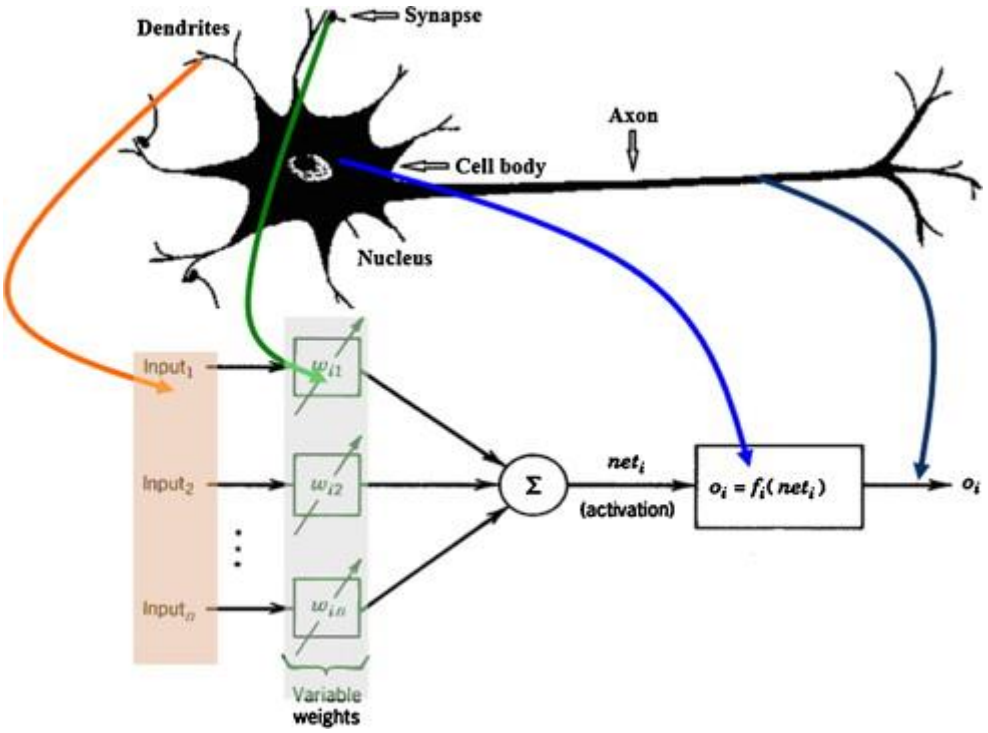
A Must-Read Introduction to Sequence Modelling (with use cases)

This article gives a brief description of Sequence Modelling with many practical examples and a thought experiment.

Introduction

Artificial Neural Networks (ANN) were supposed to replicate the architecture of the human brain, yet till about a decade ago, the only common feature between ANN and our brain was the nomenclature of their entities (for instance – neuron). These neural networks were almost useless as they had very low predictive power and less number of practical applications.

But thanks to the rapid advancement in technology in the last decade, we have seen the gap being bridged to the extent that these ANN architectures have become extremely useful across industries.



In this article, we will look at the two main advances in the field of artificial neural networks that have made these ANNs more like the human brain,

