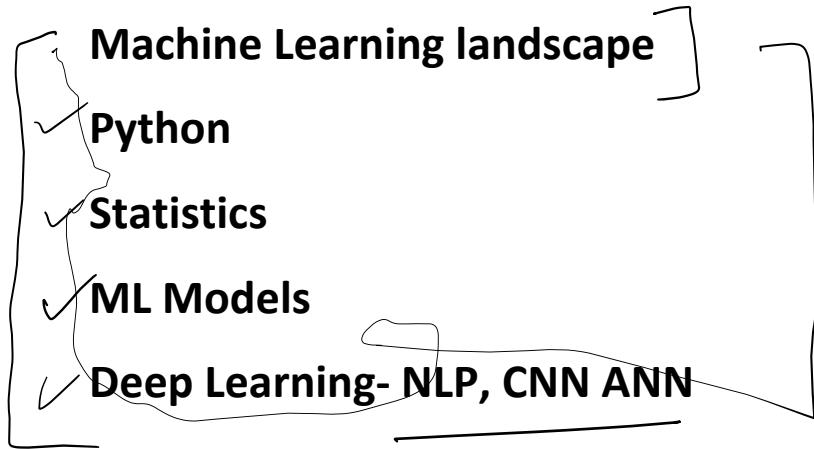


## Artificial Intelligence Feb 2024 Batch

32 hrs



## Data Science- Complete Package

**SQL, PySpark, Power BI, Cloud Computing**

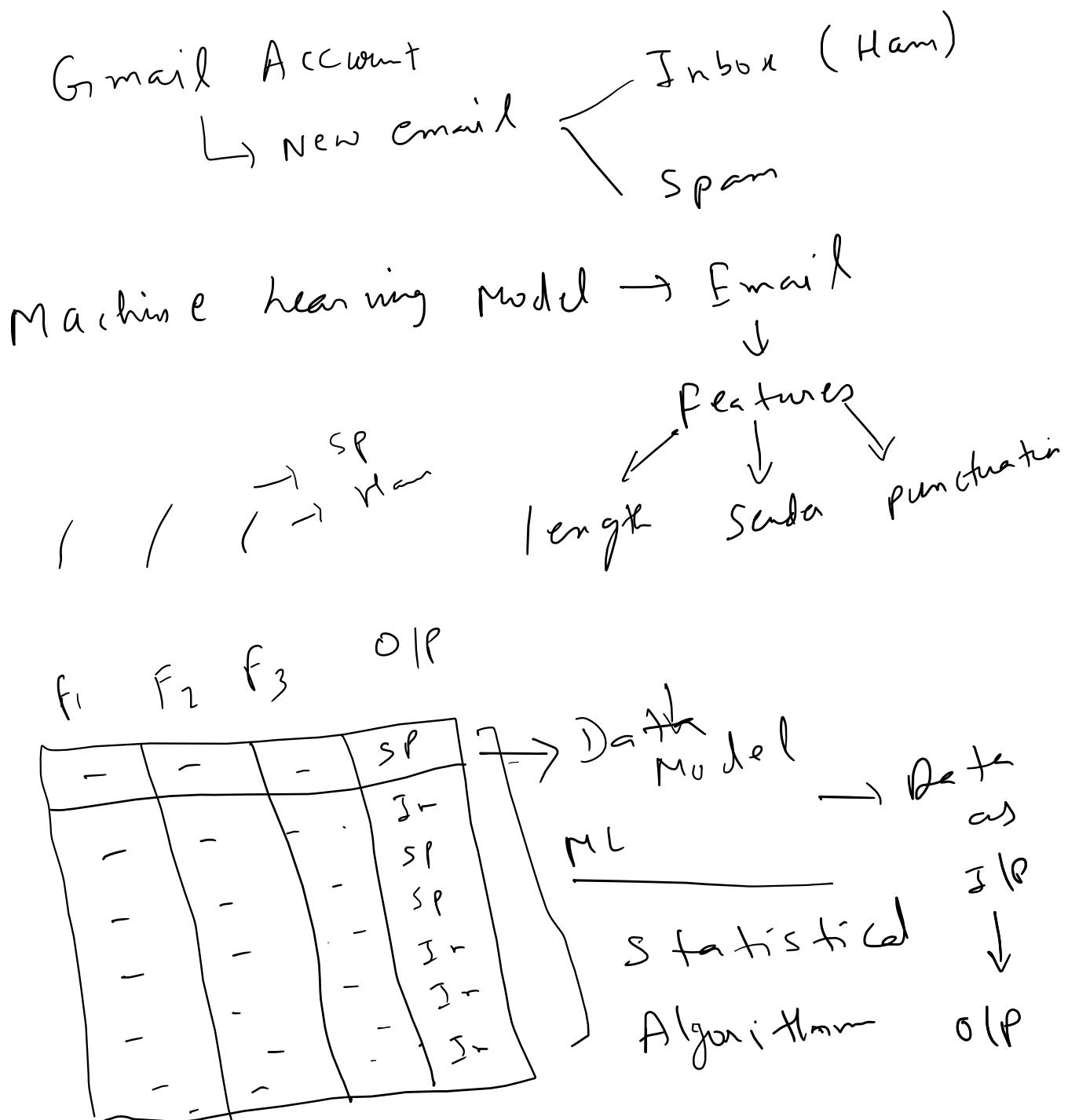
✓      ↴  
Python

- ✓
- MSFT + Azure
  - AWS + Amazon
  - GCP → Google

①  
②  
③

**Machine Learning is the science (and art) of programming computers so they can learn from data.** ✓

**Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.**



A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

