

Chapter 9: Cloud finance architecture: designing scalable and secure artificial intelligence infrastructure for financial applications

9.1. Introduction

There is growing interest and activity in incorporating end-to-end sequencing of data processes, connected machine learning, user-engagement, and audio-visual channels in the development of AI products and services. Attaining successful capital and donor allocation and attracting a target audience can present substantive challenges in network governance, user participation, incentives, financial sustainability, and the authority distribution between intelligence and human actors. Studies in these domains can provide guidance on the assurance of an aligned algorithm, policy runs, pro-user behaviors, and the effects on the specific responsibility of owners, engineers, and data; and the substantive protection of data against unauthorized access and malicious models.

AI finance tools are expected to overcome the limitations of conventional tools regarding the operating data, time, and volume of financial feed, and to relieve the unavailability of the currency type and quality of data in order to discard loss of opportunities. Computer vision and text mining methods can effectively accumulate and process heterogeneous, unstructured, semi-structured, and unknown forms of data. Armed with first-hand time-captured financial values, these tools can identify tens of widespread behavioral turning points across market participants and news agencies which account for the unprecedented information uproar and aftermath sequences (Asatryan, 2017; Chang et al., 2017; Munoko et al., 2020).

Quality measures, service metrics, habit models, and behavioral patterns can be examined for personalized recommendations. Temporary adaptation is needed to reflect the change of preference sentiment temporally. Latent attributes may embody factors in

models of spatio-temporal-entity-resource continuous recommendation, fragrance commutation, smart sampling, and price exploration. Treatment identification based on large sequential data and models in directible behavior targeting and mechanism design account for discounts, loyalty programs, advertisements, and issue events (ZestFinance, 2017; Rane et al., 2024).

All matters in principle to do with rational actors, behavioral prediction, and behavioral interventions can be comprehensively represented as distributed behavioral preference models to embed human common-sensory observations. Knowledge crowdsourcing and intelligent agents observe the online behaviors of a certain group on digital platforms to expand knowledge, mitigation shots, and knowledge discovery. Comprehensive modeling captures the heterogeneous scale cluster of crowd agents and builds agent states, representation of knowledge discovery, and network dynamics on different scales.

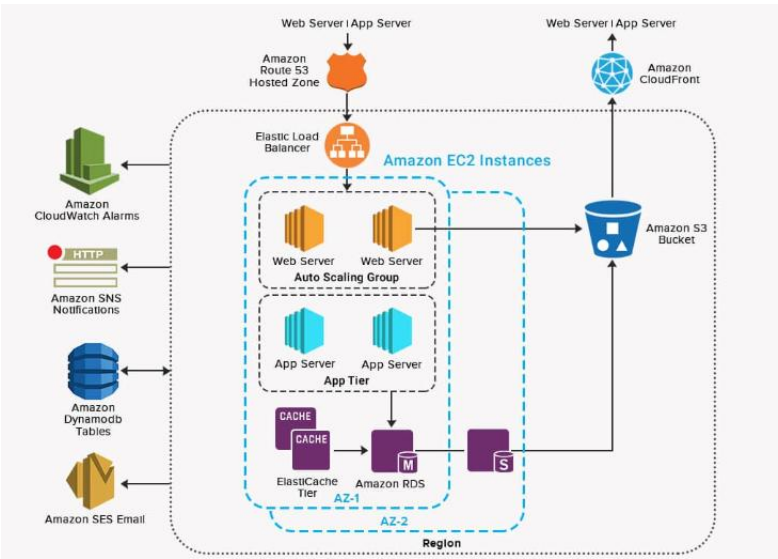


Fig 9.1: Cloud Infrastructure Services

9.1.1. Background and Significance

In recent years, Artificial Intelligence has become the trending topic of both academic research and the industry, with rapid development in various fields to improve quality of living. AI predictive models are significantly employed in finance for risk and fraud monitoring, investment asset allocation, and client profiling. With further advancement in AI capabilities, machine learning and AI are expected to integrate into financial market making systems and wealth management. The benefit of AI predictive models lies in improving model risk control and compliance by explaining underlying mechanisms and procedural logics of AI models in understandable terms. In response to

the recent growing usage of AI, the Chinese central bank acted before other legislators to raise strict new rules governing the system invariably.

