International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2021): 7.86

AI and Machine Learning for Predictive Banking: Infrastructure Challenges and Opportunities

Bharath Somu

Architect-I
Email: *bharthsomu[at]gmail.com*
ORCID ID: 0009-0008-6556-7848

Abstract: ***The volumes of data are unprecedented, both in numbers and in the variety of their formats. Such a vast amount of data has opened up many opportunities in many fields including finance. In the financial services space, the rise of high-frequency trading accounts for the explosion of data combining price, liquidity, and news indicators reserves to explore trading opportunities. On a longer horizon, the aggregation of individual investor behavior and their sentiment in association with liquidity, volatility and expected returns constitutes a potential pool for risk assessment, equity market studies, and financial marketing. Such big data also bring before any analyst a serious challenge, namely the question of how to process, analyze and extract useful information out of it in real-time at some point. A line of answers to this pressing question comes from the area of Machine Learning (ML) and more generally, Artificial Intelligence (AI) [1]. Banks face profound challenges from the new competitive landscape and the abundance of financial technology (fintech) businesses. Traditional financial institutions have become the victims of these "fast wolves," shrinking profit margins and losing clients. No matter what kind of bank it is, changes are inevitable. Banks are eager to boost the development of intelligence-related goods and services to deal with external competitiveness and the demand for internal transformation. The Industrial and Information Technology Ministry has stated that the finance industry's firms are encouraged to employ AI to improve digital levels [2]. Deep Learning (DL) has made tremendous achievements in image recognition, voice recognition, natural language processing, and many other fields related to social life. As the core technique for AI, researchers have been focusing their attention on Deep Neural Networks (DNNs) in data training. With the help of DNNs, a multitude of products have emerged. Commercial banks' data are continuous, dimension-large, and temporal-variability, which are more harmonious with the DL model. The complexity of the DL model affords commercial banks novel applications in the fields of risk management and intelligent services. Exploring the scenario application of DL technologies in financial businesses may become a sharp weapon for improving the intelligent service level.***

Keywords: AI in Banking, Machine Learning, Predictive Banking, Infrastructure Challenges, Banking Technology, Data Analytics, Real-Time Insights, Scalable Infrastructure, Cloud Computing, Financial Forecasting, IT Infrastructure, Banking Innovation, Big Data in Finance, Digital Transformation, Risk Prediction, Personalized Banking, Operational Efficiency, FinTech Solutions, Smart Banking, AI Opportunities in Finance

1. Introduction

The banking and financial services industry is facing significant change and opportunity, driven primarily by advances in artificial intelligence (AI) and machine learning (ML) . Regulatory pressures, stakeholder expectations, and digital channel expectations are just a few of the factors spurring this massive change effort. There is a race as a result of several new entrants joining the field with disruptive business models focused primarily on lending and payments. New competitive entrants are also emerging in analytics, but also infrastructure risk and compliance. The great promise of AI and ML is now viewed as an opportunity to deeply enhance revenue and efficiency both across processes and in front office applications, such as KYC, fraud detection, CRM, business intelligence, etc. A broad definition of AI would include technologies and techniques in machine vision, machine learning (supervised and unsupervised), natural language processing, robotics, and expert systems. A similar breadth is seen in ML but with the recognition that it is largely a search for patterns in data involving statistics, data mining, and optimization, with pattern recognition, neural networks, and one of many different implementations.

While there is consensus on the magnitude and meaning of the change and opportunity, there is also recognition that there are major challenges to leveraging AI and ML technology in the financial industry . An AI centered innovation strategy is seen as a major source of competitive advantage, but existing legacy processes, systems, and data constitute a risk of becoming technologically irrelevant or non-compliant. In lieu of totally begging off, the hope is to move aggressively but cautiously. While several academic literature reviews have examined AI and ML in banking, risk, and finance, practitioners are seeking a more realistic picture of the opportunities and challenges that establishes a meaningful narrative with actionable next steps. A sampling of the related research in six areas is included business models, risk modeling, credit screening, forecasting, revenue and processes. Literature reviews, too, have been done in related areas such as AI in FinTech and fraud detection. However, these and the like provide minimal investigations and little, if any, actionable insights on the industrial challenges with real world contextualization. Also lacking are comparisons drawn to other segments as now one of urgency.

2. Overview of Predictive Banking

Predictive Banking refers to undertaking banking activities by machine learning and AI led predictive modeling and statistical inferencing based predictions. In this section, key technology concepts underlying Predictive Banking are viewed, such as predictive learning models, predictive analytical frameworks and technology solutions infrastructure. Banking applications of predictive analytics such as smart onboarding, credit lending risk management, wealth and asset management, churn prediction and retention, macroeconomic scenario modeling, fraud detection and management, countering money laundering are articulated.

Figure 1: Artificial Intelligence in Finance

Predictive Learning Model. Predictive learning could be perceived as an umbrella term encompassing supervised, semi-supervised, and active learning. For a general description of customizing high-dimensional financial applications across traditional parametric approaches, machine learning tools, and hybrid solutions composed of a combination of the two, only supervised learning, i.e. the learning setting in which a set of input-output pairs (x, y) are given with input variables x and output variables y, is defined. Formally, the supervised predictive learning problem length p x and output length l y at time t, obtaining the best approximation according to a particular loss function C, and the model evaluation. A broad classification of predictive learning models according to the choice of architecture is proposed, namely, the parametric, nonparametric, and semi-parameteric models. Models in supervision learning examine learning through on potential estimators. Semi-parametric models e.g. Additive models consist of parametric and non-parametric components. High-dimensional and large-dimensional terms refer to p growing faster than n and diverging faster than n, respectively.

Predictive Analytical Framework. A high-level framework of predictive analytics comprising three stages: a model building stage, a process management stage, and a model deployment stage. The former describes data preparation and pre-processing, predictive modeling and evaluation. The latter emphasizes its service delivery and model diagnostics under a workflow process and its maintenance during the life cycle. A detailed overview of the phases comprising data acquisition, data engineering, exploratory data analysis, and business understanding within the model building stage to be considered when designing and constructing a new modelling solution.

Technology Solutions Infrastructure. Systematic taxonomy of technology solutions tools that can be used to construct banking predictive analytics are proposed, grouped into six categories: unified data staging environment, scalable data mining modules, library of predictive models or analytics, metadata, model management and governance, collaborative workspace liberated from hardware and OS dependencies.

3. Role of AI and Machine Learning in Banking

The traditional method of selecting and granting loans seems to be out-dated in this era of information explosion, and a stricter scrutiny of borrowers is required to minimize agency costs. Banks can adopt a stricter policy using a longer history of past defaults and detailed financial information. However, the annual data provided is also limited, and a conventional risk model is likely to be outperformed by a more complex deep learning model. This work focuses on predicting a borrower's first-time defaults within 3 years of bank transactions [3]. The dataset records the years of transactions and non-defaults, which is complicated temporal data, forfeiting covariate balance across time epochs and potentially leading to prediction bias.

Using conventional methods on the dataset usually leads to classification models trained with biased features. Since bank transactions are recorded only annually in the wide dataset, credit risk is predicted using the transactions and delivery kernel. For tabular data, the recurrent neural network has achieved impressive results, but it cannot take the order of transactions into account to predict bank defaults. Also, the dataset is imbalanced, which can miss the less than 1% labeled defaults in one time epoch. In this work, the temporal data problem of such imbalanced sequentially predicted defaults is formalized, and an elegant version of transforming it into hourly data is introduced, which can be applied to any other temporal datasets. It further focuses on the multitask learning framework to achieve robust deep learning on the transformed representation for credit risk evaluation. Besides, attention layers are integrated into the proposed framework to exploit the most informative features and long-term dependencies effectively.

In the last decade, machine learning has been empowering organizations in practically every profession with high-performance decision automation capabilities, making AI a key component of operational infrastructure. While AI tools are evolving rapidly across industries, banks anticipate that AI will dramatically alter their market positioning and operational models. Although the revolution of AI-driven competitive banks presents significant market opportunities, it also introduces severe and quickly increasing competitive challenges [1]. Therefore, banks urgently need to develop and implement an effective AI strategy and a robust and sustainable AI infrastructure, while leveraging opportunities across the complete AI value chain – from data sourcing, curation, modeling, and execution.

4. Infrastructure Requirements for AI in Banking

The possibility of utilizing AI+ML technology in banking decision-making scenarios extends across a wide range of circumstances. For example, intelligent credit assessment can provide tailored loans for users and dynamically adjust credit lines. AI+ML can create intelligent asset allocation models

Paper ID: MS2212141805 DOI: https://dx.doi.org/10.21275/MS2212141805

for wealth management, providing good investment suggestions, configuration options, and a reason index. In the finance market, price forecasting models can be built to calculate a probable price via suitable forecasting models. In contrast, fraud detection scenarios can utilize a history of transaction details to prevent fraudulent ATM transactions, suspicious credit card transactions, and more.

Originally, most financial institutions relied on traditional infrastructures, such as traditional relational databases, ETL tools, and data lakes. However, these infrastructures are not well suited for the AI era due to the following reasons. The amount of training data needed by AI+ML models must often reach terabytes. For financial institutions, data storage options must be expanded from traditional high-priced, high-end, performance, and reliability NAS/SAN to hybrid cloud-like storage devices that include NAS, OSD, and HDFS. These relatively cheap storage devices are scalable, easily managed and deployed, and have better I/O and performance.

