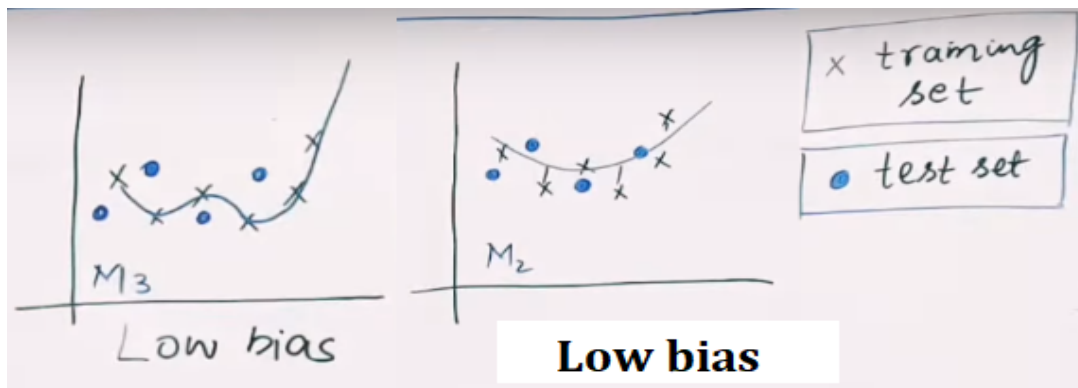


All the images and explanations learnt below, are from [Campusx tutorial](#).

Bias (The inability of a model to truly capture the relationship in the training data)

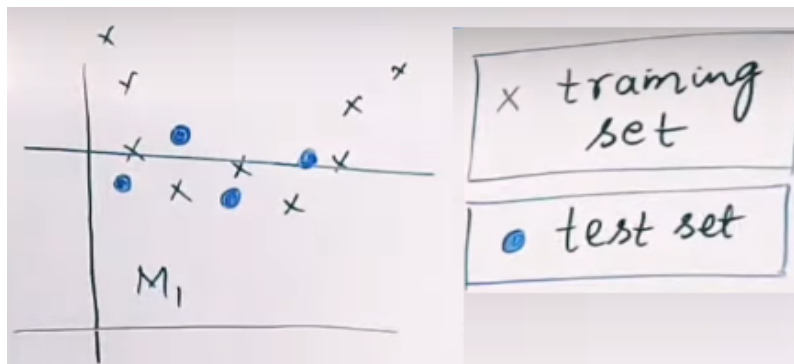
1. When the model fits the Training dataset **to perform good in training** instead of trying to understand the training dataset, **the bias is low**.

Here below, the M3 model captures the pattern of the training dataset very well. So bias is low. And the M2 model had gone through all the points closely i.e. it tried to understand the pattern of the entire dataset closely, so bias is comparatively low.



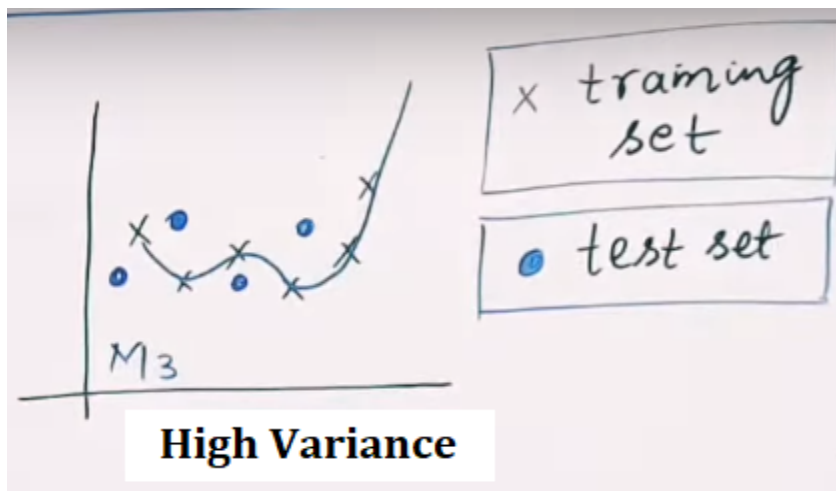
2. But when the model couldn't understand the pattern of the training dataset, the bias becomes high.

