

Conclusion

Summary of Work

- ◆ Exploration of Influencing Factors in INVASE
 - Training time
 - hyperparameter sensitivity
 - Effect of baseline model
- ◆ Propose two policies
 - Early stopping policy
 - Post-training selection policy
- ▶ Synthetic dataset A
- ◆ Exploration of Potential Improvements to INVASE
 - Modify Input data structure
 - Replace baseline model with critic using unselected features
 - Revise loss function
- ▶ Synthetic dataset A
- ◆ New IFS method proposal and Analysis
 - Propose D-ACES
 - Compare with state-of-the-art IFS methods
 - Apply to high-dimensional data (syn B)
 - Apply to text data (IMDB)
 - Analyze limitations

Conclusion

Possible Future Work

- Incorporate response variable information in a tailored manner
- Detailed exploration of hyperparameter settings
- Exploration of the loss function settings for the explainer and the reward structure for the actor
- Enhancing adversarial resilience and robustness