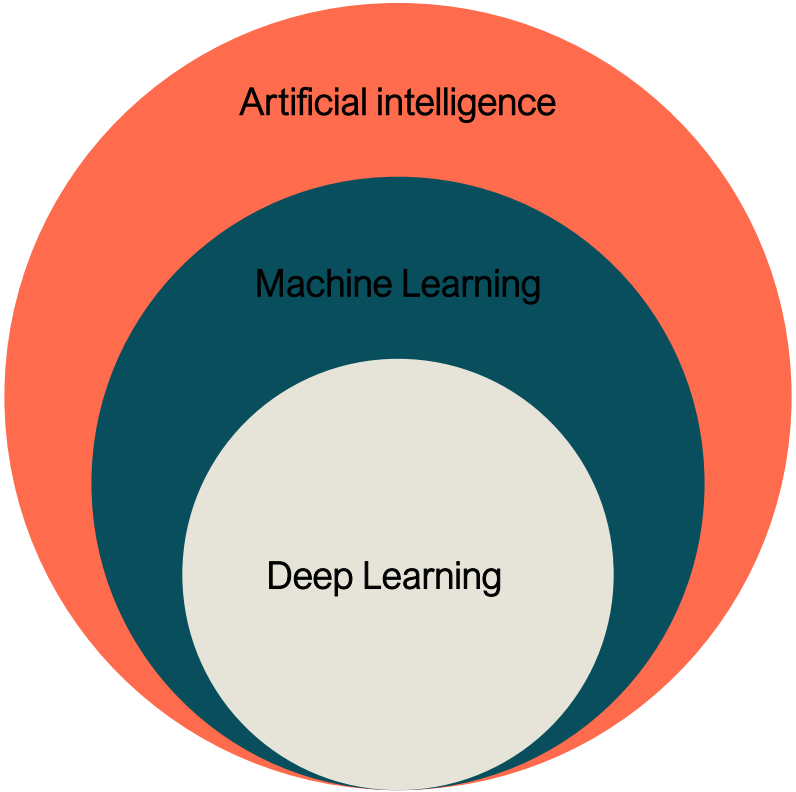
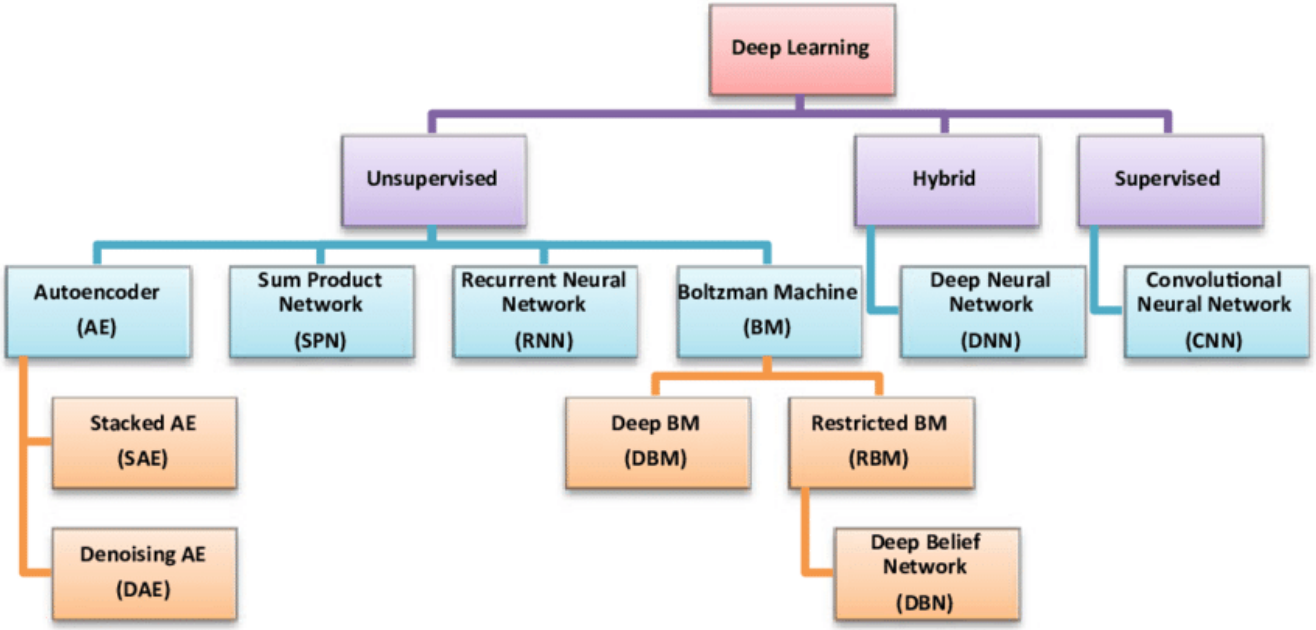
Real or Fake News Classifacation

