Predicting Temporal Patterns in Keyword Searches with Recurrent Neural Networks — Phenotyping Human Behaviour from Search Engine Usage

1st Jay Paul Morgan
Department of Computer Science
Swansea University
Swansea, United Kingdom
0000-0003-3719-362X

2nd Frederic Boy
NAIADES Research Network
Zienkiewicz Institute, Swansea University
Swansea, United Kingdom
0000-0003-1373-6634

***Abstract*—Every day, millions of people use search engines to find relevant information using various keywords and queries. These keywords may provide vital clues into our behaviour on a societal level and point to specific indicators of economics and well-being. In this work, we examine the search volume data of 61 keywords from Google Trends, finding there are three main morphological patterns of how keywords are used in searches throughout the day. The search volume for these 61 keywords are compared using Dynamic Time Warping and categorised with hierarchical clustering, while the 24-hour time series patterns are learnt by a Recurrent Neural Network (RNN). The performance of this RNN is analysed using two experiments to test its ability to generalise to different types of keywords and to different dates. If integrated with an overarching system, this RNN could track societal well-being and inform policies to tackle underlying societal issues.**

***Index Terms*—Recurrent neural network, Long short-term memory neural networks, Sequence Classification, Signal Processing**

I. Introduction

Macroeconomic indicators (*e.g.* Gross Domestic Product (GDP) and inflation) are periodically used to assess the economic performance of nations, but if they are extremely useful to policymakers and business-people, they have serious shortcomings as a substitute for evaluating societal advancement. For instance, it has been recognised for almost a century that GDP, which gauges overall economic output, fails to capture the complexity of human flourishing [25] and by only quantifying market-based transactions, omits non-pecuniary determinants of well-being like leisure time, inequalities in income distribution, access to health services, and environmental integrity [21] in advanced economies and often diverges from citizens' sense of progress. This measurement limitation has spurred calls to supplement GDP with more humanistic and holistic indicators that are able to capture experiential prosperity beyond material consumption and production output.

However, constructing metrics that align more closely with people's self-reported life satisfaction remains an ongoing challenge [13] and most existing well-being metrics depend on census survey data that lag in timely relevance and risk human response biases, undermining robustness. To solve this issue, common to many research efforts in the social sciences, and increase the temporal frequency at which well-being metrics are measured, we took to validate the use of digital epidemiology on keyword searches trends time-series on the Google search engine.

Google Trends provides a unique perspective by analysing the keywords people are using. Google Trends is a service which allows queries for search keywords or topics, returning a normalised time series index of the volume of traffic relating to these search keywords in a geographical area [6]. To date, Google Trends data have been used in a wide range of research which -examined, for instance, the impact of COVID-19 lockdowns on well-being concerns [4], [11], [24], -carried out disease surveillance [1], [5], [17], or -explored other public health related topics [10]. Recent studies investigated whether data from Google Trends could predict levels of self-reported mental well-being. [11] found that search volumes for the keyword "loneliness" in the US co-varied with self-reported loneliness from cross-sectional data covering the same period. Robust evidence that search volumes time-series for well-being related keyword queries, sampled daily and weekly, were correlated with measures of population-level well-being was provided by [3]. They replicated a series of independent correlations between national well-being surveys, the EARS AI-supported platform's measurement of Social Listening developed by the World Health Organisation (WHO) [31] and Google Trends data for specific well-being keywords (*anxiety*, *stress* and *depression*). The combination of Google Trends data with econometrics-standard objective indices (albeit some are based on self-report) has therefore the potential to afford researchers and policy-makers with a promising new means of accelerating the detection of rapid societal change.

The use of Google Trends data to facilitate the training of a Machine Learning model to perform a high frequency sampling of economic and well-being indicators provides a unique opportunity for informing policy makers. In this work, we begin this process by first analysing a set of keyword search data from Google Trends, and learning patterns of search volume with a Recurrent Neural Network (RNN). This RNN is used to model trends and patterns of keywords for hourly data sampled from Google Trends. This research then provides a foundation for future work to tightly integrate this RNN with economic and well-being indicators in a geographical monitoring framework.

The rest of this paper is organised as follows: in Section II, we look at existing works that use time series data & Google Trends. In Section III the format of the data and the architecture of the model considered in this study is explained. Section IV demonstrates the experimental framework and shows the results. We give a discussion in Section V and our conclusions in Section VI.

