Machine Learning Engineering

1. Supervised Learning :-

Regression

  - Linear Regression
  - Logistic Regression
  - Polynomial Regression
  - Ridge Regression & Lasso Regression

  Working
  Math behind the Intuition
  Learning the concepts of Coefficient and Residuals
  Cost function
  Feature scaling
  Non-Linearity and non-Invertibility

  Optimizing Linear Functions
  Standard Error
  Gradient decent intuation
  Hypothesis Representation
  Regularized Regressions
  Regularization
  L1 and L2 Regularization

  Filter method
  Wrapper method
  Embedded Method

  Decision Boundary
  Case study using SciKit Learn
Classification

KNN

Intuition
Eager and Lazy Classifiers
Other names of KNN classifiers
How to Choose k?
Distance metrics used in KNN
Mathematically Demystifying KNN Algorithm
Weighted KNN
Characteristics of KNN Algorithm
Strength and weakness
Weighted KNN
Improvements of KNN performance
Fuzzy KNN
Case Study using SciKit Learn
Applying cross validation techniques and analyzing the Algorithm behaviour.
Improvisation on the Algorithm

Support Vector Machine

Intuition
Visualize in Vector space
Large Margin Intuition
Significance of Binary Labels [+1,-1]
Inequalities and region
Maximum Margin: Formalization
Linear Support Vector machine
Non-Linear SVM
Hard Margin and Soft Margin
Kernel Tricks
C parameter?
Decision Functions
Multiclass Problem
Challenges on Multiclass classification
Polynomial Kernel
Gaussian RBF Kernel
SVR
Kernelized SVM
Tweak Performance
Upweighting
Drift Problem
Case Study using SciKit Learn
Strength and weakness

Naive Bayes Algorithm
Intuition
Demystifying Probability
Conditional Probability
Bayes Theorem
Estimation of probability for the Dataset
Likelihoods
Gaussian, Bernoulli, Multinomial.
Discriminant Functions
Expectation Maximization Algorithm – EM
Case Study using SciKit Learn
Strength and weakness
Decision Tree Classifiers

Intuition
Training and Visualization
Predictions
Estimating Class Probabilities
Computational Complexity
CART Algorithm
HUNTS Algorithm
Gini Index, Entropy and Classification Error
Bagging and Bootstrapping
Regularization Hyperparameters
Case Study using SciKit Learn
Data Fragmentation
Tree Replication

Ensemble Learning & Random Forest

Intuition
Voting Classifiers
Bagging and Pasting in Scikit-Learn
Out-of-Bag Evaluation
Random Patches and Random Subspaces
Random Forests
Boosting
AdaBoost
Gradient Boosting
Stacking
XGBoost
2. Unsupervised Learning :-

Introduction to clustering
Types of Clustering
Optimizing Objective
Data Characteristics
Prototype Based Approach
  o K Means
  Improvised K-Means Paper Implementations
Graph Based Approach
  o Hierarchical Clustering
Density Based Approach
  o DBSCAN

K MEANS

Intuition
Prototype Based Approach
Mathematically Demystifying KNN Algorithm
Expoloring K
Elbow Method
Characteristics of K-Means Clustering
Random Initialization
Data compression techniques
Distance Metrics for K Means
Strength and Weakness
Time and Space Complexity
Performance Evaluation
Improvised K-Means Implementations
Case Study using SciKit-Learn

Hierarchical Clustering
Intuition
Graph based approach
Agglomerative and Divisive
Dendrograms
Proximity Methods
Strength and Weakness
Time and Space Complexity

DBSCAN
Intuition
Density Based Approach
Mathematically Demystifying DBSCAN Algorithm
Analyzing Core points Border Point and Noise Points
Clustering Tendency
Cluster Evaluation Metrics
Cohesion and Separation
Silhouette coefficient
Time and space complexity
Strength and Weakness
Case Study using SciKit-Learn
Distance Measurements

Euclidian Distances
Squared Euclidian Distances
Manhattan Distance
Minkowski distance
Cosine measure
Jaccard distance

