AI-INVASION 2019

A 5 day Introduction to Python and Machine Learning

A Presentation by Rising Odegua

About Data Science Nigeria

FOUNDED BY

Data Science Nigeria is a non-profit driven by a vision to build a world-class Artificial Intelligence knowledge, research and innovation ecosystem that delivers high impact research, business use applications, locally-relevant AI-first start-ups, increase employability and drive for social good use cases.
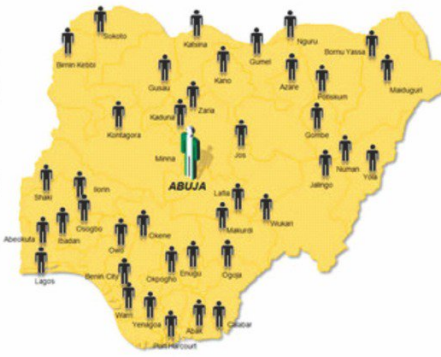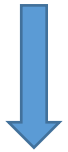
DR BAYO ADEKANMBI
CTO MTN NIGERIA

KEY MILESTONES OF DSN

- 300,000+ direct download of our Artificial Intelligence for Starter free ebook with contribution from recognized leaders like Prof Yoshua Bengio .

- 10,212 participated in the 1st ever Intercampus Machine Learning competition that involved 95 universities.

- 12,234 online participants in Data Science course

- Over 1,000 have participated in our face to face classes via fully residential all-expense-paid bootcamp, free meet-ups and weekly classes at the AI Hub.

- 203 direct job placement, project participation and internship.

- 4 innovative product ideas that are being pre-incubated

- The largest convergence of academia and industry practitioners as mentors to support real-world application of learning.

- High impact learning bootcamps, academic engagement and direct job placements (full time, freelance and internships) for young Nigerian data scientists

- Strategic partnership with leading firms and organisation. e.t.c

About Me

- I'm Human and Nigerian
- Name is Rising Osaboyen Odegua
- CS Undergrad at Ambrose Alli University (Finals, Class of 2019)
- Software Developer at IRSL Consulting Limited
- I play with Android
- Leads AI-PLUS community in AAU
- Data Scientist and Machine Learning Engineer
- Done a couple of research on Ensemble Techniques (yet to be published)
- Ranked 11th on Zindi Africa
- Won a couple Machine Learning Competition

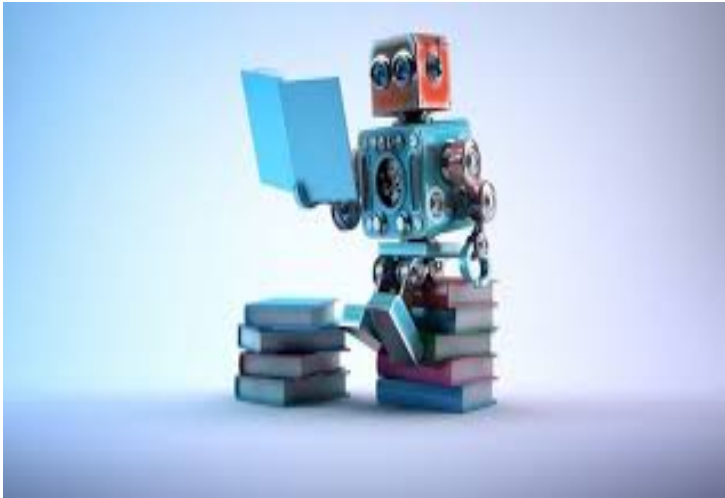 @risin_developer   @rising_developer   rising-odegua   Rising Odegua

DAY 1

CONTENT TO COVER

- Introduction to Machine Learning Part 1
- What is Machine Learning?
- What it does actually or why do we need ML or AI?
- Types of Machine Learning algorithms
- Intro to Jupyter Notebook
- Q & A

# Introduction to Machine Learning Part 1

is Machine Learning?

Machine Learning is a branch of Artificial Intelligence (AI), **which helps in making machines capable** of learning from observational data without being explicitly programmed.

"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."
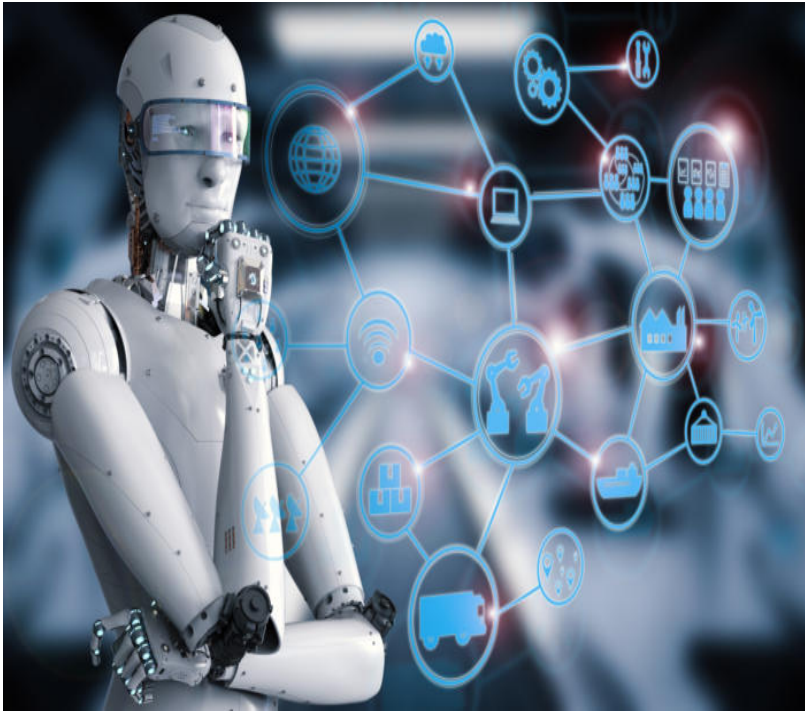
Tom Mitchell
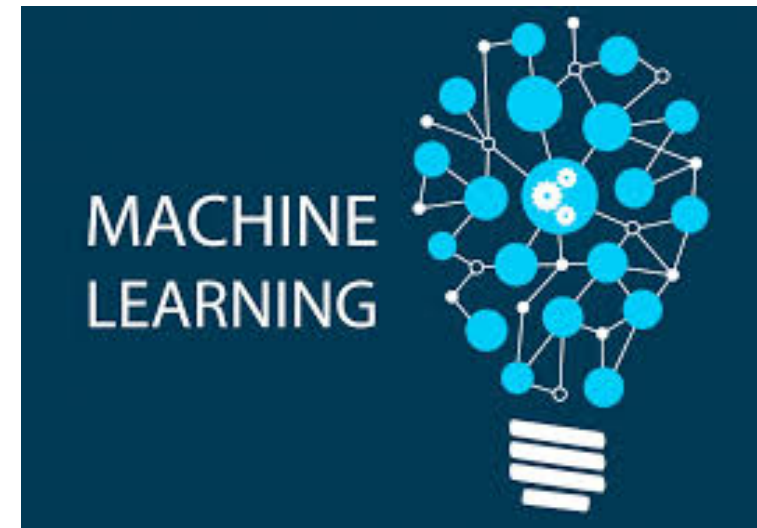
Machine learning is a component of AI !
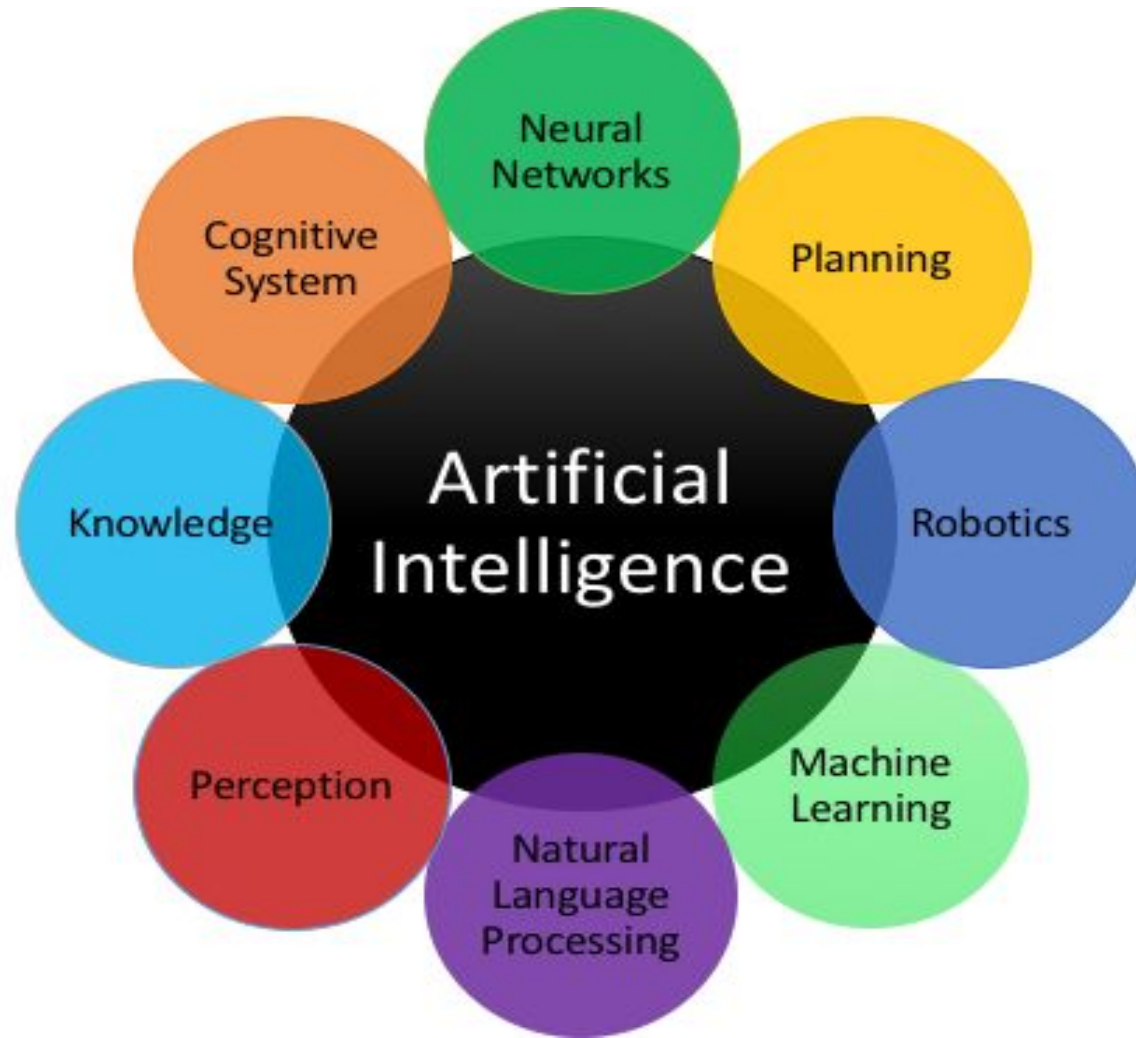
What does it actually do or why do we need ML or AI?

