MANIPAL
ACADEMY of HIGHER EDUCATION
(Institution of Eminence Deemed to be University)

Master of Engineering - ME (Artificial Intelligence and Machine Learning)

Syllabus

July 2023 Onwards

MANIPAL SCHOOL OF INFORMATION SCIENCES
MANIPAL ACADEMY OF HIGHER EDUCATION
MANIPAL – 576104, KARNATAKA, INDIA.

Program Educational Objectives / Outcomes (PEOs)

PEO 1: Acquire solid mathematical and computational skills essential for understanding, applying, and developing modern artificial intelligence and machine learning algorithms.

PEO 2: Produce industry-ready graduates with practical experience in structuring machine learning projects using state-of-the-art software.

PEO 3: Address multi-disciplinary challenges through coursework and projects by adapting to the rapidly advancing developments in artificial intelligence and machine learning approaches.

Program Objectives / Outcomes (POs)

PO 1: An ability to independently carry out research/investigation and development work to solve practical problems.

PO 2: An ability to write and present a substantial technical report/document.

PO 3: Demonstrate a degree of mastery over the area as per the specialization of the program which should be at a level higher than the requirements in an appropriate bachelor program.

PO 4: Identify appropriate and efficient algorithmic approaches for solving real-life challenges using artificial intelligence and machine learning principles and state-of-the art software prevalent in industry and academia.

PO 5: Develop teamwork and leadership skills for addressing challenges of social importance for sustainable societal development using ethical artificial intelligence and machine learning approaches.

Program Structure

ME (Artificial Intelligence and Machine Learning) – I Semester									
Course Code	Course Name	No. of Hrs./week				Duration of Exam in Hrs	Maximum Marks		
		Lecture	Tutorial	Practical	Credit		Internal 50	External 50	Total 100
BDA 5101	Algorithms and Data Structures for Big Data	3	-	-	3	3	50	50	100
AML 5101	Applied Linear Algebra	3	-	-	3	3	50	50	100
AML 5102	Applied Machine Learning	3	-	-	3	3	50	50	100
AML 5103	Applied Probability and Statistics	3	-	-	3	3	50	50	100
	Elective – I	3	-	-	3	3	50	50	100
BDA 5151	Algorithms and Data Structures for Big Data Lab	-	-	3	1	3	50	50	100
AML 5151	Applied Linear Algebra Lab	-	-	3	1	3	50	50	100
AML 5152	Applied Machine Learning Lab	-	-	3	1	3	50	50	100
AML 5153	Applied Probability and Statistics Lab	-	-	3	1	3	50	50	100
	Elective – I Lab	-	-	3	1	3	50	50	100
MPT 5100	Mini Project – I	-	-	-	4	-	10	-	100
PSD 5100	Professional Skill Development – I	-	-	-	1	-	100	-	100
Total		**15**	**-**	**15**	**25**				

ME (Artificial Intelligence and Machine Learning) – II Semester									
Course Code	Course Name	No. of Hrs. / week				Duration of Exam in Hrs	Maximum Marks		
		Lecture	Tutorial	Practical	Credit		Internal 50	External 50	Total 100
AML 5201	Advanced Applications of Probability and Statistics	3	-	-	3	3	50	50	100
AML 5202	Deep Learning	3	-	-	3	3	50	50	100
AML 5203	Machine Learning Principles and Applications	3	-	-	3	3	50	50	100
AML 5204	Reinforcement Learning	3	-	-	3	3	50	50	100
	Elective – II	3	-	-	3	3	50	50	100
AML 5251	Advanced Applications of Probability and Statistics	-	-	3	1	3	50	50	100
AML 5252	Deep Learning Lab	-	-	3	1	3	50	50	100
AML 5253	Machine Learning Principles and Applications Lab	-	-	3	1	3	50	50	100
AML 5254	Reinforcement Learning Lab	-	-	3	1	3	50	50	100
	Elective – II Lab	-	-	3	1	3	50	50	100
MPT 5200	Mini Project – II	-	-	-	4	-	100	-	100
PSD 5200	Professional Skill Development – II	-	-	-	1	-	100	-	100
TOTAL		**15**	**-**	**15**	**25**				

ME (Artificial Intelligence and Machine Learning) - III & IV Semesters									
AML 6098	Project Work	-	-	-	25				
Total Number of Credits to Award Degree							75		

List of Electives (Theory)

Elective - I		Elective - II	
Course Code	**Course Name**	**Course Code**	**Course Name**
AML 5131	Applications of Graph Theory	AML 5231	Applied Mathematics for Machine Learning
BDA 5102	Architecture of Big Data Systems	AML 5232	Convolutional Neural Networks for Computer Vision
BDA 5132	Principles of Data Visualization	AML 5233	Natural Language Processing Principles and Applications
		ENP 5230	Entrepreneurship

List of Electives (Lab)

Elective - I		Elective - II	
Course Code	**Course Name**	**Course Code**	**Course Name**
AML 5181	Applications of Graph Theory Lab	AML 5281	Applied Mathematics for Machine Learning Lab
BDA 5152	Architecture of Big Data Systems Lab	AML 5282	Convolutional Neural Networks for Computer Vision Lab
BDA 5182	Principles of Data Visualization Lab	AML 5283	Natural Language Processing Principles and Applications Lab
		ENP 5280	Entrepreneurship Lab

SEMESTER I

BDA 5101: Algorithms and Data Structures for Big Data	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Analyse recursive programs, solve a general class of recurrence relations.
2. Design programs for implementation of linked lists, stack, queues and binary search tree.
3. Design programs for dictionary, hash tables, graphs and shortest path techniques, sorting and searching.
4. Design string and text processing programs.

Unit	Topics	No. of Hours
I	**Algorithm specification and analysis techniques:** Analysis of recursive programs. Solving recurrence equations. General solution for a large class of recurrences.	3
II	Elementary data structures: Implementation of lists, stacks, queues, Trees	10
III	**Sorting & Searching, Hash Tables, Graph:** Quick sort, heap sort, merge sort. Linear search and binary search. Hashing and Dictionaries Representation of graphs. Depth First Searching. Breadth First Searching. Minimum cost spanning tree. Single source shortest paths and all-pairs shortest path	14
IV	**String and text processing techniques, Data stream algorithms:** Pattern-Matching Algorithms. Text Compression. Tries Sampling, Random Projections, Basic Algorithmic Techniques. Group Testing, Tree Method and Graph sketching.	9

References

1. Introduction to Algorithms - Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest. MIT Press.
2. Data Science Foundations: Data Structures and Algorithms Specialization (https://www.coursera.org/specializations/boulder-data-structures-algorithms)
3. Data Structures and Algorithms - Aho, Hopcroft and Ulmann. Pearson Publishers.

4. Data Structures and Algorithms in Python - Michael T. Goodrich, Roberto Tamassia, and Michael H. Goldwasser. John Wiley & Sons.
5. Data Streams: Algorithms and Applications - S. Muthukrishnan. Foundations and Trends in Theoretical Computer Science archive, Volume 1 Issue 2, August 2005, Pages 117 – 236

AML 5101: Applied Linear Algebra	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Develop a solid understanding of matrix-vector operations and relate them to real-life calculations.
2. Apply and analyse algorithms constructed using matrix-vector principles.
3. Develop models for real-life applications using the least squares technique and interpret the results from a practical perspective.

Unit	Topics	No. of Hours
I	**Vectors:** Conceptual introduction to vectors; vector addition; scalar-vector multiplication - Dot product; norm; distance - Standard deviation; standardization vs. normalization; angle between vectors - Application example: k-means clustering algorithm - Linear dependence/independence; basis - Orthonormal vectors; projections; Gram-Schmidt algorithm.	10
II	**Matrices:** Conceptual introduction to matrices; types of matrices (zero, identity, diagonal) - Addition of matrices; transpose; norm - Matrix-vector product – concept & examples - Systems of linear equations: over- & under-determined systems - Matrix-matrix product – concept & examples - QR factorization - Solving linear equations.	14
III	**Linear Least Squares:** Least squares: problem motivation and examples - Solving linear least squares problems - Least squares data fitting; validation; feature engineering - Least squares classification.	12

References
1. Introduction to Applied Linear Algebra, Vectors, Matrices, and Least Squares, Stephen Boyd & Lieven Vandenberghe, Cambridge University Press, 1st Edition, 2018. Available online at http://vmls-book.stanford.edu/vmls.pdf 2. Linear Algebra and Learning from Data, Gilbert Strang, Cambridge Uni. Press; 1st Edition, 2019. 3. Matrix Methods: Applied Linear Algebra, Richard Bronson and Gabriel B. Costa, Academic Press; 3rd Edition, 2008. 4. Matrix Methods in Data Mining and Pattern Recognition (Fundamentals of Algorithms), Lars Eldén – Society for Industrial and Applied Mathematics, 2007. 5. Introduction to computational thinking (https://computationalthinking.mit.edu/Spring21/)

AML 5102: Applied Machine Learning	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Differentiate between different types of machine learning paradigms and choose an appropriate one for a given application problem.
2. Apply different types of supervised and unsupervised machine learning algorithms to practical problems and assess their performance.
3. Understand the importance of feature engineering in machine learning applications.

