2024 27th International Conference on Computer and Information Technology (ICCIT)
20-22 December 2024, Cox's Bazar, Bangladesh

Synergizing GAN and Customized Neural Network for Enhanced Brain Stroke Prediction: A Web-based Implementation

Md. Musab Us Saber Shakin[1], Md. Minhazul Islam[2], Md. Rakib Hossain[3] S. M. Mahedy Hasan[4], Farzana Akter[5], Shykul Islam Siam[6], Azmain Yakin Srizon[7], Md. Faruk Hossain[8]
[1,2,3,5,6]*Department of Electronics & Telecommunication Engineering*
[4,7]*Department of Computer Science & Engineering*
[8]*Department of Electrical & Electronic Engineering*
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
Emails: md.shakin18@gmail.com, miinhaz14@gmail.com, rakib.ete12.ruet@gmail.com, mahedy@cse.ruet.ac.bd, farzanaeteruet@gmail.com, shykulislam120@gmail.com, azmainsrizon@gmail.com, engr.mfhossain@eee.ruet.ac.bd

***Abstract*—Brain strokes are a global health concern, demanding effective predictive models to mitigate risks and enhance preventive interventions. Timely identification of individuals at risk is pivotal, and machine learning emerges as a promising avenue for accurate stroke prediction, potentially transforming patient care. This study introduces a comprehensive approach utilizing machine learning, addressing challenges like missing data and class imbalance through Generative Adversarial Network (GAN) for data synthesis and balancing. The custom Artificial Neural Network (ANN) tailored to brain stroke prediction achieves outstanding performance with a 99.9% accuracy. Going beyond theoretical boundaries, the research translates into practical implementation through a web-based platform for real-world stroke prediction. This platform seamlessly integrates the high-performing model, providing a user-friendly interface for individuals to assess their stroke risks. Notably, the custom ANN achieves a remarkable 95.11% accuracy even without accounting for missing data or data imbalance, underscoring its robustness in practical healthcare applications. In amalgamating advanced machine learning methodologies with practical solutions, this work significantly contributes to medical research and offers a tangible preemptive healthcare solution.**

***Index Terms*—Brain Stroke, Generative Adversarial Network, Artificial Neural Network, Web Implementation**

I. Introduction

A stroke, resulting from a sudden disruption of blood flow to the brain, inflicts devastating consequences, whether caused by a clot or a bleed. The global impact is alarming, with an estimated 15 million people experiencing a stroke annually, leading to 5 million deaths and leaving another 5 million with lasting disabilities [1]. In the United States alone, strokes claim over 140,000 lives yearly, accounting for one in six cardiovascular deaths, occurring every 40 seconds, with a life lost every 3 minutes [2]. Urgent preventative measures are crucial, given that high blood pressure, smoking, and unhealthy lifestyles contribute to the stroke epidemic. Modern sedentary lifestyles and poor dietary choices exacerbate these risks, especially among younger populations [3]. Recognizing the issue's magnitude is vital for mitigation. Early detection, prompt medical intervention, and prioritizing stroke prevention through healthy lifestyles can lead to a future with fewer lives lost and fewer enduring the indelible mark of this devastating event.

The conventional approach to handling strokes often leans towards reacting to incidents rather than proactively preventing them, facing significant obstacles. This reactive tendency, compounded by fragmented healthcare systems and a lack of widespread public awareness, leads to delays in recognizing and treating strokes, hampering the achievement of optimal outcomes [4]. A fundamental paradigm shift towards a holistic and proactive model is imperative to effectively tackle the global burden of stroke. Moreover, the time-consuming nature of analyzing results with traditional methods for determining the occurrence of a previous stroke further underscores the need for a more efficient and timely approach.

