Finalized Syllabus

**Introduction to Speech and Language Processing (CS 5319)**

**Topics in Data Science (CS 4364)**

Fall 2023

Tuesdays and Thursdays 3:00 – 4:20 in Education Room 112

Instructor: Nigel Ward
Office: CCS 3.0408
Phone: 747-6827
E-mail: nigel@utep.edu
Office Hours: Wed 3:45 – 4:45, Fri 11:00 – 12:00, 12:00 – 1:00
and usually whenever the door is open; also by appointment

Speech and language processing has recently seen tremendous advances, and several core technologies are now mature. There are well-designed systems used by millions of people every day, and readily usable APIs available for both data scientists and developers of user-facing systems. At the same time, the range of viable applications is still quite limited, due to numerous challenging open problems.

This class will provide a survey and sampling of the techniques and issues in speech and language processing. Students will design, implement, and evaluate a project applying what they learn to a problem of their choosing.

**Learning Outcomes** Students having completed this course will be able to:

1a Describe the phonetic, lexical, syntactic, and pragmatic properties of human language and the associated challenges for automatic processing.

1b Explain the algorithms and representations underlying applications such as text classification, speech recognition, speech synthesis, question answering, information extraction, machine translation, search, and conversational agents.

1c Explain the nature, uses, and limitations of large language models and other pretrained models.

2a Given a user need or a business need related to natural language, identify possible technical solutions, and estimate their feasibility and likely cost.

2b Rapidly develop prototype software for typical speech and language processing tasks, using standard toolkits and development methods.

2c Apply basic machine learning concepts and techniques to speech and language processing problems.

3a Given a well-formulated problem requiring natural language processing, design and implement a solution, by: goal setting; performance metric choice; decomposition of the task into modules or training stages; selection of models, tools, and resources; implementation; and analysis of performance and limitations.

**Main Topics**

**Language:** Basic properties of human language: symbolic, feature-based, vector-space and other representations of each level (acoustic, phonetic, prosodic, morphological, syntactic,
semantic, pragmatic), with attention to differences across languages, domains, genres, and speakers.

**Representations, Models and Algorithms:** Standard and for-purpose, for various core language processing tasks, including techniques for model training.

**Tools and Technologies:** Ways to wrangle corpora and analyze text and speech data in support of discovering knowledge from data, including sentiment analysis, filtering, and various classification tasks

**Systems:** The design and development of systems for search, question-answering, conversational interaction, user state identification, information extraction, and other applications.

**Prerequisites**  Linear Algebra (Math 3323 or equivalent), Probability and Statistics (Stat 3320 or equivalent), good programming skills (CS 3331 or equivalent), and at least nascent problem-solving and systems-integration skills. Graduate students from other departments may receive prerequisite waivers; see the instructor for permission.

Also helpful will be knowledge of Python, of machine learning techniques, and of basic linguistics concepts, but these are not required.

**Format**  Lectures, student presentations, discussions, in-class design exercises, lab time, project activities, project presentations, guest speakers. There is no remote option.


**Code Resources**  Well-known toolsets include NLTK, SpaCy, and gensim. At github, ekochmar and vajjala have useful examples in Jupyter.

**Course Website**  http://www.cs.utep.edu/nigel/slp/

**Assignments**  There will be a number of structured assignments, designed to reinforce knowledge and hone skills. Many assignments will be done in teams, but individual submission is the default. Writing quality is important, and rework may be required if not up to standard. Graduate students will have two additional assignments. Late assignments will receive at most 90% credit, less when the solution has been discussed in class, decreasing by 10% per day late.

**Grading**  Approximate point values:
- 190 assignments, including presentations
- 80 project
- 230 midterms and final exam
- 90 participation and quizzes
- 590 total

Grading will be on a points-earned basis (points above zero), rather than a points-off basis (points below expectation), and everything will be challenging. However letter grades will be assigned appropriately; in the past, the A/B break has been around 80% and the B/C break around 70%.

**Conduct, etc.**  Students are expected to be punctual, and to follow the spirit and letter of the UTEP Standards of Student Conduct and Academic Integrity policy https://www.utep.edu/student-affairs/osccr/student-conduct/academic-integrity.html. In particular, cooperation among
students and among teams is encouraged, but not to the extent that it interferes with each individual’s understanding or with learning-by-doing. Help given to and received from other students and sources should be noted in the assignment write-up.

If you have or suspect a disability and need accommodation please contact CASS at 747-5148 or at cass@utep.edu or visit Room 106 Union East Building.

Tests will generally be closed-book, except that one page of hand-written notes may be used for the first test, two for the second test, and three for the final. If you leave the classroom for any reason, your test will be graded on only what you did up until that time. No make-up exams will be given except under the conditions set forth in the Catalog. Students are free to attend class or not, bearing in mind that absence may annoy other students, interfere with learning, and result in a lower grade.

