	[[59 0 0 0 0] [0 16 0 0 0] [0 0 32 0 0] [0 0 0 23 0] [0 0 0 0 31]] precision recall f1-score support									
	PRAD LUAD BRCA KIRC COAD	1.00	1.00							
W	accuracy macro avg eighted avg	1.00	1.00	1.00 1.00 1.00	161 161 161					
	Logit_ridge.C_	60 07	1.00	5440455	E4404-	1 [4:5	21.			
	rray([21.544346 Logit_lasso = I							ılti class = '	'multinomia	al",sol
;]	Logit_lasso.fit	t(X_train,	np.ravel(y_train))				_		
С	<pre>y_pred = Logit_ :\Users\matth\a s reached which</pre>	anaconda3\1	_ .ib\site-p	oackages\skl		r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
С	<pre>warnings.warn :\Users\matth\a s reached which</pre>	("The max_i anaconda3\1 h means the	ter was r ib\site-p coef_ di	reached whic backages\skl id not conve	h means " earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
а	<pre>warnings.warn :\Users\matth\a s reached which warnings.warn</pre>	anaconda3\l h means the ("The max_i	ib\site-p coef_ di ter was r	oackages\skl id not conve reached whic	earn\linea rge h means "		_			_
а	<pre>:\Users\matth\a s reached which warnings.warn :\Users\matth\a</pre>	h means the ("The max_i	coef_ di ter was r	id not conve	rge h means "		_			
a C	s reached which warnings.warn:\Users\matth\assign{a} s reached which	n means the ("The max_i anaconda3\l	coef_ di ter was r ib\site-p	d not conve reached whic backages\skl	rge h means " earn\linea		_			_
С	warnings.warn:\Users\matth\assign s reached which	("The max_i anaconda3\l h means the	ter was r ib\site-p coef_ di	reached whic backages\skl ld not conve	h means " earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
	<pre>warnings.warn :\Users\matth\a s reached which warnings.warn</pre>	anaconda3\l h means the	ib\site-p coef_ di	oackages\skl id not conve	earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
a	:\Users\matth\a s reached which warnings.warn	anaconda3\l h means the ("The max_i	ib\site-p coef_ di ter was r	oackages\skl ld not conve reached whic	earn\linea rge h means "		_			
a C	:\Users\matth\a s reached which warnings.warn :\Users\matth\a	n means the ("The max_i anaconda3\l	coef_ di ter was r ib\site-p	d not conve reached whic backages\skl	rge h means " earn\linea		_			
a C	s reached which warnings.warn:\Users\matth\as reached which	n means the ("The max_i anaconda3\l	coef_ di ter was r ib\site-p	d not conve reached whic backages\skl	rge h means " earn\linea		_			_
С	warnings.warn:\Users\matth\assign s reached which	("The max_i anaconda3\l h means the	ter was r ib\site-p coef_ di	reached whic backages\skl id not conve	h means " earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
С	warnings.warn :\Users\matth\a s reached which warnings.warn	("The max_i anaconda3\l h means the	ter was r ib\site-p coef_ di	reached whic backages\skl ld not conve	h means " earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
a	:\Users\matth\a s reached which warnings.warn	anaconda3\l h means the ("The max_i	ib\site-p coef_ di ter was r	oackages\skl id not conve reached whic	earn\linea rge h means "					
а	:\Users\matth\a s reached which warnings.warn :\Users\matth\a	h means the ("The max_i	coef_ di ter was r	id not conve	rge h means "		_			
a C	<pre>s reached which warnings.warn :\Users\matth\a</pre>	h means the ("The max_i anaconda3\l	coef_ di ter was r ib\site-p	d not convereached whic backages\skl	rge h means " earn\linea		_			
a C	s reached which warnings.warn:\Users\matth\assign{a} s reached which	h means the ("The max_i anaconda3\l	coef_ di ter was r ib\site-p	id not conve reached whic backages\skl	rge h means " earn\linea		_			
С	warnings.warn:\Users\matth\assign s reached which	("The max_i anaconda3\l h means the	ter was r ib\site-p coef_ di	reached whic backages\skl id not conve	h means " earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max ₋
	<pre>warnings.warn :\Users\matth\a s reached which warnings.warn</pre>	anaconda3\1 h means the	ib\site-p coef_ di	oackages\skl id not conve	earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
a	:\Users\matth\a s reached which warnings.warn	anaconda3\l h means the ("The max_i	ib\site-p coef_ di ter was r	oackages\skl ld not conve reached whic	earn\linea rge h means "		_			
а	:\Users\matth\a s reached which warnings.warn :\Users\matth\a	anaconda3\l h means the ("The max_i	ib\site-p coef_ di ter was r	oackages\skl ld not conve reached whic	earn\linea rge h means "		_			
a C	s reached which warnings.warn:\Users\matth\a	n means the ("The max_i anaconda3\l	coef_ di ter was r ib\site-p	d not conve reached whic backages\skl	rge h means " earn\linea		_			
a C	s reached which warnings.warn:\Users\matth\as reached which	h means the ("The max_i anaconda3\l	coef_ di ter was r ib\site-p	d not convereached whic backages\skl	rge h means " earn\linea		_			
С	<pre>warnings.warn :\Users\matth\a s reached which</pre>	("The max_i anaconda3\l h means the	ter was r ib\site-p coef_ di	reached whic backages\skl id not conve	h means " earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
