**Recurrent Neural Networks (RNNs)** process data step by step, passing hidden-state information from one time step to the next. They’re widely used for any sequence data (text, audio, time series, etc.).  
The classic RNN suffers from *vanishing/exploding gradients*, so long-range dependencies are hard to learn. Variants such as **LSTMs** and **GRUs**, or more recently **attention-based models** (e.g. Transformers), were introduced to retain information over longer contexts.  
Because RNN computations are inherently sequential, they’re slower to train and less parallelizable than feed-forward or attention-based architectures.