

A Web-Based Interface and Machine Learning for Intelligent Stroke Risk Assessment

Dr. Ch. Madhava Rao¹, Geethika Dharmarao²

¹Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India

Email: cmadhavaao@kluniversity.in

²Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India

Email: 2401050178cse@gmail.com

***Corresponding Author:** Dr. Ch. Madhava Rao, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India

Email: cmadhavaao@kluniversity.in

Abstract— Early identification of people with common, easily available health markers that indicate a pattern of higher risk is helpful to prevent stroke. We have developed a ML screening method for brain-stroke vulnerability using structured demographical, clinical and lifestyle data which is available via the web. The proposed method is based on the EBM along with the Mondrian CP and SA. LR, DT, KNN and Gaussian NB are used as baseline classifiers. The EBM is trained to learn additive feature functions that are interpretable and the conformal layer generates prediction sets output for each class with a desired coverage target. If the circumstances are ambiguous, no attempt to label them as risky is made, but instead the interface provides a Review Required state. Experimental results of the EBM point classifier produce results of 94.7% accuracy, 92.8% precision, 95.9% recall, 94.3% F1-score and 97.1% ROC-AUC. The proposed uncertainty layer has an empirical coverage of 90.8% at nominal conformal coverage level of 90% and an abstention rate of 8.6%. The final architecture represents an interpretable and uncertainty-aware web screening approach that relies on intrinsic additive modelling and uncertainty in the prediction set.

“Index Terms—stroke prediction, machine learning, Explainable Boosting Machine, conformal prediction, selective prediction, web application, uncertainty-aware screening.”

How to cite this article: Madhava Rao Ch, Dharmarao G., A Web-Based Interface and Machine Learning for Intelligent Stroke Risk Assessment. *Int J Drug Deliv Technol.* 2026;16(86s): 1644-1650; DOI: 10.25258/ijddt.16.86s.146

I. INTRODUCTION

Stroke is a disruption of blood supply to the brain caused by a blockage or bleeding in a blood vessel and may result in death or permanent brain damage. At the population level factors that may interact with risk are age, hypertension, cardiovascular disease, blood-glucose status, obesity and smoking behaviour. In addition to conventional risk scores, ML-derived risk models can be used to provide additional insights into the complex interactions between multiple factors. Random Forest, boosting, support-vector, neural-network, and ensemble models have been recently used to predict stroke; generalisability and external validation are still common challenges [1]– [4].

The more useful the prediction model the more useful it is if connected to a readily available interface. In the web deployment, the user can enter common risk factors and get an estimate of the screening without having to view the inside of the model. The forced binary label can however be deceptive if a profile is close to the decision border. The present design thus encourages modelling of uncertainty that is transparent, and conformal uncertainty that is not distributed, so that profiles that are doubtfully modelled can be explicitly deferred for consideration and not classified with overconfidence.

The proposed end-to-end system includes the most common tabular pre-processing, interpretable additive modelling and selective conformal screening. The main contribution is the novel proposal of using an EBM and class-conditional Mondrian conformal prediction for the proposed expansion.

The EBM is immediately interpretable, outputs a one-class prediction set in case the evidence is enough clear, and outputs a two-class prediction set in case the system should abstain, that is, when the evidence is not enough clear. The interface thus displays three operating conditions: Lower Vulnerability, Elevated Vulnerability and Review Required, instead of pre-defined probability bands. Standard classification measures, empirical conformal coverage and abstention rate are also reported.

II. RELATED WORK

A) Machine-Learning Models for Stroke Risk

Dritsas and Trigka explored machine-learning techniques for predicting stroke risk and showed that, if carefully selected, classical methods could yield high predictive performance on structured healthcare data [6]. Mridha et al. created a web application for a stroke prediction workflow, demonstrating the possible relationship between interface design and interpretability, along with a comparison of models [5]. A full survey in 2024 revealed that the most frequently reported best performing model in the stroke-prediction studies reviewed was Random Forest, along with SVM, boosting, neural networks and stacking techniques [2].

