

CHAPTER – 1

INTRODUCTION

Process mining technology transforms event logs into a process map capable of providing insights to better monitor and improve processes through identifying variants rework patterns, and bottlenecks based on true data.

Process mining is a technique that involves extracting insights and knowledge from event logs generated by various IT systems. It provides a visual representation of real processes, allowing organizations to analyze, optimize, and improve their operations. Through process mining, you can identify bottlenecks, deviations, and inefficiencies in workflows, ultimately leading to more informed decision-making enhanced process efficiency. Process Mining achieves this union by taking the digital footprints that are created in IT systems and using them to reconstruct and visualize process flows. Process Mining achieves this union by taking the digital footprints that are created in IT systems and using them to reconstruct and visualize process flows.

- Process Mining virtual internship is under the AICTE Organization.
- All India Council for Technical Education (AICTE) is a national level council for technical education, under the Department of Higher Education.
- AICTE is associated with Eduskills internship, it has corporate partner Celonis Process Mining academy through this platform they provided Process Mining internship.

1.1 What is a process?

A process is a series of actions or steps taken to achieve a particular result or goal. It involves a sequence of tasks that are performed in a systematic manner to produce an outcome. Processes are commonly used in various fields, such as manufacturing, software development, project management, and more. They help ensure consistency, efficiency, and effectiveness in achieving desired outcomes. Example for that could be a pizza delivery process. It starts with placing the order by calling the pizza company or via their website. Then, the order is assigned to a pizza maker. The pizzamaker bake

the pizza, the pizza, the pizza is packaged, a delivery person delivers it to the assigned address and the payment is received. The problem is this is the ideal scenario of pizza delivery process. But in practice, there are so many things that can go wrong on the way there. The pizzamaker might put the wrong ingredients, the delivery person might go to a different address or the payment fail. Therefore, we can say processes are the engine every experience. Understanding these processes and optimizing are crucial for successful businesses.

1.2 What is Process Mining?

Process Mining is the combination of two disciplines: Data Science and Business Process Management.

Process mining typically involves three main steps: data extraction from event logs, process model creation, and analysis of the generated models to uncover patterns and improvements. Process mining is a data-driven approach that involves extracting valuable insights and information from event logs recorded in various systems. It aims to understand, analyze, and improve real-world processes by visualizing their actual execution based on data. This helps organizations identify inefficiencies, bottlenecks, and opportunities for optimization within their processes.

- Process mining aims to discover, monitor and improve real processes by extracting knowledge from event logs readily available in today's information systems.
- Real process means the process that is actually taking place in an organization. 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 leverages a data-driven approach to process optimization, allowing managers to remain objective in their decision-making around resource allocation for existing processes.

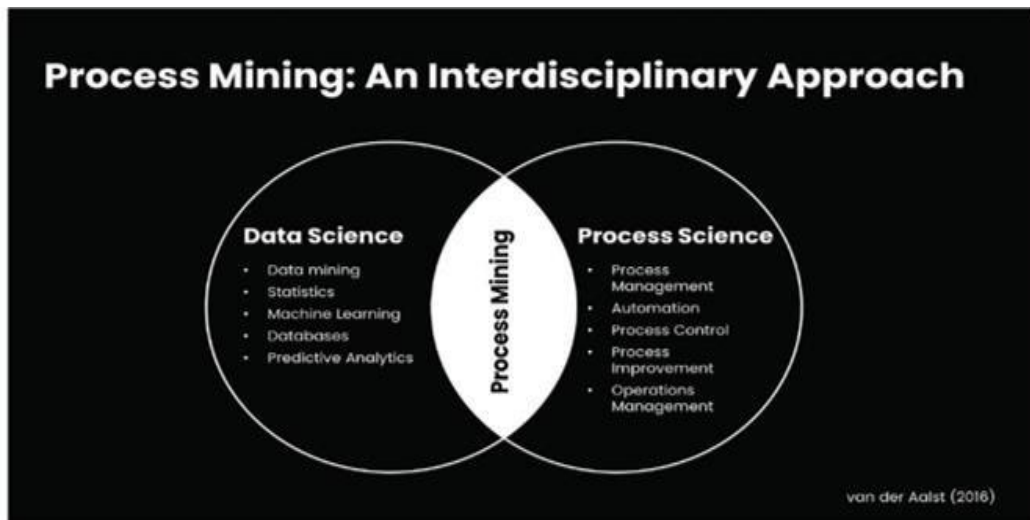


Fig. 1.1:PROCESS MINING

1.3 Importance of Process Mining

Process mining is crucial for organizations as it enables them to gain insights into their operations by analyzing event logs from various systems. It helps identify inefficiencies, bottlenecks, and opportunities for optimization in business processes. This leads to improved operational efficiency, better resource allocation, and enhanced decision making based on data-driven insights. Additionally, process mining aids in compliance monitoring, quality improvement, and overall business transparency.

