Kolhapur Institute of Technology's
College of Engineering
(Autonomous), Kolhapur

KOLHAPUR INSTITUTE OF TECHNOLOGY'S
COLLEGE OF ENGINEERING
KOLHAPUR
(EMPOWERED AUTONOMOUS)

Department of Computer Science &Engineering
Curriculum and Syllabus for
B. Tech (Hons.) Computer Science & Engineering with Specialization of Artificial Intelligence & Data Science

R.S.Sutar

Mrs. Ranjeeta Sutar
Academic Coordinator, CSE

Dr. Lingraj A. Hadimani
Head,
Department of Comp. Sci. & Engg.
KIT's College of Engg. (Autonomous)
Kolhapur

Dr. L. A. Hadimani
Head, CSE

B. Tech (Hons.) Computer Science & Engineering with Specialization of Artificial Intelligence & Data Science										
Course Code	**Course Name**	**L**	**T**	**P**	**Hrs /Week**	**Credits**	**Evaluation Scheme (Components)**			**Semester**
UCSHN0351	Mathematical Foundations for AI and ML	3	1		4	4	ESE	100	40	III
UCSHN0451	Data Engineering	3	1		4	4	ESE	100	40	IV
UCSHN0551	Artificial Intelligence and Machine Learning Programming	4		2	6	5	ESE	100	40	V
UCSHN0651	Deep Learning	3	1		4	4	ESE	100	40	VI
UCSH0701	Mini Project			2		1	ESE	100	40	VII
		13	3	4	18	18	**Total Marks: 500** **Total Credit: 18**			

Title of the Course: Mathematical Foundations for AI and ML (B. Tech Honors) Course Code: UCSHN0351	L	T	P	Credit
	3	1		4

Course Pre-Requisite: Linear algebra and a basic background in probability as well as basic experience in programming (e.g. Matlab, Python) will be required. Some basic knowledge in optimization is recommended.

Course Description: This course will cover the mathematical foundations and exact concepts behind some of the most important methods in machine learning and artificial intelligence. The emphasis in this course will be on the rigorous mathematical principles behind how and why methods work.

Course Objectives

1. To understand the mathematical concepts for AI and Machine Learning
2. Learn to implement algorithms in python
3. Understand the how the concepts extend for real world ML problems

CO	After the completion of the course the student should be able to	Bloom's Cognitive level	Descriptor
CO1	Understand the Mathematical concepts of AI ML and Modelling concepts	**2**	**Understand**
CO2	Demonstrate Linear algebra and Regression methods used as Foundations of ML	**4**	**Demonstrate**
CO3	Demonstrate statistical analysis and Multivariate analysis	**4**	**Demonstrate**
CO4	Analyze and apply appropriate mathematical techniques for solving real life problems	**2**	**Apply**

CO-PO Mapping:

CO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PSO 1	PSO 2
CO1	3		2									2	
CO2	3		2									2	
CO3	3		2									2	
CO4		2			3							1	2

Course Contents:

Unit 1. Introduction Introduction to AI & ML, Why AI & ML, Use Cases in Business and Scope, Scientific Method , Modeling Concepts.	5 Hrs.
Unit 2. Linear Algebra Vector and Matrix Norms, Vectors, Matrices, and Tensors in Python, Special Matrices and Vectors, Eigenvalues and Eigenvectors, Norms and Eigen decomposition.	7 Hrs.

Unit 3. Mathematical Foundations of ML Linear Regression method, Least squares method, Linear algebra solution to least squares problem, Examples of linear regression.	7 Hrs.
Unit 4: Statistical Analysis Initial Data Analysis, Relationship between attributes: Covariance, Correlation Coefficient, Chi Square Measure of Distribution (Skewness and Kurtosis), Box and Whisker Plot (Box Plot and its parts, Using Box Plots to compare distribution) and other statistical graphs	8 Hrs.
Unit 5: **Multivariate Analysis** Introduction to Derivatives, Basics of Integration, Gradients, Gradient Visualization, Optimization.	5 Hrs.
Unit 6: Probability Theory Introduction to Probability Theory, Probability Distributions, Expectation, Variance, and Covariance, Probability(Joint, marginal and conditional probabilities), Probability distributions (Continuous and Discrete), Density Functions and Cumulative functions	7 Hrs.
Textbooks: 1. Matrix Analysis (2nd ed.). Roger A. Horn, Charles R. Johnson. Cambridge University Press, 2013. 2. Introduction to Probability (2nd ed.). Dimitri P. Bertsekas, John N. Tsitsiklis. Athena Scientific, 2008.	

