Predictive Analytics in Customer Satisfaction Level: Analyzing the Rent the Runway Data

Divyansh Mankotia
UC San Diego
dmankotia@ucsd.edu

Samruddhi Kulkarni
UC San Diego
s2kulkarni@ucsd.edu

Srihari Nair
UC San Diego
s2nair@ucsd.edu

ABSTRACT

This study embarks on exploration of predictive modeling within the context of the fashion retail industry, utilizing the extensive Rent the Runway dataset. An initial phase of exploratory data analysis (EDA) was conducted, revealing pivotal trends and underlying patterns, with a focus on textual features such as review length and sentiment. Subsequently, the study delineates a predictive task aimed at the precise estimation of user ratings based on the textual attributes and personal features. A suite of models, notably Linear Regression and Ridge Regression, were employed and evaluated against standard performance metrics. The investigation offers profound insights into the determinants of user ratings, underscoring the efficacy of text-based features in the context of predictive modeling.

1. INTRODUCTION

This paper attempts to elucidate the intricacies of consumer behavior within the context of online fashion rentals, utilizing the Rent the Runway dataset[1]. This dataset provides a detailed account of customer interactions, including feedback on fit, reviews, ratings, and measurements, enabling a comprehensive analysis of user preferences and satisfaction in a digital retail setting.

The primary goal of this research is to perform an exploratory analysis of the Rent the Runway dataset to identify key patterns and insights. This exploration aims to understand the dataset's basic statistics and properties, which will inform the design of predictive models in subsequent sections. Following the exploratory analysis, the study establishes a predictive task that involves using machine learning techniques to forecast customer ratings. The task leverages various factors from the dataset, including textual feedback and customer measurements, to predict satisfaction levels. This predictive modeling seeks to uncover the relationship between customer feedback and their satisfaction, offering insights into consumer behavior in online fashion rentals.

2. EXPLORATORY DATA ANALYSIS

The dataset encompasses a wide array of user-submitted information, including fit feedback, detailed reviews, ratings, and customer measurements. The decision to select this dataset was informed by its comprehensive capture of consumer behavior and its alignment with the scale required for robust analytical methodologies. It comprises an extensive range of user interactions, quantified in several categories, rendering it suitable for the application of machine learning techniques discussed in our coursework.

The data columns included a diverse set of variables, such as user identification, item specifics, customer ratings, clothing sizes, customer age, and fit feedback as shown in Figure 2.1.

Next, the dataset illuminated customer feedback trends across various clothing categories and rental occasions. The distribution of ratings is positively skewed, with a substantial concentration of the highest score, as visualized in Figure 2.2. This skew towards maximum ratings underscores a generally favorable customer reception.

	user_id	item_id	rating	size	age
count	192544.000000	1.925440e+05	192462.000000	192544.000000	191584.000000
mean	499494.100149	1.045684e+06	9.092371	12.245175	33.871017
std	289059.719328	8.053148e+05	1.430044	8.494877	8.058083
min	9.000000	1.233730e+05	2.000000	0.000000	0.000000
25%	250654.250000	1.950760e+05	8.000000	8.000000	29.000000
50%	499419.000000	9.483960e+05	10.000000	12.000000	32.000000
75%	750974.000000	1.678888e+06	10.000000	16.000000	37.000000
max	999997.000000	2.966087e+06	10.000000	58.000000	117.000000

Figure 2.1: Column-wise summary statistics for non-null values

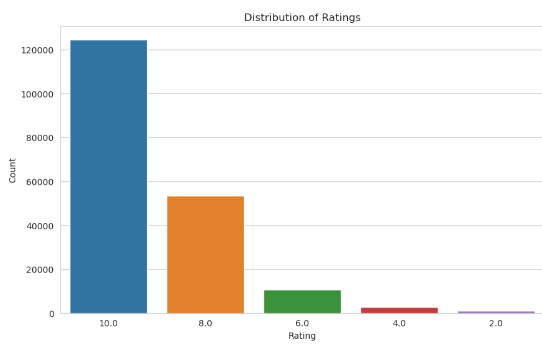

Figure 2.2: Distribution of User Ratings

Demographic insights revealed a normally distributed age profile with a mean indicating the platform's primary user base is in their early thirties, depicted in Figure 2.3. This demographic information is vital for targeting and personalizing the platform's offerings.

