Aggregated Learning for Mixed Frequency Data

Takamichi Toda
Daisuke Moriwaki
Kazuhiro Ota

ArXiv:2105.09579v1 [cs.LG] 20 May 2021

ABSTRACT

Large and acute economic shocks such as the 2007-2009 financial crisis and the current COVID-19 infections rapidly change the economic environment. In such a situation, the importance of real-time economic analysis using alternative data is emerging. Alternative data such as search query and location data are closer to real-time and richer than official statistics that are typically released once a month in an aggregated form. We take advantage of spatio-temporal granularity of alternative data and propose a Mixed-Frequency Aggregate Learning (MF-AGL) model that predicts economic indicators for the smaller areas in real-time. We apply the model for the real-world problem; prediction of the number of job applicants which is closely related to the unemployment rates. We find that the proposed model predicts (i) the regional heterogeneity of the labor market condition and (ii) the rapidly changing economic status. The model can be applied to various tasks, especially economic analysis.

CCS CONCEPTS

- Theory of computation → Semi-supervised learning; - Applied computing → Economics; - Information systems → Geographic information systems.

KEYWORDS

Nowcasting, Aggregate Output learning, Aggregate Learning, Semi-supervised learning, Mobility data, GPS, Location data, Unemployment, Mixed-frequency data, Economic policy

1 INTRODUCTION

Devastating external economic shocks such as the 2007-2009 financial crisis and COVID-19 infections rapidly change economic circumstances. Amid these situations, real-time economic analysis is essential. The real-time analysis of the economic conditions is often called Nowcasting or Economic Nowcasting. Nowcasting provides important insights into the current economic status that are essential to appropriate policy responses including monetary and fiscal stimuli. An understanding of the on-going economic situation is also needed by private companies who make decisions on investment and employment.

Nowcasting often takes advantage of alternative data, non-standard data such as search queries, location data, SNS data, and satellite images [3, 6, 8, 14, 25, 37]. These data are suitable for economic nowcasting because of their high frequency. While most of the official statistics are reported once a month, alternative data is usually recorded in real-time. Another beauty of alternative data is its granularity. While official statistics usually report state-level statistics alternative data can provide detailed information up to the individual-level.

In response to the surging demand for alternative data, several works have already emerged. Tech giants such as Apple, Google, and Facebook provide mobility statistics based on their proprietary data [2, 12, 28]. These data help us understand several aspects of economic status. However, it is hard to understand the whole picture of the economy itself because the relationship between these alternative data-based statistics and familiar economic indices such as GDP and unemployment rates is unclear. To fill this gap, many nowcasting/forecasting models that use high-frequency data such as Google search query to predict economic variables are proposed [3, 5, 6, 25, 32, 35, 37]. Combining high-frequency data like Google search query and low-frequency data such as monthly unemployment rates and quarterly GDP has been actively studied by economists [9, 10].

One caveat of these studies is that they do not fully extract the granularity of the alternative data. Target variables of these studies are often aggregated statistics such as GDP and unemployment rates. Although the predictor variables have a greater granularity, the predicted values are aggregated at the state or even national level. However, the governments, especially local governments need to deal with heterogeneity among small areas. For example, the bankruptcy of a large automaker will critically affect the local economy where the factories are located but does not affect the labor market at a national level. Authorities need to be sensitive to not only temporal changes but also regional heterogeneity.

The problem is that there is no granular data in the official statistics. That is, there is no label for the granular level. The problem is called aggregate output learning or aggregate learning and there has been only a few research papers has been published until recently [27]. However, the problem has been taking attention from an increasing number of researchers [7, 19].

We, here, combine mixed-frequency data literature in economics and aggregate learning literature in machine learning to fully utilize the richness of alternative data and provide an example of useful application. Our Mixed-Frequency Aggregate Learning (MF-AGL) model takes advantage of spatio-temporal granularity of alternative data and predicts economic indicators in smaller areas in real-time, which are not possible by standard forecasting models. The model also updates the prediction in real-time using

Daisuke Moriwaki
Kazuhiro Ota

Daisuke Moriwaki
Kazuhiro Ota
high-frequency data without high-frequency target data. To train the model, we define a novel loss function for spatio-temporally aggregated label data and granular predicted values. More specifically, we aggregate predicted values for small areas and calculate loss using an aggregated level. We also calculate the loss for each high-frequency features so that the model can learn the mixed-frequency structure of the data. That means we reuse the same label data over and over.

