Training Details:

Training on Data Structure with C
Resource Organization: ARDENT
Training Date: **17th August 2020 to 31st August 2020**
Venue: **ONLINE MODE**
Student: 3rd Year IT (6th Semester)
Students Enrolled: **28**
Students Completed Successfully: **28**
Pass Out Year: **2021**
Feedback Analysis: **Attached**
Student List: **Attached**

Introduction:

Data Structures is a concept a means of storing a collection of data. Computer Science is a concern with study of methods for effectively using a computer to solve problems. These can be solved by algorithms and data structures. Data Structures tells you what way the data as to store in computer memory and how to access the data efficiently. Many Applications are designed by data structures stack applications like page visited history in a web-browser, chain of method calls in the Java virtual machine or C++ Run-time environment etc. Queue Application Like Waiting Lines, Multi-programming etc. For many applications the choice of proper data structure is the only major decision involving the implementation. Majorly the database designing and internal implementation is done only by using Data Structures techniques through C programming language.

Training Objective:

This Course main objective for the student to understand Analysis and Designing of the Algorithms and how the different data structures are used for efficient accessing of the data and Manipulation of the data at the end of the session we can able to know different Kinds of data structures and we can able to provide different algorithms for time and space complexity.

Training Outcome:

After completed the training student will able to

- ❖ Understand the concept of data structures and its relevance in computer science.
- ❖ Familiarize with selected linear and nonlinear data structures.
- ❖ Enhance skill in programming in C.

Training Methodology:

- Hands on practice approach to training, behavioral model of training would be practiced.
- During the training, the Trainee would implement a project related to respective modules.

- Commitment to Individual growth and constant evaluation.
- Implementation of programming techniques through a Project using C language.

Summary:

The following points can be noted from the program.

- ❖ At the beginning of the training trainer has clearly described the basic Introduction Data structure and c - programming skill and its application in industries in different areas.
- ❖ Students had done many programming by themselves during the trainings.
- ❖ During the training some students raised their queries and the trainer had explained all the quarries of the students.
- ❖ At the end of the training an online exam was conducted.
- ❖ As per the feedback received from the students' end, the entire session was really fruitful
- ❖ and enjoyable and the students have learned many things about C Programming skill.

Feedback analysis for the training:

FEEDBACK ANALYSIS
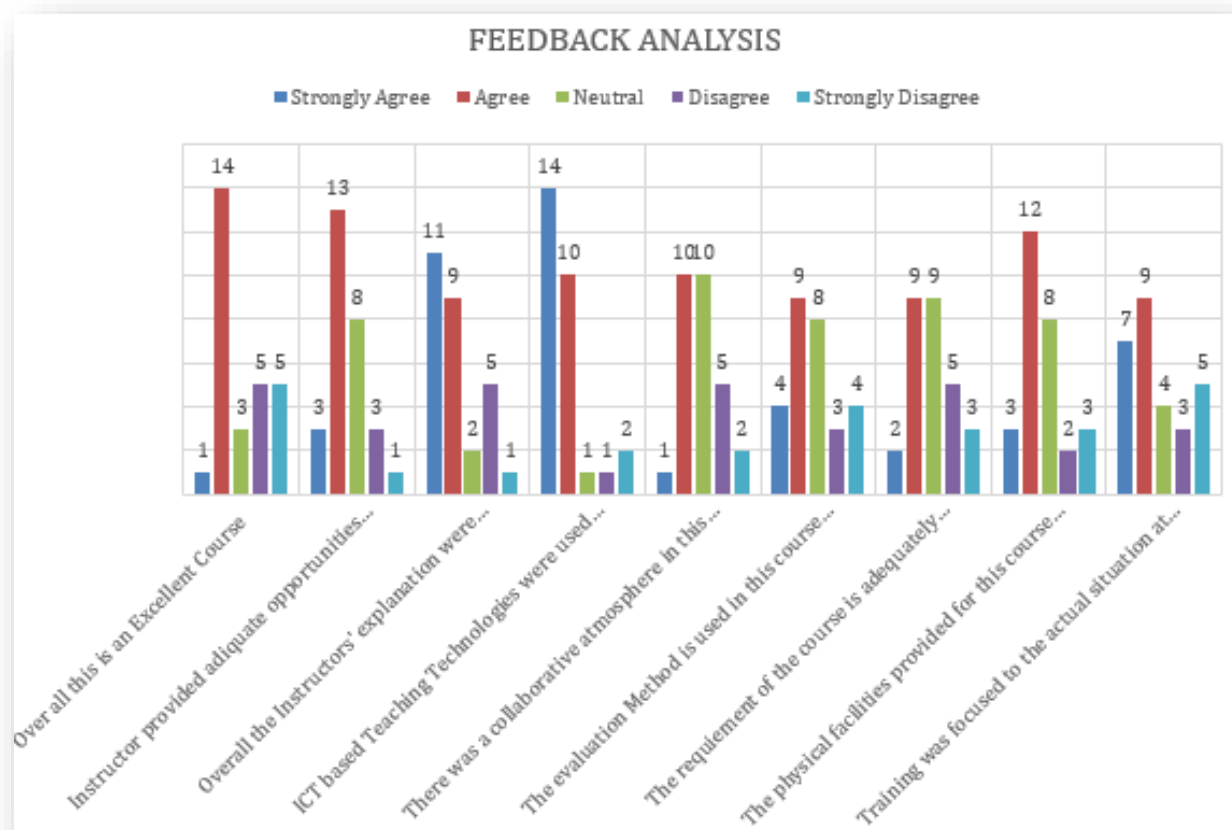
Student List:

SN	ROLLNO	NAME	Remarks
1	11900217001	TONMAY DUTTA	Successfully Completed
2	11900217002	SOUYAMA DEBNATH	Successfully Completed
3	11900217003	SOUBHIK DUTTA	Successfully Completed
4	11900217004	SHINJINI SANYAL	Successfully Completed
5	11900217006	ROHIT KUMAR BARMAN	Successfully Completed
6	11900217008	RAJANI KANTA ROUTH	Successfully Completed
7	11900217009	PRINCE KUMAR	Successfully Completed
8	11900217010	PAYEL DAS	Successfully Completed
9	11900217011	NIKITA BARA	Successfully Completed
10	11900217012	NANCY KUMARI PRASAD	Successfully Completed
11	11900217013	MD AHTESHAMUL HUSSAIN	Successfully Completed
12	11900217014	MANJIT SINGH	Successfully Completed
13	11900217015	LAKSHMINARAYAN GHOSH	Successfully Completed
14	11900217016	KISHAN BISWAKARMA	Successfully Completed
15	11900217017	JYOTIRMAY DEB	Successfully Completed
16	11900217018	INDRA NATH MUKHERJEE	Successfully Completed
17	11900217019	HIMADRI BHATTACHARYA	Successfully Completed
18	11900217020	DIPAK BARMAN	Successfully Completed
19	11900217021	CHAYAN KARMAKAR	Successfully Completed
20	11900217022	BISWAJIT SHARMA	Successfully Completed
21	11900217023	BIKRAM BARMAN	Successfully Completed
22	11900217024	BIDISHA DAS	Successfully Completed
23	11900217025	BAISHALI SAHA	Successfully Completed
24	11900217026	AYAN DUTTA	Successfully Completed
25	11900217027	ARPAN BHAKTA	Successfully Completed
26	11900217028	ANUSHKA KUMARI	Successfully Completed
27	11900217029	ADITYA HALDER	Successfully Completed
28	11900217030	ADARSH RAI	Successfully Completed

SILIGURI INSTITUTE OF TECHNOLOGY
DEPARTMENT OF INFORMATION TECHNOLOGY
Report on Industrial Training on ML With Python

Training Details:

Training on Machine Learning with PYTHON
Resource Organization: ARDENT
Training Date: **16th September 2019 to 20th September 2019**
Venue: **ONLINE MODE**
Student: 3rd Year IT (6th Semester)
Students Enrolled: **60**
Students Completed Successfully: **60**
Pass Out Year: **2020**
Feedback Analysis: **Attached**
Student List: **Attached**

Introduction:
Artificial Intelligence (AI), Machine Learning (ML) and Data Science (DS) are the pillars of the fourth industrial revolution. ML is an application of AI which allows computers to automatically learn from data without being explicitly programmed. Python has been designed with the provision for creating Machine Learning algorithms. Python is preferred as the best and robust platform for Machine Learning systems. Python also has numerous libraries for machine learning, data manipulation and analysis as well as a very active development community that continuously updates and creates new packages. It has been adopted by a wide variety of industries and applications including Data Science, Machine Learning, Data Analytics, Predictive Analytics, Business Intelligence and Web Analytics. This workshop aims to explore Python Programming right from installation, fundamentals to Machine Learning algorithms.

The Training session covered the basic algorithm that helps us to build and apply prediction functions with an emphasis on practical applications. **Training Objectives**
Main objectives of training were to learn:
- •How to determine and measure program complexity,
- •Python Programming
- •ML Library Scikit, Numpy , Matplotlib, Pandas , Theano , TensorFlow
- •Statistical Math for the Algorithms.
- •Learning to solve statistics and mathematical concepts.
- •Supervised and Unsupervised Learning
- •Classification and Regression
- •ML Algorithms
- •Machine Learning Programming and Use Cases.

The outcomes of this workshop are:

- Understand the components of a Machine Learning algorithm.
- Apply Machine Learning tools to build and evaluate predictors

- How Machine Learning uses computer algorithms to search for patterns in data
- How to uncover hidden themes in large collections of documents using topic modeling
- How to prepare data, deal with missing data and create custom data analysis solutions for different industries
- Familiarity with Python installation, syntax and design

Why Python Is a Perfect Language for Machine Learning?

1. **A great library ecosystem -** A great choice of libraries is one of the main reasons Python is the most popular programming language used for AI. A library is a module or a group of modules published by different sources which include a pre-written piece of code that allows users to reach some functionality or perform different actions. Python libraries provide base level items so developers don't have to code them from the very beginning every time. ML requires continuous data processing, and Python's libraries let us access, handle and transform data. These are some of the most wide spread libraries we can use for ML and AI:

- Scikit-learnfor handling basic ML algorithms like clustering, linear and logistic regressions,regression, classification, and others.
- Pandasfor high-level data structures and analysis. It allows merging and filtering of data, aswell as gathering it from other external sources like Excel, for instance.
- Kerasfor deep learning. It allows fast calculations and prototyping, as it uses the GPU inaddition to the CPU of the computer.
- TensorFlowfor working with deep learning by setting up, training, and utilizing artificialneural networks with massive datasets.
- Matplotlibfor creating 2D plots, histograms, charts, and other forms of visualization.
- NLTK f or working with computational linguistics, natural language recognition, and processing.
- Scikit-imagefor image processing.
- PyBrainfor neural networks, unsupervised and reinforcement learning.
- Caffefor deep learning that allows switching between the CPU and the GPU
- StatsModelsfor statistical algorithms and data exploration.

In the PyPI repository, we can discover and compare more python libraries.

2. **A low entry barrier** -Working in the ML and AI industry means dealing with a bunch of data that we need to process in the most convenient and effective way. The low entry barrier allows more data scientists to quickly pick up Python and start using it for AI development without wasting too much effort into learning the language.

In addition to this, there's a lot of documentation available, and Python's community is always there to help out and give advice

3. Flexibility- Python for machine learning is a great choice, as this language is very flexible:
   - It offers an option to choose either to use OOPs or scripting.
   - There's also no need to recompile the source code, developers can implement any
   - changes and quickly see the results.
   - Programmers can combine Python and other languages to reach their goals.

