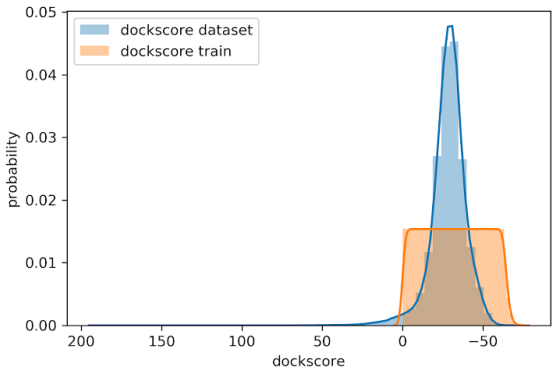
Don't know if these are particularly sensible variances and so these need to be tuned!
Cheap docking var 2.5 - 2.7
Expensive docking var 2.5 - 2.7
FEP var 0.7-1.5
Experiment: var 0.3-0.5

Oracle	Batch Size	Num Batches
MPNN	512	10 000 000
Cheap Docking	1	10 000 000
Expensive Docking	1	1 000 000
FEP	256	50
Experiment	15	5

Could you let me know more about the files you linked to last week:
Yes - the table d4_250K is obtained from: https://www.ncbi.nlm.nih.gov/pubmed/30728502

1. We took all of their 150M molecules
2. We discarded bottom 20% of molecules with very bad energies
3. We put the remaining molecules into bins from linspace(min(energy), max(energy)) and sampled uniformly from each bin
4. Here is the plot blue is the original distribution, yellow is the sampled distribution



a. This is docking score for D4 dopamine receptor? This is what we are maximising?

We are maximizing negative binding energy (ie the one with - 75 is best)

b. Are these dockscores measured in the same units as the FEP energies?

Dockscores in dock6 are not in the same range as kCal/mol. The best is to normalize dock6 score to the same variance as in Joanna's paper. I gues one needs to divide by ~4.2

c. There are 220k molecules or should I find more somewhere?

I think it could be good for toy start let’s pretend these are ground truth FEP, and we are trying to get normal FEP

d. What is the fingerpint/klabel for?

Fingerprint is Morgan Fingerprint. K-label is for K-nearest neighbors split into the test/train sets.