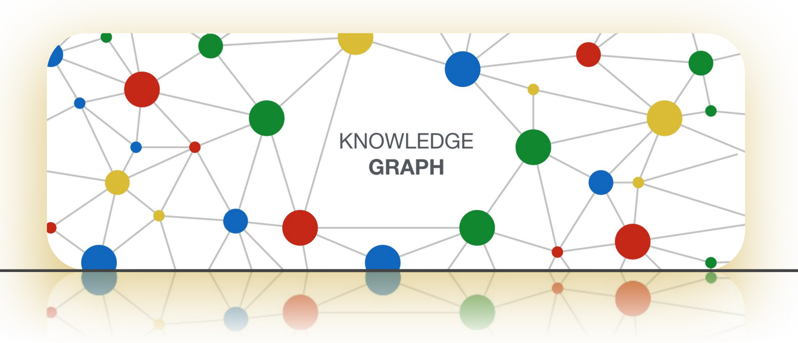
Harvesting the Web for Building Large-scale Argumentation Graphs

Anh Phuong Le



Motivation

?

What are the effects of legalizing medical marijuana?

increase
cause
lead to

- addiction
- memory loss
- depression

suppress
treat
prevent

- pain
- schizophrenia
- cancer

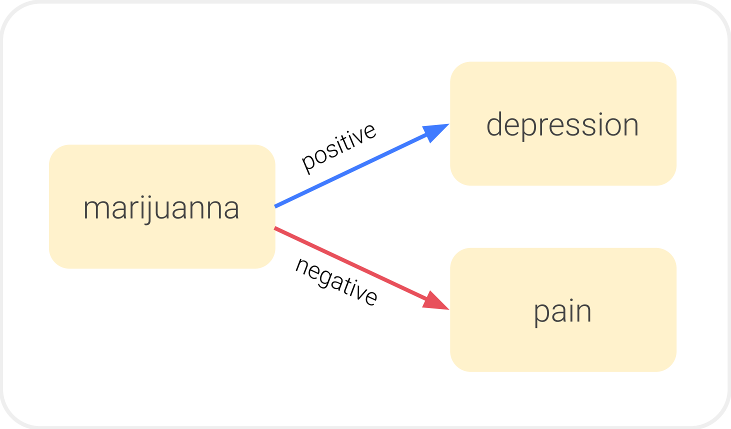
Outline

1. Background

Previous Work *(Al-Khatib et al. [2020])*

Marijuana causes higher rate of depression.

Marijuana can significantly relieve pain.



Claims from Debate Portals → Argumentation Graph

Limitation 1: Scope of Input Data

Claim: *(Al-Khatib et al. [2020])*

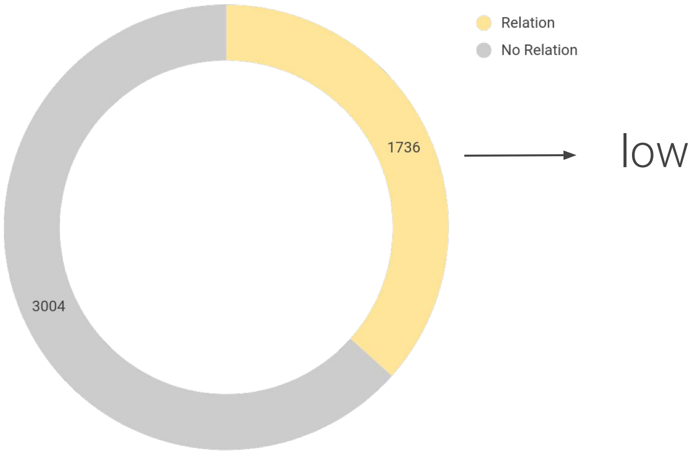
Marijuana has ability to treat cancer.

Full Arguments:

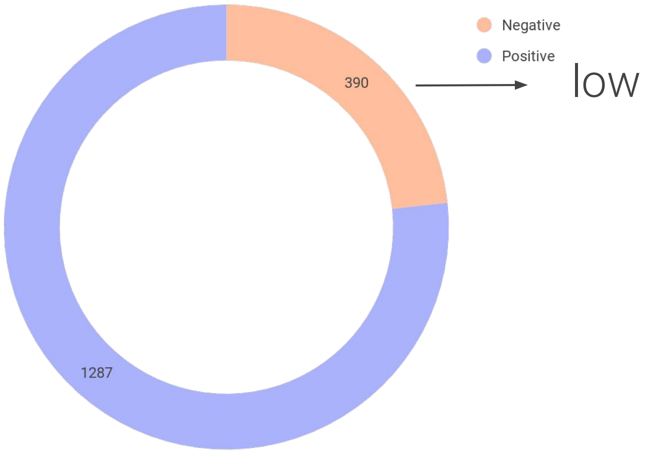
… Moreover, as I've stated before, marijuana doesn't just help with breast cancer; rather, THC (a primary chemical found in marijuana) also helps destroy brain cancer cells, and research has provided immensely compelling evidence of marijuana's ability to reduce up to 50% of tumor growth in common lung cancer, as well as prevent the spread of the cancer significantly…

→ unused

Limitation 2: Dataset Balance



Effect Relation



Relation Type

Limitation 3: Classifier Effectiveness

1. Tasks:
    - Effect Detection
    - Type Classification
2. Approach: feature engineering
3. Training data: imbalance
4. Example of failed prediction (negative relation)
    - Subsidization would damage independence of journalism.
    - Two-state solution would prevent return of Palestinian refugees.

Overview of Contribution

New Dataset

Build a dataset of annotated effect relation

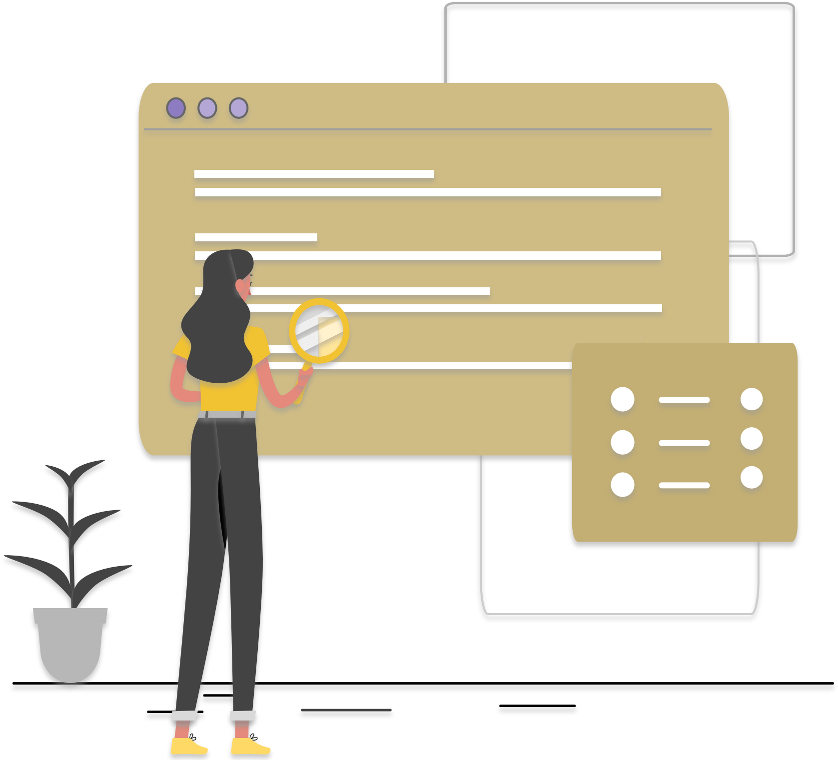
- more coverage
- more balance

New Classifier

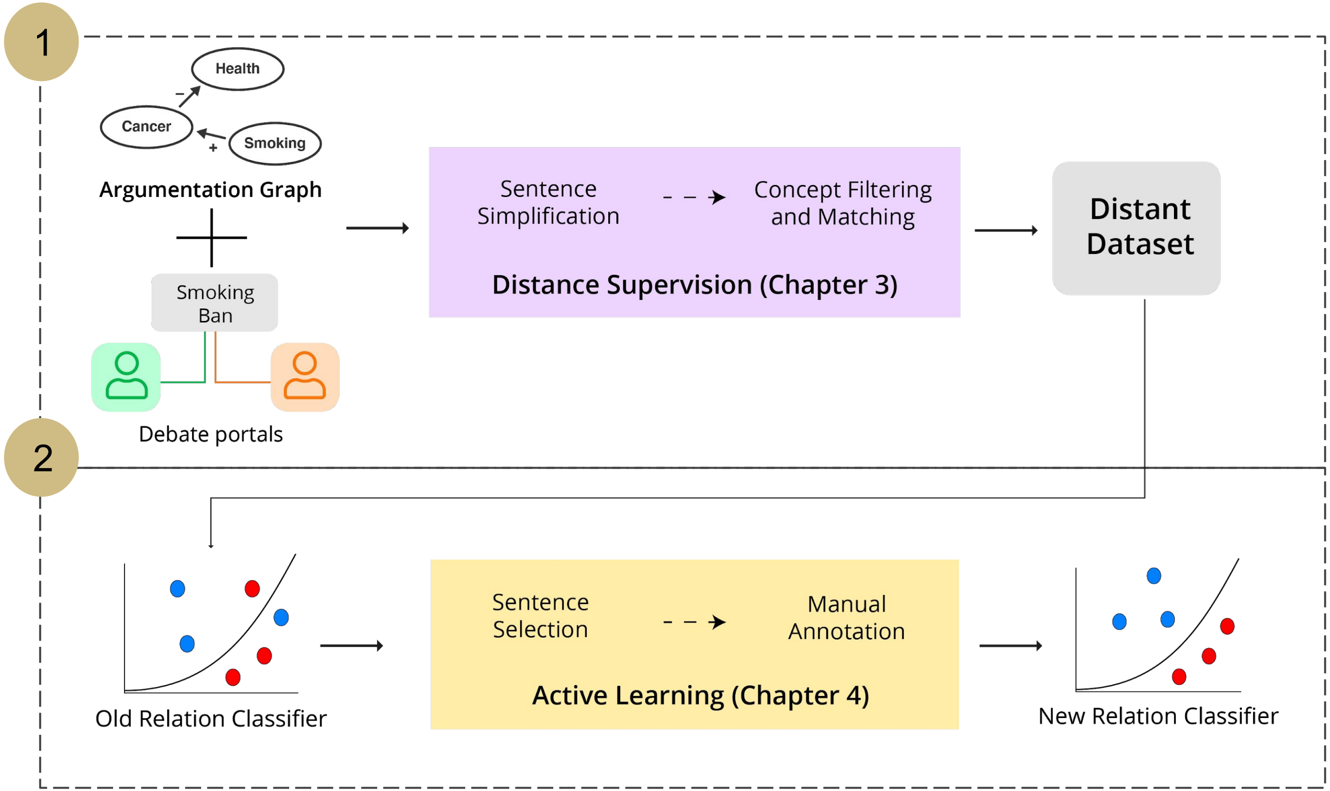
Train classifier using state-of-the-art models

