هيئة تقويم التعليم والتدريب
Education & Training Evaluation Commission

2023
TPG-153

Course Specification
(Postgraduate)

Course Title: Natural Language Processing
Course Code: 501836-3
Program: Master in Artificial Intelligence
Department: Computer Science
College: Computers and Information Technology
Institution: Taif University
Version: V2
Last Revision Date: 5 May 2024

قسم علوم الحاسب
Computer Science Department
TU
جامعة الطائف
TAIF UNIVERSITY

Table of Contents

A. General information about the course:

1. Course Identification:

1. Credit hours: (3)				
2. Course type				
A.	☐University	☐College	☒ Department	☐Track
B.	☐Required		☒ Elective	
3. Level/year at which this course is offered: (Year: 1, Level: 1)				
4. Course general Description:				
This course aims at studying theories and concepts that are required to learn computers human languages to perform useful tasks such as machine translation, text summarization, question answering and sentiment analysis. The course commences with learning text processing fundamentals, including tokenization, stemming and lemmatization. The course examines theoretical foundations of natural language processing (NLP) followed by a focus on applied aspects of NLP. The student will get familiar with machine learning techniques that obtain leading results on the problem of NLP. The students will learn how to represent words and text, how to use the well-known NLP tools for creation NLP application and use information extraction to solve real-world problems. This will be reinforced through examples and practical projects.				
5. Pre-requirements for this course (if any)**:**				
None.				
6. Co-requirements for this course (if any)**:**				
None.				
7. Course Main Objective(s):				
The objectives of this course are as follows: ▪ Understand the basic NLP techniques and their applications in different problem contexts. ▪ Ability to analyze a problem in terms of constructs and techniques from NLP. ▪ Ability to develop NLP applications for solving problems in the area of information extraction and social media analysis by using state-of-the-art tools and API's. ▪ Ability to apply statistical and machine learning techniques in NLP applications.				

2. Teaching Mode: (mark all that apply)

No	Mode of Instruction	Contact Hours	Percentage
1	Traditional classroom	**36**	**80%**
2	E-learning	**9**	**20%**
3	Hybrid ▪ Traditional classroom ▪ E-learning	**0**	**0%**

No	Mode of Instruction	Contact Hours	Percentage
4	Distance learning	**0**	**0%**

3. Contact Hours: (based on the academic semester)

No	Activity	Contact Hours
1.	**Lectures**	**45**
2.	**Laboratory/Studio**	**-**
3.	**Field**	**-**
4.	**Tutorial**	**-**
5.	**Others (Mid-Term and Final Exams)**	**-**
	Total	**45**

B. Course Learning Outcomes (CLOs), Teaching Strategies and Assessment Methods:

Code	Course Learning Outcomes	Code of CLOs aligned with program	Teaching Strategies	Assessment Methods
1.0	**Knowledge and understanding**			
1.1	Explore the basic concepts and algorithms of natural language processing.	**K1**	Lecture, Brainstorming, Self-learning	**Direct:** Quiz, Assignment **Indirect:** Survey
1.2	Understand the underlying hypothesis and state-of-the art models for semantic representations of text.	**K1**	Lecture, Brainstorming, Self-learning	**Direct:** Quiz, Assignment, Exam **Indirect:** Survey
2.0	**Skills**			
2.1	Analyze basic natural language processing problems and apply basic tools and principles to solve them.	**S1**	Lecture, Project, Self-learning	**Direct:** Quiz, Project, Exam **Indirect:** Survey
2.2	Design and implement appropriate machine learning models to build solutions for natural language processing problems.	**S2**	Lecture, Project, Self-learning	**Direct:** Quiz, Project, Exam **Indirect:** Survey
2.3	Evaluate various deep learning models for various real-world problems in natural language processing and implement potential solutions.	**S2**	Lecture, Self-learning	**Direct:** Quiz, Project, Exam **Indirect:** Survey

Code	Course Learning Outcomes	Code of CLOs aligned with program	Teaching Strategies	Assessment Methods
3.0	**Values, autonomy, and responsibility**			
3.1	Recognize professional responsibilities and make informed judgments in computing practice based on legal and ethical principles	**V1**	Discussion, Project	**Direct:** Project **Indirect:** Survey
3.2	Function effectively as a member or leader of a team engaged in activities appropriate to the field of natural language processing.	**V2**	Discussion, Project	**Direct:** Project **Indirect:** Survey

