NAAIC

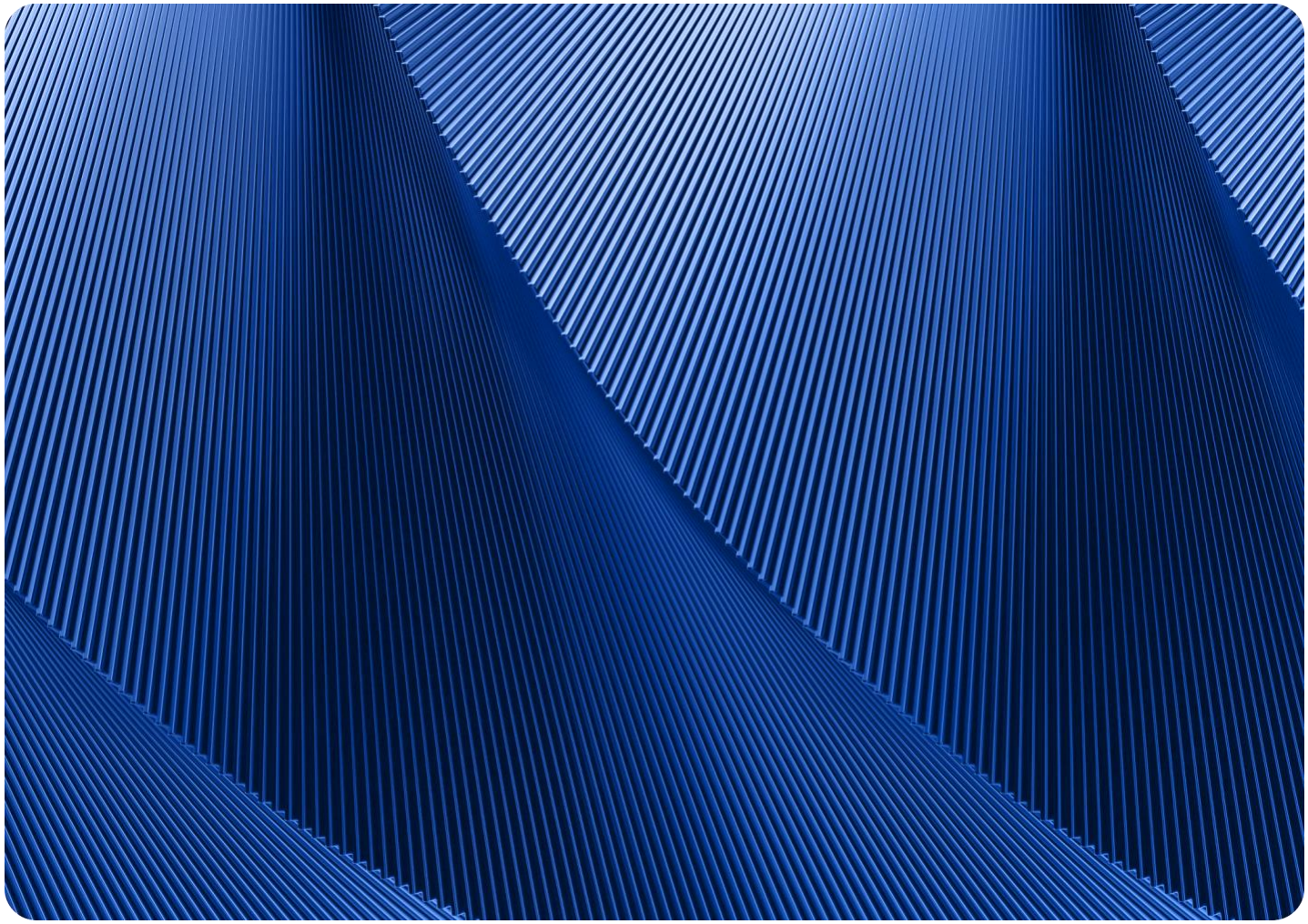
Resource provided by

HCC | HOUSTON COMMUNITY COLLEGE

Course Syllabus

Natural Language Processing

This intermediate course explores the principles, methods, and applications of Natural Language Processing (NLP) and language understanding in artificial intelligence and robotics



