# Data-Driven Business Intelligence for Operational Customer Churn Management

**Dina K. Hassan[1,*], Ahmed K. Metawee[2]**

[1]Accounting Department, Faculty of Commerce, Kafr El Sheikh University, Egypt

[2]Accounting Department, Faculty of Commerce, Mansoura University, Egypt

Emails: dina.abdelsalam@com.kfs.edu.eg; metawee68@mans.edu.eg

**Abstract**

In today's data driven world businesses face a challenge in protecting customer strategies from operational churn. This paper explores the realm of data driven business intelligence with a focus on predicting and managing customer churn through analysis of analytics methods. Recognizing that customer attrition poses a threat to business sustainability, our research aims to harness the power of methods and discriminant analysis techniques. We examine Gradient Boosting Classifier, Ada Boost Classifier and Linear Discriminant Analysis to unravel patterns in customer behavior and predict churn likelihood. By utilizing a dataset that includes details about customer services account specifics and demographics we adopt an approach. Our comparative analysis of machine learning classifiers underscores their effectiveness in identifying patterns within the dataset. Importantly our findings emphasize the potential of machine learning as a strategy for managing churn.

**Keywords**: Business Intelligence; Operation Research; Customer Churn Management; Customer Segmentation.

## 1. Introduction

In today's business landscape, where prioritizing customer needs crucial, for growth the way we handle customer churn has undergone significant changes. The combination of technology and data driven insights has introduced an approach that relies on Business Intelligence (BI) to understand the complexities surrounding customer attrition. This article explores the field of customer churn management delving into how data driven strategies within BI frameworks can not only predict but also proactively mitigate churn [1-3]. Traditionally seen as a reactive measure addressing customer churn has now become an effort grounded in data driven techniques. In a market filled with options it is essential to comprehend the factors that influence churn. Data plays a role by providing insights into customer behavior, preferences, and interactions. By analyzing both unstructured data businesses can uncover patterns that serve as the foundation for churn management strategies [4-6].

Furthermore, this paper goes beyond the aspects. Acknowledges the broader societal impacts of customer churn. In today's world that emphasizes inclusivity, fairness, and diversity the consequences of churn go beyond measures. By examining the layers of churn, we can understand how different demographics, socioeconomic backgrounds and cultural nuances intersect when it comes to customer attrition. This exploration aims to bridge the gap between approaches and the human element by recognizing the range of customers and their unique needs [7]. When businesses adopt an approach to managing customer churn, they not only protect their financial performance but also create an environment rooted in empathy and responsiveness. This paper sets out on a journey to uncover the relationship between data driven Business Intelligence and customer strategies. Ultimately this paves the way for a compassionate business landscape.



DOI: https://doi.org/10.54216/AJBOR.000205

The remainder of this study is organized as follows. Section 2 discusses the methodology of this study including the data-driven algorithms for churn prediction. In part 3, the experimental setups and the related exploratory analysis are presented. Section 4 provides an exhaustive discussion of the experimental results and their implications. Section 5 deduces the findings.

## 2. Methodology

This section delineates the blueprint guiding our exploration, detailing the comprehensive methodology crafted to decipher the complexities of customer attrition.

### 3.1. Gradient Boosting Classifier

Algorithm 1. Gradient Tree Boosting Algorithm.

1: Initialize $f_0(x) = \arg\min_\gamma \sum_{i=1}^{N} L(y_i, \gamma)$.
2: For $m = 1$ to $M$ :
3: For $i = 1,2,\ldots,N$ compute
4: $r_{im} = -\left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right]_{f=f_{m-1}}$
5: Fit a regression tree to the targets $r_{im}$ giving terminal regions $R_{jm}, j = 1,2,\ldots,J_m$.
6: For $j = 1,2,\ldots,J_m$ compute
7: $\gamma_{jm} = \arg\min_\gamma \sum_{x_i \in R_{jm}} L(y_i, f_{m-1}(x_i) + \gamma)$.
8: Update $f_m(x) = f_{m-1}(x) + \sum_{j=1}^{J_m} \gamma_{jm} I(x \in R_{jm})$.
9. Output $\hat{f}(x) = f_M(x)$.

