ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

ENHANCING THE PROCESS EFFICIENCY OF RPA USING MACHINE LEARNING

Kiran Fatima
Department of Data Science,Superior University,Lahore, Pakistan
i.mkiranali19@gmail.com
Kisra Zafar
Department of Data Science,Superior University,Lahore, Pakistan
kisrazafar@gmail.com
Nadia Shareef
Department of Data Science,Superior University,Lahore, Pakistan
nadiashareef326@gmail.com

Abstract:

The robotic process automation (RPA) has completely changed commercial processes by automating repetitive procedures and reducing human error. This study investigates how machine learning (ML) can be integrated into RPA to improve it. The paper talks about the challenges of picking the right processes for automation and exposes the advantages of RPA, like the ability to imitate human-computer interactions. It argues that decision-making, predictive maintenance, and anomaly detection can be done better by RPA systems when techniques including decision trees as well as random forests from ML are used. According to the study, it is evident that the classification algorithms with the highest accuracy were the Gradient Boosting Machine (GBM) and Random Forest Classifier (RCF) that is 98%, while the others had a lower accuracy of Decision Forest Classifier (DCF) 97%, Random Regression Forest (RRF) 93%, and lastly the Decision Random Forest (DRF) 87%. Machine learning through sophisticated technologies such as optical character recognition and natural language processing improves data extraction, processing accuracy, and user engagement. According to the findings, combining RPA with ML leads to an automated system that is more flexible, resilient and intelligent which can greatly enhance operations in a variety of industries. This comprehensive approach taken by this research will be useful for both academicians and practitioners in developing and optimizing RPA systems, finally advancing intelligent automation.

Keywords:

Robotic process automation (RPA), machine learning (ML) and intelligent process automation (IPA).

Introduction:

The software runs instead of human beings in most cases[1], and this is a game-changer in recent times for Robotic Process Automation [2]. Corporations across the globe operating in various fields use RPA to gain an edge over their competitors when it comes to automating business processes. Increasingly, on-premise environments are being improved with automation toolsets [3]. It is more important to have well-established standards within the companies that will help them benefit from this kind of technology despite the increase in its spread and adoption rates among the companies[4],[5]. Determining which procedures can be automated through RPA is one of the main challenges. Correct choice of process is essential, since if we automate a wrong one it means that failure becomes much faster while increasing inefficiency at the same time. Establishing criteria for evaluating whether or not a process lends itself to RPA will aid in choosing the best one [6]. Since RPA employs software automation technology [7]that acts like human "activity" on a computer. The mouse and keyboard are controlled by the pre-programmed execution script to finish the "human-computer interaction" simulation. Tasks classified as "personal assistant" or "unattended" fall within the general category of RPA. The objective of the so-called "personal assistant" is to automate manual interface operations, such as operating the mouse, keyboard, message, etc, and to mimic the interactive behavior of the human program interface (UI) [8]. Utilizing the application

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

interface, API, SDK, data interface, and so on via embedded code is the "unattended" task. "Personal assistant" operations, such as data entry and data reading on the SAP client or web page, are typically carried out with no manual labor or less on the user's desktop. That being said, "unattended" processes like mail processing and data format conversion are carried out on the server or in the background without an interface and the need for human participation . It automates human-computer interaction without changing the business's or organization's IT architecture[9]. Three types of RPA are recognized: attended, unintended, and hybrid. The attended approach supports each worker as a personal assistant by utilizing RPA. It functions by receiving an order from the worker. Using RPA, the unintended model automates the process without involving human workers. The hybrid model is located between the two paradigms. The hybrid paradigm works by letting users interact in areas that require cognitive decision-making while carrying out some tasks automatically [10].

