FUTURE SALES PREDICTION

PROJECT TITLE: Future Sales Prediction

PROBLEM STATEMENT:

Predicting sales of a company needs time series data of that company and based on that data the model can predict the future sales of that company or product.

So, in this research project we will analyze the time series sales data of a company and will predict the sales of the company for the coming quarter and for a specific product.

PROBLEM DEFINITION:

The problem definition of future sales prediction involves using historical sales data, market trends, and other relevant factors to forecast future sales for a specific product or service within a certain time frame.

This prediction is essential for businesses to make informed decisions regarding inventory management, marketing strategies and financial planning.

This primary goal is to develop a reliable model or algorithm that can accurately estimate future sales volumes helping businesses optimize their operations and resources.

PROCESS:

1.Data Source

2.Data Preprocessing

3.Feature Engineering

4.Model Selection

5.Model Training

6.Evaluation

DATA SOURCE IN FUTURE SALES PREDICTION:

In future sales prediction using data science, you can gather data from various sources, including.

1.Historical Sales Data: This is a fundamental source. It includes past sales records, which can reveal trends, seasonality, and patterns.

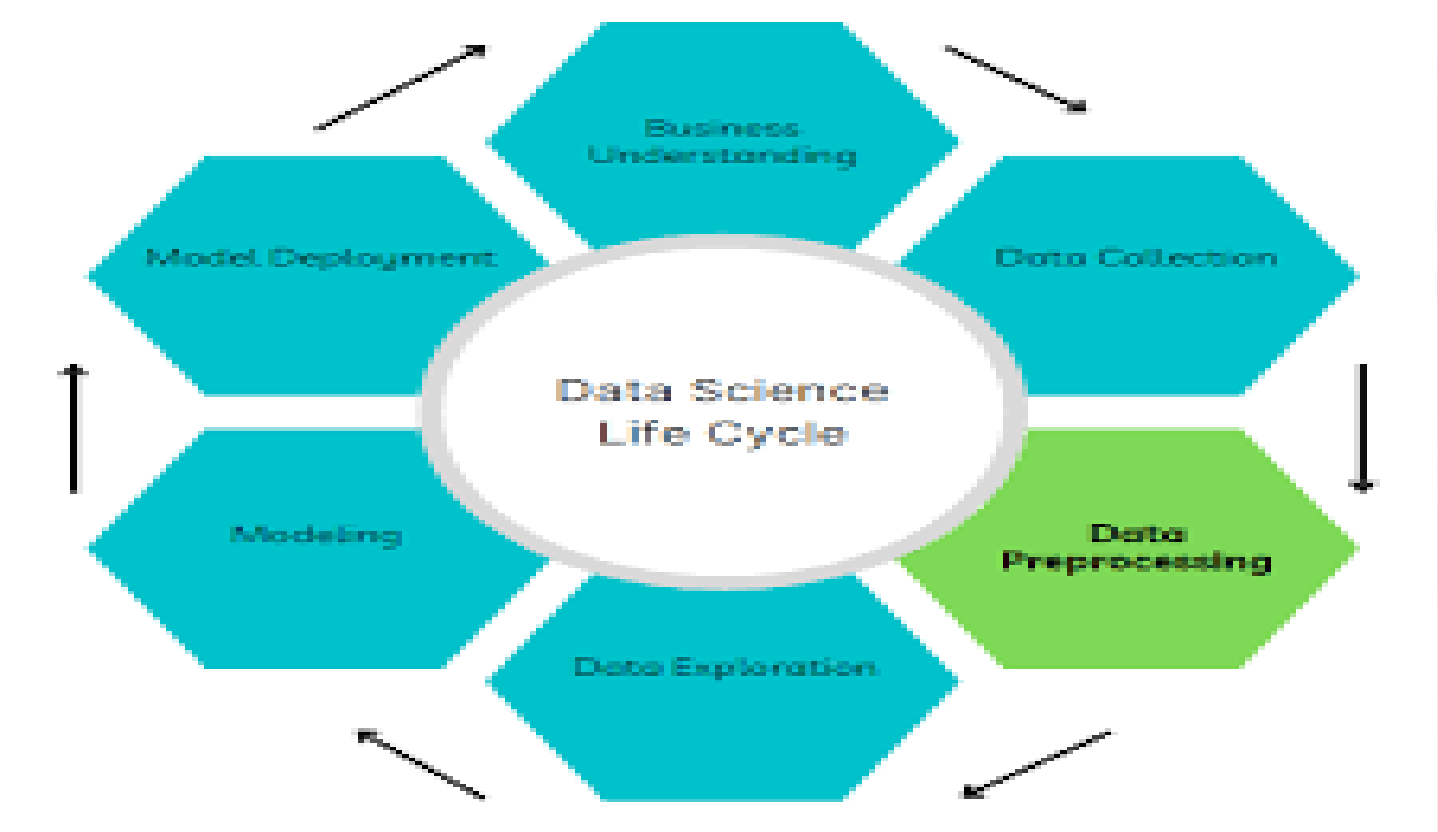
2.Market Data : Information about the overall market conditions, such as economic indicators, industry trends, and competitor data, can be valuable.