Title of the Course: Data Engineering Course Code: UCSHN0451	L	T	P	Credits
	3	1	-	4

Course Pre-Requisite: Statistics and Linear Algebra.

Course Description:

This course will cover the data pre-processing and data preparation for machine learning model.

Course Learning Objectives:

1. To understand the Core Concepts of Data Engineering
2. To design and Implement Data Pipelines
3. Work with Structured and Unstructured Data.
4. Apply Data Engineering Best Practices for Scalability and Performance

Course Outcomes:

COs	After Completion of the course, the student should be able to	Bloom's Cognitive	
		Level	Descriptor
CO1	Interpret the data properties.	2	Understand
CO2	Examine missing data and outliers.	4	Analyze
CO3	Interpret feature scaling.	2	Understand
CO4	Design feature extraction model	4	Analyze

CO-PO Mapping:

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO01	PSO02
CO1	3	2	1	1	2	1					3	1	1
CO2	3	2	2	3	2	1					3	1	1
CO3	2	2	1	3	2	1					3	1	1
CO4	2	2	2	3	3	1					3	2	3

Assessment:

Assessment	Marks
ESE	100

ESE: Assessment is based on 100% course content covered

Course Contents:

Unit 1:Data Foreseeing	**7 Hours**
Why data engineering? technical requirements of data during machine learning modeling:- identify numerical and categorical variables, missing data, determine cardinality in categorical variables, identifying linear relationship, identify normal distribution, highlighting outliers.	
Unit 2: Handling Missing Data and Data Encoding	**7 Hours**
Impute missing data: remove missing data, impute missing data by mean, mode or median, replacing missing values with value at the end of the distribution, multivariate imputation. Data Encoding: Why encoding? one hot encoding, replace categories with ordinal numbers, encoding with integers, encoding with mean of the target.	
Unit 3: Variable Discretization and Working with Outliers	**6 Hours**
Variable Discretization divide the variables into equal intervals, perform discretization followed by categorical encoding. Working with outliers: Outliers means? trimming outliers, capping the variables at arbitrary max and min values, perform zero coding.	
Unit 4: Feature Scaling	**7 Hours**
Standardizing features, mean normalization, scaling to max and min values. scaling with the median and quantiles, scaling to vector unit length, deriving new features with decision tree, carrying out PCA.	
Unit 5: Feature Creation with transaction and time series data	**7 Hours**
Aggregating transaction with mathematical operations, aggregating transaction in a time window, determining the number of local maxima and minima, deriving time elapsed between time-stamped events, creating features from transaction.	
Unit 6: Feature Extraction from text and Case study	**6 Hours**
Counting characters, words, vocabulary, estimating text complexity by counting sentences, create features with bag-of-words and n-grams, Case study: Data preprocessing @given dataset(includes exploratory data analysis(EDA), pipelining)	
Textbooks:	
1. Python Feature Engineering Cookbook by Soledad Galli. 2. Fundamentals of Data Engineering" by Joe Reis and Matt Housley	
Reference Books:	
1. Designing Data-Intensive Applications" by Martin Kleppmann Publisher: O'Reilly Media (2017) 2. Streaming Systems" by Tyler Akidau et al.Publisher: O'Reilly Media (2018) 3. Data Engineering with Python" by Paul Crickard Publisher: Packt Publishing (2020)	

Course Code:	UCSHN0551	L	T	P	Credit
Course Name:	Artificial Intelligence and Machine Learning Programming	4		2	5

Course Prerequisites:	Python Programming

Course Description:

This course aims to develop the basic concepts of Artificial Intelligence, Machine Learning Programming using Python Programming language.

Course Learning Objective:

1. The objective of this course is to provide comprehensive knowledge of python programming paradigms required for Artificial Intelligence,Machine Learning Programming.
2. To provide practical hands-on of Sci-Kit learn Python Machine learning libraries. 3. To demonstrate practical skills of huge data processing and analysis using python.

