Model: bert-base-uncased
Epochs: 5
Max token length: 512 (7 sentences per sample)

Sliding window evaluation:
Accuracy: 85.72%
F1-Score: 79.2%

	Precision	Recall	F1-score	support
0	0.88240	0.90033	0.89127	99748
1	0.80751	0.77701	0.79197	53676
accuracy			0.85719	153424
Macro avg	0.84495	0.83867	0.84162	153424
Weighted avg	0.85620	0.85719	0.85653	153424

Evaluation without sliding window:
Accuracy: 82.59%
F1-Score: 74.38%

	Precision	Recall	F1-score	support
0	0.85498	0.88186	0.86821	99748
1	0.76684	0.72204	0.74376	53676
accuracy			0.82595	153424
Macro avg	0.81091	0.80195	0.80599	153424
Weighted avg	0.82414	0.82595	0.82467	153424