Log Anomaly Detection and Incident Probability Forecasting: AI-Based Predictive Maintenance Frameworks for Financial Information Technology Infrastructure

Dr. Yasemin Şahin, Associate Professor of Electrical and Electronics Engineering, Middle East Technical University (METU), Turkey

1. Introduction

Financial institutions are highly dependent on their information systems and networks. It is critical for banks, brokers, and insurance agencies to quickly identify and repair failures in these networks in order to prevent their paralyzing impact. Modern IT solutions based on artificial intelligence, as well as event correlators and monitoring systems designed to identify fault conditions in networks, can help in this selective search for an observed problem. The artificial neural networks that have become central to the field of AI are used for predictive maintenance and failure detection. A technology allowing the use of AI to assess the technical condition of equipment based on the similarity of equipment interactions also attracts interest. Such systems seek to examine the technical state of the factory environment, its elements, and equipment in order to recognize conditions contributing to a common error.

1.1. Background and Context

1.1. Background and Context: Modern High Availability (HA) Computing Infrastructures present many benefits for financial organizations. High Availability Clusters increase the availability of critical server applications by allowing a given server to be pulled out of the cluster, brought down for maintenance, and brought back up without requiring a shutdown of the server and the application. Banks have plenty of applications with 24/7/365 demands, such as sites that customers can visit to check their accounts, ATMs, back-office applications, IT, among others. To have a standard high availability configuration, banks need at least two sites: a production site and a disaster recovery site. To implement a high availability cluster, banks must provide

support teams for each site (or many support teams in the case of local extra clusters). These support teams must have support from IT, cooling, and energy.

AI predictive maintenance may reduce datacenter energy consumption problems. Normally, banks may not have enough operational data to run time series models to do datacenter power consumption forecasts. If banks have multiple branches, the possibilities of creating a good time series model increase. Datacenter clusters have similar characteristics. Consequently, after creating the time series model at any bank, model parameters and resources spent developing the model at that bank may be shared with new banks. There are two alternative methods to develop a model. A time series model based on the matrix factorization technique will be presented using only the resources at a given bank. The alternative is to use a model available in Spark ML. Only enterprise banks have many resources to work on big data problems. A bank has a set of similar configurations to other banks. The bank has deployed many models at once. After integrating and locally testing with production data at bank i, the model (parameters and all necessary resources) can be deployed in a customizable and on-demand cluster with x number of clusters.

1.2. Significance of Predictive Maintenance in Financial IT Infrastructure

1. Introduction 1.1. Predictive Maintenance Regarding AI, the future growth engine, achieving higher accuracy in machine learning based on big data and deep learning is essential. In this context, one of the most recognized and representative fields of artificial intelligence is predictive maintenance, a core application area in which big data and deep learning can bring about immediate business results. In predictive maintenance, turbulence forecasting in aviation, service life extension for energy, and risk management in finance are cited as representative applications of AI. However, AI-based predictive maintenance accuracy technical papers have been published mainly in aerospace engineering and manufacturing areas, with minimal implementation and empirical research in the financial industry. In addition, legacy systems and other barriers have created significant challenges for the financial industry in applying predictive maintenance in actual business sites. However, timely machine failure forecasts can drive new customer value and more sophisticated financial service innovation through the maximum utilization of introduced data state. 1.2. Significance of Predictive Maintenance in Financial IT Infrastructure This study focuses on predictive

maintenance for the financial industry, specifically for the IT infrastructure that supports the financial business, to address challenges that existing research has not yet resolved. In the financial service industry, prediction and prevention of human or non-human service disruption are typically addressed using Business Continuity Management, one of the common protocols ensuring Business Continuity Planning. Traditional financial systems implement multiple layers of redundancy using network, hardware, and software technologies provided by multiple vendors per root cause analysis. However, financial institutions, like other industries, are rapidly evolving in response to the acceleration of the fourth industrial revolution and the fintech industry convergence. As a result, the integration of technology into financial business services is not as immediate. The value of predictive maintenance is clear, but the obstacles are diverse: several machine learning and data handling challenges exist, along with other difficulties such as selecting the appropriate technology and organizational leadership resistance to technological change. However, as cloud and open-source-based AI and machine learning evolve, cultural and organizational change is another challenge.