Process mining helps auditors analyze data faster and enables you to predict where compliance issues and risk factors are likely to exist. Since process mining gives you complete insights into your processes, you get a clearer picture of potential problem areas. As businesses implement digital solutions, automation, and information technology, the demand for process mining software is increasing. This software understands the operations of a business and assists them in finding inefficiencies and bottlenecks in the operation.

1.4 Digital Foot Print

A digital footprint refers to the trail of information that you leave behind while using the internet. This includes your online activities, interactions, posts, and personal data that can be collected and stored by various online platforms and services. It's important to be mindful of what you share online, as your digital footprint can impact your online reputation and privacy.

CHAPTER – 2

TECHNOLOGIES

Process discovery focuses on extracting a process model from event logs. There are various techniques for process discovery, including:

Alpha Algorithm: This technique uses the frequency of events to construct a process model.

It starts with a direct succession relationship and builds a model step by step.

Heuristic Miner: This technique infers a process model by identifying common patterns in event logs. It uses heuristics to find the most likely process flow.

Inductive Miner: This technique uses a divide-and-conquer approach to build a process model. It splits event logs into smaller subsets and constructs models for each subset, which are then combined into a final model.

Conformance Checking:

Conformance checking compares the discovered process model with the actual event logs to identify discrepancies and deviations. Main techniques include:

Token-Based Replay: This technique simulates the execution of the process model using tokens and checks whether the generated traces match the observed event logs.

Fitness Checking: It quantifies how well the process model fits the event logs by calculating the percentage of events that are correctly aligned.

Precision and Generalization Checking: These techniques measure the accuracy of the process model by considering both overfitting and underfitting.

Enhancement and Extension:

Enhancement techniques aim to improve the quality of process models or extend their capabilities. Some techniques in this class include:

Subprocess Discovery: This technique identifies frequently occurring subprocesses within a larger process, allowing for more detailed analysis and optimization.

Multi-Perspective Mining: It involves considering different aspects of the process simultaneously, such as the control flow, data perspective, and time perspective, for a more comprehensive analysis.

Predictive Analytics: By applying machine learning algorithms to historical event data, this technique can predict future process behavior, enabling proactive decision-making.

These three main classes of process mining techniques—Process Discovery, Conformance Checking, and Enhancement/Extension—provide valuable insights into understanding, analyzing, and improving business processes based on event log data.

