Assessing COVID risk based on subjective experience

JINXU GUO* and BO QIN*, Dartmouth College, USA

P2, Introduction

CCS Concepts: • **Applied computing** → **Life and medical sciences**; • **Computing methodologies** → **Machine learning algorithms**.

Additional Key Words and Phrases: Personal Traits, Beliefs, COVID prediction

ACM Reference Format:
Jinxu Guo and Bo Qin. 2022. Assessing COVID risk based on subjective experience. In . ACM, New York, NY, USA, 3 pages.

1 INTRODUCTION

In the past few years, we experienced an unprecedented global COVID-19 pandemic. According to a recent report by the World Health Organization (WHO), the spread of COVID-19 not only led to a dramatic loss of human life but also trillions of economic losses worldwide. Tens of millions of people are at risk of falling into extreme poverty and 132 million increase in the number of undernourished people is expected due to the persistent impact of the pandemic [3]. Ironically, even though COVID-19 transmission mechanism was well understood at the height of the pandemic and the government implemented scientific public health measures to curb the spreading of the virus, the incidence rate remained high across continents [11]. Other than the objective reasons (e.g.,immunosuppressive morbidity) , we believe that subjective experience may also impact people's probability of being infected, such as personal traits (introvert vs extrovert), beliefs in conspiracy theories, beliefs in science etc. It is possible that these subjective experiences motivated or discouraged compliance and safe behaviors, affecting COVID incidence accordingly.

Past literature has looked at some aspects of subjective experience on compliances and development of safe behaviors. Zajenkowski and colleagues discovered that the perception of situations and individual personalities both contribute to different rates of compliance [10]. According to Jay, collective behavior variation and endorsing non-pharmaceutical intervention is of vital significance in reducing virus transmission during a pandemic [8]. Therefore, recognizing the potential driving factors influencing compliance would be extremely conducive to health organizations and political leaders in designing and deploying more effective measures to mitigate the pandemic.

Carvalho found and confirmed that conscientiousness leads to handwashing and social distancing during the pandemic [1]. That research also manifests that extraversion would affect individual engagement regarding quarantine policies. Yildiz and colleagues discovered that agreeableness could be indicative of emotion-focused coping styles [9]. Prentice also found that emotional intelligence is highly relevant to coping strategy [7]. Several studies looked at beliefs in conspiracy theories and beliefs in government officials showed a direct positive relationship between trust

*Both authors contributed equally to this research.

Manuscript submitted to ACM

and compliance [6]. Additionally, beliefs in government implemented public health measures was also shown to be positively related with better preventive behaviors against COVID [4].

Past literature provided vast evidence that the subjective experience can lead to better compliance and safe behaviors. However, the direct relationship between the subjective experience and COVID incidence has not been investigated. Additionally, the data we used was from an international project where a wide range of questions were asked to evaluate the subjective experience of people from different countries [8]. Therefore, instead of looking at one or two traits as in past studies, our project adds value by comparing a wide range of personal traits and beliefs along with increased international generalizability.

Our primary goal is to investigate which personal attribute or belief is most predictive of COVID incidence based on the data available to us. Furthermore,leveraging the predictive power of machine learning algorithms, we want to build a predictive model to calculate the probability of getting COVID based on subjective experience. For identifying the most important predictors, we will leverage a combination of parametric and non parametric models including multilevel logistic regression and XGBoost models which can determine feature importance by coefficients or Shapley values. For building the predictive model, we will try different architectures of neural networks and present the best performing model. The key contribution of our work lies in the interpretation of the important predictors and the performance of the predictive model.

2 RELATED WORKS

Jay's team used multi-level models to conduct analysis based on the worldwide collected data discovering the contribution of collective behaviors variations toward virus transmission reduction. For example, they have found that the National Identity (NI) was associated with reduced mobility (with $r = -0.40$, $p = 0.008$)[8]. Reduced mobility is very likely to constrain the spreading of COVID virus. Zajenkowski's research group found out that both individual personalities and perceptions of current situations contribute to differences of compliance by replicating and extending previous study which links the Dark Triad Traits and Big Five Traits to patterns of perceptions of situations under the Covid-19 Background. [10]. Lucas's team found that higher scores for extroversion were in accordance with lower means of social distancing and higher scores for conscientiousness leads to higher means of social distancing and hand washing based on sample consisting of 715 adults who answered the Big Five Inventory 2 Short (BFI-2-S) and information from the Dimensional Clinical Personality Inventory 2 (IDCP-2)[1]. Another research group asked their participants to complete the Hospital Anxiety and Depression Scale, the Big Five Personality Inventory and Coping with Problems Experienced Inventory, discovering that emotion-focused and problem-focused coping style scores were positively associated with openness scores, while emotion-focused coping style scores were also positively associated with agreeableness scores[9]. The study from Prentice's team conducted confirmatory factor analysis discovering that emotional intelligence contributes to coping strategy [7].