We apply the model to the real-world problem: prediction of the number of job applicants for smaller areas in Japan. The number of job applicants reflects the condition of the labor market because job applicants are also unemployed persons. An acute increase in the number of job applicants implies deterioration of labor market conditions.

To our best knowledge, the present work is the first to propose a novel method combining mixed frequency data and aggregate learning. We also demonstrate its practical importance in a real-world application. While we applied the model to the nowcasting of the labor market, the model can be applied to any task that contains (i) infrequent and aggregated indices such as GDP and (ii) spatio-temporally granular data.

In the following sections, we first discuss the background and related works in Section 2, then state the problem setting and describe the proposed model in Section 3. We then show the experiment results with the detail of the data and data pipeline in Section 4. Finally, Section 5 concludes.

2 BACKGROUND AND RELATED WORKS

2.1 Economic Nowcasting with Mixed-Frequency Data

Nowcasting using mixed-frequency data has been an active research area [1, 4, 9, 10, 20, 24, 36]. In particular, nowcasting of labor market statistics with alternative data date back to the late 2000s [3, 6, 37]. They suggested the potential predictive power of search query data. While most of the studies utilize web data represented by Google trends, Moriwaki (2020) [25] uses smartphone GPS data to predict unemployment insurance benefits. Hence the number of visitors is providing alternative data-based indices [2, 12, 28]. The monthly statistics are usually released a month after the end of the month. Real-time data for this number is of high value in economic nowcasting.

2.2 Aggregate Output Learning

In their seminal work [27], Musicant et al. (2007) first defined the aggregate output learning problem in which the labels are only available in an aggregated form. They investigate various machine learning models that are applicable to the problem. The aggregate output learning problem has been recently re-investigated in various research [7, 19, 38].

The present work is also related to unsupervised learning for super-resolution [34] and video interpolation [33] in the computer vision literature, which aims at recovering granularity from the data without labels.

3 PROBLEM SETTING AND METHOD

3.1 Problem Setting

Let \( y^q_\tau \in \mathbb{R} \) be a target variable which is of interest of economists (e.g., unemployment rate and GDP), \( p \) stands for some larger area such as nation and state and \( t \) stands for some longer time period such as quarter and month. Let \( x^q_\tau \in \mathbb{R} \) be feature variable which is correlated with the target variable \( y^q_\tau \). \( q \) stands for some smaller area such as city and county and \( \tau \) stands for shorter time period such as day. The difference between \( t \) and \( \tau \), and \( p \) and \( q \) are the granularity. In particular, an area \( p \) is divided by multiple small areas \( q \) and time period \( t \) is divided by multiple short time period \( \tau \). We use mapping function \( \mu \) and \( \pi \) to describe the relationship. In particular, \( \mu(\tau) = t \) indicates that time \( \tau \) belongs to \( t \) and \( \pi(q) = p \) indicates area \( q \) belongs to \( p \).

Our goal is to find a predictor \( f \) which predict \( y^q_\tau \) from granular data \((x^q_\tau)|_{\tau \leq \tau_0}\), where \( \mu(\tau_0) = t_0 \), (i.e., \( \tau_0 \) belongs to \( t_0 \)). Notice that the superscript is not larger area \( p \) but small area \( q \). That is, \( f \) predicts \( y^q_\tau \) instead of \( y^p_\tau \).

Table 1 illustrates the data structure of our problem. Feature vector \( x \) is observed for City 1 and 2 and for each day. But output value \( y \) is only observed for prefecture (assume the prefecture comprise of only two cities) and each month. We want to predict monthly values for each city (i.e., \( y^1_1 \) and \( y^2_1 \)). Since feature vector \( x \) is collected in real-time, we want to update the prediction using the latest information. As shown in the Table 1 the predicted values are changing according to the feature vector. That is, \( \tilde{y}^1_{\text{Jan1}|x_{\text{Jan1}}} \neq \tilde{y}^2_{\text{Jan1}|x_{\text{Jan1}}} \). In this way, we can fully utilize the real-time and granular data for the forecast.