NAAIC

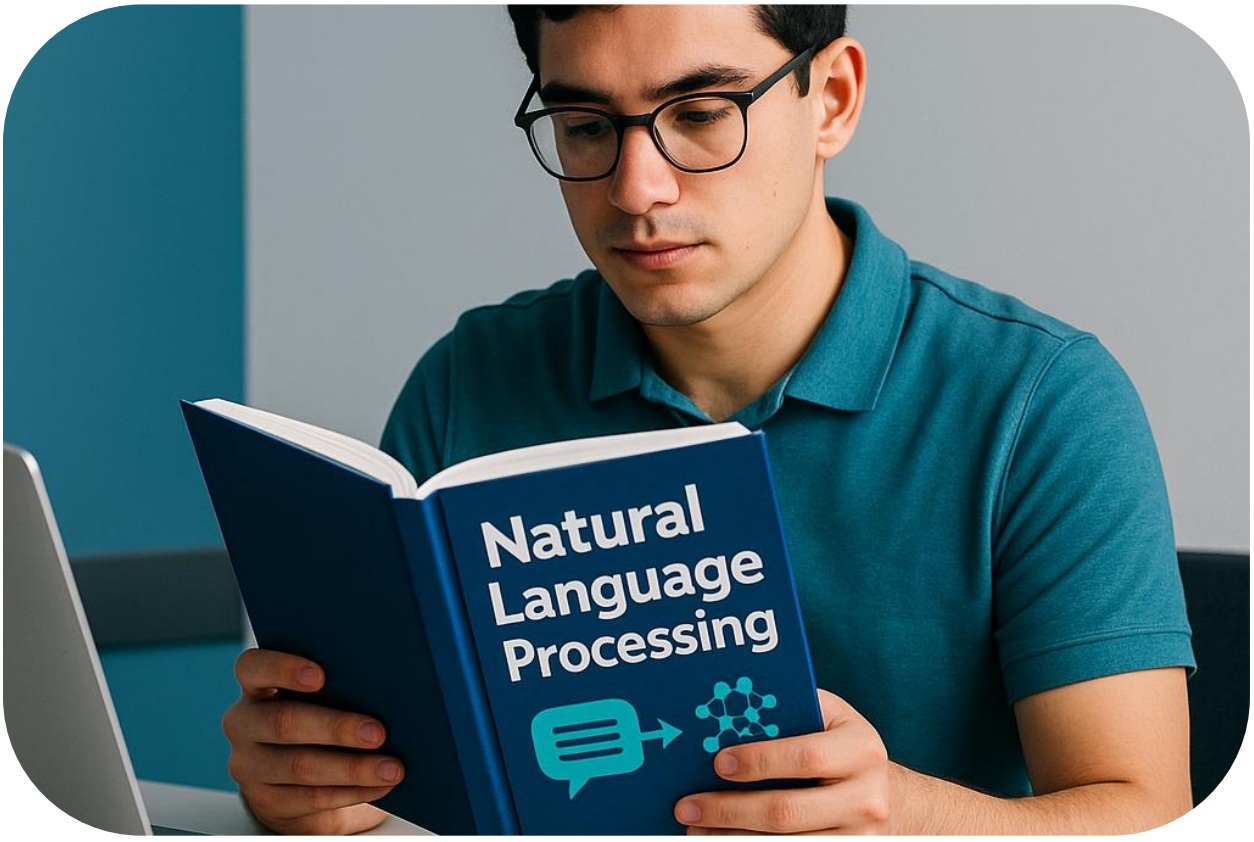
Table of Contents

NAAIC

Course Description

This intermediate course explores the principles, methods, and applications of Natural Language Processing (NLP) and language understanding in artificial intelligence and robotics. Students will investigate how machines process, analyze, and generate human language in various forms, including text and speech. Emphasis is placed on both foundational concepts and practical skills, with hands-on experience using contemporary tools and real-world datasets. The course highlights ethical considerations and the evolving role of language technologies in AI-driven systems.

Prerequisites:
Python programming language
Machine Learning course



NAAIC

Course Competencies

Competency 1
Summarize the historical evolution, foundational concepts, and industrial impact of NLP and NLU
Competency 2
Differentiate among classic NLP algorithms, NLU techniques, and their appropriate use cases in AI and robotics
Competency 3
Apply classic machine-learning methods to construct end-to-end text analytics pipelines
Competency 4
Analyze linguistic features (syntax, semantics, pragmatics) to extract structured knowledge from unstructured language data
Competency 5
Evaluate the effectiveness, fairness, and ethical implications of NLP systems using quantitative metrics and critical reflection
Competency 6
Design and create a functional NLP application (e.g., chatbot or QA system) that integrates preprocessing, modeling, and deployment best practices

NAAIC

Instructional Resources

No prescribed textbook is required due to the rapidly evolving nature of artificial intelligence and NLP. Instructors will curate up-to-date articles, papers, videos, and other resources, available via Canvas, to provide the latest advancements and diverse perspectives. This approach ensures flexibility to tailor content to class needs and incorporate emerging topics.

Platforms and Tools:
Kaggle, GitHub, Google Colab, Amazon SageMaker, VS Code, Hugging Face

Recommended Reading:
See Supplemental Books (Optional) and Supplemental Papers (Optional) sections.

