Train In Data

Instance Hardness Threshold

Instance Hardness - definition

- **Instance Hardness** is a measure of how difficult it is to classify an instance or observation correctly.
- **Hard instances** are observations that are hard to classify correctly.
- Class overlap is the principal contributor to instance hardness.

Instance Hardness - definition

The **instance hardness**, or in other words, the misclassification of an observations depends on:

- The learning algorithm used to model the task
- The observation's relation to other observations (class overlap)

Instance Hardness - illustration

The instances inside the oval represent **border points**, which have a **greater** degree of **hardness**, or in other words, are harder to classify.

Some instances are harder for some learning algorithms than for others.

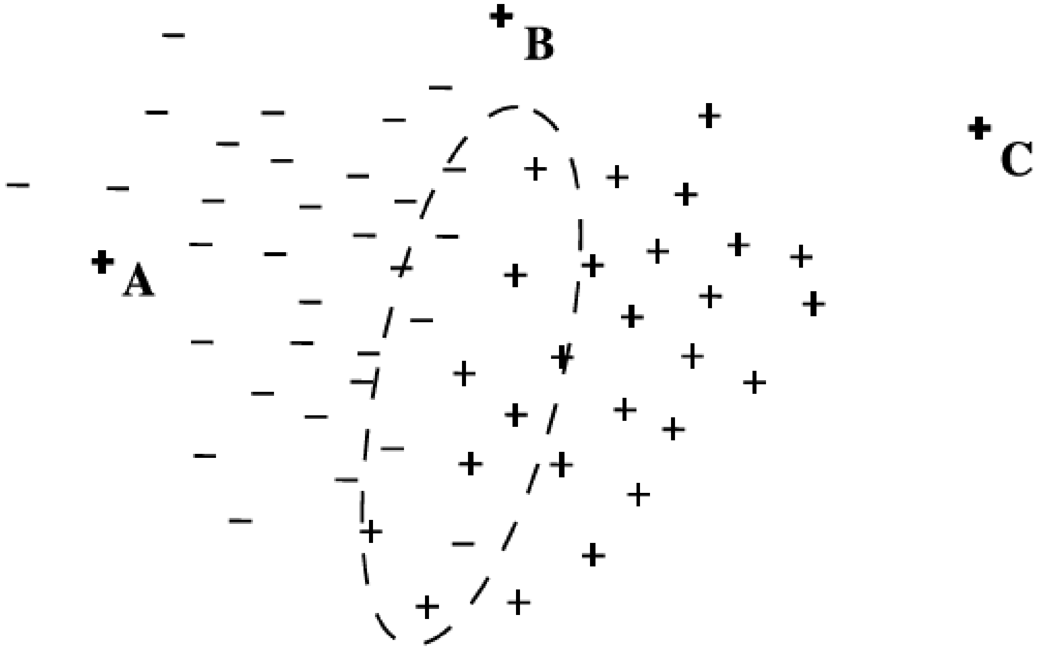


Image taken from D. Smith, et al: "An instance level analysis of data complexity." Machine learning 95.2 (2014): 225-256.

Instance Hardness - Probability

Fundamentally, instances that are hard to classify correctly are those for which the learning algorithm has a low probability of predicting the correct class label.

Probability
0.9
0.7
0.8
0.5
0.2

Instance Hardness - Probability

Fundamentally, instances that are hard to classify correctly are those for which the learning algorithm has a low probability of predicting the correct class label.

Probability
0.9
0.7
0.8
0.5
0.2

Instances hard to classify

Instance Hardness - Probability

Instance Hardness: probability of an observation of being miss-classified.

