Althea Maxene B. Viar
BSCS 3B

Activity #3

Feature engineering is very important in improving the performance of advanced machine learning models because it helps the model understand the data better. It involves selecting useful data, cleaning it, and transforming it into a format that the ML algorithm can easily learn from. Even if we use very advanced ML models, if the input data is poor, the results will also be poor.

One way feature engineering improves ML performance is by removing noise and unnecessary information. In real life datasets, data is often messy. There can be missing values, duplicate entries, or extreme values that can confuse the model. By cleaning the data, like filling missing values or removing outliers, the model can focus more on the important patterns instead of learning from errors. This makes training more stable and improves prediction accuracy.

Another reason feature engineering is important is because it improves how data is represented. Raw data is not always easy for ML models to understand. For example, text, dates, or categories need to be converted into numbers before they can be processed. Techniques like encoding categories or scaling numbers help the model process the data more effectively. When data is properly transformed, the model can detect patterns faster and more accurately.

Feature engineering also helps reduce model complexity and training time. If we only keep the most important features, the model will have fewer inputs to process. This not only makes training faster but also helps prevent overfitting. Overfitting happens when the model memorizes the training data instead of learning general patterns that can be used for new data.

Lastly, feature engineering allows us to apply real world knowledge to the model. For example, combining related variables can create more meaningful features. This helps the model make predictions that are closer to real-life situations.

Feature engineering improves ML performance by cleaning data, improving data representation, reducing unnecessary complexity, and adding domain knowledge. Even with modern advanced ML models, good feature engineering is still one of the most important steps in building accurate and reliable systems.