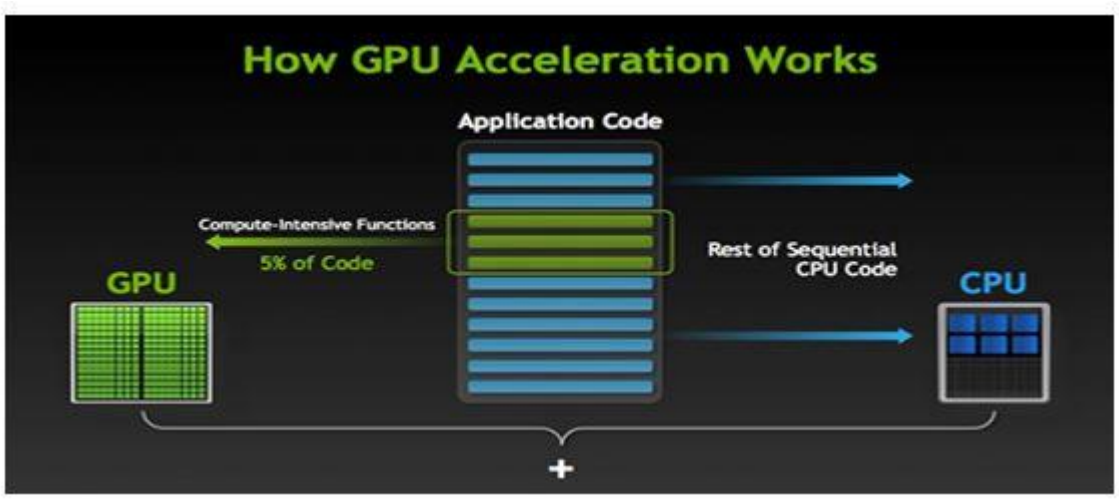
Table of Contents

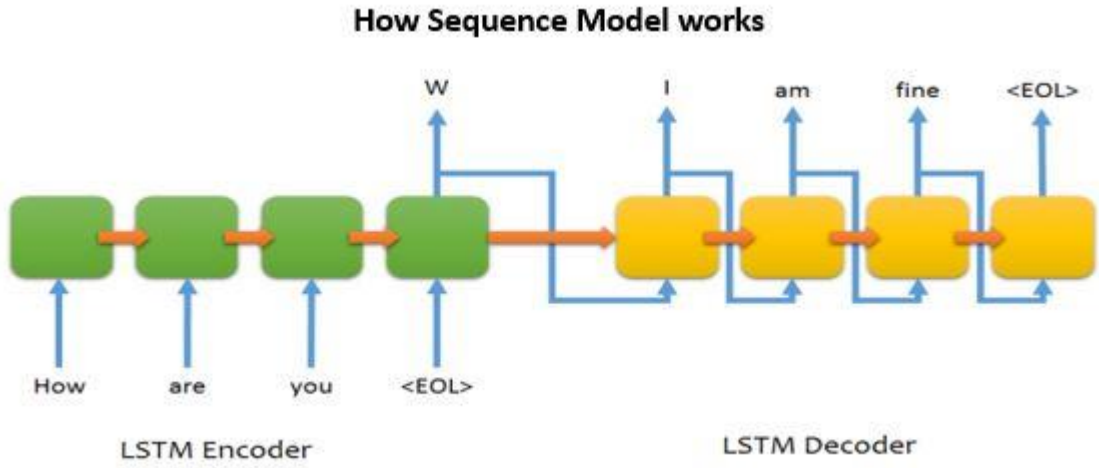
4. Sequence Generators
5. Sequence to Sequence NLP Models
6. Few More Sequence to Sequence Models that go beyond text

Two Main Advances in the Field of ANN

1. GPUs have immensely improved our computational power that now enables us to vastly increase the depth and breadth of neurons. However, we are still far away from reaching the number of neurons our brain has.
2. ANN can now process sequence data in both input and output nodes. This is how our brain works. Our brain does not solve binary classification to understand complex ideas. We formulate "Thoughts" based on a sequence of information given to us and then our brain expresses this "Thought" in understandable sequence of words.

Can we introduce this concept of "Thought" in an ANN? The answer is yes, and we will explore more about the idea in this article.





Sequence models have garnered a lot of attention because most of the data in the current world is in the form of sequences – it can be a number sequence, image pixel sequence, a video frame sequence or an audio sequence.

Over the last 10 years, we have stored 1000s of Petabytes (or more than 10 ^ 9 GBs) of unstructured sequence data for absolutely no reason as we had no way to fetch information out of such data formats. Luckily, **we now have this new family of neural network architectures called sequence models that can turn this data dump into GOLD MINES.**

The scope of this article is not to talk about all the complex mathematics that goes behind the scene in Sequence Modelling or give you some sample codes to run on sequence modelling (I will park that for some later articles), but to give you practical examples of sequence modelling implementations in the industry. These will enable you to identify business problems in your industry that might need this special tool.

To get a better understanding of what this article is about, below is a scenario which I want you to imagine. Put your analytical thinking hats on!

Thought Experiment

Walmrt has appointed you as the head of it's new vertical – WalKiosk. The company wants you to lead the development of a self servicing (human-less) store where a customer will only interact with Walmrt's Kiosk, which is very similar to a vending machine. They want to install this Kiosk in various locations across the United States.

A key difference between this Kiosk and a normal vending machine is that the Kiosk's display does not show the list of items, but simply an audio enabled Google-like search tab. The customer can literally walk up to these Kiosks, and say or type anything after the keyword "OK Walmrt, xxxxxx". Here is a sample interaction (try to evaluate if a human can do a better job than this Kiosk):

Customer says – "OK Walmrt, *I want the shoes which Leonardo DiCaprio wore in the 1st scene of the 1st movie he did with Nolan*" in any possible spoken language.