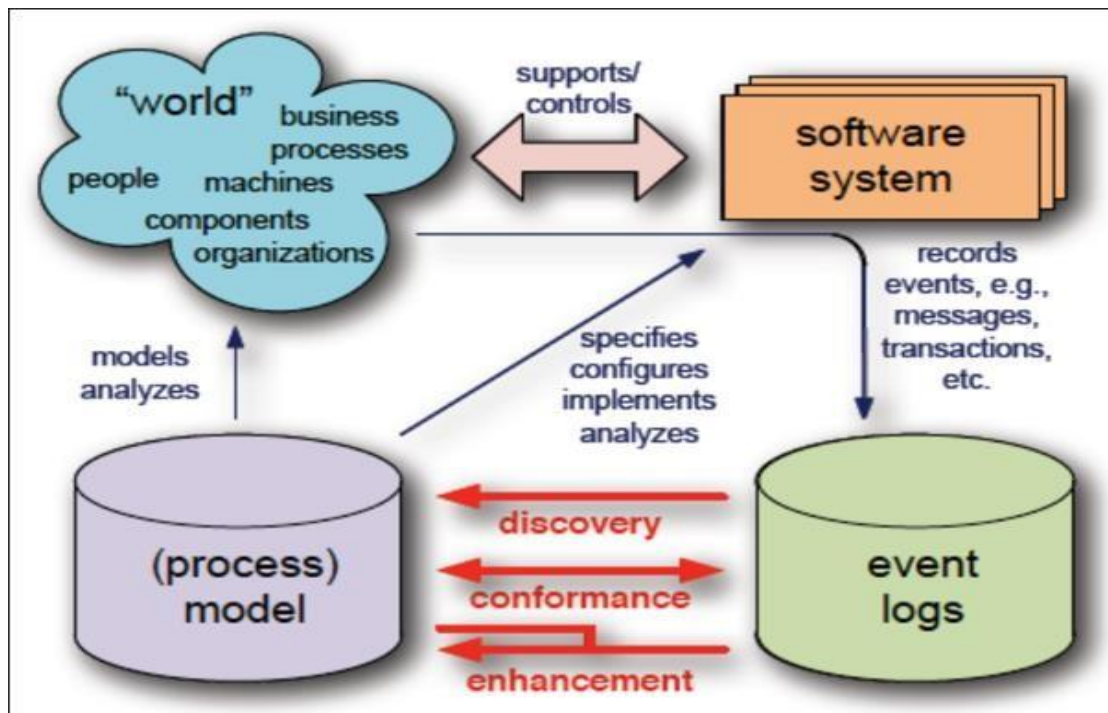


Fig. 2.1:TECNIQUES OF PROCESS MINING

Fundamentals of process mining

Process mining is a technique that involves analyzing and visualizing processes within an organization based on event data recorded in information systems. It aims to uncover insights into how processes are actually executed, identifying bottlenecks, inefficiencies, and opportunities for improvement.

The key fundamentals of process mining:

- **Event Logs:** Process mining starts with event logs, which record the sequence of activities and events in a process. These logs can come from various sources, such as IT systems, databases, or sensors.
- **Data Extraction:** Relevant data is extracted from event logs, including timestamps, activities, users, and outcomes. This forms the basis for analysis.

- **Process Discovery:** The extracted data is used to create visual process models that show how the process flows. These models can be in the form of process maps, flowcharts, or Petri nets.
- **Conformance Checking:** The actual process as depicted in the models is compared to the intended or expected process. Deviations, bottlenecks, and variations are identified, highlighting areas for improvement.
- **Enhancement and Optimization:** Process mining uncovers inefficiencies and areas for optimization. By identifying bottlenecks and deviations, organizations can make informed decisions to streamline processes and improve efficiency.
- **Visual Analytics:** Process mining employs visualizations to make complex process data understandable. These visualizations help stakeholders grasp the process flow and identify patterns.
- **Real-time Monitoring:** Advanced process mining tools can provide real-time monitoring of ongoing processes. This helps organizations identify issues as they arise and take corrective actions promptly.
- **Predictive Analysis:** Some process mining approaches incorporate predictive analytics to forecast future process behavior based on historical data. This can aid in proactive decision-making.
- **Data Privacy and Security:** As process mining involves sensitive data, ensuring data privacy and complying with regulations is crucial. Anonymization and proper data handling techniques are essential.
- **Iterative Approach:** Process mining is often an iterative process. Initial findings might lead to adjustments, refinements, and further analyses to achieve better results.

2.1 Event Logs

An event log is a structured file containing records and timings of events and activities within a computer database. They are typically used to track changes to the database, such as changes to the data structure, data entry records, and even user logins.

Event logging provides a standard and system-specific way for software applications and operating systems to record important events or changes. Event logs are common

in many IT systems, including Windows computers, Customer Relationship Management systems (CRMs) and many Enterprise Resource Planning systems (ERPs.)

	SessionID	Page	Timestamp	CookieID	DataCenter	SiteVersion
Case	487434	portal.aspx	2016-01-01 15:34:01	A	phoenix	1.12
	487434	dashboard.aspx	2016-01-01 15:34:15	A	phoenix	1.12
	487434	purchaseorderreport.aspx	2016-01-01 15:34:30	A	phoenix	1.12
Case	487435	portal.aspx	2016-01-01 14:01:10	B	phoenix	2
	487435	help.aspx	2016-01-01 14:03:23	B	phoenix	2
	487435	contactus.aspx	2016-01-01 14:04:07	B	phoenix	2
Case	487436	portal.aspx	2016-01-01 17:11:17	A	phoenix	1.12
	487436	myteam.aspx	2016-01-01 17:12:41	A	phoenix	1.12
	487436	expensereports.aspx	2016-01-01 17:12:55	A	phoenix	1.12

Fig.2.2 : EVENT LOGS EXAMPLE

Process mining assumes the existence of an event log where each event refers to a case, an activity, and a point in time. An event log can be seen as a collection of cases and a case can be seen as a trace/sequence of events. Event data may come from a wide variety of sources:

- a database system .
- a comma-separated values (CSV) file or spreadsheet,
- a transaction log (e.g., a trading system),
- a business suite/ERP system (SAP, Oracle, etc.),
- a message log (e.g., from IBM middleware)
- an open API providing data from websites or social media.

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

- **Case ID:** a unique identifier such as a purchase order item, invoice number or order number. It is a unique identifier for any business object or transaction that is tracked in event logs.
- a **activity**: the description of what has happened - for example, the creation of a purchase order or the receipt of goods. Activity - is a task or action taken within a business process, for example, “approve,” “reject” or “request.”

- **timestamp**: the date and time that the activity took place. A timestamp indicates precisely when each of the activities took place.

2.2 Data Extraction

Data extraction in process mining involves gathering raw event data from various sources, such as logs, databases, or other systems, to analyze and visualize business processes. This data usually contains timestamps, activities, actors, and other relevant information. The extracted data is then used to create process models, analyze process efficiency, identify bottlenecks, and improve overall process performance.

Data extraction is the first step in both ETL (extract, transform, load) and ELT (extract, load, transform) processes. ETL/ELT are themselves part of a complete data integration strategy. Example : Extraction of phone numbers from a digital directory which is already organized based on a logical scheme. Data that is stored in a structured

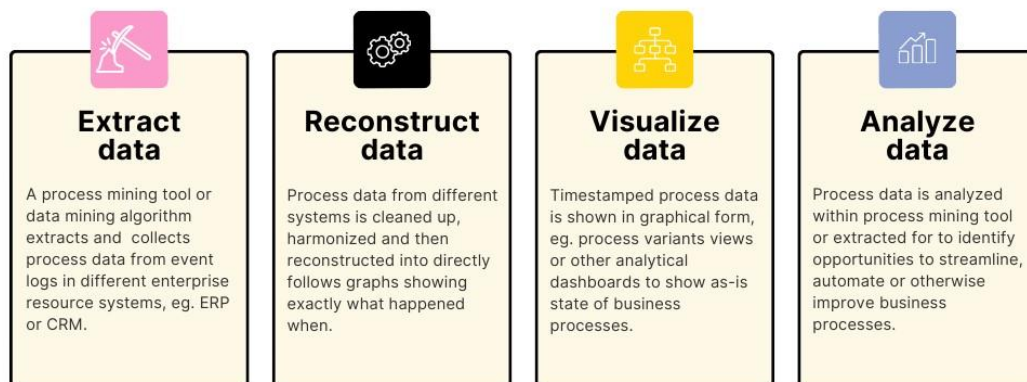


Fig 2.3 : HOW PROCESS MINING WORKS IN 4 STEPS

- process mining, data extraction involves collecting event logs from various sources, such as IT systems, to analyze and visualize processes. You'll need to specify which process you're interested in, identify relevant data sources, extract event logs, and format them in a structured way. This data can then be used for process analysis, optimization, and visualization. Process mining tools often provide functionalities to help with data extraction and preprocessing. If you need more specific guidance, feel free to ask!
- Reconstruct data refers to the process of assembling, organizing, and preparing event logs or data in a format suitable for analysis. This involves cleaning, transforming, and aggregating data to create a clear picture of the process flow

and activities. Process mining tools often include features to assist with data reconstruction, such as aligning events, handling missing data, and creating process models.

- Visualizing data in process mining is crucial for gaining insights and understanding the flow of activities within a process. You can use various visualization techniques to represent the process behavior, bottlenecks, and inefficiencies.
- In process mining, visualizing data helps you understand and analyze the flow of processes within your organization

2.3.Reconstruction of Data

Reconstructing data in process mining involves creating a representation of past events and activities within a business process. This is typically done using event logs, which record the sequence of activities, their timestamps, and other relevant information. The goal is to gain insights into how a process operates, identify bottlenecks, inefficiencies, and opportunities for improvement. Here's how data reconstruction is typically carried out in process mining:

Data Collection: Gather event data from various sources within an organization. This data should include details about activities, timestamps, case identifiers, resource assignments, and potentially other attributes depending on the process being analyzed.