C. Course Content:

No	List of Topics	Contact Hours
1.	Introduction to Natural Language Processing	**3**
2.	Regular Expressions, Text Preprocessing and Normalization. Tokenization, Stopwords, Stemming and Lemmatization.	**3**
3.	N-gram Language Models	**3**
4.	Naïve Bayes, Logistic Regression, Text Classification, Sentiment Analysis.	**6**
5.	Vector Semantics and Embeddings. TF-IDF, Pointwise Mutual Information, Word2vec, Semantic properties of embeddings.	**6**
6.	Neural Networks and Neural Language Models.	**3**
7.	Sequence Labelling for Parts of Speech and Named Entities.	**3**
8.	Transformers and Pretrained Language Models.	**6**
9.	Fine-Tuning and Masked Language Model.	**3**
10.	NLP Applications. Machine Translation, Question Answering systems, etc.	**9**
Total		**45**

D. Students Assessment Activities:

No	Assessment Activities *	Assessment timing (in week no)	Percentage of Total Assessment Score
1.	Project	Week 14	20%

No	Assessment Activities *	Assessment timing (in week no)	Percentage of Total Assessment Score
2.	Assignments	Week 3, 12	10%
3.	Quizzes	Week 4, 10	10%
4.	Mid Term Exam	Week 8	20%
5.	Final Exam	Week 16	40%

*Assessment Activities (i.e., Written test, oral test, oral presentation, group project, essay, etc.)

E. Learning Resources and Facilities:

1. References and Learning Resources:

Essential References	▪ SPEECH and LANGUAGE PROCESSING: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, written by Daniel Jurafsky and James H. Martin. Third Edition, McGraw Hill, 2023.
Supportive References	▪ Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit, by Steven Bird, Ewan Klein, and Edward Loper, (O'Reilly 2009, website 2018). http://www.nltk.org/book/ ▪ Foundations of Statistical Natural Language Processing by Chris Manning and Hinrich Schütze, MIT Press. Cambridge, MA: May 1999. https://nlp.stanford.edu/fsnlp/
Electronic Materials	▪ The world's most trusted open ecosystem for sourcing, building, and deploying data science and AI initiatives: https://www.anaconda.com/ ▪ http://www.nltk.org/
Other Learning Materials	▪ Links provided by the instructor. ▪ Handouts and Presentations Slides prepared by the instructor. ▪ Blackboard.

2. Educational and Research Facilities and Equipment Required:

Items	Resources
facilities (Classrooms, laboratories, exhibition rooms, simulation rooms, etc.)	▪ Classroom (20 students/class) ▪ Computer labs
Technology equipment (Projector, smart board, software)	▪ Video projector / data show ▪ White board
Other equipment (Depending on the nature of the specialty)	▪ To be announced during the course!

F. Assessment of Course Quality:

Assessment Areas/Issues	Assessor	Assessment Methods
Effectiveness of teaching	Students Coordinator	Indirect (Course exit survey) Indirect (Feedback from Course Coordinator)
Effectiveness of students assessment	Faculty member Coordinator	Indirect (Feedback from Faculty member) Indirect (Feedback from Course Coordinator)
Quality of learning resources	Students Faculty member Coordinator Council Curriculum Committees	Indirect (Course exit survey) Indirect (Feedback from Faculty member) Indirect (Feedback from Course Coordinator) Indirect (Feedback from council) Indirect (Feedback from Graduate Committees)
The extent to which CLOs have been achieved	Students Faculty member Coordinator Curriculum Committees	Indirect (Course exit survey) Indirect (Feedback from Faculty member/ Course Coordinator/ Graduate Committee)
Other	-	-

Assessor (Students, Faculty, Program Leaders, Peer Reviewer, Others (specify)

Assessment Methods (Direct, Indirect)

G. Specification Approval Data:

COUNCIL /COMMITTEE	**GRADUATE PROGRAMS COMMITTEE – CS DEPT.**
REFERENCE NO.	**V2**
DATE	**5/5/2024**

قسم علوم الحاسب
Computer Science Department
TU جامعة الطائف TAIF UNIVERSITY