



Model Optimization and Tuning Phase Template

Date	15 July 2024
Team ID	739699
Project Title	Telecom Customer Churn Prediction
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Hyperparameter Tuning Documentation (6 Marks):

Model	Tuned Hyperparameters	Optimal Values
-	-	-

Performance Metrics Comparison Report (2 Marks):

Optimized Metric					
accuracy macro avg weighted avg	1.00 0.00 0.50 1.00	ecall 0.80 0.00 0.40 0.80	61-score 9.89 9.00 9.80 9.44 9.89	2000 0 2000 2000 2000 2000	st)
	p 0 1 accuracy macro avg weighted avg	precision 0 1.00 1 0.00 accuracy macro avg 0.50 weighted avg 1.00	precision recall 0 1.00 0.80 1 0.00 0.00 accuracy are avg 0.50 0.40 weighted avg 1.00 0.80	precision recall f1-score	0 1.00 0.80 0.89 2000 1 0.00 0.00 0.00 0 0 0 0 0 0 0 0 0 0

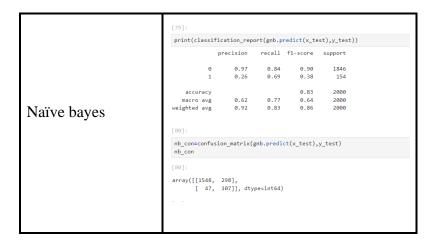




	[57]:						
	print(classific					t))	
	pr	recision					
	1	0.18	0.82	0.89 0.27	125		
	accuracy macro avg	p 57	p 70	0.81 0.58			
Logistic Regression	weighted avg	0.92	0.81	0.85	2000		
	[58]:						
	confusion_matri	x(model.pr	edict(x_te	est),y_test	t)		
	[58]:						
	array([[1542, 3 [53,	333], 72]], dtyp	e=int64)				
	[59]:						
	[61]:						
	print(classi						
		precis	ion r	recall i	f1-score	support	
	0	-	.85 .51	0.87 0.47	0.86 0.49	1562 438	
				0.4/			
	accuracy macro avg	0		0.67	0.78 0.67	2000 2000	
Decision Tree	weighted avg		.78	0.78	0.78	2000	
	[62]:						
	confusion_ma	atrix(pro	d,v tes	t)			
		- Sr(b)	,,_,=	,			
	[62]: array([[1362	2007					
		, 200], , 205]]		int64)			
	5003						
	1						
	[66]:						
	rfc_con=confu	usion_mat	rix(pred	,y_test)			
	rfc_con						
	[66]: array([[1528,	2051					
		200]],	dtype=in	nt64)			
Random Forest	[67]:						
	print(classif						
				all f1-		upport	
	0 1	0.9 0.4).88).75	0.92 0.60	1733 267	
	accuracy	- '			0.86	2000	
		0.7 0.9	3 0).82).86	0.76 0.88	2000	
	ented avg	0.5	٤	•			
	[76]:	_	_	_	_		
	print(classifi	ication_rep precision)	
	0	0.94	0.87	0.90	1728		
TZNI ' ''	1	0.43	0.64	0.51	272		
KNeighbors	accuracy macro avg weighted avg	0.68 0.87	0.75 0.83	0.83 0.71 0.85	2000 2000 2000		
C1 101	weighted avg	0.87	e.83	0.85	2000		
Classifier	[77]: knn_con=confus	ion ==+-	(knp nr- *-	ct(x +	.v tert)		
	knn_con=confus knn_con	matrix	,predi	test)	,,_vest)		
Ī	[77]:						
	array([[1496, 232], [99, 173]], dtype=int64)						
	array([[1496, [99,	173]], dty	pe=int64)				
	array([[1496, [99,	173]], dty	pe=int64)				
	array([[1496, [99,	173]], dty	pe=int64)				_ "







Final Model Selection Justification (2 Marks):

Final Model	Reasoning
	Random Forest is favored for telecom churn prediction due to its high accuracy with complex, feature-rich datasets. It excels in capturing non-linear relationships and interactions while mitigating overfitting through ensemble learning. Feature importance ranking aids in
Random Forest Classifier	identifying key predictors, and its robustness against data imbalance makes it ideal for detecting churn patterns in telecom customer data.