Sequence prediction

Sequence prediction is a problem that involves using historical sequence information to predict the next value or values in the sequence.The sequence may be symbols like letters in a sentence or real values like those in a time series of prices. Sequence prediction may be easiest to understand in the context of time series forecasting as the problem is already generally understood.In this post, you will discover the standard sequence prediction models that you can use to frame your own sequence prediction problems.

After reading this post, you will know:

- How sequence prediction problems are modeled with recurrent neural networks.
- The 4 standard sequence prediction models used by recurrent neural networks.
- The 2 most common misunderstandings made by beginners when applying sequence prediction models.

Models for Sequence Prediction

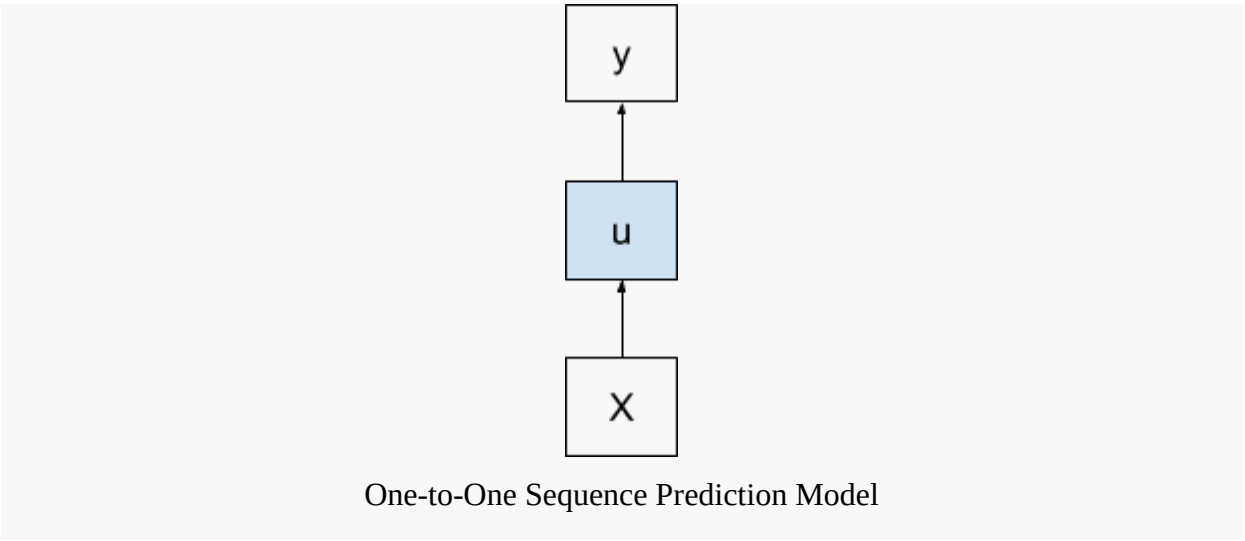
In this section, will review the 4 primary models for sequence prediction.

We will use the following terminology:

- X: The input sequence value, may be delimited by a time step, e.g. X(1).
- u: The hidden state value, may be delimited by a time step, e.g. u(1).
- y: The output sequence value, may be delimited by a time step, e.g. y(1).

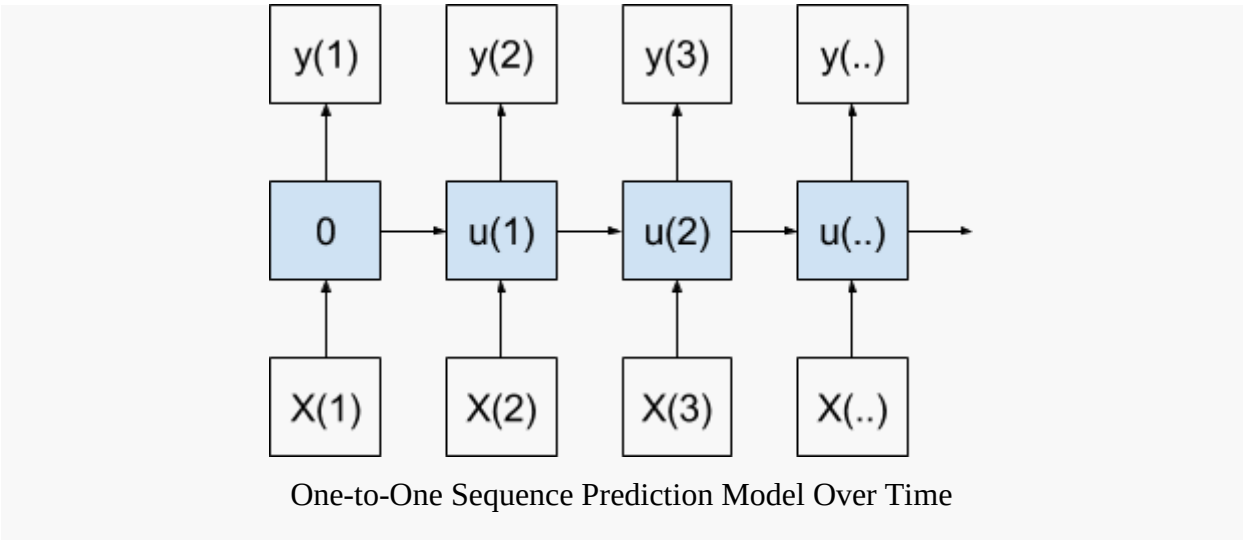
One-to-One Model

A one-to-one model produces one output value for each input value.



