STORE SALES PREDICTION

Charles Praveen Johnson

Abstract:

The COVID-19 pandemic has triggered significant transformations in various industries, with the fashion retail sector experiencing substantial challenges. Declining customer footfall, reduced purchasing power, and a diminished demand for clothing and fashion products have resulted in a notable decline in sales and profit.

This study aims to review and identify the challenges faced by fashion retail stores during the pandemic and the current challenges facing till now, focusing on communication issues, customer experience, and lack of technology adoption. The impact of the global health crisis on the fashion industry's retail stores could be solved by making awareness about implementing Machine Learning models among sellers. By recognizing the customer buying patterns we are going predict our specific target needs. The objective is to analyze the broader changes and challenges in the fashion retail industry in recent years, emphasizing the evolving nature of fashion, rise of competitors, technology adoption and making target prediction for growth to sustain.

1. Problem Statement:

The fashion retail industry, like many others, has faced significant challenges during COVID-19 pandemic. The pandemic has not only impacted sales and consumer behavior but has also forced the vendors to reevaluate their operational strategies, supply chain and overall technology facilities. The problem is to identify customer buying behavior patterns for making small scale store growth and profit prediction. The issues faced by fashion retail stores are multifaceted, encompassing communication problems, inadequate infrastructure, unique customer experience and a lack of technology adoption. By analyzing the data, we are going to develop a predictive model that enhances satisfaction between sellers and buyers. Through this model implementation, we came to know about the customer buying patterns, preferences, satisfaction and future needs.

2. Market/Customer/Business Need Assessment:

The fashion sector experienced a substantial downturn during the pandemic lockdown, affecting both retail stores and standalone sellers. Larger businesses also faced disruptions and struggled to adapt to the challenging circumstances. In the face of such natural phenomena, it is crucial for sellers and organizations to stay informed about current market trends for smoother operations. Awareness of ongoing market trends, product demand, quality, varieties, competitors, technology, customer buying patterns, and preferences is essential. This knowledge can empower store sellers to elevate their business profitably. Implementing our model technique can assist smaller businesses in growth and profit maximization.

3. Target Specifications and Characterization:

The designated target is sales revenue prediction. Store sellers should have a solid foundation for data storage, ensuring that customer data is recorded after every product purchase. Through the analysis of customer purchase history, buying patterns, and product preferences, we aim to make accurate predictions. Utilizing purchase history data based on specific timeframes, such as date, week, month, or year, enables us to forecast future sales and profits. This includes considering factors like date and demographics in the prediction process.

4. External Search (Information/References):

References:

1. India textile & Apparel data report.

https://www.ibef.org/industry/textiles

2. Conversion of Indian fashion retail outlets to AI.

https://www.emizentech.com/blog/convert-your-offline-store-into-ai-enabled-store.html

Datasets:

The Dataset used for this report

https://www.kaggle.com/code/ishanshrivastava28/clothes-and-accessories-sales-and-product-details/input

Some of the data information are,

```
In [1]: # Importing required libraries:

        import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
        import plotly.express as px
        import datetime
        from scipy.stats import zscore
        from sklearn.model_selection import train_test_split
        from sklearn.linear_model import LinearRegression
        from sklearn.ensemble import RandomForestRegressor
        from xgboost import XGBRegressor
        from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score, mean_absolute_percentage_error
        import joblib
        %matplotlib inline
```

```
In [2]: # importing data:

        df = pd.read_csv("C:\\Users\\charl\\Downloads\\Sales_Product_Details.csv",)
        df
```