2. Fundamentals of Predictive Maintenance

What is predictive maintenance? In a narrow sense, predictive maintenance is a tool for machinery and equipment servicing. In a broader sense, this term implies a system that: • demonstrates when failures can occur • allows for determining the best time for repair • makes it possible to optimize costs • helps in asset management. This approach is different from two other types of maintenance. One of them represents a technical function that is meant for troubleshooting related to the sudden destruction of systems and devices, often called breakdown maintenance. The other features predetermined intervals and changing worn-out elements, which is called the preventive kind of maintenance. In contrast to them, predictive maintenance is proactive.

Predictive maintenance is a data-driven activity that uses technology and data to help determine when equipment issues will occur and to inform maintenance scheduling and extend the life of assets. The practice requires the availability and accuracy of data, most often collected by various sensors on equipment that monitor and analyze it to derive predictive maintenance insights. It acts as a mix of insights that combine failure prediction, maintenance scheduling, and forecasting in terms of asset management. The primary tools that help build and deploy a predictive maintenance system are sensors

and devices, platforms that store and analyze information, and technologies that provide a framework for it. Still, these technologies are not secrets in themselves. They are more a matter of providing an entry to the world of data. We should clearly realize that the better the data, the more precise the algorithms, and, consequently, we will be able to predict the future with greater accuracy.

2.1. Definition and Concepts

Predictive maintenance is a technique primarily used in holistic, or proactive, asset management. It assumes that the correct proactive maintenance action can be scheduled to prevent the occurrence of failure. Predictive maintenance can assume data and information quality and has to react if something is detected. Predictive maintenance can be broken down into condition-based monitoring, where the setup of real-time multiples is compared to decision-making levels. Connecting predictive maintenance to horizon scans, where all factors are taken into account to forecast failures, a comprehensive solution is set in place as part of an organization's mission to thoroughly maintain their IT infrastructure.

In general terms, predictive maintenance can be seen as the surveillance of factors that indicate the health of an asset. To have comprehensive knowledge about asset health, a large quantity of data needs to be collected in terms of condition data, status information, sensor readings, voltages, currents, wear, efficiency, CO2, etc., to couple and calculate a resulting asset health. A relationship established between health and any of these input parameters gives rise to predictive maintenance. Hereby, big data plays an increasingly significant role as it is almost the center of machine learning, automated control, and predictive analytics. In order to see the advantages of predictive maintenance, it is of key importance to understand these factors' significance and benefits in the finance industry, where improvement of even 1% of the operation refers to several million euros. Major advantages are the documentation of trend-setting information in the field.

2.2. Traditional vs. Predictive Maintenance Approaches

Traditional methods of maintaining systems and machinery may be analogously divided into three classes. One group consists of reactive maintenance strategies that are applied after a failure occurs. Preventive maintenance schedules maintenance operations based on the predetermined time of the machinery's component life or hours of

operation. Time-based maintenance and distance-based maintenance are examples of preventive maintenance strategies. This means that some components will, therefore, be unnecessarily replaced, and critical components may be detected only after the problem has occurred. The third and final group is the predictive maintenance strategy. Predictive maintenance gives no advanced notice of equipment failure.

None of these maintenance strategies individually directly address the fact that they may negatively affect organizational resilience when it comes to financial IT infrastructure. Many studies confirm that technological advances enable predicting failures and maintenance needs more precisely and directly, in a more automated way while coping with an increasing volume of data. Predictive maintenance goes one step further, offering greater benefits in order to increase the accuracy of predictions while reducing direct and indirect costs. This can be realized through the use of technology-rich tools and advanced algorithms that consider various parameters in more detail. Implementations of predictive maintenance techniques result in a variety of beneficial effects, such as reducing maintenance expenditures, thus indirectly increasing system operational efficiency or cutting energy consumption. Current competencies allow for the well-founded choice and modification of maintenance technologies for the needs of specific sectors or infrastructure classes and their influence on specific areas of operation.

