



COVID-19 Forecasting in Japan: Fairness Analysis

Introduction

Google is committed to a core set of AI principles¹. These principles lay out how we think about using technology to solve important problems. They also identify important challenges that we have a responsibility to address clearly, thoughtfully, and affirmatively. In developing our COVID-19 Public Forecasts², we paid close attention to the disproportionate impact the disease has had globally and how this may impact our adherence to these principles, especially with respect to principle #2: “Avoid creating or reinforcing unfair bias.”

Some subgroups of the population have been affected more by COVID-19 than others. Issues such as access to healthcare, systemic bias and structural racism can lead to greater negative impact from COVID-19 (as well as many other conditions)³⁻⁸. While much has been written on this in a USA context, issues of bias⁹⁻¹⁰ and access to healthcare¹¹ have been described in Japan outside of the context of COVID-19, and a thorough analysis of how these issues might surface in our forecasting model remains important.

We have developed machine learning prediction algorithms to forecast COVID-19 cases in Japan at a prefecture level¹². Unlike our previous work in the US at a county level, forecasting at a prefecture level presents new challenges. Japanese prefectures are larger than US counties and many have fewer COVID-19 cases; the majority of prefectures have fewer than a thousand cases and fewer than ten deaths.

The overall model accuracy as a percentage of total cases remains consistent between prefectures. But this means that the mean absolute error (MAE) is larger for prefectures with a greater number of cases and the risk of over- or undercounting cases is higher (**Figure 1a**). This is important in the context of demographic groups, as the populations of certain demographics are not evenly distributed, but rather concentrated in prefectures that have higher case numbers. This was particularly notable in our US COVID-19 forecasting model. A detailed description of how MAE varied with different demographic makeups can be found in our [fairness white paper](#)². Normalising the MAE (dividing it by the death count) was found to reduce the variation in error between subgroups, indicating that the difference in absolute error across demographic groups can be largely explained by the difference in COVID-19 deaths and case counts (**Figure 1b**).

For the analysis of our Japan model, we focus on normalised mean absolute error (NMAE). Because the number of deaths are so low in Japan, we focus our analysis instead on confirmed cases. To simulate how the model would perform in practice, for the purpose of this analysis the model was trained up to September 14th 2020, and evaluated over the following 28 days. The below sections explore individual demographic groups in more detail. We hope that this approach and focus on fairness will contribute to a



broader discussion around how AI can not only support the response to COVID-19 but also support efforts to identify and address disproportionate impacts.

