SEMESTER – VIII

Savitribai Phule Pune University, Pune

Final Year Artificial Intelligence and Machine Learning (2020 Course)

418550: Natural Language Processing

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory (TH): 03 hrs/week	03 Credits	Mid_Semester: 30 Marks End_Semester:70 Marks

Prerequisite Course : Discrete Mathematics, Theory of Computation

Companion Course: Object Oriented Programming, Computer Graphics Lab and Authoring Tools

Course Objectives:

1. To be familiar with fundamental concepts and techniques of natural language processing (NLP)
2. To acquire the knowledge of various morphological, syntactic, and semantic NLP tasks
3. To develop the various language modeling techniques for NLP
4. To use appropriate tools and techniques for processing natural languages
5. To comprehend the advance real world applications in NLP domain.
6. To Describe Applications of NLP and Machine Translations.

Course Outcomes:

On completion of the course, students will be able to–

CO1. Describe the fundamental concepts of NLP, challenges and issues in NLP.
CO2. Analyze Natural languages morphologically, syntactical and semantically
CO3. Illustrate various language modelling techniques
CO4. Integrate the NLP techniques for the information retrieval task.
CO5. Demonstrate the use of NLP tools and techniques for text-based processing of natural languages.
CO6. Develop real world NLP applications.

COURSE CONTENTS

Unit I	Introduction to Natural Language Processing	(6 hrs)

Introduction: Natural Language Processing, Why NLP is hard? Programming languages Vs Natural Languages, Are natural languages regular? Finite automata for NLP, Stages of NLP, Challenges and Issues(Open Problems) in NLP

Basics of text processing: Tokenization, Stemming, Lemmatization, Part of Speech Tagging

Mapping of Course Outcomes for Unit I	CO1

Unit II	Language Syntax and Semantics	(6 hrs)

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

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

Mapping of Course Outcomes for Unit II	CO2

Unit III	Language Modelling	(6 hrs)

Probabilistic language modelling, 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 Modelling: Latent Dirichlet Allocation (LDA), Latent Semantic Analysis, Non Negative Matrix Factorization.		
Mapping of Course Outcomes for Unit III	**CO3**	
Unit IV	**Information Retrieval using NLP**	**(6 hrs)**
Information Retrieval: Introduction, Vector Space Model Named Entity Recognition: NER System Building Process, Evaluating NER System Entity Extraction, Relation Extraction, Reference Resolution, Coreference resolution, Cross Lingual Information Retrieval		
Mapping of Course Outcomes for Unit IV	**CO4**	
Unit V	**NLP Tools and Techniques**	**(6 hrs)**
Prominent NLP Libraries: Natural Language Tool Kit (NLTK), spaCy, TextBlob, Gensim etc. Linguistic Resources: Lexical Knowledge Networks, WordNets, Indian Language WordNet (IndoWordnet), VerbNets, PropBank, Treebanks, Universal Dependency Treebanks Word Sense Disambiguation: Lesk Algorithm WordNets for Word Sense Disambiguation		
Mapping of Course Outcomes for Unit V	**CO5**	
Unit VI	**Applications of NLP**	**(6hrs)**
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		
Mapping of Course Outcomes for Unit VI	**CO6**	

Textbooks:

1. Jurafsky, David, and James H. Martin Speech and Language Processing: An Introduction to Natural Language Processing , Computational Linguistics and Speech PEARSON Publication

2. Manning, Christopher D., and nrich Schütze Foundations of Statistical Natural Language Processing, Cambridge, MA: MIT Press

Reference Books:
1. Steven Bird, Ewan Klein, Edward Loper Natural Language Processing with Python Analysing Text with the Natural Langua Publication 2. Dipanjan text Analytics with Python: A Practical Real-World Approach to Gaining Actionable Insights from your Publication ISBN: 9781484223871 3. Alexander Clark, Chris Fox, and Shalom Lappin The Handbook of Computational Linguistics and Natural Language Processing, Wiley Blackwell Publications 4. Jacob Eisenstein Natural Language Processing, MIT Press 5. Jacob Eisenstein An Introduction to Information Retrieval Cambridge University Press
E Books / E Learning References:
1. https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf 2. https://www3.cs.stonybrook.edu/~cse521/L16NLP.pdf 3. https://nptel.ac.in/courses/106101007 4. https://nptel.ac.in/courses/106106211

Savitribai Phule Pune University, Pune

Final Year Artificial Intelligence and Machine Learning (2020 Course)

418551A: Elective V- (Distributed Systems)

eaching Scheme: 03 Hrs/Week	Credit Scheme:	Examination Scheme:
Theory (TH): 03 Hrs/week	03 Credits	Mid_Semester: 30 Marks End_Semester: 70 Marks

Prerequisite Courses:

Operating System, Computer Network, Data Structure and Algorithm

Companion Course, if any:

NA

Course Objectives:

1. To understand the fundamental concepts and principles of distributed systems.
2. To gain knowledge and skills in using middleware technologies for distributed systems.
3. To develop an understanding of different inter-process communication mechanisms in distributed systems.
4. To learn about replication techniques and fault tolerance mechanisms in distributed systems.
5. To explore the design and implementation of distributed file systems, multimedia systems, and web-based systems.
6. To stay updated with the latest trends and advancements in distributed systems.

Course Outcomes:

On completion of the course, students will be able to–

CO1. Analyze and evaluate the design choices and trade-offs involved in building distributed systems.
CO2. Design and implement efficient distributed systems using middleware.
CO3. Design and implement effective inter-process communication strategies in distributed systems.
CO4. Develop fault-tolerant distributed systems by implementing replication and fault tolerance strategies.
CO5. Apply distributed file, multimedia, and web-based systems to real-world scenarios.
CO6. Incorporate recent trends and technologies in the design and implementation of distributed systems.

COURSE CONTENTS

Unit I	Introduction to Distributed Systems	(6 hrs)

Introduction: Network operating System VS Distributed operating systems, Characteristics, Design goals, challenges of Distributed Systems, Examples of Distributed Systems, Trends in Distributed systems: Pervasive networking and the modern Internet, Mobile and ubiquitous computing, Focus on resource sharing

Distributed Computing Models: Physical, Architecture and Fundamental models

Case Study: Google File System (GFS)

Mapping of Course Outcomes for Unit I	CO1

Unit II	Middleware	(6 hrs)

Introduction to middleware, middleware Framework, Role of middleware, Examples of Middleware, Origins of middleware, Architecture vs Middleware, RMI, CORBA, General Approaches to adaptive software, Types of middleware-messages oriented middleware, intelligent middleware, content centric middleware, middleware protocol, middleware Services, Distributed computing Environment (DCE), middleware Issues, middleware Analyst **Case Study: -** Apache Kafka		
Mapping of Course Outcomes for Unit II	**CO2**	
Unit III	**Inter-Process Communication**	**(6 hrs)**
IPC: Introduction, Layered protocols, API for internet protocols, IPC through shared memory, external data representation and marshaling, Types of communication, inter process communication, multicast communication, message-oriented communication, MPI, network virtualization, overlay networks **Coordination:** Clock synchronization, logical clocks, mutual exclusion, election algorithms, Gossip based coordination **Case Study:** IBM WebSphere Message Queuing		
Mapping of Course Outcomes for Unit III	**CO3**	
Unit IV	**Replication and Fault Tolerance**	**(6 hrs)**
Replication: Reasons for replication, Replica management – Finding the best server location, Content replication and placement, Content distribution, Managing replicated objects Consistency protocols: Primary based protocols, replicated write protocols **Fault Tolerance:** Introduction to fault tolerance, Reliable client server communication, Reliable group communication, distributed commit, Recovery – Check pointing, Message logging **Case Study:** Amazon DynamoDB		
Mapping of Course Outcomes for Unit IV	**CO4**	
Unit V	**Distributed Files, Multimedia and Web Based System**	**(6 hrs)**
Distributed Files: Introduction, File System Architecture, Sun Network File System and HDFS. **Distributed Multimedia Systems:** Characteristics of Multimedia Data, Quality of Service Management, Resource Management **Distributed Web Based Systems:** Architecture of Traditional Web-Based Systems, Apache Web Server, Web Server Clusters, Communication by Hypertext Transfer Protocol, Synchronization, Web Proxy Caching **Case Study:** BitTorrent		
Mapping of Course Outcomes for Unit V	**CO5**	
Unit VI	**Recent Trends in Distributed Systems**	**(6 hrs)**

