ML101: Session 1

Introduction

IIT Mandi

Presented to you by

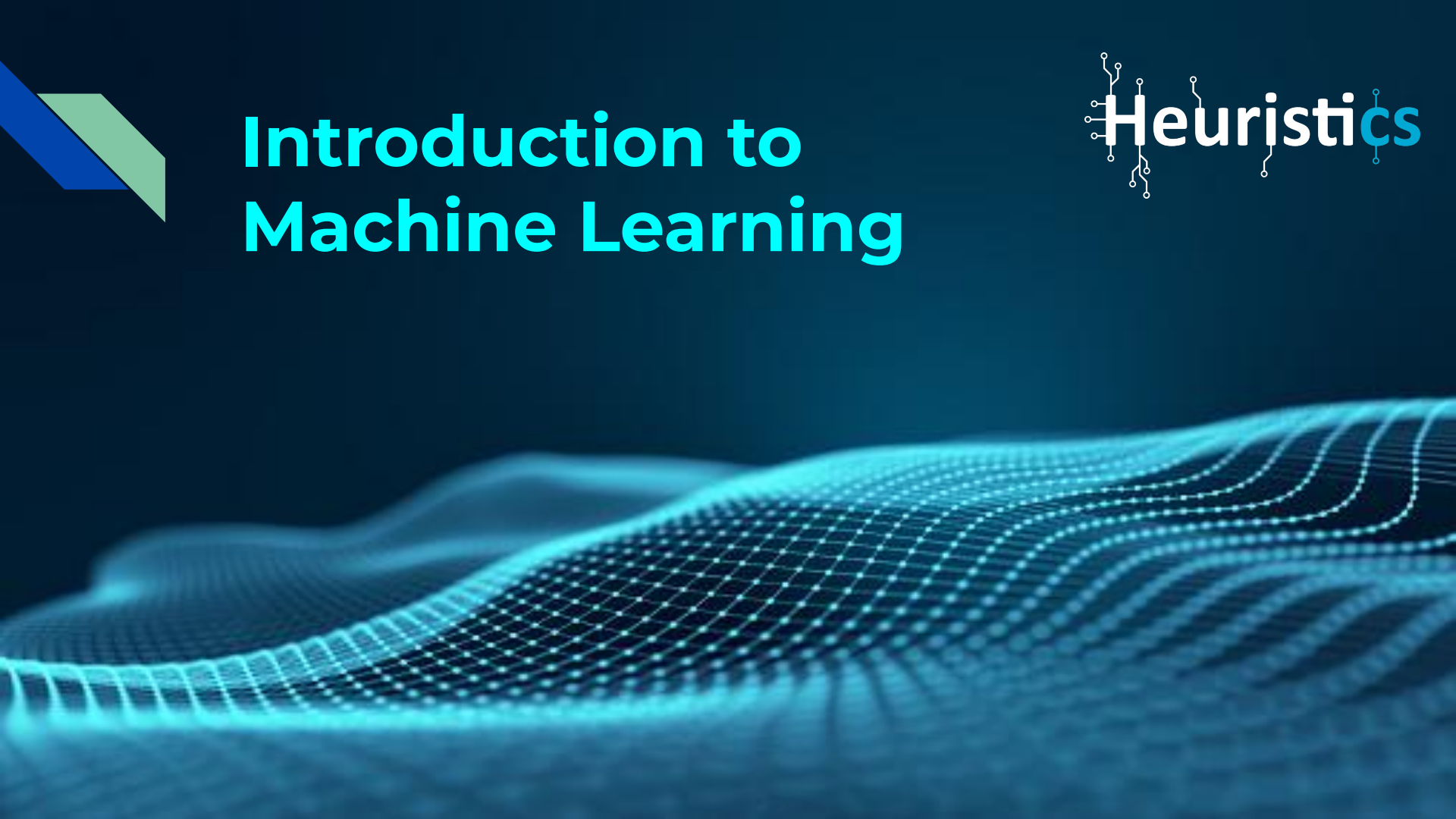
Kamand Prompt

The Programming Club of IIT Mandi


1. Intro to ML
2. Experience ML/AI
3. Basics of Python
Heuristics

Introduction to Machine Learning

Heuristics

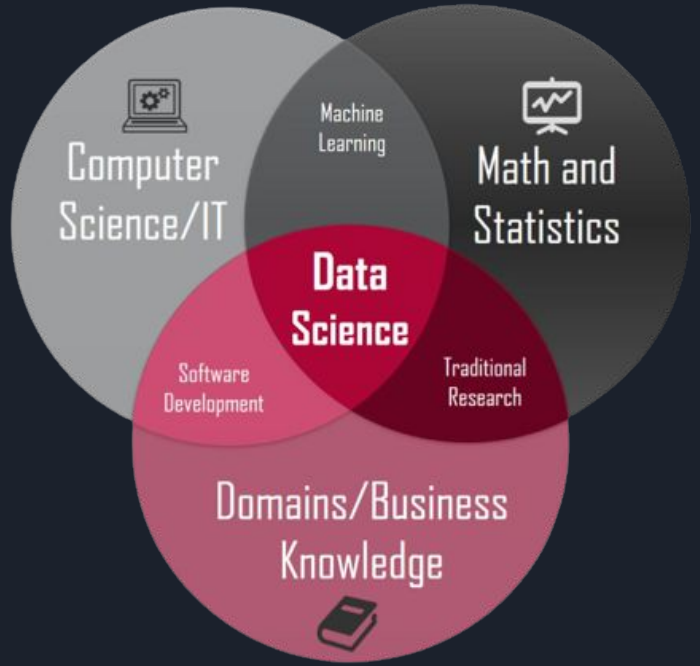


What is Data Science?

- Data Driven Decision making
- Making sense out of data
- Finding hidden patterns in the data
- Analysis using not just machine learning models but also using data visualizations, intelligent reports

Data Science- Amalgam of Multiple Fields

- Mathematics
- Statistics
- Coding
- Database management
- Data Analytics
- Predictive modelling
- Machine Learning
- Deep Learning