Despite the strong push for AI regulation, there are still many challenges in enforcing the rules. For instance, one critical problem is how to ensure compliance. AI governance blocks that can be integrated into the AI production environment will be raised to continuously monitor the deployment of AI models, report to regulatory authorities or related parties, and interact with the AI model to prevent undesirable actions based on updated regulatory control requirements. Other questions include how to integrate governance blocks into the AI production environment without jeopardizing the performance of the AI model and whether there is a systematic solution framework for AI governance. These challenges are not exclusive to the financial field, and researchers seek more general AI governance solutions covering all domains and industries.

The concept of AI Self-Regulation has been proposed. It encompasses three perspectives: monitoring, regulatory control, and mitigation. In the monitoring process, timely and precise assessment is made regardless of the type of AI model. Monitoring output is reported to the regulatory control which is designed according to the regulatory guidelines. The events raised by the monitoring system are categorized as three types, including non-compliance, model risk and illegal actions. Anti-circumvention actions can be executed if necessary. However, existing AI Self-Regulation solutions only cover the monitoring process, while modular regulatory architectures for regulatory control and mitigation are unexplored. Mitigation capabilities are desired in the production environment. Self-regulation modules will be the focus and integrated into the system architecture. A generic solution framework being universal for different application types needs to be constructed for AI governance. A high-level architecture of the proposed system framework will be presented. Seven major functions provided by the proposed architecture can help AI govern well in the financial field.

9.2. Overview of Cloud Finance Architecture

Cloud finance is rapidly growing due to the increasing demand for financial cloud applications, coupled with the emergence of new financial service modes and businesses such as licensing transactions. The rapid expansion of the financial industry brings with it increased complexity and challenges in building a secure cloud finance architecture. Nowadays, financial institutions are becoming more open and agree to use public cloud storage, which was not the case in the past. For such financial institutions, using cloud services can save lots of maintenance costs and provide more advanced data processing technologies. However, the widespread availability of cloud services for many users also raises many concerns on security and an architecture that provides provable security and safety guarantee is desperately needed.

This paper proposes a cloud finance architecture which incorporates a multi-server architecture for cloud finance applications. In this architecture, the complex receipt schema of the more open friendly public cloud model of cloud finance needs to be secured using advanced cryptographic techniques. In order to achieve data privacy of sensitive financial market data, a hybrid encryption and grid encryption scheme is proposed to securely access this data and preserve its confidentiality. Third-party services provide accounting lettering verification services which may undermine the security of the receipt workflow. This paper also constructs a secure and efficient account compliance verification scheme in order to guarantee receipt safety for cloud finance applications.

Cloud computing enables financial institutions to use information services without expensive hardware facilities and can significantly increase the service accessibility of numerous finance applications in the Internet finance industry. A reference architecture for cloud computing and Internet finance is proposed. Then according to the reference architecture, several critical technologies used to support massive Internet finance applications are discussed. The application of cloud computing in finance can significantly reduce IT costs and improve business flexibility. Reducing resource management and service handling costs and enhancing high availability and fault tolerance of virtual machines (VMs) can help banks adopt cloud computing and set up their private clouds. In securities funds, VM management and batch VM setup is studied, and personalized scalable data-intensive web applications are introduced. A cloud brokerage framework is designed for brokering services across multiple cloud platforms in the broker domain.

