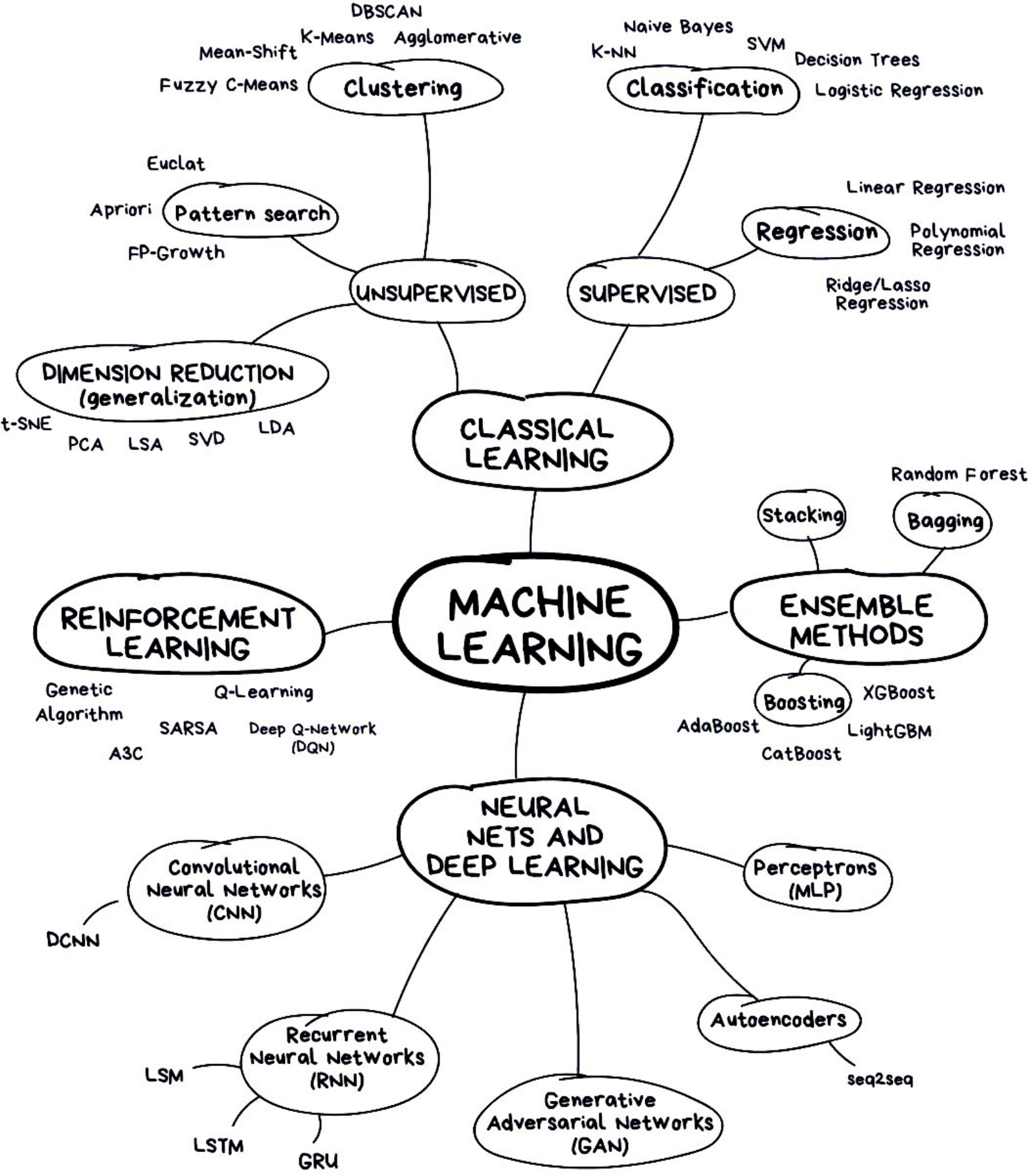

MACHINE LEARNING
CLASSICAL LEARNING
UNSUPERVISED
Clustering
DBSCAN
K-Means
Agglomerative
Mean-Shift
Fuzzy C-Means
Pattern search
Euclat
Apriori
FP-Growth
DIMENSION REDUCTION (generalization)
t-SNE
PCA
LSA
SVD
LDA
SUPERVISED
Classification
Naive Bayes
K-NN
SVM
Decision Trees
Logistic Regression
Regression
Linear Regression
Polynomial Regression
Ridge/Lasso Regression
ENSEMBLE METHODS
Stacking
Bagging
Random Forest
Boosting
XGBoost
AdaBoost
LightGBM
CatBoost
REINFORCEMENT LEARNING
Genetic Algorithm
Q-Learning
SARSA
Deep Q-Network (DQN)
A3C
NEURAL NETS AND DEEP LEARNING
Convolutional Neural Networks (CNN)
DCNN
Perceptrons (MLP)
Recurrent Neural Networks (RNN)
LSM
LSTM
GRU
Generative Adversarial Networks (GAN)
Autoencoders
seq2seq

Learn Python and Machine Learning!

