GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

Postgraduate Diploma
in
Machine Learning and Artificial Intelligence

First Semester								
Subject Code	**Subject Name**	**Subject Type (Theory/ Practical)**	**Load Allocation per Week**		**Marks Distribution**		**Total Marks**	**Credits**
			L	**P**	**Int.**	**Ext.**		
PGCS-101	Statistical Learning	Theory	3	0	40	60	100	3
PGCS-102	Applied Artificial Intelligence	Theory	3	0	40	60	100	3
PGCS-103	Python Programming for ML and AI	Theory	3	0	40	60	100	3
LPGCS-101	Statistical Learning Laboratory	Practical	0	4	30	20	50	2
LPGCS-102	Applied Artificial Intelligence Laboratory	Practical	0	4	30	20	50	2
LPGCS-103	Python Programming for ML and AI Laboratory	Practical	0	4	30	20	50	2
PGPRCS-101	Project-I*	Practical	0	8	60	40	100	8
		Total	**9**	**20**	**270**	**280**	**550**	**23**

* Project-I will be based on industrial oriented problem.

Note: Each student has to undergo Four (4) week industrial training at the end of 1st Semester. The evaluation and the credit of the industrial training will be included in 2nd Semester. The training will be purely Industry-oriented only.

Second Semester								
Subject Code	**Subject Name**	**Subject Type (Theory/ Practical)**	**Load Allocation per Week**		**Marks Distribution**		**Total Marks**	**Credits**
			L	**P**	**Int.**	**Ext.**		
PGCS-104	Machine Learning	Theory	3	0	40	60	100	3
PGCS-105	Deep Learning	Theory	3	0	40	60	100	3
LPGCS-104	Machine Learning Laboratory	Practical	0	4	30	20	50	2
LPGCS-105	Deep Learning Laboratory	Practical	0	4	30	20	50	2
LPGCS-106	AI Compute Platforms Laboratory	Practical	0	4	30	20	50	2
PGPRCS-102	Project-II*	Practical	0	8	120	80	200	8
PGTRCS-101	Industrial Training**	Practical (Training)	0	0	60	40	100	4
		Total	**6**	**20**	**350**	**300**	**650**	**24**

*The project-II will be based on industrial oriented problem.

** The marks of Industrial Training undergone at the end of 1st semester will be included in the 2nd semester. The evaluation scheme of industrial training shall be as under:

Internal: Out of 60 marks, 20 marks will be given on the basis of industrial visit made by student during 1st semester. The students have to visit the selected industry five times in the semester in such a way that in each month he/she has to make at least one visit in the same industry. The student has to maintain a diary a day for recording the report of each vist. 40 marks shall be given on the basis of his/her performance.

External: External examiner should be essentially from industry and will evaluate the student on the basis of oral viva for 40 marks

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

PGCS-101 Statistical Learning

Introduction to Statistical Learning: Definition, Need to Estimate function(f), How do we estimate function, The Trade-Off Between Prediction Accuracy and Model Interpretability , Supervised Versus Unsupervised Learning, Regression Versus Classification Problems. Assessing Model Accuracy: Measuring the Quality of Fit, The Bias-Variance Trade-Off, The Classification Setting. [6 Hours]

Linear Regression: Simple Linear Regression - Estimating the Coefficients, Assessing the Accuracy of the Coefficient, Estimates, Assessing the Accuracy of the Model. Multiple Linear Regression - Estimating the Regression Coefficients. Other considerations in the Regression Model - Qualitative Predictors, Extensions of the Linear Model, Potential Problems. [6 Hours]

Classification: Introduction, Why Not Linear Regression? Logistic Regression: The Logistic Model, Estimating the Regression Coefficients, Making Predictions, Multiple Logistic Regression, Logistic Regression for >2 Response Classes. Linear Discriminant Analysis - Using Bayes' Theorem for Classification, Linear Discriminant Analysis for p = 1, Linear Discriminant Analysis for p >1, Quadratic Discriminant Analysis. [8 Hours]

