

# Bias/Variance



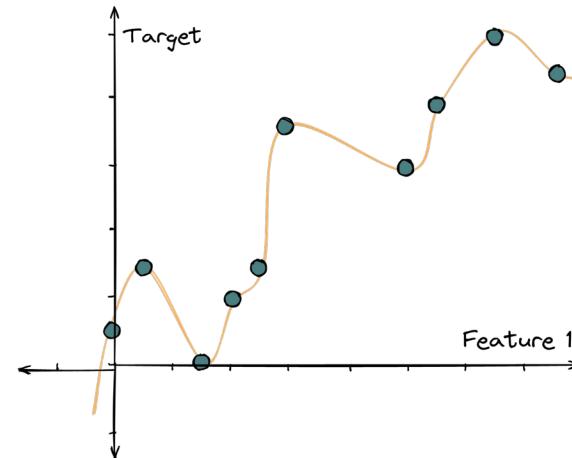
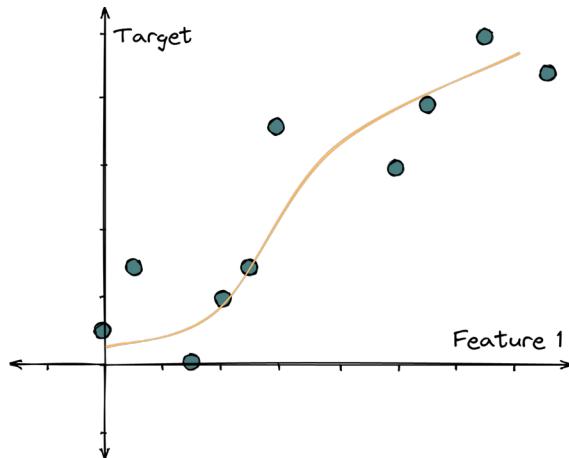
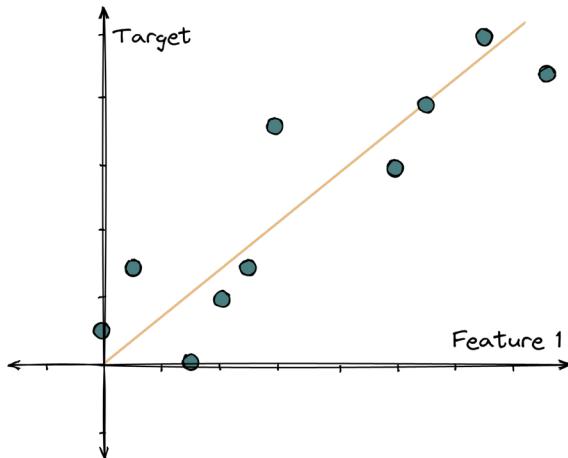
2/5



4/5



5/5



**Question:** For our training data, select the curve you find as best descriptive?  
Assume the number of stars as a quality metric.

# Bias/Variance



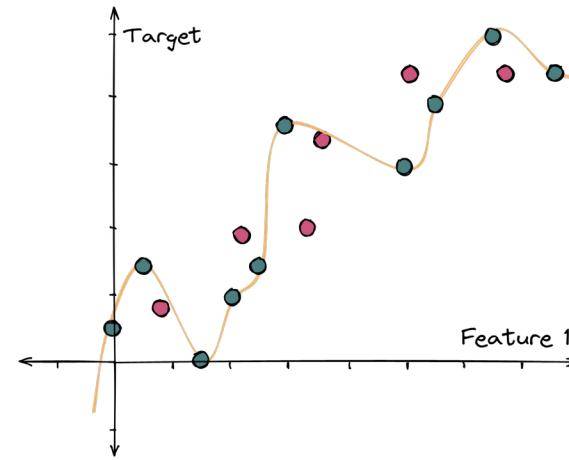
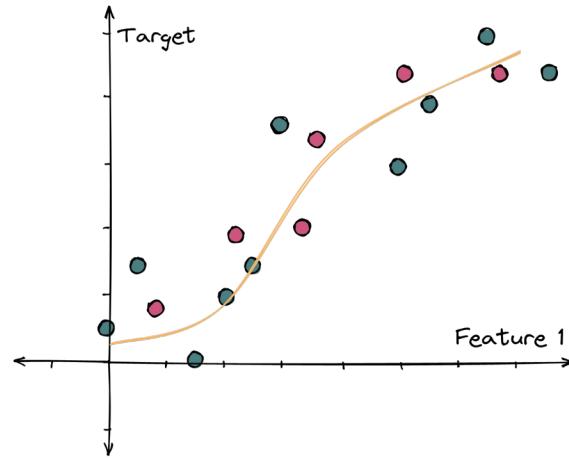
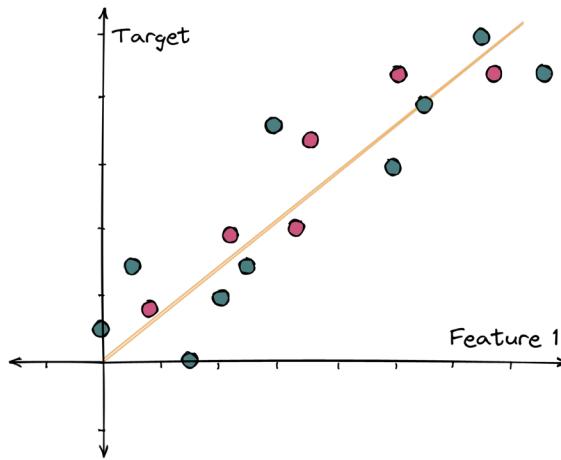
2/5



