Application of AI Techniques for Predicting Weather Conditions

Subjects: Computer Science, Artificial Intelligence Contributor: Gabriel Souto Fischer

Artificial Intelligence is the area of computing that studies intelligent entities and tries, through various techniques, to teach the computer to perform activities that previously only intelligent entities could perform. There are several approaches and ways of doing this, and one of the most used is Artificial Neural Networks. This entry presents a study on techniques for using Artificial Intelligence. Furthermore, it presents two works that use Recurrent Neural Networks and LSTM to predict weather conditions, providing an analysis of the solution proposed in each approach and a comparison between them.

ANN LSTM Neural Networks

1. Introduction

For Haugeland^[1], Artificial Intelligence (AI) studies are "*The exciting new effort to make computers think… Machines with minds, in the full and literal sense*". For Kurzweil^[2], it is "*The art of creating machines that perform functions that require intelligence when performed by people*". There are several definitions for the term, but in general the field of Artificial Intelligence (AI) is the area of computing that tries to understand intelligent entities. Thus, one of the motivations for studying AI is precisely to understand more about ourselves. However, the objective is less connected to philosophical and psychological issues but closer to the idea of building intelligent entities at the same rate as we understand them. Another reason for studying Artificial Intelligence is that these constructions can be interesting for solving problems instead of people. Currently we already have a wide variety of systems built from AI, however none with the same capacity as human intelligence. When computers reach a level of intelligence equivalent to or better than humans, it will have a major impact on everyday life and society future of civilization^[3].

2. Neural Networks

Artificial Neural Networks (ANN) are an attempt to model the information processing capacity of a nervous system^[4]. Neural Networks were originally proposed in 1943 by Warren McCulloch and Walter Pitts^[5], where they made an analogy between nerve cells and the electronic process of Formal Neurons. Since then, several models have been developed to apply the idea of Artificial Neural Networks originally conceived, trying to get closer and closer to the biological model^[6].

2.1. Artificial Neuron

The artificial neuron (**Figure 1**) is a logical structure that attempts to simulate the behavior of a biological neuron. In this way, the system's inputs are equivalent to dendrites that are connected to the artificial cell body by elements called weights, responsible for simulating synapses. The data captured by the inputs is processed by a sum function and sent to the outputs by the transfer function^[6].



Figure 1. Artificial Neuron^[7]

2.1. Artificial Neural Networks

An Artificial Neural Network is a set of neurons interconnected in several layers. The inputs of one can be connected to many other neurons, and its outputs can be connected to many others. These connections represent the contacts of the dendrites of the biological system, forming synapses. The function of the connection is basically to turn the output signal of an artificial neuron into the input signal of another or, if the same

is at the end of the network, direct the signal out of the network, as a processing response. As neurons can connect to each other in different ways in several layers, they can infinite different structures can be generated according to the needs of the network^[6].

In the ANN training phase, the weight values are determined and corrected, so that it responds as expected, minimizing the output error for the input vectors. Through the parallel information processing structure, it is possible for a certain degree of knowledge to be included in the process, so that the network is capable of detecting and classifying signals. Thus, an ANN is able to consider the knowledge acquired during training to respond to new input data more appropriately^[8].

2.3. Recurrent Neural Networks

Most training algorithms are not capable of implementing dynamic mappings, therefore, one of the ideas for temporal processing involving Neural Networks involves the use of time windows, where the network input uses portions of temporal data. However, this is not the most suitable solution for temporal processing. Thus, the big question is how to extend the structures of a network so that it is capable of to assume behavior that varies over time, being able to treat temporal signals. Therefore, for the ANN to be considered dynamic, it must have some kind of memory^[8].

One way to provide memory to an ANN is through Recurrent Neural Networks (**Figure 2**), which are those that have feedback connections to provide dynamic behavior. There are recurrent networks in which the input pattern is fixed and the output moves dynamically towards a stable state and those in which both input and output vary over time. There are numerous architectures for recurrent networks, with training algorithms that range from the simplest to the most complex and specific^[8]. According to Lipton^[9], one of the most successful Network architectures Recurrent is that of Long Short Term Memory (LSTM) Networks^[10], where the concept of Memory Cell is introduced, a computing unit that takes the place of artificial neurons in the hidden layer of the network.



Figure 2. Recurrent Neural Network 9

2.4 Long-short Term Memory Neural Networks

Recurrent Neural Networks are great for dealing with sequences, however when the objective is to learn dependencies during several stages of processing, they encounter problems. This occurs because after several iterations causing updates to the hidden layers, important information from the beginning of the sequence ends up being lost. There were several proposals to solve the problem, and LSTM cells were one of them^[11]. An LSTM

network has the same properties as a standard Recurrent Network, but it has the ability to store information for long periods of time when processing a sequence^[11].

