Data Scientist Certification Program

A **Five** month long certification programme with 10+ Industry Project.

Designed for Fresh Graduates and Professionals with **less than 2 years** of work experience.

CONTENTS

1. About the course

2. Key Features

3. Track your Learning Path

4. Program Learning outcomes

5. Who Should Enroll

6. Curriculum & Course segments

- Preparatory Course
- Essential Statistics Foundation
- Python Programming for Data Science
- Machine Leaning with python
- Essentials of Data Bases – SQL
- Data Story Telling using Tableau
- Deep Learning

7. Certification

8. Course mentors

About the Course

Technology is ever - changing and so are the industrial demands. People who are resistant to change and unwilling to expand their skill-sets have a high risk of losing their jobs.

Now is the right time to upgrade your skills and learn new technologies because Industries are openly absorbing talent that is well-versed with the latest Technologies, i.e. Data Science, Big data, Data Analytics, etc

To make candidates Industry-ready and to expand their range of skills, Business Toys has designed an industry centered Data Scientist Program. This is an online instructor-led certification program.

This program is created to balance the upskilling requirements of students and the industry demands, hence helping students and professionals to make an easy transition and upgrade in their career with the desired designation, roles and salary packages.

Key Features

150+ Hours
Instructure-led
online Session

10+ Industry
projects

100% Practical
Hands-on session

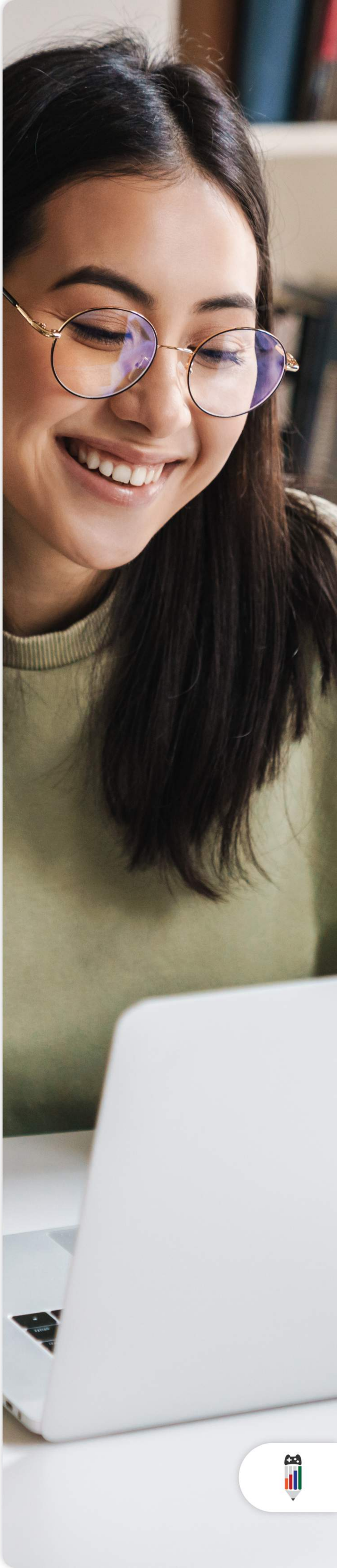
20+ In-demand
skills

No Cost EMI

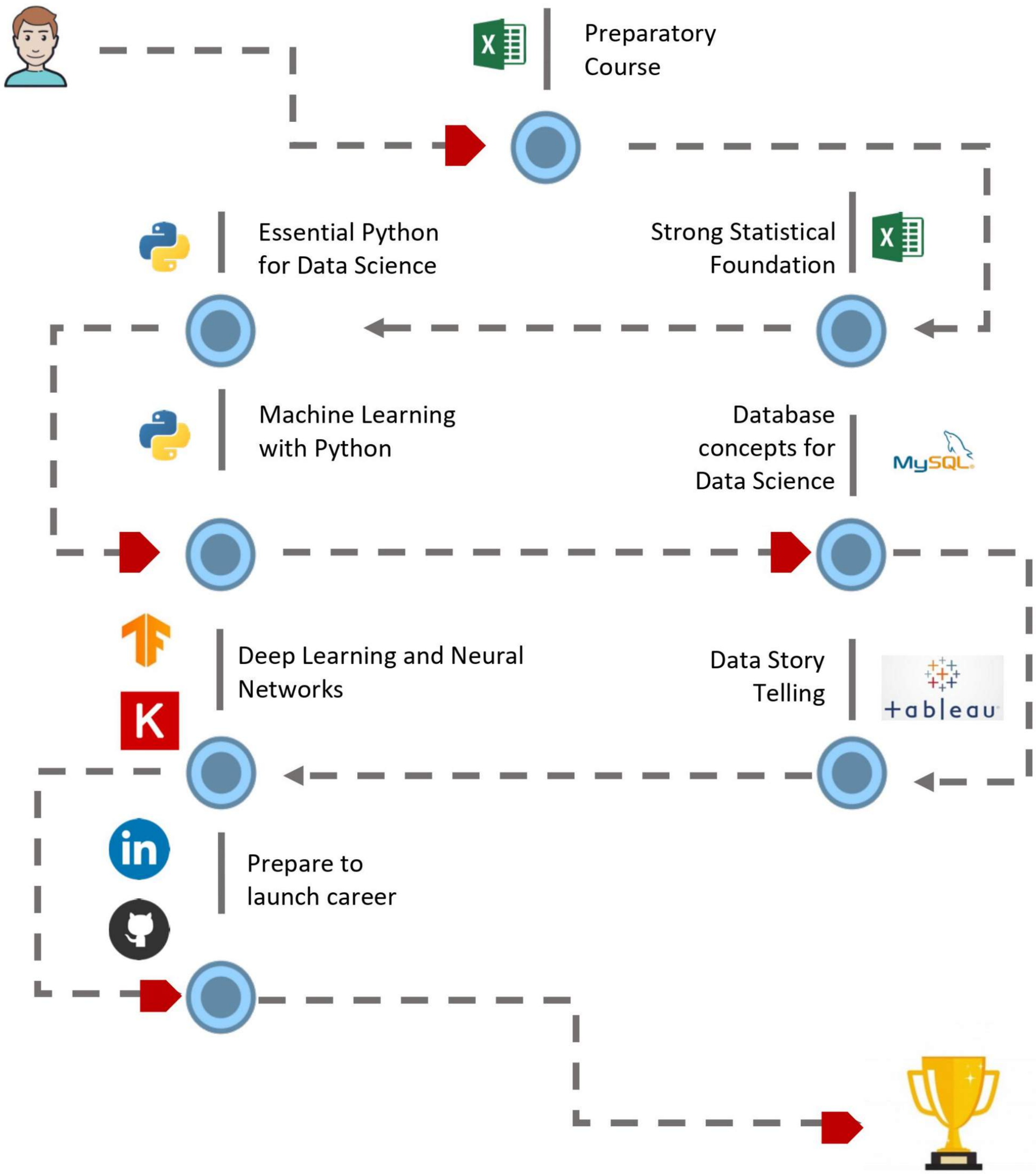