In the realm of addressing strokes, the integration of machine learning has emerged as a promising avenue, offering innovative solutions to enhance diagnosis, prognosis, and treatment planning. Researchers are leveraging advanced algorithms to analyze large datasets, including medical imaging and patient records, enabling more accurate and timely identification of stroke risk factors and early warning signs. Machine learning models contribute to personalized medicine by tailoring interventions based on individual patient profiles, optimizing outcomes and resource utilization [5], [6]. As scientists continue to pioneer breakthroughs in artificial intelligence applications for stroke care, the collaborative efforts between technology and healthcare professionals signify a concerted push toward more effective prevention, timely intervention, and improved post-stroke support.

In this study, machine learning techniques were applied to assess the risk of brain stroke, aiming to overcome limitations associated with traditional methods. The implementation of advanced algorithm seeks to provide a faster and more reliable means of predicting an individual's susceptibility to a stroke. By leveraging a good dataset and sophisticated analytics, the

979-8-3315-1909-4/24/$31.00 ©2024 IEEE

model introduced in this research offers an efficient and proactive approach, addressing the shortcomings of conventional methods. The major contributions from this study involve a set of crucial improvements implemented to surmount challenges and enhance the credibility of the approach:

- Addressing data imbalance challenges, a Generative Adversarial Network (GAN) approach has been employed to generate synthetic data. This augmentation significantly contributes to the model's performance and accuracy in predicting stroke risk by creating a more balanced and representative dataset.
- A tailored Artificial Neural Network (ANN) has been developed to optimize accuracy and results. The custom ANN is designed to effectively capture complex patterns within the data relevant to stroke risk, leading to improved predictive capabilities and aligning closely with the intricacies of the problem.
- A user-friendly interactive website has been created to streamline the prediction process. Patients can input relevant data, and the model processes the information swiftly to predict stroke risk. This platform not only enhances user accessibility but also empowers individuals to take proactive measures based on timely and accurate results.

Section II delves into a discussion of related works in the field. Section III presents the materials and methods employed in this study. Section IV provides a thorough exploration of the experimental results, featuring a detailed comparison between the performance of the proposed approach and other cutting-edge research in the field. In Section V, addressing the study's limitations. Finally, in Section VI, conclusions are drawn, and the paper is wrapped up.

II. Related Works

Recent studies in brain stroke prediction have focused on diverse machine learning techniques, emphasizing the importance of dataset balancing, deep learning, ensemble techniques for improved precision and accuracy.

Dritsas et al. [7], introduced binary classification ML models, excelling in stacking classification with an AUC of 98.9% and an accuracy of 98%. However, the study highlighted a limitation associated with relying on publicly available datasets, emphasizing the need for alternative balancing methods.

Rehman et al. [8] presented a robust tuning ensemble RXLM, combining Random Forest, XGBoost and LightGBM models which achieved an accuracy of 95.29% with a conscientious acknowledgment of potential optimistic estimates from oversampling.

In another study by Ivanov et al. [9], Support Vector Machines achieved a commendable 98% accuracy, emphasizing data quality and preprocessing. However, limitations include reliance on a single dataset and exclusion of certain variables.

Emon et al. [10] introduced the weighted voting classifier resulting in 97% accuracy but compared to recent works the method is not that advanced and also the study lacks clear explanation of the model.

Using Microsoft Azure Machine Learning, Ray et al. [11] achieved 97.6% accuracy with a boosted decision tree and the Chi-Squared test for feature selection. However, the study did not focus on all the features and also used oversampling.

Bathla et al. [12] used machine learning models and achieved accuracy of 95.23% with Random Forrest Classifier but the study delved into limitations associated with varied classifier performance and dataset imbalance.

Rahman et al. [13] achieved high accuracy in stroke prediction using ML algorithms and deep neural networks. Random Forest outperformed with 99% accuracy, followed by a 4-layer ANN (92.39%) and a 3-layer ANN (84.01%), showcasing diverse optimization considerations.

