Tokyo Workshop on Genomic Medi...eutics, and Health 20240413-02

Sat, Apr 13, 2024 10:37AM • 27:09

**SUMMARY KEYWORDS**

model, performance, snips, prior, pretty, estimate, perfect, similar, compare, assess, call, hyper parameters, method, configurations, viper, problem, prediction, update, perform, large

00:00

Probably the people who probably haven't been doing this, you have 3 million parameters. You have got 30 doing practice. And that's all the time. You have sort of a linear model. So that the way we estimate the parameters that maximize the best. The end when you got out that's all you have. For people production, you take individual, ders, and now you have to pay tax for a test phase. And now you're in five five beta. That's a popular product so that you guys assess whether high risk the hyper parameters are what we have to strategy.

01:05

So here are the first value the model by using simulation the same one as I saw you earlier, very similar, but I'm very excited for you do 101 million 1,000,003 Six. This is using the UK Java data for why do we use simulation the simulation we can assess the behavior of the model because we can set the height of the heritability accordingly, by 30 with a higher higher overview expect higher performance measured by incremental or incremental sort of complex numbers. So this is we're estimating the boxes for flow through different methods. So we have our live versus default one where the default one we estimate the hyper parameter using the answer by directly maximizing alpha and now we also have this lifers TF which is research. So essentially, we use a validation sets hyper parameters, it turns out it's actually performed a lot better than my personal real data. So you see Malaysian they perform similar because the the first two boxes are kind of hilarious. And the fact is the lasers and sometimes the Viper perform better than the law. The rest of the methods are the state of the art methods. So you might identify as a jar which is using which it also has an expected utility of sampling from our various smoke bottles. And then you have a lot of time that I'm going to talk about maybe up there is also a similar type of difference. I forgot to detail that you have an LP prior to which is also new MCMC approach and the RFC s. So I think, is also another Bayesian. There's a lot of different prior non precise tools.

03:13

Yeah, so yeah, I was thinking in terms of in terms of, you know, fine mapping, probably the last name that you introduce that the IP address seems to be the best compared to the previous models. You do it today. But I was thinking in terms of prediction of traits. I was wondering whether the simpler model may have the same promoter may have maybe may have enough you know, performance compared to the similar model. Okay, so. So, my question is whether, so, what? So, how does you know performance is different compared to the first model you do yesterday? I mean, the model container model based on that, gotcha.

04:07

Yeah, so, I think tablets, you can use of Gaussian prior as a Gaussian prior, that that as a laser. So they have done the ablation study, or what have you gotten prior was a performance. So we didn't compare that because Gaussian one is usually very low. So you I guess you might have read the LV prior to neutral paper in 2015. They have a house and car respite and so that car, very often the golf

04:38

platform, so the model considering the you know, the they're considering the LED and also, like I said, the causal causality of the causal effect of the variance. improves performance based

05:04

on our assumptions. Regarding the cover architecture sparks believe are small fires that are causing reservoir, therefore better performance, maybe perhaps in the real situation.

05:22

I was also interested in this assumption improves the plant seeds the prediction of pre processing is the prediction. Now that

05:34

we have some parameters on that we also have another different method of working on our one group and then fast on the other

05:50

Okay, so I'm sure we touched on the data for nine different qualitative traits from The Guardian, famous trader cites high productivity people protein, not body making that choice for themselves as we evaluate these same. Now we invalid based on our testing for the new for mental health for our voc for the highest and we have error bar for our viewers when they do cross validation. So your model absolutely doesn't sometimes and only computer stars version of the hardware. So because you don't really have the best performance. The runner up will be the estimates are the second best. So the visors default diapers, if you estimate the three hyper parameters using maximization despite the mathematical elegance to the prefer the actual performance during the field I'm working on it for a long time. Although this might look like a small thing. It's actually pretty surprising.

07:22

So we're waiting for that deep learning breakthrough to break out of this restaurant or Addison. Suddenly you have a much higher but that's something that we're still, you know, I know a lot of you might be familiar with and maybe more exciting for the neural network deep learning setting like in this case, when you deal with the community, on the other side, is difficult to train a neural network that can actually have a competitive performance. Because just a linear model, it has 3 million parameters. Just adding a few more didn't blow up your parameters. The training examples are not that big compared to the number of pieces that we've learned. So we have 5 million individuals we have on the other side.

