# Machine Learning

## BS/MS (Computer Science)

IQRA UNIVERSITY IU

**Lecture-01**

**14-June-2014**

**Summer Semester**

# Course Description

# Course Aims/Objectives

- Provides an overview of the state-of-the-art in the areas of pattern recognition and machine learning
- Thorough understanding of the most successful methods used in Machine learning.
- This course helps BS/MS/Ph.D. students to understand the concepts of machine learning and its application in different domains of research.

# Learning Outcomes

- After successful completion of the course, the students are familiar with strategies to solve conceptual and theoretical problems in machine learning.
- They will have the ability to apply state-of-the-art algorithms, such as: neural network, support vector machine, kernel methods, boosting and mixtures of experts, to real-world problems.
- Familiar with the advantages and disadvantages of different algorithms.

# Course: Machine Learning

## Text book

1. C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006
2. Tom Mitchell, Machine Learning, McGraw-Hill, 1997
3. Ian H. Witten, Data Mining, Practical Machine Learning Tools and Techniques, 2011

## Grading

1. Assignments
2. Presentations
3. Quizzes
4. Project
5. Mid-Term/Hourly-Exams
6. Final Exam

# Resources: Conferences

- International Conference on Machine Learning (ICML)
  - ICML05: http://icml.ais.fraunhofer.de/
- European Conference on Machine Learning (ECML)
  - ECML05: http://ecmlpkdd05.liacc.up.pt/
- Neural Information Processing Systems (NIPS)
  - NIPS05: http://nips.cc/
- Uncertainty in Artificial Intelligence (UAI)
  - UAI05: http://www.cs.toronto.edu/uai2005/
- Computational Learning Theory (COLT)
  - COLT05: http://learningtheory.org/colt2005/
- International Joint Conference on Artificial Intelligence (IJCAI)
  - IJCAI05: http://ijcai05.csd.abdn.ac.uk/
- International Conference on Neural Networks (Europe)
  - ICANN05: http://www.ibspan.waw.pl/ICANN-2005/
- ...

# Marks Distribution

1. Assignments/Quizzes 10 marks
2. BS Students: Project (Presentation & Viva) 10 marks
   MS Students: Project (Presentation & Viva)+Research Paper
3. First Hourly 15 marks
4. Second Hourly 15 marks
5. Final Examination 50 marks

# News Group

Please make use of the following newsgroup for functional and organizational questions.

**IqraMachineLearning · IU-MachineLearning**

**Post message:**

IqraML@yahoogroups.com

**Subscribe:**

IqraML-subscribe@yahoogroups.com

**Unsubscribe:**

**IqraML-unsubscribe@yahoogroups.com**

# Requirements

- Basic Requirements
  - Read materials before/after class.
  - Work on your homework individually.
    - Discussions are encouraged but don't copy others' work.
  - **Get you hands dirty**!
    - Experiment with ideas presented in class and gain first-hand knowledge!
  - Come to class and **DON'T** hesitate to speak if you have any questions/comments/suggestions!
  - Student participation is important!

# Final Project

- Groups of 4-5 students
  - Pair Programming recommended!
- Topic
  - Problem of your choice (recommend project list will be provided)
  - Should be an interesting enough (non-trivial) problem
- Suggested environment
  - MATLAb
  - WEKA

# Schedule for Project/Paper Work

- Week/Lecture 3-4: Project Introduction + Paper Review
- Week/Lecture 5-6: First Presentation + Poster Presentation
- Week/Lecture 7-8: Final Presentation + Paper Presentation
- Week/Lecture 9-10: Submit Project Report + Paper Summery
- Week/Lecture 11-12:
  - Submit Assignments/Quizzes files
  - Submit any missing files
  - VIVA

# Resources: Journals

- Journal of Machine Learning Research www.jmlr.org
- Machine Learning
- Neural Computation
- Neural Networks
- IEEE Transactions on Neural Networks
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- Annals of Statistics
- Journal of the American Statistical Association
- ...

