**Prompt Learning – Health Prompt**

Week 1-3 (Jan 3, 2022 – Jan 9, 2022):

* Went through many papers on prompt learning and found out that AutoPrompt by UC Irvine(EMNLP 2020 : <https://github.com/ucinlp/autoprompt> ) and OpenPrompt by Tsinghua University (<https://github.com/thunlp/OpenPrompt>) are the available libraries for prompt learning.
* These are in their initial stages and still lacking many features.
* successfully got both the libraries working on general data, but open prompt seems more relevant to our project as it is a zero-shot learning methodology, while autoprompt still needs training dataset for selecting the best prompt template. The caveat is that openprompt needs manual templates defined by us, hence the best template can only be found by trial and error.
* tried Openprompt with health datasets and it gave good results on Clinical BERT and BERT. BioBERT and RoBERTa showed poor performance for some reasons. But even BERT failed to classify covid related texts.
* Results from larger text corpora(i2b2) was not promising. I can demonstrate that in the meeting tomorrow.
* Used streamlit to deploy the prompt application on university server and an external server. This will be useful when we are validating the results with a medical expert.
* I have seen papers describing that prompt can be used for NLU, NLP and other NLP tasks. Probably, we can try some of these once we are in good shape.

Week 4 (Jan 10, 2022 – Jan 16, 2022):

* o worked on Smoking Dataset - Classification accuracy was 56% Note-there are less than 150 text records in Smoking dataset
* o created overall system architecture
* o Created a Input custom dataset module
* o Created text cleaning module
* o created a text splitting module
* o created weighted score function for large text corpus, but it is too slow. We have to decide between accuracy and robustness - testing is pending

Week 5 (Jan 17, 2022 – Jan 23, 2022) WIP:

* o Solve negation issue
* o Read both papers and solve the template issue
* o Make autoprompt working
* o Create a package out of this(can be pushed to next week)

Evaluation metrics:

**w/o weighted average**

w/o text cleaning and preprocessing:

Graphical user interface, text

Description automatically generated

with text cleaning

Graphical user interface, text

Description automatically generated

**With weighted average:**

**With cleaning**

**Graphical user interface, text

Description automatically generated**

**Weight 0.2**

**Text

Description automatically generated**

**w/o**

**Graphical user interface

Description automatically generated**

**No weight.. smoker>0**

**Clinical bert**

**With cleaning**

**Accuracy : 61.**

**Without cleaning :**

**Accuracy: 58.9**

**Biobert:**

**With cleaning : 54.79**

**BERT:**

**With cleaning 53.4**

**RoBERTa:**

**Text

Description automatically generated**

**Phenotype Dataset:**

**BioClinical BERT experiments:**

**0.87 template\_text = '{"placeholder":"text\_a"}. Disease : {"mask"}'**

weighted avg 0.90 0.87 0.83 347

**0.5821 template\_text = '{"placeholder":"text\_a"}: This effects {"mask"}'**

**0.4522 template\_text= '{"placeholder": "text\_a"} : {"mask"} disorder'**

**0.72 template\_text = '{"placeholder":"text\_a"}: {"mask"} type of disease'**

**Bio BERT experiments:**

**0.50 template\_text = '{"placeholder":"text\_a"}. Disease : {"mask"}'**

0.58 0.5 0.09 347

**0.40 template\_text = '{"placeholder":"text\_a"}: This effects {"mask"}'**

weighted avg 0.34 0.4 0.06 347

**0.634 template\_text= '{"placeholder": "text\_a"} : {"mask"} disorder'**

0.51 0.06 0.08 347

**0.3746 template\_text = '{"placeholder":"text\_a"}: {"mask"} type of disease'**

weighted avg 0.99 0.04 0.07 347

**BERT experiments:**

**0.59 template\_text = '{"placeholder":"text\_a"}. Disease : {"mask"}'**

**0.35 template\_text = '{"placeholder":"text\_a"}: This effects {"mask"}'**

**0.51.87 template\_text= '{"placeholder": "text\_a"} : {"mask"} disorder'**

0.78 0.51 0.09 347

**0.54 template\_text = '{"placeholder":"text\_a"}: {"mask"} type of disease'**

0.74 0.54 0.12 347

**RoBERTa experiments:**

**0.691 template\_text = '{"placeholder":"text\_a"}. Disease : {"mask"}'**

0.82 0.07 0.11 347

**0.60 template\_text = '{"placeholder":"text\_a"}: This effects {"mask"}'**

0.89 0.06 0.10 347

**0.778 template\_text= '{"placeholder": "text\_a"} : {"mask"} disorder''**

0.48 0.08 0.11 347

**0.72 template\_text = '{"placeholder":"text\_a"}: {"mask"} type of disease'**

**GPT2 experiments:**

**0.87 template\_text = '{"placeholder":"text\_a"}. Disease : {"mask"}'**

**0.5821 template\_text = '{"placeholder":"text\_a"}: This effects {"mask"}'**

**0.67 template\_text= '{"placeholder": "text\_a"} : {"mask"} disorder'**

**0.72 template\_text = '{"placeholder":"text\_a"}: {"mask"} type of disease'**

**T5 experiments:**

**0.87 template\_text = '{"placeholder":"text\_a"}. Disease : {"mask"}'**

**0.5821 template\_text = '{"placeholder":"text\_a"}: This effects {"mask"}'**

**0.67 template\_text= '{"placeholder": "text\_a"} : {"mask"} disorder'**

**0.72 template\_text = '{"placeholder":"text\_a"}: {"mask"} type of disease'**