3. Machine Learning in Predictive Maintenance

Machine learning has evolved into a powerful tool within modern predictive maintenance – the processes pivoting around sophisticated computer models for predicting asset failures. Many machine learning techniques are supported, which are divided into both semi-black algorithm groups and cryptic model-based ones. Examples of semi-black algorithms are those forming decision trees or neural networks. They are very robust and flexible, but demand careful pre-analysis and pre-selection of algorithms along with suitable parameters for the specific conditions of the incoming physical asset failure data. On the other hand, cryptic-based algorithms, which include support vector machines, remain particularly elusive as they seek to find a "best fit" failure prediction model with minimum error, without regard to the properties forcing the model.

Undoubtedly, predictive maintenance of IT equipment has become a major asset in sustaining today's enormous economic and social operations undertaken by both industrial and service site providers. Predictive asset maintenance is a data-driven strategy that aims to identify the most favorable maintenance policies based on past and real-time events, covering information not only pertaining to one single asset failure but also relating to the overall status of some or all assets. This valuable task enables operational or technical stakeholders to select an appropriate and efficient maintenance strategy, basing their selection criteria, in particular, on triggering and operating conditions. The operation of predictive maintenance, with its underpinning machine learning techniques, forms the core of the security and maintenance of IT equipment within the financial sector. The outcome of the development of the appropriate predictive model affects the business operations of a financial computer installation. Banking security experts agree that the illicit use of vulnerable machines in the financial sector infrastructure can lead to unauthorized transactions in a banking system. It is clear that too many false-positive or false-negative preemptive alerts can slow down the work of a financial institution.

3.1. Overview of Machine Learning Algorithms

Clustered analysis of the existing models and algorithms indicates that widely used models have been supervised and unsupervised learning methods. Furthermore, decision tree induction models have had the most applications due to their ease of use and ability to identify complex data patterns. Among these methods, the most prominent models have included random forests for equipment failure prediction. Neural networks and Proportional Hazards Models have recently started to attract greater research attention due to developments in deep learning. In the research literature, many studies used k-means clustering as an unsupervised learning method. As the above descriptions suggest, the research models vary according to the data categories used.

Supervised Learning Methods: Random Forest, Support Vector Machine, Neural Networks, Probabilistic Model, PHM, Logistic Regression, Naive Bayesian Model, Bernoulli Mixture Model, Feed Forward Neural Network, Extreme Learning Machine. Unsupervised Learning Methods: K-means clustering, Kohonen Map. Deep Learning Techniques: Autoencoder, Deep Neural Network. In general, the selection of models

again varies depending on the available data set and application scenario. For example, in one study, a model classification categorization based on data type was used. Given that time-to-failure data is rare in the finance dataset and that predictive maintenance is based on the usage and state of equipment, the proposed algorithm selection guide can be significant for financial institutions. Furthermore, the performance of algorithms may be altered based on data characteristics and environmental change. Thus, it is essential for financial IT maintenance organizations to establish a relevance model for the datasets in their hands and application scenarios based on supervised, unsupervised, or deep learning models. Model evaluation methods are crucial when selecting a model. This topic is addressed in the following section.

In a study of wind turbines, a PBHM that uses wind turbines' measurement data for condition-based monitoring and prediction of the future failure of power-generating components. The selection and data fusion of measurements are obtained via k-means clustering of one-year time series for a turbine's main components. The three PCAs describe generator degradation, blade degradation, and gearbox degradation. Based on the combination of the three PCAs, a PBHM is used as a predictive maintenance model for root-cause failure prediction. This model is evaluated and adjusted during periods without or with few failures. Evaluation solely occurs during periods of few failures, as this is when a higher number of turbines' equivalent operating time is available for forecasting maintenance intervention.

3.2. Data Collection and Preprocessing

The performance of predictive maintenance models depends significantly on data. Comprehensive data should be collected from various relevant sources for accurate and precise predictions of equipment failures. Sensor data, log files, and transaction records are a few examples of data sources for financial IT infrastructure. The complete process of collecting, preprocessing, and the entire safety data handling process should be carried out in compliance with financial regulations. Care should be taken to ensure clean data because inaccurate data will not lead to accurate predictive maintenance, resulting in high costs. For predictive maintenance, sensor data should be collected and integrated into a data warehouse to facilitate analytics. Data cleaning or data preprocessing eliminates inconsistencies, noise, and irrelevance from the data. Data transformation or data integration consolidates data from multiple sources to allow

efficient and coherent analytics. Modern modeling techniques combined with big data technologies integrated with data warehousing can scale to handle diverse streaming and static data for engineering features, computation of models, and customer-facing instances.

