

- Further analysis:

- Choice of α :

- If too large, we may "overshoot"

- If too small, it'll take ages to converge

- ~~- If too small, it'll take ages to converge~~

- Compare the learn. curve of our model w/ various

- Diff. α 's give different cost and hence diff. predict.

- A lower cost doesn't mean a better model (chance of overfitting, happens if train. acc. \gg test acc.)

- Recommended:

- Choose α that better minimises CF

- If overfitting, reduce overfitting (certain techniques)

- Takeaways:

- Preprocessing is important

- Each function was implemented separately, then a model was built

- Tuning α can be crucial