Project Overview: AiMlClarityDemo

Domain: The project operates within the Artificial Intelligence and Machine Learning domain, specifically focused on building intelligent applications with local vector search capabilities.

Purpose: AiMlClarityDemo serves as a foundational framework for experimenting with Retrieval-Augmented Generation (RAG) workflows in a .NET ecosystem. It allows users to ingest documents and perform semantic search and summarization using vector databases.

Key Features: - Local vector store integration using SQLite. - Document ingestion and chunking. - Vectorization using embedding models. - Fast semantic retrieval. - Plug-and-play architecture for integrating large language models (LLMs).

Technologies Used: - C# .NET 9.0 - ASP.NET Core Web API - SQLite as a local vector database - Microsoft Semantic Kernel for RAG orchestration - Azure AI for embeddings (optional) - Newtonsoft.Json and System.Text.Json for serialization - Swagger/OpenAPI for API testing

Architecture Notes: - The project follows a modular structure with separate components for ingestion, storage, search, and presentation. - Ingested documents are split into meaningful chunks and stored in a local SQLite database, where each chunk is associated with a vector embedding. - The solution is designed to allow offline or on-device usage scenarios.

Typical Use Case: - Ingest PDF, TXT, or DOCX files - Query using natural language - Receive context-rich, RAG-powered responses from the LLM

Future Extensions: - Integration with Azure Cosmos DB for cloud-scale vector search - Role-based access control (RBAC) for document permissions - UI frontend for better user experience - Integration with OpenAI or other LLM APIs

Target Audience: - AI engineers, software developers, researchers, and data scientists experimenting with hybrid search and retrieval-enhanced generation workflows

Authors & Contributors: - Ivan Vydrin and contributors from the AiMlClarity initiative

License: - MIT License

This document is intended to support RAG (Retrieval-Augmented Generation) testing and evaluation.