

Population Age Group Sensitivity for COVID-19 Infections with Deep Learning

Md Khairul Islam¹, Tyler Valentine², Timothy Joowon Sue¹, Ayush Karmacharya¹, Royal Wang², Levi Davis², Zhengguang Wang¹, Matt Manner², Luke Neil Benham¹, Judy Fox^{1,2}

¹ Computer Science Department, University of Virginia, USA.

² School of Data Science, University of Virginia, USA.

Email: {mi3se, xje4cy, und2yw, psb7wm, rjw8ng, ljd3frf, zw4re, xkv3na, lnb6grp, cwk9mp}@virginia.edu

Abstract

Sensitivity analysis is a popular feature selection approach employed to identify the important input features. This work proposes a novel approach that involves the perturbation of input features using the Morris method in the time-series transformer model. Each input feature is perturbed one at a time and the response of the model is examined to determine the feature's rank. Its efficacy in determining important features for real-world datasets is demonstrated with the COVID-19 forecasting model and this work quantifies the impact of population age groups. We calculated the sensitivity of our trained model to the age group features using the Morris Method and then ranked them based on sensitivity scores. We evaluated this rank with the case reports from the CDC. The results show that our method can predict infections and interpret the more impacted age groups. These results can inform public health policies in preventive with targeted vaccination strategies, to better control future pandemics.

Introduction

The COVID-19 pandemic has posed significant challenges for governments and healthcare systems worldwide, highlighting the need for effective measures to manage and control the virus's spread. Understanding the factors contributing to disease transmission is crucial for developing targeted public health policies and interventions. Age is a critical factor in COVID-19 transmission as shown by previous studies (Bae et al. 2021) (Dessie and Zewotir 2021).

Many prior works have studied forecasting COVID-19 infection and mortality rates using different statistical (Yang et al. 2020), auto-regressive machine learning (ArunKumar et al. 2021)(Kırbaş et al. 2020), and deep learning (Zeroual et al. 2020)(Ramchandani, Fan, and Mostafavi 2020)(Arik, Yoder, and Pfister 2022) models. The focus of these models is to make better predictions for the epidemic spread which would help take better mitigation steps. Interpreting how these models make these predictions can provide an understanding of the importance of these different input factors and their interactions (Ramchandani, Fan, and Mostafavi 2020). Temporal Fusion Transformer (TFT) (Lim et al. 2021) is a state-of-the-art forecasting model widely used for multivariate multi-horizon forecasting. TFT is particularly useful for modeling complex, non-linear relation-

ships between input features and target variables, making it an ideal tool for analyzing COVID-19 transmission patterns.

The Morris method (Morris 1991) has been widely used to perform sensitivity analysis of different models (Aumond et al. 2021)(Tsveltkova and Ouarda 2021). By calculating the output change with reference to input perturbation, the Morris method provides a simple approach to rank input factors by their sensitivity. By combining TFT with feature sensitivity analysis, we can gain a better understanding of the factors that contribute to COVID-19 transmission and develop more effective strategies to control the spread of the virus.

In this work, we collected the population by age subgroups data (Fig. 1) for each of the 3,142 US counties(US Census Bureau 2020), along with the daily vaccination rate of the population (Centers for Disease Control and Prevention 2023a) and COVID-19 case report from March 1, 2020, to Dec 27, 2021 (USA Facts 2023). While there are prior studies that observe COVID-19 trends for different age groups (Monod et al. 2021), we trained a TFT model on the dataset separately on each age subgroup as the static covariate and then calculated the sensitivity of those models using a modified Morris method (Morris 1991). We then ranked the age groups based on their Morris sensitivity scores to identify the most influential factors contributing to COVID-19 transmission. Finally, we evaluated our age group sensitivity rank with the actual COVID-19 cases reported for those age groups in the same time period (Centers for Disease Control and Prevention 2023b). Our study aims to provide insights into the age-related factors that contribute to COVID-19 transmission and inform targeted interventions to control the spread of the virus.

