Novax or Novak? Estimating Social Media Stance towards Celebrity Vaccine Hesitancy (Student Abstract)

Madhav Hota[1], **Adel Khorramrouz**[2], **Ashiqur R. KhudaBukhsh**[2]

[1]Illinois Mathematics and Science Academy
[2]Rochester Institute of Technology
mhota@imsa.edu, ak8480@rit.edu, axkvse@rit.edu

Abstract

On 15 January 2022, noted tennis player Novak Djokovic was deported from Australia due to his unvaccinated status for the COVID-19 vaccine. This paper presents a stance classifier and evaluates public reaction to this episode and the impact of this behavior on social media discourse on YouTube. We observed a significant spike of individuals who supported and opposed his behavior at the time of the episode. Supporters outnumbered those who opposed this behavior by over 4x. Our study reports a disturbing trend that following every major Djokovic win, even now, vaccine skeptics often conflate his tennis success as a fitting reply to vaccine mandates.

Introduction

Novak Djokovic is the most decorated male tennis player in the history of tennis who has publicly refused to receive the COVID-19 vaccine. In January 2022, he intended to compete in the Australian Open, but after a controversial process, was denied participation given a vaccine mandate for the country. Following a fiercely fought court trial loss, Djokovic was deported from Australia. He subsequently shared in the media his conviction that he was unwilling to be vaccinated for COVID-19 in the future. In a detailed interview with BBC[1], he reiterated how much he valued his freedom to choose and said that he was willing to pay the price by missing important tennis tournaments. This paper investigates the social media discourse on the controversial episode and examines the social media stance on Djokovic's position on the COVID-19 vaccine (Table 1 lists a few illustrative examples).

While our focus is, of course, analyzing the immediate public discourse following Djokovic's deportation, atypical to most papers on stance mining around controversial issues (Palakodety, KhudaBukhsh, and Carbonell 2020; Gautam et al. 2020), our analysis looks into the after-effect where we seek to understand how often vaccine skeptics consider Djokovic's tennis success as a fitting reply against vaccine mandates. To our knowledge, this is the first line of work that examines how global icons' vaccine stance can be used in furthering vaccine hesitancy.

[1]https://www.bbc.com/news/world-60354068

We investigate the following two research questions.

***RQ1*:** *What was the distribution of positive and negative stances toward Djokovic's vaccine position in social media discourse following his deportation from Australia?*

***RQ2*:** *Is there lingering support for Djokovic's vaccine hesitancy as he keeps breaking tennis world records?*

Dataset and Stance Classifier

Our dataset consists of comments on YouTube videos. We construct three datasets: one is an annotated dataset to train a stance classifier, and two unlabeled datasets to investigate ***RQ1*** and ***RQ2***. ***RQ1*** looks into the immediate public reaction to Djokovic's deportation following his inability to fulfill Australia's entry requirements for non-citizens. We construct a dataset, $\mathcal{D}_{vaccine}$, to investigate public discourse around Djokvic's vaccine stance and Australia's vaccine mandate consisting of 243,256 comments posted on 2,469 YouTube relevant videos. All videos in $\mathcal{D}_{vaccine}$ are published between July 1, 2021 to July 1, 2022.

For ***RQ2***, we construct a second dataset called $\mathcal{D}_{tennis}$. This dataset considers three YouTube channel sources: (1) official tennis channels; (2) news networks; and (3) opinion channels of vaccine skeptics. We consider the following tennis channels: TennisTV, ATP Tour, Wimbledon, US Open, Australian Open, and Roland Garros. For news channels, we consider the most popular cable news networks from these three countries as per the Statista report: Fox News (US); BBC News (UK); and Sky News Australia (Australia). We discard a video if any of the terms `novak, djokovic, djoker, nole` are not present in its title, description, or tags. $\mathcal{D}_{tennis}$ consists of 58,711 comments posted on 701 videos published on or after July 2, 2022.

We construct a dataset for stance classification ($\mathcal{D}_{stance}$) using active learning (Settles 2009) with the following steps:

Construct an initial seed set by randomly sampling from $\mathcal{D}_{vaccine}$. (748 positives, 376 negatives, and 876 neutral instances).

Conduct certainty sampling (Sindhwani, Melville, and Lawrence 2009) on $\mathcal{D}_{vaccine}$ (16 positives, 222 negatives, and 12 neutral instances).

Finally, conduct certainty sampling on $\mathcal{D}_{tennis}$ to tackle label distribution shift (98 positive, 42 negative, and 260 neutral instances).

