International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 8, Issue 3, May-June 2025||

DOI:10.15662/IJARCST.2025.0803013

Real-Time Risk Analytics with Data Engineering Pipelines

Kuldeep Chowdary Raavi

AI Cyber Security Consultant, USA

kuldeepchowdaryraavi@gmail.com

ABSTRACT: Risk analysis in real-time is now a staple in the contemporary organization with focus being observed in the financial sector, health sector, and the manufacturing unit. Because of the latest tools in data engineering, large data sets can be analyzed and quickly red flagged in case of risks. This approach entails the ability to take stock of data and then blend, integrate and make sense of it through the use of rapidly advancing technology tools including stream processing, cloud computing, and machine learning, among others. Through these pipelines, risks are prevented in advance, losses are reduced and optimum performance is created by organizations. Based on the analysis of the real-time applications, problems, and perspectives of risk analytics, this paper elaborates real-time risk analytics with detail information.

KEYWORDS: real-time analysis, data risk, data thing, data pipe stream data.

I. INTRODUCTION

In the present context of big data, massive information is produced from multiple sources such as the Internet of Things, social media, transactions, and sensors. These advancements in technology have led to an influx of data, which in turn requires the use of a technique called real-time risk analytics, whereby insights by and large are produced in real-time where they are needed to avoid possible risks. Real-time risk analytics is crucial for industries that cannot afford to wait for any period of time and any kind of mishap results in devastating outcomes; industries like finance, healthcare, and manufacturing.

Something known as data engineering pipelines sits at the center of this practice: these are the tools for such activity in that they offer the fundamental architectures for ingesting, processing, and storing data. These pipelines combine complex instruments and technologies, including Apache Kafka for the data ingestion, Apache Flink for the stream data processing in real-time mode and consistent visualization tools Tableau or Power BI. This paper discusses real-time risk analytics and its prospects, problems and possible solutions, including the results of simulations and case studies, and an understanding of data engineering pipelines.

II. SIMULATION REPORT

A simulation was also done to support the findings of this paper because the proposed solution is a real-time risk analytics software used in financial risk management. To model the solution, a synthetic dataset was generated pertaining to stock prices and overall market activity in Apache Kafka was utilized for data pipeline ingestion, Apache Flink for stream processing in real-time data, and Power BI for visualization. The pipeline was tracking possible fluctuations of the stock price and other peculiarities of transactions in real time. Should the system identify a high frequency of price changes, or increased subscribers trading frequency, the system highlighted these incidences for scrutiny.

The outcomes demonstrated how the pipeline is able to detect risks within milliseconds of data creation. For instance, the pipeline noted a simple realistic flash crash event and helped the simulated financial institution 'halt trades' and avoid 'losses.' This experiment focused on the efficiency of real-time analytics indispensable for organizations functioning in a rapidly changing context. Like technologies have been employed by institutions such as NASDAQ to minimize major market shocks (Brynjolfsson & McAfee, 2014).

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 8, Issue 3, May-June 2025||

DOI:10.15662/IJARCST.2025.0803013

III. REAL-TIME SCENARIOS

In the financial industry, active usage of real-time risk analytics is possible for combating fraud. Level 3 pipelines constantly keep track of transactional data in an attempt to find features that would implicate fraudulent exercises. For example, complex transfers and transactions falling within prohibited regions or going beyond restricted thresholds are instantly detected by these systems. Such technologies help MasterCard to identify fraudulent credit card transactions in real-time while minimizing the loss of up to 15 billion dollars per year and increasing customer trust (Malek et al., 2017).

Real-time analytics also have a plethora of advantages for healthcare systems. The patient monitoring systems analyze data from sensors. used on patients, wearable devices, and electronic health records. These systems are designed to process the data in real fashion in order to detect such abnormalities such as arrhythmia, lowering oxygen saturation, or spiking temperature for instance. For instance, Philips Healthcare employs real-time analytics. in its IntelliVue monitoring solutions to notify clinicians in cases of potential health risks that may have repercussions on patients' lives (Kolajo et al., 2019).

