**TITLE**

FUTURE SALES PREDICTION USING DATA SCIENCE

# INNOVATIVE TECHNIQUES DOCUMENTATION

# INTRODUCTION:

# Brief the overview of the project and the importance of data science in future sales prediction.

* Consider exploring more advanced time series forecasting techniques like Prophet or LSTM networks for improved accuracy in predicting future sales.

# FORECASTING TECHNIQUES

## LSTM NETWORKS

* Long Short-Term Memory Networks is a deep learning, sequential neural network that allows information to persist.
* It is a special type of Recurrent Neural Network which is capable of handling the vanishing gradient problem faced by RNN.
* LSTM can be implemented in Python using the keras library.

**Data Preparation:**

* Gather historical sales data, ensuring it's clean and properly structured with timestamps.
* Split the data into training and testing sets. You typically use a portion of the data for training (e.g., 70-80%) and the rest for testing to evaluate model performance.

**Feature Engineering:**

* Identify relevant features that can help the LSTM model make accurate predictions.
* This could include factors like past sales, seasonality, holidays, and external factors like economic indicators or marketing campaigns.

**Data Pre-processing:**

* Obtain a datasets (<https://www.kaggle.com//chakradharmattapalli/future-sales-prediction>) that contains relevant features for Future sales Prediction such as TV, Radio, Newspaper and Sales records.
* Normalize or scale the data to ensure that it falls within a certain range (e.g., between 0 and 1).
* This can help the model converge faster during training.
* Create sequences of data for input to the LSTM.
* For example, you might use the past N days' sales data to predict the sales for the next day.

**Model Architecture:**

* Build an LSTM model. You can use popular deep learning libraries like Tensor Flow or Py-Torch for this purpose.
* Experiment with different architectures, including the number of LSTM layers, the number of neurons in each layer, and dropout layers to prevent over fitting.

**Training:**

* Train the LSTM model using the training data.
* You can use various optimization techniques and loss functions depending on your specific problem.

**Validation and Hyper parameter Tuning:**

* Validate the model's performance on the testing data.

Fine-tune hyper parameters like learning rate, batch size, and sequence length to optimize model performance.

**Evaluation:**

* Use appropriate evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), or Root Mean Squared Error (RMSE) to assess the model's accuracy.

**Forecasting:**

* Once the model is trained and validated, you can use it to make sales forecasts for future time periods.

**Monitoring and Maintenance:**

* Continuously monitor the model's performance and retrain it periodically with new data to ensure it remains accurate over time.

**Ensemble Methods (Optional):**

* Consider combining the LSTM model with other forecasting techniques or models like ARIMA or Exponential Smoothing to improve accuracy further.

# CONCLUSION:

Keep in mind that while LSTM networks are powerful for time series forecasting, they are not a one-size-fits-all solution. The choice of model architecture and hyper parameters should be tailored to your specific sales forecasting problem and dataset. Additionally, the quality and quantity of data play a crucial role in the success of your forecasting efforts.