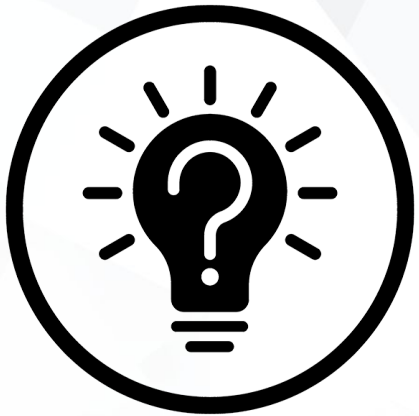
Recurrent Neural Networks

FinTech

Lesson 14.3v1.3



What Are We Going to Learn Today?

What We'll Learn Today

In this lesson, we'll cover the following:

01 RNN, the neural network that remembers the past

02 How to assess and compare models' performance

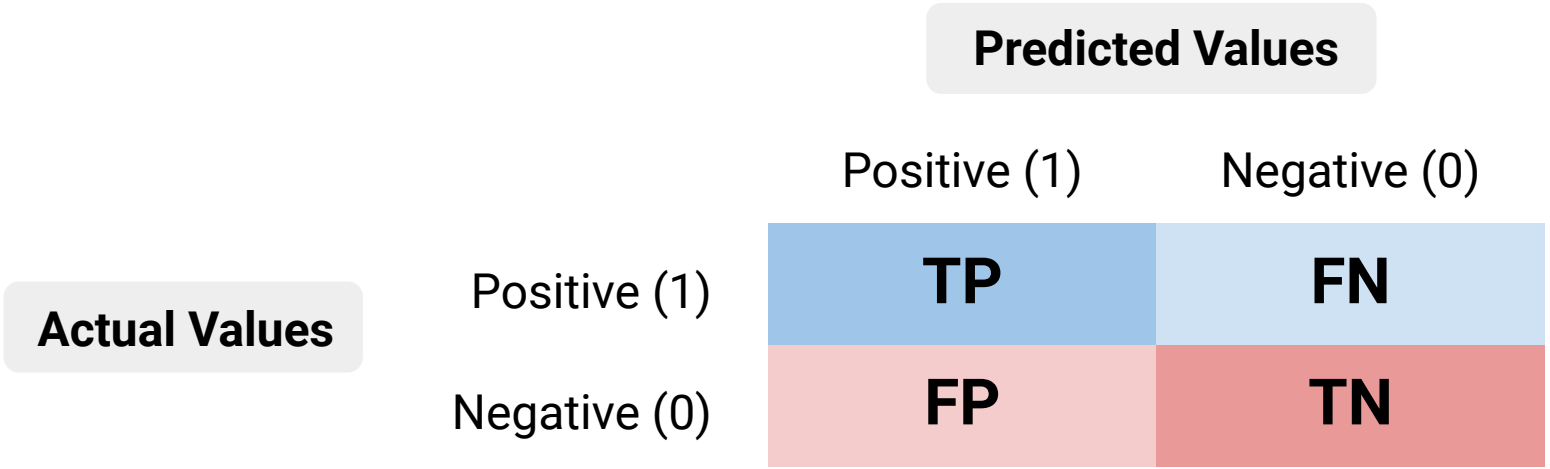
03 How to score sentiment using RNNs

04 How to use RNNs for time-series analysis

Introducing the ROC Curve and AUC

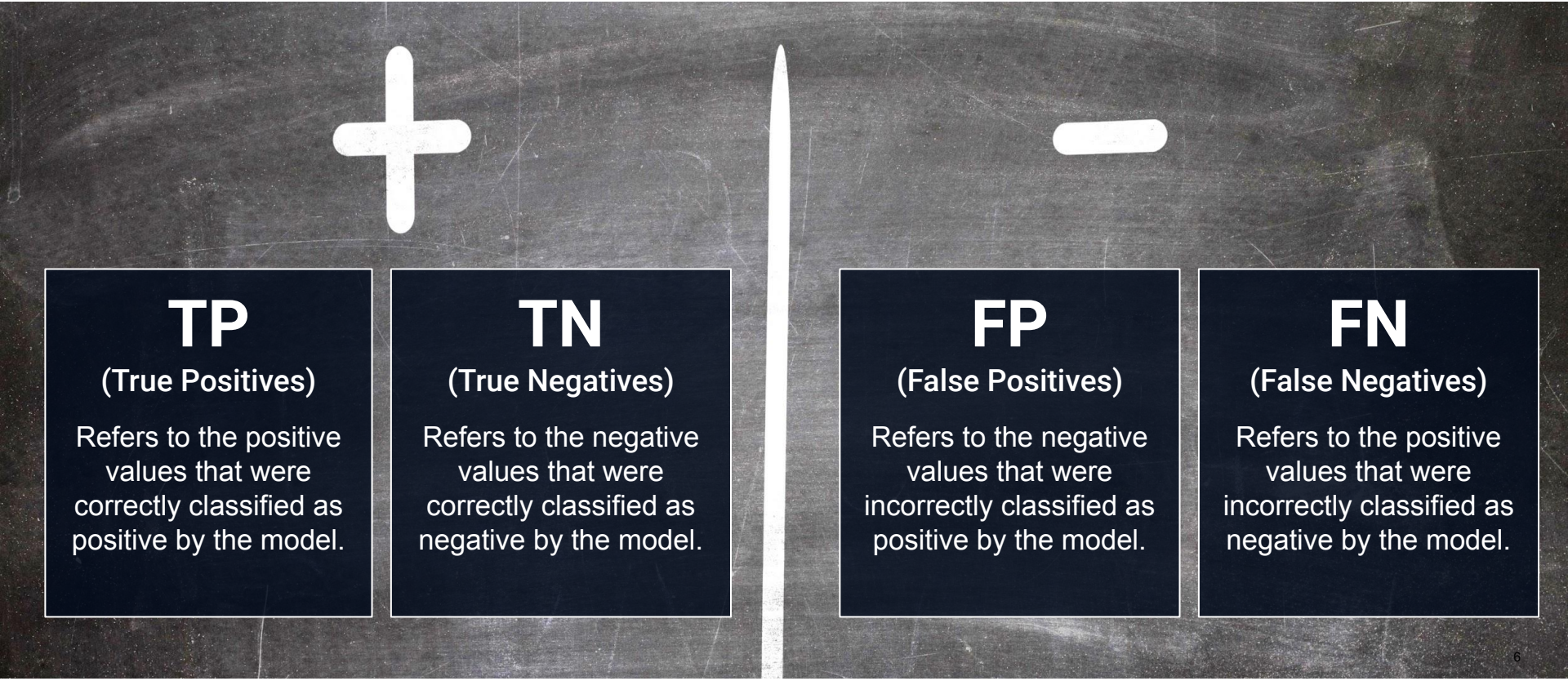
Introducing the ROC Curve and AUC

Recall that we use the confusion matrix to assess the performance of a binary classification model.



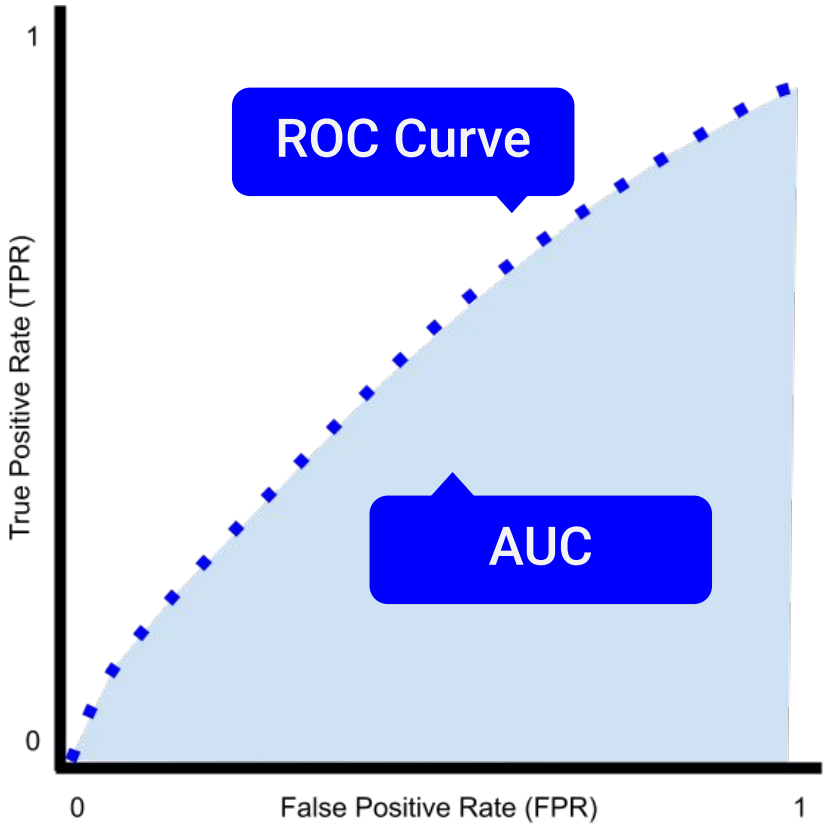
Introducing the ROC Curve and AUC

Also recall the four components of this matrix:



Introducing the ROC Curve and AUC

The ROC curve and AUC are techniques that allow us to check and visualize the performance of a classification model.



Understanding ROC

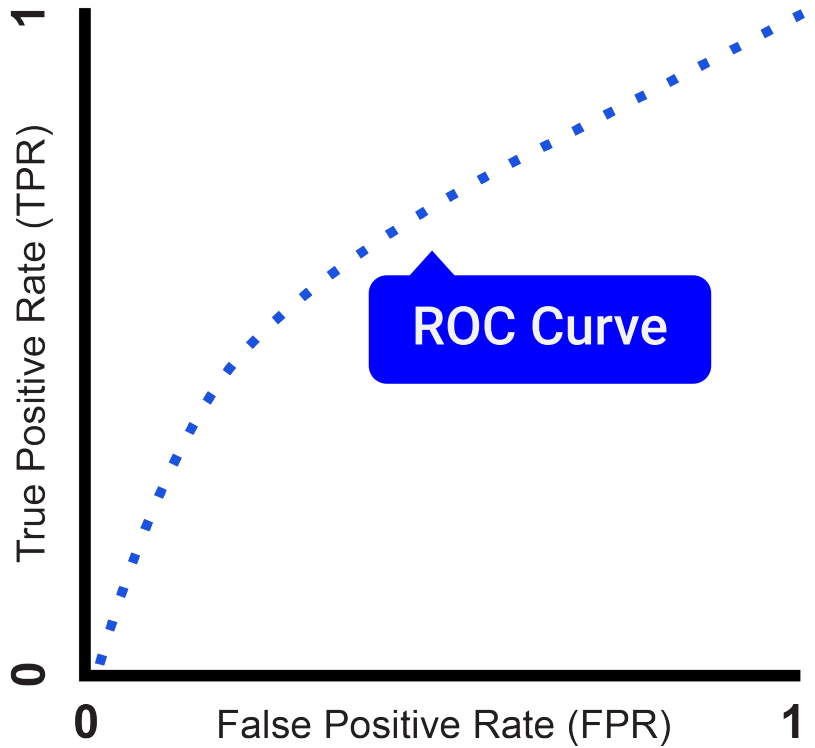
ROC

stands for

Receiver **O**perating **C**haracteristic

Introducing the ROC Curve and AUC

The ROC curve shows the performance of a classification model as its discrimination threshold is varied.



