595/498: Natural Language Processing
Syllabus

Instructor
Rada Mihalcea, BBB 3769, mihalcea@umich.edu

GSI
Laura Wendlandt, wenlaura@umich.edu
Steven Wilson, steverw@umich.edu

Meeting times
Wed 1:30-3:30pm, Fr 3-4pm, DOW 1013

Office hours
Instructor: Tue 3-4:00pm, Beyster 3769
GSIs: Thurs 1-3:00pm, Beyster 1695 (Laura)
     Mon 1-3:00pm, Beyster 1695 (Steve)

Course description
This course provides an introduction to the theory and practice of natural language processing (NLP) - the creation of computer programs that can understand, generate, and learn natural language. The course will cover the three major subfields of NLP: syntax, semantics, and pragmatics. The course will introduce both knowledge-based and statistical approaches to NLP, illustrate the use of NLP techniques and tools in a variety of application areas, and provide insight into many open research problems.

Website
Most course materials are made available on Canvas, and are considered required reading. A detailed schedule, including lecture topics, assignment due dates and exam dates, will also be available on Canvas.

Forum
We will be using Piazza to host a course forum. You are required to read this forum regularly; it is the venue we will use for important course announcements and assignment clarifications. In addition, it will be a significant source of help and hints on the assignments. We do not answer technical questions via email. In order to save everyone time, we want all students to have the benefit of seeing each question and its answer, so please use the forum. Please do not post your own solutions, project code, test cases, or output to the forum. Also, please search the forum before posting to avoid questions that have already been asked and answered.

Contact
Please direct all technical questions to our Piazza forums. For other questions, you can reach us through private posts on Piazza, or by email.

Prerequisites
For 498.007 students: EECS 281 Data Structures and Algorithms

List of Topics

Brief overview of English linguistics
Text preprocessing
Language models
Word embeddings
Part of speech tagging
Syntactic parsing
Lexical semantics
Word sense disambiguation
Subjectivity and sentiment analysis
Information extraction
Automatic summarization
Question answering
Machine translation
Natural language processing for social sciences
Other special topics

Required/recommended readings
(required) Speech and Language Processing, by D. Jurafsky and R. Martin (2nd edition)
(recommended) Foundations of Statistical Natural Language Processing, by C. Manning and H. Schutze
Additional readings will be assigned throughout the semester.

Programming Environment

For this course, you must have a CAEN account. Anyone enrolled in EECS 595/498 is eligible for one, but it takes a little while to get it set up. You will have to write your programs in Python. You are free to develop your programs on any platform you like, but they must run on the CAEN platform. We will grade your programs in the CAEN Linux environment and they must run correctly in this environment.

iClickers
The class will make use of iClickers for in-class multiple-choice questions. Please make sure you register your iClicker on Canvas.

"OneCoolThing"
Every Wednesday, three teams of students will prepare a short (1-2 slides) presentation on "One Cool Thing" - a recent exciting finding related to the field of NLP. The teams of students will be announced ahead of time on Canvas, and the teams will receive a reminder a few days in advance about their upcoming "OneCoolThing". One of the three teams will be randomly selected to present their OneCoolThing in class, during the Wednesday lecture. The other two teams will have to share their OneCoolThing presentations on Piazza, by the end of Wednesday.

Course work and grading
Programming assignments: 35%
Exam I (Oct.11): 20%
Exam II (Dec. 14 or Dec. 15): 20%
Project: 20%
Class participation: 5%

There will be four programming assignments throughout the semester. Each assignment will specify the material to be turned in. Assignments are due by 11:59pm on the due date. Assignments may be turned in up to 3 days late, with a penalty of 10% for each day late. No credit will be given after 3 days. No exceptions.

A small amount of credit (5%) will be awarded for class attendance and participation. What counts toward class participation:
- answers provided in response to in-class iClicker questions (note that the correctness of the answers will **not** count toward grading)
- answers provided to student questions on Piazza
- feedback/comments provided to OneCoolThing presentations shared on Piazza. You have to write a "substantial" (10+ words) comment or question in response to at least five OneCoolThing presentations shared on Piazza

Accommodations for Students with Disabilities
If you think you need an accommodation for a disability, please let your instructor know at the beginning of the semester or at least 2 weeks in advance. As soon as you make us aware of your needs, we can work with the Services for Students with Disabilities (SSD) office to help us determine appropriate academic accommodations. SSD (734-763-3000; http://ssd.umich.edu) typically recommends accommodations through a Verified Individualized Services and Accommodations (VISA) form. Any information you provide is private and confidential and will be treated as such.