Group Coursework Submission Form

Specialist Masters Programme

Please list all names of group members: (Surname, first name) 1. Dubey, Naveen 2. Shetty, Rakshith 3. Semaleesan, Sneha	4. Saha, Soumyajit 5. Parashar, Tridev **GROUP NUMBER:** **9**
MSc in: Business Analytics	
Module Code: SMM636	
Module Title: Machine Learning	
Lecturer: Rui Zhu	**Submission Date: 03/03/2023**

Declaration:

By submitting this work, we declare that this work is entirely our own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the coursework instructions and any other relevant programme and module documentation. In submitting this work we acknowledge that we have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. We also acknowledge that this work will be subject to a variety of checks for academic misconduct.

We acknowledge that work submitted late without a granted extension will be subject to penalties, as outlined in the Programme Handbook. Penalties will be applied for a maximum of five days lateness, after which a mark of zero will be awarded.

Marker's Comments (if not being marked on-line):

Deduction for Late Submission:

Final Mark: **%**

Table of Contents:

Seattle Police Department Customer Satisfaction

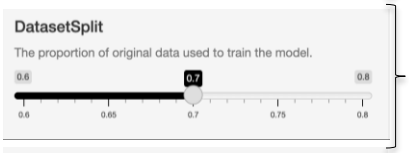
Task of the dataset:

Through this project, we intend to gauge the key characteristics that lead to a satisfactory 911 phone call made to the Seattle Police Department. The 2020 dataset[1] includes survey responses gathered from randomly selected callers who provided a combined overall score between 1 to 5 along with individual scores for each feature. This use-case could be replicated to understand and improve customer experience.

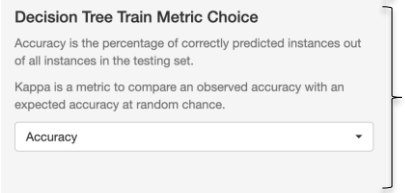
Guide to use the App:

For better understanding of the models used, we have created an interactive application using Shiny that helps the user understand how different parameters contribute to the accuracy of the model.

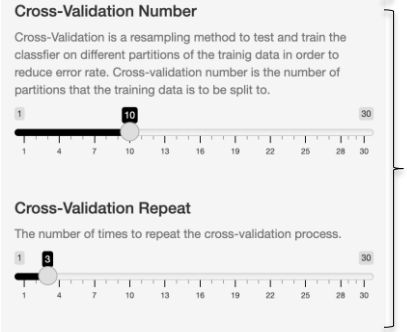
- The “Create” buttons have been provided to refresh the results after changing any of the Model Features (filters):



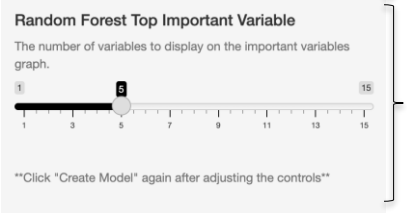
- With the “DatasetSplit,” the user can see how the performance of the model changes based on how much of the data is used to train the model.



- Since the decision tree uses Complexity Parameter (CP) to determine the minimum improvement required at each node, the relationship between CP and the accuracy/ Kappa could be observed using the “Metric Choice” dropdown.



- In order to avoid overfitting in the model, cross validation is the resampling method we have chosen with sliders for options to choose the number of partitions and repetitions.

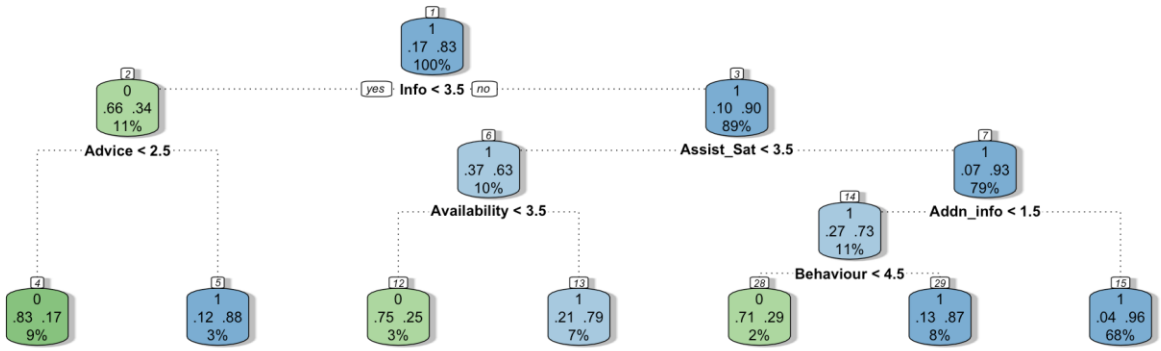


- The important variables of the model is a key feature in this project (explained below), the number of which could be determined using the last slider.

