**Weather in Szeged 2006-2016**

Hourly/daily summary with temperature, pressure, wind speed and more

**Link :**

https://www.kaggle.com/datasets/budincsevity/szeged-weather

**Content:**

The CSV file includes a hourly/daily summary for [Szeged, Hungary](https://en.wikipedia.org/wiki/Szeged?oldformat=true) area, between 2006 and 2016.

**Data available in the hourly response:**

* time
* summary
* precipType ( 517 missing values )
* temperature
* apparentTemperature
* humidity
* windSpeed
* windBearing
* visibility
* loudCover
* pressure

**Algorithms to do :**

**1)Linear regression algorithm :**

Linear Regression is one of the most fundamental and widely used supervised machine learning algorithms for predictive modeling. It is used to model the relationship between a dependent variable (target) and one or more independent variables (features) by fitting a straight line to the data.

**2)KNN regression algorithm :**

K-Nearest Neighbors (KNN) is a simple and intuitive supervised machine learning algorithm used for both classification and regression tasks. It is a non-parametric, instance-based learning technique, meaning it makes predictions based on the proximity of input data points to their nearest neighbors in the training set.

Linear regression algorithm :

A red line graph with numbers

Description automatically generated

A screenshot of a computer error

Description automatically generated

KNN regression algorithm:

A red line graph with numbers

Description automatically generated

A screenshot of a computer error

Description automatically generated

**Model Comparison Analysis:**

1. **Mean Absolute Error (MAE)**:
   * Linear Regression: **0.74**
   * KNN Regression: **0.52**  
     **Observation**: KNN Regression has a lower MAE, meaning it predicts closer to the actual values on average.
2. **Mean Squared Error (MSE)**:
   * Linear Regression: **0.89**
   * KNN Regression: **0.49**  
     **Observation**: KNN Regression also has a significantly lower MSE, suggesting fewer large errors in prediction.
3. **Root Mean Squared Error (RMSE)**:
   * Linear Regression: **0.94**
   * KNN Regression: **0.70**  
     **Observation**: RMSE confirms that KNN Regression performs better by penalizing larger errors less compared to Linear Regression.
4. **R² Score**:
   * Both models: **0.99**  
     **Observation**: Both models explain 99% of the target variance, showing strong predictive ability.

**Conclusion:**

While both models are highly accurate, **KNN Regression** slightly outperforms **Linear Regression** based on the error metrics (lower MAE, MSE, and RMSE). If computational efficiency is not a concern, KNN Regression would be the better choice. However, if interpretability or speed is critical, Linear Regression is still a strong option.