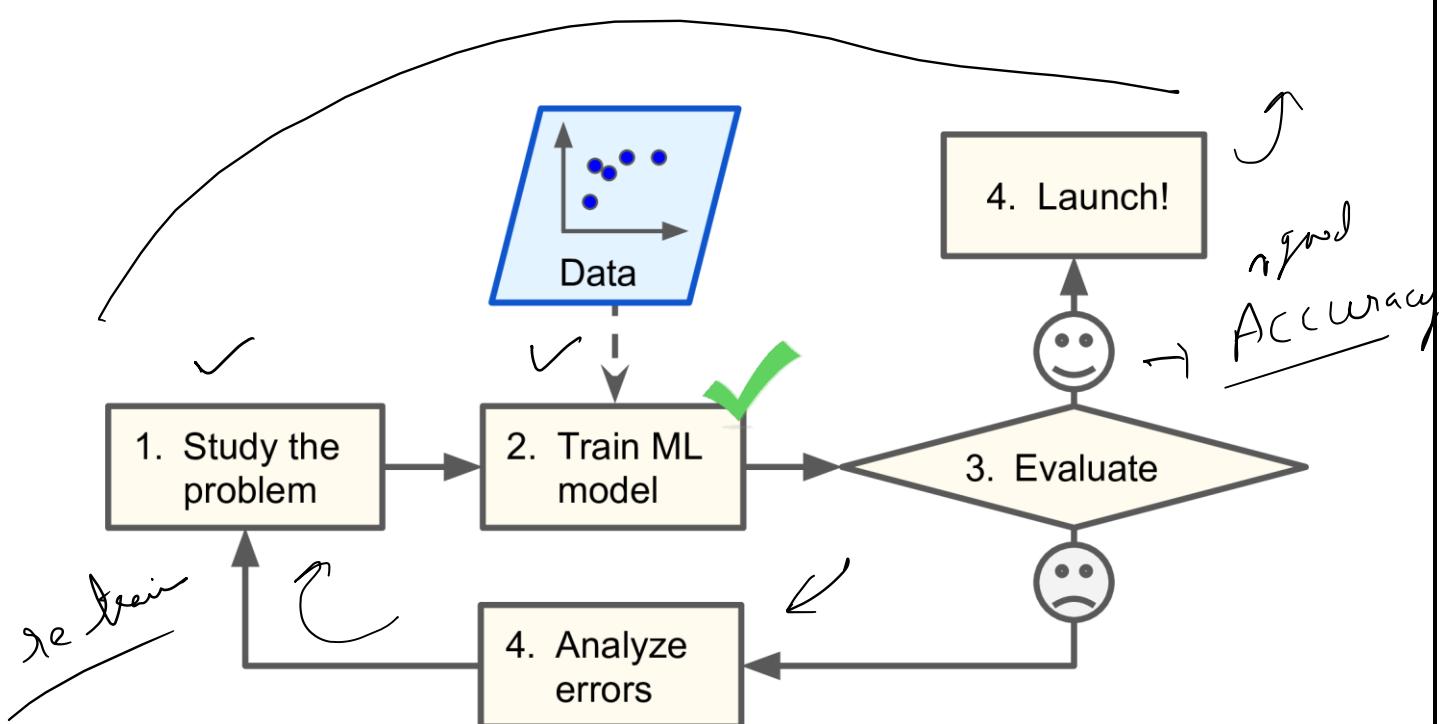
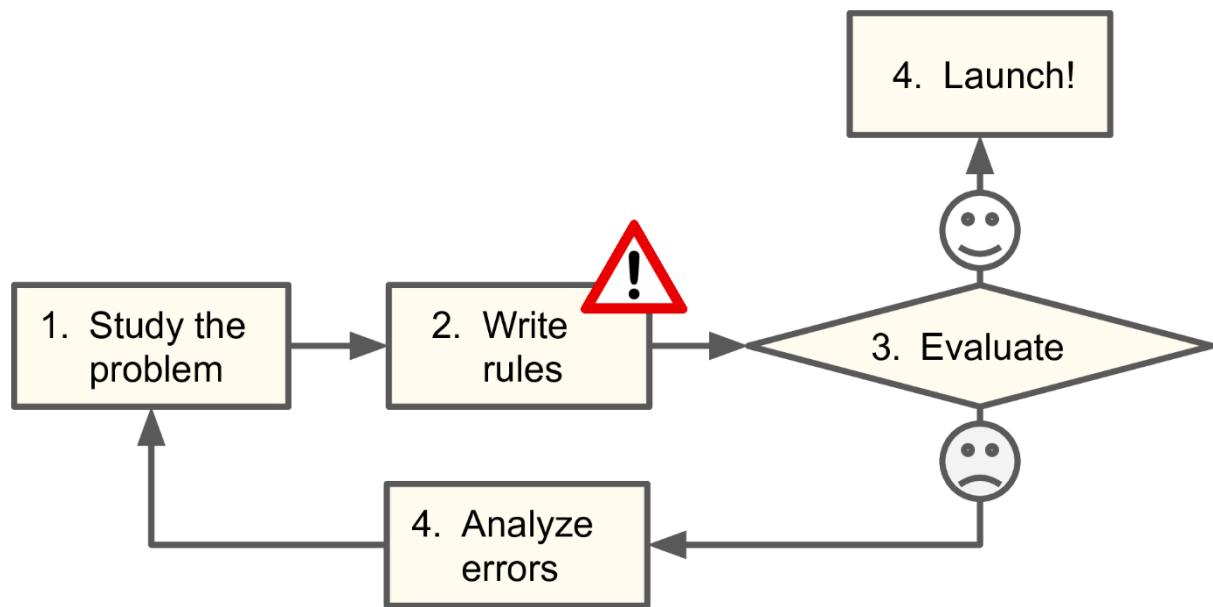
Experience E → training data  
Task T → classify email as spam/Ham  
Performance meas P → how accurately you are classifying

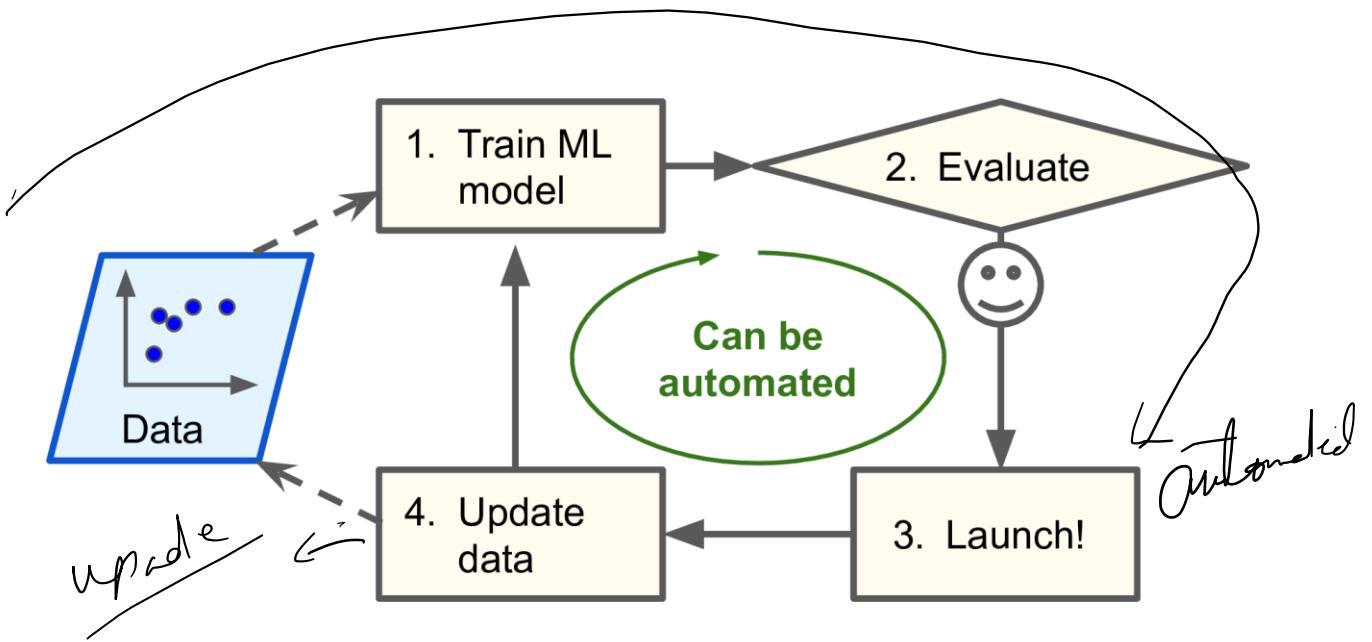
From the perspective of Spam/Ham example the experience E is the training data, the task T is to flag new/incoming email as spam/ham and the performance measure P is called as accuracy.

