

# **Dynamic Web Performance Optimization Measurement**

Using Machine Learning Analytics — Report

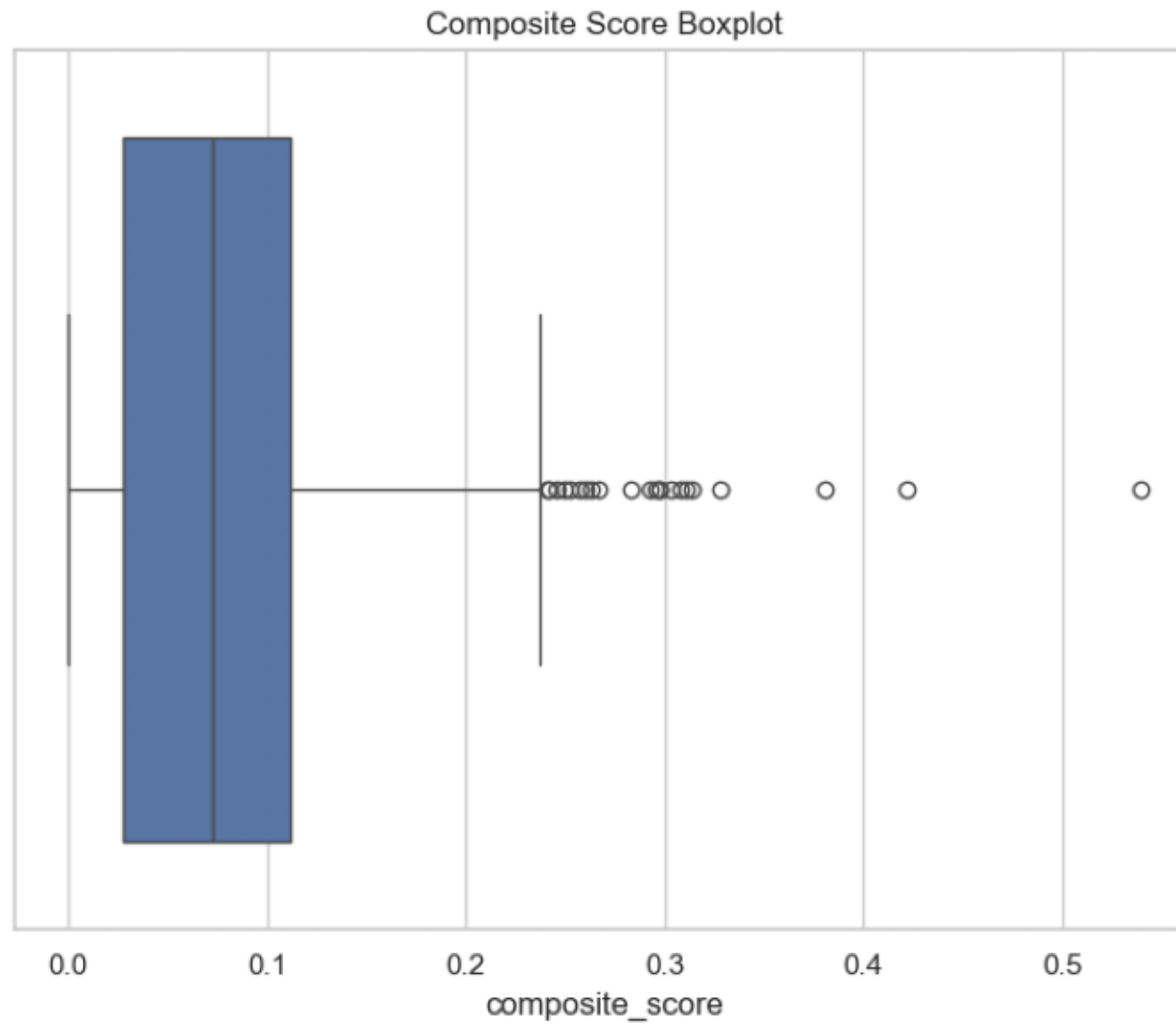
Author: (generated)

