

An integrated deep neural learning architecture with embedded feature reduction for customer lifetime value prediction in fintech firms

G. Jyothi¹, A. Gopinath², Meenu Baliyan³, Rachel Nallathamby⁴, K. Swanthana⁵, Murugesan Sundaram⁶

¹Department of MBA, International Institute of Business Studies, Bengaluru, India

²Department of Electronics with Artificial Intelligence, Dharmamurthi Rao Bahadur Calavala Cunnan Chetty's Hindu College, Chennai, India

³School of Management and Liberal Arts, SDGI Global University, Ghaziabad, India

⁴Department of CSE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

⁵Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India

⁶Department of Artificial Intelligence and Data Science, J.J. College of Engineering and Technology, Trichy, India

Article Info

Article history:

Received Jan 6, 2026

Revised Aug 12, 2026

Accepted Aug 28, 2026

Keywords:

Customer lifetime value

Deep learning

Embedded feature selection

Fintech

Temporal convolutional network

Variational graph autoencoder

ABSTRACT

The financial technology (FinTech) sector has witnessed remarkable growth in recent years, significantly transforming the way individuals and businesses enable and handle financial services. Recently, deep learning (DL) models have been beneficial in attaining appropriate data concerning customer lifetime value (CLV) for the banking experts. This paper presents a deep neural network (DNN) model with embedded feature selection (FS) for CLVP (DNNEFS-CLVP) approach. An embedded FS mechanism is primarily introduced to automatically identify the most informative attributes while eliminating redundant and less relevant data features. To further enhance predictive capability, a deep hybrid architecture is developed by integrating the temporal convolutional network, a variational graph autoencoder, and a graph convolutional network. This combined framework is expected to increase the accuracy and robustness of CLV prediction while assuring efficient feature utilization. The experimental result values of the DNNEFS-CLVP method accomplishes higher proportion of very low error values across all evaluation metrics, as evidenced by mean squared error (MSE) of 0.000081, mean absolute error (MAE) of 0.007126, root mean squared error (RMSE) of 0.009011, and mean absolute percentage error (MAPE) of 0.114715, indicating strong predictive performance and minimal deviation from actual CLV values.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

G. Jyothi

Department of MBA, International Institute of Business Studies

Bengaluru, Karnataka, India

Email: jyothiravindra20@gmail.com

1. INTRODUCTION

Over the past few decades, the fintech sector's quick development is driven by technical advances and developing consumer expectations. Fintech businesses provide novel financial services, namely payment solutions, investment platforms, and digital banking, empowering the modern tech enthusiast [1]. Customer lifetime value (CLV) is a critical idea in finance; it presents a business with insights that are priceless, which in turn presents the data regarding a customer's long-term worth to the association [2]. CLV is the estimated advantage that may be derived by an association from a customer for a specific period. CLV forecast assists in deciding which services to provide to various customer segments [3]. Therefore, predictive analytics,

which is a transformative method that combines information, statistical algorithms, and machine learning (ML) models for predicting future results, is increasingly utilized to evaluate CLV. Predictive analytics enables marketers to create informed approach decisions, optimize resource allocation, and enhance profitability and customer retention. The innovation in ML techniques has led to greater efficiency [4]. In addition, beneficial insights into the customer behavior, purchasing patterns, and preferences may be attained by precisely forecasting CLV employing ML models. Furthermore, the support vector machine (SVM), an ML method used by banking officials to detect irregularities in consumer transaction patterns, aids in achieving consumer trust with the economic health of retail banks. Thus, ML techniques are instrumental in forecasting business approximation keeping the buyer and in providing customized services to enhance business transactions. The efficient use of DL models for the prediction of CLV may assist banking in attaining maximal value from the results of this technique, enabling informed decisions on the basis of predictions [5].

