Educational Administration: Theory and Practice
2023, 29(03), 1407-1424
ISSN: 2148-2403
https://kuey.net/

Research Article

A Proficient Hospital Ratings Aware Patient Churn Prediction And Prevention System Using Abg-Fuzzy And Ner-Gfjdkmeans

Srinivasa Reddy Mukkala[1*], Vinodkumar Reddy Surasani[2], Hemasundara Reddy Lanka[3], Rajesh Kumar Kanji[4] and Vijaya Kumar Pothireddy[5]

[1*]Director of Statistical Programming and Biostatistics Episdata inc.
[2]Sr. Software Engineer RBC Wealth Management, Minneapolis.
[3]Data Architect, Publicis Sapient, Minneapolis.
[4]Lead Technical support engineer, Informatica Inc.
[5]Software Engineer, Google LLC.

***Corresponding author:** Srinivasa Reddy Mukkala
*Director of Statistical Programming and Biostatistics Episdata inc.

Citation: Srinivasa Reddy Mukkala et al. (2023), A Proficient Hospital Ratings Aware Patient Churn Prediction And Prevention System Using Abg-Fuzzy And Ner-Gfjdkmeans, *Educational Administration: Theory and Practice, 29(03), 1407-1424*
Doi: 10.53555/kuey.v29i3.9511

ARTICLE INFO

ABSTRACT

Patient churn in healthcare denotes the rate at which patients stop visiting or seeking care from a hospital. High churn represents dissatisfaction, better alternatives, or accessibility issues. However, the existing works didn't improve the healthcare churn prediction regarding the hospital ratings along with demographic information, medical history, and clinical factors. Therefore, this paper presents ABG-Fuzzy and NER-GFJDKMEANS-enabled hospital ratings-aware patient churn prediction and prevention system. Initially, the patient review dataset is taken and then pre-processed to improve the data quality. Afterward, from the pre-processed data, NLP features are extracted, and word embedding is done using BERT. Based on the extracted NLP features and word embedding outcomes, polarity identification is performed by utilizing L3STM. In the meantime, patient readmit and hospital ratings datasets are taken and then pre-processed. After that, data balancing by SBMOTE, violin plot generation, feature extraction, feature selection by QBSOA, and classification by L3STM are done. Thereafter, regarding the patient review outcomes, patient readmit, and hospital ratings, the patient chunk is predicted by using ABG-Fuzzy. Next, based on the patient review, the retain strategies are provided for the high and medium patient churn by employing NER-GFJDKMEANS. The results proved that the proposed model achieved a high accuracy of 98.3512%, which was superior to the prevailing methods.

Key words: Patient Churn Prediction, Bidirectional Encoder Representations from Transformers (BERT), Patient Readmit and Hospital Ratings, Natural Language Processing (NLP), Retain Strategy, Patient Review, Healthcare, and Missing Value Imputation (MVI).

1. INTRODUCTION

In modern healthcare systems, patient retention is an important aspect that directly affects hospital revenue, service quality, and overall patient satisfaction (Sengupta et al., 2024) (Huang et al., 2021). Patient retention refers to the ability of healthcare providers to maintain long-term relationships with patients by ensuring continuous engagement, high-quality care, and effective communication (Chauhan et al., 2020) (Albright et al., 2022). However, inadequate retention strategies can lead to patient churn, where individuals either switch providers or discontinue care, thus disrupting continuity and affecting health outcomes (Haimson et al., 2023) (Bustea et al., 2023). Moreover, churn can significantly impact the financial stability of healthcare institutions (Makene et al., 2022). Additionally, a high customer churn rate often indicates the patients' dissatisfaction with the quality of care, communication, or overall convenience of healthcare services, thus leading to client and revenue loss, reducing trust in healthcare providers, and potential reputational damages (Namakoola et al., 2024) (Teo et al., 2023). Hence, an efficient patient churn prediction system is important to identify clients at

risk of churn, thus enabling proactive measures (Arredondo et al., 2024). Various traditional Machine Learning (ML) and Deep Learning (DL)-based patient churn prediction frameworks have been implemented. Among them, ML techniques offer significant advantages like interpretability, robustness to missing data, and the ability to handle both categorical and numerical features (Ajegbile et al., 2024). However, they often struggled with capturing complex, non-linear relationships within patient data, thus limiting predictive accuracy (Bui & Moriuchi, 2021). Similarly, DL methodologies achieve superior performance in patient churn prediction by automatically extracting meaningful patterns from large-scale healthcare data (Flaks-Manov et al., 2020). Yet, these techniques require large amounts of labeled data and high computational power, thus making them less feasible for smaller healthcare institutions (Jaswal et al., 2024). Also, several conventional works predict patient churns inaccurately owing to the lack of understanding of the distribution of numerical features (Pinheiro & Cavique, 2022). Additionally, none of the traditional works concentrated on the hospital ratings data along with demographic information, medical history, and clinical factors for predicting patient churn, thus resulting in misprediction. Hence, a hospital ratings-aware patient churn prediction and prevention system is proposed using the ABG-Fuzzy and NER-GFJDKMeans technique.

1.1 Problem Statement

The limitations of several traditional methodologies are listed as follows,

- None of the traditional methodologies deemed the hospital ratings along with demographic information, medical history, and clinical factors for patient churn prediction, causing inaccurate prediction.
- Owing to the lack of analysis of patients' experiences, (Saha et al., 2024) exhibited misprediction.
- (De & Prabu, 2022) failed to inform targeted interventions to reduce the risk and improve patient engagement.
- Due to the utilization of an imbalanced dataset, (Afzal et al., 2024) attained high false positive predictions.
- Some traditional approaches failed to understand the distribution of numerical features and compare them across various categories in patient churn prediction.

1.2 Objective

- The major objectives of the proposed framework are demonstrated as follows,
- The proposed work accurately predicts the heath churn by considering the hospital ratings along with demographic information, medical history, and clinical factors.
- By identifying the polarity using the SSLAN, the proposed work predicts the health churn based on patient reviews.
- The NER-GFJDKMeans is employed to provide retention strategies.
- The SBMOTE is employed for data balancing, thus resulting in accurate predictions.
- To improve the overall performance in patient churn prediction, this framework extracts and compares the NLP and violin plot features.

The structure of the paper is outlined as follows: Section 2 demonstrates the literature survey, Section 3 depicts the proposed methodology, Section 4 illustrates the result and discussion, and Section 5 concludes the proposed work with future scope.

2. LITERATURE SURVEY

(Saha et al., 2024) presented a Deep Learning (DL) enabled churn prediction framework. Here, the SMOTE algorithm was employed to solve the class imbalance problem. In this research, the ChurnNet was used to predict the churn; here, a one-dimensional convolution layer was combined with residual block, squeeze and excitation block, and spatial attention module. The model achieved high performance and attained better outcomes regarding accuracy. However, the research failed to analyze the client experience in the corresponding environment for improving the churn prediction.

(De & Prabu, 2022) established a framework for churn prediction. In this work, a sampling-based stack model was employed for solving the imbalanced learning in churn prediction. Here, the prediction capabilities of sampling solutions were combined for stimulating the information gain of the meta-features in the ensemble. The model obtained improved performance regarding the Area Under the Curve (AUC). But, this work didn't provide targeted interventions for diminishing the risk of high churn.

(Afzal et al., 2024) propounded a framework for enhancing customer retention through ensemble models. In this research, a hybrid method, which combined K-Nearest Neighbors (KNN) and simple imputer, was employed to impute the missing values. For revealing the significant elements, Exploratory Data Analysis (EDA) was utilized. Lastly, Decision Tree, Random Forest, KNN, and eXtreme Gradient Boosting (XGBoost) classifiers were used to predict the churn. The research achieved high accuracy and provided interventions for improving customer satisfaction. Yet, the model had class imbalance issues, which increased the false positives in churn prediction.

(Ehsani & Hosseini, 2024) suggested churn prediction framework. For selecting the optimal features, this research employed permutation and SelectKBest feature selection methods. After that, the churn was predicted

by utilizing Decision Tree, Random Forest, XGBoost, Adaptive Boosting (AdaBoost), and Extra Trees classifiers. The model achieved a higher accuracy and aided as a valuable tool for detecting potential churns. However, the research provided poor performance for varying sectors with different customer behaviors and churn patterns.

(Ahmad et al., 2023) explored customer personality analysis-based churn prediction model. Here, the Conditional Tabular Generative Adversarial Networks (CTGAN) and SMOTE techniques were employed to solve the class imbalance problem. In this work, an ensemble approach was employed for churn prediction; here, the bagging approach used a random forest classifier, and the boosting approach employed XGBoost, Light GBM, and AdaBoost classifiers. The model obtained improved outcomes in terms of accuracy, precision, and recall. Yet, the model increased complexity, thus hindering interpretability.

(Joy et al., 2024) developed a model for enhancing customer retention through churn prediction. Here, optimal features were selected by using Chi-squared testing and Sequential Feature Selection (SFS). Afterward, the LSTM and Gated Recurrent Unit (GRU) were established to capture the trends in subscribers' usage patterns over time. Then, the churn forecasting was done by Light Gradient Boosting Machine (Light GBM). Lastly, the Shapley Additive Explanations (SHAP) was employed to provide an explanation about outcomes. The model was more useful for decision-makers. Nevertheless, the model didn't analyze the changes in client preferences and behaviors due to emerging trends, thus hindering the performance.

