skillspeed
for the serious learner

Machine Learning - Curriculum

Prerequisites:

- Python 2.7
- Numpy
- Plotting and visualizing with Matplotlib

Module 1. Machine Learning Foundations (4 Hours)

- Intro to problem solving using ML with focus on nature of machine learning methods and what kind of tools are appropriate for what type of problems. Intro to various approaches & their applicability - Search based, decision tree, probabilistic models, logical agents, parametric models, reinforcement learning and genetic algorithms – 1 Hour
- Introduction & applications of Probability and Probability Distributions – 1 Hour
- Practical – 2 Hours
- Digit recognition using probability distribution approach

Module 2. Introduction to Machine Learning (4 Hours)

- Big picture view of ML & AI State-of-the-art
    - Types of ML approaches – 0.5 Hour
    - Supervised Learning (Classification vs Regression), Unsupervised Learning Intro
- Fundamental concepts in ML - 2 Hours
    - Formulating the learning problem
    - Feature Engineering
    - Bias and Variance
    - Overfitting vs Under fitting (Cross Validation)
    - Regularization
- Practical - Coding a simple machine learning model from scratch to appreciate the nature of ML programs and the associated steps – 1.5 Hours

Module 3. Supervised Learning – Regression (4 Hours)

- Regression Techniques- 2 Hours
    - Regression problems and applications. Discuss interesting problems in computer vision (human pose estimation) and regression with time-series data (sequences).
- Practical - Application + Exercises – 2 Hours

Module 4. Supervised learning – Classification (4 Hours)

- Decision Trees - 1.5 Hours
    - Introducing the concept of classification using a decision tree approach. Splitting criteria based on Gini Index / Entropy.
    - Covering both Classification Trees and Regression Trees.
- Random Forests - 0.5 Hour
    - Moving from a tree to building a forest. Gives an introduction into building ensembles of model. Introducing the concept of randomization.
- Logistic Regression Intro - 2 Hours
    - MLE
    - Understanding the logistic regression output
    - Confusion matrix (and troubleshooting)
- Offline Exercise

Module 5. Unsupervised Methods (4 Hours)

- Unsupervised clustering - 2 Hours (Unsupervised)
    - k-means
    - Gaussian Mixture Models
    - hierarchical
    - agglomeration
    - Spectral Clustering
- Practical - Example using GMM to solve real world problems

Module 6. Neural Networks (4 Hours)

- Introduction to Neural Networks
    - How they work and why they work?
    - Understanding why we need multiple layers and multiple neurons in each layer
- Understanding Back-propagation with an introduction to computational graphs (Very important to understand how tensorflow / theano frameworks do the automatic backprop)
- Practical - Introduction to Keras - Using Keras to build neural network models.

Module 7. Deep Learning (4 Hours)

- Introduction to deep learning - why they are powerful and where to and where not to use them
- Details of SGD (stochastic gradient descent) and variants to give insights on what option to choose from those supported in popular deep learning frameworks.
- Concepts of pre-training, regularisation and transfer learning
- Practical - Using Keras to build a deep learning model.

Module 8. Deep Learning - CNN & RNN (4 Hours)

- Understanding Convolutional Neural Networks - The intuition, type of problems to apply this on, how they are applied in practice
- Understanding models for dealing with sequential data - Recurrent Neural Networks. The intuitive understanding of how they work and how to use them for designing models for sequential data or for data with variable length inputs.
- Practicals - Using Keras to design and build models using CNN/RNN. Probably limit to formulating various problems as CNN/RNN models than actually getting into implementing and training them.