

Data Track Syllabus

Prerequisites: Before entering into the data track you should be good at statistics, probability, calculus, linear algebra and fundamental programming skills.

The Data track course is made for the people who are curious in solving complex problems, handling the data, developing algorithms, analyzing and visualizing large amount of data.

Python:

- Introduction to Python, Packages in Python, Syntax, Special Characters, Python indentation, Keywords, Identifiers, Variables, Constants, Literals, Control flow, Data Types, Functions and common data analysis libraries like Numpy and Pandas.

Exercise: Program to display current date and time, Loops, Functions, List, Tuple, Turtle.

Data Analysis and Data Science:

- Tool: Anaconda tool, Installing Anaconda, Managing Packages and Backup environment using Anaconda.
- Data Science: Statistics, Probability, Distribution, Vectors, Calculus, Data types and Data formats, Preparing Data using Python, Data Wrangling, Data Analysis, Data Visualization and MapReduce.

Exercise: Import data from Online and Offline, Handling Missing data, Plot the Residuals, Data Analysis using python.

Artificial Intelligence:

- Introduction: Intro to Artificial Intelligence, Agent, Environment, State, Perception, Action, Cognition, Relational behaviour and Bounded optimality.
- Search: Creating the search tree, Depth first search, Uniform cost search, Local Beam search, Genetic Algorithm, Map Coloring, Constraint graph and Hypergraph, Backtracking Search, Iterative Algorithm, Proposition Logic, Limitations of propositional Logic, Truth Table and Symbols, Problem solving Vs Planning, Infinite sequence.

- Planning: Problem solving Vs Mathematics, Classical Planning, Progression Planning, Regression Planning, Bayesian Network, Naive bayes Classifier, Hidden Markov models.

Exercise: DFS, Genetic Algorithm, UCS, LBS, Map Coloring, Backtracking search, Iterative Algorithm, Proposition Logic, Planning, Bayes Rule, Pattern Recognition, Sign Language Recognition.

Machine Learning:

- Intro to Machine Learning, Model Evaluation and Validation, Supervised Learning, Unsupervised Learning and Reinforcement Learning.

Exercise: Classification, Regression, Bayesian , Naive Bayes, Decision Tree, Random forest, Clustering, Markov Decision, Q-Learning, RL models.

Deep Learning:

- Neural Network: Regression, Gradient and Stochastic gradient descent, Perceptron and Backpropagation.
- Convolutional Neural Network: CNN Architecture, Classification and Regression, Confusion Matrix, Types of errors, K- fold cross validation, Forward Propagation, Sigmoid function, Gradient and Stochastic gradient descent, Cost function, Loss function, Linear Transform, Tensorflow, On Keras, Weight Initialization, ReLU and Softmax.
- Recurrent Neural Network: RNN Intro, LSTM, Hyperparameters, Learning Rate, Word2Vec, Sentiment Prediction RNN, Reinforcement Learning and Tensorboard.
- Generative Adversarial Network: Architecture of GAN, Autoencoding, Batch Normalization, DCGAN, Deep Neural Network and One - shot learning.

Exercise: Regression, Sigmoid Function, Error Minimization, Confusion Matrix, Cost Function, Loss Function, Weight Initialization, Softmax, ReLU, LSTM, Hyperparameters, Word2Vec, Sentiment Prediction, Tensorboard, Batch Normalization.

Project 1: What's Cooking?

Cooking is an art. Since everything has become globalized, cooking also getting flourished. The trademark cooking of each country must be made familiar to all the people. In order to make every country's food familiar, we ask you to predict the category of a dish's cuisine given a list of its ingredients.

Project 2: Analysis of Kissan call center data using NLP

Call centers around the world have a massive percentage of customer support inquiries. They have too much data to explore. In this challenge, we ask you to analyse the Kissan call center data using NLP.

Project 3: Recommendation of Tamil movies

Recommender systems are one of the most popular applications of data science today. In this challenge, we ask you to create a recommender system for Tamil movie dataset and tell users which movies to watch.

Project 4: Image classification challenge

In this project, we ask you to implement an image classification application using a deep learning model for the image dataset.

Length of the Program: 3 months

Educational Objectives: At the end of this program, you will be good at

- Problem solving
- Python programming
- Data cleaning and processing
- Statistical Analysis
- Data Visualization
- Algorithm development
- Building predictive models