Data: D:\Thesis\My Thesis\All thesis data - set4.cleaned.imputed.csv

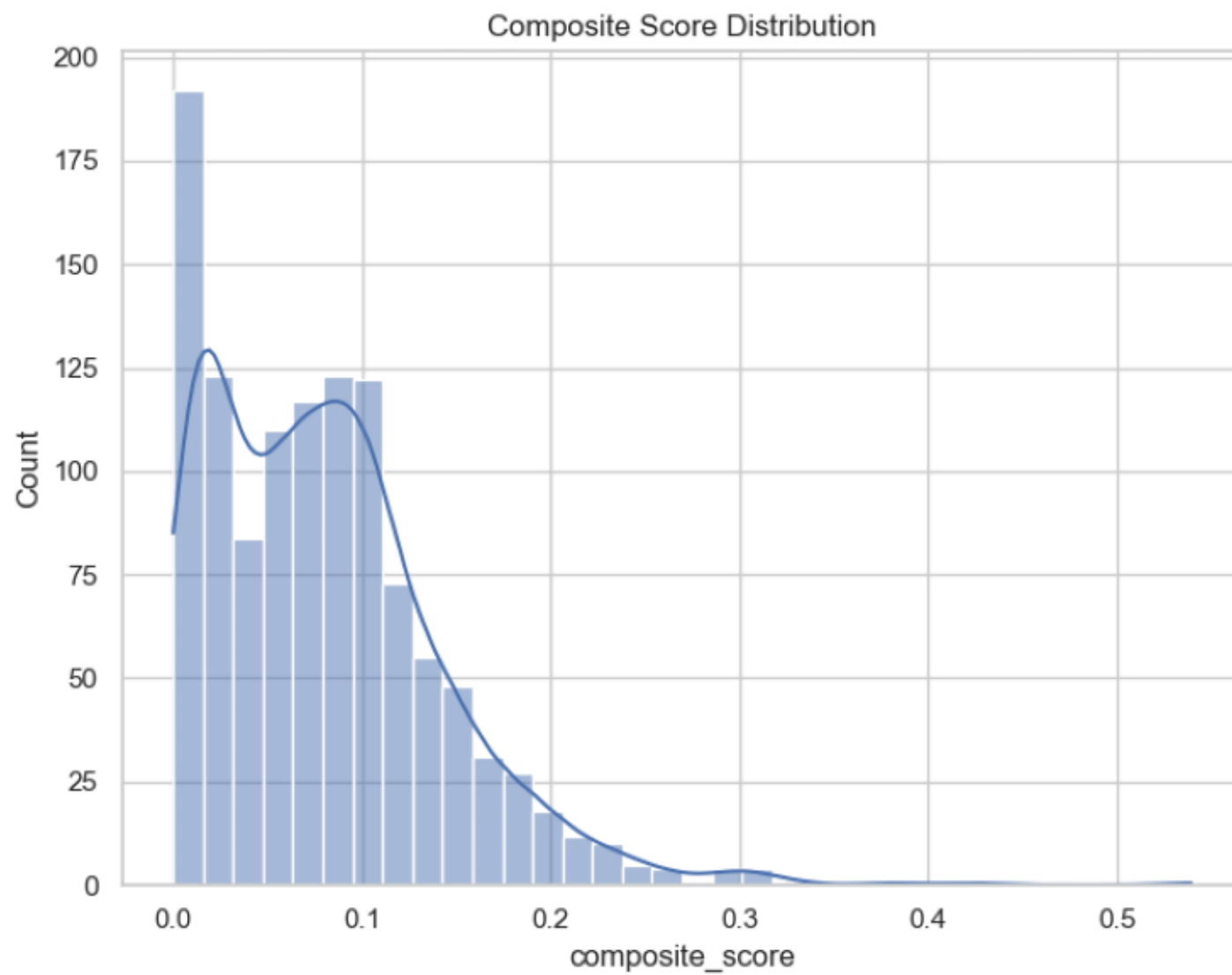
## Dataset Basic Statistics

	Unnamed: 0	count	unique	top	freq	mean	std	min	25%	50%	75%	max
	0	url	1167.0	1161https://qualcomm	3.0	nan	nan	nan	nan	nan	nan	nan
	1	Response_time_	1167.0	nan	nan	370.97	1192.173	0.0	24.0	126.0	357.5	25080.0
	2	Load_time_m	1167.0	nan	nan	4088.651	4025.015	76.0	1675.5	3308.0	5335.57	49848.0
DOM_	3	Content_Loaded	1167.0	nan	nan	2489.925	2974.745	61.0	928.0	1713.0	3047.5	35214.0
	4	First_byte_TTFB_	1167.0	nan	nan	346.755	1161.842	0.0	27.5	112.0	322.5	23972.0
	5	Total_links	1167.0	nan	nan	1039.256	2298.412	0.0	75.0	188.0	507.5	23613.0
	6	No_of_request	1167.0	nan	nan	572.236	1298.939	0.0	41.0	92.0	239.0	14600.0
	7	Byte_in_bytes	1167.0	nan	nan	1935125.727	4274507.754	0.0	269.0	597536.0	2189798.0	71211927.0
	8	Page_size_MB	1167.0	nan	nan	39.529	98.963	0.0	0.306	1.648	6.954	1051.0
Largest	9	Contentful_paint	1167.0	nan	nan	10351.592	15011.03	0.0	92.5	7738.578	15151.973	330536.188
Cumula	10	Active_Layout_S	1167.0	nan	nan	776866.249	8012336.486	0.0	0.001	0.101	6090.69	36990595.0
First_C	11	Contentful_Paint	1167.0	nan	nan	4093.012	3928.883	0.0	258.5	3736.211	5719.513	34944.798
Time	12	Time_to_interactive	1167.0	nan	nan	22528.647	17929.198	826.114	11791.618	18229.071	28389.835	334797.238
	13	Speed_Index_r	1167.0	nan	nan	6710.084	6739.612	0.0	0.078	6324.049	10019.614	40008.064
Interact	14	Time_to_Next_Pai	1167.0	nan	nan	5462.525	2442.617	800.754	4438.493	4704.264	5873.05	40308.096
Design	15	Optimization	1167.0	nan	nan	9338.848	16271.072	0.0	27.0	36.0	15727.546	104991.311
JavaScript	16	Script_Execution_	1167.0	nan	nan	18545.648	16989.548	469.336	7693.369	12493.512	22914.142	135620.152
Main	17	Thread_Work_	1167.0	nan	nan	7079.624	8355.718	0.0	1884.848	4100.646	9893.053	78840.325
CS	18	Blocking_Tim	1167.0	nan	nan	1161.931	2021.232	0.0	26.0	269.0	1654.5	27764.0
Broken	19	Broken_link_cou	1167.0	nan	nan	8481.321	16921.347	0.0	1.0	15.0	11304.818	129881.424
St	20	render_time	1167.0	nan	nan	7023.208	7378.517	0.0	3337.853	5004.288	7768.837	79090.088
Docum	21	ent_complete_	1167.0	nan	nan	3018.869	3687.513	0.0	867.0	2123.0	3947.5	49848.0