Recent Trends: Introduction, Portable and handheld Devices, Wearable devices, Devices embedded in appliances, Parallel Virtual Machine (PVM), Jini, Service Oriented Architecture, The Future of Recent Trends. **Tools for Distributed System Monitoring:** Prometheus, Zabbix, Nagios **Case Studies:** Kubernetes	
Mapping of Course Outcomes for Unit VI	**CO6**

Textbooks:

1. Distributed Systems: Concepts and Design by George Coulouris, J Dollimore and Tim Kindberg, Pearson Education, ISBN: 9789332575226, 5th Edition, 2017.
2. Distributed Systems, Maarten van Steen, Andrew S. T,ThirdeditionVersion. Andrew S. Tanenbaum, Maarten van Steen, PHI ,2nd Edition, ISBN: 978-0130888938
3. Distributed Operating Systems: Concepts and Design by P. K. Sinha, PHI, ISBN: 978-0780311190

Reference Books:

1. Distributed Computing, Sunita Mahajan and Seema Shah, Oxford University
2. Distributed Computing, Fundamentals, Simulations and Advanced topics, 2nd Edition, Hagit Attiya and Jennifer Welch, Wiley India
3. Tool for Distributed Systems Monitoring, Łukasz KUFEL, Foundation of Computing and DecisionSciences, Vol 41(4), 2016, e-ISSN 2300-3405, DOI:10.1515/fcdc-2016-0014

E Books / E Learning References:

1. http://home.mit.bme.hu/~meszaros/edu/oprendszerek/segedlet/elosztott/distributed-systems-survey.pdf
2. http://home.mit.bme.hu/~meszaros/edu/oprendszerek/segedlet/elosztott/DisSysUbiCompReport.html

Savitribai Phule Pune University, Pune

Final Year Artificial Intelligence and Machine Learning (2020 Course)

418551B: Elective V- (Software Project Management)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory (TH) : 3 hrs/week	03	Mid_Semester : 30 Marks End_Semester :70 Marks

Prerequisite Courses, if any: Software Engineering

Companion Course:---

Course Objectives:

1. To discuss the fundamentals of Software Project Management
2. To explain Project Design and Project Evaluation.
3. To acquire skill in Activity Planning and to deal with Risk Management
4. To provide platform to understand through different tools about Project Tracking, Monitoring & Control.
5. To discuss Staff Selection Process and the issues related to Staff Management.
6. To provide exposure to modern tools used for Software Project Management.

Course Outcomes:

On completion of the course, students will be able to–

CO1. Apply the practices and methods for successful Software Project Management
CO2. Use various tools of Software Project Management
CO3. Create Design and Evaluate Project
CO3. Analyze Project Schedule and calculate Risk Management with help of tools.
CO4. Demonstrate different tools used for Project Tracking, Monitoring & Control.
CO5. Analyse a case study for a distributed team and comment.
CO6. Discuss and use modern tools for Software Project Management.

COURSE CONTENTS

Unit I	Introduction to Software Project Management	(6 hrs)

Introduction to Software Project Management: Why is Software Project Management Important? What is a Project? Contract Management, Activities Covered by Software Project Management, Plans, Methods and Methodologies, Some Ways of Categorizing Software Projects, Stakeholders, Setting Objectives, Business Case, Project Success, and Failure, what is Management? Management Control, Traditional versus Modern Project Management Practices**.**

Mapping of Course Outcomes for Unit I	CO1
Case study:	Online Shopping System

Unit II	Project Planning and Project Management Tools	(6 hrs)

Project Planning: Steps for Project Planning, PERT and Gantt Charts, Gantt Project, Microsoft Project and Primavera Project Management Software, Objectives of Activity planning, Project Schedules, Activities, Sequencing and Scheduling, Network Planning Models, Formulating Network Model.

Mapping of Course Outcomes for Unit II	CO2

Case study:	**Software project plan using any tool.**	
Unit III	**Project Design Evaluation and Risk Management**	**(6 hrs)**
Project Design: Overview of UML diagrams: Use case, Class, Activity, State, Sequence, Deployment **Project Evaluation:** What is Project Evaluation? Importance of Project Evaluation, Cost Benefit Evaluation Techniques. Process Evaluation and Improvement: The Process **Improvement Process:** The Process Improvement. Cycle, Process Measurement: The GQM Paradigm, Process Analysis: Techniques of Process Analysis, Process change: The Process Change Process. **Risk Management:** Introduction, Risk Management, Risk Assessment, Risk identification, Risk Prioritization, Risk Planning, Risk control, Risk Strategies, Evaluating Risk to the schedule.		
Mapping of Course Outcomes for Unit III	**CO3,CO4**	
Case study:	**Online Shopping System, Perform Cost Benefit Analysis using Microsoft Excel**	
Unit IV	**Project Tracking and Control**	**(6 hrs)**
Introduction: Project Tracking and Control, Monitoring and Control Processes, Collection of Project data, Partial Completion Reporting. **Data Collection Methods:** Phone vs. Online vs. InPerson Interviews, Visualizing Progress, Visual Project Management, Kanban Boards, Project Calendars, Cost Monitoring, Four Steps in Project Cost Management, Earned Value Analysis, Project Tracking, Effective Approach to Track Projects. **Status Report:** Four features of a Good Status Report, Change Control, Different factors of Change Control Process, Change Process FlowDiagram, Software Configuration Management, Tasks in SCM Process, Participant of SCM Process. **Software Configuration Management Tools:** Git, Team Foundation Server, Ansible, Managing Contracts, The Stages of Contract Management, Challenges of Contract Management, Benefits of Contract Management, Types of Contracts in Software Project Management		
Mapping of Course Outcomes for Unit IV	**CO4**	
Case study:	**Online Shopping System, track different versions of a software using Git tool**	
Unit V	**Staffing and Team organization in Software Projects**	**(6 hrs)**
Managing People, Organizational behaviour, Best methods of Staff Selection, Motivation, The Oldham, Hackman job characteristic Model, Stress, Health and Safety, Ethical and Professional concerns, working in Teams, Decision Making, Organizational structures, Dispersed and Virtual Teams, Communications Genres, Communication Plans.		
Mapping of Course Outcomes for Unit V	**CO5**	
Case study:	**Team Building in Project Management with reference to academic project work.**	
Unit VI	**Applications of Software Project Management in Industry**	**(6 hrs)**

Agile Project Management with Azure DevOps: An Overview of Application Lifecycle Management & Azure DevOps, Traceability, Visibility, Collaboration, and Extensibility. Difference between Microsoft TFS and Azure DevOps. Metrics in Agile Practice: Introduction to Metrics inAgile Practice, Metrics for Project Management, Agile Project Management in Azure DevOps and TFS.	
Mapping of Course Outcomes for Unit VI	**CO6**
Case study:	**Online Shopping System.**

Text Books:

1. Bob Hughes, Mike Cotterell and Rajib Mall: Software Project Management – FifthEdition, Tata McGrawHill, New Delhi.
2. A Guide to the Project Management Book of Knowledge-Seventh Edition.
3. Walker Royce, "Software Project Management" a unified approach. Addison Wesley ISBN 0-20130958-0.
4. Robert K. Wysocki, "Effective Software Project Management", Wiley Publication, 2011

Reference Books:

1.JackMarchewka," Information Technology-Project Management", Wiley Student Version, 4th Edition, 2013.
2. Lan Somerville, Software Engineering, Fifth Edition, Addison Wesley Publications, 1996. (For Unit 2)
3. JIM Arlow, Ila Neustadt, UML 2 and the Unified Process, Pearson, Second Edition, ISBN:9788131700549
Tom Pender, UML 2 Bible, Wiley India, ISBN: 9788126504527. (For Unit 2)
4. James P Lewis, "Project Planning, Scheduling & Control", McGraw Hill, 5th Edition, 2011.
5. PankajJalote, "Software Project Management in Practice", Pearson Education, 2002.
6. Gopalaswamy Ramesh, "Managing Global Software Projects" – McGraw Hill Education (India), Fourteenth Reprint 2013.
7. Joachim Rossberg, "Agile Project Management with Azure DevOps" Apress. (For Unit 6)
8. Robert K. Wyzocki, Rudd McGary, Effective Project Management, WILEYDreamtech India Pvt. Ltd., 2000

E Books / E Learning References :

