*Article*

# Nowcasting Unemployment Using Neural Networks and Multi-Dimensional Google Trends Data

**Andrius Grybauskas** [1,*], **Vaida Pilinkienė** [1], **Mantas Lukauskas** [2], **Alina Stundžienė** [1] **and Jurgita Bruneckienė** [1]

1  School of Economics and Business, Kaunas University of Technology, 44249 Kaunas, Lithuania
2  Department of Applied Mathematics, Faculty of Mathematics and Natural Sciences, Kaunas University of Technology, 44249 Kaunas, Lithuania
*  Correspondence: andrius.grybauskas@ktu.lt

**Abstract:** This article forms an attempt to expand the ability of online search queries to predict initial jobless claims in the United States and further explore the intricacies of Google Trends. In contrast to researchers who used only a small number of search queries or limited themselves to job agency explorations, we incorporated keywords from the following six dimensions of Google Trends searches: job search, benefits, and application; mental health; violence and abuse; leisure search; consumption and lifestyle; and disasters. We also propose the use of keyword optimization, dimension reduction techniques, and long-short memory neural networks to predict future initial claims changes. The findings suggest that including Google Trends keywords from other dimensions than job search leads to the improved forecasting of errors; however, the relationship between jobless claims and specific Google keywords is unstable in relation to time.

## 1. Introduction

Due to the negative effects of the COVID-19 pandemic on the global economy and the rising rates of unemployment, the need for more accurate unemployment forecasting models has become increasingly urgent. Recent developments in the nowcasting literature indicate an attempt to provide an enhanced approximation of the unemployment situation in a timely manner, which might benefit government institutions that are preparing to take pro-active measures against a surge in joblessness. In essence, nowcasting is a technique used to predict the behavior of the economy in real time or as close as possible to real time, allowing governments to make more informed decisions about the stimulus rather than relying on lagging monthly indicators (Giannone et al. 2008; Bańbura et al. 2013). Numerous authors have already reported successes in nowcasting unemployment, car sales, GDP, household consumption with credit card transaction data, foreign arrivals, and even building permits (Choi and Varian 2009a; Banbura et al. 2010; Barreira et al. 2013; Pavlicek and Kristoufek 2015; Rusnák 2016; Coble and Pincheira 2017; Richardson and Mulder 2018; Antolini and Grassini 2018; Nymand-Andersen and Pantelidis 2018; Giovannelli et al. 2020; Aastveit et al. 2020). Although the reported effectiveness of nowcasting models differs from author to author, a general consensus has emerged that nowcasting helps improve prediction accuracy.

Despite the growing literature on nowcasting, much needs to be discussed and perfected regarding the models or variable choices themselves. For instance, one source of data that has been widely popular amongst researchers is user internet search activity, such as Google Trends data (Askitas and Zimmermann 2009; D'Amuri 2009; Choi and Varian 2009b; Chadwick and Sengul 2012; Barreira et al. 2013; Smith 2016; Naccarato et al. 2017; Nymand-Andersen and Pantelidis 2018; Nagao et al. 2019; Anttonen 2018; Dilmaghani 2019; Borup et al. 2021; Mulero and García-Hiernaux 2021). According to Ettredge et al.

(2005), the usefulness of such data is that through search activity, citizens provide useful information about their intentions, needs, wants, and interests. In many cases, people who lose their jobs use Google to help them find unemployment benefit agencies and new job openings, as indicated by the fact that 93.63% of all mobile searches are performed through Google (Netmarketshare 2022). As a result, authors such as Mulero and García-Hiernaux (2021) and Naccarato et al. (2017) claim that by using Google data, an increase in the predictive accuracy of 10–25% or a reduction in forecasting errors was achieved. Another advantage of using Google Trends is that it provides various frequencies, which benefits the nowcasting procedure.

On the other hand, authors, including Nagao et al. (2019) and Barreira et al. (2013), produced mixed results when forecasting unemployment. Google's trend data did not increase accuracy in all countries, and specific preconditions were required for search activity data to work. As Smith (2016) puts it, the challenge with search data is that one must determine which keywords are relevant and how many keywords should be incorporated to best mirror a jobless or soon-to-be-jobless person's search query that later translates to unemployment numbers. In the case of Nagao et al. (2019), only two keywords were used; by expanding the keyword dictionary, the results might have changed substantially. Thus, despite some success stories, the question of how to correctly mine the data from Google Trends is a real challenge that requires considerable future improvements.

Tangent to the keyword selection issue, there is also a lack of attempt to incorporate other socially important keywords that could be highly related to unemployment figures. For example, many authors only included keywords such as "jobs" or "job offers" but did not include other keyword dimensions, such as "online casino", "anxiety", "depression", and "leisure activities". Researchers such as Mousteri et al. (2018) and Liu et al. (2021) found that joblessness and psychological distress are interrelated; thus, the expansion of the keyword corpus could lead to a better Google Trends forecast.

Furthermore, despite recent developments in machine learning, there have been relatively few attempts to use neural networks in nowcasting unemployment. Much of the literature deploys dynamic factor models (DFM), ARIMA, or mixed data sampling (MIDAS) (D'Amuri 2009; Smith 2016; Chernis and Sekkel 2017). However, according to Hopp (2021), the long–short term memory (LSTM) neural network approach has an edge over the DFM in accuracy, speed, and volume. Thus, the machine learning approach could be explored more within the nowcasting context.

Therefore, the aim of this paper is to investigate a multidimensional approach of Google search query data in nowcasting USA initial jobless claim numbers using LSTM neural networks. To our knowledge, this is the first paper that expands Google keyword dimensions to include psychological factors, gambling, and other activities to nowcast initial jobless claims numbers using neural networks. In addition, this paper tests different feature extraction strategies, offering future researchers recommendations regarding what to focus on when conducting keyword selection to avoid common pitfalls. The findings of this study can help central banks and other government institutions expand their variable horizons to achieve better forecast accuracy. The analysis period also covers the COVID-19 timeline.

The paper is divided into four parts. In Section 2, previous authors' experience, keyword selections, and the findings of Google Trends data are reviewed. Additionally, we touch upon novel keyword opportunities that can be mined. Sections 3 and 4 cover data selection, pre-analysis, and the LSTM model architecture. Section 5 then presents the variable co-movement intricacies and Google Trends data modeling results, while Section 6 summarizes the findings. Finally, Section 7 suggests future research directions and discusses the limitations of the study.

## 2. Background

### 2.1. Forecasting Unemployment Using Google Trends

One of the first papers that considered Google Trends for unemployment nowcasting was written by Choi and Varian (2009a). The authors attempted to nowcast initial jobless claims for unemployment benefits by using the keywords "Jobs" and "Welfare & Unemployment". The results revealed that the out-of-sample mean absolute error was reduced by 15.74 and 12.90% for the long- and short-term models, respectively. In the same year, followed papers by D'Amuri (2009) and Askitas and Zimmermann (2009). The work of D'Amuri (2009) limited itself to only one keyword "jobs" whereas Askitas and Zimmermann (2009) used four different categories of keywords: "unemployment office or agency", "unemployment rate", "Personnel Consultant", and keywords that relate to the most popular job search engines in Germany. Both authors concluded that Google Trends helped improve forecasting errors.

As depicted in Table 1, Chadwick and Sengul (2012) nowcasted Türkiye's unemployment rate using keywords that related to "looking for a job", "job announcements", "CV", and "career" and provided supportive evidence for using Google Trends. Fondeur and Karamé (2013) achieved success in forecasting French youth unemployment with only one keyword "job". The RMSE was improved on average by 13.6%. Several recent papers are still using a limited range of keywords. For instance, Aaronson et al. (2022), Maas (2019), and Larson and Sinclair (2021) employed a single search term ("unemployment"), D'Amuri and Marcucci (2017) used "jobs", whereas Simionescu (2020), Simionescua and Cifuentes-Faura (2021) used two terms "unemployment" and "job offers". Despite the use of single or double keywords, Google Trends improved accuracy.

**Table 1.** A summary of Google keywords used for unemployment forecasts by previous authors.

| Authors | Keywords Used |
| --- | --- |
| Choi and Varian (2009a) | "Jobs", "Welfare & Unemployment" |
| D'Amuri (2009); Maas (2019); Larson and Sinclair (2021) | "jobs" |
| Askitas and Zimmermann (2009) | "unemployment office or agency", "unemployment rate", "Personnel Consultant", and keywords that relate to the most popular job search engines in Germany. |
| Chadwick and Sengul (2012) | looking for a job", "job announcements", "CV", and "career" |
| Fondeur and Karamé (2013) | "job". |
| Simionescu (2020); Simionescua and Cifuentes-Faura (2021) | "unemployment" and "job offers". |
| Pavlicek and Kristoufek (2015) | "work" or "jobs" |
| Nagao et al. (2019) | jobs" and "job offer". |
| Barreira et al. (2013) | "unemployment" and "unemployment benefits" |
| Schiavoni et al. (2021) | 85 keywords that strongly relate to the job process (e.g., CV, cover letter, job vacancies). |
| Yi et al. (2021) | 25 work or job-related keywords |
| Caperna et al. (2020) | Many work related keywords. |

On the other hand, some researchers claim to have achieved mixed results. Pavlicek and Kristoufek (2015) used the keywords "work" or "jobs" for Visegrad countries in multiple languages. The authors concluded that for Hungary and Slovakia, user search trails can be integrated easily; however, for Poland and Slovakia, the results were not consistent. Similar findings were achieved by Nagao et al. (2019), who limited themselves to only two terms: "jobs" and "job offer". The authors found no consistency regarding improved accuracy, particularly when considering the long-term nowcasts, whereas Barreira et al. (2013),

using "unemployment" and "unemployment benefits", found an improvement in three out of the four countries analyzed.

Unfortunately, few studies have attempted to incorporate a higher query volume. Among these papers is one by Schiavoni et al. (2021), who studied the Netherlands' job market. The authors included 85 keyword search terms but limited the query to words that strongly relate to the job process (e.g., CV, cover letter, job vacancies). A slightly different approach was adopted by Caperna et al. (2020) and Yi et al. (2021), who suggested not only examining the "unemployment" keyword but also checking the most-searched terms and, when filtering them, determining which related to work and jobs. Yi et al. (2021) obtained 25 keywords for the USA region and successfully integrated them into the forecasting model. In the case of Caperna et al. (2020), the search term queries were extremely varied; for example, for Estonia, 3 keywords were chosen and 178 for Italy, whereas for some countries, such as Luxemburg or Malta, none were found. The results of Google Trends benefits for Caperna et al. (2020) also varied, and further methods were required to find significant relationships.

Overall, the previous research suggests that authors tended to use one or several keywords to evaluate the forecasting benefits of the GTI index, with few exceptions. The keyword domain was mostly limited to job market processes, and success demonstrated a degree of variability that depended on factors such as the location, region, chosen keywords, and time horizon.

