Advanced Data Science Approaches in Fashion Retail Consumer Behavior

How can advanced Data Science approaches uncover patterns in Consumer Behavior and Guide Strategic Decision-Making in the Fashion Retail Sector?

Presented by:
Jessica Feto

TABLE OF CONTENTS

University of Westminster
School of Computer Science and Engineering

INTRODUCTION

Overview:

- changing dynamics in the fashion industry
- challenge in managing inventory and ensuring customer satisfaction
- preducting changing customer preferences and behaviuor

The importance of Data Mining and Machine Learning:

- navigating through complex data and gaining insightful knowledge
- trend forecasting and comprehending customer preferences
- optimal stock managaement and preventing significant losses

Customer Preferences:

- shifting preferences and new trends
- business strategies and driving sales
- crucial to maintain a competitive edge and business sustainibility

Jessica Feto

SIGNIFICANCE OF THE TOPIC

- Rapid change
- Cultural reflection
- Aligning preferences
- Strategic planning
- Data-Driven insights
- Strategic arrangement
- Predictive Analysis
- Adaptive innovation

Jessica Feto

METHODOLOGY USED

Objective:

- to utilise data mining and machine learning to understand and predict customer behaviors, satisfaction, and trends in fashion retail.

Data Collection	Data Exploration and Cleaning	Feature Engineering	Market Basket Analysis	Dimensionality Reduction	Modeling and Evaluation
Mock Fashion dataset from Kaggle	EDA to understand the data and cleaning to find patterns and fix any anomalies or missing parts	Created and Refined features to enhance prediction and insights: interaction features and sentiment polarity scores	Apriori algorithm to uncover frequent itemsets and make strategic decisions in product placements and promotions	PCA for variance maximisation and LDA for enhanced class separability in order to optimise the dataset for classification model performance	Used classification algorithms such as: Logistic Regression, KNN, SVM, Naive Bayes, Decision Tree and Random Forest.

Used metrics such as F1 Score, ROC curve, AUC. |

Jessica Feto

RESULTS

Dataset exploration:

- dataset with 29730 entries and 20 columns with a dominance of a younger consumer base with most products falling within the £20 to £80 range and receiving positive feedback.

Data Cleaning and Feature Engineering:

- imputed missing values and extracted unique values and interaction features to proved a clearer understanding of the data. Executed encoding techniques for the presence of well-known brands like Ralph Lauren and Burberry as they were more frequent.

Market Basket Analysis:

- used the Apriori algorithm, revealing significant associations like a strong connection between 'Style Attributes_Preppy' and 'Category_Dresses'. Those insights can optimise inventory choices and refine marketing strategies.

Dimensionality Reduciton:

- used PCA and LDA to simplify data and maximise class separability in order to increase the computational efficenct and improve model performance.

Classification Insights:

- Achieved over 90% accuracy in predicting customer satisfaction using Kernel SVM and Random Forest on PCA and LDA transformed data.Predicting customer interest was more challenging, with accuracies around 50%. Kernel SVM stood out in predicting seasonal shoppers with over 95% accuracy.
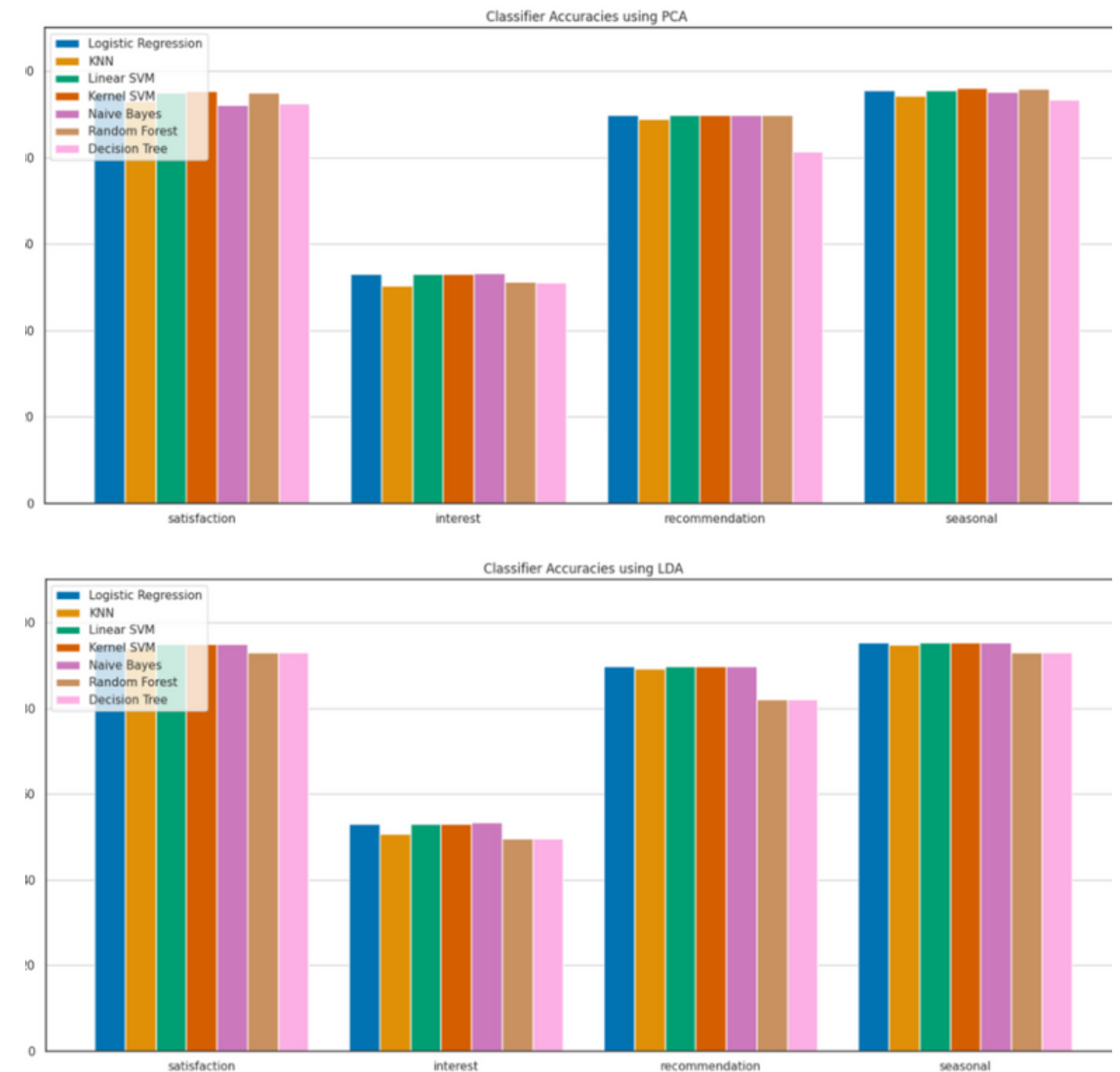
Jessica Feto

CONCLUSION

Discussion

- Dominance of younger consumer base (25-40 age bracket)
- High accuracy in predicting customer satisfaction
- Valuable insights from Market Basket Analysis for product bundling and placement
- LDA slightly superior to PCA for classification tasks due to class separation emphasis
- Challenges encountered in predicting customer interest

Limitations and Future Work:

- Dataset is a snapshot; doesn't encompass entire fashion retail landscape
- External influential factors like global events and economic conditions not considered
- Exploration of neural networks and a more holistic data approach seen as future directions

Classifier Accuracies using PCA

Classifier Accuracies using LDA

THANK YOU

Jessica Feto