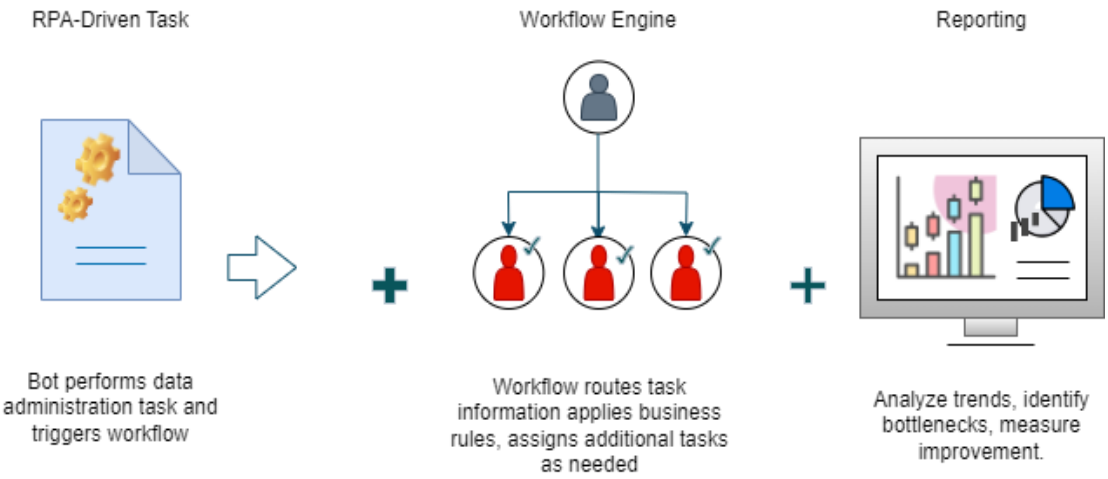


Figure 1 Work flow of RPA

The diagram in Fig 1 outlines the machine learning process starting with data acquisition where raw material in the form of raw data is obtained. This data then has to be prepared before it can be employed in the analysis; the preparation includes cleaning and pre-processing of data[11]. Then, the machine learning models[12],[13] are trained with the help of the prepared data[14] as presented by the intertwined people. Last but not least, the trained models[15],[16] are assessed and understood, as it is depicted in Figure 1. This workflow shows how data is processed systematically through the training of models to enable the processing of data to provide meaningful insights for decision-making [17],[18].

RPA can be implemented in many areas within an organization's organizational structure, such as payments and account management (managing purchase orders, generating invoices, automating payments, generating reports, managing budgets, journaling, and reconciling accounts); human resources (arranging vacations, applying for jobs, virtual assistants); and information technologies (security administration, monitoring, and management of infrastructures, managing users and accesses, integration of platforms, management of IT services) [19]. As per IBM, RPA was implemented in digital and data security processes by 64% of mid-sized businesses, general IT operations by 60%, financial management and planning tasks by 51%, marketing analytics by 49%, sales prediction by 46%, network optimization by 45%, customer experience optimization by 45%, and employee optimization by 45% . Using historical data, machine learning models can be taught to find trends that are associated with software flaws or performance problems. Next, by using these models, testing

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

efforts can be prioritized, test coverage can be optimized, and potential faults may be found early in the development lifecycle. ML can also be used to improve resource allocation; for example, it can be used to scale resources automatically based on demand patterns or to identify the best infrastructure design [20].

The uses, demands, and significance of machine learning have increased with the field's ongoing progress. In recent years, the term "big data" has gained a lot of traction; this is partly because machine learning has become more sophisticated, which makes it easier to evaluate large amounts of big data. By utilizing automatic sets of general methods in favor of more conventional statistical techniques, machine learning has also altered the way data extraction and interpretation are carried out. Machine learning operates along the lines shown in this process flow diagram.

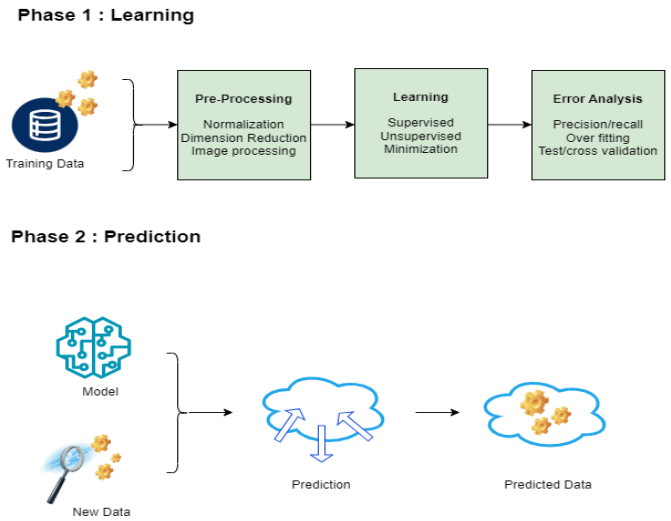


Figure 1.2 Work flow of ML

The above diagram showcases one of the most inclusive structures spanning the full machine learning process from data acquisition and preparation where data gets normalized, its dimensions reduced, and images processed among others. Next, the learning phase consists of supervised and unsupervised learning that includes the concepts of optimization and minimization. The last step is performance evaluation where metrics such as precision and recall of the model as well as concerns of over-fitting and ways of testing such as test/cross-validation. This workflow makes sure that raw data collected are pre-processed and selected properly for the development of models with the right techniques and special caution put in the assessment of the accuracies and efficiencies of the models developed .

Modern automation solutions are being used by businesses to streamline their information operations. RPA is one such crucial and practical application. With the use of software, machine learning, and artificial intelligence, this technology automates manual tasks that were previously completed by humans [21]**.** Robots can learn from their mistakes and gradually get better at what they do due to machine learning. Image and speech recognition are two examples of specialized challenges that are hard to handle using typical machine-learning techniques. Deep learning is used to address these issues [22]. Artificial Intelligence through Machine Learning is made possible by using modern RPA methodologies. The functioning of algorithms and other processes when teaching robots is heavily impacted by data inaccuracies, which emphasizes the significance of RPA governance even more since robotic processes serve as the sources of data collection **.** Advanced robotic systems can be created to carry out difficult tasks that were previously thought to be impossible by combining various technologies. In terms of analyzing and modifying sophisticated robotic systems, their link is all-encompassing. Robots

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

can detect things, navigate challenging surroundings, and even make judgments based on data that is updated in real-time with the help of AI. Using machine learning, robots can be trained to adapt to changing circumstances and learn from their experiences . With standard programming techniques, robots would not be able to accomplish difficult jobs. However, the use of DL is achievable. Programming languages including Python, C++, MATLAB, and ROS (Robot Operating System) are used extensively in robotics. Robotic system integration with AI, ML, and DL is made simpler by the libraries and tools available for these programming languages .

