

East Delta University

Department of Computer Science and Engineering

Topic
"Support Vector Machine - Basic Principle Code"

Submitted To: Arshiana Shamir, Lecturer SCHOOL OF SCIENCE, ENGINEERING AND TECHNOLOGY

Submitted By:

Mohammed Shahadat Hossain Talukder ID: 233000408

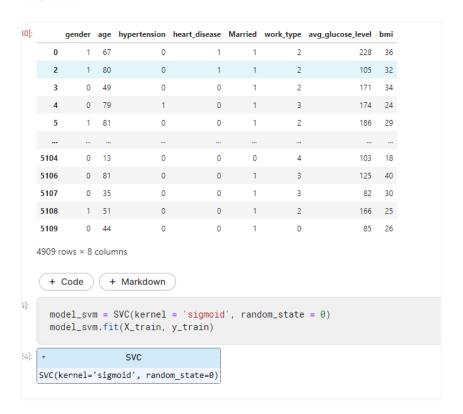
Subject: Natural Language Processing Year: 2024 Semester: 2nd (M.Sc)

Introduction: This coding prompt implements Python and Kaggle Notebook to predict using the Support Vector Machine (SVM) algorithm.

Dataset : This dataset is collected from kaggle where it contains 12 columns called ld gender,age,hypertension,heart_disease,ever_married,work_type,Residence_type, avg_glucose_level, bmi, smoking_status, stroke with 5110 Rows.

Results:

SVM Test 1: Selected Features removing "residence_type, smoking_status" and kernel "Sigmoid"





Accuracy: 95.04%

SVM Test 2: Selected Features All with kernel "linear"

[4]:		gender	age	hypertension	heart_disease	Married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status
	0	1	67	0	1	1	2	1	228	36	1
	2	1	80	0	1	1	2	0	105	32	2
	3	0	49	0	0	1	2	1	171	34	3
	4	0	79	1	0	1	3	0	174	24	2
	5	1	81	0	0	1	2	1	186	29	1
	5104	0	13	0	0	0	4	0	103	18	0
	5106	0	81	0	0	1	3	1	125	40	2
	5107	0	35	0	0	1	3	0	82	30	2
	5108	1	51	0	0	1	2	0	166	25	1
	5109	0	44	0	0	1	0	1	85	26	0

4909 rows × 10 columns

```
ij: model_svm = SVC(kernel = 'linear', random_state = θ)
    model_svm.fit(X_train, y_train)
(5]: SVC
SVC(kernel='linear', random_state=θ)
```

Accuracy: 0.9592668024439919

classification_rep_svm = classification_report(y
print(classification_rep_svm)

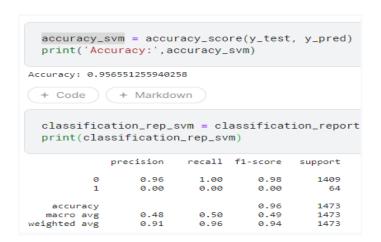
precision recall f1-score support

0	0.96	1.00	0.98	1413
1	0.00	0.00	0.00	60
accuracy macro avg weighted avg	0.48 0.92	0.50 0.96	0.96 0.49 0.94	1473 1473 1473

Accuracy: 95.92%

SVM Test 3: Selected Features removing "Married, work_type" and kernel "Poly"

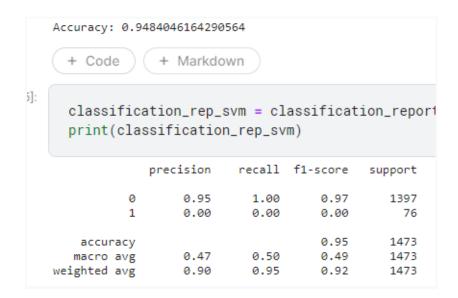
[14]:		gender	age	hypertension	heart_disease	Residence_type	avg_glucose_level	bmi	smoking_status		
	0	1	67	0	1	1	228	36	1		
	2	1	80	0	1	0	105	32	2		
	3	0	49	0	0	1	171	34	3		
	4	0	79	1	0	0	174	24	2		
	5	1	81	0	0	1	186	29	1		
	5104	0	13	0	0	0	103	18	0		
	5106	0	81	0	0	1	125	40	2		
	5107	0	35	0	0	0	82	30	2		
	5108	1	51	0	0	0	166	25	1		
	5109	0	44	0	0	1	85	26	0		
4909 rows × 8 columns + Code + Markdown											
]:	<pre>model_svm = SVC(kernel = 'poly', random_state = 0) model_svm.fit(X_train, y_train)</pre>										
5]:	*			SVC							
	SVC(k	cernel=	poly	', random_s	tate=0)						



Accuracy: 95.65%

SVM Test 4: Selected Features removing "age, hypertension, gender" and kernel "Gaussian"

19]:		heart_disease	Married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status				
	0	1	1	2	1	228	36	1				
	2	1	1	2	0	105	32	2				
	3	0	1	2	1	171	34	3				
	4	0	1	3	0	174	24	2				
	5	0	1	2	1	186	29	1				
	5104	0	0	4	0	103	18	0				
	5106	0	1	3	1	125	40	2				
	5107	0	1	3	0	82	30	2				
	5108	0	1	2	0	166	25	1				
	5109	0	1	0	1	85	26	0				
	4909 r	ows × 8 colum	ins									
	+	Code +	Markdo	wn								
]:	<pre>model_svm = SVC(kernel = 'rbf', random_state = 0) model_svm.fit(X_train, y_train)</pre>											
0]:	۳	SVC										
	SVC(r	andom_state	=0)									