Resampling Methods: Cross-Validation: The Validation Set Approach, Leave-One-Out Cross-Validation, k-Fold Cross-Validation, Bias-Variance Trade-Off for k-Fold Cross-Validation, Cross-Validation on Classification Problems. The Bootstrap. [5 Hours]

Linear Model Selection and Regularization: Subset Selection: Best Subset Selection, Stepwise Selection, Choosing the Optimal Model. Shrinkage Methods: Ridge Regression, The Lasso, Selecting the Tuning Parameter. Dimension Reduction Methods: Principal Components Regression, Partial Least Squares. [6 Hours]

Tree-Based Methods: The Basics of Decision Trees: Regression Trees, Classification Trees, Trees Versus Linear Models, Advantages and Disadvantages of Trees. Bagging, Random Forests, Boosting. [5 Hours]

Text Books

1. Gareth James, Daniela Witten, Trevor Hastie Robert Tibshirani, "An Introduction to Statistical Learning: with Applications in R".Springer
2. Pratap Dangeti, "Statistics for Machine Learning", Packt Publishing

Reference Books

1. Masashi Sugiyama, “Introduction to Statistical Machine Learning”, Morgan Kaufmann.
2. Peter Bruce, Andrew Bruce, “Practical Statistics for Data Scientists”, O'Reilly Media.

E-Books and online learning material

1. Gianluca Bontempi, “Statistical Foundations of Machine Learning”, Universite Libre de Bruxelles ,https://www.researchgate.net/publication/242692234_Handbook_on_Statistical_foundations_of_machine_learning.
2. Gareth James, Daniela Witten, Trevor Hastie Robert Tibshirani, “An Introduction to Statistical Learning: with Applications in R”.Springer, https://faculty.marshall.usc.edu/gareth-james/ISL/ISLR%20Seventh%20Printing.pdf.
3. Pratap Dangeti, “Statistics for Machine Learning”, Packt Publishing, https://www.pdfdrive.com/statistics-for-machine-learning-techniques-for-exploring-supervised-unsupervised-and-reinforcement-learning-models-with-python-and-r-e158478475.html.

Online Courses and Video Lectures

1. https://www.youtube.com/watch?v=3jQs02dbfrI
2. https://www.youtube.com/watch?v=MMN35r-zbQM3
3. https://www.youtube.com/watch?v=r-yv6GbWep4
4. https://www.youtube.com/watch?v=fQwWn4O19uw
5. https://www.youtube.com/watch?v=7frypPGKtu8
6. https://www.youtube.com/watch?v=YYGTK2vMoCQ

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

PGCS-102 Applied Artificial Intelligence

Introduction to AI: Introduction to artificial intelligence, History of AI, Proposing and evaluating AI applications. [2 Hours]

Search and Planning: Problem spaces and search, Knowledge and rationality, Heuristic search strategies, Search and optimization, Adversarial search, Planning and scheduling. [4 Hours]

Knowledge Representation and Reasoning: Logic and inference, Ontologies, Bayesian reasoning, Temporal reasoning. [4 Hours]

Machine learning: What is machine learning?, Supervised vs. unsupervised learning, Supervised methods: Regression - linear, logistic, ridge, Classification – decision trees, SVM, random forests, Model performance evaluation, Unsupervised Methods: Dimensionality reduction: PCA, Clustering – k-means, hierarchical clustering, Semi-supervised methods: Reinforcement learning, Choosing among machine learning techniques. [10 Hours]

Deep Learning: Neural networks and back-propagation, Convolutional neural networks, Recurrent neural networks and LSTMs. [5 Hours]

Image Processing: Introduction to computer vision, Image segmentation, Object and motion detection, Object classification, Use of pre-trained models (Inception). [4 Hours]

Natural Language Understanding and Interaction: Intro to natural language understanding, Sentiment analysis, Speech recognition, Hidden Markov Models, Chatbots, Natural language generation, Speech synthesis. [5 Hours]

Robotic Sensing and Manipulation: Introduction to robotics, Sensing, Manipulation, Human-robot interaction. [2 Hours]

Ethical and Legal Considerations in AI: Privacy, Bias, AI and the future of work, Appropriate uses of AI. [2 Hours]