Eqn.1: Predictive Performance Equation

$$A = M(D_q, V, R_t)$$

- M = Predictive ML model
- D_q = Quality of data (cleanliness, completeness)
- V = Volume of data
- R_t = Real-time data access
- A = Model accuracy

As for the training platform, financial institutions usually rely on traditional on-site top-shelf computational platforms. However, these platforms are too expensive and not flexible, and are not convenient for collaboration. Therefore, to build a more flexible, cheaper, and collaborative training platform, financial institutions ought to adopt partial public cloud combinations or mostly distributed cloud platforms.

For development environments, prior to the AI era, financial institutions usually relied on stable, on-premises desktop installations. However, due to diverging needs and rapid development of AI+ML techniques, a local environment cannot easily be shared. Additionally, IDEs need a constant or near-constant maintenance cost and effort. Therefore, to achieve a better development experience during the AI era, financial institutions should consider providing one or two cloud-like coding scripts and real-time execution sandbox platforms.

For evaluation platforms, prior to the AI era from either their deployment or auditive perspectives, the evaluation of models heavily relied on on-premises environments. However, with the advancement of model trading systems and collaborative environments, traditional environments became bottlenecks. Therefore, to achieve better evaluation experiences during the AI+ML era, financial institutions should consider gradually optimizing either their strategy-typing or transaction-testing environments and systems.

4.1 Data Storage Solutions

As machine learning models become more complex and data-driven inferences become more common, the quantity and variety of source data becomes wider and larger. The domain of application of Machine Learning and Estimation is seen wider than ever. Couple of use cases in a banking organization can be highlighted with focus on the computer science part of it. Use Cases under discussion in this sharing will be, Predictive Account Churn, Predicting Credit Card Holders Switching Transactions to e-Commerce Transactions Use Case 1, Predictive Model Challenging Infrastructure Options. Predictive Modeling Is Too Important to Fail Challenges. Too Many Types of Data Some Data Autodestructs. Infrastructure Options Provide complete and scalable data-life-cycle. Transform, Merge through SQL. Pre-aggregate to improve computational speed. 128-bit-algorithms to ask for a mean-age per month. Capture data-features first, then guess-predict. Ecosystem is similar to BigQuery. On-Prem Supercomputer Cluster running DB to orchestrate resources. DB already up-scaled to 40-node-relational-DB-cluster backend with untold TB of datasize is considered as future foundation.

As more modelling is done in-house, catching up action with crowd-sourced modelling is needed. Internally radiosystems are winning against vacuum systems. Saving money spent for crowd-sourcing has time-stamped. As data-to-knowledge conversion is needed in-house, off-the-shelf and keeping-in-house instruments are reconsidered. Competitive Tracking Devising homework POCs and subsequent in-house modeling in targeted domains. Closer Data Format Specifications Run SQL on Datasource. Better Machine Learning Examples Organize internal Machine Learning intelligence sharing workshops. Qualitative Data Quality assessment, winning against 60% nominal-data feed in the bootstrap of predictive modelling. Predictivity is mandatory to be automated. Data-Quality-Flyer before data-feed and other proper Test/Training/Validation-data partitioning middleware tools are considered for alternative action.

Two noteworthy tools to be used at almost every organization for predictive modeling are active learning and scalable infrastructures. Data Management and Data Preparation can be either a challenge or a bottleneck. Raw Data can be enigma data. Enigma remains enigmatical after scientist-“understanding” (seeing 0s and 1s). Data-Understanding Process is painstaking and takes a long time. The enigma information extraction could convey extra time-consuming and/or erroneous stages. You may smooth your raw data referring an off-the-shelf selecting instrument. There exist simulation-to-strategy rigorous frameworks to plan a Science-to-Society-wise (in-house predictive and prescriptive) conversion. There are also operational measures to predict the cost-iteration of a modelling pipeline-given data. Both are to save time and to maximize the output of predictive modelling [4].

4.2 Computational Power Needs

The explosion of interest in generative AI chatbot and intelligent assistant services based on large language models (LLMs) has been accelerated by the phenomenal success,

Paper ID: MS2212141805 DOI: https://dx.doi.org/10.21275/MS2212141805

market excitement, and appetite for follow-up products following the release of ChatGPT. ChatGPT demonstrated the value of conversational AI, which catalyzed massive interest and innovation across the entire AI startup ecosystem, further amplified by competitive entry by "big tech," a steady stream of investments, especially in generative art, text, and video generation, and a wider and faster dissemination of model architectures, software frameworks, and datasets [5]. There are over 170 AI-based personal assistant startups in the United States alone.

Since their invention, neural networks have allowed radical improvements in perception and decision-making tasks, but not in the sequence-to-sequence text generation task. Transformer models achieved state-of-the-art performance in practically every natural language processing (NLP) challenge. A sequential model-by-vector of a text-to-text format is trained [6]. The model is implemented with an N-gram convoluted attention mechanism for text document encoding and is finetuned with BERT-based span prediction modeling. The ability to condense and summarize with an extensive commentary or paraphrase on-demand led to the development of Condensing Recap Transformer (CoReth) and Gensim, both of which tame the ability to collapse a large text document into smaller text sent and generate abstractive or extractive summaries.

4.3 Network Infrastructure

Contemporary mass banking services depend on extensive computer networks that span multiple geographical locations. Hence, the telecommunications sector is immensely vital to the banking sector which relies heavily on complex data networks to deliver a wide range of banking services such as ATM services, Point of Sale (PoS) terminals, online banking, and mobile banking to their clients. As a result, banks needed a solid networking service that is globally scattered and fail-safe at the same time. Telecom companies must offer a resilient network infrastructure with redundancy while keeping their expenditure under control. Most of the telecommunication companies offer best-efforts networking devices that can support the leased line requirement from banks. The leased lines will be permanently reserved for the account of the banks and telecom service providers will charge the banks based on the consumption of X gigabytes per second as per the contract. In this service model, the telecom companies do not guarantee the Quality of Service (QoS) of the data machines that in turn may lead to a down-time of ATM and PoS terminals. This can result in huge financial loss for the banks.

To address the shortcomings of the telecommunication services, telecom industries have been relentlessly innovating to develop their infrastructure by adopting newer technologies, techniques, tools, and devices. The banks are also finding their solutions to enhance the safety and quality of their services and save their investment in bandwidth. The AI/ML is an umbrella term for a wide range of techniques and disciplines that are used to build and deploy intelligent systems. AI/ML models may be trained to achieve a specific objective based on data that describe their environment or train to act in response to feedback from that environment. The focus of this research agenda is primarily on machine learning, and much of the content here assumes that the reader is already aware of the types of problems that can be solved with supervised/unsupervised learning and the general concepts such as training, testing, and generalizability [7].

5. Data Management Challenges

Machine Learning has become a vital tool for predictive analytics in banking and finance, but its real-world deployment is often hampered by challenges associated with the management of data. Here, the focus is on three key challenges for predictive banking centered around data: (1) Managing data distribution shifts, (2) Dealing with missing values in the data, and (3) Detecting and handling anomalies in the data. With these challenges, several solutions that are applied to tackle these challenges in real-world predictive banking use cases are described. Recent developments in Machine Learning, particularly Neural Networks and Graphics Processing Units (GPUs), have greatly advanced the state-of-the-art for many challenging tasks such as image and speech recognition, machine translation and robotic control. However, the deployment of Machine Learning in real world industry applications has, in many instances, proven to be tedious and error-prone despite its enormous potential business insight. One such domain where ML has great potential business insight is predictive analytics in banking and finance, which entails customer churn, damage prediction, recovery prediction and many others.

In the banking industry, revenue is often recurrent. Therefore, gaining customers from competitors can yield large enrollment in revenue. However, maintaining these customers with retention opportunities can be a tedious task as consumer service needs change over time. This results in push and pull dynamics with miles of inertia and revenue lost being difficult to measure out of providing personalized service. However, in this setting, advances in AI/ML, particularly in predictive analytics, can offer real business insights in the form of lead targets and knock-off moves in the FinTech arena [8]. Predictions such as customer churn or lifetime value estimates can be incredibly valuable in making business decisions. But despite the real world side being vital, investigating in this setting remains small. As such, there is demand for a deeper understanding of the challenges to bringing predictive analytics to production in banking and finance.

Any data and associated processes in production pipelines are always prone to subtle shifts. In online banking for instance, onboarding a new group of customers is traditionally done in batches, but this is also seen in changes to customer underwriting criteria. In these cases, it is unreasonable to retrain a model from scratch as a new data source can require extensive preprocessing to be usable on the production pipeline [9]. To address distribution shifts, a more agnostic approach to data pipeline design is needed.

Figure 2: Future of Artificial Intelligence in Banking

5.1 Data Quality and Integrity

For banks that adapt data analytics solutions, the data quality and integrity is the foundation for the success of their analytics solution. If the data is populated incorrectly, or if it is outdated, broken, or conflicting, the analytics algorithm will yield incorrect and invalid insights or recommendations, which may even introduce a certain legal liability for the banks. As such, there exists a wide range of potential data quality-related risks that need to be carefully identified, and the risks must be mitigated adequately. Basic data cleaning and data validation must be ensured first.

Formal accountability from data owners should also be put in place to clarify the ownership of data and would limit moral hazards in data handling. In addition, banks currently face a mounting amount of external data flooding from social media, news articles, text entries, clickstream data, etc. It would be unfathomably and economically wasteful to include all external data even if the banks have the technical abundance to do so. As a result, the banks must prioritize the selection of the external data to be included in the analytic framework, and it includes but is not limited to the definition or typology of the external data sources, proper data pipelines and data lakes, robustness of the models reusing external data, etc.

The quality of extensive business data is often neglected in practice. For example, even though banks would credit check their borrowers, the best practices from the credit industry are still often neglected by banks. On the flip side, the emergence of AI has propelled banks to question the value of their internal business data, and how to extract value from them. Banks that adapt data analytics solutions are always accompanied by flexible, extensive, and readily available business data. Nevertheless, data governance is often disregarded, especially the access control of sensitive business data.

