UE21CS343BB2

Topics in Deep Learning

Dr. Shylaja S S
Director of Cloud Computing & Big Data (CCBD), Centre for Data Sciences & Applied Machine Learning (CDSAML)
Department of Computer Science and Engineering
shylaja.sharath@pes.edu

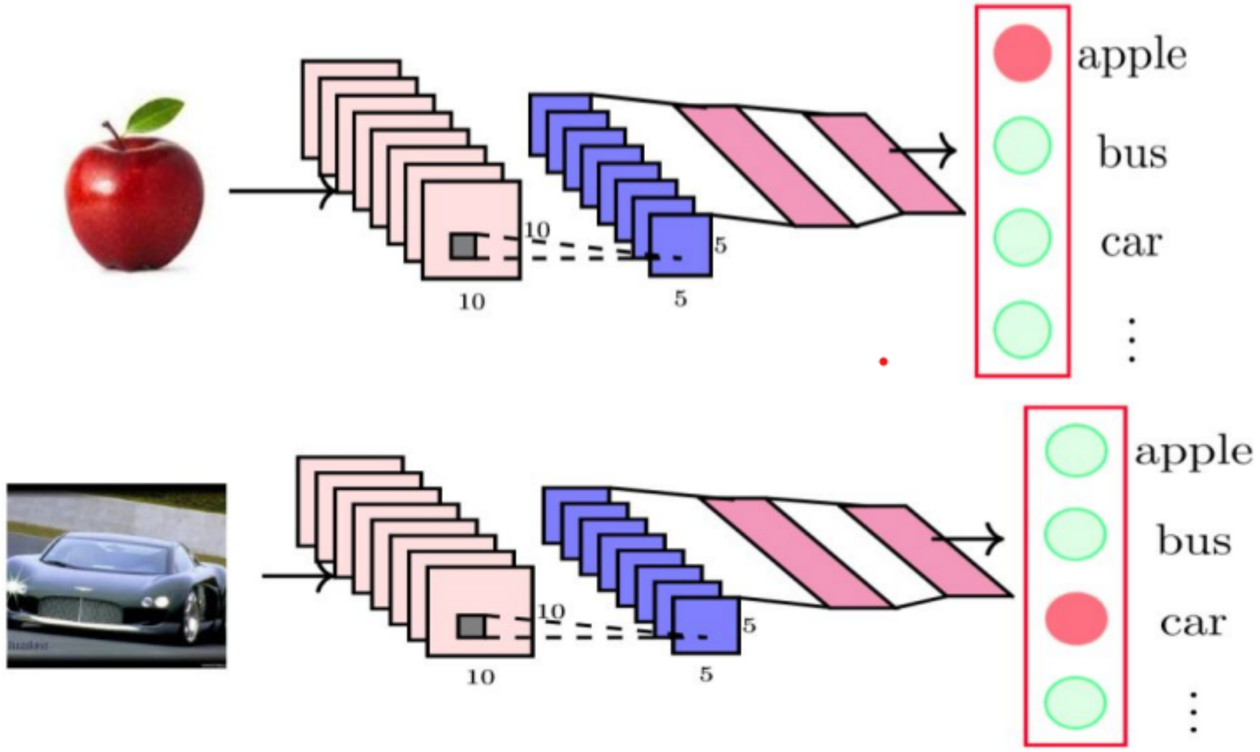
Ack:Aryan Sharma,
Teaching Assistant

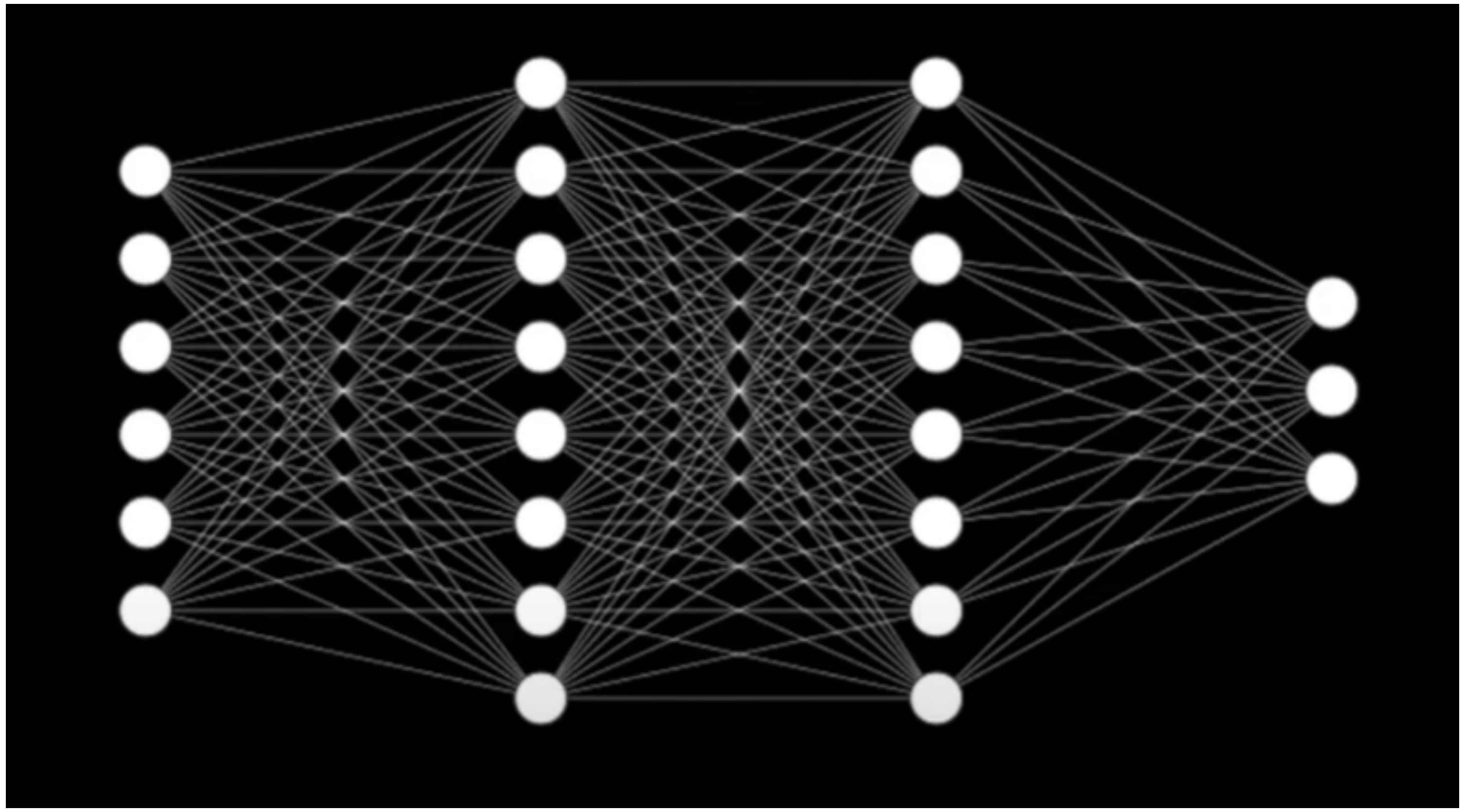