To achieve this, the model reorganizes the hidden layer of the network, where each node is replaced by a Memory Cell^[9]. Memory cells are more complex than their equivalents in traditional Recurrent Networks. They are able to store information throughout the entire sequence processing, and are also able to select which information can be lost and deleted during the process. Being able to remember information is the great advantage of an LSTM Network, as it is able to store important information from the past that may only end up making sense in the future^[11].

3. Problem Analyzed

As previously stated, recurrent Neural Networks are not efficient in storing important information from the past, unlike LSTM Networks. There are several situations in which it is necessary to know data from the past to validate things that are happening in the present moment or even predict things that will happen in the future. One of the problems that are analyzed and, in a way, solved through the analysis of conditions that have already occurred to define things that will happen is weather forecasting.

To predict what the climate will be like in the future, it is necessary not only to collect and analyze information about the current state of atmospheric data to define how the climate will evolve, but to observe how similar data has evolved in the past^[12]. The big question that has been the subject of studies for several years is how to carry out this analysis efficiently, aiming at an accurate prediction of climate conditions.

4. Techniques Used

To carry out weather forecasting, mathematical models are generally used, however, with the advancement of the Artificial Intelligence area, new approaches using LSTM Networks are being proposed to replace or complement old models. In this work, we will analyze two works proposed for Predicting Weather Conditions using LSTM Neural Networks.

To carry out this review, the following research question was created: *How to use LSTM Neural Networks to make weather forecasts?*. This study question was used as the basis for an analytical search for articles. It was decided to use the *Google Scholar* indexer, as it is capable of indexing several magazines published in our country, bringing a broad base of articles in Portuguese. To carry out the research, the search string was used: (LSTM and "Weather Prediction"). Afterwards, the classification of articles was defined in the *Google Scholar* tool based on their order of Relevance. Finally, as an exclusion criterion, it was defined that only articles published from 2016 onwards would be analyzed. Our research scope returned 51 articles, and the two most relevant among them were chosen for this analysis.

4.1. Work 1: Sequence to Sequence Weather Forecasting with Long Short-Term Memory Recurrent Neural Networks

The system proposed by Zaytar^[12] is based on the development of an LSTM Neural Network capable of analyzing recent data, which is the most important for short-term weather forecasting, and, at the same time, analyzing old data that is capable of helping the model recognize atmospheric patterns and movements that recent data alone is unable to show.

The proposed model consists of a network of several layers, two of which are LSTM, and a fully connected hidden layer with 100 neurons, created for training based on multiple experiments. The model receives as input the values of the last 24 or 72 hours and returns as output the next 24 or 72 hours of forecast. The Values that the network receives as input and returns as output are Temperature, Humidity and Wind Speed for each hour.

To train the network, climate data from nine cities in Morocco was used over a period of fifteen years of hour-byhour data. After collecting the data, it was identified that there were gaps in the tables, however, none of the gaps reached days of missing data. Therefore, to complete the \textit{Dataset}, the blank values were filled in with the actual information on the weather conditions captured prior to the gaps. After this treatment, all values were normalized to avoid problems in the training stage. The network training time was 3.2h for the 24-hour experiment and 6.7h for the 72-hour experiment.

To analyze the results, the network was fed with one day of climate data from each of the nine cities used for training and the result of the network's prediction was compared with the climate data actually observed in the following 24 and 72 hours. For analysis, the authors used the mean squared error of the test data. For the 24-hour forecast, in the case with the lowest mean squared error it was 0.00516 and 0.00839 in the case with the highest error. As for the 72-hour forecast, in the city with the lowest average squared error it was 0.00675 and 0.01053 in the city with the highest error.

The results of the article show that LSTM networks can be used for general weather forecasting with good accuracy. According to the authors, the success of the model suggests that it could be used to analyze other climatic conditions that were not addressed in the study. For the authors, an efficient model shows that in the future weather forecasting can be carried out entirely by neural networks instead of traditional methods.

4.2. Work 2: Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting

In this work^[13] seeks to address the problem of predicting rainfall for the next few hours, *Precipitation Nowcasting*, through a variation of LSTM networks called ConvLSTM. This forecast involves the use of "radar echo" images to predict a fixed period in the future in the region covered by the images. The complexity of this problem is due to the fact that the network needs to predict characteristics that relate to both space and time, in this case, the progression of rain clouds throughout the input period and the forecast period.

