MACHINE LEARNING FOR SIGNAL PROCESSING

6-1-2025

Sriram Ganapathy
LEAP lab, Electrical Engineering, Indian Institute of Science,
sriramg@iisc.ac.in

Viveka Salinamakki, Varada R.
LEAP lab, Electrical Engineering, Indian Institute of Science.
viveka.sg@gmail.com varadar2000@gmail.com

http://leap.ee.iisc.ac.in/sriram/teaching/MLSP25/

OVERVIEW

- What are the typical real-world data/signals that we deal with
  - ✓ How do we sense these signals
  - ✓ How do we represent them in digital form
- What is learning
  - ✓ Different approaches for learning from data
- Roadmap of the course
- Logistics of the course

REAL WORLD SIGNALS

- Signal in general is a function $f : X \longrightarrow V$
- 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 - TEXT

- Text data
  - ✓ How do we sense it
  - ✓ Represented as a sequence of UNICODE symbols
    - [Unicode] - an international encoding standard for use with different languages and scripts, by which each letter, digit, or symbol is assigned a unique numeric value that applies across different platforms and programs.
    - **UTF-8** — a standard that is most commonly used for encoding text
      - ➡ First 128 symbols are ASCII characters
      - ➡ Can encode more than 1M characters with variable width.

REAL WORLD SIGNALS - TEXT

- Text data
  - ✓ Discrete sequence of items

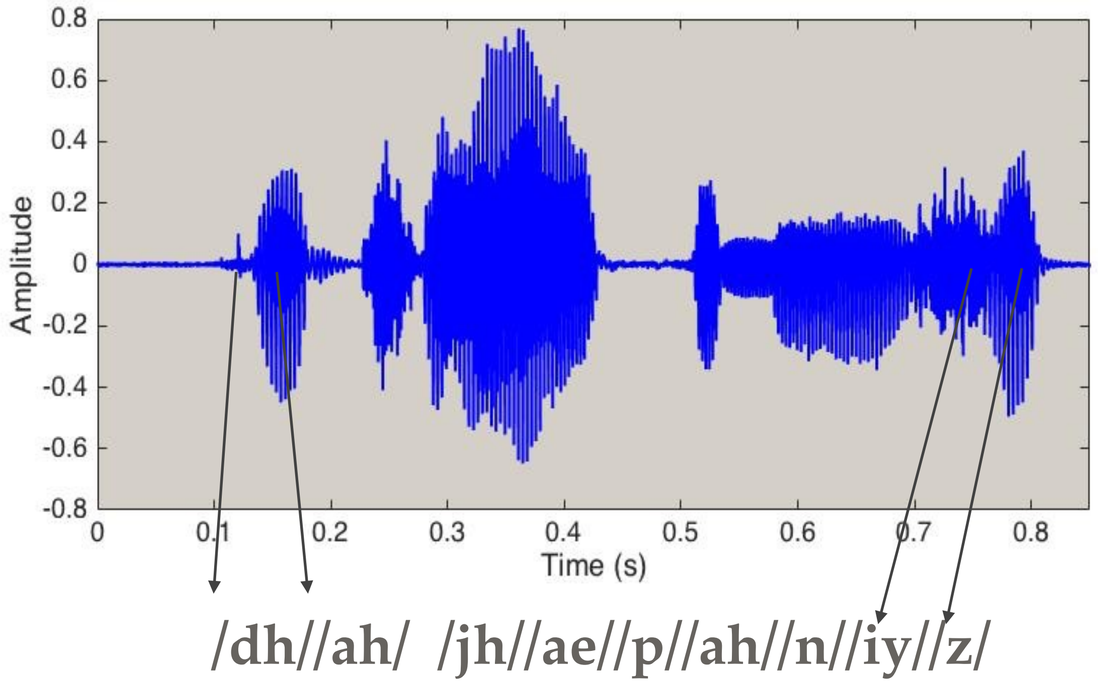
> In the last 29 years, sir has never ever said 'well played' to me because he thought I would get complacent and I would stop working hard.

