What if ? Global COVID-19 policy simulator. Web interface for public and policymakers to explore the health and societal impacts of COVID-19 policies.

Pinto Andrea

Bachelor Project under the supervision
of Mary-Anne Hartley and Prof. Martin Jaggi

EPFL
January 2021

ABSTRACT. Governments across the world have attempted to mitigate the spread and impact of the COVID-19 pandemic with various degrees of success and little evidence on which to base their decisions. This project builds an interactive web-based platform to host the ***What if...?*** Global COVID policy simulator, that allows users to simulate the effect of epidemic mitigation policies in various contexts and understand their costs and impacts.

***Index Terms*—COVID-19, Machine Learning**

I. Background

The COVID-19 pandemic has made irreparable damage to the health, social structure and economy of every country in the world. Governments implemented an array of mitigation strategies to optimize a trade-off between lives and livelihoods, however public compliance varied greatly across global inequalities and cultural norms. While a substantial amount of research has been published on how best to reduce the spread of SARS-CoV-2 or mitigate its impacts, the papers tended to be too specific in geographic, social or temporal scope and could not summarize the global heterogeneity of policy effectiveness [1], [2], [3]. Our group previously studied the global effectiveness of non-pharmaceutical epidemic mitigation policies during COVID-19 crisis [4], developing a machine learning model to infer the reproduction number (R-Value or R_0) and thus predict how it would have evolved if governments had taken other measures. It is a hybrid neural network combining an LSTM (Long-Short Term Memory) layer with an MLP (Multi-Layer Perceptron) that estimates the R_0 of the virus in 93 countries according to many non-epidemiological factors (such as GDP per capita, demographics, weather data..). This prior work is now being expanded to include further more parameters (in particular, demographic and economic) as well as use reinforcement learning to predict the optimal set of policies under certain conditions. This could then be applied to predict the future evolution of the virus and also guide the population, doctors (epidemiologists) and governments in understanding how to mitigate its spread. One could also imagine extending the model to other pathologies.

The goal of this work is to develop an intuitive web-based platform to host this model, allowing users to explore the predictions and better interpret the efficiency of policies. This naturally implies the development of the frontend of the platform, to visualize the inputs (parameters) and outputs of the model (graphs). But also the development of a solid backend allowing to call upon the useful functions of the model and thus provide the predictions to the frontend. In this context, a global API must be developed to make the two main entities communicate properly. But also another one in the backend allowing simple management of the heavy machine learning infrastructure and the large datasets. We also sought an open-source project that would allow any interested and motivated individual to contribute to this project in order to make it grow. Be it on the work of the platform or on the design of the machine learning models.

II. Aims and objectives

The aim of the project is to improve public awareness of evidence-based policy effectiveness through an interactive web tool. We would like to offer the public and policymakers the opportunity to explore the health and societal impacts of COVID-19 policies. There are many web-based tools that aim to predict and simulate the impacts of measures taken by the governments in the context of the crisis [5], [6]. However, none of them is yet optimal or offers intuitive way to visualise the impact of various policy (whether amateur or professional).

Below is an exhaustive list of the objectives that are set within the framework of the current work.

A. Intuitive frontend

The primary objective of the project is to develop a platform that is intuitive. It is mainly a page containing the *simulator*. This page must make the user understand that 1) they can choose a country, a time period and policies, and 2) that these choices have a direct effect on a graph which allows them to visualize the predicted R_0 as well as the ground truth. It is also important to create a *homepage* briefly describing the project and the use of the platform. And finally a *team page*, highlighting all the members and their involvement in the project (which will later list collaborators and incentivise contributions).

B. Connectivity with backend

When the frontend is developed, the simulator must be able to communicate with the server on which the backend is located. It is the backend that calculates the predictions and sends these new values according to the parameters that the user submits. This implies the creation of an API that allows to communicate cleanly between the backend and the frontend, and without generating too much traffic (for low latency) during data transit. The latency should be as low as possible to allow the user to see a live impact of a change in parameters on the predicted R_0.

C. Opensourcing

Because this project is aimed to be a public good, and specifically seeks to increase intrepretabilty and evidence, it is necessary to make all the content available to everyone. In doing so, we can also hope that interested and motivated people will bring their stone to the edifice and thus make this beautiful project grow. This involves making all the code open source, writing well documented code, filing a license on the main repository, and defining a procedure for all potential contributions.

D. Deployment

The initiative is good, and so are the directions. But to have something useful, you also have to think about deploying the work and putting it into production. Indeed, as we seek to make this quickly available to the public, this objective involves finding a server to host all the static files of the frontend and the machine learning models. The server must also have sufficient computing power to run the models. And finally set up a pipeline to build and deploy all the modifications that are made to the main repository.