### 2.2. Additional Keyword Opportunities

The analysis of the existing studies reveals that the authors limited themselves to one or two basic keywords, and the expansion of keywords was highly constrained to the job application procedure. This approach led some researchers to conclude that Google Trends offered no additional benefit to nowcasting models or that trends are inconsistent over longer periods. However, the pitfalls or further improvement of the models could be partially attributed to Google Trends keyword mining process. By only considering one keyword for forecasting, the authors cast aside the big picture and the nuances of behavioral economics that are profoundly related to unemployment numbers. For instance, Liu et al. (2021) and Wanberg et al. (2019) found strong links between unemployment and mental health. In times of downturn, the fear of job loss can trigger mental distress, particularly when loss of employment benefits is taken into account. The study by Breuer (2015) argued that unemployment increases the risk of suicide, while Butterworth et al. (2012) found that poor mental health leads to a longer duration of unemployment for both men and women. Furthermore, studies by Sotis (2021) and Askitas (2015) found that when labor market conditions deteriorate, searches for health symptoms proliferate; therefore, a dimension of mental health keywords (MHKs) could provide substantial benefits to model precision. Unfortunately, to our knowledge, no study is available for nowcasting unemployment that includes MHKs.

Furthermore, a contentious issue among labor economists has been to analyze the trade-offs between leisure and labor. More importantly, Havitz et al. (2004) and Goodman et al. (2016) claim that unemployed people spend more time in leisure to mitigate the stress of joblessness. In the Internet age, leisure time is frequently spent consuming free entertainment, such as YouTube, online video games, and downloading movie torrents via sites such as "The Pirate Bay" (Lehdonvirta 2013). A highly regarded study in this area was undertaken by Dilmaghani (2019) that included torrent websites as a proxy for leisure activities and managed to nowcast unemployment numbers more accurately. However, leisure-related keywords still have more to offer. For instance, Frangos et al. (2011) discovered that people involved in unemployment programs were more likely to develop a problematic pornography habit. This is because, according to Uzieblo and Prescott (2020), pornography may act as a stress reliever in times of high anxiety. Similarly, some kind of association between unemployment and gambling persists. Khanthavit (2021) found that unemployment and gambling have a circular relationship, while Pallesen et al. (2021) and

Mallorquí-Bagué et al. (2017) determined that internet gaming disorder was more prominent among single unemployed persons. Thus, an increase in online gaming, gambling platforms, and sites, such as "OnlyFans", search activity could help to improve Google Trends mining process.

Another negative behavior associated with unemployment is domestic abuse (Anderberg et al. 2016; Bhalotra et al. 2021). According to Anderberg et al. (2016), women who are at risk of job loss become more economically reliant on their partners. This reliance can then lead male partners who have a predisposition to violence to reveal their abusive tendencies. Thus, high female unemployment leads to an elevated risk of intimate partner violence. As a way out, people may use Google searches to find public shelter. A study by Berniell and Facchini (2021) reported a high Google search intensity in 11 countries for keywords such as "domestic abuse" or "domestic violence hotline" during the COVID-19 lockdowns. The strong correlation between unemployment and domestic abuse might also assist nowcasting.

Many other important keywords should also be incorporated into the nowcasting model. Some of these include "gig economy", "recipes", "hobbies", "Netflix", "payday loans", lottery tickets", "alcohol", "layoffs", "auto insurance", "hardship letter", or words related to comfort food (Dávalos et al. 2011; Askitas 2015; Agarwal et al. 2016; Chadi and Hetschko 2017; Huang et al. 2018; Gabrielyan and Just 2020; McKinsey & Company 2020; Gamze and Aydan 2021). However, due to word limitations, this study will only provide references to how particular words might be related to unemployment.

*2.3. Nowcasting and Machine Learning*

Previous studies used a wide range of methods to nowcast unemployment numbers. Mulero and García-Hiernaux (2021), Larson and Sinclair (2021), Simionescu (2020), Caperna et al. (2020), Nagao et al. (2019), Dilmaghani (2019), D'Amuri and Marcucci (2017), Pavlicek and Kristoufek (2015), Chadwick and Sengul (2012) and Choi and Varian (2009a) used some form of autoregressive models. Yi et al. (2021) suggested a PRISM approach, and Schiavoni et al. (2021) incorporated a well-known model of DFM, whereas Maas (2019) attempted the MIDAS model. Although each model has its own strengths and weaknesses, the recent developments in neural networks deserve deeper exploration. As stated by Brownlee (2018) and Bruneckiene et al. (2021), neural networks may provide a solution for capturing the non-linear patterns in the data that can be difficult to detect using conventional econometrical methods as well as offering the ability to use a higher volume of data. Unfortunately, few studies have attempted to forecast unemployment numbers using Google Trends in conjunction with LSTMs. Fenga and Son-Turan (2022) used a feed-forward neural network for counterfactual predictions, whereas Singhania and Kundu (2020) deployed the LSTM model for monthly data. According to Singhania and Kundu (2020), the authors attempted, at first, to use the VAR model but were soon faced with a major drawback as the VAR model can only be used to capture the relationship between search trends of a limited number of keywords due to growth in the complexity and parameters. As such, the authors attempted to use LSTMs as an alternative. In the end, the LSTMs significantly outperformed the VAR model but were not used for nowcasting. The latter results could be a good indication to use the LSTM model in this paper as it can help to deal with a large number of keywords in an efficient manner.

**3. Data**

Google Trends data do not provide a sheer volume of searches. Rather, they adapt a specific calculation methodology that, according to Mulero and García-Hiernaux (2021), can be expressed mathematically as:

$$RP_{(q,l,t)} = \frac{n(q,l,t)}{\sum_{q \in Q(l,t)} n(q,l,t)} \times \Pi_{(n(q,l,t) > \tau)} \tag{1}$$

where RP is relative popularity; n(q,l,t) is the number of searches for specific word q, in the location l during period t; Q(l,t) is the number of queries made from l during time t; and $\Pi_{(n(q,l,t)>\tau)}$ is made as a dummy variable, which takes the value 1 if the query is sufficiently popular (more precisely, if the absolute value of the query made n(q,l,t) exceeds $\tau$, otherwise the value is zero). The obtained number is then scaled in a range of 0 to 100 according to the proportion of a topic with respect to the total number of all search topics. Thus, the final Google Trends index is calculated as follows:

$$\text{IGT}_{(q,l,t)} = \frac{\text{RP}(q,l,t)}{\max\left\{\text{RP}\left(q,l,t\right)_{t\in 1,2,\ldots,T}\right\}} \times 100 \qquad (2)$$

where T is the last value of t.

To enhance Google Trends mining process, a total of 222 keywords were carefully selected and classified into six keyword dimensions (see Figure 1):

(a)   Job search, benefits, and application (JSBA);
(b)   Mental health (MHK);
(c)   Violence and abuse (VA);
(d)   Leisure search (LA);
(e)   Consumption and lifestyle (CL);
(f)   Disasters, war, and viruses (DWV)

The latter dimensions were developed in conjunction with the previous literature analysis findings, as well as with an effort to incorporate other unexplored territories that could assist the nowcasting model. The chosen variables constitute human behavior that relates to unemployment directly or indirectly and touches upon many new unexplored territories. The largest dimension with 97 keywords is the lifestyle section, which heavily focuses on luxury queries, such as yachts for sales, cosmetic and plastic surgery, lavish bags, wedding products, service activities, real estate searches, lottery tickets, payday loans, or other housing servicing amenities. Due to job losses and a shortage of cash, people may need to downgrade their lifestyle; hence, specific trends may be captured with the latter keywords (Reyneke et al. 2012; Chiaroni and Kaplan 2016). At the same time, to cope with unemployment effects, women might begin to use their bodies for economic gain (i.e., prostitution, escort services, and sugar baby dating), and these queries have also been included (Palomeque Recio 2021).

Mental health was the second-largest dimension, with 50 keywords that followed the recommendations of the literature findings in Section 2. Some popular medication brands were also included because many people search for the side effects of each medicine. The job search dimension contained 35 keywords. As in previous studies, job dimensions enveloped the job application process, the most popular employment websites, unemployment benefits, and general job-related queries. Meanwhile, the leisure dimension contained 25 keywords. In contrast to other authors who included only torrent services, pornographic consumption proxied by the most popular porn websites search, gambling websites, gaming platforms, and game titles were composed. As stated in the literature analysis, leisure activities could become a stress-coping mechanism for the unemployed (Pallesen et al. 2021; Dilmaghani 2019; Mallorquí-Bagué et al. 2017; Frangos et al. 2011). Lastly, abuse and turning point dimensions included 10 and 5 keywords, respectively. The abuse simply captured the most used keywords in the Google platform at the time of domestic situations, whereas turning points signal major catastrophic events, such as viral pandemics. Many of these dimensions have yet to be explored by other researchers; therefore, this research could become a reference point for future papers.

After selecting the set of keywords, a Python script was created to retrieve the data from the Google Trends website. As our intention was to nowcast unemployment, the longest weekly time series available was chosen, limiting the dataset to March 2017–March 2022. Before downloading, each keyword was inspected to eliminate time series that had

only zero values. Hence, the majority of keywords retrieved had values bigger than zero, meaning that the targeted keywords were relatively popular among Americans.
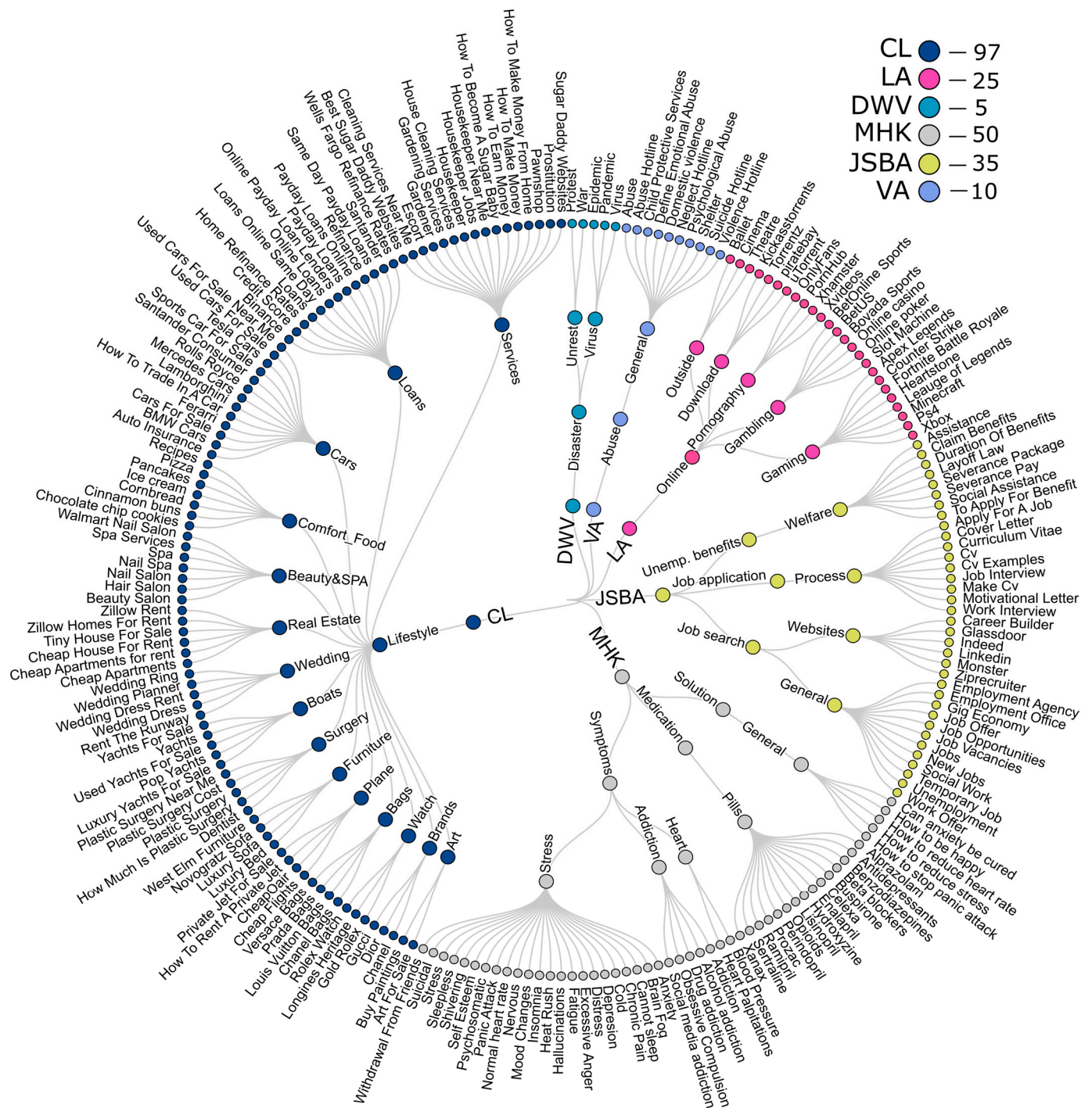


