|  |  |
| --- | --- |
| Name Of The Student | Himanshu |
| Internship Project Topic | TCS iON RIO-210: Build a Classification Model for Drug Trials Dataset |
| Name of the Organization | TCS iON |
| Name of the Industry Mentor | Himdweep Walia |
| Name of the Institute | Amity University |

|  |  |  |
| --- | --- | --- |
| Date | Day # | Hours Spent |
| 17-05-2024 | Day-25 | 5 Hours |
| **Learn about the accuracy checking of Machine learning models.**  **What Is Naive Bayes Algorithm In Machine Learning?**   * Naive Bayes is a machine learning model that is used for large volumes of data, even if you are working with data that has millions of data records the recommended approach is Naive Bayes. * It gives very good results when it comes to NLP tasks such as sentimental analysis. It is a fast and uncomplicated classification algorithm.   **Why is it called Naïve Bayes?**  The Naïve Bayes algorithm is comprised of two words Naïve and Bayes, Which can be described as:  **Naïve:**   * It is called Naïve because it assumes that the occurrence of a certain feature is independent of the occurrence of other features. * Such as if the fruit is identified on the bases of color, shape, and taste, then red, spherical, and sweet fruit is recognized as an apple. * Hence each feature individually contributes to identify that it is an apple without depending on each other.   **Bayes:**   * It is called Bayes because it depends on the principle of Bayes' Theorem.   **Bayes' Theorem**  Bayes' theorem is also known as Bayes' Rule or Bayes' law, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability.  The formula for Bayes' theorem is given as:  IMG_256  Where,   * P(A|B) is Posterior probability: Probability of hypothesis A on the observed event B. * P(B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true. * P(A) is Prior Probability: Probability of hypothesis before observing the evidence. * P(B) is Marginal Probability: Probability of Evidence.   **Types Of Naive Bayes Algorithms**     1. **Gaussian Naïve Bayes:**  * When characteristic values are continuous in nature then an assumption is made that the values linked with each class are dispersed according to Gaussian that is Normal Distribution.  1. **Multinomial Naïve Bayes:**  * Multinomial Naive Bayes is favored to use on data that is multinomial distributed. * It is widely used in text classification in NLP. * Each event in text classification constitutes the presence of a word in a document.  1. **Bernoulli Naïve Bayes:**  * When data is dispensed according to the multivariate Bernoulli distributions then Bernoulli Naive Bayes is used. * That means there exist multiple features but each one is assumed to contain a binary value. So, it requires features to be binary-valued.   **Advantages And Disadvantages Of Naive Bayes Classifier**  **Advantages:**     * It is a highly extensible algorithm that is very fast. * It can be used for both binaries as well as multiclass classification. * It has mainly three different types of algorithms that are GaussianNB, MultinomialNB, BernoulliNB. * It is a famous algorithm for spam email classification. * It can be easily trained on small datasets and can be used for large volumes of data as well.   **Disadvantages:**   * The main disadvantage of the NB is considering all the variables independent that contributes to the probability.   **Applications of Naive Bayes Algorithms**   * **Real-time Prediction:** Being a fast learning algorithm can be used to make predictions in real-time as well. * **MultiClass Classification:** It can be used for multi-class classification problems also. * **Text Classification:** As it has shown good results in predicting multi-class classification so it has more success rates compared to all other algorithms. As a result, it is majorly used in sentiment analysis & spam detection. | | |