Assessing ML Robustness to Sample selection bias

Zeeshan Khan
Z.Khan@student.tudelft.nl

Supervisor: Yasin Tapeli
Responsible Professor: Joana Gonçalves

TU Delft

Research Question:

How effective are minimax estimation techniques in mitigating sample selection bias?

Background

- Sample selection bias is a significant issue in machine learning, arising when the training and test sets originate from different distributions, impacting generalizability.
- This has led to research and implementation of domain adaptive models including minimax estimators.
- Despite the implementation of these models, it is crucial to assess their effectiveness..

Methodology

MODEL SELECTION:

- Minimax estimators RBA [1] and TCPR [2] were chosen for evaluation along with traditional models LR and SVM.

DATA GENERATION:

- Generate data and split into global (unlabeled), source (train) and target (test) sets as shown in figure 1.

BIASING:

- Induce bias in source domain.
  - Covariate Shift (figure 2b)
  - Class Imbalance
  - Survivorship Bias (figure 2c)
- Adapt using global domain.
- Test on target domain.

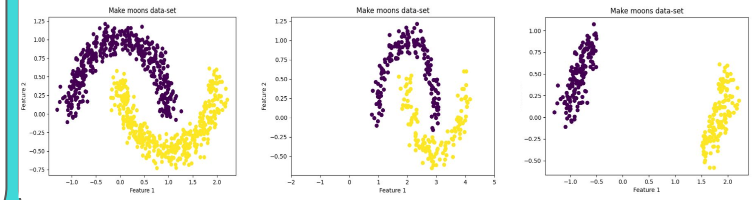
EVALUATION:

- Evaluate the chosen models using metrics.
  - F1-score
  - Log-loss,
  - AU-ROC score.

FIG 1 : DATA PARTITIONING

Splitting data into test(20%), train(30%), and unlabeled(50%) set

Fig2: Biasing the source data

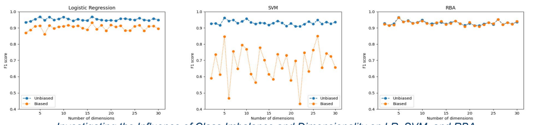


a) Unbiased data *b) Covariate shift* *c) Survivorship bias*

Results

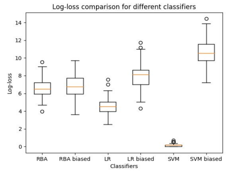
The results are shown figures 3-6

FIG3: CLASS IMBALANCE



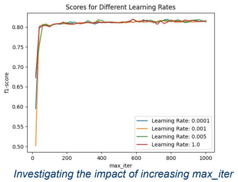
Investigating the Influence of Class Imbalance and Dimensionality on LR, SVM, and RBA

FIG4: COVARIATE SHIFT



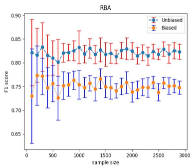
Box plots illustrating the impact of covariate shift on RBA, SVM, and LR

FIG6: PARAMETER TEST



Investigating the impact of increasing max_iter and learning rate on RBA's performance

FIG5: SURVIVORSHIP BIAS



Effect of survivorship bias and increasing number of samples on RBA performance

Discussion

- Robust Bias Aware Classifier (RBA) outperformed traditional supervised learning algorithms LR and SVM in the presence of sample selection bias.
- TCPR's performance could not be tested due to problems with its implementation
- Increasing max_iter parameter of rba beyond a threshold did not significantly improve results
- Changing the learning_rate parameter of RBA did not impact the performance of the classifier by much.

Future Work

- Testing the robustness of other minimax estimators on increasing domain distance between source and target sets.
- Testing other parameters of RBA and TCPR
- Testing the robustness of TCPR to bias on an implementation which does not throw exceptions.
- Using different evaluation metrics for evaluation

References

[1] Liu, Anqi, and Brian Ziebart. "Robust classification under sample selection bias." *Advances in neural information processing systems* 27 (2014).

[2] Kouw, Wouter M., and Marco Loog. "Target contrastive pessimistic risk for robust domain adaptation." *arXiv preprint arXiv:1706.08082* (2017).