Performance Evaluation of Regression Models on Precipitation Prediction

Submitted by - Mallem Anil Kumar *O180677*

Under the Supervision of -

LinuxWorld Informatics Pvt Ltd

Under the Guidance of -

Mr. Mallikarjuna Nandi

Project Guide

Abstract –

Predicting precipitation is important, especially for farming and preparing for disasters. In this project, we looked at different ways to make these predictions using Machine Learning models. We collected some weather data and ensured it was all ready to use. We tried out different models, like Linear Regression and Decision Trees, to see which one worked best. Turns out, the K-Nearest Neighbors model was the most accurate. This study not only helps us forecast rain better, but it also sets the stage for even better predictions in the future!

Purpose -

The purpose of this project is to improve the accuracy of predicting precipitation. This is crucial for making informed decisions in various fields, such as agriculture and disaster management. By evaluating different regression models and meticulously analyzing meteorological data, the project aims to find the most effective method for forecasting precipitation. The ultimate goal is to provide valuable insights into precipitation prediction and lay the groundwork for future advancements in this area.

WorkFlow –

Identifying the significance of accurate precipitation prediction in agriculture and disaster management.

1. Data Collection
2. Data Preprocessing
3. Model Selection
4. Model Application
5. Model Evaluation
6. Identification of Best Model
7. Insights and Future Directions

Input -

Output -

Model	Mean Square Error	R-squared score
Linear Regression	0.57671	0.42510
Decision Tree	0.05825	0.94193
Random Forest	0.03132	0.96877
K-Nearest Neighbors (KNN)	**0.03004**	**0.97004**
Gradient Boost Regressor	0.03173	0.96836
Support Vector Regressor (SVR)	1.02391	-0.02070
Stochastic Gradient Descent (SGD)	5.97188	-5.95301
Bayesian Ridge	0.57753	0.42428
XGBoost Regressor	0.03162	0.96847

Conclusion

Among the models tested, the K-Nearest Neighbors model emerged as the most accurate predictor. This finding provides valuable insights into improving precipitation forecasting. Additionally, this project lays a solid foundation for future advancements, particularly in the realms of feature engineering and real-time data integration.

Overall, the outcomes of this project not only contribute to better decision-making in crucial sectors but also pave the way for continued progress in the field of precipitation prediction.

Thank You !!