

Illuminating Process Predictions: A Visual Analytics Framework for Business Workflows

Ujjwal Yadav
y_ujjwal12@yahoo.com
Independent Researcher

Abstract

As organizations increasingly use machine learning in business process management, the need for transparent predictive models becomes critical. Predictive process monitoring (PPM) techniques deliver accurate forecasts of process outcomes, remaining time and compliance violations. However, the black-box nature of many algorithms hinders their real-world deployment. Decision-makers in regulated industries want not only accurate predictions, but also understandable explanations for those predictions. VisInterPPM is a visual analytics framework introduced in this paper to help with the interpretability of predictive process monitoring. The structure involves predictive modeling through Random Forest classifiers, post-hoc explanation creation with LIME software, and visual-to-model mapping on Business Process Model and Notation diagrams. VisInterPPM allows analysts to comprehend which activities are responsible for predictions on a per-case and cohort level by mapping local feature influence weights onto familiar representations of the process. This framework allows analysis at the global and local levels. A thorough assessment with both synthetic and real-life event logs confirms the framework's performance. Professional assessments reveal greater understandability, usefulness, ease of use, speed of insight, and confidence in VisInterPPM visual explanations compared to standard bar charts. The findings confirm that adding visual context to the process flow significantly aids business stakeholders in understanding. This research proposes a human-centred practical approach to making predictive process monitoring more accessible, actionable and trusted.

Keywords

• Process Mining • Predictive Process Monitoring • Explainable AI • Visual Analytics • Business Process Management • Decision Support

1. Introduction

Due to advances in the field of business process management, some processes and data of organisations have become available [1]. Results of the method can be predicted using machine learning techniques because it is possible. Machine learning techniques can be used for prediction of completion of process, delay in process or compliance in the process [2]. These models, however, are often black-boxes which are created using these techniques. This indicates that while the models will provide a solution, they won't be able to justify it. Machine learning models are often characterized as a black-box because they do not elaborate on their predictions [3]. The application of black-box models hinders their implementation in sectors that have financial, legal, or operational consequences. In addition, business analysts do not require these models to be fully understood internally; rather, they need to understand the rationale behind predictions to make informed decisions [4].

Domain experts often find it difficult to interpret the results produced by modern algorithms. For example, a complex machine-learning prediction algorithm may produce explanations about the relevance of various input features in terms of importance weights, often presented as a bar chart or numerical scores. Yet these explanations are typically detached from the business process flow that analysts understand and work with daily. The mental model of a process analyst is grounded in the sequence of activities (e.g., "activity X leads to outcome Y"), whereas bar charts present explanations in terms of abstract feature weights (e.g., "activity X has weight 0.75"). It is therefore necessary to present explanations through well-known objects in Business Process Management, such as BPMN diagrams, to make them more accessible and actionable [5].

A pictorial representation has more impact than words; one should say. BPMN diagrams or business process model and notation diagrams is a common language used for specifying workflows that is easily understood by analysts, managers, and technical staff. By overlaying interpretability data onto the diagram, we can represent the machine's interpretation and reasoning in a way that directly relates to the human mental model of the process. Better alignment makes it easier to understand, communicate and make decisions with greater confidence. It streamlines the process [6].

Current predictive systems that use event logs data can be used to monitor and predict the performance of business processes to prevent failure. These ex-post explanations, especially visual analytics, offer little insight with regards to process mining methods. Thus, not just performance management, but also post-hoc explanation of predictive models to assist in root cause analysis and process diagnostics [7]. This information assists analysts, planners and managers in making decisions. The proposal presents a framework for visual analytics for post-hoc explanations of predictive process monitoring models. The data is a normal event log and a process modelling notation (BPMN) process.

The workflow inference framework has three major advantages. Firstly, this enhances the credibility of predictive systems. Visualizations clarify the underlying logic of classifiers by enhancing perceived transparency. Visual representations of these algorithms show how these algorithms get to these outcomes, in other words. In addition, it speeds up the root-cause analysis process. Identify the steps in the workflow that often result in positive (or negative) outcomes, identify the components of the workflow if the dancers do not perform well, and so on. Additionally, it empowers business professionals with the ability to directly generate actionable insights from predictions. In other terms, the framework translates complex model behavior into a visual language that directly relates to the process context. Such visual interaction bridges algorithmic output and human insight, a principle demonstrated in recent work on visual interaction with machine learning models [8].

The effectiveness of the framework is shown through a comprehensive evaluation with synthetic and real-life logs. The synthetic event log implies the travel insurance claim while the real-life event log refers to the loan application process run by a financial institute [9]. We demonstrate the applicability of our framework to binary as well as multi-class prediction problems in enterprise processes through these case studies. The visual explanation's usability, clarity and usefulness are evaluated using the feedback from domain experts, which serves as part of the evaluation strategy.

