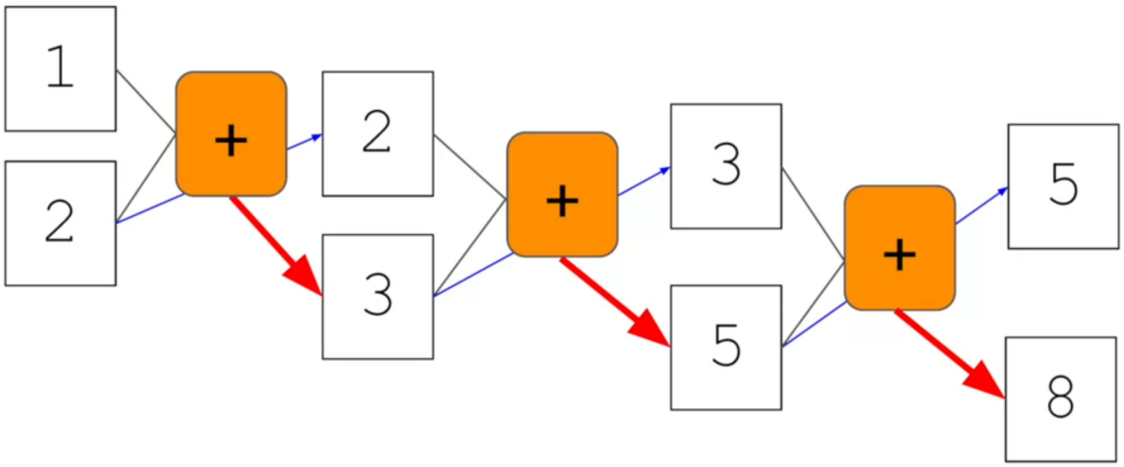
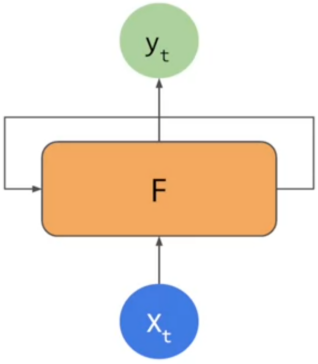
Week 3

- When words are tokenized and broken down to sub-words its hard to understand the meaning of the words
- Normal neural network does not take sequence into consideration
- The Fibonacci is the most simple example of recurrent neural network (RNN)



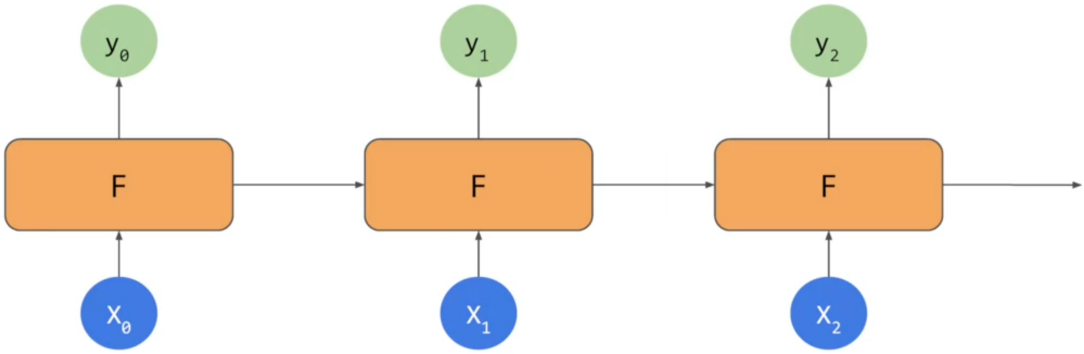
Fibonacci numbers graphed out

- It looks something like this in RNN:



A basic RNN diagram

- This becomes clear when chaining the RNN cells together:



LSTMs

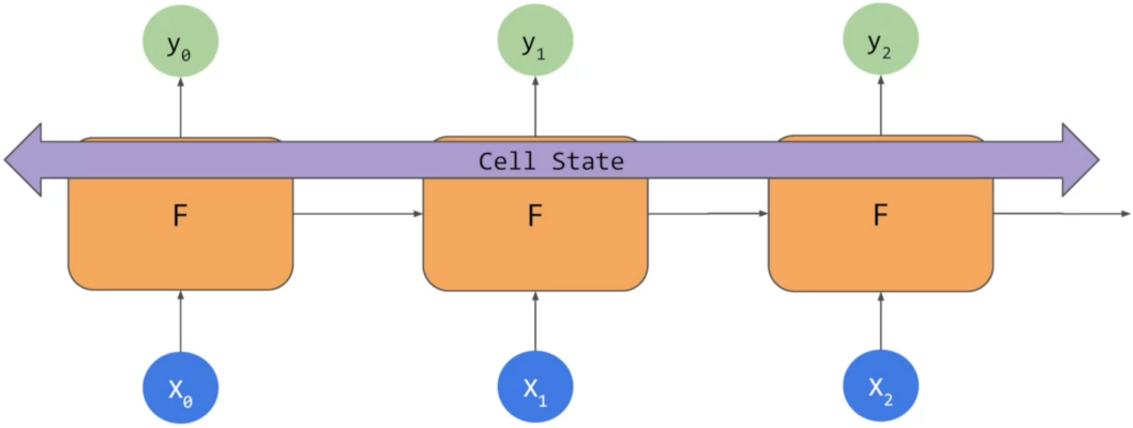
- Normal RNNs forget content when the text goes too long
- Here is an example:

I lived in Ireland, so at school they made me learn how to speak <...>

I lived in Ireland, so at school they made me learn how to speak Gaelic

- The example requires context from the first part of the sentence to answer the second part

- An update to RNN is called LSTM (long short term memory) cells are created
- LSTMs has additional pipelines called cell state that passes through the network and keeps context from earlier tokens relevant in later ones
- And the cell state could also impact earlier cells



- To implement LSTM:

```
model = tf.keras.Sequential([
  tf.keras.layers.Embedding(tokenizer.vocab_size, 64),
  tf.keras.layers.Bidirectional(tf.keras.layers.LSTM(64)),
  tf.keras.layers.Dense(64, activation='relu'),
  tf.keras.layers.Dense(1, activation='sigmoid')
])
```

- Can also feed LSTM cells into other LSTM cells
- But do have to enable `return_sequences` like this:

```
model = tf.keras.Sequential([
  tf.keras.layers.Embedding(tokenizer.vocab_size, 64),
  tf.keras.layers.Bidirectional(tf.keras.layers.LSTM(64), return_sequences=True),
  tf.keras.layers.Bidirectional(tf.keras.layers.LSTM(32)),
  tf.keras.layers.Dense(64, activation='relu'),
  tf.keras.layers.Dense(1, activation='sigmoid')
])
```

- Code using LSTM cells

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/d474043f-543a-413d-8198-efc743aee05e/imdb_multi_layer_lstm.ipynb

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/4f4b6200-92f8-4256-b364-5d6ccaa17f16/imdb_single_layer_lstm.ipynb

GRU

- Code example for GRU, a variation of RNN

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/9bf4e74e-3ef7-446d-afab-08e54ac8bd6a/multi_layer_GRU.ipynb

Exercises

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/f5e42bbc-6575-40ea-9fb3-e3afe38ea73b/exercise_1_w3.py

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/24c16927-3927-4fd5-8fb6-d4fe94003859/exercise_1_w3.ipynb