(Chang et al., 2024) established Machine Learning (ML) techniques-enabled churn behavior prediction model. In this work, ML approaches, such as Decision Trees, Boosted Trees, and Random Forests were employed for churn prediction. Afterward, SHAP and Local Interpretable Model-agnostic Explanations (LIME) techniques were used to provide explanations about churn prediction outcomes. The research excellently detected possible churns, thus allowing for timely interventions. Owing to the inadequate feature selection, the predictions were misclassified.

(Pejić Bach et al., 2021) introduced churn management framework. Here, the K-Means clustering and Chi-Square analysis were employed to identify the clusters with high churn ratios. Afterward, the churn was predicted by the CHi-squared Automatic Interaction Detector (CHAID) decision tree algorithm. Here, the predicted churn was fed back to the customer database for effectively increasing churn management. Thus, the model obtained improved performance. But, in this work, missing values, outliers, and inconsistent data could diminish the accuracy of the model.

(Thakkar et al., 2022) presented a churn prediction model. For effectively predicting the churn, the class-dependent cost-sensitive boosting algorithm known as AdaBoostWithCost was employed in this work. The research effectively reduced the misclassification cost and false negative error. However, the model had overfitting issues, thus leading to reduced performance.

(Groden et al., 2021) developed a model for improving patient retention within an intensive primary care practice. Here, the intervention solutions, including clinic signage, appointment reminder cards, and so on, were provided for patient retention. The research provided excellent outcomes and increased patient care. Nevertheless, it failed to analyze the relationship between the features, which affected the final outcomes.

3. PROPOSED ABG-Fuzzy AND NER-GFJDKMeans-ENABLED HOSPITAL RATINGS AWARE PATIENT CHURN PREDICTION AND PREVENTION SYSTEM

Here, the proposed L3STM is introduced to predict the patient readmit and hospital ratings and classify the patient review. For predicting the patient chunk, the proposed ABG-Fuzzy is established. Also, the proposed NER-GFJDKMeans is employed to provide retain strategies for high and medium patient churn. In order to select the optimal features, the proposed QBSOA is utilized. Likewise, the proposed SBMOTE is used to balance the data. The diagrammatic layout of the proposed model is displayed in Figure 1.

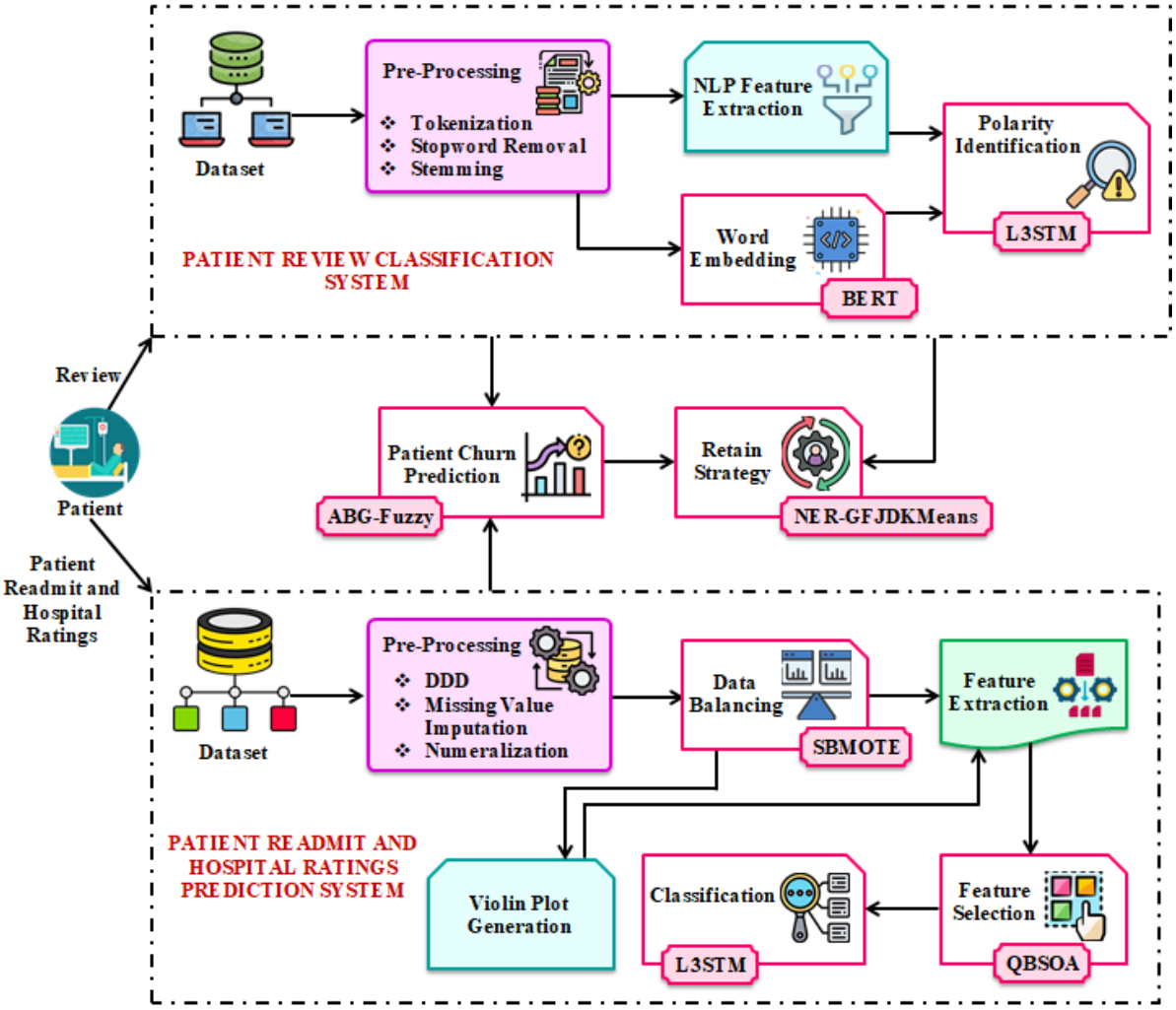

Figure 1: Diagrammatic layout of the proposed model

The proposed framework excellently predicts the patient chunks and provides interventions for avoiding client loss and revenue loss. The step-by-step working procedure of the proposed model is explained in the upcoming sections.

3.1 Patient Review Classification System

In the patient review classification system, significant processes, such as dataset, pre-processing, NLP feature extraction, word embedding, and polarity identification are done. It is described as follows,

3.1.1 Dataset

Firstly, the "Doctor review dataset (has reviews on doctors)" is collected from publicly available sources to train the patient review classification system. Here, this dataset consists of reviews of patients about the doctor and their treatment. The total k number of patient review data $\left(\wp_g^{pt}\right)$ is expressed as,

$$\wp_g^{pt} \rightarrow \left[\wp_1^{pt} + \wp_2^{pt} + \wp_3^{pt} + \ldots\ldots + \wp_k^{pt}\right] \quad (1)$$

Where, $\wp_k^{pt}$ specifies the number of $\left(\wp_g^{pt}\right)$.

3.1.2 Pre-Processing

After that, the patient review data $\left(\wp_g^{pt}\right)$ are pre-processed to improve the quality of the data. Here, the pre-processing step consists of processes, including tokenization, stop word removal, and stemming. Initially, the tokenization process is performed on $\left(\wp_g^{pt}\right)$ to break down the review sentence into words. The tokenized data is denoted as ϖ_{dt}. After that, stop words (i.e., the, is, and, in, of) are removed from ϖ_{dt} in order to focus on more meaningful words. The stop word removed data is indicated as ς_v. Subsequently, the stemming process

is performed on ς_v to reduce the words to their root form by removing suffixes and prefixes. The stemming outcomes are represented as tem. Lastly, the pre-processed data $(\Im_\kappa)$ is given as,

$$\Im_\kappa \Rightarrow \{\Im_1, \Im_2, \Im_3, \ldots\ldots, \Im_{ab}\} \quad Here \quad \kappa = (1\ to\ ab) \tag{2}$$

Where, $\Im_{ab}$ signifies the ab^{th} pre-processed data.

3.1.3 NLP Feature Extraction

From $(\Im_\kappa)$, the NLP features, such as Word Count, Character Count, Average Word Length, Unique Word Count, Type-Token Ratio (TTR), Stopword Count, Part-of-Speech (PoS), Sentence Count, Dependency Parsing, Term Frequency-Inverse Document Frequency (TF-IDF), and so on are extracted. The extracted NLP features are written as λ_y^{nlp}.

3.1.4 Word Embedding

Likewise, word embedding is performed on $(\Im_\kappa)$ by utilizing BERT to convert words into numerical representations. BERT excellently understands complex relationships between words, phrases, and sentences. The process of BERT is described as follows,

❄ Primarily, each $(\Im_\kappa)$ is converted into a vector representation. It is mathematically expressed as,

$$F(\Im_\kappa) = To + Po + Se \tag{3}$$

Where, F indicates final input embeddings, To implies token embeddings, Po specifies positional embeddings, and Se is segment embeddings.