2. Related Work

This section situates VisInterPPM within the existing literature on predictive process monitoring, explainable AI, and visual analytics for business process management.

2.1 Predictive Process Monitoring

Predictive process monitoring (PPM) leverages event logs to anticipate future states or outcomes of ongoing process instances [10]. Early work in this area focused on next-activity prediction using sequence-based models [11]. Subsequent research expanded the scope to include outcome prediction, remaining time estimation, and anomaly detection [12]. Machine learning techniques commonly employed include Random Forests, Gradient Boosting, and neural network architectures such as Long Short-Term Memory (LSTM) networks, which have demonstrated strong performance on sequential event data [13]. Beyond process mining, predictive analytics has been leveraged to improve operational efficiency in domains such as biopharmaceutical market dynamics, where forecasting directly informs decision-making [14].

Despite their predictive accuracy, these models are often considered "black boxes" due to their complex internal mechanics [15]. This opacity poses significant challenges in business environments where decision-makers require justifications for predictions that carry financial, legal, or operational consequences. The interpretability gap between model predictions and actionable business insights has motivated a growing body of research on explainable predictive process monitoring [3]. Visualization has emerged as a promising bridge between predictive model behavior and human understanding, as shown in visual interaction approaches that surface model internals to analysts [8].

2.2 Explainable AI in Process Mining

Explainable Artificial Intelligence (XAI) has emerged as a critical research area aimed at making complex machine learning models more transparent and interpretable. Post-hoc explanation methods, such as LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations), provide local explanations by approximating complex model behavior with interpretable surrogate models [16].

These methods generate feature importance scores that indicate which input features most influenced a particular prediction. Within the process mining domain, several approaches have applied XAI techniques to explain predictive process monitoring outcomes. Recent work has explored the use of LIME and SHAP to identify which activities or control-flow features drive predictions about process outcomes [17]. However, these explanations are typically presented as tabular listings or bar charts, which, while informative, remain detached from the process context that business analysts understand and navigate daily [18].

2.3 Visual Analytics for Business Process Management

Visual analytics combines automated analysis with interactive visualizations to facilitate human reasoning and decision-making [19]. In the business process management context, visual analytics has been applied to process discovery, conformance checking, and performance analysis. BPMN (Business Process Model and Notation) diagrams serve as a standardized visual language for representing workflows, making them accessible to analysts, managers, and technical staff alike [1].

Several tools have integrated predictive analytics with process visualization. For instance, dashboards that overlay key performance indicators onto process models enable monitoring of process health [20]. However, these systems primarily focus on descriptive and diagnostic analytics rather than explaining the rationale behind predictive model outputs. The integration of post-hoc explanations directly onto process diagrams remains an underexplored area [21].

2.4 Gap and Contribution of VisInterPPM

While existing work has separately addressed predictive process monitoring, explainable AI, and process visualization, there is limited research that integrates these three pillars into a cohesive framework. The key limitations of current approaches include:

1. Explanations are presented in isolation from the process context, requiring analysts to mentally map feature importance scores back to specific process activities.
2. Aggregate explanations are often favored over instance-level interpretations, limiting their utility for case-specific investigations.
3. The visual presentation of explanations rarely leverages the familiar notation of BPMN diagrams, reducing accessibility for non-technical stakeholders [6].

VisInterPPM addresses these gaps by projecting local post-hoc explanations directly onto BPMN diagrams. This approach enables analysts to interpret predictions within the process flow, supports both individual trace and cohort-level analysis, and leverages a visual language that is already familiar to business stakeholders. By bridging the gap between predictive analytics and process understanding, VisInterPPM contributes a novel visual analytics framework that makes process predictions more interpretable and actionable. This gap is emblematic of a broader need in applied analytics: ensuring that data preparedness and analytical visualization genuinely serve the downstream human insight process rather than existing as disconnected technical artifacts [22].

3. Practical Implications for Business Process Analytics

VisInterPPM addresses several practical needs in enterprise process management. By making predictive model explanations more accessible, the framework supports a range of analytical workflows and decision-making scenarios [23].

3.1 Case-Level Investigation

Business analysts frequently need to understand why a specific process instance resulted in a particular outcome [24]. For example, a loan officer investigating a rejected application must determine which activities or decision points contributed to the rejection. VisInterPPM enables this analysis by overlaying the influence weights of each activity directly onto the process path for that specific case. This visual representation allows analysts to quickly identify which steps were most influential and assess whether the model's reasoning aligns with domain knowledge.

