Master of Science in Machine Learning & AI Curriculum

PRE-PROGRAM PREPARATORY CONTENT

Module/Assignment	Session
1 INTRODUCTION TO PYTHON Build a foundation for the most in-demand programming language of the 21st century.	• Understanding the upGrad Coding Console • Data Structures in Python • Control Structure and Functions
2 PYTHON FOR DATA SCIENCE Learn how to manipulate datasets in Python using Pandas which is the most powerful library for data preparation and analysis.	• Introduction to NumPy • Operations on NumPy Arrays • Introduction to Pandas • Getting and Cleaning Data
3 VISUALIZATION IN PYTHON Humans are visual learners and hence no task related to data is complete without visualisation. Learn to plot and interpret various graphs in Python and observe how they make data analysis and drawing insights easier.	• Introduction to Data Visualization • Basics of Visualization: Plots, Subplots and their Functionalities • Plotting Data Distributions • Plotting Categorical and Time-Series Data
4 DATA ANALYSIS USING SQL (OPTIONAL) Data in companies is definitely not stored in Excel sheets! Learn the fundamentals of database and extract information from RDBMS using the structured query language.	• Basics of SQL: Data Retrieval, • Compound Functions, Relational Operators, and Sorting • Advanced SQL: Aggregate Functions, • Nested Queries, and Joins

*Syllabus is subject to change by the university

PRE-PROGRAM PREPARATORY CONTENT

Module/Assignment	Session
5 **ADVANCED SQL AND BEST PRACTICES** (OPTIONAL) Apply advanced SQL concepts like windowing and procedures to derive insights from data and answer pertinent business questions	• Stored functions and procedures in SQL • Conditional constructs in SQL • Query Optimisation
6 **DATA ANALYSIS IN EXCEL** (OPTIONAL) Taught by one of the most renowned data scientists in the country (S.Anand, CEO, Gramener), this module takes you from a beginner level Excel user to an almost professional user.	• Introduction to Excel • Data Analysis in Excel - I: Functions, Formulae, and Charts • Data Analysis in Excel - II: Pivots and Lookups
7 **ANALYTICS PROBLEM SOLVING** (OPTIONAL) This module covers concepts of the CRISP-DM framework for business problem-solving.	• The CRISP-DM Framework- Business and Data Understanding • CRISP-DM Framework - Data Preparation, Modelling, Evaluation and Deployment

*Syllabus is subject to change by the university

STATISTICS AND EDA

Module/Assignment

Session

1 EXPLORATORY DATA ANALYSIS

Learn how to find and analyse the patterns in the data to draw actionable insights.

- Data Sourcing
- Data Cleaning
- Univariate Analysis
- Segmented Univariate
- Bivariate Analysis
- Derived Metrics

2 INVESTMENT ASSIGNMENT

Learners will fill in the shoes of an analyst at an investment bank and determine where the firm should invest. They will then have to explain their recommendations in lieu of the analysis conducted.

- Problem Statement
- Evaluation Rubric
- Final Submission
- Solution

3 INFERENTIAL STATISTICS

Build a strong statistical foundation and learn how to 'infer' insights from a huge population using a small sample.

- Basics of Probability
- Discrete Probability Distributions
- Continuous Probability Distributions
- Central Limit Theorem

4 HYPOTHESIS TESTING

Understand how to formulate and validate hypotheses for a population to solve real-life business problems.

- Concepts of Hypothesis Testing - I: Null and Alternate Hypothesis, Making a Decision, and Critical Value Method
- Concepts of Hypothesis Testing - II: p-Value Method and Types of Errors
- Industry Demonstration of Hypothesis Testing: Two-Sample Mean and Proportion Test, A/B Testing

5 LENDING CLUB CASE STUDY

Determine which customers are at the risk of default and what are their charecteristics so as to avoid providing loans to similar people in the future.

- Problem Statement
- Evaluation Rubric
- Final Submission
- Solution

*Syllabus is subject to change by the university

MACHINE LEARNING - I

Module/Assignment	Session
1 LINEAR REGRESSION Venture into the machine learning community by learning how one variable can be predicted using several other variables through a housing dataset where you will predict the prices of houses based on various factors.	• Introduction to Simple Linear Regression • Simple Linear Regression in Python • Multiple Linear Regression • Multiple Linear Regression in Python • Industry Relevance of Linear Regression
2 LINEAR REGRESSION ASSIGNMENT Build a model to understand the factors car prices vary on and help a Chinese company enter the US car market.	• Problem Statement • Evaluation Rubric • Final Submission • Solution
3 LOGISTIC REGRESSION Learn your first binary classification technique by determining which customers of a telecom operator are likely to churn versus who are not to help the business retain customers.	• Univariate Logistic Regression • Multivariate Logistic Regression - Model Building • Multivariate Logistic Regression - Model Evaluation • Logistic Regression - Industry Applications - Part I • Logistic Regression - Industry Applications - Part II
4 NAIVE BAYES Understand the basic building blocks of Naive Bayes and learn how to build an SMS Spam Ham Classifier using Naive Bayes technique.	• Bayes Theorem and Its Building Blocks • Naive Bayes For Categorical Data • Naive Bayes for Text Classification
5 MODEL SELECTION Learn the pros and cons of simple and complex models and the different methods for quantifying model complexity, along with regularisation and cross validation.	• Principles of Model Selection • Model Evaluation