PYTHON BASIC

- Data types (strings, integers, floats, etc.)
- Variables
- Control flow (if/else statements, loops)
- Functions
- Modules and packages
  Input/output

LIBRARIES

- NumPy: for numerical computations and array manipulation
- Pandas: for data analysis and manipulation
- Matplotlib: for data visualization
- Scikit-learn: for machine learning algorithms and tools
- TensorFlow or PyTorch: for deep learning

DATA STRUCTURES

- Lists, tuples, and dictionaries
- Sets and frozen sets
- Stacks and queues
- Linked lists
- Trees (binary trees, AVL trees, binary search trees)
- Graphs (directed and undirected graphs, adjacency matrix, adjacency list)

MATHEMATICS

- Linear algebra (vectors, matrices, matrix multiplication)
- Calculus (differentiation, integration)
- Probability and statistics
- Optimization

DATA PREPROCESSING

- Handling missing values
- Scaling and normalization
- Encoding categorical data
- Feature selection and engineering

ML ALGORITHMS

- Linear regression
- Logistic regression
- Decision trees
- Random forests
- Support vector machines
- Naive Bayes
- K-nearest neighbors
- Clustering (K-means, hierarchical clustering)
- Dimensionality reduction (PCA)

MODEL EVALUATION

- Accuracy, precision, recall, F1 score
- Confusion matrix
- Cross-validation
- Bias-variance tradeoff

DEEP LEARNING

- Neural networks
- Convolutional neural networks (CNNs)
- Recurrent neural networks (RNNs)
- Generative adversarial networks (GANs)

PROJECTS

- Develop a machine learning project from scratch
- Use real datasets for training and testing
- Evaluate the performance of the models
- Optimize the models for better performance
- Deploy the models in production

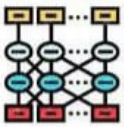
Save for later

Must Know Terms In LLM

Foundation Model

LLM designed to generate and understand human-like text across a wide range of usecases

Transformer



A popular LLM design known for its attention mechanism and parallel processing abilities

Prompting

Providing carefully crafted inputs to an LLM to generate desired outputs

Context-Length

Maximum number of input words/tokens an LLM Can consider when generating an output.

Few-Short Learning

Providing very few examples to an LLM to assis it in performing a specific task

Zero-Short Learning

Providing only task instructions to the LLM relying solely on its preexisting knowledge

RAG

Retrieval-Augmented Generation. Appending retrieved information to improve LLM response

Knowledge Base(KB)

Collection of documents from which relevant information is retrieved in RAG

Vector Database

Stores vector representations of the KB, aiding the retrieval of relevant information in RAG.

Fine-Tuning

Adapting an LLM to a specific task or domain by further training it on task-specific data.

Instruction Tuning

Adjusting an LLM's behaviour during fine-tuning by providing specific guidlines/directions

Hallucination

Tendency of LLMs to sometimes generate incorrect or non-factual information.

1. Linear Regression

Overview

Linear regression is a foundational algorithm in machine learning, used for predicting a continuous variable.

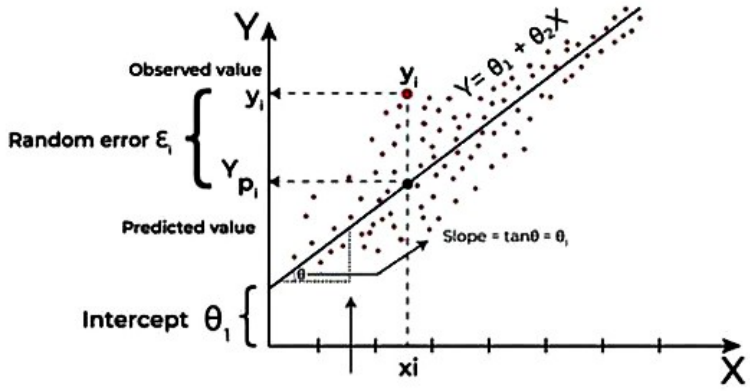
Learning Objectives

- Understand the theory behind linear regression.
- Learn to implement linear regression in Python.

Practice Questions

1. Implement linear regression to predict housing prices using a given dataset.

2. How would you evaluate the performance of your linear regression model?



Machine Learning Tools

Machine Learning Languages

Python

R

C++

Machine Learning Frameworks

Numpy

Scikit-Learn

NLTK

Big Data Tools

Apache Spark

Database

Singlestore

Data Analysis and Visualisation tools

Pandas

Matplotlib

Jupyter Notebook

Tableau

Weka

ML Frameworks for Neural Network Modelling

Pytorch

Tensorflow & Tensorboard

Keras

Some Other Tools

Accord.Net

Accord.Net is a .Net machine Learning framework combined with audio and image processing libraries written in C#.

Mlflow

Managing machine learning experiments.

Apache Mahout

It is a distributed linear algebra framework and mathematically expressive Scala DSC

Singa

This software is primarily used in Natural Language Processing (NLP) and image recognition and supports a wide range of deep learning models.

Machine Learning Algorithms

Text Analysis

Derives high-quality information from text

- Latent Dirichlet Allocation → Unsupervised topic modeling, group texts that are similar
- Extract N-Gram Features from Text → Creates a dictionary of n-grams from a column of free text
- Feature Hashing → Converts text data to integer encoded features using the Vowpal Wabbit library
- Preprocess Text → Performs cleaning operations on text, like removal of stop-words, case normalization
- Word2Vector → Converts words to values for use in NLP tasks, like recommender, named entity recognition, machine translation

Multiclass Classification

Answers complex questions with multiple possible answers

- Multiclass Logistic Regression → Fast training times, linear model
- Multiclass Neural Network → Accuracy, long training times
- Multiclass Decision Forest → Accuracy, fast training times
- One-vs-All Multiclass → Depends on the two-class classifier
- One-vs-One Multiclass → Depends on binary classifier, less sensitive to an imbalanced dataset with larger complexity
- Multiclass Boosted Decision Tree → Non-parametric, fast training times and scalable

Regression

Makes forecasts by estimating the relationship between values

- Fast Forest Quantile Regression → Predicts a distribution
- Poisson Regression → Predicts event counts
- Linear Regression → Fast training, linear model
- Bayesian Linear Regression → Linear model, small data sets
- Decision Forest Regression → Accurate, fast training times
- Neural Network Regression → Accurate, long training times
- Boosted Decision Tree Regression → Accurate, fast training times, large memory footprint

