Multimodal for Stroke Detection

Long Dang and Chaitanya Padamata

Mentor: Yugyung Lee PhD

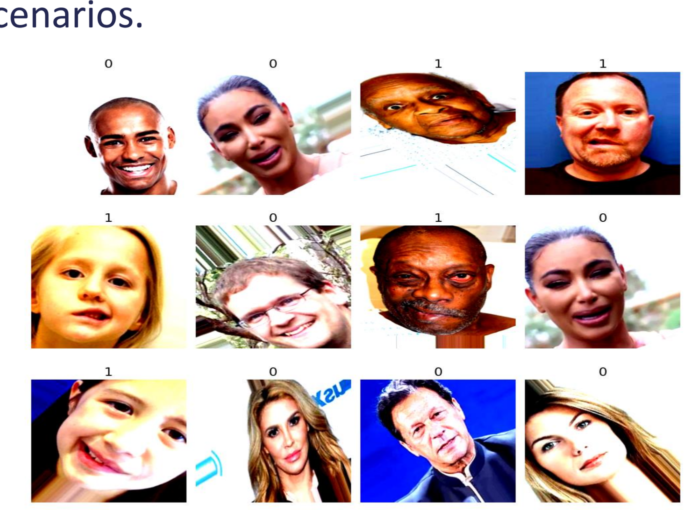
Division of Computing, Analytics & Mathematics
University of Missouri - Kansas City

Abstract

- **Motivation:** In 2022, the CDC reported that 1 in 6 deaths (**17.5%**) from cardiovascular disease in the U.S. was due to stroke. Every 40 seconds, someone in the U.S. has a stroke; every **3 minutes** and **11 seconds**, someone dies from it. Patients who arrive at the ER within three hours of symptom onset experience less disability three months after a stroke. Immediate action in stroke cases greatly improves survival rates and outcomes.
- **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

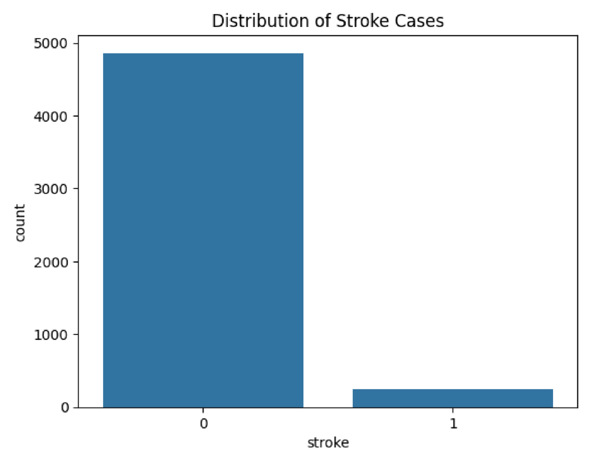
- **Facial dataset:** The dataset comprises **5029** images across two classes—acute stroke and non-stroke cases. Data augmentation techniques, such as flipping, rotation, and scaling, were applied to enhance model accuracy by diversifying and strengthening the dataset to more accurately reflect real-world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

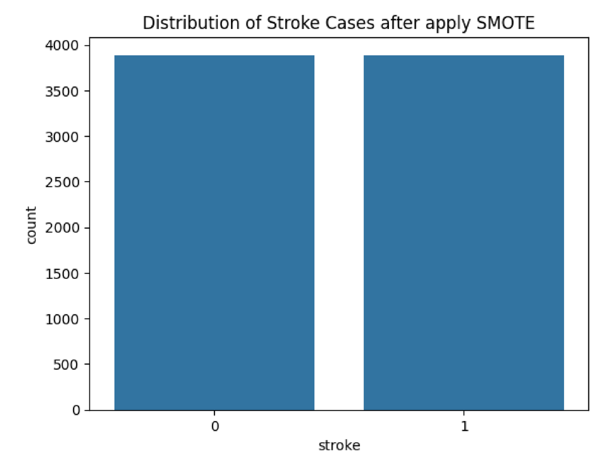


- **Medical history dataset:** This dataset contains **5110** observations with 12 features, and is used to predict whether a patient is likely to get a stroke based on input parameters like gender, age, various diseases, and smoking status

