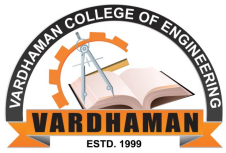

VARDHAMAN COLLEGE OF ENGINEERING

(AUTONOMOUS)

Affiliated to **JNTUH,** Approved by **AICTE,** Accredited by **NAAC** with **A++** Grade, **ISO 9001:2015** Certified

Kacharam, Shamshabad, Hyderabad - 501218, Telangana, India

www.vardhaman.org

CURRICULUM

For

Bachelor of Technology (Minors)

In

Artificial Intelligence & Machine Learning

Department of Computer Science and Engineering

VARDHAMAN COLLEGE OF ENGINEERING, HYDERABAD

An Autonomous Institute, Affiliated to JNTUH

Programme Curriculum Structure

B.Tech Minors in Artificial Intelligence & Machine Learning

S.No.	Yr/ Sem	Course Code	Course Name	Type	Credits
1	III/I	M1501	Big Data Tools	Theory	3
2	III/I	M1502	Data Analysis with Python	Practice	2
3	III/II	M1503	Machine Learning Methods	Theory	3
4	III/II	M1504	Machine Learning Laboratory	Practice	2
5	IV/I	M1505	Artificial Intelligence	Theory	3
6	IV/I	M1506	Natural Language Processing	Theory	3
7	IV/II	M1541	Mini Project in Minor Specialization	Project Work	2
Total Credits					18

Course Structure
M1501 - Big Data Tools

Hours Per Week			Hours Per Semester			Credits	Assessment Marks		
L	T	P	L	T	P	C	CIE	SEE	Total
3	0	0	42	0	0	3	30	70	100

1. Course Description

Course Overview

It serves as an introductory course for graduate students who are expecting to face Big Data storage, processing, analysis, visualization, and application issues on both work-places and research environments.Big data analytics is the use of analytic techniques against very large, diverse data sets that include structured, semi-structured and un-structured data, from different sources and in different sizes. Analysis of big data allows users to make better and faster decisions using data that was previously inaccessible or unusable. Students will gain knowledge on analysing Big Data.

Course Pre/co-requisites

The course has no specific prerequisite and co requisite.

2. Course Outcomes (COs)

After the completion of the course, the student will be able to:

M1501.1 Identify the fundamental concepts of big data analytics.
M1501.2 Select Hadoop environment and apply HDFS commands on file management tasks.
M1501.3 Make use of NoSQL databases like MangoDB to stock log data to be pulled for analysis.
M1501.4 Choose modern tools Pig and Hive for complex data flow and analysis.

3. Course Syllabus

Big Data: Classification of Digital Data, Characteristics of Data, Def-inition of Big Data, Challenges with Big Data, Traditional Business Intelligence (BI) versus Big Data, Realms of Big Data,Classification of Analytics,Few Top Analytics Tools. NoSQL:NoSQL (Not Only SQL), Types of NoSQL Databases, SQL versus NoSQL, RDBMS versus Hadoop, Distributed Computing Challenges

HDFS: Introduction to Hadoop, Hadoop Overview, Hadoop Distributors, HDFS (Hadoop Distributed File System),Working with HDFS commands,Interacting with Hadoop Ecosystem.

MONGODB: Features of MongoDB, RDBMS vs MongoDB, Data Types in MongoDB, MongoDB Query Language, CRUD operations,Count, Limit, Sort, and Skip.

Hive: Saying Hello to Hive, Seeing How the Hive is Put Together, Getting Started with Apache Hive,Examining the Hive Clients, Working with Hive Data Types, Creating and Managing Databases and Tables, Seeing How the Hive Data Manipulation Language

Works, Querying and Analyzing Data.

Pig: Hadoop Programming Made Easier Admiring the Pig Architecture, Going with the Pig Latin Application Flow, Working through the ABCs of Pig Latin, Evaluating Local and Distributed Modes of Running Pig Scripts, Checking out the Pig Script Interfaces, Scripting with Pig Latin.

4. Books and Materials

Text Books:

1. Seema Acharya, Subhashini Chellappan. Big Data and Analytics, 2nd Edition,Wiley IndiaPrivate Limited, New Delhi, 2019.

Reference Books:

1. Tom White. Hadoop - The Definitive Guide, 4th Edition, O'Reilly Publications, India,2015.
2. Judith Hurwitz, Alan Nugent, Dr. Fern Halper, Marcia Kaufman. Big Data for Dummies, John Wiley & Sons, Inc., 2013.