A summary of prior research related to CLV prediction in FinTech is presented in this section. Weng *et al.* [6] presented an innovative optimum distribution selection technique (OptDist) for CLTV forecast that uses a dynamic optimum sub distribution selection module (DSM) for enhancing the precision of complex distribution modelling. Khan *et al.* [7] introduced a method that adopts two series phases for calculating CLV that depend on their purchase behaviours. For handling the imbalance and diversity in customer information, the authors explored three decision tree (DT) selection approaches for forming an effective ensemble. Todupunuri [8] examined the knowledge regarding ML technique inclusion for predicting CLV in getting services and analyzing customer motivation for the banking industry. Cowan *et al.* [9] proposed a usual structure to model CLV that is useful to product-centric customer relationships, as well as industries with long-lasting contractual relationships, of which retail banking is an instance. Curiskis *et al.* [10] evolved an ensemble CLV technique. Segarra-Moliner and Moliner-Tena [11] examined the literature gap of the relationship among the CLV and customer citizenship behavior (CCB) in the customer engagement structure (CE). The authors in [12] proposed a new method for CLV estimation using neural network model with Monte Carlo dropout (MCD) framework. The authors in [13] derives an extensive methodology for prediction and investigation of customer behavior by employing advanced predictive techniques as well as ML approaches.

Schmittlein *et al.* [14] developed a system based on time and number of previous proceedings of customers. This system permits the probability computation that any specific consumer is still active. Thakor [15] provided a fintech definition, which examines statistics and style facts, and then examines the theory and empirical literature. Fader *et al.* [16] presented a technique, the beta-geometric/NBD (BG/NBD), revealing less difference in the behavioural “story” combined along with the Pareto/NBD but it is more convenient for the implementation. Rezaei, [17] proposed a method that is a ML-enabled technique for vital-node detection. The main concept is training the method with a smaller graph portion, say 0.5% of the nodes, and predicting with the rest of the nodes. Maheswari [18] introduced a CVLV architecture that strategically enhances churn-prediction by adjusting deep learning (DL) methods with customers’ value risk profiles. Allen *et al.* [19] explored the personal and country features combined with financial inclusion and the policies that are likely to be eliminated: young, female, rural, or poor individuals.

While previous research has examined the influence of standard ML models, sequence-based DL techniques, and graph-based learning methodologies individually for CLV predictions, they have not specifically addressed the integrated influence on getting temporal-dynamics, relational-dependencies, and latent interactions with customers in an integrated hybrid architecture for accurate estimation of CLV. Weng *et al.* [6] applied an adaptive distribution selection technique to enhance CLV prediction precision, though Khan *et al.* [7] introduced a two-stage decision tree algorithm that depends on user purchase actions to calculate both purchasing value and likelihood. Likewise, Curiskis *et al.* [10] designed an ensemble-based approach that integrates several ML methodologies to improve prediction outcomes and manage partial data.

The fast development of the fintech domain has augmented the requirement for understanding consumer value for the best personalized services and decision-making process. Standard CLV systems are frequently slow, complex, and less effective in handling consumer activities. Therefore, there is a requirement for innovative DL and ML methods to enhance customer retention, financial planning, and profitability. To overcome this, this research proposes a deep neural network (DNN) model with embedded feature selection (FS) for CLV prediction (DNNEFS-CLVP) technique. The following are the significant contributions of this study.

- Utilize an embedded FS mechanism to automatically identify the most significant and non-redundant attributes, improving model efficiency and reducing data redundancy.
- Design a deep hybrid architecture integrating the temporal convolutional network (TCN), a variational graph autoencoder (VGAE), and a graph convolutional network (GCN) to effectively capture temporal dynamics and relational dependencies for accurate CLV prediction.

- The outputs from TCN, GCN, and VGAE are combined through feature concatenation, enabling the model to jointly capture temporal, structural, and latent representations.
- Employ the proposed framework for comprehensive experimentation using a benchmark FinTech Customer LTV dataset to ensure robust evaluation across multiple performance metrics.
- Validate the experimental outcomes through broad comparison assessment, representing that the proposed technique realizes greater performance over existing approaches.

2. METHOD

In this article, the DNNEFS-CLVP model is proposed for robust prediction of CLV in the FinTech sector. To achieve that, the DNNEFS-CLVP methodology begins with data pre-processing. Then, an embedded FS approach is applied to automatically select the particular attributes while minimizing redundancy, improving the overall efficiency of the model. Finally, a deep hybrid architecture combining TCN, VGAE, and GCN is employed for accurate and robust prediction of CLV. Figure 1 describes the workflow of the DNNEFS-CLVP methodology.

