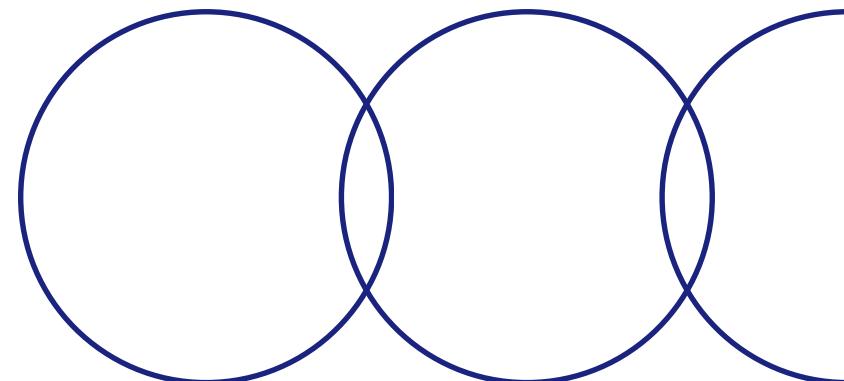


Process Mining Dashboard

Presented by Rotunno Roberto
ID: 859207



Dataset Overview

BPI Challenge Context

The dataset consists of **13,087 cases and 262,200 events**, illustrating a complex landscape of **approximately 4,000 variants** within a financial institution focused on personal loans.

Preprocessing Strategy

Key Steps for Data Preparation and Cleanup

Time Normalization

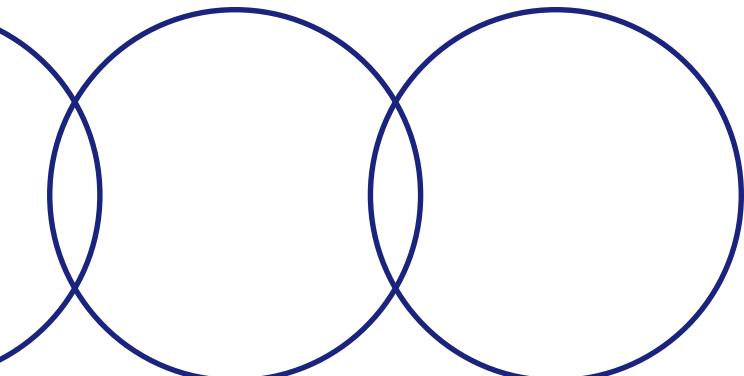
Converting all timestamps to UTC ensures consistent duration calculations across the dataset, facilitating accurate time-based analysis and comparisons.

Filtering

By retaining only COMPLETE transitions, we significantly reduce graph complexity, enhancing the clarity and interpretability of the resulting process models.

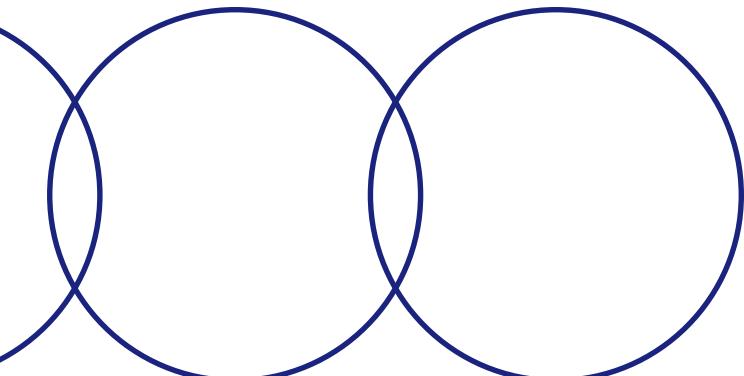
Hybrid Sampling

Utilizing a small subset for live demos allows for quicker processing, while the full dataset ensures accuracy in benchmarking and result validation.



Process Discovery Algorithms

Overview of algorithms used for process mining analysis



Alpha Miner

Alpha Miner identifies basic causal links in event logs but struggles with complex loops, leading to less informative models.

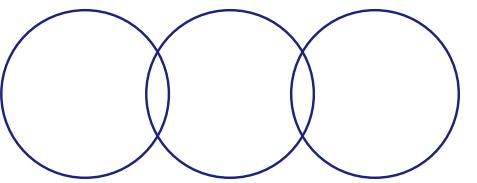
Inductive Miner

Inductive Miner produces sound process models but often underfits, resulting in overly generic representations of the actual processes.

Heuristic Miner

Heuristic Miner excels at handling noise and exceptions, providing robust models with a high degree of accuracy and interpretability.

Model Evaluation Metrics



Assessing Performance of Algorithms

Fitness Score

The fitness score indicates how well the model conforms to the actual process behavior, reflecting the percentage of observed behavior captured by the discovered model, crucial for evaluation.