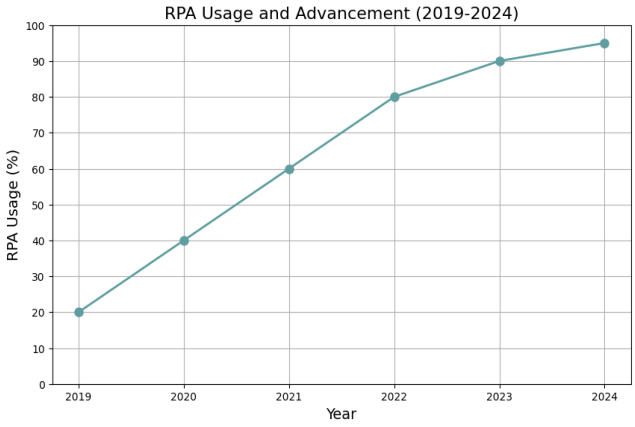
Despite RPA's widespread use, there is little scholarly literature on the technology; what is available primarily focuses on its advantages and characteristics. Due to their inherent limitations in terms of decision-making, language comprehension, and visual skills, current RPA systems are made to follow pre-established rules and workflows by encoding the graphical user interface (GUI) at the pixel level [23]. Process mining (PM), on the other hand, is a collection of methods that derive models from human behavior captured in event logs. Robotic Process Mining (RPM) is the term used to describe the junction of these domains. The goal of RPM instruments is to turn UI logs into RPA scripts. However, there exist certain distinctions between UI logs (input for RPM) and event logs (input for PM). First off, there is a difference in the degree of abstraction: event logs are generated at the task level of the business process, whereas UI logs are generated at the GUI level [24]. Using user interface (UI) logs, which document interactions between human actors and applications, RPM can gather processes in an organized manner, rank candidate processes according to how relevant it is to use RPA, and capture the specifications of the candidate processes in a way that facilitates their conversion into RPA scripts. Fully deterministic and partially deterministic processes are the two categories of processes that the authors differentiate between. Whereas the latter requires human interaction, the former can be automated without any help from outside sources [25]. In Autonomous DevOps, the integration of RPA, AI, and ML provides several advantages. By handling repetitive tasks, it increases scalability, reduces errors through intelligent decision-making, automates manual tasks, makes proactive issue detection possible through predictive analytics, and maximizes software performance and resource allocation using machine learning models. It also expedites software delivery, facilitates teamwork, and simplifies procedures . Robotic Process Automation greatly improves the automation of routine and rule-based processes but it also has some cons. It is important to identify key problems such as limited coverage and integration, poor cognitive skills, potential workforce impact, high cost of maintenance security and origination and integration/compliance issues. Due to these downsides, however, RPA systems need to be planned well ahead of deployment and thereafter maintained closely to make sure that they yield lasting gains and in - hand efficiently apportion them any hazards [26]. Combining ML algorithms with robotic process automation[27] enhances the effectiveness of RPA by means of intelligent decision making, adaptive learning and optimized processing. Using predictive maintenance, and anomaly detection ensures reliability while machine learning models like OCR or NLP provide better data extraction techniques for their processing. It uses machine learning- powered chatbots [28]. AI-driven work allocations and some sort of fraud detection that will not only make it seamless user interaction but also make sure whatever goes inside gets monitored frequently. These advanced capabilities can help organizations bring large productivity dividends in their operations automation, higher accuracy and greater operational efficiency . Advances in AI and cognitive technologies seem to converge with RPA in the future. Intelligent automation where bots have cognitive capacities, self-learning processes, and adaptive decision-making abilities is anticipated to develop from RPA. This development portends the democratization of automation technologies, allowing non-technical individuals to create and implement

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

automated workflows that are customized to meet their requirements . RPA is a technical innovation that is revolutionizing organizational procedures. RPA is positioned as a catalyst for digital transformation across industries, promising future progression toward intelligent automation because of its capacity to automate regular processes, improve operational efficiency, and embrace technological advancements .