	Date	Customer_ID	Product_ID	Quantity	Unit_Price	Sales_Revenue	Product_Description	Product_Category	Product_Line	Raw_Material	Region	Latitude	Lo
0	20210601	98	321	1	117.306016	117.306016	Cycling Jerseys	Sports	Tops	Fabrics	York	53.958332	-1
1	20210602	92	261	4	32.272403	129.089613	Casual Shirts	Menswear	Tops	Cotton	Worcester	52.192001	-2
2	20210603	92	264	1	36.193364	36.193364	Casual Shirts	Menswear	Tops	Cotton	Worcester	52.192001	-2
3	20210604	99	251	3	29.913403	89.740210	Jeans	Menswear	Trousers	Cotton	Winchester	51.063202	-1
4	20210605	66	251	1	41.843430	41.843430	Shorts	Womenswear	Trousers	Cotton	Winchester	51.063202	-1
5	20210606	97	304	3	49.887524	149.662573	Belts	Accessories	Leathers	Leather	Wells	51.209000	-2
6	20210607	45	357	2	35.416016	70.832032	Ties	Accessories	Tops	Leather	Wakefield	53.680000	-1
7	20210608	81	258	1	29.084205	29.084205	Polo Shirts	Menswear	Tops	Cotton	Wakefield	53.680000	-1
8	20210609	47	260	3	44.498077	133.494232	Tshirts	Womenswear	Tops	Cotton	Wakefield	53.680000	-1
9	20210610	24	263	3	38.497397	115.492191	Formal Shirts	Womenswear	Tops	Wool	Winchester	51.063202	-1
10	20210611	10	265	4	27.048956	108.195824	Formal Shirts	Menswear	Tops	Wool	Wakefield	53.680000	-1

```
In [6]: df.info()

        <class 'pandas.core.frame.DataFrame'>
        RangeIndex: 30 entries, 0 to 29
        Data columns (total 13 columns):
         #   Column               Non-Null Count  Dtype
        ---  ------               --------------  -----
         0   Date                 30 non-null     datetime64[ns]
         1   Customer_ID          30 non-null     int64
         2   Product_ID           30 non-null     int64
         3   Quantity             30 non-null     int64
         4   Unit_Price           30 non-null     float64
         5   Sales_Revenue        30 non-null     float64
         6   Product_Description  30 non-null     object
         7   Product_Category     30 non-null     object
         8   Product_Line         30 non-null     object
         9   Raw_Material         30 non-null     object
         10  Region               30 non-null     object
         11  Latitude             30 non-null     float64
         12  Longitude            30 non-null     float64
        dtypes: datetime64[ns](1), float64(4), int64(3), object(5)
        memory usage: 3.2+ KB
```

```
In [7]: df.isnull().sum()

Out[7]: Date                   0
        Customer_ID            0
        Product_ID             0
        Quantity               0
        Unit_Price             0
        Sales_Revenue          0
        Product_Description    0
        Product_Category       0
        Product_Line           0
        Raw_Material           0
        Region                 0
        Latitude               0
        Longitude              0
        dtype: int64
```

```
In [8]: df.describe().T
```

Out[8]:

	count	mean	std	min	25%	50%	75%	max
Customer_ID	30.0	57.400000	31.457251	9.000000	32.500000	51.000000	90.750000	99.000000
Product_ID	30.0	279.733333	24.259990	251.000000	262.250000	276.000000	287.500000	357.000000
Quantity	30.0	2.066667	1.142693	1.000000	1.000000	2.000000	3.000000	4.000000
Unit_Price	30.0	40.498759	16.742578	21.965812	32.390679	36.191911	44.341442	117.306016
Sales_Revenue	30.0	79.687953	43.008559	21.965812	36.774078	79.261696	113.761987	175.486148
Latitude	30.0	52.237571	1.449567	50.259998	51.063202	52.192001	53.680000	53.958332
Longitude	30.0	-2.270437	1.367443	-5.051000	-2.647000	-1.490000	-1.353500	-1.080278

5. Benchmarking:

In the domain of retail, the fusion of AI/ML and other technologies propels fashion and lifestyle offline stores into the area of AI-enabled establishments, reshaping the future of shopping. These stores not only streamline customer experiences but also enhance social engagement, presenting a myriad of choices to create immersive and personalized shopping environments. Despite the increasing preference for online shopping, the offline mode remains prominent; in 2022, approximately two-thirds of fashion purchases were made offline. The offline resale market in the US, valued at $101.2 billion in 2022, is projected to soar to $350 billion.

