# Harnessing AI-driven modeling for stroke sub-types classification and prediction of survival outcomes


**Suzan Nakasendwa[1*], Joan Eron Mutyoba[1], David Guwatudde[1], Joyce Nabende Nakatumba[2], Agnes Kiragga[3]**

**** Corresponding Author. Email: suzannakasendwa@gmail.com***

[1]School of Public Health, Makerere University, Kampala, Uganda
[2]College of Computing and Information Science, Makerere University, Kampala, Uganda
[3]African Population and Health Research Centre, Nairobi, Kenya


## Background

### 1. Introduction (or educational background)

Stroke is the third leading cause of mortality and disability globally and in the African region. It is the second leading cause of mortality in low-income countries.

Africa bears a significant burden of stroke – ***Figure 1***… fueled by a complex mix of factors that affect one's course of life, such as age, poor nutrition, and exposure to air pollution.

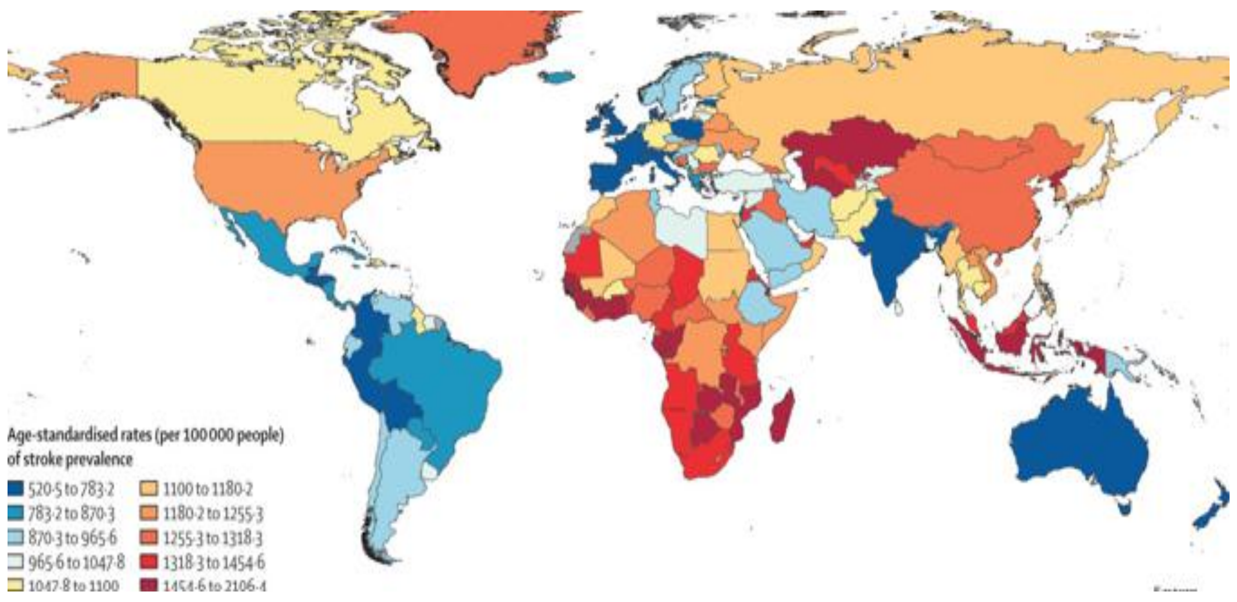

***Figure 1.*** *Global Burden of Age-Standardized Prevalence of Stroke)*

World Stroke Organization fact sheet 2025: 70% stroke burden increase (1990 to 2021), 87% deaths residing in low- and middle-income countries.

Efforts - The United Nations' sustainable development goal, target 3.4: To reduce the burden of premature deaths due to NCDs, including stroke.

### 2. Problem statement

**Critical Challenges**:
- Delayed detection & heterogeneous stroke types (Ischemic vs Hemorrhagic).
- Complex influence of comorbidities (e.g., diabetes, hypertension).
- Limited data infrastructure, especially in Africa.

**Limitations**:
- Paper-based records: Prone to missingness & outliers.
- Electronic systems: Incomplete data
- Statistical techniques struggle with large, high-dimensional, complex data.
- Limited African-focused ML research.
- Need for models addressing multimorbidity & competing risks.

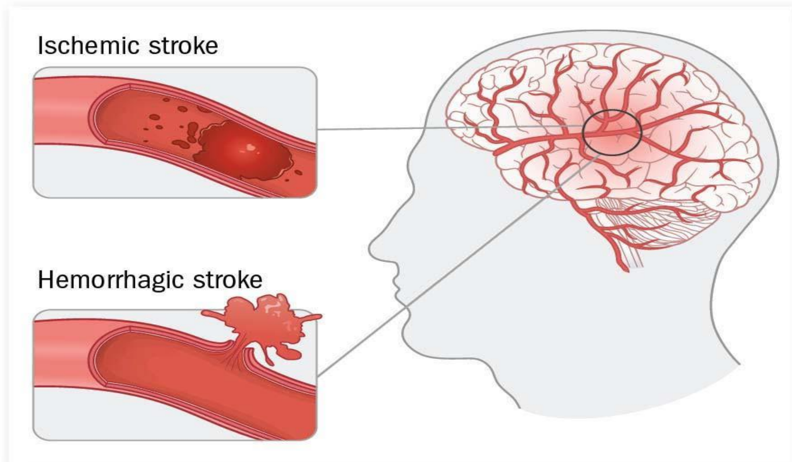


***Figure. 2*** *Stroke Sub-types*

## Current Work

### 1. Objectives

**Overall**: To develop and evaluate advanced machine learning and deep learning models that detect anomalies in health facility stroke data, classify stroke subtypes, and predict time to hospital discharge.