3.Customer Data: Understanding customer behaviour, preferences, and demographics can aid in predicting future sales. This can include data from CRM systems, surveys, and social media.

4.Product Data : Information about your products or services, including specifications, pricing, and any changes or updates, is crucial.

5.External Data: Utilize external data sources like weather data (for weather-dependent sales), social media sentiment, or news events that might impact sales.

DATA PREPROCESSING IN FUTURE SALES PREDICTION :



Data preprocessing is a critical step in future sales prediction using data science. Here are key steps to consider:

1.**Data collection**: Gather historical sales data, including transaction records, product information, customer data, and any other relevant variables.

2.**Data cleaning:** Remove or handle missing values, duplicate entries, and outliers to ensure data quality.

3.**Future Engineering**: Create new features or transform existing ones to extract valuable information, such as seasonality, trends, and customer segmentation.

4.**Data scaling and normalization**: Standardize numerical features to have a consistent scale, preventing certain features from dominating the model.

5.**Encoding Categorical data** : Convert categorical variables into numerical format using techniques like one-hot encoding or label encoding

FEATURE ENGINEERING :

Feature engineering, **in simple terms, is the act of converting raw observations into desired features using statistical or machine learning approaches.**

IDEAS:

1. Historical Sales Data: Include past sales data, such as daily, weekly, or monthly sales figures. You can calculate various statistics like moving averages, seasonality, and trends from this data.

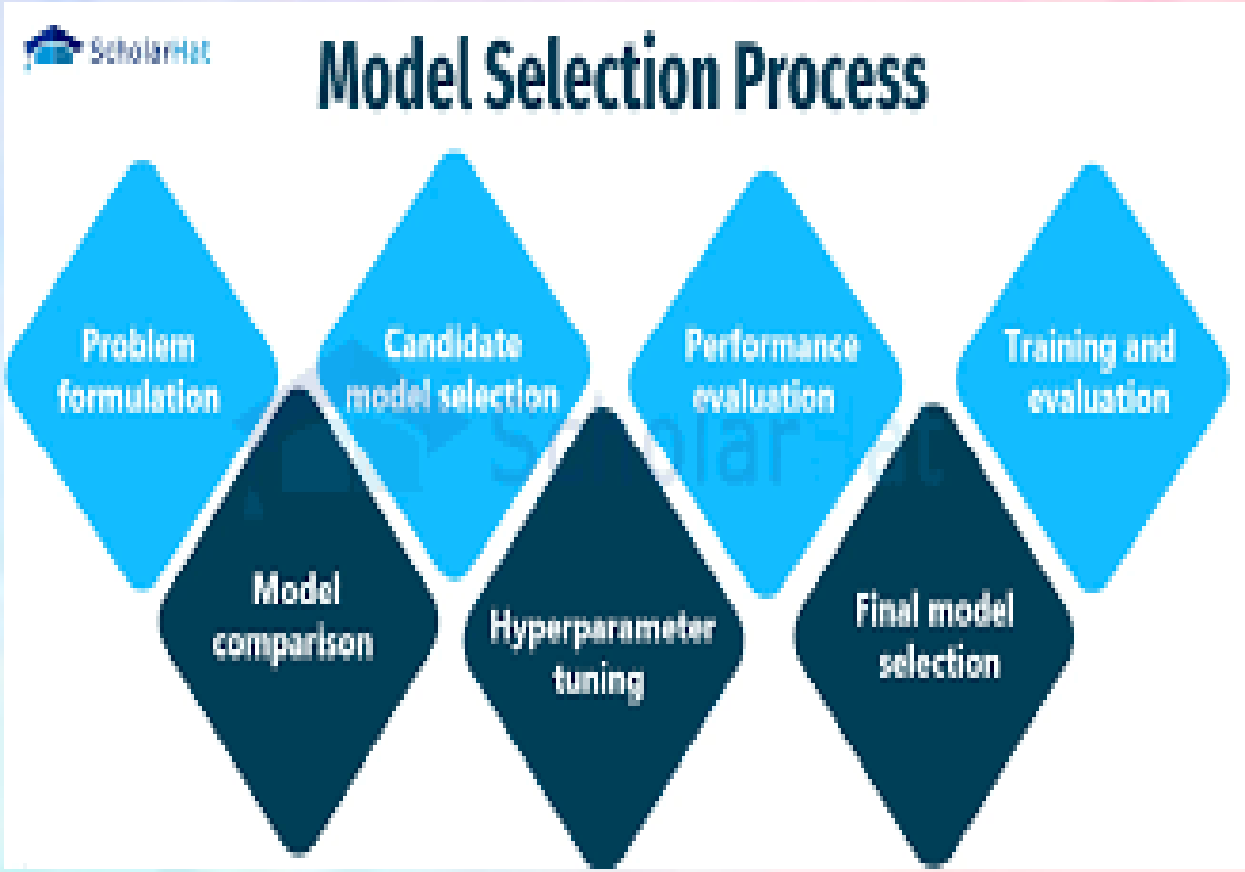
2. Calendar Features: Incorporate calendar-related features like holidays, day of the week, and month to capture seasonal trends.

