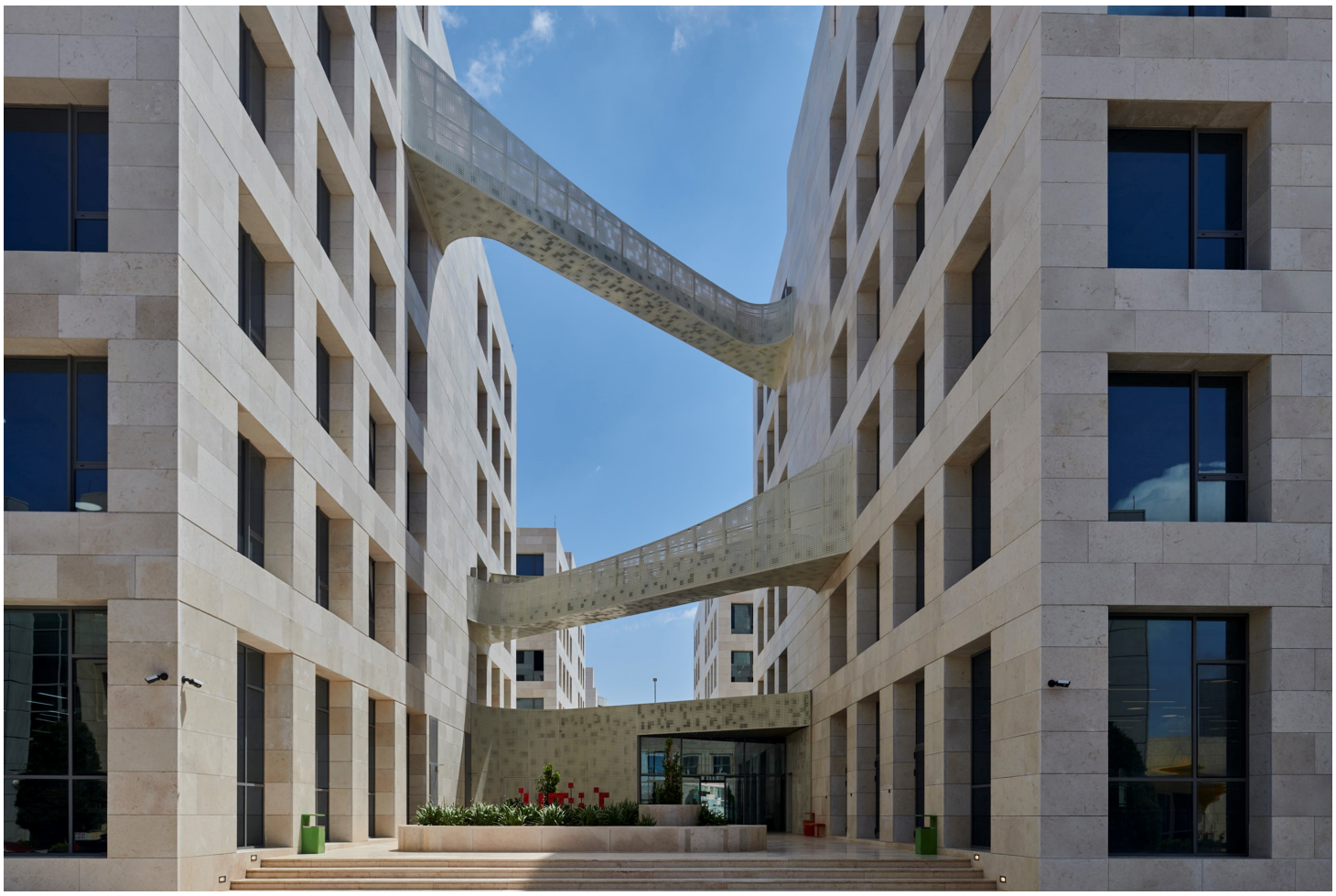
جامعة الحسين التقنية
Al Hussein Technical University

ASSIGNMENT BRIEF

HTU Course No: 10204351	**HTU Course Name:** Natural Language Processing
BTEC Unit Code:	**BTEC UNIT Name:**

Version: 3



Student Name/ID Number/Section	
HTU Course Number and Title	10204351 Natural Language Processing
BTEC Unit Code and Title	
Academic Year	2023-2024 Spring
Assignment Author	Raneem Qaddoura
Course Tutor	Raneem Qaddoura - Yara Alharahsheh
Assignment Title	Implement a natural language processing application
Assignment Ref No	1
Issue Date	14/04/2024
Formative Assessment dates	From 26/05/2024 to 30/05/2024
Submission Date	12/06/2024
IV Name & Date	Rami Ibrahim 13/04/2024

Submission Format

Part 1: In-Class Examination

- In-class closed book, closed notes examination.
- One sheet with the necessary equations, and you are allowed to use a calculator.
- Answers must be clear and coherent.
- Show detailed steps; final answers alone are not accepted.
- If answers span multiple pages, ensure your name and student number are on each paper.
- Complete and sign the student declaration form paper for the exam.