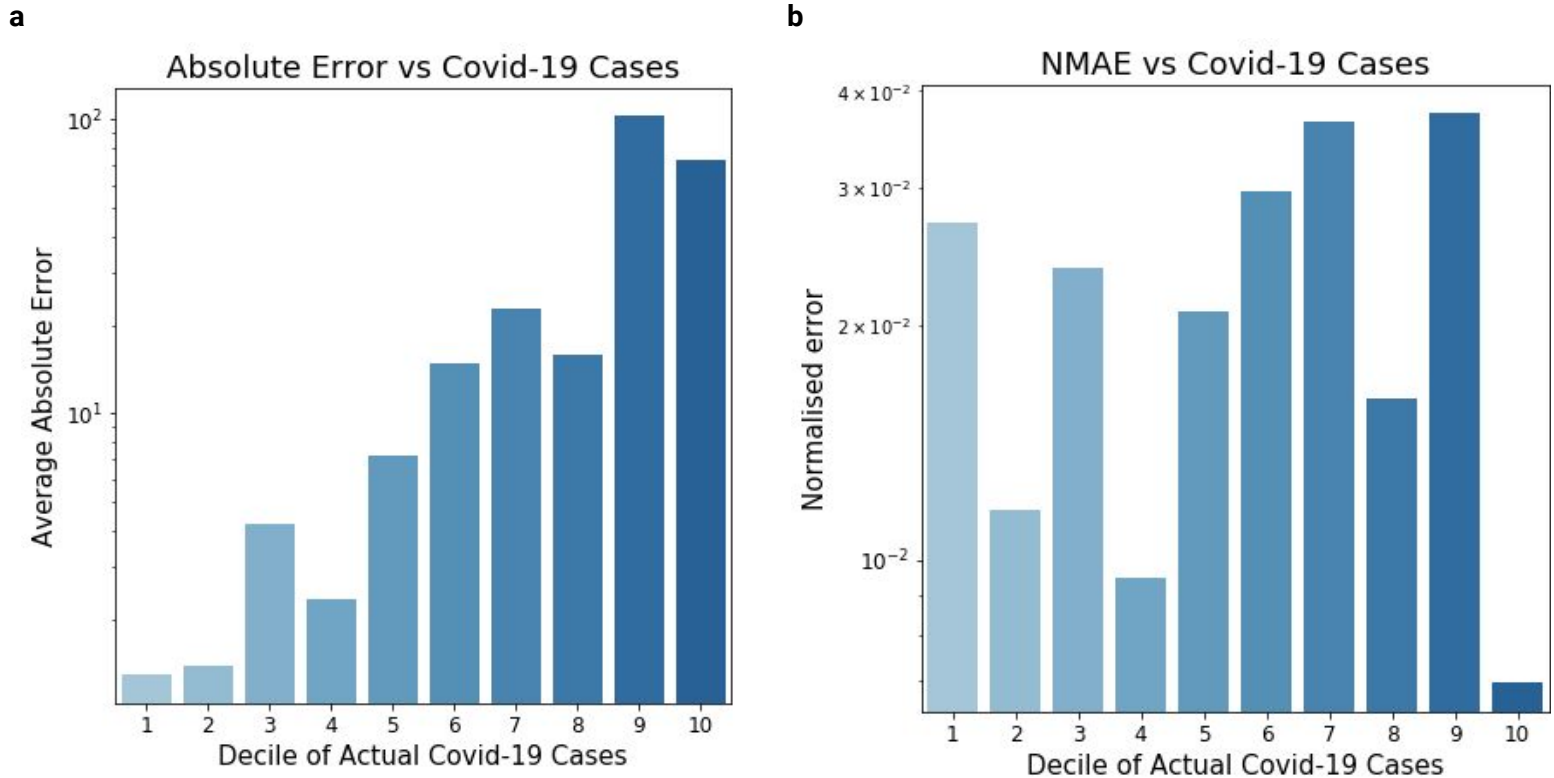


Figure 1: Absolute and normalized error and the total number of COVID-19 cases in Japan. (a) Absolute and (b) normalised error for prefectures are split into deciles, from the least number of cases to the most. In general, the model absolute error rises as the case numbers increase. After the error is normalized, a more uniform distribution of cases is observed.

Prefectural performance

Model NMAE was generally consistent across most prefectures (**Figure 2**). Higher error rates are mostly confined to prefectures with small numbers of cases or less dense populations. This can be expected, as machine learning models benefit from more training data and will perform less well where data is limited.

