

## **Coding Progress**

### **Current Progress**

We have implemented the transform model, T5, which is a text-to-text model that takes in a sentence or groups of sentences. The output will be that of grammatically correct sentences with spelling correction also considered. We are training our model on the JFLEG dataset, which is the same dataset used in applications for determining the performance of a grammar error correction system. There are three sizes of the dataset that we will be working with small (60 million), medium (220 million), and large (11 billion). For our current implementation of the model we have it being trained on the small dataset. We intend to implement the model on each of the three datasets and compare the performance of each of them.

### **Current Difficulties**

The current focus has been on implementing the transformer model known as T5 to be used for grammar-correcting applications. We are currently facing difficulties when attempting to work with the medium and large datasets. Additionally, with the small dataset, it is not having the desired performance, likely due to not having a large enough dataset.

### **Next Steps**

We are investigating the issues that are being faced when training on the larger datasets and will hopefully be able to train on both the medium and large datasets soon. After this is completed we will determine the performance of each model and will compare and contrast these different models.