IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 1, March 2025 Impact Factor: 7.67

Efficient Cloud-Based Healthcare Financial Analytics Using Snowflake and Dynamic Query Pruning

Waseem Chaudhry
Bandung Institute of Technology, Indonesia
wasim_ch2@hotmail.com

Abstract: *The rapid expansion of the Internet of Things (IoT) and rapid adoption of cloud computing have created higher demand for secure, scalable, and efficient data-processing solutions. However, existing healthcare financial analytics operations come with unique challenges, such as huge computational costs; inefficient query processing; and less security when carried in a cloud setting. Traditional methods are often rigid in terms of dynamic optimization, resulting in slower data retrieval speeds, extra storage overhead costs, and bottlenecks in performance for large-scale financial datasets. Furthermore, when it comes to encrypting financial information, conventional data encryption methods tend to compromise efficiency for the sake of security or the other way around due to limited flexibility to encrypt-on-the-fly. With the inability to enable dynamic pruning of irrelevant data during query processing, costs and time costs will be excessive, leading to overconsumption of resources. Thus, the absence of an adaptive framework for optimizing processing and cloud storage spells performance degradation in terms of both scalability and accuracy. To mitigate these challenges, this research proposes designing a Snowflake-based financial analysis system that applies the Dynamic Query Pruning (DQP) enhancement to achieve query optimization, reduce computational overhead, and support real-time analysis for healthcare financial systems. The study amalgamates hybrid cryptographic security, cloud-based storage, and machine learning-optimized solutions to allow accurate, secure, and scalable financial data analysis. From the experimental results, it can be inferred that the proposed model allows greater efficiency in data ingestion with quicker query execution and enormous cost savings in healthcare financial analytics. The findings shed light on how operability-on-cloud optimization techniques can be extended into the handling of large-scale financial datasets in the ensuring extra security and computational efficiency.*

Keywords: Internet of Things (IoT), cloud computing, healthcare financial analytics, query optimization, Dynamic Query Pruning (DQP), computational efficiency

I. INTRODUCTION

Data shared securely over the Internet of Things (IoT) system may require more novel methodologies owing to rising security threats, as conventional encryption is unable to keep up with an ever-changing network demand. The hybrid cryptographic key generation is therefore studied here and uses Super Singular Elliptic Curve Isogeny Cryptography (SSEIC) combined with Gaussian Walk Group Search Optimization (GWGSO) and Multi Swarm Adaptive Differential Evolution (MSADE) for the purpose of securing the data while trying to reduce the computational cost[1]. The rapid expansion of the IoT changed the landscape for shared data while making increased demand for security and privacy. To meet these challenges, the study applies isogeny-based hybrid cryptography using anisotropic random walks (ARW) and decentralized cultural co-evolutionary optimization (DCCO) to achieve secure data sharing for the IoT[2]. Fog computing serves well to improve performance and reduce latency in overcoming issues related to the cloud-based IoT scheme; however, the challenge remains the secure sharing of data and allocation of resources due to unstructured data flow over IoT. Such a study proposes the new setup of combining DBSCAN and fuzzy C-Means with a hybrid ABC-DE optimization that increases the accuracy and efficacy of clustering in fog computing compared to conventional security approaches[3]. Furthermore, a mixed-methodology is explained that combines PLONK for secure data sharing in IoT systems with Infinite Gaussian Mixture Models (IGMM) for dynamic load balancing. This guarantees real-time workload distribution with the smallest computational cost, thus enhancing network scalability, efficiency, and security[4]. Lastly, an analysis of financial time series becomes essential for trend forecasting, anomaly detection, and informed decision-making in dynamic markets. This study proposes a secure cloud-based framework combining DeepAR for time series forecasting, Neural Turing Machines (NTMs) for memory-based relationships, and Quadratic Discriminant Analysis (QDA) for strong classification, leveraging cloud infrastructure to combat challenges that include high dimensionality, noise, and non-linearity in financial data while providing timely, scalable, and secure real-time analytics[5].

