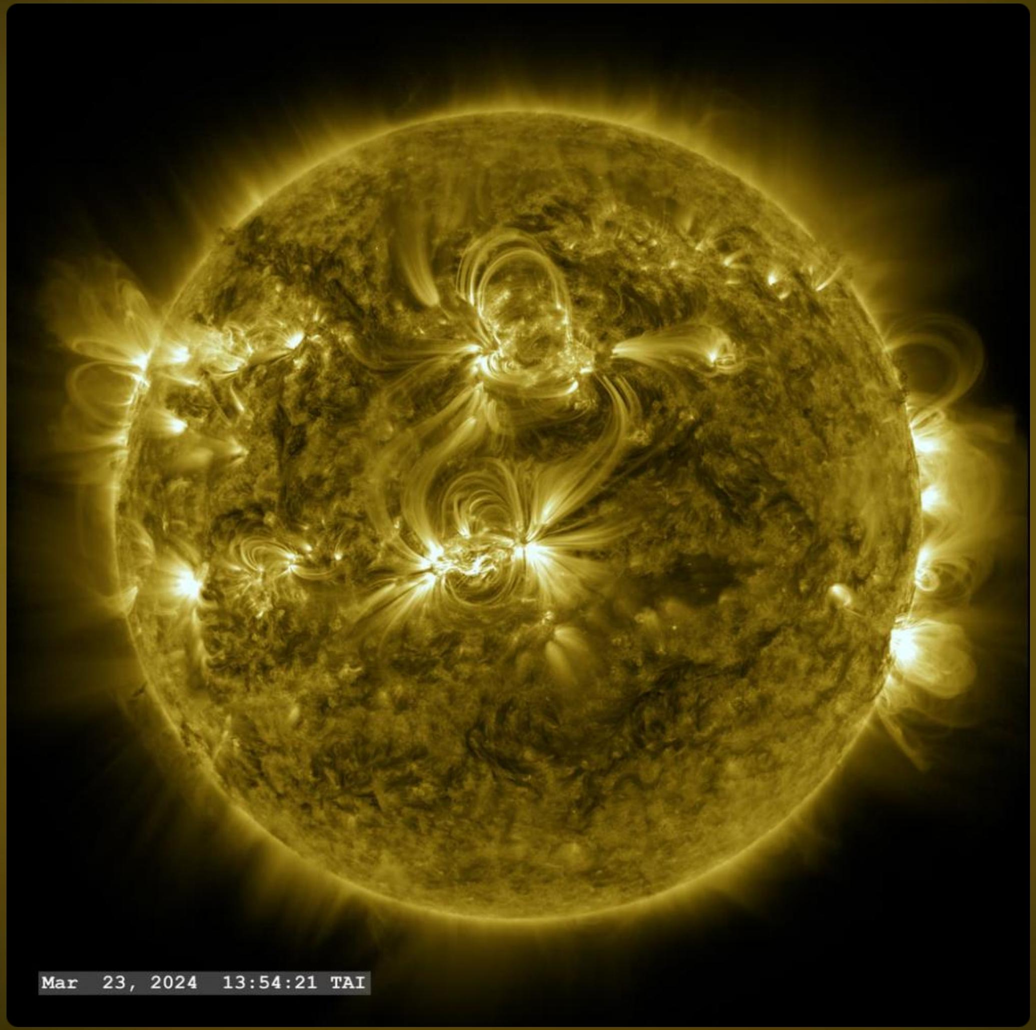
Stellar Flare Detection and Prediction Using Clustering and Machine Learning

Diksha Sen Chaudhury

University of Toronto, Department of Statistical Sciences.



Credit: NASA’s Scientific Visualization Studio

Background and Motivation

What are Stellar Flares?

- Sudden bursts of energy from stars due to magnetic reconnection.
- Play a crucial role in our understanding of the universe.

Challenges

- Irregular flare timing.
- Data lacks flare labels.

Existing Work

- Methods like HMMs and RNNs are computationally intensive.
- Forecasting flares remains unexplored.

Study Goal

Create a robust flare detection and future flare prediction model.

Data Source

Time-series of PDCSAP flux measurements from TIC 0131799991, observed by NASA's TESS satellite.

Methods and Results

1. **Preparing the Data**

   Data pre-processing, feature engineering from raw flux data.

2. **Optimizing Flare Detection Algorithm: DBSCAN**

   DBSCAN identifies flare candidates as noise points.

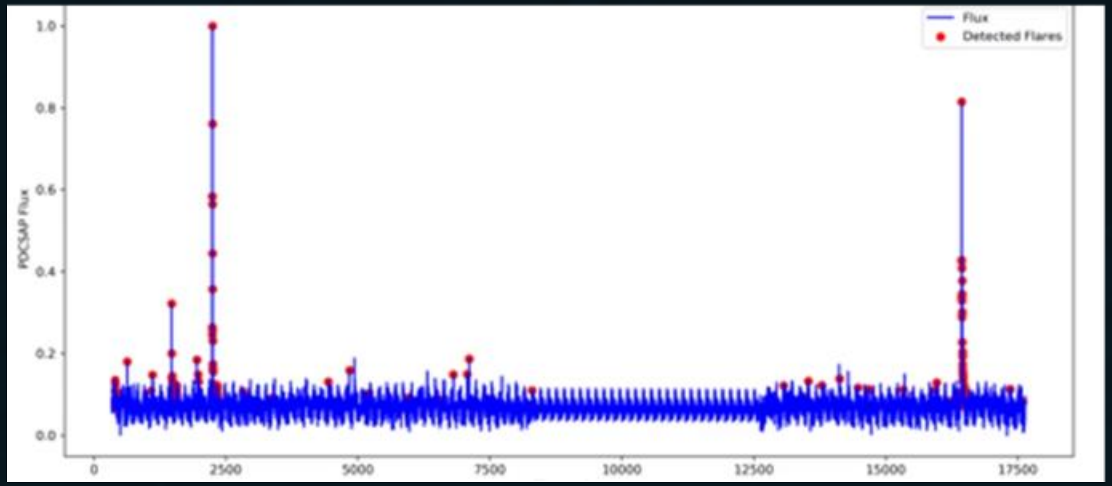
3. **Evaluate Flare Detection Algorithm: Simulations**

   Strong sensitivity in both cases; slightly reduced precision in the more realistic star-based simulation due to near-flare false positives.

4. **Predictive Model to Capture Future Flares: XGBoost**

   Use flare points flagged by DBSCAN to train a XGBoost model that can predict future flares. Model performs exceptionally well in identifying non-flare events and shows promising performance in detecting flares, despite their relative rarity.

DBSCAN Detection Results:



Simulation Results:

	Pareto-based Simulation	Realistic Star-based Simulation
Sensitivity	0.90	0.90
Precision	1.00	0.75
F1 Score	0.95	0.82

Evaluation metrics on the test data from the XGBoost model

	Not Flare	Flare
Sensitivity	1.00	0.68
Precision	1.00	0.87
F1 Score	1.00	0.76