The rest of the paper is organized as follows: Section gives the background on the problem state and theoretical foundation of our work. Section 1 presents the experimental setup and model results. Section 1 features the age group ranking by Morris sensitivity score and comparison with ground truth. Section 1 summarizes the related works. Section 1 contains the concluding remarks and potential future works.

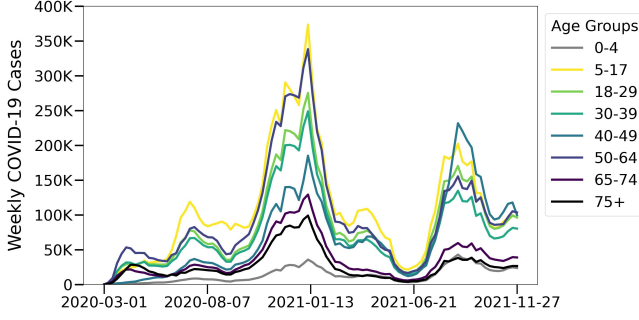


Figure 1: Ground truth: weekly COVID-19 case numbers for each of the eight age subgroups from the time period of March 1, 2020, to November 27, 2021.

Background

Sensitivity Analysis using the Morris Method

Morris method is a reliable and efficient sensitivity analysis method that defines the sensitivity of a model input as the ratio of the change in an output variable to the change in an input feature. Given a model $f(\cdot)$, and a set of k input features $X(x_1, \dots, x_k)$, the Morris sensitivity (Morris 1991) of a model input feature x_i can be defined as follows:

$$\text{Sensitivity}(\mathbf{X}, i) = \frac{f(x_1, \dots, x_i + \Delta, \dots, x_k) - f(\mathbf{X})}{\Delta} \quad (1)$$

where Δ is the small change to the input feature x_i . Since the original Morris method was for static variables, we expand it for our predictions for the high dimensional spatial and temporal datasets.

Algorithm 1 shows the implementation of the modified Morris method for our study, where we normalized the output value change by the number of input days, counties, and delta Δ . Which we call the normalized Morris index $\hat{\mu}^*$. We further scaled this index using the standard deviation (σ) of the input feature x_i , which we call the scaled Morris index ($\hat{\mu}^* \times \sigma$). This scaling takes the underlying distribution of the feature x_i when ranking the features by sensitivity. In the rest of the study, by Morris score, we refer to the scaled Morris index.

Problem Statement

In this paper, we use deep learning to study feature sensitivity for model predictions of COVID-19 infection at the county level. The workflow for the design of our experiments is illustrated in Fig. 2.

Model Training for Predictions using TFT

The Temporal Fusion Transformer (TFT) (Lim et al. 2021) is a novel, interpretable, attention-based deep learning model for multi-horizon forecasting. Its architecture is carefully designed to handle static (e.g. percentage of *age group* in a county population), past observed (e.g. *infection rate of*

cases, daily vaccination coverage), and known future (e.g. day of the week) inputs. Its architecture is specially modified for four main uses. 1) To learn both local and global time-varying relationships at different scales. 2) To filter out input noises using Variable Selection Network (VSN). 3) To incorporate static metadata into the dynamic features for temporal forecasting. 4) To efficiently use different parts of the network through a gating mechanism. TFT achieves state-of-the-art performance on a wide range of real-world tasks. Hence we use the model for our own study here.

The prediction model $f(\cdot)$ is denoted as bellow:

$$\hat{y}_i(t, \tau) = f(\tau, y_{i,t-k:t}, \mathbf{z}_{i,t-k:t}, \mathbf{x}_{i,t-k:t+\tau}, \mathbf{s}_i) \quad (2)$$

where $\hat{y}_i(t, \tau)$ is the number of cases predicted at time $t \in [0, T_i]$ for county i . τ is the forecast horizon, 15 days in our case. T_i is the length of the time series period, which for our case is the same for each county. We use the previous 13 days (lag window k) of data to forecast the next 15 days.

