Programming Assignment3 Report

1. USPS dataset (Logistic regression):

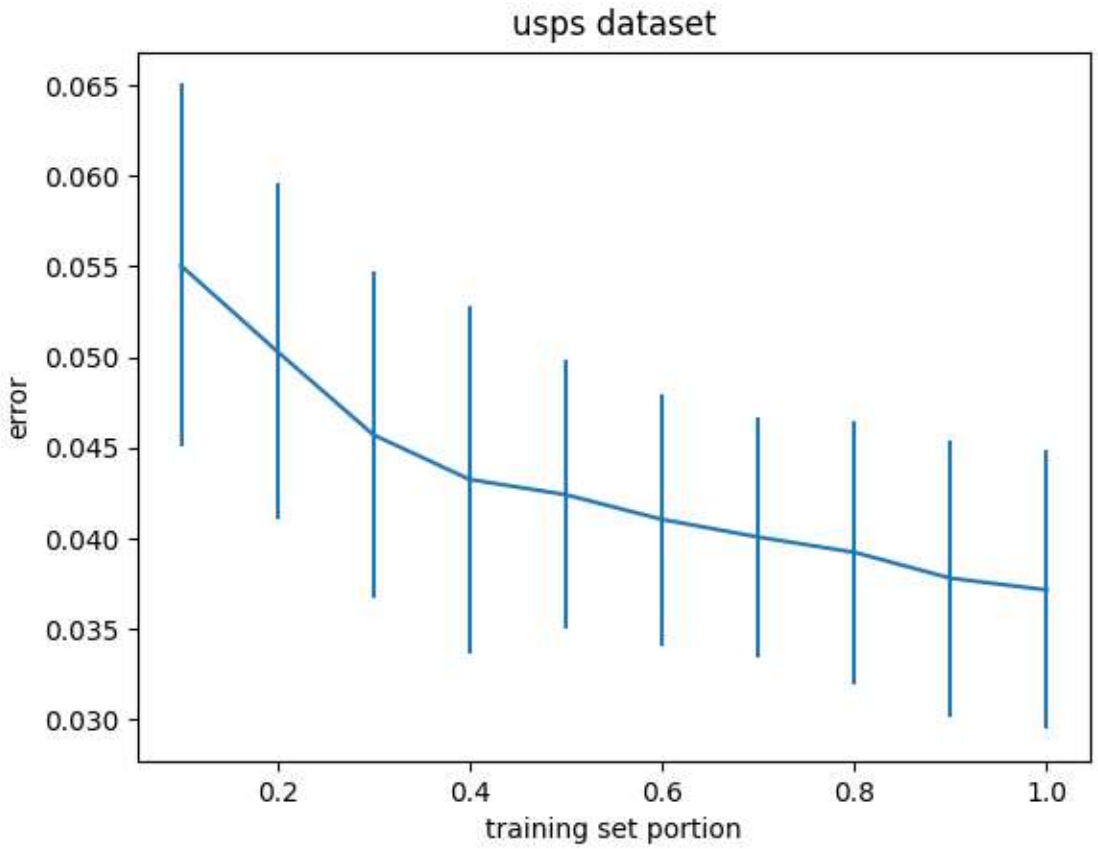

Figure 1: average of errors per training set portion for usps dataset.

Dataset portion	number of iterations average	runtime average
0.1	4.166666666666667	0.05678548812866211
0.2	5.0	0.11580134232838948
0.3	5.0	0.1937417507171631
0.4	5.0	0.2931209405263265
0.5	5.166666666666667	0.4509857177734375
0.6	5.833333333333333	0.6807629823684692
0.7	5.966666666666667	0.8579460938771566
0.8	6.0	1.122094782193502
0.9	6.0	1.424791653951009
1	6.0	1.682824937502543

