



## รายงาน

เรื่อง การเปลี่ยนแปลงของอุณหภูมิเฉลี่ยตั้งแต่ปี ค.ศ.2017 ถึง ค.ศ.2022

จัดทำโดย

เลขที่ 10 รหัสนิสิต 65360802 นายเจษฎากรณ์ จันทรโต

เลขที่ 45 รหัสนิสิต 65363933 นายภูริชัช ขาวสบาย

รายงานนี้เป็นส่วนหนึ่งในรายวิชา

**Applied Statistics**

รหัสวิชา 305231 คณะวิศวกรรมศาสตร์

มหาวิทยาลัยนเรศวร

## ที่มาและความสำคัญ

ทางผู้จัดทำได้เลือกทำรายงานเรื่องนี้เพราะว่าเห็นถึงผลกระทบจากการเปลี่ยนแปลงอุณหภูมิของประเทศต่างๆในต่างประเทศ โดยผลกระทบที่พบมี ดังนี้

ผลกระทบต่อสุขภาพ: การเพิ่มอุณหภูมิที่เร็วมากหรือไม่น่าเชื่อถืออาจทำให้เกิดปัญหาสุขภาพร้ายและสถานการณ์ฉุกเฉินทางสุขภาพ เช่น การเพิ่มอุณหภูมิสูงสุดอาจทำให้เกิดอาการทางระบบการเผาผลาญอาหารหรืออาการเสียหายของระบบทั่วไปของร่างกาย

ผลกระทบต่อสภาพแวดล้อม: การเพิ่มอุณหภูมิสามารถทำให้เกิดการเปลี่ยนแปลงทางชีวภาพและทางชนิดของพืชและสัตว์ นอกจากนี้ยังสามารถเสียหายต่อสภาพแวดล้อมทั่วไป เช่น การทำให้ทะเลทั่วไประดับสูงขึ้นอาจเป็นอันตรายต่อชุมชนชายฝั่งและพื้นที่ที่อยู่ใกล้ทะเล

ผลกระทบต่อเศรษฐกิจ: การเพิ่มอุณหภูมิอาจมีผลต่อการเกษตรและการผลิตอาหาร หากสภาพอากาศเปลี่ยนแปลงและเกษตรกรไม่สามารถปรับตัวได้ นอกจากนี้ยังสามารถมีผลกระทบต่ออุตสาหกรรมและธุรกิจต่างๆ ที่ได้รับผลกระทบจากสภาพอากาศ

และทางผู้จัดทำเห็นว่าการใช้สถิติในการแก้ปัญหาเกี่ยวกับการเปลี่ยนแปลงอุณหภูมิสามารถทำได้หลายวิธี เช่น

การวิเคราะห์แนวโน้ม: การใช้สถิติเพื่อวิเคราะห์แนวโน้มของการเปลี่ยนแปลงอุณหภูมิในระยะยาว โดยสำรวจและวิเคราะห์ข้อมูลอุณหภูมิจากปีที่ผ่านมาเพื่อทำนายแนวโน้มในอนาคต

การวิเคราะห์ความสัมพันธ์: การใช้สถิติเพื่อวิเคราะห์ความสัมพันธ์ระหว่างการเปลี่ยนแปลงอุณหภูมิกับตัวแปรอื่น ๆ เช่น ความสัมพันธ์ระหว่างการเปลี่ยนแปลงอุณหภูมิกับการเปลี่ยนแปลงในระดับน้ำทะเล

การวิเคราะห์ผลกระทบ: การใช้สถิติเพื่อวิเคราะห์ผลกระทบของการเปลี่ยนแปลงอุณหภูมิต่อสุขภาพ สภาพแวดล้อม และเศรษฐกิจ เพื่อจัดการและวางแผนในการป้องกันและปรับตัวต่อผลกระทบที่อาจเกิดขึ้นในอนาคต

ดังนั้น การใช้สถิติเป็นเครื่องมือที่มีประสิทธิภาพในการวิเคราะห์และแก้ปัญหาเกี่ยวกับการเปลี่ยนแปลงอุณหภูมิ และการนำผลวิจัยและข้อมูลทางวิทยาศาสตร์มาประยุกต์ใช้ในการจัดการสิ่งแวดล้อมและเศรษฐกิจ