Eqn.2: Infrastructure Capacity Constraint Equation

$$I = f(C, S, \frac{1}{N})$$

- C = Compute capacity (on-prem/cloud)
- S = Storage capacity
- N = Network latency
- I = Infrastructure score for AI-readiness

5.2 Data Integration from Multiple Sources

The financial ecosystem has adapted to trends such as globalization, increasing regulatory constraints, and the rise of information technologies over the last two decades. With the evolution of practices and standards in a highly competitive environment, European banks are facing the challenge of disruptive new technology companies that generate competition and collaboration developments. Banks' infrastructure is challenged with architecture agilities, development approaches, and big data and cloud infrastructures. Before money circulation, native business models had to be analyzed and aligned with artificial intelligence-based prediction models. Natural Language Processing techniques have to be applied for pattern matching and extraction. Structured query language and machine learning libraries are applied toward widespread data sets from enterprise data warehouses and targeting services, demand deposit account, and transaction history data. Crashes in analytic capabilities and burst volumes have to be matched between number analyzers and data sources. Scalable cluster technology and specialized packages are implemented. For successfully producing new partnerships and practices, the quality of both partners needs to be analyzed. The jurisdiction, the body of laws, and the application of regulation, control, and supervision aspects have to be analyzed.

This section describes the processing pipeline and showcases an application against the sector to illustrate the insights generated for predictive banking. Machine learning methods for supervised learning use given knowledge for future prediction. Structured input data points and labeled outputs are needed. The data point should consist of historical data. Label similarity for multiple points is needed. Integrity of the pipeline resilience is designed over a relay pipeline structure. The design consists of processing individuals and combining core prediction ones and a modeling component. Individual ones combine raw input sources for wider data propagation. Since they may produce soft input changeable from time to time, they are put in the relay part. The wide propagation waits until a defined waiting time or data samples exceed before enabling the prediction ones at once. The relay speeds up processing demeanor without data shuffling. The design is elaborated for a high-velocity up to hundreds of thousands records per second use case.

By adopting forecasting methods, new and alternative indicators can be created with predictive power for several minutes to weeks ahead. These methods are combined with those capable of estimating measure quality scores for ranking chosen indicators against criteria, stakeholders, and business objectives. Adversarial approaches also enable obfuscation of quantity-based measures on public dashboards, preserving the utility of insights while resisting privacy attacks. The approach exploits the way mathematics can portray complex data and relations, combining signed directed graphs with functional analysis to enhance aggregation of millions of data, visual preprocessing, and adaptive clustering. The aim is a scalable and interpretable end report predicting distributions of relevant target variables using the best aggregated measures, able to work with widely available inputs but also trained with protected datasets.

5.3 Real-time Data Processing

Real-time streaming data analytics, processing and management is a new research emphasis in the big data era. Conventional data management technologies, which have been centered on the batch-wise processing paradigm, typically struggle with the continuous nature of streaming data. Business competition in almost all domains has been shifted from handling conventional static data to taking advantage of smart real-time streaming data operations. The proactive and predictive ability based upon real-time big data is crucial for each player regardless of scale or domain. Taking the rapid growth of real-time big data into consideration, common streaming data-processing methods suffer from numerous challenges in terms of performance, flexibility, usability, and extensibility [11].

Apache spark and its related frameworks are widely adopted on big data processing. For conventional batch or iteration-based big data operations, a high-level computing abstraction, such as Resilient Distributed Dataset (RDD) and Data Frame, is sufficient for data processing resembling Map-Reduce. For real-time streaming big data operations, the common high-level computing abstractions can only be served as on-the-fly primitive computations, as no reference result is available when a computation is triggered. Hence, fresh new methods and frameworks need to be produced for real-time streaming analytics.

Spark Streaming is the most popular big data platform as a Spark library for stream-processing applications. Although Spark Streaming is one of the most mature and battle-proven stream computing frameworks, it has a fundamental disadvantage in big data environments with its micro-batch processing schema [12]. A few stream-data-processing systems adopt and construct windowing structures to segment incoming stream data into fixed size, which are then processed in the same way as static batch data. However, this architecture limits realtime performance inherent in stream processing. The streaming big data might be a few hours, days, and even weeks older (and therefore no frequently-used policies). As on the one hand, intense edge service will proliferate on a large scale, much processing capability of data will shift from the cloud environment to the edge.

Figure 3: AI solutions are boosting value in banking

6. Regulatory Compliance and Ethical Considerations

The increasing use of AI in banking requires regulators to supervise complex and opaque systems. Regulators need to get ahead of banking innovation in areas such as new business models, new products and services, and new applications of technology. The mow-utilised AI risk appetite needs to be made visible with regulatory expectations in line with innovation. The development should stay aware of decentralised finance (deFi) regulatory arbitrage opportunities due to the different nature of token issuance and crypto-asset transfer. Building a market standard approach to a common financial services regulatory cloud enables regulators to “mind the gap” in understanding emerging systemwide risks [8].

The knowledge represents AI responses, algorithms and non-linear complex functions that are not understandable to humans. These responses create ethical dilemmas, including supervisory challenges. AI instruments rebooked on Basel accord standards may apply in ways that are inconsistent with regulatory expectations. External auditability is impossible. Explainability via locally interpretable model agnostic explanation must be supported by compliance by design, best practice and explainability tools [13].

Deep learning AI cannot be re-booked as a set of business rules. Regulation, supervision, audit and compliance, including queries on input to advanced AI, known as class-of-questions problems, are impossible. Inadequate compliance creates executive decision making without human oversight or intervention. Banks using AI tools may be required to provide “unknown known” explanations. Too many explanations may break the system and insert interpretability biases.

6.1 Data Privacy Regulations

As machine learning gains adoption to mitigate risks of lending and accurately predict defaults, compliance with beyond-GDPR regulations, data privacy regulations across the globe, requires investigation into the information used for building such models. Execution plans to build advanced credit risk models using AI and explainable model architectures are available but performance evaluation with respects to user privacy and dual effects on merging and prediction datasets has not yet been carried out [14]. The challenge with sharing information for estimating and evaluating model performance is greatest in neural network context due to prohibitive compute overhead. Financial services companies have eliminate ownership of third party service implementation due to product integration processes preserving the integrity of training data. Deep learning preserve approach to sharing a trained model has been proposed to address both storage and execution overhead however its performance has not been verified in a banking context [15]. They can offer and evaluate Payout, a privacy preserving product using crypto methods implemented natively on hardware for minimally invasive data collection and transformed into their own systems for archival storage. Payout maintains privately input datasets for MT processing, each preparation step returns analyzable and user inputindependent pivot tables. These can be queried as sites as long as ruled on the rows, columns, the output of computations can only be interpreted based on the dataset transactions. The research design and development such discretion can enrich application of information system on non-intrusive privacy preserving credit risk services using data monetisation approach and provide clearer business context to data privacy regulations across world.

6.2 Ethical AI Usage in Banking

Conclusion Ethics has always been imbedded into the realms of Banking: how money is created and distributed, what price it should be, who decides? How inclusive is the access to money? Who decides? How are risks shared? What happens when the system fails? How eyes-wide-open "safe" compliance is allowed to obfuscate a regulatory regime driven by deficiencies? AI also carries its own ethics, even though the ability to see and quantify it is nascent. Ethical AI means that AI should not replicate or exacerbate biases inherent in society; that it should be transparent and accountable; that decisions by AI should not be without human oversight. AI algorithms can produce astonishing yet unexplainable predictions. Knowledge and transparency on how an AI engine works differ radically from that of a credit decision rule bound by regulation and privacy. Whistleblowers can and should exist for large legacy credit scoring systems, but it will be much more difficult as being the smartest detective on "game of thrones" a mystery play. The Goldman likely fake move on a $100,000,000 transaction would never happen with an ordinary analytics set up, even if product defences could have failed. Thus, such a business practice could be troublesome for their own reputational, regulatory or even financial perspective. However, the same suspicious switch was adopted by the fresh COO who got hired by an Israeli Bank operating in Mexico that went bust of money laundering, inside dealing and bribery between cartels: path and cost of such an analytics implementation might be uncalculatable. Then again, AI adoption in Banks generates a dual use phenomenon. Even if banks can decide lapel pin "AI, ethics & compliance," security concerns will always come first. AI seems a game of "winner-take-all" competition for private Banks. New Business as a Service (BaaS) business models represent another category of AI adoption in Banks. Embedded banking can take a plethora of BaaS options. For me, the most disruptive service will be handled as a "retail" financial supermarket where any service will be offered to any user. Similar considerations as above apply to the necessary AI infrastructure: accessibility, affordance, adaptability, accountability.

7. Opportunities for Predictive Banking

However, while AI technology affects lending businesses for credit risk assessment, fintech firms have adapted and developed the technology and models continuously to suit different emerging markets. Emerging fintech firms can benefit from the knowledge and the comparative advantage of incumbents as machine learning (ML) lending models become modular. Massive adoption of online board loans stemming from the rise of clubs, groups, and aggregators has exploded in the credit space. This regulatory unintended consequence introduces a wild-west-like stage for commercial loan evaluation. As underwriting becomes an online commodity, future challenges for regulators and lending businesses include vetting the inflow of alternative data sources.

The emergence of a new generation of bank discipline problems is expected. Loan risk models are no longer based on categorical properties, but are shaped by the extreme million-dimension matrices summarizing user behavior data. To comply closely with applications, lenders may use sampling data internal to each bank's ecosystem. However, as the original basis decays, these new categories may solely describe past behavior instead of underlying risk factors. Therefore, conditional on the same observable features, loans approved with a similar probability may still have diverse performance data, leading to heterogeneous risk distributions. It may be unclear whether a machine learning model in itself, no matter how advanced, could ensure transparency in the event of disputes [3].