Preprocessing: Clean and preprocess the data to remove duplicates, handle missing values, and ensure consistency. Data quality is essential for accurate process analysis.

Event Log Creation: Transform the raw data into an event log format, where each event corresponds to an activity instance. The event log should capture the sequence of activities, their timestamps, and any additional attributes needed for analysis.

Case Identification: Define the concept of a "case" or process instance. Each case represents a specific execution of the process and is identified using a unique case identifier.

Event Ordering: Arrange the events within each case in chronological order. This is essential for reconstructing the process flow accurately.

2.4 Visualization of Data

Visualization plays a crucial role in process mining as it helps make complex process data more understandable, enabling organizations to identify patterns, bottlenecks, and opportunities for improvement. Here are some common types of visualizations used in process mining:

Process Flow Diagrams: Process flow diagrams provide a visual representation of the sequence of activities in a process. They use nodes to represent activities and directed edges to show the flow of work from one activity to another. This type of visualization helps in understanding the overall process structure.

Event Logs: Visualizing event logs can provide insights into the distribution of activities over time. Event log visualizations might include bar charts, histograms, or line graphs showing activity frequencies, durations, and timestamps.

Process Maps: Process maps visualize the relationships between activities, showing how different activities are connected and the paths that cases take through the process. They often use symbols to represent different types of activities, decision points, and loops.

Petri Nets: Petri nets are graphical models used to represent processes. They show the interactions between activities, decision points, and synchronization points. Petri nets can provide a more detailed view of the process's behavior, including concurrency and resource allocation.

Directly-Follows Graphs: These graphs represent the direct relationships between activities in a process. Nodes represent activities, and directed edges indicate the order in which activities are typically executed. This visualization is useful for identifying frequently followed paths.

Performance Metrics Visualizations: Visualizing key performance indicators (KPIs) such as process throughput, cycle times, and resource utilization can help in assessing process efficiency and identifying areas for improvement.

2.5. Analyzation Of Data

Data analysis is a critical process of examining, cleaning, transforming, and interpreting data to discover useful information, draw conclusions, and support

decision-making. In the context of process mining, data analysis involves exploring event logs and other relevant data sources to understand the behavior and performance of business processes. Here's how data analysis is typically conducted in the context of process mining:

Data Collection: Gather event data from various sources within the organization. This data should include details about activities, timestamps, case identifiers, resource assignments, and potentially other attributes.

Data Preprocessing and Cleaning: Clean the data to remove duplicates, handle missing values, and ensure data consistency. Poor-quality data can lead to inaccurate analysis results.

Event Log Creation: Transform the raw data into an event log format, where each event corresponds to an activity instance. Organize the events by case identifiers and order them chronologically.

Descriptive Analysis:

Frequency Analysis: Determine the frequency of each activity to understand which ones are most common.

Duration Analysis: Analyze the duration of activities and identify outliers or bottlenecks.

Resource Utilization: Examine how resources are assigned to activities and assess resource efficiency.

Process Discovery: Apply process discovery techniques to generate visual representations of the process flow, such as process flow diagrams, Petri nets, or directly-follows graphs. This step helps you understand the typical sequence of activities.

Performance Measurement:

Key Performance Indicators (KPIs): Calculate KPIs like throughput, cycle time, and lead time to assess the process's efficiency and effectiveness.

Compliance Analysis: Evaluate how closely actual process executions adhere to predefined rules and guidelines.

CHAPTER – 3

APPLICATIONS

Process mining, while often associated with business and organizational contexts, can also have applications in daily life scenarios:

- 1. Personal Time Management:** Process mining can help analyze your daily routines and activities, revealing patterns and inefficiencies that could be optimized for better time management.
- 2. Fitness and Health:** By analyzing your exercise and dietary habits, process mining can provide insights into your fitness journey, helping you adjust your routines for better results.
- 3. Home Automation:** Process mining can be used to analyze your interactions with smart home devices, enabling you to optimize energy consumption and automate tasks more effectively.
- 4. Cooking and Meal Planning:** Process mining can analyze your cooking processes and meal planning habits, helping you streamline your cooking routines and discover new recipes.
- 5. Study and Learning:** Process mining can assist in analyzing your study habits, identifying areas for improvement, and suggesting effective learning strategies.
- 6. Shopping Habits:** By analyzing your online and offline shopping behaviors, process mining can offer recommendations, improve shopping efficiency, and manage impulse purchases.

