Cross Lingual Sarcasm Detection

RnD Project Presentation

Siddharth Bulia

Guide : Pushpak Bhattacharya
Aditya Joshi

Brief Overview of RnD Project

Goal

Applying concept of Cross Linguality in Sarcasm Detection

Research Part

- Sentiment Analysis.
- Cross Lingual Sentiment Analysis
- Sarcasm Detection

Read research papers, seminars about the above concepts.

Development Part

Implemented a basic Sarcasm Detector using the features described in research papers. Experimented with twitter datasets of different languages. Implemented a **Demo** for Cross Lingual Sarcasm Detection.

Contents

Introduction

Cross-lingual sentiment analysis has been reported to be worthwhile for resource-scarce languages.

The goal of this project is to validate if this holds in case of a related task: sarcasm detection.

Sarcasm detection is defined as prediction of a text as sarcastic or not.

Sentiment Analysis

Sentiment Analysis is the process of determining whether a piece of text is positive, negative or neutral.[1]

Sentiment analysis refers to the use of natural language processing, text analysis, computational linguistics to identify and extract subjective information in source material.[2]

Tagging text depending on the opinion expressed.

Drama was good.

Subjective
Positive

Drama was based
on a real incident.

Objective
Neutral

Drama was bad.

Subjective
Negative

*Images from clipart.co

[3]

Challenges

- Domain dependent
- Sarcasm
- Thwarted Expression
- Negation
- Implicit polarity

Cross Lingual Sentiment Analysis (CLSA)

Problem - Analyse sentiment in a new language.

Two ways to counter -

- In-language - Create new lexicons for the language, annotating documents for the language.

- Cross Lingual Approach - Using lexicons of another language, annotated documents of another language

Cross Lingual Sentiment Analysis

Task of predicting sentiment of corpus of T (Target) Language using the documents/resources/approaches available in S (Source Language).

Eg. Sentiment Analysis in Sanskrit using the lexicons developed in Hindi.

Sarcasm Detection in Italian using the annotated documents in English.

CLSA - Translation Methods

Mainly, two translation methods are used in cross lingual Sentiment Analysis.

1. Machine Translation - Using a well trained machine translation tool to translate data. Need to be done per sentence or per document. Examples . Google translate, Bing translate.

2. Bilingual Dictionary - Data is mapped word by word using a bilingual dictionary. Examples - Oxford Dictionary, Machine translation tool used word by word.

CLSA - Approaches

1. Learning the classifier in the source language feature space.
   + Classifier learnt over natural language.
   - Need to translate test document every time for testing.
2. Learning the classifier in the target language feature space.
   + No translation needed for any test document.
   - Classifier learnt over translated data. Error in translation can affect training.
3. Learning the classifier in common language space. Example - Using wordnet or a third language.
   + Can be used for languages where direct translation does not exist.
   - Translation required both testing and training purpose.

Sarcasm Detection

Sarcasm - ‘the activity of saying or writing the opposite of what you mean, or of speaking in a way intended to make someone else feel stupid or show them that you are angry’ (Macmillan dictionary, 2007)

‘a cutting, often ironic remark intended to express contempt or ridicule’ (the free dictionary)

‘saying the opposite of what you mean’ (Quintilien and Butler,1953)

Sarcasm Detections - Benefits

- **Personalized content** - People have mixed liking over sarcastic/non-sarcastic reviews.(the 'brilliant-but-cruel' hypothesis (Danescu-Niculescu-Mizil et al. 2009)). Identification of sarcastic reviews can therefore improve the personalization of content ranking and recommendation systems.
- **Improvement of Review Summarization and opinion mining** - Sarcasm, at its core, may harm opinion mining systems since its explicit meaning is different or opposite from the real intended meaning, thus averaging on the sentiment would not be accurate.
- **Human Computer Interface** - Human often use sarcasm in day-to-day life, hence sarcasm detection is essential in areas involving human computer interaction like digital personal assistance, etc.

Sarcasm Detection - Examples

- [I] love the cover (book)
- Wow, this thing makes me feel like I am in USA :p (GPS device)
- Are these iPods designed to die after two years? (music player)

First and second sentences appear to be compliment but on considering the context, they are examples of domain dependent sarcastic sentence.

In third sentence the sarcasm emerges from the naive-like question phrasing that assumes the general expectation that goods should last.

*Examples taken from a research paper by Tsur et al.[5]

Sarcasm Detection - Challenges

- Sarcastic sentences can use only positive words to express negative sentiment, hence normal bag of model approach can't be used.
- Most of the discussion of sarcasm detection happens in short and noisy text (tweets). The tweets are short and constrained to a length of 140 characters. The detection of sarcasm in such contextless tweets becomes very challenging.
- World knowledge is required. Example - 'You sing so nice that even Taylor swift will come to listen to you'. This can be considered as a compliment but if one knows about the Taylor swift, it is an exaggeration (hyperbole), hence it is actually an instance of sarcasm.

