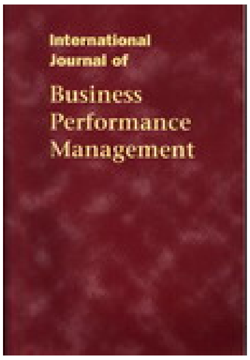


International Journal of Business Performance Management

ISSN online: 1741-5039 - ISSN print: 1368-4892
https://www.inderscience.com/ijbpm

ERP adoption prediction using machine learning techniques and ERP selection among SMEs

Aveek Basu, Rohini Jha

DOI: 10.1504/IJBPM.2023.10051659

Article History:

Received:	26 December 2021
Accepted:	05 September 2022
Published online:	01 March 2024

ERP adoption prediction using machine learning techniques and ERP selection among SMEs

Aveek Basu* and Rohini Jha

Birla Institute of Technology,
Mesra, India
Email: aveekbasu@gmail.com
Email: rohinijha@bitmesra.ac.in
*Corresponding author

Abstract: Small and medium scale industries (SMEs) have always been the backbone of a country's economy as they play a vital role in ensuring the goals such as balancing regional development, equality of income, and poverty alleviation by employment generation. However, SMEs are resistant to growth; facing challenges in sustainability in digital era. Others stay small and avoid taxation related problems. SMEs in developing nations are one of the most aggressive adopters of ERP packages. At times they have incurred huge capital expenses which ultimately raises a question mark in their survival on account of incorrect selection. ERP systems can benefit SMEs post COVID era. In this paper, machine learning techniques applied to predict adoption of ERP and multi-criteria decision-making technique (MCDM) applied for vendor and type of ERP selection viz. cloud ERP, on premise, or hybrid ERP will be suitable for SMEs and appropriate vendors.

Keywords: enterprise resource planning; ERP; cloud ERP; small and medium enterprise; SME; K-nearest neighbour; decision tree; machine learning algorithms; multi-criteria decision making.

Reference to this paper should be made as follows: Basu, A. and Jha, R. (2024) 'ERP adoption prediction using machine learning techniques and ERP selection among SMEs', *Int. J. Business Performance Management*, Vol. 25, No. 2, pp.242–270.

Biographical notes: Aveek Basu has 18+ years experience in top IT industry and currently pursuing his PhD. Currently, he has six Scopus-indexed journals are published or accepted.

Rohini Jha is an Assistant Professor at the Department of Management, BIT Mesra. She has several research works in her credit including publication in well-known journals.

1 Introduction

The small and medium enterprise (SME) sector contributes to a large extent to a nation's GDP by means of job creation, industrial output, export of quality products both in Indian and International markets, etc. SMEs had been more effective in generating much more employment opportunities as compared to large-scale industries which require huge investments. SMEs also render balanced economic development which percolates to the different sections of society and enhances equitable growth. It is evident that the future of India is the digital economy and SMEs need to join the digitisation to keep up pace with others. Enterprise resource planning (ERP) had made significant penetration in the SME sector but of late SMEs had faced a significant downturn in their business due to the above-mentioned challenges. SMEs are searching for effective ways to reduce IT costs and remain competitive in the volatile market.

SMEs are progressing in different areas like manufacturing, IT, pharma, retail, etc. and ERP can provide the necessary impetus in the path to progress. The return on investments (ROI) plays an important role in deciding whether ERP will be implemented in an SME or not. ERP helps in integrating different modules like inventory, production, finance, billing, human resources, etc. across the enterprise, and data is updated on a real-time basis. ERP provides benefits in reducing costs, efficient customer service, enhancement of productivity, improved quality, effective resource management, enriched decision making, planning, and holistic organisational development. Due to the aforesaid factors, almost all large companies have implemented ERP and different SMEs had been trying to adopt ERP to remain competitive in the market. From an SME perspective, ERP implementation is a big challenge considering the upfront investment, time, and failure risk it entails. ERP comes in packages and involves several modules.

It becomes a big challenge in decision making when SMEs decide to implement ERP. With the advent of cloud computing, the ERP industry had been undergoing a significant transformation. The on-premise ERP vendors who had shown initial reluctance in adopting the cloud are trying to leverage the Cloud computing benefits to entice customers, especially in the SME segment who had been a bit sceptical about implementing ERP due to high capital expenses. Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. Cloud computing which comes in different deployment models like public, private, and hybrid offers users to access software applications and infrastructure across the internet from anywhere.

Based on the cloud computing advantages the cloud-based ERP has emerged and there are now multiple cloud ERP vendors. Cloud-based ERP vendors are trying to provide similar solutions as an on-premise ERP. Cloud ERP is in a nascent stage as compared to cloud and ERP. Using the cloud computing benefits, cloud ERP is trying to provide customers with business process flexibility apart from other CAPEX, on demand, flexibility, etc. advantages. Cloud ERP model encompasses the features like integration of diverse information systems, improved decision making, enhanced business processes, ubiquitous accessibility along with on demand availability with payment flexibility as per usage.

A hybrid ERP system is a blend of on premise and cloud ERP systems that combines the public cloud and on-premise ERP modules. An organisation can keep the business-critical applications in on premise set up and operate other applications on cloud ERP. Additional cloud features can be plugged into the on premise ERP and can be available as per pay as you use the model. As mentioned in the research by Ruivo et al. (2015), it was observed that about 67% of the organisations are planning to move to hybrid ERP in the next 5–6 years. In terms of CAPEX, hybrid ERP scores better than on premise with reduced implementation cost.

SMEs had been facing challenges due to global competition as well as certain policy implementations in the domestic sector recently. To thrive in this critical atmosphere it's important to choose the right ERP which will ensure proper ROI and help the organisation to prosper in the competitive environment. It becomes difficult for SMEs to decide on the type of ERP and also the vendor. In developing nations where the IT infrastructure penetration is abysmally low, it becomes a challenge for SMEs to convince themselves of the utilities that ERP systems can provide. This paper applies the various machine learning techniques to determine whether the SMEs will adopt the Cloud-based ERP based on ratings of different features of ERPs provided by Gartner and PAT research. Based on the outcome of the prediction it will guide vendors and SMEs in understanding the adoption intention of different ERPs among SMEs with similar types of requirements. Once the adoption intention is predicted then the MCDM technique has been applied to select the ERP among the different vendors.

In current research work, the adoption intention of ERPs by SMEs and decision making has been analysed by different machine learning techniques and also MCDM Topsis applied in two different areas of concern viz. determining the ideal ERP suitable for SMEs among cloud ERP, on premise, hybrid ERP and then to determine the ERP vendor among the top few.

Section 2 refers to the literature review. Section 3 describes the problem statement. Section 4 describes the ERP implementation strategies and challenges among SMEs. Section 5 deals with machine learning techniques application and experimental results. Section 6 deals with fuzzy TOPSIS and experimental results. Section 7 concludes the paper.

2 Literature review

The literature review is done on reputed journal papers related to different ERPs in SME sectors. Though cloud computing research had been done over the years but cloud-based ERP research papers are comparatively fewer in number and very few research activities done on hybrid cloud ERP. Reviews of existing literature related to cloud, enterprise resource planning, cloud-based ERP, various machine learning techniques, SMEs have been conducted on acceptance analysis, selection, and application to identify the research gap.

Cloud computing has seen phenomenal evolution in the last decades in terms of adoption, acceptance, and popularity in myriad sectors across the industry. Cloud Computing has facilitated organisations with a paradigm shift from investing heavily in internally managed IT resources to a framework where enterprises can utilise the same IT resources that is scalable, ubiquitous, and have a pay-as-you-use facility managed by a cloud provider (Nayar and Kumar, 2018). Mahara (2013) has mentioned that several large organisations in India like ONGC, Tata Motors, Nerolac, Sony, etc. invested a substantial amount in ERP. Saini et al. (2012) highlighted that ERP implementation in large organisations has reached full capacity and the current ERP vendors have shifted the focus toward the SME market which comprises almost 95% of industrial units in India. Dezdar and Ainin (2011a) discussed the success of various industries after implementing ERP.

With the advent of a new paradigm called cloud computing, the ERP industry has been undergoing a significant transformation. In a recent study by Gartner in 2020, Oracle ERP Cloud has emerged as the market leader in 2019. The ERP Cloud has recently received a lot of traction in research activities and several companies are willing to take up cloud ERP. Tehrani and Shirazi (2014) found the influencing factors for cloud ERP adoption by SMEs and their study revealed that decision-makers knowledge of the cloud has significant weightage in the decision on cloud adoption.