Endless Placement
Support

Industry Recognized
Certification

Recorded video
for every session



Learning Path for Data Science



Program Outcomes

- Gain in-depth understanding about statistics – Descriptive Statistics & Inferential Statistics.
- Obtain detailed understanding of various Machine Learning Techniques such as Supervised Machine Learning, Un-supervised Machine Learning, etc.
- Learn to solve real-time industry problem using various supervised and unsupervised Machine Learning techniques.
- Obtain string understanding of how to scrap or mine the data from multiple website and convert the unstructured data into structured data.
- Get detailed understanding of covariance, co-relation, linear and non-linear regression.
- Get skilled in usage of python for Data Science and get comfortable using important libraries like pandas, numpy, scipy, sklearn, seaborn, etc.
- Learn to build unsupervised machine learning models like PCA, clustering and NMF for applications like recommender systems.
- Get a strong foundation to Deep Learning and Neural Networks and build applications like handwritten digit recognition using ANN.

Who should enroll for this Program?

Data Science & Analytics roles require a blend of domain experiences, Data Science Skills, Programming skills and skills related to tools and technologies like Excel, Tableau, etc. This career is best suited for College Graduate / Post-Graduates and experienced professionals' with less than 2 years of work experience. Candidates from any educational background with interest towards technology, data analysis and possessing analytical mindset are best suited for taking up Data Scientist Certification Program.