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

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

REAL WORLD SIGNALS - SPEECH

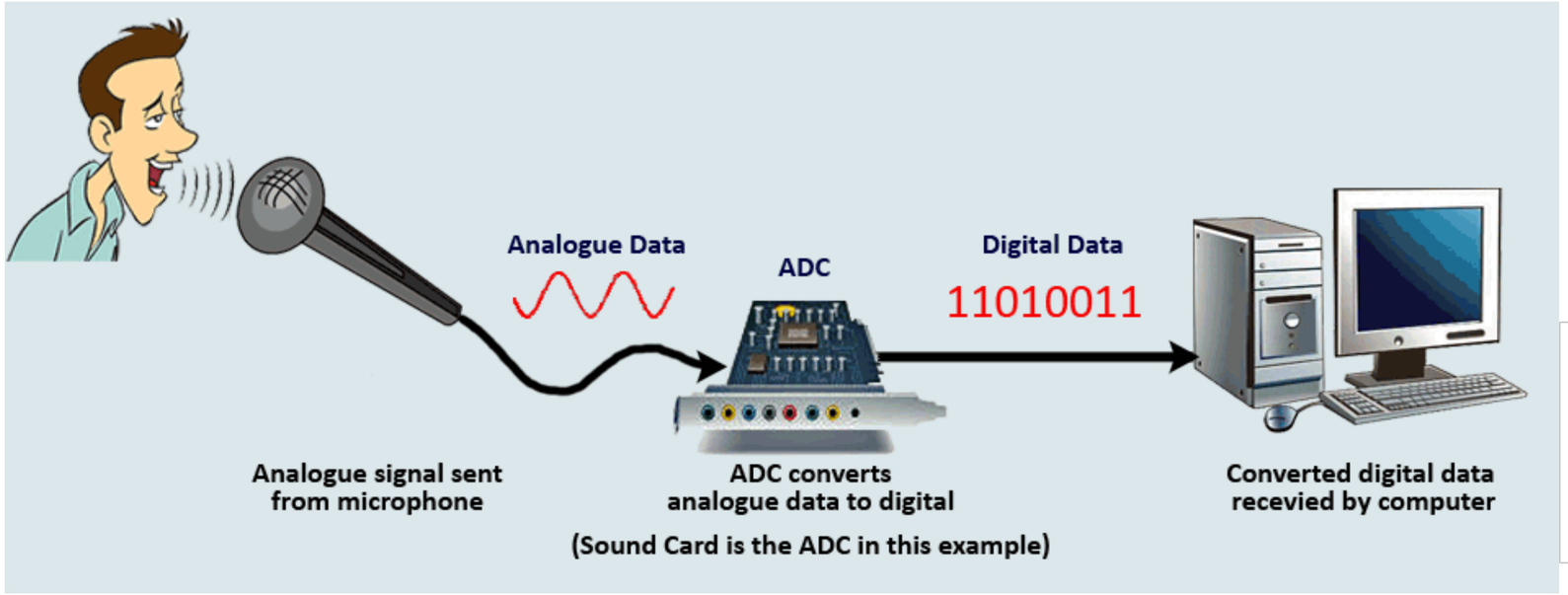
- Speech data



- ✓ Phonetic units - underlying hidden variables.

REAL WORLD SIGNALS - SPEECH

- Speech data



- ✓ Data represented as a scalar series of 16-bit integers with a specific sampling rate (typically 16kHz)

Pic - https://nandasiddhardha.medium.com/embracing-the-magic-of-data-conversion-from-analog-to-digital-88cf20bcd031

https://ccci.am/language-services/data-collection-services/audio-and-speech-data-collection/

REAL WORLD SIGNALS - IMAGE

- Images

- Basic unit is a pixel

REAL WORLD SIGNALS - VIDEO

- Image data sensing
  - ✓ Camera capture
  - ✓ Represented as 2-D grid of brightness values (8 bit/ 16 bit)
  - ✓ Can have color channels
    - Example - Data structure of 3x256x256
- Video data
  - ✓ Series of images sampled at a frame rate (e.g. 30Hz)
  - ✓ Can have an additional audio channel (sampled at higher rate).

