Capturing Discriminative Attributes

Project Proposal
HLE Fall 2022

Lucía Urcelay
Rasmus Siljander
UPC - Master in Artificial Intelligence

Motivation

1. Find methods to model semantic representations of word vectors
2. Issues with similarity approaches [1]
   a. inter-annotator agreement tends to be low
   b. the small size of some of the most popular datasets
   c. subjective similarity scores have limitations when it comes to task-specific applications
3. Goal: find another way to approach the semantic representation task other than similarity computation

UPC

[1] Manaal Faruqui, Yulia Tsvetkov, Pushpendre Rastogi, and Chris Dyer. 2016. Problems with evaluation of word embeddings using word similarity tasks. In First Workshop on Evaluating Vector Space Representations for NLP (RepEval 2016).

Semantic difference detection

- A system for basic language understanding should be able to detect when concepts are similar to each other, but also in what way concepts differ from each other
- Modelling semantic difference between two (related) words can help capture individual aspects of meaning

Semantic difference detection

- Applications of semantic difference task:
  - Automated lexicography (automatically generating features to include in dictionary definitions)
  - Conversational agents (choosing lexical items with contextually relevant differential features to create human-like dialogs)
  - Machine translation (where explicitly taking into account semantic differences between translation variants can improve the quality of the output)

Discriminative Attribute detection

SemEval 2018 Task 10: Capturing Discriminative Attributes

Word 1	Word 2	Attribute	Discriminative?
Apple	Banana	Red	

Discriminative Attribute detection

SemEval 2018 Task 10: Capturing Discriminative Attributes

Word 1	Word 2	Attribute	Discriminative?
Apple	Banana	Red	✓

Discriminative Attribute detection

SemEval 2018 Task 10: Capturing Discriminative Attributes

Word 1	Word 2	Attribute	Discriminative?
Apple	Banana	Red	✓
Gloves	Pants	Wool	

Discriminative Attribute detection

SemEval 2018 Task 10: Capturing Discriminative Attributes

Word 1	Word 2	Attribute	Discriminative?
Apple	Banana	Red	✓
Gloves	Pants	Wool	✗

Discriminative Attribute detection

SemEval 2018 Task 10: Capturing Discriminative Attributes

Word 1	Word 2	Attribute	Discriminative?
Apple	Banana	Red	
Gloves	Pants	Wool	
Spider	Elephant	Wings	

Discriminative Attribute detection

SemEval 2018 Task 10: Capturing Discriminative Attributes

Word 1	Word 2	Attribute	Discriminative?
Apple	Banana	Red	✓
Gloves	Pants	Wool	✗
Spider	Elephant	Wings	✗

Data

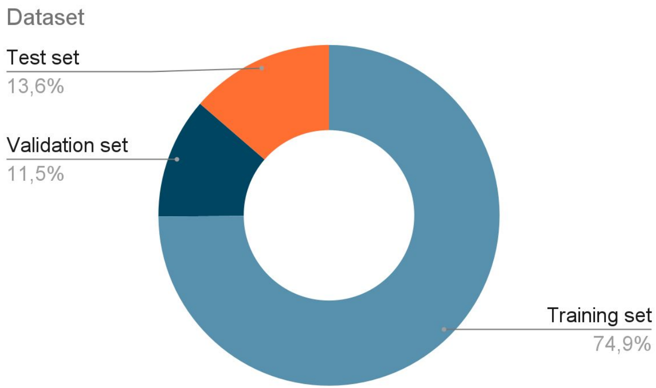
Dataset structure

Instance	Word 1	Word 2	Attribute	Label
1	Apple	Banana	Red	1
2	Gloves	Pants	Wool	0
3	Spider	Elephant	Wings	0
...	...	...	...	. . .
n				