The authors state that such problems are complex due to their high dimensionality, which favors strategies that can capture space-time dynamics. Although the work strictly addresses weather forecasting, the authors state that the problem can be generalized to a forecast of spatio-temporal sequences, including not only rain clouds but any moving object, and can be useful for object tracking in video, for example.

The constructed network determines the future state of a given cell in a grid NxN based on the current input as well as the past states of its local neighbors. To achieve this result it is necessary to use padding, that is, filling the "edges" of the grid with some initialization value which, in the case of ConvLSTM, is zero, denoting "total ignorance" regarding the future on the part of the model and at the same time helping to define the boundaries of the prediction. To test the proposed model, the authors built a new dataset based on radar echo images and compared their proposal against one of the algorithms accepted as state of the art, ROVER (Real-time Optical flow by Variational methods for Echoes of Radar) through various metrics used to evaluate rain nowcasting.

The results indicate that the proposal meets expectations and surpasses the algorithm chosen as state of the art, being more accurate in treating the borders of radar images, especially with regard to the contour of clouds in the border region. When compared to ROVER, the model presents less clear images but is more accurate and causes fewer false positives. Regarding forecast clarity, the authors state that due to the nature of the forecast task, great clarity cannot be expected when working with longer forecast time horizons. In conclusion, they state that the ConvLSTM model not only maintains the advantages of FC-LSTM but also makes it suitable for processing spatiotemporal data due to its convolutional structure.

5. Conclusion

From this work it was possible to briefly revisit basic concepts of Artificial Neural Networks, Recurrent Neural Networks and LSTM Neural Networks. Two approaches based on LSTM networks for climate prediction were also presented. Both strategies claim positive results in their fields of activity, respectively, the immediate forecast (1 to 6 hours) and the broader forecast, covering factors such as temperature, humidity and wind. According to the authors, approaches considered state of the art are becoming increasingly difficult to process with the accumulation of increasingly larger volumes of data. Thus, the works show how Artificial Intelligence, and more precisely Artificial Neural Networks, are gaining ground compared to other prediction methods both within the academic community and in the commercial environment.

References

- 1. HAUGELAND, J. Artificial Intelligence: The Very Idea. ; Massachusetts Institute of Technology: Cambridge, MA, USA, 1985; pp. 2.
- 2. KURZWEIL, R. The Age of Intelligent Machines; MIT Press: Cambridge, MA, USA, 1990; pp. 1.

- RUSSELL, S. J.; NORVIG, P. Artificial Intelligence: A Modern Approac; Prentice-Hall, Inc.: Upper Saddle River, NJ, USA, 1995; pp. 1.
- 4. ROJAS, R. . Neural Networks: A Systematic Introduction; Springer-Verlag New York, Inc.: New York, NY, USA, 1996; pp. 3.
- 5. Warren S. McCulloch; Walter Pitts; A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biol.*. **1943**, *5*, 115-133.
- TAFNER, M. A. Redes Neurais Artificiais: Aprendizado e Plasticidade.. *Revista Cérebro & Mente*. 1998, 2, 5.
- 7. Joseph Awoamim Yacim; Douw Gert Brand Boshoff; Impact of Artificial Neural Networks Training Algorithms on Accurate Prediction of Property Values. *J. Real Estate Res.*. **2018**, *40*, 375-418.
- 8. Ê. C. Segatto; D. V. Coury; Redes neurais artificiais recorrentes aplicadas na correção de sinais distorcidos pela saturação de transformadores de corrente. *null.* **2006**, *17*, 424-436.
- 9. Zachary C. Lipton, John Berkowitz, Charles Elkan. A Critical Review of Recurrent Neural Networks for Sequence Learning. *CoRR*. **2015**, *1*, 1.
- 10. Sepp Hochreiter; Jürgen Schmidhuber; Long Short-Term Memory. *Neural Comput.*. **1997**, 9, 1735-1780.
- DRUMOND, R. R. Peek: Classificação de Movimento Utilizando Dados Esparsosde Membros Superiores em Redes com LSTM. 2017. 44 p. Dissertação (Mestrado emComputação - Área de concentração: Computação Visual) — Instituto de Computação daUniverdade Federal Fluminense, Niterói, 2017.
- ZAYTAR, M. A.; AMRANI, C. E. Se. Sequence to sequence weather forecasting with long shortterm memory recurrent neural networks. *International Journal of Computer Applications*. 2016, 143, 7-11.
- SHI, X., Chen, Z., Wang, H., Yeung, D., Wong, W., WOO, W. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting; C. Cortes and N. D. Lawrence and D. D. Lee and M. Sugiyama and R. Garnett, Eds.; Curran Associates, Inc.: New York, USA, 2015; pp. 802-810.

Retrieved from https://encyclopedia.pub/entry/history/show/124933