Artificial Intelligence (AI): A Brief Overview

1. What is AI?

Artificial Intelligence refers to the development of computer systems that can perform tasks that typically require human intelligence. These tasks include reasoning, learning, problem■solving, perception, and language understanding.

2. Goals of AI

• Create intelligent agents capable of autonomous decision■making

• Develop systems that can reason, learn, and adapt

• Enable machines to understand and generate human language

• Build systems that can perceive and interpret the world

3. Types of AI

• Narrow AI – Systems designed to perform specific tasks

• General AI – Hypothetical systems capable of human■level intelligence

• Superintelligent AI – Intelligence surpassing human capabilities

4. Machine Learning

A subset of AI where systems learn patterns from data rather than being explicitly programmed. Types include supervised learning, unsupervised learning, and reinforcement learning.

5. Deep Learning

A specialized form of machine learning using neural networks with many layers. It is widely used in image recognition, language models, and speech processing.

6. Applications of AI

• Self■driving cars

• Recommendation systems

• Virtual assistants

• Medical diagnosis

• Fraud detection

7. Challenges in AI

• Data quality and availability

• Bias and fairness

• Explainability of AI decisions

• Safety and alignment concerns

8. The Future of AI

AI is expected to advance in reasoning, planning, creativity, and generalization. Research is ongoing in building aligned and safe AI systems.

This document is designed for testing Retrieval■Augmented Generation (RAG) systems. All questions must be answered strictly based on this PDF content.