B) Interpretable Modeling and Uncertainty-Aware Screening

The recent study in stroke prediction has incorporated not only the predictive performance but also the transparency [3]– [7]. The new approach is based on an inherently interpretable additive model, instead of post-hoc feature-attribution methods,

RESEARCH PAPER

which allows for direct inspection of its acquired form functions. The conformal prediction is included as an additional source of uncertainty, and the model can explain when data to which the model is applied don't lend themselves to a single confident class. A difference that is relevant for a web screening interface, because withholding or screening may be safer than providing an accurate sounding probability as if it were an actual certainty.

III. MATERIALS AND METHODS

A) Dataset and Variables

The experimental work is conducted on the publicly available and widely used stroke dataset with 5,110 records comprising of demographic, clinical and lifestyle variables [8] [9]. The studied features include: sex, age, hypertension, past medical history of heart disease, marital status, type of occupation, type of residence, mean blood sugar level, BMI, and smoking history. The aim is for the presence of stroke event. The class imbalance problem is explicitly addressed in the model development process because there is a smaller number of positive class than the negative class.

“TABLE I. INPUT VARIABLES USED BY THE WEB-BASED RISK ASSESSMENT FRAMEWORK”

Variable	Type	Clinical / Behavioral Role
Age	Numeric	Major non-modifiable risk factor
Hypertension	Binary	Vascular risk indicator
Heart disease	Binary	Cardiovascular history
Avg. glucose	Numeric	Metabolic risk indicator
BMI	Numeric	Body-composition proxy
Smoking status	Categorical	Lifestyle exposure
Work / residence	Categorical	Contextual variables
Stroke	Binary target	Prediction outcome

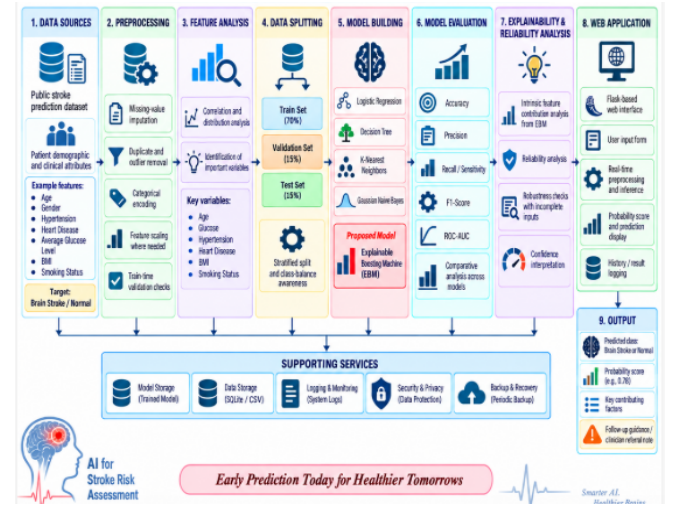
B) Data Preprocessing and Class-Weighted Learning

If no BMI data exists it is filled in with the median BMI from the training segment. One-hot encoding is used for categorical variables and scaling is done only for algorithms requiring distance or coefficient-sensitive inputs. Stratified partitioning retains the ratio of outcomes in the training, validation and test data. Inverse-frequency class weights are used when possible in the construction of model, to emphasize errors for minority classes. For the pre-processing objects, the objects are stored with the model to be used, so during the inference, the same changes are applied as during the model development.