**Figure 1.** Multiple keyword dimensions of Google Trends.

Furthermore, the leading indicator of initial jobless claims (ICs) for the USA was retrieved from the FRED website, which is released by the Bureau of Labor Statistics. The IC indicator is released on a weekly basis every Thursday (D'Amuri and Marcucci 2017). Due to the extremely large number of variables, the descriptive statistics for all the variables are provided as supplementary material.

## 4. Methodology

### 4.1. Feature Selection and Dimension Reduction

Variables were tested for stationarity using the augmented Dickey–Fuller test and were differenced as follows:

$$Y'_t = Y_t - Y_{t-1} \tag{3}$$

where $Y'_t$ is the differenced data, $Y_t$ is the data point at time t, and $Y_{t-1}$ is the data point in the previous period (Hyndman and Athanasopoulos 2018). For ADF results, please see Supplementary Material Table S6.

Before becoming involved in machine learning, it was important to explore the dataset because the insights along the way could guide the model creation process. As such, for pre-analysis, the Pearson correlation coefficient was used with the following formula:

$$r = \frac{Cov(x, y_{IC})}{S_{IC} S_x} \tag{4}$$

where $y$, $S_{IC}$ is the sample standard deviation of USA initial claims, whereas $x$ and $S_x$ are iterations of Google Trends index variables. Analysis of correlation would allow us to understand which variables had the largest co-movements in the analyzed period.

To avoid the common pitfalls[1] associated with Pearson's correlation coefficient, the maximal information coefficient (MIC) was incorporated, which has the following formula:

$$\text{MIC}(x, y) = \max\{I(x, y) / \log_2 \min\{n_x n_y\}\} \tag{5}$$

where I is mutual information, and $n_x$ $n_y$ are bins. In other words, MIC is the mutual information among the random variables X and Y normalized by their minimum joint entropy. The advantages of using MIC are the following: it can capture a wide range of linear as well as non-linear relationships, it does not make assumptions about the distribution of variables, and it is robust to outliers. However, because the coefficient values range from 0 to 1 (zero corresponding to no relationship and 1 highly correlated), the MIC does not provide directional information between two variables (Reshef et al. 2011; Kinney and Atwal 2014).

Furthermore, as a mathematical rule, models cannot handle more variables than there are data points. Thus, feature selection or dimension reduction techniques must take place. Features can be chosen according to the highest correlation coefficients or by optimizing their predictive accuracy, but an alternative approach is to reduce dimensions of the data while minimizing the information loss. One way to achieve this aim is to deploy principal component analysis (PCA). The main focus of PCA involves finding the eigenvectors and eigenvalues of the covariance matrix from the demeaned original data:

$$\text{AX} = \lambda X \tag{6}$$

where A is the covariance matrix, X is the eigenvector, and $\lambda$ is the eigenvalue. After that, eigenvectors are selected according to the highest eigenvalues because this strategy captures the most variance, and the original data are multiplied by the eigenvector. The first principal component captures the largest variance in a particular direction; similarly, the second principle component delivers the second-highest variance in a specific direction as well as being uncorrelated with the first principal component. Readers who are unfamiliar with PCA can refer to Jolliffe (2002) or Chadwick and Sengul (2012).

### 4.2. The Nowcasting Model

Nowcasting models differ from ordinary forecasting because rather than attempting to predict the future, they endeavor to predict the present. Thus, variable timestamps play an important role (Choi and Varian 2009a). As shown in Figure 2, the setup essentially considers finding Google Trends timestamps that come with "ex-ante" IC claims release. The weekly data of Google Trends indexes (GTI) come six days before the initial claims;

thus, the start of the period analyses begins with week 1, where GTI starts on 2017/03/19, whereas IC claims begin on 2017/03/25. This specification is followed consecutively until the end.
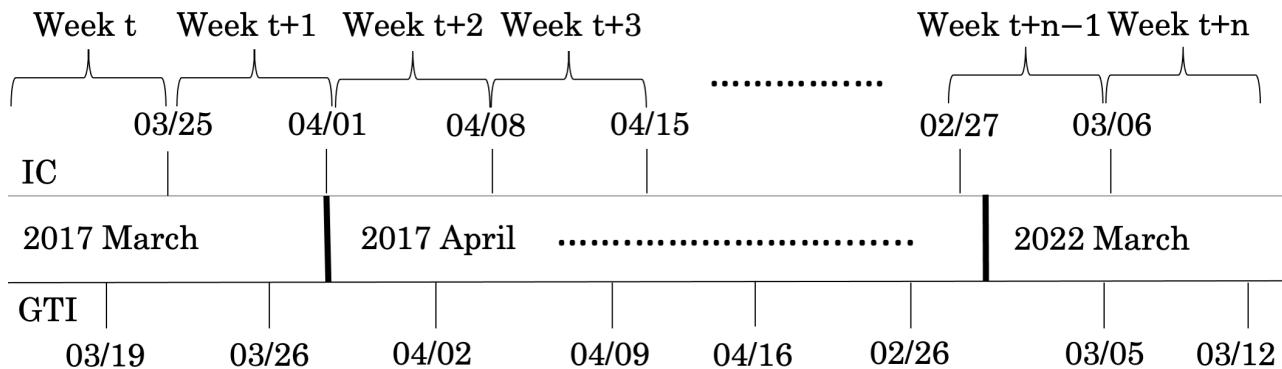


**Figure 2.** Timestamp setup of IC and Google Trends index variables.

With regard to the modeling approach, the DFM model has been extensively used in many research papers. However, the UN researcher Hopp (2021), who compared the DFM to the LSTM in GDP forecasting situations, found that speed- and accuracy-wise, the latter delivered better nowcasting results; thus, it was decided to explore the LSTM architecture in this paper.

The LSTM model was first introduced by Schmidhuber and Hochreiter (1997) and comes from a family of recurrent neural networks (RNNs). The following formulas describe the inner workings in detail:

$$f_t = \sigma_g(W_f \times x_t + U_f \times h_{t-1} + b_f) \tag{7}$$

$$i_t = \sigma_g(W_i \times x_t + U_f \times h_{t-1} + b_i) \tag{8}$$

$$o_t = \sigma_g(W_o \times x_t + U_o \times h_{t-1} + b_o) \tag{9}$$

$$\acute{c}_t = \sigma_c(W_c \times x_t + U_c \times h_{t-1} + b_c) \tag{10}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \acute{c}_t \tag{11}$$

$$h_t = o_t \cdot \sigma_c(c_t) \tag{12}$$

$$\sigma_g = \frac{1}{1 + e^{-x}} \tag{13}$$

$$\sigma_c = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{14}$$

where $f_t$ denotes the forget gate, $i_t$ is the input gate, $o_t$ denotes the output gate, $c_t$ is the cell state, $h_t$ is the hidden cell state, $\sigma_g$ is the sigmoid activation function, and $\sigma_c$ is the tanh activation function, while W and U are weights. In essence, when the input data are fed into the LSTM model, the input gate determines which information from the input sequence should be stored in the memory cell. This is undertaken by multiplying the input vector by $W_f$ and adding it to weighted hidden state plus the bias. The forget gate decides which information in the memory cell should be kept or discarded, based on the current input and the previous memory state. It also has its own weights and multiplication procedure and outputs a value between 0 and 1 which, when multiplied, will determine which information

to forget. The output gate with a similar weighted multiplication procedure determines which information should be output from the current memory state. The LSTM model also has a mechanism called "cell state", which allows it to retain long-term dependencies in the input sequence. This is achieved by selectively updating the memory cell through a combination of the input, forget, and output gates. The cell state can be thought of as the long-term memory of the LSTM model. It is important to note that · in Equations (11) and (12) denotes the Hadamard product so the hidden cell $h_t$ and cell state $c_t$ is achieved by conducting the Hadamard product between input, forget, output and cell state values (Singhania and Kundu 2020). For a better visual representation, please consult Figure 3.
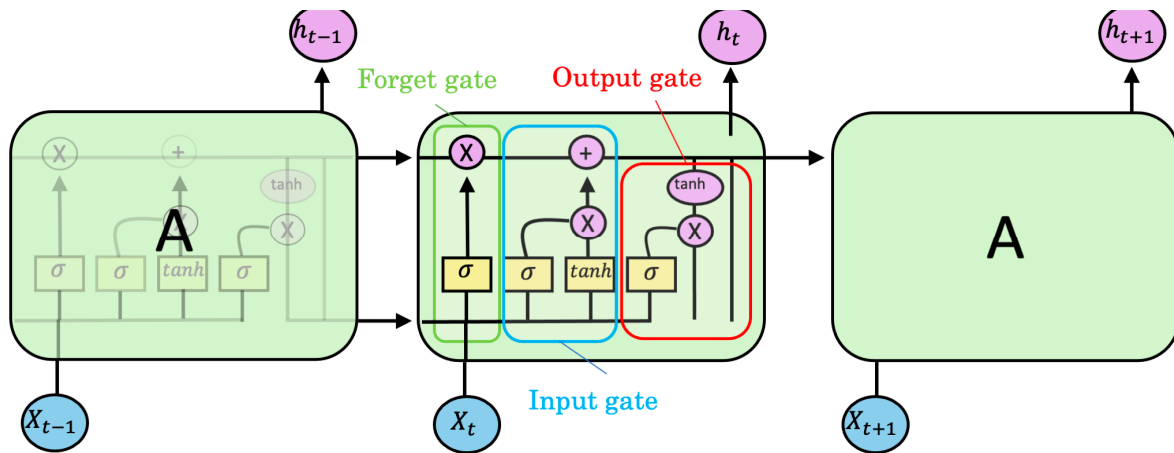


