E9 205 Machine Learning for Signal Processing

Introduction to Machine Learning of Sensory Signals

06-08-2018

Instructor - Sriram Ganapathy (sriramg@iisc.ac.in.in)

Teaching Assistant - Akshara Soman (tentative) .

Class Location - EE B303

Timings - MW 330-500pm. *Fridays (Tentative) 3-4pm.*

Overview

- What are the typical real-world signals
- What is learning
- Why should we attempt learning of such signals
- Roadmap of the course

Established 1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Real World Signals

- Signal in general is a function f : X —> V
- Real World Signals
  - which we see everyday everywhere
  - Text, Speech, Image, Videos…
  - DNA sequence, financial data, weather parameters, neural spike train…
  - Belonging to/generated by certain category of events.

Real World Signals

- Types of signals- Continuous and Discrete
- Observations from real world signals
    - Information may not be uniform.
    - Cannot be modeled deterministically.
    - Affected by noise, sensing equipments.
    - Missing or hidden variables.

Established 1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Real World Signals - Examples

- Text data
  - Discrete sequence of items

In the last 29 years, sir has never ever said 'well played' to me because he thought I would get complacent and I would stop working hard.

Items - [In] [the] **[last]** [29] **[years]** ……

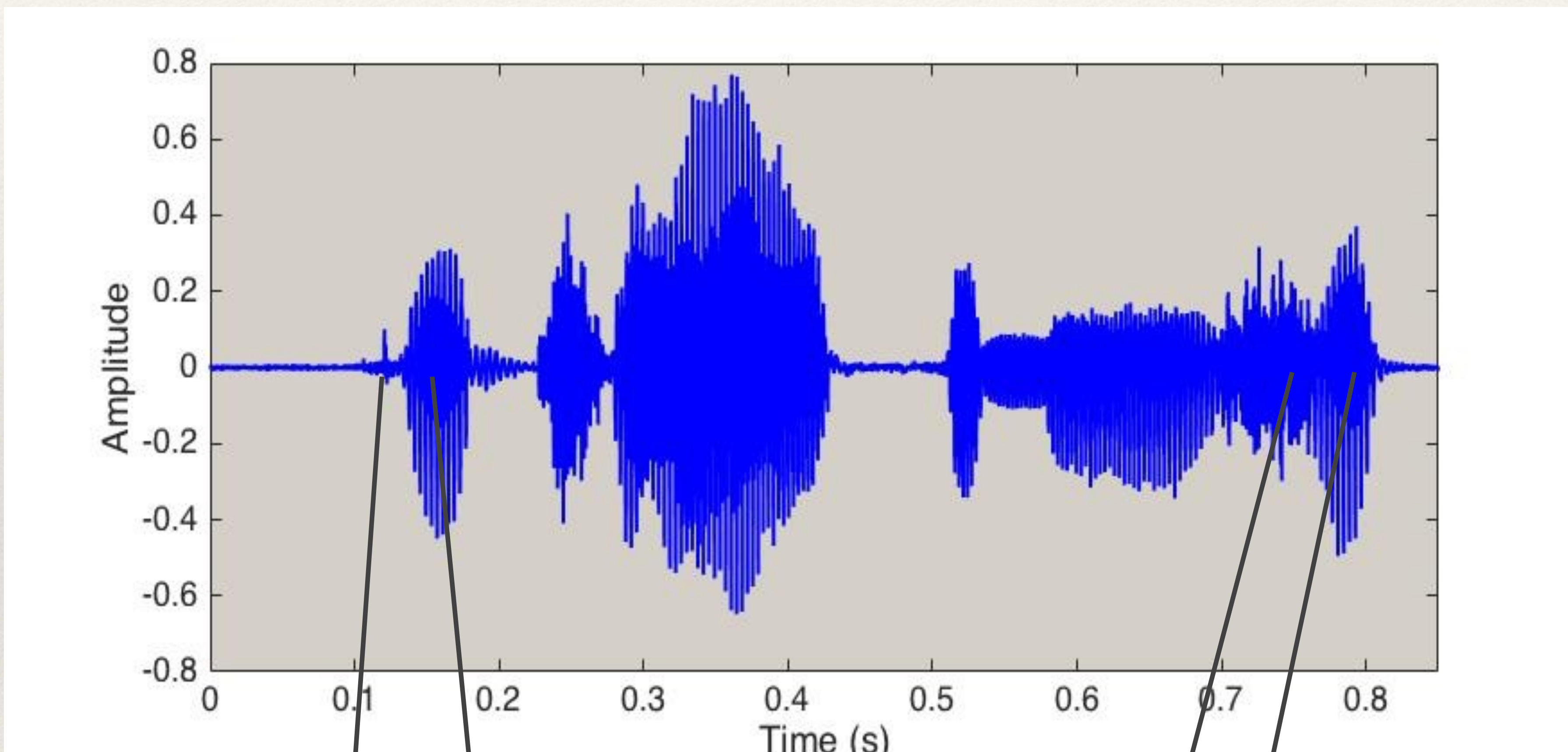
  - Some items carry more **importance** than others.

