

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

O. DEVANAND

(224G1A3214)



**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.

2024 - 2025

**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 **O. DEVANAND** bearing Roll Number **224G1A3214** in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering (Data Science)** for 10 weeks from April – June 2024.

Internship Coordinator

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

Head of the Department

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

Date:

Place: Ananthapuramu

EXTERNAL EXAMINER

PREFACE

Brief overview of the company's history:

- **Who founded it:** All India Council for Technical Education (AICTE) has initiated various activities for promoting industrial internship at the graduate level in technical institutes and Eduskills is a Non-profit organization which enables industry 4.0 ready digital workforce in India. The vision of the organization is to fill the gap between Academic and Industry by ensuring world class curriculum access to the faculties and students. Formation of the All-India Council for Technical Education (AICTE) in 19445 by the Government of India.
- **What purpose and when:** With a vision to create an industry-ready workforce who will eventually become leaders in emerging technologies, Eduskills & AICTE launches 'Virtual Internship' program on Process Mining. This field is one of the most in-demand, and this internship will serve as a primer

Company's Mission Statement: The main mission of these initiatives is enhancement of the employability skills of the students passing out from Technical Institutions.

Business Activities: The All-India Council for Technical Education (AICTE) primarily focuses on regulating and promoting technical education in India. Its business activities include accrediting institutions, approving new courses, setting quality standards, fostering research, providing policy recommendations, and ensuring the overall development of technical education across the country.

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.

O. DEVANAND
(224G1A3214)

INDEX

Contents		Page No.
List of Figures		vi
List of Abbreviations		vii
Chapter 1	Introduction to Process Mining	1-2
Chapter 2	Foundation of Process Mining	3-9
Chapter 3	Techniques in Process Mining and its types	10-13
Chapter 4	Modules Explanation	14-15
Chapter 5	Applications and Real-time Examples	16-19
Chapter 6	Learning outcomes of the Internship	20
Conclusion		21
Internship certificate		22
References		

LIST OF FIGURES

Fig. No.	Description	Page No.
1.1	Introduction to Process Mining	Error! Bookmark not defined.
2.1	Get to know celonis analyses	3
2.2	Navigate to an analysis	5
2.3	Use variant Explorer	6
2.4	Use Process Explorer	7
2.5	Selection views	7
2.6	Steps of Process mining	8
2.7	Stages of Process flow	8
3.1	Process mining Techniques	9
3.2	Stages of process mining	10
3.3	Types of Process mining	11

LIST OF ABBREVIATIONS

PM	Process Mining
RPA	Robotic process automation
CSV	comma-separated values
XML	Extensible Markup Language.
ERP	Enterprise Resource Planning
SRM	Supplier Relationship Management
MRI	Magnetic Resonance Imaging
NPS	Net Promoter Score
KPI	Key Performance Indicators

CHAPTER – 1 INTRODUCTION

World and the organizations in it are full of processes. From purchasing to order management, organizations deal with complex, global and sometimes faulty processes on a daily basis. **Frictionless processes**, on the other hand, ensure:

- that you can find the right groceries at the grocery store,
- that planes land on time,
- that patient waiting times at hospitals are kept to a minimum

1.1 Introduction to process Mining

Process Mining is often described as occupying the area between business process management and data mining. The speed, accuracy, and auditability these technologies deliver can result in significant cost savings and much faster time to market.

- 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 insights.
- Modern Process Mining technology quickly and reliably extracts information from event and transaction logs to visually depict real-time process models for current processes

Information systems, such as Enterprise Resource Planning (ERP) or Customer Relationship Management (CRM) tools, provide an audit trail of processes with their respective log data. Process mining utilizes this data from IT systems to create a process model, or process graph. From here, the end-to-end process is examined, and the details of it and any variations are outlined. Specialized algorithms can also provide insight into the root causes of deviations from the norm. These algorithms and visualizations enable management to see if their processes are functioning as intended, and if they aren't, they arm them with the information to justify and allocate the necessary resources to optimize them. They can also uncover opportunities to incorporate robotic process automation into processes, expediting any automation initiatives for a company.

