ClimateWins: Project Overview and Model Selection

Data Analyst: Ashleigh Byers

November, 2024

Project Overview

Objective:

To leverage machine learning to predict climate change outcomes based on historical weather data, supporting effective planning and action.

Hypotheses:

Extreme Weather Predictions

- If machine learning models analyze patterns in wind speed, pressure, and precipitation, then they can help identify conditions that may indicate the potential for typhoon or hurricane formation in regions where these events are possible.

Solar Energy Availability

- If patterns in global radiation and cloud cover are studied, then machine learning can forecast potential high-sunshine days, which might indicate increased solar energy availability for specific regions.

Seasonal Shifts

- If machine learning analyzes long-term trends in temperature, precipitation, and snow depth, then it can identify shifts in seasonal patterns, such as longer summers or shorter winters, and predict the resulting impact on ecosystem dynamics, like prolonged snow cover or earlier plant growth cycles.

Where Our Data Comes From

Global Temperatures and Climate Data:

- Comprehensive climate records, including daily observations of temperature, wind speed, snow, and global radiation.

Source: Collected by the European Climate Assessment & Dataset (ECA&D) project.

Geography: Data from 18 weather stations located across Europe.

Date Range: Records span from the late 1800s to 2022.

Understanding Biases and Ensuring Data Accuracy

Regional and Cultural Biases

- **Cultural Differences:** Western cultures often take an analytical approach, while Eastern cultures are more holistic in climate perception.
- **Impact:** Models may not generalize well across regions, leading to inaccurate solutions.

Data Accuracy Considerations

- **Impact:** Inaccurate data can lead to faulty weather predictions, misallocation of resources, or poor emergency response planning.

Introduced Bias

Three stations lacked data on "pleasant" vs "not pleasant" and were excluded to avoid skewing results.

- **Sampling Bias**: Excluding stations without complete data could lead to an underrepresentation of certain climates or regions, distorting the model’s predictions.
- **Coverage Bias**: By omitting these stations, the model may miss specific regional patterns or characteristics that could influence the classification of "pleasant" weather.

KPIs for Model Evaluation

Accuracy:
The percentage of correct predictions made by the model.

It tells us how often the model is right overall.

Recall:
Measures how good the model is at identifying all the actual positive cases.

For example, how many storms did it successfully find, even if it made some wrong predictions?

Macro Average:
The average F1 score across all classes, treating each class equally regardless of its size.

This gives us an overall sense of the model's performance across different categories.

Precision:
Measures how often the model is correct when it predicts a positive outcome.

For example, if it predicts a storm, how often is it actually correct?

F1 Score:
A balance between precision and recall, especially useful when we want the model to be both accurate and good at identifying all relevant cases.

This is important when there’s a trade-off between precision and recall.

KNN (K-Nearest Neighbors)

How it Works

- KNN classifies data points based on the most common category among its nearest neighbors (k), where k=3 in this case.
- **Optimization:** Performance is fine-tuned by selecting the best k-value.

Performance

- Training accuracy: 88%
- Testing accuracy: 77%

Key Model Insights

- **Moderate to strong training performance**, with a noticeable drop in accuracy on the test set, indicating potential overfitting.
- **Challenge with Station 14:** Overall balanced precision, recall, and F1 scores (macro average: 0.81), but Class 14 shows lower performance.

Key Considerations

- **Strengths:** Modest accuracy, simpler model that works well with smaller, less complex datasets.
- **Limitations**: Risk of overfitting; computationally intensive with larger datasets.

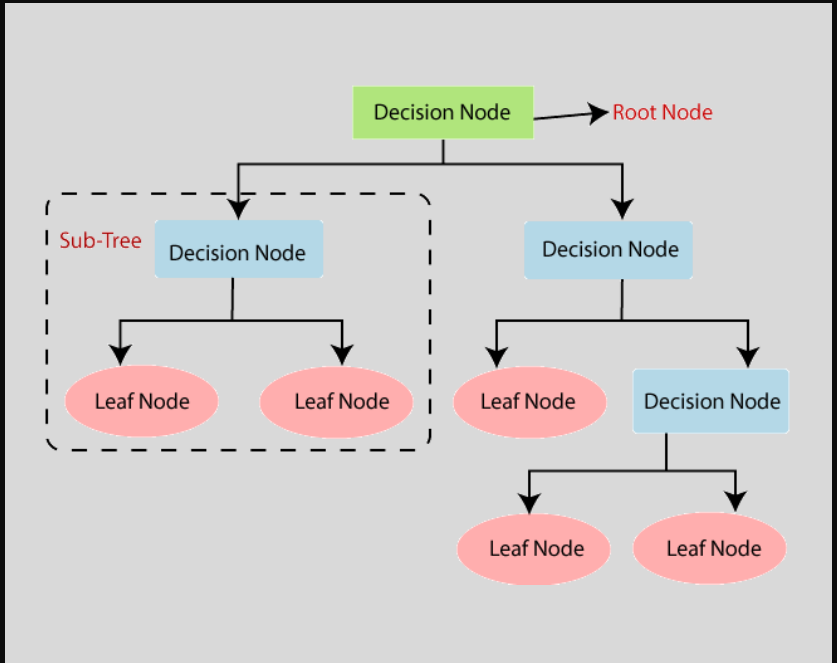
Training Set Classification Report for k=3:

	precision	recall	f1-score	support
0	0.87	0.89	0.88	3963
1	0.88	0.90	0.89	5633
2	0.87	0.91	0.89	5227
3	0.88	0.86	0.87	3144
4	0.88	0.88	0.88	3458
5	0.85	0.85	0.85	3529
6	0.87	0.84	0.85	2654
7	0.87	0.88	0.88	4519
8	0.88	0.88	0.88	3346
9	0.91	0.95	0.93	7153
10	0.85	0.85	0.85	3344
11	0.86	0.81	0.84	2544
12	0.00	0.00	0.00	0
13	0.86	0.83	0.85	2723
14	0.88	0.60	0.71	833
micro avg	0.88	0.88	0.88	52070
macro avg	0.81	0.80	0.80	52070
weighted avg	0.87	0.88	0.88	52070
samples avg	0.50	0.50	0.49	52070

Test Set Classification Report for k=3:

	precision	recall	f1-score	support
0	0.76	0.79	0.77	1701
1	0.78	0.81	0.79	2359
2	0.77	0.82	0.79	2204
3	0.75	0.73	0.74	1315
4	0.76	0.76	0.76	1476
5	0.69	0.71	0.70	1430
6	0.76	0.71	0.74	1120
7	0.76	0.79	0.77	1857
8	0.78	0.78	0.78	1420
9	0.84	0.88	0.86	3094
10	0.73	0.74	0.74	1423
11	0.69	0.67	0.68	1035
12	0.00	0.00	0.00	0
13	0.69	0.66	0.67	1171
14	0.67	0.40	0.50	341
micro avg	0.76	0.77	0.77	21946
macro avg	0.70	0.68	0.69	21946
weighted avg	0.76	0.77	0.76	21946
samples avg	0.42	0.42	0.40	21946

Decision Tree



How it Works

- The algorithm evaluates potential split points in the data to find the best separation of target classes or values at each level of the tree.
- **Optimization:** Performance is enhanced by pruning and/or setting depth limits to prevent overfitting.

Performance

- Train accuracy score: 100%
- Test accuracy score: 88%

Key Model Insights

- **Consistent Performance:** High F1-scores for most classes show that the model generalizes well.
- **Challenge with Station 14:** Lower F1-score for Station 14 (Valentia) suggests difficulty in accurately predicting this class.

Key Considerations

- **Strengths:** High accuracy and reliable performance. More efficient and faster than KNN, especially with large datasets, as it doesn't rely on distance calculations.
- **Overfitting:** The model achieves 100% accuracy on training data and 88% on test data, with a modest gap. However, pruning is unlikely to significantly improve generalization and may reduce accuracy slightly.
- **Opportunities for Improvement:** While pruning isn't necessary, exploring other algorithms may offer better ways to capture patterns, especially for underperforming classes like Station 14.

Test Set Classification Report:

	precision	recall	f1-score	support
0	0.89	0.89	0.89	1738
1	0.93	0.94	0.94	2422
2	0.92	0.93	0.92	2238
3	0.80	0.83	0.82	1349
4	0.97	0.97	0.97	1496
5	0.78	0.80	0.79	1527
6	0.82	0.84	0.83	1144
7	0.97	0.98	0.97	1945
8	0.81	0.84	0.83	1437
9	0.94	0.95	0.94	3131
10	0.86	0.89	0.88	1420
11	0.82	0.81	0.82	1116
12	0.00	0.00	0.00	0
13	0.73	0.77	0.75	1179
14	0.39	0.34	0.36	351
micro avg	0.87	0.89	0.88	22493
macro avg	0.78	0.78	0.78	22493
weighted avg	0.87	0.89	0.88	22493
samples avg	0.54	0.54	0.53	22493

ANN (Artificial Neural Network)

How ANN Works

- The model adjusts weights to minimize prediction error using gradient descent, optimizing its ability to recognize complex data patterns.
- This model has 3 layers, with 100 nodes in the first layer, 25 in the second layer, 20 in the third layer. Max Iterations: 1500