While we conduct spatial disaggregation, we directly predict the aggregated value for a longer period. That is, we predict \( y^Q_\tau \) rather than \( y^q_\tau \). This seems simple supervised learning but it is not. To see this, let \( \tau \) be Jan 15, 2020. Then \( t \) is Jan 2020. We have data only for the past. That is, we have \( y^Q_{\text{Jan15}} \sim y^Q_{\text{Jan15}} \) but do not have \( y^Q_{\text{Jan16}} \sim y^Q_{\text{Jan16}} \). Nowcasting predicts the aggregated value using incomplete time series data [10]. For economists, the monthly value \( y^Q_\tau \) is more informative than the daily value \( y^q_\tau \). \( y^Q_\tau \) represents the economic status in a very short term while \( y^q_\tau \) represents the forecast of the longer time period. For example, let \( y^Q_\tau \) be the unemployment rate for January 2020 and \( y^Q_\tau \) be that for January 1, 2020. While the daily movement of the unemployment rate is
In this section, we describe our proposed model.

### 3.2 Aggregate Learning for Mixed-Frequency Data

In this section, we describe our proposed model.

#### 3.2.1 Aggregate Function

Following the aggregate learning literature\cite{7,27}, we first define the aggregation function as,

$$
\tilde{y}_t^q = \sum_{p=1}^{P} \omega_p f(x_t^q, \phi_t^q),
$$

where $\omega_p \in [0, 1]$ is weight. The weight controls for the share of the values for each granular area in the large area. In the simplest case (including the application discussed below), $\omega_p$ is set to one. In these cases, $f$ can learn the actual values of the target from the features. However, when we only obtain normalized features such as population per acre or averaged age, we need weighted sum. In these cases, the weights are pre-determined based on real data and knowledge. Area and population are often surveyed at a fine granularity by censuses and can be used as good proxies of the share.

#### 3.2.2 Mixed-Frequency Aggregate Learning Model

Predictor $f$ predicts the outcome values for small areas. In contrast to standard supervised learning, true values for the predictor can not be observed. In another word, the predictor predicts latent values. The main feature vector $x_t^q$ is of granular information such as search queries, posts in social networking service (SNS), point of sales (POS), credit card, and mobility data. These data are often called alternative data. The vector $x_t^q$ contains the current value $x_t^q$ and its lagged values. That is, $x_t^q = (x_t^q, x_{t-\tau}^q, x_{t-2\tau}^q, \ldots)$. Forecasters want to update the prediction when new data arrive. That is, $\tau$ can be any timing. For example, Let $t$ be April 2020 and $t_1, t_2, \ldots, t_{30}$ be April 1st, 2nd, \ldots, 30th. Then, $x_{t_1}$ and $x_{t_2}$ should give different prediction. However, the label data are only available for complete data $x_t^q = (x_{t_1}^q, x_{t_2}^q, \ldots, x_{t_1}^q)$. $x_{t_1}^q$ is missing in $x_{t_2}^q$. In economics, missing data is dealt with by either (i) training model on data with the same missing structure or (ii) imputation of missing data. In the above example, a forecaster who adopts (i) uses only data from the 1st to 29th day for each month to keep the same missing structure when training their model. Naturally speaking, it causes a huge loss of information. A forecaster who resorts to (ii) needs to prepare another model to conduct imputation.

In our model, the missing-ness is treated by introducing auxiliary feature vector $\phi$ and non-linearity in the parameters. The vector $\phi_t^q$ contains one-hot encoded year, month, day, larger area $p$, and small area $q$. By doing so, predictor $f$ can learn the missing structure of the data and appropriately use the information. The non-linearity of the parameter is essential in this sense. We adopt a simple recurrent neural network (RNN) to make the predictor flexible (Fig. 3). As a result, we can fully utilize the information and make it end-to-end.

#### 3.2.3 Model Training

The predictor is trained by minimizing the MSE loss,

$$
\mathcal{L}(f) = \sum_t \sum_{\tau \mu(t) = t} \left( \tilde{y}_t^q - \sum_{q:p(x_t^q, \phi_t^q)} \omega_q f(x_t^q, \phi_t^q) \right)^2. \tag{2}
$$

At each data point $(\tau, p)$, the predictor $f$ predicts the values for small area $q$ at short term period $\tau$. The predicted values are weighted-average and the loss is calculated.

#### 3.2.4 Prediction

Trained predictor $f$ is used for the prediction using granular data. Forecaster can use real-time data $x_t^q$ to predict the smaller area’s current status as $\tilde{y}_t^q = f(x_t^q)$. The good news is that the forecaster does not need to re-train the model until new label data $y$ is released. She can re-use the same model for an extended period.