Techniques we need to know

Database Knowledge

- Database Management
- Data blending
- Querying
- Data manipulations
- ETL

Predictive Analytics & ML

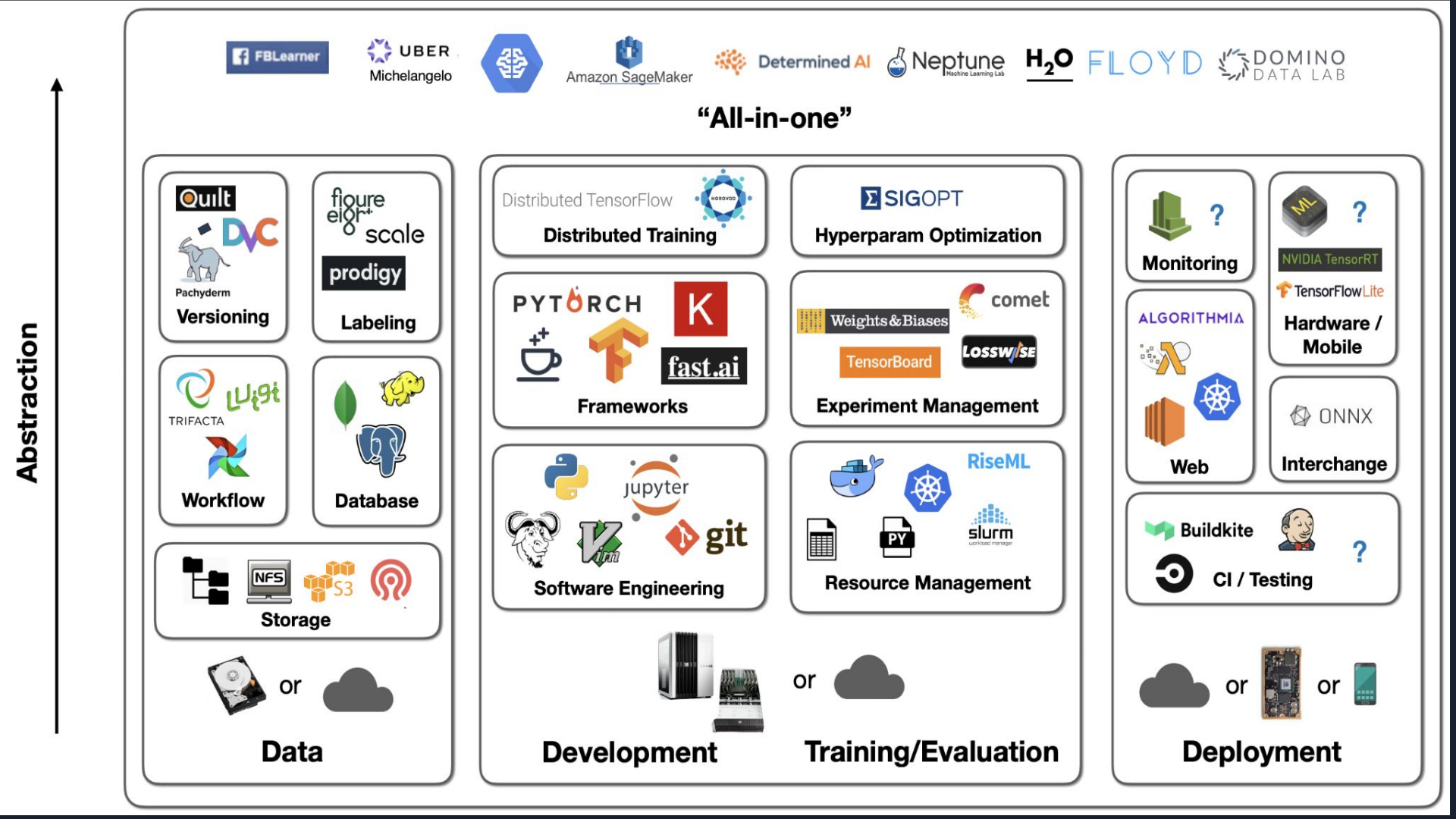
- Basic descriptive statistics
- Advanced analytics
- Predictive modelling
- Machine Learning

Big Data Knowledge

- Distributed Computing
- Big Data Analytics
- Unstructured data analysis

Presentation Skill

- Data visualizations
- Report design
- Insights presentation


FBLearner
UBER
Michelangelo
Amazon SageMaker
Determined AI
Neptune
H2O
FLOYD
DOMINO DATA LAB
"All-in-one"
Abstraction
Quilt
DVC
Pachyderm
Versioning
figure eight
scale
prodigy
Labeling
Luigi
TRIFACTA
Workflow
Database
NFS
S3
Storage
or
Data
Distributed TensorFlow
HOROVOD
Distributed Training
SIGOPT
Hyperparam Optimization
PYTORCH
K
fast.ai
Frameworks
comet
Weights & Biases
TensorBoard
Losswise
Experiment Management
jupyter
git
Vim
Software Engineering
RiseML
slurm
PY
Resource Management
or
Development
or
Training/Evaluation
Monitoring
ML
NVIDIA TensorRT
TensorFlowLite
Hardware / Mobile
ALGORITHMIA
Web
ONNX
Interchange
Buildkite
CI / Testing
or
or
Deployment

