Project Report Document

May 3, 2019

1 Report Title

Russell Conjugation

1.1 Location of This Document

A copy of the document is available here:

https://docs.google.com/document/d/1wpPzjmsaRv0960PkE7JcSzHt0kr99xIgHH31HVL1DiM/edit?usp=sharing

2 Team Composition

2.1 Sponsor

Prof. Shlomo Argamon

2.2 Illinois Tech

Prof. Shlomo Argamon

3 Project Background

3.1 Sponsor's Line of Business

Research in computational methods for style-based analysis of natural language using machine learning and shallow lexical semantic representations, exploring application in intelligence analysis, forensic linguistics, biomedical informatics, and humanities scholarship.

3.2 Description of Sponsor's Problem

The use of subtly emotive language can bias interpretation of otherwise objective and accurate characterizations of people and events. Speechwriters and rhetoricians have used careful word choice to good effect since time immemorial. Bertrand Russell memorably encapsulated the idea in pseudo-conjugations such as:

- I am firm, you are obstinate, he is a pig-headed fool.
- I am righteously indignant, you are annoyed, he is making a fuss over nothing.
- I have reconsidered the matter, you have changed your mind, he has gone back on his word.

Here, pairs of words such as 'firm', 'obstinate' and 'pigheaded' are known as Russell or Emotive Conjugates of each other. As rhetoric in all media, both political and non-, has become increasingly polarized, so has, it seems, the use of such emotive language to pre-emptively destroy one's opponents and prop-up one's heroes.

4 Project Objectives and/or Goals

To build and deploy a prototype bias-revealing browser plugin, which will reveal hidden sources of emotive bias (Russell Conjugates) in online rhetoric. The long-term goal of this project is to investigate to the extent to which Russell conjugations are used to bias rhetoric, to develop tools to make readers aware of such rhetorical tricks, and to investigate how such tools affect readers' perceptions of bias and evaluation of information. The scope of the current project is to build a browser plug-in that automatically

a) Identifies source of polarization in a rhetoric, i.e., emotively connotated words,
b) identify and present Russell Conjugates of these emotive words as an option to the user.

4.1 Delivered Objectives

a) A pipeline, with two classifiers ordered sequentially, that classifies pairs of words as Russell Conjugates or otherwise – Overall precision about 0.733. Required higher.
b) A training dataset of Russell Conjugates for Model building.
c) A validation dataset.

4.2 Researched but Not Delivered Objectives

a) A classifier, to be used in the above-mentioned pipeline, that distinguishes synonyms from Russell conjugates well. Need higher precision – Work in progress.
b) Generating Russell Conjugates using a scaled version of a word2vec vector (v1), such that: (v1 *alpha) + word = Russell Conjugates of the word – Work in progress.

5 Project Deliverables

5.1 Description of Deliverables

a) A pipeline, with two classifiers ordered sequentially, that classifies pairs of words as Russell Conjugates or otherwise: This is delivered in form of a python code, which:
    a. Loads the prepared dataset
    b. Performs data preparation, which is conversion of each word into its word2vec vector and obtaining a difference vector of each pair of words.
    c. Building a Support Vector Machine classifier, with Radial Basis kernel, for identifying denotationally and contextually similar words.
    d. The pairs flagged as positive by the above classifier is then passed to another classifier that attempts at filtering out the synonyms from the other pairs of words, which are the required Russell Conjugates essentially.
    e. The cross-validated individual results along with the result of the pipeline on a completely different test set is also included.
b) A training dataset that can be used for further research in identifying or generating Russell Conjugates. The details of this dataset is as follows:
    a. The dataset has three attributes: x1, x2 and y. x1 and x2 form the pair of words which together needs to be seen by a model and predict whether the pair of words are

Russell Conjugates or not. Y label is the target variable- 1 value means the words are Russell Conjugates, and 0 means that the words are not.
    b. Total instances in the dataset: 4,576
    c. Total positive instances: 352
    d. Total negative instances: 4404

c) A validation dataset: It is a completely exclusive set of positive and negative examples, to be used for model validation. The details of this dataset are as follows:
    a. The dataset has three attributes: x1, x2 and y. x1 and x2 form the pair of words which together needs to be seen by a model and predict whether the pair of words are Russell Conjugates or not. Y label is the target variable- 1 value means the words are Russell Conjugates, and 0 means that the words are not.
    b. Total instances in the dataset: 44
    c. Total positive instances: 15
    d. Total negative instances: 29

5.2 Availability of Deliverables

The python code, training and validation dataset are available in a GitHub repository: https://github.com/Pragya2393-mishra/Identifying-Russell-Conjugations