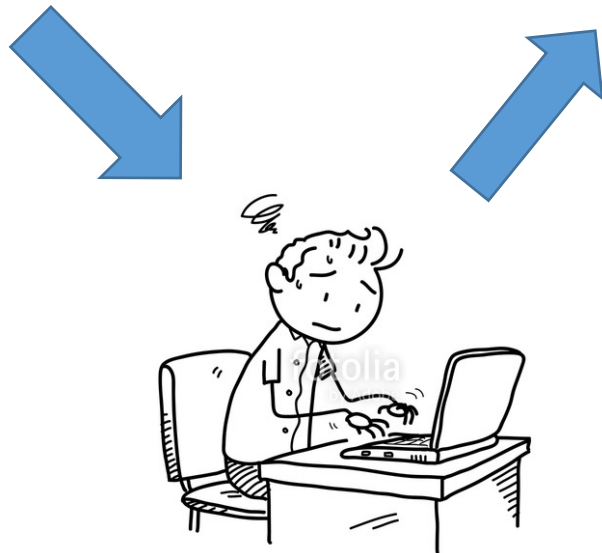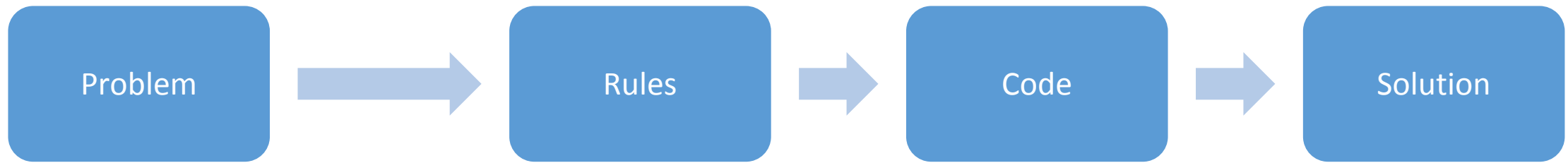**We want to** **automate** **certain processes using a computer.**

But we've been automating tasks since the invention of the computer. This raises the next question.

? Why is the classical way of programming a computer not sufficient? **?**

The Classical Way of Programming Machines

| Problem	→	Rules	→	Code	→	Solution

A developer or product manager analyzes the problem and figures out the rules/algorithm. This rules or Algorithm are then translated into programming code that can be run on a computer.

Examples are: Apps, Games, Online stores, Websites, Accounting Softrwares e.t.c

#151734702

Where The Classical Approach Fails

There is a set of problems where we are not able to identify and *explicitly* write down these rules. This concerns problems like image and speech recognition, natural language processing, or identifying complex patterns in data.
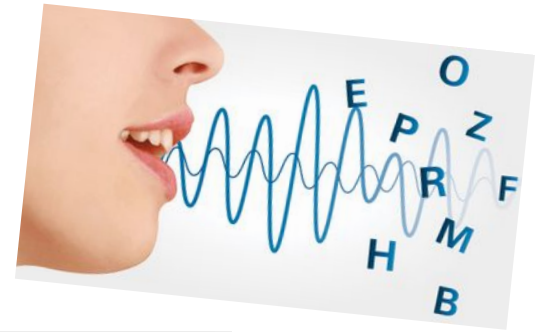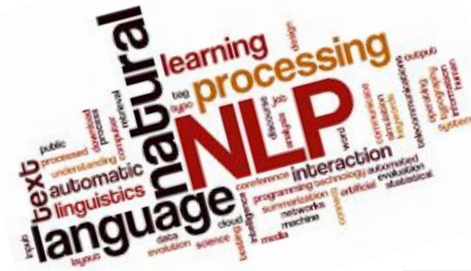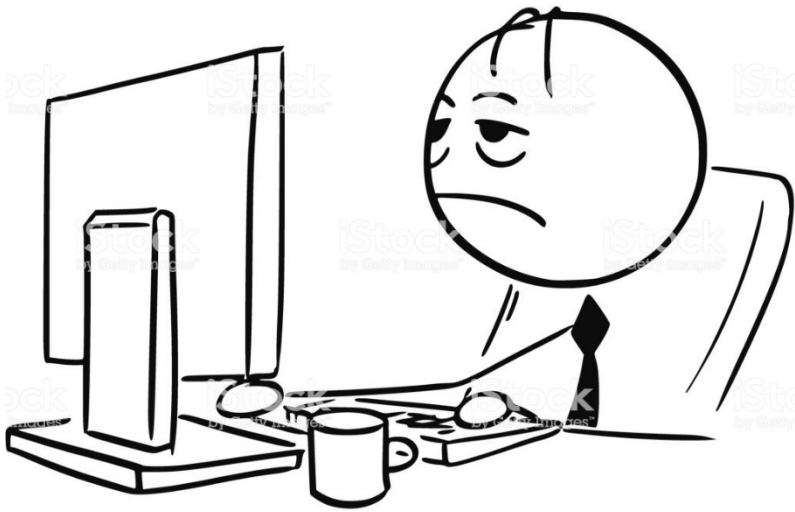
Why Machine Learning?

