



UNIVERSITY WITH A PURPOSE

Format No. QSP/7.1/01.F01 (B)

Issue No. 05, Rev. No. 5, Dated: Jan 1, 2017

UNIVERSITY OF PETROLEUM & ENERGY STUDIES

School of Computer Science

Dehradun

COURSE PLAN

Programme : B. Tech CSE-AI&ML

Course : Computational Linguistics and Natural Language Processing Lab

Subject Code : CSEG3124

No. of credits : 1

Semester : VI

Session : Jan 2022- June 22

Batch : 2019-2023

Prepared by : Rahul Kumar Singh

Email : rk.singh@ddn.upes.ac.in

Approved By

HOD/ Prog. Head

UPES Campus

“Energy Acres”

P.O. Bidholi, , Dehradun

Tel : +91-135-2770137

Fax : +91 135- 27760904

GUIDELINES TO STUDY THE SUBJECT

Instructions to Students:

1. Go through the 'Syllabus' in the Black Board section of the web-site(<https://learn.upes.ac.in>) in order to find out the Reading List.
2. Get your schedule and try to pace your studies as close to the timeline as possible.
3. Get your on-line lecture notes (Content, videos) at Lecture Notes section. These are our lecture notes. Make sure you use them during this course.
4. Check your blackboard regularly
5. Go through study material
6. Check mails and announcements on blackboard
7. Keep updated with the posts, assignments and examinations which shall be conducted on the blackboard
8. Be regular, so that you do not suffer in any way
9. **Cell Phones and other Electronic Communication Devices:** Cell phones and other electronic communication devices (such as Blackberries/Laptops) are not permitted in classes during Tests or the Mid/Final Examination. Such devices **MUST** be turned off in the class room.
10. **E-Mail and online learning tool:** Each student in the class should have an e-mail id and a pass word to access the LMS system regularly. Regularly, important information – Date of conducting class tests, guest lectures, via online learning tool. The best way to arrange meetings with us or ask specific questions is by email and prior appointment. All the assignments preferably should be uploaded on online learning tool. Various research papers/reference material will be mailed/uploaded on online learning platform time to time.
11. **Attendance:** Students are required to have minimum attendance of 75% in each subject. Students with less than said percentage shall **NOT** be allowed to appear in the end semester examination.

This much should be enough to get you organized and on your way to having a great semester! If you need us for anything, send your feedback through e-mail rk.singh@ddn.upes.ac.in. Please use an appropriate subject line to indicate your message details.

There will no doubt be many more activities in the coming weeks. So, to keep up to date with all the latest developments, please keep visiting this website regularly.

COURSE PLAN

A. PREREQUISITE:

- Basic programming skills in python, artificial intelligence and machine learning

B. PROGRAM OUTCOMES (POs) and PROGRAM SPECIFIC OUTCOMES (PSOs) for :

B1. PROGRAM OUTCOMES (POs)

1. Engineering knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. Problem analysis: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3. Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. Conduct investigations of complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. Modern tool usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
10. Communication: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend

and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

11. Project management and finance: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

B2. Program Specific Outcomes (PSOs)

1. Perform system and application programming using computer system concepts, concepts of Data Structures, algorithm development, problem solving and optimizing techniques.
2. Apply software development and project management methodologies using concepts of front-end and back-end development and emerging technologies and platforms.
3. To create & develop most efficient solutions by applying machine learning with analytical emphasis on industrial and research problems.

COURSE OUTCOMES for Cryptography and Network Security lab: At the end of this course student should be able to

CO1. Understand the basic concepts of text classification and lexical analysis.

CO2. Understanding the basic concepts of Natural Language Processing.

CO3. Understanding the basic concepts of information retrieval.

Table: Correlation of POs and PSOs v/s COs

Course Outcomes	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12	PSO 1	PSO 2	PSO 3
CO1	1	1	2									1			3
CO2	1	1	2	1	1							1			3

CO3	1	1	2	1	1							1			3
Average	1	1	2	1	1							1			3

1: Slight (Low)

2: Moderate (Medium)

3: Substantial (High)

C. PEDAGOGY

1. Power Point Presentation,

2. Experimental learning with continuous evaluation

D. COURSE COMPLETION PLAN

Total Lab sessions	10
Total Viva	03

One Session =120 minutes

E. EVALUATION & GRADING

Students will be evaluated continuously throughout the course based on the following:

- | | | |
|-------------------------|---|-----|
| 1) Performance & Record | - | 50% |
| 2) Viva Voce | - | 50% |

F1. Performance & Record: WEIGHTAGE - 50%

. The lists of activities performed under the experiments are detailed clearly in Section-F. F2F experiments have 100% weightage.

