An Internship Report on

Process Mining Virtual Internship

Submitted in partial fulfilment of the requirements

for the award of the degree of

BACHELOR OF TECHNOLOGY

in

Computer Science and Engineering (Data Science)

by

C.RUPA SREE (214G1A3288)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

(DATA SCIENCE)

SRINIVASA RAMANUJAN INSTITUTE OF TECHNOLOGY
(AUTONOMOUS)
(Affiliated to JNTUA, accredited by NAAC with 'A' Grade, Approved by AICTE, New Delhi & Accredited by NBA (EEE, ECE & CSE))
Rotarypuram village, B K Samudram Mandal, Ananthapuramu-515701.

2023 - 2024

SRINIVASA RAMANUJAN INSTITUTE OF TECHNOLOGY

(AUTONOMOUS)

(Affiliated to JNTUA, accredited by NAAC with 'A' Grade, Approved by AICTE, New Delhi & Accredited by NBA (EEE, ECE & CSE)) Rotarypuram village, B K Samudram Mandal, Ananthapuramu-515701.

Department of Computer Science & Engineering (Data Science)

Certificate

This is to certify that the internship report entitled Process Mining Virtual Internship is the bonafide work carried out by **C.RUPA SREE** bearing Roll Number 214G1A3288 in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering (Data Science)** for four months from May 2023 to July 2023

Internship Coordinator

Mr. P. Veera Prakash, M. Tech., (Ph.D.), Professor & HOD of CSE

Head of the Department

Dr. P. Chitralingappa, M. Tech., Ph.D., Assistant Associate Professor

Date:

Place: Ananthapuramu

EXTERNAL EXAMINER

PREFACE

Brief overview of the company's history:Process mining has a rich history that dates back to the early 2000s. It emerged as a field of study at the intersection of data mining and business process management. Researchers and practitioners recognized the value of analyzing event data to gain insights into business processes. Over the years, process mining techniques and tools have evolved, leading to advancements in process discovery, conformance checking, and performance analysis. Today, process mining is widely used in various industries to improve operational efficiency, identify bottlenecks, and optimize processes.

- Who founded it

  Professor Wil van der Aalst founded

- What purpose and when

The purpose of process mining is to gain insights into business processes by analyzing event data. It helps uncover the actual process flows, identify bottlenecks, measure performance, and discover areas for improvement. Process mining is used to optimize processes and make data-driven decisions. It is typically applied when organizations want to understand their processes, improve efficiency, and enhance overall operational performance.

Company's Mission Statement:

The general mission of process mining companies is to provide innovative solutions and tools that enable organizations to gain valuable insights into their business processes, optimize operations, and achieve operational excellence.

Business Activities:

Process mining involves the analysis of event logs to understand and improve business processes. Its business activities include process discovery, conformance checking, and process enhancement. It helps identifies inefficiencies, compliance issues, and areas for optimization in various industries like manufacturing, healthcare, and finance.

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of people who made it possible, whose constant guidance and encouragement crowned our efforts with success. It is a pleasant aspect that I have now the opportunity to express my gratitude for all of them.

It is with immense pleasure that I would like to express my indebted gratitude to my internship coordinator **Mr. P. Veera Prakash, Assistant Professor & HOD, Department of Computer Science and Engineering**, who has supported me a lot and encouraged me in every step of the internship work. I thank him for the stimulating support, constant encouragement and constructive criticism which have made possible to bring out this internship work.

I am very much thankful to **Dr. P. Chitralingappa, Associate Professor & HOD, Computer Science and Engineering (Data Science),** for his kind support and for providing necessary facilities to carry out the work.

I wish to convey my special thanks to **Dr. G. Balakrishna, Principal** of **Srinivasa Ramanujan Institute of Technology** for giving the required information in doing my internship. Not to forget, I thank all other faculty and non-teaching staff, and my friends who had directly or indirectly helped and supported me in completing my internship in time.

I also express our sincere thanks to the Management for providing excellent facilities and support**.**

Finally, I wish to convey my gratitude to my family who fostered all the requirements and facilities that I need.

C.RUPA SREE

(214G1A3288)

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CRM	Customer Relationship Management
DDL	Data Definition Language
DML	Data Manipulation Language
EHR	Electronic Health Records
EMS	Event Management Systems
ERP	Enterprise Resource Planning
KPI	Key Performance Indicators
PQL	Procedure Query Language
SCM	Supply Chain Management
SQL	Structured Query Language

CHAPTER – 1

INTRODUCTION

Process mining is a set of techniques for the analysis of operational processes based on event logs extracted from company's databases, information systems, or business management software such as enterprise resource planning (ERP), customer relationship management (CRM), electronic health records (EHR), etc. In simple words, it's about finding out how the processes are actually performed to discover problems and areas for improvement.

Process mining uses a data-driven strategy for process optimization, enabling managers to make decisions about the resource allocation for current processes with objectivity. It places a strong emphasis on several viewpoints, including control flow, organization, case, and time.While much of the work on process mining concentrates on the order of events or control flow, the other viewpoints also offer management teams useful information. Time perspectives may show bottlenecks by monitoring the processing times of various events within a process, while organizational perspectives can highlight the various resources inside a process, such as specific job positions or departments.

Process mining applies data science to discover, validate and improve workflows. By combining data mining and process analytics, organizations can mine log data from their information systems to understand the performance of their processes, revealing bottlenecks and other areas of improvement.

Process mining is a technique used to analyze and improve business processes. It involves extracting data from event logs and visualizing the actual process flow, identifying bottlenecks, and suggesting optimizations. It's a powerful tool for process improvement and efficiency. Let me know if you have any specific questions.

It is a technique in the field of process management that supports the analysis of business processes based on event logs and drives improved efficiency, effectiveness, and compliance through its insight.

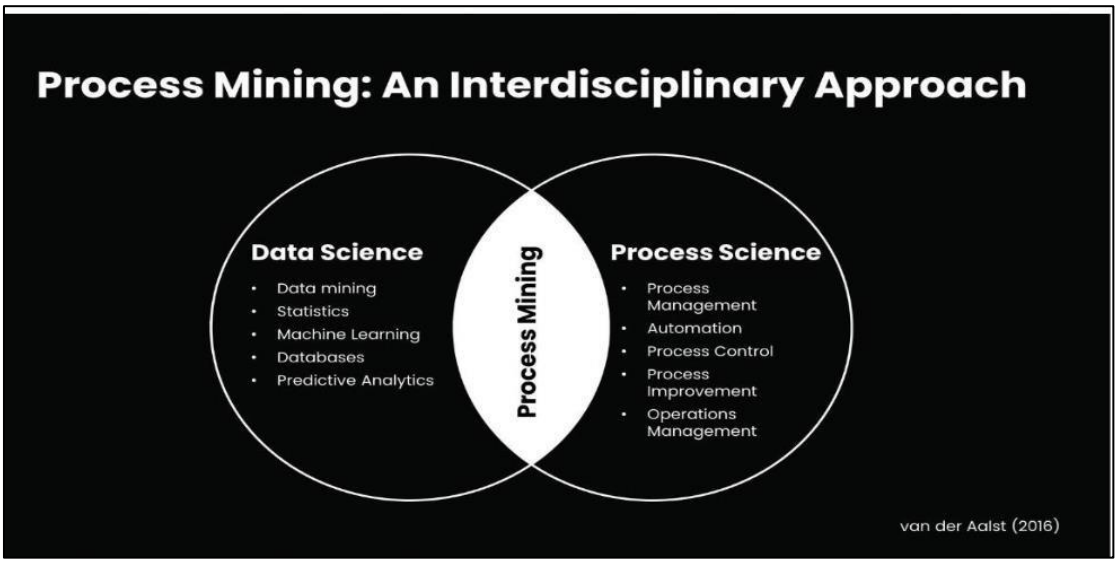

Fig. No. 1.1: Process Mining

1.2 Digital FootPrints

A digital footprint – sometimes called a digital shadow or an electronic footprint – refers to the trail of data you leave when using the internet.

