Predicting aurora borealis visibility combining satellite and citizen science data with artificial intelligence

Estevan Hernandez, Jacob Luton, Ty Woolven, and Dr. Erik Mayer, Pittsburg State University,
Dr. Tao Wu, Butler Community College, Dr. Hongsheng He, University of Alabama,
Dr. Elizabeth MacDonald, Goddard Space Flight Center

Pittsburg State University

Introduction

Our research project embarks on an innovative journey to predict the aurora borealis, or northern lights, by using a unique combination of satellite data and Aurorasaurus, a citizen science platform. The satellites provide crucial information on solar wind and geomagnetic activity while Aurorasaurus gathers public reports and observations of auroras. To predict where and when the auroras can be seen from Earth, our approach includes an advanced artificial intelligence (AI) algorithm known as logic tensor networks which combines neural networks with predicate logic.

This research is funded by the NASA EPSCoR Rapid Response Research grant Appendix F: A Neural-Symbolic Aurora Model Driven by Aurorasaurus Data in Citizen Science. The Aurorasaurus site currently uses the Ovation Prime model which uses satellite data to predict the probabilities of viewing the aurora at different locations. However, the Ovation Prime does not completely agree with citizen science reports. This research attempts to supplement the Ovation Prime model with logic tensor networks to produce more accurate predictions.

The current focus is on obtaining code for the Ovation Prime model. Plans are then to combine the probabilities generated by Ovation Prime and the Aurorasaurus sightings with logic tensor networks. The use of an autoencoder to reduce the dimensionality and noise of the satellite data is also being investigated. In addition, curriculum is being developed to introduce the technique of logic tensor networks to Electronics Engineering Technology classes such as EET 549 Advanced Microcontrollers and EET 745 Advanced Microprocessor Systems and Applications in the Spring 2024 semester.

Aurora predictions and citizen science data

Ovation Prime is a auroral precipitation model that provides a prediction of the visible aurora in both the Northern and Southern Hemispheres. It was written in Interactive Data Language (IDL), a programming language used for data analysis. However, in this case GDL (GNU Data Language), a free alternative for IDL will be used. GDL should be able to run any IDL code while being free to use. When using Ovation Prime input data from a satellite is provided, Ovation Prime then gives an output showing a prediction for visible aurora. This is illustrated in Fig. 1 using input data from OMNI.

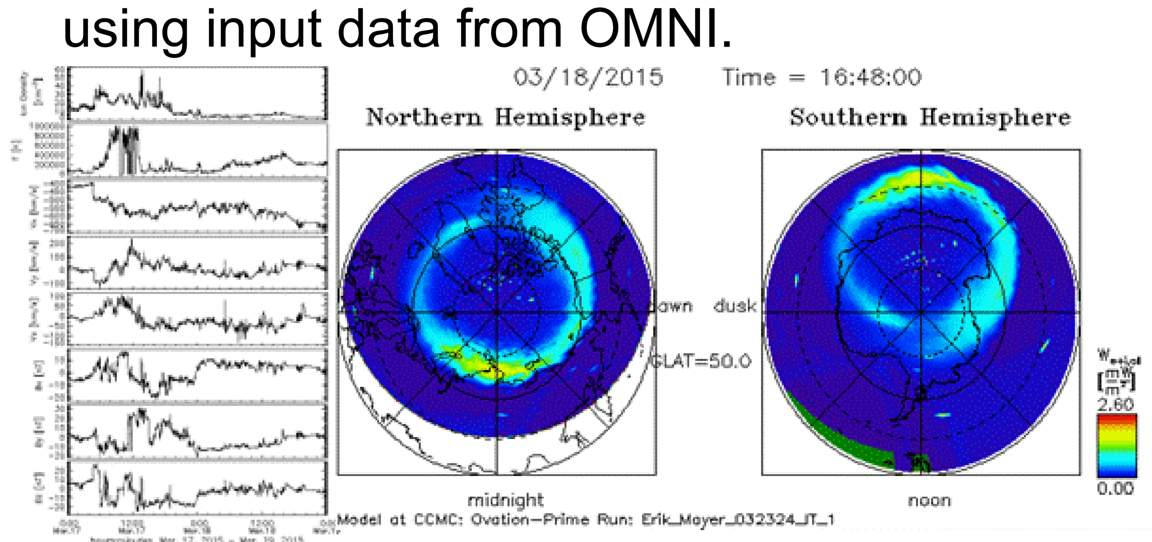

Fig. 1 **OMNI satellite data and Ovation Prime prediction (Source: https://ccmc.gsfc.nasa.gov/results/viewrun.php?runnumber=Erik_Mayer_032324_IT_1)**

Aurora citizen science data is collected from general public on visibility of the northern lights. This data is collected from Aurorasaurusas, a website where users can report sightings of the aurora. When submitting a report, users are asked to provide several pieces of data including location and length of observation, colors of the aurora, and where it was in the sky. Since the citizen science data and Ovation Prime do not always agree both sources will be used to train the new neural network in an attempt to make a model that can more accurately predict the visible aurora based on satellite data.

