CS 699

From Theory to Application: Project Roadmap

Anuj Attri (23M0808) Arnav Attri (23M0811)

October 31, 2023

*P*roject Title

Strokes Uncovered: Data Analysis, Visualization, and Predictive Insights.

1 *S*olution Blueprint: An Insight into the Proposed Solutions

1. ***Laying the Foundation***: Configuring the Environment and Libraries.
2. ***Data Narratives***: Unveiling Insights through EDA.
3. ***The Art of Data Cleaning***: Best Practices in Preprocessing.
4. ***Balancing Act***: Visualizing Data Imbalance and Sampling Techniques.
5. ***Beyond Accuracy***: In-Depth Analysis of ML Model Evaluations.
6. ***Deploying the Stroke Model***: A User-Friendly Web Application.

2 *P*reliminary Results

Up to this point, our focus has been on configuring the environment and libraries and performing the *Exploratory Data Analysis (EDA)*. We utilized ***Plotly Express (px) and Seaborn*** to *create* a range of ***visualizations*** that provide us with valuable *insights* into our dataset, enhancing our understanding of *key* factors related to stroke prediction. Moreover, the *results* of which are now available on *our GitHub Repository* in the form of a *Jupyter notebook*[1].

[1] `https://github.com/yourarnav/CS699-SW-Lab`

3 *S*emester Roadmap: A Technical Journey Towards Excellence

After successfully navigating through the phase of *Exploratory Data Analysis (EDA)*, we're now embarking on an exciting journey to address the *significant challenges* in the field of *stroke prediction and prevention*. In the coming weeks, we'll ***focus*** on the following important technical milestones:

1. **Data Preprocessing:** We'll meticulously refine and *optimize our dataset*, ensuring that it's well-prepared for the subsequent modeling phase.

2. **Balancing Data Visualization:** Using advanced *data sampling techniques*, we aim to create a balanced dataset, guaranteeing fairness and accuracy in our analysis.

3. **Modeling and Evaluation:** This is the *high point of our expedition*, where we'll carefully build and fine-tune our *machine learning models* to predict strokes with exceptional precision.

4. **Model Deployment - Building a Web Application:** The *grand finale* of our technical journey will be the creation of an *interactive web application*. This application will allow users to directly experience the outcomes of our hard work and research.