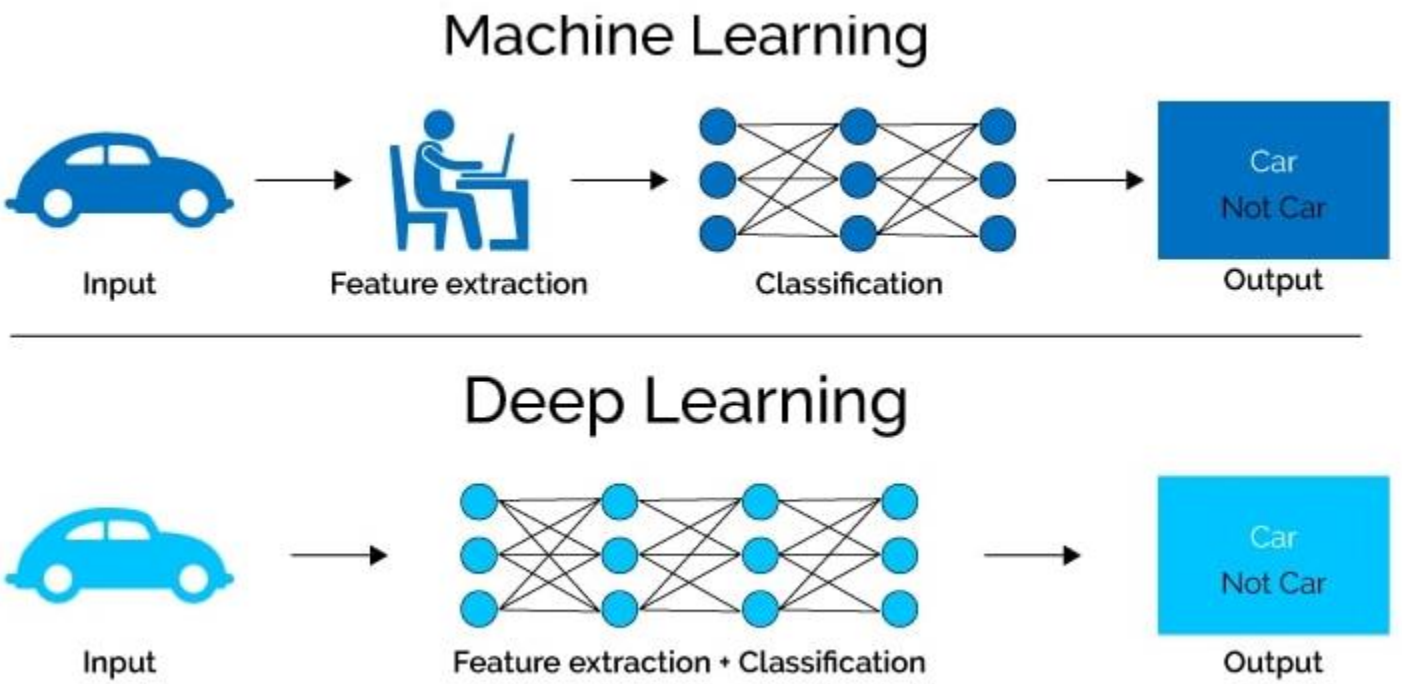
Deep Learning is a Subset of ML



Classification of Deep Learning Techniques



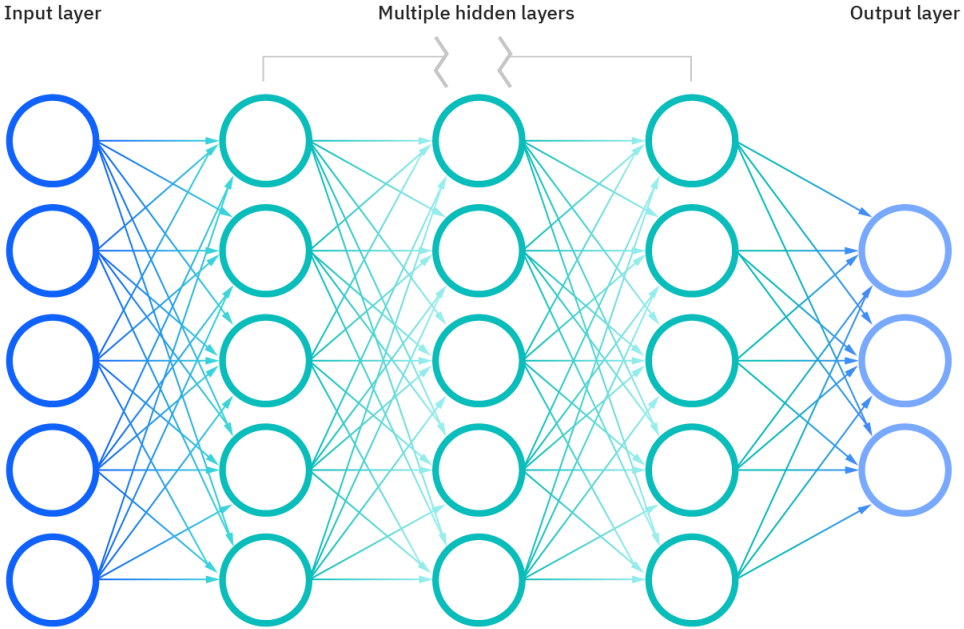
Machine Learning vs. Deep Learning



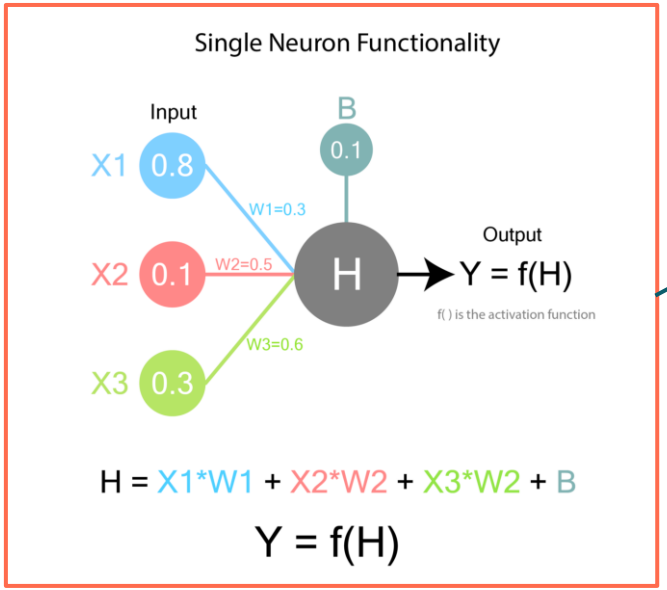
Machine Leaning vs. Deep Learning

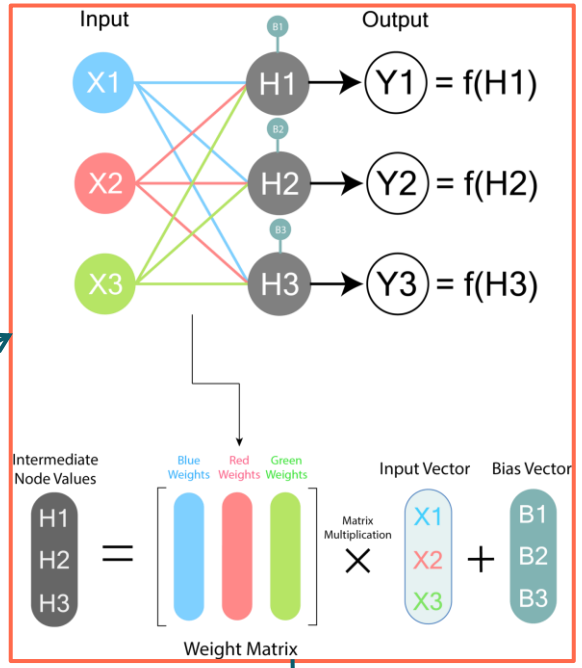
Machine learning	Deep learning
A subset of AI	A subset of machine learning
Can train on smaller data sets	Requires large amounts of data
Requires more human intervention to correct and learn	Learns on its own from environment and past mistakes
Shorter training and lower accuracy	Longer training and higher accuracy
Can train on a CPU (central processing unit)	Needs a specialized GPU (graphics processing unit) to train

