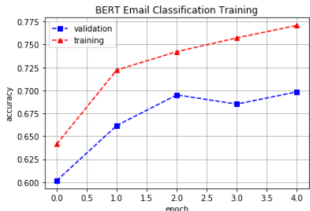
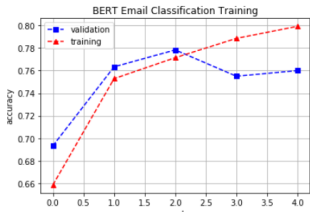
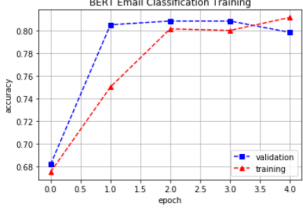
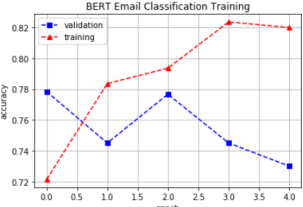


Tecnológico de Monterrey
 Inteligencia artificial avanzada para la ciencia de datos II
 Gpo 101
 BERT_2
 Saúl Francisco Vázquez del Río A01198261

Variables	BERT	Val y Loss
Nsamp = 1000, maxtokens = 50, maxtokenlen = 20		Epoch 1/5 1400/1400 [=====] - 9s 6ms/sample - loss: 0.6587 - acc: 0.6143 - val_loss: 0.5910 - val_acc: 0.6967 Epoch 2/5 1400/1400 [=====] - 4s 3ms/sample - loss: 0.5502 - acc: 0.7264 - val_loss: 0.5560 - val_acc: 0.7017 Epoch 3/5 1400/1400 [=====] - 4s 3ms/sample - loss: 0.5161 - acc: 0.7500 - val_loss: 0.5563 - val_acc: 0.7017 Epoch 4/5 1400/1400 [=====] - 4s 3ms/sample - loss: 0.5063 - acc: 0.7521 - val_loss: 0.5657 - val_acc: 0.7067 Epoch 5/5 1400/1400 [=====] - 4s 3ms/sample - loss: 0.4836 - acc: 0.7686 - val_loss: 0.5501 - val_acc: 0.7133
Nsamp = 1000, maxtokens = 100, maxtokenlen = 100		Epoch 1/5 1400/1400 [=====] - 11s 8ms/sample - loss: 0.5964 - acc: 0.6771 - val_loss: 0.5455 - val_acc: 0.7083 Epoch 2/5 1400/1400 [=====] - 7s 5ms/sample - loss: 0.4989 - acc: 0.7586 - val_loss: 0.4956 - val_acc: 0.7417 Epoch 3/5 1400/1400 [=====] - 7s 5ms/sample - loss: 0.4682 - acc: 0.7864 - val_loss: 0.5276 - val_acc: 0.7267 Epoch 4/5 1400/1400 [=====] - 7s 5ms/sample - loss: 0.4479 - acc: 0.7929 - val_loss: 0.4968 - val_acc: 0.7567

		<p>Epoch 5/5 1400/1400</p> <p>[=====]</p> <p>- 7s 5ms/sample - loss: 0.4377 - acc: 0.7929 - val_loss: 0.5122 - val_acc: 0.7533</p>
<p>Nsamp = 1000, maxtokens = 200, maxtokenken = 200</p>	 <p>The graph shows accuracy on the y-axis (0.68 to 0.80) and epoch on the x-axis (0.0 to 4.0). The training set (red dashed line with triangles) starts at ~0.68 and rises to ~0.80. The validation set (blue dashed line with squares) starts at ~0.68, peaks at epoch 1 (~0.80), and then fluctuates between 0.78 and 0.80.</p>	<p>Epoch 1/5 1400/1400</p> <p>[=====]</p> <p>- 18s 13ms/sample - loss: 0.5890 - acc: 0.6771 - val_loss: 0.5050 - val_acc: 0.7467</p> <p>Epoch 2/5 1400/1400</p> <p>[=====]</p> <p>- 15s 11ms/sample - loss: 0.4616 - acc: 0.7929 - val_loss: 0.5151 - val_acc: 0.7317</p> <p>Epoch 3/5 1400/1400</p> <p>[=====]</p> <p>- 15s 11ms/sample - loss: 0.4200 - acc: 0.8193 - val_loss: 0.4736 - val_acc: 0.7600</p> <p>Epoch 4/5 1400/1400</p> <p>[=====]</p> <p>- 15s 11ms/sample - loss: 0.3897 - acc: 0.8300 - val_loss: 0.4576 - val_acc: 0.7933</p> <p>Epoch 5/5 1400/1400</p> <p>[=====]</p> <p>- 15s 11ms/sample - loss: 0.3763 - acc: 0.8321 - val_loss: 0.4698 - val_acc: 0.7733</p>
<p>Nsamp = 1000, maxtokens = 230, maxtokenken = 200</p>	 <p>The graph shows accuracy on the y-axis (0.72 to 0.82) and epoch on the x-axis (0.0 to 4.0). The training set (red dashed line with triangles) starts at ~0.72 and rises to ~0.82. The validation set (blue dashed line with squares) starts at ~0.78, dips to epoch 1 (~0.74), and then fluctuates between 0.74 and 0.78.</p>	<p>Epoch 1/5 1400/1400</p> <p>[=====]</p> <p>- 24s 17ms/sample - loss: 0.5694 - acc: 0.7214 - val_loss: 0.4884 - val_acc: 0.7783</p> <p>Epoch 2/5 1400/1400</p> <p>[=====]</p> <p>- 18s 13ms/sample - loss: 0.4543 - acc: 0.7836 - val_loss: 0.4882 - val_acc: 0.7450</p> <p>Epoch 3/5 1400/1400</p> <p>[=====]</p> <p>- 18s 13ms/sample - loss: 0.4383 - acc: 0.7936 - val_loss: 0.4648 - val_acc: 0.7767</p> <p>Epoch 4/5 1400/1400</p> <p>[=====]</p> <p>- 18s 13ms/sample - loss: 0.3986 - acc: 0.8236 - val_loss: 0.5248 - val_acc: 0.7450</p> <p>Epoch 5/5 1400/1400</p> <p>[=====]</p> <p>- 18s 13ms/sample - loss: 0.4071 - acc:</p>

		0.8200 - val_loss: 0.5801 - val_acc: 0.7300
--	--	---

Datos pasados

Bow				
	LR	SVM	RF	XGBoost
A)	0.6666	0.67	0.67	0.6316
B)	0.7416	0.7316	0.7366	0.7166
C)	0.815	0.7933	0.7733	0.7516

TF-JDF				
	LR	SVM	RF	XGBoost
A)	0.6983	0.69	0.6666	0.64
B)	0.7316	0.7283	0.72	0.6983
C)	0.7966	0.7983	0.7433	0.78