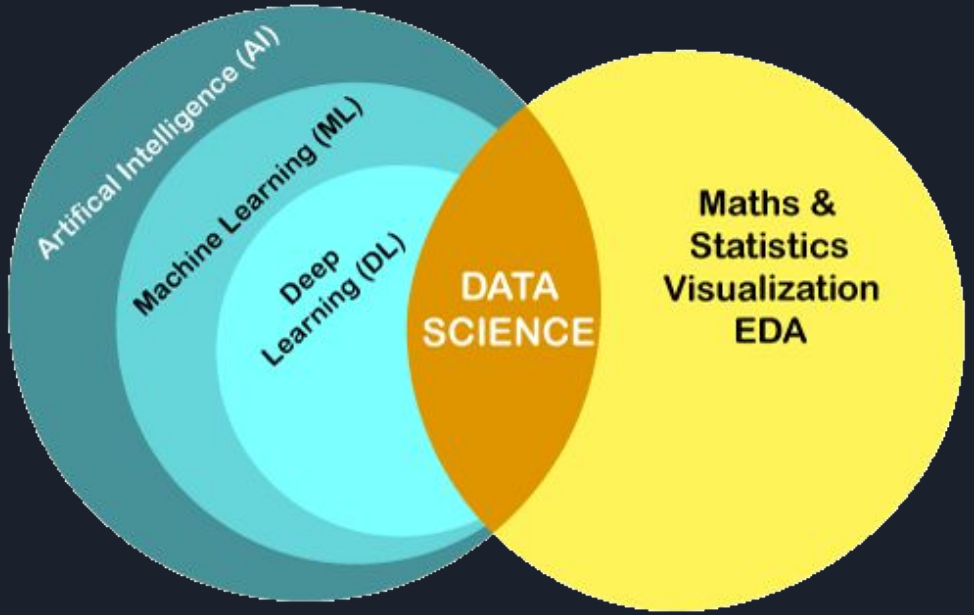
Kaggle

- Users can share their datasets and examine the datasets
- An online community platform for data scientists and machine learning enthusiasts
- Kaggle provides powerful resources on cloud and allows you to use a maximum of 30 hours of GPU and 20 hours of TPU per week.

Google Collab

- Write and execute code in Python
- Document your code that supports mathematical equations
- Create/Upload/Share notebooks
- Import/Save notebooks from/to Google Drive
- Import/Publish notebooks from GitHub
- Import external datasets e.g. from Kaggle
- Integrate PyTorch, TensorFlow, Keras, OpenCV
- Free Cloud service with free GPU

Machine Learning is a part of Data Science



Machine Learning Introduction

What is meant by learning ??


Humanoid
Quadruped
Transformer

Machine Learning

- “Machine Learning is concerned with computer programs that automatically improve their performance through experience.”

- It is all about algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instruction, relying on models and inference instead.

“Learning is any process by which a system improves performance from experience.” -Herbert Alexander Simon

Why Machine Learning?

- Personalized news or mail filter
- Data Mining: Market basket analysis.
- In simple terms
  - Using historical data to make future predictions
  - Building models on historical data to predict
  - Taking training data,building models on the training data using the models to make future predictions
  - Making the machine learn the patterns in the data

It's a Match!

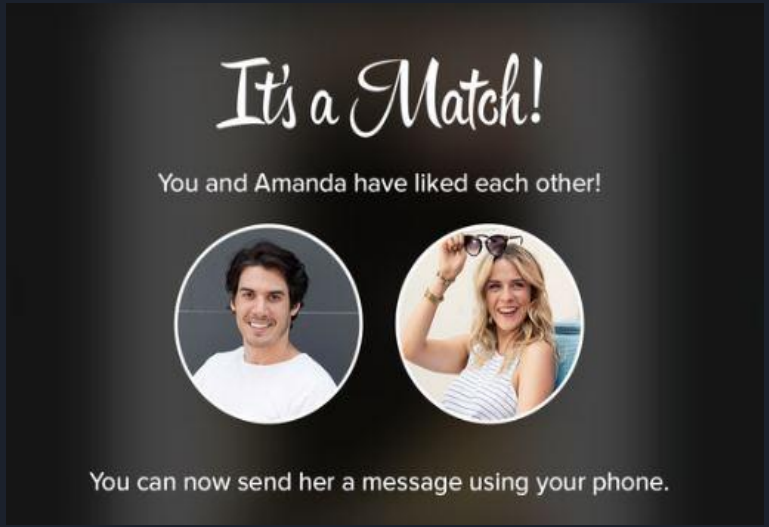
You and Machine learning have liked each other

Where is ML being used here ?

Tinder is an online dating application ~~to find the great love of your life, get married, and have children~~ have “fun”.

