
Team

• Prof. B. Aditya Prakash
• Prof. Bijaya Adhikari
• Prof. Naren Ramakrishnan
• Alexander Rodriguez
• Nikhil Muralidhar
• Anika Tabassum

Project website: cc.gatech.edu/~badityap/covid.html
### Team's Background

<table>
<thead>
<tr>
<th>Name</th>
<th>Lab interests</th>
<th>Projects</th>
</tr>
</thead>
</table>
| Prof. B. Aditya Prakash       | **Lab interests:** Data Science, ML and AI with applications to computational epidemiology/public health, urban computing, and web. | • Covid and flu forecasting  
• Epidemics over networks  
• Urban Analytics |
| Prof. Naren Ramakrishnan     | **Lab interests:** Data Science, ML applied to Comp. Epidemiology, Natural Language Processing, Urban Computing, Data-driven modeling of Cyber-Physical Systems. | • Ebola and flu forecasting  
• Forecasting disruptive events (EMBERS) |
| Prof. Bijaya Adhikari        | **Lab interests:** Data Science and ML to model dynamical processes (e.g., spread of misinformation, disease) on large networks (e.g., web, human contact networks). | • Covid and flu forecasting  
• Hospital acquired infections (HAIs) |
Our Participation in CDC Forecasting Initiatives

**Target 1:** Weighted influenza like illness count per week

**Target 2:** Weekly reported Covid Mortality

**Target 3:** Daily Covid-induced Hospitalizations

Since April End 2020

Last few years
Also in COVID-ILI (March 2020)
Introduction

• Goal:
  – Improve situational awareness for Covid and flu
  – Characterize different faces of the utility of the symptom survey data for forecasting

• Motivation:
  – A second wave of Covid is likely to coincide with the flu season and forecasting flu burden becomes even more crucial
  – Give policymakers valuable lead time to plan interventions and optimize supply chain decisions.

• Our approach is to jointly forecast Covid and flu

---

1 US CDC Director: https://www.washingtonpost.com/health/2020/04/21/coronavirus-secondwave-cdcdirector/
Tasks and Problem Formulation

- Surveillance systems are susceptible to symptomatic similarities.
- This makes it hard to recognize actual flu outbreaks.
- Task 1: Forecasting ILI in the Presence of Covid (Covid-ILI)
  - Use patterns from historical ILI
  - Leverage new data signals, e.g. symptom survey, mobility, Covid-related signals
- Task 2: Forecasting Covid Mortality and Hospitalizations
Our Approach: DeepOutbreak

• Two forecasting modules:
  – **Covid-ILI**: Steer a historical ILI model with Covid-related signals
  – **Covid**: Covid-19 forecasting using Covid-related signals

• Data sources (selected with epidemiological rationale):
  – FB Symptom Survey Data
  – Line-list based data from CDC, JHU, and CovidTracking
  – Mobility from Apple and Google
  – Testing from CovidTracking

• Approach features:
  – Deep learning-based approach allow us to omit laborious feature engineering.
  – Can ingest many heterogeneous signals that are more sensitive to what is happening on the ground
  – Robustness to noise and principled uncertainty estimation
  – Explainability module enables:
    • Epidemiological explanation of forecasts
    • Assess contribution of signal(s)
Task 1: Forecasting Covid-ILI When Historical Data Exists

- Steer an existing historical ILI model (EpiDeep, KDD 2019) with new Covid-related signals
- Goal: enable structured knowledge transfer from our historical ILI model to a spatio-temporal Covid-ILI model
- We use heterogeneous transfer learning and knowledge distillation
Task 2: Forecasting Covid-19

No historical data Available

- Unable to steer an existing model and unable to train temporal neural models (e.g. RNN)
- Use only Covid-related data sources.
- Principally propagate uncertainties in forecast from noise in data
- We use autoregressive training on bootstrap samples
Contribution of Symptom Survey Data in Overall Performance

Green (positive) represents increase in performance; brown (negative) decrease. Survey data improves performance in 29 of the 51 regions.
Task 1: Covid-ILI Forecasting

1. Forecasting models with survey data achieves better forecasting performance.
3. Plots showcase similar behavior with different regions (1,2) with varying degrees of pandemic impact.

Task 2: Covid Forecasting

Short-term forecasts: Using survey data independently is comparatively as effective as using in conjunction with other signals for COVID-19 mortality forecasting.

Long-term forecasts: Using survey data in conjunction with other signals is more effective for COVID-19 mortality forecasting.
**Contribution of Symptom Survey Data**

**Task 1: Covid-ILI Forecasting**

- FB survey data helps anticipate peak one week ahead

**Task 2: Covid Forecasting**

- Models without survey data overestimate peaks

- Models with survey data achieve better performance

- FB survey data in conjunction with other signals can forecast important changes in trend that were not possible only with the other signals

- Additional Signals

- Survey Data + Additional Signals

We also found interesting negative findings! (see white paper)
Results Summary

Facebook Survey Data Usage Notable Highlights:

• In general, **survey signals are orthogonal** to other available signals that we included in our models. We found them useful to improve our performance in the majority of geographical regions.

• We showed that survey signals help guide our forecasts to effectively **anticipate future trends**, which is the general case; however, there are some cases where it may lead to hinder some good trend predictions.

• In general, survey signals **should be used in conjunction** with others; however, we found a few interesting cases when they alone offer a different and more accurate forecasting perspective.

• Survey signals capture and help us in forecasting in regions with important **differences** such as **epidemic activity**. In particular, we found that in ILI forecasting, not using symptom survey data may lead to underestimating the epidemic curve.
Thanks!

Contact:
Alexander Rodríguez
aroDRiguezc@gatech.edu