Key Point Analysis via Contrastive Learning and Extractive Argument Summarization

Milad Alshomary, Timon Gurcke, Shahbaz Syed, Philipp Heinrich, Maximilian Spliethöver, Phlipp Cimiano, Martin Potthast, and Henning Wachsmuth

Key Point Analysis Shared Task at ArgMining 2021







Key Point Analysis Task

 Given a collection of natural language arguments, generate a set of key points and their salience scores, that summarize the collection

[----]

I do not trust that a solution shot into my child's arm can protect them. There is a risk of a serious reaction or even death. I as a parent should decide.

Forcing people to have their children vaccinated goes against free will.

I do not agree to force children without parental consent should not be fair

parents should have the freedom to decide what they consider best for their children

Routine child vaccinations isn't mandatory since children don't spread the virus

The child population has a low degree of vulnerability, so vaccination is not urgent yet

[----]

Key Point Analysis Task

 Given a collection of natural language arguments, generate a set of key points and their salience scores, that summarize the collection

[----]

I do not trust that a solution shot into my child's arm can protect them. There is a risk of a serious reaction or even death. I as a parent should decide.

Forcing people to have their children vaccinated goes against free will.

I do not agree to force children without parental consent should not be fair

parents should have the freedom to decide what they consider best for their children

Routine child vaccinations isn't mandatory since children don't spread the virus

The child population has a low degree of vulnerability, so vaccination is not urgent yet

[----]

1) Key point generation

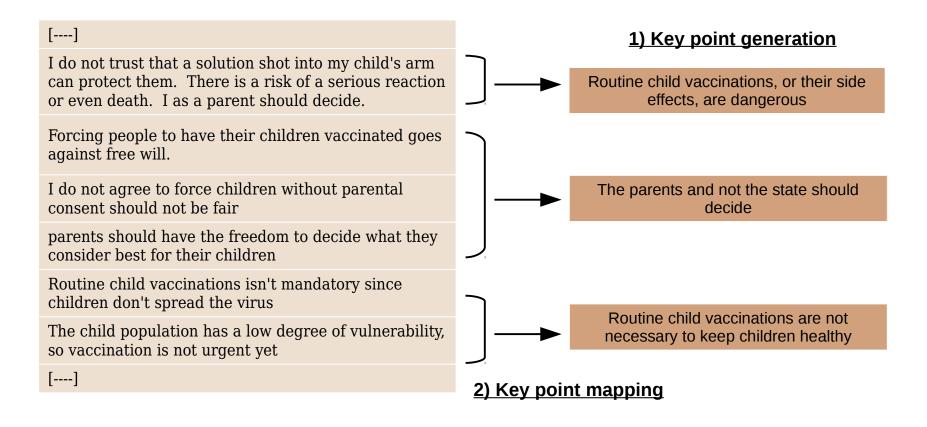
Routine child vaccinations, or their side effects, are dangerous

The parents and not the state should decide

Routine child vaccinations are not necessary to keep children healthy

Key Point Analysis Task

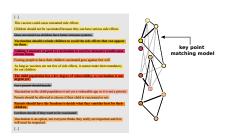
 Given a collection of natural language arguments, generate a set of key points and their salience scores, that summarize the collection



Approach

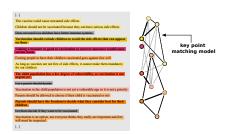
Approach

1) Key point generation: A graph based approach to extract key points from the input arguments

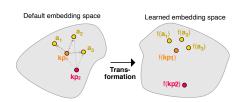


Approach

1) Key point generation: A graph based approach to extract key points from the input arguments



2) Key point matching: A contrastive learning approach to learn a new embedding space that represent the argument key point matching criterion



• **Key Idea:** Extractive summarization by modeling *centrality* and *argumentativeness* in sentences (Alshomary et al. 2020)

 Key Idea: Extractive summarization by modeling centrality and argumentativeness in sentences (Alshomary et al. 2020)

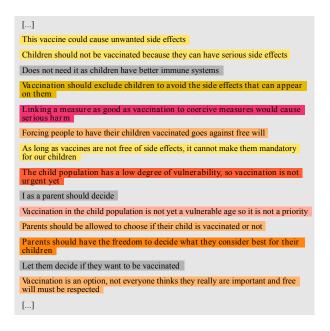
Procedure:

1) Arguments to a collection of sentences

[...] This vaccine could cause unwanted side effects Children should not be vaccinated because they can have serious side effects Does not need it as children have better immune systems Vaccination should exclude children to avoid the side effects that can appear Linking a measure as good as vaccination to coercive measures would cause Forcing people to have their children vaccinated goes against free will As long as vaccines are not free of side effects, it cannot make them mandatory The child population has a low degree of vulnerability, so vaccination is not urgent yet I as a parent should decide Vaccination in the child population is not yet a vulnerable age so it is not a priority Parents should be allowed to choose if their child is vaccinated or not Parents should have the freedom to decide what they consider best for their Let them decide if they want to be vaccinated Vaccination is an option, not everyone thinks they really are important and free will must be respected [...]