Two-Class Classification

Answers simple two-choice questions, like yes or no, true or false

- Two-Class Support Vector Machine → Under 100 features, linear model
- Two-Class Averaged Perceptron → Fast training, linear model
- Two-Class Decision Forest → Accurate, fast training
- Two-Class Logistic Regression → Fast training, linear model.
- Two-Class Boosted Decision Tree → Accurate, fast training. large memory footprint
- Two-Class Neural Network → Accurate, long training times

Anomaly Detection

Identifies and predicts rare or unusual data points

- One Class SVM → Under 100 features, aggressive boundary
- PCA-Based Anomaly Detection → Fast training times

Recommenders

Predicts what someone will be interested in

- Use the Train Wide & Deep Recommender module → Hybrid recommender, both collaborative filtering and content-based approach
- SVD Recommender → Collaborative filtering, better performance with lower cost by reducing dimensionality

Image Classification

Classifies images with popular networks

- ResNet → Modern deep learning neural
- PCA-Based Anomaly Detection → Network

Clustering

Separates similar data points into intuitive groups

- K-Means → Unsupervised learning

Roadmap To Learn Machine Learning

Mathematics

- Probability
- Discrete
- Statistics

Programming

- Python
- R

Databas

- MySQL
- MongoDB

ML Algorithms

- Linear Logistic Regression
- KNN
- K-means
- Random forest & more!

Machine Learning

- ML Libraries and Non- ML Libraries

Machine Learning

- Scikit learn Supervised learning
- Unsupervised learning
- Reinforcement learning

Deep Learning

- TensorFlow
- Keras
- Neural Networks
- CNN, RNN, GAN, LSTM

Data Visualization Tools

- Tableau
- Qlikview
- PowerBI

ML Engineer

Types of Machine Learning Algorithms

Supervised Learning

Algorithms learn from labeled data. The input data comes with corresponding correct outputs, allowing the algorithm to make predictions on unlabeled data.

Example

- Decision Tree
- Support Vector Machine (SVM)
- K-Nearest Neighbor (KNN)
- Linear Regression

Real-life Analogy

Imagine a teacher supervising a class. The teacher knows the correct answers and guides students to learn them. The algorithm learns from a training dataset and improves through feedback.

Unsupervised Learning

Algorithms train on unlabeled data without knowledge of the output, seeking to find patterns in the data.

Example

- K-Means Clustering
- Linear Discriminant Analysis (LDA)
- Principal Component
- Analysis (PCA) Mean Shift

Real-life Analogy

Without a teacher, students are left to learn on their own. There is no specific answer to find; the algorithm explores the data to uncover underlying structures

Semi-Supervised Learning

Combines labeled and unlabeled data for training, using a small amount of labeled data to guide learning from a larger set of unlabeled data.

Example

- Self-Training
- Co-Training
- Graph-Based Methods
- Google Photos

Real-life Analogy

A teacher helps students with books while also trying to teach those without. The algorithm learns from the limited labeled dataset and applies insights to the larger dataset.

Reinforcement Learning

Algorithms interact directly with their environment, receiving rewards for correct actions and penalties for incorrect ones. They learn from mistakes to improve performance.

Example

- AlphaGo
- Deep Q-Network (DQN)
- Autonomous Vehicles
- Robotic Arm Manipulation

Real-life Analogy

Students learn from their experiences, adjusting their actions based on past outcomes. The algorithm uses trial and error to determine optimal actions that maximize future rewards.

Top 10 ML Algorithms

1.Linear Regression

Used for predicting continuous outcomes

- Models the relationship between
- dependent and independent variables. Fits a linear equation to observed data.
- Equation: $Y = aX + b$, where: Y is the dependent variable, X is the independent variable, a is the slope, and b is the intercept.

2. Logistic Regression

Used for binary classification tasks (e.g., predicting yes/no outcomes

- Estimates probabilities using a logistic function.
- Can be enhanced by including interaction terms and and regularization techniques.
- Often used in medical diagnosis and credit scoring.

3. Decision Tree

Classifies data based on feature values

- Divides the population into homogeneous sets based on significant attributes.
- Easy to interpret and visualize.
- Prone to overfitting if not properly pruned.

4. Support Vector Machine (SVM) 4

Primarily used for classification tasks.

- Effective in high-dimensional spaces.
- Utilizes hyperplanes to separate classes. Can handle both linear and non-linear
- classification through the kernel trick.

5. Naive Bayes Algorithm

Primarily used for classification tasks.

- Assumes independence among features
- when calculating probabilities. Fast and efficient, especially for large datasets.
- Commonly used in spam detection and sentiment analysis.

6. K-Nearest Neighbors (KNN)

Can be applied to both classification and regression tasks

- Classifies new data points based on the majority class of their k nearest neighbors. Simple to implement and intuitive.
- Computationally expensive for large datasets; requires distance normalization,

7. K-Means Clustering

Unsupervised learning algorithm for clustering tasks.

- Groups data into a specified number of clusters (k).
- Iteratively assigns data points to the nearest centroid and recalculates centroids.
- Sensitive to initial centroid selection;
- may converge to local minima.

8. Random Forest Algorithm

An ensemble method for classification and regression

- Combines multiple decision trees to improve accuracy and prevent overfitting.
- Uses majority voting for classification tasks.
- Robust to noise and capable of handling large datasets with high dimensionality

9.Dimensionality Reduction Algorithms

Reduce the number of features in a dataset while retaining important information.

- Techniques like PCA help visualize data and improve model performance.
- Reduces computational cost and mitigates the curse of dimensionality.
- Useful in preprocessing steps for machine learning models.

10. Gradient Boosting & AdaBoosting Algorithms

Enhance prediction accuracy through ensemble learning techniques.