Experimental Setup

Computational Resources

We implement our TFT model with PyTorch (PyTorch 2024). Then we conducted a performance evaluation of the model training on HPC clusters including the GPU nodes in Table 1. Each training epoch takes an average of 50 minutes on a GPU node with at least 32GB of RAM. Each Morris runs with a trained model, and with additional feature analysis that takes around 35 minutes per epoch.

Table 1: Runtime environment and hardware specification.

Driver	CUDA	Processor	NVIDIA GPU
470.82.01	11.4	Intel Xeon	A100-SXM4-40GB
			Tesla P100-PCIE
			Tesla V100-SXM2
			Tesla K80

Input Data and Features

The data set consists of daily COVID-19 cases for over 3,142 US counties with eight static features and one dynamic feature. The static features signify the percentage of all people in one of the eight age subgroups (0-4, 5-17, 18-29, 30-39, 40-49, 50-64, 65-74, and 75 and older) for all counties. The dynamic feature indicates the percentage of people who are fully vaccinated for all counties.

The data set was split into training, validation, and testing sets. The training set includes the dates March 1, 2020, through November 27, 2021, the validation set includes the dates November 28, 2021, through December 12, 2021, and the testing set includes the dates December 13, 2021, through December 27, 2021.

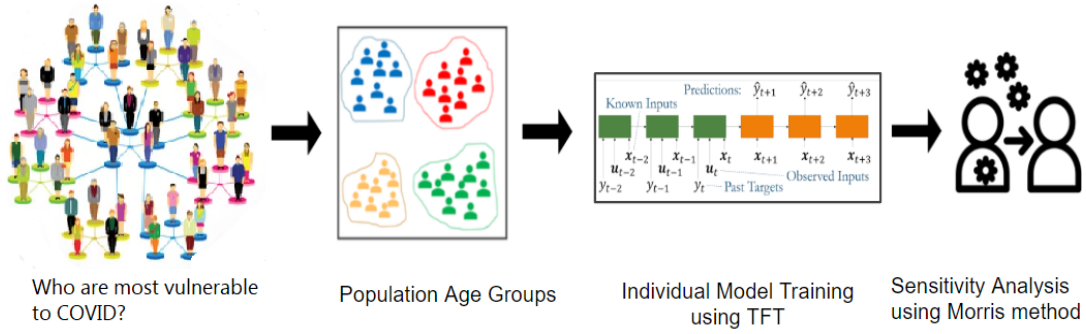


Figure 2: Workflow of Population Age Group Sensitivity Analysis for COVID-19 Forecasts.

Algorithm 1: Modified Morris score calculation for **spatio-temporal** data

Input: $\mathbf{X} = \{x_1, \dots, x_k\}$, target feature $x_i \in \mathbf{X}$ with dimension $[C, T]$, model f , Δ
 // \mathbf{X} is a set of k input features, Δ is the change to x_i
 $\mathbf{Y} = f(\mathbf{X})$
 $\mathbf{Y}_\Delta = f(x_1, \dots, x_i + \Delta, \dots, x_k)$
 $G, t = 0, 0$
while $t < T$ **do** /* Temporal */
 // Loop through 637 Days
 $c = 0$
 while $c < C$ **do** /* Spatial */
 // Loop through 3142 US Counties
 $G = G + \mathbf{Y}_\Delta[c][t] - \mathbf{Y}[c][t]$ /*Total Change*/
 $c = c + 1$
 $t = t + 1$
 $c = 0$
 // Calculate normalized Morris Index $\hat{\mu}^*$
 $\hat{\mu}^* = G / (C * T * \Delta)$
 // Get standard deviation of the feature
 $std = \text{standard_deviation}(x_i)$
return $\hat{\mu}^* \times std$

Model Training and Prediction

A separate TFT model was trained for each age group, for a total of eight models. Since all static variables are presented together as context vectors, the separation ensures the identification of a specific subgroup’s contribution without the interference of other correlated static variables. In addition to the static age feature, all TFT models include a dynamic vaccination status feature. An additional variable SinWeekly is also included, which represents weekly trends for changes in the number of new cases. The target variable for each model was the daily number of cases for each county. The daily predictions for each county were determined for the train-

ing, validation, and testing time periods. These predictions were compared to the observed values to obtain RMSE values.