**Figure 3.** A repeating module in LSTM unit.

The LSTM models can also be multi-layered meaning that the same procedure that is depicted in Figure 3 is undertaken multiple times and multiple cell and hidden cell states are calculated. In this paper's case, it was decided to construct two LSTM variations: the vanilla LSTM and encoder–decoder LSTM model (both of which are depicted in Figure 4). The vanilla LSTM has only one layer of LSTM followed by a dense layer. The dense layer is a fully connected layer that maps the output of LSTM to the desired shape and can be described using the following formula:

$$y = Wx + b \tag{15}$$

where $y$ is the output, $W$ are the weights for the dense layer, $x$ is the input coming from the previous layer, and $b$ is the bias.
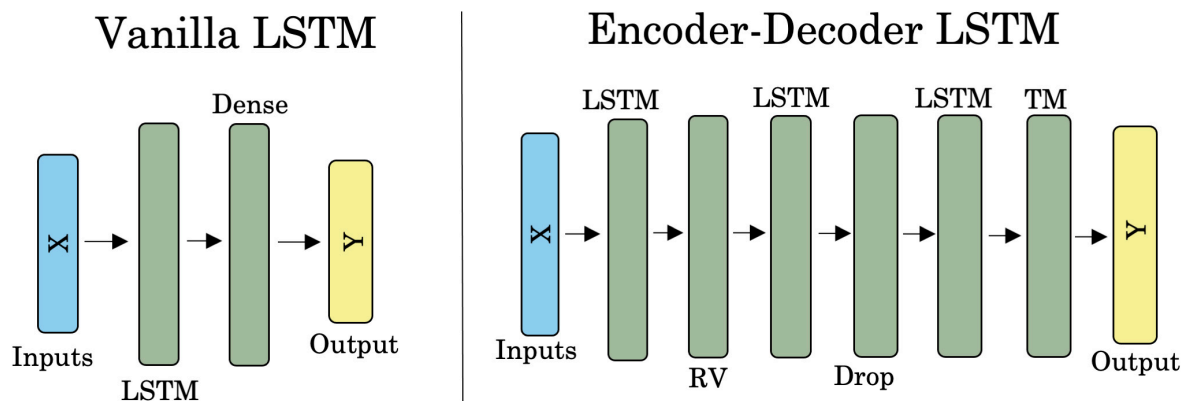


**Figure 4.** LSTM model structure. Note: RV stands for repeat vector, and TM stands for time-distributed layer, while Dense is a fully connected layer.

The vanilla LSTM model does a good job of capturing time-series dynamics; however, a recent paper by Laptev et al. (2017) introduced an auto-encoder layering for extreme event forecasting on Uber data. As our timeframe included turbulence due to pandemic, it was decided to adopt a similar approach with specific modifications where different LSTM layers were used to encode and decode data, which helps to accommodate outlier effects. Some differences in the encoder–decoder model depicted in Figure 4 are that it uses three LSTM layers, where the first layer has 50 units (dimensionality of the output space), whereas the second LSTM layer has 8 units and last LSTM layer has 50 units again. This was undertaken so that the neural network would recreate the most important parts of the time series avoiding noise. The drop layer stands for a dropout layer, which is a regularization technique in neural networks that randomly drops out (or "turns off") a fraction of the neurons in the layer during training. This helps to prevent overfitting. Lastly, the time-distributed layer applies a layer to every timestep of a sequence independently as if it was processing a single input at a time.

With regard to variable specification, the following models were constructed:

$$IC_t = F(IC_{t-1}) \tag{16}$$

$$IC_t = F(IC_{t-1}; GTI_{t,Unemp,TF}) \tag{17a}$$

$$IC_t = F(IC_{t-1}; GTI_{t,v\_iter,TF}) \tag{17b}$$

$$IC_t = F(IC_{t-1}; GTI_{t,iter,TF} = \begin{pmatrix} x_{t,unemployment,TF} \\ x_{t,v\_iter,TF} \\ \vdots \\ \vdots \\ x_{t,v\_iter}, TF \end{pmatrix}) \tag{18a}$$

$$IC_t = F(IC_{t-1}; GTI_{t,iter,TF} = \begin{pmatrix} x_{t,v\_iter,TF} \\ x_{t,v\_iter,TF} \\ \vdots \\ \vdots \\ x_{t,v\_iter}, TF \end{pmatrix}) \tag{18b}$$

$$IC_t = F(IC_{t-1}; GTI_{t,JSBA,TF}) \tag{19}$$

$$IC_t = F(IC_{t-1}; X = \begin{pmatrix} X\_PCA_{t,first,JSBA,TF} \\ X\_PCA_{t,first,MHK,TF} \\ \vdots \\ \vdots \\ x \end{pmatrix}) \tag{20}$$

$$IC_t = F(IC_{t-1}; X = \begin{pmatrix} X\_PCA_{t,first,JSBA\_SUBS,TF} \\ X\_PCA_{t,first,MHK\_SUBS,TF} \\ \vdots \\ \vdots \\ x \end{pmatrix}) \tag{21}$$

where F is a function that maps inputs to outputs (in this paper's case, the different LSTM models); t symbolizes timesteps; TF denotes timeframes; iter refers to the optimization algorithm, where v_iter (variable iteration) is replaced in search of a better one; $X\_PCA_{first}$ denotes the first principal component that had the largest eigenvalue, whereas $X\_PCA_{JSBA\_SUBS}$ refers to the 27 sub-dimensions of each category as depicted in Figure 1;

JSBA and MHK stand for the 5 root dimensions described in Section 3, and $IC_t$ is the dependent variable.

The logic behind the variable specification is as follows: Equation (16) creates a persistence model, which is the benchmark that must be beaten. It is one of the simplest models and attempts to predict the next timestep value only by considering the lagged value of the dependent variable. Afterward, because many authors achieved success by only using the "unemployment" keyword, Equation (17a) attempts to replicate previous success stories with the latter keyword (Maas 2019; Larson and Sinclair 2021). However, a question lingers regarding whether the "Unemployment" term is the best keyword to continue building the model forward; thus, in Equation (17b), it was decided to construct an optimization algorithm that iterates through each keyword and all timeframes and retrieves the best RMSE[2] score. Model (18a) optimizes forward on the pre-specified "Unemployment" keyword in model (17a), whereas model (18b) optimizes forward on model (17b) keyword. Finally, Equations (19)–(21) attempt to extract the most important variance using PCA dimension reduction techniques. This reduction is undertaken to the JSBA, the 6 major GTI dimensions, and additionally to 27 sub-dimensions to avoid the dimensionality reduction side effects, such as reducing and eliminating excessive information from the data, thus losing granularity. If adding GTI variables does not improve the final prediction score, the GTI index does not then carry valuable information for the prediction process (D'Amuri and Marcucci 2017; Chadwick and Sengul 2012; Mulero and García-Hiernaux 2021).

All variables prior to modeling were scaled to the range of $(-1, 1)$ using the scikit-learn min–max scaler library:

$$z(x) = \frac{2 \times (x - \min(x))}{(\max(x) - \min(x)) - 1} \tag{22}$$

where $x$ is the original value, $\min(x)$ is the minimum value of $x$ in the dataset, and $\max(x)$ is the maximum value of $x$ in the dataset.

Lastly, the model timeframes were divided into three categories: full, pre-pandemic, and pandemic timeframes.[3] This division was carried out intentionally to test how each feature variables and models perform under economic growth and volatile conditions and was tested on all model and variable specifications. The model prediction performance was for 20 timesteps of the test data, and these steps were split from the original dataset. For model implementation, we used the Python Keras framework.

## 5. Empirical Results

The initial screening of the variables in the pre-analysis immediately revealed a challenging modeling problem. After differencing the IC variable in the pre-pandemic timeframe, the time series was mostly stationary. However, on 21 March 2020, due to COVID-19, the surge in jobless claims quickly went from an average of 225,000 to approximately 6 million and then swiftly returned to its original 248,000 average (see Figure 5, sides A and B). Hence, the pandemic timeframe is an outlier, or the so-called extreme event forecasting problem, which is similar to something that Laptev et al. (2017) encountered in their UBER data during the holiday surge.

Similarly, particular search queries, such as "Unemployment", "Online poker", "Online casino", "Severance Pay", "Layoff Law", and "To Apply For Benefit", exhibited low or stable search volumes ("Unemployment", "Online poker", and "Online casino" had a mean of 4–8 index points in the pre-pandemic period) and when dramatically rushed alongside the IC variable during the pandemic timeframe, neared the ceiling of 100 index points, as shown in Figure 5, side C. One must point out a rather important observation that the surge in terms such as "Unemployment" and "Severance Pay" occurred prior to the IC variable surge. Similarly, virus-related search terms were found to surge before hospital-registered virus outbreaks (Ginsberg et al. 2009). Hence, this situation can be a powerful predictor for the nowcasting model.
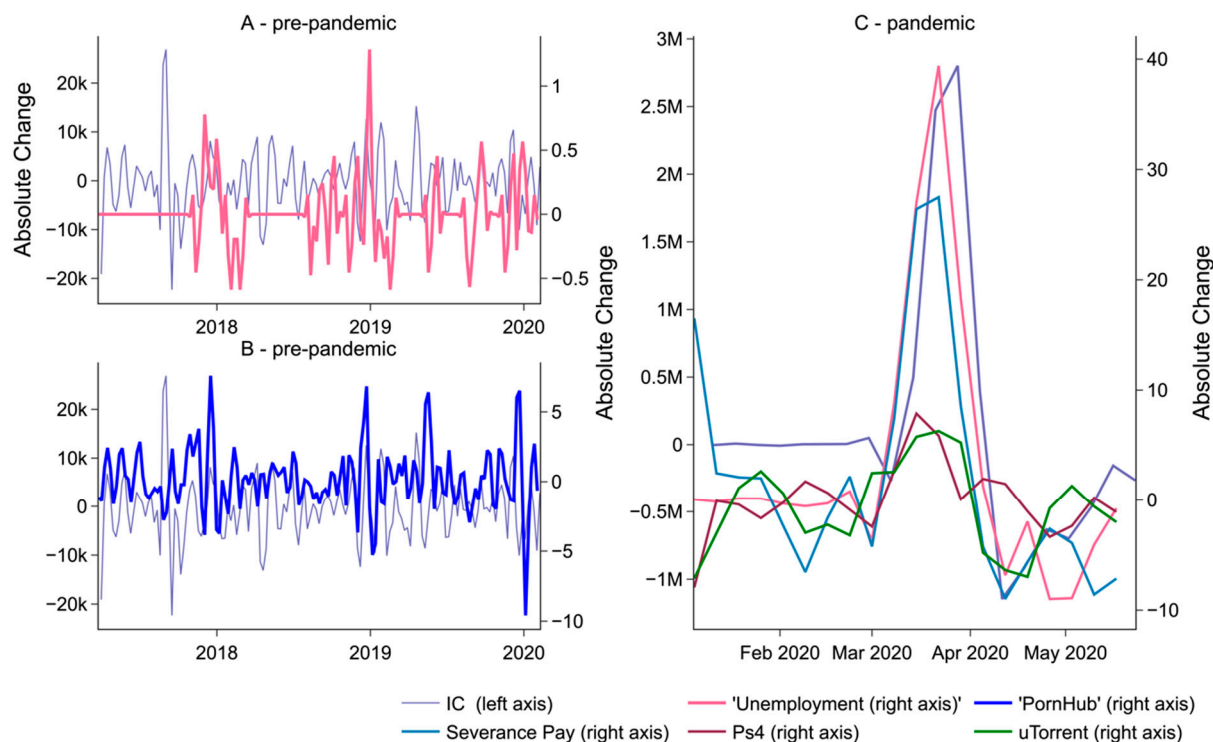
**Figure 5.** Side (**A**,**B**) depicts the pre-pandemic timeframe of differenced time series for the IC, "PornHub", and" Unemployment" variables. Side (**C**) depicts the pandemic timeframe, with differenced time series of IC, "Ps4", "Severance Pay", and "Unemployment". Note: due to the large number of keywords, only a few have been depicted; for more detailed information about all variables, see Supplementary Material 1 and 2 as well as Figure 6.

