# Stroke Disease Prediction Dataset

## Overview

This dataset contains information about individuals and their health attributes, collected for the purpose of predicting the likelihood of stroke occurrence. Stroke is a critical medical condition that can lead to serious health issues, and early prediction can aid in preventive healthcare.

## Content

The dataset includes the following features:

* **id:** Unique identifier for each individual.
* **gender:** Gender of the individual (Male, Female).
* **age:** Age of the individual in years.
* **hypertension:** Whether the individual has hypertension (0 - No, 1 - Yes).
* **heart\_disease:** Whether the individual has heart disease (0 - No, 1 - Yes).
* **ever\_married:** Whether the individual has ever been married (Yes, No).
* **work\_type:** Type of occupation (Private, Self-employed, Govt\_job).
* **Residence\_type:** Type of residence (Urban, Rural).
* **avg\_glucose\_level:** Average glucose level in the blood.
* **bmi:** Body mass index (BMI) of the individual.
* **smoking\_status:** Smoking status of the individual (formerly smoked, never smoked, smokes, Unknown).
* **stroke:** Whether the individual had a stroke or not (0 - No, 1 - Yes).

## Usage

This dataset can be used for various purposes such as:

* Exploratory data analysis (EDA) to understand patterns and correlations.
* Building predictive models to predict the likelihood of stroke based on health attributes.
* Research on stroke risk factors and prevention strategies.

## Source

The dataset was sourced from Kaggle.

Feel free to adjust the sections and details as per your actual dataset and its characteristics. This README file provides an overview, content details, potential usage scenarios, source information, citation instructions, and licensing information for the stroke disease prediction dataset.