Background:

Prior research has shown that robotic process automation is growing in Western economies because it can help with one of the main business issues facing today's corporate world: the need to handle ever-growing amounts of data. However, since its inception, this automated system has encountered several important problems. They concentrate on the three primary domains of technological, organizational, and economic variables as indicated in the sidebar note . Because of the recent large increase in data volume, many industries are trying to leverage this new resource for their service offerings. ML is advancing significantly in almost every industry. Furthermore, several companies use machine learning models in their information systems to increase process productivity or offer new products and services. Some key challenges that have recently emerged in research and practice are the scalability of high-dimensionality model datasets, distributed computing, and the real-time use of machine learning on streaming data . These findings imply that to reduce human involvement, cut down on mistake rates, manage the processing time, and preserve cost-effectiveness, we need an alternative to the manual processing of documents. These elements have contributed to a recent increase in the importance of document processing research. RPA has been used in business process optimization, so the concept is not new. RPA is a type of automation technique that replaces human intervention at the beginning point to an endpoint with integration, by developing software robots for repetitive high-volume tasks targeting processes based on rule. By pairing ML with RPA, many of the limitations of traditional RPA are addressed. The quality and accuracy of decision-making improves because of machine learning, providing greater adaptability for the system; the capability to handle more complex and unstructured data is enhanced; self-healing capacity minimizes maintenance costs; security and compliance are improved due to better monitoring; easier integration and communication with other systems are enabled; and machine learning helps to augment the capacities of human workers by automating repetitive workloads and enabling more time for humans to perform more strategic and complex tasks . Across industries, the breadth of use cases and the effectiveness of RPA are both enhanced by this combination, resulting in a more resilient, adaptable and intelligent automation system. The aim of this research, therefore, is to gather important issues such as benefits derived from RPA; as well as criteria needed for determining when a given process can be effectively automated using these techniques between others.



ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Figure 1.3 RPA applications and its development

The use of RPA has been on the rise as depicted in the graph from 2019 to 2024 the usage of the same is expected to rise from approximately 20% in 2019 to nearly 100% in 2024. This also works to show that there has been a progressing demand and evolution of RPA technologies in these years .

Literature Review:

Since RPA is a relatively new and unexplored area of research, conducting a literature review is sufficient because it can highlight gaps in the body of knowledge regarding RPA and suggest areas for future study. It can also serve as a framework for positioning new research areas. Additionally, as the review is structured around concepts and is presented and analyzed based on the most pertinent concepts connected to RPA, this research adheres to the concept-centric approach put forward [21]**.** Machine learning is a tool used in RPA to identify and respond to things on a screen. In most cases, APIs and specially written scripts are utilized when a programmer wishes to automate a list of tasks for a certain goal as a workflow. Using a Graphical User Interface (GUI), RPA's automation mechanism watches and records manual execution to replicate the activities of a coder. Irrespective of the menus' placement on the screen, the RPA records each menu click and key, and then repeats the sequence as many times as required [13], [21]. This contrasts with conventional methods, which use a key and a record of each click's location for future replication. Fixed recorded scripts won't work as intended or remain valid if menus change in location, shape, screen resolution, operating system, or even if an error happens [22]**.** RPA is always changing and adopting new technology and is increasingly used in conjunction with cutting-edge technologies like cognitive computing, AI, ML, and NLP. Due to this evolution, intelligent automation is now possible, allowing bots to perform more complicated jobs than just following rules. They can now learn from interactions, adapt to changing surroundings, and judge based on context [18]. In addition, ensemble learning techniques have been used to study performance prediction and optimization. An ensemble model combining multiple machine learning methods was presented to predict the scalability of software systems. Overall, studies revealed they could handle different workloads and estimate system performance more accurately [6]. The overwhelming consensus in the literature is that combining AI/ML techniques with performance analysis and recommendation systems has much potential to improve software application performance. Further research is required to develop comprehensive frameworks for the intricate relationships between system configurations, optimization strategies, and performance indicators [23]**.**

ML Approach:

In RPA, machine learning-based predictions are produced by training algorithms on large datasets to find patterns. After that, these algorithms are incorporated into RPA systems to automate challenging tasks that typically require human participation [20]. It comprises some steps, i.e. Compiling historical data from several sources, preparing it for analysis, and evaluating its caliber are all part of data collection. Exploratory analysis can aid in understanding the nuances of the data, which is important for analytical techniques. Train and test subsets of the data are separated apart, with the former being utilized to construct the model [24]. The algorithm's decision is made after test data is used to assess the correctness of the model. Changes to the model or the addition of new variables may improve performance. It takes a lot of time and work to complete this process [5].

Proposed Methodology:

Selecting the appropriate algorithm may seem intimidating. Numerous machine learning algorithms, both supervised and unsupervised, exist, and they all approach learning in different ways. There's no one-size-fits-all solution or ideal method. Even highly skilled data scientists can't predict whether an algorithm would work without testing it, so choosing the best one often

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

comes down to trial and error. The amount and kind of data you're working with, the insights you hope to derive from it, and the intended application of those insights all play a role in the algorithm selection process [5], [9]**.** Model performance and accuracy depend heavily on training them on pertinent datasets. Maintaining the efficiency and dependability of machine learning models requires constant observation and assessment [25]. The machine learning methodologies used in this paper are decision classification tree, decision regression tree, random classification forest, random regression forest and gradient boosting machine.