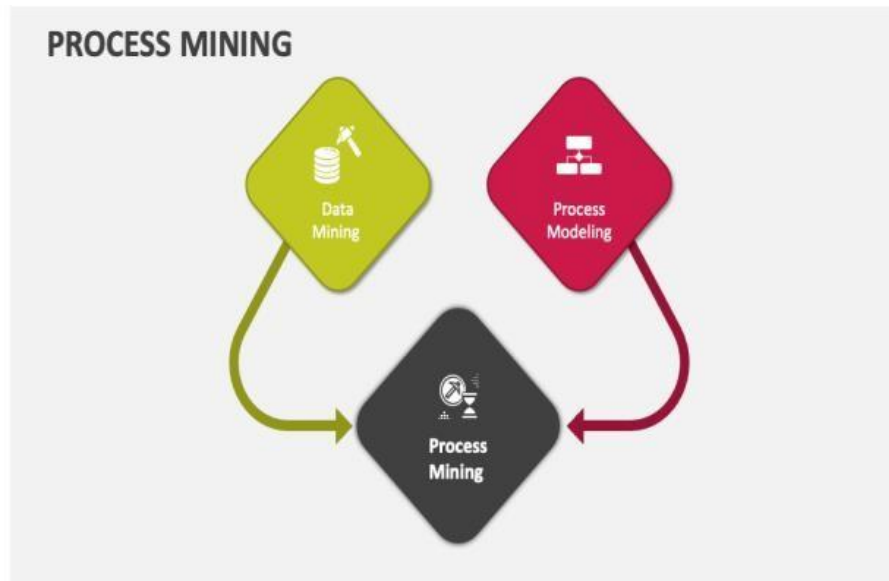


Fig. No. 1.1: Process Mining Model

The real-world examples of process mining include procurement, order management, compliance, intelligent automation, digital transformation, KPI reporting, accounts payable, accounts receivable, auditing, IT development, service management, logistics, and many more.

Understanding Process Mining:

Process mining is a dynamic discipline that involves extracting meaningful insights from event logs to uncover, analyze, and enhance real-world processes within an organization. By bridging the gap between data science and process management, process mining enables us to understand the intricacies of how processes are executed and identify opportunities for improvement.

CHAPTER – 2

FOUNDATION OF PROCESS MINING

Process mining is a set of techniques used for obtaining knowledge and extracting insights from processes by the means of analyzing the event data, generated during the execution of the process.

- This training track provides both the theoretical and applied foundations around Process Mining.
- Process mining reads this data, converts it into an event log, and then creates visualizations of the end-to-end process, along with insightful analytics.
- An event log contains each step performed during the process (the activity), the time at which the event occurred (the timestamp), and for which instance of the process (the case ID).

Using this event log, algorithms generate a process model that shows the process as it really is including the timing of each step and all variations. Process: A Series of linked steps taken in order to achieve a goal. Case: An item or object you follow through the process. Activity: Events that take place during a process.

2.1 Get to know Celonis Analysis:

The Review and Interpret Analyses training track is designed for data and business analysts, process experts, and process improvement specialists. Keep in mind, this track is focused mainly on product know-how and less so on business acumen.

If you'd like to complement your own experience in strategically identifying and prioritizing process inefficiencies, and planning for and implementing improvement measures, then we recommend you take a look at the Deliver Business Value with Celonis training track after completing this one.

Here's a sneak peak of what you'll experience in the Review and Interpret Analyses training track.

