Slang Detection and Identification

Zhengqi Pei (zhengqi.pei@mail.utoronto.ca)
Engineering Science Program, University of Toronto
HAETAK Institute of Machine Intelligence

Zhewei Sun (zheweisun@cs.toronto.edu)
Department of Computer Science
University of Toronto

Yang Xu (yangxu@cs.toronto.edu)
Department of Computer Science
University of Toronto

Introduction

- Informal language such as slang is ubiquitous yet notoriously difficult for natural language systems.
- Downstream processing of slang necessitates the detection and identification of such linguistic phenomena in text.
- We take an initial step at automatic detection and identification of slang from natural sentences using established deep learning methods and show how linguistic features can be incorporated to offer both performance and interpretability.

- The system is able to learn and cope with two main categories of slang coinage:

1. **Newly extended senses**: existing words in the lexicon that develop novel slang senses distinct from their conventional senses.

That is a sick rock band.

2. **Newly created words**: novel word forms that do not exist in the standard lexicon.

That rock band is so gucci.

Tasks:

1. **Slang detection**:
   - Given a sentence, determine whether it contains at least one slang usage.
   - Posed as a **binary classification** problem.

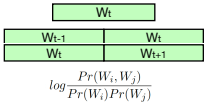
That salsa has quite a **kick** to it. ✓
The boy wants to **kick** the ball. ✗

2. **Slang identification**:
   - Given a sentence, determine the exact positions of slang usage.
   - Posed as a **sequence labeling** problem.

That salsa has quite a *kick* to it.
The boy wants to *kick* the ball.

Linguistic Features

- We incorporate a comprehensive set of linguistic features as input to our model to facilitate interpretable learning.
  - Word embeddings
  - Bigram embeddings (forward and backward)
  - Co-occurrence Statistics (e.g. PMI)
  - Part-of-Speech (POS) features:
    - POS embeddings (POS)
    - POS distribution (POS-Prob)
    - POS-Transfer (POSt)

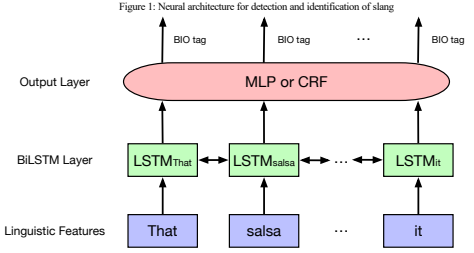


- Word embeddings are obtained from pre-trained Word2Vec vectors, while co-occurrence statics are estimated using Penn Treebank.

Model

- Formulate slang detection as a sequence labeling problem, where each token is labeled as slang or non-slang.
- Tags are defined following the BIO convention, where we create separate tags for existing and novel slang word forms.

Figure 1: Neural architecture for detection and identification of slang



Data:

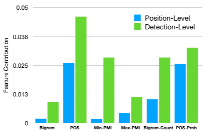
- **Online Slang Dictionary (OSD)**:
  - 15,000 example sentences from slang lexical entries as *positive examples* that contain slang usage.
  - 10,000 of which contain slang word forms that are not covered by the standard lexicon.
- **Wall Street News (2011-2016)**:
  - 15,000 formal newswire sentences as *negative examples* that do not contain slang.
  - Each sentence contains at most 20% Out-of-Vocabulary (OOV) tokens.

Results

Table 1: F1-statistics for detection and identification of slang

Model (Features)	Detection / Identification		
	Precision	Recall	F1 Score
Random Guess	0.500 / 0.026	0.500 / 0.483	0.500 / 0.050
Baseline - MLP	**0.989** / **0.624**	0.481 / 0.317	0.647 / 0.421
Full Features - MLP	0.943 / 0.542	0.684 / **0.461**	0.793 / **0.499**
Baseline - CRF	0.987 / 0.567	0.597 / 0.371	0.744 / 0.449
Full Features - CRF	0.952 / 0.550	**0.686** / 0.450	**0.797** / 0.495

Figure 2: Summary of Feature Ablation Analysis



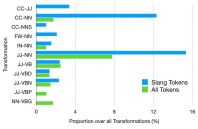
- Excluding Part-of-Speech (POS) features cause the most significant degradation in model performance while syntactic co-occurrence features also show significance when performing detection.

Table 2: Identification accuracy by slang coinage type

Model (Features)	Number of identified slangs	
	Existing word form	Novel word form
Baseline - MLP	194/523	35/199
Full Features - MLP	**267/523**	66/199
Baseline - CRF	227/523	41/199
Full Features - CRF	240/523	**83/199**

- Part-of-Speech transformations occur much more frequently in slang (53.94%) than ordinary usage (25.74%), and that the specific types of transformation is indicative of slang usage.

Figure 3: Common POS shift patterns for slang



Conclusion

- We present an interpretable neural framework for slang detection and identification by incorporating linguistic features and find Part-of-Speech (POS) to be prominent features for slang usage.
- Our work provides a basis for locating slang from its flexible and unconventional syntactic word uses and offers opportunities for slang processing in downstream tasks in natural language processing.