ML Pipeline
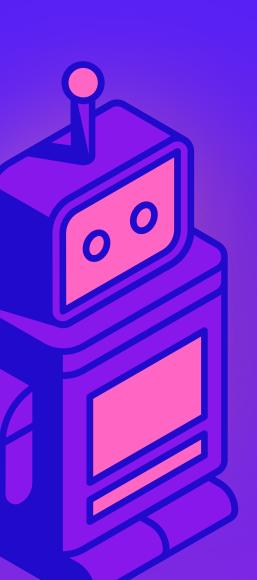
Steps on Machine Learning Pipeline

Data Collection

01

This step gathers raw data from various sources like files, databases, or sensors. The goal is to collect enough relevant and accurate data to represent the problem.
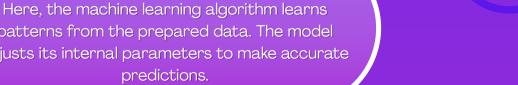
Preprocessing

02

The collected data is cleaned and transformed to make it suitable for analysis. This includes handling missing values, removing errors, and converting data into a usable format.

Training

03

Here, the machine learning algorithm learns patterns from the prepared data. The model adjusts its internal parameters to make accurate predictions.

Evaluation

04

The trained model is tested on new data to measure its performance. Evaluation helps ensure the model can generalize well and isn't just memorizing training examples.

Deployment

05

The final model is integrated into a real-world system or application for use. It continuously makes predictions and may be updated as new data becomes available.

BSCS3B