Data Mining

Association Analysis –Apriori Algorithm
Anomaly Detection

Natural Language Processing

Python Texts
Working with Text PDF and Other Files
Spacy Basics, NLTK
Stemming, Lemmatization, Stop Words
Speech Tagging and Named Entity Recognition
Text Classification
Semantics and Sentimental Analysis
Topics Modeling’s

Case Study using TensorFlow

Case Studies And Implementation Of NLP:

Sentimental Analysis
Chabots
Deep Learning Engineering

Artificial Neural Network - ANN

- Transforming Biological neuron to Artificial Neurons
- Logical Computations with Neurons
- Single Layer Perceptron
- Sequential Modelling
- Multi-Layer Perceptron
- Activation Functions
- Loss functions
- Vanishing/Exploding Gradients Problems
- Batch Normalization
- Learning Rates
- Train Test and Validation
- Overfitting and Underfitting Problems
- Dealing with Data Augmentation
- One Hot Encoding
- Dropout
- Gradient Clipping

Convolution Neural Network - CNN

- Intuition
- Zero Padding
- Convolution Layers
- Max Pooling
- Back Propagation
- Weights and its Importance
- Classification MLPs
- Backpropagation
Dealing with Augmented Data
Reusing Pretrained Layers
Transfer Learning with Keras
Unsupervised Pretraining
Faster Optimizers
Momentum Optimization
Batch Size
Max-Norm Regularization
Fine Tuning
CNN Architectures
Self-Organizing Maps
Boltzmann Mechanism
Autoencoders

Recurrent Neural Network and LSTM

Intuition
RNN
Bidirectional RNN’S
LSTM
Memory Requirements

Deep Diving into Neural Network

Discussion on LeNet-5
Discussion on AlexNet
Discussion on GoogLeNet
Discussion on VGGNet
Discussion on ResNet
Using Pretrained Models from Keras
Pretrained Models for Transfer Learning
Classification and Localization
Object Detection
Fully Convolutional Networks (FCNs)
You Only Look Once (YOLO)
Time Series Analysis
Generative Adversarial Network
Deploying Deep Learning Models using Django

Exploring Neural Network using TensorFlow and Keras

TensorFlow and Keras Initiation
Tensors and Operations
Tensors with NumPy
Placeholders
Type Conversions
Variables
Data Structures Indepth
TensorFlow Functions and Graphs
Building an Image Classifier
Using Sequential API to build Regression MLP
Using Sequential API to build Complex Models
Using the Subclassing API Saving and Restoring a Model
Implementing Callbacks
Visualization Using Tensor Board
Fine-Tuning Neural Network Hyperparameters
Hidden Layers
Learning Rate, Batch Size and Other Hyperparameters
Customizing Metrics, Layers, Training Loops, Models and Training Algorithms
Custom Loss Functions
Autograph and Tracing
TF Function Rules

=================================================================
Fine Tuning Data Science Algorithm

Demystifying Diagnostic Report for Learning Models

Feature Engineering and Model Selection
Underfitting and overfitting
Bias, Variance Trade-off / F1scores
Confusion matrix
Accuracy metrics
Univariate, Bivariate, Multivariate Dataset
Evaluating machine learning model
ROC Curves
Hyper parameter tuning
Importance of Data and its quality
Attributes Types
Feature selection and Feature extraction
Stepwise Selection
Loss Function
Curse of Dimensionality
ChiSquare Test
Impact on Outliers
Cohen’s D Statistics
Error Analysis
General Distance metrics
Graph analysis on Datasets
Regularization
MSE, RMSE, MSE
Feature Slicing
Correlation and Causation
Training /Validation /Testing Data
Learning Rate
Confidence Intervals
Degree of Freedom
Dimensionality Reduction Techniques

1. PCA
2. LDA
3. QDA

Intuition on Dimensionality Reduction
Geometrical intuition.
Alternative formulation of PCA: distance minimization
Eigenvalues and eigenvectors.
PCA for dimensionality reduction and visualization.
Visualize MNIST dataset.
Limitations of PCA
Ts-SNE Estimator for Dimensionality Reduction
Impact on Algorithm

Cross Validation Techniques

Holdout Method
K-Fold Cross Validation
Stratified K-Fold Cross Validation
Leave-One-Out Cross Validation