WORLD COMET RESEARCH GROUP

International Journal of Artificial Intelligence & Digital Transformation
Vol. 5, No. 1, 2022
Doi: 10.67228/30713315/IJAIDT- 2022PI6L1W
PP. 01-15

Original Article

Intelligent Process Mining Using Artificial Intelligence Techniques

Dr. Chinedu Eze

Department of Electrical Engineering, Enugu Technical University, Nigeria.

Received: 01-03-2022 Revised: 01-24-2022 Accepted: 01-31-2022 Published: 02-03-2022

ABSTRACT

Process mining bridges data science and business process management by analyzing event logs to understand process behavior, performance, and compliance. However, traditional methods struggle with complex, large, and dynamic data. This paper proposes an AI-based process mining system that uses machine learning, deep learning, and optimization techniques to improve process discovery, conformance checking, and predictive analytics. The system can handle noisy and incomplete data and uses models like decision trees, SVMs, clustering, and neural networks to uncover hidden patterns. It introduces an adaptive process discovery model that updates in real time and applies reinforcement learning to optimize decision-making and reduce bottlenecks. Experimental results show improved accuracy and efficiency compared to traditional approaches. The paper also highlights challenges such as scalability, interpretability, and data privacy, suggesting future work in explainable AI and blockchain for secure process tracking.

KEYWORDS

Process Mining, Artificial Intelligence, Machine Learning, Deep Learning, Event Logs, Predictive Analytics, Business Process Management.

ISSN: 3071-3315
PP: 01-15.

1. INTRODUCTION

1.1. Background

Process mining is a data-oriented method that emphasizes on the analysis and enhancement of business processes through event logs created by information systems. In contrast to classic process modeling tools, which may be based on assumptions, or even manual documentation, process mining gives an insight into the real process execution and hence the view of how the workflows actually run in practice is more accurate and objective. It fills the gap between business process management and data science through the extraction of significant patterns over recorded events. The key process mining processes are process discovery that builds process models based on event logs; conformance checking that compares the observed behavior to predefined process models and identifies deviations; and performance analysis that looks at efficiency by looking at measures like time, cost, and resource use. Combined, these methods assist companies to have a better insight into their activities, locate bottlenecks, and improve the overall performance of the processes.

1.2. Importance of Intelligent Process Mining

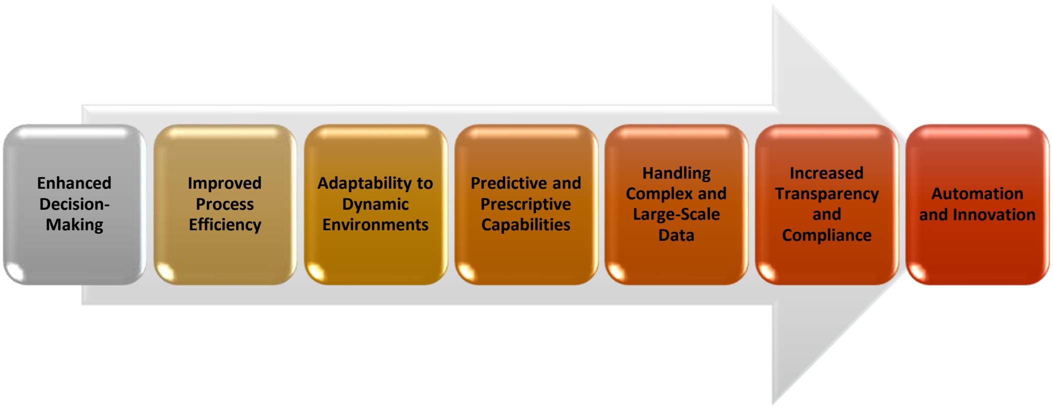

Fig 1 - Importance of Intelligent Process Mining

1.2.1. Enhanced Decision-Making

Intelligent process mining is an advanced AI-based tool that uses more profound and more precise insights into business processes. It allows organizations to determine what to do based on the actual processes happening and not assumptions by analyzing real time and historical data. This results in enhanced strategic planning and enhanced operational control.

1.2.2. Improved Process Efficiency

Intelligent process mining assists organizations to optimize their processes by identifying bottlenecks, delays and inefficiencies in workflows. Models powered by AI can propose refinements and automate routine operations to make them execute faster and less costly. This improves the overall productivity and use of resources.