Finally, below is the final mock-up of the platform we wanted and that we realized in August 2020. The goal is to get as close as possible to this mock-up. The idea is to have the global parameters (country and time-frame) on the left side of the graph that render the R_0 as a function of time $f(t) \rightarrow R_0|_t$. The graph is himself above a component that we soberly call the *policy box* in which each of the inner small boxes corresponds to the policies visible on the left (in line) fragmented on the time window. These boxes are interactive and allow users to change the corresponding policy's stringency over the time period and thus see direct consequences on the graph.

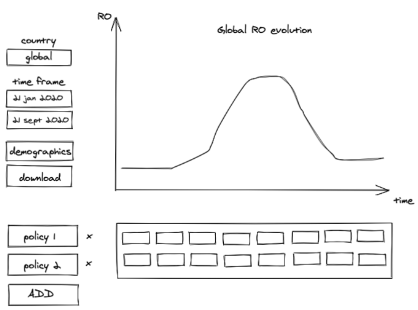

Figure 1. Global COVID-19 policy simulator mock-up

III. Methods

Detailed methods and information of the development environment, server access, deployment and other technicalities can be found in the README of the main GitHub repository **covid-what-if** [7]. Two other READMEs hosted in the frontend and backend folders, explicitly give the architectures and code structures of the frontend and backend respectively. This can help to understand the code and allows to quickly find information and different modules through the project.

The main resources used are listed below:

– ReactJS. Main framework for the frontend.
– Python. Main language for the backend.
– Flask. Library for the communication API.
– Tailwind. For CSS and frontend design.
– D3JS. For graph rendering on the frontend.
– NGINX. For deployment on the server.

IV. Results

After a long work to think about the different technologies as well as the setting up of the whole infrastructure. The final result is currently available online at **covid-what-if.io** [8]. Note that the domain name may changes soon or later as we might migrate to EPFL servers and change the domain extension.

It offers the user the choice of the country first as well as the choice of the time window (optional). When the country is entered by the user, the data (on the time window) is imported and displayed on the graph. The ground truth from historical data is rendered in black and the model prediction in blue. The graph also shows the orange line $R_0 = 1$ which is the threshold at which the growth and evolution of the virus becomes exponential.

The prediction is always inferred by taking as input the distribution of the policies and their stringency. At the time of first data import, the prediction displayed is therefore the one calculated with the policy mix that have actually been taken historically. From there, the user is free to play with the policies: he can add or remove them and can modify their stringency on time samples. The darker a small box is, the more strongly the policy is implemented on the time interval. On the contrary, it is gray if the policy is not implemented.

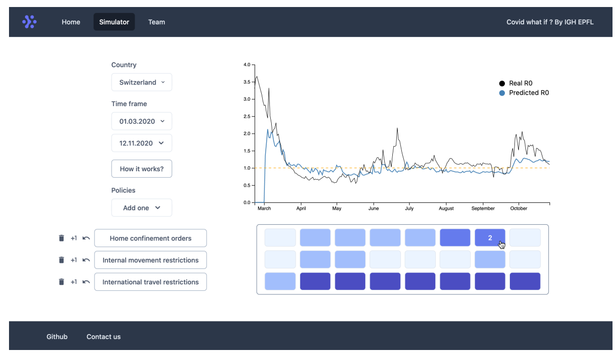

Figure 2. Global COVID-19 policy simulator

The simulator looks very much like the mock-up we originally designed and we are very happy with it.

V. Future Work

To ensure the success of the project, each of the objectives listed previously should be expanded and continuously updated. Several point are enhanced.

Performance and design. At the platform level, it is always possible to improve the UX/UI to make it more intuitive and visually attractive. However, performance and robustness should be prioritized over these cosmetic concerns. It would be interesting to improve the website's performance serving less CSS to the user by using purgeCSS. This is related to the use of TailwindJS and should be prioritized.

Deployment. At this stage, the project is running on my private AWS server. It would be useful to quickly contact the EPFL to have a dedicated server for the project. Making the transfer is quite easy, it is more of an administrative process. This would also allow us to have a larger computing power and thus a larger bandwidth. It would also be useful to set up the scripts on GitHub to deploy automatically whenever there is a commit (this is a bonus, specifically useful when there is a lot of traffic and many commits moving forward).

Outsourcing. The repository and READMEs are set up so that people can contribute quite easily. Nevertheless, it would be a good idea to extend this to supervise the running of tasks and commits from external collaborators. It would also be more efficient and productive to set priorities on a shared resource. That way, contributors would be updated on modifications.

New features. Whether directly on the platform (visible) or on the models and their capabilities, many product improvements can be imagined. During the work, we had some ideas, which are also detailed :

A. Peer review

It is useful to bring a critical eye on the project, and this can be done by implementing peer review. To begin with, it could simply be a page allowing users of the website to submit comments, and to file a constructive and unmoderated opinion on the project. On the one hand, this would make it possible to highlight positive opinions and thus market the tool, but on the other hand, it would also make it possible to handle potential errors or improvements desired by users directly.

