Explainable AI for Time Series Forecasting

Exposé

Marisa Lange, Laura Plodek

March 15, 2024

Introduction

With the main focus on applying explainable AI techniques to predict time series data, our collaboration partner BASF has directed us to 50-year annual records of canola yield in the Saskatchewan region, Canada. We will use this dataset to assess various explanatory methods regarding the feature contribution to individual observations in the context of machine learning models.

Explainable AI (XAI) methods are immensely popular as they enable a comprehensive understanding of how AI systems arrive at their conclusions. The transparency resulting from these methods fosters increased trust and confidence in models. Another advantage lies in their ability to produce additional insights into the data through the AI's decision-making process. This aspect, in particular, will be the focal point of our project.

Our aim is to explore various methods of feature importance specifically at a local level. This involves examining explanations for specific predictions, or in the context of time series, understanding the significance of specific features in particular segments of the time series. Additionally, we emphasize the utilization of agnostic methods, which are applicable to diverse machine learning models rather than being limited to one specific type. Two widely used methods meeting these criteria are SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) values. Stemming from the principles of game theory, SHAP values employ Shapley values to determine the contribution of each feature to a given prediction and can be constructed locally or globally, while LIME values approximate the model exclusively locally with simpler, less computationally expensive interpretable models (Molnar, 2018).

For the main part of our investigation, we plan to explore the behavior of the two methods when applied to the obtained dataset and evaluate the conclusions and results derived from them. Depending on our analysis results, we may intend to explore potential extensions to these methods.

Data

The dataset supplied by our collaboration partner pertains to agricultural crop records, specifically the cultivation of Canola in the Saskatchewan region of Canada since 1970. As one product line of BASF consists of crop protection supplies, this use case is highly relevant for the company.

Our goal is to identify and analyze the influence of weather conditions, as well as other external factors, on the yield of Canola. To achieve this, we begin by merging two separate datasets. The first set contains information on crop yields across various rural municipalities in Saskatchewan, accessible online through the official website of the Saskatchewan region (The Government of Saskatchewan, 2023). It provides details on yields measured in tons for each municipality over the specified period, often dating back to 1970. The second set comprises weather-related information for the corresponding time span and is also publicly available via the CDS Climate Copernicus website (Copernicus Climate Change Service, 2023).

Our aspiration is to gain insights into the influence of weather conditions on the average yield in this region, with a particular emphasis on local feature contribution. The overarching goal of the project is to utilize these findings to understand the specific consequences of various weather effects on Canola cultivation. Given the increasingly extreme changes in weather expected as a result of climate change, it is crucial to gather as much knowledge as possible about the influence of specific meteorological indicators on agricultural output. Understanding and explaining effects observed in the past is an indispensable part of sparking changes in agriculture as we move to adapt to the weather conditions likely to be experienced in the future.

Methodology

In the first step, a suitable target variable is constructed in the context of a relevant use case. To ensure high data quality, an authentic public source is consulted for consistent and correct information on the variable over a sufficiently long time span. Following the check of the data for potential pre-processing requirements, such as filling missing values or parsing timestamps, relevant features contributing to explaining the target variable are defined. To do so, basic contextual knowledge will be acquired through a literature review, while domain experts from within the collaboration partner's network will be consulted. Corresponding data will also be obtained from authentic public sources and analyzed exploratorily, e.g., through correlation analyses, to construct variables for modeling.

In the following step, appropriate models are selected. We plan to train 2-3 machine learning models, such as XGBoost or certain types of neural networks, as these are suitable choices for time series forecasting purposes Phan et al., 2020. We will also train a more traditional statistical model, e.g., ARIMA, as a reference model for the quality assessment of metrics used to quantify feature importance. In the next step, 2-3 suitable metrics, such as SHAP and LIME values, will be defined. The models will be trained, and feature importance information extracted using the metrics defined.

As the main focus of our project work is on exploring the quality, i.e., the accuracy and robustness of these metrics in application to machine learning models, values will be calculated for the trained models under different specifications in an extensive interpretability analysis.

Any project-related coding work is carried out using Python in Jupyter Notebook files in a Kaggle environment, which will provide sufficient computing resources. For version control and collaboration purposes, a GitHub repository is set up and maintained.

References

Copernicus Climate Change Service. (2023). Cds.climate.copernicus.eu [Accessed: 2023-13-11].

Molnar, C. (2018). A guide for making black box models explainable. *URL: https://christophm. github. io/interpretable-ml-book*, *2*(3).

Phan, Q.-T., Wu, Y.-K., & Phan, Q.-D. (2020). A comparative analysis of xgboost and temporal convolutional network models for wind power forecasting. *2020 International Symposium on Computer, Consumer and Control (IS3C)*. https://doi.org/10.1109/IS3C50286.2020.00113

The Government of Saskatchewan. (2023). Saskatchewan.ca [Accessed: 2023-13-11].