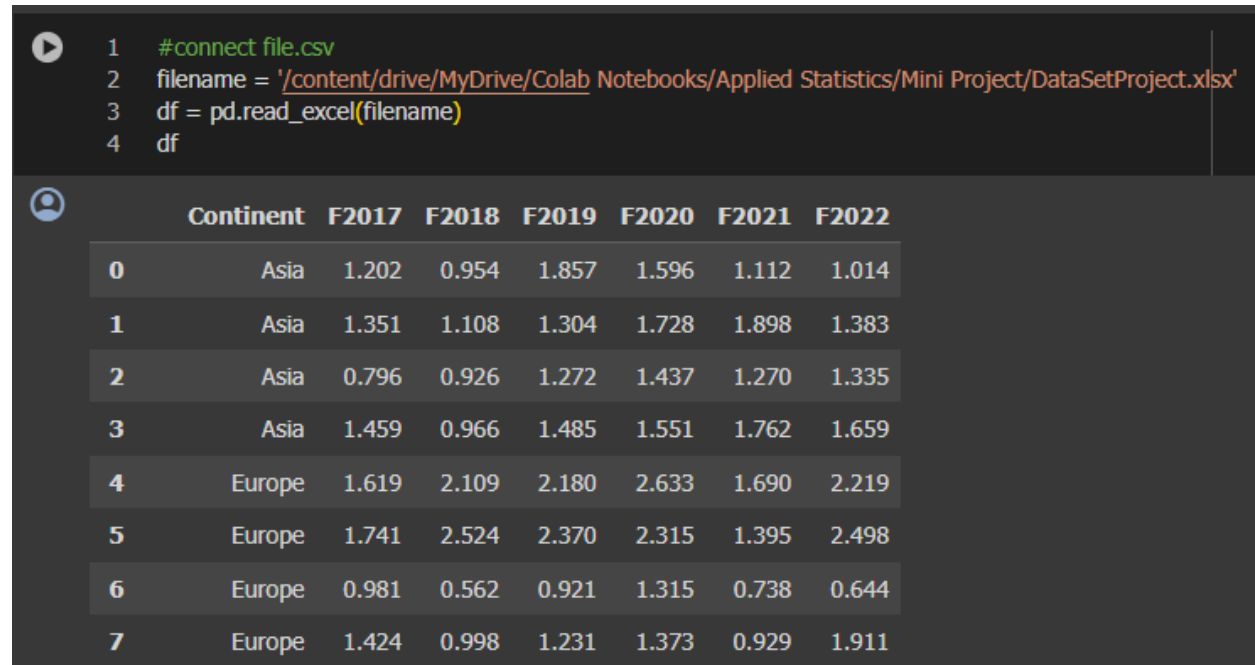
## หลักการและทฤษฎีที่เกี่ยวข้อง

**K-Nearest Neighbors** หรือที่เรียกว่า KNN หรือ k-NN เป็นวิธีการแบ่งคลาสสำหรับใช้จัดหมวดหมู่ข้อมูล (Classification) โดยมีหลักการนำข้อมูลอื่นๆมาเปรียบเทียบกับตัวข้อมูลที่สนใจ ว่ามีความใกล้เคียงกันมากแค่ไหน หากข้อมูลที่สนใจอยู่ใกล้กับข้อมูลใดมากที่สุด ระบบจะให้คำตอบเป็นเหมือนคำตอบของข้อมูลที่อยู่ใกล้ที่สุด โดยส่วนใหญ่จะใช้สำหรับแก้ปัญหาที่เราู้จำนวนกลุ่มที่แน่นอนอยู่แล้ว แต่มีบางข้อมูลบางตัวที่เราไม่สามารถบอกได้ว่าข้อมูลนั้นอยู่กลุ่มไหน เราก็จะใช้ระบบนี้เข้ามาช่วยเลือก ซึ่งมันจะคล้ายๆกับการโหวตเสียงข้างมาก

