**CS 5316 – Natural Language Processing**

Spring 2019-2020

## 

|  |  |
| --- | --- |
| Instructor | Asim Karim |
| Room No. | 9-G18A |
| Office Hours | TBA |
| Email | [akarim@lums.edu.pk](mailto:akarim@lums.edu.pk) |
| Telephone | 8205 |
| Secretary/TA | Ishtiaq A Bhatti |
| TA Office Hours | TBA |
| Course URL (if any) | On LMS |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Course Basics | | | | |
| Credit Hours | 3 hours | | | |
| Lecture(s) | Nbr of Lec(s) Per Week | 2 | Duration | 75 Minutes |
| Recitation/Lab (per week) | Nbr of Rec(s) Per Week |  | Duration |  |
| Tutorial (per week) | Nbr of Tut(s) per Week | 1 (optional) | Duration | 50 Minutes |

p

|  |  |
| --- | --- |
| Course Distribution | |
| Core |  |
| Elective | This is an elective course. |
| Open for Student Category | Senior & Graduate |
| Close for Student Category | Freshmen, Sophomore |

|  |
| --- |
| COURSE DESCRIPTION |
| This course will cover the foundations of natural language processing (NLP) from textual content processing to corpus understanding. It is designed for develop the syntactic and semantic concepts of NLP and introduce the computational techniques for analyzing and understanding textual content. In addition to foundations, the course will also introduce significant application areas of NLP such as sentiment analysis, machine translation, and question-answering/conversational agents. A strong programming background is required for the course, and it is recommended for CS seniors and graduate students. |
|  |

|  |  |
| --- | --- |
| COURSE PREREQUISITE(S) | |
|  | (CS 300 – Advanced Programming AND MATH 230 - Probability) OR Graduate standing  Proficiency in Python; CS 535 – Machine Learning will help but is not required |

|  |  |
| --- | --- |
| COURSE OBJECTIVES | |
|  | To introduce the fundamental concepts and techniques in textual natural language processing  To provide experience in the implementation and evaluation of NLP algorithms  To introduce NLP resources and application areas |

|  |  |
| --- | --- |
| Learning Outcomes | |
| 1.  2.  3.  4. | Process and segment textual content for natural language processing  Extract syntactic and semantic structure from text  Understand and apply probabilistic and neural network sequence modeling techniques for text processing  Use language resources and corpora to implement NLP solutions  Process multi-lingual and informal-language text in NLP solutions |
| Grading Breakup and Policy | | |
| Assignment(s) (~ 4): 15%  Project: 10%  Quiz(s) (~ 8): 12%  Midterm Examination: 30%  Final Examination: 33% | | |
|  | | |

|  |  |  |
| --- | --- | --- |
| Examination Detail | | |
| Midterm Exam | Yes/No: Yes  Combine Separate:  Duration: 2 hours  Preferred Date:  Exam Specifications: Closed books and notes. Help sheet and calculator allowed. |
| Final Exam | Yes/No: Yes  Combine Separate:  Duration: 2 hours  Exam Specifications: Closed books and notes. Help sheet and calculator allowed. |

     Course Outline

|  |  |  |  |
| --- | --- | --- | --- |
| # | Topic | Lecture | Reading |
| 1 | Introduction and Motivation | 1 | 1:1 |
| 2 | Basic Text Processing: normalization, regular expression, min. edit distance | 2-4 | 1:2 |
| 3 | Language Modeling with N-grams: MLE, add-k smoothing, backoff, Kneaser-Ney smoothing | 5-8 | 1:3 |
| 4 | Text Classification and Sentiment Analysis: naïve Bayes, logistic regression, feature engineering | 9-11 | 1:4, 1:5 |
| 5 | Vector Semantics and Embeddings: co-occurrence, word2vec, glove | 12-14 | 1:6 |
| 6 | Midterm exam | 15 |  |
| 7 | Neural Network and Neuro Language Modeling: feedforward network, learning, evaluation | 16-17 | 1:7 |
| 8 | Sequence Modeling: HMM, RNN, GRU, LSTM | 18-21 | 1:9 |
| 9 | POS Tagging and Named Entity Recognition | 22 | 1:8, 1:18 |
| 10 | Semantics and Textual Resources: lexicons, WordNet | 23-24 | 1:19 |
| 11 | Machine Translation and Language Generation: encoder-decoder architecture, evaluation | 25-26 | 1:10, 1:11 |
| 12 | Question-Answering and Chatbots | 27-28 | 1:25, 1:26 |

|  |  |
| --- | --- |
| Textbook(s)/Supplementary Readings | |
| 1. Speech and Language Processing, 3rd Edition, Jurafsky and Martin, Stanford University, https://web.stanford.edu/~jurafsky/slp3/, 2019. 2. Natural Language Processing with Python, Bird and Klein, O’Reilly Media, 2009. |