Cloud and Data Transformation in Banking: Managing Middle and Back Office Operations Using Snowflake and Databricks

Venugopal Tamraparani

Vice President, Marlabs Piscataway , NJ , USA, Email: Venugopal.tp@gmail.com

Received: 17.10.2021 Revised: 18.11.2021 Accepted: 20.12.2021

ABSTRACT
To improve their data transformation procedures used in their back-office and middle-office activities, leading financial institutions are slowly but surely adopting cloud computing. With a focus on cloud-based technologies and platforms like Snowflake and Databricks, this article aims to explore the program management approaches used in transformation initiatives. This study focuses on the complexities of handling large transitions. The complications include data migration, system integration, and regulatory compliance issues. It also underlines how cloud-native solutions improve operational efficiency, scalability, and data accessibility. Case studies from the banking industry show that managing cloud migration requires careful preparation, cross-departmental coordination, and adaptable techniques. The essay focuses on the key findings and proposes an approach that banks and other financial institutions may use to improve their data architecture and optimize cloud operations.

Keywords: Data Transformation, Cloud Computing, Data Mitigation, Banking, Financial.

1. INTRODUCTION

In response to the pandemic, the expectations of digital native customers, and the necessity to make the most of current investments, many organizations have been pursuing digital transformation over the past ten years. When businesses are able to access the enterprise core and experience concrete benefits, they are able to create stronger innovations.Despite the importance of modern technology in driving digital transformation, just 7% of core banking workloads worldwide have moved to the cloud as of 2022 (link appears to be external to IBM). Although digital front ends and middle-office apps have received most of the attention so far, it is clear that these systems also require modernization [1].

We think updating these back-and middle-office systems is where digital transformation really shines. Financial organisations rely on these systems heavily; updating them can boost their competitiveness, fficiency, resilience, security, and compliance in the face of ever-changing regulations [2].By migrating its trading platform to the cloud, one of the top financial trading organisations was able to increase performance by a factor of three. This allowed their clients to take advantage of short-term profit possibilities on a platform that was both secure and easy to use, while also safeguarding their money.

Particularly for long-standing businesses or those operating in highly regulated sectors, updating back- and middle-office tasks can be a daunting task. In addition to the intricacy of migrating to the cloud and the inherent regulatory and security concerns in these sectors, a thorough comprehension of the current operations and the company's future goals is essential. Bringing the middle and back office systems up to date will have more positive effects than negative ones, in our opinion.

A. Businesses typically face difficulties in four main areas

The expense of a safe, resilient, and regulatory-compliant multicloud infrastructure is something that many of our clients have expressed concern about. A private computing environment that includes computation, containers, databases, encryption, and key management is essential for a secure and robust platform. The expense of maintaining a controls framework is a burden, but it should be used to support an enterprise's regulatory compliance obligations through an industry-specific common-controls framework [3].

- We also think that cloud-concentration risk, which affects application resilience and performance, increases when you depend on only one cloud provider. The complexity of programs used in the middle and back offices presents a substantial challenge that organisations often face.
- Factors adding to this complexity include processing transactions quickly and in large quantities, ensuring data consistency, processing in batches, and dealing with complicated business rules. Even

though these apps have been around for a while, there aren't many resources that are familiar with the underlying business principles and technology [4].

- Data: Data is an additional crucial area of focus. Data encryption, consistency, latency, and sprawl are crucial considerations. The user experience can be affected by latency that occurs between cloud front ends and on-premises middle- and back-office applications. Workload placement decisions may be impacted by data sovereignty rules, which could hinder innovation.
- When companies undergo modernisation, they may face difficulties in adjusting their operating models to accommodate the new technology, procedures, and ways of working. To succeed, you need the right people with the right talents. For example, to convert applications, you'll need a Site Reliability Engineer (SRE), full-stack developers, and cloud engineers. In order to meet the demands of cloud-native development, companies should think about providing their employees with the skills they need to succeed in the ever-evolving digital world.

If you want your digital transformation journey in the front, middle, and back office to be a success, you need a plan and vision that includes things like placing workloads optimally, modernising with a focus on business processes, making sure everything is secure and compliant, and using a 360-degree operating model.Ascertain the most efficient distribution of tasks: Tomaximise business value, it is recommended to reevaluate your current application and data portfolio in light of a hybrid multicloud approach. This will help you select the optimal modernisation pattern and the "fit-for-purpose" landing zone. When deciding where to put your workload, keep in mind its total cost of ownership, performance, security, resilience, and compliance. When it comes to optimising business processes and gaining efficiency through application modernisation, this orientation is crucial. The reality is that various tasks require varying levels of efficiency [5].

B. Focus on the business processes

Find the best way to update front-, middle-, and back-office data and apps by analysing important end-to-end business processes. In order to provide smooth and efficient experiences, businesses should also evaluate their digital supply chain, which includes things like business rules, process automation, and workflows. It is recommended to modernise the front office simultaneously with the strategy for middle and back-office applications. Prototypes can be used to evaluate the approach. In order to revamp middle-office apps, businesses must establish new workflows, enable customisable business rules, and update the integration layer using messaging, eventing, APIs, and other technology. Modern technologies such as Linux and Java can be utilised to upgrade existing back-office applications on the mainframe. This allows for core banking functions and high-throughput transaction processing [6].

Integrate modernization-related security and compliance features into the design process: Hybrid multicloud environments present unique security and compliance difficulties due to factors such as complexity, lack of end-to-end observability, API vulnerabilities, and ever-changing regulatory requirements. Over time, we've seen that security and compliance frameworks that allow users to "roll their own" have proven to be quite expensive to maintain [7]. Operations may be made easier and cloud adoption can be accelerated with compliance capabilities that are in line with industry norms. In our next blog post, we will delve deeper into the topic of security.

Create an all-encompassing operational model: Improving agility, establishing a consistent automated delivery pipeline with continuous monitoring and feedback, and enabling collaboration and prioritisation of data-driven investments are all benefits of implementing an operating model that addresses platform governance, DevSecOps, and FinOps. For IT to undergo a transformation, the emphasis must be on culture and cloud-native development abilities, with mechanisms and structures put in place to promote the exchange of ideas and information.

2. Program management approaches based on transformation initiatives

As a result of incorporating cutting-edge technology with long-standing banking procedures, the financial sector is undergoing a sea change with the digital transformation of banking operations. Not only is this change altering the way banks function, but it is also changing the expectations of customers and the dynamics of competition in the business [7]. More efficient, nimble, and customer-centric service delivery is made possible by digital transformation, which is the strategic adoption of digital technology throughout all operations of financial institutions. Digital transformation offers both opportunities and problems for banks in developing nations that are trying to stay competitive and increase their market reach. These economies still have a long way to go in terms of infrastructure and financial inclusion.

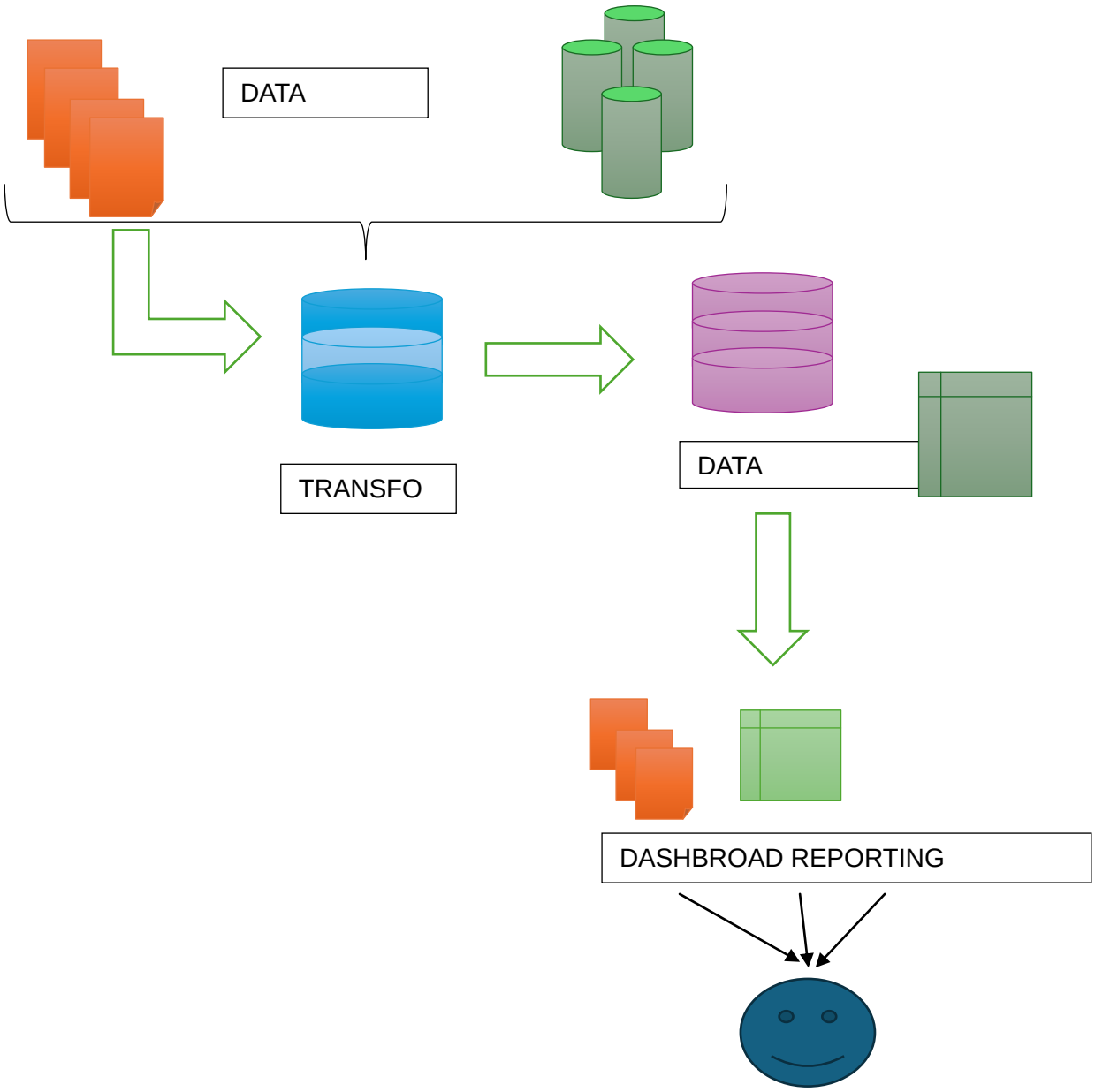

Figure 1. Data Warehouse Architecture using DATA Process

The use of state-of-the-art technology to improve operational efficiency and propel innovation is fundamental to digital transformation [8]. This technology includes blockchain, artificial intelligence (AI), big data analytics, and fintech solutions. Banks can automate operations, improve decision-making, and meet consumers' ever-changing demands with personalised services made possible by these technologies. Machine learning algorithms aid financial institutions in detecting fraud and managing risk, while AI-powered chatbots and virtual assistants are seeing increased usage in handling consumer enquiries. [9] found that these innovations have improved the accessibility, reliability, and security of banking services, especially in areas without typical banking infrastructure. Traditional banks are now up against nimble fintech firms that provide specialised financial goods and services, thanks to the change towards digital banking as in Figure 1.

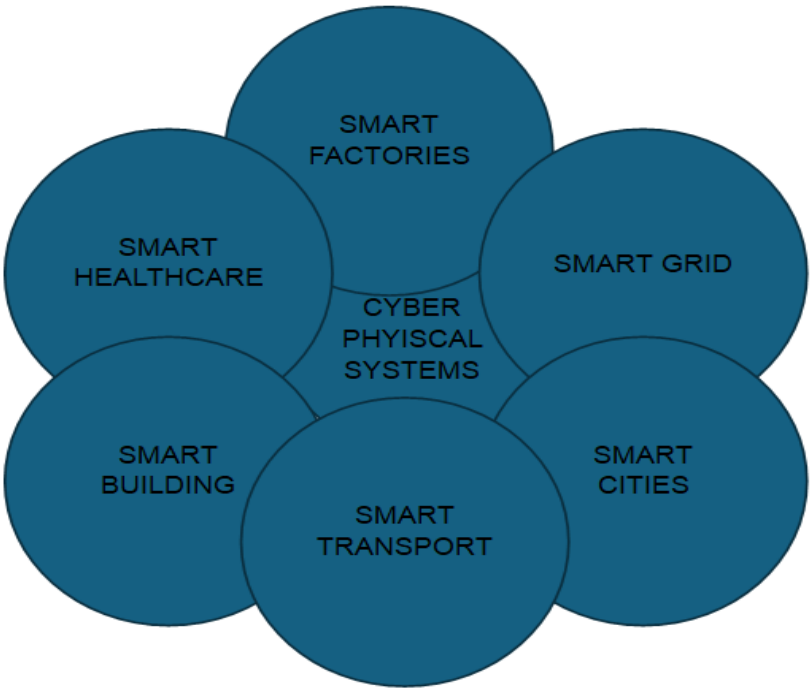

Figure 2. CPS Applications

According to [10], financial technology companies offer alternatives to traditional banking services, such as digital wallets, mobile payments, and peer-to-peer lending, using digital platforms. These alternatives are frequently more convenient and cost-effective. With the advent of fintech, underprivileged communities in developing economies can now access financial services more quickly through digital channels, speeding up the process of financial inclusion. New business models and alliances are reshaping the financial environment in these locations as a result of the convergence of banking and fintech. Adopting blockchain technology, which improves the security, transparency, and efficiency of financial transactions, is also strongly related to digital transformation in banking. By eliminating middlemen and facilitating instantaneous settlements, blockchain technology has the potential to significantly cut down on transaction fees. Blockchain technology provides a trustworthy and open alternative to traditional banking in developing nations where people may be wary of putting their money in institutions.

Many underdeveloped nations rely on remittances and cross-border transactions for their economic survival; this technology is especially pertinent to these processes. To further facilitate data-driven decision-making and the provision of tailored financial solutions, big data analytics is an essential part of digital transformation. Banks can learn about their customers' habits and preferences by sifting through mountains of data. This capability boosts operational efficiency by minimising operational risks and optimising resource allocation, which in turn promotes customer satisfaction. According to [11], emerging nations generally deal with inconsistent and fragmented consumer data. However, big data analytics can help improve financial services by integrating and analysing diverse data sources. The path to digital transformation is not without its obstacles, especially for developing economies, despite the many advantages.

