



# CPSC 5310 MACHINE LEARNING WINTER 2026

DR. DIALA EZZEDDINE

**Week 1 Session 1**

# AGENDA

- Introductions
- Course expectations
- Intro to ML
- Python setup
- In-Class activity I
- Practice: EX I ( Homework)

# INTRODUCTIONS

## **Dr. Diala Ezzeddine**

- PhD in Computer Science
  - MS in Statistics
  - BS in Pure Mathematics

## Experience

- Data scientist, research scientist, product manager
- Teaching DS, ML, CS, Statistics, Analytics

# INTRODUCTIONS

- Name
- Where are you from?
- Work or student or both
- Major
- Prior knowledge in:
  - Statistics and Linear algebra
  - Programming language/s Python
  - Machine Learning (give an example)



# COURSE EXPECTATIONS



# COURSE DESCRIPTION

This course introduces machine learning foundations, concepts, and algorithms and their applications in analyzing massive amounts of data to find interesting patterns that can be used to assist decision-making or provide predictions. Topics include decision trees, Bayesian classification, clustering, sequence clustering, association rules, time series analysis, and neural networks. Students are expected to analyze real-world data.

# COURSE DESCRIPTION

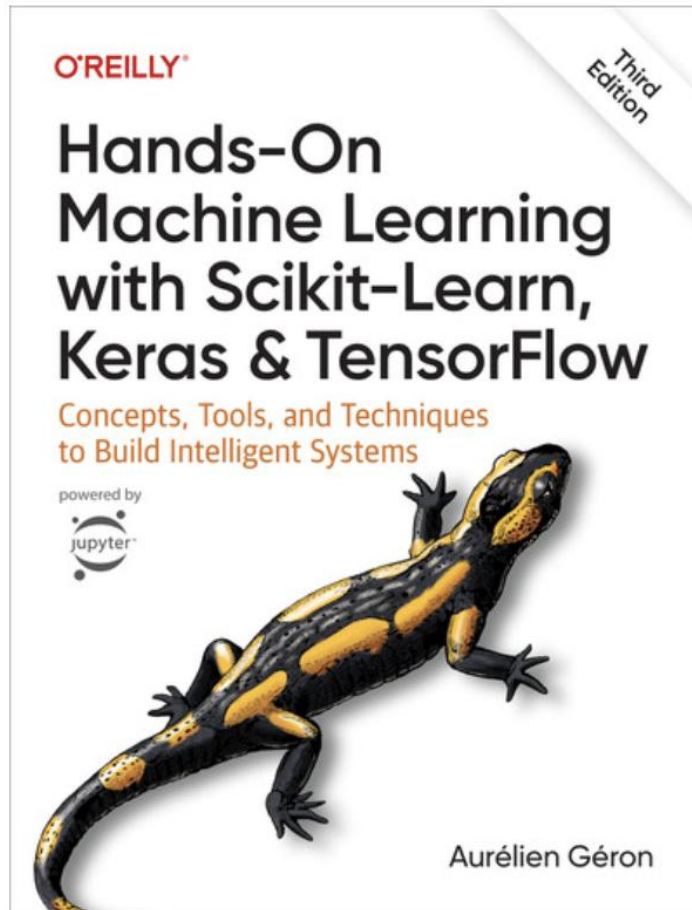
- This course focuses on the foundations of machine learning and concludes with an introduction to neural networks. Students will learn a broad range of supervised and unsupervised ML techniques, covering core algorithms, evaluation strategies, and practical tooling.
- To ground the concepts in real-world practice, we will walk through a complete machine learning project end-to-end. The goal is to demonstrate the key stages of an ML workflow and support students in planning their own quarter-long group ML project.
- The group project is a central component of the course. Working in instructor-assigned teams of two, students will select a dataset from a curated set of four provided options and define an approved problem to investigate. Each group will apply relevant ML algorithms and evaluation methods to develop and validate solutions and will present their findings at the end of the quarter.

# CLASS OVERVIEW

- **Canvas is the official location of everything!**
  - Syllabus
  - Lectures
  - Grading
  - Assignments
  - Discussions
  - Textbook and Resources



# TEXTBOOK



- Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow by Aurélien Géron, 3rd edition, O'Reilly Media.

# GRADING

Evaluation will be based on class activities, homework assignments, quizzes, and project deliverables:

- Project: 30%
- Data Camp courses: 15%
- Write-ups of assigned readings: 15%
- Quizzes: 25%
- Participation: 15%

# COURSE POLICIES AND EXPECTATIONS

- **Attendance:** Attendance are required. If you can't attend, you need to inform your instructor as soon as possible.
- **Assignments:** Late work will not be accepted.
- **Using AI Applications in Written Assignments:** Using AI/Generative AI (GenAI) applications to assist with written assignments and project deliverables is permitted, as long as it does not result in plagiarism. You are responsible for ensuring that all submitted work is original, accurate, and properly attributed.
- Course Policies

# OUR PRINCIPLES

- We will do our best to create a good learning experience.
- We will support each other, as humans.
  - Recognizing that many of us may be going through challenging situations.
- We will prioritize clear communication, accessible content and resources.
- We will remain flexible and adjust to the situation.

# BUILD A GREAT COMMUNITY

- **Collaborative learning can enhance our learning**
  - which involves joint intellectual effort by students and instructors together.
  - Help out your peers through discussions during class and outside class!
  - Still “being able to work independently” is important for your professional career!
- **Be mindful of the tone you use**
  - be respectful and supportive, help everyone feel at home.
- **Watch out for implicit bias; catch yourself before acting on it.**
  - Someone’s gender, race, ethnicity, sexual orientation, etc. do NOT have anything to do with how awesome they will be in this class.
  - Having a ton of programming experience will help some with projects, but does NOT give anyone an edge on how well they can understand the material and how highly they can score on the exams.