Self-placed Reading and Video Demos

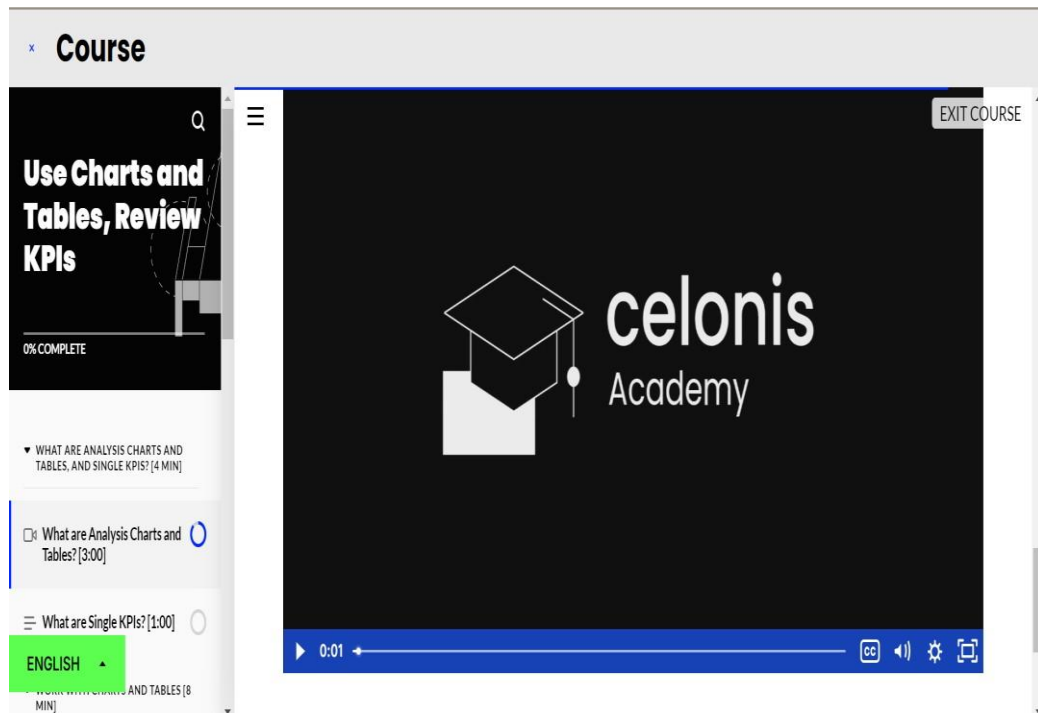


Fig. No. 2.1: Module of Use of Chart and Tables

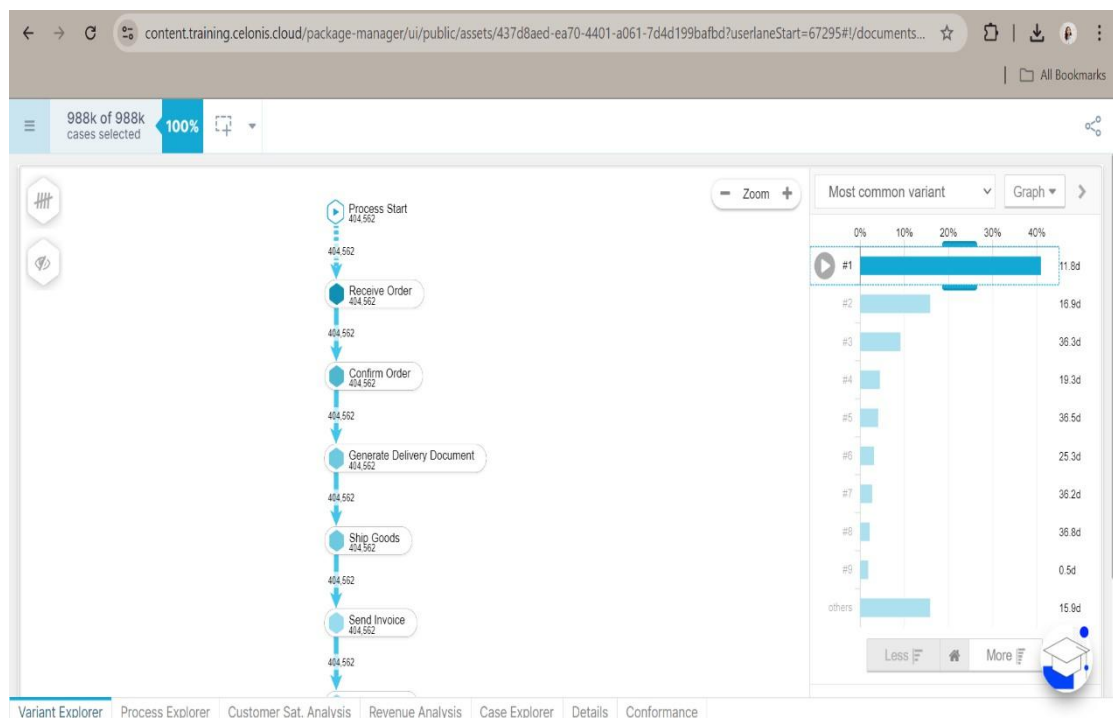


Fig. No. 2.2: Variant Explorer

Process-Specific Examples:

By the end of the training track, you'll have gotten to know the most widely used Celonis Analysis components. Celonis Analysis tools work the same way regardless of the process but since every organization's process is unique with its own nuances, you'll need to apply critical thinking on your part to "connect the dots" so to say, between this training and your own process. We'll help you along the way though, prompting you to reflect and take notes on your ideas to refer to on the job.

2.2 Navigate to an Analysis:

In the concept of navigate to analysis there are three words in it, that are space, package, analysis. These are arranged in hierarchically.

These are used to represent the data in pictorially. By seeing that and graph related to it. We can analyze the data.

These two screenshots are from two public Analysis demos. Depending on which training track you're completing, Review and Interpret Analyses or Monitor KPIs in Analysis Dashboards, you'll actually get to interact with one of these two demos

Guided Learning Tours:

At certain points throughout the courses, you'll be prompted to complete guided tours in a public Analysis demo (no login required) that we've created just for this training.

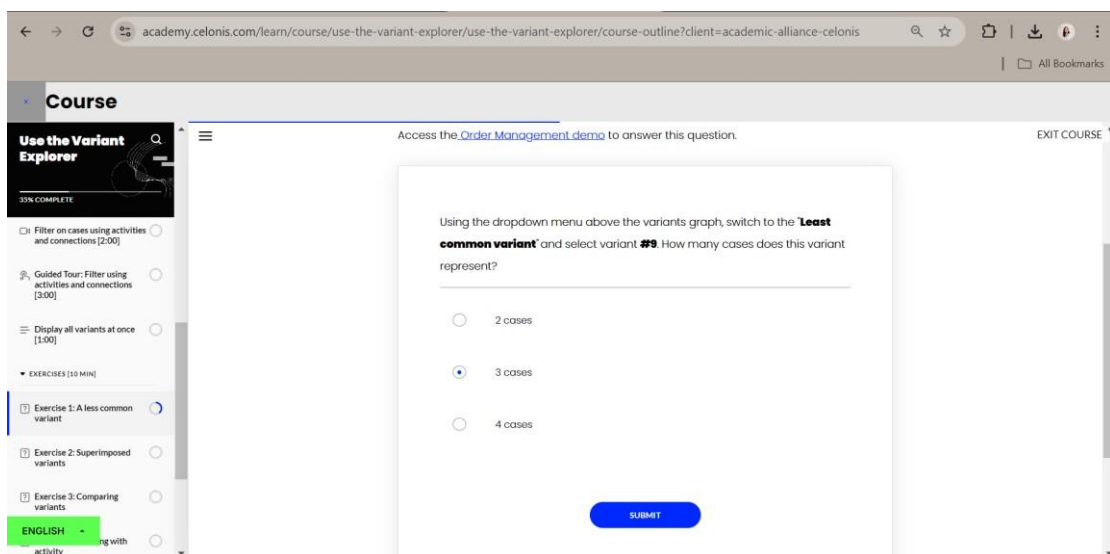


Fig 2.3: Exercise of process mining

2.3 USE VARIANT EXPLORER:

These two screenshots are from two public Analysis demos. Depending on which training track you're completing, Review and Interpret Analyses or Monitor KPIs in Analysis Dashboards, you'll actually get to interact with one of these two demos. Variant Explorer is a Celonis EMS Analysis tool that helps you explore how a specific process flows through your organization.