	<pre>warnings.warn :\Users\matth\a s reached which warnings.warn</pre>	anaconda3\l n means the	ib\site-p coef_ di	oackages\skl id not conve	earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
a	:\Users\matth\a s reached which warnings.warn	anaconda3\l h means the ("The max_i	ib\site-p coef_ di ter was r	oackages\skl ld not conve reached whic	earn\linea rge h means "		_			
a	<pre>:\Users\matth\a s reached which warnings.warn :\Users\matth\a</pre>	h means the ("The max_i	coef_ di ter was r	ld not conve	rge h means "		_			
a C	s reached which warnings.warn:\Users\matth\a	n means the ("The max_i anaconda3\l	coef_ di ter was r ib\site-p	d not conve reached whic backages\skl	rge h means " earn\linea		_			
С	<pre>s reached which warnings.warn :\Users\matth\a s reached which</pre>	("The max_i anaconda3\l h means the	ter was r ib\site-p coef_ di	reached whic backages\skl id not conve	h means " earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
С	warnings.warn:\Users\matth\as reached which warnings.warn	("The max_i anaconda3\l h means the	ter was r ib\site-p coef_ di	reached whic backages\skl ld not conve	h means " earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
а	:\Users\matth\a s reached which warnings.warn	anaconda3\l h means the ("The max_i	ib\site-p coef_ di ter was r	backages\skl id not conve reached whic	earn\linea rge h means "		_			
a C	:\Users\matth\a s reached which warnings.warn :\Users\matth\a	n means the ("The max_i anaconda3\l	coef_ di ter was r ib\site-p	d not conve reached whic backages\skl	rge h means " earn\linea		_			_
a C	s reached which warnings.warn:\Users\matth\a	n means the ("The max_i anaconda3\l	coef_ di ter was r ib\site-p	d not conve reached whic backages\skl	rge h means " earn\linea		_			_
С	<pre>s reached which warnings.warn :\Users\matth\a s reached which</pre>	("The max_i anaconda3\l h means the	ter was r ib\site-p coef_ di	reached whic backages\skl ld not conve	h means " earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
С	warnings.warn:\Users\matth\as reached which warnings.warn	("The max_i anaconda3\l h means the	ter was r ib\site-p coef_ di	reached whic backages\skl id not conve	h means " earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
	<pre>warnings.warn :\Users\matth\a s reached which warnings.warn</pre>	anaconda3\1 h means the	ib\site-p coef_ di	oackages\skl id not conve	earn\linea rge	r_model_	_sag.py:328	: Convergence	Warning: Tl	he max_
	print(confusion	n_matrix(y_	test, y_p	ored))		s=y.Class	.unique())			
[[59 0 0 0 0 ([0 16 0 0 (0]	_		_					
	[0 0 32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0]	recall	f1-score	support					
	PRAD LUAD	1.00	1.00	1.00	59 16					
	BRCA KIRC COAD	1.00 1.00 1.00	1.00 1.00 1.00	1.00 1.00 1.00	32 23 31					
W	accuracy macro avg eighted avg	1.00	1.00	1.00 1.00 1.00	161 161 161					
]	Logit_lasso.C_									
a	array([0.04641589, 0.04641589, 0.04641589, 0.04641589])									
	coef = pd.DataF	Frame(data	= Logit_l	.asso.coef_,	columns =	X.column	ns, index =	y.Class.uniqu	ne())	
	coef = coef.loc	c[:, (coef	!= 0).any	/(axis=0)]						
	coef.iloc[[0]].							none 257	1652	740
	gene_ 7964 -0.119404	gene_15589 0.119164	-0.109951	gene_18746 0.086577		-0.077042		gene_357 gene_9 0.072482 0.067		
	0.115404		.argsort(-np.abs(coe	f.values[1]))[:20]]				
	coef.iloc[[1]].	.iloc[:, np		gene 11652 g		ne_ 3532 g		ene_ 2037 gene_ 7 0.028416 0.02		
P	coef.iloc[[1]].		ene_14845 0.046956	0.041404	0.040591					
P	coef.iloc[[1]].	gene_ 4804 go	0.046956	0.041404))[:20]]				
P	<pre>gene_4805 g UAD 0.050317 coef.iloc[[2]].</pre>	gene_4804 g 0.048228 .iloc[:, np	0.046956 .argsort(0.041404 -np.abs(coe	f.values[2]			ne_219 gene_161	30 gene_161	
: P	<pre>gene_4805 g UAD 0.050317 coef.iloc[[2]].</pre>	gene_4804 g 0.048228 .iloc[:, np	0.046956 .argsort(0.041404 (-np.abs(coe	f.values[2]		gene_450 gei	ne_ 219 		105 ger
P	gene_4805 (UAD	gene_4804 gc 0.048228 .iloc[:, np gene_14114 0.075337	0.046956 .argsort(gene_16131 0.074638 .argsort(0.041404 -np.abs(coe 1 gene_220 g 3 0.069651 -np.abs(coe	f.values[2] ene_16173 g 0.065259 f.values[3]	ene_4429 0.052744 ())[:20]]	gene_ 450 gel	D44649 0.0371	89 0.0343	105 ger 318
P	gene_4805 (UAD	gene_4804 gc 0.048228 .iloc[:, np gene_14114 0.075337	0.046956 .argsort(gene_16131 0.074638 .argsort(0.041404 -np.abs(coe 1 gene_220 g 3 0.069651 -np.abs(coe gene_15898	f.values[2] ene_16173 g 0.065259 f.values[3]	ene_4429 0.052744 ())[:20]]	gene_450 ger 0.04643 0.0 6 gene_11352	0.0371 044649 0.0371 gene_11903 g	89 0.0343	1 05 ger 318 ene_ 1516
P Ltu	gene_4805 (UAD	gene_4804 gc 0.048228 .iloc[:, np gene_14114 0.075337 .iloc[:, np gene_15895 0.10961	0.046956 .argsort (gene_16131 0.074638 .argsort (gene_13639 0.083694	0.041404 -np.abs(coe 1 gene_220 g 3 0.069651 -np.abs(coe gene_15898 0.073549	f.values[2] ene_16173 g 0.065259 f.values[3 gene_15591 0.066675	ene_4429 0.052744 ()))[:20]] gene_15896 0.064456	gene_450 ger 0.04643 0.0 6 gene_11352 6 0.049966	0.0371 044649 0.0371 gene_11903 g	89 0.0343 ene_ 7058 g e	105 ge n 318