Process Mining achieves this union by taking the digital footprints that are created in ITsystems and using them to reconstruct and visualize process flows. From here, Process Mining technology can identify patterns and deviations and ultimately eliminate bottlenecks. A digital footprint grows in many ways – for example, posting on social media, subscribing to a newsletter,leaving an online review, or shopping online.

Sometimes, it's not always obvious that you are contributing to your digital footprint. For example, websites can track your activity by installing cookies on your device, and apps can collate your data without you knowing it. Once you allow an organization to access your information, they could sell or share your data with third parties. Worse still, your personal information could be compromised as part of a data breach.

g

Fig. No. 1.2: Digital FootPrint

1.3 Event Logs

Event Logs are the format in which we can retrieve our digital footprints from the underlying IT systems. They're essentially the log books that IT systems keep to record what events take place for each Case ID and at what time.

The Event Log information can be retrieved from several types of IT systems such as Enterprise Resource Planning (**ERP**), Supply Chain Management (**SCM**) or Customer Relationship Management (**CRM**) systems. These systems typically generate and store Event Log information in real time. Event Log information might also be retrieved in various situations and contexts from automated payment to customer journeys.

An Event Log contains each of the three key pieces of information that our digital footprint have:

- **Case ID**: a unique identifier such as a purchase order item, invoice number or order number.
- **An Activity**: the description of what has happened - for example, the creation of a purchase order or the receipt of goods.
- **Timestamp**: the date and time that the activity took place.

order number	activity	timestamp	user	product	quantity
9901	register order	22-1-2014@09.15	Sara Jones	iPhone5S	1
9902	register order	22-1-2014@09.18	Sara Jones	iPhone5S	2
9903	register order	22-1-2014@09.27	Sara Jones	iPhone4S	1
9901	check stock	22-1-2014@09.49	Pete Scott	iPhone5S	1
9901	ship order	22-1-2014@10.11	Sue Fox	iPhone5S	1
9903	check stock	22-1-2014@10.34	Pete Scott	iPhone4S	1
9901	handle payment	22-1-2014@10.41	Carol Hope	iPhone5S	1
9902	check stock	22-1-2014@10.57	Pete Scott	iPhone5S	2
9902	cancel order	22-1-2014@11.08	Carol Hope	iPhone5S	2
...	...	...	...	...	...

Fig. No. 1.3: Sample Event Log

1.4 Techniques of Process Mining

There are three key steps they are Discovery,Conference,Enhancement.

Discovery: Describes the direct insights into the as-is process. In contrast to traditional Process Management techniques (for example, human-driven modeling or sampling) Process Mining provides objective, data-based insights into the as-is process in its full complexity.

Conformance: Refers to the comparison of the as-is process to an ideal model. Conformance can tell you how far and in how many cases the as-is process is deviating from the desired model. Furthermore, Discovery can help to reconstruct a model that is realistic and captures the complexity behind the process.

Enhancement: Refers to all actions and measures of improvements taken to bring the as-is process closer to the desired model. It encompasses any feedback from the real world including predictions and recommendations. As a research field, Enhancement opens up a completely new world for the adaptation and use of Process Mining within organizations and how it helps to create actual business value as well as insights.

CHAPTER - 2

TECHNOLOGY

Technology plays a crucial role in process mining, which is a discipline within the field of business process management that focuses on discovering, monitoring, and improving business processes using data from various information systems. Process mining leverages technology to extract valuable insights from event logs and other data sources to understand how processes are actually executed within an organization. Here's how technology is involved in the process mining process:

1. Data Extraction and Collection: Technology is used to extract relevant data from various sources such as enterprise systems, databases, transaction logs, and more. This data is then collected and consolidated for analysis.

2. Event Log Generation: Event logs are generated by recording the activities and events that occur within a business process. These logs capture information such as timestamps, activities performed, resources involved, and the sequence of events. Technology is employed to capture and store this data accurately.

3. Data Preprocessing: Raw event data often needs to be preprocessed to clean and transform it into a suitable format for analysis. Technology tools are used to remove duplicates, handle missing data, and standardize data formats.

4.Processing Discovery: Process discovery involves using technology to analyze event logs and automatically generate process models, such as process flowcharts or process maps. These models visualize the sequence of activities and their relationships, providing insights into the actual process flows.

5.Conformance Checking: Technology helps compare the discovered process model with the actual event data to identify deviations or non-conformance. This allows organizations to understand how closely their real-world processes align with intended processes.

6.Performance Analysis: Through process mining technology, organizations can assess process performance by analyzing metrics like cycle times, bottlenecks, and resource utilization. This analysis helps identify areas for optimization.

7.Process Enhancement: Technology helps compare the discovered process model with the actual event data to identify deviations or non-conformance. This allows organizations to understand how closely their real-world processes align with intended processes.

.

8.Process Mining: Process mining technology enables real-time monitoring of ongoing processes. By continuously analyzing event data, organizations can identify issues, predict potential bottlenecks, and make informed decisions to ensure smooth process execution.e

9.Data Visualization: Technology tools are used to create visual representations of process data and analysis results. Interactive dashboards and visualizations help stakeholders understand complex insights and make data-driven decisions.

10.Integration with other Systems: Process mining tools often need to integrate with other enterprise systems and databases to access relevant data. Technology facilitates seamless integration to ensure accurate and up-to-date information is used for analysis.

11.Machine Learning And AI: Advanced process mining solutions may incorporate machine learning and AI techniques to improve the accuracy of process discovery, predict process behavior, and provide more sophisticated insights.

12.Automation of Analysis: As the field advances, there's a trend toward automating more aspects of process mining, allowing organizations to continuously analyze and optimize processes without extensive manual intervention.

Overall, technology is the foundation of process mining, enabling organizations to gain actionable insights from their operational data and drive continuous process improvement.

