School of Computer Science and Engineering, VIT Chennai.

BCSE209P Machine Learning Lab

Lab-1 Python : Numpy and Pandas

Faculty : Dr. R. Jothi

**Due Date : 07/01/2022**

Submit your python code (Jupyter notebook): with output for all the questions.

(Use appropriate library functions)

1. Create a 2-D Array 3 rows and 4 cols. Find transpose of the matrix. Also extract 2nd row of the transposed matrix,
2. Create a 4 x 4 identity matrix
3. Create a 5 x 4 matrix. Find descriptive statistics (min, max, std. deviation) about the array.
4. Load iris dataset. Print names of all the features of the dataset. Also print shape of the dataset (rows and columns)
5. Count number of distinct elements in species (i.e. the 3 different classes ) also count the number of samples in each of three classes
6. Checking if there is any inconsistency in the DataSet. (use df.info())
7. Extract all the rows pertaining to 'setosa' into a dataframe named as df_setosa. Similarly extract for 'versicolor', and 'virginica classes.
8. Visualize distribution of 3 different classes in the iris dataset.
9. Plot the relationship between three different iris classes with respect to petal length and petal width.
10. Visualize the correlation between sepal length and sepal width.