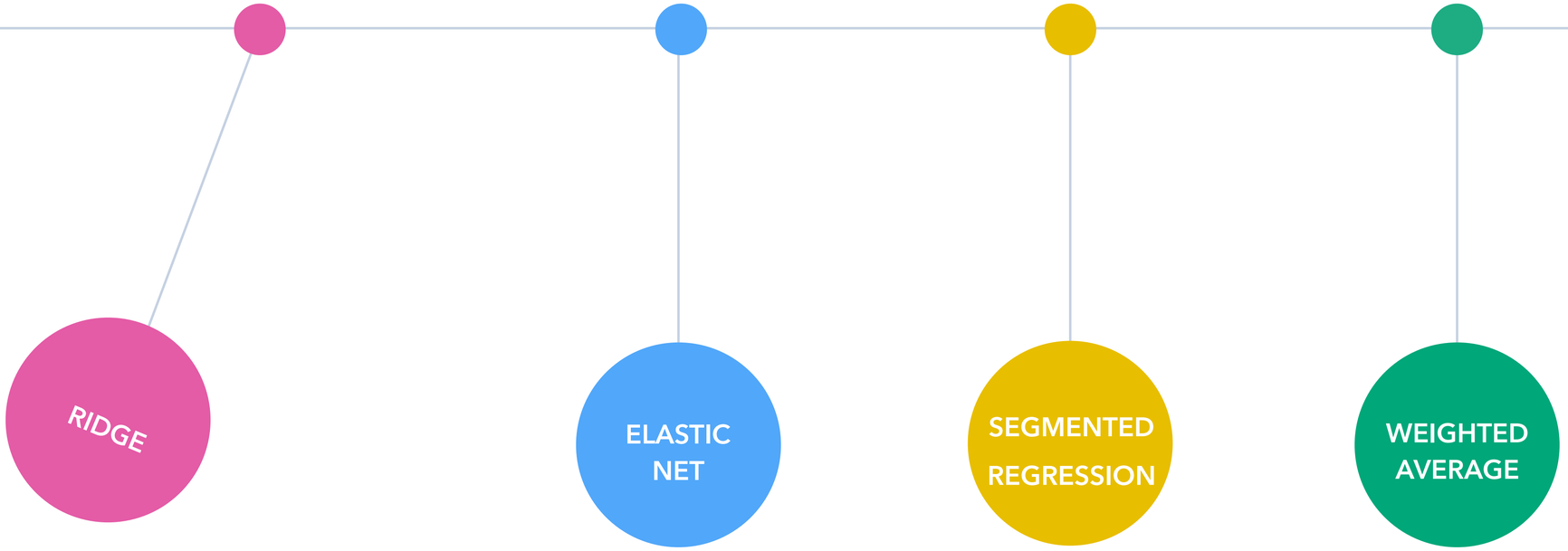


RIDGE BENCHMARK

We start with the ridge regression benchmark.

ELASTIC NET

Our first approach is replacing ridge with elastic net.

SEGMENTED REGRESSION

We improve our model with segmented regression.

WEIGHTED AVERAGE

The final model is taking a weighted average over Y.

BENCHMARK RIDGE

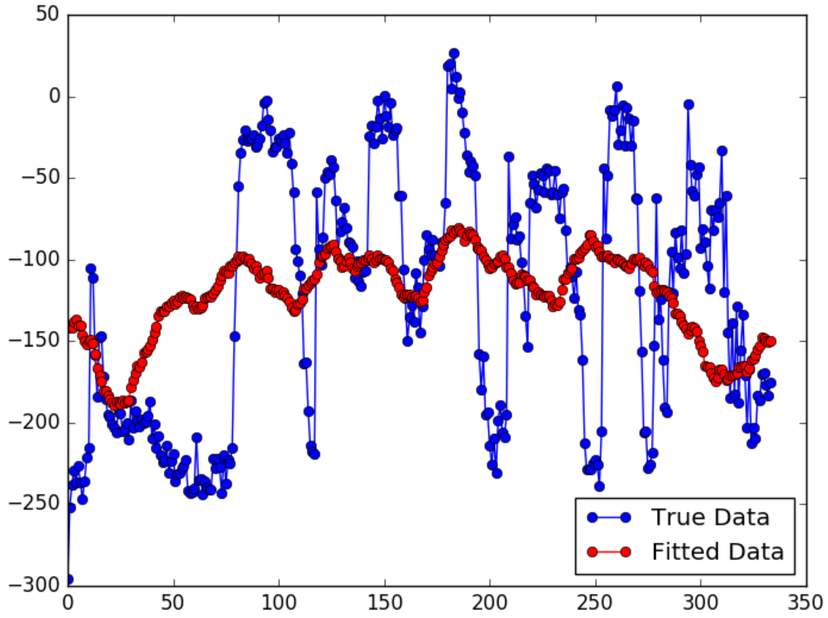
15.78168

OBSERVATION:

01 The data is made up of many sentences. The starting points and ending points are given as break points. We should take each sentence as an unit and fit the sentences.

02 For better visualization, we apply PCA on Y data and plot PC1 of Y.

03 To test our method, we use a 5-fold cross validation. In the plot, the fitted data is generated in the cross validation.



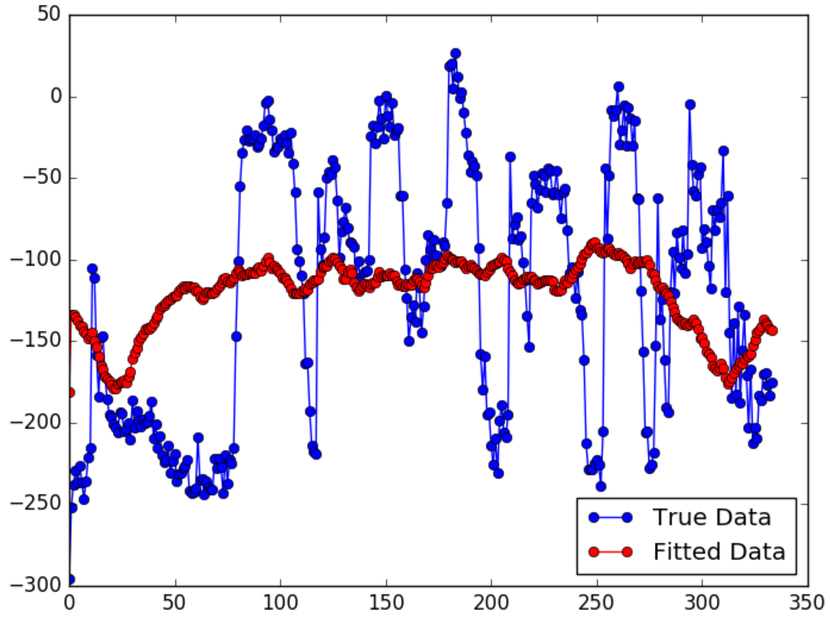
PC1 of 1st sentence in training data,
and its corresponding Ridge fitting.

OBSERVATION & IMPROVEMENT:

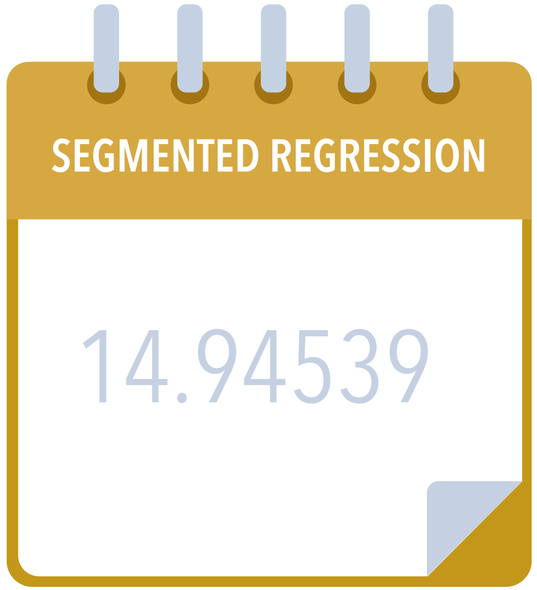
01 We first replace Ridge with Elastic-net. The additional Lasso penalty selects only the important features of X.

02 Non-linearity is simply added in the linear regression to improve the results.

03 PCA is applied to Y in the fitting, and parameters are tuned separately for each PC of Y.



PC1 of 1st sentence in training data, and its corresponding elastic net fitting.

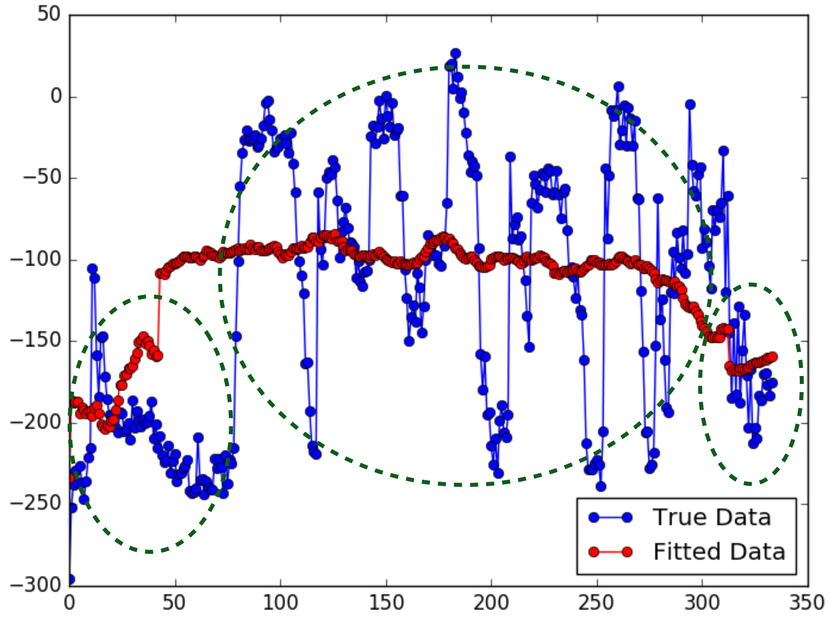


OBSERVATION & IMPROVEMENT:

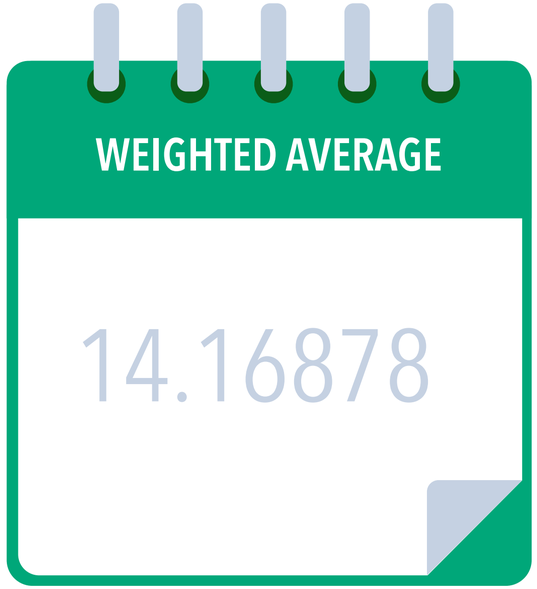
01 We notice that in each sentence Y can be classified into three groups: the beginning, the middle and the end. Therefore, we apply segmented regression to the model.

02 We apply the elastic net to the beginning, the middle and the end of sentences separately.

03 The beginning of each sentence may correspond to the period when voice is low or muted.



PC1 of 1st sentence in training data, and its corresponding segmented regression.

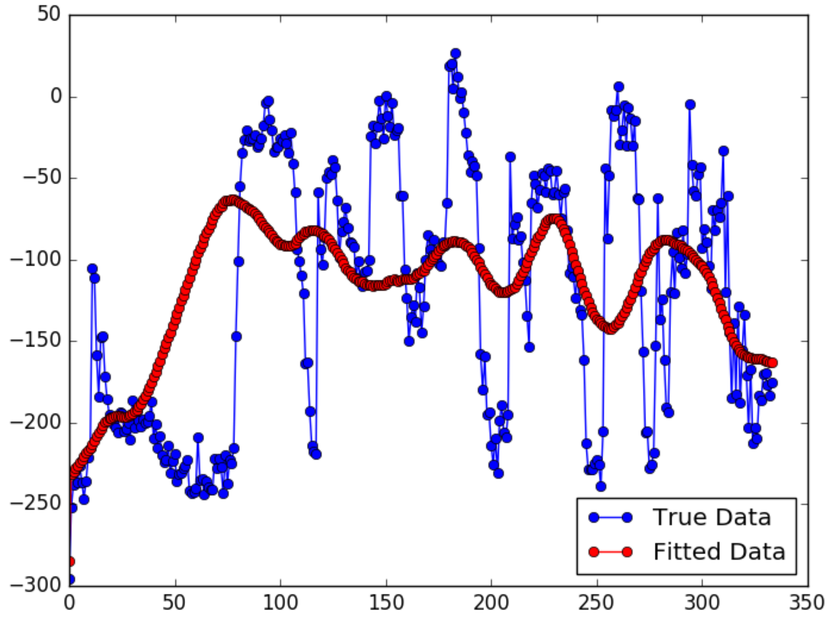


OBSERVATION & IMPROVEMENT:

01 We now let the information of Y dominate the prediction. For each time point in a testing sentence, we take all the corresponding time points in the training sentences and compute a weighted average of Y.

02 The weight is given by the correlation on some projection of X.

03 This method works because it makes our prediction more like a “sentence” rather than a simple function of X.



PC1 of 1st sentence in training data, and its corresponding weighted average result.