III. Research Methodology

This research proposes a methodology for predicting brain stroke using machine learning models, particularly focusing on a novel Artificial Neural Network (ANN) approach. The workflow begins with data acquisition and preprocessing, including label encoding, feature scaling, missing value handling, and dataset balancing with Generative Adversarial Network (GAN). Subsequently, the data is split into 70% for training and 30% for testing. Various machine learning models, including the proposed ANN, are trained and evaluated for their classification performance in predicting brain stroke. Finally, a web interface is implemented to offer a user-friendly platform for utilizing the model and predicting individual stroke risk. This approach is visually summarized in Figure 1 for a quick overview.

A. Dataset Description

The research utilized an openly accessible dataset obtained from Kaggle [14]. Within this dataset, there were a total of 12 columns and 5110 rows. The ”stroke” column exhibited 249 rows with a value of 1, indicative of instances where a stroke occurred and 4861 rows featured a value of 0, denoting the absence of a stroke. More of the attributes are delineated as follows:

- **Age (years):** Describing the age of participants aged 18 and above.
- **Gender:** Categorizing participants by gender, with 1260 men and 1994 women.
- **Hypertension:** Indicating the hypertensive status of participants, with a prevalence of 12.54
- **Heart disease:** Identifying participants with heart disease, with an incidence rate of 6.33
- **Ever married:** Reflecting the marital status, with 79.84% of participants being married.
- **Work type:** Classifying participants' employment status into four categories (private 65.02%, self-employed 19.21%, govt-job 15.67%, and never-worked 0.1%).
- **Residence type:** Categorizing participants' living arrangements into urban (51.14%) and rural (48.86%) areas.
- **Avg glucose level (mg/dL):** Quantifying the average glucose level of participants.

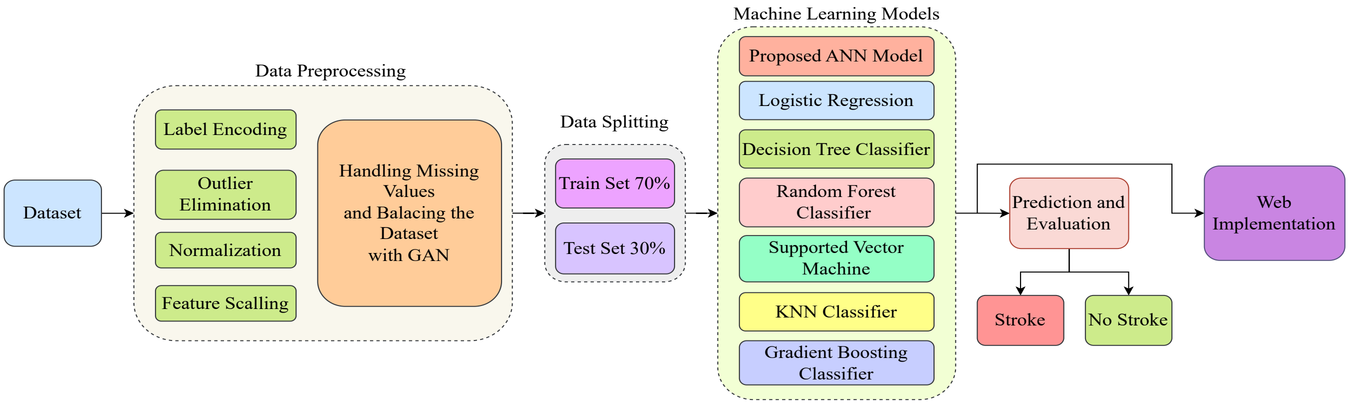

Fig. 1: Workflow of the Proposed Methodology

- **BMI (Kg/m2):** Expressing the body mass index of participants.
- **Smoking Status:** Differentiating participants' smoking habits across three categories (smoke 22.37%, never smoked 52.64%, and formerly smoked 24.99%).
- **Stroke:** Determining if participants had experienced a stroke previously, with an incidence rate of 5.53%.

