MACHINE LEARNING ADVANCED

By Dr. Vishwanath Rao

Description:

■ This a basic session outlining the field of Artificial Intelligence. The idea behind the training is to give the partcipants a solid demonstration of State-of-the-art Artificial Intelligence. This will serve as an inspiration for those who want to pursue this field and paint with broad strokes for those who want to casually know basic concepts . The rest of the training is devided between Engineering(building data processing pipeline from scratch) and Algorithms(Linear and Logistic Regression).

Key Skills:

- Solid foundation of the basic Engineering that goes behind Machine Learning
- An idea of what State-of-the-art Artificial Intelligence can achieve
- Solid foundation of some of the Algorithms
  Prerequisites:
- Basic idea of Differential Calculus will be an advantage
- Working knowledge of python
- Basic idea of Lenear Algebra will be an advantage
  Instructional Method:
- *Introduction*
  - Model selection
  - Supervised learning
  - Discovering graph structure
  - Types of machine learning
  - Machine learning: what and why?
  - Parametric vs non-parametric models

    - No free lunch theorem
    - Linear regression
    - Some basic concepts in machine learning
    - Discovering clusters
    - Classification
    - Regression
    - Matrix completion
    - Parametric models for classification and regression
    - Logistic regression
    - The curse of dimensionality
    - Overfitting
    - Unsupervised learning
    - Discovering latent factors
    - A simple non-parametric classifier: K-nearest neighbors
- *Case Studies and Live Demos*

■ Hidden Markov Models
- Evaluation of an HMM
- Generating an observation sequence
- Extensions of HMM
- Tree-structured HMM
- The HMM
- Markov models

- Factorial HMMs
- State space models
- Parameterization of HMM

■ Transformer
• Speech Recognition • Visualization
- Neural Machine Translation
    - German to English
    - Hindi to English
    - French to English
    - Visualization
- Autonomous Driving
  Engineering
- Error-based Learning
    - Setting the Learning Rate Using Weight Decay
    - Error Surfaces
    - Multinomial Logistic Regression
    - Modeling Non-linear Relationships
    - Handling Categorical Descriptive Features
    - Interpreting Multivariable Linear Regression Models
    - Simple Linear Regression
    - Big Idea
    - Handling Categorical Target Features: Logistic Regression
    - Extensions and Variations
    - Fundamentals

    - Choosing Learning Rates and Initial Weights
    - Standard Approach: Multivariable Linear Regression with Gradient Descent
    - Gradient Descent
    - Multivariable Linear Regression
    - Measuring Error
- Similarity-based Learning
- Standard Approach: The Nearest Neighbor Algorithm
- Predicting Continuous Targets
- Fundamentals
- Other Measures of Similarity
- Extensions and Variations
- Data Normalization
- Feature Space
- Big Idea
- Measuring Similarity Using Distance Metrics
- Feature Selection
- Handling Noisy Data
- Efficient Memory Search
- Machine Learning for Predictive Data Analytics
    - Predictive Data Analytics Tools
    - How Does Machine Learning Work?
    - The Road Ahead

    - What Can Go Wrong with Machine Learning?
    - The Predictive Data Analytics Project Lifecycle: CRISP-DM
    - What Is Machine Learning?
    - What Is Predictive Data Analytics?
- Data to Insights to Decisions
    - Different Types of Data
    - Different Types of Features
    - Designing the Analytics Base Table
    - Designing and Implementing Features
    - Assessing Feasibility
    - Converting Business Problems into Analytics Solutions
    - Case Study: Motor Insurance Fraudmotor
    - Implementing Features
    - Handling Time
- Data Exploration • Outliers
- Handling Missing Values
- Handling Outliers
- Missing Values
- Irregular Cardinality
- Handling Data Quality Issues
- The Data Quality Report
- The Normal Distribution

- Identifying Data Quality Issues
- Getting to Know the Data

■ Advanced Data Exploration
- Measuring Covariance and Correlation
- Visualizing Relationships Between Features
- Binning
- Data Preparation
- Normalization
- Information-based Learning
    - Shannon’s Entropy Model
    - Handling Continuous Descriptive Features
    - Decision Trees
    - Predicting Continuous Targets
    - Extensions and Variations
    - Fundamentals
    - Information Gain
    - Big Idea
    - Standard Approach: The ID Algorithm
    - Tree Pruning
    - Alternative Feature Selection and Impurity Metrics
- Evaluation
    - Performance Measures: Prediction Scores
    - Designing Evaluation Experiments

    - Evaluating Models after Deployment
    - Performance Measures: Multinomial Targets

- Extensions and Variations
- Fundamentals
- Performance Measures: Continuous Targets
- Performance Measures: Categorical Targets
- Big Idea
- Standard Approach: Misclassification Rate on a Hold-out Test Set
  Algorithms
- Linear regression
    - Regularization effects of big data
    - Bayesian inference when ?^2 is unknown *
    - Model specification
    - Numerically stable computation *
    - Computing the posterior
    - Geometric interpretation
    - Convexity
    - Connection with PCA *
    - Maximum likelihood estimation (least squares)
    - Bayesian linear regression
    - Derivation of the MLE
    - Computing the posterior predictive

    - EB for linear regression (evidence procedure)
    - Ridge regression
    - Basic idea
    - Robust linear regression *
    - Introduction
- Logistic regression
    - Residual analysis (outlier detection) *
    - Generative vs discriminative classifier
    - Multi-class logistic regression
    - Online learning and regret minimization
    - Iteratively reweighted least squares (IRLS)
    - Quasi-Newton (variable metric) methods
- Newton’s method
- Bayesian logistic regression
- A Bayesian view
- Laplace approximation
- l2 regularization
- Gaussian approximation for logistic regression
- Approximating the posterior predictive
- Derivation of the BIC
- Steepest descent
- Introduction

- MLE
- Model specification
- Online learning and stochastic optimization
- Dealing with missing data
- Fisher’s linear discriminant analysis (FLDA) *
- Model fitting
- Stochastic optimization and risk minimization
- Pros and cons of each approach
- The LMS algorithm
- Logistic regression
- The perceptron algorithm