Cross Lingual Sarcasm Detection: First attempt

Cross Lingual Sarcasm Detection can be defined as using resources/documents/approaches available in Source language(S) to detect sarcasm in Target Language(S).

At first it seems same as Cross Lingual Sentiment Analysis. So to check about the validity of this assumption, we tried a few experiments over it.

Cross Lingual Sarcasm Detection - Need

- Cross Lingual Sarcasm Detection is as important as Cross Lingual Sentiment Analysis.
- With the increasing use of sentiment analysis in non-popular languages, problem arising in them need to be addressed one of which Sarcasm Detection.
- From analysis point of few, finding which languages are primarily used for sarcasm is a interesting topic.
- Developing complex NLP tools in other languages is a tough task.
- Majority of research done in Sarcasm Detection is in English language. To use that pre-existing research done in this field, cross linguality is required.

Cross Lingual Sarcasm Detection - Challenges

- Cross Lingual Sarcasm Detection bring up the problems of both areas, Cross Lingual Sentiment Analysis and Sarcasm Detection.
- Translation methods used for translation may not be reliable.
- Finding sarcasm in natural language is itself a tough job. Translation can completely change the structure/meaning of sentence.
- Idioms used in Sarcasm. Idioms used for sarcasm in language may not be used in other languages.
- Words used to show sarcasm in one language might not present in other one.
- Examples of challenges face while translating :
  - *“Yeah, today's morning is as awesome as that of 9/11.” - “हाँ , आज की सुबह 9/11 की है कि जैसे ही भयानक है।”*
  - *“I love being ignored.” - “मैं अनदेखा किया जा रहा प्यार करता हूँ।”*
  - *Person A - “Oh! Power is back.”. Person B : “Oh? Really? But the tube-light still took time to light-up!” Person A - ओह! बिजली आ गई है।. Person B : “ओह? वास्तव में? लेकिन ट्यूब प्रकाश अभी भी प्रकाश -अप करने के लिए समय ले लिया !”*

Development Work

Implemented an experimental setup and Web based User Interface.

1. Implemented In Language and Cross Lingual Sarcasm Detector various NLP and ML tools.
2. Developed a Web application which can be used directly from the browser to use the classifier learnt using the above setup. Web Application allow the user to give input in various languages including English, Italian, French, Hindi, etc.

Experimental Setup - Classifier and tools

Classifier - SVM with RBF kernel, C = 0.3 and gamma = 1.0.

Machine Translation tool* - Google Translate per sentence.

Bilingual Mapping tool* - Google Translate word by word.

Sentiment Analyzer - TextBlob library built over nltk.

Twitter API - To download Tweets given tweet IDs.

Stemmer - nltk and sumy library to stem each word.

Scikit-learn - Machine Learning tool to implement classifier.

*Used open source program translate shell which in turn uses google api.

Experimental Setup - Pre processing

- Replaced all kinds of emoticons with their corresponding meanings. Emoticons with same meaning are clubbed together.
- Words carrying same type of meaning are replaced with their common meaning. Example: yay, yaay, yaay are replaced with good.
- Improper words used in tweets in tweets are replaced by their corresponding full words. Example: ’r’ goes to ’are’, “don’t” goes to ”do not”,etc.
- Each exclamation and question mark symbol replaced with ” Exclamation ” and ” QuestionMark ”.
- After above steps, stemming is done word by word.

Experimental Setup - Features

- 1-gram, 2-gram
- Emoticons
- Contradiction in statement
- Number of question mark
- Number of Exclamation mark
- Subjectivity of the tweet
- Length of the tweet

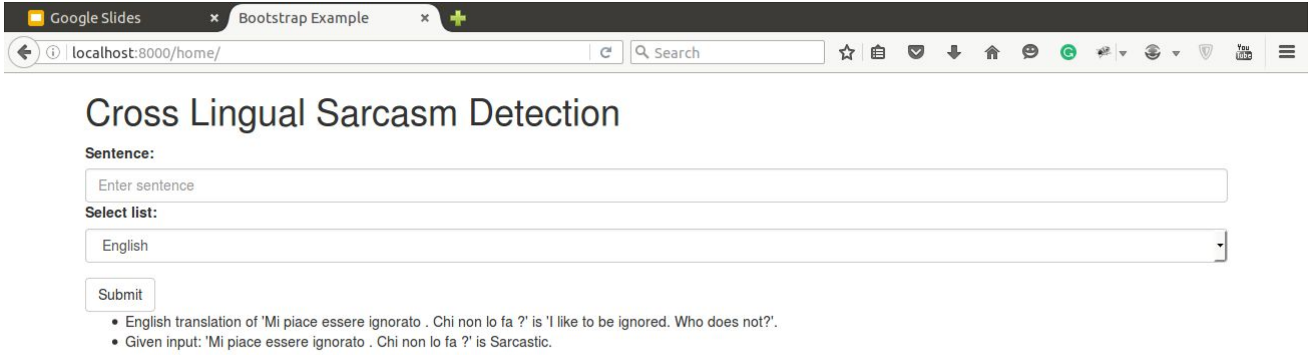
Experimental Setup - Translation methods

