Codeicon

PROBLEM STATEMENT

Business Case: Tender Acceptance Hit or Miss Prediction

Business Case Overview: Transportation is a major part of any Food Manufacturing Company. Generally, companies use 3rd Party Logistic Partners to transport raw materials, finished goods and its like from one location to another. These locations can be plants, warehouses, store etc. The path from one location to another is called a lane. Our company has a contract with one Primary carrier and multiple secondary carriers for every lane. Whenever there is a need for transportation at a particular lane, a communication is sent by the company to Primary carrier. The Primary carrier can then accept or reject the communication. If the Primary carrier does not accept the communication, it is then cascaded to the secondary carriers. The same process is repeated with secondary carriers. In-case none of the secondary carriers accept the communication, it is then broadcasted to all the carriers. The tender is then awarded to the first carrier who accepts. When the primary or secondary carrier does not accept, there is higher cost incurred by the company as the transportation rates are not contracted. The given dataset contains events on Tender Acceptance for the year 2018-2019. The Hit and Miss column tells us whether it was the primary partner who fulfilled the transportation needs.

Created for illustration purpose

Transportation Team

3P Logistic Partner/s

Tender Event Communication

PROBLEM STATEMENT (Cont.)

Questions:

- Can we predict tender for which shipment will result in a hit-or-miss?
- Calculate the F1 score for the final model.
- What are the key drivers of hit-or-miss?
- Any other valuable insights you can draw from the data?

Predictive Analytics Guidelines:

Kindly follow the guidelines below to build a predictive model on the given dataset:

- Use the Data Dictionary to get more information about the data columns.
- Use the 'data_flag' column to identify the training and testing data, marked as '**train**' and '**test**' respectively.
- The final submission file should contain all the **test** '***shipment number***' sorted in the **ascending order** and the corresponding predicted '**hit_or_miss**' column. Please refer to the example submission file 'Sample Submission.xlsx'

Join Conditions:

1. tender_metric_data.csv to accept_event.csv => *shipment number = shipment number*
2. tender_metric_data.csv to new_tender_event.csv => *shipment number = shipment number*
3. tender_metric_data.csv to target_pcnt_identifier.csv => *concat(source_location_cd,'_',destination_location_cd,'_',award_carrier_cd,'_',creation_ts) = join_tag*

F1 Score Calculation:

$$F_1 = 2 \cdot \frac{\mathrm{precision} \cdot \mathrm{recall}}{\mathrm{precision} + \mathrm{recall}} = \frac{\mathrm{TP}}{\mathrm{TP} + \frac{1}{2}(\mathrm{FP} + \mathrm{FN})}$$

TP = number of true positives
FP = number of false positives
FN = number of false negatives

DATA

Location:

1. tender_metric_data.csv
2. accept_event.csv
3. new_tender_event.csv
4. target_pcnt_identifier.csv
5. Data Dictionary.xlsx
6. Sample Submission.xlsx

JUDGING CRITERIA

Pre-Screening:

1. The number of exact matches of the final output predicted by the participant vs the actual output.
2. The F1-score of the model
3. Quiz scores

Final Selection:

1. Business case explanation
2. Insights gathered from the data
3. Approach used to get to the final output

GENERAL INSTRUCTIONS

1. All common instructions and files will be shared in the General Channel of the Codeicon Team. Keep checking the channel for announcements and instructions.
2. A quiz will be circulated at 12pm on the General Channel. The quiz link will be open for 1hr. Only the SPOC of a Team should submit the quiz.
3. Participants can find the Data Files, Data Dictionary and Sample Submission file in the Files tab on MS Teams.
4. Use the Teams Private Chat, named 'Team – *Team_name*', to post your queries.
5. The final submission should be in the form of a CSV to be uploaded to each team's folder located on the Files tab on MS Teams.