AE Section

Ruben Heuer

5/27/2021

Autoencoders

Method

To further validate our results, we conducted a second analysis over the same dataspace using autoencoders for dimensionality reduction. To this end, we first trained an autoencoder with two hidden layers and 8 nodes each on a random subset consisting of 80% of observations of the ANES feeling-thermometer data. We then compared the overall distribution and degree of separation between training and test set. The autoencoder was trained over 1000 epochs using a hyperbolic tangent activation function. Subsequently we conducted our analysis on the 2 dimensions of the autoencoder's 2-node output layer. We chose our neural network architecture on the assumption that the feeling-thermometer space is likely hierarchically structured. The layered architecture accordingly allowed us to examine different levels of abstraction in order to infer candidate ideological dimensions underlying the data.

To numerically guide our exploration of the autoencoder embeddings, we deployed association rule mining on the media space. In this way, we obtained promising combinations of media items allowing us to extend our analysis beyond the inspection of single formats.

Initial Analysis

Inspecting relative feature importance, we see that feelings towards certain politicians play the most significant role in determining the structure of the autoencoder space. Among these are former presidents Obama and Trump. In addition, feelings towards the Republican and Democrat party as well as the Black Lives Matter movement seem to explain a large portion of the variation in the reduced dataspace.

Just as in the plots obtained in the LLE analysis, we can see similar degrees of separation in terms of ideology as well as party affiliation with moderates again slightly more dispersed than liberals and conservatives and an analogous tendency regarding the partisan aspect.

Comparing training and testing set, we obtained a similar picture with respect to the spatial distribution of ideology and party, lending more face validity to our previous results.

Combination of Media Items

In a next step, we conducted association rule mining on the media space to guide our exploration of the autoencoder embeddings, restricting the consequent of each rule to be either party affiliation or self-reported ideology associated with each observation. Within the restricted set of rules, we inspected those with a lift value higher than 1.5. for each ideology and party affiliation indicating a marked improvement over a random choice model.A set of high-confidence association rules extracted in this way for a conservative consequent takes the following shape:

Interestingly, a search over moderate consequents yields less useful rules, indicated by considerably lower lift values:

In our analysis we used association rules mainly to provide us with interesting but unintuitive combinations of consumed media associated with a given ideology or party affiliation. This idea is illustrated in the figure

Relative Feature Importance
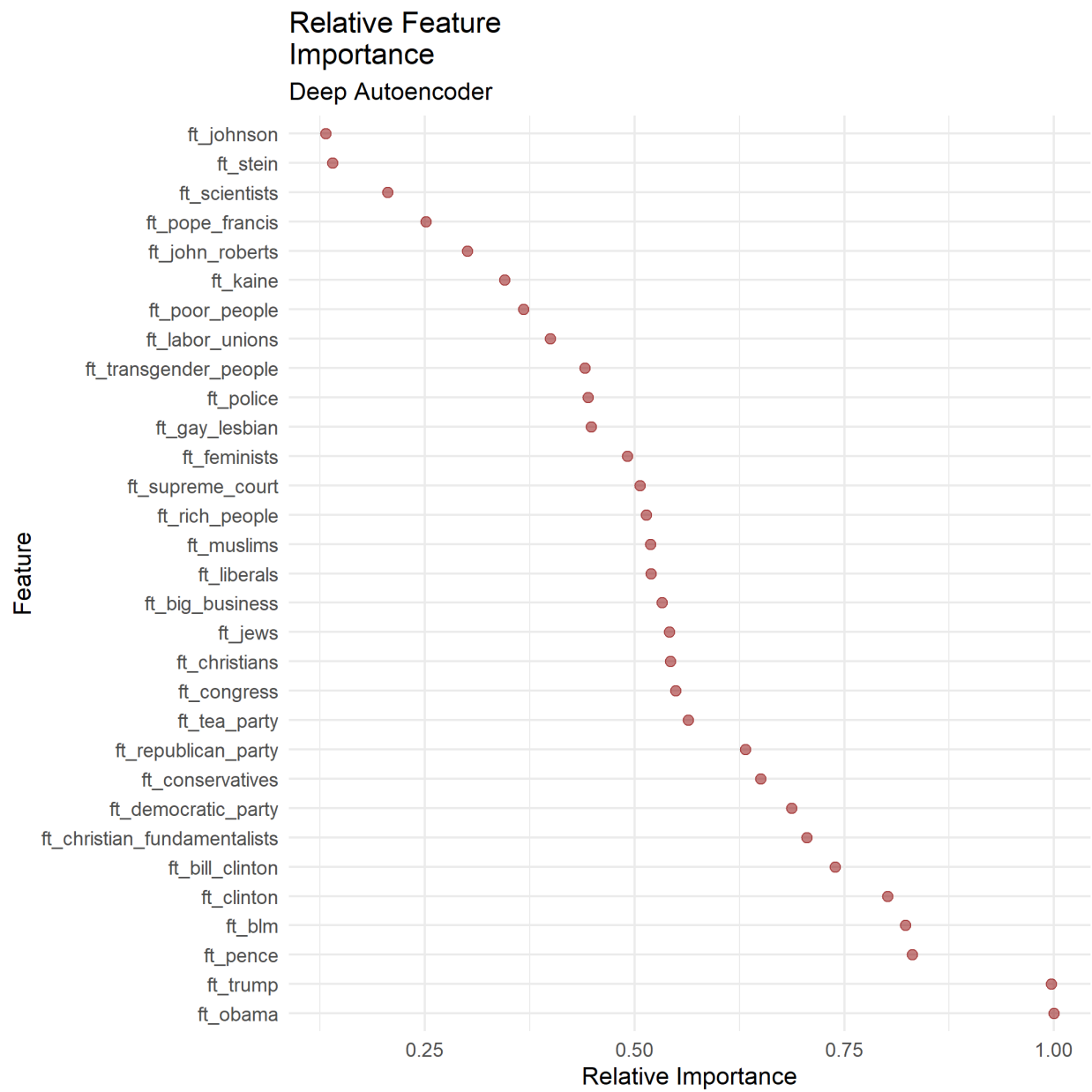
Deep Autoencoder

