Steering a Historical Disease Forecasting Model Under a Pandemic: A Case of Flu and COVID-19

Authors:

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

*Equal contribution
Outline

• Motivation
• Problem Formulation
• Approach
• Results and discussion
• Conclusion and future work
Flu in Our Society

• Every year
  – Millions get infected
  – Hundreds of thousands get hospitalized
  – Thousands die
• Surveillance and forecasting methods are key
  – Planning
  – Designing countermeasures
• Only way to directly surveillance flu is through a virological test
  – Costly, and very few testing stations
• Instead, we use Influenza like illness (ILI) reported by hospitals
  – Symptomatic data
  – ILI = fever (temperature of 100°F [37.8°C] or greater) and a cough and/or a sore throat without a KNOWN cause other than influenza.
• What is weighted influenza like illness (wILI)?
  – Department of Health and Human Services (HHS) divides the country into 10 regions.
    • Each region has a separate wILI incidence count, which is a weighted average
    • 1 national region depicting overall wILI trend

• Effect of COVID: contamination of COVID in the flu due to symptomatic similarities

• March 2020: Region 2, 9, 10 initially emerged as COVID-19 hot-spots
A Novel Forecasting Setting

• Symptomatic similarities between these two illnesses and change in patient's behavior affects our current surveillance systems.

• wILI counts may be affected by
  – COVID “contamination”
  – Shift in healthcare seeking behavior during the pandemic

• This new scenario lead us a novel forecasting problem: COVID-ILI forecasting

Rodríguez, et al., 2020
How Forecasting COVID-ILI is Useful?

• Forecasts the actual burden to hospitals
  – Helpful for resource allocation and healthcare worker deployment

• Can also be used to help with indirect COVID surveillance (Castrofino et al. 2020; Boëlle et al. 2020)
  – Especially useful at the early stages of the pandemic, when there were no well-established surveillance mechanisms for COVID

• Disambiguate trends between historical strains and new emerging strains during a flu season

Rodriguez, et al., 2020
Challenges

• Historical data alone is inadequate to represent the current scenario
• Traditional ILI models are unable to adapt
• We propose to:
  – Use patterns from historical ILI
  – Leverage COVID-related signals (limited in size)

Rodríguez, et al., 2020
Outline

• Motivation
• Problem Formulation
• Approach
• Results and discussion
• Conclusion and future work
COVID-ILI Forecasting

• **Given**
  – Historical wILI seasons
  – Partially observed COVID-ILI incidence curve \( Y_c = \{y_c^1, y_c^2, \ldots, y_c^t\} \) till week \( t \) for each region \( r \).
  – A set of COVID-related features observed till week \( t \) for all regions.

• **Predict**
  – Future incidence for next four observations \( y_c^{t+1}, y_c^{t+2}, y_c^{t+3}, y_c^{t+4} \) for each region \( r \).
Technical Challenges

- Covid related data is very sparse
- Historical wILI data is rich but does not represent any of the effects of COVID
- How to exploit spatial correlation in exogeneous signals?
- How to leverage both the rich historical wILI data and highly informative sparse features for COVID-ILI forecasting?
- Historical wILI and COVID-related signals are asynchronous
Outline

• Motivation
• Problem Formulation
• Approach
• Results and discussion
• Conclusion and future work
Covid-Augmented ILI Forecasting Network (CALI-Net)

• 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 heterogenous transfer learning and knowledge distillation losses
We propose to address this problem as a heterogeneous transfer learning (HTL) problem, we adapt the HTL framework of Moon and Carbonell, 2017. Knowledge extracted from historical wILI and from COVID-related signals are projected to a shared latent space. Use of denoising autoencoder to improve representations.
COVID-Augmented Exogenous Model (CAEM)

- Global model with joint modeling of data from all regions (10 HHS regions + National)
- Region specific embeddings (one-hot encoding + autoencoder)
- Laplacian regularization exploiting regional inter-dependencies
- Recurrent architecture to model temporal evolution
Attentive Knowledge Distillation (KD) Losses

- To better structure the knowledge transfer, we propose to incorporate KD losses; they encourage positive transfer between the COVID model (CAEM) and the historical model (EpiDeep).
- Attention in the KD losses (Saputra et al. 2019) automatically prevent negative transfer.

\[ \mathcal{L}_{KD} = \alpha \frac{1}{n} \sum_{i=1}^{n} \Phi_i \left( \left\| \hat{y}_s - \hat{y}_t \right\|_i^2 + \Phi_i \left( \left\| \Psi_s - \Psi_t \right\|_i^2 \right) \right) \]

Calculating Attention in KD Loss

\[ \Phi_i = \left( 1 - \frac{\left\| \hat{y}_s - y \right\|_i^2}{\eta} \right) \]

\[ \eta = \max (e_T) - \min (e_T) \]

\[ e_s = \left\{ \left\| y - \hat{y}_s \right\|_j^2 : j = 1, \ldots, N \right\} \]
Outline

• Motivation
• Problem Formulation
• Approach
• Results and discussion
• Conclusion and future work
Experiment Setup

• We divide the forecasting period in two:
  – T1: period of non-seasonal rise of wILI due to contamination by COVID-19 (EWs 9-11)
  – T2: period when COVID-ILI trend is declining more in tune with the wILI pattern (EWs 12-15)

• Metric: RMSE

• Results presented for next incidence prediction, more in appendix

Rodríguez, et al., 2020
Datasets

- wILI data collected by CDC and publicly available
- COVID-related signals collected from multiple public sources

Table 1: Overview of COVID-Related Exogenous Data.

