

Advancing Solar Flare Detection: Integrating LSTM and Convolutional Layers in Neural Networks

Melvin Adkins, Fehmi Neffati, Nikolas Santamaria

Data Collection

- GOES-R Space Weather L2 Data
 - XRS 1-Minute Averages: 16, 17, 18
 - XRS Flare Summary: 16, 17, 18
- Combine the 1-Minute Averages with respective Summaries
- Combine summarized 1-Minute Averages from 3 satellites.

Binary Labeling:

EVENT_PEAK

1

Everything else:

0

Data Preprocessing

Data Structuring:

Wide format

Flare to Non flare Proportion control

Example window sizes: 20, 30, 50, 70, 80, 90, 100, 120, 150

Example Proportions: 0 (all non-flares) 1 2 3 4 and 5

Model Architecture

Bidirectional LSTM layer:

- Captures information from past/future flux measurements, makes predictions at each step.
 - Flux sequences processed in forward/backward direction simultaneously.
 - Outputs merged as tensor, form of feature extraction
 - Useful layer for sequential/time series data.

Convolutional 1D layer and Max Pooling:

- Used when spatial information is sequential
- Filters and stride are specified to produce feature maps capturing different patterns in the data.
- Max Pooling down samples the feature maps, reducing model complexity

Dropout:

- Regularization technique to improve model generalization
- Prevents memorization
- Enables model to learn new features

Flatten and Dense layers:

- Final output is a single value for Binary Classification, so Flatten and Dense are needed
- Flatten Layers transforms convolutional multidimensional output into one-dimensional output
- Final Dense layer with 1 neuron and sigmoid activation function for Binary Classification

Model Performance

- **Accuracy**

- $(TP + TN) / (TP + TN + FP + FN)$
- This metric tells us how well the model predicts overall, but is highly influenced by imbalanced data

- **Precision**

- $TP / (TP + FP)$
- *High Precision*: When the model predicts a solar flare, it is likely to be a solar flare (low FP rate)
- *Low Precision*: Indicates that the model is possibly finding undiscovered flares.

- **Recall**

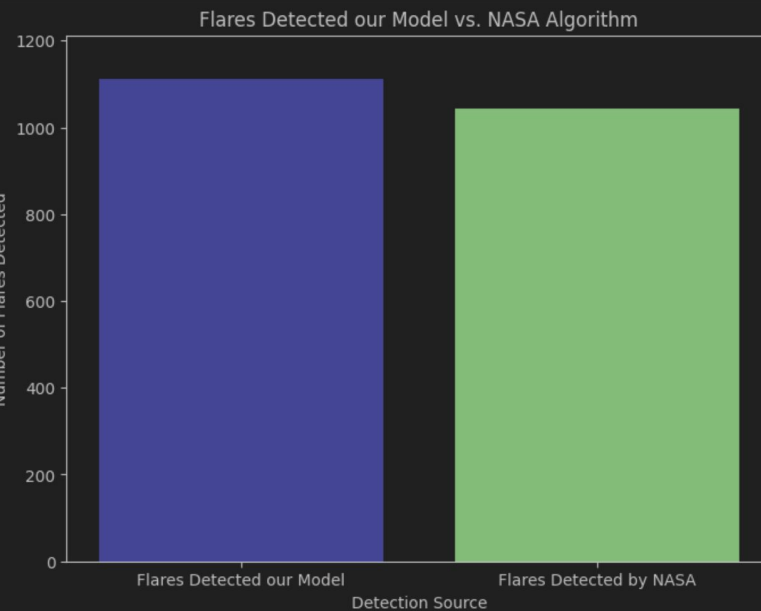
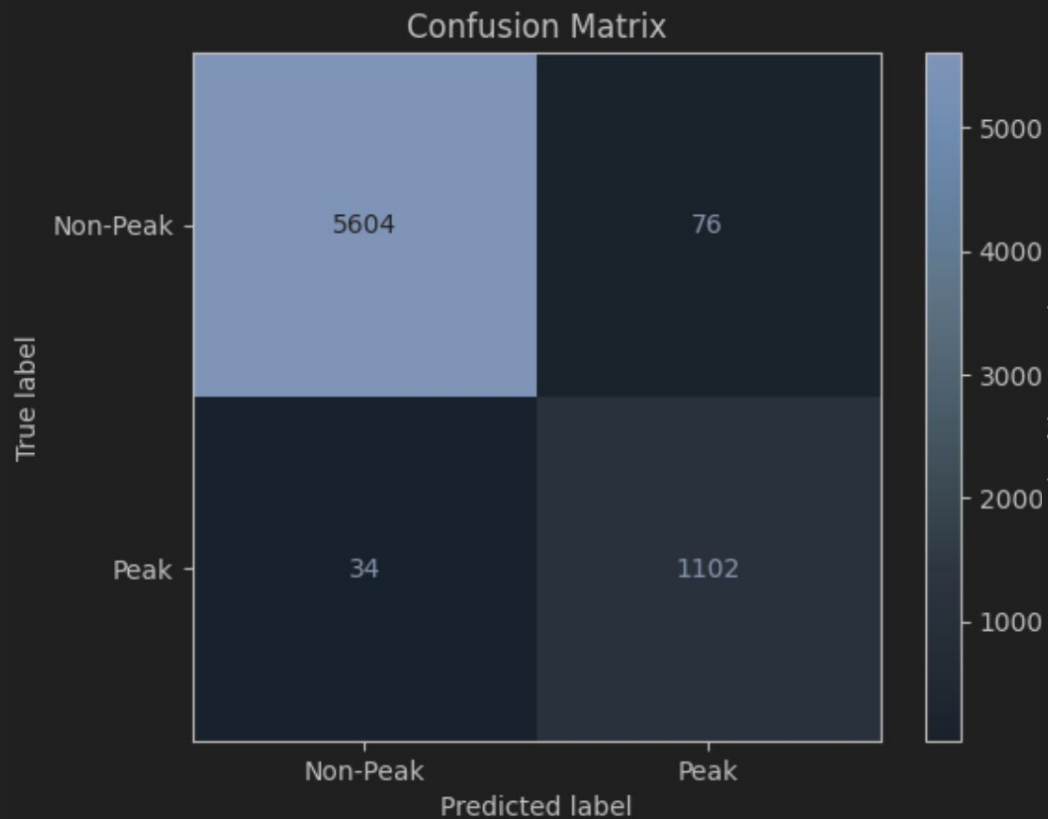
- $TP / (TP + FN)$
- *High Recall*: Model is predicting most of the solar flares in the data (low FN rate ~ model is not incorrectly rejecting possible flares)

