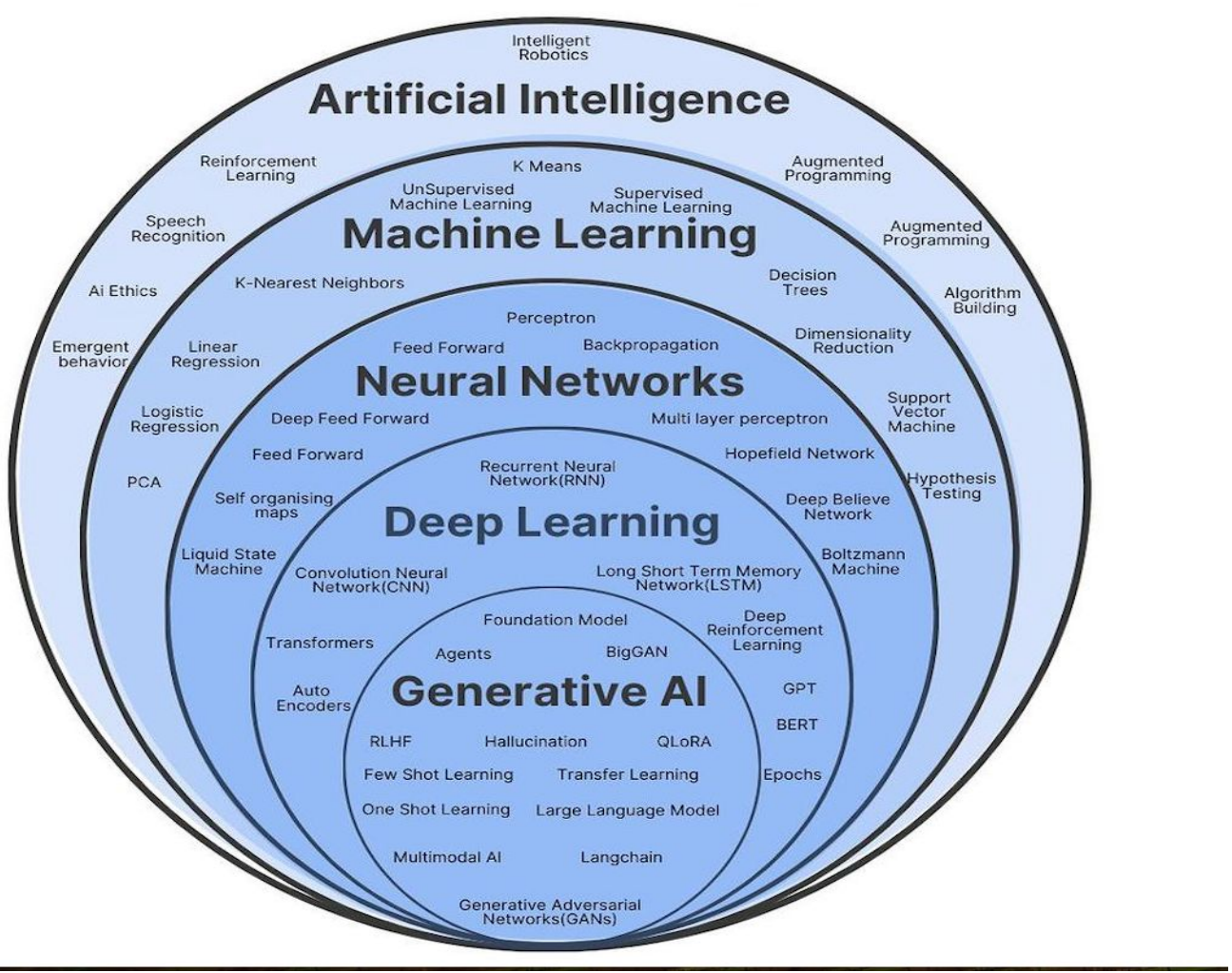
Machine Learning Terminologies

Week 02

Instructor: Dr. Anam Qureshi


Intelligent Robotics
Artificial Intelligence
Reinforcement Learning
K Means
Augmented Programming
UnSupervised Machine Learning
Supervised Machine Learning
Speech Recognition
Machine Learning
Augmented Programming
Ai Ethics
K-Nearest Neighbors
Decision Trees
Algorithm Building
Perceptron
Emergent behavior
Linear Regression
Feed Forward
Backpropagation
Dimensionality Reduction
Neural Networks
Logistic Regression
Deep Feed Forward
Multi layer perceptron
Support Vector Machine
Feed Forward
Hopefield Network
Recurrent Neural Network(RNN)
PCA
Hypothesis Testing
Self organising maps
Deep Believe Network
Deep Learning
Liquid State Machine
Convolution Neural Network(CNN)
Long Short Term Memory Network(LSTM)
Boltzmann Machine
Foundation Model
Deep Reinforcement Learning
Transformers
Agents
BigGAN
Auto Encoders
Generative AI
GPT
BERT
RLHF
Hallucination
QLoRA
Few Shot Learning
Transfer Learning
Epochs
One Shot Learning
Large Language Model
Multimodal AI
Langchain
Generative Adversarial Networks(GANs)

Key Terminologies