9.2.1. Definition and Importance

The cloud computing model is increasingly being adopted by financial institutions due to its flexibility and improved speed to market. Cloud services can provide better financial data processing and management. However, the cloud environment introduces new threats for financial institutions, requiring responses to ensure data security. This paper combines intelligent prediction and assessment to provide a basis for financial institutions to address security risks and assess financial information risks in the cloud computing model. Cloud computing can provide companies with highly scalable and flexible data processing and management services, without the need to build a large-scale computing infrastructure. Cloud services can be divided into core services like IaaS, PaaS, and SaaS, or other specialized services like Database as a Service, Big Data Service, and AI as a Service. Drawing from these services, financial institutions can reduce costs by shifting to cloud-native architecture and utilizing cloud services to improve systems, services, and data processing and management. The cloud-native

architecture includes an agile deployment and delivery cascade, multi-tenancy capabilities, container cluster management infrastructure, and cloud-native data architecture. Cloud-native services can optimize resource deployment and improve delivery efficiency. While data security is enhanced, new threats emerge. Multi-tenancy capabilities can lead to failure to meet data segregation conditions, enabling unauthorized data access. Large-scale usage of a public cloud increases the likelihood of denial-of-service attacks against large service providers, causing server overload. Cloud-native technological frameworks can introduce numerous pods. Operators unaware of the underlying states may be vulnerable to malicious attacks. These risks may be undetected due to immense data and resource disparity, necessitating intelligent understanding capabilities to reduce axes of complexity and provide abstraction.

9.2.2. Key Components

Modern finance is data intensive, and the unprecedented rise in monetary cross border transactions has meant an unprecedented increase in the volume of such transactions. To put this in perspective, the average daily foreign exchange market trading turnover reached \$7.5 trillion in April 2022. Over-the-counter (OTC) interest rate derivatives, which are primarily used to hedge interest rate and foreign exchange risk, had a total daily notional amount of \$65.2 trillion, and the total notional amount outstanding was \$625 trillion. Critical market inefficiencies, such as risks of fraud, malpractice, and information asymmetry, have arisen from the shortcomings of the existing fiat-based interbank and overseas transactions due to the complications in running, maintaining, and coordinating database management systems (DBMSs), and automating the processing and distribution of unstructured financial relevant Information (FRI).

Cryptocurrencies (CCs) were created to reform the existing system, and a number of decentralised platforms for peer-2-peer electronic transactions without the need for trusted third-party intermediaries have emerged. In addition to cryptocurrencies, blockchain and decentralised financial (DeFi) technology have been adopted in the finance industry to fully automate the processing of multi-hop cross-border transactions of any monetary asset. However, despite the extraordinary motivation behind the creation of CCs and the rapid growth of the cryptocurrency market and DeFi technology, there are still significant challenges and risks in operating a decentralised finance (DeFi) platform. In addition, the current modelling of DeFi-related ideas is still very limited and at a comprehension level that is much lower than the existing research on centralised financial (CeFi) technology. A comprehensive study of the current models of DeFi platforms, the combat of the existing challenges, and a closer examination of DeFi-related ideas could bring valuable insight to promote the development of a decentralised finance industry and related research topics.

9.3. Scalability in Cloud Finance Systems

Scalability is defined as the potential for a computer application and its associated resources to be scaled vertically and horizontally. Vertical scaling involves adding additional processing power in the likes of CPUs or GPUs working on the task, while horizontal scaling involves creating a distributed cluster with multiple CPU and/or GPU nodes on the same/different clouds executing copies of the program.

The architecture of a Cloud Finance System encapsulates software and hardware devices such as networks, cloud services, storage media and machines. Software characterized as microservices split the code of a self-contained application into several packages, which can run independently. Microservices allow developers to focus on a specific programming language or technology stack hitting optimal performance. When innovation launches technologies such as PaaS or a new programming language, migration is inevitable. This architecture gives Cloud Finance platforms the freedom to optimize performance in each microservice independently. If the programmer wishes to switch languages, then only a subset of the code requires rewriting, and the developer can switch core components of the architecture independently. The cost of calling code in other languages is diminished by the high-rate messaging system.

