Multimodal for Stroke Detection

Long Dang and Chaitanya Padamata

Mentors: Yugyung Lee PhD

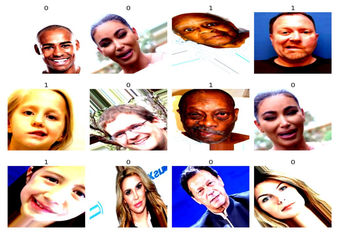
Division of Computing, Analytics & Mathematics
University of Missouri - Kansas City

Abstract

- **Motivation:** Immediate action in stroke cases greatly improves survival rates and outcomes. Early recognition and quick access to emergency care can significantly reduce the risk of severe disability or death in stroke patients.
- **Problem Statement:** Stroke is a critical medical emergency that requires prompt treatment, but many stroke victims face delays in care due to late recognition of symptoms and challenges in accessing emergency services. This lack of timely intervention remains a significant barrier to improving stroke outcomes.
- **Proposal:** Leveraging recent advancements in AI and machine learning, we propose a platform for early stroke detection and fast emergency response. This platform uses AI-driven facial analysis and medical history data to detect potential stroke symptoms, alert users, and direct them to nearby emergency facilities, ensuring faster intervention and better patient outcomes.

Dataset

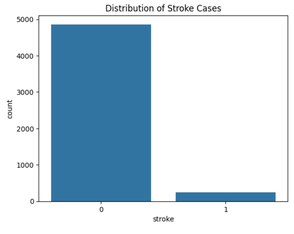
- **Facial dataset:** The dataset comprises **5029** images across two classes—acute stroke and non-stroke cases. Data augmentation techniques, such as flipping, rotation, and scaling, were applied to enhance model accuracy by diversifying and strengthening the dataset to more accurately reflect real-world scenarios.

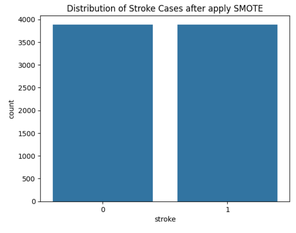


- **Medical history dataset:** This dataset contains **5110** observations with 12 features, and is used to predict whether a patient is likely to get a stroke based on input parameters like gender, age, various diseases, and smoking status

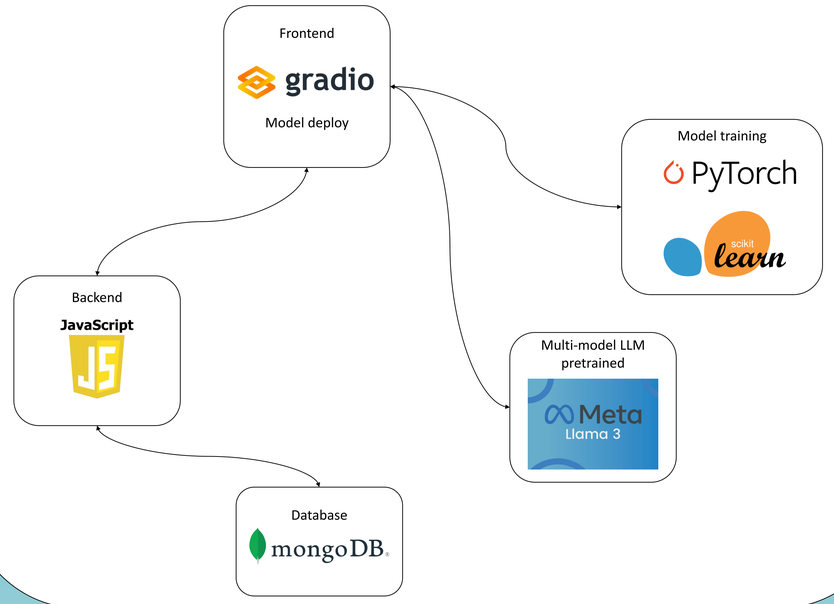
Categorical variables	Continuous variables
1.Gender 2.Hypertension 3.Heart disease 4.Ever married (drop) 5.Work type (drop) 6.Residence type (drop) 7.Smoking status 8.Stroke	1.Age 2.ID (drop) 3.BMI 4.Average glucose level

Note: Since the dataset consists of more non-stroke than stroke cases; hence, the data-balanced technique (SMOTE) was applied.





Architecture



Result

1. Facial stroke prediction

- We study various CNN architectures—ResNet50, ResNet34, EfficientNet, and ConvNeXt—for facial stroke detection.

CNN architectures	Learning rate: 0.001	Learning rate: 0.01	# of iteration until converges
ResNet34			
ResNet50	99.48%		34
EfficientNet			
ConvNeXt			

2. Stroke based medical history prediction

- We used various Machine Learning algorithms – Random Forest, Gradient Boosting, XGBoost, and SVM – for stroke prediction-based medical history.

ML algorithms	Learning rate: 0.05	Learning rate: 0.1	N_estimators: 100	N_estimators: 200
RandomForest				
Gradient Boosting				
XGBoost				
SVM				

3. Total prediction

- We combined the 2 best models with 60% of facial prediction and 40% of medical history prediction.

$$P_{final} = 0.6 \cdot P_{ResNet50} + 0.4 \cdot P_{GradientBoosting}$$

Continues & Conclusion

Limitation & Future Work

References

[1] Contactdoctor/Bio-Medical-MultiModal-Llama-3-8B-V1 · hugging face (no date) *ContactDoctor/Bio-Medical-MultiModal-Llama-3-8B-V1 · Hugging Face*. Available at: https://huggingface.co/ContactDoctor/Bio-Medical-MultiModal-Llama-3-8B-V1 (Accessed: 07 November 2024).