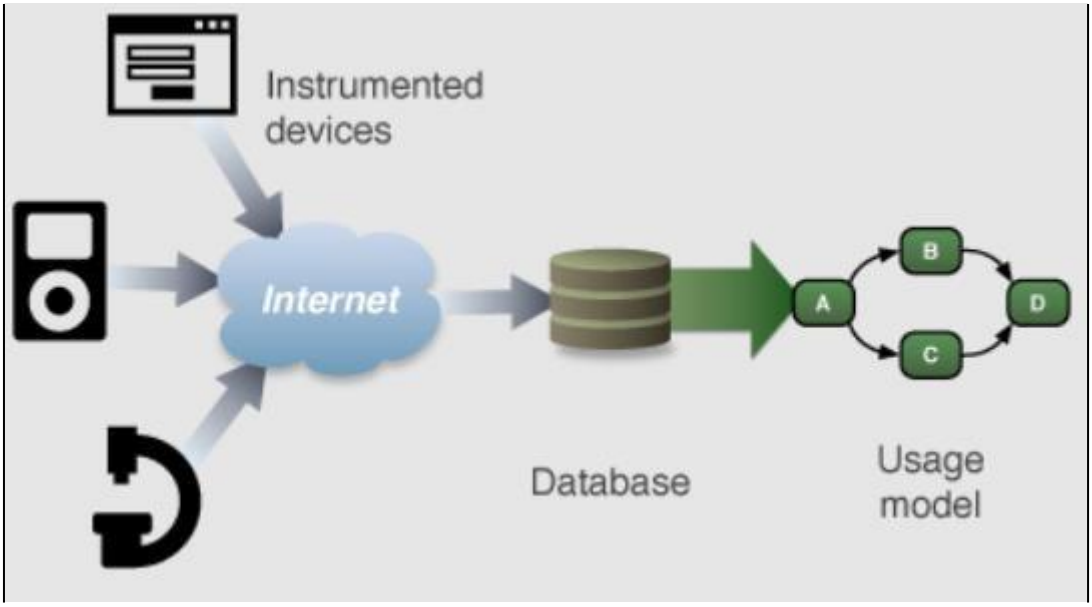

Fig. No. 2.1: Technology of Process Mining

CHAPTER - 3

APPLICATIONS

Process mining is a data-driven technique that involves analyzing event logs to extract insights and discover patterns from business processes. It has a wide range of applications across various industries. Some common applications of process mining include:

1.Process Discovery: Process mining can be used to automatically generate process models based on the event logs from various IT systems. This helps organizations understand how processes are actually executed and identify deviations from expected behavior.

2.Process Conformance Checking: Organizations can compare the actual execution of processes with the intended process models to identify discrepancies and noncompliant activities. This helps in improving process adherence and compliance.

3.Process Analysis: Process mining can provide insights into process performance metrics such as cycle times, bottlenecks, and resource utilization. This information helps organizations optimize their processes for efficiency.

4.Root Cause Analysis: When issues or inefficiencies arise in a process, process mining can be used to analyze event logs and identify the root causes of these problems. This enables organizations to address underlying issues and improve process effectiveness.

5.Process Enhancement and Redesign: By analyzing event logs, organizations can identify areas where processes can be improved or streamlined. This information informs process redesign efforts aimed at achieving better outcomes.

6.Resource Allocation: Process mining can help organizations allocate resources more effectively by understanding how resources are used throughout a process. This ensures optimal resource allocation and reduces unnecessary costs.

7.Risk Management: Process mining can be used to detect and mitigate risks by identifying deviations from expected process behavior that could lead to errors, delays, or compliance violations

8.Customer Journey Analysis: Organizations can use process mining to understand the customer journey and identify pain points or areas for improvement in customer interactions and experiences.

9.Supply Chain Analysis: In supply chain management, process mining can provide insights into the movement of goods, identify delays, and improve overall supply chain efficiency.

10.Healthcare Process Optimization: Process mining can be applied to healthcare to analyze patient pathways, identify bottlenecks in treatment processes, and optimize resource utilization in hospitals.

11.IT Service Management: IT service processes can be analyzed using process mining to identify areas of improvement, reduce response times, and enhance service delivery.

12.Audit and Compliance: Process mining can assist auditors in understanding how processes are executed and verifying compliance with regulations and policies.

Fig. No. 3.1: Applications of Process Mining

14.Fraud Detection: By analyzing patterns of behavior in event logs, process mining can help identify anomalous activities that may indicate fraudulent behaviour**.**

15.Environment Impact Assessment: Process mining can be used to analyze resource consumption and waste generation in industrial processes, aiding in environmental impact assessment and sustainability efforts.

16.Manufacturing Process Optimization: Process mining can optimize manufacturing processes by identifying production bottlenecks, reducing downtime, and improving resource allocation.

These are just a few examples of the many applications of process mining. The technique's ability to extract insights from data about how processes are executed makes it a valuable tool for process improvement and optimization across a wide range of industries and domains.

CHAPTER - 4

MODULES

4.1 Fundamentals of Process mining

Modules in process mining refer to distinct areas or topics of study within the field. These modules help break down the comprehensive study of process mining into manageable segments, allowing learners to delve deeper into specific aspects.

What is an Analysis?

Once you have accessed the analysis, you may see anywhere from one to several sheets in it. The person building the analyses creates each analysis with the specific user(s) needs in mind.

Analysis in process mining involves examining event data to gain insights into how processes are actually executed, identifying bottlenecks, inefficiencies, compliance issues, and opportunities for improvement.

1.Data Preparation: Before analysis can begin, event data needs to be collected, cleaned, and prepared. This involves extracting relevant information from various sources, ensuring data quality, and structuring the data into event logs.

2.Process Discovery: Process discovery is the initial step where a visual representation of the process is created based on the event logs. Various algorithms are used to generate process models, such as process maps, flowcharts, Petri nets, or BPMN diagrams. This helps understand the actual flow of activities, decisions, and paths taken by cases.

3.Conformance Checking: Conformance checking involves comparing the discovered process model with the actual event data. Deviations between the model and the real data are highlighted. This analysis can help identify instances of non-compliance, inefficiencies, and variations from the expected process flow.

4.Performance Analysis: Performance analysis involves assessing process performance metrics using the event data. Key Performance Indicators (KPIs) like cycle times, waiting times, resource utilization, and throughput rates are calculated. This provides insights into process efficiency and potential areas for optimization.

5.Root Cause Analysis: Process mining can be used to identify the root causes of process issues. By analyzing the event data, you can trace back the factors contributing to bottlenecks, delays, or deviations. This information is valuable for making targeted improvements.

6.Variant Analysis: Process mining can reveal different process variants that occur within a single process. This is particularly useful in scenarios where processes have

variations based on context, exceptions, or different paths taken by cases. Analyzing variants helps understand the complexity of the process and identify opportunities for standardization.

7.Resource Analysis: Resource analysis focuses on how resources (people, equipment, systems) are utilized within the process. This helps identify resource constraints, overutilization, and potential areas for resource optimization.

8.Predictive analysis: Some advanced process mining tools incorporate predictive analysis using machine learning techniques. This enables organizations to forecast process behavior, predict potential bottlenecks, and suggest optimal paths for cases.

9.Continuous Monitoring: Process mining is an ongoing practice. Continuous monitoring of processes using event data helps identify trends, shifts, and emerging issues in real time. This enables organizations to maintain process efficiency and responsiveness.

10.Decision Making and Improvement: The insights gained from process mining analysis directly impact decision-making and process improvement initiatives. Organizations can prioritize changes, allocate resources effectively, and implement datadriven improvements.

Finally, analysis in process mining transforms raw event data into actionable insights that guide process optimization and innovation. It plays a crucial role in bridging the gap between designed processes and their real-world execution, leading to more efficient, effective, and adaptable operations.

4.2 Variant Explorer

The Variant Explorer is a tool within Celonis that allows users to analyze and visualize different variants or paths that process instances take through a process flow. In many business processes, there are multiple ways a process can be executed based on different conditions, exceptions, or decisions made during the process.

what is an Variant Explorer ?

Variant Explorer is a feature or tool within process mining software that allows users to explore and visualize different paths, sequences, and variations within a process. It provides insights into the diverse ways that instances of a process can unfold based on real-world event log data.