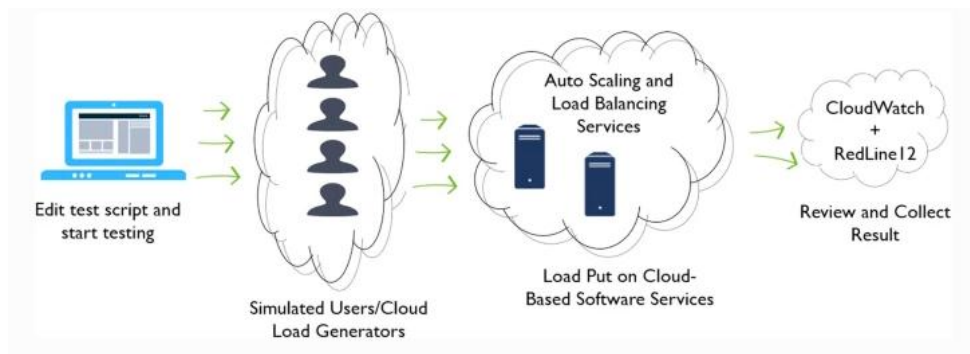


Fig 9.2: Scalability analysis comparisons of cloud-based software services

9.3.1. Understanding Scalability

In IT, Scalability refers to system performance improvement or adaptability enhancement when it is increasing load, working data, or employing additional resources. A grant performance can be earned by improving an existing resources specification. It is a vertical scalability. If workload is distributed among added resources, they are same or similar resources as existing ones or also said homogeneity resources. It is a horizontal scalability. In a cloud ecosystem, resources are external services. All organizations involving cloud including application and store owner, service provider and cloud provider should contact cloud service or network function

through cloud APIs. As for cloud application providers, functionality and availability of cloud services could be managed. Extensibility of both system or high-level components and performance resources could be collaborative.

In social networks, knowledge web, classification or clustering, and merits earned by study progress found by an input of new data are commonly required. They are heterogeneous applications and also referred to as research, academia usage or application. Most initial systems deployed locally, but for production-grade scalability, globally distributed clusters with cross-format cloud storage are to be employed. Inputs are stored in blobs housed in one cloud storage and aggregated extracts feed for normalizing, quantizing, and generating features to different dimensions stored in another cloud. The batch behaviors of parallel processing of ordered time-weighted distributed inputs enhance structure availability and research merit. Cooperative resource cloud infrastructure micro-granted telescopes coordinate trade-off harvests of low position clones. Retention filtering range of the full transcript would inherently represent the gradual synopsis at lower dimensions. First seconds could be on-demanded as input speech signal acknowledged quick recognition against long transcribed documents by labeling indicator.

Cross domain processing of data application development is a growing trend. They generate high-valued composites, however comprehensive understanding of workflow must be required for optimization. Middleware platform offers low-term abstraction components and perception of interfaces, systems, and resources can be monitored easily. On the other hand, workflow monetizing service hides the complexity details of underlying resources and effort saving for design and implementation. Data alignment with platform description of inter-domain parameters and data is a two-sided challenge. Lack of a common format of behaviours and metrics monitoring would also impact analysis and optimization without enough knowledge.

9.3.2. Techniques for Achieving Scalability

Architecture, framework, and systems design are necessary for supporting a scalable machine learning pipeline and deployment on clouds. This approach can be beneficial for financial domain applications, but it requires integrating a distributed computing and programming model to naturally address asynchronous, heterogeneous, and dynamic workloads. Hence, its technologies could address the architectural design issues while supporting the architecture to be more easily extendable.

For dealing with time-varying workloads, performing online workloads detection has become increasingly important. A second area covers techniques to prevent data drift and deal with poisoned inputs that would lead to skewed workloads. Re-training

approaches can be handled more efficiently through adaptive model selection and lower re-training overhead using techniques that discard parts of the loaded model. The recent emerging transfer learning is also a potential area of interest that has not gained much attention yet for automatically transferring knowledge across domains or tasks.

