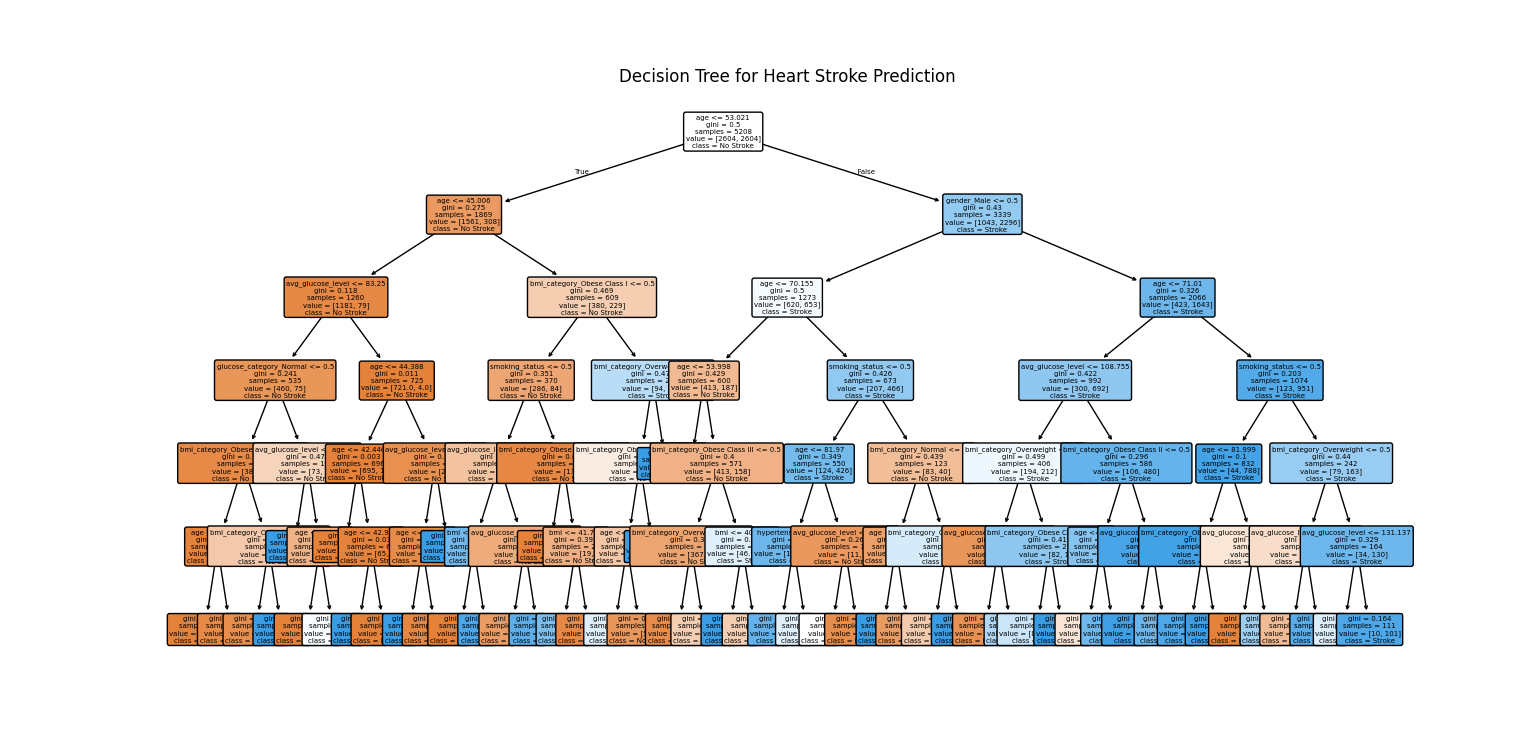
Explanation of Results and Visualizations

# Decision Tree Model

The decision tree model was trained to predict stroke occurrence based on features such as age, BMI, average glucose levels, and smoking status. The tree has a maximum depth of 6 to balance complexity and interpretability. While the overall model accuracy is 78.57%, the classification report highlights significant class imbalance:

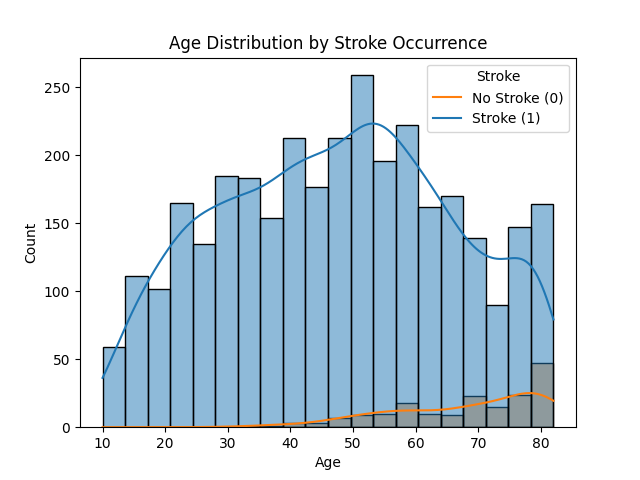
* Class 0 (No Stroke): The model performs well, with high precision (0.95) and recall (0.81).
* Class 1 (Stroke): Performance is limited due to the small sample size of positive stroke cases, with low precision (0.12) and recall (0.39). Oversampling via SMOTE improves recall but doesn't fully address the imbalance.

\*\*The model uses thresholds (e.g., avg\_glucose\_level <= 98.5) to differentiate between stroke and non-stroke cases effectively. Notably, the binary feature smoking\_status is split using a threshold of 0.5, distinguishing between smokers and non-smokers.