*Syllabus is subject to change by the university

MACHINE LEARNING - II

Module/Assignment	Session
1 ADVANCED REGRESSION Understand generalised regression and different feature selection techniques along with the perils of overfitting and how it can be countered using regularisation.	• Generalized Linear Regression • Regularized Regression
2 ADVANCED REGRESSION ASSIGNMENT Build a model to understand the factors house prices vary on and help an American company enter the Australian housing market.	• Problem Statement • Evaluation Rubric • Final Submission • Solution
3 SUPPORT VECTOR MACHINE (OPTIONAL) Learn how to find a maximal marginal classifier using SVM, and use them to detect spam emails, recognise alphabets and more!	• SVM - Maximal Margin Classifier • SVM - Soft Margin Classifier • Kernels
4 TREE MODELS Learn how the human decision-making process can be replicated using a decision tree and other powerful ensemble algorithms.	• Introduction to Decision Trees • Algorithms for Decision Tree • Construction • Truncation and Pruning • Random Forests
5 MODEL SELECTION - PRACTICAL CONSIDERATIONS Given a business problem, how do you choose the best algorithm? Learn a few practical tips for doing this here.	• Model Selection - Best Practices

*Syllabus is subject to change by the university

MACHINE LEARNING - II

Module/Assignment	Session
6 BOOSTING Learn how weak learners can be 'boosted' with the help of each other and become strong learners using different boosting algorithms such as Adaboost, GBM, and XGBoost.	• Introduction to Boosting and AdaBoost • Gradient Boosting
7 UNSUPERVISED LEARNING: CLUSTERING Learn how to group elements into different clusters when you don't have any pre-defined labels to segregate them through K-means clustering, hierarchical clustering, and more.	• Introduction to Clustering • K Means Clustering • Executing K Means in Python • Hierarchical Clustering • Other Forms of Clustering
8 UNSUPERVISED LEARNING: PRINCIPAL COMPONENT ANALYSIS Understand important concepts related to dimensionality reduction, the basic idea and the learning algorithm of PCA, and its practical applications on supervised and unsupervised problems.	• Principal Component Analysis • PCA in Python
9 TELECOM CHURN CASE STUDY Solve the most crucial business problem for a leading telecom operator in India and southeast Asia - predicting customer churn.	• Problem Statement • Evaluation Rubric • Final Submission • Solution

*Syllabus is subject to change by the university

NATURAL LANUAGE PROCESSING

Module/Assignment	Session
1 LEXICAL PROCESSING Do you get annoyed by the constant spams in your mail box? Wouldn't it be nice if we had a program to check your spellings? In this module learn how to build a spell checker & spam detector using techniques like phonetic hashing, bag-of-words, TF-IDF, etc.	• Introduction to NLP • Basic Lexical Processing • Advanced Lexical Processing
2 SYNTACTIC PROCESSING Learn how to analyse the syntax or the grammatical structure of sentences with the help of algorithms & techniques like HMMs, Viterbi Algorithm, Named Entity Recognition (NER), etc.	• Introduction to Syntactic Processing • Parsing • Information Extraction • Conditional Random Fields
3 SYNTACTIC PROCESSING -ASSIGNMENT Build a POS tagger for tagging unknown words using HMMs and modified Viterbi algorithm.	• Problem Statement • Evaluation Rubric • Final Submission • Solution
4 SEMANTIC PROCESSING Learn the most interesting area in the field of NLP and understand different techniques like word-embeddings,LSA, topic modelling to build an application that extracts opinions about socially relevant issues (such as demonetisation) on social media platforms.	• Introduction to Semantic Processing • Distributional Semantics • Topic Modelling • Social Media Opinion Mining • Semantic Processing Case Study
5 BUILDING CHATBOTS WITH RASA Imagine if you could make restaurant booking without opening Zomato. Build your own restaurant-search chatbot with the help of RASA - an open source framework and deploy it on Slack.	• Problem Statement • Evaluation Rubric • Final Submission • Solution