PATTERNS IN REAL WORLD SIGNALS

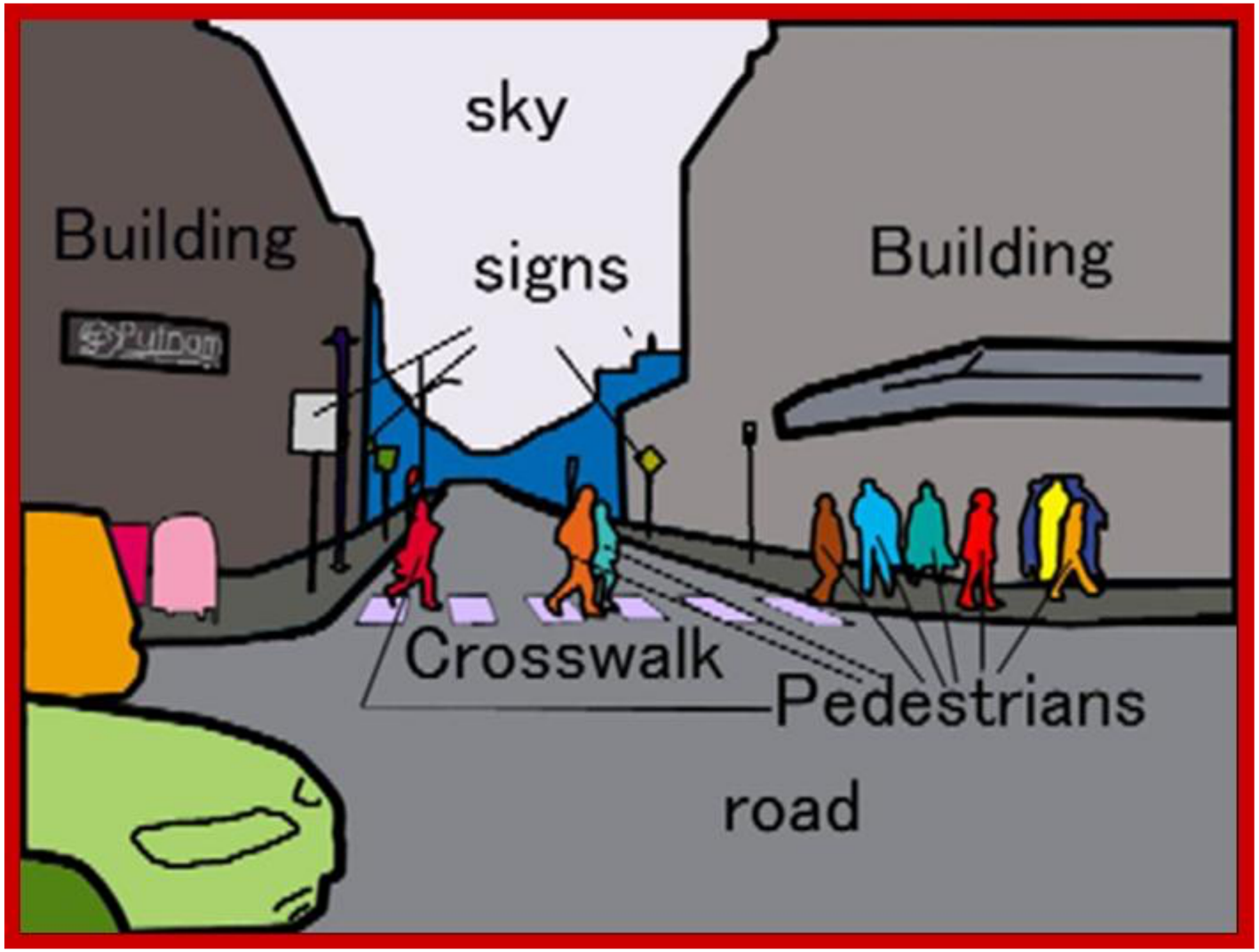
- Patterns in real world signals
  - ✓ Caused by various generation processes in the real-world signals.
  - ✓ Hidden from the observation.
  - ✓ Value patterns and geometric patterns.
  - ✓ May be hierarchical in nature.
  - ✓ Manifested as pure patterns or transformed/distorted versions.

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.
  - ✓ In some cases, learning is motivated by human abilities

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, Regression and Generation.

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 ?

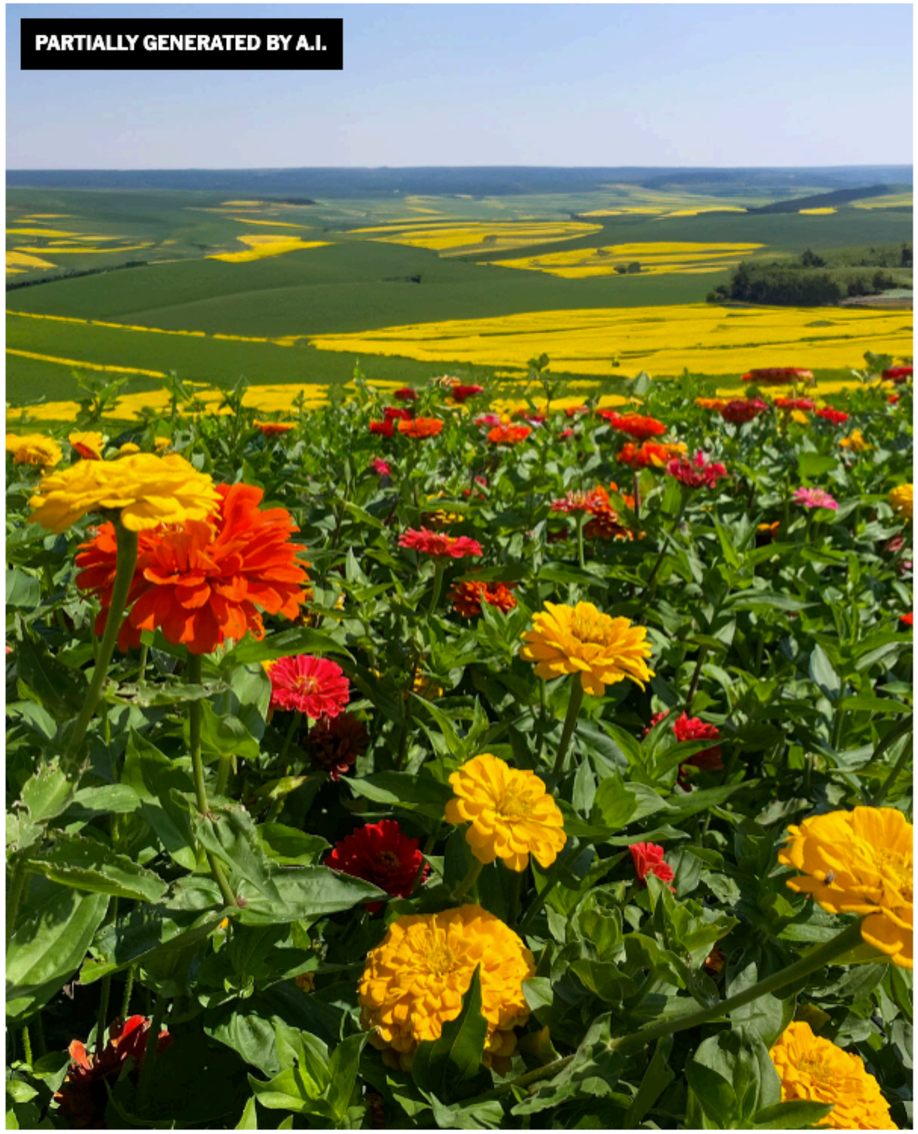
"A shallow magnitude 4.7 earthquake was reported Monday morning five miles from Westwood, California, according to the U.S. Geological Survey. The temblor occurred at 6:25 AM, Pacific time at a depth of 5.0 miles."

