

LEARNING TO ACT: NOVEL INTEGRATION OF ALGORITHMS AND MODELS FOR EPIDEMIC PREPAREDNESS

Sekou L. Remy & Oliver E. Bent

IBM Research Africa

Nairobi, Kenya

sekou@ke.ibm.com, oliver.bent2@ibm.com

ABSTRACT

In this work we present a framework which may transform research and praxis in epidemic planning. Introduced in the context of the ongoing COVID-19 pandemic, we provide a concrete demonstration of the way algorithms may learn from epidemiological models to scale their value for epidemic preparedness. Our contributions in this work are two fold: 1) a novel platform which makes it easy for decision making stakeholders to interact with epidemiological models and algorithms developed within the Machine learning community, and 2) the release of this work under the Apache-2.0 License. The objective of this paper is not to look closely at any particular models or algorithms, but instead to highlight how they can be coupled and shared to empower evidence-based decision making.

1 INTRODUCTION

This work provides a concrete example of a new paradigm to inform decision support processes in a public health context. Public health professionals are expected to engage in evidence-informed decision making to advise their practice (Petticrew et al., 2004; Yost et al., 2014). As indicated in Nutbeam (2001), evidence is required from a variety of sources including expert knowledge, existing domestic and international research, stakeholder consultation, and even assessment of existing policies. While it is rational to expect evidence to be used, challenges in evidence based decision-making are ever present. Today, very few decisions have complete information, the simplest decisions often have innumerable outcomes which are ultimately uncertain. For high impact decision making, such as epidemic planning, though of great importance, it is not possible to evaluate all possible options, or completely characterize their uncertainty. Many turn to modeling as a mechanism to generate insights about what could potentially happen, and consider model output as informative in the decision making process. It is well known that all models can be considered to be ‘wrong’ (Ioannidis et al., 2020), and in some cases have led to ill conceived actions being considered, but it has also been shown that when properly contextualized, even these incorrect models can be useful (Holmdahl & Buckee, 2020).

In this work we demonstrate an approach to harness epidemiological models to generate evidence which can inform decision making. We utilize Machine learning algorithms from a range of classes (from RL, to Optimization, to Planning). Our aim is not to focus on model validity as defined in Petticrew et al. (2004), nor the non-technical issues related to adoption of the insights generated from these models Liverani et al. (2013), nor even to adjudicate which model (or algorithm) is ‘best’. Our goal is to show the community how models can be connected with these algorithms in a flexible manner which will support integration of multiple classes and implementations of models. What we champion here will also provide a mechanism to realize a repeatable infrastructure to generate insights, not just for COVID-19, but for any decision making process which can be informed by multiple models and data sources. The resources we describe in this work have been released via a public GitHub source repository¹ which serves as both an installation point and templates for continued extension.

¹<https://github.com/IBM/ushiriki-policy-engine-library>

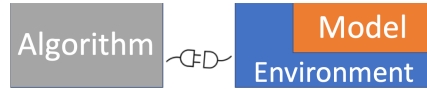


Figure 1: The key abstraction presented in this work wraps domain models in OpenAI Gym environments, exposing the STEP method as a connection between models and algorithms.

2 METHODOLOGY

In this project we utilize the framework presented in Wachira et al. (2020) to access models at scale with transparency and trust. This framework enables API mediated access to a suite of models. Containerized models are deployed on-demand, and configured at runtime to use the input parameters defined by the user. The framework marshals the use of the model and the associated data in a manner that gives both the users and owners of these assets confidence in the output which is generated. Moreover, the framework permits data collected by different algorithms, and possibly even at different times, to be shared. Every additional execution of an algorithm contributes to the pool of knowledge which exists about the models. Finally, the framework also supports the flexible integration of input into a format which can be consumed by the specific models which are available.

For the larger project in which this work is situated, we select models provided by modelers from different communities. In this manuscript we will only highlight the use of two types of models: compartmental models (Brauer, 2008; Edlund et al., 2011) and agent based models (Macal & North, 2009; Kerr et al., 2020) as they are both widely used by epidemiologists across a variety of domains. For these instances of COVID-19 models referenced in this work, the countermeasures (or interventions) are captured as sequences of parameter values. In the first framing, we permit the actions (e.g. physical distancing, mask wearing, vaccination) to directly change model parameters like transmission rate, recovery rate, and the death rate. In an alternate framing, we also link model parameters to the Stringency Index (Hale et al., 2020), a metric which captures the impact that a particular set of countermeasures has on disease progression. These framings represent distinct implementations of models specifically either for model predictions or assessing the impact of control policies.