If we think about a process as a road trip, each process variant would be a potential route. Each activity within a process would be a waypoint along a route, and the connections between activities are like the roads that connect the stops. And, each trip a person makes along a particular route would be a case.

In short, Variant Explorer gives you a quick way to see whether most process cases follow an acceptable flow of activities or not and helps you develop your first analysis questions.

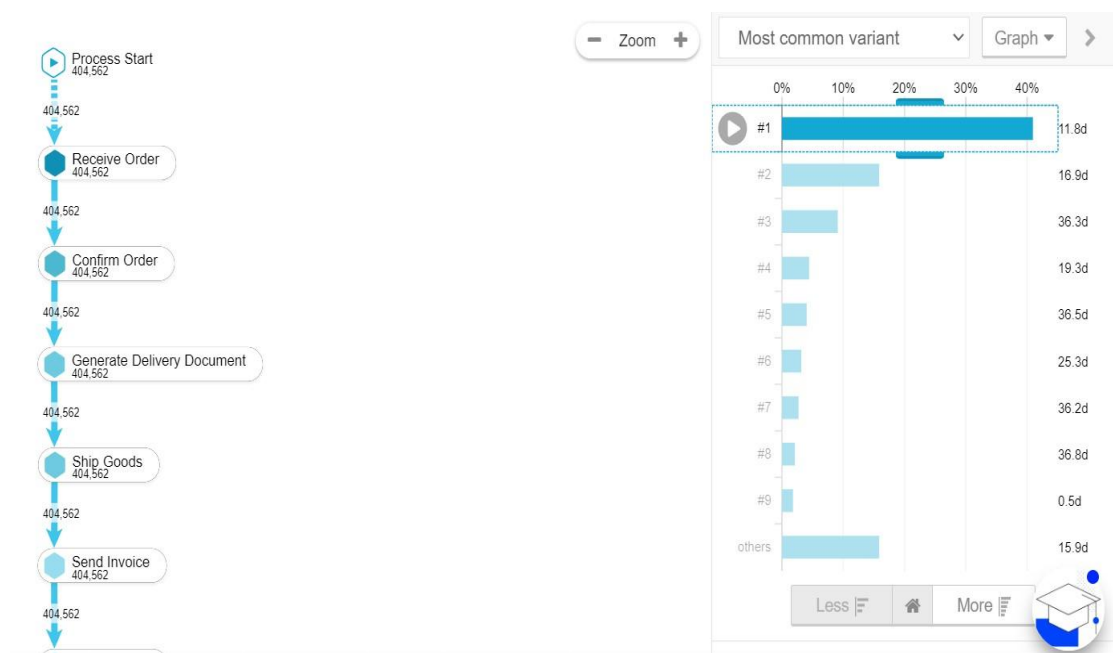


Fig 2.4: Example of Variant Explorer (DDA)

2.4 USE PROCESS EXPLORER:

Process Explorer is a Celonis EMS Analysis tool that helps you explore how process activities are connected. Instead of showing us specific process variants, it shows the most common activities and connections.

Using our road trip analogy again, Process Explorer isn't showing the different routes (variants) that people actually took on a given trip (case). Process Explorer shows us which waypoints (activities) and roads (connections) are the most common along the journey.

Process Explorer is extremely useful for finding infrequent activities, which can be difficult to spot using Variant Explore as these rare activities may not appear in common variants.

2.5 SELECTION VIEWS:

Selection Views offer a more comprehensive set of options to filter on cases as compared to filtering you can do using the components in analysis sheets.

Selection Views Button

You can access the six Selection Views from anywhere in the analysis by clicking on the Selection Views button located in the analysis toolbar.

