Language development in ASD, part 3

By Lena, Sebastian, Lisa, Astrid & Frederik

1. **How did your models from last time perform? In this exercise you have to compare the results on the training data () and on the test data. Report both of them. Compare them. Discuss why they are different.**

The models

We tested out three different models:

1) The simple model: A mixed effect model predicting CHI_MLU from Visit-diagnosis interaction, and with random effects consisting subject for random intercept and visits (in other words, the effect of time) as random slope.
2) The quadratic model: As the simple model, with the exception that visits is squared
3) The complex model: Our complex model ended up being the quadratic model plus Verbal IQ and amount of different unique words (types_CHI).

   We found very high correlation (types_CHI and token_CHI), so it wouldn't make sense to use them both as predictors as they would capture the variance. We picked types_CHI, because we assume that it isn't as dependent on the mood of the kid on the day of visit, as the amount of words.

The simple model

The prediction power of the simple model on the training data as measured in root-mean square error is 0.35; when tested on the test (novel) data the mean error is 0.986.

The quadratic model

Testing the quadratic model's prediction power on the training data resulted in a mean error (root-mean square error) of 0.32; as for the test data, this number is 0.991.

The complex model

The prediction power of the complex model on the training data is 0.30, and for the test data 0.980.

The prediction power of the three models were quite similar – for the training data as well as for the test data. Generally, however, they are not as accurate on the test data as on the training data. This is expected. This difference can be explained in part by the random effects: These effects are

specific to the subjects of the training data, and the explanatory power of these are lost when transferring the model to the novel test data set.

The other thing accounting for the poorer prediction power of the models on the test data, is that we have overfitted the model in our hunt for significant results for the training data.

2. Which model would you select based on cross-validation: the basic model or your fancy one? Explain. Then try to identify the best possible model to predict CHI_MLU using cross-validation.

The three models we created last week performed about equally well.

When based on cross-validation, however, the results are different. The prediction power is improved for all three models, but the simple model can no longer keep up with the other two models: The simple model had a 0.74 error on the test-data, the quadratic 0.44 error, and the complex 0.45. While the predictions of the test-data improved, the predictions of the training-data improved.

The quadratic model and the complex model performed equally well, as based on the rmse-scores. Following the principle of Ochard's Razor, we follow the simpler of the two models, the quadratic model.

In our attempt of improving the model, we kept the framework of the quadratic model, but with a few more fixed effects: Diagnosis-visit-interaction as primary effect, followed by gender, CHI_LUstd,s unique words, verbal IQ.

This model resulted in the prediction power of 0.34 mean error on the test-data, and 0.22 mean error for training-set.

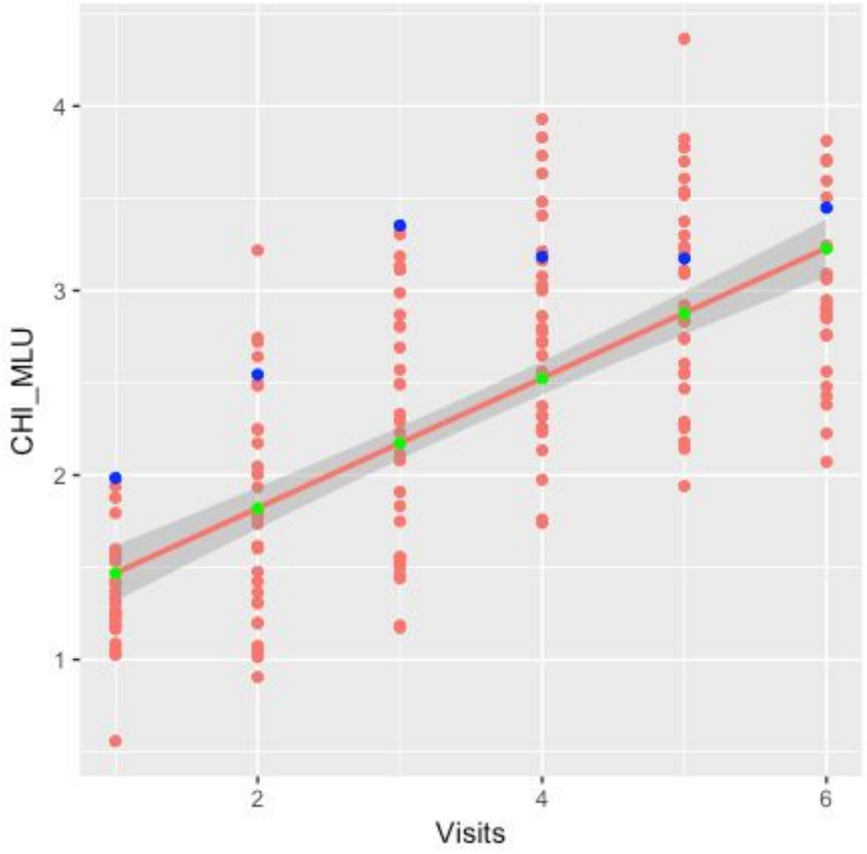
3. How does Bernie fare in ChildMLU compared to the average TD child at each visit? Define the distance in terms of absolute difference between this Child and the average TD.

Visits	Bernie - mean(CHI_MLU$Visits(n))
1	0.663291
2	0.7598606
3	1.093328
4	0.4074076
5	0.1770933

6	0.5320438

Bernie lies above the TD (normal healthy) children's average language capabilities (as measured in mean length of utterance) on all six of his visits.

How does Bernie fare compared to the model predictions at Visit 6? Is the child below or above expectations?



Plot 1: Red = TD children's data points, Blue = Bernie's actual performance, Green = Model's estimate

Bernie performs 0.22 CHI_MLU scores better on his 6th Visit than the model predicted. However, Bernie's performance at visit 1 was also higher than the model's estimate.

Find us on
https://gitlab.com/superpowerLisa/almost_online

https://github.com/coolasacucumber/upgraded-lamp/blob/master/3_LangASD_prediction_seb.Rmd