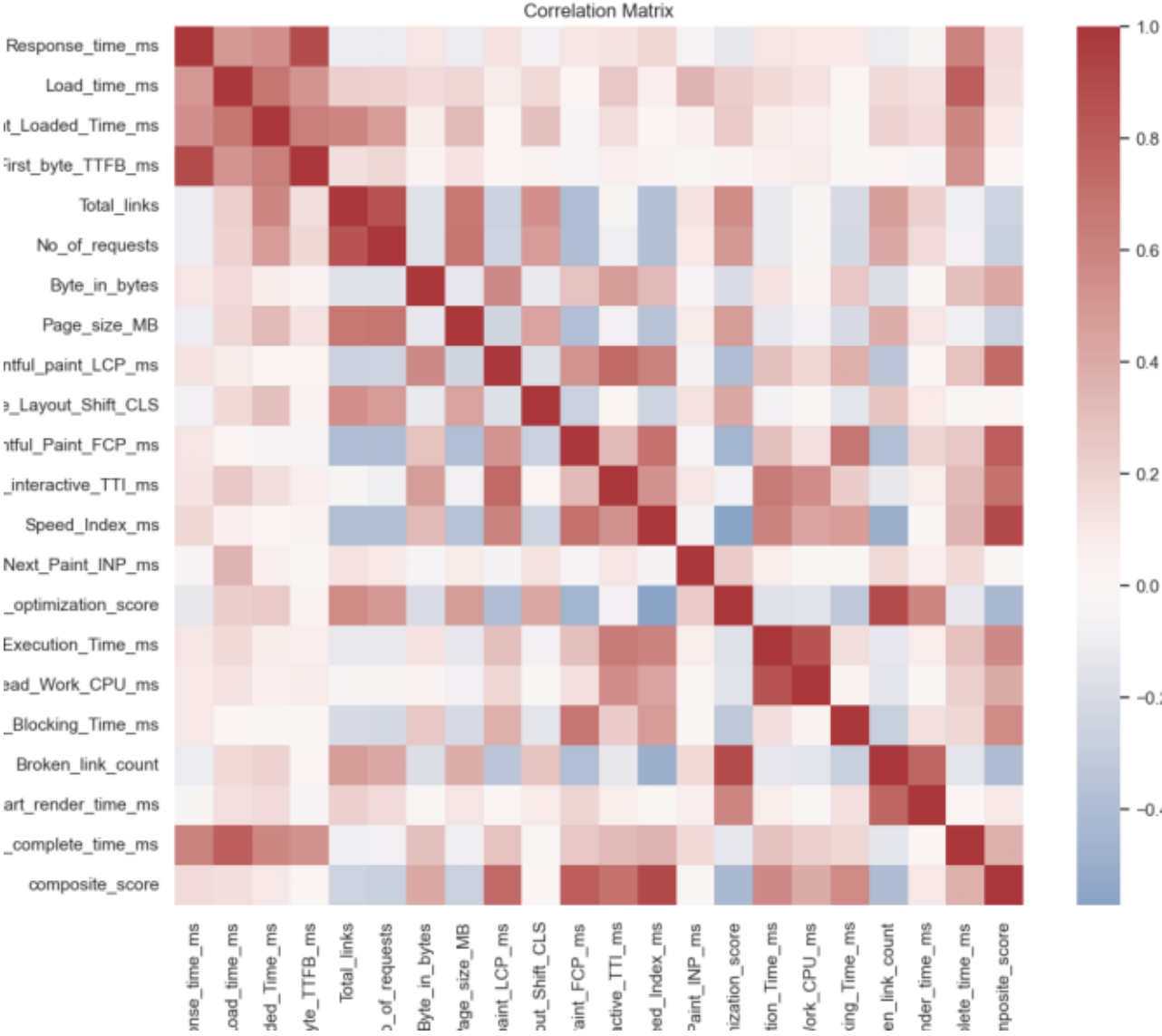
# Composite Boxplot



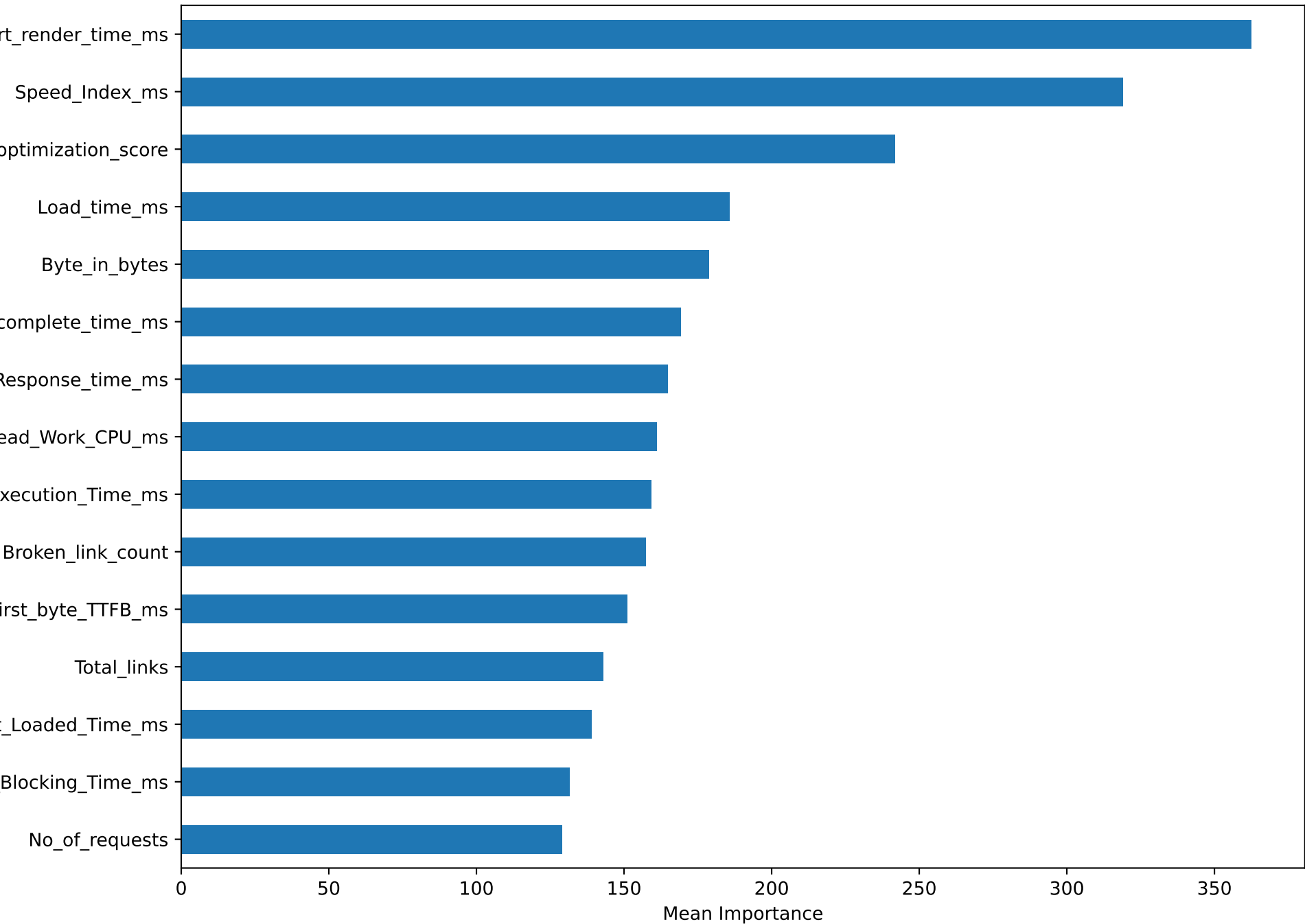
# Composite Hist



# Correlation Matrix



Top 15 Features by Mean Importance



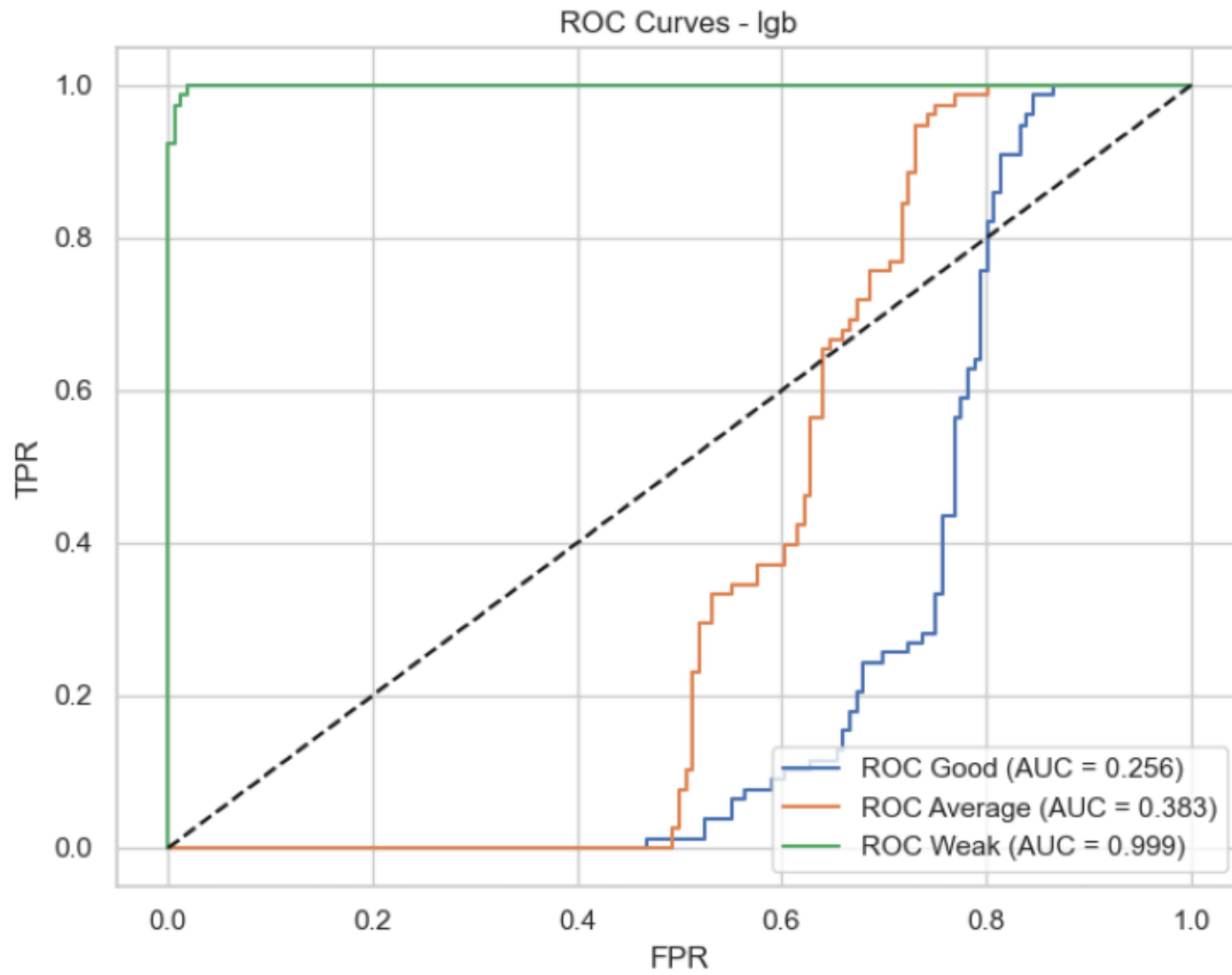
Model	MAE	RMSE
rf	185807.42	1277758.89
lgbm	190320.49	1085288.86
keras	204530.30	1392463.05

## Classifier F1 Macro Scores

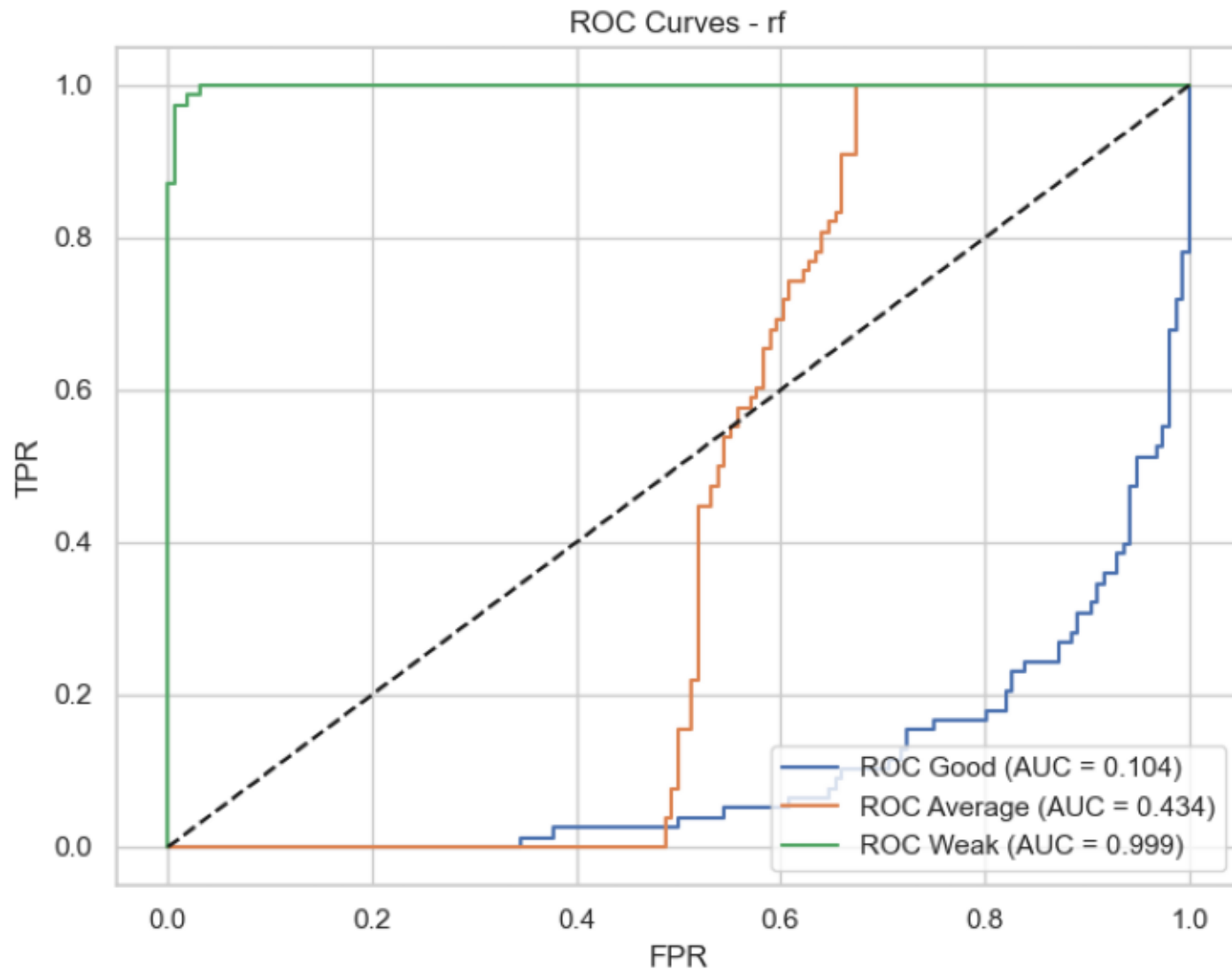
strategy	model	f1_macro
label_tertiles	rf	0.9572
label_tertiles	lgbm	0.9571
label_tertiles	keras	0.9699
label_weighted	rf	0.9438
label_weighted	lgbm	0.9485
label_weighted	keras	0.9658
label_kmeans	rf	0.9755
label_kmeans	lgbm	0.9847
label_kmeans	keras	0.9021