1. https://www.inflectra.com/SpiraPlan/(for Unit 3)
2. https://www.techtarget.com/searchsecurity/definition/governance-risk-management-andcompliance-GRC(for Unit 3)
3. https://www.softwaretestinghelp.com/risk-management-tools/#3_Risk_Management_Studio
1. (For Unit 3)
4. NPTEL: https://nptel.ac.in/courses/106101061/29
5. https://onlinecourses.nptel.ac.in/noc17_mg01/preview
6. Coursera: https://www.coursera.org/learn/uva-darden-project-management
7. http://managementhelp.org/evaluation/program-evaluation-guide.htm.
8. https://nptel.ac.in/courses/106105218 (NPTEL)
9. Virtual Labs:- Software Engineering-
    i http://vlabs.iitkgp.ernet.in/se/3/
    ii http://vlabs.iitkgp.ernet.in/se/5/
    iii http://vlabs.iitkgp.ernet.in/se/6/
    iv http://vlabs.iitkgp.ernet.in/se/7/

Savitribai Phule Pune University, Pune

Final Year Artificial Intelligence and Machine Learning (2020 Course)

418551C: Elective V- (Computer Vision)

Teaching Scheme	Credit Scheme	Examination Scheme
Theory (TH): 3 hrs/week	03 Credits	Mid_Semester: 30 Marks End_Semester: 70 Marks

Prerequisite Courses:

1. Students should know vectors, linear algebra (i.e., matrix operations, solution of linear equations).
2. Programming language (e.g., C, Matlab, Python etc).

Companion Course, if any:

Course Objectives:

1. To review image processing techniques for computer vision.
2. To understand shape and region analysis.
3. To understand three-dimensional image analysis techniques.
4. To understand motion detection techniques.
5. To study some applications of computer vision algorithms.

Course Outcomes:

By the end of the course, students should be able to

CO1. Implement fundamental image processing techniques required for computer vision.
CO2. Apply feature extraction techniques.
CO3. Apply Hough Transform for line, circle, and ellipse detections.
CO4. Implement three-dimensional analysis techniques.
CO5. Implement Motion detection and object tracking techniques.
CO6 Develop skills to implement diverse computer vision applications.

COURSE CONTENTS

Unit I	Fundamentals of Digital Image Processing	(6 hrs)

Introduction to Computer Vision?, Fundamentals Of Image Formation, **Review of Digital image processing:** Introduction, Origin, Applications and Examples of Digital Image Processing, Fundamental steps in Digital Image Processing, Components of Digital Image Processing, Elements of Visual Perception, Image Sensing and Acquisition, Image Sampling and Quantization, Basic Relationship between pixels, **image processing techniques:** classical filtering operations, Thresholding techniques, edge detection techniques, corner and interest point detection, texture Analysis

Mapping of Course Outcomes for Unit I	CO1

Unit II	SHAPES And REGIONS	(6 hrs)

Binary shape analysis, Connectedness, object labelling and counting, size filtering, distance functions and their uses, skeletons and thinning, Other Measures for Shape Recognition, Boundary pattern analysis: Boundary Tracking Procedures, Centroidal Profiles, Tackling the Problems of Occlusion, Accuracy of Boundary Length Measures, Object segmentation and shape models, Active Contours, Shape Models

Mapping of Course Outcomes for Unit II	CO2

Unit III	FEATURE DETECTION AND MATCHING	(6 hrs)

Points and patches: Feature detectors, Feature descriptors, Feature matching, Feature tracking Application: Performance-driven animation, **Edges:** Edge detection, Edge linking, Application: Edge editing and enhancement, Vanishing points, **Application:** Rectangle detection		
Mapping of Course Outcomes for Unit III	**CO3**	
Unit IV	**HOUGH TRANSFORM**	**(6 hrs)**
Line detection – Hough Transform (HT) for line detection, the foot-of-normal method, Using RANSAC for Straight Line Detection, Hough-Based Schemes for Circular Object Detection, The Problem of Unknown Circle Radius, Overcoming the Speed Problem, Ellipse Detection, Applications, and case study: Human Iris Location, The Generalized Hough Transform (GHT), Use of the GHT for Ellipse Detection, A Graph-Theoretic Approach to Object Location, Possibilities for Saving Computation, Using the GHT for Feature Collation		
Mapping of Course Outcomes for Unit IV	**CO4**	
Unit V	**3D VISION AND MOTION**	**(6 hrs)**
The three-dimensional world, Methods for 3D vision, projection schemes for 3D vision, Shape from X : shape from shading, Photometric Stereo, Shape from texture, Share from focus, The Assumption of Surface Smoothness, Shape from Texture, Use of Structured Lighting, 3D Reconstruction, active range finding, surface representations, point-based representation, volumetric representations, Structure from motion: triangulation, bundle adjustment, Dense motion estimation: translational alignment, parametric motion, spline based motion, Optical flow layered motion		
Mapping of Course Outcomes for Unit V	**CO5**	
Unit VI	**COMPUTER VISION APPLICATIONS**	**(6 hrs)**
Application: Photo album – Object detection, Face detection, Pedestrian detection, Face recognition: Eigen faces, Active appearance and 3D shape models, Application: Personal Photo Collections, Category Recognition, Intelligent Photo Editing, Image Search, **Application: Surveillance** – The basic geometry, foreground-background separation, particle filters, Chamfer Matching, Tracking, and Occlusion, combining views from multiple cameras, License Plate Location, Occlusion Classification for Tracking, Human Gait Analysis, In-vehicle vision system: Locating the Roadway, Location of Road Markings, Location of Road Signs, Location of Vehicles, Locating Pedestrians		
Mapping of Course Outcomes for Unit VI	**CO6**	
Textbooks:		
1. Simon J. D. Prince, "Computer Vision: Models, Learning, and Inference", Cambridge University Press, 2012. 2. Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", 3rdEdition, Pearson, ISBN: 978-81-317-2695-2		

Reference Books:

1. R. Davies, “Computer & Machine Vision”, Fourth Edition, Academic Press, 2012.
2. R. Szeliski, “Computer Vision: Algorithms and Applications”, Springer 2011.
3. Mark Nixon and Alberto S. Aquado, “Feature Extraction & Image Processing for Computer Vision”,Third Edition, Academic Press, 2012.
4. D. L. Baggio et al., “Mastering OpenCV with Practical Computer Vision Projects”, Packt Publishing, 2012.
5. Jan Erik Solem, “Programming Computer Vision with Python: Tools and algorithms for analyzing images”, O'Reilly Media, 2012.
6. Sudha Challa, “Fundamentals of Object Tracking”, Cambridge University Press, 2011.

Online references:

1. http://kercd.free.fr/linksKCD.html
2. http://www.cs.ubc.ca/spider/lowe/vision.html
3. http://www.visionscience.com/
4. https://www.fritz.ai/object-detection/
5. https://viso.ai/deep-learning/object-tracking/
6. https://www.pearson.com/us/higher-education/program/Gonzalez-Digital-Image-Processing-4th-Edition/PGM241219.html?tab=resources

Savitribai Phule Pune University, Pune
Final Year Artificial Intelligence and Machine Learning (2022 Course)
418552A: Elective VI- (Reinforcement Learning)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory (TH) : 03 hrs/week	03 Credits	Mid_Semester : 30 Marks End_Semester : 70 Marks

Prerequisite Courses, if any: Linear algebra, multivariable calculus, Basic machine learning knowledge

Companion Course, if any: https://nptel.ac.in/courses/106106143

Course Objectives:

1. To familiarize the students with the basic concepts in deep reinforcement learning.
2. To Compare and contrast different learning algorithms with parameters.
3. To Examine the nature of a problem at hand and find the appropriate reinforcement learning algorithms and its parameters that can solve it efficiently enough.
4. To Design and implement of deep and reinforcement learning approaches for solving real-life problems.

Course Outcomes:

On completion of the course, students will be able to–

- **CO1:** Describe about theories and process in a reinforcement learning problem
- **CO2:** Understand and apply basic Reinforcement Learning algorithms for simple sequential decision making problems in uncertain conditions
- **CO3:** Evaluate the performance of the solution and find optimal strategy.
- **CO4:** Understand how to fine tune the target to have better learning performance.
- **CO5:** Learn approximation methods and algorithms for optimizing the problem.
- **CO6:** Understand to decompose a reinforcement learning problem into hierarchy of sub problems or sub tasks.