- IT Professionals
- Banking professionals
- Financial Analysts
- Research Analysts
- Marketing Managers
- College Graduates
- Business Analyst

Preparatory Course

This course is specifically designed for new-bees in data science with not prior experience with data and statistics. Course is completely focused on use of Microsoft Excel for Analyzing Data because learners will be focused on learning important aspects of statistics without much of hassle to deal with programming aspects. This course prepares you to understand semantics of the data, data analysis framework and provides a string foundation to statistics for Data Science.

Course curriculum:

- Exploratory Data Analysis for a Mobile & Electronics company (A practical use-case)
  - Understanding problem statement
  - Understanding types of Data
  - Analysis of categorical data
  - Univariate Analysis
  - Histograms, boxplots, density plots
  - Cross Tables
  - 3M Analysis of numerical data
  - Concept of percentiles and ranges
  - Univariate analysis and multivariate analysis
  - Creating dashboards using Excel
- Introduction to statistics, Descriptive statistics – detailed understanding of central tendency, dispersion, kurtosis, skewness, normal distribution and its importance, visualizing distributions.
- Different types of probability distributions, Bernoulli and Binomial Distribution, Log normal distribution, power law distribution, Box cox transform.
- Covariance and correlation, correlation vs causation
- Inferential Statistics – population and sample, central limit theorem, confidence interval, discrete and continuous distributions, Hypothesis testing, t-test and different types, Anova, CHI Square test.

» **Number of assignments: 3**

Preparatory Course

This course will introduce the learner to the basics of the python programming environment, including fundamental python programming techniques such as lambdas, reading and manipulating csv files, and the numpy library. The course will introduce data manipulation and cleaning techniques using the popular python pandas data science library and introduce the abstraction of the Series and DataFrame as the central data structures for data analysis, along with tutorials on how to use functions such as groupby, merge, and pivot tables effectively. By the end of this course, students will be able to take tabular data, clean it, manipulate it, and run basic inferential statistical analyses.

This course should be taken before any of the other Applied Data Science with Python courses: Applied Plotting, Charting & Data Representation in Python, Applied Machine Learning in Python, Applied Text Mining in Python, Applied Social Network Analysis in Python.

Course curriculum:

- Introduction to python IDEs – Anaconda Distribution: Jupyter Notebook, VSCode, Spyder.
- Why learn python, keywords and identifiers, commends & indentations, variables and data types, operators.
- Control flow: If and else statements, While Loop, For Loop, Break and continue.
- Data structures: Lists, Tuples, Sets, Dictionary, Strings.
- Functions: Introduction to python user defined functions, Lambda functions, Modules, Packages, Exception handling, Debugging.
- Numpy and Pandas: Why Numpy and pandas, different operations and functions associated with Numpy and Pandas.
- Other important packages – Matplotlib, seaborn, scipy, sys, etc.

» **Number of assignments: 8**

Essential Database Concepts

Database handling using SQL:

- SQL: Introduction to Databases, Why SQL, Execution of an SQL statement, SQL Dataset, Installing MySQL.
- SQL Commands – USE, DESCRIBE, SHOW TABLES, SELECT, LIMIT, OFFSET, ORDER BY, DISTINCT, WHERE, Comparison operators, NULL, Logical operators.
- Aggregate functions : COUNT, MIN, MAX, AVG, SUM, GROUP BY, HAVING, Order of keywords, Join and Natural joins, Inner, Left, Right, Outer join, SubQueries/Nested Queries/Inner Queries.
- DML: INSERT, UPDATE, DETELE, CREATE TABLE, ALTER, ADD, MODIFY, DROP, DROP TABLE, TRUNCATE, DELETE, DCL: GRANT, REVOKE, Learning resources.

Database handling using SQL:

- Introduction to MongoDB & NoSQL framework
- Structure of data in MongoDB
- CRUD Operations and Queries in MongoDB [Create, Read, Update, Delete]

» **Number of assignments: 8**

Machine Learning (Track – I)

In this module, you will learn about the most effective machine learning techniques, and gain practice implementing them and getting them to work for yourself. More importantly, you'll learn about not only the theoretical underpinnings of learning, but also gain the practical know-how needed to quickly and powerfully apply these techniques to new problems. Finally, you'll learn about some of best practices in innovation as it pertains to machine learning and AI.