Text Books

1. Stuart Russell and Peter Norvig, “Artificial Intelligence: A Modern Approach”, Prentice Hall.

2. Saroj Kaushik, “Artificial Intelligence”, Cengage Learning India.

Reference Books

1. Elaine Rich and Kevin Knight, “Artificial Intelligence”, Tata McGraw Hill.
2. Trivedi, M.C., “A Classical Approach to Artifical Intelligence”, Khanna Publishing House, Delhi.
3. David Poole and Alan Mackworth, “Artificial Intelligence: Foundations for Computational Agents”, Cambridge University.

E-Books and online learning material

1. Handbook of Artificial Intelligence Edited by Avron Barr and Edward A. Feigenbaum, Computer Science Department, Stanford University.
2. https://stacks.stanford.edu/file/druid:qn160ck3308/qn160ck3308.pdf

Online Courses and Video Lectures

1. https://www.coursera.org/courses?query=artificial%20intelligence Accessed on August, 24, 2020
2. https://nptel.ac.in/courses/106/105/106105077/ Accessed on August, 24, 2020
3. https://nptel.ac.in/courses/106/102/106102220/ Accessed on August, 24, 2020
4. https://www.youtube.com/watch?v=bV4t4r3SGuI Accessed on August, 24, 2020
5. https://www.youtube.com/watch?v=iF1tOCEXLXY Accessed on May, 20, 2020

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

PGCS-103 Python Programming for ML and AI

Introduction: Introduction to Artificial Intelligence, Machine Learning and Deep Learning, History of Artificial Intelligence, Difference between Artificial Intelligence, Machine Learning and Deep Learning, Applications of Machine Learning. [4 hours]

Python Basics: Why Python Programming, Data Types and Operators, Control Flow, Functions, Scripting, Classes, Python setup - Anaconda, Jupyter Notebook, NumPy, Pandas. [6 hours]

Data Handing in Python: Working with NumPy and Pandas, Statistics - Basics of Statistics, Central Tendency, Covariance, Correlation, Standard Deviation, Z-Score, Bayes' Theorem, Probability - Joint Probability, Conditional Probability, Distribution of Data, Linear Algebra – Introduction, Vectors, Linear Combination, Linear Transformation and Metrics, Calculus Essentials – Introduction, Derivatives Through Geometry, Chain Rule and Dot Product, Limits and Integration, Information Entropy. [10 hours]

Data Visualization: Introduction to Visualization tools – Matplotlib, Seaborn, Area Plot, Histogram, Bar Charts, Pie Charts, Box Plots, Scatter Plots, Waffle Charts, Word Clouds, Seaborn and Regression Plots, Creating maps and Geo spatial data(Folium, Maps with Markers) [6 hours]

Machine Learning: Supervised Learning, Unsupervised Learning, and Reinforcement learning. [3 hours]

Regression, Classification and Clustering: Simple linear regression, Multiple linear regression, Model evaluation, Evaluation metrics, Nonlinear regression, Ridge regression. K-nearest neighbors, Evaluation metrics, Decision tree, Random Forest Algorithm, Naïve Bayes Algorithm, Logistic regression, Logistic vs linear regression, Support vector machine, K- Means, Hierarchical Clustering, DBSCAN Clustering**.** [9 hours]

Recommendation System: Recommender systems, Content based, Collaborative filtering. [2 hours]

Text Books:

1. Philip C. Jackson Jr., "Introduction to Artificial Intelligence" Dover Publications
2. Stuart Russell, "Artificial Intelligence: A Modern Approach", Pearson

Reference Books:

1. Anirban DasGupta, "Probability for Statistics and Machine Learning", Springer
2. Sebastian Raschka, "Python for Machine Learning", Packt Publishing
3. Andreas C. Müller, "Introduction to Machine Learning with Python: A Guide for Data Scientists", O'Reilly Media

Online Courses and Video Lectures

1. https://www.coursera.org/specializations/python? Accessed on July.28, 2020
2. https://www.coursera.org/learn/machine-learning-with-python Accessed on July.28, 2020
3. https://nptel.ac.in/courses/106/102/106102220/# Accessed on July. 28, 2020

