Name: Nguyen Dang Hoang Khang
Student ID: 109006269

Assignment 2: Hospital Death and Diabetes Mellitus Classification
Report

---

1. **Briefly describe why we take log when implementing the Bayesian classifier? (1%)**
   We take the log to prevent underflow, which happens when the probability becomes too small to represent accurately with floating-point numbers. Therefore, by taking the log, we turn the multiplication into addition, which is computationally more stable

2. **Briefly describe the difference between Naïve Bayesian and Gaussian Naïve Bayesian classifiers? (1%)**
   The main difference between Naive Bayesian and Gaussian Naive Bayesian is the likelihood part. For Gaussian NB, we can use the probability density function (PDF) of the Gaussian distribution. The reason we need to use Gaussian is that when the data type is continuous numbers instead of discrete numbers, we can't build a table by just counting all the possible cases. However, we can assume the data distribution follows a Gaussian (or Normal) distribution by calculating its mean and variance.

3. **Difficulties and solutions:**
   My main difficulty with this assignment is to understand the Bayes theorem and what is the difference between all kinds of probability. In order to understand them, I decided to read some materials from Probability books about Bayes theorem.

4. **Reflections:**
   The assignment provided a deep understanding of Bayesian classifiers, including the Naive Bayesian and Gaussian Naive Bayesian models. I learned how these classifiers make decisions based on probabilistic principles