Across such diverse fields as DNA analysis, physics, computer science, and medicine, the theoretical models on complex networks have formed space for very many analyses. Fractal geometry and graph theory are applied in interpreting DNA sequences with nucleotide conversion, graph construction, Hurst exponent estimation, and network property computation in the fields where DNA has its application [6]. It is no longer possible to overemphasize the contribution of these technologies in the new health architecture in diagnosis and real-time monitoring, be it by hybrid learning models or neural fuzzy systems, while minimizing uncertainty in the big pool of medical data collected from IoT devices. Just as the e-commerce and finance sectors have made cloud computing and smart networks address scalability, security, and efficiency questions associated with managing a large volume of data, so also will blockchain answer the same challenges Heart failure is, thus far, a health challenge of massive proportions across low-and-middle income countries and also in the developed world, characterized by high morbidity, mortality, and healthcare costs. Predictive models have to be advanced sufficiently toward improving the outcomes. The same artificial intelligence (AI), machine learning (ML), and blockchain are transforming human resource management (HRM) by improving data security, predictive analytics, and human resources automation in solving weaknesses of traditional centralized HR systems and enabling safe, scalable, and intelligent management of employee data

II. LITERATURE REVIEW

The mixed-methods approach, or data analysis, regression-based case studies on mobile finances, debit card availability, and cloud-facilitated application for financial literacy-from, were intended at understanding income equality as well as improvements in access across rural-urban areas. On the securing and efficient IoT Data clustering, that combined Adaptive Clustering AP to MQC in making strong encryption where AP clusters encrypted data on organizing critical documents safely. Performance indicators such as scalability, security, clustering efficiency, and computing overhead were used to validate the system. text, t-SNE for clustering, and Genetic Algorithms for optimization, combined by cloud technology, for real-time financial analysis with scalability. Another measure under which the performance of all components was optimized includes precision and scalability, which also incorporated built-in processes for smooth operation and accuracy.

Quantitative study incorporating statistical regression, econometric modeling, and machine learning to find an impact of mobile internet accessibility and financial inclusion on rural economic outcomes, significantly increasing income levels, entrepreneurship, and business growth.Thus internet access in rural areas becomes a bridge between the urban and rural divide and into the Global development goals healthcare system data analytics framework incorporating stochastic gradient boosting, generalized additive models, latent Dirichlet allocation, and regularized greedy forests, whereby data is preprocessed and aggregated into an ensemble prediction for high accuracy an IoMT-powered real-time surgical monitoring and control system using DCGANs and reinforcement learning. They introduced image synthesis and preprocessing for improving tissue segmentation and tool accuracy. GrabCut segmentation, bilateral filtering, SIFT feature extraction, SVM classification are used of segmentation, where DCGAN is used to generate synthetic surgical images, and reinforcement learning is used to optimize tool control for precision. a method for clustering and classifying electronic health records (EHR) and wearable sensors data using sequential mining in terms of trend detection and improved predictions, targeted toward reducing misdiagnosis in cardiovascular diseases and personalizing treatments with 93% accuracy, outperforming past methods for efficiency and timeliness, which used to have high error rates of 37%.

Data-driven strategy for the examination of regional inequality indicators with demonstration that financial inclusion through Cloud IoT significantly reduces income level disparity and propels economic development. the methodology based on deep learning and procedural models, bringing in datasets for experimentation, evaluation metrics, and existing methodologies mainly for comparison. crowdsourcing to gathering patient-centric data and provided input to optimize the therapeutic pathway in cardiology using decision trees (DTs) and measured the performance using indicators of data accuracy and prediction accuracy to assess standard versus novel methods.

III. PROBLEM STATEMENT

Regional income disparities that are increasing, alongside inefficiencies of deep learning-based evaluation methodologies, as well as the need to optimize therapeutic pathways in cardiology are pertinent challenges to economic development, data-driven analysis, and health care this study addresses such issues by exploring financial inclusion through Cloud IoT to tackle income inequality, deep learning methodologies for enhanced evaluation and comparison, and crowdsourcing using decision trees to optimize the treatment pathways in cardiology with higher data and predictive accuracies.