- Combines multiple weak learners to create a strong predictive model. Gradient boosting builds trees in a sequential manner, minimizing errors from previous models.
- AdaBoost adjusts weights of misclassified instances, focusing on hard-to-classify examples.

Performance and Efficiency of Machine Learning Algorithms

Fastest Machine Learning Algorithms

Some of the fastest machine learning algorithms include:

- Linear Regression
- K-Nearest Neighbors (KNN)
- Naive Bayes
- Stochastic Gradient Descent
- (SGD) Decision Trees

Most Commonly Used Machine Learning Models

The most frequently used machine learning models are:

- Linear Regression
- Logistic Regression
- Random Forests
  Decision Trees

Best Algorithm for Prediction

For predictive model building, Linear Regression is widely used. Other effective algorithms for prediction include:

- Decision Trees
- Support Vector Machines
- (SVM)
- Neural Networks
- Gradient Boosting Methods

Efficient Machine Learning Algorithms

Efficiency varies by task, but some of the fastest and most efficient machine learning algorithms include:

- Random Forests
- XGBoost
- Linear Regression
- K-Nearest Neighbors (KNN)

BEST FREE RESOURCES TO LEARN AI

YOUTUBE

- 3Blue Brown
- Dirk Zee
- CS Dojo
- Analytics Vidhya
- Two Minute
- Papers
- Sentdex
- Corey Schafer
- Alex The Analyst
- Matt Wolfe
- Data Science Dojo
- Abhishek Thakur
- Lex Fridman
- StatQuest With Josh Starmer

BLOGS

- distill.pub
- machinelearningisfun.com
- machinelearningmastery.com
- fastml.com
- ai.googleblog.com
- towardsai.net
- kdnuggets.com
- analyticsvidhya.com
- towardsdatascience.com
- openai.com/blog

WEBSITES FOR DATASET

- Kaggle
- UCI Machine Learning Repository
- OpenML
- Google Cloud AI Platform Datasets
- Microsoft Azure Open Datasets
- Amazon SageMaker Open Data Registry
- Papers with Code
- Hugging Face Datasets
- OpenMLDB
- MachineHack Datasets

WEBSITES FOR COURSES

- mygreatlearning.com
- classcentral.com
- dirkzee.com
- simplilearn.com
- edx.org
- freecodecamp.org
- udacity.com
- deeplearning.ai
- ai.google
- pil.harvard.edu

COURSES FROM GOOGLE

- AI for Everyone
- Machine Learning Crash Course
- Elements of Artificial Intelligence
- Building TensorFlow Lite Applications Introduction to TensorFlow for Artificial Intelligence, Machine Learning, and Deep Learning
- Fundamentals of Reinforcement Learning
- Generative Adversarial Networks (GANS)

COURSES FROM MICROSOFT

- AI for Beginners
- Introduction to Artificial intelligence
- Azure AI Fundamentals
- Machine Learning for Beginners
- Responsible AI
- transform Your Business with AI
- Career Essentials in Generative AI
- Building AI Solutions on Azure

EXCEL VS SQL VS PYTHON

TASK	EXCEL	SQL	PYTHON (PANDAS)
Load Data	Open Excel file or use File > Open	SELECT FROM table_name;	df = pd.read_csv("file.csv")
Filter Rows	=FILTER(A2: B10, B2:B10>100)	SELECT * FROM table WHERE column > 100;	df [df['column'] > 100]
Select Columns	Use column letters (e.g., A, B)	SELECT column1, column2 FROM table;	df[['column1', 'column2']]
Short Data	Data > Sort	SELECT * FROM table ORDER BY column DESC;	df.sort_values(by='column', ascending=False)
Group By/ Aggregate	Use Pivot Table	SELECT dept, COUNT(*) FROM emp GROUP BY dept;	df.groupby('dept').size) or agg
Count Rows	=COUNTA(A2:A100)	SELECT COUNT(*) FROM table;	len(df) or df.shape[0]
Average/ Mean	=AVERAGE (B2:B100)	SELECT AVG(salary) FROM emp;	df['salary'].mean()
SUM	=SUM(B2:B100)	SELECT SUM(sales) FROM data;	df['sales'].sum()
Remove Duplicates	Data > Remove Duplicates	SELECT DISTINCT column FROM table;	df.drop_duplicates()
Join Tables	Use VLOOKUP or XLOOKUP	SELECT * FROM A JOIN B ON A.id = B. id;	pd.merge(df1, df2, on='id')
Create New Column	=B2 * 0.1 in new column	SELECT salary, salary*0.1 AS bonus FROM emp;	df['bonus'] = df['salary'] 0.1
Rename Column	Rename Manually	SELECT column AS new_name FROM table;	df.rename(columns={'old': 'new'}, inplace=True)
Handle Missing Data	=IF(ISBLANK (A2), "N/A", A2)	Depends on DB: use IS NULL, COALESCE()	df.fillna('N/A') or df.dropna()
Export Data	File Save As (CSV/XLSX)	Use tool (e.g., SSMS) or INTO OUTFILE	df.to_csv("output.csv", index=False)
Data Visualization	Insert > Charts	Not native; use BI tools	df.plot(kind='bar'), seaborn, matplotlib

MACHINE LEARNING PROJECTS

Basic Projects

- Iris Flower Classification
- Titanic Survival Prediction
- House Price Prediction
- Spam Email Detection
- Basic Image Classification
- Student Performance Prediction
- Weather Forecasting
- Heart Disease Prediction
- Wine Quality Prediction
- Loan Approval Prediction
- Breast Cancer Classification
- K-Means Clustering
- MNIST Digit Classification
- Customer Segmentation
- Fake News Detection
- Credit Card Fraud Detection
- Stock Price Prediction (Basic)
- Handwritten Character Recognition
- Chess Move Prediction
- Number Plate Recognition
- Basic Recommender System
- Age Prediction From Facial Image
- Music Recommendation
- Sentiment Analysis
- Handwritten Digit Recognition

