Injecting External Metadata into Foundational Time-Series Models For Epidemiological Forecasting

Arjun Verma
Georgia Institute of Technology
Atlanta, Georgia, USA
averma332@gatech.edu

Ben Zabriskie
Georgia Institute of Technology
Atlanta, Georgia, USA
bzabriskie3@gatech.edu

Harsha Kamarthi*
Georgia Institute of Technology
Atlanta, Georgia, USA
hkamarthi3@gatech.edu

Abstract

Time series forecasting is crucial for epidemiological decision-making, particularly in the early stages of disease outbreaks when quick and accurate predictions can significantly impact public health outcomes. While foundational time-series models (FMs) offer flexibility and generalizability, they often lack domain-specific knowledge that could enhance their predictive accuracy for epidemiological applications. This paper presents a novel approach to improving FM performance by integrating external metadata, specifically exploring how text-based epidemiological data can guide and enhance the forecasting capabilities of Amazon's Chronos model for influenza-like illness (ILI) prediction. We develop a pipeline that combines time series encodings from Chronos with text encodings from SentenceTransformers using cross-attention fusion, incorporating metadata such as Google search trends and temperature data. Our experimental results demonstrate that this metadata-integrated approach outperforms both traditional ARIMA models and fine-tuned Chronos baselines, achieving a 19.1% improvement in mean squared error compared to ARIMA and a 56.4% improvement over the fine-tuned Chronos model without metadata. These results suggest that integrating relevant epidemiological metadata into foundational time series models can significantly improve their forecasting accuracy while maintaining their inherent flexibility and rapid deployability.

Keywords

Machine Learning, Time Series, NLP, Epidemiology

ACM Reference Format:
Arjun Verma, Ben Zabriskie, and Harsha Kamarthi. 2024. Injecting External Metadata into Foundational Time-Series Models For Epidemiological Forecasting. In . ACM, New York, NY, USA, 6 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction/Motivation

Disease outbreaks require quick decisions from public health officials and policy makers. Quick and appropriate actions are critical in the early stages of an outbreak to minimize its impacts. To properly make these decisions, decision makers must have quality information. The best information they can have is an accurate forecast of how bad the outbreak will be, where and when. In other words, good time series forecasting of how a disease outbreak will progress has the potential to save lives and inform decision makers. However, at the early stages of outbreaks, when good decisions can have the biggest impact and benefit, data and information is often limited. There may not be sufficient understanding of a new disease or variant, the data may be limited, etc. Thus, we seek to develop an approach that is quickly usable and accurate without careful and time intensive modeling to adapt it to a specific disease. Specifically, we seek to improve the forecasting of influenza like illness (ILI) time series.

2 Response to Milestone Comments

Comment: *What qualitative analysis/case studies will you perform to study how the models used different sources of data? Ablation studies on this will be very useful*

Response: Due to noisiness in some of our text metadata, only two sources of data were used. These were the Google search trends and temperature data described in Section 5.3.2. Unfortunately, limited time and computational resources meant we were only able to train the model using both of these sources and no ablation studies were able to be completed to better isolate their individual effects. This is discussed as an avenue of future work in Section 7.

3 Problem Definition

Epidemiological forecasting is a difficult problem where guiding inputs such as text have not been fully applied. There is a need for a highly accurate and generalizable model that can forecast epidemiological data with the guidance of flexible text inputs. Such a model would be able to be used in the early stages of a disease or outbreak, when specific modeling of that disease is infeasible and another approach is needed. The need for such a model was evinced by the COVID-19 pandemic, where a model that could have, for example, responded to news articles regarding COVID's early rise would have been incredibly useful. With this, our goal is to integrate relevant text metadata into an existing general time series forecasting model to quickly and adaptively improve its disease forecasting capabilities. As we will discuss in the sections that follow, we will modify existing foundational time-series models (FM).