- ✓ Data Analyst- SQL, Python, Power BI or Tableau, Basic Statistics, Cloud Computing
- ✓ Data Scientist- SQL, Python, Power BI or Tableau, Advance Statistics, Machine Learning, Deep Learning, Cloud Computing
- ✓ ML Engineer- Production
- ✓ Data Engineer- ETL (Extract, Transform & Load) or ELT(Extract, Load& Transform)
- Cloud Computing- Microsoft Azure, GCP(Google Cloud Platform), AWS(Amazon Web-Services)

Model → Model Pipeline (Automated)  
↑ SF

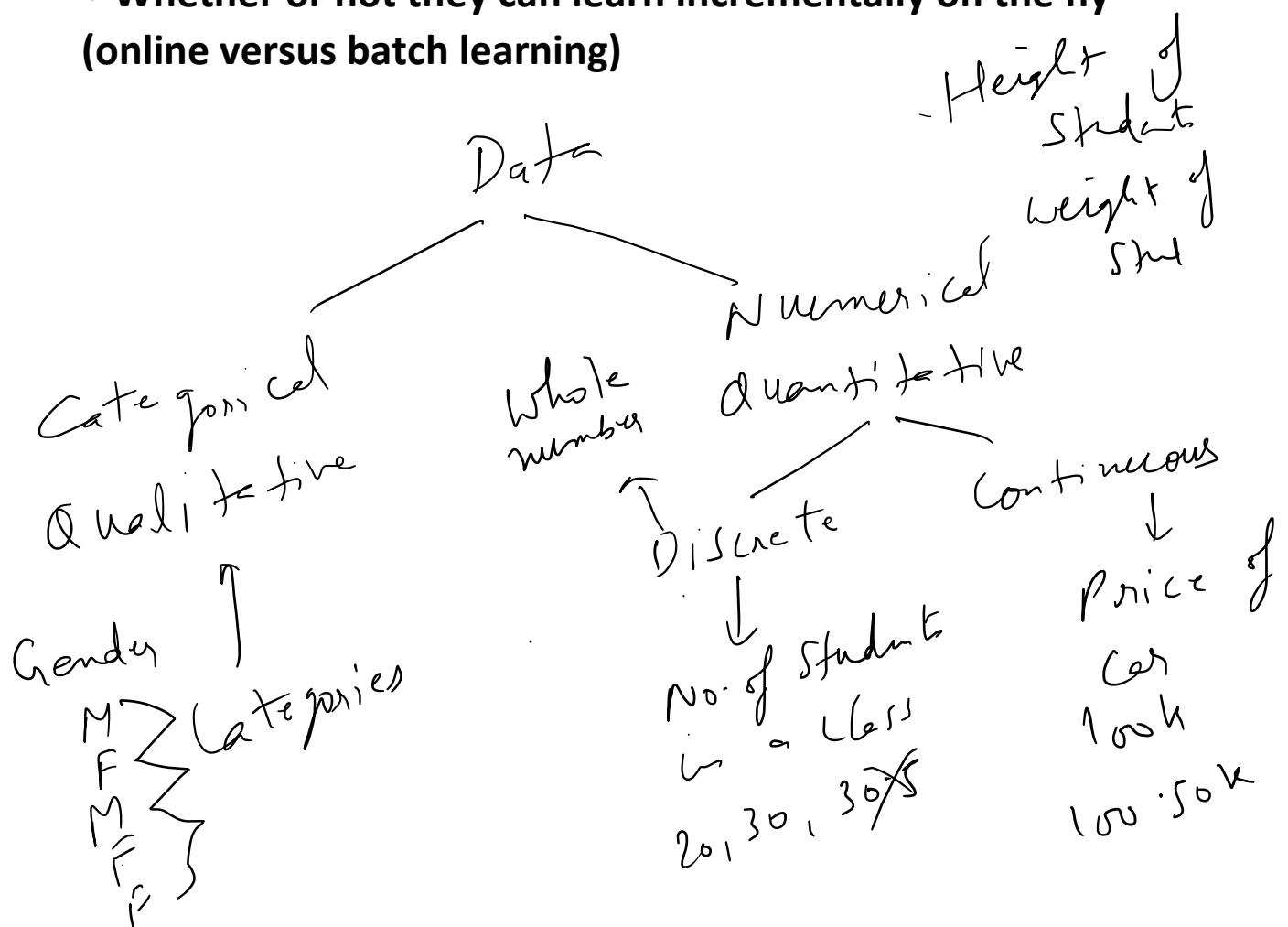
↓ processed Data → *Agree*  
↓ CSV files → DB → *Read*





## Types of Machine Learning Systems

- How they are supervised during training (supervised and unsupervised machine learning)
- Whether or not they can learn incrementally on the fly (online versus batch learning)