- deal with new scope
- better effectiveness

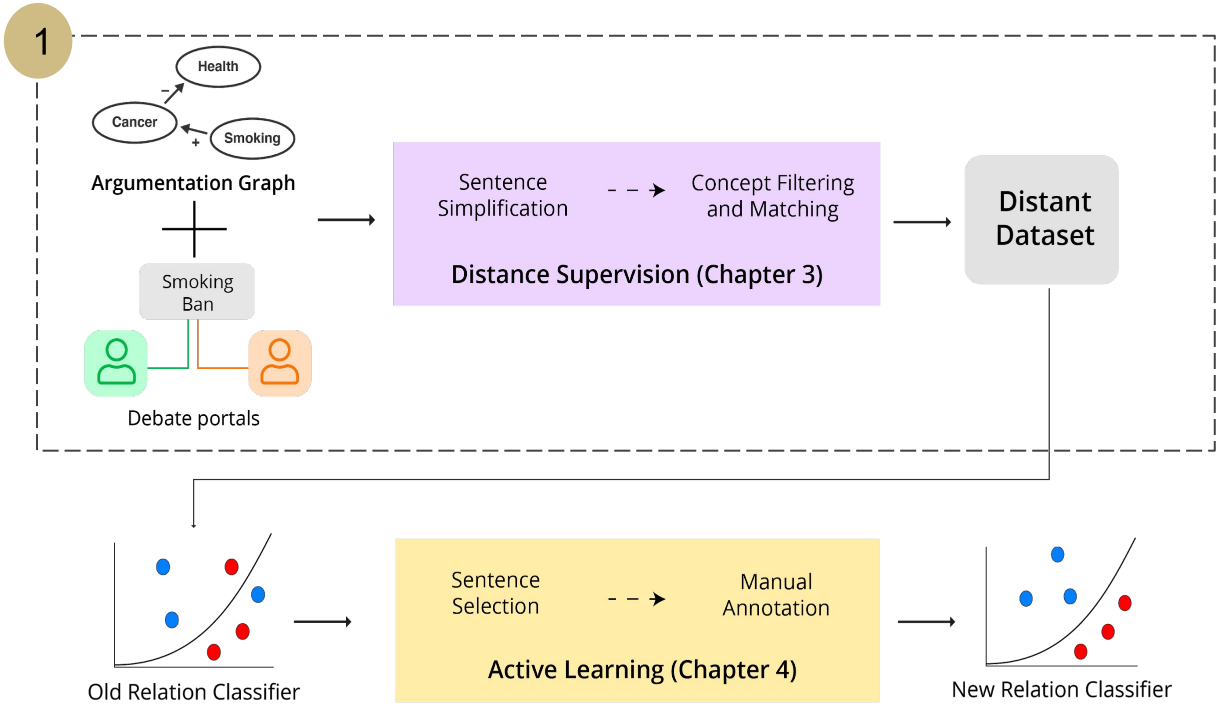
2.
Effect Relation Extraction



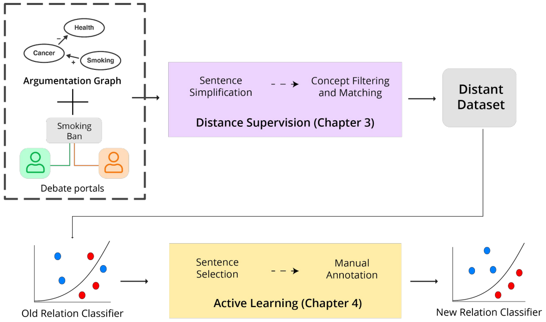
Our Approach

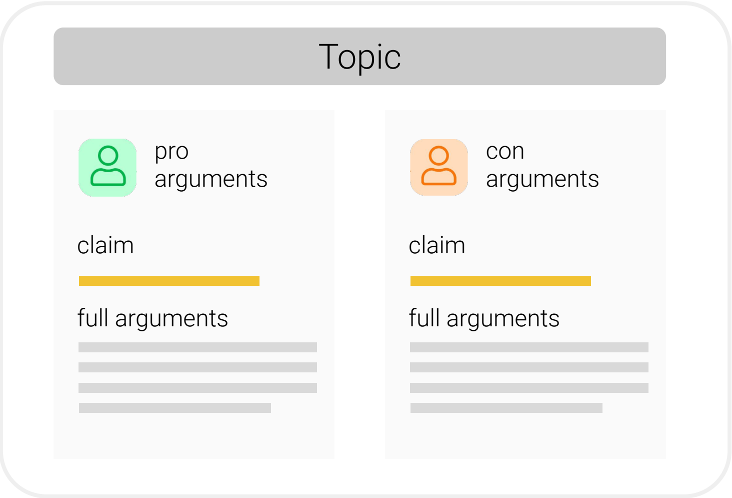


Distant Supervision

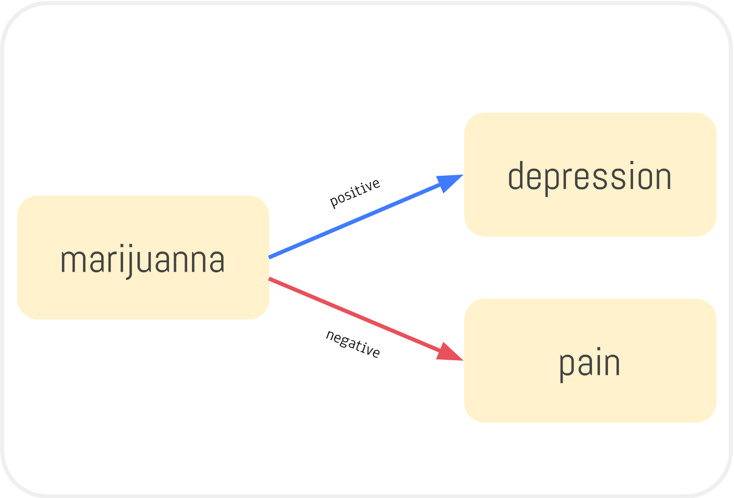


Input: Argumentation Graph & Debate Portals Arguments



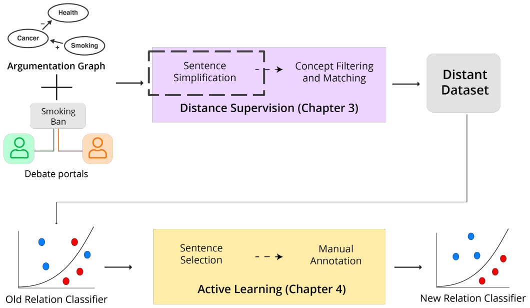


args.me dataset
(Ajjour et al. [2019])



Argumentation Graph
(Al-Khatib et al. [2020])

Sentence Simplification



… Moreover, as I've stated before, marijuana doesn't just help with breast cancer; rather, THC (a primary chemical found in marijuana) also helps destroy brain cancer cells, and research has provided immensely compelling evidence of marijuana's ability to reduce up to 50% of tumor growth in common lung cancer, as well as prevent the spread of the cancer significantly...

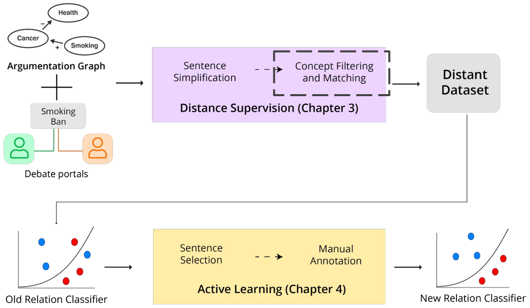
Graphene
(Cetto et al. [2018])

- Marijuana doesn't just help with breast cancer.
- THC (a primary chemical found in marijuana) also helps destroy brain cancer cells.
- Research has provided immensely compelling evidence of marijuana's ability to reduce up to 50% of tumor growth in common lung cancer, as well as prevent the spread of the cancer significantly.

Arguments from args.me dataset

Simple sentences

Concepts Expansion



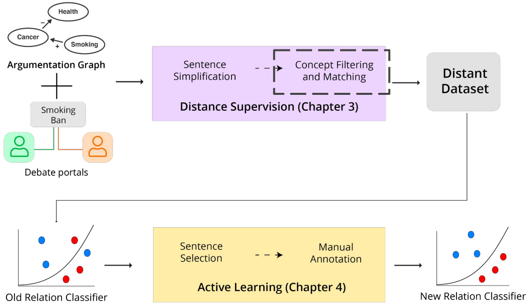
Concepts in Argumentation Graph
(Al-Khatib et al. [2020])

marijuanna = cannabis

headscarf = veil-wearing = face-covering veil = burqa

Individual Concepts → Group of concepts

Concept Matching

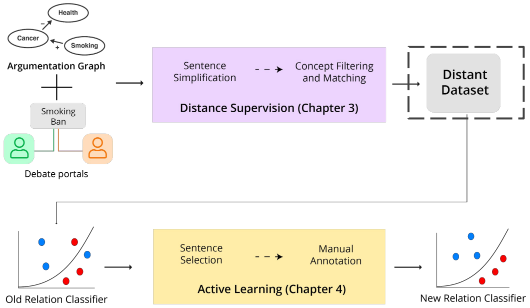


marijuanna → cancer *(Alkhatib et al. [2020])*

- Marijuana doesn't just help with breast cancer.
- THC (a primary chemical found in marijuana) also helps destroy brain cancer cells.
- Research has provided immensely compelling evidence of marijuana's ability to reduce up to 50% of tumor growth in common lung cancer, as well as prevent the spread of the cancer significantly.

