



# Article The Deep Learning LSTM and MTD Models Best Predict Acute Respiratory Infection among Under-Five-Year Old Children in Somaliland

Mohamed Yusuf Hassan

Department of Statistics, College of Business, United Arab Emirates University, Al Ain P.O. Box 15551, United Arab Emirates; myusuf@uaeu.ac.ae; Tel.: +971-50-743-4612

**Abstract:** The most effective techniques for predicting time series patterns include machine learning and classical time series methods. The aim of this study is to search for the best artificial intelligence and classical forecasting techniques that can predict the spread of acute respiratory infection (ARI) and pneumonia among under-five-year old children in Somaliland. The techniques used in the study include seasonal autoregressive integrated moving averages (SARIMA), mixture transitions distribution (MTD), and long short term memory (LSTM) deep learning. The data used in the study were monthly observations collected from five regions in Somaliland from 2011–2014. Prediction results from the three best competing models are compared by using root mean square error (RMSE) and absolute mean deviation (MAD) accuracy measures. Results have shown that the deep learning LSTM and MTD models slightly outperformed the classical SARIMA model in predicting ARI values.

**Keywords:** artificial intelligence; training data; Pearson correlation; Dickey–Fuller test; long short-term memory; machine learning

# 1. Introduction

Acute respiratory infections (ARI) are the most common causes of both illness and mortality in children under five regardless of where they live or what their economic situation is [1,2]. ARIs, particularly lower respiratory tract infections (LRTI), are a major cause of death among children under five years of age. It is estimated that between 1.9 and 2.2 million children die annually because of this infection. Other estimates indicate that some 10.8 million children die each year [3]. Studies conducted by [4] suggested that more than 95% of all cases attributed to clinical pneumonia in young children worldwide occur in developing countries. Ref. [5] estimated that 4 million child deaths occur due to pneumonia each year, including 2.6 million infants and 1.4 million children aged 5-14 years. Despite the fact that forty-two percent of these ARI related deaths occur in Africa, the epidemiology and pathogenesis of LRTI remains understudied [6]. Accurate information on specific diseases in African children dying from respiratory illnesses is scarce [7]. Refs. [8–10] indicated the incidence of ARI is closely associated with the nutritional status of the child, socio-economic level of the family, and the family size. Studies conducted by [11,12] have shown that children who are under one year old are more likely to have ARI compared to older children. Some studies have associated environmental and sociodemographic factors with ARI. These include age of the mother, family size, and unhealthy environments [13–15]. Similar results from studies conducted in India showed a higher prevalence in urban areas [16]. Other studies demonstrated that children born to mothers younger than 20 years old had higher incidence of ARI infection compared to those with mothers older than 20 years [17]. Results from other studies have also shown increased incidence of ARI among children from homes that use unclean cooking fuel such as charcoal and firewood [18]. Ref. [19] suggests that improving nutrition and parental literacy may contribute to lowering the incidence of acute lower respiratory infections. Ref. [20] noted that other factors contributing to the incidence



Citation: Hassan, M.Y. The Deep Learning LSTM and MTD Models Best Predict Acute Respiratory Infection among Under-Five-Year Old Children in Somaliland. *Symmetry* **2021**, *13*, 1156. https://doi.org/10.3390/ sym13071156

Academic Editor: Juan Luis García Guirao

Received: 18 May 2021 Accepted: 21 June 2021 Published: 28 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). and severity of lower respiratory infections in developing countries include crowding in

households, high birth rate, vitamin A deficiency, and population. The World Health Organization (WHO) estimates under-five mortality in Somalia at 200 deaths per 1000 births, which is one of the highest in the world. Approximately one third of these are neonatal deaths, occurring during the first month of life. Pneumonia and diarrhea are the main killers, each contributing 20–25 percent of all under-five mortality [21]. Since different regions of the world may not experience outbreaks of communicable diseases at the same time, it is difficult to make worldwide joint predictions for the seasonal outbreaks of these diseases. Despite this, many studies have been conducted to predict the spread of these diseases using classical statistics and deep learning techniques. Deep learning models do not only play a prominent role in disease prediction, but these models are also used in analyzing and segmenting images [22], detection of tumors and lesions in medical images [23,24], and computer-aided diagnostics [25,26].