Gradient Boosting Classifier is a technique that combines learners usually decision trees to create a powerful predictive model. It works by adding models one, after another to correct the mistakes made by the existing models. Each new model in the sequence focuses on the errors or discrepancies of the one gradually reducing prediction errors by giving importance to areas where the model doesn't perform well. This iterative process helps minimize prediction error and results in an accurate predictive model[8-11].

### 3.2. Ada Boost Classifier

Algorithm 2. Adaptive Boosting classifier

1: $H_0(x) = 0 \forall x$ → Initialization of classifier.
2: $w_i = \frac{1}{n} \forall i$ →Initialization of weights.
3: for t = 1 to T do
4: $h = \mathbb{A}(D_1(w_1, \ldots w_n))$ →Find the best weak learner for current weighting.
5: $\epsilon = \sum_{i:h(x_i) \neq y_i} w_i$ →Weighted error of $h$.
6: if $\epsilon > \frac{1}{2}$ then break;
7: end if →If worse than random, stop.
8: $\alpha = \frac{1}{2} \ln\left(\frac{1-\epsilon}{\epsilon}\right)$ →Compute learning rate.
9: $H_t = H_{t-1} + \alpha h$ →Take gradient step.
10: $\forall i w_i \leftarrow \frac{w_i e^{-\alpha h(x_i) y_i}}{2\sqrt{\epsilon(1-\epsilon)}}$ →Compute new (normalized) weights.
11: end for
12: return $H_T$

On the hand AdaBoost (short for Adaptive Boosting) is another method for learning that trains a series of weak learners sequentially. However, unlike Gradient Boosting AdaBoost assigns weights to each instance in the dataset and concentrates on samples that were misclassified before giving them weight. Subsequent weak learners then prioritize



DOI: https://doi.org/10.54216/AJBOR.000205

these misclassified instances with the goal of correcting them in iterations [12 15]. This adaptive nature allows Ada Boost to continuously improve its performance by concentrating on areas of difficulty within the dataset.

### 3.3. Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis is a classification technique that seeks to find linear combinations of features that best separate different classes in the dataset. It works by projecting the data into a lower-dimensional space while maximizing the separation between classes. LDA aims to maximize the ratio of between-class variance to within-class variance, finding the linear discriminants that best preserve class-specific information. This method is particularly useful when dealing with multi-class classification problems and assumes that the features are normally distributed within each class [16-17].

## 3. Experimental Design

This section delineates the architecture of our experimentation, outlining the deliberate design choices aimed at validating and refining the efficacy of our proposed methodologies.

3.1. Case study

Our case study revolves around utilizing customer data to anticipate behaviors that contribute to keeping customers. Utilizing the dataset provided by IBM Sample Data Sets we gather a wealth of information, about each customer, including the services they have subscribed to account details such as how they have been with us their payment methods and billing preferences. Additionally we collect information such as gender, age range and familial connections. Our analysis focuses on understanding and predicting churn behavior – studying patterns and correlations, in the services customers use (such as phone and internet subscriptions, security measures, support systems and entertainment options) – in order to develop tailored retention programs backed by analytics that cater to specific customer needs.

3.2. Data exploration

In our pursuit to comprehensively explore and interpret the rich landscape of customer behavior, we deploy a multifaceted visualization strategy, encapsulating the essence of our statistical analyses. Table 1 stands as a testament to the depth of our statistical scrutiny, offering a structured view of key metrics, trends, and summaries derived from our dataset. This tabular representation serves as a digestible repository, encapsulating variables, measures of central tendency, and variability, providing stakeholders with a concise yet informative snapshot of our analytical findings.