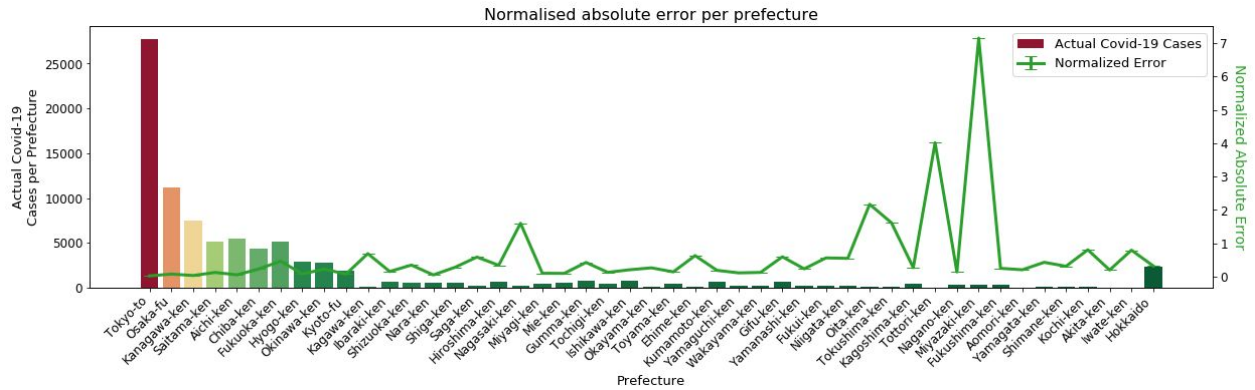


Figure 2: Model performance for each prefecture. Total number of cases (at the time of model assessment) and normalised error for all 47 Japanese prefectures. For each prefecture the green line indicates NMAE and the bar shows the total number of cases. Each bar is colored according to population density (densely populated regions in red, sparsely populated regions in green). For clarity, the prefectures are also ordered by population density, from the highest density on the left to the lowest on the right. Both the highest individual errors and case counts are seen in rural prefectures.

Subgroup analysis

In this section we present the model NMAE results for individual demographic groups: age, sex, ethnicity and median income. To better demonstrate trends across demographic groups, we bin the prefectures into four groups of an equal number of prefectures according to the percentage prevalence of that demographic group. We use self reported data from the [2015 Japan Population Census](#) for sex and ethnicity; for the later we report results over the three most populous groups (Japanese, Korean, Chinese). We use [MHLW statistics](#) for age (following the existing grouping of 0-14 years, 15-64 years and >65 years) and Japanese [Cabinet Office statistics](#) for median income.

Figures 3, 4 and 5 shows the total number of cases in each quartile, along with the respective model absolute and normalised error and their credible intervals. A clear correlation is observed with case count. After normalization this correlation is less clear, and in all groups there is substantial overlap in the credible intervals between prefectures with greater or lesser populations of different demographic subgroups.