For disease prediction, some of the previous studies about the prediction of ARI that have used classical time series and deep learning techniques were conducted by [27–33].

## 2. Materials and Methods

# 2.1. Data

The data used in this study were collected from five regions in Somaliland from 2011–2014. The numbers of reported cases of acute respiratory infection and pneumonia during this period were 246,349 and 61,599, respectively. Ethical approval was provided by the institutional ethics review board of Somaliland's Ministry of Health (degree # 2020-7-027). The requirement for informed consent was waived because the data used in the study contained only the historical time series values for the number of admitted patients to the hospitals. Statistical and deep learning techniques are utilized to understand the trends of these diseases. SARIMA, MTD, and LSTM techniques are employed to assess and predict the spread of these diseases.

#### 2.2. Long Short Term Memory (LSTM) Model

Artificial neural networks are designed to perform tasks in a similar fashion to human neurons. Many variates of these algorithms are currently used in machine learning. Recurrent neural networks (RNN) are special cases of the general neural networks used to model patterns in sequential data. These networks are capable of capturing the dependency of sequential data like the data collected from sensors and stock markets over time. One of the limitations for these models is a vanishing or exploding gradient that has a negative impact on their convergence [34,35].

Ref. [35] introduced long short-term memory (LSTM), which is a type of recurrent neural network that overcomes the problem of vanishing/exploding gradients. These models can be trained in both short and long-term dependences. For some of the most recent studies about LSTM and ARIMA, see [36–40]. LSTM is made up of a memory cell, an input gate, an output gate, and a forget gate. The memory cell has the information of the past filtration, while the gates control the amount of memory needed to pass to the next stage given the new information and the past filtration. Unlike a feedforward neural network, LSTM has a looping mechanism that feeds information cyclically in a loop. This allows the use of both current information input and what has been learned previously. Figure 1 displays LSTM architecture.

$$f_t = \sigma \Big( W_f[h_{t-1}, x_t] + b_f \Big)$$
$$i_t = \sigma (W_i[h_{t-1}, x_t] + b_i)$$
$$\widetilde{C}_t = Tanh(W_c[h_{t-1}, x_t] + b_c)$$
$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t$$
$$o_t = \sigma (W_o[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * Tanh(C_t)$$

where  $C_{t-1}$  is the memory cell for past filtration, and  $f_t$ ,  $i_t$ , and  $o_t$  are the forget gate, input gate, and the output gate, respectively.  $\sigma$  is a logistic sigmoid function and Tanh is a hyperbolic activation function, whereas W and b are the matrix of the parameters and the bias, respectively.



Figure 1. LSTM [35].

#### 2.3. Mixture Transition Distribution (MTD) Model

Another promising technique for the analysis and the prediction of time series data is the MTD models introduced in 1985 by [41] for the modeling of high-order Markov chains with a finite state space. Since then, it has been successfully employed in a wide range of research problems. Ref. [42] proposed a univariate MTD with Gaussian components. For bivariate cases, Ref. [43] investigated a class of bivariate mixture transition distribution (BMTD) models using mixtures with components of different probability distributions. For non-Gaussian distributions, Ref. [44] considered MTD for high Markov chains and non-Gaussian time series. For more recent references about MTD, see, for example, [45–47].

In this study, a univariate version of the MTD model that has Gaussian components is used. The model has the following form:

$$\phi\left(y_t \middle| y^{t-1}, \varphi\right) = \sum_{j=1}^p \alpha_j \phi_j\left(y_t \middle| y^{t-1}, \varphi_j\right),$$

where  $\phi_j$  is a probability density function for the jth component of the mixture for j = 1, 2, 3...,  $p, y^{t-1}$  is the past filtration,  $\alpha_j$  is the weight for the *j*th component, and  $\varphi$  is the set of the model parameters.

