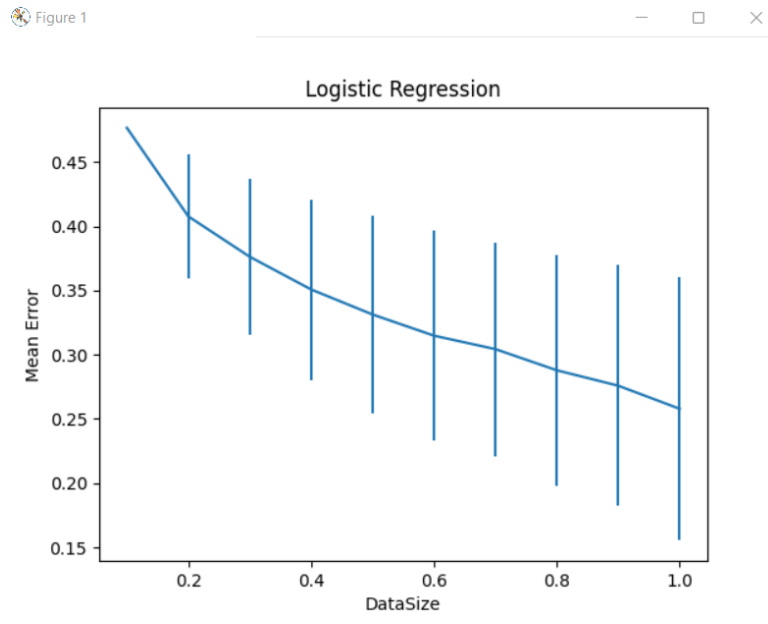


ML PP3 Report

1. Logistic Regression

Dataset: A.csv

Plot:



Runtime and Iterations:

```
"C:\Users\Abhishek Sharma\PycharmProjects\GLM\venv\Scripts\python.exe" "C:/Users/Abhishek Sharma/PycharmProjects/GLM/main.py"  
Average No of Iterations 85  
Runtime: 7.753549337387085 seconds
```

Discussion of the result:

It can be seen from the plot that as the train dataset size increases, the overall error rate decreases. Each fragment of the dataset takes about 85 iterations before W converges and the avg runtime for the whole process is around 7.7 seconds.

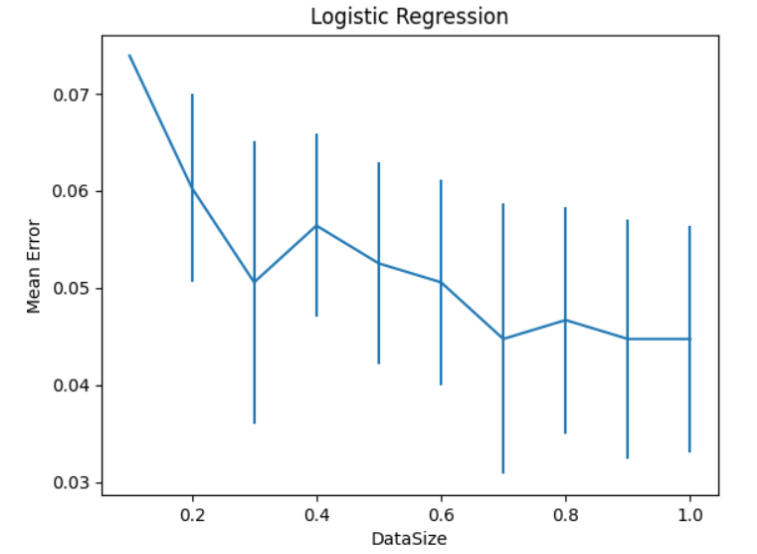
The highest mean of the error goes upto 0.45 and the error goes to a low of around 0.28 at full fraction of the train Dataset.

P.S: I acknowledge the unusually high standard deviation but I couldn't resolve the issue as I ran out of the time. I acknowledge that the std should not be as high as it shows in the plots.

Dataset: USPS

Plot:

Figure 1



Runtime and Iterations:

```
"C:\Users\Abhishek Sharma\PycharmProjects\GLM\venv\Scripts\python.exe" "C:/Users/Abhishek Sharma/PycharmProjects/GLM/main.py"
Average No of Iterations100
Runtime: 16.506453275680542 seconds
```

Discussion of the result:

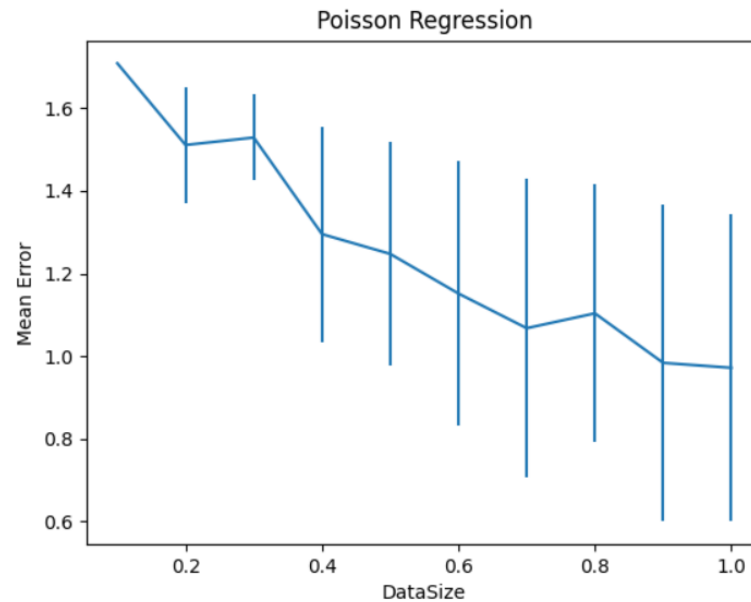
It can be seen from the plot that as the train dataset size increases, the overall error rate decreases. Each fragment of the dataset takes about 100 iterations before W converges and the avg runtime for the whole process is around 16.5 seconds.

