Vaccine Attitude Detection in Social Media via Disentangled Learning of Stance and Aspect Topics

Lixing Zhu, Zheng Fang, Gabriele Pergola, Robert Procter, and Yulan He

September 14th, 2022
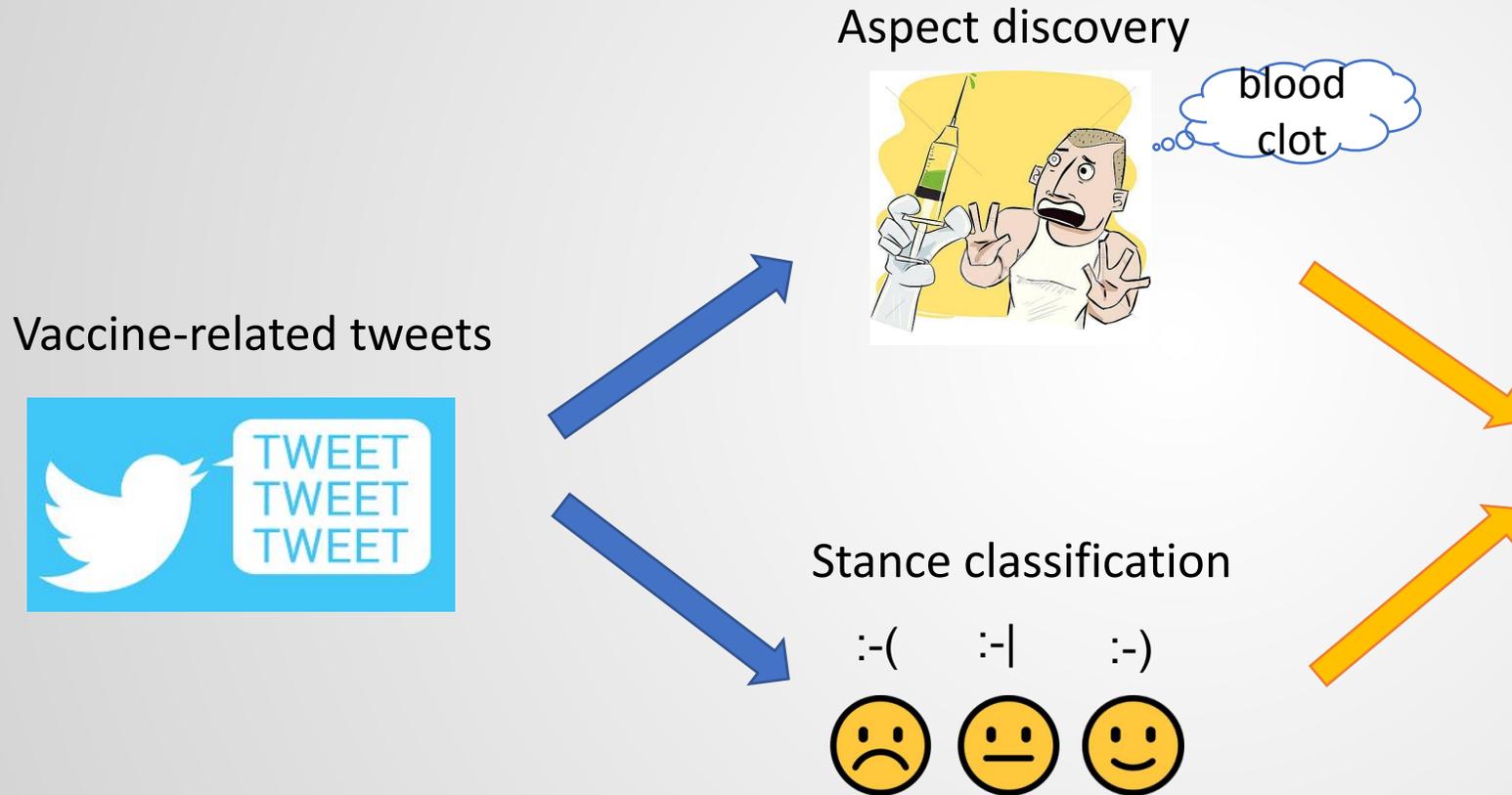
CONTENTS

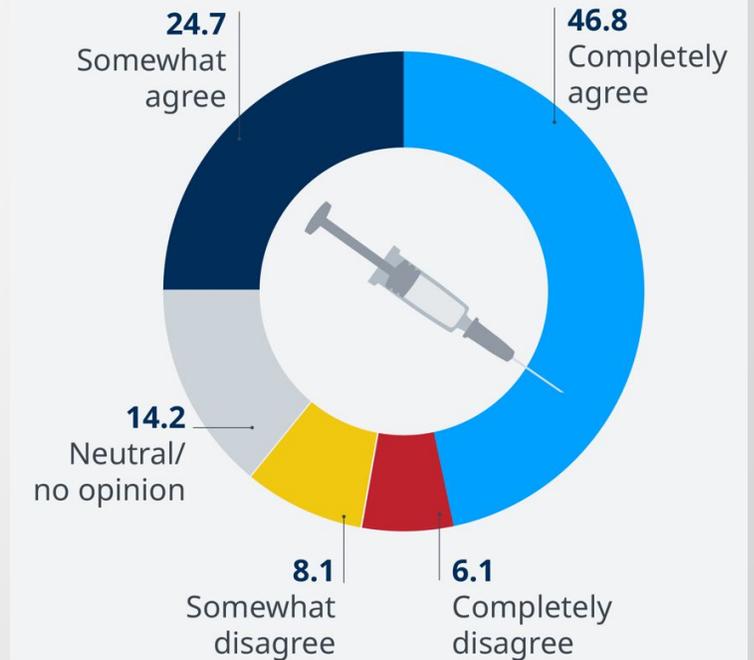**1** Motivation

2 Contribution

3 Architecture

4 Experiments

01
Motivation

Vaccine attitude detection: what is it?