*Idea originally presented by Harsha

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/XXXXXXX.XXXXXXX

4 Related Work and Survey

4.1 Epidemiological Time-Series Models

As stated previously, the current state of the art (SOTA) in epidemiological forecasting does not utilize the power of FMs and thus misses out on a lot of data that could make it more accurate and generalizable. Additionally, these models are carefully calibrated to certain diseases and do not necessarily transfer well, and also cannot handle complex guiding metadata inputs such as text. We will consider a breadth of papers that were published in 2023/2024 on time series analysis for epidemiology. We first consider the LSTM developed by Wood for Varicella infection prediction, which can only handle a relatively small dataset and is trained specifically for Varicella, without being generalizable [18]. As well as this, the RF algorithms employed by Vaughn et. al. in 2023 to do time series forecasting on COVID infections using wastewater metadata required very careful tuning of sampling frequency, dataset size and other factors to get a result on specifically COVID forecasting [17]. Multiple other Epidemiological time series models are fraught by similar issues caused by their inability to train on the massive amounts of data that foundational models reap the benefits of as well as being unable to handle live input of changing metadata [12] [1].

4.2 Foundational Time-Series Models

As noted previously, this project aims to improve the performance of foundational time-series models for epidemiology. FMs are generic models that are intended to be used to analyze a wide-range of time series. They are often pre-trained on large datasets, using time-series data of varying types and from various domains. This makes them flexible and allows them to be generically applied to new time-series problems readily [13]. This flexibility could be quite useful for, as an example, forecasting an outbreak of a young disease that has not been thoroughly studied. In the early stages of the outbreak, public health decisions must be made quickly to respond and slow or stop the outbreak of the disease. However, properly developing accurate models for a given disease takes time and potentially large amounts of data, both of which are unlikely to be available. In this case, FM models can be applied straight from the box to provide a baseline prediction.

However, this flexibility means that FMs may not be taking into account important domain knowledge. For our interest of epidemic forecasting, this domain knowledge is critical in ensuring our predictions are as accurate as they can be. While pre-training using a wide-range of time series data may produce reasonably good predictions, this may not be sufficient for public health applications where forecasts help inform literally life-critical decisions. Therefore, any way to improve these predictions should be explored thoroughly. For instance, by integrating epidemiological metadata.

4.3 Existing Time Series FMs

There already exist a variety of FM models that can be used as a baseline for our work.

Chronos [7], developed in collaboration with Amazon, uses a transformer-based, self-supervised approach to create pre-trained probabilistic time series models for general time-series use. It works by tokenizing time series values through scaling and quantization, then training the models on these tokenized series by using cross-entropy loss. The models have been pre-trained on a large amount of both publicly available datasets and synthetic data generated via Gaussian processes. The creators of CHRONOS assert that their models have comparable or even better zero-shot performance on new datasets relative to methods trained specifically for those datasets.

Google's TimesFM [6] is a decoder-only model that uses input patching to split time series into a decodable sequence of tokens. It was again trained using a very large set of time-series data from both real and synthetic time series. As with Chronos, the creators of TimesFM show that their model can consistently produce accurate zero-shot forecasts on previously unseen time-series data from various domains.

There are other time-series FM models, including MOMENT, which claims to be the first to provide these open-source, pre-trained time series models. Additionally, MOMENT [10] makes available the Time Series Pile, a collection of publicly available time-series datasets they used to train their models.

4.4 Multimodal Input

Multimodal input to FMs is a crucial aspect of our approach - we want to be able to feed epidemiological metadata such as EHRs, population dynamics, etc. to our FM to increase its forecasting accuracy. Ideally, more information will lead to a better model. Guo et. al describe one of the most prevalent ways of adding multimodal input to a model to be creating a latent representation of the auxiliary input and adding that as input to the original model [11]. However, they also describe how the latent representations of the different inputs may all lie in different subspaces, reducing the overall utilization of the data by the model - this is known as the heterogeneity gap.

Multiple different approaches to closing the heterogeneity gap across different domains have been introduced, such as video2vec [2] and CLIP [15]. In particular, CLIP embeds images and text into the same space by contrastively training images to have the same embedding as their captions. We could potentially explore this style of contrastive learning in order to enforce our latent metadata representation to sit in the same subspace as the latent time series tokens.