2. A dataset (Logistic regression):

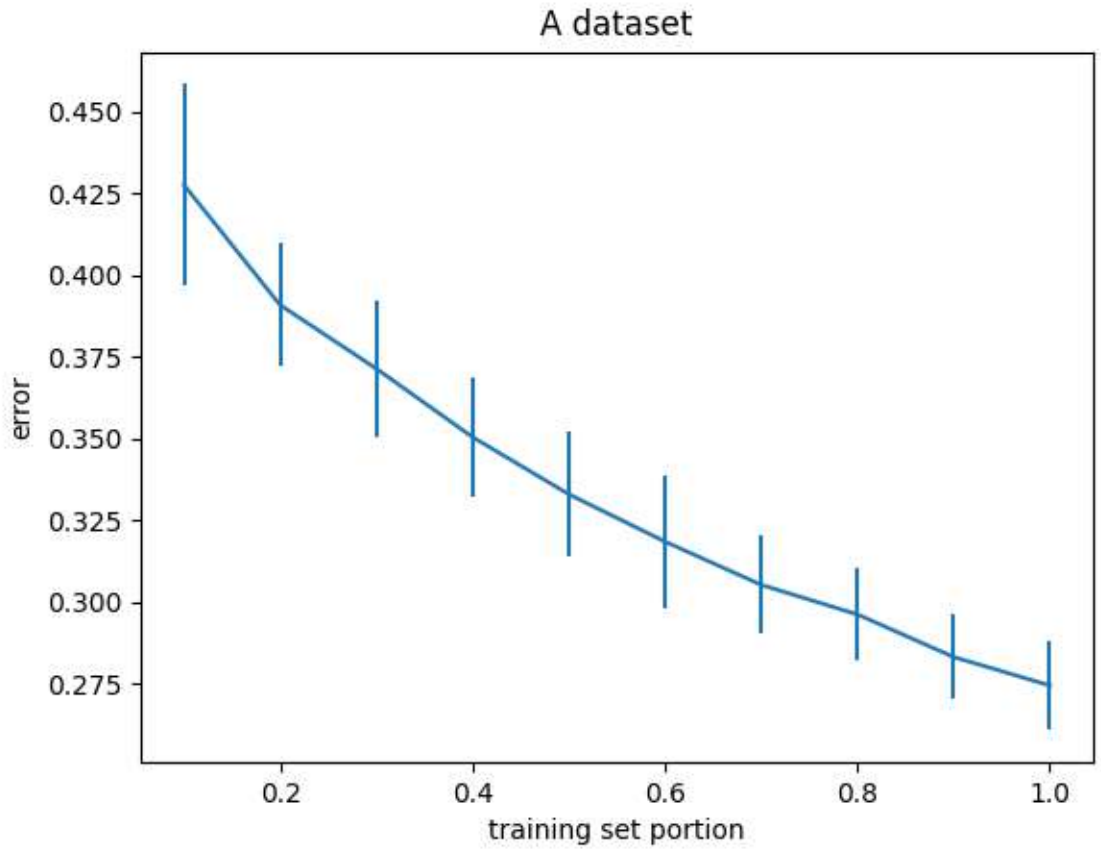

Figure 2: average of errors per training set portion for A dataset

Dataset portion	number of iterations average	runtime average
0.1	2.0	0.04541617234547933
0.2	2.0	0.04541617234547933
0.3	2.0	0.09040830930074056
0.4	2.0	0.1533329725265503
0.5	2.0	0.23295534451802571
0.6	2.0	0.33303030331929523
0.7	2.0	0.45542680422465004
0.8	2.0	0.5855802218119304
0.9	2.0	0.7262588262557983
1	2.0	0.928596830368042