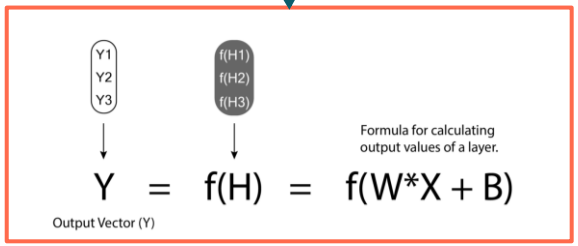
Deep Neural Network



Deep Neural Network

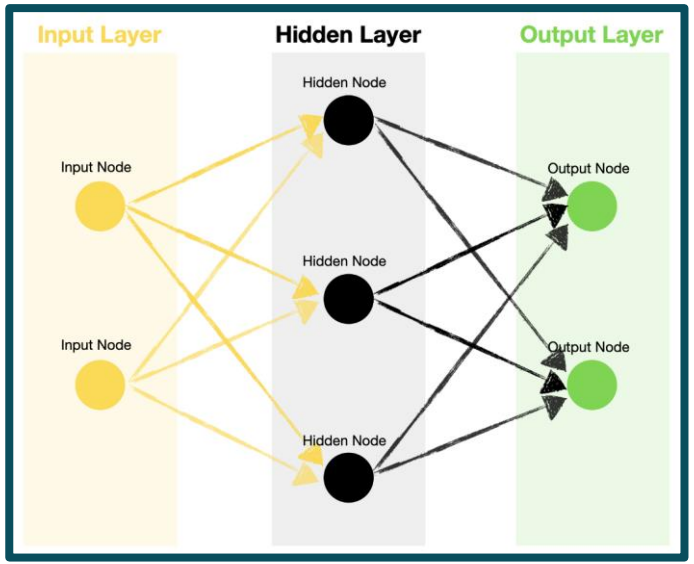




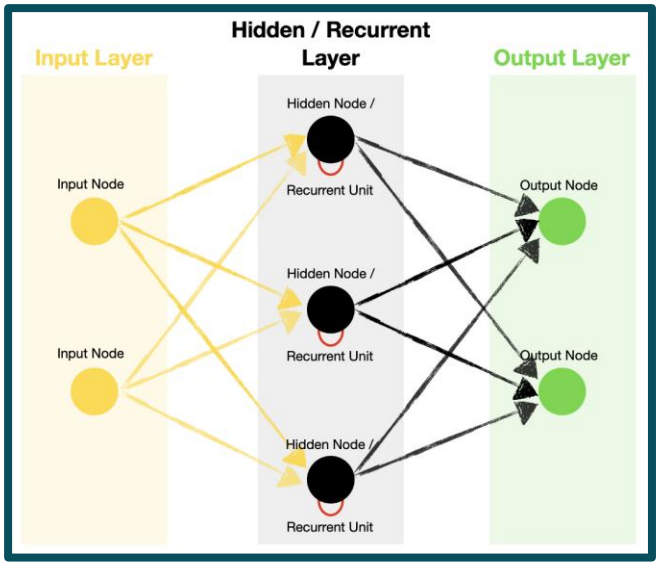


The Structure of Recurrent Neural Networks (RRNs)

Feed Forward Neural Network

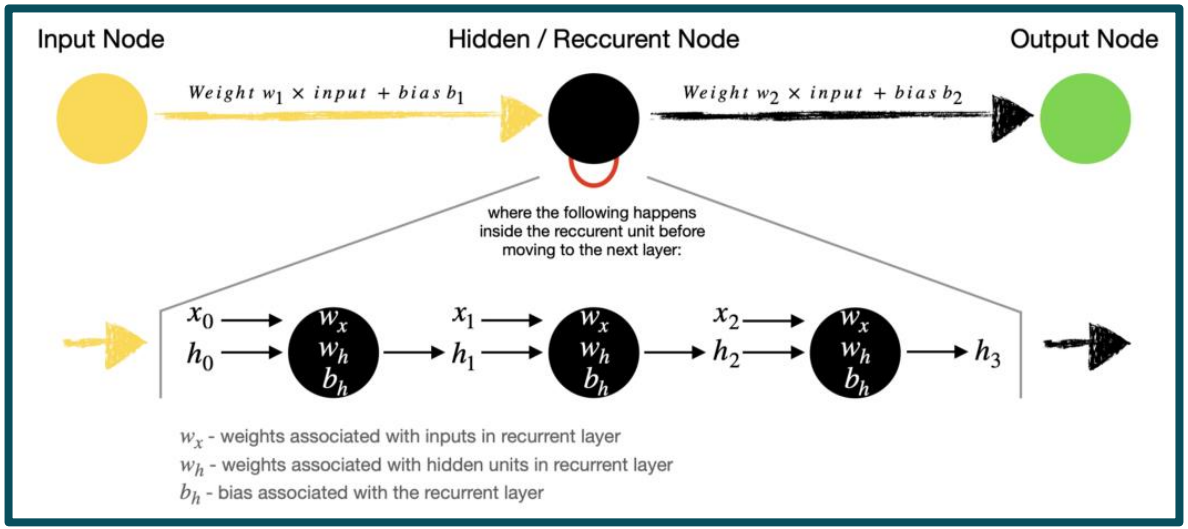


Recurrent Neural Network



The Structure of Recurrent Neural Networks (RRNs)

Timestep



Activation Function

$$Y = W_1 * X_1 + W_2 * X_2$$

represents a linear relationship between Y and X1,X2.

if we look at real world data, we realize this is not desirable because data often has nonlinear relationships between the input and output variables.