B. Dataset Preprocessing

Preprocessing the data is an important part as it leads to better model performance. Selecting the features which have impact on the actual stroke cases is also very crucial. So understanding how features are connected to one another is made easier through feature selection. The target feature exhibits positive correlations with several key attributes including Gender, Age, Hypertension, Average Glucose Level, Heart Disease, Ever Married, BMI and Job types. The column for the 'bmi' had no values in 201 rows. Instead of getting rid of these rows synthetic data generation was used with the proposed GAN method for generating bmi values for these values based on the entire dataset. Another issue to be noticed is that the dataset had 249 case of stroke and 4861 case of no stroke. So training the model on an imbalanced dataset like this will result in better prediction for one class only. Instead of using oversampling methods, GAN was used to balance the dataset by generating synthetic data which will be discussed in the next part.

C. Proposed GAN Model for Handling Missing Data and Imbalanced Dataset

This segment introduces an inventive approach that customizes a Generative Adversarial Network (GAN) model to simultaneously address two common challenges: missing data imputation and class imbalance within a dataset. The approach leverages the inherent capabilities of GANs to generate synthetic data, aiming to enhance dataset completeness and ensure a balanced representation of classes. GANs are chosen for their innate ability to capture intricate data distributions and generate synthetic instances that closely emulate authentic data. In datasets affected by missing entries and uneven class distribution, GANs serve as a potent solution by learning and replicating the underlying data structure.

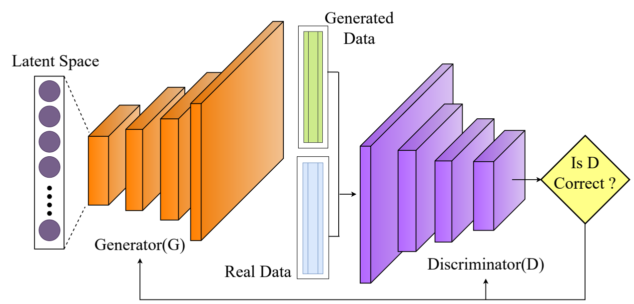

Fig. 2: Architecture of The Proposed GAN Model

This, in turn, facilitates the imputation of missing values and the creation of synthetic samples to rectify class imbalances. The GAN framework operates as a minimax game, involving a generator (G) and a discriminator (D). The optimization objective, expressed as a minimization-maximization problem, is encapsulated in the following equation [15]:

$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})}[\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\text{latent}}(\mathbf{z})}[\log(1 - D(G(\mathbf{z})))] \quad (1)$$

Here, $\mathbf{x}$ signifies real data, $\mathbf{z}$ represents latent space samples, $p_{\text{data}}(\mathbf{x})$ denotes the true data distribution, and $p_{\text{latent}}(\mathbf{z})$ signifies the distribution of latent space samples. The GAN model encompasses a generator (G) and a discriminator (D). The generator transforms random noise from a latent space into synthetic samples, while the discriminator distinguishes between real and synthetic instances. The interplay involves G minimizing the generator loss and D minimizing its own loss.

$$G^* = \arg\min_G \max_D V(D, G) \quad (2)$$

Fig. 2 depicts the architecture of the proposed GAN model. The model's generator accepts a 10-dimensional latent vector as input, navigating it through two dense layers with Rectified Linear Unit (ReLU) activation, ultimately generating synthetic samples through a linear activation output layer. In contrast, the discriminator intricately processes both authentic and synthetic samples by employing layers with Leaky ReLU activation, concluding with a sigmoid activation output layer designed for binary classification. The Generative Adversarial Network (GAN) harmoniously integrates these components, refining the generator's ability to produce authentic synthetic data while concurrently honing the discriminator's aptitude for distinguishing between genuine and synthetic instances. Notably, during training, the discriminator undergoes temporary freezing, facilitating a dynamic interplay that enhances the model's overall performance. This architectural design proves particularly effective in mitigating challenges such as missing data and class imbalance within datasets. Unlike simple oversampling, the proposed Generative Adversarial Network (GAN) intelligently generates diverse and realistic synthetic data to balance dataset classes. By dynamically learning the underlying distribution, the GAN produces novel samples, mitigating overfitting risks associated with direct replication. The GAN's training balance, achieved through temporary discriminator freezing, ensures the generation of high-quality synthetic data, making it a superior approach for addressing class imbalance compared to conventional oversampling methods.

