

CS4317, Language Systems (3-2) -- Syllabus

Catalog description

This course introduces the computational aspects of processing human languages with attention to military text systems. Topics include lexicography, morphology, grammars, parsing, semantics, deep neural-network architectures, large language models, language generation, language translation, connecting to databases, AI agents and agentic workflows. Prerequisite: CS3021 or two courses in Python programming, and CS3310.

Course objectives

The course supports the Educational Skill Requirements 3 (Programming and Systems), 4 (Practices and Fundamental Principles), and 5 (Program/Project Management) for the Computer Science degree. In particular:

From ESR 3: “The officer will be proficient in programming and programming languages, and in analyzing large software-intensive systems. The officer will demonstrate advanced knowledge and understanding of programming practices and programming paradigms.... The officer will demonstrate advanced knowledge and understanding of system modules, interfaces, testing and validation, risk factors, and project management.”

From ESR 4: “Officers will have competence in computer science practices and principles including design and mechanics of computing and their applications to military needs. They will be proficient in core technologies including algorithms, architectures, operating systems, data structures, databases, software engineering, reverse translation, intelligent systems, and human-computer interactions.”

From ESR 5: “This includes, but is not limited to, planning and implementing a major programming project and developing the appropriate technical and acquisition documentation, performing financial, cost-benefit and tradeoff analyses and performing required planning, programming and budgeting actions in order to develop means to exploit technology advantages in a network-centric environment to achieve operational objectives.”

Learning objectives

- Apply the key concepts and terms in computational natural-language processing to particular problems.
- Identify syntactic and semantic constraints between sense pairs of natural-language words.
- Understand how large language models are used for generative AI.
- Implement language systems using packages exploiting large language models.
- Implement retrieval-augmented generation from databases to support large language models.
- Implement agentic workflows and understand their requirements.
- Identify metrics for large language models and plan their application.
- Understand methods for optimization of large language models for new applications.

Course topics and schedule

- Shallow linguistic processing (1 week)
 - Dictionaries, word lists, and thesauri
 - Morphology and stemming
 - Unigrams, bigrams, and trigrams; text classification
 - Word senses, vector semantics
 - Topic classification
 - Sentiment classification
 - Language identification
 - Speech
- Language processing with neural networks (3 weeks)
 - Recurrent neural networks for natural-language processing
 - The concept of context, large language models
 - Transformers for natural-language processing
 - Generative AI for natural language tasks
 - Prompt engineering
 - Choosing foci of discourse
 - Templates for linguistic structures
 - Language generation
 - Readability and comprehensibility planning
 - Machine translation
- Retrieval-augmented generation (2 weeks)
 - Large language model context length
 - Vector databases
 - Retrieval-augmented generation architectures
 - Performance metrics
- Agentic workflows (2 weeks)
 - Agentic architectures
 - Preconditions, postconditions, and reasons
 - Defining goal-oriented agents
 - Agentic tool usage
 - Performance metrics
- Optimization of large language models (1 week)
 - Specialized datasets
 - Methods for optimizing a model for a dataset
 - Metrics for tuning
- Class project presentations (1 week)

Course assignments

Three homework assignments including laboratory work, three tests, and a class project of at least 20 hours of work. Laboratory work will involve programming tasks with language tools. There will be a task with shallow processing, a task with large language models, a task with training and test a language generator based on large language models, a task with setting up retrieval-augmented, and a task on speech processing. Programming assignments will use Python.

Homework must be done by each student on their own without consulting anyone besides the instructor. The class project can be done in groups or individually. The project topic must be something related to

language systems and is approved by the instructor. The penalty is 15% for assignments received after the due date.

Textbook

Raj, Arun, R. (2024) Mastering large language models with Python (English edition). Orange Education.

Justification

This course was previously taught from approximately 1995 to 2015, and we are asking to reinstate it, as it was removed from the catalog in 2022. We are reinstating as a required course for the new Master's degree in Artificial Intelligence, with substantial updates to cover recent developments. It will also be an elective for the M.S. degree in Computer Science for the Artificial Intelligence specialization.

Duplication

No similar courses are taught at NPS.

Scheduling

No additional instructors are required to teach this once a year.

Scheduling

Every Summer quarter starting in 2025.