Introducing the ROC Curve and AUC

To plot a ROC curve, we use two parameters:

TPR

True Positive Rate
(also known as recall)

The **TPR** is calculated as follows:

$$TPR = \frac{TP}{TP + FN}$$

FPR

False Positive Rate

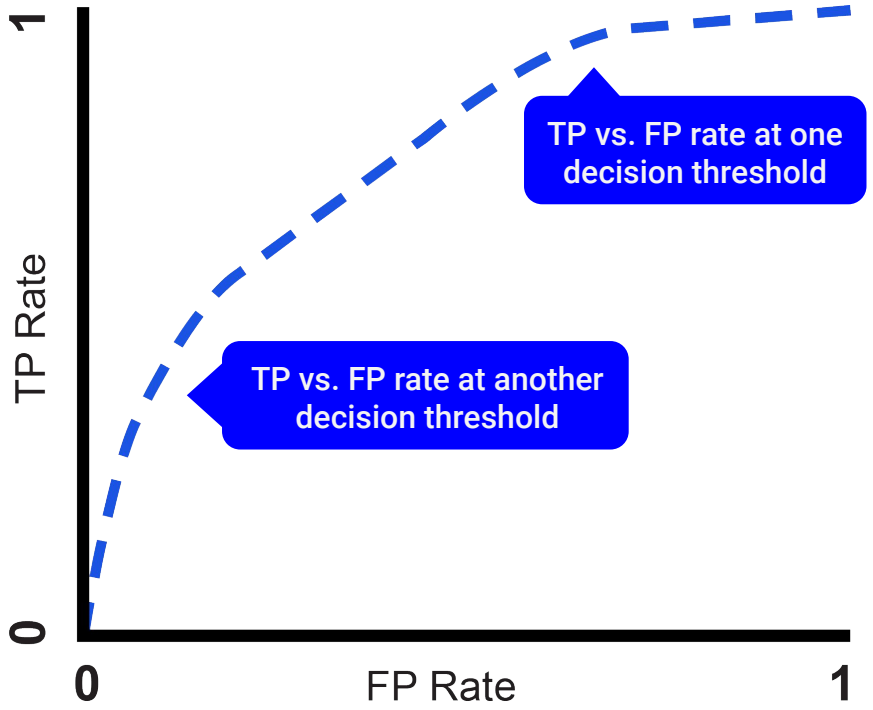
The **FPR** is calculated as follows:

$$FPR = \frac{FP}{FP + TN}$$

Introducing the ROC Curve and AUC

Every point in the ROC curve represents the TPR vs. FPR at different thresholds.

This image shows a typical ROC curve.



Understanding AUC

AUC

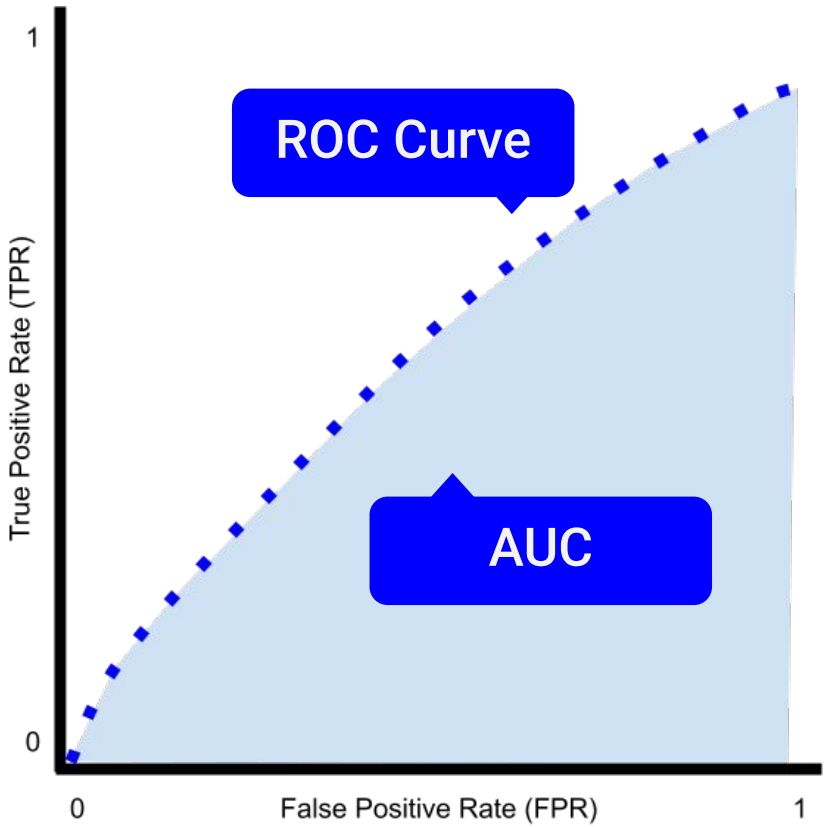
stands for

Area **U**nder the ROC **C**urve

Introducing the ROC Curve and AUC

Interpreting the ROC curve can be challenging.

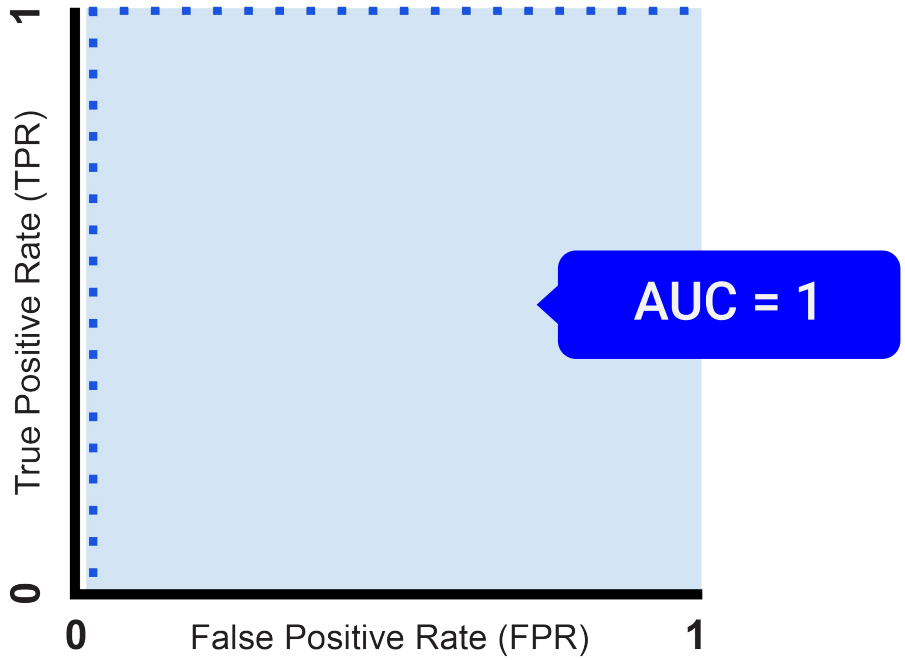
Fortunately, we have the AUC that measures the area below the entire ROC curve from (0,0) to (1,1).



The value of AUC ranges from 0 to 1.

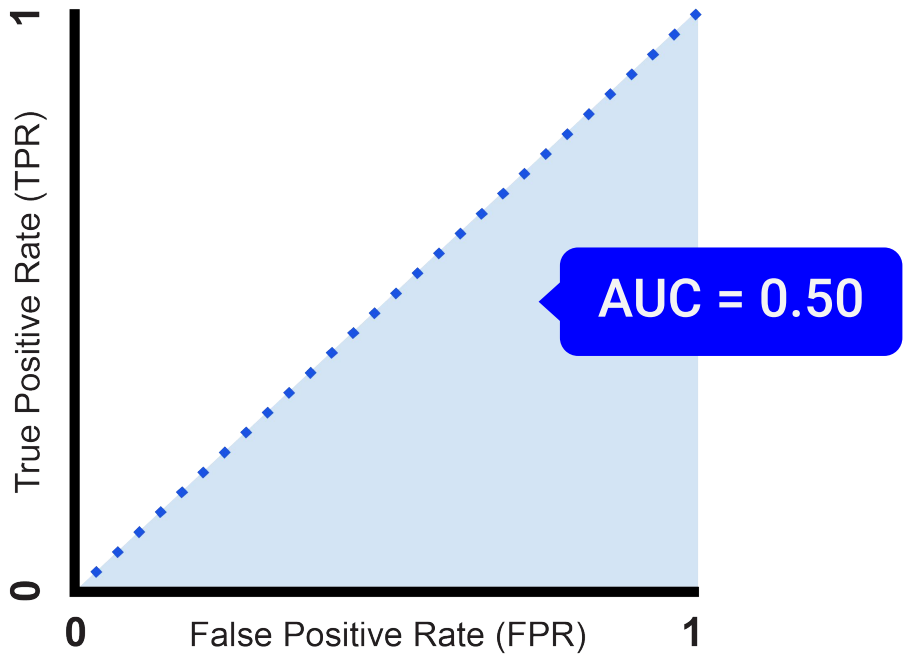
Introducing the ROC Curve and AUC

AUC = 1 is a paradox: it indicates that the model is perfect, but you may not trust your model because it might be overfitted.