Simple sentences → Matching sentences

Distant Dataset



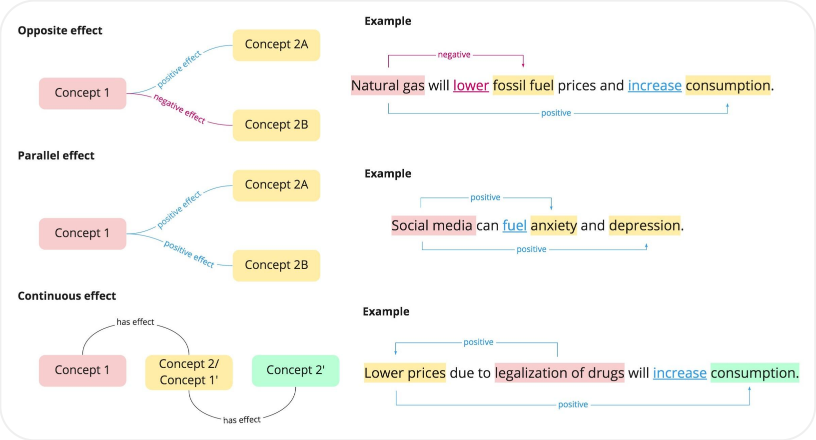
1. Filter out noisy sentences from matched

   ⟶ **10,000** sentences

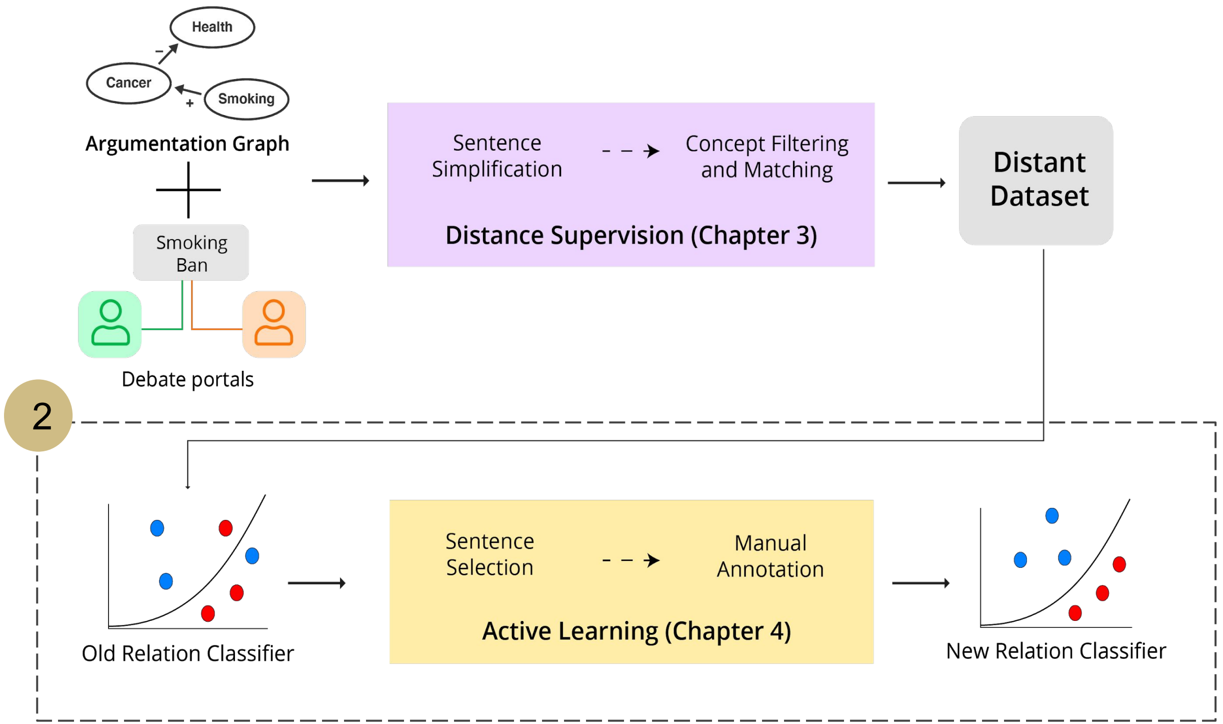
2. Manually inspect 100 sentences,

   ⟶ 70% effect relation

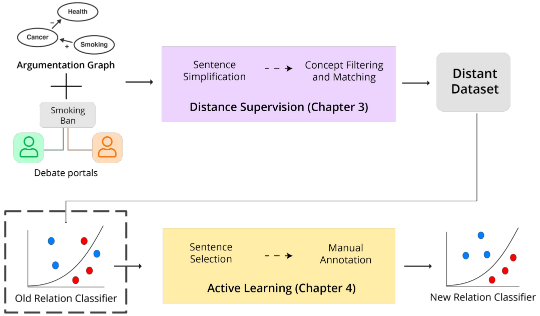
3. Found complex effect relations ⟶



Active Learning

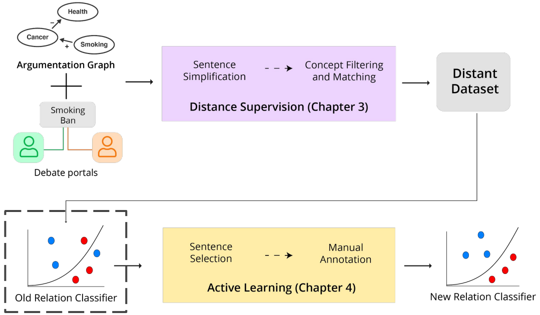


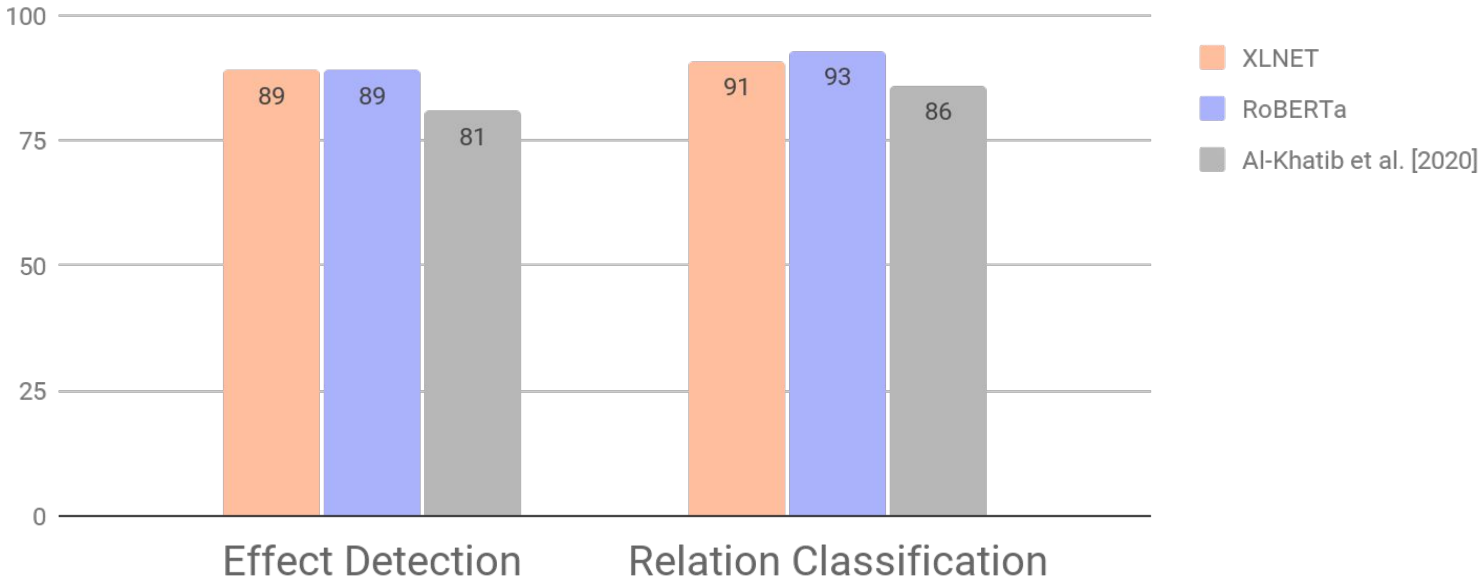
Old Relation Classifier: Training using Deep Learning



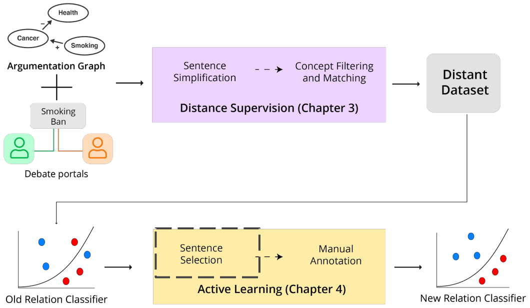
- Tasks
  - Detecting 'Effect Relation' in sentences
  - Classifying whether the detected effect is positive or negative
- Training datasets
  - Old annotated dataset *(Alkhatib et al. [2020])*
- Approach
  - Different neural-based models (Hugging Face library - *Wolf et al. [2019]*)
  - Features: sentence embedding

Old Relation Classifier: Results (F1 score)



