(E.EMAIL, '@gmail.com') AS EMPLOYEE_EMAIL ,
                                                 CITY NVARCHA

    ⊞ dbo.locations

□ □ Data_Phuvadol

                                                 STATE_PROVI

    ⊞ dbo.regions

    Database Diagrams

                                                                                                         REPLACE(E.PHONE_NUMBER,
                                                                                                                                  .', '-') AS EMPLOYEE_PHONE_NUMBER ,
                                                 COUNTRY_NAME
       E Views
                                                                     ☐ ■ Tables
                                                                                                         E.HIRE_DATE as HIRE_DATE ,
                                                 REGION_NAME.

    ■ External Resources

                                                                       System Tables
                                                                                                         ISMULL(CONVERT(NVARCHAR, JH.END_DATE, 121), 'NA') AS LAST_DATE,
                                                 ETL_DATE DAT

⊕ ■ Synonyms

                                                                        3.308_TITLE as 308_TITLE
                                                 ETL_LAST_UPL
                                                                                                         COALESCE(E.COMMISSION_PCT, 0) AS EMPLOYEE_COMMISSION,

    ■ Programmability

    ■ Graph Tables

                                                                                                         E.MANAGER_ID as MANAGER_CD,
       H Query Store
                                                                                                         (SELECT CONCAT(TRIM(E.FIRST_NAME), ' ', TRIM(E.LAST_NAME))

⊞ ■ Service Broker

                                                                                                         FROM employees M
                                                                        ⊞ m dbo.D DEPARTMENT O

    ■ Storage

                                                                                                         WHERE E.MANAGER_ID- M.EMPLOYEE_ID )AS MANAGER_NAME,
                                                                        ⊞ dbo.D_EMPLOYEE_ORA
                                             USE Data_Phuvado

    ■ Security

                                                                                                         GETDATE() AS ETL_DATE,
                                             CREATE TABLE F

    ⊞ dbo.departments

                                                                                                         GETDATE() AS ETL_LAST_UPDATE
     ⊞ ■ test
                                                 SEQ INT,
                                                                        FROM EMPLOYEES E

    ⊞ Security

                                                 YEAR CD INT
                                                                        LEFT JOIN JOBS J ON E.JOB_ID - J.JOB_ID

    ■ Server Objects

                                                 MONTH CD IN
                                                                        # III dbo.job_history
                                                                                                         LEFT JOIN JOB_HISTORY JH ON E.EMPLOYEE_ID = JH.EMPLOYEE_ID;
   EMPLOYEE_CD
                                                                        USE Data_Phuvadol;
                                                                        CREATE OR ALTER VIEW D_DEPARTMENT_ORA_V AS
                                                                     E Synonyms