Listing 1: Sample code instantiating an environment with access of the STEP method on line #4.

```

1 import gym
2 env = gym.make(name, userID=userID, baseuri=baseuri)
3 env.reset()
4 observation, reward, done, _ = env.step(env.action_space.sample())

```

The crux of the contribution of this work lies in encapsulating the epidemiological model into an OpenAI Gym environment. In this case, when the step method is provided with an action (see Listing 1), the model is able to run and provides information indicating what occurred (the observation), and a metric capturing the goodness of the action (the reward). Other metadata is provided as well, but it is out of the scope of this paper. Depending on the definition of the environment, the action can be a single intervention or a sequence of interventions.

In this project we have successfully harnessed algorithms from multiple domains of the Machine Learning spectrum. From Reinforcement Learning (Watkins & Dayan, 1992; Keerthi & Ravindran, 1994; Mnih et al., 2013) to Optimization (Srinivas et al., 2009; Contal et al., 2013), they have all been implemented utilizing the *step* method of the environment to access the reward function (and in some cases the observations of the model states) when presented with an action/intervention to evaluate. An example of these algorithms is seen in the pseudocode for Bayesian optimization presented in Algorithm 1. Due to space constraints, we omit discussion of the algorithm itself, but refer the curious reader to Contal et al. (2013) for further details.

3 RESULTS

What are the most cost effective ways to “flatten the curve”? If there is a fixed amount of money which can be allocated to interventions over the next three months, what sequence would be best? If you wish to simplify your policies which reduced sets of interventions are most impactful? To address any of these questions using model derived evidence, one must first calibrate the model to

Result: Policy $\pi_{greedy} = \operatorname{argmax}(\hat{v}_{\pi}(\theta^{gp}))$, where $\hat{v}_{\pi}(\theta^{gp}) \approx v_{\pi}$
 Initialise \mathcal{GP} priors, θ_0^{GP}
for $i = 1, 2, \dots, \text{end}$ **do**
 if $i \leq k$ **then**
 | $\text{select } \pi_i = \text{rand}(\pi)$
 else
 | $\text{select } \pi_i = \operatorname{argmax}(\alpha(\hat{v}_{\pi}(\theta_{i-1}^{gp})))$, where α is the acquisition function
 end
 $R_i = \text{step}(\pi_i)$
end
 Update \mathcal{GP} Posterior $\hat{v}(\theta_i^{gp})$; mean $m_i(\pi)$; co-variance function $k_i(\pi)$;

Algorithm 1: An Implementation of Bayesian Optimization using the STEP method.

data for the location in question, and then one can utilize an environment which provides the reward functions relevant to the particular question.

As indicated previously, the same algorithms can be applied with the multiple algorithm types and model types (through the use of the respective environments). Moreover, for each model, there were environments for calibration and optimization with various objective functions. A subset of the combinations is included herein for brevity. Figure 2 shows the results of calibrating the parameters three different models to the same data, utilizing Algorithm 1. In each case the algorithm was able to identify reasonable fits, even though the structures of the models were quite different.

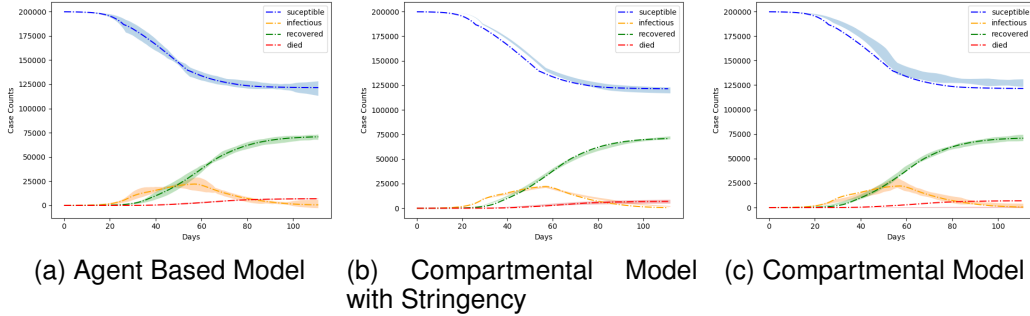


Figure 2: Calibration of three environments to the same case data (in this case using more data than is typically available). In each calibration, the results consider learning 20 sets of parameters.