Intermediate Projects

- Movie Recommendation
- Stock Price Prediction
- Fraud Detection System
- Voice Assistant
- Customer Churn Prediction
- News Classification
- Disease Prediction
- Face Recognition System
- Credit Scoring Model
- Email Spam Classification
- Text Summarization
- Recommender System
- Sales Forecasting
- Time Series Forecasting
- Image Captioning.
- Text Classification With NLP
- Multi-Class Classification
- Loan Default Prediction
- Music Genre Classification
- Human Activity Recognition
- Image Super-Resolution
- Collaborative Filtering
- Named Entity Recognition
- Autonomous Vehicles
- Real-Time Object Detection

AI Concepts Explained

Meta Prompts

- Adaptive tasking strategies.
- Customized content generation.
- Efficient training optimization.

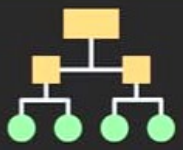
Prompt Chaining

- Altering input conditions.
- Tailoring AI responses.
- Customizing model behavior.

Agents

- Autonomous decision-making.
- Intelligent system entities.
- Interactive user interfaces.

Classification Model



- Categorizing data points.
- Predictive modeling applications.
- Efficient decision-making support.

Supervised Learning

- Labeled training data.
- Predictive modeling accuracy.
- Targeted output training.

Unsupervised Learning

- No labeled data.
- Clustering patterns discovery.
- Anomaly detection applications.

Reinforcement Learning

- Trial and error learning.
- Game strategy optimization.
- Autonomous system training.

Neural Learning

- Mimicking brain functions.
- Dynamic pattern recognition.
- Cognitive computing approaches.

Transfer Learning

- Knowledge transferability.
- Pre-trained model application.
- Improved model training.

Generative Adversarial Networks (GANs)

- Synthetic data generation.
- Image and content creation.
- Creative model applications.

Attention Mechanism

- Focus on specific information.
- Image and text tasks.
- Enhanced sequence modeling.

Neural Language Processing (NLP)

- Natural language understanding.
- Sentiment analysis applications.
- Language generation tasks.

Computer Vision

- Image and video analysis.
- Object detection capabilities.
- Visual data interpretation.

Recurrent Neural Networks (RNNs)

- Sequential data processing.
- Time series predictions.
- Natural language tasks.

Convolutional Neural Networks (CNNs)

- Image recognition tasks.
- Visual pattern extraction.
- Feature learning in images.

Deep Reinforcement Learning

- Enables complex decision-making tasks.
- Used in advanced gaming AI.

Explainable AI (XAI)

- Enhances AI transparency and interpretability.
- Improves trust in AI decisions.

Swarm Intelligence

- Collective behavior from decentralized systems.
- Used in optimization problems.

Few-shot Learning

- Learns from limited training examples.
- Reduces data dependency significantly.

AI Ethics

- Ensures fair AI system development.
- Addresses bias and discrimination issues.

0 to Data Scientist Complete Roadmap

0% - 10%: MATHEMATICS & STATISTICS

- Algebra, calculus, probability, and basic statistics.

10% - 20%: EXCEL

- Data manipulation, cleaning, and basic visualization.

20% - 30%: SQL

- Querying databases, joins, aggregations.

30% - 40%: PROGRAMMING (PYTHON/R)

- Basic syntax, data structures, libraries (Pandas, NumPy).

40% - 50%: DATA VISUALIZATION

- Creating charts/graphs with tools like Matplotlib, Seaborn, or Tableau..

50% - 60%: EXPLORATORY DATA ANALYSIS (EDA)

- Identifying patterns, data summarization

60% - 70%: MACHINE LEARNING BASICS

- Supervised/unsupervised learning, model evaluation.

70% - 80%: ADVANCED MACHINE LEARNING

- Deep learning, NLP, CV, model deployment.

80% - 90%: BIG DATA TOOLS

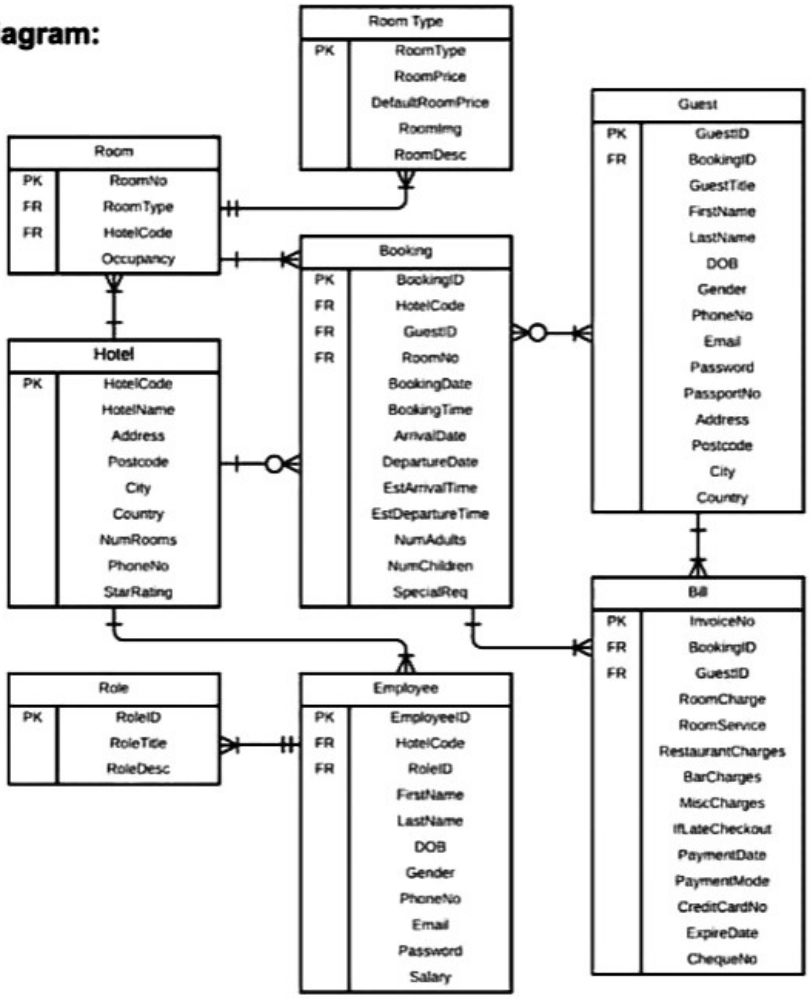
- Hadoop, Spark, cloud platforms.

