Daniel, Jesper, Pernille KJ og Astrid

23. September 2020

Portfolio 3

Language development in ASD - making predictions

Here is the link to our full code

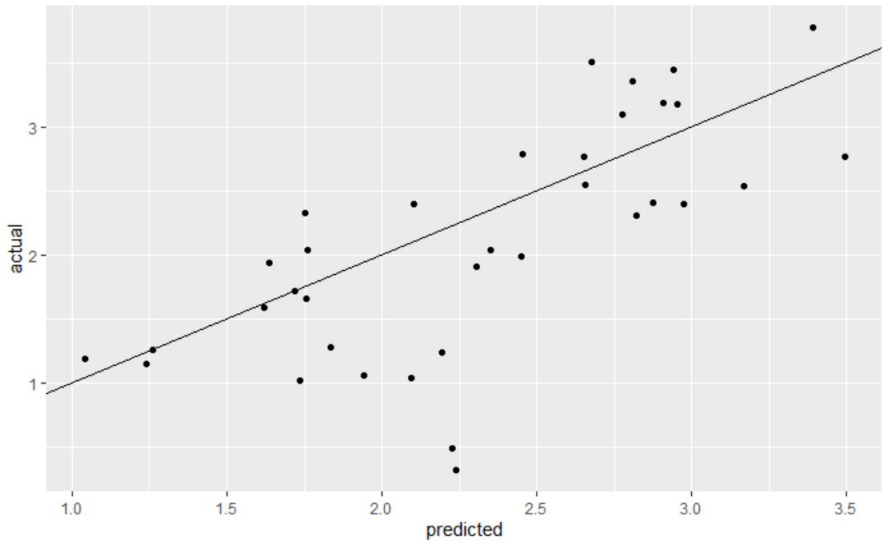
Discuss the differences in performance of your model in training and testing data

Our model:
lmer(CHI_MLU ~ poly(Visit,2) * Diagnosis +verbalIQ1 +(1|Child.ID), data = all)

To evaluate the performance of our model on the training data we calculated the root mean square. The root mean square of the model on the training data is 0.4. Since the children's lengths of utterance are generally quite small, this is RMSE does not indicate a great performance of the model on the training data. However, it is even higher for the test data (1.1).

It makes more sense to evaluate a model using the test data, this makes for a more generalizable model.

Here's a plot showing the predicted values of our model (built on the training data) compared to the observed (actual) value of the test data.



This makes sense, since the model has been built on the training data it more closely follows this data.

Which individual differences should be included in a model that maximizes your ability to explain/predict new data?

When we compare our old model to the base model we get two sets of RMSE values. For the base value the RMSE on the test data has a mean of 0.77 and for our model: 0.52. Therefore, our model has a better fit. This is because the predicted values vary the least from the observed values.

Our model includes diagnosis and verbal IQ. To find the best possible model using any combination of the measures of individual difference, we would repeat the steps done for the base model and our model, looking at all theorietcally justified combinations possible of variables and select the model with the lowest RMSE.

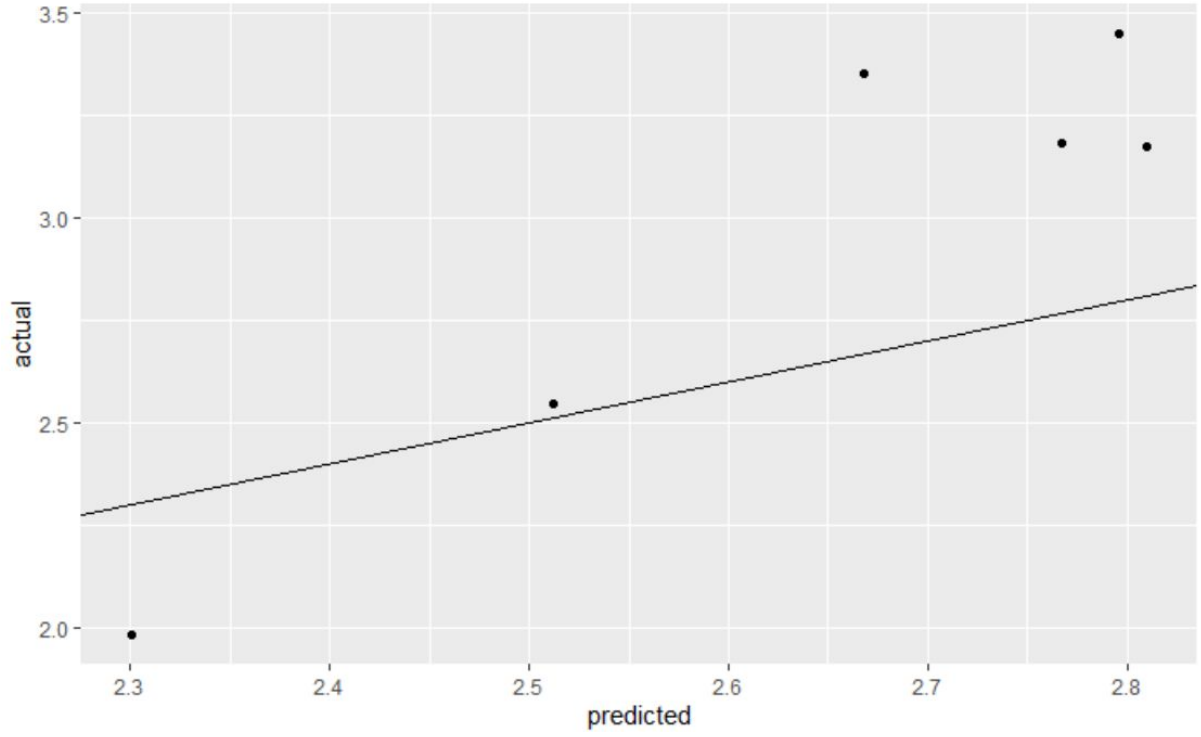
Predict a new kid's performance (let's call him Bernie) and discuss it against the expected performance of the two groups

Here's a table showing Bernies MLU at visits 1 through 6 compared to that of the average child. We also included a column showing the abolsute difference between Bernie and the average child.

	Diagnosis	Visit	TDmean	meanbernie	abs
1	TD	1	1.311766	1.984456	0.6726900
2	TD	2	1.761914	2.544444	0.7825305
3	TD	3	2.226604	3.353191	1.1265878
4	TD	4	2.731307	3.183099	0.4517913
5	TD	5	2.992527	3.173252	0.1807254
6	TD	6	2.910988	3.448413	0.5374250

Bernie does better than the average child in all visits.

We also used our model to predict how well a child with the predispositions of Bernie is expected to do. Bernie did better than our model predicted We can see this in the plot below. The plot compares the predicted perfomance of Berine to his observed (actual) perfomance.



Daniel, Jesper, Pernille KJ og Astrid

23. September 2020

Who did what?

We all worked on the code on our individual computers in different states of mess. This time we used Jespers code for the hand-in. We all discussed the answers for the questions in person and on chat. Astrid wrote the sum-up of the discussion in this document and reported the measures from Jespers code.