❄ In the next step, each token's embeddings are converted into Query (Qr), Key (Key), and Value (Vl) matrices and are defined as,

$$\begin{aligned} Qr &= To\,\varpi^{Qr} \\ Key &= To\,\varpi^{Key} \\ Vl &= To\,\varpi^{Vl} \end{aligned} \tag{4}$$

Here, ϖ defines the weight matrix. After that, the attention score (Att) is estimated as,

$$Att = soft\max\left(\frac{QrKey^{Tr}}{d}\right)Vl \tag{5}$$

Where, $soft\max$ is the softmax function, d is the dimension of the embeddings, and Tr implies the transpose matrix. Then, the outcomes of (Att) are concatenated to obtain multiple attention heads. Subsequently, each token is passed through a feed-forward layer (Ff) and is determined as,

$$Ff(F) \Rightarrow \mathrm{Re}\,LU(F\varpi_1 + B_1)\varpi_2 + B_2 \tag{6}$$

Where, $\mathrm{Re}\,LU$ signifies Rectified Linear Unit function, and B_1 indicates the bias term. Lastly, the word embedding outcomes are denoted as W_ε.

3.1.5 Polarity Identification

Based on W_ε and λ_y^{nlp}, the polarity identification is done by employing Long Smish tanhexp Stochastic depth Short Term Memory (L3STM). Generally, Long Short Term Memory (LSTM) has the ability to capture long-term dependencies in sequential data. Also, it aids in eliminating irrelevant past information, leading to better model efficiency and performance. Nevertheless, LSTM has computational complexity due to an improper activation. Likewise, LSTM has overfitting issues. To address the computational complexity, the Smish TanhExp activation function is employed. Similarly, in order to overcome the overfitting problem, the Stochastic Depth regularization technique is used in LSTM. The structural diagram of the L3STM classifier is shown in Figure 2.

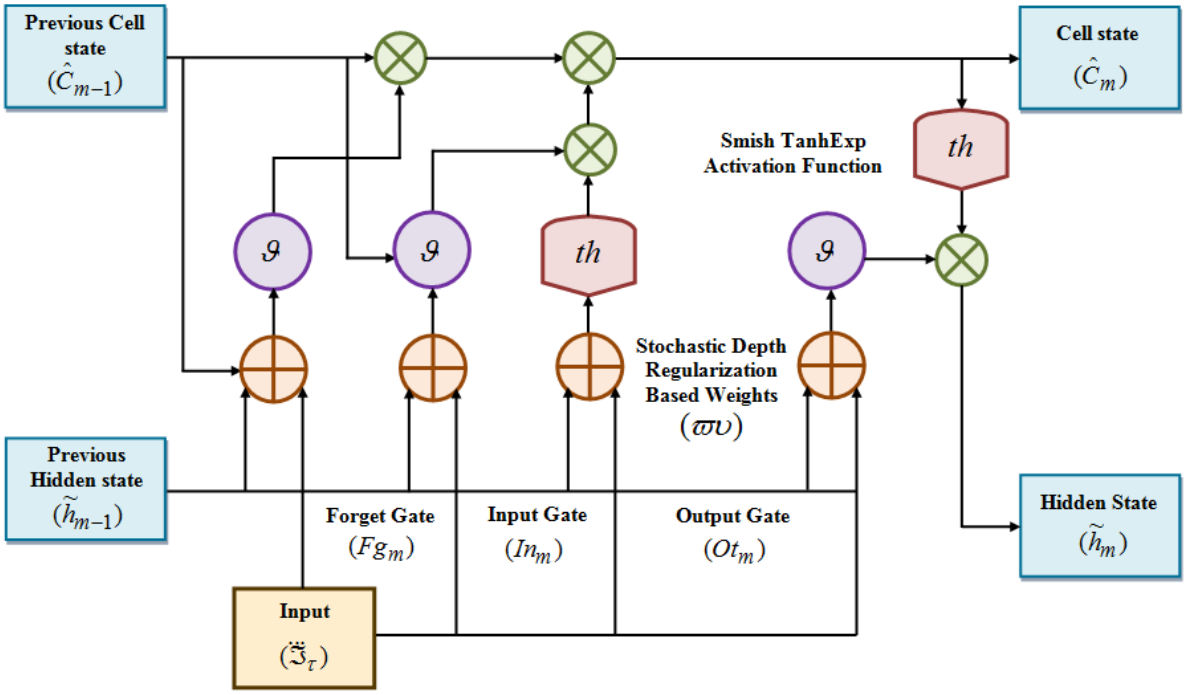

Figure 2: Structural diagram of L3STM classifier

Here, the L3STM consists of forget gate, input gate, candidate cell state, cell state, and output gate. Here, the inputs W_ε and λ_y^{nlp} are considered as $\ddot{\Im}_\tau$. The step-by-step mathematical expression of L3STM is given below,

➲ ***Forget Gate***

In the first step, the forget gate operation is performed; here, the forget gate $\left(Fg_m\right)$ removes the irrelevant information from the gates and is mathematically expressed as,

$$Fg_m = \vartheta \times \left\langle \left\{ \varpi\upsilon \times \left| \ddot{\Im}_\tau \cup \tilde{h}_{m-1} \right| \right\} + Bs \right\rangle \quad (7)$$

$$\vartheta = t\left(\ddot{\Im}_\tau\right)^2 \cdot th\left(\ln\left(1+\sigma\exp\left(u\ddot{\Im}_\tau\right)\right)\right) \quad (8)$$

$$\varpi\upsilon = \ddot{\Im}_\tau + sp \cdot fn\left(\ddot{\Im}_\tau, Wi\right) \quad (9)$$

Where, Bs implies the bias term, $\tilde{h}_{m-1}$ represents the previous hidden state, ϑ indicates the Smish TanhExp activation function, which is employed to avoid the computational complexity, $\varpi\upsilon$ signifies the Stochastic Depth regularization-based weights, which is used to solve the overfitting issue, t and u signify the parameters, th denotes hyperbolic tangent function, σ is the logistic sigmoid, exp depicts the exponential function, $fn\left(\ddot{\Im}_\tau, Wi\right)$ defines the transformation function, and sp denotes the survival probability of the layer.

➲ ***Input Gate***

After that, the input gate discovers the significant information that is needed to pass through the gates. The input gate operation is defined as,

$$In_m = \left\langle \left[\left\{ \varpi\upsilon \times \left| \ddot{\Im}_\tau \cup \tilde{h}_{m-1} \right| \right\} + Bs \right] * \vartheta \right\rangle \quad (10)$$

Where, In_m signifies the input gate operation.

➲ ***Candidate Cell State***

Subsequently, the candidate cell state is computed and is defined as,

$$C_m = \left\langle \left[\left\{ \varpi\upsilon \times \left| \ddot{\Im}_\tau \cup \tilde{h}_{m-1} \right| \right\} + Bs \right] * th \right\rangle \quad (11)$$

Where, $\left(C_m\right)$ denotes the candidate cell state operation.

➲ ***Cell State***

Afterward, in the cell state, the forget gate and input gate are concatenated, and it is represented as,

$$\hat{C}_m = \left(Fg_m \bullet \hat{C}_{m-1}\right) + \left(In_m \bullet C_m\right) \quad (12)$$

Where, $\left(\hat{C}_m\right)$ indicates the cell state operation, $\hat{C}_{m-1}$ is the previous cell state, and $\bullet$ denotes the element-wise multiplication.

➲ ***Output Gate***

Lastly, the decisions about the polarity identification are made by the output gate $\left(Ot_m\right)$, and it is indicated as,

$$Ot_m = \left\langle \left[\left\{ \varpi \upsilon \times \left| \ddot{\Im}_\tau \cup \tilde{h}_{m-1} \right| \right\} + Bs \right] * \vartheta \right\rangle \quad (13)$$

Likewise, the present hidden state $\left(\tilde{h}_m\right)$ is given as,

$$\tilde{h}_m = Ot_m \bullet \vartheta\left(\hat{C}_m\right) \quad (14)$$

The identified polarity $\left(Pol\right)$ is defined as,

$$Pol = \left\langle N, P \right\rangle \quad (15)$$

Where, N signifies negative review, and P indicates positive review. The pseudo-code for L3STM is depicted below,

Pseudocode for L3STM

Input: Word Embedding Outcomes $\left(W_\varepsilon\right)$ and Extracted NLP Features $\left(\lambda_y^{nlp}\right)$

Output: Identified Polarity $\left(Pol\right)$

Begin

Initialize cell state $\left(\hat{C}\right)$, hidden state $\left(\tilde{h}\right)$, and bias $\left(Bs\right)$

Consider $\left(W_\varepsilon\right)$ and $\left(\lambda_y^{nlp}\right)$ as $\ddot{\Im}_\tau$

For each $\ddot{\Im}_\tau$

Estimate forget gate

$$Fg_m = \vartheta \times \left\langle \left\{ \varpi \upsilon \times \left| \ddot{\Im}_\tau \cup \tilde{h}_{m-1} \right| \right\} + Bs \right\rangle$$

Compute

$$\vartheta = t\left(\ddot{\Im}_\tau\right)^2 \cdot th\left(\ln\left(1 + \sigma \exp\left(u\ddot{\Im}_\tau\right)\right)\right)$$

Discover Stochastic Depth regularization-based weights

$$\varpi \upsilon = \ddot{\Im}_\tau + sp \cdot fn\left(\ddot{\Im}_\tau, Wi\right)$$

Perform input gate operation

$$\left(In_m\right)$$

Evaluate

$$C_m = \left\langle \left[\left\{ \varpi \upsilon \times \left| \ddot{\Im}_\tau \cup \tilde{h}_{m-1} \right| \right\} + Bs \right] * th \right\rangle$$

Find cell state

$$\hat{C}_m = \left(Fg_m \bullet \hat{C}_{m-1}\right) + \left(In_m \bullet C_m\right)$$

Perform output gate $Ot_m = \left\langle \left[\left\{ \varpi \upsilon \times \left| \ddot{\Im}_\tau \cup \tilde{h}_{m-1} \right| \right\} + Bs \right] * \vartheta \right\rangle$

Compute present hidden state

End For

Obtain $Pol = \left\langle N, P \right\rangle$

End

Then, the $\left(Pol\right)$ are given for patient churn prediction, which is explained in further sections.

