Varnika Mulay

ML basics assignment

1. Clustering Vs Classification

S.No.	Clustering	Classification
1.	Clustering is a technique that comes under unsupervised Machine Learning.	Classification is a technique that comes under supervised Machine Learning.
2.	Clustering doesn't necessarily require labeled data.	Classification requires labeled data for analysis purposes.
3.	In clustering, similar data points are grouped to form clusters.	The new and unseen data points are classified into predefined labels, or labels are predicted for the same.
4.	Suitable for detecting outliers in the dataset.	It doesn't help in identifying outliers.
5.	Some commonly used methods are K-Means, DBSCAN, etc.	Some widely used methods are Decision Tree, SVM, regression, etc.
6.	**Ex:** If a school wants to add a few new co-curricular activities and wants to analyze which activities to introduce, they can perform clustering based on the student's hobbies and decide accordingly.	**Ex:** If a school wants to print report cards of each student they can use classification to analyze each student's marks and decide whether they passed or failed and classify them accordingly.

2. Regression Vs Classification

S.No.	Regression	Classification
1.	Regression is a widely used supervised Machine Learning technique.	Classification is also a widely used supervised Machine Learning technique.
2.	The relationship between the dependent and independent variables is analyzed in regression.	The new and unseen data points are classified into predefined labels, or labels are predicted for the same in classification tasks.
3.	The data used in regression is continuous in nature.	The data used in classification is discrete in nature.

4.	Methods: Linear regression, Ridge and lasso regression, etc.	Methods: SVM, Decision Tree, Random Forest, etc.
5.	**Ex:** Predicting weather forecast of the day or week could be an example of regression analysis.	**Ex:** Email spam detection could be an example of classification.