

Course Competency

CAI 2300C Introduction to Natural Language Processing

Course Description

Students will learn the fundamental concepts of Natural Language Processing (NLP) and text processing. In addition, focus will be on knowledge and skills necessary to create a language recognition application. Recommended preparation: COP1047C or equivalent knowledge of Python programming language. (2 hr. lecture, 2 hr. lab)

Course Competency	Learning Outcomes
<p>Competency 1:Students will describe common techniques in Natural Language Processing and associated applications by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data 2. Critical thinking 3. Information Literacy 4. Ethical Issues 5. Computer / Technology Usage
<ol style="list-style-type: none"> 1. a) Exploring AI For NLP, Applications of NLP, NLP data processing, BOW, Algorithms in NLP. b) Processing textual data by sentence segmentation, tokenization, lemmatization, stop word removal, etc. c) Applying data preprocessing techniques like document similarity, Word Vectors, Cosine similarity, etc. d) Distinguishing between NLP models and algorithms. 	
<p>Competency 2:Students will describe the data acquisition process in NLP by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data 2. Critical thinking 3. Information Literacy 4. Ethical Issues 5. Computer / Technology Usage
<ol style="list-style-type: none"> 1. a) Comparing different types of NLP datasets. b) Identifying and examining various data storage methods. c) Examining different sources of curated 	

<p>data and interpreting their usage in NLP domain. d) Downloading and processing data using the NLTK library. e) Applying data visualization techniques specific to NLP.</p>	
<p>Competency 3:Students will explore NLP Data Preprocessing by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data 2. Critical thinking 3. Ethical Issues 4. Computer / Technology Usage
<ol style="list-style-type: none"> 1. a) Utilizing proprietary and open-source libraries and data visualization techniques. b) Exploring and applying various vectorization techniques. c) Exploring and applying the methods of document similarity and vector visualization. d) Distinguishing between various distance measurement techniques. e) Defining and understanding the various processes associated with NLP data pipeline. 	
<p>Competency 4:Students will describe, compare, and train different machine learning models by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data 2. Critical thinking 3. Information Literacy 4. Ethical Issues 5. Computer / Technology Usage
<ol style="list-style-type: none"> 1. a) Describing and applying NLP classifiers to train machine learning models. b) Describing neural networks and their working principles. c) Understanding various language models. d) Defining and summarizing various Neural Language Models, N-gram Models and Sequential Models. e) Defining and demonstrating Recurrent Neural Networks and Named Entity Recognition (NER) models through various activities and use cases. 	
<p>Competency 5:Students will explore NLP Model deployment by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data

	<ol style="list-style-type: none"> 2. Critical thinking 3. Information Literacy 4. Ethical Issues 5. Computer / Technology Usage
<ol style="list-style-type: none"> 1. a) Identifying and exploring various machine learning model deployment platforms. b) Describing and classifying various types of chatbots by their applications. c) Implementing Language Detection, Transliteration, Translation, Sentiment Analysis for different language scenarios. d) Using different tools to create and deploy chatbots, pre-existing chatbot frameworks, and ChatterOn. e) Using cosine similarity in neural networks to train chatbots. 	
<p>Competency 6:Students will discuss and describe advanced models in NLP by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data 2. Critical thinking 3. Information Literacy 4. Ethical Issues 5. Computer / Technology Usage
<ol style="list-style-type: none"> 1. a) Exploring the most recent developments in the NLP space. b) Explaining the workings of LSTM, Transformers, and BERT. c) Examining several NLP pretrained models. d) Comparing the workings, performance, and architectures of DistilBERT, RoBERTa, GPT, GPT -2, BERT, and BART models of NLP. e) Summarizing various types of learning fast learning techniques, Zero-shot, One-shot, and Few- shot. f) Evaluating various ethical issues in language models. 	

Updated: SUMMER TERM 2023