ASSIGNMENTS

**1. Differentiate between GIT’s and GIT hub**

**Ans: Git:** Git is a distributed version control system for tracking changes in source code during software development. It is designed for coordinating work among programmers, but it can be used to track changes in any set of files. Its goals include speed, data integrity, and support for distributed, non-linear workflows.

**GitHub:** GitHub is a web-based Git repository hosting service, which offers all of the distributed revision control and source code management (SCM) functionality of Git as well as adding its own features.

**2.What is version control system?**

**Ans:** Version control systems are a category of software tools that helps in recording changes made to files by keeping a track of modifications done in the code. As we know that a software product is developed in collaboration by a group of developers they might be located at different locations and each one of them contributes to some specific kind of functionality/features. So in order to contribute to the product, they made modifications to the source code (either by adding or removing). A version control system is a kind of software that helps the developer team to efficiently communicate and manage(track) all the changes that have been made to the source code along with the information like who made and what changes have been made. A separate branch is created for every contributor who made the changes and the changes aren’t merged into the original source code unless all are analyzed as soon as the changes are green signaled they merged to the main source code. It not only keeps source code organized but also improves productivity by making the development process smooth.

**3.What is GIT repository?**

**Ans:** A **Git repository** tracks and saves the history of all changes made to the files in a Git project. It saves this data in a directory called **.git**, also known as the repository folder.

Git uses a version control system to track all changes made to the project and save them in the repository. Users can then delete or copy existing repositories or create new ones for ongoing projects.

**4. What are the types of GIT repository?**

### Ans: 1. Bare repository:- Software development teams use ****bare repositories**** to share changes made by team members. Individual users aren't allowed to modify or create new versions of the repository.

### 2.Non-Bare repository:- With ****non-bare repositories****, users can modify the existing repository and create new versions. By default, the cloning process creates a non-bare repository.

**5. Differentiate between local and remote repository**

**Ans:** **The local repository**is a Git repository that is**stored on your computer.**

**The remote repository** is a Git repository that is **stored on some remote computer**

The remote repository is usually used by teams as **a central repository** into which everyone pushes the changes from his local repository and from which everyone pulls changes to his local repository.

When you are finished with doing changes into **your workspace**, you can add them to **staging area**and from there you can commit the changes to your **local repository**. This can be done even when you are disconnected from the internet and **nobody else can see the changes** in your local repository.

**6.What is a branch? How do you create one in GIT hub?**

**Ans:Branch:** Branches allow you to develop features, fix bugs, or safely experiment with new ideas in a contained area of your repository. You always create a branch from an existing branch. Typically, you might create a new branch from the default branch of your repository.

**Creating a branch**

1. At the top of the app, click Current Branch and then in the list of branches, click the branch that you want to base your new branch on.
2. Click New Branch.
3. Under Name, type the name of the new branch.
4. Use the drop-down to choose a base branch for your new branch.
5. Click create Branch.

**7. What is GIT hub desktop?**

**Ans:** GitHub Desktop is an application that enables you to interact with GitHub using a GUI instead of the command line or a web browser. GitHub Desktop encourages you and your team to collaborate using best practices with Git and GitHub.

**8.Differentiate between classification and regression?**

**Ans:**

|  |  |
| --- | --- |
| **Classification Algorithm** | **Regression Algorithm** |
| The mapping function is used for assigning values to predefined groups. | The mapping function is used for the assignment of values to continuous output. |
| In Classification, the output element must be a discrete attribute. | In Regression, the output element must be of the constant type of real value. |
| The role of the classification algorithm is to map the input value(x) with the discrete output variable(y). | The role of the regression algorithm is to map the continuous output variable(y) with the input value (x). |
| Classification Algorithms are used for discrete data. | Regression Algorithms are used for continuous data. |
| In Classification, we strive to locate the judgment limit, which may split the dataset into different classes. | In Regression, we strive to find the best match rows, which can forecast the performance more accurately. |
| Classification Algorithms may be used to solve classification problems such as Voice Recognition, Identification of spam emails, Identification of cancer cells, etc. | Regression algorithms may be used to solve the regression problems such as House price prediction, Weather Prediction,  etc. |
| The Classification algorithms can be classified into Multi-class Classifier and Binary Classifier. | The regression Algorithm can be further separated into Non-linear and Linear Regression. |

**9.Write short note on 1.Nominal data 2.Categorical data**

**Ans:Nominal Data:** is data that can be labelled or classified into mutually exclusive categories within a variable. These categories cannot be ordered in a meaningful way.

For example, for the nominal variable of preferred mode of transportation, you may have the categories of car, bus, train, tram or bicycle.

**2.Categorical data:**

Categorical data is a type of data that is used to group information with similar characteristics, while numerical data is a type of data that expresses information in the form of numbers. Example of categorical data: gender. Fake data for ML, synthesized from production.

**10.Explain OHE(One hot ending)**

**Ans:** Basic of one hot encoding using various ways: **numpy, sklearn, Keras etc**. The machine cannot understand words and therefore it needs numerical values so as to make it easier for the machine to process the data. To apply any type of algorithm to the data, we need to convert the categorical data to numbers