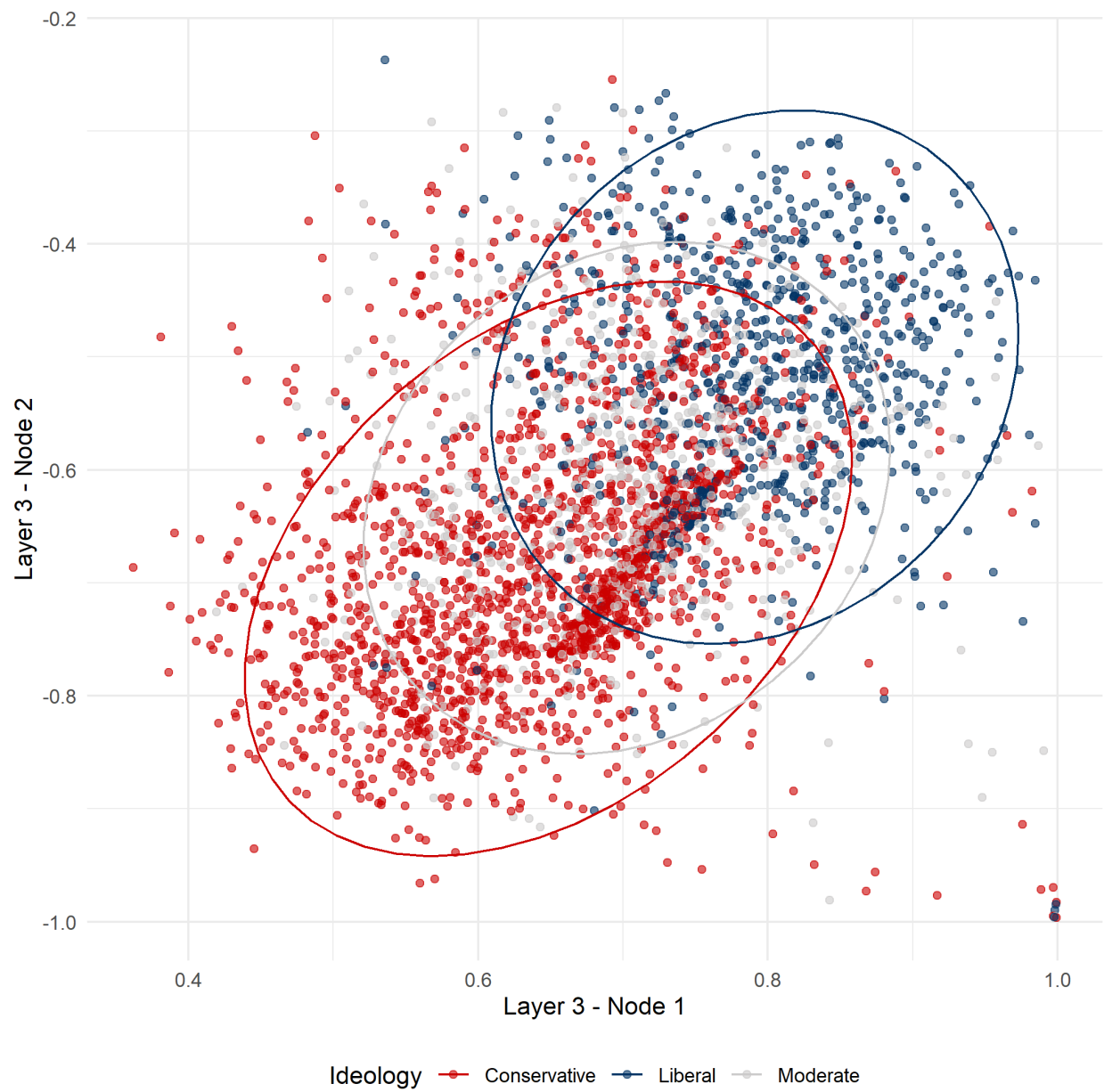Figure 1: Relative feature importance of feelig-thermometer data.

