أكاديمية سدايا
SDAIA Academy

Rain in Australia

Abstract

The climate of Australia varies widely across its eight states and territories; most of the country has four seasons, with a wet and dry season in the tropical north. Rainfall in Australia is extremely varied, owing to the region's large-scale atmospheric and oceanic causes. In this project, we train classification models on the features to predict next-day rain.

2. Design

What are countries with highest degree?

What are countries with lowest degree?

Which countries have a lot of rain?

3. Data

For this project, we are using the public dataset available at Kaggle: https://www.kaggle.com.

The dataset Before cleaning 145460 rows and 23 columns.

After cleaning 56420 rows and 22 columns.

Spliting The Data to (Train, Valdation and Test) datasets

4. Algorithms

1. Problem understanding

2. Data collection

3. Data preparation

4-Build the model.

5-Train the model

6-Evaluate the model

7-choose the right model(conclussion)

Tools

- Software Platform:
  - Jupyter Notebook
- Programming Language:
  - Python
- Python Libraries:
  - Data manipulation libraries:
    - Pandas
    - Numpy
    - sklearn
  - Visualization libraries
    - Matplotlib
    - Seaborn

6. Conclusion

Models	Training score	validation score
Logistic Regression	0.850	0.853
K-Nearest Neighbor	0.902	0.839
Naive Bayes	0.802	0.809
Decision Tree	0.861	0.845
SVM	0.779	0.788
XGboost	0.858	0.858
Random Forst	0.993	0.853
Stacking	0.999	0.867

Communication ➔ Slides containing visualizations