There have also been direct attempts to add multimodal input to time series forecasting models. Zhijan et. al recently introduced TGForecaster, a robust model that fuses textual and time series inputs with cross attention to create a text guided time series model [19]. Text is an incredibly rich domain with multiple different well defined encoders [14] and using it as a medium to integrate metadata into our model would provide a functional way to achieve our goal.

5 Proposed Method

5.1 Method

Our proposed method is to utilize Chronos, Amazon's foundational time series transformer, and SentenceTransformers to encode both time series and text input, combine them with cross attention and decode the fused encoding to get a forecasting output. Insight into

the specifics of our method and the reasoning behind it can be found in section 5.3.

5.2 Intuition

Our proposed method seeks to combine the strengths of more traditional epidemiological modeling with the previously described FM time series models. Time series FMs are trained on huge amounts of time series data. This makes them effective at general time series forecasting. However, this data is from many different kinds of time series and is not specific to disease spread, much less the spread of a specific disease. Incorporating relevant disease-specific metadata will allow the FM to leverage both its general time series training, but also more appropriately take into account the most relevant data for the specific problem (i.e flu forecasting). As well as this, it will allow our model to respond dynamically to major changes in the epidemiological landscape that are reflected by metadata indicators (such as news articles about the initial coming of COVID). Preliminary results from our data provide support for this theory.

A lag correlation analysis was performed between the base time series data we are forecasting and some of the metadata we collected (Google search term trends). For the particular search terms evaluated, a PCC of 0.75 was achieved with a two week lag. The time series of the base ILI data and the search term trend with a 2 week lag applied are shown in Figure 1.

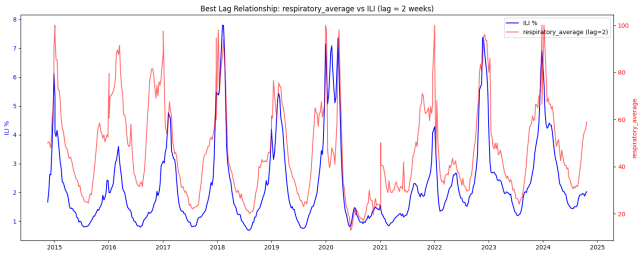

Figure 1: Two week lage of Google search terms overlayed on ILI

These initial results show that the metadata we have collected provide early indicators of ILI trends. Thus, if given this and other relevant metadata, our model should be able to more accurately preemptively predict trends in ILI, as the metadata preemptively predicts peaks in ILI.

5.3 Description

5.3.1 Overall Approach Summary

Our overall approach was to integrate text metadata into Chronos through cross attention in the latent space and then train the model on ILI data and Epidemiological metadata. This approach is supported by prior work in text integrated time series forecasting [19] and also provides a transparent and viable way to integrate text metadata. The following sections will detail the specifics of our data collection, processing, model integration and training processes.

5.3.2 Data

We have two categories of data. The first category is the base time series ILI data. This data was downloaded from the CDC Fluview dashboard [3] as a .csv file. This dataset includes weekly ILI numbers for the U.S. since 1997. Data can also be collected easily by HHS region or by state for future work. However, our work here focuses on forecasting at the national level.

The second category of data is the metadata used to guide the FM with the goal of improving forecasting accuracy. The following metadata was collected:

Temperature Data: By accessing the National Oceanic and Atmospheric Administration's (NOAA) Climate Data Online API [8], we accessed daily temperature readings from all NOAA stations. Stations were selected from across the contiguous U.S. The average daily temperature data for each of these stations was collected from the API since 2014. These results were then averaged over each station to obtain a measure for the average U.S. temperature for a given day. These results were then averaged across weeks to match the weekly ILI data.

Google Search Trends: By accessing the Google Search Trends dataset [9], we collected search trend data for various keywords and groups of keywords. The pytrends [20] API was used to collect the trends results and returns the results relative to the peak search volume in the period selected on a scale of 1 to 100. Additionally, it only offers weekly data granularity if the data is pulled one year at a time. Therefore, the search data was collected per year. When training our model, we used search trend data for the year the training point is in. The search terms collected include 'flu', 'cough', 'sick', 'fever', and 'shortness of breath'. Search data for each term since 2014 was collected. The model can be trained using data for individual search terms, or the terms can also be grouped into categories. For example, 'cough' and 'fever' could be grouped into a symptoms category.

New York Times Headlines: By accessing the New York Times (NYT) API [16], we collected data on headlines from the NYT article archive. We collected headlines that include the terms 'flu' and 'influenza'. The data collects the headline title itself as text, as well as the date of the article, its keywords and a one or two sentence snippet from the article. We can use this data both to track the number of related articles published as a time series, as well as inputting the headlines and text data itself into our model.

CDC Reports: In addition to pulling the baseline time series ILI data from the CDC, we wrote a Playwright based web-scraper to pull the text from the weekly CDC ILI reports [4]. The pdfs or html pages of the reports were saved as JSON files and then put as text into a .csv file. In addition to weekly ILI reports, we also scraped the weekly CDC Nationally Notifiable Infectious Diseases and Conditions (NNDSS) reports for other infectious diseases [5]. We scraped the data for all the available diseases, but may want to test subsets of diseases to get a better sense of which diseases are the best predictors and why.

Finally, to actually use this data we have obtained and incorporate it into the FM, we wrote a PyTorch dataloader. This dataloader will allow us to pull the ILI time series and metadata for each time period to train and evaluate the model.

5.3.3 Model Pipeline

Our pipeline takes a list of ILI values for a set of epiweeks and the associated text metadata for each week. It then passes the ILI values into the Chronos model encoder to get a time series encoding and the text values into SentenceTransformers' "all-MiniLM-L6-v2" model to get a text encoding. We chose to use Chronos because it achieved similar metrics to TimesFM and Moment while also having a simpler, easier to integrate architecture (basic transformer). We chose to use SentenceTransformers because it is the current SOTA in open source sentence encoding and had the best support in Python. After getting the time series and text encodings, we combine them with CrossAttention and pass the combined encoding through the decoder portion of Chronos in order to get our forecasting output. CrossAttention is not the only fusion module we implemented - we also consider a basic sum of encodings as well as a concatenation of encodings. We then train this pipeline on our collected and preprocessed data to get a trained model. Details of the training process and selected encoding fusion can be found in section 6.

The first step in creating our forecasting pipeline was developing the latent fusion modules. We developed a simple cross attention model using PyTorch's MultiHead Attention model, a simple sum model with PyTorch, and a simple concatenation with PyTorch. From here, we had to edit Chronos's internal implementation in order to integrate our modality fusion module. This process introduced a slew of technical difficulties, as Chronos's internal workings are opaque and poorly documented. After careful debugging and analysis, we found that Chronos utilized Google's T5Stack, a modular set of transformers designed for NLP tasks. We were able to extract and edit the encoder portion of the T5 transformer, adding our module to the class and its output to the forward pass of the encoder. This integration allowed the model pipeline for one forward pass to work perfectly. From here, we had to create a custom training script for our model based off the Chronos training script. We did so by integrating our text encoding data into the base Chronos time series dataset such that the relevant text data would be returned alongside the ILI time series at each training iteration. This concludes the description of our implementation of the model pipeline.

6 Experiments/Results

6.1 Experimental Questions and Testbed

Our experiments are designed to answer whether or not text inputs can improve the accuracy of foundational time series models on ILI and also to determine whether or not foundational time series models are applicable and outperform baselines in epidemiological forecasting. Our testing baselines include a basic ARIMA forecasting model and a fine tuned (purely on ILI, not text metadata) version of Chronos. Our baseline experiment calculates MSE (mean squared error) over the entire test set with a context length of 16 weeks and a prediction length of 16 weeks using all three models to determine which model can achieve the lowest MSE. We chose these lengths as they reflect a relevant epidemiological task of predicting far future ILI (16 weeks) based off previous data. Our goal is to outperform both the fine tuned and ARIMA baselines and prove that text metadata can improve the accuracy of foundational time series models in ILI forecasting and prove their efficacy in the field. Training of both the metadata integrated and base Chronos models was done on an NVIDIA H100 GPU accessed through Georgia Tech's PACE-ICE Cluster, with each model receiving 15000 epochs of training with an initial learning rate of 0.0001, batch size of 8, context length of 16, prediction length of 16 and 4096 total tokens in the embedding space. These hyperparameters were selected as they achieved best performance over our hyperparameter selection process.

