Introduction to Artificial Intelligence

What is Artificial Intelligence?

Artificial Intelligence (AI) is a branch of computer science that aims to create intelligent machines capable of perf

Machine Learning is a subset of AI that enables computers to learn and improve from experience without being e

Deep Learning is a further subset of machine learning that uses neural networks with multiple layers to model an

Key Concepts:

1. Supervised Learning: The algorithm learns from labeled training data, making predictions or decisions based o

2. Unsupervised Learning: The algorithm finds hidden patterns in data without labeled examples, such as cluster

3. Reinforcement Learning: An agent learns to make decisions by interacting with an environment and receiving

Natural Language Processing

Natural Language Processing (NLP) is a field of AI that focuses on the interaction between computers and huma

Key NLP Tasks:

- Text Classification: Categorizing text into predefined classes
- Named Entity Recognition: Identifying entities like names, locations, dates
- Sentiment Analysis: Determining the emotional tone of text
- Machine Translation: Translating text from one language to another
- Question Answering: Answering questions based on context

Modern NLP uses transformer architectures, such as BERT and GPT, which have achieved remarkable results i

Embeddings are vector representations of words or sentences that capture semantic meaning. They allow mach

Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation, or RAG, is a technique that combines information retrieval with generative AI m

How RAG Works:

1. Document Ingestion: Documents are processed, chunked, and converted into embeddings that are stored in a

2. Query Processing: When a user asks a question, the query is converted into an embedding and used to searc

3. Context Retrieval: The most relevant chunks are retrieved from the vector database based on semantic simila

4. Response Generation: The retrieved context is provided to the language model along with the user's question

Benefits of RAG:

- Up-to-date information: Can access current information not in training data
- Source citations: Can cite specific documents and pages
- Reduced hallucinations: Grounded in actual documents
- Domain-specific knowledge: Can be customized for specific domains

This approach is particularly useful for building chatbots that can answer questions about specific documents, su