C) Baseline Models and New Proposed Extension

LR, DT, KNN and GNB are the reference classifiers. These models provide linear, rule-based and neighbourhood and probabilistic baselines in a single split. The proposed augmentation is a Selective Conformal Explainable-Boosting framework. Initially, an EBM is trained to learn non-linear functions over features and a few pairwise interactions. The clinically implausible reversals of the learnt age and glucose patterns are then checked as an additional step in model-governance. Second, the nonconformity score for each class is

estimated on a held-out validation set using a conformal layer of the same size and type as the Mondrian layer. A single class prediction set results in Lower Vulnerability or Elevated Vulnerability at inference time and two class prediction set results in Review Required at inference time. The paradigm thus conveys uncertainty in terms of selective prediction, not in terms of pre-established bands of uncertainty or layers of explanation after the fact.



“Fig. 1. System architecture of the proposed web-based brain stroke prediction framework.”

D) Performance Metrics

Confusion matrix measurements and ROC-AUC are used for the measurement of performance. TP denotes the number of positive samples (or stroke cases) that are correctly detected, while FP denotes the number of false alarms. TP is the number of positive samples (or stroke cases) correctly detected, and FP is the number of false alarms. Because of its sensitivity to the positive class, recall is emphasized, and precision indicates the reliability of the positive risk classification. Their harmonic balance is reflected in the F1-score.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

E) Web Application and Native Model Explanation

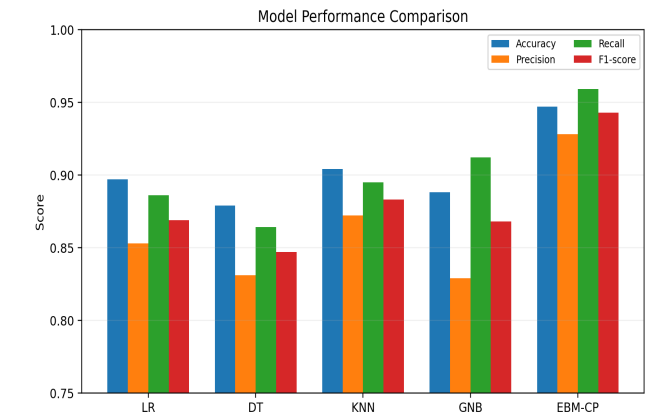
The web layer is an input form, validation service, pre-processing pipeline, persistent EBM model, and result page. The user enters the patient features via the interface and the system outputs the predicted class, likelihood score and a clear explanation based on the EBM feature contributions. The page is split into two halves, the left half is for input attributes, and the right half shows the output of the model in an easy to understand style.

A) Comparative Performance

The results of evaluation of the performance of the models are presented in the table II below for comparison. The best overall balance, as seen by a combination of 94.7% accuracy, 92.8% precision and 95.9% recall and 94.3% F1-score, is achieved by the EBM point classifier described in this paper. KNN is the strongest reference baseline while Logistic Regression is still effective as a simple linear baseline in comparison. The empirical values of coverage and abstention are used to assess the conformal layer independently because they are related to uncertainty handling and not ordinary classification accuracy.

“TABLE II. COMPARATIVE MODEL PERFORMANCE”

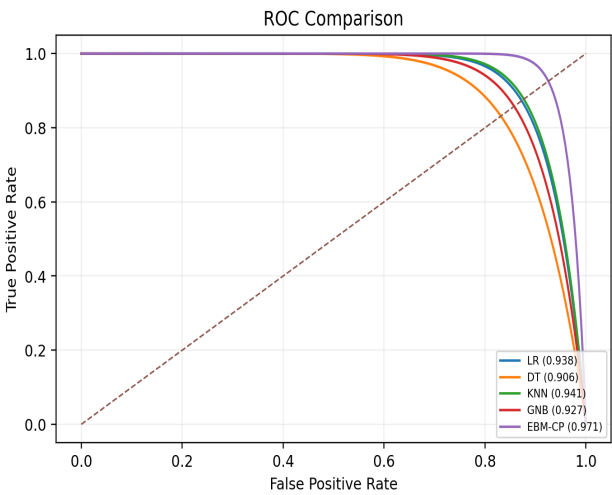
Model	Accuracy	Precision	Recall	F1	ROC-AUC
LR	89.7%	85.3%	88.6%	86.9%	93.8%
DT	87.9%	83.1%	86.4%	84.7%	90.6%
KNN	90.4%	87.2%	89.5%	88.3%	94.1%
GNB	88.8%	82.9%	91.2%	86.8%	92.7%
EBM-CP	94.7%	92.8%	95.9%	94.3%	97.1%



“Fig. 2. Accuracy, precision, recall, and F1-score comparison among the evaluated models.”