3.1 Objective

This project shows real-time financial analytics in health care by optimizing data-driven decision making with dynamic query pruning for better efficiency, accuracy, and real-time analytics [7]. The research also combines crowdsourcing and a decision tree to teach the application of healthcare analytics in optimizing cardiology treatment pathways with better predictive reliability[8]. It ensures scalability and security by adopting Snowflake's cloud infrastructure, enhanced encryption, and low computational overhead brought by dynamic query pruning for speedy processing that would generate insights for economics and healthcare applications [9].

IV. PROPOSED SNOWFLAKE AND DYNAMIC QUERY PRUNING (DQP) FOR FINANCIAL ANALYSIS SYSTEM

The proposed methodology supplements the current snowflake cloud-based data warehousing with dynamic query pruning to further the healthcare financial analytics. DQP, as an optimization technique to query execution, removes irrelevant data early, redundancies, thus improving processing efficiency. In this regard machine learning will identify trends and risks associated with finances. [10] Automated data pipeline transformation along the real-time data ingestion are available here. Advanced Encryption and Access control heighten security and compliance. In terms of scalability, accuracy, and efficiency of computation, it generally enhances the financial landscape in hospitals.

4.1 Data Collection

The processes of data collection include the factors in gathering the hospital financial records such as patient billings, insurance claims, medical expenses, and operational costs, portraying the primary database for analyzing finances. The data is collected and then prepared for further processing which involves tackling missing values and normalizing the data to achieve consistent quality of input for later snowflake-based finance analytics and query optimization.

4.2 Data Preprocessing

Imputation techniques or exclusion methods are used to keep data complete. It normalizes the financial records to bring about uniformity and enhance accuracy in data analysis in Snowflake for query processing towards reliable financial insights.

4.2.1 Handle Missing Value

Techniques used to cope with missing data in hospital financial records include mean imputation, median substitution, or deletion, which ensures that the data is complete. Mean imputation replaces missing data from a column with the mean of that column:

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, March 2025
Impact Factor: 7.67

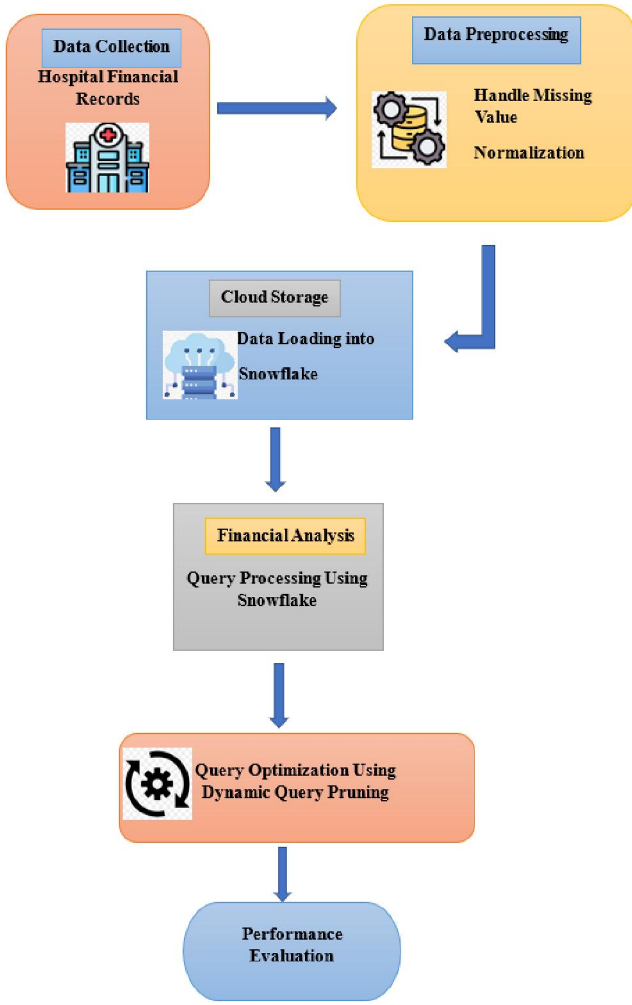

Figure 1: Snowflake and Dynamic Query Pruning (DQP) for Financial Analysis System

Equation for Handle Missing Value:

$$X_{\text{new}} = \frac{\sum X}{N} \tag{1}$$

where X represents available values, and N is the total count of non-missing values. This ensures consistency in Snowflake-based financial analysis.