Established 1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Real World Signals - Examples

- Speech data



/dh//ah/ /jh//ae//p//ah//n//iy//z/

- Phonetic units - underlying hidden variables.

Real World Signals - Examples

- Images

- Measurement artifacts - noise.

Patterns in Real World Signals

- Patterns in real world signals
  - Caused by various generation processes in the real-world signals.
  - Hidden from the observation.
  - Value patterns and geometric patterns.
  - May be hierarchical in nature.
  - Manifested as pure patterns or transformed/distorted versions.

Established 1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

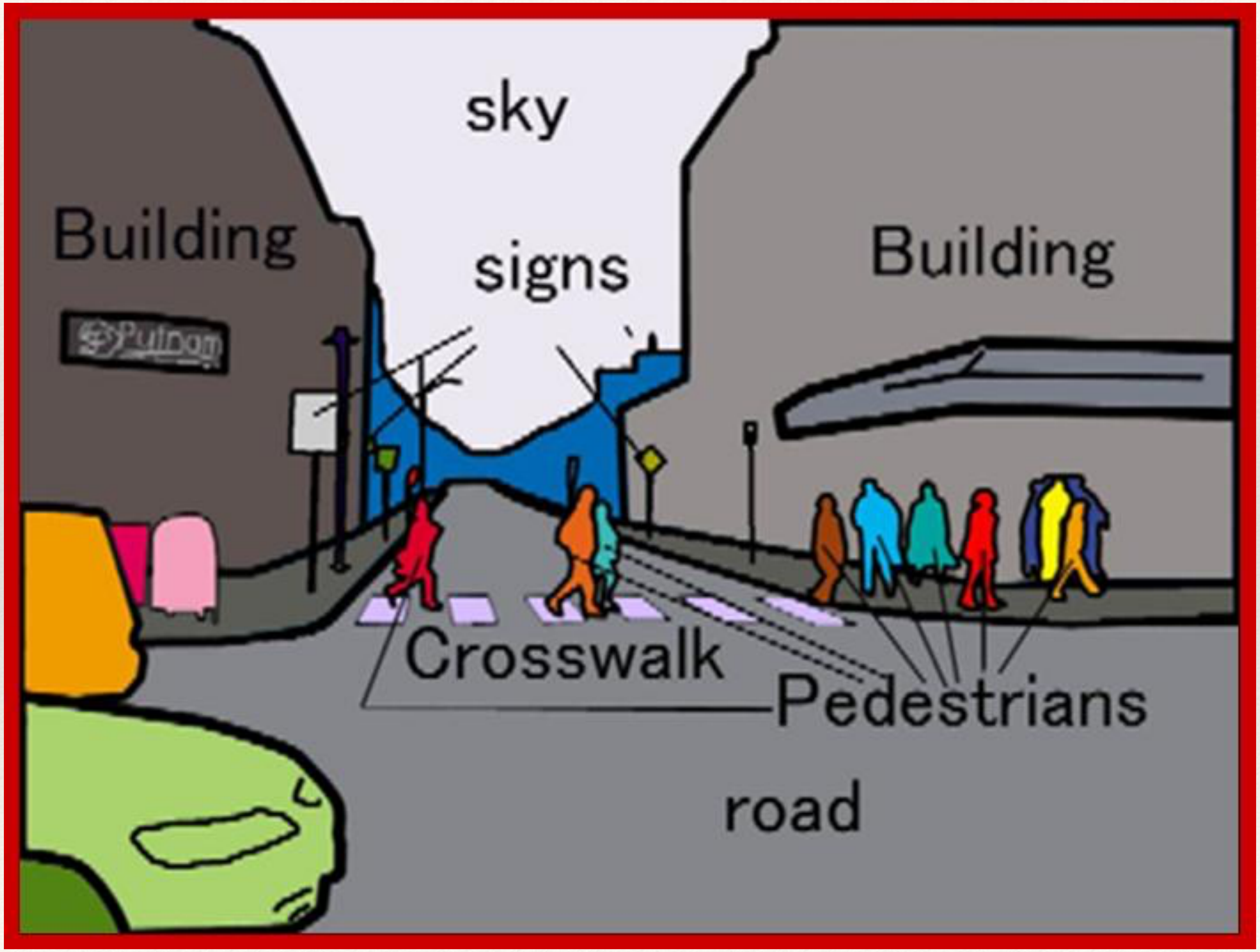
What is Learning

- Learning
  - Process of describing or uncovering the pattern.
  - Understanding the physical process of generation.
  - Generalization for prediction, classification, decision making.
  - Using the data to learn the underlying pattern.
- Humans are fundamentally trained to learn and recognize patterns.

Established 1911

What is Learning

Object Recognition



What is Learning

Facial Identification

Topic Summarization