3.2 Patient Readmit and Hospital Ratings Prediction System

In the meantime, the patient readmit and hospital ratings prediction systems are trained based on the steps, such as pre-processing, data balancing, violin plot construction, feature extraction, feature selection, and classification. It is explained in further sections.

3.2.1 Dataset

Here, the "Patient Readmit dataset" and "Hospital Ratings dataset" are gathered to train the patient readmit and hospital ratings prediction system. In these datasets, the readmitting status of patients and ratings about the hospitals are available. The data in the dataset can be expressed as,

$$\left(\Re_h, H_z\right) \to \left[\left(\Re_1, H_1\right), \left(\Re_2, H_2\right), \ldots\ldots, \left(\Re_l, H_q\right)\right] \quad (16)$$

Where, $\Re_h$ depicts the Patient Readmit data, H_z signifies the hospital ratings data, $\Re_l$ indicates the number of $\Re_h$, and H_q is the number of H_z.

3.2.2 Pre-Processing

Thereafter, the pre-processing is done on $\Re_h$ to enhance the quality of the data. Initially, the duplicate data in $\Re_h$ are removed in the Data DeDuplication (DDD) process, and the remaining data is denoted as ∂d. After that, the missing values in ∂d are imputed by using the mean formula, and it is expressed as,

$$M(\partial d) = \frac{\sum \partial d}{\upsilon} \tag{17}$$

Where, M implies the missing value imputed data, and υ demonstrates the number of ∂d. Next, in the numeralization process, the M are converted into numerical values, and it is indicated as $\ddot{N}$. Lastly, the pre-processed patient readmit data $(A_{y'})$ is given as,

$$A_{y'} = \{A_1 + A_2 + A_3 + \ldots\ldots + A_{pq}\} \quad where \quad y' = (1,2,\ldots, pq) \tag{18}$$

Where, $y' = (1,2,\ldots, pq)$ defines the number of $(A_{y'})$. Also, the H_z is pre-processed based on the above steps, and it is represented as $\Xi_{q'}$.

3.2.3 Data Balancing

Next, data balancing is performed on $(A_{y'})$ by utilizing the Synthetic Brahmagupta Minority Oversampling TEchnique (SBMOTE) to increase the data size. Generally, the Synthetic Minority Oversampling TEchnique (SMOTE) is a widely used oversampling technique designed to handle a class imbalance in machine learning datasets. Although SMOTE reduces overfitting compared to simple oversampling, it can still lead to overfitting when applied to small datasets. To address this problem, Brahmagupta's interpolation technique is added in SMOTE. The step-by-step process of SBMOTE is given below,

In SBMOTE, firstly, for oversampling the minority classes, the synthetic instances are created. Here, the minority and majority classes are denoted as $\min(A_{y'})$ and $maj(A_{y'})$, correspondingly. After that, from the $\min(A_{y'})$, a minority instance $(\aleph_a)$ is randomly selected and is expressed as,

$$\aleph_a = (\aleph_1, \aleph_2, \aleph_3, \ldots\ldots, \aleph_i) \quad Where \quad a = 1,2,\ldots, i \tag{19}$$

Where, i indicates the number of minority instances. In the next step, the nearest neighbor (ϕ) of the minority instance $(\aleph_a)$ is discovered. Here, Brahmagupta's interpolation technique is employed to avoid the overfitting issue. It is given as,

$$\phi(\aleph_a, \min(A_{y'})) = \min(A_2) + \frac{(\aleph - \aleph_2)(\min(A_3) - \min(A_1))}{\aleph_3 - \aleph_1} \tag{20}$$

Here, $(\aleph_1, \min(A_1))$, $(\aleph_2, \min(A_2))$, and $(\aleph_3, \min(A_3))$ are known data points, and $\aleph$ signifies the value at which interpolation is performed. Thereafter, the random neighbor $(\tilde{\phi})$ is selected and is indicated as $\tilde{\phi} \in \min(A_{y'})$. Afterward, the new instance $(n\varepsilon)$ is created as,

$$n\varepsilon = \aleph_a + (\tilde{\phi} - \aleph_a) \times rd \tag{21}$$

Where, rd depicts the random number. Then, $(n\varepsilon)$ is added to $A_{y'}$ for minority class. Lastly, the balanced data is indicated as Bl_π. Likewise, the $\Xi_{q'}$ is balanced and is denoted as $\hbar$.

3.2.4 Violin Plot Generation

For Bl_π, a violin plot is generated to understand the distribution of numerical features and compare them across different categories, and it is signified as δ_{vi}. Also, the violin plot is generated for the $\hbar$ and is implied as ζ_{mn}.

3.2.5 Feature Extraction

From δ_{vi}, the features, including Mean, Median, IQR, Skewness, Kurtosis, and Number of Outliers, are extracted. Likewise, the features, including race, gender, age, weight, admission_type_id, discharge_disposition_id, admission_source_id, time_in_hospital, metformin, and so on, are extracted from the Bl_{π}. These extracted features (V_j) are signified as,

$$V_j \Rightarrow [V_1, V_2, , V_3, \ldots\ldots, , V_r] \quad (22)$$

Where, V_r depicts the r^{th} extracted feature. Similarly, from the ζ_{mn}, the features, including Mean, Median, IQR, Skewness, Kurtosis, and Number of Outliers, are extracted. Also, the features, such as Provider ID, Hospital Name, Address, City, State, ZIP Code, County Name, Phone Number, Hospital Type, Hospital Ownership, Emergency Services, and so on are extracted from the $\hbar$. These extracted features are represented as K.

3.2.6 Feature Selection

After that, the optimal features are selected from (V_j) by employing the Quokka Bernoulli Swarm Optimization Algorithm (QBSOA) for efficient patient readmit prediction. Generally, the Quokka Swarm Optimization Algorithm (QSOA) provides excellent outcomes even though inconsistent values are presented in the input data. In QSOA, the position of the population is mainly updated based on the drought function. However, the drought function is chosen randomly between [0, 1], which results in premature convergence. To address this problem, the Bernoulli chaotic map function is included in QSOA. The working of QBSOA is derived below,

Primarily, the population of quokka is initialized regarding its position. Here, the extracted features (V_j) are considered as the initialized population. The initialized population (D_p) is represented as,

$$D_p = (D_1 + D_2 + D_3 + \ldots\ldots + D_{pq}) \quad (23)$$

Where, D_{pq} defines the pq^{th} population member. After that, the fitness function (f) is estimated; here, the maximum classification accuracy $(\max(Acc))$ is considered as the fitness function. It is written as,

$$f \rightarrow \max(Acc) * f(D_p) \quad (24)$$

$$(D_p)_{best} \Rightarrow Best(f) \quad (25)$$

Where, $(D_p)_{best}$ depicts the best candidate solution, and $Best(f)$ represents the best fitness value.

Subsequently, the position update of each quokka within a group (J^{new}) is formulated as follows,

$$J^{new} = \frac{(T + mi)}{0.8 \times Bou} + \Delta w \times rnd \times \Delta(D_p) \quad (26)$$

$$(D_p)^{new} = (D_p)^{old} + J^{new} \times tro \quad (27)$$

$$Bou = \begin{cases} \dfrac{D_p}{1-ct} & 0 < D_p \leq 1-ct \\ \dfrac{D_p - (1-ct)}{ct} & 1-ct < D_p < 1 \end{cases} \quad (28)$$

Here, mi elucidates the ratio of the humidity, T indicates the ratio of the temperature, Δw is the differences of weight between the leader and the quokka, rnd represents the random number with interval (0 to 1), $(D_p)^{old}$ is the old position of quokkas, tro indicates the ratio of nitrogen, $(D_p)^{new}$ implies the updated position of quokkas, $\Delta(D_p)$ implies the differences of position between the leader and quokka, Bou defines Bernoulli chaotic map function, which is employed to overcome the pre-mature convergence issue, and ct specifies the control parameter. This position updation is continued until discovering the optimal solution. The selected features (S_ω) are given as,

$$S_\omega \rightarrow (S_1 + S_2 + S_3 + \ldots\ldots + S_n) \quad Here \quad \omega = (1,2,\ldots,n) \quad (29)$$

Where, S_n implies the number of selected features. The pseudocode for QBSOA is given as follows,

Pseudocode for QBSOA

Input: Extracted features (V_j)

Output: Selected features (S_ω)

Begin

Initialize (V_j)

For each (V_j)

Initialize (D_p)

Find fitness function

$$f \rightarrow \max(Acc) * f(D_p)$$

$$(D_p)_{best} \Rightarrow Best(f)$$

Implement position update of each quokka within a group

$$J^{new} = \frac{(T + mi)}{0.8 \times Bou} + \Delta w \times rnd \times \Delta(D_p)$$

Update new position

$$(D_p)^{new} = (D_p)^{old} + J^{new} \times tro$$

Discover $Bou = \begin{cases} \frac{D_p}{1-ct} & 0 < D_p \leq 1-ct \\ \frac{D_p - (1-ct)}{ct} & 1-ct < D_p < 1 \end{cases}$

Continue until discovering the optimal solution

End For

Obtain (S_ω)

End

Likewise, based on the above process, the optimal features are selected from K for efficient hospital ratings prediction, and it is denoted as $\ddot{\chi}$.