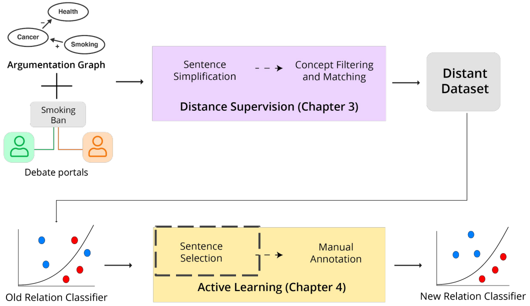


Sentence Selection

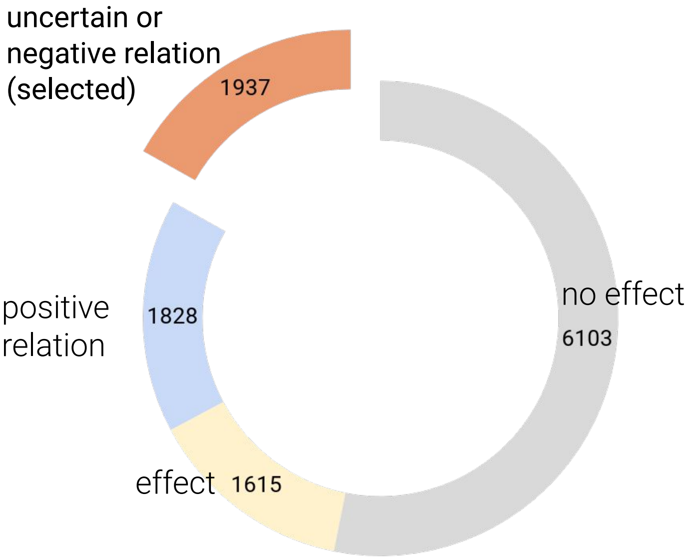


- Objective
  - Select most informative sentences
  - Get more negative relations
- Approach
  - Apply old classifiers to distant dataset
  - Distinguish based on
    - Uncertainty Sampling
    - Most Disagreement

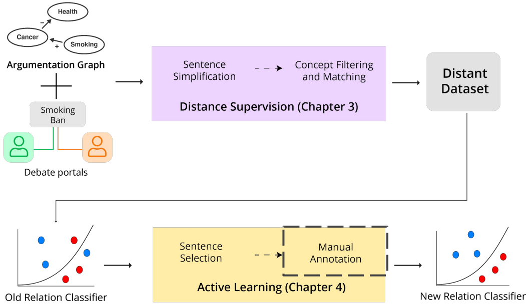
Sentence Selection



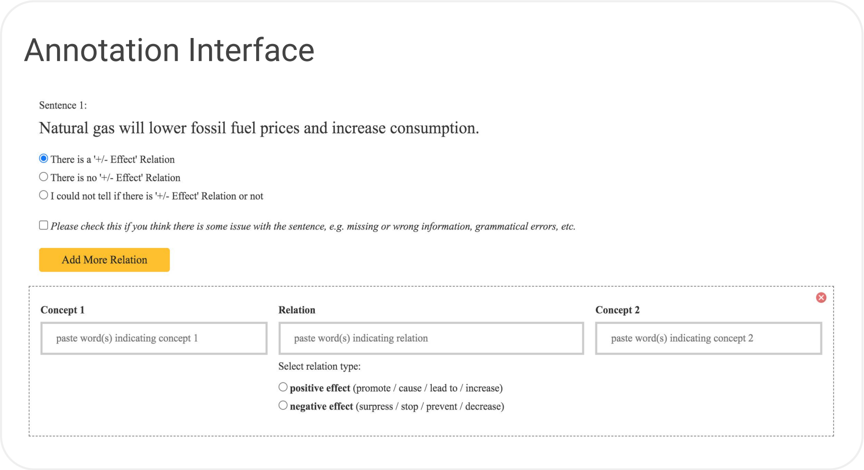
- Filter out sentences with high confidence of
  - Effect: 6,103
  - No Effect: 1,615
  - Positive Relation: 1,828
- Select the rest: 1,937



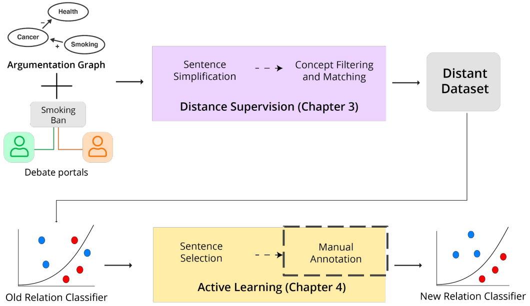
Crowd-sourcing: Task

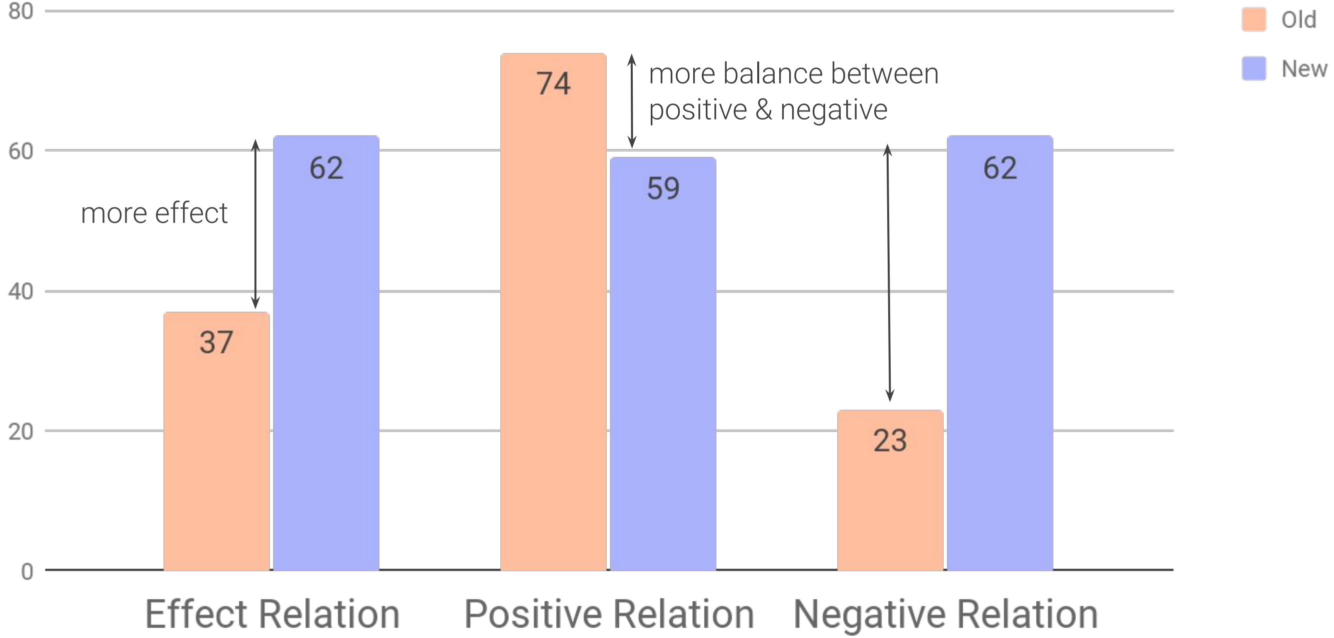


- Input
  - Selected sentences from Distant Dataset
- Task
  - 3 people label concepts, relations
- Output
  - Annotation of the sentences
- Aggregation of Annotation
  - Majority Vote

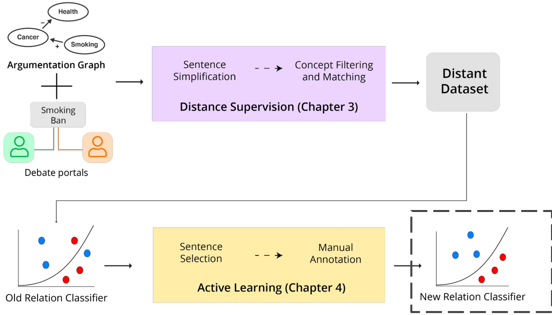


New dataset


Argumentation Graph
Health
Cancer
Smoking
Smoking Ban
Debate portals
Sentence Simplification
Concept Filtering and Matching
Distance Supervision (Chapter 3)
Distant Dataset
Old Relation Classifier
Sentence Selection
Manual Annotation
Active Learning (Chapter 4)
New Relation Classifier


80
60
40
20
0
74
more balance between positive & negative
62
59
62
more effect
37
23
Effect Relation
Positive Relation
Negative Relation
Old
New

Classifiers with New dataset & Combined dataset

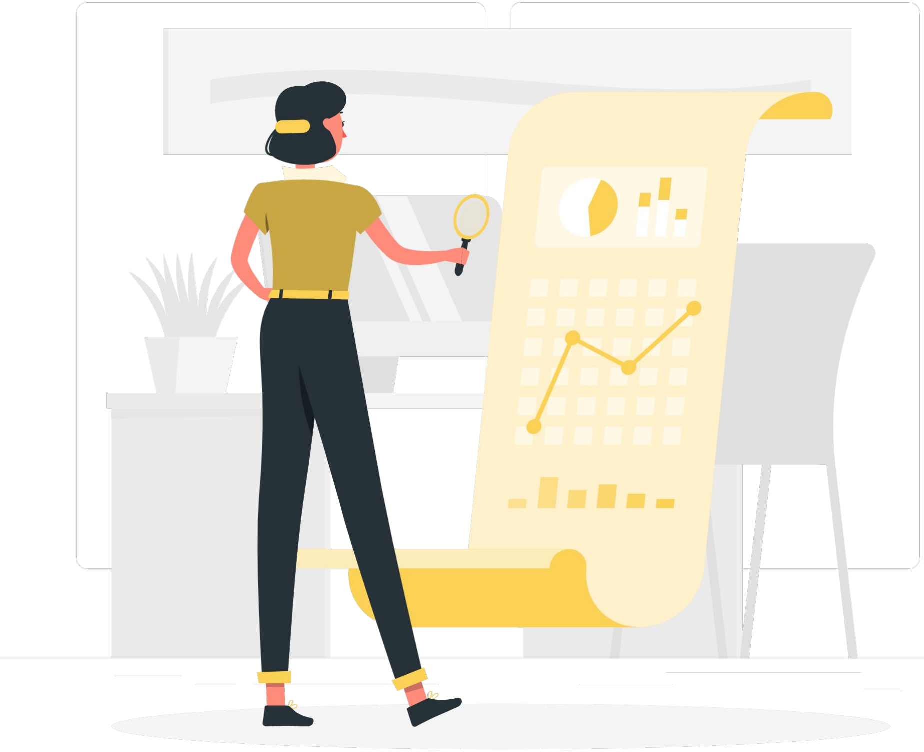


- Tasks
  - Detecting 'Effect Relation' in sentences
  - Detecting positive relation
  - Detecting negative relation

due to multiple relation

- Classifier type 1:
  - Trained on new annotated dataset
- Classifier type 2:
  - Trained on old *(Alkhatib et al. [2020]) combined* with new dataset