**การวิเคราะห์การถดถอยโลจิสติก (Logistic Regression Analysis)** เป็นเทคนิคการวิเคราะห์ตัวแปร เชิงพหุที่มีวัตถุประสงค์เพื่อประมาณค่าหรือทำนายเหตุการณ์ที่สนใจว่าจะเกิดหรือไม่เกิดเหตุการณ์นั้นภายใต้ อิทธิของตัวปัจจัย แบบจำลองโลจิสติกประกอบด้วยตัวแปรตาม (หรือตัวแปรเกณฑ์) ที่ต้องเป็นตัวแปรแบบทวินาม (Dichotomous Variable) กล่าวคือมีได้สองค่า เช่น “เกิด” กับ “ไม่เกิด” หรือ “เสี่ยง” กับ “ไม่เสี่ยง” เป็นต้น และตัวแปรอิสระ (หรือตัวแปรทำนาย) ที่อาจมีตัวเดียวหรือหลายตัวที่เป็นได้ทั้งตัวแปรเชิงกลุ่ม (Categorical Variable) หรือตัวแปรแบบต่อเนื่อง (Continuous Variable) การวิเคราะห์การถดถอยโลจิสติก เกี่ยวข้องกับทฤษฎีความน่าจะเป็นทวินามถูกเรียกว่า Binomial Logistic Regression ถ้าตัวแปรตามเป็นพหุนามจะเรียกว่า Multinomial Logistic Regression

## ขั้นตอนการแก้ไข้ตามหลักการของ Data Science

### Step 1: Data Set



The image shows a Google Colab interface. At the top, a code cell contains four lines of Python code to connect to a file and read an Excel spreadsheet. Below the code cell, a preview of the data is shown as a table with 8 rows and 7 columns. The first column is an index from 0 to 7. The second column is 'Continent' with values 'Asia' and 'Europe'. The remaining columns are 'F2017', 'F2018', 'F2019', 'F2020', 'F2021', and 'F2022', containing numerical values.

```
1 #connect file.csv
2 filename = '/content/drive/MyDrive/Colab Notebooks/Applied Statistics/Mini Project/DataSetProject.xlsx'
3 df = pd.read_excel(filename)
4 df
```

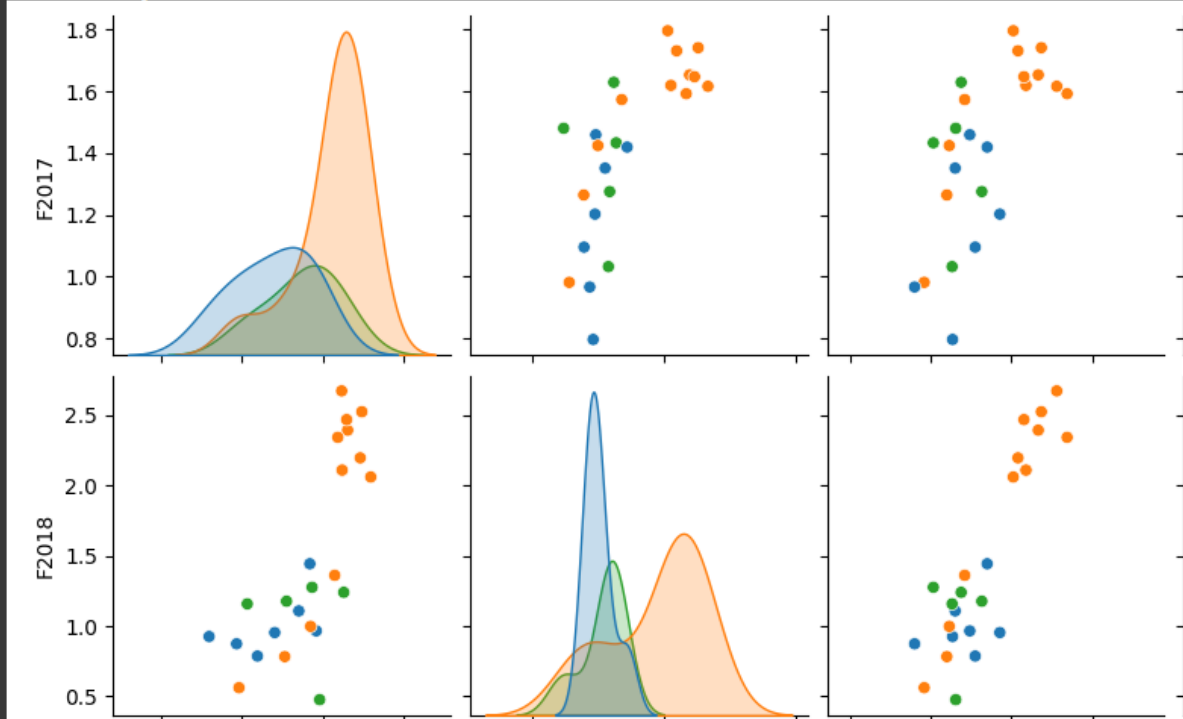
	Continent	F2017	F2018	F2019	F2020	F2021	F2022
0	Asia	1.202	0.954	1.857	1.596	1.112	1.014
1	Asia	1.351	1.108	1.304	1.728	1.898	1.383
2	Asia	0.796	0.926	1.272	1.437	1.270	1.335
3	Asia	1.459	0.966	1.485	1.551	1.762	1.659
4	Europe	1.619	2.109	2.180	2.633	1.690	2.219
5	Europe	1.741	2.524	2.370	2.315	1.395	2.498
6	Europe	0.981	0.562	0.921	1.315	0.738	0.644
7	Europe	1.424	0.998	1.231	1.373	0.929	1.911