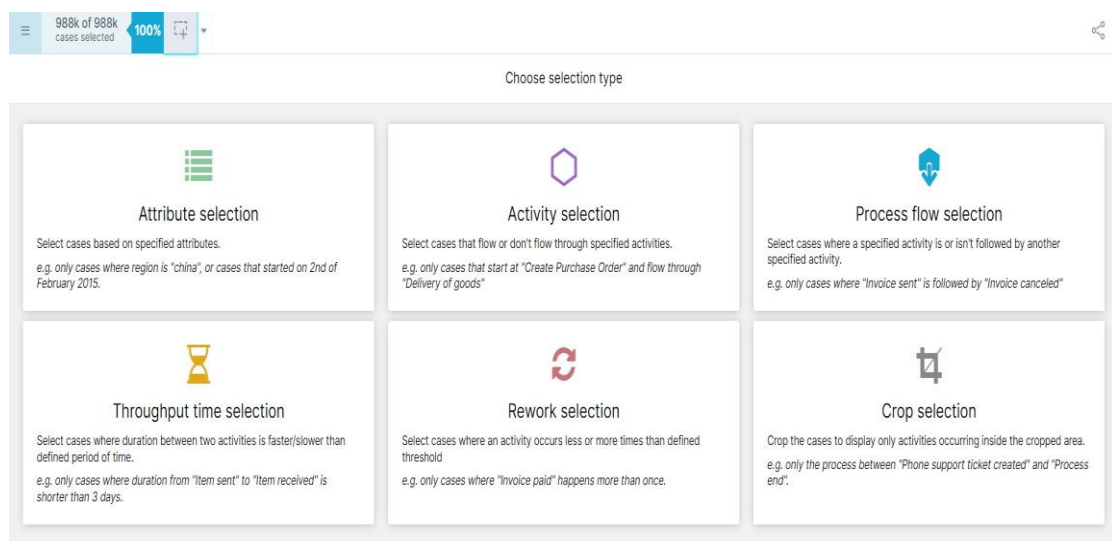


Fig. 2.5 Six selection types

THERE ARE FOUR IMPORTANT STEPS IN PROCESS MINING

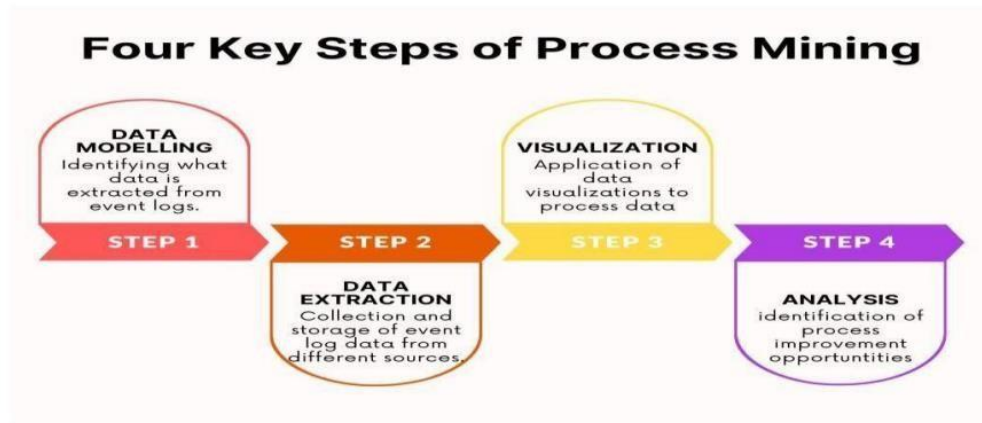


Fig. 2.6 Steps of process Mining

1.Data Modeling:

Identifying the data from which you extracted it from event log(where each event refers to a case, an activity, and a point in time)

2.Data Extraction:

Collection and storage of event log data from different sources

3.Visualization:

To visualize the data, such that it would be helpful.

4.Analysis:

Identification of process improvement opportunities.

Process Flow

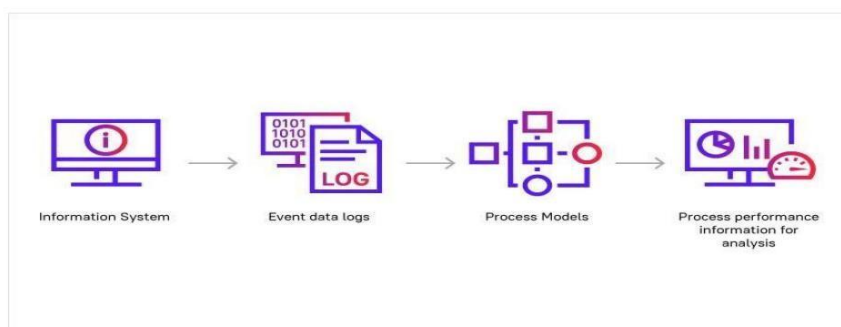


Fig 2.7: Stages of process flow

CHAPTER 3

TECHNIQUES IN PROCESS MINING

Process mining techniques are able to extract knowledge from event logs commonly available in today's information systems. These techniques provide new means to discover, monitor, and improve processes in a variety of application domains. There are two main drivers for the growing interest in process mining. On the one hand, more and more events are being recorded, thus, providing detailed information about the history of processes.

3.1 Techniques:

This manifesto is created by the IEEE Task Force on Process Mining and aims to promote the topic of process mining. Moreover, by defining a set of guiding principles and listing important challenges, this manifesto hopes to serve as a guide for software developers, scientists, consultants, business managers, and end-users. The goal is to increase the maturity of process mining as a new tool to improve the (re)design, control, and support of operational business processes.