Decision tree a machine learning technique can be employed by SMEs to address the dilemma of the choice of ERP implementation. Akhondzadeh-Noughabi et al. (2016) discussed performance management related to the decision tree and investigated different machine learning techniques in HRM. The decision tree analysis is a schematic representation of diverse decisions accompanied by numerous probabilities of occurrence. Decision tree analysis can be described as a tree-shaped graphical pattern of decisions that aid in analysing probable outcomes (Zhang et al., 2014a).

Decision trees are applicable in resolving regression and classification problems and can be relevant in handling quantitative and categorical values. Decision trees techniques can handle attribute missing values and can be leveraged in filling up the probable values. Thus the current research is suitable considering that there can be some missing data. Abdelaziz et al. (2018) discussed the optimal use of virtual machines for enhancing system performance using machine learning techniques. Fertiliser quality prediction and sharing info of the change of fertiliser in the market. Applying the ranking method fertilisers standard is determined in research done by Gladence et al. (2021). Banerjee and Roy (2017) developed a model based on data mining techniques like LDA, classification tree, etc. to determine the best model among the used cars.

In studies by Dwivedi (2018) six different ML techniques are discussed viz. k-nearest neighbour (kNN), support vector machine (SVM), LR, classification tree, artificial neural network (ANN), and naïve Bayes on CAD data. Pandya et al. (2022) discussed the readiness of the young generation in the domain of artificial intelligence. The results are analysed using receiver operating characteristic (ROC) curves. ANN and LR had the highest accuracy as per the results. A study was conducted by Ayatollahi et al. (2019) on 1,324 heads at AJA University. SVM and ANN algorithms were applied for the normalisation of data. Different machine learning methods like SVC, linear SVM, and nuSVM were tested by Abdar et al. (2019). In the medical domain, Akella and Akella (2021) worked on six different ML algorithms to determine patients with heart-related issues. All the six methods had good accuracy and the neural network performed the best among the six. Cüvitoğlu and Işik (2018) applied principal component analysis (PCA) and different machine learning techniques while testing data during feature selection. Kutrani and Eltalhi (2018) applied machine learning techniques in the dataset of 1,770 individuals collected from Benghazi Heart Center. In the study, it has been found that KNN and SVM classifications are better compared to other methods. Similar studies on heart ailments have been done by Naushad et al. (2018) on 648 patients by applying ML algorithms. The specificity and sensitivity of each algorithm were analysed.

Linear regression or LR is applied when the research technique is centred on whether or not an event happened, rather than when it happened. It is specifically applicable for models concerning the state of disease (healthy or deceased) and in making decisions and thus is extensively used in research in the health sciences. According to Hosmer et al. (1989), LR does not presume a linear relationship between dependent and independent variables. Prediction of postoperative sepsis and kidney injuries can be effectively done by applying logistic regression according to Thottakkara et al. (2016). LR has been successfully applied in forecasting bankruptcy among Slovakian companies by Kovacova et al. (2018). Failure of the machine and degradation can be successfully predicted using LR as per research done by Caesarendra et al. (2010). Logistic regression has been applied in determining groundwater contamination by Mair and El-Kadi (2013). Discrimination between shallow and deep-induced micro-earthquakes using the LR technique has been done in studies by Mousavi et al. (2016). Palvanov et al. (2018) analysed the execution of logistic regression models in real-time to determine their efficiency. Bos et al. (2014) studied prediction on encrypted data using LR. The study was conducted on an existing trained model and thus it didn't consider the model training process. Slavkovic et al. (2007) applied secured logistic regression on horizontal and vertical partitioned datasets by applying multi-party computational techniques. The data quality aspect was not taken into consideration in this study. Bootstrapping and homomorphic encryption were employed to train the logistic regression model using encrypted data (Han et al., 2019). A proposed scheme that is computationally intensive was tested to predict encrypted data.

The naïve Bayes classification technique has been widely explored since the 1950s, and is a type of probability classifier that applies the Bayes theorem, and presumes autonomy between features (Park, 2016). It has the capability in judging spam mail and analysing medical images of patients. The naïve Bayes classifier is remarkably efficient in practice as its categorisation decision may frequently be accurate even if its probability estimates are incorrect (Narang, 2013). Lately, in the weather categorisation area, weighted naïve Bayes classifiers accomplished improved results compared to traditional naïve Bayes. It is a technique of using covariance as a weight to enhance the theory of autonomy between characteristics. Farid et al. (2014) suggested a hybrid algorithm for a naïve Bayes classifier to enhance categorisation accuracy in multi-class categorisation tasks. In the hybrid naïve Bayes classifier, a decision tree is employed to discover a subset of critical qualities for classification, with the subsequent weights providing exponential parameters in determining the conditional probability of the class. Huang et al. (2014) performed a different approach to resolve rule conflicts in the naïve Bayesian model. A different approach, associative classification with Bayes (AC-Bayes), has been used to resolve rule conflicts in the naïve Bayesian model (Huang et al., 2014]). Hadi et al. (2018) suggested a new-found hybrid AC algorithm (HAC) in which the naïve Bayes algorithm was utilised to lessen the number of classification rules demonstrating each of the attribute values, thus enhancing the classification precision.

Extensive research work has been carried out on ERP till date but none of the research papers have utilised the various machine learning techniques in prediction and decision making. In research by Salum and Rozan (2016) the challenges and drivers related to cloud ERP adoption among SMEs are mentioned. Mehta et al. (2018) performed a brief analysis of several research papers related to cloud ERP adoption and challenges faced by SMEs and relevant factors responsible for cloud ERP adoption. Hedau et al. (2013) discussed on the SME's contribution in India and describes the suitability of cloud ERP for SMEs. Johansson et al. (2014) analysed the opportunities and challenges of cloud ERP for SMEs and large companies. Cloud ERPs may not be beneficial to large companies as compared to SMEs. Mahara (2013) developed a framework to determine the opportunities and threats based upon three parameters viz. economical, technical and human aspects for cloud-based ERP in SMEs. Sharma et al.

(2010) in their paper compared the per-user cost between on-premise and cloud ERP on annual basis and also compared the complexity level while adapting cloud ERP and on-premise ERP by SMEs. Al-Johani and Youssef (2013) proposed a framework related to Cloud-based ERP. It provides a percentage of reduction of different costs like technical support, testing effort, overall expenses, etc. for cloud-based ERP. Gupta et al. (2018a) conducted an empirical test to determine organisation performance with cloud-based ERP using a theoretical framework based on the contingent resource-based view. Saini et al. (2014) compared cloud-based ERP and on premise ERP using two hypotheses developed in the research. T-tests result in an annual cost per user is much in cloud-based ERP as compared with on premise for SMEs. Seethamraju (2015) carried out a study on study organisations to determine the benefits and challenges in adopting SaaS-based ERP as also some important determining factors like customer support in the product lifecycle, software vendors' reputation, etc. Elmonem et al. (2016) followed the systematic literature review research method while finding the key benefits like reduced cost, scalability, availability, upgrades, etc. and challenges such as security issues, SLA, data ownership, hidden costs etc. Taghipour et al. (2020) discussed the assessment and risk management in an ERP implementation project in Iran that will aid in reducing time and effort. Gupta et al. (2018b) applied resource dependency theory (RDT) to realise the relationship between cloud ERP service providers and 208 SMEs using structural equation modelling. Menon (2019) analysed the critical challenges faced by organisations during ERP implementation. In 2019, a quantitative study was done on ERP related to risk mitigation. Critical success factors of ERP implementation are discussed in the research performed by Kiran and Reddy (2019). Bjelland and Haddara (2018) presented research work on cloud ERP evolution and system updates. Estefania et al. (2018) published a literature review on the integration of inter and intra organisational information system and ERP. Panarama in 2018 provided a detailed analysis including best practices of the latest ERP implemented systems vis. S4 Hana, MS Dynamics, Oracle Cloud and Infor CloudSuite. Hietala and Päivärinta (2021) presented a case study of the various benefits obtained by the organisation after implementation ERP system. Hasan et al. (2019) applied Smart-PLS package in analysing the factors responsible for success in post ERP implementation phase. Lee and Chang (2020) discussed the capabilities driving the adaptation of ERPs in post implantation phase.

