innovate achieve lead

BITS Pilani

# Machine Learning (IS ZC464) Session 16: Review Session

# Session 1

- Learning – experience from past
- Learning -- humans vs. machines
- Why machine learning?
- Machine learning and Artificial intelligence
- Intelligent systems
- Examples – vehicle recognition, fruit recognition and face recognition: attributes and their significance
- etc.

# Session 2

- Training and testing in learning systems
- Prediction – training and testing
- (X,Y) – input output pair
- Function approximation : given training (x,y) pairs
- Best hypothesis-minimum error tradeoff
- Learning parameters defining a hypothesis
- Line fitting – slope and intercept
- Squared mean error
- etc.

# Session 2 (continued)

- Error surface
- Local and global optima
- Generalization in function approximation
- Traditional vs. machine learning
- Glimpse of learning through artificial neural network
- etc.

# Session 3

- Uncertainty handling in real world using probability theory
- Types of uncertainty
- Types of random variables
- Probability theory
- Prior and posterior probability
- Joint and conditional probability
- Inference using full joint probability distribution
- Marginal probability and conditioning
- etc.

# Session 4

- Bayes' theorem
- Hypothesis in different contexts
- Best hypothesis
- Bayesian learning : example of sounds of words as classes
- Posterior probability computation: sound example-selection of maximum probability based class
- Maximum a posteriori hypothesis (MAP)
- etc.

# Session 5

- Bayes' optimal classifier
- Gibbs algorithm
- Naïve bayes' classifier
- Problem solving as discussed – bayes' theorem, MAP, entropy, etc.
- etc.

# Session 6

- Regression: understanding
- Linear models for regression
- Parametric form of the hypothesis
- Error surface construction from the parameters [ parameters→ hypothesis (prediction)→error (actual – predicted)]
- Visualization of error surface as contour plots
- Gradient descent algorithm
- Hypothesis for regression

  $y(x,w)= w_0 + w_1 x_1+ w_2 x_2+ w_3 x_3+ .....+w_D x_D$
- etc.

# Session 7

- Understanding of classification process
- Decision regions
- Binary classification
- Decision boundary
- Linearly separable data and Non linearly separable data
- Discriminant functions
- Work out on classification of a test sample (slide 11, 12)
- Decision boundary as $y = w^T x$
- Binary versus multiclass classification
- Linear versus circular boundaries
- Classification as an optimization process – (1) find W that gives minimum error (2) Find W that maximizes the separation of the classes
- etc.

# Session 8

- Decision trees
- Its limitations
- Attributes and their significance in decision tree based classification
- Leaf and non leaf nodes of a decision tree
- Information content – entropy of the given set of observations
- Splitting the observations based on decision key and attribute selected for splitting
- Gain and remainder computations
- etc.

# Session 9

- Human nervous system : some inspiration
- Artificial neuron
- Simulation  of a single neuron to perform logical AND-weight, threshold, weighted input, activation function, training and testing, unseen patterns for testing
- XOR simulation, why non linearly separable data, why two neurons, threshold, weights, activation functions for the required neurons
- etc.

# Session 10

- Neural network simulation for XOR and related computations
- Fruit recognition example and it neural network simulation-output modeling, weights etc
- Feed forwards neural network architecture
- Significance of layers
- Learning factors
- etc.

# Session 11

- Delta rule of weight update in gradient descent method
- Learning rate
- Weights as knowledge – classification of unseen pattern using acquired weights
- Chain rule of partial derivatives – used for computing the delta of weight + understanding of the derivation of the term for weight update
- etc.

# Session 12

- Evolutionary algorithms and their applications
- Population based algorithm
- Chromosome as bit string
- Representing rules as bit strings
- Search as combinatorial problem
- Fitness value
- Crossover and mutation
- etc.

# Session 13

- Roulette wheel
- Fittest parent-pool of parents
- Exploration and exploitation
- Application of genetic algorithm in unsupervised clustering, clustering as an optimization problem, fitness function for clustering
- Limitation of GA
- etc.

# Session 14

- Instance based learning
- Lazy learning
- K-Nearest Neighbor Learning
- Classification as mapping
- Locally Weighted Regression
- Case based reasoning – basic idea
- Generalization in classification
- Application of covers theorem in mapping data in higher dimensional space
- Radial basis functions neural networks
- etc.

# Session 15

- Support vector machines
- Support vectors – understanding
- Margin and decision boundary
- Computation of the margin
- What to optimize in SVM?
- Kernel functions
- etc.

# Syllabus for Comprehensive examination

- All topics and the corresponding depth of concepts covered during lecture sessions 1 to 16.