Accuracy: 94.84%

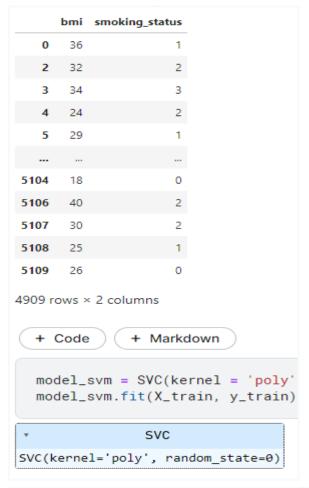
SVM Test 5: Selected Features removing "' avg_glucose_level, heart_disease" and kernel "sigmoid"

	gender	age	hypertension	Married	work_type	Residence_type	bmi	smoking_status
0	1	67	0	1	2	1	36	1
2	1	80	0	1	2	0	32	2
3	0	49	0	1	2	1	34	3
4	0	79	1	1	3	0	24	2
5	1	81	0	1	2	1	29	1
5104	0	13	0	0	4	0	18	0
5106	0	81	0	1	3	1	40	2
5107	0	35	0	1	3	0	30	2
5108	1	51	0	1	2	0	25	1
5109	0	44	0	1	0	1	26	0
4909 r	ows × 9	colun	nns					
	Code		Markdown					
mod mod	del_svr	n.fi	t(X_train,	y_train				
woods	cernel=	'sigm		om_state:				
svc(k	cernel=	'sign	SVC moid', rando	om_state:				
woods	curacy:	0.90	svc moid', rando 660556687033 + Markdo	om_state= 266 own	_{=Ø)}	ation_report		
svc(k	curacy:	0.90	svc moid', rando 660556687033 + Markdo	om_state 266 own svm = c n_rep_s	_{=Ø)}	-		
SVC(k	curacy:	0.90	svc moid', rando 660556687033 + Markdo ation_rep_s	om_state 266 own svm = c n_rep_s	lassifica vm)	e support		
SVC(k	curacy:	0.90	svc moid', rando 660556687033 + Markdo ation_rep_s ssification precision	om_state: 266 own svm = c n_rep_s recall	lassifica vm) f1-score	support		
svc(k	curacy: + Code classi	o.90	svc moid', rando 660556687033 + Markdo ation_rep_s ssification precision 0.97	om_state: 266 own svm = c n_rep_s recall 1.00	lassifica vm) f1-score 0.98 0.00	support 3 1423 0 50		
svc(k	curacy: + Code classi print(0.90 e ifica (class	svc moid', rando m	om_state= 266 own svm = c n_rep_s recall 1.00 0.00	lassifica vm) f1-score 0.98 0.00	support 3 1423 5 50 7 1473		
SVC(k	curacy: + Code classi	o.90 e ifica (clas	svc moid', rando 660556687033 + Markdo ation_rep_s ssification precision 0.97	om_state: 266 own svm = c n_rep_s recall 1.00	lassifica vm) f1-score 0.98 0.00	support 1423 50 7 1473 9 1473		

Accuracy : 96.60%

SVM Test 6: Selected Features removing

('age','work_type','gender','hypertension','Married','Residence_type','avg_glucose_level','heart_disease') and kernel "poly"



Accuracy: 0.9511201629327902 classification_rep_svm = classification_repo print(classification_rep_svm) precision recall f1-score support 0.95 0 1.00 0.97 1401 1 0.00 0.00 0.00 72 0.95 1473 accuracy macro avg 0.48 0.50 weighted avg 0.90 0.95 0.49 1473 1473 0.93

Accuracy: 95.11%

SVM Test 7: Selected Features all and kernel "gaussian"

	gender	age	hypertension	heart_disease	Married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	
0	1	67	0	1	1	2	1	228	36	1	
2	1	80	0	1	1	2	0	105	32	2	
3	0	49	0	0	1	2	1	171	34	3	
4	0	79	1	0	1	3	0	174	24	2	
5	1	81	0	0	1	2	1	186	29	1	
5104	0	13	0	0	0	4	0	103	18	(
5106	0	81	0	0	1	3	1	125	40	2	
5107	0	35	0	0	1	3	0	82	30	2	
5108	1	51	0	0	1	2	0	166	25	1	
5109	0	44	0	0	1	0	1	85	26	0	
4909 rows × 10 columns + Code + Markdown											
<pre>model_svm = SVC(kernel = 'rbf', random_state = 0) model_svm.fit(X_train, y_train)</pre>											
•	SV	С									
SVC(r	andom_	state	9=0)								

Accuracy: 0.9619823489477257

classification_rep_svm = classification_report
print(classification_rep_svm)

	precision	recall	f1-score	support
0	0.96	1.00	0.98	1417
1	0.00	0.00	0.00	56
accuracy			0.96	1473
macro avg	0.48	0.50	0.49	1473
weighted avg	0.93	0.96	0.94	1473

Accuracy: 96.19%