Analytica- ECommerce Market Analyzer

Introduction

The success of any business depends on its ability to understand the market and make data-driven decisions. However, the sheer volume of data available can make it challenging to extract meaningful insights. And in today's fast-paced world, online shopping has become the go-to option for most people. However, with so many sellers and products available online, it can be challenging for average buyers to make informed purchasing decisions. This is where our market and product analysis tool comes in, developed using Python, React, NLP, statistical analysis, BI, data mining, and AWS as the database. Following are the modules of my tool:

Data Collection and Storage

Our tool starts by collecting data from various websites, online marketplaces, including product listings, seller information, and customer reviews. We used python's package "**beautiful soap**" to extract the data from the online marketplaces and also found datasets from various ongoing research projects.We have used AWS as the database to store the collected data, ensuring that it is easily accessible and secure.

Natural Language Processing (NLP)

The collected data is then processed using NLP techniques to extract insights from the text data. The NLP module is used to analyze the text data and extract information such as **sentiment analysis** and **keyword extraction**. This helps us to identify the most relevant and helpful reviews for each product.

```
for review in reviews:
    tokens = word_tokenize(review)
    tokens = [token.lower() for token in tokens if token.isalpha()]
    tokens = [token for token in tokens if token not in stop_words]
    review_tokens.append(tokens)

# Count the frequency of each token across all reviews
all_tokens = [token for tokens in review_tokens for token in tokens]
fdist = FreqDist(all_tokens)

# Identify the most frequent negative words
negative_words = ["bad", "dislike", "hate", "problem", "difficulty", "issue"]
negative_tokens = [token for token in all_tokens if token in negative_words]
negative_fdist = FreqDist(negative_tokens)

# Print the most frequent negative tokens
print("Most frequent negative tokens:")
for token, freq in negative_fdist.most_common(10):
    print(f"{token}: {freq}")
```

We have used the “**nltk**” python package to tokenize and to extract keywords from the customer reviews, “**FreqDist**” function is used to find the most recurring negative words i.e. problems that customers are facing. This will help the business to find out about the shortcomings of the product before they jump into its market and it also raises awareness among the common buyers about the problems of the product and will help them in finding a better alternative.

Statistical Analysis

The statistical analysis module is used to identify trends and patterns in the data. This includes identifying customer preferences, product demand, and market trends. The module utilizes statistical models and techniques such as regression analysis, time-series analysis, and clustering.**Times Series** analysis will help us in finding and analyzing the past sale trends of the product .

Business Intelligence (BI)

The BI module is used to provide businesses with actionable insights and visualizations. The module provides interactive dashboards and reports that allow businesses and buyers to explore and analyze the data. The dashboards are generated by collecting the

findings of past-sales trends; future sales predictions and product shortcomings. These can be customized to meet the specific needs of the businesses and the buyers.

Data Mining

The data mining module is used to identify patterns and relationships in the data that can be used to predict future trends. The module utilizes a technique such as **predictive modeling**.This will help us in identifying the futures sale pattern and forecast the sale .
We have used predictive model i.e. **Auto Arima** to predict the future sales of the product.
ARIMA is a class of models that uses past values of a time series to predict its future values. It is a combination of three components: the Autoregressive (AR) component, the Integrated (I) component, and the Moving Average (MA) component. The AR component uses past values of the time series to predict future values, while the MA component uses past forecast errors to predict future values. The I component is used to remove any trends or seasonality present in the data.

Auto ARIMA works by systematically searching through different combinations of ARIMA models and selecting the model that provides the best fit for the data based on a chosen criteria such as Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC).

Product and Seller Listing

In addition to market and product analysis, our tool also provides a product and seller listing for average buyers. This listing includes the best-selling products, the most trustworthy sellers, and the customer reviews for each product. This helps buyers to make informed purchasing decisions.**Regression analysis and clustering** will help us to identify the best-selling products and the most trustworthy sellers.

Technology Stack

Our tool was developed using a combination of Python and React. Python was used for the backend development and data analysis modules, while React was used for the

frontend development and BI modules. AWS was used as the database to store the collected data.

Customization

Our tool is designed to be customized to meet the specific needs of the businesses and the buyers . The dashboards and reports can be tailored to provide insights into the markets and the products. Additionally, the statistical analysis and data mining modules can be adjusted to address your unique business challenges and the needs of the buyers.

Benefits

Using our market and product analysis tool provides a range of benefits for businesses, including:

- Understanding customer preferences and product demand.
- Identifying market trends and predicting future demand.
- Improving decision-making by providing data-driven insights.
- Enhancing the effectiveness of marketing and sales efforts.
- Making informed purchasing decisions based on customer reviews and ratings.
- Finding the best-selling products and the most trustworthy sellers.
- Saving time and effort in researching and comparing products and sellers.
- Enjoying a seamless and user-friendly online shopping experience.

Conclusion

Our market and product analysis tool provides businesses with the ability to extract meaningful insights from the vast amounts of data available. By combining Python, React, NLP, statistical analysis, BI, data mining, and AWS as the database, our tool provides businesses with actionable insights that can be used to make data-driven decisions.
Our tool also provides buyers with the ability to make informed purchasing decisions based on customer reviews, ratings, and market insights. The customizable nature of the tool ensures that it can be tailored to meet the specific needs of your business and the common buyer.

Join us in revolutionizing online shopping and enjoy a seamless and informed shopping experience and Embrace the power of our tool and unlock the potential of your business today!