EFG – SU Competition Guidlelines

Competition Use Case:

Every business wants to optimize for customers remaining loyal to their services. They typically calculate their churn rate and strive to minimize it through enhancing their services, products, and marketing campaigns. EFG provides seamless financial services, one of which is a stock market trading application called EFG Hermes ONE. In this use case, we aim to predict customer churn for retail Egyptian clients trading on the ONE app, to retain them before actual churn, and explore internal or external factors to reduce churn in the brokerage market for EFG. EFG is considered your client, and you are attempting to sell them a final product based on the data available.

Data Description:

You have two dummy data clients_data, orders_data. Clients' data mainly consists of clients' information and demographics, while orders' data mainly comprises the clients' orders on the app.

Some features explanation:

For client's data:

1. Each client has a unique Client ID, but can have multiple Account IDs. This means that a client can have multiple accounts.
2. Individual clients are retail clients.
3. Risk rate is the risk assigned to the client upon entry based on their KYC (Know Your Customer data). It doesn't have to do something with their actual behavior.
4. Hermes Securities Brokerage (HSB) is the company we are interested in.
5. Is dormant: clients who are not trading for more than one year are labelled as dormant.
6. A Closed account is an account that is voluntarily closed by the client.
7. A suspended client/account is an client/account that was suspended by EFG due to compliance issues or suspicious actions etc.
8. OpenDate is the account opening date.

For the orders data:

1. Security ID: security is the stock itself. Each stock has a unique ID.
2. Order Via: We have multiple channels: Online (Hermes ONE), Call Center etc.
3. Is Complete: An Order can be completed or not based on supply and demand on the stock. (I want 10 stocks @ price x, but the market didn't meet x)
4. Is Canceled: the client can cancel the order or it can be cancelled from the app.
5. Expire Date: The expiration date of the order for which it will be cancelled if it doesn't meet the customers expectations.

6. Execution: The order might be executed or partially executed or not executed (I want 10 stocks @ price x but 5 only are available)
7. Sector Name: It is the sector of the stock traded.
8. Price & Quantity are obvious.

Tips & Tricks:

1. A good data science project starts with defining a good objective that you can extract from the Competition Use Case paragraph.
2. You will be frustrated with the quality of data. This is normal in a real-life example project.
3. You will need to define what customer churn means. Are these the clients who have not traded for 6 months, 1 year, or 2 years etc.? Are they the ones who closed their accounts or have suspended accounts? Define your label wisely.
4. Always ask yourself why I'm doing this? For example, why am I visualizing this data?
5. It's a great opportunity to learn some coding skills. This is the main direction currently in industries. This is how you can deal with big data!
6. A good data science project is one that creates value at the end of the day. If we have the best churn prediction model with no business solutions, then we have done nothing. After completing your data science part, consider how you will help your client achieve lower churn rates and what corrective actions they can take with their clients.
7. Many questions arise when you explore the data. You can approach me anytime, and I'll answer your questions. (mrahman@efghldg.com)

Evaluation

- You will be evaluated on how you approach the problem, including defining the business objective, cleansing the data, exploratory data analysis, extracting useful insights, exploring the power of machine learning model is a major plus, presentation of results, selling your product, and most importantly, the added value and recommendations to the business. It's subjective, but this is how we do business!
- All efforts are appreciated.

Deliverables

You will submit the following:

- A final PowerPoint presentation convincing the panel about your solution & recommendations.
- A Zipped file contains the project (preprocessing approaches, visualizations, models, etc.)