Table 1: statistical summary for the numerical variables.

| | tenure | MonthlyCharges | TotalCharges |
|---|---|---|---|
| count | 7032 | 7032 | 7032 |
| mean | 32.42179 | 64.79821 | 2283.3 |
| std | 24.54526 | 30.08597 | 2266.771 |
| min | 1 | 18.25 | 18.8 |
| 25% | 9 | 35.5875 | 401.45 |
| 50% | 29 | 70.35 | 1397.475 |
| 75% | 55 | 89.8625 | 3794.738 |
| max | 72 | 118.75 | 8684.8 |



DOI: https://doi.org/10.54216/AJBOR.000205

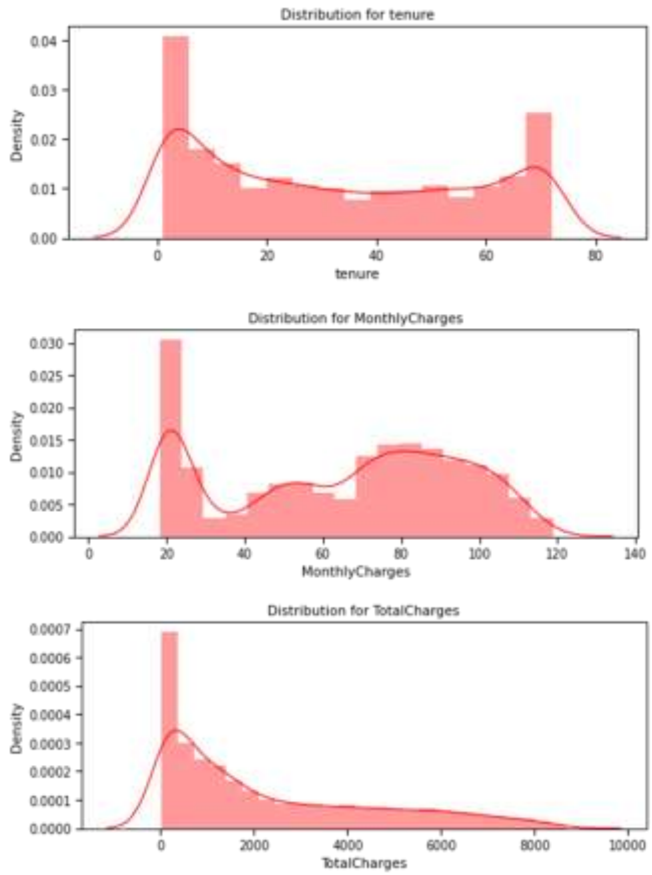


Figure 1: distributional analysis of custom churn variables

Figure 1 takes on the mantle of elucidating variable distributions, breathing life into the numbers by rendering them in visually engaging formats. Through histograms, density plots, or box plots, this graphical representation paints a vivid picture of the distribution patterns across various customer attributes—be it service subscriptions, account information, or demographic specifics. By harnessing the power of visual storytelling, Figure 1 serves as a visual gateway, inviting observers to discern trends, outliers, and nuances that might otherwise remain obscured in numerical abstractions, fostering a more intuitive understanding of the dataset's intrinsic characteristics.

Furthermore, Figure 2 emerges as the canvas upon which correlation analyses come to life, weaving intricate webs of connections between different customer attributes. Through correlation matrices, heatmaps, or network graphs, this visual depiction unveils the interplay and dependencies among variables, spotlighting relationships that hold significance in understanding customer behavior and churn. Beyond mere numbers, this visualization engenders a holistic comprehension of how various factors intertwine, empowering decision-makers to discern influential drivers and potential predictors of customer churn, thereby fostering more targeted and effective retention strategies.

3.3. Performance metrics

The evaluation process is performed using the following performance metrics through our experiments.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall(Sensitivity)} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (3)$$

$$\text{F1} - \text{score} \ \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$



DOI: https://doi.org/10.54216/AJBOR.000205

$$\mathrm{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)\times(TP+FN)\times(TN+FP)\times(TN+FN)}} \tag{5}$$

$$kappa = \frac{p_0 - p_e}{1 - p_e} \text{ with } p_0 = \frac{TP+TN}{TP+TN+FP+FN} \text{ and } p_e = \frac{(TP+FN)\times(TP+FP)+(FP+TN)\times(FN+TN)}{(TP+TN+FP+FN)^2} \tag{6}$$