- **Dataset**: A collection of data used for training and testing a model.
- **Feature**: An individual measurable property or characteristic of data.
- **Label**: The output or target value in supervised learning.
- **Model**: The mathematical representation of a problem.
- **Inference**: The process of making predictions using a trained model. Happens after deployment.
- **Hypothesis**: A formal representation or function proposed by the learning algorithm to approximate the true relationship between input features and output labels

Types of Learning

- **Supervised Learning**:
  - Input data is labeled.
  - Example: Predicting house prices.
- **Unsupervised Learning**:
  - Input data is not labeled.
  - Example: Customer segmentation.
- **Reinforcement Learning**:
  - Learning by interacting with an environment.
  - Example: Game playing bots.

Training, Testing, and Validation

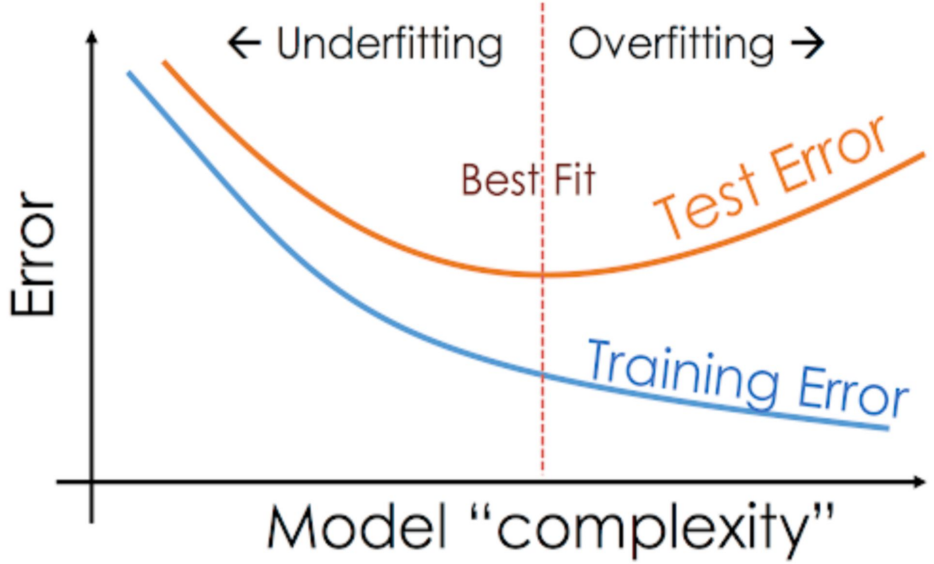
- **Training Set**:
  - The subset of data used to train the model.
- **Validation Set**:
  - Used to tune model parameters.
- **Test Set**:
  - Evaluates the final performance of the model.