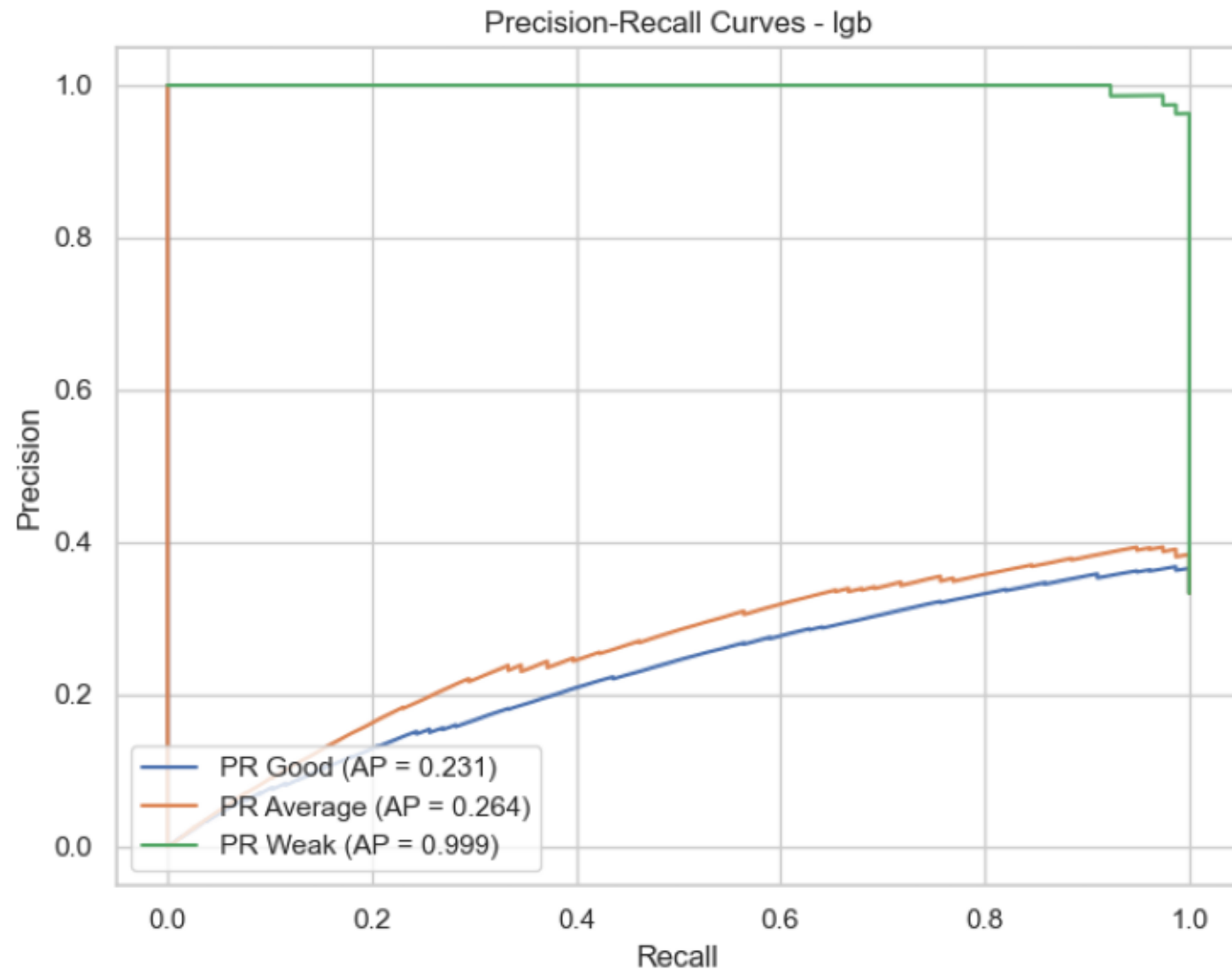
roc\_lgb.png

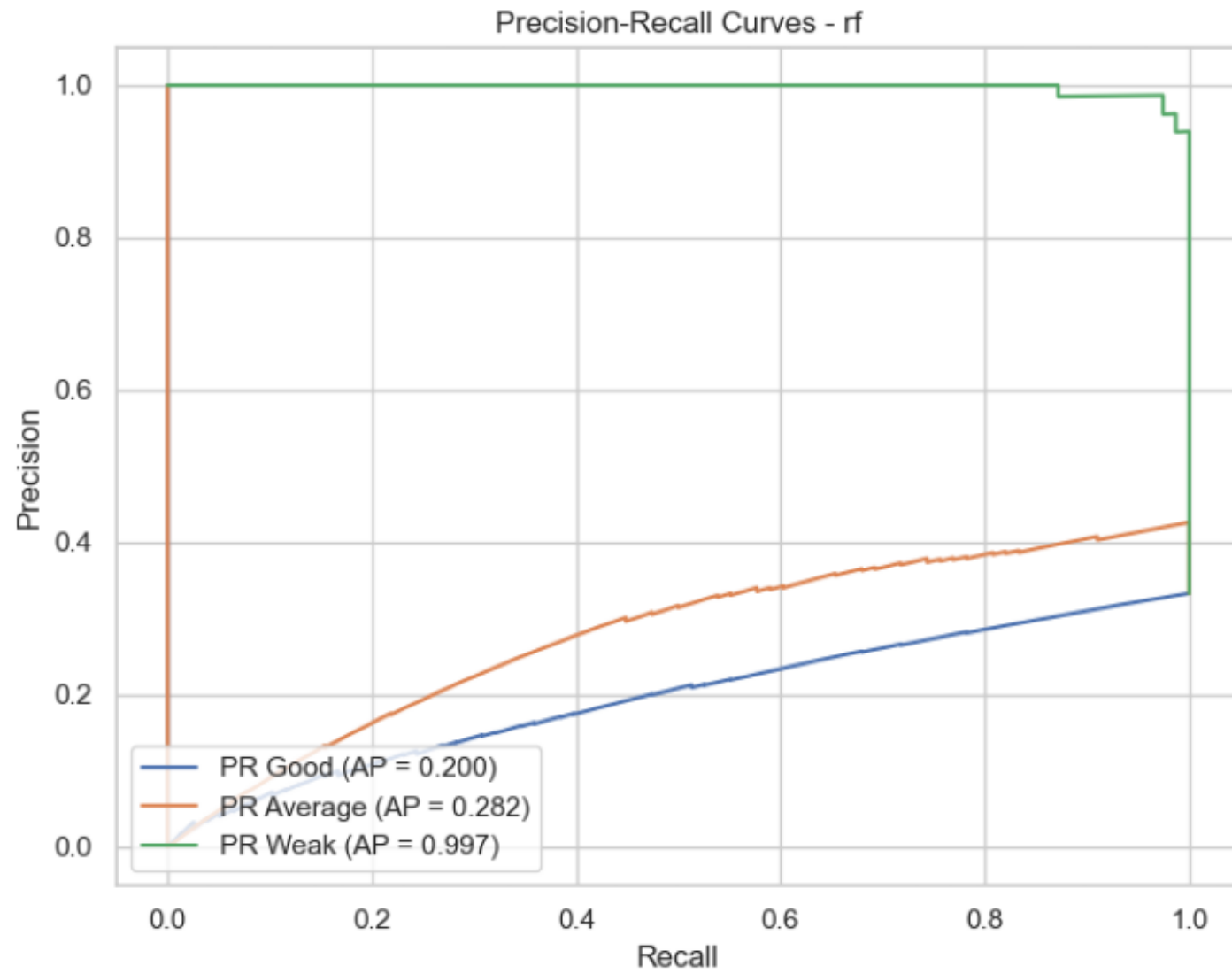


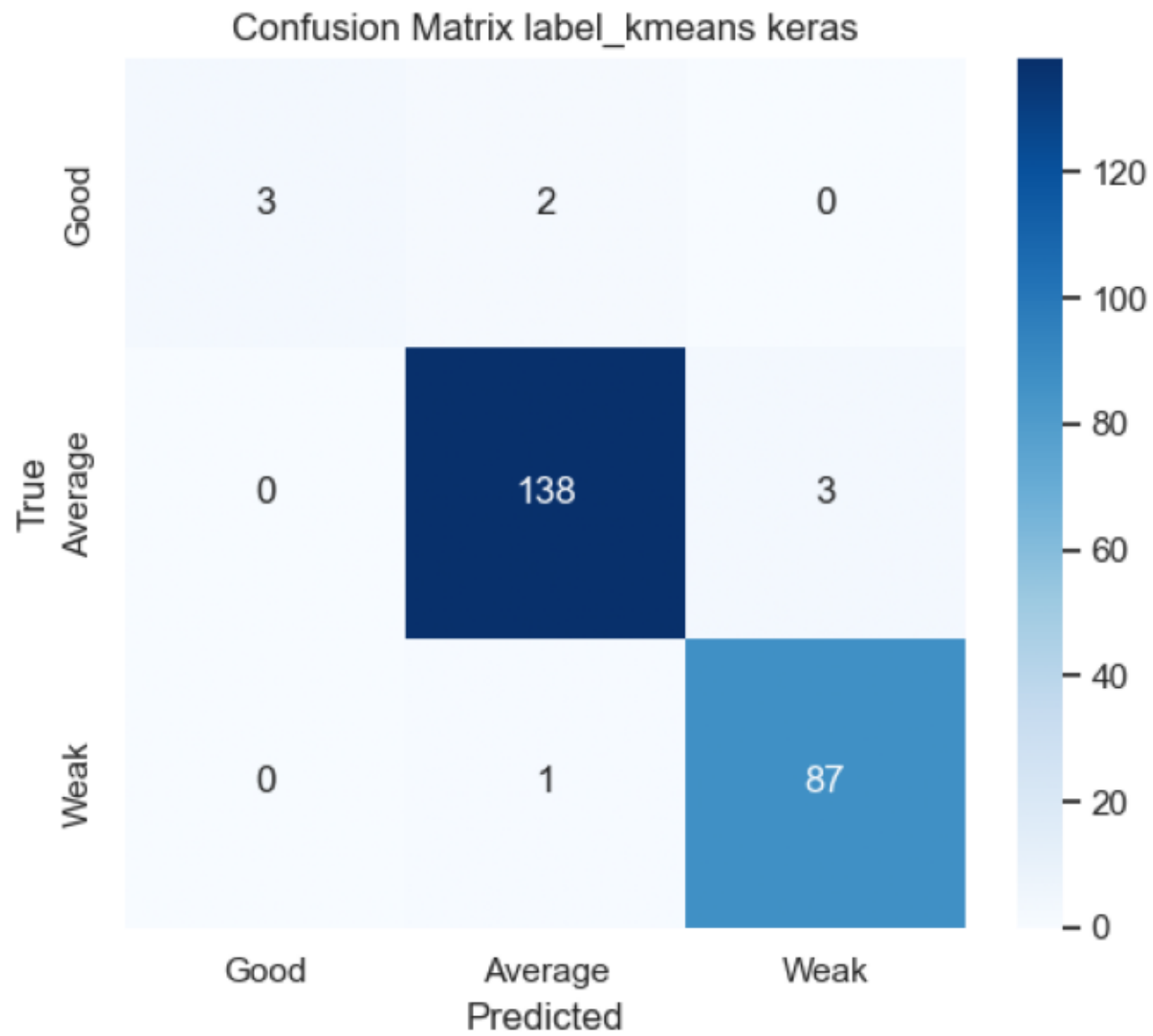
roc\_rf.png