The reward function in the case of calibration is the error between the data and the observed model output. The actions permit the algorithms to set the values of the calibration parameters. By changing the reward function to represent different properties, for example, the number of cases, and changing the actions to modify the transmission rate, another environment can be created permitting the same models to help consider the problem of flattening the curve. In this case, the algorithm’s role is to find the best set of interventions (the sequence of transmission rates) which minimize the number of individuals who contract the disease (see Figure 3).

In this case, the environment’s states are the values from each of the compartments (or the fraction of the population in a given state); the rewards produced are the cumulative incidence for the two week period which each step is assessed for; and the actions are integers in the range $[0, 99]$. The overall goal is thus to minimize the total number of cases by learning when to vary the stringency of the interventions implemented. In Figure 3, we show the state trajectories (incidence) for two sequences of actions. The first was a very stringent disease management response which was learned, while the other may be considered lenient. In both cases there are individuals in the population who contract the disease, however under the stringent policy, there are much less who do. Also, under this policy, the peak number who are infected at the same time is also lower. This might be considered as an example of the impact of “flattening the curve”. Using this type of approach, and by changing the reward function, it would be possible to raise the question of whether there are other intervention programs (sequences of actions) which could have flattened the curve with a lower societal cost.

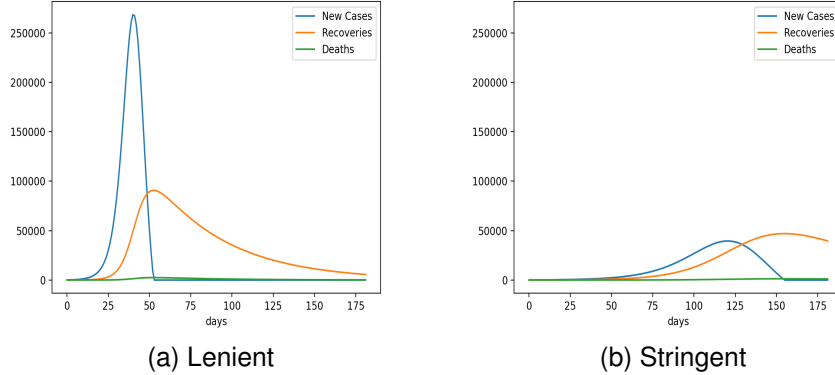


Figure 3: Two examples of deterministic policies of a single countermeasure for the simulated timeframe. Could a reactive policy have performed “better” according to a chosen reward function?

As alluded, by changing the reward function but keeping all the actions the same (which changes the response from the STEP function) we can also consider the impact of sequences of actions under a different rubric. In this case we quantify the impact of actions on the community. Executing interventions have a cost, but there’s also a cost if individuals in the populations get infected. Individuals can be infected even when the most stringent interventions are put into place. Health economists have characterized the notion of Years of effective Life Lost as one such measure to quantify the impact of an illness. In this work we implement what they have outlined and also include the cost of an intervention inferred from the newspaper sources.

4 SUMMARY

In proposing a new paradigm to inform decision support for epidemic planning, we have detailed extensions of epidemiological models to learning environments and the demonstration of learning algorithm performance on such environments. We posit that there is much to consider in the rich space of reward functions which arise from capturing the complexity of decisions in the real-world which are full of trade-offs and balancing competing needs. It is left to be seen whether the algorithms which will be most successful in this problem domain will be similar to those which are dominant in domains such as backgammon, chess, or even those which have physics based dynamics like the cart-pole or walking robots. COVID-19 has shown us that there is still need for research advancement in the process of decision-making being supported by models.

Our primary contribution in this work is the demonstration of how decision makers can access self-contained resources curated by the relevant experts (e.g. epidemiological modelers and ML experts). The secondary, and equally as valuable contribution is the presentation of the open sourced, OpenAI Gym environments typifying the encapsulation of epidemiological models in a manner which will be illustrative and enable stakeholders to play with the assets themselves, and subsequently to extend them to include models/data/algorithms which are relevant for their context.