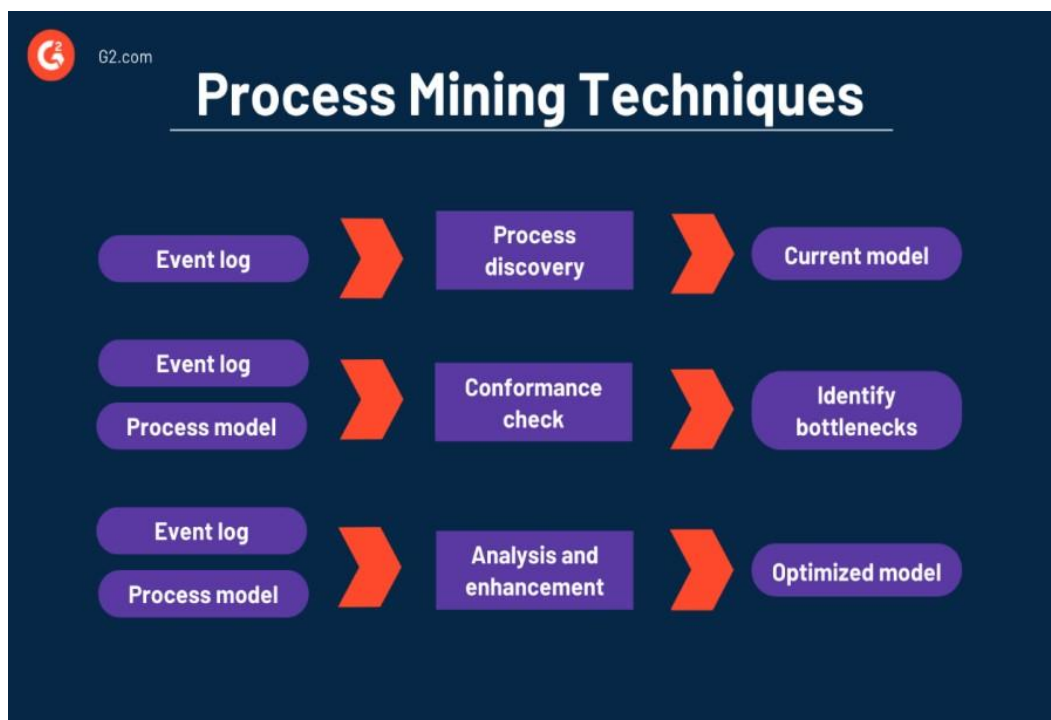


Fig 3.1: Process mining techniques

3.2 Main stages of process mining:

As people (and software) interact with business IT systems, their actions are captured by these systems and can then be transformed into event logs and visualized with the help of process mining.

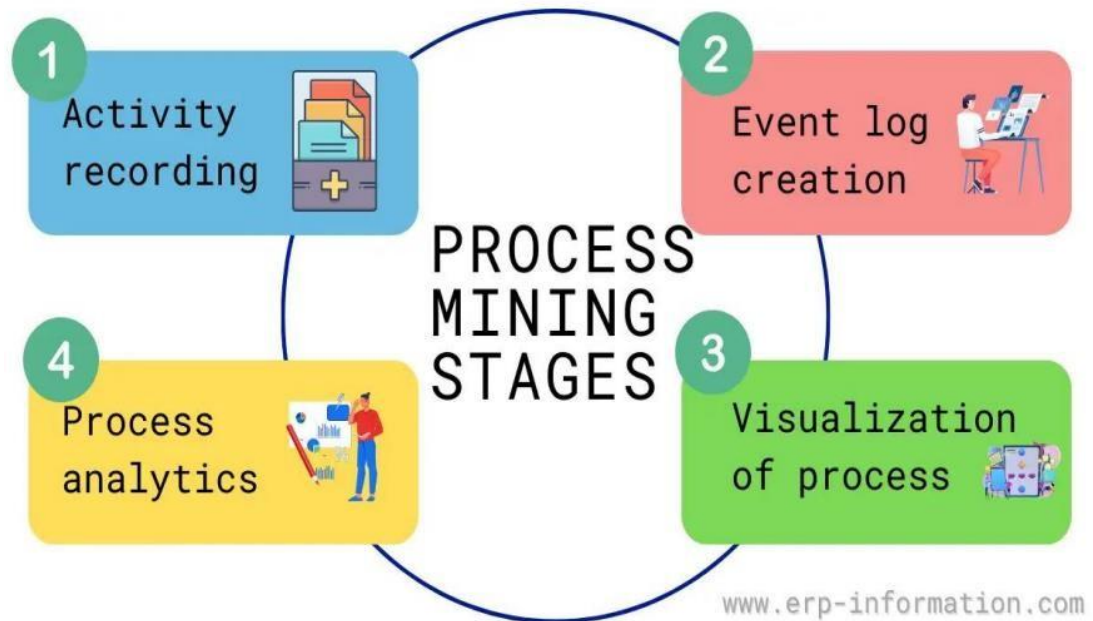


Fig 3.2: Stages in process mining

1. Activity recording:

Activity recording in process mining refers to the process of capturing and logging the activities performed during a particular process. These activities are recorded in the form of event logs which serve as the basis for process analysis and improvement.

2. Event log creation:

Event log creation in process mining involves collecting and recording data about the activities, events and attributes associated with a particular process. This data is then transformed into an event log, which serves as the foundation for process analysis and discovery in process mining.

