MLCC DAY 2

MACHINE LEARNING AND STATISITICS

TYPES OF MACHINE LEARNING

- Supervised Learning – Train me
  - Classification
  - Regression
    - Linear & Polynomial
- Unsupervised Learning – Self sufficient in learning
  - Clustering
  - Association Analysis
  - Hidden Markov Model
- Reinforcement Learning – My lyf mah rulz (Hit & Trial)

REGRESSION PROBLEMS -

- Linear Regression
- Polynomial Regression

- Regression is a way of finding pattern in known data and predict unforeseen data

CLASSIFICATION

- K – Nearest Neighbors
- Logistic Regression
- Decision Trees
- Ada Boost (Boosted Decision Tree)
- Support Vector Machines (SVMs)
- Random Forests
- Neural Networks

CLASSIFICATION

- 2 Class Classification – Only two labels to predict or only two possible outcomes
- Multi class classification – More than two possible outcomes

STATISTICAL LEARNING THEORY
OCKHAM'S RAZOR

- The best models are simple models that fit the data well
  - We need a balance between simplicity and accuracy
- William of Ockham (1287 – 1347) stated –
- "Among the hypothesis that predict equally well, we should choose one with the fewest assumptions

MEAN, MEDIAN, STANDARD DEVIATION

- Mean –
  - it is an average of all the values in the given column of data
- Median –
  - It is the middle value in the distribution
- Standard Deviation -
  - a quantity expressing by how much the members of a group differ from the mean value for the group. a quantity expressing by how much the members of a group differ from the mean value for the group.

VARIANCE

- It is the expectation of squared deviation of a random variable from the mean value
- It is square of Standard Deviation