Vaccine-related tweets
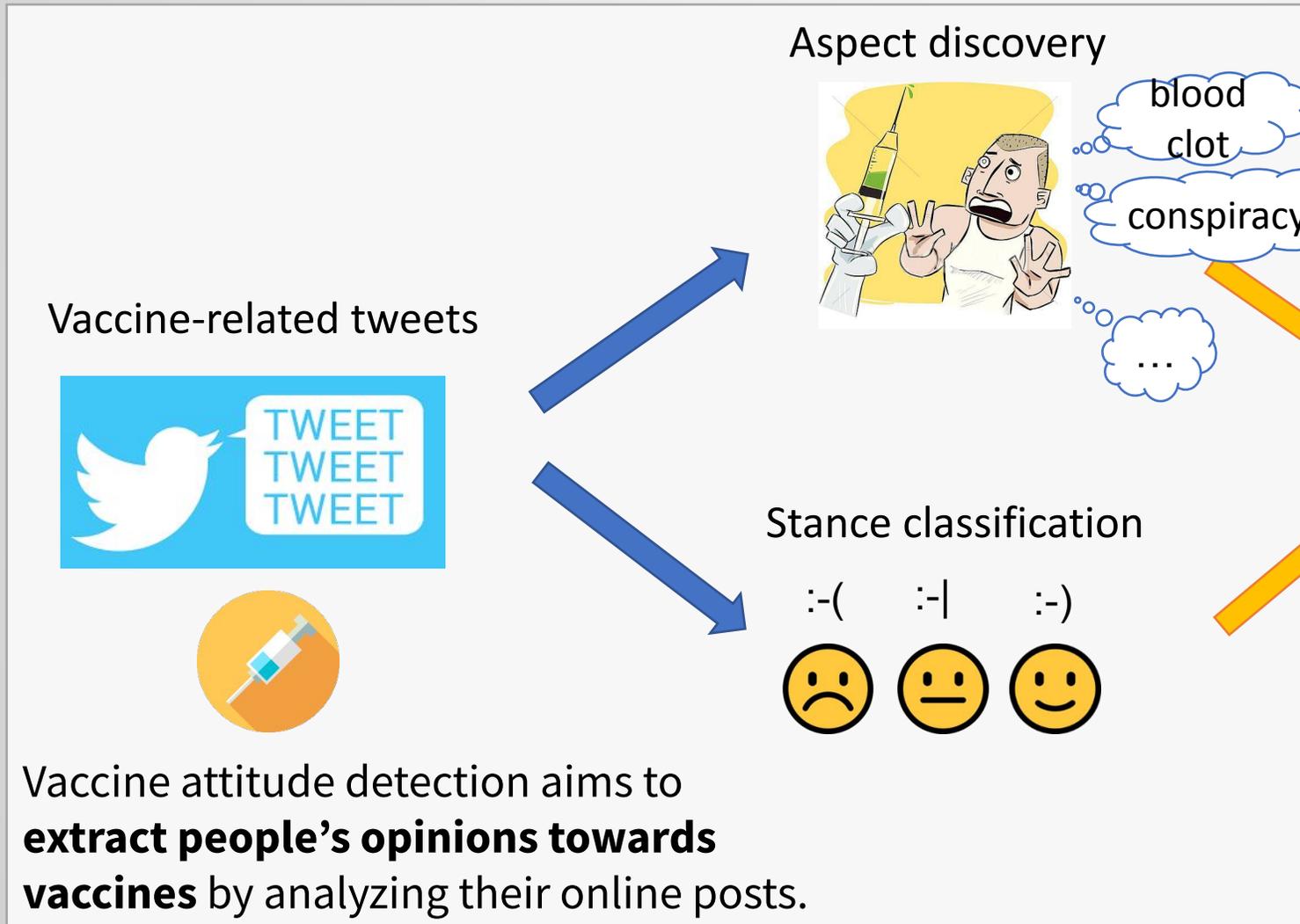
Aspect discovery

blood clot

Stance classification

:-(:-| :-)

Aspect-based stance classification

If a COVID-19 vaccine is shown to be safe and effective and is available to me, I will take it

24.7 Somewhat agree

46.8 Completely agree

14.2 Neutral/ no opinion

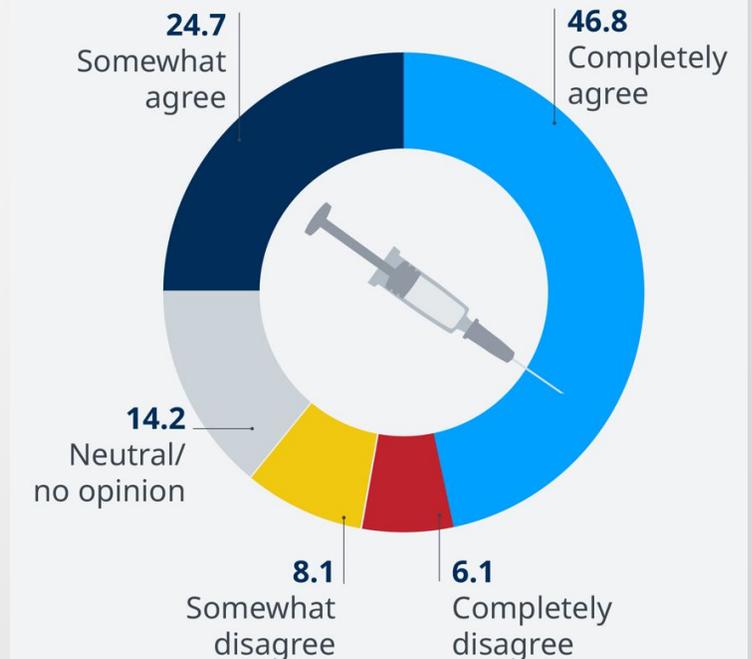**8.1** Somewhat disagree

6.1 Completely disagree

1

1. Survey: Majority of people would get vaccinated. **Nature (June 2020) global survey of 13,000+ people across 35 countries most affected by COVID-19**. https://www.nature.com/articles/s41591-020-1124-9

Vaccine attitude detection: what is it?

Aspect discovery

blood
clot

conspiracy

...

Vaccine-related tweets

TWEET
TWEET
TWEET

Stance classification

:-(:-| :-)

☹ 😐 🙂

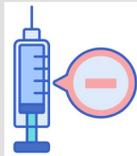Vaccine attitude detection aims to **extract people's opinions towards vaccines** by analyzing their online posts.

Aspect-based stance classification

If a COVID-19 vaccine is shown to be safe and effective and is available to me, I will take it

24.7 Somewhat agree

46.8 Completely agree

14.2 Neutral/ no opinion

8.1 Somewhat disagree

6.1 Completely disagree

1

1. Survey: Majority of people would get vaccinated. **Nature (June 2020) global survey of 13,000+ people across 35 countries most affected by COVID-19.** https://www.nature.com/articles/s41591-020-1124-9

Vaccine attitude detection: why is it hard?

Vaccine attitude detection: why is it hard?

Vaccine attitude detection aims to **extract people's opinions towards vaccines** by analyzing their online posts.

- User attitude emerges from positive or negative **stances** related to a wide spectrum of **aspects**:
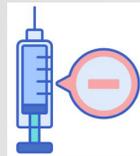
Individual and group influences
Vaccine safety
Lack of information
Low risk/severity of disease
Vaccines not effective
Mistrust in health institutions
Healthy bodies belief
Social norms
Vaccination not a priority
Against vaccination in general
Alternative prevention methods
Diseases are beneficial
Fear of injection
Previous negative experiences
Humans too weak to fight vaccines
Responsibility if something bad happens

Vaccine and vaccination specific issues
No medical need
Access
Financial cost
Lack of recommendation from providers
Vaccine novelty
Inconsistent advice from providers

Contextual influences
Conspiracy theories
Religious fatalism
Negative exposure to media
Violation of human rights

2

2. **Rapid literature review on motivating hesitant population groups in Europe to vaccinate**. https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/vaccination-motivating-hesistant-populations-europe-literature-review.pdf

Vaccine attitude detection: why is it hard?