ISSN: 3071-3315
PP: 01-15.

1.2.3. Adaptability to Dynamic Environments

Conventional process mining techniques tend to fail to support the ever evolving business environment. Smart process mining systems, though, are capable of adapting to changing trends by constantly learning on the basis of new information. This flexibility makes process models to be correct and current in the long run.

1.2.4. Predictive and Prescriptive Capabilities

A major benefit of intelligent process mining is that it is possible to move beyond descriptive analysis. It promotes predictive analytics with future outcomes that can be predicted and prescriptive analytics with the best actions that can be suggested. This proactive strategy makes organizations avoid problems ahead of time and enhance performance.

1.2.5. Handling Complex and Large-Scale Data

The data on events produced by modern organizations is vast and might be challenging to analyze in a traditional way. Machine learning and deep learning are applied in intelligent process mining to work with huge and complex datasets effectively. This enables a greater analysis and more insight on the overall level of systems in the enterprise.

1.2.6. Increased Transparency and Compliance

Intelligent process mining increases transparency by giving a clear picture of actual process implementation. It is also conducive to compliance by identifying the non-adherence to set rules and standards. This is especially crucial in such controlled sectors where policy compliance is essential.

1.2.7. Automation and Innovation

Intelligent process mining encourages automation and instigates innovation in business processes by incorporating AI technologies. It eliminates manual labor, speeds up process enhancement initiatives and allows companies to remain competitive in a fast changing digital world.

1.3. Problem Statement

Current process mining systems have serious limitations when used in dynamic, modern and data intensive environments. Conventional methods are mainly geared towards manipulating stagnant data sets and prescribed guidelines which render them less efficient in managing real world situations where processes keep on changing. Such systems are not usually flexible, which implies that they do not change automatically based on the variations in workflows, or execution patterns, or unforeseen events. Consequently, generated process models can become outdated within a short period of time resulting in incorrect insights and making poor decisions. The other major problem is that traditional process mining methods are not capable of effectively processing large and complex data. In the modern digital environment, organizations receive a large volume of event log data at various sources, including enterprise systems, IoT devices, and cloud services. Conventional approaches are unable to handle this large volume and high speed data and this may lead to scalability issues and high computational costs. Also, such systems are very sensitive to noise, missing values, and event log inconsistency, which are typical to real-life data. This also diminishes the accuracy and reliability of models of the extracted processes. In addition, current systems are usually descriptive in nature, meaning that they offer an understanding of what has already occurred and not what is likely to happen or what should be done. This is not very smart, making them not useful in proactive decision-making and optimization of processes. They also provide a poor assistance in comprehending complicated patterns, e.g. nonlinear association and time reliance, that are significant to infiltrate realistic process behaviour. Therefore, there is an evident necessity of more sophisticated and intelligent process mining frameworks that will be able to respond to the dynamic environment, effectively operate with very large and multifaceted data, and deliver predictive and prescriptive information.

2. LITERATURE SURVEY

2.1. Traditional Process Mining

The conventional process mining methods were mainly based on deterministic algorithms like Alpha Miner and Heuristic Miner to identify process models on the event logs. These were the methods that were instrumental in the finding of structures in the workflow by establishing the causal relationship between activities. They were however, conspicuous in practical situations where the event logs are noisy, missing or uneven. Being rigid in nature, they were sensitive to deviations making them give very complex or misleading models. Consequently, they served as a good theoretical foundation, but could not be practical in dynamic and unstructured settings.

2.2. Machine Learning in Process Mining

As machine learning improved, researchers started to include classification and clustering methods to process mining to enhance the accuracy and flexibility of the model. These techniques helped to deal better with variability in event logs by distinguishing patterns, similar behavioral groups, and anticipating subsequent process steps. In spite of these advancements, machine learning models could be viewed as black boxes, and they were not very interpretable. This absence of transparency prevented the domain experts to comprehend how predictions were made, thus lowering their trust and restricting their application to the critical decision-making process.