The Karnataka government is planning to start an aviation school to help students from lower economic and rural backgrounds become pilots.

Machine Learning

- Machine Learning
  - Automatic discovery of patterns.
  - Motivated by human capabilities to process real world signals.
  - Mimicking/Extending/Replacing human functions.
  - Branch of artificial intelligence.
  - Classification and Regression.

Established 1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Machine Learning - Examples

Domain Identification - Blog v/s Chat ?

"I tried these Butterscotch Muffins today and they turned out so good. I had half the pack of butterscotch chips that I bought long back so wanted to use it up."

"Hey, it's Geoff from yesterday. How's it going? Hi there. Don't wanna bother you long, but you saw this video?"

Machine Learning - Examples

Did a Human or Machine write this ?

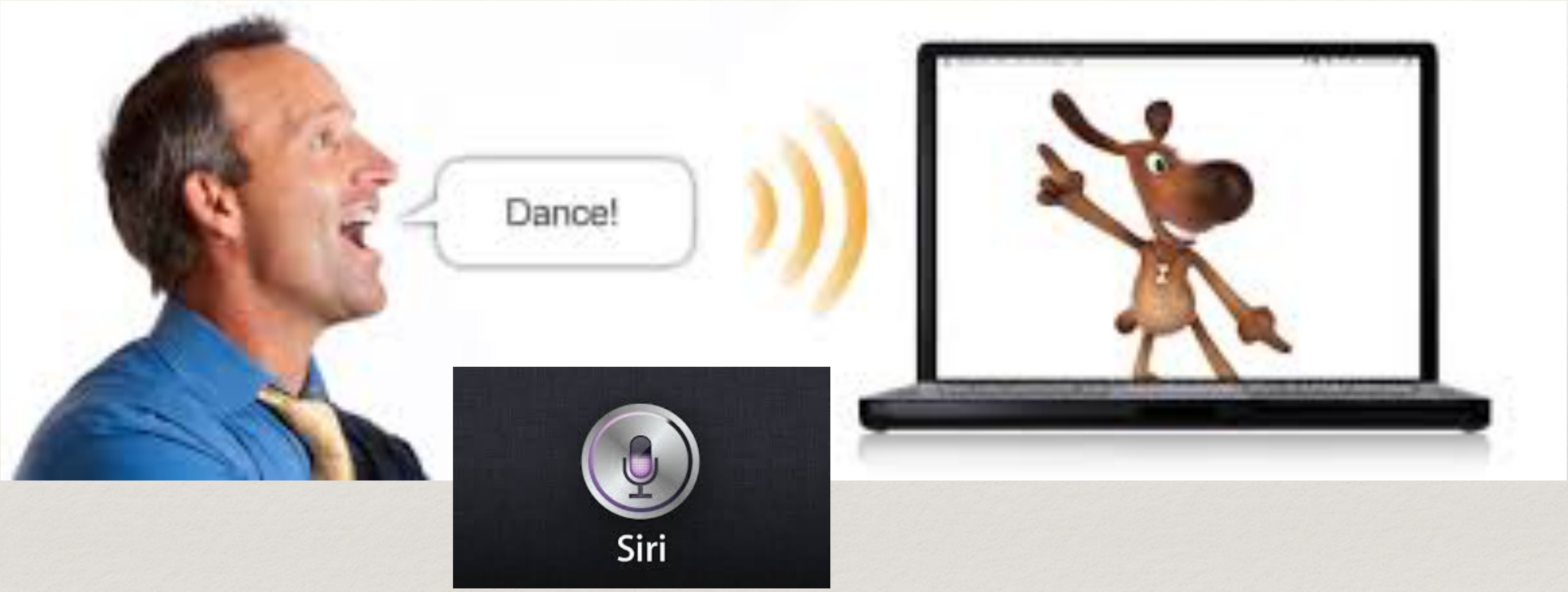
"A shallow magnitude 4.7 earthquake was reported Monday morning five miles from Westwood, California, according to the U.S. Geological Survey. The temblor occurred at 6:25 AM, Pacific time at a depth of 5.0 miles."

