Meteorologic Real-time Extreme Learning Machine for Pressure Prediction

Anagha Ram
Columbia University
New York, NY, USA
ar4462@columbia.edu

Kate Keahey (advisor)
Argonne National Laboratory
Lemont, IL, USA
keahey@mcs.anl.gov

Alicia Esquivel Morel (advisor)
University of Missouri
Columbia, MO, USA
ace6qv@mail.missouri.edu

Abstract

Significant advances in weather prediction stemmed primarily from combining observational data, sophisticated modeling techniques, and analysis of simulated or historical weather data. This study builds on prior research using an Extreme Learning Machine (ELM) approach to detect weather anomalies in real-time, enhancing existing predictive systems. Our model is implemented on IBIS, an adaptable edge computing framework for multi-sensor data collection. The ELM model, applied to detect real-time weather anomalies, offers fast training and operational efficiency. Data on atmospheric phenomena, including pressure and wind, was generated and stored in a time series database. The model was then trained on 80% of 550,000 records. Our experiments demonstrated a 92% R^2 score, supporting its effectiveness. Our work within IBIS represents a cost-effective and scalable solution for collecting, monitoring, and predicting hazardous atmospheric conditions.

CCS Concepts: • **Computing methodologies** → ***Artificial Intelligence; Machine Learning***.

Keywords: Meteorologic Real-time Extreme Learning Machine Pressure Prediction on the Edge

ACM Reference Format:
Anagha Ram, Kate Keahey (advisor), and Alicia Esquivel Morel (advisor). 2024. Meteorologic Real-time Extreme Learning Machine for Pressure Prediction. In *Proceedings of ACM Student Poster Competition (SC '24).* ACM, New York, NY, USA, 3 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Weather prediction and meteorological data collection can be challenging [8]. Our goal is to simplify this process, providing accurate Extreme Learning Machine (ELM) model for real-time analysis and anomaly detection. Our ELM operates within IBIS [1], a framework for edge and multi-sensor data collection. IBIS gathers meteorologic pressure and wind data, which feeds into our model. ELM, introduced by Huang et al. in 2006, is valued for its fast learning, ease of implementation, good generalization, and suitability for real-time analysis [6]. Unlike feed-forward neural networks (FFN), which require multiple iterations and gradient-based back-propagation, ELM computes output layer weights in a single step using the Moore-Penrose inverse of the hidden layer's output matrix [6]. Although machine learning on the edge has been explored [2, 3, 10, 11], ELM has not been a major focus. ELM has been used in weather prediction, but it has not been applied on edge devices. For example, Petros Karvelis et al. proposed a lightweight onboard system for real-time wind prediction using training-less machine learning [7] and Mimoun Moussaoui deployed a model based on FIRMS datasets to IoT devices [5]. Our work extends their research by implementing ELM on edge devices for weather prediction, with promising results.

SC '24, Nov 12-17, 2024, Atlanta, GA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/XXXXXXX.XXXXXXX

2 Research Methodology

2.1 Data collection and preparation

Data Ingestion: Data from January 1, 2024 to January 10, 2024 is queried from InfluxDB, encompassing readings from two pressure sensors and one wind sensor, as illustrated in **Figure 1**.

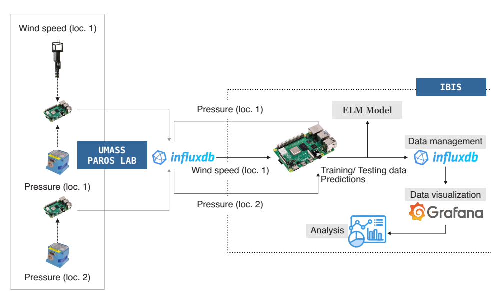


Figure 1. Comprehensive End-to-End Pipeline for ELM Model Training deployed on IBIS; Sensors located in Paros Center for Atmospheric Research [4]

The dataset comprises three comma-separated files: pressure data from location 1, wind speed data from location 1, and pressure data from location 2.