3.Process analytics:

Process analytics in process mining refers to the analysis and interpretation of event logs to gain insights into the performance, efficiency, and compliance of a process. It involves using various techniques and algorithms to extract valuable information from the event logs, such as process models, performance metrics, bottlenecks, and deviations from expected behavior. These insights can then be used to optimize and improve the process.

4.Visualization of process:

Visualization of process in process mining involves representing the analyzed process data in a visual format. It helps to understand the flow, patterns, and behavior of the process. Visualizations can include process maps, flowcharts, graphs, and diagrams, which provide a clear and intuitive representation of the process and its various activities, paths, and dependencies. This visual representation aids in identifying bottlenecks, inefficiencies, and opportunities for process improvement.

3.3 Types in Process Mining:

The four basic types of process mining: (1) process discovery, (2) conformance checking, (3) process reengineering (changing the process model), and (4) operational support (influencing the process without reengineering it).



Fig 3.3: Types of process mining

1.Process discovery:

Process discovery in process mining refers to the automatic generation of process models based on event logs. It involves analyzing the recorded activities, events, and attributes in the event log to infer the underlying process flow. Process discovery techniques can create various types of process models, such as Petri nets, BPMN diagrams, or process trees, which provide a visual representation of the process and its activities.

2.Process reengineering:

Process reengineering in process mining involves using insights and findings from process mining analysis to redesign and optimize existing processes. It aims to identify inefficiencies, bottlenecks, and areas for improvement in the current process and then implement changes to enhance its performance, efficiency, and effectiveness. Process reengineering can involve modifying process steps, redefining roles and responsibilities, automating tasks, or introducing new technologies to streamline the process and achieve better outcomes.

3.Conformance checking:

Conformance checking in process mining involves comparing the actual execution of a process, as recorded in event logs, with the expected or modeled process. It aims to identify and quantify deviations, such as missing or extra activities, incorrect ordering of activities, or violations of control flow rules. Conformance checking helps to assess how well the observed process aligns with the intended process model, highlighting areas where there may be non-compliance or inefficiencies.

4.Operational support:

Operational support in process mining refers to the use of process mining techniques and insights to improve real-time decision-making and operational efficiency. It involves applying process mining in a live or ongoing operational environment to monitor and analyze process execution, identify potential issues or bottlenecks, and provide actionable recommendations for process improvement.

CHAPTER 4

MODULES EXPLANATION

1.Data Extraction and Collection:

- In this initial phase, relevant data is extracted from various sources, such as ERP systems, databases, and application logs.
- The data collected usually includes timestamps, activity names, case IDs, and other relevant attributes that capture the process execution.

2.Data Preprocessing:

- Raw data often requires cleaning and preprocessing to ensure accuracy and consistency.

Data preprocessing involves tasks like handling missing values, removing duplicates, and transforming data into a suitable format for analysis.

3.Event Log Preparation:

- The preprocessed data is structured into event logs, which contain chronological records of activities associated with case IDs.
- Each event in the log consists of a timestamp, activity name, and case ID, providing a comprehensive overview of process execution.

4.Process Discovery:

- Process discovery aims to create a visual representation of the process based on the event logs.
- Various algorithms are used to generate process models, such as Petri nets, process trees, and flowcharts, reflecting the actual flow of activities.

5.Conformance Checking:

- Conformance analysis compares the actual process execution recorded in the event logs with the expected process model.

- Deviations, bottlenecks, and variations are identified to pinpoint areas where the actual process diverges from the intended model.

6.Performance Analysis:

- This module focuses on evaluating process performance metrics like cycle time, throughput, and resource utilization.
- By analyzing these metrics, inefficiencies and areas for improvement can be identified.

7.Enhancement and Optimization:

- Using insights gained from the previous modules, organizations can optimize their processes.
- Process optimization involves redesigning workflows, reallocating resources, and improving overall efficiency based on data-driven recommendations.

8.Visualization and Reporting:

- Effective visualization tools are used to represent process models, performance metrics, and conformance analysis results.
- Clear visualizations help stakeholders understand complex process dynamics and make informed decisions.

9.Predictive Analysis:

- Some advanced process mining tools incorporate machine learning techniques to predict future process behaviors based on historical data.
- Predictive analytics can be used to forecast potential bottlenecks or delays and make proactive adjustments.

10.Continuous Monitoring:

- After implementing process improvements, continuous monitoring ensures that the optimized processes are maintained over time.

CHAPTER 5

APPLICATIONS AND REAL-TIME EXAMPLES

Celonis:

Explanation

Celonis is a leading process mining platform that helps organizations visualize and analyze their processes using data from various systems. It offers process discovery, conformance checking, and performance analysis.

Features

- **User-Friendly Interface:** Designed for ease of use and accessibility, making it suitable for various users.
- **Process Visualization:** Generates process maps and diagrams that visually represent actual process flows.
- **Root Cause Analysis:** Helps identify underlying causes of process deviations and bottlenecks.
- **Benefits:** Disco provides a simple yet powerful tool set for discovering insights from event data and driving process improvements.

PROM:

Explanation

Prom is an open-source process mining tool set used in both research and industry.

