Introduction to Machine Learning

VIJAY JAISANKAR

Before we start…

- **Official Github Repository:**

https://github.com/vijay-jaisankar/ML_TA_IIITB_2022

- **Official communications:**

Slack and LMS

- **Find Good Resources?**

Make a PR! https://github.com/zense/helpful-resources

FAQs

- Do I need a fancy computer? No.

- Do I need to know programming? Yes.

- Do I need to know a lot of maths? It's complicated.

Lecture Plan

Some gradient descent.
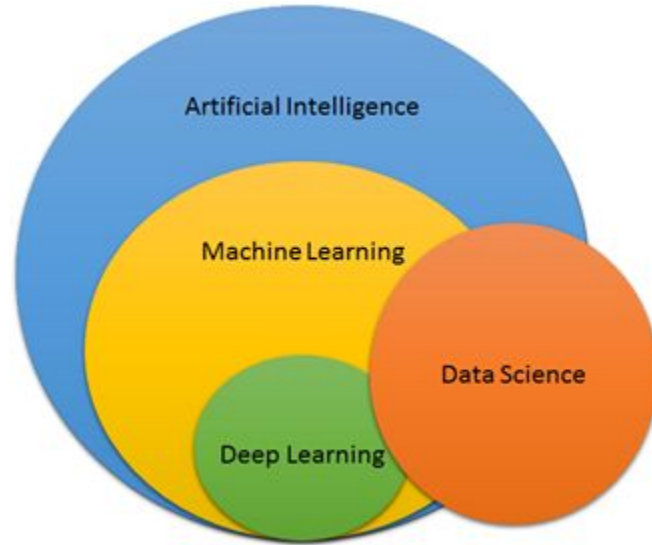
Honest advice and tools to get you started, too!

So, what actually is ML?

What is AI?

Are they the same?

Is it a scam?

What is life?

Introductory Example (Also: PTSD)

Given the `root` of a binary tree, *determine if it is a valid binary search tree (BST)*.

A **valid BST** is defined as follows:

- The left subtree of a node contains only nodes with keys **less than** the node's key.
- The right subtree of a node contains only nodes with keys **greater than** the node's key.
- Both the left and right subtrees must also be binary search trees.

How would you approach this problem?

Are there solid and defined rules?

What rules will you define?
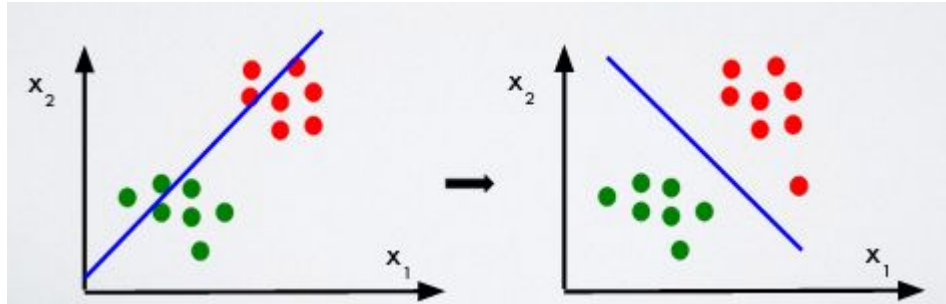
What rules can we define for this problem?
Where do we stop?

What test-cases do we check our solution's correctness with?

What if?

What if someone or something can analyse the data and get the rules for you?

Analogy: Leetcode Discuss

Machine Learning

ML consists of a set of algorithms
that allow software applications to become more accurate at
predicting outcomes **without being explicitly programmed**
to do so.

Basic Components

- Tasks

- Models

- Features

- Datasets

???

A little more detail

- Tasks == Problems you wish to apply Machine Learning on; clear declaration and definition of inputs and outputs

- Models == Algorithms run on data that generate insights

- Features == Filtered and Processed Inputs

- Datasets == "Raw" Data

Pandas, Numpy, Matplotlib, Kaggle, Colab

Introduction

- Pandas - Dataset operations

- Numpy - Mathematical functions

- Matplotlib - Plots and Charts

- Kaggle - Hosts Datasets, Notebooks, and Contests

- Colab - Interface Google Drive with Jupyter

Jupyter Notebooks

- Interactive Python

- Local server on browser

- VSCode Plugin

- Kaggle/Colab

Let's get our hands dirty!