ในบรรทัดที่ 2 เรียกไฟล์จากไดรฟ์ และ บรรทัดที่ 3 คือ อ่านไฟล์ที่เรียกมาในรูปตาราง

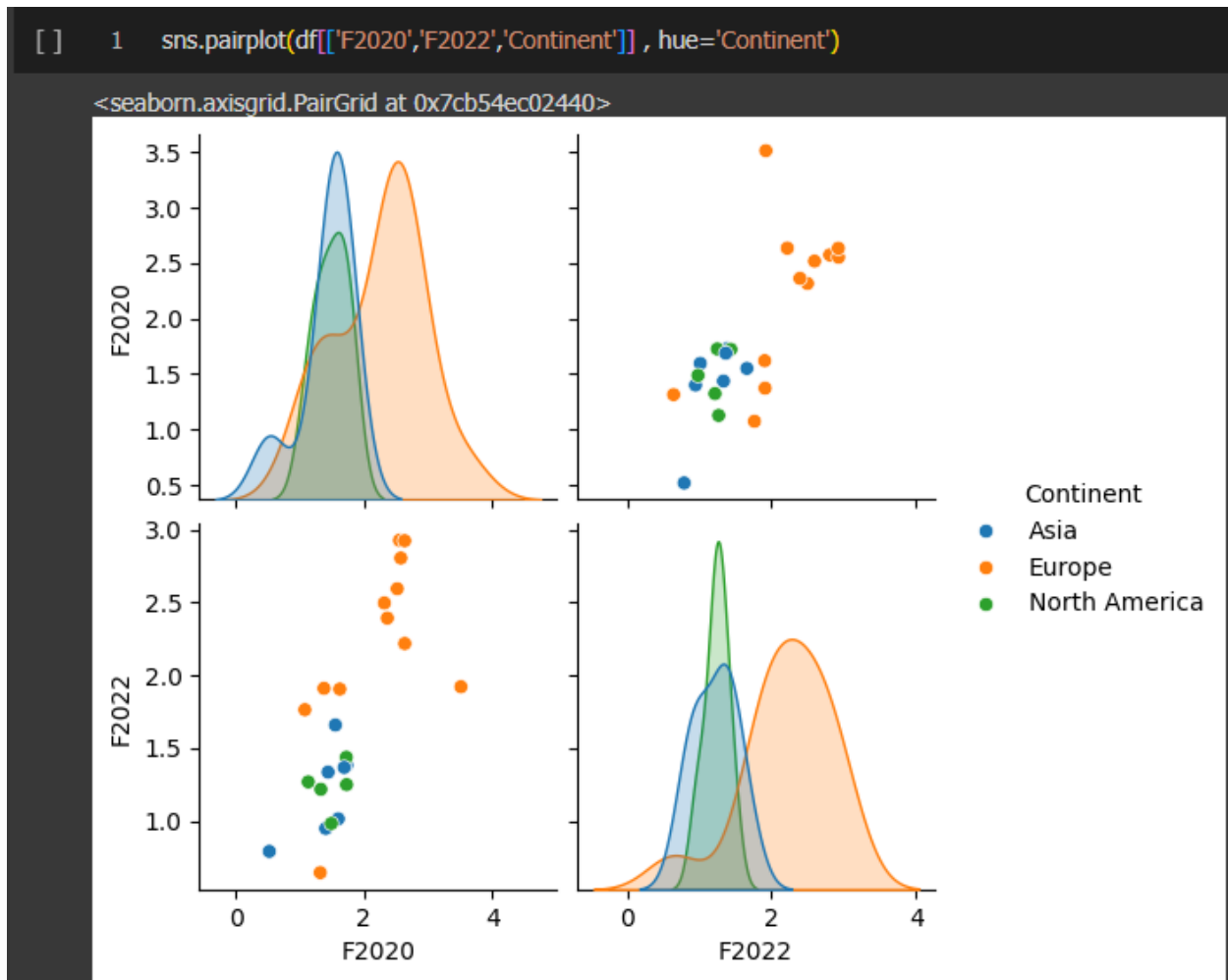
## Step 2: Data preparation

```
[ ] 1 #  
    2 sns.pairplot(df, hue='Continent')
```