<table>
<thead>
<tr>
<th>Type of signal</th>
<th>Description</th>
<th>Signals</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>(DS3) Crowdsourced symptoms based</td>
<td>Crowdsourced symptomatic data from personal devices</td>
<td>14. Digital thermometer readings;</td>
<td>(Miller et al. 2018)</td>
</tr>
<tr>
<td>(DS4) Social media</td>
<td>Social media activity</td>
<td>15. Health Related Tweets</td>
<td>(Dredze et al. 2014)</td>
</tr>
</tbody>
</table>
Baselines

Recent historical wILI models (Reich et al. 2019):
- **Delta-Density**
  - Kernel conditional density estimation, a non-parametric statistical methodology that is a distribution-based variation on nearest-neighbors regression
- **Empirical Bayes**
  - Model past seasons’ epidemic curves as smoothed versions plus noise.
  - Construct prior for the current season’s epidemic curve by considering sets of transformations of past seasons’ curves
- **SARIMA**
  - Autoregressive Integrated Moving Average model with seasonality

Also, **HIST**, a persistence baseline based on weekly average of the historical seasons
Experimental Questions

**Transfer Learning**
• Q1. Is CALI-NET able to achieve successful positive transfer to model the contamination of wILI values?
• Q2. Does CALI-NET prevent negative transfer by automatically recognizing when wILI and COVID-19 trends deviate?

**Forecasting Performance**
• Q3. Does CALI-NET’s emphasis on transfer learning sacrifice overall performance with respect to state-of-the-art methods?

**Ablation Studies**
• Q4. How does each facet of CALI-NET affect COVID-ILI forecasting performance?
• Q5. What data signals are most relevant to COVID-ILI forecasting?
Transfer Learning Results (compare vs historical ILI model)

Leverage positive transfer

Forecasting performance during period of increasing COVID-ILI leading to unseasonal peak

Prevent negative transfer

Forecasting performance during period of declining COVID-ILI trend i.e., return to traditional dynamics.

Lower is better
Overall Performance: Emphasis in Adaptation Doesn't Compromise It

- Overall model performance across period $T_1 + T_2$
- CALI-NET outperforms other models in 5 out of 11 regions, on par with DeltaDensity, a SOTA model
- CALI-NET yields competitive performance across the entire course $T_1 + T_2$

Histogram of Number of Regions where each Model is the Best Performing One
Focusing on Period T1 (uptake)

Performance Characterization in Period T1

CALI-Net outperforms all models in 9 out of 11 regions for positive transfer phase T1 where COVID-19 contamination of wILI is the greatest.
Table 1: Per-region RMSE performance characterization of the CALI-NET model when different components of CAEM architecture are deactivated.

<table>
<thead>
<tr>
<th>Regions</th>
<th>Our Method</th>
<th>CALI-NET w/o GRU</th>
<th>CALI-NET w/o Laplacian</th>
<th>CALI-NET w/o Regional Recon.</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>0.9196</td>
<td>22.0574</td>
<td>0.9118</td>
<td>0.9161</td>
</tr>
<tr>
<td>R2</td>
<td>2.6869</td>
<td>9.2662</td>
<td>2.6843</td>
<td>2.6977</td>
</tr>
<tr>
<td>R3</td>
<td>1.293</td>
<td>13.2952</td>
<td>1.3647</td>
<td>1.2965</td>
</tr>
<tr>
<td>R4</td>
<td>1.6605</td>
<td>6.9054</td>
<td>1.7944</td>
<td>1.7345</td>
</tr>
<tr>
<td>R5</td>
<td>1.5879</td>
<td>16.1975</td>
<td>1.687</td>
<td>1.6532</td>
</tr>
<tr>
<td>R6</td>
<td>2.93</td>
<td>7.8045</td>
<td>3.0516</td>
<td>2.951</td>
</tr>
<tr>
<td>R7</td>
<td>2.2805</td>
<td>5.7593</td>
<td>2.4184</td>
<td>2.322</td>
</tr>
<tr>
<td>R8</td>
<td>1.3774</td>
<td>14.9026</td>
<td>1.4898</td>
<td>1.3949</td>
</tr>
<tr>
<td>R9</td>
<td>1.8225</td>
<td>4.7056</td>
<td>1.8099</td>
<td>1.8714</td>
</tr>
<tr>
<td>R10</td>
<td>1.2069</td>
<td>6.2994</td>
<td>1.2578</td>
<td>1.2262</td>
</tr>
<tr>
<td>National</td>
<td>1.3308</td>
<td>9.9319</td>
<td>1.4597</td>
<td>1.4141</td>
</tr>
</tbody>
</table>
Effect of Knowledge Distillation

Knowledge distillation (KD) is helpful in most of the regions/weeks, especially in short term forecasting and in T2 (i.e., helping to prevent negative transfer).

KD helps improve predictions
KD is not helping
Data Ablation

DS1: Line list data
DS2: Testing data
DS3: Crowdsourced symptomatic data
DS4: Social media

Line list data is the most helpful, followed by crowdsourced and testing. Social media is the least helpful
Outline

• Motivation
• Problem Formulation
• Approach
• Results and discussion
• Conclusion and future work
Conclusions and Future Work

- We proposed CALI-Net, a novel framework for principled transfer of relevant knowledge from an existing forecasting model (based on rich historical data) to a one relying on relevant but limited recent exogenous signals.

- Characterized CALI-Net performance at different stages of the wILI season and showcase effectiveness of its transfer learning capabilities.

- Compared CALI-Net to SOTA and showcase comparable (and in many cases superior) performance of CALI-Net.

- Moving forward, we wish to:
  - Automatically differentiate outbreaks of COVID and flu.
Thanks!

Pre-print:

Code:
https://github.com/AdityaLab/CALI-Net

Contact:
Alexander Rodríguez
arodriguezc@gatech.edu