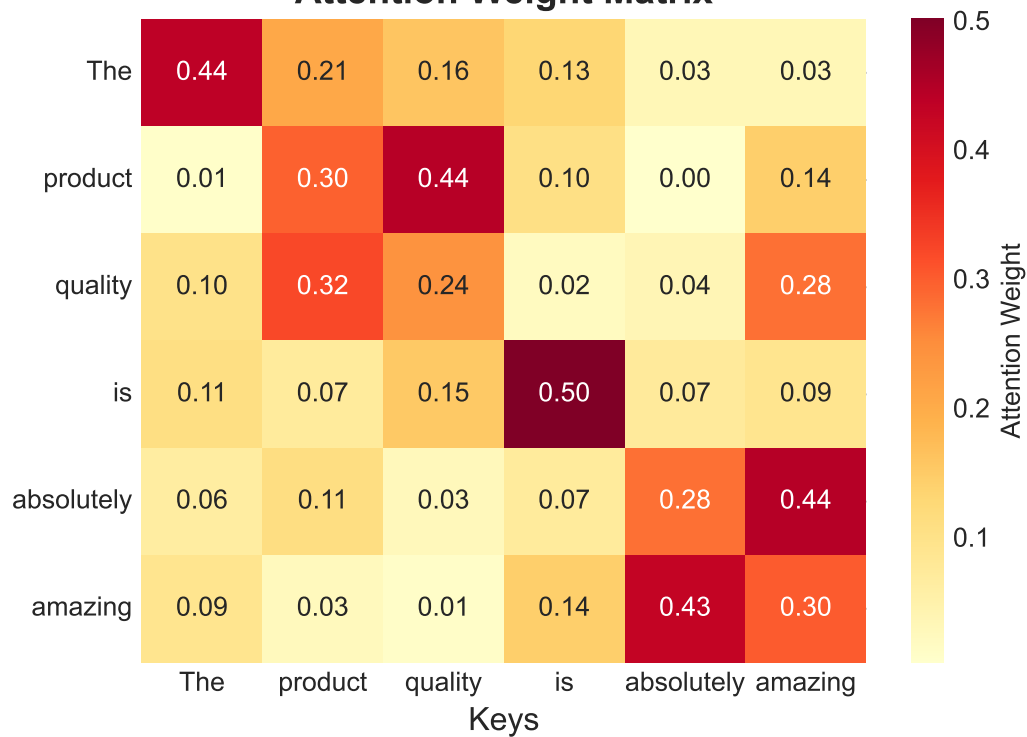


# Transformer Attention Mechanisms

## Self-Attention Mechanism in Transformers



### Attention Weight Matrix



### Multi-Head Attention