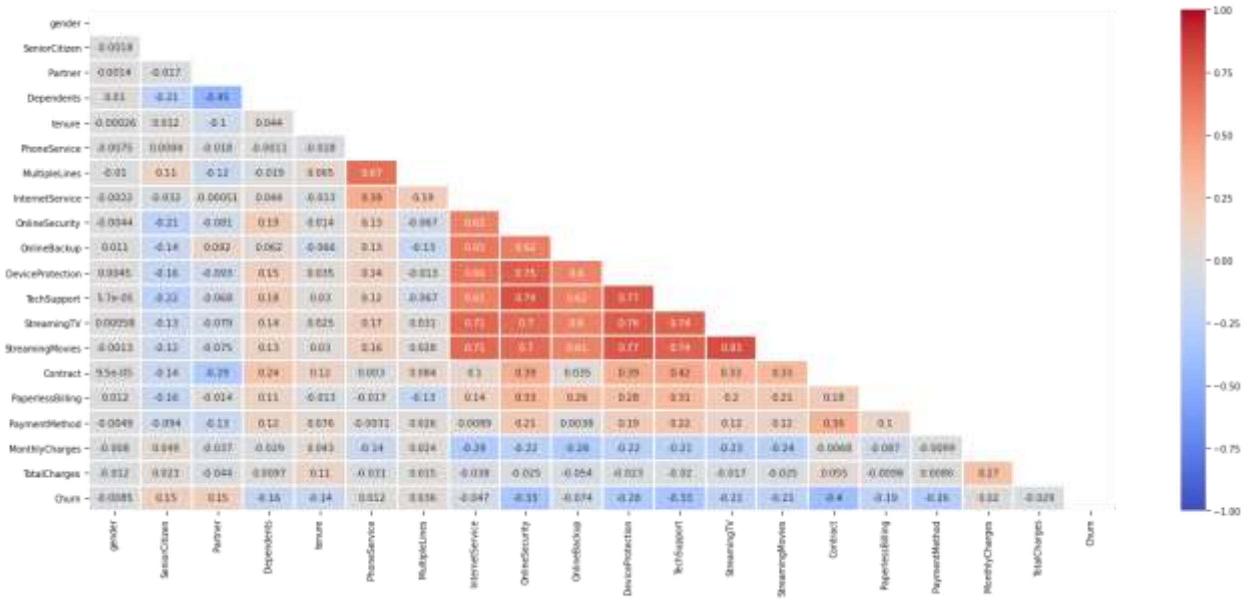


Figure 2: visual correlation analysis of custom churn variables

## 4. Results and Discussion

In this section, we unveil the empirical findings derived from our experiments and analyses, juxtaposing these revelations against established benchmarks and theoretical constructs.

Table 2 emerges as the synthesis of our methodological prowess, distilling complex algorithmic performances into a digestible format that empowers decision-makers with actionable insights. By presenting a side-by-side comparison, this tabular representation unveils the strengths and limitations of each classifier, offering a roadmap for selecting the most adept model suited to the specific nuances of customer behavior within the dataset. Moreover, it steers the narrative toward informed decision-making, fostering a data-driven approach in identifying the optimal predictive model for operationalizing customer churn management strategies within business intelligence frameworks. Through this comparative lens, Table 2 stands as a testament to our commitment to precision and efficacy in the realm of machine learning-driven customer churn prediction.

Table 2: Comparison of ML Classifier Performance for Predicting Customer Churn

| Model | Accuracy | AUC | Recall | Prec. | F1 | Kappa | MCC |
|---|---|---|---|---|---|---|---|
| Logistic Regression | 0.7969 | 0.8465 | 0.5499 | 0.6471 | 0.5941 | 0.46 | 0.463 |
| Gradient Boosting Classifier | 0.793 | 0.8455 | 0.5154 | 0.6503 | 0.5737 | 0.4398 | 0.4457 |
| Ada Boost Classifier | 0.7922 | 0.843 | 0.5364 | 0.638 | 0.5817 | 0.4453 | 0.4488 |
| Linear Discriminant Analysis | 0.7902 | 0.8409 | 0.5551 | 0.6273 | 0.5885 | 0.4485 | 0.4503 |
| CatBoost Classifier | 0.7867 | 0.8402 | 0.5169 | 0.6307 | 0.5669 | 0.4276 | 0.432 |
| Random Forest Classifier | 0.7835 | 0.8282 | 0.4824 | 0.6321 | 0.5459 | 0.4076 | 0.4146 |



