

Semi-supervised New Event Type Induction and Description via Contrastive Loss-Enforced Batch Attention

Carl Edwards and Heng Ji
Department of Computer Science, University of Illinois Urbana-Champaign

INTRODUCTION

- Many existing resources for event extraction may cover a limited number of types.
- This is especially true for domain-specific documents, such as within a scientific field which uses unique jargon and processes.
- In order to understand these, it is necessary to automatically discover these rare domain-specific events in these papers.
- We propose a novel semi-supervised event type induction approach using a semi-supervised contrastive loss-enforced batch attention mechanism.

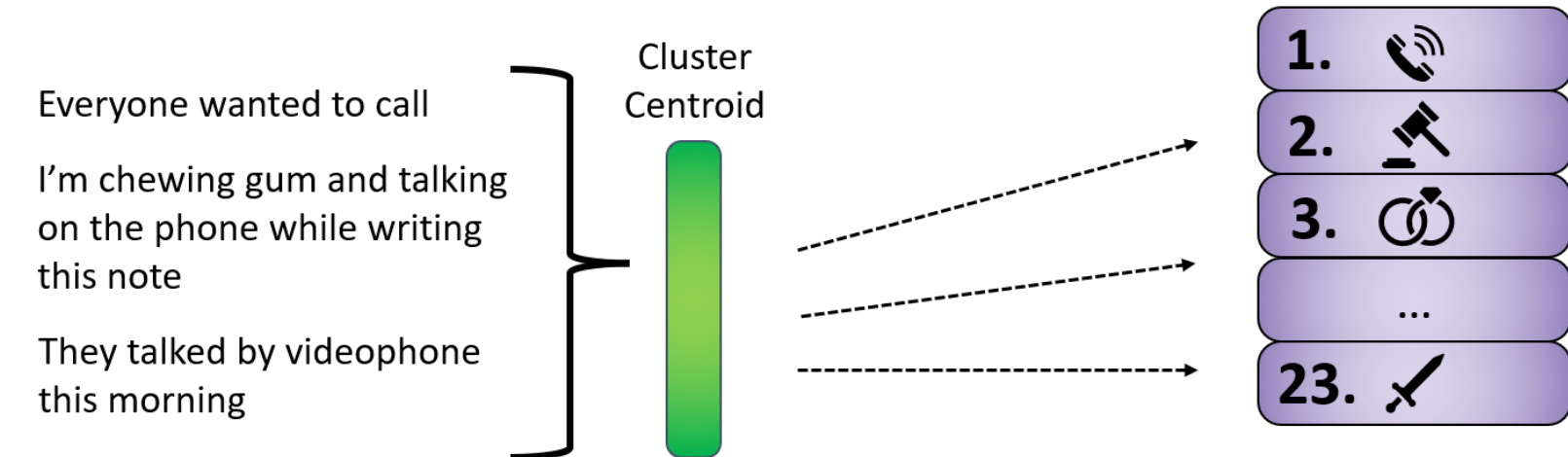
Tasks

Semi-supervised Event Type Induction

- Given the event mentions in a dataset with n types, the goal is to discover $n - m$ unknown types given m known types.
- This is essentially a special case of semi-supervised clustering.
- We evaluate on the ACE2005 dataset with 10 known types and 23 unknown types.