Figure 2: Ideological separation in the training set.
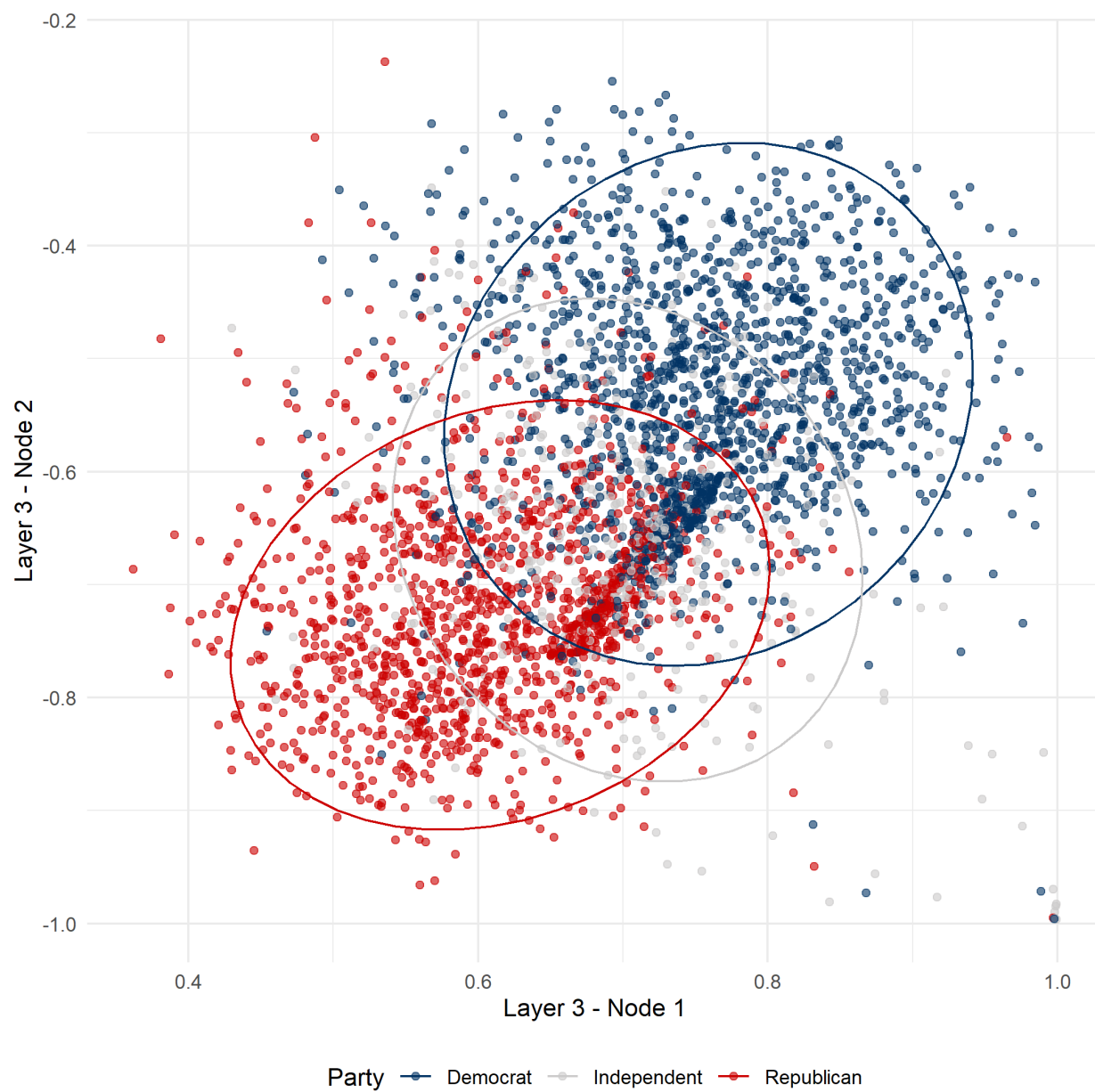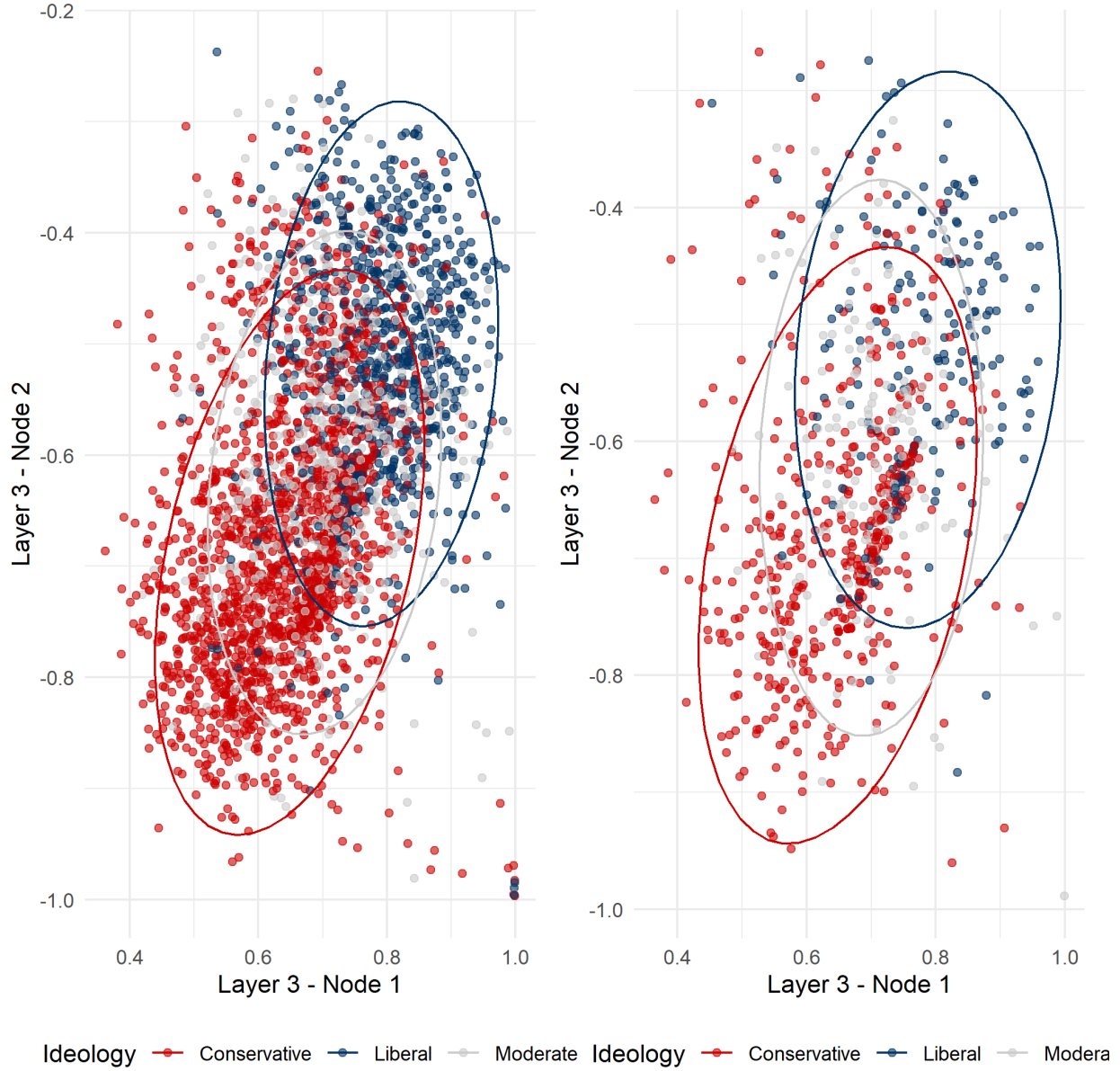
Figure 3: Party separation in the training set.