Meanwhile, addictive behavior searches, which include terms such as "Online Casinos", "Online Poker", and "PornHub" and searches related to specific stress symptoms, such as "Panic Attack" and "Distress", also increased prior to the jobless claims surge. Again, this result could indicate stress mitigation activity or stress-induced symptoms, whereby an employee already anticipates an upcoming lay-off and searches for ways to overcome difficult times. Searches for specific medications, such as "Lisinopril", were found to increase weeks or, in the case of "Alprazolam", months before.

By the same token, an inverse surge relationship was detected among variables such as "Cinema", "Theatre", "Dentist", "Nail Spa", "Cheap Flights", and "Spa Services" in the pre-pandemic period, with an average of 50–74, which rapidly decreased to near 0–8 and bounced back to their normal position (for more variable inverse surges, please refer to Figure 6, where a negative correlation is depicted). Perhaps such behavior originated from the COVID-19 lockdowns, which forbade people to participate in leisure activities and, as a consequence, search queries for tickets or other events decreased. Similarly, for the "Unemployment" keyword, many of the inverse surge variables were reduced prior to the job claims surge.

Somewhat analogous dynamics were also detected using the Pearson and MIC coefficients, which were calculated for all three timeframes and are depicted in Figure 6. First, by analyzing the pre-pandemic period, one can immediately observe that out of 222 keywords, only 10 achieved a Pearson correlation coefficient higher than |0.2| and only 23 for MIC, representing only 4 and 10 percent, respectively. However, of the latter keywords, none surpassed the |0.3| Pearson or MIC mark, meaning that the co-movement was negligible at best. It seems that during the stable economic growth periods with unemployment close to the natural unemployment rate, the keywords become heterogeneous in their movements from the IC numbers. Even though most of the search queries were relatively active, with the mean concentrating at around 50 index points, it is probable that the searches were not

solely motivated by attempts to find jobs, and people's change of professions did not result in widespread stress. The situation was found to be similar for the full timeframe, where only 12 for Pearson and 7 for MIC bypassed the |0.2| mark but remained well below the |0.3| mark.
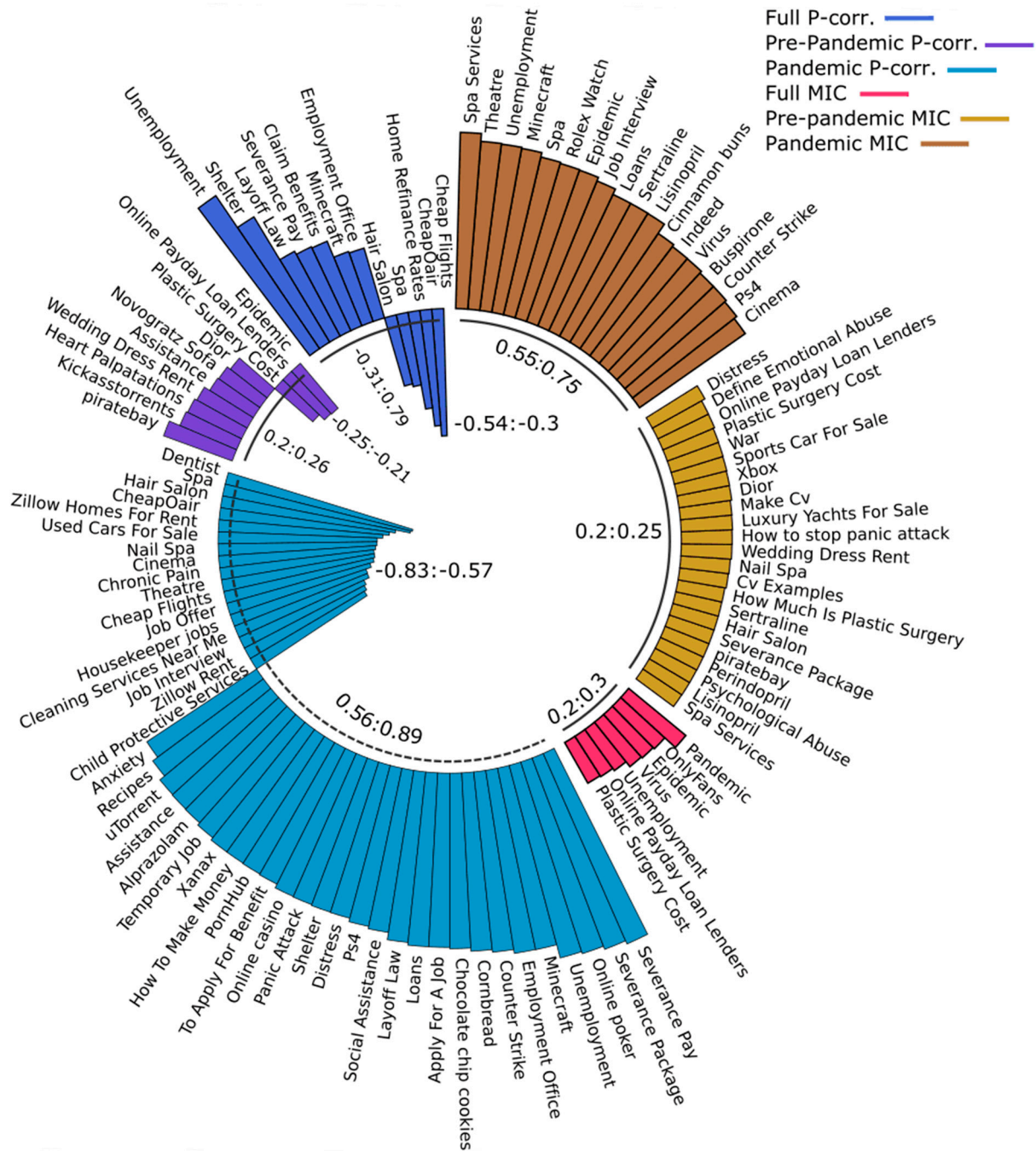


**Figure 6.** Pearson and MIC coefficients of GTI indexes for three timeframes. IC was held as the dependent variable.

Nonetheless, correlations drastically change if one considers only the pandemic time-frame. As shown in Figure 6, the number of Pearson coefficients that were greater than |0.56| increased to 45 and 18 for MIC. Specific variables, such as "Severance Pay", "Online Poker", and "Unemployment", even reached a Pearson correlation of approximately 0.8, suggesting a strong relationship that coincides with the surge behavior analyzed in the previous paragraphs. Some interesting discrepancies were also observed when comparing

Pearson with MIC coefficients. For instance, in the full-time frame analysis, the unemployment Pearson coefficient, as depicted in Supplementary Material in Table S7, showed a high correlation, while MIC was rather small. For further analysis, we depicted scatter plots in Supplementary Material to see why that might be the case. For the unemployment keyword case, several outliers were recorded, which could have contributed to the MIC becoming small. Furthermore, certain variables in the pandemic timeframe had a small Pearson coefficient, but their MIC values increased significantly, e.g., in the "Epidemic" keyword case, the Pearson coefficient was recorded at −0.28, while MIC increased to 0.69. When examining the scatter plot in Supplementary Material in Figure S1, part B, it would seem that large fluctuations could have significantly contributed to such MCI results, although visually the pattern is not clear.

Despite the increase in co-movement, one must be careful when interpreting these variables because many of them do not have a causal relationship but rather are likely to have some sort of confounding factor that relates them to the IC variable or, in a worst-case scenario, are merely a byproduct of randomness. For instance, high correlations with IC and a rapid surge to ~90 index points among particular games, such as "Counter-strike" or "Minecraft", might have originated from young children staying at home during the pandemic timeframe and from taking part in remote learning, although the number of adults playing video games in recent decades has increased, and this is not an uncommon leisure activity among the adult population; hence, this rise could also indicate a stress-coping mechanism as well. Another issue with interpretation is that one must be careful to avoid the trap of brand explosions (e.g., "OnlyFans"). Although the surge occurred within a similar timeframe as IC, the "OnlyFans" keyword surge could be related more to product popularity and extreme growth of a startup rather than unemployment numbers. On the other hand, it could be considered a similar addictive coping mechanism. Contrary to Dilmaghani (2019), the torrent website names in our paper showed little significance in co-movement; however, the uTorrent keyword, which is software used to download pirated files, demonstrated a high correlation.

Lastly, the modeling results are reported in Table 2 according to the outlined methodology. By looking at different time periods, it is immediately obvious that the COVID-19 pandemic caused some noise to appear in the model as the errors increased significantly when comparing pre-pandemic with pandemic timeframes. The error increase can be explained by severe shutdowns that required borders, shops, events, and factories to close, and, as a consequence, the usual consumer activities were limited. The latter resulted in massive layoffs between industries and a sudden decline in GDP growth. Therefore, the performance of the model during the pandemic period should not be assessed solely based on its parameters but should also be viewed in the context of the COVID-19 disruption.

Model 16, which reports the persistence model, is considered the benchmark model to beat. The next-in-line model (17a) is used as a single GTI keyword "Unemployment". As one can see, the latter model had difficulty beating the persistence scores in the full and pandemic timeframes and was off by 3.8 and 43 percent, respectively, although it offered minuscule improvements for the pre-pandemic case of 0.4 percent in the case of the vanilla LSTM model. As a counter to that, we attempted to find an alternative keyword before proceeding with building model extensions using the optimization iteration procedure. However, the challenge was to deal with the number of permutations (e.g., optimization can be considered along with timeframes, different hyper-parameters or model layers, and keyword combinations). Thus, for model (18a), one would have $3 \times 10 \times$ (222 keywords in position one $\times$ 221 keywords in position two), amounting to 1,471,860 model routes just for a two-GTI keyword model. Due to computational resource limitations and according to the findings in the pre-analysis, we decided to optimize models (17b), (18a), and (18b) only, according to vanilla LSTM layering and all timeframes.

Beginning with the optimization for the pandemic period, the optimized model (17b) produced promising results with 14 variables, resulting in pandemic RMSE scores lower than 50,000. The terms were relatively dispersed and came from different dimensions. To name

a few and their respective RMSE scores: "Gardener"—52,548, "Stress"—54,089, "Career Builder"—53,681, "Best Sugar Daddy Websites"—54,781, "Nervous"—54,276, "Luxury Yachts For Sale"—52,653, and "Rolex Watch"—46,531. Some of the terms come from the JSBA dimension and affirm prior researchers' term use as good starting points, but others touched upon anxiety escort services, stress symptoms, or particular luxury brands (Fondeur and Karamé 2013; Maas 2019; Larson and Sinclair 2021; D'Amuri and Marcucci 2017).

