e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

ARTIFICIAL INTELLIGENCE IN FINANCE: COFFEE COMMODITY TRADING BIG DATA FOR INFORMED DECISION MAKING

Ms. Ayushi Gautam*[1], Prof. Rupali Maske*[2], Prof. Rutika Shah*[3]

*[1,2,3]Department Of Computer Engineering, Trinity College Of Engineering And Research, Pune, India.

DOI : https://www.doi.org/10.56726/IRJMETS63299

ABSTRACT

Coffee, the second largest soft commodity globally, can benefit from thorough analysis of both daily and historical market data to make better informed trading choices. Sophisticated ICT and data mining technologies have the potential to alter the operation of the trading market. The current systems face limitations like lack of complete data, inadequate documentation for storage, and the need for a scalable infrastructure for big data analytics, such as a data warehouse or data lake house. In order to tackle this problem, the paper showcases a coffee commodity trading big data warehouse design and implementation that can analyze crucial parameters to support decision-making. Initially, the system is able to autonomously gather data on coffee trading by collecting New York Arabica coffee futures prices from various global reports and financial data portals. Then, the ETL process is utilized to input data from coffee futures trading web scraping into the 3 layers data warehouse. In the end, the analytics system will identify and display specific important factors that impact coffee futures prices across various time frames and viewpoints. Therefore, we develop a model of a coffee trading data repository using the gathered information from January 2000 to October 2022 and present patterns in coffee futures prices using the acquired data for making well-informed decisions. The construction system can effectively handle and manage high amounts of transaction data. This document will serve as a useful reference and decision-making tool for coffee trading businesses and will help advance the creation of predictive models.

I. INTRODUCTION

Coffee has become the product with a high level of consumption worldwide in recent years, with over 2.25 billion cups of coffee consumed daily and playing an essential role in many countries. For example, in Vietnam, over the, years, coffee has been the main export and become the second most exported goods in export turnover after rice. Every year, coffee contributes up to 10% of the country's total export turnover. Trading coffee through an exchange is not a new type of transaction. However, traders and investors will still be exposed to many risks if they do not adequately understand the market and the trading methods, price trends, or market data analysis. Price fluctuations, exchange rates, and export prices of big exporting countries can affect coffee prices in domestic countries. Price fluctuation in world coffee trading and the changes in the exchange market should be monitored carefully in every session and compared to the trading historical periods. Coffee trading exchange data is presented as time-series data with on-stop fluctuations in each trading session. Forecasting the price trends to minimize the risk of investing can become a guaranteed technique in trading markets. In coffee commodity markets, the amount of coffee transaction data is generated, increasing in volume and variety over the years.

This allows data gathering, integration, and analysis demands on a scale not seen earlier. Some reports have been published and provided as coffee world trade statistics such as the report Coffee: World Markets and Trade , Coffee Market Report , Global Market Report: Coffee, World Bank Commodities Price Data (Pink Sheet) , etc. Typically, these reports will provide readers with up-to-date information on coffee world markets and various units with pre-calculated formulas. However, current systems face specific limitations, including insufficient data, inadequate documentation for storage, and the need for scalable infrastructure to facilitate big data analytics, such as a data warehouse or data lake house.. The data warehouse transforms and categorizes data using the Extract – Transform – Load (ETL) process from different data sources before importing it into the repository.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

II. LITERATURE SURVEY

The integration of fog computing, big data analysis, and optimization techniques can enhance decision-making in coffee commodity trading. Studies have shown that smart ant colony optimization can improve task offloading efficiency (Kishor & Chakarbarty, 2022), while big data analysis can extract valuable insights from large datasets (Le, 2022). Effective data management using ETL tools (Singh, 2022) and reliable data sources such as World Bank Commodities Price Data (World Bank, 2023) and International Coffee Organization reports (ICO, 2022) support accurate market forecasting. These findings imply improved efficiency, reduced costs, and better risk management in coffee trading. Future research directions include integrating AI, blockchain, and real-time monitoring systems to further enhance transparency and accuracy in coffee commodity trading.

The international coffee market involves complex supply chains, and risk management strategies mitigate market volatility. Market forecasting models predict price fluctuations, and sustainable practices ensure long-term coffee supply. Research addresses global coffee market dynamics, regional trends, and stakeholder collaboration.