The digital gap, which includes things like low levels of digital literacy, insufficient digital infrastructure, and limited internet access, is one of the main hurdles (Barroso & Laborda, 2022). Financial institutions need to put money into programs that raise awareness and educate the public in areas where these barriers are particularly severe. Also, many poor countries' regulatory frameworks are still in the process of being refined, so they may still have regulations and laws that aren't up to speed with the intricacies of digital banking [12]. Consequently, the rate of digital adoption and innovation can be slowed down by concerns about regulatory ambiguity and compliance. Additionally, with banking operations undergoing digital transition, cybersecurity is still a major worry. According to [13], cyberattacks, data breaches, and fraud are growing threats to banks as they move towards digitising their services. As a result of using antiquated technologies and without strong cybersecurity procedures, developing countries are more at risk.

Investing in secure infrastructure, continuously monitoring, and collaborating with cybersecurity professionals are all necessary to address these dangers. If these problems aren't fixed, digital

transformation projects can hit a wall and lose customers' trust [14]. Beyond enhancing operations, digital transformation affects both the expectations and experiences of customers. Customers expect services that are fast, convenient, and tailored to their lifestyles in today's digital-first environment (Osei et al., 2023). Whether they're using a mobile app, an online platform, or a branch, customers of banks that have effectively transformed digitally can have a consistent experience across all of these channels. Customers in developing nations have various demands, and access to digital services can vary greatly based on characteristics like geography and socioeconomic status. To fulfil these diverse needs, an omnichannel approach is important.

3. Cloud And Data Transformation Process

Assessing the performance can be achieved based on the organization's achievement with certain tasks and activities. The organization's achievement and performance can be calculated and analyzed to represent the operational impact, which may be positive or negative [13]. Here comes the role of big data technology as it plays a challenging role while the top companies like Google, Facebook, etc. consider the unstructured data as it needs to be managed and utilized based on vast amount of large data [14]. This big data takes the major role in the field of information technology as it makes the process of managing and analysing the data. As the large volume of data to be process will be much more complex with the standard technologies, big data takes the process of managing and analysing those data [10-12]. The information technology purely depends on data utilization and managing them as the entire organization takes the lead as it dependence on the data. Those organization takes the role of organizing the tool to managing the data through application and their features are represented in Figure 2.

Based on the business organization, data warehouse plays a significant role as it makes the various processes, technologies and tools are performed along with the data as it gathers certain knowledge and useful data as it results in the profitable information based on certain organization activities. In order to utilize the emerging opportunities, there are various integration of data warehouse, mining data, analysing the data in multi dimension with certain graphical representation [15]. Then these data takes certain analyses to determine the growth and enhancement of the organization. Here the business intelligence includes, data warehouse, data mining, analysis the multi dimension, etc. In this data warehouse technology, data can be managed from various multiple sources as it gives improved insights on the business. It is the repository to store large volumes of information especially as the data gets generated by the business organization. In this data are processed, transformed, and absorbed by the users based on the tools being used by the business organization.

The DATA process is being studied based on the data warehouse on sales as it is represented using pentaho tools [16]. In this study, the sales in the data warehouse are processed to form useful information as it gets analysed and processed to determine the performance of the sales to make effective decision making. [17] has analyzed the data store of the United States from 2014-2017 as it takes the representation of Microsoft office doc as it transforms the data into MYSQL database. The data warehouse is analyzed on certain applications of PHI Minimart using pentaho tools to analyze the data warehouse sales with the help of an DATA system with information sales. The data warehouse is analyzed using OLAP process of data creation as schema. The data sales of minimart sales 2008 [18] as it is represented as sales chart and their total sales for each department in the organization.

The online market datasets and based on that the data warehouse is designed and it is represented as the schema of snowflake schema. The data modeling of this snakeflake schema takes various 9 steps of methodology to model the data. The data platform is used to process the data in order to represent the data, which is informative are useful. Business intelligence is applied to determine the effectiveness of decision-making and improved data quality.

In the proposed approach, the multi-dimensional model is deployed as it helps to construct the data warehouse as the modelling process

A. Multi-Dimensional Model

The data's are modelled based on the process of data warehouse with kimbell's approach as, Initial step helps to identify the business process as analyse the students entry, their registering for the new course, analyse the student grade obtained, student payment and their graduation count of students. Then data can be grained and declared based on the data measure as it is integrated with the data granularity. The data measurement is based on the student information with every year, information of the student who is registering, distribution of grade of the student, graduation data and payment done by the student. The data dimension is identified based on role associated with the student.

"Rajesh (**Gender**), Semester, Adgami (**Religion**), Mahesh (**Student**), Mahesh Reg (**Student Registration**), Model, …….. Bank Name (**Name of the Bank**), Delhi (**Payment**), STAS (**High School**), ……."

