

# Predicting Loneliness

Ian Shryock

## Contents

<b>Contents</b>	<b>1</b>
<b>List of Tables</b>	<b>2</b>
<b>1 Research Problem</b>	<b>3</b>
<b>2 Descriptions of the Data</b>	<b>3</b>
<b>3 The Models</b>	<b>5</b>
3.1 Logistic Regression . . . . .	5
3.2 Elastic Net . . . . .	5
3.3 Random Forest . . . . .	5
<b>4 Comparing Model Fits</b>	<b>6</b>
<b>5 Plots</b>	<b>8</b>
5.1 Model Fits . . . . .	8
5.2 Comparing Coefficients/Importance . . . . .	14
<b>6 Discussion</b>	<b>15</b>

## List of Tables

1	Item Descriptives . . . . .	3
2	Model Fit Statistics . . . . .	6

# 1 Research Problem

In this project, I follow up on the analyses conducted by Beck and Jackson (2022) and create nomothetic models for predicting loneliness. A succinct phrasing of the overarching research question is: do models derived from populations or from a single individual provide better predictions of individual behaviors and feelings?

Positivist science has largely endeavored to uncover general principles, and psychology is no exception. We use data from samples to describe how populations typically behave and feel. Clinicians, however, work in the opposite direction – they try and understand individuals using general scientific principles supplemented with specific information they have learned about those individuals. This leaves them susceptible to the ecological fallacy, which is essentially the idea that a general description of how people behavior on average may not be directly applicable to how a single person behaves (Molenaar 2004). Whereas the nomothetic approach answers the question “what treatment will be best for the population” the idiographic approach asks “what is the best treatment for this person.” What we need, then, are models for understanding how single individuals behave that can complement these population based models and enable us to personalize treatment (Stumpp and Sauer-Zavala 2021).

In order to demonstrate the utility of an idiographic approach, Beck and Jackson (2022) created numerous models that use a single person’s data to predict their own future affect and behavior. However, they left out an important comparison: do these models outperform models created with population data?

In the current study, I create some admittedly naive nomothetic models to compare to their measures of model performance.

# 2 Descriptions of the Data

The final data set has 44 total participants which have responded to an average of 44.59, and a minimum of 30 surveys. The outcome measure was a measure of loneliness binarized to enable the use of classification models.

Because this was previously used in machine learning analyses, it comes preprocessed with many features we learned about in this course. For example, cyclic time variables, and various measures of cumulative time on linear, quadratic, and cubic scales. In terms of substantive variables, there are both measures of general personality Traits (self-reports about how people feel and behave in general, variable names are preceded by the big five trait they are associated with such as openness\_, extraversion\_, etc.), as well as contemporaneous personality States (about how they have felt and behaved in the past 4 hours). There are also measures of situational features like whether they had any social interaction, watched TV, or used the internet in the last four hours. These data have no missing values. See below for descriptives and note that categorical situational variables were summed within-person and means were taken across participants.

Table 1: Item Descriptives

Measure Type	Variable	Mean	SD
Trait	agree	2.69	0.82
	open	2.89	0.59
	neurot	2.91	0.55
	extra	2.98	0.64
	consc	3.08	0.69
Momentary	afraid	2.27	1.03
	angry	1.99	0.91
	attentive	3.47	0.99

Table 1: Item Descriptives (*continued*)

Measure Type	Variable	Mean	SD
Situation and Time	content	3.48	1.00
	excited	3.13	1.02
	goaldir	3.44	1.00
	guilty	2.46	1.09
	happy	3.37	0.98
	proud	2.89	0.92
	purposeful	3.39	0.96
	adversity	1.10	0.37
	deception	1.08	0.31
	duty	2.16	0.85
	intellect	1.72	0.79
	mating	1.37	0.71
	negativity	1.79	0.79
	positivity	2.23	0.71
	sociability	2.17	0.85
	anx_s_wk	11.86	9.86
	argument	3.64	1.35
	brd_s_wk	6.57	6.13
	class	6.73	3.66
	exc_s_wk	4.34	2.40
	frgt_smthng	3.43	1.11
	interacted	25.45	10.86
	internet	12.36	10.10
	late	3.95	2.03
	lost_smthng	3.14	0.55
	music	10.25	8.06
	sick	5.59	4.13
	sleeping	8.20	5.12
	study	12.27	7.28
	tired	16.36	9.25
	tv	5.91	4.30
	evening	38.07	10.88
	fri	6.09	2.55
	midday	16.25	6.37
	mon	6.75	2.28
	sat	5.43	2.28
	sun	5.20	2.61
	thu	6.80	1.95
	tue	7.27	2.39
	wed	7.05	1.79
Outcome	lonely	5.36	3.97

## 3 The Models

Broadly, I planned to use Accuracy to evaluate these models as that is what the original authors did. I also planned to look at AUC to give me a sense of how well these models perform above chance.