Innovative technologies transform coffee trading, enhancing supply chain efficiency and sustainability. By exploring these themes, researchers and practitioners can optimize coffee commodity trading, ensuring a resilient and prosperous coffee industry.

III. RESEARCH METHODOLOGY

1. **Mixed-Methods Approach**: Combining quantitative (data analysis, simulation models) and qualitative (surveys, interviews, case studies) methods to explore the application of fog computing, big data analysis, and optimization techniques in coffee commodity trading.
2. **Data Collection:** Utilizing secondary data (World Bank, International Coffee Organization reports) and primary data (surveys, interviews with coffee traders, producers, and industry experts) to gather comprehensive insights.
3. **Data Analysis:** Employing quantitative techniques (descriptive statistics, inferential statistics, time-series analysis) and qualitative techniques (content analysis, thematic analysis) to analyze data and identify trends.
4. **Tools and Techniques:** Leveraging fog computing simulation models, big data analytics tools (data mining, machine learning), and optimization algorithms (linear programming, dynamic programming) to develop and test models that enhance coffee commodity trading efficiency and decision-making.

IV. PROPOSED SYSTEM

Key Components:

1. Data Collection Module:

- Collects data from various sources (World Bank, ICO, CFTC, USDA)
- Integrates IoT sensors for real-time coffee production and supply chain data

2. Fog Computing Module:

- Utilizes fog computing for efficient task offloading and real-time processing
- Employs smart ant colony optimization for optimal resource allocation

3. Big Data Analytics Module:

- Applies machine learning algorithms for market trend analysis and forecasting
- Uses data mining techniques for pattern detection and anomaly identification

4. Optimization Module:

- Employs linear programming and dynamic programming for optimal trading decisions
- Integrates risk management strategies for mitigating market volatility

5. Decision Support System (DSS) Module:

- Provides real-time insights and recommendations for traders and producers
- Offers data visualization and reporting tools for informed decision-making

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

V. SYSTEM ARCHITECTURE

The AI-driven coffee commodity trading system architecture consists of six layers. The Data Ingestion Layer collects and integrates coffee market data from various sources, including World Bank Commodities Price Data, International Coffee Organization reports, market feeds, and IoT sensor data. The Data Processing and Storage Layer employs big data analytics tools for data cleaning, preprocessing, warehousing, and real-time processing. The Predictive Analytics Engine utilizes machine learning algorithms for price forecasting, market trend analysis, risk management, and supply chain optimization.

System Architecture:

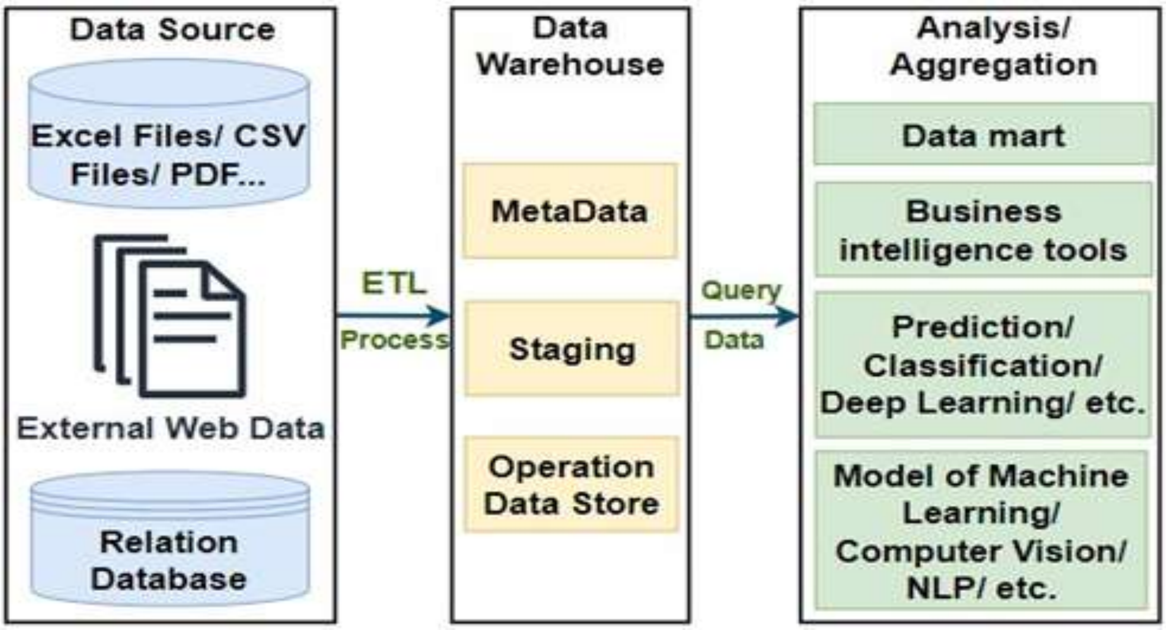
1. Cloud-based infrastructure for scalability and flexibility
2. Fog computing nodes for edge processing and real-time analytics
3. Big data storage solutions for handling large datasets
4. API integration for seamless data exchange between modules

