

Hackathon Problem Statement: Credit Card Call Centre Call Volume Forecast

Category: AIML

Assumptions:

1. Participants can clean data and categorize the credit card types accurately for time series forecast
2. Participants include the current exogeneous variables
3. The submission should be test period call volume and judgement should be on MAPE score

Use Cases:

Participants are expected to forecast the call volume accurately to reduce the abandoned calls and establish better staffing needs to manage all queries.

Project Description:

Wells Fargo has call center operation teams based out of US and Manila, where Customer Support Officers (CSO) will be responding to client queries for credit cards and all other commercial products.

There are 3 types of credit cards:

1. Commercial Card
2. Retail Card
3. Student Card

Commercial Card is used by clients' employees to pay for expenses related to their employee's official travel bookings, supplier payments etc. Whenever employees of clients face any issues on card usage like:

- A. Card decline – Card can be declined due to fraud activity, password or chip issues
- B. Temporary card block – Customer has availed a feature of not using the card temporarily, but he swiped the card unknowingly.
- C. Statements, payment history, etc.

Employees call the service desk for resolving these issues either by calls (Inbound), emails or chats, where CSO's team handle significant volume of calls on a daily basis to address these issues 24*7 entire year. **Participants are expected to forecast the call volume accurately to reduce the abandoned calls and establish better staffing needs to manage all queries.**

Problem Statement:

I

From the 3 observations for each timestamp, identify the observation corresponding to Commercial Card for each timestamp. Create a time series dataset of only entries corresponding to Commercial card.

II

Build a quantitative model to forecast the number of calls offered to support Capacity plan/Workforce Management to manage the work load of call handling. All teams are expected to use the available data to come up with a logical and scientifically derived solution for the problem to build a robust capacity plan to predict future call volume. Teams are allowed and encouraged to download and use any relevant open source data that might impact the incoming call volume.

The teams are free to use any framework, technique or models to come up with their solution to the following questions:

- 1) Short term forecast - Forecast call volume for the **coming week** b/w 5/1/2021 and 5/7/2021 (Call volume for each day in the week)
- 2) Long term forecast - Forecast call volume for the **next 3 months – May, June and July 2021** (Call volume accumulated at monthly level)

Data

This is a time series data, with 4 fields, “Date”, “No_of_calls_from_commercial_card”, “No_of_calls_from_retail_card”, “No_of_calls_from_student_card”. Series starts from **01-January-2017** and ends at **30-April-2021**.

For the period 01-January-2017 to 31-December-2017, the observations corresponding to the 3 types of cards are correct. For the remaining time period, the 3 observations have been mixed up. The participants must correctly sort the 3 entries first and then proceed to forecast the call volume for commercial cards.

Key Objectives

Round 1

Solution accuracy will be evaluated based on Mean Absolute Percentage Error (MAPE) of predictions for both problems. Final evaluation metric will be calculated by averaging the MAPE of both solutions.

MAPE calculation

Round 2

In this round, finalists from the first round, will present their solution to the leadership. Teams are expected to present the following details:

- 1) Solution design and methodology
 - 2) Innovation
-