4/5

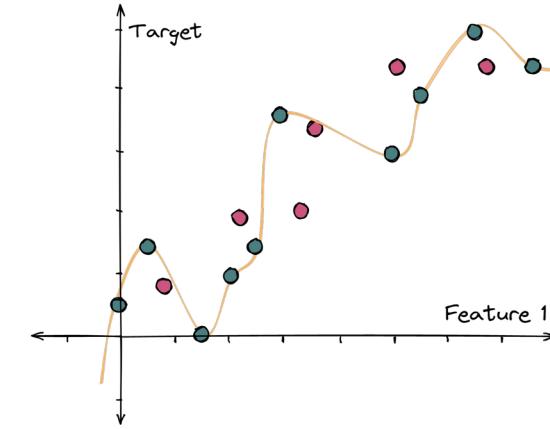
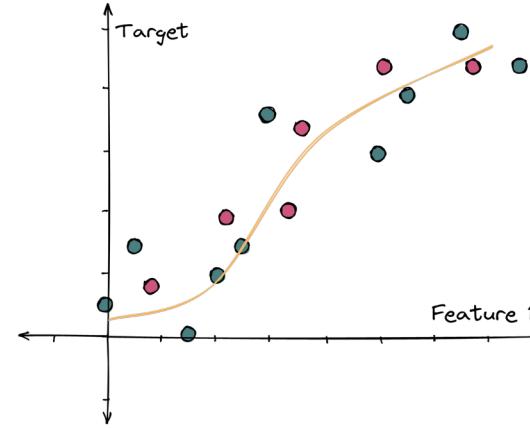
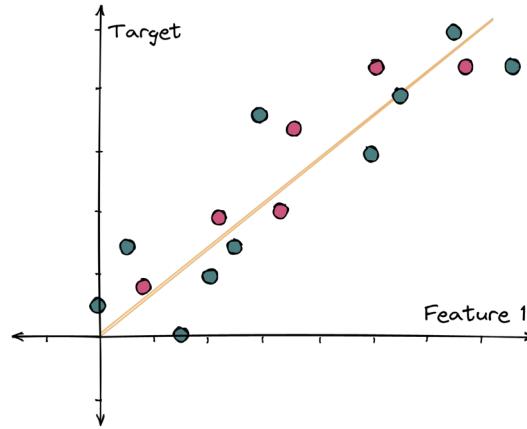
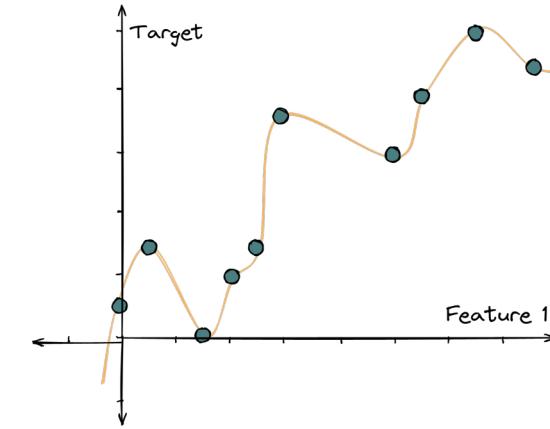
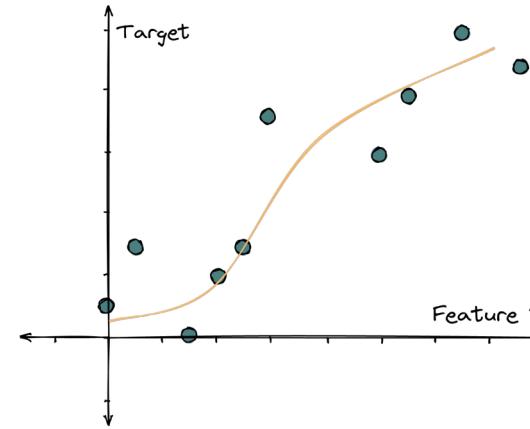
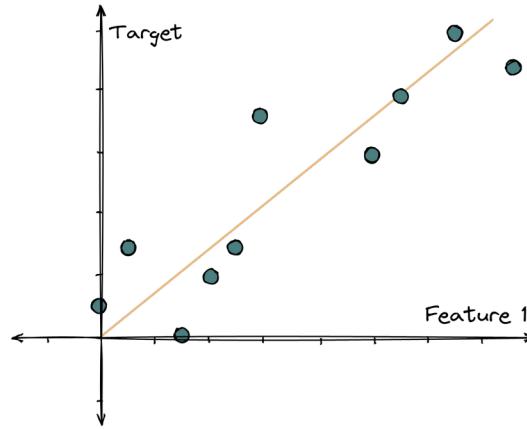