4. **Good Visualization Options-** For AI developers, it's important to highlight that in artificial intelligence, deep learning, and machine learning, it's vital to be able to represent data in a human-readable format. Libraries like Matplotlib allow data scientists to build charts, histograms, and plots for better data comprehension, effective presentation, and visualization. Different application programming interfaces also simplify the visualization process and make it easier to create clear reports.

5. **Community Support-** It's always very helpful when there's strong community support built around the programming language. Python is an open-source language which means that there's a bunch of resources

open for programmers starting from beginners and ending with pros. A lot of Python documentation is available online as well as in Python communities and forums, where programmers and machine learning developers discuss errors, solve problems, and help each other out. Python programming language is absolutely free as is the variety of useful libraries and tools.

6. **Growing Popularity-**As a result of the advantages discussed above, Python is becoming more and more popular among data scientists. According to Stack Overflow, the popularity of Python is predicted to grow until 2020, at least. This means it's easier to search for developers and replace team players if required. Also, the cost of their work maybe not as high as when using a less popular programming language Data Preprocessing, Analysis & Visualization Machine Learning algorithms don't work so well with processing raw data. Before we can feed such data to an ML algorithm, we must preprocess it. We must apply some transformations on it. With data preprocessing, we convert raw data into a clean data set.

To perform data this, there are 7 techniques –

1. Rescaling Data -For data with attributes of varying scales, we can rescale attributes to possess the same scale. We rescale attributes into the range 0 to 1 and call it normalization. We use the Min Max Scaler class from scikit-learn. This gives us values between 0 and 1.

2. Standardizing Data -With standardizing, we can take attributes with a Gaussian distribution and different means and standard deviations and transform them into a standard Gaussian distribution with a mean of 0 and a standard deviation

3. Normalizing Data -In this task, we rescale each observation to a length of 1 (a unit norm). For this, we use the Normalizer class.

4. Binarizing Data -Using a binary threshold, it is possible to transform our data by marking the values above it 1 and those equal to or below it, 0. For this purpose, we use the Binarizer class.

5. Mean Removal-We can remove the mean from each feature to center it on zero.

6. One Hot Encoding -When dealing with few and scattered numerical values, we may not need to store these. Then, we can perform One Hot Encoding. For k distinct values, we can transform the feature into a k-dimensional vector with one value of 1 and 0 as the rest values.

7. Label Encoding -Some labels can be words or numbers. Usually, training data is labelled with words to make it readable. Label encoding converts word labels into numbers to let algorithms work on them

Machine Learning Algorithms:

There are many types of Machine Learning Algorithms specific to different use cases. As we work with datasets, a machine learning algorithm works in two stages. We usually split the data around 20%-80% between testing and training stages. Under supervised learning, we split a dataset into a training data and test data in Python ML. Followings are the Algorithms of Python Machine Learning -

1. Linear Regression-Linear regression is one of the supervised Machine learning algorithms in Python that observes continuous features and predicts an outcome. Depending on whether it runs on a single variable or on many features, wecan call it simple linear regression or multiple linear regression .This is one of the most popular Python ML algorithms and often under-appreciated. It assigns optimal weights to variables to create a line ax+b to predict the output. We often use linear regression to estimate real values like a number of calls and costs of houses based on continuous variables. The regression line is the best line that fits Y=a*X+b to denote a relationship between independent and dependent variables.

2. Logistic Regression -Logistic regression is a supervised classification is unique Machine Learning algorithms in Python that find sits use in estimating discrete values like 0/1, yes/no, and true/false. This is based on a given set of independent variables. We use a logistic function to predict the probability of an event and this gives us an output between 0 and 1. Although it says 'regression', this is actually a classification algorithm. Logistic regression fits data into a logit function and is also called logit regression.

3. Decision Tree -A decision tree falls under supervised Machine Learning Algorithms in Python and comes of use for both classification and regression- although mostly for classification. This model takes an instance, traverses the tree, and compares important features with a determined conditional statement. Whether it descends to the left child branch or the right depends on the result. Usually, more important features are closer to the root. Decision Tree, a Machine Learning algorithm in Python can work on both categorical and continuous dependent variables. Here, we split a population into two or more homogeneous sets. Tree models where the target variable can take a discrete set of values are called classification trees; in these tree structures, leave represent class labels and branches represent conjunctions of features that lead to those class labels. Decision trees where the target variable can take continuous values (typically real numbers) are called regression trees.

4. Support Vector Machine (SVM)-SVM is a supervised classification is one of the most important Machines Learning algorithms in Python, that plots a line that divides different categories of your data. In this ML algorithm, we calculate the vector to optimize the line. This is to ensure that the closest point in each group lies farthest from each other. While you will almost always find this to be a linear vector, it can be other than that. An SVM model is are presentation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. When data are unlabeled, supervised learning is not possible, and an unsupervised learning approach is required, which attempts to find natural clustering of the data to groups, and then map new data to these formed groups.

5. Naïve Bayes Algorithm - Naive Bayes is a classification method which is based on Bayes' theorem. This assumes independence between predictors. A Naive Bayes classifier will assume that a feature in a class is unrelated to any other. Consider a fruit. This is an apple if it is round, red, and 2.5 inches in diameter. A Naive Bayes classifier will say these characteristics independently contribute to the probability of the fruit being an apple. This is even if features depend on each other. For very large data sets, it is easy to build a Naive Bayesian model. Not only is this model very simple, it performs better than many highly sophisticated classification methods. Naïve Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Maximum-likelihood training can be done by evaluating a closed-form expression, which takes linear time, rather than by expensive iterative approximation as used for many other types of classifiers.