- Recommendation system

Tinder implements a machine learning-based algorithm to generate personalized recommendations.

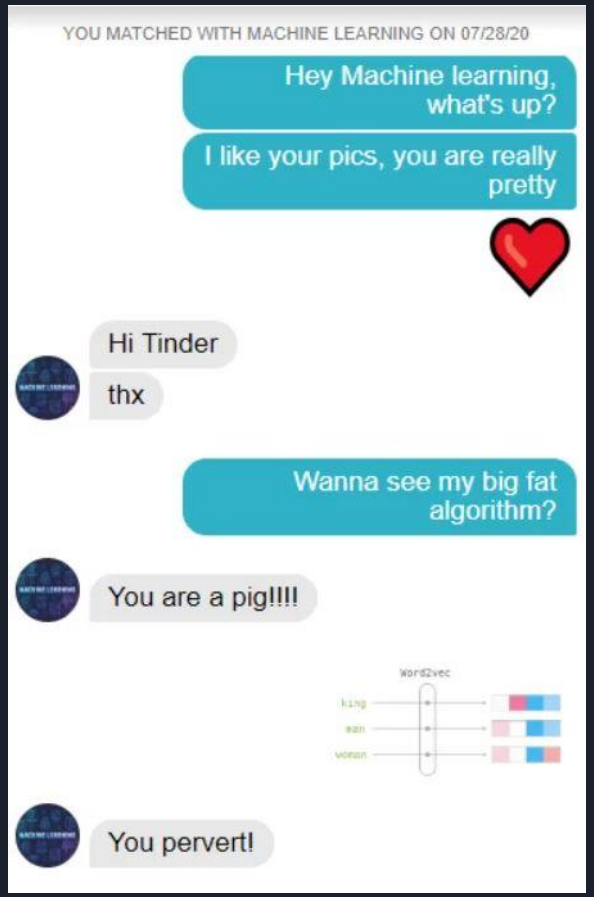


- Harassment checking

The app uses ML to automatically screen for potentially offensive messages.

- “Smart Photos” feature

Machine learning helps users identify which profile pics are the most successful.



Commonly used ML algorithms

- Linear regression
- Logistic regression
- Naive Bayes algorithm
- KNN algorithm
- K-means clustering
- Support Vector Machines
- DBSCAN

Data is in different forms

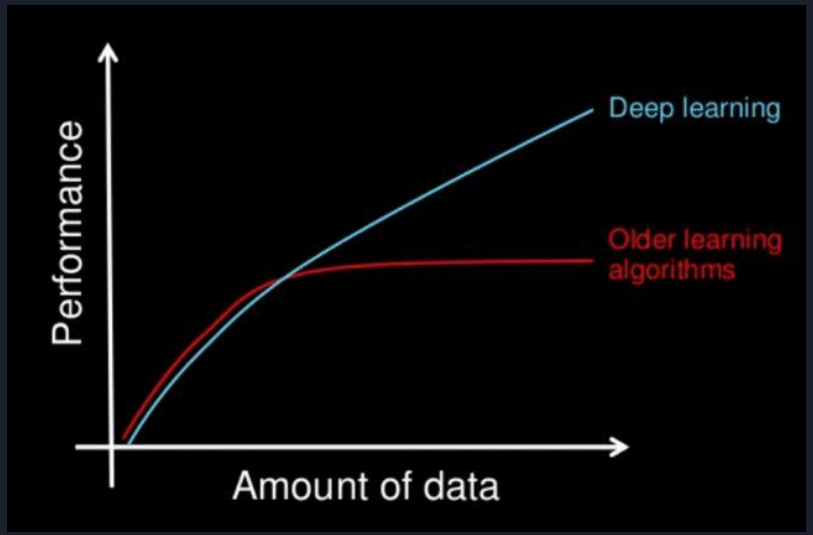
- Numerical Data
- Image data(pixel intensities)
- Video data(frames per second)
- Sound Data(waves)
- Text data(tweets , comments , feedback)

What is Deep Learning?

Deep Learning is a subfield of machine learning

It is inspired by the structure and function of the brain called artificial neural networks.

Deep learning requires substantial computing power.

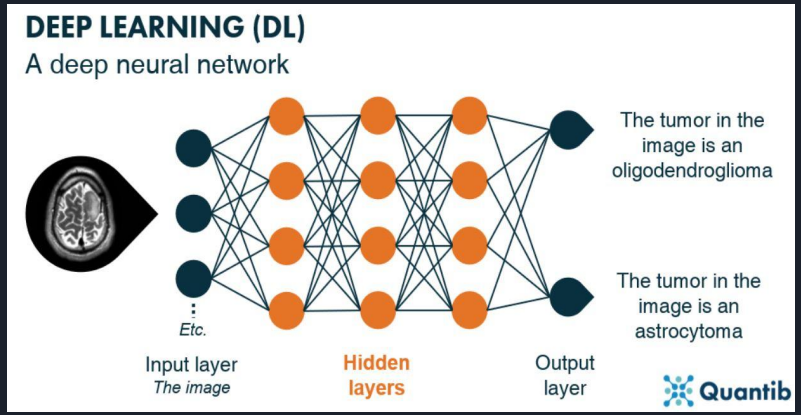


ANN

- ANN - Artificial Neural Network
- ANN has input layer , hidden layer and output layer
- For a really complex and non linear datasets we need several hidden layers
- ANN with multiple hidden layers is known as deep neural network

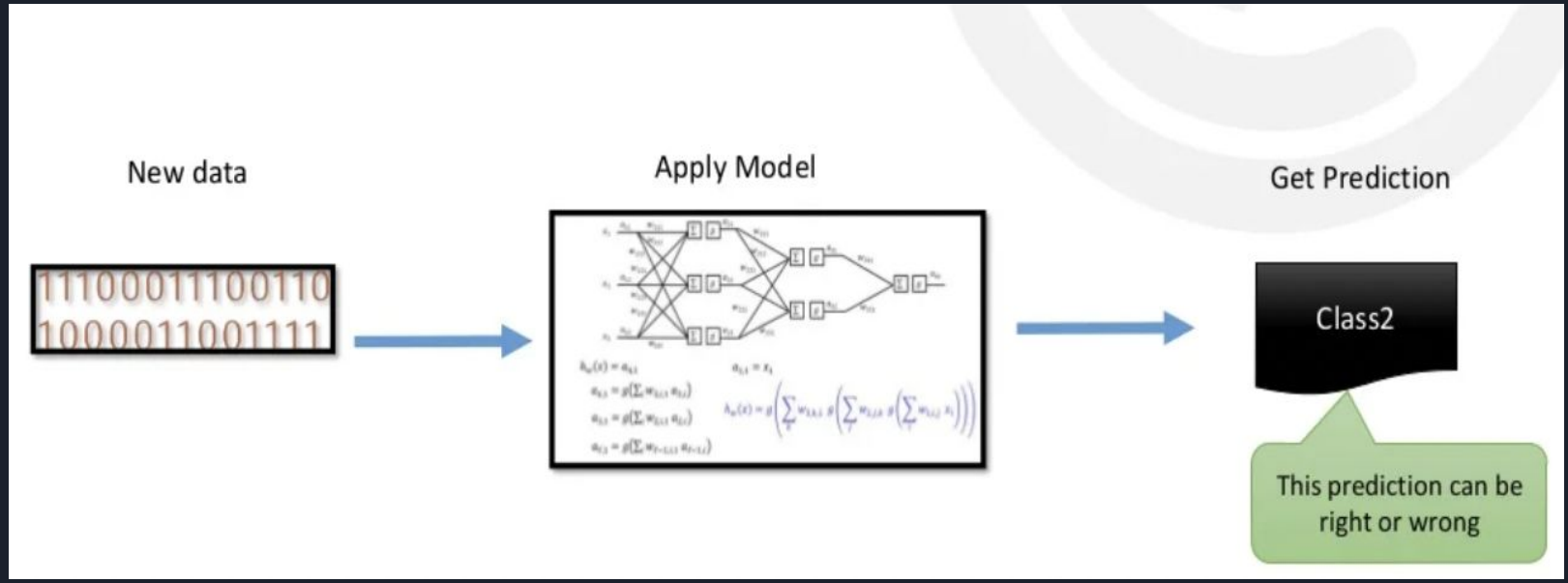
Deep Learning

- ANN with a single layer is known as shallow network
- ANN with multiple hidden layers is known as deep neural network
- Not just multiple hidden layers sometimes