CogALex-V Shared Task: GHHH – Detecting Semantic Relations via Word Embeddings

Mohammed Attia[1], Suraj Maharjan[2], Younes Samih[3], Laura Kallmeyer[3], and Thamar Solorio[2]

[1]Google Inc.
New York City
attia@google.com

[2]Department of Computer Science
University of Houston
smaharjan2@uh.edu, solorio@cs.uh.edu

[3]Department of Computational Linguistics
Heinrich Heine University
{samih, kallmeyer}@phil.hhu.edu

Tasks

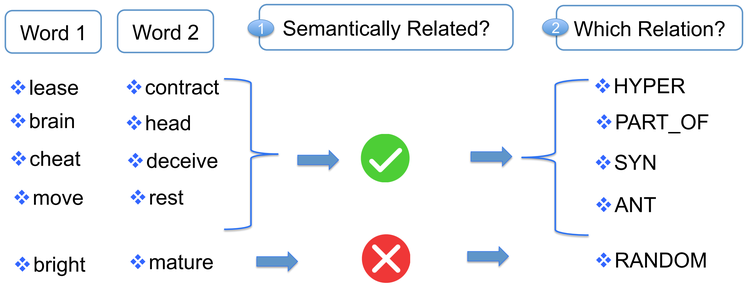

Fig 1: Tasks Overview

Dataset

Dataset	Task 1		Task 2					Word Pairs	Target words
	TRUE	FALSE	HYPER	PART_OF	SYN	ANT	RANDOM		
Training	826	2,228	255	163	167	241	2,228	3,054	318
Test	1,201	3,059	382	224	235	360	3,059	4,260	429