"Kitty couldn't fall asleep for a long time. Her nerves were strained as two tight strings, and even a glass of hot wine, that Vronsky made her drink, did not help her. Lying in bed she kept going over and over that monstrous scene at the meadow."

http://www.nytimes.com/interactive/2015/03/08/opinion/sunday/algorithm-human-quiz.html

Machine Learning - Examples

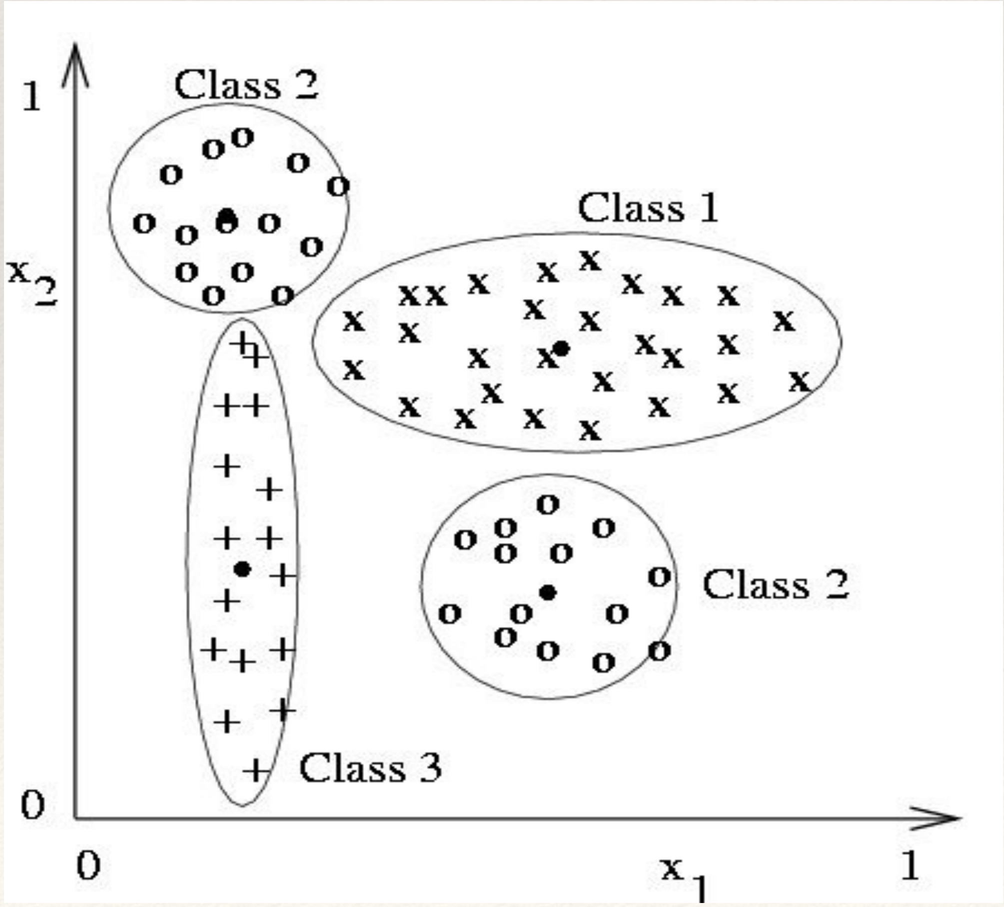
