For research papers:

1. Researchgate
2. arxiv.org
3. paperswithcode
4. IEEE.org -> copy doi number -> search in scihub

For paid paper

1. find doi of research paper
2. Paste in scihub

To read Research Paper

1. Read Abstract
2. Read Conclusion
3. Read body if abstract and conclusion matches or concept.

Application of Machine Learning

1. Fraud Detection system
2. Customer churn prediction
3. Association Mining
4. Handwritten Optical Character Recognition
5. Regression Analysis: If raining predicts mm of water (continuous prediction)
6. Recommendation system
7. Chatbots (virtual assistant)
8. Machine Translation (for e.g., Google translation)

Traditional Programming vs Machine Learning

1. Traditional: Knows (Input, Rule), finds output
2. Machine: Knows (Input, output), finds rule

Types of learning

1. Supervised learning

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | 2 | 3 | Result |
| features | | | Label/features |

For example: check criminal history, condition, credit score etc. to provide loan to a person by a bank.

It can be divided in 2 classes:

* 1. Classification: Weather to give a loan

1. Logistic Regression
2. K-Nearest Neighbor
3. Decision Trees
4. Support Vector Machine
   1. Regression: What amount of loan to be given
5. Unsupervised learning
6. Similar features can be found out using distance
7. Distance close to zero shows similarity
8. Find distance and cluster

Correlation: one variable impact on another

Change in x implies changes in y

 Correlation tells you if a relationship exists between two variables and, if so, its strength and direction. It quantifies how much two variables tend to move together.

 Regression goes a step further. It aims to model that relationship with an equation, allowing you to predict the value of one variable based on the value of another (or several others). It helps understand *how* one variable affects another.

Crisp logic vs fuzzy logic

1. Reinforcement learning

RL by David Silver (YouTube) (Q Learning, DQN, DDQN)

Challenges of Machine Learning:

1. Insufficient Data: Noise in data (for e.g., rare disease)
2. Irrelevant features
3. Hyperparameter tuning and model selection

Parameter: model learns itself by learning and adapting

Hyperparameter: which algo to be chosen? What set of rules? What type of data? (This has to be explicitly defined by the programmer)

1. Complex Business model
2. Computational needed for experiments.

Tasks to do:

Personalized vs Hyper Personalized

Machine Learning Workflow:

1. Gathering data
2. Data preparation (via Data Lake, Data warehouse)
3. Data wrangling
4. Analyze Data
5. Train the model
6. Test the model
7. Deployment