90% - 100%: REAL-WORLD PROJECTS

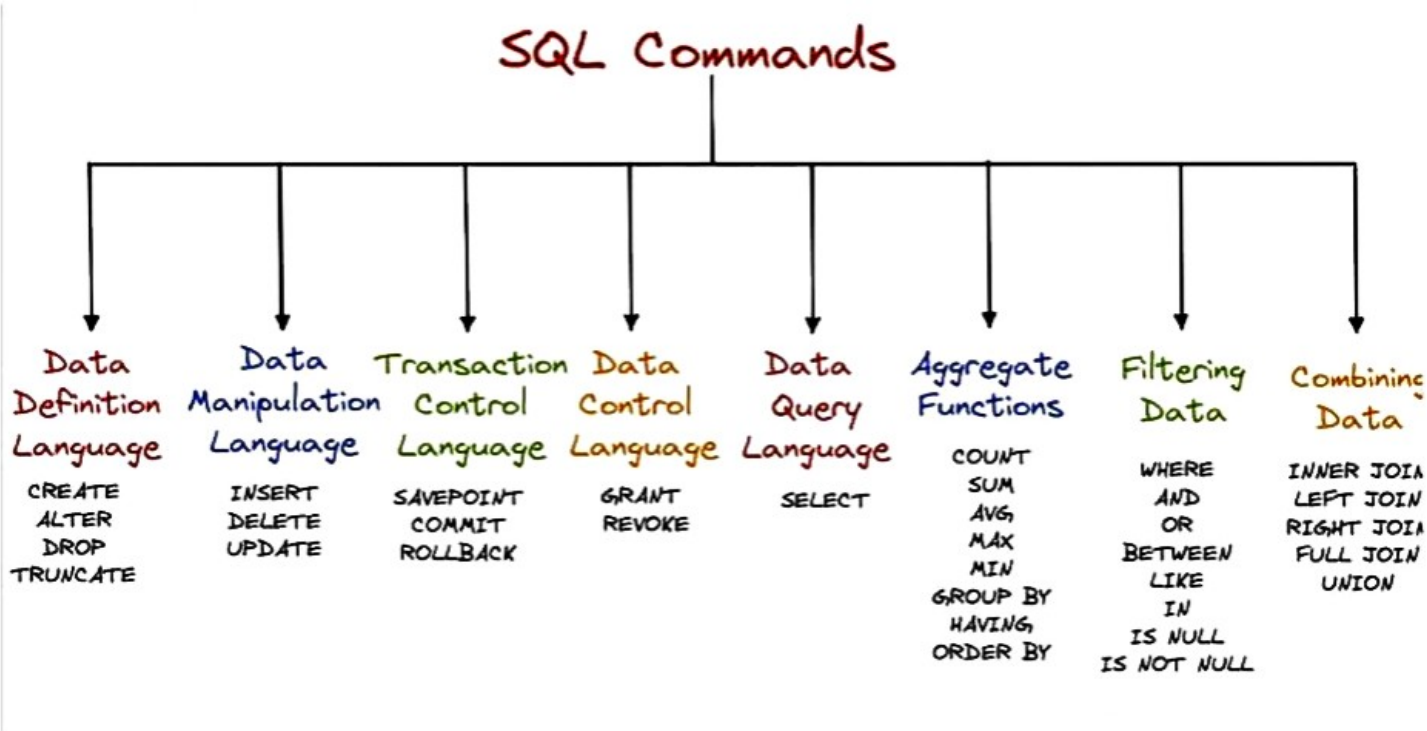
- End-to-end projects, case studies, portfolio building.

SQL Basics For Data Scientists

Entity Relationship Diagram:



SQL Commands:



Machine Learning Algorithms

Text Analysis

Derives high-quality information from text

- Latent Dirichlet Allocation → Unsupervised topic modeling, group texts that are similar
- Extract N-Gram Features from Text → Creates a dictionary of n-grams from a column of free text
- Feature Hashing → Converts text data to integer encoded features using the Vowpal Wabbit library
- Preprocess Text → Performs cleaning operations on text, like removal of stop-words, case normalization
- Word2Vector → Converts words to values for use in NLP tasks, like recommender, named entity recognition, machine translation

Multiclass Classification

Answers complex questions with multiple possible answers

- Multiclass Logistic Regression → Fast training times, linear model
- Multiclass Neural Network → Accuracy, long training times
- Multiclass Decision Forest → Accuracy, fast training times
- One-vs-All Multiclass → Depends on the two-class classifier
- One-vs-One Multiclass → Depends on binary classifier, less sensitive to an imbalanced dataset with larger complexity
- Multiclass Boosted Decision Tree → Non-parametric, fast training times and scalable