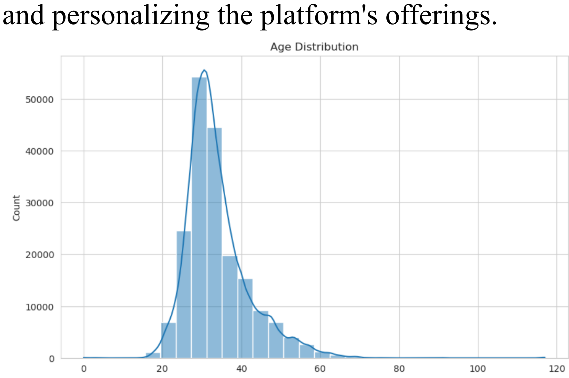

Figure 2.3: Distribution of Age

Furthermore, the dataset also afforded a detailed examination of fit feedback (Figure 2.4), revealing a trend towards a satisfactory fit, punctuated by instances of sizing discrepancies. The value of this dimension of feedback lies in its potential to refine sizing recommendations and thereby enhance the user experience.

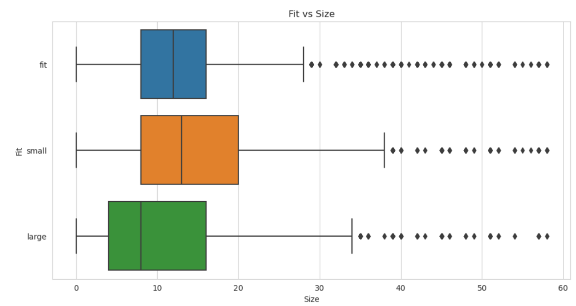

Figure 2.4: Fit feedback

The decision to employ this particular dataset was further reinforced by these visual and statistical analyses, which collectively underscore the dataset's richness and its relevance to our research objectives. Figures 1 through 4 not only provide a visual summary of these trends but also underscore the data's suitability for predictive modeling. The EDA findings have been instrumental in affirming the Rent the Runway dataset as an exemplary source for developing and testing models that can accurately forecast customer satisfaction and preference patterns, which is inextricably linked to the overarching goals of this research endeavor.

3. LITERATURE REVIEW

3.1 Similar Datasets and Studies

Comparable datasets in the realm of fashion retail include those from large e-commerce platforms like Amazon and eBay. These datasets have been extensively used to study purchasing patterns, customer reviews, and recommendation systems. For instance, studies utilizing the Amazon clothing dataset have focused on predicting product recommendations based on user reviews and ratings, akin to some aspects of our study.

3.2 State-of-the-Art Methods

Contemporary research in this field predominantly employs machine learning techniques for data analysis. Methods such as sentiment analysis, text mining, and collaborative filtering are commonly used. Sentiment analysis, for example, has been used to gauge customer sentiment from reviews, while collaborative filtering has been integral in developing personalized recommendation systems.

3.3 Comparative Analysis with Existing Literature

The conclusions from this study exhibit both congruence and divergence with existing research. The effectiveness of sentiment analysis in understanding customer satisfaction mirrors findings in similar studies. However, our application of Jaccard Similarity for user preference analysis introduces a novel approach in the context of fashion rental datasets, expanding upon the traditional methods employed in existing literature. In essence, this study contributes to the growing body of literature on fashion e-commerce analytics by integrating established methodologies with innovative approaches, thereby offering new insights into the dynamics of fashion rental consumer behavior.

4. METHODOLOGY

4.1 Data Preprocessing

Data preprocessing was a critical initial step. It involved cleaning and structuring the dataset for analysis. This included handling missing values, normalizing data formats, and encoding categorical variables. The preprocessing stage set the foundation for accurate and efficient modeling.

