# Example of Input and Output for Event Log Mining

## Target Variable (y)

The target variable y represents the analytical outcome derived from process mining on system event logs.   
In this project, y denotes the system process performance status or conformance result, indicating whether a process instance behaves as expected.  
This is typically a binary or categorical variable:  
• y = 1 (or “Conformant”) – The process instance follows the expected workflow.  
• y = 0 (or “Non-conformant”) – The process instance contains deviations, delays, or anomalies.

## Feature Vector (X)

X is the collection of all log attributes used for event correlation and performance evaluation.   
Each event within the log contains features that describe the system state and operational context. Typical attributes include:  
• Timestamp – Event occurrence time.  
• Component – Source service, module, or application.  
• Event Type – Category of event (INFO, WARN, ERROR).  
• Activity Name – Specific action or operation performed.  
• Case ID – Unique identifier for process instance (session ID, request ID, etc.).  
• Resource/User – Actor or subsystem triggering the event.  
• Duration – Time difference between start and completion of related events.  
• Severity Level – Indicates importance or criticality of event.

For a given process instance, a complete feature vector X might appear as follows:  
X = [Timestamp, Component, Event Type, Activity Name, Case ID, Resource, Duration, Severity]

## Examples of (X, y)

Below are two examples showing how system event logs are represented for process mining and how the model interprets their outputs.

Example 1: A process instance that conforms to the expected flow  
• X (Features): { Timestamp: “2025-10-14T10:15:24Z”, Component: “AuthService”, Event Type: “INFO”, Activity: “User Login Success”, Case ID: “REQ-4587”, Resource: “UserID-102”, Duration: 0.35s, Severity: “Low” }  
• y (Target): Process Status: “Conformant” (1)

Example 2: A process instance containing deviations  
• X (Features): { Timestamp: “2025-10-14T10:16:02Z”, Component: “PaymentService”, Event Type: “ERROR”, Activity: “Transaction Timeout”, Case ID: “REQ-4592”, Resource: “UserID-108”, Duration: 12.4s, Severity: “Critical” }  
• y (Target): Process Status: “Non-conformant” (0)

## Objective

The model aims to learn relationships between log attributes (X) and the resulting process status (y) to automatically detect anomalies,   
improve trace reconstruction, and optimize system performance through intelligent event log analysis.