

DATA SCIENCE, DEEP LEARNING AND AI

By Dr. Vishwanath Rao

Deep Learning: A revolution in Artificial Intelligence

-
-
- Limitations of Machine Learning

What is Deep Learning?

- Need for Data Scientists
- Foundation of Data Science
- What is Business Intelligence
- What is Data Analysis
- What is Data Mining

What is Machine Learning?

Analytics vs Data Science

- Value Chain
- Types of Analytics
- Lifecycle Probability
- Analytics Project Lifecycle
- Advantage of Deep Learning over Machine learning
- Reasons for Deep Learning
- Real-Life use cases of Deep Learning
- Review of Machine Learning

Data

- Basis of Data Categorization
- Types of Data
- Data Collection Types
- Forms of Data & Sources
- Data Quality & Changes
- Data Quality Issues
- Data Quality Story
- What is Data Architecture
- Components of Data Architecture

- OLTP vs OLAP
- How is Data Stored?

Big Data

- What is Big Data?
- 5 Vs of Big Data
- Big Data Architecture
- Big Data Technologies
- Big Data Challenge
- Big Data Requirements
- Big Data Distributed Computing & Complexity
- Hadoop
- Map Reduce Framework
- Hadoop Ecosystem

Data Science Deep Dive

- What Data Science is
- Why Data Scientists are in demand
- What is a Data Product
- The growing need for Data Science
- Large Scale Analysis Cost vs Storage
- Data Science Skills
- Data Science Use Cases
- Data Science Project Life Cycle & Stages
- Data Acquisition
- Where to source data
- Techniques
- Evaluating input data
- Data formats
- Data Quantity
- Data Quality
- Resolution Techniques
- Data Transformation
- File format Conversions
- Anonymization

Python

- Python Overview
- About Interpreted Languages
- Advantages/Disadvantages of Python pydoc.
- Starting Python
- Interpreter PATH
- Using the Interpreter
- Running a Python Script
- Using Variables
- Keywords
- Built-in Functions
- StringsDifferent Literals
- Math Operators and Expressions
- Writing to the Screen
- String Formatting
- Command Line Parameters and Flow Control.
- Lists
- Tuples
- Indexing and Slicing
- Iterating through a Sequence
- Functions for all Sequences

Operators and Keywords for Sequences

- The xrange() function
- List Comprehensions
- Generator Expressions
- Dictionaries and Sets.

Numpy & Pandas

- Learning NumPy
- Introduction to Pandas
- Creating Data Frames
- GroupingSorting
- Plotting Data
- Creating Functions
- Slicing/Dicing Operations.

Deep Dive – Functions & Classes & Oops

- Functions
- Function Parameters
- Global Variables
- Variable Scope and Returning Values. Sorting
- Alternate Keys
- Lambda Functions
- Sorting Collections of Collections
- Classes & OOPs

Statistics

- What is Statistics
- Descriptive Statistics
- Central Tendency Measures
- The Story of Average
- Dispersion Measures
- Data Distributions
- Central Limit Theorem
- What is Sampling
- Why Sampling
- Sampling Methods
- Inferential Statistics
- What is Hypothesis testing
- Confidence Level
- Degrees of freedom
- what is pValue
- Chi-Square test
- What is ANOVA
- Correlation vs Regression
- Uses of Correlation & Regression

Introduction

- ML Fundamentals
- ML Common Use Cases
- Understanding Supervised and Unsupervised Learning Techniques

Clustering

- Similarity Metrics
- Distance Measure Types: Euclidean, Cosine Measures

- Creating predictive models
- Understanding K-Means Clustering
- Understanding TF-IDF, Cosine Similarity and their application to Vector Space Model
- Case study

Implementing Association rule mining

- What is Association Rules & its use cases?
- What is Recommendation Engine & it's working?
- Recommendation Use-case
- Case study

Understanding Process flow of Supervised Learning Techniques

Decision Tree Classifier

- How to build Decision trees
- What is Classification and its use cases?
- What is Decision Tree?
- Algorithm for Decision Tree Induction
- Creating a Decision Tree
- Confusion Matrix
- Case study

Random Forest Classifier

- What is Random Forests
- Features of Random Forest
- Out of Box Error Estimate and Variable Importance
- Case study

Naive Bayes Classifier.

- Case study

Problem Statement and Analysis

- Various approaches to solve a Data Science Problem

- Pros and Cons of different approaches and algorithms.

