Course Description Form

1. Course Name:

Data Coding and Data Compression

2. Course Code:

3. Semester / Year:

Second Semester/ 2025-2026

4. Description Preparation Date:

31march 2026

5. Available Attendance Forms:

Attendance Study

6. Number of Credit Hours (Total) / Number of Units (Total)

30

7. Course administrator's name (mention all, if more than one name)

Name: Assistant Prof. Abeer Khalid Al-Mashhadany

Email: abeer.khalidahmed@nahrainuniv.edu.iq

8. Course Objectives

Course Objectives	• Natural Language Processing (NLP) is a branch of Artificial Intelligence that studies how ma that can make sense of text and perform tasks like translation, summarization, gran mar ch Language Models, Naive Bayes and Sentiment Classification, Vector Semantics and Em Constituency Grammars, Logical Representations of Sentence Meaning, Machine Tra slator and Dialogue Systems and Chatbots.

9. Teaching and Learning Strategies

Strategy	• **Books, theoretical lectures and references to helpful websites** • **Providing the electronic presentation of the lecture…as an aid to clarifi ation** • **Explanation on the board and solving examples with integrated details** • **Design and implementation of practical programs for programming algo ithms** • **Solve multiple examples and engage the student by providing quick exer motivating students to follow up and correct each other.** • **Homework examples from the article** • **In addition to homework that focuses on strengthening the student's pr gram**

10. Course Structure

Week	Hours	Required Learning Outcomes	Unit or subject name	Learn

Week No.	Date	Title	Sub-Titles	Refere
1		**Introduction: purposes and tasks**		

2		N-gram Language Models	• The role of language models. • Simple N-gram models. • Estimating parameters and smoothing. • Evaluating language models.	
3		Naive Bayes and Sentiment Classification	• Naive Bayes Classifiers • Optimizing for Sentiment Analysis • Naive Bayes for other text classification tasks • Test sets and Cross-validation	
4				**Idle**
5		Vector Semantics and Embeddings	• Lexical Semantics • Vector Semantics • Words and Vectors • TF-IDF: Weighing terms in the vector • Optional: Pointwise Mutual Information (PMI) • Semantic properties of embeddings • Bias and Embeddings • Evaluating Vector Models	
6		**Mid 1**		
7		Part-of-Speech Tagging	• English Word Classes • The Penn Treebank Part-of-Speech Tagset • Part-of-Speech Tagging • HMM Part-of-Speech Tagging • Maximum Entropy Markov Models • Part-of-Speech Tagging for Morphological Rich Languages	
8		Constituency Grammars	• Context-Free Grammars • Grammar Rules for English • Treebanks • Grammar Equivalence and Normal Form • Lexicalized Grammars	

9		Training on Constituency Grammars	• Construct Grammar and train to solve NLP problems	
10		**idle**		
11		Logical Representations of Sentence Meaning	• Computational Desiderata for Representations • Model-Theoretic Semantics • First-Order Logic • Event and State Representations • Description Logics	
12		**Mid 2**		
13		Logical Representations of Sentence Meaning	Construct model from real information and train to solve NLP problems	
14		Real NLP problem	• Apply all given NLP tools to solve the problem	
15		**Seminar**		

11. Course Evaluation

Project	10
Seminar	05
Homeworks	05
Mid 1	05
Mid 2	05
Final Exam	70 marks

12. Learning and Teaching Resources

Required textbo (curricular books, if any	- "Speech and Language Processing", by Daniel Jurafsky, Jar es H 2024. - Natural Language Processing with Python, by Steven Bird, Ew 2024.
Main references (sourc	- "Speech and Language Processing", by Daniel Jurafsky, Ja Copyright c 2024.

	- Natural Language Processing with Python, by Steven Bird © 2024.
Recommended books and references (scientific journals, reports…)	- "Speech and Language Processing", by Daniel Jurafsky, Ja Copyright c 2024. - Natural Language Processing with Python, by Steven Bird © 2024.