AI Literacy – Study Notes

1. Defining AI Literacy

AI literacy refers to the foundational knowledge and critical understanding required to engage with artificial intelligence responsibly and effectively. Just as traditional literacy enabled societies to function in a written culture, AI literacy equips individuals to function in a digital society where intelligent systems influence decision-making, information access, and daily interactions.

Key point: AI literacy does not require becoming a data scientist. Instead, it emphasizes understanding the principles, limitations, and social implications of AI.

2. Everyday Applications of AI

Artificial intelligence systems already permeate everyday life, often invisibly:

- **Information Retrieval**: Search engines use ranking algorithms to determine the order of results.

- **Entertainment and Media**: Platforms such as Netflix and Spotify recommend content based on viewing or listening history.
- **E-commerce**: Online retailers employ recommendation systems to suggest additional products.
- **Social Media**: Recommendation feeds, image filters, and content moderation are powered by AI.
- **Digital Assistants**: Systems such as Siri, Alexa, and Google Assistant rely on natural language processing.

Reflection: Most individuals already interact with AI multiple times per day without explicit awareness.

3. The AI Ecosystem

AI can be understood as a hierarchical system:

- **Artificial Intelligence (AI)**: Broad field concerned with building systems that exhibit behaviors associated with intelligence.
- **Machine Learning (ML)**: Subfield in which systems improve performance through data-driven learning.
- **Deep Learning (DL)**: A subset of ML employing multi-layered neural networks for complex tasks involving text, images, or speech.
- **Generative AI**: Models designed to produce original outputs, including text, images, code, or music.

Mnemonic: AI is the umbrella; ML, DL, and Generative AI represent increasingly specialized approaches.

4. Types of Intelligence and AI Counterparts

Human intelligence is multi-faceted. AI

attempts to replicate aspects of it but not the whole spectrum:

- **Creative intelligence**: Human creativity mirrored in systems such as DALL·E or MidJourney.
- **Kinetic intelligence**: Embodied skill reflected in robotics or drone systems.
- **Linguistic intelligence**: Paralleled by natural language models such as ChatGPT.
- **Mathematical-logical intelligence**: Implemented in scientific computing and predictive analytics.

Observation: AI simulates discrete dimensions of intelligence but lacks holistic cognition or emotional depth.

5. Human Learning vs. Machine Learning

- **Human learning**: Based on limited examples, social interaction, and contextual understanding.
- **Machine learning**: Requires vast datasets,

algorithmic training, and numerical optimization.

Example: A child may recognize cats after seeing only a few examples. A machine learning model may require thousands of labeled images.

6. Pattern Recognition as the Core of AI

Pattern recognition underlies most AI applications. While humans are adept at intuitive pattern recognition with limited data, machines excel in analyzing massive and complex datasets.

- **Simple patterns**: Numeric sequences (e.g., squares of numbers).
- **Complex patterns**: Irregularities in financial transactions or large-scale genomic data.

Conclusion: AI's comparative advantage lies in scale and computational intensity.

7. Case Study: Loan Prediction and Bias

Loan approval systems illustrate the risks of biased AI. Features such as age, salary, and geographic region may be appropriate predictors. However, if gender or race is disproportionately weighted due to biased historical data, the system perpetuates discrimination.

Notable incident: Apple's credit card algorithm faced criticism in 2019 for allegedly assigning lower credit limits to women compared to men with similar financial profiles.

Lesson: AI systems inherit biases from data and design. Critical oversight is required to ensure fairness.

8. Why Deep Learning Matters

Traditional machine learning performs effectively with structured datasets (numerical or tabular). Deep learning, however, is suited to unstructured data such as images, video, or natural language.

Analogy: Machine learning is akin to a calculator handling defined problems, whereas deep learning is closer to the brain's ability to interpret sensory information.

9. Neural Networks

Neural networks are computational models inspired by the human brain.

- **Perceptron**: The basic unit, which processes weighted inputs and produces an output.
- **Layers**: Multiple perceptrons organized into layers create complex networks capable of sophisticated tasks.
- **Scale**: Modern generative models contain billions of parameters, making them powerful but also opaque.

Challenge: As networks grow in scale, they become increasingly difficult to interpret or explain.

10. The Black Box Problem

Many deep learning systems operate as “black boxes”: users see inputs and outputs but not the reasoning pathway.

Concerns include:

- Inability to explain why a specific decision was made.
- Risk of perpetuating hidden biases.
- Difficulty in auditing fairness and accountability.

Approaches to mitigation:

- **Explainable AI (XAI)** methods such as LIME and SHAP.
- **Human oversight** in high-stakes decision-making.
- Analysis of feature importance to reveal influential factors.

11. Discriminative vs. Generative AI

- **Discriminative AI**: Focused on classification tasks, such as distinguishing spam from legitimate emails.
- **Generative AI**: Produces novel content, such as generating an essay, image, or software code.

Trend: The field has shifted from primarily discriminative systems in the 2010s to widespread generative applications in the 2020s.

12. Challenges and Ethical Considerations

Key issues in contemporary AI deployment include:

- **Data Privacy**: The trade-off between convenience and personal data exposure.
- **Bias and Fairness**: The risk of embedding social inequalities in automated systems.
- **Labor Market Disruption**: Automation displacing some jobs while creating new categories requiring reskilling.

- **Misinformation**: Deepfake technologies challenging truth and authenticity in media.

Principle: Ethical design must precede technical implementation to ensure AI benefits society equitably.

13. Current and Future States of AI

- **Narrow AI (current)**: Systems specialized in specific tasks (e.g., facial recognition, chatbots).
- **General AI (aspirational)**: A system with human-level cognitive flexibility across diverse tasks.
- **Superintelligence (speculative)**: A system surpassing collective human intelligence, currently theoretical.

Status: Contemporary AI remains within the boundaries of narrow intelligence.

14. Summary of Key Insights

- AI primarily functions through pattern recognition at scale.
- The quality of input data determines the fairness and accuracy of outputs.
- Current AI systems are specialized rather than conscious or general.
- Ethical safeguards and human oversight are non-negotiable.
- The most productive path forward lies in collaboration between human expertise and AI capabilities.