

# CHAPTER 1

## INTRODUCTION



**Fig. 1.1: Prof. Wil Van Der Aalst**

### **1.1 History of Process Mining:**

The term Process mining was first coined in a research proposal written by the Dutch computer scientist . Thus began a new field of research that emerged under the umbrella of techniques related to data science and process science at the Eindhoven university in 1999. In the early days, process mining techniques were often convoluted with the techniques used for workflow management. In the year 2000, the very first practically applicable algorithm for process discovery, alpha miner was developed. The very next year, in 2001, a much similar algorithm based on heuristics called "Heuristic miner" was introduced in the research papers. Further along the link more powerful algorithms such as inductive miner were developed for process discovery.

As the field of process mining began to evolve, conformance checking became an integral part of it. The year 2004 earmarked the development of performance

analysis for conformance checking purposes. Apart from the mainstream techniques of process discovery and conformance checking, process mining branched out into multiple areas leading to the discovery and development of “performance analysis” “Decision mining” and “Organizational mining” in the year 2005 and 2006 respectively. In the year 2007, the first-ever commercial process mining company "Futura Pi" was established. The "IEEE" task force on PM, a governing body was formed in the year 2009 that began to overlook the norms and standards related to process mining. Further techniques were developed for conformance checking which led to the publishing of "Alignment based conformance checking" in the year 2010. In 2011, the first-ever Process mining book was published. Further along in 2014, a MOOC course was offered by coursera on Process mining. By the year 2018, nearly 30+ commercially available process mining tools were in the picture. The year 2019 earmarked the first Process mining conference. Today we have over 35 vendors offering tools and techniques for process discovery and conformance checking.

Process mining should be viewed as a bridge between data science and process science. Process mining focuses on transforming event log into a meaningful representation of the process which can lead to the formation of several data science and machine learning related problems. Process mining is a powerful and innovative approach to understanding and improving business processes by analyzing event data recorded during their execution. It provides valuable insights into how processes actually work in practice, helping organizations identify inefficiencies, bottlenecks, and areas for improvement. Process mining combines data science, data analytics, and business process management to provide a comprehensive view of processes, enabling data-driven decision-making and process optimization.

## **1.2 Key Concepts:**

- 1. Event Logs:** Process mining relies on event logs, which are chronological records of activities, decisions, and data points within a process. These logs capture the details of each event, including timestamps, activity names, actors, and outcomes.

2. **Process Discovery:** This is the initial step in process mining, where algorithms analyze event logs to automatically construct process models. These models visualize the sequence of activities, decision points, and variations in the process flow.
3. **Conformance Checking:** Once a process model is created, it can be compared to the actual event data to identify deviations or discrepancies. This helps in pinpointing areas where the actual process doesn't adhere to the expected or desired flow.
4. **Enhancement and Optimization:** By analyzing event data, process mining can uncover bottlenecks, inefficiencies, and opportunities for improvement within a process. This information is crucial for optimizing processes to increase efficiency and reduce costs.
5. **Metrics and Performance Analysis:** Process mining provides insights into various process metrics, such as cycle times, execution times, waiting times, and frequency of activities. This information helps in measuring and evaluating process performance.
6. **Process Compliance:** Process mining can be used to ensure that processes adhere to regulatory standards, industry norms, and internal policies. It helps identify deviations and non-compliance instances.
7. **Predictive Analysis:** Advanced process mining techniques can predict future process behavior based on historical data, aiding in decision-making and resource allocation.

## CHAPTER 2

### TECHNOLOGY

#### 2.1 Technology of Process Mining:

Process mining relies on various technologies to collect, analyze, and visualize event data from business processes. These technologies work together to provide insights into process execution, bottlenecks, inefficiencies, and opportunities for improvement. Here are some key technologies used in process mining:



**Fig. 2.1: Technology**

- **Event Logs:** Event logs are at the core of process mining. They consist of records detailing each activity, decision, and data point within a process. Event logs can be generated by various systems such as enterprise resource planning (ERP) software, customer relationship management (CRM) systems, and more.
- **Data Extraction and Integration:** Process mining tools need to extract data from different sources and integrate them into a unified event log format. This involves techniques like data extraction, transformation, and loading (ETL) to ensure data consistency and compatibility.
- **Data Warehousing:** Storing event logs efficiently is crucial. Data warehousing technologies are used to store and manage large volumes of data for analysis and reporting.