Unit	Topics	No. of Hours
I	**Introduction to Machine Learning; Introduction to Supervised Learning; Decision Trees:** Overview of Supervised (regression and classification), unsupervised (clustering and dimensionality reduction), semi-supervised, and reinforcement learning with practical examples - Machine learning nomenclature: raw data, types of features and outputs, feature vector. Computing distances and similarities - Prototype based classification - K-nearest neighbours - Over- and under-fitting -Introduction to cross validation.	12

	Decision tree model of learning - Classification and regression using decision trees - Splitting criteria: entropy, information gain, Gini impurity - Building a decision tree.	
II	**Linear Models; Feature Selection; Introduction to Unsupervised Learning:** Linear model for regression and classification - Simple linear regression: model, estimation and interpretation of coefficients - Introduction to bias/variance trade-off - Regularized linear regression. Filter, wrapper, and embedded methods. Clustering vs. classification - Hierarchical clustering: dendrogram construction, types of linkage - Dimension reduction using principal component analysis (PCA)	12
III	**Probabilistic Models for Supervised Learning; Support Vector Machine; Ensemble Methods:** Probabilistic modelling of data using parameters - Introduction to maximum likelihood estimation (MLE) of parameters - Naive Bayes model for classification - Logistic regression for binary classification Classification using linear SVM - Dealing with nonlinearly separable data Bagging: classification using random forest - Boosting	12

References

1. Module: Introduction to Machine Learning (https://www.intel.com/content/www/us/en/developer/tools/oneapi/training/academic-program/educators/intro-machine-learning-training-kit.html)
2. Module: Get started with AI on Azure (https://learn.microsoft.com/en-us/training/modules/get-started-ai-fundamentals/)
3. Module: Microsoft Azure AI Fundamentals: Get started with artificial intelligence (https://learn.microsoft.com/en-us/training/paths/get-started-with-artificial-intelligence-on-azure/)
4. Learning path: Understand data science for machine learning (https://learn.microsoft.com/en-us/training/paths/understand-machine-learning/)
5. Grokking Machine Learning, Luis G. Serrano, Manning Publications; 1st Edition, 2019 - Online resource from Manning Publications available at https://www.manning.com/books/grokking-machine-learning
6. A Course in Machine Learning, Hal Daumé III – Online resource available at http://ciml.info/
7. An Introduction to Statistical Learning with Applications in R, Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer; 1st Edition, 2013, Corr. 7th printing 2017 Edition.
8. Mathematics for Machine Learning, Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong – Online resource from Cambridge University Press available at https://mml-book.github.io/book/mml-book.pdf.

AML 5103: Applied Probability and Statistics	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Model random phenomena using random variables.
2. Construct Bayesian models for quantifying uncertainty in practical problems.
3. Use sample information and perform hypothesis-test analysis using an appropriate statistical technique to explain attributes of a population.

Unit	Topics	No. of Hours
I	**Counting; Probability Concepts; Conditional Probability:** Multiplication rule; permutation; combination - Sampling: with/without replacement and order matters/does not matter - Binomial & multinomial coefficients - Distribution problems; Set theory; sample space; outcomes; events - Frequency based definition of probability - Equally likely vs. not equally likely outcomes - Axioms of probability Conditional probability; probability tree model; chain rule - Decomposition and the law of total probability - Bayes' rule - intuition, dependence/independence of events.	8
II	**Random Variables:** Modelling using discrete random variables: Bernoulli, geometric, binomial, negative binomial, hypergeometric, and Poisson distributions - Probability mass function and cumulative distribution function - Expectation and variance: discrete case - Modelling using continuous random variables: uniform, normal, log-normal, exponential, and beta distributions; probability density function - Expectation and variance: continuous case - Functions of random variables.	16
III	**Sampling and Parameter Estimation:** Population and sample - Statistic & sampling distribution - Sample mean and variance - Central limit theorem – intuition and applications Point estimation - Standard error - Interval estimation: interpretation of confidence interval - Hypothesis testing: p-values, significance level and their interpretations, application to analysis of one- /two-sample mean and paired data.	12

References

1. Digital Dice: Computational Solutions to Practical Probability Problems, Paul Nahin, Princeton University Press

2. Applied Data Science with R Specialization from IBM (Courses 1, 3, and 4, https://www.coursera.org/specializations/applied-data-science-r)
3. Introduction to Probability, Charles M. Grinstead, American Mathematical Society; 2nd Revised Edition 1997. Available online at https://chance.dartmouth.edu/teaching_aids/books_articles/probability_book/amsbook.mac.pdf
4. A First Course in Probability, Sheldon Ross, 9th Edition, Pearson Education India; 9th Edition, 2013
5. Statistics without Tears: An Introduction for Non-Mathematicians (Paperback), Derek Rowntree, Penguin UK
6. Biostatistics Open Learning textbook – Online resource from University of Florida available at https://bolt.mph.ufl.edu/6050-6052/
7. All of Statistics: A Concise Course in Statistical Inference, Larry Wasserman – Springer

ELECTIVES - SEMESTER I

AML 5131: Applications of Graph Theory	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Develop a thorough understanding of fundamental graph theoretic concepts and apply them to understanding practical problems.
2. Relate a real-life problem to an appropriate graph theoretic setup.
3. Describe how graph theory can be used for machine learning applications.

Unit	Topics	No. of Hours
I	**Graphs; Euler Tours and Hamilton Cycles:** Graphs and their representations - Incidence and adjacency matrices - Vertex degrees - Paths and connection - Cycles - Directed graphs - Subgraphs and supergraphs - The shortest path problem - Forests and trees, Cayley's formula.	14
II	**Flow in Networks; Matchings; Colouring Problems:** Flows and cuts - Max-flow min-cut theorem and its applications. Matchings and coverings in bipartite graphs - Perfect matchings - Applications of matchings. Edge colouring & Vertex colouring	14

III	**Random walks and Applications; Spectral Clustering and Applications**	8

References

1. Introduction to Graph Theory, Richard J. Trudeau, Dover Publications Inc.: 2nd Revised Edition, 1994.
2. Pearls in Graph Theory: A Comprehensive Introduction, Nora Hartsfield and Gerhard Ringel, Dover Publications, 2003.
3. Graph Theory, Adrian Bondy, M. Ram Murty, Springer Publications,1st Edition, 2008.

BDA 5102: Architecture of Big Data Systems	**L**	**T**	**P**	**C**	**Total hours**
	3	**0**	**0**	**3**	**36**

Course Outcome

1. Apply various techniques to examine different types of data and understand lambda architecture.
2. Apply different tools and frameworks of Hadoop eco-system.
3. Apply Spark engine to process real-time data.
4. Design applications to handle batch and streaming data using Hadoop and Spark tools.