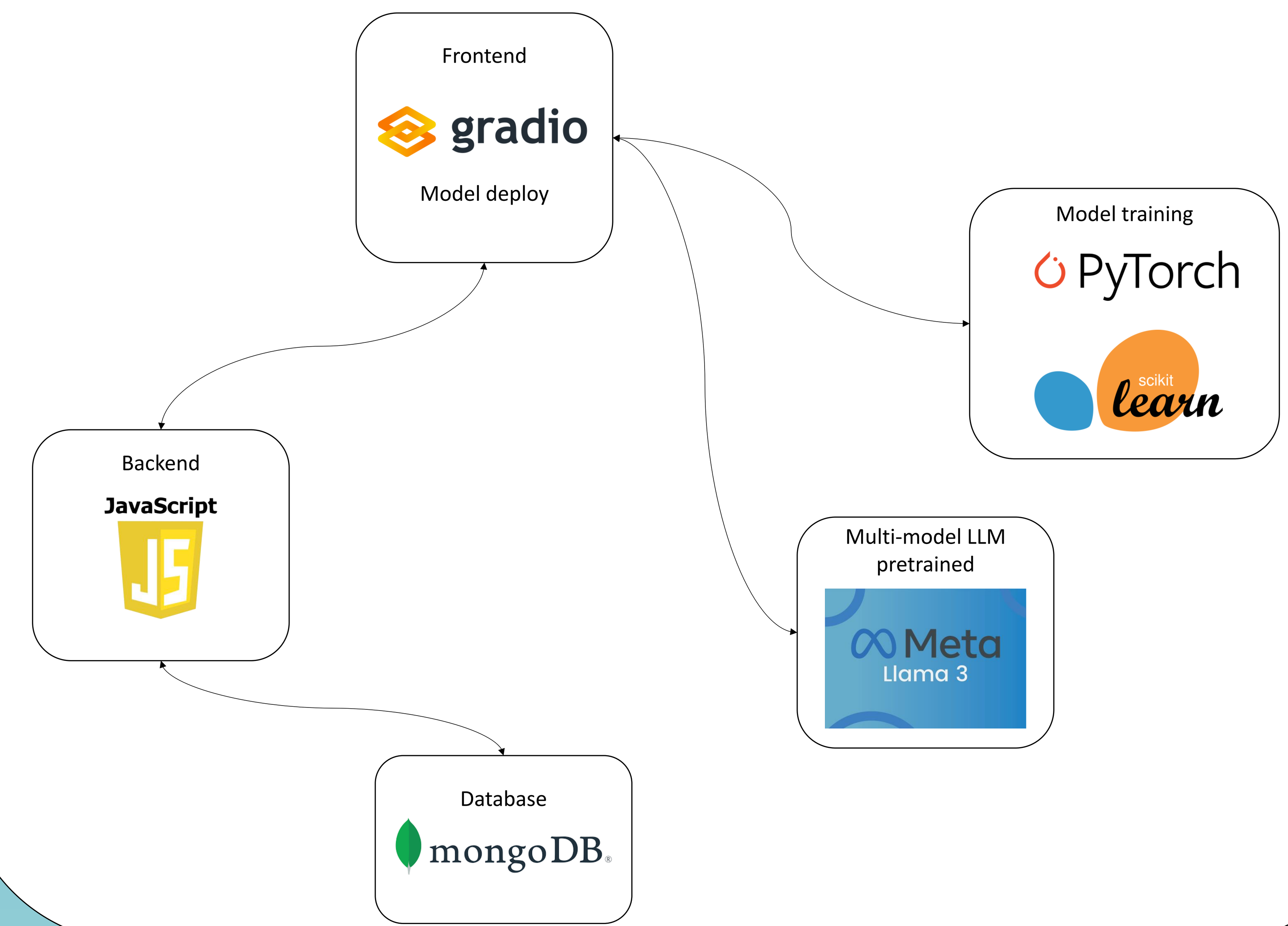
Categorical variables	Continuous variables
1.Gender 2.Hypertension 3.Heart disease 4.Ever married (drop) 5.Work type (drop) 6.Residence type (drop) 7.Smoking status 8.Stroke	1.Age 2.ID (drop) 3.BMI 4.Average glucose level

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





Architecture



Result

1. **Facial stroke prediction**

- We studied various CNN architectures—ResNet34, ResNet50, EfficientNet, and ConvNeXt—for facial stroke detection, where AdamW optimizer was used.