## 3. Results

After the data were collected, they were randomly split into training and test sets with 70% for training and 30% for testing. Training data are fitted to the MTD, SARIMA, and LSTM models, and they are then validated with the test data to compare the performance of the models. RMSE and MAD accuracy measures are used to choose the best model for predicting the spread of the disease.

ARI accounted for 80% of the 307,948 combined cases for both diseases. To glean insights from the data, a qualitative assessment of the data is presented; Figures 2 and 3 display the time series trends of ARI and pneumonia for those four years. These trends show that both ARI and pneumonia are increasing over time, but the increase in ARI is relatively higher.



Figure 2. Pneumonia Trends.



Figure 3. ARI Trends.

It is observed from Figure 4, that the shape of the ARI distribution appears to be bimodal with a peak in November through January and another from March to May. This demonstrates that the disease is common through the seasons from winter to spring in Somaliland. The bar chart in Figure 5 also explicitly shows that increase.

Quantitative summary statistics of the data are depicted in Tables 1 and 2. From Table 1, for the last three years, the infection rate of ARI has increased by 18%, 19%, and 8%, respectively, whereas the annual increase of pneumonia incidences were 31%, 3%, and 10% during the same period. In addition, Table 2 displays 95% confidence intervals for the true monthly means of ARI and pneumonia incidences. Additionally, it has been observed that pneumonia is highly correlated with ARI. The Pearson correlation coefficient between ARI and pneumonia is 0.866, with a *p*-value of (<0.001). Since ARI and pneumonia are highly correlated and the majority of the cases are acute respiratory infections, we restrict our study to the investigation of the ARI trends.

Table 1. Yearly Percentage Increase.

Disease	11/12	12/13	13/14
ARI	18%	19%	8%
Pneumonia	31%	3%	0%

Disease	95% CI
ARI	(4757–5507)
Pneumonia	(1205–1362)



Figure 4. Average Monthly Cases.



Figure 5. Average Yearly Cases.

The original ARI data were not stationary, and its sample autocorrelation and partial autocorrelation functions are displayed in Figure 6a,b. The autocorrelation function dies down very slowly and the partial autocorrelation cuts off after the first lag, which is clearly an indication that the series is non-stationary. These results reveal that data transformation

is needed to make the time series stationary. The autocorrelation and the partial autocorrelation functions of the series after the first differencing are presented in Figure 6c,d respectively. Both functions die down very quickly after the first differencing.



Figure 6. Autocorrelation and Partial Autocorrelation of ARI.

The Dickey–Fuller test statistic for stationarity was computed and a value of -3.9 was obtained with a *p*-value of 0.022. The above results indicate the stationarity of the time series after the first differencing. Machine learning algorithms and Box–Jenkins forecasting methods are employed to predict the spread of the disease. The three competing models were chosen from the deep learning long short term memory (LSTM), EM-algorithm based mixture transition distribution (MTD), and the seasonal autoregressive integrated moving averages (SARIMA) techniques. The data were fitted to each one of these models and their results were compared to understand the patterns of the disease. 70% of data are used for training and the remaining 30% for testing. Root mean square error (RMSE) and mean absolute deviation (MAD) accuracy measures are used to choose the best model for predicting the future spread of the disease.

The best LSTM model identified by the two accuracy measures (RMSE and MAD) is the model with a learning rate of 5%, hidden-dim = 3, activation function = Tanh, loss function = MSE, optimizer = Adam, and the number of epochs = 300. The model is trained very well with the data and the training error rate has fallen very sharply, as shown in Figure 7. The best MTD model is the second order model with the parameter estimates ( $\alpha_1, \alpha_2$ ) = (0.1, 0.43), ( $\mu_1, \mu_2$ ) = (0.93, 2.4), and ( $\sigma_1, \sigma_2$ ) = (14.8, 1.39). Finally, the best SARIMA model is SARIMA (1, 1, 0)(1, 0, 0)<sub>12</sub>.



Figure 7. Training Error for 300 epochs.

Table 3 shows the estimated parameters of the SARIMA model with the test statistics and *p*-values. Both *p*-values are less than 0.02. Ljung–Box Chi-Square statistics of the SARIMA are also calculated for lags 12 and 24 and presented in Table 4 to check the adequacy of the fitted ARIMA model. RMSE and MAD indicated that the results of the three fitted models are very close with slight differences as shown in Table 5. However, MTD and LSTM performed better than the well-known and the more commonly used SARIMA model.

Table 3. Parameter Estimates.

Туре	Coefficient	<b>T-Statistic</b>	<i>p</i> -Value
ARI <sub>1</sub>	-0.415	-2.54	0.016
SAR <sub>12</sub>	0.961	7.74	0.000

**Table 4.** Ljung-Box Chi-Square statistics.

Туре	12	24
<b>Chi-Square</b>	16.28	20.69
DF	10	22
<i>p</i> -Value	0.092	0.540

Table 5. Accuracy Measures.

MODEL	RMSE	MAD
LSTM	0.307	0.302
MTD	0.503	0.433
SARIMA	0.60	0.476

# 4. Discussion

In this study, time series trends of acute respiratory infection and pneumonia among Somaliland children are investigated. Both the qualitative and the quantitative assessments of the data sets are performed. LSTM, MTD, and SARIMA models are employed to predict the spread of the ARI. Two accuracy measures, RMSE and MAD, were used to identify the best model that could be applied to predict acute respiratory infection. Results obtained in the study demonstrate that implementing several different machine learning and classical time series predictive modeling techniques could aid in the search for the best model that can predict the spread of diseases with high accuracy. Stationarity and data normalization improved the predictive accuracy. Furthermore, those predictive modeling techniques are varied in their performance depending on the given data, and on the stability of their convergence. Thus, finding a method that could generate predictive models that outperform all other models for any given time series data is unlikely, and instead, the implementation of multiple techniques is the standard practice in model development.

There is a large body of literature on predicting the spread of acute respiratory infection among different regions of the world. They generally use one of the predictive modeling techniques, but we are not aware of any modeling methods that compare the performance of the classical techniques, including but not limited to, Box–Jenkins, exponential smoothing, and decomposition with MTD and artificial intelligence. The comparisons implemented here are, as far as we know, the first attempt to compare deep learning sequential techniques with MTD and classical time series methods.

There are a number of limitations to this study. First, the time series data sample size was relatively small. Although the number of observations is 246,349, the data collection period was 48 months. Second, each of the three employed techniques has its own drawbacks. For example, parameter estimates of the MTD are based on the EM algorithm that has reproducibility problems if not carefully chosen. Major disadvantages of the LSTM algorithm include longer training times and more memory to train. For the Box–Jenkins models, the main weakness is the satisfaction of the residual assumptions and model stability. Finally, to overcome the limitations of those techniques, domain knowledge could be useful to select the initial values of the parameters.

In conclusion, the results of this study showed that LSTM and MTD both performed better than the SARIMA model. LSTM and MTD models demonstrated their flexibility and competitiveness in the study, which may lead to their being considered viable alternatives to the existing time series models.

### 5. Conclusions

In this study, we have compared the prediction performance of MTD, seasonal autoregressive integrated moving averages (SARIMA), and the long short term memory (LSTM) deep learning models. The data used in the study were monthly ARI and Pneumonia cases among under-five year old children that were collected from five regions of Somaliland from 2011–2014. To get reliable results from the above models, a number of changes has been made to the original observations of the data before comparing competing models. First, the data are transformed by differencing to make them stationary; the autocorrelation function (ACF), partial autocorrelation (PACF), and the Dickey-Fuller test for stationarity were computed to check the stationarity of the series. The results of these measures indicated that the time series is stationary. Second, data were randomly split into training and testing sets with a ratio of 70:30 (70% training and 30% testing). Third, the independent variables in the training data were normalized by subtracting their means and dividing them by their standard deviations. Data normalization in the preprocessing stage is needed to implement machine learning algorithms to reduce the variability among the different variables. Training data are fitted to the three models, and then validated with the test data to compare the performance of the competing models. RMSE and MAD accuracy measures are used to choose the best model in predicting the spread of the disease.