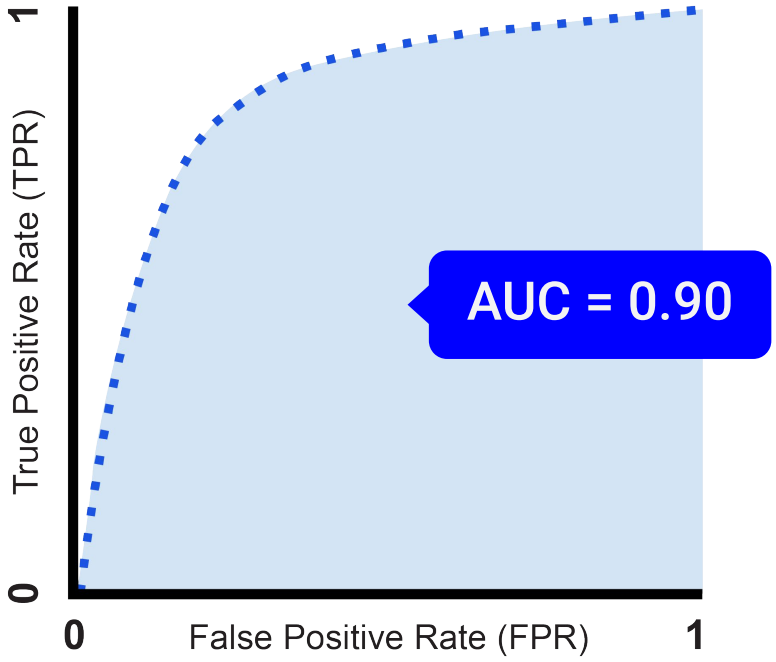
Introducing the ROC Curve and AUC

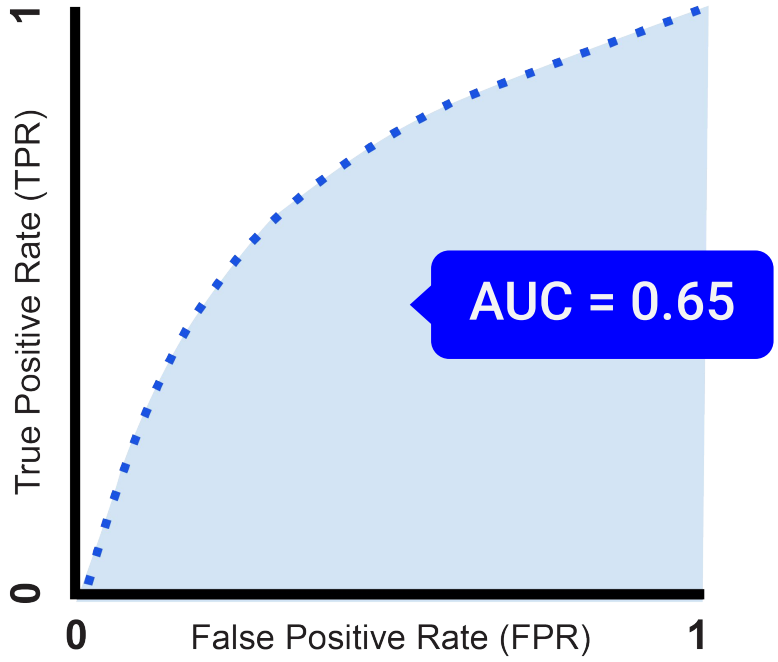
`AUC = 0.50` means that the model is unable to distinguish between positive and negative classes.



Introducing the ROC Curve and AUC

Ideally, we want AUC values ranging between 0 and 1. The higher the AUC, the better the model is at predicting 0s as 0s and 1s as 1s.





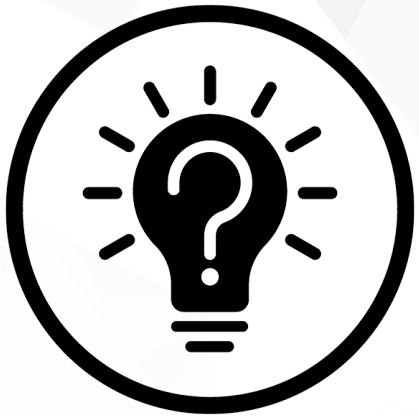
So, a model with an `AUC=0.90` may be better than a model with an `AUC=0.65`.

Intro to Recurrent Neural Networks (RNNs)

Intro to Recurrent Neural Networks (RNNs)

RNNs are able to remember the past, and their decisions are influenced by what they have learned in the past.

(Animated Cat)



What Are Recurrent Neural Networks Used For?

What Are Recurrent Neural Networks Used For?

RNNs are used for the following:

01 Natural language processing (NLP)

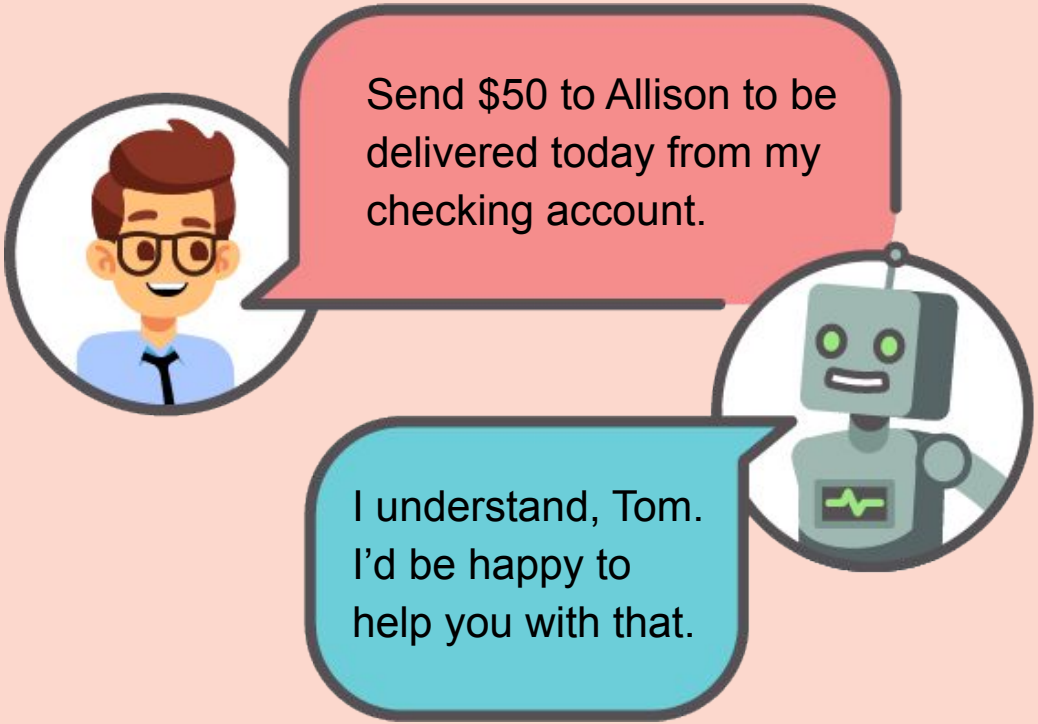
02 DNA sequences

03 Time series data

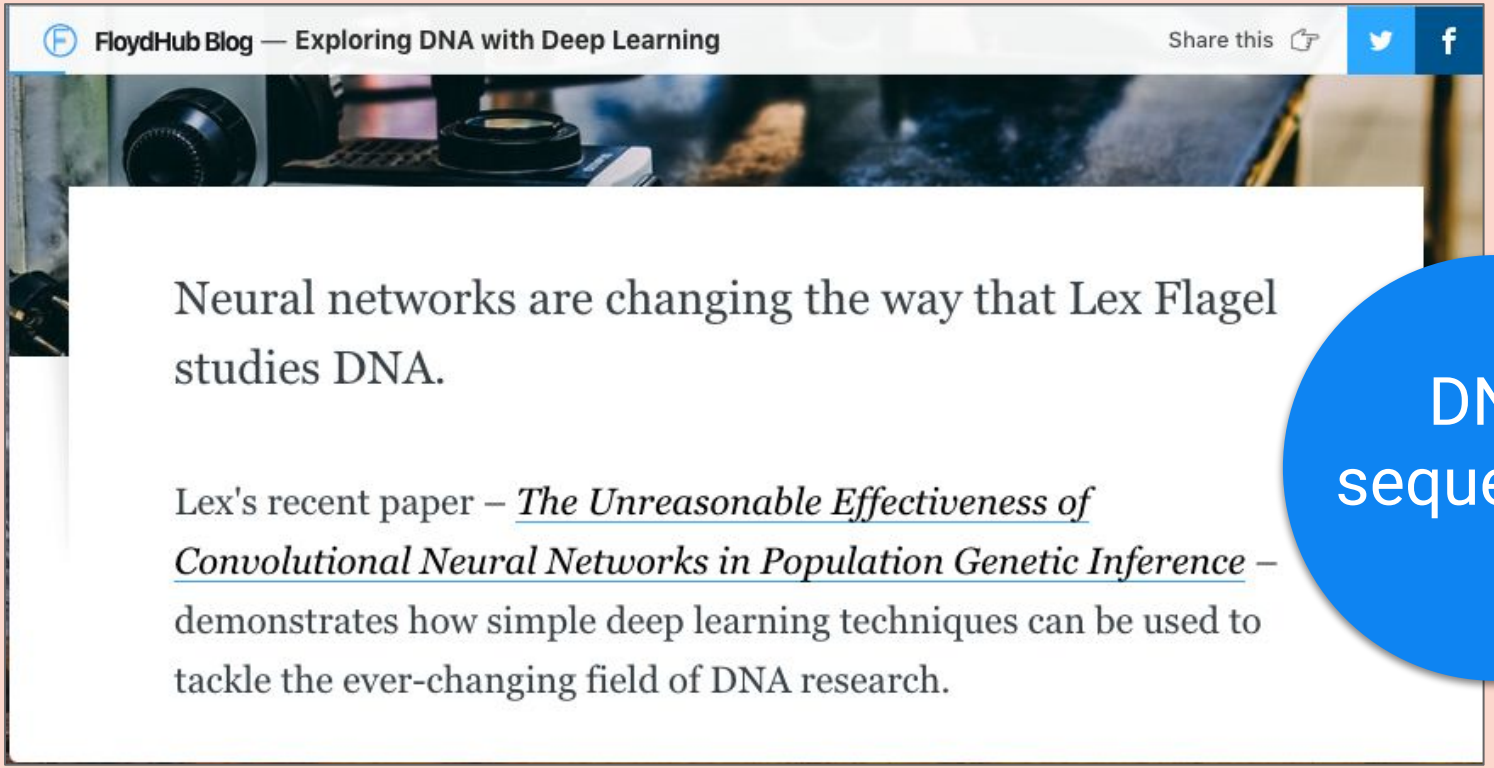
04 Music composition

What Are Recurrent Neural Networks Used For?

