AI Agent for Automated Financial Insights from the Malawi Stock Exchange

Overview

We propose the development of an **AI-powered data extraction and analysis agent** designed to monitor, collect, and interpret financial data from the **Malawi Stock Exchange (MSE)**.
The agent will automate the process of **scraping financial PDFs**, **storing structured data**, and **analyzing both quantitative and qualitative indicators** to estimate a company's fair stock price and future performance trends.

Key Features

1. **Automated Data Scraping**

- The agent will continuously monitor the **Malawi Stock Exchange website** MSE website for newly uploaded PDFs such as annual reports, trading summaries, and financial disclosures.
- Using advanced **document parsing (OCR + NLP)**, it will extract tables of financial data, CEO statements, and key performance indicators (KPIs).

2. **Structured Data Storage**

- Extracted data will be cleaned, normalized, and stored in a **PostgresSQL database** as the system's primary backend, this is server-hosted.
- Parallel exports in **Excel (.csv)** format will be generated for end users, enabling seamless financial analysis and visualization for local users. Ideally, only at query time - so Python Django or Flask could be used to serve the data a bit more efficiently without much overhead.

3. **Sentiment and Confidence Analysis**

- The agent will scrape **CEO statements**, **LinkedIn posts**, and **news articles** to assess leadership tone and sentiment regarding the company's financial outlook. Primarily, we want to look at Zodiak, NyasaTimes and Malawi24.
- Using **NLP techniques** (e.g., sentiment analysis, topic modeling), the agent will assign a **market-side confidence metric** to each company based on leadership communication and market sentiment.

4. **Quantitative & Qualitative Insights**

- Merging financial ratios, growth trends, and sentiment scores, the agent will deliver both **quantitative** and **qualitative evaluations** of listed companies.
- These insights will be used to estimate:
    - **Fair stock prices**
    - **Volatility forecasts**
    - **Predictive indicators** for investor decision-making

Technical Stack

Component	Technology
Web Scraping	Python (BeautifulSoup, Selenium, PyPDF, OCR APIs)
NLP & Sentiment Analysis	OpenAI GPT APIs, HuggingFace Transformers
Data Storage	PostgreSQL / MySQL
Export Layer	Pandas → Excel (.csv)
Visualization (Optional)	Streamlit / Power BI Connector
Hosting	AWS / Azure with scheduled crawlers

Impact

This AI agent bridges the **data accessibility gap** in the Malawian financial ecosystem, empowering investors, analysts, and regulators with:

- Reliable, real-time data collection
- Enhanced transparency in company reporting
- Predictive insights for **fair stock valuation**

By combining **machine intelligence** with **financial domain expertise**, this project aims to set a new standard for **AI-driven financial analysis in emerging markets**.

Next Steps

1. Develop a prototype for PDF scraping and table extraction.
2. Set up SQL + CSV storage pipelines.
3. Integrate sentiment analysis and confidence scoring.
4. Test predictive models for fair stock pricing.

Pitch Summary:

> *An AI agent that automatically scrapes, analyzes, and interprets financial and qualitative data from the Malawi Stock Exchange to deliver deep insights, improve transparency, and predict company valuation metrics.*

Glossary:

- **OCR**: Optical Character Recognition, technology to convert different types of documents into editable and searchable data.
- **NLP**: Natural Language Processing, a field of AI focused on the interaction between computers and human language.
- **KPI**: Key Performance Indicator, a measurable value that demonstrates how effectively a company is achieving key business objectives.

- **CSV**: Comma-Separated Values, a simple file format used to store tabular data.
- **API**: Application Programming Interface, a set of rules that allows different software entities to communicate with each other.