<seaborn.axisgrid.PairGrid at 0x7cb58d8271c0>



นำข้อมูลทั้งหมดมาจัดรูปเป็นข้อมูลที่ใช้ได้ให้อยู่ในรูปกราฟ



เลือกข้อมูลที่เราจะใช้ในการคำนวณแล้วเรียกมาอยู่ในกราฟเดียวกัน

### Step 3: Model Building

```
[ ] 1 #Data preparation, x and y data
    2 x = df.drop(['Continent'],axis=1)
    3 y = df['Continent']

[ ] 1 # Prepare data : Train-70%, Test-30%
    2 X_train, X_test, y_train, y_test = train_test_split(x, y, test_size=0.3, random_state=1)

[ ] 1 # Train model with k=5
    2 k=5 #***
    3 knn = KNeighborsClassifier(n_neighbors=k).fit(X_train,y_train)
```

ในบล็อกที่ 1 นำข้อมูลมาพล็อตใน x และ y

ในบล็อกที่ 2 นำข้อมูลมาตรวจสอบโดยเป็นการฝึก 70% และการทดสอบ 30%

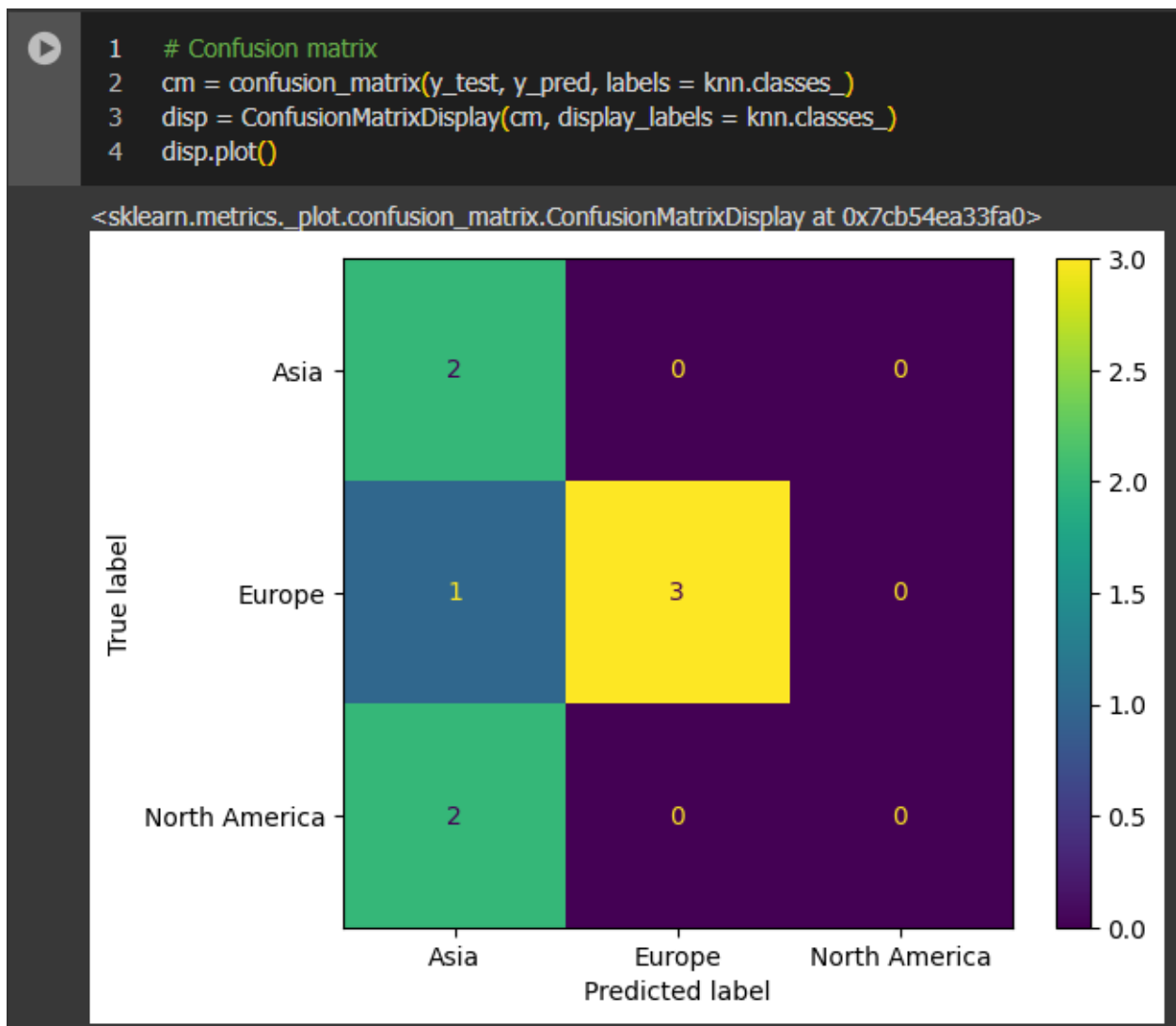
ในบล็อกที่ 3 ฝึกโมเดลโดยให้ค่า k เท่ากับ 5

#### Step 4: Result and Evaluation

```
1 # Predict the outcome for the test dataset
2 y_pred = knn.predict(X_test)
3 #y_pred
4 print(accuracy_score(y_test,y_pred))
```

0.625

ทำการตรวจเช็คความแม่นยำ



นำข้อมูลลง matrix



```
[ ] 1 print(classification_report(y_test, y_pred, target_names=knn.classes_))