Unit	**Topics**	**No. of Hours**
I	**Classifying Big Data Characteristics and Big Data processing - the Lambda architecture:** Analysis type, Processing methodology, Data Types, Data sources Different data storing and processing layers and architecture: Batch layer, Serving layer and Speed layer	12
II	**Batch layer, Serving layer and Speed layer:** Choosing a storage solution for the batch layer: Distributed file systems, Vertical partitioning. MapReduce: a paradigm for Big Data computing. Performance metrics for the serving layer. Speed layer.	9
III	**Spark: Alternatives to MapReduce:** Spark Architecture, Spark Session, DataFrame, Transformations and Actions, Spark SQL, Resilient Distributed Datasets (RDDs)	9

IV	**Stream Processing and Machine Learning using Spark:** Advantages and challenges of stream processing, Stream Processing Design Points, Streaming APIs, Structured Stream Processing High Level M-Lib concepts and M-Lib in Action.	6

References

1. Big Data: Principles and best practices of scalable real-time data systems - Nathan Marz and James Warren. Manning Publisher.
2. Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale – Tom White, O'Reilly Publication 4th Edition.
3. Spark: The Definitive Guide: Big Data Processing Made Simple – Bill Chambers, Matei Zaharia, O'Reilly Publication 1st Edition.
4. Explore core data concepts (https://learn.microsoft.com/en-us/training/modules/explore-core-data-concepts/)
5. Explore fundamentals of real-time analytics (https://learn.microsoft.com/en-us/training/modules/explore-fundamentals-stream-processing/)

BDA 5132: Principles of Data Visualization	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Implement web scrapping techniques to extract data from websites.
2. Organize raw data for analysis using data manipulation techniques.
3. Use Power BI for preparation and modelling of data for analysis.
4. Interpret data using various data visualization techniques.
5. Report data for analytics and to manage workspace using Power BI

Unit	Topics	No. of Hours
I	**Introduction to Web scraping:** Web scraping models and techniques - Case study BeautifulSoup, Scrapy, Selenium.	8

II	**Data Analysis: Data structures for analysis:** numpy, pandas. Data Wrangling - Clean, Transform, Merge, Reshape, Data Aggregation and Group Operations. Prepare data for analysis using Power BI, Model data in Power BI	12
III	**Data Visualization:** Visualization techniques: time series, statistical distributions, maps - Data visualization for web. Visualize data in Power BI, Data analysis in Power BI, Manage workspaces and datasets in Power BI, Create and use analytics reports with Power BI, Manage workspaces and datasets in Power BI.	16

References

1. Website Scraping with Python: Using BeautifulSoup and Scrapy, Gábor & Hajba, APRESS Publications, 1st Edition, 2018.
2. Web Scraping with Python: Collecting More Data from the Modern Web, Ryan Mitchell Shroff, O'Reilly, 2nd Edition, 2018.
3. Designing Data Visualizations, Julie Steele and Noah Iliinsky; O'Reilly Media; 1st Edition, 2011.
4. Python for Data Analysis, Wes McKinney; Shroff; O'Reilly; 2nd Edition, 2018.
5. https://learn.microsoft.com/en-us/certifications/exams/pl-300/

BDA 5151: Algorithms and Data structures for Big Data Lab	**L**	**T**	**P**	**C**	**Total hours**
	0	**0**	**3**	**1**	**36**

Course Outcome

1. Design programs for implementation of linked lists, stack, and queues.
2. Design programs for implementation of binary search tree, sorting and searching, dictionary and Hash Table
3. Design programs for graphs and shortest path techniques.

Unit	**Topics**	**No. of Hours**
I	Implement single linked and double linked list Implement array based and linked list-based Stack Implement array based and linked list-based Queue	15
II	Implement binary tree and BST Design and Implement Hash functions, implement open and closed Hash Tables	6

III	Implement Graph algorithms and shortest path	15

References

1. Introduction to Algorithms - Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest. MIT Press.
2. Data Science Foundations: Data Structures and Algorithms Specialization (https://www.coursera.org/specializations/boulder-data-structures-algorithms)
3. Data Structures and Algorithms - Aho, Hopcroft and Ulmann. Pearson Publishers.
4. Data Structures and Algorithms in Python - Michael T. Goodrich, Roberto Tamassia, and Michael H. Goldwasser. John Wiley & Sons.
5. Data Streams: Algorithms and Applications - S. Muthukrishnan. Foundations and Trends in Theoretical Computer Science archive, Volume 1 Issue 2, August 2005, Pages 117 – 236

AML 5151: Applied Linear Algebra Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Develop solid skills in using Python's legacy libraries for coding matrix-vector operations.
2. Implement algorithms constructed using matrix-vector principles.
3. Implement models for real-life applications using the least squares technique and interpret the results from a practical perspective.

Unit	Topics	No. of Hours
I	**Vectors:** Understand how to perform vector operations using Python; Visualize vectors and relate them to their geometric description; Implement the K-means algorithm from scratch using vector operations; Implement and interpret the output of the Gram-Schmidt algorithm.	10
II	**Matrices:** Understand how to perform matrix operations using Python; Implement and interpret matrix-vector operations using block-matrix operations; Understand how to solve	14

	linear systems of equations using Python; Code practical applications of QR factorization of matrices.	
III	**Linear Least Squares:** Solve linear least squares problems using Python and interpret the results; Implement and fine-tune feature extraction using least squares for practical problems; Implement least squares classification.	12

References

1. Introduction to Applied Linear Algebra, Vectors, Matrices, and Least Squares, Stephen Boyd & Lieven Vandenberghe, Cambridge University Press, 1st Edition, 2018. Available online at http://vmls-book.stanford.edu/vmls.pdf
2. Linear Algebra and Learning from Data, Gilbert Strang, Cambridge Uni. Press; 1st Edition, 2019.
3. Matrix Methods: Applied Linear Algebra, Richard Bronson and Gabriel B. Costa, Academic Press; 3rd Edition, 2008.
4. Matrix Methods in Data Mining and Pattern Recognition (Fundamentals of Algorithms), Lars Eldén – Society for Industrial and Applied Mathematics, 2007
5. Introduction to computational thinking (https://computationalthinking.mit.edu/Spring21/)

AML 5152: Applied Machine Learning Lab	**L**	**T**	**P**	**C**	**Total hours**
	0	**0**	**3**	**1**	**36**

Course Outcome

1. Differentiate between different types of machine learning paradigms and choose an appropriate one for a given application problem.
2. Apply different types of supervised and unsupervised machine learning algorithms to practical problems and assess their performance.
3. Understand the importance of feature engineering in machine learning applications.

Unit	Topics	No. of Hours
I	**Introduction to Machine Learning; Introduction to Supervised Learning; Decision Trees:** Program data, perform data wrangling, understand the data matrix, and differentiate between sample and feature; Investigate over- and underfitting concepts using the K-nearest neighbour algorithm; Implement and interpret results of cross-validation; Implement decision tree models in Python, fine-tune model parameters, and interpret results.	12
II	**Linear Models; Feature Selection; Introduction to Unsupervised Learning:** Implement linear models in Python and interpret model coefficients for practical problems; Implement and visualize bias-variance trade-off using linear regression as a basis; Compare, and contrast different feature engineering approaches for practical problems; Visualize the output of hierarchical clustering and PCA algorithms and interpret the results.	12
III	**Probabilistic Models for Supervised Learning; Support Vector Machine; Ensemble Methods:** Implement maximum likelihood estimation for a simple model; Analyse the performance of the Naive Bayes model for practical problems; Apply the SVM algorithm for linearly- and not-linearly separable data, compare the performance; Through coding, understand how ensemble methods in machine learning work.	12

References

1. Module: Introduction to Machine Learning (https://www.intel.com/content/www/us/en/developer/tools/oneapi/training/academic-program/educators/intro-machine-learning-training-kit.html)
2. Module: Get started with AI on Azure (https://learn.microsoft.com/en-us/training/modules/get-started-ai-fundamentals/)
3. Module: Microsoft Azure AI Fundamentals: Get started with artificial intelligence (https://learn.microsoft.com/en-us/training/paths/get-started-with-artificial-intelligence-on-azure/)
4. Learning path: Understand data science for machine learning (https://learn.microsoft.com/en-us/training/paths/understand-machine-learning/)
5. Grokking Machine Learning, Luis G. Serrano, Manning Publications; 1st Edition, 2019 - Online resource from Manning Publications available at https://www.manning.com/books/grokking-machine-learning
6. A Course in Machine Learning, Hal Daumé III – Online resource available at http://ciml.info/

7. An Introduction to Statistical Learning with Applications in R, Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer; 1st Edition, 2013, Corr. 7th printing 2017 Edition.
Mathematics for Machine Learning, Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong – Online resource from Cambridge University Press available at https://mml-book.github.io/book/mml-book.pdf.

AML 5153: Applied Probability and Statistics Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Visualize probability concepts through frequency-based interpretations.
2. Simulate discrete and continuous random variables for modelling random phenomena.
3. Design and apply hypothesis tests followed by interpretation of results.