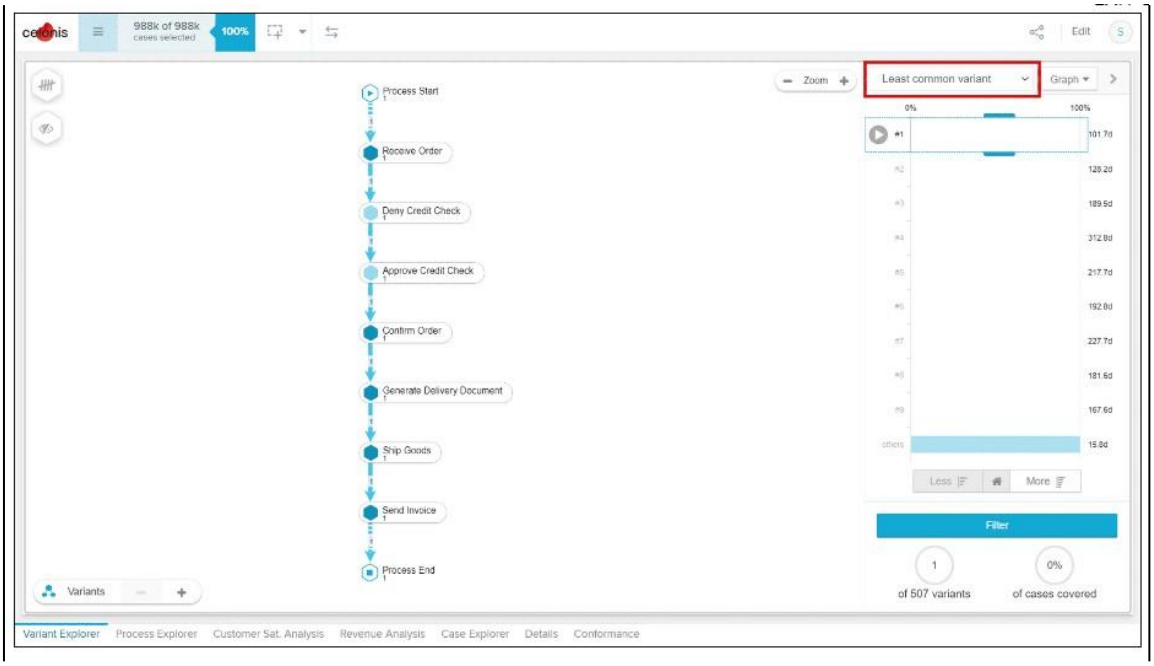

Fig. N0. 4.2: Variant Explorer

4.3 KPI(Key Performance Indicates)

KPI stands for Key Performance Indicator. KPIs are specific metrics or measures used to evaluate the performance of a business, organization, process, project, or individual. They provide a quantifiable way to assess whether goals and objectives are being met effectively. KPIs are widely used in various industries to track progress, make informed decisions, and identify areas for improvement. Here are some key points about KPIs:

1.Purpose: KPIs are used to measure and assess the performance and success of various aspects of a business or process. They help answer the question: "Are we achieving our desired outcomes?"

2.Quantifiable: KPIs are typically expressed as numerical values or ratios, making them measurable and comparable. They provide a clear and objective way to evaluate performance.

3.Relevance: KPIs should be directly tied to the goals and objectives of the entity being measured. They should align with the organization's overall strategy and priorities.

4.Monitoring: KPIs are regularly monitored over time to track trends, patterns, and changes in performance. This monitoring helps in identifying deviations and making timely adjustments.

5.Actionable: Effective KPIs drive action. When a KPI indicates a problem or an opportunity for improvement, organizations can take specific actions to address the situation.

6.Different Levels: KPIs can be applied at various levels, such as organizational, departmental, project, or individual levels. Each level might have different KPIs relevant to its goals.

7.Examples: Some common examples of KPIs include financial metrics like revenue growth and profit margins, operational metrics like customer satisfaction and on-time delivery, and process metrics like cycle time and defect rate.

8.Small Criteria: To be effective, KPIs often follow the SMART criteria: Specific, Measurable, Achievable, Relevant, and Time-bound. This ensures that KPIs are welldefined and practical.
Finally,KPIs provide a structured way to assess performance and guide decision-making based on data-driven insights. They are an essential tool for measuring progress, driving improvement, and ensuring that efforts are aligned with strategic goals.

In the context of process mining and process analysis, "Case Frequency" and "Activity Frequency" are key performance indicators (KPIs) that provide insights into how often specific cases (process instances) and activities occur within a process. These KPIs help in understanding process behavior, identifying patterns, and highlighting potential areas for improvement. Here's what each of these KPIs means:

Case Frequency:

Case Frequency, also known as Case Count or Case Instances, refers to the number of times a specific process case or instance occurs in the given dataset. In other words, it's a count of how many times a particular sequence of activities is executed. Case Frequency helps you understand the prevalence of different scenarios or scenarios that repeat frequently. For example, in a customer order processing process, if you have 100 customer orders and 10 of them experience delays, the Case Frequency for the "delayed orders" variant would be 10.

Activity Frequency:

Activity Frequency, also known as Activity Count, represents the number of times a specific activity is performed within the process. This KPI provides insight into the usage and frequency of each individual activity in the process flow. For instance, in an order fulfillment process, if the "Check Inventory" activity occurs 500 times out of 1000 total cases, the Activity Frequency for the "Check Inventory" activity would be 500

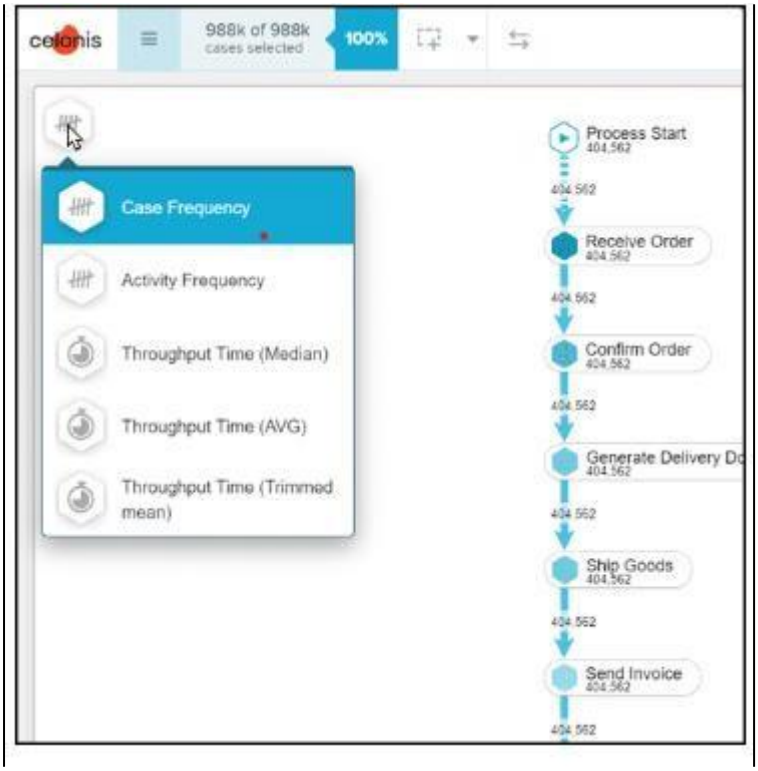
celonis
988k of 988k
cases selected
100%
Case Frequency
Activity Frequency
Throughput Time (Median)
Throughput Time (AVG)
Throughput Time (Trimmed mean)
Process Start
404,562
Receive Order
Confirm Order
Ship Goods
Send Invoice