Problem	→	Rules	→	Code	→	Solution
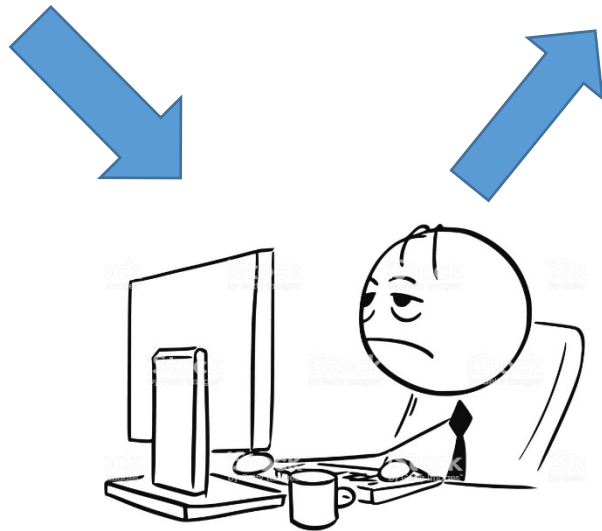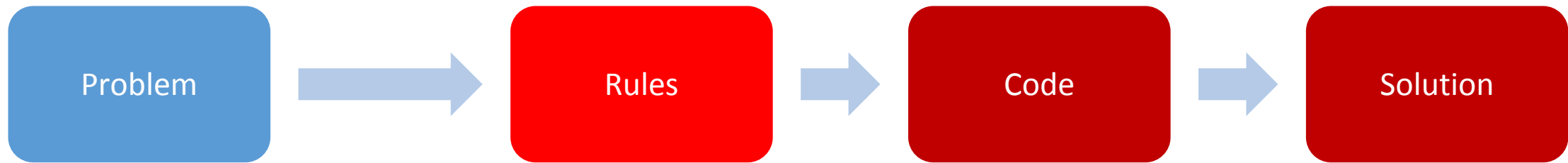
We are able to identify objects with our eyes, talk to and understand other people, read and write text, understand languages. e.t.c

But we don't know how our brain does it, or what rules it follows.

Nature is too complex, too ambiguous, varies a lot, hence cannot be described by a finite set of explicit rules.

RE-DEFINING MACHINE LEARNING

- Machine learning is basically an **automated** approach, where a machine (computer) analyzes data (images, audio, texts, etc.) for certain patterns.

- The difference is that the machine figures out the **rules** on its own, i.e which patterns to look for (usually by analyzing many examples).

- The machine learns the rules which would be impossible for a programmer to **explicitly** write down.
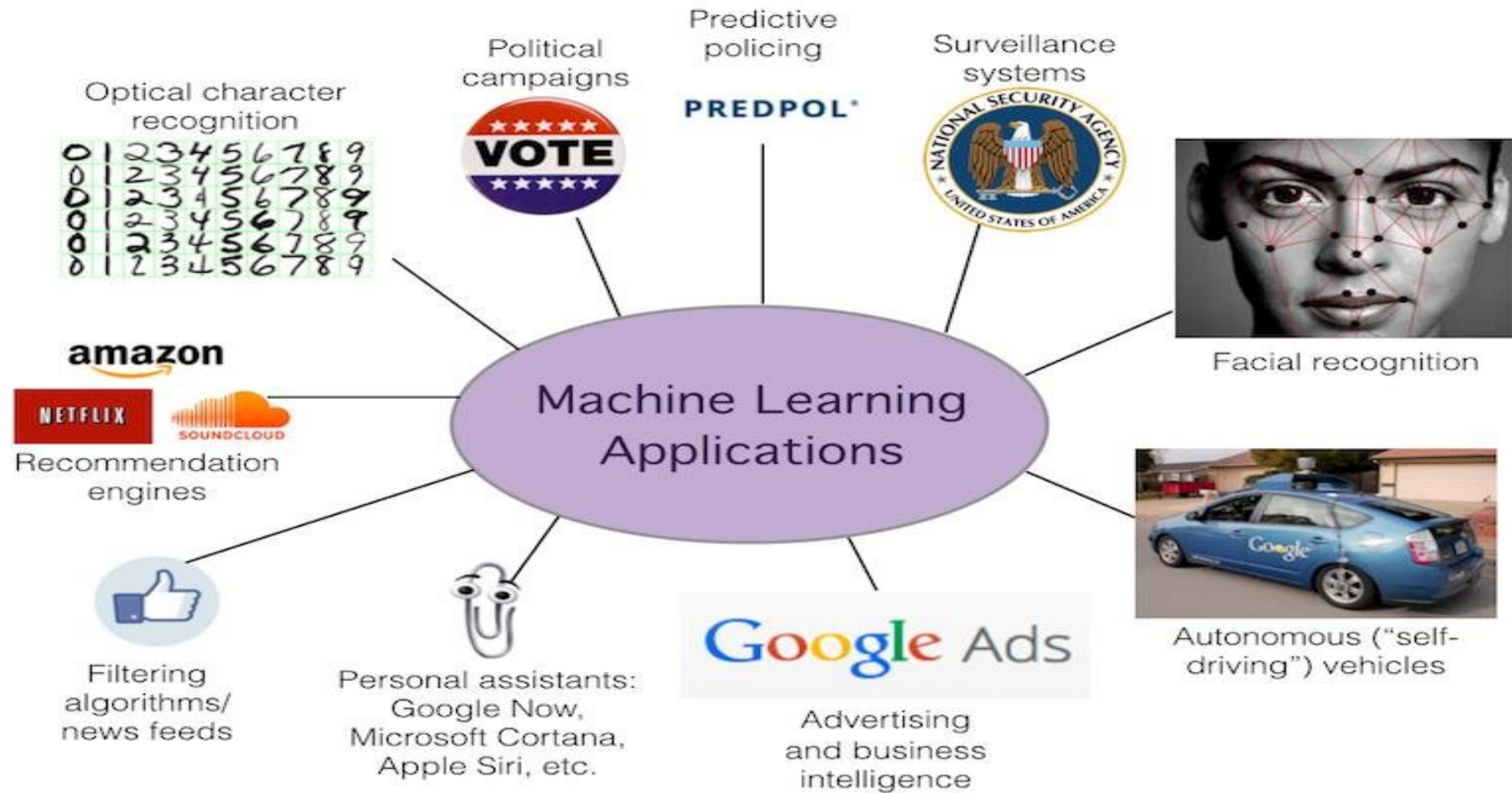
Where should you use Machine Learning?

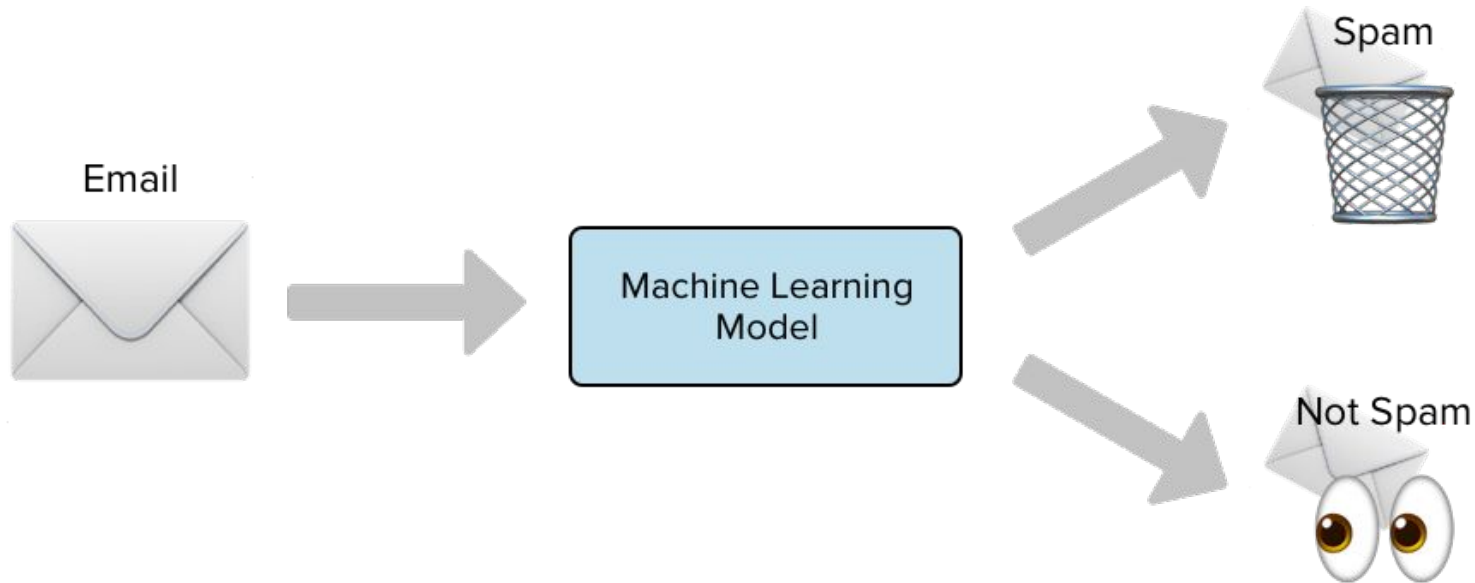Not all task requires ML!

