Group Members: Krishna Dubey (211020425), Pankaj Kumar Giri (211020430), Nayandeep (211000033)

STOCK SAGE: AI-DRIVEN EQUITY RESEARCH TOOL

PROBLEM DEFINITION:

Equity research analysts often face the challenge of sifting through vast amounts of financial news articles, earnings reports, and other relevant documents to gather insights and make informed decisions. The current manual approach to analyzing these documents is time-consuming, tedious, and prone to human error. Analysts require a tool that can efficiently process multiple articles, extract relevant information, and provide concise answers or summaries based on their queries.

OBJECTIVES:

- Automate Information Retrieval: Develop an AI-driven tool that can automatically retrieve, process, and analyze large volumes of financial news articles and reports.
- Enhance Decision-Making: Enable equity research analysts to make faster and more informed decisions by providing them with accurate and relevant data from trusted sources.
- Reduce Costs: Minimize the manual effort involved in searching for and analyzing information, thereby reducing the time and cost associated with equity research.
- Improve Accuracy: Ensure that the extracted information is precise and contextually relevant, reducing the risk of misinterpretation.

RESEARCH OUTCOMES AND DELIVERABLES:

- **AI-Driven Research Tool**: A web-based application capable of retrieving, analyzing, and summarizing data from financial news articles and reports using AI techniques.
- **Customizable Knowledge Base**: A dynamic knowledge base that aggregates information from multiple trusted sources, allowing for targeted and efficient querying.
- **Semantic Search Functionality**: Implementation of semantic search to understand context, allowing the tool to differentiate between similar terms based on their use (e.g., "Apple" as a fruit vs. "Apple" as a company).
- **Optimized Cost Structure**: Smart chunking and retrieval methods to minimize API call costs when using large language models (LLMs).
- **User-Friendly Interface**: A simple, intuitive interface built with Streamlit for initial deployment, followed by more robust, scalable architecture for long-term use.

Group Members: Krishna Dubey (211020425), Pankaj Kumar Giri (211020430), Nayandeep (211000033)

Week 1-2: Planning & Initial Setup

- Task: Finalize project requirements, set up the development environment, and gather necessary resources (e.g., APIs, libraries).
- Expected Outcome: Project plan approval, tools and environment ready for development.

Week 3-4: POC Development

- Task 1: Implement document loader to ingest news articles from URLs.
- Task 2: Develop chunking algorithm to split articles into relevant sections.
- Task 3: Build a basic chatbot interface using Streamlit to answer queries.
- Expected Outcome: Functional POC with basic query-answer capability.

Week 5-6: POC Testing & Validation

- Task 1: Test the POC with sample financial articles to ensure correct chunking and data retrieval.
- Task 2: Validate the accuracy of the responses generated by the chatbot.
- Exp. Outcome: Refined POC with validated results, ready for further enhancement.

Week 7-8: Semantic Search Integration

- Task 1: Implement semantic search using word embeddings and vector databases.
- Task 2: Integrate with the existing chatbot to improve query accuracy.
- Exp. Outcome: Enhanced tool with context-aware query processing.

Week 9-10: API Optimization & Cost Management

- Task 1: Optimize chunk selection to minimize OpenAI API call costs.
- Task 2: Test different chunking strategies to find the most cost-effective approach.
- Exp. Outcome: Optimized API usage reducing operational costs.

Week 11: Feature Enhancement

- Task 1: Add summarization and aggregated response features to the chatbot.
- Task 2: Extend the tool's capability to handle multiple articles simultaneously.
- Exp. Outcome: Feature-rich tool with advanced querying options.

Week 12: Short-Term Deployment & Testing

- Task 1: Deploy the enhanced tool for internal testing with a broader user base.
- Task 2: Collect feedback and make final adjustments before wider release.
- Exp. Outcome: Ready-to-deploy version of the tool, prepared for the next phase of development.