3. AP dataset (Poisson regression):

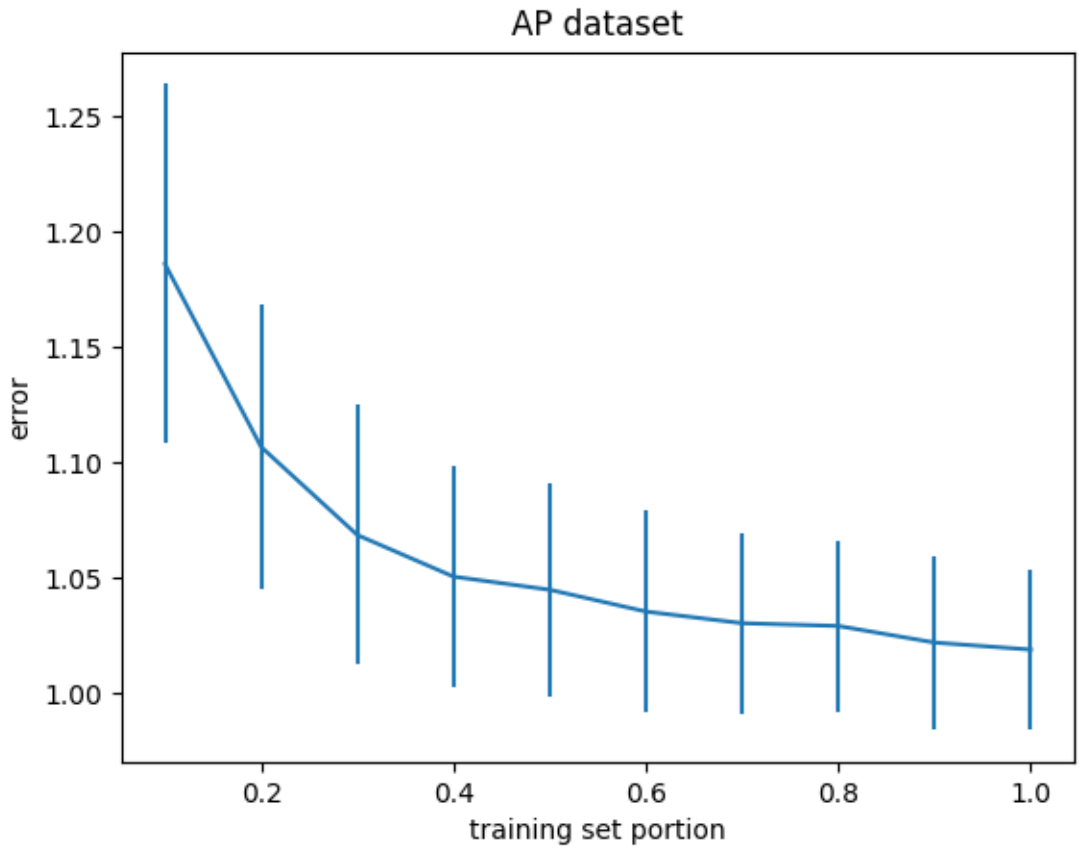

Figure 3: average of errors per training set portion for AP dataset

Dataset portion	number of iterations average	runtime average
0.1	6.433333333333334	0.049626795450846355
0.2	6.633333333333334	0.1471654176712036
0.3	6.7	0.2899428447087606
0.4	6.633333333333334	0.5114253044128418
0.5	6.633333333333334	0.690285587310791
0.6	6.633333333333334	0.9788608551025391
0.7	6.633333333333334	1.29704798857371
0.8	6.566666666666666	1.6291708072026572
0.9	6.533333333333333	1.9432049512863159
1	6.433333333333334	2.4554223934809367