Figure 4: Ideological separation in training and test set.

Association Rules with RHS = Conservative

Rules	Conf.	Cov.	Lift	Count
{tv_hannity,tv_sunday_night_football,tv_on_the_record_with_greta_van_susteren} => {ideologyConservative}	0.92	0.02	1.68	69
{radio_the_glenn_beck_program,tv_the_o_reilly_factor} => {ideologyConservative}	0.92	0.01	1.67	57
{tv_hannity,tv_the_kelly_file,tv_sunday_night_football} => {ideologyConservative}	0.91	0.02	1.66	91
{tv_hannity,tv_on_the_record_with_greta_van_susteren,tv_the_o_reilly_factor} => {ideologyConservative}	0.90	0.04	1.64	142
{tv_hannity,tv_the_kelly_file,tv_shark_tank} => {ideologyConservative}	0.90	0.01	1.63	52
{tv_hannity,tv_dancing_with_the_stars} => {ideologyConservative}	0.89	0.01	1.62	57
{tv_hannity,tv_shark_tank,tv_the_o_reilly_factor} => {ideologyConservative}	0.88	0.02	1.61	69
{radio_the_mark_levin_show,tv_the_o_reilly_factor} => {ideologyConservative}	0.88	0.02	1.61	61
{tv_hannity,tv_sunday_night_football,tv_the_o_reilly_factor} => {ideologyConservative}	0.88	0.03	1.61	119
{web_fox_news,tv_sunday_night_football,tv_the_o_reilly_factor} => {ideologyConservative}	0.88	0.02	1.60	74

Figure 5: Association rules with conservative ideology as RHS.

5

Association Rules with RHS = Moderate

Rules	Conf.	Cov.	Lift	Count
{tv_60_minutes,tv_the_voice,tv_good_morning_america} => {ideologyModerate}	0.35	0.04	1.68	56
{radio_no_radio,tv_60_minutes,tv_dancing_with_the_stars} => {ideologyModerate}	0.34	0.04	1.62	60
{tv_sunday_night_football,tv_60_minutes,tv_good_morning_america} => {ideologyModerate}	0.34	0.04	1.61	60
{radio_no_radio,tv_abc_world_news_with_david_muir,tv_dateline_nbc,tv_good_morning_america} => {ideologyModerate}	0.33	0.04	1.59	55

Figure 6: Association rules with moderate ideology as RHS.

below. Inspected in isolation, the viewership of the politically conservative TV show *Hannity* clearly clusters in the ideologically conservative portion of the autoencoder space while *NCIS*, at face value an unpolitical entertainment format, is spread over its entire range.

In response to relatively higher confidence values of the association rule *Hannity + NCIS -> Ideology: conservative* compared to *Hannity* alone as an antecedent, we once more surveyed the autoencoder space, this time coloring the intersection of *Hannity* and *NCIS* viewers. As expected, a number of outliers that cannot be unambiguously associated with the conservative portion of the autoencoder space cease to be part of the selection. This, however, comes at the cost of reducing the number of observations present in our visualization.

In order to probe the utility of intersections of different media items based on our intuitions about combinations of partisan media, we examined *Sean Hannity* and shows commonly associated with conservative ideology. As expected, the intersection of *Hannity*, *The O-Reilly Factor* and *On the Record with Greta Van Susteren* clearly picks out a cluster of observations in the conservative range of our dimension-reduced dataspace.