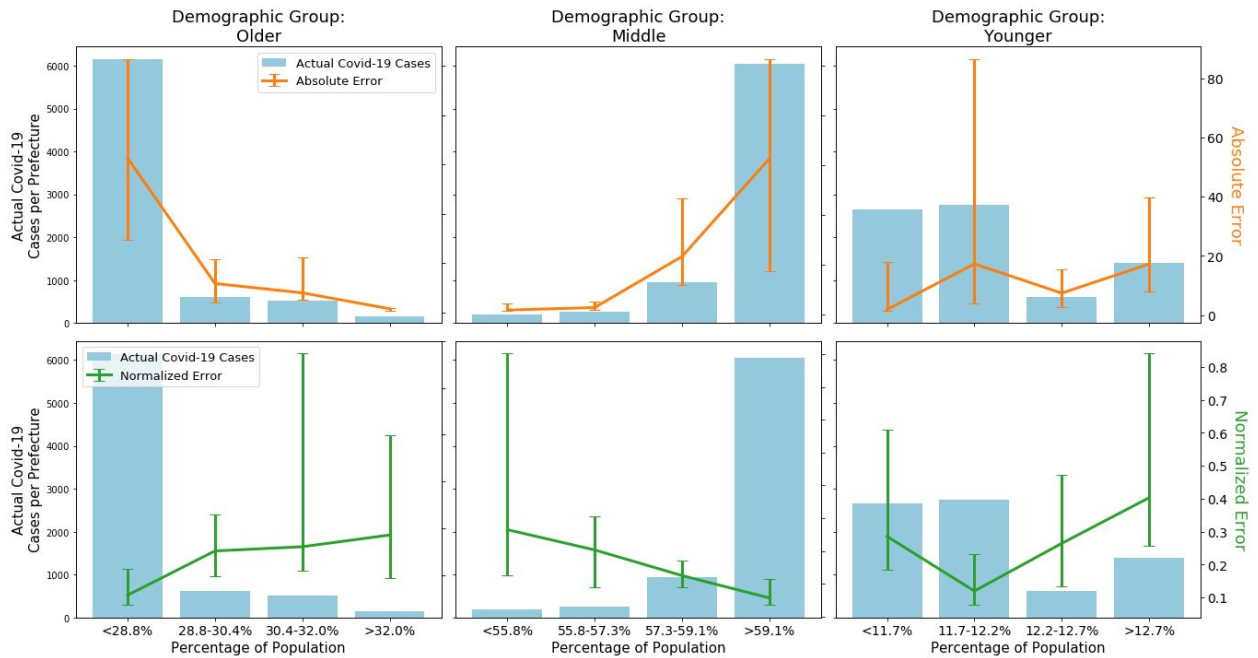


Figure 3: Normalised error across different age groups in Japan. Prefectures are split into quartiles by the percentage of the population that fall into younger (<14 years old), middle (15-64 years old) and older (>65) age groups. For each quartile, the total number of COVID-19 cases are shown (blue bars), along with the absolute error (orange line) a normalised error (green). The orange and green bars indicate the 25% and 75% credible intervals for absolute and normalised error respectively.

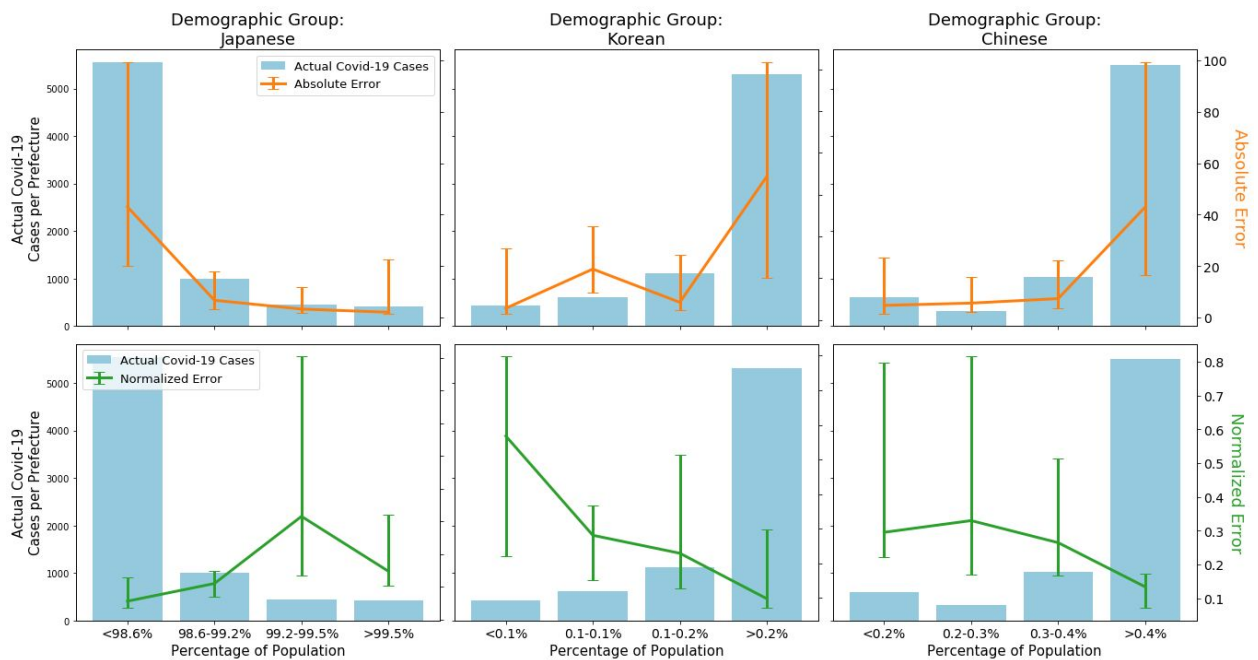




Figure 4: Normalised error across different ethnicity groups in Japan. Prefectures are split into quartiles by the percentage of the population that fall into the three most populous ethnicity groups according to the 2015 Census (Japanese, Korean and Chinese). For each quartile, the total number of COVID-19 cases are shown (blue bars), along with the absolute error (orange line) a normalised error (green). The orange and green bars indicate the 25% and 75% credible intervals for absolute and normalised error respectively.