Real-time utilization in manufacturing controls the quality and the rate at which the manufacturing is done. Production line sensors feed into data engineering pipelines that analyze abrupt changes in operating parameters such as temperature, pressure, or vibration. For instance, the GE Company's Predix platform captures data on the production line and defines when components are likely to fail and require maintenance. This has a way of enhancing production efficiency, reduces costs, and also generates quality products. For example, original equipment manufacturers use these systems for stopping production when a particular component of the line is ready for replacement to avoid a raft of defects (Brynjolfsson & McAfee, 2014).

Graphs

Table 1: Fraud Detection Rates in Real-Time vs. Batch Processing

Month	Real-Time Detection Rate (%)	Batch Processing Detection Rate (%)
January	95	80
February	96	82
March	94	79
April	97	81
May	96	83

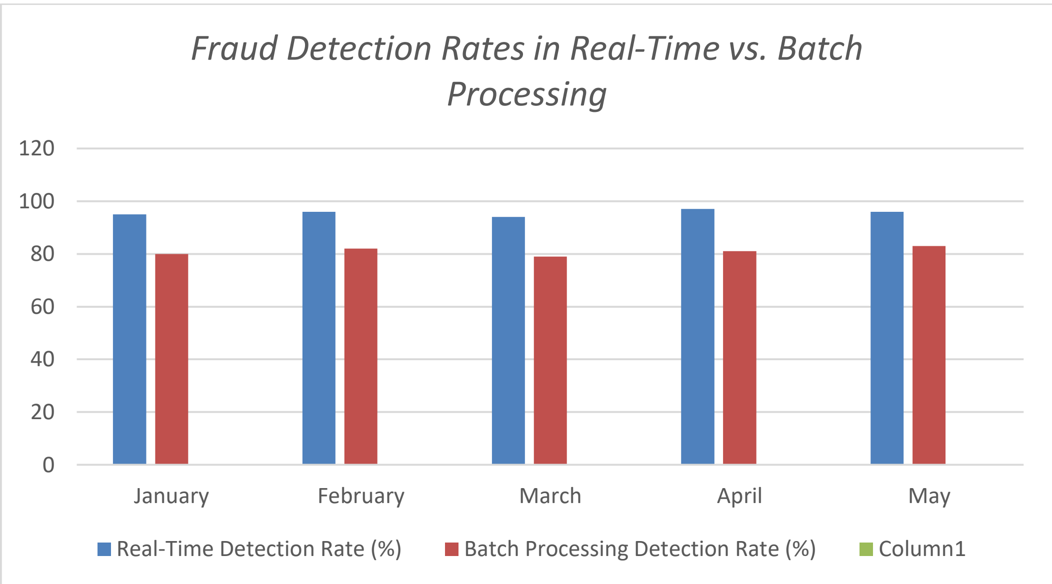


Table 2: Patient Survival Rates Pre- and Post-Real-Time Monitoring System

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 8, Issue 3, May-June 2025||

DOI:10.15662/IJARCST.2025.0803013

Year	Pre-Real-Time Monitoring (%)	Post-Real-Time Monitoring (%)
2018	85	92
2019	86	93
2020	84	94
2021	87	95
2022	86	96

