AI Basics - Study Notes

What is Artificial Intelligence?

AI refers to computer systems that can perform tasks typically requiring human intelligence, such as learning, reasoning, problem-solving, perception, and language understanding.

Types of AI

Narrow AI (Weak AI)

- Designed for specific tasks
- Current AI systems fall into this category
- Examples: voice assistants, recommendation systems, image recognition

General AI (Strong AI)

- Hypothetical AI with human-level intelligence across all domains
- Can understand, learn, and apply knowledge like humans
- Does not currently exist

Superintelligence

- Theoretical AI that surpasses human intelligence
- Still in the realm of speculation and future possibilities

Key AI Technologies

Machine Learning (ML)

- Subset of AI where systems learn from data without explicit programming
- Algorithms improve performance through experience
- Foundation for most modern AI applications

Deep Learning

- Subset of ML using artificial neural networks
- Mimics how human brain processes information
- Excellent for pattern recognition in images, speech, and text

Natural Language Processing (NLP)

- Enables computers to understand and generate human language

- Powers chatbots, translation services, and text analysis
- Includes sentiment analysis, text summarization, and language generation

Computer Vision

- Allows machines to interpret and understand visual information
- Applications in facial recognition, medical imaging, autonomous vehicles
- Processes images and videos to extract meaningful information

Common AI Applications

Everyday Use

- Search engines (Google, Bing)
- Social media algorithms (Facebook, Instagram feeds)
- Streaming recommendations (Netflix, Spotify)
- Virtual assistants (Siri, Alexa, Google Assistant)

Professional Applications

- Medical diagnosis and drug discovery
- Financial fraud detection and trading
- Autonomous vehicles and transportation
- Manufacturing and quality control
- Customer service chatbots

How AI Systems Learn

Supervised Learning

- Learns from labeled training data
- Examples: email spam detection, image classification
- Requires input-output pairs for training

Unsupervised Learning

- Finds patterns in data without labels
- Examples: customer segmentation, anomaly detection
- Discovers hidden structures in data

Reinforcement Learning

- Learns through trial and error with rewards/penalties
- Examples: game playing (AlphaGo), robotics
- Agent learns optimal actions through interaction with environment

Key Concepts

Algorithm

- Set of rules or instructions for solving problems
- Foundation of all AI systems

Training Data

- Information used to teach AI systems
- Quality and quantity of data significantly impacts AI performance

Neural Networks

- Computing systems inspired by biological neural networks
- Consist of interconnected nodes (neurons) that process information

Bias in AI

- Unfair or prejudiced outcomes from AI systems
- Often reflects biases present in training data
- Important consideration for ethical AI development

Current Limitations

Data Dependency

- AI systems require large amounts of quality data
- Performance degrades with poor or insufficient data

Lack of Common Sense

- AI struggles with context and real-world understanding
- Can make obvious mistakes humans wouldn't make

Explainability

- Many AI systems are "black boxes"
- Difficult to understand how they reach decisions

Generalization

- AI systems may not perform well on data different from training examples
- Can struggle with novel situations

Future Directions

Emerging Trends

- More efficient and smaller AI models
- Improved human-AI collaboration
- Better AI safety and alignment
- Integration with robotics and IoT devices

Potential Impact

- Transformation of various industries
- Changes in employment and job markets
- Enhanced scientific research and discovery
- Improved healthcare and education