CHAPTER – 4

MODULES EXPLANATION

The Process Mining Fundamentals for Students training track gives learners insights into both the theoretical and applied foundations around Process Mining.

There are three sublevels in Process Mining Fundamentals. They are:

1. Review and Interpret Analyses
2. Build Analyses
3. Case study

1. Review and Interpret Analyses :

The learner can conceptualize in terms of activities and cases. They can use the Variant Explorer for an overview of all the ways the process is flowing and the Process explorer to quickly identify undesired activities and see how cases are flowing into and out of individual activities. They are able to verify assumptions about how the process is performing and to focus on attributes of cases reflecting inefficiency. There are several stages in Review and Interpret Analyses. They are:

- **Get to know Celonis Analysis**
- **Navigate to an Analysis :** Back on the job, in your organization's Celonis EMS, to access the analyses made available to you, you'll need to log in to your Celonis team and navigate to the 'Apps' service. 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.
- **Use the Variant Explorer :** As the name implies, using the Variant Explorer, you can discover all the process variants that is all the different ways the process flows in your organization. The Variant Explorer is one of the Analysis tools to help you take an "exploratory" approach to find out how your process is performing.

- **Use the Process Explorer** : The Process Explorer 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. You'll be nicely surprised by all that you can accomplish with it.
- **Use Selection Views** : Selection Views offer a more comprehensive set of options to filter on cases as compared to what you can do in the components in analysis sheets. Process Mining Virtual Internship
- **Use the Case Explorer** : The Case Explorer is useful once you've narrowed down the analysis to a few cases that you want to investigate further. You can view specific case details such as timestamp of activities, user type, possibly even user name, and other useful pieces of info.
- **Use the conformance Checker** : Every organization has an optimal process in mind that they want to achieve. With the conformance checker, you can see how far away the organization is from reaching that goal and investigate common patterns for inefficiency.
- Save and Share Analysis Selection, Export Data
- Wrap-up: Review and Interpret Analysis

2.Build Analysis

- 1.Get Started. Get Ready to Build Analyses. Required.
- 2.Build the Heart of the Analysis. Configure Tables and Charts in Analysis.
- 3.Build KPIs with the visual Editor. Configure Standard Process KPIs in the visual Editor.
- 4.Take it to the Next Level. Configure a Conformance Checker Sheet.

5.Wrap-up.Wrap -up: Build Analyses

Celonis Academy offers you a unique learning experience to achieve that mission: The Rising Stars Program.

Ignite your learning journey with foundational knowledge about Process Mining.

Then specialize in your focus area:

- Business
- Technical
- Automation
- Research

Upon successful completion , you'll earn a Rising Star Badge which will set you apart when applying for process mining jobs.

4.1 Pathway to Rising Business Star

1.Deliver Business Values with celonis :

- Frame value by identifying and qualifying improvement opportunities.
- Use Celonis to derive the root causes of improvement opportunities and to quantify the business impact of resolving them.
- Develop an action plan for relevant improvement actions.
- Monitor realized value and adoption and communicate successes.

2.Create and Delivar Demos:

Design, plan and structure a demo around a clear value-focused story with the customer's use cases at its center.

- Deliver a business value oriented demo adapting your presentation to the audience at hand.
- Locate , deploy and tailor relevant demos in the Try Celonis platform.
- Determine when to use the Demo Gallery.

- Build custom demos in the Demo Generator Tool
- Tailor product demos by configuring tables, charts, single KPIs, selections, and design components. Pathway to Rising Technical Star - >Build Analyses.

4.2 Process Query Language(PQL)

Celonis Process Query Language (PQL) is a specialized query language designed for analyzing and querying data within the Celonis platform. It's used to extract insights and information from business process data, allowing users to create custom reports, visualizations, and analyses. PQL enables users to perform complex data manipulations, transformations, and aggregations to uncover process inefficiencies and opportunities for improvement. It's a core component of the Celonis platform, which focuses on process mining and process improvement within organizations.