3. Lagged Variables: Create lag features, representing past sales for different time periods (e.g., lag of 1 month, 3 months) to capture autocorrelation.

4. Product Attribute: Include information about the products, such as category, brand, price, and any promotions or discounts.

5. Store Information: Consider store-specific details like location, size, and foot traffic. This can help account for variations in different store locations

MODEL SELECTION:



Statistical Methods:

ARIMA (Auto Regressive Integrated Moving Average): Suitable for stationary time series data. It can capture auto-regressive and moving average patterns in the data. ARIMA models are effective when the data has a clear trend and seasonality.

State Space Models:

Kal man Filter: Suitable for dynamic systems with linear relationships. It can be used for state-space modeling and time series prediction.

Deep Learning with Transformers:

Transformer-based Models: These models, like GPT (Generative Pre-trained Transformer), have shown promise in time series forecasting tasks, especially when dealing with sequences of varying lengths and complex patterns

MODEL TRAINING:

Model training is the phase in the data science development lifecycle where practitioners try to fit the best combination of weights and bias to a machine learning algorithm to minimize a loss function over the prediction range.

Steps:

1. Data Collection: Gather historical sales data, including information on products, sales channels, time periods, and any relevant external factors like marketing campaigns or economic indicators.

2. Data Preprocessing: Clean and prepare the data by handling missing values, outliers, and converting categorical variables into numerical representations through techniques like one-hot encoding.

3. Feature Engineering: Create relevant features that can help the model better understand the data. This might include creating lag features, aggregating data by time periods, or incorporating external data sources.

4. Splitting Data: Divide your dataset into training, validation, and test sets to evaluate the model's performance accurately.

5. Selecting a Model: Choose an appropriate machine learning model for sales prediction. Common choices include linear regression, decision trees, random forests, or more advanced techniques like neural networks.

EVALUATION:


Engage Stakeholders
Describe the program
Focus Evaluation Design
Gather Credible Evidence
Justify Conclusions
Ensure Use and Share Lessons
Standards
Utility
Feasibility
Propriety
Accuracy

Definition:

Evaluation is the process of assessing or appraising something based on certain criteria or standards to determine its value, effectiveness, quality, or significance

1. Performance Evaluation: Assessing an individual's or a group's performance in a specific role or task, often used in workplaces to determine promotions, bonuses, or areas for improvement.
2. Program Evaluation: Analyzing the effectiveness and efficiency of a program, project, or intervention, usually to determine whether it achieves its intended goals and if resources are used appropriately.
3. Product Evaluation: Assessing the quality, functionality, and overall performance of a product or service to determine its suitability for a particular purpose.

1.Mean Absolut Error: ���=1�∑�=1�|��−�^�|MAE=n1∑i=1n|Yi−Y^i|

2. Mean Absolute Percentage Error (MAPE):
Formula:����=100�∑�=1�|��−�^��|��MAPE=n100∑i=1nYi|Yi−Y^i|

3.Mean Bias Deviation (MBD):

Formula: ���=1�∑�=1�(�^�−��)MBD=n1∑i=1n(Y^i−Yi)

THANKING YOU