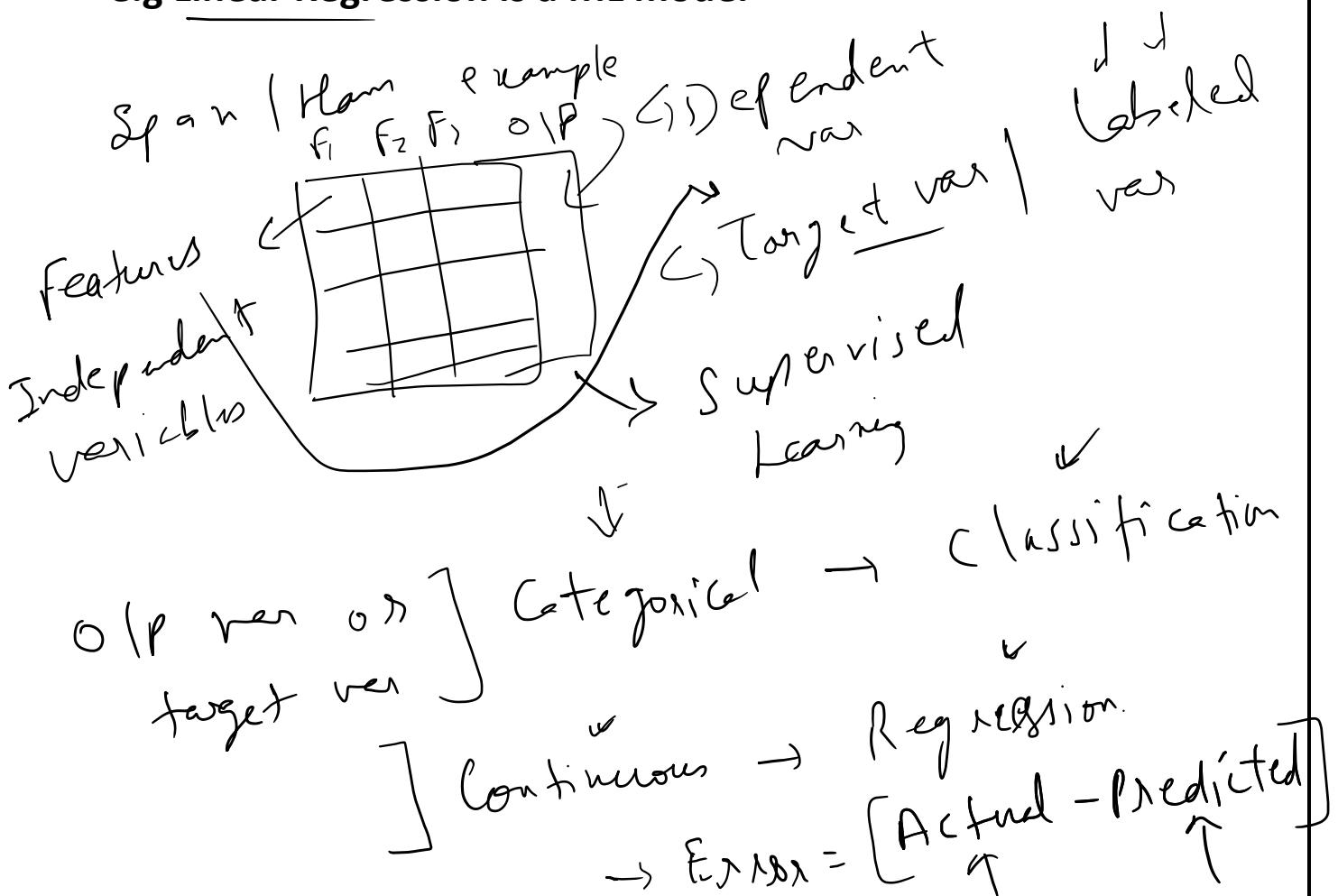


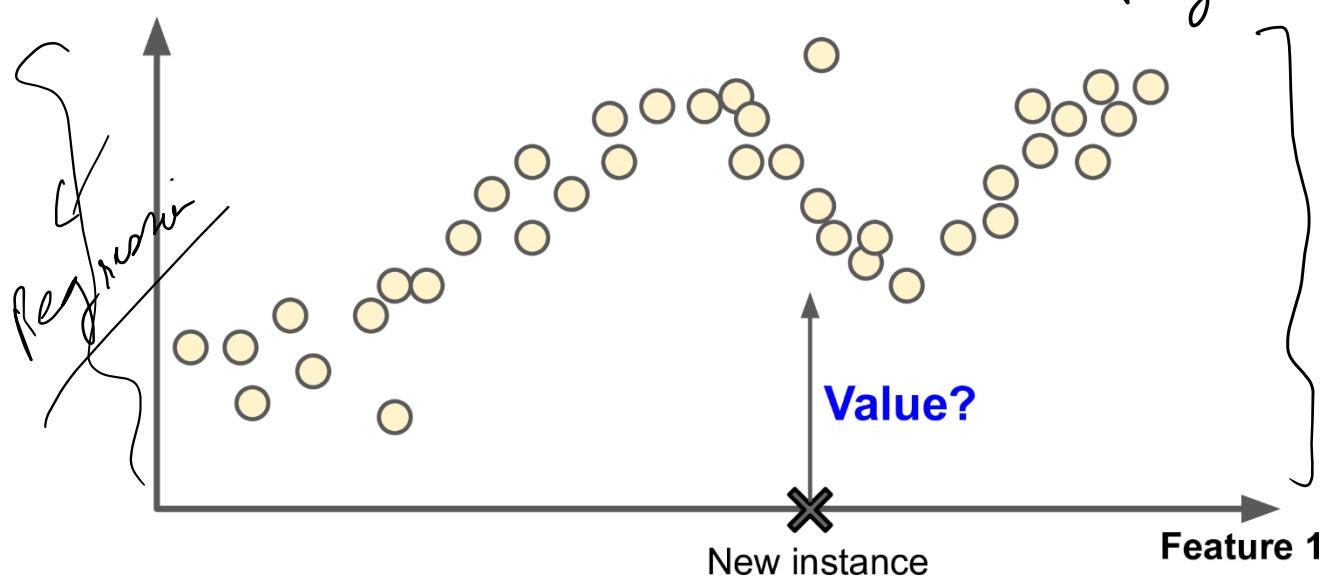
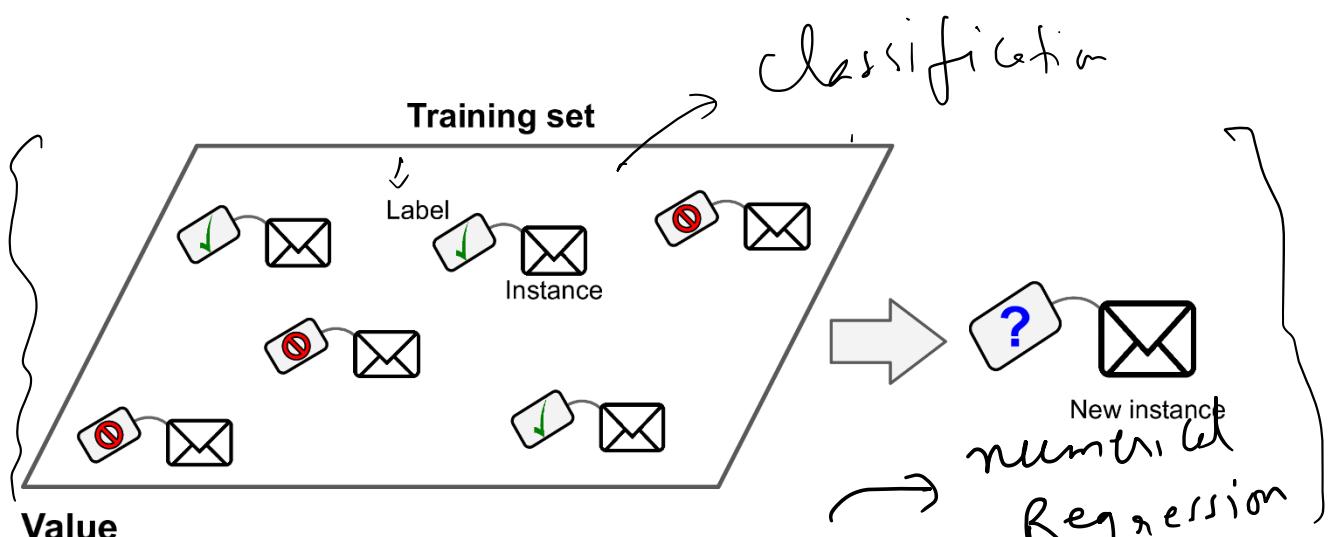
**Training Supervision-** ML systems can be classified according to the amount and type of supervision they get during training.

**Supervised Learning-** In supervised learning, the training set you feed to the algorithm includes the desired solutions called labels.