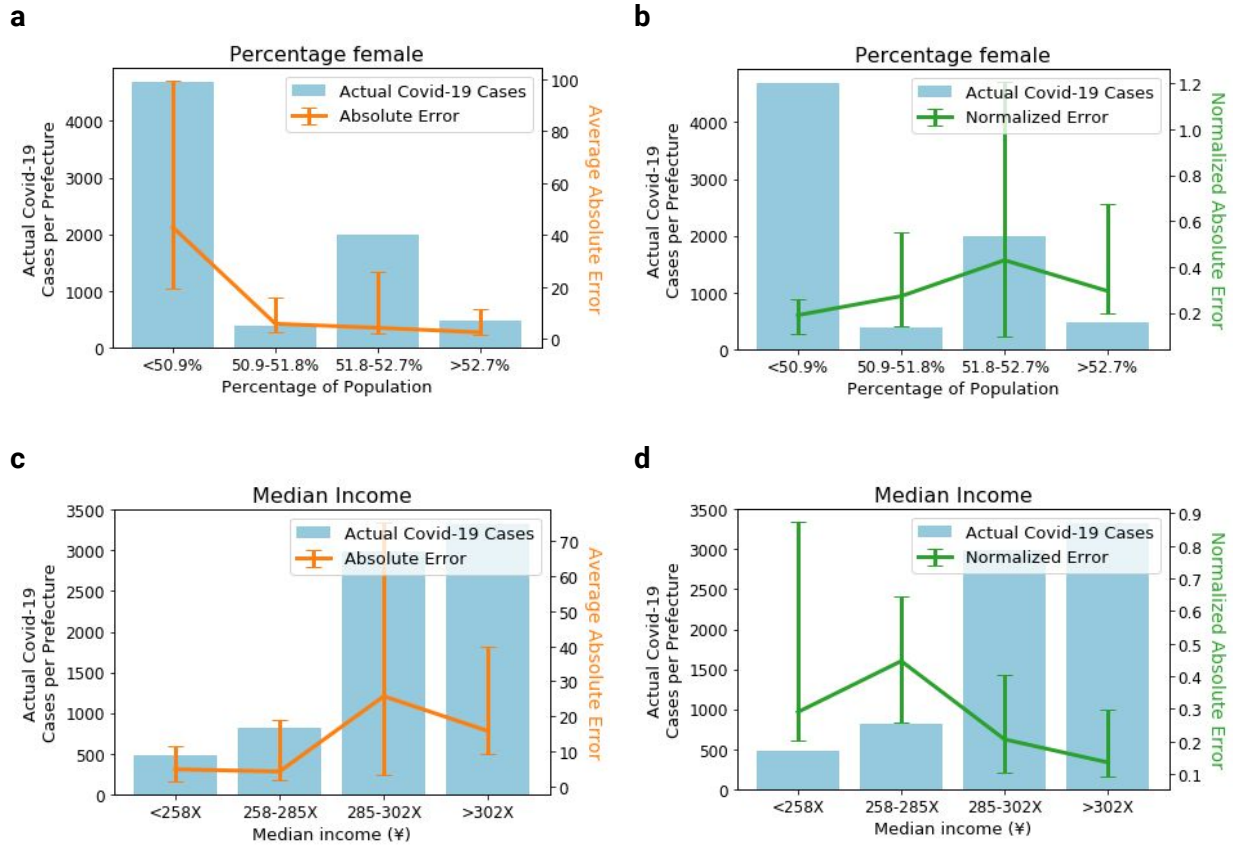


Figure 5: Normalised error across sex and income groups in Japan. Prefectures are split into quartiles by (a, b) the percentage of the population that are female, and (c, d) the medium income of the population. For each quartile, the total number of COVID-19 cases are shown (blue bars), along with the absolute error (orange line) a normalised error (green). The orange and green bars indicate the 25% and 75% credible intervals for absolute and normalised error respectively.