Galileo Galilei
@TheREALGalileo

Follow

Have felt for the past 24 hours that I've been run over by three double decker buses after the AstraZeneca vaccine yesterday morning. Starting to feel a little normal now but it's not been nice!

💬 10   🔁 189   ♡ 964   ⬆

Nicolaus Copernicus
@TheREALCopernicus

Follow

The AstraZeneca one is rough for up to 48 hours; after that you may still be a bit swollen but you'll basically feel fine. I've had that and the virus, and the vaccine is far less unpleasant.

Vaccine attitude detection: why is it hard?

Vaccine attitude detection aims to **extract people's opinions towards vaccines** by analyzing their online posts.

- User attitude emerges from positive or negative **stances** related to a wide spectrum of **aspects.**

A high (and growing) number of aspects

Limited availability of annotated data

02
Contribution

VAD<small>ET</small>: Vaccine-Attitude Detection Model

The goal of our work is to :

- **detect the stance** expressed in a tweet (i.e., 'pro-vaccination', 'antivaccination', or 'neutral'),

- **identify a text span** that indicates the concerning **aspect** of vaccination,

- **cluster** tweets into groups that share similar aspects.

VADᴇᴛ: Vaccine-Attitude Detection Model

The goal of our work is to :

- **detect the stance** expressed in a tweet (i.e., 'pro-vaccination', 'antivaccination', or 'neutral'),

- **identify a text span** that indicates the concerning **aspect** of vaccination,

- **cluster** tweets into groups that share similar aspects.

Vaccine Attitude Detection (**VADᴇᴛ**) Model

1. Initially trained on a large amount of **unannotated** Twitter data to learn latent **topics**

2. Then, it is fine-tuned on a small amount of Twitter data **annotated with stance labels** and **aspect text spans**

VADᴇᴛ: Vaccine-Attitude Detection Model

The goal of our work is to :

- **detect the stance** expressed in a tweet (i.e., 'pro-vaccination', 'antivaccination', or 'neutral'),

- **identify a text span** that indicates the concerning **aspect** of vaccination,

- **cluster** tweets into groups that share similar aspects.

Vaccine Attitude Detection (**VADᴇᴛ**) Model

1. Initially trained on a large amount of **unannotated** Twitter data to learn latent **topics**

2. Then, it is fine-tuned on a small amount of Twitter data **annotated with stance labels** and **aspect text spans**
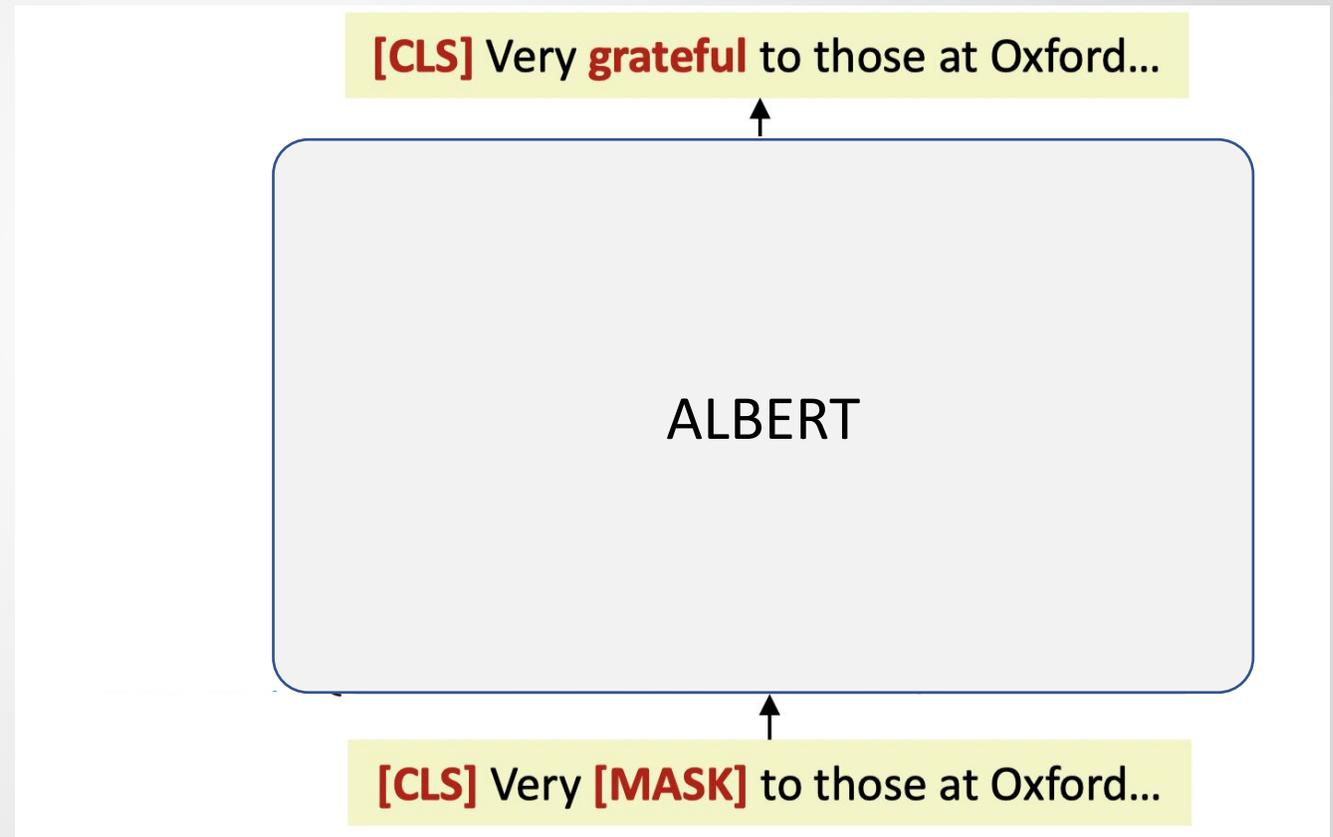
Vaccine Attitude Dataset (**VAD**)

