

Name of the Programme: MCA

Course Code: CSA-512

Title of the Course: Machine Learning Lab

Number of Credits: 2 (0L+0T+ 2P)

Effective from AY: 2022-23

<u>Prerequisites for the course</u>	Course: Mathematics for Computer Science and Programming language background.	
<u>Objectives:</u>	The objective is to learn to build the different machine learning models by doing a set of assignments and mini projects.	
<u>Content:</u>	Introduction to python libraries for machine learning - scikit learn, tensor flow, keras, pytorch, pandas, matplotlib, seaborn, numpy and other relevant libraries.	5 hours
	Four branches of machine learning-supervised, unsupervised, self-supervised, reinforcement, Evaluating machine learning models, Data pre-processing, feature engineering and feature learning, overfitting and underfitting - Numerical Programming fundamentals-finding nearest neighbours via Euclidean distance-splitting data sets into training and testing.	10 hours
	Regression, cross validation and regularization-polynomial regression -model selection on a fixed validation set -Polynomial Regression - Model Selection with Cross-Validation-Polynomial Regression with L2 Regularization - Model Selection with Cross-Validation-Comparison of methods on the test set. Evaluating Binary Classifiers and Implementing Logistic Regression-Binary Classifier for movies reviews-classifying newswires-predicting house prices -Computing the Loss for Logistic Regression without Numerical Issues	10 hours
	Neural Networks and Stochastic Gradient Descent-MLPs with L-BFGS: What model size is effective?-MLPs with SGD: What batch size and step size?-Producing your own figure comparing batch size and learning rate.	10 hours
	Trees and Random Forests for Bag of Words-Code Implementation of Decision Tree Regression-Decision Trees for Review Classification - Random Forests for Review Classification -Comparing Trees to Linear Models for Review Classification.	10 hours
	Implementation of CNN, RNN, LSTM, Implementation of Boltzmann machine and Transformers (BERT, GPT3) .Generative deep learning (GAN).	10 hours
	Project discussions -Classifying Images with Feature Transformations- Classifying Sentiment from Text Reviews-Recommendation Systems via Matrix Factorization-Text summarization - language Translation - Sentimental analysis- speech to text translatioXiv, Explore the keras ecosystem.	5 hours
<u>Pedagogy:</u>	Programming in lab and practical exercises	
<u>References/Readings</u>	1. Hands on machine learning with scikit learn by Aurielien 2. Deep learning with python by Francois 3. Text Analytics with Python: A Practitioner's Guide to Natural Language Processing by dipanjan sarkar. 4. keras: the python deep learning API	

	5. https://www.cs.tufts.edu/comp/135/2020f/assignments.html	
	6. Python library reference	
<u>Course Outcomes</u>	Students will be able – 1. to collect data and preprocess them 2. choose the suitable machine learning model and 3. study its performance and able to carry out mini project	