

# Data Science Certification Program

## 1. Data Science Foundation:

- Python for Data Analysis: Get acquainted with Data structures, Object Oriented Programming, Data Manipulation and Data Visualization in Python.
- Mathematics for Machine learning:
  - Linear Algebra
  - Multivariate calculus
  - Probability theory and Probability Distributions
  - Matrices, Eigen Vectors and their application for Data Analysis.
- Computer Science & Algorithms: Matrix implementation/ Data Structures & Algorithms
- Introduction to Cloud Computing and AWS
- Functions and Graphs
- Statistics and it's applications
- Hypothesis Testing

## 2. Getting Started with Data Science/Machine Learning

- What is Data Science – Examples and Applications
- Numpy and Pandas Tutorial
- Scikit Learn Tutorial
- Machine Learning Algorithms
- Cost Function
- Metrics for Model Evaluation and Validation
- Training and Testing
- Model Overfitting - Underfitting
- Bias & Variance
- Gradient Descent Optimization & Learning Rate
- Bias & Variance
- Hyper-parameters Tuning & Model Optimization
- S2 Mini-Projects to understand and implement Machine Learning Basics

## 3. Data Exploration - Data Preprocessing & Feature Engineering

- Data Extraction, Transformation, and Loading
- Data Wrangling and Data Exploration.
- Data Pre-processing
- Data Visualization
- Feature Selection
- Feature Transformations
- Outlier Detection and Handling
- Handling Missing Values

## 4. Supervised Learning

- Introduction to Supervised Learning
- Linear Regression
- Logistic Regression
- Decision Trees
- Random Forests
- Naïve Bayes Classifier
- Bayesian Statistics and Inference
- K-Nearest Neighbor
- Support Vector Machine – SVM
- One mini project hands-on for each algorithm

## 5. Unsupervised Learning

- Introduction to Unsupervised Learning
- K-Means Clustering
- Agglomerative Hierarchical Clustering
- Clustering using DBSCAN
- Expectation–Maximization (EM) Clustering using Gaussian Mixture Models (GMM)
- Clustering Mini-Project

## 6. Amazon Web Services

- IAM/S3/EC2
- Machine Learning on AWS Machines

## 7. Deep Learning

- Introduction to Deep Learning
- Machine Learning VS Deep Learning
- Introduction to Neural Networks
- TensorFlow/Theano/Keras
- Deep Neural networks
- Forward propagation
- Back Propagation Learning
- Activation Functions
- Loss Functions
- Hyperparameter Tuning
- DropOut Regularization
- Batch Normalization
- Early Stopping
- Vanishing & Exploding Gradient Descent
- Optimization Algorithms
  - Gradient descent with momentum
  - RMS Prop
  - Adam

## 8. Computer Vision

- Introduction to Computer Vision
- Convolutional Neural Networks CNN
- Residual Networks- ResNets
- Long Short Term Memory LSTM
- Image identification
- Object detection and Localization
- YOLO Algorithm
- Introduction to DarkNet Project
- Case Studies
  - Le-Net5
  - AlexNet
  - VGG16
  - Google Net ( Inception Networks)
- Transfer learning (Resnet, ResNext, Xception, EffecientNet, DenseNet etc.)
- Data Augmentation - Keras Generators
- Data Augmentation- OpenCv
- Ensembling, Concatenation of models
- Stacking of Models
- Keras Functional API
- Face Recognition
- Siamese Networks
- Kaggle Project on Computer Vision

## 9. Natural Language Processing NLP

- Introduction to Natural language Processing
- Application of NLP
- NLTK
- Tokenization
- Stemming
- Lemmatization
- Stop Words
- Similarity Functions - TFIDF
- Word Representations
- One Hot Encoding
- Word Embedding
  - Learning word Embedding
    - - Word2Vec
    - - SkipGram , N-Gram
    - - Negative Sampling
    - - GloVe
- Beam Search
- Recurrent Neural Networks- RNN
- BiDirectional RNN
- Gated Recurrent Unit GRU

- Long Short Term Memory LSTM
- Deep RNNs
- Kaggle Project on NLP

## 10. Dimensionality Reduction

- PCA
- LDA
- Kernel PCA
- SMOTE

## 11. Ensemble, Bagging & Boosting

- k-fold Cross Validation
- Grid Search
- Bagging & Boosting
- ADA boost
- XGBoost
- Light GBM
- Ensembling Techniques
- Stacking

## 12. Major projects covered during the course

- All ML/DL algorithms will be covered with hands-on mini projects
- 15-20 mini projects
- Deep Learning Projects:-
  - Sentiment Analysis
  - Handwritten Digits recognition
  - Chatbot from Scratch
  - Named Entity Recognition using Deep Learning Algorithms
- **Introduction to Kaggle Platform and other Data Science Competitions**

## 13. 8 weeks E2E Industry Project

### Challenge for Industry professionals:

- How to align current experience with Data Science?
- How to crack Data Science job interviews?
- What is the surety that after Data Science training, I would be able to get a better job offer?

### Solution:

- You have to complete a domain-specific project to solve a real-world problem using different Machine Learning/Deep Learning.

- After Data Science course completion, student need to identify and business problem with help of Mentors at IT Bodhi and implement solution end to end using ML.
- You need to present this project in Data Science interview as white paper to showcase your in depth knowledge of giving ML/DS solutions to customer problems.
- This project will help in aligning current domain experience to machine learning knowledge and would pave the way to crack ML job interviews.
- Time duration to complete the project: Depends on the complexity of the problem but generally it takes 2-3 month approx.

## 14. Interviews Preparation

### Acing Data Science interviews

- -Booklet of all ML interview questions
- -Showcasing and presenting ML projects in interviews?
- -Presenting E2E Industry project in interviews?
- -How to align current industry experience with Data Science learning?
- -How to make a big Impact with Domain knowledge + Data Science learning?
- -Handling E2E business problem solving questions
- -Handling Project management questions data sciences
- -Do's-Dont's in interviews
- -Interviews preparation

### Resume Preparation

- -How to make a impressive resume?
- -Mention right ML Projects in the resume
- -A good & a Bad resume