Speech Recognition



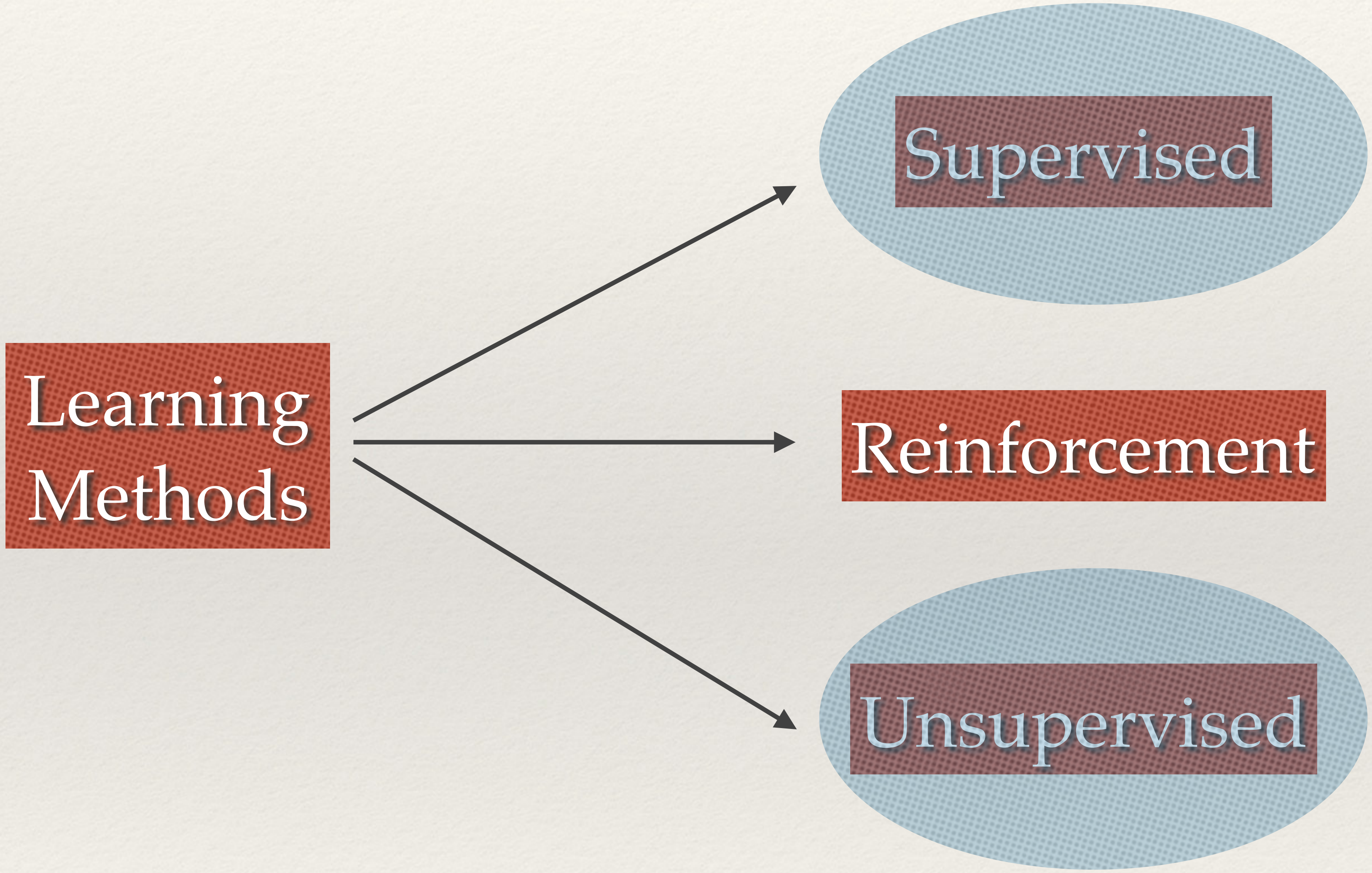
Sound Synthesis http://news.mit.edu/2016/artificial-intelligence-produces-realistic-sounds-0613

Machine Learning

- Traditional approaches to Machine Learning
  - Rule and heuristic based methodologies
  - Using small amounts of data.
- Recently, most problems are addressed as statistical pattern recognition problem with big data.



