Implementación de regresión logística para evaluación de dataset MNists

Autor 1*: Juan Sebastián Villanueva* Quintero*,* Autor 2*: Héctor Javier Rojas Silva,* Estudiantes Maestría Inteligencia Artificial, Pontificia U. Javeriana

## Diseño

Para el diseño de un sistema de reconocimiento de los dígitos a mano alzada contenidos en el dataset Mnist usando la regresión logística, hemos escogido usar 10 regresiones logísticas, una por cada digito en consecuente tendremos 10 regresiones logísticas actuando dentro del modelo.

## Procedimiento

Durante la predicción cada número se pasa por cada una de las regresiones y se obtendrá un numero entre cero y uno que se almacenará en un arreglo, este número significa la probabilidad que sea ese número, y se seleccionará el índice de la probabilidad más alta.

Después de realizar la implementación del código para leer el dataset de mnist, procesos de transformación para separar datos de entrenamiento y test y la aplicación de estandarización, como la regularización en el método fit de la clase del gradiente descendiente hicimos distintos entrenamientos los cuales detallamos a continuación.

## Resultados

### Sin estandarización ni regularización

Este es nuestro experimento base con el que vamos a medir la eficacia y eficiencia de los siguientes enfoques. Primero, realizamos el entrenamiento con una tasa de aprendizaje de 0.01 y 10000 iteraciones para todo el conjunto de datos. A continuación, se muestran las matrices de confusión resultado del proceso de evaluación en train y test.

#### Matriz de confusión de entrenamiento

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Matriz de Confusión | | | | | | | | | | | | | | |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **Precisión** | **Recall** | **F1-score** | **Total** |
| **0** | 4,692 | 1 | 29 | 8 | 9 | 30 | 31 | 6 | 39 | 8 | 93% | 97% | 95% | 4,853 |
| **1** | 2 | 5,279 | 51 | 28 | 2 | 42 | 9 | 13 | 77 | 14 | 95% | 96% | 95% | 5,517 |
| **2** | 53 | 41 | 4,321 | 88 | 64 | 32 | 57 | 47 | 161 | 12 | 84% | 89% | 86% | 4,876 |
| **3** | 31 | 25 | 222 | 4,334 | 14 | 142 | 24 | 38 | 129 | 55 | 82% | 86% | 84% | 5,014 |
| **4** | 21 | 15 | 60 | 32 | 4,342 | 23 | 44 | 24 | 71 | 166 | 87% | 90% | 89% | 4,798 |
| **5** | 84 | 27 | 62 | 320 | 94 | 3,493 | 103 | 26 | 202 | 65 | 81% | 78% | 79% | 4,476 |
| **6** | 58 | 15 | 107 | 17 | 65 | 115 | 4,412 | 4 | 41 | 1 | 93% | 91% | 92% | 4,835 |
| **7** | 27 | 28 | 92 | 111 | 73 | 26 | 5 | 4,488 | 41 | 159 | 91% | 89% | 90% | 5,050 |
| **8** | 42 | 86 | 177 | 220 | 83 | 337 | 52 | 33 | 3,609 | 101 | 79% | 76% | 77% | 4,740 |
| **9** | 31 | 27 | 40 | 113 | 264 | 92 | 9 | 254 | 210 | 3,801 | 87% | 79% | 82% | 4,841 |

Accuracy: 87% Total: 49000

#### Matriz de confusión de Test

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Matriz de Confusión | | | | | | | | | | | | | | |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **Precisión** | **Recall** | **F1-score** | **Total** |
| **0** | 2,009 | - | 18 | 5 | 5 | 10 | 9 | 3 | 15 | 2 | 93% | 97% | 95% | 2,076 |
| **1** | 1 | 2,153 | 22 | 12 | 1 | 25 | 4 | 4 | 22 | 6 | 95% | 96% | 95% | 2,250 |
| **2** | 28 | 22 | 1,840 | 38 | 26 | 17 | 29 | 15 | 63 | 5 | 83% | 88% | 86% | 2,083 |
| **3** | 9 | 9 | 98 | 1,890 | 9 | 69 | 17 | 15 | 64 | 20 | 82% | 86% | 84% | 2,200 |
| **4** | 11 | 7 | 31 | 15 | 1,836 | 9 | 11 | 13 | 31 | 68 | 87% | 90% | 89% | 2,032 |
| **5** | 41 | 12 | 30 | 136 | 30 | 1,482 | 41 | 8 | 85 | 22 | 80% | 79% | 79% | 1,887 |
| **6** | 16 | 6 | 45 | 4 | 26 | 50 | 1,906 | 1 | 15 | 2 | 93% | 92% | 93% | 2,071 |
| **7** | 15 | 5 | 42 | 57 | 31 | 8 | 3 | 1,961 | 21 | 59 | 91% | 89% | 90% | 2,202 |
| **8** | 14 | 35 | 71 | 102 | 28 | 139 | 22 | 14 | 1,608 | 50 | 80% | 77% | 78% | 2,083 |
| **9** | 10 | 16 | 24 | 45 | 125 | 43 | 4 | 121 | 93 | 1,635 | 87% | 77% | 82% | 2,116 |

Accuracy: 87% Total: 21000

### Con estandarización y sin regularización:

Ejecutamos un proceso de estandarización sobre los datos crudos tanto para entrenamiento como para test y a continuación realizamos la ejecución del entrenamiento con una tasa de aprendizaje de 0.01 y 10000 iteraciones. A continuación, se muestran las matrices de confusión tanto para entrenamiento como para test.

#### Matriz de confusión de entrenamiento

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Matriz de Confusión | | | | | | | | | | | | | | |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **Precisión** | **Recall** | **F1-score** | **Total** |
| **0** | 4,194 | 66 | 97 | 67 | 81 | 49 | 118 | 55 | 101 | 38 | 67% | 86% | 75% | 4,866 |
| **1** | 27 | 4,816 | 112 | 71 | 70 | 57 | 40 | 74 | 119 | 55 | 68% | 89% | 77% | 5,441 |
| **2** | 330 | 292 | 2,753 | 282 | 235 | 107 | 300 | 250 | 235 | 100 | 63% | 56% | 59% | 4,884 |
| **3** | 220 | 243 | 278 | 2,975 | 179 | 221 | 128 | 274 | 282 | 157 | 59% | 60% | 60% | 4,957 |
| **4** | 110 | 274 | 136 | 201 | 2,938 | 142 | 187 | 127 | 199 | 478 | 59% | 61% | 60% | 4,792 |
| **5** | 429 | 230 | 150 | 521 | 286 | 1,607 | 270 | 252 | 478 | 210 | 58% | 36% | 45% | 4,433 |
| **6** | 298 | 121 | 198 | 110 | 123 | 83 | 3,625 | 48 | 113 | 108 | 72% | 75% | 73% | 4,827 |
| **7** | 221 | 330 | 140 | 196 | 188 | 58 | 95 | 3,554 | 64 | 279 | 63% | 69% | 66% | 5,125 |
| **8** | 217 | 496 | 246 | 356 | 393 | 280 | 162 | 184 | 2,239 | 239 | 56% | 47% | 51% | 4,812 |
| **9** | 231 | 249 | 260 | 234 | 512 | 165 | 126 | 829 | 204 | 2,053 | 55% | 42% | 48% | 4,863 |

Accuracy: 63% Total: 49000

#### Matriz de confusión de Test

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Matriz de Confusión | | | | | | | | | | | | | | |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **Precisión** | **Recall** | **F1-score** | **Total** |
| **0** | 1,763 | 29 | 47 | 24 | 29 | 16 | 42 | 35 | 42 | 10 | 67% | 87% | 75% | 2,037 |
| **1** | 9 | 2,162 | 52 | 33 | 34 | 30 | 20 | 29 | 40 | 27 | 70% | 89% | 78% | 2,436 |
| **2** | 128 | 101 | 1,211 | 120 | 98 | 42 | 119 | 116 | 112 | 59 | 63% | 58% | 60% | 2,106 |
| **3** | 111 | 104 | 136 | 1,291 | 88 | 97 | 50 | 99 | 131 | 77 | 59% | 59% | 59% | 2,184 |
| **4** | 43 | 117 | 66 | 78 | 1,248 | 55 | 81 | 63 | 87 | 194 | 56% | 61% | 59% | 2,032 |
| **5** | 182 | 106 | 51 | 210 | 145 | 700 | 127 | 97 | 170 | 92 | 60% | 37% | 46% | 1,880 |
| **6** | 142 | 57 | 95 | 50 | 60 | 29 | 1,499 | 21 | 47 | 49 | 71% | 73% | 72% | 2,049 |
| **7** | 83 | 122 | 62 | 81 | 94 | 20 | 43 | 1,513 | 33 | 117 | 63% | 70% | 66% | 2,168 |
| **8** | 84 | 197 | 88 | 173 | 190 | 112 | 74 | 90 | 890 | 115 | 54% | 44% | 49% | 2,013 |
| **9** | 90 | 89 | 100 | 120 | 235 | 72 | 49 | 348 | 97 | 895 | 55% | 43% | 48% | 2,095 |

Accuracy: 63% Total: 21000

### Con estandarización y con regularización (LASSO)

Ejecutamos un proceso de estandarización sobre los datos crudos tanto para entrenamiento como para test y a continuación realizamos la ejecución del entrenamiento con una tasa de aprendizaje de 0.01, 10000 iteraciones y regularización LASSO con una lambda de 0.0001. A continuación, se muestran las matrices de confusión tanto para entrenamiento como para test.