B. Updatable backend

ML models that infer R_0 must be trained to perform well. Ideally, they should always be up to date with the latest data and be freshly available. This involves developing an infrastructure using scripts (e.g. executed with crontab) directly on the server that would sequentially perform the following operations. Fetch the latest data from sites and reference sources. Update our respective databases. And then train the ML models for each country with the new data, if any, to be able to make them available on the same day or within 24 hours maximum following an update.

C. Policy mix implementation

Currently, models allow the inference of R_0 based on certain parameters and a mix of health policies. The initial goal was also to provide a functionality to display to the user the optimal policy mix according to the parameters he gave. It would be nice to develop this feature, involving reinforcement learning models, in order to make it available. It would also be very interesting to extend the capabilities of the model not only to infer historical data and compare them with reality, but also to make projections and thus predict future developments based on the same parameters and policy mix.

D. Adding demographics

To provide the user with demographic information about the country they are analyzing for better understanding and contextualization. Also to allow the user to possibly change some of these parameters and make a hypothetical population.

E. Create own country

In the current layout, a user can select a country and then play with parameters. However, it would be very interesting to either import data from another source to plot it and run the ML model on this given data (this could be an entire feature alone) or create your own country by giving all the required information e.g. population size, climate, policies in place, etc, and then run the ML model with this construction. This would be particularly useful for modeling countries that have little or no reliable data. It would then be up to professionals or users to make estimates and thus visualize predictions about their country (imaginary or real) based on all these characteristics. You could also imagine a dashboard where you could save all your constructions to be able to reuse them in the future.

F. Download reports

It would be interesting to add the possibility to download a snapshot of the visualized data and the current parameters. This would allow the users to create reports with a certain mix of policies to save a given setup or event and thus facilitate sharing and discussion.

VI. Conclusion

Initially, I thought that my work would be limited to developing the frontend of a platform and then slightly modifying the code of the different ML models. But in the end, the biggest job (which unexpectedly took me quite a bit of time) was 1) to design all the infrastructure that could be scalable as well as 2) the whole communication pipeline between the backend and frontend entities. It was necessary to build two APIs allowing to communicate between two completely different languages and to transmit highly optimized requests through the Internet.

Looking back, the primary objective of the project was to realize from scratch a platform that would allow users to play with the parameters of the models we developed and thus increase their awareness of the measures taken by various governments in the context of the COVID-19 crisis. I arrived at the end of the semester with a functional platform, allowing us to effectively use the models and thus visualize the changes. Nevertheless, there is still a lot of room for improvement. Whether it be on the new features mentioned in future work or simply maintenance. I am however fully confident in the ability of the platform and repository to grow and scale well. And I will be extremely happy to continue developing this project within the team as it was very instructive.

Acknowledgment

I would like to thank all the Intelligent Global Health Laboratory team and especially Mary-Anne Hartley for giving me the opportunity to work on such a project and guiding me all the way through. Thanks to Prof. Martin Jaggi and Dr. Prakhar Gupta for supervising the project. Also huge thanks to Thierry Bossy and Lucas Massemin who helped me and advised me all along the path. Final thanks to Lucas Meier, a friend who also supported during the project.

References

[1] V. Kompella, R. Capobianco, S. Jong, J. Browne, S. Fox, L. Meyers, P. Wurman, and P. Stone, "Reinforcement Learning for Optimization of COVID-19 Mitigation Policies," Sony AI, Sapienza University of Rome, The University of Texas at Austin, Tech. Rep., 2020. [Online]. Available: https://arxiv.org/pdf/2010.10560.pdf

[2] Q. Zhaozhi and A. Ahmed, “When and How to Lift the Lockdown? Global COVID-19 Scenario Analysis and Policy Assessment using Compartmental Gaussian Processes,” University of Cambridge, UCLA, Tech. Rep., 2020. [Online]. Available: https://proceedings.neurips.cc/paper/2020/file/79a3308b13cd31f096d8a4a34f96b66b-Paper.pdf
[3] M. Douglas, S. Katikireddi, M. Taulbut, M. McKee, and G. McCartney, “Mitigating the wider health effects of covid-19 pandemic response,” University of Edinburgh, Tech. Rep., 2020. [Online]. Available: https://www.bmj.com/content/bmj/369/bmj.m1557.full.pdf
[4] T. Bossy, “Adaptive Mitigation: Identification of the Dynamic Drivers of Effective Policy during the COVID-19 Pandemic,” EPFL, Tech. Rep., 2020. [Online]. Available: https://github.com/epfl-iglobalhealth/COVID-SocioLegalEpi/blob/master/report/thesis.pdf
[5] T. Ayer, J. Chhatwal, B. Linas, “Covid-19 simulator,” https://analytics-tools.shinyapps.io/covid19simulator01/.
[6] R. Neher, I. Aksamentov, N. Noll, J. Albert, R. Dyrdak, “Covid-19 scenarios,” https://covid19-scenarios.org.
[7] A. Pinto, “Covid-what-if,” https://github.com/andreakiro/covid-what-if.
[8] ——, “Covid-what-if,” https://covid-what-if.io.