Unit	Topics	No. of Hours
I	**Counting; Probability Concepts; Conditional Probability:** Understand the basic principles of the R programming language; Develop short code snippets to understand the basic principles of sampling and probability; Visualise and interpret probability concepts through a frequency-based approach; Program and analyse Bayesian models for practical problems.	8
II	**Random Variables:** Understand and apply R functions to simulate discrete and continuous random variables; Using sampling, compute and interpret different attributes of random variables; Visualise and interpret histograms and probability mass/density functions of random variables using state of the art visualisation libraries in R; Develop codes to model random phenomena using appropriate random variables.	16
III	**Sampling and Parameter Estimation:** Visualise sample data through histograms; Compute estimates of population parameters using samples and communicate the uncertainty in the estimates; Use R in-built functions for performing hypothesis tests; Interpret and communicate the results of hypothesis tests.	12

References

1. Digital Dice: Computational Solutions to Practical Probability Problems, Paul Nahin, Princeton University Press

2. Applied Data Science with R Specialization from IBM (Courses 1, 3, and 4, https://www.coursera.org/specializations/applied-data-science-r)
3. Introduction to Probability, Charles M. Grinstead, American Mathematical Society; 2nd Revised Edition 1997. Available online at https://chance.dartmouth.edu/teaching_aids/books_articles/probability_book/amsbook.mac.pdf
4. A First Course in Probability, Sheldon Ross, 9th Edition, Pearson Education India; 9th Edition, 2013
5. Statistics without Tears: An Introduction for Non-Mathematicians (Paperback), Derek Rowntree, Penguin UK
6. Biostatistics Open Learning textbook – Online resource from University of Florida available at https://bolt.mph.ufl.edu/6050-6052/
7. All of Statistics: A Concise Course in Statistical Inference, Larry Wasserman – Springer

AML 5181: Applications of Graph Theory Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Visualize graphs and graph models using Python.
2. Implement appropriate algorithms for solving graph theoretical problems.
3. Implement graph theoretic approaches for machine learning applications.

Unit	Topics	No. of Hours
I	**Graphs; Euler Tours and Hamilton Cycles:** Visualize graphs using Python; Compute structural properties of graphs such as connectivity using the adjacency matrix.	14
II	**Flow in Networks; Matchings; Colouring Problems:** Visualize network flow problems through applications; Implement practical applications of matching, edge- & vertex-colouring.	14
III	**Random walks and Applications; Spectral Clustering and Applications:** Create a random walk and analyse its properties; Model multidimensional data as similarity graph; Apply spectral clustering to practical problems.	8

References

1. Introduction to Graph Theory, Richard J. Trudeau, Dover Publications Inc.: 2nd Revised Edition, 1994.

2. Pearls in Graph Theory: A Comprehensive Introduction, Nora Hartsfield and Gerhard Ringel, Dover Publications, 2003.
3. Graph Theory, Adrian Bondy, M. Ram Murty, Springer Publications,1st Edition, 2008.

BDA 5152: Architecture of Big Data Systems Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Use data extraction tools to ingest various types of data into big data systems.
2. Experiment with different tools and frameworks of Hadoop eco-system.
3. Experiment with Spark Engine to process real-time data.
4. Design applications to handle batch and streaming data using Hadoop and Spark tools.

Unit	Topics	No. of Hours
I	Using HDFS commands to get familiarize with HDFS environment Transfer structured data to and from HDFS Creating Hive tables in HDFS and running SQL statements Developing and running MapReduce programs	15
II	Creating RDDs in spark Creating Spark DataFrames, loading, transforming and performing actions in Spark environment Handling streaming data using structured streaming	12
III	Developing applications using Hadoop tools Develop applications using Spark environment	9

References

1. Big Data: Principles and best practices of scalable real-time data systems - Nathan Marz and James Warren. Manning Publisher.
2. Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale – Tom White, O'Reilly Publication 4th Edition.
3. Spark: The Definitive Guide: Big Data Processing Made Simple – Bill Chambers, Matei Zaharia, O'Reilly Publication 1st Edition.

4. Explore core data concepts (https://learn.microsoft.com/en-us/training/modules/explore-core-data-concepts/)
5. Explore fundamentals of real-time analytics (https://learn.microsoft.com/en-us/training/modules/explore-fundamentals-stream-processing/)

BDA 5182: Principles of Data Visualization Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Experiment web scrapping techniques to extract data from websites.
2. Implement NumPy and Pandas for data science operations with examples.
3. Organize data for visualization using data manipulation techniques.
4. Experiment different visualization techniques.
5. Use power BI for analytics and to manage workspace.

Unit	Topics	No. of Hours
I	Design programs to dynamically extract data from web.	12
II	Understand and integrate various data structures for data analysis process. Create various techniques to clean and handle missing data. Design data filtering and transformation techniques.	9
III	Describe what is the purpose of Visualization and ways of classifying visualization. Create visualization for time series data, statistical distributions. Create visualization for maps, Hierarchical data and network data. Incorporate Power BI to plot visualization	15

References

1. Website Scraping with Python: Using BeautifulSoup and Scrapy, Gábor & Hajba, APRESS Publications, 1st Edition, 2018.
2. Web Scraping with Python: Collecting More Data from the Modern Web, Ryan Mitchell Shroff, O'Reilly, 2nd Edition, 2018.

3. Designing Data Visualizations, Julie Steele and Noah Iliinsky; O'Reilly Media; 1st Edition, 2011.
4. Python for Data Analysis, Wes McKinney; Shroff; O'Reilly; 2nd Edition, 2018.
5. https://learn.microsoft.com/en-us/certifications/exams/pl-300/

MPT 5100: Mini Project - I	L	T	P	C	Total hours
	0	0	0	4	48

Course Outcome

1. Identify the real-world and socially relevant problems and perform feasibility analysis for finding solutions.
2. Organize work effectively as a member in a team, examine, experiment, and communicate technical information constructively.
3. Develop and implement solutions to the identified problems by applying research methodology and development life cycle with appropriate documentation by incorporating ethical standards.

Unit	Topics	No. of Hours
I	Problem identification, literature survey, formation of detailed specifications.	48
II	Design and implementation of the proposed system architecture.	
III	Demonstrate an ability to present and defend project work carried out to a panel of experts.	

References

1. Research articles and Online Resources.

PSD 5100: Professional Skill Development - I	L	T	P	C	Total hours
	0	0	0	1	12

Course Outcome

1. Identify and synthesize important themes in the field of engineering which transform socio-economic ecosystem.
2. Develop competence to communicate effectively in oral and written forms.
3. Effective management of time, involve in reflective learning and adhere to the professional code of conduct.

Unit	Topics	No. of Hours
I	Report writing involves identifying the topic of interest from current issues in the domain of engineering and technology or inter disciplinary domains, then framing the order in the report, writing abstract, deciding on the content itself, conclusion and future scope of the topic and properly citing the references from bibliography.	12
II	Presenting in classroom to audience where content spoken, the conceptual knowledge and presentation skills (like audibility, eye contact, memory) of speaker is assessed.	

References

1. Research articles and Online Resources.

SEMESTER II

AML 5201: Advanced Applications of Probability and Statistics	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Apply linear and logistic regression models for practical problems and assess model performance.
2. Interpret the output of principal component analysis (PCA) applied to multivariate data for dimension reduction.
3. Identify multivariate data with mixed data type features and cluster using an appropriate technique.
4. Understand the basics of time series modelling and apply to real-life problems.

Unit	Topics	No. of Hours
I	**Multivariate Distributions:** Mean vector, covariance and correlation – population vs. sample - The multivariate Gaussian – joint-, marginal-, and conditional distributions, Mahalanobis distance and outliers - Properties of the multivariate Gaussian - Parameter estimation: maximum likelihood estimation (MLE) and maximum aposteriori estimation (MAP).	6
II	**Linear and Logistic Regression:** Simple linear regression – regression model, estimating and interpreting coefficients, accuracy of coefficient estimates and model, ANOVA, R2 statistic - Multiple linear regression – estimating coefficients, qualitative predictors, interaction effects, potential problems - Logistic regression – binary and multinomial logistic regression models, estimating and interpreting coefficients, assessing model calibration and discrimination, area under the ROC.	12
III	**Principal Component Analysis; Cluster Analysis:** Geometric intuition of principal components - Maximum variance perspective – algebraic setup, eigenvectors and eigenvalues of sample correlation matrix - Interpretation and application of principal components for dimension reduction.	10

	Dissimilarity measures for mixed data types - Partition around medoids (PAM) vs. K-means algorithms - Selecting the number of clusters.	
IV	**Bootstrapping; Time Series Analysis:** Time series concepts: stationarity, trend, seasonality, autocorrelation - Autoregressive moving average (ARMA) models - Resampling, smoothing, windowing, and rolling average - First and second order differencing - Validating time series predictions.	8

References

1. An Introduction to Statistical Learning with Applications in R, Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer; 1st Edition, 2013, Corr. 7th printing 2017 Edition.
2. Applied Multivariate Statistical Modeling from Swayam (NPTEL, https://onlinecourses.nptel.ac.in/noc23_ma26/preview)
3. An Introduction to Applied Multivariate Analysis with R, Brian Everitt and Torsten Hothorn- Springer Publications,1st Edition, 2011.
4. Machine Learning - A Probabilistic Perspective, Kevin P. Murphy, The MIT Press; 1st Edition, 2012.
5. Mathematics for Machine Learning, Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong, Cambridge University Press, 2020. – Online resource from Cambridge University Press available at https://mml-book.github.io/book/mml-book.pdf

AML 5202: Deep Learning	**L**	**T**	**P**	**C**	**Total hours**
	3	**0**	**0**	**3**	**36**

Course Outcome

1. Gain a solid understanding of the mathematical basis of neural networks.
2. Build and analyse deep learning models for application problems.
3. Devise techniques for improving the way neural networks learn.
4. Develop skills to choose an appropriate deep learning model.