Fig. No. 4.3: KPI in Variant Explorer

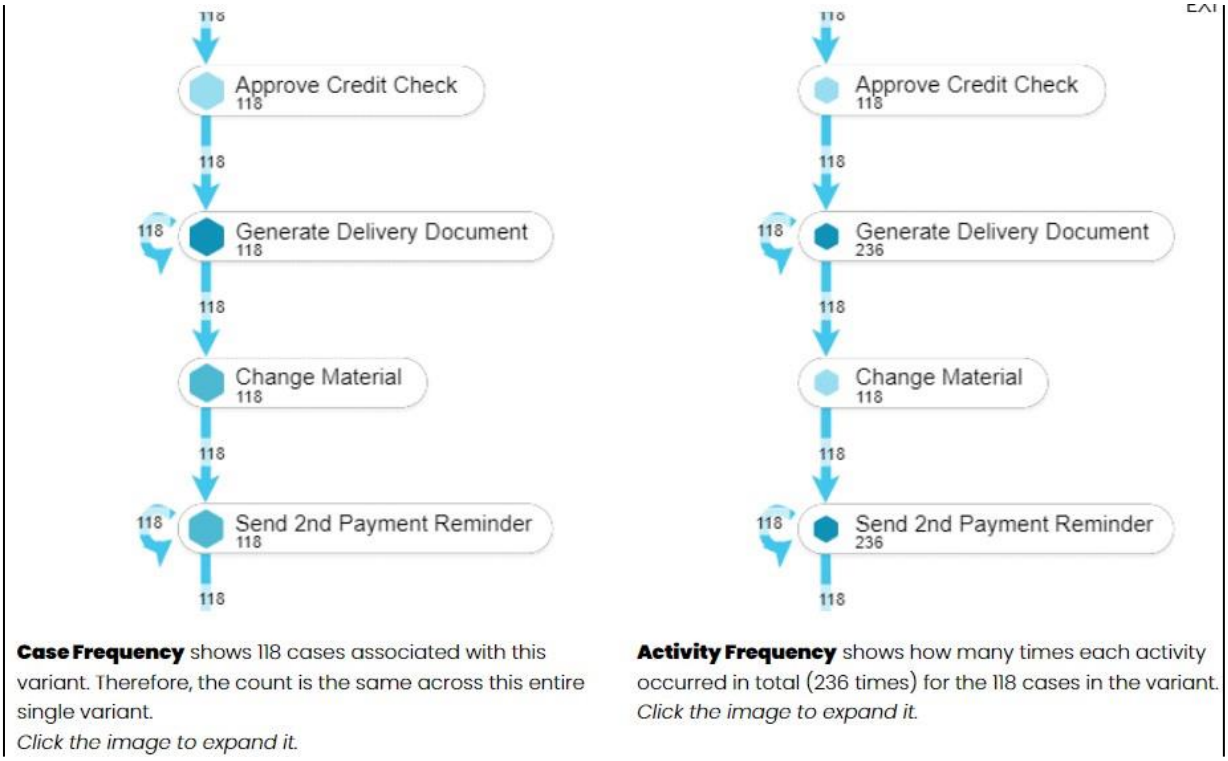

Fig. No. 4.3 Difference Between Case Frequency and Activity Frequency

Above is a side-by-side comparison of the same variant with case Frequency and Activity Frequency KPIs. Notice the difference in the count for the "Generate Delivery Document" activity.

4.4 Compare Multiple Variants

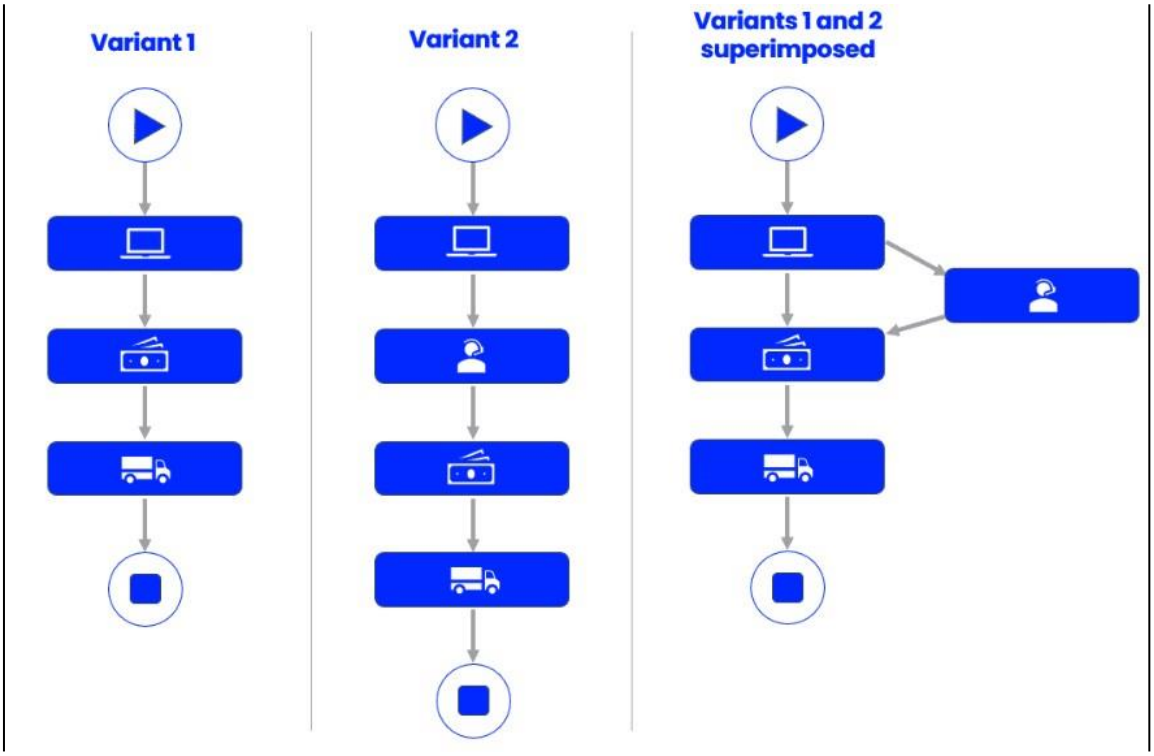

Fig. No. 4.4: Compare Multiple Variants

Comparing multiple variants in process mining involves analyzing and contrasting the different paths or scenarios that process instances take through a process. This comparison helps to identify patterns, bottlenecks, and areas for improvement.

4.5 Process Explorer

The process Explorer is another analysis tool to use when taking an exploratory approach. It's especially useful for quickly revealing activities beyond the most common ones. It also allows you to narrow your focus on a single activity, for example an undesired activity, to see which activities cases typically come from and which activities they're going to.

Celonis Process Explorer is a component within the Celonis platform that allows users to:

1.Visualize Process: provides visual representations of processes based on event data. These visualizations include process maps, flowcharts, and diagrams that illustrate the flow of activities and paths taken by process instances.

2.Explore Variants: Users can explore different variants or paths that process instances follow. This helps identify common patterns, exceptions, and variations in process execution.

3.Analyze Activity: The tool enables users to analyze individual activities within the process. This includes insights into the frequency, timing, and relationships between activities.

4.Performance Metrics: Celonis Process Explorer offers performance metrics such as cycle times, waiting times, and other key indicators. These metrics provide insights into process efficiency and effectiveness.

5.Root cause Analysis: Users can trace back the sequence of events to identify root causes of process issues, bottlenecks, and deviations from the expected flow.

6.Continuous Monitoring: Celonis Process Explorer might support real-time or continuous monitoring of processes. This allows users to track changes and deviations as they occur.

7.Decision-Making and Improvement: By utilizing Celonis Process Explorer, organizations can make informed decisions for process optimization and improvement. The insights gained help prioritize changes and drive improvements.

