

Walchand Institute of Technology, Solapur
Computer Science and Engineering
Honors in Artificial Intelligence and Machine Learning

*Structure of S. Y. B. Tech. Computer Science and Engineering
(W.E.F. 2022-2023)*

Semester- IV

Course Code	Theory Course Name	Engagement Hours			Credits	FA	SA		Total
		L	T	P		ESE	ISE	ICA	
21CSU4HA1T	Foundations and Applications of Machine Learning	3	-	-	3	60	40	--	100
	Laboratory:								
21CSU4HA1L	Foundations and Applications of Machine Learning	--	--	2	1			25	25
	Grand Total	3	-	2	4	60	40	25	125

Walchand Institute of Technology, Solapur
Computer Science and Engineering
Honors in Artificial Intelligence and Machine Learning

*Structure of T. Y. B. Tech. Computer Science and Engineering
(W.E.F. 2023-2024)*

Semester- V

Course Code	Theory Course Name	Engagement Hours			Credits	FA		SA		Total
		L	T	P		ESE	ISE	ICA		
21CSU5HA1T	Reinforcement Learning	3	--	--	3	60	40	--	100	
	Laboratory:									
21CSU5HA1L	Reinforcement Learning	--	--	2	1	--	--	25	25	
	Grand Total	3		2	4	60	40	25	125	

*Structure of T. Y. B. Tech. Computer Science Engineering.,
(W.E.F. 2023-2024)*

Semester- VI

Course Code	Theory Course Name	Engagement Hours			Credits	FA		SA		Total
		L	T	P		ESE	ISE	ICA		
21CSU6HA1T	Natural Language Processing	3	--	--	3	60	40	--	100	
	Laboratory:									
21CSU6HA1L	Natural Language Processing	--	--	2	1	--	--	25	25	
	Grand Total	3		2	4	60	40	25	125	

Walchand Institute of Technology, Solapur
Computer Science and Engineering
Honors in Artificial Intelligence and Machine Learning

*Structure of Final Year B. Tech. Computer Science and Engineering
(W.E.F. 2024-2025)*

Semester- VII

Course Code	Theory Course Name	Engagement Hours			Credits	FA	SA		Total
		L	T	P		ESE	ISE	ICA	
21CSU7HA1T	Deep Learning	3	--	--	3	60	40		100
	Laboratory:					OE			
21CSU7HA1L	Deep Learning	--	--	2	1	--	--	25	25
21CSU7HA2L	Mini Project	--	--	4	2	50	--	50	100
	Grand Total	3		6	6	110	40	75	225

Note: -These courses are to be completed by the student in addition to the courses of B.Tech. Computer Science and Engineering.



**WALCHAND INSTITUTE OF TECHNOLOGY
(AUTONOMOUS INSTITUTE)**

Name of the Faculty: Science and Technology

CHOICE BASED CREDIT SYSTEM

Structure and Syllabus : Computer Science & Engineering

Name of the Course: T.Y. B. Tech. Sem V and Sem VI

Hons. Degree : Artificial Intelligence & Machine Learning

(Syllabus to be implemented from July 2023)

WALCHAND INSTITUTE OF TECHNOLOGY, SOLAPUR

**Faculty of Science and Technology
Third Year B.Tech. (CSE & IT)
Honors Degree in AI & Machine Learning
SEMESTER - V**

21CSU5HM2 : REINFORCEMENT LEARNING

Teaching Scheme

Lectures – 3 Hours/week, 3 Credits

Practical – 2 Hour/week, 1 Credit

Examination Scheme

ESE – 60 Marks

ISE – 40 Marks

ICA – 25 Marks

Introduction :

Reinforcement learning is an area of machine learning, where an agent or a system of agents learns to archive a goal by interacting with their environment. In recent years there has been success in reinforcement learning research in both theoretical and applied fields. This course primarily focuses on training students to frame reinforcement learning problems and to tackle algorithms from dynamic programming, Monte Carlo and temporal-difference learning.

Pre-requisite:

A basic course on Artificial Intelligence & Machine learning

COURSE OUTCOMES :

At the end of the course students will be able to

1. Demonstrate the fundamental mathematical models and algorithms in the field of NLP.
2. Apply these mathematical models and algorithms in applications of software design and implementation for NLP.
3. Use tools to analyze language resource annotation and apply to data for acquiring intended information.
4. Design and implement various NLP applications.