3.2.7 Classification

Based on S_ω, the patient readmit prediction is performed using L3STM. The working of L3STM is described in Section 3.1.5. The patient readmits prediction outcomes (PR) are signified as,

$$PR = \langle X, \ell, \lambda \rangle \quad (30)$$

Where, X signifies no readmission, ℓ denotes readmission of patients within 30 days of discharge, and λ exemplifies readmission of patients after 30 days of discharge.

Similarly, the hospital ratings are predicted based on the $\ddot{\chi}$ by employing L3STM. The predicted hospital ratings (Hr) are expressed as,

$$Hr = \{1,2,3,4,5\} \quad (31)$$

Next, the (PR) and (Hr) are given for patient churn prediction.

3.3 Patient Churn Prediction

Regarding the (Pol), (PR), and (Hr), the patient churn is predicted. Here, Asymmetric Bézier Gaussian-based Fuzzy (ABG-Fuzzy) is employed for patient churn prediction. Fuzzy algorithm is easy to implement, and it has the ability to provide effective solutions for complex problems. However, the Fuzzy system has tuning difficulty due to the membership function, which causes improper decisions. To address this problem, the Asymmetric Bézier Gaussian (ABG) membership function is employed in Fuzzy. The working of ABG-Fuzzy is derived below,

Firstly, if-then conditions are employed to compute the fuzzy rules (χ) and are expressed as,

$$\chi \xrightarrow{Pol,PR,Hr} \begin{cases} if\ Pol = N \ \&\& \ PR = \ell \ \&\& \ Hr \le 2 & Then \quad hg \\ if\ Pol = N \ \&\& \ PR = \lambda \ \&\& \ Hr \le 3 & Then \quad mm \\ if\ Pol = P \ \&\& \ PR = X \ \&\& \ Hr \ge 4 & Then \quad lw \end{cases} \tag{32}$$

Afterward, to address the tuning difficulty due to the membership function, the ABG membership function (Θ) is utilized. This divides the input space into multiple linear segments and allows for a more accurate representation of systems. It is equated as,

$$\Theta = \begin{cases} \exp^{-\left(\frac{(Pol,PR,Hr)-e}{Y(e-b)}\right)} & b \le (Pol,PR,Hr) < e \\ \exp^{-\left(\frac{(Pol,PR,Hr)-e}{Z(c-e)}\right)} & e \le (Pol,PR,Hr) \le c \end{cases} \tag{33}$$

Here, b, c, and e are control points, and Y and Z are shape parameters. Here, the decision-making unit performs interference operations, such as fuzzification and defuzzification. In fuzzification interference $(\Im\mu)$, the crisp data $(c\Re)$ are converted into fuzzy data (χ) using Θ, and it is formulated as,

$$\Im\mu = (c\Re \to \chi) \tag{34}$$

Likewise, in defuzzification interference $(\partial\varepsilon)$, the inverse operation of converting the fuzzy data (χ) into crisp data $(c\Re)$ is performed and is given as,

$$\partial\varepsilon = (\chi \to c\Re) \tag{35}$$

The predicted patient churn is signified as,

$$\Omega = \{hg, mm, lw\} \tag{36}$$

Where, hg indicates high churn, mm signifies medium churn, and lw demonstrates low churn.

3.4 Retain Strategy

For hg and mm, the retain strategy is provided regarding the patient review $(\wp_g^{pt})$. Firstly, the interventions (ψ_μ) are clustered under different categories like personalized patient engagement (regarding automated appointment reminders, personalized health check-ups, and follow-up calls), improving patient experience and service quality (regarding reducing wait times, enhancing doctor-patient communication, and offering virtual consultations), and so on. Here, Named Entity Recognition-based Greedy Forgy Jeffreys Damerau K-Means (NER-GFJDKMeans) is employed for estimating retain strategy. In general, K-Means effectively ensures that all the data are uniformly distributed within each cluster. However, K-Means have trouble in clustering varying densities of clusters. To overcome this problem, Jeffreys Damerau Distance is employed in K-Means. Also, K-Means has issues with random centroid initialization. To address this issue, Greedy Forgy initializers are utilized in K-Means.

Step 1: Primarily, by employing Greedy Forgy initializers (β), the centroids are initialized, and it is formulated as,

$$\beta = \arg \max_{\psi_\mu}(\psi_\mu) \tag{37}$$

Where, arg signifies argument, and $\max_{\psi_\mu}$ defines the maximum value of ψ_μ. The initialized centroids (G_s) are given as,

$$G_s \xrightarrow{\beta} (G_1, G_2, G_3, \ldots\ldots, G_x) \quad Here \quad s = (1,2,\ldots,x) \tag{38}$$

Where, G_x represents the x^{th} initialized centroids.

Step 2: Subsequently, the Jeffreys Damerau distance is computed between the (G_s) and (ψ_μ). It is defined as,

$$Ds(G_s, \psi_\mu) = \sum_{s,\mu} (G_s - \psi_\mu) \log\left(\frac{G_s}{\psi_\mu}\right) \tag{39}$$

Where, Ds defines the Jeffreys Damerau distance, and $\log$ depicts the logarithmic function.

Step 3: After that, for all data that belong to each cluster, the average is found. Thereafter, the new centroid is estimated for each cluster and is written as,

$$\Phi = \frac{\sum_{s=1,\mu=1}^{x,L} Avg_{s\mu}\psi_{\mu}}{\sum_{s=1,\mu=1}^{x,L} \psi_{\mu}} \tag{40}$$

Where, Φ demonstrates the new centroid, L implies the number of ψ_{μ}, and $Avg_{s\mu}$ indicates the average of data. Lastly, each node is reassigned to the new closest centroid of each cluster. The above-described steps are continued until optimal clusters are found. The grouped interventions are denoted as η_{v}.

Then, NER is performed for patient review $\left(\wp_{g}^{pt}\right)$. From that, keywords and corresponding entity names are obtained. Then, these NER outcomes are matched with grouped interventions $\left(\eta_{v}\right)$ for obtaining a retain strategy. The obtained retain strategy is indicated as R_{σ}. Thus, the proposed framework excellently performed patient churn prediction and provided corresponding interventions.

4. RESULT AND DISCUSSION

In this section, the proposed framework's performance in patient churn prediction is depicted by comparing it with several conventional methodologies. The experimental analysis of this proposed work is implemented in the working platform of PYTHON.

4.1 Dataset Description

This proposed work uses three datasets, namely Doctor Review Dataset, Hospital Review Dataset, and Diabetes 130-US hospitals for the years 1999-2008 dataset. The source link to access these datasets is demonstrated in the reference section. Here, the patients' reviews about doctors are collected using the doctor review dataset, which contains the review of patients, labels, and tags. Likewise, the hospital review dataset helps in comparing the quality of care at over 4000 Medicare-certified hospitals across the country. Similarly, the patient readmits are predicted using the Diabetes 130-US hospitals for years 1999-2008 dataset, which includes the hospital records of patients diagnosed with diabetes, who underwent laboratory, medications, and stayed up to 14 days. From these datasets, the proposed framework uses 80% data for training and 20% data for testing.

4.2 Performance Analysis for the proposed L3STM

In this section, the L3STM's performance is validated by comparing it with several conventional methodologies, namely LSTM, Recurrent Neural Network (RNN), Deep Belief Network (DBN), and Deep Neural Network (DNN) regarding patient review classification, patient readmits prediction, and hospital rating prediction.