There are two kinds of supervised learning:

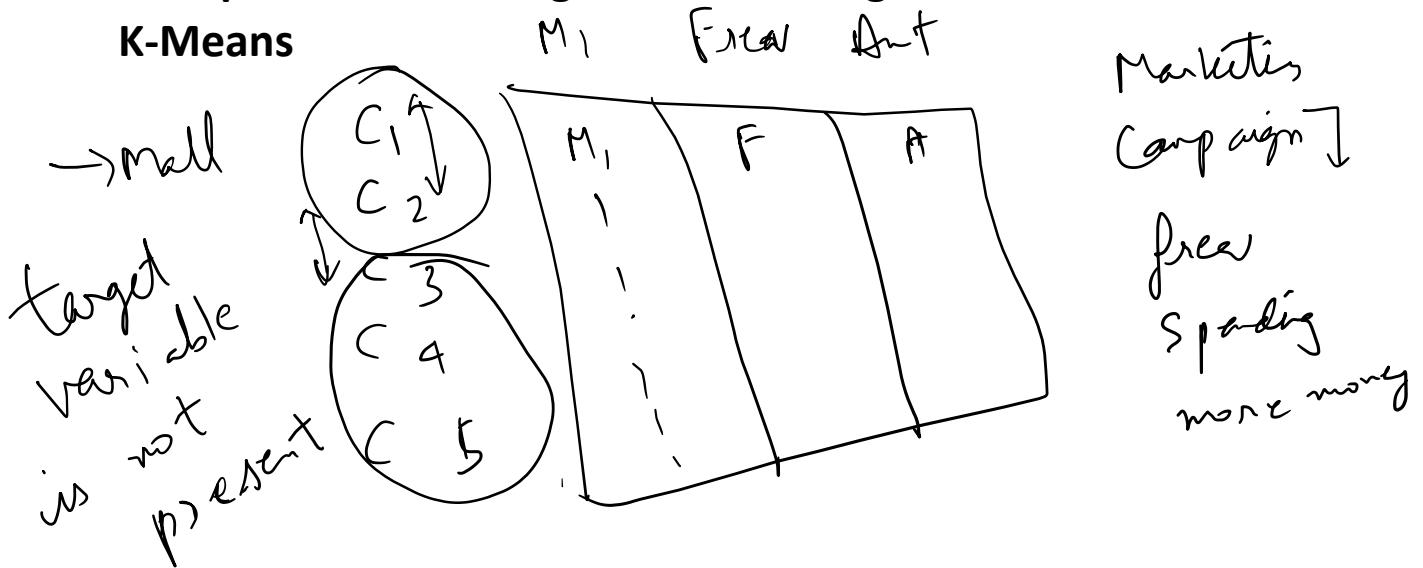
- ✓
  - 1. Classification- Spam/Ham model is a classification task where target variable is discrete (0 or 1) e.g. Logistic Regression is a ML model
  - 2. Regression- The target variable is continuous in nature e.g Linear Regression is a ML model