3/5

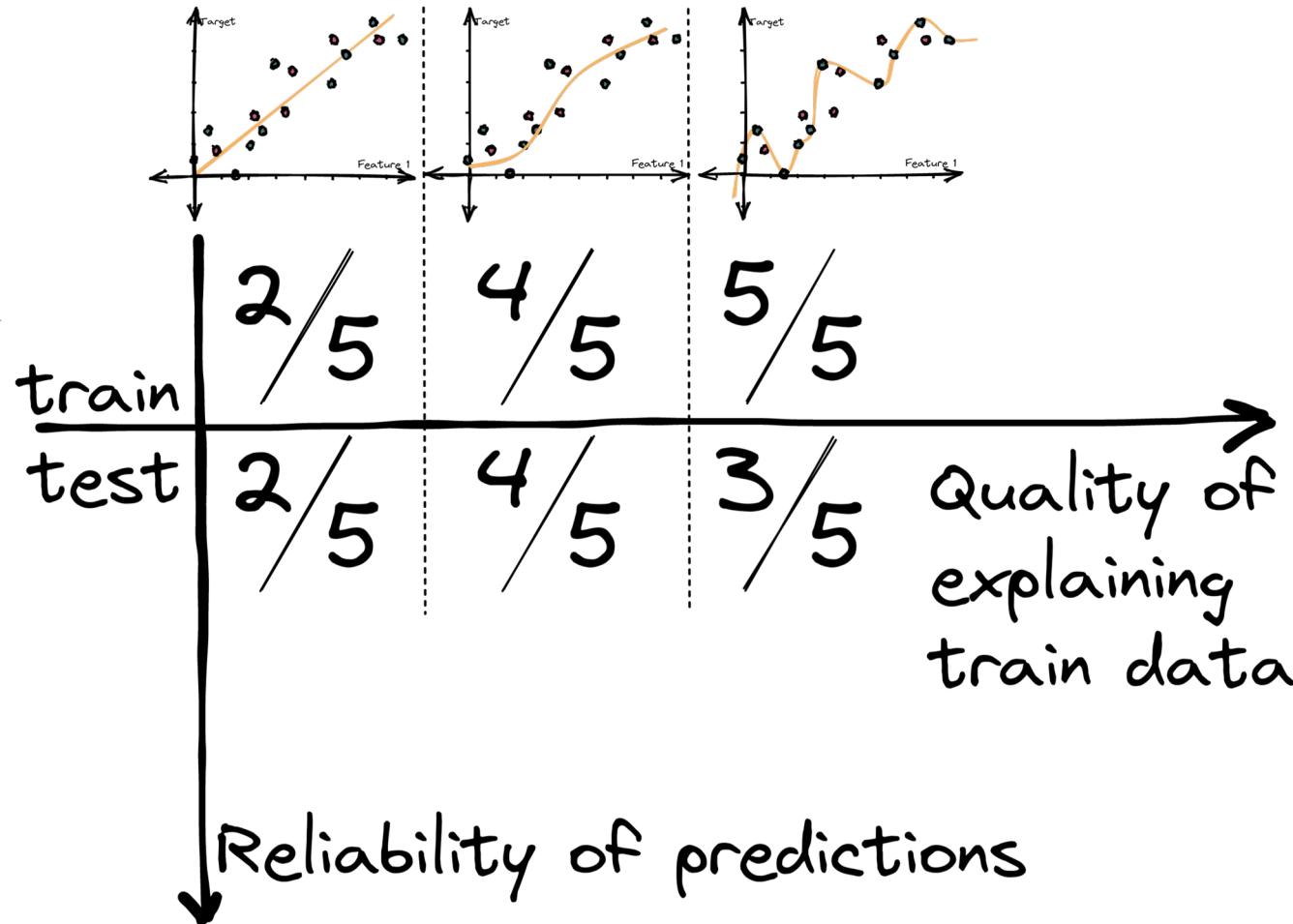


Now we add the test data.

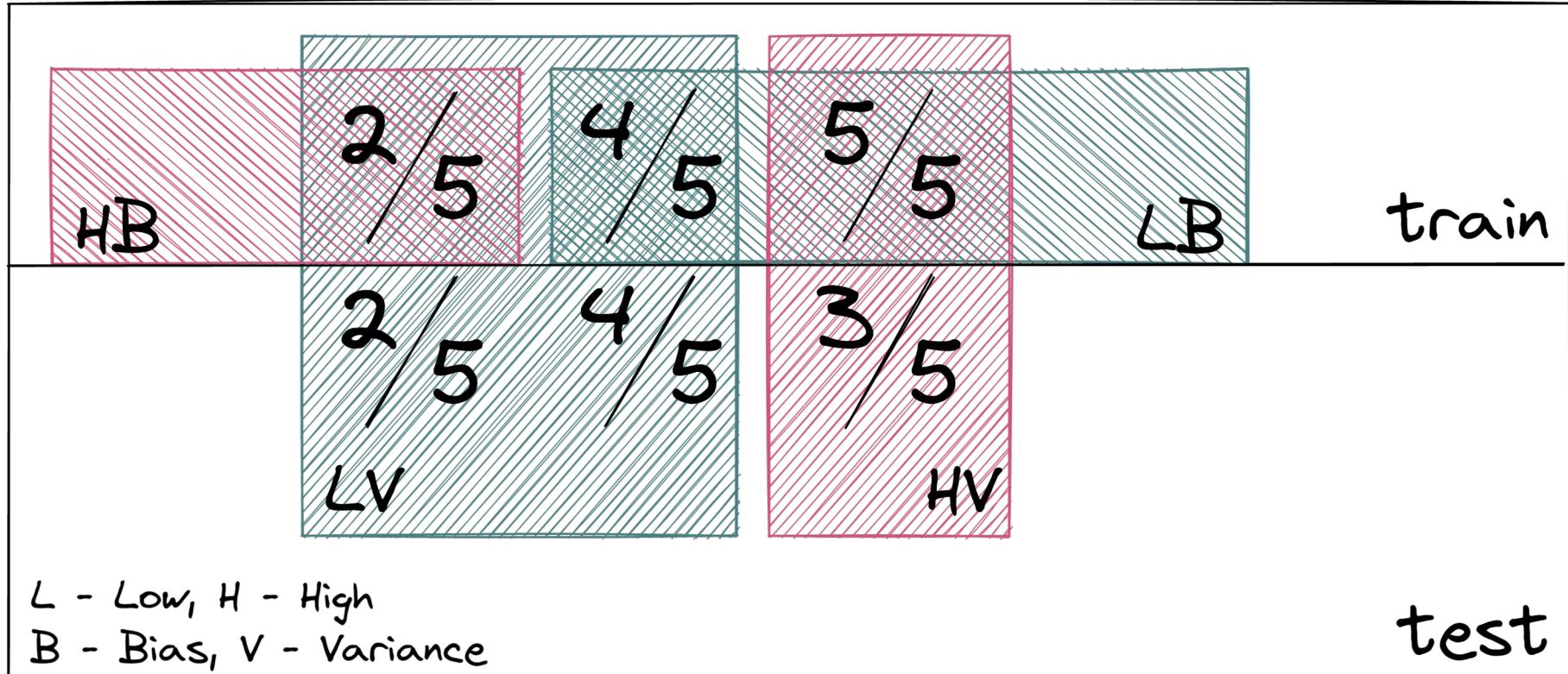
**Question:** How good is the selected model at describing test data?



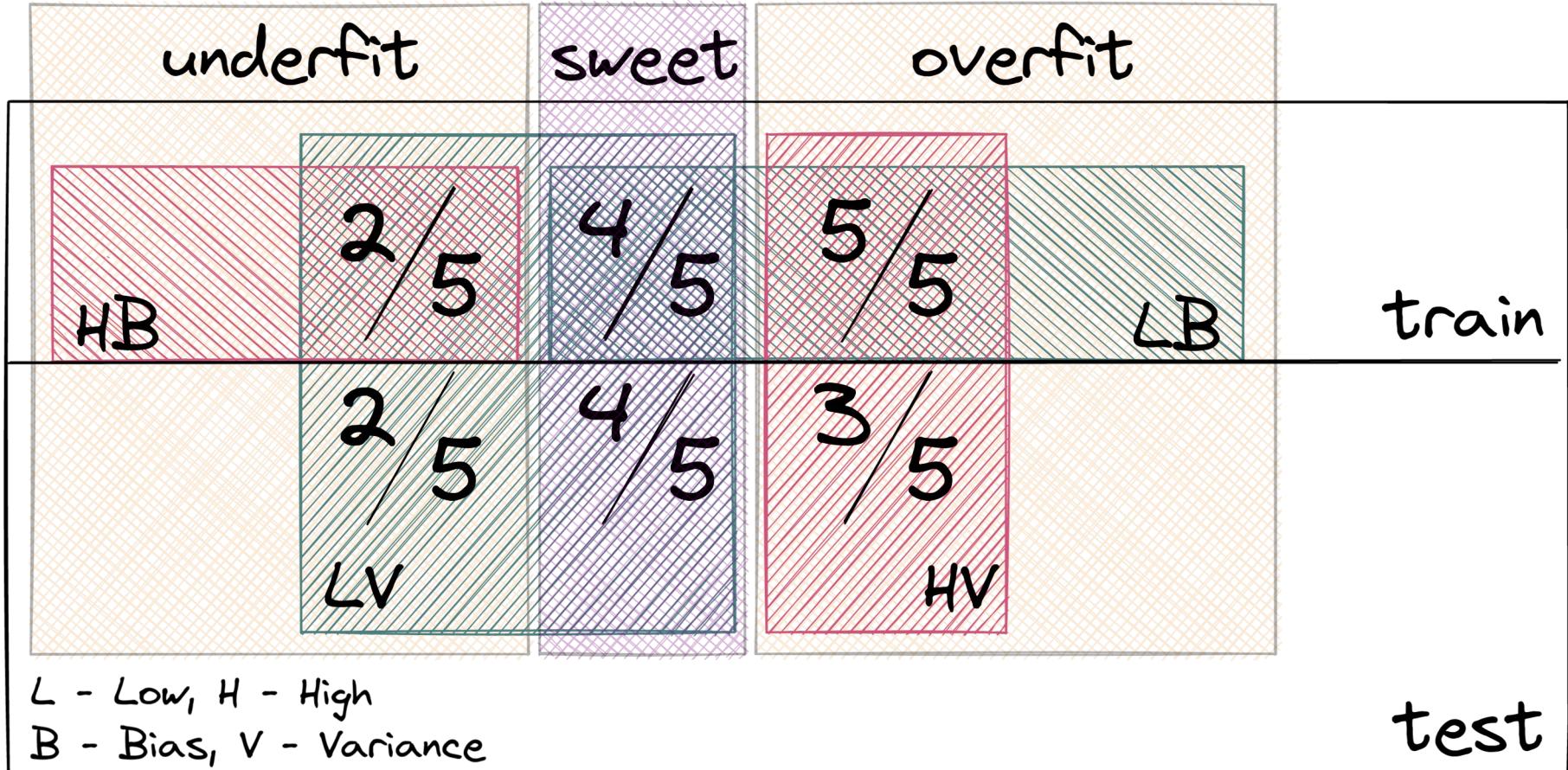
# Bias/Variance



# Bias/Variance



# Bias/Variance



# Bias/Variance

**Bias** is responsible for the quality of the model. It is how well you can describe your training data.

**Variance** is responsible for the reproducibility of your model on the test dataset.

In supervised ML, real bias/variance terms can't be determined.

Yet, the trade-off helps analyze ML algorithms and control their predictive performance.