Types of Learning



Camstra, Vinciarelli, "Machine Learning for Audio, Image and Video Analysis" 2007.

Unsupervised Learning

- Data is presented without associated output targets
  - Extracting structure from the data.
  - Examples like clustering and segmentation.
  - Concise description of the data - dimensionality reduction methods.

Established 1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Reinforcement Learning

- Dynamic environment resulting in triplets - state/action/reward.
  - No optimal action for a given state
  - Algorithm has to learn actions in a way such the expected reward is maximized over time.
  - May also involve minimizing punishment.
  - Reward/punishment could be delayed - learning based on past actions.

Sutton, Barto, "Reinforcement Learning: An Introduction." MIT Press, 1998.

Established 1911

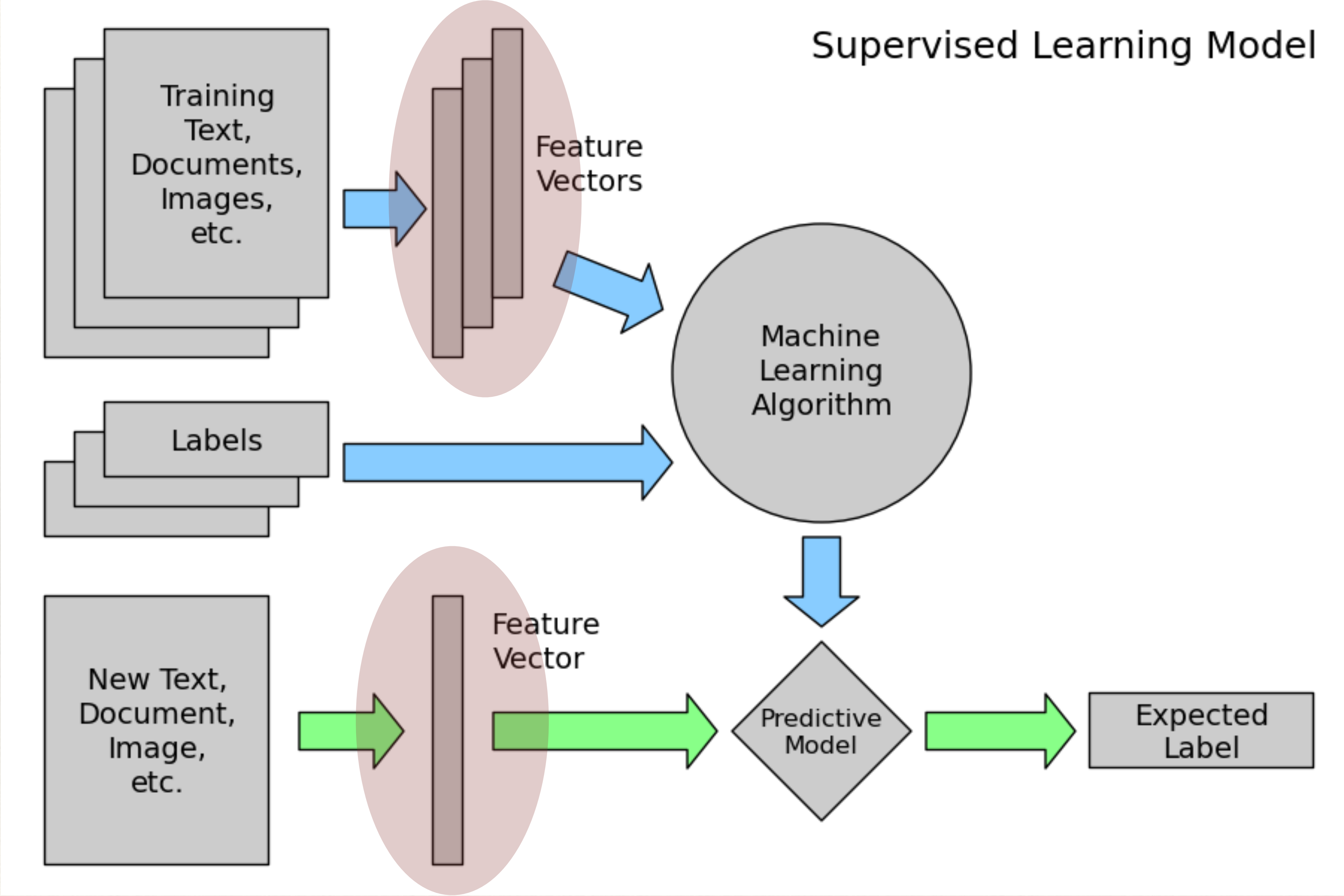
INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Supervised Learning