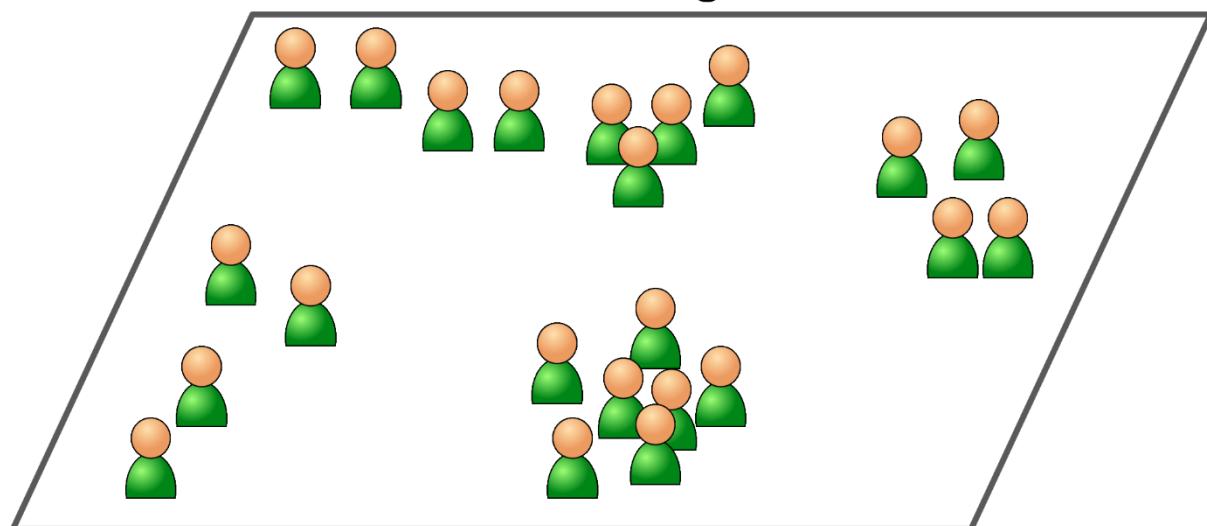


**Unsupervised Learning: In this the target data is unlabelled.**

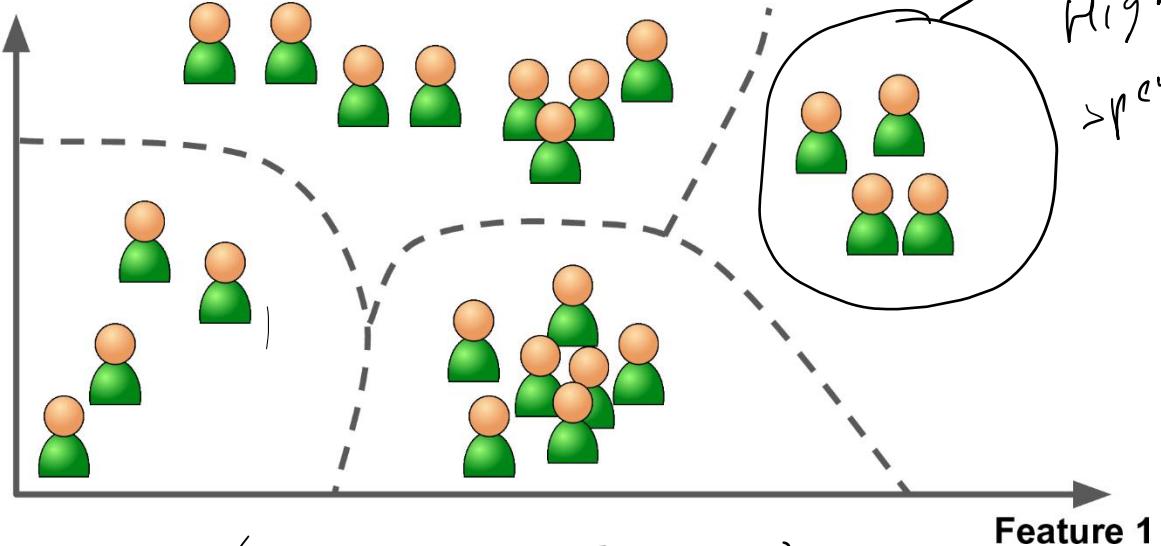
**K-Means**



**Training set**



Feature 2

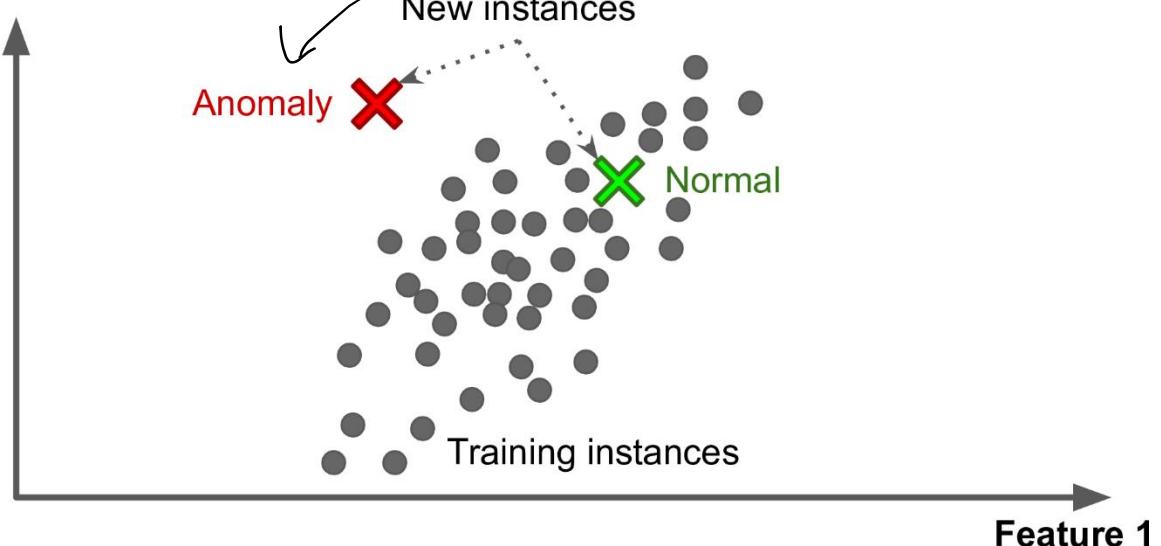


Feature 1

1000 → 50 customers

CC → 1 lac (10 - 15 k) 80k ← call blocked

Feature 2

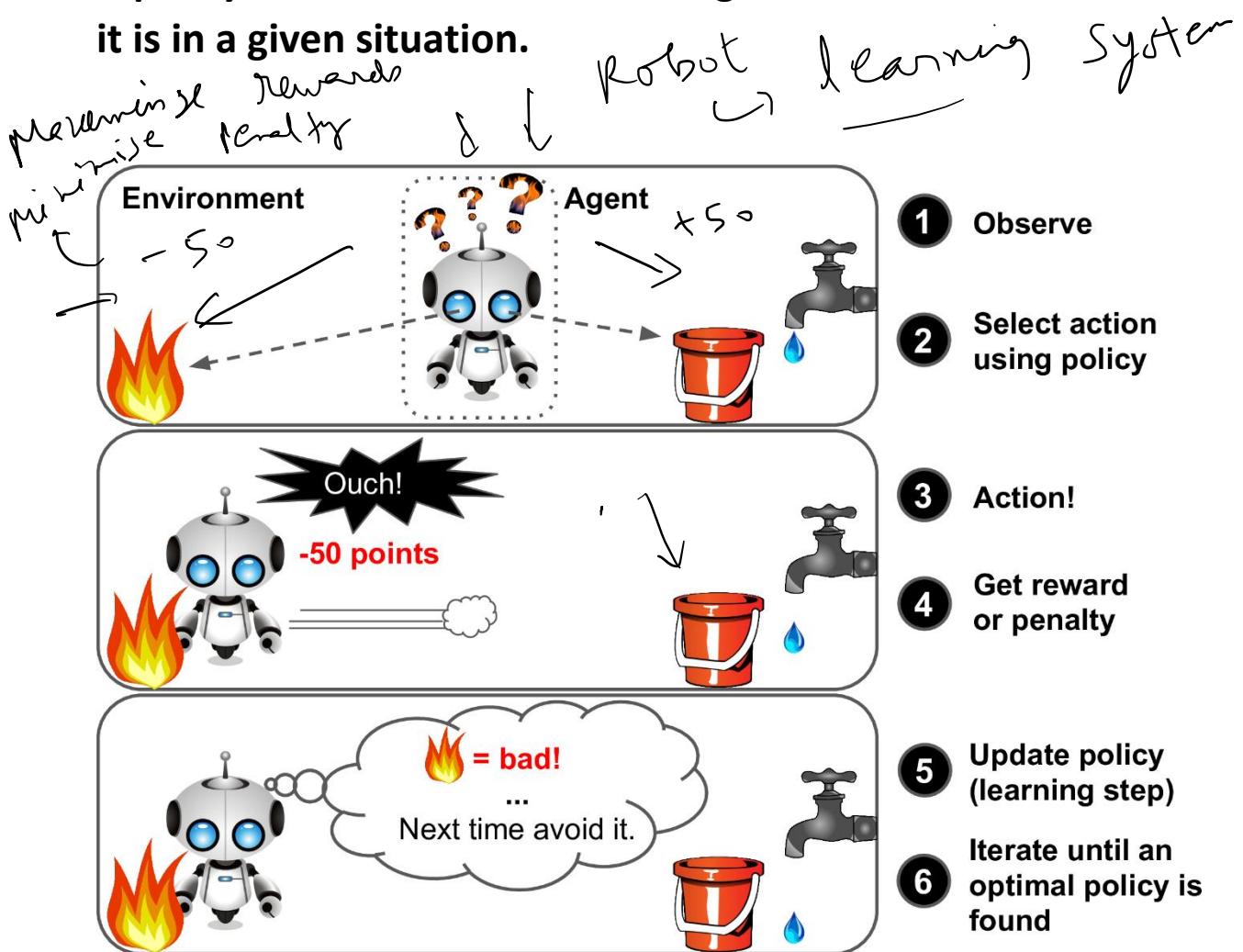


Feature 1

**Reinforcement Learning**- The learning system, called an agent in this context, can observe the environment, select and performs actions and get rewards in return (or penalties in the form of negative rewards).

It must then learn by itself what is the best strategy, called a policy to get the most rewards over time.

A policy defines what action the agent should choose when it is in a given situation.



