Store Sales Forecasting Using Machine Learning & Deep Learning

1. Introduction
This blog presents a complete sales forecasting project using Machine Learning and Deep Learning (MLP Neural Network). Deep Learning receives special emphasis, as required for evaluation. **2. Dataset Overview**
Datasets used:
- train.csv (historical sales, customers, promo)
- store.csv (store metadata)
- test.csv (sales to predict)

3. Exploratory Data Analysis
Key findings:
- Strong correlation between Sales and Customers.
- Promotional days significantly increase sales.
- Weekly and monthly seasonality is present.
- StoreType influences performance.

4. Feature Engineering
Features added:
- Year, Month, Day
- WeekOfYear, IsWeekend
- CompetitionDistance (missing filled)
- Ensured Customers=0 in test.csv

5. Machine Learning Model
A Decision Tree Regressor was used as baseline ML model. It learned simple rules from the historical data. **6. Deep Learning Model (MLP Neural Network)**
Due to TensorFlow hardware limitations, a Scikit-Learn MLPRegressor was used.

Deep Learning Architecture:
- Dense Layer: 128 neurons (ReLU)
- Dense Layer: 64 neurons (ReLU)
- Output Layer: linear
- Optimizer: Adam
- Inputs scaled with StandardScaler

MLP captured nonlinear patterns better than the ML model. Deep Learning was the best-performing model. **7. Prediction Generation**
Two final prediction files created:
- simple_predictions.csv (Decision Tree)
- dl_predictions.csv (Deep Learning)

8. Conclusion
Deep Learning outperformed traditional ML models and provided a more accurate sales forecast. The project pipeline includes EDA, feature engineering, ML + DL modeling, and final predictions. **9. References**
- Scikit-Learn Documentation
- Pandas & Matplotlib Documentation
- Kaggle Rossmann Competition