

## Lab 8: NLP

### Spring 2023 – April 10

### Twitter US Airline Sentiment: HuggingFace

---

In this lab, we will continue with the data from the previous labs, but this time we will be using one of the HuggingFace pretrained models and fine-tune it on our data. Here is the link for the sentiment analysis models:

<https://huggingface.co/models?other=sentiment-analysis>

There is a deck of slides for HuggingFace tutorial provided for you. In addition to this, two base codes are shared with you. One code is a sample code of fine-tuning with IMDB dataset is provided for you (You may first run this code to familiarize yourself with HuggingFace).

Then, there is another code for fine-tuning any model (you can change the name of pre-trained model) for Twitter data. You do not need to handle the twitter data, as the code is provided for you. It is strongly recommended that you run the code on GPU for speed (In Colab, go to RunTime and then “Change run type”). After fine-tuning your model, you can check the test accuracy. Your final model, will be run on an unseen test data and you will be graded based on the ranking of accuracy as follows:

Team Rank	Grade
1	100
2	96
3	92
4	88
5	84
6	80
7	76

**Derivable:** In your submission, provide a PDF file, on the top of the file the link to your GitHub repo, and then the accuracy of on your validation and test set.