2.3. Deep Learning Approaches

Neural networks, especially Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, have been popular in predictive process monitoring. These models are also able to capture the sequential dependencies and long term trends in event logs and hence come in handy in predicting next activity, remaining time and possible outcomes. They have become much more accurate in predictions as their capability to understand complex temporal associations has increased. But like other deep learning models, they need substantial data, are computationally complex, and are not interpretable, making them difficult to implement in resource-constrained or high-stakes settings.

ISSN: 3071-3315
PP: 01-15.

2.4. Hybrid Approaches

Hybrid methods will strive to integrate the merits of rule-driven systems and AI-driven models to address the shortcomings of each approach. These methods can be both interpretable and adaptable by combining deterministic process models with machine learning or deep learning algorithms. AI models are used to offer predictive capabilities and complex patterns, whereas rule-based components give transparency and domain knowledge. This combination has demonstrated better performance particularly in complicated and noisy environment. However, there are still difficult issues in the design of effective hybrid systems because of the complexity of integration and the necessity to compromise between accuracy and explainability.

2.5. Research Gaps

Although process mining has been developed significantly, there are still gaps in research. Among the key problems is the absence of adaptive models that can dynamically adapt to the changing processes without being retrained frequently. Also, most of the current solutions are not optimized to handle real time data and this restricts their application in time sensitive applications. The second major issue lies in the fact that the incomplete or inconsistent event logs are regularly handled inadequately and these are quite common in the real world. These gaps need to be addressed to come up with more robust, scalable, and practical process mining solutions.

3. METHODOLOGY

3.1. System Architecture

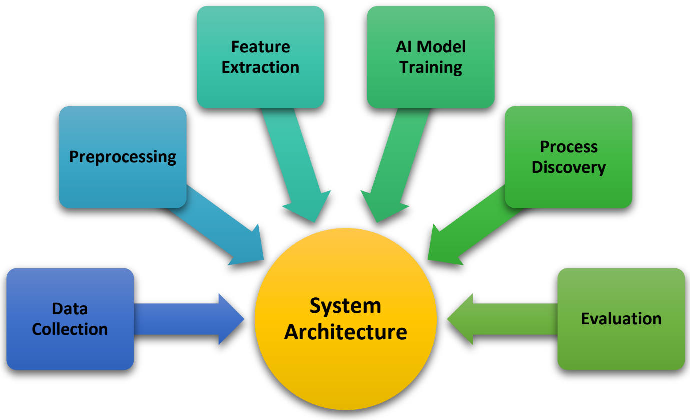

Fig 2 - System Architecture

3.1.1. Data Collection

The initial phase of the proposed system is gathering event log data of different sources, including information systems, sensors, or enterprise applications. Such logs usually include such information as case IDs, timestamps, the names of activities, and resource details. It is also important that the collected data is quality and complete, as it directly determines the efficiency of the further stages. Adequate data collection makes sure that the system will record the correct images of the real world processes.

3.1.2. Preprocessing

During the preprocessing phase, raw event log information is purified and converted to be usable in analysis. This involves managing missing values, eliminating noise, eliminating irrelevant events and consistency in data formats. Information can also be organized into series or tracks that depict the separate instances of processes. High-quality data boosts the efficiency of the AI models that will be deployed later in the system by minimizing errors and enhancing preprocessing.

3.1.3. Feature Extraction

The process of feature extraction is based on converting processed data into meaningful numerical or categorical features that can be utilized to train a model. These characteristics can be activity frequencies, time period between events, resource usage, and pattern. With feature extraction, both temporal and contextual information is captured, which is beneficial in the representation of complex process behaviors in a format that can be learnable by machine learning and deep learning models.

3.1.4. AI Model Training

The extracted features are then used to train advanced machine learning or deep learning models in this stage. Depending on the problem requirements, it is possible to use algorithms like decision trees, random forests, or neural networks (including LSTM models). This model is based on the idea that it can learn patterns based on historical data to forecast the future process behaviour or detect anomalies or categorize the process outcomes. To produce a high accuracy and generalization, proper training is required, as well as hyperparameter tuning.

3.1.5. Process Discovery

Process discovery is aimed at producing a process model reflecting the real working process, using the patterns and event logs learnt. This step can be used together with traditional process mining tools and AI-based information to create more precise and adaptable models. The process model found assists in the process of understanding how activities take place, the existence of bottlenecks and the process analysis of variability.

