

Evolution of Sequence Modeling: From N-grams to Transformers

