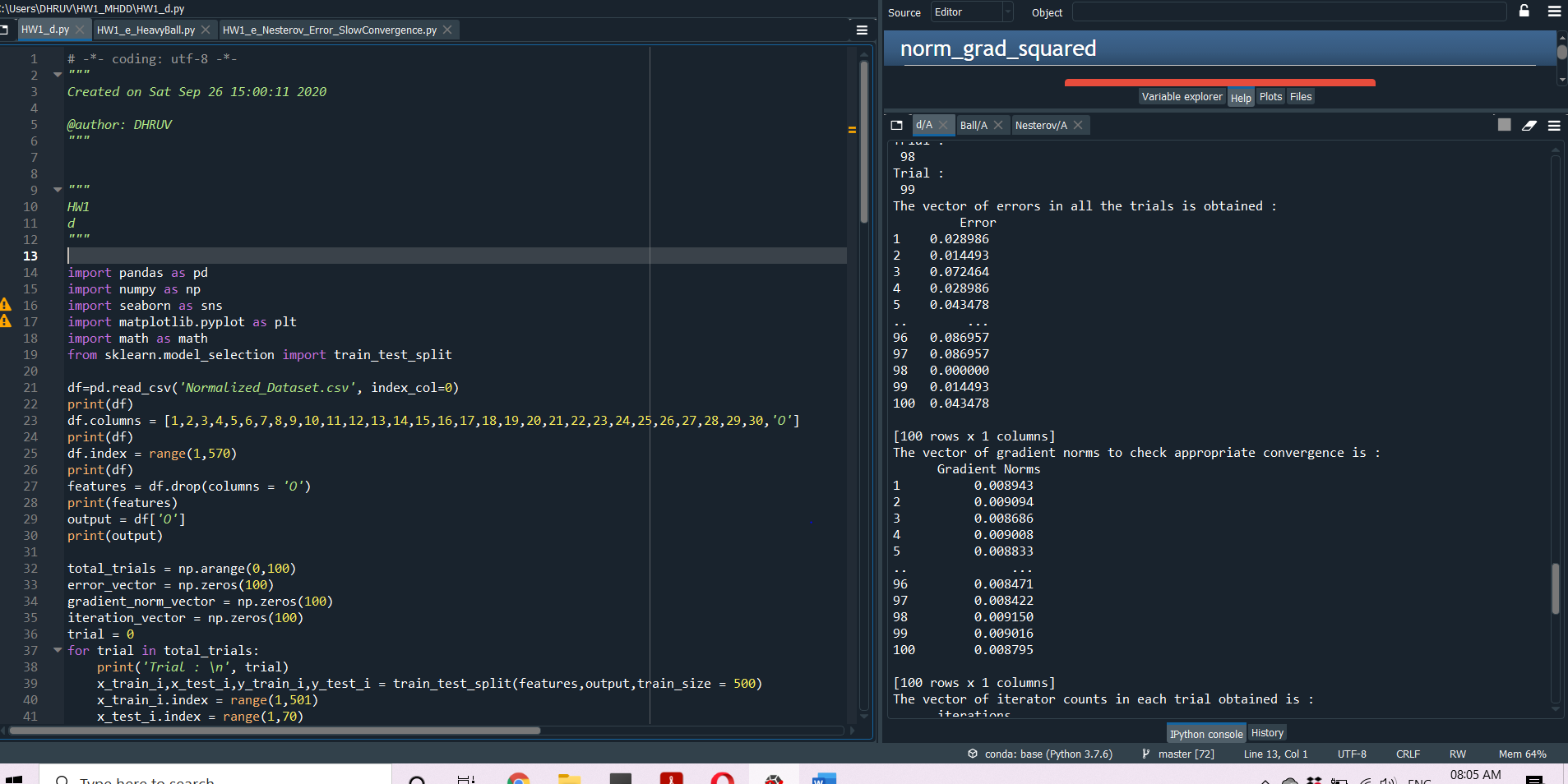
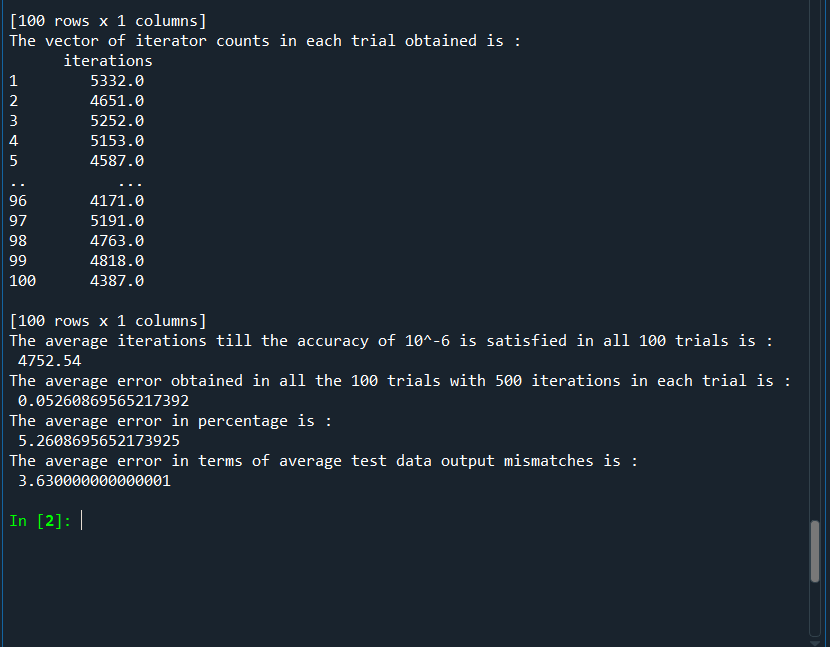
THESE DIAGRAMS ARE FOR PART d.