II. Related Works

RNNs are primarily used for time-series forecasting and sequence modelling tasks due to their architecture, which is optimised for processing sequential data [14]. They have been used in a range of contexts, as in language modelling [27], stock price prediction [18], activity recognition or classification [16], and anomaly detection [8]. RNNs have also been used successfully for a variety of other niche tasks. For example, [20] used a combination of convolutional layers and Gated Recurrent Units (GRU) to recognise the emotions from speech wave, forms. [19] use Long Short-Term Memory Units (LSTM) in an encoder-decoder architecture to compress video into a lower file size without compromising on video quality. [23] used an RNN to perform denoising and up-sampling.

Few works have used RNN in combination with Google Trends. [26] selected a series of keywords from 2004-2019, and after using the Shapely Additive explanations (SHAP) algorithm to reduce to a smaller set of useful keywords, trained an LSTM to forecast and predict unemployment claims in the US. Their results show that an LSTM is able to outperform the typical Vector Autoregressive Models on this task. [15] used an LSTM to predict influenza infections. Other work show that a can LSTM replaced by a Seq2Seq+Attention model for better performance in influenza infection prediction using Google Trends data [12]. More recent work has continued this trend of infection prediction, where [22] used an LSTM to forecast cases of COVID-19 in India, the UK, and the US. Other work investigates the migration task, [9] used an LSTM to predict migration patterns using 67 migration related keywords.

While the previous state of the art mainly focused on trend forecasting and prediction, it motivates the present attempt to use RNN for classifying temporal patterns of search engine keyword queries. To the best of our knowledge, this is the first attempt to characterise phenotypical patterns of online information search at an hourly granularity. We foresee the use of these patterns to abstract away from the specific keywords, and therefore may be able to transfer knowledge and extend methodologies to different languages and countries

III. Data & Methods

Google Trends is a publicly available data service provided by Google Inc. that allows users to access the temporal dynamics of keyword-based internet search volume. Google Trends provides access to a single measurement: the relative Search Volume (RSV), a normalised metric relative to the temporal- and geographical remits of a specific interrogation of the database, such that the values range from 0 to 100 over the dates and geography the user has requested. Due to limitations provided by Google Trends it is not possible to query for significantly large time frames of search volumes, therefore the user will often perform multiple queries for different times and perform a post-processing method to make the normalisation consistent across queries such as performed in [4], [5]. This post-processing stage is necessary to analyse changes in search volume across large time frames. However, unlike [4], [5], the present research is not concerned by change of search volume across time, but rather, by the morphology of search volume over shorter 24-hour or 168-hour intervals. Therefore, in the present case, scaling methods are not necessary.

A. Dataset

To facilitate the training of an RNN to predict patterns in the search volumes of keywords, a set of 61 keywords are sampled from Google Trends[1] at an hourly frequency over 168-hours from over a period of 5 years from 2018-01-01 to 2023-01-29 23:00:00.

The keyword selection is based on established theoretical frameworks in affective psychology and well-being research. Specifically, we draw upon the seminal works of [30] on mood and affect, [28] research on emotional experiences, and [7] studies on well-being measures. This grounding in recognised theories ensure our keyword choices are academically rigorous and relevant to well-being indicators. Our study uses a comprehensive set of 61 emotion-related and neutral words in the English language, derives from multiple validated psychological assessment tools. The emotion words are drawn from instruments such as the Positive and Negative Affect Schedule Extended Scale (PANAS-X) [29], the International Positive and Negative Affect Schedule Short Form Scale (I-PANAS-SF) [28], and the Scale of Positive and Negative Experience (SPANE) [7], among others. This diverse selection ensures a broad coverage of affective states, enhancing the robustness of our analysis in capturing various dimensions of emotional experience.

An example of the search volume for the keyword *Why* is presented in Fig. 1. Here, we observe a common pattern of increased search frequency during certain hours of the day (Fig. 1 (b)). These patterns appear to be consistent over the week (Fig. 1 (a)) with some differences for the weekend. These patterns of search frequency are not limited to just the keyword *Why* but are seen across all keywords in our dataset.

Averaging the volume data over the week, we find three morpholigical patterns of search volume emerge for all keywords. In Fig. 2 we show the three patterns in the search volume data. In one pattern (labelled B), the search volume increases during the midday period. While for the two other

[1]For sampling the data, we use the gtrendsR library (https://github.com/PMassicotte/gtrendsR).

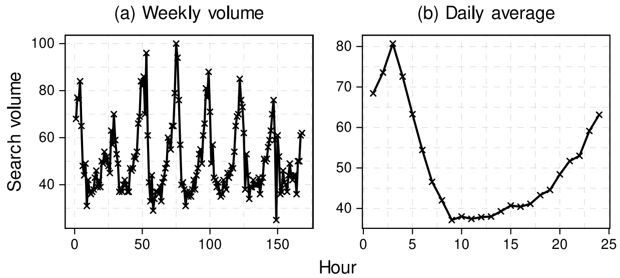

Fig. 1. Search volume for the keyword *Why*. In (a), we show the search volume over the entire week, whereas in (b) the daily average.

shapes (labelled A and C), the most traffic occurs during nighttime. There is a key difference between A & C, however. First, with pattern A, the volume increases mostly in the early hours of the morning and slowly increases during the nighttime hours. With pattern C, the search volume increases during the evening, resulting in high traffic during the nighttime hours.

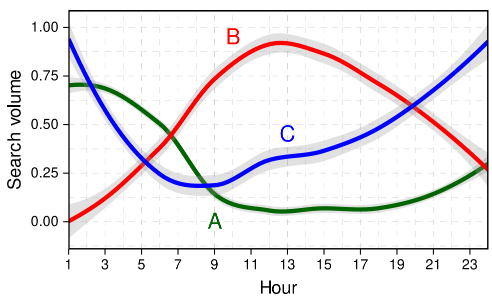

Fig. 2. Three patterns of search trends.

The existence of these morphological patterns in search volume data reinforced the notion of seasonality to which people search using particular keywords, and motivated the use of an RNN to learn from noisy data (Fig. 1 (a)) to predict one of the three patterns in Fig. 2.

To compute the associated search volume pattern for each keyword, i.e. to which of the three morphological patterns does each keyword correspond to, we first impute missing values using a combination of linear interpolation and closest value pairing. This first step addresses missing values through linear interpolation using adjacent data points. For gaps at the time-series' start or end, we use the nearest available value (see Fig. 3)[2]

Next, using the Dynamic Time Warping (DTW) algorithm [2], we generate of a pairwise distance matrix of similarity scores between daily patterns for each keyword. The DTW algorithm measures the similarity between sequences (where lower values indicate a higher degree of similarity). While DTW is an effective method to account for differences in temporal variation between two time-series, DTW can also accurately measure similarities when the there is no temporal variation, such as in our data where the sampling frequency

[2]We have released the source code under an open source licence. This code can be found at (anonymised).

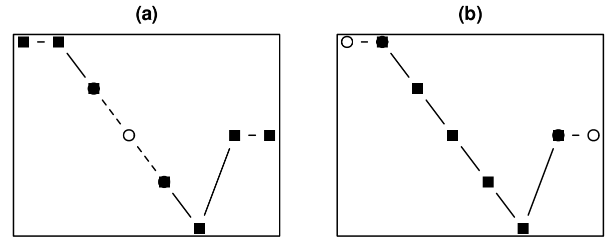

Fig. 3. Imputation of missing values. The missing values visualised by a white-filled circle is interpolated using the surrounding black-filled squares. In (a), the interpolation considers the previous and next values. In (b), the closet non-missing value is used.

TABLE I
NUMBER (PROPORTION) OF SAMPLES PER PATTERN.

Pattern	N = 113,155
A	66,780 (59%)
B	18,550 (16%)
C	27,825 (25%)

is consistent across keywords. Using DTW, keywords with a similar morphological pattern will have a smaller distance, while different patterns will have a larger distance. DTW is applied to the mean daily average of each keyword, against every other keyword, thereby generating a 61×61 distance matrix.

Using the pairwise distance matrix, we perform hierarchical clustering using the distance matrix. Keywords with small distances to one another (and thereby by extension will have a similar pattern) are clustered/grouped together. Finally, to label each of the keywords into one of the three patterns, we selected three clusters as the cut-off value for the hierarchical clustering algorithm.

Following this process, for each of the 61 keywords a corresponding label (A, B, & C) is assigned, resulting in the dendrogram shown in Fig. 4. We observe there is a common theme between clustered keywords. For example, pattern A consists of many terms associated with emotions such as *Angry*, *Amazing*, and *Annoyed*. Pattern B contains economic related terms such as *HMRC*, *Money*, and *Mortgage*. This pattern also includes many neutral keywords such as *Plant*, *Table*, *Flower*. Meanwhile, pattern C contains more negatively associated terms such as *Depression*, *Sad*, *Fear*, and *Cancer*. These negative terms, in particular, are increasingly searched for in the early hours of the morning and night, providing some insight into the condition of those individuals who are using these terms as part of their internet searches. There appears to be some overlap in themes between patterns A & C, which is consistent with the similarity in the pattern morphology. Therefore, in our experiments, we consider both a 3-class prediction task (A, B, & C), and a 2-class task where keywords labelled C are collapsed into A, thereby predicting between B and not B (A+C, & B).

