Deloitte AI Academy Bootcamp – IIT Kharagpur

MARCH 3-10, 2022

About IIT Kharagpur

First

Largest

Most Diverse

Deep Expertise in AI

About the Boot Camp

Date: March 3-10 (Thursday-Thursday), 2022

Time: 9:00-13:00 and 14:00-18:00 hrs

Lectures (Interactive)

Hands On and Programming Sessions

15min Quiz at end of day

One Test

Case Study Project

Teachers (Faculty Members of IIT Kharagpur):

Pabitra Mitra (Coordinator), CSE

Debasis Samanta, CSE

Saptarshi Ghosh, CSE

Debasis Sen, ECE

Jiaul Paik, GSST

Alok K Deb, EE

Debdoot Sheet, EE

Sujoy Bhattacharya, VGSOM

Swagata Chatterjee, VGSOM

PhD Students (Teaching Assistants)

Program: Day 1-2

9:00-10:00	Introduction	Pabitra Mitra
10:00-11:00	Regression - I	Debasis Samanta
11:00-13:00	K-Nearest Neighbor	Swagato Chatterjee
14:00-16:00	Regression - II	Debasis Samanta
16:00-18:00	Clustering	Saptarshi Ghosh

9:00-11:00	Feature Engineering and Data Preprocessing	Pabitra Mitra
11:00-13:00	Regression Hands On	Debasis Samanta
14:00-16:00	Text Analytics	Jiaul Paik
16:00-18:00	Lucene Hands On	

Program: Day 3-4

9:00-11:00	Decision Tree	Sujoy Bhattacharya
11:00-13:00	Recommender System	Pabitra Mitra
14:00-16:00	Recommender System Hands On	
16:00-18:00	Ensemble Methods	Sujoy Bhattacharya

9:00-11:00	SVM	Alok Deb
11:00-13:00	Time Series Analysis	Sujoy Bhattacharya
14:00-16:00	SVM Hands On	
16:00-18:00	Neural Networks - I	Alok Deb

Program: Day 5-6

9:00-11:00	Test	
11:00-13:00	Neural Networks - II	Alok Deb
14:00-16:00	Bayes Classifier	Pabitra Mitra
16:00-18:00	Deep Learning - I	Debasis Sen

9:00-11:00	Deep Learning - II	Debdoot Sheet
11:00-13:00	Case Study Project Discussion	Pabitra/Sujoy
14:00-18:00	Case Study Project	Pabitra Mitra

Introduction to Machine Learning

PABITRA MITRA

COMPUTER SCIENCE AND ENGINEERING

IIT KHARAGPUR

PABITRA@CSE.IITKGP.AC.IN

Machine Learning

A branch of Artificial Intelligence, concerned with the design of algorithms that allow computers to evolve behaviors (learn) based on empirical data (experience)

Formally: 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.

Experience

Adaptation

Generalization

Why do we need to learn?

