1.1(a)

1.2(a)

(b)

(c)

Basically, the time complexity will be . But if is very sparse, it will very close to 1 pair of (x,y). Then the time complexity will be close to .

(d)

1. Using large of will take less times of iteration, which will have less computational cost but may miss the convergence point.

2. Using small of will not miss the convergence point but will take more times of iteration at the same time, which will have more computational cost.

(e)

The time complexity will keep as .

2.1

|  |  |  |
| --- | --- | --- |
| Metric | Deceased patients | Alive patients |
| Event Count 1. Average Event Count 2. Max Event Count 3. Min Event Count | 1027.7385  16829  2 | 683.1553  12627  1 |
| Encounter Count 1. Average Encounter Count 2. Median Record Count 3. Max Encounter Count 4. Min Encounter Count | 24.8393  14  375  1 | 18.6955  9  391  1 |
| Record Length 1. Average Record Length 2. Median Record Length 3. Max Record Length 4. Min Record Length | 157.0419  25  5364  0 | 194.7028  16  3103  0 |
| Common Diagnosis | DIAG320128  DIAG319835  DIAG313217  DIAG197320  DIAG132797 | DIAG320128  DIAG319835  DIAG317576  DIAG42872402  DIAG313217 |
| Common Laboratory Test | LAB3009542  LAB3023103  LAB3000963  LAB3018572  LAB3016723 | LAB3009542  LAB3000963  LAB3023103  LAB3018572  LAB3007461 |
| Common Medication | DRUG19095164  DRUG43012825  DRUG19049105  DRUG956874  DRUG19122121 | DRUG19095164  DRUG43012825  DRUG19049105  DRUG19122121  DRUG956874 |

2.3(b)

For default (, ). For value1(, ).

A close up of a map

Description automatically generatedA close up of a map

Description automatically generated

For value3(, ). For value4(, ).

A screenshot of a cell phone

Description automatically generated

A close up of a map

Description automatically generated

For value5(, ). For value6(, ).

A close up of a map

Description automatically generatedA close up of a map

Description automatically generated

When learning rate gets larger, ROC curve becomes smoother and area becomes larger.

When mu gets larger, ROC curve becomes non-smoother and area become smaller.

2.4(b)

A close up of a map

Description automatically generated

(c) Compared to the previous logistic regression model with default value, ensemble model with the same default value does not perform better, which is a little surprise. The reason could be that at least one of the five models in ensemble model performs really bad, which make the average lower than just one better model.

Or it could be not enough models together in ensemble model because current ensemble model is high bias and low variance. The way to lower the bias and upper variance is to increase the number of models in ensemble model. The current number of models is 5. Maybe increasing to 10 will get better performance.