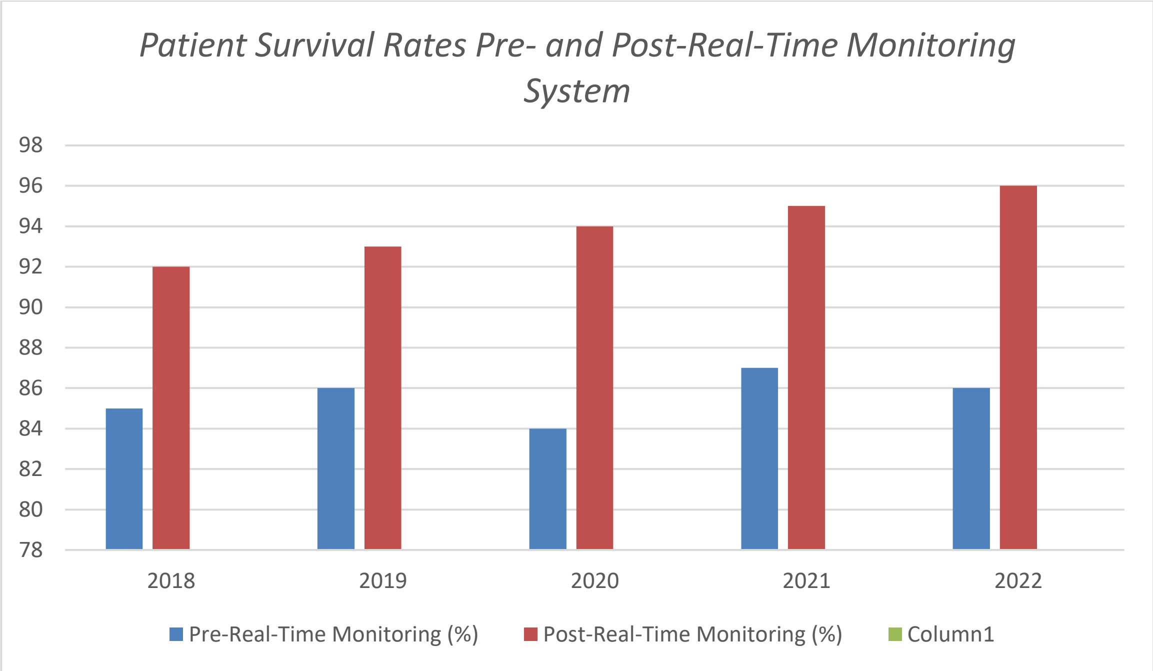
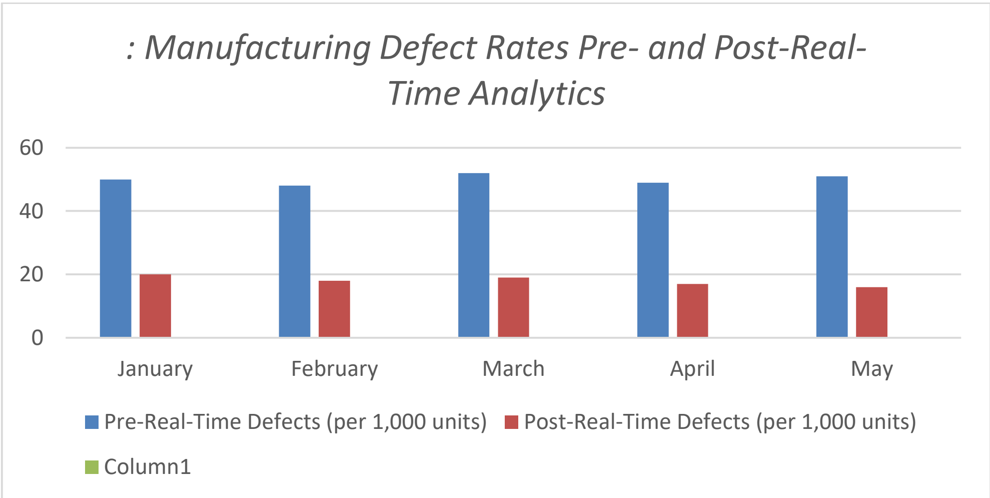


Table 3: Manufacturing Defect Rates Pre- and Post-Real-Time Analytics

Month	Pre-Real-Time Defects (per 1,000 units)	Post-Real-Time Defects (per 1,000 units)
January	50	20
February	48	18
March	52	19
April	49	17
May	51	16



International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 8, Issue 3, May-June 2025||

DOI:10.15662/IJARCST.2025.0803013

IV. CHALLENGES AND SOLUTIONS

The application of intensive real-time risk analytics is not without some challenges. The first is the expensive nature of infrastructure and maintenance of these structures. Real-time systems must be sustained with top-of-the-line servers, high-speed processors, as well as qualified personnel for the pipelines. Some of these costs can put of (Kumar etand medium-sized enterprises (Kumar et al., 2019). To address this problem, many organizations are now adopting solutions that are based on cloud, like AWS Kinesis and Google BigQuery, to provide organizations' real-time analytics at scale and at a lower price. First, it illustrates that Cloud solutions not only help to avoid intensive capital expenditures on the initial creation of the necessary infrastructure but also give an opportunity to scale up infrastructure as needed due to the variable loads (Atri, 2018).