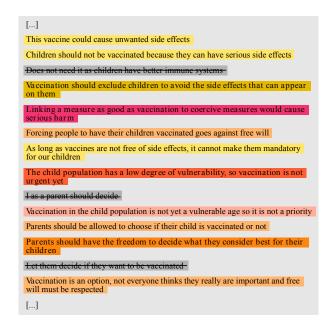
 Key Idea: Extractive summarization by modeling centrality and argumentativeness in sentences (Alshomary et al. 2020)

- 1) Arguments to a collection of sentences
- 2) Argumentative quality scores collected via *Project Debater's API*



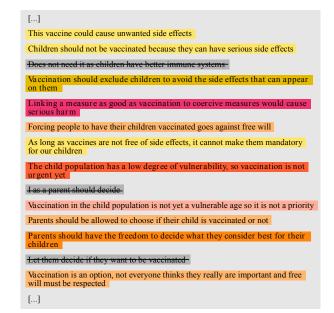
 Key Idea: Extractive summarization by modeling centrality and argumentativeness in sentences (Alshomary et al. 2020)

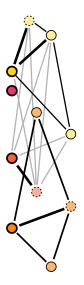
- 1) Arguments to a collection of sentences
- 2) Argumentative quality scores collected via *Project Debater's API*
- 3) Filter out sentences with argument quality score less than a minimum threshold



• **Key Idea:** Extractive summarization by modeling *centrality* and *argumentativeness* in sentences (Alshomary et al. 2020)

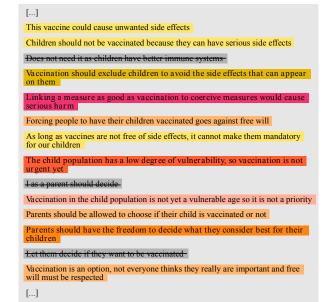
- 1) Arguments to a collection of sentences
- 2) Argumentative quality scores collected via *Project Debater's API*
- 3) Filter out sentences with argument quality score less than a minimum threshold
- 4) Build a graph with edges drawn between sentences with a minimum matching score

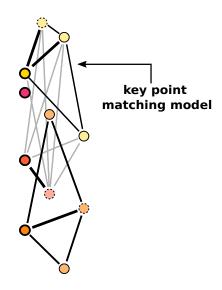




 Key Idea: Extractive summarization by modeling centrality and argumentativeness in sentences (Alshomary et al. 2020)

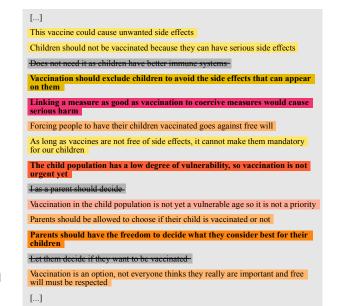
- 1) Arguments to a collection of sentences
- 2) Argumentative quality scores collected via *Project Debater's API*
- 3) Filter out sentences with argument quality score less than a minimum threshold
- 4) Build a graph with edges drawn between sentences with a minimum matching score

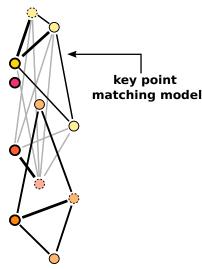




• **Key Idea:** Extractive summarization by modeling *centrality* and *argumentativeness* in sentences (Alshomary et al. 2020)

- 1) Arguments to a collection of sentences
- 2) Argumentative quality scores collected via *Project Debater's API*
- 3) Filter out sentences with argument quality score less than a minimum threshold
- 4) Build a graph with edges drawn between sentences with a minimum matching score
- 5) Run PageRank → Take top N sentences → Filter duplicates to ensure diversity

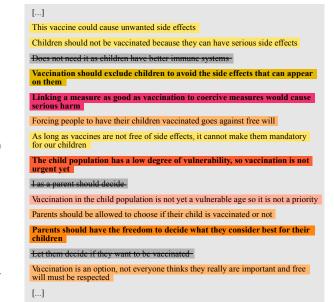


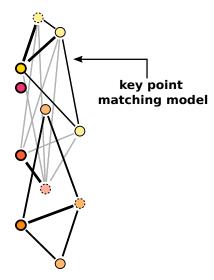


 Key Idea: Extractive summarization by modeling centrality and argumentativeness in sentences (Alshomary et al. 2020)

Implementation:

- Spacy for sentence splitting → sentences with a minimum of 5 and a maximum of 20 tokens are kept
- 2) Optimized on the training set to select best parameters: quality threshold: 0.2, Matching threshold: 0.8, and damping factor: 0.2
- 3) Post-processing to remove candidate key points that has a matching score higher than 0.8

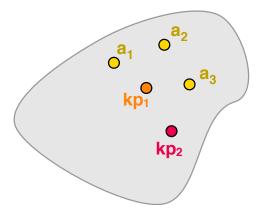




 Key Idea: mapping pairs of argument and key point into a new embedding space, where matching pairs become closer otherwise farther

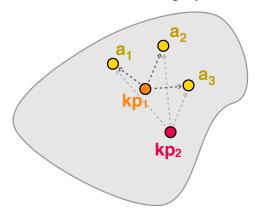
- Key Idea: mapping pairs of argument and key point into a new embedding space, where matching pairs become closer otherwise farther
- Procedure:
 - 1) Mapping arguments and key points into a default semantic space

Default embedding space



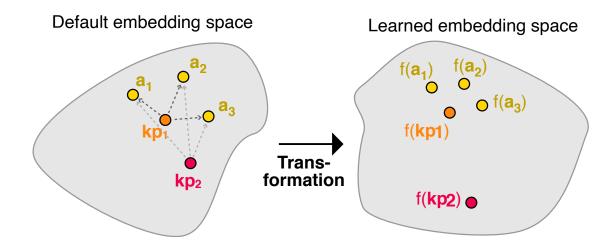
- Key Idea: mapping pairs of argument and key point into a new embedding space, where matching pairs become closer otherwise farther
- Procedure:
 - 1) Mapping arguments and key points into a default semantic space

Default embedding space



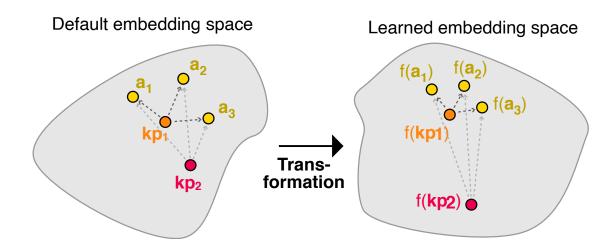
 Key Idea: mapping pairs of argument and key point into a new embedding space, where matching pairs become closer otherwise farther

- 1) Mapping arguments and key points into a default semantic space
- 2) Transform to a new embedding space reflecting argument and key point matching criterion

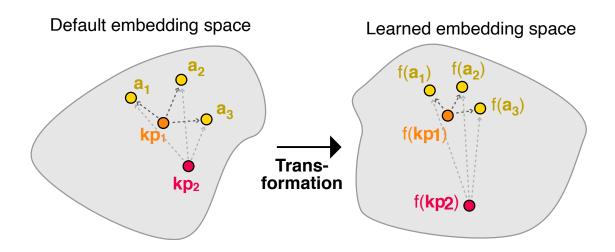


 Key Idea: mapping pairs of argument and key point into a new embedding space, where matching pairs become closer otherwise farther

- 1) Mapping arguments and key points into a default semantic space
- 2) Transform to a new embedding space reflecting argument and key point matching criterion



- Key Idea: mapping pairs of argument and key point into a new embedding space, where matching pairs become closer otherwise farther
- Implementation:
 - Two inputs: An argument and concatenation of the topic and the key point
 - Roberta-large (Liu and Lapata, 2019) as an encoder
 - Mean-pooling layer to map encoder's outputs into a sentence embedding
 - Cosine similarity as a similarity function



Results

TRACK1: Key points matching

Model	Rank	Strict mAP (r)	Relaxed mAP (r)	Mean of mAP (r)	Strict p@50% (r)
SMatchToPR *	1	0.789 (1)	0.927 (4)	0.858 (1)	0.848 (1)
NLP@UIT *	2	0.746 (3)	0.93(3)	0.838 (2)	0.827 (3)
ModrnTalk *	3	0.754(2)	0.902 (6)	0.828 (4)	0.806 (5)
Enigma *	4	0.739 (5)	0.928 (4)	0.833 (3)	0.828 (2)
MatchTStm *	5	0.745 (4)	0.902 (6)	0.824 (5)	0.808 (4)

TRACK2: Key points generation and matching

Model	Rank	mAP (r)	KP set rank
BarH	1	0.885 (1)	1 / (1,1,2)
SMatchToPR	2	0.818 (2)	2 / (2,1,2)
Enigma	3	0.491 (3)	4 / (2,4,3)
XLNet	4	0.443 (4)	3 / (4,3,1)
pa_vincent	5	0.096 (5)	None
mozhiwen	6	0 (6)	None