Multiple criteria decision-making (MCDM) method has been applied in various areas of decision making in the last couple of years and it has been proven effective and efficient. Over the period of past several years and in diverse research articles the MCDM techniques have been employed to solve numerous problems. Sandhya et al. (2018) applied a multi-attribute decision-making problem along with a distance-based approximation method in selecting the most suitable CSP and successfully compared the results with AHP methods. Talbi and Haqiq (2020) developed a cloud broker architecture based on a rough set algorithm and multi-criteria mathematical model in determining the most suitable cloud service provider. Trueman and Narayanasamy (2018) investigated the selection of cloud service providers which will benefit small and medium enterprises by applying the three-pronged technique considering that there is a lack of technical acumen among the enterprises in selecting the best suited service provider. Md and Vijaykumar (2019) established that the existing trust assessment methods of CSPs disregard all attributes of quality of service (QoS) and it transforms regularly.

3 Problem statement

The impact of COVID had been phenomenal across different industries and particularly the SME sector has suffered a huge blow in their progress. Post-COVID era the turnaround in economies can happen through the SME sector and everywhere governments have earmarked huge capital to aid the ailing SMEs. In order to keep the momentum going SMEs need to utilise information technology in an effective way for efficient delivery. ERP software can significantly help SMEs in dealing with these obstacles and provide the desired ROI in short term provided the right type of ERP is chosen for a particular organisation. Due to budget constraints, some SMEs might find it difficult in implementing on-premise ERP due to high capital expenses. SMEs can be enticed towards cloud ERPs due to lower capital, operating expenses, and other advantages. But cloud ERP comes with certain disadvantages like security, customisation, service level agreement (SLA), etc. Hybrid ERP which is a blend between on-premise and cloud ERP can mitigate certain disadvantages of both on-premise and cloud ERP. Hybrid ERP is in its nascent stage and it can provide benefits to SMEs where already on-premise ERP is in place and new components like data analytics, HR applications, etc. can be plugged into the existing ERP instead of outright new purchases by replacing the existing one. It becomes difficult for SMEs to decide on the right ERP which will be tailor-made for their organisation. The research papers published to date have discussed different advantages and disadvantages of on-premise and cloud ERP but none of the papers were able to provide a framework on which decisions can be taken. It becomes difficult to decide which type of ERP will be suitable for an SME based on the different advantages and disadvantages of the three types of ERP. In this paper, we have applied machine learning techniques to determine the adoption intention of different ERPs among SMEs and also implemented a multi-criteria decision-making technique to find out the suitable ERP among on-premise, cloud and hybrid for SMEs. Once the type of ERP has been determined then we will be applying the MCDM technique to determine the suitable service provider for that selected ERP.

4 ERP implementation strategies, challenges, threats and benefits towards Indian SMEs

SMEs play a significant role in an economy, and they help in poverty alleviation by employment generation. With the advent of globalisation, SMEs are facing challenges from different sectors in the international market. ERP can help SMEs by integrating the different business processes and divisions into one centralised repository. ERP has the capability of boosting productivity, reducing operating costs, and agility, and enhancing decision making and long-term planning.

ERP implementation involves a good amount of planning, time, and effort. The implementation is done in phases starting from requirement gathering, blueprinting, planning, developing, building, testing, deploying, go-live, hyper care, and post-go-live support. The implementation strategy plays a crucial role in ERP implementation as any wrong decision can lead to a disastrous outcome. The implementation strategies encompass the following areas like defining requirements, evaluation of requirements gathered, documentation of process and procedures, single or phased roll-out, strategy on existing legacy systems, customisations as per business needs, selecting the project leader to steer the project, training of existing support staffs, etc.

ERP implementation comes with a set of challenges. The challenges related to capital expenses are very high for SMEs. Apart from the implementation cost, the ERP needs to be upgraded from time to time to ensure they receive proper support from the vendor. There had been numerous examples where ERP implementation had been a failure and it had a cascading impact on the SME.

The SME sector to remain competitive requires ERP for the different benefits it provides. Integration of information flow among different departments and business processes can be streamlined post-ERP implementation. This in turn will benefit the SMEs in strategic decisions with enhanced financial implications, delivery of finished products, inventory and warehouse management, timely procurement of raw materials, invoicing, and other related business processes. Productivity gets a significant boost as there is timely coordination among the different departments. With the development of business, complex business processes can be managed by ERP only as its difficult to maintain data in excel or in locally available accounting software. ERPs provide flexibilities and agility which helps SMEs to transform as per change imposed by government or policy change. Decision making forms an important aspect in the ever-changing current world scenario. The ERP will help SMEs in taking fast decisions with proper precision. SMEs with ERP will have stringent security features which in turn will help in gaining market share and long-term planning.

5 Experimental results – implementing machine learning techniques

Machine learning is a branch of computer science and artificial intelligence that emphasises the application of algorithms and data to replicate the way humans learn gradually and improve precision. It allows software applications to predict outcomes with accuracy without explicit programming. Machine learning algorithms apply historical data as input while predicting new output values. Supervised, Unsupervised, and reinforcement learning are different types of machine learning techniques.

5.1 Comparative studies between the different techniques

5.1.1 Logistic regression vs. decision tree

- Collinearity is better handled in DT as compared to LR.
- The significance of features can be derived in LR unlike in DT.
- DT performs well compared to LR in terms of categorical values.

5.1.2 Logistic regression vs. neural network

- LR cannot support nonlinear solutions but NN can.
- LR will hang in local minima due to convex loss function property but NN might hang.
- NN requires significant training data but LR performs best with a large number of features and less training data.

5.1.3 Logistic regression vs. KNN

- LR is a parametric model whereas KNN is a non-parametric model.
- LR is comparatively faster than KNN.
- LR works on linear solutions only but KNN supports nonlinear solutions.
- KNN can only yield the labels whereas LR can produce confidence levels related to its predictions.

5.1.4 KNN vs. naive Bayes

- Due to KNNs real time execution, it is much slower compared to naïve Bayes algorithm.
- KNN is non-parametric but naïve Bayes considers parametric values.

5.1.5 Decision tree vs. KNN

- DT and KNN are non-parametric methods.
- KNN wont support automatic feature interaction like DT.
- KNN's expensive real-time execution makes it slower as compared to DT.

5.1.6 Decision tree vs. naive Bayes

- Naive Bayes is a generative model but DT is a discriminative model.
- Naïve Bayes are more complex compared to DT.
- Decision tree pruning may disregard certain crucial values in training data, which may lead to some inaccuracy.

5.2 Contributions of different Machine learning techniques

Decision trees can be applied as a useful method to extract logical rules from data that can be employed to support prediction as well as selection. Accuracy, support, and complexity of the rule are considered in this process. Decision tree algorithms are applied in assessing potential growth prospects for business as per historical records. Business strategies can be altered based on the outcome of the decision tree when historical sales data is provided as input which will lead to the growth and expansion of the business. Decision trees can be applied to demographic data to find prospective clients. Proper decisions can be done on the target market relevant to the current business. In the domain of logistics and strategic management, a decision tree can be successfully applied while achieving the intended goals. Decision trees can predict the probability of loan default by customers based on past records. Customers' creditworthiness using decision trees can help organisations in loss prevention. The interpretation of the decision tree is simple and there is no need to have a statistical background. It's been found that decision trees take less effort as compared to other decision techniques. Classification of data can be done without computing complex calculations. Decision trees can be combined with other methods to perform complex calculations. Decision trees outcome is not affected much if there are missing values and outliers in data. Once the variables are created in the decision tree then less data cleaning is needed.

In KNN algorithm the data is classified by determining the K closest data points. As per the distance and majority voting between neighbours the data class prediction is done. The performance of KNN algorithm is determined based on the value of K and the type of distance. It's observed that the average accuracy of the KNN algorithm is very high but a bit lesser compared to the decision tree technique. In the case of prediction where a rapid decision needs to be taken, it's observed that KNN doesn't perform well. This algorithm is preferred while trying to determine similarities among instances. The major disadvantage of the KNN algorithm is due to high space and time complexity. The time that a model takes to determine the class of query points refers to time complexity. Space complexity refers to the memory taken up by the algorithm. KNN can be applied in detecting outliers as in credit card fraud detection. Semantically similar objects identified as vectors can be searched using KNN. KNN can be applied in recommendation systems as a basic approach and not for high dimensional values. KNN algorithm in certain cases fails to identify outliers if data points from the existing classes are present far away.

