Naphon Santisukwongchot

Profile summary

Student

Thammasat business school

Business administration: Finance

Aug 2017 - May 2021

Present

Associate account manager

N-Squared eCommerce, Bangkok Oct 2021 - May 2023

Seeking a career transition into data science. Excellent understanding and proficiency of platforms for effective data analysis, including Excel, Python, R, and SQL. Strong communication, organizational and analytical skills.

Technical strengths

Business Intelligence: Looker, Power BI, Tableau

Data Analysis: Pandas, NumPy

Data Visualization: Matplotlib, Seaborn

Machine Learning: Scikit-Learn

Microsoft Office: Excel, PowerPoint, Word

Programming: Python, R, SQL

Skills

- ♦ Attention to Detail
- ♦ Collaboration
- ♦ Problem Solving
- ♦ IELTS 6 ♦ Regression , Classification, Clustering
- Business Acumen
- ♦ Critical Thinking

SpaceX Falcon 9 (I)

X.head(100)

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns

from sklearn import preprocessing
from sklearn.model_selection import train_test_split
from sklearn.model_selection import GridSearchCV
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from sklearn.tree import DecisionTreeClassifier
from sklearn.neighbors import KNeighborsClassifier
```

Orbit_ES-Orbit_GEO Orbit_GTO Orbit_HEO Orbit_ISS ... Serial_B1058 FlightNumber PayloadMass Flights Block ReusedCount 1.0 6104.959412 0.0 525.000000 1.0 1.0 0.0 0.0 ... 0.0 677.000000 1.0 0.0 0.0 0.0 1.0 ... 0.0 500.000000 0.0 ... 5.0 3170.000000 1.0 1.0 0.0 0.0 0.0 1.0 0.0

Import library

♦ Import frameworks

pandas, numpy: Data manipulation matplotlib, seaborn: Data visualization Sklearn: Machine learning

♦ Exploring data: df.head(), df.info(), df.describe

SpaceX Falcon 9 (II)

```
transform = preprocessing.StandardScaler()
X = transform.fit_transform(X)
```

Feature Selection

- ♦ 'Class' column : Target
- ⋄ All columns (except 'Class') : Feature

Standardization

♦ Ensuring all features contribute equally and improves convergence

Data splitting

♦ Split the data into training and test set, 20%

```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=2)
```

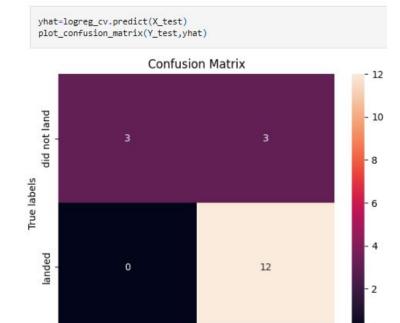
SpaceX Falcon 9 (III)

accuracy: 0.8464285714285713

Conducting model (1) -

- ♦ Conduct Logistic regression model
- ♦ GridSearchCV, cv=10 : find the best parameters
 - -C = 0.01
 - penalty: '12' (L2 regularization)
 - solver: 'lbfgs'
- ♦ Accuracy best params = 0.846

SpaceX Falcon 9 (IV)



Predicted labels

land

did not land

Confusion matrix (1)

♦ Accuracy test data = 0.83

♦ Precision = 1.00

♦ Recall = 0.50

SpaceX Falcon 9 (V)

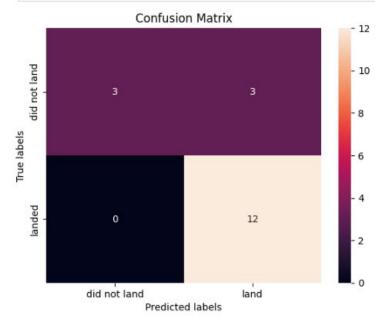
accuracy: 0.8482142857142856

Conducting model (2) -

- ♦ Create support vector machine
- ♦ GridSearchCV, svm_cv=10 : find the best parameters
 - -C := 1.00
 - -gamma = 0.032
 - kernel: 'sigmoid'
- ♦ Accuracy best params = 0.848

SpaceX Falcon 9 (VI)





Confusion matrix (2)

♦ Accuracy test data = 0.83

♦ Precision = 1.00

♦ Recall = 0.50

SpaceX Falcon 9 (VII)

Conducting model (3) -

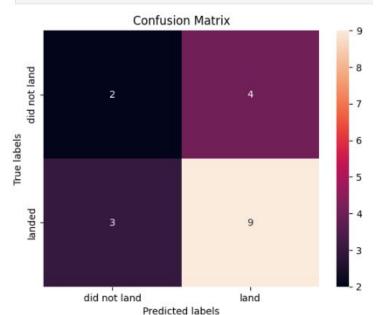
- Create decision tree classifier
- ♦ GridSearchCV, tree_cv=10 : find the best parameters
- ♦ Accuracy best params = 0.9

```
print("tuned hpyerparameters :(best parameters) ",tree_cv.best_params_)
print("accuracy :",tree_cv.best_score_)

tuned hpyerparameters :(best parameters) {'criterion': 'entropy', 'max_depth': 4, 'max_features': 'sqrt', 'min_samples_lea
f': 4, 'min_samples_split': 2, 'splitter': 'random'}
accuracy : 0.9
```

SpaceX Falcon 9 (VIII)





Confusion matrix (3)

♦ Accuracy test data = 0.61

♦ Precision = 0.40

♦ Recall = 0.33

SpaceX Falcon 9 (IX)

print("accuracy :",knn cv.best score)

accuracy: 0.8482142857142858

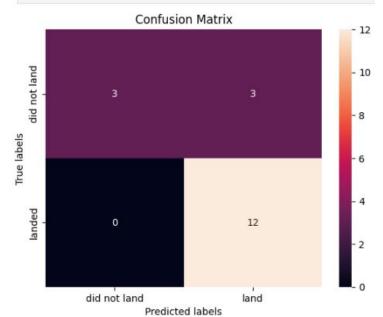
tuned hpyerparameters : (best parameters) { 'algorithm': 'auto', 'n_neighbors': 10, 'p': 1}

Conducting model (4) -

- Create k nearest neighbors classifier
- ♦ GridSearchCV, tree_cv=10 : find the best parameters
- ♦ Accuracy best params = 0.848

SpaceX Falcon 9 (X)





Confusion matrix (4)

♦ Accuracy test data = 0.83

♦ Precision = 1.00

| ♦ Recall = 0.50

Contact

Naphon Santisukwongchot

emoney euro@hotmail.com (+66)89 738 3632

https://www.linkedin.com/in/naphon1999/ https://github.com/naphon1999 https://www.datacamp.com/portfolio/naphon1999 https://drive.google.com/drive/folders/1-3x -Xmho0 3z5u3PA6VKZi2-nY90oixK?usp=sharing

Data Source

https://drive.google.com/file/d/1arTdtj3HaGGeVVVU 6CgdzYKqkC__VaFk/view?usp=drive_link

https://drive.google.com/file/d/16aoYQ-j5AAbTy445 m4_3LmysiAZRi_vD/view?usp=drive_link

Certifications & Developments

Data Science Bootcamp 10: DataRockie

Data Analyst in SQL & Python: DataCamp

Google Advanced Data Analytics : Google

IBM Data Science: IBM

Machine Learning: DeepLearning.Al