- Training data is provided with along with target values (ground truth).
    - Goal - to learn the mapping function from data to targets.
    - Use the mapping function to predict unseen/test data samples.
- Two types based on the structure of the labels.
    - Classification - discrete number of classes or categories.
    - Regression - continuous output variables.

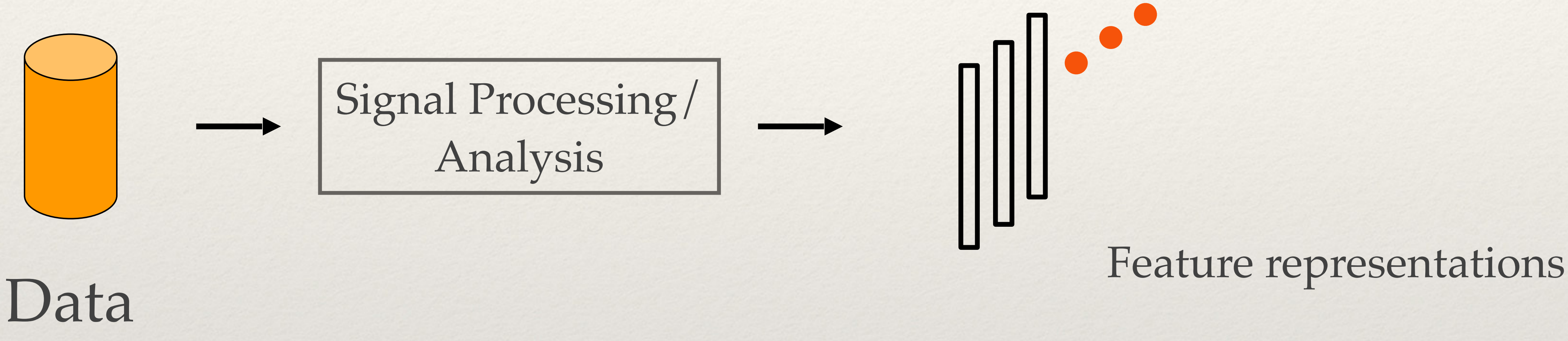
Established
1911

Supervised Learning



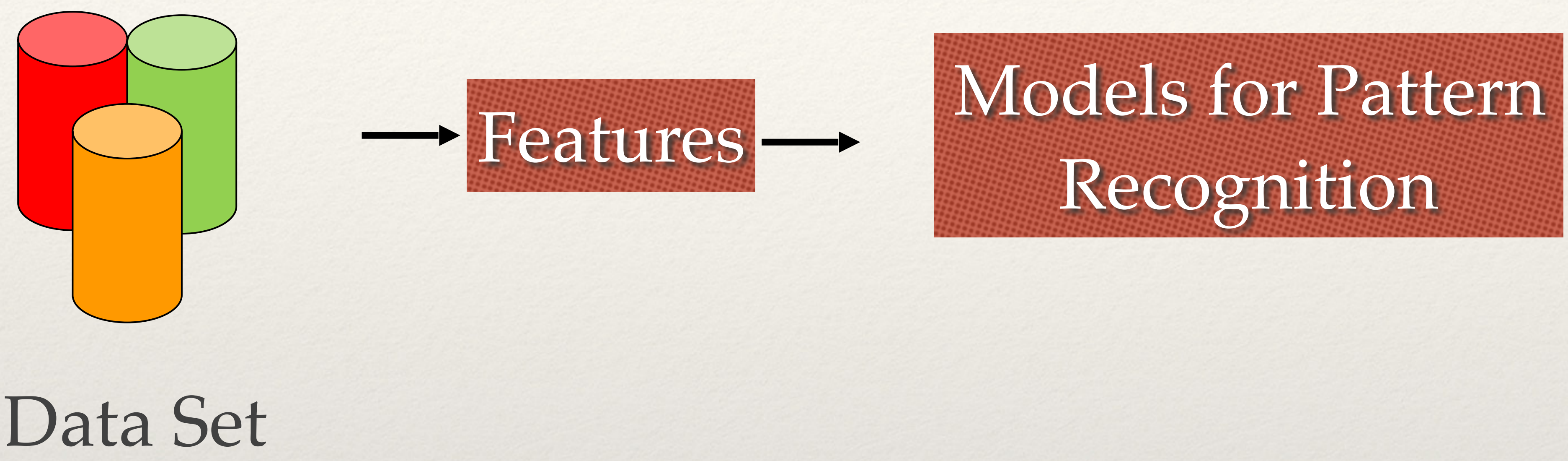
http://www.astroml.org/sklearn_tutorial/auto_examples/plot_ML_flow_chart.html