COURSE CONTENTS

Unit I	Reinforcement Learning Problem	(06 hrs)
The Reinforcement Learning problem: evaluative feedback, nonassociative learning, Rewards and returns, Markov Decision Processes, Value functions, optimality and approximation		
Mapping of Course Outcomes for Unit I	**CO1**	
Unit II	**Bandit Problems & Dynamic Programming**	**(06 hrs)**
Bandit Problems: Explore-exploit dilemma, Binary Bandits, Learning automata, exploration schemes Dynamic programming: value iteration, policy iteration, asynchronous DP, generalized policy iteration		
Mapping of Course Outcomes for Unit II	**CO2**	
Unit III	**Monte Carlo Methods and Temporal Difference Learning**	**(06 hrs)**

Monte-Carlo methods: policy evaluation, roll outs, on policy and off policy learning, importance sampling Temporal Difference learning: TD prediction, Optimality of TD(0), SARSA, Q-learning, R-learning, Games and after states

Mapping of Course Outcomes for Unit III	**CO3**	
Unit IV	**Eligibility Traces**	**(06 hrs)**

Eligibility traces: n-step TD prediction, TD (lambda), forward and backward views, Q(lambda), SARSA(lambda), replacing traces and accumulating traces

Mapping of Course Outcomes for Unit IV	**CO4**	
Unit V	**Function Approximation**	**(06 hrs)**

Function Approximation: Value prediction, gradient descent methods, linear function approximation, Control algorithms, Fitted Iterative Methods Policy Gradient methods: non-associative learning - REINFORCE algorithm, exact gradient methods, estimating gradients, approximate policy gradient algorithms, actor-critic methods

Mapping of Course Outcomes for Unit V	**CO5**	
Unit VI	**Hierarchical Reinforcement Learning**	**(06 hrs)**

Hierarchical RL: MAXQ framework, Options framework, HAM framework, Option discovery algorithms Case studies: Elevator dispatching, Samuel's checker player, TDgammon, Acrobot, Helicopter piloting, Computational Neuroscience

Mapping of Course Outcomes for Unit VI	**CO6**

Text Books:

1. R. S. Sutton and A. G. Barto. Reinforcement Learning - An Introduction. MIT Press. 1998.
2. Csaba Szepesvari. Algorithms for Reinforcement learning. Morgan & Claypool Publishers.
3. Marco Wiering and Martijn van Otterlo, Eds. Reinforcement Learning: State-of-the-Art. Sprinkler.
4. Stuart J. Russell and Peter Norvig. Artificial Intelligence: A Modern Approach. Pearson.
5. Ian Goodfellow, Yoshua Bengio,and Aaron Courville. Deep Learning. MIT Press.

Reference Books:

1. Reinforcement Learning: State-of-the-Art, Marco Wiering and Martijn van Otterlo, Eds.(Reinforcement Learning: State-of-the-Art | SpringerLink)
2. Artificial Intelligence: A Modern Approach, Stuart J. Russell and Peter Norvig. (Artificial Intelligence: A Modern Approach, 4th US ed. (berkeley.edu))
3. Deep Learning, Ian Goodfellow, Yoshua Bengio, and Aaron Courville. (Deep Learning (deeplearningbook.org))
4. David Silver's course on Reiforcement Learning

E Books / E Learning References:
1. https://nptel.ac.in/courses/106106143 2. http://cse.iitkgp.ac.in/~aritrah/course/theory/RL/Autumn2022 3. https://onlinecourses.nptel.ac.in/noc20_cs74/preview 4. https://www.davidsilver.uk/teaching/

Savitribai Phule Pune University, Pune Final Year Artificial Intelligence and Machine Learning (2020 Course) 418552B: Elective VI- (BigData Analytics)		
Teaching Scheme:	**Credit Scheme:**	**Examination Scheme:**
Theory (TH) : 3 hrs/week	**03 Credits**	**Mid_Semester : 30 Marks** **End_Semester : 70Marks**

Prerequisite Courses, if any: Should have knowledge of one Programming Language (Java preferably), Practice of SQL (queries and sub queries), exposure to Linux Environment.

Companion Course, if any: Big Data Computing by PROF. RAJIV MISRA, Dept. of Computer Science and Engineering, IIT Patna.

Course Objectives:

1. Understand the Big Data Platform and its Use cases
2. Provide HDFS Concepts and Interfacing with HDFS
3. Understand Map Reduce Jobs
4. Provide hands on Hodoop Eco System
5. Exposure to Data Analytics with R
6. Understand Future Emerging tools for Data Analytics.

Course Outcomes:

On completion of the course, students will be able to–

CO1 : Identify Big Data and its Business Implications.
CO2 : List the components of Hadoop and Hadoop Eco-System
CO3 : Manage Job Execution in Hadoop Environment
CO4 : Develop Big Data Solutions using Hadoop Eco System
CO5 : Apply Machine Learning Techniques using R.
CO6 : Analyze Infosphere BigInsights Big Data Recommendations.

COURSE CONTENTS

Unit I	**Introduction To Big Data and Bigdata Analytics**	**(6 hrs)**

Introduction to big data: Data, Characteristics of data and Types of digital data: Unstructured, Semi-structured and Structured, Sources of data, Working with unstructured data, Evolution and Definition of big data, Characteristics and Need of big data, Challenges of big data, Data environment versus big data environment

Big data analytics: Data science and Analytics, Meaning and Characteristics of big data analytics, Need of big data analytics, Classification of analytics, Challenges to big data analytics, Importance of big data analytics, Basic terminologies in big data environment

Mapping of Course Outcomes for Unit I	**CO1**	
Unit II	**HDFS (Hadoop Distributed File System)**	**(6 hrs)**

History of Hadoop, Apache Hadoop, Analysing Data with Unix tools, Analysing Data with Hadoop, The Design of HDFS, HDFS Concepts, Command Line Interface, Hadoop file system interfaces, Data flow, Data Ingest with Flume and Scoop and Hadoop archives, Hadoop I/O: Compression, Serialization, Avro and File-Based Data structures.

Mapping of Course Outcomes for Unit II	**CO2**	
Unit III	**Mongodb and Mapreduce Programming**	**(6 hrs)**

Anatomy of a Map Reduce Job Run, Failures, Job Scheduling, Shuffle and Sort, Task Execution, Map Reduce Types and Formats, Map Reduce Features. Introduction to MongoDB and its needs, Characteristics of MongoDB.

Mapping of Course Outcomes for Unit III	**CO3**	
Unit IV	**Hadoop Eco System**	**(6 hrs)**

Pig : Introduction to PIG, Execution Modes of Pig, Comparison of Pig with Databases, Grunt, Pig Latin, User Defined Functions, Data Processing operators. **Hive :** Hive Shell, Hive Services, Hive Metastore, Comparison with Traditional Databases, HiveQL, Tables, Querying Data and User Defined Functions. **Hbase :** HBasics, Concepts, Clients, Example, Hbase Versus RDBMS. Big SQL : Introduction

Mapping of Course Outcomes for Unit IV	**CO4**	
Unit V	**Data Analytics with R**	**(6 hrs)**

Big Data Analytics with BigR. Exploratory Data Analysis, Linear (Multiple Regression) Models and Analysis of Variance.

Graphical Data Analysis with R :Various types of plots drawn in R programming, Appropriate Graph in R, R Graphical Models, Types , Conditional Independence in Graphs

Mapping of Course Outcomes for Unit V	**CO5**	
Unit VI	**Recent Trends in Big data : Spark, Cassandra, Xplenty**	**(6 hrs)**

Streaming Analytics, Rise of AI-Powered BigData Analytics, DataOps for Data, Dark Data, Real-time Analytics.
Tools: Spark, Cassandra, Xplenty.

Mapping of Course Outcomes for Unit VI	**CO6**

Text Books:

1. Tom White " Hadoop: The Definitive Guide" Third Edit on, O'reily Media, 2012.
2. Seema Acharya, Subhasini Chellappan, "Big Data Analytics" Wiley 2015. References
3. Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer, 2007.
4. Jay Liebowitz, "Big Data and Business Analytics" Auerbach Publications, CRC press (2013)
5. Tom Plunkett, Mark Hornick, "Using R to Unlock the Value of Big Data: Big Data Analytics with Oracle R Enterprise and Oracle R Connector for Hadoop", McGraw-Hill/Osborne Media (2013), Oracle press.

Reference Books:

1. Anand Rajaraman and Jef rey David Ulman, "Mining of Massive Datasets", Cambridge University Press, 2012. • Bill Franks, "Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics", John Wiley & sons, 2012.
2. Glen J. Myat, "Making Sense of Data", John Wiley & Sons, 2007
3. Pete Warden, "Big Data Glossary", O'Reily, 2011.
4. Michael Mineli, Michele Chambers, Ambiga Dhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley Publications, 2013.
5. ArvindSathi, "BigDataAnalytics: Disruptive Technologies for Changing the Game", MC Press, 2012 • Paul Zikopoulos ,Dirk DeRoos , Krishnan Parasuraman , Thomas Deutsch , James Giles , David Corigan , "Harness the Power of Big Data The IBM Big Data Platform ", Tata McGraw Hill Publications, 2012.

