Hachion

Machine Learning Course Content

Course Description:

Machine learning is a branch of artificial intelligence technology that combines a set of algorithms and statistical models that are used for future predictions and identifying patterns in data. Machine learning enables computer systems to learn automatically with less human involvement. ML is leading technology applicable in the fields of data science, analytics, business intelligence, search engine, and e-commerce.

Machine learning with AI and data mining technology makes a more effective way to analyze huge quantities of data. ML applies complex calculations to big data to produce reliable, faster, and accurate results. Machine Learning is a fast-growing technology that has much future scope with many job opportunities.

Hachion Machine Learning online training offers in-depth knowledge of all basic and advanced topics. Our ML tutorial includes all key modules such as python, algorithms, supervised & unsupervised learning, statistics & probability, decision trees, random forests, linear & logistic regression. Beginners can also learn ML easily to get placed as a Machine Learning Engineer.

Course Content:

Basics of Python

- Keywords and Identifiers
- Comments, Indentations, and Statements
- Variables and Data Types in Python
- Standard Input and Output
- Operators Control Flow: If Else
- Control Flow: While Loop
- Control Flow: For Loop
- Control Flow: Break and Continue
- Lists, Tuples Part
- Tuples Part 2: Sets
- Dictionary Strings
- Types of Functions and Function Arguments
- Recursive Functions, Lambda Functions, Modules
- Packages, File Handling
- Exception Handling, Debugging Python

Numpy

- Numpy Introduction
- Numerical Operations on Numpy

Exploratory Data Analysis

- Need and Use of EDA
- Exploring the IRIRS Dataset
- 2D Scatter Plot
- 3D Scatter Plot
- Pair Plots, Histogram
- PDF, Univariate Analysis using PDF
- CDF - Cumulative Distributive Function
- Mean, Variance, and Standard Deviation, Median
- Percentiles and Quantiles IQR (Inter Quartile Range)
- MAD (Mean Absolute Deviation)
- Box- Plot with Whiskers, Violin Plots
- Univariate, Bivariate and Multi-Variate Analysis
- Multivariate Probability Density, Contour Plot

Dimensionality Reduction and Visualisation

- Introduction to Dimensionality Reduction
- Representing Datasets using Row and Column Vectors
- Representation of Datasets as a Matrix
- Data Pre Processing - Feature Normalization
- Mean of Data Matrix
- Column Standardization
- Co-Variance of Data Matrix
- PCA, PCA with a Code Example

Principle Component Analysis

- Introduction and use of PCA, Geometric Intuition of PCA
- The mathematical objective function of PCA
- Distance Minimization
- Eigen Values and Eigen Vectors (PCA): Dimensionality Reduction
- PCA for Dimensionality Reduction and Visualization
- Limitation of PCA and PCA with a Code Example
- Supervised Learning

Linear Regression

- Geometric Intuition of Logical Regression
- Squashing using Sigmoid Function
- Objective Function mathematical formulation
- Weight Vector
- L2 Regularization: Overfitting and Underfitting
- L1 Regularization and sparsity
- Probabilistic Interpretation: Gaussian Naive Bayes Loss minimization representation
- Hyperparameter Search: Grid Search and Random Search

- Column Standardization
- Feature Importance and Model Interpretability
- Collinearity of Features
- Test/ Run Time Space and Time Complexity
- Real-World Cases
- Non-Linearly separable data and Feature Engineering

Logistic Regression

- GridSearchCV, RandomSearchCV
- Extensions to Logistic Regression: Generalised Linear Models

Neural Networks

- Working of Biological Neurons
- Growth of Biological Neural Networks
- Diagrammatic representation: Logistic Regression and Perceptron, Multi-Layered Perceptron (MLP)
- Notation, Training a Single-Neuron Model
- Training an MLP: Chain Rule Training an MLP: Memorization, Back Propagation, Activation Functions
- Vanishing Gradient Problem
- Bias-Variance tradeoff, Decision Surfaces, Playground

Decision Trees

- Axis Parallel Hyperplanes, Sample Decision Tree
- Building a Decision Entropy
- Building a Decision Tree: Information Gain
- Building a Decision Tree: Gini Impurity
- Building a Decision Tree: Constructing a DT
- Building a Decision Tree: Splitting Numerical Features, Features Standardization
- Building a Decision Tree: Categorical Features with many possible values
- Overfitting and Underfitting
- Train and Run Time Complexity
- Regression using Decision Trees
- Cases, Code Samples

Naive Bayes

- Conditional Probability
- Independent Vs. Mutually Exclusive Events
- Bayes Theorem with Examples
- Exercise Problems on Bayes Theorem
- Naive Bayes Algorithm
- Toy Example: Train and Test stages
- Naive Bayes on Test Data
- Laplace/ Additive Smoothing

- Log Probabilities for Numerical Stability
- Bias and Variance Tradeoff
- Feature Importance and Interpretability
- Imbalanced Data, Outliers, Missing Values
- Handling Numerical Features (Gaussian NB)
- Multiclass Classification, Similarity or Distance Matrix
- Large Dimensionality, Best and Worst Cases

Support Vector Machines

- Geometric Intuition
- Why we take values of +1 and -1 for Support Vector Planes
- Mathematical derivation
- Loss Function (Hinge Loss) based Interpretation
- Dual Form of SVM Formulation
- Kernel Trick, Polynomial Kernel, RBF-Kernel
- Domain-specific Kernels
- Train and Run Time Complexities
- nu-SVM: Control Errors and Support Vectors
- SVM Regression Cases

Unsupervised Learning

- Clustering
- What is Clustering?
- Unsupervised Learning
- Applications
- Metrics for Clustering
- K-Means: Geometric Intuition
- Centroids
- K-Means: Mathematical Formulation: Objective Function
- K-Means: Algorithm
- How to initiate K-Means++.
- Failure Cases/ Limitations
- K-Medoids?
- Determining the Right K
- Code Samples
- Time and Space Complexity

Hierarchical Clustering

- Agglomerative and Divisive
- Dendrograms
- Agglomerative Clustering
- Proximity Methods: Advantages and Limitations

- Time and Space Complexity
- Limitations of Hierarchical Clustering
- Code Sample

DBSCAN Technique

- Density-based Clustering
- MinPts and Eps: Density
- Core
- Border and Noise Points
- Density Edge and Density Connected Points
- DBSCAN Algorithm
- Hyper Parameters: MinPts and Eps