In closing, we acknowledge the limitations of models but we still advocate for their use in principled and pragmatic ways. We specifically assert that whether considered in the traditional sense or as we’ve proposed in this work, any model usage must be in line with the intended manner defined by the model creators, as well as those defined by the users who will consume the output of the model. While we connect models to algorithms by extending the model abstractions to environments, we take care not to violate the underlying assumptions and definitions prescribed by their creators. In our future work we will perform user studies with the aim of characterizing the utility of these resources in practice, and to better understand how to amplify the decision making power of policy makers. Further, we will use in silico means to consider the impact of data integrity on the calibration and optimization results which can be generated using our approach.

REFERENCES

- Fred Brauer. Compartmental models in epidemiology. In *Mathematical epidemiology*, pp. 19–79. Springer, 2008.
- Emile Contal, David Buffoni, Alexandre Robicquet, and Nicolas Vayatis. Parallel Gaussian process optimization with upper confidence bound and pure exploration. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8188 LNAI(PART 1):225–240, 2013. ISSN 03029743. doi: 10.1007/978-3-642-40988-2_15.
- Stefan Edlund, James Kaufman, Justin Lessler, Judith Douglas, Michal Bromberg, Zalman Kaufman, Ravit Bassal, Gabriel Chodick, Rachel Marom, Varda Shalev, et al. Comparing three basic models for seasonal influenza. *Epidemics*, 3(3-4):135–142, 2011.
- Thomas Hale, Anna Petherick, Toby Phillips, and Samuel Webster. Variation in government responses to covid-19. *Blavatnik School of Government Working Paper*, 31, 2020.
- Inga Holmdahl and Caroline Buckee. Wrong but useful—what covid-19 epidemiologic models can and cannot tell us. *New England Journal of Medicine*, 2020.
- John PA Ioannidis, Sally Cripps, and Martin A Tanner. Forecasting for covid-19 has failed. *International journal of forecasting*, 2020.
- S Sathiya Keerthi and B Ravindran. A tutorial survey of reinforcement learning. *Sadhana*, 19(6): 851–889, 1994.
- Cliff C. Kerr, Robyn M. Stuart, Dina Mistry, Romesh G. Abeysuriya, Gregory Hart, Katherine Rosenfeld, Prashanth Selvaraj, Rafael C. Nunez, Brittany Hagedorn, Lauren George, Amanda Izzo, Anna Palmer, Dominic Delport, Carrie Bennette, Bradley Wagner, Stewart Chang, Jamie A. Cohen, Jasmina Panovska-Griffiths, Michal Jastrzebski, Assaf P. Oron, Edward Wenger, Michael Famulare, and Daniel J. Klein. Covasim: an agent-based model of covid-19 dynamics and interventions. *medRxiv*, 2020. doi: 10.1101/2020.05.10.20097469. URL <https://www.medrxiv.org/content/early/2020/05/15/2020.05.10.20097469>.
- Marco Liverani, Benjamin Hawkins, and Justin O Parkhurst. Political and institutional influences on the use of evidence in public health policy. a systematic review. *PloS one*, 8(10):e77404, 2013.
- Charles M Macal and Michael J North. Agent-based modeling and simulation. In *Proceedings of the 2009 Winter Simulation Conference (WSC)*, pp. 86–98. IEEE, 2009.
- Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.
- Don Nutbeam. Evidence-based public policy for health: matching research to policy need. *Global Health Promotion*, pp. 15, 2001.
- Mark Petticrew, Margaret Whitehead, Sally J Macintyre, Hilary Graham, and Matt Egan. Evidence for public health policy on inequalities: 1: the reality according to policymakers. *Journal of Epidemiology & Community Health*, 58(10):811–816, 2004.
- Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design. 2009. ISSN 00189448. doi: 10.1109/TIT.2011.2182033. URL <http://arxiv.org/abs/0912.3995>{%}7B{%}25{%}7D0Ahttp://dx.doi.org/10.1109/TIT.2011.2182033.
- Charles Wachira, Sekou L Remy, Oliver Bent, Nelson Bore, Samuel Osebe, Komminist Welde-mariam, and Aisha Walcott-Bryant. A platform for disease intervention planning. In *2020 IEEE International Conference on Healthcare Informatics (ICHI)* - to appear. IEEE, 2020.
- Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8(3-4):279–292, 1992.
- Jennifer Yost, Maureen Dobbins, Robyn Traynor, Kara DeCorby, Stephanie Workentine, and Lori Greco. Tools to support evidence-informed public health decision making. *BMC public health*, 14(1):728, 2014.