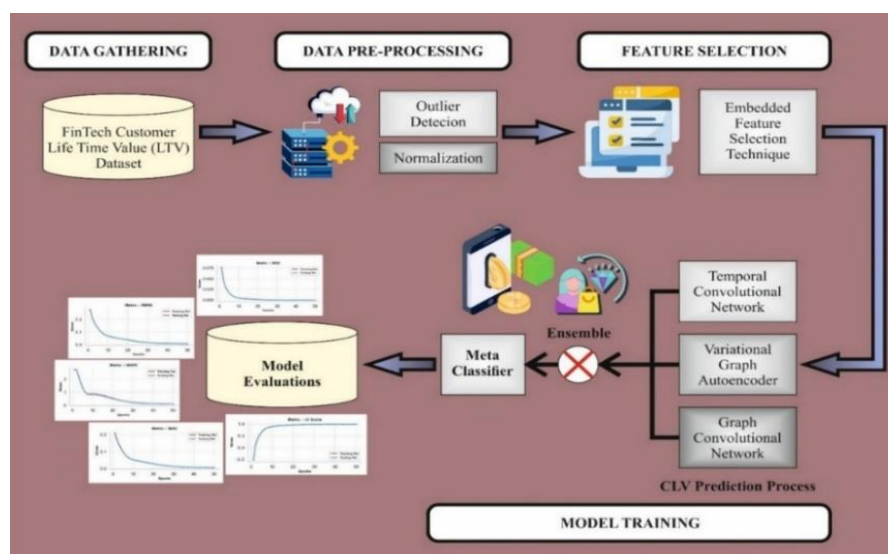


Figure 1. Workflow of DNNEFS-CLVP methodology

2.1. Input data preparation

An outlier detection and normalization are employed as input data pre-processing techniques to ensure data reliability and consistency. To identify outliers and improve the sturdiness of the model, the Z-score technique has been employed [20]. It helps to discover concealed patterns to pinpoint observations that significantly vary. This approach has a probable model for identifying outliers. This model was extensively leveraged for recognizing outliers. It aids in decreasing the consequence of huge values and bringing data more closely to a normal distribution that is helpful for regression approaches and statistical analysis.

2.2. Embedded FS methodology

An embedded FS method is utilized for identifying related features while decreasing data redundancy. Embedded FS models comprise the selection procedure in the model training itself, being a compromise between filter and wrapper models. In such methods, the learning algorithm automatically picks the most informative attributes [21] in contrast to Wrapper models, which consider dissimilar feature subsets throughout the selection procedure, embedded models like the best features in the learning procedure, which is commonly faster and less prone to overfitting.

The LASSO/Ridge for dimensionality reduction is added as a pre-processing stage for reducing redundancy of features, improving the data clarity, and enhancing the stability of the model before training the deep-hybrid framework. Irrelevant features are filtered and reduce the computing complexity, which allows the further VGAE, GCN, and TCN components to concentrate on more inputs.

LASSO: LASSO is an embedded method of linear regression that incorporates both FS and regularisation. It decreases the sum of squares focused on a restriction on the number of absolute values of

the model coefficients. This disadvantage is recognized as an L1 penalty, which means a sparse method, meaning that the terms of a few coefficients will end up being precisely 0, which in turn can aid in picking significant features and make classifiers more interpretable.

The LASSO loss of the function is described in mathematical formulation,

$$L1 = \sum_{j=1}^m (y_i - \sum_{i=1}^q x_{ji} \beta_i)^2 + \vartheta \sum_{i=1}^q |\beta_i| \dots \quad (1)$$

by comparison, ridge regression, another regularisation model, employs the L2 penalty, exhausting the extent of coefficients in (2).

$$L2 = \sum_{j=1}^m (y_i - \sum_{i=1}^q x_{ji} \beta_i)^2 + \vartheta \sum_{i=1}^q |\beta_i|^2 \quad (2)$$

Here, m denotes the no. of training instances, q is the number of features, y_i represents a true target value for the j^{th} sample, x_{ji} refers to the value of feature i for the j^{th} training sample; β_i refers to the coefficient of the i^{th} feature, and ϑ is the regularisation parameter.