By categorizing data, the **Decision Classification Tree** facilitates automated decision-making and expedites processes like anticipating the next course of action or detecting exceptions. This expedites task completion and lessens the need for manual intervention. On the other hand, the **Decision Regression Tree** deals with predicting proportions and is useful where continuous outcomes might be desirable in establishing cost, time and resource regression. It's possible to schedule better, allocate resources better, and load balance which in turn decreases operating expenses. Again, an ensemble learning technique that is powerful compared to other classification techniques, particularly for complicated classification problems like spam detection and document classification, the **Random Classification Forest** is an alloy of several decision trees. This robustness ensures that the number of classification errors is minimized and thus affirmatively the level of accuracy in the automated processes is highly enhanced. Likewise, the **Random Regression Forest** model is critical for extrapolating the KPIs such as the throughput rates or the customer satisfaction indices because such a model generates continuous values by providing the average of the predictions of numerous regression trees. Further, accurate predictions generated with the help of this algorithm enhance predictive & preventative maintenance, capacity forecasting, and implementing dynamic changes by increasing the efficiency of automated work scheduling and advanced optimization.

Gradient Boosting Machines is an ensemble learning algorithm that works by building an accurate and robust model by adding decision trees incrementally. It is generally appreciated for the higher precision and the capacity to work in different data sets' patterns.

When combined, these algorithms enhance the flexibility and operation speed of RPA systems; they also enhance the decision-making capability and the capacity to predict the consequences of the system. They optimize resource use, reduce managerial intervention and hasten procedures by filtering out complex decisions and making exact predictions. Thus, human output increases, operating costs decline, and the nature of the automation structure evolves and expands. Decision trees as well as ensemble methods are a good combination in which interpretability and robustness are balanced, thus the effectiveness and efficiency of RPA systems are improved for working with a set of tasks and processes.

Data Set:

The data set used in this paper contains **11,055 rows** and **50 columns** and is intended to help differentiate between reputable and phishing websites. Numerous characteristics taken from URLs are included, including the quantity of dots, subdomain tiers, path complexity, and the existence of specific symbols. Additional characteristics encompass the length of various URL elements, the existence of insecure forms, and the utilization of HTTPS. These characteristics are essential for spot-spotting patterns frequently linked to phishing attempts. The main goal of the data set is to make it easier to train and assess machine learning models that better protect users from cyber threats by correctly identifying phishing websites.

1. **Results of Decision Classification Tree:**

Accuracy: 96.73% (The model correctly classified 96.73% of the instances in the test set)

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Confusion Matrix:

TN = 1414	FP = 49
FN = 49	TP = 148E

Classification Report:

	precision	recall DRF	f1-score	support
0	0.97	0.97	0.97	1463
1	0.97	0.97	0.97	1573
accuracy			0.97	3000
macro avg	0.97	0.97	0.97	3000
weighted avg	0.97	0.97	0.97	3000

Precision: Precision for both classes (0 and 1) is 0.97, meaning 97% of the predicted positives are true positives.

Recall: The recall for both classes is 0.97, indicating that 97% of the actual positives were correctly identified by the model.

F1-score: The F1-score for both classes is 0.97, a harmonic mean of precision and recall, indicating a good balance between the two metrics.

Interpretation:

With high accuracy, precision, recall, and F1 scores, the Decision Tree Classifier has done well on this data set. Both classes perform similarly, with few false positives and false negatives, according to the confusion matrix.

Results of Decision Regression Tree:

Evaluation of the model's performance using metrics such as Mean Squared Error (MSE) and R^2 score.

DRF	
MSE	0.0327
R^2	0.8693
Accuracy	87%

Interpretation:

- The MSE of the Decision Tree Regression model is comparatively low, suggesting that the predictions are reasonably close to the observed values.
- Regression models perform rather well when they have an R^2 score of 0.8693, which indicates that the model explains a substantial proportion of the variance in the target variable.

Results of Random Forest Classification:

Accuracy: 98.1% (The model accurately identified 98.1% of the cases in the test set. This remarkably high accuracy demonstrates how well the software distinguishes between legitimate and fictitious URLs.)

Confusion Matrix:

TN = 1437	FP = 26
FN = 31	TP = 1506

The confusion matrix indicates good performance because it shows relatively few false positives and false negatives.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Classification Report:

	precision	recall	f1-score	support
0	0.98	0.98	0.98	1463
1	0.98	0.98	0.98	1573
accuracy			0.98	3000
macro avg	0.98	0.98	0.98	3000
weighted avg	0.98	0.98	0.98	3000

Precision: 98% of the predicted positives are real positives or 0.98 for both classes (0 and 1).
Recall: The model successfully detected 98% of the real positives, with a recall of 0.98 for both classes.
F1-score: A harmonic mean of precision and recall, or 0.98 for both classes, indicates a fair balance between the two metrics.

Interpretation:

The classifier model does a remarkable job of differentiating between phishing and legitimate URLs in terms of precision, recall, and F1 scores. It is highly accurate with a 98.1% accuracy rate and excellent precision, recall, and F1 scores for both classes. The confusion matrix shows minimal classifications, indicating strong performance.

