Data - Learned Prediction of Space Weather

Sukumar Hakhoo (sh5661)
MS Student, Computer Science,
New York University

Agnit Mukhopadhyay [External Collaborator]
PhD Candidate in Space Sciences,
University of Michigan

IDEA

The project will utilize Machine Learning to predict upper-atmospheric (ionospheric) electrical conductance in the aurora/polar regions. The auroral conductance is severely affected by space weather phenomena like solar flares and coronal mass ejections (CMEs). This project will enable significant improvements in present predictive investigations into the physics and estimation of the auroral conductance.

PROBLEM

Global magnetohydrodynamic (MHD) models are used predominantly in space weather prediction. In almost all such global models, the estimation of the ionospheric conductance is conducted using outdated empirical models. This poses a massive challenge, when predicting quantities like the conductance in the aurora/polar region, which are driven mostly by electromagnetic interactions with charged particles, that MHD cannot predict accurately.

AIM

We intend to use machine learning to predict this quantity using inputs from in-situ satellite data and ground-based instruments. This ML-based predictive model will be designed to take multiple inputs and return the conductance as an output. Once operational, the model will be installed into the University of Michigan's Space Weather Modeling Framework (SWMF) to study improvements in space weather predictive skill.

DATA

DATASET

QUANTITY	DATASET	COVERAGE
Field Aligned Currents	Assimilative Mapping of Ionospheric Electrodynamics (AMIE)	2000 - 2010

FEATURES

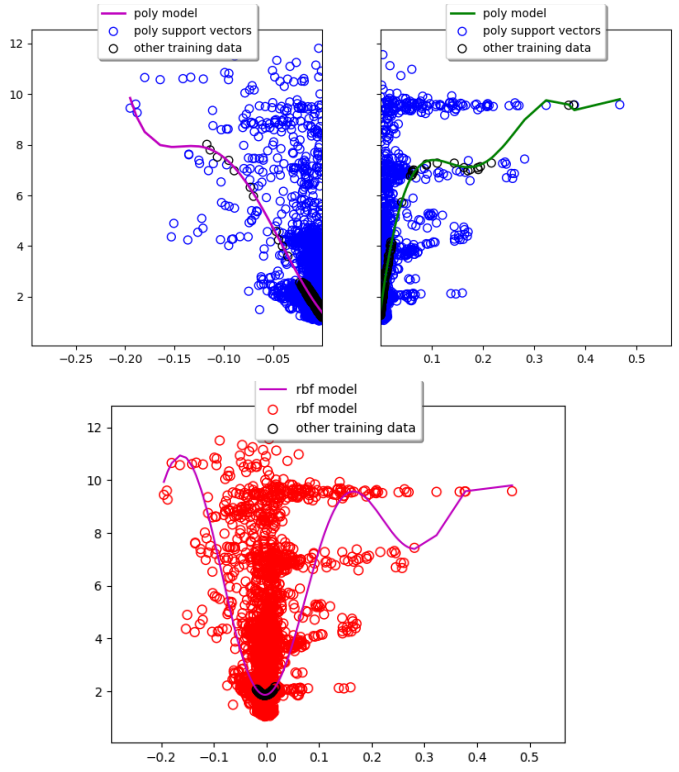
MODEL	LABELS		FEATURES		
Single Feature	Hall Conductance	Pedersen Conductance	Field Aligned Currents		
Multi-feature	Hall Conductance	Pedersen Conductance	Field Aligned Currents	Latitude	MLT

OBJECTIVE

MODEL TRAINING

SUPPORT VECTOR REGRESSION

The visualization of data suggested a considerable difference between a model that fits all the data in and two models trained on parts of the data ($x > 0$ and $x < 0$). The objective of classifying the data into two classes, one with input $x > 0$ and the other with $x < 0$ and training separate models on each, is to increase the accuracy of the entire model while not impacting its ability to generalize on new data.



SUPPORT VECTOR REGRESSION

The visualization of data suggested a considerable difference between a model that fits all the data in and two models trained on parts of the data ($x > 0$ and $x < 0$). The objective of classifying the data into two classes, one with input $x > 0$ and the other with $x < 0$ and training separate models on each, is to increase the accuracy of the entire model while not impacting its ability to generalize on new data.

RESULTS