SECTION I

Unit 1 Introduction

(05)

Reinforcement Learning, Examples, Elements of Reinforcement Learning, History of Reinforcement Learning

Unit 2 Evaluative Feedback

(05)

A k-armed Bandit Problem, Action-value Methods, The 10-armed Test-bed, Incremental Implementation

Unit 3 The Reinforcement Learning Problem

(06)

The Agent–Environment Interface, Goals and Rewards, Returns, Unified Notation for Episodic and Continuing Tasks, Value Functions, Optimal Value Functions, Optimality and Approximation

Unit 4 Finite Markov Decision Processes

(6)

The Agent–Environment Interface, Goals and Rewards, Returns and Episodes, Unified Notation for Episodic and Continuing Tasks, Policies and Value Functions.

SECTION II

Unit 5 Dynamic Programming

(5)

Policy Evaluation (Prediction), Policy Improvement, Policy Iteration, Value Iteration, Asynchronous Dynamic Programming, Generalized Policy Iteration, Efficiency of Dynamic Programming, Introduction to Monte Carlo Methods.

Unit 6 Temporal-Difference Learning**(5)**

TD Prediction, Advantages of TD Prediction Methods, Optimality of TD(0), SARSA: On-policy TD Control, Q-learning: Off-policy TD Control.

Unit 7 Planning and Learning**(6)**

Models and Planning, Dyna: Integrating Planning, Acting, and Learning , When the Model Is Wrong , Prioritized Sweeping, Expected vs. Sample Updates.

Unit 8 Applications and Case Studies**(6)**

TD-Gammon, Samuel's Checkers Player, Watson's Daily-Double Wagering, Mastering the Game of Go and AlphaGo.

Internal Continuous Assessment (ICA) :

Analysis and implementation of

1. Flappy Kernel Markov Decision Process
2. Implementation of Performance Difference Lemma.
3. Implementation of Pong with Deep Q Learning.
4. Estimation of Warfarin Dose
5. Implementing Bayesian regret bound for Thomson Sampling

Text Books:

1. Reinforcement Learning: An Introduction (Second edition + Upcoming Edition) by: Richard S. Sutton and Andrew G. Barto, MIT Press Publication

(The book is available at <http://incompleteideas.net/book/the-book-2nd.html> Upcoming edition's January 1 2018 draft available at <http://incompleteideas.net/book/bookdraft2018jan1.pdf>)

Reference Books:

1. Reinforcement Learning: With Open AI, TensorFlow and Keras Using Python By Abhishek Nandy, Manisha Biswas. Apress Publication
2. Reinforcement Learning: State-of-the-Art, Marco Wiering and Martijn van Otterlo, Eds.
3. Artificial Intelligence: A Modern Approach, Stuart J. Russell and Peter Norvig.
4. Deep Learning, Ian Goodfellow, Yoshua Bengio, and Aaron Courville.

WALCHAND INSTITUTE OF TECHNOLOGY, SOLAPUR

Faculty of Science and Technology

Third Year B.Tech. (CSE & IT)

Honors Degree in AI & Machine Learning

SEMESTER - VI

21CSU6HM3 : NATURAL LANGUAGE PROCESSING

Teaching Scheme

Examination Scheme

Lecture: 3 Hours /Week, 3 Credits
Practical: 2 Hours/Week, 1 Credit

ESE – 60 Marks
ISE – 40 Marks
ICA - 25 Marks

Introduction:

Natural Language Processing (NLP) is essentially a method for teaching computers how to understand human languages and interpret text. This course covers fundamentals of NLP including language morphology and language modelling, syntax analysis, semantic analysis, and the use of NLP for information retrieval. It also covers the most popular and effective current techniques, strategies, and toolkits for NLP to develop the various real-world NLP applications.

Pre-requisite:

Basic mathematics, algorithms and programming skills, Theory of computation and parsing.

COURSE OUTCOMES:

At the end of the course students will be able to

1. Understand the fundamentals of Natural Language Processing.
2. Analyze how the words are formed morphologically and how they are related to each others.
3. Develop strategies for language modeling, syntax and semantic analysis.
4. Design and implement and analyze the Natural Language Processing algorithms for real word applications.

SECTION-I**Unit 1 Introduction to Natural Language Processing****(6)**

Introduction to NLP, Machine Learning and NLP, why NLP is hard? Programming languages Vs Natural Languages, Are natural languages regular? Finite automata for NLP, Stages of NLP, challenges (Open Problems) in NLP.

Basics of Text Processing: Tokenization, Stemming, Lemmatization, Part of Speech Tagging.