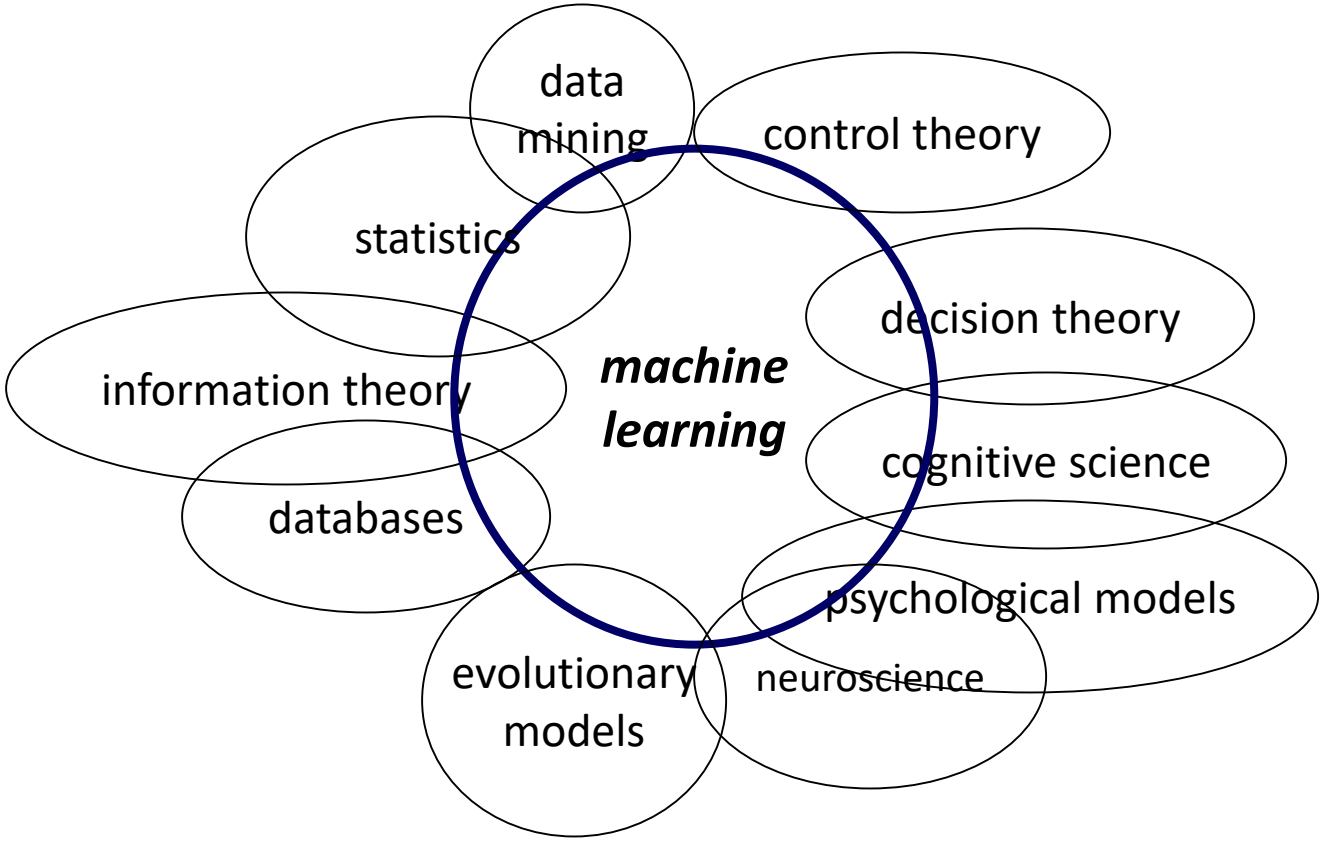
Incomplete information about the environment

A changing environment

Use the sequence of percepts to estimate the missing details

Hard for us to articulate the knowledge needed to build AI systems – e.g. try writing a program to recognize visual input like various types of flowers

Related Fields



ML in Practice

Understanding domain, prior knowledge, and goals

Data integration, cleaning, pre-processing, selection

Feature selection – some features are relevant

Learning models

Evaluating performance and interpreting results

Consolidating and deploying discovered knowledge

Learning Paradigms

Supervised learning- Given a set of input/output pairs, learn to predict the output if faced with a new input

- Classification
- Regression

Unsupervised Learning- Learning patterns in the input when no specific output values are supplied.