Classification Report

Window Size: 70 Proportion: 5:1

Ratio 5:1

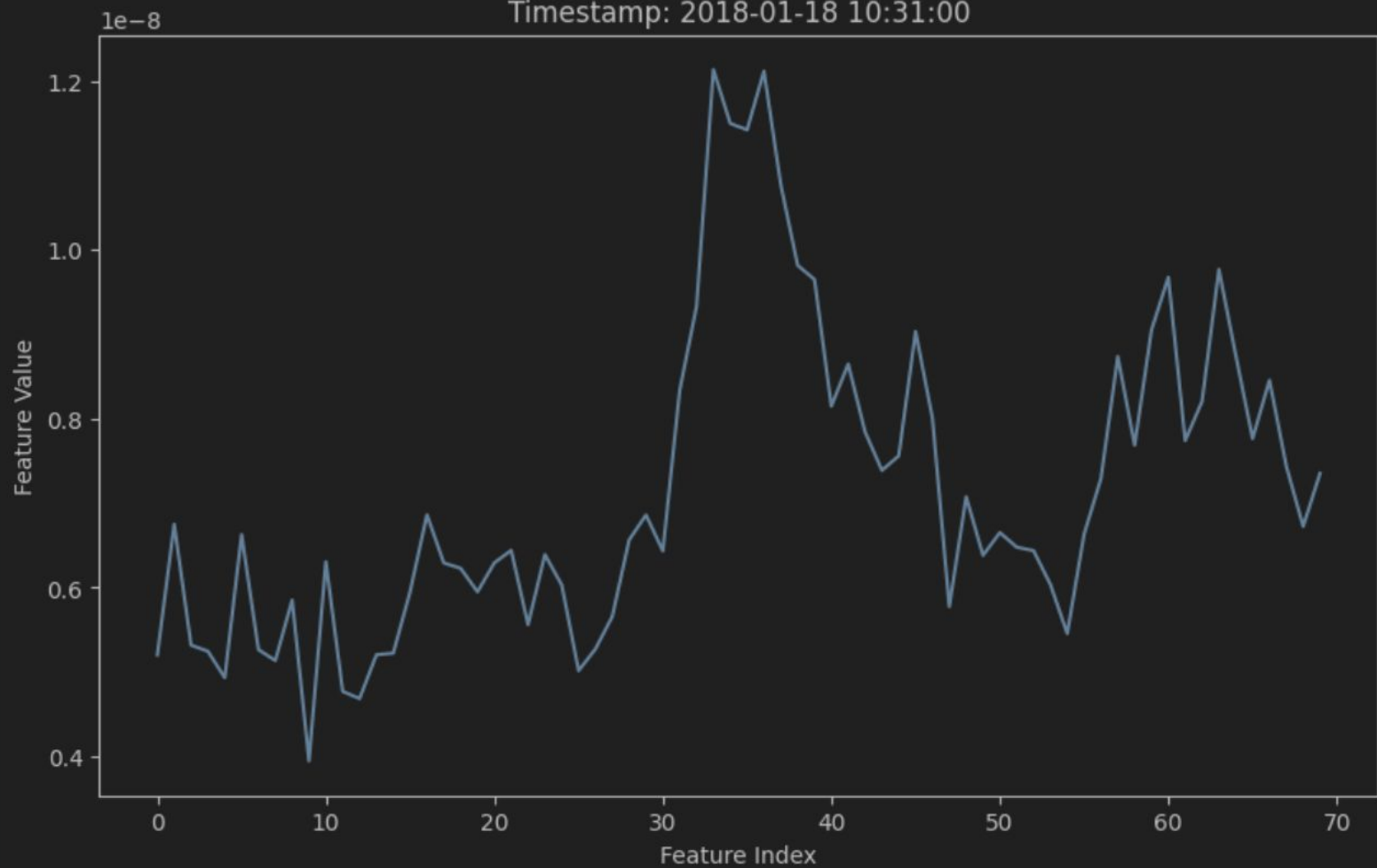
	precision	recall	f1-score	support
0	1.00	0.98	0.99	5680
1	0.92	0.98	0.95	1136
accuracy			0.98	6816
macro avg	0.96	0.98	0.97	6816
weighted avg	0.98	0.98	0.98	6816