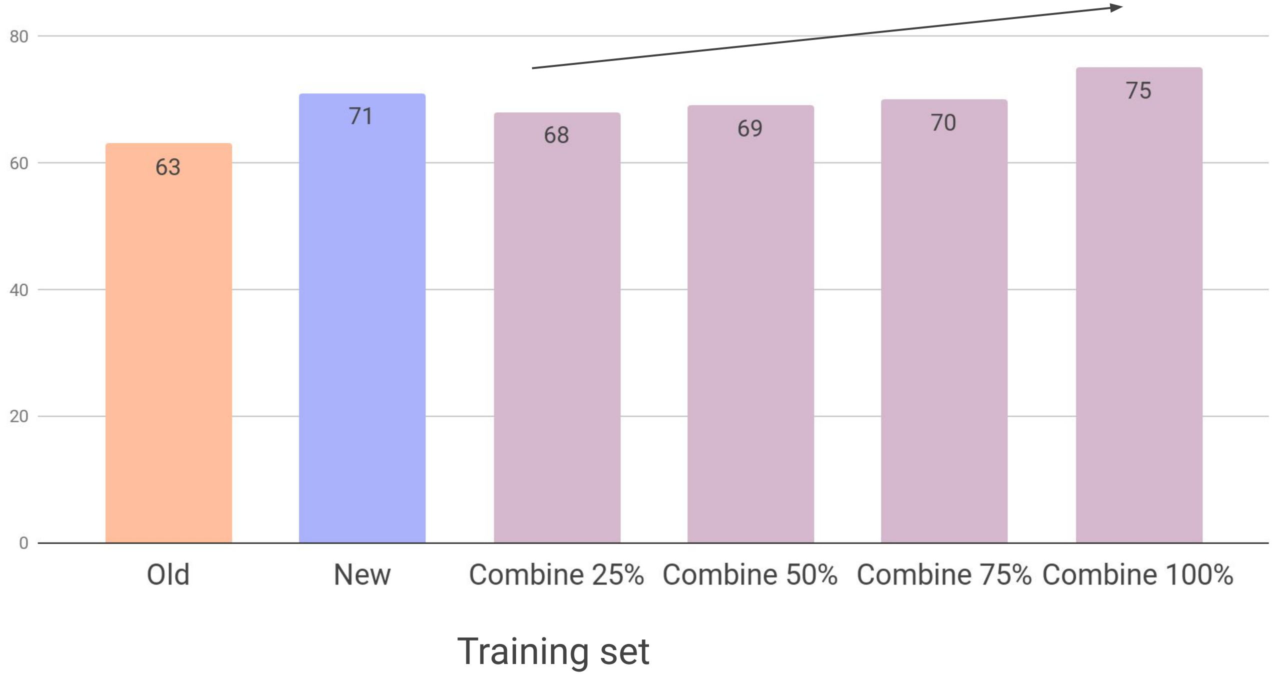
3.
Evaluation



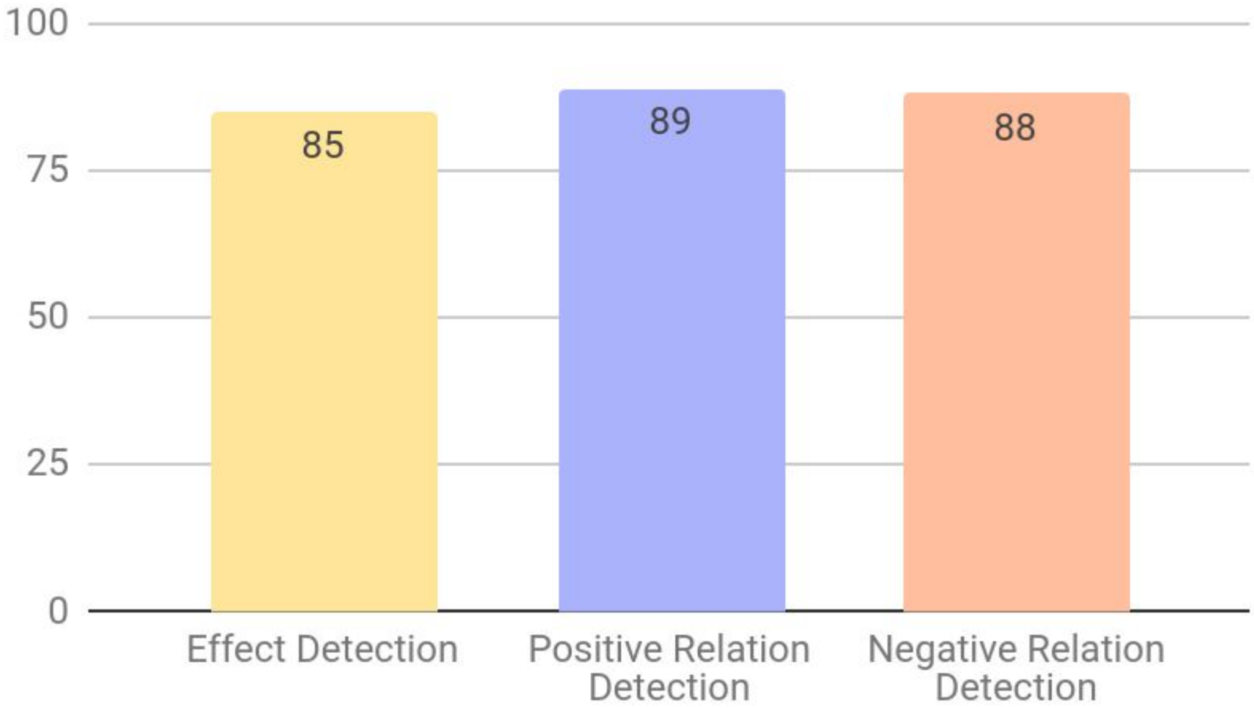
Experiment Setting

- Training and testing
  - Old annotated dataset *(Alkhatib et al. [2020])*
  - New annotated dataset
  - Combine
  - Split: 80% training, 20% testing

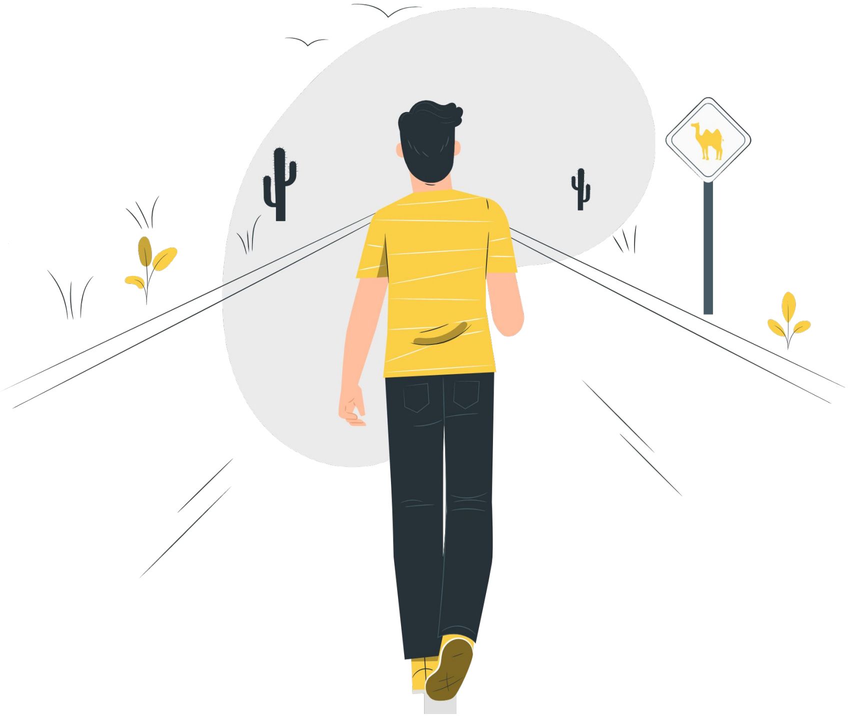
Effect Detection: Testing on New Dataset



Training and testing on Combined Dataset



4. Conclusion & Future Work



Contribution

New Dataset

Build a dataset of annotated effect relation

- more coverage
  - full arguments
- more relation
  - 63%
- more balance of relation types

New Classifier

Train classifier using state-of-the-art models

- more reliable
  - deal with complex sentences
- effectiveness
  - 85% for effect detection
  - 89% for positive / negative relation detection

Future Work

- Applying new effect relation classifiers on big dataset to build large-scale argumentation graph
- Multi-task learning classifier (relation + concept)
- Using effect relations for question-answering system

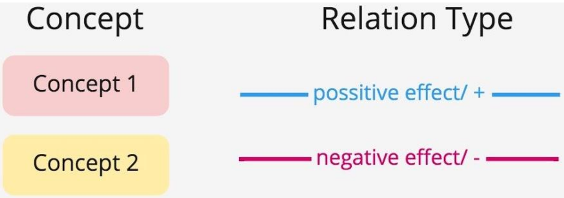
Question & Answer

Thank you!

source illustration image: https://stories.freepik.com/

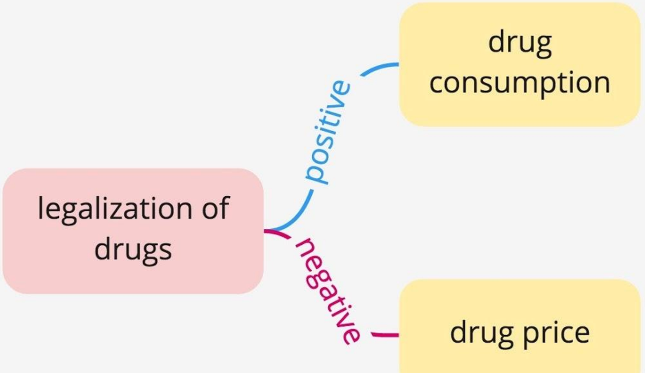
Figures

Sentence with positive / negative effect relation
=> Relation Triple: (*Concept 1, Relation Type, Concept 2*)

