

# An Integrative Paradigm for Enhanced Stroke Prediction: Synergizing XGBoost and xDeepFM Algorithms

Weinan Dai  
wdai22@my.trine.edu  
Trine University  
Phoenix, USA

Chengjie Mou  
cmou22@my.trine.edu  
Trine University  
Phoenix, USA

Yifeng Jiang  
yjiang8@bu.edu  
Boston University  
Boston, USA

Chongyu Zhang  
chongyu.zhang@tum.de  
Technical University of Munich  
Munich, Germany

## ABSTRACT

Stroke prediction plays a crucial role in preventing and managing this debilitating condition. In this study, we address the challenge of stroke prediction using a comprehensive dataset, and propose an ensemble model that combines the power of XGBoost and xDeepFM algorithms. Our work aims to improve upon existing stroke prediction models by achieving higher accuracy and robustness. Through rigorous experimentation, we validate the effectiveness of our ensemble model using the AUC metric. Through comparing our findings with those of other models in the field, we gain valuable insights into the merits and drawbacks of various approaches. This, in turn, contributes significantly to the progress of machine learning and deep learning techniques specifically in the domain of stroke prediction.

## CCS CONCEPTS

• **Computing methodologies** → **Ensemble methods; Neural networks; Classification and regression trees.**

## KEYWORDS

Stroke prediction, ensemble model, XGBoost, xDeepFM

## 1 INTRODUCTION

Stroke is a severe medical condition caused by the abrupt disruption of blood flow to the brain. This interruption leads to significant neurological impairments and can result in long-term disability or even death. Stroke, a leading cause of morbidity and mortality globally, places a substantial burden on healthcare systems and society at large. Early detection and accurate prediction of stroke risk are crucial for implementing preventive measures, initiating timely interventions, and improving patient outcomes.

In recent years, machine learning techniques have garnered significant attention due to their potential in enhancing stroke prediction. These approaches leverage large-scale datasets and sophisticated algorithms to identify patterns and risk factors that contribute to the occurrence of strokes. By analyzing various demographic, behavioral, and medical features, these models aim to provide reliable and timely predictions of stroke risk, enabling healthcare professionals to take proactive measures and allocate resources effectively.

In this study, we address the challenge of stroke prediction using a comprehensive dataset obtained from the Kaggle Stroke Prediction Dataset<sup>1</sup>. Our objective is to develop a model that surpasses existing approaches in terms of accuracy and robustness. To achieve this, we propose an ensemble approach that combines the strengths of two powerful algorithms: **XGBoost** and **xDeepFM**.

XGBoost is a popular gradient boosting algorithm known for its excellent performance in structured data analysis. It effectively captures complex relationships between features and provides robust predictions. xDeepFM, on the other hand, is a deep learning model that excels in modeling feature interactions in sparse datasets, which is particularly relevant in the context of stroke prediction where certain risk factors may be infrequent but highly informative.

By leveraging the complementary strengths of XGBoost and xDeepFM, our ensemble model aims to achieve enhanced predictive accuracy and a more comprehensive understanding of stroke risk factors. Through rigorous experimentation and evaluation, we validate the effectiveness of our model using the widely accepted evaluation metric, AUC. The AUC metric quantifies the model's capacity to differentiate between individuals who will experience a stroke and those who will not, providing a comprehensive evaluation of its predictive performance.

## 2 RELATED WORK

Stroke prediction using machine learning techniques has gained significant attention in the healthcare domain in recent years. Several studies have been conducted to explore the effectiveness of different approaches for stroke prediction. In this section, we provide a comprehensive review and analysis of relevant literature pertaining to stroke prediction using machine learning methods.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
ICBDT 2023, September 22–24, 2023, Qingdao, China

© 2018 Association for Computing Machinery.  
ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00  
<https://doi.org/XXXXXXXXXXXXXX>

