**Main Challenges of Machine Learning:**

**Insufficient Quantity of Training Data:**

In Machine Learning (ML), models require large amounts of training data to perform well. Unlike humans, who can recognize an apple after seeing just a few, ML algorithms need thousands or even millions of examples. For instance, image recognition systems require vast datasets to accurately classify objects.

**Nonrepresentative Training Data:**

For a Machine Learning model to make accurate predictions, its training data must represent real-world cases well. If important data is missing or biased, the model may fail. Even large datasets can be flawed if the sampling method is incorrect. This problem is called **sampling bias** and can lead to inaccurate predictions.

**Sampling Bias:**

**Sampling bias** occurs when the training data does not accurately represent the real-world cases the model needs to predict. This leads to incorrect or biased results.

**Example:** If a survey on smartphone usage is only conducted among college students, the results will not represent older adults, leading to biased conclusions. Similarly, if an ML model for disease prediction is trained only on hospital patients, it may not work well for the general population.

**Poor-Quality Data:**

Poor-quality data, like errors, outliers, or missing values, can reduce a Machine Learning model’s accuracy. Cleaning the data improves performance.

**Solutions:**

* Remove or fix incorrect data.
* Handle missing values wisely (e.g., fill with averages).

**Irrelevant Features:**

If a dataset has too many useless features, the model may not learn well. Feature engineering helps by selecting, combining, or creating important features to improve learning.

**Overfitting the Training Data:**

**Overfitting** happens when a model learns too much from the training data, including noise, and fails to perform well on new data. It's like judging all taxi drivers as bad based on one bad experience.

**Underfitting the Training Data:**

**Underfitting** happens when a model is too simple and fails to capture important patterns in the data, leading to poor predictions.

**Fixes for underfitting:**

* Use a more complex model.
* Improve the input features.
* Reduce model restrictions (e.g., lower regularization).