**Specific Objectives**
1. To apply and compare data anomaly detection techniques for quality assurance of routine stroke health facility stroke data

2. To develop an explainable multi-model stroke subtype classifier by integrating structured data using ensemble and deep learning models.

3. To develop and validate a joint modeling framework for survival analysis under competing risks, integrating deep survival learning and conventional models with multimorbidity adjustments.

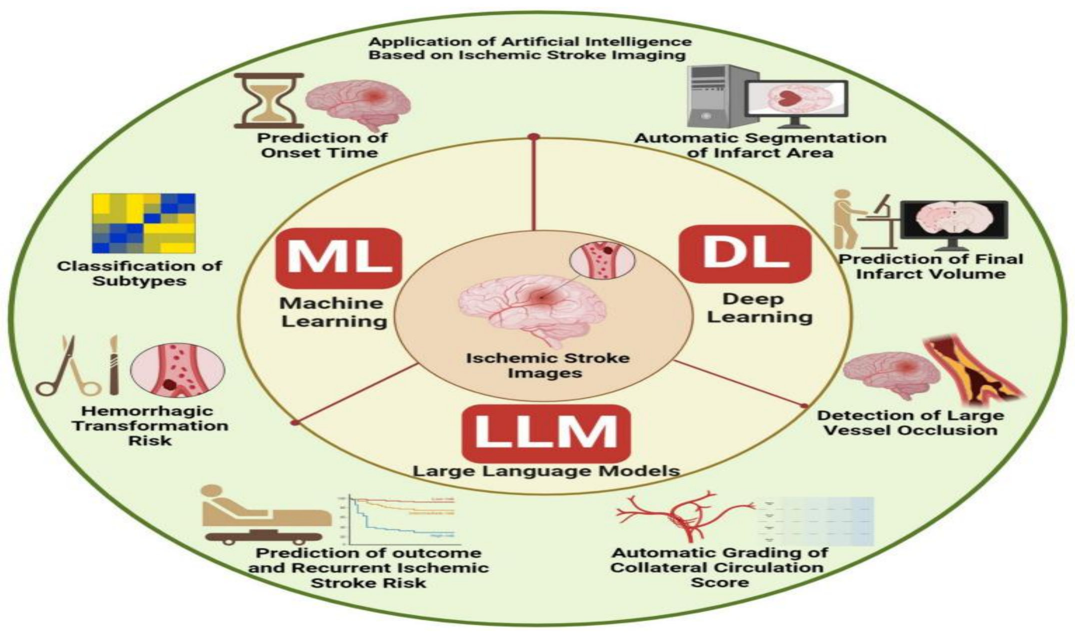


***Figure 3****: Application of AI in stroke research*

### 2. Methodology

**Study design and setting:**
The study will be conducted in Uganda and Cameroon at five (5) health facilities: Lubaga Hospital, Mulago National Referral Hospital, Uganda Heart Institute, Mildmay Uganda Hospital, and Douala General Hospital in Cameroon.

**Study population:** Hospitalized stroke patients seeking care at a health facility.
**Data sources:** Electronic systems and patient files
**Data Collection tool: Redcap**

**Analysis Methods**
**Objective 1**: Machine learning detection algorithms – combining clustering and distance metrics: Euclidean, Canberra, Mahalananobis, and Manhattan

**Objective 2**: Machine learning (logistic regression, decision trees, random forest, support vector machines (SVM), K-Nearest Neighbors (KNN), gradient boosting, AdaBoost, XGBoost), and artificial neural networks (ANN). Shapley Additive exPlanations, and feature importance

**Objective 3**: Compare statistical () and machine learning (Random survival forest, DeepHit, PLANNCR) methods. Discriminative ability will be assessed using time-dependent AUC. The Brier score --- quantify overall prediction error

## Conclusion & Expectations

### 1. Conclusion

This research is of major public health importance as it enhances stroke diagnosis and survival prediction using AI-driven models. It enables timely, accurate, and equitable care in resource-limited African settings while supporting data-informed clinical and policy decision-making.

### 2. Expectations

- A validated AI model capable of distinguishing between ischemic and hemorrhagic stroke using clinical and imaging data.
- A survival prediction tool tailored to diverse African healthcare settings, enhancing prognosis estimation.
- Scalable and adaptable tools to support clinicians and policymakers in improving stroke management and resource allocation.

## Acknowledgment