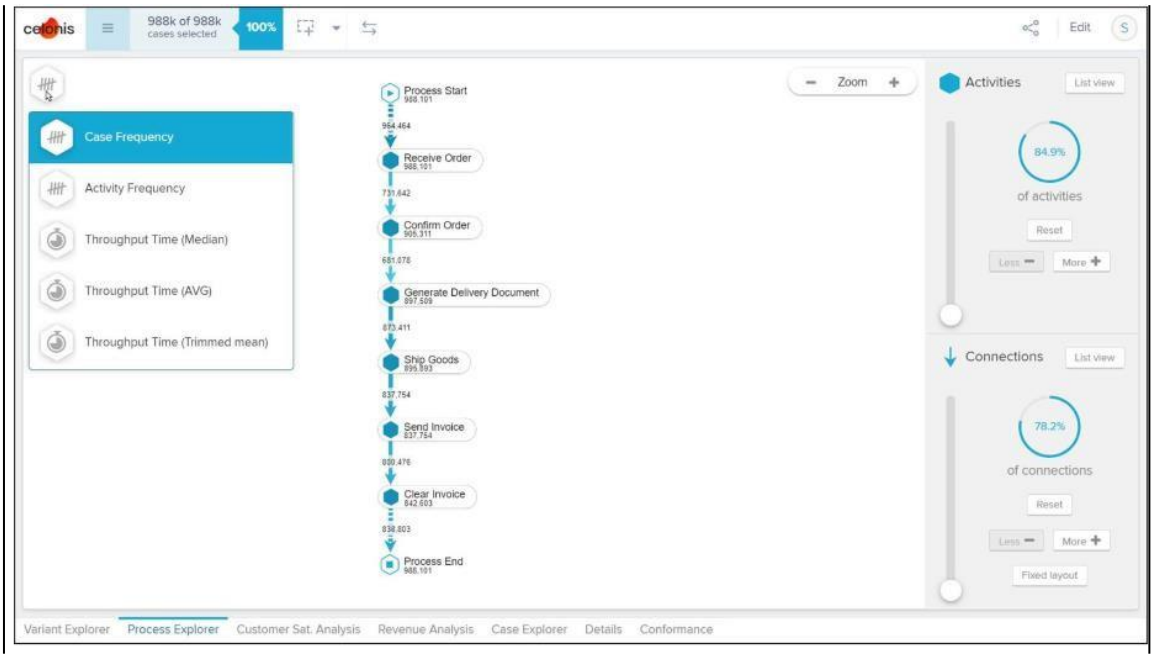

Fig. No. 4.5: Process Explorer

4.6 MODULE 2:Rising star Technical

"Rising star" is a term often used to describe individuals who are on a trajectory of rapid growth, development, and success in their field or career. In the context of "Rising star

Technical in process mining," it refers to someone who is quickly emerging as a skilled and influential professional in the technical aspects of process mining

It mainly consists of Two pats:

1.Write PQL Queries

2.Get Data into the EMS

Write PQL Queries:

Writing Process Query Language (PQL) queries in process mining involves creating specific queries to extract insights and information from event log data. PQL is a query language designed to interact with event logs and process models, enabling users to retrieve relevant process information for analysis. Here's an explanation of how to write PQL queries in process mining:

In the course of digitization, an increasing number of log data is recorded in IT systems of companies worldwide. This data is precious, as it represents how business processes are running inside a company. Process Mining comprises data-driven methods to discover, enhance and monitor processes based on such data. The heart of Process Mining are the Event Logs. Those Event Logs are a collection of process events that can be described by the following attributes:

1.Case: The case attribute indicates which process instance the event belongs to. A process instance is called a case, usually consisting of multiple events.

Let's consider an example: Imagine you are running a restaurant with food delivery. Each order has a specific number, the order number. This number is the unique case ID, and all related activities are assigned to this ID.

2.Activity: The activity attribute describes the action that is captured by the event.

In our food delivery example, these are all the steps an order has passed through, from receiving the order, to cooking the meal, delivery and payment.

3.Timestamp: Each activity leaves a digital footprint with a timestamp, indicating precisely when each event took place

With the help of timestamps, we know precisely in which chronological order the different activities have runoff.

4.7 Executable Queries in Process Mining

To gain valuable process insights, it is essential for Process Mining users to formalize their process questions as executable queries. For this purpose, we present the Celonis Process Query Language (Celonis PQL), which is:

1.a domain-specific language
2.tailored towards a particular process data model and
3.designed for business users.

It translates process-related business questions into queries and executes them on a custom-built query engine, the Celonis PQL Engine.

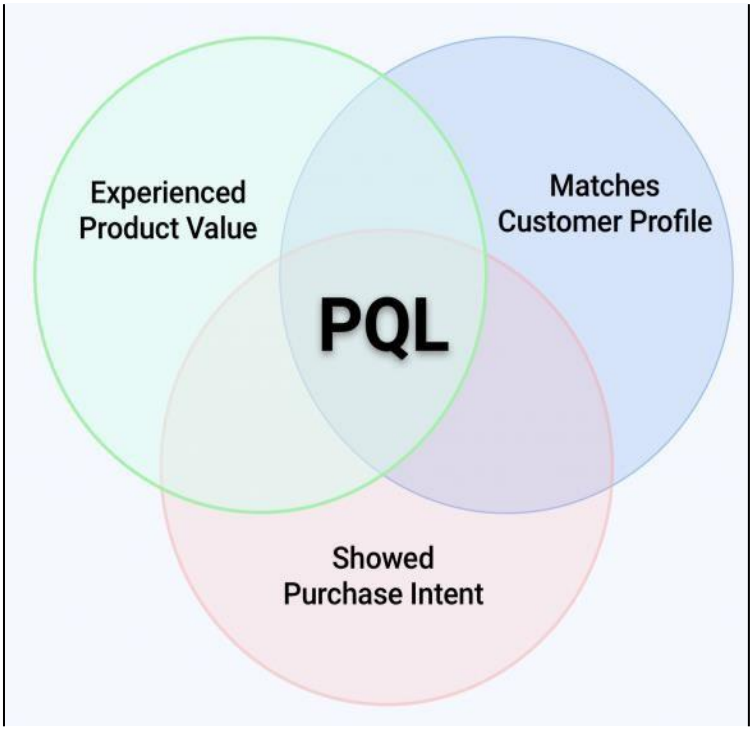

Fig. No. 4.7: PQL Queries in Process Mining

4.8 SQL VS PQL

Even though Celonis PQL is inspired by SQL, there are major differences between the two query languages.

On a high level, Celonis PQL varies along four key dimensions:

1.Language Scope:elonis PQL does not support all operators that are available in SQL.

This is because customer requirements drive the development of the language, and only operators needed for the target use cases are implemented

2.Data Manipulation Language: Second, Celonis PQL is not supported by a data manipulation language (DML).

As all updates in the Process Mining scenario should come from the source systems, there is no need to manipulate and update the data through the query language directly.

3.Data Definition Language: Furthermore, Celonis PQL does not provide any data definition language (DDL).

As the data model is created by a visual data model editor and stored internally, there has not been any need for creating and modifying database objects.

4.Domain-Specific: In contrast to SQL, Celonis PQL is domain-specific and offers a wide range of Process Mining operators not available in SQL.

Consequently, Celonis PQL seamlessly integrates the data with the process perspective.