3.2 Aggregated Process Diagnosis

Beyond individual cases, analysts often seek to diagnose systemic issues within a process. By aggregating explanations across all instances that share a particular outcome such as rejected loan applications VisInterPPM reveals patterns at the process level. Activities that consistently exhibit strong positive or negative influence across a cohort highlight potential bottlenecks, quality issues, or process design flaws. This aggregated view supports root-cause analysis and informs process improvement initiatives. Aggregated visual diagnosis also benefits from perturbation-driven techniques that probe how internal representations respond to systematic changes, a strategy transferable to cohort-level process analysis [25].

3.3 Collaboration Between Data Scientists and Domain Experts

A recurring challenge in deploying predictive models in business settings is the communication gap between data scientists, who develop the models, and domain experts, who must act on their predictions. VisInterPPM facilitates collaboration by providing a shared visual vocabulary. Domain experts can interpret the visual explanations without needing technical expertise in machine learning, while data scientists can validate model behavior against business expectations. This alignment fosters trust in predictive systems and accelerates the adoption of analytics-driven decision-making [26].

3.4 Regulatory Compliance and Auditing

In regulated industries such as banking, healthcare, and insurance, organizations must often provide justifications for automated or algorithm-assisted decisions. VisInterPPM supports compliance by making the factors behind predictions transparent and auditable. The visual explanations serve as documentation that can be reviewed by internal auditors, regulators, or other stakeholders, demonstrating that decision-making processes are grounded in interpretable and reasonable criteria [5].

3.5 Integration with Existing Process Analysis Workflows

VisInterPPM is designed to complement existing business process management practices. The framework accepts standard event logs and BPMN models as inputs, making it compatible with established process mining tools and repositories. The visual outputs can be incorporated into existing dashboards, reports, or presentation materials, allowing organizations to augment their current analytical capabilities without requiring significant changes to their workflows.

3.6 Democratization of Predictive Analytics

By lowering the barrier to understanding predictive models, VisInterPPM contributes to the democratization of analytics within organizations. Business stakeholders with limited data science training can engage meaningfully with predictive outputs, ask informed questions, and contribute to data-driven decision-making. This broader engagement enhances the overall analytical culture of the organization and ensures that predictive insights are effectively translated into action [27].

4. Foundational Concepts

Process mining is the technique to extract process-related information from event logs that are maintained by information systems during business process execution. An event log refers to a case of a process in a company. For example, a specific customer order, a specific loan application [11]. A case consists of events a event refers to unique and single activity e.g. "Submit Application", "Verify Documents", etc. happening at a particular point of time performed by an organizational resource. A process model essentially represents a "process instance" in the sense of the possible activities that may be performed along with their control-flow dependencies. Formal representation of process models can be done using various notations like BPMN [2].

Predictive Process Monitoring leverages event data to anticipate the future states or the ultimate result of an ongoing case within a process. Outcome prediction is the most common PPM problem [12]. This is done by training a machine learning model on past completed process traces and predicting the final outcome for a currently executing case. For example, suppose a certain loan application will be accepted,

rejected, or cancelled. Will an insurance claim be paid soon or be delayed? Forecasts like these can help in taking early action to influence the 'case' more likely to result in an undesired outcome or next better manage business resources [23].

The predictive models, particularly Random Forests or Gradient Boosting machines, may be very accurate, but they are usually not interpretable. These models can deliver an accurate prediction on whether a unit would fail or not, but they do not explain how they got to that prediction - they are black boxes. When reviewing how AI is impacting business, we see the importance of having intelligent systems that are not only accurate but also interpretable, enabling business stakeholders to act on predictions with confidence [4].

In a business environment, since the cost of making such decisions is high, prediction is not sufficient. CEOs and analysts will want to understand not only what the model predicts, but also why it made that prediction, particularly when decisions have significant business implications [6].

Explainable AI is known for a set of techniques that makes it easy to comprehend the outcome of a complex model. Methods of post-hoc explanation utilize a model that has already been trained and whose predictions they seek to explain. LIME (Local Interpretable Model-agnostic Explanations) is one of the widely used post-hoc methods. For example, LIME can explain which input features were important for the model to make a given prediction for an instance. Even though LIME is quite capable, the explanation package it provides is (most often) a weighted list of features [16]. A business user is, for example, more interested in understanding which process activities contributed to a prediction than in examining raw feature importance scores.

Visual analytics techniques leverage the synergy of visual and computational representations to promote perception, cognition, and reasoning in analysis. Visual analytics can help users to explore predictions as well as explanations in PPM at the same time. It allows us to use visual representations of the process and prediction interactively to develop hypotheses, validate assumptions and understand insights that might not be obvious from textual or tabular explanations [19].