D. Proposed Custom ANN Model

The advanced Artificial Neural Network (ANN) model, tailored for refining datasets following Generative Adversarial Network (GAN) processing, is characterized by a sophisticated architecture aimed at optimizing pattern recognition, minimizing overfitting, and bolstering generalization performance. This model adopts a stacked structure featuring six dense layers, with the final layer utilizing a sigmoid activation function for binary classification. Ranging from 16 to 256 neurons, these dense layers operate as effective feature extractors, systematically learning intricate hierarchical representations within the data. Notably, this last dense layer is composed of a single neuron.

To enhance training stability and prevent overfitting, each dense layer is interspersed with batch normalization and dropout layers. Batch normalization ensures consistent training dynamics by normalizing inputs between layers, while dropout layers randomly deactivate 50% of neurons during training, thus mitigating overfitting risks. Additionally, the model incorporates L2 regularization with a coefficient of 0.01, effectively regulating complexity and further fortifying against overfitting. This amalgamation of techniques not only adeptly captures nuanced relationships within the dataset but also excels in generalization, outperforming traditional machine learning methods. Such intentional inclusion of regularization mechanisms becomes crucial, particularly when confronted with datasets rich in intricate patterns and potential noise, ensuring the model's robust performance on previously unseen data.

E. Experimental Setup

The entire experiment was conducted in Python-3 on Kaggle Notebooks, leveraging hardware specifications of 13GB of RAM and a 16GB P100 GPU. The dataset was partitioned into training (70%) and testing (30%) sets for model development. During the training of the Generative Adversarial Network (GAN) aimed at generating synthetic data, a learning rate of 0.0002 was employed for both the generator and discriminator, with the adam optimizer and a batch size of 32. This specialized approach was integrated to enhance the model's capabilities in generating synthetic data.

In addition to conventional GAN training, a custom artificial neural network (ANN) was trained using the adam optimizer with a learning rate of 0.0001 and binary crossentropy as the loss function. To comprehensively evaluate the proposed model, various statistical measures, including accuracy, precision, recall, F1-score, ROC, and AUC were considered. This multifaceted evaluation strategy ensures a robust assessment of the model's performance, taking into account both GAN and custom ANN training components.

IV. Results and Discussion

A. Performance Analysis and Comparison

The result analysis section commences by displaying the performance of the proposed model before balancing the dataset. This is followed by the proposed GAN model training for balancing the dataset. Next, performance of the model is evaluated on the balanced dataset. In comparison to several machine learning models were also examined in the study. To conclude this section, a performance assessment is performed by comparing the proposed model with the methods used in previous studies within the field.

TABLE I: Performance comparison of the proposed model with other models on imbalanced dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
LR	94.60	89.50	94.60	91.98	85.50
DR	90.43	91.07	90.43	90.74	56.69
RF	94.60	89.50	94.60	91.98	80.19
SVM	94.60	89.50	94.60	91.98	67.55
KNN	94.60	92.28	94.60	92.17	65.25
NB	37.47	95.03	37.47	48.70	82.27
XGBoost	94.50	91.82	94.50	92.29	76.97
Proposed ANN	**95.11**	**90.46**	**95.11**	**92.73**	**84.19**

Table I shows the performance of the proposed ANN model on the imbalanced dataset with the missing data. Compared to other machine learning techniques the proposed custom ANN has the best performance.