Unit	**Topics**	**No. of Hours**
I	**Introduction to Deep Learning; Matrix Calculus; Logistic Regression:**	9

	Sigmoid neurons - The architecture of neural networks. Derivatives in one dimension - Derivative in multiple dimensions: gradient and Jacobian matrices - Rules of matrix calculus: product and chain rules - Optimizing using the gradient descent method– intuition and principle. Binary classification using logistic regression: cost function, gradient descent, and vectorization.	
II	**Shallow Neural Network:** One hidden layer neural network: architecture and notation - The role of activation functions and their derivatives - Forward propagation using matrix-based approach - Cost/loss function: intuition and setup - Gradient descent: backpropagation intuition and vectorized setup using matrix-based approach - Random initialization of network parameters.	10
III	**Deep Neural Network; Improving the Way neural Networks Learn:** Deep L-layer neural network: architecture, notation, and building blocks - Forward and backward propagation in a deep neural network using matrix-based approach - The importance of deep representations - Parameters vs. hyperparameters. The cross-entropy cost function - The learning slowdown problem - Overfitting and regularization: L1/L2, dropout - Weight initialization.	11
IV	**Hyperparameter Tuning; Recurrent Neural Networks; Advanced Deep Learning Models** Random initialization using appropriate scales - Batch normalization. Recurrent neural network: architecture and notation - Forward and backward propagation through time - Different types of recurrent neural networks and their applications. LSTM networks, transformers, autoencoders, and functional deep learning models.	6

References

1. Neural Networks and Deep Learning, Michael Nielsen – Determination Press – Available online at http://neuralnetworksanddeeplearning.com/index.html
2. Lecture slides of Prof. Andrew Ng – Stanford University – Available online at https://cs230.stanford.edu/syllabus/
3. Tensor Flow for Deep Learning Paperback, Reza Zadeh & Bharath Ramsundar, O'Reilly, 2018.

4. Deep Learning, Ian Goodfellow, Yoshua Bengio, and Aaron Courville – MIT Press – Available online at http://www.deeplearningbook.org/
5. Learning path: Microsoft Azure AI Fundamentals: Explore natural language processing (https://learn.microsoft.com/en-us/training/paths/explore-natural-language-processing/)
6. Essentials of SYCL (https://www.intel.com/content/www/us/en/developer/tools/oneapi/training/dpc-essentials.html#gs.uhlycs)

AML 5203: Machine Learning Principles and Applications	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Differentiate between discriminative and generative algorithms for supervised machine learning.
2. Evaluate machine learning algorithms for accuracy and performance.
3. Devise techniques for dealing with practical difficulties in applying machine learning techniques to real-life problems.
4. Develop low dimensional models of application problems with mixed data type features.

Unit	Topics	No. of Hours
I	**Kernel Methods; Linear Regression:** Kernels as feature maps - Kernel functions: types, hyperparameters - Kernel matrix: interpretation and properties - Kernel (nonlinear) SVM. Least mean squares (LMS) algorithm: cost function - Gradient descent algorithm: learning rate, batch and stochastic gradient approaches - Probabilistic interpretation of linear regression: MLE and MAP estimates.	10
II	**Generative Learning Algorithms; Regularization, Model Selection, & Evaluation:** Gaussian discriminant analysis (GDA) - Naive Bayes algorithm: MLE estimates, Laplace smoothing. Grid search for best hyperparameters - Cross validation: types and practical approaches – Feature selection: forward/backward search, wrapper model & filter feature selection - Metrics for evaluating supervised & unsupervised machine learning algorithms.	11

III	**Imbalanced Data; Expectation Maximization; Dimension Reduction; Independent Component Analysis:** Modifying the training data: over- and under-sampling - Modifying the loss function. Clustering with a mixture of Gaussians - Expectation maximization (EM) framework. Factor analysis (FA) - Generalized low rank models (GLRM). Independent Component Analysis (ICA)	15

References

1. Learning path: Microsoft Azure AI Fundamentals: Explore visual tools for machine learning (https://learn.microsoft.com/en-us/training/paths/create-no-code-predictive-models-azure-machine-learning/)
2. Learning path: Microsoft Azure AI Fundamentals: Explore decision support (Learning path: Microsoft Azure AI Fundamentals: Explore decision support (https://learn.microsoft.com/en-us/training/paths/explore-fundamentals-of-decision-support/)
3. Learning paths: Microsoft Azure AI Fundamentals: Explore knowledge mining (https://learn.microsoft.com/en-us/training/paths/explore-fundamentals-of-knowledge-mining/)
4. A Course in Machine Learning, Hal Daumé III – Online resource available at http://ciml.info/
5. Machine Learning: A Probabilistic Perspective, Kevin Murphy, MIT Press,2017.
6. An Introduction to Statistical Learning with Applications in R, Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer Publications, 2017.
7. Lecture slides of Prof. Andrew Ng – Stanford University – Available online at http://cs229.stanford.edu/syllabus.html
8. Mathematics for Machine Learning, Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong – Online resource from Cambridge University Press available at https://mml-book.github.io/book/mml-book.pdf
9. Pattern Recognition and Machine Learning, Christopher Bishop, Springer Publications, 2017.

AML 5204: Reinforcement Learning	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Define the key features of reinforcement learning that distinguishes it from AI and non-interactive machine learning.
2. Understand how ideas such as temporal difference learning and dynamic programming fit in the framework of learning from interaction to achieve goals.
3. Decide if an application problem can be formulated as a reinforcement learning problem and choose an appropriate algorithm.
4. Understand and implement commonly used reinforcement learning algorithms.

Unit	Topics	No. of Hours
I	**Introduction to the Reinforcement Learning Problem; Reinforcement Learning Framework; Dynamic Programming:** Examples and elements of reinforcement learning - Limitations and scope of reinforcement learning - History of reinforcement learning. n-Armed bandit problem: action-value methods - Finite Markov decision process: the agent–environment interface, goals and rewards, returns, Markov decision processes, value functions, and optimal value functions.	15
II	**Model Free Reinforcement Learning:** Generalized policy iteration - Importance of exploration - Monte Carlo control - Temporal difference methods for control.	11
III	**Approximate Solution Methods; Policy Based Methods:** Value prediction with function approximation - Gradient-descent methods - Linear methods - Control with function approximation. Policy gradient - Actor–critic methods - Policy-based vs. value-based methods - Integrating supervised & reinforcement learning.	10

References
1. Deep Reinforcement Learning Hands-On, Maxim Lapan, 2nd Edition, Packt, 2020. 2. Reinforcement Learning: An Introduction, Richard S. Sutton and Andrew G. Barto, MIT Press, 2nd Edition – Available online at https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf 3. Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play, David Foster – O'Reilly, 1st Edition, 2019.

ELECTIVES - SEMESTER II

AML 5231: Applied Mathematics for Machine Learning	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Develop a solid understanding of fundamentals of matrix decomposition techniques and apply them to practical problems.
2. Describe the role of derivatives in machine learning and understand different methods for computing them.
3. Acquire solid foundation in understanding the principles behind state-of-the-art optimization algorithms used in machine learning libraries.

Unit	Topics	No. of Hours
I	**Matrix Decompositions and Applications:** Matrix and tensor products - Determinant and trace - Eigen-decomposition and diagonalization - Cholesky decomposition - Singular value decomposition - Nonnegative matrix factorization.	14
II	**Computing Derivatives:** Differentiability - Symbolic differentiation - Finite differences - Automatic differentiation.	8

III	**Continuous Optimization:** Optimization using gradient descent - Constrained optimization and Lagrange multipliers - Convex optimization – Sub gradients - Stochastic gradient descent - Momentum methods.	14

References

1. Mathematics for Machine Learning by Marc Peter Deisenroth, Aldo Faisal, and Cheng Soon Ong – Online resource from Cambridge University Press available at https://mml-book.github.io/book/mml-book.pdf
2. Matrix Computations, Gene H. Golub and Charles F. Van Loan, Hindustan Book Agency; 4th Edition, 2015.
3. Deep Learning, Ian Goodfellow, Yoshua Bengio, and Aaron Courville, MIT Press,2017. – Available online at http://www.deeplearningbook.org/
4. Understanding Machine Learning: From Theory to Algorithms (UML), Shai Shalev-Shwartz and Shai Ben-David, Cambridge University Press, 1st Edition, 2014.

