

|  |  |  |  |
| --- | --- | --- | --- |
| Metric | Deceased patients | Alive patients | Function to complete |
| Event Count  1. Average Event Count 2. Max Event Count 3. Min Event Count | 8635, 1, 982.014 | 12627, 1, 498.118 | event count metrics |
| Encounter Count  1. Average Encounter Count 2. Max Encounter Count 3. Min Encounter Count | 203, 1, 23.038 | 391, 1, 15.452 | encounter count metrics |
| Record Length  1. Average Record Length 2. Max Record Length 3. Min Record Length | 1972, 0, 127.532 | 2914, 0, 159.2 | record length metrics |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Accuracy | AUC | Precision | Recall | F-Score |
| Logistic Regression | 0.954 | 0.945 | 0.987 | 0.899 | 0.941 |
| SVM | 0.994 | 0.995 | 0.988 | 0.997 | 0.993 |
| Decision Tree | 0.776 | 0.747 | 0.792 | 0.601 | 0.683 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Accuracy | AUC | Precision | Recall | F-Score |
| Logistic Regression | 0.738 | 0.737 | 0.680 | 0.733 | 0.706 |
| SVM | 0.738 | 0.739 | 0.677 | 0.744 | 0.709 |
| Decision Tree | 0.671 | 0.656 | 0.632 | 0.555 | 0.591 |

Strategies:

1. Do parameter tuning. The parameters in the algorithms are using the default values. Tuning the parameters will give a better performance to the algorithm.
2. Use another algorithm. Generally speaking, the performance of deep learning is better than that of traditional machine learning methods. Algorithms like LSTM will provide a better performance.

|  |  |  |
| --- | --- | --- |
| CV strategy | Accuracy | AUC |
| K-Fold | 0.725 | 0.710 |
| Randomized | 0.738 | 0.719 |

My Model:

New features: I selected the counting of DIAG, DRUG and LAB as features and used SVM to train the model. This time, I tuned the parameter C to be 0.01 to give a better performance, reaching the AUC 0.69.