The results of this study have shown that no model is a panacea over the other two models, but they demonstrated that the deep learning LSTM and MTD models slightly outperformed the classical SARIMA model in predicting ARI values.

Although there is a large body of literature dealing with the comparison of Box– Jenkins and machine learning techniques, we are not aware of any studies that compare the MTD, the SARIMA, and the LSTM models. Perhaps one of the most important outcomes of this study is the performance of the MTD model. The study illustrated the utility and the efficacy of the MTD model, which is not familiar to many researchers. Besides, this study is the first to attempt to show that MTD could be a highly competitive and flexible predictive model that can challenge other machine learning algorithms.

Funding: We received no financial support for this paper.

Institutional Review Board Statement: Not applicable.

**Informed Consent Statement:** Ethical approval was provided by the institutional ethics review board of Somaliland's Ministry of Health (degree # 2020-7-027).

Data Availability Statement: The study did not report any data.

**Acknowledgments:** The authors would like to thank the Somaliland Ministry of Health for the data collection.

Conflicts of Interest: The author declares no conflict of interest.

## References

- Tupasi, T.E.; Velmonte, M.A.; Sanvictores, M.E.G.; Abraham, L.; De Leon, L.E.; Tan, S.A.; Miguel, C.A.; Saniel, M.C. Determinants of Morbidity and Mortality Due to Acute Respiratory Infections: Implications for Intervention. J. Infect. Dis. 1988, 157, 615–623. [CrossRef]
- Kamath, K.R.; Feldman, R.A.; Sundar, P.R.; Webb, J.K.G. Infection and Disease in a Group of South Indian Families. *Am. J. Epidemiol.* 1969, 89, 375–383. [CrossRef]
- 3. Bryce, J.; Boschi-Pinto, C.; Shibuya, K.; Black, R.E. WHO estimates of the causes of death in children? *Lancet* 2005, 365, 1147–1152. [CrossRef]
- 4. Black, R.E.; Morris, S.; Bryce, J. Where and why are 10 million children dying every year? Lancet 2003, 361, 2226–2234. [CrossRef]
- Rudan, I.; Tomaskovic, L.; Boschi-Pinto, C.; Campbell, H. Global estimate of the incidence of clinical pneumonia among children under five years of age. *Bull. World Health Organ.* 2005, *82*, 895–903.
- Leowski, J. Mortality from acute respiratory infections in children under 5 years of age: Global estimates. World Health Stat. Q. 1986, 39, 138–144.
- 7. Williams, B.G.; Gouws, E.; Boschi-Pinto, C.; Bryce, J.; Dye, C. Estimates of world-wide distribution of child deaths from acute respiratory infections. *Lancet Infect. Dis.* 2002, *2*, 25–32. [CrossRef]
- Chintu, C.; Mudenda, V.; Lucas, S.; Nunn, A.; Lishimpi, K.; Maswahu, D.; Kasolo, F.; Mwaba, P.; Bhat, G.; Terunuma, H.; et al. Lung diseases at necropsy in African children dying from respiratory illnesses: A descriptive necropsy study. *Lancet* 2002, 360, 985–990. [CrossRef]
- 9. Singh, M.P.; Nayar, S. Magnitude of acute respiratory infections in under five children. J. Commun. Dis. 1996, 28, 273–278.
- 10. Oyedeji, G.A. The effect of socio-economic factors on the incidence and severity of gastroenteritis in Nigerian children. *Niger. Med. J.* **1987**, *4*, 229–232.