Name and FrameNet Linking

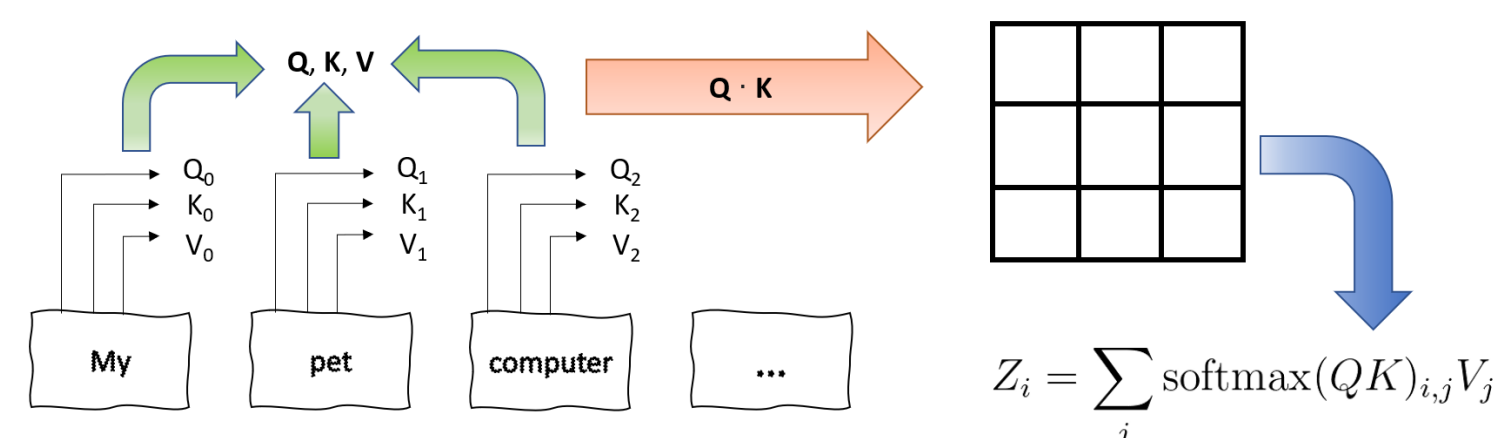
- We design two retrieval tasks for our discovered clusters.
- Cluster descriptors are retrieved from:
 - The 23 unknown cluster names (e.g. "Injure").
 - The 1,221 FrameNet frame definitions (e.g. "The words [...] describe situations in which an Agent or a Cause injures a Victim [...]").



Background

How does a Transformer usually work?

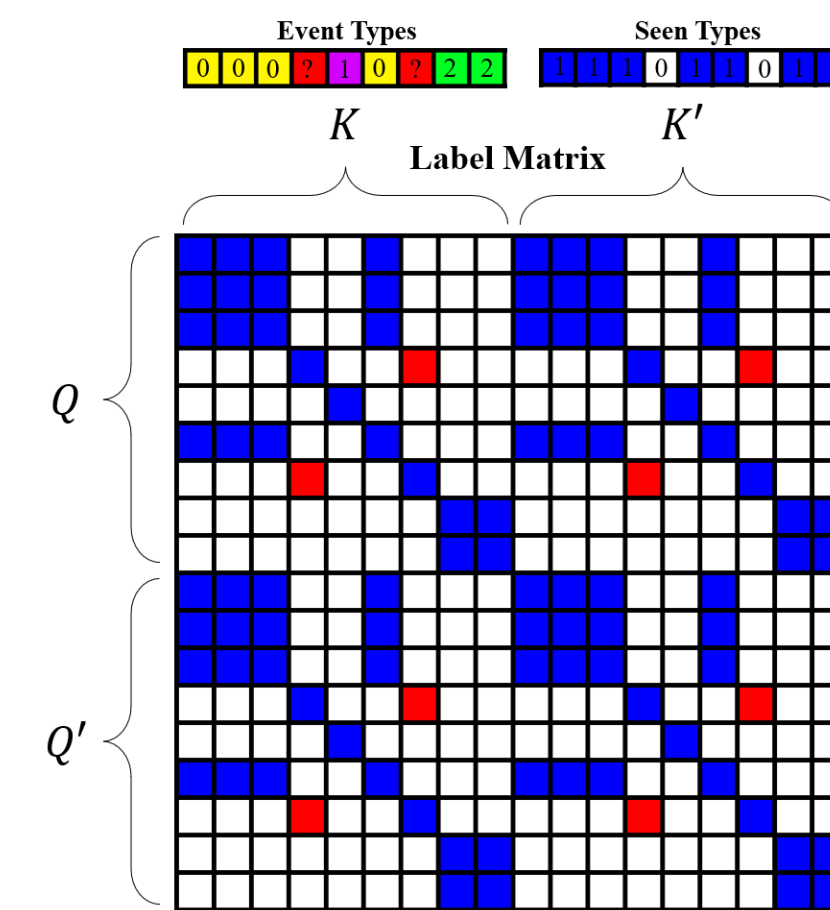
- A Transformer layer learns query, key, and value representations for a sequence of tokens.
- The dot product between the query and key representations is used to create a convex combination of the value representations.