"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."

https://www.nytimes.com/2020/07/29/opinion/gpt-3-ai-automation.html

MACHINE LEARNING - EXAMPLES

https://www.nytimes.com/interactive/2024/12/27/technology/artificial-intelligence-generative-fill-photoshop-openai.html

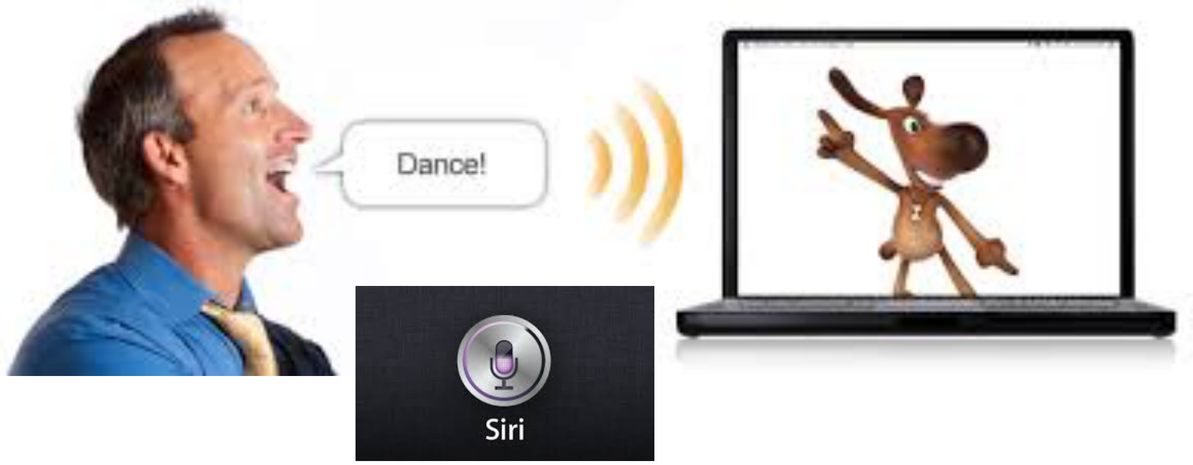


MACHINE LEARNING - EXAMPLES

https://www.nytimes.com/interactive/2024/12/27/technology/artificial-intelligence-generative-fill-photoshop-openai.html