6.2 Experimental Results

6.2.1 Initial Experiments

We initially tested multiple different combinations of metadata inputs and fusion techniques, however all of them except for the one utilized failed (order of magnitude higher than the baseline ARIMA MSE error). The combination that gave us a viable result was utilizing only the Google search and temperature metadata and the CrossAttention fusion. We postulate that the NYT and CDC report data was too sparse (only existed for about 30 percent of the training data) and too noisy for the model to learn considering the limited amount of data we were training on. This idea is supported by comparing the top dimensions of the text encoding of the CDC reports to the corresponding ILI data, as can be seen in Figure 2. This plot compares the encoding values of the top 5 dimensions of the text encoding of the CDC ILI reports against the ILI time series. These reports should represent the same thing as the ILI time series, just in text format and with additional information. However, as can be seen in Figure 2, the text encoding is very noisy and has no discernible connection to the base ILI time series.

We further postulate that the sum and concatenation fusion modules failed to generate good results because they did not foster a synthesis that allowed for complex interactions between the string metadata and time series encodings.

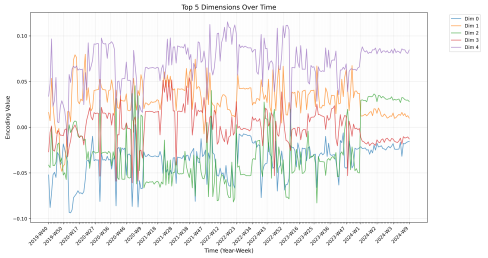

Figure 2: Enter Caption

6.2.2 Quantitative Results

As stated previously, we found that the pipeline that worked best empirically was one that only considered the Google Search and Temperature data and used CrossAttention for the encoding fusion. The results of our optimal Text Metadata Integrated Model compared to the base ARIMA and Fine Tuned Model can be seen below in Table 1:

Model	MSE	% Difference
Text Metadata Integrated Chronos	**3.46**	–
ARIMA	4.12	+19.1%
Fine Tuned Chronos (no text metadata)	5.41	+56.4%

Table 1: Model MSE Scores on Test Set

Our metadata integrated model outperformed both models, outperforming the ARIMA baseline by 19.1 percent and significantly outperforming the fine tuned model by 56.4 percent. Interestingly, we see that the ARIMA baseline actually outperforms the fine tuned Chronos model - this may be due to the fact that our training set was too small to fine tune on properly (i.e. the fine tuning overfit). We expected the text metadata model to outperform the fine tuned baseline, however it may outperform it to this high degree due to the fact that the text metadata added extra size and variance to the training set, thus discouraging overfitting. Regardless, our hypothesis has been proven correct - adding relevant epidemiological text metadata to foundational time series models improves their performance on the prediction of ILI.

6.2.3 Qualitative Results

Our metadata integrated also tends to produce plots more aligned with the ground truth ILI than the fine tuned baseline. Two plots of predicted ILI on two randomly sampled sections of the test epiweeks that compare the results of the basefine tuned and text integrated models can be seen below:

The base fine tuned model tends to get "confused", predicting a large spike when the ILI is trending downward and shooting down when the ILI is supposed to spike. The text integrated model tends to do much better, more closely following the trend of the true result in both cases. This may be the case because certain text metadata such as Google Search trends are "predictors" for when ILI is going to spike or drop, giving the model more insight than just raw time series input. As well as this, the text guided model is seemingly able predict complex trends in the data, for instance being able to peak when necessary and drop afterward in example (c).