```

	precision	recall	f1-score	support
Asia	0.00	0.00	0.00	2
Europe	0.50	1.00	0.67	4
North America	0.00	0.00	0.00	2
accuracy			0.50	8
macro avg	0.17	0.33	0.22	8
weighted avg	0.25	0.50	0.33	8

ทำการแสดงผลหรือคำตอบ โดยข้อมูลนี้เหมาะกับการทำงาน regression

## Step 5: Deployment

```
1 # Save model
2 import pickle
3
4 filename = 'finalized_model_knn.sav'
5 pickle.dump(knn, open(filename, 'wb'))

[ ] 1 # load model
2 filename_2 = 'finalized_model_knn.sav'
3 loaded_model = pickle.load(open(filename_2, 'rb'))
4
5 y_pred = loaded_model.predict(X_test)
6
7 print(accuracy_score(y_test, y_pred))

0.5

[ ] 1 print(loaded_model.classes_)

['Asia' 'Europe' 'North America']
```

ในบล็อกที่ 1 ทำการบันทึกโมเดลที่สร้างมา

ในบล็อกที่ 2 เรียกใช้โมเดลและป้อนความแม่นยำ

ในบล็อกที่ 3 เรียกคลาสของโมเดล

## สรุปผลและวิเคราะห์

การใช้ข้อมูลภาวะโลกร้อนเพื่อลงโมเดล KNN (K-Nearest Neighbors) สามารถช่วยในการทำนายหรือจำแนกประเภทของภาวะโลกร้อนในพื้นที่ต่าง ๆ ผลสรุปที่ได้จากการใช้ข้อมูลภาวะโลกร้อนในโมเดล KNN อาจมีดังนี้:

**การจำแนกประเภทของภาวะโลกร้อน:** โมเดล KNN สามารถช่วยในการจำแนกประเภทของภาวะโลกร้อนได้ เช่น การแยกประเภทของพื้นที่ที่มีความร้อนสูง

**การทำนายอนุกรมของภาวะโลกร้อน:** KNN สามารถใช้ในการทำนายอนุกรมของภาวะโลกร้อนในพื้นที่ต่าง ๆ ได้ เช่น การทำนายการเปลี่ยนแปลงของอุณหภูมิในอนาคต

**การตรวจจับความผิดปกติ:** โมเดล KNN สามารถช่วยในการตรวจจับความผิดปกติของภาวะโลกร้อนในพื้นที่ต่าง ๆ

**การวิเคราะห์ข้อมูลเชิงพื้นที่:** โมเดล KNN สามารถช่วยในการวิเคราะห์ข้อมูลเชิงพื้นที่ของภาวะโลกร้อนในพื้นที่ต่าง ๆ ได้ เช่น การวิเคราะห์ลักษณะการกระจายของการเปลี่ยนแปลงของอุณหภูมิในพื้นที่ เพื่อให้สามารถวางแผนการจัดการภาวะโลกร้อนให้เหมาะสมในแต่ละพื้นที่ได้

การนำข้อมูลภาวะโลกร้อนมาลงโมเดล KNN สามารถช่วยให้เราเข้าใจและวิเคราะห์ภาวะโลกร้อนในพื้นที่ต่าง ๆ ได้อย่างลึกซึ้งและมีประสิทธิภาพมากยิ่งขึ้น ทำให้สามารถวางแผนและดำเนินการป้องกันหรือลดความเสี่ยงต่อภัยพิบัติที่เกิดขึ้นจากภาวะโลกร้อนได้อย่างมีประสิทธิภาพและเป็นระบบ

## แหล่งที่มา

<https://climatedata.imf.org/pages/climatechange->

[data?fbclid=IwAR387XBsU1swoCEIHCq4153N1T1J7gmxdpz1OqaihkG-b\\_N-fZjMwk9RFec](https://climatedata.imf.org/pages/climatechange-data?fbclid=IwAR387XBsU1swoCEIHCq4153N1T1J7gmxdpz1OqaihkG-b_N-fZjMwk9RFec)

<https://neptune.ai/blog/knn-algorithm-explanation-opportunities->

[limitations?fbclid=IwAR2w2ot5kXomGth0aai5xdGAIy4hRTJs8ZNTMreimlWN6q\\_zabONj8j1bYA](https://neptune.ai/blog/knn-algorithm-explanation-opportunities-limitations?fbclid=IwAR2w2ot5kXomGth0aai5xdGAIy4hRTJs8ZNTMreimlWN6q_zabONj8j1bYA)

<https://digi.data.go.th/blog/what-is-k-nearest->

[neighbors/?fbclid=IwAR12l8xx5WPCPCBFLi5hFsw9GyUwY\\_wcvP2sr\\_n\\_uQAUj7UmsKvi\\_Pd4BMk](https://digi.data.go.th/blog/what-is-k-nearest-neighbors/?fbclid=IwAR12l8xx5WPCPCBFLi5hFsw9GyUwY_wcvP2sr_n_uQAUj7UmsKvi_Pd4BMk)

[chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://forest-](https://efaidnbmnnnibpcajpcglclefindmkaj/https://forest-)

[admin.forest.ku.ac.th/304xxx/?q=system%2Ffiles%2Fbook%2F5\(2018\)%20Logistic%20Regression.pdf&fbcl](https://forest-admin.forest.ku.ac.th/304xxx/?q=system%2Ffiles%2Fbook%2F5(2018)%20Logistic%20Regression.pdf&fbcl)

[id=IwAR20hIa4CCjVfP42vS\\_V-Ry900MvGW\\_wj8waCZ6rH07KkDwmkRB73R-6Di4](https://forest-admin.forest.ku.ac.th/304xxx/?q=system%2Ffiles%2Fbook%2F5(2018)%20Logistic%20Regression.pdf&fbclid=IwAR20hIa4CCjVfP42vS_V-Ry900MvGW_wj8waCZ6rH07KkDwmkRB73R-6Di4)