Furthermore, because the subsequent models (18a) and (18b) will be optimized by holding the 17th model keyword variable fixed, we decided to ground the first keyword in model (17b) in the JSBA dimension because this search query is more likely to be related to jobless claims in its entirety than medications or stress symptoms. Hence, the "Career Builder" term was selected as model (17b)'s first keyword and is reported in Table 2. From this point forward, models (18a) and (18b) attempted to further reduce forecasting errors for the pandemic timeframe. Through trial and error by including "Unemployment", "Home Refinance Rates", and "Withdrawal From Friends" as keywords for model (18a) and by including "Career Builder" and "Luxury Yachts For Sale" for model (18b) alongside lagged IC for both models, the errors were reduced by 2.67 and 24.39% from the persistence model for the pandemic period, respectively. Although further improvements could have been achieved by extensive optimization, the initial results suggest that additional keyword dimensions act as decent complementary variables for forecasting models and that "unemployment" is not the optimal choice of keyword.

**Table 2.** RMSE test scores retrieved from different LSTM models for different optimization targets.

| | | | **Vanilla LSTM** | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Optimized For Pandemic | | | | | |
| TF | Model 16 | Model (17a) | Model (17b) | Model (18a) | Model (18b) | Model (19) | Model (20) | Model (21) |
| Full | 38,577.35 | 40,065.89 | 37,825.99 | 59,880.21 | 42,081.86 | 50,979.00 | 99,616.56 | 196,009.32 |
| Pre-p. | 7548.92 | 7516.13 | 7562.90 | 7461.71 | 7705.44 | 8036.815 | 7837.147 | 159,54.240 |
| Pand. | 58,539.20 | 84,259.30 | **54,983.85** | **56,974.43** | **44,256.68** | 97,397.90 | 110,090.8 | 259,531.99 |
| | | | Optimized For Full | | | | | |
| Full | | | **34,539.50** | **36,792.24** | **28,842.48** | | | |
| Pre-p. | | | 7780.78 | 7564.31 | 10,095.02 | | | |
| Pand. | | | 64,079.87 | 140,741 | 117,158.69 | | | |
| | | | Optimized For Pre-Pandemic | | | | | |
| Full | | | 130,737.6 | 36,706.20 | 42,986.68 | | | |
| Pre-p. | | | **6768.696** | **6284.63** | **5456.54** | | | |
| Pand. | | | 95,662.77 | 85,302.66 | 142,165.29 | | | |
| | | | **Encoder–Decoder Style LSTM Layering** | | | | | |
| | | | Variables Used For Pandemic | | | | | |
| Full | 28,173.06 | 28,179.74 | 28,149.95 | 28,168.96 | 28,118.63 | 28,202.63 | 28,202.41 | 28,203.59 |
| Pre-p. | 7830.775 | 7828.18 | 7831.052 | 7830.53 | 7831.084 | 7830.88 | 7831.01 | 7830.82 |
| Pand. | 52,079.45 | 51,476.21 | 52,095.41 | 51,239.92 | 51,217.42 | 52,100.72 | 52,084.75 | 52,107.12 |
| | | | Variables Used For Full | | | | | |
| Full | | | 28,158.31 | 28,202.03 | 28,126.43 | | | |
| Pre-p. | | | 7830.871 | 7830.54 | 7831.04 | | | |
| Pand. | | | 52,299.02 | 51,613.26 | 51,774.72 | | | |
| | | | Variables Used For Pre-Pandemic | | | | | |
| Full | | | 28,154.71 | 28,174.80 | 28,114.49 | | | |
| Pre-p. | | | 7830.25 | 7830.618 | 7829.38 | | | |
| Pand. | | | 50,987.53 | 51,801.75 | 51,911.86 | | | |

**Note:** The numbers near the model names represent formulas that were presented in Section 4, and optimized numbers are bolded. The following GTI keywords are used alongside the lagged IC variable: model (17a)—"Unemployment", for pandemic optimization; model (17b)—"Career Builder"; model (18a)—"Unemployment", "Home Refinance Rates", "Withdrawal From Friends"; model (18b)—"Career Builder", "Luxury Yachts For Sale"; optimized for full period model (17b)—"Glassdoor"; model (18a)—"Chronic Pain", "Layoff Law", "Unemployment"; model (18b)—"Glassdoor", "Epidemic", "Plastic surgery near me"; optimization for the pre-pandemic period model (17b)—"Assistance"; model (18a)—"Assistance", "Plastic Surgery Cost", "Unemployment"; model (18b)—"Assistance", "Anxiety", "Employment Office". Due to time-consuming optimizations, the same specification of variables of LSTM layering was used for the encoder–decoder model.

In a similar manner, optimizations took place for the full and pre-pandemic time-frames. For the full timeframe, model (17b) iteration resulted in 11 keywords with an RMSE score lower than 36,000 (e.g., "Insomnia", "Glassdoor", "Epidemic", "Domestic Violence", "LinkedIn", "Gig Economy"). For the pre-pandemic period, 15 variables resulted in scores lower than 7400 (e.g., "Plastic Surgery Cost", "Home Refinance Rates", "Layoff Law", "Excessive Anger", "Ramipril", "Assistance"). Both latter scores represented an improvement on the persistence model for the respective timeframes and offered choices of keywords from multi-dimensions. To ground the first variables close to the JSBA dimension, "Glassdoor" and "Assistance" were selected for full and pre-pandemic timeframes. Extending model (18a) optimized for the full timeframe using the keywords "Chronic Pain" and "Layoff Law" helped further improve the accuracy by 4.62 percent from the persistence model, whereas extending the pre-pandemic model (18a) with the "Assistance" and "Plastic Surgery Cost" keywords led to a 16.7 percent improvement. Analogously, extending the full timeframe model (18b) with "Epidemic" and "Plastic surgery near me" improved forecasting errors by 25.23 percent, whereas using the pre-pandemic model (18b) extension with "Anxiety" and "Employment Office" resulted in an improvement of 27.7 percent over the persistence model. However, the improvements in forecasting errors came at a cost of other periods, which ultimately worsened predictions. Some consistencies to be mentioned are that although the dimensions and keywords varied, the JSBA dimension was always available to improve the prediction errors, although the predictive power differed across timeframes.

Furthermore, because the choice of variables fluctuated, an attempt was made to use PCA to reduce the dimensions. Hence, rather than dealing with long optimization processes for individual keywords, we incorporated all the dimensions at once. Unfortunately, whether considering the six major dimensions or their sub-dimensions, the scores were considerably worse. Even when considering the JSBA dimension on its own—something that authors such as Schiavoni et al. (2021), Yi et al. (2021), and Caperna et al. (2020) limited themselves to—the model was the best out of three PCA specifications but still worse by 32.14, 6.46, and 66.38 percent for the full, pre-pandemic, and pandemic timeframes compared with the persistence model.

The same keyword models were deployed with encoder–decoder layers that attempted to deal with extreme event situations more efficiently, although they were not optimized according to model layers due to computational limitations. The results were more conservative and did not surpass the optimized vanilla models regarding accuracy; this result originated as a consequence of dealing with outliers, although specific improvements can be seen in models (18a) and (18b) for the full timeframe compared with the persistence encoder–decoder model. It is most likely that individual optimizations for encoder–decoder models could have benefited the models more; however, out of the box, they were more resilient, particularly when considering the PCA models.

With regard to authors such as Pavlicek and Kristoufek (2015), Nagao et al. (2019), and Barreira et al. (2013), all of whom achieved mixed results using GTI, the empirical results of this paper support the latter authors' claims. The use of GTI is not as straightforward because the GTI keyword relationship is dynamic with the jobless claims indicator. This dynamism may be one of the reasons why long-term forecasting was found to be ineffective by some researchers, especially if it was considered during economic growth periods. An analogous situation was encountered by Caperna et al. (2020) in which specific Google keyword topics showed no significance in forecasting; thus, the careful selection of keywords was necessary.

## 6. Conclusions

This study proposed the need to expand the GTI keyword search queries to multiple dimensions in order to predict the number of initial jobless claims in the USA. During the literature analysis, it was found that many research papers limit themselves to only a few keywords that relate to the job market, mostly "Unemployment", "Jobs", and "Job offer",

thus excluding many other keyword opportunities, which could enhance the forecasting models and help government bodies to be better prepared and take appropriate action before the crisis snowballs. As such, it was decided that it was necessary to create six GTI dimensions with 222 keywords in total that better encompass human behavior from different perspectives. For the purpose of prediction, the recurrent neural networks (LSTM) with different layers were tested with multiple variable specifications that were optimized using the variable MSE scores or chosen according to other authors' experiences.

The empirical results suggest that Google Trends keywords do carry predictive power.

Consistent with previous studies, work-related keywords such as "unemployment", "career builder", and others have contributed significantly to many model specifications. The findings also suggest that the consumption lifestyle dimension searches, e.g., "luxury yacht for sale" or "plastic surgery cost", were important variables in increasing prediction accuracy, confirming previous authors' findings that there is a shift in how people spend their time and money. Models were also significantly improved with mental health keywords such as "anxiety" and "Withdrawal From Friends" affirming a strong link between unemployment and mental health issues and validating the findings of other authors. Although not all model parameters were tested, the violence, abuse, and gambling dimensions were not the first model choices for accuracy improvement, although, in certain cases, strong correlations were detected.

However, the difficulty is in determining whether the intentions behind the keyword search query are related to job market processes, as opposed to when the search queries just coincidently match. Furthermore, during economic growth periods, searches for medications or particular game brands might not necessarily be related to being laid off from work but can have other purposes, and, as a consequence, they showed little to no correlation in the pre-pandemic timeframe. On the other hand, during economic downturns, the stress searches might be induced by being laid off or anticipating being fired from work, and, therefore, they begin to correlate and co-move. By considering different optimization pathways and model layers, we were able to reduce the RMSE errors by 16.7, 25.23, and 27.7% for the pandemic, full, and pre-pandemic timeframes compared to the persistence model. Nonetheless, the error improvement possibilities were not exhausted to their fullest due to computer resource limitations. It was also discovered that the "unemployment" keyword was not among the best keywords to choose from for increasing prediction accuracy.

## 7. Policy Recommendations

Five directions in policy-relevant unemployment forecasting work can be suggested. First, as the world is becoming more abundant in data, government bodies, and commercial and central banks should aim to explore alternative datasets that can help to further increase accuracy in unemployment number predictions. Google keywords could be one choice out of the many that could be used not only for unemployment but for GDP or inflation forecasts as well. Second, for forecasting institutions, we recommend they craft their own sets of important Google keywords that correlate with countries' unemployment changes. However, as the literature analysis and empirical results indicate, there is substantial evidence to go beyond typical keywords related to the job process, such as "jobs" and "work" and to encompass mental health, violence and abuse, leisure search, consumption and lifestyle, disasters, and war and virus related keywords. Third, the constructed keyword datasets will be highly sensitive to timing. As our analysis indicates, in stable economic conditions, the correlation of Google keywords with unemployment significantly drops compared to economic recessions. Thus, if a policy decision is to tackle both periods with the same set of keywords, a work-related dimension might produce the most consistent results; however, this may be at the cost of accuracy. Thus, two sets of datasets are encouraged: one with the work-related keywords only and the second that includes other dimensions discussed in the methodology section. Fourth, as the number of variables increases, more efficient modeling procedures will need to be adopted. We advise

forecasting institutions to explore deep learning methods for forecasting related matters as evidence shows tremendous potential at being more efficient than VAR or factor models. Lastly, there is also a trade-off between LSTM model specifications. A vanilla LSTM model can be pushed further to provide better accuracy results as it has fewer constraints; however, it is more sensitive to outliers compared to auto-encoders. Depending on the aim, policy bodies need to be aware of whether accuracy or consistency is the aim and choose accordingly.