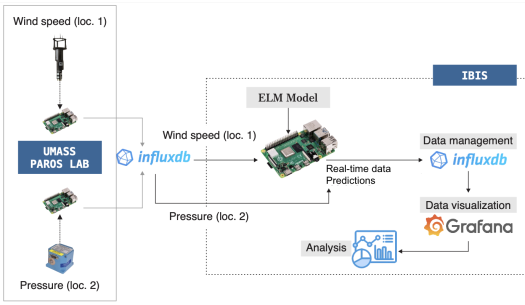


Figure 2. Comprehensive End-to-End Pipeline for ELM Model Real-time Runtime Phase deployed on IBIS

Data Preparation We corrected fractional timestamp discrepancies in the ingested data by standardizing the index using wind data timestamps and re-indexed the Pandas pressure data frames, interpolating as necessary. The synchronized data frame, containing pressure (loc. 1), pressure (loc. 2), and wind data, with about 550,000 records, was used for model creation and uploaded to InfluxDB for visualization.

Model Creation In the model, *pressure_2* and *wind speed* are treated as independent variables, while *pressure_1* serves as the dependent variable. The dataset, consisting of 550,000 records, is split into training and testing sets with an 80/20 ratio. The ELM model is trained with 96 hidden units and a sigmoid activation function

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Since this is a regression problem, all data is handled as floating-point numbers. Training typically takes a few minutes, and model performance is evaluated using the R^2 score.

2.2 ELM Real-time Prediction

The ELM model, written in Python, is deployed in a ARM architecture Docker container on the Chameleon testbed [9] for data storage and IBIS [1] for weather instrumentation. The real-time program queries InfluxDB for pressure and wind speed data, which is standardized on timestamps. The trained model predicts *pressure_1* from *pressure_2* and *wind speed*, with predictions and anomalies sent to InfluxDB for visualization.

2.3 Visualization

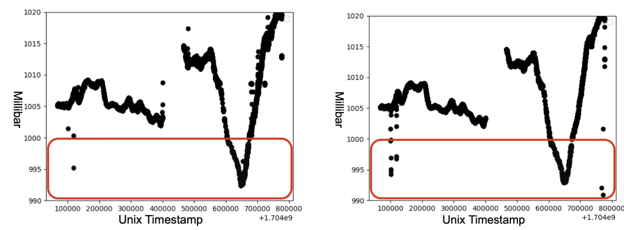
The Grafana dashboard was used to connect to InfluxDB, enabling the visualization of various data sets. The following time series were displayed: measured *pressure_1*, measured *pressure_2*, measured wind data, and *pressure_1* predicted by the ELM model.

3 Evaluation

With 80% of the 550,000 records used for training and 96 hidden units with a sigmoid activation function, the ELM model achieved a high accuracy of 92%. This accuracy was measured using the R^2 score, as shown in **Equation 1**.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1}$$

Figures 3 a and **b** present plots generated using the Matplotlib library, comparing the measured pressure 1 data with the predicted pressure 1 data for the testing period. Both figures exhibit similar patterns, aligning with the high R^2 score. Additionally, **Figures 4**, **5**, **6**, and **7** display Grafana dashboard graphs, facilitating both historical and real-time data visualization.



(a) Predicted pressure (b) Measured pressure

Figure 3. Time series plots of the predicted and measured pressures at Location 1 over time with boxed anomalies (y-axis: millibars, x-axis: Unix timestamp).

Figure 4. Time series plot in Grafana dashboard for measured pressure at Location 1 over time (y-axis: millibars, x-axis: timestamp).

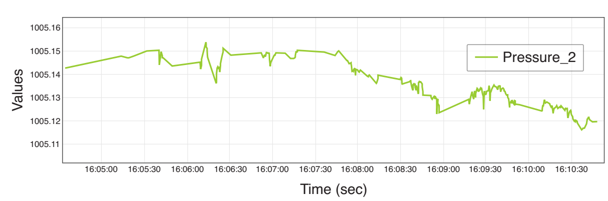


Figure 5. Time series plot in Grafana dashboard for measured pressure at Location 2 over time (y-axis: millibars, x-axis: timestamp).

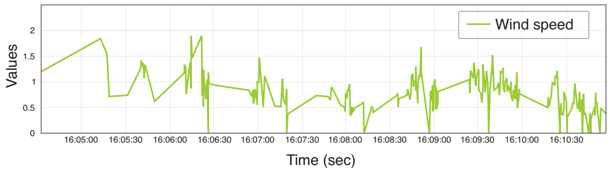


Figure 6. Time series plot in Grafana dashboard for measured wind speed over time (y-axis: m/s, x-axis: timestamp).

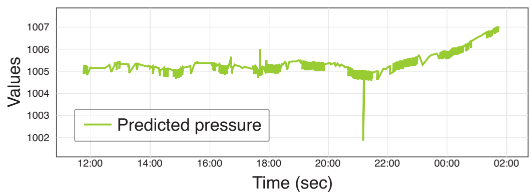


Figure 7. Time series plot in Grafana dashboard for predicted pressure at location 1 over time (y-axis: millibar, x-axis: timestamp).