- **Big Data Processing:** As organizations deal with massive amounts of data, big data processing technologies like Apache Hadoop and Apache Spark can be employed to handle and process event logs efficiently.
- **Process Discovery Algorithms:** These algorithms automatically construct process models from event logs. They can be based on various techniques such as alpha algorithms, inductive mining, and genetic algorithms to create visual process representations like Petri nets or BPMN diagrams.
- **Process Conformance Checking:** Technologies like alignment algorithms are used to compare the discovered process model with the actual event log to identify deviations and conformance issues.
- **Predictive Analytics:** Advanced analytics and machine learning techniques can be employed to predict future process behavior based on historical data, aiding in decision-making and resource allocation.
- **Visualization Tools:** These tools convert process models and event data into visual representations, such as flowcharts, Gantt charts, and heatmaps, making complex process insights more accessible.
- **Simulation:** Simulation technologies allow organizations to model and test changes to processes before implementing them, predicting how modifications will impact process efficiency.
- **Machine Learning:** ML techniques can be used to uncover hidden patterns, anomalies, and correlations within event data, providing deeper insights into process behavior.
- **Natural Language Processing (NLP):** In cases where event logs include unstructured text data (e.g., customer feedback), NLP can be used to extract valuable insights from textual information.
- **Cloud Computing:** Cloud platforms can provide the scalability and resources needed to process and analyze large event log datasets.
- **Real-Time Process Monitoring:** Technologies that allow for real-time monitoring and analysis of ongoing processes, enabling quick response to deviations or anomalies.

- **Integration with BPM Tools:** Some process mining tools can integrate with Business Process Management (BPM) suites to bridge the gap between process design and execution.
- **Security and Privacy Technologies:** Given the sensitivity of process data, technologies like encryption, access controls, and anonymization are crucial to ensure data security and privacy.

Process mining technologies are constantly evolving as new methods are developed and existing techniques are refined. Organizations can choose tools and technologies based on their specific needs, the complexity of their processes, and the volume of event data they need to analyze.

## 2.2 Tools of Process Mining:

**Ui Path** is a global provider of an end-to-end automation platform that combines robotic process automation (RPA) with artificial intelligence, process mining, and cloud technologies to help companies scale their digital business operations. Its process mining product suite comprises process analytics, app templates, automated alerts, built-in data transformation, intelligent software robots, task mining, and other capabilities for analyzing data from business applications. UiPath is available as a cloud-based and on-premises solution.



**Fig. 2.2: Tools of Process Mining**

**iGrafx** Business Transformation Platform enables transformation by connecting strategy to execution while mitigating risk, ensuring compliance, and providing a framework for governance, resiliency, business continuity, and

continuous improvement. iGrafx's process mining technology combines artificial intelligence (AI) and machine learning to help users capture up-to-date details about how processes are working, identify opportunities for improvement, standardize processes, and assess which strategies could benefit from a shared services approach.

**Celonis** is a global provider of execution management solutions that help companies improve how they run their business processes. With Celonis' suite of process and task mining capabilities, companies across industries can improve visibility into their operations, identify bottlenecks, and streamline efficiencies. Those capabilities—powered by machine learning and industry-standard process query language (PQL)—include analytic visualizations, drag-and-drop customization tools, task mining, extensible data models, multi-event logs, best-practice benchmarking, and tools for identifying processes that could benefit from automation.

**Bizagi** is a leader in digital business process automation software. The vendor offers three tiers of solutions, including Bizagi Engine, Bizagi Studio, and Bizagi Modeler. Bizagi's process mining capabilities are included via the Enterprise model of Bizagi Modeler, which equips companies with the process mining tools they need to understand their processes. Other capabilities available with Bizagi Modeler Enterprise include value-chain diagrams, Single Sign-On, model sharing, private cloud storage, real-time notifications, and more.

