CIS 4190/5190 Applied Machine Learning
Fall 2024 Course Project Handout:
Topic 2 - News Source Classification

Project TA: Zelong Wang, Haorui Li, Wendy Deng, Chandler Cheung

1 Model and Leaderboard Submission

Upload both the provided dataset and your own data (if you would like to use additional data) to Huggingface. This way, everyone in your team, as well as the TAs, can access your data using a simple API call. After training your model using this data, you will also need to upload your model to Huggingface as part of your deliverables by following the instructions below. Please make sure that each model and dataset has a clear and unambiguous name. Additionally, in your model's repository, update the 'README.md' file to include a simple code example demonstrating how to load your model and perform inference on our released dataset. You **must** ensure that your model is capable of running inference on our released datasets; otherwise, we will not be able to evaluate your model.

Ideally, we should be able to run your model using a simple Hugging Face API. If you used a custom model class, please ensure that its implementation is included in the 'README.md' as well.

Additionally, make sure that your dataset is private (visible only to your team members) and that manual review is enabled for your model weights. Do not approve access for anyone other than your team members and the TAs. You can enable manual review by navigating to **Settings** → **Enable Access Requests** → **Manual Review**.

1.1 Submission Details

The first step is for one of your teammates (as the admin) to create an organization on Huggingface, and then invite your teammates and TAs to join the organization (go to Organization Settings → Members → Add New Member → Set Role to 'Write' → Fill in the username).

After training your model and when you are ready to submit it for the leaderboard, push your model weights to Huggingface by following this notebook. You are able to use any models on Huggingface, but make sure you modify the CustomModel class in the notebook.

After uploading your model, make sure to enable access requests and select "Manual request" for new requests. This ensures that people from other teams **will not** have access to your model weights.

Please add the TAs as members of your Hugging Face organization by filling out this Google form. (Only one submission per team is required.) **It is crucial to ensure that the TAs have the correct access** and it is your responsibility to verify that we have the proper access. Please remind us if we accidentally missed the invite. The TA Hugging Face usernames are as follows:

1. haoyund

2. haoruili

3. ZelongWang

4. cheungch

Due to limited computational resources, the leaderboard will not update in real-time. The TA team will evaluate your progress over the weekend and update the ranking file here. Each week's rankings will be available in a separate tab.

Each group must make at least one leaderboard submission before the project deadline to receive credit.

2 Resources

1. **Upload Datasets and Model to Hugging Face:** link here