Feature selection and engineering are vital to developing accurate models. Supervised learning, unsupervised learning, and reinforcement learning are the methodologies to recognize or develop assets from sensor or maintenance log data. It is clear that feature selection and engineering are data preprocessing steps that are required for model development. In conclusion, data warehousing and real-time data processing allow analysts to track vital infrastructure components from sensor data, maintenance logs, repair logs, and historical transactions. While there are some underlying limitations around cost, scalability, and technology, the foundation for applying ML/AI in this critical area of banking infrastructure is strong. Those limitations can be mitigated with a phased strategic approach.

4. Implementation and Case Studies

Many predictive maintenance solutions operate as additional components to integrate into existing systems. The high level of complexity and strict regulations in the financial sector can, however, pose a challenge when trying to incorporate these solutions into existing systems. The various teams and departments work with different software and systems that need to be included in any predictive maintenance solution. When conducting the implementation, the different teams and departments actively work together from the outset, including input about any different requirements they may have. This level of collaboration between the different stakeholders is instrumental in subsequently gathering data.

A number of different considerations must be accounted for when implementing predictive maintenance within IT infrastructure, ranging from selecting the right tool that is compatible with existing systems, databases, and formats, the required level of training of staff, and the availability of support. Beyond the technical aspects, practitioners need to ensure full compliance with privacy and legal regulations, as well as corporate security policies. What becomes clear is the technicalities; the tool must provide the team with a significant look into the future and require little time to install and understand, given that innovators simply would not have had spare time. Several

case studies demonstrate the successful implementation of predictive maintenance. In the energy sector, predictive maintenance has been used on solar rooftop panels to anticipate faults, reduce downtime, improve worker safety, and save money. In early 2022, the first implementations of predictive maintenance have proven successful in some use cases within the financial sector. Financial innovators have successfully used predictive maintenance in the trading system to predict and prevent problems before they happen, ensuring business as usual in volatile markets. In a different use, predictive maintenance tools are used to monitor a mission-critical application as part of a required extra level of checks before any step up in access is granted. Predictive maintenance was used to foresee a potential issue, resulting in a manual check being made, which highlighted an apparent issue. This has demonstrated the capability to prevent potential financial loss by monitoring and predicting the health of the systems using AI-based predictive maintenance.

4.1. Integration of Predictive Maintenance Tools

Maintenance tools help enterprises identify infrastructure problems and make quick decisions by predicting possible outages. With an increase in the number of predictive maintenance tools and the functionalities they offer, it is essential to position, enhance, and maintain them in the operational environment in financial institutions. In this subsection, we discuss several maintenance tools available in the market, the range of functionalities they offer, and how they can be embedded within the IT infrastructure with the existing tool systems.

The purchased system decision based on available tools is followed by planning in a way that fits the operational environment according to environment readiness, software installation, infrastructure setup, reconciliation with existing business applications, resource allocation, and delivery schedules. Deployment encompasses the execution of software within the operational environment; this phase includes installation, customization, and configuration. After the executable integration, the entire application unit is tested or validated to verify the functionality of the maintenance tool, including the infrastructure components. Once integrated, financial IT users should become accustomed to leveraging the output of these tools as their day-to-day base through proper user exposure and user training schedules. Organizations could face significant challenges related to data silos and bottlenecks in the integration of tools in various

environments. Some possible practices for the smooth implementation of predictive maintenance tools are proposed, following a certain predefined roadmap.

Successful integration can help financial firms improve the real-time view of the operational environment and infrastructure, make quicker decisions during concealed events or actions, and implement curative pre-steps during infrastructure degradation situations. The roadmap includes tools capable of AR, VR, simulation, neuronal analytics, data analytics, anomaly detection, and AI-based prediction and prescription of infrastructure breakdown leading to a lower level of acceptable service.