DOI: https://doi.org/10.54216/AJBOR.000205

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Light Gradient Boosting Machine | 0.7831 | 0.8349 | 0.5266 | 0.6181 | 0.5675 | 0.4243 | 0.4274 |
| Extra Trees Classifier | 0.7725 | 0.8057 | 0.4846 | 0.5985 | 0.5342 | 0.3865 | 0.3908 |

In Figure 3, we unveil the intrinsic workings of our machine learning classifiers by spotlighting the feature importance across varied predictive models. This visual representation stands as a beacon illuminating the significance and impact of individual attributes in predicting customer churn. By rendering the importance of features within each classifier—be it decision trees, random forests, gradient boosting, or other models—this figure unveils the hierarchy of attributes that wield the most influence in determining churn behavior. Through bar plots, heatmaps, or hierarchical diagrams, Figure 3 encapsulates the narrative of feature importance, offering stakeholders a holistic view of which customer attributes wield the most weight in our predictive models. In Figure 4, we show the prediction performance under different resampling mechanisms.

## 5. Conclusion and Future work

This paper navigates the intricate terrain of operational customer churn management through the lens of data-driven business intelligence. Our exploration underscores the pivotal role of predictive analytics, machine learning, and comprehensive data analysis in fortifying businesses against customer attrition. By amalgamating diverse customer attributes, services subscribed to, demographic information, and account specifics, our research unveils a roadmap for proactive churn management. The comparative analysis of machine learning classifiers showcased in Table 2 and the elucidation of feature importance in Figure 3 underscore the efficacy of data-driven predictive models in deciphering and predicting customer behavior. Our findings advocate for a shift from reactive approaches to proactive interventions rooted in nuanced predictive insights, fostering a customer-centric business landscape anchored in empathy and responsiveness. Looking ahead, future work in this domain could delve deeper into dynamic modeling techniques that consider temporal trends and evolving customer behaviors. Incorporating real-time data streams, sentiment analysis from customer interactions, and integrating external factors such as market trends or social dynamics could enhance the predictive accuracy and relevance of churn models. Furthermore, exploring advanced AI-driven strategies like reinforcement learning or ensemble techniques might offer heightened precision in identifying early warning signals and tailoring retention programs. Additionally, a focus on ethical considerations surrounding data privacy and fairness in deploying these predictive models within business strategies remains crucial for building trust and ensuring responsible implementation in the evolving landscape of customer-centric business practices.

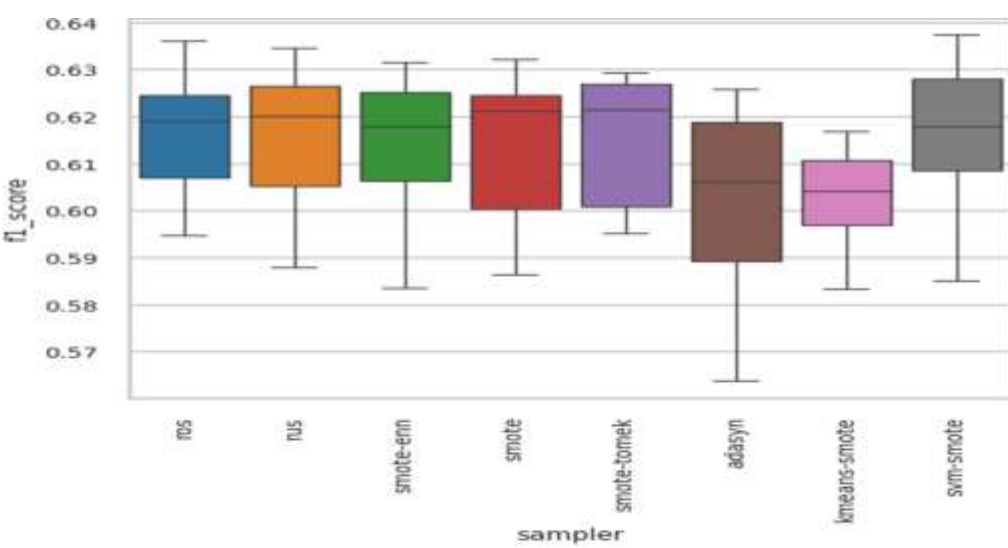


Figure 3: Comparison of impacts of different resampling methods on performance for Customer Churn prediction.