Regression

Makes forecasts by estimating the relationship between values

- Fast Forest Quantile Regression → Predicts a distribution
- Poisson Regression → Predicts event counts
- Linear Regression → Fast training, linear model
- Bayesian Linear Regression → Linear model, small data sets
- Decision Forest Regression → Accurate, fast training times
- Neural Network Regression → Accurate, long training times
- Boosted Decision Tree Regression → Accurate, fast training times, large memory footprint

Two-Class Classification

Answers simple two-choice questions, like yes or no, true or false

- Two-Class Support Vector Machine → Under 100 features, linear model
- Two-Class Averaged Perceptron → Fast training, linear model
- Two-Class Decision Forest → Accurate, fast training
- Two-Class Logistic Regression → Fast training, linear model.
- Two-Class Boosted Decision Tree → Accurate, fast training. large memory footprint
- Two-Class Neural Network → Accurate, long training times

Anomaly Detection

Identifies and predicts rare or unusual data points

- One Class SVM → Under 100 features, aggressive boundary
- PCA-Based Anomaly Detection → Fast training times

Recommenders

Predicts what someone will be interested in

- Use the Train Wide & Deep Recommender module → Hybrid recommender, both collaborative filtering and content-based approach
- SVD Recommender → Collaborative filtering, better performance with lower cost by reducing dimensionality

Image Classification

Classifies images with popular networks

- ResNet → Modern deep learning neural
- PCA-Based Anomaly Detection → Network

Clustering

Separates similar data points into intuitive groups

- K-Means → Unsupervised learning

Roadmap to **Machine Learning**

Basics

- Choose Python or R
- Learn basic data types, variables

Math & Stats

- Study probability, linear algebra
- Learn calculus basics

Grasp ML Concepts

- Supervised vs. unsupervised
- Regression, SVM, K-means

ML Libraries

- Master NumPy, Pandas, Matplotlib
- Explore Scikit-learn, TensorFlow

Advanced Topics

- Deep learning (CNNs, RNNs)
- Learn NLP, reinforcement learning

Model Evaluation

- Use precision, recall, F1
- Apply cross-validation techniques

Practical Applications

- Work on Kaggle, GitHub
- Solve real-world ML problems

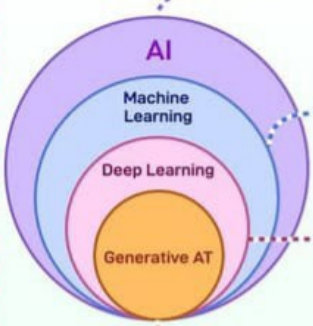
Collaborations

- Engage with the ML community
- Create a portfolio of projects

Continuous Learning

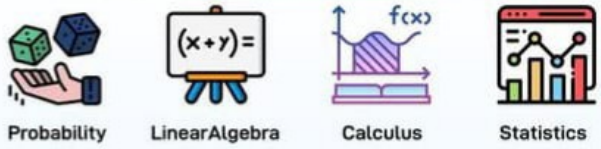
- Stay updated with courses
- Follow ML experts on YouTube

1 What is Generative AI



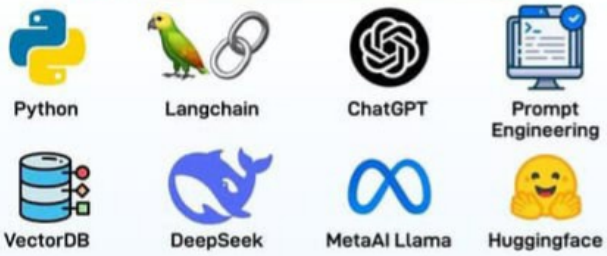
- Broad field fotused on creating mochines capable of doing tasks that need human intelligence
- Subset of AI involves algorithms that enable computers to learn from experience
- Subset of ML based on artificial neural networks where algorithms learn to identify patterns
- GenAI is a type of AI that creates new content based on what it has learned from existing content

2 Important Concepts

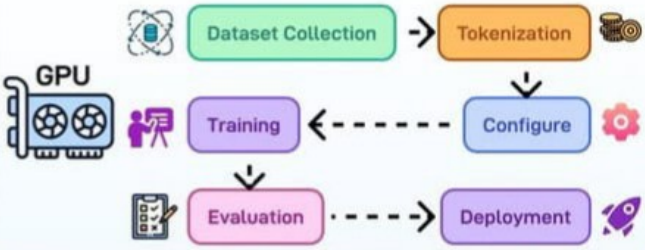


3 Foundation Models

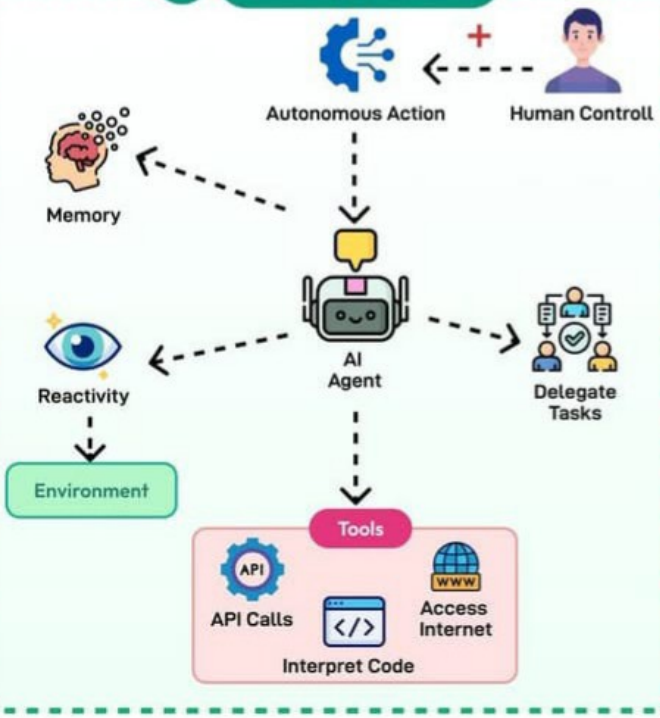
4 GenAI Development Stack



5 Training a Foundation Model



6 Building AI Agents



7 GenAI Models for Computer Vision

8 GenAI Learning Resources

Roadmap To Learn Machine Learning

Mathematics

- Probability
- Discrete
- Statistics

Programming

- Python
- R

Databas

- MySQL
- MongoDB

ML Algorithms

- Linear Logistic Regression
- KNN
- K-means
- Random forest & more!

