A Note on Naive Bayes

The Naive Bayes code can return NaN as the class conditional probabilities of the two classes $P(\text{female}|x)$ and $P(\text{male}|x)$ given a new test input x.

Notice that the function corresponding to $P(x|\text{female})$ and $P(x|\text{male})$ needs to be changed since we are now dealing with continuous (real-valued) data as opposed to discrete (categorical) data which was originally the case in the code.

You are required to compute the probability of each of the attributes of the input feature x assuming that these arise out of a Gaussian distribution. As stated in the question, this would involve first computing the mean and standard deviation of each of the attributes for each of the two classes, i.e. for the distributions $\mathcal{N}\left(\mu_{height}, \sigma^2_{height}|female\right), \mathcal{N}\left(\mu_{height}, \sigma^2_{height}|male\right)$ (and the equivalent distributions for the other two attributes). Following this, you will be able to compute $P(x|female)$ and $P(x|male)$ for using these 6 distributions (over 2 classes, and 3 attributes) which you can then apply to the Naive Bayes equation as

$$P(female|x) = \frac{P(x|female)P(female)}{\sum_{gender\in\{female,male\}} P(x|gender)P(gender)}$$

and,

$$P(male|x) = \frac{P(x|male)P(male)}{\sum_{gender\in\{female,male\}} P(x|gender)P(gender)} \ .$$

The values of $P(x|female)$ and $P(x|male)$ according to the Gaussian density (i.e. continuous) function can be approximated by binning (or discretising) the Gaussian function over a large number of points uniformly drawn from the distribution. This is similar to calculating, by using the Matlab function `cdf`, the area under the curve as the probability that x will assume approximately a given value. Alternatively, the Matlab function `normpdf` can be used to produce density values from the Gaussian, which can be numbers greater than one (not to be confused with probabilities), which can be used directly with the Maximum A Posteriori (MAP) decision rule, that is, argmax.