What is surprising, however, is the similarity between the combinations of three conservative shows and that of *Hannity* and a seemingly unpolitical entertainment format. It is possible that the two shows appeal to a subset of the population that
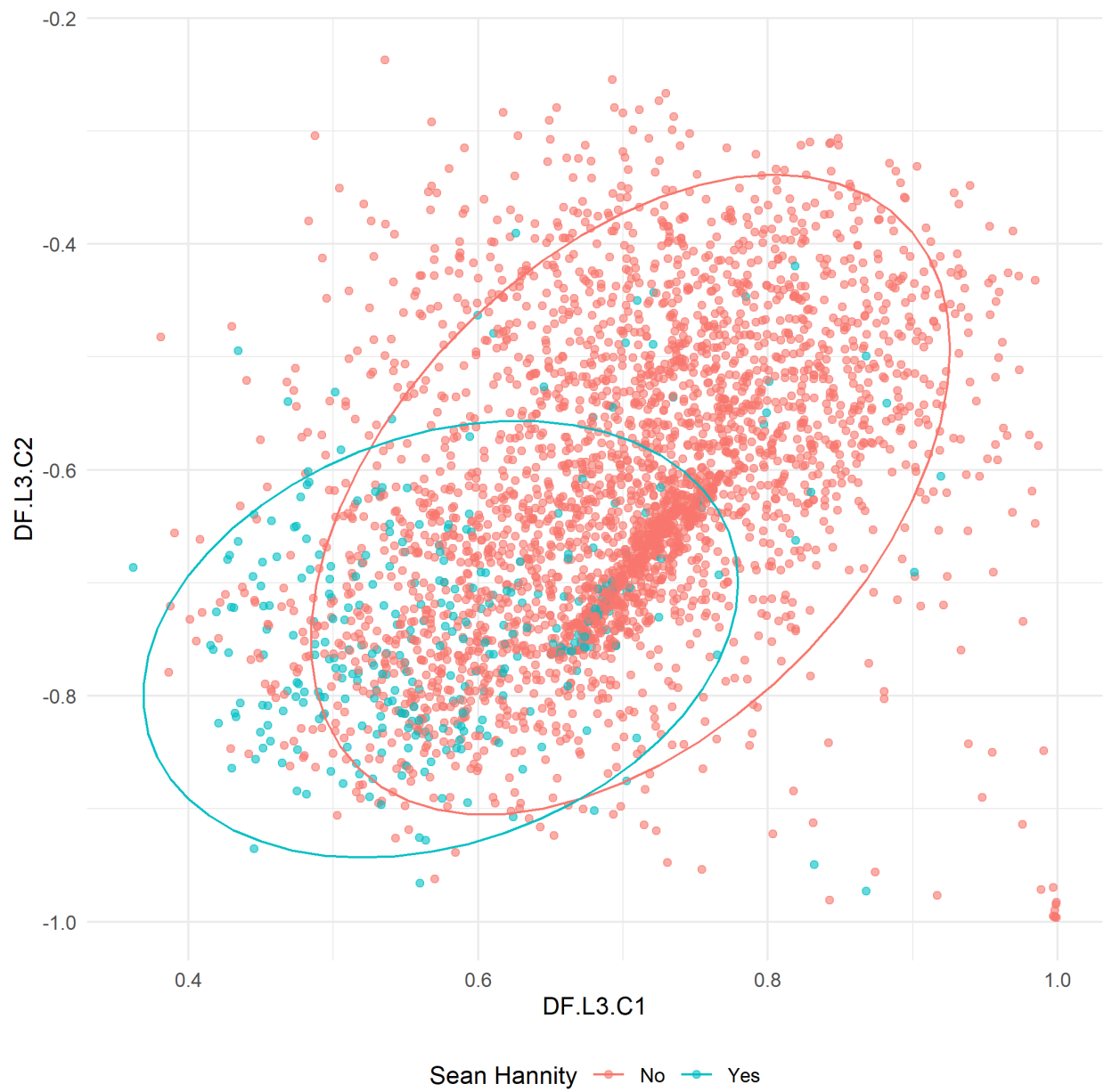
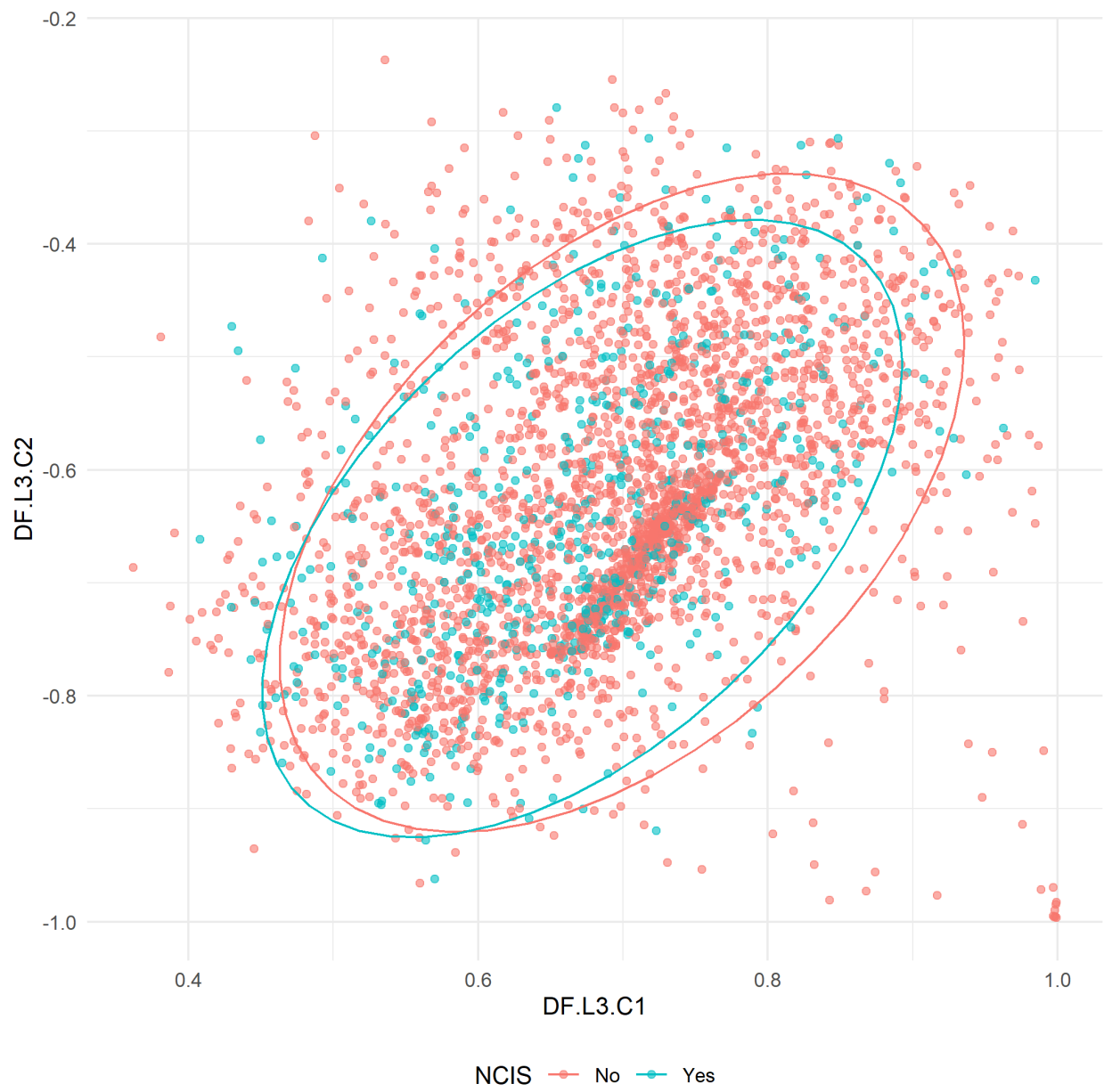Figure 7: Separation in TV shows watched: Sean Hannity.