#### Matriz de confusión de entrenamiento

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Matriz de Confusión | | | | | | | | | | | | | | |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **Precisión** | **Recall** | **F1-score** | **Total** |
| **0** | 4,763 | 2 | 8 | 7 | 11 | 15 | 24 | 9 | 27 | 1 | 95% | 98% | 97% | 4,867 |
| **1** | 1 | 5,301 | 32 | 12 | 4 | 17 | 6 | 8 | 50 | 4 | 94% | 98% | 96% | 5,435 |
| **2** | 39 | 48 | 4,462 | 62 | 47 | 23 | 53 | 63 | 107 | 18 | 93% | 91% | 92% | 4,922 |
| **3** | 21 | 25 | 102 | 4,441 | 6 | 131 | 21 | 57 | 92 | 55 | 91% | 90% | 90% | 4,951 |
| **4** | 8 | 22 | 19 | 9 | 4,482 | 11 | 31 | 13 | 37 | 133 | 92% | 94% | 93% | 4,765 |
| **5** | 41 | 28 | 22 | 133 | 58 | 3,905 | 90 | 18 | 94 | 56 | 90% | 88% | 89% | 4,445 |
| **6** | 34 | 18 | 26 | 4 | 27 | 76 | 4,638 | 6 | 22 | 3 | 94% | 96% | 95% | 4,854 |
| **7** | 17 | 34 | 53 | 14 | 45 | 4 | 6 | 4,815 | 16 | 151 | 93% | 93% | 93% | 5,155 |
| **8** | 41 | 124 | 53 | 109 | 48 | 117 | 41 | 23 | 4,131 | 70 | 89% | 87% | 88% | 4,757 |
| **9** | 29 | 21 | 13 | 79 | 154 | 45 | 2 | 167 | 43 | 4,296 | 90% | 89% | 89% | 4,849 |

Accuracy: 92% Total: 49000

#### Matriz de confusión de Test

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Matriz de Confusión | | | | | | | | | | | | | | |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **Precisión** | **Recall** | **F1-score** | **Total** |
| **0** | 1,970 | - | 9 | 3 | 2 | 13 | 17 | 3 | 16 | 3 | 94% | 97% | 95% | 2,036 |
| **1** | 2 | 2,359 | 13 | 9 | 5 | 17 | 2 | 4 | 27 | 4 | 94% | 97% | 95% | 2,442 |
| **2** | 17 | 23 | 1,826 | 35 | 31 | 11 | 26 | 35 | 56 | 8 | 91% | 88% | 89% | 2,068 |
| **3** | 8 | 10 | 54 | 1,927 | 12 | 78 | 8 | 18 | 50 | 25 | 89% | 88% | 88% | 2,190 |
| **4** | 4 | 11 | 18 | 11 | 1,870 | 8 | 17 | 9 | 24 | 87 | 90% | 91% | 91% | 2,059 |
| **5** | 28 | 10 | 6 | 78 | 18 | 1,613 | 35 | 10 | 46 | 24 | 87% | 86% | 87% | 1,868 |
| **6** | 16 | 8 | 18 | 2 | 9 | 31 | 1,926 | 3 | 8 | 1 | 94% | 95% | 95% | 2,022 |
| **7** | 14 | 15 | 26 | 13 | 25 | 5 | 4 | 1,978 | 7 | 51 | 92% | 93% | 92% | 2,138 |
| **8** | 30 | 67 | 36 | 61 | 22 | 60 | 11 | 9 | 1,729 | 43 | 88% | 84% | 86% | 2,068 |
| **9** | 16 | 16 | 8 | 38 | 78 | 16 | 4 | 83 | 13 | 1,837 | 88% | 87% | 88% | 2,109 |

Accuracy: 91% Total: 21000

### Con estandarización y con regularización (RIDGE)

Ejecutamos un proceso de estandarización sobre los datos crudos tanto para entrenamiento como para test y a continuación realizamos la ejecución del entrenamiento con una tasa de aprendizaje de 0.01, 10000 iteraciones y regularización RIDGE con una lambda de 10000. A continuación, se muestran las matrices de confusión tanto para entrenamiento como para test.

