**Master’s Project**

**Title: US Patent Phrase To Phrase Matching: Detect Novel Semantic Similarity (NSS) To Extract Meaningful Information From US Patent Applications**

**Abstract: The U.S. Patent and Trademark Office (USPTO) offers one of the world’s largest repositories of scientific, technical, and commercial information through its Open Data Portal (ODP). Patents are a form of intellectual property granted in exchange for the public disclosure of new and valuable inventions. The motivation of this undertaking is to detect/establish “novel semantic similarity (NSS).” NSS between phrases is a critical part of the patent evaluation process to recognize if an invention has been described before. Spurious and duplicate patent applications clog the pipeline and lead to unnecessary time and financial wastage. The scope of this project goes beyond simple phrase-to-phrase matching. Contextual establishment of phrases is also essential. For example, while the terms "bird" and "Cape Cod" may have low semantic similarity in everyday language, the likeness of their meaning is much closer if considered in the context of "house" or “scenery.”**