6. k NN Algorithm -This is a Python Machine Learning algorithm for classification and regression- mostly for classification. This is a supervised learning algorithm that considers different centurions and uses a usually Euclidean function to compare distance. Then, it analyzes the results and classifies each point to the group to optimize it to place with all closest points to it. It classifies new cases using a majority vote of k of its neighbors. The case it assigns to a class is the one most common among its K nearest neighbors. For this, it uses a distance function. k-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification.
k-NN is a special case of a variable- bandwidth, kernel density "balloon" estimator with a uniform kernel.

7. K-Means Algorithm -k-Means is an unsupervised algorithm that solves the problem of clustering. It classifies data using a number of clusters. The data points inside a class are homogeneous and heterogeneous to peer groups.
k-means clustering is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. k -means clustering aims to partition n observations into k-clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.
k-means clustering is rather easy to apply to even large data sets, particularly when using heuristics such as Lloyd's algorithm. It often is used as a preprocessing step for other algorithms, for example to find a starting configuration.

The problem is computationally difficult(NP-hard). k-means originates from signal processing, and still finds use in this domain. In cluster analysis, the k-means algorithm can be used to partition the input data set into k partitions (clusters).

k-means clustering has been used as a feature learning (or dictionary learning) step, in either(semi-)supervised learning or unsupervised learning.

8.Random Forest - A random forest is an ensemble of decision trees. In order to classify every new object based on it s attributes, trees vote for class- each tree provides a classification. The classification with the most votes win sin the forest. Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

Feedback Analysis:
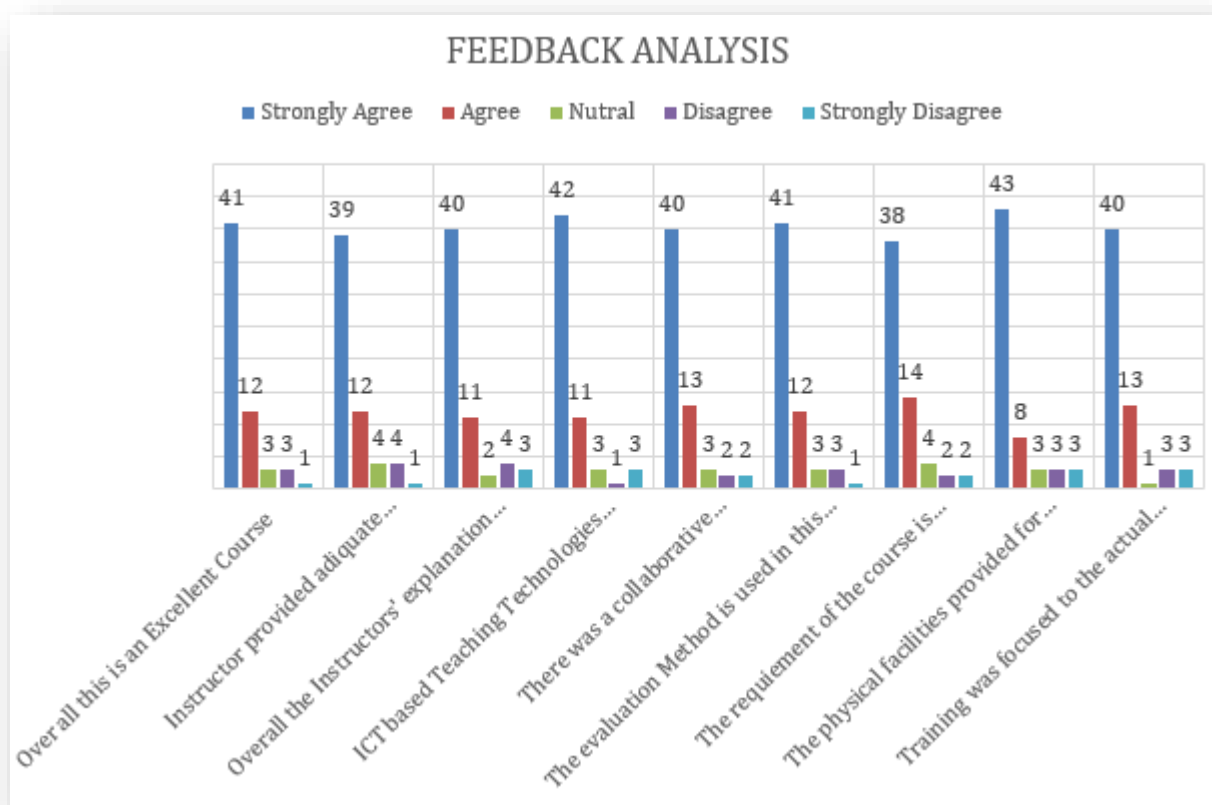
Student List:

SN	ROLL NO	NAME	Remarks
1	11900216001	VISAL HAZRA	Successfully Completed
2	11900216002	UJJAYINEE MAJUMDER	Successfully Completed
3	11900216003	TUHIN BANERJEE	Successfully Completed