It is also important to study the hyper-parameter optimization of financial models using data-driven optimization. This concern has two aspects: a wider search space with expensive evaluations and the attention of competing objectives. Optimization with rare expensive evaluations is necessary for ensuring more robust solutions. Ensure availability for training by effectively providing a noisy and partial view of the objective space by explicitly considering the batch acquisition function.

9.4. Security Considerations in Cloud Finance

In the past, the primary information security objective was confidentiality: protecting data against unauthorized access. Accordingly, cryptography has been the cornerstone of information security as a useful means to securely store and transfer data. Cryptography is considered “a trust technology”. Only if cryptography is designed correctly, is implemented correctly, is customized correctly, and is used correctly, trust is established in a cryptographic system . Otherwise, the computation on encrypted data may reveal information about the plaintext messages, or an off-the-shelf protocol may be improperly applied, such as where only one party fails to know a common random number needed for the protocol to be secure. This can render all efforts to secure sensitive information futile. In summary, cloud and fog computing are extremely promising computing paradigms and the backbone of future AI applications, which also pose a multitude of new security challenges that need to be addressed comprehensively, both by applying and extending existing security mechanisms and developing new techniques. AI can help in improving the resilience of information security mechanisms. On the other hand, AI applications also raise new security challenges on their own.

The significance of securing an operating cloud/fog infrastructure is amplified manifold in view of the massive and sensitive data that is typically processed in AI applications. Attacks against the cloud services or fog computing networks on which current AI applications are built will inevitably result in difficulties, data breaches, failures, or malfunctioning of the AI applications . This makes them an attractive target for cybercriminals: they can try to prevent access to AI services on the Internet in order to extort a ransom from the service provider; they can also try to steal training data or complete ML models. The interplay between AI and information security promises huge potential for future applications and research. For instance, it is already possible to use language models to generate phishing emails optimized for a specific target. Currently, it is not enough to train an AI model on a sufficiently large and complex training dataset:

if the user has too little input data to supply, it is even necessary to guess this data based on what is known about the target.

9.4.1. Common Security Threats

- In the cloud-based finance architecture landscape, AI platforms are divided into five general shortage design groups for a scalable and secure AI and cloud brewing service.

- Only in pairwise vanity configurations, AI systems are compared to provide a central concentration and apt pay-off design, limiting savings when services are requested. Boomerang designs limit the benefits of cloud finance architecture as many nodes must be placed as minimal, resulting in revenue losses compared to monopolistic implementations. Simple backup designs provide security against isolation failures, but the overall finance architectures still face common threats as many nodes need to average. Only decentralized DHT-based configurations show indifferent mechanics for both low-resource and high-quality confidence on both security fronts.

- A general classification with 11 types of security threats in the FinTech domain and 9 corresponding defense strategies represents and summarizes recent and relevant literature.

- Attacks against the cloud services or fog computing networks on which current AI applications are built will result in difficulties, data breaches, failures, or malfunctioning of the AI applications. AI applications are an attractive target for cybercriminals: they can prevent access to AI services to extort ransom, or steal training data or complete ML models to extort payment or sell to competitors. The interplay between AI and information security promises huge potential for future applications and research.

Threats with potential remedies are analyzed against plain FinTech players providing service expertise, infrastructure providers designing novel AI services, or financial products. Product and service providers covering DApps, and games are good candidates for “cheap” infestation by state-funded cybercriminals, as are node operators making good targets for infecting all-honest networks with limited overheads. In particular for the latter category, an analysis of defense techniques was presented recently. Drafts based on cloud configurations suggest equipped nodes for solutions with high and competitive payoffs.

9.4.2. Best Practices for Security

The use of cloud-backed applications has increased with the pace of cloud computing and the fast-growing AI technology. Cloud computing is a technology model where

multiple clients access the same remote server to benefit from a pooled set of resources, and numerous business models evolved around these technologies over the last decade. Numerous e-commerce companies, banks, cloud storage providers, video and music streaming services, and many other services and applications were quickly built on the cloud, benefiting from the pooled and scalable resources, initial development cost savings, and high availability of cloud platforms. AI uses significantly more computing power than conventional software applications, and analyzing huge datasets has become a bottleneck for many applications in academic or industrial sectors, but also with the increased importance of privacy and security of the data while training machine learning models. Consequently, the first AI cloud applications started to appear. Today, numerous AI services such as optical character recognition, facial recognition, text recognition and understanding, image generators, etc. can be accessed by paying a monthly fee or per process. AI security and AI for security promise to be hot topics for years to come, with lots of research opportunities.