Functionalities:

1. Real-time market monitoring and forecasting
2. Optimized trading decision support
3. Risk management and mitigation strategies
4. Supply chain optimization and monitoring
5. Data visualization and reporting

Benefits:

1. Improved trading efficiency and profitability
2. Enhanced decision-making through real-time insights
3. Reduced risks through optimized trading strategies
4. Increased transparency and accountability in coffee trading
5. Better supply chain management and sustainability



1. Data Integrity Verification Layer

Before the data is used for training, a Provable Data Possession (PDP) algorithm is employed to ensure that the data is authentic, untampered, and secure. This layer applies cryptographic methods to maintain the integrity of the data.

2. Distributed Machine Learning Layer

Data is processed in a distributed machine learning framework, where training occurs across multiple nodes or workers in parallel. A parameter server system is used to update and synchronize model parameters across all nodes, ensuring efficient and scalable model training.

3. Predictive Analytics Engine

This layer uses historical data and machine learning algorithms to predict fraudulent behavior. Supervised and unsupervised learning models are used to detect anomalies in financial transactions. This engine continuously learns and updates the fraud detection models based on new data.

4. Fraud Detection and Prevention Module

Real-time fraud detection models monitor incoming transactions. Anomalies or suspicious activities are flagged, and alerts are generated. The system also initiates preventive actions such as blocking transactions, flagging accounts, or notifying financial institutions for further investigation.

5. Feedback and Update Layer

The system is self-learning, where the results from the fraud detection module are used to update the predictive models. This layer ensures that the system adapts to new fraud patterns, improving accuracy over time. 6. Feedback and Update Layer

Updates predictive models and trading strategies:

- Continuous learning from market data
- Model retraining and validation
- Adaptation to changing market conditions

VI. ALGORITHM AND DESCRIPTION

The AI-driven coffee commodity trading algorithm utilizes machine learning and big data to predict market trends, generate trading recommendations, and optimize portfolio allocation. By continuously learning from market data, this algorithm enables informed decision-making, improves trading accuracy, and enhances management.

Algorithm:

Step 1: Data Collection

- Input: Historical coffee market data, weather data, supply chain data, and market sentiment data
- Description: Gather data from various sources, including World Bank Commodities Price Data, International Coffee Organization reports, weather APIs, and market news feeds

Step 2: Data Preprocessing

- Input: Collected data
- Process: Clean, transform, and normalize data for analysis
- Output: Preprocessed data

Step 3: Predictive Analytics Engine

- Input: Preprocessed data
- Process: Train machine learning models (e.g., LSTM, GRU, ARIMA) to forecast coffee prices and detect market trends
- Output: Trained models

Step 4: Feature Engineering

- Input: Trained models
- Process: Extract relevant features from data, including technical indicators and sentiment analysis
- Output: Engineered features

Step 5: Decision Support System

- Input: Engineered features
- Process: Generate trading recommendations based on predicted prices and market trends

- Output: Trading signals

Step 6: Portfolio Optimization

- Input: Trading signals
- Process: Optimize portfolio allocation using optimization algorithms (e.g., linear programming, genetic algorithms)
- Output: Optimized portfolio

