Ciphix Machine Learning Case

CIPHIX

Problem Description

At Ciphix Health, you are part of a team that helps doctors & other medical employees do their job more efficiently. Recently, you started a project with your team to reduce the administrative workload off of your client.

Every day, your client processes a great number of medical documents. In doing so, your client has to scan these documents, extract relevant information and fill out paperwork based on this information.

You, being a knowledgeable AI developer, know exactly what this client needs. Given the structured format of the paperwork, you know filling out forms is easy to be automated. The only challenge lies with extracting the right information from unstructured medical text documents that need to be filled in.

You decide to pitch this idea to your client and decide to build a demo of this technology in order to showcase its potential.

Assignment

Your assignment boils down to the following:

- Use a medical entity extraction dataset of your choosing – think about this choice carefully given the needs of your client which are identify diseases, extract bloodwork values etc. from unstructured medical documents.
- Build & train an entity extraction model
- Deploy this model as an API (local deployment using Flask is sufficient for this step)
- Build a front-end that is understandable and appealing to your client to use during the demo. This demo should showcase the potential of this solution. For example, allow a client to upload a medical document, use the model for prediction in the background and show tagged entities in the original text.

Resources

Below you can find some resources to get you started:

- https://arxiv.org/pdf/1901.08746v4.pdf
- https://arxiv.org/pdf/1910.07419.pdf
- https://aws.amazon.com/comprehend/medical/

Requirements

1. Use Python.
2. Submit a train.py file that is used to train the model / share or make sure that the data can be accessed.
3. Submit the trained model alongside with script(s) to run the deployed service / demo application.
4. Make sure to use virtual environments & submit your settings in order for us to reproduce the development environment.
5. Submit a small instruction README that explains how to run your code / how to do the demo (include test data to be used in the demo as well).

Evaluation of assignment

We evaluate your assignment on based on the following points:

1. Quality of your code (structure, readability, etc.)
2. Efficiency and scalability of your solution
3. Technical skills (machine learning algorithms, technical approach etc.)
4. Creativity on presenting your results

Submission of assignment

1. Track your progress on GitHub
2. and share your repo link with batav@ciphix.io when you are done.
3. If you need to share large files as part of your submission – you can use WeTransfer to batav@ciphix.io to do so.

Good luck!