**Hope AI (Week - 3) Questions**

**Thanigaivel G**

1. **Difference Between Training and Testing Dataset**

Dataset is a collection of large amount of data, which is used to train our ML Model. We split this dataset into two proportions, training dataset and testing dataset. We use this training dataset to **train our model** for a particular problem. Once the model is trained and ready, we use testing dataset which contains different set of data to **test our model** and find its **accuracy**. For building a best model, we need to train the model with large amount of data and to test it appropriately.

1. **Name Three Evaluation Metrics for Regression Models.**

Regression Model is one of the type of Machine Learning Model in the Artificial Intelligence. We use Regression Model when our problem statement can be solved using Supervised Learning. We Select Regression Model when our prediction is a continuous numeric value. There are several metrics available to evaluate our Regression Models.

* Mean Absolute Error (MAE)
* Mean Squared Error (MSE)
* R2 Squared

1. **Why Feature Scaling Important in Regression**

Feature scaling is important, since when the feature (independent) is similar, then we our model is learns it fast and It will predict the value with better accuracy. When the feature is similar, there will not be large difference of values between them, hence it improves the accuracy.

1. **What Type of Data is Unsuitable for Regression**

Regression Model is one of the Machine Learning Model and It is used when our problem statement is used to predict the continuous **numerical values**. The Categorical data is not suitable for Regression, we need to convert those data to numeric (0 or 1) to build our ML Model.

1. **How does Mean Absolute Error is different from Mean Squared Error**

Both Mean Absolute Error and Mean Squared Error were the evaluation metrics used to evaluate the Regression Models to finds its **accuracy**. Mean Absolute Error is the absolute difference between the actual and predicted value whereas the Mean Squared Error is the squared difference between the actual and predicated value.

1. **How to interpret coefficients of a Regression Model**

The Coefficients of the Regression Model, tells the relationship between the feature and the outcome. If the coefficient is positive number, then the feature increases and the outcome also increases, If the coefficient is negative number, then the feature increases but the outcome decreases.

1. **What is residual in Regresssion**

Residual is the **difference between the actual and the predicted value**. If the Residual is low then we can say that our model is performing better and if the residual is more then we need to fine tune our model with different hyper tuning parameters or to train with large amount to data.

1. **Why Cross Validation is important in Regresssion**

Cross Validation is important in regression model, since in cross validation multiple times the dataset is **splits into training dataset and testing dataset in different proportions**, so it guarantees the accurate result of model prediction and we can used this technique to evaluate our model.

9) **How would you interprete R2 value of 0.85**

R2 Squared is one of the evaluating metrics used to evaluate the accuracy of our regression model. If the R2 value is 0.85, we can conclude that our model is **above average** and we can try to improve the model by using the several combination of hyper tuning parameters and to train the model with large set of data.

10) **What steps would you take if your regression model is too slow to converge**

When the Model is too slow, then we preprocess the dataset by using standardization techniques and we can change the hyper tuning parameters or to change the algorithm which is suitable to our dataset by building and training different algorithm.

11) **A Regression Model is underperforming, What steps can you take**

If the Trained Regression Model is underperforming, that is the accuracy of the model is low, then we try with the different hyper tuning parameters to create that model and to **train the model with large datasets**, again if the accuracy is not up to mark, then we can try other regression models.