UE21CS343BB2: Topics in Deep Learning

Introduction to RNNs

Introduction

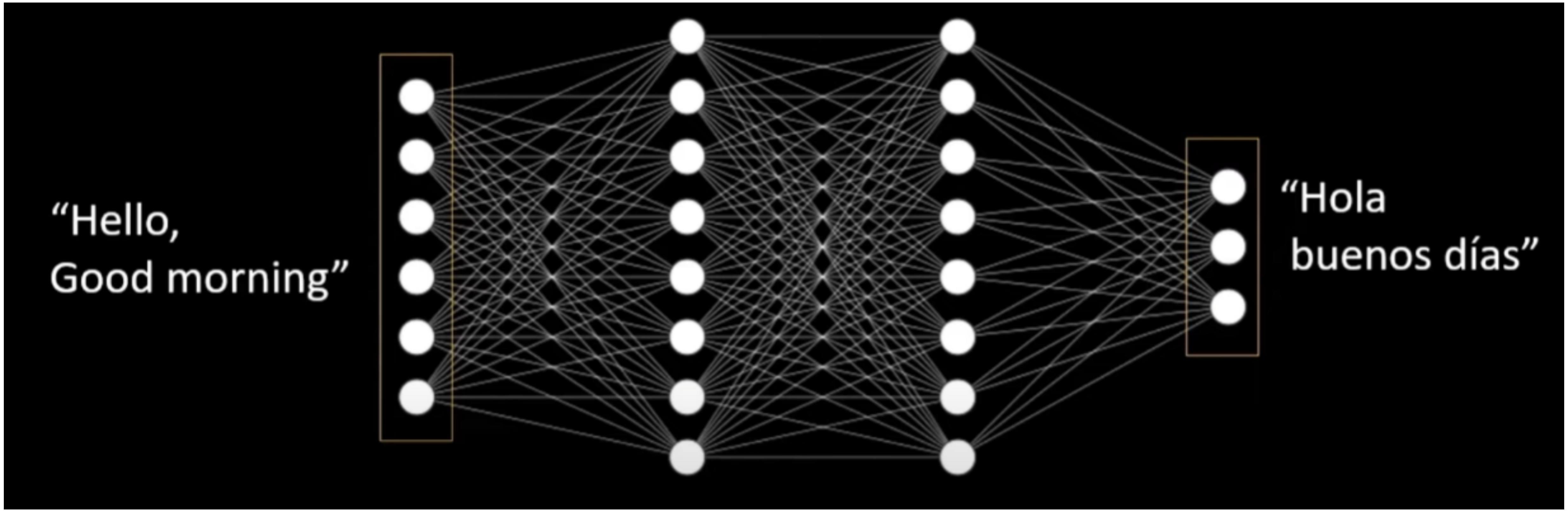
- Artificial Neural Networks and Convolutional Neural Networks work very well in predicting outputs to specific sized inputs.
- Here the inputs being certain images resized to fit a particular ratio and a fixed set of classes or outputs.