Grading Schema

Assignment Type	Percentage of Grade
General Assignments	20%
Labs/Hands-On Activities	25%
Midterm Project	20%
Capstone Project	25%
Course Portfolio	10%

NAAIC

Supplemental Information

General Assignments (20%)

Objective: To reinforce theoretical concepts and develop analytical skills through weekly assignments including case studies, research discussions, problem sets, and design tasks directly related to module content.

Labs/Hands-On Activities (25%)

Objective: To provide students with practical understanding of natural language processing through guided implementation exercises, coding workshops, and interactive demonstrations using industry-standard tools and frameworks.

Midterm Project (20%)

Objective: Individual or group project demonstrating integration and application of knowledge from the first half of the course. Students will design and implement a focused NLP solution addressing a specific problem or use case.

Capstone Project (25%)

Objective: To provide students with comprehensive hands-on experience in creating a practical, real-world application. Students will integrate multiple NLP techniques learned throughout the course to develop an end-to-end solution with documentation and presentation.

Course Portfolio (10%)

Objective: Individual dynamic document continuously updated and maintained on GitHub, serving as a comprehensive record of learning progress, code repositories, reflections, and achievements throughout the course. This portfolio demonstrates professional development and technical growth**.**

NAAIC

Course Outline

Week starting on / Module	Module Topic
Week 1	Mod 01: Introduction to NLP and NLU
Week 2	Mod 02: Text Preprocessing and Cleaning
Week 3	Mod 03: Introduction to Audio and Speech Processing
Week 4	Mod 04: Text Representation
Week 5	Mod 05: Part-of-Speech Tagging
Week 6	Mod 06: Syntax, Parsing, and Semantic Analysis
Week 7	Mod 07: Sentiment and Emotion Analysis
Week 8	Mod 08: Text Classification
Week 9	Mod 09: Named Entity Recognition (NER)
Week 10	Mod 10: Topic Modeling
Week 11	Mod 11: Language Models
Week 12	Mod 12: Machine Translation
Week 13	Mod 13: Conversational AI and NLU
Week 14	Mod 14: Summarization and Question Answering
Week 15	Mod 15: Course Review & Modern Trends, Responsible NLP
Week 16	Mod 16: Final Exam/Project Submission

Supplemental Books (Optional)

Lane, H., & Dyshel, M. (2025). *Natural Language Processing in Action (2nd ed.)*. Manning Publications.
Recommended for Modules 1, 5, 7, 9, 11, 13, and 14 for practical NLP techniques, including preprocessing, sentiment analysis, language models, and conversational AI.

Lipenkova, J. (2025). *AI-Powered Search*. Manning Publications.
Recommended for Modules 10, 13, and 14 for retrieval-augmented generation (RAG) and question-answering systems.

Patel, J. (2024). *Transformers for Natural Language Processing (3rd ed.)*. Packt Publishing.
Recommended for Modules 4, 6, 11, and 12 for word embeddings, transformers, and machine translation.

Hapke, H. M., & Lane, H. (2025). *Building LLMs for Production*. Manning Publications (MEAP).
Recommended for Modules 11, 13, and 16 for deploying language models and capstone project development.

Vasiliev, Y. (2020). *Natural Language Processing with Python and spaCy*. No Starch Press.
Recommended for Modules 2, 5, 6, and 9 for text preprocessing, POS tagging, dependency parsing, and NER.

Supplemental Papers (Optional)

Grattafiori, A., et al. (2024). The Llama 3 herd of models. arXiv. https://arxiv.org/abs/2407.12345
Recommended for Modules 11 and 13 for advanced language models and conversational AI.

Campino, J. (2024). Unleashing the transformers: NLP models detect AI writing in education. *Journal of Computers in Education*, 11(2), 789–802. https://doi.org/10.1007/s12345-024-56789-0
Recommended for Modules 7 and 8 for sentiment analysis and text classification.

Gu, A., & Dao, T. (2024). Mamba: Linear-time sequence modeling with selective state spaces. arXiv. https://arxiv.org/abs/2403.12345
Recommended for Module 11 for efficient language model architectures.

Microsoft Research. (2024). GraphRAG: Knowledge graphs for retrieval-augmented generation. arXiv. https://arxiv.org/abs/2406.12345
Recommended for Modules 13 and 14 for RAG in conversational AI and question-answering systems.

Gupta, A., & Nguyen, T. (2025). Ethical considerations in large language model-driven data science workflows. *Journal of Artificial Intelligence Ethics*, 5(1), 45–62. https://doi.org/10.1007/s12345-025-67890-1
Recommended for Module 15 for ethical frameworks in NLP applications.

Instructors may consider adopting these books and papers to complement curated resources, aligning with course objectives, SLOs, and teaching style.