Finally, the facts are identified based data tables being deployed by the individual user as represented as,

Student Table | Student Count | Facts
Table Student Registration | Student Count
Grade table Student | CGPA | GPA
Payment Student Table | Payment Student | Status

B. Modified DATA Process

In the DATA process, the server, mainframe, source production is taken as the input as it get the initial process of extraction as takes the unstructured data and perform extraction on the data stage and generalize the data as structured one. Then it performs data cleaning as it cleans the unwanted data into useful one. Then data conforms and deliver make the data to be analysed and processed. Finally it is sent back to the use application by the end user as it operates several processes such as, scheduling the data, handling the data exception, data recovery, and data restart, checking the data quality and reliability with certain user support as in Figure 3.

Algorithm 1: Modified DATA Process for Data Integration

Input: Server; Mainframe; Data Source

Output: Data Integration

1. Identify the data source to perform data extraction
2. The data tables are structured and their explanation
    - Using a merging strategy, create the master table
        a. Table 1 → DB1
        b. Table 1 → DB2
        c. Merger (Table 1 from different DB1 and 2)
    - Using Merge Union Strategy, Create the master table
        a. Table 1 → DB1
        b. Table 1 → DB2
        c. Table 1 not influence on Table 2
        d. Apply Merge strategy with Union Action
    - Using Union Strategy, create a transaction table
        a. Large table → Target Table
        b. Target Table = {Pilot Table}
3. Heterogeneous Data Source
4. DB = DB {Merge and Union Strategy}
5. Perform Data identification and analyze the data source
    a. hostname (Domain name or IP address of the database server)
    b. Database name (The schema or other database identifiers)
    c. Port Username and password to access the data source)
6. IFNULL( ) and NULL Values Expression
7. Creating the data dimension tables.

The process of data integration is performed and represented in Algorithm 1.

C. Data Cleaning and Conforming:

In the cleaning process, the data errors are detected and removed as inconsistent data's are identified as it results in enhancing the data quality.

- Analysis the data
- Refining the data
- Verify the data

Based on the data rule and value, re-analyze the data based on DATA process to perform data confirmation, data join and association. Then the data helps to create conform data dimension as it takes data hierarchy and dimension

D. Data Delivery and Loading:

In this data table, only user analysis data's are present and for each data table, the key is generated. Here the start schema being used as it helps to de normalize the data and then deliver the data as the targeted data table.

- Slow Changing Dimension (SCD)
- Loading the Fact Table

E. DATA Process Testing

The main objective of DATA process is to test the data by identifying and collecting the data errors as it occurs before the process of data analytical and reporting. Various data testing are performed i.e. validate the data completion, testing the Meta data and DATA test incrementing. All loaded data are measured as it helps to validate the objective if data completion. The table defined are verified while performing the Meta data testing. Data document can be mapped based on checking the data type, length of the data, checking

the data index between the source and destination table. Then the unwanted data duplication can be determined based on the incremented DATA testing as in Figure 4.

Finding scholarly research on the topics of big data and finance is the goal of this project. In order to complete this research, secondary data sourceswere utilized to acquire pertinent data. The study used the online databases Scopus, the web of science, and Google scholar to find additional data. Big data finance, big data and finance, big data and the stock market, big data in banking, big data management, and big data and FinTech are the keywords of this study. The researcher occasionally looked at publications on the Internet that had not been published in academic or peer-reviewed journals, but the search was primarily limited to those. Search engine information can occasionally aid in understanding a subject. Big data has already been studied, but the amount of data on it in finance is