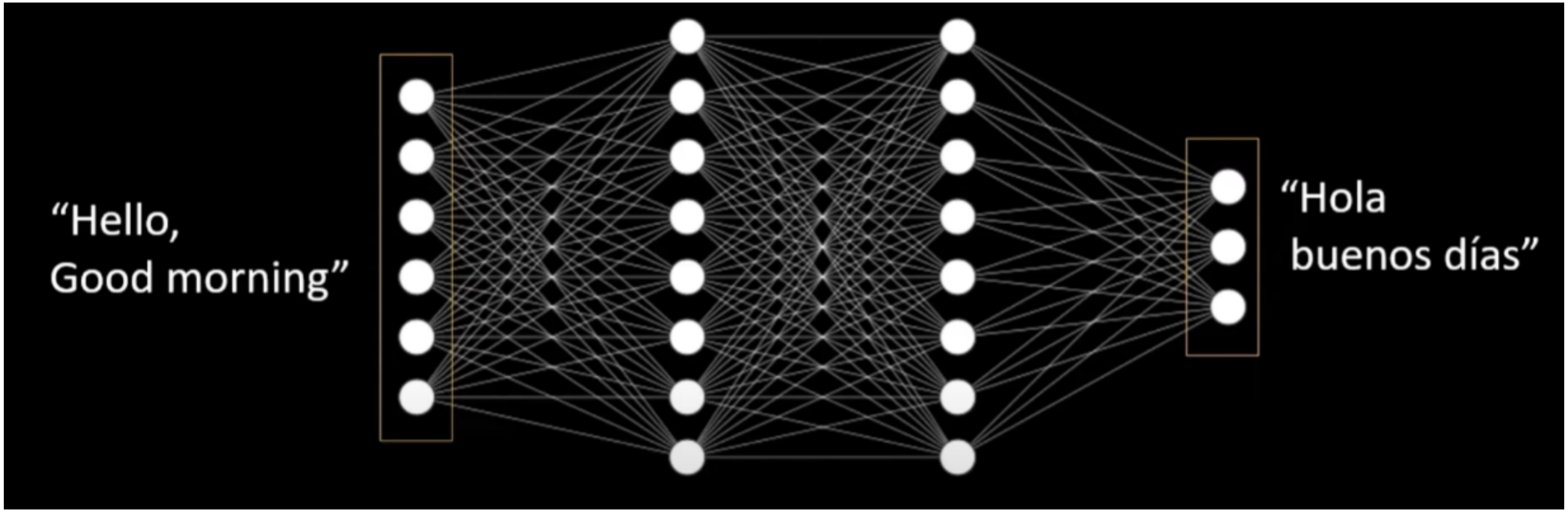
Introduction

- To understand its limitations, let's take an example of a use case.
- Let's say we want to translate a sentence from English to another language.
- Consider the sentence - “Hello good morning” must be translated to spanish.



Introduction

- The first step would be to convert the words into numbers, specifically vectors which is a format the computer can understand.
- For example - Let's use One Hot Encoding.
  Hello = [1,0,0]; Good = [0,1,0]; Morning = [0,0,1].
- But in the case of there being a longer sentence we have taken an ANN with 6 input nodes. What would the rest of the nodes get as an input?