- 11. Harerimana, J.-M.; Nyirazinyoye, L.; Thomson, D.R.; Ntaganira, J. Social, economic and environmental risk factors for acute lower respiratory infections among children under five years of age in Rwanda. *Arch. Public Health* **2016**, *74*, 19. [CrossRef] [PubMed]
- 12. Geberetsadik, A.; Worku, A.; Berhane, Y. Factors associated with acute respiratory infection in children under the age of 5 years: Evidence from the 2011 Ethiopia demographic and health survey. *Pediatric Health Med. J.* **2015**, *6*, 9.
- 13. Abuka, T. Prevalence of pneumonia and factors associated among children 2–59 months in Wondo genet district sidama zone SNNPR Ethiopia. *Curr. Pediatric* 2017, 21, 19–25.
- Gothankar, J.; Doke, P.; Dhumale, G.; Pore, P.; Lalwani, S.; Quraishi, S.; Murarkar, S.; Patil, R.; Waghachavare, V.; Dhobale, R.; et al. Reported incidence and risk factors of childhood pneumonia in India: A community-based cross-sectional study. BMC Public Health 2018, 18, 1111. [CrossRef]
- Jroundi, I.; Mahraoui, C.; Benmessaoud, R.; Moraleda, C.; Tligui, H.; Seffar, M.; El Kettani, S.E.-C.; Benjelloun, B.S.; Chaacho, S.; Muñoz-Almagro, C.; et al. Risk factors for a poor outcome among children admitted with clinically severe pneumonia to a university hospital in Rabat, Morocco. *Int. J. Infect. Dis.* 2014, 28, 164–170. [CrossRef]
- 16. Kumar, S.G.; Majumdar, A.; Kumar, V.; Naik, B.N.; Selvaraj, K.; Balajee, K. Prevalence of acute respiratory infection among under-five children in urban and rural areas of Puducherry, India. *J. Nat. Sci. Biol. Med.* **2015**, *6*, 3–6. [CrossRef]
- 17. Ujunwa, F.; Ezeonu, C. Risk Factors for Acute Respiratory Tract Infections in Under-five Children in Enugu Southeast Nigeria. *Ann. Med. Health Sci. Res.* **2014**, *4*, 95–99. [CrossRef]
- 18. Fekadu, G.A.; Terefe, W.M.; Alemie, G.A. Prevalence of pneumonia among under five children in Este town and the surrounding rural Kebeles, Northwest Ethiopia: A community based cross sectional study. *Sci. J. Public Health* **2014**, *2*, 150–155. [CrossRef]
- 19. Pawliska-Chmara, R.; Wronka, I. Assessment of the effect of socioeconomic factors on the prevalence of respiratory disorders in children. J. Physiol. Pharmacol. 2007, 58, 523–529.
- 20. Berman, S. Epidemiology of Acute Respiratory Infections in Children of Developing Countries. *Clin. Infect. Dis.* **1991**, *13*, S454–S462. [CrossRef]
- World Health Organization (WHO). 2012. Available online: http://www.emro.who.int/images/stories/somalia/documents/ layoutchildhealth-9mar.pdf?ua=1 (accessed on 15 November 2020).
- 22. Liu, N.; Wan, L.; Zhang, Y.; Zhou, T.; Huo, H.; Fang, T. Exploiting Convolutional Neural Networks with Deeply Local Description for Remote Sensing Image Classification. *IEEE Access* **2018**, *6*, 11215–11228. [CrossRef]
- 23. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* 2017, 42, 60–88. [CrossRef]
- 24. Brunetti, A.; Carnimeo, L.; Trotta, G.F.; Bevilacqua, V. Computer-assisted frameworks for classification of liver, breast and blood neoplasias via neural networks: A survey based on medical images. *Neurocomputing* **2019**, *335*, 274–298. [CrossRef]