Only Use ML when:

- When classical approach fails.
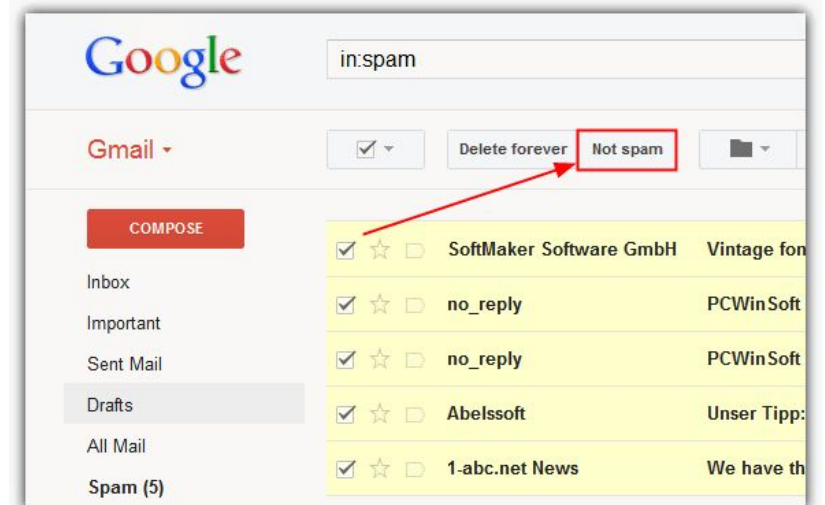- When classical approach works but is computational intractable.
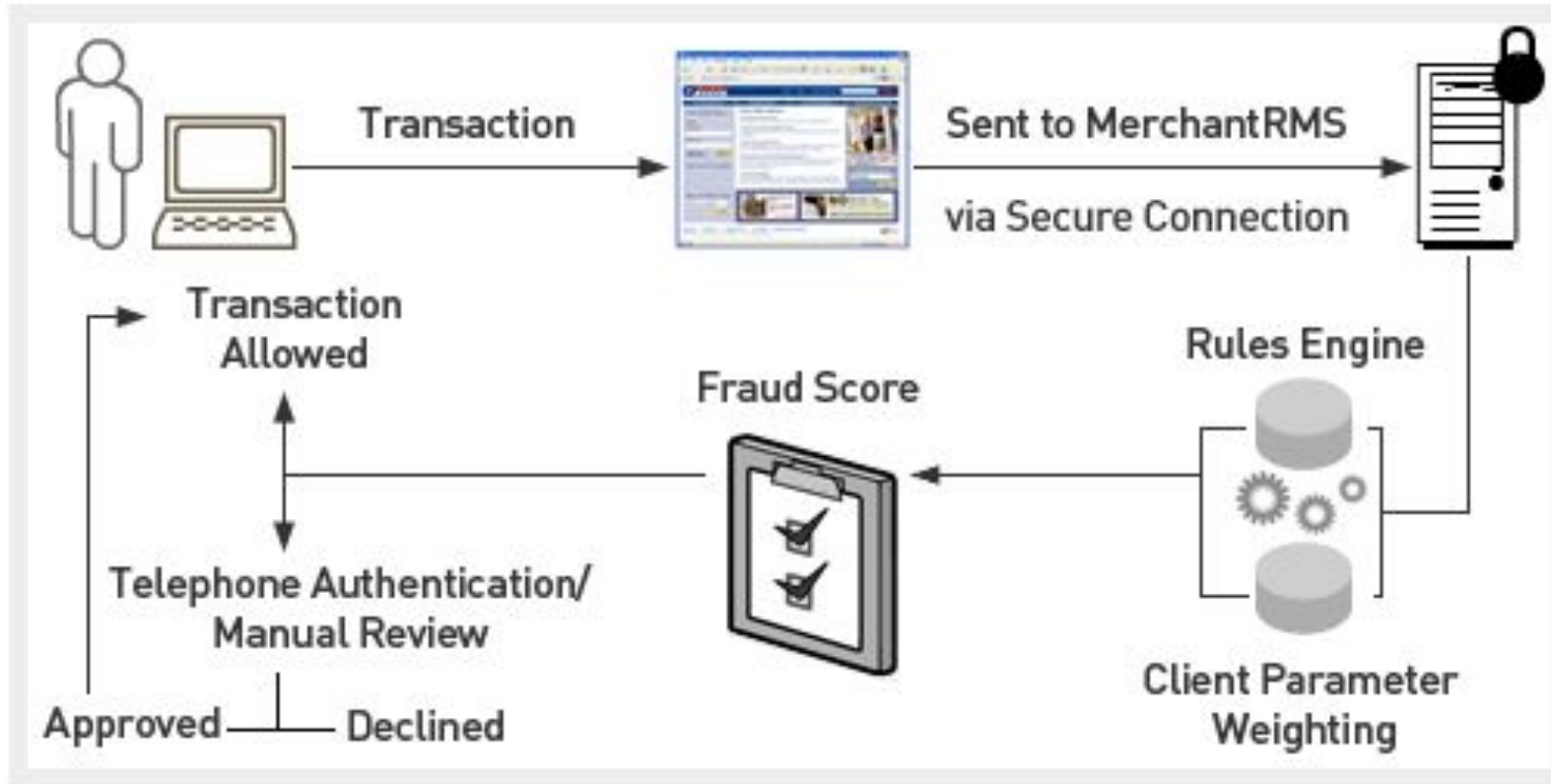
PRODUCTS THAT USE ML

Spam Detection_products that uses ML

When Google Mail detects your spam mail automatically, it is as a result of applying machine learning techniques.

Credit Card Fraud_products that uses ML

Used by Payment Institutions to detect unusual or anomalous activities regarding customer card transactions.
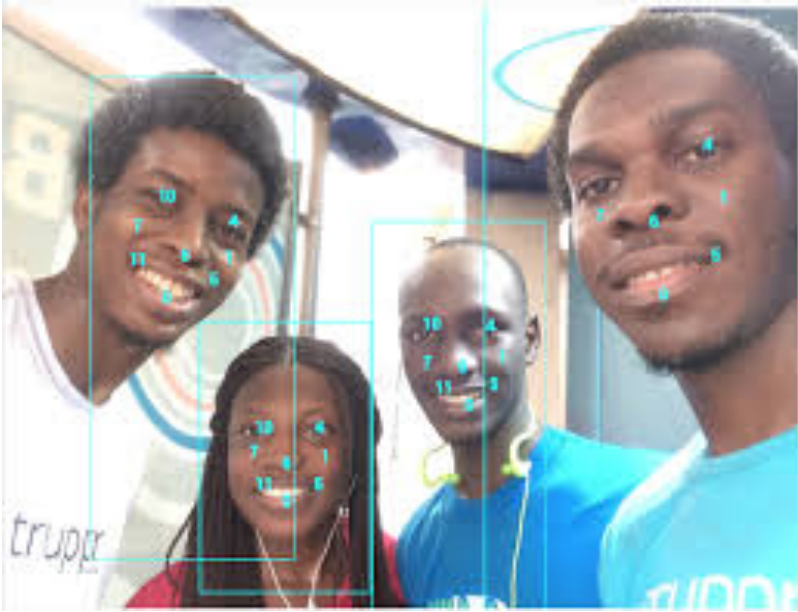
# Product Recommendation_products that uses ML

When Netflix suggest movies to you, when Amazon suggest you books, or when YouTube recommend videos to you. It's not magic, its machine learning.
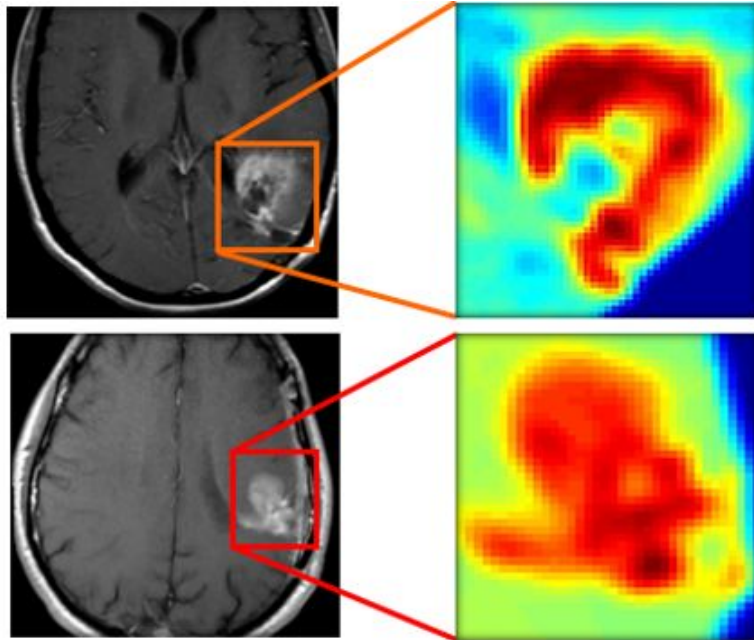
Face Detection_products that uses ML

When Facebook automatically recognizes faces of your friends in a photo, a machine learning process is what's running in the background.

THINKSTOCK/CNN ILLUSTRATION
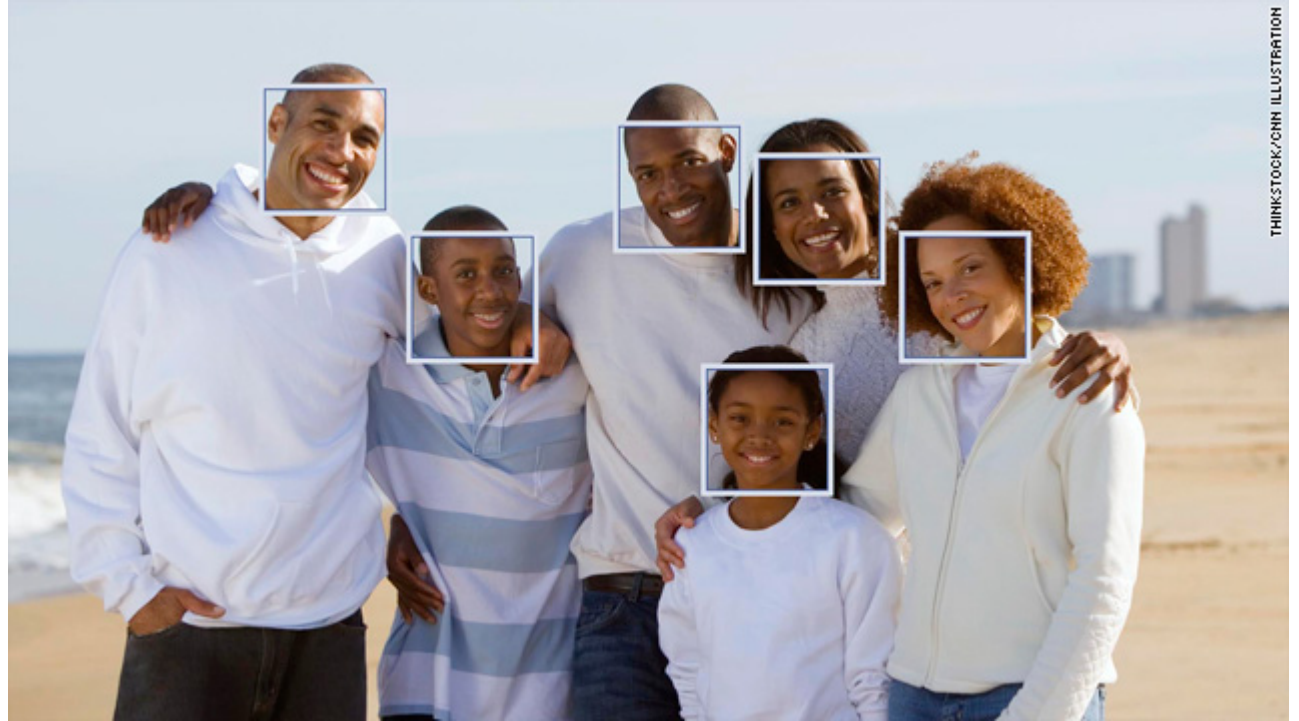
Medical Treatment/Diagnosis_products that uses ML

- IBM Watson Genomics
- Babylon Health for Analysis
- Careskore
- Google Deepmind Health for fast-track medical information
- Splice Machine & QIAGEN Clinical Insight (QCI) for precision medicine
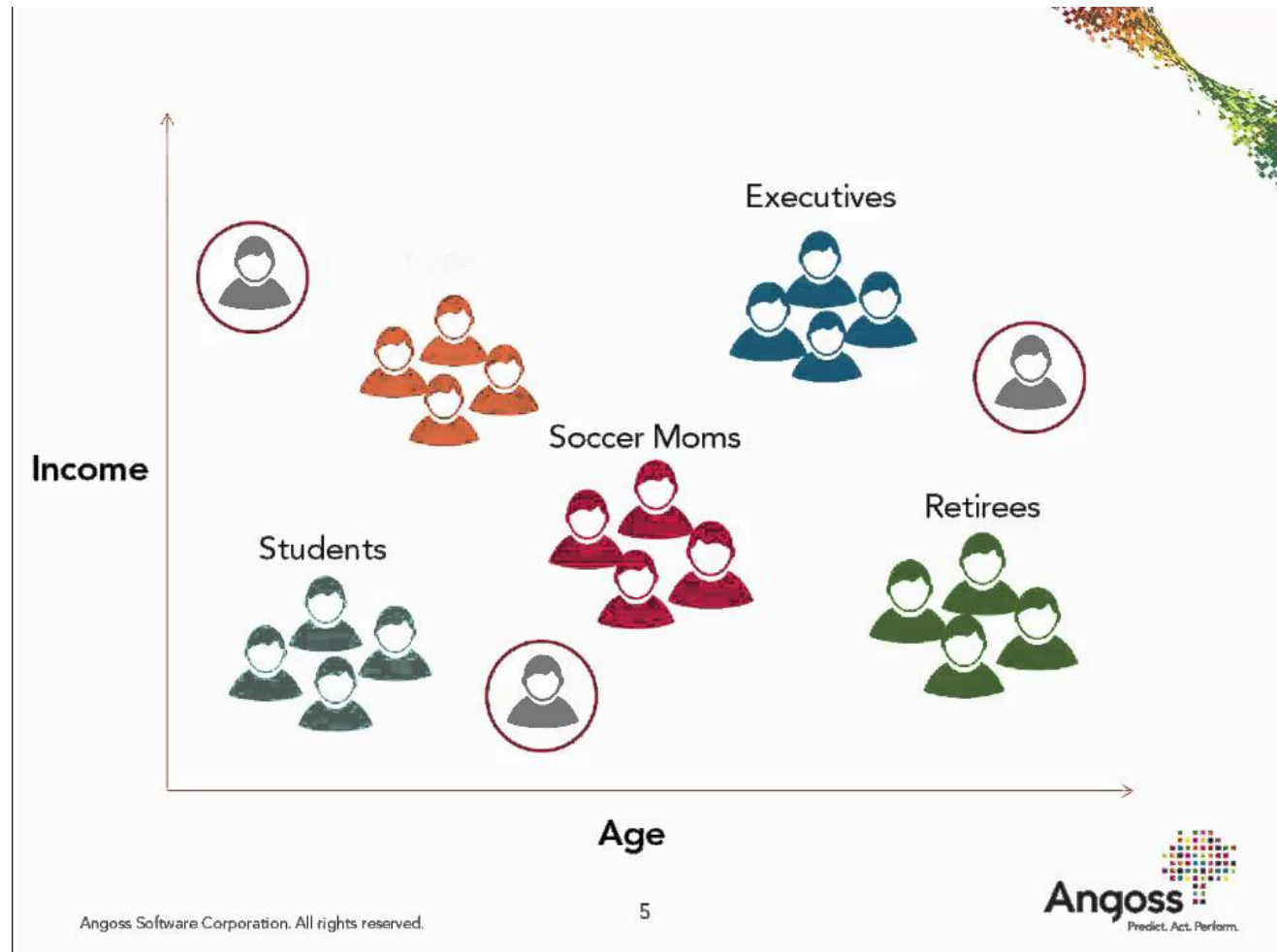- Zephyr Health for Treatment design e.t.c

Using a database of symptoms and treatments of patients, a popular machine learning problem is to predict if a patient has a particular illness.