# WHAT TO EXPECT IN EACH WEEK

- Weekly lectures to explain and discuss the weekly topic/Chapter
- Weekly lab/coding exercises
- Weekly readings and Discussions
- Weekly online course/modules through DataCamp

# COURSE TENTATIVE SCHEDULE

Week 1: Overview of Machine Learning

Week 2: End to end ML project

Week 3: Classification

Week 4: Training Models

Week 5: Unsupervised Learning

Week 6: Decision Trees

Week 7: Ensemble Methods

Week 8: Support Vector Machine

Week 9: Dimensionality Reduction

Week 10: Neural Networks

Evaluation Week: Final presentation (6:00 - 7:50 pm on Tuesday March 17)

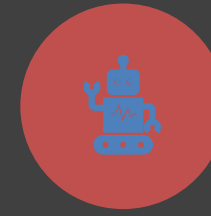
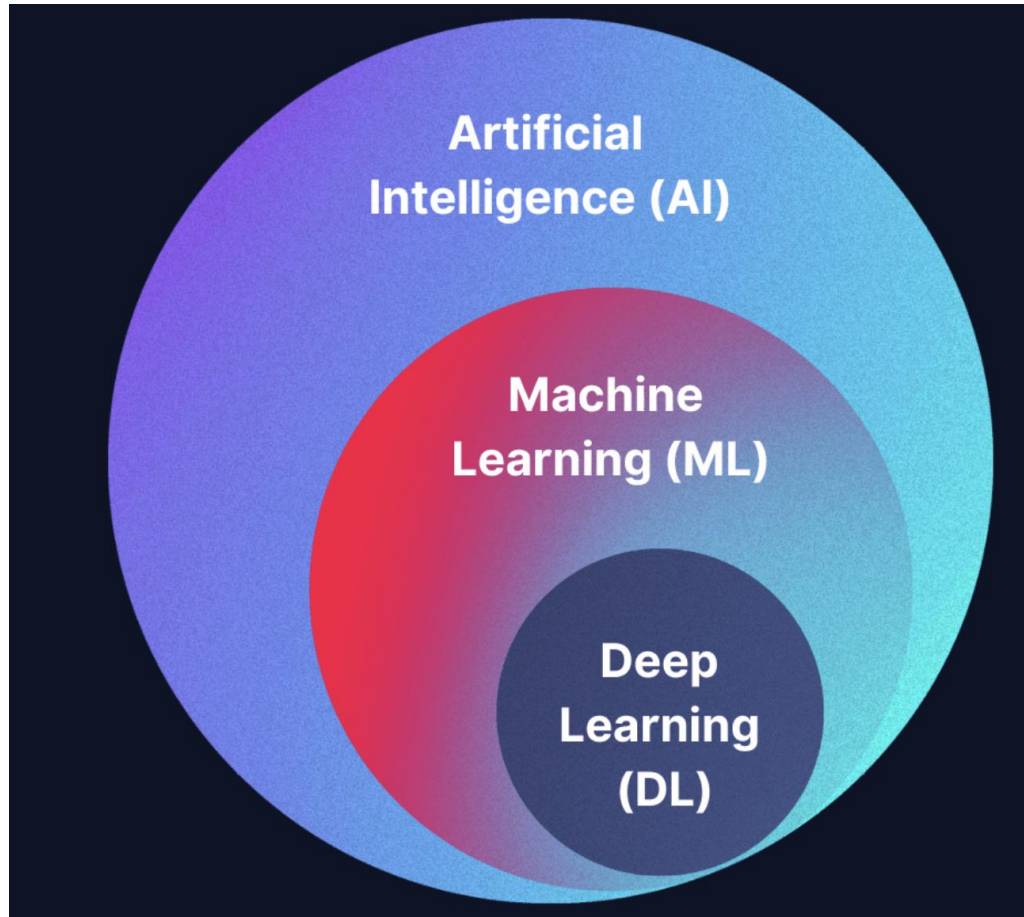


# INTRO TO MACHINE LEARNING

Reference:

Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow by Aurélien Géron, 3rd edition, O'Reilly Media.





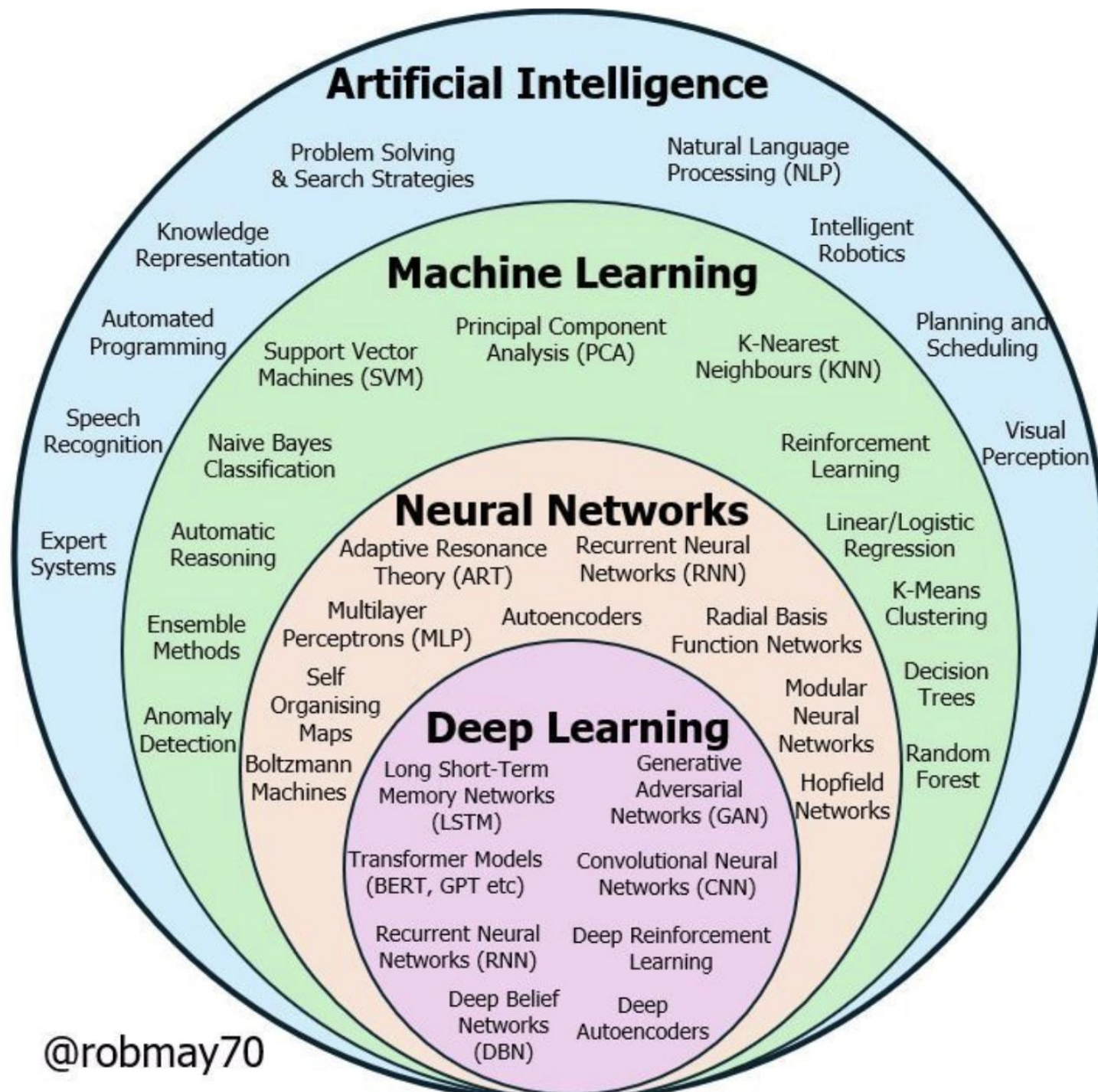
AI IS THE STUDIES OF  
HOW TO CREATE  
COMPUTER  
SOFTWARE THAT ARE  
CAPABLE OF  
INTELLIGENCE  
BEHAVIOR.



MACHINE LEARNING  
IS A BRANCH OF  
ARTIFICIAL  
INTELLIGENCE.



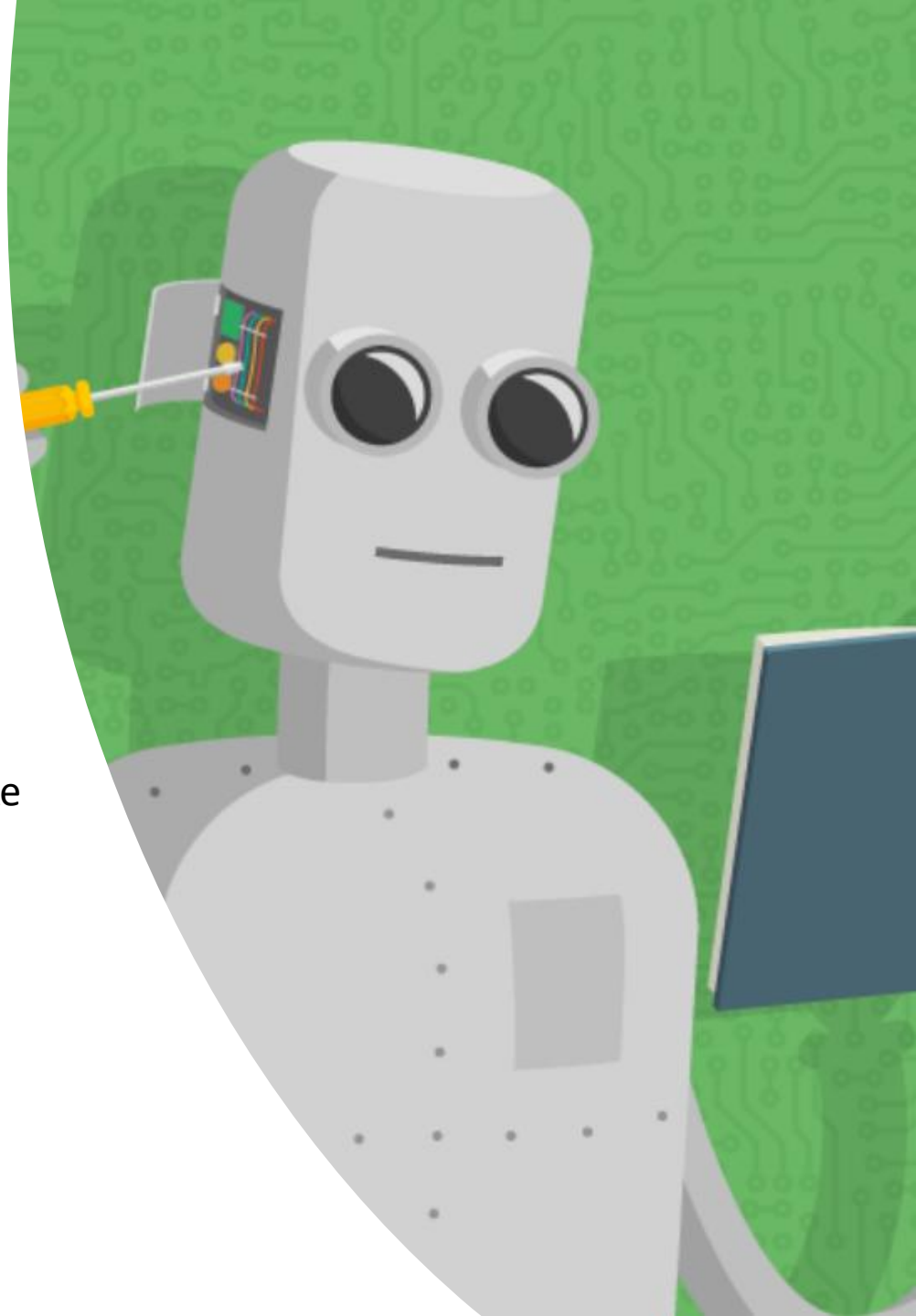
DL IS PART OF A  
BROADER FAMILY OF  
MACHINE LEARNING  
METHODS BASED ON  
ARTIFICIAL NEURAL  
NETWORKS.



# Machine Learning Definitions

---

- ML is known as **the science of learning**.
- **ML is the study, the design and the development of the algorithms that give computers the capability to learn without being explicitly programmed.**
- Machine learning is a way of teaching computers to make prediction based on some data. As more data is added and used in this process, the prediction improves.
- ML is the study of computer algorithms that improve automatically through experience.
- ML is a field of CS that uses statistical techniques to give computer systems the ability to learn with data, without being explicitly programmed.(wiki)
- ...







# Machine Learning

- **Arthur Samuel (1959)** “Machine learning: Field of study that gives computers the ability to learn without being explicitly programmed”
- **Tom M. Mitchell(1997)** ” Well-posed Learning problem: 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  $E$ .”

# MACHINE LEARNING

Based on Mitchell's definition "*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  $E$ .*",

Can you identify what is the task  $T$ , the experience  $E$  and the performance measure  $P$  in the following examples?

1. Suppose your email program watches which emails you do or do not mark as spam and based on that learns how to better filter spam.  
(Andrew Ng)
2. Suppose you build a model that learns from historical data and updates itself as new daily data arrives to improve future predictions.

# WHY USE MACHINE LEARNING?

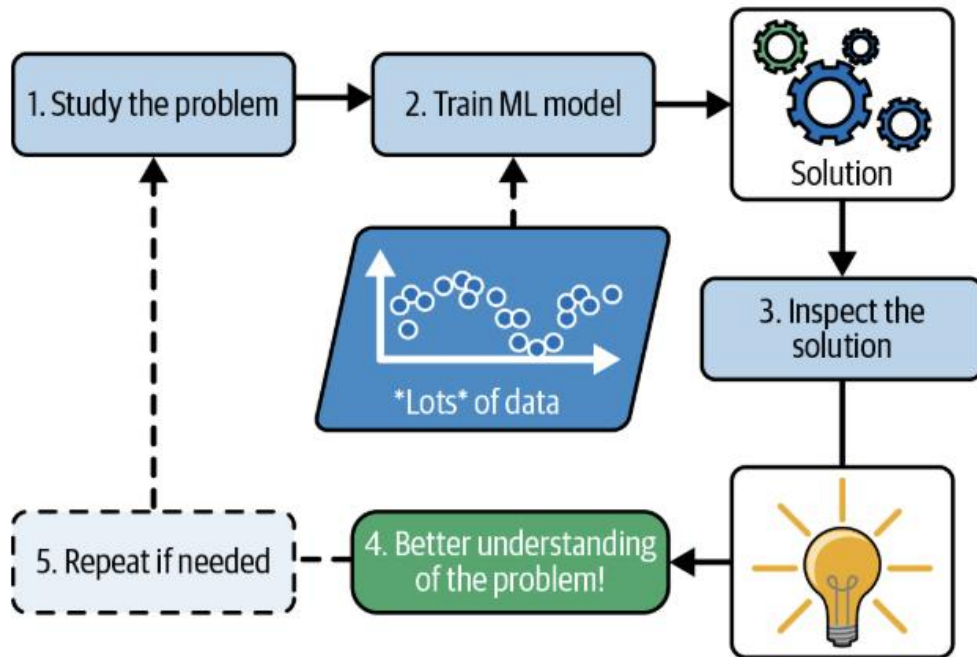


Figure 1-4. Machine learning can help humans learn

Machine Learning is great for:

- Problems for which existing solutions require a lot of work and maintenance, such as long lists of rules.
- Complex problems for which using a traditional approach yields no good solution.
- Fluctuating environments: A machine learning system can easily be retrained on new data, always keeping it up to date.
- Getting insights about complex problems and large amounts of data.

# Application of ML

We all use ML on a daily basis.

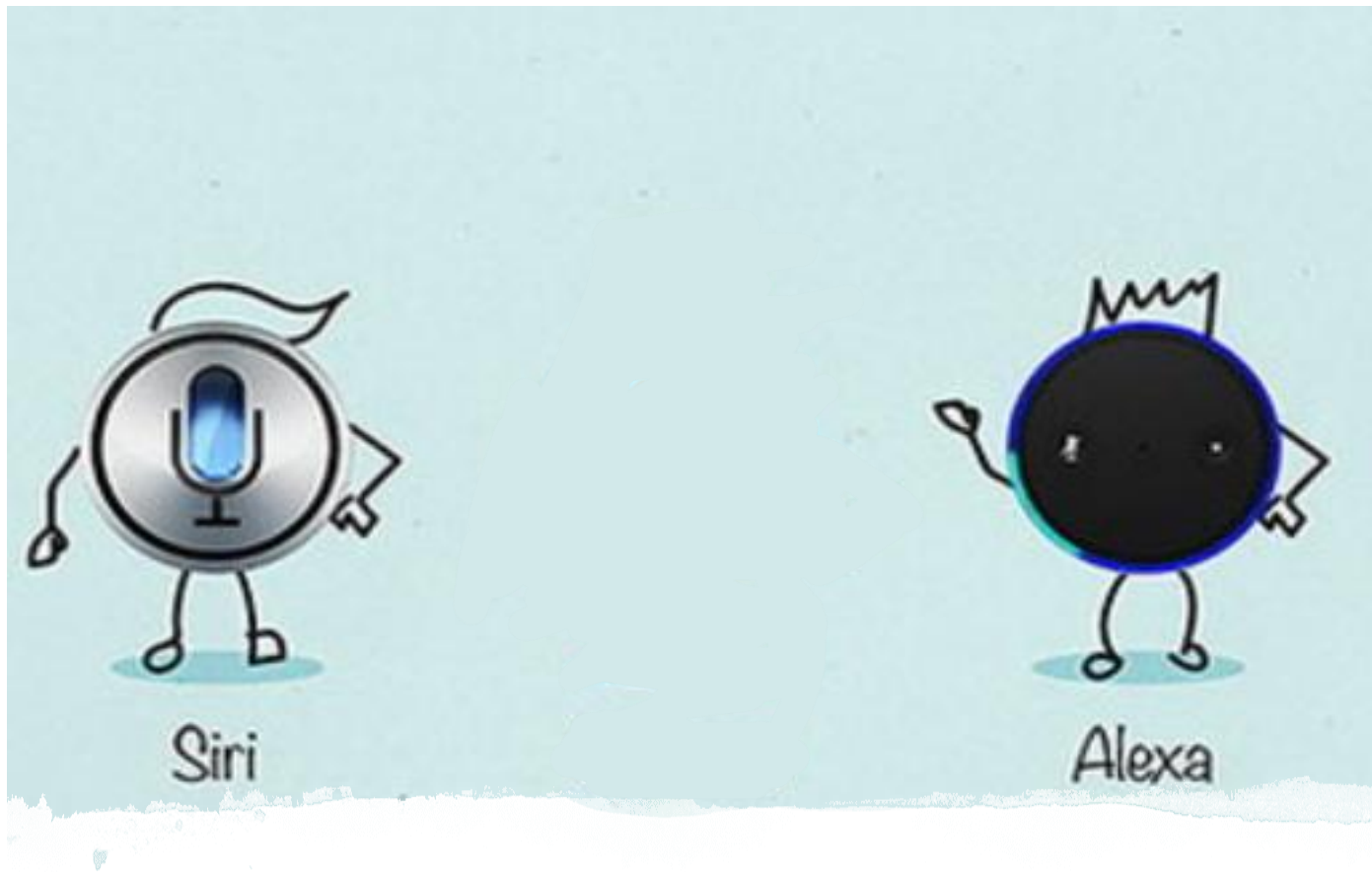
Examples:

- Spam filtering
- Credit card fraud detection
- Digit recognition on checks, zip codes
- Detecting faces in images
- MRI image analysis
- Recommendation system
- Search engines
- Handwriting recognition
- Scene classification
- ...

# EXAMPLES OF ML APPLICATIONS FROM THE TEXTBOOK

- Analyzing images of products on a production line to automatically classify them
- Detecting tumors in brain scans
- Automatically classifying news articles
- Automatically flagging offensive comments on discussion forums
- Summarizing long documents automatically
- Estimating a person's genetic risk for a given disease by analyzing a very long DNA sequence
- Creating a chatbot or a personal assistant
- Forecasting your company's revenue next year, based on many performance metrics
- Making your app react to voice commands
- Detecting credit card fraud
- Segmenting clients based on their purchases so that you can design a different marketing strategy for each segment
- Representing a complex, high-dimensional dataset in a clear and insightful diagram
- Recommending a product that a client may be interested in, based on past purchases
- Building an intelligent bot for a game





## Application of ML: **Virtual assistance**

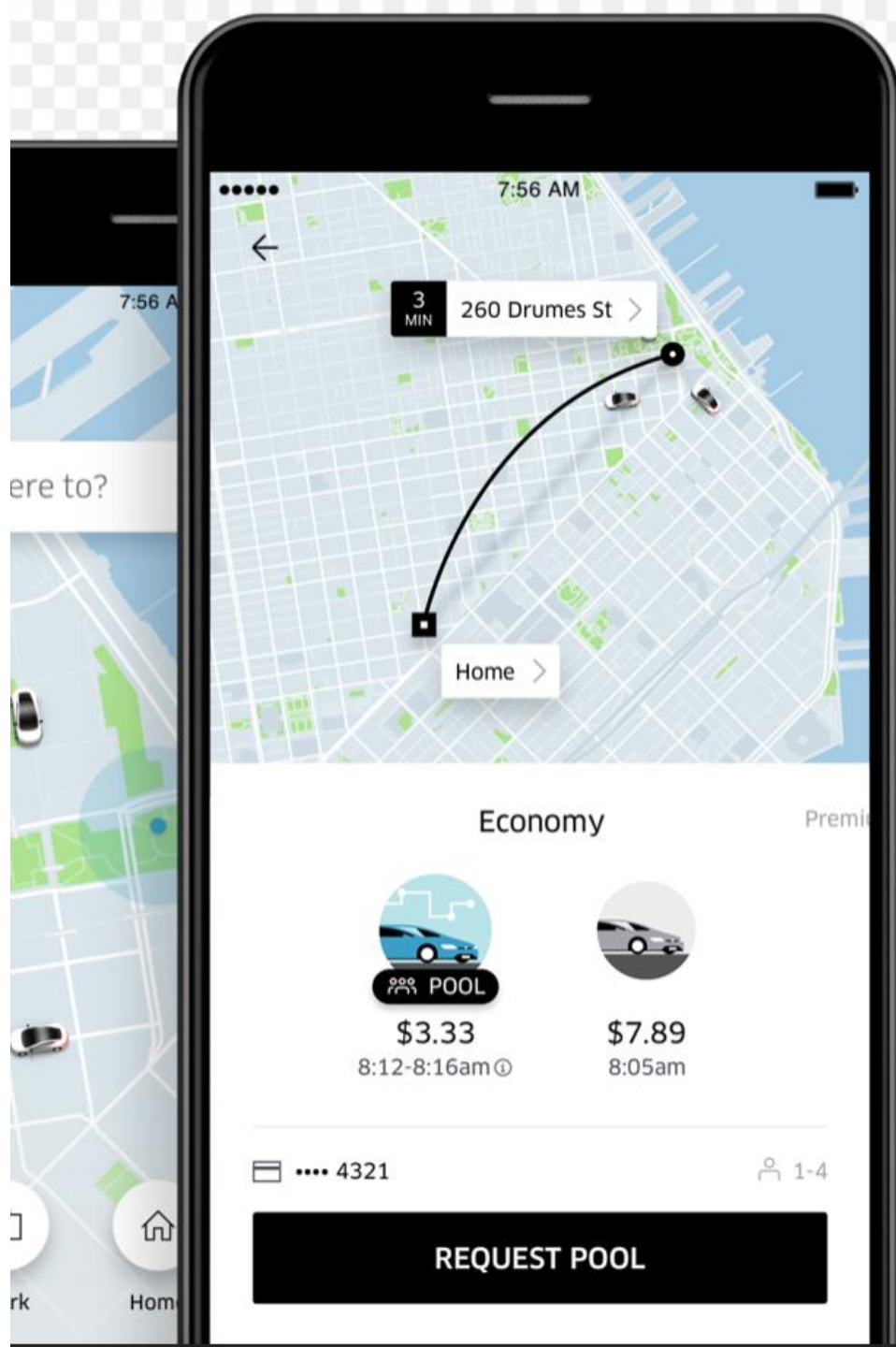
Siri and Alexa are using ML when they collect data and refine the information on the basis of previous involvement with them.

# Application of ML: Traffic predictions

---

ML help estimating the regions where congestion can be found on the basis of daily experiences using data collected from us using a GPS navigation system.





# Application of ML: **Online Transportation Networks**

Uber use ML to define price surge hours by predicting the rider demand.



# Application of ML: **Video Surveillance**

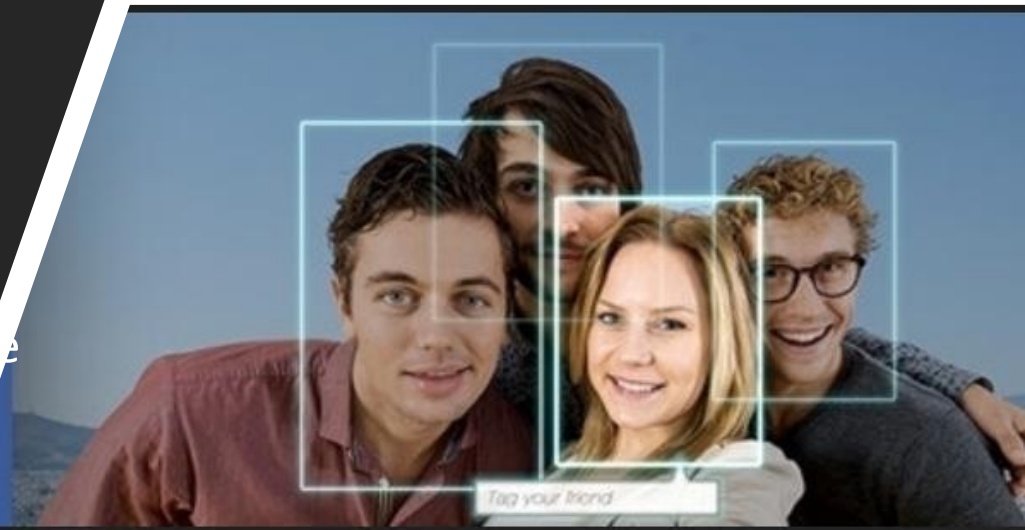
ML is used to detect crime  
before they happen.



