|  |  |
| --- | --- |
| 125  Flat Data Science Vector Icons ~ Icons ~ Creative Market  Machine Learning with Python  Beginner Level Training | Objective  A course on Introduction to Python, Basic Statistics, Machine Learning Algorithms, Building and validating machine learning models in python  Venkata Reddy  Data Scientist / Corporate Trainer/ Author |

Contents

[Machine Learning with Python 2](#_Toc176559)

[What will you learn at the end of this training? 2](#_Toc176560)

[Pre-requisites for the training 2](#_Toc176561)

[Training Details and Features 2](#_Toc176562)

[Trainer Profile 2](#_Toc176563)

[Course Curriculum 3](#_Toc176564)

[Day-1: Python Programming Basic Statistics & Regression 3](#_Toc176565)

[Day-2: Classification using Logistic Regression and Trees 3](#_Toc176566)

[Day-3: Model Validation Techniques and Neural Networks 4](#_Toc176567)

[Appendix 4](#_Toc176568)

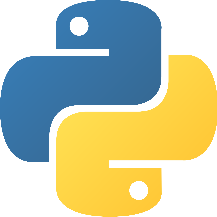
[List of Case Studies used in the course 4](#_Toc176569)

# Machine Learning with Python

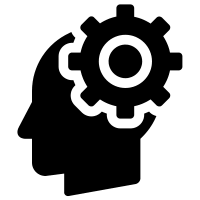
|  |  |
| --- | --- |
| C:\Users\StatInfer\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\4A018B81.tmp | What will you learn at the end of this training?  * Python programming basics & Data Handling in Python * Basic Statistics and Reporting * Machine Learning tools and techniques * Basic and advanced algorithms in Machine Learning * Model building and Validation techniques |
| C:\Users\StatInfer\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\AB371445.tmp | Pre-requisites for the training  * Minimum BSc/ BTech * Basic mathematical skills (12th class mathematics) * Knowledge on statistics is **NOT** necessary  Training Details and Features  * 100% hands on sessions * Duration –24 Hours – 3Days * Learning through problem solving and case studies |
| C:\Users\StatInfer\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\C4B0A4F9.tmp | Trainer Profile  * Venkata Reddy Konasani * 10+ years – Data Scientist / 5+ years – Corporate Training * Conducted 5000+ hours training and 60+ corporate batches * Author of the book “Practical Business Analytics using SAS” * Rich industry experience as applied Data Analyst and Data Scientist * Experience in credit risk model building, market response model building, social media analytics, revenue forecasting and machine learning * Post Graduate in Applied Statistics and Informatics from IIT Bombay * <https://statinfer.com/venkat-profile/> |

# Course Curriculum

### Day-1: Python Programming Basic Statistics & Regression

1. Introduction to Machine Learning
   1. Introduction to Machine Learning
   2. Machine Learning tools and Techniques
2. Data Handling, Data Validation and Graphs
   1. Important packages used in Machine Learning
   2. Data importing
   3. Working with datasets
   4. Descriptive statistics
   5. Central Tendency
   6. Variance
   7. Percentiles
   8. Outlier detection
   9. Variable distribution charts
3. Regression Analysis
   1. Correlation
   2. Simple Regression models
   3. R-Square
   4. Multiple regression
   5. Multicollinearity
   6. Individual Variable Impact
   7. Air passenger’s data case study
   8. SAT score data case study

### Day-2: Classification using Logistic Regression and Trees

* Logistic Regression
  + Need of logistic Regression
  + Logistic regression models
  + Validation of logistic regression models
  + Multicollinearity in logistic regression
  + Individual Impact of variables
  + Confusion Matrix
  + Service Provider Attrition data case study
* Decision Trees
  + Segmentation
  + Entropy
  + Information gain
  + Building Decision Trees
  + Validation of Trees
  + Pruning the trees
  + Fine tuning the trees
  + Prediction using Trees
  + Customer retention case study
* Cluster Analysis
  + Supervised vs unsupervised learning
  + Need of Cluster Analysis
  + K- Means clustering algorithm
  + The theory behind cluster Analysis
  + Building and interpreting clusters

### Day-3: Model Validation Techniques and Neural Networks

* Model Selection and Cross validation
  + How to validate a model?
  + What is a best model?
  + Types of data
  + Types of errors
  + The problem of over fitting
  + The problem of under fitting
  + Bias Variance Trade-off
  + Cross validation techniques
* Neural Networks
  + Neural network Intuition
  + Neural network and vocabulary
  + Neural network algorithm
  + Math behind neural network algorithm
  + Building the neural networks
  + Validating the neural network model
  + Neural network applications
  + Image recognition using neural networks
* Course Conclusion and Q&A
  + Couse conclusion
  + Reference books, videos and blogs
  + Next steps
  + Final Q&A
  + Fina assessment (optional)

## Appendix

### List of Case Studies used in the course

1. Online Sales data case study – Data Handling in Python
2. USA Census Income data analysis – Basic statistics
3. Air passenger prediction and driver analysis -Regression
4. SAT score prediction and driver analysis -Regression
5. E-com Website sales prediction case study -Regression
6. Product sales analysis – Logistic Regression
7. Customer attrition analysis -Logistic Regression
8. Customer Survey Segmentation and Drivers – Decision Trees
9. Internet service provider customer segmentation – Decision Trees
10. Retail customer segments – Cluster Analysis
11. Customer attrition analysis – Model selection and cross validation
12. Productivity data -Neural networks
13. Image recognition -Neural networks