Course Roadmap



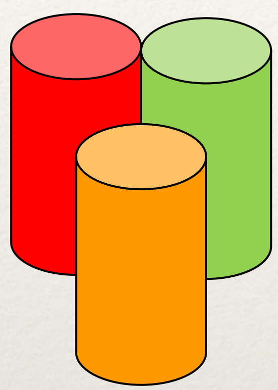
- Feature Extraction from Text, Speech, Image/Video signals (first 3 lectures).

Course Roadmap



- Between features and pattern recognition
  - Feature selection, dimensionality reduction.
  - Representation learning.

Course Roadmap



Data Set → Features → **Models for Pattern Recognition**

- Modeling the generation of data
  - Gaussian, Mixture Gaussian, Hidden Markov Models etc.
- Modeling the separation of data
  - Support Vector Machines, Deep Neural Networks etc.

Established 1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Course Structure (Rough Schedule)

- Signal analysis and processing (1st week)
  - Audio/Speech - spectrograms
  - Text - TF/IDF, Image feature extraction
- Basics of Pattern Recognition (2nd week).
  - Dimensionality reduction, factorization and feature selection.
- Generative modeling (next 3 weeks)
  - Gaussian and mixture Gaussian modeling, hidden Markov modeling.
- Discriminative modeling - Support vector machines (next 2 weeks)
- Deep Learning (next 6 weeks)
- Unsupervised learning from Deep Models (last 2 weeks)

Housekeeping

Requisite

- Must
  - Probability/Random process/Stochastic Models
  - Linear Algebra/Matrix Analysis
- Preferred
  - Intro to Signal Processing
- Preferred
  - Coding in Python

Grading

- Assignments - Theory + Implementation (20%)
- Mid-terms (20%)
- Project (25%)
- Finals (35%)

Established
1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Housekeeping

Project and Coding Assignments

- Coding and submissions
  - Preferred Language - Python.
- In class demos and example recipes in python.
- Final Project - GPU platform can be setup

Resources

- Textbooks -
  - PRML (Bishop), NN (Bishop).
  - Deep Learning (Goodfellow)
- Online resources (papers and other textbooks listed in webpage).

Course Webpage

www.leap.ee.iisc.ac.in/sriram/teaching/MLSP_18

Established 1911

Dates of Various Rituals

- 5 Assignments spread over 3 months (roughly one assignment every two weeks).
- September 1st week - project topic announcements.
- September 3rd week - 1st Midterm
- September 4th week - project topic and team finalization. [1 and 2 person teams].
- October 1st week - Project Proposal
- October 3rd week - 2nd MidTerm
- November 3rd week - Project MidTerm Presentations.
- December 1st week - Final Exams
- December 2nd week - Project Final Presentations.

Established 1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Content Delivery

Theory and Mathematical Foundation

Intuition and Analysis

Implementation and Understanding

Lecture and Beyond

- Teaching Assistant - Akshara
- Additional lecture slot on Friday (time ?)
- Industry research lectures (1-2)

Established 1911

Housekeeping

Questions/Comments ?

Established
1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान