Parameter Estimations -- Multivariate

thetaPosTrue
```
 0.4045584   0.51566952 0.55698006 0.69088319 0.90598291 0.25925926
 0.8960114   0.61538462 0.502849   0.82193732 0.36752137 0.8019943
 0.44871795 0.53703704 0.26495726 0.74216524 0.94159544 0.62535613
 0.90883191 0.59259259 0.54985755 0.48148148 0.58404558 0.38034188
 0.37321937 0.71225071 0.53846154 0.38176638 0.47578348 0.76068376
 0.60968661 0.48575499 0.39316239 0.33618234 0.52136752 0.4045584
 0.31054131 0.3974359  0.43162393 0.34045584 0.33333333 0.46153846
 0.37179487 0.35754986 0.42592593 0.46296296 0.45299145 0.33333333
 0.46438746 0.38034188 0.53846154 0.40598291 0.58974359 0.6039886
 0.41452991 0.37606838 0.44586895 0.55698006 0.5954416  0.53276353
 0.3988604  0.39316239 0.73504274 0.57977208 0.31196581 0.74358974
 0.35185185 0.52564103 0.32763533 0.56695157 0.35327635 0.55840456
 0.39316239 0.31908832 0.5042735  0.35470085 0.43589744 0.33760684
 0.45156695 0.33760684 0.37891738 0.34045584 0.24786325 0.33048433
 0.44017094 0.55698006 0.34188034 0.42165242 0.39316239 0.37037037
 0.45726496 0.36609687 0.36039886 0.42877493 0.30911681 0.4045584
 0.40598291 0.35327635 0.38176638 0.52564103
```

thetaNegTrue
```
 0.51851852 0.52564103 0.55270655 0.72222222 0.87464387 0.29202279
 0.86324786 0.52706553 0.34188034 0.79344729 0.39458689 0.84615385
 0.33333333 0.41737892 0.5042735  0.7037037  0.95868946 0.66809117
 0.89316239 0.59259259 0.51994302 0.51994302 0.50712251 0.23931624
 0.36609687 0.71794872 0.52706553 0.36182336 0.55413105 0.80626781
 0.58262108 0.47863248 0.45014245 0.33190883 0.43874644 0.43162393
 0.33903134 0.38319088 0.42592593 0.31054131 0.35612536 0.46153846
 0.33475783 0.34900285 0.37464387 0.39316239 0.46438746 0.31054131
 0.48148148 0.34188034 0.58974359 0.39031339 0.62535613 0.6951567
 0.35042735 0.31054131 0.45868946 0.51139601 0.47863248 0.48575499
 0.4031339  0.38176638 0.6965812  0.52421652 0.33048433 0.76638177
 0.38176638 0.48148148 0.30769231 0.54985755 0.31339031 0.50712251
 0.36324786 0.34045584 0.36324786 0.27920228 0.39458689 0.31196581
 0.48433048 0.32051282 0.35470085 0.35754986 0.28062678 0.29202279
 0.35754986 0.50569801 0.29344729 0.41595442 0.33190883 0.35754986
 0.31908832 0.39173789 0.3960114  0.26923077 0.36467236 0.46581197
 0.39173789 0.3048433  0.34615385 0.37464387
```

Bernoulli Classification Accuracy

BNBC classification accuracy = 0.5966666666666667

Scikit-learn Multinomial Classification Accuracy
Sklearn MultinomialNB accuracy = 0.7416666666666667

Parameter Estimation -- Multinomial

thetaPos
```
 9.65210903e-06 7.92902883e-05 2.32008102e-05 2.57389574e-05
 1.09033084e-05 1.43351693e-05 1.63728368e-05 1.21545077e-05
 1.97332007e-05 2.79911162e-05 2.98142923e-05 2.46307523e-05
 2.09129029e-05 1.17112256e-04 7.84680716e-05 6.26314631e-05
 1.55541950e-04 3.03505206e-05 1.45496607e-05 1.73022992e-05
 1.42279237e-05 1.43673430e-04 1.66588252e-05 1.21187591e-05
 2.58819516e-05 1.20830106e-05 2.55959632e-05 3.41041186e-05
 3.22094453e-05 1.17255250e-05 1.02240859e-05 2.23070964e-05
 3.87514303e-05 5.04769554e-05 2.47737465e-05 1.49428947e-05
 1.69448136e-05 1.40134324e-05 3.18519598e-05 2.10201486e-05
 6.38826624e-05 1.99119435e-05 4.05388579e-05 1.95187094e-05
 2.57747060e-05 4.11108348e-05 1.92327210e-05 1.80172702e-05
 1.33342099e-05 2.89205786e-05 1.40491809e-05 9.68785759e-06
 3.09224975e-05 1.42636722e-05 2.34153015e-05 1.59438542e-05
 1.36916954e-05 2.42375182e-05 2.69544082e-05 1.50858889e-05
 1.41564266e-05 1.42636722e-05 2.93138126e-05 1.42636722e-05
 3.44616041e-05 1.19400164e-05 5.02624641e-05 1.98761949e-05
 1.44424150e-05 1.86249956e-05 1.96259550e-05 1.37989411e-05
 2.19853595e-05 2.52742262e-05 1.63370883e-05 1.88394869e-05
 1.50143918e-05 1.57293629e-05 1.14395366e-05 2.43447639e-05
 2.34867986e-05 1.57293629e-05 1.39776838e-05 1.34414555e-05
 1.28694787e-05 1.69090651e-05 2.38085356e-05 1.84462528e-05
 1.47284034e-05 1.43709179e-05 1.70878078e-05 1.12250453e-05
 2.14491312e-05 1.24762446e-05 2.13776341e-05 1.20472620e-05
 1.43351693e-05 1.24762446e-05 1.60868484e-05 2.02336804e-05
 9.97426712e-01
```
thetaNeg
```
 2.88693832e-05 7.99645676e-05 2.66548559e-05 2.85875342e-05
 1.32871638e-05 1.37300693e-05 1.99307457e-05 1.30455790e-05
 2.23063295e-05 3.78482847e-05 3.90562087e-05 2.30310839e-05
 2.89096473e-05 1.36052504e-04 1.02834595e-04 6.88516669e-05
 1.40602351e-04 2.13802545e-05 1.61459172e-05 1.81188597e-05
 1.48977291e-05 1.11531648e-04 1.71927847e-05 1.28845225e-05
```

```
2.39571590e-05  1.47769367e-05  2.42390079e-05  3.61571912e-05
2.76211950e-05  1.63875020e-05  1.44145595e-05  1.97696892e-05
4.89611854e-05  6.00338219e-05  2.93525527e-05  2.02931229e-05
1.40521823e-05  1.24013529e-05  3.41842487e-05  2.39571590e-05
6.25301981e-05  1.77967467e-05  4.05862458e-05  1.30455790e-05
3.04396843e-05  4.26397165e-05  2.33934611e-05  1.02270897e-05
1.53808987e-05  2.35142535e-05  1.03478821e-05  1.50185215e-05
3.29763247e-05  1.22402963e-05  2.26687067e-05  1.49782574e-05
1.79175391e-05  2.55677243e-05  1.48574650e-05  1.69914640e-05
1.67096151e-05  1.24416170e-05  1.96086326e-05  1.51795780e-05
2.74601385e-05  1.51393139e-05  5.07730713e-05  2.11789338e-05
1.45353519e-05  1.68304075e-05  1.84409728e-05  1.51795780e-05
2.23063295e-05  2.50845547e-05  1.50185215e-05  1.79578032e-05
1.44145595e-05  1.44145595e-05  1.35690127e-05  2.16621034e-05
1.66693509e-05  2.19439523e-05  1.39716541e-05  1.32468997e-05
1.19987115e-05  2.75406668e-05  2.21452730e-05  1.55419552e-05
1.85215010e-05  1.24818811e-05  2.41584796e-05  1.32871638e-05
1.45756160e-05  1.23208246e-05  1.67096151e-05  1.38508617e-05
1.32871638e-05  1.65888227e-05  1.75954260e-05  1.06297310e-05
9.97365679e-01
```

MNBC Classification Accuracy
MNBC classification accuracy = 0.7416666666666667

Comparison to MNBC
The MNBC handwritten classification algorithm would appear to have identical accuracy to the scikit-learn algorithm. As a result, I must conclude that the multinomial classification algorithm implemented by sklearn is essentially identical to the one we have implemented, probably without the regularizations and smoothing that we implemented.

BERT

training loss and validation loss vs epoch number
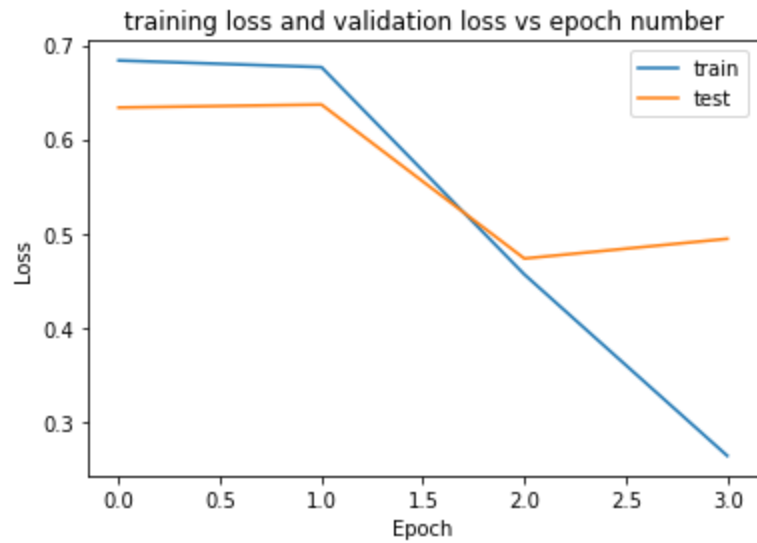
BERT accuracy while using validation split: 0.7833

For BERT:
'batch_size': 8,
 'validation_split': 0.1
    'epoch':4,
     'learning_rate': 0.00001

BERT accuracy training on entire training set without validation split: 0.8183

training loss and validation loss vs epoch number