Machine Learning Fundamentals

A Quick Introduction

This document provides a brief overview of key machine learning concepts that are commonly used in modern AI applications.

Contents:
1. Supervised Learning
2. Neural Networks
3. Transformers

Machine learning is a subset of artificial intelligence that enables systems to learn and improve from experience without being explicitly programmed. It focuses on developing computer programs that can access data and use it to learn for themselves.

1. Supervised Learning

Supervised learning is a type of machine learning where the model is trained on labeled data. The algorithm learns from the training dataset and makes predictions based on that learning.

Key concepts:
- Training data: Input-output pairs used to train the model
- Features: Input variables used to make predictions
- Labels: Output variables the model learns to predict
- Loss function: Measures the difference between predictions and actual values

Common algorithms:
- Linear Regression: For continuous output prediction
- Logistic Regression: For binary classification
- Decision Trees: For both classification and regression
- Support Vector Machines: For classification with margin optimization

Example: Email spam detection
- Features: Word frequencies, sender information, email metadata
- Label: Spam (1) or Not Spam (0)

2. Neural Networks

Neural networks are computing systems inspired by biological neural networks. They consist of interconnected nodes (neurons) organized in layers.

Architecture:
- Input Layer: Receives the raw input data
- Hidden Layers: Process information through weighted connections
- Output Layer: Produces the final prediction

Key components:
- Weights: Parameters that are learned during training
- Activation Functions: Non-linear functions (ReLU, Sigmoid, Tanh)
- Backpropagation: Algorithm for computing gradients
- Gradient Descent: Optimization algorithm for updating weights

Deep Learning extends neural networks with many hidden layers, enabling the learning of complex hierarchical representations from data.

Applications:
- Image classification and object detection
- Natural language processing
- Speech recognition
- Game playing (AlphaGo, etc.)

3. Transformers

Transformers are a neural network architecture introduced in the paper "Attention Is All You Need" (Vaswani et al., 2017). They have become the foundation for modern large language models.

Key innovation: Self-Attention Mechanism
- Allows the model to weigh the importance of different parts of the input
- Enables parallel processing (unlike RNNs)
- Captures long-range dependencies effectively

Architecture components:
- Multi-Head Attention: Multiple attention mechanisms in parallel
- Positional Encoding: Adds position information to embeddings
- Feed-Forward Networks: Process attention outputs
- Layer Normalization: Stabilizes training

Popular Transformer models:
- BERT: Bidirectional encoding for understanding
- GPT: Generative pre-training for text generation
- T5: Text-to-text transfer transformer

Transformers have revolutionized NLP and are now being applied to computer vision (Vision Transformers), audio processing, and more.