- Cross Lingual Sarcasm detection using Bilingual Mapping
- Cross Lingual Sarcasm Detection using Machine Translation

Experimental Setup - All approaches

- Learning the classifier in Source Language.
- Learning the classifier in Target Language.

Web Application Interface



Datasets

English Tweets: 5949 sarcastic tweets, 5912 non-sarcastic tweets

Source: Contextualized Sarcasm Detection on Twitter, David Bamman and Noah A. Smith, ICWSM 2015

Italian Tweets: 3131 sarcastic tweets, 3131 non-sarcastic tweets

Source: Italian Irony Detection in Twitter: a First Approach, Francesco Barbieri, Francesco Ronzano, Horacio Saggion, CLICIT 2014

Czech Tweets: 274 sarcastic tweets, 274 non-sarcastic tweets

Source: Sarcasm Detection on Czech and English Twitter, Tom´aˇs Pt´aˇcek, Ivan Habernal and Jun Hong, COLING 2014

Results

- A total of 18 experiments were conducted. 9 for each target language(Czech, Italian).

  For each target language:

- Skyline : In Language Sarcasm Detection experiment.
- Rest experiments differ from each other in terms of source language (English, Czech, Italian), Approach(Classifier learnt in source language or target) and type of translation used (Machine Translation, Bilingual Mapping).

Results - Target Language Italian

Skyline : In Language		
Precision	Recall	F-Score
0.80	0.79	0.79

*Approach : Classifier learnt in Source Language/Target Language

Source	Approach	Bilingual Mapping			Machine Translation		
		Precision	Recall	F-Score	Precision	Recall	F-Score
English	Target	0.56	0.68	0.62	0.51	0.82	0.63
English	Source	0.53	0.89	0.67	0.51	0.91	0.65
Czech	Target	0.44	0.66	0.53	0.44	0.67	0.53
Czech	Source	0.46	0.74	0.57	0.48	0.62	0.62

Results - Target Language Czech

Skyline : In Language		
Precision	Recall	F-Score
0.63	0.60	0.62

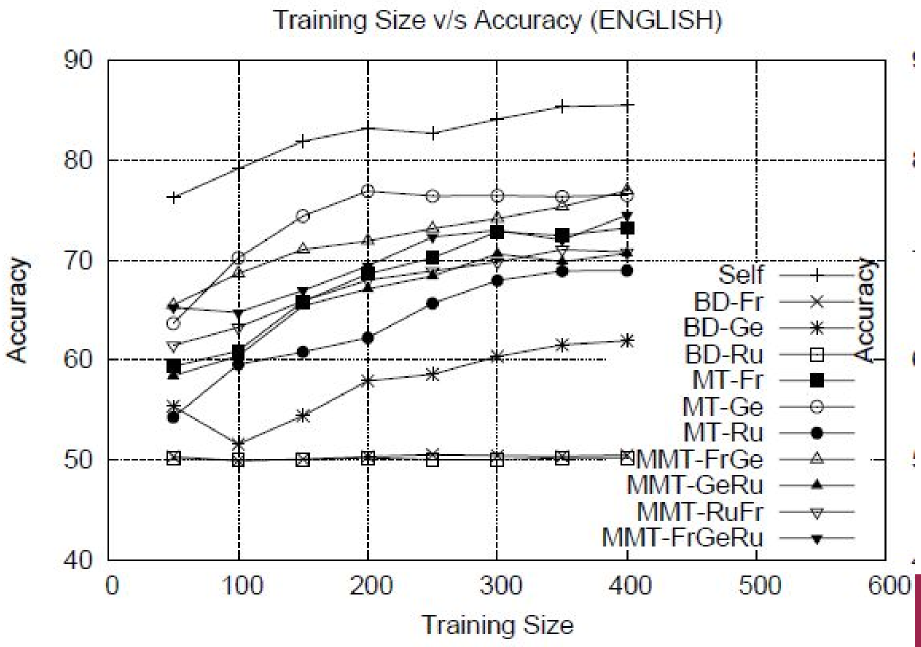
*Approach : Classifier learnt in Source Language/Target Language