In addition to the demographic subgroups above, we also look at the effect of population density. A strong correlation between MAE and case count is again seen, which flattens out after normalization (Figure 6). Although population density is closely related to case count in Japan, in prefectures such as Hokkaido with lower population density but higher case counts the model NMAE remains relatively low (Figure 1).

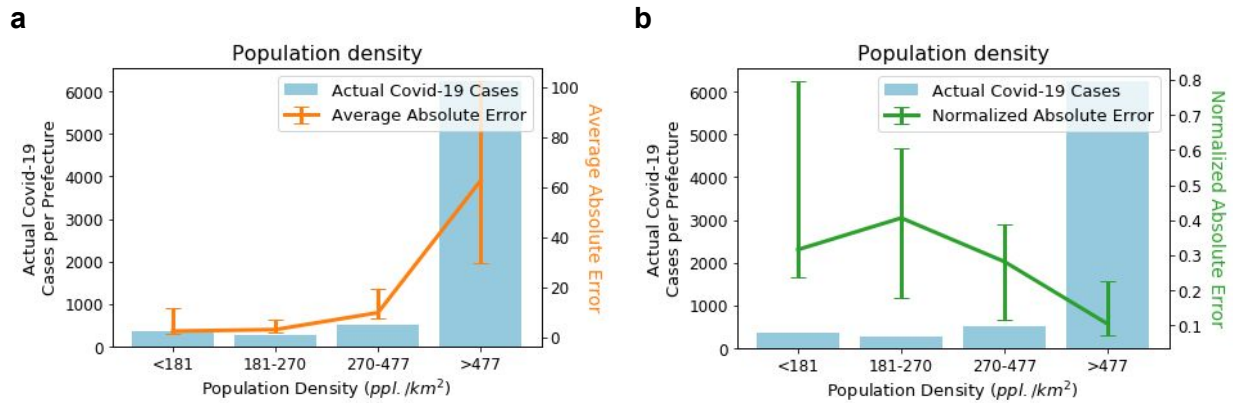


Figure 6: Normalised error across population density quartiles in Japan. Prefectures are split into quartiles by the average population density. For each quartile, the total number of COVID-19 cases are shown (blue bars), along with the absolute error (orange line) and normalised error (green). The orange and green bars indicate the 25% and 75% credible intervals for absolute and normalised error respectively.

Discussion

Our model optimizes for accuracy across all Japanese prefectures to provide the best overall forecast for most communities. After normalisation by case counts, the model error was consistent across all groups studied. Unlike our US COVID-19 forecasting work, model NMAE was lower or similar in areas with higher proportions of ethnic minority groups, despite previous reports of potential biases¹¹.

Performance appeared to be correlated with population density, though the correlation was markedly reduced by normalization. There may be several reasons for this. Firstly, these areas have the smallest number of cases, meaning there are fewer positive cases for a model to learn from making prediction more difficult. Secondly, the spread of COVID-19 will likely be slower and less predictable in sparsely populated areas than urban areas, which compounds the issue of lower data for these regions. Finally, while we didn't see evidence for it in this analysis, access to specialist care may be more difficult in rural areas that lead to health inequalities that may be represented in the data itself.

Our work has several limitations. As mentioned above, the numbers of cases were very small in many prefectures, presenting challenges both to model development and to drawing reliable conclusions in an analysis. In addition, due to the limited available duration for model training, we only report model performance at one point in time. As the factors influencing spread of disease may vary considerably over time, the results for a single time point may be less generalisable. Future work should be done to evaluate the model prospectively once launched to address potential limitations around training data and generalisability. Finally, it is important to note that for Non-Japanese ethnicities the proportions of individuals are very small (<1% of the population) in the majority of prefectures. This makes comparison difficult, particularly as no data was available on the number of cases in these communities.

