

BERT Attention Patterns: What the Model Focuses On