<sup>1</sup><https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>

Khosla et al. [1] propose an integrated model that combines various machine learning techniques to improve prediction accuracy. The study demonstrates improved performance compared to individual models. Two other studies also focused on integrated models from different perspectives. BotShape [2], a social bots detection system, demonstrates the high effectiveness of ensemble tree-based models to improve the accuracy compared with traditional classifiers. Wu et al. [3] created a system integrating an ensemble classifier to combine the information from different aspects of features for boosting accuracy. Emon et al. [4] conduct a comprehensive analysis of different machine learning algorithms specifically in the context of stroke prediction. The study evaluates their performance, highlighting decision trees as effective models. However, scalability may be a concern when dealing with large datasets. Liu et al. [5] tackle the issue of imbalanced data in stroke prediction, presenting a hybrid approach to address this challenge. The proposed hybrid approach shows promise in handling imbalanced datasets, although overfitting remains a potential issue. Wu et al. [6] proposed BotTriNet, a unified embedding framework by leveraging metric learning to refine raw embeddings, effectively tackling the great challenge posed by imbalanced datasets. Another medical research work, FineEHR [7], a novel clinical note embedding refining framework based on metric learning, successfully broke the performance bottleneck in an imbalanced medical data set. Wu's research demonstrates that metric learning can be widely efficient in solving data-imbalanced problems in various domains. Y Wu et al. [8] specifically focuses on stroke prediction among older Chinese individuals, showcasing the effectiveness of machine learning methods within this demographic group. However, further investigation is needed to assess generalizability to other populations. Weinan Dai et al. [9] employed the Twins-PCPVT, a vision transformer, for DR detection, underscoring the potential of deep learning in this domain. Their approach proved more efficient than relying solely on human expertise.

M Rajora et al. [10] propose a machine learning framework for stroke prediction in a distributed environment, addressing challenges related to data privacy and communication overhead. Their work lays the foundation for practical implementation. Daghistani et al. [11] contrasted statistical Logistic Regression with the RandomForest machine learning approach in predicting diabetes. The results underscored the nuances of different predictive methodologies in healthcare analytics. MS Sirsat et al. [12] provide a comprehensive overview of machine learning techniques used for stroke prediction, discussing various approaches and providing insights into their strengths and limitations. However, a more in-depth analysis of specific algorithms would enhance the comprehensiveness of the review. M Goyal et al. [13] concentrate on using deep learning models for stroke prediction, demonstrating high prediction accuracy. However, reliance on large labeled datasets may limit practical implementation. Y Xie et al. [14] explore the use of deep neural networks for stroke prediction based on electrocardiograms, showing promising results and suggesting the potential of electrocardiogram data in stroke prediction. YA Choi et al. [15] propose a deep learning-based system for stroke disease prediction using real-time biosignals, demonstrating the potential of real-time biosignals in early stroke detection, but further validation is required. Ye et al. [16] introduced MedLens, an intelligent mortality prediction

system based on electronic health records. It implements a vital medical signs selection approach for significant feature mining from medical big data and designs a robust interpolation algorithm to solve a universal challenge of missing data. IR Chavva et al. [17] investigate the applications of deep learning in acute stroke management, highlighting the potential of deep learning techniques and emphasizing the need for additional clinical validation. R Feng et al. [18] provide an in-depth review of clinical applications of deep learning in stroke management, highlighting the potential of deep learning techniques in various stages of stroke management, including diagnosis, treatment, and rehabilitation. N Stier et al. [19] explore deep learning approaches for extracting tissue fate features in acute ischemic stroke.

The existing literature on stroke prediction using machine learning methods has demonstrated promising outcomes in terms of predictive accuracy and the potential of deep learning models. However, certain challenges, including interpretability, computational requirements, and generalizability, still need to be addressed. Our proposed model addresses these challenges by utilizing an integrated machine learning approach that combines the strengths of multiple algorithms. This strategy allows us to achieve a balance between high prediction accuracy and factors such as interpretability, scalability, and practical implementation.

## 3 METHODOLOGY

In this section, we present an ensemble approach that combines the predictions of the XGBoost and xDeepFM models. Both models are trained separately on the same dataset, and their individual predictions are combined using ensemble coefficients determined through grid search. We also delve into the loss functions employed in XGBoost and xDeepFM, as well as the evaluation metric used to evaluate the performance of the ensemble model. We executed our experiments on a streamlined setup, highlighting our model's efficiency. The system was powered by an Intel<sup>®</sup> Core i5-7400 CPU and an NVIDIA GeForce GTX 1050 Ti GPU. This simple configuration emphasizes the adaptability and resource-friendliness of our model, ensuring its applicability across various scenarios without necessitating high-end hardware.

### 3.1 XGBoost

XGBoost [20] is an optimized implementation of the gradient boosting machine learning algorithm. It utilizes a boosting ensemble technique, where weak prediction models, typically decision trees, are combined to create a strong predictive model. XGBoost aims to minimize a loss function by iteratively adding new trees to the ensemble. The final prediction is derived by aggregating the predictions made by each individual tree in the model. XGBoost incorporates regularization terms to control overfitting and provides flexibility in defining custom loss functions.