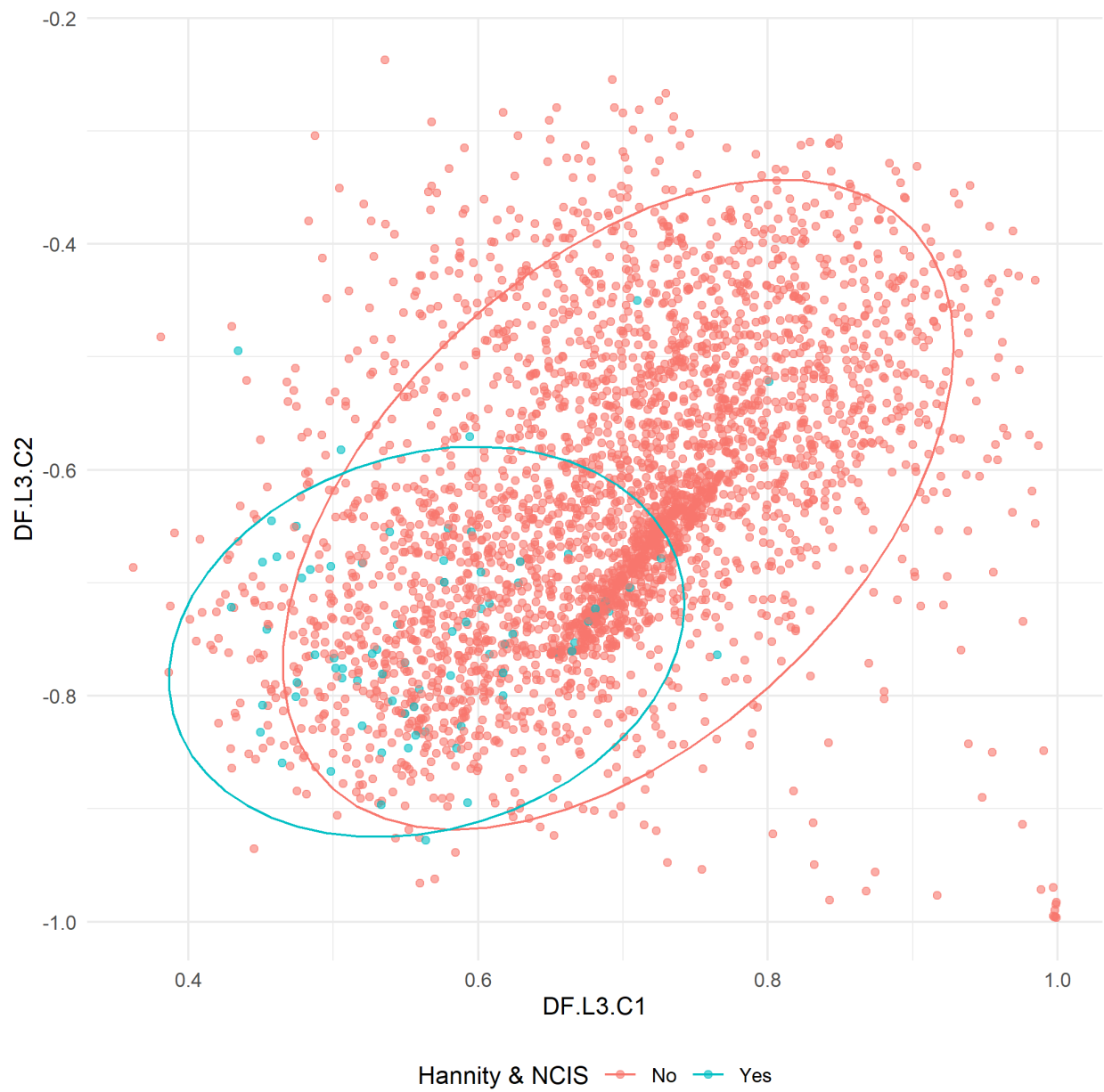Figure 8: Separation in TV shows watched: NCIS.