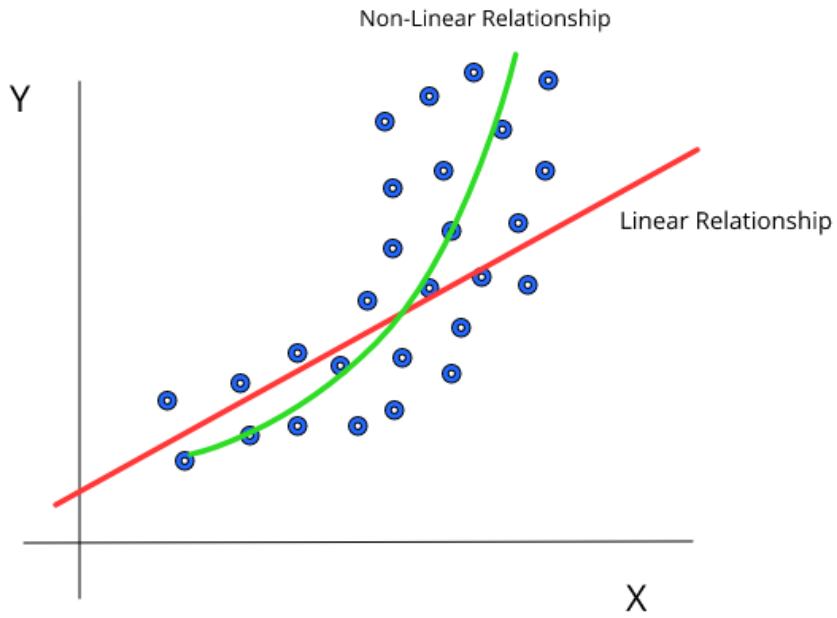
DEEP NEURAL NETWORK

Activation Function

The diagram represents a dataset which shows a non-linear relationship.

If we try to fit a linear relationship on the data, we will end up with the red line, which is not a very accurate representation of the data.

If our relationship can be non-linear, we are able to get the green line, which is much better



Activation Function

Without Activation Function

$$y = \sum_{i=0}^{n}(W_i * X_i) + B$$

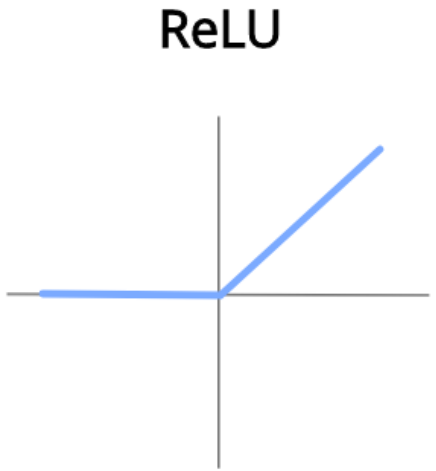
With Activation Function

$$y = f(\sum_{i=0}^{n}(W_i * X_i) + B)$$

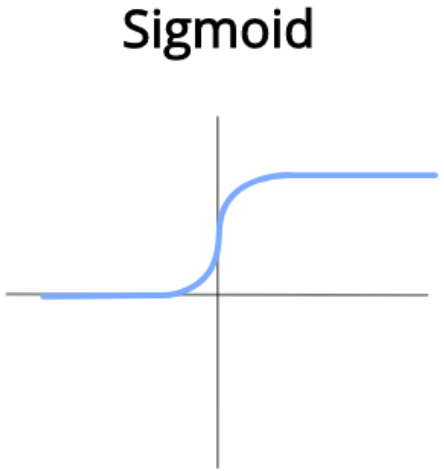
IF the activation function is itself nonlinear. Hence all we have to do is keep some nonlinear function as the activation function for each neuron and our neural network is now capable of fitting on nonlinear data.

Activation Function

ReLU: stands for Rectified Linear Unit. It essentially becomes an identity function (y = x) when x ≥ 0 and becomes 0 when x < 0. This is a very widely used activation function because its a nonlinear function and it is very simple.



Sigmoid: is essentially a function bounded between 0 and 1. It will become 0 for values which are very negative and 1 for values which are very positive. Hence this function squishes values which are very high or very low to values between 0 and 1.This function is usually used at the last layer when we need values which are binary (0 or 1).



Why Is It Important To Initialize Our Weights?

1. If we don't initialize our weights, our neural network will not be able to learn anything and will be unable to make any predictions.

2. If we initialize our weights randomly, our neural network will be able to learn faster and be more accurate when making predictions.

Different Weight Initialization Techniques

1- Zero Initialization (Initialized all weights to 0)

The derivative weight loss function is the same for every weight in W[I], thus all weights have the same value. This makes hidden layers symmetric, and this process continues for all the n iterations. Thus, initialized weights with zero make your network no better than a linear model.

Note: setting biases to 0 will not create any problems as non-zero weights take care of breaking the symmetry and even if bias is 0, the values in every neuron will still be different.

The Symmetry Problem

When some machine learning models have all the weights initialized to the same value, it can be difficult or impossible for the weights to differ as the model is trained. This is the “symmetry”.

