

Detecting Emerging Drug Slang and Code Language in Social Media Posts

Damir Cavar, Shruthi Nagappan, Kumar Koushik Telaprolu, Amy Stafford, Maria A. Parker, Danny Valdez

NLP-Lab, Indiana University, Bloomington, IN 47405, USA
School of Public Health, Indiana University, Bloomington, IN 47405, USA

Introduction

- Drug involved deaths are 250% higher than they were two decades ago.
- National surveillance systems identified social media conversations around emerging drugs including synthetic opioids (e.g., nitazenes), opioid-like compounds (e.g., kratom), and polydrug-use combinations (e.g., benzodiazepines, kratom, or ketamine combined with fentanyl).
- Issues:
 - Social Media Posts use coded language and (regional) slang undetectable with current NLP and ML models. Emergence of new drugs not part of the initial training data may escape detection before harm occurs.
- Prior Work Highlighting Surveillance Issues:
 - Tweets scraped prior to the closure of the Twitter API; YouTube Comments; BlueSky; Reddit and analyzed with BERTopic and LDA had high observed noise levels as slang terms typically comprised multiple meanings
- Observed need to identify:
 - New street names for drugs
 - New adverse reactions
 - Identification through surrounding context
- Solution:
 - Enhance LLMs with dynamic ontologies derived from domain seed knowledge for rapid drug triaging and identification.
- Seed Knowledge: NIDA A-Z Drug Dictionary
- Seed Knowledge: DEA Slang Terms and Code Words for Drugs

Objective

Build an OWL ontology of NIDA drug terminology and slang to improve LLM-based detection and classification of substance-abuse content on social media, detecting new slang terms.

Methods

Drugs A to Z

Drugs A to Z gives basic information on drugs with addictive potential, including how they are used, how they make people feel, and their health effects, including risk for substance use disorder. Treatment options for substance use disorders related to these drugs are also included.

Cannabis (Marijuana) Cocaine Fentanyl Heroin LSD (Acid) Meth

Jump to:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

```
1 drug:Opioid
2 drug:hasEffect effect:Analgesia ;
3 drug:hasEffect effect:Euphoria ;
4 drug:hasEffect effect:RespiratoryDepression
```

[Opioids] → Analgesia
 → Euphoria
 → Respiratory Depression

Results

- NIDA Drug Ontology V. 1
 - Full hierarchical organizations of concepts by hypernyms
- Task 1 Social Media Post Classifiers
 - Feature Selection and Identification using Logistic Regression
 - Binary classifiers: +/- is about drugs
 - BERT Tweet: F1 score .935
- Corpus of social media posts
 - 890,000 social media posts

Discussion and Future Work

- Experiments with LLMs and Ontology
 - Identification of all new slang words and emoji for drugs
 - Discover newly reported adverse reactions in social media
 - Identify new drugs based on context (e.g., Kratom)
 - Dynamically expand the NIDA Ontology
- Add additional languages to NIDA Ontology

References

- NIDA A-Z Drugs: <https://nida.nih.gov/research-topics/drugs-a-to-z>
- The Drug Ontology: <https://bioportal.bioontology.org/ontologies/D RON>
- DEA Slang Terms and Code Words: <https://www.dea.gov/sites/default/files/2018-07/DIR-022-18.pdf>

Acknowledgements

- PIs
 - Dr. Damir Cavar (NLP-Lab)
 - Dr. Danny Valdez (Public Health)
 - Shruthi Nagappan (NLP-Lab)
 - Anuj Prakash (NLP-Lab)