Course curriculum:

- Introduction to Machine Learning and Artificial Intelligence
- Exploratory Data Analysis and Data Manipulation
- Supervised Machine Learning:

  Regression & Classification Problems:

  - Univariate regression techniques, Multivariate Linear Regression, Assumptions of linear models, Homoscedastic, Residual analysis, Normality and Multicollinearity, evaluation and validation of regression models, analysis and interpretation of regression parameters. Analysis of factors impacting the target variable.

  - Optimization of linear models using polynomial regression, Quantile Regression, Regularization – Ridge & Lasso Regression.

  - Logistic Regression – Introduction to classification problems, validation of logistic regression – confusion matrix, accuracy, precision, recall, F1 score, ROC curve, AIC and AUC.

Number of assignments: 4

Machine Learning (Track – II)

Course curriculum: Other Supervised Machine Learning Algorithms

- Decision Trees: What is decision tree? Decision criterion for spitting: Entropy, Information Gain, Gini Impurities, constructing decision trees, splitting numerical values, Overfitting and under-fitting. Solving regression problems using decision trees.

- Ensemble Techniques: What are ensembles? Bootstrapped Aggregation (Bagging) Intuition, Random forest and their construction, Bias and variance trade-off, Gradient Boosting, XGBoost, AdaBoost.

- Support Vector Machine: Geometric representation, vector planes and selection of vector planes, loss functions, Dual form of SVM Formulation, Kernel Tricks, Polynomial Kernel, RBF Kernel, Domain Specific Kernels, SVM Classifiers and Repressors.

- K-Nearest Neighbour – How it works? Visual representation of KNN, Failure of KNN, Distances – Euclidean (L2), Manhattan (L1), Minkowski, Hamming, measuring effectiveness of KNN, KNN limitations, overfitting and under fitting, how to get best K value?

- Naïve Bayes – Introduction to conditional probability, Independent and Mutually Exclusive events, Bayes Theorem, Using Naïve Bayes in best possible way.

» **Number of assignments: 5**

Machine Learning (Track – III)

Course curriculum: Un-supervised Machine Learning Algorithms

- **Clustering:**

  Concept and intuition behind clustering, Significance of clustering in Data Science, Different types of clustering: K-Means Clustering, Agglomerative Hierarchical Clustering, Mean Shift clustering.

- **Dimensionality Reduction Techniques:**

  What is Dimensionality reduction? Row Vector and Column Vector, how to represent a data set? How to represent a dataset as a Matrix, Principal Component Analysis - Data Pre-processing: Feature Normalization, Mean of a data matrix, Data Pre-processing: Column Standardization, Co-variance of a Data Matrix.

  Concept of Linear Discriminant Analysis (LDA), advantages and disadvantages of PCA and LDA.

- **Recommendation Engine:**

  Intuition behind concept of recommendation engines. Non-Negative Matrix Factorization (NMF), Decomposition of matrix, concept of similarities- Euclidean distance and cosine similarity, optimization of time complexity in recommender systems.

» **Number of assignments: 3**

Machine Learning (Track – IV)

Course curriculum:

Feature Engineering:

- Dealing with categorical variables – one-hot-encoding and label encoding, outlier treatment, missing value treatment – imputations strategies, handling numerical features, Multiclass classification, feature selection and feature reduction.

Hyper-parameter Optimization:

- Hyper-parameter Optimization
- Understanding hyper-parameters for different Machine Learning algorithms
- Grid Search Cross Validation
- Randomized Search Cross Validation

Model Selection

- Selection of best Machine Learning Model – Time complexity approach, Bias and variance trade-off, Learning Curves, Analysis of learning curves to understand bias and variance in the models.

» **Number of assignments: 10**

Deep Learning

In this course, you will learn the foundations of Deep Learning, understand how to build neural networks, and learn how to lead successful machine learning projects. You will learn about basics working of Artificial Neural Networks and you will be able to build neural networks for solving classification and regression use cases.

Course curriculum:

- **Artificial Neural Networks:** History of Neural Network, How Biological Neurons work? Growth of Biological Neural networks, Perceptron, Multi layered perceptron, Notations, Training a single-neuron model, Feed forward, back propagation, Activation functions, Gradient Descent.