2. **Result of Random Forest Regression:**

RRF	
MSE	0.0175832
R^2	0.9296243803451967
Accuracy	93 %

Interpretation:

Along with its outstanding accuracy, the regression model has a high R^2 score of 0.9296 and a low MSE of 0.0176. In other words, the variation of the target variable is mostly explained by the model, whose predictions are highly similar to the actual values.

3. **Result of Gradient Boosting Machine:**

Accuracy: 98.1% (The model had a test accuracy of 98.1% so this means that the model makes correct predictions with an average probability.)

Confusion Matrix:

TN = 966	FP = 22
FN = 16	TP = 996

Classification Report:

	precision	recall	f1-score	support
0	0.98	0.98	0.98	988
1	0.98	0.98	0.98	1012
accuracy			0.98	2000
macro avg	0.98	0.98	0.98	2000
weighted avg	0.98	0.98	0.98	2000

Interpretation:

As the results show, we observed that our GBM model is very good at discerning between real URLs and phishing URLs characterized by its high Precision as well, recallF1-scores for both

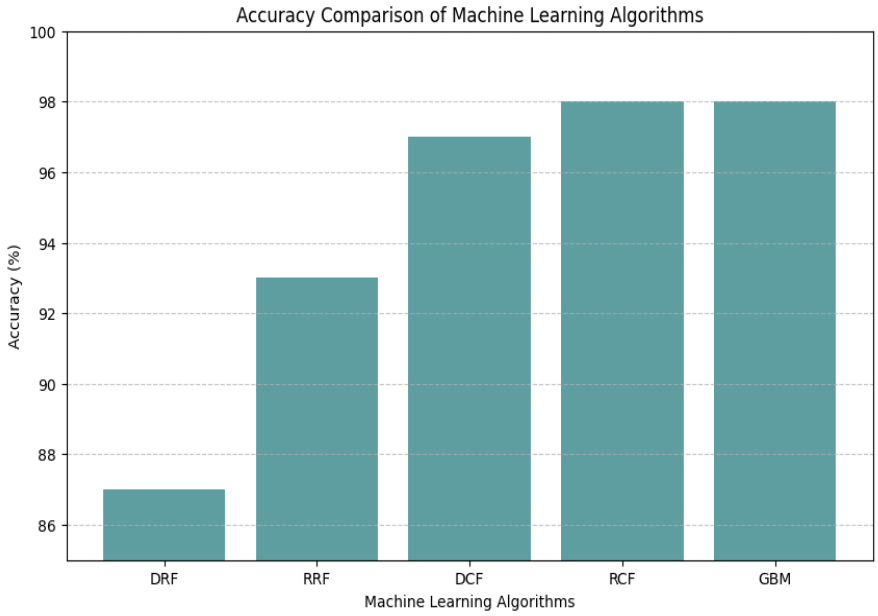
ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

classes. Due to its high detection accuracy of phishing and legitimate URLs, the model is proposed as a potent mechanism for this classification process.

Collective Results:

	DRF	RRF	DCF	RCF	GBM
Accuracy	87%	93 %	97%	98%	98%

Below is a bar chart representing the achieved accuracy of various types of machine learning algorithms. It depicts the performance of each of the algorithms and shows that in terms of accuracy, RCF and GBM were the best with 98% while DCF followed with 97%, RRF with 93% and last was DRF with 87%.



Moreover, the implications and lessons learned from these findings can serve as a pivotal roadmap for RPA practitioners to address critical barriers that prevent the deeper incorporation of machine learning in automation processes [24].

Research Gaps and Opportunities:

Even with abundant research on RPA, some holes in the field still need to be filled. Few studies have been done on the socio-economic effects of RPA adoption on a large scale. Nor have there been many ethical issues and ramifications of democratizing RPA for citizen developers. A thorough understanding of robotic process automation's (RPA) current state, applications, advantages, difficulties, and future directions can be gained from the literature now in publication. However, there are also gaps in our knowledge of its wider societal ramifications and how RPA may be made more accessible to non-technical users. Filling in these gaps will provide a more comprehensive comprehension of RPA's industry-transforming potential.

A literature review with gap analysis is shown in Table [18], [23], [26].

References	Summary	Research Gap
Syed, R., et al. (2020)	Investigated modern issues and difficulties in robotic process automation.	Slightly discussed are the moral ramifications of RPA adoption and how it affects job duties.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Pramod, D. (2022)	Examined the level of RPA adoption, its advantages and disadvantages, and suggested a study agenda.	Neglecting the difficulties associated with scaling and maintaining RPA systems over the long run.
Gotthardt, M., et al. (2020)	Analyzed the difficulties of integrating intelligent RPA with accounting and auditing.	Minimal attention was paid to integration and interoperability difficulties with intelligent RPA with current frameworks.
Chakraborti, T., et al. (2020)	Investigated developments between RPA and IPA (intelligent process automation).	Insufficient understanding of the real-world implementation difficulties encountered while switching from RPA to IPA.
Patri, P. (2021)	Investigated the banking industry's RPA problems and solutions.	Insufficient attention is paid to the security and compliance issues unique to the banking industry in RPA.
Santos, F., et al. (2020)	Investigated the deployment of RPA from end to end.	Insufficient conversation about how RPA adoption will affect society on a socioeconomic level.
Agostinelli, S., et al. (2021)	Examined the difficulties with RPA's automated segmentation.	Not enough attention is paid to how technology developments are resolving RPA's segmentation issues.
Kämäräinen, T. (2018)	Examined the potential and difficulties of federated governance for managing the RPA model.	There is insufficient focus on the scalability issues with large-scale RPA governance structures.