Table 2: RMSE values for all TFT models.

Model	Train	Validation	Test
0-4	56.49	82.34	192.30
5-17	60.65	80.25	199.40
18-29	56.89	82.24	193.80
30-39	56.95	82.22	199.10
40-49	52.45	80.65	197.70
50-64	56.70	82.53	200.40
65-74	56.06	81.52	199.40
75+	60.72	83.80	204.40

The RMSE values are shown in Table 2 for each of the eight models. The "Model" column indicates the age subgroup used as the static variable for that model. As expected, all models performed well when predicting the test data set and have comparable loss values.

Sensitivity Analysis

Next, our Modified Morris Method was applied to each model. This involved perturbing the age feature in each model by a range of positive and negative delta values. The chosen values ranged from -0.01 to 0.01 with steps of 0.001, disregarding the delta value equal to zero. The Morris indices for each of the delta values are shown below (Fig. 3). Positive and negative values of delta indicate an increase and decrease in the age feature, respectively. Investigating both an increase and decrease in these features allowed for an understanding of how the Morris indices change between different delta values. Our results show that the Morris values are approximately stable for each age group across a range of values.

Another observation about these results is the magnitude of the Morris values. With the delta and Morris values shown at the same scale of 10^{-4} , the Morris values remain about two to three orders of magnitude smaller than the delta values. This means that the changes caused by the perturbation of

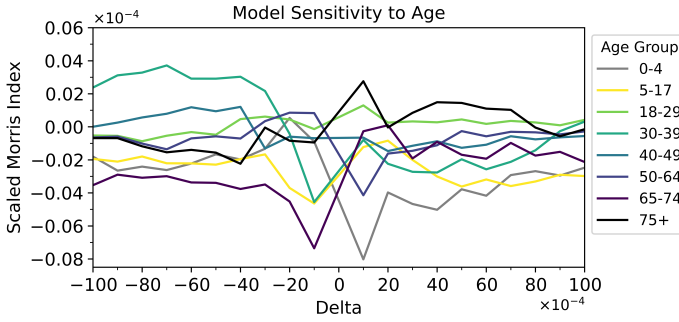


Figure 3: Morris indices as a function of delta for all eight models, calculated at intervals of 0.001, excluding zero.

the static variable result in small changes in the predictions of the models, indicating a non-linear relationship between the static input and the output.

Despite the relatively small values, there is a clear distinction between which age groups have higher Morris indices on average. This separation allowed us to rank the age groups based on their sensitivity to the models. Before this was done, it was necessary to obtain a ground truth ranking to use for verification of our ranking derived from the Morris method.

Data was obtained from the CDC for the total number of COVID-19 cases reported for each age group within the time period of the training set, as shown previously (Fig. 1). Data from the US Census were also obtained to provide an estimate of the total population of each age group, with the estimate made for April 1, 2020. Having both the cases and the estimated population allowed us to calculate the true ranking. This was done by calculating the percentage of people in each age group who were infected. For example, 10,018,923 cases were reported for the age group 18-29, with an estimated population of 54,992,661 for that age group. This results in an infection rate of 18.2 percent. The infection rates for all age groups are shown below (Fig. 3).

Table 3: Total number of cases, total population, and infection rate of each age subgroup according to CDC and US Census data.