# Resources: Conferences

- International Conference on Machine Learning (ICML)
  - ICML05: http://icml.ais.fraunhofer.de/
- European Conference on Machine Learning (ECML)
  - ECML05: http://ecmlpkdd05.liacc.up.pt/
- Neural Information Processing Systems (NIPS)
  - NIPS05: http://nips.cc/
- Uncertainty in Artificial Intelligence (UAI)
  - UAI05: http://www.cs.toronto.edu/uai2005/
- Computational Learning Theory (COLT)
  - COLT05: http://learningtheory.org/colt2005/
- International Joint Conference on Artificial Intelligence (IJCAI)
  - IJCAI05: http://ijcai05.csd.abdn.ac.uk/
- International Conference on Neural Networks (Europe)
  - ICANN05: http://www.ibspan.waw.pl/ICANN-2005/
- ...

# Recommended Projects

1. Machine learning application to predict fault prone program source files
2. Machine learning application for bug triaging (text mining)
3. Machine learning application for Image Processing
4. Machine learning application for voice recognition and signal processing
5. Machine learning application in Robotics
6. Machine Learning Application for Autopilot system
7. Machine learning application in Networking (Network Intrusion Detection System)

# Tools

- MATLAb
- WEKA (Java APIs)
- SPSS or R (statistical tools)

# Course Layout

1. Introduction to Machine Learning
2. Overview of Machine Learning Algorithms
3. Data Analysis Methods
   - Scatter Plotting/Correlation Analysis
   - Principal Component Analysis
4. Supervised Machine Learning
   - Statistical Regression Methods
   - Artificial Neural Network
   - Decision Tree
   - Support Vector Machine
5. Unsupervised machine Learning
   - Clustering (k-means clustering, mixture models, hierarchical clustering)
   - Self-Organizing Map
   - Expectation Maximization Algorithm
6. Bayes Theorem and Bayesian Belief Network
7. Hidden Markov Model
8. Ensemble Learning Algorithms:
   - Bagging
   - Boosting
9. Pattern Mining
   - Association Rules
   - Apriori Algorithms
10. Information Search and Retrieval Methods
    - Vector Space Model
    - Latent Semantic Indexing
11. Application of Machine Leaning
    - Robotic/Image Processing/Fault Prediction

# Lecture-01

**Q. What is Intelligence?**

**Q. What is Artificial Intelligence?**

**Q. What is Learning?**

**Q. What is Machine Learning?**

# Q. What is Intelligence?

**A.** Intelligence is the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines.

# Q. What is Artificial Intelligence?

**A.** It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence.

# What is Learning

To gain knowledge or understanding of, or skill in by study, instruction or experience"

- Learning a set of new facts
- Learning HOW to do something
- Improving ability of something already learned

Simon's ddefinition:

``Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the same task or tasks drawn from the same population more effectively the next time"

-- Machine Learning I, 1993, Chapter 2.

# Why "Learn" ?

- **Learning is used when:**
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition)
  - Solution changes in time (routing on a computer network)
  - Solution needs to be adapted to particular cases (user biometrics)

# Why We Need CI/AI/ML ?

The objective is to build an intelligent system/agent which can think like human and should be adoptive in nature.

To build an intelligent system we build models. These models are based on AI/CI/ML techniques.

Models are trained to solve the problems by using some experience/data

# Types of Problems

1. Decision Making Problem
2. Optimization Problems

To provide the solution, we have to train the model by using a data set

Training/testing datasets may have two or more attributes/parameters. These parameters may be classified as dependent or independent parameters.

# Machine Learning: A Definition

**Definition:** A computer program is said to *learn* from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.

# Machine Learning (ML):

- ML is programming computers to optimize a performance criterion using example data or past experience.
- Learn models from example data
  - A massive amount of data
  - A moderate amount of data
  - A small amount of data
- After learning, apply the learned models for inference.
  - Accuracy
  - Time/Space Complexity

# Examples of Successful Applications of Machine Learning

- Learning to recognize spoken words (Lee, 1989; Waibel, 1989).
- Learning to drive an autonomous vehicle (Pomerleau, 1989).
- Learning to classify new astronomical structures (Fayyad et al., 1995).
- Learning to play world-class backgammon (Tesauro 1992, 1995).
- ARVIN vehicle based on ANN

# Why is Machine Learning Important?

- Some tasks cannot be defined well, except by examples (e.g., recognizing people).
- Relationships and correlations can be hidden within large amounts of data. Machine Learning/Data Mining may be able to find these relationships.
- Human designers often produce machines that do not work as well as desired in the environments in which they are used.