Precision Metric

Precision measures the accuracy of the model's predictions. It helps determine how many of the predicted behaviors are accurate, providing insights into the reliability of the algorithm used.

F1-Score

The F1-Score combines precision and recall, providing a single metric that balances the trade-off between false positives and false negatives, allowing for an overall assessment of model effectiveness.



Control Panel

1. Caricamento Dati

Upload XES File

Drag and drop file here

Limit 200MB per file • XES, XML

Browse files

File Default rilevato

2. Ottimizzazione

Max Casi (Sampling)

13087

Filtro Varianti (Semplificazione)

1.00

Totale Casi nel file: 13087

Status: Ready

Process Mining Dashboard | BPI Challenge 2012

Stats Process Discovery Prediction Chat AI

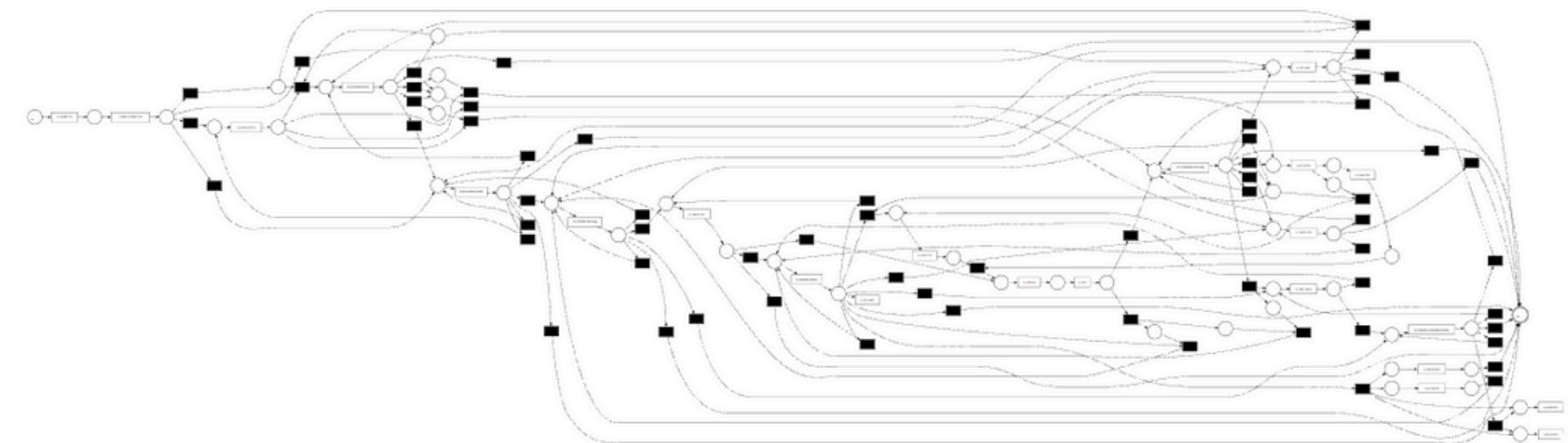
Process Discovery & Evaluation

Algoritmo: [?](#)