Course Structure
M1502 - Data Analysis with Python

Hours Per Week			Hours Per Semester			Credits	Assessment Marks		
L	T	P	L	T	P	C	CIE	SEE	Total
0	0	4	0	0	56	2	30	70	100

1. Course Description

Course Overview

Python is a powerful programming language used for Artificial Intelligence Machine Learning applications. In this course learners will be able to implement various python libraries Numpy and Pandas to work with multi dimensional arrays and data sets of real time problems. NumPy and Pandas libraries are essential and fundamental requirement for building machine learning models.

Course Pre/co-requisites

The course has no specific prerequisite and co requisite.

2. Course Outcomes (COs)

After the completion of the course, the student will be able to:

M1502.1 Identify the process of data analysis for a given data set.
M1502.2 Make use of Numpy functions to perform various mathematical operations on arrays.
M1502.3 Use Pandas to perform data analysis in machine learning applications.
M1502.4 Construct various plots on real data to identify useful insights.

3. Course Syllabus

Introduction to Data Analysis: Understanding the Nature of the Data, The Data Analysis Process, Data Extraction, Data Preparation, Data Exploration/Visualization, Predictive Modeling, Model Validation, Deployment, Quantitative and Qualitative Data Analysis, Open Data, Python and Data Analysis.

Introduction to NumPy: NumPy arrays, special numeric values, creating numpy arrays, creating ndarray, Numpy operators. Operations on NumPy Arrays: Selecting elements explicitly,Slicing arrays,Expanding arrays,Arithmetic and linear algebra with arrays,Employing array methods and functions.

NumPy Arrays: Computation on NumPy Arrays, Universal Functions, Aggregations,Min, Max, and Everything in Between, Computation on Arrays: Broadcasting Comparisons, Masks, and Boolean Logic, Fancy Indexing, Sorting Arrays,Structured Data: NumPy's Structured Arrays,

Pandas and Opearations: Exploring series and Dataframe objects, creating series, creating dataframes, adding data, saving data frames, subsetting your data, indexing methods, Arithmetic, Function Application and Mapping with pandas: arithmetic operations with dataframes,vectorization with dataframes,Dataframe function application,handling missing data in pandas DataFrame.

Managing, Indexing and Plotting: Using pandas for Data Analysis, Index sorting, Hierarchial indexing,plotting with pandas.

List of Programs for Practice

1. Program on Numpy Aggregations: Min, Max, and etc.. Example What Is the Average Height of prime ministers of india?
2. Program using NumpyComparisons, Masks, and Boolean Logic example: Counting Rainy Days
3. Program using Numpy Fancy Indexing example: Selecting Random Points.
4. Write a NumPy program to create a 3x3 identity matrix.
5. Write a NumPy program to create a vector of length 10 with values evenly distributed between 5 and 50.
6. Program using Pandas to Combining Datasets: Join.
7. Program using Pandas on Pivot Tables.
8. Program using Pandas to Vectorized String Operations.
9. Program using Pandas to Working with Time Series Example: Visualizing Seattle Bicycle Counts.
10. Write a NumPy program to swap rows and columns of a given array in reverse order.
11. Write a NumPy program to compute the mean, standard deviation, and variance of a given array along the second axis.
12. Write a NumPy program to sort the student id with increasing height of the students from given students id and height. Print the integer indices that describes the sort order by multiple columns and the sorted data.

4. Laboratory Equipment/Software/Tools Required

1. A computer System with Ubuntu Operating System.
2. Python 3.x (Any Latest Version).
3. Anakonda , Jupyter Notebook

5. Books and Materials

Text Books:

1. Jake VanderPlas. Python Data Science Handbook, 1st Edition, Published by O'Reilly Media, 2017.
2. Curtis Miller., Hands-On Data Analysis with NumPy and pandas, Packt Publishing, India,2018.

Reference Books:

1. Jake VanderPlas., Python Data Science Handbook,1st*Edition*,O'Reilly Publications Media, 2017.
2. Fabio Nelli., Python Data Analytics with Pandas, NumPy, and Matplotlib, 2nd Edition, Apress, 2018.

Course Structure
M1503 - Machine Learning Methods

Hours Per Week			Hours Per Semester			Credits	Assessment Marks		
L	T	P	L	T	P	C	CIE	SEE	Total
3	0	0	42	0	0	3	30	70	100

1. Course Description

Course Overview

This course provides a concise introduction to the fundamental concepts in machine learning and popular machine learning algorithms. The students will be introduced to the foundations of machine learning along with a slew of popular machine learning techniques. This will also give insights on how to apply machine learning to solve a new problem. Students will learn the algorithms which under pin many popular Machine Learning techniques, as well as developing an understanding of the theoretical relationships between these algorithms.