The highest mean of the error goes up to 0.07 and the error goes to a low of around 0.05 at full fraction of the train Dataset.

P.S: I acknowledge the unusually high standard deviation but I couldn't resolve the issue as I ran out of the time. I acknowledge that the std should not be as high as it shows in the plots.

2. Poisson Regression

Plot:



Runtime and Iterations:

```
"C:\Users\Abhishek Sharma\PycharmProjects\GLM\venv\Scripts\python.exe" "C:/Users/Abhishek Sharma/PycharmProjects/GLM/main.py"
Runtime: 4.8341827392578125 seconds
Average No of Iterations 96
```

Discussion of the result:

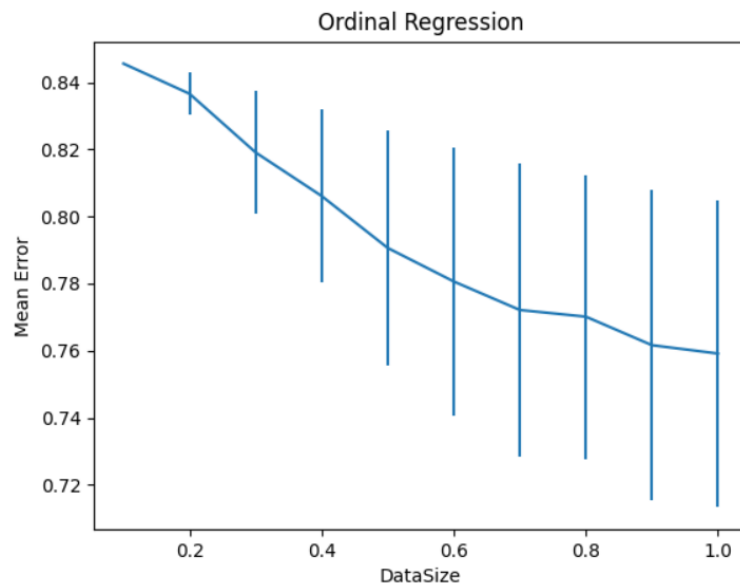
It can be seen from the plot that as the train dataset size increases, the overall error rate decreases. Each fragment of the dataset takes about 96 iterations before W converges and the avg runtime for the whole process is around 4.9 seconds.

The highest mean of the error goes up to 1.7 and the error goes to a low of around 1 at full fraction of the train Dataset.

P.S: I acknowledge the unusually high standard deviation but I couldn't resolve the issue as I ran out of the time. I acknowledge that the std should not be as high as it shows in the plots.

3. Ordinal Regression:

Plot:



Runtime and Iterations:

```
"C:\Users\Abhishek Sharma\PycharmProjects\GLM\venv\Scripts\python.exe" "C:/Users/Abhishek Sharma/PycharmProjects/GLM/main.py"  
Runtime: 57.68153405189514 seconds  
Average No of Iterations 99
```

Discussion of the result:

It can be seen from the plot that as the train dataset size increases, the overall error rate decreases. Each fragment of the dataset takes about 99 iterations before W converges and the avg runtime for the whole process is around 57.6 seconds.

The highest mean of the error goes up to 0.84 and the error goes to a low of around 0.76 at full fraction of the train Dataset.

P.S: I acknowledge the unusually high standard deviation but I couldn't resolve the issue as I ran out of the time. I acknowledge that the std should not be as high as it shows in the plots.

Q. Are the learning curves as expected?

The error rate is expected to go down as the learning process progresses further. Also, a bigger train data size is expected to enhance the learning which would result in lower error rates. As is evident from the plots, the learning curves follow the expected behavior.

Q. How does learning time vary across datasets for classification?

The time taken by the USPS dataset of the Logistic Likelihood model was of about 16.7 seconds, whereas the same model takes only about 7 seconds on the A.csv dataset. This shows that the learning time depends on the size of the datasets.

The dataset used in the Poisson takes the least amount of time out of all the other datasets.

Q. How does learning time vary across the likelihood models?

The learning time varies widely across different likelihood models. For example, the shortest time was taken by the Poisson Likelihood model of about 4.9 seconds, whereas the highest time was taken by the Ordinal Likelihood Model, around 57.6 seconds.

The logistic Likelihood model takes around 7.6 seconds for the A.csv dataset and around 16.7 seconds for the USPS dataset.

Since the steps involved in the Ordinal Model are more complex than the other two models, the effect of this can be easily seen in the runtime of this model, which is the highest of all.

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

The main costs that affect the number of iterations can be thought to be as follows:

1. The complexity of the Model:

It can be observed that the more complexly designed model results in higher amount of time taken as compared to those models that are simpler in comparison.

For eg: The Ordinal model involves some complex steps in creating the y matrix and in calculating the d and R matrices and these complex steps result in higher execution time.

2. The learning labels provided:

The learning labels provided with the datasets can also affect the number of iterations and the time per iteration. This is because the labels are used in the calculation of the d and R matrices which affect the convergence of W .