AML 5232: Convolutional Neural Networks for Computer Vision	**L**	**T**	**P**	**C**	**Total hours**
	3	**0**	**0**	**3**	**36**

Course Outcome

1. Identify and apply appropriate image processing operation for a problem-solving task.
2. Analyze a real-life problem involving computer vision and solve it using CNNs.
3. Develop and design CNN architectures using the basic building blocks of CNN
4. Investigate use of an existing CNN architecture for an application problem.

Unit	**Topics**	**No. of Hours**
I	**Introduction to Computer Vision; Features; Neural Networks Basics:** Computer vision overview - Historical context and applications - Image processing vs. computer vision Role of convolution in image processing, neighbourhood operations for image enhancement, edge detection, colour models – RGB, HSI, CIE LAB, YUV.	6

II	**Convolutional Neural Networks (CNN):** Network layers: pre-processing, convolutional layers, pooling layers, nonlinearity, fully connected layers, region of interest pooling - Loss functions: hinge loss, squared hinge loss, cross-entropy loss, Euclidean loss, L1 error.	10
III	**CNN Training:** Weight initialization – Regularization - Tackling overfit and underfit situations, Activation functions, Gradient based learning: batch-, stochastic-, and mini-batch gradient descent, understanding loss and accuracy curves, skip connections.	10
IV	**CNN Architectures; Applications of CNNs in Computer Vision:** Image classification, Segmentation and Object detection and localization.	10

References

1. Learning path: Microsoft Azure AI Fundamentals: Explore computer vision (https://learn.microsoft.com/en-in/training/paths/explore-computer-vision-microsoft-azure/)
2. A Guide to Convolutional Neural Networks for Computer Vision, Salman Khan, Hossein Rahmani, Syed Afaq Ali Shah, and Mohammed Bennamoun, Morgan & Claypool Publishers, 2018.
3. Lecture slides of Prof. Fei-Fei Li – Stanford University – Available online at http://cs231n.stanford.edu/
4. Neural Networks and Deep Learning, Michael Nielsen, Determination Press – Available online at http://neuralnetworksanddeeplearning.com/index.html
5. Computer Vision: Algorithms and Applications, Richard Szeliski, Springer, 2011 – Online resource from Springer available at http://szeliski.org/Book/
6. Lecture slides of Prof. Fei-Fei Li – Stanford University – Available online at http://cs231n.stanford.edu/

AML 5233: Natural Language Processing Principles and Applications	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Develop an in-depth understanding of both algorithms for processing linguistic information and the underlying computational properties of natural languages.
2. Analyse word-level, syntactic, and semantic processing from both a linguistic and an algorithmic perspective.
3. Formulate deep learning approaches for natural language processing tasks.

Unit	Topics	No. of Hours
I	**Introduction to Natural Language Processing (NLP); Regular Expressions; N-gram Language Models:** Terminology - Probability and NLP. Introduction to regular expressions - Information extraction using regular expressions. Probabilistic language model - Chain rule and Markov assumption - Evaluating language models – Smoothing.	14
II	**Naive Bayes and Sentiment Classification; Vector Semantics and Embeddings:** Vector semantics - Words and vectors - Cosine for measuring similarity - TF-IDF vector model - Word2Vec & GloVe models - Visualizing embeddings.	8
III	**NLP with Deep Learning; Applications of Natural Language Processing:** Neural language models - Introduction to PyTorch -Sequence processing with recurrent neural networks.	14

References
1. Speech and Language Processing, Dan Jurafsky and James H. Martin, Pearson; 3rd Edition (draft) – Available online at https://web.stanford.edu/~jurafsky/slp3/ 2. Natural Language Processing with Python. – Analysing Text with the Natural Language Toolkit, Steven Bird, Ewan Klein, and Edward Loper, ISTE Ltd., 1st Edition, 2017 - Available online at https://www.nltk.org/book/ 3. A Primer on Neural Network Models for Natural Language Processing, Yoav Goldberg – Available online at http://faculty.cse.tamu.edu/huangrh/Spring18/nnlp.pdf 4. Natural Language Processing with PyTorch, Delip Rao & Brian McMahan, O'Reilly, 1st Edition, 2019.

ENP 5230: Entrepreneurship	L	T	P	C	Total hours
	3	0	0	3	36

Course Outcome

1. Explain the importance of entrepreneurship and entrepreneurial development model, social responsibilities of business.
2. Describe Entrepreneurial Traits and Factors affecting Entrepreneurship process.
3. Discuss Business Start-up Process.
4. Summarize a business and marketing plan for entrepreneurs.

Unit	Topics	No. of Hours
I	**Introduction to Entrepreneurship:** Meaning and Definition of Entrepreneurship-Employment vs Entrepreneurship, Theories of Entrepreneurship, approach to entrepreneurship, Entrepreneurs VS Manager	6
II	**Entrepreneurial Traits:** Personality of an entrepreneur, Types of Entrepreneurs	5

III	**Process of Entrepreneurship:** Factors affecting Entrepreneurship process	6
IV	**Business Start-up Process:** Idea Generation, Scanning the Environment, Macro and Micro analysis	7
V	**Business Plan writing:** Points to be considered, Model Business plan	6
VI	**Case studies:** Indian and International Entrepreneurship	6
References		
1. NVR Naidu and T. Krishna Rao, "Management and Entrepreneurship", IK International Publishing House Pvt. Ltd 2008. 2. Mohanthy Sangram Keshari, "Fundamentals of Entrepreneurship", PHI Publications, 2005		

AML 5251: Advanced Applications of Probability and Statistics Lab	**L**	**T**	**P**	**C**	**Total hours**
	0	**0**	**3**	**1**	**36**

Course Outcome

1. Build and assess linear and logistic regression models for practical problems.
2. Perform principal component analysis (PCA) for dimension reduction in multivariate data.
3. Cluster multivariate data with mixed data types.
4. Apply time series modelling to real-life problems.

Unit	**Topics**	**No. of Hours**
I	**Multivariate Distributions:** Compute descriptive statistics of multivariate data; Perform exploratory data analysis of multivariate data; Identify outliers in multivariate data; Visualise and understand the properties of multivariate Gaussian data.	6
II	**Linear and Logistic Regression:** Use in-built functions in R to build linear models for practical problem; Compute different performance metrics to assess model performance; Interpret model coefficients and investigate the effect of input variables on output through sensitivity	12

	analysis; Use in-built functions in R to build logistic regression models for practical binary classification problems and assess model performance.	
III	**Principal Component Analysis; Cluster Analysis:** Visualise the geometric interpretation of principal component analysis (PCA); Use in-built functions in R to perform PCA on multivariate data; Compare and contrast PCA for variance maximization vs. clustering of multivariate data; Cluster multivariate data with mixed data types using in-built functions in R.	10
IV	**Bootstrapping; Time Series Analysis:** Apply bootstrapping on a practical data set and assess performance; Understand and apply in-built functions in R for time series modelling; Apply time series modelling to practical problems; Interpret the results of times series model predictions.	8

References

1. An Introduction to Statistical Learning with Applications in R, Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer; 1st Edition, 2013, Corr. 7th printing 2017 Edition.
2. Applied Multivariate Statistical Modeling from Swayam (NPTEL, https://onlinecourses.nptel.ac.in/noc23_ma26/preview)
3. An Introduction to Applied Multivariate Analysis with R, Brian Everitt and Torsten Hothorn– Springer Publications,1st Edition, 2011.
4. Machine Learning - A Probabilistic Perspective, Kevin P. Murphy, The MIT Press; 1st Edition, 2012.
5. Mathematics for Machine Learning, Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong, Cambridge University Press, 2020. – Online resource from Cambridge University Press available at https://mml-book.github.io/book/mml-book.pdf

AML 5252: Deep Learning Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Develop practical experience with state-of-the-art deep learning tools and libraries.
2. Implement deep learning models for application problems.
3. Implement techniques for improving the way neural networks learn.