Natural Language Processing (NLP)



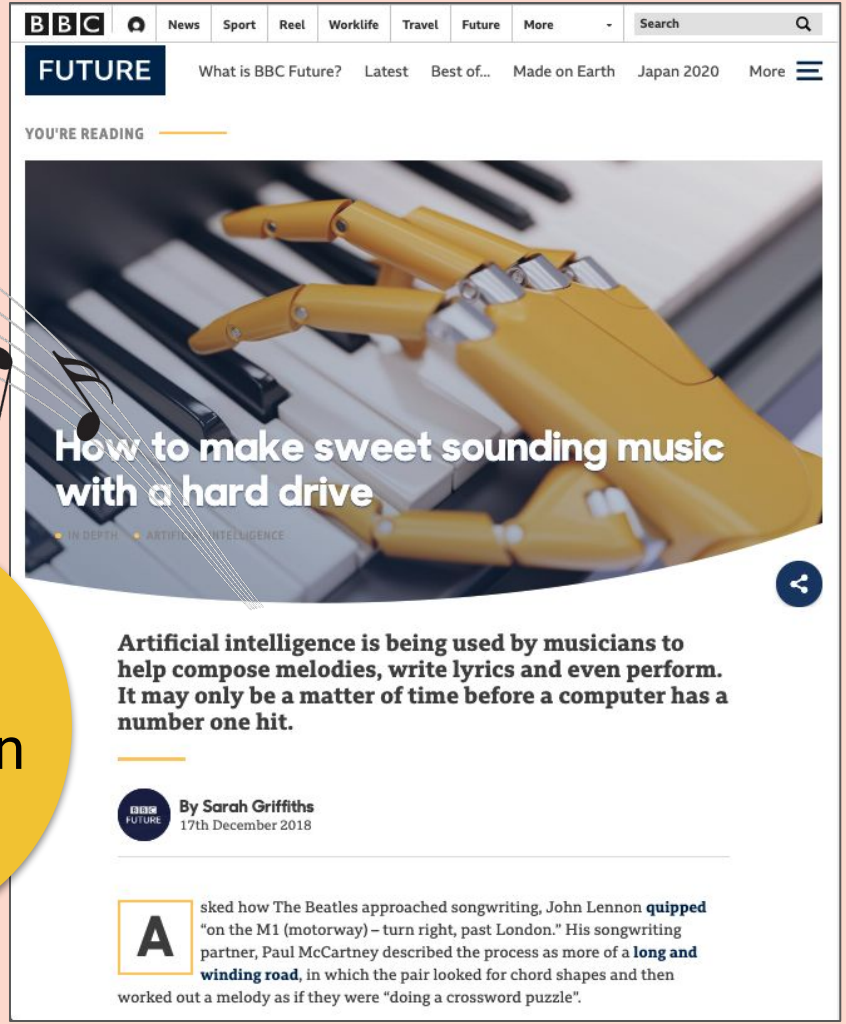
What Are Recurrent Neural Networks Used For?



DNA sequences

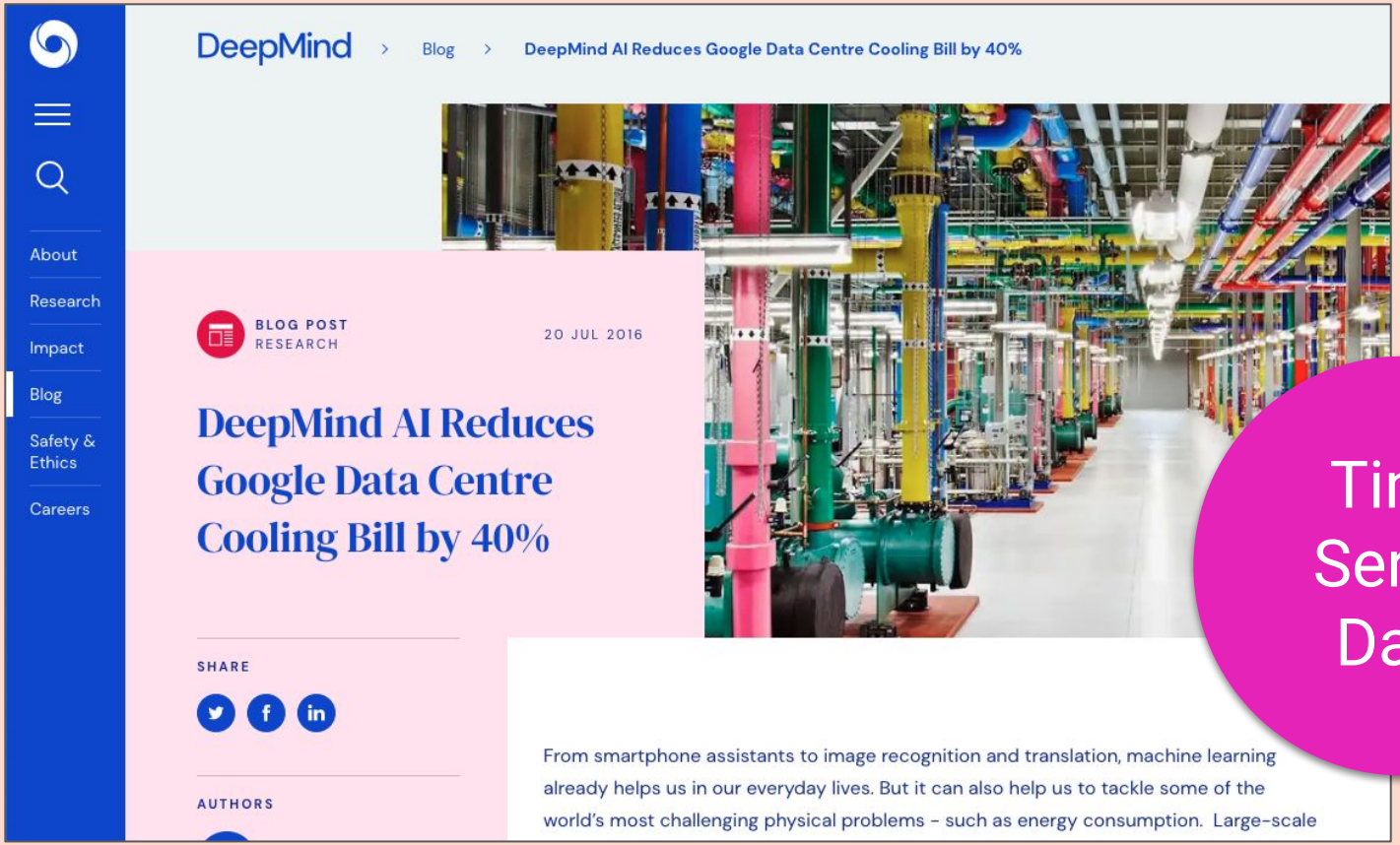
What Are Recurrent Neural Networks Used For?

Music Composition



www.bbc.com/future/

What Are Recurrent Neural Networks Used For?