Below, the M1 model just made a straight line that specially couldn't go through the upper training points at all, so the bias is high.

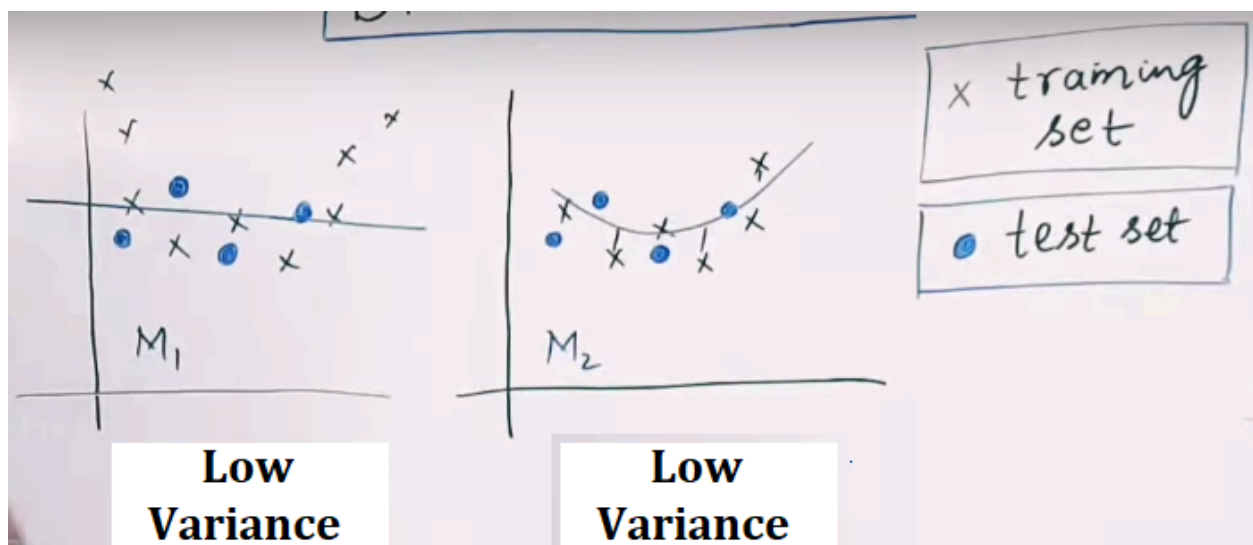


Variance(It measures how much the predictions of a model vary for different training datasets.)

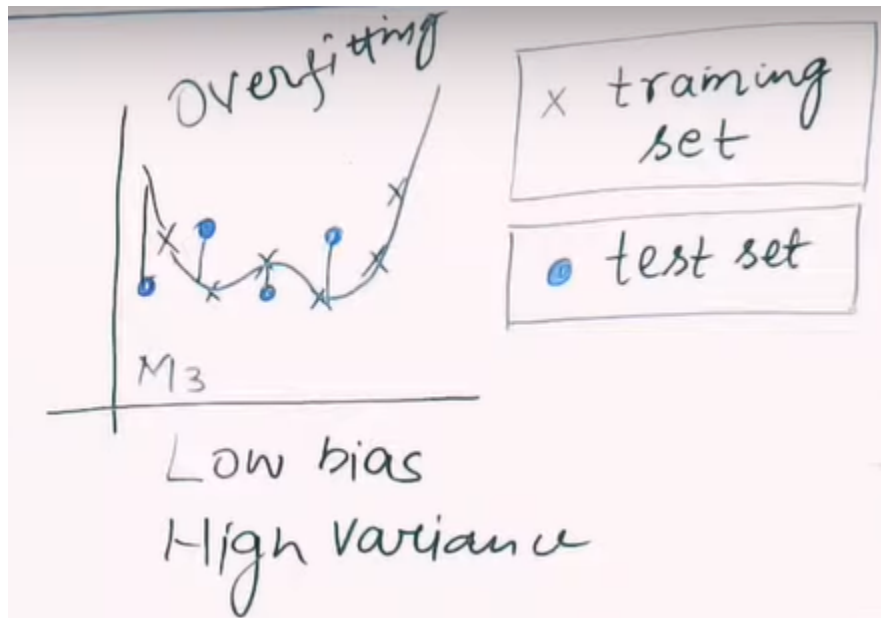
1. When the model performed well (by capturing the pattern of the training data very closely) in training dataset, the model performs less in new(testing) dataset which it never saw and that's when the **variance(the difference of error in Training and Testing dataset of a Model)** becomes high.



2. But when the model performed almost well or less well in the training dataset but tried to understand the training data as much as possible, the model performs relatively well in new(testing) dataset which it never saw and that's when the variance becomes less.

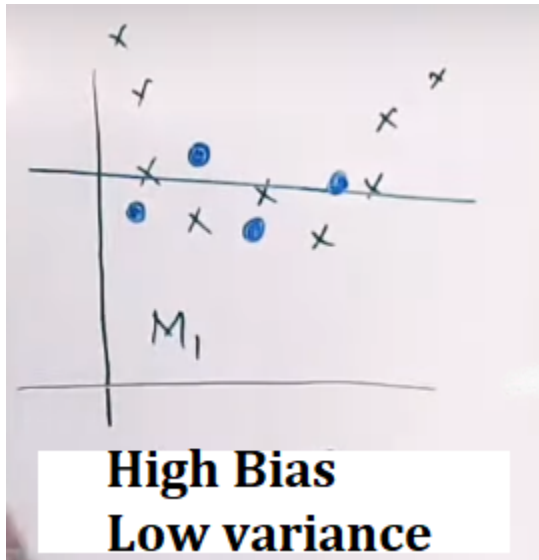


Overfitting



When the model captures the whole pattern of the entire dataset i.e. fitting the training dataset without a single thing, that's when the model performs very less in new (testing) dataset which it never saw before

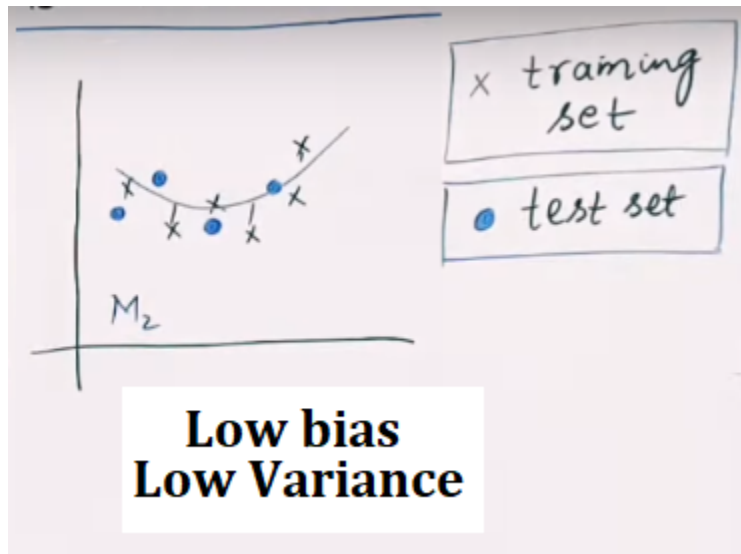
Underfitting



When the model couldn't understand the pattern of the training dataset, how can it perform well in a new (testing) dataset which it has never seen before?

Final Decision

As you can see we want a model which performs well in training dataset i.e. understand the pattern of the training dataset and also performs well in a new(testing) dataset which is M2 Model :



How to achieve the M2 Model? By 3 ways :

- 1) Regularization
- 2) Bagging
- 3) Boosting