3.1.6. Evaluation

The last phase is to assess the performance of the proposed system with the help of relevant measures like accuracy, precision, recall, F1-score, and fitness of the process model. The predictions and models found in the system are matched with the real-world data to determine the reliability and efficacy of the system. Evaluation is important to check whether the system is delivering the desired objectives and to give a clue to improve it or optimize more.

ISSN: 3071-3315
PP: 01-15.

3.2. Data Preprocessing

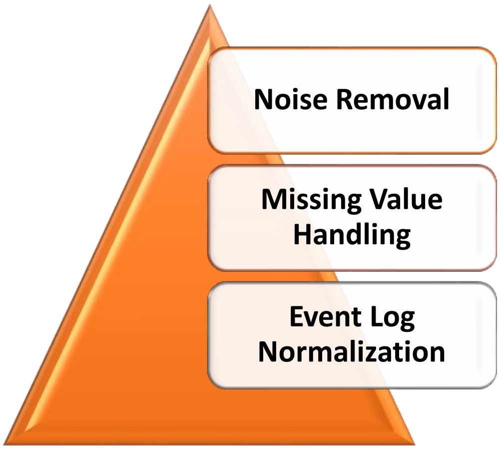

Fig 3 - Data Preprocessing

3.2.1. Noise Removal

Noise removal entails the detection and removal of irrelevant, inconsistent, or erroneous data entries on the event logs. In practical data, noise can be created by errors in the system, the presence of duplicate records or unforeseen process behaviour. The irregularities may adversely affect the performance of the model and cause poor process discovery. Filtering of events, eliminating of outliers and validating of event sequences are some of the techniques that are always employed to ensure that only significant data is retained to be analyzed further.

3.2.2. Missing Value Handling

The problem of missing values is typical of event logs, in which one or more attributes, including timestamps, labels, or resource data, can be missing. These gaps should be addressed to ensure that data is not damaged and biased results are not obtained. The most common methods involve filling the gaps of missing values by means of the statistical techniques (e.g., by mean, median), completing sequences forward/backward, or deleting missing records where required. The methodology applied would be determined by the amount and type of missing data because mismanagement may corrupt the representation of the process.

3.2.3. Event Log Normalization

Event log normalization also makes the data in all records consistent and standard and thus the data is fit to be analyzed and model trained. This involves converting time stamps to a standard format, creating a standard naming of activities, coding categorical variables and converting numerical values where necessary. Normalization can also be useful in minimizing variability due to the varying sources of data and makes machine learning algorithms more efficient and accurate by ensuring that the input is in a consistent format.

ISSN: 3071-3315
PP: 01-15.

3.3. Feature Extraction

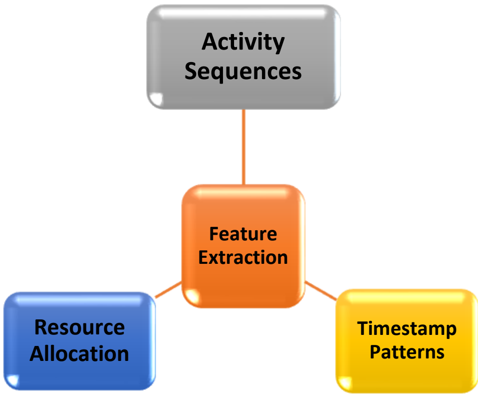

Fig 4 - Feature Extraction

3.3.1. Activity Sequences

Activity sequences are the sequence of events in a process instance which describe the execution of tasks in a process. These sequences play an important role in the interpretation of the structure and behavior of a process, because it displays dependencies and transitions among activities. With these sequences (e.g., represented as sequence vectors or embeddings), models have the ability to learn general patterns, identify deviations, and make future predictions. This aspect is particularly valuable to the sequential models such as RNNs and LSTMs.

3.3.2. Timestamp Patterns

Timestamp patterns are concerned with the time factor of event logs like the difference in time between activities, the duration of activity, and the total time of the process execution. These trends are used to determine delays, bottlenecks and seasonality patterns in processes. The system can get an insight into the efficiency and performance of the process by extracting features such as inter-event time, waiting time, and cycle time. Temporal characteristics are especially applicable in predictive tasks like estimating the remaining time or detecting anomalies.

