Explainable Weather Forecasts Through an LSTM-CBR Twin System

Craig Pirie\textsuperscript{1*}, Malavika Suresh\textsuperscript{1}, Pedram Salimi\textsuperscript{1}, Chamath Palihawadana\textsuperscript{1} and Gayani Nanayakkara\textsuperscript{1}

\textsuperscript{1}School of Computing, The Sir Ian Wood Building, Garthdee Road, Robert Gordon University, Aberdeen, Scotland, United Kingdom, AB10 7GJ

Abstract
In this paper, we explore two methods for explaining LSTM-based temperature forecasts using previous 14 day progressions of humidity and pressure. First, we propose and evaluate an LSTM-CBR twin system that generates nearest-neighbors that can be visualised as explanations. Second, we use feature attributions from Integrated Gradients to generate textual explanations that summarise the key progressions in the past 14 days that led to the predicted value.

Keywords
LSTM, XCBR, NLG, Integrated Gradients, Forecasting, Visualisation

1. Introduction
Inaccurate weather forecasting can have a serious impact to life, property and businesses. For instance, farmers rely on weather forecasting to plan activities such as ploughing, harvesting and cultivation [1]. In 2018, the inability of weather forecasts to detect an early monsoon indirectly trapped 12 Thai children and their football coach in a cave for 18 days — and caused the death of two of their rescuers \textsuperscript{1}. Efficiently predicting weather can help to minimise these losses in the future [2]. Howbeit, weather is stochastic and its patterns are non-linear, making it a challenge to obtain precise predictions. The problem lends itself well to neural networks for their ability to model non-linearity. Particularly, weather forecasting suits sequential methods such as the Long Short-Term Memory (LSTM) [3], to allow for the utilisation of time-series data in the prediction process [2]. By making use of the recurrent nature of the LSTM, information can be fed back through different steps along a time-series, capturing trends, seasonality and more, which all help increase predictive power.
Although weather forecasts are common in everyday life, we seldom seek for deeper insights to explain how a prediction was made. On the contrary — if we did ask for further explanation — we could pass better judgement on the validity of predictions. Explainable AI (XAI) has become increasingly prevalent in recent times. This is because despite deep learning methods now excelling in performance in a variety of domains, they often remain as black-boxes; we are unable to ‘look inside’ to understand why a prediction was made. Methods such as Integrated Gradients [4] have been useful for producing saliency maps in images to highlight the important pixels that weigh heavily on the outcome of a classification. They have also been applied in a time-series setting but are often unsophisticated and can be difficult to interpret — presenting a need for dedicated time-series solutions [5, 6]. Literature shows that some (albeit few) attempts have been made to explain weather forecasts in a time-series setting such as by score-maximisation and occlusion analysis visualisations [7], or applying LIME [8]. However, none of these methods propose the use of case-based reasoning (CBR) to provide explanations. Using an approach grounded in similarity, such as CBR, a posteriori knowledge can be drawn upon to help both prediction and explanation. Past experiences can be drawn upon to make arguments for why a decision was made. We propose an approach that takes the learned LSTM embeddings from the prediction stage and uses these to build a case-base to provide explanations that can be visualised. We will also use Integrated Gradients and NLG to provide further explanation by means of a report.

2. Methodology

2.1. Predicting Temperature with LSTM

We use the default prediction model provided for the challenge which consists of two bidirectional LSTM layers.

2.2. Building the Twin LSTM-CBR System

We use the Clood CBR system as the twin. Clood CBR is a distributed CBR system which supports high scalability. The system has a micro-service architecture which splits an application into a set of smaller and interconnected services that scale to meet varying demands [9]. The output embeddings from the final LSTM layer are extracted for each input row and are fed into the Clood CBR system to form the case base. For a given row and LSTM prediction that we want to explain, we query the Clood system using the nearest-neighbors approach. This gives us the top three similar rows based on the embedding similarity. To establish the quality of the twin, we compare our approach to a baseline twin system that uses the raw features instead of the embeddings.

We use Mean Absolute Error (MAE) as the evaluation metric to compare the two twin systems as it is one of the common measures of forecasting error in time-series analysis. MAE measures the absolute difference $|y_i - x_i|$ between a prediction $y_i$ and an actual observation $x_i$ where the individual differences share the equal weight. The 3 nearest neighbours for a prediction are retrieved from the case-base and the mean MAE is calculated between their predicted values.

2 https://github.com/rgu-computing/cloud
Comparing the mean gives an indication of whether the cases are better aligned through the use of embeddings — or not.

\[ MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \]  

(1)

2.3. Generating Explanations

A template-based Natural Language Generation approach to generate explanations was taken. Based on Integrated Gradients, we extracted feature importance attributions for each day over three different time frames. These were as follows: the entire two-week period; the final week; and the final three days. The prediction, feature attributions and general information from the dataset were then slotted into the template to produce a weather forecast report (as seen in Figure 1). Furthermore, to aid interpretability, we add visualisations (emojis, graphics and colour) to our textual explanation to indicate the weather prediction.

Figure 1: An example weather report showing the prediction, an explanation of the most important features over two-week, one-week and 3-day periods. A thermometer with a high reading, a hot-faced emoji and red-text are included to help visualise the results.

3. Results

3.1. LSTM-CBR Twin Evaluation

The results in Table 1 prove that the twin LSTM-CBR system based on embeddings performs better than a twin modelled around raw features. The embedded approach achieved an MAE of 0.066 degrees whereas the raw features approach achieved an MAE of 0.133 degrees. This means that using embeddings, the twin system is able to find better nearest-neighbours, therefore can produce better explanations.

3.2. Nearest Neighbors Explanations Using LSTM-CBR Twin

In Figures 2 and 3, we visualize the progression of the average humidity and pressure, respectively, over the previous 14 days for a sample query row, top 3 nearest neighbors and 3 randomly selected rows. As can be seen, the query and the nearest neighbors have a similar progression of the predictor variables as well as close predictions for the subsequent day temperature.
Table 1
LSTM-CBR MAEs (Embedded & Raw) compared with the original LSTM predictions.

<table>
<thead>
<tr>
<th>Embedded Features</th>
<th>Raw Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.066</td>
<td>0.133</td>
</tr>
</tbody>
</table>

Figure 2: Nearest Neighbors Explanation (Average Humidity): Relation=1 depicts the query, Relation=2 depicts the 3 nearest neighbors from the lstm-cbr twin, Relation=3 depicts 3 randomly chosen data points (best viewed in colour)

Figure 3: Nearest Neighbors Explanation (Average Pressure): Relation=1 depicts the query, Relation=2 depicts the 3 nearest neighbors from the lstm-cbr twin, Relation=3 depicts 3 randomly chosen data points (best viewed in colour)

4. Conclusion

In this paper we explore two methods for explaining LSTM predictions of weather forecast. We first propose an LSTM-CBR twin system that explains a prediction using nearest-neighbors and then visualises the explanations through parallel coordinate plots. We show that using the LSTM embeddings can result in better twinning compared to raw features. As a second explanation method, we make use of feature attributions from Integrated Gradients in text
templates to generate text explanations. We note that a current limitation of the twin system is that it considers all days as equally important when retrieving the nearest-neighbors. Therefore, future work could explore different weighting strategies (such as weights derived from feature attributions) for strengthening the twin system. A second limitation lies with the need to provide a template for the NLG-based explanation, which would differ between domains. Further work could explore the application of more advanced NLG techniques that allow the inference of templates in a new domain.

Acknowledgments

We would like to thank Professor Nirmalie Wiratunga for motivating us to take part in this challenge and for her advise throughout the competition.

References