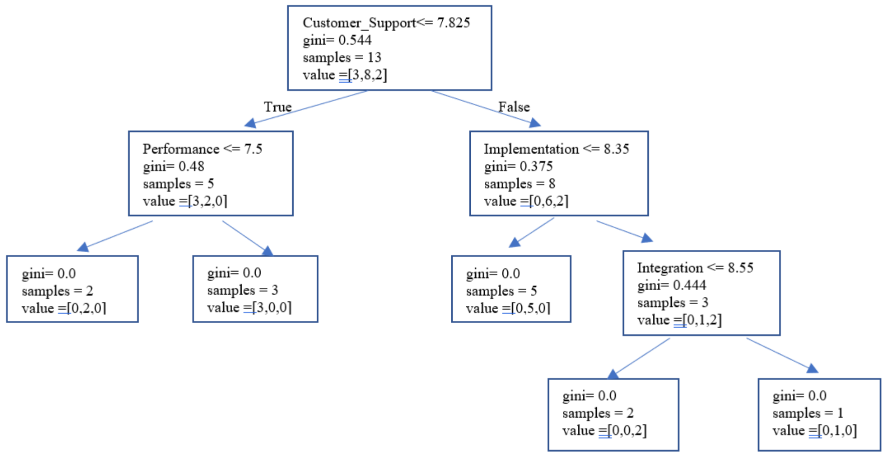
In logistic regression, there are fewer chances of overfitting apart from high dimensional datasets. The problem of overfitting can be lessened by applying the regularisation technique. This technique performs well on basic datasets that can be separated linearly. Compared to other techniques, Logistic regression is simple to understand, apply and train. In this technique, unlike Decision trees, models can be changed to incorporate the most recent data. With adequate training instances, logistic regression is less probed to overfitting. Performance increases when dataset characteristics are linearly separable. Logistic regression almost resembles neural networks. A neural network appears to be a collection of tiny logistic regressions assembled together. Compared to Artificial Neural Network and other sophisticated algorithms, the logistic regression method training time is relatively shorter because of straightforward probabilistic interpretation.

Naïve Bayes classifier is a machine learning technique that employs the Bayes theorem for probabilistic classification. By examining the input values of a given set of parameters or features, characterised as B in an equation, the Naïve Bayes classifier is capable to determine the probability of the input data belonging to a certain class A. It can handle high-dimensional data and big data in a better way. When the dataset is small, the performance of this algorithm is better than other complicated models.

Table 1 Source dataset

Ease of use	*Features and functionalities*	*Advanced features*	*Integration*	*Performance*	*Customer support*	*Implementation*	*Renew and recommend*	*OutClass/Variable*
8.25	8.55	8.35	8.3	6.45	8.5	8.0	8.3	SAP Business ByDesign (C)
8.4	7.65	8.25	7.85	8.5	7.85	8	9.4	Dynamics 365 (C)
8.0	7.85	8.0	8.1	6.85	6.65	5.5	4.9	Net Suite ERP (C)
8.45	8.35	8.6	8.35	8.6	8.2	8.9	9.3	JD Edwards Enterprise One (H)
8.45	8.15	8.5	7.6	8.4	7.8	8.2	8.0	Epicor ERP (O)
8.3	8.1	8.3	7.75	8.15	7.85	8.0	8.4	IQMS (C)
8.0	8.6	8.45	8.3	8.25	8.4	7.8	7.9	IFS ERP (C)
7.75	8.15	8.2	8.1	7.55	7.25	8.0	8.5	SAP ERP (O)
7.35	7.55	7.95	8.45	8.45	8.35	8.7	7.5	SAP S/4 Hana (H)
8.1	7.95	8.15	8.1	7.45	7.75	8.0	8.0	SAP S/4 Hana Cloud (C)
8.35	8.75	7.25	8.65	8.2	8.5	8.9	8.2	RAMCO ERP (C)
8.0	8.4	8.3	8.4	8.15	8.35	8.0	7.7	Acumatica Cloud ERP (C)
8.1	8.05	8.25	7.8	7.8	7.8	8.0	9.0	Deltek Vision (O)

Figure 1 Decision tree output (see online version for colours)



5.3 *Decision tree technique to determine adoption intention of ERP type based on features*

The application of supervised learning techniques like decision tree has been applied in different real-life scenarios and across various industry domains. The decision tree consists of the following:

- starting node or root node
- internal node consisting of multiple outgoing edges and one incoming edge
- leaf node or terminal node with no out-going edge.

The approach outline is given in the following:

1. the best feature selected in every stage or node as per test condition
2. based on possible outcomes, the nodes are split into forming internal nodes
3. till all test conditions are exhausted into terminal node the above steps are repeated.

In the current research the Gini index has been implemented to ascertain the root node. The Gini coefficient is having a value between 0 and 1. Measure of inequality of distribution can be done using Gini coefficient. The binary split of each attribute is done by Gini index. The attribute with minimum Gini index determines the splitting attribute.

To construct the model the initial rows of the above are used as a training dataset and the accuracy is cross-checked by the final record. Built-in classes like Tree and DecisionTreeClassifier are applied in Python to develop the decision tree.

The source data has been collected from Gartner (Service-Centric Cloud ERP Solutions Reviews 2021 | Gartner Peer Insights) and B2B platform PAT research.

The steps involved are:

1. The continuous variables are converted into categorical variables.

2 The categorical variables are then converted into groups like >= 7 && < 8 is considered as 0, range >= 8 && < 9 is considered as 1 and range >= 9 && < 10 is taken as 2.

3 The ERP providers are segregated into groups where 0 represents on premise, 1 represents cloud ERP and hybrid ERP is 2.

4 As per the table we have three on-premise ERPs, eight cloud ERPs, and two hybrid ERPs.

5.3.1 Discussion on the results

The root node is decided is the customer support according to the Gini index and it has all the datasets of 13 different types of ERP. The splitting node is customer support which is divided into two groups. If Customer_Support value <= 7.825 is True then it branches further till the Gini value is 0. The next splitting attribute is selected based on the Gini index value and based on the performance attribute. If the performance value <= 7.5 is true then the type of ERP chosen is cloud-based ERP. If the value is False then it's on-premise ERP. No further branching will be performed on the left part as the final leaf node has been formed with a Gini value equal to 0. A similar iteration is performed based on splitting attribute implementation <= 8.35.

Table 2 Adoption rule sets from decision tree model

Rule no.	*Rule description*	*Rule explanation*
Rule 1	If (*customer support* <= 7.825), and (*performance* <= 7.5) then the ERP type is cloud ERP.	If (*customer support* is less than 7.825) along with *performance* value less than 7.5 the SMEs will adopt cloud-based ERP.
Rule 2	If (*customer support* <= 7.825), and (*performance* > 7.5) then the ERP type is on premise ERP.	If (*customer support* is less than 7.825) along with *performance* value greater than 7.5 the SMEs will adopt on premise ERP
Rule 3	If (*customer support* > 7.825) and *implementation* <= 8.35) then the ERP type is cloud ERP	If (*customer support* greater than 7.825) and *implementation* less than 8.35) then the SMEs will adopt cloud ERP.
Rule 4	If (*customer support* > 7.825) and *implementation* > 8.35 and *integration* <= 8.55) then the type of ERP is hybrid ERP.	If (*customer support* greater than 7.825) and *implementation* greater than 8.35 and *integration* less than 8.55) then SMEs will opt for hybrid ERP.
Rule 5	If (*customer support* > 7.825) and *implementation* > 8.35 and *integration* > 8.55) then type is cloud ERP.	If (*customer support* greater than 7.825) and *implementation* greater than 8.35 and *integration* greater than 8.55) then type is cloud ERP.