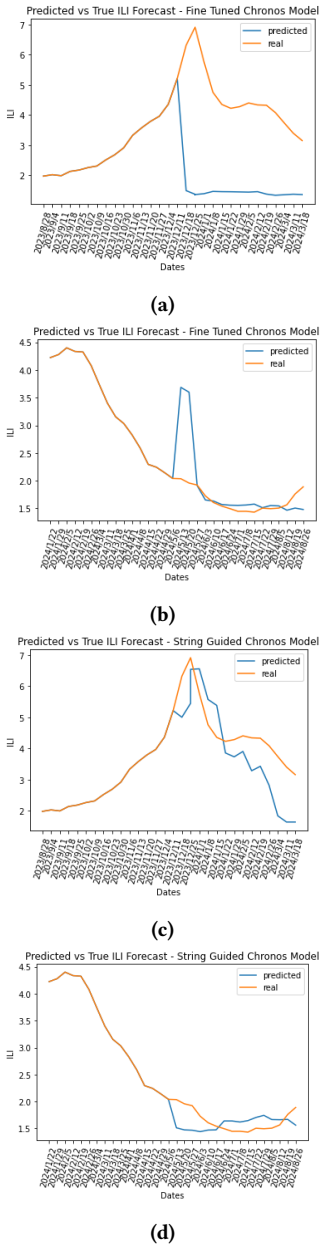

Figure 3: Base Fine Tuned Model vs Metadata (String) Guided Model

7 Conclusion

Overall, our work has shown that the integration of encoded text metadata into the Chronos FM improves its forecasting accuracy for ILI. The string guided Chronos model outperformed both the baseline Chronos model fine tuned on only ILI data and an ARIMA model. Thus, we have achieved our initial goal of integrating relevant text metadata into an existing general time series forecasting model in a way that improves its forecasting accuracy. Our work serves as a proof of concept for this approach and shows it may be a useful avenue of future work. This work could include:

- Improving Metadata: As we discussed, some of the metadata we tried to incorporate did not improve forecasting in its current form. This could be due to small data sizes, noisiness, etc. In the case of noisiness, as was the case for the CDC ILI report data, efforts can be made to better clean the text data and reduce its noisiness. This can be done as a data pre-processing step prior to being integrated into the FM or can be embedded as a step with the FM integration itself. In other cases, as for the NNDSS data on other infectious diseases, the scale of the data available was too large for our case. We did not have the necessary resources to test its integration.

- Additional Metadata: In this study, only a limited number of metadata sources were used. Additional sources could be beneficial. These could include Tweets, more detailed information from healthcare systems, mobility data in various forms, etc. We attempted to find much of this data, but were unsuccessful for a variety of reasons (paywalls, time intensive to collect, lack of access, etc.). Exploring these other forms of metadata could be beneficial.
- Ablation: As mentioned in the milestone comments section, ablation studies to isolate which text metadata had which effect would be very helpful in this case. However, we were unable to do such a study given our time and computational constraints. Ablation studies to understand the most effective metadata sources to use would be a beneficial next step. These studies can be done at a high level for categories of metadata (ie Google search trends versus temperature), but also at a more granular level. For example, now that search trends appear to be an effective predictor, more analysis can be done to understand the most predictive search terms for ILI or another disease.
- Other Base Time-Series Models: We used Chronos, a general time-series FM as our base forecasting model that we augmented with metadata. It is possible that using a similar approach for a different underlying general time-series model would yield even better results. This could include a different FM that we discussed previously in this report, a custom-made FM trained on only disease data, etc.
- Other Diseases: We have only used this model and approach for ILI. It would be useful to see if it extends to other diseases, particularly ones that have significant differences from ILI. First, it would be interesting to see if the same metadata we used on a different disease's time series would yield positive results. Then, it could be analyzed if differing metadata sources more relevant to that disease might be used. A goal could be to develop a set of 'general' metadata that yields positive results for most or all known diseases. Therefore, this set of metadata could be used as the default at the beginning of a new disease outbreak until more information is known.

