CS229 - Solar Flare Prediction

Paul Warren
pwarren@stanford.edu

Gabriel Bianconi
bianconi@stanford.edu

December 2014

1 Problem

Solar flares, the release of energy the equivalent of 160,000,000,000 megatons of TNT in the form of high energy particles, occur when unstable magnetic field lines reconnect into a series of loops, releasing energy that accelerates high energy particles into space, often disrupting satellite operations and damaging astronaut health. Although we understand the general process, we don't know enough details to forecast flares with any reasonable degree of reliability. Large frames are always preceded by magnetic field activity, but magnetic field activity isn't always followed by large flares - it's common for two extremely similar magnetic fields to lead to two different results.

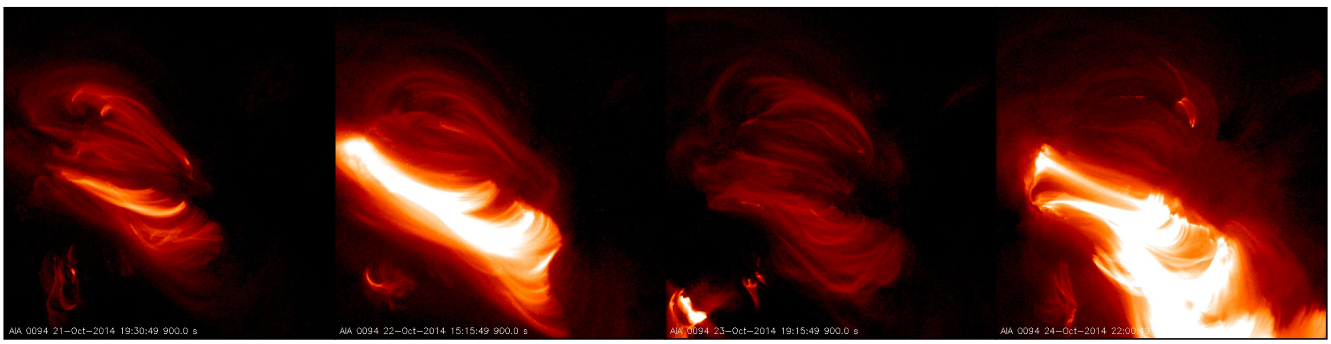

Figure 1: An image of the Sun in Fe XVIII, which shows plasma at temperatures of 4-8 megakelvin (MK). The very intense brightenings are solar flares. Frame 1 is more active than Frame 3, but Frame 3 leads to a larger much flares 24 hours later. It's also common to see two similar Frames where one leads to a flare and the other doesn't.

2 Literature Review

Machine learning is new to most of astrophysics, and solar physics is no exception: a search for the keywords "machine learning" on the Astrophysics Data System paper library returns only 180 abstracts. What little previous work exists focuses on vector magnetic field data or line-of-sight magnetograms [Bobra 2014]. After consulting a solar physicist, we decided to combine line-of-sight photospheric magnetic field data and images that show the high temperature corona, what we believe to be a novel feature set, to predict solar flares 24 hours in advance. The temporally resolved image data should give insight into stresses in the magnetic field that lead to the flares.

3 Data

The NASA Solar Dynamics Observatory (SDO) has been recording solar temperature and magnetic field data via the Atmospheric Imaging Assembly (AIA) and Helioseismic and Magnetic Imager (HMI), respectively, since May 2010. The NOAA Satellites (GOES) have been recording solar radiation data via the X-ray Sensors (XRS) since the 1970s.

AIA images give us information about the high-temperature corona in the form of snapshots of different types of iron plasma, which appear at various temperatures, moving across magnetic field lines. Fe XVIII, for example, is 4-8 million degrees hot, and Fe XXI (131) is 10-20 million degrees hot. Literature review suggested most large flares were preceeded by large amounts Fe XVIII and Fe XXI, although large amounts of Fe XVIII and 131 don't always lead to large flares. HMI images are magnetograms of the photosphere. GOES data measures the amount of X-rays emitted from the sun and tells us if a flare happened somewhere on the sun.

GOES data is taken for every 2 seconds. AIA data is collected every 900 seconds. HMI data is collected every 900 seconds. 4.5 day chunks of GOES, AIA, and HMI data were collected for the regions that produced each of the 50 largest flares within 30 degrees of the central meridian (the area we have high quality data from) since May 2010. This data was processed into 8 AIA snapshots (one every 12 hours) with the GOES and HMI data for a 24 hour period. Some files were corrupted or incomplete. We were left with 396 24 hour chunks, of which 22% had Class M2 or above flares (a max 131 value above 2.68e+07). All data was scaled by log10 for crunchability and normalized by the mean and standard deviation for comparability.

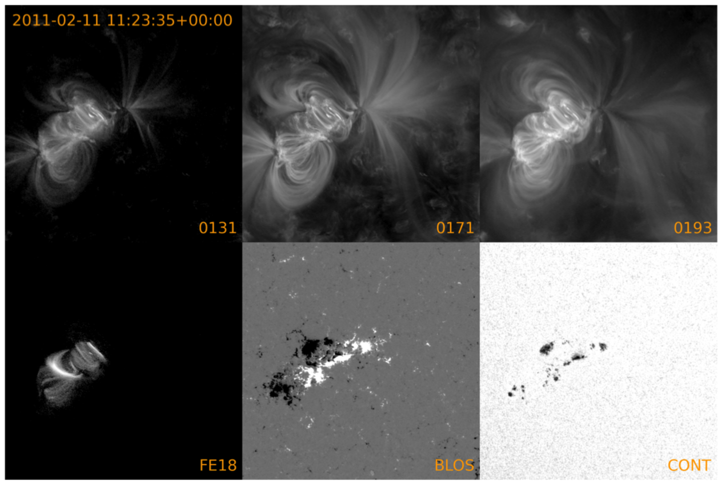

Figure 2: Intensities in the AIA images show the amount of Fe IX 171, Fe XII 193, Fe XVIII, and Fe XXI 131 on the sun. These emission lines appear at 1 million degrees, 1.5 million degrees, 4-8 million degrees, and 10-20 million degrees, respectively. BLOS and CONT are HMI images that show the photospheric magnetic flux and sunspots on the white light continuim, respectively. The challenge is to use these snapshots and the time evolution over the past 24 hours to predict if a flare will occur in the next 24 hours.

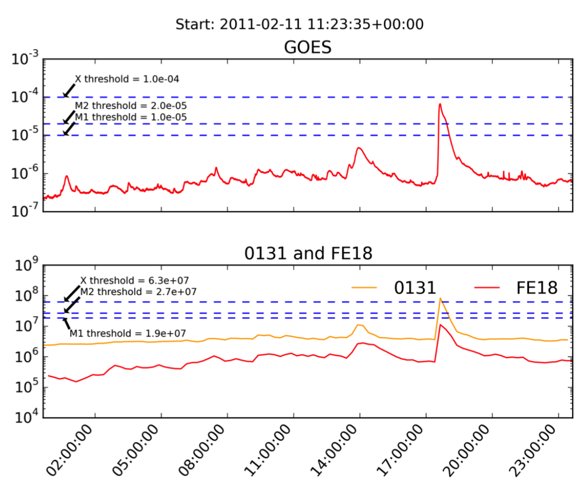

Figure 3: Lightcurves of GOES, 131, and Fe XVIII for 24 hours after the AIA images are taken

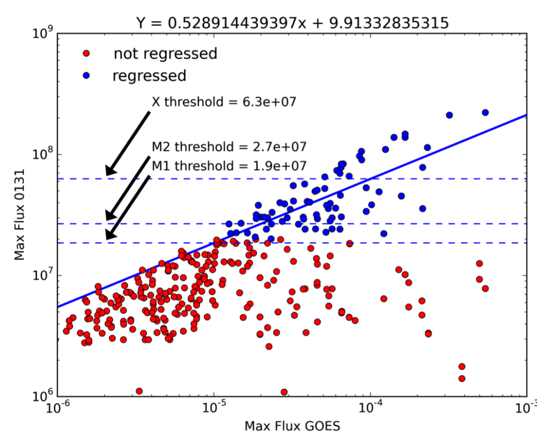

Figure 4: GOES tells us if there's a flare, but doesn't say where. We used linear regression to map the global GOES flare threshold to a local 131 flare threshold to determine when a flare happened in the region we were looking at.