Linear Regression

- Case study
- Introduction to Predictive Modeling
- Linear Regression Overview
- Simple Linear Regression
- Multiple Linear Regression

Logistic Regression

- Case study
- Logistic Regression Overview
- Data Partitioning
- Univariate Analysis
- Bivariate Analysis
- Multicollinearity Analysis
- Model Building
- Model Validation
- Model Performance Assessment AUC & ROC curves
- Scorecard

Support Vector Machines

- Case Study
- Introduction to SVMs
- SVM History
- Vectors Overview
- Decision Surfaces
- Linear SVMs
- The Kernel Trick
- Non-Linear SVMs
- The Kernel SVM

Time Series Analysis

- Describe Time Series data
- Format your Time Series data

- List the different components of Time Series data
- Discuss different kind of Time Series scenarios
- Choose the model according to the Time series scenario
- Implement the model for forecasting
- Explain working and implementation of ARIMA model
- Illustrate the working and implementation of different ETS models
- Forecast the data using the respective model
- What is Time Series data?
- Time Series variables
- Different components of Time Series data
- Visualize the data to identify Time Series Components
- Implement ARIMA model for forecasting
- Exponential smoothing models
- Identifying different time series scenario based on which different Exponential Smoothing model can be applied
- Implement respective model for forecasting
- Visualizing and formatting Time Series data
- Plotting decomposed Time Series data plot
- Applying ARIMA and ETS model for Time Series forecasting
- Forecasting for given Time period
- Case Study

Machine learning algorithms Python

- Various machine learning algorithms in Python
- Apply machine learning algorithms in Python

Feature Selection and Pre-processing

- How to select the right data
- Which are the best features to use
- Additional feature selection techniques
- A feature selection case study
- Preprocessing
- Preprocessing Scaling Techniques
- How to preprocess your data
- How to scale your data
- Feature Scaling Final Project

Which Algorithms perform best

- Highly efficient machine learning algorithms

- Bagging Decision Trees
- The power of ensembles
- Random Forest Ensemble technique
- Boosting – Adaboost
- Boosting ensemble stochastic gradient boosting
- A final ensemble technique

Model selection cross validation score

- Introduction Model Tuning
- Parameter Tuning GridSearchCV
- A second method to tune your algorithm
- How to automate machine learning
- Which ML algo should you choose
- How to compare machine learning algorithms in practice

Text Mining& NLP

- Sentimental Analysis
- Case study

PySpark and MLlib

- Introduction to Spark Core
- Spark Architecture
- Working with RDDs
- Introduction to PySpark
- Machine learning with PySpark – Mllib

Deep Learning & AI using Python

Deep Learning & AI

- Case Study
- Deep Learning Overview
- The Brain vs Neuron
- Introduction to Deep Learning

Introduction to Artificial Neural Networks

- The Detailed ANN
- The Activation Functions
- How do ANNs work & learn
- Gradient Descent
- Stochastic Gradient Descent
- Backpropagation
- Understand limitations of a Single Perceptron
- Understand Neural Networks in Detail
- Illustrate Multi-Layer Perceptron
- Backpropagation – Learning Algorithm
- Understand Backpropagation – Using Neural Network Example
- MLP Digit-Classifer using TensorFlow
- Building a multi-layered perceptron for classification
- Why Deep Networks
- Why Deep Networks give better accuracy?
- Use-Case Implementation
- Understand How Deep Network Works?
- How Backpropagation Works?
- Illustrate Forward pass, Backward pass
- Different variants of Gradient Descent

Convolutional Neural Networks

- Convolutional Operation
- Relu Layers
- What is Pooling vs Flattening
- Full Connection
- Softmax vs Cross Entropy
- " Building a real world convolutional neural network
- for image classification"

What are RNNs – Introduction to RNNs

- Recurrent neural networks rnn
- LSTMs understanding LSTMs
- long short term memory neural networks lstm in python

Restricted Boltzmann Machine (RBM) and Autoencoders

- Restricted Boltzmann Machine
- Applications of RBM
- Introduction to Autoencoders

- Autoencoders applications
- Understanding Autoencoders
- Building a Autoencoder model

Tensorflow with Python

- Introducing Tensorflow
- Introducing Tensorflow
- Why Tensorflow?
- What is tensorflow?
- Tensorflow as an Interface
- Tensorflow as an environment
- Tensors
- Computation Graph
- Installing Tensorflow
- Tensorflow training
- Prepare Data
- Tensor types
- Loss and Optimization
- Running tensorflow programs

Building Neural Networks using

Tensorflow

- Tensors
- Tensorflow data types
- CPU vs GPU vs TPU
- Tensorflow methods
- Introduction to Neural Networks
- Neural Network Architecture
- Linear Regression example revisited
- The Neuron
- Neural Network Layers
- The MNIST Dataset
- Coding MNIST NN

Deep Learning using

Tensorflow

- Deepening the network
- Images and Pixels
- How humans recognise images
- Convolutional Neural Networks
- ConvNet Architecture
- Overfitting and Regularization
- Max Pooling and ReLU activations
- Dropout
- Strides and Zero Padding
- Coding Deep ConvNets demo
- Debugging Neural Networks
- Visualising NN using Tensorflow
- Tensorboard

Transfer Learning using Keras and TFLearn

- Transfer Learning Introduction
- Google Inception Model
- Retraining Google Inception with our own data demo
- Predicting new images
- Transfer Learning Summary
- Extending Tensorflow
- Keras
- TFLearn
- Keras vs TFLearn Comparison