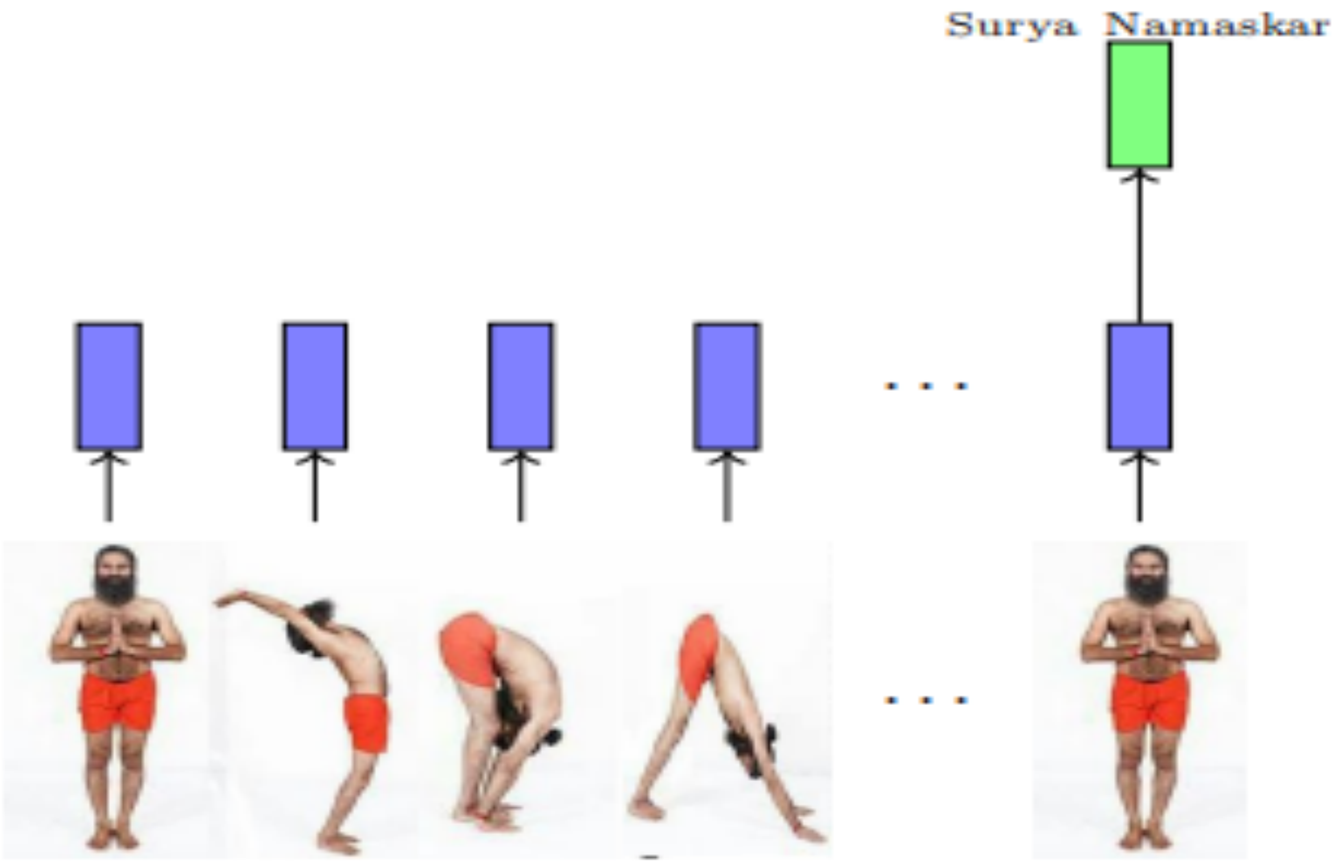


Introduction

- We could zero encode it in case of a shorter input. This is okay considering our vocabulary is only 3 words.
- In real cases, if each root word in the language or at least the vocabulary in some specific case would need a vector, that would mean a vector whose dimensions are 10,000 or more!
- But, in cases of a large vocabulary (words and their respective vectors), the size of the vectors would lead to unnecessary overhead.
- **Problem #1- As you can see above, ANNs cannot accept variable length inputs.**

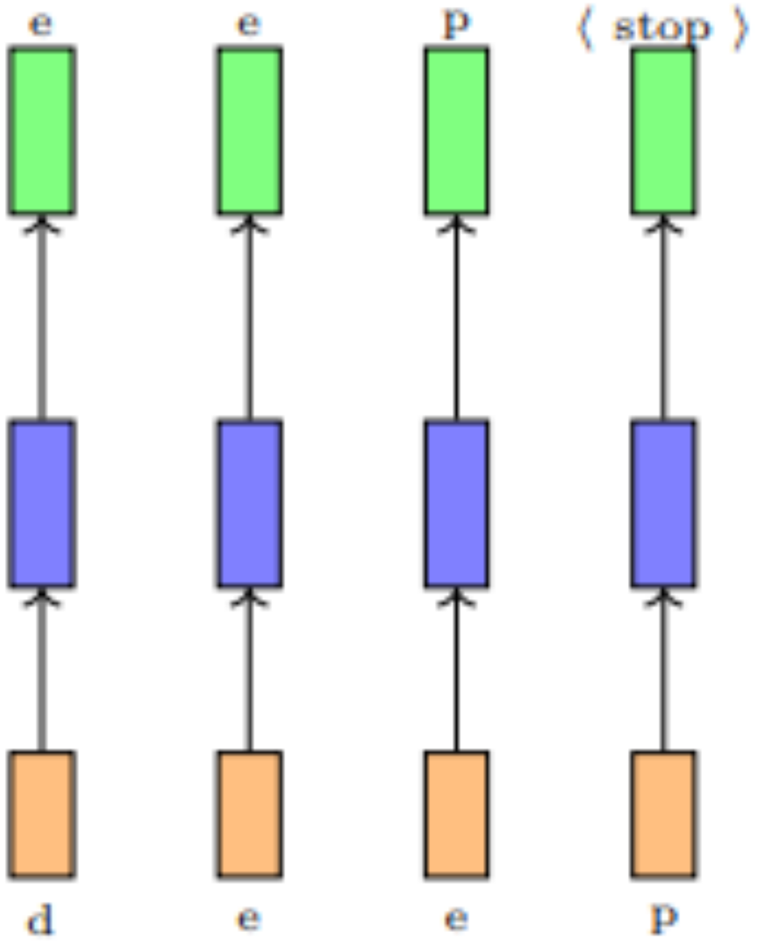
Introduction

- Now let’s take another example, video captioning.
- The video will be broken into individual images which are passed to the network.
- The order here becomes very important for the caption to make sense.