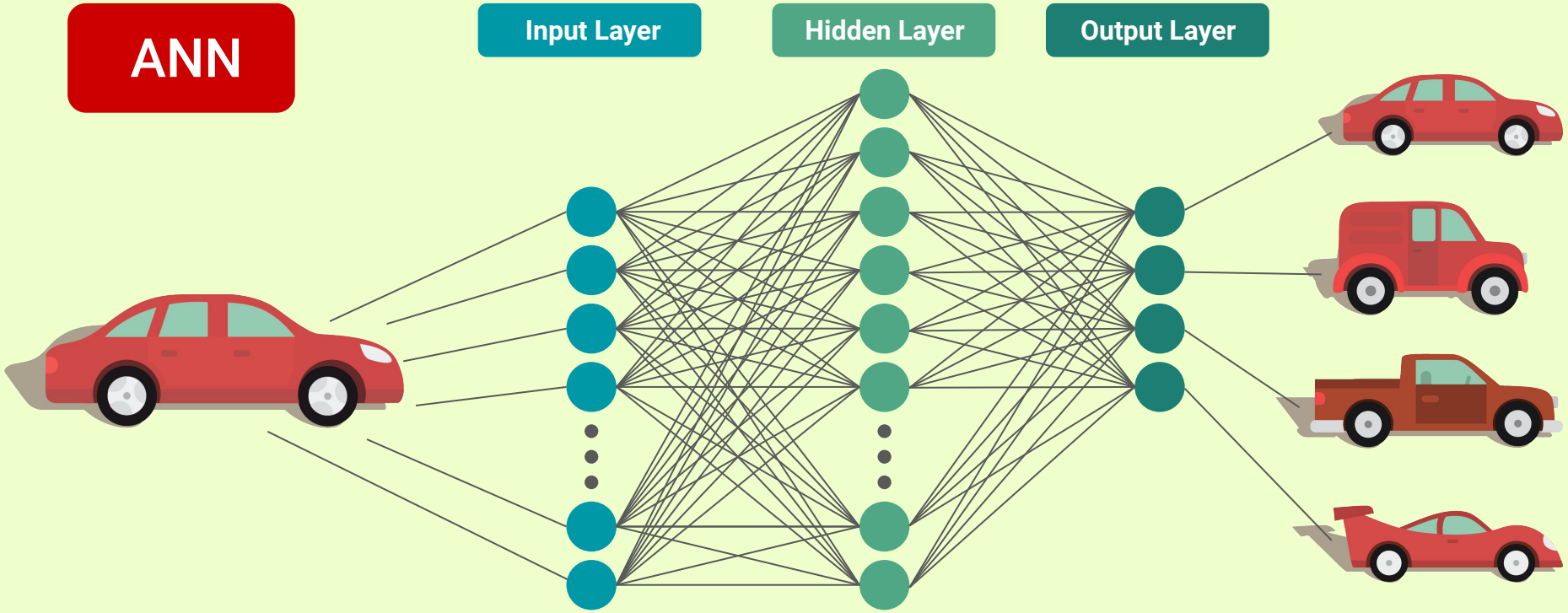


deepmind.com

ANNs vs. RNNs

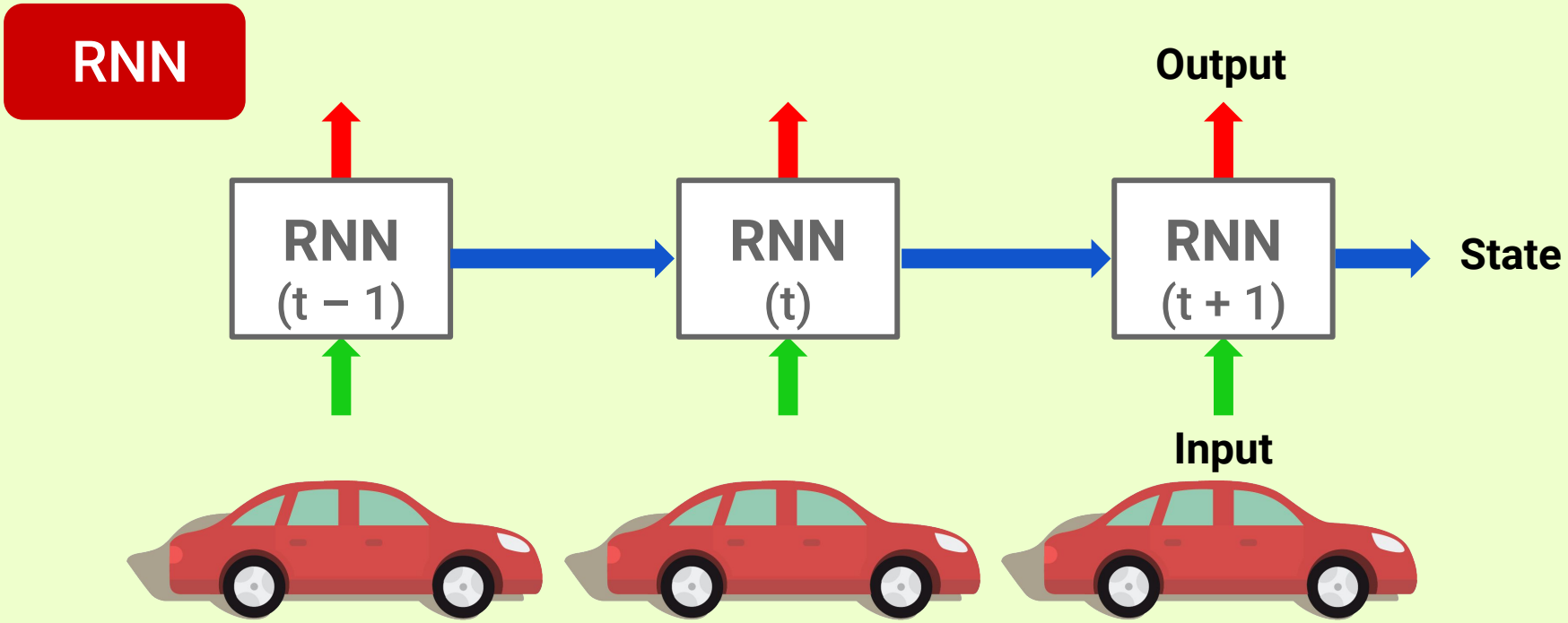
ANNs vs. RNNs

We can use **ANNs** to identify the type of car from a still image.
But can we predict the direction of a car in movement?



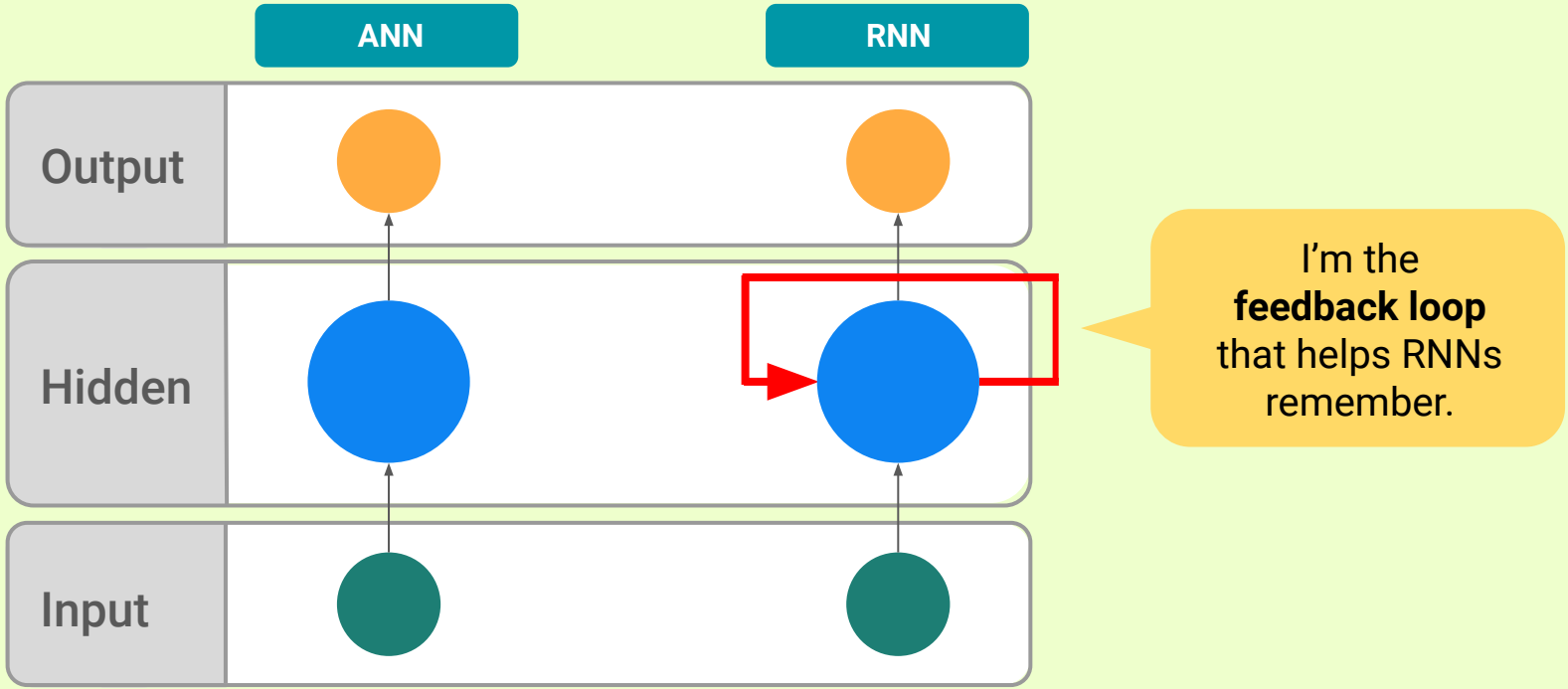
ANNs vs. RNNs

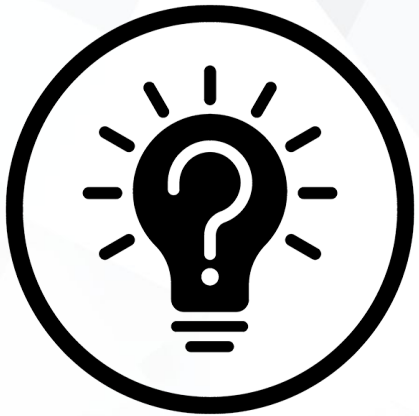
RNNs are good at modeling sequence data thanks to their **sequential memory.** Using RNNs, we can predict that the car is moving to the right.



ANNs vs. RNNs

RNNs are good at modeling sequence data thanks to their **sequential memory.** Using RNNs, we can predict that the car is moving to the right.

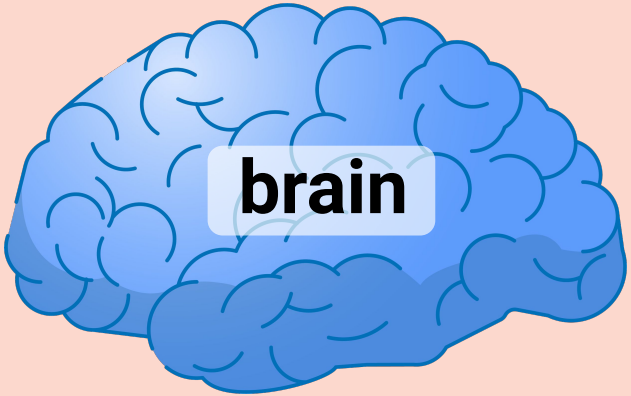




How Do RNNs Work?

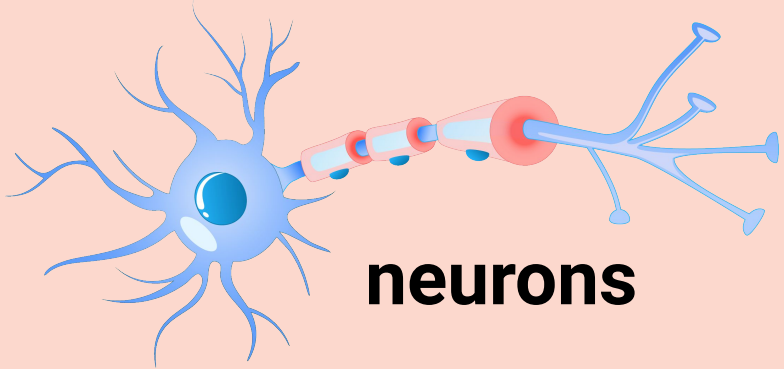
How Do RNNs Work?

When you read this sentence, your



is able to decode it and understand . . .