Machine Learning

- ML Libraries and Non- ML Libraries

Machine Learning

- Scikit learn Supervised learning
- Unsupervised learning
- Reinforcement learning

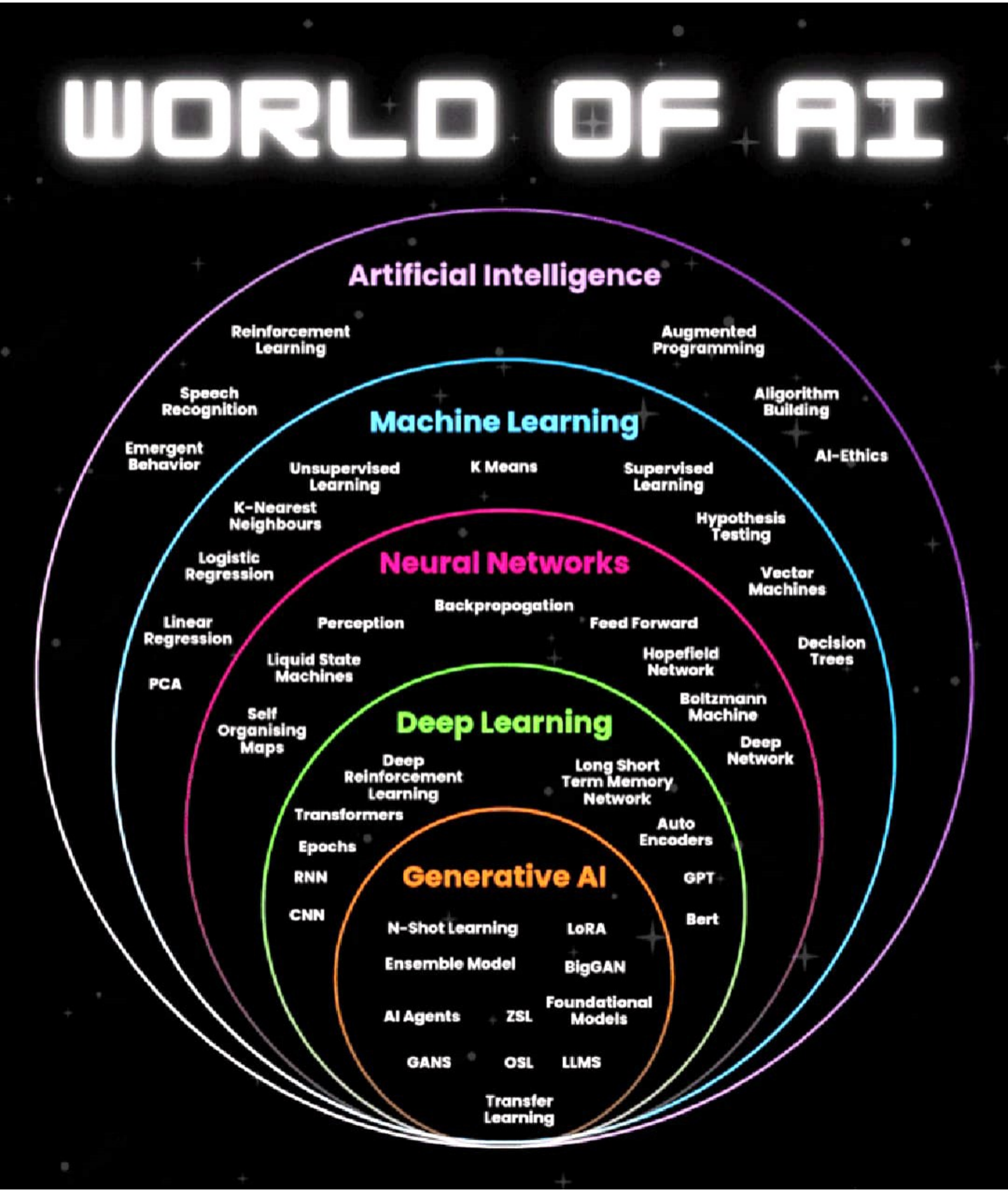
Deep Learning

- TensorFlow
- Keras
- Neural Networks
- CNN, RNN, GAN, LSTM

Data Visualization Tools

- Tableau
- Qlikview
- PowerBI

ML Engineer


WORLD OF AI
Artificial Intelligence
Reinforcement Learning
Augmented Programming
Speech Recognition
Aligorithm Building
Machine Learning
Emergent Behavior
AI-Ethics
Unsupervised Learning
K Means
Supervised Learning
K-Nearest Neighbours
Hypothesis Testing
Logistic Regression
Neural Networks
Vector Machines
Backpropogation
Linear Regression
Perception
Feed Forward
Decision Trees
Liquid State Machines
Hopefield Network
PCA
Boltzmann Machine
Self Organising Maps
Deep Learning
Deep Network
Deep Reinforcement Learning
Long Short Term Memory Network
Transformers
Auto Encoders
Epochs
RNN
Generative AI
GPT
CNN
Bert
N-Shot Learning
LoRA
Ensemble Model
BigGAN
AI Agents
ZSL
Foundational Models
GANS
OSL
LLMS
Transfer Learning