*Syllabus is subject to change by the university

DEEP LEARNING

Module/ Assignment	Session
1 INTRODUCTION TO NEURAL NETWORKS Learn the most sophisticated and cutting-edge technique in machine learning - Artificial Neural Networks or ANNs.	• Structure of Neural Networks • Feed Forward in Neural Networks • Backpropagation in Neural Networks • Modifications to Neural Networks • Hyperparameter Tuning in Neural Networks
2 CONVOLUTIONAL NEURAL NETWORKS -INDUSTRY APPLICATIONS Learn the basics of CNN and OpenCV and apply it to Computer Vision tasks like detecting anomalies in chest X-Ray scans, vehicle detection to count and categorise them to help the government ascertain the width and strength of the road.	• Building CNNs with Python and Keras • CNN Architectures and Transfer Learing • Style Transfer and Object Detection • Industry Demo:Using CNNs with Flowers Images • Industry Demo:Using CNNs with X-ray Images
3 NEURAL NETWORKS - ASSIGNMENT Build a neural network from scratch in Tensorflow to identify the type of skin cancer from images.	• Problem Statement • Evaluation Rubric • Final Submission • Solution
4 RECURRENT NEURAL NETWORKS Ever wondered what goes behind machine translation, sentiment analysis, speech recognition etc.? Learn how RNN helps in these areas having sequential data like text, speech, videos, etc.	• What Makes a Neural Network Recurrent? • Variants of RNNs • Building RNNs in Python
5 NEURAL NETWORKS PROJECT - GESTURE RECOGNITION Make a Smart TV system which can control the TV with the user's hand gestures as the remote control.	• Problem Statement • Evaluation Rubric • Final Submission • Solution

*Syllabus is subject to change by the university

REINFORCEMENT LEARNING

Module/ Assignment	Session
1 CLASSICAL REINFORCEMENT LEARNING Ever wondered how Alpha Go beat the best GO player or how Boston Dynamics made robots that can run. Start your journey with the classical RL algorithms like dynamic programming, Monte Carlo methods, Q Learning to train the state value and action value functions of the policy.	• Markov Decision Process • Fundamental Equations in RL • Model-Based Method - Dynamic Programming • Model-Free Methods • Inventory Management Demo
2 ASSIGNMENT -CLASSICAL REINFORCEMENT LEARNING Train an agent that'll beat you in the game of numerical tic-tac-toe everytime you play.	• Problem Statement • Evaluation Rubric • Final Submission • Solution
3 DEEP REINFORCEMENT LEARNING Want to build your own Atari Game? Learn the Q-function or policy using the various Deep Reinforcement Learning algorithms: Deep Q Learning, Policy Gradient Methods, Actor - Critic method.	• Architectures of Deep Q Learning • Deep Q Learning • Policy Gradient Methods • Actor-Critic Methods
4 REINFORCEMENT LEARNING PROJECT Improve the recommendation of the the rides to the cab drivers by creating a RL based algorithm using vanilla Deep Q-Learning (DQN) to maximize the driver's profits and inturn help in retention of the driver on the cabaggregator service.	• Problem Statement • Evaluation Rubric • Final Submission • Solution

*Syllabus is subject to change by the university

CAPSTONE

Module/ Assignment

Session

1 **DEPLOYMENT**

Learn how to productionize your model and deploy it on the server.

- Introduction to machine learning pipleline
- Intro to Git and github
- Deploy machine learning models in production as API
- Deploying machine learning models on heroku using flask
- Deploying ML models on AWS
- Deploying machine learning models on flask and docker.
- Deployment on the edge

2 **CAPSTONE**

Choose from a range of real-world industry woven projects on advanced topics like Recommendation Systems, Fraud Detection, Emotion Detection from faces, Social Media Listening, Speech Recognition among many others.

- Problem Statement
- Evaluation Rubric
- Final Submission
- Solution

*Syllabus is subject to change by the university

RESEARCH METHODOLOGY

1 WHAT IS RESEARCH?

Familiarise yourself with different aspects of research

- Introduction to research
- Importance of research
- Criticism in research and its importance
- Peer reviews in research and its importance

2 TYPES OF RESEARCH

Develop an understanding of various research designs and techniques

- Descriptive vs Analytical
- Applied vs Fundamental
- Quantitative vs Qualitative
- Bayesian vs Frequentist Approach

3 RESEARCH PROCESS

Learn about the different steps in the research process and how to evaluate a literature review

- Research question
- Hypothesis and aims
- Formulating a Problem
- Literature review

4 RESEARCH PROJECT MANAGEMENT

Learn how to plan the project timelines and arrange for data & software

- Understand the different steps involved in a project cycle
- Project Requirements on Data
- Identifying the milestones in a research project
- Learn how to track the progress using Gantt Chart

5 REPORT WRITING AND PRESENTATION

Master good scientific writing and proper presentation skills

- Art of writing papers
- Parts of a paper
- Tools to write papers
- Publishing papers: Journals + Seminars

6 SCIENTIFIC ETHICS

Develop an understanding of the ethical dimension in research

- Citation Methods and Rules
- Honor Code
- Research Claims

*Syllabus is subject to change by the university

MASTER'S DISSERTATION

1

MASTER'S THESIS

Research, articulate & present your project.

- Monthly Checkpoints
- Submission

Examples of Project Outlines:

- Investigate the risk factors for eye disease from complex longitudinal datasets
- Investigate a diagnosis of eye diseases using imaging ophthalmic data
- Multi-task learning for drug design and discovery
- Using stacking for brain tumour discrimination
- Investigate dietary patterns and metabolite fingerprints of takeaway (fast) food consumers using PCA and Clustering methods
- Longitudinal studies to investigate the complex link between corporate environment engagement, green disclosure, business model transformation and supply chain performance
- Preventing credit card fraud through pattern recognition
- Developing a recommender system for a Media giant
- Using social media feed to place tweets regarding natural disasters on a map

*Syllabus is subject to change by the university