In XGBoost, we employ both L1 and L2 regularization techniques to manage the complexity of the model and mitigate overfitting. L1 regularization, or Lasso regularization, incorporates the absolute values of the weights into the loss function, encouraging sparsity in the learned feature weights. The L1 regularization term is defined as:

$$\Omega^{L1}(f_1) = \lambda_1 \sum_{j=1}^J |w_j| \quad (1)$$

where  $\lambda_1$  is the regularization parameter,  $J$  is the number of leaves in the tree, and  $w_j$  is the weight associated with the  $j$ -th leaf.

In contrast, L2 regularization, also referred to as Ridge regularization, involves adding the squared values of the weights to the loss function. This regularization technique encourages the weights to be small and helps in preventing extreme values. The L2 regularization term is defined as:

$$\Omega^{L2}(f_2) = \lambda_2 \sum_{j=1}^J ||w_j||^2 \quad (2)$$

where  $\lambda_2$  is the regularization parameter,  $J$  is the number of leaves in the tree, and  $w_j$  is the weight associated with the  $j$ -th leaf.

By incorporating L1 and L2 regularization terms into the XGBoost loss function, the model can effectively control the complexity and prevent overfitting, leading to improved generalization performance.

Once the model is trained, we can obtain the feature importances, as depicted in Figure 1.

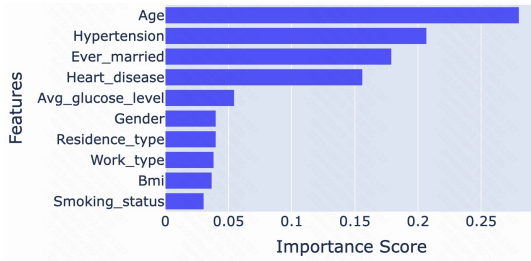


Figure 1: The Top 10 Features Importance of XGBoost

### 3.2 xDeepFM

xDeepFM [21] is a hybrid model that merges the advantages of factorization machines and deep neural networks. By employing factorization machines, it effectively captures high-order feature interactions, while utilizing deep neural networks to learn intricate feature representations. The model can handle both sparse and dense features as inputs and captures interactions among them.

The architecture of xDeepFM comprises several essential components, including an embedding layer, a stacking layer, a cross-network, and a deep network, as depicted in Figure 2.

The embedding layer maps categorical features into low-dimensional dense vectors, allowing the model to learn meaningful representations for categorical variables. Let's assume we have  $N$  categorical features and each feature has  $M$  distinct values. The embedding layer converts each categorical feature into a  $K$ -dimensional dense embedding vector, resulting in an input matrix  $E \in \mathbb{R}^{N \times K}$ .

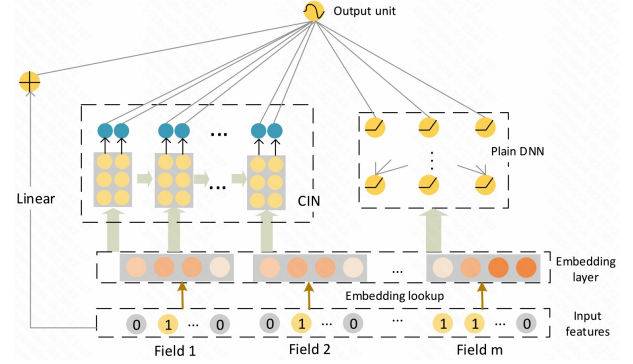


Figure 2: The architecture of xDeepFM based on the pipeline from Lian et al. [21]

The stacking layer combines the dense embedding vectors from the embedding layer with the original dense features. It concatenates the embedding vectors with the dense features and feeds the concatenated input to the subsequent layers.

The cross-network component of the model captures feature interactions through cross-product operations, allowing it to learn both low-order and high-order interactions. It consists of multiple cross layers, each of which computes the outer product of the input with a learnable weight matrix. Let  $H_l \in \mathbb{R}^{N \times (K+D)}$  denote the input to the  $l$ -th cross layer, where  $D$  represents the dimensionality of the dense features. The output of the  $l$ -th cross layer can be computed as:

$$H_{l+1} = H_l \cdot W_l + b_l + H_l \cdot C_l \cdot H_0^T \quad (3)$$

Within the cross-network, the output of the  $l$ -th cross layer is calculated by taking the dot product between the input  $H_l$  and the weight matrix  $W_l$ . The resulting vector is then combined with the bias vector  $b_l$ . Additionally, the weighted interaction between the  $l$ -th cross layer and the input layer, represented by the weight matrix  $C_l$ , is included in the computation.