Conclusion & Future Work:

This research paper aims to integrate machine learning techniques to increase robotic process automation process efficiency. Robotic Process Automation has had a significant impact on the service industry by automating repeatable, rule-driven tasks once carried out painstakingly. But currently, the benefits RPA provides are mostly restricted by its absence of cognitive skills and inflexibility. The research suggests that machine learning algorithms, such as decision regression trees or random classification forests overcome the mentioned limitations. Better decision-making and outcome prediction also help make the models more intelligent in managing very complex classification jobs while enhancing RPA. The random classification

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

forest and the decision classification tree improve automation by making accurate automated decisions, forgoing human intervention with maximum efficiency. On the other hand, random regression forest and decision forward project Bayesian model averaging help to resource allocation, load balancing as well as cost estimation for continuous outcome prediction. Secondly, system analysis and data capture will need to become more sophisticated while incorporating strict security measures – especially about unstructured content that could also continually evolve due to ML. Automation systems become more resilient and flexible when these three advantages are combined: improved workflow efficiency, reduced maintenance costs, and predictive maintenance capabilities. The primer showed that systems using ML algorithms in RPA can increase output, improve operations efficiency and reduce costs. This paper also described the possibility of RPA becoming a form of intelligent automation where processes can be performed through self-learning bots making decisions will serve to democratize automation technologies, making them available to a larger number of non-technical consumers and inspiring future innovations in digital transformation in other fields. Emphasizes that the next wave of progress in RPA will be a continuous innovation and the integration of the newest AI and cognitive technologies will be the key to unlocking the hidden potential of RPA.

References:

[1] Akbar, Salman. "DESIGNING ACCESS, NOT JUST SCHOOLS (OUT OF SCHOOL CHILDREN): EQUITY-DRIVEN LESSONS FOR PAKISTAN FROM THE US IN ADDRESSING EDUCATIONAL EXCLUSION." Contemporary Journal of Social Science Review 4, no. 1 (2026): 154-172.

[2] D. Kedziora and S. Hyrynsalmi, “Turning Robotic Process Automation onto Intelligent Automation with Machine Learning,” 2023.

[3] Khan, Ali Raza A., Muhammad Ismaeel Khan, and Aftab Arif. "AI in Surgical Robotics: Advancing Precision and Minimizing Human Error." Global Journal of Computer Sciences and Artificial Intelligence 1, no. 1 (2025): 17-30.

[4] Mim, Sadia Afrin, Fatemeh Vares, Andrew Meenly, and Brittany Johnson. "An investigation into open source fairness tool sustainability." In Proceedings of the 1st International Workshop on Responsible Software Engineering, pp. 21-28. 2025.

[5] McQueary, Wren, Sadia Afrin Mim, Md Nishat Raihan, Justin Smith, and Brittany Johnson. "Py-holmes: Causal testing for deep neural networks in python." In Companion Proceedings of the 32nd ACM International Conference on the Foundations of Software Engineering, pp. 602-606. 2024.

[6] F. Santos, R. Pereira, and J. B. Vasconcelos, “Toward robotic process automation implementation: an end-to-end perspective,” Bus. Process Manag. J., vol. 26, no. 2, pp. 405–420, Jan. 2019, doi: 10.1108/BPMJ-12-2018-0380.

[7] Khan, Hira Haroon, Saleem Zubair, Fawad Nasim, Shamim Akhter, Muhammad Naushad Ghazanfar, and Sumbul Azeem. "Role of Kubernetes in DevOps Technology for the Effective Software

[8] “Managing RPA Scalability: An Explanatory Case Study - ProQuest.” Accessed: Aug. 21, 2024.

[9] Mim, Sadia Afrin, and Brittany Johnson. "Inside Fairness Tools: What Academic Practitioners Really Experience." In 2025 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), pp. 444-446. IEEE, 2025.