### 3.1 Logistic Regression

I decided to start out with a basic unregularized logistic regression to see how poorly these models could perform. An unregularized model is typically undesirable when a large number of variables are included as predictors. Essentially, regularization techniques remove predictors from the model if they do not meaningfully contribute to prediction. This model does not do that, and instead includes every predictor I give it regardless of its utility for predicting loneliness. So, what you'll see here can be interpreted as a basic logistic regression can – as straightforward linear combinations of the predictor variables. Logistic regression models do not have hyperparameters to tune.

### 3.2 Elastic Net

The elastic net model is a regularized logistic regression. So, linear combinations of the predictor variables are used to predict the outcome. However, because these are regularized, not all predictors make it in the final model. This is helpful to reduce issues due to multicollinearity (when multiple predictors are nearly identical).

Elastic net is a useful compromise between using either a ridge or LASSO regression and allows us to choose which balance between their penalty functions optimizes prediction. The different penalty functions have differential utility for some prediction problems, so optimizing  $\alpha$  allows us to determine what balance of the penalty term we ought to apply.

We also need to decide what the proper amount of regularization/shrinkage to apply (i.e. what improves prediction) by optimizing  $\lambda$ , which essentially decides how strongly to penalize predictors. High values of  $\lambda$  strongly penalize non-contributing predictors while a value of zero means there will be no shrinkage applied.

I use a pretty basic tuning grid that searches  $\alpha$  from 0-1 in increments of 0.01 and values of  $\lambda$  from 0.001-0.75 in increments of 0.01.

### 3.3 Random Forest

To replicate the analyses in the original paper, I included a random forest model. As opposed to the first two approaches that use regression, random forest uses decision trees. Decision trees turn classification problems into a series of binary categorization questions. For example, a very naive model might ask “is it currently Saturday night” and decide that when it is Saturday night, it will predict that people are lonely. It might go further and include multiple variables, for example concluding that people are lonely when it is both Saturday night and they are alone. As you can imagine, there are a lot of decisions these trees might attempt to make. Random forests approach this problem by creating many different combinations of splits on different variables to explore which splits optimize prediction.

The `mtry` option allows us to select the number of predictors that we include in the models. In this case, I set it to simply 20. Is this justified? Probably not, but I wanted to ensure that the models were relatively quick to run, I know that the maximum number of predictors is  $\sim 80$ , and wanted to compare the top 20 predictors from each modeling technique.

I did, however, look into the effect of different numbers of trees. Essentially, I trained the model using different amounts of bootstrapped information, looking at the effect that number of trees had on the prediction. Rather than aggregating across all of these models as is conventional, I selected the best performing model as my final mode.

## 4 Comparing Model Fits

Here's where things get awkward. These models are...shall we say...abysmally bad. That's a bit of an overstatement when comparing these models to the original paper, but they're pretty bad by conventional standards.

The elastic net model does the best in terms of Accuracy, at 0.83, compared to the mean of 0.87 in the original (although the median is 0.91). The AUC, too, is worse at 0.61 in the current models compared to 0.69 (median = 0.74) in the original.

The random forest model is the best model in terms of AUC 0.65 and is comparable to 0.62 (median = 0.58) in the original.

