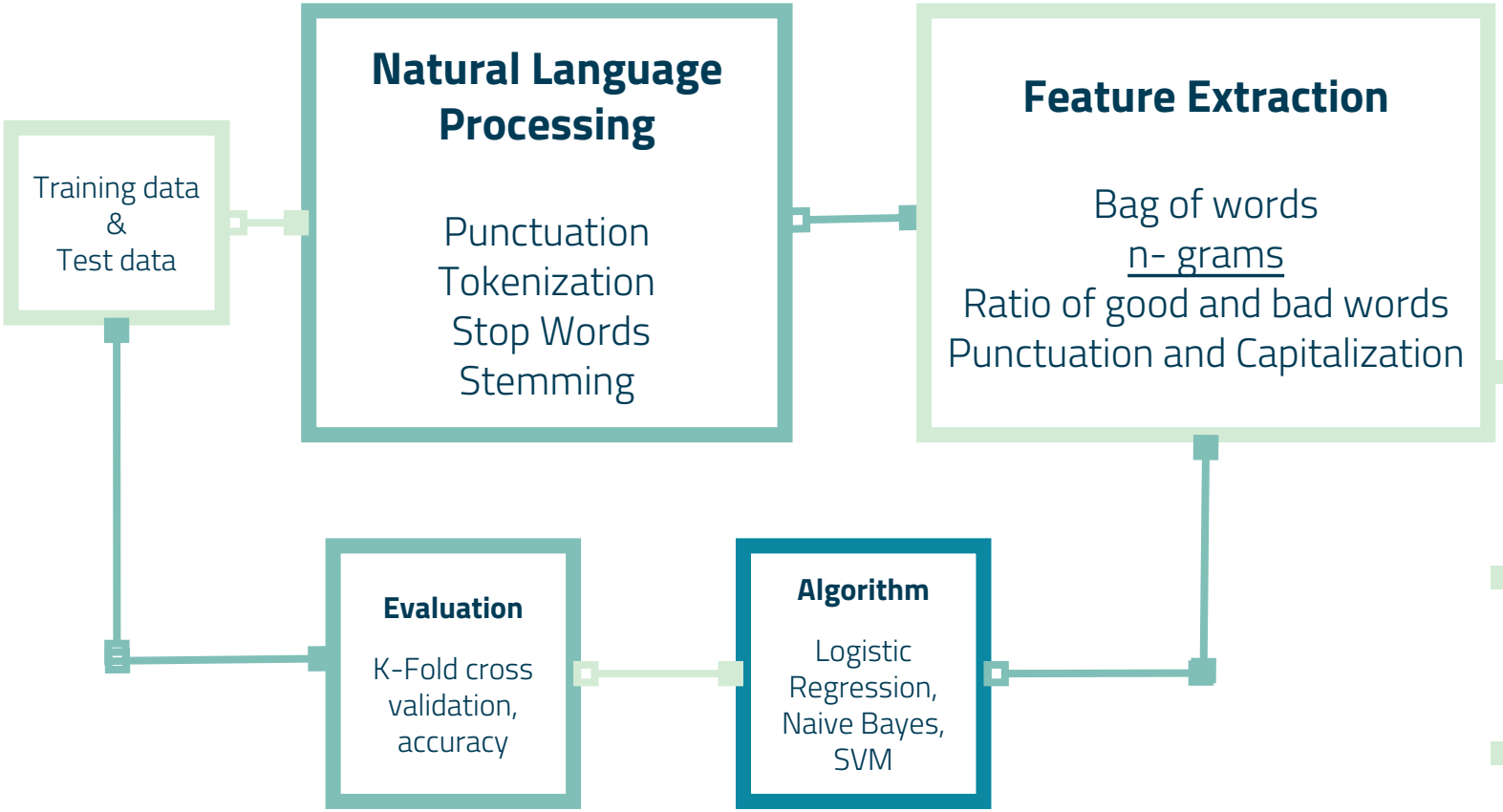
Insult Detection in Social Commentary

Natural Language Processing:

" NLP is a way for computers to analyze, understand, and derive meaning from human language in a smart and useful way "

Swastika Bhat & Xi(Lucas) Huang

ML Process


Training data
&
Test data
Natural Language Processing
Punctuation
Tokenization
Stop Words
Stemming
Feature Extraction
Bag of words
n- grams
Ratio of good and bad words
Punctuation and Capitalization
Evaluation
K-Fold cross validation, accuracy
Algorithm
Logistic Regression, Naive Bayes, SVM

Results

Bag of Word	Logistic Regression		Naive Bayes		SVM	
	K-Fold	Test Data	K-Fold	Test Data	K-Fold	Test Data
Uni grams	0.7234	0.7185	0.6975	0.6306	0.8206	0.702
1-6 grams	0.6944	0.6667	0.6944	0.6667	0.7861	0.7038

Logistic Regression	Ratio of Good words and Bad Words	Punctuation	Capitalization and Punctuation
K -Fold	0.7887	0.7185	0.6944

Tools, Open-Source and References

- Programming Languages: Python and Matlab
- Libraries: graphlab, nltk, numpy, pandas, scikit
- Dataset: Kaggle, Google list of bad words, list of positive words

https://github.com/Lucashuang0802/simple_detection_of_insults

http://cs229.stanford.edu/proj2013/Heh-DetectionOfInsultsinSocialCommentary.pdf