With a pattern label assigned to the keywords, each 168-hour sample is split into 24-hour windows, resulting in 113,155 samples. The number and representative proportion for each of the three pattern types in shown in Table I.

Fig. 4. Hierarchical clustering of keywords.

B. Model Architecture

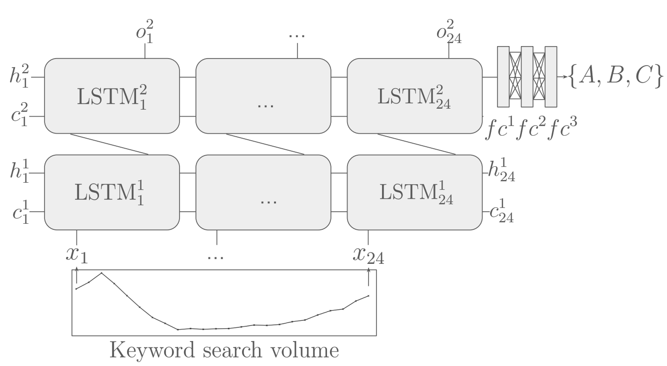

Fig. 5. Architecture of the RNN.

To predict the search volume pattern given a 24-hour window, we construct an RNN as illustrated in Fig. 5. Our RNN is composed of two stacked layers of LSTM cells that process each hour of the 24-hour window sequentially while maintaining a hidden and cell state. After processing the entire sequence, the final hidden state of the last LSTM layer is projected to a classification label (A, B, or C) via a series of fully-connected layers with ReLU activation functions.

The RNN is built using R 4.4.1 with torch 0.13.0, running on an Nvidia 3070 GPU. The network is trained for 10 epochs with a batch size of 512. The model weights are optimised using the Adam optimiser with its default parameters. At the end of each epoch, the model is evaluated on a validation set. The model weights with the lowest loss on the validation set are used for evaluation on the testing set.

This architecture serves as a baseline for assigning search volume patterns to daily search frequencies. We do not perform hyperparameter tuning, therefore there may be some changes to the training which yield better model performance in some experiments. While this model may not give the best possible results, it may be iterated upon in future work.

IV. Experiments

In this work, we perform two main experiments to evaluate the RNN under a range of conditions and scenarios. The first experiment investigates how well an RNN generalises to different keywords. The second experiment evaluates the model's ability to generalise to different dates. This experiment may be pertinent given that a subset of the data was sampled during the COVID-19 pandemic, which had a measured effect on the economic and behavioural indicators [4].

A. Generalisation to Different Keywords

While we have a wide range of different types of keywords, generalising to different keywords is a desirable property of any Machine Learning model. Therefore, in this experiment, we perform a 5-fold cross validation on the keywords (*i.e.* the search volume data for a single keyword is either in the training or testing set, but not both). We use stratified sampling for the splitting of folds to ensure a consistent proportion of pattern types are in the testing set and minimise skewed results as a result of under-representation. The results are then averaged over the 5 folds and, where applicable, are formatted as: mean (standard deviation). We report both the Accuracy (%) and F_1 score.

We report the results for training and testing on different keywords in Table II. Here we observe the RNN achieves an accuracy of 77.64% (Table II row 1) when predicting over 3-classes (A, B & C). If we take the most commonly predicted pattern over all 24-hour windows for each keyword, the accuracy score increases to 86.79% (Table II row 2) which means out of the 61 keywords predicted (over all folds), then roughly 53 keywords are predicted correctly while 8 are not. Most notably, we see through the visualisation of the Receiver-Operating Curves (ROC) (Fig. 6) that the RNN has more difficulty in predicting for classes A & C in one of the folds (fold 4). If we investigate these failure cases for each fold (Table III) we see the majority of misclassifications are between the A & C patterns. This result may be expected, as patterns A and C share a similar morphological profile, with

TABLE II
ACCURACY AND F_1 SCORES FOR OUR RNN TRAINED OVER 5 K-FOLDS. RESULTS ARE SHOWN FOR BOTH 3-CLASS CLASSIFICATION OF THE 3 PATTERNS, AND FOR THE BINARY CLASSIFICATION WHERE THE A & C PATTERNS ARE COMBINED INTO A SINGLE CLASSIFICATION LABEL. ROW NUMBERS 1 TO 8 HAVE BEEN ADDED TO HELP FOR REFERENCE IN THE TEXT.