The rapid proliferation of ML models in lending has raised several regulatory concerns. Information asymmetries in user behavior elicitation raise questions about the fairness of imputing classification weights. It remains to be seen whether a biased weighting scheme would cast uncertainty over loan performance during litigation. The emergence of a new generation of bank discipline problems was also expected. ML threat models may be built to replicate and attack existing models, rendering monitoring of compliant models futile. How regulatory technology (regtech) can adapt to this disruption is an open question.

7.1 Enhanced Customer Experience

Digital banks use AI-based sentiment analysis to study users and classify them based on responses. User responses can be categorized as positive, negative, or neutral. Various features contributing to sentiment impact and response can also be tracked. Automatic screening and processing of comments and reviews reduce costs and enhance authenticity compared to manual analysis. Banks can utilize Machine learning to trace the history of users' transactions and maintain a personalized profile for every user. A customized FAQ with a solutions-oriented approach can classify users based on their interaction patterns; chat bubbles can guide users with different capabilities by identifying their sentiments. The bank can use a combination of various technologies to package a solution for their specific business requirements [1].

The bank stores past conversations of both users and business representatives. Using Natural language processing techniques, it can detect what users are more interested in. A state-of-the-art recommendation system can group multiple user personalities based on their interests and respond to questions accordingly. Adequate advertisement targeting in user language and writing style can encourage users. Banks need to consider the impact of chatbots on customer professionalism; a probabilistic theory of linguistic damage can guide the bank in implementing them. Large models need to be customized, improved, and fine-tuned before deploying them. Evaluating the performance of friendly, responsive, and kind chatbots is critical [16].

Thus new-age banking leverages data at the very core of their issuance. Data handling and comprehensive capabilities to get right insights lead to better decision-making. To be an identity maker in the banking system this requires a complete generational infrastructure stack, which covers Cloud, Data storage, AI, Reporting, Security, Connectivity, Performance Management, and Real-time engines which are supported by a robust network and controlled by a completely scalable IOT which can control all performance metrics. Cloud or Data Infrastructure is responsible for ingesting data and systems connected to basic structure, storage, and ETL. Security and financial services are like life and soul, for every piece of information AI and data, advanced security or cyber security should be at the utmost next level to maintain integrity.

7.2 Risk Management and Fraud Detection

The use of sophisticated analytics and others machine based learning implementation technologies in financial institutions has led to the possibility of better understanding the evolving behavior of the product and customer base. Accurate forecasts help businesses map out scenarios and strategies and better equip companies to respond and benefit from opportunities, changes, risks and future challenge. It also helps understand product mix to suggest formulating right strategies; foresee customer behavior; meet vulnerability; gauge market feasibility; enable proactive risk management etc. Certain business challenges such as upturn and downturn prediction on large datasets with millions of rows, predictive challenge on new applications with limited sample space, and just in time fraud detection are being addressed. The article proposes presented work in these domains.

Traditionally, the analytics methodologies adopted across the industry were limited to data mining and insights visualization. Now a day's empirical machine learning techniques are increasingly being embraced but the paradigm shift which increased data availability and volume due to change in transaction processing technologies and customer interaction mediums, matures them into machine learning (ML) based automated analytics [17]. The challenge for the financial industry from a predictive modeling perspective is the sheer size of the datasets it handles from a single customer interaction, although its transaction space is predominant to textual semi structured data, which may grow algebraically larger due to newer aspects being recorded presenting therein. The prediction horizon is also small and in fact the model may need to be built heuristically on new banks, hence the samples available for training may be limited, capturing only affirmative cases as these fraud instances increases. This challenge of needing enough evidence for inferencing etc. for latest behavioral pitch by reducing data induces the need of speedy analytics.

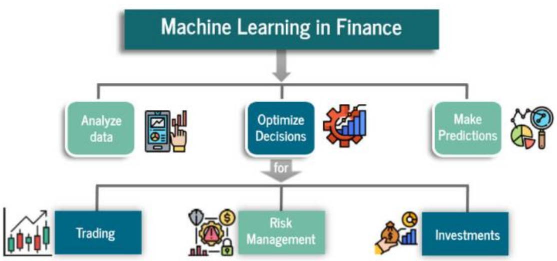

Figure 4: Machine Learning in Finance

7.3 Personalized Financial Services

Personalized FinTech services have emerged as one of the hottest topics in the finance industry in recent years. Major technology companies and hundreds of FinTech startups are vying to exploit customer data for personalized services. Traditional financial institutions, however, have been slow to adopt big data technologies and efforts to offer customer-oriented personalized financial services. In contrast to the technology giants, at least in the financial services sector, traditional financial powerhouses seem to be at risk of having an 'unfair advantage'. FinTech companies are inherently good at building strong customer relationships, because of extensive experience in personalized services. Traditional financial institutions have rich, long-established customer bases and relationships, but these are not easy to leverage into personalized services [13]. Today, with increased competition from internet companies in China, traditional financial institutions are contemplating how to strengthen their knowledge of customers and increase the value and stickiness of the relationships by offering personalized financial services. People in both industrial and academic circles have realized that "big data" holds great potential in the personalized services arena, especially among financial institutions, to improve the accuracy and usability of financial data. Traditional financial institutions have accumulated and retained unique data for every trading counterparty or corporate client over many years or even decades. With the development of big data technologies, however, drawing on an extensive pool of historical records and understanding the new real-time transactions have become feasible for constructing personalized financial services. To balance potential compliance risk, unlike user profiles for advertisement recommendation, financial data for personalized services have to believably characterize customer behaviors and intent without tampering with privacy. By examining only the modeling variables and strictly enforcing confidentiality control, new personalized banking services have been built. However, traditional data mining solutions face enormous infrastructure challenges due to both the huge data scale and high concurrency intensity of new financial data. These unique challenges drastically enlarge the engineering effort and total cost of safeguarding, operating, and maintaining the system. The state of the art in big data has seen much progress; however, the requirements of the initialized and evolving software and hardware are different from statistics machine learning and general IT systems [1]. Thus, many details are not yet fully covered in previous literature.

8. Case Studies of Successful Implementations

Advances in artificial intelligence (AI) methodology and technology, as well as their implementation and adoption in businesses and organizations, are badly neglected in the financial sector. However, this sector regards technology as a means of processing or storing data. The third subarea of focus investigates the concept, technology, implementation, and application of AI and its subheading technologies in various functions of banks. The subarea of task-based prediction is primarily explored using big data, with model construction and analysis being the main research work. Finally, the last part demonstrates how AI and its subheading technologies can be applied to businesses and organizations in the financial sector. Commercial banks are chosen as the research background here. Commercial banks and their dispersion and mass data, multi-scenario application modes, and at least two transformation motivations—external competitive pressure and internal transformation needs—make the banking industry a suitable domain to introduce AI technologies.

As a prominent bank in India, the State Bank of India (SBI) has 450 million subscribers, 1,200 million transactions, and 60 lakh complaints each month. As a result, it employs AI-powered tools like chatbots to assist clients in obtaining banking services. SIA is a noteworthy initiative that is being used by the SBI for task automation in need-oriented chats. The study focuses on determining how well SBI-SIA works as an AI-based virtual assistant capable of doing novel tasks. Adjustable models would cope with a complaint like, "Will icici bank be open tomorrow" and lead clients smoothly from intent recognition and parameter extraction to be liable. Previously categorizing active alerts into attention, information, transaction, or service was considered manual effort. An NLP-based model can now automatically classify approximately 3 million notifications into 4 labels with a precision of 78% or above [1].

Furthermore, task automation according to need is based on a hybrid approach. With confidence scores above 0.8, task-oriented QAs would pass to the banking QA, which with NLP would pass to RGSASK. Another technique that will streamline banking procedures is NATOTR, which is built with clarification questions [16]. This would automate transactions using the existing system. However, the target audience would require engaging interfaces, such as voice and text chat. Although commercial banks were among the earliest adopters of non-banking chatbots, innovative ideas are progressively being implemented. Additionally, alternative forms of AI, such as squads, can be considered for digital banking. Implementing bot-based engagement engines would help gather off-site customer insights, while digital onboarding tools would make it easier for new customers to open accounts.

8.1 Large Banking Institutions

Large banking institutions understand the advantages of predictive banking. They have made strategic investments in using artificial intelligence (AI) and machine learning (ML) for customer acquisition, advertising, and pricing. However, large banks must overcome data and analytics infrastructure challenges to realize the full promise of AI and ML in risk mitigation and modeling future states that could hurt profitability or customer participation. Top management needs to understand the importance of data and analytics infrastructure in making AI and ML scalable across the enterprise. Instead of purchasing platforms, banks need to build flexible enterprise infrastructures for data and analytics. This requires a long-term commitment of strategic direction, funding, and resources [1].

Historically, financial institutions used legacy environments to consolidate data and create an enterprise-wide repository. This became a source of record and data for managing operations, reporting, and regulatory compliance. Over time, more advanced BI tools were acquired for data access, reporting, and analytics. These early infrastructures served banks well and created investments that are well exploited. However, they have become an impediment. Traditionally, banks relied on batch architectures where data were cleansed, standardized, and stored in a data warehouse. BI tools relied on pre-calculated answers using a centralized data mart. While these approaches made sense in simpler environments, they are no longer effective.