4 Features

1. Total Unsigned Magnetic Fluc
2. Total FE XVIII
3. Standard Deviation FE XVII Past 12 Hours (if available)
4. Standard Deviation GOES X-Ray Past 12 Hours (if available)
5. Standard Deviation FE XXI 131 Past 12 Hours (if available)
6. Standard Deviation FE XVII Past 24 Hours (if available)
7. Standard Deviation GOES X-Ray Past 24 Hours (if available)
8. Standard Deviation FE XXI 131 Past 24 Hours (if available)
9. Standard Deviation FE XVII Next 24 Hours (if available)

Features were selected from previous work. Feature 1 gives us information about the photospheric magnetic field. Feature 2 gives us information about the high temperature corona, which is a good proxy for information about the coronal magnetic field; Features 3-5 tell us information about the solar activity over the past 12 hours for the 337 data points that have that information; Features 6-8 tell us information about the solar activity over the past 24 hours for the 290 data points that have that information; and Feature 9 tells us information about the magnetic flux over the next 24 hours, the period when we're trying to predict the solar flares.

5 Analysis

All machine learning algorithms were evaluated using two metrics: accuracy and True Skill Score (TSS). Accuracy may be skewed as a result of our data set being skewed towards flaring regions, but the TSS takes into account wrong guesses and goes from -1 to 1.

$$\text{True Skill Score} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} - \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}}$$

5.1 SVM

We used scikit-learn to determine the true skill score of an SVM for each pair of features. We trained and tested a two-dimensional SVM on each pair of features. We tried splitting the 396 data points into a testing and a training set, but our data set is so small our TSS on the training set varied from -1 and .4 from run to run, so all TSSs are calculated using the same training and testing set.

To make sure our SVM worked, we predicted flares over the next 24 hours using data about the variability of the plasma for the next 24 hours and, as expected, got a high result (a true skill score of >.8). Because the data was so linearly separable, a linear kernel was comparable to an RBF kernel. Most data points were thoroughly mixed, though, so RBF kernels usually outperformed linear kernels. Most SVMs performed similarly, with a TSS around .35. The best one being the total amount of Fe 18 and the standard deviation in 131 for the past 12 hours. This makes sense because both values relate to instability, and literature review suggests large flares are often preceded by instability (although instability doesn't always lead to large flares). The TSS are so low because of the high number of false positives that result from the fact that we're trying to predict data that's really difficult to categorize – if we had more regions without large flares, as previous work has, the TSS would likely be much higher. It's possible there are so many non-flaring data points near flaring data points because those regions had recently flared and active regions are unlikely to have large flares twice in a row. A lot of data points (e.g. 131 and Fe 18) seemed to be linearly correlated, and so wouldn't provide additional useful information if added to a multi-dimensional input learning algorithm.

SVM Feature Pairs	Accuracy	True Skill Score	SVM Feature Pairs	Accuracy	True Skill Score
2, 8	.75	.37	4, 9	.92	.84
2, 7	.71	.31	4, 3	.70	.32
2, 9	.91	.86	4, 1	.60	.27
2, 6	.74	.36	9, 7	.90	.82
2, 1	.70	.39	9, 6	.91	.83
5, 2	.74	.40	3, 2	.72	.36
5, 4	.68	.27	3, 9	.91	.84
5, 9	.91	.84	6, 7	.68	.33
5, 3	.70	.33	1, 2	.71	.36
5, 1	.73	.34	1, 8	.66	.27
8, 7	.62	.27	1, 7	.68	.25
8, 9	.91	.83	1, 9	.89	.83
8, 6	.66	.31	1, 3	.70	.34
4, 2	.71	.30	1, 6	.64	.36