Solution: Initializing the model to small random values breaks the symmetry and allows different weights to learn independently of each other → Symmetry breaking

Different Weight Initialization Techniques

2- Random Initialization (Initialized weights randomly)

- This technique tries to address the problems of zero initialization since it prevents neurons from learning the same features of their inputs since our goal is to make each neuron learn different functions of its input and this technique gives much better accuracy than zero initialization.

- In general, it is used to break the symmetry. It is better to assign random values except 0 to weights.

What happens if the weights initialized randomly can be very high or very low?”

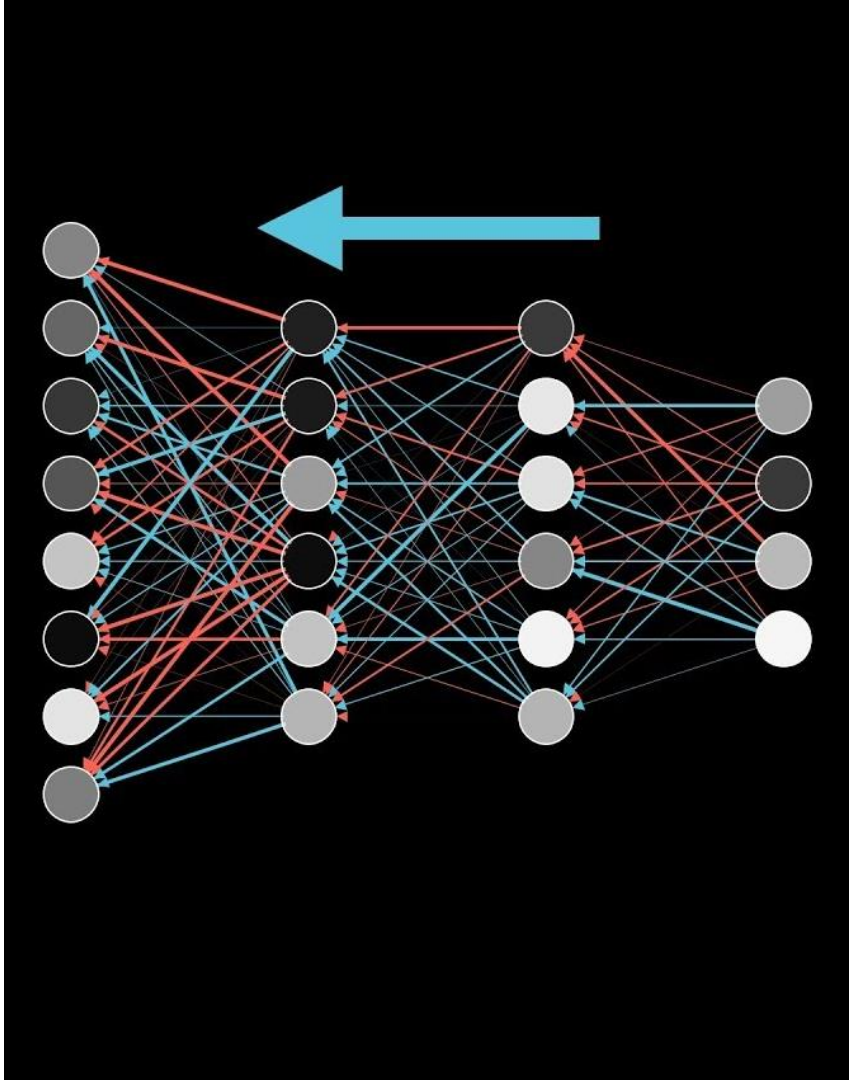
The problems are Vanishing and Exploding gradient descent. To solve these issues, one solution could be to initialize the parameters carefully.

MODEL BUILDING

Backpropagation

Back propagation: allows the data to flow from the output backward while updating the parameters (weights and biases)

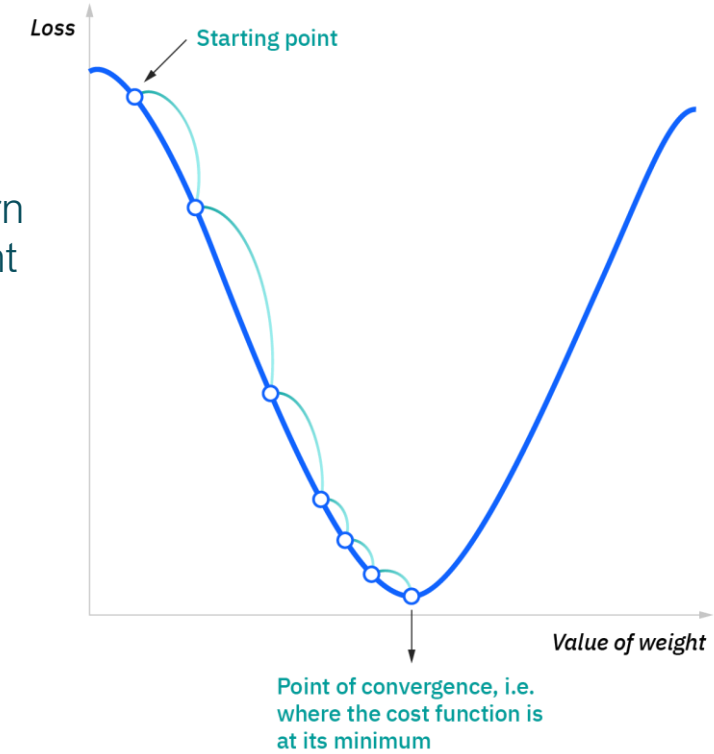
Definition: Back-Propagation is a method for supervised learning used by NN to update parameters to make the networks' predictions more accurate. The parameter optimization process is achieved using an optimization algorithm called gradient decent.