#### Matriz de confusión de entrenamiento

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Matriz de Confusión | | | | | | | | | | | | | | |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **Precisión** | **Recall** | **F1-score** | **Total** |
| **0** | 4,687 | - | 12 | 11 | 15 | 35 | 31 | 13 | 27 | 3 | 93% | 97% | 95% | 4,834 |
| **1** | 1 | 5,339 | 32 | 14 | 6 | 31 | 7 | 11 | 57 | 10 | 93% | 97% | 95% | 5,508 |
| **2** | 45 | 55 | 4,348 | 75 | 64 | 19 | 57 | 62 | 121 | 10 | 93% | 90% | 91% | 4,856 |
| **3** | 41 | 37 | 107 | 4,434 | 21 | 137 | 22 | 52 | 89 | 60 | 90% | 89% | 89% | 5,000 |
| **4** | 15 | 30 | 27 | 15 | 4,395 | 17 | 39 | 17 | 39 | 179 | 90% | 92% | 91% | 4,773 |
| **5** | 54 | 39 | 23 | 147 | 56 | 3,898 | 91 | 23 | 106 | 50 | 89% | 87% | 88% | 4,487 |
| **6** | 49 | 23 | 24 | 2 | 29 | 63 | 4,626 | 4 | 25 | 3 | 94% | 95% | 95% | 4,848 |
| **7** | 24 | 45 | 49 | 18 | 68 | 10 | 5 | 4,719 | 12 | 140 | 93% | 93% | 93% | 5,090 |
| **8** | 62 | 154 | 52 | 129 | 69 | 129 | 40 | 17 | 4,047 | 79 | 89% | 85% | 87% | 4,778 |
| **9** | 43 | 35 | 13 | 74 | 172 | 44 | 5 | 154 | 46 | 4,240 | 89% | 88% | 88% | 4,826 |

Accuracy: 91% Total: 49000.

#### Matriz de confusión de Test

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Matriz de Confusión | | | | | | | | | | | | | | |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **Precisión** | **Recall** | **F1-score** | **Total** |
| **0** | 1,998 | 4 | 6 | 5 | 6 | 12 | 15 | 4 | 17 | 2 | 93% | 97% | 95% | 2,069 |
| **1** | - | 2,297 | 16 | 10 | 2 | 5 | 2 | 3 | 33 | 1 | 92% | 97% | 94% | 2,369 |
| **2** | 30 | 27 | 1,864 | 44 | 33 | 14 | 26 | 38 | 43 | 15 | 91% | 87% | 89% | 2,134 |
| **3** | 18 | 15 | 61 | 1,843 | 11 | 82 | 9 | 26 | 54 | 22 | 88% | 86% | 87% | 2,141 |
| **4** | 4 | 15 | 11 | 5 | 1,878 | 8 | 16 | 11 | 20 | 83 | 89% | 92% | 90% | 2,051 |
| **5** | 33 | 10 | 13 | 63 | 23 | 1,562 | 33 | 9 | 57 | 23 | 86% | 86% | 86% | 1,826 |
| **6** | 16 | 12 | 18 | 4 | 25 | 39 | 1,897 | 5 | 10 | 2 | 94% | 94% | 94% | 2,028 |
| **7** | 14 | 23 | 28 | 18 | 36 | 9 | 3 | 2,004 | 8 | 60 | 91% | 91% | 91% | 2,203 |
| **8** | 22 | 86 | 29 | 53 | 21 | 67 | 21 | 16 | 1,697 | 35 | 87% | 83% | 85% | 2,047 |
| **9** | 18 | 20 | 9 | 47 | 77 | 17 | 1 | 97 | 15 | 1,831 | 88% | 86% | 87% | 2,132 |

Accuracy: 90% Total: 21000.

## Resumen Comparativo

Para el grafico se comparan los accuracy de cada clase y el global de todas las métricas obtenidas en todos los experimentos realizados y usando solamente los datos arrojados por el datase de Test.

## Conclusiones

* Los mejores resultados se obtuvieron usando en conjunto estandarización y regularización.
* Los peores resultados se dieron cuando solamente se usó estandarización.
* Los tiempos mas altos de entrenamiento se dieron sin usar regularización ni estandarización.
* La clase más susceptible a fallos es 5 con todas las técnicas excepto con regularización únicamente en ese caso fue el 4.
* Las clases con mejor desempeño son 0, 1 y 6 en todas las ocasiones.