The idea is for the Kiosk to do a quick search and if it finds a convincing answer, it should reply, in the same language as the customer's query, something like – "*Leonardo DiCaprio wore black colored Nike shoes of model xxxxx. Click the link on the kiosk to watch a video cut of the scene you asked me to look at. Great news – we currently have the exact same shoe with the same size as you are wearing, and it's cost is $200. As you are a loyal customer of Walmrt, I have found a steal deal for you! The new price of the shoe, if you buy it immediately, is $150 for you"*.

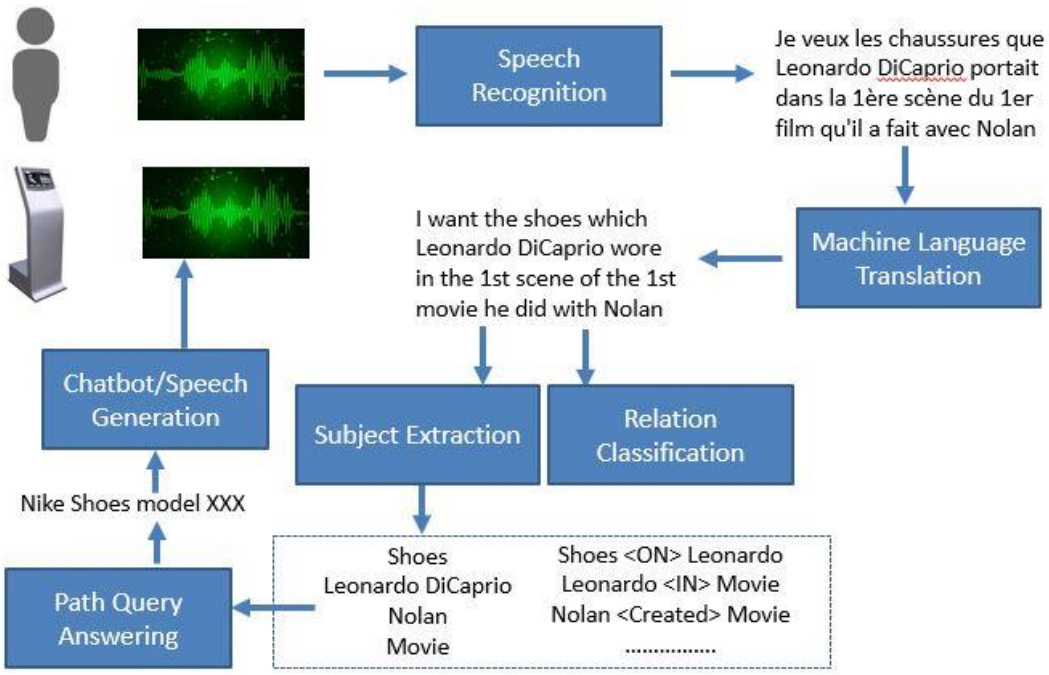
If the customer says "*I want to buy it"*, the Kiosk dispenses the shoe once the customer makes the payment.

Kiosk finally replies – "*Thanks Mr. XYZ for shopping with us today. Please give your valuable feedback for us to improve our service further.*" Customer writes or says the feedback of this transaction and leaves.

This simple transaction, that will probably take a good chunk of your time in today's world, will be resolved in less than 2 minutes (if everything works, that is).

Sounds futuristic? Here's a spoiler – all the fancy next gen functional skills you need to build in this Kiosk will be done mainly by a single architecture – sequence modelling. Here is a small list of tasks the Kiosk needs to do:

1. **Speech Recognition** to understand what the customer is saying
2. **Machine Language Translation** from source language to a known language (say English)
3. **Name entity/Subject extraction** to find the main subject of the customer's query translated in step 2
4. **Relation Classification** to tag relationships between various entities tagged in step 3
5. **Path Query Answering** (Similar to Google search) on entity-relationship found in step 3 & 4 using core knowledge graph
6. **Speech Generation** to generate answers for the customer with all the relevant information found in step 5
7. **Chatbot skill** to have conversational ability and engage with customers just like a human
8. **Text Summarization** of customer feedback to work on key challenges/pain points
9. **Product Sales Forecasting** to replenish stock



The skills required to create WalKiosk are not limited to these nine steps, but they are good enough to bring out the core idea. Each of these nine skills can be modeled by a single architecture – Sequence Modelling (but you already knew this).

You can imagine sequence modelling as a black box which stays almost the same; all you need to change is the input and target data for each of the nine skill sets. Leveraging the idea that all the model architectures in each step is the same, we can take this a step further and create a single model that takes input in any language and completes the self service process/reporting process/inventory management process all together.

If this was not enough to make you Google all about sequence modelling, let's look at an exhaustive list of all functions sequence modelling is capable of.

Practical Applications of Sequence Modelling

To make sure we cover most of the possible applications of sequence modelling, we will categorize them based on the type of input and output sequences. Inputs and outputs can be one of the following: Scalar, Trend, Text, Image, Audio or Video. If each of these six can be both input and output, we have 36 categories in total. However, not each of these pairs has been explored in depth yet.

Before moving to the list, pause for a moment and create your own list of applications (you can use our thought experiment as a reference).

Here goes the list:

Input		Target		Use Cases
Type	Elements	Type	Elements	
Scalar	One	Trends	Many	Pattern generation
		Audio	Many	Music Generation
		Text	Many	Text Generation
		Image	Many	Image generation
Trends	Many	Scalar	One	Stock Trading decisions
				Forecasting KPI for fixed duration
		Trends	Many	DNA Sequence analysis
				Time series forecasts
Text	Many	Scalar	One	Sentiment Classification
				Topic Classification
				Answer Selection
		Text	Many	Text Summarization
				Machine translation
				Chatbots
				Name Entity Recognition
				Subject Extraction
				Part of Speech Tagging
				Textual Entailment
				Relation Classification
		Trends	Many	Path Query Answering
		Audio	Many	Speech Generation
Image	Many	Scalar	One	Facial expression tagging
				Entity classification
		Text	Many	Image Captioning
		Image	Many	Image Modification
Audio	Many	Scalar	One	Sentiment Classification
				Number of speaker tagging
				Topic Classification
		Text	Many	Speech Recognition
				Conference Summarization
		Audio	Many	Speech Assistant
Video	Many	Scalar	One	Activity Recognition
		Text	Many	Subtitle generation

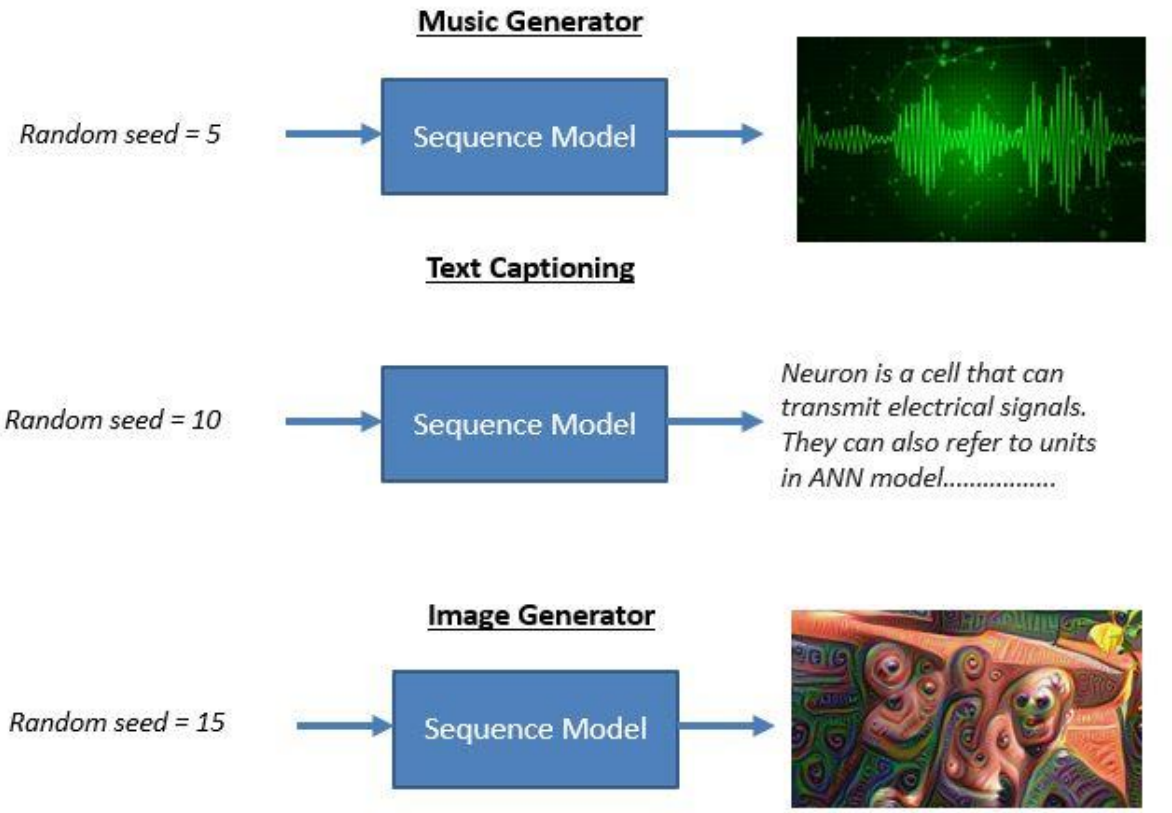
Reading the table is fairly straight forward:

- **Type** is the category of input/target
- **Elements** are the number of elements in input/target series
- **Use Cases** are the possible applications in the category

We will review a few of these use cases in order to get a grasp of the superpowers that our sequence model possess.

First, let's talk about the easiest of the lot – Sequence Generators

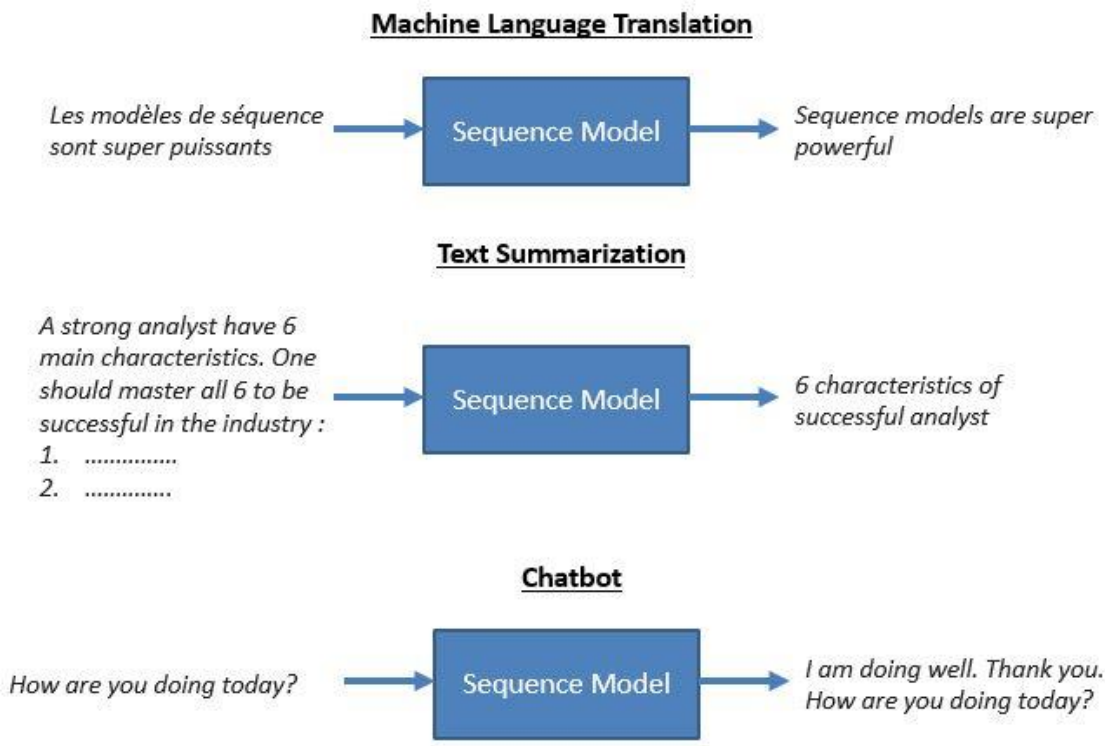
These generators generally take scalar inputs. The scalar input can be any random seed/number. Following are a few examples of generators:



Note that we can train our model on any specific type of data. For instance, if we train our text generator on a Harry Potter book, it is highly likely that you will get a text which is full of imagination/magic with the main character as Harry Potter. If you were lucky, you might get a chapter that makes sense and you can enjoy this privileged chapter that no one has access to!

Another example – if you train the model on Jazz music, you can create new songs in the same genre using this model. Yet another example – if you train the model on images of animals, you might see how cross breeds might look like.