Unit 2: Syntax and semantic Analysis**(8)**

Morphological Analysis: What is Morphology? Types of Morphemes, Inflectional morphology & Derivational morphology, Phonetics, HMM, Morphological parsing with Finite State Transducers (FSD. Syntactic Analysis: Syntactic Representations of Natural Language, Parsing Algorithms, Probabilistic context-free grammars and Statistical parsing.

Lexical Semantic, Relations among lexemes & their senses, Homonymy, Polysemy, Synonymy, Hyponymy, Word-Net, Word Sense Disambiguation (WSD), Dictionary based approach, Latent Semantic Analysis.

Unit 3: Language Modeling**(08)**

Probabilistic language modeling, Markov models, Generative models of language, Log-Liner Models, Graph-based Models.

N-gram Models: Simple n-gram models, Estimation parameters and smoothing, evaluating language models.

Word Embeddings/Vector Semantics: Bag-of-words, TFIDF, word2vec, doc2vec, Contextualized representations (BERT).

Topic Modeling: Latent Dirichlet Allocation (LDA), Latent Semantic Analysis, Non –Negative Matrix Factorization.

SECTION-II

Unit 4: Information Retrieval using NLP

(06)

Information Retrieval: Introduction, Vector Space Model.

Named Entity Recognition: NER System building process, Evaluating NER System. Entity Extraction, Relation Extraction, Reference Resolution, Co reference resolution, Cross Lingua information retrieval (CLIR).

Unit 5: NLP Tools and Techniques

(07)

Prominent NLP libraries: Natural Language Toolkit (NLTK), spaCy, TextBlob, Gensim. Linguistic Resources: Lexical Knowledge Networks, WordNets, Indian Language WordNet, (IndoWordnet), VerbNets, PropBank, Treebanks, Universal Dependency Tree banks.

Word Sense Disambiguation: Lesk Algorithm Walker's algorithm, WordNets for Word Sense Disambiguation.

Unit 6: Applications of NLP

(07)

Machine Translation: Rule based techniques, Statistical Machine Translation (SMT), Cross Lingual Translation.

Sentiment Analysis, Question Answering, Text Entailment, Discourse Processing, Dialog and Conversational Agents, Natural Language Generation.

Internal Continuous Assessment (ICA) :

ICA shall include at least eight of the following assignments. The assignment's objective should align with course's outcomes and focus on higher order bloom's cognitive levels.

- 1) Perform Exploring and processing of text data.
- 2) Perform language modeling using one-hot encoding, N-grams, count vectorizer, TFID vectorizer and word embedding techniques.
- 3) Study and implementation of Noun-phrase extraction, text-similarity, part of speech tagging, chunking and named entity recognition.
- 4) Write a simple English to French translation algorithm using pre-computed word embeddings and locality sensitive hashing to relate words via approximate k-nearest neighbor search.
- 5) Study and implementation of sentimental analysis
- 6) Study and implementation of SPAM-HAM email classification
- 7) Write a better auto-complete algorithm using an N-gram language model
- 8) Study and implementation of text to speech and speech to text conversion
- 9) Study of Stanford Parser and POS Tagger. <https://nlp.stanford.edu/software/lex-parser.html>
- 10) Study of language modeling for Indian languages
- 11) Study and implementation of Multi-class classification using NLP and Naïve Bayes classifier
- 12) Study and implementation of text summarization and document clustering.

Text Books:

1. Dr. Smita M. Chaudhari, Devika A. Verma, Nitin N. Sakhare, "Natural Language Processing", Nirali Prakashan, Pune.
2. Allen, James, "Natural Language Understanding", Second Edition, Benjamin/Cumming, 1995.

Reference Books:

1. Bird, S., Klein, E., Loper, E. (2009). "Natural Language Processing with Python". Sebastopol, CA: O'Reilly Media.
2. Akshay Kulkarni, Adarsha Shivananda, "Natural Language Processing Recipes Unlocking Text Data with Machine Learning and Deep Learning using Python", Apress, Bangalore
3. Radford, Andrew et. al., "Linguistics, An Introduction", Cambridge University Press, 1999.

4. Charniack, Eugene, “Statistical Language Learning”, MIT Press, 1993.
5. Joseph D. Booth Foreword by Daniel Jebaraj “Natural Language Processing Succinctly”, Succinctly EBook series, SyncFusion.
6. Github link for practical: <https://github.com/kb1907/Natural-Language-Processing-Specialization>