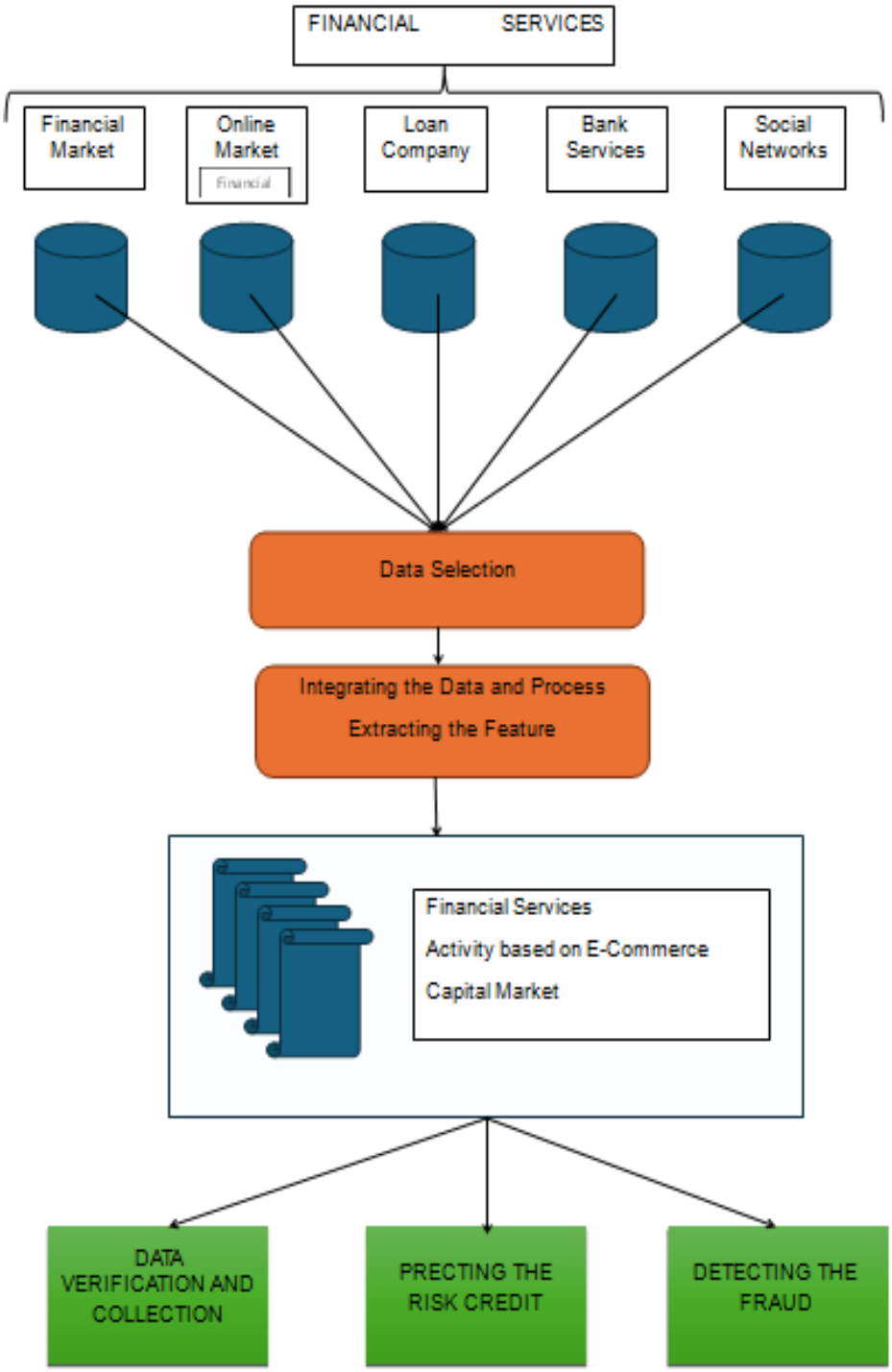

Figure 3. Financial Data Processing

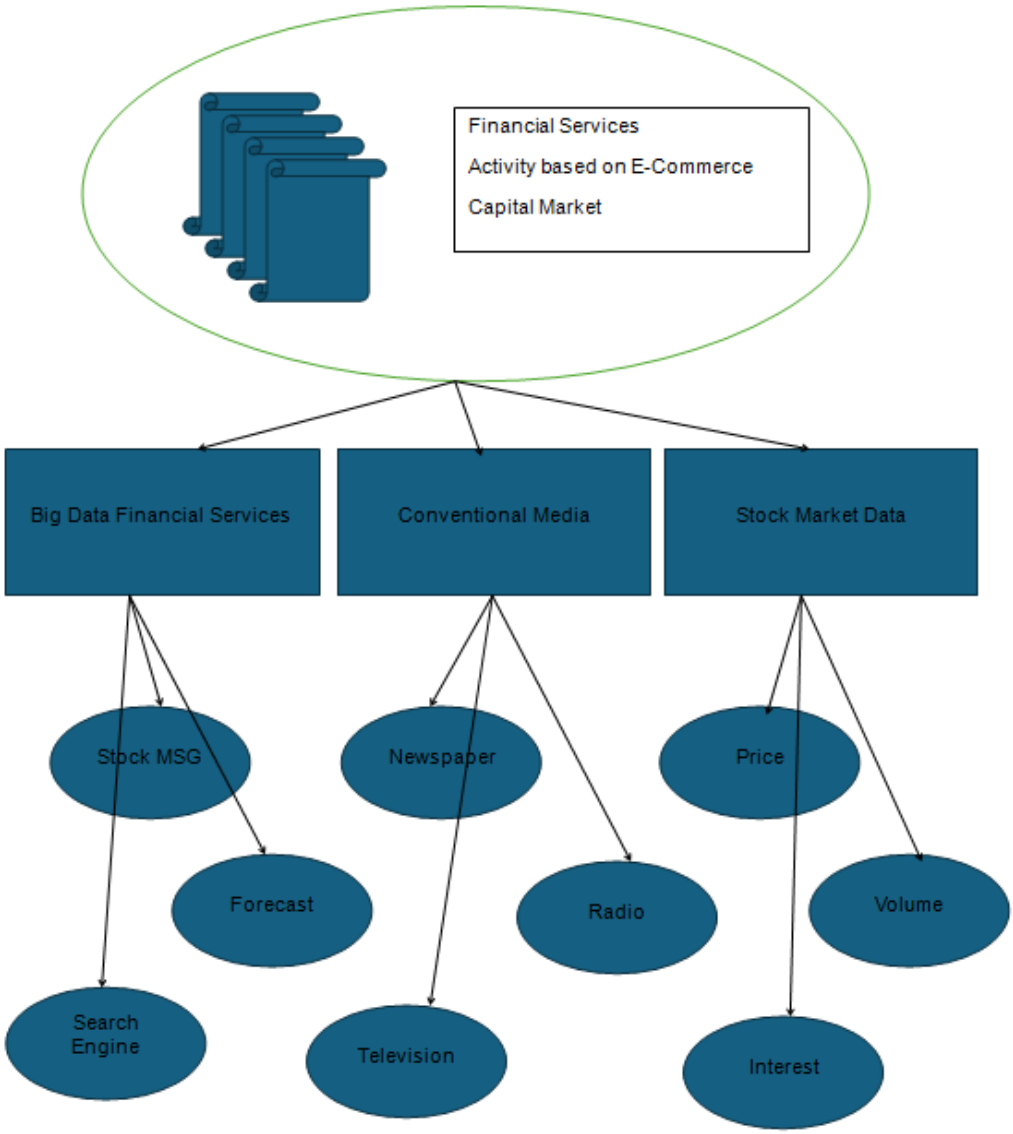

Figure 4. Financial Services Offered

not as great. For this reason, we did not restrict our search to a specific time frame, as doing so could have narrowed the field of this study's study.

After reviewing the literature, this study discovered that big data is largely associated with the financial market, specifically Internet finance. Credit Service Company, financial service administration, financial software, etc. The four main categories of the financial industry—financial markets, online marketplaces, lending companies, and banks—are the focus of the statistics. These businesses generate billions of pieces of data every day as a result of their regular transactions, user accounts, data updates, account modifications, and other operations.

These businesses analyses zettabytes of data and utilise it to forecast customer preferences based on past behaviour as well as the degree of credit risk associated with each user. On the basis of these data, financial institutions aid in decision-making. However, various financial institutions are processing big data and receiving assistance for verification and collection, predicting credit risk, and detecting fraud. As billions of data are being generated from many sources, missing data, data quality, and data reliability are all critical issues. Regardless, the concept of the significance of economic large-scale data is starting to shape, in which the sources of financial market information are cited as stock market data (e.g., stock prices, stock trading volume, interest rates, etc.), social media (e.g., Facebook, Twitter, newspapers, advertising, television, etc.). Market return prediction, volatility forecasting, position valuation, spotting excessive trading volume, risk analysis, stock price changes, option pricing, algorithmic trading, idiosyncratic volatility, and many other financial market applications rely heavily on this information.

Based on the Figure 3 and 4, some financial outcomes are identified and summarized below,

- Trading Algorithmic
- Pricing Option
- Analysis the Risk
- Evaluating the Market
- Forecast the Financial