the type of hidden layer is also different.
- This concept of solving problems with multiple hidden layers is known as **deep learning.**



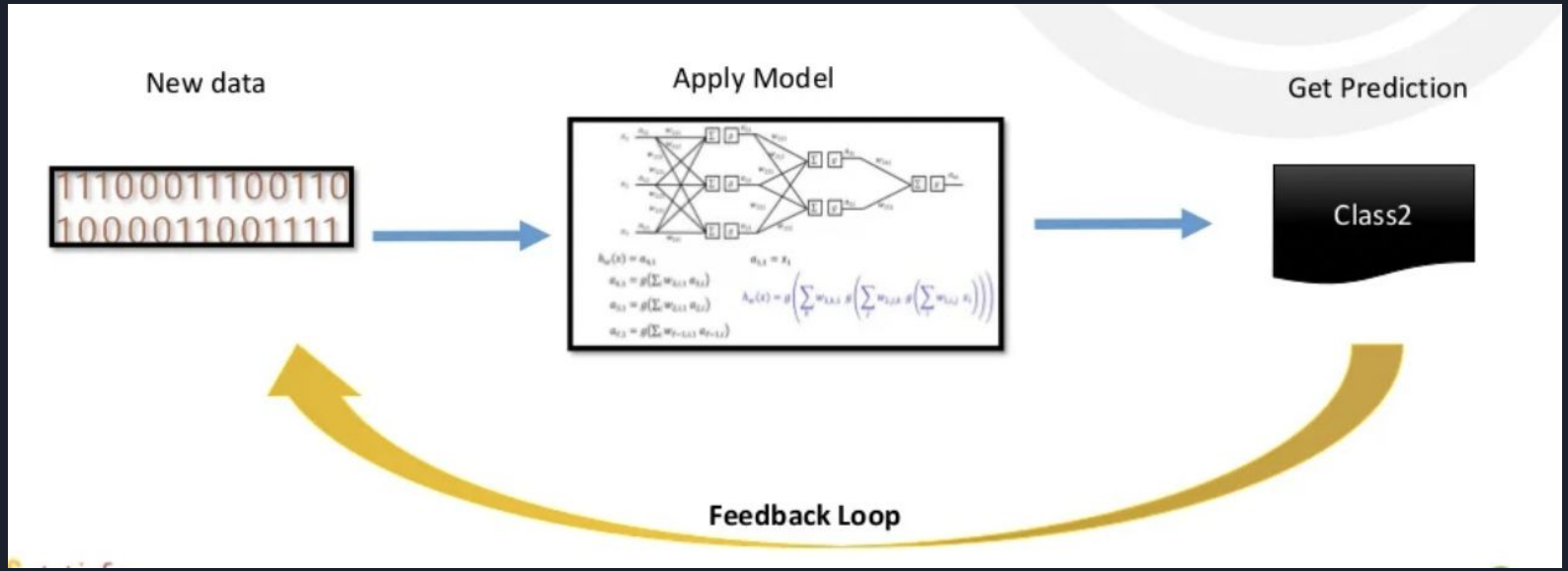
What is Artificial Intelligence?