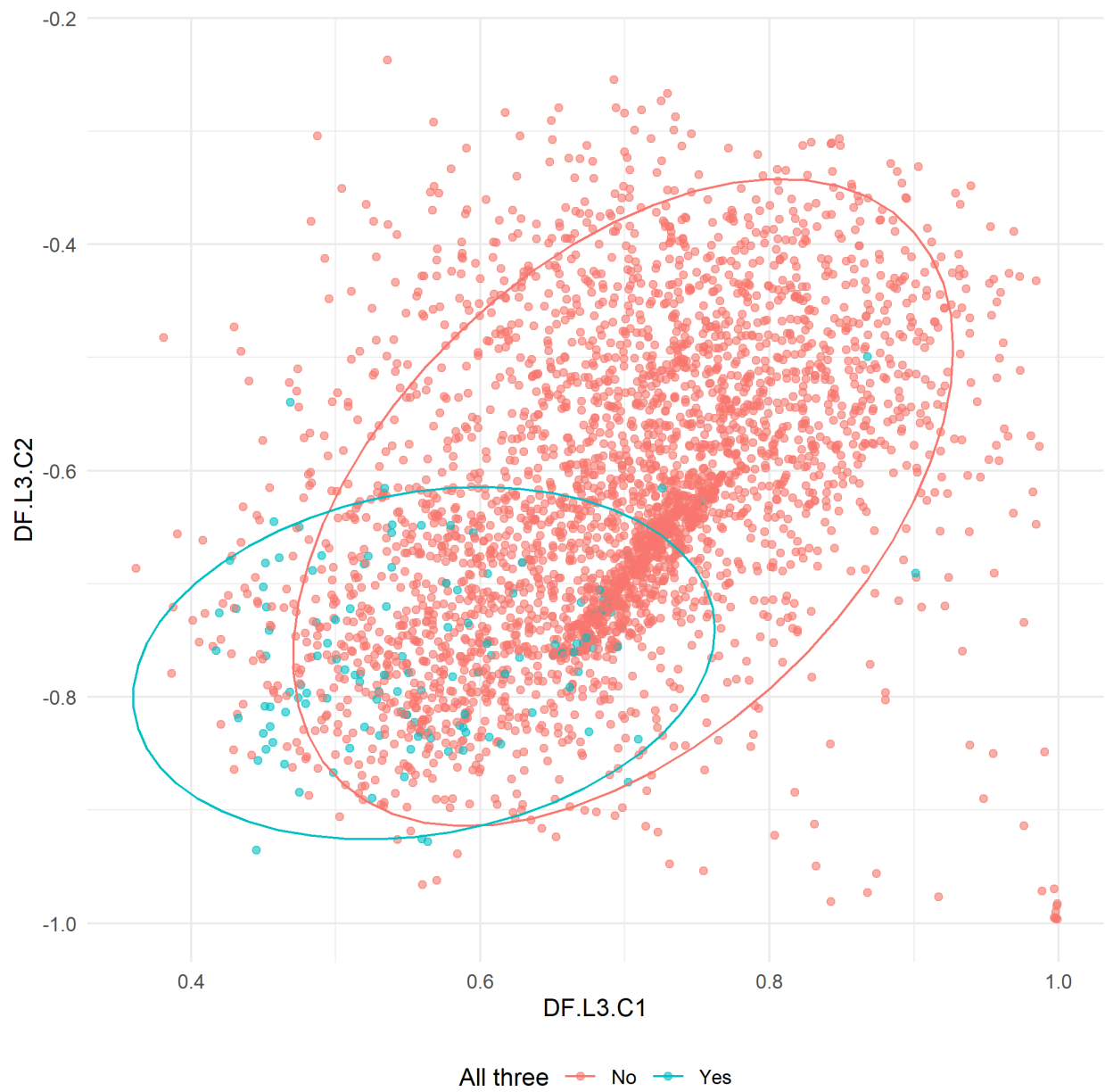Figure 9: Separation in TV shows watched: Hannity & NCIS.

Figure 10: Separation in TV shows watched: Hannity, The O'Reilly Factor & OTR with Greta Van Susteren.