Supportive	Not Supportive	Neutral
Novak Djokovic has been fearlessly brave showing the world that having rights to bodies is an individuals choice.	Do the major sponsors of the tournament really want their trophy held aloft by an anti-vaxxer? This in front of a fully vaccinated audience and televised in a country where 95% of the adults are vaccinated.	Everyone can agree that both Novak and the Border Force could and should have been more transparent.

Table 1: Examples YouTube comments supportive, not supportive, and neutral to Novak Djokovic's vaccine position.

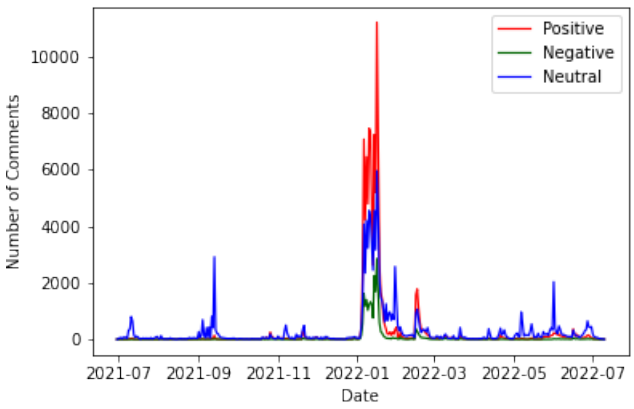

Figure 1: Positive, negative, and neutral comments to Djokovic's vaccine status from July 1, 2021 to July 1, 2022.

We conduct annotation using third objective instance (Gao and Huang 2017) with two human annotators and `GPT4`[2]. We train a stance classifier using `BERT` (Devlin et al. 2019), and obtain a Macro F1 score of 78.0 ± 0.4.

Results and Analyses

RQ1**:** *What was the distribution of positive and negative stances toward Djokovic's vaccine position in social media discourse following his deportation from Australia?*
As shown in Figure 1, our findings indicate that support for Djokovic's vaccine stance outweighed opposition to it. It is understandable that Djokovic being an iconic tennis player, and on the cusp of breaking the tennis record for most grand slams won by a male tennis player, public frustration was high that he couldn't play. In addition, the Australian government's messy handling of his immigration possibly garnered him additional support.
RQ2**:** *Is there lingering support for Djokovic's vaccine hesitancy as he keeps breaking tennis world records?*
Figure 2 summarizes the support across different sources. We observe that Tennis channels, which have mostly tennis fans as primary audience, have largely moved on with the vaccine controversy. While there still exists a small fraction of support for Djokovic's vaccine status, discussions are largely neutral around this issue. However, we note that this is not the case when we consider news networks and opinion pieces posted by vaccine skeptics. We notice that still a large fraction of comments laud Djokovic's stance of not taking the vaccine indicating that it is his freedom of choice.

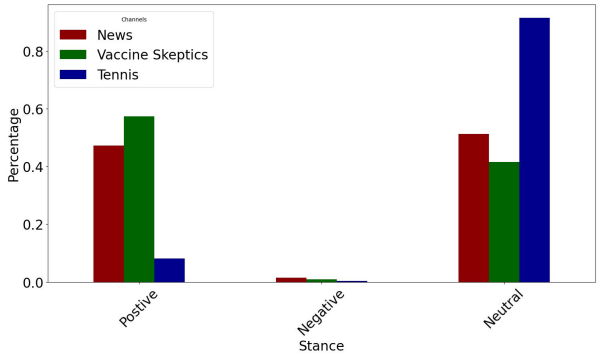

Figure 2: Lingering discussion regarding Novak Djokovic's vaccine stance from news channels, tennis channels, and vaccine skeptic videos.

[2]https://openai.com/gpt-4

References

Devlin, J.; Chang, M.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of NAACL-HLT*, 4171–4186. Association for Computational Linguistics.

Gao, L.; and Huang, R. 2017. Detecting Online Hate Speech Using Context Aware Models. In Mitkov, R.; and Angelova, G., eds., *RANLP 2017*, 260–266. INCOMA Ltd.

Gautam, A.; Mathur, P.; Gosangi, R.; Mahata, D.; Sawhney, R.; and Shah, R. R. 2020. # metooma: Multi-aspect annotations of tweets related to the metoo movement. In *Proceedings of ICWSM*, volume 14, 209–216.

Palakodety, S.; KhudaBukhsh, A. R.; and Carbonell, J. G. 2020. Hope Speech Detection: A Computational Analysis of the Voice of Peace. In *ECAI 2020 - 24th European Conference on Artificial Intelligence*, 1881–1889.

Settles, B. 2009. *Active learning literature survey*. University of Wisconsin-Madison Department of Computer Sciences.

Sindhwani, V.; Melville, P.; and Lawrence, R. D. 2009. Uncertainty sampling and transductive experimental design for active dual supervision. In *ICML*, 953–960.