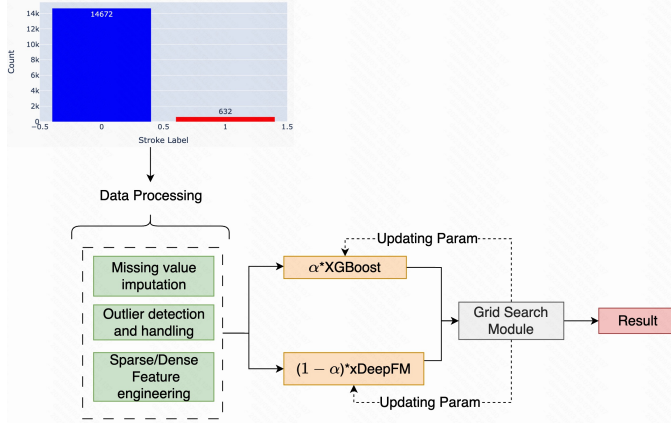
The deep network, on the other hand, is responsible for capturing intricate feature interactions and learning high-level representations. It comprises multiple fully connected layers that utilize non-linear activation functions like ReLU or Sigmoid. The output of the cross-network and the dense features from the stacking layer are combined as inputs to the deep network. The final prediction is obtained by aggregating the outputs of the deep layers.

By combining the strengths of factorization machines and deep neural networks, xDeepFM achieves powerful feature learning and interaction modeling. This approach effectively captures both low-order and high-order feature interactions, making it applicable to a diverse range of recommendation and classification tasks.

### 3.3 Ensemble Approach

To leverage the strengths of both XGBoost and xDeepFM, we propose an ensemble approach [22]. The models are trained separately on the same dataset, and their predictions are combined using ensemble coefficients, obtained through grid search, to determine the contribution of each model to the final prediction.

The ensemble approach aims to benefit from the complementary strengths of XGBoost and xDeepFM, improving the overall predictive performance compared to using either model individually. The architecture is shown in Figure 3.



**Figure 3: Workflow: Ensemble Model.** Stroke data, split into two classes (0-blue and 1-red), is processed and then fed into XGBoost and xDeepFM models. The output is weighted by coefficients  $\alpha$  and  $1 - \alpha$ . A grid search module refines the optimal  $\alpha$  for the ensemble’s final prediction.

### 3.4 Grid Search

Grid search is a systematic approach used to find the optimal values of ensemble coefficients that maximize the overall performance of the ensemble model. It involves defining a grid of possible values for the coefficients and evaluating the performance of the ensemble using each combination of these values. The performance metric, such as AUC, is computed for each combination, and the combination that yields the highest performance is selected as the optimal set of coefficients.

Grid search eliminates the need for manual tuning and guesswork, providing a systematic and automated way to explore different combinations of ensemble coefficients. It ensures that the optimal values are identified objectively, enhancing the performance of the ensemble model.

### 3.5 Loss Function Selection

The selection of a suitable loss function is critical in training predictive models as it determines the objective the model aims to optimize during the learning process. In the case of the ensemble of XGBoost and xDeepFM, the primary loss function used is the Binary Cross-Entropy (BCE) loss. The BCE loss function penalizes the discrepancy between the predicted probabilities and the true labels, thereby encouraging improved calibration and discrimination performance. Additionally, the BCE loss is differentiable, which makes it well-suited for training deep learning models like xDeepFM.

Mathematically, the BCE loss is defined as following:

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)) \quad (4)$$

, where  $N$  is the number of training examples,  $y_i$  is the true binary label, and  $\hat{y}_i$  is the predicted probability.

By employing the BCE loss during the training process, the models are optimized to minimize the discrepancy between the predicted probabilities and the actual labels. This objective promotes enhanced classification accuracy and the generation of reliable probability estimates for stroke prediction.

### 3.6 Evaluation Metric: AUC

In our evaluation of the ensemble model, we utilize the area under the receiver operating characteristic (ROC) curve (AUC) as the performance metric, as proposed by Lobo et al. in their study [23]. The ROC curve is a graphical representation that illustrates the relationship between the true positive rate (sensitivity) and the false positive rate (1-specificity) at various classification thresholds. The AUC is a single-value metric that summarizes the overall performance of the model across all possible threshold settings. Its value ranges from 0 to 1, with higher values indicating a better ability to discriminate between positive and negative instances. The AUC metric is particularly useful when dealing with imbalanced datasets or when the costs associated with false positives and false negatives differ.