Next, let's talk about the favorites – Sequence to sequence NLP Models

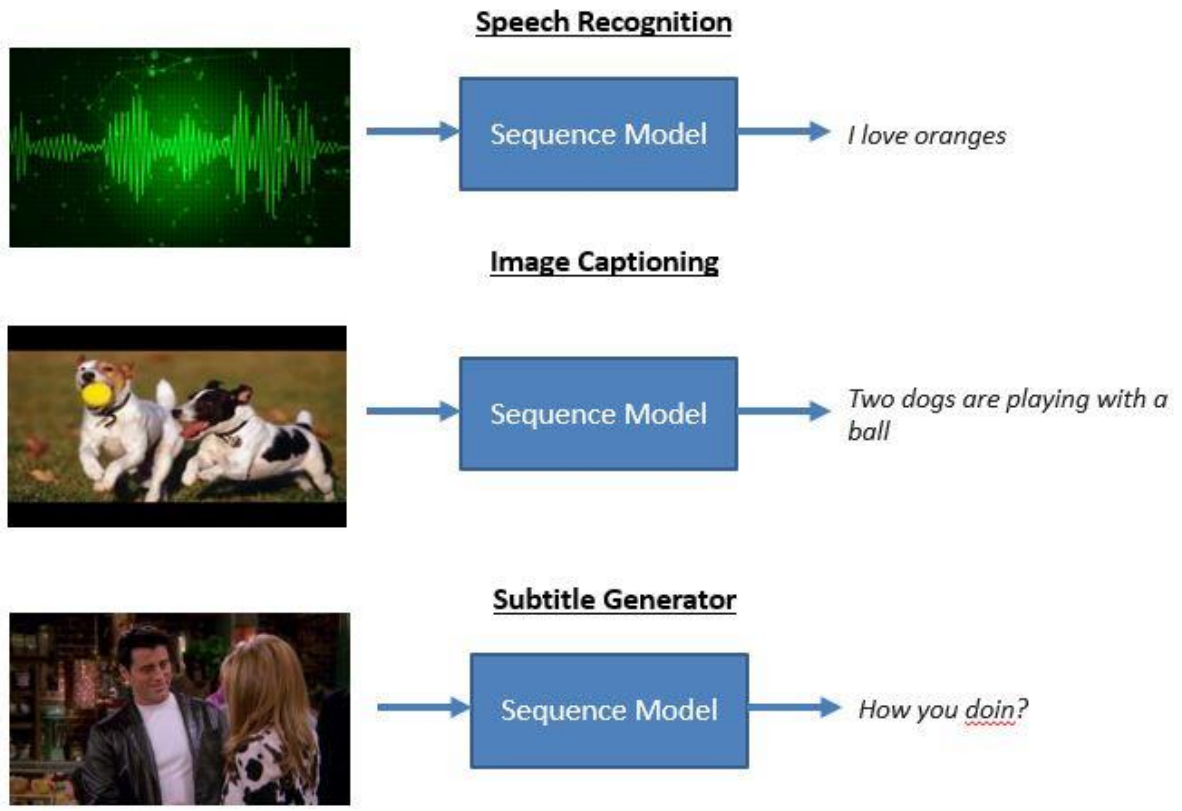


Machine Language Translation has reached new heights and is now competing strongly with human translators. Today, you can find real-time translating machines which are based on the core concept of sequence to sequence models.

Text summarization is another important use case of sequence models. Text summarization can significantly reduce the task of manually reading lengthy customer complaints, monitoring compliance based call/chat monitoring, and reviewing customer feedback on product etc.

Chatbot is yet another important application and is now being widely used in Operations/Call Centers/Chat Centers/Personal assistants like Siri/Google Home/Alexa.

Finally, we will talk about a few more sequence to sequence models that go beyond text



Speech recognition is currently the category which has absorbed the maximum investment in terms of money. Speech recognition is extremely important in tools like personal AI assistants (Alexa, Google Home, etc.) and call center speech recording tools.

Currently we have billion dollar companies whose sole competency is speech recognition. Speech recognition also uses sequence to sequence models extensively. Image Captioning is one of the hottest research fields which has a wide application in the social media industry. Subtitle generation has not reached the stage of production yet, but is being actively researched.

End Notes

A lot of the data science talent today focuses its effort on solving problems that already exist. An equally important task, for any successful data scientist or analyst, is to identify and create new tasks that can be solved analytically. The latter is a very different exercise and does not need a lot of coding experience or mathematically background. All you need to know is what is possible and what is not, using a given tool.

Problem identification is a skill set that is a “must” for any senior analytics professional. I hope this introductory article on sequence learning gave you strong motivation to start searching for new problems in your industry that can be solved using this method.