Ridge: ridge is a regularisation model. The objective procedure provides a squared magnitude of the standard parameter as a penalty term. Ridge penalizes coefficient estimates near zero but not surely zero, distinct from the LASSO model that presents a sparse exhibition of the function. The significant discrepancy between Ridge and LASSO plans is that LASSO reduces the coefficient of irrelevant variables to 0, efficiently recompensing them.

2.3. Ensemble DL approaches

A deep hybrid architecture integrating a TCN, VGAE, and GCN is applied for the accurate prediction of CLV. Every individual node in the graph refers to an individual customer from the FinTech Customer LTV dataset. In the derived graph, the nodes signify to particular customer records. The edges between the nodes are derived by the use of feature similarity relationships derived by the use of customer behavioral and financial attributes. A similarity-based adjacency matrix is generated and normalized before being supplied to the VGAE-GCN framework for representation learning.

2.3.1. TCN

DNNs are especially designed for processing temporal details [22]. The design depends on convolutions rather than repetitive units, as in the RNN framework. This technique has been beneficial for similar calculations and attains greater training consistency. The TCN method utilizes convolutional layers for processing temporal details. Initially, it uses a 1D CNN that substitutes a repetitive framework that aids in enhancing training efficacy. Next, causal convolutions certify that the present input does not affect future values throughout training.

2.3.2. VGAE

Stimulated by the latent representation of a node, the VGAE is constructed to build complete usage of topological data and effective attributes for network representation learning. In VGAE, the encoding consists of three layers to incorporate data of anchor nodes and their neighborhood [23]. Then, latent representation Y' is molded through presenting a standard normal distribution. Its objective function was expressed.

$$q(Y'|Y, Q) = \prod_{i=1}^n q(y_i|Y, Q) = \prod_{i=1}^n N(y_i | \mu_i, \text{diag}(\sigma_i^2)) \quad (3)$$

Whereas μ denotes the mean, the symbol σ represents variance vectors, respectively; $N(\cdot)$ means a standard Gaussian distribution. Here, Y aid as initial inputs, and data were assembled into graph Q as an adjacency matrix. The GCN in the encoding is formulated,

$$f_{gcn}(Y^{(l)}, G, W^{(l)}) = \text{ReLU}(D^{-\frac{1}{2}} Q D^{-\frac{1}{2}} Y^{(l)} + \sum_{(m,n) \in \epsilon} D_{mn}^{-\frac{1}{2}} D_{nm}^{-\frac{1}{2}} f(e_{mn})) \quad (4)$$

Here, D_{mn} and D denote a degree matrix linked with edges and nodes, respectively; e_{mn} represents an extent connection, and $f(\cdot)$ indicates a linear transformation of edges. By presenting the edge features that rank higher, the reliability interaction network in structures enlarges prediction accuracy.

2.3.3. GCN

Classical CNN acquires local spatial data but is limited to Euclidean space and lacks the capability of managing topologies. GCN has an expansion of CNN and usually the spatial feature remover manages a topological graph with correlation determined by edges and vertices [24]. The norm of GCN must be followed: in the graph with N nodes and every node with S attributes, an $N \times S$ dimension feature matrix X and an $N \times N$ adjacency matrix A have been formed, the basics of adjacency among each pair of nodes; subsequently, X and A are utilized as input.

$$H^{l+1} = \sigma D^{-\frac{1}{2}} \hat{A} D^{-\frac{1}{2}} H^{(l)} \theta^{(l)} \quad (5)$$

Where $\hat{A} = A + I_N$ signifies a matrix with auto-correlated edges added to every node within the graph, I_N represents the identity matrix, $\sigma(\cdot)$ indicates the sigmoid function of a non-linear model, and $\tilde{D} = \sum_j \tilde{A}_{ij}$ refers to the degree matrix that signifies the outcome of 1, $\theta^{(1)}$ comprises the parameters. Table 1 provides the hyperparameter settings of the presented DNNEFS-CLVP method.