4	11900216004	TAPAS KR PRAMANIK	Successfully Completed
5	11900216005	SWATI KUMARI	Successfully Completed
6	11900216006	SWAGATA SARKAR	Successfully Completed
7	11900216007	SUYASH KUMARI	Successfully Completed
8	11900216008	SUKIRTI KUMARI	Successfully Completed
9	11900216009	SUBHRAJIT SAHA	Successfully Completed
10	11900216010	SUBHANKAR BHOWAL	Successfully Completed
11	11900216011	SOURAV BAKSHI	Successfully Completed
12	11900216012	SIDDHARTH PANDEY	Successfully Completed
13	11900216013	SIBANGI DEB	Successfully Completed
14	11900216014	SHASWAT SINGH	Successfully Completed
15	11900216015	SAURAV GOSWAMI	Successfully Completed
16	11900216016	SANANDA CHATTERJEE	Successfully Completed
17	11900216017	ROHAN MITRA	Successfully Completed
18	11900216018	RASHMI PRASAD	Successfully Completed
19	11900216019	RAMAKANT PRASAD	Successfully Completed
20	11900216020	RAJ ROY	Successfully Completed
21	11900216021	RAINAK BASAK	Successfully Completed
22	11900216022	RAINA CHOUDHURY	Successfully Completed
23	11900216023	RAHUL KUMAR	Successfully Completed
24	11900216024	RAHUL CHOUDHURI	Successfully Completed
25	11900216025	PURBITA PAL	Successfully Completed
26	11900216027	PRATIK BOSE	Successfully Completed
27	11900216028	POOJA SINGH	Successfully Completed
28	11900216029	PANKAJ CHAUDHURI	Successfully Completed
29	11900216030	NIKITA GHOSH	Successfully Completed
30	11900216031	NIKESH KUMAR SAH	Successfully Completed
31	11900216032	NAMRATA DUTTA	Successfully Completed
32	11900216033	MEGHA BHATTACHARJEE	Successfully Completed
33	11900216034	MANISHA JHA	Successfully Completed
34	11900216035	MADHUSHREE MAJUMDAR	Successfully Completed
35	11900216036	KAUSTAV DAS	Successfully Completed
36	11900216037	KARTIK ROY	Successfully Completed
37	11900216038	GITASHREE SAHA	Successfully Completed
38	11900216039	DREEK GUHA	Successfully Completed
39	11900216040	DIVYANSHU PRAKASH PUNJ	Successfully Completed
40	11900216041	DIPAYAN GHOSH	Successfully Completed
41	11900216042	DIPANKAR SAHA	Successfully Completed
42	11900216044	DEBJYOTI JHA	Successfully Completed
43	11900216045	DEBANJAN DAS	Successfully Completed
44	11900216046	DEBANGSHU SAHA	Successfully Completed
45	11900216047	CHAITALI SAHA	Successfully Completed

46	11900216048	BIKASH GURAGAI	Successfully Completed
47	11900216049	AYUSH ANAND	Successfully Completed
48	11900216050	AYAN MUKHOPADHYAY	Successfully Completed
49	11900216051	AVISHEK DAS	Successfully Completed
50	11900216052	AUROSHREE MUKHERJEE	Successfully Completed
51	11900216053	ASHMITA BASU MAZUMDAR	Successfully Completed
52	11900216054	ARYA ROY	Successfully Completed
53	11900216055	ARVI HASSAN	Successfully Completed
54	11900216056	ARUP DAS	Successfully Completed
55	11900216057	ARITRA BHATTACHARJEE	Successfully Completed
56	11900216058	ARINDAM PODDER	Successfully Completed
57	11900216059	ARGHYADEEP PANDIT	Successfully Completed
58	11900216061	ANKITA PAUL CHOWDHURY	Successfully Completed
59	11900216062	AKASH KUMAR	Successfully Completed
60	11900216063	ABHISHEK PRASAD SINGH	Successfully Completed

SILIGURI INSTITUTE OF TECHNOLOGY
DEPARTMENT OF INFORMATION TECHNOLOGY
Industrial Training Report on Big Data/Hadoop

Training Details:

Training on Big Data / Hadoop
Resource Organization: I & WE
Training Date: **18th January 2018 to 28th January 2018**
Venue: **SIT, OT&UML Lab**
Student: 3rd Year IT (6th Semester)
Students Enrolled: **41**
Students Completed Successfully: 41
Pass Out Year: **2019**
Feedback Analysis: **Attached**
Student List: **Attached**

Introduction

Hadoop is an open-source framework that allows to store and process big data in a distributed environment across clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage. In the Training provides a quick introduction to Big Data, Map Reduce algorithm, and Hadoop Distributed File System.

Training Objective:

- Upon completion of this course, participants will be able to:
- Understand fundamentals of Concepts in Bigdata and hadoop etc
- Understand fundamentals of Hadoop etc.
- Be able to use the HDFS file system, debug and run simple Java programs for hdfs.
- Be aware of the important topics and principles of software development and write better &more maintainable code
- Be able to program using advanced Java topic like JDBC, Servlets and JSP .

What is Big Data?

Big data means really a big data, it is a collection of large datasets that cannot be processed using traditional computing techniques. Big data is not merely a data, rather it has become a complete subject, which involves various tools, techniques and frameworks.

Advantages of Hadoop:

- Hadoop framework allows the user to quickly write and test distributed systems. It is efficient, and it automatic distributes the data and work across the machines and in turn, utilizes the underlying parallelism of the CPU cores.
- Hadoop does not rely on hardware to provide fault-tolerance and high availability (FTHA), rather Hadoop library itself has been designed to detect and handle failures at the application layer.