E Books / E Learning References :

1. https://www.geeksforgeeks.org/exploratory-data-analysis-in-r-programming/
2. https://data-flair.training/blogs/graphical-data-analysis-with-r/
3. https://online.umich.edu/series/data-analytics-in-the-public-sector-with-r/
4. https://cran.r-project.org/doc/contrib/usingR.pdf
5. https://archive.nptel.ac.in/courses/106/104/106104189/
6. https://www.coursera.org/learn/introduction-to-big-data-with-spark-hadoop

Savitribai Phule Pune University, Pune
Final Year Information Technology (2019 Course)
418552C: Elective VI- (Artificial Intelligence using R programming)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory (TH) : 3 hrs/week	03	Mid_Semester : 30 Marks End_Semester : 70 Marks

Prerequisite Courses, if any: Python programming , Fundamentals of AI

Companion Course, if any: Statistics, Machine Learning

Course Objectives:

1. To understand the basics in R programming in terms of constructs, control statements, string functions.
2. To be able to appreciate and apply the R programming from a statistical perspective.
3. To understand the concept of regression
4. To implement Machine learning algorithms using R

Course Outcomes:

On completion of the course, students will be able to–

CO1: Understand the use of R programming language.
CO2: Use programming structures like loops, functions, exceptions in R.
CO3: Understand the basic terminologies of statistics used in AI.
CO4: Understand the basic terminologies of probability used in AI.
CO5: To understand the concept of regression.
CO6: To implement Machine learning algorithms using R.

COURSE CONTENTS

Unit I	Introduction to R	(6 hrs)

Getting Started : Obtaining and Installing R from CRAN, Opening R for the First Time: Console and Editor Panes, Comments, Working Directory , Installing and Loading R Packages, Help Files and Function Documentation, Third-Party Editors, Workspaces, Scripts, Conventions: Coding, Math and Equation References

Numerics, Arithmetic assignment and vectors : R for Basic Math, Arithmetic, Logarithms and Exponentials, E-Notation, Assigning Objects, Vectors : Creating a Vector, Sequences, Repetition, Sorting, and Lengths, Subsetting and Element Extraction, Vector-Oriented Behavior

Conditions and loops : if Statements, Stand-Alone Statement, else Statements, Using if else for Element-wise Checks, Nesting and Stacking Statements, the switch Function .

Mapping of Course Outcomes for Unit I	CO1

Unit II	Programming in R	(6 hrs)

Coding Loops : for Loops, while Loops ,Implicit Looping with apply , Declaring break or next ,The repeat Statement

Writing functions : The function Command, Function Creation ,Using return, Arguments ,Lazy Evaluation, Setting Defaults ,Checking for Missing Arguments ,Dealing with Ellipses ,Specialized Functions , Helper Functions ,Disposable Functions ,Recursive Functions

Exceptions, timings and Visibility : Exception Handling , Errors and Warnings ,Catching Errors with try Statements ,Progress and Timing , Textual Progress Bars, Measuring Completion Time

Mapping of Course Outcomes for Unit II	**CO2**

Unit III	**Statistics**	**(6 hrs)**

Elementary statistics: Describing Raw Data , Numeric Variables, Categorical Variables, Univariate and Multivariate Data

Summary Statistics : Centrality: Mean, Median, Mode , Counts, Percentages, and Proportions, Quantiles, Percentiles, and the Five-Number Summary , Spread: Variance, Standard Deviation, and the Interquartile Range ., Covariance and Correlation , Outliers

Mapping of Course Outcomes for Unit III	**CO3**

Unit IV	**Probability**	**(6 hrs)**

Data Visualization : Barplots and Pie Charts, Building a Barplot , A Quick Pie Chart , Histograms , Box-and-Whisker Plots, Stand-Alone Boxplots, Side-by-Side Boxplots, Scatterplots, Single Plot, Matrix of Plots

Probability : What Is a Probability? ,Events and Probability , Conditional Probability, Intersection ,Union, Complement, Random Variables and Probability Distributions, Realizations, Discrete Random Variables ,Continuous Random Variables ,Shape, Skew, and Modality

Mapping of Course Outcomes for Unit IV	**CO4**

Unit V	**Simple Linear Regression**	**(6 hrs)**

General Concepts : Definition of the Model, Estimating the Intercept and Slope Parameters , Fitting Linear Models with lm , Illustrating Residuals,

Statistical Inference : Summarizing the Fitted Model Regression, Coefficient Significance Tests, Coefficient of Determination

Prediction : Confidence Interval or Prediction Interval? , Interpreting Intervals, Plotting , Interpolation vs. Extrapolation, Understanding Categorical Predictors, Binary Variables: k = 2 , Multilevel Variables: k > 2, Changing the Reference Level , Treating Categorical Variables as Numeric

Mapping of Course Outcomes for Unit V	**CO5**

Unit VI	**Multiple Linear Regression**	**(06 hrs)**

Terminology , **Theory** : Extending the Simple Model to a Multiple Model, Estimating in Matrix form **Implementing in R and Interpreting** : Additional Predictors, Interpreting Marginal effects, Visualizing the Multiple Linear Model , Finding Confidence Intervals , Omnibus F-Test, Predicting from a Multiple Linear Model **Machine Learning in Action** : Packages, Dataset , Data partitioning , Linear model , Prediction , Logistic regression , Residuals, Least squares regression , Relative importance , Stepwise regression , The k-nearest neighbor classification , Naïve Bayes , **The train Method** : Support vector machines, K-means clustering , Decision trees , AdaBoost , Neural network ,Random forests

Mapping of Course Outcomes for Unit VI	**CO6**

Text Books:

1. "The book of R : A first course in programming and statistics ", Tilman A. Davies, No Starch press
2. "R for Data Science: Learn and explore the fundamentals of data science with R", Dan Toomey, Packt Publishing

Reference Books:

1. Jared P. Lander, "R for Everyone: Advanced Analytics and Graphics", Addison-Wesley Data & Analytics Series, 2013.
2. Mark Gardener, " Beginning R – The Statistical Programming Language", Wiley, 2013
3. Robert Knell, "Introductory R: A Beginner's Guide to Data Visualisation, Statistical Analysis and Programming in R", Amazon Digital South Asia Services Inc, 2013.'

Savitribai Phule Pune University, Pune Final Year Artificial Intelligence and Machine Learning (2020 Course) 418553: Startup and Entrepreneurship		
Teaching Scheme:	**Credit Scheme:**	**Examination Scheme:**
Tutorial (TUT) : 03 hrs/week	**03 Credits**	**TW: 50 Marks**

Prerequisite Courses, if any:

Course Objectives:

1. To encourage students to build new technology, knowledge system based on innovations and can address local challenges.
2. Creating environment to innovate and build products towards sustainable development goals.
3. To provide platform for speedy communication and market reach of technology/ product developed by students.
4. To have start up ecosystem by bridging the gap between academia, industries and financial institutions, government support

Course Outcomes:

On Completion of Course students will be able to:-

CO1. understand key concepts and framework of innovation and start-up ecosystem.
CO2. gain knowledge of how to develop start up ecosystem, its key components and how to influence and managedynamics between them and increase the productivity of ecosystem.
CO3. understand the role of different stakeholders in ecosystem in building and supporting growth of start-ups.
CO4. have insight into global trend in start-up ecosystem and product development.
CO5. mapping different start-up ecosystems and developing performance indicators.