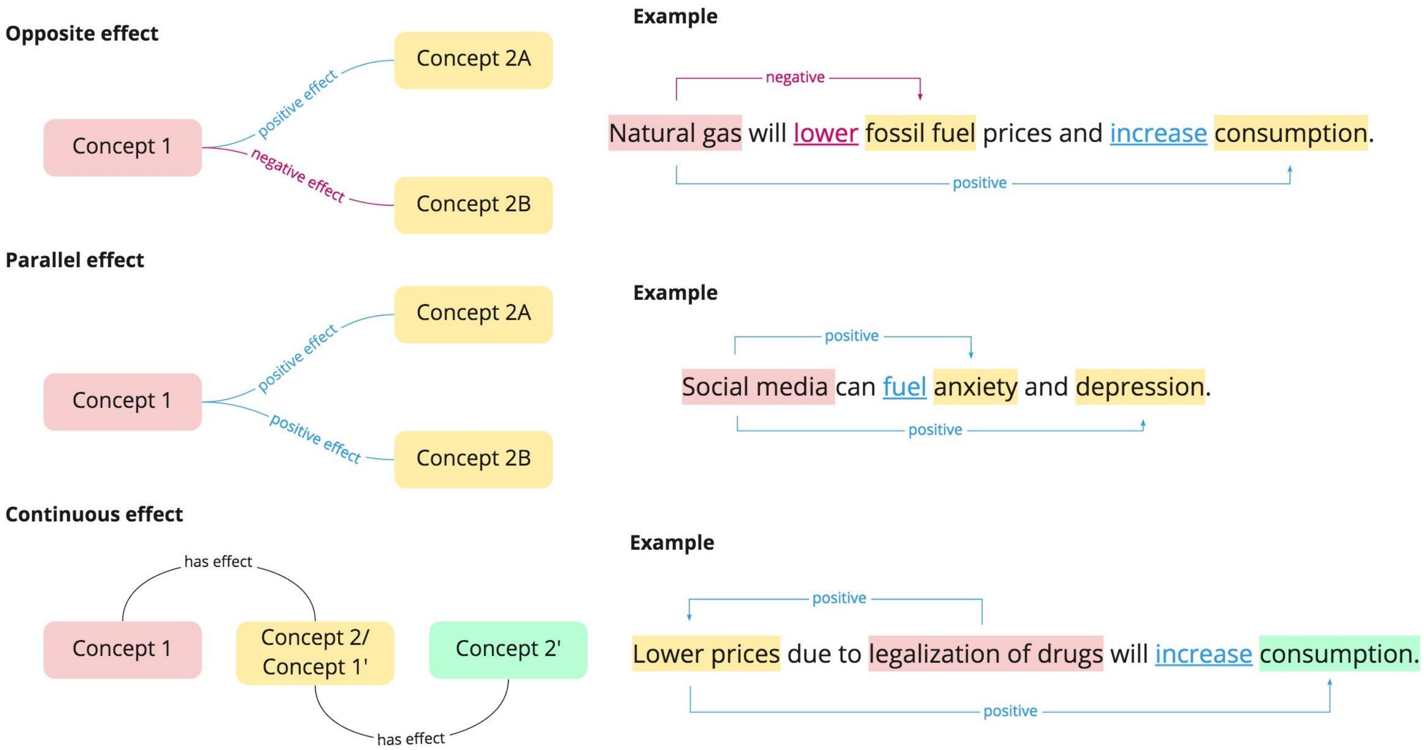


Legalization of drugs increases drug consumption.
=> (*legalization of drugs, positive effect, drug consumption*)

Legalization of drugs lowers drug price.
=> (*legalization of drugs, negative effect, drug price*)



Figures



Opposite effect
Concept 1
positive effect
negative effect
Concept 2A
Concept 2B
Example
negative
Natural gas will lower fossil fuel prices and increase consumption.
positive
Parallel effect
Concept 1
positive effect
positive effect
Concept 2A
Concept 2B
Example
positive
Social media can fuel anxiety and depression.
positive
Continuous effect
has effect
Concept 1
Concept 2/ Concept 1'
Concept 2'
has effect
Example
positive
Lower prices due to legalization of drugs will increase consumption.
positive

Predictions

```
s = "Subsidization would damage independence of journalism"
```
executed in 4ms, finished 20:48:41 2020-10-14

```
forward_calculate_probs(s, model)
```
executed in 43ms, finished 20:48:41 2020-10-14

```
prob:  [0.0462454  0.95375454]
1
```

```
predictor.explain(s)
```
executed in 2m 9s, finished 20:50:52 2020-10-14

y=1 (probability **0.984**, score **4.131**) top features

Contribution?	Feature
+5.825	Highlighted in text (sum)
-1.694	<BIAS>

subsidization would damage independence of journalism

```
s = "Two-state solution would prevent return of Palestinian refugees."
```
executed in 4ms, finished 20:53:47 2020-10-14

```
forward_calculate_probs(s, model)
```
executed in 61ms, finished 20:53:50 2020-10-14

```
prob:  [0.02058323 0.9794168 ]
1
```

```
predictor.explain(s)
```
executed in 1m 56.1s, finished 20:55:49 2020-10-14

y=1 (probability **0.974**, score **3.609**) top features

Contribution?	Feature
+4.137	Highlighted in text (sum)
-0.528	<BIAS>

two-state solution would prevent return of palestinian refugees.

Figures

due to cinemas and movie theaters closing, the global box office has dropped by billions of dollars, and streaming has become more popular, while the stock of film exhibitors has also dropped dramatically.

y=1 (probability **0.908**, score **2.287**) top features

3

Contribution?	Feature
+3.586	Highlighted in text (sum)
-1.299	<BIAS>

beyond remittances, however, migrants and diaspora contribute to countries of origin and destination economically in many more ways - through labour force participation, entrepreneurship and self-employment, small-scale investments including real estate/portfolio markets, nostalgia/ cross border trade, and the transfer of social and technological capital.

y=1 (probability **0.883**, score **2.017**) top features

5

Contribution?	Feature
+3.267	Highlighted in text (sum)
-1.249	<BIAS>

most governments around the world have temporarily closed educational institutions in an attempt to reduce the spread of covid-19.

y=1 (probability **1.000**, score **9.212**) top features

7

Contribution?	Feature
+10.508	Highlighted in text (sum)
-1.296	<BIAS>

no time to die was the first film to change its planned release outside of china because of the coronavirus outbreak, and has opened discussions of dramatic implications on the film economy: many other productions had avoided scheduling releases at the same time as the 25th bond film, and its new november date is in the busy holiday release period, leading to low box office intake in march/april and uncertain intake in november.

y=1 (probability **0.996**, score **5.539**) top features

4

Contribution?	Feature
+6.830	Highlighted in text (sum)
-1.291	<BIAS>

cineworld, which is the second biggest cinema chain in the world, warned on march 12, when multiple films pushed back their releases, that extended disruption and continuing falling stock could cause the company to collapse.

y=1 (probability **0.999**, score **7.178**) top features

6

Contribution?	Feature
+8.495	Highlighted in text (sum)
-1.317	<BIAS>

the actions were criticized for creating a potential superspreader event as the social nature of the festival could increase the risk for covid-19 transmission.

y=1 (probability **0.999**, score **6.974**) top features

8

Contribution?	Feature
+8.988	Highlighted in text (sum)
-2.013	<BIAS>

Figures

9

y=0 (probability **0.947**, score **-2.889**) top features

Contribution?	Feature
+2.128	Highlighted in text (sum)
+0.762	<BIAS>

55% of lower-income adults say the outbreak is a major threat to their finances, compared with 32% of middle-income adults and 24% of upper-income adults.

11

y=0 (probability **1.000**, score **-11.962**) top features

Contribution?	Feature
+11.355	Highlighted in text (sum)
+0.607	<BIAS>

the analysis also showed that more downstream cases were linked to spread in social settings such as weddings and restaurants than to household spread.

10

y=0 (probability **0.705**, score **-0.872**) top features

Contribution?	Feature
+0.841	<BIAS>
+0.030	Highlighted in text (sum)

in a study that looked specifically at school closure in the united states, closure of schools was associated with 1.37 million fewer cases and 40,600 fewer deaths from covid-19 in a six-week study period.

9

y=1 (probability **0.994**, score **5.189**) top features

Contribution?	Feature
+6.566	Highlighted in text (sum)
-1.377	<BIAS>

55% of lower-income adults say the outbreak is a major threat to their finances, compared with 32% of middle-income adults and 24% of upper-income adults.

11

y=1 (probability **0.711**, score **0.903**) top features

Contribution?	Feature
+1.840	Highlighted in text (sum)
-0.937	<BIAS>

the analysis also showed that more downstream cases were linked to spread in social settings such as weddings and restaurants than to household spread.

10

y=1 (probability **0.997**, score **5.659**) top features

Contribution?	Feature
+6.941	Highlighted in text (sum)
-1.282	<BIAS>

in a study that looked specifically at school closure in the united states, closure of schools was associated with 1.37 million fewer cases and 40,600 fewer deaths from covid-19 in a six-week study period.

Figures

1

y=1 (probability **0.981**, score **3.921**) top features

Contribution?	Feature
+4.752	Highlighted in text (sum)
-0.831	<BIAS>

people are asked to self-isolate or isolate to help prevent the spread of covid-19 to others.

2

y=1 (probability **0.998**, score **6.002**) top features

Contribution?	Feature
+6.836	Highlighted in text (sum)
-0.833	<BIAS>

most governments around the world have temporarily closed educational institutions in an attempt to reduce the spread of covid-19.

3

y=0 (probability **0.998**, score **-6.347**) top features

Contribution?	Feature
+5.756	Highlighted in text (sum)
+0.591	<BIAS>

55% of lower-income adults say the outbreak is a major threat to their finances, compared with 32% of middle-income adults and 24% of upper-income adults.

4

y=0 (probability **0.622**, score **-0.498**) top features

Contribution?	Feature
+0.803	<BIAS>
-0.305	Highlighted in text (sum)

rapid assessment can identify areas to increase efficiencies, prioritize initiatives and reduce spending

Figures

y=1 (probability **0.689**, score **0.797**) top features

3

Contribution?	Feature
+0.541	<BIAS>
+0.257	Highlighted in text (sum)

55% of lower-income adults say the outbreak is a major threat to their finances, compared with 32% of middle-income adults and 24% of upper-income adults.

y=1 (probability **0.943**, score **2.797**) top features

3

Contribution?	Feature
+3.264	Highlighted in text (sum)
-0.467	<BIAS>

[mask] say [mask] is a major threat to [mask], compared with [mask] and [mask].

y=0 (probability **0.980**, score **-3.868**) top features

4

Contribution?	Feature
+3.341	Highlighted in text (sum)
+0.527	<BIAS>

rapid assessment can identify areas to increase efficiencies, prioritize initiatives and reduce spending

y=1 (probability **0.905**, score **2.251**) top features

4

Contribution?	Feature
+2.408	Highlighted in text (sum)
-0.157	<BIAS>

[mask] can identify [mask] to increase [mask], prioritize [mask] and reduce [mask]

y=1 (probability **0.786**, score **1.299**) top features

3

Contribution?	Feature
+2.603	Highlighted in text (sum)
-1.304	<BIAS>

[mask] say [mask] is a major threat to [mask], compared with [mask] and [mask].

y=0 (probability **0.311**, score **0.797**) top features

4

Contribution?	Feature
+0.655	<BIAS>
-1.452	Highlighted in text (sum)

rapid assessment can identify areas to increase efficiencies, prioritize initiatives and reduce spending

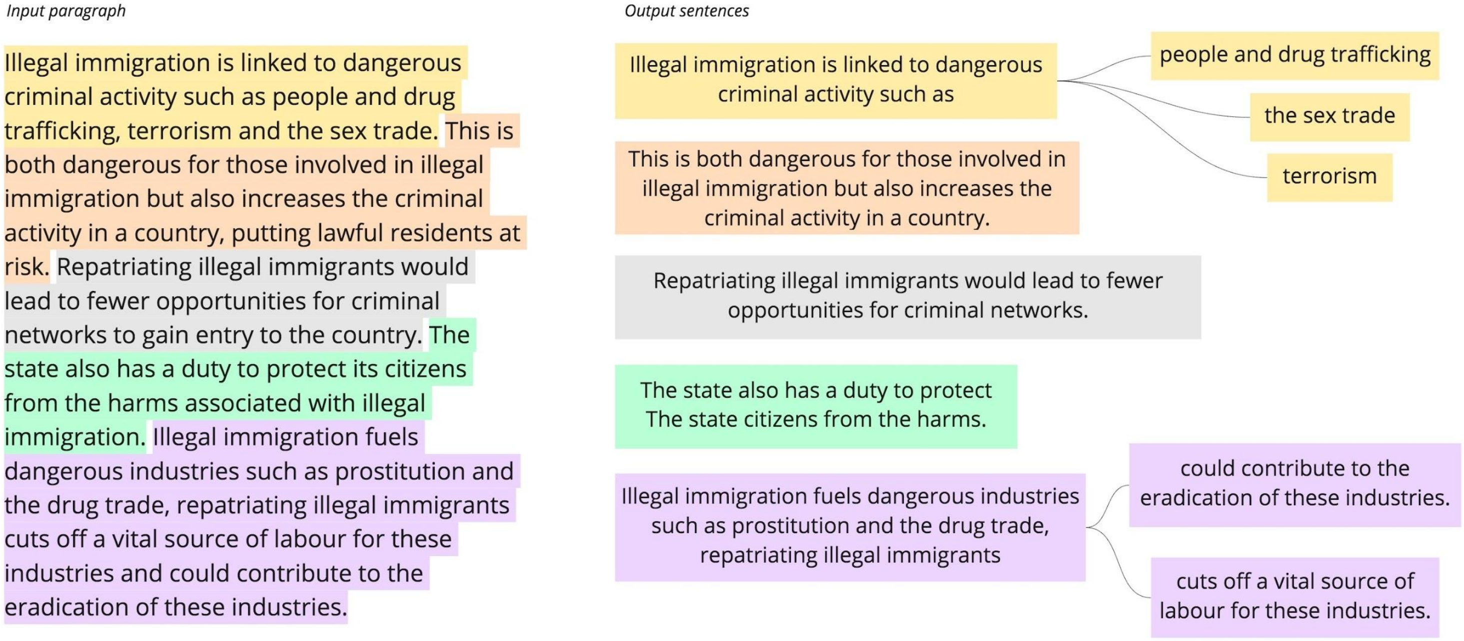
y=1 (probability **0.905**, score **2.251**) top features

4

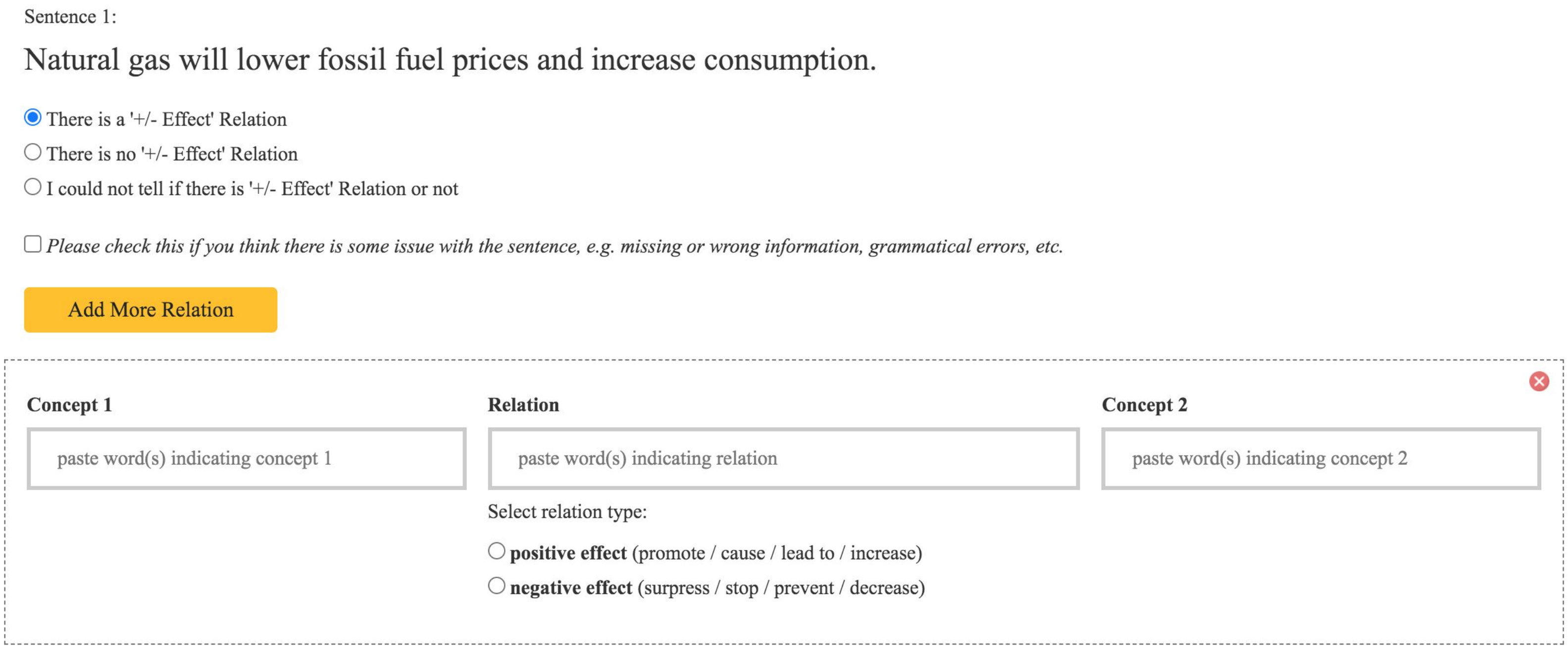
Contribution?	Feature
+2.408	Highlighted in text (sum)
-0.157	<BIAS>

[mask] can identify [mask] to increase [mask], prioritize [mask] and reduce [mask]

Figures


Input paragraph
Illegal immigration is linked to dangerous criminal activity such as people and drug trafficking, terrorism and the sex trade. This is both dangerous for those involved in illegal immigration but also increases the criminal activity in a country, putting lawful residents at risk. Repatriating illegal immigrants would lead to fewer opportunities for criminal networks to gain entry to the country. The state also has a duty to protect its citizens from the harms associated with illegal immigration. Illegal immigration fuels dangerous industries such as prostitution and the drug trade, repatriating illegal immigrants cuts off a vital source of labour for these industries and could contribute to the eradication of these industries.
Output sentences
Illegal immigration is linked to dangerous criminal activity such as
people and drug trafficking
the sex trade
terrorism
This is both dangerous for those involved in illegal immigration but also increases the criminal activity in a country.
Repatriating illegal immigrants would lead to fewer opportunities for criminal networks.
The state also has a duty to protect The state citizens from the harms.
Illegal immigration fuels dangerous industries such as prostitution and the drug trade, repatriating illegal immigrants
could contribute to the eradication of these industries.
cuts off a vital source of labour for these industries.

Figures



Sentence 1:
Natural gas will lower fossil fuel prices and increase consumption.
There is a '+/- Effect' Relation
There is no '+/- Effect' Relation
I could not tell if there is '+/- Effect' Relation or not
Please check this if you think there is some issue with the sentence, e.g. missing or wrong information, grammatical errors, etc.
Add More Relation
Concept 1
paste word(s) indicating concept 1
Relation
paste word(s) indicating relation
Concept 2
paste word(s) indicating concept 2
Select relation type:
positive effect (promote / cause / lead to / increase)
negative effect (surpress / stop / prevent / decrease)

Figures

Identify '+/- Effect' relation in a given sentence!

If this is your first HIT, please, read the task description and the examples carefully before working on the task!
We will validate your submission base on our requirement.

Task | Description | Examples | Comments

1. **Identify** if a sentence contains an effect relation between pairs of concepts mentioned in the sentence.

Example:

Social media helps to nurture your relationships.

- *Note:* Only annotate if the text **explicitly** supports the effect relation (either positive or negative) between 2 concepts, i.e. not using background knowledge or inference.

2. **Concept**: a phrase that expresses an entity (*Donald Trump*), event (*smoking in streets*), or an abstract principle/idea (*society*).

- *Note:* Demonstrative pronouns (this, that, these, those) or indefinite pronouns (something, everywhere, anybody, no-one) should not be considered as concrete concepts.
- *Note:* Be careful of positions of concept 1 and concept 2.
For example, in passive sentence, concept 1 comes **after** concept 2.

The greenhouse gases were produced by humans.