In [1]:

In [2]:

In [3]:

In [4]:

In [5]:

Out[5]:

In [6]:

In [7]:

In [8]:

In [9]:

X.head()

gene_0

2 -0.213757 0.341956

3 -0.213757 0.487482

5 rows × 20221 columns

import pandas as pd
import numpy as np

from sklearn import metrics

X = X.dropna(axis='columns')

scaler = StandardScaler()

gene_1

1 -0.213757 -1.797344 -1.403860

from sklearn.preprocessing import StandardScaler
from sklearn.preprocessing import PowerTransformer
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegressionCV

X = pd.read_csv("gene_data.csv", index_col = 0)
y = pd.read_csv("gene_labels.csv", index_col = 0)

pt = PowerTransformer(method = "yeo-johnson")

gene_2

1.162020

1.336004

Logit_ridge.fit(X_train, np.ravel(y_train))

y_pred = Logit_ridge.predict(X_test)

print(confusion_matrix(y_test, y_pred))

X = X.drop(X.var()[X.var() < 0.001].index.values, axis=1)

X = pd.DataFrame(scaler.fit transform(X), columns = X.columns)

gene_3

gene_4

4 -0.213757 -0.397980 -0.267585 -0.191536 -0.108244 -0.765225 -0.080037 -0.142766 -0.070844 -0.949409 ...

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.20, random_state = 101)

print(classification_report(y_test, y_pred, target_names=y.Class.unique()))

0.012614 0.773066

gene_6

1.300770 -0.343099 -0.547056 -1.125183 -0.142766 -0.070844

gene_7

Logit_ridge = LogisticRegressionCV(cv=10, scoring='accuracy', n_jobs=-1, multi_class = "multinomial", max_iter

gene_8

0.135235 -0.142766 -0.070844

gene_9

0.460090 ...

-0.949409 ...

-0.949409 ...

gene_10 ... gene_20521 gene_20522 gei

-0.921700

-2.431478

-1.062246

0.054274

0.217140

-1.259121

-1.601084

-1.045556

0.146152

0.029564

X = pd.DataFrame(pt.fit_transform(X), columns = X.columns)

from sklearn.metrics import classification_report, confusion_matrix