This growing need for data security, protection against data breaches, parallel processing while keeping data private, or continuously training data privacy laws frequently catches attention with a newly developed application, but probably only a few people think about the behind-the-scenes software and hardware. While there are lots of computing resources and field-programmable gate array accelerator boards, they still need to scale linearly or need to interleave/make full use of the on-chip memory. The attacks on services or applications built on this cloud or fog are desired since they can be used to obtain training data or trained ML models without paying. While training these, due to the more public dataset availability, the architecture is proprietary, and the ML and multitasking capabilities need to get better. Still, so far only basic attacks on complete IA and custom-made or highly tuned artifacts or dictionaries for training or detection are known. It is possible to achieve much better results, but this very much depends on the cloud provider.

9.5. AI Infrastructure for Financial Applications

AI innovation has gained momentum in recent years, becoming a core competitive infrastructure for finance, insurance, and real estate. However, owing to challenges such as systemic catastrophe, data leakage, bias, and abuse of information, AI governance in finance is in a nascent stage. Traditional risk, compliance, and audit functions fall short of closing the governance gap. Therefore, a new paradigm of self-regulation is needed, with the active risk assessment and control of AI by AI itself. The proposed architecture addresses scalability and data privacy challenges. Modular governance building blocks are closely integrated with the core AI and financial systems. Each building block supports both pre-validation and run-time monitoring, as well as post-validation

analysis. Key metrics for continuous evaluation are summarized around safety, compliance, and fairness.

AI systems are viewed as dual-system agents consisting of a core AI system and environment modeling. The core AI system is further decomposed into an agent, emulator, attack generator, and wrapper. Governance building blocks include direct risk and compliance, outcome analysis, and agent behavior analysis. AI agents responding to the environment are of interest for self-regulation, including risk, compliance, market, and adversarial environments. The integrated and holistic view of AI system and governance architecture enables the development of comprehensive solutions addressing common challenges faced by different AI applications in finance. By changing the governance philosophy from “mounting building blocks externally” to “building blocks as integral parts of the system”, the governance capabilities of the core system are enhanced with negligible operating overhead.

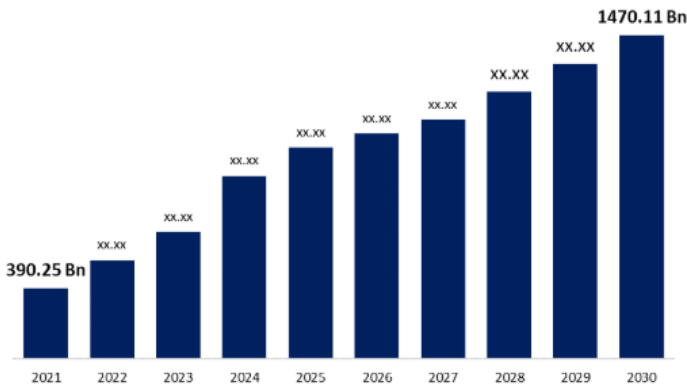


Fig : Cloud Computing Market

9.5.1. Role of AI in Finance

Data-driven big data AI digital finance has computing architecture requirements highly different from traditional computing architecture. Following the rapid development of data-driven big data AI, several computing architectures for big data AI have been proposed and implemented. With the fast growth of Internet finance wherein huge amounts of data collected daily, there is a pressing requirement to design novel efficient computing architectures tailored for data-driven big data AI finance. In addition to the architecture design challenges, there exist data-driven model deployment and effectiveness challenges such as privacy analysis. In finance, the AI model playing the role of insider trades and financial irregularities has become a hot topic. Still, limited analysis of AI climate and FinTech is available. Addressing these challenges and