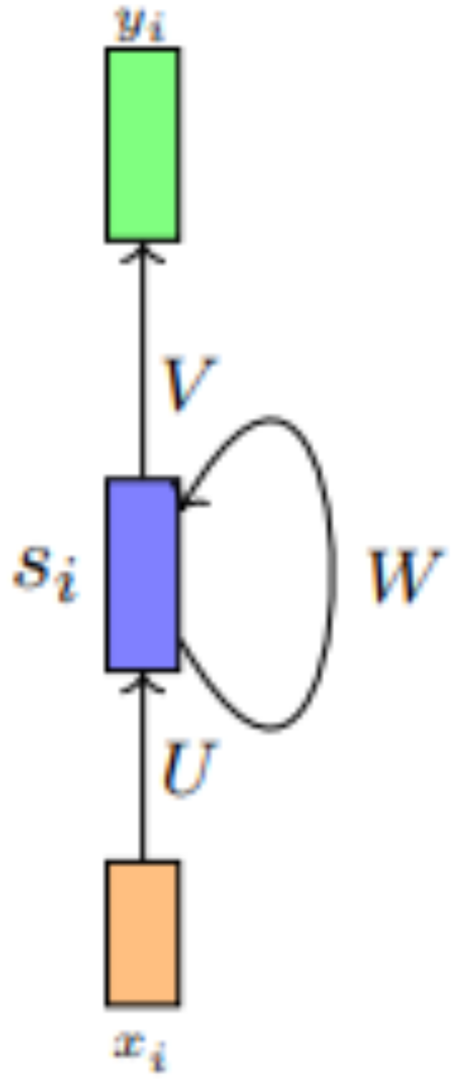
Introduction

- Let’s take more examples, consider the task of auto completion
- Given the first character ‘d’ you want to predict the next character ‘e’ and so on.