6 Project Complications and Constraints

6.1 Complications

1. One major complication that we faced during the problem was building a classifier that could filter synonyms out of the list of contextually and denotationally paired words. The two approaches adapted for solving this problem were:

- Word2vec-vector based classifier: The idea behind this was to feed a classifier a set of data with pairs of words that were either synonyms or Russell conjugates. The classifier worked decently in filtering out the synonyms, but went ahead classifying some actual Russell conjugates as synonyms. This is a work-in-progress approach and the goal is to get precision for identifying Russell conjugates as high as possible.
- Sentiment lexicon-based identification: The idea behind this approach was to use a sentiment lexicon to assign emotional valence to each word in the pair, and then compare the difference of this score for the 2 words against a threshold to classify the pair of words as Synonyms or not. There were multiple issues with this approach:
    - Some sentiment-lexicons were not comprehensive enough to include most words
    - Lexicons like Senti-Wordnet were comprehensive but their emotional valence scoring was not reliable. For example, the word “riot” had an overall positive score.
    - Also, each word had different emotional valence score depending upon the context that they were used in. Identification of which context of the word to use, would need more work, and even if we get around it, the valence scores associated with each context of the words were not reliable as stated earlier.

2. Because of the above problem, a final end-to-end classifier, that identifies a pair of words as Russell Conjugates or not, is still in progress. Because of this, the planned Russell Conjugate

database, that would be formed by randomly pairing words in a thesaurus and passing through the classifier, could not be completed in time.

6.2 Constraint

Some primary constraints associated with this project were:

a) No prior computational work on Russell Conjugates. Thus, there is lack of references.
b) No prior dataset present to be used as reference.
c) Dependency on third party packages like Scikit-learn, Matplotlib etc.
d) Applicable to only English language.
e) Dependency on pre-trained word vectors by Google using word2vec

7 Project Execution

7.1 Agile Development: Epic, Story, Task Completed

Epic is a sub-project, Story is a collection of related tasks, and Task is the tangible work that you are going to complete.

Epics	Stories	Tasks	Progress
Phase 1	Literature survey	Literature survey on Computational work on emotive conjugations	
		Literature survey on media bias, its type and influence	
		Literature survey of Word Embeddings	
	Dataset creation	Collecting examples of Russell Conjugates by comparing different news coverage on the same topic	
		Filtering the examples to create positive examples of pairs of Russell Conjugates:	
		1. Ensuring pairs of words to have same POS and participle.	
		2. Include same of pairs of positive examples but with their order switched (for example firm, obstinate and then obstinate, firm)	
		Creating negative examples by randomly combining words from dictionary	
		Keeping the ratio of positive to negative example as 1:10 to increase variety of negative examples	
Phase 2	Data Preparation	Using pretrained word vectors using Google's word2vec	
		Converting each word pair from data into a difference vector using the pre-trained vectors	
	Modeling and Validation	Creating baseline linear classification models	
		Using oversampling to reduce bias due to unequal proportion of positive and negative examples	
		Performing dimensionality reduction using PCA to get a sense of highest variant directions for each class	
		Building classifier 1 - identifying contextually & denotationally similar pair of words	
		Validation and fine-tuning classifier 1	
		Building classifier 2- filtering out synonyms	
		Validation and fine-tuning classifier 2	
		Research other approaches to identify/generate Russell Conjugates	

Phase 3	Creating Russell Conjugate database	Decide a thesaurus to be used	
		Use the model to classify pairs of words in thesaurus to build a database	
Phase 4	Web Plug-in (Optional)	Create a module that analyses a web page's text to highlight emotionally connotated words	
		Use the database of Russell Conjugates created earlier to provide users with alternate words	
		Deploy the above two functionalities as a Web Plug-in	

Phase 1: The two stories in phase 1 were literature survey and dataset creation. Literature survey included looking computational work done previously on emotive conjugations. Apart from one paper [1], describing these conjugations from a literary and philosophical point of you, there was nothing in any research journals on this topic. And specifically, computationally, there was nothing. Our literature survey also included studying media bias , its type and influence to get a sense of how widely Russell Conjugates are used in practice and what are its impacts. This survey showed multiple news sources adopting Russell conjugation in implementing spin bias to their articles (mostly article headings), which affirms the importance of researching on this issue, and identifying and producing neutral alternatives automatically. Literature survey also included researching previous work done on word embeddings. Since the approach of this project was heavily based on usage of word embeddings, doing this survey made sense. This survey exposed some really good work done in the field of word representation. Methods like GLOVE, ELMo and BERT were studied along with their variations that included contextual and emotional representations too. Apart from literature survey, we all collected pairs of words that were Russell conjugates. This was done using multiple ways, and one of them were using the examples that we collected during literature survey on media bias. Other helpful ways were looking at previous interviews of Bertrand Russell, historical coverage of Bertrand Russell's interviews in various newspaper were other examples were added. After this, synonyms of these words were looked at to get more Russell conjugates. When all these positive examples were collected, there were reversed in order and added to the dataset as other instances of Russell conjugates. This is because "firm" and "obstinate" are as much Russell conjugates as "obstinate" and "firm". In creating the dataset, negative examples were needed too. Random words were clubbed together from an online dictionary to create negative samples. The ratio of positive and negative examples were kept to 1:10 to increase variety of negative examples.