### 4 EXPERIMENT WITH JAPANESE JOB APPLICATION DATA

We apply the MF-AGL model to analyze the real-world economy. In particular, we apply the model to predict the number of job applicants in Japan. The Japanese government releases official statistics on the number of persons who file job applications to public employment offices on monthly basis. In Japan, unemployed persons need to file a job application to the public employment office to take up the unemployment insurance benefits. The number of job applicants is counted at 544 public employment offices, the official statistics summarizes the number for 47 prefectures. The number of job applications is a good proxy for the number of unemployed persons. This is very similar to the unemployment insurance claims statistics in the U.S., which is considered one of the most important economic indicators by economists.

The reasons we chose the problem are the following. First, unemployment is a huge tragedy; It leads to loss of income and also loss of contact with society, which causes economic and mental hardship. Real-time analysis of labor market conditions is essential for a swift policy response. Second, the prefecture-level data provided by the reports are too rough for the appropriate policy response. In Japan,
We use three datasets as follows. We need a real-time update for the data. The points are artificially generalized and does not represent real human location history.

The whole process is summarized in Fig. 2. We first extract feature vector $z$ from raw GPS readings taken from mobile apps. Then the visit predictor predicts $x^v$, the number of visits to public unemployment offices located in each city. Then another predictor $f$ predicts $y^j$, the number of job applicants for each office. The visit predictor is trained on the different domain and transferred to the task. The transfer learning is explained in Section 4.3.

### 4.1 Datasets

We use three datasets as follows.

1. **Reports on Employment Service by Ministry of Health, Labour, and Welfare.** The reports on employment service is a monthly official statistics released by the Ministry of Health, Labour and Welfare (MHLW). The reports contain the monthly number of job applicants by prefecture. The data is publicly available on the webpage of MHLW [23].

2. **GPS data from Smartphones.** GPS data is taken from various mobile apps from Jan 1, 2016, to Oct 31, 2020. The data is anonymized and not related to privacy. The number of users ranges from several hundred thousand to several million. The data contains tuples (latitude, longitude, timestamp). The data contains no private information.

3. **Locations of Public Employment Offices.** The public employment offices are established by the Japanese government based on the ILO treaty. The location of the offices is publicly available on the webpage of the MHLW. We use the location with the GPS trajectories to extract feature values [22].

### 4.2 Feature Extraction from GPS Readings

The crucial challenge is to accurately count visitors from noisy GPS readings. GPS is very noisy especially for the smartphone because the logs are usually very sparse. For example, many apps log the GPS records when the phone is moved. This algorithm aims at minimizing battery consumption. As a result, the records are not recorded at the same intervals. Also, the accuracy of location deteriorates inside buildings as the signals from satellites are not reached. One possible solution is to rely on the machine learning techniques that denoise the data and extract visited POIs (Point-of-Interests). There are various approaches to do this job [11, 16, 29, 30]. We find that the visited point extraction methods typically rely on (i) the number of stay points, (ii) the stop location, (iii) stop duration, and (iv) speed. We follow their approach and extract extensively rich information from location data to achieve high performance. The list of the features is presented in Table 2.

The visit predictor in Fig. 2 uses location trajectory to classify visit/non-visit to some POIs.

### 4.3 Visits Prediction using Transfer Learning

1. **Transfer of Visit Predictor.** Another challenge for visit prediction is that there is no true label for public employment office visitors. We transfer the visit predictor trained on the different proprietary source and transfer to the visit prediction for the main task; visits to public employment offices.

   The visits predictor uses the extracted features as in Section 4.2. The predictor is implemented using LightGBM [15]. The hyperparameter tuning is done with the LightGBM tuner. These features are powerful. The visit predictor achieved an AUC of 0.86 for the classification task for the original task (train: test = 9,091 : 4,479).

2. **Visits Prediction for Public Employment Offices.** Unemployed persons visit nearest public employment office to file an unemployment insurance claim. There are 544 offices in Japan. We divide the entire country into 544 regions based on the location of the offices.
assuming each region is covered by the nearest office. We use the transferred predictor to predict the visit count to each office.