## 8. Future Directions and Limitations

The major limitation of the proposed Google dimension expansion is the stability and indication of keyword uses. As of now, it is difficult to consider one luxury brand as indicative of an upcoming crisis because, in the next ten years, the brand might lose its client base or popularity, and, therefore, the keyword may no longer be indicative. Similar issues may arise with other queries (such as mental health medication from anxiety) where the brand name might change under new management, thus making the keyword obsolete. As a suitable approach for such situations, future researchers may be advised to do the following:

(a) Consider gathering many keywords over multiple crises and extract the most indicative keywords that are prevalent in all crises and all periods;

(b) Further explore keyword dimension reduction techniques to avoid omitting useful keywords that might have a high prediction impact on one crisis but not the other.

Additionally, a method of quantifying the degree to which keyword intention relates to job market searching could dramatically assist in understanding why many of the keywords during periods of economic growth carried little to no information regarding the IC variable. Perhaps case studies of individuals would help to understand the searching behavior of an unemployed or soon-to-be unemployed person, thus helping to avoid coincidental keyword matches with initial claims growth. Future studies could also attempt to include other variables, such as inflation and business surveys, as well as analyze how Google Trends queries in parallel with other economic variables can reduce error rates.

Furthermore, the PCA models delivered poor results; however, important limitations need to be acknowledged. For instance, the PCA assumes a linear relationship, which may be the source of PCA model inaccuracies. We encourage other authors to explore other dimensionality techniques, such as isomap or t-distributed stochastic neighbor embedding (t-SNE), that are more advanced and take into account non-linear relationships. The t-SNE model, due to heavy-tailed t-distribution, is also less prone to outliers, which might be another source of the significant drawbacks for the PCA model.

Lastly, because there are more than a million variables and parameter permutations, future researchers who can access cloud computing could fully exhaust the potential for error improvements. Future optimizations may further consider memory cell numbers, different lags or layers, or even various neural nets that could be a combination of CNNs (convolution neural networks) and LSTMs.

**Author Contributions:** Conceptualization, A.G., M.L., V.P., J.B. and A.S.; methodology, A.G., M.L., V.P., J.B. and A.S.; software, A.G.; validation, V.P., J.B., A.S. and A.G.; formal analysis, A.G.; investigation, M.L., V.P. and A.G.; resources, M.L., V.P., J.B., A.S. and A.G.; data curation, A.G.; writing—original draft preparation, M.L. and A.G.; writing—review and editing, A.G., V.P., J.B. and A.S.; visualization, A.G.; supervision, V.P., J.B. and A.S.; project administration, V.P., J.B. and A.S.;

## Notes

[1] According to Speed (2011) and Edelmann et al. (2020), the Pearson correlation coefficient suffers from the drawback of only being able to identify the linear dependence while leaving non-linear patterns undetected (Speed 2011). Furthermore, Pearson requires a roughly normal distribution with no extreme outliers. These serious limitations may miss the dynamism that could exist among the dependent variable and Google Trends.

[2] In this paper, the neural networks were compiled using the MSE loss functions. Thus, RMSE is a better-suited metric for accuracy evaluation than MAPE or MAE. Furthermore, optimization was not possible using MAPE due to differencing the data, which leads to infinite values because the original target values in certain cases come close to zero.

[3] The full timeframe corresponds to the period of 2017/03/19 to 2021/01/16; the pre-pandemic timeframe corresponds to 2017/03/19 to 2020/02/01; and the pandemic timeframe corresponds to 2020/02/08 to 2021/01/16.

## References

Aaronson, Daniel, Scott A. Brave, R. Andrew Butters, Michael Fogarty, Daniel W. Sacks, and Boyoung Seo. 2022. Forecasting unemployment insurance claims in realtime with Google Trends. *International Journal of Forecasting* 38: 567–81. [CrossRef]

Aastveit, Knut Are, Tuva Marie Fastbø, Eleonora Granziera, Kenneth Sæterhagen Paulsen, and Kjersti Næss Torstensen. 2020. Nowcasting Norwegian Household Consumption with Debit Card Transaction Data. Norges Bank, Working Paper 17/2020. Available online: https://hdl.handle.net/11250/2722899 (accessed on 4 August 2022).

Agarwal, Sumit, Tal Gross, and Bhashkar Mazumder. 2016. How Did the Great Recession Affect Payday Loans? Economic Perspective. Available online: https://fraser.stlouisfed.org/files/docs/historical/frbchi/economicperspectives/frbchi_econper_2016n2.pdf (accessed on 4 August 2022).

Anderberg, Dan, Helmut Rainer, Jonathan Wadsworth, and Tanya Wilson. 2016. Unemployment and Domestic Violence: Theory and Evidence. *Economic Journal, Royal Economic Society* 126: 1947–79. [CrossRef]

Antolini, Fabrizio, and Laura Grassini. 2018. Foreign arrivals nowcasting in Italy with Google Trends data. *Quality & Quantity* 53: 2385–401. [CrossRef]

Anttonen, Jetro. 2018. *Nowcasting the Unemployment Rate in the EU with Seasonal BVAR and Google Search Data*. ETLA Working Papers, No. 62. Helsinki: The Research Institute of the Finnish Economy (ETLA).

Askitas, Nikolaos, and Klaus F. Zimmermann. 2009. *Google Econometrics and Unemployment Forecasting*. IZA Discussion Paper. Bonn: Institute for the Study of Labor, p. 4201.

Askitas, Nikolaos. 2015. *Google Search Activity Data and Breaking Trends*. IZA World of Labor. Bonn: Institute for the Study of Labor (IZA), p. 206. [CrossRef]

Banbura, Marta, Domenico Giannone, and Lucrezia Reichlin. 2010. Nowcasting (November 30, 2010). ECB Working Paper No. 1275. Available online: https://ssrn.com/abstract=1717887 (accessed on 4 August 2022).

Bańbura, Marta, Domenico Giannone, Michele Modugno, and Lucrezia Reichlin. 2013. Now-Casting and the Real-Time Data Flow. Working Paper Series. Available online: https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1564.pdf (accessed on 4 August 2022).

Barreira, Nuno, Pedro Godinho, and Paulo Melo. 2013. Nowcasting unemployment rate and new car sales in south-western Europe with Google Trends. *NETNOMICS: Economic Research and Electronic Networking* 14: 129–65. [CrossRef]

Berniell, Inés, and Gabriel Facchini. 2021. COVID-19 lockdown and domestic violence: Evidence from internet-search behavior in 11 countries. *European Economic Review* 136: 103775. [CrossRef]

Bhalotra, Sonia, Diogo G. C. Britto, Paolo Pinotti, and Breno Sampaio. 2021. Job Displacement, Unemployment Benefits and Domestic Violence. CEPR Discussion Paper No. DP16350. Available online: https://ssrn.com/abstract=3886839 (accessed on 4 August 2022).

Borup, Daniel, David E. Rapach, and Erik Christian Montes Schütte. 2021. Mixed-Frequency Machine Learning: Now- and Backcasting Weekly Initial Claims with Daily Internet Search-Volume Data (July 28, 2021). Available online: https://ssrn.com/abstract=3690832 (accessed on 4 August 2022).

Breuer, Christian. 2015. Unemployment and Suicide Mortality: Evidence from Regional Panel Data in Europe. *Health Economics* 24: 936–50. [CrossRef]

Brownlee, Jason. 2018. Deep Learning for Time Series Forecasting. Machine Learning Mastery. Available online: https://machinelearningmastery.com/deep-learning-for-time-series-forecasting/ (accessed on 4 August 2022).

Bruneckiene, Jurgita, Robertas Jucevicius, Ineta Zykiene, Jonas Rapsikevicius, and Mantas Lukauskas. 2021. Quantum Theory and Artificial Intelligence in the Analysis of the Development of Socio-Economic Systems: Theoretical Insights. In *Developing Countries and Technology Inclusion in the 21st Century Information Society*. Hershey: IGI Global.

Butterworth, Peter, Liana S. Leach, Jane Pirkis, and Margaret Kelaher. 2012. Poor mental health influences risk and duration of unemployment: A prospective study. *Social Psychiatry and Psychiatric Epidemiology* 47: 1013–21. [CrossRef] [PubMed]

Caperna, Giulio, Marco Colagrossi, Andrea Geraci, and Gianluca Mazzarella. 2020. *Googling Unemployment During the Pandemic: Inference and Nowcast Using Search Data*. JRC Working Papers in Economics and Finance. Brussels: Joint Research Centre, European Commission.

Chadwick, Meltem Gülenay, and Gönül Sengul. 2012. *Nowcasting Unemployment Rate in Turkey: Let's Ask Google*. Ankara: Central Bank of the Republic of Turkey.

Chadi, Adrian, and Clemens Hetschko. 2017. Income or Leisure? On the Hidden Benefits of (Un-)employment. Available online: https://www.researchgate.net/publication/329152413_Income_or_Leisure_On_the_Hidden_Benefits_of_Un-employment (accessed on 4 June 2022).

Chernis, Tony, and Rodrigo Sekkel. 2017. A dynamic factor model for nowcasting Canadian GDP growth. *Empirical Economics* 53: 217–34. [CrossRef]

Chiaroni, Caroline, and Greg Kaplan. 2016. Do Households Substitute Among Luxury Goods? The Impact of the Great Recession on Fragrance Consumption. Available online: https://dataspace.princeton.edu/handle/88435/dsp01g732dc44x (accessed on 4 August 2022).

Choi, Hyunyoung, and Hal Varian. 2009a. *Predicting Initial Claims for Unemployment Benefits*. Google Technical Report. Mountain View: Google Inc.

Choi, Hyunyoung, and Hal Varian. 2009b. Predicting Present with Google Trends. Available online: https://static.googleusercontent.com/media/www.google.com/en//googleblogs/pdfs/google_predicting_the_present.pdf (accessed on 4 August 2022).

Coble, David, and Pablo M. Pincheira. 2017. Now-Casting Building Permits with Google Trends (February 1, 2017). Available online: https://ssrn.com/abstract=2910165 (accessed on 5 August 2022). [CrossRef]

D'Amuri, Francesco. 2009. *Predicting Unemployment in Short Samples with Internet Job Search Query Data*. MPRA Paper No: 18403. Boston: Statistical Association.