PQL enables the user to translate process-related business questions into queries, which are then executed by a custom-built query engine. PQL covers a broad set of operators, ranging from process-specific functions to aggregations and mathematical operators. Its syntax is inspired by SQL, but specialized for process-related queries.



Fig 4.1 : PQL QUERIES BADGE

4.3.Execution Management System(EMS):

Business Analyses Advanced

- Get to know PQL and how to work with it in the analysis code editor.

- Get to know use cases that can only be solved with PQL's PU-functions.
- Create static background filters at analysis-, sheet-, and component-level.
- Create benchmarking analysis sheets to compare attributes (such as regions, vendors, or customers) side-by-side. – Configure interactive, dynamic drilldown tables that hold several dimensions and display one at a time based on user's selection by using variables.
- Configure interactive, dynamic drilldown tables that hold several dimensions and display one at a time based on user's selection by using variables.

Get Data into EMS

- Connect to systems.
- Extract data.
- Transform data.
- Set up and load a data model.
- Schedule regular data loads.
- Monitor and validate data.
- Apply SQL performance best practices.



Fig.4.2.CELONIS RISING STAR BUSINESS BADGE

CHAPTER – 5

REAL TIME APPLICATION

Hospital Patient Care Optimization:

- **Patient Admission:** When a patient arrives, their information is entered into the hospital system. Process mining tools analyze this step to ensure timely admission and proper allocation of resources.
- **Treatment Workflow:** As patients move through various treatment stages (tests, consultations, procedures), real-time process mining monitors the workflow to identify bottlenecks or delays.
- **Medication Administration:** Process mining analyzes the administration of medication, ensuring that doses are given on time and tracking any deviations from prescribed protocols.
- **Nurse and Staff Allocation:** The system monitors the workload of nurses and staff, reallocating resources in real-time to ensure balanced work distribution.
- **Emergency Response:** If a critical situation arises (e.g., sudden increase in patient arrivals), process mining helps hospital administrators adjust resources and prioritize cases accordingly.
- **Discharge Process:** Real-time analysis ensures efficient discharge processes, minimizing waiting times for patients ready to leave the hospital.

By using real-time process mining, the hospital can streamline its operations, reduce patient waiting times, enhance staff efficiency, and ultimately provide better patient care. This approach allows the hospital to adapt quickly to changing conditions and make data-driven decisions to optimize its processes.

CHAPTER – 6

LEARNING OUTCOMES

After you complete this training, you should be able to:

After you complete this training, you should be able to:

- ⑦ Understand what process mining is and the basics of how it works.
- ⑦ Summarize what an event log is and why we need it for processing.
- ⑦ Identify business use cases for process mining.
- ⑦ Learn how to find training courses to get started.
- ⑦ Understanding how to discover, analyses, and improve business process using data driven techniques.
- ⑦ will learn to extract insights from event logs, identify bottlenecks, inefficiencies, and opportunities for optimization.
- ⑦ And also, you will learn to extract to create visual representations of processes to aid decision making and process improvement efforts.
- ⑦ You will gain skills in using process mining tools and interpreting the results to enhance organizational efficiency and effectiveness.

CONCLUSION :

Overall, process mining can be a powerful tool for improving the efficiency and effectiveness of business processes, and is increasingly used by organizations to drive process improvement efforts. It offers objective, fact-based insights, derived from actual data, to help you audit, analyze, and improve your existing business processes. Process mining aims to extract information from event logs, which are recorded from running business processes

INTERNSHIP CERTIFICATES



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

JYOTHI YANAMALA

Srinivasa Ramanujan Institute of Technology

has successfully completed 10 weeks
Process Mining Virtual Internship
during May - July 2023

Supported By 



Jerome Geyer-Klingenberg
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 :00c9b3a16a975882d77e2b814a412889
Student ID :STU622f60cfc7ba81647272143



REFERENCES:

- <https://academy.celonis.com/learn/dashboard>
- . IEEE CIS Task Force on Process Mining. Process Mining Manifesto
- van der Aalst, Wil (2016). Process Mining: Data Science in Action.
- Burattin Andrea. Applicability of Process Mining techniques in Business Environments
- D.Harel and R.Marely. Come, Let's play: Scenari-Based Programming Using LSC's and
- Play-Engine. Springer, Berlin, 2003