**Course name**: Intro to Artificial Intelligence

**Professor**: Saidi Elmi

**Group 2**

**Title**: Predicting weather patterns using Decision tree

**Group members**:

Krishna Sai manas karthik Bhiravajosyula

Lohithkumar Bollampally

Apurva Ravikar Chavan

Lalitha Sravani Gannavarapu

Lavanya Gurram

**Abstract**

Seattle, renowned for its distinctive weather patterns, experiences a mix of rainy and dry conditions throughout the year. Notably, the period from November to March witnesses particularly adverse weather conditions, marked by heavy rainfall and strong winds. Amongst these months, January emerges as the coldest, while November stands out as the wettest, characterized by intense rainstorms. Conversely, April to October offers a more favourable climate, conducive to travel, with pleasant and mild weather prevailing. Leveraging this rich dataset, this project utilizes decision tree classifier algorithms to predict weather patterns, aiming to provide valuable insights into Seattle's dynamic climate.

**Introduction**

The unpredictable and varied weather patterns of Seattle present both challenges and opportunities for residents, tourists, and businesses alike. Understanding and predicting these weather phenomena are crucial for effective planning and decision-making. This study aims to leverage machine learning techniques, particularly decision tree classifier algorithms, to forecast Seattle's weather patterns with greater accuracy and reliability. By analysing a comprehensive dataset encompassing historical weather data, including precipitation levels, temperature variations, and wind patterns, this research seeks to develop a robust predictive model. Such a model not only enhances our understanding of Seattle's dynamic climate but also provides valuable insights for urban planning, disaster preparedness, and tourism management. Through the application of advanced computational methods, this project endeavours to contribute to the broader discourse on weather prediction and its implications for societal resilience and sustainability.

**Methods**

In this study, we employed the decision tree classifier algorithm to predict weather patterns based on historical data collected from Seattle. The dataset was preprocessed to remove any null values and split into features (X) and the target variable (y), representing weather conditions. Utilizing the train\_test\_split function from the sklearn.model\_selection module, we divided the dataset into training and testing sets, with a split ratio of 0.65 for training and 0.35 for testing. Subsequently, we performed hyperparameter tuning to enhance the model's accuracy by identifying the optimal combination of parameters using grid search cross-validation. The parameters grid encompassed variations in maximum depth, minimum samples per leaf, and the criterion for node splitting. Upon determining the best parameters through grid search, the decision tree classifier was trained on the training data using these parameters. The model's performance was evaluated using accuracy metrics, both on the training and testing sets, to assess its predictive capability. Finally, we implemented a Flask web application to facilitate user interaction for weather prediction based on input variables such as precipitation, maximum and minimum temperature, and wind speed, leveraging the trained decision tree model to generate predictions in real-time.

**Results**

* Best Parameters: {'criterion': 'gini', 'max\_depth': 3, 'min\_samples\_leaf': 8}
* Training Accuracy: 0.8566912539515279
* Testing Accuracy: 0.8515625

Criterion 'gini' measures node impurity in decision tree splits, emphasizing balanced class distribution. Max depth of 3 limits tree complexity, mitigating overfitting. Min samples leaf of 8 ensures leaf nodes have sufficient data, preventing overfitting.

The training accuracy of the model, which represents the accuracy of the model on the training data, is approximately 85.67%. This indicates that the model correctly predicts the weather conditions for about 85.67% of the data points in the training set.

The testing accuracy of the model, which represents the accuracy of the model on unseen data (testing data), is approximately 85.16%. This suggests that the model generalizes well to new, unseen data, as the testing accuracy is close to the training accuracy.

Visualization 1:

A screenshot of a graph

Description automatically generated

Figure visualizes the count of weather categories such as ‘rain’, ‘snow’, ‘sun’, ’fog’ using a bar chart. This gives us clear understanding about weather pattern in a complete year.

**Visualization 2:**

A graph of a graph showing the temperature

Description automatically generated with medium confidence

This Figure illustrates the maximum and minimum temperature in an year in Seattle which is very crucial because the temperature determines the seasons and the amount of precipitation in the atmosphere. We find relationships in our Data using scatter plot.

**Discussion or Analysis**

The achieved testing accuracy of approximately 85.16% demonstrates the effectiveness of the decision tree classifier in accurately predicting Seattle's weather patterns. The selected parameters, including a maximum depth of 3 and a minimum of 8 samples per leaf, contribute to a well-balanced model that avoids overfitting while maintaining predictive power. Notably, the model's performance on unseen data closely mirrors its accuracy on the training set, indicating robust generalization capabilities. However, further analysis could explore additional features or alternative algorithms to potentially enhance prediction accuracy and provide more granular insights into Seattle's dynamic climate. Additionally, evaluating the model's performance under various weather conditions and geographical regions could offer valuable perspectives for real-world applications, such as urban planning and disaster preparedness.

**Code Snipets:**

**Seattleweather.py:**

**A screen shot of a computer program

Description automatically generated**

**A screen shot of a computer program

Description automatically generated**

**A computer screen shot of a program

Description automatically generated**

**App.py:**

**A screen shot of a computer program

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

**A computer screen shot of a program code

Description automatically generated**

**Home.html:**

**A screen shot of a computer

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

**A screen shot of a computer program

Description automatically generated**

**Results.html:**

**A screen shot of a computer program

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

**A screen shot of a computer program

Description automatically generated**

**Conclusion:**

In conclusion, this project demonstrates the efficacy of employing decision tree classifier algorithms to predict Seattle's weather patterns with a high degree of accuracy. By leveraging historical data and optimizing model parameters, such as criterion, max depth, and min samples leaf, we have developed a reliable predictive model capable of capturing the city's dynamic climate trends. The achieved testing accuracy of approximately 85.16% underscores the model's robustness in generalizing to unseen data, thus presenting valuable insights for various applications, including urban planning, tourism management, and disaster preparedness. Moving forward, further refinements and explorations into alternative methodologies could contribute to even more precise weather predictions, ultimately fostering resilience and sustainability in the face of evolving environmental challenges.

**References**

**Seattle weather Data set**[**:** Seattle-weather (kaggle.com )](:%20%20Seattle-weather%20(kaggle.com%20))[**https://www.kaggle.com/datasets/mahdiehhajian/seattle-weather**](https://www.kaggle.com/datasets/mahdiehhajian/seattle-weather)

**Decision tree and Applications: 1.** [Decision Trees Explained With a Practical Example – Towards AI](https://towardsai.net/p/programming/decision-trees-explained-with-a-practical-example-fe47872d3b53)