Master of Engineering - ME (Big Data Analytics)

Course Name : Fundamentals of Machine Learning Lab

--- --- ---

Course Code : BDA 5153

Academic Year : 2024 - 25

Semester : I

Name of the Course Coordinator : Dr. Arockiaraj S

Name of the Program Coordinator : Dr. Prathviraj N

Course File

Signature of Program Coordinator Signature of Course Coordinator

--- ---

with Date with Date

Program Education Objectives (PEOs)

The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for **ME (Big Data Analytics)**, program are as follows.

PEO No.	Education Objective
PEO 1	Develop in depth understanding of the key technologies in data engineering, data science and business analytics.
PEO 2	Practice problem analysis and decision-making using machine learning techniques.
PEO 3	Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and applied research experiences.

Program Outcomes (POs)

By the end of the postgraduate program in **ME (Big Data Analytics)**, graduates will be able to:

PO1	Independently carry out research /investigation and development work to solve practical problems.
PO2	Write and present a substantial technical report/document.
PO3	Demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program.
PO4	Develop and implement big data analysis strategies based on theoretical principles, ethical considerations, and detailed knowledge of the underlying data.
PO5	Demonstrate knowledge of the underlying principles and evaluation methods for analyzing data for decision-making.

1. Course Plan

1.1 Primary Information

Course Name : Fundamentals of Machine Learning Lab [BDA 5153]
L-T-P-C : 0-0-3-1
Contact Hours : 36 Hours

Pre-requisite : Basic Programming with Python

Core/ PE/OE : Core

1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping

CO At the end of this course, the student should be able to: No. of Contact Program Outcomes BL

--- --- --- --- ---

Hours (PO's)

CO1 Apply different types of supervised and unsupervised 8 PO3 3

machine learning algorithms to practical problems.

CO2 Analyse different types of machine learning paradigms. 12 PO4 4

CO3 Evaluate the performance of machine learning algorithms. 16 PO5 5

1.3 Assessment Plan

Components Lab test Flexible Assessments End semester/ Makeup

--- --- --- ---

(2 - 3 in number) examination

Duration 90 minutes To be decided by the faculty. 180 minutes

Weightage 0.3 0.2 0.5

Typology of Applying; Analyzing and Evaluating. Applying; Analyzing. Applying; Analyzing;

questions Evaluating. Evaluating.

Assignment: Solving problems by

Pattern Answer all the questions. Maximum marks applying, analyzing and evaluating Answer all the questions.

30. Generative AI use cases. Maximum marks 50.

[To be decided by the faculty.]

Schedule To be decided by the faculty Assignment submission: (To be decided by the faculty) To be decided by the faculty

Introduction to Machine Learning; Decision Comprehensive examination

Topics Trees- Linear Model: K-nearest Neighbours Lab assignment: Implement pre-training, covering the full syllabus.

fine-tuning, and evaluation of large

covered Algorithm- Cross-validation - Dimension language models. Students are expected to

Reduction. answer all questions.

1.4 Lesson Plan

L. No.	TOPICS	Course Outcome Addressed
L0	Course delivery plan, Course assessment plan, Course outcomes, Program outcomes, CO-PO mapping, reference books	---
L1	Program data ingestion, perform data wrangling, understand the data matrix, and differentiate between sample and feature.	CO1
L2	Implement decision tree models in Python, fine-tune model parameters, and interpret results.	CO1
L3	Implement linear models in Python and interpret model coefficients for practical problems.	CO2
L4	Implement and visualize bias-variance trade-off using linear regression as a basis; Implement the K-nearest neighbours algorithm.	CO2
L5	Implement, visualize, compare, contrast, and interpret the results of dimension reduction applied to practical data using PCA, MDS, and t-SNE.	CO2
IT1	Internal lab test	CO1 & CO2
L6	Through coding, understand how ensemble methods in machine learning work.	CO3
L7	Implement maximum likelihood estimation for a simple model.	CO3