- **Keras:** Keras overview, GPU and CPU for deep learning, installing tensor flow, SoftMax Classifier, Initialization, Sigmoid, ReLU, Batch Normalization, Dropout, MNIST Classification in Keras, Hyper-parameter Tuning in Keras.

» **Number of assignments: 4**

Data Storytelling using Tableau

In this course, you will be able to master various aspects of Data Visualization and reporting. You will learn to build various types of advanced visualizations, providing user interactivity, building dashboards and stories. Moreover, you will also learn to map statistical concepts and data connections. It is very much essential for every Data Scientists to be able to present their analysis to decision makers in much easier way, for which this module will very helpful.

Course curriculum:

- Getting familiar with Tableau interface
- Establishing Database connections
- Simple exploratory data analysis using simple charts
- Data manipulations and creating calculated fields
- Exploratory Data Analysis and Data Manipulation
- Dashboards and stories
- Building advanced visualizations like Heat maps, Tree maps, Waterfall, Pareto, Gantt Chart, Sankey Chart, etc.

» **Number of assignments: 3**

Certificate

Certificate of Achievement

This is to Certify that

Ruchi Sharma

has successfully completed all prescribed requirements established
by Business Toys Pvt. ltd. for professional attainment of
Data Scientist Certification Program
for a period of 5 Months
and also implemented projects in relevant domain and therefore awarded certificate of

Data Scientist Certification Program

This 13 day of February 2021
UCN: BTDS21209878738

Digitally signed & Authorized by
Mr. Omkar Raikar
Director, Business Toys Pvt. Ltd.

Data Scientist Certification Program

- Internationally recognized certificates
- Unique Certificate identification number for authentication
- Real time validation check through the website
- Accepted by top multinationals

Chief Course Mentors

Omkar Raikar

Omkar heads Global Operations & Delivery of Analytics Products and solutions at Basal Analytics. He is also a co-founder at Business Toys Pvt. Ltd. Omkar has overall 9 years of Data Science experience with expertise in advanced Data Science applications like Natural Language Processing, Machine Learning, AI Product Development.

Manikanta Prasad

Mr. Manikanta Prasad is a Vice President at Business Toys Private Limited. He has 5+ years of experience working as a Data Scientist and has mentored 1000+ professionals to make a career transition in Data Science. He is an expert in Decisions Science using data for Businesses across sectors like FMCG, Retail and E-commerce.

Harika Samala

Harika Samala is a Data Science expert with 4+ years of experience working with top FinTech companies in India. Her areas of expertise are Natural Language Processing and Computer vision. She loves doing research in the field of advanced computing and Artificial Intelligence and is contributing a lot in educating Data Science enthusiasts in building their dream career.

Student Reviews

Arshdeep Kaur

amazon

I got to know about Business Toys from my friend. It's a wonderful platform to get the best mentor support and handholding. I have completed Data Scientist Certification Program from Business Toys and course has helped me get my first job at Amazon. Mentors always provided me an unbiased and frank feedback about my learning progress which made me work on my skills in right direction. Me being from non-technical graduation background, I was initially facing difficulties with python programming but with support of trainers and extra assignments I could adapt to it.

Ritesh Kumar

Kaseya

I have completed Data Science Certification Program from Business Toys. And they had helped me to get job as a Business Analyst at Kaseya. After 6 months of experience I learnt that Data Science should really be helpful for me to grow in my career so I got myself enrolled for Data Science Certification Program. I started demonstrating my Data Science skills at work and my work started getting recognized by senior management. I recently got promoted to a position of Sr. Business Analyst and I am working of Data based automations at Kaseya. I highly recommend Business Toys they for any kind of upskilling programs.

Abhishek Mishra

I have always been very passionate about Data Science and Data Analysis. In fact I love working with Data. That is why I joined Business Toys finishing school to do a professional course in Data Science. What I loved about Business Toys is that both Mani sir and Omkar sir teach practically using lots of case studies. In fact they teach every topic with cases. I am really thankful to Business toys and the mentors for helping me with resume development and interview preparations. Now, I am working as a data analyst in Analytical wizards, US based healthcare Industry.

Business Toys Pvt. Ltd.
#237, 4th Cross, 60 ft. Road, AGB Layout,
Chikkabanavara, Hesareghatta Main Road,
Bangalore, Karnataka (India)
560 090

Phone No : +91 9019 366 831
Email : contact@businesstoys.in