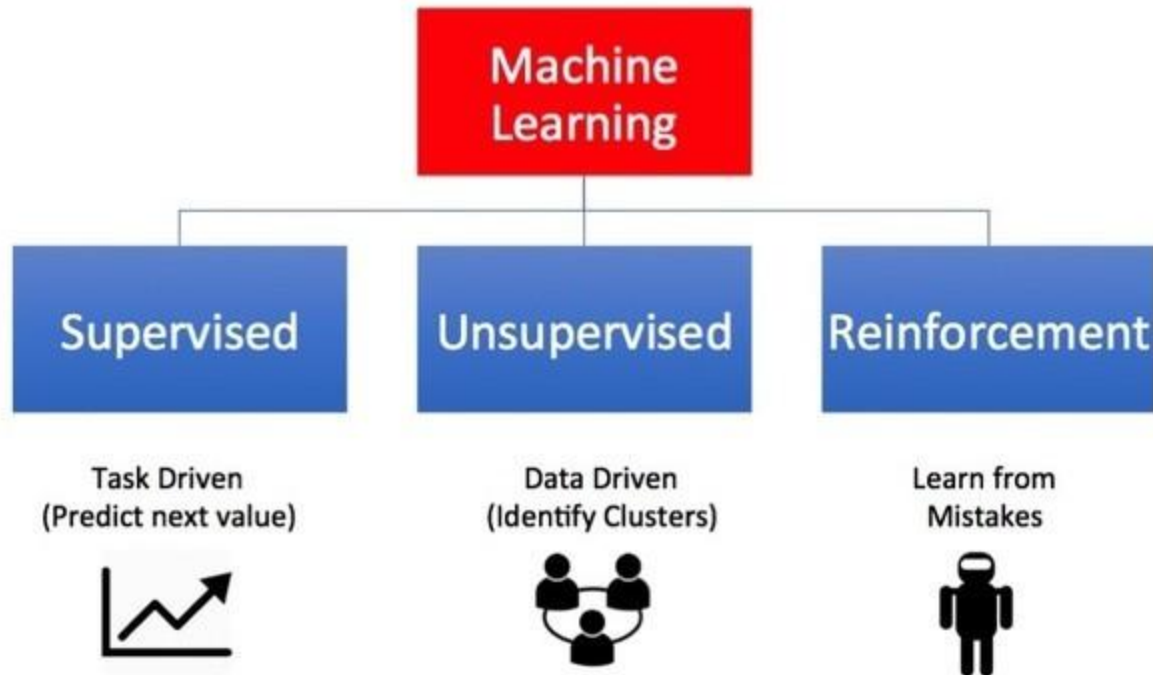# Customer Segmentation_products that uses ML

Using Income and Age data we could cluster/ segment customer into the following categories.

This is called Segmentation.

TYPES OF MACHINE LEARNING

- Supervised Learning: classification and regressions problems.

- Unsupervised Learning : Clustering problems

- Reinforcement Learning

Types of Machine Learning

Machine Learning

| Supervised | Unsupervised | Reinforcement |

Task Driven (Predict next value) — Data Driven (Identify Clusters) — Learn from Mistakes

Supervised Learning
- Labeled data
- Direct feedback
- Predict outcome/future

Unsupervised Learning
- No labels
- No feedback
- Find hidden structure in data

Reinforcement Learning
- Decision process
- Reward system
- Learn series of actions
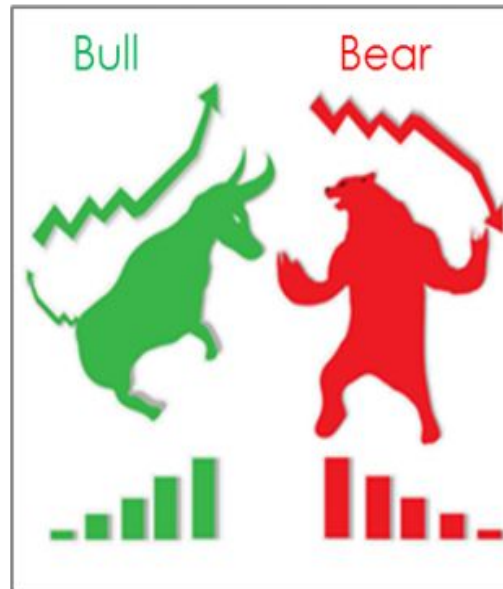
SUPERVISED LEARNING _TYPES OF MACHINE LEARNING

- It consists of Classification and Regression Problems.
- The target variable is known.

Regression   vs   Classification

Prediction vs Real Return

Bull   Bear
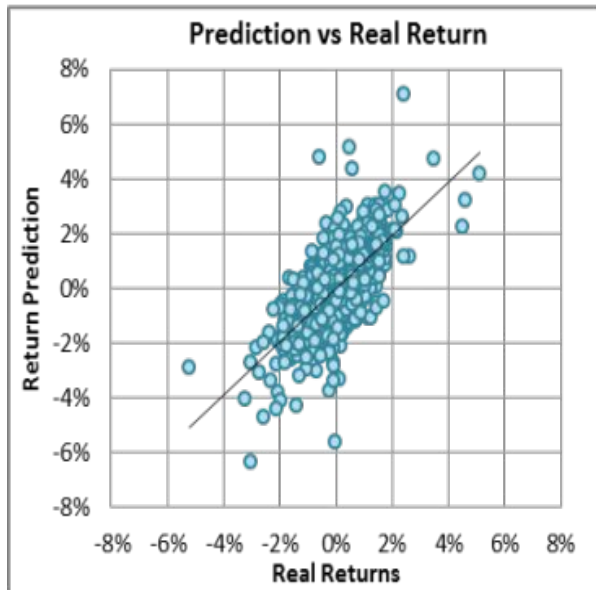
Classification vs Regression

- Classification means to group the output into a class.
- classification to **predict** the type of tumor i.e. harmful or not harmful using training data
- if it is discrete/categorical variable, then it is classification problem

- Regression means to predict the output value using training data.
- regression to **predict** the house price from training data
- if it is a real number/continuous, then it is regression problem.

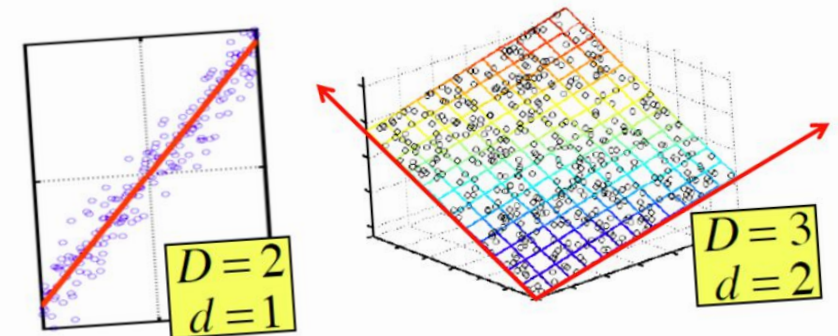# UNSUPERVISED LEARNING _TYPES OF MACHINE LEARNING