The demand for answers has dramatically expanded in frequency, speed, and breadth. Performance management answers need to be generated constantly. As access to raw data has broadened, the use of those data has outstripped the controls that were implemented. Regulations increasingly require firms to store information for a longer time to reconstruct their deeds. Data are dusky, disparate, and at varying states of cleanliness, making them difficult to extract from and use. Data environments also have deep analytic and IT dependencies that are difficult to change. In the changing landscape, the importance of data and analytics is increasing. Consequently, banks are expending significant funds developing predictive banking capabilities. As opportunities and demand for analytical answers expand, banks are buying a set of analytics-specific platform software applications that perform tasks such as data preparation, model development and management, and real-time algorithm execution. These investments are not scalable because the requirements of analytic environments are fundamentally different than for data and reporting. Banks need to rethink their data and analytics infrastructures to produce scalable, effective predictive banking environments.

8.2 Fintech Startups

AI startups are growing exponentially. Companies ranging in size from hedge funds to the financial department of Tesla are implementing AI technologies. A slew of articles are published every day, heralding the forthcoming AI revolution in all aspects of human life. Many finance companies are minting money on investments in AI-powered startups. New blog posts and social media videos appear almost every day about which AI company to invest in and how quickly that company will turn a dime. A plethora of AI startups are entering the market every month. Anything from a small feature to tons of data is augmented with the word "AI" in whatever product they advertise. Keeping these facts in mind, understanding what is true in this mess is currently one of the hottest issues in finance [19]. Traditional finance is a convoluted organism with structured compartments populated by talented, motivated humans. On the other hand, FinTech is chaotic, unstructured, and packed with great expectations—often exaggerated. Enhanced market efficiency and reduced costs of trading or banking have been topics of heated debate ever since algo-trading first appeared on the radar. However, the increasing role of AI in banking is often neglected. The banking sector of finance is competing, or rather cooperating, with the exploding world of FinTech. There are thousands of startups out there vying to get a slice (or bite) of the financial pie. It is a pale shadow of traditional banking but alive, virulent, and unregulated. The perception of personal finances is rapidly changing—ex-military hedge fund managers are no longer the ones to watch, and investment clubs with no financial know-how are ballooning on social media.

9. Future Trends in AI and Machine Learning for Banking

Artificial intelligence (AI) and machine learning (ML) are increasingly becoming part of everyday life. As a result, predictive banking holds enormous opportunities for improved consumer servicing and behavior analysis, fraud detection and risk assessment, and regulatory compliance. This text sheds light on the infrastructure challenges and opportunities of adopting predictive banking. It also provides important theoretical and personal insights for future research projects. AI and ML are increasingly popular in the financial services sector for many purposes, including consumer servicing and behavior prediction, fraud detection, risk assessment, regulatory compliance, and trading [1]. Recent corporate case studies shed light on the technological aspects, opportunities, and limitations of predictive banking as a new service sector that combines data analytics and financial services. Substantial opportunities await developments in this area, but at least as many significant challenges will need to be addressed. It represents a significant theoretical contribution from an information technology (IT) innovation perspective, focusing on the perceptions of non-expert users of predictive banking and the associated infrastructure challenges [20]. For many financial services and sectors, integrating advanced data analytics into operations is no longer a novel research topic. For instance, research on analytical banking, HR analytics, and risk analytics has been conducted for some time and contributes significantly to academic knowledge. In contrast, predictive banking as a new domain or sector that includes aspects of both data analytics and financial services has won little attention so far. Yet, it presents considerable academic and practical opportunities and challenges. To seek relevant literature and published articles, two keywords were used: “predictive banking” and “predictive bank” with an asterisk for derivable forms. Searches were performed in four databases: the Financial Times, Social Science Research Network, and the Web of Science. Academic journal articles, industry reports, and practitioners’ comments were covered from these sources. In addition, empirical insight was gathered from the personal professional experience of at least ten interviews with financial services executives and IT experts.

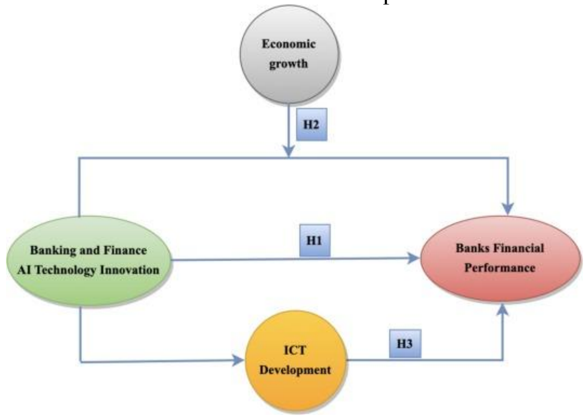

Figure 5: Transforming banking

9.1 Emerging Technologies

In the world of personal finance, no development has been as impactful as the advent of smartphones. Being able to monitor and set achievable savings goals from the comfort of one’s bed is no small feat. However, equally astonishing are the rising numbers of payment wallets, payment apps and gateways, and many other new business models in retail banking. These developments have given customers options that were previously not imagined. Competitors from various industries are all working to make life easier for consumers and vice versa. With advances in technology, the banking industry must simultaneously invest in new infrastructure and take a closer look at the legacy systems that are already in place [13]. Outdated legacy systems and deeply ingrained legacy practices can be a source of friction in the finance and banking business. Too many systems ensure efficiency; yet information remains fragmented across siloed organisations. This fragmentation is a huge hindrance, as is the need to comply with existing processes, which often results in lost potential and opportunities. Inadequate and incompatible systems are causing frustration for customers and employees alike. Furthermore, even if they don’t create friction, these systems greatly inhibit banks’ ability to innovate when it comes to competing with tech firms. Hence, banks face the challenge of aligning their existing systems with the new emerging technologies.

Financial inclusion is an important agenda for policy and decision-makers globally. Technology financing through crowd funding platforms and digital contract enforcement, lending through peer-to-peer lending platforms and payment banks have emerged as alternative solutions for less privileged users. However, given the choice between a new bank and a technology solution for transition, consumers still prefer to persist with the existing form of transactions. It is important to ensure the effective integration of technology solutions and delivery through the wide network of existing banks, non-banking financial companies, and micro-finance institutions. Digital transformation is not only concerned with applying existing processes to a digital platform but also about transformation in thinking. Digital banking requires thinking beyond banking and regulators, central banks, and banking duty-holders must move together towards an integrated approach for regulating FinTech enterprises [2]. The financial industry has been quick to adopt cloud services to cut costs whilst meeting rigorous compliance demands. However, emerging technology is still viewed with some caution. In the face of accelerating change driven by new risk factors, heightened regulatory complexities, and digital disruptions, banks are seeking to identify a pragmatic path to transformation. This calls for a major reassessment of new technology adoption by institutional lenders and regulators alike to establish which technologies present genuine promises of business growth, and discuss which emerging technologies such as blockchain and distributed ledgers, P2P lending, regulatory technology, smart contracts, payment gateways, and augmented reality are most poised to displace or extinguish traditional banking’s existing sources of revenue.

9.2 Predictions for the Next Decade

Artificial intelligence (AI) technology is changing people's lives and is widely used in many industries, particularly in financial technology (fintech). There are many forms of instrumental AI technology being used, and numerous consumer-facing products are being reported widely. Further, many also believe that predictive banking capability, which is generally understood as a financial service that anticipates a customer's need instead of the customer having to trigger it, is no longer sci-fi fantasy but will be a reality. Giants such as Amazon and Google have been engaged in this area to challenge traditional banks since several years ago, and traditional banks have also started to explore this capability in financial services. Nevertheless, predictive banking capability will be new for most traditional banks, especially in the operationalisation of this capability. In particular, the wider use of machine learning (ML)-based predictive banking is associated with huge technology infrastructure challenges for traditional banks. In light of this observation, this article aims to address these technology infrastructure challenges and provide opportunities on how technology infrastructure can take advantage of predictive banking.

Banks, as highly regulated financial institutions, are very different from non-financial institutions. The internal and external factors threaten and hinder their adoption of ML-based predictive banking capability. Traditional banking infrastructure is also found to be a core challenge hindering the development of predictive banking capability. This consists of the poor internal data landscape, difficult integration of non-banking data and legacy systems and insufficient internal ML research and development capabilities. The encouraging signs on cloud technology and GTP-3 deployment are also observed. This article further sheds light on opportunities on how banks can overcome technology infrastructure challenges and embrace predictive banking capability [3]. SNS banks are typically digital leaders also with best-in-class technology infrastructure, which serves as the strong foundation on which predictive banking capability is constructed. The collective technology stack of cloud technology, the architecture compatible with MLOps, robust data integration, enrichment and governance tools is suggested to be leveraged to develop predictive banking capability incrementally. Further, an ambition roadmap for rolling out ML-based predictive capabilities in various business lines is presented [1].

10. Challenges in Implementing AI Solutions

AI has proven to be a significant opportunity for the banking industry, offering innovative ways to gain an informational edge and improve efficiency. However, the more AI systems enter into critical processes of the financial industry, the more essential it is to have monitoring methods that ensure their robustness and compliance. Regulatory bodies have begun the initial steps of developing AI model governance regulations and stressed the importance of implementing monitoring solutions compliant with regulations. The nature of AI systems presents a number of challenges for their model governance. It is unclear whether the existing monitoring practices for conventional statistical models are also effective for complex AI models like deep neural networks [8]. Beyond the existing complex evaluation methods, not fully developed or explored one-size-fits-all solutions might be subject to regulatory scrutiny. As bank financial monitoring is the most heavily regulated industry, financial institutions will likely adopt the strongest monitoring practice across different industries. Those banks and financial institutions early in the adoption of generative AI models face a contradictory challenge. Existing monitoring methods like documenting datasets and technical papers which are mandatory required by financial regulators, are impossible to implement due to the growing complexity of the models, the exponential increase of training dataset sizes and the EULA. Beyond the need of AI monitoring practices, almost all banking institutions lack the infrastructure that is necessary to implement the AI model governance framework, namely model inventory system, logging equipment and even recording pipelines. AI Model monitoring is system-wide, infrastructure supported and requires the inclusion of various databases such as the model performance database, the monitoring report database and the model configuration change source. Those databases are not required for conventional statistics models. Most current financial institutions' practices are simply a series of procedures implemented by model owners without any system support or collaboration, making them difficult to be consistent, repeatable and efficient. Team structures and model owners also clash with systems implemented in Banks to use ML because controlling teams capturing business users via an extensive process for datasets will not be compliant with rapid changing businesses that require on-the-fly data preparation.