In order to address these challenges, we build on previous knowledge of process mining, machine learning and XAI approaches to form a comprehensive conceptual framework. The event logs and BPMN (Business Process Model and Notation) process models used for process mining data are known [1]. The machine learning capability that we mostly use allows us to make predictions on the outcome of a process. At last, the explanations of these predictions are made possible due to XAI approaches. Connecting the pillars of the conceptual framework is an important insight of the model.

5. Proposed Visual Analytics Framework

The proposed framework, termed the Visual Interpretability for Predictive Process Monitoring (VisInterPPM) framework, is designed as an end-to-end pipeline that transforms raw event data into interpretable visual overlays on process models. The architecture is modular, comprising four primary stages: Data Preprocessing and Feature Engineering, Predictive Model Training, Explanation Generation, and Visual Mapping. Each stage is detailed in the workflow depicted in Figure 1 and elaborated in the subsequent paragraphs [21].

The first step in the predictive analysis to predict next activities require event log processing to make patterns ready for event classification using machine learning. Categorical and temporal data make up raw event logs, and the attributes must be extracted in a feature vector format to train any classification algorithm. By far this is the most important step as feature quality can impact prediction. To add, the Categorical data such as activity and resource names were encoded. Next, control-flow characteristics,

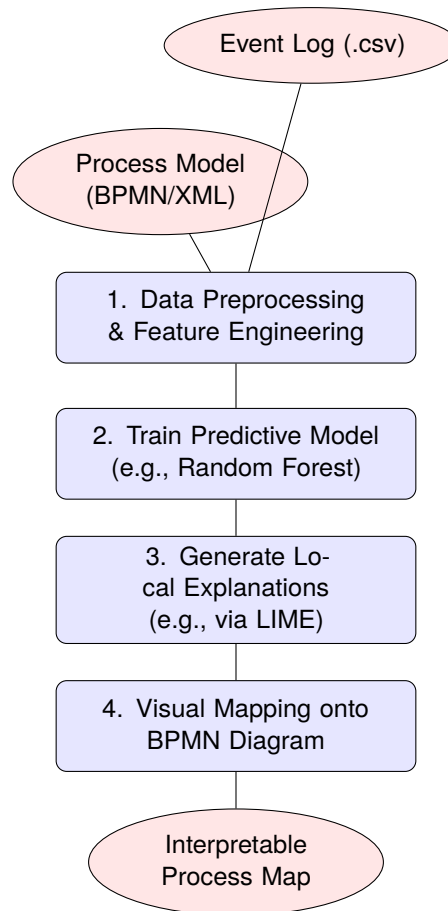


Figure 1: High-level workflow of the VisInterPPM framework.

such as the activities that occur before a certain moment in a process, were extracted. Ultimately, we create aggregate and time-based features such as the time elapsed since the onset of the case or the number of occurrences of a specific activity [9]. This emphasis on data preparation echoes findings that data preparedness and insightful visualization jointly determine the success of downstream analytical modeling [22].

The next step is to create a predictive model. Using the feature-engineered data, the algorithm becomes capable of predicting the outcome of a case of the process. The presented structure does not depend on the choice of a model. For this experiment, we chose a Random Forest classifier as this model is shown to be effective with tabular data and often used for this type of predictive problem [15]. The model is trained on past completed cases and evaluated based on standard metrics to ensure a good model before generating explanations.

The third stage focuses on interpretability. To generate an explanation of the outcome for a given case, we employ a local post-hoc explanation technique such as LIME. This approach approximates the behavior of the complex model in the vicinity of the case instance using an interpretable surrogate model (e.g., a linear model). The surrogate model subsequently explains the decision for that specific case by returning a weight for each feature—representing each process activity or derived attribute—indicating its contribution to the predicted outcome. Features with positive weights push the prediction toward the observed class, while features with negative weights push the prediction away from it [17].

The latest phase of visual mapping is a fourth phase of the model. The feature weighting list is assigned colors, with the resultant list projected on the BPMN process model that contains the classified

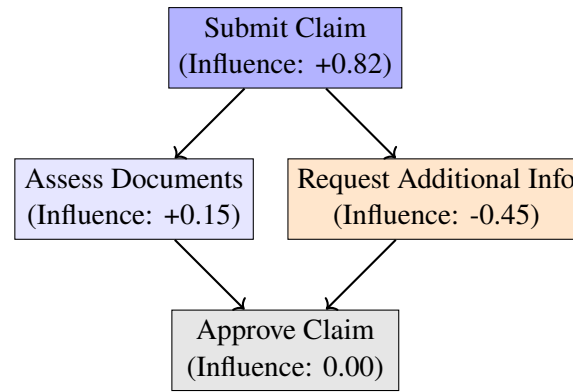


Figure 2: Example visual mapping for a single case. Activity colors reflect influence on a 'Rejection' prediction. Blue shades indicate activities supporting rejection, orange indicates activities reducing its likelihood, and gray shows a neutral activity.