**Skana** is an AI-powered process intelligence platform designed to help modern enterprises understand the telemetry of their digital work and use data to identify opportunities for improvement. The company's hybrid approach to task and process mining enables businesses to improve how they plan digital transformation initiatives, streamline their automation efforts, develop strategic employee training initiatives, and maintain organization-wide process improvement.

**IBM** Process Mining product suite uses data-driven process insights to help companies across markets improve processes and make faster, more informed decisions. IBM's process mining tools can be applied in use cases like intelligent automation, customer onboarding, procure-to-pay (P2P), accounts

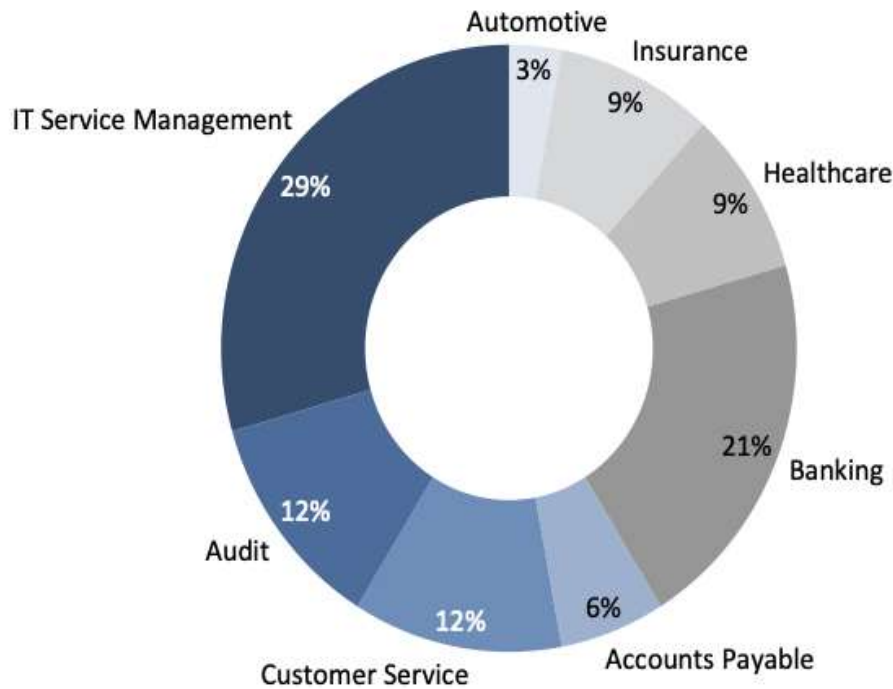
payable, IT incident management, and order-to-cash. Features include automated robotic process automation (RPA) generation, fact-based process models, AI-powered process simulations, conformance checking, task mining, and seamless integrations with leading software SAP, Oracle, and other IBM products.



## CHAPTER 3

### APPLICATIONS

#### 3.1 Applications of Process Mining:



**Fig. 3.1: Applications of Process Mining**

##### 1. Financial Services:

Because of the rise in transaction volume and the digitization of more industries, aberrant activity is harder to detect using manual methods. Companies in the financial services sector have the chance to continually and thoroughly identify issues within high-volume processes thanks to process mining, which is a solution to the increased regulatory and audit requirements.

##### 2. Telecommunications:

As subscriber quantities increase and activations become more and more automated, there is a greater danger of unsuccessful activations. When telecom companies get more orders, process mining gives them the chance to identify pricey issues and client blowback in their Order-to-Activation processes.

##### 3. Healthcare:

The risks associated with preserving population health and achieving individual patient journey objectives rise as data about patient experiences and results keep growing. Process mining supports the delivery of effective and high-quality end-to-end patient journeys for healthcare organizations dealing with the exponential growth of data, from before a first doctor appointment through treatment regimens to closed treatment cases.

#### **4. Retail:**

Due to technology or process problems, retail businesses have seen expensive consumer fallout from complicated e-commerce operations. Process mining assists merchants in ensuring that consumers can complete transactions efficiently and without issues despite rising transaction volumes.