	Target Patterns	Criteria	Accuracy (%) Mean (SD)	Accuracy (%) Range	F_1 Mean (SD)	F_1 Range
Experiment 1 - Generalisation to keywords						
1	3-class (A, B, & C)	24-window	77.64 (8.208)	0.6757 - 0.8874	69.68 (14.314)	0.5042 - 0.8569
2		By keyword	86.79 (9.590)	0.7500 - 1.0000	86.95 (11.044)	0.7026 - 1.0000
3	2-class (A+C, & B)	24-window	93.56 (4.189)	0.8766 - 0.9940	91.89 (4.767)	0.8764 - 0.9923
4		By keyword	86.79 (9.590)	0.7500 - 1.0000	86.95 (11.044)	0.7026 - 1.0000
Experiment 2 - Generalisation to dates						
5	3-class (A, B, & C)	24-window	84.12 (1.139)	0.8255 - 0.8559	80.39 (1.706)	0.7800 - 0.8250
6		By keyword	92.46 (1.466)	0.9016 - 0.9344	91.57 (1.990)	0.8916 - 0.9344
7	2-class (A+C, & B)	24-window	96.21 (1.358)	0.9441 - 0.9772	95.55 (1.361)	0.9381 - 0.9705
8		By keyword	92.46 (1.466)	0.9016 - 0.9344	91.57 (1.990)	0.8916 - 0.9344

TABLE III
INCORRECTLY PREDICTED KEYWORDS OVER 5 FOLDS.

Keyword	Predicted	Target
Cool	B	C
Depression	A	C
Disease	A	C
Fear	A	C
Flower	B	C
Sad	A	C
Sleep	C	A
Tree	A	B

the main difference being the amount of increased searches in the evening/night hours for pattern C.

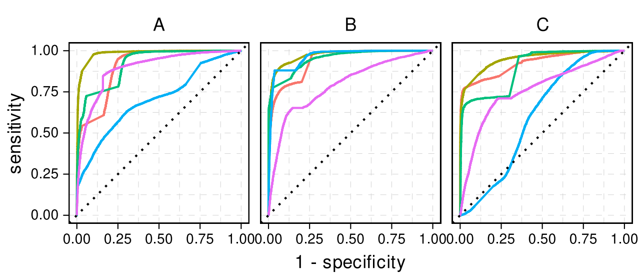

Fig. 6. ROC curves over 5-fold cross validation.

In one failure case, *i.e.* the *Cool* & *Flower* keywords, the RNN predicts pattern B, whereas the calculated pattern is C. Looking at the average daily search volume for these keywords (Fig. 7 (a) & (b)), it is clear these curves may not easily be classified into one of the three patterns due the very quick decrease of searches at hour 5 and the sharp increase thereafter. These two curves only slightly resemble pattern C, but also contains the high search index during the daytime. Therefore, it is understandable why the RNN has made a misclassification in these situations. In other cases where the RNN has made incorrect predictions, such as Fig. 7 (c), the RNN does not correctly distinguish between patterns C & A since they are very close in similarity. If we were to collapse the classification C into A to form a single class (thereby performing a binary classification task between A+C and B) we get the performance increase considerably for the 24-hour window up to 98.38% (Table II row 3). This model is of course more accurate due to there being a much more clear distinction between pattern B and the other two patterns.

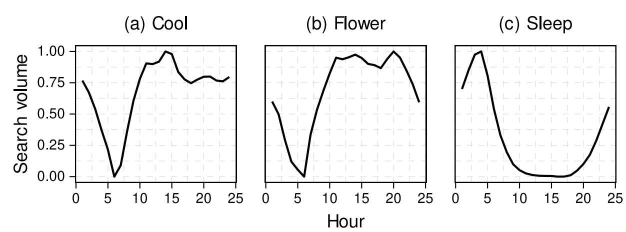

Fig. 7. Example search index curves for 3 different misclassified keywords: (a) Cool, (b) Flower, and (c) Sleep.

In general, we see the RNN is able to predict the pattern well given a 24-hour window, especially for the 2-class task. However, in some cases for neutral keywords (such as *Cool* and *Flower*), where this pattern is not easily distinguishable.

B. Generalisation to different dates

In this experiment, we train the RNN on all keywords and the train/test split is made upon the date. We perform 5-fold cross validation, meaning about 53 sequential weeks are used for testing, while the rest of the dates are used for training.

The results for this experiment are presented in Table II for both 3-class and 2-class tasks (rows 5-8). When predicting using a 24-hour window (row 5), the RNN achieves an accuracy score of 84.12% which is roughly 223 weeks out of the total 265. If we collapse pattern C into A (row 7), the accuracy increases to 96.21% (256 weeks out of 265). While these incorrectly predicted 24-hour windows do not occur sequentially, it does indicate that the RNN is able to predict well for a significant proportion of the data.