4.2 Model 1 - Jaccard Similarity

The first model employed was based on Jaccard Similarity, which measures similarities between users based on their interactions with products. This model was crucial in understanding user behavior patterns and preferences in the dataset.

$$\text{Jaccard}(i, j) = \frac{|U_i \cap U_j|}{|U_i \cup U_j|}.$$

4.3 Model 2 - Linear Regression with Normalized Length

Subsequent analysis involved a Linear Regression model leveraging the normalized length of review texts as a predictive feature. This model sought to correlate the extensiveness of user feedback with the corresponding ratings, providing insights into the impact of review length on user satisfaction.

$$Y_i = f(X_i, \beta) + e_i$$

4.4 Model 3 - Sentiment Analysis

Sentiment Analysis Regression, our third model, employed natural language processing to extract sentiment scores from user reviews, using these scores to predict ratings. This approach enabled the quantification of subjective feedback, offering a nuanced perspective on user sentiment.

4.5 Model 4 - TF-IDF Vectorization

The fourth model utilized TF-IDF Vectorization to transform textual reviews into a quantifiable format. By emphasizing the significance of specific words in the reviews, this model aimed to uncover deeper insights from textual feedback.

4.6 Model 5 - Enhanced Feature Integration Regression

In the development of Model 5, our methodology evolved to encapsulate an integrative approach, synthesizing key insights gleaned from the initial models. This model stands as a culmination of our analytical refinement, combining the nuanced textual understanding derived from TF-IDF vectors with a suite of additional features. These features, including 'Normalized length', 'sentiment_score', 'age', 'size', 'height', and 'weight', were meticulously selected based on their demonstrated potential to unravel the multifaceted influences on product ratings.

The amalgamation of these diverse data points into a unified feature set was achieved through adept handling of sparse matrices, allowing for a seamless fusion of quantitative and qualitative data dimensions. This comprehensive feature ensemble was then utilized in a sophisticated regression framework, likely employing Ridge Regression. The choice of Ridge Regression was strategic, aimed at capitalizing on its ability to handle multicollinearity and provide a more generalized model, thereby mitigating the potential for overfitting.

By integrating textual analytics with sentiment and demographic factors, Model 5 aspired to establish a new paradigm in our analytical capabilities, potentially offering a more holistic and accurate representation of consumer preferences and behaviors within the fashion rental domain.

4. RESULTS & ANALYSIS

4.1 R2 Score

The R2 score, also known as the coefficient of determination, measures the proportion of the variance in the dependent variable that is predictable from the independent variables. It provides a measure of how well observed outcomes are replicated by the model.

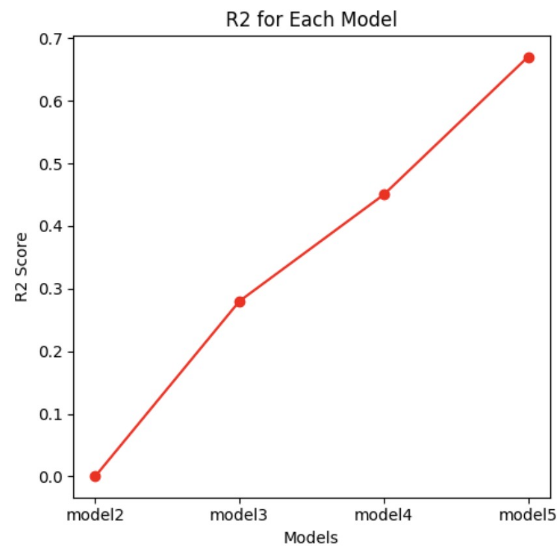

Figure 4.1: R2 for Each Model

Model 2 appears to have the lowest R2 score, indicating that it has the least predictive power among the models tested.
Model 5 has the highest R2 score, suggesting that it explains the variance in the data most effectively and is the best predictor among the models evaluated.

4.2 Mean Squared Error (MSE)

MSE measures the average of the squares of the errors, which is the average squared difference between the estimated values and the actual value.