#### **5. Digital Transformation:**

Process mining is frequently used in larger-scale digital transformation initiatives because it can give you the precise insights needed for process improvement, allowing systems to run more quickly, smoothly, and efficiently, as well as objective data-driven insights into the causes of delays and inefficiencies within business processes.

#### **6. Public works:**

Process mining has been used to streamline the invoice process for public works projects, which involve various stakeholders, such as construction companies, cleaning businesses, and environmental bureaus.

#### **7. Software Development:**

Since engineering processes are typically disorganized, process mining can help to identify a clearly documented process. It can also help IT administrators monitor the process, allowing them to verify that the system is running as expected.

## CHAPTER 4

### MODULES

#### 4.1: Steps of Process Mining:

Process mining typically involves several key modules or steps that collectively allow for the analysis, discovery, and optimization of business processes. Here's an explanation of the main modules often present in process mining:



Fig. 4.1: Steps of Process Mining

##### 1. Data Collection and Preprocessing:

- **Event Log Collection:** Gather event data from various sources within the organization, such as IT systems, databases, and application logs.
- **Data Preprocessing:** Clean, transform, and consolidate the event data into a standardized format suitable for analysis. This might involve handling missing data, removing outliers, and aligning timestamps.

##### 2. Data Integration and Transformation:

- **Data Integration:** Combine event logs from multiple sources to create a unified dataset for analysis.
- **Data Transformation:** Convert raw event data into a structured format that process mining tools can interpret. This might involve mapping activities, timestamps, and actors to standard identifiers.

##### 3. Process Discovery:

- **Automated Discovery:** Algorithms analyze event data to automatically construct process models, such as Petri nets, BPMN diagrams, or process flowcharts.
- **Interactive Discovery:** Analysts can interact with the process mining tool to fine-tune and validate the automatically generated process models.

#### 4. Conformance Checking:

- **Alignment Analysis:** Compare the discovered process model with the actual event log data to identify deviations and discrepancies.
- **Fitness Analysis:** Quantify how well the process model fits the actual execution data, providing insights into process adherence and efficiency.

#### 5. Performance Analysis:

- **Key Performance Indicators (KPIs):** Calculate metrics like cycle times, waiting times, and execution times to evaluate process efficiency.
- **Bottleneck Identification:** Identify activities that significantly contribute to delays and inefficiencies within the process.

#### 6. Variants and Patterns Analysis:

- **Variant Analysis:** Identify different paths and variations within a process, showing the frequency and sequence of activities.
- **Pattern Recognition:** Discover recurring patterns and sequences of activities that might indicate best practices or areas for improvement

#### 7. Simulation and What-If Analysis:

- **Process Simulation:** Simulate process execution based on historical data to test changes or optimizations before implementing them.
- **What-If Analysis:** Explore the potential impact of process changes on KPIs and performance metrics.

#### 8. Social Network Analysis:

- **Role and Interaction Analysis:** Visualize how individuals and roles interact within a process, revealing communication patterns and collaboration dynamics.

#### 9. Resource Allocation and Utilization:

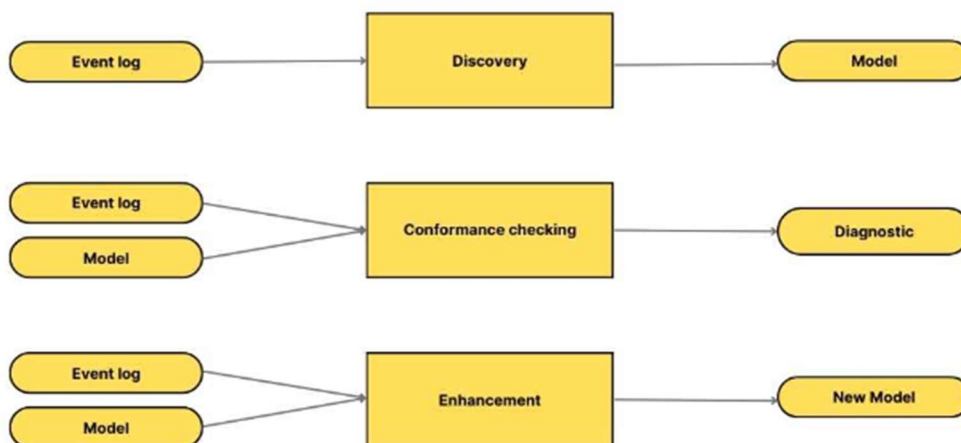
- **Resource Analysis:** Evaluate how resources (human or technical) are assigned and utilized throughout the process.
- **Resource Bottleneck Identification:** Identify resource-related bottlenecks and allocation issues.