To calculate the AUC, we compute the integral of the ROC curve, which represents the probability that a randomly chosen positive example will be ranked higher than a randomly chosen negative example. By optimizing the ensemble model using AUC, we aim to achieve superior performance in terms of both sensitivity (the ability to correctly identify positive cases) and specificity (the ability to correctly identify negative cases), while considering the trade-off between true positive and false positive rates.

The choice of AUC as the evaluation metric enables a comprehensive assessment of the ensemble model’s discriminative power and robustness, facilitating reliable model comparison and selection.

### 3.7 Data Preprocessing

The Stroke Prediction Dataset<sup>2</sup> from Kaggle is utilized in this study to facilitate further analysis and modeling.

To prepare the data for modeling, several preprocessing steps are performed. Missing values in the dataset are addressed using appropriate imputation techniques. Categorical variables are encoded into numerical representations using methods such as one-hot encoding or label encoding. Feature scaling is applied to ensure compatibility with the modeling algorithms. Next, the dataset is divided into training and testing sets while ensuring that consistent transformations are applied to both sets. This ensures that the data distribution remains consistent between the training and testing phases.

By executing these preprocessing steps, the dataset is appropriately prepared to facilitate ensemble modeling using XGBoost and xDeepFM algorithms.

<sup>2</sup><https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>

### 3.8 Experimental Results

The performance of the ensemble model and individual models in stroke prediction is assessed using the area under the receiver operating characteristic curve (AUC) metric. In Table 1, we provide the AUC scores obtained by each model, emphasizing the notable improvements achieved.

**Table 1: AUC Scores of Stroke Prediction Models**

Model	AUC
xDeepFM	0.8792
XGBoost	0.8783
DeepFM	0.8785
FiBiNET	0.8773
DCNMix	0.8789
WDL	0.8786
LightGBM	0.8681
<b>Ensemble Model</b>	<b>0.8841</b>

The ensemble model, combining XGBoost and xDeepFM, achieved the highest AUC score of 0.8841. This represents a substantial improvement in predictive accuracy compared to the other models.

The findings highlight the effectiveness of the combined model, which leverages the individual strengths of XGBoost and xDeepFM. This leads to a notable enhancement in the precision of stroke prediction, thereby reducing the likelihood of duplicate results. The ensemble model’s superior performance suggests its potential for more accurate stroke risk assessments and better decision-making in clinical settings.

## 4 CONCLUSION

In this study, we proposed an integrative paradigm for enhanced stroke prediction by synergizing the XGBoost and xDeepFM algorithms. Our study makes a valuable contribution to the progress of machine learning and deep learning methodologies in the field of stroke prediction, providing valuable knowledge for healthcare professionals in identifying individuals at higher risk of stroke and enabling timely interventions and personalized care. Future research directions include further optimization and refinement of the ensemble model, exploring the integration of additional algorithms, and evaluating the model’s performance on diverse datasets. Furthermore, to enhance the interpretability of the ensemble model, we can incorporate explainability techniques. These techniques enable healthcare professionals to gain insights into the key factors contributing to stroke risk, thereby improving the model’s transparency and facilitating informed decision-making.