- Heuristic Miner
- Inductive Miner
- Alpha Miner

Genera e Valuta Modello

Analisi Completata



Metriche di Valutazione

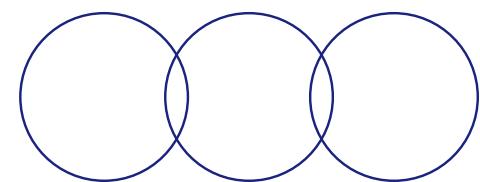
Fitness: 90.89%

Precision: 91.58%

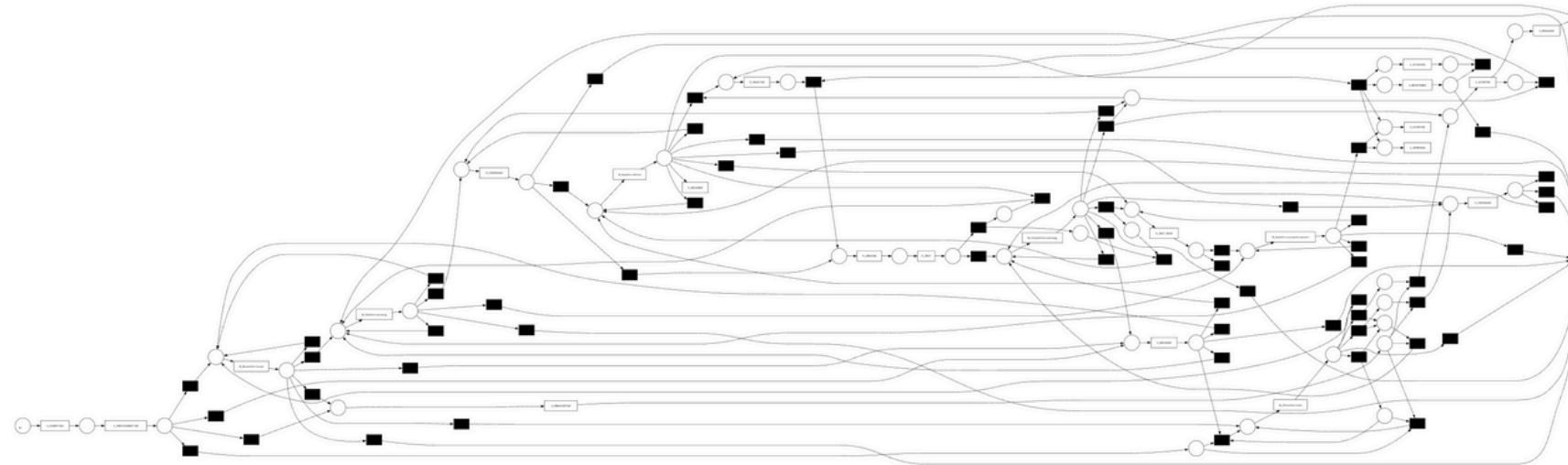
F1-Score: 91.24%

Process Discovery Algorithms:

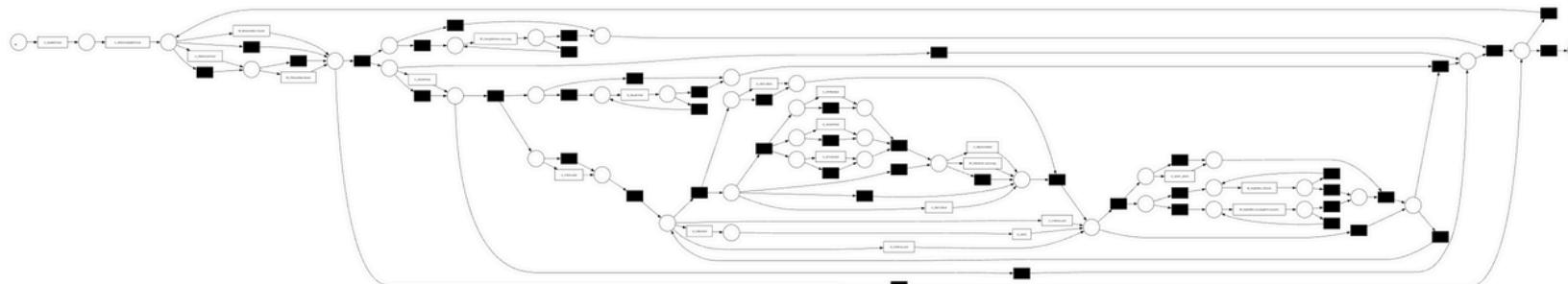
Visual Comparison



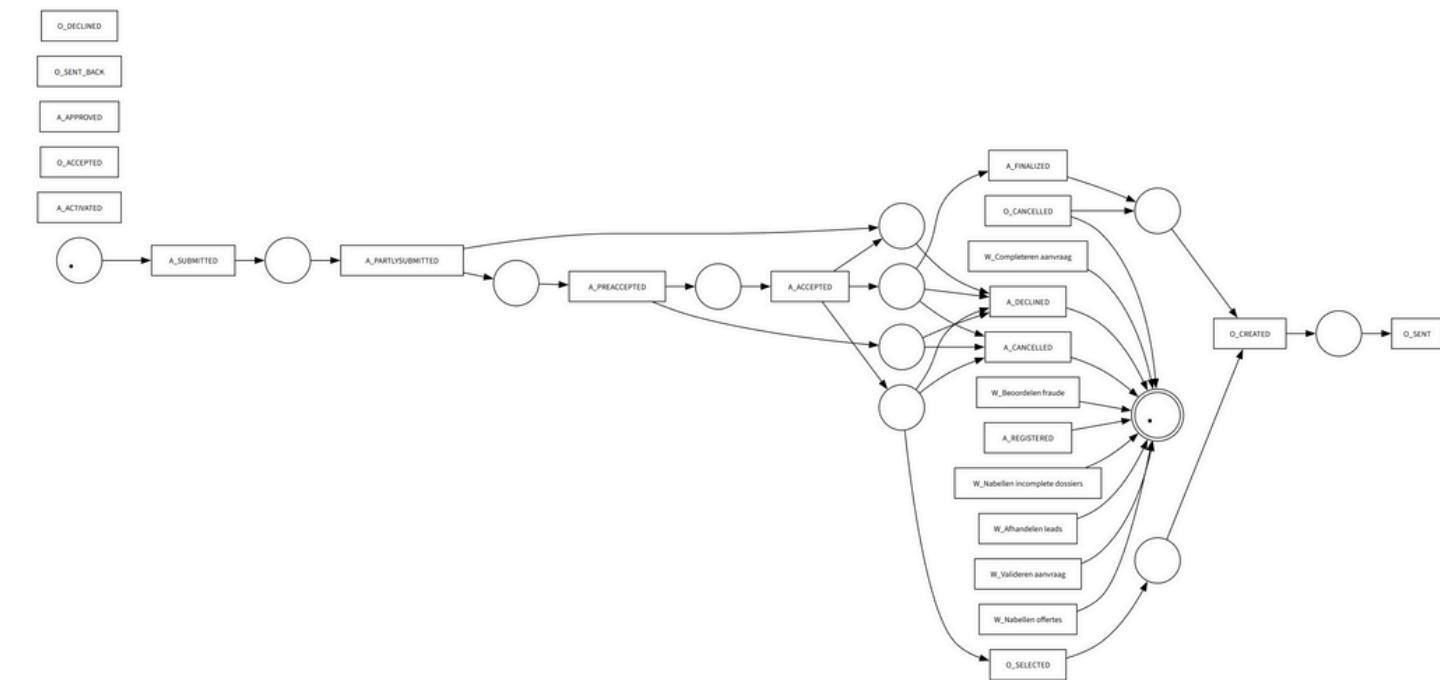
Heuristic miner



Inductive miner