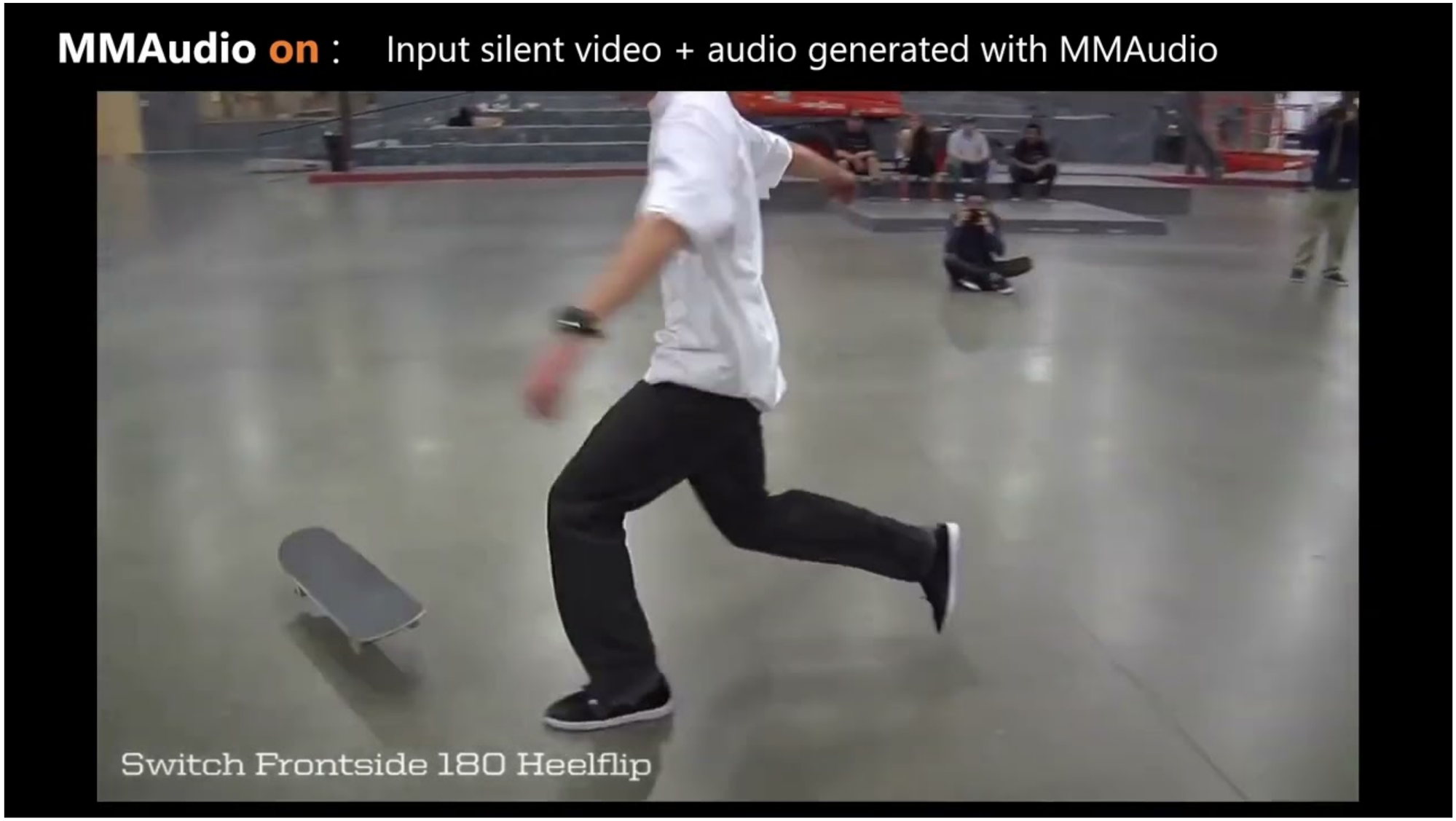
MACHINE LEARNING - EXAMPLES

Speech Recognition



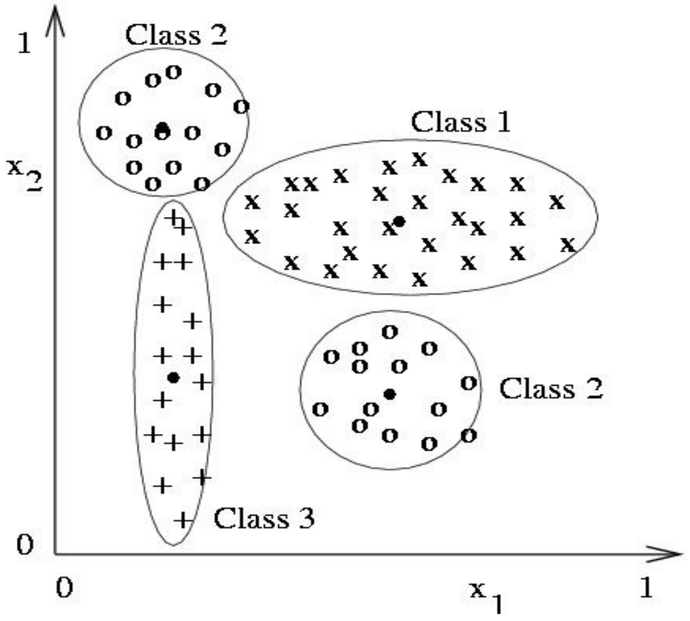
Sound Synthesis

MACHINE LEARNING VIDEO TO AUDIO GENERATION



MACHINE LEARNING

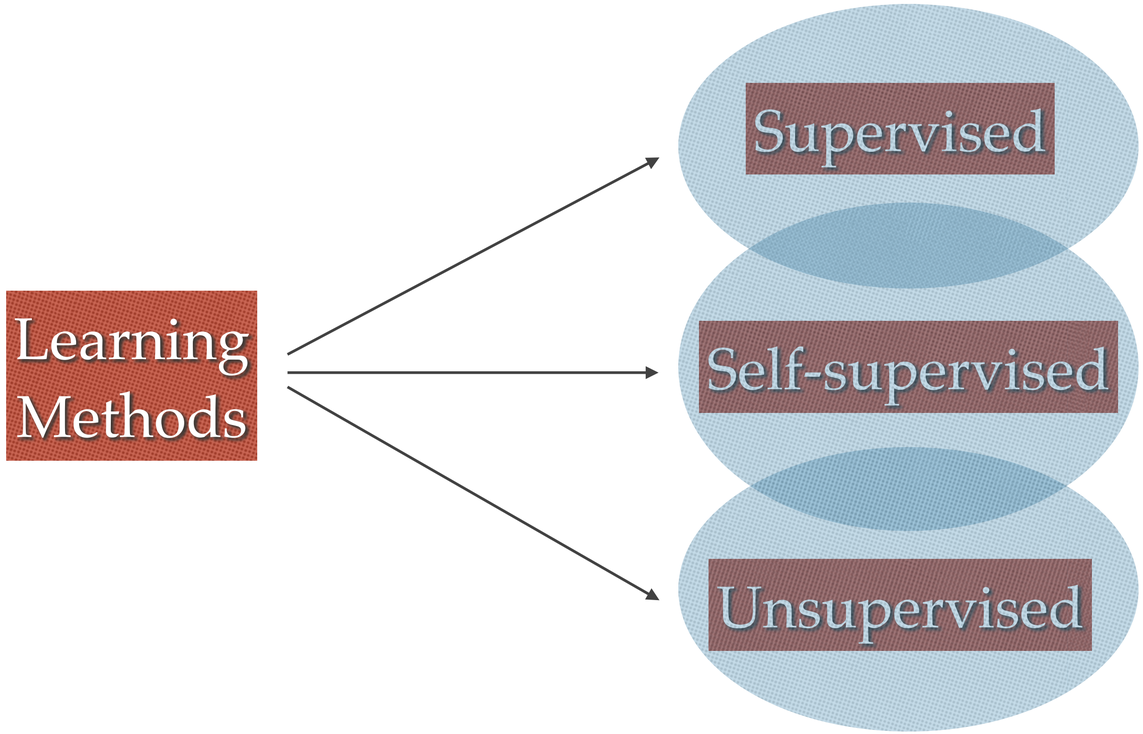
- Traditional approaches to Machine Learning
  - ✓ Rule and heuristic based methodologies
  - ✓ Using small amounts of data.
- Recently, most problems are addressed as statistical pattern recognition problem with big data.



OVERVIEW

- What are the typical real-world data/signals that we deal with
  - ✓ How do we sense these signals
  - ✓ How do we represent them in digital form
- What is learning
  - ✓ Different approaches for learning from data
- Roadmap of the course
- Logistics of the course

TYPES OF LEARNING



Reinforcement 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.

