|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Woldia University**  **Institute of Technology**  **School of Computing**  **Software Engineering Program** | | | | | |
| **Program** | Software Engineering, Regular | | | | |
| **Course Code** | SEng3092 | | | | |
| **Course Title:** | Fundamentals of Machine Learning | | | | |
| **Degree Program** | BSc. in Software Engineering | | | | |
| **Module Name/ No.** | AI, Machine Learning and Big Data/ 09 | | | | |
| **Course Instructor** | Prince M Thomas | | | | |
| **ECTS (Cr. Hrs.)** | 5 (3) | | | | |
| **Contact Hours**  **(per week)** | ***Lecture*** | ***Tutorial*** | ***Lab/Practical*** | ***Home Study*** | ***Total*** |
| 2 |  | 3 |  |  |
| **Target Group:** | 3rd year | | | | |
| **Year /Semester** | Year: III, Semester: II | | | | |
| **Pre-requisites** | SEng2022 | | | | |
| **Course Category** | Compulsory | | | | |
| **Course Description** | Machine learning techniques enable us to automatically extract features from data so as to solve predictive tasks, such as speech recognition, object recognition, machine translation, question-answering, anomaly detection, medical diagnosis and prognosis, automatic algorithm configuration, personalization, robot control, time series forecasting, and much more. Learning systems adapt so that they can solve new tasks, related to previously encountered tasks, more efficiently. This course will introduce the field of Machine Learning, in particular focusing on the core concepts of supervised and unsupervised learning. In supervised learning we will discuss algorithms which are trained on input data labelled with a desired output, for instance an image of a face and the name of the person whose face it is, and learn a function mapping from the input to the output. Unsupervised learning aims to discover latent structure in an input signal where no output labels are available, an example of which is grouping web-pages based on the topics they discuss. 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. The practical‘s will concern the application of machine learning to a range of real-world problems. | | | | |

|  |  |
| --- | --- |
| **Course Objective** | The course aims to provide an introduction to the basic principles, techniques, and applications of Machine Learning. Programming assignments are used to help clarify basic concepts. The course covers the  principles, design and implementation of learning programs that improve their performance on some set of tasks with experience. Upon successful completion of the course, students will have a broad understanding of machine learning algorithms and their use in data-driven knowledge discovery and program synthesis. Students will have designed and implemented several machine learning algorithms in Python. Students will  also be able to identify, formulate and solve machine learning problems that arise in practical applications. |
| **Learning Outcomes** | The program provides a knowledge and understanding of the following:  ● Be able to formulate machine learning problems corresponding to different applications.  ● Understand a range of machine learning algorithms along with their strengths and weaknesses.  ● Understand the basic theory underlying machine learning.  ● Be able to apply machine learning algorithms to solve problems of moderate complexity.  ● Apply machine learning algorithm for classification and problems in health-related problem domains like medicine, genetics, medical diagnosis. |
| **Course Outline** | **Chapter 1: Introduction to Machine Learning**  Artificial Intelligence and Machine Learning  - *What is Artificial Intelligence?*  *- Relationship between AI, Machine Learning, and Data Science*  What is Machine Learning?  - *Definition and Key Concepts (Features, Labels, Models, and Predictions)*  *- Machine Learning Pipeline (Data Collection, Preprocessing, Model Training)*  Machine Learning Problems and Applications  - *Regression, Classification Problems*  *- Applications in Healthcare, Agriculture, Finance, and Robotics*  Types of Machine Learning  *- Supervised Learning*  *- Unsupervised Learning*  *- Reinforcement Learning*  **Chapter 2: Supervised Machine Learning: Linear Models and Fundamentals**  Linear Regression  - *Simple Linear Regression and Multiple Linear Regression*  *- Cost Function (Mean Squared Error)*  *- Gradient Descent for Linear Regression*  Logistic Regression (Linear Classification)  - *Sigmoid Function and Decision Boundaries*  *- Binary Classification vs. Multi-class Classification*  Gradient Descent Algorithm  - *Stochastic Gradient Descent (SGD) vs. Batch Gradient Descent*  *- Learning Rate and Convergence Issues*  Overfitting and Underfitting  *- Bias-Variance Trade-off*  *- Cross-Validation Techniques*  Regularization Techniques (L1, L2)  *- Lasso and Ridge Regression*  *- High Dimensional Data*  *- Curse of Dimensionality and Feature Selection*  Multivariate Methods  - *Covariance Matrix*  *- Linear Models for Multivariate Regression*  Parametric vs. Non-parametric Methods  *- Differences, Examples, and Use Cases*  **Chapter 3: Supervised Learning: Nonlinear Models**  K-Nearest Neighbors (K-NN)  *- Distance Metrics (Euclidean, Manhattan)*  *- K-Value Selection and Cross-Validation*  Neural Networks and Multilayer Perceptrons (MLPs)  *- Structure of Neural Networks (Input, Hidden, and Output Layers)*  *- Activation Functions (ReLU, Sigmoid, Softmax)*  *- Backpropagation and Gradient Descent in Neural Networks*  Decision Trees  - *Splitting Criteria: Gini Index, Entropy, and Information Gain*  *- Overfitting in Decision Trees and Pruning*  Random Forests (Introduction to Ensembles)  - *Bagging and Random Subspace Sampling*  Boosting Techniques (e.g., AdaBoost)  - *Concept of Weak Learners*  *- Boosting Algorithms: AdaBoost, Gradient Boosting*  Stacking and Voting Methods  *- Model Combination Techniques*  *- Hard vs. Soft Voting*  **Chapter 4: Unsupervised Learning and Graphical Models**  Clustering Methods  - *K-Means Clustering: Centroid Calculation and Convergence*  *- Hierarchical Clustering: Agglomerative and Divisive Methods*  Dimensionality Reduction  - *Principal Component Analysis (PCA): Eigenvectors and Eigenvalues*  *- Linear Discriminant Analysis (LDA)*  Frequent Pattern Mining  - *Apriori Algorithm and Association Rules*  *- Applications in Market Basket Analysis*  Graphical Models  - *Bayesian Networks*  *- Markov Networks*  Hidden Markov Models (HMMs)  - *States, Transitions, and Observations*  *- Forward Algorithm and Viterbi Algorithm*  *- Applications in Speech Recognition and NLP*  **Chapter 5: Reinforcement Learning**  Introduction to Reinforcement Learning  - *Definitions: Agent, Environment, Action, Reward, and Policy*  *- Exploration vs. Exploitation*  Reinforcement Learning Problem Setup  - *Markov Decision Processes (MDPs)*  *- Value Function, Policy Function, and Bellman Equation*  Q-Learning Algorithm  - *Q-Table and Update Rule*  *- Temporal Difference Learning* |

|  |  |
| --- | --- |
| **Assessment** | Mid Exam = 30%  Assignment(Individual) = 10%  Project (Group) = 20%  Final Exam = 40% |
| **Reference:** | **Text Books**   * Kevin P. Murphy (2012): Machine Learning, A Probabilistic Approach, MIT Press. * Christopher M. Bishop (2006): Pattern Recognition and Machine Learning.   **Other References:**   * Han J. and Kamber M. (2006): Data Mining: Concepts and Techniques, Second Edition * Python and R codes: [http://www.analyticsvidhya.com/blog/2015/09/full-cheatsheet-](http://www.analyticsvidhya.com/blog/2015/09/full-cheatsheet-machine-learni%20ng-algorithms/) [machine-learni ng-algorithms/](http://www.analyticsvidhya.com/blog/2015/09/full-cheatsheet-machine-learni%20ng-algorithms/) * Introductory Primer to Machine Learning: [http://www.toptal.com/machine-learning/machine-learning-](http://www.toptal.com/machine-learning/machine-learning-theory-an-introductory-primer)   [theory-an-introductory-primer](http://www.toptal.com/machine-learning/machine-learning-theory-an-introductory-primer) |