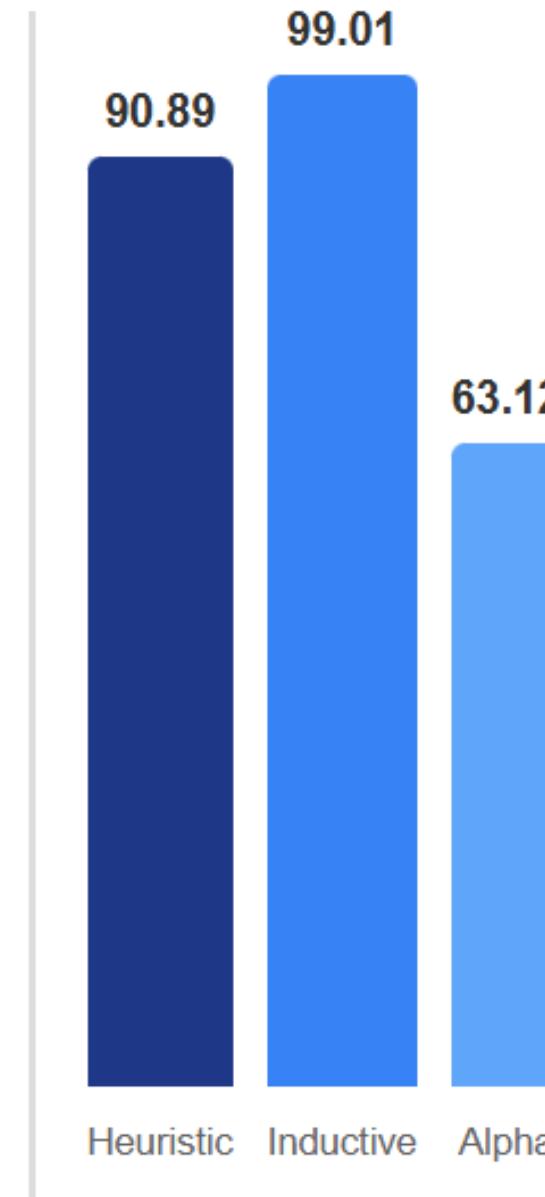


Alpha miner

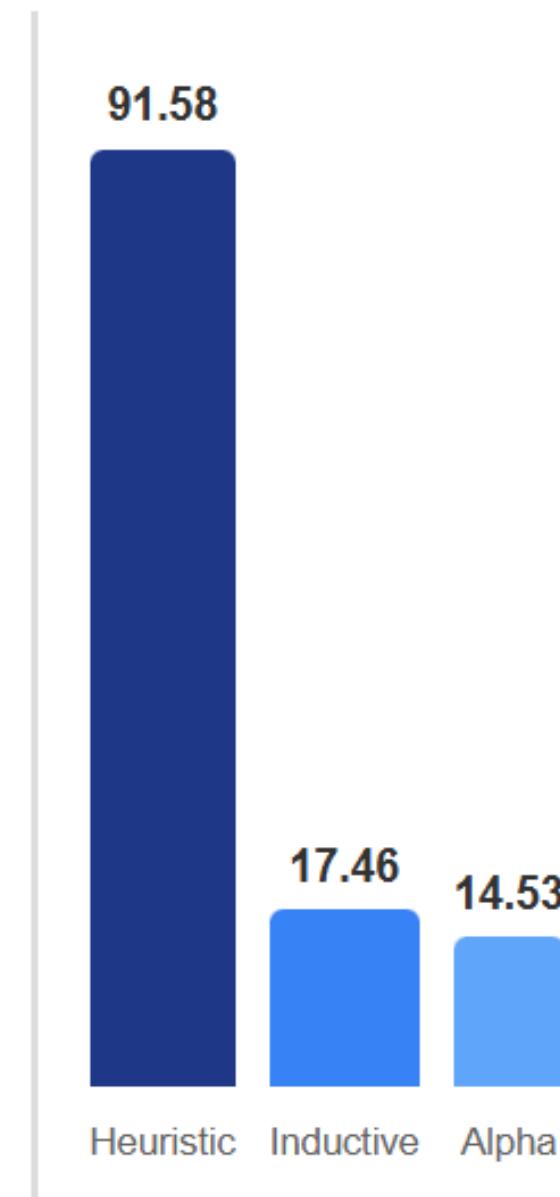


Benchmark Results

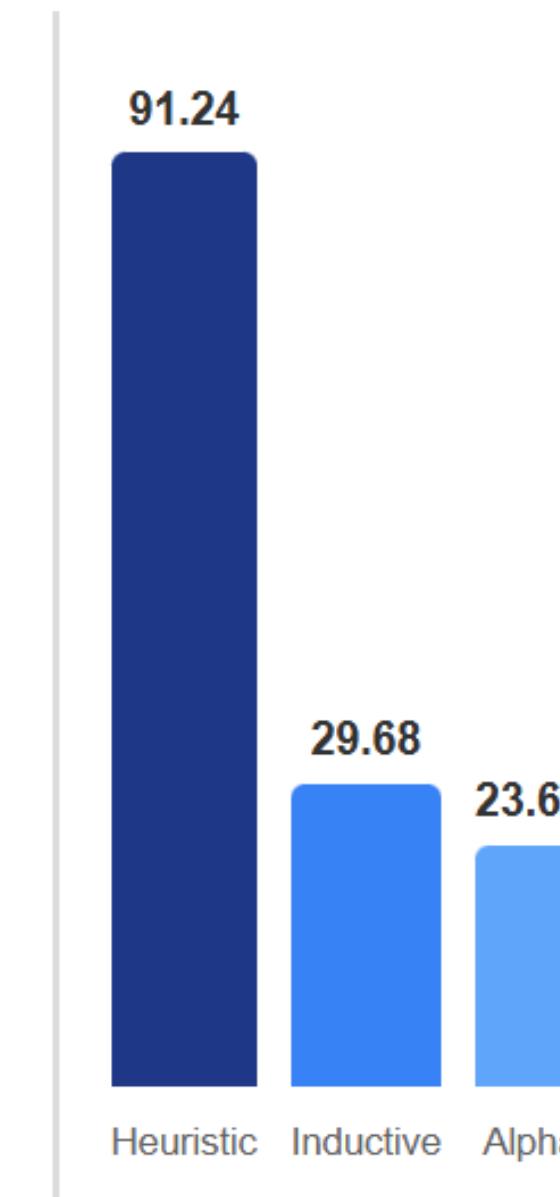
Fitness (%)