- Managing the portfolio

4. Performance Analysis

In the performance analysis, Table 1 takes the Data value of 50 and 100 and determine the Precision, Recall, F1 Score and Data Support. Table 2, Determine the data accuracy.

Table 1. Attribute metrics based on mask value

Data_Value	Precision	Recall	F1-Score	Support
50	0.93	0.99	0.96	382
100	0.98	0.87	0.92	208

Table 2. Attribute metrics based on mask value (Accuracy; Macro Avg; Weighted Avg)

Mask_Value	Precision	Recall	F1-Score	Support
Accuracy	-	-	0.95	590
Macro Avg	0.96	0.93	0.94	590
Weighted Avg	0.95	0.95	0.95	590

5. CONLCUSION

Big data, machine learning, artificial intelligence, and cloud computing are driving the finance industry's digitization. Large firms are adopting these technologies to implement digital transformation, enhance profitability, manage losses, and meet consumer demand. Most companies are gathering new and useful data, but it's not clear how this will impact the financial industry. In the future, all financial services will be exceptionally sophisticated and will regard data with the same importance as the circulation of blood.

The findings of this study provide sufficient evidence to assert that big data has fundamentally transformed the financial sector, specifically in terms of immediate stock market intelligence that has altered trading and investment approaches, the identification and prevention of fraudulent activities, and accurate risk assessment using machine learning. These services have a positive effect by increasing sales and customer satisfaction, improving the process of making a purchase, making workflow more efficient and reliable, analysing financial performance, and managing growth. There are still some significant big data issues in the finance sector, even with these revolutionary service transmissions. The protection of data privacy is a critical concern for big data services. In addition to data quality, regulatory regulations are often perceived as significant challenges. Even though many financial products and services use data, big data and finance research is still young. Given this, it becomes useful to examine this work in order to locate potential areas for further research.