Table 1. Hyperparameter settings of the presented DNNEFS-CLVP approach

Module	Hyperparameter	Parameterized value
Embedded FS (LASSO + Ridge)	LASSO regularization (α)	0.01
	Ridge regularization (λ)	0.01
	Feature importance threshold	0.2
Temporal convolutional network (TCN)	Learning rate	0.001
	Batch size	32
	Epochs	25
	Number of TCN layers	3
	Number of filters	64
	Kernel size	3
	Dilation rate	2
	Dropout rate	0.2
Variational graph autoencoder (VGAE)	Latent dimension	64
	Encoder layers	3
	Hidden units	128
	Reconstruction weight	0.5
Graph convolutional network (GCN)	Number of GCN layers	3
	Hidden dimension	128
	Activation function	ReLU
Training configuration	Initial learning rate	0.001
	Weight decay	1e-4
	Early stopping patience	10

3. RESULTS AND DISCUSSION

In the following section, the study of the DNNEFS-CLVP method is inspected under the FinTech customer life time value (LTV) dataset. The experimental study of the DNNEFS-CLVP method is inspected under the FinTech customer LTV Dataset [25]. The dataset assists in predicting the LTV for users of online payments, particularly targeting platforms such as Khalti and PayTM. The dataset consists of about 7,000 examples with 20 rich features, but only 16 are chosen.

Table 2 represents the comparative study of the DNNEFS-CLVP method with current models under several measures like MSE, RMSE, MAE, and MAPE [12], [13]. Under MSE, the DNNEFS-CLVP method has achieved a minimal MSE of 0.000081. In contrast, the existing approaches, like MLP, DCNv2, Ziln, Stacking, DNN, Catboosting, and XGboost, have reached a superior MSE of 0.004871, 0.004181, 0.003411, 0.002661, 0.002071, 0.001471, and 0.000831, respectively.

Table 2. Comparative analysis of DNNEFS-CLVP model with existing approaches

Method	MSE	RMSE	MAE	MAPE
MLP	0.004871	0.069793	0.029100	0.483500
DCNv2	0.004181	0.064661	0.049530	0.422600
Ziln	0.003411	0.058404	0.038180	0.750000
Stacking	0.002661	0.051585	0.009643	0.442670
DNN	0.002071	0.045508	0.023530	0.281940
Catboosting	0.001471	0.038354	0.012234	0.427620
XGboost	0.000831	0.028827	0.017432	0.372090
DNNEFS-CLVP	0.000081	0.009011	0.007126	0.114715

Table 3 indicates the Ablation study of proposed DNNEFS-CLVP model. The TCN only, achieved RMSE of 0.056011, MSE of 0.004701, MAE of 0.059226 and MAPE of 0.573715 as well as VGAE only providing a MSE of 0.003941, RMSE of 0.050611, MAE of 0.053526 and MAPE of 0.498715. Likewise, the TCN + VGAE, attained a MSE of 0.002541, RMSE of 0.038211, MAE of 0.039026 and MAPE of 0.361715. Also, the TCN + GCN achieves a MSE of 0.001851, RMSE of 0.030911, MAE of 0.030826 and MAPE of 0.310715. Meanwhile, the TCN + VGAE + GCN (without EFS) better performance of, MSE of 0.000641, RMSE of 0.017711, MAE of 0.013726 and MAPE of 0.177715. At the same time, the DNNEFS-CLVP (EFS + TCN + VGAE + GCN) model, better outputs with MSE of 0.000081, RMSE of 0.009011, MAE of 0.007126 and MAPE of 0.114715, emphasising the efficiency of the components.

