1. Can you think of a few applications for a sequence-to-sequence RNN?

Applications are **speech recognition, machine translation, image captioning and question answering**.

What about a sequence-to-vector RNN, and a vector-to-sequence RNN?

* Sequence to Sequence Learning, RNN is trained to map an input sequence to an output sequence which is not necessarily of the same length.
* a vector-to-sequence RNN **is a model that takes a single vector as input and produces a sequence as output**. An example of these models can be image to sentence model, which takes an image(consider it as a vector) and then produces a sentence to describe that image

1. How many dimensions must the inputs of an RNN layer have?

The RNN input needs to have **3 dimensions**.

**What does each dimension represent?**

Typically it would be batch size, the number of steps and number of features

**What about its outputs?**

Outputs are the last layer

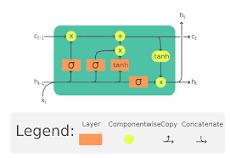
1. If you want to build a deep sequence-to-sequence RNN, which RNN layers should have return\_sequences=True? What about a sequence-to-vector RNN?
2. **Suppose you have a daily univariate time series, and you want to forecast the next seven days. Which RNN architecture should you use?**

* LSTM - Long short term memory networks is applied in time series forecasting. This Model solve various problem of univariate time series forecasting

1. What are the main difficulties when training RNNs? How can you handle them?

* There are two widely known issues with properly training Recurrent Neural Networks, **the vanishing and the exploding gradient problems.**
* The vanishing gradient problem is caused by the derivative of the activation function used to create the neural network. The simplest solution to the problem is to **replace the activation function of the network**. Instead of sigmoid, use an activation function such as ReLU.

1. **Can you sketch the LSTM cell’s architecture?**

The LSTM architecture aims to provide a short-term memory for RNN that can last thousands of timesteps, thus "long short-term memory". **A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate.**

1. Why would you want to use 1D convolutional layers in an RNN?

1D CNN **can perform activity recognition task from accelerometer data**, such as if the person is standing, walking, jumping etc.

1. **Which neural network architecture could you use to classify videos?**

Video classification is similar to image classification, in that the algorithm uses feature extractors, such as **convolutional neural networks (CNNs)**, to extract feature descriptors from a sequence of images and then classify them into categories.

1. Train a classification model for the SketchRNN dataset, available in TensorFlow Datasets.