PDF-XChange Click to buy NOW! www.docu-track.com

We've made a few improvements to the forums. You can read read more on the blog.

Forums / Data Analysis Assignment 2

Black-Box vs White-Box Models

Subscribe for email updates.

Sort replies by: Oldest first Newest first Most popular

No tags yet.+ Add Tag

Jonathon Khoo · 11 days ago

Ok, this is going to be a long post, so the TL;DR is right here below. Was going to post it in response to this thread, but thought it deserved its own thread for debate >.<

TL;DR
Random Forest = Black-Box, more accurate, hard to explain and explore further (which means possibly lower marks)
Simple Classification Trees = White-Box, less accurate, easier to explain and explore further (which means possibly higher marks)
IRL, it really depends which you use. (Personally, I prefer white-box methods)

I think part of the issue is that this course does not really mention much about the whole debate of black-box and white-box model building. Personally, I settled on using a regular classification tree in the end after using random forest as it just made more sense when explaining the conclusion. While the course did explain how to create these models, we have not covered how to explain them well and also the reality that IRL, when people go to deploy their models in a production environment, many times they will only use a few factors in a less accurate but less computationally expensive model. This is exactly what happened with the Netflix Prize. From all the crazy algorithms that were developed, only a small subset of the combined models was actually used in production.

Most people don't know what is going into the trees created by the random forest, and thus can only comment that the results they have were accurate, but if a regular manual classification tree was used, we can actually see what variables were eventually used and notice and explain some interesting patterns within the classifications. This was what I personally didn't like about most people's random forest write-ups, but as I mentioned, this isn't really their fault. However, this should at least be mentioned as a limitation of using it.

This varies largely from the first assignment in which we are much more clear of what goes in and allows us to actually try to understand why things happen the way they do and are not confounders. This is also why most likely few would actually talk about confounders as well in this assignment.

The limitation is that this is an academic exercise without a very clear idea of how our model would be applied once created besides meeting the grading criteria. In a real-life setting, the ultimate goal will always dictate how complex the model we design should be.

The volume of data can make a difference when choosing which type of model to use. Monitors attached to an ICU patient might not have the time or ability to run an extremely complex model every few seconds looking for anomalies, so it might only focus on a few key factors and alert the doctor to make a decision. Alternately, a bank checking credit scores for potential loan applicants on a batch basis at the end of the day might be able to run a huge complex model with hundreds of factors.

Also very important is how actionable the model is. If we know what the key-factors are, it makes it easier for us to act on them in specific ways. If the goal is to only act on the final result of the model, a black-box will work, but what this action is might not be so well defined. If we want to work on individual touch-points that might affect the result, it isn't so useful.

Then there is the whole problem of rebuilding the model after some kind of intervention and the new data that comes with it… etc.

In conclusion, it really depends from situation to situation which type of modeling is more applicable. Personally, I like to know what I am using and why things happen. What do you guys think, especially those of you who are also in this field as a profession?

9

Richard James Wilmot · 11 days ago

I agree. A very good post. I think that Jeff clearly made this point that accurate methods are sometimes not that interpretable. I think that in order to trust a model, you need to be able to understand it. And to be able to understand it, it's got to be interpretable. I think that few of us could clearly explain in 2000 words why a random forest predictor is the best tool for the job, even though it might be. Put that in a professional context, and explaining why the company should bet on a black box is a big ask...

3

+ Add New Comment


Thia, Kai Xin · 11 days ago

Well, if this was a real world assignment, I will just deploy a grey box: Push random forest to max accuracy, find the top 10-15 factors that can explain 90% of the variance and plot a simple decision tree using the top 10-15 factors (R random forest library even come with build in importance() function). Ironically, one of the papers I marked did that while I didn't for my own =P

5

PDF-XChange Click to buy NOW! www.docu-track.com

+ Add New Comment

Paul S. Hewitt · 11 days ago

Personally, I feel more comfortable with a white box method, because I can understand the decision tree that gives the answer. Dr. Leek did mention that this is only an issue when people are expected to use the model to make predictions. When computers are required to run the model, you can go as complex as you like, for example random forests.

Despite my preference, I went with the random forest, because it provided the best predictions and it was clear (to me, at least) that no human would be able to make all of the calculations necessary to use even a simple tree decision model. Note that the readings from the gyroscope and accelerometer would have to be taken, then standardized, and several of the 559 variables would have to be calculated. Then, the human classifier would have to compare the current observation variables to the ones in the decision tree, to many decimal places. Simply not practical. So, if a computer is required to process the data and make the decision, we might as well go with the model that predicts best.

Even though it is difficult to interpret the random forest model, we can say something about the most important variables that contributed to the classification (of each activity, too).

1

+ Add New Comment

New post

Bold *Italic* Bullets Numbers Link Image Math <HTML>

- [ ] Make this post anonymous to other students
- [x] Subscribe to this thread at the same time

Add post

PDF-XChange Click to buy NOW! www.docu-track.com