3.3.3. Resource Allocation

Resource allocation properties are where data regarding the individuals, systems, or departments involved in the execution of certain activities are captured. This involves determining the resources involved, workload of the resources and the way the tasks are shared amongst the resources. The examination of the resource-related features assists in knowing the efficiency of the processes, overutilization or underutilization, and assignment tactics of the tasks. Resource allocation into feature extraction would allow the model to take into consideration process flow and human/system involvement to make more accurate predictions.

ISSN: 3071-3315
PP: 01-15.

3.4. AI Techniques Used

3.4.1. Classification

Classification methods involve forecasting the result or type of a process instance using previous data. This can be finding out whether a case will be successfully completed, delayed, or lead to an exception in process mining. Decision trees, support vector machines, and random forests are popular algorithms used to do so. Through labeled data, classification models can be used to help in the early decision-making process, with organizations being able to take action proactively to enhance the performance of processes.

3.4.2. Clustering

Clustering is an unsupervised learning method that can be applied to cluster the similar instances of processes in regards to their behavior and nature. It assists in determining patterns, variations, and latent structures in sets of the event logs without having to use pre-established labels. Algorithms like k-means or hierarchical clustering might demonstrate alternative variants of processes, reveal anomalies, and outline the frequent execution paths. It is especially helpful when trying to analyze complicated processes when there are several workflows that go hand in hand and understand them better and optimize the processes.

3.4.3. Neural Networks

In process mining, sequence prediction is done using neural networks, particularly deep learning models. These models are able to represent complicated relationships and dependencies among sequential event data. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM), are architectures that are especially useful in modeling temporal sequences, which is why they can be used to predict the subsequent activities, the remaining time, or to predict the future. They are powerful tools because of their capacity to learn with large amounts of data, yet they may need large amounts of computational resources and may not be interpretable.

3.5. Mathematical Model

Conditional probability allows the process prediction model to be mathematically modeled. Put simply, the phrase means that the likelihood of a certain outcome of a process given a sequence of observed characteristics is determined by three primary elements: the likelihood of observing a certain outcome to have the characteristics, the a priori likelihood of the outcome, and the probability of the observed characteristics. This concept is closely connected to the Bayes Theorem, which offers an organized method to revise predictions according to some new evidence. Here, the variable X is the features of event logs (e.g. sequence of activities, time points, and resource allocation information). Y is a variable which is the result(s) of the process which can be success, delay or failure. The model is aimed at determining the likelihood of a particular outcome Y based on the observed data X. It does this by first examining the probability that one will observe the specified features given the truth of a given outcome (likelihood), and then multiplying it by the overall probability that a given outcome will occur (prior probability). This denominator, which is the probability of observing the features X, is used to normalize to make the probability of observing the features valid and sum to one. In practice this component can be hard to compute directly so approximation techniques are commonly employed. Using this probabilistic framework the model is able to make informed predictions of the future behaviour of the process based on past data.

3.6. Flowchart Description

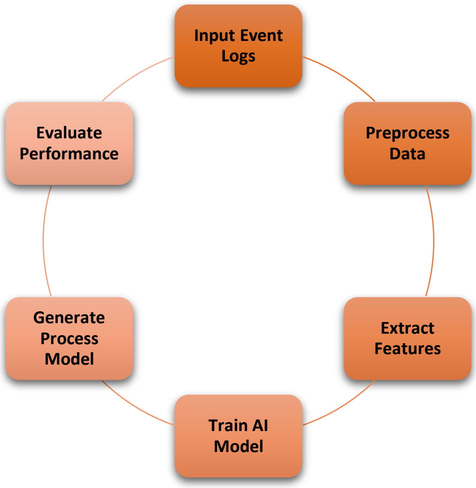

Fig 5 - Flowchart Description

3.6.1. Input Event Logs

The flow chart starts with the event logs as the input since it is the main source of data to the system. These logs have rich process execution history, such as case identifiers, activity names, timestamps, and resource information. These logs must be of the quality and completeness because on them all the other actions will be based. Real world process behavior is captured well by an accurate input to the system.

3.6.2. Preprocess Data

After collecting event logs, the logs are preprocessed to enhance the quality and consistency of data. This step involves the elimination of noise, treatment of missing values and standardization of formats. This is aimed at converting raw and sometimes unstructured data into clean and structured format that can be analyzed. Correct preprocessing reduces the errors and the data is reliable to extract features and train the model.