# Application of ML: Social Media

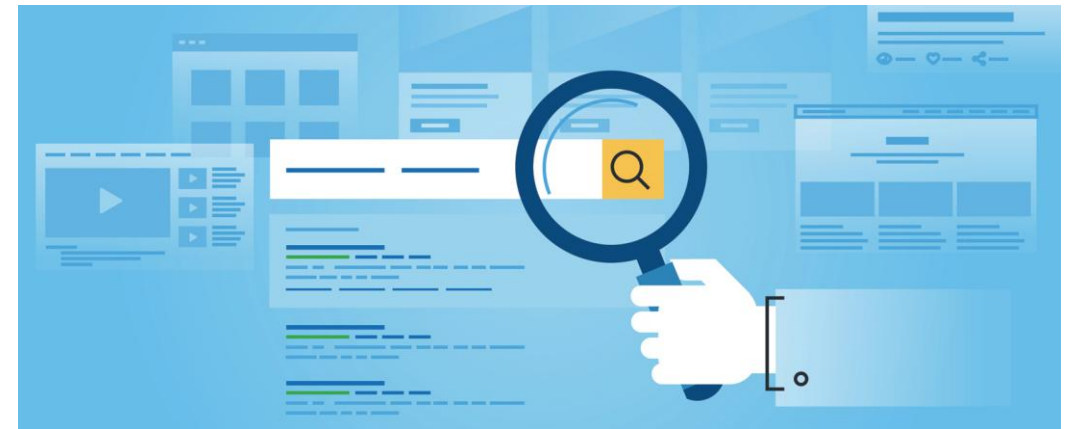
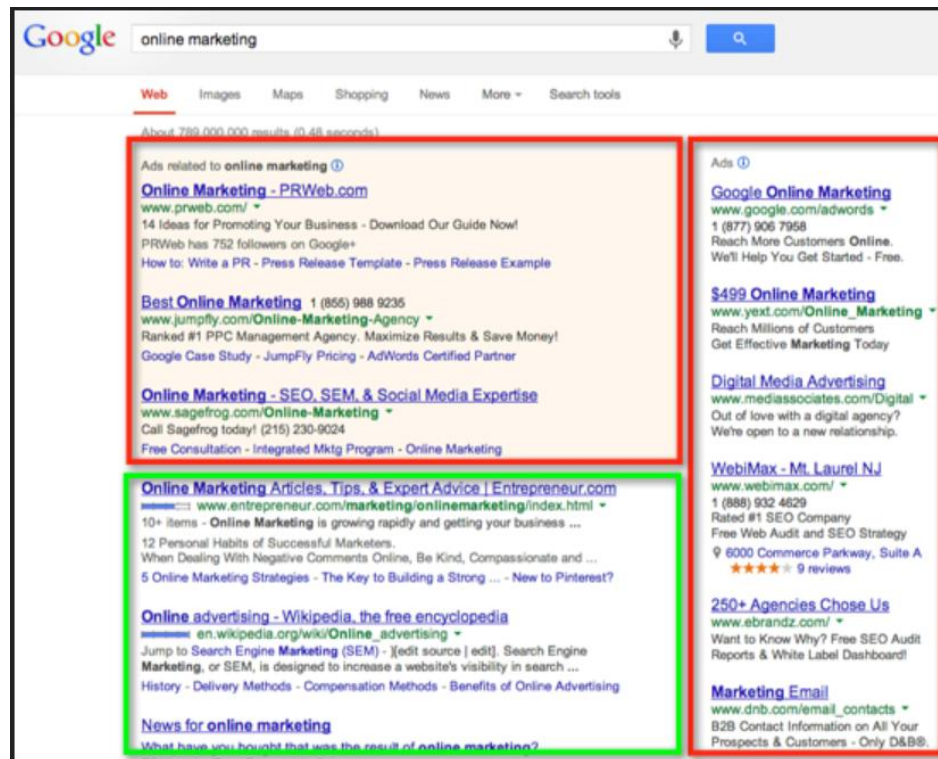
## Facebook

- **Friend suggestion:** Facebook continuously notices the friends that you connect with, the profiles that you visit very often, your interests, workplace, or a group that you share with someone etc. A list of Facebook users are suggested that you can become friends with based on continuous learning.
- **Face recognition:** Facebook checks the poses and projections in the picture, notice the unique features, and then match them with the people in your friend list.



# Application of ML: Search Engines Google

Google use machine learning to improve the search results for you.



# Application of ML: Product Recommendations Amazon

Amazon use ML to recommend new product that you may need to buy.







amazon [Try Prime](#) [Your Amazon.com](#) [Today's Deals](#) [Gift Cards](#) [Sell](#) [Help](#) [Celebrate 4<sup>th</sup> of July](#) [Shop now](#)

Shop by Department [Clothing & Accessories](#) [Go](#) Hello, Sign in [Your Account](#) Try Prime [Cart](#) [Wish List](#)

[Clothing, Shoes & Jewelry](#) [Clothing](#) [Shoes](#) [Handbags](#) [Accessories](#) [Luggage](#) [Jewelry](#) [Watches](#) [New For You](#)

## NEW DRESSES

Fit-and-flare styles, floral picks, and more. [See more](#)



### MAGGY LONDON

#### Maggy London Women's Cap Sleeve Draped Dress

[Be the first to review this item](#)


Price: **\$53.10** & **FREE Shipping**. [Details](#)  
List Price: \$98.00  
You Save: \$44.90 (45%)

**Size:**  
[Select](#) [Sizing info](#)


**Color:** lawn

- 95% Polyester/5% Spandex
- Imported
- Hand Wash
- Top lined in tricot; skirt unlined
- No zipper


**Wear it With**




STEVEN by Steve  
Madden  
\$50.21 - \$97.99



FRYE  
\$260.96



Lucky Brand  
\$52.99



XOXO  
\$19.99

[Share](#) [Email](#) [Facebook](#) [Twitter](#) [Pinterest](#)

To buy, select **Size**  
Choose from options to the left

[Add to Cart](#)