D'Amuri, Francesco, and Juri Marcucci. 2017. The predictive power of Google searches in forecasting US unemployment. *International Journal of Forecasting* 33: 801–16. [CrossRef]

Dávalos, María E., Hai Fang, and Michael T. French. 2011. Easing The Pain of An Economic Downturn: Macroeconomic Conditions And Excessive Alcohol Consumption. *Health Economics* 52: 1318–35. [CrossRef]

Dilmaghani, Maryam. 2019. Workopolis or The Pirate Bay: What does Google Trends say about the unemployment rate? *Journal of Economic Studies* 46: 422–45. [CrossRef]

Edelmann, Dominic, Tamás F. Móri, and Gábor J. Székely. 2020. On relationships between the Pearson and the distance correlation coefficients. *Statistics & Probability Letters* 169: 108960. [CrossRef]

Ettredge, Michael, John Gerdes, and Gilbert Karuga. 2005. Using Web-based Search Data to Predict Macroeconomic statistics. *Communications of the ACM* 48: 87–92. [CrossRef]

Fenga, Livio, and Semen Son-Turan. 2022. Forecasting youth unemployment in the aftermath of the COVID-19 pandemic: The Italian case. *International Journal of Scientic and Management Research* 5: 75–91. [CrossRef]

Fondeur, Yannick, and Frédéric Karamé. 2013. Can Google data help predict French youth unemployment? *Economic Modelling* 30: 117–25. [CrossRef]

Frangos, Christos C., Constantinos C. Frangos, and Ioannis Sotiropoulos. 2011. Cyberpsychology, Behavior, and Social Networking. *Cyberpsykologi, Beteende och Sociala Nätverk* 14: 51–58. [CrossRef]

Gabrielyan, Gnel, and David R. Just. 2020. Economic shocks and lottery sales: An examination of Maine State lottery sales. *Applied Economics* 52: 3498–511. [CrossRef]

Gamze, Gamze Bayın, and Seda Aydan. 2021. Association of COVID-19 with lifestyle behaviours and socio-economic variables in Turkey: An analysis of Google Trends. *International Journal of Health Planning and Management* 1: 20. [CrossRef]

Giannone, Domenico, Lucrezia Reichlin, and David Small. 2008. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics* 55: 665–76. [CrossRef]

Ginsberg, Jeremy, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant. 2009. Detecting influenza epidemics using search engine query data. *Nature* 457: 1012–1014. [CrossRef]

Giovannelli, Tommaso, Alessandro Giovannelli, Ottavio Ricchi, Ambra Citton, Christían Tegami, and Cristina Tinti. 2020. Nowcasting GDP and Its Components in a Data-Rich Environment: The Merits of the Indirect Approach (May 29, 2020). CEIS Working Paper No. 489. Available online: https://ssrn.com/abstract=3614110 (accessed on 4 August 2022).

Goodman, William K., Ashley M. Geiger, and Jutta M. Wolf. 2016. Leisure Activities Are Linked to Mental Health Benefits by Providing Time Structure: Comparing Employed, Unemployed and Homemakers. *Journal of Epidemiology and Community Health* 71: 4–11. Available online: https://jech.bmj.com/content/71/1/4.short (accessed on 6 August 2022). [CrossRef] [PubMed]

Havitz, Mark E., Peter A. Morden, and Diane M. Samdahl. 2004. *The Diverse Worlds of Unemployed Adults: Consequences for Leisure, Lifestyle, and Well-Being*. Waterloo: Wilfrid Laurier University Press.

Hopp, Daniel. 2021. *Economic Nowcasting with Long Short-Term Memory Artificial Neural Networks (LSTM)*. UNCTAD Research Paper No. 62. Genève: UNCTAD.

Huang, Ni, Gordon Burtch, and Paul Pavlou. 2018. Local Economic Conditions and Worker Participation in the Online Gig Economy. Paper presented at the Thirty Ninth International Conference on Information Systems, San Francisco, CA, USA, December 13–16; Available online: http://metadataetc.org/gigontology/pdf/Huang%20et%20al.%20%20Local%20Economic%20Conditions%20and%20Worker%20Participation.pdf (accessed on 4 August 2022).

Hyndman, Rob J., and George Athanasopoulos. 2018. *Forecasting: Principles and Practice*, 2nd ed. Melbourne: OTexts. Available online: OTexts.com/fpp2 (accessed on 6 August 2022).

Jolliffe, Ian T. 2002. *Principal Component Analysis*. Springer Series in Statistics; New York: Springer.

Khanthavit, Anya. 2021. A Causality Analysis of Lottery Gambling and Unemployment in Thailand. *The Journal of Asian Finance, Economics and Business* 8: 149–56.

Kinney, Justin B., and Gurinder S. Atwal. 2014. Equitability, mutual information, and the maximal information coefficient. *Proceedings of the National Academy of Sciences* 111: 3354–59. [CrossRef]

Laptev, Nikolay, Slawek Smyl, and Santhosh Shanmugam. 2017. Engineering Extreme Event Forecasting at Uber with Recurrent Neural Networks. Available online: https://eng.uber.com/neural-networks/ (accessed on 4 August 2022).

Larson, William D., and Tara M. Sinclair. 2021. Nowcasting unemployment insurance claims in the time of COVID-19. *International Journal of Forecasting* 38: 635–47. [CrossRef]

Lehdonvirta, Vili. 2013. A history of the digitalization of consumer culture. In *Digital Virtual Consumption*. Edited by Mike Molesworth and Janice Denegri-Knott. Abingdon-on-Thames: Routledge, pp. 18–35.

Liu, Shuyan, Stephan Heinzel, Matthias N. Haucke, and Andreas Heinz. 2021. Increased Psychological Distress, Loneliness, and Unemployment in the Spread of COVID-19 over 6 Months in Germany. *Medicina* 57: 53. [CrossRef]

Maas, Benedikt. 2019. Short-term forecasting of the US unemployment rate. *Journal of Forecasting* 39: 394–411. [CrossRef]

Mallorquí-Bagué, Nuria, Fernando Fernández-Aranda, María Lozano-Madrid, Roser Granero, Gemma Mestre-Bach, Marta Baño, Amparo Del Pino-Gutiérrez, Mónica Gómez-Peña, Neus Aymam, José M. Menchón, and et al. 2017. Internet gaming disorder and online gambling disorder: Clinical and personality correlates. *Journal of Behavioral Addictions* 6: 669–77. [CrossRef]

McKinsey & Company. 2020. How the Coronavirus Could Change US Personal Auto Insurance. Available online: http://dln.jaipuria.ac.in:8080/jspui/bitstream/123456789/1531/1/How-the-coronavirus-could-change-us-personal-auto-insurance.pdf (accessed on 4 August 2022).

Mousteri, Victoria, Michael Daly, and Liam Delaney. 2018. The scarring effect of unemployment on psychological well-being across Europe. *Social Science Research* 72: 146–69. [CrossRef]

Mulero, Rodrigo, and Alfredo García-Hiernaux. 2021. Forecasting Spanish unemployment with Google Trends and dimension reduction techniques. *SERIEs* 12: 329–49. [CrossRef]

Naccarato, Alessia, Stefano Falorsi, Silvia Loriga, and Andrea Pierini. 2017. Combining official and Google Trends data to forecast the Italian youth unemployment rate. *Technological Forecasting and Social Change* 130: 114–22. [CrossRef]

Nagao, Shintaro, Fumiko Takeda, and Riku Tanaka. 2019. Nowcasting of the U.S. unemployment rate using Google Trends. *Finance Research Letters* 30: 103–109. [CrossRef]

Netmarketshare. 2022. Google Search Engine Total Market Share. Available online: https://netmarketshare.com/search-engine-market-share (accessed on 4 August 2022).

Nymand-Andersen, Per, and Emmanouil Pantelidis. 2018. *Google Econometrics: Nowcasting Euro Area Car Sales and Big Data Quality Requirements*. ECB Statistics Paper, No. 30. Frankfurt a. M.: European Central Bank (ECB). ISBN 978-92-899-3359-9. [CrossRef]

Pallesen, Ståle, Rune Aune Mentzoni, Arne Magnus Morken, Jonny Engebø, Puneet Kaur, and Eilin Kristine Erevik. 2021. Changes Over Time and Predictors of Online Gambling in Three Norwegian Population Studies 2013–2019. *Frontiers in Psychiatry* 12: 597615. [CrossRef] [PubMed]

Palomeque Recio, Rocío. 2021. 'I have bills to pay!' Sugar dating in British higher education institutions. *Journal of Gender and Education* 34: 545–60. [CrossRef]

Pavlicek, Jaroslav, and Ladislav Kristoufek. 2015. Nowcasting Unemployment Rates with Google Searches: Evidence from the Visegrad Group Countries. *PLoS ONE* 10: e0127084. [CrossRef]

Reshef, David N., Yakir A. Reshef, Hilary K. Finucane, Sharon R. Grossman, Gilean McVean, Peter J. Turnbaugh, Eric S. Lander, Michael Mitzenmacher, and Pardis C. Sabeti. 2011. Detecting Novel Associations in Large Data Sets. *Science* 334: 1518–24. [CrossRef]

Reyneke, Mignon, Alexandra Sorokáčová, and Leyland Pitt. 2012. Managing brands in times of economic downturn: How do luxury brands fare? *Journal of Brand Management* 19: 457–66. [CrossRef]

Richardson, Adam, and Thomas Mulder. 2018. Nowcasting New Zealand GDP Using Machine Learning Algorithms. *International Journal of Forecasting* 37: 941–48. [CrossRef]

Rusnák, Marek. 2016. Nowcasting Czech GDP in real time. *Economic Modelling* 54: 26–39. [CrossRef]

Schiavoni, Caterina, Franz Palm, Stephan Smeekes, and Jan van den Brakel. 2021. A dynamic factor model approach to incorporate Big Data in state space models for official statistics. *Journal of the Royal Statistical Society Series A-Statistics in Society* 184: 324–53. [CrossRef]

Schmidhuber, Jürgen, and Sepp Hochreiter. 1997. Long-short-term memory. *Neural Computation* 9: 1735–80.

Simionescu, Mihaela. 2020. Improving unemployment rate forecasts at regional level in Romania using Google Trends. *Technological Forecasting and Social Change* 155: 120026. [CrossRef]

Simionescua, Mihaela, and Javier Cifuentes-Faura. 2021. Can unemployment forecasts based on Google Trends help government design better policies? An investigation based on Spain and Portugal. *Journal of Policy Modeling* 44: 1–21. [CrossRef]

Singhania, Rajshekhar, and Sourav Kundu. 2020. Forecasting the United States Unemployment Rate by Using Recurrent Neural Networks with Google Trends Data (June 18, 2020). *International Journal of Trade, Economics and Finance* 11. Available online: https://ssrn.com/abstract=3801209 (accessed on 6 August 2022).

Smith, Paul. 2016. Google's MIDAS Touch: Predicting UK Unemployment with Internet Search Data. *Journal of Forecasting* 35: 263–84. [CrossRef]

Sotis, Chiara. 2021. How do Google searches for symptoms, news and unemployment interact during COVID-19? A Lotka–Volterra analysis of google trends data. *Quality & Quantity* 55: 2001–16. [CrossRef]

Speed, Terry. 2011. A Correlation for the 21st Century. *Science* 334: 1502–503. [CrossRef] [PubMed]

Uzieblo, Kasia, and David Prescott. 2020. Online pornography use during the COVID-19 pandemic: Should we worry? Part I. *Sexual Abuse-Blogspot* 40: 1080–1089. [CrossRef]

Wanberg, Connie R., Edwin A. J. van Hooft, Karyn Dossinger, Annelies E. M. van Vianen, and Ute-Christine Klehe. 2019. How Strong Is My Safety Net? Perceived Unemployment Insurance Generosity and Implications for Job Search, Mental Health, and Reemployment. *Journal of Applied Psychology* 105: 209. [CrossRef]

Yi, Dingdong, Shaoyang Ning, Chia-Jung Chang, and Supeng Kou. 2021. Forecasting Unemployment Using Internet Search Data via PRISM. *Journal of the American Statistical Association* 116: 1662–73. [CrossRef]