10.1 Skill Gaps in the Workforce

Despite increasing interest in AI technologies for predictive banking, tools are oftentimes not utilized as intended. This is especially concerning in the case of predictive tools that rely on trained ML models, where human analysts are supposed to build new models, enhance current ones, and monitor which models are applied for which tasks. The training of ML models typically requires individuals with computer programming backgrounds who are familiar with algorithms and data wrangling, yet a large share of analysts still employs traditional statistical tools.

In a survey of the most widely used data mining tools by bachelor's-degree-holding analysts, SAS Enterprise Miner was the most commonly used tool, with 19% usage share. This was followed by the free software R (10%), SPSS Modeller (8%), and Weka (6%). If university education is examined, the most widely used tools are SAS Enterprise Miner (34%) and SPSS Modeller (18%). Additionally, the usage of non-academic statistical packages by Ph.D.s not only remained below that of traditional statistical packages, especially SAS and SPSS, but showed little to no increase in the past decade. Hence by 2010, well over 90% of users of non-academic statistical packages were judged to be completely unaware of their availability. These findings suggest what a research institution should prioritize in addition to tool development. It is suggested to conduct programming boot camps tailored to analysts, covering SQL data wrangling and Python R/MATLAB script writing tailored to building ML models [9]. Further, the original ML framework points towards two additional initiatives for

Paper ID: MS2212141805 DOI: https://dx.doi.org/10.21275/MS2212141805

maintaining institutional knowledge in ML tool use, with one being a public-access knowledge base for storing ML tool use cases, underlying data, and performance metrics, and the other being a subscription-tiered consulting service that pairs expert ML practitioners with analysts seeking assistance.

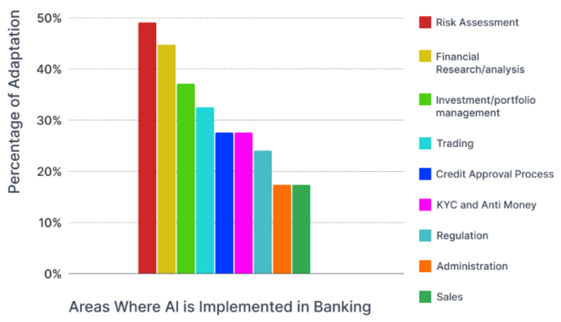

Figure: Financial sector with the integration of AI in banking

This discussion focuses on skill gaps concerning the use of AI tools. On the other side of the AI knowledge divide, there exist opportunities for banks because AI tools predictive privacy and interpretability. Researchers collaboration types are categorized within a mutual engagement framework ranging from teaching-collaboration to problem-collaboration [22]. Problem-collaboration refers to a researcher being presented with a problem and producing a solution, and teaching-collaboration refers to a researcher proactively leading a project with scientists/analysts as bystanders.

10.2 Resistance to Change

All that is needed to extend the use of AI in financial regulation and supervision, essentially that of the private sector currently in use, is some technical know-how. This poses problems of its own, but there is no clear reason why it cannot be done. What is missing is that know-how and the data-sharing efforts. All AI banks use the benefits of technology, everything else is the same. Technologies can be used and developed without ethics, and without government regulation. The authorities have some pre-existing data problems of their own that come on top of that [20].

Money Laundering Reporting Officers (MLRO) would rather turn a blind eye to a customer than treat them with suspicion and end up in the papers. The same with Democratic, Non-for-Profit, and Social Banks. All MLROs across banks could share information, but doing so would get on everyone's bad side. No one trust, especially the high net worth individuals. For every due diligence requirement placed on clients, there is a work-around. Alternatively, banking engagement could turn fractious, and banks lose clients [13].

Customer profiling and loan approval decisions are similar. Banks would benefit greatly from ubiquitous and granular customer information categories. This could have knock-on effects, but it is difficult to move the information. Instead, competition is maintained. Challengers, neo-banks, do not have legacy systems and technology. Customers can sign up to an app quickly and be cleared to make faster transactions, obtain cheaper loans, and get survey bins across a much greater area. Some homogeneous equity projects existed. With those there are greater costs in terms of waiting times, and completed transactions before the purchase. It is starker in the public sector, as authorities have ignored fast-transacting AI systems.

Eqn.3: Model Latency Equation (Challenge)

$$\Delta L = L_m - L_t$$

- L_m = Model inference latency
- L_t = Tolerable latency for user experience
- ΔL = Latency gap

10.3 Cost of Implementation

Big Data and Data Science are two terms that go hand-in-hand in modern times. Big Data refers to massive datasets that cannot be analyzed using traditional methods or on conventional platforms. These datasets contain large volumes of structured, unstructured, and semi-structured data. When using technologies that are capable of handling this large volume of data, it must be stored in Big Data's best practices formats so that data scientists can retrieve the dataset easily and proficiently to run faster and better algorithms on it. Data Science is a combination of skills, tools, and technology to analyze big datasets and create insights or lift or boost the business. Data Science is the shadow of Big Data and relies on machine learning and statistical models to learn from the data and make predictions or forecasts taking into consideration all known variables [23]. Banking is one of the industries that has been revolutionized by the application of data science and machine learning models. After the global financial and economic crisis in 2008, banks have ramped up efforts to address the onslaught of new legislation to counter issues that may potentially result in the next explosion of subprime mortgages and toxic securities. Volumes of transactions have prompted the rise of Big Data [24]. Banks have specialized units to record, track, and manage all risks a bank has exposure to. Large- and mid-size financial institutions have specialized teams for risk data aggregation (RDA) and risk reporting (RR) to ensure that available data is sufficient, relevant, accurate, and timely for risk analysis purposes. The RDA and RR infrastructure can provide a broader view that crosses risk types. However, there are trade-offs among risk types, risk areas, and risks in terms of infrastructure. Clarity and visualization of the trade-offs would help senior management make informed decisions on priorities and investments given the limited budget.

11. Collaboration Between Banks and Tech Firms

The profound development of information technology has surged in banks and other financial institutions, leading to data to be considered as the fundamental competitive resource to these kinds of institutions. Meanwhile, if banks did not grasp AI technology, they would lose big data control, and information banks run the risk of turning into traditional monopolies. The development of internet finance is like a storm that challenges conventional banks. Immediately, conventional banks began to realize the importance of AI technology, and gather funds to recruit data experts, access to large data platforms, and develop AI technology through mergers and acquisitions, joint ventures and self-development [1]. In these efforts, CMB gets millions of praises as it is the first to establish its intelligent banking. CMB has constructed its infrastructure including computing platform, data warehouse and deep learning platform, and has provide deep learning service on core services such as credit card underwriting, anti-money laundering, risk controlling and intelligent marketing. However, as AI technology has playing a more critical role after its establishment, there is also inevitable challenges. The out-migration of core talents from the core risk control department causes intangible loss for the bank. Substantial resources have been invested to build deep learning infrastructure, but the return will only show in three to five years [16]. It is fair to say that it means the limit of interdisciplinary results exploitation and inequality of resources distribution between these two departments. Although there is an increasing amount of cases of FinTech, the research using the business and technology factors is still scant, especially for educational institutions. This work assesses the financial activities in terms of FinTech and innovation domains by taking an overall look at banks' network and investigating the factors that drive the growth of FinTech banks. Given the powerful effects of these variables, policy makers and banking authorities can be better guided to establish Fintech-friendly environment and more supports can be provided to encourage innovative development of other banks. Generation of novel processes, services or products that is gained by analyzing business data is labelled as innovation.

12. Conclusion

The foundation layer houses the cloud infrastructure on which other layers reside. The cloud infrastructure can either be public or private. The foundation layer consists of key cloud ecosystem players like vendors, financial services institutions, third party aggregators, and regulators. The foundation layer plays host to multiple core banking systems. Core banking systems should be deployed on hypervisors that can abstract their workloads across cloud infrastructures. Ecosystem players like financial institutions and third-party aggregators offload variable workloads to the public cloud, generating savings on cost, energy, and carbon footprint build. The foundation layer is also home to a cloud-native API gateway to securely publish, discover, and consume APIs. A services mesh technology running on a platform is embedded to determine smart routing and access control to microservice endpoints. This aids in subtle service interoperability which increases ease of integration points to AI-enabled microservices. Services cost and SLA control policy automation using a cloud cost-management system includes cloud cost visibility dashboards and policy automation tools. The centre of enablement is an ecosystem of cloud-native tools to quickly develop and deploy AI-enabled microservices. It is an environment for knowledge workers. Cloud-native tools include no-code/low-code tools that provide quick wins. AI-enabled microservices can address easy use cases with low-hanging fruit. Infrastructure comprises container orchestration tools and DevOps tools to enable CI/CD pipelines. CI/CD tools are needed to deploy scalable, observable, and manageable microservices. Skilled resources include data scientists and software engineers who understand software development best practices. The marketplace hosts every AI-enabled microservice to be easily reused. Built microservices can be validated by other financial institutions at the preproduction level under strict protocols. AI-enabled microservices can be deployed in a multi-cloud setup, avoiding vendor lock-in and promoting resource efficiency.