### 4.4 Job Applicants Predictor and Aggregation Function

With the predicted visits count, we train the job applicants predictor that predicts the number of job applicants for each public employment office. Input to the model is described in Table 3. The predictor uses past visit count and dummy variables extracted from the date of making a prediction.

**Table 3: List of Input Variables in Job Applicants Predictor**

| feature             | description                                         |
|---------------------|-----------------------------------------------------|
| visit counts        | Predicted daily visit counts for past 31 days.      |
| year dummy          | One-hot encoded year dummy                          |
| month dummy         | One-hot encoded month dummy                         |
| day dummy           | One-hot encoded day dummy                           |
| prefecture dummy    | One-hot encoded prefecture dummy                    |
| public office dummy | One-hot encoded office dummy                        |

The training/prediction model is a simple recurrent neural network, and Fig. 3 shows the network architecture. This model uses an LSTM layer and a multi-layer perceptron to predict a daily visit count of small areas, and each prediction is aggregated to predict a monthly visit count of large areas according to Eq. (1). We use Adam [17] solver for optimization with $\beta_1 = 0.9$, $\beta_2 = 0.999$, initial learning rate = 0.0001, no weight decay and no learning rate decay. We train our models for a total of 600 epochs with a batch size of 256. The other finding is that the AR and RF models are not good at prediction. To see this, we first aggregate the prediction by MF-AGL model at prefecture-level and calculate Mean Absolute Percentage Error (MAPE), a standard metric to evaluate forecasting models.

$$\text{MAPE} = \frac{1}{|P|} \sum_{p \in P} \frac{|y^p_t - \hat{y}^p_t|}{y^p_t}$$

The reason we use MAPE is that the number of job applications in each prefecture is proportional to the population. Other metrics such as mean squared error (MSE) and mean absolute error (MAE) are more affected by an error in the populous prefecture while MAPE treats each prefecture equally.

The other finding is that the AR and RF models are not good at prediction. To see this, we first aggregate the prediction by MF-AGL model at prefecture-level and calculate Mean Absolute Percentage Error (MAPE), a standard metric to evaluate forecasting models.

$$\text{MAPE} = \frac{1}{|P|} \sum_{p \in P} \frac{|y^p_t - \hat{y}^p_t|}{y^p_t}$$

The results are shown in Table 5. The MF-AGL obtained the lowest error. Although the prediction performance at the aggregate
Figure 3: Our aggregate learning architecture for job applicants prediction. We use a simple recurrent neural network composed of an LSTM layer and a multi-layer perceptron as a predictor. Each daily prediction is aggregated as monthly visits count of large area. The loss value is computed by the MSE criterion on the training phase.

Figure 4: Year-over-year change in the number of job applicants in Japan. Due to the limitation of the space, we exclude Okinawa islands from the map. The color represents good (lighter)-bad (darker) conditions of labor market. Ground Truth shows the actual data from official statistics; AR shows the prediction made by the AR model with 11 lags; RF shows the predictions made by Random Forest model with 11 lags and year, month and prefecture dummy as input; MF-AGL shows the prediction made by MF-AGL model. MF-AGL model shows the much granular predictions.

Table 4: The schedule of data availability

|        | Oct 1  | Oct 2, · · · , Oct 29 | Oct30 | Oct31 |
|--------|--------|------------------------|--------|--------|
| output | $y_p \cdot \cdot \cdot , y_{ily}$ | $y_p \cdot \cdot \cdot , y_{Aug}$ | $y_p \cdot \cdot \cdot , y_{Sep}$ |        |
| input  | $x_{Oct} \cdot \cdot \cdot , x_{Oct1}$ | $x_{Aug} \cdot \cdot \cdot , x_{Aug1}$ | $x_{Sep} \cdot \cdot \cdot , x_{Sep1}$ |        |
| AR/RF  | $y_{Oct} \cdot \cdot \cdot , y_{ily}$ | $y_{Oct} \cdot \cdot \cdot , y_{Aug}$ | $y_{Oct} \cdot \cdot \cdot , y_{Sep}$ | $y_{Oct} \cdot \cdot \cdot , y_{Sep}$ |
| MF-AGL | $x_{Oct} \cdot \cdot \cdot , x_{Oct1}$ | $x_{Aug} \cdot \cdot \cdot , x_{Aug1}$ | $x_{Sep} \cdot \cdot \cdot , x_{Sep1}$ | $x_{Oct30} \cdot \cdot \cdot , x_{Oct31}$ |

level is not the priority of the MF-AGL model, this result highlights the robustness of the prediction of the model.