DOI: https://doi.org/10.54216/AJBOR.000205

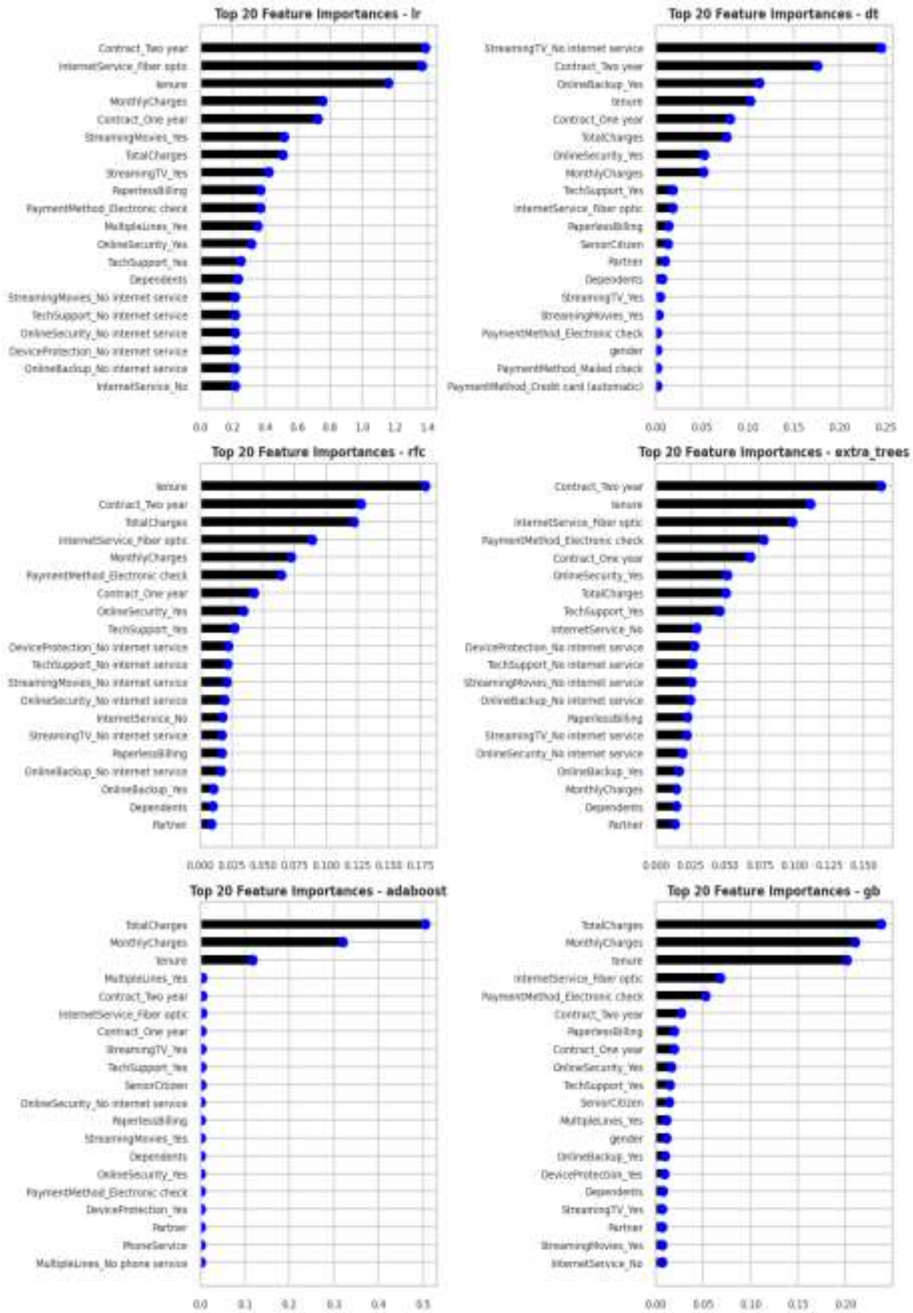


Figure 4: Visualization of feature importance for different ML classifiers

## References

[1] Osman, C. C. (2019). Robotic Process Automation: Lessons Learned from Case Studies. *Informatica Economica*, *23*(4).