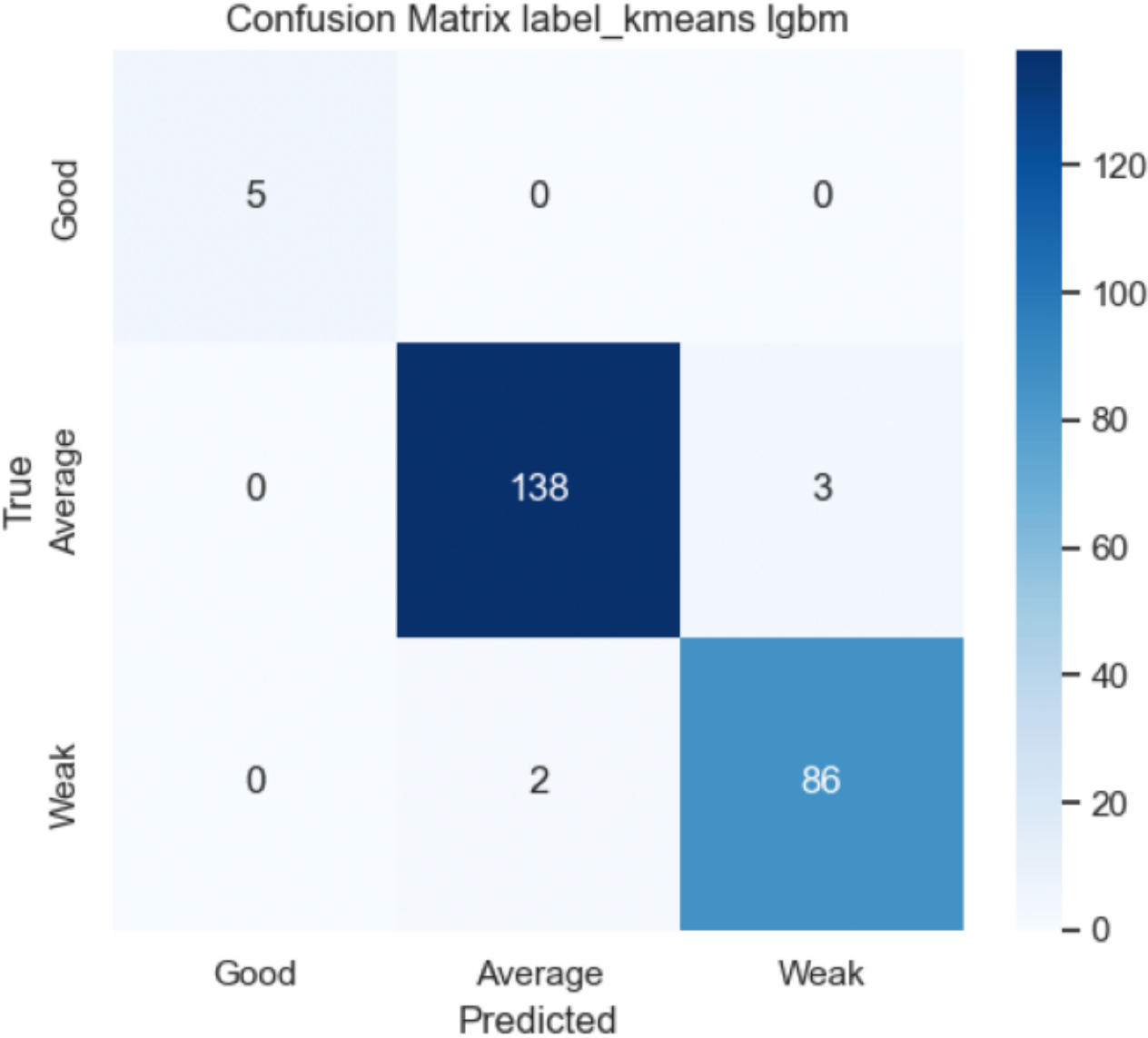


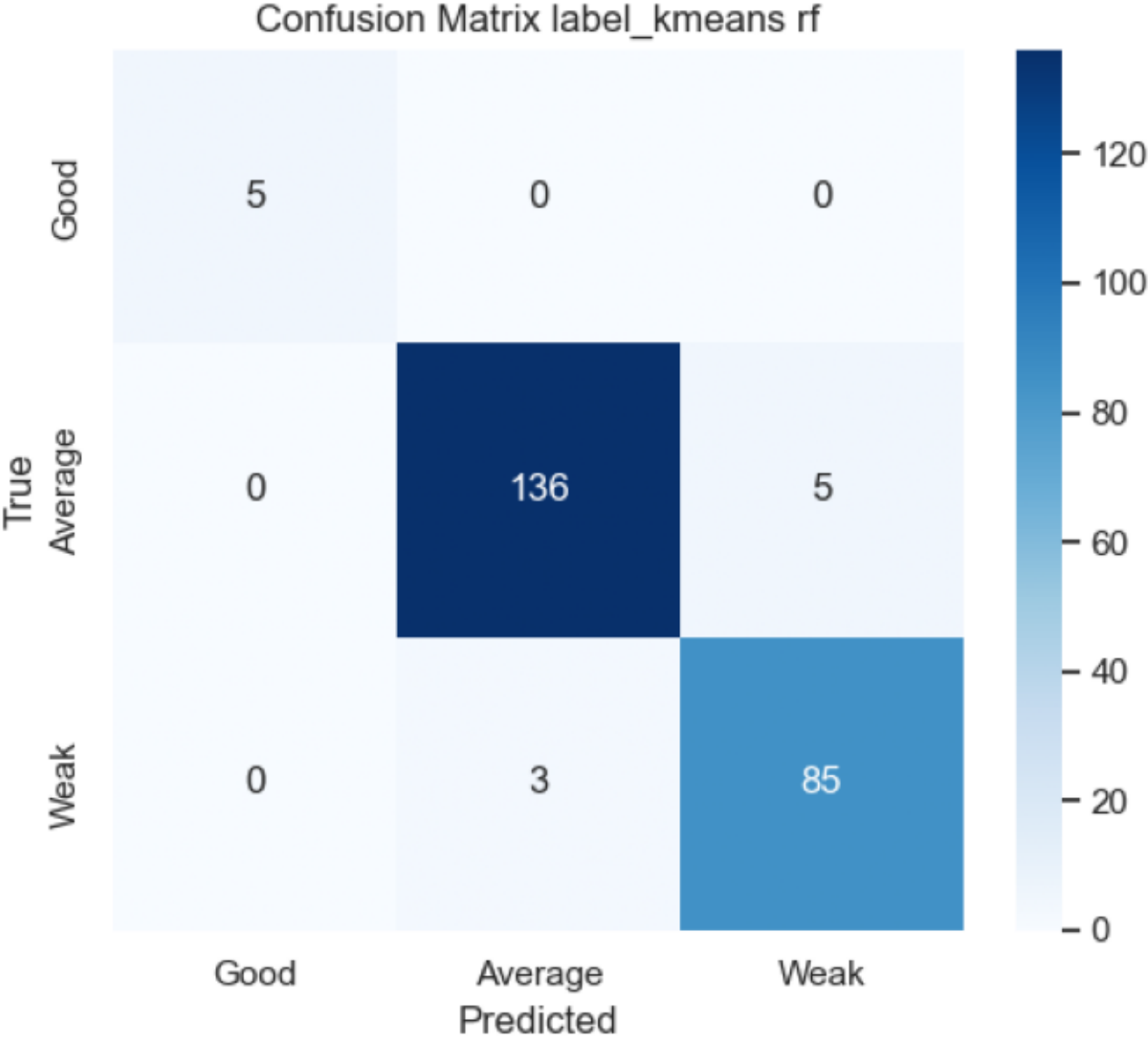
pr\_lgb.png

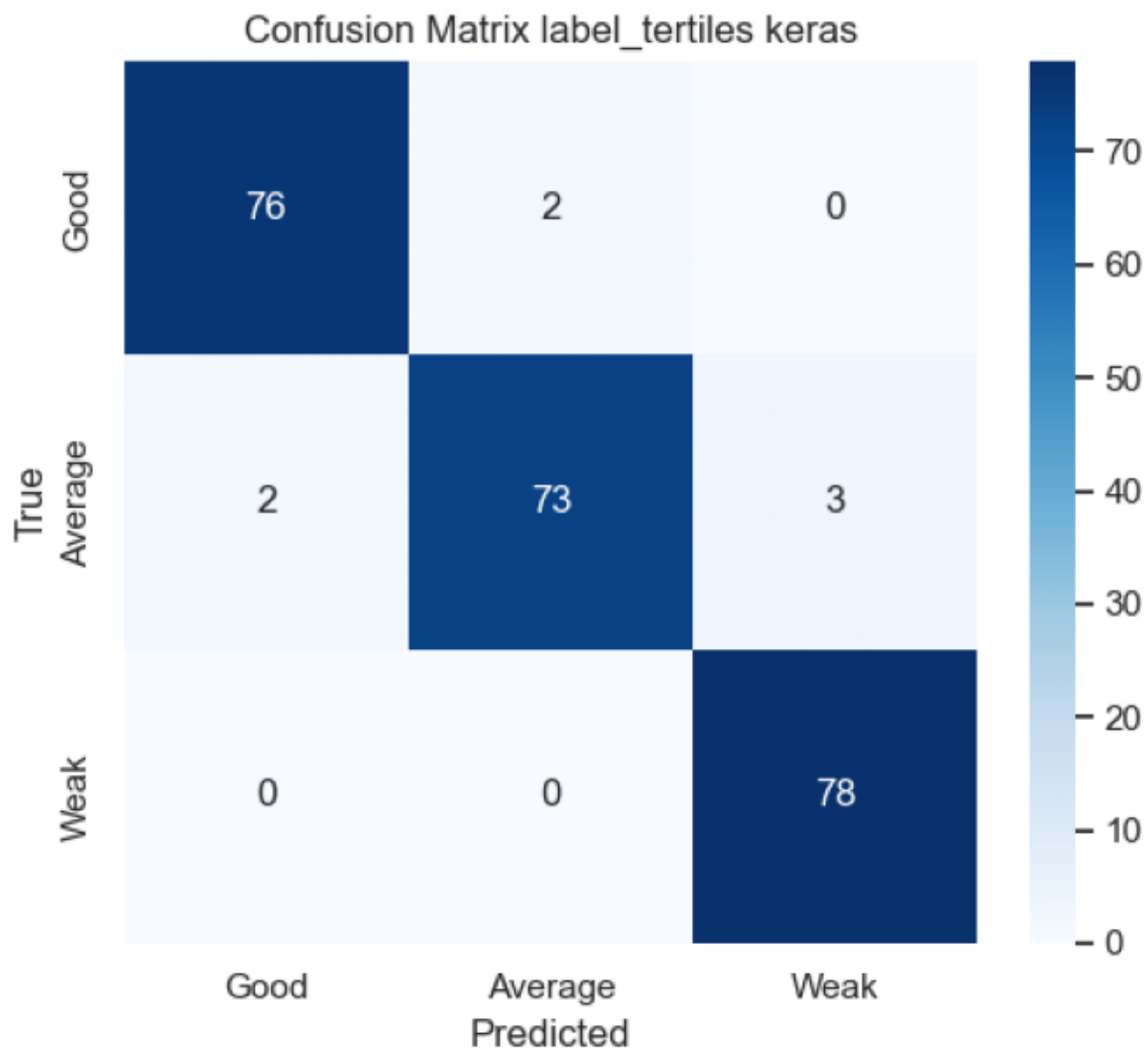




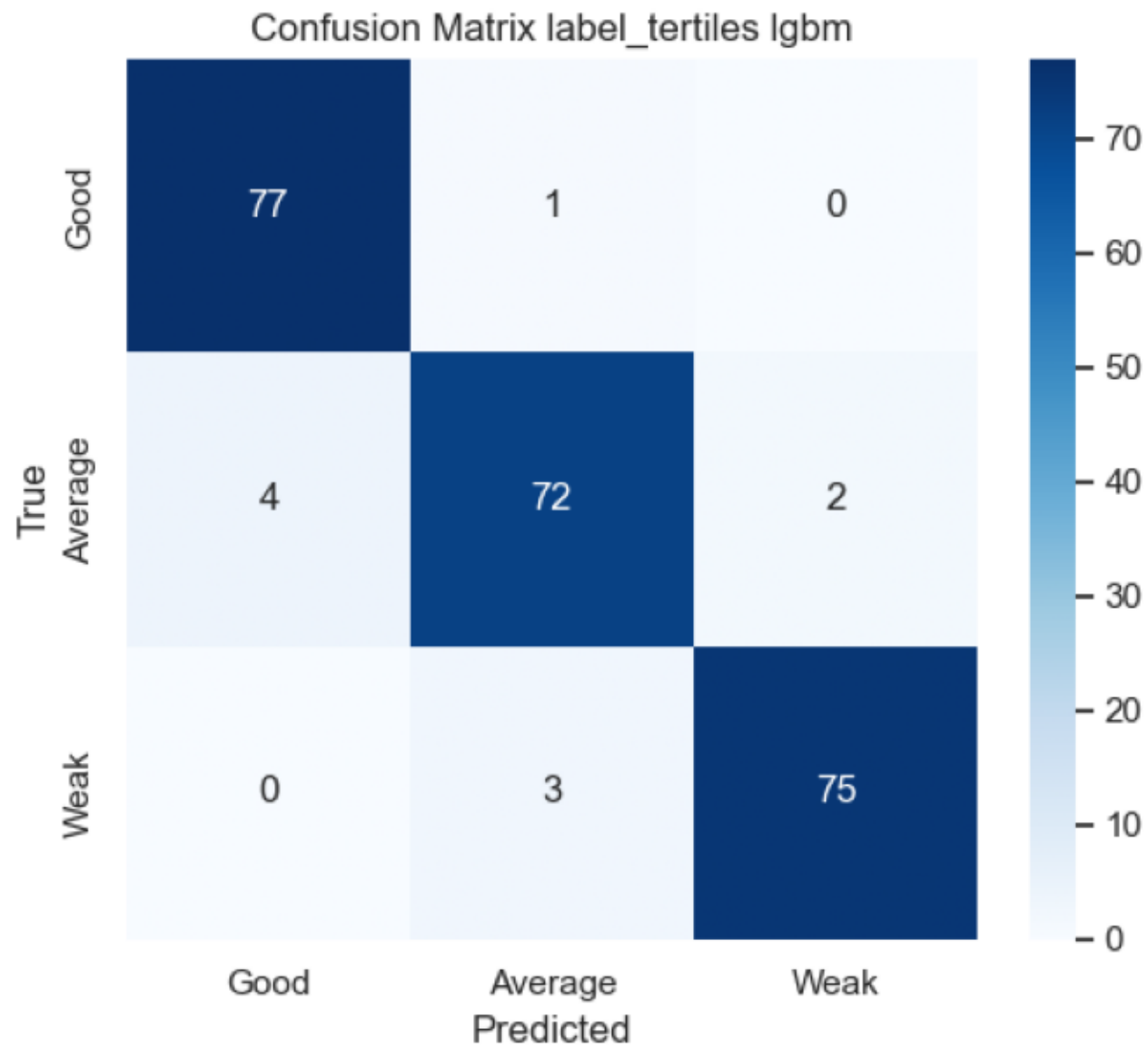