5.4 Distance-based method – nearest neighbour technique

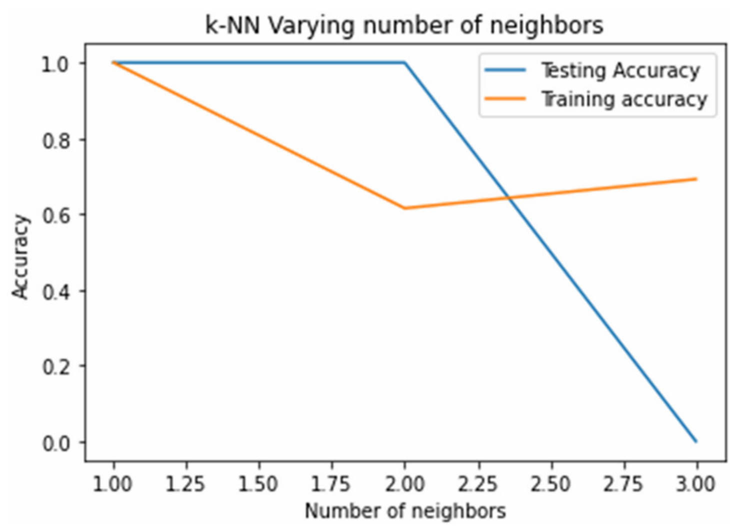
KNN or K-nearest neighbour is a technique in supervised learning that accumulates all available training data as long as new data enters for testing. Categorisation of the new case is performed by similarity measure that may represent a distance function. In regression and classification problems the KNN technique can be utilised. In the current research, this is applied.

The steps of the technique are as follows:

- The training dataset is categorised in the form of a feature vector which comprises the input to the KNN model and they form unique data points.
- The majority vote of nearest K neighbours determines the output in KNN classification and the object assigned to the class which is very common among them. An object is assigned to the class of single nearest neighbour when $K = 1$.
- According to the KNN regression, the new object output value is determined by the average of the nearest K neighbour's output values.
- In determining the adoption of ERP for SMEs the KNN algorithm can be successfully implemented likewise it is applied in other research areas as well.

The concept is applied in detecting outliers including the style applications of sensors and surveillance (Knorr et al., 2000). In these scenarios, the patterns forming the majority are not as valuable and noteworthy as the exceptions. It is the onus of the system to determine the outliers or the path that are deviating from the usual scenario and generate alerts.

Figure 2 KNN model graphical output (see online version for colours)



5.4.1 Discussion on the results

KNN algorithm based on Euclidean distance of nearest neighbours works best when the data size is small. In the above graph, the training dataset is formed based on the table source dataset. Once the model is trained using the data provided in the table source dataset, the same is tested on the last row of the Table and it has been found to provide the expected outcome. In Figure 2, its observed that there is a change in the pattern of the training line and testing at value approx. to 2.1 value along the X-axis. Thus it validates the accuracy of the model. Based on the different features SMEs can take decisions on the type of ERP to be adopted.

5.5 Logistic regression

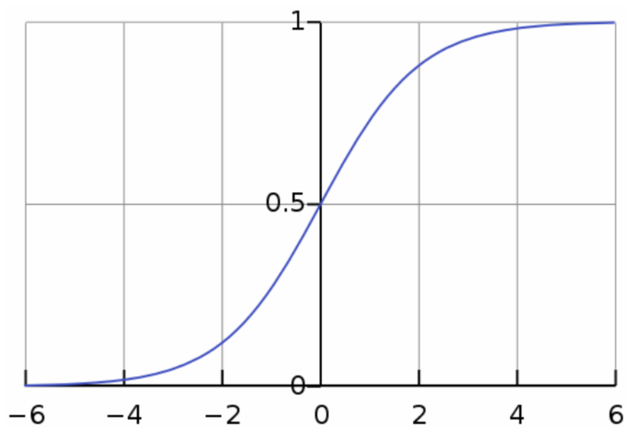
In a scenario where the dependent variable is discrete, the logistic regression technique is applied. Example of logistic regression is true or false, 0 or 1, etc. signifying only two possible values of the target variable. A sigmoid curve indicates the relationship between the independent variable and the target variable.

The relationship between the independent and target variable is determined by the logit function in logistic regression. The equation is given below:

$$\text{logit}(p) = \ln(p/(1-p)) = b0 + b1X1 + b2X2 + b3X3 + \ldots + bkXk$$

where p is the probability of occurrence of the feature.

Figure 3 Logistic regression (see online version for colours)



Logistic regression can be suitably applied when the volume of data is large and there is an equal occurrence of values in target variables. The data should be free from multicollinearity, i.e., in the dataset there should not be any correlation between independent variables.

5.5.1 Discussion on the results

Initially data analysis has been done to check there is no junk value or missing value in the data. Next data wrangling, i.e., cleaning data and validation of data done to check if there is any null value. Here as per the algorithm, the value lies between 0 and 1. Any value above 0.5 is considered 1 and anything below 0.5 is considered 0. The training dataset is formed from the source table and it is verified with the test data and expected result obtained. The logistic regression will assist to determine based on different features whether an SME will adopt cloud ERP or not. A classification report is generated to identify the precision.

5.6 Naïve Bayes algorithm

This is a type of classification technique where predictors are assumed to be independent. In the naïve Bayes technique, a particular feature in a class is assumed to be unrelated to the occurrence of any other feature.

As an example, fruit can be considered guava if it is round, green and two inches in diameter. These features might depend on each other but they independently promote the probability that this fruit is guava and hence it is known as naïve.

The technique is represented below:

$$P(c|x) = \{P(x|c)P(c)\}/P(x)$$

where

- P(c|x) is posterior probability of class (c, target) given predictor (x, attributes)
- P(x) is predictor prior probability
- P(c) is class prior probability
- P(x|c) is likelihood.

$$P(c|x) = P(x_1|c) * P(x_2|c) * \ldots * P(x_n|c) * P(c)$$

Steps involved:

Step 1 Formation of frequency table after converting the dataset.

Step 2 Creation of likelihood table.

Step 3 Apply naïve Bayesian theorem to calculate the posterior probability of each class. The highest probability class is the prediction outcome.

5.6.1 Discussion and results

This technique is applied in problems with multiple classes and in text classification. It is also applied in real-time prediction, multi-class features, sentiment analysis, etc. Here in our current research, this is applied to train the model and then apply it to the test data to determine the accuracy. The training dataset was formed from the top 12 rows of the source dataset and it was verified against the last row and the expected outcome was there.

6 Applying MCDM techniques by TOPSIS

MCDM method has been employed in various areas of decision making in the last couple of years and it has been proven effective and efficient. Over the period of past several years and in diverse research articles the MCDM techniques have been employed to solve numerous problems. One needs to find out the most suitable option from available alternatives. With conflicting situations, constraints, and unforeseen end results decision making becomes a challenging task in a real-world scenario.

6.1 MCDM techniques and comparative studies

Table 3 ERP Benefits discussed in research journals over a period

MCDA compare	*TOPSIS*	*ELECTRE*	*VIKOR*	*SAW*
Calculation process	Medium	Complex	Medium	Easy
Features	Compromise method	Non-rank method	Compromise method	Scoring method
	Applies the shortest distance from the ideal solution and the remotest distance from the counter-ideal solution for the selected option	Parallel comparison of criteria and exclusion of defeated criteria.	The method considers the importance of the optimal distance to the best and worst case in calculating the distances of the options.	After weighing the applied variable, ranking is executed.
Output results	The nearer the coefficients to one, the more significant, and the closer the coefficients to zero, its less significant.	Priority is determined based on more wins and losses.	The closer the coefficients are to zero, the more significant, and the closer the coefficients are to one, the less significant.	It is more significant if coefficients are closer to one and less if it's close to zero.

6.2 Outline of TOPSIS technique

TOPSIS is an acronym and it refers to technique for order preference by similarity to ideal solution. Ideal positive and ideal negative are two hypothesised alternatives applied in TOPSIS. TOPSIS method determines the alternative that is farthest from the negative ideal solution and nearest to the positive ideal solution.

