预后模型

一、不使用Lasso回归进行特征选择，直接使用下面24个特征进行模型构建

```
0    Age                                        400 non-null    int64
1    Gender                                     400 non-null    int64
2    Hypertension                               400 non-null    int64
3    Diabetes                                   400 non-null    int64
4    CHF                                        400 non-null    int64
5    COPD                                       400 non-null    int64
6    Self-Care                                  400 non-null    int64
7    5-MFI                                      400 non-null    int64
8    Osteoporosis                               400 non-null    int64
9    BMI                                        400 non-null    float64
10   PNI                                        400 non-null    float64
11   Albumin                                    400 non-null    float64
12   Absolute Lymphocyte Count                  400 non-null    float64
13   CRP                                        400 non-null    float64
14   Hb                                         400 non-null    int64
15   Lactate                                    400 non-null    float64
16   Surgery Duration                           400 non-null    int64
17   Anesthesia Type                            400 non-null    int64
18   ASA                                        400 non-null    int64
19   Hospital Stay Duration                     400 non-null    int64
20   Perioperative Blood Product Preparation    400 non-null    int64
21   Hospitalization Cost                       400 non-null    int64
22   ADL Preoperative                           400 non-null    int64
23   ADL Postoperative                          400 non-null    int64
```
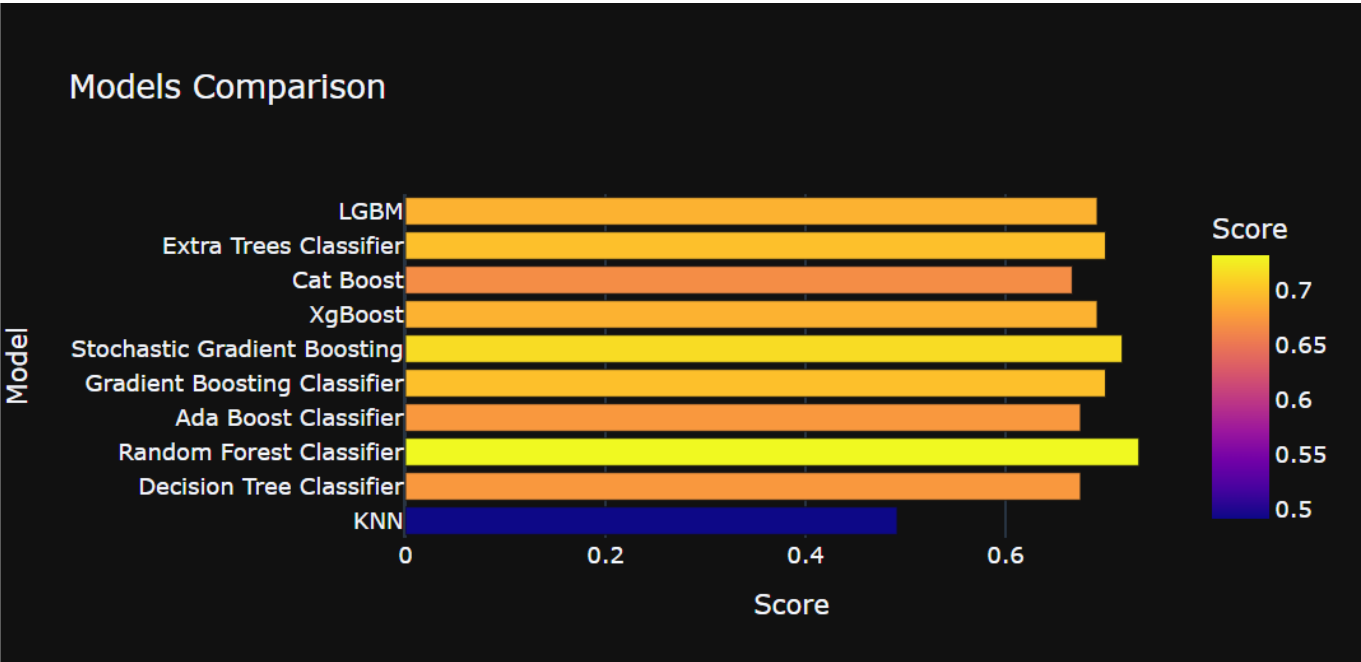
模型准确性结果如下

	Model	Score
2	Random Forest Classifier	0.733333
5	Stochastic Gradient Boosting	0.716667
4	Gradient Boosting Classifier	0.700000
8	Extra Trees Classifier	0.700000
6	XgBoost	0.691667
9	LGBM	0.691667
1	Decision Tree Classifier	0.675000
3	Ada Boost Classifier	0.675000
7	Cat Boost	0.666667
0	KNN	0.491667