Machine Learning Models

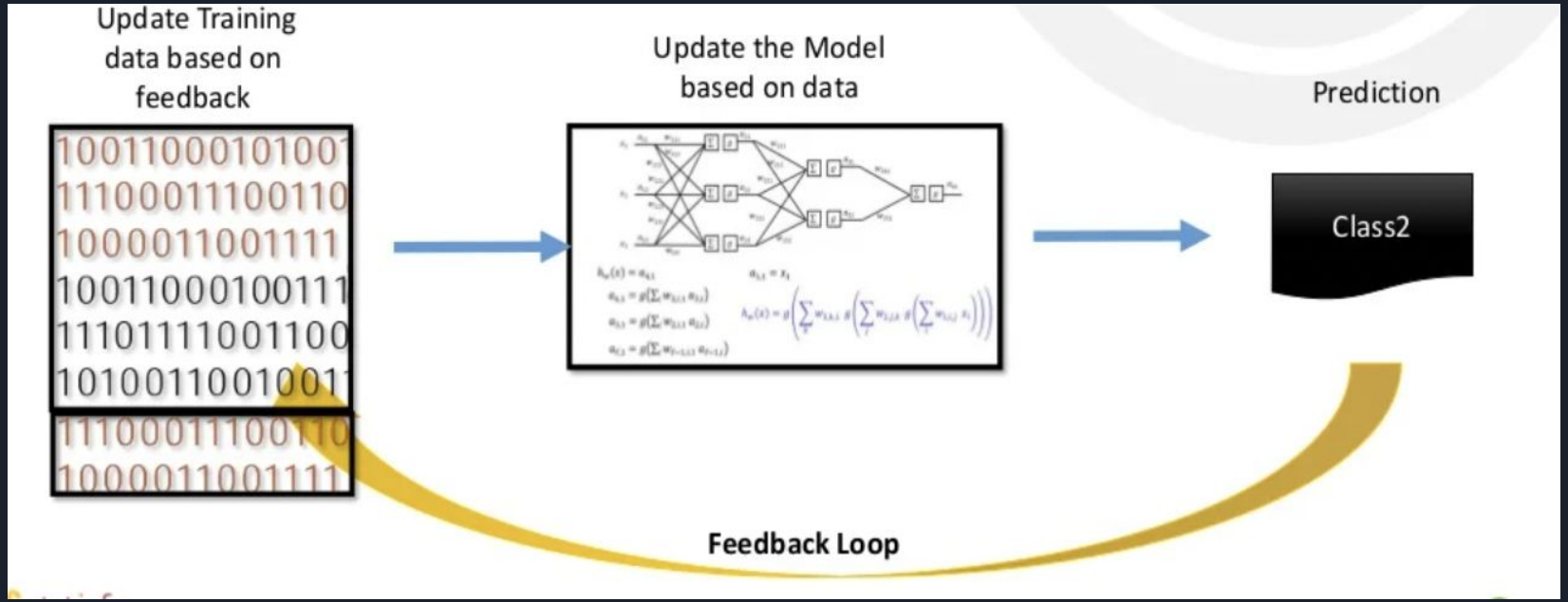


One way Models

AI = Machine Learning Models + Feedback Loop



AI = Machine Learning Models + Feedback Loop



People with no idea about AI saying it will take over the world:

My Neural Network:

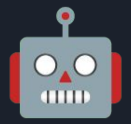
Applications of AI

- Self driving cars
- SIRI/Ok-google
- Alexa /Google home
- Recommendation systems
- Image recognition
- Speech recognition
- Spam filtering

When someone asks you if you like ML ?

Jobs in Data Science

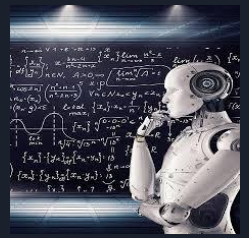
Career Prospects

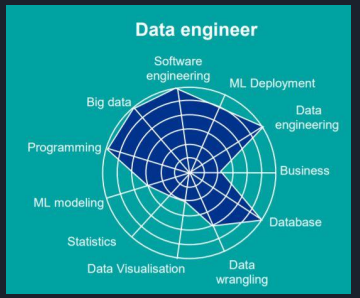



ACHHE
DIN
AA GAYE
HAIN

Career Prospects

- Machine Learning Engineer
- Data Scientist
- Data Analyst
- Data Engineer
- Data Architect
- Business Intelligence Analyst
- Statistician





Experience ML