4.2.2 Normalization

Normalization standardizes values in hospital financial data by reducing them into a single range. This increases consistency and accuracy in analysis via Snowflake. One way to do it is by Min-Max Normalization, whereby all data is transformed to a new range [0,1] through the application of the formula.

Equation for Normalization:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{2}$$

where X is the original value, $X_{\min}$ is the minimum value, and $X_{\max}$ is the maximum value. This ensures uniformity and enhances query optimization. A hybrid CNN-RNN framework for secure cloud-based healthcare data processing in banking achieving high accuracy but facing latency under user load was proposed by Chaitanya Vasamsetty et al.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 1, March 2025 Impact Factor: 7.67

(2024).[11] This informs developed approach to enhance healthcare financial analytics, focusing on scalable performance and efficient query execution.

4.3 Cloud Storage

Processing cloud-based storage fills hospital finance data light Speedy Efficient Loading into Snowflake [12]. Snowflake's architecture is scale-up/down to allow an agile and seamless return of structured and semi-structured data; integrated with fast accesses [13]. Within the framework of improved query processing and financial analysis, the data undergo staging, validation, and transforming [14].

4.4 Cloud-Based Healthcare Financial Analytics Using Snowflake

Snowflake is known for processing financial data queries efficiently through columnar storage in parallel execution [15]. Due to fast retrieval and analysis of hospital financial data, resource computation allocation dynamically optimizes query execution. Plus, it refines the unnecessary data scans with Dynamic Query Pruning (DQP).

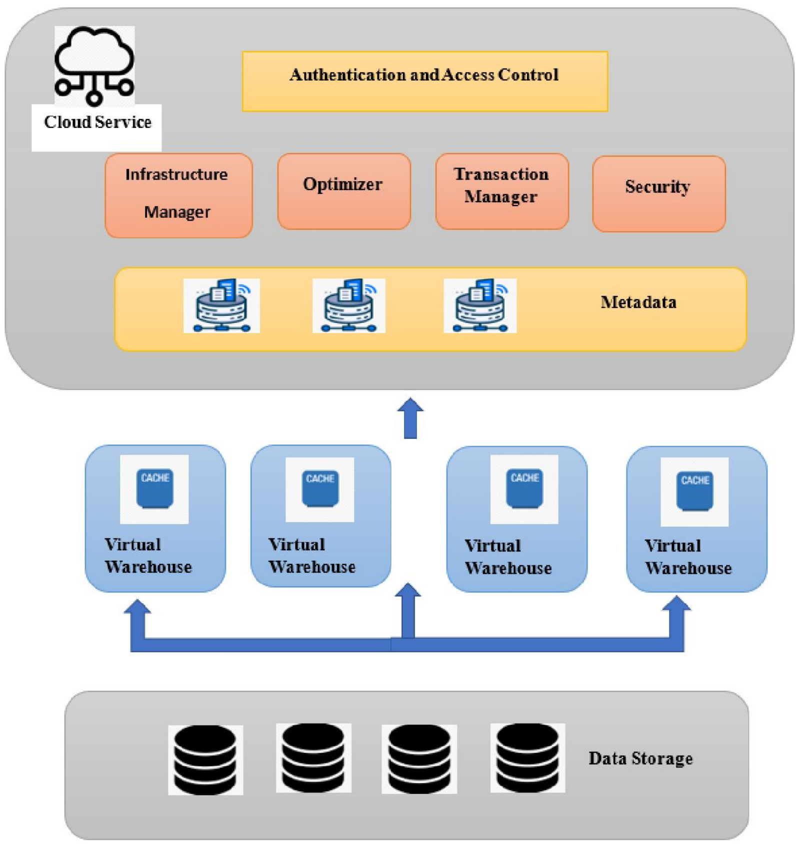

Figure 2: Cloud-Based Healthcare Financial Analytics Using Snowflake

A common query execution cost formula in Snowflake is:

$$\text{Total Query Cost} = \sum_{i=1}^{n} \left(\text{Compute Cost}_i + \text{Storage Cost}_i \right) \quad (3)$$

where compute cost depends on processing time and resources, and storage cost is based on data volume. This ensures efficient financial analysis [16].