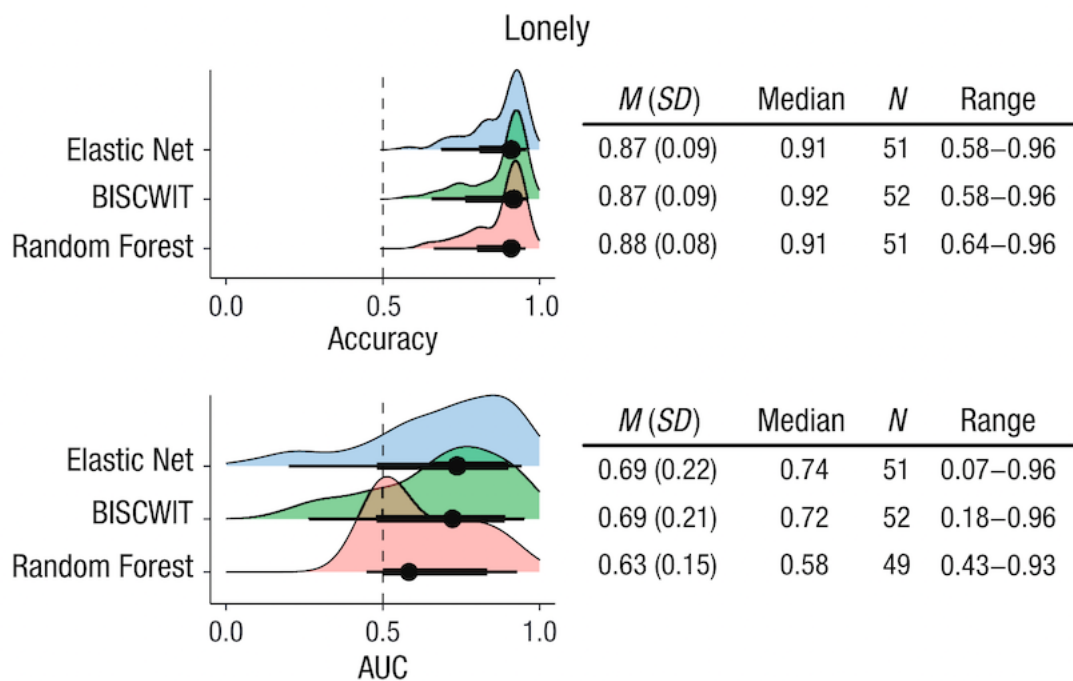
Surprisingly, the basic logistic regression is not all that far behind the others, with an AUC of 0.61 and accuracy of 0.76.

These models were pretty conservative, and had low True Positive Rates. This likely has a lot to do with the cutpoints I chose. Because the cutpoint function was unable to find an optimal cutpoint, I poked around manually. At 0.5, none of the models predicted a single instance of loneliness. To keep things interesting, I decreased the value until they were making at least a handful of predictions, though the False Positive Rate started increasing quickly. Much of these models' success seemed to come from not making predictions of loneliness, which was true for nearly 93% of observations. Could this be a low base rate issue? There is a very good chance.

Table 2: Model Fit Statistics

Model	TN	FN	FP	TP	LL	AUC	ACC	TPR	TNR	PRE
logistic	140	13	24	3	0.39	0.61	0.79	0.19	0.85	0.11
elastic	148	15	16	1	0.35	0.61	0.83	0.06	0.90	0.06
rforest	133	12	31	4	0.36	0.65	0.76	0.25	0.81	0.11

See below for the results from the original paper, and note that because they created a single model for each person, these distributions are of models of different individuals.

