***Licensed to the Apache Software Foundation (ASF) under one or more contributor license agreements.  See the NOTICE file distributed with this work for additional information regarding copyright ownership.  The ASF licenses this file to you under the Apache License, Version 2.0 (the "License"); you may not use this file except in compliance with the License.  You may obtain a copy of the License at***

[***http://www.apache.org/licenses/LICENSE-2.0***](http://www.apache.org/licenses/LICENSE-2.0)***. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.  See the License for the specific language governing permissions and limitations under the License.***

**COVID CASE PROJECTION MODEL DEVELOPMENT AND STRATEGY**

**COVID-19 - Case projections Model:**

**Introduction**

This COVID case projection model is an attempt to build, in the short-term, a model to project potential upticks in COVID-19 cases, taking into accounts social distancing-related containment measures such as - education institution closures, non-essential business closures and a complete lockdown with stay-at-home orders. By including these key public health and safety measures into the projection, we believe allows for more up-to-date evaluation of the impact of these measures, as well as a more accurate assessment of whether measures such as social distancing will continue to be effective and warrant extension.

**Model Development:**

The model considered here is based on the classic Susceptible - Infectious - Recovered (SIR) epidemiological model combined with Bayesian inferencing. This model allows analyses of the number of COVID-19 positive cases in the region and then shows how the infection rates change over time as various social distancing measures (mild distancing, strong distancing and total lockdown) come into effect. Based on a research conducted on the daily number of confirmed positive COVID-19 cases in Germany (Dehning, et al., 2020), the model assumes that the infection rate in the region begins to decline over time. The rate of decline projected will depend on when the social distancing measures went into effect, and the type of social distancing measures that were put into place for that region(infection rate would reduce drastically in the case of a total lockdown and would reduce slowly in the case of only mild distancing). Thus, along with the daily count of confirmed cases in the area, the model can take into account the dates at which certain social distancing measures were initiated and whether those measures are still in place. .. Using Bayesian Markov-Chain-Monte-Carlo sampling, the model leverages the data inputs to estimate a temporal change in the virus transmission rate. This estimated transmission rate is then fed into the SIR Model, which provides a projection of & estimation of the cumulative COVID positive cases in a region for each of the social-distancing scenario, over a period of approximately 28 days.

In order to create a model that is more accurately representative of the actual reported cases, we assume the infection rate and recovery rate follow a dynamic distribution and change over time. For this process, we apply Bayesian Markov Chain Monte Carlo Sampling in order to infer a temporal infection rate from the reported cases over time, and leverage the same for our projections. Furthermore, we assume that at each level of Social Distancing implemented, the infection rate changes to follow a more representative distribution with respect to the shape of the temporal infection rate.

**Theories and Assumptions for Model Development Approach:** The creation of this model occurred in three different phases:

1. **Iteration 1**: Here, we assumed the default prior distribution to be a Half-Cauchy distribution with the infection rate and recovery rate following Log-Normal Distributions. This was a baseline model created for a short-term projection of 35 days. Extending the projection time window to a longer time period, however, did not yield good results as the default distribution applied was not sensitive enough to account for a peak in the infections in time. In other words, if we increased the projection time-window to 120 days instead of 35, the projections were seen to be continuously exponentially or linearly increasing, without a peak in the number of infections.
2. **Iteration 2**: In order to increase the projection time window, we looked at making the default distribution a little more sensitive, and redeveloped the algorithm with the Default Priors following a Half-Normal distribution instead of a Half-Cauchy Distribution. This allowed the default distribution
3. **Iteration 3**: The final iteration involved changing the Default Prior Distribution to unique shapes more unique and suitable to the Vulnerability of each State. As a result, the Social Vulnerability Index value corresponding to each State (which represents socio-economic factors for vulnerability such as transportation, accommodation, ethnicity, and other factors) was incorporated into the Default Distribution, in order to create more drastic increases in the infection rates for more vulnerable states. This created infection distributions more uniquely representative for each State.

The model runs on pymc3, where each sampling chain performs 600 steps in order to approximate a posterior distribution. As a result, we obtain three distinct projections for each scenario with respect to each location. These three projections can then be visualized to show a basic what-if scenario on each social distancing measure from the time of projection, which can also then be compared to the actual cases being reported in real-time.

**Estimated Hospital Projections:**

For hospital projections, we have not used or built any specific model to create the projection profiles.. Instead, we leveraged the projection numbers from the simulation model. This is essentially done by analyzing the rate of change in reported hospitalizations and establishing a relationship between the first order derivative of the hospitalization rates and that of the projected confirmed cases. This allows us to extend the expected rate of hospitalizations to the projection time-window as a component of the projected cases.

**Cautions when utilizing this Projection Model:**

* The accuracy and confidence limits of the projections developed by this COVID Case Projection Model depend on the use of the most up to date, publicly available data associated with the area being evaluated. As information about positive case rates for a particular region and social distancing measures in place may change over time, projections will change.
* This COVID Case Projection Model has not been endorsed, assessed or cleared for use by any public health or regulatory authorities for its proposed use.

.

**Disclaimer:**

The models developed here are simulation models and their outputs should be taken as simulated projections, not actual model predictions. Validation of the models have been conducted by comparing the projections (along with their confidence intervals) with actual reported cases in the same time period. However, because this is only a simulation model, Model may not accurately project cases due to to the anomalous changes in infection rates caused by non-adherence to social distancing, reporting errors, and other factors.

Additionally, it is to be noted that the current version of the code takes upwards of 18 hours to complete all iterations. This is primarily because the current version of the code leverages Theano (support for this open code has now stopped).

Note that updates to this COVID Case Projection Model may be made without prior notice. Always review the READ ME files prior to running this Model.

# **Reference**

Dehning, J., Zierenberg, J., Spitzner, F. P., Wibral, M., Neto, J. P., Wilczek, M., & Priesemann, V. (2020, March 29). *Inferring change points in the COVID-19 spreading reveals the eﬀectiveness of interventions .* Retrieved from arxiv.org: https://arxiv.org/pdf/2004.01105.pdf