Data

Total n. of instances: 22884

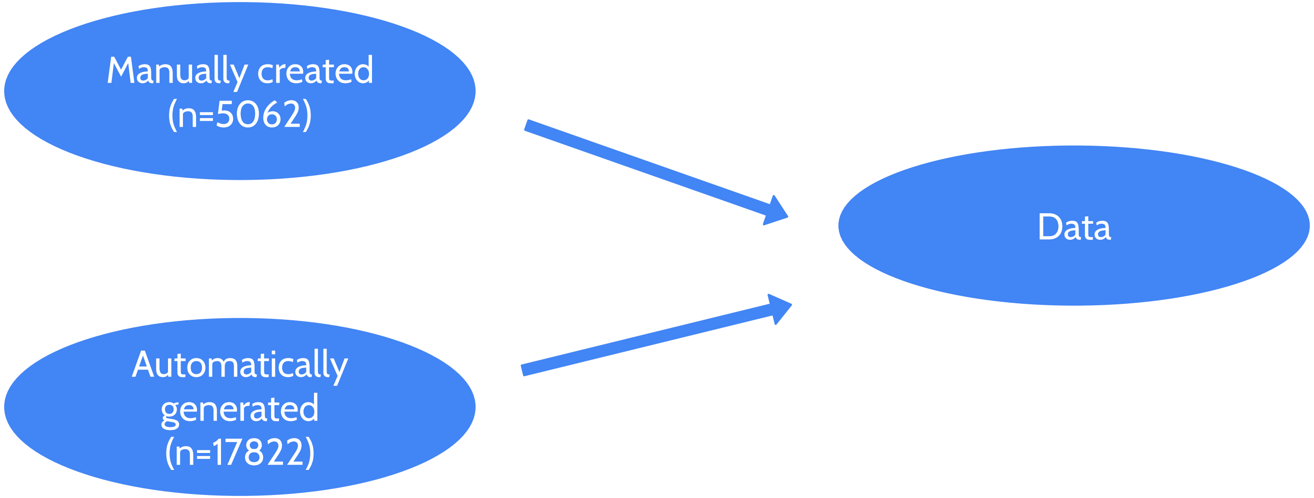
- Training set: 17782
- Validation set: 2722
- Test set: 2340



	training	validation	testing
positive	6591	1364	1047
negative	11191	1358	1293
total	17782	2722	2340

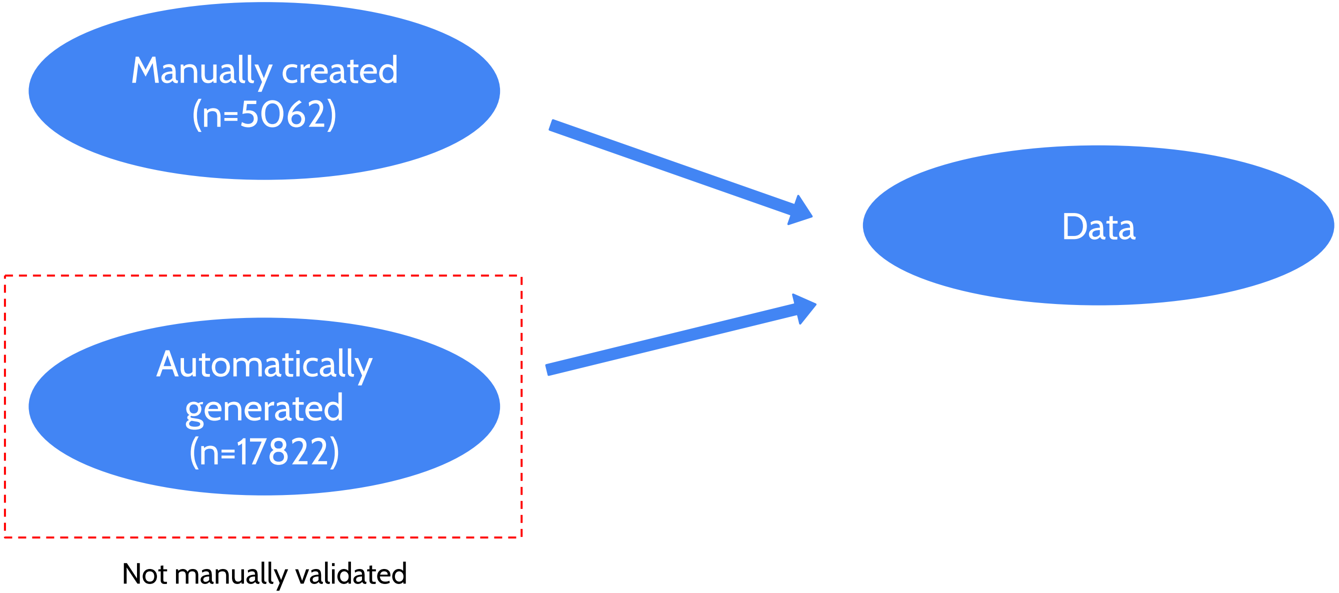
Table 4: Total size of the final dataset.

Data



Manually created (n=5062)
Automatically generated (n=17822)
Data

Data


Manually created
(n=5062)
Automatically
generated
(n=17822)
Not manually validated
Data