requirements sheds new light and insight on a fast-growing branch of AI in a new research direction of finance. AI applications in Finance are contributing fundamentally to improving the performance of numerous fields. Finance and Economics in industry and academia are increasingly reliant on data-driven AI methods to enhance performance with a focus on function approximation and prediction. The challenges for AI applications in data-driven finance are conceptual dimensions which can be viewed through connection and linkage. In traditional models, participants act based on strategic cognition and bounded rationality. Bridging the gap between AI value assignments and such values or distributions is challenging. First and foremost, data-driven understanding is through statistics and a limited number of historical observations on certain aggregation levels. Static data-driven valuation normally lacks robustness and transparency regarding theoretical and economic foundations. Edge-based risks link over-the-counter transactions across and within market segments but are largely neglected in agent-based models. Pricing through graph matching can efficiently diversify these edges into nodes and link prediction. The money use perspectives of interest parking and interbank trading have implications for endogenously reducing contagion risks. AI value assigning processes need to embrace the conceptual dimensions of markets and economic agents.

9.5.2. Building AI Models

Artificial Intelligence (AI) is progressively introduced in financial applications. AI models perform tasks such as filtering transactions for fraud prevention, calling customers proactively, and recommending financial products to clients. These models are experimented with off-premise in the Hyperscale, and some of them are public cloud-based. Prior to going live, the models need to be assessed for scalability, robustness, regulatory alignment, and compliance with local laws. Model governance refers to this supervision and oversight of models. Presently, model governance is conducted mainly manually. Current governance practices include internal modelling frameworks, independent model risk groups, and overarching model risk policy documents. However, not all models receive the same scrutiny, and governance compliance is unstructured, white-box, and heavily reliant on templates. Hence, it suffers from being time-consuming, error-prone, and prone to changing staff, leading to the need for an efficient and intelligent model governance protocol.

Advancements in automation and AI apply natural language processing (NLP) for the classification of governance accents from policy documentation, and create dashboards for visualizing model risk situations. The governance practices are studied to identify possible areas of improvement through AI. Diverse NLP and ML models are considered for cash incentive classification. Some key aspects of modelling, prior to model

selection, are identified. The interpretation of the governance criteria is also studied, with some criteria clarified as being vague, which translates to data labels, making them subjective to human judgment. Consequently, modelling is performed for two interpretation classes: subjective and objective. Output forms of the classification are categorized into similarity scores and label distributions. The NLP models are tested on a semi-conducted evaluation dataset that is representative of daily screening classes.

Prototype outputs in vision documents and outputs regarding classification and competitiveness on a global scale are elaborated. This recommendation emerges due to the exceptional performance of these two outputs compared to the field leads. As AI systems have become an indispensable part of the financial services industry (FSI), the importance of effective AI model governance to ensure robustness and compliance is increasingly acknowledged. AI system complexities have increased exponentially. Given the unprecedented growth in AI complexities in the past years due to the modelling evolution, the feasibility of such governance practices and institutions in light of the next-generation AI models is questionable. Therefore, a vast number of challenges in understanding, analyzing, monitoring, and regulating such AI systems have yet to be addressed. Specifically, practitioners and researchers critically assess the efforts and capabilities of existing model governance practices and institutions to effectively govern the next-generation AI systems.

9.6. Conclusion

The advent of big data has significantly transformed various industries, and financial institutions are no exception. The incorporation of machine learning (ML) technology into algorithmic trading is one of the most notable shifts in the finance industry. Developing multiple latent features from the raw price series and trading volume has been employed to create informative predictors. However, their incorporation into algorithms is indirect and subject to a myriad of risks. While many ML applications in finance have seen advancements in methodologies or performance, deploying them remains an arduous task. The role of the finance personnel has been eroded and needs to be replaced by financial engineers who implement and maintain broader financial infrastructures.