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, March 2025
Impact Factor: 7.67

4.5 Query Optimization Using Dynamic Query Pruning

Dynamic Query Pruning (DQP) enables scattered performance of queries in snowflakes using runtime data elimination. It dynamically filters partitions according to the query conditions with lowered I/O operations and processing times. This proves most helpful with large datasets concerning finances and makes sure of fast and cost-efficient analyses. Rahul Jadon et al. (2024) [17] developed a high-accuracy plant disease detection model using EfficientNet with HOG features and data augmentation, achieving 98.8% accuracy. Adopting a similar strategy, proposed framework incorporates performance-efficient in healthcare financial analytics, where proposed framework mirrors the approach to scalable, accurate systems.

A common equation representing query pruning efficiency is:

$$\text{Optimized Query Time} = \frac{\text{Initial Query Time} \times \text{Scanned Partitions}}{\text{Total Partitions}} \quad (4)$$

where fewer scanned partitions lead to lower execution time, improving financial data retrieval efficiency.

V. RESULTS AND DISCUSSION

Dynamic Query Pruning (DQP), developed in Snowflake, heightens query speed, cost savings, and overall query response time during processing data [18]. It reduces retrieval time, storage overhead, execution costs, and provides accurate yet scalable hospital financial analytics treatment [19].

Cloud Storage and Data Size

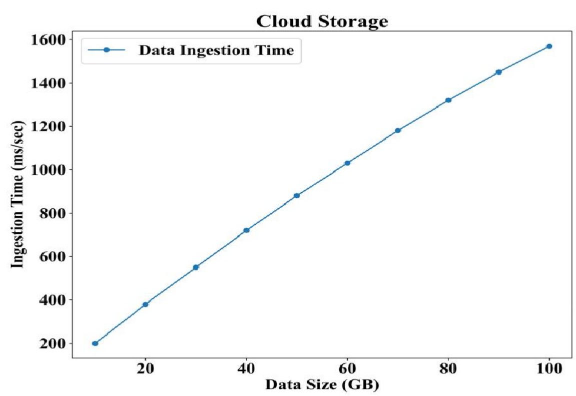

Figure 3: Cloud Storage and Data Size

In Figure 3, These have brought a change in the trust regarding the data ingestion time [20] where in cloud storage there will be reasonable linearity in terms of linear data ingestion time against data size,[21] which clearly shows that larger data sizes would eat time additionally processed. It clearly indicates the need for efficient data loading and optimization into the systems such as Snowflake [22]. Deep learning models for threat detection, as demonstrated in the work of Venkat Garikipati (2024) [23], address challenges like scalability and data preprocessing. Leveraging this approach, proposed study applies similar strategies in healthcare financial analytics using Dynamic Query Pruning for adaptive, secure, and efficient query processing.

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 1, March 2025 Impact Factor: 7.67

Performance Evaluation

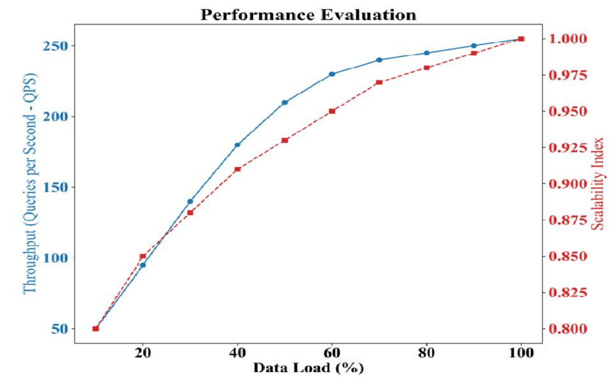

Figure 4: Performance Evaluation

Figure 4 shows that the depiction evaluates performance throughput represented by QPS and efficiency index compared to the data load, which indicates the percentage of the load. With increased data load, throughput rises but eventually plateaus, while consistently going up for scalability which further means the improved query contains all possible scenarios for the adaptability of the system [24].