COURSE CONTENTS

Unit I	**Start-up Opportunity**	**(3 hrs)**
Identify business opportunity with problem identification, market size, existing pains for customers, existing alternatives, customer psychology, willingness to pay, customer segments.		
Mapping of Course Outcomes for Unit I	**CO1**	
Unit II	**Product/ Service Proposal**	**(3 hrs)**
Value Proposition Canvas, problem-solution fit, brainstorming, competition analysis, creating competitive advantage, sustainable differentiation.		
Mapping of Course Outcomes for Unit II	**CO2**	
Unit III	**Business model**	**(3 hrs)**
Types, Lean canvas, Risky assumptions related to product, market, business, and execution capabilities		
Mapping of Course Outcomes for Unit III	**CO3**	

Unit IV	Minimum Viable Product (MVP)	(3 hrs)
Create and iterate, testing of MVP, customer feedback, validate risky assumptions, solution-market fit		
Mapping of Course Outcomes for Unit IV	**CO4**	
Unit V	**Financial Plan**	**(3 hrs)**
Manpower, Sales, Expenses, profitability projections, reality check, Funding plan, Pitch deck		
Mapping of Course Outcomes for Unit V	**CO5**	
Unit VI	**Marketing strategy**	**(3 hrs)**
Importance of brand and branding strategy, positioning, market penetration strategy/ plan, digital marketing, use of social media, customer acquisition Use of technology: for business scalability, effective execution, growth plan		
Mapping of Course Outcomes for Unit VI	**CO6**	
E Books / E Learning References:		

1. https://www.forbes.com/sites/palomacanterogomez/2019/04/10/how-to-frame-a-problem-to-find-the-right-solution/?sh=13af54355993
2. https://hbswk.hbs.edu/item/how-entrepreneurs-can-find-the-right-problem-to-solve
3. https://www.youtube.com/watch?v=6y3WIrgp_NY
4. https://hbr.org/2014/07/what-you-need-to-know-about-segmentation
5. https://www.youtube.com/watch?v=ReM1uqmVfP0
6. https://www.youtube.com/watch?v=w62zW30PKms
7. https://www.youtube.com/watch?v=FULiFueLGzE
8. https://www.youtube.com/watch?v=7o8uYdUaFR4
9. https://steveblank.com/2021/04/20/the-secret-to-the-minimum-viable-product/
10. https://www.youtube.com/watch?v=1hHMwLxN6EM
11. https://www.youtube.com/watch?v=4uGx14UVWPc
12. https://www.youtube.com/watch?v=OVnN4S52F3k
13. https://www.entrepreneur.com/article/251687
14. https://www.forbes.com/sites/forbesbusinessdevelopmentcouncil/2020/09/14/13-key-steps-to-developing-a-go-to-market-strategy/?sh=53023c476fc1
15. https://www.garyfox.co/business-model/business-model-channels/
16. https://www.forbes.com/sites/allbusiness/2019/05/25/small-business-website-tips/?sh=2c551a0421ad
17. https://www.forbes.com/sites/forbesagencycouncil/2020/10/08/digital-marketing-best-practices- for-startups/?sh=2e55af9e3ded

Savitribai Phule Pune University, Pune

Final Year Artificial Intelligence and Machine Learning (2020 Course)

418554 : Lab Practice V (Natural Language Processing)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Practical: 2 Hours/Week	01 Credit	Term Work: 50 Marks

Prerequisite Course : Discrete Mathematics, Theory of Computation

Companion Course: Python Programming

Course Objectives:

- To understand the fundamental concepts and techniques of natural language processing (NLP)

Course Outcomes:

On completion of the course, students will be able to–

Course Outcomes:

On completion of this course, the students will be able to

CO1: Apply basic principles of elective subjects to problem solving and modeling.

CO2: Use tools and techniques in the area of software development to build mini projects

CO3: Design and develop applications on subjects of their choice.

CO4: Generate and manage deployment, administration & security.

Natural Language Processing

Any 5 Assignments and 1 Mini Project are mandatory

Group 1

1. Perform tokenization (Whitespace, Punctuation-based, Treebank, Tweet, MWE) using NLTK library. Use porter stemmer and snowball stemmer for stemming. Use any technique for lemmatization.
2. Perform bag-of-words approach (count occurrence, normalized count occurrence), TF-IDF on data. Create embeddings using Word2Vec.
3. Perform text cleaning, perform lemmatization (any method), remove stop words (any method), label encoding. Create representations using TF-IDF. Save outputs.
4. Create a transformer from scratch using the Pytorch librar
5. Morphology is the study of the way words are built up from smaller meaning bearing units. Study and understand the concepts of morphology by the use of add delete table

Group 2 Mini Project

1. Fine tune a pre-trained transformer for any of the following tasks on any relevant dataset of your choice:
   - Neural Machine Translation
   - Classification
   - Summarization
2. POS Taggers For Indian Languages
3. Feature Extraction using seven moment variants
4. Feature Extraction using Zernike Moments

Savitribai Phule Pune University, Pune
Final Year Artificial Intelligence and Machine Learning (2020 Course)
418555 : Lab Practice VI (Big data analytics)

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Practical (PR): 02 hrs/week	01	Term Work: 25 Oral : 50

Prerequisite Courses, if any: Database Management Systems

Companion Course, if any: Practical hands on is the absolute necessity as far as employability of the learner is concerned. The presented course is solely intended to enhance the competency by undertaking the laboratory assignments of the core courses.

Course Objectives:

1. To develop problem solving abilities using Mathematics
2. To apply algorithmic strategies while solving problems
3. To develop time and space efficient algorithms
4. To study algorithmic examples in distributed, concurrent and parallel environments

Course Outcomes:

On completion of the course, students will be able to–

CO1: Write case studies in Business Analytic and Intelligence using mathematical models

CO2: Present a survey on applications for Business Analytic and Intelligence

CO3: Provide problem solutions for multi-core or distributed, concurrent/Parallel environments

Guidelines for Laboratory Conduction

- List of recommended programming assignments and sample mini-projects is provided for reference.
- Referring these, Course Teacher or Lab Instructor may frame the assignments/mini-project by understanding the prerequisites, technological aspects, utility and recent trends related to the respective courses.
- Preferably there should be multiple sets of assignments/mini-project and distribute among batches of students.
- Real world problems/application based assignments/mini-projects create interest among learners serving as foundation for future research or startup of business projects.
- Mini-project can be completed in group of 2 to 3 students.
- Software Engineering approach with proper documentation is to be strictly followed.
- Use of open source software is to be encouraged.
- Instructor may also set one assignment or mini-project that is suitable to respective course beyond the scope of syllabus.
- Operating System recommended :- 64-bit Open source
- Programming Languages: PYTHON/R
- Programming tools recommended: Front End: Java/Perl/PHP/Python/Ruby/.net,
- Backend :MongoDB/MYSQL/Oracle,
- Database Connectivity: ODBC/JDBC, Additional Tools: Octave, Matlab, WEKA.

Guidelines for Instructor's Manual
The instructor‟s manual is to be developed as a hands-on resource and as ready reference. The instructor's manual need to include prologue (about University/program/ institute/ department/foreword/ preface etc), University syllabus, conduction and Assessment guidelines, topics under consideration-concept, objectives, outcomes, set of typical applications/assignments/ guidelines, references among others.
Guidelines for Student Journal
The laboratory assignments are to be submitted by student in the form of journal. Journal may consists of prologue, Certificate, table of contents, and handwritten write-up of each assignment (Title, Objectives, Problem Statement, Outcomes, software and Hardware requirements, Date of Completion, Assessment grade/marks and assessor's sign, Theory- Concept in brief, Algorithm/Database design, test cases, conclusion/analysis). **Program codes with sample output of all performed assignments are to be submitted as softcopy.** As a conscious effort and little contribution towards Green IT and environment awareness, attaching printed papers as part of write-ups and program listing to journal may be avoided. Use of digital storage media/DVD containing student's programs maintained by lab In-charge is highly encouraged. For reference one or two journals may be maintained with program prints at Laboratory.
Guidelines for Lab /TW Assessment
Continuous assessment of laboratory work is to be done based on overall performance and lab assignments performance of student. Each lab assignment assessment will assign grade/marks based on parameters with appropriate weightage. Suggested parameters for overall assessment as well as each lab assignment assessment include- timely completion, performance, innovation, efficient codes, punctuality and neatness reserving weightage for successful mini-project completion and related documentation.
Guidelines for Practical Examination
• Both internal and external examiners should jointly frame suitable problem statements for practical examination based on the term work completed. • During practical assessment, the expert evaluator should give the maximum weightage to the satisfactory implementation of the problem statement. • The supplementary and relevant questions may be asked at the time of evaluation to test the student‟s for advanced learning, understanding of the fundamentals, effective and efficient implementation. • Encouraging efforts, transparent evaluation and fair approach of the evaluator will not create any uncertainty or doubt in the minds of the students. So adhering to these principles will consummate our team efforts to the promising boost to the student's academics.