Autoencoder neural network layers

Autoencoders begin by reducing the size and complexity of the input data. This "compression" is done by the encoder, which takes detailed data and transforms it into a more compact form. This is analogous to summarizing a detailed report into key points.

The decoder then takes this compressed data and works to reconstruct the original dataset as accurately as it can as shown in Fig. 2. The goal is for the autoencoder can understand and reproduce the key features of the data, despite having only the "summary" to work from.

After these steps, the compressed data is now a simplified essence of the original. This concentrated version of the dataset retains the most important information in a more manageable form. It is planned to fed his compact data into neural network layers utilizing logic tensor networks. Because it is less complex, it is easier for these networks to process the data, making them more efficient.

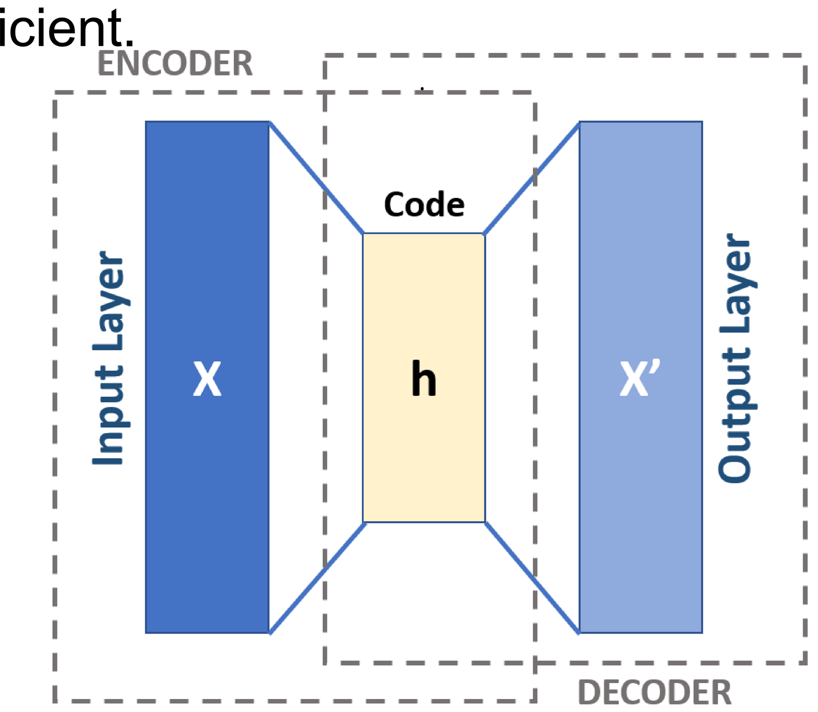

Fig. 2 Autoencoder training (Source: Autoencoder – Wikipedia)

Logic tensor network layers and curriculum development

Predicate logic or first-order logic consists of formulas that can be evaluated as true or false. For example, the formula "If the location is x and the solar wind measurements are y, then the aurora will be seen" can be generated from Aurorasaurus observations. In comparison, fuzzy logic uses formulas that use continuous values between 0 and 1. An example is the formula for the probability of seeing aurora as predicted by the Ovation Prime model.

Logic tensor networks integrate predicate logic and fuzzy logic with the use of neural networks. This makes them a viable solution to combine Aurorasaurus observations with the Ovation Prime model to obtain a model which will more accurately predict the possibility of seeing the aurora.

In addition to using logic tensor networks for aurora prediction, curriculum on logic tensor networks will be developed which will provide background material on fuzzy logic and predicate logic.

Acknowledgments

This material is based upon work supported by the National Aeronautics and Space Administration under Cooperative Agreement No. 80NSSC 22M0256. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration nor of Wichita State University.

Simulation results have been provided by the Community Coordinated Modeling Center (CCMC) at Goddard Space Flight Center through their publicly available simulation services (https://ccmc.gsfc.nasa.gov). The OVATION Prime 2.3 Model was developed by Patrick Newell and co-workers at the Johns Hopkins Applied Physics Laboratory.