

Mortality Weekly Report (11/14)

Rustom Ichhaporia

Last week, I registered for HAL access so that I could run the hyperparameter optimization script remotely because my computer overheated and could not run it for the appropriate number of trials. I was able to upload my files and run the first half of the script, but unfortunately when running the `fmin` optimization function, the program crashes after the first of 150 loops with the error:

```
lightgbm.basic.LightGBMError: Cannot change  
bin_construct_sample_cnt after constructed  
Dataset handle.
```

After googling the error, the only information I could find was a GitHub issue in which someone faced the same error message¹. Although the issue had been resolved and merged into the master branch of LightGBM nearly a year ago, one commenter explained:

Looks like the issue occurs when `n_jobs`(number of threads) of a model not in `[1, #cores]`

Previously, I had been using the option `n_jobs=-1` to maximize the use of available processing power, but in response to this post, I tried changing the option to 4 and then to 1. I also attempted to run both interactive and batch jobs, as well as use both CPU and GPU partitions to run the script. None of these options helped the script to run, so I don't currently know how to run the script on more than 10-20 loops because my personal computer cannot handle it and it crashes on the remote server. I would appreciate any suggestions on how to resolve the error or proceed otherwise.

I also pasted this in the Slack, but I did not receive any suggestions.

As I was not able to make progress in that regard, I spent some time cleaning up the code on my local machine, so that it will be more polished after I am able to run the script.

¹<https://github.com/microsoft/LightGBM/issues/2696>