Table 3. Ablation study of DNNEFS-CLVP model

Method	MSE	RMSE	MAE	MAPE
TCN only	0.004701	0.056011	0.059226	0.573715
VGAE only	0.003941	0.050611	0.053526	0.498715
GCN only	0.003271	0.044611	0.044826	0.419715
TCN + VGAE	0.002541	0.038211	0.039026	0.361715
TCN + GCN	0.001851	0.030911	0.030826	0.310715
VGAE + GCN	0.001261	0.023711	0.021926	0.246715
TCN + VGAE + GCN (without EFS)	0.000641	0.017711	0.013726	0.177715
DNNEFS-CLVP (EFS + TCN + VGAE + GCN)	0.000081	0.009011	0.007126	0.114715

Table 4 represents the results of the 10-fold cross-validation executed to evaluate the model effectiveness. The achieved MSE values range from 0.000089 to 0.000152, when the RMSE values vary among 0.007221 and 0.007812 across different folds. Similarly, the MAE values are observed between 0.111423 and 0.117302, and the MAPE values range from 0.02200 to 0.08000. The outcomes show that the model attains stable and consistent performance across all folds with very small variations in error values.

Table 4. K-fold cross validation of DNNEFS-CLVP model

K-fold cross validation	MSE	RMSE	MAE	MAPE
Fold 1	0.000105	0.007572	0.113151	0.02200
Fold 2	0.000090	0.007365	0.116273	0.07900
Fold 3	0.000089	0.007664	0.114424	0.08000
Fold 4	0.000108	0.007812	0.117302	0.07200
Fold 5	0.000150	0.007221	0.112280	0.06200
Fold 6	0.000100	0.007358	0.115536	0.04900
Fold 7	0.000152	0.007224	0.114522	0.05800
Fold 8	0.000108	0.007466	0.116835	0.05800
Fold 9	0.000135	0.007694	0.111423	0.02700
Fold 10	0.000104	0.007706	0.115395	0.05900

The proposed method presents an effectual DL model for CLV forecast in the FinTech field. It integrates an embedded FS system to detect the more appropriate features though eliminating repeated attributes. A deep hybrid framework integrating TCN, VGAE, and GCN is applied to extracts both temporal patterns and relational dependencies in user information. This combination improves the methods predictive accurateness, robustness, and feature utilization efficacy. Eventhough the dataset is comparatively small; regularization methods and validation techniques help to reduce the overfitting and improves reliability of the model. Cross validation outcomes also show that this model is consistent over various data splits.

4. CONCLUSION

This article has presented a DNNEFS-CLVP technique to foresee CLV in the FinTech field through advanced DL techniques. The DNNEFS-CLVP initially undergoes data pre-processing steps using outlier detection and normalisation to ensure data reliability and consistency. Next, an embedded FS process has been implemented to recognize the significant features while reducing data redundancy automatically. In the last stage, a deep hybrid architecture combining a TCN, VGAE, and GCN is implemented to perform precise and efficient CLV prediction. Our study suggests that lower MSE values are strongly associated with improved prediction performance in CLV estimation, indicating better model accuracy. The proposed DNNEFS-CLVP method achieves a significantly lower error MSE of 0.000081 compared to baseline models such as MLP, DCNv2, ZILN, Stacking, DNN, CatBoost, and XGBoost, MSE of 0.004871, 0.004181, 0.003411, 0.002661, 0.002071, 0.001471, and 0.001471.

Even though the proposed method demonstrates robust results, it may affect from higher computation complexity owing to its hybrid DL framework and may involve larger datasets for effectual training. It may have inadequate interpretability due to the incorporation of distinct deep strategies. Future work can emphasis on emerging most lightweight and effective framework to minimize computation cost and enhance scalability. Moreover, combining explainable AI algorithm and computing the system on distinct real-time datasets can improve transparency and generalization solutions.

FUNDING STATEMENT

The authors received no funding for this study.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
G. Jyothi	✓	✓	✓											
A. Gopinath			✓	✓	✓									
Meenu Baliyan					✓	✓						✓		
Rachel Nallathamby								✓	✓			✓		
K. Swanthana										✓	✓	✓		
Murugesan Sundaram													✓	✓

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

DATA AVAILABILITY

The data that support the findings of this study are openly available in the Kaggle repository at <https://www.kaggle.com/datasets/harunrai/fintech-customer-life-time-value-ltv-dataset>, reference number [24].

REFERENCES

[1] M. H. Rahman *et al.*, “Transforming customer retention in fintech industry through predictive analytics and machine learning,” *The American Journal of Engineering and Technology*, vol. 06, no. 10, pp. 150–163, Oct. 2024, doi: 10.37547/tajet/Volume06Issue10-17.