In Fig. 8 (a) the most commonly incorrect keywords are from pattern C, with *Sad* having the highest frequency of

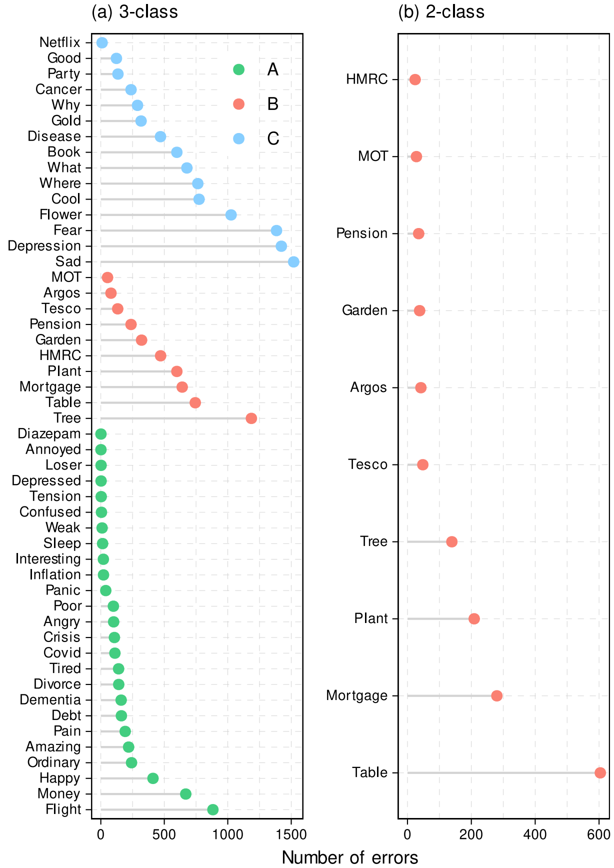

Fig. 8. Number of incorrect pattern prediction per keyword. Figure (a) shows the incorrect predictions for the 3-class task (A, B, C), while figure (b) shows the incorrect predictions with 2-classes (A+C, B).

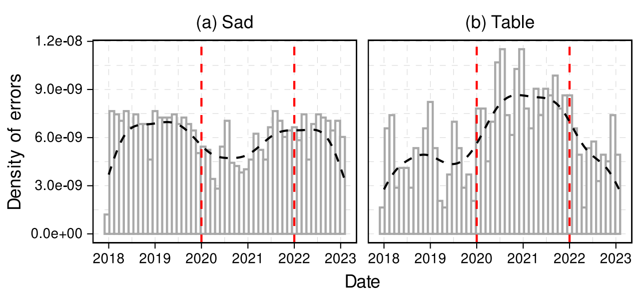

Fig. 9. Density of misclassifications for (a) *Sad* and (b) *Table* keywords. Red dashed lines shows the start/end of heightened period of COVID-19 activity.

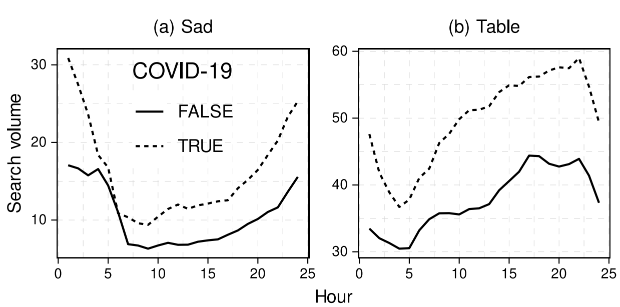

Fig. 10. Comparison of search index for the keyword (a) *Sad* and (b) *Table* during COVID-19 lock-downs and otherwise.

incorrect predictions. Fig. 9 (a) shows that the *Sad* keyword is more accurately predicted during the COVID-19 pandemic than other dates. While the search volume pattern of the *Sad* keywords looks similar if we compare the volume of searches for during and outside the pandemic lock-downs (Fig. 10), the range of values is much higher during these lock-downs. With the 2-class scenario (Fig. 8 (b)), it is apparent that the only incorrect predictions are made for the B pattern, most prominently is the *Table* keyword. Fig. 10 (b) shows an increase for searches using *Table* during COVID-19 lock-downs than other dates, while the morphology of the search index remains the same. Fig. 9 (b) shows there are more incorrect predictions for the *Table* keyword from 2020-2021. From both of these cases, it becomes more clear that the RNN has difficulty in predicting the patterns when the level of search indexes increases or decreases, even if the pattern is consistent.