# Areas of Influence for Machine Learning

- **Statistics:** How best to use samples drawn from unknown probability distributions to help decide from which distribution some new sample is drawn?
- **Brain Models:** Non-linear elements with weighted inputs (Artificial Neural Networks) have been suggested as simple models of biological neurons.
- **Adaptive Control Theory:** How to deal with controlling a process having unknown parameters that must be estimated during operation?

# Practical Application of ML: Data Mining

- Retail: Market basket analysis, Customer relationship management (CRM)
- Finance: Credit scoring, fraud detection
- Manufacturing: Optimization, troubleshooting
- Medicine: Medical diagnosis
- Telecommunications: Quality of service optimization
- Web mining: Search engines
- Text Mining
- …

# Designing a Learning System:

1. Problem Description
2. Choosing the Training Experience
3. Choosing the Target Function
4. Choosing a Representation for the Target Function
5. Choosing a Function Approximation Algorithm
6. Final Design

# Machine Learning

A computer program is said to *learn* from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.

# What is Learning

To gain knowledge or understanding of, or skill in by study, instruction or experience"

- Learning a set of new facts
- Learning HOW to do something
- Improving ability of something already learned

Simon's ddefinition:

``Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the same task or tasks drawn from the same population more effectively the next time"

-- Machine Learning I, 1993, Chapter 2.

# Major Paradigms of Machine Learning

- **Rote Learning** – One-to-one mapping from inputs to stored representation. "Learning by memorization." Association-based storage and retrieval.
- **Induction** – Use specific examples to reach general conclusions .
- **Clustering** – Unsupervised identification of natural groups in data

# How Machine Learning Methods Operate

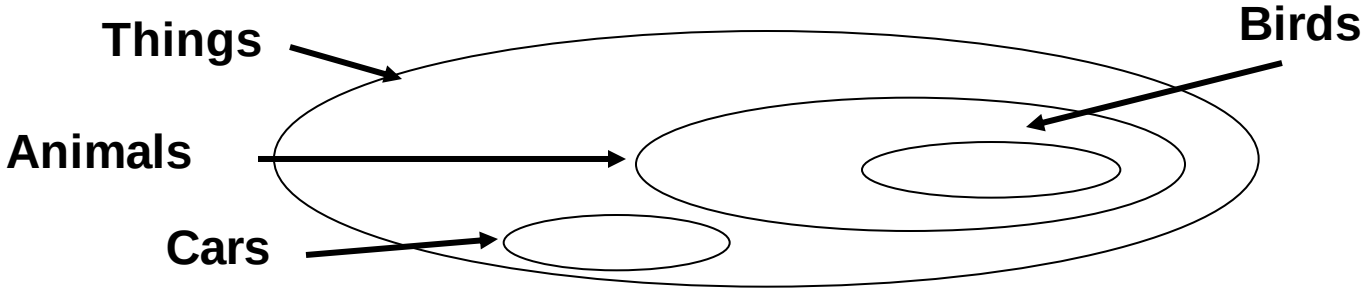
- Inputs:
  - What's a Concept? What is concept learning?
  - What is an Example/Instance?
  - What is attribute?
- Outputs (Knowledge Representation):
  - Decision Tables
  - Decision Trees
  - Classification Rules
  - Association Rules
  - Instance based Representation
  - Clusters

# Styles of Learning

- Four basically different styles of learning appear in data mining applications.
  - ***classification learning,*** the learning scheme is presented with a set of classified examples from which it is expected to learn a way of classifying unseen examples.
  - In ***association learning,*** any association among features is sought, not just ones that predict a particular *class* value.
  - In ***clustering,*** groups of examples that belong together are sought.
  - In ***numeric prediction,*** the outcome to be predicted is not a discrete class but a numeric quantity.
- Regardless of the type of learning involved, we call the thing to be learned the ***concept*** and the output produced by a learning scheme the ***concept description***.