activities. According to their impact weight, the activities are colored using a colorblind-accessible diverging palette. Activities with a strong positive influence on the predicted outcome are mapped to blue (or a blue-orange diverging scheme), while activities with a strong negative influence are mapped to orange. Activities that are not features or have weak impact are mapped to gray or light gray. This color scheme is designed to be distinguishable for readers with common forms of color-vision deficiency. Additionally, textual labels displaying the influence weights are provided for each activity to ensure redundant encoding of the information. The process model is such that the more visually prominent the activity is, the more effect it has on the prediction [1]. The design of this visual channel draws on broader lessons from time-series and machine learning visualization, where encoding choices strongly shape how quickly analysts extract actionable signals [28].

The framework supports two complementary analytical views: local and global. The local view, exemplified in Figure 2, provides an explanation for a single, specific process instance. This is crucial for case-specific investigations, such as understanding why a particular loan was denied. The global view aggregates explanation data across many instances (e.g., all rejected loans) to highlight patterns. Here, activities might be colored based on their average influence across the selected cohort, revealing which steps are consistently problematic or beneficial for a given outcome [18].

The technical implementation involves a Python-based pipeline. Libraries are used for process mining operations, machine learning, and explanation generation. The BPMN visualization is handled by libraries capable of parsing BPMN XML and rendering diagrams with programmatically styled elements. The final output is an interactive or static image where business experts can immediately perceive the "hotspots" and "coldspots" within their process that drive predictive outcomes, thereby closing the loop between data-driven prediction and actionable process knowledge.

6. Experimental Design and Datasets

To empirically validate the utility and effectiveness of the VisInterPPM framework, we conducted a structured evaluation using two distinct event logs. The selection criteria aimed to cover different process complexities and prediction scenarios [9]. The first dataset is a synthetically generated event log representing a controlled environment, while the second is a real-world log from a financial institution, presenting inherent noise and complexity. The characteristics of these datasets are summarized in Table I. Similar evaluation designs have proven valuable in adjacent application domains, where combining

Table 1: Characteristics of the Evaluation Datasets

Feature	Synthetic Log	Real-World Log
Domain	Insurance Claims	Loan Application
Source	Generated	BPI 2012
Cases	1,500	13,087
Events	9,842	164,506
Activities	9	24
Prediction	Binary	Multi-class
Balance	Balanced	Imbalanced
Perspective	Control-flow	Control-flow, Resource

analytical modeling with expert-facing visualization supports strategy and performance assessment [29].

A travel agency that handles health insurance claims is shown below in a synthetic event log. An event log was created from a BPMN model. One of the authors has used structured approach to created model so that control-flow patters are made clean and structured. Each case correspond to a singular claim. The activity sequence encompasses submission of a claim, verification of the claim and decision on approval or rejection of the claim. The label (outcome) can be either of the two modes approved or rejected [2].

The model's structuredness and the event log's low noise makes the log a perfect candidate for the first validation of the frameworks.

The real-life Log originates from a publicly available benchmark related to the loan application process of a Dutch financial institute. The most recent event log possesses greater complexity than the prior one which was utilized. In particular, the new event log contains a bigger number of activities and multiple resources, along with more variability in the case paths. The expected outcome is also a tri-class where application was approved, rejected or cancelled by the customer. The resulting class imbalance makes the event log challenge greater [10]. Aside from that, the application process of loan involves the cases' unstructured behaviour, which will test the framework ability in providing meaningful explanation under more realistic and messy setting.

Applying the framework to existing event logs requires minimal additional effort beyond standard data preprocessing. The event log was processed. After the extraction of events belonging to incomplete cases from the CSV file, prefixes of event sequences were extracted. For the world log, the categorical resource information was handled apart and the timedeltas between subsequent events were also retrieved. Feature vectors were then built for each of the case prefixes based on what activities had been performed up to that moment [13].

An intermediate local surrogate model was used to generate the explanation. For each test case, we simulated random perturbations of the case feature vector, querying the trained Random Forest to fit a linear model to the local variations of the prediction. The execution of this operation produces a weighting for each feature (activity), which that is the raw explanation data. Once the weights were filtered and normalised, the graph only shows the most influential activities per case to avoid cluttering the image [17].

