# CHAROTAR UNIVERSITY OF SCIENCE & TECHNOLOGY
# FACULTY OF TECHNOLOGY AND ENGINEERING
## Department of Information Technology

**Subject Name:** Machine Learning & Applications **Semester:** B.Tech VI
**Subject Code**: IT377 **Academic Year**: 2019-20

Note: The laboratory will emphasize the use of Python, Python Packages, Machine Learning and its applications.
Instructions:

1. **All Practical must be performed individually and all experimental results must be uploaded on your respective EDMODO account.**
2. **All Practical will be evaluated at the end of laboratory by concern Lab Teacher.**
3. **Each practical session should be have handwritten notes to be verified and submitted as part of term work.**
4. **Students will require giving presentation as and when requested by teacher.**

## Practical List

| Sr. No. | Aim of the Practical | Hrs | COs | POs | PEOs |
|---|---|---|---|---|---|
| Pre Req. | Introduction to Python Programming. How python used in machine learning? Discuss python with Google Colab. | - | 4,5 | 1,3,6 | 1,2,4 |
| 1. | Numpy<br>- Creating blank array, with predefined data, with pattern specific data<br>- Slicing and Updating elements,<br>- Shape manipulations<br>- Looping over arrays.<br>- Reading files in numpy<br>For Help:<br>https://www.dataquest.io/m/289-introduction-to-numpy/<br>https://cloudxlab.com/blog/numpy-pandas-introduction/<br>Pandas<br>- Creating data frame<br>- Reading files<br>- Slicing manipulations<br>- Exporting data to files<br>- Columns and row manipulations with loops<br>For Help:<br>https://www.hackerearth.com/practice/machine-learning/data-manipulation-visualisation-r-python/tutorial-data-manipulation-numpy-pandas-python/tutorial/<br>Matplotlib<br>- Importing matplotlib<br>- Simple line chart<br>- Correlation chart<br>- Histogram<br>For Help: | | | | |

| | | | | | |
|---|---|---|---|---|---|
| | https://towardsdatascience.com/data-visualization-using-matplotlib-16f1aae5ce70 | | | | |
| 2. | Linear Regression<br>Select Dataset of your choice and respond to following questions.<br>- Why you want to apply regression on selected dataset? Discuss full story behind dataset.<br>- How many total observations in data?<br>- How many independent variables?<br>- Which is dependent variable?<br>- Which are most useful variable in estimation? Prove using correlation.<br>- Quantify goodness of your model and discuss steps taken for improvement (RMSE, SSE, R2Score).<br>- Prepare presentation for this work in group of 5<br>**For help: refer following free course on datacamp.**<br>**Regression models: fitting them and evaluating their performance** | 2 | 4,5 | 1,3,6 | 1,2,4 |
| 3. | Two Class Classification (Logistic Regression)<br>Select Dataset of your choice and respond to following questions.<br>- Why you want to apply classification on selected dataset? Discuss full story behind dataset.<br>- How many total observations in data?<br>- How many independent variables?<br>- Which is dependent variable?<br>- Which are most useful variable in classification? Prove using correlation.<br>- Quantify goodness of your model and discuss steps taken for improvement (Accuracy, Confusion matrices, F-measure).<br>- Prepare presentation for this work in group of 5.<br>**For Help:**<br>1. https://medium.com/@anishsingh20/logistic-regression-in-python-423c8d32838b<br>2. https://www.datacamp.com/community/tutorials/understanding-logistic-regression-python<br>3. **https://towardsdatascience.com/logistic-regression-python-7c451928efee**<br>4. https://towardsdatascience.com/building-a-logistic-regression-in-python-step-by-step-becd4d56c9c8<br>5. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html | 2 | 4,5 | 1,3,6 | 1,2,4 |
| 4. | Multi Class Classification (KNN)<br>Select Dataset of your choice and respond to following questions.<br>- Why you want to apply classification on selected dataset? Discuss full story behind dataset.<br>- How many total observations in data?<br>- How many independent variables?<br>- Which is dependent variable? | 4 | 4,5 | 1,3,6 | 1,2,4 |

| | | | | | |
|---|---|---|---|---|---|
| | - Which are most useful variable in classification? Prove using correlation.<br>- Quantify goodness of your model and discuss steps taken for improvement.<br>- Can we use KNN for regression also? Why / Why not?<br>- Discuss drawbacks of algorithms such as KNN.<br>- Prepare presentation for this work in group of 5.<br>For Help:<br>https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/ | | | | |