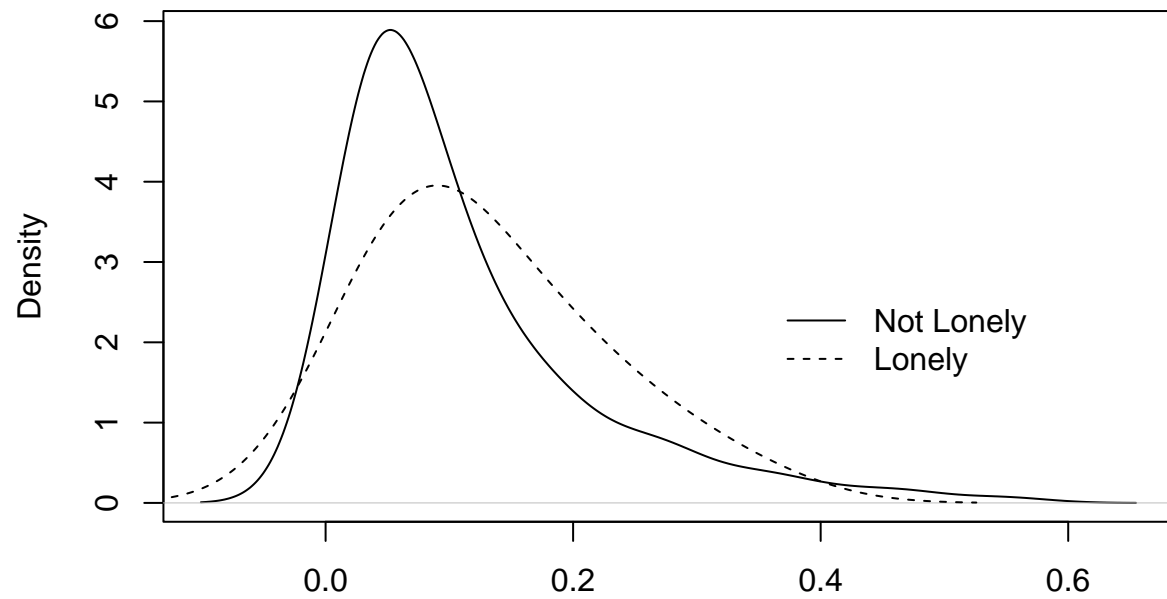


## 5 Plots

### 5.1 Model Fits

#### 5.1.1 Logistic

See...not a lot of separation in the predicted values for lonely versus lonely moments. This pattern will repeat itself across models.

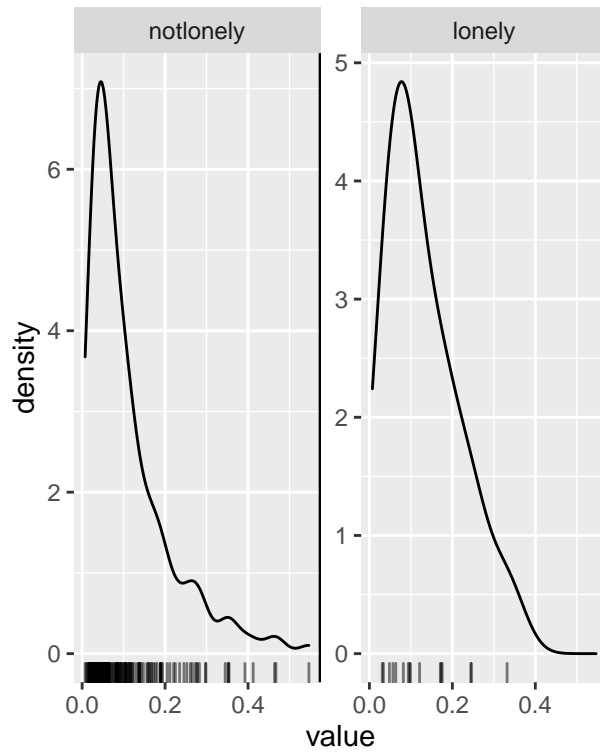




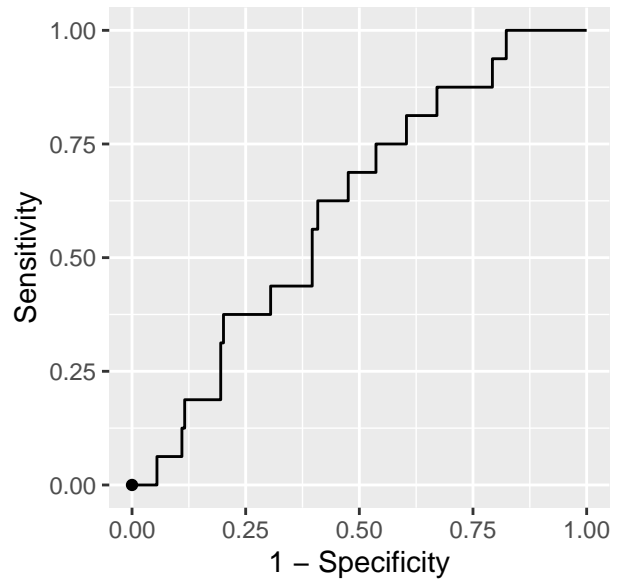
The AUC curve on the right shows a model that does not do much better than would be expected by chance.