[2] F. Xiao, S. X. Chen, Z. yu Chen, Y. N. Wang, and J. Q. Wang, “An interpretable data-driven approach for customer purchase prediction using cost-sensitive learning,” *Engineering Applications of Artificial Intelligence*, vol. 138, 2024, doi: 10.1016/j.engappai.2024.109344.

[3] J. Bauer and D. Jannach, “Improved customer lifetime value prediction with sequence-to-sequence learning and feature-based models,” *ACM Transactions on Knowledge Discovery from Data*, vol. 15, no. 5, 2021, doi: 10.1145/3441444.

[4] G. Zhou *et al.*, “Deep interest evolution network for click-through rate prediction,” *33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*, pp. 5941–5948, 2019, doi: 10.1609/aaai.v33i01.33015941.

[5] R. Ranganath, A. Perotte, N. Elhadad, and D. Blei, “Deep survival analysis,” in *Proc. Mach. Learn. Healthcare Conf. (MLHC)*, 2016, pp. 101–114, [Online]. Available: <http://arxiv.org/abs/1608.02158>.

[6] Y. Weng *et al.*, “OptDist: learning optimal distribution for customer lifetime value prediction,” *International Conference on Information and Knowledge Management, Proceedings*, pp. 2523–2533, 2024, doi: 10.1145/3627673.3679712.

[7] Z. Khan, A. Ali, S. Aldahmani, H. Nordmark, and B. Lausen, “Customer lifetime value modelling via two stage selected trees ensembles,” *IEEE Access*, vol. 13, pp. 115236–115247, 2025, doi: 10.1109/ACCESS.2025.3584709.

[8] A. Todupunuri, “Develop machine learning models to predict customer lifetime value for banking customers, helping banks optimize services,” *International Journal of All Research Education & Scientific Methods*, vol. 12, no. 10, pp. 1254–1259, 2024, doi: 10.56025/ijaresm.2024.1210241254.