Step 1	Evolution matrix is developed using n criteria and m alternatives, using the intersection of each criterion and alternative given as , and then we have a matrix $(x_{ij})_{m\times n}$
Step 2	The matrix is formed by normalisation of $(x_{ij})_{m\times n}$. R = $(r_{ij})_{m\times n}$ using the normalisation method $r_{ij} = \frac{X_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}}$, $i = 1, 2, \ldots, m; j = 1, 2, \ldots., n$
Step 3	Then the weighted normalised decision matrix is calculated $t_{ij} = r_{ij} \cdot w_j, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n$ where $w_j = W_j \Big/ \sum_{j=1}^{n} W_j$, $j = 1, 2, \ldots, n$ so that $\sum_{j=1}^{n} w_j = 1$ and W_j is the original weight given to the indicator v_j, j = 1, 2, …, n
Step 4	The best alternative (and the worst alternative ( ) is determined. $A_w = \{(\max(t_{ij}\|i = 1, 2, \ldots, m)\|j \in J_-), (\min(t_{ij}\|i = 1, 2, \ldots, m)j \in J_+)\}$ $\equiv \{t_{wj}\|j = 1, 2, \ldots, n\}$,

$A_b = \{(\min(t_{ij}|i = 1, 2, \ldots, m)|j \in J_-), (\max(t_{ij}|i = 1, 2, \ldots, m)j \in J_+)\}$
$\equiv \{t_{bj}|j = 1, 2, \ldots, n\}$,

where

$J_+ = \{j = 1, 2, \ldots, n|j$ associated with the criteria having a positive impact

$J_- = \{j = 1, 2, \ldots, n|j$ associated with the criteria having a negative impact

Step 5 Calculate the L2 – the distance between the alternative i and the best ideal condition A_b

$$d_{ib} = \sqrt{\sum_{j=1}^{n} \left(t_{ij} - t_{bj}\right)^2}, \; i = 1, 2, \ldots, n$$

and distance between the target alternative i and the worst condition A_w

$$d_{iw} = \sqrt{\sum_{j=1}^{n} \left(t_{ij} - t_{wj}\right)^2}, \; i = 1, 2, \ldots, m$$

where d_{ib} and d_{iw} are L2 – distances from the target alternative i to the best and the worst conditions, respectively.

Step 6 Calculate the similarity to the worst condition:

$$s_{iw} = d_{iw} / \left(d_{iw} + d_{ib}\right), 0 \leq s_{iw} \leq 1, \; i = 1, 2, \ldots, m$$

$s_{iw} = 1$ if and only if the alternative solution has the best condition

$s_{iw} = 0$ if and only if the alternative solution has the worst condition.

Step 7 Rank the alternative according to s_{iw} ($i = 1, 2, \ldots, m$).

6.3 Source data

The ERP benefits and challenges of different types of ERP had been discussed in different journal papers. The following data has been collected from by Elmonem et al. (2016). Here it deals with 31 journal papers over a period of 6 years and important factors influencing the decision on whether to select cloud-based ERP, hybrid or on premise. There are multiple benefits and challenges associated with different ERPs but from this we have deduced the critical factors. If we consider the Benefits table then we can determine the areas which SMEs will consider during decision making. The leading benefits can be determined from the percentage in the agree column. Lower upfront costs, lower operating costs, rapid implementation, improved accessibility, mobility and usability, etc. have here percentage value and hence these should be considered first before other factors in benefits. Likewise from challenges table we should consider security risks, performance risks, customisation and integration, data ownership, compliance are critical factors to be considered during decision making.

Table 4 ERP Benefits discussed in research journals over a period

ERP benefits	*Number of papers*	*Agree %*
Lower upfront costs	30	96.77419
Lower operating costs	26	83.87097
Rapid implementation	21	67.74194
Scalability	20	64.51613

Table 4 ERP Benefits discussed in research journals over a period (continued)

ERP benefits	*Number of papers*	*Agree %*
Focus on core competencies	10	32.25806
Using advanced technology	7	22.58065
Rapid updates and upgrades	13	41.93548
Improved accessibility, mobility, and usability	21	67.74194
Easier integration with cloud services	5	16.12903
Improved system availability and disaster recovery	12	38.70968
Cost transparency	2	6.451613
Sales automation	1	3.225806
Using security standards	1	3.225806
Free trials	1	3.225806

Table 5 Challenges discussed in research journals over a period

Challenges	*Number of papers*	*Agree %*
Subscription expenses	4	12.9032258
Security risks	23	74.1935484
Performance risks	19	61.2903226
Customisation and integration limitations	19	61.2903226
Strategic risks	5	16.1290323
Compliance risks	7	22.5806452
Loss of IT competencies	4	12.9032258
Functionality limitations	5	16.1290323
SLA issues	11	35.483871
Sensitivity of the information	4	12.9032258
Control over cloud ERP	7	22.5806452
Hidden costs in the contract	2	6.4516129
Loss of technical knowledge	1	3.22580645
Data ownerships	8	25.8064516
Need for ERP as service standards and regulations	4	12.9032258
Knowledge about the cloud	3	9.67741935
Startup support	1	3.22580645
Organisational challenges	2	6.4516129
Choosing between cloud ERP systems	1	3.22580645

The following subsection describes the salient benefits and challenges attributes in brief.

- *Upfront costs:* the upfront cost is related to capital expenses that an organisation requires during the start of business.

- *Operating costs:* the cost is incurred by an enterprise in its operation process on regular basis.
- *Implementation duration:* ERP implementation is a time bound activity due to the complexity of the software and duration depends on the type of ERP chosen.
- *Scalability:* the scalability attribute is related to the resource utilisation which can scale up during peak period and might scale down during the non-peak period.
- *Improved accessibility, mobility and usability:* this refers to the ERP system accessible from different locations apart from the Enterprise location. Mobility and usability grow once the accessibility increases for an organisation.
- *Customisation and complex integration:* ERPs are not tailor made for a particular enterprise. As per business needs the ERP needs to undergo customisation. For proper integration with the existing business
- *Upgrades:* ERPs systems need to undergo upgrades from time to time. Through upgrades new features are incorporated into the ERP systems.
- *Security risks:* ERP deals with all data related to an enterprise. Security is one of the vital factors in decision making while choosing type of ERP.
- *Performance risks:* ERP service providers are developing the software on regular basis to improve the system performance. Companies will face challenges if the ERP performance is not up to the mark.
- *Customisation and integration:* ERPs are delivered in packages which requires lot of customisation post implementation in an organisation. The existing legacy systems or any external system needs to be integrated with the implemented ERP on various occasions.
- *SLA:* service level agreement or SLA is an important factor related to terms and conditions like service quality, delivery, maintenance etc. between ERP vendor, customer and third party broker.
- *Data ownership:* data ownership refers to the legal right of an owner to store, process, modify, share etc. a set of data. Data owners can allow the above mentioned functions to other third party vendors to manage data on their behalf.

Table 6 is developed based on the data collected from research by Elmonem et al. mentioned above and also from Duan et al. (2013). The same is used as the input of the TOPSIS algorithm to find the most effective one.

The eight criteria mentioned above are not having same importance. The relative weights of the criteria are determined as per priority given in existing research papers. Costs like upfront costs, operating costs, security risks, performance, customisation are given higher priority as compared to other features like SLA, usability, etc. Table 7 represents the same.

Table 6 Comparison of features among various ERP types

ERP types	*Cost*	*Rapid implementation*	*Scalability*	*Usability*	*Security risks*	*Performance*	*Customisation*	*Service level agreement*
Cloud ERP	9	7	6	7	3	4	4	6
On premise ERP	4	3	3	5	9	8	7	7
Hybrid ERP	7	4	8	6	4	6	4	6

Table 7 Determining weights for the criteria

Cost	*Rapid implementation*	*Scalability*	*Usability*	*Security risks*	*Performance*	*Customisation*	*Service level agreement*
9	7	7	7	8	8	8	5

Table 8 Various organisations and their feature weightage

Features ERP providers	*Ease of use*	*Features and functionalities*	*Advanced features*	*Integration*	*Performance*	*Customer support*	*Implementation*	*Renew and recommend*
SAP Business ByDesign	8.25	8.55	8.35	8.3	6.45	8.5	8.0	8.3
Dynamics 365	8.4	7.65	8.25	7.85	8.5	7.85	8	9.4
Net Suite ERP	8.0	7.85	8.0	8.1	6.85	6.65	5.5	4.9
JD Edwards Enterprise One	8.45	8.35	8.6	8.35	8.6	8.2	8.9	9.3
Epicor ERP	8.45	8.15	8.5	7.6	8.4	7.8	8.2	8.0
IQMS	8.3	8.1	8.3	7.75	8.15	7.85	8.0	8.4
IFS ERP	8.0	8.6	8.45	8.3	8.25	8.4	7.8	7.9

Table 9 Determining weights for the criteria

Ease of use	*Features and functionalities*	*Advanced features*	*Integration*	*Performance*	*Customer support*	*Implementation*	*Renew and recommend*
7	9	7	9	9	7	8	7

6.4 Experimental results

By implementing the TOPSIS methodology and using Tables 6 and 7 are as inputs, following values are received; based on which the effectiveness of the model is decided.