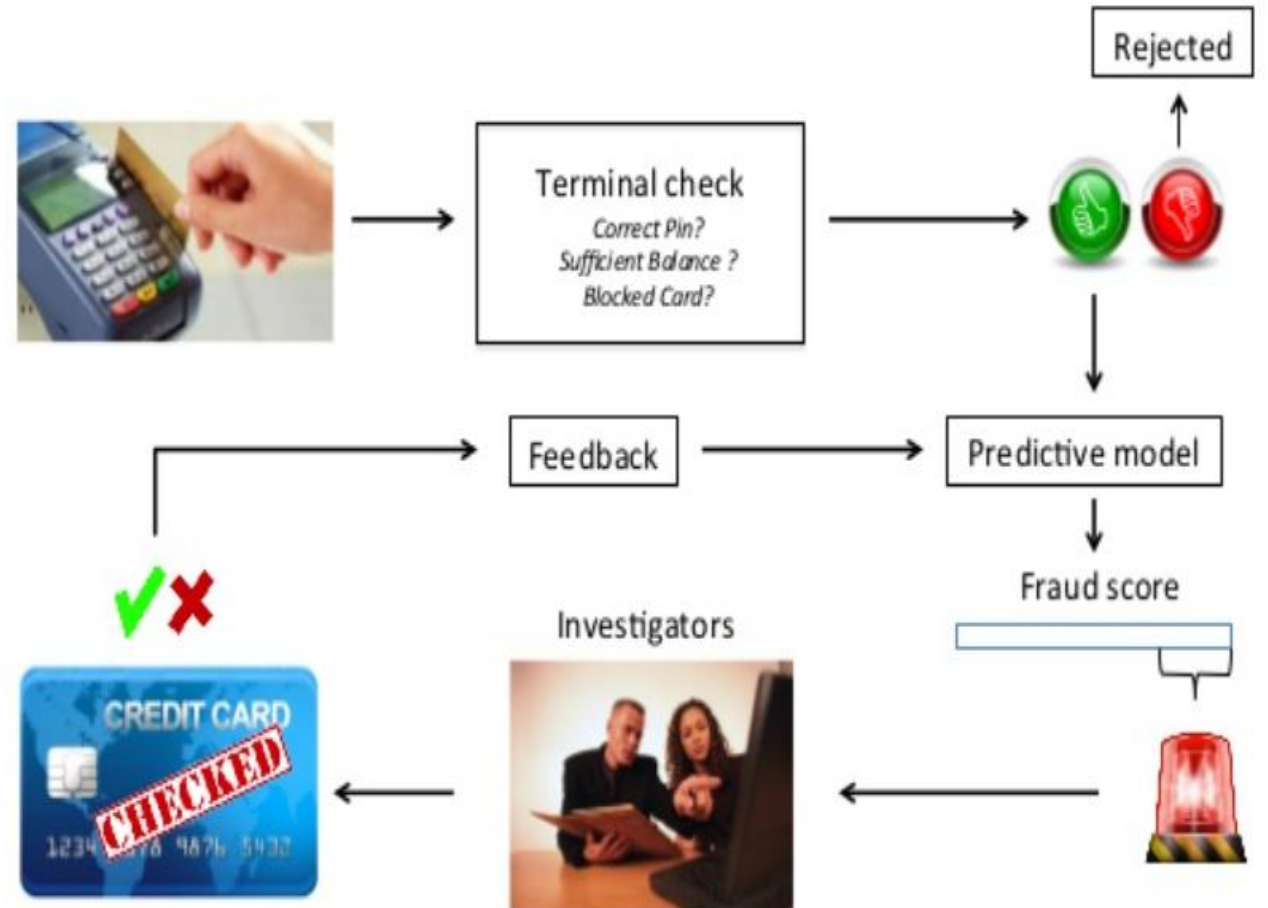
[Add to Wish List](#)



## Application of ML: Fraud detection Paypal

**Paypal** uses a set of tools that helps them to compare millions of transactions and distinguish between legitimate or illegitimate transactions taking place between the buyers and sellers.