- Asiri, N.; Hussain, M.; Al Adel, F.; Alzaidi, N. Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey. *Artif. Intell. Med.* 2019, 99, 101701. [CrossRef] [PubMed]
- 26. Zhou, T.; Thung, K.-H.; Zhu, X.; Shen, D. Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. *Hum. Brain Mapp.* **2019**, *40*, 1001–1016. [CrossRef]
- 27. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* **2018**, *172*, 1122–1131. [CrossRef]
- Rahaman, M.; Li, C.; Yao, Y.; Kulwa, F.; Rahman, M.A.; Wang, Q.; Qi, S.; Kong, F.; Zhu, X.; Zhao, X. Identification of COVID-19 samples from chest X-Ray images using deep learning: A comparison of transfer learning approaches. *J. X-ray Sci. Technol.* 2020, 28, 821–839. [CrossRef]
- 29. Lin, Y.; Chen, M.; Chen, G.; Wu, X.; Lin, T. Application of an autoregressive integrated moving average model for predicting injury mortality in Xiamen, China. *BMJ Open* **2015**, *5*, e008491. [CrossRef]
- 30. Bahadori, M.T.; Lipton, Z.C. Temporal-Clustering Invariance in Irregular Healthcare Time Series. arXiv 2019, arXiv:1904.12206.
- 31. Hassan, M. Trends of Tuberculosis in Somaliland's Young Children after the Conflict and the Role Khat Marfishes Play Its Transmission. *J. Biom. Biostat.* **2018**, *9*, 1–6. [CrossRef]
- Olsavszky, V.; Dosius, M.; Vladescu, C.; Benecke, J. Time Series Analysis and Forecasting with Automated Machine Learning on a National ICD-10 Database. Int. J. Environ. Res. Public Health 2020, 17, 4979. [CrossRef]
- Perna, D.; Tagarelli, A. Deep Auscultation: Predicting Respiratory Anomalies and Diseases via Recurrent Neural Networks. *arXiv* 2019, arXiv:1907.05708.
- 34. El Hihi, S.; Bengio, Y. Hierarchical recurrent neural networks for long term dependencies. *Adv. Neural Inf. Process. Syst.* **1995**, *8*, 493–499.
- 35. Hochreiter, S.; Schmidhuber, J. LSTM Can Solve Hard Long-Time Lag Problems; MIT Press: Cambridge, MA, USA, 1997.
- Chen, K.; Zhou, Y.; Dai, F. A LSTM-based method for stock returns prediction: A case study of China stock market. In Proceedings of the 2015 IEEE International Conference on Big Data (Big Data), Santa Clara, CA, USA, 29 October–1 November 2015; pp. 2823–2824.
- Ji, L.; Zou, Y.; He, K.; Zhu, B. Carbon futures price forecasting based with ARIMA-CNN-LSTM model. *Procedia Comput. Sci.* 2019, 162, 33–38. [CrossRef]
- Temür, A.; Akgün, M.; Temür, G. Predicting Housing Sales in Turkey Using Arima, LSTM And Hybrid Models. J. Bus. Econ. Manag. 2019, 20, 920–938. [CrossRef]
- Kijewski, M.; Slepaczuk, R. Predicting Prices of S&P500 Index Using Classical Methods and Recurrent Neural Networks; University of Warsaw: Warsaw, Poland, 2020.
- 40. Roondiwala, M.; Patel, H.; Varma, S. Predicting stock prices using LSTM. IJSR 2017, 6, 1754–1756.
- 41. Raftery, A.E. A Model for High-Order Markov Chains. J. R. Stat. Soc. Ser. B 1985, 47, 528–539. [CrossRef]
- 42. Le, N.D.; Martin, R.D.; Raftery, A.E. Modeling Flat Stretches, Bursts, and Outliers in Time Series Using Mixture Transition Distribution Models. J. Am. Stat. Assoc. 1996, 91, 1504.
- 43. Hassan, M.Y.; Lii, K.-S. Modeling Marked Point Processes via Bivariate Mixture Transition Distribution Models. J. Am. Stat. Assoc. 2006, 101, 1241–1252. [CrossRef]
- 44. Berchtold, A.; Raftery, A. The Mixture Transition Distribution Model for High-Order Markov Chains and Non-Gaussian Time Series. *Stat. Sci.* **2002**, *17*, 328–356. [CrossRef]
- 45. Bolano, D.; Berchtold, A. General framework and model building in the class of Hidden Mixture Transition Distribution models. *Comput. Stat. Data Anal.* **2016**, *93*, 131–145. [CrossRef]
- 46. Zheng, X.; Kottas, A.; Sanso, B. On Construction and Estimation of Stationary Mixture Transition Distribution Models. *arXiv* 2020, arXiv:2010.12696.
- 47. Heiner, M.; Kottas, A. Estimation and selection for high-order Markov chains with Bayesian mixture transition distribution models. *arXiv* 2019, arXiv:1906.10781.