You receive a stream of data from a Citizen Science web project. Each row in the data contains a user *HIT*.

1. A *user HIT* consists of a user ID and a time-stamp that indicates when a user made a contribution to the project.
2. A *user session* is a sequence of *user HITs* for the same user where consecutive *HIT*s are not separated by more than 30 minutes. I.e., if a specific user *HIT* is more than 30 minutes after her previous *HIT*, this user is assumed to be in a new session.

You need to predict, after receiving each user *HIT*, if this user is going to disengage from the current session within the next 5 minutes.

You are given real data collected from a citizen science project. You need to separate this data to a training and testing sets (in a 75/25 ratio) and develop the requested solution.

You should develop a solution in **python** for the above problem. You should train your solution (**only)** on the training data and report measurements when running your solution on the testing data. Your submission should include:

1. A pdf file explaining your solution, the results and your recommended next steps.
2. A pdf file of your notebook with your running code.

Link to data:
https://docs.google.com/spreadsheets/d/1jcck83BBshlg8gbHXRY4GepV6k7U-5YUl87cDTRg7gM/edit?usp=sharing