

Contents

1. Backgrounds & Related Work
- 2. Research Questions**
3. Preliminaries
4. Analysis of Performance Influencing Factors in INVASE
5. Exploration of Potential Improvements to INVASE
6. D-ACES Framework Proposal and Analysis
7. Conclusion

Research Questions (1)

Analysis of Performance Influencing Factors in INVASE

1. Within the framework of INVASE, how do the following factors influence the explainer's ability to identify relevant features while discarding irrelevant ones?
 - (a) Does longer training improve the explainer's performance? If not, how does it affect the results?
 - (b) If longer training may lead to worse performance, what policy could be proposed to prevent selecting a suboptimal explainer?
 - (c) How sensitive is the explainer's performance to the choice of hyperparameter λ ?
 - (d) Does the presence of a baseline model consistently enhance the explainer's performance?