4. Computationally analyse deep learning models and select the best model.

Unit	Topics	No. of Hours
I	**Introduction to Deep Learning; Matrix Calculus; Logistic Regression:** Implement a sigmoid neuron from scratch; Implement forward and backward propagation for a sigmoid neuron; Implement gradient descent for a sigmoid neuron; Implement cost function for binary classification using logistic regression using vectorized approach.	9
II	**Shallow Neural Network:** Visualize different nonlinear activation functions; Implement forward and backward propagation for a shallow neural network using matrix-based approach; Implement gradient descent method for a shallow neural network; Numerically investigate the effect of random initialization of network parameters.	10
III	**Deep Neural Network; Improving the Way neural Networks Learn:** Visualise architecture of a deep neural network; Implement forward and backward propagation for a deep neural network using matrix-based approach; Implement deep neural networks using in-built libraries for real-life problems; Implement different regularization approaches and compare their advantages and disadvantages.	11
IV	**Hyperparameter Tuning; Recurrent Neural Networks:** Fine tune hyperparameters; Numerically investigate the effect of random initialization in deep neural networks; Visualise the architecture of a recurrent neural network; Implement recurrent neural network models for real-life problems.	6

References

1. Neural Networks and Deep Learning, Michael Nielsen – Determination Press – Available online at http://neuralnetworksanddeeplearning.com/index.html
2. Lecture slides of Prof. Andrew Ng – Stanford University – Available online at https://cs230.stanford.edu/syllabus/
3. Tensor Flow for Deep Learning Paperback, Reza Zadeh & Bharath Ramsundar, O'Reilly, 2018.
4. Deep Learning, Ian Goodfellow, Yoshua Bengio, and Aaron Courville – MIT Press – Available online at http://www.deeplearningbook.org/

5. Learning path: Microsoft Azure AI Fundamentals: Explore natural language processing (https://learn.microsoft.com/en-us/training/paths/explore-natural-language-processing/)
6. Essentials of SYCL (https://www.intel.com/content/www/us/en/developer/tools/oneapi/training/dpc-essentials.html#gs.uhlycs)

AML 5253: Machine Learning Principles and Applications Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Practically apply state of the art machine learning tools and libraries.
2. Evaluate machine learning algorithms for accuracy and performance for practical problems.
3. Implement different strategies for selecting features and dealing with missing data.
4. Implement machine learning models for real-life data with mixed datatype features.

Unit	Topics	No. of Hours
I	**Kernel Methods; Linear Regression:** Implement and compare different kernels for feature mapping; Implement kernel SVM and investigate the effects of model parameters through visualization; Implement gradient descent for least mean squares algorithm and investigate the effects of hyperparameters on performance; Compare linear regression applied to practical problems with and without regularization.	10
II	**Generative Learning Algorithms; Regularization, Model Selection, & Evaluation:** Implement probabilistic models of data; Perform grid search to identify best model hyperparameters; Perform feature engineering for real-life problems; Evaluate machine learning algorithms using well established performance metrics.	11
III	**Imbalanced Data; Expectation Maximization; Dimension Reduction; Independent Component Analysis:**	15

	Implement and compare different approaches for dealing with missing data in real-life problems; Implement and interpret low dimensional models of data; Implement models for analysing mixed datatype data; Compare and contrast different techniques for dimension reduction and their practical implications.	

References

1. Learning path: Microsoft Azure AI Fundamentals: Explore visual tools for machine learning (https://learn.microsoft.com/en-us/training/paths/create-no-code-predictive-models-azure-machine-learning/)
2. Learning path: Microsoft Azure AI Fundamentals: Explore decision support (Learning path: Microsoft Azure AI Fundamentals: Explore decision support (https://learn.microsoft.com/en-us/training/paths/explore-fundamentals-of-decision-support/)
3. Learning paths: Microsoft Azure AI Fundamentals: Explore knowledge mining (https://learn.microsoft.com/en-us/training/paths/explore-fundamentals-of-knowledge-mining/)
4. A Course in Machine Learning, Hal Daumé III – Online resource available at http://ciml.info/
5. Machine Learning: A Probabilistic Perspective, Kevin Murphy, MIT Press,2017.
6. An Introduction to Statistical Learning with Applications in R, Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer Publications, 2017.
7. Lecture slides of Prof. Andrew Ng – Stanford University – Available online at http://cs229.stanford.edu/syllabus.html
8. Mathematics for Machine Learning, Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong – Online resource from Cambridge University Press available at https://mml-book.github.io/book/mml-book.pdf
9. Pattern Recognition and Machine Learning, Christopher Bishop, Springer Publications, 2017.

AML 5254: Reinforcement Learning Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Understand the trade-off between exploration vs. exploitation approaches in solving reinforcement learning tasks.

2. Use dynamic programming approach to solve reinforcement learning tasks.
3. Model real-life problems using Markov decision processes.
4. Compare and contrast several methods for solving reinforcement learning tasks.

Unit	Topics	No. of Hours
I	**Introduction to the Reinforcement Learning Problem; Reinforcement Learning Framework; Dynamic Programming:** Implement building blocks for solving a reinforcement learning task; Solve an n-Armed bandit problem using different exploration strategies; Implement Markov decision process models.	15
II	**Model Free Reinforcement Learning:** Implement iterative policy evaluation; Implement Monte Carlo methods for solving reinforcement learning tasks; Implement temporal difference methods for solving reinforcement learning tasks.	11
III	**Approximate Solution Methods; Policy Based Methods:** Implement function approximation methods for value prediction; Implement linear value function approximators; Explore integration of supervised and reinforcement learning.	10

References

1. Deep Reinforcement Learning Hands-On, Maxim Lapan, 2nd Edition, Packt, 2020.
2. Reinforcement Learning: An Introduction, Richard S. Sutton and Andrew G. Barto, MIT Press, 2nd Edition – Available online at https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf
3. Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play, David Foster – O'Reilly, 1st Edition, 2019.

AML 5281: Applied Mathematics for Machine Learning Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Assess applicability of matrix decomposition techniques for practical problems.

2. Implement and compare different methods for computing derivatives.
3. Understand the implementations of state-of-the-art optimization algorithms used in machine learning libraries.

Unit	Topics	No. of Hours
I	**Matrix Decompositions and Applications:** Implement matrix decompositions using block matrix representations; Implement and compare exact and approximate decompositions; Implement codes to understand the optimization-centric view to matrix factorization; Interpret the factors arising out of matrix factorizations for real-life problems.	14
II	**Computing Derivatives:** Visualize differentiability concepts in 3D; Implement symbolic differentiation for computing derivatives exactly; Implement finite difference methods for approximating derivatives; Implement automatic differentiation and compare it with other approaches.	8
III	**Continuous Optimization:** Solve continuous optimization problems using state of the art libraries; Visualize constrained optimization problems and solutions in 3D; Implement and visualize solutions of gradient descent method and its extensions for continuous optimization; Understand implementations of continuous optimization algorithms used in state-of-the-art libraries.	14

References

1. Mathematics for Machine Learning by Marc Peter Deisenroth, Aldo Faisal, and Cheng Soon Ong – Online resource from Cambridge University Press available at https://mml-book.github.io/book/mml-book.pdf
2. Matrix Computations, Gene H. Golub and Charles F. Van Loan, Hindustan Book Agency; 4th Edition, 2015.
3. Deep Learning, Ian Goodfellow, Yoshua Bengio, and Aaron Courville, MIT Press,2017. – Available online at http://www.deeplearningbook.org/
4. Understanding Machine Learning: From Theory to Algorithms (UML), Shai Shalev-Shwartz and Shai Ben-David, Cambridge University Press, 1st Edition, 2014.

AML 5282: Convolutional Neural Networks for Computer Vision Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Design computer vision techniques for problem solving.
2. Develop and design CNN architectures using the basic building blocks of CNN.
3. Design and develop a CNN architecture for an application problem.