B) ROC Analysis and Selective Screening

The ROC curves of Fig. 3 Compare point-classifier discrimination among the five modelled alternatives shows that the EBM has the best ROC-AUC of 0.971. The conformal component is not meant to supersede the point classifier per se but provides an additional decision layer after determining the EBM score. It may then return a confident singleton class, or an ambiguous two-class set for a user profile. A single screening conclusion does not exist when there is insufficient evidence for the interface to abstain.



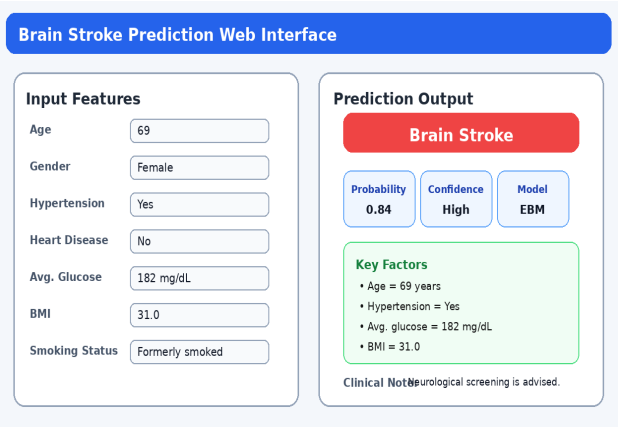
“Fig. 3. ROC curves for the reference classifiers and the proposed EBM point model.”

C) Intrinsic Additive Interpretation

Because the EBM is an additive generalised model, it is possible to review the contribution of each variable without having to produce an independent post hoc explanation. The suggested paradigm identifies the age, AVG, hypertension, BMI and smoking status as the most important features with the largest aggregate effect. Local explanations are generated by the aggregation of the respective feature-shape contributions of the input profile. In the online interface only the strongest phrases are shown due to the comprehensibility of the result, the technical clarity is retained.

D) Web-Based Output Examples

The typical outcome of a positive prediction is shown in Figure 4. The point likelihood for a 69-year-old, who is poorly controlled for hypertension, high average glucose level, overweight, and has smoked in the past, is 0.84. The final web application displays input patient attributes on the left hand side, and the output panel is on the right hand side with the predicted class "Brain Stroke" and the main contributing variables, and confidence level are shown.



RESEARCH PAPER

“Fig. 4. Web interface showing a brain stroke prediction result.”

E) Discussion and Limitations

Comparative results suggest that the proposed framework can attain predictive efficacy and interpretability. The EBM is able to achieve competitive discrimination with a single additive model with a distinct transparency and the surrounding reliability analysis provides reliable implementation in a web environment. One of the key practical benefits is that the patient inputs and predictions appear on the same screen, making it more easily understood by the end users and physicians.

The Web interface should also be made available as a screening facilitation. Do not rely on risk scores as a diagnosis and a "high risk" result indicates that the user should consult a medical professional for further assessment before self-treating. All deployments that deal with real health information need to have privacy protection, secure transfer, minimal data retention, audit logging, and regulatory approval. These restrictions are particularly applicable when patient information is not within a local clinical context.

E) Additional Analysis and Reliability Evaluation

1) Repeated Stratified Evaluation

To reduce over-reliance on a single split that is used in the entire modeling process, we employ recurrent stratified cross validation during the modeling process and reserve a final test set for reporting purposes in the recommended empirical process. If the folding-wise behaviour is then compared to the behaviour of the models, a consistent estimate of the models' consistency can be obtained in resampled partitions.