The final step was visualising it. Each process model utilized a BPMN of the "reference" process. For a specific case study, the framework relied on the trace of that single case to identify one path in the BPMN model. Subsequently, the colour mapping described earlier was used to style each activity node on the path using the normalized influence weight from the explanation step. It resulted in a more interpretable process map. Hundreds of maps have been generated for both datasets which will be used in the following expert evaluation. The maps span diverse outcome classes and influence patterns.

Table 2: Aggregate Expert Ratings (5-point scale) for Explanation Formats.

Metric	Bar Chart	VisInterPPM	Improvement
Understandability	2.9	4.5	+1.6
Usefulness for Decision	3.1	4.4	+1.3
Ease of Use	3.4	4.6	+1.2
Speed of Insight	2.8	4.3	+1.5
Confidence in Explanation	3.0	4.2	+1.2
Average	3.04	4.40	+1.36

7. Results and Expert Evaluation

When assessing a framework, especially its performance beyond algorithms, we must adhere to its context of use. For instance, within our VisInterPPM framework, it would be odd to request the evaluation of algorithmic components since business-process experts will be the end-users. We will use a combination of quantitative methods, a survey, as well as qualitative interview. The eight experts were picked because they have experience with business process management, data science and the financial domain [26]. We have two evaluation cycles, both associated with the visualisation output of the framework after the adjustments. The synthetic insurance log is used to visualise the output of the first cycle.

The baseline explanation format was first shown to participants. It was a bar chart representing activity names, displaying their influence weights for a single-case prediction. Subsequently, the participants were presented with the VisInterPPM output, a graphical representation of the same explanation. With every explanation format, we posed participants the same fixed set of questions which aimed at measuring the understandability, usefulness and speed of insight.

Survey results were aggregated in a quantitative manner on a 5-point Likert scale across different dimensions. The results evidently show a preference for a visual and BPMN-based explanation. Table II summarizes the quantitative results in a nutshell. The VisInterPPM visualization consistently received much better ratings than the bar chart for understandability, usefulness for decision-making, and ease of use. All of the differences were statistically significant. What the experiment shows is that the explanation offering visual contextualization improves the interpretability experience significantly [16].

Data scientists also found the viz useful, but for other reasons. Experts pointed out that the bar charts show the exact weights; however, the process map added an extra layer of validation. In other words, the participants were able to reason each of the activity's members to be plausible in the situations. In principle, requiring the explanation to make sense in the process flow is a sanity check for the model or the artefact. The data scientists noted that the visual format helped to better reflect the process context and fostered improved dialogues with business people [23].

An interesting observation has been made at the global aggregated level. The visualization of an event log showed events colored by their average influence across all of the rejected cases. This allowed assessors to readily identify bottlenecks in the business processes as well as system-level issues. Take the loan application event log, for example, where one of the verification activities is strongly negative on average for approved cases. This suggests that doing that verification activity seems bad for getting cases approved. In a similar fashion, the correspondence activity that has a strong negative influence for approved cases in the loan application log indicates that to not have to perform that step helps achieve approval [12].

In the second evaluation cycle, the experts faced further extensive scrutiny with complex real-world

log. They verified that notwithstanding the data being much messier and the traces being much longer, the visual explanation remains entirely doable. While they can't make sense of everything at once, they can filter it so the view only contains the subset of activities that occur in a given case, e.g., a "Canceled" application.

It does not get overloaded with same info. Moreover, we also confirmed that the multi-class explanation was comprehensible. The specialists were able to correctly interpret the effect pattern of "Canceled" vs. "Rejected". The visual differences helped them understand that often deviate occur different outcomes [3].

Some useful feedback was collected which can make further improvements. Some experts wanted interactivity; for example, they wanted to be able to click on an action to see its exact numerical impact weight or filter to the cases where that action was especially impactful. Some respondents suggested additional applications of the time-based visual channels. For example, the width of an arrow could indicate the average delay associated with that path [20]. Recommendations like these show that our framework has the potential to serve as a more comprehensive interactive visual analytics system. The expert evaluation really strongly supports the overall thesis. If modeling interpretability is added to the process model, the usefulness will greatly increase and hence, the predictive process monitoring's adoption in business settings.

8. Conclusion and Future Work

Predictive analytics is being rapidly pressed into an operational business process management activity over the last few years. This process is called predictive process monitoring (PPM). On the basis of predictive modeling of business processes, this development can recommend impactful actions. Because impactful decisions require understanding the why and how the model prediction is obtained, human domain experts must have an understanding of PPM model rationale [11]. It poses a challenge as still complex PPM models need exact technical data science expertise.

