# CIS4930/5930: Machine Learning

## Introduction to ML

Alan Kuhnle

Florida State University

Slides adapted from Mehryar Mohri

# This Lecture

- *Basic definitions and concepts*x
- Introduction to the problem of learning
- Probability tools

# Machine Learning

- Definition: computational methods using experience to improve performance
- Experience: data-drive task, thus statistics, probability, and optimization
- Computer science: learning algorithms, analysis of complexity, theoretical guarantees
- Example: use document word counts to predict its topic

# Examples of Learning Tasks

- Text: document classification, spam detection
- Speech: recognition, synthesis, verification
- Image: annotation, face recognition, OCR, handwriting recognition
- Games (e.g. chess, go)
- Unassisted control of vehicles
- Medical diagnosis, fraud detection, network intrusion

# Some Broad ML Tasks

- Classification: assign a category to each item
- Regression: predict a real value for each item
- Ranking
- Clustering
- Dimensionality reduction

# General Objectives of ML

- Theoretical questions
  - what can be learned, under what assumptions?
  - are there learning guarantees?
  - analysis of learning algorithms

# General Objectives of ML

- Theoretical questions
  - what can be learned, under what assumptions?
  - are there learning guarantees?
  - analysis of learning algorithms
- Algorithms
  - more efficient and more accurate algorithms
  - handle large-scale problems
  - deal with avariety of different learning scenarios

# This Course

- Theoretical foundations
    - learning guarantees
    - analysis of algorithms

# This Course

- Theoretical foundations
  - learning guarantees
  - analysis of algorithms
- Algorithms
  - present major, mathematically well-studied algorithms
  - discussion of extensions

# This Course

- Theoretical foundations
  - learning guarantees
  - analysis of algorithms
- Algorithms
  - present major, mathematically well-studied algorithms
  - discussion of extensions
- Applications
  - illustration of their use

# Topics

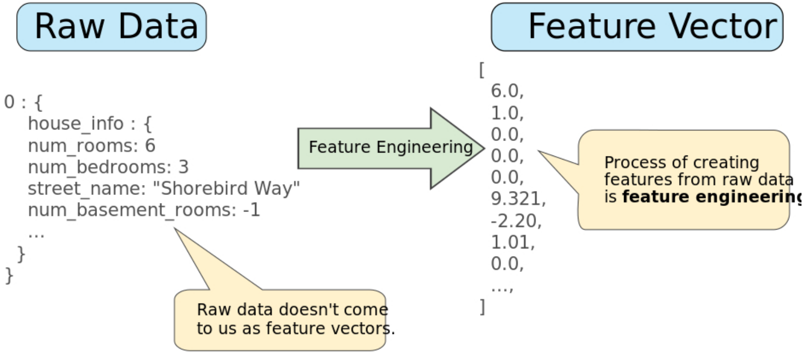
- PAC learning framework
- Rademacher Complexity & VC Dimension
- Model Selection
- Support vector machines
- Kernel methods
- Online learning
- Regression
- Dimensionality reduction
- Reinforcement learning
- Deep Feedforward Networks
- Optimization for Training Deep Models

## Definitons and Terminology

- Example: item, instance of the data used. Often drawn from underlying (unknown) probability distribution
- Features: attributes associated to an example, which may be used for learning. Often represented as a vector

# Definitons and Terminology

- Example: item, instance of the data used. Often drawn from underlying (unknown) probability distribution
- Features: attributes associated to an example, which may be used for learning. Often represented as a vector



# Definitions and Terminology

- Labels: May be categorical (classification) or real values (regression) associated to an item. Labels are what we are trying to infer
- Data: Set of examples drawn from underlying distribution
  - training data (typically labeled)
  - test data (labeled, but labels are not seen)
  - validation data (labeled, may be used for tuning parameters)

# General Learning Scenarios

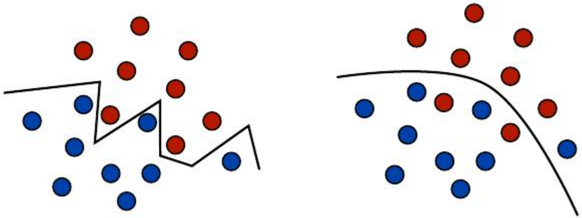
- Settings: *batch* vs. *online*
- Queries: *active* vs. *passive*

# Standard Batch Scenarios

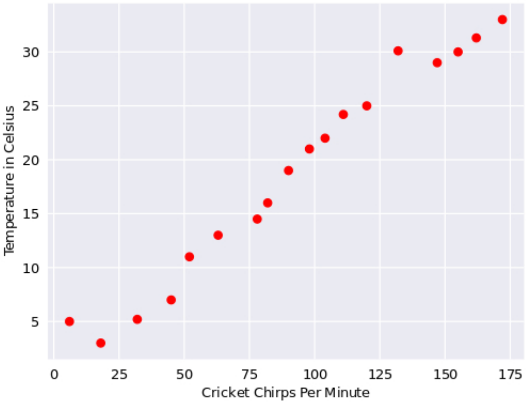
- Unsupervised learning
- Supervised learning
- Semi-supervised learning

# Example – SPAM Detection

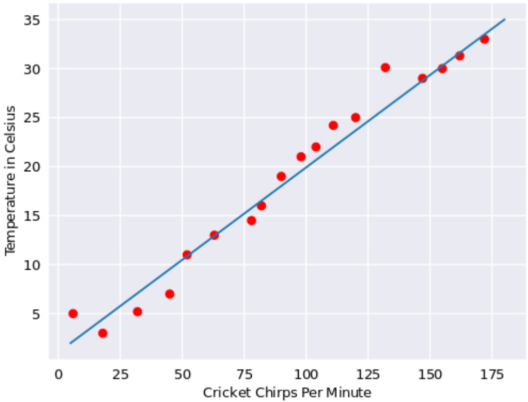
- Problem: classify each e-mail message as SPAM or non-SPAM
- Potential data: large collection of SPAM and non-SPAM messages



# Example – Linear regression



# Example – Linear regression




$$y = mx + b$$

# Learning Stages