Future financial data management systems will need to overcome several types of technical difficulties while trying to realise the benefits of big data. Financial analysts and researchers should concentrate on the challenges of managing enormous data volumes when trying to produce game-changing solutions. Therefore, it is imperative to emphasise the importance of effectively handling extensive data sets for major corporations in contrast to small businesses. The persistent issue is that databases grow larger as industries grow. Managing such large data sets can be costly, and accessing them can frequently be extremely difficult. Oftentimes, small organisations and individuals lack direct access to large-scale datasets. Hence, future studies could focus on enhancing the accessibility of large datasets for small enterprises. It is crucial to prioritise the analysis of how big data impacts financial markets, products, and services. Further investigation is necessary into the security risks posed by big data in financial services. Furthermore, it is necessary to expand the formal and thorough procedure of implementing big data strategies in financial institutions. Specifically, the impact of big data on the stock market should be studied further. It is important to experimentally highlight the increasing concerns regarding big data in finance that were addressed in this study, in future research.

REFERENCES

[1] Loshin, David. Business Intelligence: The Savvy Manager's Guide. 2nd Editio, Elsevier Science, 2012, https://www.google.co.id/books/edi-tion/Business_Intelligence/L7SLNlS1ao8C?hl=id&gbpv=1&dq=business+intelligence+is&printsec =frontcover.

[2] Qalam, Yance Ibnu. "Hubungan Data Warehouse Dengan Business Intel-ligence Dan DATA." Kepo.Co, 2020, https://kepo.co/hubungan-data-ware-house-dengan-business-intelligence-dan-Data/.

[3] Wijaya, Wayan M. Teknologi Big Data: Sistem Canggih Di Balik Google, Ya-Hoo!, Facebook, IBM (Teori Hingga Tutorial). Nilacakra, 2019, https://www.google.co.id/books/edition/Teknologi_Big_Data/VeNDwAAQBAJ?hl=id&gbpv=1&dq=big+data+ada-lah&printsec=frontcover.

[4] Nordeen, Alex. Learn Data Warehousing in 24 Hours. Guru99, 2020, https://www.google.co.id/books/edition/Learn_Data_Warehousing_in_24_Hours/wgf9DwAAQBAJ?hl=id&gbpv=0.

[5] Saeed K Rahimi, F.S Haug. Distributed Database Management System–A Practical Approach. New Jersey: John Wiley & Sons Inc. 2010:1.

[6] M.T Ozsu, P. Valduriez. Principles of Distributed Database–Third Edition. New York: Springer. 2011:3.

[7] T.Connolly, C. Begg. Database System. A Practical Approach to Design, Implementation and Management. Fourth Edition. Essex: Pearson Education. 2005: 695.

[8] Kimball, Ralph., Ross, Margy. The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, Third Edition. Indianapolis: John Wiley & Sons Inc. 2013:38.

[9] Igor Mekterovic, Ljiljana Brkic, and Mirta Baranovic. Improving the DATA process of higher education information system data warehouse. Proceedings of the 9th WSEAS International Conference on Applied Informatics and Communications (AIC'09). Moscow.2009: 265-270.

[10] Vishal Gour, S.S. Sarangdevot, G.S. Tanwar, A. Sharma. Improve Performance of Extract, Transform and Load (DATA) in Data Warehouse. Int. Journal on Comp. Sci. and Eng. 2010; 2(3):786-789

[11] Abid Ahmad, Muhammad Zubair. Using Distributed Database Technology to simplify the DATA Component of Data Warehouse. Proceedings of WSEAS International Conference on Applied Computer Science (ACS'10), Iwate. 2010; 61-65.

[12] Tute E, Steiner J. Modeling of DATA-Processes and Processed Information in Clinical Data Warehousing. Stud Health Technol Inform. 2018; 248 204-211. PMID: 29726438.

[13] Sonali Vyas & Pragya Vaishnav. A comparative study of various DATA process and their testing techniques in data warehouse, Journal of Statistics and Management Systems. 2017; 20(4): 753-763.

[14] C. Adamson. Mastering Data Warehouse Aggregates. Solutions for Star Schema Performance. Indianapolis: Wiley Publishing Inc. 2006:20.

[15] W.D Back, N. Goodman,J Hyde. Mondrian in Action. Open Source Business Analytics. New York: Manning Publications Co. 2014:195.

[16] A. Meadows, A.S. Pulvirenti, M.C. Roldan. Pentaho Data Integration Cookbook. Birmingham: Packt Publishing, 2013:11.

[17] Rahm, E., H. H. Do, Data cleaning: Problems and current approaches. IEEE Data Eng. Bull. 2000; 23(4): 313.

[18] R. Bouman, J.V. Dongen. Pentaho Solutions: Business Intelligence and Data Warehousing with Pentaho and MySQL. Indianapolis: Wiley Publishing, Inc. 2009:160.