References

[1] L Alnaji. 2024. Machine Learning in Epidemiology: Neural Networks Forecasting of Monkeypox Cases. (May 2024). https://doi.org/10.1371/journal.pone.0300216

[2] T. Cees G. M. Snoek Amirhossein Habibian, Mensink. 2017. Video2vec Embeddings Recognize Events When Examples Are Scarce. (Oct. 2017). https://ieeexplore.ieee.org/document/7740886

[3] CDC. 2020. *National, Regional, and State Level Outpatient Illness and Viral Surveillance.* https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html

[4] CDC. 2024. *FluView: Past Weekly Influenza Surveillance Reports.* https://www.cdc.gov/fluview/surveillance/past-reports.html

[5] CDC. 2024. *National Notifiable Diseases Surveillance System (NNDSS) Weekly Tables.* https://wonder.cdc.gov/nndss/nndss_weekly_tables_menu.asp

[6] Kong W. Sen R. Zhou Y. Research-G Das, A. 2024. A DECODER-ONLY FOUNDATION MODEL FOR TIME-SERIES FORECASTING A PREPRINT. (April 2024). https://arxiv.org/pdf/2310.10688

[7] Stella L. Turkmen C. Zhang X. Mer-Cado P. Shen H. Shchur O. Rangapuram S. Pineda Arango S. Kapoor S. Zschiegner J. Maddix D. Ma-Honey M. Torkkola K. Wilson A. Bohlke-Schneider M. Wang Y. Fatir Ansari, A. 2024. Chronos: Learning the Language of Time Series. (May 2024). https://arxiv.org/pdf/2403.07815

[8] NOAA National Centers for Environmental Information. 2024. *Climate Data Online (CDO).* https://www.ncei.noaa.gov/cdo-web/

[9] Google. 2024. *Google Trends.* https://trends.google.com/trends/

[10] Szafer K. Choudhry A. Cai Y. Li S.- Dubrawski A Goswami, M. 2024. MOMENT: A Family of Open Time-series Foundation Models. (May 2024). https://arxiv.org/pdf/2402.03885

[11] et al. Guo, W. 2019. Deep Multimodal Representation Learning: A Survey. (2019). https://doi.org/10.1109/access.2019.2916887

[12] et al Igwama, G. 2024. Big data analytics for epidemic forecasting: Policy Frameworks and technical approaches. (July 2024). https://www.researchgate.net/profile/Chukwudi-Maha/publication/382466518_Big_data_analytics_for_epidemic_forecasting_Policy_Frameworks_and_technical_approaches/links/669fa5e4705af5364494bde0/Big-data-analytics-for-epidemic-forecasting-Policy-Frameworks-and-technical-approaches.pdf

[13] Wen H. Nie Y. Jiang Y. Jin M.-Song D. Pan S. Wen Q Liang, Y. 2024. Foundation Models for Time Series Analysis: A Tutorial and Survey. (June 2024). https://arxiv.org/pdf/2403.14735

[14] Boit S. Gudivada V. Nandigam J. Patil, R. 2023. A Survey of Text Representation and Embedding Techniques in NLP. (April 2023). https://ieeexplore.ieee.org/abstract/document/10098736

[15] Kim J. Hallacy C. Ramesh A. Goh G.-Agarwal S. Sastry G. Askell A. Mishkin P. Clark J. Krueger G. Sutskever I Radford, A. 2021. Learning Transferable Visual Models From Natural Language Supervision. (Feb. 2021). https://arxiv.org/pdf/2103.00020

[16] The New York Times. 2024. *Article Search API.* https://developer.nytimes.com/docs/articlesearch-product/1/overview

[17] et al Vaughan, Liam. 2023. An Exploration of Challenges Associated with Machine Learning for Time Series Forecasting of COVID-19 Community Spread Using Wastewater-Based Epidemiological Data. (Jan. 2023). https://doi.org/10.1016/j.scitotenv.2022.159748

[18] David A. Wood. 2023. Weeks-Ahead Epidemiological Predictions of Varicella Cases from Univariate Time Series Data Applying Artificial Intelligence. (Aug. 2023). https://doi.org/10.1097/id9.0000000000000096

[19] Bian Y. Zhong J. Wen X. Xu-Q. Xu, Z. 2024. Beyond Trend and Periodicity: Guiding Time Series Forecasting with Textual Cues. (May 2024). https://doi.org/10.48550/arxiv.2405.13522

[20] Taylor York. 2024. *Pytrends: Unofficial API for Google Trends.* https://pypi.org/project/pytrends/

Received 08 October 2024