Query Optimization

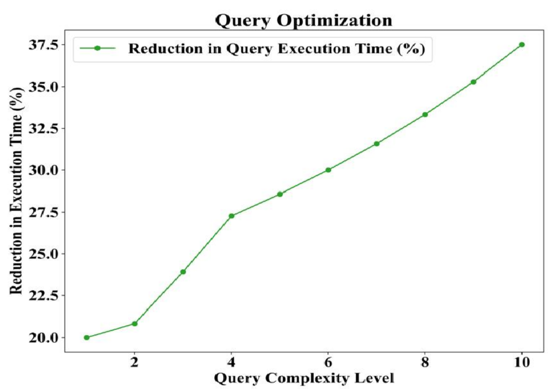

Figure 5: Query Optimization

Figure 5 explains the execution time reduction (%) as the complexity of queries increases. Thus, for the most complicated queries, the execution time benefits would be the highest. This could be cited as the strength of optimization techniques regarding query performance improvements.

VI. CONCLUSION

Snowflake based financial intelligence system that is using Dynamic Query Pruning (DQP), which improves how data is processed, executed, and optimized in queries. Results show there is an increase in data ingestion and scalability with less query execution time which translates into better performance in financial analytics. Cloud Integration presents healthcare records in terms of a financial database thus enhancing decision-making and accuracy. Overall, the approach used balances resource usage against delivering high-performance data analytics.

REFERENCES

[1]. Shamim, S., Yang, Y., Zia, N. U., & Shah, M. H. (2021). Big data management capabilities in the hospitality sector: Service innovation and customer generated online quality ratings. *Computers in Human Behavior*, *121*, 106777.

[2]. Allam, K., Ankam, M., & Nalmala, M. (2023). CLOUD DATA WAREHOUSING: HOW SNOWFLAKE IS TRANSFORMING BIG DATA MANAGEMENT. *Journal of Computer Engineering and Technology (IJCET)*, *14*(3), 156-162.

[3]. Kumar, Y., Marchena, J., Awlla, A. H., Li, J. J., & Abdalla, H. B. (2024). The AI-Powered Evolution of Big Data. *Applied Sciences*, *14*(22), 10176.

[4]. Machireddy, J. R. (2024). Integrating Machine Learning-Driven RPA with Cloud-Based Data Warehousing for Real-Time Analytics and Business Intelligence. *Hong Kong Journal of AI and Medicine*, *4*(1), 98-121.

[5]. Ibrahim, F., & Aoun, M. (2022). Improving query efficiency in heterogeneous big data environments through advanced query processing techniques. *Journal of Contemporary Healthcare Analytics*, *6*(6), 40-64.

[6]. Ramamoorthi, V. (2023). Applications of AI in Cloud Computing: Transforming Industries and Future Opportunities. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *9*(4), 472-483.

[7]. Kumar, R. M. (2024). Integrating data science with cloud computing: Opportunities and challenges. *Journal of Recent Trends in Computer Science and Engineering*, *12*(2), 40-53.

[8]. Pendyala, S. K. (2024). Optimizing Cloud Solutions: Streamlining Healthcare Data Lakes For Cost Efficiency. *Technology (IJRCAIT)*, *7*(2).

[9]. Ogeawuchi, J. C., Akpe, O. E., Abayomi, A. A., Agboola, O. A., Ogbuefi, E. J. I. E. L. O., & Owoade, S. A. M. U. E. L. (2022). Systematic review of advanced data governance strategies for securing cloud-based data warehouses and pipelines. *Iconic Research and Engineering Journals*, *6*(1), 784-794.