■ Programmability

                                                                                                         D.DEPARTMENT_ID as DEPARTMENT_CD

    ⊕ Query Store

                                                                                                         D.DEPARTMENT_NAME as DEPARTMENT_NAME,
                                                                                                         LEFT(STREET_ADDRESS, 30)as STREET_ADDRES
                                                                     E Service Broker
                                                                                                         COALESCE(L.POSTAL_CODE, 'NA') as POSTAL_CODE,
                                                                      Storane
    USE Data_Phuvadol;
                                                                                                         L.city AS city,
                                                                                                         COALESCE(L.STATE_PROVINCE, 'NA') AS STATE_PROVINCE,
                                                                                                         C.COUNTRY_NAME,
  CREATE OR ALTER VIEW F_SALARY_ORA_V AS
                                                                                                         REGION_NAME as REGION_NAME_EN,
    SELECT
                                                                                                        GETDATE() AS ETL_DATE,
GETDATE() AS ETL_LAST_UPDATE
       YEAR (GETDATE()) AS YEAR_CD,
       CONVERT(VARCHAR(6), GETDATE(), 112) AS MONTH_CD,
       EMP.EMPLOYEE_ID AS EMPLOYEE_CD,
       J. JOB_ID AS JOB_CD,
        ISNULL(D.DEPARTMENT_ID, 0) AS DEPARTMENT_CD,
       ISMULL(CONVERT(VARCHAR, EMP.SALARY), '0') AS EMPLOYEE_SALARY, ISMULL(CONVERT(VARCHAR, MAX_SALARY), '0') AS MAX_JOB_SALARY, ISMULL(CONVERT(VARCHAR, MIN_SALARY), '0') AS MIN_JOB_SALARY,
            SELECT SALARY
           FROM EMPLOYEES
            WHERE EMP.MANAGER_ID - EMPLOYEE_ID
        ) AS MANAGER_SALARY,
        DENSE_RANK() OVER (PARTITION BY EMP.JOB_ID ORDER BY EMP.SALARY DESC) AS JOB_SALARY_RANKING,
       SYSDATETIMEOFFSET() AS ETL_DATE,
       SYSDATETIMEOFFSET() AS ETL_LAST_UPDATE
   FROM
       EMPLOYEES EMP
       JOBS J ON EMP. JOB_ID - J. JOB_ID
       DEPARTMENTS D ON EMP.DEPARTMENT_ID = D.DEPARTMENT_ID;
```

insert into d_employee_ora SELECT * FROM D_EMPLOYEE_ORA_V;

insert into D_DEPARTMENT_ORA SELECT = FROM D_DEPARTMENT_ORA_V;

INSERT INTO F_SALARY_ORA (
YEAR_CD,
MONTH_CD,
EMPLOYEE_CD,
JOB CD.

ETL - Python

```
import pandas as pd
       from datetime import datetime, timedelta
       import random
       import os
       # สร้างโฟลเตอร์หากยังใงเมื
       folder_path = r'D:\Phuvadol\skill\ETL\file_excel'
       if not os.path.exists(folder_path):
            os.makedirs(folder_path)
       for i in range(1, 21):
                 start_date = datetime(2023, 1, 1)
                  end_date = datetime(2024, 3, 30)
                  date_range = [start_date + timedelta(days=random.randint(0, (end_date - start_date).days))for _ in range(30)]
                  values = [random.randint(9, 999999) for _ in range(30)]
                 data = {'Date': date_range, 'Value': values}
                  df = pd.DataFrame(data)
                  # บันทึกเป็นไฟล์ Excel
                  file_path = os.path.join(folder_path, f'data_{i}.xlsx')
                 df.to_excel(file_path, index=False)
                  print(f"ไฟล์ '{file_path}' ถูกสร้างเวียบร้อยแล้ว")
                 print(f"เกิดข้อผิดพลาดในการสร้างไฟล์ที่ '{file_path}': {e}")
           OUTPUT DEBUG CONSOLE PORTS TERMINAL POSTMAN CONSOLE
ไฟล์ 'D:\Phuvadol\skill\ETL\file excel\data_15.xlsx' ดูกสร้างเรียกร้อยแล้ว
    'D:\Phuvadol\skill\ETL\file_excel\data_16.xlsx' ถูกเข้างเรียกต่อนแล้ว
'D:\Phuvadol\skill\ETL\file_excel\data_17.xlsx' ถูกเข้างเรียกต่อนแล้ว
'D:\Phuvadol\skill\ETL\file_excel\data_18.xlsx' ถูกเข้างเรียกต้อนแล้ว
Twb
    'D:\Phuvadol\skill\ETL\file_excel\data_19.xlsx' ถูกสร้างเรียกส่อนแล้
'D:\Phuvadol\skill\ETL\file_excel\data_20.xlsx' ถูกสร้างเรียกส่อนแล้ว
```