CNN architectures	Learning rate: 0.001	Learning rate: 0.01	# of iteration until convergences (0.001)	# of iteration until convergences (0.01)
ResNet34	98.09%	97.74%	30	24
ResNet50	99.48%	93.02%	34	39
EfficientNet	**99.83%**	**99.13%**	**21**	29
ConvNeXt	99.48%	91.15%	43	**20**

2. **Stroke based medical history prediction**

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

ML algorithms	Accuracy	Precision	Recall	F1 Score	AUC (ROC)
RandomForest	89.56%	0.901%	89.58%	89.53%	89.57%
Gradient Boosting	**96.5%**	**96.53%**	**96.5%**	**96.5%**	**96.5%**
XGBoost	94.19%	94.21%	94.19%	94.19%	94.19%
SVM	79.38%	79.62%	79.39%	79.35%	87.75%

3. **Total prediction**

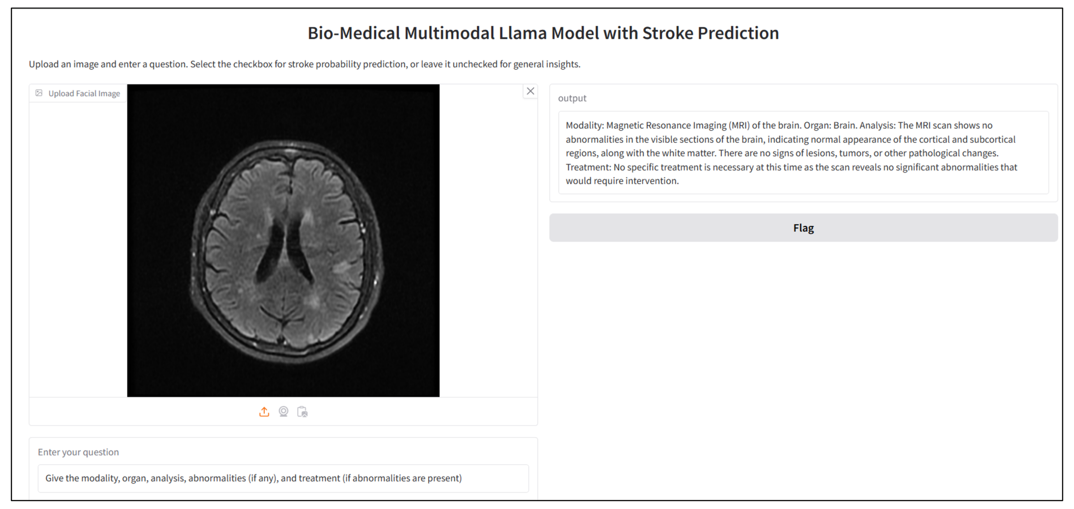
- We combined the 2 best models with 60% of facial prediction and 40% of medical history prediction.

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

Continues & Conclusion

4. **Multimodal LLM**

- We incorporated multimodal Llama 3 – 8b from Hugging face.



5. **Conclusion**

- Successfully implementing a platform to assist users with medical questions and stroke prediction represents a significant innovation in AI-driven healthcare solutions.
- By leveraging advanced machine learning models, including multimodal architectures, the platform offers a robust tool for early stroke detection and provides users with timely, accurate medical insights.

Limitation & Future Work

- Build an emergency switch whenever the prediction is above **95%** chance of having a stroke. In that scenario, a map will be displayed with nearby **ERs** with the wait time sorted from shortest to longest.
- Advanced the platform by not only taking facial and medical history; but also considering every possible scenario such as detecting stroke based on **speech**, **electrocardiogram (EKG)**, and the story of the victim of that day.
- Advanced the front end into more detailed frameworks such as HTML, CSS, and JavaScript instead of built-in Gradio. Also, incorporate modern deep learning frameworks to improve reliability.
- Upscaled the platform by expanding the database and API to make sure the data transfers front and back smoothly.

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).

[2] Jameel, D. (2023) *Face images of acute stroke and non acute stroke, Kaggle*. Available at: https://www.kaggle.com/datasets/danish003/face-images-of-acute-stroke-and-non-acute-stroke (Accessed: 08 November 2024).

[3] Fedesoriano (2021) *Stroke prediction dataset, Kaggle*. Available at: https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset (Accessed: 08 November 2024).

GitHub: Video: Application: