



M E T H O D L O G Y

Predicting Diarrhoea Outbreaks with Climate Change

Tassallah Abdullahil, Geoff Nitschke, Neville Sweijd

April 19, 2022

MACHINE LEARNING

a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so.

Deep Learning DL Techniques;

Convolutional Neural Networks (CNNs)

Long-Short term Memory Networks (LSTMs);

Support Vector Machine (SVM).

Generative Adversarial Networks (GANs)

Relevance Estimation and Value Calibration (REVAC)

Sensitivity Analysis

Four triangles are positioned around the central text: an orange triangle pointing left at the top left, an orange triangle pointing right at the top right, a white triangle pointing left at the bottom left, and a white triangle pointing right at the bottom right.

DEEP LEARNING TECHNIQUES

Uses artificial neural networks to perform sophisticated computations on large amounts of data.

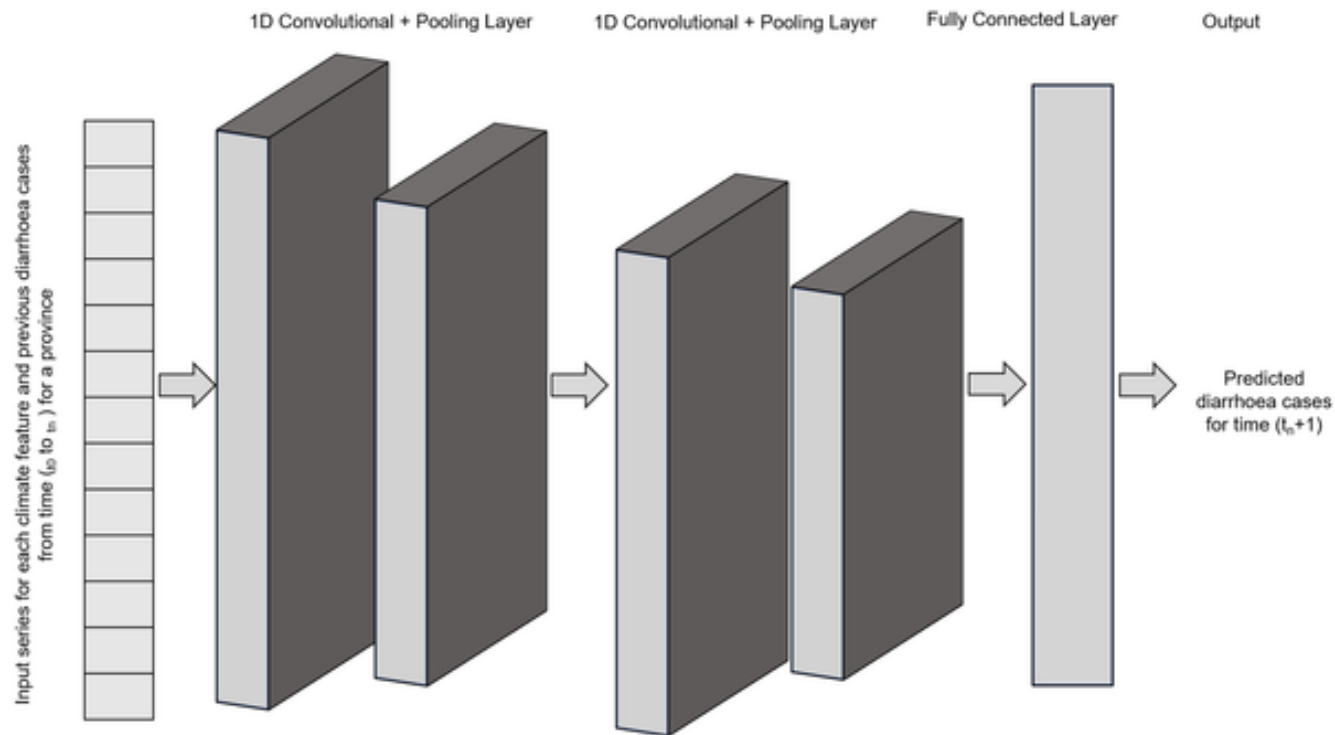
It works based on the structure and function of the human brain.



Convolutional Neural Networks (CNNs)

CNNs are a class of feed forward, deep neural network that consist of multiple convolutional and activation layers, pooling layers, and a fully connected layer as shas shown in Fig 2.

Fig 2. Basic architecture of the Convolutional Neural Network (CNN) with two convolution and pooling layers.



Abdullahi T, Nitschke G, Sweijd N (2022) Predicting diarrhoea outbreaks with climate change. PLOS ONE 17(4): e0262008.

<https://doi.org/10.1371/journal.pone.0262008>

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0262008>

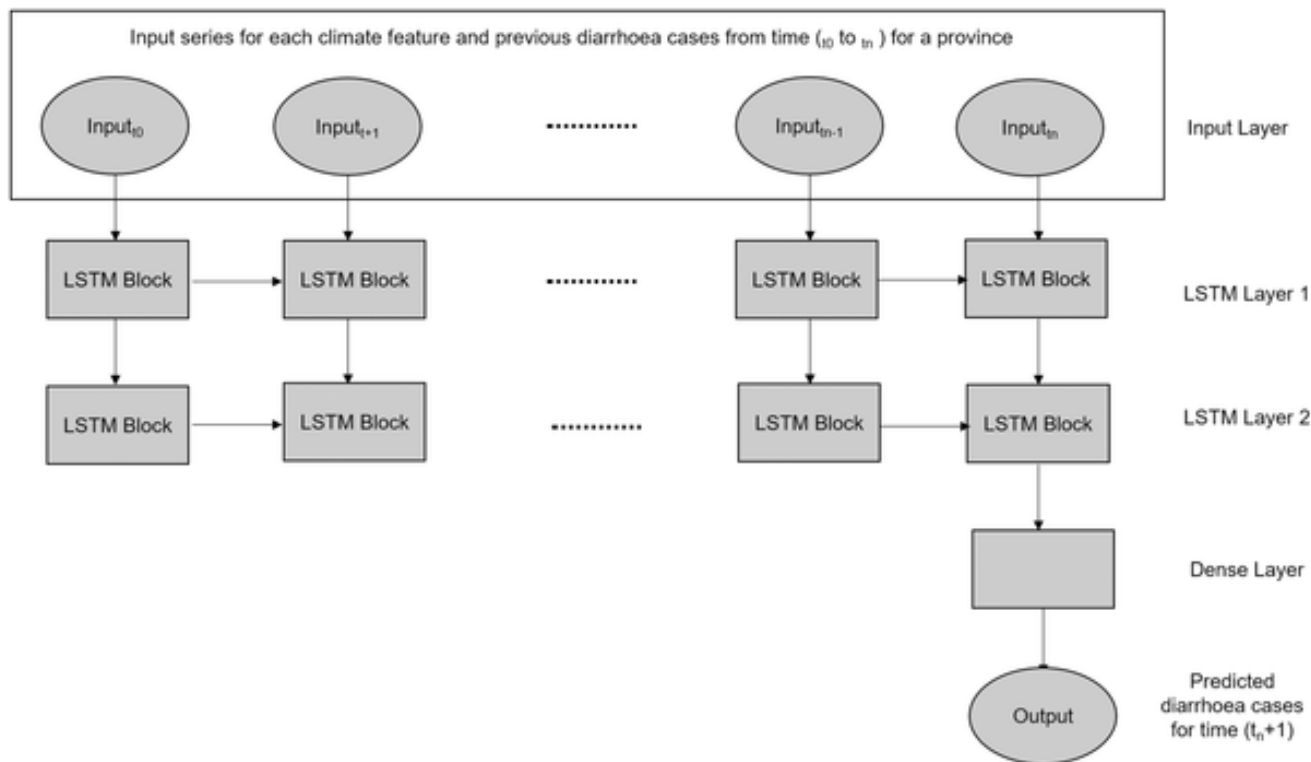
PLOS ONE



Long-Short term Memory Networks (LSTMs)

LSTMs as shown in Fig 3 are examples of Neural Networks under the category of Recurrent Neural Networks (RNNs) that address the issue of exploding and vanishing gradients.

Fig 3. Basic structure of the Long-term Short Term (LSTM) method with two LSTM layers.



Abdullahi T, Nitschke G, Sweijd N (2022) Predicting diarrhoea outbreaks with climate change. PLOS ONE 17(4): e0262008.