The most traditional way of predicting infectious diseases outcome is through the susceptible-infected-resistant (SIR) model. The SIR model takes into account the number of people who are susceptible to the diseases, the number of people who are transmitting the diseases and the number of people who are recovering from the diseases simultaneously and construct an ordinary differential equation system to solve for transmission rate and predict number of transmission. In 2020, Ian Cooper applied the SIR model to the COVID infection data from seven countries (USA, China, India etc.) and discovered that the expected time for the pandemic to wane should be around 150 days [2]. Apparently, the estimate was not very accurate. SIR model was not able to take into account policy changes or new variants that can cause surges of infections. Additionally, it provided aggregate level of infection risk but provide minimal estimation for individual

risk. Past literature has also looked at the ability of using machine learning models to predict COVID incidence using various predictors. At the onset of COVID 19 pandemic, one study estimated county-level COVID-19 occurrences using XGBoost algorithm [3]. The model considered demographics and morbidity information as predictors for county-level COVID incidence and achieved a sensitivity over 71% and specificity over 94%. The high specificity and low sensitivity may due to the lack of incidence at the beginning of the pandemic. Another study conducted later in the year used and compared several machine learning models (logistic regression, decision trees, support vector machine etc.) to predict COVID incidence. The predictors similarly included demographics and clinical features. They discovered that the best performing model is decision tree model with a highest accuracy of 94.99% [5].

Previous work is abundant on the impact of subjective experience on COVID incidence and COVID incidence prediction model. However, no work has looked at the impact of several subjective experiences and their contribution to COVID incidence. Additionally, no prediction model was built using these subjective experiences. Thus, our study will close the gap by first looking at the combined impact of several subjective experiences and then building a high-performing broadly-applicable machine learning model based on the results.

REFERENCES

[1] Lucas de F. Carvalho, Giselle Pianowski, and André P. Gonçalves. 2020. Personality differences and COVID-19: are extroversion and conscientiousness personality traits associated with engagement with containment measures? *Trends in Psychiatry and Psychotherapy* 42 (April 2020), 179–184. https://doi.org/10.1590/2237-6089-2020-0029 Publisher: Associação de Psiquiatria do Rio Grande do Sul.

[2] Ian Cooper, Argha Mondal, and Chris G. Antonopoulos. 2020. A SIR model assumption for the spread of COVID-19 in different communities. *Chaos, Solitons, and Fractals* 139 (Oct. 2020), 110057. https://doi.org/10.1016/j.chaos.2020.110057

[3] Suparna Ghosh-Jerath, Ridhima Kapoor, Ayushi Dhasmana, Archna Singh, Shauna Downs, and Selena Ahmed. 2022. Effect of COVID-19 Pandemic on Food Systems and Determinants of Resilience in Indigenous Communities of Jharkhand State, India: A Serial Cross-Sectional Study. *Frontiers in Sustainable Food Systems* 6 (2022). https://www.frontiersin.org/articles/10.3389/fsufs.2022.724321

[4] Sunhee Kim and Seoyong Kim. 2020. Analysis of the Impact of Health Beliefs and Resource Factors on Preventive Behaviors against the COVID-19 Pandemic. *International Journal of Environmental Research and Public Health* 17, 22 (Nov. 2020), 8666. https://doi.org/10.3390/ijerph17228666

[5] L. J. Muhammad, Ebrahem A. Algehyne, Sani Sharif Usman, Abdulkadir Ahmad, Chinmay Chakraborty, and I. A. Mohammed. 2020. Supervised Machine Learning Models for Prediction of COVID-19 Infection using Epidemiology Dataset. *SN Computer Science* 2, 1 (Nov. 2020), 11. https://doi.org/10.1007/s42979-020-00394-7

[6] Irena Pavela Banai, Benjamin Banai, and Igor Mikloušić. 2022. Beliefs in COVID-19 conspiracy theories, compliance with the preventive measures, and trust in government medical officials. *Current Psychology* 41, 10 (Oct. 2022), 7448–7458. https://doi.org/10.1007/s12144-021-01898-y

[7] Catherine Prentice, Susan Zeidan, and Xuequn Wang. 2020. Personality, trait EI and coping with COVID 19 measures. *International Journal of Disaster Risk Reduction* 51 (Dec. 2020), 101789. https://doi.org/10.1016/j.ijdrr.2020.101789

[8] Jay J. Van Bavel, Aleksandra Cichocka, Valerio Capraro, Hallgeir Sjåstad, John B. Nezlek, and Pavlović. 2022. National identity predicts public health support during a global pandemic. *Nature Communications* 13, 1 (Jan. 2022), 517. https://doi.org/10.1038/s41467-021-27668-9 Number: 1 Publisher: Nature Publishing Group.

[9] M. Yıldız, Ş. Şahin, S. Batmaz, E. Songur, and F. Kutlutürk. 2017. The relationship between depression, anxiety, personality traits and coping strategies of patients with euthyroid Hashimoto's Thyroiditis. *The European Journal of Psychiatry* 31, 3 (July 2017), 113–118. https://doi.org/10.1016/j.ejpsy.2017.06.001

[10] Marcin Zajenkowski, Peter K. Jonason, Maria Leniarska, and Zuzanna Kozakiewicz. 2020. Who complies with the restrictions to reduce the spread of COVID-19?: Personality and perceptions of the COVID-19 situation. *Personality and Individual Differences* 166 (Nov. 2020), 110199. https://doi.org/10.1016/j.paid.2020.110199

[11] Renyi Zhang, Yixin Li, Annie L. Zhang, Yuan Wang, and Mario J. Molina. 2020. Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proceedings of the National Academy of Sciences* 117, 26 (June 2020), 14857–14863. https://doi.org/10.1073/pnas.2009637117

Received 14 October 2022