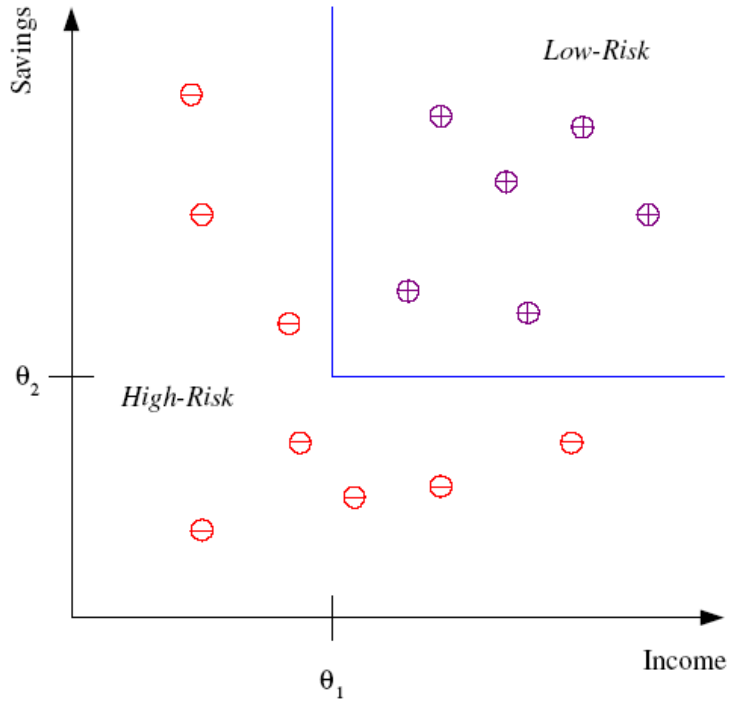
- Clustering
- Anomaly detection

Reinforcement Learning- Learn to interact with the world from the reinforcement you get.

Classification

Example: Credit scoring

Differentiating between low-risk and high-risk customers from their *income* and *savings*



Discriminant: IF *income* $> \theta_1$ AND *savings* $> \theta_2$
THEN low-risk ELSE high-risk

Model

Classification Models

Bayes Classifier

Decision Tree

K-Nearest Neighbor

Support Vector Machine

Neural Networks

Classification: Applications

Face recognition: Pose, lighting, occlusion (glasses, beard), make-up, hair style

Character recognition: Different handwriting styles.

Speech recognition: Temporal dependency.

- Use of a dictionary or the syntax of the language.
- Sensor fusion: Combine multiple modalities; eg, visual (lip image) and acoustic for speech

Medical diagnosis: From symptoms to illnesses

Web Advertising: Predict if a user clicks on an ad on the Internet.

Prediction: Regression

Example: Price of a used car

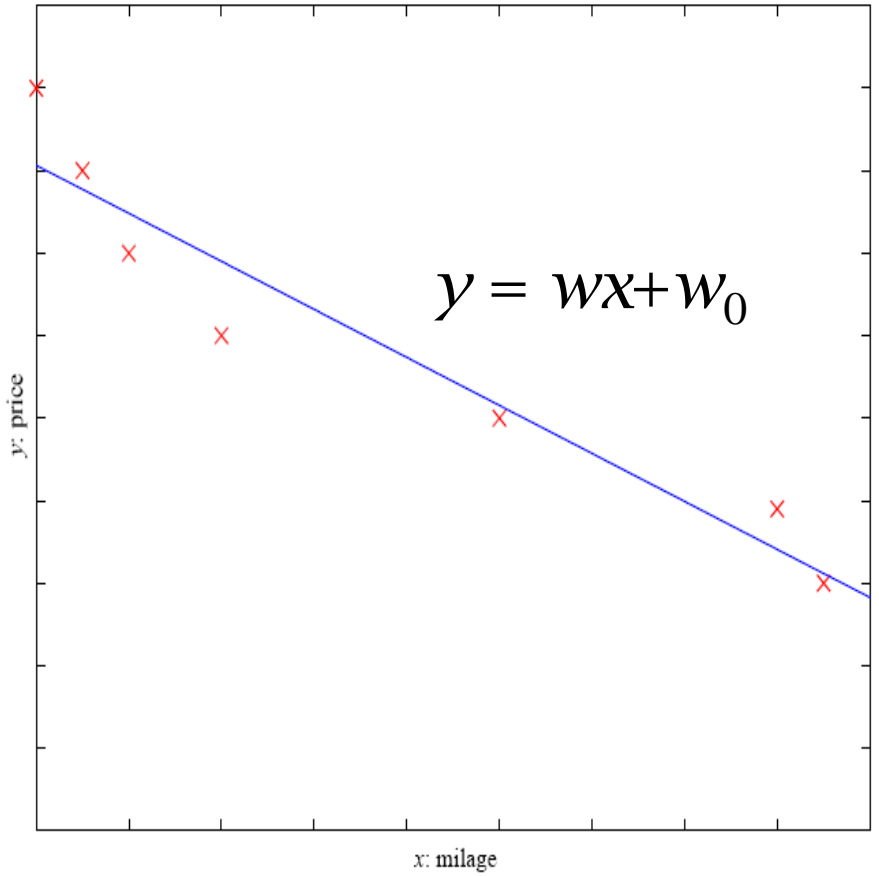
x : car attributes

y : price

$$y = g(x \mid \vartheta)$$

$g(\ )$ model,

ϑ parameters



Unsupervised Learning

Learning “what normally happens”

