* What is Hierarchical clustering algorithm?

It is an unsupervised machine learning algorithm is used to group unlabelled dataset into a single group is called as cluster. It is known as HCA. In this we try to create the hierarchy of cluster in the form of tree it is called as dendogram.

* What are types of hierarchical clustering?

1. Agglomerative
2. Divisive

* What is Agglomerative hierarchical clustering?

It is bottom-up approach. In which algorithm start taking all data points as single cluster and merging them until one cluster is left.

* What is divisive hierarchical clustering?

It is top-down approach

* What we can do with the outlier?

1. We can remove outlier
2. Don’t remove outlier

* What is confusion matrix?

It is table that is used to estimate the performance of the model. It tabulates the actual values and predicted values in a 2x2 matrix.

* What is bias in data science?

When algorithms is not strong enough to capture the underlying pattern in data. This may lead to lower accuracy because of underfitting.

* What is dimensionality reduction?

Dimension reduction is the process of converting a dataset with high number of dimensions to a dataset with a lower number of dimensions.

* What is a normal distribution

Mean = Median = Mode

* What is P value in hypothesis testing?

If p value is more than the critical value then fall to reject H0

If p value = 0.5 -Strong evidence

If p value =0.05 Weak evidence

If p value less than critical value then we can reject H0

If p value =0.05-weak evidence

If p value = 0.5 strong weak evidence

* How we can add two dict?

A={a:25}

B={l:’hi’}

C=print(\*\*{A}\*\*{B})

* What is Selection bias?

Selection bias is the bias that occurs during the sampling of data.

* What is Precision?

Precision = True positive/ True positive + False positive