Unit	Topics	No. of Hours
I	**Introduction to Computer Vision; Features; Neural Networks Basics:** Implement basic computer vision and image processing tasks; Implement feature extraction using convolution from images; Implement and understand various colour model representations.	6
II	**Convolutional Neural Networks (CNN):** Visualize and understand the building blocks of a CNN; Implement different types of CNN layers and understand their utility. Implement different nonlinear activation functions, compare and contrast them; Implement and understand the role of different types of loss functions used in a CNN.	10
III	**CNN Learning; Visualizing and Understanding CNNs:** Implement random initialization of weights in a CNN and compare it with a non-random initialization; Implement regularization to prevent overfitting in CNNs; Implement different gradient-based approaches for optimization; Implement efficient gradient computations in CNNs.	10
IV	**CNN Architectures; Applications of CNNs in Computer Vision:** Explore the building blocks of state-of-the-art CNN architectures; Explore applications of CNNs to real life problems.	10

References

1. Learning path: Microsoft Azure AI Fundamentals: Explore computer vision (https://learn.microsoft.com/en-in/training/paths/explore-computer-vision-microsoft-azure/)
2. A Guide to Convolutional Neural Networks for Computer Vision, Salman Khan, Hossein Rahmani, Syed Afaq Ali Shah, and Mohammed Bennamoun, Morgan & Claypool Publishers, 2018.
3. Lecture slides of Prof. Fei-Fei Li – Stanford University – Available online at http://cs231n.stanford.edu/

4. Neural Networks and Deep Learning, Michael Nielsen, Determination Press – Available online at http://neuralnetworksanddeeplearning.com/index.html
5. Computer Vision: Algorithms and Applications, Richard Szeliski, Springer, 2011 – Online resource from Springer available at http://szeliski.org/Book/

AML 5283: Natural Language Processing Principles and Applications Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Develop an in-depth understanding of both algorithms for processing linguistic information and the underlying computational properties of natural languages.
2. Formulate deep learning approaches for natural language processing tasks.
3. Develop practical experience with state-of-the-art natural language processing tools and libraries.

Unit	Topics	No. of Hours
I	**Introduction to Natural Language Processing (NLP); Regular Expressions; N-gram Language Models:** Understand the basics of NLP and role of probability in it; Understand how to use and apply regular expressions; Develop the idea of a probabilistic language model; Understand how to evaluate and compare language models.	14
II	**Naive Bayes and Sentiment Classification; Vector Semantics and Embeddings:** Implement sentiment classification using real-life datasets; Implement building blocks for vector representation of words; Implement and compare vector models for words; Visualize word embeddings.	8
III	**NLP with Deep Learning; Applications of Natural Language Processing:** Implement neural models for NLP applications; Gain experience in using PyTorch; Implement recurrent neural network models for NLP applications; Explore practical applications of NLP.	14

References
1. Speech and Language Processing, Dan Jurafsky and James H. Martin, Pearson; 3rd Edition (draft) – Available online at https://web.stanford.edu/~jurafsky/slp3/ 2. Natural Language Processing with Python. – Analysing Text with the Natural Language Toolkit, Steven Bird, Ewan Klein, and Edward Loper, ISTE Ltd., 1st Edition, 2017 - Available online at https://www.nltk.org/book/ 3. A Primer on Neural Network Models for Natural Language Processing, Yoav Goldberg – Available online at http://faculty.cse.tamu.edu/huangrh/Spring18/nnlp.pdf 4. Natural Language Processing with PyTorch, Delip Rao & Brian McMahan, O'Reilly, 1st Edition, 2019.

ENP 5280: Entrepreneurship Lab	L	T	P	C	Total hours
	0	0	3	1	36

Course Outcome

1. Study of prominence of entrepreneurship.
2. Develop use cases for building a business.
3. Evaluation of factors influencing business venture.

Unit	Topics	No. of Hours
I	Study of use cases for need and prominence of entrepreneurship, associated decision making process.	6
II	Study of report by the National Knowledge Commission on the importance of entrepreneurship in economic development.	6
III	Develop use cases for identifying and evaluating opportunities, developing business plan, assessment of resources, project appraisal and feasibility plan.	9
IV	Creating and starting venture includes legal requirements, marketing strategies, financial plans and human resources management.	9
V	Design a Case studies of Indian and International Entrepreneurship.	6

References
1. Management and Entrepreneurship, NVR Naidu and T. Krishna Rao, IK International Publishing House Pvt.Ltd, 2008.
2. Fundamentals of Entrepreneurship, Mohanthy Sangram Keshari, PHI Learning Pvt. Ltd., 2005.

MPT 5200: Mini Project - II	L	T	P	C	Total hours
	0	0	0	4	48

Course Outcome

1. Identify the real-world and socially relevant problems and perform feasibility analysis for finding solutions.
2. Organize work effectively as a member in a team, examine, experiment, and communicate technical information constructively.
3. Develop and implement solutions to the identified problems by applying research methodology and development life cycle with appropriate documentation by incorporating ethical standards.

Unit	Topics	No. of Hours
I	Problem identification, literature survey, formation of detailed specifications.	48
II	Design and implementation of the proposed system architecture.	
III	Demonstrate an ability to present and defend project work carried out to a panel of experts.	

References

1. Research articles and Online Resources.

PSD 5200: Professional Skill Development - II	L	T	P	C	Total hours
	0	0	0	1	12

Course Outcome

1. Develop the skills needed for approaching technical and HR interviews.
2. Use mathematical, reasoning, and domain specific skills to solve objective questionnaires in time.
3. Demonstrate depth of knowledge in the chosen field of study.

Unit	Topics	No. of Hours
I	Peer interviews, mock interviews.	12
II	Logical reasoning, mathematical aptitude, domain specific problem-solving skills.	
III	Conduction of domain specific knowledge test.	

References

1. R S Aggarwal. Quantitative Aptitude for Competitive Examinations. S Chand, 2017.
2. McDowell, Gayle Laakmann. Cracking the coding interview: 189 programming questions and solutions. CareerCup, LLC, 2015.
3. Domain specific tools and online resources.

AML 6098: Project Work	L	T	P	C	Total hours
	0	0	0	25	300

Course Outcome

1. Undertake innovative industry/research-oriented projects and perform feasibility analysis for finding solutions.
2. Implement and test the proposed design using appropriate framework, programming language and tools.
3. Demonstrate an ability to present and defend project work carried out to a panel of experts.

Unit	Topics	No. of Hours
I	Problem identification, literature survey, formation of detailed requirement specification document.	300
II	Design and implementation of the proposed modules with specific test cases.	
III	Detailed report of the work carried out, present, and defend the project work carried out to a panel of experts.	
References		
1. Research articles, domain specific tools and online resources.		

Program Outcome and Course Outcome Mapping

Sl. No.	Course Code	Course Name	Credits	PO1	PO2	PO3	PO4	PO5
1	BDA 5101	Algorithms and Data Structures for Big Data	3			*	*	
2	AML 5101	Applied Linear Algebra	3	*			*	
3	AML 5102	Applied Machine Learning	3	*			*	*
4	AML 5103	Applied Probability and Statistics	3			*		*
5	AML 5131	Applications of Graph Theory	3	*		*	*	
	BDA 5102	Architecture of Big Data Systems	3			*	*	*
	BDA 5132	Principles of Data Visualization	3			*		*
6	BDA 5151	Algorithms and Data Structures for Big Data Lab	1			*	*	
7	AML 5151	Applied Linear Algebra Lab	1	*			*	
8	AML 5152	Applied Machine Learning Lab	1	*			*	*
9	AML 5153	Applied Probability and Statistics Lab	1			*		*
10	AML 5181	Applications of Graph Theory Lab	1	*		*	*	
	BDA 5152	Architecture of Big Data Systems Lab	1			*	*	*
	BDA 5182	Principles of Data Visualization Lab	1			*		*
11	MPT 5100	Mini Project – I	4	*	*	*	*	*
12	PSD 5100	Professional Skill Development – I	1	*	*			
13	AML 5201	Advanced Applications of Probability and Statistics	3			*		*
14	AML 5202	Deep Learning	3	*			*	*

15	AML 5203	Machine Learning Principles and Applications	3	*			*	*
16	AML 5204	Reinforcement Learning	3	*			*	*
17	AML 5231	Applied Mathematics for Machine Learning	3	*		*		
	AML 5232	Convolutional Neural Networks for Computer Vision	3	*			*	*
	AML 5233	Natural Language Processing Principles and Applications	3	*			*	*
	ENP 5230	Entrepreneurship	3			*	*	
18	AML 5251	Advanced Applications of Probability and Statistics Lab	1			*		*
19	AML 5252	Deep Learning Lab	1	*			*	*
20	AML 5253	Machine Learning Principles and Applications Lab	1	*			*	*
21	AML 5254	Reinforcement Learning Lab	1	*			*	*
22	AML 5281	Applied Mathematics for Machine Learning Lab	1	*		*		
	AML 5282	Convolutional Neural Networks for Computer Vision Lab	1	*			*	*
	AML 5283	Natural Language Processing Principles and Applications Lab	1	*			*	*
	ENP 5280	Entrepreneurship Lab	1			*	*	
23	MPT 5200	Mini Project – II	4	*	*	*	*	*
24	PSD 5200	Professional Skill Development – II	1			*	*	*
25	AML 6098	Project Work	25	*	*	*	*	*