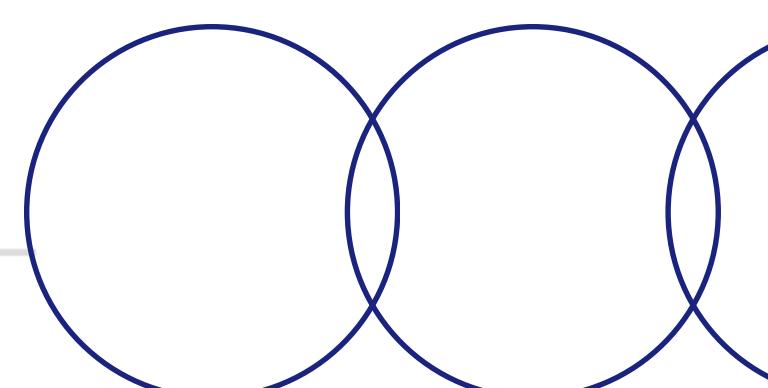
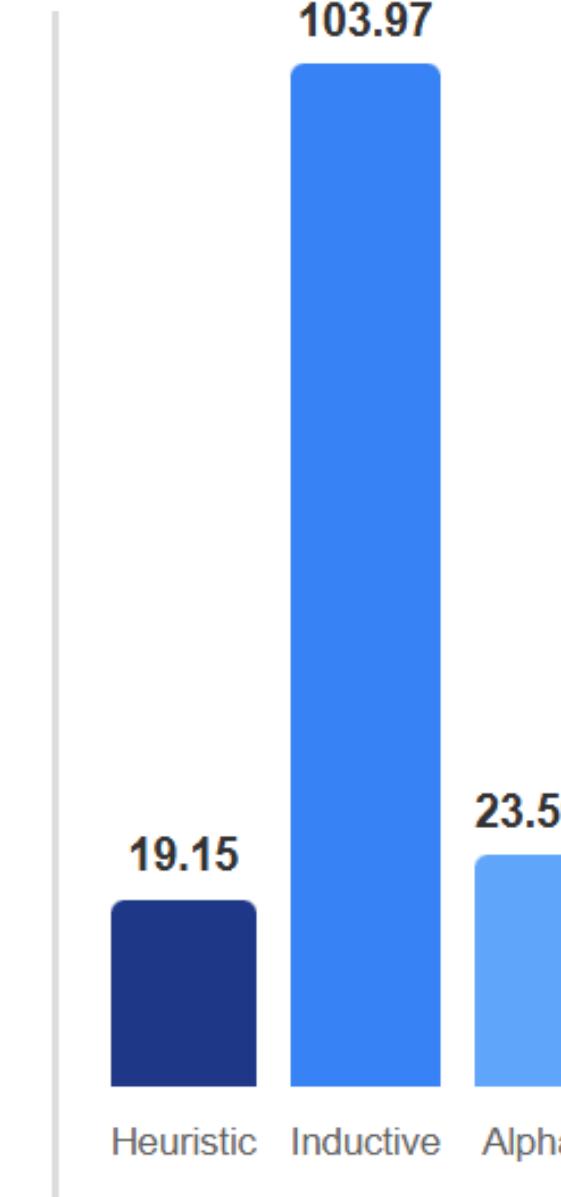
Precision (%)



F1-Score (%)



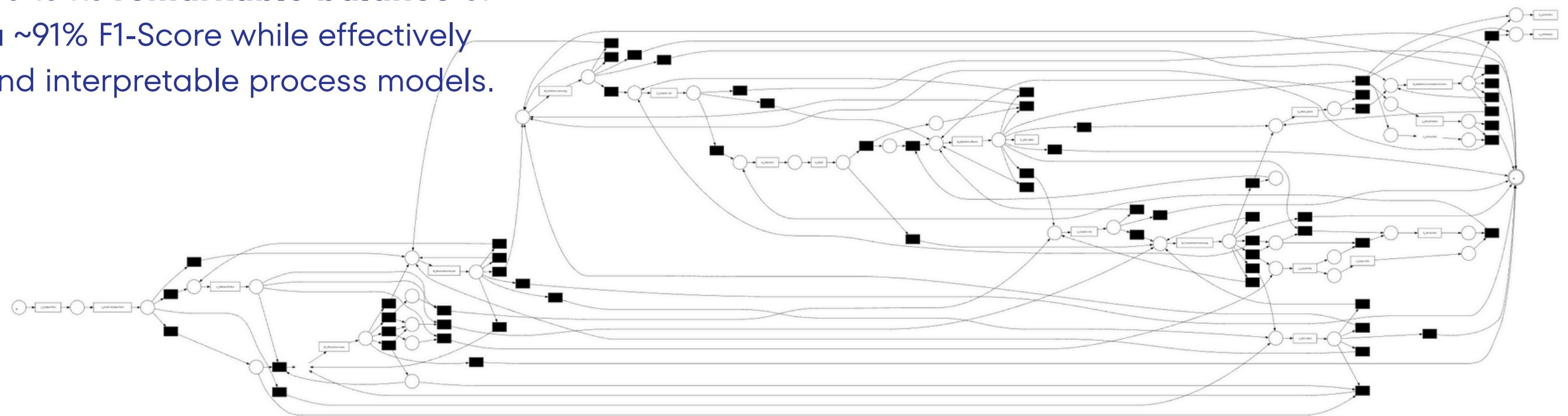
Calculation Time (s)



Best Model: Heuristic Miner

Optimal Choice for Analysis

The Heuristic Miner stands out due to its **remarkable balance** of fitness and precision, achieving a ~91% F1-Score while effectively filtering noise, resulting in clear and interpretable process models.



Dashboard Features

Built with Python

The dashboard is developed using Python, ensuring flexibility and robust performance for data processing and visualization tasks.