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

LPGCS-101 Statistical Learning Laboratory

1. Introduction to simple basic R commands.
2. Fit Simple Linear Regression model on some suitable dataset for prediction.
3. Fit a multiple linear regression model using least squares on suitable dataset.
4. Fit a multiple linear regression model using least squares on suitable dataset that contains Qualitative Predictors.
5. Fit a logistic regression model using least squares on suitable dataset.
6. Fit a Linear Discriminant Analysis on suitable dataset.
7. Fit a Quadratic Discriminant Analysis model on suitable dataset.
8. Estimate the test error associated with fitting any Regression model using Leave-One-Out Cross-Validation on above used dataset.
9. Estimate the test error associated with fitting any Regression model using k-Fold Validation on above used dataset.
10. Apply Best Subset Selection method on some suitable dataset for prediction purpose.
11. Apply Ridge Regression and the Lasso on some suitable dataset for prediction purpose.
12. Apply Principal Components Regression on some suitable dataset for prediction purpose.

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

LPGCS-102 Applied Artificial Intelligence Laboratory

1. **Review of Important Python Concepts**
   Classes, String, Tuples, Lists, Dictionaries, sorting, handling exceptions, using iPython
2. **Machine Learning Algorithms with Scikit-learn**
   Pandas Library, Using Scikit-Learn for Logistic Regression, Support Vector Machines, Building Neural Networks,
3. **Introduction to Tensor Flow**
   Concept of Computational Graph and Nodes, Virtual Environment and Anaconda, Installing TensorFlow with GPU support on a Linux System, TF Data types, Placeholders, TF Variables, TF Session, Softmax, One Hot Encoding, Dropout, building hidden layers, Batching, Stochastic Gradient Descent, Building an Optimizer, Training and displaying results
4. **Building a Neural Network with Tensor Flow**
   Using inbuilt TensorFlow functionality to build a Neural Network and train on MNIST Dataset for classification
5. **Introduction to Image Processing**
   Overview, Smoothing, Image Morphology, Flood Fill, Resize, Image Pyramids, Thresholding operation, Image Transforms: Convolution, Gradients and Sobel Derivatives, Laplace, Canny, Stretch, Shrink, Warp, and Rotate
6. **Tracking and Motion**
   The Basics of Tracking, Corner Finding, Subpixel Corners, Invariant Features, Optical Flow, Mean-Shift & Camshift Tracking, Motion Templates, Estimators
7. Introduction to NLTK, Python 3 and the Jupyter Notebook.

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

LPGCS-103 Python Programming for ML and AI Laboratory

1. Getting started with Python -

a) Setting up an environment, any IDE (Anaconda)
b) Installation of Python/Code Editor and Jupyter Notebook
c) Working with basic concepts of Python – Functions, Modules and Built-in Data Structures.

2. Working with Data -

a) Working with Number, Strings, objects and Creating list
b) Implementation of Loops, execution of Iterating over sequence of objects, Arranging to pause execution, Generating random integers.

3. Working with Structured Data -

a) Working with Dictionary - insertion, deletion, updating, iterating over dictionary
b) Experiment with Ranges
c) Working with Random library for generating random data

4. Working with NumPy and Pandas -

a) Introduction to NumPy Array (1-D, 2-D and N-D)
b) Dot and cross product of arrays of N-d arrays
c) Working with data series, data frames and headers of data frames in Pandas
d) Implementation of statistical operations

5. Working with Data Visualization – Metplolib, Seaborn, Matplotlib, Seaborn, Area Plot, Histogram, Bar Charts, Pie Charts, Box Plots, Scatter Plots, Waffle Charts, Word Clouds, Seaborn and Regression Plots, Creating maps and Geo spatial data(Folium, Maps with Markers)

6. Working on regression problems

a) House price prediction
b) Car price prediction

7. Working with classification problems

a) Handwritten digit recognition
b) Binary Classification example Dog – Cat recognition

