Ayush Ranjan

[Email address]

Abstract

In this continual learning endeavor, i trained an adaptive model on medical text data, freezing key insights to retain knowledge. Evolving through datasets, model's F1 score mastery ensures it adapts dynamically while preserving past expertise.

Continual Learning

A brief report on continual learning experiment

**Continual Learning Model Report**

**Objective**The objective of the task was to implement a continual learning model for named entity recognition. The goal was to train an initial model (model1) on the first dataset (data1\_train), extend the training on the second dataset (data2\_train) and 100 samples from the first dataset such that the trained model does not forget the learning from first dataset (model2), and finally train the model on third dataset (data3\_train) along with 100 samples from each dataset1 and dataset2.

**Implementation Steps:**

* Loaded the dataset, cleaned the tags column and pre-processed the medical text datasets: G1, G2, and G3
* Split each dataset into training and testing sets (data\_train and data\_test)
* Sampled 100 examples from data1\_train and data2\_train
* Defined a neural network architecture (model1) for the initial training
* Trained model1 on data1\_train and evaluated its performance on data1\_test for each tags
* Loaded model1 and extended the training on data2\_train and 100 samples from data1\_train
* Evaluated the updated model (model2) on combined data of data1\_test and data2\_test
* Loaded model2 and extended the training on data3\_train and 100 samples from each data1\_train and data2\_train
* Trained a model on combined train dataset of G1, G2 and G3
* Created a function which takes dataset and model as input and gives back the f1 score for all the tags
* Tested this function in a separate notebook
* Put the trained model on Huggingface hub

**Freezing Layers:**

* Used the technique of freezing layers to retain knowledge learned from the previous datasets
* Freezed the layers of the loaded models (model1 and model2) before extending the training

**Evaluation:**

* Utilized F1 score as a metric for model evaluation on specific tags: 'treatment', 'cancer', 'allergy\_name', 'chronic\_disease'
* Evaluated each model on the different test datasets

**Result:**

* The freezing of layers helped prevent forgetting of knowledge from the initial training on data G1
* Performance of trained model on Allergy\_name is very minimal because of few data points with Allergy\_name tag to learn