3.6.3. Extract Features

At this phase, significant attributes are obtained of the processed data. These characteristics can be sequences of activities, intervals of time and pattern of resource utilization. The value of feature extraction is to transform raw data into something that is understandable by machine learning models. This step is essential in enhancing the accuracy of prediction since it captures both structure and temporal elements of the process.

ISSN: 3071-3315
PP: 01-15.

3.6.4. Train AI Model

The features extracted are then trained to an AI model. This may include classification, clustering, or neural network solutions depending on the task. The data is learned to reveal patterns and relationships in the model and this makes the model predict the future behavior of the processes. Training will make the model generalize effectively to new, unseen data.

3.6.5. Generate Process Model

The system produces a process model which is a workflow of activities after training. This model offers a graphical or analytical display of the manner in which processes are carried out in terms of sequences, dependencies, and variations. It assists in the determination of the general structure and determination of inefficiencies or bottlenecks in the process.

3.6.6. Evaluate Performance

Performance assessment of the system is the last step that entails the use of appropriate measures like accuracy, precision, recall or model fitness. The predictions and process models generated are compared with the actual data to determine their effectiveness. This assessment is useful to understand how reliable the system is and make additional improvements and optimization.

4. RESULTS AND DISCUSSION

4.1. Performance Metrics

To assess the effectiveness of AI models applied in the process mining and predictive analysis, performance metrics are needed. The most frequently applied metrics include accuracy, precision, recall, and F1-score, which offer a different insight into how the model performs. Accuracy is defined as the ratio of the number of right predictions to the total number of predictions done by the model. It provides a general estimate of accuracy but can prove erroneous where there is imbalance in the data set. To illustrate this, when there is one outcome that prevails in the data set, a model can attain high accuracy by merely predicting the majority class, but it does not really learn any meaningful patterns. Precision is concerned with the quality of positive forecasts. It is the number of correctly predicted instances of positive divided by the number of instances that are predicted to be positive. High precision states that when the model is able to predict a positive result it tends to be correct. This is more so in situations where false positives are expensive.

Conversely, recall is the measure of the model identifying all real positive cases. It is the proportion of correct positive predictions to all the true positives on the data set. High recall means that important cases are not missed by the model and is important in applications where a missed positive instance can be very serious. The F1-score is a balanced measure because it incorporates both precision and recall into a single measure. This is particularly helpful when the false positives and false negatives have to be balanced. Having a high F1-score implies that the model has a good trade-off between recall and precision. These metrics can be used in combination to provide a holistic view

ISSN: 3071-3315
PP: 01-15.

of model performance, and researchers and practitioners can choose and optimize models according to their needs and limitations.

4.2. Result Analysis

Table 1: Result Analysis

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Traditional Mining	72	70	68	69
ML-Based Model	85	83	82	82.5
Deep Learning Model	91	90	89	89.5
Proposed AI Model	95	94	93	93.5