Predictions

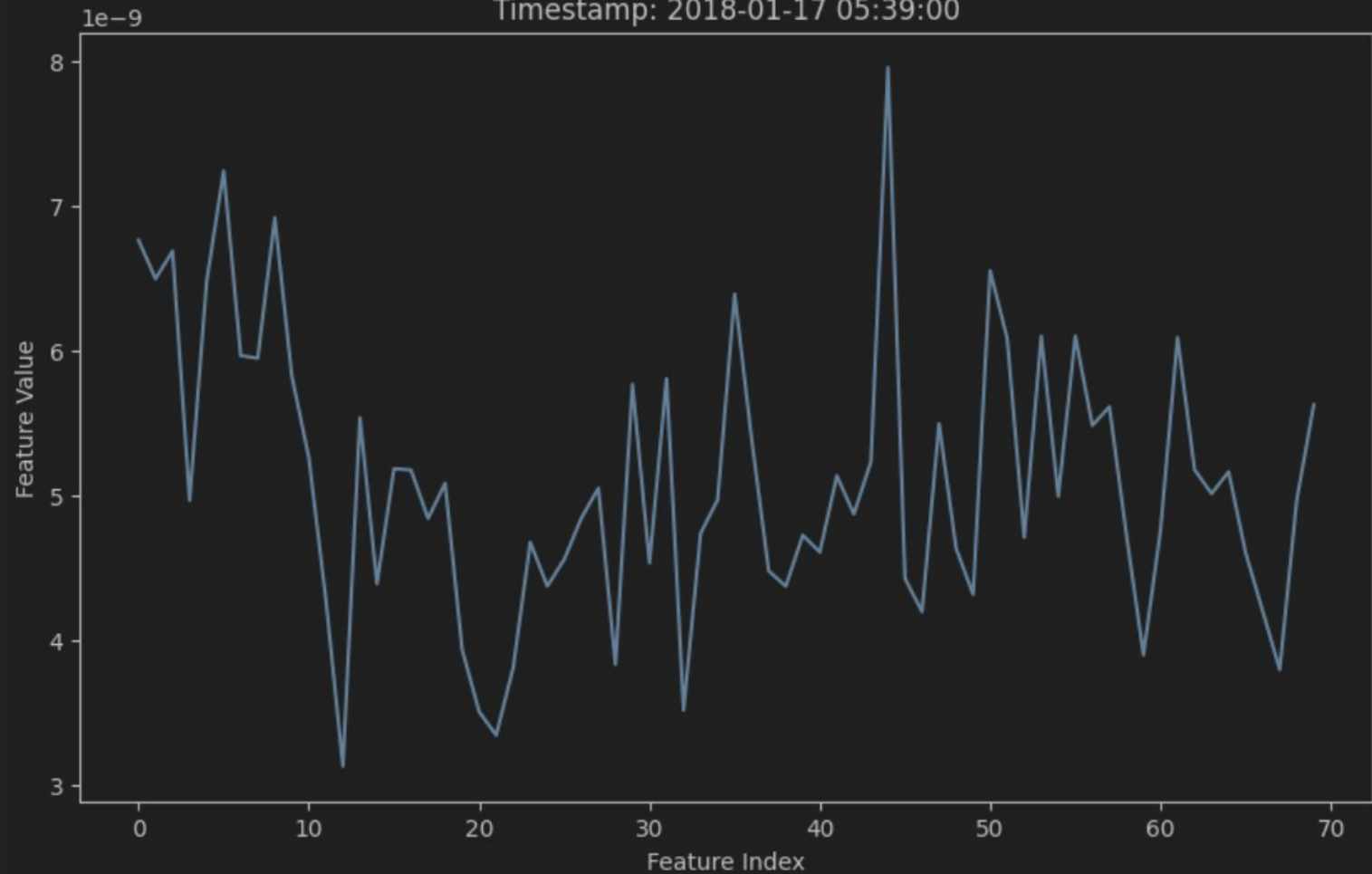
Ones that look nice

Timestamp: 2018-01-18 10:31:00

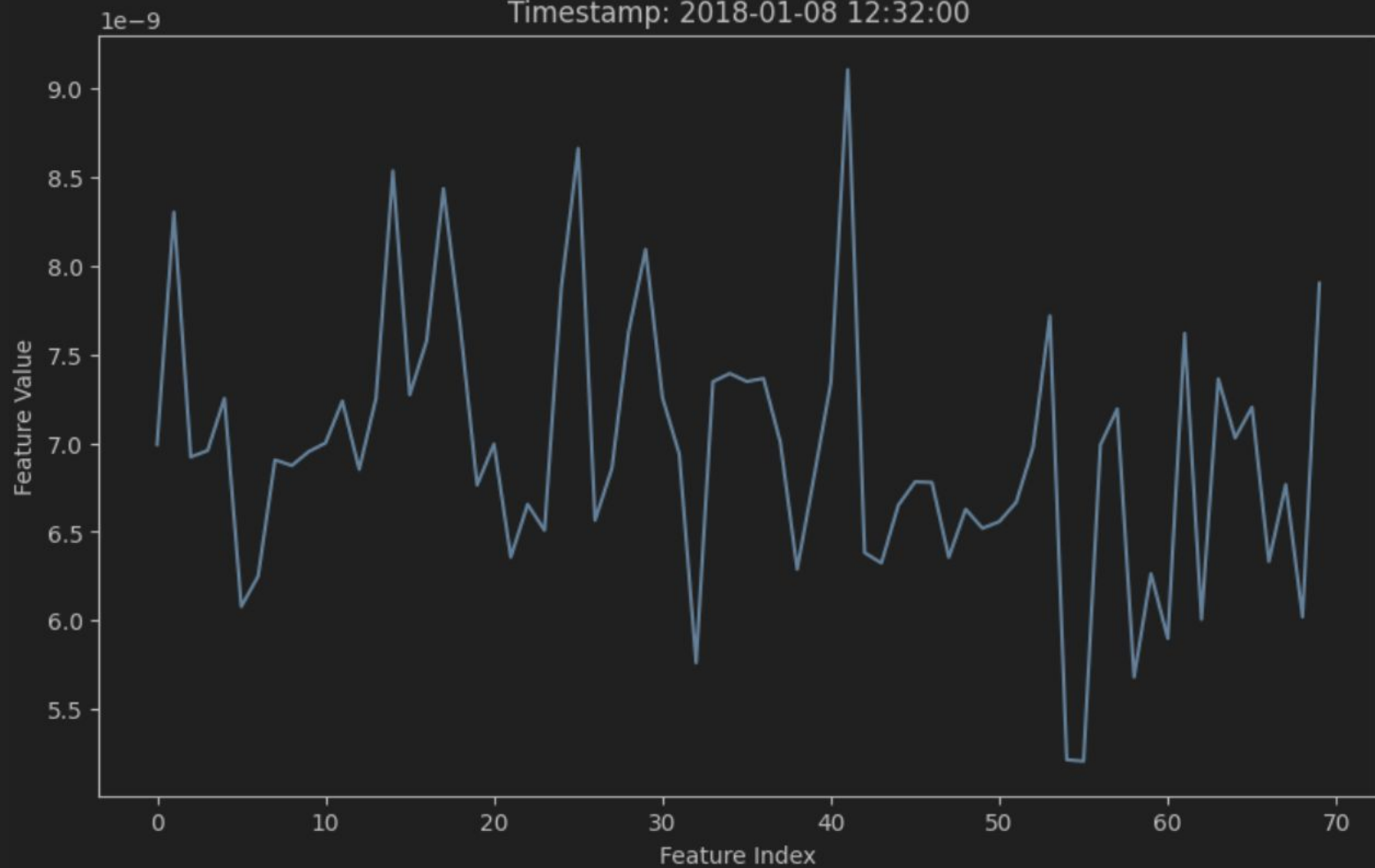


Ones that are controversial

Timestamp: 2018-01-17 05:39:00



Timestamp: 2018-01-08 12:32:00



Questions?