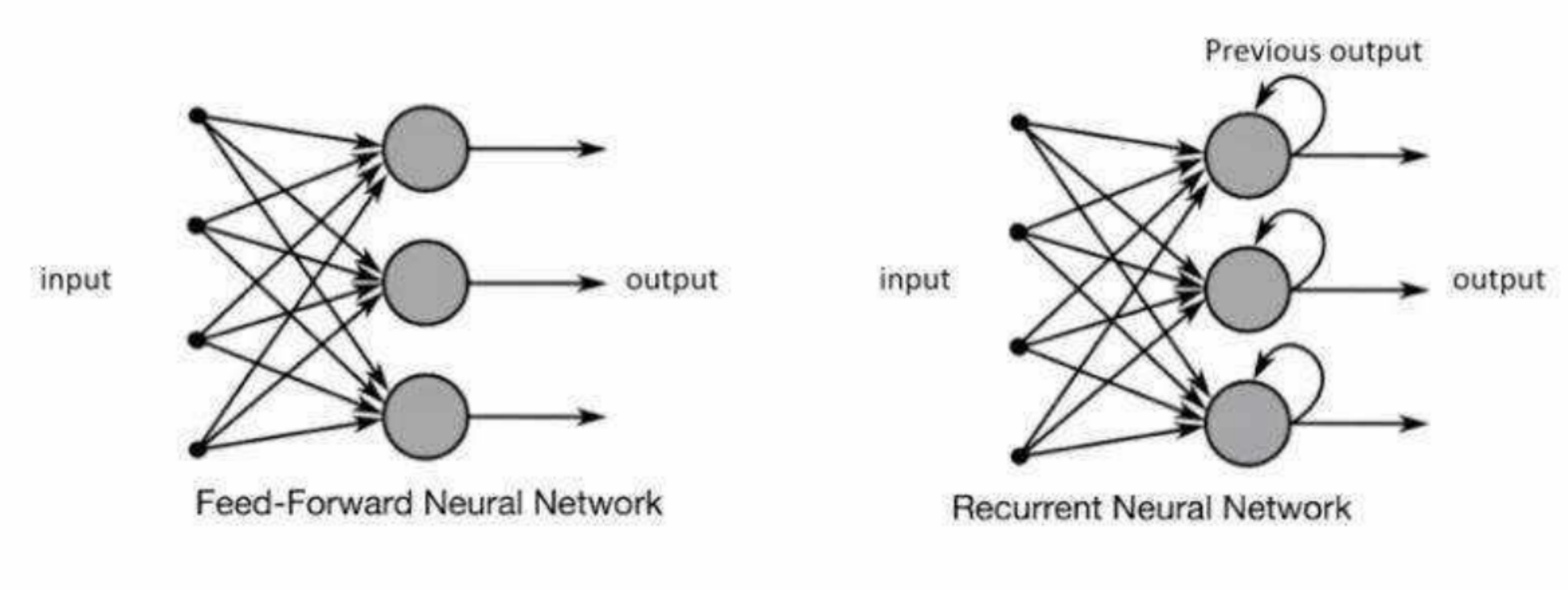


- As seen here, the sequence matters but also, the word can be anything other deep too! It can say detail, desperate, destroy or anything else after the ‘d’ and ‘e’.
- The ANN would not be able to handle such cases.
- In ANNs, the order the input is passed into the network is arbitrary.
- **Problem #2- Thus, in ANNs the sequence is not taken into consideration.**

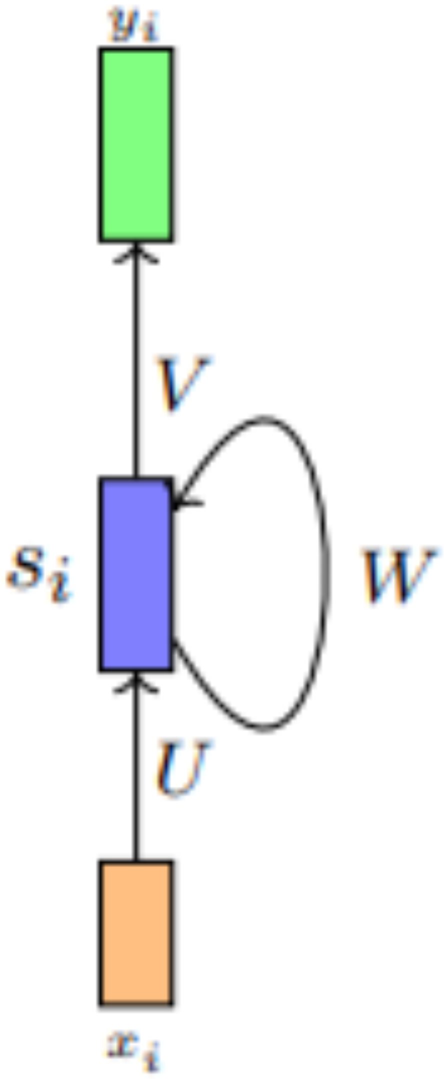
Recurrent Neurons - Memory Cells



- In the given diagram, a chunk of neural network, A, looks at some input x_t and outputs a value y_t. A loop allows information to be passed from one step of the network to the next.



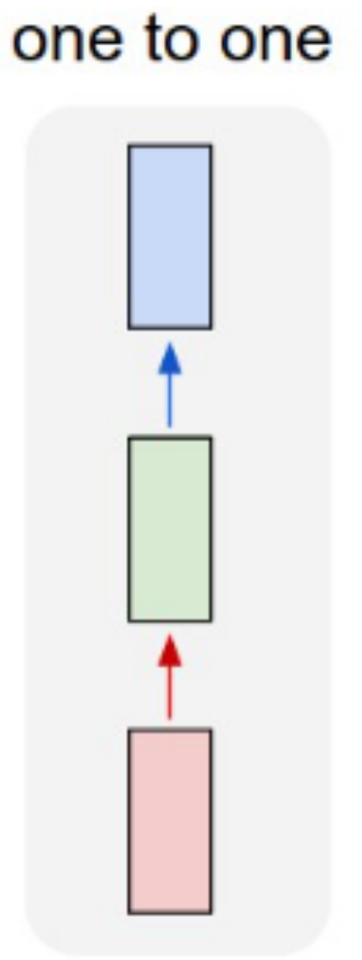
Recurrent Neurons - Memory Cells



- This information is stored in the ‘hidden state’ which constitutes its memory and passed in the ‘hidden layer’.
- A recurrent neural network can be thought of as multiple copies of the same network(A), each passing a message to a successor.

Variable-Length Input, Output Sequences

- Each rectangle is a vector and arrows represent functions (e.g. matrix multiply). Input vectors are in red, output vectors are in blue and green vectors hold the RNN's state (more on this soon).

