Capstone 2 Project Proposal

Writing cover letters is often a time consuming and carefully crafted task which a person tries to accomplish with a formulation detailed on a particular job. I have started to write these cover letters and have found that it is in fact time consuming and I have a lot to accomplish and wish to maximize my time. Therefore, I have decided to create a cover letter generator; This particular project would not only help me to maximize my time but can be used by all job searchers to ensure they have a well crafted cover letter specific towards a job description and outlining the skills required for the job they are interested in pursuing.

In order to require the data for a cover letter generator, the web will be searched and the sample cover letters will be scraped from the internet to gather many cover letters to include in a training corpus. The cover letters should be carefully chosen to ensure that they are communicating effective descriptions of a job seeker and should contain a contemporary format. I will have to search websites that display impressive samples, historically accepted, and attention grabbing cover letters that will be reflected in the cover letter generator's output. Scrappy is an easy, fast and powerful web scraping/crawling framework. It allows one to implement task specific functionality and is used by many professionals within the Natural Language Processing and Understanding world. After scraping the data from the internet, the format should be preserved and cleaned to fit the expected format of the generating sequence to sequence model. Such model has finally been released by openAI after much debate about a powerful text generating falling into the wrong hands; it is called GPT-2. This is a great time to be involved in NLP because recent advancements have propelled this area of artificial intelligence. One advancement using these new and state of the art models was just released last month by a company called hugging face. Hugging face has created a pytorch architecture that uses transformers from 30 different SOTA models for Natural Language Processing specifically Natural Language Understanding and Natural Language Generation. The transformers are easy to use and have fantastic results from fine tuning these neural networks for a specific task. It took days on many TPU's to develop these transformers but I think it is amazing that we can use them now on our projects and fine tune them with our own data. These pre-trained weights will be used to understand a job description and create a cover letter by providing it a seed. This seed will be based on the job title and ideally will generate a cover letter that uses the job description's requirements and preferences to generate a specific, original and effective cover letter. A deliverable for this project would ultimately lead to a website that can be used by all job seekers but for the sake of time, because I have to find a job now, will only generate cover letters from the quick and easy to use Jupyter Notebook.