4.6 Real-time Labor Market Analysis using MF-AGL Model

Finally, we demonstrate another practical usefulness of our model. As discussed in Section 3.1, the beauty of our model is real-time updates of forecasting using granular data. Fig. 5 shows how the model updates the prediction using the real-time data. The model shows the year over year for each city and for each day. That is, we first predict the disaggregate number of job applicants on the same day of the previous year and calculate the year-over-year change in the number of job applicants. The figure implies the rapid improvement in the labor market during October 2020. The
possible reason is the peak-out of the second wave of the COVID-19 pandemic. As shown in Fig. 6, the number of cases was dramatically decreased in September and stable in October.

5 CONCLUSION

In this work, we proposed a novel aggregate learning method that adopts to the mixed-frequency data. The model predicts spatio-temporal changes in the economic indices without granular label data. We proposed an LSTM-based simple architecture and the loss function for the training.

We applied the model to a real-world task that is the prediction of the number of job applicants in Japan. Our MF-AGL model well predicts the regional heterogeneity at the sub-prefecture level and also the rapid change in the labor market condition during a month.

The present model can be applied to broad areas including GDP prediction, labor market prediction, and industrial production prediction. The direction of the future work can be the extension to the other domain and more improvement of the model using more fine-grained architectures.

ACKNOWLEDGMENTS
Figure 5: Changes in the number of job applicants in Japan. Due to the limited space of the paper, we exclude Okinawa islands from the map. The color represents good (lighter)-bad(darker) condition of the labor market from the previous year. The prediction is done by MF-AGL model uses the input available at each day.