Experiment Evaluation (10 marks per experiment)	
Algorithm Design	3 Marks
Coding Syntax	2 Marks
Execution / Bug Finding	2 Marks
Records (submitted before the very next turn.)	3 Marks

F2. Viva Voce: WEIGHTAGE - 50%

The preparation of the students would be evaluated based on viva-voce or quiz examination in periodic schedules.

It is mandatory for the students to attend the above said continuous evaluation. Students who do not attend will lose their marks. Continuous Internal Assessment Record Sheet will be displayed at the end of the semester.

F3. GRADING:

The overall marks obtained at the end of the semester comprising the above two mentioned shall be converted to a grade.

Student(s), who have met the qualifying criteria of individual practical subject but not met qualifying criteria of SGPA, will not be allowed to re-appear for improvement. Students, who wish to re-appear in the practical subject, shall be required to pay the prescribed fee per subject as notified by the University. The student with Grade “F” only will be eligible to *repeat continuous evaluation* of that respective practical subject (s) during summer vacation (June-July).

Grade shall be awarded on the performance of the student(s). The Grade will be capped as per the rules mentioned in student Bulletin. There will be no capping of SGPA for the students re-appeared for Practical Subject. All Other rules and regulations such as requirement of passing, etc. will remain same as mentioned in rules & regulations.

F. DETAILED SESSION PLAN

S. No	Lab Exercises	Contents
1	Experiment No 1	Text Retrieval
2	Experiment No 2	Processing Data
3	Experiment No 3	Spacy Language Processor
4	Experiment No 4	Language Processor
5	Experiment No 5	Accessing text corpus
6	Experiment No 6	Processing raw text
7	Experiment No 7	Categorizing and tagging words
8	Experiment No 8	POS tagger
9	Experiment No 9	Parsing
10	Experiment No 10	Text Similarity

BOOKS:

- T1. Book provided by IBM- Computational Linguistics and Natural Language Processing (Course code GAI06SG1 V1.0)

REFERENCES:

- Artificial Intelligence: A Modern Approach by Stuart Russell and Peter Norvig, c 1995 Prentice-Hall, Inc.
- J. Craig, Introduction to Robotics Mechanics and Control, Pearson, 2018
- B. Ripley, Pattern Recognition and Neural Networks, Cambridge University Press, 1996.
- Steven Bird, Ewan Klein, Edward Loper, Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit (O'Reilly 2009, website 2018) <http://www.nltk.org/book/>
- Dipanjan Sarkar, Text Analytics with Python (Apress/Springer, 2016) <https://link-springer-com.proxy.uchicago.edu/book/10.1007%2F978-1-4842-2388-8>

- Stanford University CS224n: Natural Language Processing with Deep Learning
<http://web.stanford.edu/class/cs224n/>
- Paul Vierthaler's Stylometric PCA and Network Data Explorer
<https://www.pvierth.com/pca>

G. GUIDELINES:

Cell Phones and other Electronic Communication Devices: Cell phones and other electronic communication devices (such as Blackberries/Laptops) are not permitted in classes during Tests or the Mid/Final Examination. Such devices **MUST** be turned off in the class room.

E-Mail and online learning tool: Each student in the class should have an e-mail id and a password to access the e-labs system regularly. Regularly, important information – Date of conducting class tests, guest lectures, via online learning tool. The best way to arrange meetings with us or ask specific questions is by email and prior appointment. All the assignments preferably should be uploaded on online learning tool. Various research papers/reference material will be mailed/uploaded on online learning platform time to time.

Attendance: Students are required to have **minimum attendance of 75%** in each subject. Students with less than said percentage shall **NOT** be allowed to appear in the end semester examination.

Passing criterion: Student has to secure minimum 40% marks of the “highest marks in the class scored by a student in that subject (in that class/group class)” individually in both the ‘End-Semester examination’ and ‘Total Marks’ in order to pass in that paper.

- Passing Criterion for B. Tech: minimum 40% of the highest marks in the class
- Passing Criterion for M. Tech: minimum 40% of the highest marks in the class

H. Course outcome assessment: To assess the fulfilment of course outcomes two different approaches have been decided. Degree of fulfillment of course outcomes will be assessed in different ways through direct assessment and indirect assessment. In Direct Assessment, it is measured through quizzes, tests, assignment, Mid-term and/or End-term examinations. It is suggested that each examination is designed in such a way that it can address one or two outcomes (depending upon the course completion). Indirect assessment is done through the student survey which needs to be designed by the faculty (sample format is given below) and it shall be conducted towards the end of course completion. The evaluation of the achievement of the Course Outcomes shall be done by analyzing the inputs received through Direct and Indirect Assessments and then corrective actions suggested for further improvement.