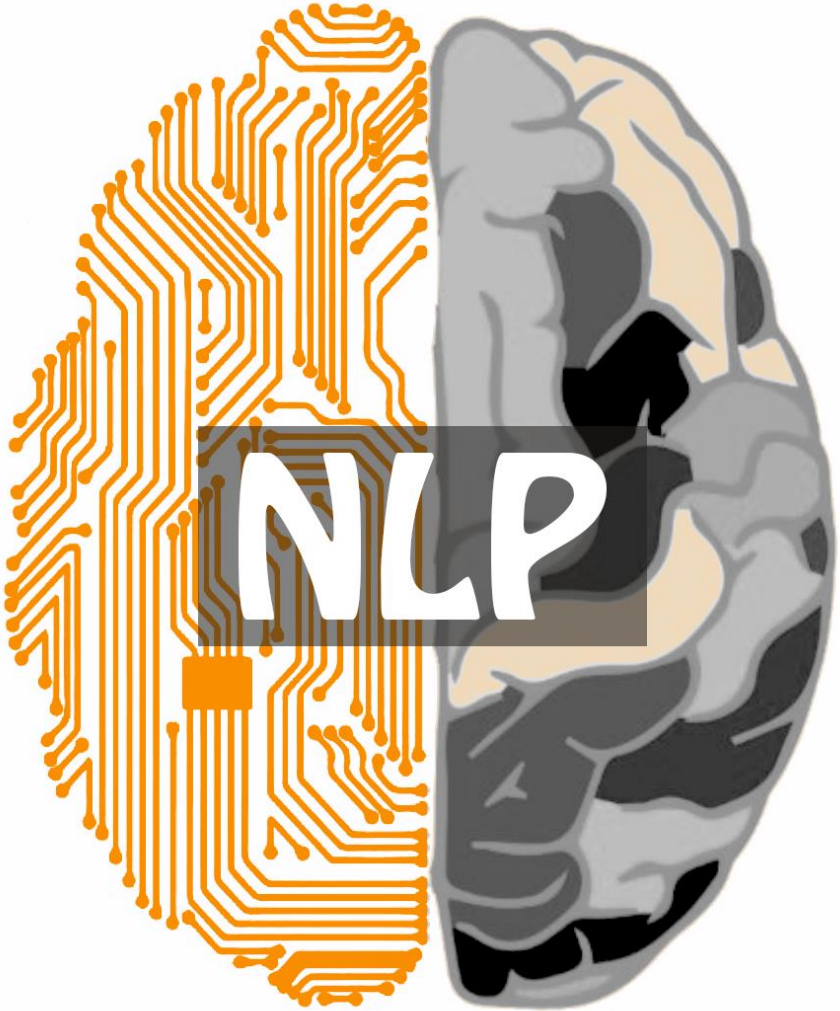
MODEL BUILDING

Gradient Decent

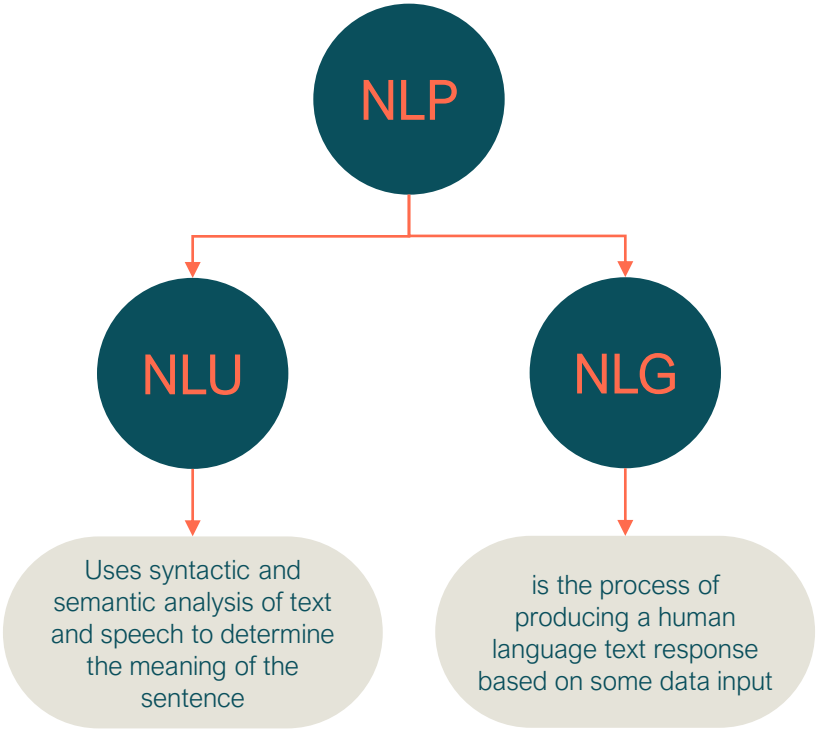
Gradient descent is an optimization algorithm that is commonly-used to train machine learning models and neural networks. Training data helps these models learn over time, and the cost function within gradient descent specifically acts as a barometer, gauging its accuracy with each iteration of parameter updates. Until the function is close to or equal to zero, the model will continue to adjust its parameters to yield the smallest possible error.



Natural language processing concerned with giving computers the ability to understand text and spoken words in much the same way human beings can.



Aim of Natural Language Processing



NLP Challenges

Words can have different meanings. Slangs can be harder to put out contextual. And certain languages are just hard to feed in, owing to the lack of resources. Despite being one of the more sought-after technologies, NLP comes with the following rooted and implementation AI challenges.

- Lack of Context
- Ambiguity

NLP Steps

1- Data collection

2- Data exploration and visualization

3- Data pre-processing

4- Model building

5- Model evaluation

Fake Dataset

The dataset contain a set of fake news with 4 columns and 23481 rows:

1- title

2- text

3- subject

4- date

title	text	subject	date
Donald Tr	Donald Trump just couldn t wish all Americans a H	News	December 31, 2017
Drunk Bra	House Intelligence Committee Chairman Devin Nu	News	December 31, 2017
Sheriff Da	On Friday, it was revealed that former Milwaukee	News	December 30, 2017
Trump Is	On Christmas day, Donald Trump announced that	News	December 29, 2017
Pope Fran	Pope Francis used his annual Christmas Day mess	News	December 25, 2017
Racist Ala	The number of cases of cops brutalizing and killin	News	December 25, 2017
Fresh Off	Donald Trump spent a good portion of his day at	News	December 23, 2017
Trump Sa	In the wake of yet another court decision that der	News	December 23, 2017