Moreover, the expert assessment outcomes prove beneficial regarding the framework's practicality. More specifically, the participants judged the visual BPMN-based explanations significantly higher than the bar charts on all metrics considered (understandability, usefulness, ease of use, speed of insight, and confidence). Therefore, through experiments with a synthetic and a real-life event log, we validate the main hypothesis that visual contextualization is key to effective interpretability in the specific business process context [15]. The experiments conducted for both the binary and multi-class prediction problem illustrate the capability of the framework to produce intuitive visual explanations both at the single case level and at the level of overall process behaviour patterns.

Any firm looking to harness the power of A.I to improve their working processes can make use of this work. The platform first democratises predictive analytics for business processes through a low-code and engaging environment for business stakeholders [1]. Our framework's explanations allow data scientists and process experts to work together to validate the AI model predictions. They further support human-in-the-loop decision making by being transparent and reaching the right stakeholders by explaining the process at the right level of detail. A key benefit of our work is that it enables one to interrogate AI and justify its predictions from an ethics and AI perspective. Such interrogation and justification are important in regulated industries such as banking and healthcare [5]. Ultimately, our visual clarifications enhance.

An essential path for advancing research is the framework that expands in several ways. To begin with, we intend to extend the framework with the help of other processes. To illustrate the influence of resources (from an organizational standpoint), lanes in pools being colored can be visualised. Additionally, one can see how temporal features impact the performance perspective by altering visual features such as the

border width or the animation speed for delays. These will explain process outcomes more holistically if they are complementary and integrated.

The visualization might be enhanced with interactivity. A user should be able to interact with the dashboard to select a cohort of cases, for example all late deliveries, and see not only the aggregate influence diagram but also drill into specific cases (the actual shipments) and toggle parameters interactively, for example the explanation method and the classification threshold. Such interaction will help in explorative and investigative workflows [19]. It would change your framework from a reporting tool to an analytical workbench. Perturbation-based interaction methods for probing model internals offer one promising direction for making such workbench capabilities more systematic and reproducible [25].

We also want to experiment with different techniques for generating and visualizing explanations. While the current method is a viable option, alternative methods may achieve different tradeoffs in fidelity, stability, and cost. To demonstrate the outputs, we developed these visual metaphors we will test the other approach. The framework will get stronger both theoretically as well as practically [7]. We also want to apply the framework to additional industrial case studies from various sectors. The design principles and best practices embedded in the framework will help us to generalize them. Cross-domain applications, such as sports analytics and demand forecasting, provide useful testbeds for evaluating how well our visual explanation principles transfer beyond traditional enterprise processes [29].

VisInterPPM is thus a key step towards human-centric interpretable predictive process monitoring. An enabler for organizations to not just see the future for their processes, but also shape the future for their processes [12].

References

- [1] Abdul Azeem, Md Abul Kalam, and R. Md. Ai-enhanced process mining in business analysis: Driving operational excellence by smart insights. *American Journal of Engineering, Mechanics and Architecture*, 2(11):143–170, 2024.
- [2] W. M. Van der Aalst. Process mining: a 360 degree overview. In *Process mining handbook*, pages 3–34. Springer International Publishing, Cham, 2022.
- [3] P. Ceravolo, M. Comuzzi, J. De Weerd, C. Di Francescomarino, and F. M. Maggi. Predictive process monitoring: concepts, challenges, and future research directions. *Process Science*, 1(1):2, 2024.
- [4] D. Paschek. Business process management using artificial intelligence—an important requirement, success factor and business need for industry 5.0. 2020.
- [5] Harish Kasireddy. Responsible ai in fintech interfaces: Balancing innovation with ethical user experience design. *Journal of Data Analytics and Artificial Intelligence*, 2025. doi: 10.46632/jdaai/4/3/9.
- [6] I. Beerepoot, C. Di Ciccio, H. A. Reijers, S. Rinderle-Ma, W. Bandara, A. Burattin, et al. The biggest business process management problems to solve before we die. *Computers in Industry*, 146: 103837, 2023.
- [7] D. A. Neu, J. Lahann, and P. Fettke. A systematic literature review on state-of-the-art deep learning methods for process prediction. *Artificial Intelligence Review*, 55(2):801–827, 2022.