8. Working with clustering problems

a) Voice Clustering

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

PGCS-104 Machine Learning

Introduction: Well defined learning problems, Defining a learning system, perspectives and issues in machine learning, the concept learning task, concept learning as search, Find-S: finding a maximally specific hypothesis, version spaces and the candidate elimination algorithm, Inductive bias. [4 Hours]

Supervised Learning: Basic methods: Distance based methods, Nearest- Neighbors, Decision Trees, Naive Bayes, and Linear models: Linear regression, Logistic Regression, Generalized Linear Models, Support Vector Machines, Nonlinearity and kernel Methods. [10 Hours]

Unsupervised Learning: Clustering: k-means/ kernel k-means, Dimensionality Reduction: PCA and kernel PCA, Matrix Factorization and Matrix Completion, Generative models (mixture models and latent factor models). [5 Hours]

Decision Tree Learning: Introduction, Decision tree representation, appropriate problems for decision tree learning, the basic decision tree learning algorithm, hypothesis space search in decision tree learning, Inductive bias in decision tree learning, issues in decision tree learning. [4 Hours]

Artificial Neural Networks: Introduction, Neural network representation, appropriate problems for neural network learning, perceptrons, gradient descent and the delta rule, Adaline, Multilayer networks, Derivation of Back propagation rule, back propagation algorithm. [6 Hours]

Bayesian Learning: Introduction, Bayes theorem and concept learning, Maximum likelihood and least squared error hypothesis for predicting probabilities, minimum description length principle, Bayes optimal classifier, Naive bayes classifier, Bayesian belief networks. [5 Hours]

Genetic Algorithms: Motivation, Genetic algorithms, an illustrative example, hypothesis space search, genetic programming, models of evolution and learning. [6 Hours]

Text Books:

1. Tom M. Mitchell, Machine Learning, McGraw Hill, First Edition.

2. Ethern Alpaydin, Introduction to Machine Learning, MIT Press, 3rd Edition.

3. Aditya Dwivedi, Machine Learning Textbook, Kindle Edition, Dec 2019

Reference Books:

1. Chris Bishop, Pattern Recognition and Machine Learning, Springer.

2. Kevin Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2nd Edition

E-Books and Online Learning Material

1. Introduction to Machine Learning by Nils J. Nilsson

https://ai.stanford.edu/~nilsson/MLBOOK.pdf Accessed on: March 20,2020

2. Lecture Notes on Machine Learning by Sebastian Raschka

https://sebastianraschka.com/pdf/lecture-notes/stat479fs18/01_ml-overview_notes.pdf

Accessed on: March 20, 2020

3. https://www.tutorialspoint.com/machine_learning/machine_learning_tutorial.pdf

Accessed on: March 20,2020

Online Courses and Video Lectures:

1. https://nptel.ac.in/courses/106106139/ Accessed on: March 20,2020
2. https://nptel.ac.in/courses/106106213/ Accessed on: March 20,2020
3. **https://www.cs.ox.ac.uk/people/nando.defreitas/machinelearning/** Accessed on: March 20,2020

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

PGCS-105 Deep Learning

Introduction: Deep Learning definition, why Deep Learning, history of Deep Learning, Biological Neuron, Idea of computational units, McCulloch–Pitts unit and Thresholding logic, Linear Perceptron, Perceptron Learning Algorithm, Convergence theorem for Perceptron Learning Algorithm. [5 Hours]

Feedforward Networks: Multilayer Perceptron, Representation power of Feedforward Neural Networks, Backpropagation Gradient Descent, Empirical Risk Minimization, autoencoders. [4 Hours]

Deep Neural Networks: Difficulty of training deep neural networks, Greedy layerwise training, Newer optimization methods for neural networks (Adagrad, adadelta, rmsprop, adam, NAG), second order methods for training, Saddle point problem in neural networks, Regularization methods (dropout, drop connect, batch normalization). [5 Hours]

Recurrent Neural Networks: Back propagation through time, Long Short Term Memory, Gated Recurrent Units, Bidirectional LSTMs, Bidirectional RNNs. [5 Hours]