V. Discussion

Economic and well-being indicators, which provide crucial insights into a nation's prosperity and associated citizens' quality-of-life, are notoriously expensive to measure accurately. This high cost leads to infrequent data collection, often resulting in outdated or incomplete information for policymakers and researchers. As detailed in Section I, this limitation poses significant challenges for real-time economic analysis and social policy formulation, potentially hindering effective decision-making and timely interventions. Google Trends and keyword search volume analysis offer policymakers access to high-frequency data, enabling more informed and timely decision-making. This innovative approach provides a continuous stream of real-time insights into public interests and concerns. This data-driven approach empowers policymakers to craft agile, evidence-based strategies that respond effectively to shifting societal dynamics and emerging challenges. This research establishes a foundation for developing a Machine Learning framework to examine keyword trends. Our RNN model possesses demonstrable proficiency in categorising 24-hour periods into three distinct morphological patterns. This approach offers a scalable method for automated trend classification in large-scale search data. The developed RNN model can function as an anomaly detection system by identifying deviations from established pattern classifications for known keywords. This capability enables the detection of unusual trends or shifts in search behaviour, potentially alerting analysts to emerging issues or changes in public interest. Such an application could enhance real-time monitoring and early warning systems across various domains. The present analysis shows that while the RNN performs well, it struggles with changes in search volume. More work is needed to improve the model's ability to handle these fluctuations. This improvement would make the RNN more reliable for real-world use.

VI. Conclusion

Our study analyses hourly search volumes for 61 keywords on Google Trends over a five-year period. The analysis reveals three primary patterns of daily seasonality in the search data.

Using the DTW algorithm, we categorise each of the keywords into one of the three identified morphological patterns. We then train an RNN to learn from and classify these patterns automatically. This approach enables efficient categorisation of search trends and lays the groundwork for automated pattern recognition in large-scale search data. We assess the performance of the RNN through two experiments, demonstrating its ability to generalise across diverse keywords and date periods. Our evaluations reveal that the RNN effectively categorises two distinct patterns (A+C and B). However, distinguishing between patterns A and C proved more challenging due to their morphological similarities. This insight highlights both the strengths and limitations of our current approach to pattern classification in search trend data. The challenge of distinguishing between patterns A and C is further compounded by significant fluctuations in search volume, such as those observed during the peaks of the COVID-19 pandemic (e.g., lock-downs in the UK). These volume changes can lead to increased misclassifications. Future research may examine techniques to enhance the models' robustness against volume fluctuations, potentially improving classification accuracy during periods of exceptionally atypical search behaviour.

Acknowledgements

This work was funded by Swansea University. For the purpose of Open Access, the author has applied a CC BY licence to any Author Accepted Manuscript (AAM) version arising from this submission.

References

[1] Kevin M. Bakker, Micaela Elvira Martinez-Bakker, Barbara Helm, and Tyler John Stevenson. Digital epidemiology reveals global childhood disease seasonality and the effects of immunization. *Proceedings of the National Academy of Sciences*, 113:6689 – 6694, 2016.

[2] Richard Bellman and Robert Kalaba. On adaptive control processes. *IRE Transactions on Automatic Control*, 4(2):1–9, 1959.

[3] Frederic Boy. Google trends™ dynamics for the guidance of open-source intelligence: Augmentation of social-media and survey surveillance of population mental health. In *2023 IEEE International Symposium on Technology and Society (ISTAS)*, pages 1–9. IEEE, 2023.

[4] Abel Brodeur, Andrew E. Clark, Sarah Fleche, and Nattavudh Powdthavee. COVID-19, lockdowns and well-being: Evidence from Google Trends. *Journal of Public Economics*, 193:104346, 2021.

[5] Herman Anthony Carneiro and Eleftherios Mylonakis. Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks. *Clinical Infectious Diseases*, 49(10):1557–1564, 2009.

[6] Hyunyoung Choi and Hal Varian. Predicting the Present with Google Trends. *Economic Record*, 88(s1):2–9, 2012.

[7] Ed Diener, Derrick Wirtz, William Tov, Chu Kim-Prieto, Dong-won Choi, Shigehiro Oishi, and Robert Biswas-Diener. New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social indicators research*, 97:143–156, 2010.

[8] Tolga Ergen and Suleyman Serdar Kozat. Unsupervised anomaly detection with lstm neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 31(8):3127–3141, 08 2020. Conference Name: IEEE Transactions on Neural Networks and Learning Systems.

