Model Types	Dataset	CoLA	MRPC	STS-B	RTE
Normal	Bert-base-uncased	56.82	90.24/86.27	88.29/87.96	68.23
Regularized Models	$\alpha = -1.0$	0.0	79.01/68.87	57.23/60.16	52.71
	$\alpha = -0.5$	61.42	83.17/74.02	85.28/85.22	57.04
	$\alpha = -0.1$	58.06	89.50/85.05	88.71/88.27	65.70
	$\alpha = 0.1$	58.34	90.59/86.76	88.12/87.81	64.98
	$\alpha = 0.5$	27.01	83.56/73.28	85.25/85.03	59.93
	$\alpha = 1.0$	0.0	81.22/68.38	52.07/55.60	47.29
	Local Best	61.42	90.59/86.76	88.71/88.27	65.70
Augmented Models	$g_1 / c = 0.2$	59.85	88.11/83.33	88.56/88.22	68.59
	$g_1 / c = 1$	56.51	90.88/87.01	88.96/88.60	71.12
	$g_2 / c = 0.2$	56.29	87.65/82.60	88.60/88.24	68.59
	$g_2 / c = 1$	58.85	87.74/82.60	88.68/88.32	70.40
	$g_1 \& g_2 / c = 0.2$	57.32	89.62/85.29	88.48/88.19	71.12
	$g_1 \& g_2 / c = 1$	58.30	90.40/86.52	88.83/88.45	68.95
	Local Best	59.85	90.88/87.01	88.96/88.60	71.12
Negative Models	$r=0.1~/~\beta=0.1$	56.22	88.54/83.82	88.25/87.91	65.34
	$r=0.2~/~\beta=0.1$	57.92	90.00/85.78	88.22/87.84	66.06
	$r=0.3~/~\beta=0.1$	57.92	89.31/84.80	88.26/87.90	67.15
	$r=0.1~/~\beta=0.2$	58.92	87.90/83.33	88.34/88.11	63.54
	$r=0.2~/~\beta=0.2$	57.13	87.87/83.09	88.59/88.27	64.98
	$r=0.3~/~\beta=0.2$	58.14	88.97/84.56	88.64/88.33	66.79
	Local Best	58.92	90.00/85.78	88.64/88.33	67.15
	Reg & Aug	56.56	88.54/83.82	88.86/88.60	68.59
Combined Models	$\operatorname{Reg} \& \operatorname{Neg}$	58.11	88.19/83.33	88.41/88.17	69.31
	Aug & Neg	59.07	90.49/86.76	88.59/88.21	70.76
	Reg & Aug & Neg	58.92	88.39/83.58	88.32/88.01	67.87
	Local Best	59.07	90.49/86.76	88.86/88.60	70.76
	Global Best	61.42	90.88/87.01	88.96/88.60	71.12

Table 1: Partial GLUE test results of different modifications. "Local Best" is used to display the best results for each modification type, where bolded results indicate the performance superior to the original model. "Global Best" is used to showcase the best results among all modifications. Matthews correlation, F1 scores/accuracy, Pearson/Spearman correlation, accuracy are reported for CoLA, MRPC, STS-B, RTE respectively.



Figure 1: Performance of different combinations on linear (Left), exponential (Center), and trigonometric (Right) tasks.