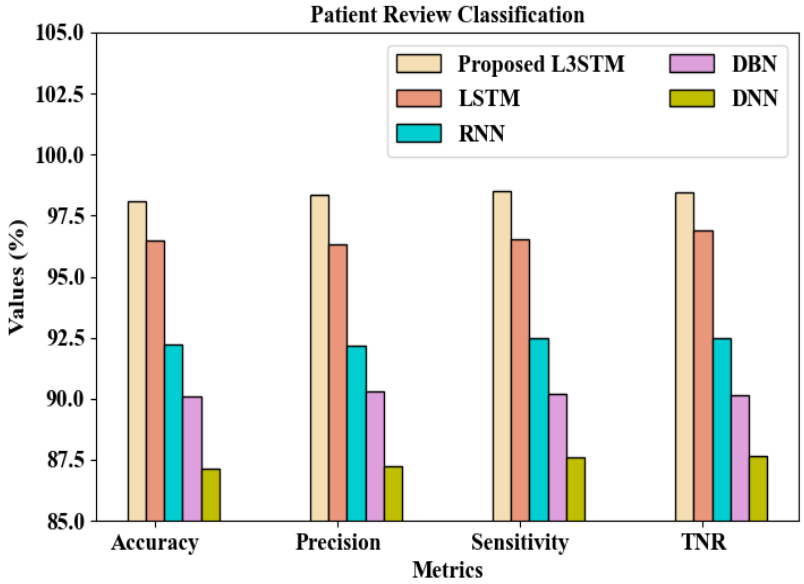

Figure 3: Efficiency analysis for patient review classification

Figure 3 illustrates the L3STM's performance in patient review classification. Here, the L3STM attains higher accuracy (98.1045%), precision (98.3256%), sensitivity (98.5148%), and True Negative Rate (TNR) (98.4578%). Thus, by integrating the smish tanhexp activation function and stochastic depth regularization technique, the proposed L3STM enhances the model's efficiency by reducing computational complexity and mitigating overfitting. Hence, the L3STM outperforms the other conventional methodologies regarding accurate patient review classification.

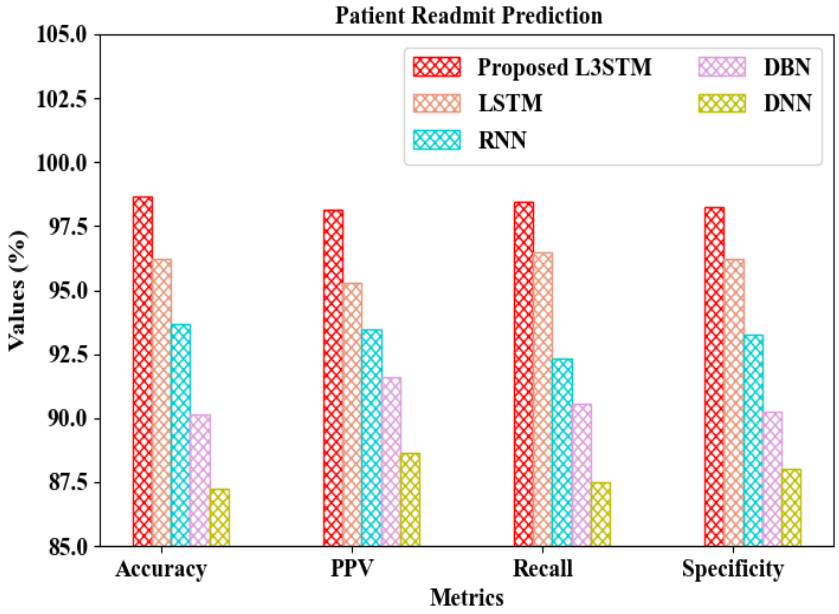

Figure 4: Effectiveness validation for patient readmits prediction

The L3STM's effectiveness in patient readmit prediction is validated in Figure 4. Here, the L3STM includes the Smish TanhExp activation function and Stochastic Depth regularization technique for accurate patient readmitting prediction. Thus, by using these approaches, the L3STM efficiently handles the very long sequence of patient records while reducing computational complexity. As a result, the L3STM generalized well in patient readmit prediction by attaining 98.65% accuracy, 98.15% Positive Predictive Value (PPV), 98.45% recall, and 98.26% specificity. Therefore, it is noted that the L3STM outperforms the other traditional approaches.

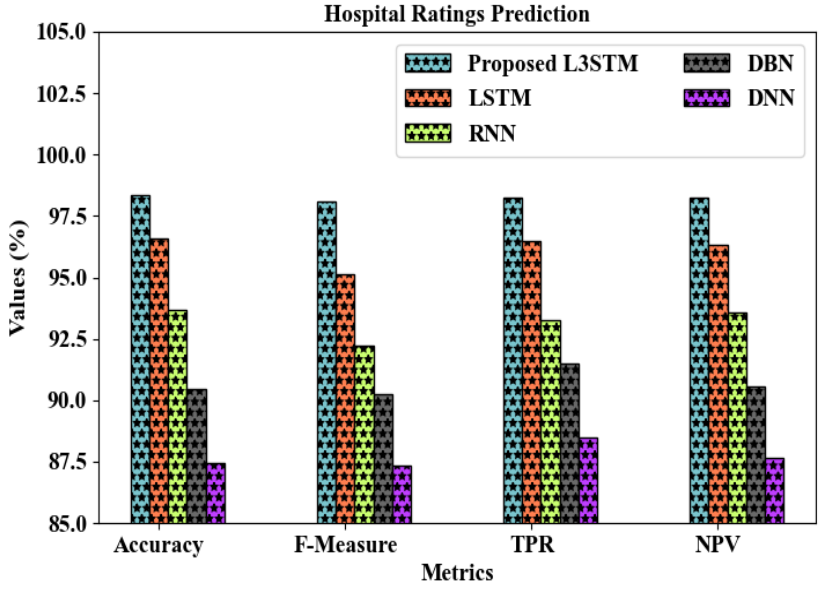

Figure 5: Performance evaluation for hospital rating prediction

Figure 5 demonstrates the L3STM's efficiency in hospital rating prediction for accurate patient churn prediction. From Figure 5, it is noted that the L3STM generalizes better in predicting the hospital rating than the other prevailing techniques. The L3STM attains 98.35% accuracy, 98.104% f-measure, 98.24% True Positive Rate (TPR), and 98.23% Negative Predictive Value (NPV). However, traditional approaches like LSTM, RNN, DBN, and DNN attain lower accuracy of 96.56%, f-measure of 92.23%, TPR of 91.48%, and NPV of 87.63%, respectively.

4.3 Effectiveness Evaluation for the proposed ABG-Fuzzy

In this section, the ABG-Fuzzy's efficiency in patient churn prediction is demonstrated by comparing it with several conventional methodologies like FLS, Triangular Fuzzy (TF), Sigmoid Fuzzy (SF), and Gaussian Fuzzy (GF).

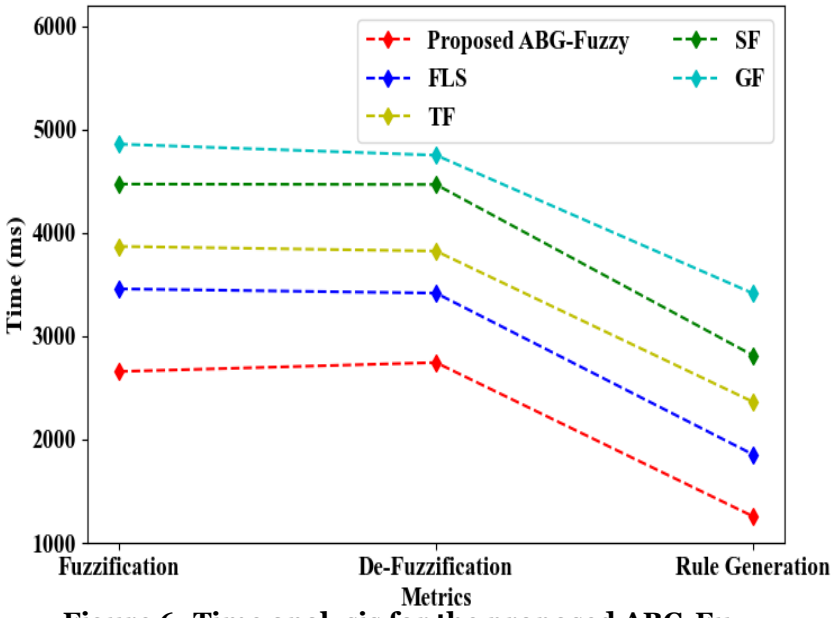

Figure 6: Time analysis for the proposed ABG-Fuzzy

Figure 6 depicts the ABG-Fuzzy's efficiency regarding its fuzzification time, defuzzification time, and rule generation time for accurate patient churn prediction. Here, the ABG-Fuzzy converges faster than the other conventional methodologies in predicting patient churn. The ABG-Fuzzy takes 2658ms, 2745ms, and 1257ms of time for fuzzification, defuzzification, and rule generation. Hence, by using the Asymmetric Bezier Gaussian (ABG) membership function, the ABG-Fuzzy efficiently divides the input space into multiple linear segments. This improves decision-making accuracy and reduces computational overhead, thus leading to faster convergence.

4.4 Efficacy Validation for the proposed NER-GFJDKMeans

This section elaborates on the NER-GFJDKMeans's efficacy by comparing it with several prevailing approaches like K-Means, Clustering for LARge Application (CLARA), Partitioning Around Medoids (PAM), and Fuzzy C-Means (FCM) for accurate patient churn prediction.

