Deception Detection

Group 66: **Akshat Chaw Parmar** 2022050 **Rishi Pendyala** 2022403 **Vimal Jayant Subburaj** 2022571

Abstract

Deception is a key strategy in Diplomacy, a negotiation-based game. This paper explores detecting deceptive messages using machine learning models. We preprocess game conversation data and evaluate Logistic Regression with TF-IDF embeddings and LSTM with GloVe embeddings for deception classification.

1 Problem Statement

Deception detection involves identifying misleading statements in communication. In Diplomacy, a negotiation-based game, deception is a key strategy where players use misleading messages to gain an advantage. Detecting such deception helps analyze strategic interactions and linguistic patterns. This research aims to develop a machine learning system to classify Diplomacy messages as truthful or deceptive based on linguistic and contextual features.

2 Plan and Approach

We follow a structured approach:

- Data Preprocessing: Cleaned messaged by removing punctuation and emojis while tokenizing text.
- Feature extraction: We convert text into numerical representations using TF-IDF and GloVe embeddings.
- Model Training: We train and compare two models:
  - Logistic Regression with TF-IDF
  - LSTM with GloVe embeddings
- **Evaluation:** We assess model performance using test accuracy and macro F1 score.

3 Results

The performance of our models is summarized as follows:

- **Test Accuracy:** 90.84
- **Macro F1-score:** 49.51

4 Results and Discussion

LSTM with GloVe embeddings outperforms Logistic Regression, capturing deeper contextual meaning. However, deception detection remains challenging due to subtle linguistic cues.

5 Future Work

- Enhance text representation with advanced preprocessing techniques.
- Address data imbalance using resampling, augmentation, and class weighting.
- Refine the LSTM-based approach used in the referenced literature for improved performance.
- Incorporate additional linguistic features and the referenced lexicon-based enhancements.
- Experiment with advanced deep learning architectures.
- Explore transformer models like BERT for better contextual understanding.

References

[1] Denis Peskov, Benny Cheng, Ahmed Elgohary, Joe Barrow, Cristian Danescu-Niculescu-Mizil, Jordan Boyd-Graber. 2020. *It Takes Two to Lie: One to Lie, and One to Listen*. University of Maryland, Cornell University.

[2] Sheila Mms. Diplomacy-NLP. *GitHub Repository*. Available: https://github.com/shielamms/Diplomacy-NLP

[3] Denis Peskov. 2020 ACL Diplomacy. *GitHub Repository*. Available: https://github.com/DenisPeskoff/2020_acl_diplomacy/tree/master