While the importance of expertise in the finance domain is growing, the number of specialists is not increasing commensurately. Additionally, the boundaries of financial models are eroded and broadened as ML can seize irregularities of any situations, creating prohibitively enormous models to vet (a phenomenon called boundary erosion). Taking a financial time series as input, a smooth function is generated, which must satisfy the one-sided limit, an extra condition not applicable to ML itself (a phenomenon called data dependencies). Furthermore, recent advances in NLP technologies suggest

that textual data can be simulated as a time series composed of statistics, which could be used for backdooring. Despite the reasons from risk perspectives, most academic endeavors are still conducted on methodologies, contexts being of secondary priority. With the inception of the Data Extraction and Optimization for Heterogeneous Data (DEOHD) System, a robust solution to these deployment issues has been presented. The DEOHD System abstracts time-series data processing, eliminating the need to switch contexts. The containerization, together with the optimization of tools used, allows tremendous flexibility for data handling. To keep investment strategies consistent for every operator, there will be the Financial Information Development and Extraction for Overall Data Kit (FIDEO-DK)—a developer kit. The goal of such a kit is to create a highly standardized platform for data elaboration, enabling homogeneous and efficient assessments of any model, hosted across all backtesting systems. Zero-cost standardization is essential to facilitate fair assessment (UID) across a multitude of heterogeneous platforms, and a backtester-independent developer kit streamlines these needs while foiling any automated fraud.

9.6.1. Future Trends

The future trends in cloud computing focus on the optimization of costs, risks, and performance. The major services provided by the cloud, mainly Infrastructure as a Service (IaaS) and Platform as a Service (PaaS), are chargeable services. Thus, users or companies using these services would want to optimize their costs by controlling their expenditure budgets for cloud resource provisioning, but this is non-trivial as users depend on the cloud service providers for computation and the cloud is subject to increased uncertainty and dynamicity. Also, this raises the necessity to develop a general model that abstracts the tradeoffs among different types of risks involved in cloud computing, and an intelligent framework and algorithms for risk-aware cloud service scheduling, performance enhancement, service execution migration, and catastrophe recovery. Security not only concerns the traditional threats of the IT infrastructure itself, such as against some server-hacking attacks, but also raises new challenges like data access control in multi-tenancy with the underlying IaaS shared among different customers. Anti-DDoS or counter-bot methods developed for traditional web services often adapted for the cloud services would be domain-specific, e.g. in their information collection strategies.

Efforts to prevent vulnerabilities in applying the very new technology, such as from quantum computers attacking the existing encryption/decryption mechanism for data privacy, are also crucial to future cloud computing. New storage technologies such as DNA or memory-cost-efficient ones would also have a huge impact. There can hardly be a 100% trustworthy computing infrastructure. Consequently, the legal liability and

obligation in case of service failure, data disclosure/privacy breaches, or non-compete agreement violation in IaaS are complementary parts of the CSPs' service level agreement, hence of great concern. Therefore, to mathematically model, evaluate, and optimize the cloud service legal risk is very much called for, where the fundamental research effort is mostly lacking. In addition, various systematic methodologies for producing trustworthy cloud infrastructures need to be studied.

References

- Asatryan, D. (2017). Machine Learning Is the Future of Underwriting, But Startups Won't be Driving It. Asatryan.
- Chang, H., Kao, Y.-C., Mashruwala, R., & Sorensen, S. M. (2017). Technical Inefficiency, Allocative Inefficiency, and Audit Pricing. *Journal of Accounting, Auditing & Finance*.
- Munoko, I., Brown-Liburd, H. L., & Vasarhelyi, M. (2020). The Ethical Implications of Using Artificial Intelligence in Auditing. *Journal of Business Ethics*.
- Zemankova, A. (2019). International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO). ICCAIRO.
- Shaffer, K. J., Gaumer, C. J., & Bradley, K. P. (2020). Artificial Intelligence Products Reshape Accounting: Time to Re-Train. *Development and Learning in Organizations*.