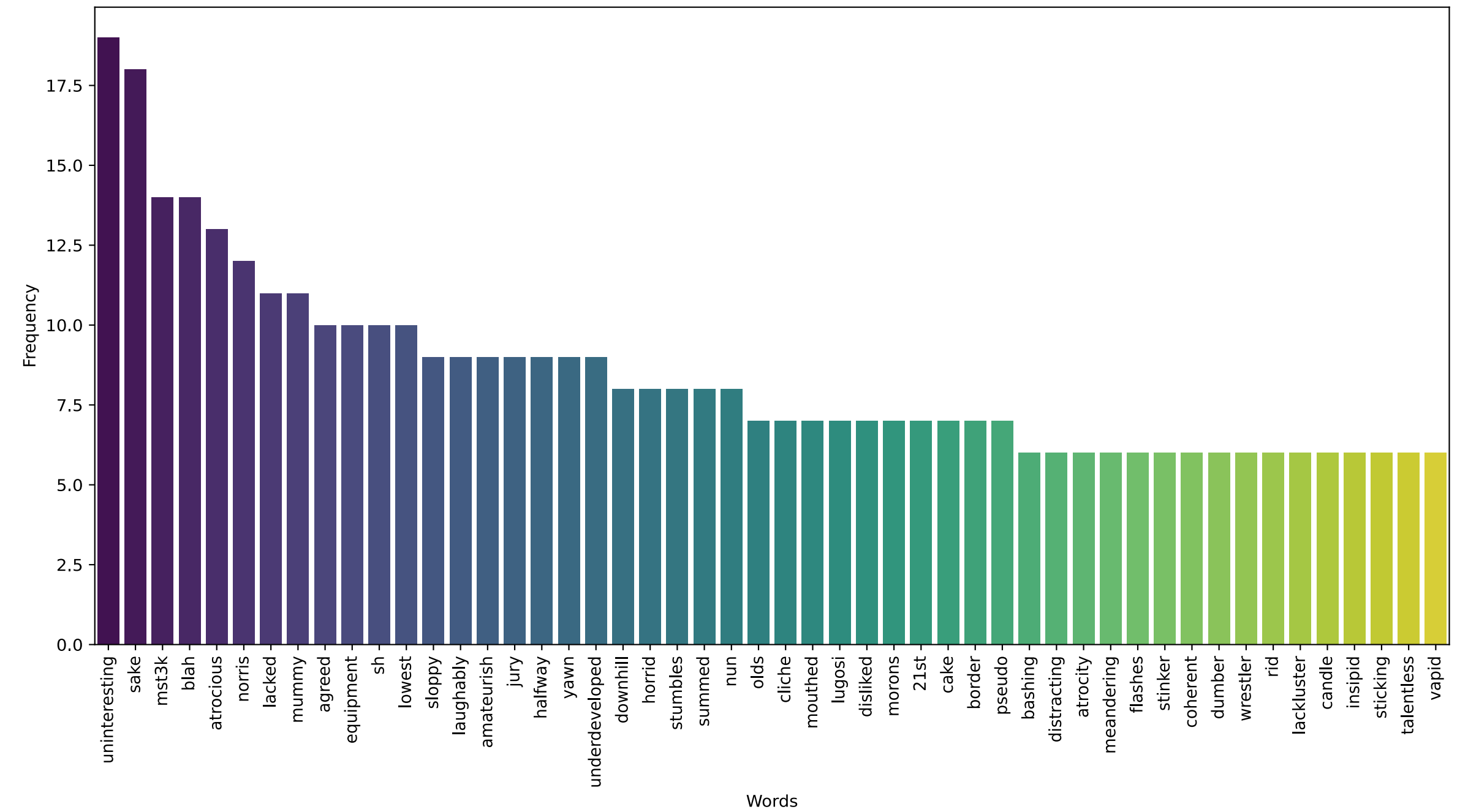
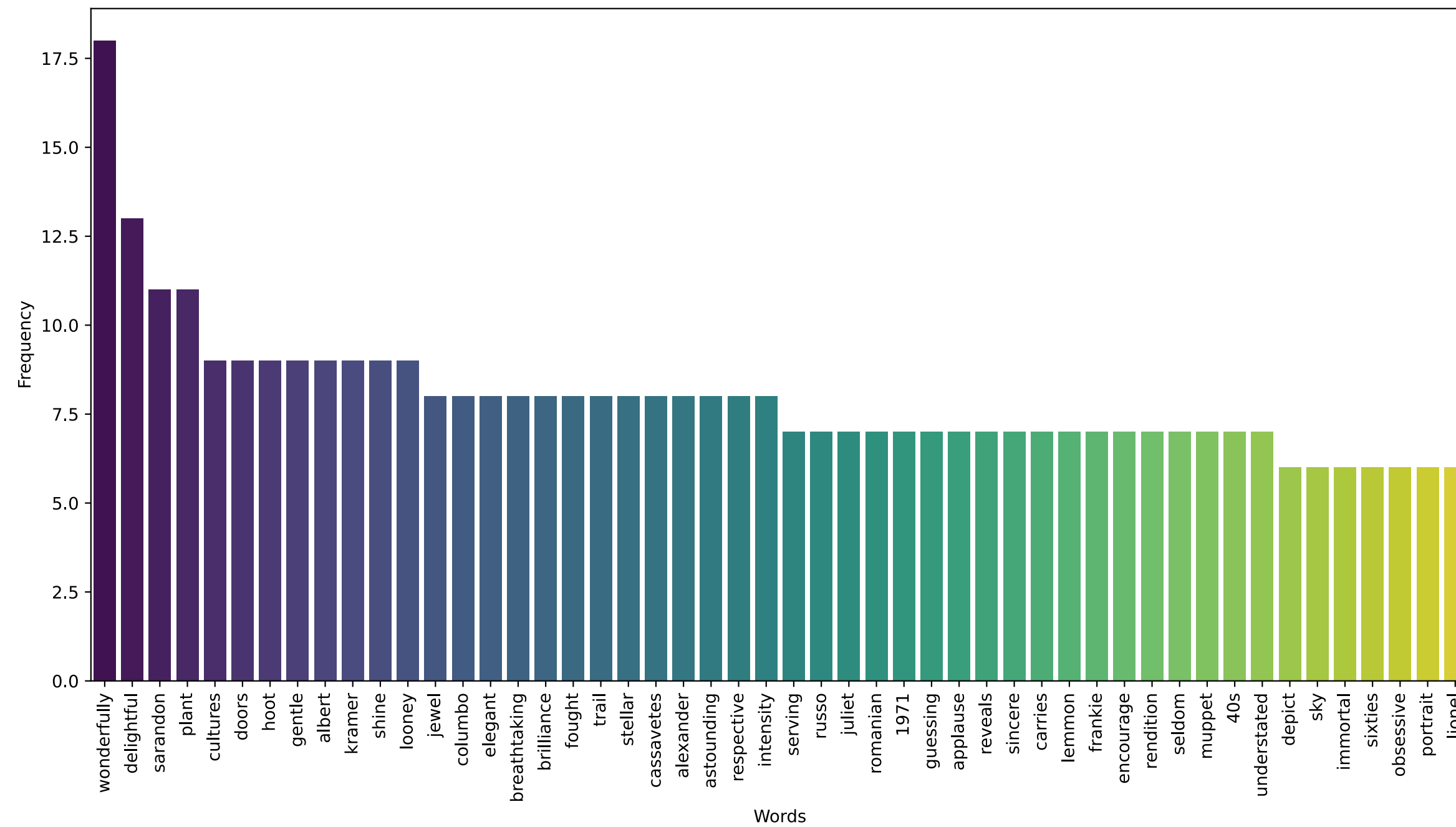


D-ACES Framework Proposal and Analysis

Application to IMDB



✓ D-ACES performs meaningful on text datasets, effectively learned meaningful distinctions between positive and negative reviews.

D-ACES Framework Proposal and Analysis

Limitation Analysis

- ◆ Lack of knowledge on hyperparameter settings
 - Know the trend of how to set
 - Still need several attempts to find an appropriate range
- ◆ Limited exploration of the loss function settings for the explainer and the reward structure for the actor
- ◆ Absence of considerations for adversarial resilience and robustness of the framework
 - As they are important properties in real-life application