08:13

So and then we divided on the binary. So the binary prediction we were kind of Reformed seminary by the sandwich seminary to each other. And now we often value the runtime. So how fast because again, the various prosthetic recovers in sort of less than maybe 10 minutes. When you actually do this, yeah, we reserve to take longer because it has a service, all that type of practice. Each time we take into account the time they take for the model converge on a specific time decided even that because we're actually pretty competitively we actually got about one hour we twin it sequentially. We cleaned this Viper is a now we trained sort of 22 Vipers individual protocols and then they do generosity some crazy because they're from different from No they're not. So sort of the client essentially sum up all the quantities so we click on sequences well from right now we're working on sort of how the last version massively paralyzes. Right this is our independent LT blocks to see if there are different chromosomes. You can make use of a multi thread for CPU to really bring down the time for one hour. I know it was the first time you achieved maybe a couple of minutes so in here before this for the rest of the model performance, they have much longer separate an awful lot of time and similar to coordinate sense the same time that performance is not made up there is a cost to that really, it's kind of fast with no theory. So speaking, all that discussion that we just had, we test how much you know, how do we perform? Let me explain. That's a white British and then divided on this individual from different assets. But you said oh, well do you have this report and you were drawn out from our wife, but then you'll have the federal individual, like these percentages, from different athletes backwards? And we want to assess how well did you do? Will you train on my British and they'll be passed on, let's say in Korean, Indian, Chinese. And so nowadays, it's all about it online and different sites. So you can see that the one that performs the book faster that shoots a pretty substantial performance. So you can kind of have a similar afternoon. Whether it's performance start dropping the population happy backgrounds that start becoming more and more distinct from that the libraries are highly similar. So now that you can see that we again we compare it across to the people method, you'd say as a top contender will be through this type of recorded zero songs. It was only 21 years that one athlete will be just middle to fine tuning we will continue to learn from these other groups that by the transfer ability is so that we want to because the model is very efficient, so we scaled it up to 9.6 million, including when and obviously we assess how much can we gain by increasing this? 1 million 103 to 9.6 million to happy to do it for our highest potential. This is comparing Viper trader and my first screen on my quite big so we compare between them and then this is why first Twister right the first one, we have a brand new one. And then on the left. Here you also have especially for the massively polygenic like fights are facing you know, pretty much there are a lot of columns that always feel a small fraction of the overall majority but by embracing you know from successful seminar for ACL for the amount of robbing place, you know sometimes you might not get that right depending on the quality of music. So in summary, you know we interview to phase in Platinum so that you can use your various nutrients to estimate in fact, even the Bible says sophisticated prior museological new friends will make use of some resistances to trend by first right you will have accepted you do and then we assess the convergence by elbow as the lower bound is highly scalable. I can show you how we can still approximate this. So the model predictions for the future. So we have a still ongoing work paralyzing the Vipers, we have another new methods and mixture of experts classes. So these are all pretty much Saudis, the main driver of the project, we also have a crop population. So there are several instances that we're currently working on. Obviously we have a limited bandwidth and you guys some of you are interested in some of these projects, contact me. So, before I close, I do want to raise a caveat. So I'm actually doing a project because as a data from my personal finance, so now there is a problem.

14:32

So now if you do not say what type of scenario based simulations can actually be simulated, where your snips are almost perfect. So once they will follow. You compare with the exact users, right, it will simulate small numbers and maybe similarly, you can just use often enumerate all the configurations. That's exactly correct. So now we can see the next time you have a slip of water. That's one of them. So in the back, of course, the tip is pretty much split by six. Now if you run something like Vipers with the music in one neutralization, because they are different that you want to model things first, right. Second one is you don't have a pie chart, the exact both these are the configurations. So both both called oil to 23%. They run in multiple times. And then second quarter of the year is pretty much the wild third, for each job. The configurations are as a variation on users. Send me 17% And then the first one is probably a 2% That's pretty high. Despite the fact that first of all, a perfect clerk will now have to run it again. Then you have to post it stay safe is awesome. All right, all the configuration are out of 700 learn the various new fundamental saying that first one is now called the second law the law if you hadn't thought you see that at first, you missed the two holidays so the Viper or the murders on New Zealand. The tip that London on the Smith, rednecks has verified with a call, but it's a third law and that in this case it wouldn't call the line Florida. It has a really tight LDP with everybody. The model will be overhauled so that's the caveat that I wanted to do find that we don't recommend to use. But there's a difference between having a very accurate Cherise first, you want to have a very accurate car so you don't have to choose the correct color. If you exclude this name that has the perfect LD with a causal, you're gonna get exactly pretty much the same with predation happens on the fly but if you want to say okay, can I use the backer kit to do experimental validation founders believe that the payments well calibrated. So now you have this. So now given what we've covered, how you identified a closet, in the in the museum, remember recycled at one time. So here's the problem over here. So now the kids is one of the key points. Are you perfect it was awesome once marked. Perfect. So that means you basically cannot do anything, but two features are identical. Why the models sometimes seem overconfident that happens very rarely. But sometimes it seems good. likes to call it sometimes it is it is complicated right sometimes conflict.

