

10/15/2016 06:41 PM

I also thought a little more about the lasso feature selection. I think we can do it just fine as

long as we use a different loss function: namely- use $y - \text{sign}(\theta * \phi)$ instead of just $j - \theta * \phi$. The problem we might run into, however, is calculating ϕ . I'm still not so sure I'm thinking about this the right way. I totally get how it works with regression but I'm not sure how to implement it with classification. I have read about other people using it for this, however, so I trust that it will work and that you know what you're doing. In either case, I do see how we can ignore the θ from this stage farther down the pipeline if we just get the subset of features, which is what I believe you were trying to tell me the last time we spoke. I recently reviewed the information analysis approach introduced in the same lecture and I see how this can easily be implemented, so I may try to build a sub-module that does this. It can't hurt to try more than one thing here and report on the results right? **Yup, doesn't hurt, infact looks good, since it's what true ML work in the project is, rather than implementation of code. I am not sure if I got your query right for the first part of this paragraph, but I'll read this again tomorrow (Saturday) once I have done some research.**

I can't remember when we planned on talking next. Is it tomorrow or Sunday? I'm free either day and the afternoon would be better for me in either case. I'm also free all day Monday. Whatever is best for you. **Let's talk Sunday, so I have sufficient time to catch up on the backlog and so that we both can discuss what we did.**

Cheers,
Phil