- It consists of Clustering and Dimension Reduction Problems.
- The target variable is unknown.
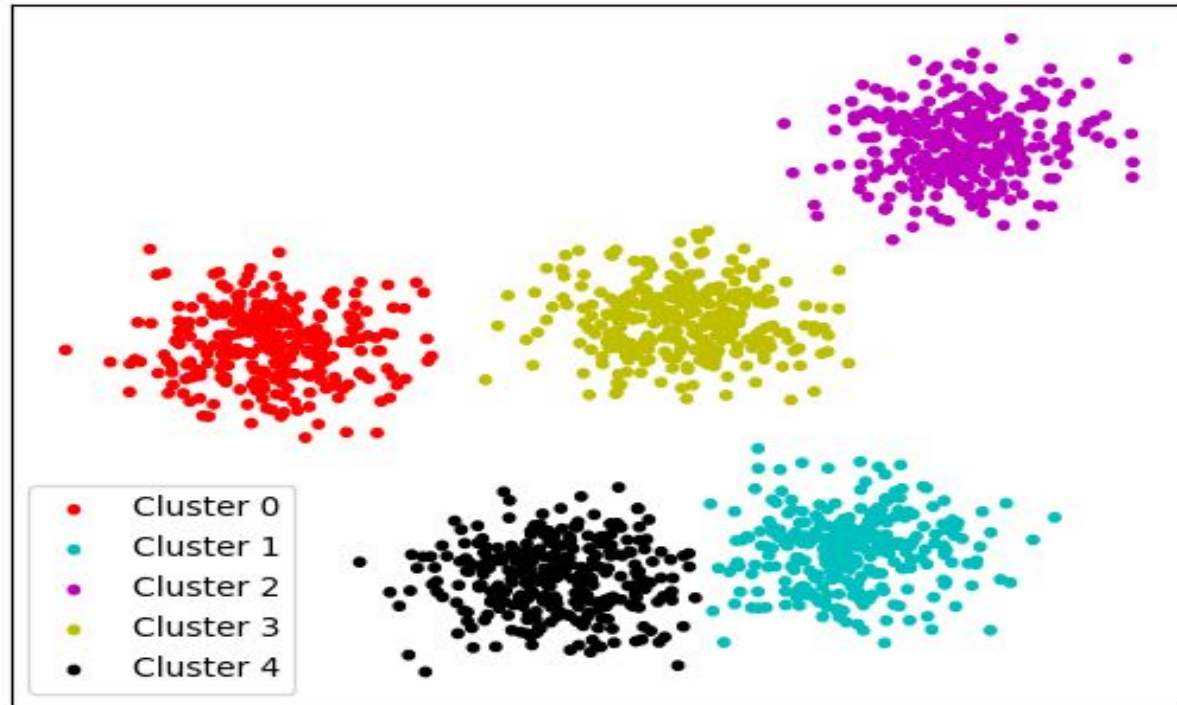
Clustering

Dimensionality reduction
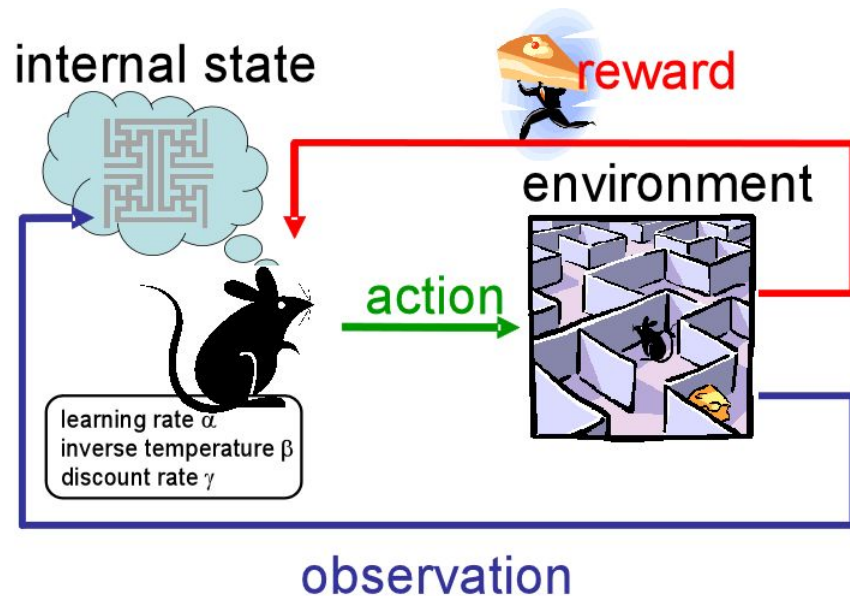
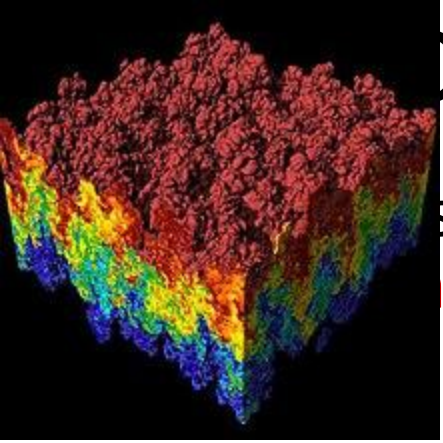# REINFORCEMENT LEARNING _TYPES OF MACHINE LEARNING

- Reinforcement learning refers to goal-oriented algorithms, which learn how to attain a complex objective (goal) or maximize along a particular dimension over many steps.

RL-type Problems

- game of chess, GO, Space Invaders
- balancing a unicycle
- investing in stock market
- running a business
- making fast food
- life…!

DATA ANALYTICS/ANALYSIS

...s is a process of inspecting, cleansing, transforming, ...**g** data with the goal of discovering useful information, ...nclusions, and supporting decision-making.

Techniques of Data Analysis:

1. Data Mining (Predictive Data Analysis)
2. Business Intelligence (Actionable Business Information)
3. Data Visualization (Representation of Data)

Data

Useful knowledge (patterns)
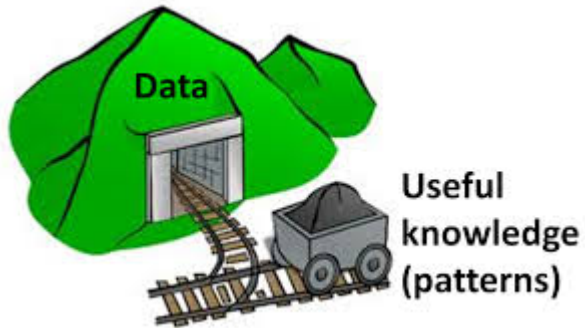
Data Analysis uses **MACHINE LEARNING** for **Modelling!**

Typical Process of Data Analytics

Decision Making Problem

Requirement Understanding → Data Understanding → Data Preparation → Data Exploration → Insight Development → Decision Making

Modeling & Evaluation → Deployment

★ The most time-consuming part

★ The most exciting part

Modeling Problem
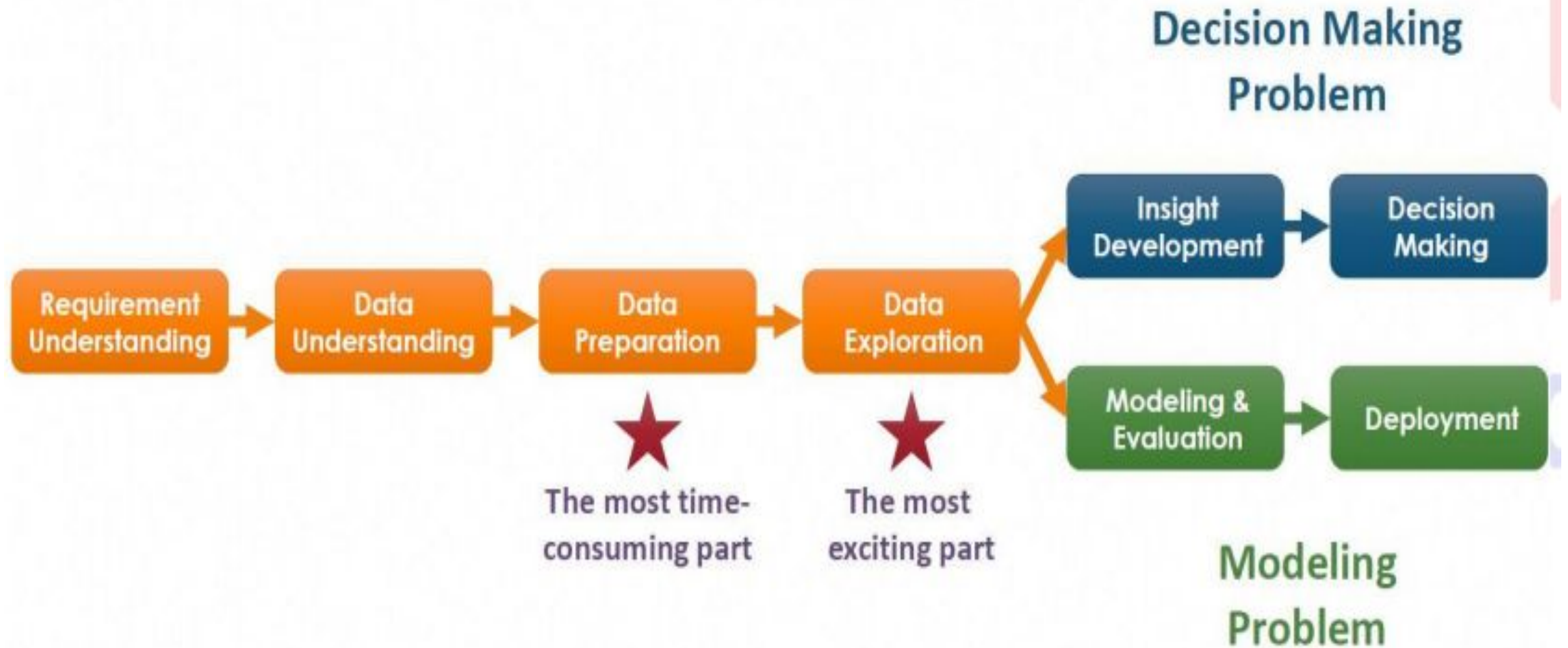
TYPES OF DATA ANALYTICS