Important Dates
- August 29  Class begins
- September 28  Test 1 (tentative)
- November 2  Test 2 (tentative)
- November 23  Thanksgiving
- December 14, 4:00-6:45  Final Exam

Schedule

A. Introduction (1 day)
   - Overview of Language Applications
   - Review of Bayes Law and other Basic Mathematics
     - Exercise 1: Observe Language in Use (2 pts)
     - Exercise 2: Link Analysis (5 pts)

B. Rules, Features, and Classification (4 days)
   - Rules and Tendencies
   - Feature Design
   - Text Classification with a Linear Model
   - Model Evaluation
     - Assignment B (1,2,3): Simple Predictions (Surnames) (30 points)

C. Sequences, Context, Ngrams, and Language Modeling (3 days)
   - Edit Distances
   - The Noisy Channel Model
   - Bigrams and Beyond
   - Tagging
   - Sequence-to-Sequence Mapping
   - Language Modeling
   - Generative Modeling
     - Assignment C (1,2): Sequence Modeling (Surnames, again) (20)

D. Pattern Matching and Regular Expressions (1 day)
   - Regular Expressions in Python
   - Tokenization

Test 1 (5 days)

E. Words and Representations of Meaning
   - Bag-of-Words
   - Lexical Disambiguation
   - Vector-Space Similarity
d. Information Retrieval
   Assignment G (1, 2, 3, 4): Sentiment Analysis (10)

F. Grammatical Structures and Logical Forms (2 days)
   a. Dependencies
   b. Constituency, Context-Free Grammars, and Syntactic Ambiguity
   c. Chunks
   d. Logic-Based and Graph-Based Meaning Representations

G. Pretraining Language Models
   a. Word Embeddings and Context Vectors
   b. Large Language Model Training
   c. Generative Models, revisited
      Assignment F: Using Word Embeddings (20)
      Exercise 20: Prompting Large Language Models (5)

H. Sound, Phonetics, and Prosody (5 days)
   a. Acoustic Phonetics and Spectral Representations
   b. Articulatory Phonetics and Phonemes (Guest Lecture: Dr. Mueller)
   c. The Noisy Channel Model, again
   d. Speech Recognition
   e. Speech Synthesis
   f. Inferring Speaker States and Traits
      Exercise 4: Phonetic Observations (5)

I. Dialog Structure and Dialog Flow (3 days)
   a. Finite-state Dialog Management
   b. Retrieval-Based Dialog and Chatbots
   c. Endpointing and Turn Taking
   d. Pragmatics, Dialog Acts, and User Intentions
   e. Natural Language Generation
   f. Advantages and Disadvantages of Natural Language Interfaces
      Exercise 6: Dialog States and Flow Authoring (5)

J. Other Applications (1 day)
   a. Search (guest lecture: Chris Mendoza)
   b. Spelling Correction
   c. Spam Filtering
   d. Machine Translation
   e. Call-Center Analytics
   f. Information Extraction (Entity Extraction and Linking)
   g. Question Answering
   h. Language Proficiency Assessment
   i. Tutoring and Training
   j. Summarization
      Exercise 9: Ethical Issues (5)

Other Assignments
   Assignment P: Final Project (80)
   Assignment X: Present a Research Paper (graduate students only) (30)
   Assignment Y: Research-Project Mini-Proposal (graduate students only) (20)
   Exercise 10: A Question for the final exam (5)
Readings

Jurafsky and Martin, *Speech and Language Processing*, 3rd edition, January 2023 version
- Chapter 2, except byte-pair encoding  (Regular Expressions, etc.)
- Chapter 3 through 3.5.1  (Language Modeling)
- Chapter 4 through 4.5, 4.7-4.8  (Classification, Sentiment)
- Chapter 6 through 6.5, 6.8 - 6.12  (Vector Space Models, Embeddings)
- Chapter 8 through 8.3  (Part-of-Speech Tagging)
- Chapter 10: introduction  (Pretrained Models)
- Chapter 11: introduction  (Fine Tuning)
- Chapter 13 through 13.1, historical notes  (Machine Translation)
- Chapter 14 through 14.1.4, introduction to 14.2, 14.3  (Question Answering, Info. Retrieval)
- Chapter 15 first 10 pages, 15.4.5, 15.6  (Dialog Systems and Chatbots)
- Chapter 16: 16.5, 16.7  (Speech Recognition Evaluation, other Speech Tasks)
- Chapter 17: first 5 pages and 17.5  ( Constituency, Context-Free Grammars, and Parsing)
- Chapter 18 through 18.1  (Dependency Structures)
- Chapter 19 through 19.3  (Logical Representations of Sentence Meaning)
- Chapter 21 through 21.2.2  (Information Extraction)
- Chapter 23: first 10 pages  (Word Senses)
- Chapter 25 through 25.2, 25.6  (Lexicons for Sentiment)
- Chapter 28  (Phonetics)

Manning, Raghavan, Schutze: *Introduction to Information Retrieval*
- Chapter 19, through 19.4  (Web Search Basics)
- Chapter 21 through 21.2.2  (Link Analysis)

Giangola, Balogh, and Cohen, *Voice User Interface Design*
- Chapter 3  (Overview of the Methodology)

Sarikaya: The Technology behind Personal Digital Assitants, pages 72-72 (reactive assistance)

Links for most are on the course homepage.