1. XY\_train.csv in folder ‘’Task 1’’ contains a dataset of **Views** against 6 input variables.
   1. Select and train a model to predict the views for a given video, given the features provided.
   2. Explain your choice of model and why it is suitable for this problem
   3. Comment on your model’s performance as well as the significance of each feature.
   4. Provide predictions for views of videos within X\_test.csv, as well as an estimation of your out-of-sample performance.
2. Using the audience retention graphs of the videos within the folder ‘’Task 2 & 3’’ predict the audience retention % of Video A at positions 0, 1, 2, 3, 4, 5 by inputting your answers in the yellow fields. Explain your reasoning and how you got to this number.
3. Using the scripts provided, come up with more effective wording for the intro of Video A to increase audience retention. Predict the audience retention with this new intro and explain your reasoning.

**Submission Instructions**

* Provide a repository containing your code for each point above
* Send us your predictions for Task 1
* Send us your filled in “X\_test” and “Video A” files for Tasks 2 and 3