准确率最高的模型是Random Forest Classifier 准确性为0.733333

ROC曲线如下
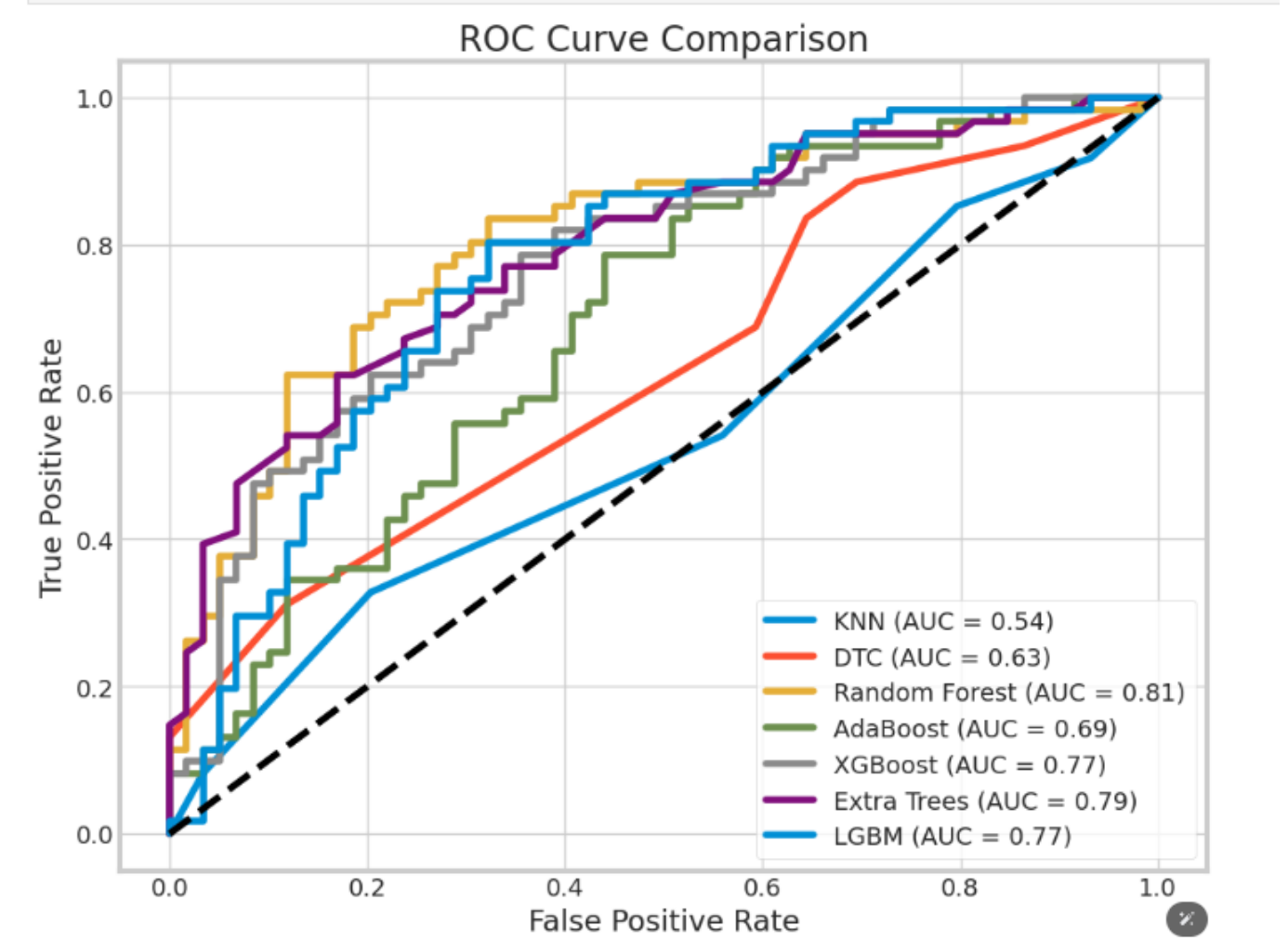
二、使用Lasso回归进行特征选择

Lasso回归结果
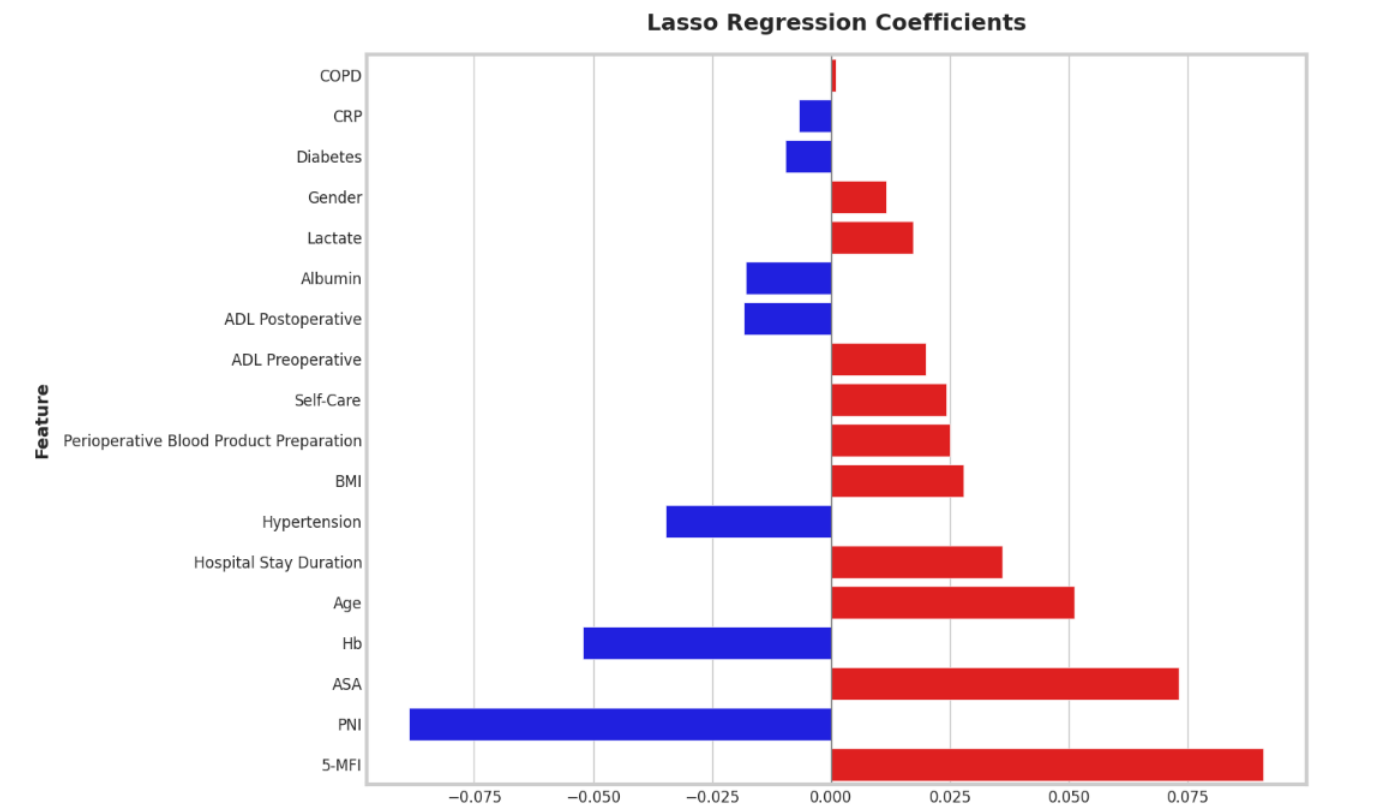
	Feature	Coefficient
0	Age	0.051091
1	Gender	0.011562
2	Hypertension	-0.034805
3	Diabetes	-0.009626
4	COPD	0.000871
5	Self-Care	0.024228
6	5-MFI	0.090868
7	BMI	0.027908
8	PNI	-0.088656
9	Albumin	-0.017930
10	CRP	-0.006837
11	Hb	-0.052245
12	Lactate	0.017152
13	ASA	0.073025
14	Hospital Stay Duration	0.035960
15	Perioperative Blood Product Preparation	0.025026
16	ADL Preoperative	0.019883
17	ADL Postoperative	-0.018379

模型准确性结果如下

	Model	Score
7	Cat Boost	0.758333
3	Ada Boost Classifier	0.733333
5	Stochastic Gradient Boosting	0.725000
2	Random Forest Classifier	0.716667
6	XgBoost	0.716667
8	Extra Trees Classifier	0.700000
4	Gradient Boosting Classifier	0.691667
9	LGBM	0.675000
0	KNN	0.633333
1	Decision Tree Classifier	0.600000

```
     Model                              Score
7    Cat Boost                          0.758333
3    Ada Boost Classifier               0.733333
5    Stochastic Gradient Boosting       0.725000
2    Random Forest Classifier           0.716667
6    XgBoost                            0.716667
8    Extra Trees Classifier             0.700000
4    Gradient Boosting Classifier       0.691667
9    LGBM                               0.675000
```

```
0    KNN                                    0.633333
1    Decision Tree Classifier               0.600000
```

ROC曲线如下