Bryan Leonardo 2201745004

Rio Nagano 2201767232

Ronaldo Kenny Chandra 2201763234

1. Can you name four of the main challenges in Machine Learning?

* Training data sedikit
* Data yang tidak berkualitas seperti error, noise dan data asing
* Fitur yang tidak relevan
* Training data yang overfitting atau underfitting

1. What is a test set and why would you want to use it?

Test set adalah bagian dari dataset yang digunakan untuk mengukur keakuratan dari machine learning. Test set digunakan setelah machine learning mempelajari pattern dari training set.

1. What is the purpose of a validation set?

Tujuan dari validation set adalah memilih model dan hyperparameters yang dapat bekerja dengan baik. Validation set menghentikan proses training ketika ditemukan beberapa error pada training dataset.

1. What is cross-validation and why would you prefer it to a validation set?

Cross validation adalah metode statistik yang mengevaluasi kinerja algoritma terhadap proses learning data dan validasi data. Cross validation lebih diminati untuk validation set karena memiliki waktu komputasi yang lebih sedikit namun tetap menjaga keakuratan estimasi.