Course Pre/co-requisites

M1502 - Data Analysis with Python

2. Course Outcomes (COs)

After the completion of the course, the student will be able to:

M1503.1 Identify the design process and applications of machine learning to solve a real time problem.
M1503.2 Select appropriates pre processing techniques to extract useful insights from real time data.
M1503.3 Examine supervised learning methods on a given data set.
M1503.4 Make use of clustering techniques for several hidden patterns in data.
M1503.5 Identify the need of neural network in bulding a model

3. Course Syllabus

Introduction to Machine Learning: Programs vs learning algorithms, Machine Learning definition, Components of a learning, Different Types of Learning, Applications of machine learning, case studies on machine learning, Life Cycle of Machine Learning model.

Data Preprocessing: Overview of Data preprocessing, Data cleaning, Data reduction, modeling and Evaluation: Selecting a Model, Training a Model, Model Representation and Interpretability.

Supervised Learning: Introduction to Supervised Learning and Regression, Statistical Relation between variables, Steps to establish Linear Regression, Evaluation of model estimators. Introduction to logistic regression, Building logistic regression model, Decision trees and support vector machines, Bayes' Theorem, Naïve Bayes Classifier, Applications of Naïve Bayes Classifier.

Unsupervised Learning: Introduction, Unsupervised vs Supervised Learning, Application of Unsupervised Learning, Different Types of Clustering ,Clustering K-Means,

K-Mediods.

Artificial Neural Networks: Artificial Neural Network- Basic Architecture, Non linear model of a neuron, Network Architectures- Single Layer Feed Forward, Multi Layer Feed Forward and Recurrent Networks. Activation Functions, Cost Functions and Optimizers, Backpropagation.

4. Books and Materials

Text Books:

1. Anuradha Srinivasaraghavan, Vincy Joseph., Machine Learning, Wiley ,2020.
2. Simon Haykin., Neural Networks: a Comprehensive Foundation, 2nd Edition, Pearson Education, India, 1999.

Reference Books:

1. Tom M.Mitchell., Machine Learning, India Edition, McGraw Hill Education, 2013.

Course Structure
M1504 – Machine Learning Laboratory

Hours Per Week			Hours Per Semester			Credits	Assessment Marks		
L	T	P	L	T	P	C	CIE	SEE	Total
3	0	0	42	0	0	3	30	70	100

1. Course Description

Course Overview

This course provides a concise introduction to the fundamental concepts in machine learning and popular machine learning algorithms. This undergraduate-level course, students will be introduced to the foundations of machine learning along with a slew of popular machine learning techniques. This will also give insights on how to apply machine learning to solve a new problem.Students will learn the algorithms which underpin many popular Machine Learning techniques, as well as developing an understanding of the theoretical relationships between these algorithms

Course Pre/co-requisites

M1502- Data Analysis with Python
M1503- Machine Learning Methods

2. Course Outcomes (COs)

After the completion of the course, the student will be able to:

M1504.1. Implement Supervised learning algorithms on a given data set.
M1504.2. Examine the performance of various unsupervised algorithms on real time data.
M1504.3. Implement the domain specific case studies using machine learning algorithms

3. Course Syllabus

List of Programs

- Implement and demonstrate the FIND-S algorithm for finding the most specific hypothesis.
- Implement and demonstrate the Candidate Elimination algorithm for finding the most specific hypotheses.
- Implement Linear Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw graphs.
- Write a program to demonstrate the working of the decision tree based ID3 algorithm.
- Write a program to implement Support Vector Machine (SVM).
- Write a program to implement the naïve Bayesian classifier and compute accuracy.
- Write a program to construct a Bayesian network by considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set.

- Write a program to implement the KNN classifier for a sample training data sets.
- Write a program to implement K-Means Clustering algorithm.
- Write a program to implement agglomerative hierarchical Clustering algorithm.
- Case Study-1 (Domain Specific)
- Case Study-2 (Domain Specific)
- Case Study-3 (Domain Specific)

4. Books and Materials

Text Books:

1. Tom M.Mitchell., Machine Learning, India Edition, McGraw Hill Education, 2013.
2. Anuradha Srinivasaraghavan, Vincy Joseph., Machine Learning, Wiley ,2020.

Reference Books:

1. Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar., Introduction to Data Mining, 2nd Edition, , Pearson Publication, 2016.

Course Structure
M1505 - Artificial Intelligence

Hours Per Week			Hours Per Semester			Credits	Assessment Marks		
L	T	P	L	T	P	C	CIE	SEE	Total
3	0	0	42	0	0	3	30	70	100

1. Course Description

Course Overview

Artificial intelligence (AI) is a research field that studies how to realize the intelligent human behaviors on a computer. The ultimate goal of AI is to make a computer that can learn, plan, and solve problems autonomously.This course is to acquire the ability to design intelligent solutions to problems in a variety of domains and business applications such as natural language Processing, text mining, and robotics, reasoning and problem-solving. AI will focus on problem solving, reasoning, planning and gaming. Through learning problem solving skills can be acquired.