```
folder_path = r'D:\Phuvadol\skill\ETL\file_excel'
output_file_oath = r'O:\Phuvadol\skill\ETL\file_excel\output.xlsx'
data = []
# วนดูปผ่านไฟล์ในโฟลเดอร์
 for file_name in os.listdir(folder_path):
     if file_name.endswith('.xlsx'):
         file_path = os.path.join(folder_path, file_name)
                                                                                                                            File Name
                                                                                                                                                  Date
                                                                                                                                                                     Value
                                                                                                                                                                                      Number
         df = pd.read_excel(file_path)
latest_date_row = df[dff['Date'] == df['Date'].max()]

value = latest_date_row['Value'].values[0]

data.append({'File Name': file_name, 'Date': latest_date_row['Date'].values[0], 'Value': value'

result_df = pd.DataFrame(data)
                                                                                                                                             2024-03-25 12:0
                                                                                                                                                                         42,542 1
                                                                                                                         data 1.xlsx
                                                                                                                          data 2.xlsx
                                                                                                                                             2024-02-25 12:0
                                                                                                                                                                        167,569 2
                                                                                                                          data 3.xisx
                                                                                                                                             2024-03-26 12:0
                                                                                                                                                                       559,429 3

    พึ่งก็ขึ้นสำหรับการตรวจสอบว่าเลขในชื่อไฟส์ถูกต้องหรือไม่

                                                                                                                     5
                                                                                                                          data 4xisx
                                                                                                                                             2024-03-22 12:0
                                                                                                                                                                       410,521 4
def is_number_in_filename(file_name):
                                                                                                                          data_5.xlsx
                                                                                                                                             2024-03-24 12:0
                                                                                                                                                                        133,807 5
    match = re.search(r'\d+', file_name)
     return match.group() if match else None
                                                                                                                                             2024-03-13 12:0
                                                                                                                                                                        540,340
                                                                                                                          data 6.xlsx
result_df('Number') = result_df('File Name'].apply(is_number_in_filename)
                                                                                                                          data 7.xlsx
                                                                                                                                             2024-02-28 12:0
                                                                                                                                                                       808,734 7
                                                                                                                     9
                                                                                                                          data_8.xlsx
                                                                                                                                             2024-03-27 12:0
                                                                                                                                                                        301,296 8
result_df_invalid = result_df[result_df['Mumber'].isna()]
                                                                                                                    10
                                                                                                                          data_9.xlsx
                                                                                                                                             2024-03-22 12:0
                                                                                                                                                                       511,233 9
  แสดงให่ล้ที่ไม่ผ่านการตรวจสอบ
print("Invalid files:", result_df_invalid)
                                                                                                                          data_10.xlsx
                                                                                                                                             2024-03-02 12:0
                                                                                                                                                                       616,982 10
                                                                                                                                             2024-03-26 12:0
                                                                                                                                                                       194.369 11
                                                                                                                    12
                                                                                                                          data_11.xlsx
if not result_df_invalid.empty:
                                                                                                                    13
                                                                                                                          data 12.xlsx
                                                                                                                                             2024-03-25 12:0
                                                                                                                                                                       915,112 | 12
    print("Some files contain invalid numbers in their names. Please check and correct them.")
                                                                                                                                            2024-03-21 12:0
                                                                                                                                                                         71,794 13
                                                                                                                    14
                                                                                                                          data_13.xlsx
                                                                                                                    15
                                                                                                                         data_14.xlsx
                                                                                                                                             2024-03-25 12:0
                                                                                                                                                                        160,240 14
     def extract_number(file_name):
         return int(re.search(r'\d+', file_name).group())
                                                                                                                    16
                                                                                                                          data 15.xlsx
                                                                                                                                             2024-03-26 12:0
                                                                                                                                                                       498.812

    เรียงสาดบดามเลขในชื่อไฟล์

                                                                                                                    17
                                                                                                                          data 16.xlsx
                                                                                                                                             2024-03-04 12:0
                                                                                                                                                                        772.618 16
    result_df_sorted = result_df.sort_values(by='Number', key=lambda x: x.map(extract_number))
result_df_sorted.to_excel(output_file_path, index=False)
                                                                                                                    18
                                                                                                                          data_17.xlsx
                                                                                                                                             2024-03-28 12:0
                                                                                                                                                                         40,168 17
     print("Data has been written to:", output_file_path)
                                                                                                                          data_18.xlsx
                                                                                                                                             2024-03-27 12:0
                                                                                                                                                                        459,792 18
                                                                                                                    20
                                                                                                                          data_19.xlsx
                                                                                                                                             2024-03-26 12:0
                                                                                                                                                                       907,617
                                                                                                                                                                                  19
                                                                                                                                             2024-03-13 12:0
                                                                                                                                                                       998,538 20
                                                                                                                    21
                                                                                                                          data_20.xlsx
```

Web Scraping Python