## Independent variable

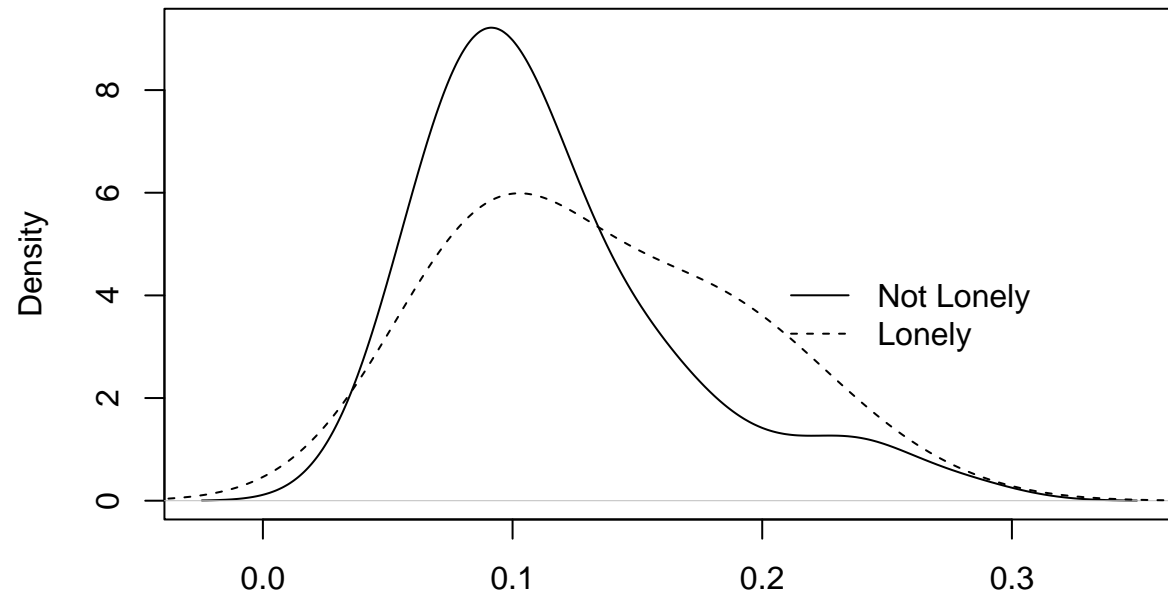
optimal cutpoint and distribution by class