[10] Nasim, M. F., M. Anwar, A. S. Alorfi, H. A. Ibrahim, A. Ahmed, A. Jaffar, S. Akram, A. Siddique, and H. M. Zeeshan. "Cognitively inspired sound-based automobile problem

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

detection: A step toward explainable AI (XAI)." International Journal of Advanced and Applied Sciences 12, no. 8 (2025): 1-15.
[11] Azam, Muhammad, Fawad Nasim, Jawad Ahmad, and Sohail Masood Bhatti. "A Security Framework for Data Migration over the Cloud." Journal of Computing & Biomedical Informatics 7, no. 02 (2024).
[12] Nasim, Fawad, Sohail Masood, Arfan Jaffar, Usman Ahmad, and Muhammad Rashid. "Intelligent Sound-Based Early Fault Detection System for Vehicles." Computer Systems Science & Engineering 46, no. 3 (2023).
[13] Tariq, Muhammad Arham, Muhammad Ismaeel Khan, Aftab Arif, Muhammad Aksam Iftikhar, and Ali Raza A. Khan. "Malware Images Visualization and Classification With Parameter Tunned Deep Learning Model." Metallurgical and Materials Engineering 31, no. 2 (2025): 68-73.https://doi.org/10.63278/1336.
[14] Mim, Sadia Afrin, Fatema Tuz Zohra, Justin Smith, and Brittany Johnson. "An Exploratory Analysis of Available Fairness Interventions in Open Source." In 2025 IEEE International Conference on Software Analysis, Evolution and Reengineering-Companion (SANER-C), pp. 17-24. IEEE, 2025.
[15] Jabed, Mohammad Majharul Islam, Ahmed Sohaib Khawer, Sharmin Ferdous, Danial Hadid Niton, Amit Banwari Gupta, and Mohammed Shafeul Hossain. "Integrating Business Intelligence with AI-Driven Machine Learning for Next-Generation Intrusion Detection Systems." International Journal of Research and Applied Innovations 6, no. 6 (2023): 9834-9849.
[16] Ankhi, Rokeya Begum. "Leveraging Business Intelligence and AI-Driven Analytics to Strengthen US Cybersecurity Infrastructure." International Journal of Engineering & Extended Technologies Research (IJEETR) 7, no. 2 (2025): 9637-9652.
[17] Arif, Aftab, Fadia Shah, Muhammad Ismaeel Khan, Ali Raza A. Khan, Aftab Hussain Tabasam, and Abdul Latif. 2023. “Anomaly Detection in IoHT Using Deep Learning: Enhancing Wearable Medical Device Security.” Migration Letters 20 (S12): 1992–2006.
[18] Hameed, Usman, Sohail Masood, Fawad Nasim, and Arfan Jaffar. "Exploring the Accuracy of Machine Learning and Deep Learning in Engine Knock Detection." Bulletin of Business and Economics (BBE) 13, no. 3 (2024): 203-210.
[19] Jabed, M. M. I., & Ferdous, S. (2024). Integrating Business Process Intelligence with AI for Real-Time Threat Detection in Critical US Industries. International Journal of Research and Applied Innovations, 7(1), 10120-10134.
[20] “The Power Trio: RPA, AI, and ML: Autonomous DevOps Leverages a Powerful Combination of Technologies | Request PDF.” Accessed: Aug. 21, 2024.
[21] J. Siderska, L. Aunimo, T. Süße, J. von Stamm, D. Kedziora, and S. N. B. M. Aini, “Towards Intelligent Automation (IA): Literature Review on the Evolution of Robotic Process Automation (RPA), its Challenges, and Future Trends,” Eng. Manag. Prod. Serv., vol. 15, no. 4, pp. 90–103, Dec. 2023, doi: 10.2478/emj-2023-0030.
[22] Zainab, Hira, A. Khan, Ali Raza, Muhammad Ismaeel Khan, and Aftab Arif. "Integration of AI in Medical Imaging: Enhancing Diagnostic Accuracy and Workflow Efficiency." Global Insights in Artificial Intelligence and Computing 1, no. 1 (2025): 1-14.
[23] A. Jain, S. Paliwal, M. Sharma, L. Vig, and G. Shroff, “SmartFlow: Robotic Process Automation using LLMs,” May 21, 2024, arXiv: arXiv:2405.12842. Accessed: Aug. 21, 2024.
[24] D. Moraru, A. Sterca, V. Niculescu, M.-C. Chisa, and C.-C. Osman, “Conceptual Model Detection in Business Web Applications Within the RPA Context,” 2024.
[25] “A rule-based method to effectively adopt robotic process automation - Bédard - Journal of Software: Evolution and Process - Wiley Online Library.” Accessed: Aug. 21, 2024.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

[26] M. Mashtakov, S. Shirokova, and M. Bolsunovskaya, "Application of RPA technology in management and decision-making processes," 2023, doi: 10.57809/2023.2.1.4.3.
[27] Din, Ahmad, Basilio Bona, Joel Morrissette, Moazzam Hussain, Massimo Violante, and M. Fawad Naseem. "Embedded low power controller for autonomous landing of UAV using artificial neural network." In 2012 10th International Conference on Frontiers of Information Technology, pp. 196-203. IEEE, 2012.
[28] "No Need to Cry over Spilt Milk: A Workflow for Regenerating Graph Data Using Robotic Process Automation | SpringerLink." Accessed: Aug. 21, 2024.
[29] L. Baier, F. Jöhren, and S. Seebacher, CHALLENGES IN THE DEPLOYMENT AND OPERATION OF MACHINE LEARNING IN PRACTICE. 2019.