## 10. Visualization and Reporting:

- **Process Visualization:** Represent process models, data, and analysis results in visual formats like flowcharts, Gantt charts, and heatmaps.
- **Report Generation:** Create comprehensive reports that summarize process insights, KPIs, and optimization recommendations.

Each module contributes to the overall process mining workflow, from gathering and preprocessing data to generating insights and actionable recommendations for process improvement.

## 4.2 Types of process mining:



**Fig. 4.2: Types of Process Mining**

- **Discovery:** Process discovery uses event log data to create a process model without outside influence. Under this classification, no previous process models would exist to inform the development of a new process model. This type of process mining is the most widely adopted.
- **Conformance:** Conformance checking confirms if the intended process model is reflected in practice. This type of process mining compares a process description to an existing process model based on its event log data, identifying any deviations from the intended model.

- **Enhancement:** This type of process mining has also been referred to as extension, organizational mining, or performance mining. In this class of process mining, additional information is used to improve an existing process model. For example, the output of conformance checking can assist in identifying bottlenecks within a process model, allowing managers to optimize an existing process.

### 4.3 Why is process mining important?

Increasing sales isn't the only way to generate revenue. Six sigma and lean methodologies also demonstrate how the reduction of operational costs can also increase your return-on-investment (ROI). Process mining helps businesses reduce these costs by quantifying the inefficiencies in their operational models, allowing leaders to make objective decisions about resource allocation. The discovery of these bottlenecks can not only reduce costs and expedite process improvement, but it can also drive more innovation, quality, and better customer retention. However, since process mining is still a relatively new discipline, it still has some hurdles to overcome. Some of those challenges include:

- **Data Quality:** Finding, merging and cleaning data is usually required to enable process mining. Data might be distributed over various data sources. It can also be incomplete or contain different labels or levels of granularity. Accounting for these differences will be important to the information that a process model yields.
- **Concept drift:** Sometimes processes change as they are being analyzed, resulting in concept drift.

### 4.4 Benefits of Process Mining:

1. **Visibility:** Process mining provides a clear, visual representation of complex processes, making it easier to understand how activities interact and where improvements can be made.

2. **Data-Driven Insights:** Process mining enables organizations to base their decisions on real data rather than assumptions, leading to more effective and informed strategies.
3. **Continuous Improvement:** By identifying inefficiencies and bottlenecks, organizations can continuously refine and optimize their processes for better performance.
4. **Quick Wins:** Process mining often reveals low-hanging fruit – easy-to-fix issues that can result in immediate improvements.
5. **Compliance and Risk Mitigation:** Ensuring processes comply with regulations and standards becomes easier with the ability to track and analyze events.
6. **Collaboration:** Process mining fosters collaboration between business analysts, data scientists, and process stakeholders, aligning everyone towards process improvement goals. Process mining has applications across various industries, including manufacturing, finance, healthcare, logistics, and more. As technology advances, process mining continues to evolve, offering increasingly sophisticated tools and techniques to analyze, optimize, and innovate business processes.



**Fig. 4.3: Benefits of Process Mining**

## **CHAPTER 5**

### **REAL TIME APPLICATIONS**

#### **5.1 Sales:**

➤ **Lead-To-Order:**

- **Reduced sales cycle time:** Lead-to-order processes can take a long time. This causes the payback time of marketing investments to increase. Companies can uncover the reasons behind this issue and, take action to reduce sales cycle time.
- **Increased conversion rate:** Converting marketing strategies into sales is critical for companies. With a process mining tool, companies can discover if they have proper strategies for increasing conversion rates.

