Part b)

Lakshita Sethi (21BDA54) Assignment 3

**Write the difference between the following:**

1. Gausian Naive Bayes: This type of classification assumes that features follows a normal distribution. In other words, it should represent a bell shaped normal curve.
2. Multinomial Naive Bayes: It is used for the discrete count. We consider Bernoulli trials here, say the number of times x is observed in n number of trials.
3. Complement Naive Bayes: As the word “Complement” suggests, we consider the probability of an item to all the classes instead of calculating the probability of the item belonging to one specific class.
4. Bernoulli Naive Bayes: This model is useful when our feature vectors are binomial, 0’s and 1’s. It’s main application is text classification where, 0’s is “Word doesn’t occur in the document”, 1’s are “Word occurs in the document”.
5. Categorical Naive Bayes: The categorical Naive Bayes classifier is suitable for classification with discrete features which are categorically distributed.
6. Out-of-core Naive Bayes model fitting: Naive Bayes models can be used to tackle large scale classification problems for which the full training set might not fit in memory. To handle this case, [MultinomialNB](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html" \l "sklearn.naive_bayes.MultinomialNB" \o "sklearn.naive_bayes.MultinomialNB), [BernoulliNB](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.BernoulliNB.html" \l "sklearn.naive_bayes.BernoulliNB" \o "sklearn.naive_bayes.BernoulliNB), and [GaussianNB](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.GaussianNB.html" \l "sklearn.naive_bayes.GaussianNB" \o "sklearn.naive_bayes.GaussianNB) expose a partial\_fit method that can be used incrementally as done with other classifiers.

**What is Jaccard and Cosine Similarity?**

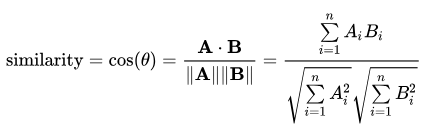
**Jaccard Similarity :**

[Jaccard similarity](https://en.wikipedia.org/wiki/Jaccard_index) or intersection over union is defined as size of intersection divided by size of union of two sets.Jaccard similarity is used for two types of binary cases:

1. Symmetric, where 1 and 0 has equal importance (gender, marital status,etc)
2. Asymmetric, where 1 and 0 have different levels of importance (testing positive for a disease)

# Cosine Similarity:

[Cosine similarity](https://en.wikipedia.org/wiki/Cosine_similarity) calculates similarity by measuring **the cosine of angle between two vectors**. This is calculated as:



With cosine similarity, we need to convert sentences into vectors. Cosine similarity is usually used in the context of text mining for comparing documents or emails. If the cosine similarity between two document term vectors is higher, then both the documents have more number of words in common.

References:

1. <https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#:~:text=There%20are%20three%20types%20of,have%20a%20text%20classification%20problem>
2. <https://scikit-learn.org/stable/modules/naive_bayes.html#:~:text=1.9.-,6.,might%20not%20fit%20in%20memory>
3. <https://towardsdatascience.com/overview-of-text-similarity-metrics-3397c4601f50#:~:text=Jaccard%20similarity%20takes%20only%20unique,term%20frequency%20or%20tf%2Didf>