Bayes error calibrated models

Evaluation dataset