```
import requests
from bs4 import BeautifulSoup
import pandas as pd
# ดึงข้อมลจากเว็บไซต์
response = requests.get("https://www.myhora.com/%E0%B8%AB%E0%B8%A7%E0%B8%A2/%E0%B8%AA%E0%B8
soup = BeautifulSoup(response.text, 'html.parser')
# ดึงข้อมูลที่ต้องการจากเว็บไซต์
data = []
temp_row = []
a = soup.find_all('div', class_='colx')
del a[:28]
for i, div in enumerate(a, start=1):
  temp_row.append(div.text.strip())
  if i % 12 == 0:
     data.append(temp_row)
     temp_row = []
if temp_row:
  data.append(temp_row)
df = pd.DataFrame(data)
df.to_excel("หวย.xlsx", index=False, header=False)
```

	Α	В	C	D	E	F	G	Н	I	J	K	L		М
1	1	เมษายน	M.E.	2567	67	803481	81	481	90	122 809	5 90	122 809	559 94	47
2	16	มีนาคม	มี.ค.	2567	67	997626	26	626	78	509 571	3 78	509 571	329 79	94
3	1	มีนาคม	มี.ค.	2567	67	253603	03	603	79	900 975	3 79	900 975	382 70	03
4	16	กุมภาพันธ์	n.w.	2567	67	941395	95	395	43	056 330	3 43	056 330	375 58	37
5	1	กุมภาพันธ์	n.w.	2567	67	607063	63	063	09	454 943	5 09	454 943	544 59	91
6	17	มกราคม	ม.ค.	2567	67	105979	79	979	61	429 931	1 61	429 931	196 63	35
7	30	ธันวาคม	5.A.	2566	66	625544	44	544	89	600 648	4 89	600 648	456 88	32
8	16	ชั้นวาคม	5.A.	2566	66	356757	57	757	85	058 410	5 85	058 410	584 96	54
9	1	ธันวาคม	5.9.	2566	66	251097	97	097	91	055 265	0 91	055 265	092 28	80
10	16	พฤศจิกายา	W.E.	2566	66	557990	90	990	14	346 412	7 14	346 412	778 96	51
11	1	พฤศจิกายา	W.U.	2566	66	743951	51	951	63	335 913	0 63	335 913	019 34	49
12	16	ตุลาคม	ด.ค.	2566	66	931446	46	446	44	167 398	2 44	167 398	272 97	70
13	1	ดูลาคม	ด.ค.	2566	66	727202	02	202	66	324 355	4 66	324 355	426 61	15
14	16	กันยายน	ก.ย.	2566	66	320812	12	812	46	037 699	0 46	037 699	057 34	44
15	1	กันยายน	ก.ย.	2566	66	915478	78	478	91	521 596	2 91	521 596	291 69	92
16	16	สิงหาคม	ส.ค.	2566	66	471782	82	782	67	431 739	7 67	431 739	737 74	42
17	31	กรกฎาคม	ก.ค.	2566	66	260453	53	453	11	268 708	3 11	268 708	387 60	01
18	16	กรกฎาคม	ก.ค.	2566	66	169530	30	530	62	261 384	0 62	261 384	066 78	80
19	1	กรกฎาคม	ก.ค.	2566	66	922605	05	605	16	281 867	4 16	281 867	491 94	47
20	16	มิถุนายน	រាំ.ប	2566	66	264872	72	872	30	519 628	2 30	519 628	202 87	74
21	1	มิถุนายน	มิ.ย	2566	66	125272	72	272	09	000 681	3 09	000 681	386 97	71
22	16	พฤษภาคม	W.A.	2566	66	132903	03	903	99	678 739	6 99	678 739	693 73	31
23	2	พฤษภาคม	W.A.	2566	66	843019	19	019	65	500 780	1 65	500 780	187 26	59
24	16	เมษายน	tal.tl.	2566	66	984906	06	906	71	670 678	5 71	670 678	551 79	97
25	1	เมษายน	131. 21.	2566	66	087907	07	907	99	111 914	2 99	111 914	290 69	98
26	16	มีนาคม	มี.ค.	2566	66	025873	73	873	73	420 800	3 73	420 800	355 54	44
27	1	มีนาคม	มี.ค.	2566	66	417652	52	652	55	577 919	7.55	577 919	748 98	34
28	16	กุมภาพันธ์	n.w.	2566	66	590417	17	417	80	195 664	3 80	195 664	377 52	23
29	1	กุมภาพันธ์	ก.พ.	2566	66	297411	11	411	92	181 789	1 92	181 789	101 66	54
30	17	มกราคม	ม.ค.	2566	66	812519	19	519	47	389 443	5 47	389 443	564 84	49
31	30	ธันวาคม	5.9.	2565	65	157196	96	196	58	007 522	2 58	007 522	250 42	25
32	16	ชันวาคม	6.A.	2565	65	845093	93	093	14	411 912	5 14	411 912	593 85	55
33	1	ธันวาคม	5.9.	2565	65	375805	05	805	08	170 786	4 08	170 786	409 42	21
34	16	พฤศจิกายา	W.U.	2565	65	121789	89	789	64	532 722	1 64	532 722	157 97	73
35	1	พฤศจิกายา	W.E.	2565	65	913106	06	106	70	722 839	3.70	722 839	343 92	22
36	16	ตุลาคม	A.A.	2565	65	613106	06	106	15	037 158	6 15	037 158	606 79	99

Association Rule - Python