Convolutional Neural Networks: LeNet, AlexNet, ZF-Net, VGGNet, GoogLeNet, ResNet [4 Hours]

Generative models: Restrictive Boltzmann Machines (RBMs), Introduction to MCMC and Gibbs Sampling, gradient computations in RBMs, Deep Boltzmann Machines. [5 Hours]

Recent trends: Variational Autoencoders, Auto-encoders and unsupervised learning, Stacked auto-encoders and semi-supervised learning, transfer learning, multi-model learning, Generative Adversarial Networks, Multi-task Deep Learning, Multi-view Deep Learning. [5 Hours]

Applications: Vision, NLP, Speech, etc. [3 Hours]

Textbook

1. Deep Learning, Ian Goodfellow and Yoshua Bengio and Aaron Courville, MIT Press, 2016.

Reference Books:

1. Neural Networks: A Systematic Introduction, Raúl Rojas, 1996
2. Pattern Recognition and Machine Learning, Christopher Bishop, 2007

E-books and online learning material:

1. http://deeplearning.net/tutorial/deeplearning.pdf
2. http://neuralnetworksanddeeplearning.com/index.html
3. https://d2l.ai/d2l-en.pdf

Online lectures:

1. https://nptel.ac.in/courses/106/106/106106184/
2. https://nptel.ac.in/courses/108/105/108105103/
3. https://nptel.ac.in/courses/106/106/106106201/

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

LPGCS-104 Machine Learning Laboratory

List of Practicals:

1. Introduction to Machine Leaning Tools with Python language.

2. Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.

3. Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.

4. Write a program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set.

5. Write a program to implement k-Nearest Neighbour algorithm to classify a standard data set. Print both correct and wrong predictions.

6. Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same data set for clustering using k-Means algorithm. Compare the results of these two algorithms and comment on the quality of clustering.

7. Develop machine learning method for classifying i) the incoming mails. ii) how people rate the movies, books, etc.

8. Build an Artificial Neural Network by implementing the Backpropagation algorithm and test the same using appropriate data sets.

9. Implementation of Genetic Algorithm for optimization of decision variables in specific problems.

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

LPGCS-105 Deep Learning Laboratory

List of Practicals:

1. Introduction to Deep Learning platforms (such as H2O.ai, Dato GraphLab) and Deep Learning Software libraries (such as Tensorflow, Theano, Deeplearning4j, Torch, Caffe)
2. Introduction to tensor flow, setting up the Computational Graphs and MLPs, Linear Regression, Logistic Regression, Activation functions, Optimization, Regularization and Batch normalization
3. Setting up Artificial Neural Network: various classifiers, neural networks and genetic algorithm
4. Implementation of Convolutional Neural Network: Learning Objectives, Use Dataset, CNN classification
5. Implementation of Recurrent Neural Network: Learning Objectives, Use Dataset, Seq2seq model, RNN model, LSTM model, RNN for Language/Character Modelling with LSTMs/GRUs
6. Implementation of Generative model or Setting up Restricted Boltzmann Machine: Learning Objectives, Use Dataset, Collaborative filtering with RBM
7. Implementation of Autoencoder or Setting up Autoencoder: Use Autoencoder with RNN for language modelling
8. Setting up Deep Neural Network, train and effectively regularise it (for real time application)

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA
Department of Computer Science & Engineering

LPGCS-106 AI Compute Platforms Laboratory

List of Practicals:

1. Introduction to AI/Cognitive platforms.
2. Study of Google Cloud platform, AWS, AZURE
3. Installation and setup of PySpark on Hydoop environment.
4. Running a Function in Parallel with Python by using multiprocessing module.
5. Introduction and installation of PySpark.
6. Study of basic concepts of SparkContext and RDD (Resilient Distributed Dataset).
7. Implement basic operations with RDD & PySpark such as count(), collect(), foreach(f), filter(f), reduce(f).
8. Study of basic data structures and dimensions of TensorFlow.
9. Implementation of convolutional neural network (CNN) using TensorFlow.
10. Train and Build models (such as a fraud detection model) and making predictions with TensorFlow Estimator.