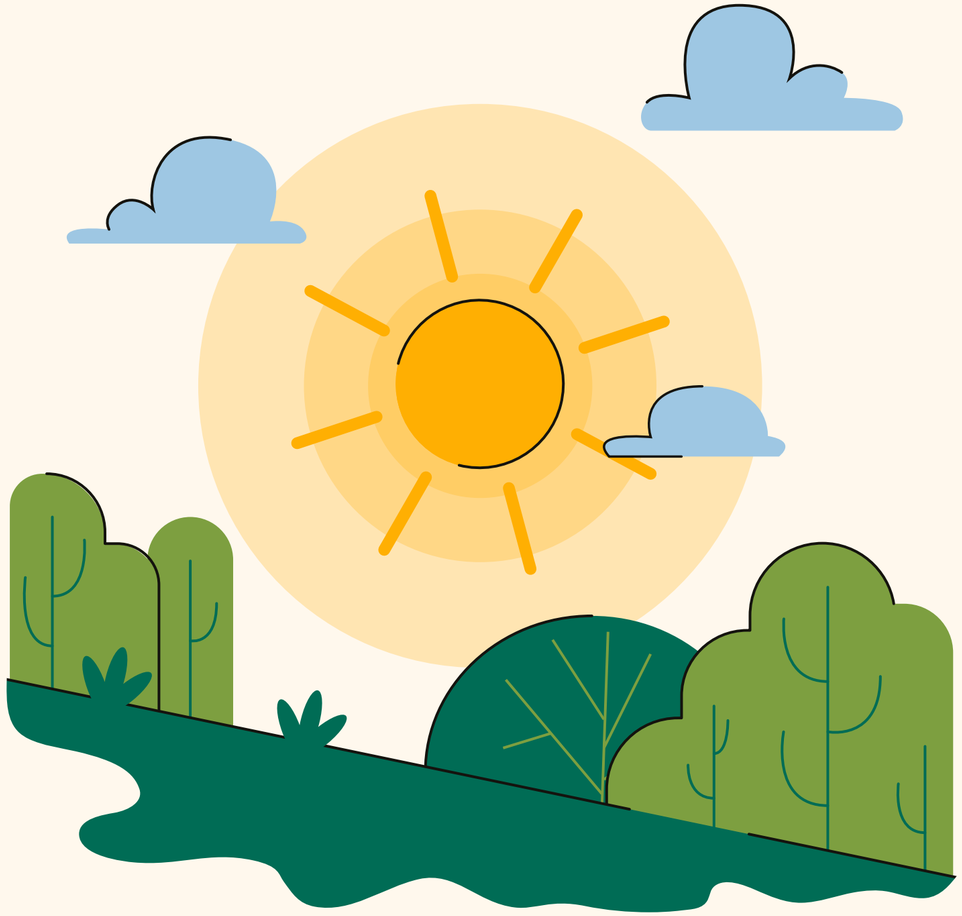


Weather classification & prediction

F20DL UG Group 7

Introduction

A

Rainfall
classification

B

Sunshine amount
prediction

C

Weather image
classification

Analysis of chosen datasets

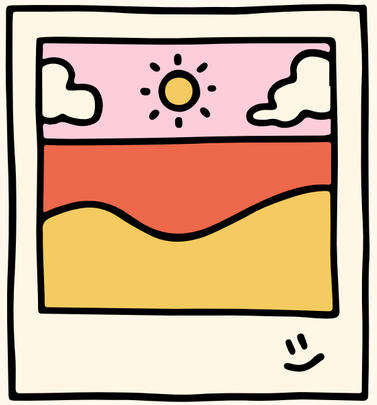
We chose 3 datasets. 2 tabular datasets and 1 image dataset.

Australia Weather Data

10 years, 22 attributes

London Weather Data

42 years, 10 attributes



Weather Image Recognition

6,862 images, 11 classes

PCA for feature reduction
+
Correlation Analysis

PREPROCESSED

Handled missing values
and class imbalances.
Oversampling + Undersampling

MODELS EXPLORED

Regression

Explored models such as:

- Linear
- Ridge
- Lasso
- Neural Networks

Classification

Explored models such as:

- Decision Trees
- kNN
- Naïve Bayes
- Neural Networks

Clustering

Explored models such as:

- K-Means
- Gaussian Mixture Models

IMPROVED THEIR PERFORMANCE BY

- Feature Engineering
- Hyperparameter Tuning
- Optimal Clustering
- Regularization in NN to prevent overfitting

RESULTS!

DATASET	BEST NEURAL NETWORK	BEST NON-NEURAL NETWORK MODEL	BEST OVERALL MODEL
AUSTRALIA WEATHER DATA	Adjusted Neural Network: Accuracy: 95.67%.	CatBoost: Accuracy: 92.67%	Adjusted Neural Network
LONDON WEATHER DATA	Deep Neural Network: RMSE: 3.47, R^2: 0.26	XGBoost: RMSE: 3.48, R^2: 0.25	Deep Neural Network
WEATHER IMAGE RECOGNITION	Five-Layer CNN – 1: Accuracy: 94.09%	XGBoost (with PCA): Accuracy: 80.46%	Five-Layer CNN - 1

In order to verify our results we carried out 5 fold cross-validation.

Conclusions

1. Preprocessing, feature engineering, and experimentation were crucial.
2. Neural Networks excel at handling complex patterns.
3. Ensemble methods like CatBoost and XGBoost provided competitive alternatives.

Advanced ML techniques significantly improve weather-related predictions and classification tasks.