03

Architecture

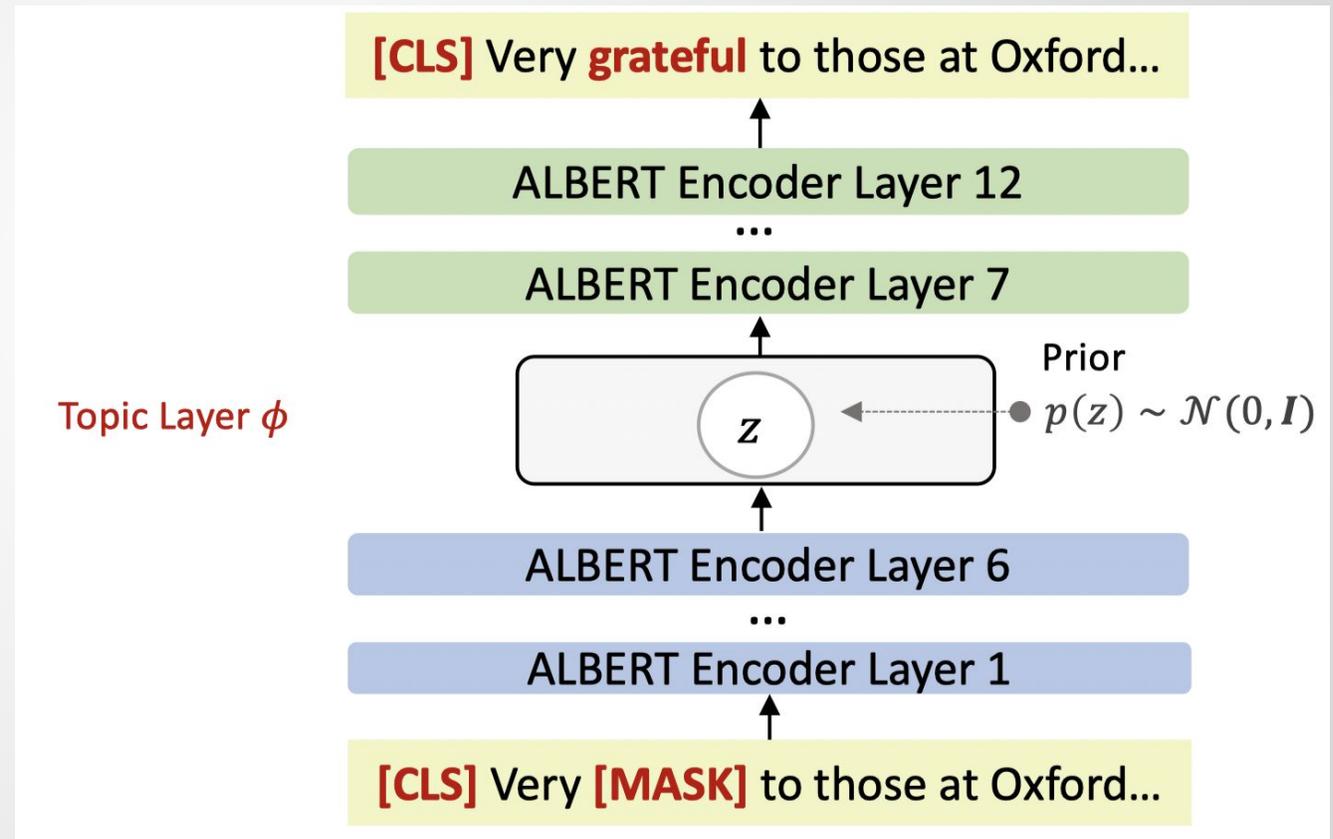# VADᴇᴛ: Vaccine-Attitude Detection Model

1. Unsupervised Step

Language model enrichment via
a Topic Layer

[CLS] Very **grateful** to those at Oxford…

ALBERT

[CLS] Very **[MASK]** to those at Oxford…

VADᴇᴛ: Vaccine-Attitude Detection Model

1. Unsupervised Step

Language model enrichment via
a Topic Layer

[CLS] Very **grateful** to those at Oxford…

ALBERT Encoder Layer 12
…
ALBERT Encoder Layer 7

Topic Layer ϕ

z

Prior
$p(z) \sim \mathcal{N}(0, I)$

ALBERT Encoder Layer 6
…
ALBERT Encoder Layer 1

[CLS] Very **[MASK]** to those at Oxford…

1. Unsupervised Step

Language model enrichment via a Topic Layer

- Training Objective: Maximize the Evidence Lower Bound (ELBO):

$$\mathbb{E}_{q_\phi(z|\psi(\mathbf{w}))}[\log p_\theta(\mathbf{w}|z, \psi(\mathbf{w}))] - \mathrm{KL}[q_\phi(z|\psi(\mathbf{w}))||p(z)]$$

[CLS] Very **grateful** to those at Oxford…

$p_\theta(\boldsymbol{w^H}|z, \psi(\boldsymbol{w}))$ — ALBERT Encoder Layer 12
···
VAE Decoder θ — ALBERT Encoder Layer 7

Topic Layer ϕ — z — Prior $p(z) \sim \mathcal{N}(0, \boldsymbol{I})$

$q_\phi(z|\psi(\boldsymbol{w}))$ — ALBERT Encoder Layer 6
···
VAE Encoder ψ — ALBERT Encoder Layer 1

[CLS] Very **[MASK]** to those at Oxford…
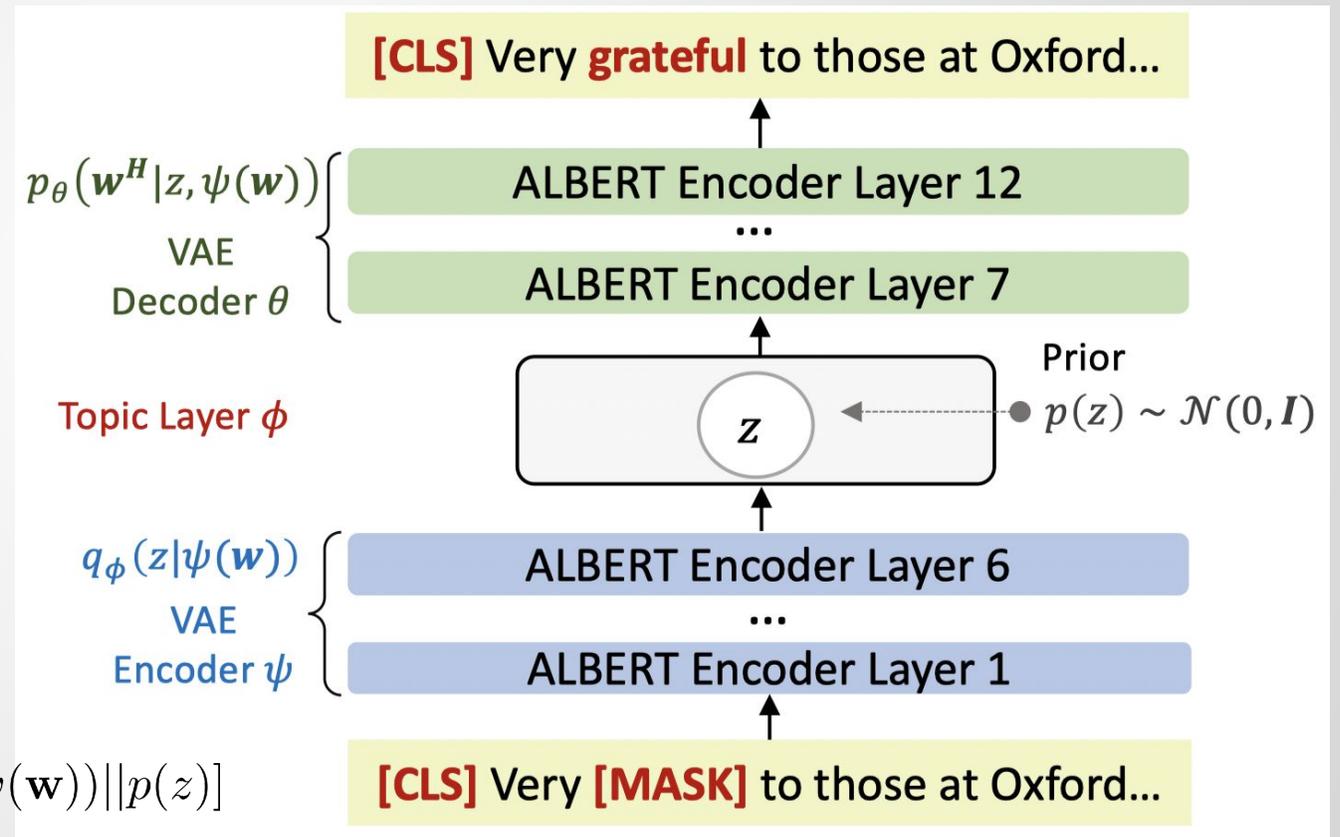
2. Supervised step

- A pre-trained language model is applied to the aspect span.