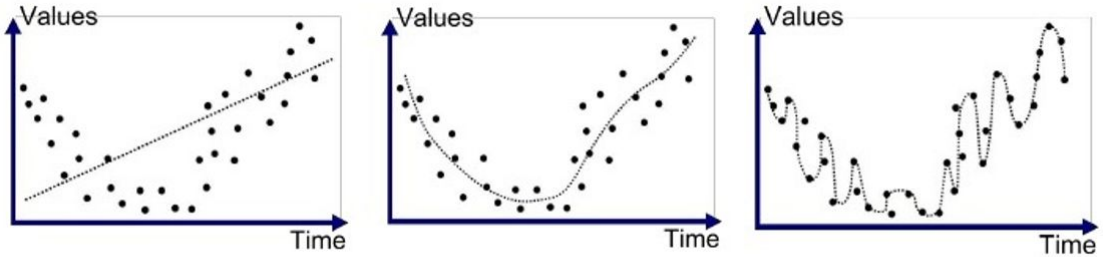
Testing vs. Inference:

- **Testing**: Measures model performance on a test dataset during development.
- **Inference**: Applies the trained model to real-world, unseen data after deployment.

Overfitting and Underfitting

- **Overfitting**:
  - Model performs well on training data but poorly on unseen data.
  - Example: Memorizing data.
- **Underfitting**:
  - Model is too simple to capture the underlying patterns.
  - Example: High bias.

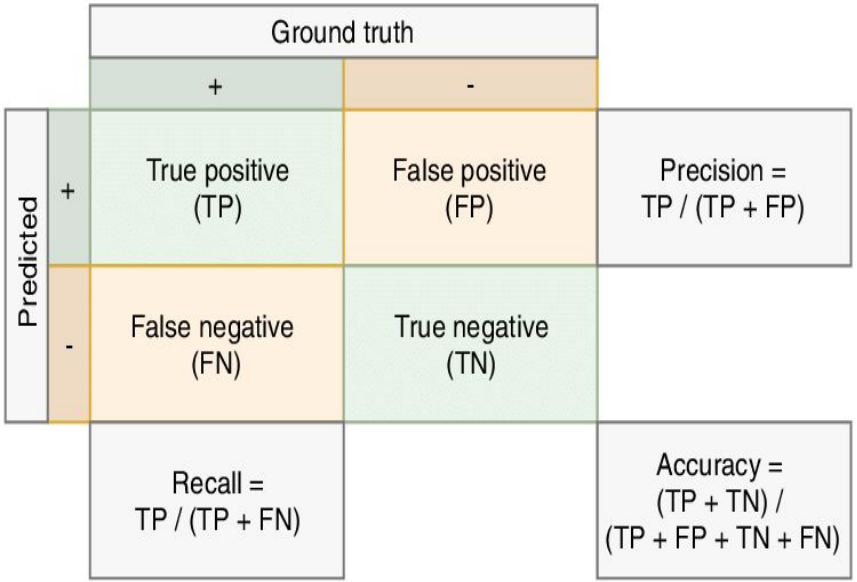




Underfitted | Good Fit/Robust | Overfitted

Key Performance Metrics

- **Accuracy**: Percentage of correct predictions.
- **Precision**: True positives / (True positives + False positives).
- **Recall**: True positives / (True positives + False negatives).
- **F1 Score**: Harmonic mean of precision and recall.
- **Loss or Error Function**: A mathematical function that quantifies the difference between predicted and actual values. Common examples include Mean Squared Error (MSE) and Cross-Entropy Loss.



Algorithms Overview

- **Regression**: Predicts continuous values.
- **Classification**: Predicts discrete categories.
- **Clustering**: Groups similar data points.
- **Dimensionality Reduction**: Reduces the number of features.

Some common terminologies

- **Hyperparameters**: Parameters set before training.
- **Weights**: Values adjusted during training.
- **Learning Rate**: Controls how much weights are updated.
- **Epoch**: One complete pass through the training dataset.