- [8] Aby Abraham. Bridging big data machine learning and human insight with visual interaction. In *IEEE*. URL <https://ieeexplore.ieee.org/author/549808137525098>. Transparent Boosting Tree (TBT).
- [9] S. K. Pradhan, M. Jans, and N. Martin. Getting the data in shape for your process mining analysis: An in-depth analysis of the pre-analysis stage. *ACM Computing Surveys*, 57(6):1–37, 2025.
- [10] C. Di Francescomarino and C. Ghidini. Predictive process monitoring. In *Process mining handbook*, pages 320–346. Springer International Publishing, Cham, 2022.
- [11] Mostafa Abbasi, Rahnuma Islam Nishat, Corey Bond, John Brandon Graham-Knight, Patricia Lasserre, Yves Lucet, and Homayoun Najjaran. A review of ai and machine learning contribution in predictive business process management (process enhancement and process improvement approaches). *arXiv preprint arXiv:2407.11043*, 2024.
- [12] Mostafa Abbasi, Rahnuma Islam Nishat, Corey Bond, John Brandon Graham-Knight, Patricia Lasserre, Yves Lucet, and Homayoun Najjaran. A review of ai and machine learning contribution in business process management (process enhancement and process improvement approaches). *Business Process Management Journal*, 31(4):1414–1452, 2025.
- [13] E. Rama-Maneiro, J. C. Vidal, and M. Lama. Deep learning for predictive business process monitoring: Review and benchmark. *IEEE Transactions on Services Computing*, 16(1):739–756, 2021.
- [14] Jennifer Joseph. Leveraging predictive analytics for enhanced biopharmaceutical market dynamics and operational efficiency. *J. Carcinog.*, 24(7s):628–634. doi: 10.64149/J.Carcinog.24.7s.628-634. URL <https://doi.org/10.64149/J.Carcinog.24.7s.628-634>.
- [15] W. Kratsch, J. Manderscheid, M. Röglinger, and J. Seyfried. Machine learning in business process monitoring: A comparison of deep learning and classical approaches used for outcome prediction. *Business & Information Systems Engineering*, 63(3):261–276, 2021.
- [16] Fethi Rabhi, Amin Beheshti, and Asif Gill. Editorial: Business transformation through ai-enabled technologies. *Frontiers in Artificial Intelligence*, 8:1577540, 2025. doi: 10.3389/frai.2025.1577540.
- [17] Pramath Parashar. Self-evolving ai workflows: A formalized feedback model for autonomous optimization. *International Journal of Emerging Research in Engineering and Technology*, 2025. doi: 10.63282/3050-922X.IJERET-V6I3P104.
- [18] E. Hasan. Big data-driven business process optimization: Enhancing decision-making through predictive analytics. *Authorea Preprints*, 2025.
- [19] Y. Pan and L. Zhang. A bim-data mining integrated digital twin framework for advanced project management. *Automation in Construction*, 124:103564, 2021.
- [20] M. Achouch, M. Dimitrova, K. Ziane, S. Sattarpanah Karganroudi, R. Dhouib, H. Ibrahim, and M. Adda. On predictive maintenance in industry 4.0: Overview, models, and challenges. *Applied Sciences*, 12(16):8081, 2022.

- [21] A. Kokala. Business process management: The synergy of intelligent automation and ai-driven workflows. *International Research Journal of Modernization in Engineering Technology and Science*, 6, 2024.
- [22] Tushita Agarwal. Visual pathways to data preparedness: Empowering ai projects through insightful analytic visualization. Technical report, TechRxiv. URL <https://doi.org/10.36227/techrxiv.175832480.02114963/v1>.
- [23] D. Martino, C. Perlangeli, B. Grottoli, L. La Rosa, and M. Pacella. A knowledge-driven framework for ai-augmented business process management systems: Bridging explainability and agile knowledge sharing. *AI*, 6(6):110, 2025. doi: 10.3390/ai6060110.
- [24] V. Nasir and F. Sassani. A review on deep learning in machining and tool monitoring: Methods, opportunities, and challenges. *The International Journal of Advanced Manufacturing Technology*, 115(9):2683–2709, 2021.
- [25] Eeshwar Pasula. Perturbation-driven visual analytics for probing internal representations in neural language models. *JPR*. doi: 10.67231/jpr65e97. URL <https://doi.org/10.67231/jpr65e97>.
- [26] A. Nguyen, Y. Hong, B. Dang, and X. Huang. Human-ai collaboration patterns in ai-assisted academic writing. *Studies in Higher Education*, 49(5):847–864, 2024.
- [27] M. Soori, B. Arezoo, and R. Dastres. Internet of things for smart factories in industry 4.0, a review. *Internet of Things and Cyber-Physical Systems*, 3:192–204, 2023.
- [28] Gnanendra Bachu. Visualizing energy demand: Enhancing power forecasting through time-series and machine learning techniques. In *IEEE*. URL <https://ieeexplore.ieee.org/document/11280613/>.
- [29] Pratik Dhameliya. Data-driven basketball strategy: Enhancing team performance through analytics and visualization. In *ICICDS*. doi: 10.1109/ICICDS70526.2026.11604669. URL <https://doi.org/10.1109/ICICDS70526.2026.11604669>.