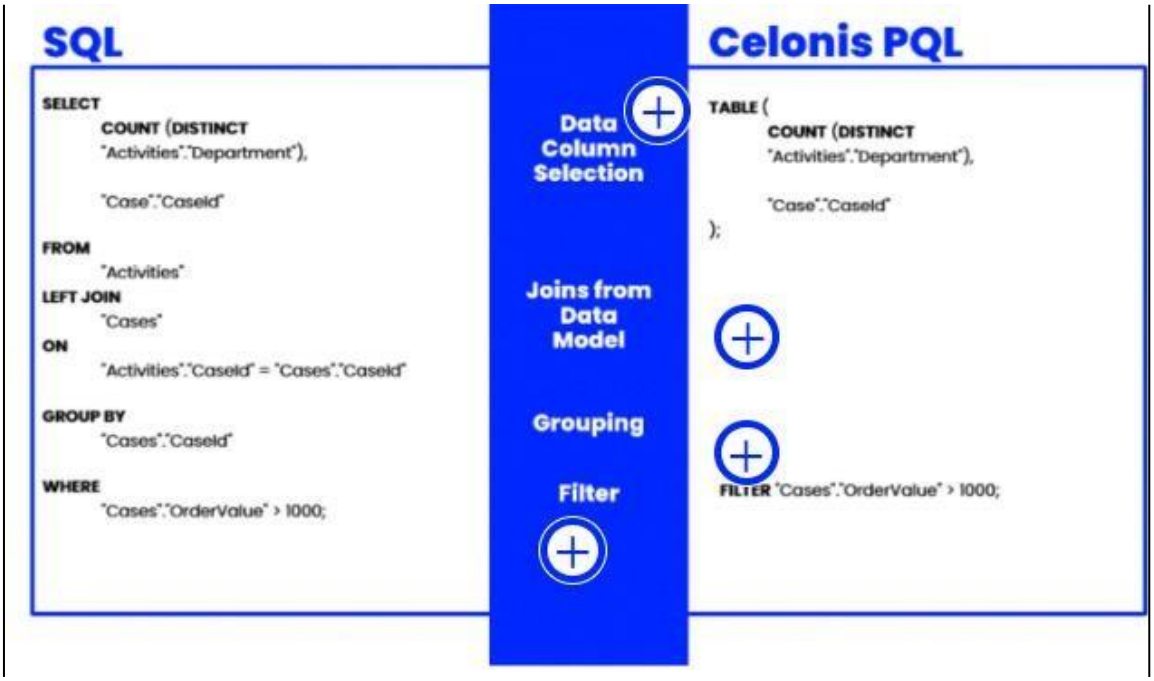

Fig. No. 4.8: Difference between SQL and Celonis PQL

4.9 Get Data Into the EMS

In process mining, "EMS" could refer to "Event Management System" or another specific context related to process mining tools or platforms. If you're referring to an Event Management System, the process of getting data into it for process mining analysis involves collecting and storing event log data from various sources. Here's an explanation of how data is typically collected and integrated into an Event Management System for process mining:

1.Event Data Source: Event log data is the foundation of process mining analysis. This data comes from various sources, such as:
Operating Systems: These are the systems where actual business processes are executed, such as CRM systems, ERP systems, workflow systems, etc.
Sensor Data: For manufacturing or industrial processes, data from sensors and IoT devices can be important sources.
IT Services: Log files from IT systems can also contribute to understanding process behavior.

2.Data Extraction: Involves retrieving relevant event data from the sources. This could be done using various methods, including APIs, direct database queries, log file parsing, or data connectors provided by process mining tools.

3.Data Transferring and Cleansing: The extracted data might be in different formats and structures. Data transformation involves converting the data into a consistent format suitable for process mining analysis. This step might also include cleansing the data to remove inconsistencies, duplicates, and irrelevant information.

4.Data Integration: Once the data is transformed and cleansed, it needs to be integrated into the Event Management System. This could involve loading the data into a dedicated database or storage system used by the process mining tool.

5.Mapping Event Attributes: Each event in the event log data contains attributes like timestamps, activity names, resource IDs, etc. These attributes need to be mapped to the corresponding attributes recognized by the Event Management System.

In summary, the process of getting data into an Event Management System for process mining involves extracting event data from various sources, transforming and cleansing it, integrating it into the system, and mapping event attributes. This data serves as the basis for process mining analysis, enabling organizations to gain insights into their business processes and drive optimization efforts.

CHAPTER - 5

REAL TIME EXAMPLES

Process mining involves analyzing event logs to gain insights into business processes. Here are some examples of how process mining can be applied across different industries and domains:

1.Supply Chain Management:

Order Fulfillment Process: Analyzing event logs from order processing, inventory management, and shipping systems to identify bottlenecks and optimize order delivery times.
Vendor Performance Analysis: Tracking interactions between a company and its suppliers to evaluate supplier performance, lead times, and order accuracy.

2.Healthcare:

Patient Journey Analysis: Studying patient data to understand the flow of patients through different departments, identifying delays, and improving resource allocation for better patient care.
Claim Processing: Analyzing insurance claims processing to identify inefficiencies, reduce processing times, and ensure accurate billing.

3.Finance:

Accounts Payable Process: Analyzing invoices, payment records, and approvals to streamline the accounts payable process, reduce payment delays, and identify potential fraud.
Loan Approval Process: Mapping out the steps involved in loan applications, approvals, and disbursements to optimize the loan approval process and reduce turnaround times.

4.Manufacturing:

Production Line Analysis: Analyzing data from production machines and quality control systems to identify production bottlenecks, improve product quality, and optimize resource allocation.
Equipment Maintenance: Tracking maintenance logs to predict equipment failure, schedule preventive maintenance, and minimize downtime.

5.Retail:

Customer Purchase Journey: Analyzing customer transactions, online behavior, and interactions with sales staff to optimize store layouts, improve customer experience, and enhance cross-selling opportunities.
Inventory Management: Studying inventory movement data to optimize stock levels, reduce carrying costs, and prevent stockouts.

6.Telecommunications:

Service Action Process: Analyzing the end-to-end process of activating new services for customers to identify delays, improve activation times, and enhance customer satisfaction.
Network Fault Analysis: Studying network logs to identify patterns of network failures, diagnose root causes, and optimize network maintenance.

7.Logistics:

Shipment Tracking: Analyzing data from tracking systems to visualize the movement of shipments, identify potential delays, and improve delivery accuracy.
Route Optimization: Studying vehicle movement data to optimize delivery routes, reduce fuel consumption, and improve overall fleet efficiency.

8.Human Resources:

Employee Onboarding: Analyzing the onboarding process to identify areas for improvement, reduce paperwork, and ensure a smoother transition for new employees.
Performance Review Process: Mapping out the steps in the performance review process to identify bottlenecks, streamline evaluations, and ensure timely feedback.

