

Stellingen

behorende bij het proefschrift
Misclassification Bias in Statistical Learning
door Quinten Meertens

1. In contrast to what is sometimes claimed, classification errors do not cancel out by aggregation (Ch. 1).
2. Under the double sampling scheme, the calibration estimator outperforms the misclassification estimator as a correction method for misclassification bias, regardless of the base rate and the accuracy of the classification algorithm (Ch. 2).
3. When manually annotated data is scarce, the misclassification estimator often reduces misclassification bias more effectively than the calibration estimator (Ch. 3).
4. The estimation error of misclassification probabilities should not be neglected when correcting for misclassification bias (Ch.4).
5. Cross-border Internet purchases within the European Union cannot be estimated accurately by means of consumer surveys, as these surveys result in a severe underestimation (Ch. 5).
6. Smoothed variants of the AUC, such as softAUC, should not be regarded as poor model selectors for algorithms that rank (Ch. 6).
7. Domain experts may never be replaced by machine learning algorithms.
8. The economy is an organism. Therefore, national statistical institutes should aim for more than merely counting and weighting its cells.
9. An intern or a student should not be considered a cheap employee.
10. Practice makes perfect, but only with appropriate feedback.

Amsterdam, 28 April 2021