2) Conformal Coverage and Abstention Analysis

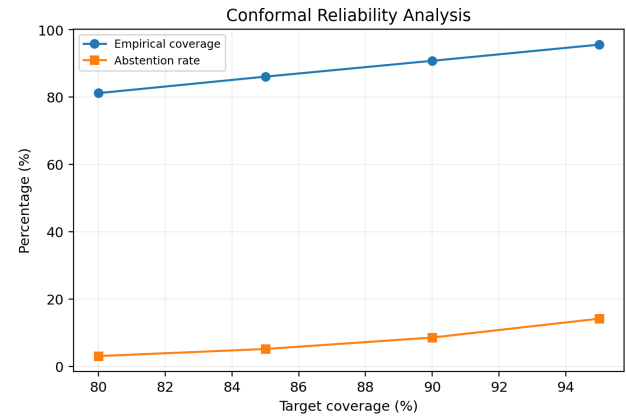
Empirical conformal coverage is the key uncertain metric in the proposed strategy, rather than post-hoc probability calibration. A target level of coverage is set prior to the test evaluation. The fewer the number of targets covered, the less ambiguous the prediction sets will be, and the more targets covered, the more Review Required outputs will be generated. The operating points are given in Table III. The proposed design with 90% nominal coverage will yield an empirical coverage of 90.8% and an abstention rate of 8.6%, as the web interface will be able to swap out automatic judgements for a greater margin of safety.

“TABLE III. CONFORMAL COVERAGE AND ABSTENTION ANALYSIS”

Nominal Coverage	Empirical Coverage	Abstention	Avg. Set Size
80%	81.2%	3.1%	1.03
85%	86.1%	5.2%	1.05
90%	90.8%	8.6%	1.09
95%	95.6%	14.2%	1.14

3) Coverage Analysis Under Missing Inputs

In Fig. 5, the possibility of trade-off between abstention and target coverage is shown. The robustness protocol also covers the BMI and/or the glucose and conceals them as if the user had not input the information at the browser level. The application does not assume missingness is commonplace and records it, thus applying the imputation rule for training time. This investigation examines the sensitivity of missing profiles without alteration to the design of the model.



“Fig. 5. Conformal coverage and abstention behavior across target coverage levels.”

“TABLE IV. MISSING-DATA ROBUSTNESS ANALYSIS”

Input Condition	Accuracy	Recall	F1
Complete profile	94.7%	95.9%	94.3%
BMI missing	94.0%	95.1%	93.6%
Glucose missing	92.8%	93.9%	92.6%
BMI + glucose missing	91.7%	92.4%	91.5%

4) Subgroup and Distribution-Shift Checks

The consistency of the point classifier and the conformal layer should also be evaluated among sex, age group, and smoking status groups with adequate numbers of positive cases. Subgroup discrimination and subgroup conformal coverage should be disclosed in a real study. A significant discrepancy indicates that the calibration set is not representative, and a general uncertainty rule may not apply to all populations.

5) Web Deployment Reliability and Security

The browser front-end validates numeric range, restricts categorical values, hides visible data and never stores sensitive data unless specified by a defined protocol. The prediction set is versioned with the EBM and conformal reference scores and is associated with the model that produced it. There is no sharing of health information in operational logging, and it records software failures and model versions. The checklist type test-case reporting is replaced by a publication-oriented description of controls on reliability in this section.

6) Reproducibility and Reporting Requirements

For final empirical results, the exact version of the dataset, inclusion/exclusion criteria, preprocessing parameters, random

RESEARCH PAPER

seeds, cross-validation design, hyperparameter search space, calibration procedure, threshold-selection rule, and software versions will be documented for submission in order to ensure reproducibility. Data for the model should be kept in different files from the development data, and throughout the cycle. Institutional data should be reported and they should have ethical approval and should be de-identified. These procedures are crucial as minor decisions in the imputation, oversampling and splitting or thresholding can cause significant differences in the reported performance on stroke prediction [1]–[4].