Manually Collected Data

McRae Triples

Manually Collected Data

McRae Triples

SimLex

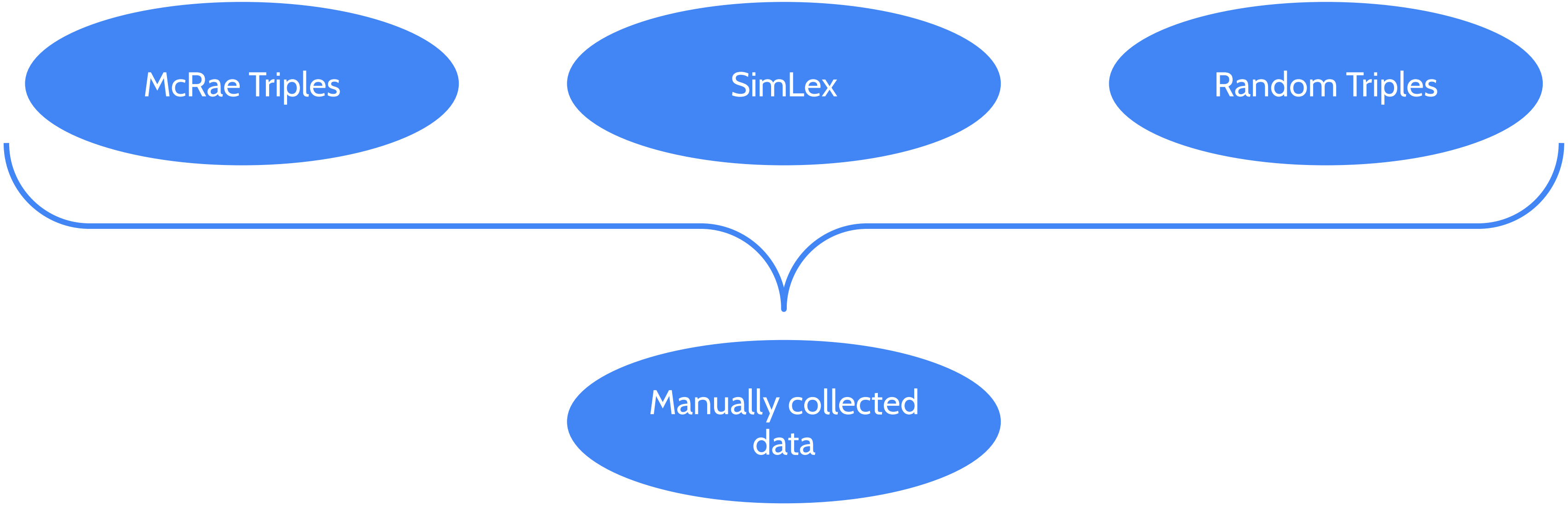
Manually Collected Data

McRae Triples

SimLex

Random Triples

Manually Collected Data



Evaluation

- Binary classification task evaluated on F1 score
- Baselines:
  - Random classifier: 0.517 (test set is slightly unbalanced)
  - Vector cosine baseline: 0.607
- Human upper bound:
  - Positive cases: 0.89
  - Negative cases: 0.91

SemEval 2018 Task 10 Top Performances

Rank	Team	Features	Classifier	F1 Score
1	SUNNYNLP	Pre-trained word embeddings + Probase information	SVM	0.75
2	Luminoso	ConceptNet word-embeddings + ConceptNet information	LinearSVC	0.74
3	BomJi	Pre-trained word embeddings + info. from graph-based distributional model (JoBimText)	XGBoost	0.73
4	NTU NLP	Pre-trained word embeddings + point wise mutual info. + ConceptNet edge information	MLP	0.73

SemEval 2018 Task 10 Observations

- Exploiting information from knowledge base resources like WordNet does improve the performance on average
- Traditional machine learning systems that entered competition were much more likely to make use of knowledge bases
- Many different systems and knowledge bases were used

Resource type	Average F1
WE + KB	0.678
WE	0.638

UPC

SemEval 2018 Task 10 Observations

- Exploiting information from knowledge base resources like WordNet does improve the performance on average
- Traditional machine learning systems that entered competition were much more likely to make use of knowledge bases
- Many different systems and knowledge bases were used

Resource type	Average F1
WE + KB	0.678
WE	0.638

combining neural approaches with knowledge bases may well lead to improved performances

Limitations

Data

Word 1	Word 2	Attribute	Label
Apple	Banana	Red	1
Gloves	Pants	Wool	0
Spider	Elephant	Wings	0

Tabular Data (isolated words)

Lack of contextual information (no sentences)

Proposed Approach

Feature Computation

- **Word Embeddings** (GloVe, Word2Vec)
- **Knowledge Bases**
  - WordNet
  - ConceptNet
- **Other features** (POS Tags, lemmas, distance, dissimilarity and similarity measures...)

System Type

1. **Support Vector Machine (SVM)**
2. **Convolutional Neural Network (CNN)**
   a. Different feature combinations
   b. Different architectures
   c. + Attention layers

Capturing Discriminative Attributes

Final Project - Proposal
HLE Fall 20222

Lucía Urcelay
Rasmus Siljander
UPC - Master in Artificial Intelligence

Observations

1. “Easy” triples were easy for both the human and model to discriminate.
2. “Hard” triples were harder for the models than it was for humans.
3. There were triplets that were failed by all systems.

Observations

1. “Easy” triples were easy for both the human and model to discriminate
2. “Hard” triples were harder for the models than it was for humans
3. There were triplets that were failed by all systems

→ There is room for improvement even for the best performing models