Table 1: SVM results on each feature pair

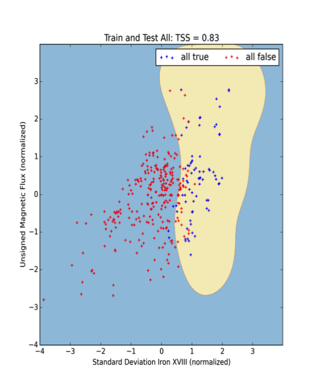

Figure 7. Predicting using future data to test algorithm, RBF kernel

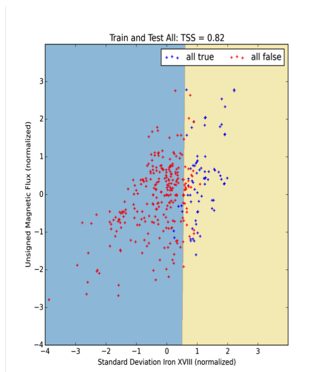

Figure 8. Predicting using future data to test algorithm, linear kernel

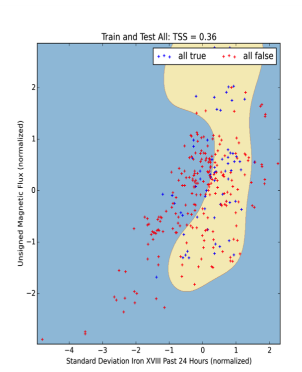

Figure 9. Predicting using past data, RBF kernel

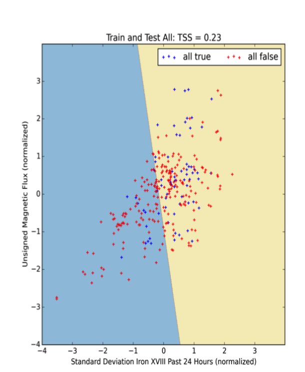

Figure 10. Predicting using past data, linear kernel

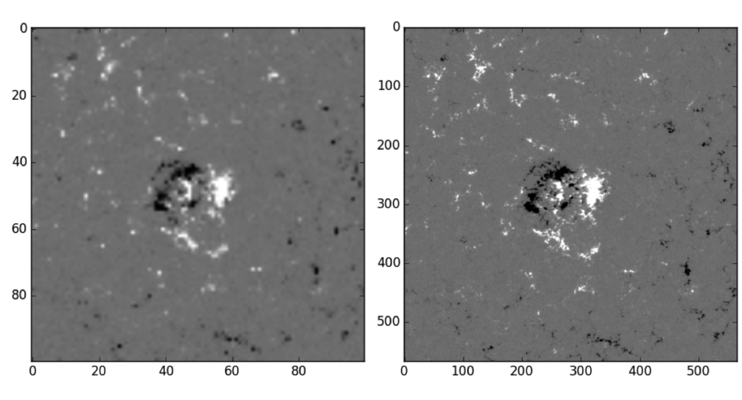

Figure 5: BLOS magnetograms were reduced from 594 x 594 to 100 x 100 and fed into a Neural Net

5.2 Neural Net

We also reduced the amount of data in a BLOS magnetic field data image and fed it to a neural net using the Fast Artificial Neural Net (FANN) library. The ANN had 10,000 input neurons, 50 hidden neurons, and 1 output neuron. The 100% prediction rate of the training data suggests overfitting, but the 74% of the testing set makes that seem unlikely. The TSS for the training set was .25, below the usual TSS for the SVMs, likely because this

	Accuracy	True Skill Score	MSE
Training Set (198 images)	1.00	1.00	.0004
Testing Set (198 images)	.74	.25	.2054

Table 2: Neural Net Results

neural net has no concept of data over time.

6 Conclusion

The goal of the project was to predict flares 24 hours in advance. Our SVMs did so with around 70% accuracy and our Neural Net did it with 74% accuracy. This is better than the 61% accuracy Bobra achieved, but the results may be biased because of how we constructed our data set. The similarities in accuracy aren't reflected in the respective TSSes - SVMS TSS were higher than the ANN TSS by .2, likely because the SVMS incorporated temporal data and the ANN didn't.

7 Future Work

Future work includes working with multidimensional SVMs, experimenting with the type and settings of the ANN, experimenting with new flare thresholds, using new features like variability in the magnetic field, and collecting higher cadence data on regions without large flares.

8 Acknowledgements

Dr. Harry Warren of the Naval Research Laboratory helped significantly with the literature review and feature selection.

9 References

1. Bobra, M.G. & Couvidat, S. 2014, Astrophysics Journal , TBD, TBD