Table 1: Clustering time analysis

Methodologies	Clustering Time (ms)
Proposed NER-GFJDKMeans	14875
K-Means	18693
CLARA	23254
PAM	27418
FCM	32569

The clustering time of the NER-GFJDKMeans is analyzed in Table 1. From this table, it is proved that the NER-GFJDKMeans converges faster in clustering the data for accurate patient churn prediction than the other conventional approaches. This is because of the integration of Keffreys Damerau distance and Greedy Forgy initializer. Thus, by enhancing the clustering accuracy and mitigating the randomness, the NER-GFJDKMeans ensures uniform data distribution within the clusters and converges faster. The NER-GFJDKMeans takes a minimum clustering time of 14875ms.

4.5 Performance Validation for the proposed SBMOTE

The SBMOTE's performance in data balancing is illustrated by comparing it with several traditional approaches like SMOTE, SMOTE for Nominal (SMOTEN), SMOTE for Nominal and Continuous (SMOTENC), and ADAptive SYNthetic (ADASYN).

Table 2: Precision validation for the proposed SBMOTE

Techniques	Precision (%)
Proposed SBMOTE	98.0248
SMOTE	95.3621
SMOTEN	92.2014
SMOTENC	89.4518
ADASYN	87.4015

Table 2 demonstrates the SBMOTE's performance in data balancing for accurate patient churn prediction regarding precision. As per the result, the SBMOTE generates more diverse and realistic synthetic samples by mitigating the overfitting using Brahmagupta's interpolation technique. This improves class balance, thus leading to higher precision in patient churn prediction. Here, the SBMOTE balances the data efficiently by attaining a higher precision of 98.0248%. However, the conventional SMOTE, SMOTEN, SMOTENC, and ADASYN attains lower precision of 95.3621%, 92.2014%, 89.4518%, and 87.4015%, respectively.

4.6 Efficiency Analysis for the proposed QBSOA

Now, the QBSOA's efficiency in feature selection for accurate patient churn prediction is depicted by comparing it with several existing methodologies, such as the Quokka Swarm Optimization Algorithm (QSOA), Tuna Swarm Optimization Algorithm (TSOA), Slap Swarm Optimization Algorithm (SSOA), and Particle Swarm Optimization Algorithm (PSOA).

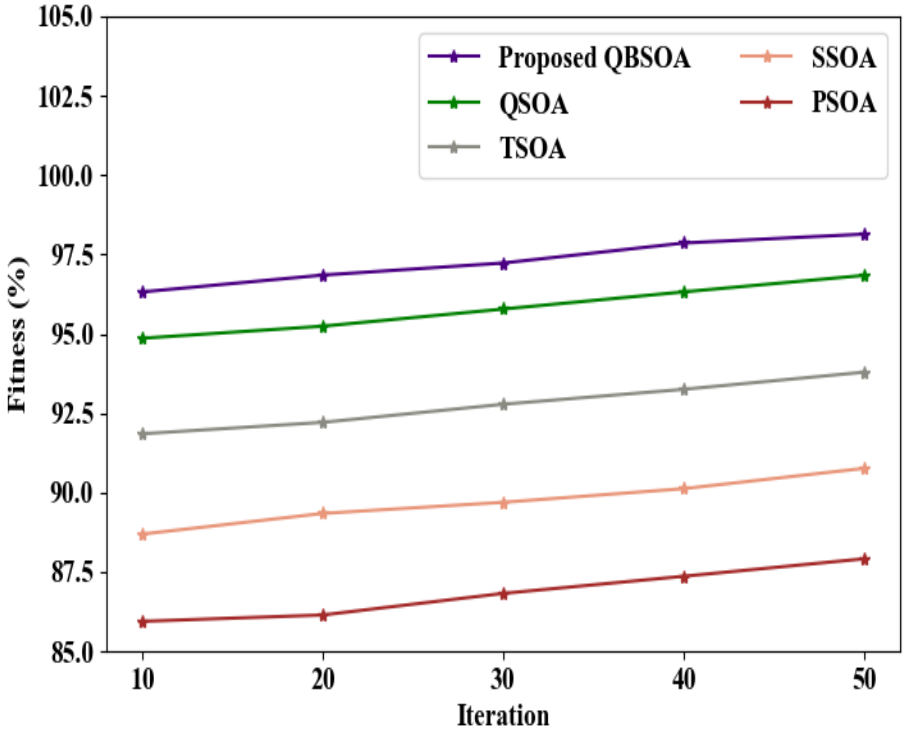

Figure 7: Fitness Validation

Figure 7 illustrates the QBSOA's performance in feature selection based on its fitness value. Here, the QBSOA attains 96.32% fitness for 10 iterations and 98.14% for 50 iterations, thus indicating that the QBSOA's fitness increases as the number of iterations for selecting features also increases. This is because of the integration of the Bernoulli Chaotic map in the exploitation phase, which improves the exploration and exploitation balance. Hence, the QBSOA outperforms the other prevailing methodologies by enhancing the search efficiency and preventing QBSOA from getting stuck in local optima, thus achieving higher fitness values and ensuring feature selection.

4.7 Comprehensive Analysis

The comprehensive analysis of the proposed framework in patient churn prediction is demonstrated in this section.

Table 3: Comparative analysis for the proposed framework in patient churn prediction

Authors' Name	Objectives	Methodologies	Accuracy (%)
Proposed Framework	Proficient hospital-aware patient churn prediction and prevention model	L3STM	98.65
(Ruiz de San Martín et al., 2024)	Local predictive model for early individual unscheduled hospital readmission identification	Interpretable ML approaches	74.5
(Panchangam et al., 2024)	Readmission risk prediction using patient's electronic medical record data	eXtreme Gradient Boosting version 1.7.6 (XGBoost v1.7.6)	82
(Jajam et al., 2023)	Customer churn prediction in the insurance industry	Stacked Bidirectional LSTM and RNN for Arithmetic Optimization Algorithm (AOA-SBLSTM-RNN)	97.89
(Jamjoom, 2021)	Customer churn prediction in Business-to-Business environment	Decision Tree (DT)	90
(Luo et al., 2023)	Readmissions prediction in a united states' healthcare system	XGBoost	71.3

Table 3 demonstrates the comparative analysis of the proposed framework in patient churn prediction by comparing it with several conventional works. As per the results, the proposed framework outperforms the other prevailing works by attaining higher accuracy (98.65%). Although the conventional approaches like interpretable ML technique, XGBoost v1.7.6, AOA-SBLSTM-RNN, DT, and XGBoost provide advantages like better interpretability, scalability, and efficiency, they often struggle to capture complex temporal dependencies and contextual inherent in patient churn prediction. For instance, the interpretable ML technique, XGBoost v1.7.6, AOA-SBLSTM-RNN, DT, and XGBoost attains lower accuracy of 74.5%, 82%, 97.89%, 90%, and 71.3%, respectively.

5. CONCLUSION

In this paper, a proficient hospital ratings-aware patient churn prediction and prevention system was proposed using the ABG-Fuzzy and NER-GFJDKMeans. Initially, the proposed framework trained the patient review classification model for polarity identification. Here, the proposed framework accurately identified the patient polarity by processing the patient review data using the L3STM, which attained 98.1045% accuracy. Meanwhile, the patient readmits and hospital ratings prediction models were trained by collecting the data from the patient readmit and hospital rating datasets, respectively. Here, the class balance was improved using the SBMOTE with a precision of 98.02%. Further, the optimal features were selected using the QBSOA by attaining 98.14% fitness. Also, by using the L3STM, the proposed framework accurately predicted the patient readmits and hospital rating with an accuracy of 98.65% and 98.35%, respectively. Moreover, the ABG-Fuzzy accurately predicted the patient churn within a minimum fuzzification and defuzzification time. Finally, the proposed framework efficiently provided the prevention mechanisms based on the patient reviews using the NER-GFJDKMeans. Thus, the proposed framework outperformed the other prevailing approaches by integrating advanced methodologies, which ensured precise hospital ratings, accurate patient readmit predictions, and effective prevention strategies.

Future Scope: Yet, the proposed work failed to implement patient churn prediction by designing to understand complex client relationships and identifying hidden patterns in patient data. Future work will focus on addressing these capabilities for enabling efficient client segmentation and targeting.