- Servers can be added or removed from the cluster dynamically and Hadoop continues to operate without interruption.
- Another big advantage of Hadoop is that apart from being open source, it is compatible on all the platforms since it is Java based.
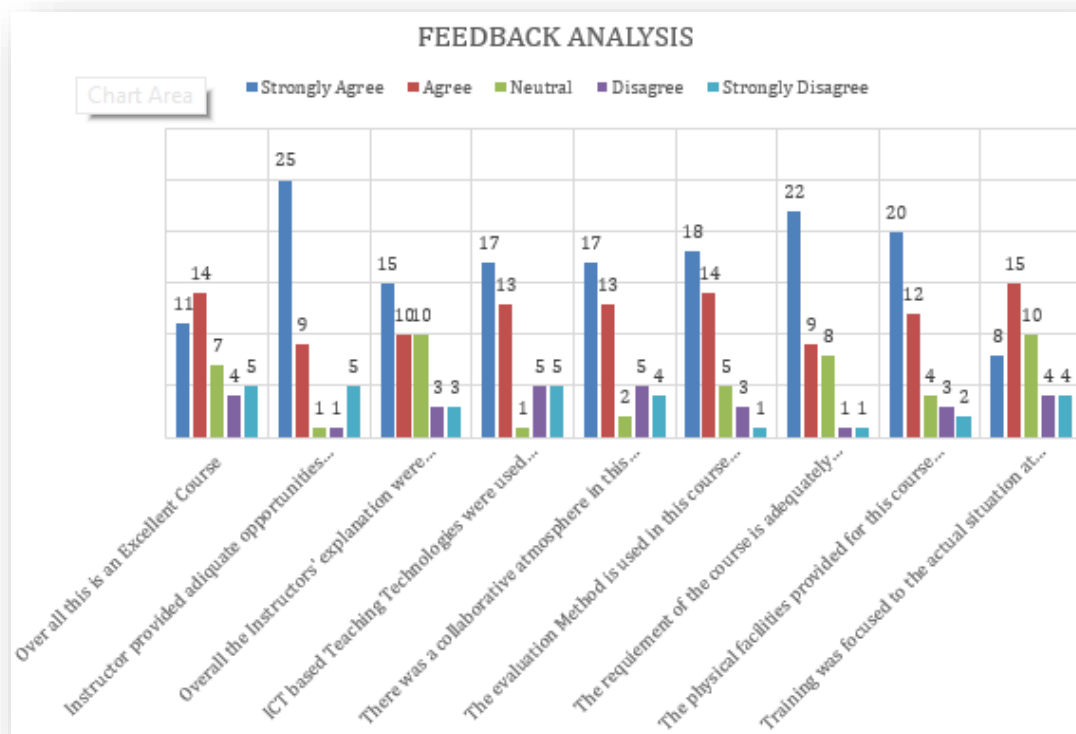
Training Methodology:

- Hands on practice approach to training, behavioral model of training would be practiced.
- During the training, the Trainee would implement a project related to respective modules.
- Commitment to Individual growth and constant evaluation.
- Implementation of programming techniques through a Project.

Summary of the program:
The following points can be noted from the program.

- ❖ At the beginning of the training trainer has clearly described the basic Introduction to java its application in industries in different areas.
- ❖ Students had done many data analysis algorithm by themselves during the trainings.
- ❖ During the training some students raised their queries and the trainer had explained all the quarries of the students.
- ❖ At the end of the training an online exam was conducted.
- ❖ As per the feedback received from the students end, the entire session was really fruitful
- ❖ and enjoyable and the students have learned framework of Hadoop.

Feedback analysis for the training:

FEEDBACK ANALYSIS

Student List:

SN	ROLLNO	NAME	Remarks
1	11900215001	ABHISHEK CHOUDHARY	Successfully Completed
2	11900215002	ABIR NANDY	Successfully Completed
3	11900215003	ADRIJA ROY	Successfully Completed
4	11900215004	AIJURA KSHIAR	Successfully Completed
5	11900215005	AKASH DAS	Successfully Completed
6	11900215006	AKASH MANNA	Successfully Completed
7	11900215007	ALOK SARKAR	Successfully Completed
8	11900215008	AMARENDRA NATH SAHA	Successfully Completed
9	11900215009	ANKIT CHHABRA	Successfully Completed
10	11900215010	ARINDAM SAHA	Successfully Completed
11	11900215011	ARTI KUMARI	Successfully Completed
12	11900215012	ARZU DAS	Successfully Completed
13	11900215013	ASTIK KUMAR	Successfully Completed
14	11900215015	BISWAJIT MANDAL	Successfully Completed
15	11900215017	DITIPRIYA BHADRA	Successfully Completed
16	11900215018	GARGI BHATTACHARYA	Successfully Completed
17	11900215019	GARGI NAG	Successfully Completed
18	11900215020	GOPAL KUMAR DAS	Successfully Completed
19	11900215021	GOURAB KUMAR GUPTA	Successfully Completed
20	11900215022	GUNJAN CHHETRI	Successfully Completed
21	11900215023	IFFAT ZAIDI	Successfully Completed
22	11900215024	KALLOL ROY	Successfully Completed
23	11900215025	MANISH ANAND	Successfully Completed
24	11900215026	MAYURI ROY	Successfully Completed
25	11900215027	MD MUKHLESUR RAHAMAN	Successfully Completed
26	11900215028	MRIGANKA HEMBRAM	Successfully Completed
27	11900215029	NITA SARKAR	Successfully Completed
28	11900215030	OM KUMARI PRADHAN	Successfully Completed
29	11900215032	POULAMI DAS	Successfully Completed
30	11900215033	PRATIKCHA KHAWAS	Successfully Completed
31	11900215034	RAJARSHI DEB	Successfully Completed
32	11900215035	RAJIV JAJODIA	Successfully Completed
33	11900215036	RASHMI KUMARI	Successfully Completed
34	11900215037	SANTOSH KUMAR	Successfully Completed
35	11900215038	SHASHI KUMAR SAHA	Successfully Completed
36	11900215039	SHIMPI PAUL	Successfully Completed
37	11900215040	SOMJYOTI HORE	Successfully Completed
38	11900215041	SRISHTI KUMARI	Successfully Completed
39	11900215042	SWAGATA SAHA	Successfully Completed
40	11900215043	TANFEEZ AHSAN	Successfully Completed
41	11900215044	TUHIN ROY	Successfully Completed

Training Details:

Training on Big Data / Hadoop
Resource Organization: I & WE
Training Date: **16th July 2017 to 17th July 2018**
Venue: **SIT, OT&UML Lab**
Student: 3rd Year IT (6th Semester)
Students Enrolled: **16**
Students Completed Successfully: 16
Pass Out Year: **2018**
Feedback Analysis: **Attached**
Student List: **Attached**

Introduction

Hadoop is an open-source framework that allows to store and process big data in a distributed environment across clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage. In the Training provides a quick introduction to Big Data, Map Reduce algorithm, and Hadoop Distributed File System.

