Probability of Using Overrule to Evaluate Overlap in Causal Inference

Author: Shukun Cheng
Supervisor: Rickard Karlsson
Responsible professor: Jesse Krijthe
Delft University of Technology

Background

- **Rule-based classification**: Classification method based on IF-Then statements.

Research Question

What is the performance of Overrule on identifying overlap for different types of datasets?

- How does Overrule work with identifying overlap?
- How to find feasible hyperparameters for Overrule?
- Is Overrule sensitive to outliers?
- How does Overrule scale as the number of features and samples increases in the dataset?

Methodology

Overrule[1] package is used as the start of the project.

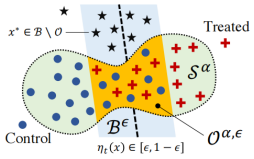
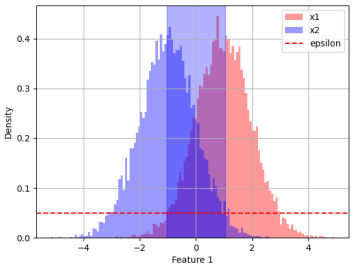


Figure 1. Overlap of two 1d normal-distributed classes

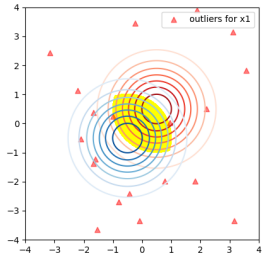
1. One-Class Support Vector Machine [2] to define the classes distribution.
2. Boolean Rule to extract the rules and evaluate the overlap.

Experimental Setup

- Simulated data is used.
- `scipy.stats` is used to calculate the data distribution in order to find true overlap.
- Confusion matrix and Intersection over Union (IoU) is used to test the performance.
- Grid search for choosing hyperparameters.
- Performance of Overrule with different numbers of samples.
- Uniform distributed data is used to simulate outliers.
- Test with Iris dataset.

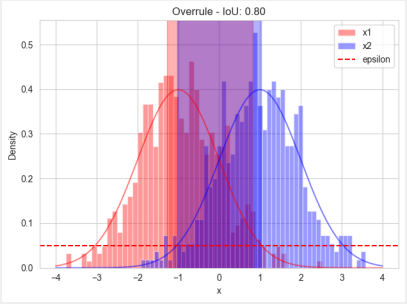


(a) True overlap estimator in 2d



(b) Outliers in 2d dataset

Results



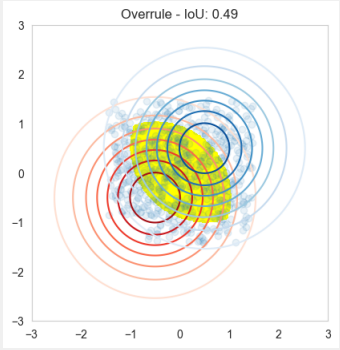
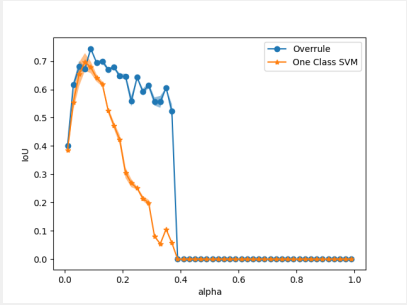
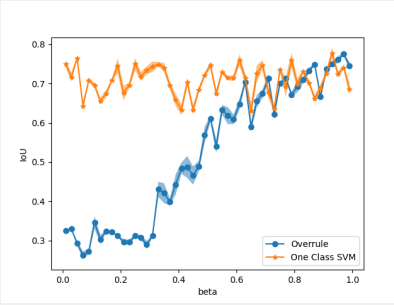


Figure 3. Basic experiments with default parameters

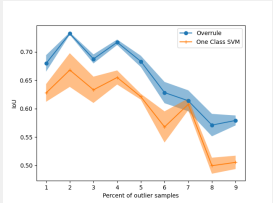


(a) α

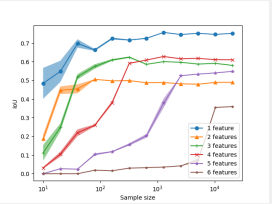


(b) β

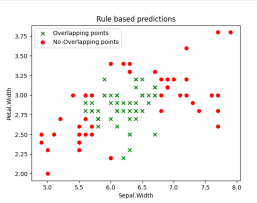
Figure 4. Grid search for each hyperparameter in 1d dataset



(a) IoU of Overrule with a different fraction of ourliers



(b) IoU with different numbers of samples and features (α = 0.1, β = 0.9)



(c) Overrule for Iris

Conclusion and future work

- α = 0.1, β = 0.9 works for most cases.
- Overrule can handle outliers effectively.
- More samples are needed for more features.

- Test with datasets with different distributions.
- Test with more than two classes and different variable types
- Test with real datasets.

References

[1] Michael Oberst, Fredrik Johansson, Dennis Wei, Tian Gao, Gabriel Brat, David Sontag, and Kush Varshney. Characterization of overlap in observational studies. In *International Conference on Artificial Intelligence and Statistics*, pages 788–798. PMLR, 2020.

[2] Bernhard Schölkopf, John C Platt, John Shawe-Taylor, Alex J Smola, and Robert C Williamson. Estimating the support of a high-dimensional distribution. *Neural computation*, 13(7):1443–1471, 2001.