L8 Analyse the performance of the Naive Bayes and logistic regression models for practical problems using appropriate performance metrics. CO3

L9 Lab assignment: Implement pre-training, fine-tuning, and evaluation of large language models. CO3

--- --- ---

L10 Lab assignment: Implement pre-training, fine-tuning, and evaluation of large language models. CO3

L11 Lab assignment: Implement pre-training, fine-tuning, and evaluation of large language models. CO3

L12 Lab assignment: Implement pre-training, fine-tuning, and evaluation of large language models. CO3

1.5 References

- 1. Module: Introduction to Machine Learning (https://www.intel.com/content/www/us/en/developer/tools/oneapi/training/academicprogram/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/getstarted-with-artificial-intelligence-on-azure/)
- 4. Learning path: Understand data science for machine learning (https://learn.microsoft.com/en-us/training/paths/understand-machinelearning/)
- 5. Module: Generative AI with Large Language Models (https://www.coursera.org/learn/generative-ai-with-llms)
- 6. 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
- 7. A Course in Machine Learning, Hal Daumé III Online resource available at http://ciml.info/

- 8. An Introduction to Statistical Learning: with Applications in Python (Springer Texts in Statistics), Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, and Jonathan Taylor, 1st Edition, 2023 - Online resource available at https://www.statlearning.com/

1.6 Other Resources (Online, Text, Multimedia, etc.)

- 1. Web Resources: Blog, Online tools and cloud resources.

- 2. Journal Articles.

1.7 Course Timetable

1 st Semester Big Data Analytics Lab:

--- --- --- --- --- --- --- --- ---

9-10 10-11 11-12 12-1 1-2 2-3 3-4 4-5

MON

TUE FML LAB

WED

THU

FRI

SAT

1.8 Assessment Plan

COs Marks & Weightage

--- --- --- --- --- ---

CO No. CO Name Mid semester Assignment End Semester CO wise

(Max. 50) (Max. 20) (Max. 100) Weightage

CO1 Apply different types of supervised and unsupervised 25 5 20 0.30

machine learning algorithms to practical problems.

Analyse different types of machine learning

25 5 40 0.40

CO2 paradigms.

Evaluate the performance of machine learning

CO3 - 10 40 0.30

algorithms.

Marks (weightage) 0.3 0.2 0.5 1.0

Note:

- In-semester Assessment is considered as the Internal Assessment (IA) in this course for 50 marks, which includes the performances in class participation, assignment work, class tests, mid-term tests, quizzes etc.

- End-semester examination (ESE) for this course is conducted for a maximum of 100 and the same will be scaled down to 50.

- End-semester marks for a maximum of 50 and IA marks for a maximum of 50 are added for a maximum of 100 marks to decide upon the grade in this course.

Weightage for CO1 = (mid semester marks for CO1 / 1.6666 + Assignment marks for CO1/1.0 + ESE marks for CO1 / 2)/100

1.9 Assessment Details

The assessment tools to be used for the Current Academic Year (CAY) are as follows:

SI. Tools Weightage Frequency Details of Measurement (Weightage/Rubrics/Duration, etc.)

--- --- --- --- --- --- ---

No.

- Performance is measured using internal test attainment level.

- Reference: question paper and answer scheme. 1 Internal Test 0.3 1

Each internal test is assessed for a maximum of 50 marks and scaled down to 40 -

marks.

- Performance is measured using assignments/quiz attainment level. 2 Assignments 0.2 2

- Assignments/quiz are evaluated for a maximum of 10 marks.

- Performance is measured using ESE attainment level.

3 ESE 0.5 1 - Reference: question paper and answer scheme.

- ESE is assessed for a maximum of 100 marks and scaled down to 50 marks.

1.10 Course Articulation Matrix

CO	PO1	PO2	PO3	PO4	PO5
CO1	Y				
CO2		Y			
CO3			Y		
Average Articulation Level	Y	Y	Y		