18:49

But is it related to the effect size so once of the snip snip has a larger effect size than the other? Well if they're perfect for your computer, the markup factor is the same. You have two identical each other on the other?

19:17

Does this depend on your spotting values and also the order that you will be updating. So if you update one first thing, do you have a large estimate in and when you come to the second one? There's context for the first one

19:38

that exactly depends on the board or not you run this sort of iterative update. Let's say I have this forum. So now at the beginning my newsletter can be variable. So now first I decide first update this is the causal marketing process of our large although is perfectly correlated, non causal. So this part is going to be small. Let's say we have only one policy, right according to the rest of sense artists. There are no problems. Now if that is the order and you're lucky, you will have a larger, healthier immune disease the first thing you update is the logs are unlocked. If the first day of your update is a non follow, there's no problem other than the fact that for a compromised not to be very low rent because that's how you use it. You always want to call it again. Now you've got to have a very large marketing process for nonprofits. Then the second one you will need the combo for now you're gonna have a very large effect size for that no problem because you update it first. And now when you subtract it from the marginal effect size of the combo and that's the last right because there are perfect accordion. So it really depends on the war. Right so yeah, so that's the problem. So how do we solve the problem? What's the perfect what's the what is the ideal pricing for this answer when the snips are perfectly the core, what do you think the pips?

21:42

So let's for the interest of time let's go through this. Now I can reveal what anger for example. All right, awesome. People will find that being across areas tradable sets, what is credible, credible sites that we rent in the Bitcoin water and then you add in the sniff until you reach a certain percentage, but Well, that means you're you're adding this sort of snips so that the some of the tips over that you have added it's a percentage of the total. So that gives you the sort of so called value set so I'll try to remove 95% You know, when you add the only a few steps are very successful. Okay. Now you don't have to add that many snips which let's say 95%. If the results are low, you already feel okay. That means your mold is very common for the furnace. On the other hand if your snips are in turmoil is not competent. So you snips are sort of kind of similar between uniform distribution, right, so Juwan 990 5%, particle size, you pretty much have to add 95% of all the snips right over three or fewer particles. That's sort of the value model. So so that's that's a traditional finding. Landed only on one single but that doesn't really solve the problem that we've just talked about. Right? So you have a cool person in Florida is another wrestler, maybe you know that they're nonprofit, in this locus, while you aren't necessarily impressed that answer. So, for example, I say consider this scenario so you have a two column here that says they wanted four. So one is completely correlated with this new two. And then four is completely according to a tweet recently is that the three and the two are passengers. So now you have a two independent is called the second one four. So x one is perfect for three and four, but not where you want to try this. So one point was that I have a one and two you will always minus Kabo because there are an animal. The other one is the Bible Saturday contains to inform

24:05

well, how will you infer what the pin we didn't? What's the what? What's the be the pin for this? For 5.5 Because that's your best bet. Because these are identical. There's no way you can tell which one is but that's actually the best uncertainty also for the the other partners that will pass the differential payments. So for us out of time I wouldn't get into the there's two papers one is called the Susy which is actually right now is the most well known find marking method. Right so we also use a Rhian formula for our pro so we show that Lucy actually also offered on nusil Various remover, wrote a fabulous sr 30. These credible sites you know exactly what calls but because you have all kinds of stuff you can have multiple independent policy. When you only have one call this guy does something exactly one of the new things a lot easier for me because in terms of school normalization for your kid, you have a need for someone to have exactly one policy but it will require solving from the outset when you activate the the underlying right that's similar to correlate the sense of sorry acids, but you'll cycle through the Bible as a cycle. If the last phase you didn't ask the operator, you can use it in the SPAR Pro that we had. Obviously in the last year. We also leveraging the functional implications so prior information about the accuracy and locations on the relation is kind of a separate line of evidence we used to hold the prior so I'll have to acknowledge sorry. for that. Also starting to see a project and also practical job

26:20

so I think we're gonna take a break when you can offer like, I do have a second part but I could also, you know, I think there's a lot of information I could also kind of present maybe using a bit of afternoon time but I think there's probably my first following

26:49

resume live at the macro level.

27:01

Call me