

Sample PDF for RAG Chatbot Testing

Chapter 1: Introduction to Machine Learning

Machine learning is a subset of artificial intelligence that focuses on the development of algorithms and statistical models that enable computer systems to improve their performance on a specific task through experience.

Key concepts in machine learning include:

- Supervised learning: Learning from labeled training data
- Unsupervised learning: Finding patterns in data without labels
- Reinforcement learning: Learning through interaction with an environment

Chapter 2: Natural Language Processing

Natural Language Processing (NLP) is a branch of artificial intelligence that deals with the interaction between computers and humans through natural language.

Common NLP tasks include:

- Text classification
- Named entity recognition
- Sentiment analysis
- Machine translation
- Question answering systems

Chapter 3: Vector Databases and Embeddings

Vector databases are specialized databases designed to store and query high-dimensional vectors efficiently. They are particularly useful in machine learning applications where data is represented as dense numerical vectors.

Chapter 4: Retrieval-Augmented Generation (RAG)

RAG is a framework that combines retrieval-based and generation-based approaches for natural language processing tasks. It works by first retrieving relevant information from a knowledge base using vector similarity search.