CSE 5523: Machine Learning and Statistical Pattern Recognition

Linear Models and SGD on Brain Image Data

CHANDRASEKAR SWAMINATHAN (.42)

Logistic loss

Learning rate	Lambda	Accuracy	
0.1	1	0	Overflow
0.1	0.3	0	
0.1	0.1	0	
0.01	1	0	
0.01	0.3	0	
0.01	0.1	0	
0.001	1	0.8	
0.001	**0.3**	**0.9**	**Best model**
0.001	0.1	0.9	
0.0001	1	0.9	
0.0001	0.3	0.9	
0.0001	0.1	0.9	

Accuracy (Logistic Loss): 0.794117647059

We can see that for most cases the accuracy is 0 due to overflow during the computation. The accuracy is the same for the remaining choices of parameters except for the case of 0.001 and 1. Learning rate of **0.001** was chosen over 0.0001 because it takes lesser time to train the dataset using the former. Regularization parameter of **0.3** was chosen over 0.1 to make sure that the weights are not very high. The accuracy of these parameters over the test data set is approximately **79%**.

The loss numbers at the end of each iteration for the best model is available in the **logistic_loss_numbers.txt** file included in the submission. The accuracy numbers shown in the above table are available in the **logistic_loss_accuracy_numbers.txt**

Hinge Loss

Learning rate	Lambda	Accuracy
0.1	1	0.7
0.1	0.3	0.55
0.1	0.1	0.7
0.01	1	0.75
0.01	0.3	0.9
0.01	0.1	0.8
0.001	1	0.9
0.001	0.3	0.9
0.001	0.1	0.9
0.0001	**1**	**0.95**
0.0001	0.3	0.9
0.0001	0.1	0.9

Accuracy (Hinge Loss): 0.823529411765

We can see how the accuracy value has changed for different values of learning rate and lambda in the above table. There is only one combination of parameters which give the highest value of accuracy **0.95**. The learning rate of **0.0001** and regularization parameter of **1** was chosen as the best model and the accuracy of this model on the test data is **82%**.

We can see that using the hinge loss results in a better accuracy over logistic loss.

The loss numbers at the end of each iteration for the best model is available in the **hinge_loss_numbers.txt** file included in the submission. The accuracy numbers in the above table are available in the **hinge_loss_accuracy_numbers.txt** file.

ROI Analysis

Region Name	Number of Voxels	No of positive weights	No of negative weights	% of positive weights	Sum of all weights
LFEF	109	64	45	**0.587**	0.00169861416687
LT	305	176	129	**0.577**	0.00463728039273
LIT	287	161	126	**0.561**	0.00256354864715
RIT	278	153	125	**0.55**	0.00249688092389
RFEF	68	37	31	**0.544**	0.00155600785342
LTRIA	113	60	53	**0.531**	0.00034720130957
LIFG	282	144	138	**0.511**	-0.00035542635673
RT	284	142	142	0.5	0.00043004701056
LSGA	6	3	3	0.5	0.00012904817015
LOPER	169	84	85	0.497	-0.0007026276663
CALC	318	153	165	0.481	-0.00063598064636
LIPS	236	113	123	0.479	-0.00128117593341
LIPL	134	64	70	0.478	-0.00122196852899
LPPREC	153	70	83	0.458	-0.00127799215382
ROPER	181	82	99	0.453	-0.00368530921501
RIPS	166	75	91	0.452	-0.00217111610541
RPPREC	144	64	80	0.444	-0.00171541603951
SMA	215	93	122	0.433	-0.00210008932206
LSPL	308	132	176	0.429	-0.00551763984187
RSPL	252	107	145	0.425	-0.00514572679539
RIPL	92	39	53	0.424	-0.00353780964927
RSGA	34	13	21	0.382	-0.00135195854489
RTRIA	57	21	36	0.368	-0.00169931465748
LDLPFC	440	155	285	0.352	-0.0146420816174
RDLPFC	349	121	228	0.347	-0.0153424671666

The **weight_vector** with 258391 values is stored in the **final_W.txt** file. The first 258390 values in the vector are reshaped to a 55x4698 matrix, where 55 is the maximum number of images that are present in one trial and 4698 is the number of voxels in each of those 55 images. The average of 55 rows are calculated to get a single weight vector. We then group the weight vector values by the regions of the brain and determine which regions have more than 50% of the weight vector values greater than zero. These regions are highlighted in darker shade of green in the above. The regions **LFEF , LT , LIT , RIT , RFEF , LTRIA , LIFG** have more than 50% voxels with positive weights. A positive weight to a voxel indicates that it contributes more towards the prediction that the image shown to the user contains a picture.