CS 419 : Introduction to Machine Learning **Spring 2022**

Lecture 8: February 9

Lecturer: Abir De *Scribe: Dhruv Shah*

8.1 Multi-Class Classification

Let y = 1,2,...,k be the labels of each class. Then using the probabilistic approach is easier as compared to using SVM. One would have,

$$P(y = i) = p_i = f({w_i}^T x) = \frac{e^{{w_i}^T x}}{\sum_k e^{{w_k}^T x}}$$

Now in the binary case we would have,

$$P(y = 1) = \frac{1}{1 + e^{-{w_1}^T x}}$$

$$P(y = -1) = \frac{1}{1 + e^{-{w_{-1}}^T x}}$$

Now as $P(y = 1) + P(y = -1) = 1$ we would have ${w_1}^T x = -{w_{-1}}^T x$
Hence in the binary case we model y as

$$P(y|x) = \frac{1}{1 + e^{-w^T xy}}$$

8.2 Stability of Classifiers

Let A be an algorithm which outputs a vector. Let S be the data set which would be the input for the algorithm. Then, $A(S)$ is called stable if

$$||A(S) - A(S')|| = O(\frac{1}{|S|})$$

where S' is the same data set S with only one element changed, which means that $|S| = |S'|$.

Now, if we use $y = w^T x + b$ as a classifier then we can see that by changing a point in the data set such that it no longer belongs to the convex hull of the remaining points, the classifier has a finite drift. This would mean that $y = w^T x + b$ is not a stable classifier. This would pose the following problems

- Not robust to outliers
- Generalisation Issues
- Privacy Issues

CS 419 : Introduction to Machine Learning **Spring 2022**

Lecture 8: February 9

Lecturer: Abir De *Scribe: Dhruv Shah*

Instability of the classifier can cause privacy issues because, by changing a single entry in the data set and observing the new and old model learned by the system, one could reverse engineer the new entry in the data set and hence breach the privacy.

We also observe the following points:
If $b \neq 0$

- Improves training accuracy
- Not necessarily improves test accuracy

If $b = 0$

- Does not improve training accuracy
- may improve test accuracy
- will be stable with regularizers