Training Objective:

- Upon completion of this course, participants will be able to:
- Understand fundamentals of Concepts in Bigdata and hadoop etc
- Understand fundamentals of Hadoop etc.
- Be able to use the HDFS file system, debug and run simple Java programs for hdfs.
- Be aware of the important topics and principles of software development and write better &more maintainable code
- Be able to program using advanced Java topic like JDBC, Servlets and JSP.

What is Big Data?

Big data means really a big data, it is a collection of large datasets that cannot be processed using traditional computing techniques. Big data is not merely a data, rather it has become a complete subject, which involves various tools, techniques and frameworks.

Advantages of Hadoop:

- Hadoop framework allows the user to quickly write and test distributed systems. It is efficient, and it automatic distributes the data and work across the machines and in turn, utilizes the underlying parallelism of the CPU cores.
- Hadoop does not rely on hardware to provide fault-tolerance and high availability (FTHA), rather Hadoop library itself has been designed to detect and handle failures at the application layer.

- Servers can be added or removed from the cluster dynamically and Hadoop continues to operate without interruption.
- Another big advantage of Hadoop is that apart from being open source, it is compatible on all the platforms since it is Java based.
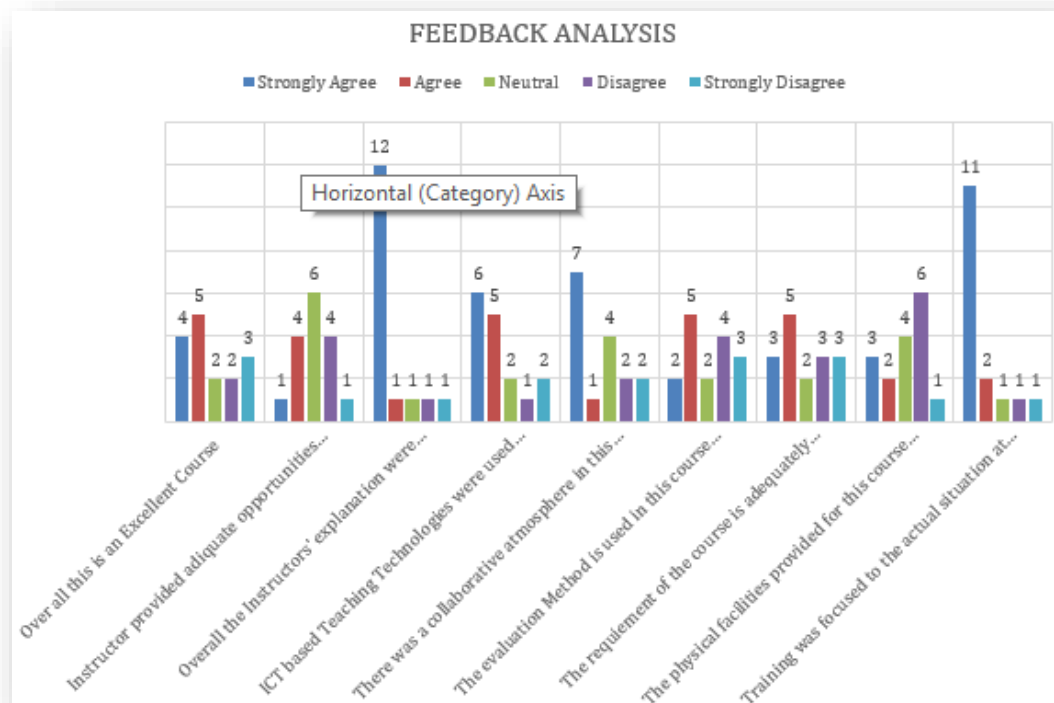
Training Methodology:

- Hands on practice approach to training, behavioral model of training would be practiced.
- During the training, the Trainee would implement a project related to respective modules.
- Commitment to Individual growth and constant evaluation.
- Implementation of programming techniques through a Project.

Summary of the program:

The following points can be noted from the program.

- ❖ At the beginning of the training trainer has clearly described the basic Introduction to java its application in industries in different areas.
- ❖ Students had done many data analysis algorithm by themselves during the trainings.
- ❖ During the training some students raised their queries and the trainer had explained all the quarries of the students.
- ❖ At the end of the training an online exam was conducted.
- ❖ As per the feedback received from the students' end, the entire session was really fruitful
- ❖ and enjoyable and the students have learned framework of Hadoop.

Feedback analysis for the training:

FEEDBACK ANALYSIS

Student List:

SN	ROLL NO.	NAME	Remarks
1	11900214001	ARABINDA ROY	Successfully Completed
2	11900214002	ARPAN PAUL	Successfully Completed
3	11900214003	BINEETA MAJUMDER	Successfully Completed
4	11900214004	BIPUL SARKAR	Successfully Completed
5	11900214005	DEEPAK SINGH	Successfully Completed
6	11900214006	JYOTI KUMARI GUPTA	Successfully Completed
7	11900214007	NAIRITH DAS	Successfully Completed
8	11900214008	NAYAN KUMAR	Successfully Completed
9	11900214009	PRABHAT PUSHKAR	Successfully Completed
10	11900214010	PUJA KUMARI	Successfully Completed
11	11900214011	SAHELI PYNE	Successfully Completed
12	11900214012	SOUBHIK DAS	Successfully Completed
13	11900214013	SUSHRI PAUL	Successfully Completed
14	11900214015	VINEET KUMAR	Successfully Completed
15	11900214016	YEAKUTUN NESSA	Successfully Completed
16	11900215045	DEBOJIT PAUL	Successfully Completed

Training Details:

    Training on Big Data / Hadoop

    Resource Organization: NSIC

    Training Date: **1st August 2016 to 12th August 2016**

    Venue: **SIT, Programming Lab I/ SIT, Programming Lab II/OT&UML Lab**

    Student: 3rd Year IT (6th Semester)

    Students Enrolled: **17**

    Students Completed Successfully: **17**

    Pass Out Year: **2017**

    Feedback Analysis: **Attached**

    Student List: **Attached**

Introduction:

Apart from University requirement, Java is also a pre-requisite for learning latest technologies like Android and Big Data. In order to prepare and make students ready for industry Computer science department has carved out a course that specifically aligns with industry requirements and conducted by industry experts.