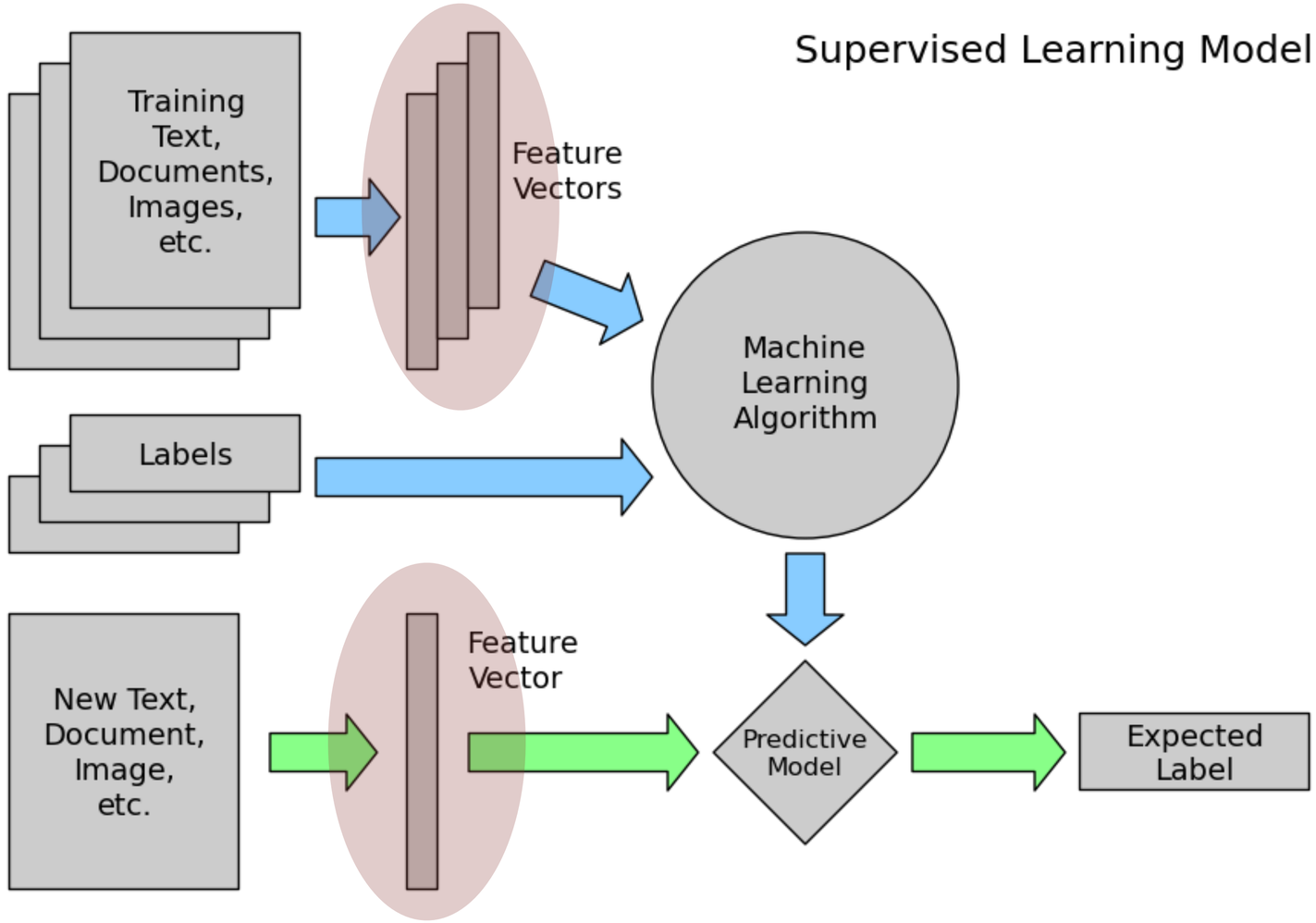
REINFORCEMENT LEARNING

- 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.

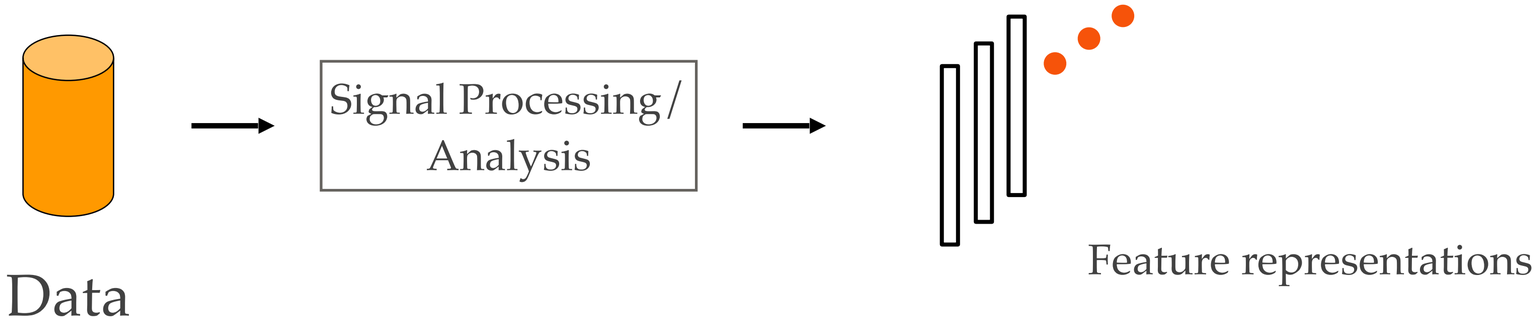
SUPERVISED LEARNING

- 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.

SUPERVISED LEARNING

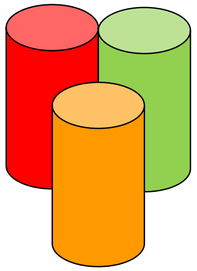


COURSE ROADMAP



- Feature Extraction from Text, Speech, Image/Video signals.

COURSE ROADMAP

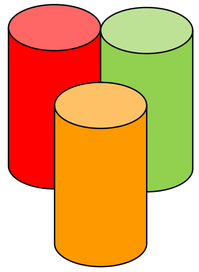


Data Set

→ Features → Models for Pattern Recognition

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

COURSE ROADMAP



Features

Models for Pattern Recognition

Data Set

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

COURSE STRUCTURE (ROUGH SCHEDULE)

- Introduction to real world data and signals - text, speech, image, video.
- Dimensionality reduction - principal components, linear discriminants.
- Decision theory for pattern recognition, ML and MAP methods, Bias-variance trade-off, model assessment, cross-validation, estimating generalization error.
- Generative modeling and density estimation - Gaussian and mixture Gaussian models, kernel density estimators, hidden Markov models. Expectation Maximization.
- Linear regression and kernel methods. Regularization methods.
- Discriminative modeling - support vector machines, decision trees and random forest classifiers, bagging and boosting.
- Neural networks: gradient descent optimization and back propagation, regularization in neural networks, dropout. normalization methods.
- Introduction to deep learning - feedforward, convolutional and recurrent networks, practical considerations in deep learning.
- Introduction to transformer models - self and cross attention, encoder and decoder architectures, autoregressive decoding.

HOUSEKEEPING

Requisite

- Must
  - ✓ Probability/Random process/Stochastic Models
  - ✓ Linear Algebra/Matrix Analysis
- Must
  - ✓ Coding in Python

Grading

- Assignments - Theory + Implementation (15%)
- Mid-terms (20%)
- Project (25%)
- Finals (40%)

HOUSEKEEPING

Project and Coding Assignments

- Coding and submissions
  - ✓ Preferred Language - Python.
- In class demos and example recipes in python.

Resources

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

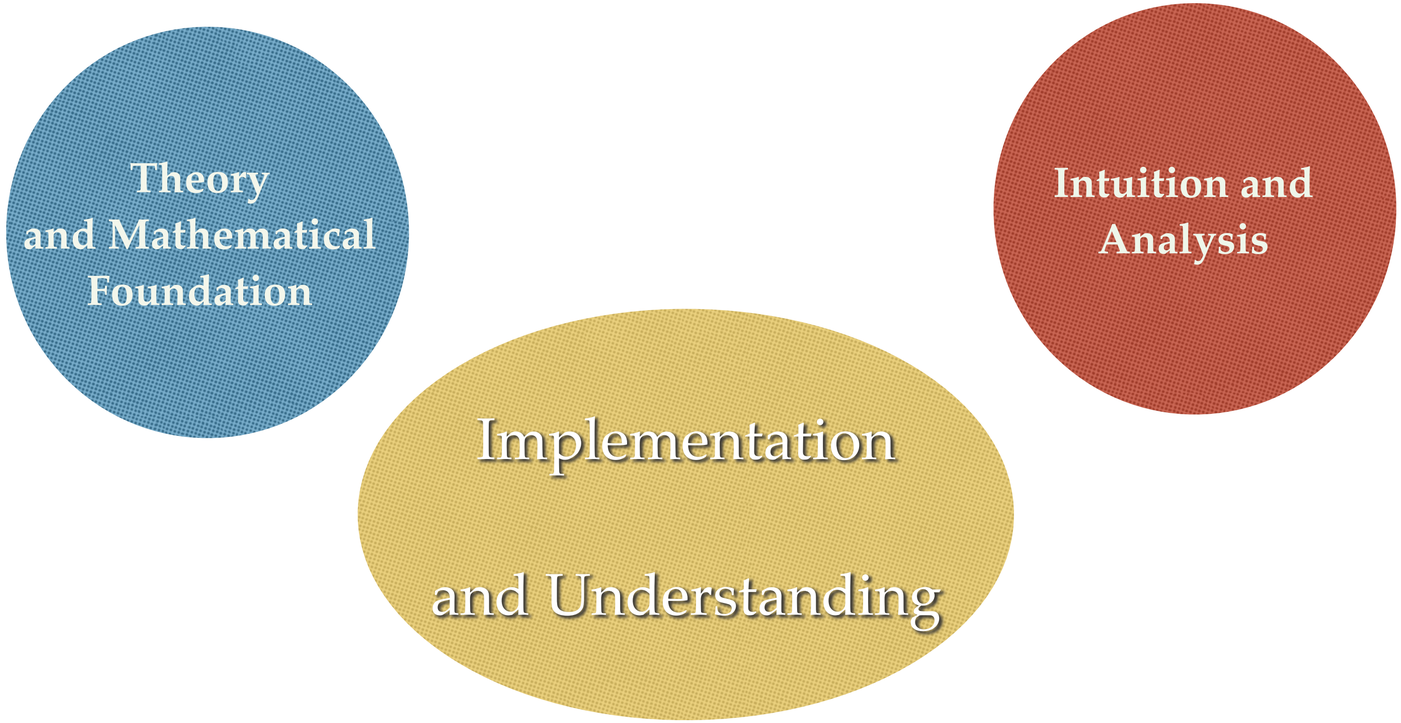
Course Webpage www.leap.ee.iisc.ac.in/sriram/teaching/MLSP25

Course Enrollment https://docs.google.com/forms/d/e/1FAIpQLSedVL4HCA5GlvGFiWHPHgDG5sR1ZR7spF4ZzDkJp9bskOJVCw/viewform

DATES OF VARIOUS RITUALS

- 3 Assignments spread over 3 months (roughly one assignment every 3 weeks).
- February second half - Midterm
- February 4th week - project topic and team finalization and proposal submission. [1 and 2 person teams].
- March 3rd week - Project MidTerm Presentations.
- April 3rd week - Final Exam
- April last week - Project Final Presentations

CONTENT DELIVERY



Lecture and Beyond

- Teaching Assistant - More will be added
- Course time frozen (M/W 11:00am - 12:30pm)
- Industry research lectures (1-2)

THANK YOU

Sriram Ganapathy and TA team
LEAP lab, C328, EE, IISc
sriramg@iisc.ac.in