## ROC curve

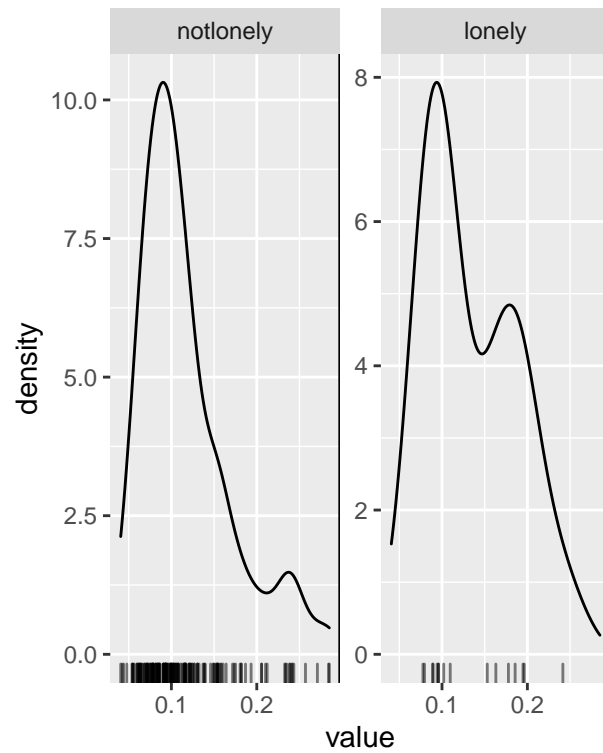


### 5.1.2 Elastic Net

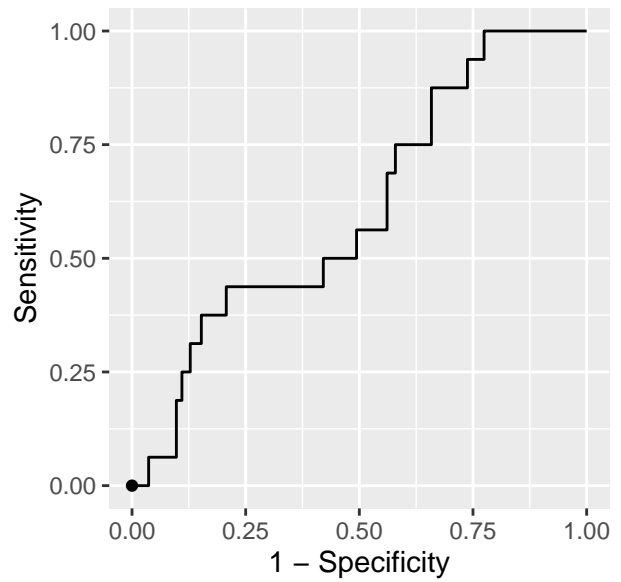


## Independent variable

optimal cutpoint and distribution by class

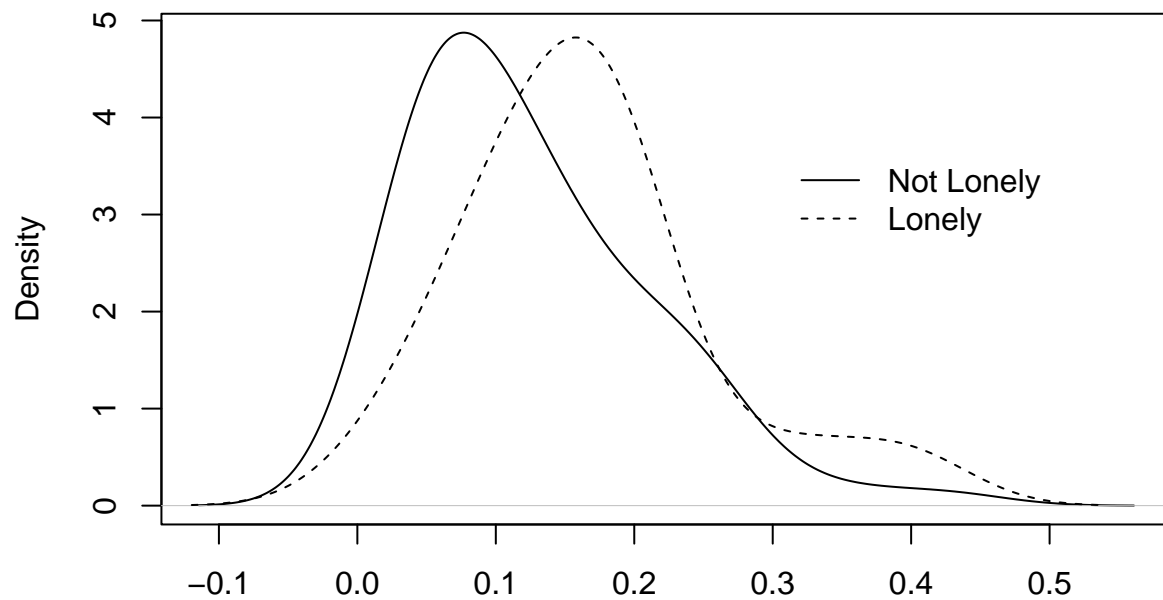


## ROC curve



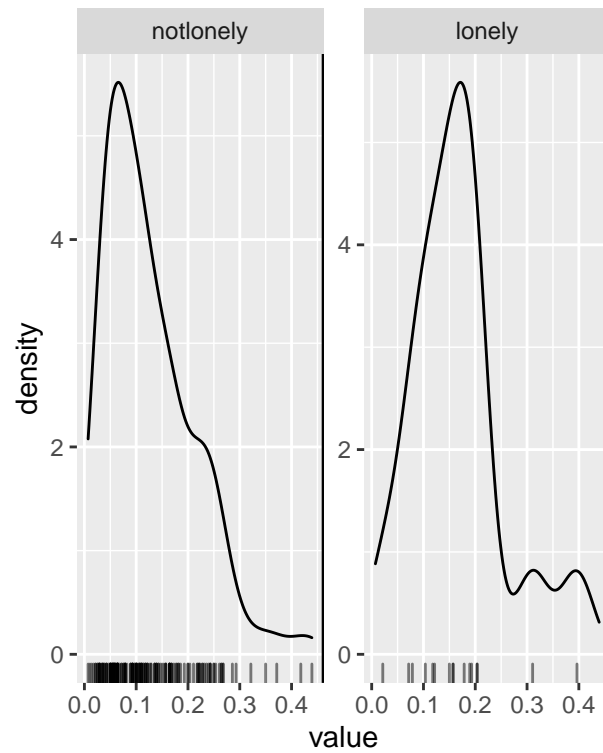
### 5.1.3 Random Forest

Here, at least, we see two non-overlapping peaks.