Data Analytics

| Descriptive Analytics	Predictive Analytics	Prescriptive Analytics
What has happened or is happening?	What could happen in the future?	What should we do to make that happen or not happen?
"How has the population been changing?"	"How will the population change over the next ten years?"	"What actions should be taken in order to avoid the demographic cliff?"

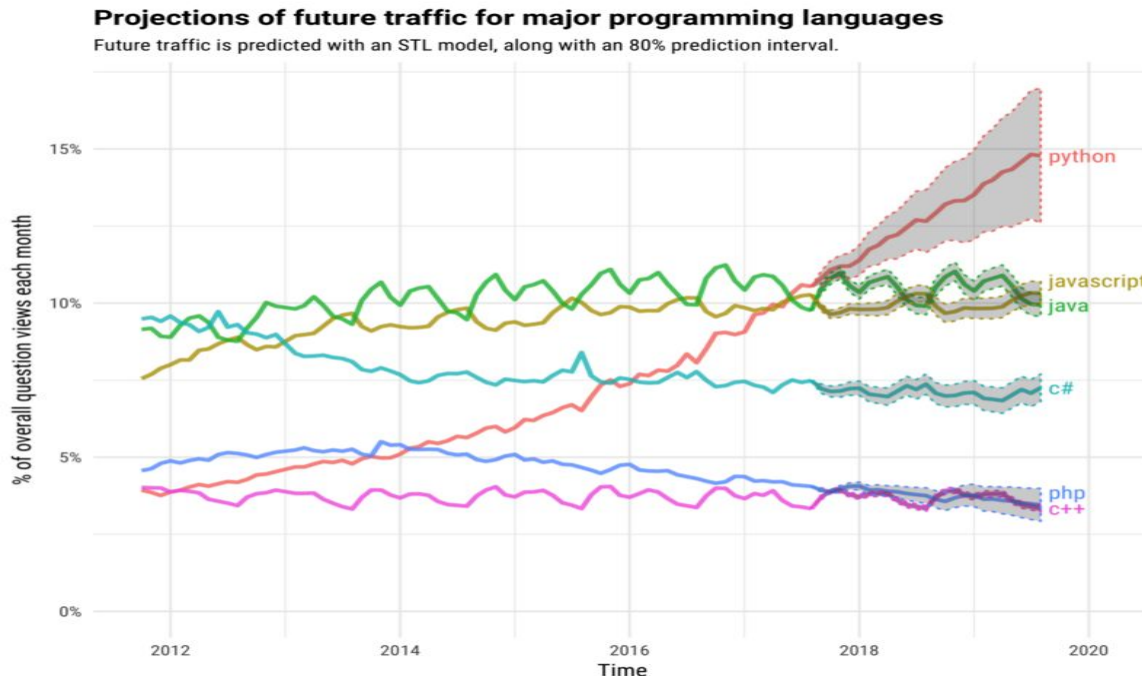END OF INTRO TO MACHINE LEARNING

NEXT UP—INTRO TO PYTHON FOR ML/DS

WHY PYTHON

- Python was built for readability and has les complexities.
- Python is easy to learn.
- Lots of packages already built for Machine Learning.

  - Want to work with images — numpy, opencv, scikit
  - Want to work in text — nltk, numpy, scikit
  - Want to work in audio — librosa
  - Want to solve machine learning problem — pandas, scikit
  - Want to see the data clearly — matplotlib, seaborn, scikit
  - Want to use deep learning — tensorflow, pytorch, theano
  - Want to do scientific computing — scipy
  - Want to integrate web applications — Django
  - Want to take a fly …. Import antigravity
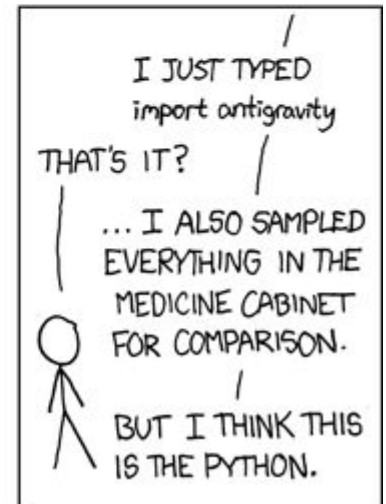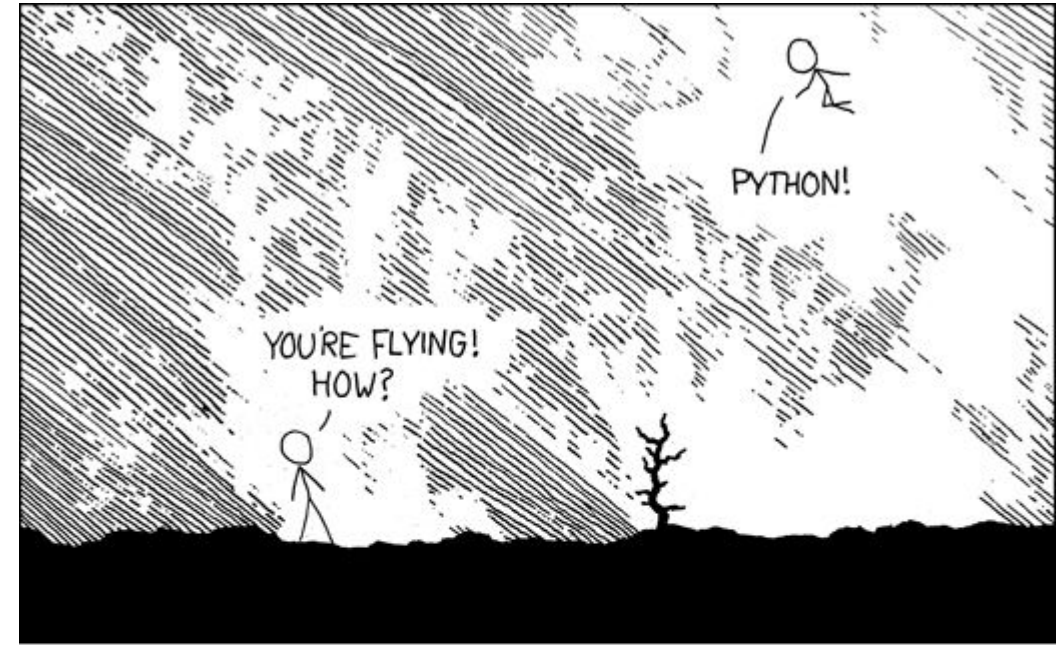
WHY PYTHON_contd

- Shorter syntax.
- Uses whitespaces instead of braces.
- Future predictions show increased demand and use.

Projections of future traffic for major programming languages
Future traffic is predicted with an STL model, along with an 80% prediction interval.

"Hello, World"

- C

```
#include <stdio.h>

int main(int argc, char ** argv)
{
    printf("Hello, World!\n");
}
```

- Java

```
public class Hello
{
    public static void main(String argv[])
    {
        System.out.println("Hello, World!");
    }
}
```

- now in Python

```
print "Hello, World!"
```

Monday, June 14, 2010

Source: stackoverflow developer survey

INTRO TO JUPYTER NOTEBOOK AND JUPYTER LAB

END OF DAY 1
Thank you for Listening