(It has a rather high run time)

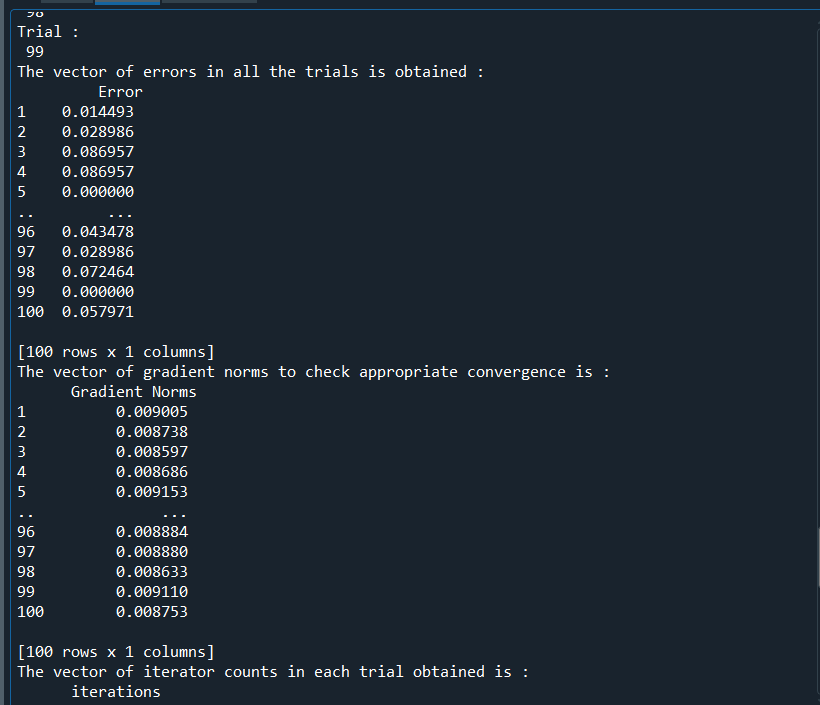
Few Results for Gradient Descent

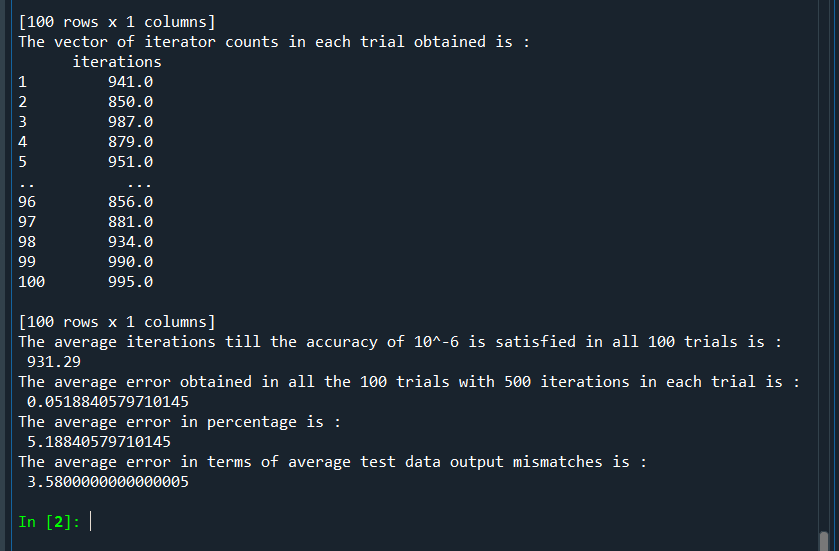






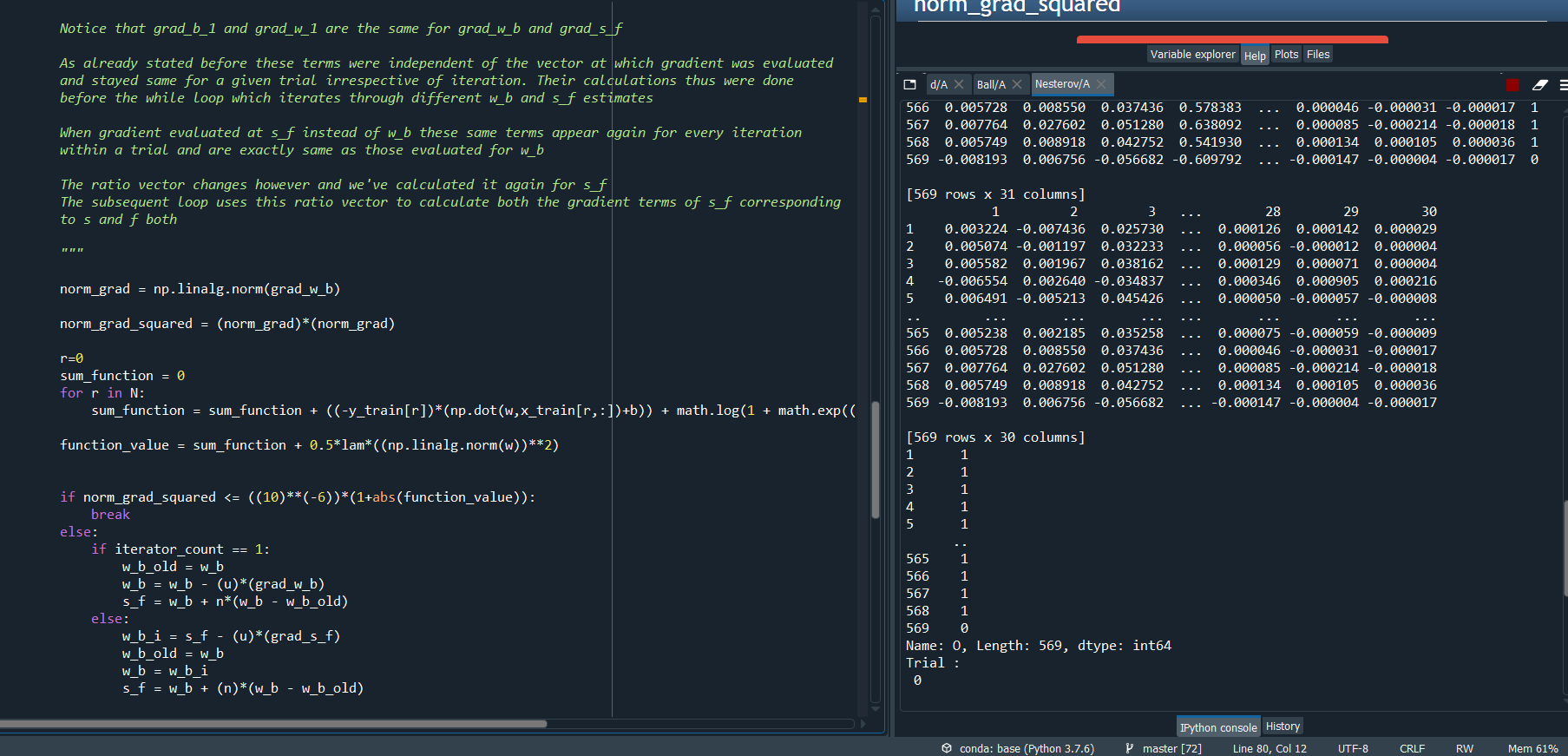
Few Results for Heavy Ball





As can be seen Heavy Ball converges much faster than normal Gradient Descent

For Nesterov



The run times for Nesterov Model were rather high and so in the few trials I could observe I obtained

approximate iterations of ~1000

Again it was, due to large run times, not possible for me to tune the momentum term to ensure

a quick convergence

CURVES

Plots have been submitted in the folder

Which model would I select?

I would select the Heavy Ball one as it showed the best results for me, despite the Nesterov method being the most preferred one. Due to run time issues I couldn’t tune the parameter of momentum in Nesterov, which weakened it’s results