METHODOLOGY

Semi-supervised, masked contrastive loss

- The following shows an example label matrix for the contrastive loss.
- Blue is positive, white is negative, and red is masked.
- Q' and K' are created using data augmentation.

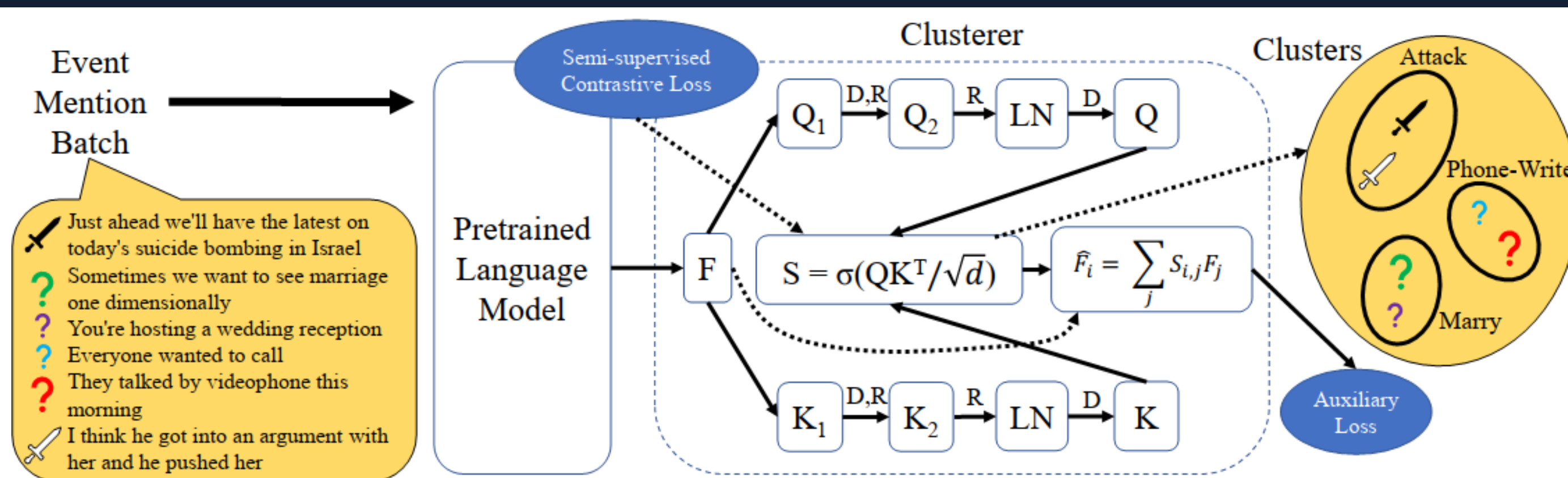


RESULTS

Method	Clusters	Geometric NMI	Fowlkes Mallows	Completeness	Homogeneity	V-Measure	ARI
One Cluster	1	0.00	25.58	100.00	0.00	0.00	0.00
SS-VQ-VAE w/o VAE (Huang and Ji, 2020)	500	33.45	25.54	42.76	26.17	32.47	-
SS-VQ-VAE (Huang and Ji, 2020)	500	40.88	31.46	53.57	31.19	39.43	-
SS-VQ-VAE + SBERT	33	19.08	19.45	25.80	15.13	19.08	7.54
SBERT	Agglo	23	50.71	34.35	57.05	45.07	50.36
SBERT	Manifold	23	48.75	36.02	51.32	46.30	48.68
Attn-Cosine	Agglo	23	46.40	34.60	49.82	43.24	46.27
Attn-DotProduct	Agglo	23	50.17	37.48	53.50	47.06	50.06
Attn	Manifold	23	54.83	42.77	55.00	54.82	38.74
FT-SBERT	Manifold	23	60.28	50.63	60.19	60.28	47.24
Attn-DotProduct	Affinity	49-68	56.87	35.64	49.58	56.33	30.02
Attn-Cosine	Affinity	50-69	56.54	33.00	48.72	55.91	27.04
E-Attn-DotProduct	Agglo	23	56.50	43.26	59.62	53.54	56.41
E-Attn	Agglo	23	59.00	46.19	58.36	59.00	42.56
E-FT-SBERT	Manifold	23	63.56	52.10	63.11	64.01	63.56
E-Attn-DotProduct	Affinity	63	60.00	38.41	51.32	70.15	31.78