## REFERENCES

- [1] Aditya Khosla, Yu Cao, Cliff Chiung-Yu Lin, Hsu-Kuang Chiu, Junling Hu, and Honglak Lee. An integrated machine learning approach to stroke prediction. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 183–192, 2010.
- [2] Jun Wu, Xuesong Ye, and Chengjie Mou. Botshape: A novel social bots detection approach via behavioral patterns. In *12th International Conference on Data Mining & Knowledge Management Process*, 2023.
- [3] Jun Wu and Xuesong Ye. Fakeswarm: Improving fake news detection with swarming characteristics. *arXiv preprint arXiv:2305.19194*, 2023.
- [4] Minhaz Uddin Emon, Maria Sultana Keya, Tamara Islam Meghla, Md Mahfujur Rahman, M Shamim Al Mamun, and M Shamim Kaiser. Performance analysis of machine learning approaches in stroke prediction. In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pages 1464–1469. IEEE, 2020.
- [5] Tianyu Liu, Wenhui Fan, and Cheng Wu. A hybrid machine learning approach to cerebral stroke prediction based on imbalanced medical dataset. *Artificial intelligence in medicine*, 101:101723, 2019.
- [6] Jun Wu, Xuesong Ye, and Yanyuet Man. Bottrinet: A unified and efficient embedding for social bots detection via metric learning. In *2023 11th International Symposium on Digital Forensics and Security (ISDFS)*, pages 1–6. IEEE, 2023.
- [7] Jun Wu, Xuesong Ye, Chengjie Mou, and Weinan Dai. Fineehr: Refine clinical note representations to improve mortality prediction. In *2023 11th International Symposium on Digital Forensics and Security (ISDFS)*, pages 1–6. IEEE, 2023.
- [8] Yafei Wu and Ya Fang. Stroke prediction with machine learning methods among older chinese. *International journal of environmental research and public health*, 17(6):1828, 2020.
- [9] Weinan Dai, Chengjie Mou, Jun Wu, and Xuesong Ye. Diabetic retinopathy detection with enhanced vision transformers: The twins-pecvt solution. In *2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI)*, pages 403–407. IEEE, 2023.
- [10] Maihul Rajora, Mansi Rathod, and Nenavath Srinivas Naik. Stroke prediction using machine learning in a distributed environment. In *Distributed Computing and Internet Technology: 17th International Conference, ICDCIT 2021, Bhubaneswar, India, January 7–10, 2021, Proceedings 17*, pages 238–252. Springer, 2021.
- [11] Tahani Daghistani and Riyadh Alshammari. Comparison of statistical logistic regression and random forest machine learning techniques in predicting diabetes. *Journal of Advances in Information Technology Vol*, 11(2):78–83, 2020.
- [12] Manisha Sanjay Sirsat, Eduardo Fermé, and Joana Câmara. Machine learning for brain stroke: a review. *Journal of Stroke and Cerebrovascular Diseases*, 29(10):105162, 2020.
- [13] Pattanapong Chantamit-O-Pas and Madhu Goyal. Prediction of stroke using deep learning model. In *Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14–18, 2017, Proceedings, Part V 24*, pages 774–781. Springer, 2017.
- [14] Yifeng Xie, Hongnan Yang, Xi Yuan, Qian He, Ruitao Zhang, Qianyun Zhu, Zhenhai Chu, Chengming Yang, Peiwu Qin, and Chenggang Yan. Stroke prediction from electrocardiograms by deep neural network. *Multimedia Tools and Applications*, 80:17291–17297, 2021.
- [15] Yoon-A Choi, Se-Jin Park, Jong-Arm Jun, Cheol-Sig Pyo, Kang-Hee Cho, Han-Sung Lee, and Jae-Hak Yu. Deep learning-based stroke disease prediction system using real-time bio signals. *Sensors*, 21(13):4269, 2021.
- [16] Xuesong Ye, Jun Wu, Chengjie Mou, and Weinan Dai. Medlens: Improve mortality prediction via medical signs selecting and regression. In *2023 IEEE 3rd International Conference on Computer Communication and Artificial Intelligence (CCAI)*, pages 169–175. IEEE, 2023.
- [17] Isha R Chavva, Anna L Crawford, Mercy H Mazurek, Matthew M Yuen, Anjali M Prabhat, Sam Payabvash, Gordon Sze, Guido J Falcone, Charles C Matouk, Adam de Havenon, et al. Deep learning applications for acute stroke management. *Annals of Neurology*, 92(4):574–587, 2022.
- [18] Rui Feng, Marcus Badgeley, J Mocco, and Eric K Oermann. Deep learning guided stroke management: a review of clinical applications. *Journal of neurointerventional surgery*, 10(4):358–362, 2018.
- [19] Noah Stier, Nicholas Vincent, David Liebeskind, and Fabien Scalzo. Deep learning of tissue fate features in acute ischemic stroke. In *2015 IEEE international conference on bioinformatics and biomedicine (BIBM)*, pages 1316–1321. IEEE, 2015.
- [20] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.
- [21] Jianxun Lian, Xiaohuan Zhou, Fuzheng Zhang, Zhongxia Chen, Xing Xie, and Guangzhong Sun. xdeepfm: Combining explicit and implicit feature interactions for recommender systems. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 1754–1763, 2018.
- [22] Joao Mendes-Moreira, Carlos Soares, Alípio Mário Jorge, and Jorge Freire De Sousa. Ensemble approaches for regression: A survey. *Acm computing surveys (csur)*, 45(1):1–40, 2012.
- [23] Jorge M Lobo, Alberto Jiménez-Valverde, and Raimundo Real. AUC: a misleading measure of the performance of predictive distribution models. *Global ecology and Biogeography*, 17(2):145–151, 2008.