4.2. Real-world Applications and Success Stories

Predictive Maintenance at Rabobank: The objective of this predictive maintenance initiative is to reduce the number of incidents in population sizing, improve time to repair and availability, and save on maintenance costs. Key takeaway: Statistical failure analysis showed that aging and inspections put high pressure on system reliability, resulting in high failure rates. We addressed this issue by using tools such as reliability testing, failure modes and effects analysis, modeling, and predictive maintenance strategies. As a result, the time needed to repair is reduced by 38%. Customer downtime was reduced by 44%. The modeling tool was the first to combine object-oriented fault tree modeling and Monte Carlo simulation, predicting the future failure rate under the usage of maintenance efforts. We are the first in the world to accomplish this for a multiple vendor open system platform environment divided into shared components and customer products! The tool is also very useful for benchmarking components by setting them for 100% utilization and then comparing the predicted results with another component.

Predictive Maintenance for FATplus at ABN AMRO: Key takeaway: While both Rabobank and ABN AMRO were dealing with similar problems, i.e., minimizing downtime due to releases and maintenance efforts for data centers and the FATplus products, their solutions were different. While Rabobank focused on the thermal value of statistics to improve the availability and downtime of FATplus, ABN AMRO focused on the packages being used by both IT systems and the behavior of those systems. Interestingly, both financial institutions implemented advanced IT predictive maintenance tools.

5. Challenges and Future Directions

Now, we discuss the challenges and future research directions that require further investigation and innovation for enabling the real-focused application of AI-based project management. The typical financial network is composed of proprietary and COTS software and mainstream but closed-spec hardware. The complexity, heterogeneity, and closed structure assembly of financial IT infrastructure pose several challenges for developing an automated project management solution. Facing those challenges requires later machines and AI methods to mimic human IT operations and support the knowledge-based customer project management at scale. We summarize those challenges and future research directions and some suggestions for immediate and actionable further directions of interest to the industry and academia. The inefficiency of current AI-based project management is challenging for the stakeholders who rely on hidden real-time and mission-critical project management of the daily operations of unstable transmission of digital information. This is unconsidered in the investment and insulated IT infrastructure, which is rigid in evaluating the performance of financial institutions, unable to set and adjust customer maintenance management in the overall operations strategy, unable to tune the evaluation objectives more focused on the performance of the business, unable to directly model the environment without knowledge, and unable to optimize algorithms tailored to the financial sector with narrow, insufficient, and painstaking real problems. Therefore, developing AI methods that cater to the needs of practical and immediate action is a crucial and stipulative direction that is desperately in need of attention.

5.1. Key Challenges in Implementing AI-Based Predictive Maintenance

Key Challenges in Implementing AI-Based Predictive Maintenance of Financial IT Infrastructure

The following typical challenges occur when applying AI techniques for predictive maintenance of banking IT infrastructure.

Lack of large amounts of labeled data. The shortage in the number of labeled objects is always an issue for any area of machine learning. In the case of predictive maintenance for IT infrastructure problems, this problem is typical, since, in many cases, the volume of observable negative examples is scarce or contains only a 1-1 correspondence with each other.

Selection of the appropriate types of labeled data. The given maintenance problem can be solved using different types of labeled data. For instance, on the lower functional layers of IT services, it is easy to gather status and health metric descriptors and use this data for model development. In the case of higher-level functional layers of services, predictions could be developed using a service-oriented architectural description of business processes with service components and service level objectives as the target model. However, such solutions have their own drawbacks: the chosen labels may not hold performance relations or may not make any sense for IT operations and system administrators to use.

5.2. Emerging Trends and Technologies

There are emerging trends, technologies, and methodologies in the areas of predictive maintenance and fault detection that can be ported to the domain of financial IT infrastructure. These include concepts such as dynamic Bayesian networks, a combination of inherent and observed characteristics of the system, evidence theory, adopted KPIs, early diagnostics, and early warning systems. Additional focus should be put on representing, evaluating, storing, and handling the results of predictive maintenance systems that deal with financial tasks.

In the field of financial predictive maintenance and risk modeling, it is important to design, develop, implement, and economize the systems of financial predictive maintenance and database solutions for the predictive modeling of financial risk. Particularly, this refers to risk management and Basel III requirements. In database and knowledge base system management, it is important to go beyond traditional management and establish the best practices guiding novel databases and big data systems. The implementation of data management of business rules is crucial, particularly with respect to Basel-related procedures. The active support to financial institutions for Basel compliance is crucial in today's society. Data management of results from Bayesian, decision, lattice, similarity, and topological logics is becoming routine in financial institutions.