The compliance layer enhances model transparency, explainability, and robustness. Continuous data feeding, retraining, and performance monitoring determine controllability. An adversarial detect, retrain, and revalidate approach triggers compliance checks when performance drops. Regulators applying adverse attacks on shielded models are possible. Challenge-response APIs for end-to-end model explainability and regular performance reporting are available to regulators on-demand. The compliance framework impacts microservices/financial products as little as possible. Conversation-flow history maps inputs, system transformations, and outputs for problem, causal, and risk-centric traces. Annotated regulatory reports are automatically generated. Trained domain experts perform on-demand thorough investigations into on-coverage/false negatives, and off-coverage triggers.

References

[1] Kommaragiri, V. B., Preethish Nanan, B., Annapareddy, V. N., Gadi, A. L., & Kalisetty, S. (2022). Emerging Technologies in Smart Computing, Sustainable Energy, and Next-Generation Mobility: Enhancing Digital Infrastructure, Secure Networks, and Intelligent Manufacturing. Venkata Narasareddy and Gadi, Anil Lokesh and Kalisetty, Srinivas.

[2] Pamisetty, V., Dodda, A., Singireddy, J., & Challa, K. (2022). Optimizing Digital Finance and Regulatory Systems Through Intelligent Automation, Secure Data Architectures, and Advanced Analytical Technologies. Jeevani and Challa, Kishore, Optimizing Digital Finance and Regulatory Systems Through Intelligent Automation, Secure Data Architectures, and Advanced Analytical Technologies (December 10, 2022).

[3] Paleti, S. (2022). The Role of Artificial Intelligence in Strengthening Risk Compliance and Driving Financial Innovation in Banking. International Journal of Science and Research (IJSR), 11(12), 1424–1440. https://doi.org/10.21275/sr22123165037

[4] Komaragiri, V. B. (2022). Expanding Telecom Network Range using Intelligent Routing and Cloud-Enabled Infrastructure. International Journal of Scientific Research and Modern Technology, 120–137. https://doi.org/10.38124/ijsrmt.v1i12.490

[5] Pamisetty, A., Sriram, H. K., Malempati, M., Challa, S. R., & Mashetty, S. (2022). AI-Driven Optimization of Intelligent Supply Chains and Payment Systems: Enhancing Security, Tax Compliance, and Audit Efficiency in Financial Operations. Tax Compliance, and Audit Efficiency in Financial Operations (December 15, 2022).

[6] Mashetty, S. (2022). Innovations In Mortgage-Backed Security Analytics: A Patent-Based Technology Review. Kurdish Studies. https://doi.org/10.53555/ks.v10i2.3826
[7] Kurdish Studies. (n.d.). Green Publication. https://doi.org/10.53555/ks.v10i2.3785
[8] Motamary, S. (2022). Enabling Zero-Touch Operations in Telecom: The Convergence of Agentic AI and Advanced DevOps for OSS/BSS Ecosystems. Kurdish Studies. https://doi.org/10.53555/ks.v10i2.3833
[9] Kannan, S. (2022). AI-Powered Agricultural Equipment: Enhancing Precision Farming Through Big Data and Cloud Computing. Available at SSRN 5244931.
[10] Suura, S. R. (2022). Advancing Reproductive and Organ Health Management through cell-free DNA Testing and Machine Learning. International Journal of Scientific Research and Modern Technology, 43–58. https://doi.org/10.38124/ijsrmt.v1i12.454
[11] Nuka, S. T., Annapareddy, V. N., Koppolu, H. K. R., & Kannan, S. (2021). Advancements in Smart Medical and Industrial Devices: Enhancing Efficiency and Connectivity with High-Speed Telecom Networks. Open Journal of Medical Sciences, 1(1), 55-72.
[12] Meda, R. (2022). Integrating IoT and Big Data Analytics for Smart Paint Manufacturing Facilities. Kurdish Studies. https://doi.org/10.53555/ks.v10i2.3842
[13] Annapareddy, V. N., Preethish Nanan, B., Kommaragiri, V. B., Gadi, A. L., & Kalisetty, S. (2022). Emerging Technologies in Smart Computing, Sustainable Energy, and Next-Generation Mobility: Enhancing Digital Infrastructure, Secure Networks, and Intelligent Manufacturing. Venkata Bhardwaj and Gadi, Anil Lokesh and Kalisetty, Srinivas, Emerging Technologies in Smart Computing, Sustainable Energy, and Next-Generation Mobility: Enhancing Digital Infrastructure, Secure Networks, and Intelligent Manufacturing (December 15, 2022).
[14] Phanish Lakkarasu. (2022). AI-Driven Data Engineering: Automating Data Quality, Lineage, And Transformation In Cloud-Scale Platforms. Migration Letters, 19(S8), 2046–2068. Retrieved from https://migrationletters.com/index.php/ml/article/view/11875
[15] Kaulwar, P. K. (2022). Securing The Neural Ledger: Deep Learning Approaches for Fraud Detection and Data Integrity In Tax Advisory Systems. Migration Letters, 19, 1987-2008.
[16] Malempati, M. (2022). Transforming Payment Ecosystems Through The Synergy Of Artificial Intelligence, Big Data Technologies, And Predictive Financial Modeling. Big Data Technologies, And Predictive Financial Modeling (November 07, 2022).
[17] Recharla, M., & Chitta, S. (2022). Cloud-Based Data Integration and Machine Learning Applications in Biopharmaceutical Supply Chain Optimization.
[18] Lahari Pandiri. (2022). Advanced Umbrella Insurance Risk Aggregation Using Machine Learning. Migration Letters, 19(S8), 2069–2083. Retrieved from https://migrationletters.com/index.php/ml/article/view/11881
[19] Paleti, S., Burugulla, J. K. R., Pandiri, L., Pamisetty, V., & Challa, K. (2022). Optimizing Digital Payment Ecosystems: Ai-Enabled Risk Management, Regulatory Compliance, And Innovation In Financial Services. Regulatory Compliance, And Innovation In Financial Services (June 15, 2022).
[20] Singireddy, J. (2022). Leveraging Artificial Intelligence and Machine Learning for Enhancing Automated Financial Advisory Systems: A Study on AIDriven Personalized Financial Planning and Credit Monitoring. Mathematical Statistician and Engineering Applications, 71 (4), 16711–16728.
[21] Paleti, S., Singireddy, J., Dodda, A., Burugulla, J. K. R., & Challa, K. (2021). Innovative Financial Technologies: Strengthening Compliance, Secure Transactions, and Intelligent Advisory Systems Through AI-Driven Automation and Scalable Data Architectures. Secure Transactions, and Intelligent Advisory Systems Through AI-Driven Automation and Scalable Data Architectures (December 27, 2021).
[22] Sriram, H. K. (2022). Integrating generative AI into financial reporting systems for automated insights and decision support. Available at SSRN 5232395.
[23] Koppolu, H. K. R. (2021). Leveraging 5G Services for Next-Generation Telecom and Media Innovation. International Journal of Scientific Research and Modern Technology, 89–106. https://doi.org/10.38124/ijsrmt.v1i12.472
[24] End-to-End Traceability and Defect Prediction in Automotive Production Using Blockchain and Machine Learning. (2022). International Journal of Engineering and Computer Science, 11(12), 25711-25732. https://doi.org/10.18535/ijecs.v11i12.4746
[25] Chaitran Chakilam. (2022). AI-Driven Insights In Disease Prediction And Prevention: The Role Of Cloud Computing In Scalable Healthcare Delivery. Migration Letters, 19(S8), 2105–2123. Retrieved from https://migrationletters.com/index.php/ml/article/view/11883
[26] Sriram, H. K., ADUSUPALLI, B., & Malempati, M. (2021). Revolutionizing Risk Assessment and Financial Ecosystems with Smart Automation, Secure Digital Solutions, and Advanced Analytical Frameworks.
[27] Avinash Pamisetty. (2021). A comparative study of cloud platforms for scalable infrastructure in food distribution supply chains. Journal of International Crisis and Risk Communication Research , 68–86. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/2980
[28] Gadi, A. L., Kannan, S., Nanan, B. P., Komaragiri, V. B., & Singireddy, S. (2021). Advanced Computational Technologies in Vehicle Production, Digital Connectivity, and Sustainable Transportation: Innovations in Intelligent Systems, Eco-Friendly Manufacturing, and Financial Optimization. Universal Journal of Finance and Economics, 1(1), 87-100.
[29] Dodda, A. (2022). The Role of Generative AI in Enhancing Customer Experience and Risk Management in Credit Card Services. International Journal of Scientific Research and Modern Technology, 138–154. https://doi.org/10.38124/ijsrmt.v1i12.491
[30] Gadi, A. L. (2022). Connected Financial Services in the Automotive Industry: AI-Powered Risk Assessment

and Fraud Prevention. Journal of International Crisis and Risk Communication Research, 11-28.

[31] Pamisetty, A. (2022). A Comparative Study of AWS, Azure, and GCP for Scalable Big Data Solutions in Wholesale Product Distribution. International Journal of Scientific Research and Modern Technology, 71–88. https://doi.org/10.38124/ijsrmt.v1i12.466

[32] Adusupalli, B. (2021). Multi-Agent Advisory Networks: Redefining Insurance Consulting with Collaborative Agentic AI Systems. Journal of International Crisis and Risk Communication Research, 45-67.

