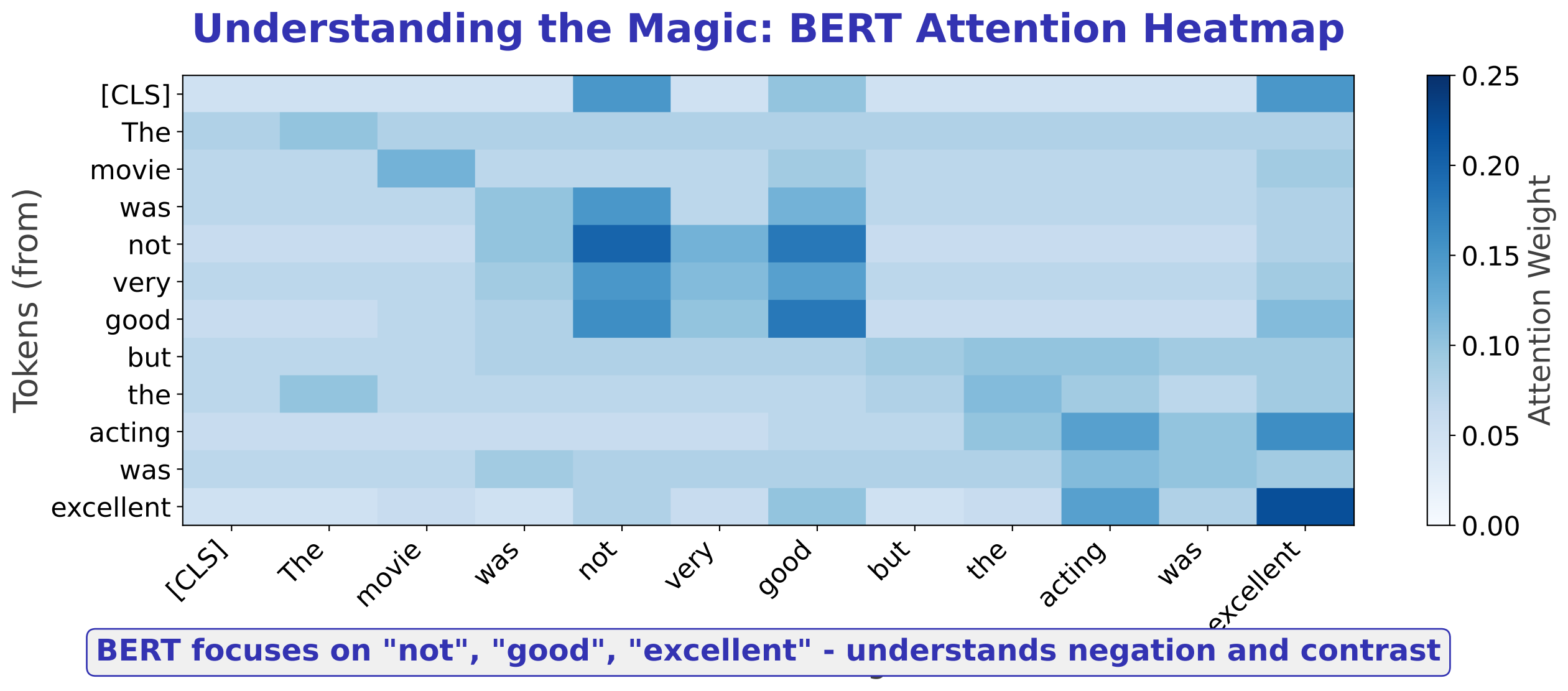
Understanding the Magic: BERT Attention Heatmap
Tokens (from)
[CLS]
The
movie
was
not
very
good
but
the
acting
was
excellent
Attention Weight
0.25
0.20
0.15
0.10
0.05
0.00
BERT focuses on "not", "good", "excellent" - understands negation and contrast