Meeting April 4 2024

Thursday, April 4, 2024 10:55 AM

(1)Multi-task Same input:

(a) no robust y-scaling
Impacted by training dataset outlier

(b) robust y-scaling
Not affected by training dataset outlier

MAE, RMSE,… mean prediction relevant metrics are the best

However, explained variance too small, to achieve global optimization, the variance in mean prediction is too small

(2)Simple-task Robust y-scaling
Best in explained variance
Worst in uncertainty

(3)Simple-task standard y-scaling
Best in uncertainty
Worse in mean prediction metrics

LVGP only takes single output variables
In
```
def lvgp_nll(p_quant, p_qual, lvs_qual, n_lvs_qual, dim_z, X_quant, X_qual, Y, min_eig, k, M):
Y = Y.reshape(-1, 1)
```

Singular input matrix problem

How to jointly reduce the dimension of numerical and categorical variables based on the response variable