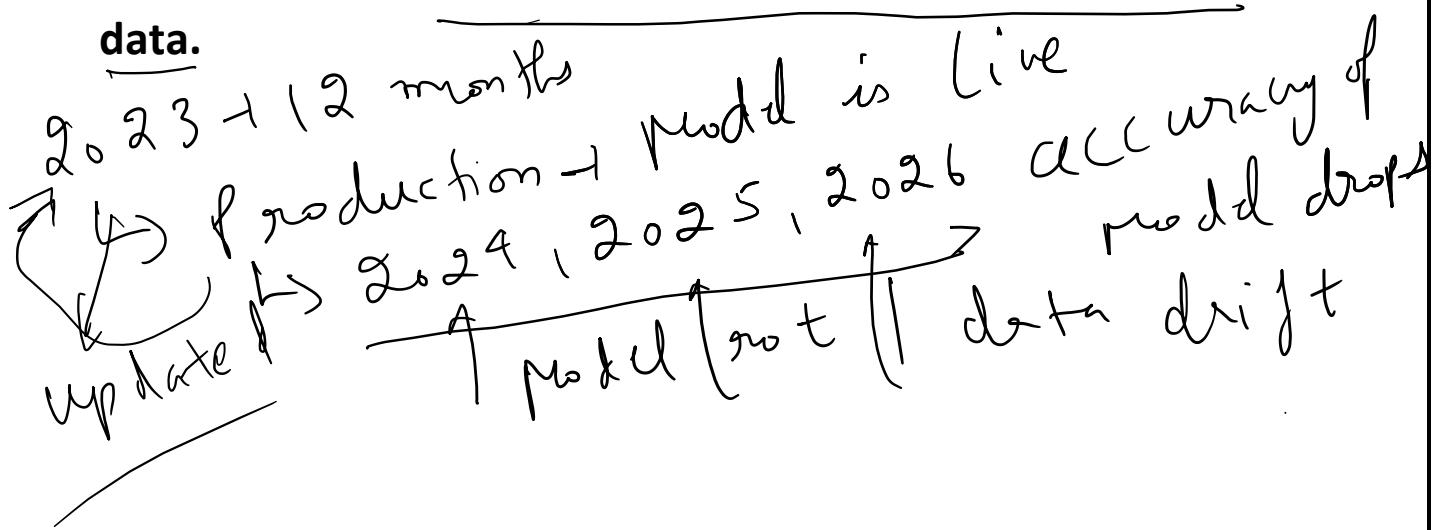
## Batch versus Online Learning

All data is used

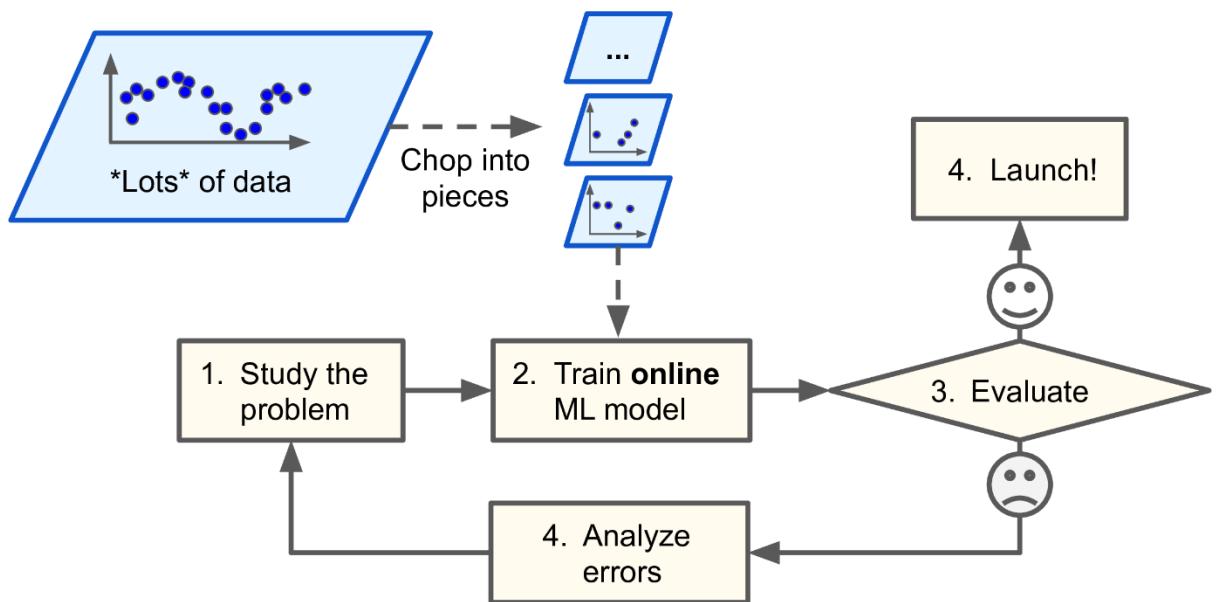
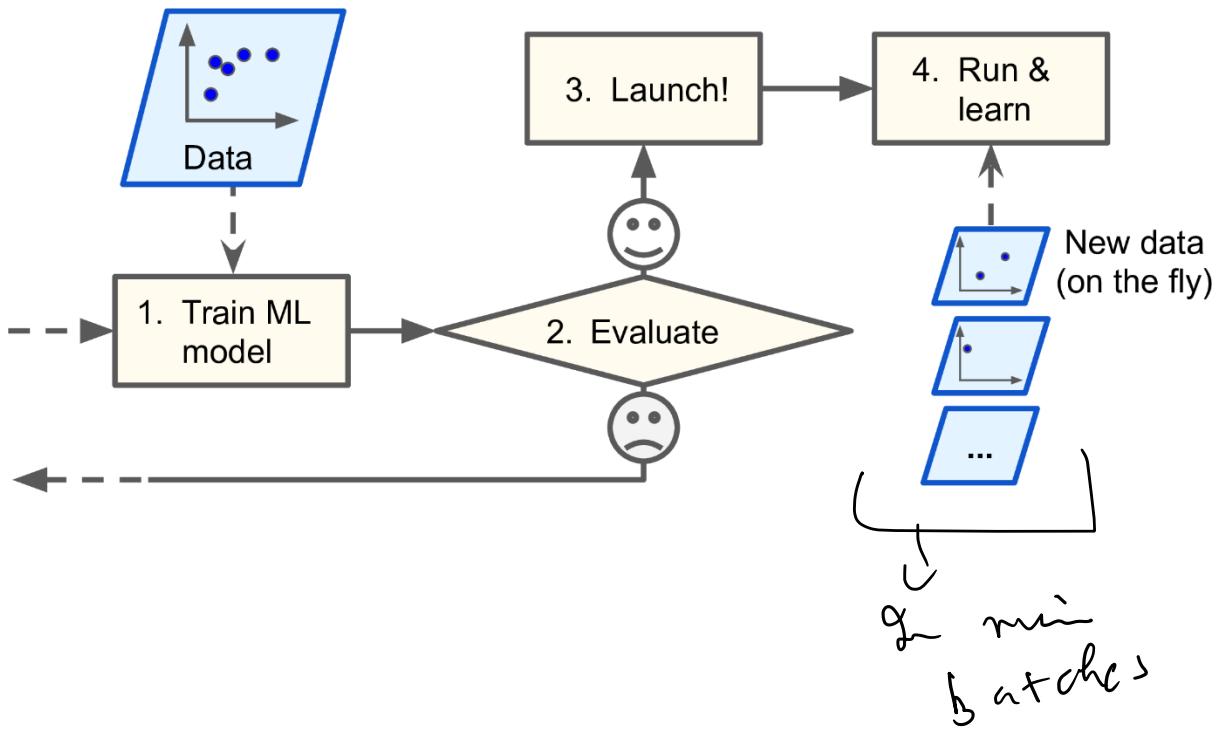
**Batch Learning:** In batch learning the system is incapable of learning incremental, it must be trained using all the available data. This is called as offline learning. A model's performance tends to decay slowly over time, simply because the data changes with time.

