| Data Splitting | Data Prep | Training | Evaluation | Tuning |

Model Deployment Workflow

Data Splitting

This allows you to have isolation between datasets before doing data preparation and prevent any issues where data preparation on the entire dataset can alter model performance.

Data Preparation

Data needs to be in a format the model can use for training. Most data is dirty, so it will need to be cleaned in a variety of ways so the model can use it. During the EDA process, some data may already be in a prepared state.

Training

Model training works best with clean data. Data splitting and data prep feed into training

Evaluation

Evaluate trained models on unseen data. The outcome of evaluation leads to product decisions.

Tuning

Tuning changes how the model is trained. Optimizations in tuning feedback into previous steps. Workflow is code, allows repeatable iteration.

This workflow can iterate over itself, allowing the outcome of evaluation and tuning to affect how data preparation, training, evaluation, and tuning operate.