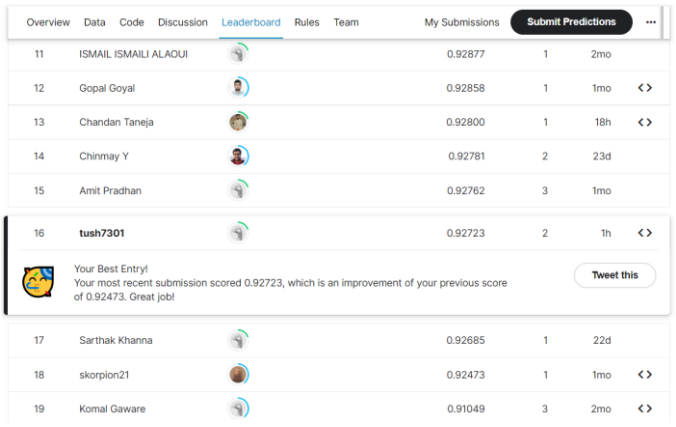
UCS663 CONVERSATIONAL AI, DATA SCIENCE
LAB EVALUATION-1

SUBMITTED TO: DR. SAHIL SHARMA
SUBMITTED BY: TUSHAR MITTAL
3CS10
101916042

Problem Name	CONTRADICTORY, MY DEAR WATSON Detecting contradiction and entailment in multilingual text using TPUs
Problem Link	https://www.kaggle.com/c/contradictory-my-dear-watson/overview
Problem Type	CLASSIFICATION
GitHub Link	https://github.com/tush7301/LAB-EVALUATION-DATA-SCIENCE
Kaggle Link	https://www.kaggle.com/tush7301/watson
Libraries used	Pandas, Numpy, Seaborn, Matplotlib, Transformers, Tensorflow
Model Implemented	Pretrained Model on HUGGING FACE CLASSIFICATION 'joeddav/xlm-roberta-large-xnli'
Kaggle Rank Achieved	16/53
Proof of Rank	

Series of Steps Followed	1. Importing the libraries and Loading the dataset. 2. Removing unwanted columns from the dataset. 3. Applying the pre-trained model 'xlm-robberta-large-xnli'. The model was chosen since it is very effective against text classification, such as with Hugging Face Classification. The model has been pre trained on 100 different languages and has thus shown effectiveness in classifying and doing the above question. (0.92723 score was achived with a rank 16/53) 4. Next, the data was tokenized and a mask column was added. 5. Then model was built using Keras from Tensorflow. 6. And finally predictions were made on test set and recorded in submission.csv.
SUBMISSION FILE	https://drive.google.com/file/d/13hxbNl1HAdkospL_vtCQf0Vs7XW11oKv/view?usp=sharing