Leverage based sampling for classification

DTU

Abstract

We validate the results of leverage based sampling for LS-regression, shown by Ma et al. [1]. We explore the possibility of using the leverage based sampling distribution from LS-regression on 2 class classification, and introduce two new approaches for sampling form an leverage distribution (important points).

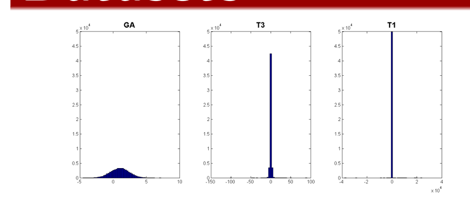
Motivation

The importance of sampling methods are initiated by very large datasets where it is not feasible to use all of the available data. This is illustrated by the rise in online access to video data. These data contain many frames that are basically the same and therefore redundant.

Research Questions

- Will the regression based sampling distribution improve our performance in classification?
- Can leverage based sampling be generalized to classification?

Datasets



Leveraging in general

LS-based Distribution

We wish to find the effect that a datapoint's class has on the predicted class for that datapoint.

$$\frac{\delta \hat{Y}_n}{\delta Y_n} \quad (1)$$

There is a closed form solution which is linear in y

$$\hat{\beta}_{OLS} = \left(X^T X\right)^{-1} X^T y$$

And prediction is also linear

$$\hat{Y}_n = X_n * \hat{\beta}$$

Therefore (1) is the coefficient

$$\frac{\delta \hat{Y}_n}{\delta Y_n} = X\left(X^T X\right)^{-1} X^T$$

Test results

T1.png

N=1000

p=10

Sensitivity Based Distribution

The target is again (1)

$$\hat{Y}_n = p(y|\bar{x},\bar{w}) \quad \bar{w} \text{ s.t. } \frac{\delta L}{\delta \bar{w}} = 0$$

Since 1 depends both directly and indirectly on y we see that

$$\frac{\delta}{\delta y}\frac{\delta L}{\delta \bar{w}} = 0$$

$$\Downarrow$$

$$\frac{\delta^2 L}{\delta y \delta \bar{w}} + \frac{\delta^2 L}{\delta \bar{w} \delta \bar{w}^T}\frac{\delta \bar{w}}{\delta y} = 0$$

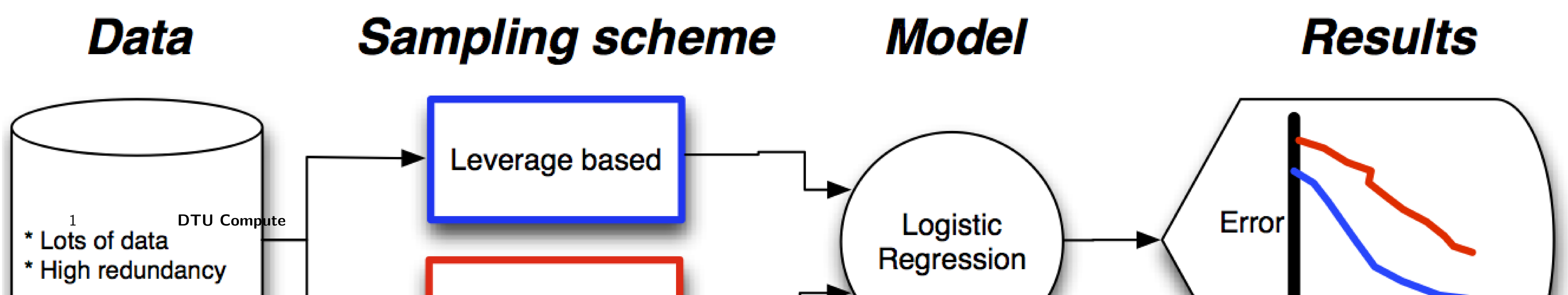
and from this we can get our leverage score (1)

$$\frac{\delta p(y|\bar{x}_n,\bar{w})}{\delta \bar{w}^T}\frac{\delta \bar{w}}{\delta y} = -\frac{\delta p(y|\bar{x}_n,\bar{w})}{\delta \bar{w}^T}\left[\frac{\delta^2 L}{\delta \bar{w}\delta \bar{w}^T}\right]^{-1}\frac{\delta^2 L}{\delta y \delta \bar{w}}$$

Uncertainty Based Distribution

bla

Process



Conclusion

itr doesnt work because the classification doesnt gain much from the collective distribution of both classes.