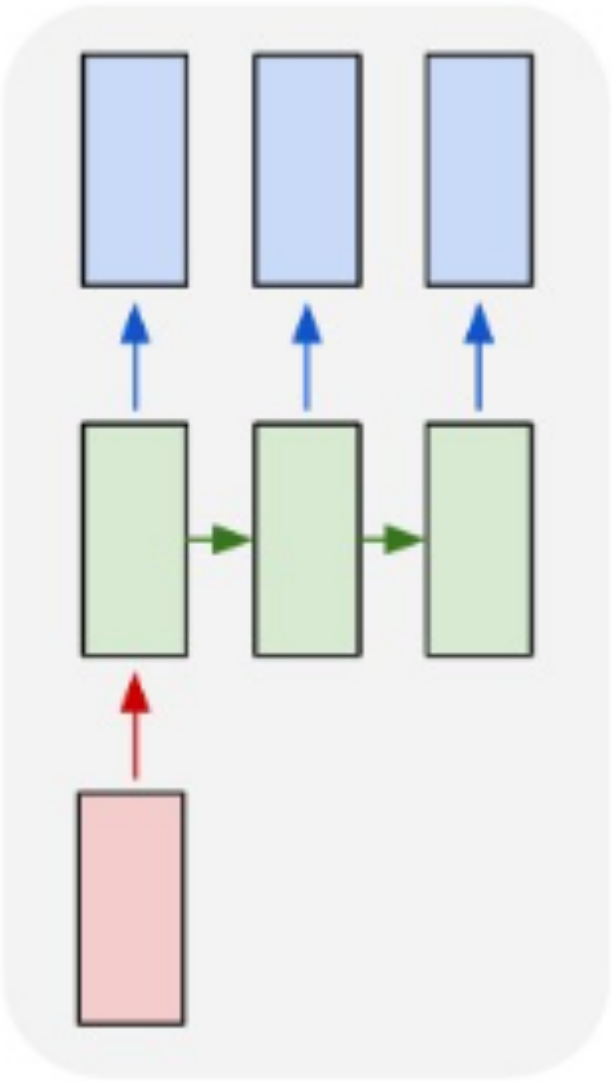


- Vanilla mode of processing without RNN, from fixed-sized input to fixed-sized output (e.g. image classification).

Variable-Length Input, Output Sequences

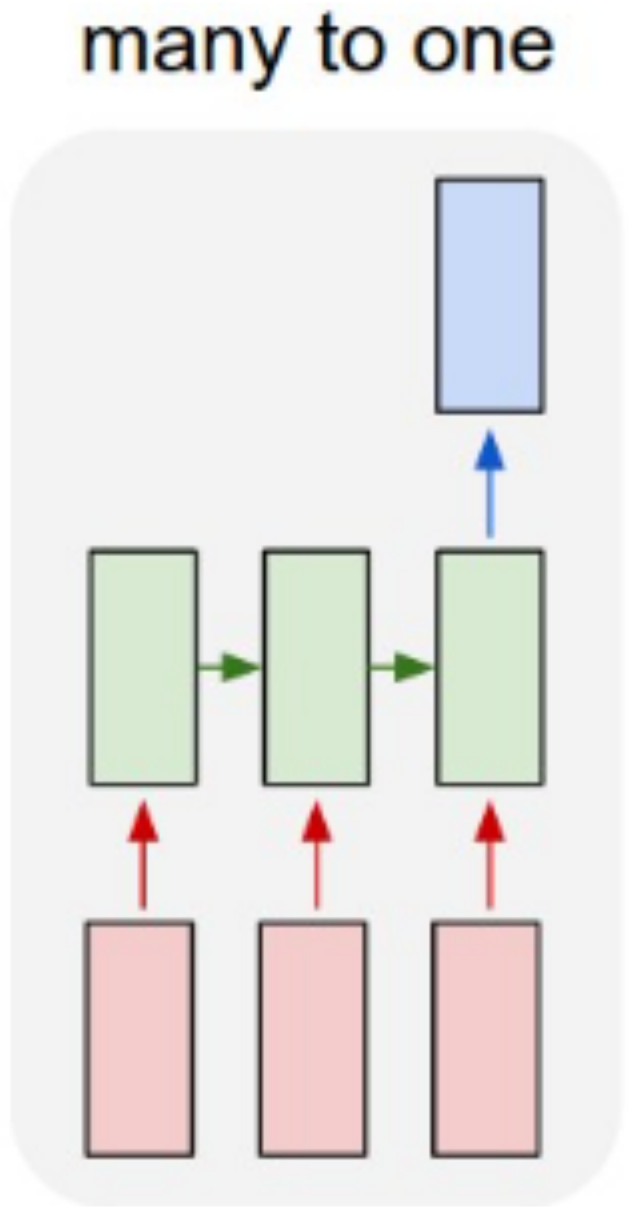
- Sequence output (e.g. image captioning takes an image and outputs a sentence of words).