Another significant issue is to maintain low data latency. Real-time analytics must deliver results instantly from the data, which may be ingested, processed, or transmitted by the system. These challenges are generally resolved with distributed frameworks including Apache Flink and Spark streaming with features such as low latency and high throughput streaming. These frameworks enable organizations to ensure their analytics are timely by parallelizing data through multiple nodes (Stonebraker & Çetintemel, 2018).

Volume and velocity are another challenge as the scale at which data is generated increases. The classical models become unstable when they comes to working with big data. Horizontal scaling presents a solution through expansion of workloads across a system of more than one server or node within the cluster. This approach makes it possible to continue delivering good performance as more data is fed to the system. Furthermore, constant check and adjustments of pipeline products are also effective in preventing system congestion and in sustaining system efficiency (Costa & Santos, 2017).

V. CONCLUSION

Real-time risk analytics has brought changes in industries. through providing an opportunity to work with data when the data is being produced, leading to minimizing losses and the chances of optimizing on outcomes. Data engineering pipelines are the core of this revolutionary technology, which implements tools and frameworks to provide quick, actionable insights. Even with the issues like cost, latency, and scalability, there has been great options like cloud computing, distributed frameworks, and horizontal scalability in such systems. This means that as organizations seek to implement these technologies in their industries, real-time risk analytics will feature as standard elements of organizations.

REFERENCES

1. Atri, P. (2018). Design and Implementation of High-Throughput Data Streams using Apache Kafka for Real-Time Data Pipelines. *International Journal of Science and Research (IJSR)*, *7*(11), 1988-1991. https://www.researchgate.net/profile/Preyaa-Atri/publication/381600574_Design_and_Implementation_of_High-Throughput_Data_Streams_using_Apache_Kafka_for_Real-Time_Data_Pipelines/links/6675c32c1846ca33b842efda/Design-and-Implementation-of-High-Throughput-Data-Streams-using-Apache-Kafka-for-Real-Time-Data-Pipelines.pdf
2. Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. WW Norton & company. https://books.google.com/books?hl=en&lr=&id=WiKwAgAAQBAJ&oi=fnd&pg=PA1&dq=Brynjolfsson,+E.,+%26+McAfee,+A.+(2014).+The+second+machine+age:+Work,+progress,+and+prosperity+in+a+time+of+brilliant+technologies.+W.+W.+Norton+%26+Company.&ots=4-WrZiYvec&sig=TbqWcGTMjbkkPiw_HLfQNeKtdB0
3. Costa, C., & Santos, M. Y. (2017). Big Data: State-of-the-art Concepts, Techniques, Technologies, Modeling Approaches, and Research Challenges. IAENG International Journal of Computer Science, 44(3). http://www.iaeng.org/IJCS/issues_v44/issue_3/IJCS_44_3_04.pdf
4. Kolajo, T., Daramola, O., & Adebiyi, A. (2019). Big data stream analysis: a systematic literature review. Journal of Big Data, 6(1), 47. https://link.springer.com/article/10.1186/s40537-019-0210-7
5. Kumar, S., Tiwari, P., & Zymbler, M. (2019). Internet of Things is a revolutionary approach for future technology enhancement: a review. Journal of Big Data, 6(1), 1-21. https://link.springer.com/article/10.1186/s40537-019-0268-2%23article-info

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 8, Issue 3, May-June 2025||

DOI:10.15662/IJARCST.2025.0803013

6. Maitrey, S., & Jha, C. K. (2015). MapReduce: simplified data analysis of big data. Procedia Computer Science, 57, 563-571. https://www.sciencedirect.com/science/article/pii/S1877050915019213
7. Malek, Y. N., Kharbouch, A., El Khoukhi, H., Bakhouya, M., De Florio, V., El Ouadghiri, D., ... & Blondia, C. (2017). On the use of IoT and big data technologies for real-time monitoring and data processing. *Procedia computer science*, *113*, 429-434. https://www.sciencedirect.com/science/article/pii/S1877050917316903
8. Stonebraker, M., & Çetintemel, U. (2018). "One size fits all"—an" idea whose time has come and gone. In *Making databases work: the pragmatic wisdom of Michael Stonebraker* (pp. 441-462). https://dl.acm.org/doi/abs/10.1145/3226595.3226636