Architecture de Données pour un Assistant IA Conseiller Financier

# 🎯 Objectif

Construire une architecture robuste, modulaire et évolutive permettant à une IA d'analyser en continu la situation financière d'un utilisateur, de surveiller le marché et d’émettre des recommandations personnalisées.

# 🧩 1. Types de Données à Intégrer

## A. Données Client

- Transactions bancaires : Dépenses, virements, prélèvements

- Comptes & Épargne : Solde, type, fréquence

- Profil utilisateur : Revenus, âge, objectifs, charges

- Documents financiers : PDF, factures, relevés

## B. Données Marché

- Cours boursiers / crypto : Prix en temps réel / historique

- Indices économiques : Inflation, taux d’intérêt, PIB

- News financières : Flux RSS, API actu

- Données fiscales : Lois, barèmes

## C. Données Comportementales du Marché

- Investisseurs majeurs : Portefeuilles de fonds/insiders

- Volume anormal : Activité brutale sur un actif

- Réseaux sociaux / forums : Sentiment du public

- Décisions d'entreprise : Dividendes, rachats, fusions

# 🏗️ 2. Architecture Technique de Données

## 2.1. Collecte (Data Ingestion)

- Données bancaires client : API Open Banking (Plaid, Tink, Salt Edge)

- Documents : Upload + OCR (Tesseract, DocTR, LayoutLM)

- Données boursières : APIs temps réel (Alpha Vantage, Yahoo Finance, Binance)

- Indices économiques : Open Data scraping (INSEE, IMF, Eurostat)

- News & sentiment : NewsAPI, Reddit API, spaCy/BERT

- Activité marché : WhaleWisdom, SEC, Binance WebSocket

## 2.2. Stockage (Data Storage)

- Données client : PostgreSQL / MySQL

- Historique / OCR : MongoDB / Elasticsearch

- Flux : Kafka / S3 / BigQuery

- Index sémantiques : FAISS / Qdrant / Weaviate

## 2.3. Prétraitement (Cleaning & Normalisation)

- Nettoyage, normalisation, NLP, extraction PDF

## 2.4. Traitement et Analyse

- Agrégation, scoring, suivi portefeuille, clustering, anomalies

- Outils : Scikit-learn, XGBoost, Prophet, LSTM

## 2.5. Recommandation et IA Générative

- Recommandations : LightFM, RecBole

- Assistant IA : GPT, LangChain, RAG

- Prédictions : LSTM, Prophet

- Analyse de sentiment : BERT, FinBERT, Vader

## 2.6. Visualisation et Interaction Utilisateur

- Dashboard (Streamlit, Dash, Grafana, React + FastAPI)

# ✅ Conclusion

Cette architecture permet une vision complète du profil utilisateur, une compréhension du marché en temps réel, et des recommandations dynamiques et personnalisées, le tout avec une structure scalable et modulaire.