

Large-Scale Numerical Analysis for the Prediction of Flares using Support Vector Machines

R. Qahwaji, T. Colak and M. Al-Omari
University of Bradford

In this work a computer platform that uses support vector machines for the large scale analysis of sunspots and flares data is introduced. This work represents the ongoing improvements to our recent work reported in (Qahwaji and Colak 2007, Solar Physics). We aim to design more reliable prediction models with better generalization capabilities by considering additional sunspots features and more flare classes (i.e. C-class flares are considered in this work in addition to the X and M-classes).

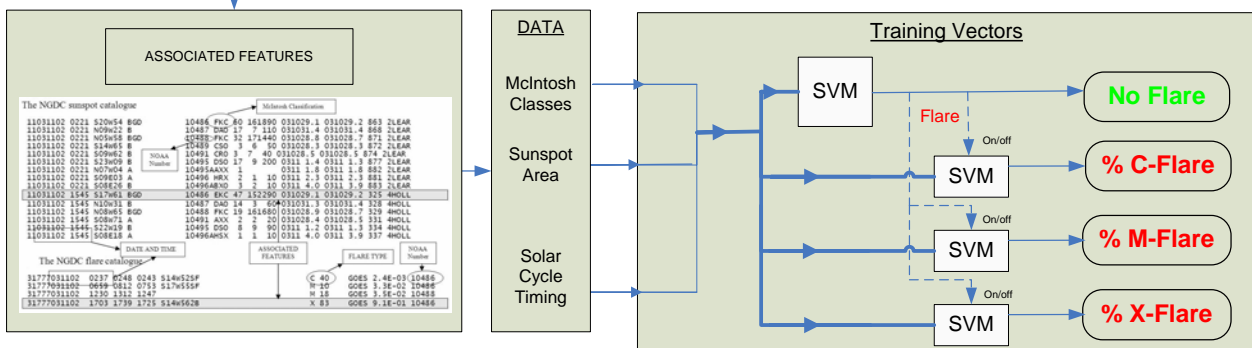
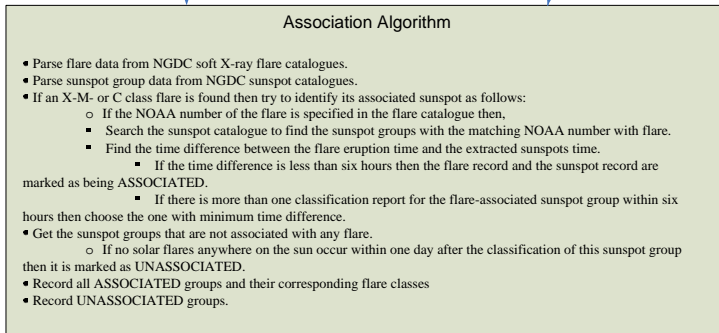
NGDC Sunspots Catalogue	
11031103 1516 209617 A	10486 AXZ 2 2 10 0311 2.4 0311 2.3 338 480LL
11031103 1516 209618 A	10486 AXZ 2 2 10 0311 2.9 0311 2.9 337 300LL
11031104 0347 218075 B	10486 EKC 22 141200 031029.4 031029.2 863 12LAR
11031104 0347 218076 B	10487 CSD 3 8 30 03101.4 03101.4 868 22AR
11031104 0347 218082 B	10488 FPC 11 14 520 031029.3 031028.7 871 12LAR
11031104 0347 222615 B	10489 CSD 3 8 30 0311 1.3 0311 1.3 877 22AR
11031104 0347 209611 A	10490 AXZ 2 2 10 0311 1.7 0311 1.8 886 22AR
11031104 0347 209622 A	10486 AXZ 1 1 10 0311 2.4 0311 2.3 885 22AR
11031104 0347 218077 A	10490 AXZ 1 1 10 0311 4.3 0311 4.1 888 22AR
11031104 0347 212011 B	10486MDD 2 3 10 0311 4.1 0311 4.1 887 22AR
11031104 0312 218077 A	10488 MXX 5 8 120 031029.3 031029.2 264 20YTF
11031104 0312 209622 A	10487 MXX 3 3 30 03101.4 03101.4 237 20YTF
11031104 0312 223619 A	10488 DDO 4 10 170 0311 3.5 0311 3.2 244 20YTF
11031104 1528 209646 B	10488 DDO 12 13 140 031029.4 031029.2 282 400LL
11031104 1528 219658 B	10487 CSD 2 3 30 03101.3 03101.4 328 400LL
11031104 1528 222641 B	10489 CSD 10 10 90 0311 1.3 0311 1.3 334 400LL
11031104 1528 219642 B	10490 AXZ 3 8 10 0311 1.5 0311 1.8 340 400LL
11031104 1528 219630 B	10490 AXZ 1 1 10 0311 6.9 0311 6.9 339 400LL
11031105 0357 219641 A	10487 AXZ 7 3 20 03101.4 03101.4 868 42AR
11031105 0357 212648 B	10490 AXZ 2 7 100 0311 1.3 0311 1.3 877 42AR

NGDC Flares Catalogue	
31777031103 0109 0145 0130 310W832B	X 27 GOES 3.6E-01 10488
31777031103 0943 1019 0955 308W772F	X 39 GOES 5.6E-01 10488
31777031103 1526 1543 1512 315W975F	M 39 GOES 2.5E-02 10486
31777031103 1951 1957 1954	C 44 GOES 1.5E-03
31777031103 2031 2041 2037	C 54 GOES 2.9E-03
31777031103 2238 2240 2236	C 31 GOES 2.7E-03
31777031104 0404 0419 0411	C 50 GOES 4.1E-03
31777031104 0543 0507 0556	M 26 GOES 2.5E-02
31777031104 0940 0950 0943	C 28 GOES 1.6E-03 10486
31777031104 1011 1033 1022	M 30 GOES 2.7E-02 10488
31777031104 1115 1125 1119	C 57 GOES 2.6E-03 10486
31777031104 1343 1401 1349	M 11 GOES 9.3E-03 10486
31777031104 1929 2006 1950 319W833B	X 289 GOES 2.3E-02 10486
31777031105 0237 0245 0241 319W995F	M 16 GOES 5.2E-03 10486
31777031105 0754 0803 0759	C 47 GOES 2.1E-03
31777031105 1046 1056 1052 316W995F	M 53 GOES 1.7E-02 10486
31777031105 1629 1709 1648	C 19 GOES 3.8E-03
31777031106 1105 1138 1110	B 52 GOES 3.8E-04

Learning Mode

A C++ platform was created correlate the NGDC sunspots and flares catalogues and to associate all the X, M and C class flares with their corresponding sunspot groups. The correspondence was determined based on the location (i.e., same NOAA number) and timing information (maximum six hours time difference between the erupting flare and the time for which the classification of the associated sunspot group is reported). After finding all the associations, a numerical dataset is created for the machine learning algorithms.

	AVERAGE AFTER 20 EXPERIMENTS
FLARING PREDICTION	78.0%
C- FLARE PREDICTION	79.5%
M FLARE PREDICTION	73.8%
X FLARE PREDICTION	80.8%



Real-Time Mode