FrameNet Linking Results

Clustering	Representation	Mean Rank	Hits@1	Hits@5	Hits@10	Hits@50	Hits@100	MRR	Average Purity	Type Representation
Base-23	Untuned	95.9	4.3%	21.7%	26.1%	30.4%	34.8%	0.128	25%	47.8%
FF-Base-23	Finetuned	156.9	30.4%	30.4%	34.8%	43.5%	47.8%	0.336	57.4%	65.2%
Ensemble-23	Untuned	72.7	17.4%	30.4%	39.1%	47.8%	65.2%	0.264	68.6%	69.6%
Ensemble-23	Finetuned	115.7	21.7%	34.8%	34.8%	43.5%	52.2%	0.308	68.6%	69.6%
Perfect-23	Untuned	15.9	26.1%	39.1%	52.2%	65.2%	73.9%	0.374	100%	100%
Perfect-23	Finetuned	42.7	47.8%	56.5%	60.9%	69.6%	69.6%	0.539	100%	100%



- Injure:** If those weren't gunshot wounds to cause the broken bones, do they know what caused the fractures
 - Injure:** More than 40 were injured
 - Injure:** There was no information on the identity of the injured person
 - Injure:** Sergeant Chuck Hagel was seriously wounded twice in Vietnam
- Charge-Indict:** 56-year-old forry drake has been charged with interstate transport of a minor
 - Charge-Indict:** Ocalan, being tried in absentia, was indicted for entering the country illegally, a
 - Convict and Charge-Indict:** convicted oklahoma city bombing conspirator terry nichols will stand trial again on state murder charges
 - Appeal:** in the african nation of nigeria, an islamic court delayed the appeal of a woman condemned to death by stoning
- Marry:** My wife and I were guests at a wedding on the Carnival Legend on New Years Eve 2003
 - Marry and Divorce:** Giuliani, 58, proposed to Nathan, a former nurse, during a November business trip to Paris - five months after he finalized his divorce from Donna Hanover after 20 years of marriage
 - Merge-org:** So Oracle and Peoplesoft , who spent the last 18 months insulting one another in every imaginable way, are finally tying the knot
- Declare-Bankruptcy:** You need to speak to a bankruptcy attorney pronto; this is a bankruptcy matter, not a tax matter
 - Declare-Bankruptcy:** despite operating under bankruptcy laws, united posted the best on time performance
 - Declare-Bankruptcy:** That means that he received the shares while he was still in bankruptcy, which means that the shares were potentially assets that the trustee could use to pay off creditors
- Start-Org:** Kiichiro Toyoda founded the automaker in 1937, transforming the loom manufacturer started by his father into an automaker
 - Merge-Org:** I believe any neutral management consultant worth his or her salt would recommend a merger of the two organizations
 - End-Org:** It's a dying organization, and this will be just the jolt it needs for another couple decades of somnambulant staggering before being ultimately replaced by far more efficient companies
- Marry:** Either its bad or good
 - End-Org:** i felt t7ire was something else too, much history behind silver cross to end is now
 - Trial-Hearing:** Yeah, we're a pretty small town, so our newspaper covers it a lot
 - Trial-Hearing:** Yeah, because I was really -- I wasn't really following it that much because I was
 - Start-Position:** then when they're ready to breed they go to the wb

CONCLUSIONS

- We propose a novel framework for new event type induction which uses a masked contrastive loss to enforce an attention mechanism over data minibatches. This framework is also potentially applicable for semi-supervised clustering and classification problems in other settings where a pretrained model exists.
- We use the "clustered" features produced by our model to extend new event type induction to two novel downstream tasks: type name prediction and FrameNet linking.

ACKNOWLEDGEMENTS

This research is based upon work supported in part by U.S. DARPA KAIROS Program No. FA8750-19-2-1004. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of DARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.