Features

- **Plugin-Based Architecture:** Offers a wide range of plugins for different process mining tasks, allowing users to choose specific techniques.
- **Flexibility:** Allows customization and experimentation with various process mining algorithms.

ProcessGold:

Explanation

ProcessGold offers process mining solutions with a focus on understanding and optimizing processes.

Features

- **Advanced Analytics:** Combines process mining with advanced analytics for deeper insights and predictive capabilities.
- **Process Discovery:** Automatically creates visual process models from event data, helping organizations understand their operations.
- **Benefits:** ProcessGold empowers organizations to enhance their processes, identify improvement opportunities, and predict future outcomes.

Minit:

Explanation

Minit specializes in simplifying process discovery and analysis for organizations.

Features

- **User-Friendly Interface:** Designed for non-technical users, making it accessible for process analysts and business users.
- **Performance Metrics:** Provides performance indicators and metrics to measure process efficiency.
- **Benefits:** Minit aims to make process mining accessible to a broader audience, enabling process optimization and efficiency gains

QPR Process Analyzer:

Explanation

QPR Process Analyzer offers process mining combined with advanced analytics capabilities.

Features

- **Analytics Integration:** Combines process data with advanced analytics to provide in-depth insights and predictions.
- **Real-Time Monitoring:** Monitors processes in real time to identify deviations and opportunities for improvement.
- **Benefits:** QPR Process Analyzer allows organizations to gain a holistic view of their processes, make data-driven decisions, and enhance operational performance.

Each of these process mining applications offers unique features and benefits. The choice of the right application depends on factors like your organization's specific goals, technical requirements, and the complexity of the processes you intend to analyze and optimize.

Some more real time applications are:

1. **Supply Chain Optimization:** Process mining can be used to analyze and optimize supply chain processes in real-time. It helps identify bottlenecks, inefficiencies, and deviations from the ideal process flow, enabling organizations to make informed decisions and adjustments on the fly.
2. **Healthcare Process Improvement:** In healthcare, process mining can be applied to analyze patient treatment pathways, identify variations, and optimize resource allocation in real-time. This can lead to improved patient care and reduced wait times.
3. **Manufacturing Process Analysis:** Process mining can monitor and analyze manufacturing processes in real-time to ensure that production is running smoothly, detect anomalies or deviations from the standard process, and make immediate adjustments to prevent defects or disruptions.
4. **Financial Transaction Monitoring:** Process mining can be used in the financial sector to monitor and detect fraudulent activities in real-time. By analyzing transaction logs, it can identify patterns of suspicious behavior and trigger alerts for further investigation.

5. **IT Service Management:** Process mining can help IT departments optimize their service management processes by monitoring the flow of IT service requests, identifying bottlenecks, and streamlining incident resolution and request fulfillment.
6. **Logistics and Transportation:** Real-time process mining can be applied to logistics and transportation operations to monitor the movement of goods, track delivery routes, and optimize the overall supply chain to ensure timely and efficient deliveries.
7. **Customer Journey Analysis:** Process mining can provide insights into customer interactions and behaviors across various touch points in real-time. This information can be used to improve customer experiences and optimize marketing and sales strategies.
8. **Energy Management:** Process mining can monitor energy consumption patterns in real-time to identify energy wastage, optimize energy usage, and reduce operational costs for industrial facilities and buildings.
9. **Emergency Response and Disaster Management:** During emergency situations, process mining can help emergency responders analyze data in realtime to allocate resources effectively, optimize response times, and make informed decisions to manage crises.
10. **Compliance Monitoring:** Process mining can assist organizations in real-time compliance monitoring by identifying deviations from regulatory processes and providing alerts to prevent potential compliance violations

CHAPTER 6

OUTCOMES

After completing this Training Track, you will be able to:

- Interpret process visualizations and leverage analyses to identify process inefficiencies.
- Conceptualize your process in terms of activities and cases.
- Save an analysis selection for future reference and share it with your team; export visualizations and process data.
- Perform the basic tasks necessary to build Celonis analyses.
- Become familiar with Analysis Settings and Permissions.
- Publish analyses using best practices in version control.
- Put your knowledge about the theoretical foundations of Process Mining into practice.

CONCLUSION

In conclusion, process mining is a powerful and versatile technology that offers valuable insights into the inner workings of organizational processes. By analyzing event data generated during the execution of processes, process mining uncovers hidden patterns, identifies inefficiencies, and provides actionable recommendations for process optimization. This technology has the potential to drive improvements across a wide range of industries, including manufacturing, healthcare, finance, logistics, customer service, and more.

Process mining's ability to visualize process flows, detect bottlenecks, and pinpoint deviations from the ideal path enables organizations to make informed decisions aimed at enhancing efficiency, reducing costs, and improving overall performance. The real-time applications of process mining are particularly noteworthy, as they empower businesses to respond promptly to changing circumstances, address issues as they arise, and ensure that processes operate at their optimal levels.

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.

INTERNSHIP CERTIFICATE



REFERENCES

- **Website:** <https://academy.celonis.com/learn/dashboard>
- **Execution management link:** <https://tinyurl.com/2pvuuzmd>
- **Login page:** <https://academy-login.celonis.com/s/login/>