REFERENCE

1. https://www.kaggle.com/datasets/avasaralasaipavan/doctor-review-dataset-has-reviews-on-doctors
2. https://www.kaggle.com/datasets/center-for-medicare-and-medicaid/hospital-ratings
3. https://archive.ics.uci.edu/dataset/296/diabetes+130-us+hospitals+for+years+1999-2008
4. Afzal, M., Rahman, S., Singh, D., & Imran, A. (2024). Cross-sector application of machine learning in telecommunications: enhancing customer retention through comparative analysis of ensemble methods. *IEEE Access*. https://doi.org/10.1109/ACCESS.2024.3445281
5. Ahmad, N., Awan, M. J., Nobanee, H., Zain, A. M., Naseem, A., & Mahmoud, A. (2023). Customer Personality Analysis for Churn Prediction Using Hybrid Ensemble Models and Class Balancing Techniques. *IEEE Access*. https://ieeexplore.ieee.org/abstract/document/10322869/

6. Ajegbile, M. D., Olaboye, J. A., Maha, C. C., Igwama, G. T., & Abdul, S. (2024). The role of data-driven initiatives in enhancing healthcare delivery and patient retention. *World Journal of Biology Pharmacy and Health Sciences*, *19*(1), 234-242. https://doi.org/10.30574/wjbphs.2024.19.1.0435
7. Albright, B. B., Nitecki, R., Chino, F., Chino, J. P., Havrilesky, L. J., Aviki, E. M., & Moss, H. A. (2022). Catastrophic health expenditures, insurance churn, and nonemployment among gynecologic cancer patients in the United States. *American journal of obstetrics and gynecology*, *226*(3), 384-e1. https://doi.org/10.1016/j.ajog.2021.09.034
8. Arredondo, K., Hughes, A. M., Lester, H. F., Pham, T. N., Petersen, L. A., Woodard, L., ... & Hysong, S. J. (2024). Churning the tides of care: when nurse turnover makes waves in patient access to primary care. *BMC nursing*, *23*(1), 739. https://doi.org/10.1186/s12912-024-02389-8
9. Bui, Q. N., & Moriuchi, E. (2021). Economic and social factors that predict readmission for mental health and drug abuse patients. *Sustainability*, *13*(2), 531. https://doi.org/10.3390/su13020531
10. Bustea, C., Tit, D. M., Bungau, A. F., Bungau, S. G., Pantea, V. A., Babes, E. E., & Pantea-Roşan, L. R. (2023). Predictors of Readmission after the First Acute Coronary Syndrome and the Risk of Recurrent Cardiovascular Events—Seven Years of Patient Follow-Up. *Life*, *13*(4), 950. https://doi.org/10.3390/life13040950
11. Chang, V., Hall, K., Xu, Q. A., Amao, F. O., Ganatra, M. A., & Benson, V. (2024). Prediction of Customer Churn Behavior in the Telecommunication Industry Using Machine Learning Models. *Algorithms*, *17*(6), 231. https://doi.org/10.3390/a17060231
12. Chauhan, S., Saini, S., Bathla, R., & Rana, A. (2020). Application of Machine Learning to Predict Hospital Churning. In *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)* (pp. 33-37). IEEE. https://drajayrana.com/wp-content/uploads/2020/11/paper-19.pdf
13. De, S., & Prabu, P. (2022). A sampling-based stack framework for imbalanced learning in churn prediction. *IEEE Access*, *10*, 68017-68028. https://ieeexplore.ieee.org/abstract/document/9803037/
14. Ehsani, F., & Hosseini, M. (2024). Customer churn prediction using a novel meta-classifier: an investigation on transaction, Telecommunication and customer churn datasets. *Journal of Combinatorial Optimization*, *48*(1), 7. https://doi.org/10.1007/s10878-024-01196-w
15. Flaks-Manov, N., Srulovici, E., Yahalom, R., Perry-Mezre, H., Balicer, R., & Shadmi, E. (2020). Preventing hospital readmissions: healthcare providers' perspectives on "impactibility" beyond EHR 30-day readmission risk prediction. *Journal of General Internal Medicine*, *35*, 1484-1489. https://doi.org/10.1007/s11606-020-05739-9
16. Groden, P., Capellini, A., Levine, E., Wajnberg, A., Duenas, M., Sow, S., ... & Kishore, S. (2021). The success of behavioral economics in improving patient retention within an intensive primary care practice. *BMC Family Practice*, *22*(1), 253. https://doi.org/10.1186/s12875-021-01593-8
17. Haimson, C. E., Simes, J. T., Eason, J. M., & Zhang, J. (2023). The impact of carceral churn and healthcare organizations on HIV/AIDS incidence in Arkansas. *SSM-Population Health*, *21*, 101355. https://doi.org/10.1016/j.ssmph.2023.101355
18. Huang, Y., Talwar, A., Chatterjee, S., & Aparasu, R. R. (2021). Application of machine learning in predicting hospital readmissions: a scoping review of the literature. *BMC medical research methodology*, *21*, 1-14. https://doi.org/10.1186/s12874-021-01284-z
19. Jajam, N., Challa, N. P., Prasanna, K. S., & Deepthi, C. V. S. (2023). Arithmetic optimization with ensemble deep learning SBLSTM-RNN-IGSA model for customer churn prediction. *Ieee Access*, *11*, 93111-93128. https://ieeexplore.ieee.org/iel7/6287639/10005208/10216284.pdf
20. Jamjoom, A. A. (2021). The use of knowledge extraction in predicting customer churn in B2B. *Journal of Big Data*, *8*(1), 110. https://doi.org/10.1186/s40537-021-00500-3
21. Jaswal, N., Goel, S., Upadhyay, K., Pathni, A. K., Bera, O. P., & Shah, V. (2024). Factors affecting patient retention to hypertension treatment in a North Indian State: A mixed-method study. *The Journal of Clinical Hypertension*, *26*(9), 1073-1081. https://doi.org/10.1111/jch.14866
22. Joy, U. G., Hoque, K. E., Uddin, M. N., Chowdhury, L., & Park, S. B. (2024). A Big Data-Driven Hybrid Model for Enhancing Streaming Service Customer Retention through Churn Prediction Integrated with Explainable AI. *IEEE Access*. https://ieeexplore.ieee.org/abstract/document/10530632/
23. Luo, A. L., Ravi, A., Arvisais-Anhalt, S., Muniyappa, A. N., Liu, X., & Wang, S. (2023). Development and Internal Validation of an Interpretable Machine Learning Model to Predict Readmissions in a United States Healthcare System. In *Informatics*, *10*(2), pp. 33. https://doi.org/10.3390/informatics10020033
24. Makene, F. S., Ngilangwa, R., Santos, C., Cross, C., Ngoma, T., Mujinja, P. G., ... & Mackintosh, M. (2022). Patients' pathways to cancer care in Tanzania: documenting and addressing social inequalities in reaching a cancer diagnosis. *BMC Health Services Research*, *22*(1), 189. https://doi.org/10.1186/s12913-021-07438-5
25. Namakoola, I., Moyo, F., Birungi, J., Kivuyo, S., Karoli, P., Mfinanga, S., ... & Garrib, A. (2024). Long-term impact of an integrated HIV/non-communicable disease care intervention on patient retention in care and clinical outcomes in East Africa. *Tropical Medicine & International Health*, *29*(8), 723-730. https://onlinelibrary.wiley.com/doi/pdf/10.1111/tmi.14026

26. Panchangam, P. V., BU, T., & Maniaci, M. J. (2024). Machine Learning-Based Prediction of Readmission Risk in Cardiovascular and Cerebrovascular Conditions Using Patient EMR Data. In *Healthcare, 12*(15), pp. 1497. https://doi.org/10.3390/healthcare12151497
27. Pejić Bach, M., Pivar, J., & Jaković, B. (2021). Churn management in telecommunications: Hybrid approach using cluster analysis and decision trees. *Journal of Risk and Financial Management*, *14*(11), 544. https://doi.org/10.3390/jrfm14110544
28. Pinheiro, P., & Cavique, L. (2022). Telco customer churn analysis: Measuring the effect of different contracts. In *World Conference on Information Systems and Technologies* (pp. 112-121). Cham: Springer International Publishing. https://www.researchgate.net/publication/360627430_Telco_Customer_Churn_Analysis_Measuring_the_Effect_of_Different_Contracts?enrichId=rgreq-f598e29e9e22dd5e3fe87c5782a6a3e5-XXX&enrichSource=Y292ZXJQYWdlOzM2MDYyNzQzMDtBUzoxMTQzMTI4MTI5NTg4MDQ4NkAxNzMzNzc2OTQ5Njg4&el=1_x_3&_esc=publicationCoverPdf
29. Ruiz de San Martín, R., Morales-Hernández, C., Barberá, C., Martínez-Cortés, C., Banegas-Luna, A. J., Segura-Méndez, F. J., ... & Hernández-Morante, J. J. (2024). Global and Local Interpretable Machine Learning Allow Early Prediction of Unscheduled Hospital Readmission. *Machine Learning and Knowledge Extraction*, *6*(3), 1653-1666. https://doi.org/10.3390/make6030080
30. Saha, S., Saha, C., Haque, M. M., Alam, M. G. R., & Talukder, A. (2024). ChurnNet: Deep Learning Enhanced Customer Churn Prediction in Telecommunication Industry. *IEEE Access*. https://ieeexplore.ieee.org/abstract/document/10380579/
31. Sengupta, A., Mukherjee, A., & VanderMeer, D. (2024). Impact of Perceived Barriers of Electronic Health Information Exchange on Physician's Use of EHR: A Normalisation Process Theory Approach. *Information Systems Frontiers*, 1-15. https://doi.org/10.1007/s10796-024-10524-8
32. Teo, K., Yong, C. W., Chuah, J. H., Hum, Y. C., Tee, Y. K., Xia, K., & Lai, K. W. (2023). Current trends in readmission prediction: an overview of approaches. *Arabian journal for science and engineering*, *48*(8), 11117-11134.https://doi.org/10.1007/s13369-021-060405
33. Thakkar, H. K., Desai, A., Ghosh, S., Singh, P., & Sharma, G. (2022). Clairvoyant: AdaBoost with Cost-Enabled Cost-Sensitive Classifier for Customer Churn Prediction. *Computational intelligence and neuroscience*, *2022*(1), 9028580. https://doi.org/10.1155/2022/9028580