- Aspect span is generated with

$$\mathcal{L}_A = \mathbb{E}_{q_\phi(z_a|\psi(w_{a:b}))}[\log p_\theta(w_{a:b}|z_a, \psi(w_{a:b}))] \\ -\mathrm{KL}[q_\phi(z_a|\psi(w_{a:b}))||p(z_a)],$$
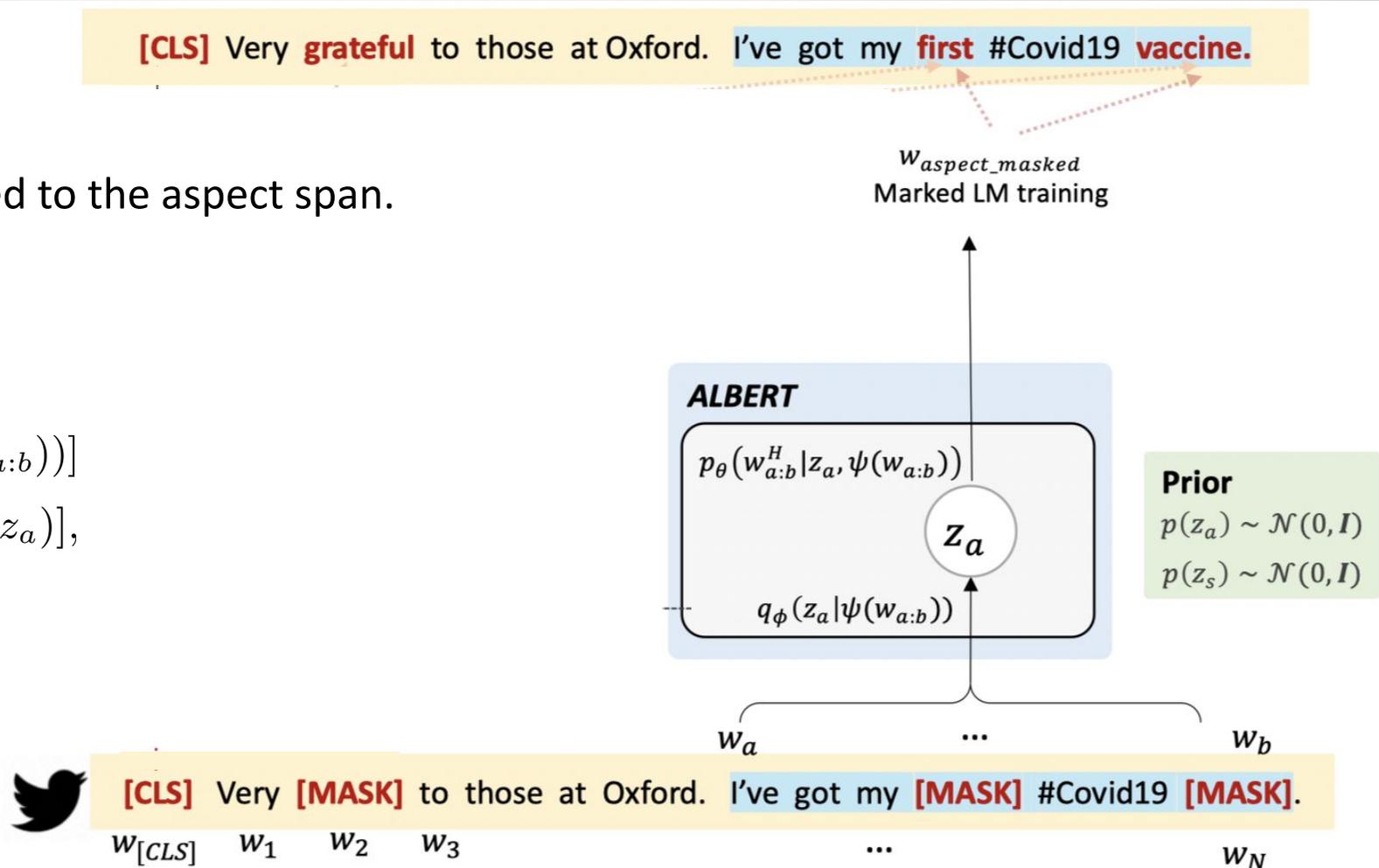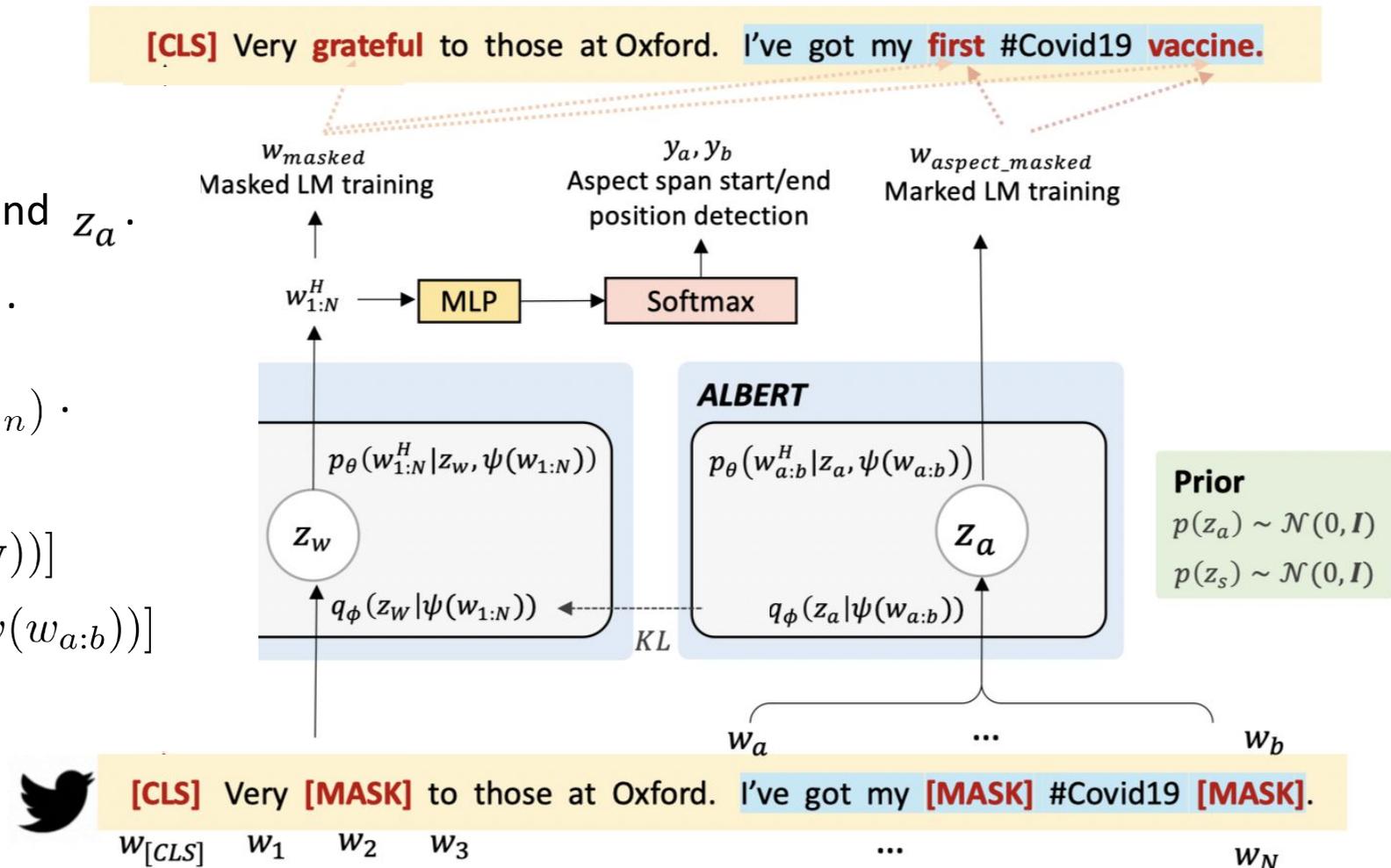
2. Supervised step