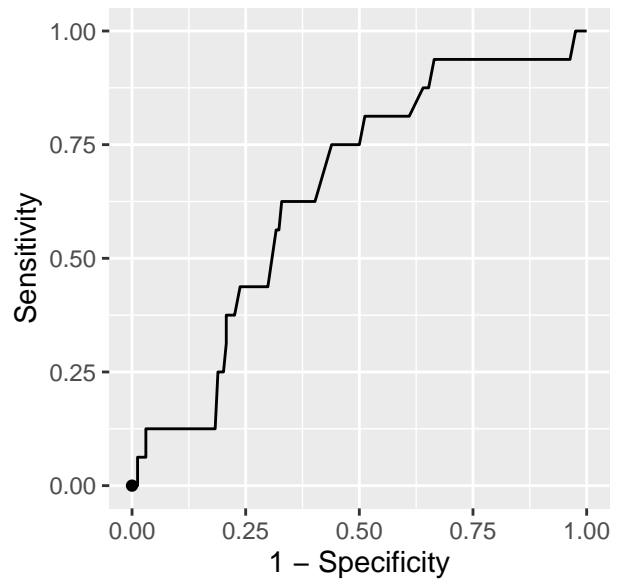


## Independent variable

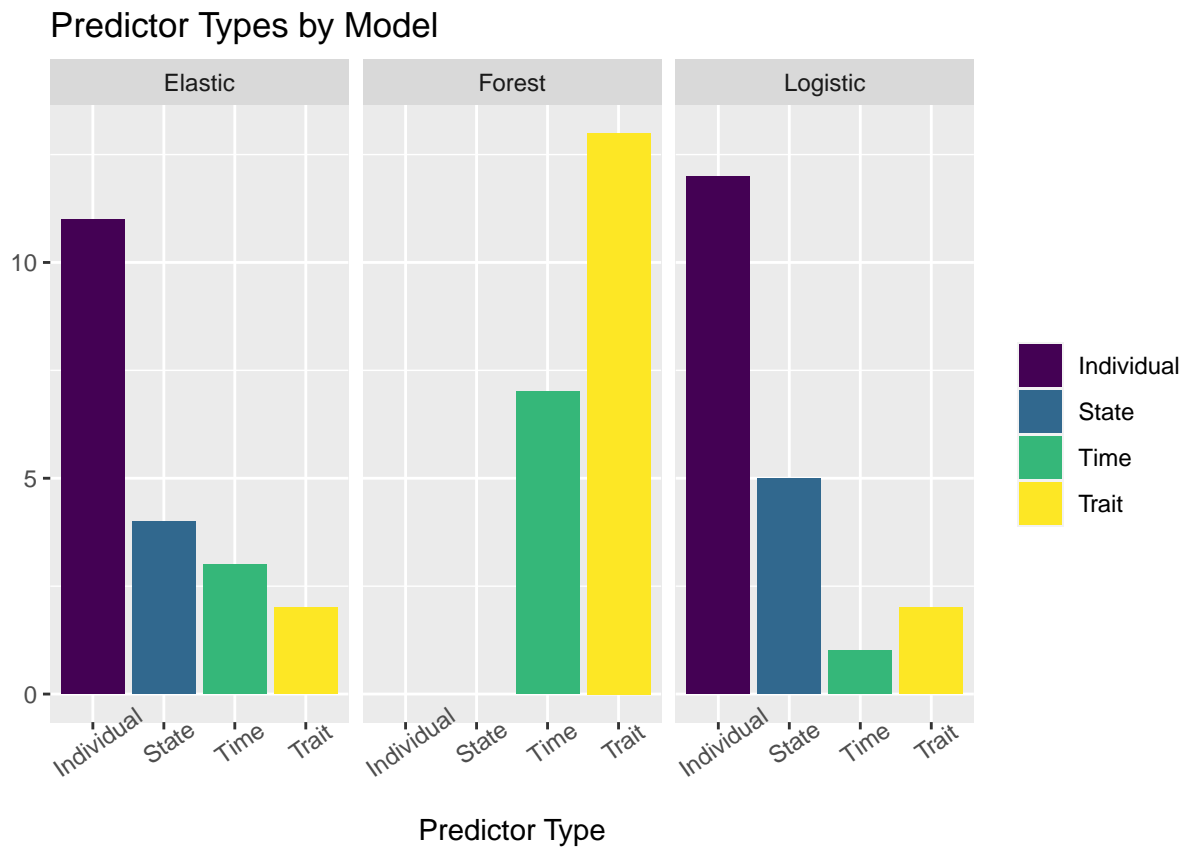
optimal cutpoint and distribution by class