Performance

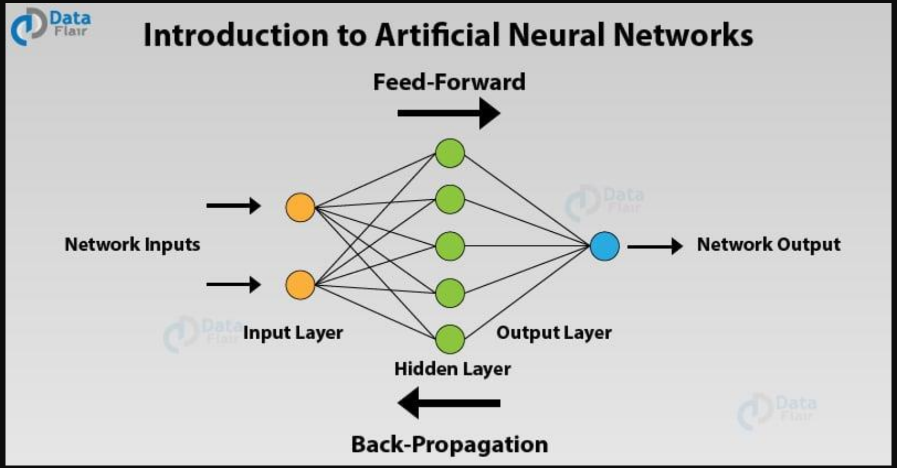
- **Training Accuracy**: 97%
- **Test Accuracy**: 89% (indicating strong generalization with minimal overfitting)

Key Model Insights

- **High F1 Scores:** Consistent, balanced recall and precision across training and test sets, indicating reliable performance.
- **Improved Predictions for Station 14:** Test accuracy for Station 14 is the highest across all models so far.

Key Considerations

- **Strengths**: High accuracy and reliable, balanced performance across all weather stations.
- **Resource Needs**: Requires more computational resources due to the complexity of the model, which may increase processing time and costs.



Training Set Classification Report:

	precision	recall	f1-score	support
0	0.96	0.99	0.98	3926
1	0.99	1.00	0.99	5570
2	0.99	0.99	0.99	5193
3	0.93	0.97	0.95	3110
4	0.97	0.96	0.96	3438
5	0.95	0.99	0.97	3432
6	0.97	0.97	0.97	2630
7	0.98	0.99	0.98	4431
8	0.93	0.96	0.95	3329
9	1.00	0.99	1.00	7116
10	0.99	0.97	0.98	3347
11	0.94	0.96	0.95	2463
12	0.00	0.00	0.00	0
13	0.93	1.00	0.96	2715
14	0.91	0.85	0.88	823
micro avg	0.97	0.98	0.97	51523
macro avg	0.90	0.91	0.90	51523
weighted avg	0.97	0.98	0.97	51523
samples avg	0.58	0.59	0.59	51523

Test Set Classification Report:

	precision	recall	f1-score	support
0	0.87	0.89	0.88	1738
1	0.93	0.93	0.93	2422
2	0.92	0.90	0.91	2238
3	0.83	0.86	0.84	1349
4	0.88	0.88	0.88	1496
5	0.86	0.91	0.88	1527
6	0.86	0.85	0.86	1144
7	0.88	0.87	0.87	1945
8	0.85	0.90	0.87	1437
9	0.95	0.94	0.95	3131
10	0.86	0.86	0.86	1420
11	0.85	0.86	0.86	1116
12	0.00	0.00	0.00	0
13	0.85	0.95	0.90	1179
14	0.70	0.74	0.72	351
micro avg	0.88	0.90	0.89	22493
macro avg	0.81	0.82	0.81	22493
weighted avg	0.88	0.90	0.89	22493
samples avg	0.52	0.53	0.52	22493

Algorithm Review

KNN (K-Nearest Neighbors):

- **Performance:** Training Accuracy: 88%, Test Accuracy: 77%
- **Key Insights:** Moderate training performance with a drop in test accuracy, suggesting potential overfitting. Station 14 (Valentia) shows lower performance.

Decision Tree:

- **Performance:** Training Accuracy: 100%, Test Accuracy: 88%
- **Key Insights:** Strong overall generalization with high F1 scores, but struggles with Station 14 (Valentia).

ANN (Artificial Neural Network):

- **Performance:** Training Accuracy: 97%, Test Accuracy: 89%
- **Key Insights:** High, balanced performance across both training and test sets, with particularly improved predictions for Station 14.

Conclusion & Next Steps

Hypotheses:

Extreme Weather Predictions:

Machine learning can identify patterns in wind speed, pressure, and precipitation that may signal typhoon or hurricane conditions.

Solar Energy Availability:

Analyzing global radiation and cloud cover can help forecast high-sunshine days for improved solar energy planning.

Seasonal Shifts:

Patterns in temperature, precipitation, and snow depth can reveal shifts in seasonal cycles, impacting ecosystems.

Chosen Method:

The ANN model is ClimateWins' best option, achieving 97% training and 89% test accuracy with minimal overfitting and strong, reliable performance.

Next Steps:

Enhance Model Accuracy and Versatility:

Tune ANN hyperparameters to reduce prediction errors further.

Explore Broader Patterns with Unsupervised Learning:

Discover uncategorized patterns and anomalies in weather data.

Develop a hybrid model pipeline to capture both short-term weather events and long-term climate trends.

Thank You

Ashleigh Byers

November, 2024