▪ Enforce an association between z_w and z_a.

▪ Make $w_{[\text{CLS}]}$ solely dependent on z_s.

▪ Make z_w solely dependent on $\psi(w_{1:n})$.

$$\mathcal{L}_S = \mathbb{E}_{q_\phi(z_w)}\mathbb{E}_{q_\phi(z_s)}[\log p_\theta(\mathbf{w}|z, \psi(\mathbf{w}))]$$
$$- \text{KL}\left[q_\phi(z_w|\psi(w_{1:n}))||q_\phi(z_w|\psi(w_{a:b}))\right]$$
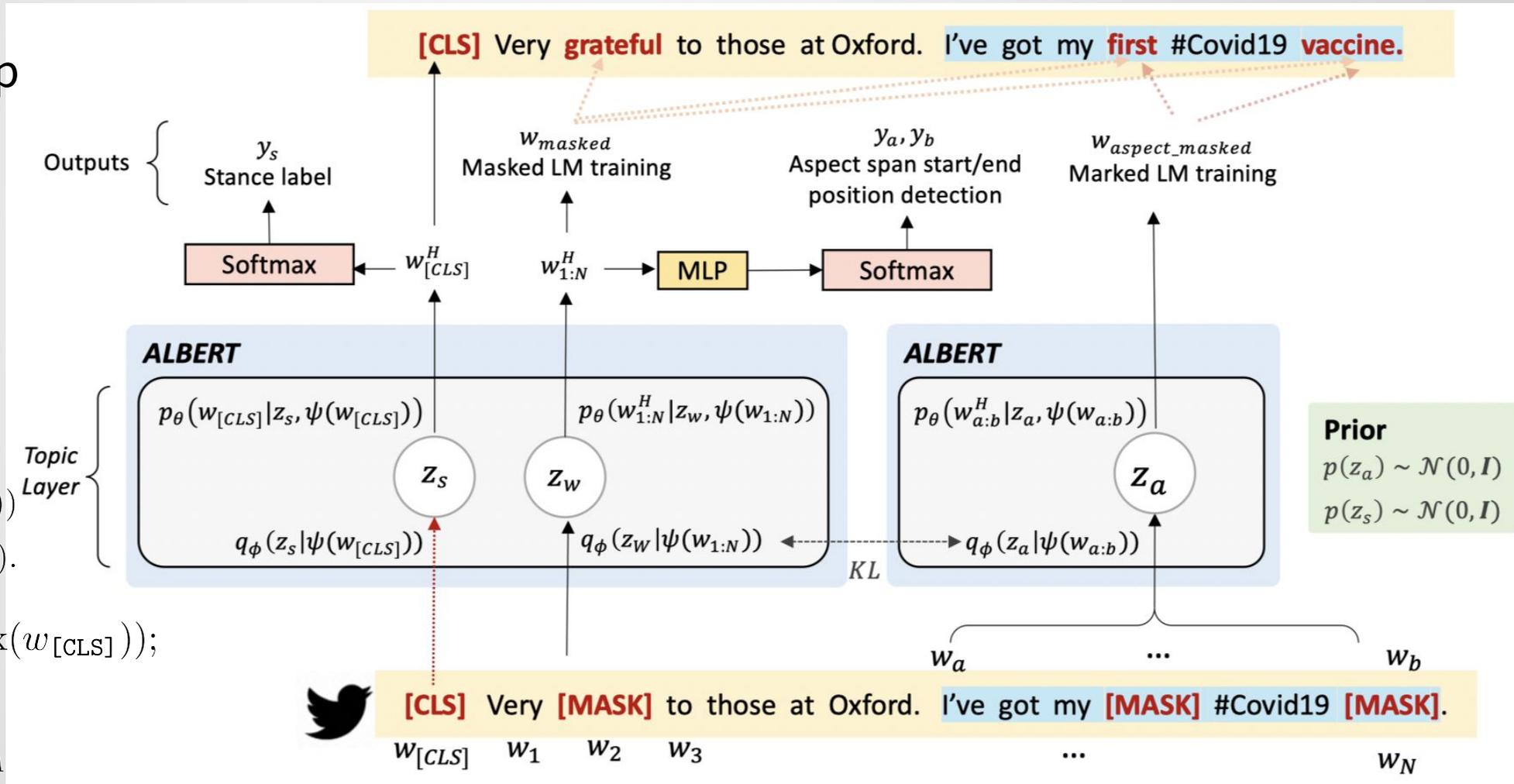$$- \text{KL}[q_\phi(z_s|\psi(\mathbf{w})))||p(z_s)].$$

2. Supervised step

- Use text span and stance labels: predict the starting and ending position

$$\mathcal{L}_a = -\log p(y_a|\mathrm{MLP}(w_{1:n}^H))$$
$$-\log p(y_b|\mathrm{MLP}(w_{1:n}^H)).$$

$$\mathcal{L}_s = -\log p(y_c|\mathrm{Softmax}(w_{[CLS]}));$$

$$\mathcal{L} = \mathcal{L}_s + \mathcal{L}_a - \mathcal{L}_S - \mathcal{L}_A$$

Experimental Assessment:
The proposed VAD Dataset

- **VAD: Vaccine Attitude Dataset:**
    - 1.9 million English tweets collected between February 7th and April 3rd, 2021 using 60 predefined keywords relating to COVID-19 vaccines
    - 2,800 tweets have been annotated in which 2,000 are used for training and the remaining 800 are used for testing.
        - Each tweet is annotated with a stance label and a text span.
- An example of annotation:

[1]: "An overview of the BioASQ largescale biomedical semantic indexing and question answering competition", Tsatsaronis et al., BMC Bioinformatics 2015
[2]: "The Covid-19 open research dataset", L. Wang et al. 2020
[3]: "Rapidly bootstrapping a question answering dataset for COVID-19" R. Tang et al. 2020

Experimental Assessment:
The proposed VAD Dataset

- **VAD: Vaccine Attitude Dataset:**
  - *1.9 million English tweets* collected between February 7th and April 3rd, 2021 using 60 predefined keywords 3 relating to COVID-19 vaccines
  - *2,800 tweets have been annotated* of which 2,000 are used for training and the remaining 800 are used for testing.
    - Each tweet is annotated with a stance label and a text span.

- **VC** (Morante et al., 2020) is a **v**accination **c**orpus consisting of 294 web documents about online vaccine debate annotated with events, 210 of which are annotated with opinions (in the form of text spans) towards vaccines.
  - An example of annotation:

  - The Hepatitis B vaccine is considered [one of the safest and most effective vaccines ever made].

  - The establishment media is [desperately pushing the myth]$_{expression\_neg}$ that vaccines are completely safe and effective.

2. Roser Morante, Chantal van Son, Isa Maks, and Piek Vossen. 2020. **Annotating perspectives on vaccination**. In LREC, pages 4964–4973, Marseille, France. European Language Resources Association.

VADET – Classification and Clustering

- **Baselines**

  - BertQA (Li et al., 2018b) – a pre-trained LM well-suited for relevant span detection. We rely on its HuggingFace4 (Wolf et al., 2020) implementation. We employ ALBERT (Lan et al., 2020) as the backbone language model for both BertQA and VADET.

  - ASTE (Peng et al., 2020) – a pipeline approach consisting of aspect extraction (Li et al., 2018b) and sentiment labelling (Li et al., 2018a).

  - Evaluation Metric – exact match accuracy and Macro-averaged F1 score (Rajpurkar et al., 2016).

VADET – Classification and Clustering

■ Results on stance classification and aspect span extraction.

Table 1

Model \ Dataset	VAD		VC	
	Acc.	F1	Acc.	F1
BertQA	0.754	0.742	0.719	0.708
ASTE	0.723	0.710	0.704	0.686
VAD$_{ET}$	0.763	0.756	0.727	0.713

■ Results on aspect span extraction.

Table 2

Model \ Dataset	VAD		VC	
	Acc.	F1	Acc.	F1
BertQA	0.546	0.722	0.525	0.670
ASTE	0.508	0.684	0.467	0.652
VAD$_{ET}$	0.556	0.745	0.541	0.697

VADET – Cluster Evaluation

@user **AstraZeneca vaccine proven to reduce hospitalisation of over 80s by 80% after just one dose** The opposition are full of _pile_of_poo_

Queensland senator and leader of the One Nation party Pauline Hanson is fighting back against **mandatory vaccines** in Australia . #NoMandatoryVaccines #MyBodyMyChoice #VaccinePassports #Australia #MedicalApartheid

**personal freedom to choose in relation to vaccines**

Immunity level

@user I am not a Putin fan at all , but this **Sputnik vaccine seems to be one of the best in the business** . Certainly better than AstraZeneca's jab .

economic effect

A year ago **I tried the Moderna vaccine to see if it was safe** . Spoiler : It is ! Now , on my #COVIDvaccine anniversary , I??m happy to share that I just got a 3rd dose . This booster experiment will reveal 1 if strainadapted vaccines boost immunity 2 whether they are safe .

personal experiences

(adverse) side effects

The MHRA has published new vaccine adverse reaction data , so expect more in this series soon : ??Covid19 Vaccine?? Adverse Reactions; Part Two : **AstraZeneca Is Not A Safe Option** #Budget #BudgetSpeech2021 #Budget21 #RishiSunak #PMQs #COVID19Vaccine

religion, conspiracy or moral attitudes

@user Nothing cures or protects against viruses . **This is a trial vax , gene therapy** . mRNA changes DNA . This is mass depopulation/sterilisation there??s warning on vax websites , no sex for a month after a shot incase of sterilisation . It??s a politicised 99 . 98% survival rate FLU
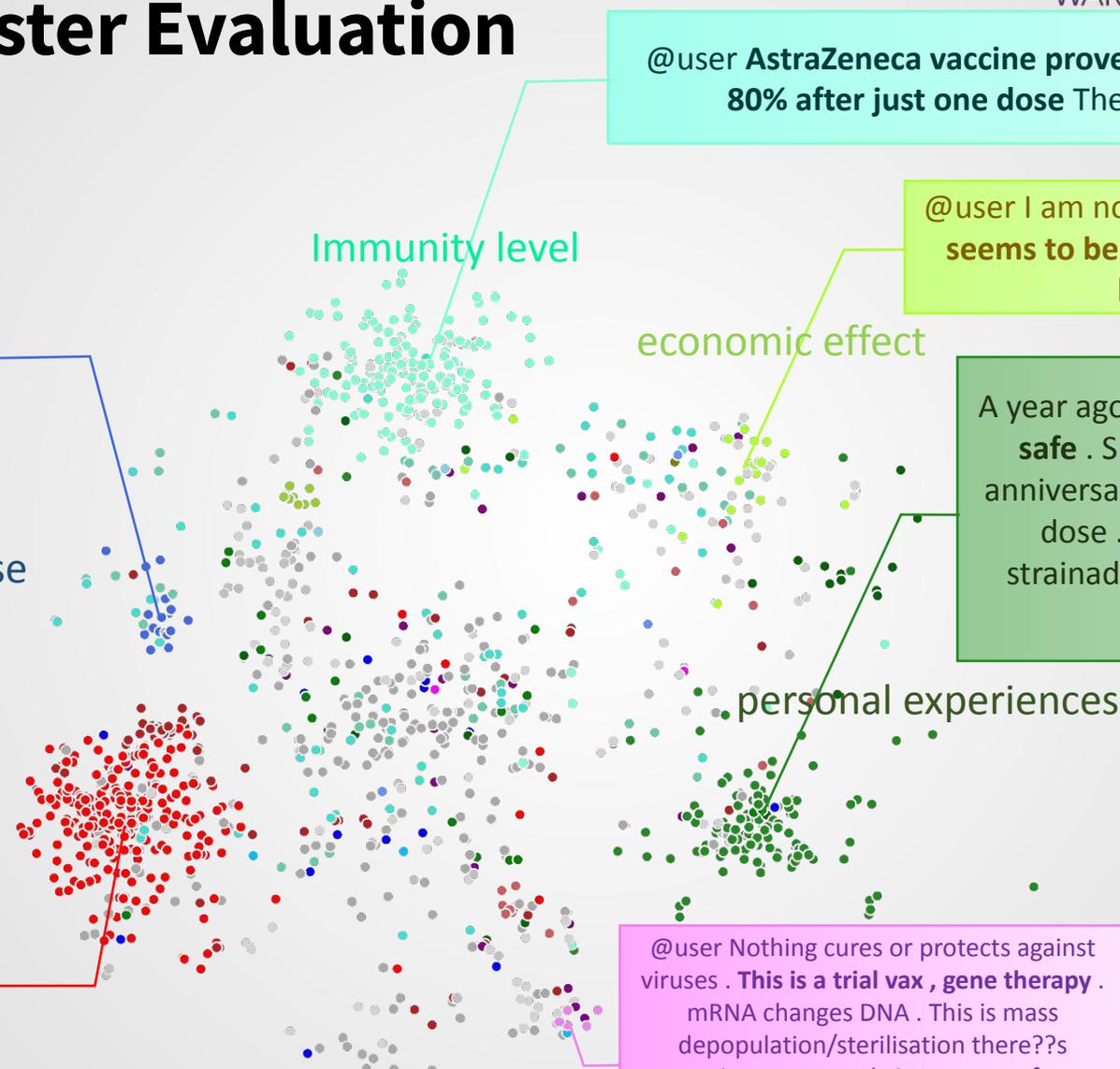
05
Conclusion

