CS249 Advanced Seminar: Learning from Text – Spring 2019

Time and Place

• **Hours**: Monday and Wednesday, 10:00 AM - 11:50 AM
• **Location**: Public Affairs 2250
• **Web site**: http://oak.cs.ucla.edu/classes/cs249/

Instructor

• **Name**: Junghoo “John” Cho
• **Email**: cho@cs.ucla.edu
• **Office**: 3531H Boelter Hall
• **Office hour**: Tuesday 2:30PM - 3:30PM

Course Description

Humans use language to perceive, describe, communicate, and record the world around them, so text is the primary “data” generated by humans. Despite the importance of text data written in natural languages, making computers “understand” them is still a challenging task. In this class, we will go through key papers in natural language processing (NLP) throughout the quarter and learn (1) how NLP problems can be formulated as a computational problem and (2) the basic algorithms that solve the formulated problems.

Course Organization

The course consists of paper readings, presentations, and projects. Students will have to read research papers throughout the quarter and write short summaries. Every student group will also have to give a two-hour “lecture” on a paper once. Finally, students have to finish an independent project on an NLP task of their choice, give short presentation on their idea and result, and submit a final report.
Course Goal

There are three high-level goals behind this class:

1. Learn the basic theory: By going over key research papers with other students, students will learn the basic NLP theory.
2. Learn to read research papers: A big part of a software engineer or research scientist’s task is to read and understand research papers. We will learn this skill by reading and discussing more than 20 key research papers this quarter.
3. Learn to present ideas: Being able to explain one’s ideas in a way that others can understand is an important skill in any career. Students will practice and learn how to present through three presentations – once in a lecture-style and twice in a conference style.
4. Perform mini “research”: Through an independent project, students will have a chance to come up with an idea of their own and evaluate it experimentally.

Prerequisites

This class requires what is expected from a typical graduate student in Computer Science: Solid understanding of core CS theory, competence in at least one popular programming language, lots of programming experience. In addition, this class requires the following prior knowledge:

- Understanding of basic machine learning
- Advanced Calculus, in particular understanding of partial derivatives
- Familiarity with basic optimization techniques, including the greedy approach, linear programming, and gradient-based optimization
- Understanding of basic probability theory including the Bayes rule
- Understanding of basic linear algebra
- Understanding of basic information theory, including entropy, mutual information, etc

The papers covered in this paper will be extremely math heavy. Strong math background is an absolute requirement. Students should also have access to a computer on which they wrote code, conduct experiments, and access the Internet.
Books

The class does not have a required textbook, but students may find the following books helpful for reference and in-depth learning:

- Foundations of Statistical Natural Language Processing by Christopher D. Manning and Hinrich Schutze
- Speech and Language Processing by Daniel Jurafsky and James H. Martin
- Deep Learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville
- Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit by Steven Bird, Ewan Klein, and Edward Loper

Grading Plan

For final grading, coursework will be weighted as follows:

- 10% Attendance
- 20% Paper summaries
- 20% Paper presentation
- 50% Project

Class Communication

Important class announcements will be done through online class forum on Piazza. If you have any questions regarding class materials, they also need to be asked on Piazza. Please make sure to sign up for CS249 Piazza forum at http://piazza.com/ucla/spring2019/cs249