[1] SPD 911 Customer Satisfaction Survey Data

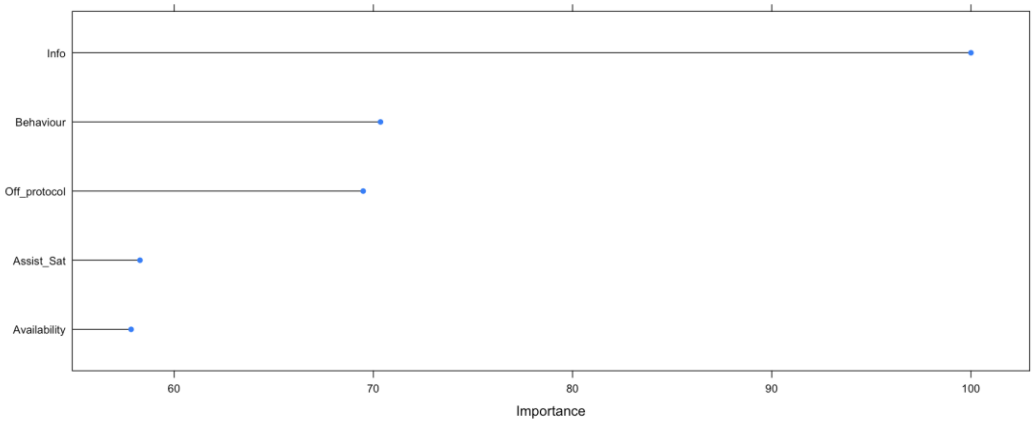
Analysis:

Decision trees help in choosing between alternative courses of action by branching decisions which end in outcomes, creating a tree-like structure. A decision tree would split the features using a cost function which could be further optimised by removing the branches with irrelevant features (pruning).



Overall Accuracy: 84.87%

The decision tree considers the optimum features at each split and does not consider the model as a whole. Hence, a more robust solution is the Random Forest as it considers all possible iterations of decision trees with each tree randomly choosing features at each split (wherein the number of features in each split remains the same across). Since it is an optimised model, it does not require to be pruned and shows the influential features (i.e., Variable Importance) that help in accurately predicting the outcomes. Since the variables are positively correlated (as shown in the correlation matrix), therefore the top variables show the features that make a satisfactory phone call. The following table shows the top 5 important features for a helpful 911 call that the department could focus on to improve customer experience:



Both the models show an accuracy of 84-85% and we can observe that- adequate information (*Info*), good Behaviour, Satisfaction from assistance provided (*Assist_sat*) and on-time Availability, has led to satisfactory 911 calls.

Appendix:

1. Data Lexicon:

Name (in dataset)	Description
ID	Unique ID
CallDate	CallDate and timestamp
Overall_sat	Overall, how satisfied are you with this experience with the Seattle Police Department -from calling 9-1-1 on to all contacts you had with the Police Department as a result of that call?
Assist_Sat	How satisfied are you with the assistance provided by the 9-1-1 operator over the telephone?
Info	The officer provided you the information you needed.
Off_protocol	The officer clearly explained procedures and requirements
Behaviour	The officer was professional and courteous.
Procedure	The officer told you what would happen next.
Advice	The officer gave you tips on preventing future crimes.
Addn_info	The officer provided information about other crimes or problems in your area.
Availability	The Department is available when you need them.
Resource	The Department is a good resource for information about preventing crime.
Dept_protocol	The Department clearly explains its procedures and requirements.
Dept_beh	Department personnel are professional and courteous.
Safety	The Department focuses on the public safety issues that concern you.
Feedback	Would you like to have someone from the Police Department contact you so you can discuss further the service you received after your call to 9- 1-1?

2. Correlation Matrix