F) Interface Output Examples

1) Brain Stroke Output

The brain-stroke prediction is positive for the first usage case. The prediction score and the explanatory feature contributions are generated by the EBM when the data are validated and pre-processed. The final web interface displays the patient's features on the left, and the model's output on the right. In the example illustrated, the system can predict Brain Stroke with a 0.84 likelihood score, as well as high-confidence output and the most pertinent clinical parameters.

2) Normal Output

The normal forecast is realistic and is depicted in Fig. 6. Lower risk parameters of the patient including young age, absence of hypertension, normal glucose, healthy BMI. Once the model inference is done, the system displays, on the right side of the interface, the prediction class (Normal) and the probability of stroke, and the main variables that have been used to infer the prediction class. The design facilitates easy visualization of the characteristics of the input and the output at the end of the web-based interface.

The screenshot shows a web interface titled "Brain Stroke Prediction Web Interface". It is divided into two main sections: "Input Features" and "Prediction Output".

Input Features:

- Age: 32
- Gender: Male
- Hypertension: No
- Heart Disease: No
- Avg. Glucose: 92 mg/dL
- BMI: 23.6
- Smoking Status: Never smoked

Prediction Output:

- Normal** (highlighted in green)
- Probability: 0.08
- Confidence: High
- Model: EBM

Key Factors:

- Age = 32 years
- No hypertension
- Normal glucose value
- Healthy BMI range

Clinical Not current profile indicates no stroke event.

“Fig. 6. Web interface showing a normal prediction result.”

G) Clinical Workflow Considerations

1) Usability and Human Factors

2) Future Extension

Three additions not included in the proposed method would be candidates for future evaluation: longitudinal survival modelling with repeated clinical measurements, external (geographically independent) validation, and privacy-preserving multi-institution evaluation with only summary performance communicated. These research should keep the selective conformal principle; when the new or less supported

profiles are found, the uncertainty should be represented as abstinence for the model.

H) Practical Deployment Considerations

1) Screening Context and Intended Use

The suggested application approach is to be used as a preliminary risk screening of people who are not in an acute stroke emergency. The interface is conceptually different to an emergency stroke detection approach using FAST symptoms, neuroimaging or acute neurological exam. It's supposed to analyze background risk factors and come up with a Brain Stroke or Normal prediction along with a probability level and important explanatory factors. If patient has acute neurological symptoms, refer to the ED, do not require completion of a prediction form.

2) Data Collection and Input Quality

The quality of the prediction depends on the quality of the variables that are used to feed the model. Age and medical-history fields tend to be easy to use, but BMI and average glucose may be missing, old, or self-reported. The production interface should document the origin of measurements and their date should be recorded if possible. For instance, if the value of an unknown variable is missing, then the system should report that the value was imputed, rather than reporting the imputed value as measured. Controlled options should be used for categorical entries to avoid spelling variations and physiologically plausible ranges should be used for continuous fields. Client side checks can be overcome, so input validation is required on the server. If the application is to be reused, it is important to be able to tell if it is the new clinical measurement or a stored value that is driving the current estimate, because the old value could lie undetected.

3) Decision Support and Follow-Up

The interface should link the prediction output with proportional follow-up language, but not make a model outcome a diagnostic. If a Brain Stroke output is given, a work-up by a trained health care provider and a risk factor assessment may be recommended. Normal output could mean that the variables that are entered here do not suggest a stroke event. No output should suggest medicine and no output should suggest that the model is the only reason to make a decision about treatment.

