

Word	Description
Early Stopping	<p>Early stopping is a technique for avoiding overfitting when training a machine learning model with iterative method. We set the early stopping in such a way that when the performance has stopped improving on the held-out validation set, the model training stops.</p> <p>For example, in XGBoost, as you train more and more trees, you will overfit your training dataset. Early stopping enables you to specify a validation dataset and the number of iterations after which the algorithm should stop if the score on your validation dataset didn't increase.</p>
EDA	<p>EDA or exploratory data analysis is a phase used for data science pipeline in which the focus is to understand insights of the data through visualization or by statistical analysis.</p> <p>The steps involved in EDA are:</p> <ol style="list-style-type: none">1. Variable IdentificationIn this step, we identify the data type and category of variables <div><div><div>Type of Variable</div><div>Predictor Variable<ul style="list-style-type: none">- Gender- Prev_Exam_Marks- Height- WeightTarget Variable<ul style="list-style-type: none">- PlayCricket</div><div>Data Type</div><div>Character<ul style="list-style-type: none">- StudentID- GenderNumeric<ul style="list-style-type: none">- PlayCricket- Prev_Exam_Marks- Height- Weight</div><div>Variable Category</div><div>Categorical<ul style="list-style-type: none">- Gender- PlayCricketContinuous<ul style="list-style-type: none">- Prev_Exam_Marks- Height- Weight</div></div><ol style="list-style-type: none">2. Univariate analysis3. Multivariate analysis</div>