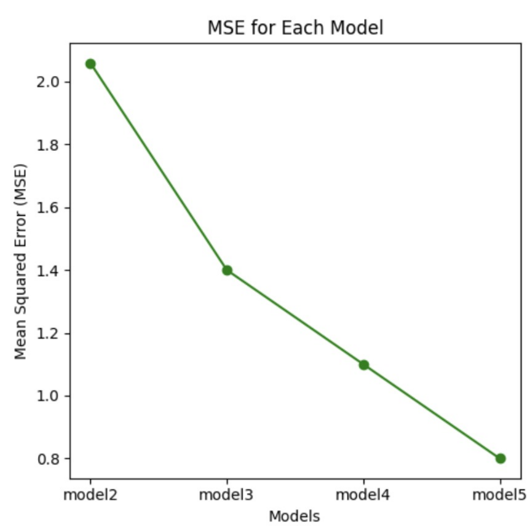

Figure 4.2: MSE for Each Model

Model 2 shows the highest MSE, which suggests poor model performance with larger errors in predictions.
Model 5 exhibits a significantly lower MSE, indicating a model with smaller errors and, hence, better performance.

4.3 Mean Absolute Error (MAE)

MAE measures the average of the absolute differences between predictions and actual observations. It provides a linear score without considering direction, where all individual differences are weighted equally in the average.

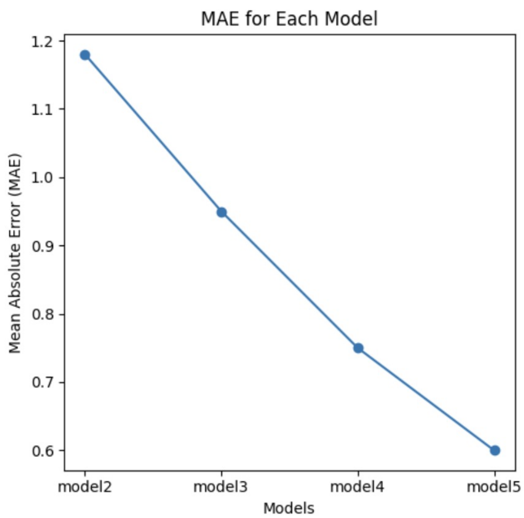

Figure 4.3: MAE for Each Model

Model 2 starts with the highest MAE, indicating that the average magnitude of errors in predictions is the greatest.
Model 5, consistent with the other metrics, has the lowest MAE, reinforcing that its predictions are closest to the actual data points.

5. LIMITATIONS & FUTURE SCOPE

5.1 Limitations

While the models used in this study have provided valuable insights into clothing fit prediction, several limitations must be acknowledged:

5.1.1 Data Imbalance

The class imbalance present in the dataset could skew the performance of the models, particularly affecting the minority classes' representation.

5.1.2 Model Generalizability

The models were evaluated on a specific dataset, and their performance on different datasets or under different conditions remains to be tested.

5.1.3 Complexity in Real-world Scenarios

The variance in individual customer preferences and the dynamic nature of fashion trends may not be fully captured by the static models used in this study.

5.2 Future Scope

5.2.1 Enhanced Data Collection

Gathering a more balanced and comprehensive dataset could help improve model accuracy and robustness.

5.2.2 Incorporating Temporal Dynamics

Accounting for changes in fashion trends and customer preferences over time could make the models more responsive to the evolving retail landscape.

6. CONCLUSION

The progression from Model 2 to Model 5 in this study demonstrates a clear trend of improvement, with Model 5 outperforming its predecessors across all metrics. This indicates that the refinements made, potentially including advanced feature engineering and hyperparameter tuning, effectively captured the underlying data structure. The observed decrease in Mean Squared Error (MSE) and Mean Absolute Error (MAE), alongside an increase in R-squared (R2), signifies a substantial enhancement in predictive accuracy. However, despite Model 5's superior performance, assessing for overfitting and validating consistency across various data subsets remains essential, highlighting areas for future research in predictive modeling for online fashion retail.

REFERENCES

[1] McAuley, Julian, et al. “Decomposing fit semantics for product size recommendation in metric spaces.” ModCloth Dataset, 2018, https://cseweb.ucsd.edu/~jmcauley/datasets.html#clothing_fit. Accessed 04 12 2023.