Concept 1	Relation	Concept 2
humans	produce	greenhouse gases

Select relation type
- ◉ positive effect
- ○ negative effect

3. **Effect relation types**: there could be two relation types between concept 1 and concept 2.

Positively (+) correlated:
Concept 1 'promotes / causes / leads to / increases / generates / protects etc.' **Concept 2**.
Example: "Smoking causes cancer."

Negatively (-) correlated:
Concept 1 'suppresses / stops / prevents / decreases etc.' **Concept 2**.
Example: "Sport prevents sickness."

- *Note :* **Neutral** relation is not considered as positive or negative effect relation.
For instance, you should choose **"No +/- Effect Relation"** for the following sentence:

Certain financial decision will have a big impact on our work.

- *Note:* **Negated statement** is not considered as positive or negative effect relation.
For example, you should also choose **"No +/- Effect Relation"** for the following sentence:

Smoking doesn't cause cancer

4. **Complex effect relation**: A compound-complex sentence may include multiple effect relations.

- Parallel effect relation

Social media can fuel anxiety and depression.

Concept 1	Relation	Concept 2
social media	fuel	anxiety

Select relation type
- ◉ positive effect
- ○ negative effect

Concept 1	Relation	Concept 2
social media	fuel	depression

Select relation type
- ◉ positive effect
- ○ negative effect

- Opposite effect relation

Natural gas will lower fossil fuel prices and increase consumption.

Concept 1	Relation	Concept 2
natural gas	lower	fossil fuel prices

Select relation type
- ○ positive effect
- ◉ negative effect

Concept 1	Relation	Concept 2
natural gas	increase	consumption

Select relation type
- ◉ positive effect
- ○ negative effect

- Continuous effect relation

Lower prices due to legalization of drugs will increase consumption.

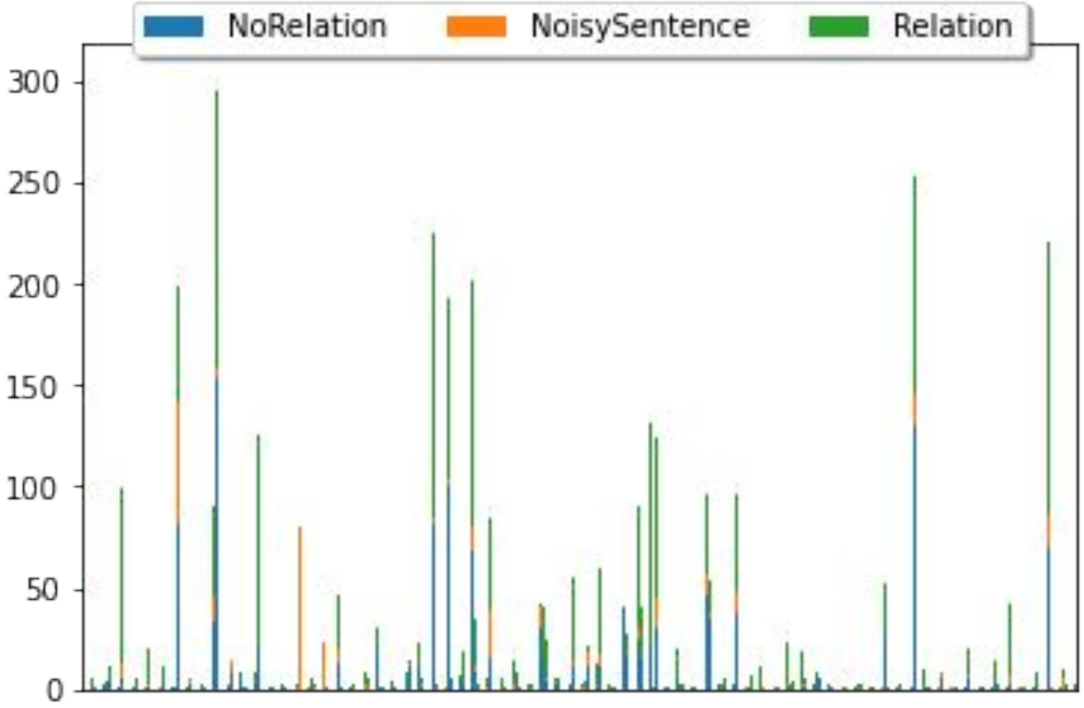
Concept 1	Relation	Concept 2
legalization of drugs	due to	lower prices

Select relation type
- ◉ positive effect
- ○ negative effect

Concept 1	Relation	Concept 2
lower prices	increase	consumption

Select relation type
- ◉ positive effect
- ○ negative effect

Figures



on the full 'manual annotated dataset'. Following are the types of instances that are filtered from our 'matching' sentences.

- Those with high confidence of *no effect* (agreed by *masking* and *non-masking effect detection* classifiers): **6103** sentences
- Those with high confidence of *effect* (agreed by *masking* and *non-masking effect detection* classifiers): **1615** sentences
- Those with some *positive effect relation* for sure (agreed by *masking* and *non-masking effect detection* classifiers and best *relation type* classifiers): **1828** sentences

With this filtering methods, we acquire in total **1,937** sentences left for crowd-

Tables

Number of matching sentences						
	debateorg	debatepedia	debatewise	idebate	parliament	sum
full	24,064	2,650	466	831	2	**27,793**
two-third	47,257	1,660	312	465	0	49,694
half	133,1995	40,171	23,654	32,743	257	1,428,820

Tables

Table 3.2: Concept matching after noise reduction statistics

Number of *full matched* sentences after noise reduction					
debateorg	debatepedia	debatewise	idebate	parliament	sum
9,302	613	241	173	0	10,329

Tables

	F_1 score						
Effect Detection	DistilBERT	ALBERT	BERT	RoBERTa	XLNET	NBSVM	Fasttext
non-Masking	0.88	0.88	0.88	**0.89**	**0.89**	0.81	0.79
Masking	0.84	0.62	0.85	**0.86**	**0.86**	0.79	0.79
Relation Type	DistilBERT	ALBERT	BERT	RoBERTa	XLNET	NBSVM	Fasttext
non-Masking	0.90	0.79	0.90	**0.93**	**0.91**	0.88	0.86
Masking	0.89	0.79	0.79	0.79	**0.86**	**0.88**	0.87

Tables

Table 4.2: Annotation Agreement

	Krippendorff Agreement scores			
	Effect Detection	Positive Relation Detection	Negative Relation Detection	Multiple Relation Detection
Expert	0.34	0.66	0.70	0.28
Public	0.27	0.31	0.36	0.03

Tables

Comparison of annotation results						
	Old dataset		New dataset			
			Experts		Public	
	#	%	#	%	#	%
Effect Detection						
Overall	4740	100	80	100	1324	100
Relation	1736	37	48	60	819	62
No Relation	3004	63	32	40	505	38
Relation Type						
Overall	1736	100	48	100	819	100
If Positive	1287	74	29	60	486	59
If Negative	390	23	29	60	507	62
Multiple Relation						
Overall	-	-	48	100	819	100
Single	-	-	34	71	607	75
Multiple	-	-	14	29	202	25

Tables

Table 5.1: Comparison of Effect Detection Classifiers

F_1 score		Test Set: Old x	Old M	New x	New M	Combined x	Combined M
Training Set: Old	x	0.88	0.84	0.63	0.57	0.82	0.78
	M	0.85	0.83	0.62	0.58	0.80	0.77
New	x	0.74	0.77	0.71	**0.75**	0.74	0.77
	M	0.67	0.69	0.62	0.70	0.66	0.70
Old + 25% New	x	0.86	0.83	0.68	0.59	0.82	0.78
	M	0.87	0.84	0.66	0.68	0.83	0.80
Old + 50% New	x	0.87	0.83	0.69	0.60	0.83	0.78
	M	0.87	**0.85**	0.65	0.68	0.82	0.81
Old + 75% New	x	0.88	0.80	0.70	0.61	0.84	0.76
	M	0.87	0.83	0.67	0.67	0.82	0.79
Old + 100% New	x	**0.89**	0.83	**0.75**	0.62	**0.85**	0.78
	M	0.88	**0.85**	0.70	0.70	0.84	**0.82**
Majority Class Baseline		0.64	0.64	0.53	0.53	0.62	0.62
Al-Khatib et al. [2020]		0.81	-	-	-	-	-

Tables

Table 5.2: Comparison of Positive Relation Detection Classifiers

F_1 score		Test Set: Old x	Old M	New x	New M	Combined x	Combined M
Training Set: Old	x	0.91	0.90	0.64	0.61	0.82	0.81
	M	0.90	**0.91**	0.72	0.74	0.84	0.85
Training Set: New	x	0.81	0.81	0.77	0.78	0.80	0.80
	M	0.74	0.87	**0.86**	0.79	0.78	0.85
Training Set: Old + New (Single)	x	0.90	**0.91**	0.77	0.77	0.86	0.86
	M	**0.92**	0.89	0.83	**0.84**	**0.89**	0.87
Training Set: Old + New (Single + Multiple)	x	0.91	**0.91**	0.74	0.71	0.86	0.84
	M	0.91	**0.91**	0.83	0.82	**0.89**	**0.88**
Majority Class Baseline		0.79	0.79	0.69	0.69	0.78	0.78
Al-Khatib et al. [2020]		0.86	-	-	-	-	-

Tables

Table 5.3: Comparison of Negative Relation Detection Classifiers

F_1 score			Test Set Old x	Old M	New x	New M	Combined x	Combined M
Training Set	Old	x	0.90	0.90	0.79	0.76	0.86	0.85
		M	0.90	**0.91**	0.73	0.78	0.85	0.87
	New	x	0.77	0.80	0.75	0.79	0.77	0.80
		M	0.77	0.81	0.81	0.76	0.79	0.79
	Old + New (Single)	x	**0.92**	0.90	0.74	0.74	0.86	0.85
		M	0.90	0.90	0.82	**0.81**	0.87	**0.87**
	Old + New (Single + Multiple)	x	0.91	0.89	**0.83**	0.79	**0.88**	0.85
		M	**0.92**	**0.91**	0.72	0.75	0.86	0.86
Majority Class Baseline			0.79	0.79	0.66	0.66	0.77	0.77
Al-Khatib et al. [2020]			0.86	-	-	-	-	-