One-to-One Sequence Prediction Model

The internal state for the first time step is zero; from that point onward, the internal state is accumulated over the prior time steps.



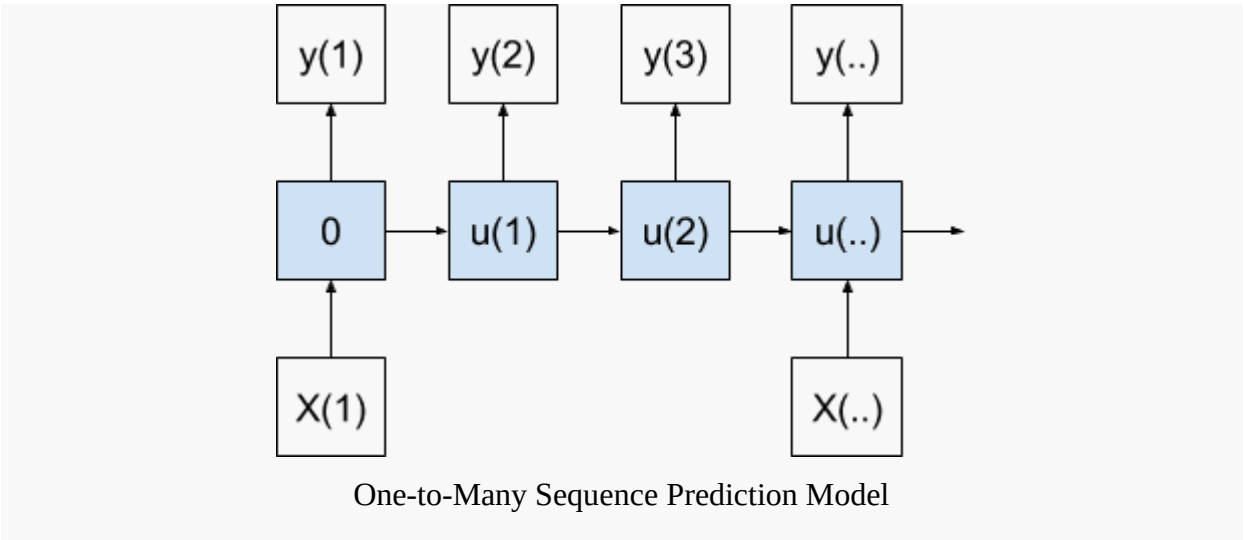
One-to-One Sequence Prediction Model Over Time

In the case of a sequence prediction, this model would produce one time step forecast for each observed time step received as input.

This is a poor use for RNNs as the model has no chance to learn over input or output time steps (e.g. BPTT). If you find implementing this model for sequence prediction, you may intend to be using a many-to-one model instead.

One-to-Many Model

A one-to-many model produces multiple output values for one input value.