Course Outcomes:

COs	After completion of the course the student will be able to	Bloom's Cognitive Level	Bloom's Cognitive Descriptor
CO1	language	4	Demonstrate
CO2	Demonstrate significant experience with program development environment	4	Demonstrate
CO3	Pandas and MatplotLib modules	4	Demonstrate
CO4	solving real based data analytics techniques for solving rea	3	Apply

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2
CO1	3		2									2	
CO2	3		2									2	
CO3	3		2									2	
CO4		2			3							1	2

Assessment Scheme:

SN	Assessment	Marks	Remark
4	End Semester Examination (ESE)	100	100% course contents

Course Contents:

Unit 1	Introduction To Python - Sequence Data Types	6 Hours
Sequences, Tuple, Sets, Mapping and Sets- Dictionaries Introduction to Regular Expressions using "re" module. Exercises: 1. Demonstrate Tuples and Sets 2. Demonstrate Dictionaries		

Unit 2	Using NumPy	8 Hours

Basics of NumPy - Computation on NumPy- Aggregations -Computation onBasics of NumPy - Computation on NumPy- Aggregations -Computation on Arrays Comparisons, Masks and Boolean Arrays-Fancy Exercises: 1. Demonstrate Aggregation
2. Demonstrate Indexing and Sorting Indexing-Sorting Arrays- Structured Data: NumPy's Structured Array.

Unit 3	Data Manipulation with Pandas -I	6 Hours

Introduction to Pandas Objects-Data indexing and Selection-Operating on Data in Pandas, Handling Missing Data-Hierarchical Indexing - Combining Data Sets,Data Pre-processing, Data Reduction
Exercises:
1. Demonstrate handling of missing data
2. Demonstrate hierarchical indexing

Unit 4	Data Manipulation With Pandas -II	8 Hours

Aggregation and Grouping-Pivot Tables-Vectorized String Operations -Working with Time Series-High Performance Pandas- eval() and query()
Exercises:
1. Demonstrate usage of Pivot table
2. Demonstrate use of eval() and query()

Unit 5	Machine learning techniques with Sci-kit Learn libraries and Keras, Tensorflow Libraries	8 Hours

Supervised Learning techniques Regression, Linear Regression, Logistic Regression,Classification Trees, Support Vector Machines, Ensemble Methods: Random Forest, Decision trees, K Means Clustering, Artificial Neural Networks, Deep learning.K Nearest Neighbours
Unsupervised Learning techniques: K-meansClustering, Associative Rule Mining, Big data analysis
Application development using Keras and Tensorflow Libraries Exercises:
1. Students will be deploying various Machine learning based predictive models based on python programming

Unit 6	Visualization using MatPlotLib	6 Hours

Customizing Plot Legends, Colour Bars- Three Dimensional Plotting in Matplotlib.
Exercises:
1. Demonstrate Scatter Plot
2. Demonstrate 3Dplotting

Text Books:

1. Hands-On Machine Learning with Scikit-Learn and TensorFlow :Concepts, Tools, and Techniques to Build Intelligent Systems by Aurélien Géron,O'reilly publications.
2. Python for Data Analysis,First edition, by Wes McKinney

Reference Books:

1. Hands-on Deep Learning Algorithms with Python - SudharsanRavichandran

Title of the Course: Deep Learning **Course Code: UCSHN0651**	**L**	**T**	**P**	**Credits**
	3	-	-	3

Course Pre-Requisite: Linear Algebra, Probability and Information Theory, Numerical Computation

Course Description:

The purpose of this course is to provide the students with the advanced knowledge of Machine learning. It aims to enable the students to understand the design of various Deep Learning models and application.

Course Learning Objectives:

1. To introduce the idea of artificial neural networks and their architecture
2. To introduce techniques used for training artificial neural networks
3. Understanding the working of Convolutional Neural Networks and RNN in decision making.
4. Illustrate the strength and weaknesses of many popular deep learning approaches.

Course Outcomes:

COs	**After Completion of the course, the student should be able to**	**Bloom's Cognitive**	
		Level	**Descriptor**
CO1	Understand Core Concepts of Deep Learning	2	Understand
CO2	Illustrate idea of artificial neural networks, their architecture and applications	2	Illustrate
CO3	Explain convolution neural network, deep sequence model	2	Explain
CO4	Analyse different applications of deep learning	3	Analyse

CO-PO Mapping:

CO	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PSO01**	**PSO02**
CO1	3	3	2	2	2	1					3	1	1
CO2	3	3	2	3	2	1					3	1	1
CO3	2	3	2	3	2	1					3	1	1
CO4	2	2	2	3	3	1					3	2	3

Assessment:

Assessment	Marks
ESE	100

ESE: Assessment is based on 100% course content covered

Course Contents:

Unit 1: Introduction to Artificial Neural Network	**2 Hours**
Supervised Learning, unsupervised Learning, what is a Neural Network? The Human Brain, Models of a Neuron, Training dataset, validation dataset, test dataset, cross validation, bias, variance underfitting, overfitting, Hyperparameters, generalization error, regularization.	
Unit 2: Artificial Neural Network Architecture	**2 Hours**
Perceptron, Neural Network Architecture, NN with One Hidden Layer, NN with One Hidden Layer and Multiple Outputs, Neural Network Hyperparameter, forward propagation, loss functions, backward propagation, chain rule of differentiation, vanishing gradient problem, Gradient Descent.	
Unit 3: Deep Neural Network Architecture	**2 Hours**
Hyper-parameters in Deep Neural Networks, Activations functions: Sigmoid, Tanh, ReLu, Leaky ReLU, ELU, PReLU, Optimizers: Gradient Descent, SGD, Mini-batch SGD, SGD Momentum, Adagrad, RMSPROP, Adam Optimizers, Learning Rate, dropout.	
Unit 4: Convolution Neural Network	**2 Hours**
Motivation and Applications, Dense Layers to Convolutions, pooling layers, stride, zero padding, CNN Architectures (Alex Net, VGG, NiN, GoogLeNet, ResNet, DensNet), Application in Image segmentation, Automated Object Detection models.	
Unit 5: Deep Sequence Models	**2 Hours**
Sequence Modelling Problems, Motivation and Applications, Traditional Models: Recurrent Neural Networks, Modern Recurrent Neural Networks: Gated Recurrent Units, Long Short Term Memory (LSTM), Bidirectional LSTM	
Unit 6: Applications of Deep learning	**2 Hours**
Deep learning applications are making an impact across many different industries. Case studies like Fraud detection , Financial services, Natural language processing, Facial recognition ,Self-driving vehicles, Recommender systems, Health care.	

Textbooks:

1. Deep Learning , Author " Ian Goodfellow and Yoshua Bengio and Aaron Courville". Publisher MIT Press Edition 2017.
   https://www.deeplearningbook.org/lecture_slides.html
2. Simon Haykin, Neural networks and Learning Machines, Third Edition, Pearson, 2016.

Reference Books:

1. Machine Learning: An Algorithmic Perspective, Second Edition , Author Stephen Marsland Publisher Chapman and Hall/CRC
2. Aurélien Géron, "Hands-On Machine Learning with Scikit- Learn and TensorFlow", O'Reilly, 2017.
3. Nikhil Ketkar, "Deep Learning with Python: A Hands-on Introduction", Apress, 2017.

Course Code:	UCSH0701
Course Name:	Mini Project

L	T	P	Credit
		2	1

Course Prerequisites:

Project Based Learning

Course Description:

The students shall apply the course knowledge and project-based learning skills for solving real world problems. The students shall use the concepts they have learned during B.Tech program (III-VI) to develop a solution to the considered problem statement.

Course Objectives :

1.To occupy the knowledge of Project based learning 2.
To design and Implement Real world problem solutions
3. Work with State of art technologies
4. Apply Engineering Best Practices to develop the projects

Course Outcomes:

COs	Description	Bloom's Cognitive	
		Level	Descriptor
CO1	Identify the real-world problems to be solved applied computer science knowledge	2	Identify
CO2	Explain the proposed solution for problem by carrying survey and analysis.	4	Explain
CO3	Implement the proposed solution using state of art technologies	2	Implement
CO4	Build a detailed project report.	4	Build

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2
CO1	3	1			3	3	2	3	2	3	3	3	3
CO2	2	2	3		3	3	2	3	2	3	3	3	3
CO3	2	2	3		3	3	3	3	2	3	3	3	3
CO4	2	1	1		3	3	3	3	3	2	3	3	3

Assessment Scheme:

SN	Assessment	Marks	Remark
1	End Semester Examination (ESE)-POE	100	Assessment is based on 100% course content covered

Guidelines for Mini Project **7 Hours**

The primary objective of the mini project is to attain multi course project Based learning. Course Instructor shall form a team of 2-3 students. Each team shall apply the knowledge learned in previous semesters to identify the real-World problem and consider state of art technologies as part of the solution. The students shall be graded based on the skills demonstrated to identify the problem statement & design a proposed methodology. The students shall be graded based on the project implementation and submission of detailed project report which shall include the. technical aspects of the project. It is recommended to consider a common project report format and common evaluation process. Course instructors shall discuss the sample case studies to help them students understand the mini project deliverables.