Age Group	Cases	Population	Infection Rate (%)
0 - 4	1.25 M	19.4 M	6.4
5 - 17	6.18 M	55.0 M	11.2
18 - 29	10.0 M	53.0 M	18.9
30 - 39	7.76 M	45.0 M	17.2
40 - 49	6.77 M	41.0 M	16.5
50 - 64	8.82 M	63.9 M	13.8
65 - 74	3.29 M	32.4 M	10.2
75+	2.51 M	21.8 M	11.5

Based on the infection rate within each age subgroup, a ranking was made ranging from 1, the age group with the lowest infection rate, to 8, the age group with the highest infection rate. A ranking of the sensitivity of the model to each

age feature was also determined by first ranking the relative Morris values for each age group for a given delta to obtain a ranking from 1 (lowest) to 8 (highest) value. After ranking each subgroup for all 20 delta values, the average ranking was determined. Both of these rankings, along with the difference between them, are shown in Fig. 4.

Table 4: Ranking of the contribution of each age group to the total number of COVID-19 cases based on Infection rate or based on the Morris ranking (Morris), along with the difference between both ranks (Difference).

Age Group	Infection Rank	Morris Rank	Difference
0 - 4 Years	8	8	0
5 - 17 Years	6	7	1
18 - 29 Years	1	1	0
30 - 39 Years	2	3.5	1.5
40 - 49 Years	3	2	1
50 - 64 Years	4	5	1
65 - 74 Years	7	6	1
75+ Years	5	3.5	1.5

Since the average Morris ranking for the age groups 30-39 and 75+ were found to be identical, we chose to break the tie arbitrarily. We noticed that the 5-17 age group ranking of 2 was significantly different from the ground truth ranking of 7. However, our results indicate that the Morris ranking closely matches the ground truth ranking for the other seven age groups, as most ranks are identical or have a difference of 1 to 3. This is interpreted as indicating that our predictions of which age groups contribute most to the total COVID-19 case numbers during the time period of the data are approximately equal to the true values verified by CDC and US Census data.

Related Work

Numerous works have been done on COVID-19 forecasting using deep learning and other AI or statistical models (Zeroual et al. 2020; Clement et al. 2021). Understanding the feature importance and interaction of the input factors is crucial to the adaptation of control measures based on the changing dynamics of the pandemic.

Prior work involved (Malik et al. 2022) simulation of the spread of COVID-19 using a fractal-fractional model and used sensitivity analysis to assess the impact of model parameters such as the transmission rate, recovery rate, and vaccination rate on the outcomes of the model. Other work (Ni et al. 2022) investigated the impact of key model parameters on the outcomes of a Susceptible-Carrier-Exposed-Infected-Recovered (SCEIR) model predicting COVID-19 infection. Similarly, additional work (Monod et al. 2021) analyzed the demographic patterns of COVID-19 cases during the recent resurgence of the pandemic in the United States. The data, consisting of 948 counties, studies the age distribution of cases and their relationship with changes in COVID-19 incidence.

Conclusion and Future Work

Typically, deep learning lacks sufficient interpretability to understand the decisions a model makes to improve predictive performance. However, our work demonstrates the successful application of the Morris method. By analyzing model sensitivity to changes in the age input feature, we explained the impact of COVID-19 on various age group populations in the US between the time period of March 1, 2020 to November 27, 2021. Small perturbations in the static features showed that our models were most sensitive to the 18-29, 30-39, and 40-49 age groups, indicating that these groups contributed most to the total number of COVID-19 cases. These results were then verified by comparison to the ground truth from CDC and Census data.

After demonstrating the effectiveness of the Morris method applied to deep learning models for understanding the impact of COVID-19 on different age groups, we intend for future experiments to center on exploring the temporal component of these predictions. CDC data show that the ranking of COVID-19 infection rates by age group changes over time. By training new TFT models for a variety of time periods, we can better understand how the impact of COVID-19 varies throughout the pandemic and can further strengthen the reliability of our methods.

Acknowledgment

This work is supported by NSF grant CCF-1918626 Expeditions: Collaborative Research: Global Pervasive Computational Epidemiology, and NSF Grant 1835631 for CINES: A Scalable Cyberinfrastructure for Sustained Innovation in Network Engineering and Science.

