Welcome to the Mistplay Data Science Challenge!

Today you embark on one of perhaps the most daring adventures you will ever experience. A challenge not even the greatest warriors of modern history could have dreamed of accomplishing. And upon the completion of this challenge - if it does not destroy you first - you will rise a champion. In this challenge, we will test every aspect of your data science skills, as well as tackle your intuition on a typical business problem.

Some context to your mission
Gamers on Mistplay are rewarded for playing, engaging and buying stuff in mobile games. All these games are freemium, where a user can make in-app purchases to unlock levels or buy extra features within the game. Your primary goal today is to build a model to predict whether a user will be spending for one specific game or not (here's a binary classification problem just for you). But could you go further? Explore more options such as building a model to predict when a user will spend for the first time, or how much will a user will be spending in total...

The Data
You are provided with three fictional datasets :

- The user profile (*user_table.csv)* : you will find information on the user (user_id, age category, date he installed Mistplay…) as well as the date of the game install
- The in-app purchase table (*user_purchase_events.csv)* : the in-app purchase events are recorded in this table such as user_id, date and amount (by category)
- The user's apps statistics : extra statistics on the user's apps (number of Google's Play Store top grossing apps, number of shopping apps, total number of apps)

Deploying model to "production"
Once you're satisfied with your model(s), you'll need to create a portal for us to test. Please create a basic webpage with an API that will take our test data set and return the results from your model. The webpage should have a text box where we can simply input our data set (in the same format as the original training set, i.e. the original three tables) and then a button that will send the data to the endpoint and then return the results as a list on the webpage.

<u>Extra information on the datasets</u>

user_table.csv

- Epoch times are given in ms (**installed_Mistplay**, **installed_Mistplay_timezone**, **game_install_timezone**)
- **Source** indicates the network the user used to find Mistplay. If null, this refers to an 'organic' user who downloaded Mistplay directly from the Play Store

user_purchase_events.csv

- **amount_spend** (in arbitrary $):
  Rookie : 1
  Casual : 3
  Player : 5
  Whale : 10

---- Good luck Spartans!