Perform any 4 assignment and mini-project is compulsory

1. Download the Iris flower dataset or any other dataset into a DataFrame. (eg https://archive.ics.uci.edu/ml/datasets/Iris) Use Python/R and Perform following –
   - How many features are there and what are their types (e.g., numeric, nominal)?
   - Compute and display summary statistics for each feature available in the dataset.(eg. minimum value, maximum value, mean, range, standard deviation, variance and percentiles
   - Data Visualization-Create a histogram for each feature in the dataset to illustrate the feature distributions. Plot each histogram.
   - Create a boxplot for each feature in the dataset. All of the boxplots should be combined into a single plot. Compare distributions and identify outliers.
2. Download Pima Indians Diabetes dataset. Use Naive Bayes‟ Algorithm for classification
   - Load the data from CSV file and split it into training and test datasets.
   - Summarize the properties in the training dataset so that we can calculate probabilities and make predictions.
   - Classify samples from a test dataset and a summarized training dataset.
3. Write a Hadoop program that counts the number of occurrences of each word in a text file.
4. Write a program that interacts with the weather database. Find the day and the station with the maximum snowfall in 2013
5. Use Movies Dataset. Write the map and reduce methods to determine the average ratings of movies. The input consists of a series of lines, each containing a movie number, user number, rating, and a timestamp: The map should emit movie number and list of rating, and reduce should return for each movie number a list of average rating.
6. Trip History Analysis: Use trip history dataset that is from a bike sharing service in the United States. The data is provided quarter-wise from 2010 (Q4) onwards. Each file has 7 columns. Predict the class of user. Sample Test data set available here https://www.capitalbikeshare.com/trip-history-data
7. Bigmart Sales Analysis: For data comprising of transaction records of a sales store. The data has 8523 rows of 12 variables. Predict the sales of a store. Sample Test data set available here https://datahack.analyticsvidhya.com/contest/practice-problem-big-mart-sales-iii/
8. Twitter Data Analysis: Use Twitter data for sentiment analysis. The dataset is 3MB in size and has 31,962 tweets. Identify the tweets which are hate tweets and which are not. Sample Test data set available here https://datahack.analyticsvidhya.com/contest/practice-problemtwitter-sentiment-analysis/
9. Time Series Analysis: Use time series and forecast traffic on a mode of transportation. Sample Test data set available here https://datahack.analyticsvidhya.com/contest/practice-problemtime-series-2/

Reference Books:

1. David Dietrich, Barry Hiller, “Data Science and Big Data Analytics”, EMC education services, Wiley publications, 2012, ISBN0-07-120413-X.
2. Ashutosh Nandeshwar , “Tableau Data Visualization Codebook”, Packt Publishing, ISBN 978-1-84968-978-6.
3. Maheshwari Anil, Rakshit, Acharya, “Data Analytics”, McGraw Hill, ISBN: 789353160258.
4. Mark Gardner, “Beginning R: The Statistical Programming Language”, Wrox Publication, ISBN: 978-1-118-16430-3
5. Luís Torgo, “Data Mining with R, Learning with Case Studies”, CRC Press, Talay and Francis Group, ISBN9781482234893
6. Carlo Vercellis, “Business Intelligence - Data Mining and Optimization for Decision Making”, Wiley Publications, ISBN: 9780470753866.

Savitribai Phule Pune University, Pune

Final Year Artificial Intelligence and Machine Learning (2020 Course)

418556: Project Stage-II

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Practical: 10 hrs/week	05 Credits	Term Work : 100 Marks Oral : 50 Marks

Prerequisite Courses, if any: Project Phase-I (B.E. (AI & ML) Final Year Semester-I)

Companion Course, if any: NA

Course Objectives:

1. To enable the student to extend further the investigative study taken up under Project stage-I, either fully theoretical/practical or involving both theoretical and practical work, under the guidance of a Supervisor from the Department alone or jointly with a Supervisor drawn from R&D laboratory / Industry.
2. To build up exposure of implementation and hence develops analysis of results by considering performance measures.
3. To expose students to product development environment using industrial experience, use of state of art technologies.
4. To encourage and expose students with funding agency for sponsored projects.
5. To generate the opportunities to experience practically the facts learned in various fields together.
6. To improve overall communication skill, Teamwork and Leadership Qualities, professionalism.
7. Evaluate the various validation and verification methods.
8. Analyzing professional issues, including ethical, legal and security issues, related to computing projects.
9. To evaluate alternative approaches, and justify the results obtained.

Course Outcomes:

On completion of the course, students will be able to–

1. To apply engineering and mathematical knowledge to investigate / select proper technology / Algorithm suitable to solve the problem in hand.
2. To apply knowledge of statistics for analysis of results and express conclusion and justification for the same.
3. To design and conduct experiments, as well as to analyze and interpret data or develop prototype model of the application.
4. To communicate effectively.
5. Get broad education which is necessary to understand the impact of engineering solutions in a global, economic, environmental, ethically and societal context.
6. Recognition of the need for, and an ability to engage in life-long learning.

Introductory Information:

BE Project Phase-II is the continuation of Project Phase-I for implementation, and analysis of results to arrive a valid conclusion with justification.

Guidelines to Faculty and Students:

1. Preferably same review committee needs to continue for Project Phase-II.
2. There shall be **TWO** reviews in Project phase –II (in semester-II) by the review committee.
3. The Project Review committee will be responsible for evaluating the timely progress of the projects. It is suggested to evaluate the skills learned by the students in their PBL (in their previous years).
4. Student needs to justify the Algorithm / Model used for implementation.
5. Every student of the project group shall make presentation on the progress made by them before the committee during each reviews. Each student/group is required to give presentation as part of review for 10 to 15 minutes followed by a detailed discussion and query session.
6. Students need to note down the queries raised during review(s) and comply the same in the next review session.
7. The record of the remarks/suggestions of the review committee (project dairy) should be properly maintained in continuation of Project Phase-II and should be made available at the time of university examination.
8. Project group needs to present / publish **TWO** papers (One in each semester, at least one paper should be in **UGC – Care journal**).
    a. Paper must be checked for Plagiarism by any open software.
    b. One paper during second semester which includes Methodologies / Algorithms implemented, Results obtained, Analysis of results and conclusion.
9. Project report must also be checked for Plagiarism.
10. The examinee will be assessed by a panel of examiners of which one is necessarily an external examiner. The assessment will be broadly based on work undergone, content delivery, presentation skills, documentation, question-answers and report.

Review 3: Implementation –

Points to be covered:

1. Detailed study of Algorithm(s) / Model / Hardware specification (As applicable).
2. Confirmation of Data set used (As applicable)
3. Detailed ER Diagram / DFD diagrams.
4. Detailed UML Diagrams.
5. Sample results (module based).

Review 4: Testing and Result Analysis.

Points to be covered:

1. Appropriate test cases and results of test cases.
2. Representation of results with analysis.
3. Conclusion over performance parameters (as applicable)
4. Conclusion and future work suggested.
5. Knowledge of references utilized.

Evaluation Criteria:

Following criteria and weightage is suggested for evaluation of Project-Phase II Term Work.

1. Availability of standard Data set / Input parameters: 10%
2. Depth of Understanding of implemented Technology / Algorithm / Domain / Model: 40%
3. Test cases / Validation and Verification process: 10%
4. Justification of Algorithm / Model / Architecture / System: 10%
5. Analysis of results and conclusion: 10%
6. Presentation Skill: 10%
7. Report preparation and Paper publication: 10%

Project report contains the details as Follows:

It is suggested to have only one Project report which includes work carried at Project Phase-I as well. Project report must have:

i. Certificate from the institute.
ii. Certificate sponsoring organization (If any).
iii. Acknowledgement.
iv. Abstract.
v. Contents.
vi. List of Abbreviations (As applicable).
vii. List of Figures (As applicable).
viii. List of Graphs (As applicable).
ix. List of Tables (As applicable).

1) Introduction and aims/motivation and objectives.
2) Literature Survey (with proper citation).
3) Problem Statement/definition.
4) Software Requirement Specification (In SRS Documentation only).
5) Flowchart
6) Project Requirement specification.
7) Proposed system Architecture.
8) High level design of the project (DFD , UML , ER Diagrams).
9) System implementation-code documentation: Algorithm style, Description of detailed methodologies, protocols used etc..as applicable.
10) Test cases.
11) GUI/Working modules and Experimental Results in suitable format.
12) Project Plan.
13) Analysis and Conclusions with future work.
14) Bibliography in IEEE format.

Appendices

a) Plagiarism Report of Paper and Project report from any open source tool.
b) Base Paper(s) [If any].
c) Tools used / Hardware Components specifications [If any].
d) Published Papers and Certificates (Both Papers).