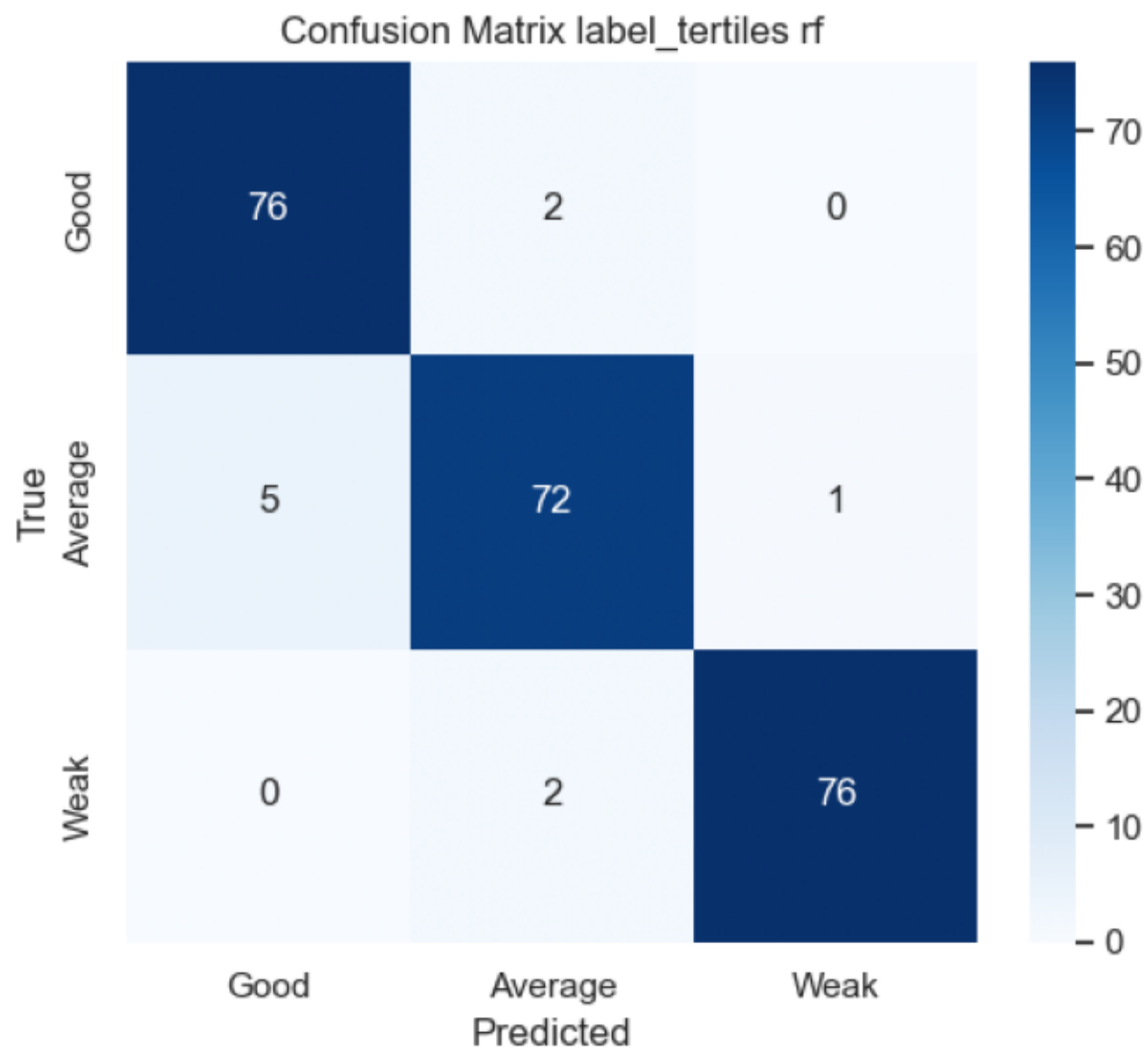


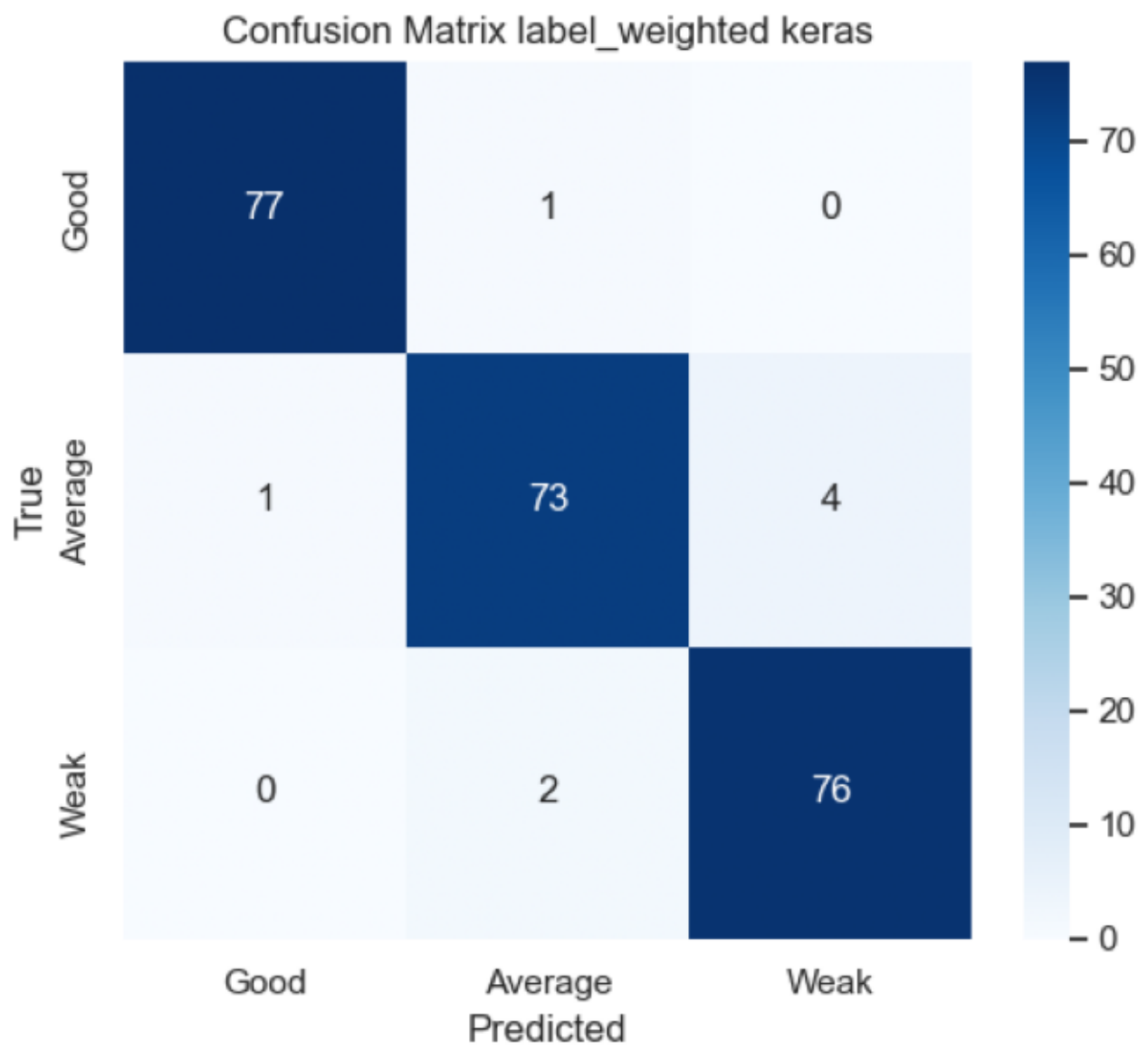


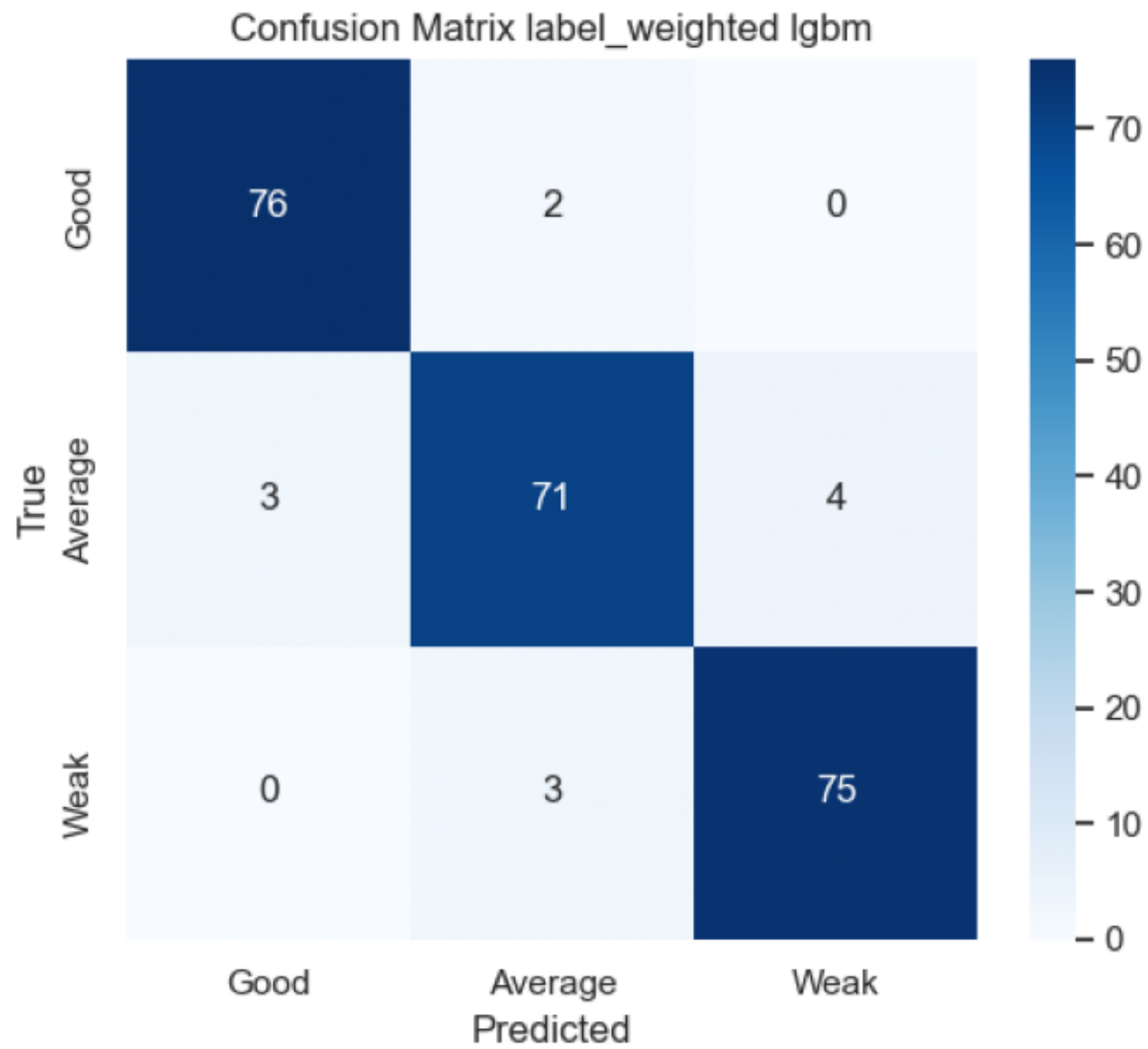


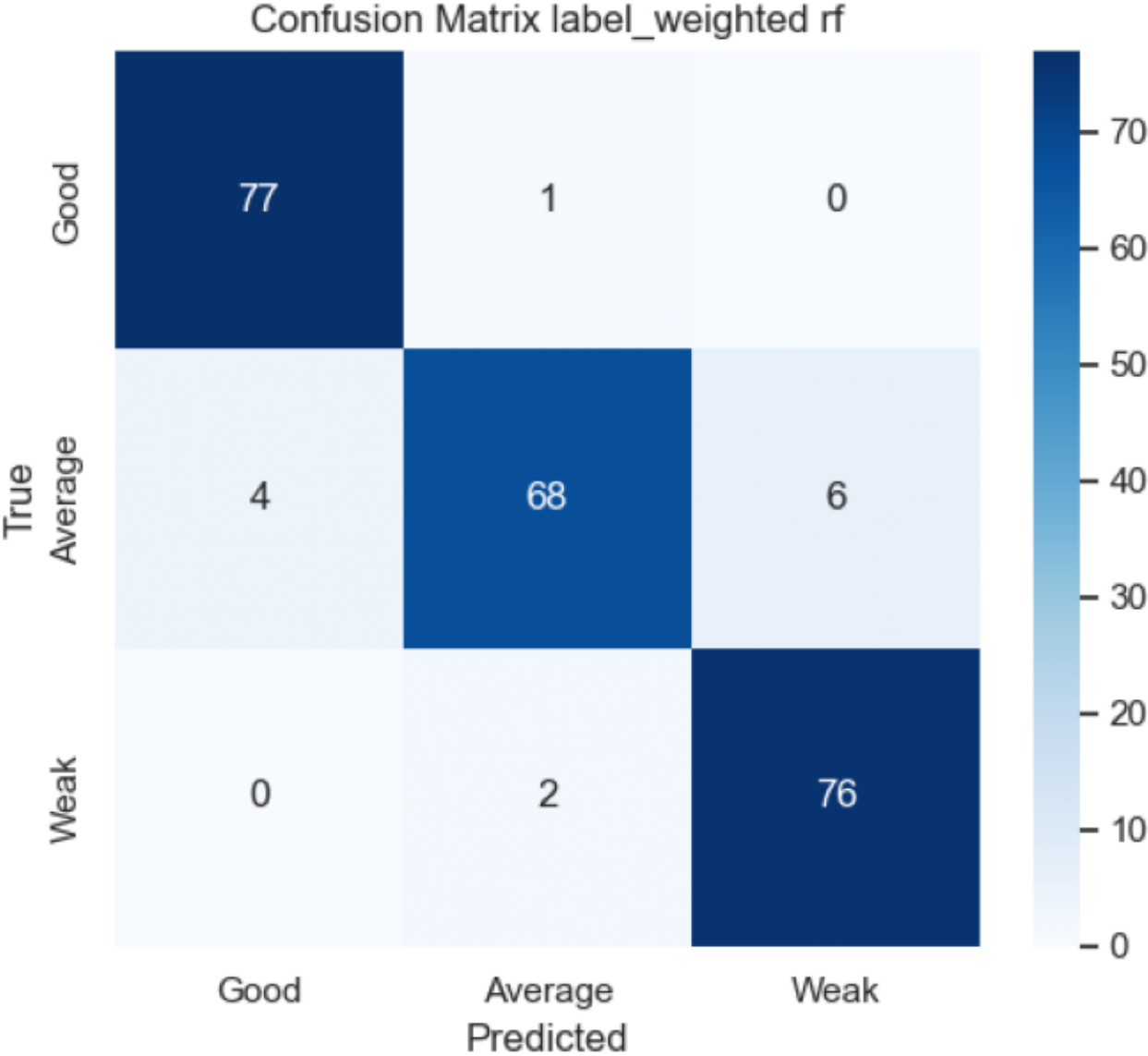