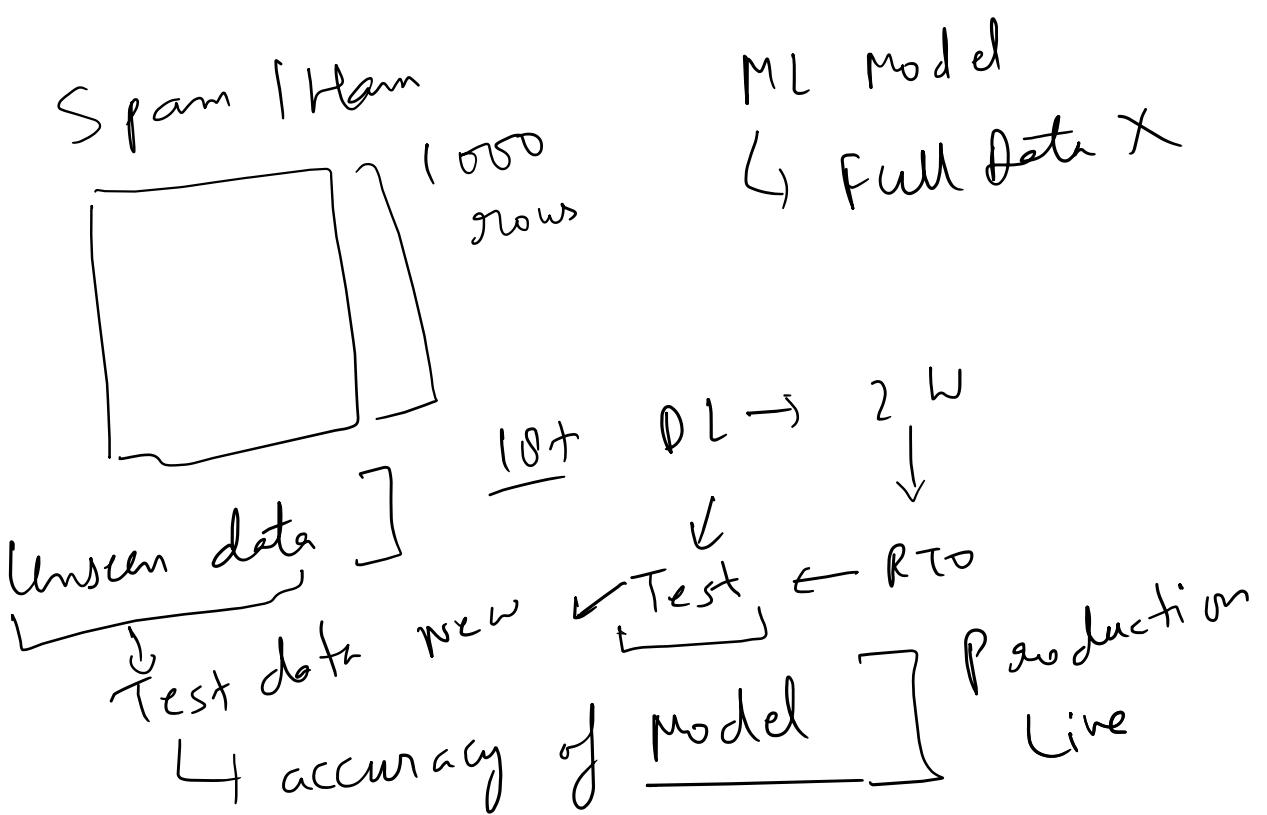
This phenomenon is often called **model rot** or **data drift**.

The solution is to regularly retrain the model on up-to-date data.



**Online Learning-** In online learning, you train the model incrementally by feeding it instances either individually or in small groups called mini-batches.





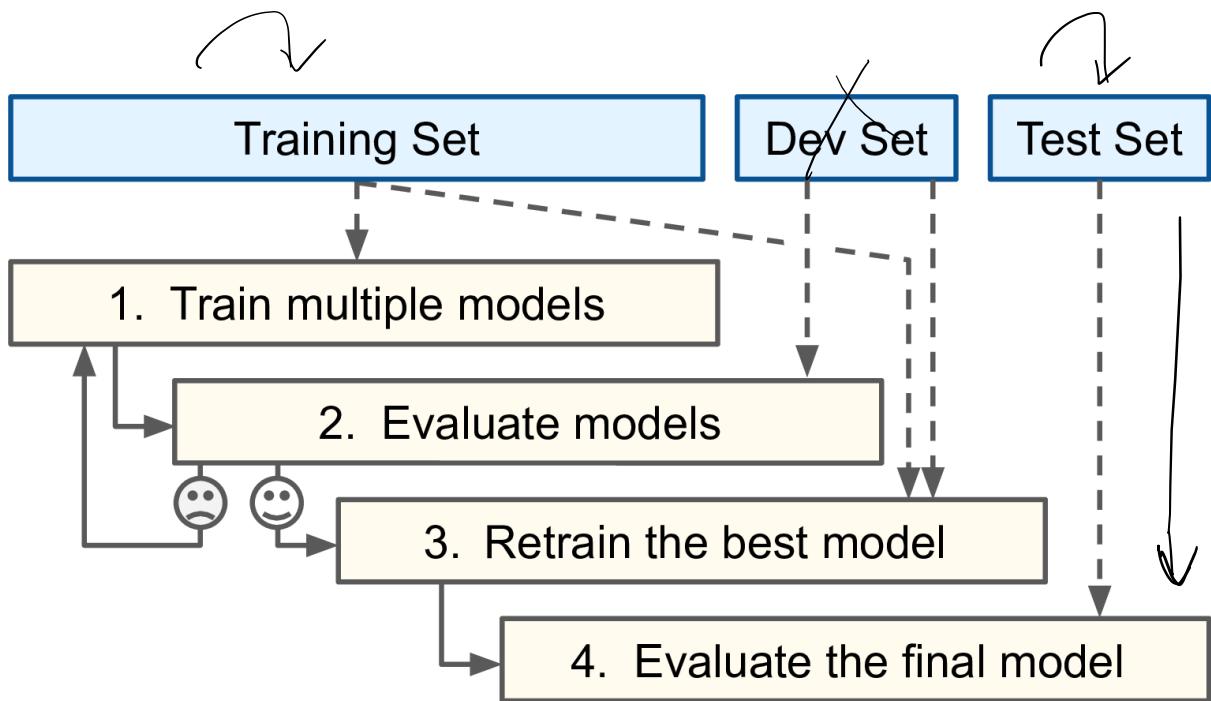
Divide the whole dataset into

train & test set

1000  $\rightarrow$  70% : 30%  $\rightarrow$  test

train  $\leftarrow$  80% : 20%  $\rightarrow$  test

Run model on train dataset  
 2 validate the accuracy on  
test dataset.



ML, DL → Test date to evaluate  
the model