In Fig. 3, the generator and discriminator loss of the proposed GAN model for addressing missing data and balancing the dataset are illustrated. The model underwent training for 200 epochs on the dataset. Notably, both losses exhibit a decreasing trend over the epochs. The discriminator loss is

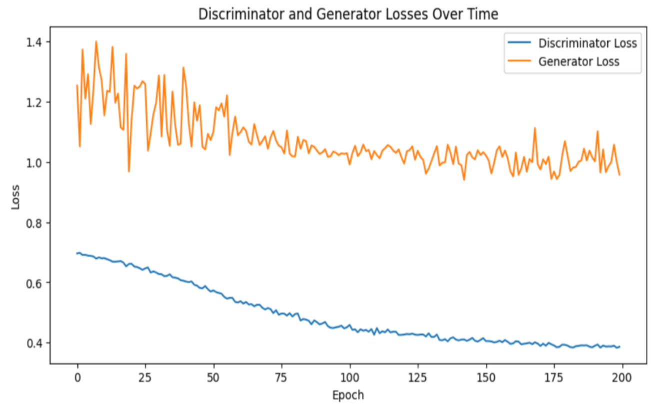

Fig. 3: Generator and Discriminator loss curve of the proposed GAN model.

consistently low, suggesting the model's adeptness in distinguishing between real and generated data. Consequently, the decreasing generator loss signifies the model's improvement in generating more realistic data over time, thereby challenging the discriminator. This dynamic interaction results in the stabilization of the discriminator loss after a certain period.

After handling the missing data and balancing the dataset the proposed model significantly improved it's performance resulting in 99.99% accuracy which can be seen in Table II.

TABLE II: Performance comparison of the proposed ANN with Imbalanced and Balanced Dataset

Model	**Accuracy** (%)	**Precision** (%)	**Recall** (%)	**F1-Score** (%)	**AUC** (%)
Proposed ANN with Imbalanced Data	95.11	90.46	95.11	92.73	84.19
Proposed ANN with Balanced Data	**99.99**	**100**	**99.99**	**100**	**99.99**

Receiver Operating Characteristic (ROC) curves were employed to visually assess the proposed model's ability to differentiate between positive and negative classes. Figure 4 illustrates the ROC curve for the model, both before and after dataset balancing, showcasing excellent performance in terms of accuracy and ROC score. The AUC (Area Under the Curve) for the imbalanced data is 0.80, while for the balanced data, it reaches a perfect score of 1.00. This notable difference implies that the model's accuracy significantly improves when operating on a balanced dataset, emphasizing the importance of data balancing for enhanced performance.

Finally, Table III provides a comprehensive performance comparison with state-of-the-art research, unequivocally demonstrating that the proposed model surpasses its counterparts. The results underscore the superior efficacy of the developed model in the realm of predicting brain strokes, showcasing its advancements over existing methodologies. This comparative analysis serves to highlight the notable contributions of the proposed approach, reinforcing its position as a leading solution in the field.

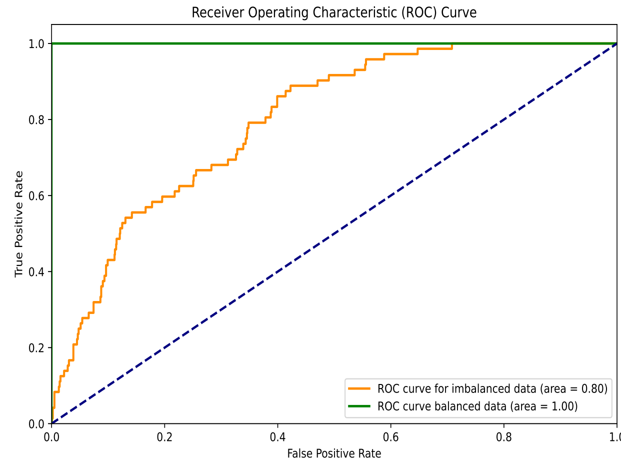

Fig. 4: ROC curve on the imbalanced vs balanced dataset for the proposed model

TABLE III: Performance Comparison with State of the Art Works

State of the Art Works	**Accuracy** (%)	**Precision** (%)	**Recall** (%)	**F1-Score** (%)	**AUC**
[7]	98	97.4	97.4	97.4	98.9
[8]	95.29	94.36	96.38	95.35	99.18
[9]	98.76	95.50	94.90	-	-
[10]	97	93	100	97	93
[11]	97.6	-	-	-	-
[12]	98	-	-	-	-
[13]	99	100	98	99	100
Proposed Model	**99.99**	**100**	**99.99**	**100**	**99.99**

