Answer for question 3:

From the comparisons of the accuracy for decision stump and Naïve Bayes shown above, it is shown that decision stump generally has a lower accuracy than Naïve Bayes. This could be because the decision stump only takes one attribute into consideration in the decision-making process, and is not sensitive to the interaction between different attributes. For the datasets that do not contain large number of classes, such as cmc. csv, car. csv and hepatitis. csv, the difference between two accuracies are not too significant.

However, the decision stump accuracy for primary-tumor. csv is shockingly low. Primary-tumor has more classes than attributes and has a large amount of missing values. This quality has caused low accuracy for both classifiers, but decision stump is extremely low. It indicates that decision stump does not work well when there are too many classes or missing values.

An example of another unfortunate situation the decision stump could encounter is shown above. The most frequent classes for each attribute value are the same, resulting in every instance getting classified into the same class. This could be another causation of the low accuracy.

Answer for question 1:

Since the accuracy function and train function here do not delete the instances with missing values, the accuracies for predictions of the dataset with missing values are underestimated. Therefore, the datasets with missing values are not taken into consideration in this specific analyzation. As the comparison between average information gain and accuracy of the Naïve Bayes classifier shown above, it appears that in general, higher information gain leads to higher accuracy for the predictions. Since the information gain calculates the reduction of entropy after knowing the attribute value, a higher information gain implies a less random sample after knowing the attribute value. Naïve Bayes classifier predicts classes basing on the probability of a certain class with the attribute values known. Therefore if the sample is less random with the knowledge of the attribute value, it is more predictable, thus leads to a higher Naïve Bayes classifier accuracy.