Course Code	Course Title					Course Type
CS3202	Machine Learning					PCC
Prerequisite	Contact Hours per Week			Internal	External	Credits
	L	T	P			
Knowledge on statistical methods	3	0	0	40	60	3

Course Objectives

- To provide an introduction to the field of machine learning and its applications
- To teach the fundamental concepts of linear regression, decision trees, instance-based learning, and clustering
- To enable students to understand the basics of artificial neural networks and support vector machines
- To introduce students to ensemble learning techniques and recommender systems
- To provide hands-on experience in implementing machine learning algorithms using Python

Course Outcomes

1. Understand the basic concepts of machine learning and its applications
2. Apply linear regression and decision tree learning techniques to real-world problems
3. Understand and apply instance-based learning and clustering algorithms
4. Implement artificial neural networks and support vector machines
5. Apply ensemble learning techniques and develop recommender systems
6. Use Python to implement machine learning algorithms and evaluate their performance

Detailed Contents

UNIT I

Introduction: Introduction to Machine Learning: Introduction. Different types of learning, Hypothesis space and inductive bias, Evaluation. Training and test sets, cross validation, Concept of over fitting, under fitting, Bias and Variance.

Linear Regression: Introduction, Linear regression, Simple and Multiple Linear regression, Polynomial regression, evaluating regression fit.

UNIT II

Decision tree learning: Introduction, Decision tree representation, appropriate problems for decision tree learning, the basic decision tree algorithm, hypothesis space search in decision tree learning, inductive bias in decision tree learning, issues in decision tree learning, Python exercise on Decision Tree.

Instance based Learning: K nearest neighbor, the Curse of Dimensionality, Feature Selection: forward

search, backward search, univariate , multivariate feature selection approach, Feature reduction (Principal Component Analysis) , Python exercise on kNN and PCA.
Recommender System: Content based system, Collaborative filtering based.

UNIT III
Probability and Bayes Learning: Bayesian Learning, Naïve Bayes, Python exercise on Naïve Bayes, Logistic Regression.
Support Vector Machine: Introduction, the Dual formulation, Maximum margin with noise, nonlinear SVM and Kernel function, solution to dual problem.

UNIT IV
Artificial Neural Networks: Introduction, Biological motivation, ANN representation, appropriate problem for ANN learning, Perceptron, multilayer networks and the back propagation algorithm,

UNIT V
Ensembles: Introduction, Bagging and boosting, Random forest, Discussion on some research papers.
Clustering: Introduction, K-mean clustering, agglomerative hierarchical clustering, Python exercise on k-mean clustering.

Text Books

- Kevin P. Murphy, “Machine Learning: A Probabilistic Perspective”, MIT Press, 2012.

References

1. T. Mitchell, "Machine Learning," First Edition, McGraw-Hill, 1997.
2. E. Alpaydin, "Introduction to Machine Learning," MIT Press, 2020.
3. C. Bishop, "Pattern Recognition and Machine Learning," Springer, 2007.

Course Code	Course Title					Course Type
CS3802	Machine Learning Lab					PCC
Prerequisite	Contact Hours per Week			Internal	External	Credits
	L	T	P			
	0	0	3	40	60	1.5

Course Objectives

- The objective of this lab is to get an overview of the various machine learning techniques and can able to demonstrate them using python.

Course Outcomes

After the completion of the course the student can able to:

1. Understand complexity of Machine Learning algorithms and their limitations;
2. Understand modern notions in data analysis-oriented computing;
3. Be capable of confidently applying common Machine Learning algorithms in practice and implementing their own;
4. Be capable of performing experiments in Machine Learning using real-world data.

List of Experiments

1. Basic exercises on Python Machine Learning Packages such as Numpy, Pandas and matplotlib.
2. Given a dataset. Write a program to compute the Covariance, Correlation between a pair of attributes. Extend the program to compute the Covariance Matrix and Correlation Matrix.
3. Given a set of sample points in N dimensional feature space. Write a program to fit the points with a hyper plane using Linear Regression. Calculate sum of residual error.
4. Write a program that provides option to compute different distance measures between two points in the N dimensional feature space. Consider some sample datasets for computing distances among sample points.
5. Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.
6. Write a program to implement k-Nearest Neighbour algorithm to classify the iris data set. Print both correct and wrong predictions. Python ML library classes can be used for this problem.
7. Write a program to implement feature reduction using Principle Component Analysis
8. Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.
9. Given a dataset for classification task. Write a program to implement Support Vector Machine and estimate its test performance.
10. Write a program to implement perceptron for different learning tasks.
11. Write programs to implement ADALINE and MADALINE for a given learning task.
12. Build an Artificial Neural Network by implementing the Back propagation algorithm and test the

same using appropriate data sets.

13. Write a program to implement K means clustering algorithm. Select your own dataset to test the program. Demonstrate the nature of output with varying value of K.