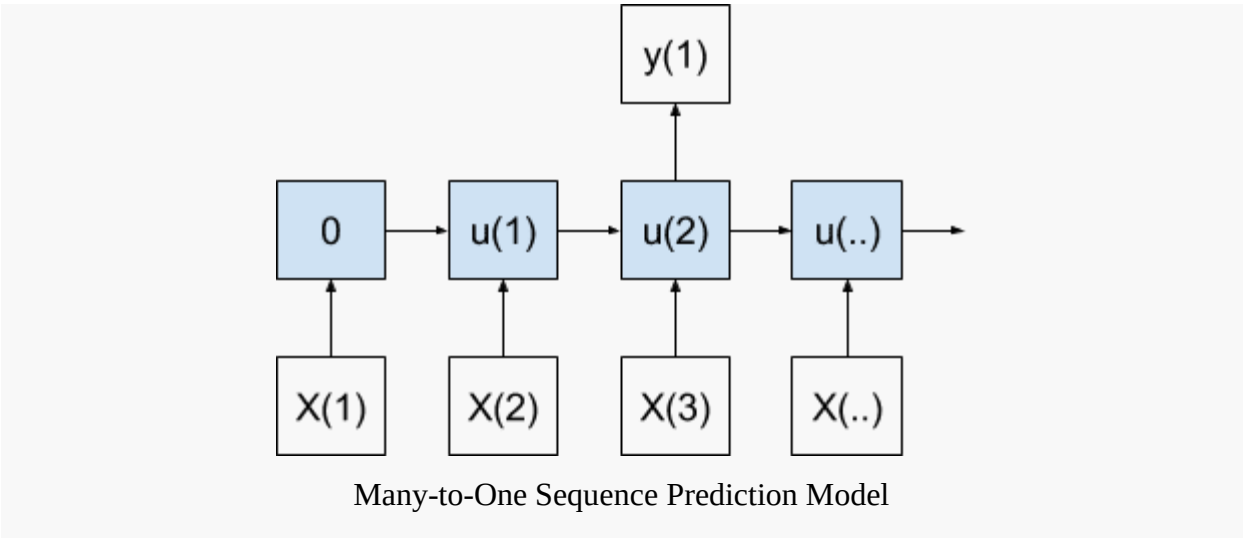
One-to-Many Sequence Prediction Model

The internal state is accumulated as each value in the output sequence is produced.

This model can be used for image captioning where one image is provided as input and a sequence of words are generated as output.

Many-to-One Model

A many-to-one model produces one output value after receiving multiple input values.



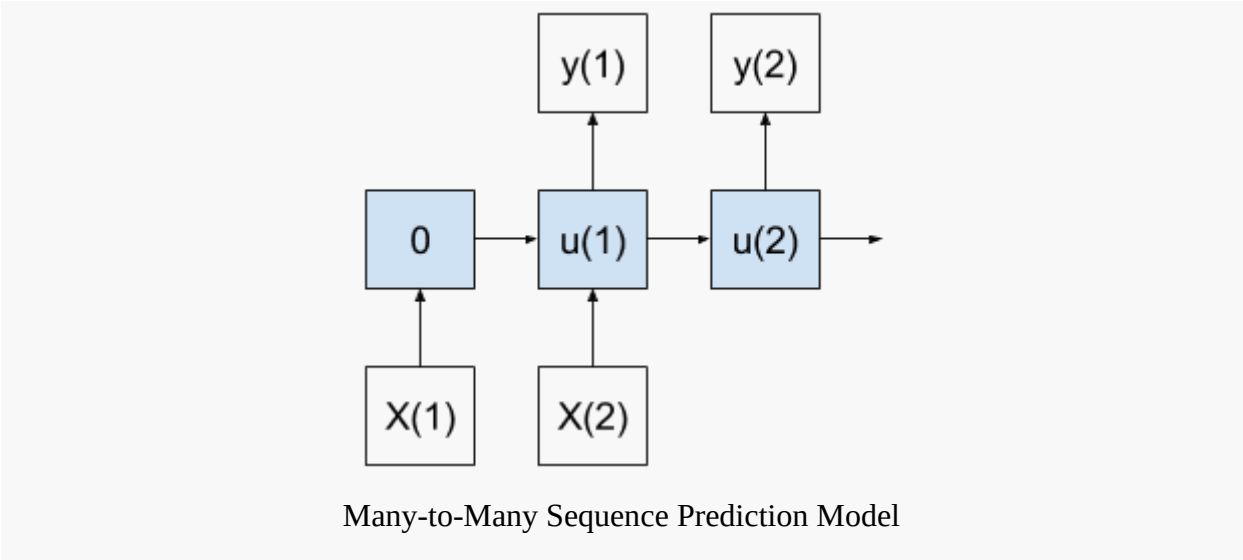
Many-to-One Sequence Prediction Model

The internal state is accumulated with each input value before a final output value is produced.

In the case of time series, this model would use a sequence of recent observations to forecast the next time step. This architecture would represent the classical autoregressive time series model.

Many-to-Many Model

A many-to-many model produces multiple outputs after receiving multiple input values.



Many-to-Many Sequence Prediction Model

As with the many-to-one case, state is accumulated until the first output is created, but in this case multiple time steps are output.

Importantly, the number of input time steps do not have to match the number of output time steps. Think of the input and output time steps operating at different rates.

In the case of time series forecasting, this model would use a sequence of recent observations to make a multi-step forecast.

In a sense, it combines the capabilities of the many-to-one and one-to-many models.

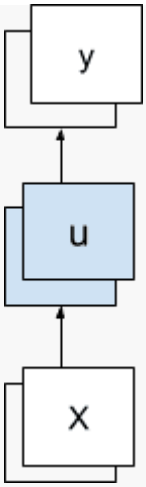
Cardinality from Timesteps

A common point of confusion is to conflate the above examples of sequence mapping models with multiple input and output features.

A sequence may be comprised of single values, one for each time step.

Alternately, a sequence could just as easily represent a vector of multiple observations at the time step. Each item in the vector for a time step may be thought of as its own separate time series. It does not affect the description of the models above.

For example, a model that takes as input one time step of temperature and pressure and predicts one time step of temperature and pressure is a one-to-one model, not a many-to-many model.



Multiple-Feature Sequence Prediction Model

The model does take two values as input and predicts two values, but there is only a single sequence time step expressed for the input and predicted as output.

The cardinality of the sequence prediction models defined above refers to time steps, not features (e.g. univariate or multivariate sequences).