Gene-Level Performance Across Datasets

(a) auPRC

Model	Train	Val	Test
xgb_new	0.95	0.68	0.83
deep_residual	0.88	0.81	0.81
xgb_tuned	0.97	0.97	0.81
optimized_residual	0.91	0.83	0.79
baseline_mlp	0.92	0.80	0.77
dgit_super	0.89	0.83	0.73
xgb_paper	0.87	0.64	0.71
transformer	0.84	0.69	—

(b) AUC

Model	Train	Val	Test
xgb_new	1.00	0.99	1.00
deep_residual	1.00	0.96	1.00
xgb_tuned	1.00	1.00	1.00
optimized_residual	1.00	0.97	1.00
baseline_mlp	1.00	0.95	0.99
dgit_super	1.00	0.99	0.99
xgb_paper	1.00	0.99	0.96
transformer	1.00	0.99	0.99

(c) Precision

Model	Train	Val	Test
xgb_new	0.80	0.64	0.65
deep_residual	0.96	0.99	0.93
xgb_tuned	0.81	0.82	0.79
optimized_residual	0.92	0.90	0.81
baseline_mlp	0.96	0.96	0.98
dgit_super	0.97	0.96	0.97
xgb_paper	0.92	0.68	0.86
transformer	0.94	0.93	0.89

(d) Recall

Model	Train	Val	Test
xgb_new	0.93	0.73	0.75
deep_residual	0.57	0.71	0.53
xgb_tuned	0.96	0.96	0.80
optimized_residual	0.81	0.77	0.63
baseline_mlp	0.72	0.72	0.49
dgit_super	0.66	0.70	0.52
xgb_paper	0.74	0.65	0.72
transformer	0.85	0.75	0.53

(e) F1

Model	Train	Val	Test
xgb_new	0.86	0.68	0.69
deep_residual	0.72	0.82	0.68
xgb_tuned	0.88	0.88	0.79
optimized_residual	0.86	0.83	0.71
baseline_mlp	0.82	0.82	0.65
dgit_super	0.79	0.81	0.68
xgb_paper	0.82	0.66	0.78
transformer	0.89	0.83	0.66

(f) Specificity