**CSCI-6401-01 BLACK SHARKS**

**WEATHER PREDICTION**

Team name: **Black Sharks**

Team members:

1. Sudheer Kumar Reddy Reddymalla [sredd1@unh.newhaven.edu](mailto:sredd1@unh.newhaven.edu) 00719922
2. Durga Guna Sekhar Reddy Savanam [dsava2@unh.newhaven.edu](mailto:dsava2@unh.newhaven.edu) 00755099
3. Vamsi Krishna Regalla [vrega1@unh.newhaven.edu](mailto:vrega1@unh.newhaven.edu) 00730353

**Research Question:**

To predict the weather based on the data previously collected using data mining techniques.

**Dataset:**

The dataset we have chosen is “Weather Prediction”. This data set is taken from Kaggle website (https://www.kaggle.com/datasets/ananthr1/weather-prediction). This dataset is about the chance of Weather and what factors effect like rain, sun, etc.,

**Exploration techniques:**

No-one knows how to handle raw data. Need to use data algorithm techniques to handle raw data to develop insights. So, we are using data exploration techniques on our dataset.

**Our dataset:**

Table

Description automatically generated

**Data Exploration:**

**Unique value count:** Examining how many distinct values there are in categorical columns is one of the first things that might be helpful during data exploration. This provides insight into the data's subject matter. A unique value count of categorical columns in the automobile’s dataset is presented here.

Chart, bar chart

Description automatically generated

**Scatter plot:** It has points that shows the relationship between weather and wind looks as:

**A picture containing scatter chart

Description automatically generated**

**Histogram:** The method determines how frequently different values appear in a column. Here, let us compare rain and sun with the wind.

**Chart, bar chart

Description automatically generated**

**Pareto analysis:** A unique method for concentrating on what matters is Pareto analysis. The Pareto 80-20 rule is a useful tool for data exploration. We may use Pareto analysis in this dataset in between rain and maximum temperatures.

A picture containing diagram

Description automatically generated

**Conclusion:**

From the above data exploration techniques, we can easily extract the insights in no time with the help of them. These techniques help a lot to understand the data much better. Exploratory data analysis has always taken a very graphical approach. Histograms, box plots, scatter plots, and many others are examples of common plots used to discover distributions, correlations, outliers, trends, and other data features

**GitHub:**