B. Web Implementation

In order to bring the proposed model to real-world application, a dedicated website [16] was meticulously crafted and implemented. Python Flask [17] served as the backbone framework for this innovative platform, seamlessly integrating the trained model into the website's robust back-end. Leveraging HTML, CSS and JavaScript a visually appealing and user-friendly interface is made which allows users to input data similar to the dataset, which is then converted into the appropriate format for feeding into the model. The output is presented as a percentage, indicating the likelihood of an individual experiencing a brain stroke in the future based on the input data. To ensure widespread accessibility, the website is hosted using Render [18], making it available to users across diverse locations. This comprehensive strategy not only exemplifies the tangible implementation of the model but also underscores the convergence of cutting-edge technology and a user-centric design ethos to confront and solve real-world challenges. Fig. 5 shows the website where two case can be seen for stroke and no stroke for user data input.

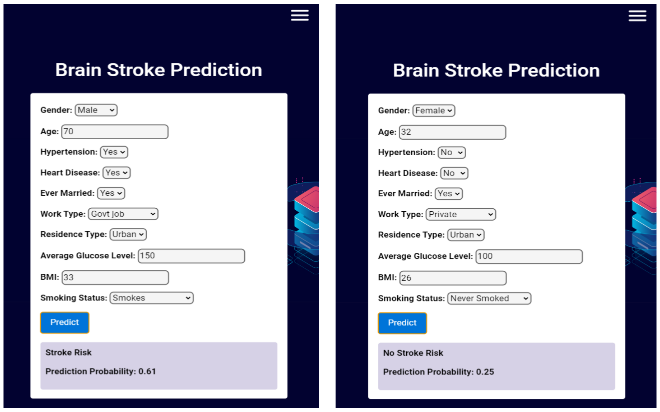

Fig. 5: Website implementation of the proposed model

V. Threats to Validity

In this study, a balanced dataset was used to develop the proposed model, trained on specific data. The findings may be limited when applied to different data types like cholesterol levels, genetic history, alcohol consumption, obesity, and drug abuse. A customized ANN model was used, but there is room for improvement. GAN was employed to balance the dataset, equalizing stroke and non-stroke cases, generating realistic data. The GAN model's simplicity allows future enhancements. Future research aims to explore advanced GAN models for improved synthetic data generation and utilize larger, more diverse datasets to expand applicability and accuracy.

VI. Conclusion

This study introduces an innovative solution to address the critical challenge of predicting Brain Strokes, emphasizing the necessity for early intervention in the medical domain. In recognizing prevailing limitations, the research makes a significant contribution by incorporating Generative Adversarial Network (GAN)-based data handling and the development of a deep learning model. The implementation of this model on a website further extends its applicability to real-world scenarios.

The research methodology centers on the utilization of synthetic data generated through GAN for data preprocessing, providing a more realistic alternative compared to conventional oversampling techniques. Additionally, a customized Artificial Neural Network (ANN) tailored to the dataset exhibits exceptional performance. Rigorous empirical evaluations on the dataset underscore the credibility of the proposed approach, revealing an impressive accuracy rate of 99.99%. This achievement surpasses the performance of contemporary state-of-the-art methods, emphasizing the methodology's efficacy and competitiveness in the realm of brain stroke prediction.