### THE FRAUD DETECTION PROCESS

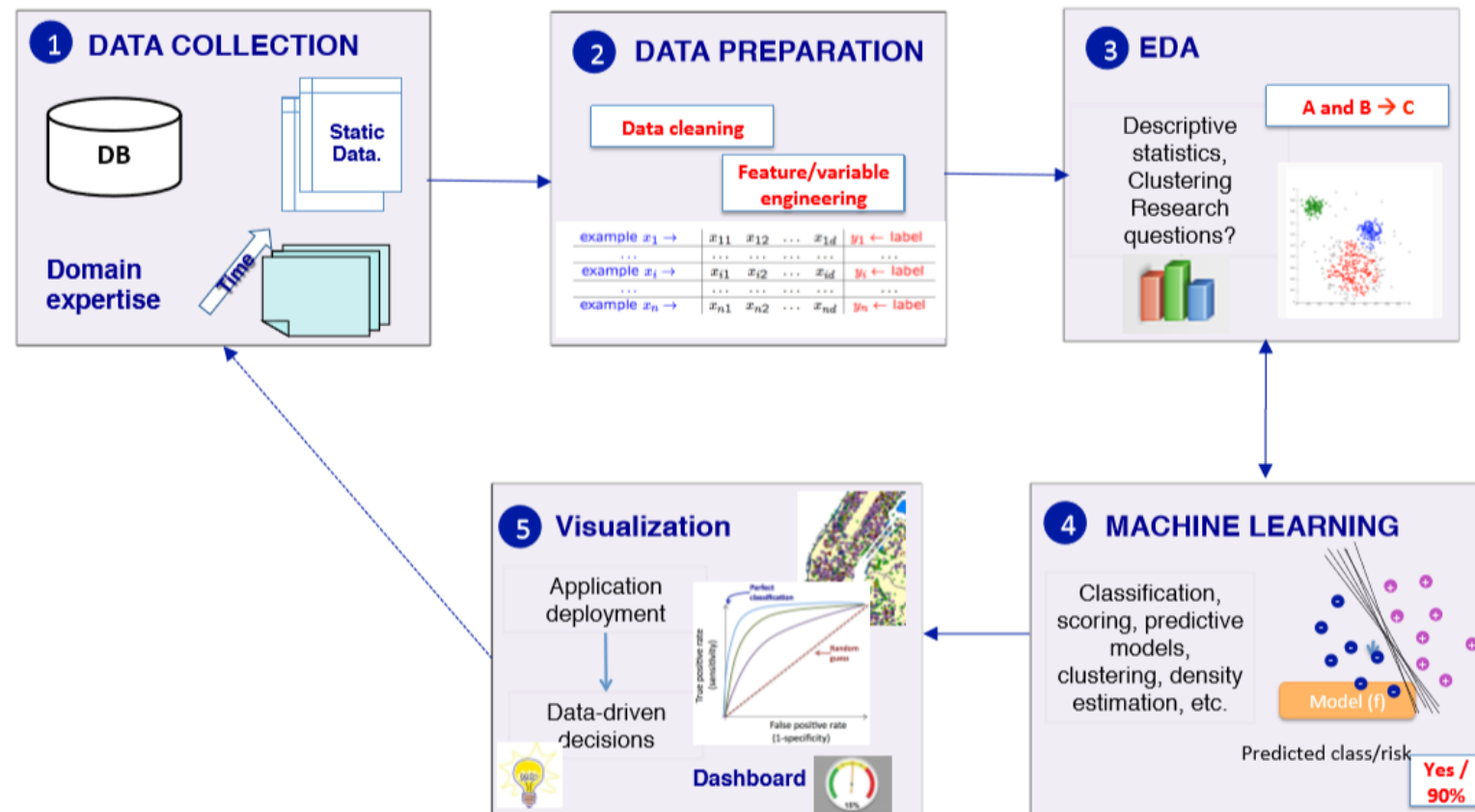




# VIDEO ABOUT MACHINE LEARNING

- ML Andrew Ng
- ML Oxford U

# Simplified ML workflow





# BREAK

Please use the break to introduce yourself to your classmates.



# PYTHON SETUP



# WHY PYTHON?

- Python is a scripting language, can be used to quickly write small programs, or scripts to automate other tasks.
- Python is generally the second-best language for everything.
- Python has developed a large and active scientific computing and data analysis community.
- Millions of users who are happy to offer advice or suggestions when you get stuck on something.
- Python offers a wide range of powerful libraries which have made it a popular choice for data analysis tasks. E.g., NumPy, SciPy, Pandas, Matplotlib and scikit-learn
- Data scientists in almost every sector will find packages already tailored to their needs freely available.

# JUPYTER NOTEBOOK

- It is a Web based, interactive computing notebook environment.
- It Edit and run human readable documents while describing the data analysis.
- The Jupyter notebook combines two components:
  - **A web application:** a browser-based tool for interactive authoring of documents which combine explanatory text, mathematics, computations and their rich media output.
  - **Notebook documents:** a representation of all content visible in the web application, including inputs and outputs of the computations, explanatory text, mathematics, images, and rich media representations of objects.
- The notebook extends the console-based approach to interactive computing in a qualitatively new direction, providing a web-based application suitable for capturing the whole computation process: **developing**, **documenting**, and **executing code**, as well as communicating the results.

# JUPYTER NOTEBOOK: MAIN FEATURES

- In-browser editing for code, with automatic syntax highlighting, indentation, and tab completion.
- The ability to execute code from the browser, with the results of computations attached to the code which generated them.
- Displaying the result of computation using rich media representations, such as HTML, LaTeX, PNG, SVG, etc. For example, publication-quality figures rendered by the [matplotlib](#) library, can be included inline.
- In-browser editing for rich text using the [Markdown](#) markup language, which can provide commentary for the code, is not limited to plain text.
- The ability to easily include mathematical notation within markdown cells using LaTeX and rendered natively by MathJax.


# JUPYTER DOCUMENTATION

- Markdown: [Jupyter\\_Markdown Documentation](#)
- Latex math symbols: <https://www.caam.rice.edu/~heinken/latex/symbols.pdf>
- Everything about Jupyter notebook: <https://docs.jupyter.org/en/latest/>



localhost:8888/notebooks/BUAN670/First%20session.ipynb#Text-font




 **jupyter** First session Last Checkpoint: 6 minutes ago (autosaved)




Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted

Python 3 



```
In [9]:  # Dev: Diala Ezzeddine  
#Date: 1/10/2010  
#Desc: My first code using Python  
from math import pi  
print("pi value is ", pi)  
print(pi**2)  
age = input("how old are you?")  
print(age)  
name = input("what is your name?")  
print(name.replace("d", "a"))
```

```
pi value is  3.141592653589793  
9.869604401089358  
how old are you?4  
4  
what is your name?diala  
aiala
```

# IN-CLASS ACTIVITY I

Open a new Jupyter Notebook file and name it: activity-I.ipynb and try to answer the following questions in python code:

## Exercise 1 — Input + Math + Strings

Start with the initial comment: to get personal information from the user

Write code to:

- Ask for name, age, height (cm), and weight (kg)
- Print lowercase, uppercase, and the length of the name
- Convert height to meters and print  $BMI = \text{weight} / \text{height}^2$
- Print weight divided by  $\pi$
- Print age in days
- Split the name into characters and print them one per line

## Exercise 2 — Two Number Calculator

Ask the user to input two numbers, then print:

- Sum, difference, product, quotient (division result)
- The results formatted like:

Sum: 10

Product: 24

# IN-CLASS ACTIVITY I

## Exercise 3 — List Processing

Prompt the user to enter 10 integers and store them in a list, then print:

- Mean (average)
- Min and Max
- Count of numbers > 50
- All odd numbers
- All numbers divisible by 3
- The list sorted (ascending and descending)
- The list reversed

## Exercise 4 — Function Practice

Create 3 functions:

```
def add(a, b): return a + b
```

```
def multiply(a, b): return a * b
```

```
def power(a, b): return a ** b
```

Then write a main program that:

1. Takes 2 integers from the user
2. Calls all 3 functions
3. Prints results

## Exercise 5 — Dictionary Storage

Create a program that asks for:

5 student names

1 score for each student

Store them in a dictionary like:

```
{"Alice": 88, "Bob": 92, ...}
```

Then print:

- Class average
- Student with highest score
- All students who scored above 90

# EXI-PANDAS DATAFRAMES

- You will be working on a set of problems for Python Series and Pandas.

**Part 0:** plot\_ex.ipynb

Practice Pandas Series and DataFrame with the given examples.

**Part I:** df\_ex.ipynb

Complete the two tasks given in the notebook.

- Submit a one paragraph write-up of what you learn from this exercise and the completed Notebook.



COMING UP...  
IN THE NEXT EXCITING EPISODE



# NEXT CLASS

- Week 1 Session 2: ML Types, Challenges and Testing
- Homework
  - Review the slides from today
  - DCI: Data camp course 1 due on Thursday Jan 8 at midnight.
  - Submit EX1- Pandas DataFrames