4 Conclusion

The project assesses the real-time performance of our Extreme Learning Machine (ELM) model on IBIS, an edge computing platform. Results are promising, with the ELM achieving an R^2 score of approximately 92%, demonstrating its effectiveness as a real-time model. IBIS was selected for its cost-effectiveness and scalability. The next step is to replace InfluxDB in Figures 1 and 2 with a direct data pipeline from the two IBIS devices, connected to three sensors.

Acknowledgments

Results presented in this paper were obtained with support from the National Science Foundation through the Chameleon testbed (Award Number 2027170) and the NDCC supplement (Award Number XXXXXXX). This material is based upon work supported by the U.S. Department of Energy, Office of Science, under contract number DE-AC02-06CH11357.

References

[1] K. Keahey Z. Murry T. Sitzmann J. Zhou A. Esquivel Morel, M. Powers and P. Calyam. 2024. IBIS — An Infrastructure Management Framework for Adaptable, Multi-Sensor Data Collection in Scientific Research. *ISC High Performance 2024 International Workshops* (2024). https://doi.org/10.13140/RG.2.2.11758.22082 Preprint available on ResearchGate.

[2] N. Acharya, N. A. Shrivastava, B. K. Panigrahi, and U. C. Mohanty. 2014. Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: an application of extreme learning machine. *Climate Dynamics* 43, 5–6 (2014), 1303–1310. https://doi.org/10.1007/s00382-013-1942-2

[3] L. Chen, B. Han, X. Wang, J. Zhao, W. Yang, and Z. Yang. 2023. Machine Learning Methods in Weather and Climate Applications: A Survey. *Preprints* 202309.1764.v2 (2023). https://doi.org/10.20944/preprints202309.1764.v2

[4] Paros Center for Atmospheric Research. [n. d.]. *Paros Center for Atmospheric Research*. https://www.umass.edu/paros-atmospheric-center/

[5] M. Grari, I. Idrissi, M. Boukabous, O. Moussaoui, M. Azizi, and M. Moussaoui. 2022. Early Wildfire Detection Using Machine Learning Model Deployed in the Fog/Edge Layers of IoT. *Indonesian Journal of Electrical Engineering and Computer Science* 27, 2 (Aug. 2022), 1062–1073. https://doi.org/10.11591/ijeecs.v27.i2.pp1062-1073

[6] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew. 2006. Extreme learning machine: Theory and applications. *Neurocomputing* 70, 1-3 (2006), 489–501.

[7] P. Karvelis, D. Mazzei, M. Biviano, and C. Stylios. 2020. PortWeather: A Lightweight Onboard Solution for Real-Time Weather Prediction. *Sensors* 20, 11 (Jun. 2020), 3181. https://doi.org/10.3390/s20113181

[8] ŞM Kaya, B İşler, AM Abu-Mahfouz, J Rasheed, and A AlShammari. 2023. An Intelligent Anomaly Detection Approach for Accurate and Reliable Weather Forecasting at IoT Edges: A Case Study. *Sensors (Basel)* 23, 5 (2023), 2426. https://doi.org/10.3390/s23052426

[9] Kate Keahey, Jason Anderson, Zhuo Zhen, Pierre Riteau, Paul Ruth, Dan Stanzione, Mert Cevik, Jacob Colleran, Haryadi S. Gunawi, Cody Hammock, Joe Mambretti, Alexander Barnes, François Halbach, Alex Rocha, and Joe Stubbs. 2020. Lessons learned from the Chameleon testbed. In *Proceedings of the 2020 USENIX Conference on Usenix Annual Technical Conference (USENIX ATC'20)*. USENIX Association, USA, Article 15, 15 pages.

[10] Z. Prathyusha, Z. Savya, T. N. Alex, and S. C. C. 2021. A Method for Weather Forecasting Using Machine Learning. In *2021 5th Conference on Information and Communication Technology (CICT)* (Kurnool, India). 1–6. https://doi.org/10.1109/CICT53865.2020.9672403

[11] Mahmut Yazici, Shadi Basurra, and Mohamed Medhat Gaber. 2018. Edge Machine Learning: Enabling Smart Internet of Things Applications. *Big Data and Cognitive Computing* 2, 3 (2018), 26. https://doi.org/10.3390/bdcc2030026