➤ **Order-to-Cash:**

- **Order-to-cash (O2C)** is the step-by-step process from when a customer places an order until when the company gets paid and ships out the goods. Process mining can help identify all the little bottlenecks that might exist in the process that hamper a smooth operation.

➤ **Increasing on-time delivery:**

- For customer satisfaction, on-time deliveries are essential. Companies can use process mining to uncover the reasons behind late deliveries.

➤ **Identifying reasons that hurt monthly revenue:**

- Companies may lose a portion of their earnings during this process due to long-running holds or order cancellations.

➤ **Locating key regions:**

- With process mining, the companies can detect their high-value customers and critical areas to focus on those areas.

➤ **Identifying root causes of order changes:**

- Customers sometimes change their orders which causes processes to take a longer time. The un clarity of pre-order stages might cause these



order changes. Companies prefer to decrease these order changes to stabilize their processes.

➤ **Benchmark the amount of returned goods:**

- Companies can discover the value of returned goods by using process mining. According to this insight, they can focus on improving their order-to-cash processes.

## **5.2 Finance:**

Here we cover applications of process mining in the finance function of companies (not financial services industry specific applications):

### **1. Purchase-To-Pay :**

➤ **Identifying manual steps to automate:**

- Mistakes and manual interventions made in purchase-to-pay processes increase lead time. Companies can use process mining as a solution to find compatible areas of automation. With that, the company can improve its first-time-right rate by diminishing mistakes and reduce its rework time. In some process mining tool can increase automation by 35% and decrease rework time by 52%.

➤ **Eliminate maverick buying:**

- Companies can mine their purchase-to-pay processes to reduce maverick buying. If companies have a specific problem on maverick buying, they may find certain areas to improve the use of frame agreements with process mining. Process mining vendors claim that they can detect maverick buying by following the rules below:
- A receipt shouldn't be generated before a purchase order is created
- All invoices should be created after a PO
- PO without a contract should not exist (especially if the order is large in quantity and happens regularly)

➤ **Uncovering root causes for delays:**

- Process mining enables companies to pinpoint which suppliers, products or departments cause delays. By taking relevant actions, they can achieve more on-time deliveries internally.

## **2. Accounts Receivable:**

➤ **Discover actions to encourage on-time payments:**

- Customers don't always pay on time. Companies can't collect their receivables on-time in the end, and this might affect other processes. Process mining can identify the causes of this problem and find appropriate solutions.

➤ **Quicker invoicing:**

- Billing your customers is another process that can become expensive and complicated from time to time. Process mining discovers the bottlenecks in the invoicing process and may find ways to automate it. As a result, it is possible to diminish invoice costs and provide quicker invoicing.

## **3. Accounts Payable:**

- **Reducing late payments:** Companies can mine their business processes to uncover the reasons for their late payments. By fixing these inefficiencies, companies can diminish late payments and improve cash discounts.
- **Identifying real reasons behind incorrect invoices:** Mistakes on the invoices or duplicate payments are common issues that cause extra workload. Companies can identify the reasons for these cases with process mining

## **5.3 Audit:**

- **Compare “before” and “after”:** When a company makes a change in its process, verifying the improvement may be a challenge. For

consultants, process mining enables consultants to relate “before” and “after” of the processes.

- **Improve response time:** While traditional process discovery may take months, process mining is faster. As a result, consultants like EY can complete end-customer process analysis within days using process mining tools
- **Risk identification:** Process mining assures data-based information to consultants. With insights, consultants can identify risks and advise companies accurately.

#### 5.4 Customer Service:

- **Cross-channel analysis to identify anomalies:** Process mining software can help analyze process steps across different channels to identify compliance issues and inefficiencies.
- **Map customer journey:** Process mining tool can illustrate the customer journey in a given channel by extracting data from CRM and ticket systems). By doing so, process mining facilitates tracking customer experience, challenges that customers face, and interactions between responsible agents and customers.

## CHAPTER 6

### LEARNING OUTCOMES

Participating in a process mining internship can provide us with valuable skills, insights, and experiences that contribute to our personal and professional development.