The focus of the database and knowledge base community should be concentrated on the following suggestions: storability of database and knowledge base implementations designed for financial IT infrastructure, data integrity from the designed financial maintenance operations, and knowledge base management systems collected from the

implemented database and knowledge base business tasks. We would like to point out that an important orientation of the database community should be the creation of subject-related networks among researchers in predictive maintenance, financial risk modeling, financial IT infrastructure, and databases. These networks would be based on knowledge exchange and joint creation of high-quality software tools. They will allow us to promote high-quality research to emerging leaders, researchers, and non-academic users who need efficient cooperation in integrating software maintenance. The proposed emerging community will provide a forum integrating database and knowledge base operation with their archiving, which represents a fruitful basis for many insights useful to the academic society. In this way, real assistance will be offered for creating strong potential milestones of software maintenance engineering.

6. Future Direction

The above concept of predictive maintenance will evolve as machine learning algorithms and applications are developed further for uses such as IoT and blockchain. The role of machine learning models and various packages and libraries will be profound and on a larger scale. Predictive analytics will commence from the root level to replace preventive maintenance with the help of future off-the-shelf IT infrastructure experts. Later, these analytics will replace these experts, so organizations will start converting their experts to data scientists.

Future machine learning models and predictive maintenance IT infrastructure will reduce spare parts costs, increase service life, and lower maintenance costs. Extensive adoption of PLL seems to be better for smaller databases and financial IT infrastructure with smaller components. Developing long short-term memory networks will have the ability to predict trends in the long run. Applications will use blockchains to maintain historical data and future predictor IT infrastructure. It seems that tools and calls will be developed to report tool and fleet performance and any issues. With new companies entering predictive analysis, the importance of the regulatory landscape will increase. The introduction of new statistical and machine learning models should make them user-friendly. It is expected that new predictive maintenance models will support infrastructure security researchers by providing data from the tools, allowing researchers to develop percentile thresholds to study IT security.

The growth of machine learning and sophisticated sensor technologies is rapid, and the availability of a large number of data analytics tools will be present in the near future with successful case studies. These will come from the tools side, accuracy of the data, knowledge repositories, diagnostics, prognostics, and also on the economic value of the data. The future goal would not be limited to the use of locator-based solutions that connect models, as models vary in scope and specificity. It is expected that progress will be made in integrating these diverse approaches to create more holistic applications that allow users to select the most respected basic infrastructure with various options to tailor the analysis.

7. Conclusion

The end of this essay review for the special issue of Sustainable Operations and Computers in Finance is going to provide common results and give an overview of the developments in machine learning and AI, as well as advance data analytics more generally for redesigning financial IT operations and delivering new predictive maintenance strategies. The review has checked wide implications of unsuitable maintenance models for the finance and financial IT sectors for disruption recovery, digitization, new financial markets, and soft controls, as well as for the reduction of costs of capital and hence the WACC. In addition, they are suitable to reverse the trend of increasing time for IPO processes, which means that their use has also societal and regulatory implications. Moreover, it has been shown how the massive integration of machine learning and advanced analytics in predictive and prescriptive maintenance strategies for physical assets has produced a profound transformation for the sector, allowing the organization to address problems that have been for years under the famous Pareto. Almost synergetically with the rise of predictive maintenance, we finally proposed preventive maintenance from data for the so-called intangible assets starting from IT infrastructure, which is the base of operations for all other products and services, and we will see a shift towards new financial products and services. The need for wide use of this kind of data-driven approach has also been reported by several organizations that promote the successful use of data or by those consulting organizations collaborating to arrive at trustworthy AI. Above this overview, the review discusses both the main current trends and the open challenges, offering direct investigation opportunities, including those that, due to managerial inconvenience, like bias in criticality assessment, strongly condition the performance of final strategies.

Finally, the essay pointed out how carrot and stick techniques can be suitable roadmaps for organizational change in the direction of predictive maintenance, allowing for the overcoming of the spendthrift maintenance heuristic, developing a common understanding of the future priorities and strategic direction of the organization(s), and getting managerial strategies for exploiting AI at its best in predictive IT maintenance. In conclusion, regardless of the underpinning techniques, context, and final aim, the continuous, pervasive, and permanent change of information systems and IT infrastructure makes continuous maintenance mandatory and desirable.