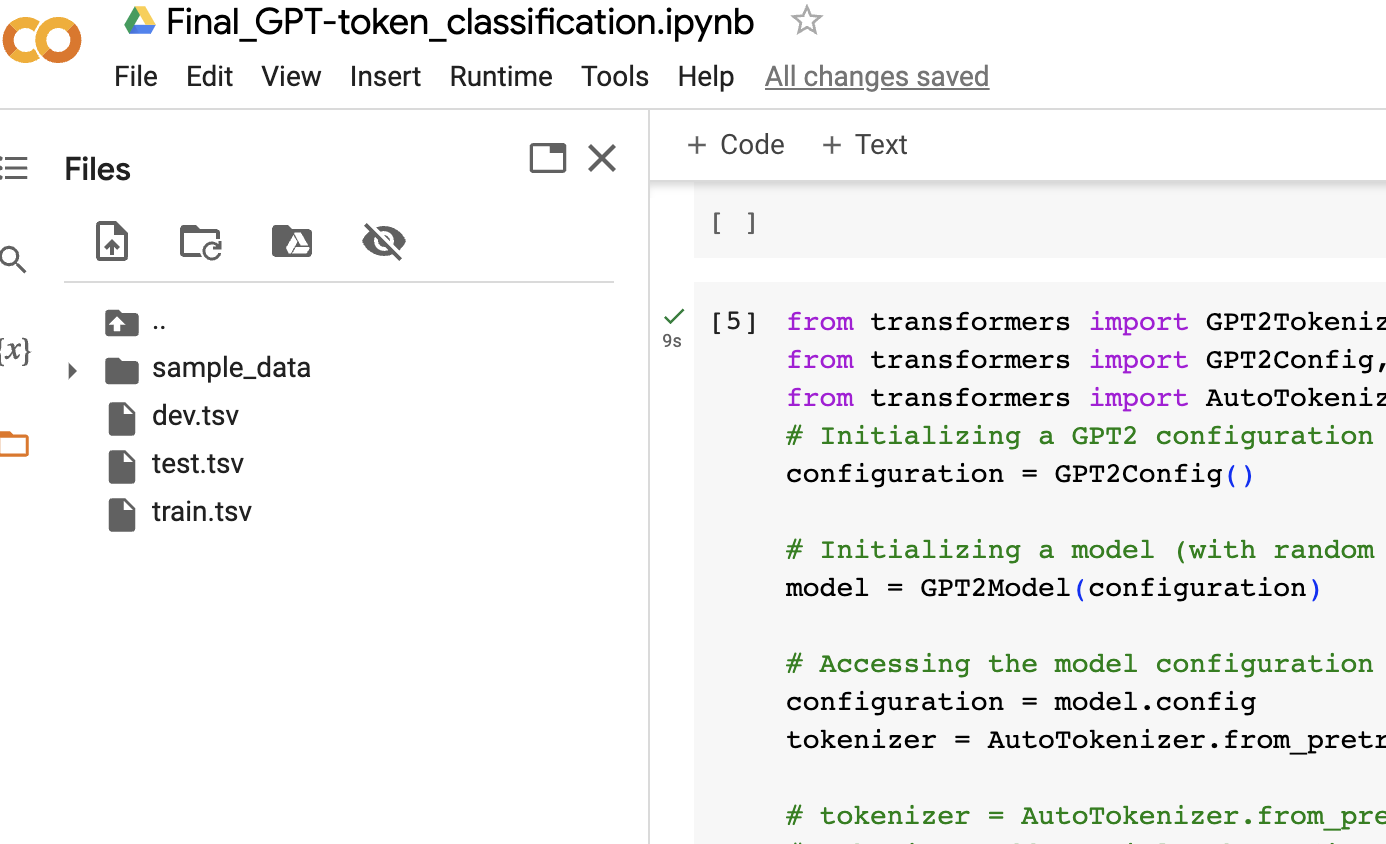
**Datasets:**

We have DataSet folder which contains datasets for different languages.

**\*\*\*Only English-FCE dataset can be used with all the notebooks. For other languages, only the BERT-MultiLingual-GED.ipynb can be used.\*\*\***

For running on the English-FCE dataset we have 3 files(train.tsv, test.tsv and dev.tsv). All these files need to be uploaded in the root directory of the python notebook. This English-FCE dataset can be used with any of the python notebooks as shown in the following image on the left.



Similarly, for running the BERT-MultiLingual-GED.ipynb notebook, we require train.tsv and dev.tsv that can be taken from a single language that you can select(Czech, swedish, italian, English-FCE). Both the train.tsv and dev.tsv have to uploaded into the root directory inside the python notebook.

**Running the python notebooks:**

Before running the cells in each of the notebook, it requires us to have a Hugging Face account as the notebook would ask for a token in case of uploading the models into hub of Hugging Face. The models that we trained have been made public so that you do not encounter any errors while running the evaluation steps in the notebook.

In case you face any errors, you need to create a write access token in Hugging Face. After creating the token, the username can be given in places where the usernames have been mentioned in lines such as model="Varunreddy/bert-token-class".

Here ‘Varunreddy’ can be changed to your username.

Just click on Runtime->Run All.

In case the training step fails in the training step due to ‘Partial State’ error, you can click on Runtime->Restart Runtime and click on Runtime->Run All again.

**Note**: Most of our fine-tuned models can be accessible on hugging face w.r.t the model name in the colab notebook. You can evaluate the data directly on these models incase you dont want to run the fine-tuning code.