Course Pre/co-requisites

The course has no specific prerequisite and co requisite.

2. Course Outcomes (COs)

After the completion of the course, the student will be able to:

M1505.1 Identify the need and design challenges of AI in real time.
M1505.2 Use the searching strategies to design AI problems.
M1505.3 Make use of various planning techniques for finding solution to a problem.
M1505.4 Utilize various game playing approaches and Expert Systems in AI to design solution to complex problems.

3. Course Syllabus

Introduction: Introduction to AI and Turing Test, Weak AI, Strong AI, Structure of AI, Advantages of AI, Defining AI techniques, Using Predicate Logic and Representing Knowledge as Rules, Representing simple facts in logic, Computable functions and predicates, Procedural vs Declarative knowledge, Logic Programming. AI Technique, General Problem Solving.

Solving problem by searching: Knowledge-Based Agent, the Wumpus World, Logic, Propositional Logic. First-Order Logic: Syntax and Semantics of First-Order Logic, Extensions and Notational Variations. Inference in First-Order Logic: Inference Rules Involving Quantifiers, Generalized Modus Ponens, Forward and Backward Chaining.

Intelligent Agents and Planning: Knowledge-Based Agent, the Wumpus World, Logic, Propositional Logic. First-Order Logic: Syntax and Semantics of First-Order Logic, Extensions and Notational Variations. Inference in First-Order Logic: Inference Rules Involving Quantifiers, Generalized Modus Ponens, Forward and Backward Chaining.

Game Playing and Expert Systems: : Introduction- Games as a Search problem, Perfect Decisions in Two-Person Games, Imperfect Decisions, Alpha-Beta Pruning.

Expert Systems: Introduction to expert system, the expert system Development process, knowledge acquisition, Expert System Tools.

4. Books and Materials

Text Books:

1. Stuart Russel, Peter Norvig., Artificial Intelligence – A Modern Approach, 3nd Edition, Pearson Education, 2010.
2. Tom Taulli., Artificial Intelligence Basics: A Non-Technical Introduction, Apress, 2019.

Reference Books:

1. Saroj Kaushik. Artificial Intelligence, Cengage Learning India Private Limited, 2011.

Course Structure
M1506 - Natural Language Processing

Hours Per Week			Hours Per Semester			Credits	Assessment Marks		
L	T	P	L	T	P	C	CIE	SEE	Total
3	0	0	42	0	0	3	30	70	100

1. Course Description

Course Overview

A major portion of communication now is through text and any organization has most of its content in the unstructured form. Natural Language Processing (NLP), an important part in Artificial Intelligence, is one of the technology that would help in activities such as classification, retrieving and extraction of information, identifying important documents, etc. Students will gather Knowledge in the fundamentals of NLP, methods and techniques and gain skills to use them in practical situations. This course also deals with grammar, semantics and pragmatics. The course enables the student to know in detail the various applications of NLP in real time.

Course Pre/co-requisites

M1503 - Machine Learning Methods
M1505 - Artificial Intelligence

2. Course Outcomes (COs)

After the completion of the course, the student will be able to:

M1506.1 Identify the NLP pre-processing techniques for analysing text.
M1506.2 Inspect Speech recognition approach for converting speech into text.
M1506.3 Make use of grammar rules while parsing text.
M1506.4 Examine the various semantics and pragmatics in processing text.
M1506.5 Compare statistical approaches for various NLP applications.

3. Course Syllabus

Introduction: Words - Regular Expressions and Automata - Words and Transducers -N-grams - Part-of- Speech – Tagging - Hidden Markov and Maximum Entropy Models.

Speech: Speech, Phonetics, Speech Synthesis, Automatic Speech Recognition, Speech Recognition Advanced Topics, Computational Phonology.

Syntax: Formal Grammars of English - Syntactic Parsing - Statistical Parsing - Features and Unification - Language and Complexity.

Semantics and Pragmatics: The Representation of Meaning - Computational Semantics - Lexical Semantics - Computational Lexical Semantics - Computational Discourse.

Applications: Information Extraction, Question Answering and Summarization, Dialogue and Conversational Agents, Machine Translation.

4. Books and Materials

Text Books:

1. Daniel Jurafsky, "Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech", Pearson Publication, 2014.
2. Steven Bird, Ewan Klein and Edward Loper, Natural Language Processing with Python, 1st Edition, O'Reilly Media, 2009.

Reference Books:

1. Nitin Indurkhya and Fred J. Damerau., Handbook of Natural Language Processing, 2nd Edition, Chapman and Hall/CRC Press,2010.