

Machine Learning Lab	
Course Code: ISL65	Credits: : : 0:0:1
Pre – requisites: Python	Contact Hours: 28
Course Coordinator: Mrs. Shruthi G	

Sl.NO	List of Experiments
	Part-A
1	Model Measurement Analysis: Using <i>any dataset</i> , calculate TP, TN, FP ,FN and different metrics (Accuracy, Precision, Recall(Sensitivity), F1-Score, MCC, Specificity, Negative Predictive Value) by defining your own functions. Compare your values with scikit-learn's library functions. Get the result of Confusion Matrix using sklearn. Using sklearn, plot the ROC & AUC Curves for your test data and random probabilities. Using sklearn, calculate the AUC of your test data and of random probabilities. Interpret the results. Write the inference/analysis of each output.
2	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. Interpret the results. Write the inference/analysis of each output.
3	Design an experiment to investigate the impact of varying the number of trees in a Random Forest classifier on its performance for a given dataset. Write Python code to implement the Random Forest algorithm with different numbers of trees and evaluate its classification performance using appropriate evaluation metrics.
4	Supervised Learning Algorithms - Linear Regression: Assume the dataset to Create a Simple Linear Regression model. Predict the scores on the test data and output RMSE and R Squared Score. Include appropriate code snippets to visualize the model. Interpret the result.
5	Supervised Learning Algorithms - Logistic Regression (Univariant): Implement logistic regression and test it using any dataset. Give new test data and predict the classification output. Print the confusion matrix, accuracy, precision, recall, MSE , RMSE etc. Analyze and write the inference.
6	Probabilistic Supervised Learning - Naive Bayes: Create a dataset and Perform the necessary pre-processing steps. Train the model using Naive Bayes Classifier. Give new test data and predict the classification output. Analyze and write the inference.
7	Build a KNN model for predicting if a person will have diabetes or not with a high accuracy score. Perform some appropriate Pre-Processing steps on the given dataset for better results. Implement the KNN algorithm on your own. Try other possible processes that can be done to dataset and tuning the model to increase accuracy such as Increase K value, Normalization and Different Distance Metrics.
8	Un-Supervised Learning Algorithms - K-Means Clustering: Build a K-Means Model for any dataset. Assume K value as 2,3,4 .Compare and interpret the results of different clusters..
9	Un-Supervised Learning Algorithms - Hierarchical Clustering: Using any dataset implement Hierarchical Clustering (AGNES and DIANA). Plot the Dendrogram for Hierarchical Clustering and analyze your result. Plot the clustering output for the same dataset using these two hierarchical techniques. Compare the results. Write the inference.

10	Design an experiment to compare the performance of AdaBoost and XGBoost algorithms on a binary classification task using a dataset of your choice. Write Python code to implement both algorithms and evaluate their performance using appropriate evaluation metrics.
----	--

Course outcomes:

At the end of the course the student will be able to:

1. Understand the implementation procedures for the machine learning algorithms. (PO – 1, 2, 3, 4, 5) (PSO – 1, 3)
2. Design Java/Python programs for various Learning algorithms. (PO – 1, 2, 3, 4, 5) (PSO -1, 3)
3. Apply appropriate data sets to the Machine Learning algorithms. (PO - 1, 2, 3, 4, 5) (PSO – 1, 3)
4. Identify and apply Machine Learning algorithms to solve real world problems. (PO - 1, 2, 3, 4, 5) (PSO – 1, 3)