[9] Nicolas Golenvaux, Pablo Gonzalez Alvarez, Harold Silvère Kiossou, and Pierre Schaus. An lstm approach to forecast migration using google trends. *arXiv preprint arXiv:2005.09902*, 2020.

[10] Scott Houghton, Frederic Boy, Alexander Bradley, Richard James, Heather Wardle, and Simon Dymond. Tracking online searches for gambling activities and operators in the united kingdom during the covid-19 pandemic: A google trends™ analysis. *Journal of Behavioral Addictions*, 12(4):983 – 991, 2023.

[11] Duleeka Knipe, David Gunnell, Hannah Evans, Ann John, and Daisy Fancourt. Is google trends a useful tool for tracking mental and social distress during a public health emergency? a time–series analysis. *Journal of Affective Disorders*, 294(July):737–744, 2021.

[12] Kenjiro Kondo, Akihiko Ishikawa, and Masashi Kimura. Sequence to sequence with attention for influenza prevalence prediction using google trends. In *Proceedings of the 2019 3rd International Conference on Computational Biology and Bioinformatics*, pages 1–7, 2019.

[13] Simon Kuznets. National income, 1929-1932, 1934.

[14] Bryan Lim and Stefan Zohren. Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194):20200209, 04 2021.

[15] Liyuan Liu, Meng Han, Yiyun Zhou, and Yan Wang. *LSTM Recurrent Neural Networks for Influenza Trends Prediction*, volume 10847, pages 259–264. Springer International Publishing, Cham, 2018.

[16] Igor Matias and Katarzyna Wac. Dejà vu: Recurrent neural networks for health wearables data forecast. In *2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 1724–1731. IEEE, 2022.

[17] Amaryllis Mavragani and Gabriela Ochoa. Google trends in infodemiology and infoveillance: Methodology framework. *JMIR Public Health and Surveillance*, 5(2):1–15, 2019.

[18] Adil Moghar and Mhamed Hamiche. Stock market prediction using lstm recurrent neural network. *Procedia computer science*, 170:1168–1173, 2020. Publisher: Elsevier.

[19] Zahra Montajabi, Vahid Khorasani Ghassab, and Nizar Bouguila. Recurrent neural network-based video compression. In *2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 925–930. IEEE, 2022.

[20] Alaa Nfissi, Wassim Bouachir, Nizar Bouguila, and Brian L Mishara. Cnn-n-gru: end-to-end speech emotion recognition from raw waveform signal using cnns and gated recurrent unit networks. In *2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 699–702. IEEE, 2022.

[21] Marcin Piekałkiewicz. Why do economists study happiness? *The Economic and Labour Relations Review*, 28(3):361–377, 2017.

[22] Sikakollu Prasanth, Uttam Singh, Arun Kumar, Vinay Anand Tikkiwal, and Peter HJ Chong. Forecasting spread of covid-19 using google trends: A hybrid gwo-deep learning approach. *Chaos, Solitons & Fractals*, 142:110336, 2021. Publisher: Elsevier.

[23] Aditya Ranganath, Omar DeGuchy, Fabian Santiago, Mukesh Singhal, and Roummel Marcia. Recurrent nerual imaging: An evolutionary approach for mixed possion-gaussian image denoising. In *2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 484–489. IEEE, 2022.

[24] Archi Roy, Soudeep Deb, and Divya Chakarwarti. Impact of covid-19 on public social life and mental health: A statistical study of google trends data from the usa. *Journal of Applied Statistics*, 2023.

[25] Caroline Saunders and Paul Dalziel. Twenty-five years of counting for nothing: Waring's critique of national accounts. *Feminist Economics*, 23(2):200–218, 2017.

[26] Rajshekhar Singhania and Sourav Kundu. Forecasting the united states unemployment rate by using recurrent neural networks with google trends data. *International Journal of Trade, Economics and Finance*, 11(6), 2020.

[27] Daniel Soutner and Luděk Müller. *Application of LSTM Neural Networks in Language Modelling*, volume 8082, pages 105–112. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[28] Edmund R Thompson. Development and validation of an internationally reliable short-form of the positive and negative affect schedule (panas). *Journal of cross-cultural psychology*, 38(2):227–242, 2007.

[29] David Watson and Lee Anna Clark. The panas-x: Manual for the positive and negative affect schedule-expanded form. *Unpublished manuscript, University of Iowa*, 1994.

[30] David Watson, Lee Anna Clark, and Auke Tellegen. Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54(6):1063, 1988.

[31] World Health Organisation. EARS - Early AI-supported Response with Social Listening, 2022.