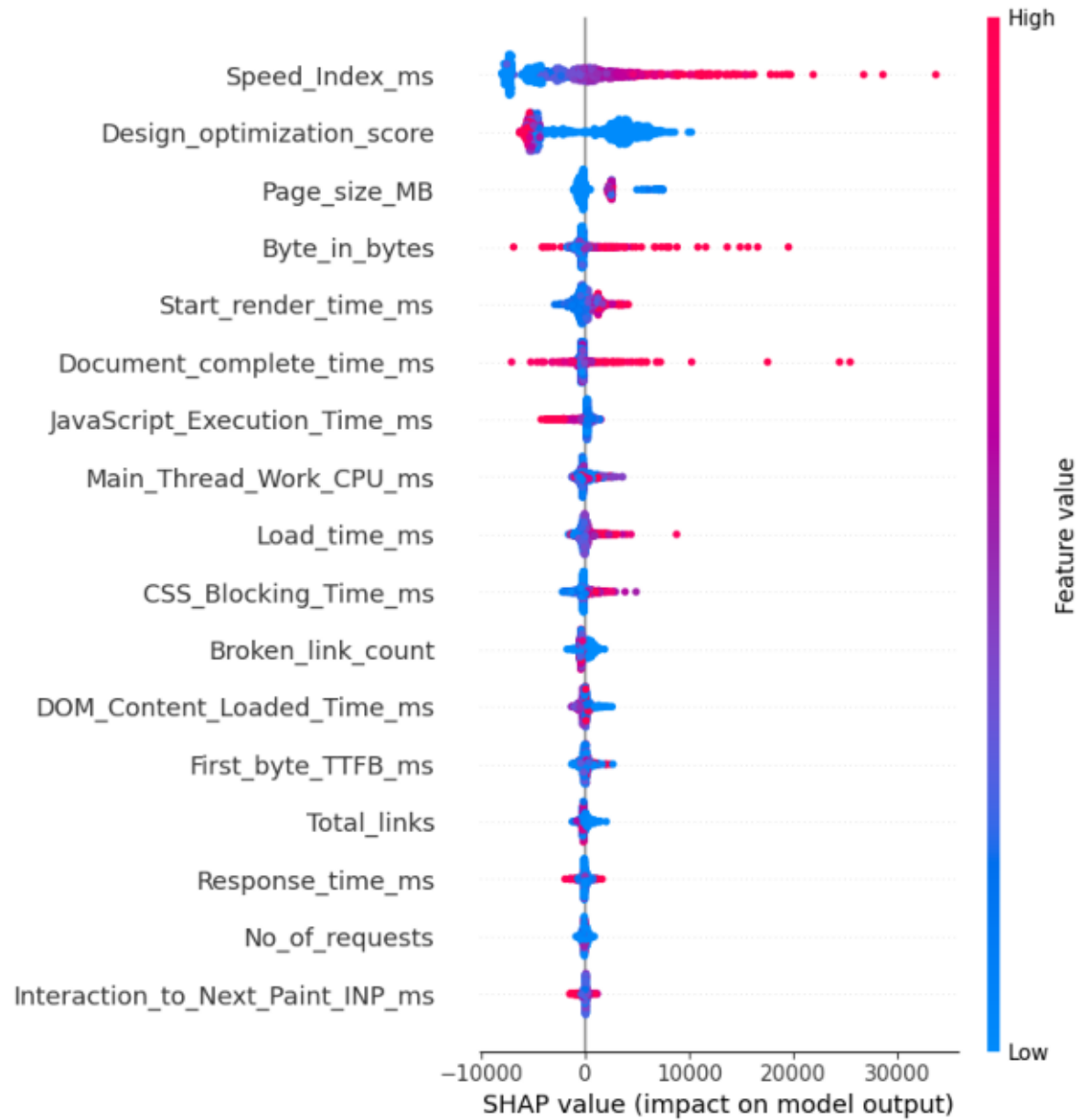








## SHAP summary (example)



## **Conclusions:**

- K-means labeling produced best classifier f1 (see classifiers tuning).
- LightGBM and Keras performed well; check per-target regression metrics for LCP/FCP/TTI.
- SHAP explanations available via API /explain to identify feature causes and suggested fixes.