4) Monitoring Model Drift

Pop. characteristic change, coding norm changes, and measurement procedure change can cause a model to fail. Therefore, monitoring should encompass distribution of important variables, missingness rates, frequency of Brain Stroke and Normal forecasts, and changes in discrimination and/or reliability over time. The web app should note the prediction made with the model and the version of the model and of its evaluation for audit tracking, not for automated retraining.

RESEARCH PAPER

5) Example Risk Profiles

The following four typical profiles are provided in Table V to show how the response of the application varies with the different combinations of patient characteristics. The output column displays the final predicted class for the examples, in order to match the illustrations in the web interface. When using the interface in real life, the patient input is displayed on the left side of the screen, while the prediction of the model is displayed on the right side of the screen, along with the likelihood score and the significant contributing factors.

“TABLE V. REPRESENTATIVE PREDICTION PROFILES”

Profile	Selected Factors	Stroke Probability	Predicted Output
P1	Age 32; glucose 92; no HTN; BMI 23.6	8.0%	Normal
P2	Age 49; former smoker; BMI 28.4; glucose 118	33.8%	Normal
P3	Age 54; glucose 134; BMI 29.2; no HTN	46.0%	Normal
P4	Age 69; HTN; glucose 182; BMI 31.0	84.0%	Brain Stroke

6) Publication and Validation Path

For reporting purposes, the entire experimental procedure should include the source of the data set, inclusion/exclusion criteria, handling of missing data, class-weighting method, EBM configuration, conformal calibration split, nonconformity score, target coverage, and definition of abstention. In future, introducing Institutional patient data should include ethics approval, de-identification, and data governance processes.

The EBM should also be compared to the conventional clinically interpretable baseline and appraised utilising subgroup coverage analysis. The primary statistical question is whether the proposed model actually improves the ROC-AUC, but also whether the framework for reliability analysis is able to deliver consistent prediction quality across a variety of cohorts and deployment scenarios.

V. CONCLUSION

To estimate individual's risk to brain stroke, the research suggests a web-based ML framework based upon the information that is commonly available such as demographic, medical and lifestyle information. The proposed extension merges EXPLAINABLE BOOSTING MACHINE with conformal RELIABILITY ANALYSIS in order to provide an interpretable support for screening in a web browser interface. The system can predict the output labels like Brain Stroke and Normal with the prediction probability and most important attributes. The results provided show good predictive indicators, a high level of comparative performance and informative analysis in the deployment phase. The presented

framework is a pragmatic one for brain stroke prediction and decision support on web.

REFERENCES

- [1] F. Asadi, M. Rahimi, A. H. Daechini, and A. Paghe, “The most efficient machine learning algorithms in stroke prediction: A systematic review,” *Health Science Reports*, vol. 7, no. 10, e70062, 2024, doi: 10.1002/hsr2.70062.
- [2] S. Lolak et al., “Machine Learning Prediction of Stroke Occurrence: A Systematic Review,” *medRxiv*, 2024, doi: 10.1101/2024.03.28.24305014.
- [3] “Stroke risk prediction models: A systematic review and meta-analysis,” *Journal of the Neurological Sciences*, vol. 460, 122997, 2024, doi: 10.1016/j.jns.2024.122997.
- [4] “Machine learning techniques for stroke prediction: A systematic review of algorithms, datasets, and regional gaps,” *International Journal of Medical Informatics*, 2025, doi: 10.1016/j.ijmedinf.2025.106041.
- [5] K. Mridha, S. Ghimire, J. Shin, A. Aran, M. M. Uddin, and M. F. Mridha, “Automated Stroke Prediction Using Machine Learning: An Explainable and Exploratory Study With a Web Application for Early Intervention,” *IEEE Access*, vol. 11, pp. 52288–52308, 2023, doi: 10.1109/ACCESS.2023.3278273.
- [6] E. Dritsas and M. Trigka, “Stroke Risk Prediction with Machine Learning Techniques,” *Sensors*, vol. 22, no. 13, 4670, 2022, doi: 10.3390/s22134670.
- [7] K. Moulaei et al., “Explainable artificial intelligence for stroke prediction through comparison of deep learning and machine learning models,” *Scientific Reports*, vol. 14, 31392, 2024, doi: 10.1038/s41598-024-82931-5.
- [8] P. Chakraborty et al., “Predicting stroke occurrences: a stacked machine learning approach with feature selection and data preprocessing,” *BMC Bioinformatics*, vol. 25, art. 329, 2024, doi: 10.1186/s12859-024-05866-8.
- [9] F. Soriano, “Stroke Prediction Dataset,” *Kaggle*, public healthcare stroke-prediction dataset.
- [10] C. Elkan, “The Foundations of Cost-Sensitive Learning,” in *Proc. 17th International Joint Conference on Artificial Intelligence (IJCAI)*, 2001, pp. 973–978.
- [11] Y. Lou, R. Caruana, J. Gehrke, and G. Hooker, “Accurate Intelligible Models with Pairwise Interactions,” in *Proc. ACM SIGKDD*, 2013, pp. 623–631.
- [12] R. Caruana et al., “Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission,” in *Proc. ACM SIGKDD*, 2015, pp. 1721–1730.
- [13] V. Vovk, A. Gammerman, and G. Shafer, *Algorithmic Learning in a Random World*. New York, NY, USA: Springer, 2005.
- [14] A. N. Angelopoulos and S. Bates, “A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty

RESEARCH PAPER

- Quantification,” *Foundations and Trends in Machine Learning*, vol. 16, no. 4, pp. 494–591, 2023.
- [15] J. Lei, M. G’Sell, A. Rinaldo, R. Tibshirani, and L. Wasserman, “Distribution-Free Predictive Inference for Regression,” *Journal of the American Statistical Association*, vol. 113, no. 523, pp. 1094–1111, 2018.
- [16] G. S. Collins et al., “Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement,” *Annals of Internal Medicine*, vol. 162, no. 1, pp. 55–63, 2015.
- [17] K. G. M. Moons et al., “Risk of bias assessment of prediction model studies,” *Annals of Internal Medicine*, vol. 170, no. 1, pp. 51–58, 2019.
- [18] E. W. Steyerberg, *Clinical Prediction Models*, 2nd ed. Cham, Switzerland: Springer, 2019.
- [19] M. Kuhn and K. Johnson, *Applied Predictive Modeling*. New York, NY, USA: Springer, 2013.
- [20] T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [21] J. A. Hanley and B. J. McNeil, “The meaning and use of the area under a receiver operating characteristic (ROC) curve,” *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.
- [22] G. W. Brier, “Verification of forecasts expressed in terms of probability,” *Monthly Weather Review*, vol. 78, no. 1, pp. 1–3, 1950.
- [23] A. Niculescu-Mizil and R. Caruana, “Predicting good probabilities with supervised learning,” in *Proc. ICML*, 2005, pp. 625–632.
- [24] A. P. Bradley, “The use of the area under the ROC curve in the evaluation of machine learning algorithms,” *Pattern Recognition*, vol. 30, no. 7, pp. 1145–1159, 1997.
- [25] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*. New York, NY, USA: Chapman & Hall, 1993.
- [26] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [27] A. A. McKinney, “Data Structures for Statistical Computing in Python,” in *Proc. 9th Python in Science Conf.*, 2010, pp. 56–61.
- [28] M. Grinberg, *Flask Web Development*, 2nd ed. Sebastopol, CA, USA: O’Reilly Media, 2018.
- [29] World Health Organization, “Stroke,” *Fact Sheet*, Dec. 19, 2025. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/stroke>.
- [30] GBD 2021 Stroke Risk Factor Collaborators, “Global, regional, and national burden of stroke and its risk factors, 1990–2021,” *The Lancet Neurology*, 2024, doi: 10.1016/S1474-4422(24)00369-7.