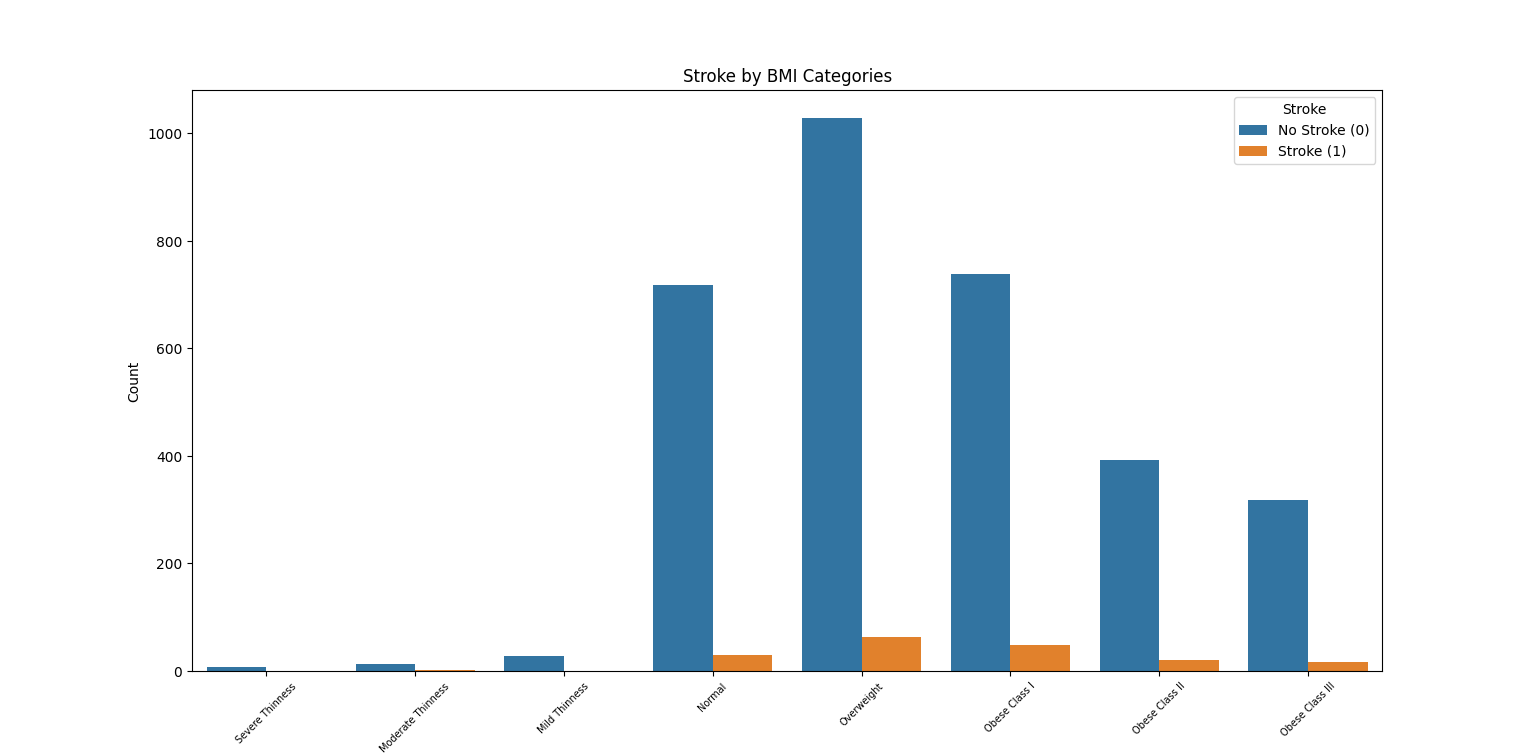
# Age-Stroke Distribution

Stroke incidence significantly increases with advancing age, especially among individuals in their 50s. This highlights age as a critical risk factor, likely due to the accumulation of conditions such as hypertension and vascular issues. Preventative measures should focus on older populations.



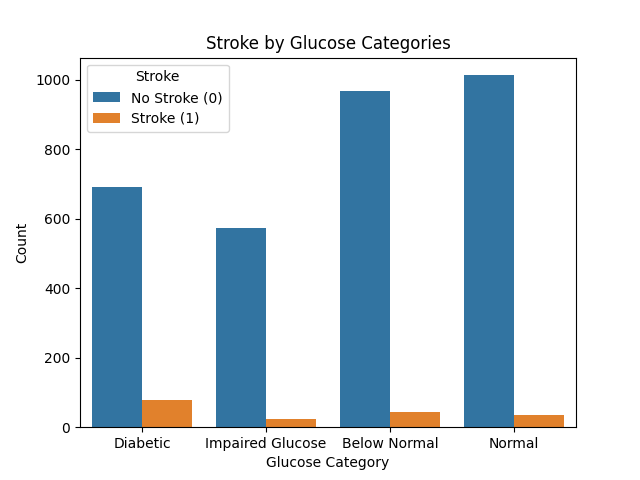
# BMI-Stroke Distribution

The data shows a modest rise in stroke cases among overweight and obese individuals. However, BMI alone is not a strong predictor, suggesting the need to consider other health factors like metabolic health and physical activity when assessing stroke risk.



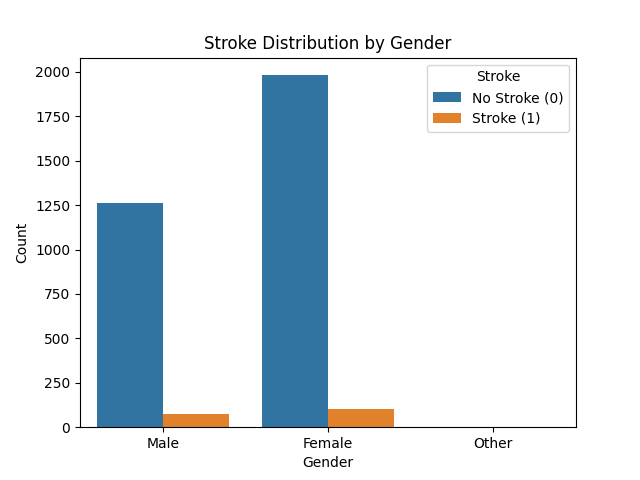
# Glucose-Stroke Distribution

Elevated glucose levels are strongly associated with higher stroke incidence. Individuals with "Diabetic" or "Impaired Glucose" levels face notably higher risks, emphasizing the importance of glucose control through diet, exercise, and medical management.



# Gender-Stroke Distribution

The analysis reveals no significant differences in stroke occurrence between males and females. This indicates that prevention strategies should be equally targeted across genders.



# Marriage-Stroke Distribution

Married individuals show slightly higher stroke rates, potentially linked to age or lifestyle factors. These findings suggest the need for nuanced prevention strategies that consider marital status alongside other variables.