No output

Clustering: Grouping similar instances

Other applications: Summarization, Association Analysis

Example applications

- Customer segmentation in CRM
- Image compression: Color quantization
- Bioinformatics: Learning motifs

Reinforcement Learning

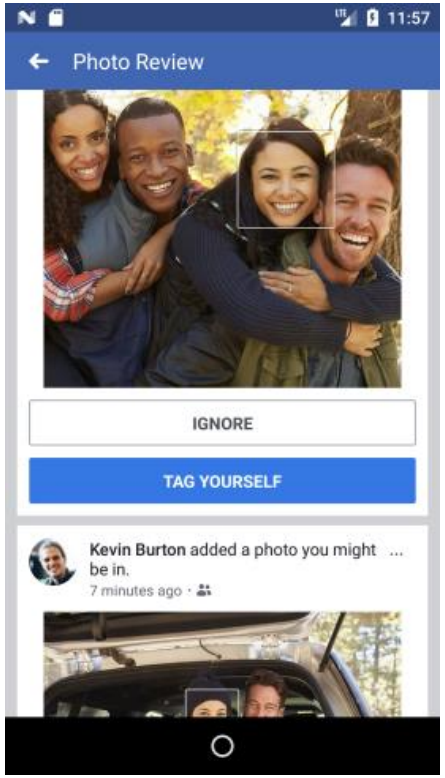
No supervised output but delayed reward

- Policies: what actions should an agent take in a particular situation
- Utility estimation: how good is a state (→used by policy)

Applications:

- Game playing
- Robot in a maze
- Multiple agents, partial observability, ...

Facebook Face Recognition

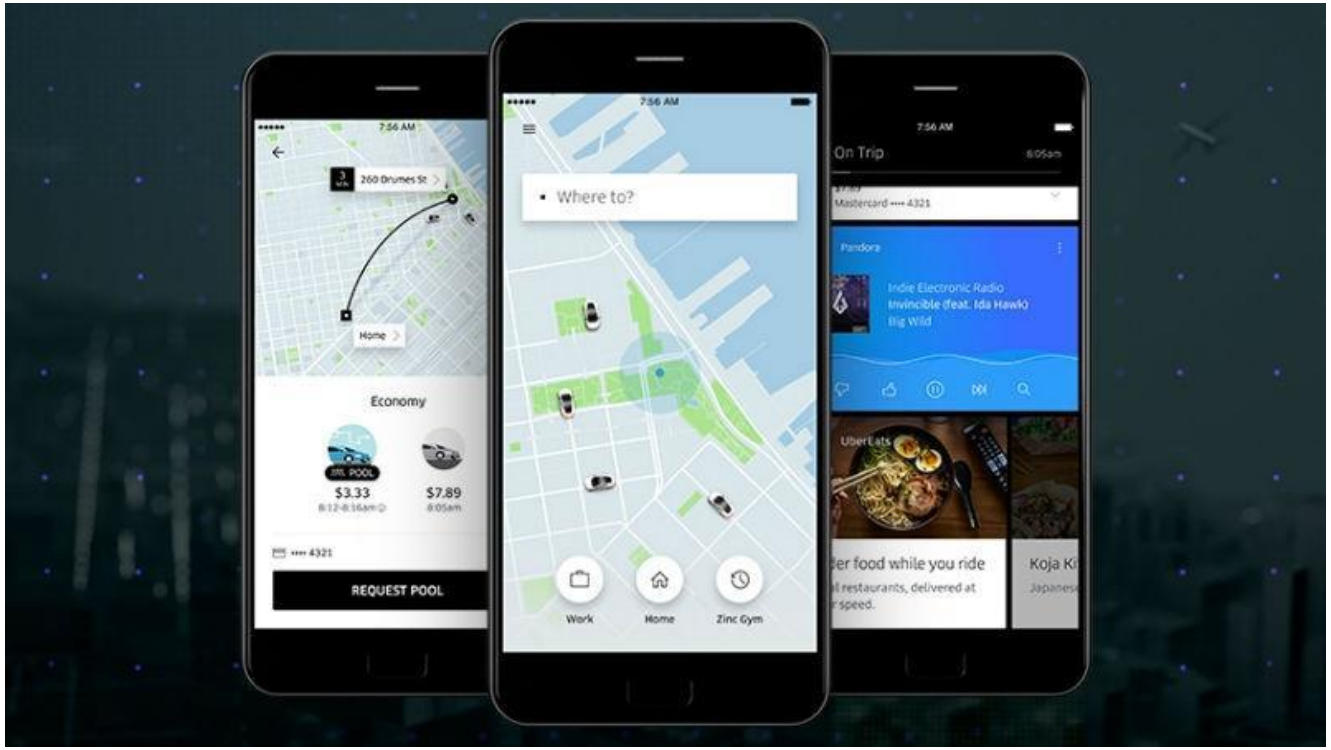


Deep neural network model: Convolutional Neural Network

Google Voice Recognition

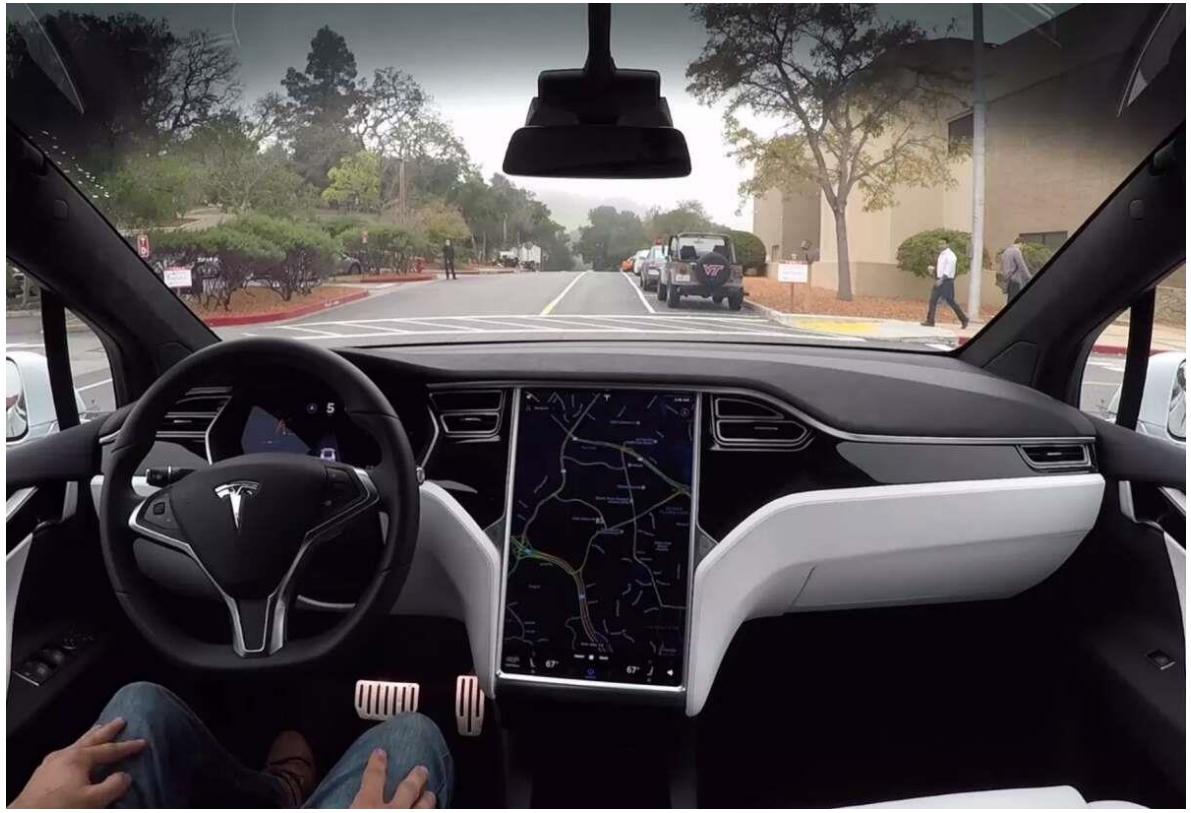
Deep neural network: Recurrent Neural Network

Uber Route Suggestions



Mobility models

Tesla Autonomous Cars

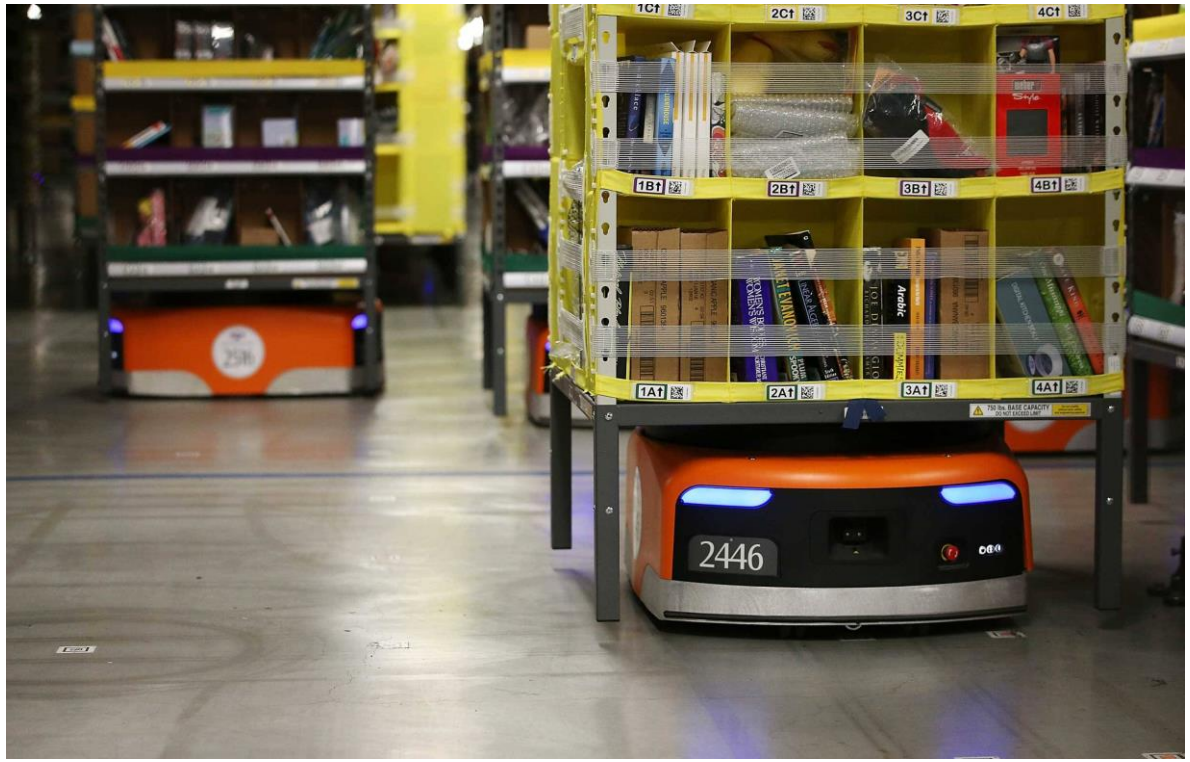


Challenge in machine learning

Amazon Recommendation Systems

Statistical models of buyers

Robotic Warehouse Delivery



Reinforcement learning

Cyclone Track Prediction

Prediction problem

Deep Dream Painting Generator

Generative Adversarial Networks

Further Reading

Text Book: Machine Learning – Tom Mitchell

Thank You!

Many Interesting Challenges Ahead!