```
import numpy as np
from mlxtend.frequent_patterns import apriori, association_rules
<>:1: DeprecationWarning: invalid escape sequence '\P'
:1: DeprecationWarning: invalid escape sequence '\P'
                                        rnel_13800\1690963744.py:1: DeprecationWarning: invalid escape sequence '\P'
 df = pd.read csv("D:\Phuvadol\Portfolio ETC\Data Set\Market Basket Optimisation.csv")
                                    vegetables
                                                                                                          green
                                                                                                                                                               frozen
noothie
                                                   green
                                                                         cottage
                                                                                         tomato
                                                                                                  low fat
                                                            weat
flour
                almonds avocado
                              eggs
       chutney
                    NaN
                              NaN
                                          NaN
                                                    NaN
                                                            NaN
                                                                  NaN
                                                                           NaN
                                                                                    NaN
                                                                                            NaN
                                                                                                    NaN
                                                                                                           NaN
                                                                                                                   NaN
                                                                                                                         NaN
                                                                                                                                  NaN
                                                                                                                                           NaN
                                                                                                                                                       NaN
                                                                                                                                                                 NaN
                                                                                                                                                                          NaN
                                                                                                                                                                                NaN
                              NaN
                                          NaN
                                                                                                                                                                                NaN
        turkey
                  avocado
                                                    NaN
                                                                  NaN
                                                                           NaN
                                                                                    NaN
                                                                                            NaN
                                                                                                    NaN
                                                                                                           NaN
                                                                                                                   NaN
                                                                                                                         NaN
                                                                                                                                  NaN
                                                                                                                                           NaN
                                                                                                                                                       NaN
                                                                                                                                                                 NaN
                                                                                                                                                                          NaN
                             energy
bar
                                         whole
                                                green tea
                                                                                    NaN
                                                                                                                         NaN
                                                                                                                                           NaN
                                     wheat rice
                    NaN
                              NaN
                                          NaN
                                                    NaN
                                                            NaN
                                                                  NaN
                                                                            NaN
                                                                                    NaN
                                                                                            NaN
                                                                                                     NaN
                                                                                                           NaN
                                                                                                                   NaN
                                                                                                                         NaN
                                                                                                                                  NaN
                                                                                                                                           NaN
                                                                                                                                                       NaN
                                                                                                                                                                 NaN
                                                                                                                                                                          NaN
                                                                                                                                                                                NaN
 7495
        butter
               light mayo
                                          NaN
                                                    NaN
                                                            NaN
                                                                  NaN
                                                                           NaN
                                                                                    NaN
                                                                                            NaN
                                                                                                    NaN
                                                                                                           NaN
                                                                                                                   NaN
                                                                                                                         NaN
                                                                                                                                           NaN
                                                                                                                                                                 NaN
                                                                                                                                                                                NaN
                                                                                                                                  NaN
                                                                                                                                                                          NaN
                             bread
                                    french fries
                                                                  NaN
                                                                                                     NaN
                                                                                                                                           NaN
                                                                                                                                                                 NaN
       burgers
                              eggs
                                                                            NaN
                                                                                    NaN
                                                                                            NaN
                                                                                                           NaN
                                                                                                                   NaN
                                                                                                                         NaN
                                                                                                                                  NaN
                                                                                                                                                                          NaN
                                                                                                                                                                                NaN
                    NaN
                              NaN
                                                                                                                                           NaN
                                                                                                                                                                  NaN
                                                                                                                                                                                NaN
                                          NaN
                                                    NaN
                                                                  NaN
                                                                            NaN
                                                                                    NaN
                                                                                                     NaN
                                                                                                                                                                          NaN
                 green tea
                                          NaN
                                                                                                                                           NaN
                   frozen
                             yogurt
                                        low fat
                                                    NaN
                                                                  NaN
                                                                           NaN
                                                                                    NaN
                                                                                            NaN
                                                                                                     NaN
                                                                                                           NaN
                                                                                                                         NaN
                                                                                                                                           NaN
```

```
df = df.apply(lambda row: ','.join(row.dropna().astype(str)), axis=1 )
   df = pd.DataFrame(df, columns=["Products"])
   df
                                             Products
                                burgers, meatballs, eggs
                                              chutney
                                       turkey, avocado
       mineral water, milk, energy bar, whole wheat rice...
    4
                                         low fat yogurt
 7495
                          butter, light mayo, fresh bread
 7496
        burgers, frozen vegetables, eggs, french fries, ma...
 7497
                                               chicken
 7498
                                    escalope, green tea
        eggs,frozen smoothie,yogurt cake,low fat yogurt
7500 rows × 1 columns
```

fro a = a_d df	m mlxtend. Transacti ata = a.fi	preproces ionEncoder it(data).t rame(a_da	transform(dat uta,columns=a	Transaction (a)	nEncoder															
√ 12s	asparagus	almonds	antioxydant juice	asparagus	avocado	babies food	bacon	barbecue sauce	black tea	blueberries	turkey	vegetables mix	water spray	white wine	whole weat flour	whole wheat pasta	whole wheat rice	yams	yogurt cake	
																				0
																				0
					True						True									0
																	True			0
																				a
7495																				0
7496																				0
7497																				0
7498																				0
7499																			True	
7500 ro	ws × 120 col	lumns																		

```
true_counts = df.sum()
   sorted_counts = true_counts.sort_values(ascending=False)
   print(sorted counts)
mineral water
                 1787
                 1348
eggs
spaghetti
                1306
french fries
                 1282
chocolate
                 1229
bramble
                  14
cream
napkins
                                   df = df.replace(False,0)
water spray
 asparagus
Length: 120, dtype: object
```

df = pd.DataFrame(a_data,columns=a.columns_)
df = df.replace(False,0)

df = apriori(df, min_support = 0.05, use_colnames = True, verbose = 1)
df = df.sort_values(by='support', ascending=False)
df['support_percent'] = df['support'] * 100
df

Processing 6 combinations | Sampling itemset size 3 2 <u>c:\Users\athit\AppData\Local\Programs\Python\Python311\Lib\site-packages\mlxt</u> warnings.warn(

	support	itemsets	support_percent
16	0.238267	(mineral water)	23.826667
6	0.179733	(eggs)	17.973333
21	0.174133	(spaghetti)	17.413333
8	0.170933	(french fries)	17.093333
3	0.163867	(chocolate)	16.386667
12	0.132	(green tea)	13.2
15	0.1296	(milk)	12.96
13	0.098267	(ground beef)	9.826667
10	0.095333	(frozen vegetables)	9.533333
18	0.095067	(pancakes)	9.506667
0	0.0872	(burgers)	8.72
1	0.081067	(cake)	8.106667
4	0.0804	(cookies)	8.04
7	0.079333	(escalope)	7.933333
14	0.0764	(low fat yogurt)	7.64
19	0.071333	(shrimp)	7.133333
22	0.0684	(tomatoes)	6.84
1			

```
df_ar = association_rules(df, metric = "confidence", min_threshold = 0.06)
df_ar_sorted = df_ar.sort_values(by='confidence', ascending=False)
df_ar_sorted = df_ar_sorted.round(2)
df_ar_sorted
```

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
0	(spaghetti)	(mineral water)	0.17	0.24	0.06	0.34	1.44	0.02	1.16	0.37
2	(chocolate)	(mineral water)	0.16	0.24	0.05	0.32	1.35	0.01	1.12	0.31
4	(eggs)	(mineral water)	0.18	0.24	0.05	0.28	1.19	0.01	1.06	0.19
1	(mineral water)	(spaghetti)	0.24	0.17	0.06	0.25	1.44	0.02	1.10	0.40
3	(mineral water)	(chocolate)	0.24	0.16	0.05	0.22	1.35	0.01	1.07	0.34
5	(mineral water)	(eggs)	0.24	0.18	0.05	0.21	1.19	0.01	1.04	0.21

Certificate







หลักสูตรวิทยาการการจัดการข้อมูล มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ ขอมอบเกียรติบัตรฉบับนี้ให้ไว้เพื่อแสดงว่า

นาย ภูวดล วรบุตร

ผ่านการอบรม หัวข้อ

Google Data Studio กับการวิเคราะห์ข้อมูลเชิงธุรกิจ

โครงการพัฒนาสมรรถนะวิชาชีพนักศึกษาประจำหลักสูตร

วันที่ ๑๐ กุมภาพันธ์ พ.ศ ๒๕๖๖



(ดร.รุจิรา คงนุ้ย) ประชานหลักสูตรวิทยาการการจัดการข้อมูล ที่

Made for free with Certify'em



Certificate

OF COMPLETION

ภูวดล วรบุตร



has successfully completed Get started with Programming รวมสิ่งที่ต้องรู้ ก่อนเดินสู่เส้นทางโปรแกรมเมอร์

an online course offered by FutureSkill 03/04/2024



Futureskill has confirmed the identity of this individual and their participation in the course

Futureskill.co



Certificate

OF COMPLETION

ภูวดล วรบุตร



has successfully completed

UpSkill Python Programming เส้นทางสู่ Developer

an online course offered by FutureSkill 11/01/2024



Futureskill.co

Futureskill has confirmed the identity of this individual and their participation in the course



Certificate

OF COMPLETION

ภูวดล วรบุตร



has successfully completed จัดการฐานข้อมูลอย่างมีประสิทธิภาพ พร้อมวิเคราะห์ข้อมูลที่มีในไม่กี่อึดใจด้วย SQL Command

an online course offered by FutureSkill 19/12/2023



Futureskill has confirmed the identity of this individual and their participation in the course

Futureskill.co



แพลดฟอร์มเพื่อการเรียนรู้ออนใลน์ตลอดชีวิต | กระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม

ภูวดล วรบุตร

ได้ผ่านการเรียนออนไลน์ตามเกณฑ์การวัดผลในรายวิชา

วิทยาการข้อมูลและการประยุกต์ใช้ (30 ชั่วโมงการเรียนรู้)

พัฒนารายวิชาโดย จุฬาลงกรณ์มหาวิทยาลัย

ผศ.ดร.ณัฐาฒี หมูใพโรจน์ ผู้ช่วยศาสตรจารย์ ภาควิชาวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย



แพลตฟอร์มเพื่อการเรียนรู้ออนไลน์ตลอดชีวิต | กระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัดกรรม

ภูวดล วรบุตร

ได้ผ่านการเรียนออนไลน์ตามเกณฑ์การวัดผลในรายวิชา

การวิเคราะห์เชิงลึกอภิมหาข้อมูล (10 ชั่วโมงการเรียนรู้)

พัฒนารายวิชาโดย มหาวิทยาลัยรังสิต

ผศ. ดร. เชฏฐเนติ ศรีสอ้าน คณบดี, ว. นวัตกรรมดิจิทัลและเทคในโลยีสารสนเทศ มหาวิทยาลัยรังส็ต



แพลดฟอร์มเพื่อการเรียนรู้ออนไลน์ดลอดชีวิต | กระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม

ภูวดล วรบุตร

ได้ผ่านการเรียนออนไลน์ตามเกณฑ์การวัดผลในรายวิชา

ความรู้พื้นฐานเกี่ยวกับบิ๊กดาตา (10 ชั่วโมงการเรียนรู้)

พัฒนารายวิชาโดย มหาวิทยาลัยสุโขทัยธรรมาธิราช

ผศ.ดร. วฤษาย์ ร่มสายหยุด สาขาวิชาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช



แพลตฟอร์มเพื่อการเรียนรู้ออนไลน์ตลอดชีวิต | กระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม

ภูวดล วรบุตร

ได้ผ่านการเรียนออนไลน์ตามเกณฑ์การวัดผลในรายวิชา

ภาษาอังกฤษสำหรับเทคโนโลยีสารสนเทศ (10 ชั่วโมงการ เรียนรู้)

พัฒนารายวิชาโดย มหาวิทยาลัยขอนแก่น

ผศ.ดร.กรวิภา พูลผล

รองคณบดีฝ่ายแผนและสารสนเทศ คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยขอนแก่น



สำนักวิทยบริการและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

ขอมอบวุฒิบัตรให้ไว้เพื่อแสดงว่า

นาย ภูวดล วรบุตร

เข้าร่วมอบรมแนะนำการใช้งาน Google Workspace for Education Plus

ให้ไว้ ณ วันที่ 30 มกราคม 2566



ดร.เอกชัย เนาวนิช ผู้อำนวยการสำนักวิทยบริการและเทคโนโลยีสารสนเทศ







หลักสูตรวิทยาการการจัดการข้อมูล มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ ขอมอบเกียรติบัตรฉบับนี้ให้ไว้เพื่อแสดงว่า

ฎวดล วรบุตร

ผ่านการอบรม

กิจกรรมที่ ๑ หัวข้อ คุณธรรมการดำเนินชีวิตในยุคปัจจุบัน โครงการพัฒนาสมรรถนะนักศึกษาตามกรอบศตวรรษที่ ๒๑ วันที่ ๒๖ กรกฎาคม พ.ศ ๒๕๖๖



(ดร.รุจิรา คงนุ้ย) ประชานหลักสูตรวิทยาการการจัดการข้อมูล



สำนักวิทยบริการและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

ขอมอบวุฒิบัตรให้ไว้เพื่อแสดงว่า

นาย ภูวดล วรบุตร

ได้เข้าร่วมเสริมสร้างความฉลาดรู้ด้านดิจิทัลในอาเซียน : ASEAN DIGITAL LITERACY PROGRAMME (ADLP) จำนวน 3 ชม. ให้ไว้ ณ วันที่ 25 มกราคม พ.ศ. 2566

TO

นายเอกชัย เนาวนิช ผู้อำนวยการสำนักวิทยบริการและเทคโนโลยีสารสนเทศ



thank you