- [9] M. K. Baidya, B. Maity, and K. Ghose, "Innovation in marketing strategy: a customer lifetime value approach," *Journal of Business and Management*, vol. 25, no. 2, pp. 71–97, 2019, doi: 10.1504/jbm.2019.141272.
- [10] S. Curiskis, X. Dong, F. Jiang, and M. Scarr, "A novel approach to predicting customer lifetime value in B2B SaaS companies," *Journal of Marketing Analytics*, vol. 11, no. 4, pp. 587–601, 2023, doi: 10.1057/s41270-023-00234-6.
- [11] J. R. Segarra-Moliner and M. A. Moliner-Tena, "Engaging in customer citizenship behaviours to predict customer lifetime value," *Journal of Marketing Analytics*, vol. 12, no. 2, pp. 307–320, 2024, doi: 10.1057/s41270-022-00195-2.
- [12] X. Cao, Y. Xu, and X. Yang, "Customer lifetime value prediction with uncertainty estimation using Monte Carlo dropout," 2024, [Online]. Available: <http://arxiv.org/abs/2411.15944>.
- [13] A. M. Haddadi and H. Hamidi, "A hybrid model for improving customer lifetime value prediction using stacking ensemble learning algorithm," *Computers in Human Behavior Reports*, vol. 18, 2025, doi: 10.1016/j.chbr.2025.100616.
- [14] D. C. Schmittlein, D. G. Morrison, and R. Colombo, "Counting your customers: who are they and what will they do next?," *Management Science*, vol. 33, no. 1, pp. 1–24, 1987, doi: 10.1287/mnsc.33.1.1.
- [15] A. V. Thakor, "Fintech and banking: What do we know?," *Journal of Financial Intermediation*, vol. 41, 2020, doi: 10.1016/j.jfi.2019.100833.
- [16] P. S. Fader, B. G. S. Hardie, and K. L. Lee, "'Counting your customers' the easy way: An alternative to the Pareto/NBD model," *Marketing Science*, vol. 24, no. 2, pp. 275–284, 2005, doi: 10.1287/mksc.1040.0098.
- [17] A. A. Rezaei, J. Munoz, M. Jalili, and H. Khayyam, "A machine learning-based approach for vital node identification in complex networks," *Expert Systems with Applications*, vol. 214, 2023, doi: 10.1016/j.eswa.2022.119086.
- [18] U. M. Gurusamy, M. Anantharaman, R. P. Selvamani, and S. Vijayarajan, "A customer lifetime value-aware framework for strategic churn prediction using deep learning," *Advanced Engineering Research (Rostov-on-Don)*, vol. 26, no. 1, 2026, doi: 10.23947/2687-1653-2026-26-1-2211.
- [19] F. Allen, A. Demircuc-Kunt, L. Klapper, and M. S. M. Peria, "The foundations of financial inclusion: Understanding ownership and use of formal accounts," *Journal of Financial Intermediation*, vol. 27, pp. 1–30, 2016, doi: 10.1016/j.jfi.2015.12.003.
- [20] M. Alanazi and T. N. Alharby, "Computational evaluation using machine learning for analysis of membrane desalination process powered by solar energy," *Scientific Reports*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-20291-4.
- [21] A. K. Kareem, M. Al-Mahdawi, and A. A. Nafea, "A review of feature selection techniques for cancer diagnosis and prediction with machine learning and deep learning: current trends and future directions," *Al Rafidain Journal of Engineering Sciences*, vol. 3, no. 2, Sep. 2025, doi: 10.61268/q78nda33.
- [22] M. I. Habadi, S. AL-qahtani, H. I. Hariry, and M. Alshehri, "A comparative analysis of statistical and deep learning models for global temperature anomalies forecasting," *Journal of Reliability and Statistical Studies*, vol. 18, no. 2, pp. 371–398, 2025, doi: 10.13052/jrss0974-8024.1825.
- [23] S. Chen, A. Zheng, W. Yu, and C. Zhan, "DCMF-PPI: a protein-protein interaction predictor based on dynamic condition and multi-feature fusion," *BMC Bioinformatics*, vol. 26, no. 1, 2025, doi: 10.1186/s12859-025-06272-4.
- [24] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, 2017.
- [25] H. Rai, "FinTech customer life time value (LTV) dataset," *https://www.kaggle.com/Datasets/Harunrai/Fintech-Customer-Life-Time-Value-Ltv-Dataset*, vol. 2, no. 1, pp. 1–2, 2025, [Online]. Available: <https://www.kaggle.com/datasets/harunrai/fintech-customer-life-time-value-ltv-dataset>.

BIOGRAPHIES OF AUTHORS



G. Jyothi     is an associate professor in the Department of MBA at the International Institute of Business Studies, Bengaluru, Karnataka, India. Her academic interests include management studies, business administration, and interdisciplinary research. She can be contacted at email: jyothiravindra20@gmail.com.



A. Gopinath     is an assistant professor in the Department of Electronics with Artificial Intelligence at Dharmamurthi Rao Bahadur Calavala Cunnan Chetty's Hindu College, Chennai, Tamil Nadu, India. His research focuses on electronics, artificial intelligence, and emerging technologies. He can be contacted at email: gopinath@drbccchinducollege.edu.in.



Meenu Baliyan    is a professor in the School of Management and Liberal Arts, SDGI Global University, Ghaziabad, Uttar Pradesh, India. Her areas of expertise include management, organizational studies, and academic research. She can be contacted at email: meenu_baliyan@yahoo.co.in.



Rachel Nallathamby    is an assistant professor in the Department of Computer Science and Engineering at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India. Her interests include computer science, software technologies, and innovative research practices. She can be contacted at email: drachelnallathamby@veltech.edu.in.



K. Swanthana    is an assistant professor in the Department of Computer Science and Engineering at Koneru Lakshmaiah Education Foundation, Hyderabad, Telangana, India. Her research interests include computer science, data-driven technologies, and engineering applications. She can be contacted at email: kswanthana@gmail.com.



Murugesan Sundaram    is a professor, Department of Artificial Intelligence and Data Science, J.J. College of Engineering and Technology, Trichy, Tamil Nadu, India. His academic and research interests include artificial intelligence, data science, and advanced computing technologies. He can be contacted at email: bsmurugesan@gmail.com.