[33] Dwaraka Nath Kummari. (2022). Iot-Enabled Additive Manufacturing: Improving Prototyping Speed And Customization In The Automotive Sector . Migration Letters, 19(S8), 2084–2104. Retrieved from https://migrationletters.com/index.php/ml/article/view/11882

[34] Data-Driven Strategies for Optimizing Customer Journeys Across Telecom and Healthcare Industries. (2021). International Journal of Engineering and Computer Science, 10(12), 25552-25571. https://doi.org/10.18535/ijecs.v10i12.4662

[35] Adusupalli, B., Singireddy, S., Sriram, H. K., Kaulwar, P. K., & Malempati, M. (2021). Revolutionizing Risk Assessment and Financial Ecosystems with Smart Automation, Secure Digital Solutions, and Advanced Analytical Frameworks. Universal Journal of Finance and Economics, 1(1), 101-122.

[36] AI-Based Financial Advisory Systems: Revolutionizing Personalized Investment Strategies. (2021). International Journal of Engineering and Computer Science, 10(12). https://doi.org/10.18535/ijecs.v10i12.4655

[37] Karthik Chava. (2022). Harnessing Artificial Intelligence and Big Data for Transformative Healthcare Delivery. International Journal on Recent and Innovation Trends in Computing and Communication, 10(12), 502–520. Retrieved from https://ijritcc.org/index.php/ijritcc/article/view/11583

[38] Challa, K. (2022). The Future of Cashless Economies Through Big Data Analytics in Payment Systems. International Journal of Scientific Research and Modern Technology, 60–70. https://doi.org/10.38124/ijsrmt.v1i12.467

[39] Pamisetty, V., Pandiri, L., Annapareddy, V. N., & Sriram, H. K. (2022). Leveraging AI, Machine Learning, And Big Data For Enhancing Tax Compliance, Fraud Detection, And Predictive Analytics In Government Financial Management. Machine Learning, And Big Data For Enhancing Tax Compliance, Fraud Detection, And Predictive Analytics In Government Financial Management (June 15, 2022).

[40] Innovations in Spinal Muscular Atrophy: From Gene Therapy to Disease-Modifying Treatments. (2021). International Journal of Engineering and Computer Science, 10(12), 25531-25551. https://doi.org/10.18535/ijecs.v10i12.4659

[41] Kaulwar, P. K. (2022). Data-Engineered Intelligence: An AI-Driven Framework for Scalable and Compliant Tax Consulting Ecosystems. Kurdish Studies, 10 (2), 774–788.

[42] Operationalizing Intelligence: A Unified Approach to MLOps and Scalable AI Workflows in Hybrid Cloud Environments. (2022). International Journal of Engineering and Computer Science, 11(12), 25691-25710. https://doi.org/10.18535/ijecs.v11i12.4743

[43] Nandan, B. P., & Chitta, S. (2022). Advanced Optical Proximity Correction (OPC) Techniques in Computational Lithography: Addressing the Challenges of Pattern Fidelity and Edge Placement Error. Global Journal of Medical Case Reports, 2(1), 58-75.

[44] Raviteja Meda. (2021). Machine Learning-Based Color Recommendation Engines for Enhanced Customer Personalization. Journal of International Crisis and Risk Communication Research , 124–140. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/3018

[45] Rao Suura, S. (2021). Personalized Health Care Decisions Powered By Big Data And Generative Artificial Intelligence In Genomic Diagnostics. Journal of Survey in Fisheries Sciences. https://doi.org/10.53555/sfs.v7i3.3558

[46] Implementing Infrastructure-as-Code for Telecom Networks: Challenges and Best Practices for Scalable Service Orchestration. (2021). International Journal of Engineering and Computer Science, 10(12), 25631-25650. https://doi.org/10.18535/ijecs.v10i12.4671

[47] Vamsee Pamisetty, Lahari Pandiri, Sneha Singireddy, Venkata Narasareddy Annapareddy, Harish Kumar Sriram. (2022). Leveraging AI, Machine Learning, And Big Data For Enhancing Tax Compliance, Fraud Detection, And Predictive Analytics In Government Financial Management. Migration Letters, 19(S5), 1770–1784. Retrieved from https://migrationletters.com/index.php/ml/article/view/11808

[48] Someshwar Mashetty. (2020). Affordable Housing Through Smart Mortgage Financing: Technology, Analytics, And Innovation. International Journal on Recent and Innovation Trends in Computing and Communication, 8(12), 99–110. Retrieved from https://ijritcc.org/index.php/ijritcc/article/view/11581

[49] Srinivasa Rao Challa,. (2022). Cloud-Powered Financial Intelligence: Integrating AI and Big Data for Smarter Wealth Management Solutions. Mathematical Statistician and Engineering Applications, 71(4), 16842–16862. Retrieved from https://philstat.org/index.php/MSEA/article/view/2977

[50] Paleti, S. (2022). Fusion Bank: Integrating AI-Driven Financial Innovations with Risk-Aware Data Engineering in Modern Banking. Mathematical Statistician and Engineering Applications, 71(4), 16785-16800.

[51] Pamisetty, V. (2022). Transforming Fiscal Impact Analysis with AI, Big Data, and Cloud Computing: A Framework for Modern Public Sector Finance. Big Data, and Cloud Computing: A Framework for Modern Public Sector Finance (November 30, 2022).

[52] Kommaragiri, V. B., Gadi, A. L., Kannan, S., & Preethish Nanan, B. (2021). Advanced Computational Technologies in Vehicle Production, Digital Connectivity, and Sustainable Transportation: Innovations in Intelligent Systems, Eco-Friendly Manufacturing, and Financial Optimization.

[53] Annapareddy, V. N. (2022). Integrating AI, Machine Learning, and Cloud Computing to Drive Innovation in Renewable Energy Systems and Education Technology Solutions. Available at SSRN 5240116.

[54] Transforming Renewable Energy and Educational Technologies Through AI, Machine Learning, Big Data Analytics, and Cloud-Based IT Integrations. (2021). International Journal of Engineering and Computer Science, 10(12), 25572-25585. https://doi.org/10.18535/ijecs.v10i12.4665

[55] Venkata Bhardwaj Komaragiri. (2021). Machine Learning Models for Predictive Maintenance and Performance Optimization in Telecom Infrastructure. Journal of International Crisis and Risk Communication Research , 141–167. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/3019

[56] Paleti, S. (2021). Cognitive Core Banking: A Data-Engineered, AI-Infused Architecture for Proactive Risk Compliance Management. AI-Infused Architecture for Proactive Risk Compliance Management (December 21, 2021).

[57] Harish Kumar Sriram. (2022). AI-Driven Optimization of Intelligent Supply Chains and Payment Systems: Enhancing Security, Tax Compliance, and Audit Efficiency in Financial Operations. Mathematical Statistician and Engineering Applications, 71(4), 16729–16748. Retrieved from https://philstat.org/index.php/MSEA/article/view/2966

[58] Chava, K., Chakilam, C., Suura, S. R., & Recharla, M. (2021). Advancing Healthcare Innovation in 2021: Integrating AI, Digital Health Technologies, and Precision Medicine for Improved Patient Outcomes. Global Journal of Medical Case Reports, 1(1), 29-41.

[59] Data Engineering Architectures for Real-Time Quality Monitoring in Paint Production Lines. (2020). International Journal of Engineering and Computer Science, 9(12), 25289-25303. https://doi.org/10.18535/ijecs.v9i12.4587

[60] Pallav Kumar Kaulwar. (2021). From Code to Counsel: Deep Learning and Data Engineering Synergy for Intelligent Tax Strategy Generation. Journal of International Crisis and Risk Communication Research, 1–20. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/2967

[61] Pandiri, L., & Chitta, S. (2022). Leveraging AI and Big Data for Real-Time Risk Profiling and Claims Processing: A Case Study on Usage-Based Auto Insurance. Kurdish Studies. https://doi.org/10.53555/ks.v10i2.3760

[62] Kummari, D. N. (2022). AI-Driven Predictive Maintenance for Industrial Robots in Automotive Manufacturing: A Case Study. International Journal of Scientific Research and Modern Technology, 107–119. https://doi.org/10.38124/ijsrmt.v1i12.489

[63] Gadi, A. L. (2022). Cloud-Native Data Governance for Next-Generation Automotive Manufacturing: Securing, Managing, and Optimizing Big Data in AI-Driven Production Systems. Kurdish Studies. https://doi.org/10.53555/ks.v10i2.3758

[64] Dodda, A. (2022). Secure and Ethical Deployment of AI in Digital Payments: A Framework for the Future of Fintech. Kurdish Studies. https://doi.org/10.53555/ks.v10i2.3834

[65] Gadi, A. L. (2021). The Future of Automotive Mobility: Integrating Cloud-Based Connected Services for Sustainable and Autonomous Transportation. International Journal on Recent and Innovation Trends in Computing and Communication, 9(12), 179-187.

[66] Dodda, A. (2022). Strategic Financial Intelligence: Using Machine Learning to Inform Partnership Driven Growth in Global Payment Networks. International Journal of Scientific Research and Modern Technology, 1(12), 10-25.

[67] Just-in-Time Inventory Management Using Reinforcement Learning in Automotive Supply Chains. (2021). International Journal of Engineering and Computer Science, 10(12), 25586-25605. https://doi.org/10.18535/ijecs.v10i12.4666

[68] Srinivasa Rao Challa. (2021). From Data to Decisions: Leveraging Machine Learning and Cloud Computing in Modern Wealth Management. Journal of International Crisis and Risk Communication Research , 102–123. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/3017

[69] Kommaragiri, V. B. (2021). Enhancing Telecom Security Through Big Data Analytics and Cloud-Based Threat Intelligence. Available at SSRN 5240140.