[10]. Nambiar, A., & Mundra, D. (2022). An overview of data warehouse and data lake in modern enterprise data management. *Big data and cognitive computing*, *6*(4), 132.

[11]. Vasamsetty, C., Nippatla, R. P., Kadiyala, B., Alavilli, S. K., Boyapati, S., & Karthick, M. (2024). AI-powered healthcare services in banking: A hybrid deep learning framework for secure cloud-based medical data processing. International Journal of HRM and Organizational Behavior, 1(1).

[12]. Machireddy, J. R., & Devapatla, H. (2022). Leveraging robotic process automation (rpa) with ai and machine learning for scalable data science workflows in cloud-based data warehousing environments. *Australian Journal of Machine Learning Research & Applications*, *2*(2), 234-261.

[13]. Jaiswal, S., Shaheen, N., Chinta, D. U., Singh, N., Goel, O., & Chhapola, A. (2024). Modernizing Workforce Structure Management to Drive Innovation in US Organizations Using Oracle HCM Cloud. *International Journal of Research Radicals in Multidisciplinary Fields, ISSN: 2960-043X, 3 (2), 269*, *293*.

[14]. Elzein, N. M., Majid, M. A., Hashem, I. A. T., Ibrahim, A. O., Abulfaraj, A. W., & Binzagr, F. (2023). Jqpro: Join query processing in a distributed system for big RDF data using the hash-merge join technique. *Mathematics*, *11*(5), 1275.

[15]. Joseph, A. (2023). A Holistic Framework for Unifying Data Security and Management in Modern Enterprises. *International Journal of Social and Business Sciences*, *17*(10), 602-609.

[16]. Shaheen, N., Jaiswal, S., Mangal, A., Singh, D. S. P., Jain, S., & Agarwal, R. (2024). Enhancing Employee Experience and Organizational Growth through Self-Service Functionalities in Oracle HCM Cloud. *Journal of Quantum Science and Technology (JQST)*, *1*(3), 247-264.

[17]. Jadon, R., Srinivasan, K., Chauhan, G. S., Budda, R., Gollapalli, V. S. T., & Mekala, R. (2024). A novel deep learning framework for plant disease detection using EfficientNet model. Indo-American Journal of Pharma and Bio Sciences, 22(4)

[18]. Chhapola, A., Shrivastav, A., Ravi, V. K., Jampani, S., Gudavalli, S., & Goel, P. (2022). Cloud-native DevOps practices for SAP deployment. *International Journal of Research in Modern Engineering and Emerging Technology*, *10*(2), 95-116.

[19]. Ravalji, Vardhansinh Yogendrasinnh, and M. S. R. Prasad. "Advanced .NET Core APIs for financial transaction processing." *International Journal for Research in Management and Pharmacy (IJRMP), 13 (10), 22. https://www. ijrmp. org* (2024).

[20]. Garudasu, S., Kumar, A., Joshi, A., Goel, O., Kumar, L., & Jain, A. (2024). Implementing Row-Level Security in Power BI: Techniques for Securing Data in Live Connection Reports. *International Journal of Progressive Research in Engineering Management and Science (IJPREMS)*, *4*(4), 2187-2204.

[21]. Gupta, A. K., & Khan, S. (2024). Streamlining SAP Basis Operations to Improve Business Continuity in Modern Enterprises. *International Journal of Computer Science and Engineering (IJCSE)*, *13*(2), 923-954.

[22]. Iliadis, A., & Acker, A. (2024). The Palantir Files: public interest archives for platform accountability. *Information, Communication & Society*, *27*(13), 2343-2365.

[23]. Garikipati, V., Ubagaram, C., Dyavani, N. R., Jayaprakasam, B. S., Mandala, R. R., & Mekala, R. (2024). AI-powered cyber threat detection and response: Leveraging network, textual, and relational data. Indo-American Journal of Pharma and Bio Sciences, 22(2).

[24]. Gupta, H., & Saxena, D. N. (2024). Leveraging Machine Learning for Real-Time Pricing and Yield Optimization in Commerce. *International Journal of Research Radicals in Multidisciplinary Fields*, *3*(2), 501-525