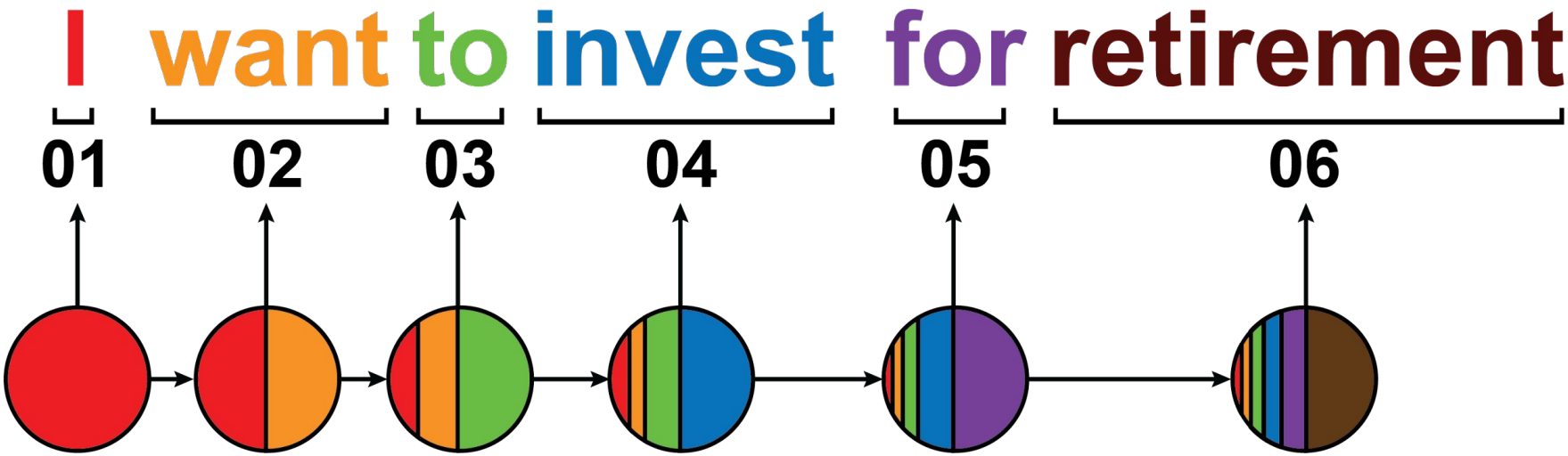
. . . because our



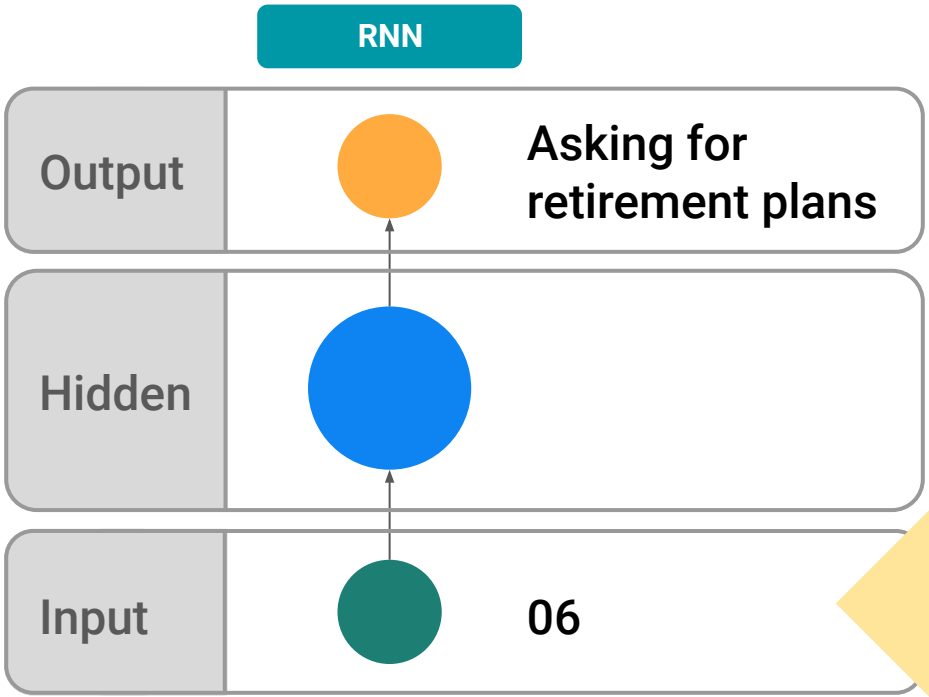
have memory like RNNs.

How Do RNNs Work?

The sentence is split into individual words. RNNs work sequentially, so we feed it one word at a time. By the final step, the RNN has encoded information from all the words in previous steps.



How Do RNNs Work?

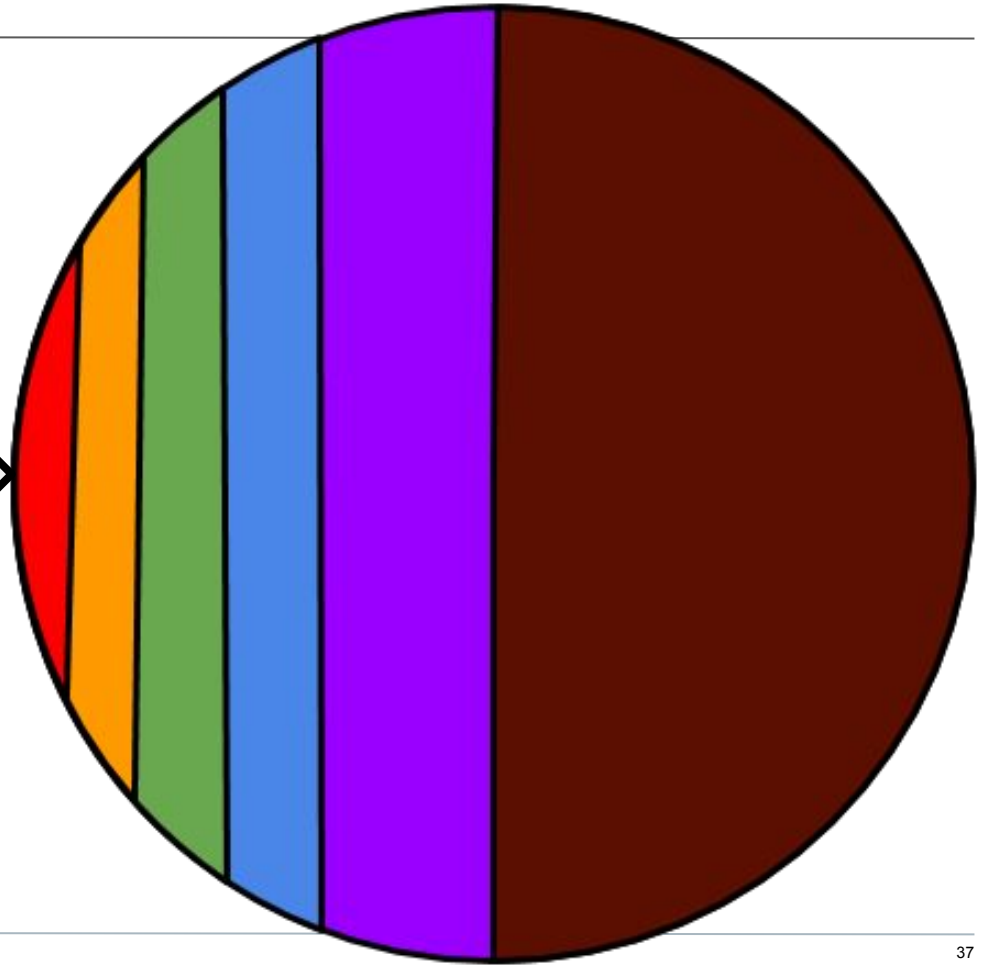


The final output was created from the rest of the sequence. To predict what the phrase means, we take the input and pass it to the feed-forward layer of the RNN to classify the intent.

RNNs Are Forgetful

RNNs only “remember”
the most recent few steps.

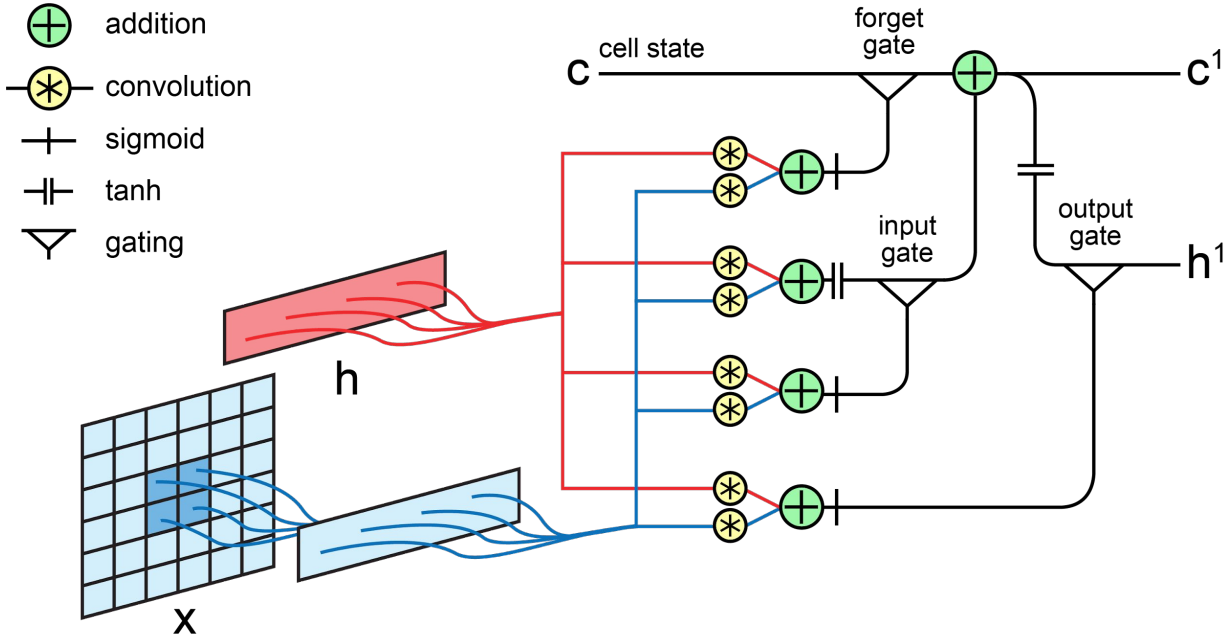
The vanishing gradient in the hidden states illustrates an issue with RNNs: **short-term memory**.