Figure 6: The number of cases (PCR positive) of COVID-19 in Japan.
REFERENCES

[1] Katharine G. Abraham, Ron S. Jarmin, Brian Moyer, and Matthew D. Shapiro. 2019. Big Data for 21st Century Economic Statistics. (March 2019).
[2] Apple Inc. 2020. COVID-19 - Mobility Trends Reports - Apple. https://webcache.googleusercontent.com/search?q=cache:9t5e-S9-MxVDYJ covid19.apple.com/mobility#1cd1kja3lrc5hncgdpj-gp. 
[3] Nikos Askitas and Klaus F. Zimmermann. 2009. Google Econometrics and Unemployment Forecasting. (2009).
[4] Jennie Bai, Eric Ghysels, and Jonathan H. Wright. 2013. State Space Models and MIDAS Regressions. Econometric Reviews 32, 7 (June 2013), 799–813. https://doi.org/10.1080/07474938.2012.690675
[5] Hyunyoung Choi and Hal Varian. 2012. Predicting the Present with Google Trends. Economic Record 88, s1 (2012), 2–9. https://doi.org/10.1111/j.1475-4932.2012.00869.x
[6] Francesco D’Amuri and Juri Marcucci. 2010. ‘Google It!’ Forecasting the US Unemployment Rate with A Google Job Search Index. SSRN Electronic Journal (2010). https://doi.org/10.2139/ssrn.1594132
[7] Guillaume Derval, Frédéric Docquier, and Pierre Schaus. 2020. An Aggregate Learning Approach for Interpretable Semi-Supervised Population Prediction and Disaggregation Using Ancillary Data. In Machine Learning and Knowledge Discovery in Databases (Lecture Notes in Computer Science), Ulf Brefeld, Elisa Fromont, Andreas Hotho, Arno Knobbe, Marloes Maathuis, and Céline Boraulet (Eds.). Springer International Publishing, Cham, 672–687. https://doi.org/10.1007/978-3-030-46333-1_40
[8] Jaquenon K. Galimberti. 2020. Forecasting GDP Growth from Outer Space. Oxford Bulletin of Economics and Statistics 82, 4 (2020), 697–722. https://doi.org/10.1111/obes.12361
[9] Eric Ghysels. 2016. Macroeconomics and the reality of mixed frequency data. Journal of Econometrics 193, 2 (Aug. 2016), 294–314. https://doi.org/10.1016/j.jeconom.2016.04.008
[10] Eric Ghysels, Arthur Sinko, and Rossen Valkanov. 2007. MIDAS Regressions: Further Results and New Directions. 26, 1 (2007), 53–90. https://doi.org/10.1097/01.jeconometrics.2007.07.020
[11] Lei Gong, Takayuki Morikawa, Toshiyuki Yamamoto, and Hitomi Sato. 2020. COVID-19 Mobility Data Network. 2020. Facebook Data for Good Mobility Dashboard. https://www.covid19mobility.org/dashboards/facebook-data-for-good/
[12] Kyosuke Nishida, Hiroyuki Toda, and Yoshimasa Koike. 2017. Extracting Arbitrary-shaped Stay Regions from GeospatiaL Trajectories with Outliers and Missing Points. In Proceedings of the 8th ACM SIGSPATIAL International Workshop on Computational Transportation Science - IWCTS’15. ACM Press, Seattle, WA, USA, ı–6. https://doi.org/10.1145/3084388.3084984
[13] Kyosuke Nishida, Hiroyuki Toda, Takeshi Kurashima, and Yoshihiko Suhara. 2014. Probabilistic identification of visited point-of-interest for personalized automatic check-in. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - Ubicomp ’14 Adjunct. ACM Press, Seattle, Washington, 631–642. https://doi.org/10.1145/2632048.2632092
[14] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Advances in Neural Information Processing Systems 32 (H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett (Eds.)). Curran Associates Inc., 8024–8035. https://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf
[15] Jaroslav Pavliček and Ladislav Kristoufek. 2015. Nowcasting Unemployment Using Deep Internal Learning. In Proceedings of the Fourth International ACM Workshop on Managing and Mining Regional Science and Urban Economics (NIPS’17). Curran Associates Inc., 15076–15086. https://proceedings.neurips.cc/paper/2019/hash/64517d8435994992e682b3e4aa0a0661-Abstract.html
[16] David R. Musicant, Janara M. Christensen, and Jamie F. Olson. 2007. Super-Regional Decision Tree. In Proceedings of the National Conference on Data Mining (ICDM 2007). Springer International Publishing, Cham, 21–33. https://doi.org/10.1007/978-3-030-38861-6_3
[17] Matteo Mogliani. 2019. Bayesian MIDAS Penalized Regressions: Estimation, Selection, and Prediction. arXiv:1903.08025 [cond-mat] (March 2019). http://arxiv.org/abs/1903.08025
[18] Daisuke Morikawa. 2020. Nowcasting Unemployment Rates with Smartphone GDP Data: In Multiple-Aspect Analysis of Semantic Trajectories (Lecture Notes in Computer Science), Konstantinos Tserpes, Chiara Renso, and Stan Matwin (Eds.). Springer International Publishing, Cham, 21–33. https://doi.org/10.1007/978-3-030-38861-6_3
[19] Daisuke Morikawa, Soichiro Harada, Jiyun Schneider, and Takahiro Hoshino. 2020. Nudging Preventive Behaviors in COVID-19 Crisis: A Large Scale RCT using Smartphone Advertising. Technical Report 2020-021. Institute for Economics Studies, Keio University. https://ideas.repec.org/p/keo/dpapper/2020-021.html
[20] Tanya Suhoy. 2008. Query Indices and a 2008 Downturn: Israeli Data. (2008), 34. https://doi.org/10.1145/1897494.1897503
[21] MARKETPLACE. 2020. We Need to Change Our Economic Indicators to Keep up with the Crisis. https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/keizai/keizai.html
[22] Ministry of Health, Welfare, and Labour. [n.d.]. Location of Hellowek. https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/keizai/keizai.html
[23] Kyosuke Nishida, Hironobu Toda, and Yoshimasa Koike. 2017. Extracting Arbitrary-shaped Stay Regions from Geospatial Trajectories with Outliers and Missing Points. In Proceedings of the 8th ACM SIGSPATIAL International Workshop on Computational Transportation Science - IWCTS’15. ACM Press, Seattle, WA, USA, 1–6. https://doi.org/10.1145/3084388.3084984
[24] Matteo Mogliani. 2019. Bayesian MIDAS Penalized Regressions: Estimation, Selection, and Prediction. arXiv:1903.08025 [cond-mat] (March 2019). http://arxiv.org/abs/1903.08025
[25] Hyunyoung Choi and Hal Varian. 2012. Predicting the Present with Google Trends. Economic Record 88, s1 (2012), 2–9. https://doi.org/10.1111/j.1475-4932.2012.00869.x