References

- Arik, S. O.; Yoder, N. C.; and Pfister, T. 2022. Self-Adaptive Forecasting for Improved Deep Learning on Non-Stationary Time-Series. *arXiv preprint arXiv:2202.02403*.
- ArunKumar, K.; Kalaga, D. V.; Kumar, C. M. S.; Chilkoor, G.; Kawaji, M.; and Brenza, T. M. 2021. Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA). *Applied soft computing*, 103: 107161.
- Aumond, P.; Can, A.; Mallet, V.; Gauvreau, B.; and Guillaume, G. 2021. Global sensitivity analysis for road traffic noise modelling. *Applied acoustics*, 176: 107899.
- Bae, S.; Kim, S. R.; Kim, M.-N.; Shim, W. J.; and Park, S.-M. 2021. Impact of cardiovascular disease and risk factors on fatal outcomes in patients with COVID-19 according to age: a systematic review and meta-analysis. *Heart*, 107(5): 373–380.
- Centers for Disease Control and Prevention. 2023a. COVID-19 Vaccinations in the United States Counties.
- Centers for Disease Control and Prevention. 2023b. COVID-19 Weekly Cases and Deaths by Age, Race/Ethnicity, and Sex.
- Clement, J. C.; Ponnusamy, V.; Sriharipriya, K.; and Nandakumar, R. 2021. A survey on mathematical, machine learning and deep learning models for COVID-19 transmission and diagnosis. *IEEE reviews in biomedical engineering*, 15: 325–340.
- Dessie, Z. G.; and Zewotir, T. 2021. Mortality-related risk factors of COVID-19: a systematic review and meta-analysis of 42 studies and 423,117 patients. *BMC infectious diseases*, 21(1): 855.
- Kırbaş, İ.; Sözen, A.; Tuncer, A. D.; and Kazancıoğlu, F. Ş. 2020. Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches. *Chaos, Solitons & Fractals*, 138: 110015.
- Lim, B.; Arik, S. Ö.; Loeff, N.; and Pfister, T. 2021. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4): 1748–1764.
- Malik, A.; Alkholief, M.; Aldakheel, F.; Khan, A. A.; Ahmad, Z.; Kamal, W.; Gatasheh, M. K.; and Alshamsan, A. 2022. Sensitivity analysis of COVID-19 with quarantine and vaccination: A fractal-fractional model. *ScienceDirect*, 1–16.
- Monod, M.; Blenkinsop, A.; Xi, X.; Hebert, D.; Bershan, S.; Tietze, S.; Baguelin, M.; Bradley, V. C.; Chen, Y.; Coupland, H.; et al. 2021. Age groups that sustain resurging COVID-19 epidemics in the United States. *Science*, 371(6536): eabe8372.
- Morris, M. 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2): 161–174.
- Ni, G.; Wang, Y.; Gong, L.; Ban, J.; and Li, Z. 2022. Parameters Sensitivity Analysis of COVID-19 Based on the SCEIR Prediction Model. *MDPI*, 1–19.
- PyTorch. 2024. Temporal Fusion Transformer.
- Ramchandani, A.; Fan, C.; and Mostafavi, A. 2020. Deep-covidnet: An interpretable deep learning model for predictive surveillance of covid-19 using heterogeneous features and their interactions. *Ieee Access*, 8: 159915–159930.
- Tsvetkova, O.; and Ouarda, T. B. 2021. A review of sensitivity analysis practices in wind resource assessment. *Energy Conversion and Management*, 238: 114112.
- US Census Bureau. 2020. County Population by Characteristics: 2010-2020.
- USA Facts. 2023. (US COVID-19 cases and deaths).
- Yang, Z.; Zeng, Z.; Wang, K.; Wong, S.-S.; Liang, W.; Zanin, M.; Liu, P.; Cao, X.; Gao, Z.; Mai, Z.; et al. 2020. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of thoracic disease*, 12(3): 165.
- Zeroual, A.; Harrou, F.; Dairi, A.; and Sun, Y. 2020. Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. *Chaos, Solitons & Fractals*, 140: 110121.