Input:

Features values received for various ERP types – present in Table 6

Weights assigned for every feature – present in Table 7

Technique:

TOPSIS

Tool:

Python 3

Output:

Cloud ERP – CC1 = 0.4862315267760994

On premise – CC2 = 0.47502734844444733

Hybrid ERP – CC3 = 0.5089748608173186

The ranking order is now determined based on the closeness coefficient and its found hybrid ERP > cloud ERP > on premise ERP. Hence the best alternative is the hybrid ERP for an SME.

Once the type of ERP is determined, it becomes imperative for an SME in determining the top ERP that will be suitable for a particular organisation.

Table 8 is the source data that has been collected from Predictive Analytics Today which is a B2B discovery platform and Software Advice which is part of Gartner. This platform guides different organisations in proper decision making when choosing the appropriate technology for an organisation.

The different features are having various weights of importance. As per the literature review on research journals discussed above, it is observed that the highest weightage is given to features and functionalities, integration, performance followed by implementation and then customer support, advanced features, ease of use, and renew and recommendations while determining the top ERP vendor. Table 9 represents the same.

Applying TOPSIS on the top 7 ERP providers, it was found that JD Edwards Enterprise One which is ERP in a hybrid cloud landscape is the best among other top ERPs considered in our research followed by IQMS, Dynamics 365, IFS ERP, Epicor ERP, SAP Business ByDesign and Net Suite ERP.

Input:

Features values received for various organisations – present in Table 8

Weights assigned for every feature – present in Table 9

Technique:

TOPSIS

Tool:

Python 3

Output:

SAP Business ByDesign – 0.5389601817619295

Dynamics 365 – 0.5922383754467504

Net Suite ERP – 0.4594907143089015

JD Edwards Enterprise One – 0.6043250475775317

Epicor ERP – 0.5747204928568965

IQMS – 0.5929717994196178

IFS ERP – 0.5834725854443694

7 Conclusions and future scope

The economy has been facing extraordinary challenges in the COVID era. Due to the downturn in the economy, SMEs are facing huge challenges to survival. Since SMEs are the backbone of any nation's economy hence is imperative that SMEs will be playing a vital role when the situation improves. With the aid of ERP, SMEs can improve drastically in IT infrastructure and it can benefit in signification contribution to future progress. On-premise ERP and cloud-based ERP hold certain benefits and challenges and hybrid ERP can provide advantages in certain aspects. But every organisation faces a huge challenge in deciding the type of ERP and the ERP vendor that will be suitable for an organisation.

The present research work contributes in diverse ways. The machine learning algorithms are applied to predict the adoption intention of Indian SMEs related to different ERPs. From the diverse machine learning model analysis, the SMEs can take a decision on cloud ERP adoption. The graphical representation also provides a clear view which will aid in the appropriate prediction of the ERP. Decision making can also be done via MCDM. The benefits and challenges of different types of ERP are synthesised and based on critical factors multi-criteria decision analysis has been implemented to determine the suitable type of ERP for an SME. Now after determining the type of ERP, the next challenge faced by an SME is to find out a suitable ERP service provider among the top providers. Here as well the TOPSIS has been applied in developing a framework that will guide an SME in decision making and selecting the suitable ERP service provider. Hybrid ERP is developed by considering the best of on-premise and cloud ERP and in the current study, it is evident that the hybrid ERP is the most suitable for an organisation. The research work is novel considering that hybrid ERP which is any emerging trend but is in nascent stage is considered in the paper. The research work will act as a harbinger in future research work in hybrid ERP area.

References

Abdar, M., Acharya, U.R., Sarrafzadegan, N. and Makarenkov, V. (2019) 'NE-nu-SVC: a new nested ensemble clinical decision support system for effective diagnosis of coronary artery disease', *IEEE Access*, Vol. 7, pp.167605–167620.

Abdelaziz, A., Elhoseny, M., Salama, A.S. and Riad, A.M. (2018) 'A machine learning model for improving healthcare services on cloud computing environment', *Measurement*, Vol. 119, pp.117–128.

Akella, A. and Akella, S. (2021) 'Machine learning algorithms for predicting coronary artery disease: efforts toward an open source solution', *Future Science OA*, Vol. 7, No. 6, p.FSO698.

Akhondzadeh-Noughabi, E., Amin-Naseri, M.R., Albadvi, A. and Saeedi, M. (2016) 'Human resource performance evaluation from CRM perspective: a two-step association rule analysis', *International Journal of Business Performance Management*, Vol. 17, No. 1, pp.89–102.

Al-Johani, A.A. and Youssef, A.E. (2013) 'A framework for ERP systems in SME based on cloud computing technology', *International Journal on Cloud Computing: Services and Architecture*, Vol. 3, No. 3, pp.1–14.

Ayatollahi, H., Gholamhosseini, L. and Salehi, M. (2019) 'Predicting coronary artery disease: a comparison between two data mining algorithms', *BMC Public Health*, Vol. 19, No. 1, pp.1–9.

Banerjee, D. and Roy, B. (2017) 'Application of data mining with R programming to implement ERP for digitization of valuation reports on commercial vehicles and passenger cars', *International Journal of Management*, Vol. 8, No. 6.

Bjelland, E. and Haddara, M. (2018) 'Evolution of ERP systems in the cloud: a study on system updates', *Systems*, Vol. 6, No. 2, p.22.

Bos, J.W., Lauter, K. and Naehrig, M. (2014) 'Private predictive analysis on encrypted medical data', *Journal of Biomedical Informatics*, Vol. 50, pp.234–243.

Caesarendra, W., Widodo, A. and Yang, B.S. (2010) 'Application of relevance vector machine and logistic regression for machine degradation assessment', *Mechanical Systems and Signal Processing*, Vol. 24, No. 4, pp.1161–1171.

Cüvitoğlu, A. and Işik, Z. (2018) 'Classification of CAD dataset by using principal component analysis and machine learning approaches', in *2018 5th International Conference on Electrical and Electronic Engineering (ICEEE)*, IEEE, May, pp.340–343.

Duan, J., Faker, P., Fesak, A. and Stuart, T. (2013) 'Benefits and drawbacks of cloud-based versus traditional ERP systems', *Proceedings of the 2012-13 course on Advanced Resource Planning*.

Dwivedi, A.K. (2018) 'Analysis of computational intelligence techniques for diabetes mellitus prediction', *Neural Computing and Applications*, Vol. 30, No. 12, pp.3837–3845.

Elmonem, M.A.A., Nasr, E.S. and Geith, M.H. (2016) 'Benefits and challenges of cloud ERP systems – a systematic literature review', *Future Computing and Informatics Journal*, Vol. 1, Nos. 1–2, pp.1–9.

Estefania, T.V., Samir, L., Robert, P., Patrice, D. and Alexandre, M. (2018) 'The integration of ERP and inter-intra organizational information systems: a literature review', *IFAC-PapersOnLine*, Vol. 51, No. 11, pp.1212–1217.

Farid, D.M., Zhang, L., Rahman, C.M., Hossain, M.A. and Strachan, R. (2014) 'Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks', *Expert Systems with Applications*, Vol. 41, No. 4, pp.1937–1946.

Gartner, I. (2021) *Cloud ERP for Service-Centric Enterprises Reviews 2021*, Gartner Peer Insights [online] https://www.gartner.com/reviews/market/cloud-erp-for-service-centric-enterprises (accessed 1 October 2021).

Gladence, L.M., Reddy, K.R., Reddy, M.P. and Selvan, M.P. (2021) 'A prediction of crop yield using machine learning algorithm', in *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)*, IEEE, June, pp.1072–1077.

Gupta, S., Kumar, S., Singh, S.K., Foropon, C. and Chandra, C. (2018a) 'Role of cloud ERP on the performance of an organization', *The International Journal of Logistics Management*.

Gupta, S., Misra, S.C., Kock, N. and Roubaud, D. (2018b) 'Organizational, technological and extrinsic factors in the implementation of cloud ERP in SMEs', *Journal of Organizational Change Management*.

Hadi, W.E., Al-Radaideh, Q.A. and Alhawari, S. (2018) 'Integrating associative rule-based classification with Naïve Bayes for text classification', *Applied Soft Computing*, Vol. 69, pp.344–356.

Han, K., Hong, S., Cheon, J.H. and Park, D. (2019) 'Logistic regression on homomorphic encrypted data at scale', in *Proceedings of the AAAI Conference on Artificial Intelligence*, July, Vol. 33, No. 1, pp.9466–9471.