Use appropriate plagiarism tools, reference managers, Latex for efficient and effective project writing.

Savitribai Phule Pune University, Pune

Final Year Artificial Intelligence and Machine Learning (2020 Course)

418557A: Audit Course 8

Functional Programming in Haskell

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory (TH): 01 hrs/week	Non-Credit	Audit Course

Prerequisite Courses: Programming using any high-level language.

Course Objectives:

1. To understand the paradigm of programming.
2. To develop insight about 'lazy' execution.
3. To learn the syntax and semantics of the Haskell programming language.
4. To learn 'idioms' of Haskell programming

Course Outcomes:

On completion of the course, students will be able to–

CO1. Understand the correctness of programs.

CO2. Make use of higher-order functions.

CO3. Make use of the data encapsulation and parametric polymorphism for functional programming.

CO4. Comprehend the importance of the 'type checking' of values/functions to develop programs relatively faster.

COURSE CONTENTS

Unit I	Introduction	(3 hrs)
Types and Values, Running Haskell Programs, Lists, Strings, Tuples. Introduction to ghci interpreter		
Mapping of Course Outcomes for Unit I	**CO1**	
Unit II	**Functions**	**(3 hrs)**
Functions, Type Inference, Recursion, Higher-order Functions, Polymorphic Types, Lambda Functions. Computation as rewriting, lazy evaluation and infinite data structures		
Mapping of Course Outcomes for Unit II	**CO2, CO3**	
Unit III	**Data Types**	**(3 hrs)**
User defined Data Types, Abstract data types, Recursive Data Types-Binary search trees		
Mapping of Course Outcomes for Unit III	**CO4**	
Unit IV	**Arrays and IO**	**(3hrs)**
Arrays, Input / Output		
Mapping of Course Outcomes for Unit IV	**CO4**	

Textbooks:
1. Brian O'Sullivan, John Goerzen and Don Stewart, 'Real World Haskell', O'reilly. **2.** MiranLipovača, 'Learn You a Haskell for Great Good!', No Starch Press. **3.** Graham Hutton, "Programming in Haskell", Cambridge University Press. **4.** https://nptel.ac.in/courses/106106137
Evaluation
Students should select any one of the topic in a group of 3 to 5. Students should submit a written report and make a presentation on the topic. The task should not be repeated among students. Report will be evaluated by the faculty as per rubrics defined by him/her/them at start of course.

Savitribai Phule Pune University, Pune

Final Year Artificial Intelligence and Machine Learning (2020 Course)

418557B: Audit Course 8

Cyber Laws And Use Of Social Media

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory (TH): 01 hrs/week	Non-Credit	Audit Course

Prerequisite Courses: Programming using any high-level language.

Course Objectives:

1. To understand and be aware of Cyber laws which focus on protecting the privacy of users from organizations and other users.
2. To know the cyber threats happening around them and to help them stay secure in the daily use of Cyberspace.

Course Outcomes:

On completion of the course, students will be able to–

CO1. Understand the importance of the IT Act.

CO2. Understand the significance of cyber laws and their practices.

CO3. Identify and Analyze software vulnerabilities and security solutions to reduce the risk of exploitation.

CO4. To study various privacy and security concerns of Online social media.

COURSE CONTENTS

Unit I	Introduction to the IT Act	(03 hrs)

Evolution of the IT Act, Genesis and Necessity Various authorities under IT Act and their powers: Penalties & Offences, amendments. Traditional Principals of Jurisdiction, Extra-terrestrial Jurisdiction, Case Laws on Cyber Space Jurisdiction

Mapping of Course Outcomes for Unit I	CO1

Unit II	Cyber Law: International Perspective	(03 hrs)

EDI: Concept and Legal Issues, UNCITRAL Model Law, Electronic Signature Laws of Major Countries, Cryptography Laws, Cyber Laws of Major Countries, EU Convention on Cyber Crime

Mapping of Course Outcomes for Unit II	CO2

Unit III	Cyber Forensic and Computer Crimes	(03 hrs)

Types, Crimes targeting Computers: Definition of Cyber Crime & Computer-related crimes.
Classification & Differentiation between traditional crime and cyber-crimes.
Cyber-crimes and cyber terrorism: -
a) Cyber-crimes and the categories of crime i) Cyber frauds ii) Cyber thefts iii) Cyber stacking
b) Cyber Terrorism. c) Hacking, Viruses, Trojans, worms etc.

Mapping of Course Outcomes for Unit III	CO3

Unit IV	Use of Social Media	(03 hrs)

Elements of Social Networks, Social Media Outlets. (Facebook, Twitter, etc.): How the differences impact how to use them. Videos: Broadcasting to peers, many to many, friends and followers, apps, pages, pseudonyms of good and evil Focused Networks (Flickr, Linked In, YouTube, etc.) networks that focus on specific topics or activities	
Mapping of Course Outcomes for Unit IV	**CO4**
Textbooks:	
1. The Information Technology Act, 2000, Bare Act-Professional Book Publishers, New Delhi. **2.** Aparna Viswanathan, "Cyber Law- Indian and International Perspectives On Key Topics Including Data Security, E-Commerce, Cloud Computing and Cyber Crimes". **3.** First Responder's Guide to Computer Forensics by Richard Nolan et al. Carnegiei Mellon, 2005. **4.** https://nptel.ac.in/courses/106106146	
Evaluation	
Students should select any one of the topics in a group of 3 to 5. Students should submit a written report and make a presentation on the topic. The task should not be repeated among students. The report will be evaluated by the faculty as per rubrics defined by him/her/them at start of course.	

Savitribai Phule Pune University, Pune

Final Year Artificial Intelligence and Machine Learning (2020 Course)

418557C: Audit Course 8

Constitution Of India

Teaching Scheme:	Credit Scheme:	Examination Scheme:
Theory(TH): 01 hrs/week	Non-Credit	Audit Course

Prerequisite Courses, if any:

Course Objectives:

1. Understand the premises informing the twin themes of liberty and freedom from a civil rights perspective.
2. To address the growth of Indian opinion regarding modern Indian intellectuals' constitutional role and entitlement to civil and economic rights.
3. To address the role and functions of local administration.

Course Outcomes:

On completion of the course, students will be able to–

CO1. Understand the Principles of the Indian Constitution.

CO2. Understand and identify the growth of the demand for civil rights in India.

CO3. Understand the organizations of governance.

CO4. Understand the role and functions of local administration.

COURSE CONTENTS

Unit I	History of Making of the Indian Constitution	(03 hrs)
History Drafting Committee, (Composition & Working), Philosophy of the Indian Constitution: Preamble, Salient Features		
Mapping of Course Outcomes for Unit I	**CO1**	
Unit II	**Contours of Constitutional Rights & Duties**	**(03 hrs)**
Fundamental Rights, Right to Equality, Right to Freedom against Exploitation, Right to Freedom of Religion, Cultural and Educational Rights, Right to Constitutional Remedies, Directive Principles of State Policy, Fundamental Duties.		
Mapping of Course Outcomes for Unit II	**CO2**	
Unit III	**Organs of Governance:**	**(03 hrs)**
Parliament, Composition, Qualifications and Disqualifications, Powers and Functions, Executive, President, Governor, Council of Ministers, Judiciary		
Mapping of Course Outcomes for Unit III	**CO3**	
Unit IV	**Local Administration and Election Commission**	**(03 hrs)**

District's Administration head: Role and Importance, Municipalities: Introduction, Mayor and role of Elected representative, CEO of Municipal Corporation. Pachayati raj: Introduction, PRI: Zila Pachayat. Elected officials and their roles, CEO Zila Pachayat: Position and role. Block level: Organizational Hierarchy (Different departments), Village level: Role of Elected and Appointed officials, Importance of grass root democracy. Election Commission: Role and Functioning	
Mapping of Course Outcomes for Unit IV	**CO4**

Textbooks:

1. The Constitution of India, 1950 (Bare Act), Government Publication.
2. Dr. S. N. Busi, Dr. B. R. Ambedkar. Framing of of Indian Constitution, 1st Edition, 2015.
3. M. P. Jain, Indian Constitution Law, 7th Edition, Lexis Nexis, 2014.
4. D.D. Basu, Introduction to the Constitution of India, Lexis Nexis, 2015.
5. https://nptel.ac.in/courses/129106003

Evaluation:

Students should select any one of the topics in a group of 3 to 5. Students should submit a written Report. Make a presentation on the topic. The report will be evaluated by the faculty as per rubrics defined by them at start of course.