

Review: Transformers

The *Transformer* is a Neural Network architecture for processing sequences.

Among the many sequences of interest to us

- Word sequences: Natural Language Processing (NLP)
- Timeseries

The Transformer is an important advance over other architectures for sequences

- Recurrent Neural Network (RNN)
- LSTM

It derives a lot of its power from the *Attention* mechanism.

Attention

- [Attention \(Intro to Attention.ipynb\)](#)
- [Implementing Attention \(Attention Lookup.ipynb\)](#) Deferred: will do at same time as reviewing code for Transformer

Advanced

- [What does BERT look at? \(Attention What Does BERT Look At.ipynb\)](#)

Transformers

- [Transformer \(Transformer.ipynb\)](#).