4. AO dataset (Ordinal regression):

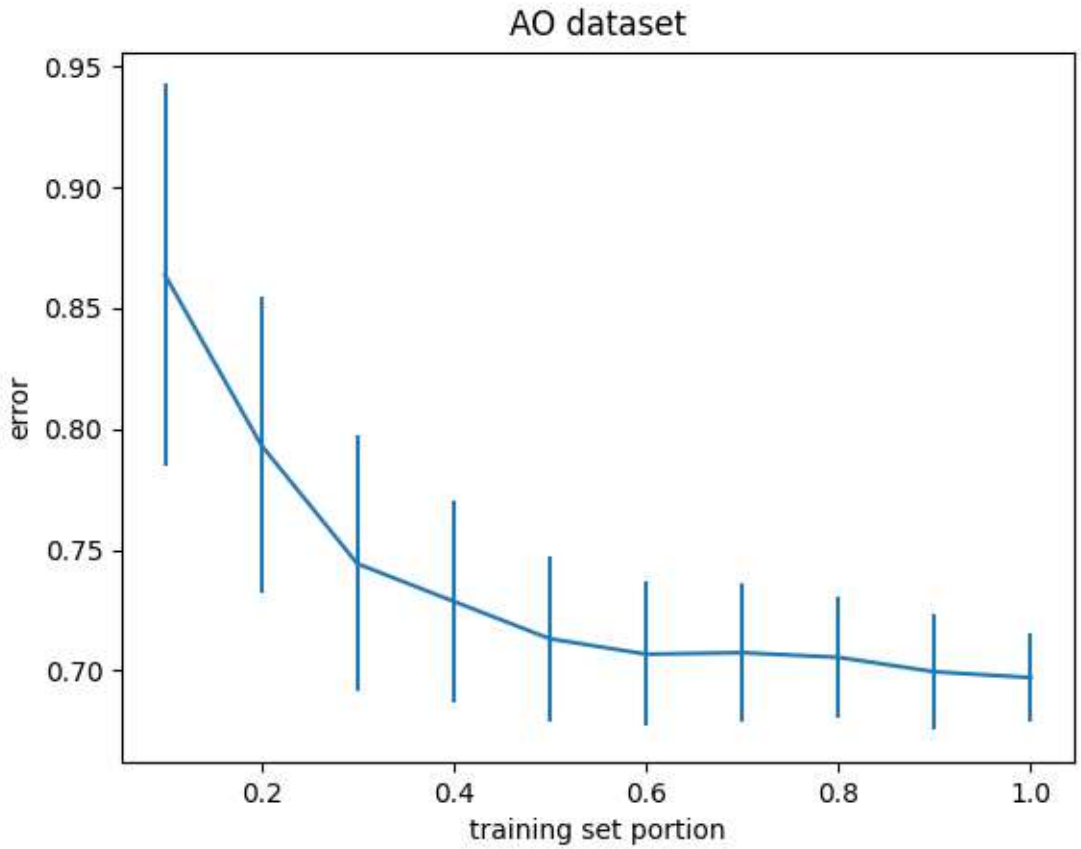

Figure 4: average of errors per training set portion for AO dataset

Dataset portion	number of iterations average	runtime average
0.1	2.1	0.049626795450846355
0.2	2.033333333333333	0.1471654176712036
0.3	2.0	0.2899428447087606
0.4	2.0	0.5114253044128418
0.5	2.0	0.690285587310791
0.6	2.0	0.9788608551025391
0.7	2.0	1.29704798857371
0.8	2.0	1.6291708072026572
0.9	2.0	1.9432049512863159
1	2.0	2.4554223934809367

provide a short discussion of the results: by increasing the training set we get more data to learn from so the error rate will decrease. Also, the running time will increase because we have more data and the computational cost will be more so it needs more time to calculate the w map.

Are the learning curves as expected? Yes, because by increasing the training set size we have more data to learn from so our result could be more accurate

how does learning time vary across datasets for classification? and across the likelihood models?

Ordinal regression likelihood takes more time because it has more computational cost comparing to others. Then we can say that after that Poisson regression is time consuming and at the end logistic regression.

For Logistic regression we have 2 datasets and we see that the running time for usps dataset is more than A dataset and it is because of that usps has more features so it takes more time to calculate w map comparing to A dataset.

what are the main costs affecting these (time per iteration, number of iterations)?

As we see in the tables number of iterations does not change that much but the running time increased respect to the training set size so time per iteration will change respect to the size of the training dataset.

Extra credit:

Implement cross validation. It is slower than evidence approximation but based on the results of previous assignment it is more accurate for small datasets.

I just check alphas from 1 to 10.

usps

error for alpha from 1 to 10: [0.848127613021332, 0.8477152347564697, 0.7367217063903808, 0.7321239709854126, 0.7256992196573795, 0.7404907798304141, 0.7388916587366642, 0.7309962605967105, 0.740990505172211, 0.7578519057301641]

best error is: 0.7256992196573795

best alpha is: 5

~~~~~~
~~~~~~

A

error for alpha from 1 to 10: [0.5236170523202242, 0.4440978836657396, 0.4647356129404324, 0.48153664140558955, 0.3207183691992689, 0.29922616339441555, 0.2995341948608854, 0.30193768188134945, 0.3053160332921726, 0.31320475037418194]

best error is: 0.29922616339441555

best alpha is: 6

~~~~~~

AP

error for alpha from 1 to 10: [0.9943084898279674, 1.0172196384686143, 1.0019056555050523, 1.048877596499315, 1.0844082405318074, 1.0388829704540878, 1.085309605811959, 1.0490374960116486, 1.0310986490391973, 1.0574857206486945]

best error is: 0.9943084898279674

best alpha is: 1

~~~~~~

AO

error for alpha from 1 to 10: [0.516790295358914, 0.5174875875017536, 0.4356811530554472, 0.37991827672986844, 0.39371184057264186, 0.3606320971873269, 0.3678947801020608, 0.3660946959879861, 0.3762492877333911, 0.35705911935265383]

best error is: 0.35705911935265383

best alpha is: 10

~~~~~~
~~~~~~