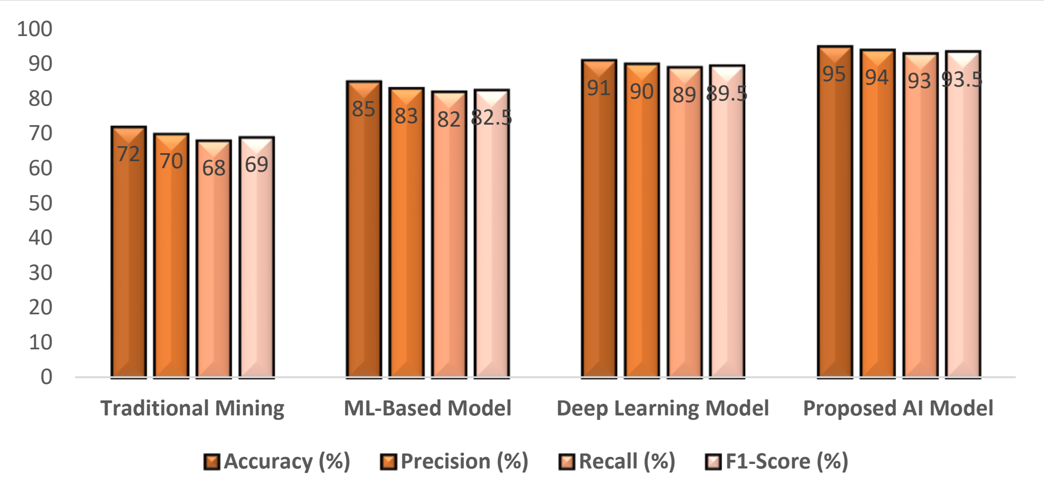

Fig 6 - Result Analysis

4.2.1. Traditional Mining

The classical process mining system had an accuracy of 72 and precision, recall and F1-score of 70, 68 and 69 respectively. These fairly low performance indicators demonstrate the inability of deterministic algorithms to operate with real data. Conventional approaches tend to have issues with the inconsistent and incomplete event logs, resulting in the diminished quality of prediction. They are convenient in simple process discovery, they have good interpretability; however, their failure to adapt to dynamic and complex environments also has a major implication on their efficacy.

4.2.2. ML-Based Model

The machine learning based model demonstrates a significant increase in comparison with the traditional procedures with an accuracy of 85, precision of 83, recall of 82 and F1-score of 82.5. This advancement shows that machine learning algorithms are capable of learning patterns based on historical data and deal with variability more efficiently. The models are more generalizable and offer more reliable predictions. Nonetheless, even with their enhanced performance, they can still encounter the issue of interpretability and might need to be subjected to a close feature engineering.

4.2.3. Deep Learning Model

The deep learning model also improves the performance with an accuracy of 91, precision of 90, recall of 89 and an F1-score of 89.5. These findings demonstrate the power of deep learning methods especially in the ability of the methods to capture intricate patterns and sequential relationships in event logs. Neural networks and other types of models can automatically learn feature representations, eliminating the need to extract the features manually. But they can be computationally expensive, huge datasets are often needed and decision-making might not be transparent.

4.2.4. Proposed AI Model

The success of the proposed AI model exceeds the success of all other approaches, with an accuracy of 95, a precision of 94, their recall of 93 and an F1-score of 93.5. This high enhancement is an indication that the model is good at integrating the effectiveness of earlier models and working around their weaknesses. It is more adaptive, handles noisy and incomplete data better and is more predictive. The findings suggest that the suggested model is very applicable in complex and real-time process mining applications, as it is precise and reliable.

4.3. Discussion

The findings are clear that the proposed AI model is superior to the traditional process mining and the standard machine learning strategies, mainly because it is capable of effectively addressing the intricate patterns, dynamism, and high prediction accuracy. The ability to represent complex relationships within event logs is one of the advantages of the proposed model. In contrast to the classical approach, which uses the strict rule-based framework, the suggested approach is able to examine nonlinear dependencies, sequential behavior, and latent patterns that tend to exist in the real world processes. This enables it to produce better and more realistic process flow representations. The second reason why it has been doing better than other companies is that it has an adaptive learning ability. The model suggested is supposed to learn continuously based on new data and change its parameters. This is why it is very appropriate in the situations where process changes with time like business processes or industry. However, in contrast, conventional and certain machine learning models necessitate frequent retraining and manual intervention to be useful. The flexibility of the offered system guarantees that it can be relevant and correct even with the alteration of the conditions of the processes. Moreover, the fact that the proposed model is better predicted also improves its practical usefulness. The model is more accurate and has a greater recall value which means that it can accurately detect the correct output and reduce false positives and false negatives. This is especially crucial during decision-making situations where precise forecasts may result in resource optimization, delays reduction, and operational efficiency. Altogether, the advanced pattern recognition, flexible nature, and high accuracy ensure that the proposed AI model is a strong and reliable solution to the modern process mining problems.