- 1. Understanding Business Processes:** You'll gain a deeper understanding of how businesses operate by directly working with real-world processes and data. You'll see how different activities, decisions, and interactions come together to achieve specific goals.
- 2. Data Analysis Skills:** Process mining involves working with event data, which hones your data analysis skills. You'll learn how to clean, preprocess, and analyze data to derive meaningful insights.
- 3. Process Modeling:** Through process discovery, you'll learn how to create visual process models using tools like BPMN diagrams or Petri nets. This skill is valuable in various industries, from business to IT.
- 4. Conformance Checking:** You'll understand how to compare expected process models with actual event data to identify deviations and areas of improvement, enhancing your ability to ensure process compliance.
- 5. Performance Evaluation:** Process mining often involves calculating key performance indicators (KPIs) to assess process efficiency. This teaches you how to measure and quantify process performance.
- 6. Bottleneck Identification:** Identifying bottlenecks and inefficiencies is a crucial skill learned during process mining. You'll learn how to pinpoint areas that need optimization.
- 7. Data Visualization:** Process mining results are often presented visually. You'll gain experience in creating informative and engaging visualizations to communicate insights effectively.
- 8. Problem-Solving:** As you analyze processes and discover inefficiencies, you'll develop problem-solving skills by devising solutions to address these challenges.
- 9. Communication Skills:** You'll learn to convey complex process insights to both technical and non-technical stakeholders. Clear communication is key to implementing process improvements.

- 10. Collaboration:** Process mining often involves working with cross-functional teams, such as business analysts, data scientists, and IT professionals. You'll enhance your ability to collaborate effectively.
- 11. Process Improvement:** Through real-world scenarios, you'll understand how process mining leads to actionable improvements. You'll see firsthand how data-driven decisions can enhance operations.
- 12. Hands-on Experience:** Internships provide hands-on experience that complements theoretical knowledge gained in classrooms. You'll apply concepts in a practical setting.
- 13. Professional Networking:** Internships expose you to professionals in the field, offering opportunities to build a network of contacts who can offer guidance and potential future opportunities.
- 14. Resume Building:** Completing a process mining internship enhances your resume by demonstrating practical skills and relevant experience to potential employers.
- 15. Industry Insights:** Depending on the organization's sector, you might gain insights into specific industries like finance, healthcare, manufacturing, etc., broadening your knowledge base.

Overall, a process mining internship can provide a strong foundation for a career in data analysis, business process management, and related fields. It offers a unique chance to apply theoretical knowledge in a real-world context and develop practical skills that are highly sought after in today's data-driven business landscape.

## CONCLUSION

During my process mining internship, I gained invaluable insights into the world of data-driven process analysis and optimization. Throughout this experience, I was able to apply theoretical knowledge to real-world scenarios, collaborating with professionals and utilizing cutting-edge tools. One key takeaway from this internship is the significance of data quality and proper preprocessing in ensuring accurate process models. I learned to handle messy and incomplete data, employing techniques to clean and transform it into meaningful insights.

Moreover, the internship highlighted the importance of cross-functional communication. I interacted with various teams and stakeholders to understand their perspectives, which proved essential in creating comprehensive process models that cater to the needs of different departments. Working with process mining tools showcased their power in revealing bottlenecks, inefficiencies, and opportunities for improvement within processes. By analyzing event logs, I was able to identify areas for optimization, ultimately leading to enhanced process efficiency and cost savings.

The experience not only bolstered my technical skills but also improved my critical thinking and problem-solving abilities. I am excited to carry these lessons forward as I continue to explore the realms of data-driven process optimization in my future endeavors.

## REFERENCES

- [1] <https://in.indeed.com/career-advice/career-development/process-mining>
- [2] [https://www.softwareag.com/en\\_corporate/resources/process-mining/article/process-mining.html](https://www.softwareag.com/en_corporate/resources/process-mining/article/process-mining.html)
- [3] <https://www.g2.com/articles/process-mining>
- [4] <https://www.zoho.com/inventory/guides/basic-guide-to-order-management.html>