Customer Satisfaction Prediction for Delivery Service

Richard Han
rickyhan24@gmail.com

Abstract

In this project, I perform exploratory data analysis on a dataset consisting of customer survey data to identify the most significant attributes of the data. Secondly, I apply several classification models and identify the best-performing models in regard to accuracy. The delivery service company can identify which features of their service or product are most significant for customer satisfaction in order to focus on projects that matter. The company can also gauge how well it is doing in various aspects of their service or product.

1 Problem Statement

The task is to build a model to predict customer satisfaction given survey data about the customer's experience using the delivery service such as whether the order was delivered on time. I need to identify which attributes are significant for prediction and to use those attributes to build a model that will classify a given customer as happy or not.

2 Data

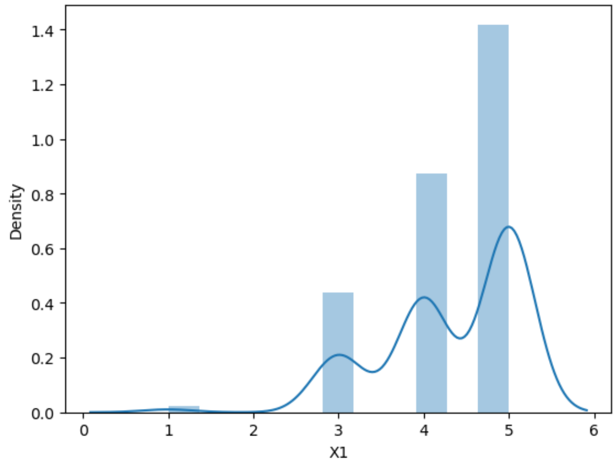
The data consists of 126 rows of completed customer surveys and 7 columns. The first column indicates whether the customer is happy or not. There are 6 attributes X1,..., X6 for the remaining 6 columns. Each attribute is rated 1 to 5 by the customer. X1 indicates the extent to which the order was delivered on time, X2 indicates the extent to which the contents of the order were as expected, X3 indicates the extent to which the customer ordered everything they wanted to order, X4 indicates the extent to which the customer paid a good price for the order, X5 indicates the extent to which the customer was satisfied with the courier, and X6 indicates the extent to which the app made it easy to order.

There are 57 unhappy customers and 69 happy customers.

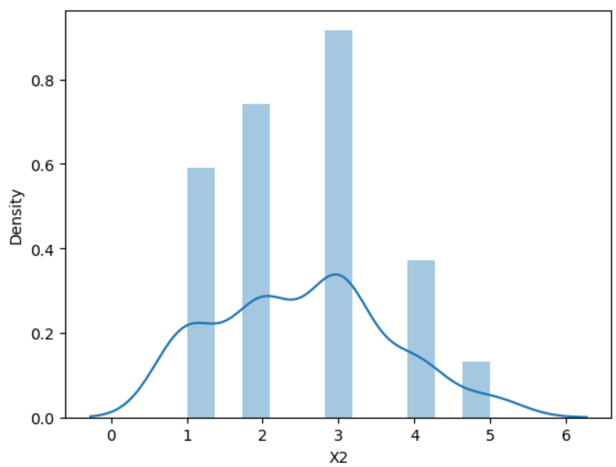
3 Exploratory Data Analysis

3.1 Results and Evaluation

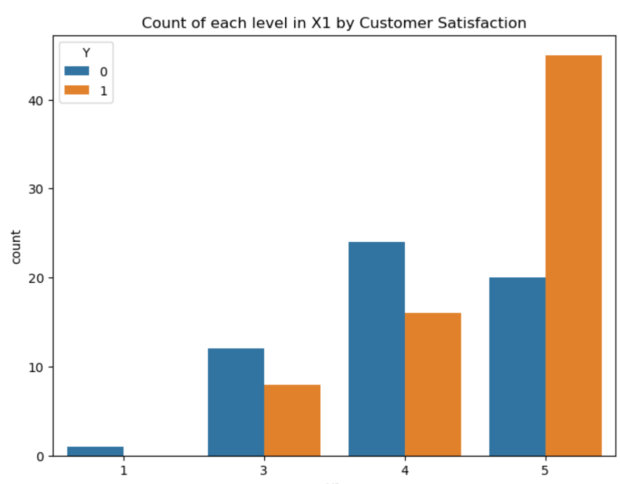
To get a sense of how the company is doing on each of the attributes, we can plot histograms for each attribute. For example, here is the histogram for X1:



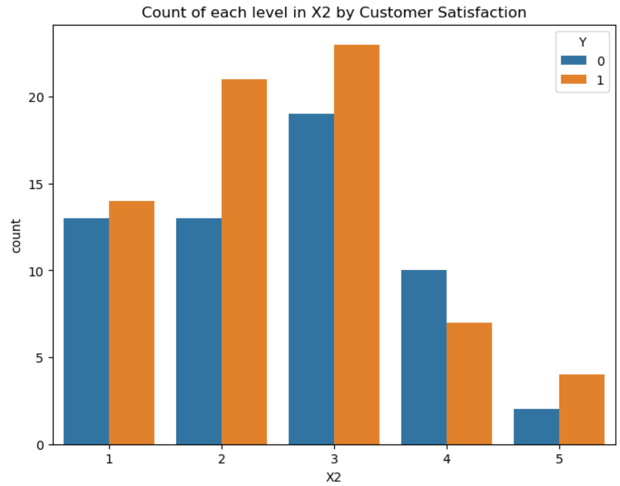
The company seems to be doing okay on delivery time since most of the ratings are 3, 4,or 5. Similarly, the company appears to be doing okay on X3, X4, X5 and X6. On the other hand, the company seems to be struggling with X2, the content of orders being as expected, since there are a lot of 1's and 2's:



To get a sense of how each attribute might affect the satisfaction of a customer, we can look at bar plots for each attribute. For instance, here's the bar plot for X1:

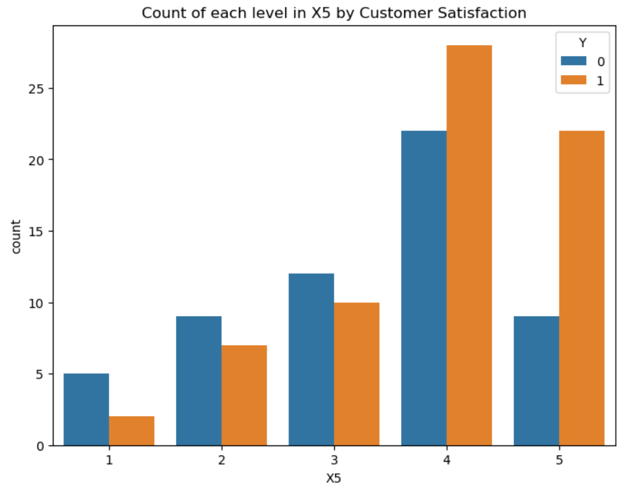


As the delivery time score goes up, so does the number of satisfied customers; for the score of 5, there are about twice as many happy customers as opposed to unhappy customers. So X1 seems to be a significant attribute. On the other hand, X2–the extent to which the order is as expected–appears to not make much of a difference in satisfaction since the numbers of happy vs unhappy customers for each score appear commensurate.

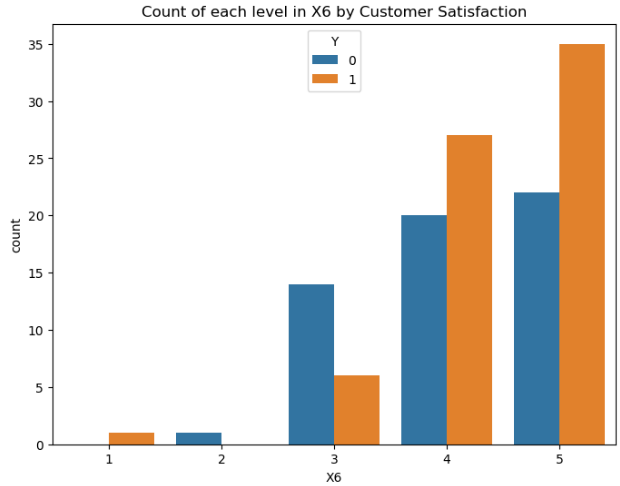


Similarly, X3 and X4 appear to make little difference to satisfaction.

For X5, as the score increases, both the number of happy and the number of unhappy customers generally increase until we reach the score of 5, in which case the number of happy customers greatly outweighs the number of unhappy customers; so X5–the satisfaction with the courier–seems to make a difference.



For X6, as the score increases, both the number of happy and the number of unhappy customers increase, but generally the higher X6 is, the more satisfied customers there are.



Finally, I also performed a Mann-Whitney U test on each attribute and confirmed the significance of X1, X5, and X6.

4 Classification

4.1 Results and Evaluation

I applied multiple classification models using all of the attributes.

Here is a summary table of the performance of all the models I considered. For each model, it gives the test accuracy:

Model	Test accuracy
Logistic regression	61.54%
KNN	53.85%
Linear SVM	57.69%
Kernel SVM	57.69%
Naive Bayes	57.69%
Decision Tree	50%
Random Forest	65.38%
AdaBoost	61.54%
Gradient Boosting	61.54%
LASSO	61.54%

Random Forest performed the best at 65.38%.

To improve the accuracy, I re-applied all the models using only the attributes X1, X5, and X6–which are the ones found to be significant in our exploratory data analysis section. Here are the results:

Model	Test accuracy
Logistic regression	61.54%
KNN	57.69%
Linear SVM	61.54%
Kernel SVM	76.92%
Naive Bayes	57.69%
Decision Tree	73.08%
Random Forest	73.08%
AdaBoost	65.38%
Gradient Boosting	76.92%
LASSO	61.54%

The highest accuracy achieved was 76.92% by both Kernel SVM and Gradient Boosting. We were able to achieve higher accuracy using only the significant attributes.

5 Conclusion

Attributes X1, X5, and X6 were found to be significant in predicting customer satisfaction. X1 measures the extent to which the order is delivered on time, X5 measures the extent of satisfaction with the courier, and X6 measures the extent to which the app made ordering easy. The company was doing okay on all attributes with the exception of X2 which measures the extent to which the contents of the order was as expected; this is something the company can try to improve on.

The company may be able to get rid of survey questions corresponding to attributes X2, X3, and X4 in future surveys since these don't seem significant in predicting satisfaction.

Finally, I was able to achieve a test accuracy of 76.92% using only the significant attributes X1, X5, and X6.