Source	Approach	Bilingual Mapping			Machine Translation		
		Precision	Recall	F-Score	Precision	Recall	F-Score
English	Target	0.47	0.82	0.60	0.47	0.83	0.60
English	Source	0.47	0.90	0.62	0.47	0.88	0.61
Italian	Target	0.49	0.91	0.63	0.48	0.87	0.62
Italian	Source	0.49	0.94	0.64	0.48	0.95	0.64

CLSA vs CLSD (Target language: Italian, English)

*Comparing English CLSA with Italian CLSD.
Making assumption as Italian is close to english, the result should be close.

CLSA (English)		CLSD (Italian)	
In language	Cross Lingual	In Language	Cross Lingual
0.83 (approx)	0.78	0.79	0.67



*English CLSA results from balamurali et al. 2013

Observations

- The results obtained in Cross Lingual Sarcasm Detection are lower as compared to the results obtained through In Language Sarcasm Detection.
- Results obtained using Bilingual Mapping and Machine Translation are almost equal.
- Results obtained in Italian are better than the results obtained in Czech.
- On training Classifier on source language gives better results.
- CLSD is not able to perform as good as CLSA.

Error Analysis

Sentence : *“giugn arriv commodor 64. dann poste. #sarcastic”*

Translated to English using Bilingual Mapping : *“June arrives the Commodore 64. damned posed.”*
Classified as Non Sarcastic.

Word/Occurances	Sarcastic	Non Sarcastic
arriv*	2	5
damn*	24	30
pose*	0	3
Commodor*	0	0

Occurances of words more in Non Sarcastic Sentences than sarcastic sentence.
Hence, classifier classified it as non sarcastic

Error Analysis

Sentence : *“#Sport - Mondiali, la disfatta mai vista del Brasile. #non_sarcastic”*

Translated to English using Bilingual Mapping : *“#Sport - World Cup defeat ever in Brazil.”* Classified as Non Sarcastic.

Word/Occurances	Sarcastic	Non Sarcastic
world*	62	40
cup*	10	3
defeat	2	0
ever*	53	34

Occurances of words more in Sarcastic Sentences than non sarcastic sentence.
Hence, classifier classified it as sarcastic

Conclusion

- CLSD follow CLSA in terms of comparison with In Language experiments.
- Close language give better results in CLSD as compared to languages which are not close.
- Training on Natural language is better as compared to training of translated language.
- Techniques used for Cross Language Sentiment Analysis is not adequate for Cross Lingual Sarcasm Detection.

Future Work

The field of Cross Lingual Sarcasm Detection is untouched, there is definitely a lot of scope for improvement in this approach.

Applying best methods of both worlds, Cross Lingual Sentiment Analysis and Sarcasm Detection to get better results.

Extending the experiments over more languages, can give a better insight about closeness of languages.

References

1. Lexanalytics.com
2. Wikipedia.org
3. Clipart.co
4. Aditya Joshi, Sentiment Analysis: The World Within, Lecture in IIIT Hyderabad (https://www.cse.iitb.ac.in/~adityaj)
5. Oren Tsur et al, 2010. ICWSM - A great catchy name: Semisupervised recognition of sarcastic sentences in online product reviews. AAAI Conference.
6. B. Pang and L. Lee, “Opinion mining and sentiment analysis,” Foundations and Trends in Information Retrieval, 2(1-2),2008, pp. 1–135.

References

7. Aditya Joshi, Vinita Sharma and Pushpak Bhattacharyya, *Harnessing Context Incongruity for Sarcasm Detection*, **ACL 2015**, Beijing, China, July 26-31, 2015.
8. Aditya Joshi, 2013. Cross-lingual Sentiment Analysis, Seminar Report, IIT Bombay.
9. Ellen Riloff, Ashequl Qadir, Prafulla Surve, Lalindra De Silva, Nathan Gilbert, and Ruihong Huang. 2013. Sarcasm as contrast between a positive sentiment and negative situation. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pages 704–714. Association for Computational Linguistics.
10. Bing Liu. Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies, 5(1):1–167, 2012

References

11. AR Balamurali, Mitesh M. Khapra, and Pushpak Bhattacharyya. Lost in translation: Viability of machine translation for cross language sentiment analysis. In Alexander F. Gelbukh, editor, CICLing (2), volume 7817 of Lecture Notes in Computer Science,pages 38–49. Springer, 2013. ISBN 978-3-642-37255-1. URL http://dx.doi.org/10.1007/978-3-642-37256-8.