Conclusion, On-Going and Future Work

- We presented a **semi-supervised** model, **VADET,** to detect user attitudes and distinguish aspects of interest about vaccines on social media.

- The experimental assessment has shown benefits of such an approach for attitude detection and aspect clustering over two vaccine corpora.

- Fine-grained annotations on user experiences and stances

- One of the current limitations is that the model is mainly focused on **short text** of social media data. A future extension would explore the application to online debate corpora, where multiple arguments and opinions may appear in a sentence

Thank you

Questions are Welcomed

VADᴇᴛ – Cluster Evaluation

Fig. 2

Immunity level

economic effect

personal experiences

personal freedom to choose in relation to vaccines

(adverse) side effects

religion, conspiracy or moral attitudes

Queensland senator and leader of the One Nation party Pauline Hanson is fighting back against **mandatory vaccines** in Australia . #NoMandatoryVaccines #MyBodyMyChoice #VaccinePassports #Australia #MedicalApartheid
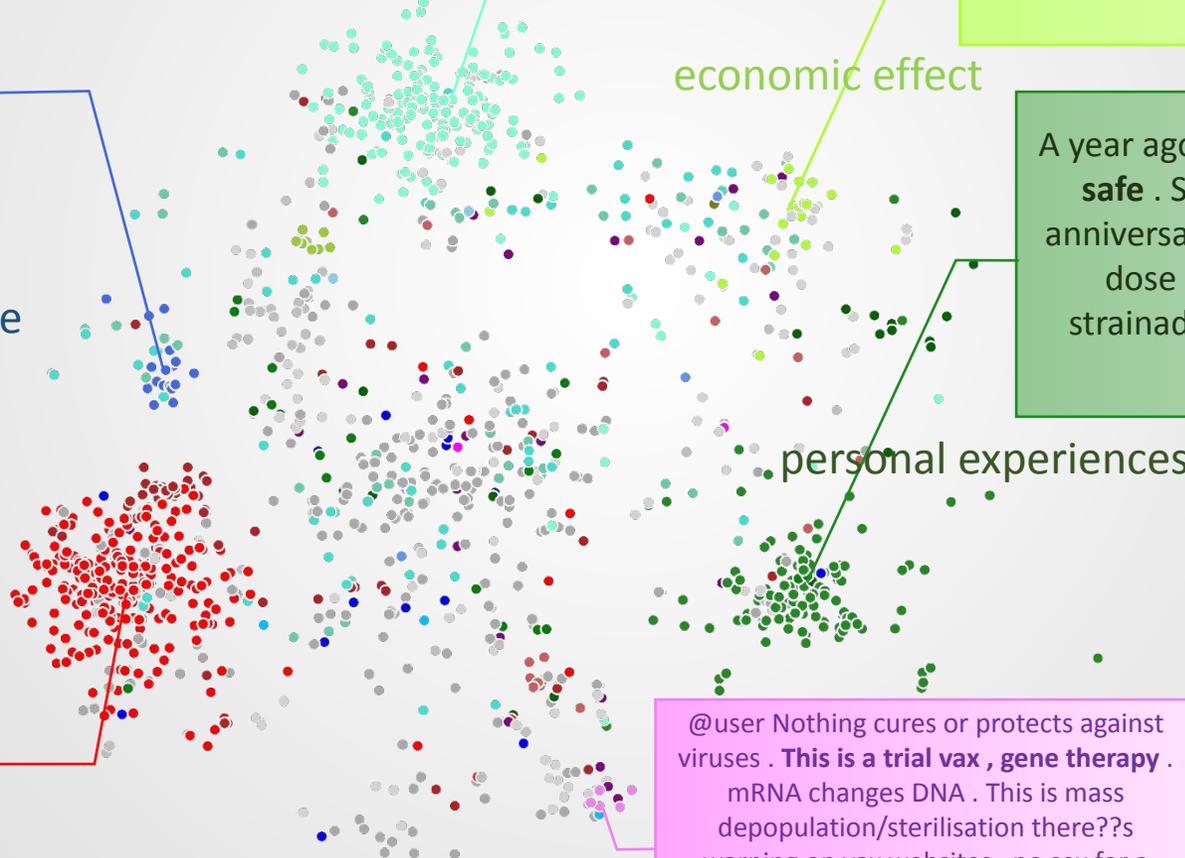
@user **AstraZeneca vaccine proven to reduce hospitalisation of over 80s by 80% after just one dose** The opposition are full of _pile_of_poo_

@user I am not a Putin fan at all , but this **Sputnik vaccine seems to be one of the best in the business** . Certainly better than AstraZeneca's jab .

A year ago **I tried the Moderna vaccine to see if it was safe** . Spoiler : It is ! Now , on my #COVIDvaccine anniversary , I??m happy to share that I just got a 3rd dose . This booster experiment will reveal 1 if strainadapted vaccines boost immunity 2 whether they are safe .

The MHRA has published new vaccine adverse reaction data , so expect more in this series soon : ??Covid19 Vaccine?? Adverse Reactions; Part Two : **AstraZeneca Is Not A Safe Option** #Budget #BudgetSpeech2021 #Budget21 #RishiSunak #PMQs #COVID19Vaccine

@user Nothing cures or protects against viruses . **This is a trial vax , gene therapy** . mRNA changes DNA . This is mass depopulation/sterilisation there??s warning on vax websites , no sex for a month after a shot incase of sterilisation . It??s a politicised 99 . 98% survival rate FLU

WARWICK
Queen Mary
Alan Turing Institute
UK research

0	8	16
1	9	17
2	10	18
3	11	19
4	12	20
5	13	21
6	14	22
7	15	23