Hasan, N., Miah, S.J., Bao, Y. and Hoque, M.R. (2019) 'Factors affecting post-implementation success of enterprise resource planning systems: a perspective of business process performance', *Enterprise Information Systems*, Vol. 13, No. 9, pp.1217–1244.

Hedau, V., Malviya, A. and Chakraborty, N. (2013) 'Cloud based ERP for small and medium scale enterprises', in *International Journal of Engineering Research and Technology*, ESRSA Publications.

Hietala, H. and Päivärinta, T. (2021) 'Benefits realisation in post-implementation development of ERP systems: a case study', *Procedia Computer Science*, Vol. 181, pp.419–426.

Hosmer, D.W., Jovanovic, B. and Lemeshow, S. (1989) 'Best subsets logistic regression', *Biometrics*, pp.1265–1270.

Huang, Z., Zhou, Z. and He, T. (2014) 'Resolving rule conflicts based on Naïve Bayesian model for associative classification'.

Johansson, B., Alajbegovic, A., Alexopoulos, V. and Desalermos, A. (2014) 'Cloud ERP adoption opportunities and concerns: a comparison between SMES and large companies', in *Pre-ECIS 2014 Workshop 'IT Operations Management' (ITOM2014)*, June, pp.1–13.

Kiran, T. and Reddy, A. (2019) 'Critical success factors of ERP implementation in SMEs', *Journal of Project Management*, Vol. 4, No. 4, pp.267–280.

Knorr, E.M., Ng, R.T. and Tucakov, V. (2000) 'Distance-based outliers: algorithms and applications', *The VLDB Journal*, Vol. 8, No. 3, pp.237–253.

Kovacova, M., Kliestik, T., Kubala, P., Valaskova, K., Radišić, M. and Borocki, J. (2018) 'Bankruptcy models: Verifying their validity as a predictor of corporate failure', *Polish Journal of Management Studies*, Vol. 18.

Kutrani, H. and Eltalhi, S. (2019) 'Cardiac catheterization procedure prediction using machine learning and data mining techniques', *IOSR Journal of Computer Science*, Vol. 21, No. 1, pp.86–92.

Lee, N.C. and Chang, J. (2020) 'Adapting ERP systems in the post-implementation stage: dynamic IT capabilities for ERP', *Pacific Asia Journal of the Association for Information Systems*, Vol. 12, No. 1, p.2.

Mahara, T.N. (2013) 'Indian SMEs perspective for election of ERP in cloud', *Journal of International Technology and Information Management*, Vol. 22, No. 1, p.5.

Mair, A. and El-Kadi, A.I. (2013) 'Logistic regression modeling to assess groundwater vulnerability to contamination in Hawaii, USA', *Journal of Contaminant Hydrology*, Vol. 153, pp.1–23.

Md, A.Q. and Vijayakumar, V. (2019) 'Dynamic ranking of cloud services for web-based cloud communities: efficient algorithm for rating-based discovery and multi-level ranking of cloud services', *International Journal of Web Based Communities*, Vol. 15, No. 3, pp.248–270.

Mehta, S., Oberoi, A. and Bhatia, S.S. (2018) *A Cloud based Assurance Framework for Implementation of ERP in SMEs: A Literature Survey*.

Menon, S. (2019) 'Critical challenges in enterprise resource planning (ERP) implementation', *International Journal of Business and Management*, Vol. 14, No. 7.

Mousavi, S.M., Horton, S.P., Langston, C.A. and Samei, B. (2016) 'Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression', *Geophysical Journal International*, Vol. 207, No. 1, pp.29–46.

Narang, R.K. (2013) *Inside the Black Box: A Simple Guide to Quantitative and High Frequency Trading*, Vol. 846, John Wiley & Sons.

Naushad, S.M., Hussain, T., Indumathi, B., Samreen, K., Alrokayan, S.A. and Kutala, V.K. (2018) 'Machine learning algorithm-based risk prediction model of coronary artery disease', *Molecular Biology Reports*, Vol. 45, No. 5, pp.901–910.

Nayar, K.B. and Kumar, V. (2018) 'Cost benefit analysis of cloud computing in education', *International Journal of Business Information Systems*, Vol. 27, No. 2, pp.205–221.

Palvanov, A. and Im Cho, Y. (2018) 'Comparisons of deep learning algorithms for MNIST in real-time environment', *International Journal of Fuzzy Logic and Intelligent Systems*, Vol. 18, No. 2, pp.126–134.

Pandya, B., Patterson, L. and Ruhi, U. (2022) 'The readiness of workforce for the world of work in 2030: perceptions of university students', *International Journal of Business Performance Management*, Vol. 23, Nos. 1–2, pp.54–75.

Park, D.C. (2016) 'Image classification using naïve Bayes classifier', *Int. J. Comput. Sci. Electron. Eng.*, Vol. 4, No. 3, pp.135–139.

Ruivo, P., Rodrigues, J. and Oliveira, T. (2015) 'The ERP surge of hybrid models-an exploratory research into five and ten years forecast', *Procedia Computer Science*, Vol. 64, pp.594–600.

Saini, I., Khanna, A. and Kumar, V. (2012) 'ERP systems: problems and solution with special reference to small & medium enterprises', *International Journal of Research in IT & Management*, Vol. 2, No. 2, pp.715–725.

Saini, I., Khanna, A. and Peddoju, S.K. (2014) 'Cloud and traditional ERP systems in small and medium enterprises', in *2014 International Conference on Information Systems and Computer Networks (ISCON)*, IEEE, March, pp.138–141.

Salum, K.H. and Rozan, M.Z.A. (2016) 'Exploring the challenge impacted SMEs to adopt cloud ERP', *Indian Journal of Science and Technology*, Vol. 9, No. 45, pp.1–8.

Salum, K.H. and Rozan, M.Z.A.R. (2015) 'Barriers and drivers in cloud ERP adoption among SMEs', *Journal of Information Systems Research and Innovation*, Vol. 9, No. 1, pp.9–20.

Sandhya, Garg, R. and Kumar, R. (2018) 'Computational MADM evaluation and ranking of cloud service providers using distance-based approach'. *International Journal of Information and Decision Sciences*, Vol. 10, No. 3, pp.222–234.

Seethamraju, R. (2015) 'Adoption of software as a service (SaaS) enterprise resource planning (ERP) systems in small and medium sized enterprises (SMEs)', *Information Systems Frontiers*, Vol. 17, No. 3, pp.475–492.

Sharma, M., Mehra, A., Jola, H., Kumar, A., Misra, M. and Tiwari, V. (2010) 'Scope of cloud computing for SMEs in India'.

Slavkovic, A.B., Nardi, Y. and Tibbits, M.M. (2007) ''Secure' logistic regression of horizontally and vertically partitioned distributed databases', in *Seventh IEEE International Conference on Data Mining Workshops (ICDMW 2007)*, IEEE, October, pp.723–728.

Small, R. (2016) 'Factors affecting the adoption of enterprise resource planning (ERP) on cloud among small and medium enterprises (SMES) in Penang, Malaysia', *Journal of Theoretical and Applied Information Technology*, Vol. 88, No. 3.

Taghipour, M., Shabrang, M., Habibi, M.H. and Shamami, N. (2020) 'Assessment and analysis of risk associated with the implementation of enterprise resource planning (ERP) project using FMEA technique (including case-study)', *Management*, Vol. 3, No. 1, pp.29–46.

Talbi, J. and Haqiq, A. (2020) 'A cloud broker architecture for cloud service selection based on multi-criteria decision making and rough set theory', *International Journal of Computer Aided Engineering and Technology*, Vol. 13, No. 4, pp.448–464.

Tehrani, S.R. and Shirazi, F. (2014) 'Factors influencing the adoption of cloud computing by small and medium size enterprises (SMEs)', *International Conference on Human Interface and the Management of Information*, June, Springer, Cham, pp.631–642.

Thottakkara, P., Ozrazgat-Baslanti, T., Hupf, B.B., Rashidi, P., Pardalos, P., Momcilovic, P. and Bihorac, A. (2016) 'Application of machine learning techniques to high-dimensional clinical data to forecast postoperative complications', *PloS One*, Vol. 11, No. 5, p.e0155705.

Trueman, T.E. and Narayanasamy, P. (2018) 'A three-pronged technique for the intelligent ranking of cloud service providers', *International Journal of Business Information Systems*, Vol. 29, No. 3, pp.387–402.