## ROC curve



## 5.2 Comparing Coefficients/Importance



## 6 Discussion

My takeaway is that models can have somewhat reasonable fit statistics (i.e. accuracy) without being all that predictive. I created three models that were hesitant to make a decision that a person would be lonely. On one hand, the low base rate of loneliness makes this a reasonable tendency. On the other, it inflated the false negative rate.

These models underperformed idiographic models that were trained on much smaller datasets (since each person's data was used in isolation), which is to the credit of the idiographic approach. However, none of these models (mine or theirs) had fit statistics worthy of real-world implementation.

Of course, this is not an apples-to-apples comparison. One major limitation of my work has less to do with the models (though there are limitations there, naturally) than with the cross-validation. In the original research, each person had a rolling window where 15 measurements would be used to predict the 16th. This allows the models to incorporate more behavioral/affective inertia, which are likely very important in predicting affect. Time constraints prevented me from hacking together a solution, but I have a strong feeling this would improve the models.

I would also be curious to see how a leave-one-participant-out approach would fare, since that's a stronger test of the initial question about the relative merits of nomothetic versus idiographic models. The current models include some personalized data, but a more common situation would be to have an existing model trained on a sample, and apply it to a new person's data that is not present in the training data.

An interesting pattern in these data can be seen in my final figure, which shows predictor types preferred by each model. Both the elastic net and logistic regression show that individuals were the most common of the top 20 strongest coefficients. This is substantively interesting because the main question of this line of research is how personalized should models be. Unfortunately, the sample size is so small that I have a hard time deciding if this is because more people would be preferable, or that an idiographic approach is warranted because of the outsized influence of individuals. The fact that State affect/behavior is the second most common predictor type in these models is not particularly surprising.

On the other hand, the important variables in the random forest model are perplexing. These are not coefficients, but were extracted using the `importance()` function, and give a metric of how influential the variables were in the final model. Here, we see Traits and Time as the most important. If I were to speculate wildly, I might say that  $\text{Traits} \times \text{SpecificTimes} = \text{States}$ . In other words, people that are extraverted tend to be out on Friday evenings, being extraverted. But I really don't think these models are strong enough to start reading into them.

Ultimately, it seems that these models, like researchers, decide what variables are important in a very idiosyncratic process (Bastiaansen et al. 2020). This work is very much a first draft but was an interesting foray into the trials and tribulations of predictive modeling. I haven't drawn any substantive conclusions, but have started to wrap my head around the various challenges that confront researchers hoping to make predictions about human emotions.

- Bastiaansen, Jojanneke A., Yoram K. Kunkels, Frank J. Blaauw, Steven M. Boker, Eva Ceulemans, Meng Chen, Sy-Miin Chow, et al. 2020. "Time to Get Personal? The Impact of Researchers Choices on the Selection of Treatment Targets Using the Experience Sampling Methodology." *Journal of Psychosomatic Research* 137 (October): 110211. <https://doi.org/10.1016/j.jpsychores.2020.110211>.
- Beck, Emorie D., and Joshua J. Jackson. 2022. "Personalized Prediction of Behaviors and Experiences: An Idiographic Person–Situation Test." *Psychological Science* 33 (10): 1767–82. <https://doi.org/10.1177/09567976221093307>.
- Molenaar, Peter C. M. 2004. "A Manifesto on Psychology as Idiographic Science: Bringing the Person Back Into Scientific Psychology, This Time Forever." *Measurement: Interdisciplinary Research & Perspective* 2 (4): 201–18. [https://doi.org/10.1207/s15366359mea0204\\_1](https://doi.org/10.1207/s15366359mea0204_1).
- Stumpp, Nicole E., and Shannon Sauer-Zavala. 2021. "Evidence-Based Strategies for Treatment Personalization: A Review." *Cognitive and Behavioral Practice*, November, S1077722921001334. <https://doi.org/10.1016/j.cbpra.2021.10.004>.