We remain committed to ensuring that our AI models are both valuable to the broader public and fair to vulnerable populations. We hope our model proves to be a valuable resource to the larger efforts to



respond to the global COVID-19 pandemic. Further work should be undertaken to understand the impact of the model when deployed in practice.

References

1. Google (2018). 'AI at Google: our principles.' blog.google. Jun 7, 2018. Available at: <https://www.blog.google/technology/ai/ai-principles/> (Accessed: 2020/11/03).
2. Google Cloud (2020). 'COVID-19 Public Forecasts: Fairness Analysis'. Google Cloud, Aug 3, 2020. Available at: <https://storage.googleapis.com/covid-external/COVID-19ForecastFairnessAnalysis.pdf> (Accessed: 2020/11/03).
3. Laurencin, Cato T., and Aneesah McClinton. "The COVID-19 pandemic: a call to action to identify and address racial and ethnic disparities." *Journal of racial and ethnic health disparities* (2020): 1-5.
4. Tai, Don Bambino Geno, et al. "The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States." *Clinical Infectious Diseases* (2020).
5. Aldridge, Robert W., et al. "Black, Asian and Minority Ethnic groups in England are at increased risk of death from COVID-19: indirect standardisation of NHS mortality data." *Wellcome Open Research* 5.88 (2020): 88
6. Baqui, Pedro, et al. "Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study." *The Lancet Global Health* 8.8 (2020): e1018-e1026.
7. Steyn, Nicholas, et al. "Estimated inequities in COVID-19 infection fatality rates by ethnicity for Aotearoa New Zealand." *medRxiv* (2020).
8. Labgold, Katie, et al. "Widening the gap: greater racial and ethnic disparities in COVID-19 burden after accounting for missing race/ethnicity data." *Medrxiv* (2020).
9. UN (United Nations) (2006). Report of the Special Rapporteur on Contemporary Forms of Racism, Racial Discrimination, Xenophobia and Related Intolerance, Doudou Diène : addendum. E/CN.4/2006/16/Add.2. Available at <https://digitallibrary.un.org/record/566139> (Accessed: 2020/11/03).
10. ERD Net (2018). Joint Civil Society Report on Racial Discrimination in Japan to the Committee on the Elimination of Racial Discrimination (CERD) for its Consideration of the Tenth and Eleventh Combined Periodic Report of Japan. CERD/C/JPN/10-11. Available at https://tbinternet.ohchr.org/Treaties/CERD/Shared%20Documents/JPN/INT_CERD_NGO_JPN_31918_E.pdf (Accessed: 2020/11/03).
11. Yasukawa, Kosuke, et al. "Health-care disparities for foreign residents in Japan." *The Lancet* 393.10174 (2019): 873-874.
12. Arik, Sercan O., et al. "Interpretable Sequence Learning for COVID-19 Forecasting." *arXiv preprint arXiv:2008.00646* (2020).

Acknowledgements

Special thanks to those on the Google Cloud, Google Japan and AI Fairness teams who worked on this project, including Andrew Max, Ben Hutchinson, Emilio Garcia, Fergal Daly, Hiroki Kayama, Ivor Horn, Joe Ledsam, Joel Shor, Jinsung Yoon, Junichi Kawai, Kaho Kobayashi, Karen Ouk, Kris Pependorf, Madeleine Elish, Mike Dusenberry, Nanako Yamaguchi, Natalie Wei, Nate Yoder, Peter Fitzgerald, Raj Sinha, Ryu Hirayama, Sakura Tominaga, Sercan Arik, Takahide Kato, Timnit Gebru, Tomas Pfister, Tomohiko Kikuchi, Vik Menon.