Additionally, it is noteworthy that the customized Artificial Neural Network (ANN) showcases an impressive accuracy of 95.11%, even in the absence of considerations for missing data or data imbalance. This underscores its resilience and effectiveness in practical healthcare applications. By effectively tackling a critical medical challenge, this research not only contributes to addressing pressing issues but also lays a solid foundation for the progression of predictive modeling within the healthcare domain, making a substantial impact on the field.

Acknowledgement

This research was funded by University Grants Communication (UGC), Bangladesh and Office of the Director of Research and Extension (DRE), Rajshahi University of Engineering & Technology, under project No. (2024-25/28). The authors, therefore, gratefully acknowledge UGC and DRE technical and financial support.

References

[1] World Health Organization, https://www.emro.who.int/health-topics/stroke-cerebrovascular-accident/index.html, accessed on December, 2023.
[2] Centers for Disease Control and Prevention, https://www.cdc.gov/stroke/facts.htm, accessed on December, 2023.
[3] American Stroke Association, https://www.yalemedicine.org/news/high-blood-pressure-hypertension, accessed on December, 2023.
[4] V. L. Feigin, G. A. Roth, M. Naghavi, P. Parmar, R. Krishnamurthi, S. Chugh, G. A. Mensah, B. Norrving, I. Shiue, M. Ng *et al.*, "Global burden of stroke and risk factors in 188 countries, during 1990–2013: a systematic analysis for the global burden of disease study 2013," *The Lancet Neurology*, vol. 15, no. 9, pp. 913–924, 2016.
[5] E. Dritsas and M. Trigka, "Stroke risk prediction with machine learning techniques," *Sensors*, vol. 22, no. 13, p. 4670, 2022.
[6] M. Alruily, S. A. El-Ghany, A. M. Mostafa, M. Ezz, and A. A. El-Aziz, "A-tuning ensemble machine learning technique for cerebral stroke prediction," *Applied Sciences*, vol. 13, no. 8, p. 5047, 2023.
[7] E. Dritsas and M. Trigka, "Stroke risk prediction with machine learning techniques," *Sensors*, vol. 22, no. 13, p. 4670, 2022.
[8] A. Rehman, T. Alam, M. Mujahid, F. S. Alamri, B. Al Ghofaily, and T. Saba, "Rdet stacking classifier: a novel machine learning based approach for stroke prediction using imbalance data," *PeerJ Computer Science*, vol. 9, 2023.
[9] I. G. Ivanov, Y. Kumchev, and V. J. Hooper, "An optimization precise model of stroke data to improve stroke prediction," *Algorithms*, vol. 16, no. 9, p. 417, 2023.
[10] M. U. Emon, M. S. Keya, T. I. Meghla, M. M. Rahman, M. S. Al Mamun, and M. S. Kaiser, "Performance analysis of machine learning approaches in stroke prediction," in *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 2020, pp. 1464–1469.
[11] S. Ray, K. Alshouiliy, A. Roy, A. AlGhamdi, and D. P. Agrawal, "Chi-squared based feature selection for stroke prediction using azureml," in *2020 Intermountain Engineering, Technology and Computing (IETC)*. IEEE, 2020, pp. 1–6.
[12] P. Bathla and R. Kumar, "Artificial intelligence based model for brain stroke prediction," in *2022 IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, 2022, pp. 1–6.
[13] S. Rahman, M. Hasan, and A. K. Sarkar, "Prediction of brain stroke using machine learning algorithms and deep neural network techniques," *European Journal of Electrical Engineering and Computer Science*, vol. 7, no. 1, pp. 23–30, 2023.
[14] "Stroke prediction dataset," https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset, accessed: January 2024.
[15] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems (NIPS)*, 2014. [Online]. Available: https://arxiv.org/abs/1406.2661
[16] "Stroke prediction website," https://brain-stroke-pred.onrender.com/, accessed: January 2024.
[17] *Flask*, accessed: January 2024. [Online]. Available: https://flask.palletsprojects.com/en/3.0.x/
[18] Render. Accessed: January 2024. [Online]. Available: https://render.com/