In the field of fashion retail, AI and ML bring forth automated solutions that monitor customers' shopping activities, observe their behavior, and discern their preferences and dislikes. The implementation of AI/ML models in offline retail stores is becoming increasingly feasible. For example, Azorte, a fashion retail store, leverages AI to analyze customer browsing behavior, preferences, and purchase history, tailoring product offerings, discounts, and promotions to align with individual tastes. Arvind Fashions is similarly dedicated to enhancing customer satisfaction and retention by employing AI/ML models that deliver personalized in-store shopping experiences. It's crucial, however, to strike a balance as an overreliance on technology may risk diminishing human judgment in decision-making, potentially leading to errors or misinterpretations of data.

6. Applicable Patents:

1. Using machine learning to predict retail business volume.

2. Review of artificial intelligence with retailing sector

7. Applicable Regulations:

Generally, when implementing new technologies in any industry, including fashion retail, organizations should consider various legal and ethical aspects. Here are some potential areas of concern:

1. Government laws and Regulations for small Businesses.

2. Data protection and Privacy regulations of customers.

3. The Information Technology Rules.

4. Rules against fake marketing.

5. Employment and Labor regulations.

6. Environmental acts and regulations.

8. Applicable Constraints:

The constraints for AI in fashion retail stores outlined in the provided information include:

1. Copyright concerns arise with the use of generative model techniques in fashion retail store.

2. Complexities in registering hybrid works with both human and tech elements.

3. Diminished recognition of human authorship in the face of predictive model.

4. Ethical and quality assurance concerns with potential employee oversight.

5. Lack of awareness among vendors about new technology limitations.

6. Store sellers face complex decision-making integrating generative model.

7. Competitive pressure for small businesses to adopt technology in response to changing consumer preferences.

9. Business Opportunity:

Here are some ways businesses can leverage using tech models in fashion retail stores:

Customer Segmentation:

Employ clustering algorithms to segment customers based on their behavior, demographics, and preferences. Tailor marketing strategies and promotions for different customer segments.

Customer Personalization:

Use AI to analyze customer data, including purchase history and preferences. Provide personalized product recommendations to customers based on their individual tastes and buying patterns.

Inventory Management:

Implement ML algorithms to optimize inventory levels based on historical data and current trends. Predict demand for different products, minimizing overstock or stockouts.

Visual Merchandising Optimization:

Use computer vision algorithms to analyze in-store customer behavior. Optimize product placements and visual merchandising based on customer engagement and preferences.

Dynamic Pricing:

Implement dynamic pricing models based on real-time factors such as demand, inventory levels, and competitor pricing. Adjust prices to maximize revenue and competitiveness.

Fraud Detection and Security:

Employ ML models for fraud detection in transactions, enhancing security in offline transactions. Identify unusual patterns in payment behavior to prevent fraudulent activities.

Social Media Integration:

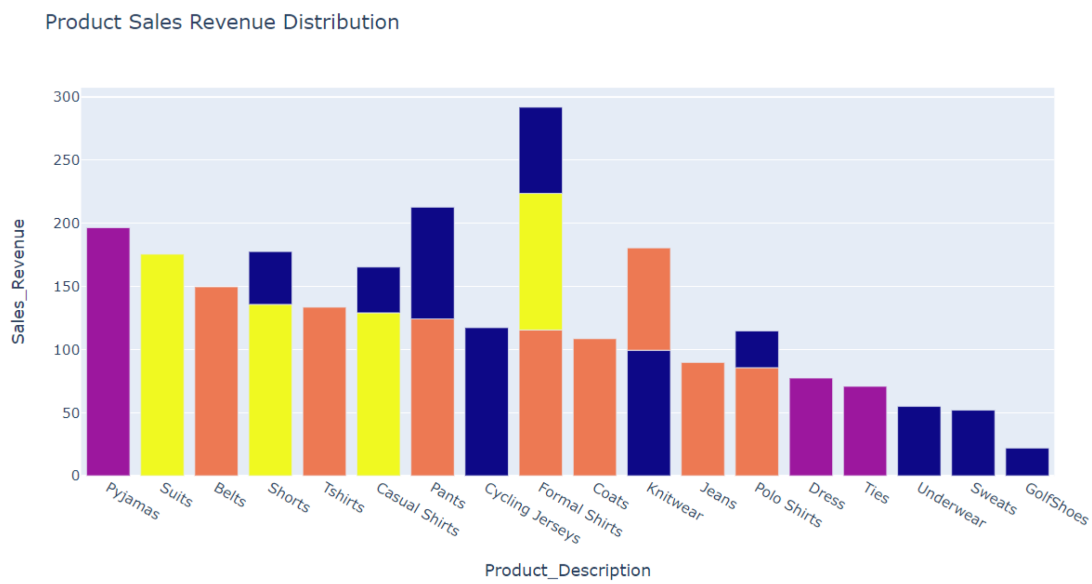
Use AI to analyze social media data for insights into current fashion trends. Align inventory and marketing strategies with popular trends identified through social media.