Phase 2: Data preparation, modeling and validation were part of this phase. Data preparation here essentially included converting every word into word2vec vectors and then calculating difference vectors for each pair of word. The idea behind using word-vectors was to identify if there is a hyperplane that could divide Russell Conjugate pairs from non-Russell Conjugate pairs. After this, the need for a bias correction, like oversampling, was realized. This is because the number of negative samples were ten times as many as positive examples. If the data were to be

given without any bias correction, the model would be heavily biased. After oversampling the data, an SVM classifier with radial basis kernel and gamma=0.3 was used to build the classifier 1. There were many other classifiers tried, including logistic regression, Naïve Bayes, and SVM with a linear kernel, but the one that worked the best was the SVM with the radial basis kernel. This classifier, although the best amongst other classifiers, could not distinguish between Russell conjugates and synonyms. So, another level of filter was needed. This filter could be another classifier or it can be built directly using lexicon-based approach. Due to multiple complications, discussed in the "complications" section of the report, lexicon-based approach did not work at all. The classifier approach worked relatively better. This classifier, let's call it classifier 2, was again an SVM based classifier with a radial basis kernel. The optimum gamma value for this classifier was 0.75 and a separate dataset, with only Russell conjugate pairs and synonym pairs, was used to build this model. There were other models tried too, logistic, Random Forests, Gradient-Boosting, and Naïve Bayes, but SVM with RBF again worked the best. Please find below a figure summarizing this approach:

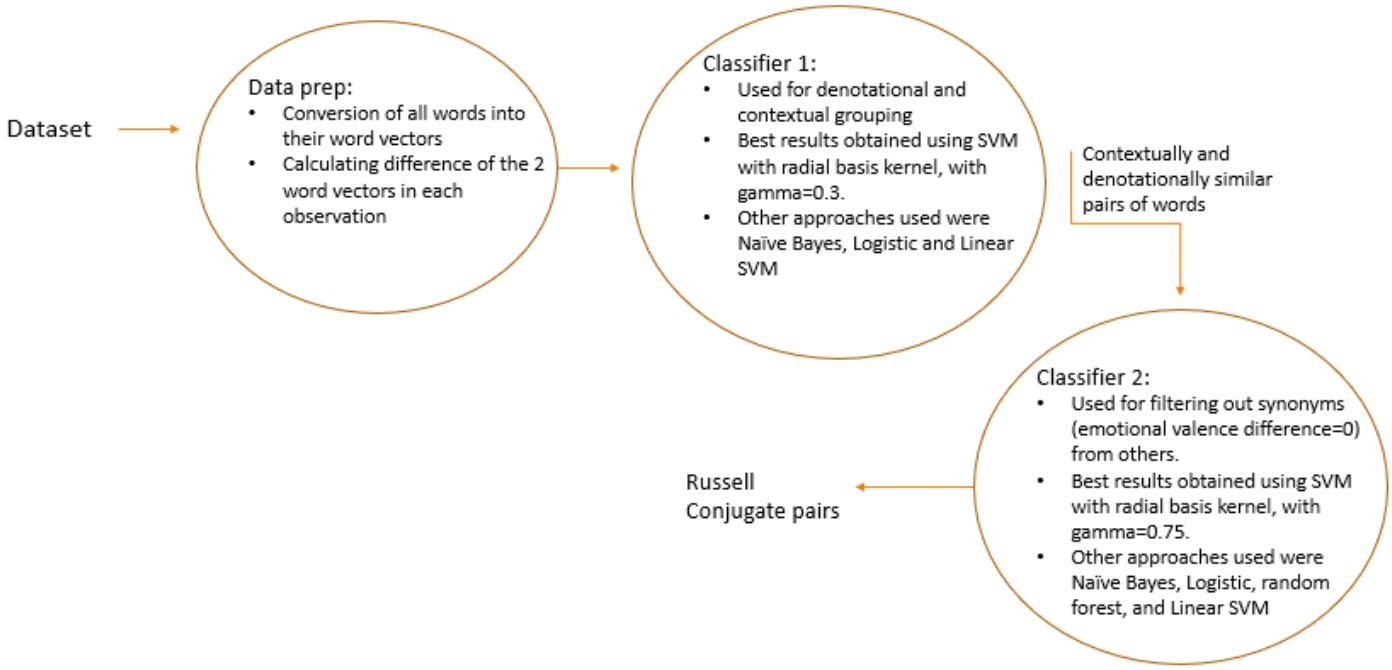

Figure 1: Approach 1

The above method described is referred as approach 1. Classifier 2 was able to filter out synonyms decently but it also wrongly classified some Russell Conjugates as synonyms. This was a trigger for looking at different options of identifying/generating Russell conjugates. And this is where we started exploring a method where we can explore and narrow down upon a vector v1, such that when scaled with a constant alpha and added to a word w1, provided Russell conjugates of the word w1. This method would be referred as approach 2.

Therefore, (v1* alpha) + word (w1) = Russell Conjugates of w1

To begin with, we started using the vector difference between the words “good” and “evil” as v1. Then we scaled alpha from 0.1 to 5 in steps of 0.2 and used the combination on a set of 50 words. The value of alpha that works relatively best is 0.1. The output of words in this case has a combination of potential good quality Russell conjugates, antonyms, modified original words, and other words with same lemma as the root word but different tense or participle. These outputs look promising and next step would include filters to remove other kinds of words generated such that only good quality Russell conjugates would be the output.

Phase 3: The main story in this phase was of building the Russell conjugate database. Since phase 2 is still progressing, the planned Russell Conjugate database, that would be formed by randomly pairing words in a thesaurus and passing through the classifier, could not be completed in time. This can definitely be attempted once the phase 2 is completed with desired results. The idea behind building a Russell Conjugate database was to serve as a backend database for the web plugin, that would eventually take a webpage as its input, identify emotionally loaded words and then provide a more neutral word (Russell conjugate) as an alternative for users to be able to read a more unbiased version of online news and rhetoric.

Phase 4: This was an optional phase where the main story was to develop the web plugin. This web plug-in would essentially be built to work on webpages and provide users with a more neutral versions of online news and rhetoric. Since the database built in phase 3 would have been the database on which this web-plugin would operate, the completion of this phase requires the completion of phase 3.

7.2 Findings, Results

The metric of importance in this project was precision. It was important to identify good quality Russell Conjugates than not finding any, for a particular word. The best results achieved so far, given this metric was the pipeline built using approach 1 (discussed in the previous section). Please find the results below:

Classifier-1 10-fold cross-validated results:

	Models			
	Logistic	**Naïve Bayes**	**SVM-Linear**	**SVM-RBF**
Accuracy	0.473881029	0.892365379	0.559169891	0.940789662
F-1	0.326444474	0.685981865	0.361925812	0.791424944
Precision	0.432427529	0.664648373	0.444875423	0.926099522

Recall	0.275879996	0.722130443	0.342845562	0.741192416

Table 1: Classifier-1 10-fold cross-validated results

Classifier-2 10-fold cross-validated results:

	Models				
	Naïve Bayes	**SVM-Linear**	**GradientBoosting**	**RandomForest**	**SVM - RBF**
Accuracy	0.646948718	0.526809955	0.569769231	0.58741629	0.706348265
F-1	0.640339407	0.435208772	0.541194063	0.535540457	0.601182652
Precision	0.655236225	0.425913644	0.547334208	0.551725413	0.608676886
Recall	0.642592114	0.475737628	0.54043166	0.536755369	0.600713012

Table 2: Classifier-2 10-fold cross-validated results

Results of the complete pipeline on the Validation dataset:

- Accuracy: 0.818
- F-1 score: 0.733
- Precision: 0.733
- Recall: 0.733

8 Project Support

Since this research concludes alongside the end of my graduate studies, my options to support this project may be limited. Nevertheless, I intend to keep supporting this project in any way I can.

9 Project Retrospective

Some of the most important lessons learned during the project are :

1. Never giving up and powering through- sometimes the desired results may be evading us. But if we keep trying, with different approaches or perspective, we can eventually get to our goal

2. Summarizing results: The practice of summarizing and updating the work and results of a particular, every week, was a very good practice of essentially using business language in talking highly technical content.
3. Power of literature survey: Before this project, I had not ever done such an extensive literature survey on any topic. Such efforts helped me develop the practice of efficiently and quickly doing my research on any topic, before working on it.

10 References

[1] R. Alexander, "Three in a Row: or, from Evaluative Lexis to Conjugating Adjectives," *AAA: Arbeiten aus Anglistik und Amerikanistik,* vol. 38, no. 2, pp. 203-211, 2013.

11 Document Revision History

5/3/2019 – Pragya – Initial Version