Step 7: Continuous Learning

- Input: Market data and trading results
- Process: Retrain models and update trading strategies based on new data
- Output: Updated models and strategies

VII. CONCLUSION

Coffee, the second-largest global soft commodity, can take advantage of a comprehensive mining of daily and historical market data for more effective informed trading decisions. The paper presents a design and implementation of a coffee commodity trading Big Data warehouse supporting informed decision-making for trading coffee of small and mid-size enterprises (SMEs), business households, farmers, and others concerned through the visualization analysis dashboards. Specifically, we collected data from different sources, including the coffee market data from the New York Market of Intercontinental Exchange (ICE) from the portal Yahoo Finance, the Commitment of Traders (COT) report, and the

Commodity Index Trader (CIT) report from January 2000 to October 2022. The Extract, transform, and load (ETL) process is adopted to ingest coffee futures trading crawled data into the 3 layers data warehouse. Finally, the analytical system will extract and visualize selected key dimensions that influence coffee futures prices within different observation windows and perspectives. Our research focused our visualization on four main parameters: market price, volume, trailing over 250 days, and the highest/lowest 40 days. As a result, we implement a prototype of a coffee trading data warehouse on the crawled data and visualize trends in coffee futures prices.

Future work will be needed to collect data from the other markets and compare the correlation between different markets. We also want to define more parameters that can affect the coffee price. We tend to build an Android app based on cloud computing and use model machine learning, model computer vision, or natural language processing to build a real-time prediction application. The market for derivatives develops in parallel with the need to manage price changes of critical commodities. According to modern financial theory, introducing a derivative market in a spot market helps to complete the process of market-based pricing and prevent the risk of price fluctuations in the spot market at a low cost. In research authors synthesized a complete case study of commodity futures and agricultural option markets regarding the inter-time price relationship of stocks, speculation, price behavior, and institutional problems. The authors raise controversial issues and challenges for the studies focusing on risk management and market strategy, price and volatility behavior, e-transactions and trading funds, and international framework. In the coffee derivatives market, these problems will be affected by similar effects and problems. In the future, we will focus on data analysis and discovery and build effective prediction models for the coffee derivative market based on deep learning or machine learning algorithms.

VIII. FUTURE SCOPE

The future scope of AI-driven coffee commodity trading holds immense potential for growth and innovation. Integration with emerging technologies such as blockchain, Internet of Things (IoT), and natural language processing (NLP) can further enhance transparency, security, and decision-making capabilities. Expanding the system to incorporate real-time weather forecasting, soil health monitoring, and farm-level data can improve predictive accuracy. Additionally, exploring applications in other commodity markets, such as cocoa and sugar, can broaden the system's impact. Advancements in explainable AI (XAI) and transfer learning can also improve model interpretability and adaptability. Furthermore, integrating with trading platforms and exchanges can

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

enable seamless execution of AI-generated trading recommendations, revolutionizing the coffee commodity trading landscape.

IX. REFERENCES

[1] A. Kishor and C. Chakarbarty, ''Task offloading in fog computing for using smart ant colony optimization,'' Wireless Pers. Commun., vol. 127, no. 2, pp. 1683–1704, Nov. 2022

[2] World Bank. (2023). World Bank Commodities Price Data (The Pink Sheet). Accessed: 2024. [Online]. Available: https://www.worldbank.

[3] M. M. Singh, ''Extraction transformation and loading (ETL) of data using ETL tools,'' Int. J. Res. Appl. Sci. Eng. Technol., vol. 10, no. 6, pp. 4415–4420, Jun. 2022.

[4] N. B. V. Le, ''Building coffee commodity trading data warehouse: Decision making support applying big data analysis,'' M.S. thesis, Dept. Data Inform., Nat. Korea Maritime Ocean Univ., Busan, South Korea, 2022, pp. 1–40.

[5] Coffee: World Markets and Trade (Report), Foreign Agricult. Service/USDA, Global Market Anal. United States Dept. Agricult., Dec. 2022. [7] Monthly Coffee Markets Report, Int. Coffee Org., Apr. 2022

[6] Commodity Futures Trading Commission. About Commodity Futures Trading Commission. Accessed: 2024. [Online]. Available: https://cftc.gov/About/About The Commission.