In this training session students learned basic object oriented concepts such as inheritance, encapsulation, and abstraction. They learn how to create and use simple Java classes containing arrays, loops, and conditional constructs. They also learn to use and manipulate object references, and to write simple error handling code. They also learned some advance topic like JDBC connectivity, JSP, Servlates.

Training Objective:

Upon completion of this course, participants will be able to:

• Understand fundamentals of Java programming such as variables, conditional and iterative execution, methods, etc

• Understand fundamentals of object-oriented programming using Java, including defining classes, invoking methods, using class libraries, etc.

• Be able to use the Java SDK environment to create, debug and run simple Java programs

• Be aware of the important topics and principles of software development and write better &more maintainable code

• Be able to program using advanced Java topic like JDBC, Servlets and JSP .

Training Methodology:

- Online on approach to training, behavioral model of training would be practiced.
- During the training, the Trainee would implement a project related to respective modules.
- Commitment to Individual growth and constant evaluation.
- Implementation of programming techniques through a Project.

Summary of the program:

The following points can be noted from the program.

❖ At the beginning of the training trainer has clearly described the basic Introduction to, OOPs programming and java its application in industries in different areas with the students.
❖ Students had done many programming by themselves during the trainings.
❖ During the training some students raised their queries and the trainer had explained all the quarries of the students.
❖ At the end of the training an online exam was conducted.
❖ As per the feedback received from the students end, the entire session was really fruitful
❖ and enjoyable and the students have learned many things about Java

Feedback analysis for the training:
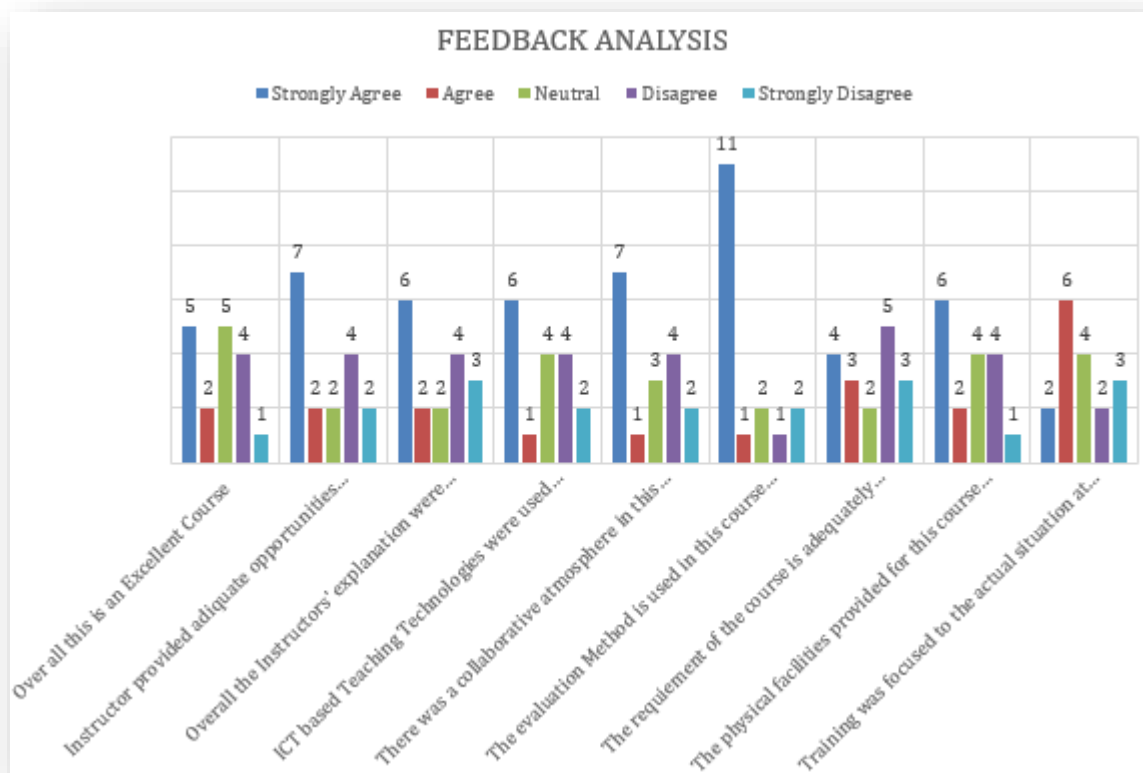
Student List:

SN	ROLL NO.	NAME	Remarks
1	11900213001	ABHISHEK KUMAR LAL	Successfully Completed
2	11900213002	AMIT KUMAR ROY	Successfully Completed
3	11900213003	ANAND KESHAV	Successfully Completed
4	11900213004	ANINDITA BHATTACHARJEE	Successfully Completed
5	11900213005	KUMARI RIAH	Successfully Completed
6	11900213006	MANISH THAKUR	Successfully Completed
7	11900213007	MUKESH KUMAR BURNWAL	Successfully Completed
8	11900213008	POONAM SONAR	Successfully Completed
9	11900213009	RAHUL KUMAR SHAW	Successfully Completed
10	11900213010	RAMU CHHETRI	Successfully Completed
11	11900213011	RITESH SHAW	Successfully Completed
12	11900213012	RITUPARNA DAS	Successfully Completed
13	11900213013	SANCHITA MONDAL	Successfully Completed
14	11900213014	SUBHANKAR BISWAS	Successfully Completed
15	11900213015	SUSHMITA ROY	Successfully Completed
16	11900213016	SWAGAT BHATTACHARJEE	Successfully Completed
17	11900213018	YASHASBI GUPTA	Successfully Completed