one to many



Variable-Length Input, Output Sequences

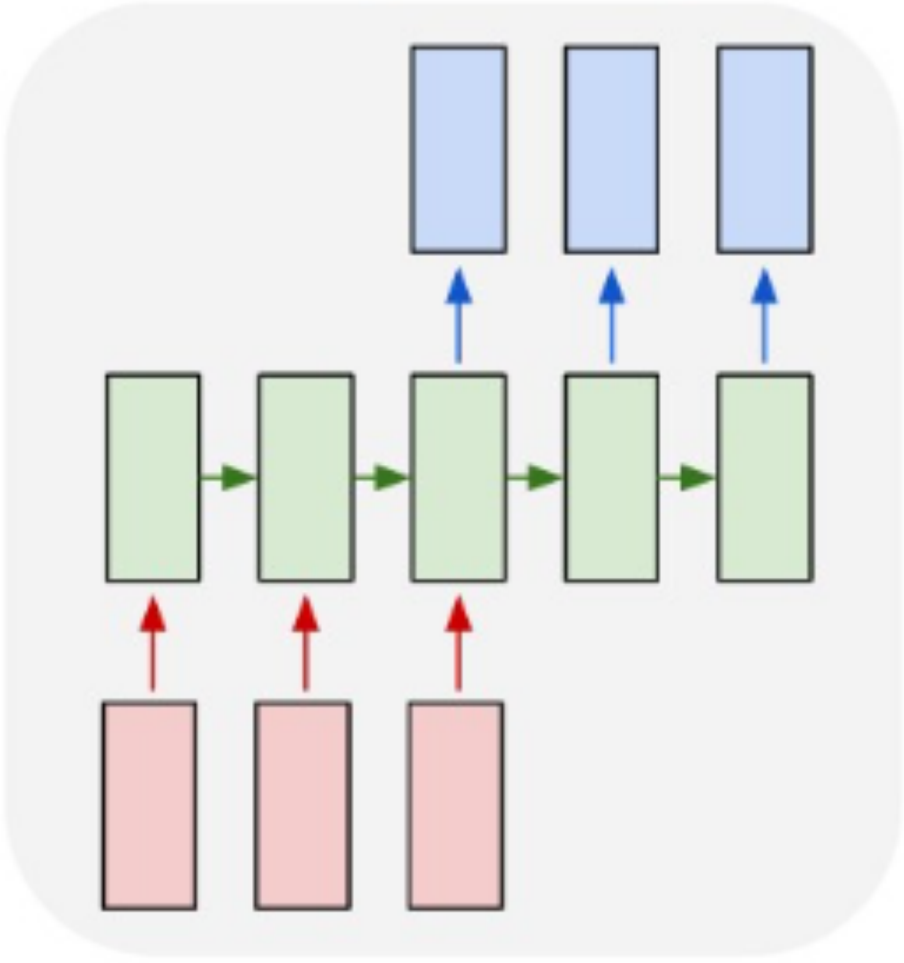
- Sequence input (e.g. sentiment analysis where a given sentence is classified as expressing positive or negative sentiment).



Variable-Length Input, Output Sequences

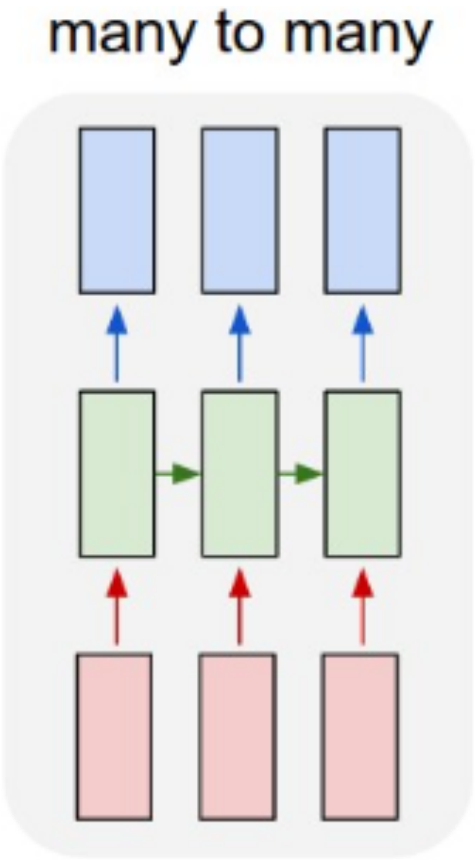
- Sequence input and sequence output (e.g. Machine Translation: an RNN reads a sentence in English and then outputs a sentence in French).

many to many



Variable-Length Input, Output Sequences

- Synced sequence input and output (e.g. video classification where we wish to label each frame of the video). Notice that in every case are no pre-specified constraints on the lengths sequences because the recurrent transformation (green) is fixed and can be applied as many times as we like.



Acknowledgements & References

- http://www.cse.iitm.ac.in/~miteshk/CS7015/Slides/Handout/Lecture14.pdf
- https://www.youtube.com/watch?v=7ZQgK_MV_t0
- https://karpathy.github.io/2015/05/21/rnn-effectiveness/

UE21CS343BB2

Topics in Deep Learning

Dr. Shylaja S S
Director of Cloud Computing & Big Data (CCBD), Centre for Data Sciences & Applied Machine Learning (CDSAML)
Department of Computer Science and Engineering
shylaja.sharath@pes.edu

Ack:Aryan Sharma, Teaching Assistant