<https://doi.org/10.1371/journal.pone.0262008>

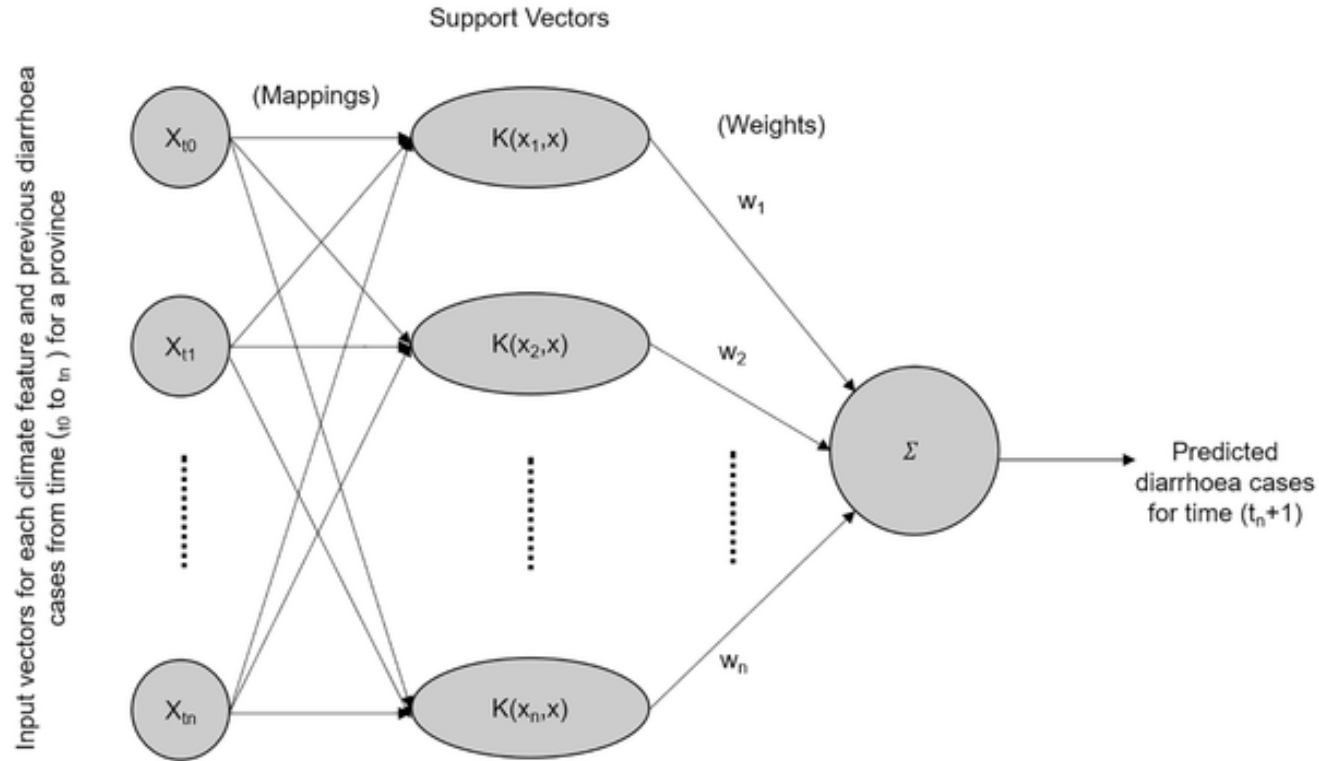
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0262008>



Support Vectore Machine (SVM)

SVMs are mathematical models whose main function is to find hyperplanes capable of creating margins that separates data points in a high dimensional feature space with the smallest structural risk using kernel functions as shown in Fig 4.

Fig 4. Structure of the Support Vector Machine (SVM) regression method.



Abdullahi T, Nitschke G, Sweijd N (2022) Predicting diarrhoea outbreaks with climate change. PLOS ONE 17(4): e0262008.

<https://doi.org/10.1371/journal.pone.0262008>

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0262008>



Generative Adversarial Networks (GANs)

GAN is a deep neural network framework which is able to learn from a set of training data and generate new data with the same characteristics as the training data.



Relevance Estimation and Value Calibration (REVAC)

REVAC is a method for calibrating the parameters of an evolutionary algorithm (EA) which is a laborious task. The highly stochastic nature of an EA typically leads to a high variance of the measurements.



Sensitivity Analysis

Sensitivity analysis is a procedure that can help discover how the effectiveness of your datasets varies with less or more data included. It is useful for model optimization and time and cost-effectiveness.



Root Mean Square Error (RMSE)

RMSE measures the average difference between a statistical model's predicted values and the actual values. To compare and evaluate the performance of ML methods, the Root Mean Square Error (RMSE) was used since it is widely adopted in many prediction studies.



Root Mean Square Error (RMSE)

RMSE is the square root of the mean of the squared differences between actual outcomes and the predictions made by a given method. It is calculated using the equation below:

$$RMSE = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^n (x_i - y_i)^2}$$

x_i is the actual value while

y_i is the predicted value and

n is the total number of observations to be analyzed.



T H A N K Y O U

JUDISMA SALI