Visualization

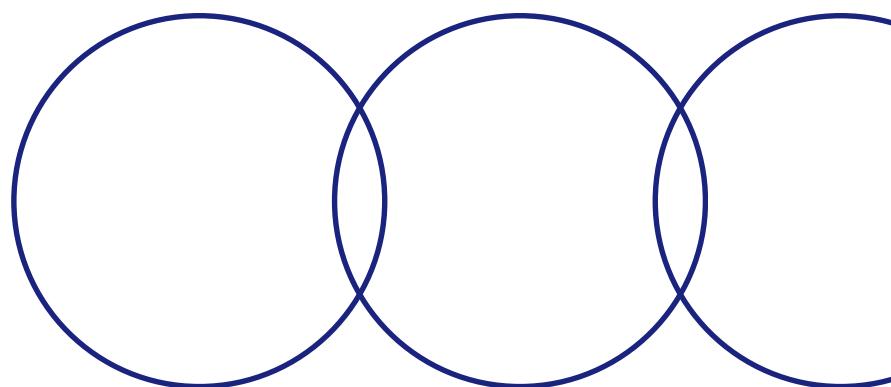
Visualization elements include zoomable Petri nets, providing clear insights into process flows and enabling effective analysis of variations.

Control Panel

An intuitive control panel allows users to dynamically filter data and adjust sampling settings for tailored analytics experiences.

Real-Time Analytics

The dashboard features real-time analytics with frequency charts, allowing users to monitor trends and changes in process performance seamlessly.





Control Panel

1. Caricamento Dati

Upload XES File

Drag and drop file here

Limit 200MB per file • XES, XML

[Browse files](#)

File Default rilevato

2. Ottimizzazione

Max Casi (Sampling)

13087

Filtro Varianti (Semplificazione)

1.00

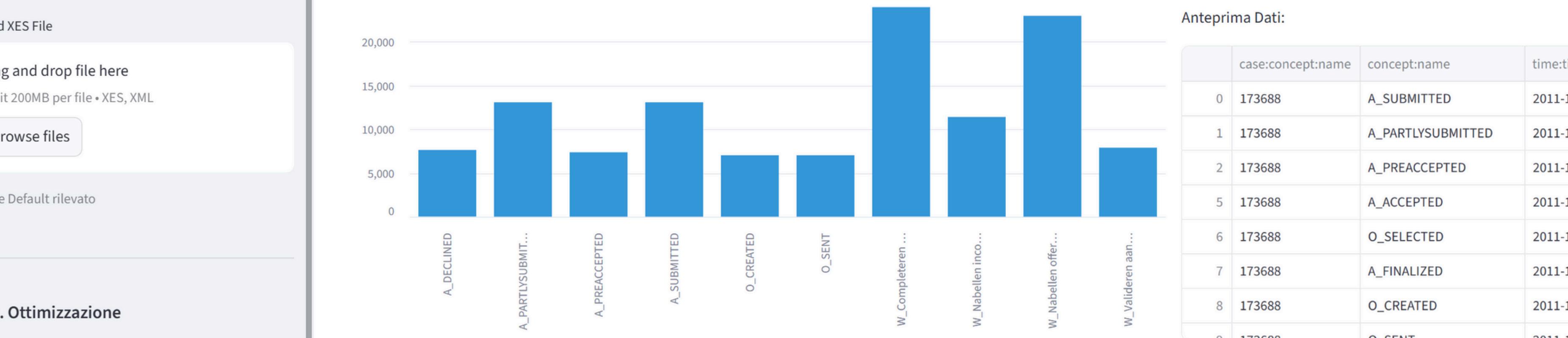
Totale Casi nel file: 13087

Status: Ready

Process Mining Dashboard | BPI Challenge 2012

[Stats](#) [Process Discovery](#) [Prediction](#) [Chat AI](#)

Analisi delle Frequenze



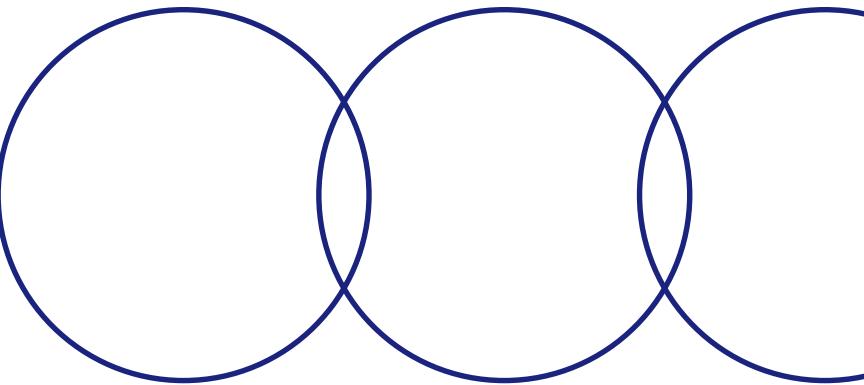
Anteprima Dati:

	case:concept:name	concept:name	time:ti...
0	173688	A_SUBMITTED	2011-1...
1	173688	A_PARTLYSUBMITTED	2011-1...
2	173688	A_PREACCEPTED	2011-1...
5	173688	A_ACCEPTED	2011-1...
6	173688	O_SELECTED	2011-1...
7	173688	A_FINALIZED	2011-1...
8	173688	O_CREATED	2011-1...
9	172600	O_SENT	2011-1...

AI Reasoning Integration

Utilizing Advanced Technologies

The integration of **Google Gemini API** enhances data analysis by extracting comprehensive statistics and generating insights, facilitating effective decision-making through automated context injection and AI-driven recommendations.



Gemini

AI Use Cases

Leveraging AI for Enhanced Financial Analysis Insights

Anomaly Detection

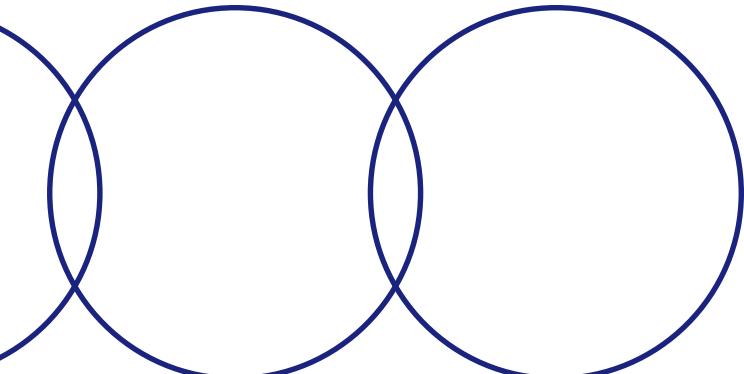
AI algorithms identify excessive variants in loan processes, flagging them as potential **process drift** that requires immediate attention for resolution.

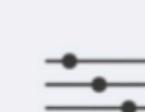
Optimization Suggestions

By analyzing repetitive manual steps, AI can recommend automation opportunities, significantly improving efficiency and reducing operational costs in the process.

Qualitative Summaries

AI translates complex Petri nets and process data into **easier-to-understand text**, providing stakeholders with clear insights into performance and operational trends.





Control Panel

1. Caricamento Dati

Upload XES File

Drag and drop file here

Limit 200MB per file • XES, XML

[Browse files](#)

File Default rilevato

2. Ottimizzazione

Max Casi (Sampling)

13087

Filtro Varianti (Semplificazione)

1.00

Totale Casi nel file: 13087

Status: Ready

Process Mining Dashboard | BPI Challenge 2012

Stats Process Discovery Prediction Chat AI

Consulente AI Interattivo



give me some optimization suggestions



Sulla base dei dati estratti dal log di processo (presumibilmente un processo di richiesta prestito/finanziamento), ecco alcune raccomandazioni strategiche di ottimizzazione:

1. Riduzione della Variabilità (Standardizzazione)

Il rapporto tra Varianti (4.336) e Casi (13.087) è estremamente elevato (~33%). Questo indica un processo "spaghetti", scarsamente standardizzato.

- Azione:** Identificare le "Happy Path" (le 10-20 varianti più frequenti) e analizzare le deviazioni. L'obiettivo è ridurre le eccezioni manuali che frammentano il flusso.

2. Miglioramento della Qualità in Ingresso (First Time Right)

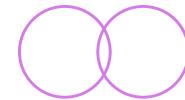
L'attività `w_Nabellen incomplete dossiers` (11.407 eventi) indica che quasi ogni caso richiede un ricontatto per documentazione mancante.

- Azione:** Ottimizzare l'interfaccia utente (UI/UX) durante le fasi `A_SUBMITTED` e `A_PARTLYSUBMITTED`. Introdurre controlli di validazione obbligatori e caricamento documenti guidato per minimizzare il lavoro di recupero manuale.

Chiedi un'analisi...

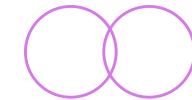


Conclusion and Deliverables



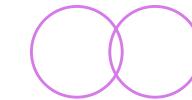
01

Analyzed 13,000 real financial loan cases for insights.



02

Validated Heuristic Miner as the most effective model.



03

Developed a hybrid GenAI-powered dashboard for enhanced analysis.

Thank you!