These examples showcase the versatility of process mining in various industries and processes. Process mining helps organizations gain valuable insights into their operations, streamline processes, and make data-driven decisions for continuous improvement

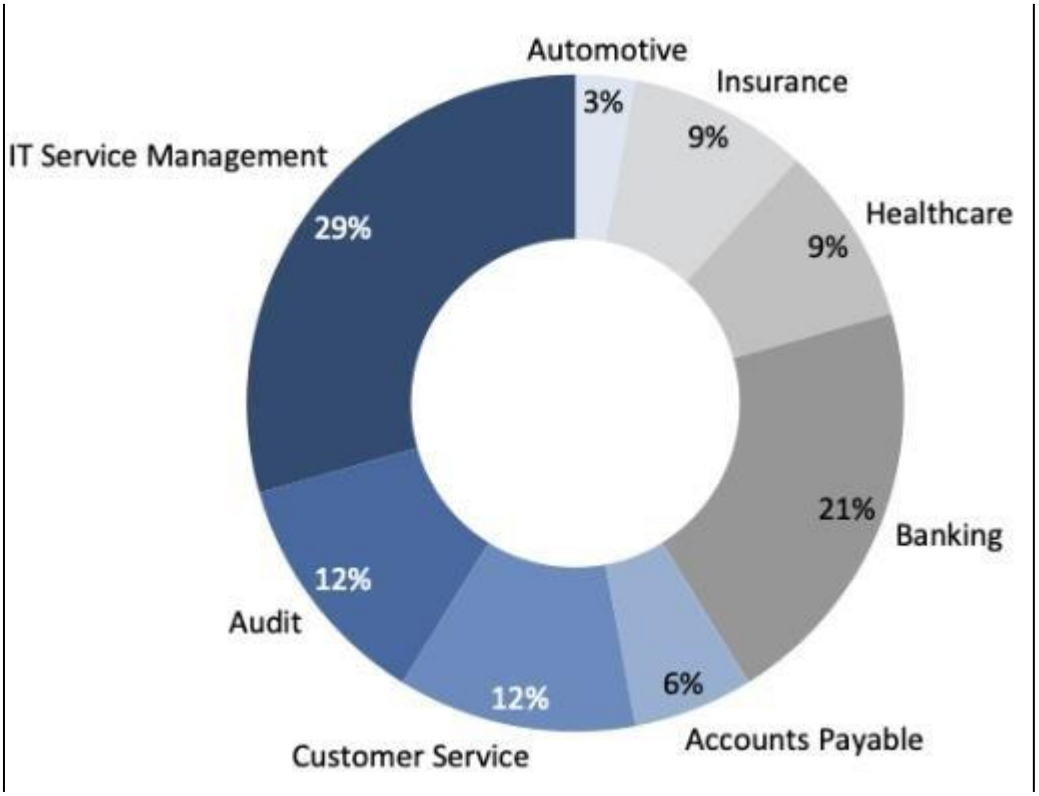

Fig. No. 5.1: Real Time Examples of Process Mining.

CHAPTER - 6

LEARNING OUTCOMES

Some of the outcomes on Process Mining include:

Improves Efficiency: By analyzing event logs and process data, process mining can identify bottlenecks, inefficiencies, and deviations in business processes. This information can help organizations optimize their processes and improve overall efficiency.

Process mining techniques can automatically discover and visualize the actual process flows based on event logs. This helps organizations gain a better understanding of their processes, identify process variants, and uncover hidden patterns or deviations.

Process Monitoring: Process mining allows organizations to monitor and measure process performance indicators, such as cycle times, throughput, and resource utilization. This helps in identifying performance issues, evaluating process improvements, and making data-driven decisions.

Conformance Checking: Process mining can compare discovered process models with predefined process models or business rules to assess compliance. It can identify deviations, non-compliance, and potential risks, enabling organizations to take corrective actions.

➢Understand what process mining is and the basics of how it works.

➢Understanding how process mining helps in Business world.

➢you will learn to extract to create visual representations of processes to aid decision making and process improvement efforts

➢Process mining can help ensure compliance with regulations and internal policies

by identifying deviations and non-compliances in process execution.

➢Analyzing process performance metrics and KPIs.

➢Summarize what an event log is and why we need it for processing.

➢You will gain skills in using process mining tools and interpreting the results to

enhance organizational efficiency and effectiveness.

➢Understanding how process mining helps to uncover inefficiencies and methods

for insufficient process.

➢Predicting process outcomes and future behaviour.

➢After completing this course you will learn about process query language.

➢Understanding process behaviour and it's applications in day life

CONCLUSION

Process mining enables organizations to gain comprehensive insights into their operational processes. By analyzing event logs, it captures the actual sequence of activities, decisions, and interactions that occur during process executions. This level of Stransparency goes beyond traditional process documentation, providing a deep understanding of how processes truly function.

The insights generated by process mining are rooted in real-world data, allowing for data-driven decision-making. Organizations can base their strategies, improvements, and resource allocation on empirical evidence rather than assumptions. This lends credibility to decisions, increases their chances of success, and supports more effective allocation of resources.

A key takeaway from process mining is the concept of continuous improvement. By constantly analyzing processes and acting on insights, organizations can create a culture of ongoing enhancement. Processes can be iteratively refined, leading to higher efficiency, shorter cycle times, and ultimately, better customer satisfaction.

Process mining highlights the importance of ethical considerations and data privacy. Working with sensitive process data demands responsible handling to safeguard privacy and comply with regulations. This ethical approach ensures that insights are generated in a manner that respects individual rights and organizational integrity.

As technology continues to advance, process mining techniques are likely to become even more sophisticated and integrated with other data-driven approaches, further enhancing their ability to drive process excellence. However, successful implementation of process mining requires a comprehensive understanding of both the technology and the underlying business processes. Organizations that embrace process mining stand to gain a competitive edge by harnessing the power of data-driven insights to continuously refine their operations and achieve higher levels of efficiency and effectiveness.

N·E·A·T
प्रौद्योगिकी के लिए राष्ट्रीय शैक्षणिक सहयोग
National Educational Alliance for Technology

अखिल भारतीय तकनीकी शिक्षा परिषद्
All India Council for Technical Education

EduSkills®
Nation Building Through Skills

Certificate of Virtual Internship

This is to certify that

Rupasree Chillu

Srinivasa Ramanujan Institute of Technology

has successfully completed 10 weeks

Process Mining Virtual Internship

during May - July 2023

Supported By celonis

Jerome Geyer-Klingeberg
Head of Academic Alliance
Celonis

Shri Buddha Chandrasekhar
Chief Coordinating Officer (CCO)
NEAT Cell, AICTE

Dr. Satya Ranjan Biswal
Chief Technology Officer (CTO)
EduSkills

Certificate ID :845feef8d7cdbf2e7a7e15c11fe17273
Student ID :STU62307fa818a9c1647345576

Andhra Pradesh State
Council of Higher Education

EduSkills®
Nation Building Through Skills

Virtual Internship Completion Certificate

This is to certify that

Rupasree Chillu

has successfully completed 10 weeks

Process Mining Virtual Internship

during May - July 2023

Supported By celonis

Prof. K. Hemachandra Reddy
Chairman, Andhra Pradesh State
Council of Higher Education

Jerome Geyer-Klingeberg
Head of Academic Alliance
Celonis

Dr. Satya Ranjan Biswal
Chief Technology Officer (CTO)
EduSkills

Certificate ID :414377896687defbeae37b2e8d2ec01c

REFERENCES

- The Reference of this internship was done in the Celonis platform link: https://academy.celonis.com/learn/dashboard
- Different academic disciplines and institutions have specific formatting guidelines for how references should be presented. These guidelines are known as citation styles. Some common citation styles include APA (American Psychological Association), MLA (Modern Language Association), Chicago, IEEE (Institute of Electrical and Electronics Engineers), and more. Each style has specific rules for how to format citations and references.
- A typical reference entry includes several components, which can vary depending on the type of source (book, journal article, website, etc.) and the citation style you're using. Common components include the author's name, title of the work, publication date, publisher, and source location (page numbers, URL, etc.).
- When you use information, ideas, data, or any content from a source (such as a research paper, book, website, etc.), you need to cite it within your own work. This citation points your readers to the full reference in your reference list.
- IEEE CIS Task Force on Process Mining .
- Burattin Andrea. Applicability of Process Mining techniques in Bussiness Environments.