Table 1: Data Distribution

Projection of Word vectors

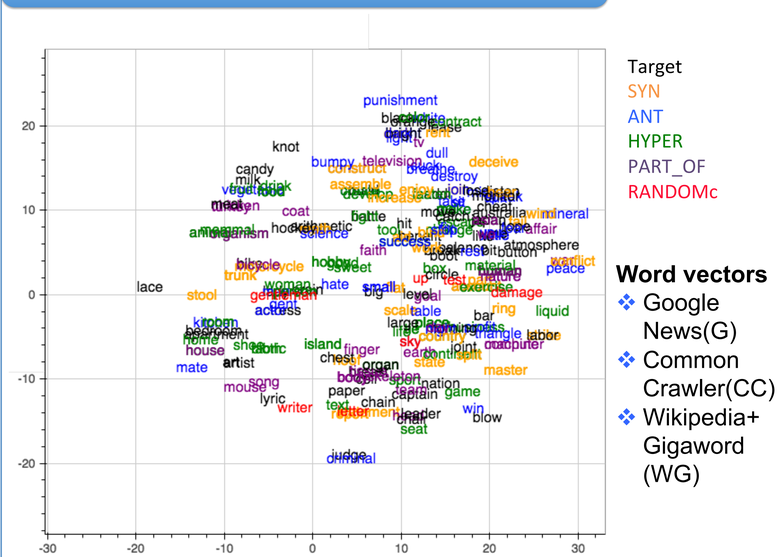

Fig 2: 2D Projection of Word Vectors

Methods

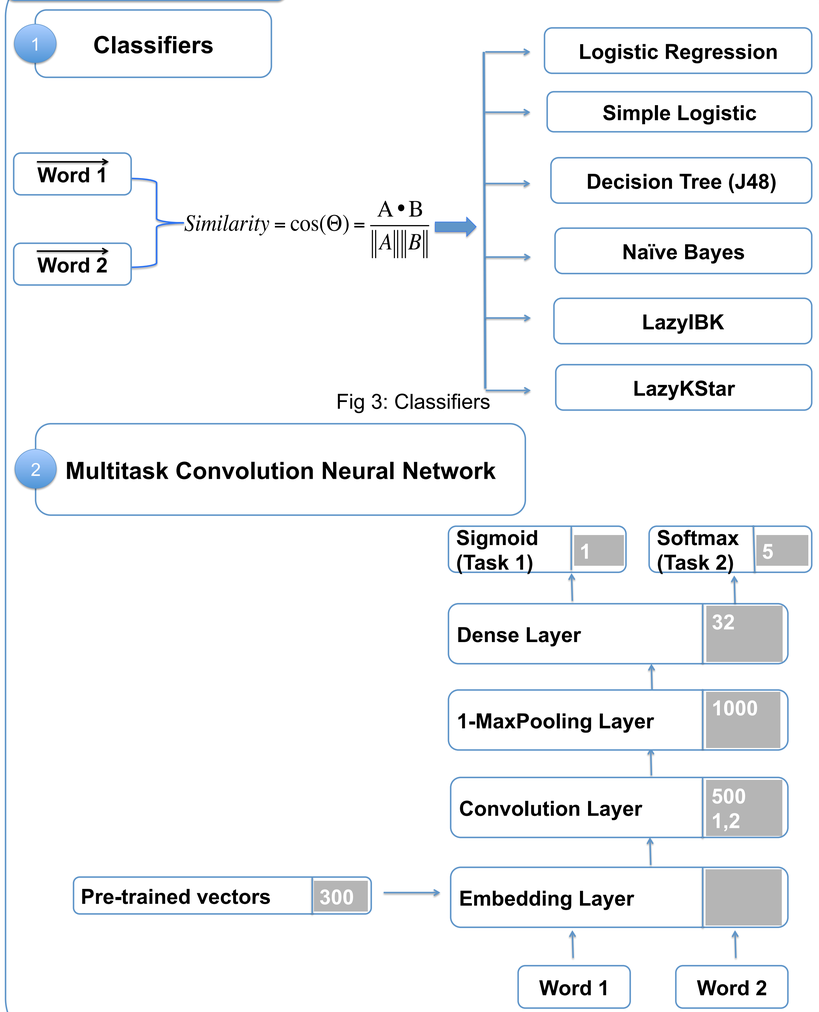

Fig 3: Classifiers

Fig 4: Multitask CNN Architecture

Comparison of Classifiers

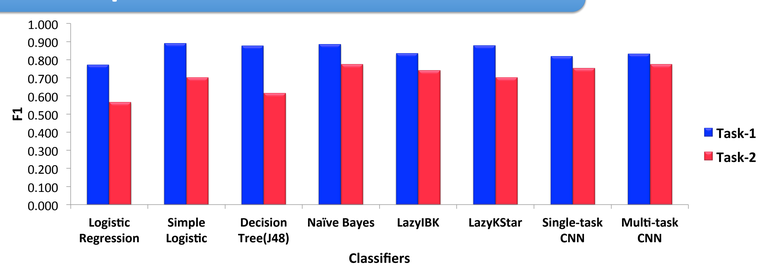

Fig 5: Comparison of ML Classifiers

Comparision of Word Vectors

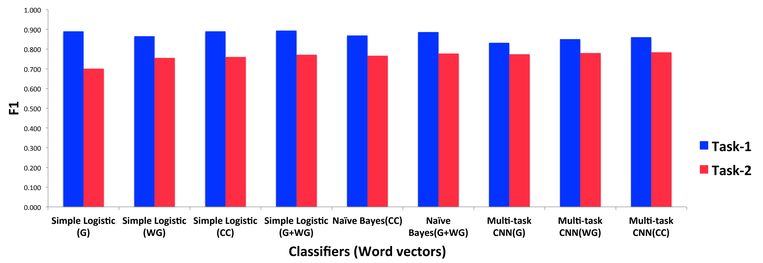

Fig 6: Comparison of Word Vectors

Effect of Unrelated Pairs

Limit	TRUE(F1)	FALSE(F1)	Average	Diff
1	0.918	0.795	0.883	0.123
2	0.891	0.865	0.881	0.026
3	0.866	0.891	0.880	0.025
4	0.836	0.901	0.875	0.065
5	0.822	0.914	0.882	0.092
No Limit	0.793	0.931	0.894	0.138

Table 2: Results for Different Limits of Unrelated Pairs

Shared Task Performance

Method	Task-1 (F1)	Task-2(F1)
Simple Logistic	0.790	0.287
Multi-task CNN	0.710	0.423

Table 3: Final Results on Test Set

Teams	GHHH	Mach5	LexNET	ROOT18	LOPE	HsH-Supervised	CGSRC
Task 1 (F1)	0.790	0.778	0.765	0.731	0.713	0.585	0.431
Task 2 (F1)	0.423	0.295	0.445	0.262	0.247	_	0.252

Table 4: Shared Task Results (1st, 2nd, 3rd)

Conclusions

- Cosine similarity scores between word vectors can reveal semantic relationship between them
- Multitask CNN approach is better than single task CNN approach
- Ambiguity makes the task difficult

CogALex-V , COLING Workshop
December 12, 2016
Osaka, Japan