Part 2: Assignment Submission

- Submit to the university's eLearning system via https://elearning.htu.edu.jo by the specified deadline.
- The submission is a Source code file (ipynb) and a presentation file (pptx).
- No compressed files or folders (no .zip or .tar extensions).
- Plagiarism will result in course failure.
- Signed declaration Form (Word Document).

Unit Learning Outcomes

LO1 Understand the basic concepts of natural language processing and morphological analysis

LO2 Analyze various text encoding and vectorization techniques for natural language processing tasks.

LO3 Investigate and apply a range of natural language processing traditional techniques across different applications.

LO4 Demonstrate the application of neural language models across various natural language processing tasks.

Assignment Brief and Guidance

Part 1: In-Class Examination (Understanding Natural Language Processing Techniques)

As an NLP engineer at a company, it is essential to possess a thorough understanding of various NLP concepts and techniques. This includes exploring the historical evolution, diverse techniques, and wide-ranging applications of natural language processing while analyzing associated bias concerns. Additionally, you need to explore a range of linguistic subfields to analyze and describe diverse aspects of language. This part requires you to perform morphological analysis incorporating various linguistic elements. This also includes understanding word representation techniques employed in text encoding for

natural language processing tasks and understanding the traditional NLP models and neural language models and their applications across various domains and tasks.

An in-class exam is scheduled for **Sunday, June 2, 2024 at 3:30 PM**.

Part 2: Assignment Submission (Real-World NLP Application)

As an NLP engineer entrusted with the development of a comprehensive application and corresponding presentation to show your work. Your task involves working with the News agency that aims to classify the news into different categories based on the title of the news.

You need to submit a project that includes the following steps:

- Implement preprocessing methods and relevant packages.
- Apply diverse word representation techniques to address real-life applications.
- Implement traditional NLP models and neural language models to address real-life applications.

You also need to submit a presentation that includes the following:

- Show the different preprocessing methods and packages that you have used in your project and evaluate the effectiveness of the selected methods and packages in solving real-world applications.
- Compare and contrast the different word representation techniques used in your project.
- Conduct a comparative analysis of different traditional NLP models and neural language models to discern their respective strengths and weaknesses.
- Evaluate the effectiveness of word representation techniques in addressing a real-life application.

The assignment submission is scheduled for **Wednesday, June 12, 2024 at 10:00 AM**.

Learning Outcomes and Assessment Criteria			
Learning Outcome	**Pass**	**Merit**	**Distinction**
LO1 Understand the basic concepts of natural language processing and morphological analysis	**P1** Explore the historical evolution, diverse techniques, and wide-ranging applications of natural language processing, while analyzing associated bias concerns. **P2** Explore a range of linguistic subfields to analyze and describe diverse aspects of language.	**M1** Perform morphological analysis incorporating various linguistic elements. **M2** Implement preprocessing methods and relevant packages to address real-life applications within natural language processing.	**D1** Evaluate the effectiveness of the selected preprocessing methods and packages in solving real-world applications.
LO2 Analyze various text encoding and vectorization techniques for natural language processing tasks.	**P3** Understand word representation techniques employed in text encoding for natural language processing tasks. **P4** Apply diverse word representation techniques to address real-life applications.	**M3** Compare and contrast different word representation techniques used in text encoding for natural language processing tasks.	**D2** Evaluate the effectiveness of word representation techniques in addressing real-life applications.
LO3 Investigate and apply a range of natural language processing traditional techniques across different applications.	**P5** Understand the traditional NLP models and their applications across various domains. **P6** Implement traditional NLP models to address real-life applications.	**M4** Conduct comparative analysis of different traditional NLP models to discern their respective strengths and weaknesses.	

LO4 Demonstrate the application of neural language models across various natural language processing tasks.	**P7** Acquire comprehension of neural language models and their applications across diverse natural language processing tasks. **P8** Implement neural language models to solve real-life applications effectively.	**M5** Conduct comparative analysis of different neural language models to discern their respective strengths and weaknesses.	