[2] Dey, S., & Das, A. (2019). Robotic process automation: assessment of the technology for transformation of business processes. *International Journal of Business Process Integration and Management*, *9*(3), 220-230.



DOI: https://doi.org/10.54216/AJBOR.000205

[3] Devarajan, Y. (2018). A study of robotic process automation use cases today for tomorrow's business. *International Journal of Computer Techniques*, *5*(6), 12-18.

[4] Romao, M., Costa, J., & Costa, C. J. (2019, June). Robotic process automation: A case study in the banking industry. In *2019 14th Iberian Conference on information systems and technologies (CISTI)* (pp. 1-6). IEEE.

[5] Carden, L., Maldonado, T., Brace, C., & Myers, M. (2019). Robotics process automation at TECHSERV: An implementation case study. *Journal of Information Technology Teaching Cases*, *9*(2), 72-79.

[6] Flechsig, C., Lohmer, J., & Lasch, R. (2019). Realizing the full potential of robotic process automation through a combination with BPM. In *Logistics Management: Strategies and Instruments for digitalizing and decarbonizing supply chains-Proceedings of the German Academic Association for Business Research, Halle, 2019* (pp. 104-119). Springer International Publishing.

[7] Mishra, S., Sree Devi, K. K., & Badri Narayanan, M. K. (2019). Technology dimensions of automation in business process management industry. *International Journal of Engineering and Advanced Technology*, *8*(6), 1919-1926.

[8] Zhang, N., & Liu, B. (2019). Alignment of business in robotic process automation. *International Journal of Crowd Science*, *3*(1), 26-35.

[9] Anagnoste, S. (2018). Setting up a robotic process automation center of excellence. *Management Dynamics in the Knowledge Economy*, *6*(2), 307-332.

[10] Ansari, W. A., Diya, P., Patil, S., & Patil, S. (2019, April). A review on robotic process automation-the future of business organizations. In *2nd International conference on advances in science & technology (ICAST)*.

[11] Leshob, A., Bourgouin, A., & Renard, L. (2018, October). Towards a process analysis approach to adopt robotic process automation. In *2018 IEEE 15th international conference on e-business engineering (ICEBE)* (pp. 46-53). IEEE.

[12] Torkhani, R., Laval, J., Malek, H., & Moalla, N. (2018, September). Intelligent framework for business process automation and re-engineering. In *2018 International Conference on Intelligent Systems (IS)* (pp. 624-629). IEEE.

[13] Mendling, J., Decker, G., Hull, R., Reijers, H. A., & Weber, I. (2018). How do machine learning, robotic process automation, and blockchains affect the human factor in business process management?. *Communications of the Association for Information Systems*, *43*(1), 19.

[14] Asatiani, A., & Penttinen, E. (2016). Turning robotic process automation into commercial success–Case OpusCapita. *Journal of Information Technology Teaching Cases*, *6*(2), 67-74.

[15] Willcocks, L., Lacity, M., & Craig, A. (2017). Robotic process automation: strategic transformation lever for global business services?. *Journal of Information Technology Teaching Cases*, *7*(1), 17-28.

[16] Habib, Muhammad Nauman, and Attaullah Shah. "Business process reengineering: Literature review of approaches and applications." In *Proceedings of 3rd Asia-Pacific Business Research Conference*, pp. 25-26. 2013.

[17] Zhang, Q., & Cao, M. (2002). Business process reengineering for flexibility and innovation in manufacturing. *Industrial Management & Data Systems*, *102*(3), 146-152.




DOI: https://doi.org/10.54216/AJBOR.000205