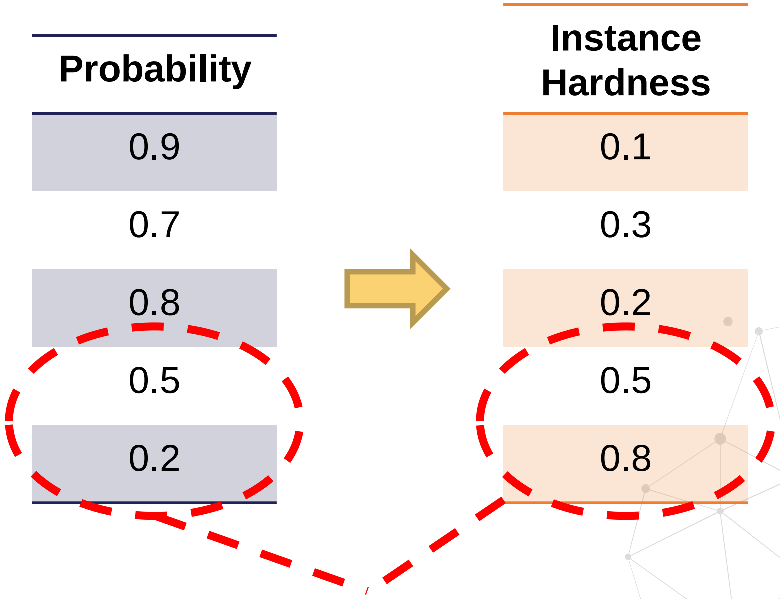
Instance hardness = 1 - probability

Probability		Instance Hardness
0.9		0.1
0.7		0.3
0.8	→	0.2
0.5		0.5
0.2		0.8

Instance Hardness - Probability

Hard instances:

- Hard instance metric is high
- The class probability is low



Probability
0.9
0.7
0.8
0.5
0.2

Instance Hardness
0.1
0.3
0.2
0.5
0.8

Instances hard to classify

Instance Hardness Filtering

- Simple idea: **remove hard instances** from data to reduce class overlap and thus, **increase class separation.**

- Remove instances with an instance hardness greater than a threshold.

How do we find the threshold?

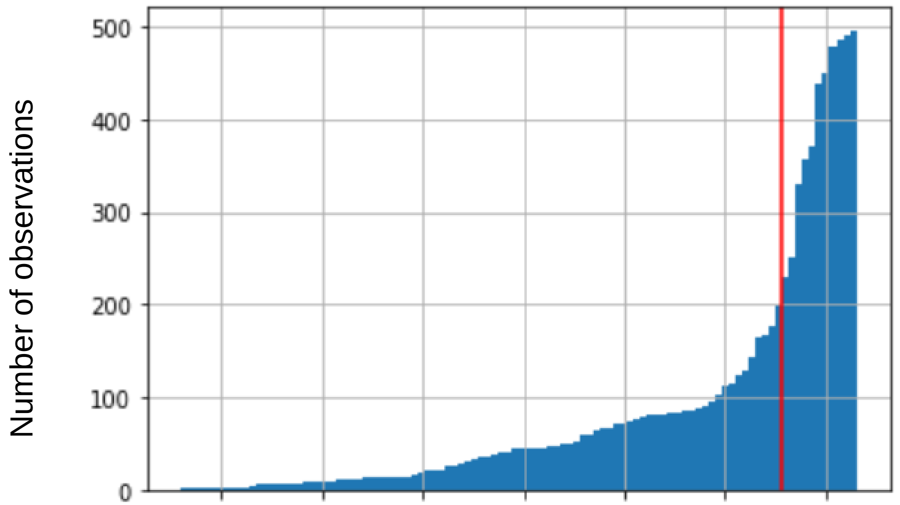
Instance Hardness Threshold

- Determine the threshold arbitrarily, like the authors of the method.

- Find threshold to match a desired balancing ratio, like imbalanced-learn

Instance Hardness Threshold

Cumulative distribution



Probability (of class to undersample)

To select as many observations from the majority, as those from the minority:

- Perc = $\left(1-\frac{desired\ \#\ observations}{\#\ observations\ of\ majority\ class}\right)\times 100$

- If desired # obs is10 and # obs majority class is 90 ➔ perc = (1 – 10/90) x 100 = 88.89

Instance Hardness Threshold

- If X(min)=10 and X(maj)=90 ➔ percentile = (1 – 10/90) x 100 = 88.89
- np.percentile(vector_of_probabilities, percentile)
- np.percentile(vector_of_probabilities, 88.8)
- The threshold is a probability
- Select observations from the majority class which probability > threshold

In summary

Instance Hardness Filtering

- Train a machine learning algorithm

- Determine the instance hardness

- Remove observations with high instance hardness (or equivalently, with low probability of class)

- If more than 2 classes ➔ 1 vs Rest approach to determine hardness

Instance Hardness Filtering

- Filter with the same algorithm that you intend to train.
- The beauty of instance hardness is that various tresholds can be used and compared.
- Instance hardness filtering was designed to improve classifier performance in general, not just for imbalanced datasets.

Train In Data

THANK YOU

www.trainindata.com