Appendix: Parameter Sensitivity

Each of our non-linear extensions adds one additional hyper parameter to the LMNN problem. In section 6, we set these parameters by evaluation on a hold-out set. Here we explicitly examine their effect on the learned metric. For GB-LMNN, the new hyper-parameter is the regression tree depth. Figure 3(left) compares depths $4-7$ for several of the datasets evaluated in section 6. The figure depicts the ratio of $k$NN classification error for each depth setting to the $k$NN error of linear LMNN. GB-LMNN appears to be largely insensitive to tree depth within range.

For $\chi^2$-LMNN, the additional hyper-parameter is the size of the large margin. Figure 3(right) examines several margin values: 0.01, 0.05, 0.10, 0.15 and 0.20. The figure depicts the ratio of $k$NN classification error for each margin setting to the $k$NN error of the $\chi^2$ distance baseline. For all but two settings, the transformation learned by $\chi^2$-LMNN improves over the $\chi^2$ baseline, generally by a large extent. However, the margin size parameter is clearly important to achieving the best performance. Fortunately, the parameter seems to be well-behaved and easily set by evaluation on a hold-out set.