

Homework 2 comments

Comment on role of k on ease of optimization and accuracy

- A neural network with the hidden layer has more parameters than the simple linear regression model.
- For smaller k, the model is able to optimize with larger learning rate. Similarly, models with higher k needs to be trained with a shorter learning rate. This could be due to low complexity of the models(number of parameters) with shorter k and hence is less likely to have the deviation from the direction of minima.
- The models with smaller k tend to saturate early. That is their accuracy stops improving while there is still a scope to improve. The models with higher k keeps on improving accuracy upto a better point. With the higher k, models gain the ability to learn more features that could indeed help in achieving better accuracy.

Comment on the difference between linear model and neural net.

- Linear models are easier to train than the neural networks. They have only one solution and hence the training configuration does not need to be very specific to achieve best results. The neural network due to their complexity have many possible solutions but not all of them are best. It is critical to choose good hyperparameters and initialization in order to get better results.
- Neural networks due to their complexity can give better solutions than the linear models. They can learn representations that can never be learned by linear models. This makes neural nets a better performer than the linear model.

Comment on the differences between logistic and quadratic loss in terms of optimization and test/train accuracy.

- Quadratic loss is able to optimize better over logistic loss. It gives a better gradient that works with a larger learning rate. This is evidence that gradients generated from quadratic loss function are more precise.
- The accuracy achieved with quadratic loss is better than the logistic loss.
- The main reason for this could be quadratic loss is more precise than the logistic loss.