10. Concept Generation:

The process of predicting sales revenue through machine learning (ML), data is collected from various sources, including historical sales, customer preference, customer behavior, and relevant factors. Following data preprocessing to ensure accuracy and consistency, the most pertinent features, such as product details and customer demographics, are selected. An appropriate ML algorithm is then chosen, ranging from regression models to time series analysis. The algorithm undergoes training using historical sales data, enabling it to recognize patterns for accurate predictions. Validation and testing verify the model's performance and its ability to be applied to new data. Post-validation, the model is deployed into business applications, like sales forecasting tools, for predicting future sales. Continuous improvement is achieved through iterative refinement method, involving retraining with new data, ensuring the model's adaptability to evolving market conditions. This complete process optimizes sales prediction accuracy and effectiveness in dynamic business environments.

11. Concept Development:

The development of a predictive model involves using specific Machine Learning algorithms after collecting domain-specific data. Prior to data collection, understanding basic domain knowledge is essential to determine the necessary information to be stored. Following collection, essential data preprocessing steps are undertaken, including the analysis of null or missing values and their handling. Exploratory Data Analysis (EDA) is then performed, visualizing the data to extract valuable insights. Following that, the model is trained, integrating the target variables based on visualization reports and domain knowledge. After training, model analysis is conducted, followed by parameter tuning to enhance accuracy and predictions. The optimized model is then deployed as a predictive model app in a web application. Regular storage and updates of data are important to ensure continual improvement in future prediction results.



Dark Blue = 1 qty

purple = 2 qty

orange = 3 qty

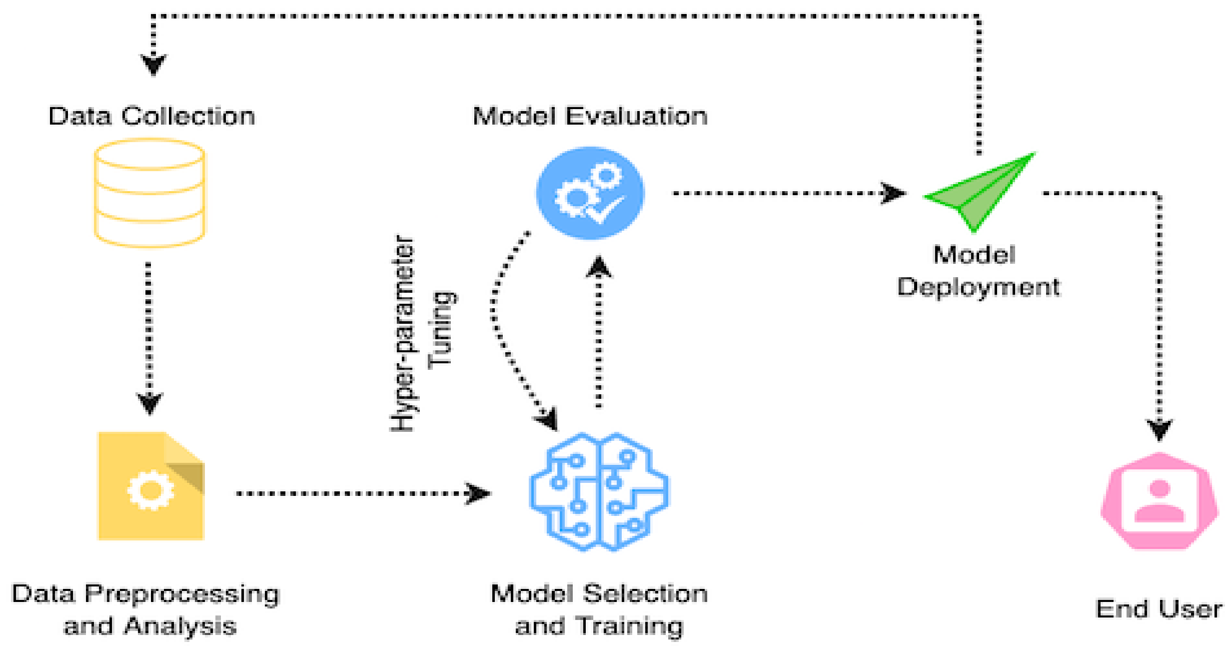
yellow = 4 qty

Here is the developed model code implementation link,

GitHub Code Link: https://github.com/CharlesGthub17/Store_Sales_Prediction

12. Final Product Prototype:

The outcome of our model is to offer a service to owners or vendors through the development of a Machine Learning model made for their small businesses or organizations. Retail vendors can leverage this model to make informed decisions by analyzing the data and generating predictions based on their specific business requirements. Our focus lies in creating web applications specifically designed for fashion retail offline stores. The deployment of the Machine Learning model into a web app is facilitated by using Python frameworks such as Django or Flask, enhancing the reliability and functionality of the web application. The installation of this web app is swift, requiring only a minimal number of days. Once implemented, it exhibits sustained performance over decades, delivering consistently high accuracy and precise predictions.



Here the prediction webapp model can offer several benefits to a fashion retail offline store:

Sales Prediction: The model can predict future sales based on historical data, enabling the store to optimize inventory levels, reduce overstock or stockouts, and enhance overall sales performance.

Customer Behavior Analysis: By analyzing past customer behavior, the model can provide insights into preferences, buying patterns, and trends. This information allows the store to give its offerings to meet customer expectations.

Personalized Marketing: Understanding customer preferences allows for targeted and personalized marketing campaigns, increasing the effectiveness of promotions and advertisements.

Trend Forecasting: The model can analyze industry trends and consumer preferences to help the store stay ahead of fashion trends, ensuring that the inventory aligns with current market demands.

Customer Retention: By understanding customer preferences, the store can implement strategies to enhance customer satisfaction, loyalty, and retention.

13. Product Details:

The development of the required model involves collecting customer clothing purchase history, storing this data, and analyzing it to identify purchase patterns and customer preferences. This data serves as the foundation for creating in-store preferences and arrangements. The process includes training the stored data, targeting specific variables for predictions, evaluating models, and refining parameters to enhance accuracy.

Before initiating model development, essential libraries need installation, such as pandas for data manipulation, visualization libraries like seaborn and matplotlib, and sklearn for model training with algorithms like regression. Post-development, the model is deployed into an app. This involves creating an environment, constructing an app model, and interpreting the ML model with front-end and back-end HTML work. This combines the use of templates and frameworks to execute and process results effectively.

Professionals need:

- Data Scientist.
- ML Engineer.
- Full Stack Developer.
- Web Developer.

14. Conclusion:

The integration of an AI/ML prediction model into fashion retail offline stores represents a framework, transforming both the customer experience and operational strategies. Through machine learning algorithms, retailers gain valuable insights into customer behaviors, preferences, and emerging market trends. This data-driven approach not only facilitates personalized in-store arrangements and inventory management but also enables accurate sales revenue predictions. The predictive model empowers retailers to make informed decisions, respond quickly to evolving consumer demands, and stay ahead in a dynamic industry. The collaboration between AI/ML technology and fashion retail creates a seamless and enriched shopping environment. Moreover, the iterative refinement process ensures adaptability, allowing retailers to navigate changing market dynamics effectively. As the industry adopts this technological transformation, the integration of AI/ML promises increased efficiency, customer satisfaction, and sustainable growth, initiating in a new era of innovation and competitiveness.