Interpretability of Machine Learning Applications in Hydrology

Asude Konuk

Freie Universität Berlin

Abstract

The interpretability of machine learning in the natural sciences has received considerable attention in recent years. The main reasons for using machine learning in the natural sciences are to achieve scientific understanding, infer causal relationships from observational data, and gain new scientific insights. To increase interpretability, we need to discover possible techniques to apply and determine which ones are most effective in this context. This summary will be achieved through machine learning applications in hydrology and primarily LSTM (Long-Short Term Memory) and results from applications with different machine learning methods.

Literature Summary

Interpretability of Machine Learning in Natural Science

Interpretability of machine learning in natural science refers to the ability to understand and explain the decisions made by a machine learning model in the context of natural scientific research. Natural science research often involves the analysis of complex data sets, which can be challenging for humans to understand and interpret on their own. By using explainable machine learning, scientists can gain insights into the underlying mechanisms of the data and gain a better understanding of the relationships between different variables. This can help to generate new hypotheses and drive scientific discovery. Overall, the use of machine learning in natural science has the potential to greatly advance our understanding of the world and help us solve complex problems in fields like earth science, environmental science, and more. However, ensuring the interpretability of these models is crucial to ensuring their accuracy and usefulness [1].

The Availability and Applicability of LSTM

Long Short Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) that is able to learn long-term dependencies in sequential data. LSTM-RNNs are well-suited for time series prediction and are able to handle non-linearity and large amounts of data [2]. The main concept behind LSTM is the use of memory cells that enable constant error flow during training. This allows LSTM to avoid the vanishing gradient problem, where training errors decrease as they are propagated back through time [3]. In the field of natural sciences, LSTM networks have been applied to a range of tasks, such as predicting the properties of chemical compounds and analyzing time-series data in ecology and meteorology. One notable example is the use of LSTM networks to analyze and forecast time-series data in the field of meteorology, where the networks can be trained on large datasets of weather data to make highly accurate predictions of weather patterns [4].

The Suitability of Using LSTM for Hydrology Applications

LSTM networks are a useful tool for making predictions about time series data and have many potential applications in the fields of hydrology and environmental science [1]. Examples of common formats for storing hydrology data include tabular, time-series formats, and spatial formats. These formats are well-suited for different types of hydrology data and can be used with a variety of software tools for analysis and visualization. One example of a technique used to explain the predictions of deep neural networks is layer-wise relevance propagation (LRP) is a method that assigns a score to each of the input variables for a given neural network prediction, indicating the extent to which they contributed to the prediction [3]. Especially if we consider hydrological applications LSTM is a particularly effective network because it can model the evolution of states over time and map them to a specific output, which makes them effective at modeling complex hydrological systems [1].

Different Methods to Improve Interpretability

Other than LSTM, ANN (Artificial Neural Network) is also efficient in many examples for increasing explainability in research. Although LSTM is commonly used in hydrology, the choice of solution depends on the specific problem being addressed. In particular, ANN has popularity for solving problems related to rainfall-runoff prediction, stream-flow prediction, and groundwater hydrology [5]. The key ability of ANN models is their ability to forecast future events. LSTMs are able to capture patterns and dependencies in time-series data that are not easily modeled by ANNs. LSTMs have several advantages over traditional ANNs, including their ability to learn long-term dependencies and their flexibility with different data types. [2]. However, they can also be more difficult to train, which can be a disadvantage in some cases. One the other hand, the type of data they can process should be considered. ANNs are typically used for processing static, structured data, such as images, text, or tabular data [5]. In contrast, LSTMs are designed to process time-series data, such as sequences of words in natural language, time-series sensor data, or other types of data that have a temporal component [1], [2], [3]. Both ANN and LSTM can be used to increase interpretability, and there are many examples of their use for this purpose. There are some other factors that should be considered that can have a significant impact on the success of the chosen technique, and should be taken into account when making a selection.

Conclusion

Explainable machine learning techniques can help scientists to gain insights into the underlying patterns and trends in the data but it is crucial to select the appropriate technique in order to obtain the desired solution [4]The use of LSTM networks in some tasks in hydrology offers the benefit of being able to make a single, multi-step forecast. This is particularly useful for tasks such as rainfall-runoff forecasting [1]. There are several techniques that can be used for tasks like rainfall-runoff forecasting in hydrology [2]. However, the choice of which technique to use should be based not only on the solution and the task at hand, but also on the structure of the data. The type and complexity of the data, as well as the specific goals of the project, can all influence the decision of which technique to use.

References

[1] F. Kratzert, M. Herrnegger, D. Klotz, S. Hochreiter and G. Klambauer, "NeuralHydrology – Interpreting LSTMs in Hydrology," in *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen and K. Müller, Eds., Cham, Springer International Publishing, 2019, p. 347–362.

[2] Y. Sudriani, I. Ridwansyah and H. A. Rustini, "Long short term memory (LSTM) recurrent neural network (RNN) for discharge level prediction and forecast in Cimandiri river, Indonesia," *IOP Conference Series: Earth and Environmental Science,* vol. 299, p. 012037, July 2019.

[3] L. Arras, J. Arjona-Medina, M. Widrich, G. Montavon, M. Gillhofer, K.-R. Müller, S. Hochreiter and W. Samek, "Explaining and Interpreting LSTMs," in *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen and K. Müller, Eds., Cham, Springer International Publishing, 2019, p. 211–238.

[4] R. Roscher, B. Bohn, M. F. Duarte and J. Garcke, "Explainable Machine Learning for Scientific Insights and Discoveries," *IEEE Access,* vol. 8, pp. 42200-42216, 2020.

[5] R. Tanty and T. Desmukh, "Application of Artificial Neural Network in Hydrology- A Review," *International Journal of Engineering Research and,* vol. V4, June 2015.