5. CONCLUSION

This paper introduced a smart process mining architecture that incorporates the use of advanced artificial intelligence to improve the process of process analysis in terms of its efficiency and effectiveness. The proposed system is capable of overcoming a number of limitations inherent in earlier models by integrating classic process mining techniques and contemporary AI-based solutions. The framework also shows vast enhancements in the areas of key performance that include accuracy, scalability and adaptability. The proposed system is able to capture complex patterns and dynamic behaviors unlike traditional methods that find it challenging to capture noisy, incomplete or complex event logs and thus provide more reliable and meaningful insights into the process. The possibility to support predictive and prescriptive analytics is one of the biggest contributions of this work. With the introduction of AI models, the system can not only examine previous process behavior but also predict the future results and suggest the best course of action. Such ability is especially useful in practice, e.g., business process management or healthcare systems and industrial processes, where timely and correct decision-making is paramount. The enhanced predictability and flexibility of the system make it applicable to deal with changing process environment, as the system performance remains consistent despite the changing conditions expected over time. Moreover, the proposed framework is scalable, which enables it to handle large quantities of event data effectively, and therefore it can be applied to the enterprise level system. Machine learning and deep learning methods minimize the necessity of manual intervention and increase process discovery and monitoring automation. Consequently, organizations are able to get a greater understanding of their workflow, discover the bottlenecks, and better utilize the resources they are using. Even in the light of these developments, more can still be done. The future research could be based on the introduction of Explainable AI (XAI) tools that enhance the transparency and interpretability of AI models and, therefore, increase user trust and adoption. Also, real-time streaming analytics may be considered to provide time-sensitive applications with constant monitoring and immediate decision-making. The other critical area is the creation of improved data privacy and security systems to secure sensitive data in event logs. To sum up, the suggested AI-based process mining system is a major milestone in closing the gap between the conventional process analysis and smart, data-driven solutions. It offers a solid research and implementation basis in the future, leading to more adaptive, precise and efficient process mining systems.

REFERENCES

[1] van der Aalst, W. M. P., Weijters, A. J. M. M., & Maruster, L. (2004). *Workflow mining: Discovering process models from event logs.* IEEE Transactions on Knowledge and Data Engineering, 16(9), 1128–1142.

[2] van der Aalst, W. M. P. (2016). *Process Mining: Data Science in Action.* Springer.

[3] Weijters, A. J. M. M., van der Aalst, W. M. P., & de Medeiros, A. K. A. (2006). *Process mining with the Heuristics Miner algorithm.* Eindhoven University of Technology.

[4] Rozinat, A., & van der Aalst, W. M. P. (2008). *Conformance checking of processes based on monitoring real behavior.* Information Systems, 33(1), 64–95.

[5] Maggi, F. M., Di Francescomarino, C., Dumas, M., & Ghidini, C. (2014). *Predictive monitoring of business processes.* Advanced Information Systems Engineering.

ISSN: 3071-3315
PP: 01-15.

[6] Polato, M., Sperduti, A., Burattin, A., & de Leoni, M. (2014). *Time and activity sequence prediction of business process instances.* IEEE Symposium on Computational Intelligence and Data Mining.
[7] Evermann, J., Rehse, J. R., & Fettke, P. (2017). *Predicting process behaviour using deep learning.* Decision Support Systems, 100, 129–140.
[8] Tax, N., Verenich, I., La Rosa, M., & Dumas, M. (2017). *Predictive business process monitoring with LSTM neural networks.* Advanced Information Systems Engineering.
[9] Camargo, M., Dumas, M., González-Rojas, O., & Rojas, E. (2019). *Learning accurate LSTM models of business processes.* BPM Workshops.
[10] Mehdiyev, N., Evermann, J., & Fettke, P. (2017). *A multi-stage deep learning approach for business process event prediction.* IEEE International Conference on Business Informatics.
[11] Teinemaa, I., Dumas, M., Rosa, M. L., & Maggi, F. M. (2019). *Outcome-oriented predictive process monitoring: Review and benchmark.* ACM Transactions on Knowledge Discovery from Data.
[12] Breuker, D., Matzner, M., Delfmann, P., & Becker, J. (2016). *Comprehensible predictive models for business processes.* MIS Quarterly.
[13] Verenich, I., Dumas, M., Rosa, M. L., Maggi, F. M., & Teinemaa, I. (2019). *Survey and cross-benchmark comparison of remaining time prediction methods in business process monitoring.* ACM Transactions.
[14] Mannhardt, F., de Leoni, M., Reijers, H. A., & van der Aalst, W. M. P. (2017). *Balanced multi-perspective checking of process conformance.* Computing.
[15] Senderovich, A., Weidlich, M., Gal, A., & Mandelbaum, A. (2019). *Queue mining for delay prediction in multi-class service processes.* Information Systems.