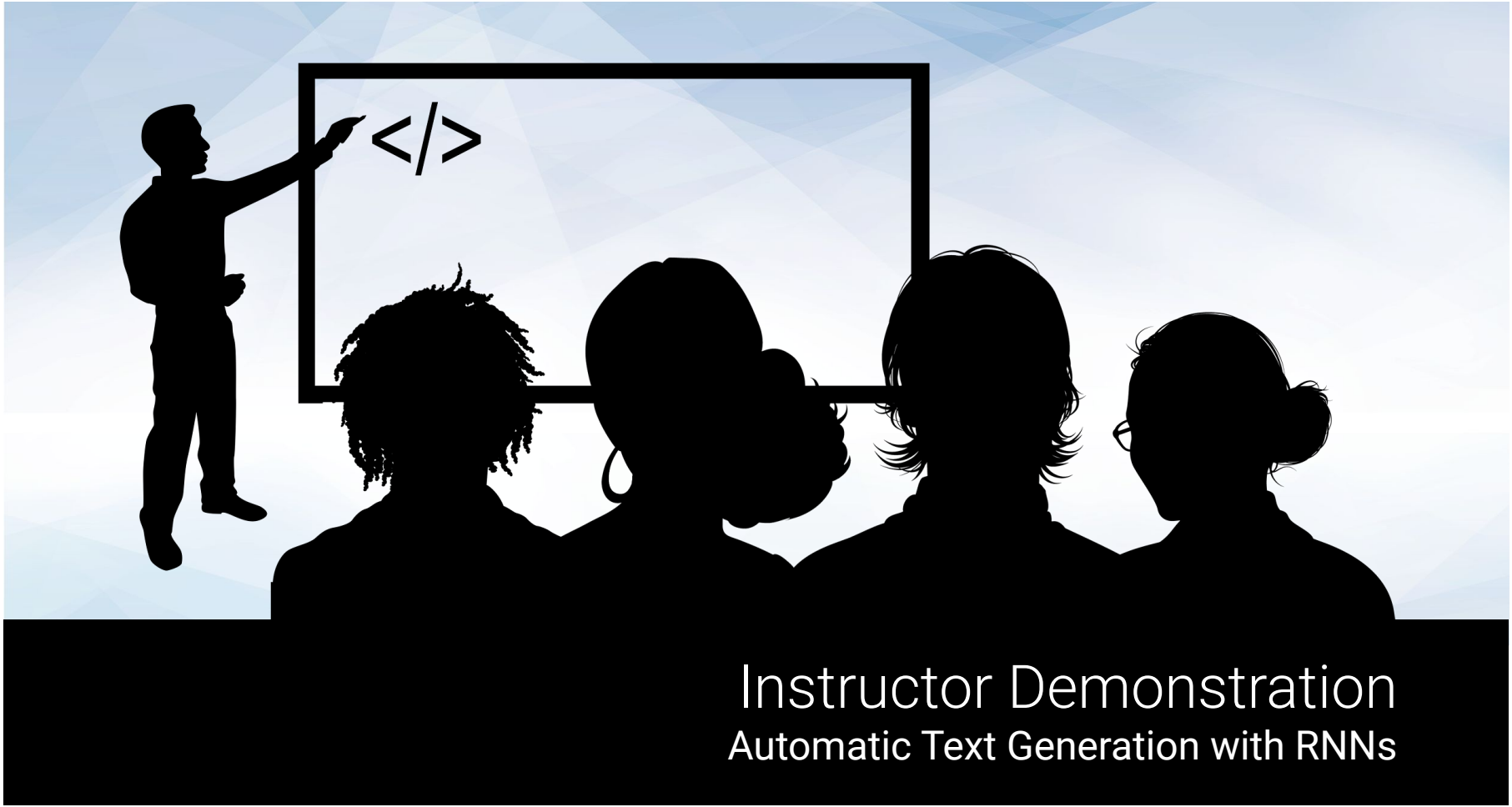
Long Short-Term Memory (LSTM)

LSTMs to the Rescue

LSTM (Long Short-Term Memory) RNNs are one solution for longer time windows. An LSTM RNN works like an original RNN, but it selects which types of longer-term events are worth remembering and which are okay to forget.



(Source)

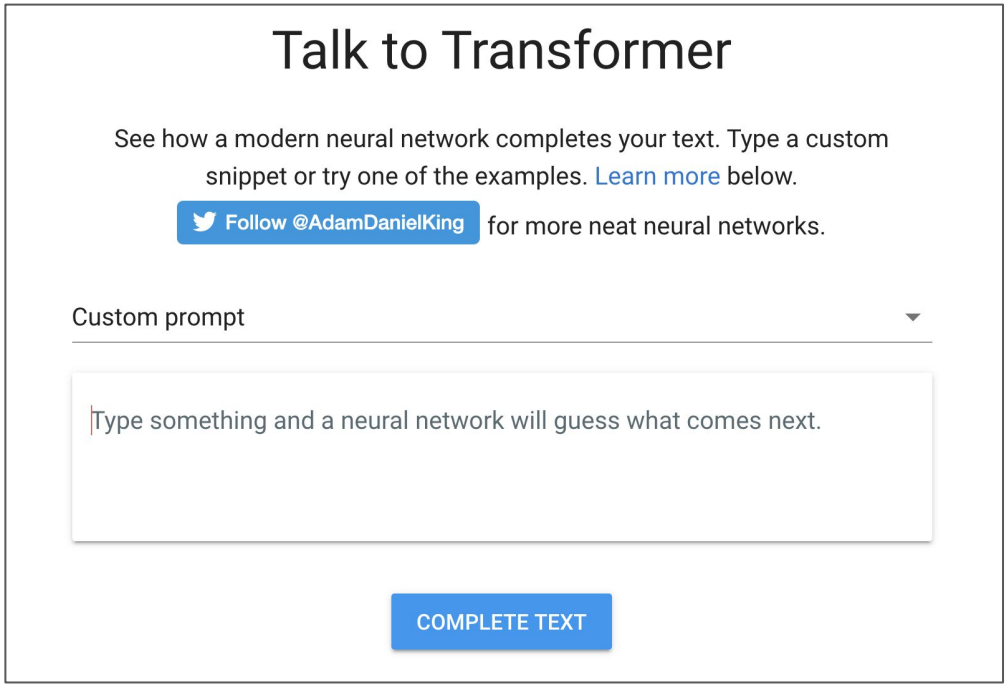


Instructor Demonstration

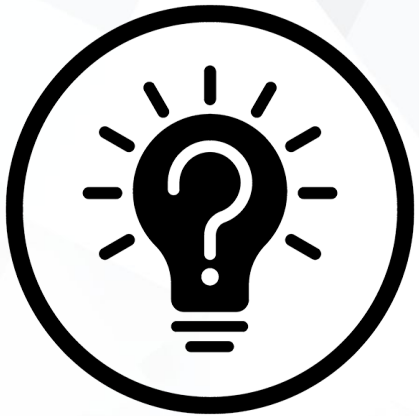
Automatic Text Generation with RNNs

Automatic Text Generation with RNNs

In this demo, we will explore how an RNN can be used to automatically generate text.



(Talk to Transformer)



Want to Learn More About RNNs?

Take a Look at This Recurrent Neural Networks Cheat Sheet

Shervine Amidi About Projects Teaching Blog About Afshine Amidi

Recurrent Neural Networks

- Overview
  - Architecture structure
  - Applications of RNNs
  - Loss function
  - Backpropagation
- Handling long term dependencies
  - Common activation functions
  - Vanishing/exploding gradient
  - Gradient clipping
  - GRU/LSTM
  - Types of gates
  - Bidirectional RNN
  - Deep RNN
- Learning word representation
  - Notations
  - Embedding matrix
  - Word2vec
  - Skip-gram
  - Negative sampling
  - GloVe
- Comparing words
  - Cosine similarity
  - t-SNE

View PDF version on GitHub

Would you like to see this cheatsheet in your native language? You can help us translating it on GitHub!

CS 230 - Deep Learning

English فارسی Français 日本語 한국어 Türkçe

Convolutional Neural Networks | Recurrent Neural Networks | Tips and tricks

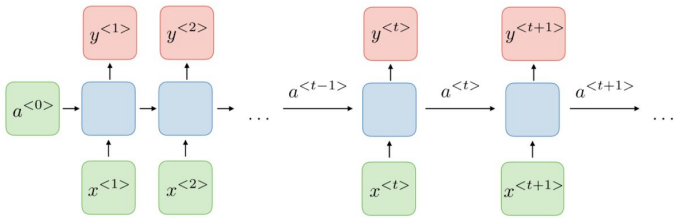
Recurrent Neural Networks cheatsheet

★ Star 3,793

By Afshine Amidi and Shervine Amidi

Overview

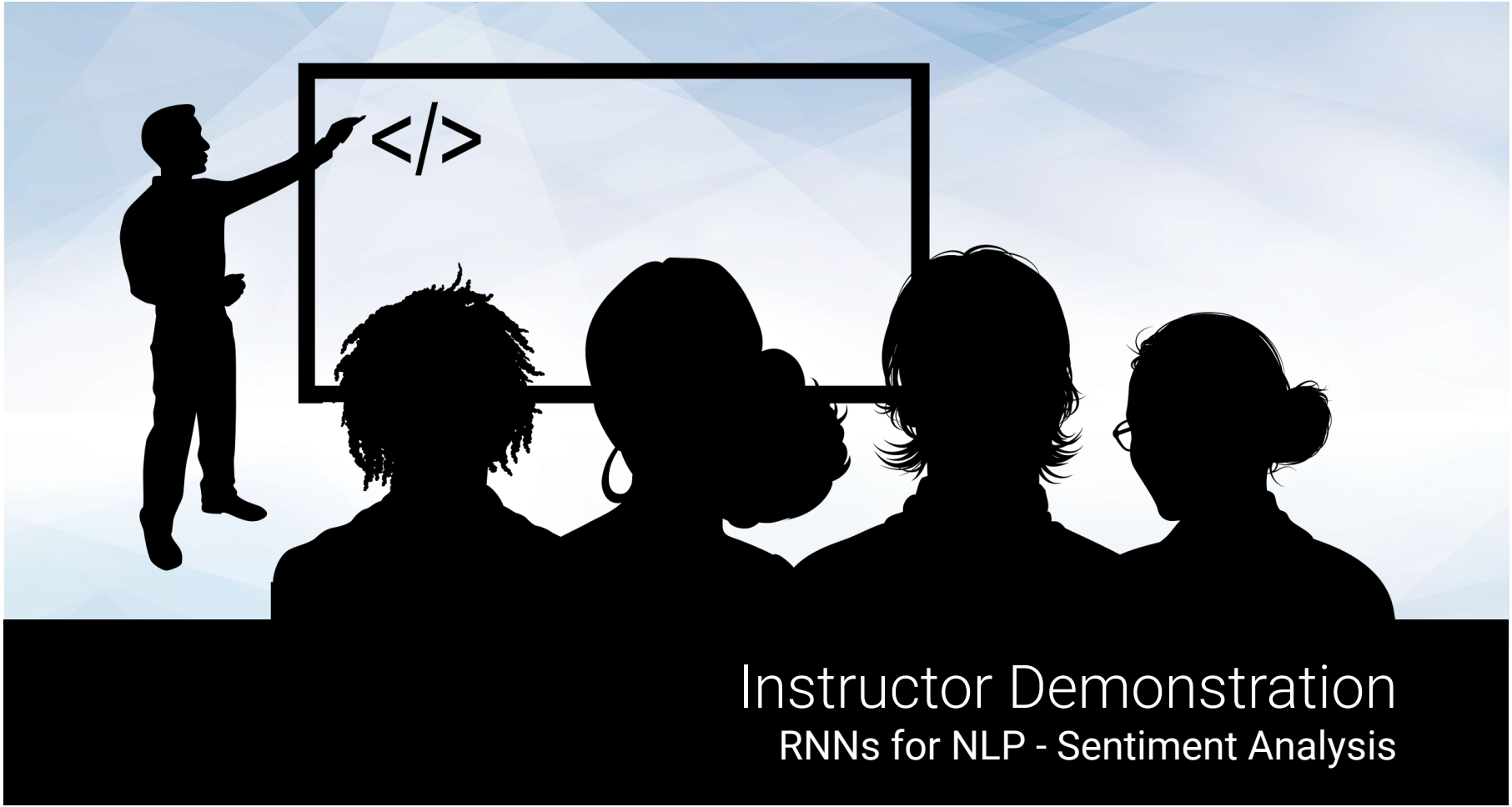
Architecture of a traditional RNN — Recurrent neural networks, also known as RNNs, are a class of neural networks that allow previous outputs to be used as inputs while having hidden states. They are typically as follows:



For each timestep t, the activation $a^{<t>}$ and the output $y^{<t>}$ are expressed as follows:

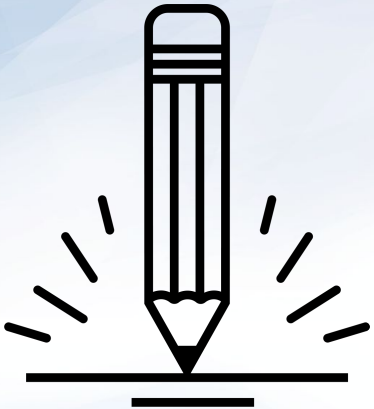
$$a^{<t>} = g_1(W_{aa}a^{<t-1>} + W_{ax}x^{<t>} + b_a) \quad \text{and} \quad y^{<t>} = g_2(W_{ya}a^{<t>} + b_y)$$

(Source)



Instructor Demonstration

RNNs for NLP - Sentiment Analysis



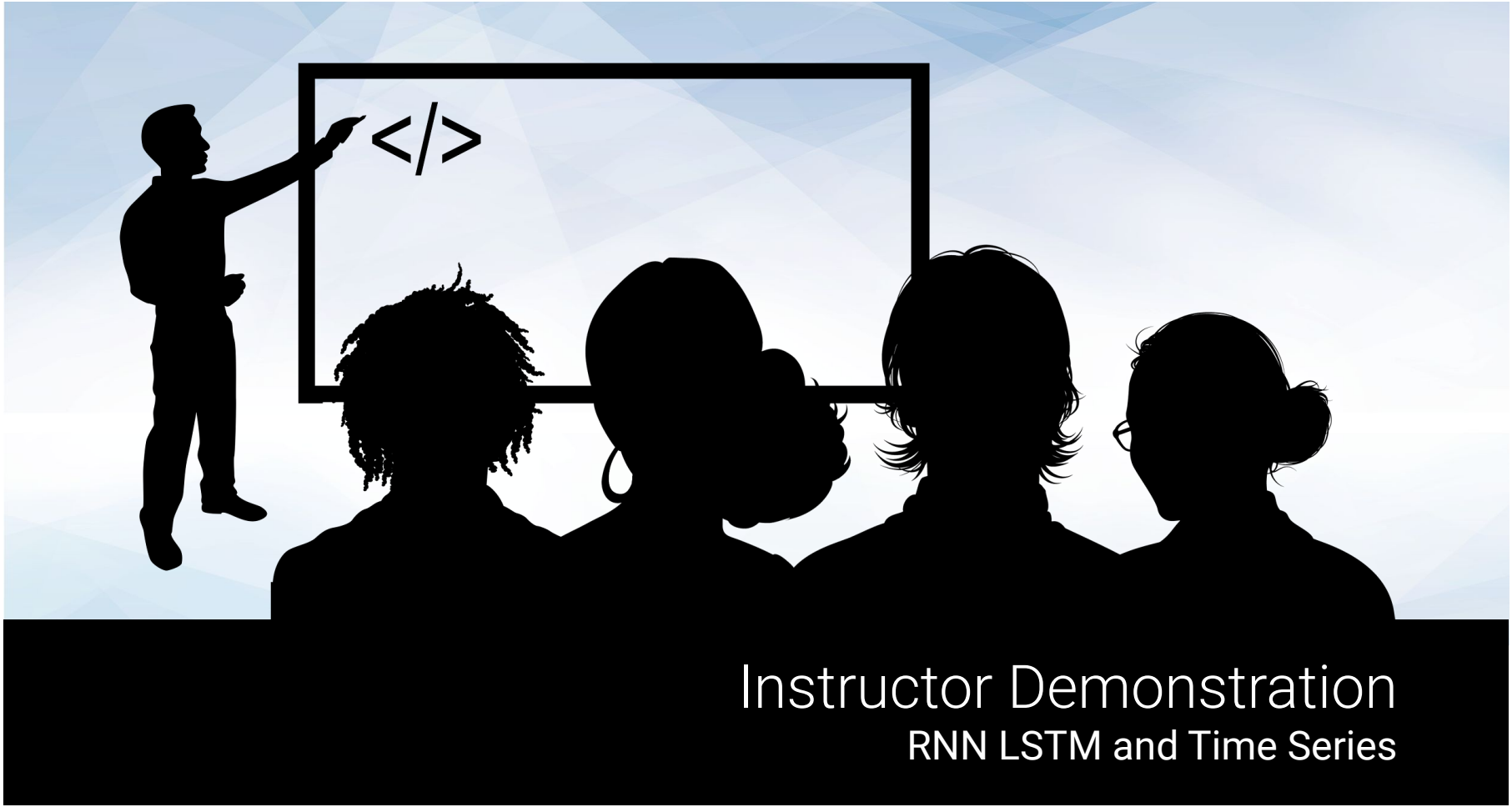
Activity: RNNs Vs. Vader

In this activity, you will use two different models to score sentiment.

Suggested Time:
40 minutes

Time’s Up! Let’s Review.


Break



Instructor Demonstration

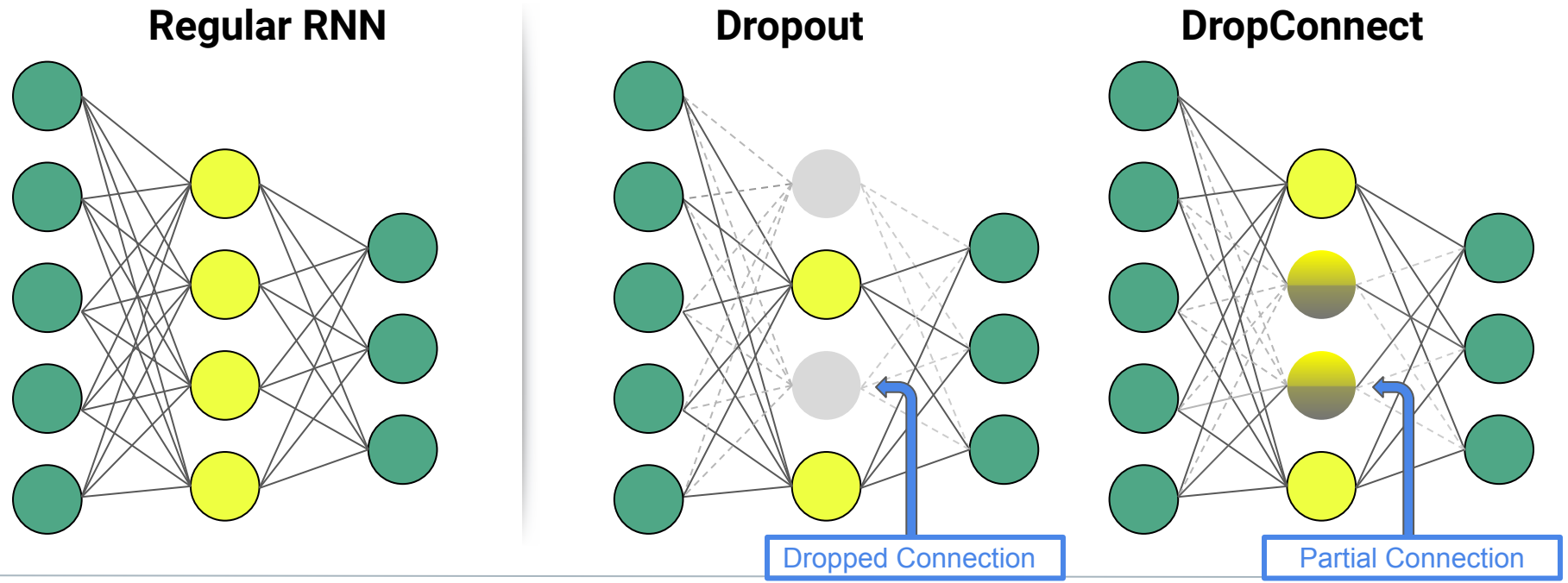
RNN LSTM and Time Series

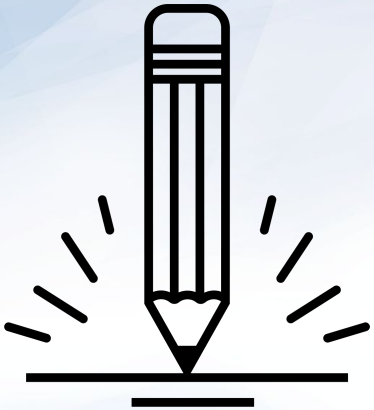
Introduction to Dropout

Introduction to Dropout

Dropout consists of removing units from the hidden layers by randomly selecting a fraction of the hidden nodes and setting their output to zero, regardless of the input.

A different subset of units is randomly selected every time we feed a training example.





Activity: Predicting Gold Closing Prices

In this activity you will gain hands-on experience building an RNN LSTM model for time-series data predicting gold closing prices.

Suggested Time:
30 minutes

Time’s Up! Let’s Review.

Questions?


The End.