

UCS663 CONVERSATIONAL AI, DATA SCIENCE LAB EVALUATION-1

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

Problem Name	CONTRADICTION, 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	<div><div>OverviewDataCodeDiscussionLeaderboardRulesTeamMy SubmissionsSubmit Predictions...</div><table><tr><td>11</td><td>ISMAIL ISMAILI ALAOU</td><td></td><td>0.92877</td><td>1</td><td>2mo</td></tr><tr><td>12</td><td>Gopal Goyal</td><td></td><td>0.92858</td><td>1</td><td>1mo <></td></tr><tr><td>13</td><td>Chandan Taneja</td><td></td><td>0.92800</td><td>1</td><td>18h <></td></tr><tr><td>14</td><td>Chinmay Y</td><td></td><td>0.92781</td><td>2</td><td>23d</td></tr><tr><td>15</td><td>Amit Pradhan</td><td></td><td>0.92762</td><td>3</td><td>1mo</td></tr><tr><td>16</td><td>tush7301</td><td></td><td>0.92723</td><td>2</td><td>1h <></td></tr></table><div><div></div><div>Your Best Entry! Your most recent submission scored 0.92723, which is an improvement of your previous score of 0.92473. Great job!</div><div>Tweet this</div></div><table><tr><td>17</td><td>Sarthak Khanna</td><td></td><td>0.92685</td><td>1</td><td>22d</td></tr><tr><td>18</td><td>skorpion21</td><td></td><td>0.92473</td><td>1</td><td>1mo <></td></tr><tr><td>19</td><td>Komal Gaware</td><td></td><td>0.91049</td><td>3</td><td>2mo <></td></tr></table></div>	11	ISMAIL ISMAILI ALAOU		0.92877	1	2mo	12	Gopal Goyal		0.92858	1	1mo <>	13	Chandan Taneja		0.92800	1	18h <>	14	Chinmay Y		0.92781	2	23d	15	Amit Pradhan		0.92762	3	1mo	16	tush7301		0.92723	2	1h <>	17	Sarthak Khanna		0.92685	1	22d	18	skorpion21		0.92473	1	1mo <>	19	Komal Gaware		0.91049	3	2mo <>
11	ISMAIL ISMAILI ALAOU		0.92877	1	2mo																																																		
12	Gopal Goyal		0.92858	1	1mo <>																																																		
13	Chandan Taneja		0.92800	1	18h <>																																																		
14	Chinmay Y		0.92781	2	23d																																																		
15	Amit Pradhan		0.92762	3	1mo																																																		
16	tush7301		0.92723	2	1h <>																																																		
17	Sarthak Khanna		0.92685	1	22d																																																		
18	skorpion21		0.92473	1	1mo <>																																																		
19	Komal Gaware		0.91049	3	2mo <>																																																		

<p>Series of Steps Followed</p>	<ol style="list-style-type: none"> 1. Importing the libraries and Loading the dataset. 2. Removing unwanted columns from the dataset. 3. Applying the pre-trained model 'xlm-roberta-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 achieved 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.
<p>SUBMISSION FILE</p>	<p>https://drive.google.com/file/d/13hxbNI1HAdkospL_vtCQf0Vs7XW11oKv/view?usp=sharing</p>