Sales Prediction Model Documentation

Objective:

Predict the sales of 50 different items at 10 different stores for 3 months using 5 years of historical sales data.

Dataset Description:

- **Train Dataset:** Contains columns - date, store, item, sales.
- **Test Dataset:** Contains columns - id, date, store, item.

Workflow:

1. **Data Preprocessing:**
    - Loaded train and test datasets using Pandas.
    - Converted date columns to datetime format.
    - Extracted time series features like day, month, year, and day of the week.
2. **Model Building:**
    - Used an LSTM (Long Short-Term Memory) neural network for time series prediction.
    - Scaled the numerical features using MinMaxScaler from sklearn.
    - Prepared data for LSTM by creating sequences of input-output pairs.
    - Built an LSTM model with input shape and appropriate activation functions.
    - Trained the model using historical sales data.
3. **Prediction and Evaluation:**
    - Prepared the test data in the required format for the trained model.
    - Made predictions using the trained LSTM model.
    - Inverse scaled the predicted values to get actual sales predictions.
    - Visualized the predicted sales data for comparison.
4. **Model Optimization:**
    - Performed hyperparameter tuning using GridSearchCV from scikit-learn.
    - Tuned parameters such as epochs, batch size, and LSTM layer parameters.
5. **Model Deployment:**
    - Saved the trained model to a file for future use without retraining.
    - Created a readme file explaining the project, dataset, and usage of the model.

Algorithms and Techniques Used:

- **Long Short-Term Memory (LSTM)** Neural Network: For time series prediction due to its ability to capture long-term dependencies.
- **MinMaxScaler:** Scaled numerical features to a range suitable for training the LSTM model.
- **GridSearchCV:** Used for hyperparameter tuning to optimize model performance.

- **Pandas:** Utilized for data loading, manipulation, and feature extraction.
- **Keras (with TensorFlow backend):** Built and trained the LSTM model.

Sales Prediction Workflow Summary

The task involved predicting sales for 50 different items across 10 stores over a span of 3 months, based on historical sales data. The process began with a thorough understanding of the problem statement.

The train and test datasets, containing date, store, item, and sales information, were loaded and preprocessed. This involved converting date columns to datetime format and extracting key time series features such as day, month, year, and day of the week.

To build the predictive model, an LSTM (Long Short-Term Memory) neural network was employed due to its efficacy in capturing temporal patterns in time series data. Numerical features were scaled using MinMaxScaler to ensure consistency and enhance model performance.

The LSTM model architecture was constructed and trained using historical sales data. Subsequently, predictions were generated for the test dataset, allowing for an assessment of future sales.

Efforts were made to optimize the model's performance through hyperparameter tuning using GridSearchCV. Various parameters such as LSTM units, epochs, batch size, and optimizer configurations were iteratively experimented with to enhance predictive accuracy.

The LSTM model demonstrated its ability to effectively forecast future sales by leveraging historical patterns and relevant features. Comprehensive preprocessing, feature engineering, and the LSTM's capacity to capture long-term dependencies in data contributed significantly to accurate predictions.

The tools and techniques utilized encompassed Pandas for data manipulation, TensorFlow and Keras for modeling, MinMaxScaler for feature scaling, and GridSearchCV for hyperparameter optimization.

The workflow underscored the importance of preprocessing, model selection, and optimization techniques in achieving robust and accurate predictions for sales forecasting.