# What is a Concept?

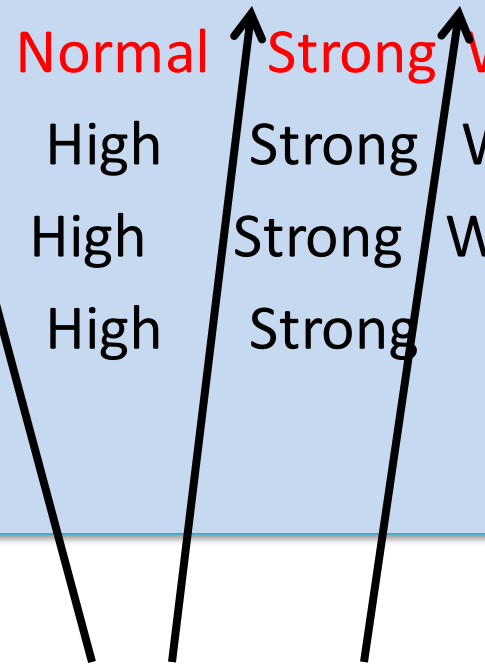
- A Concept is a subset of objects or events defined over a larger set [Example: The concept of a *bird* is the subset of all objects (i.e., the set of all *things* or all *animals*) that belong to the category of bird.]




- Alternatively, a concept is a boolean-valued function defined over this larger set [Example: a function defined over all *animals* whose value is *true* for birds and *false* for every other animal].

# What is an Example/Instance ?

| *Day* | *Sky* | *AirTemp* | *Humidity* | *Wind* | *Water* | *Forecast* | *WaterSport* |
|---|---|---|---|---|---|---|---|
| 1 | Sunny | Warm | Normal | Strong | Warm | Same | Yes |
| 2 | Sunny | Warm | High | Strong | Warm | Same | Yes |
| 3 | Rainy | Cold | High | Strong | Warm | Change | No |
| 4 | Sunny | Warm | High | Strong | Cool | Change | Yes |



Attributes

# What is Concept-Learning?

Given a set of examples labeled as members or non-members of a concept, *concept-learning* consists of automatically inferring the general definition of this concept.

In other words, *concept-learning* consists of approximating a boolean-valued function from training examples of its input and output.

# Example of a Concept Learning task

- **Concept:** Good Days for Water Sports (values: Yes, No)
- **Attributes/Features:**
  - Sky (values: Sunny, Cloudy, Rainy)
  - AirTemp (values: Warm, Cold)
  - Humidity (values: Normal, High)
  - Wind (values: Strong, Weak)
  - Water (Warm, Cool)
  - Forecast (values: Same, Change)
- **Example of a Training Point:**

<Sunny, Warm, High, Strong, Warm, Same, Yes>

**class**

# Example of a Concept Learning task

**Database:**

| *Day* | *Sky* | *AirTemp* | *Humidity* | *Wind* | *Water* | *Forecast* | *WaterSport* |
|---|---|---|---|---|---|---|---|
| 1 | Sunny | Warm | Normal | Strong | Warm | Same | Yes |
| 2 | Sunny | Warm | High | Strong | Warm | Same | Yes |
| 3 | Rainy | Cold | High | Strong | Warm | Change | No |
| 4 | Sunny | Warm | High | Strong | Cool | Change | Yes |

**class**

## Chosen Hypothesis Representation:

**Conjunction of constraints on each attribute** where:

- "?" means "any value is acceptable"
- "0" means "no value is acceptable"

**Example of a hypothesis: <?,Cold,High,?,?,?>**
(If the air temperature is cold and the humidity high then it is a good day for water sports)

# Example of a Concept Learning task

- **Goal:** To infer the "best" concept-description from the set of all possible hypotheses ("best" means "which best generalizes to all (known or unknown) elements of the instance space".)
- **Most General Hypothesis:** Everyday is a good day for water sports <?,?,?,?,?,?>
- **Most Specific Hypothesis:** No day is a good day for water sports <0,0,0,0,0,0>