

The example predictions file you are looking at is a valid submission file. You could upload it right now and it would be accepted. So that's what it is supposed to be like – it includes ALL of the rows in the “tournament” data file (i.e. submit predictions for that entire file.) But the only new data each week is the “live” data, so if your model doesn't change from week to week you can work out your system so you don't need to predict the entire thing every time. Just depends how fast your model runs and the resources it needs if you need to fuss about that. (Last week's live data is added to the end of the tournament file each week as a new test era.)

As far as your predictions, they should be in the range of 0-1 just like in the example file. But yes, the training data only uses 5 discrete values/buckets for the targets. Nevertheless, your predictions should be real valued and ideally not contain any ties (they will be broken by row order, i.e. essentially randomly).

Oh, and in about a week we are getting new validation eras and moving to a target with a different distribution (but still only has discrete training values) so what I've just said about the file being relatively stable won't be true on the round that starts Nov 14. On that round, you'll see some test eras disappear and some new validation eras be added. Keep an eye on this forum and the rocketchat for the latest changes.