Former Cl	Many people have raised the alarm regarding the	News	December 22, 2017
WATCH: E	Just when you might have thought we d get a brea	News	December 21, 2017
Papa John	A centerpiece of Donald Trump s campaign, and r	News	December 21, 2017
WATCH: F	Republicans are working overtime trying to sell th	News	December 21, 2017
Bad News	Republicans have had seven years to come up wit	News	December 21, 2017
WATCH: L	The media has been talking all day about Trump a	News	December 20, 2017
Heiress Ta	Abigail Disney is an heiress with brass ovaries wh	News	December 20, 2017
Tone Dea	Donald Trump just signed the GOP tax scam into l	News	December 20, 2017
The Interi	A new animatronic figure in the Hall of Presidents	News	December 19, 2017
Mueller S	Trump supporters and the so-called president s fa	News	December 17, 2017
SNL Hilari	Right now, the whole world is looking at the shock	News	December 17, 2017
Republica	Senate Majority Whip John Cornyn (R-TX) thought	News	December 16, 2017
In A Hear	It almost seems like Donald Trump is trolling Ame	News	December 16, 2017
KY GOP St	In this #METOO moment, many powerful men are	News	December 13, 2017
Meghan M	As a Democrat won a Senate seat in deep-red Ala	News	December 12, 2017
CNN CALL	Alabama is a notoriously deep red state. It s a pla	News	December 12, 2017
White Ho	A backlash ensued after Donald Trump launched	News	December 12, 2017
Despicabl	Donald Trump is afraid of strong, powerful wome	News	December 12, 2017
Accused C	Ronald Reagan is largely seen as the Messiah of th	News	December 11, 2017
WATCH: F	Judge Jeanine Pirro has continued her screamy r	News	December 10, 2017
Liberal Gr	Donald Trump held a rally for Alabama Senate ca	News	December 9, 2017
Don Jr. Tr	When Sen. Al Franken (D-MN) announced his plan	News	December 7, 2017
BREAKING	In America, we have been having a conversation a	News	December 7, 2017
Watch: Is	New questions are being asked about President D	News	December 7, 2017

Real Dataset

The dataset contain a set of real news with 4 columns and 21417 rows :

1- title

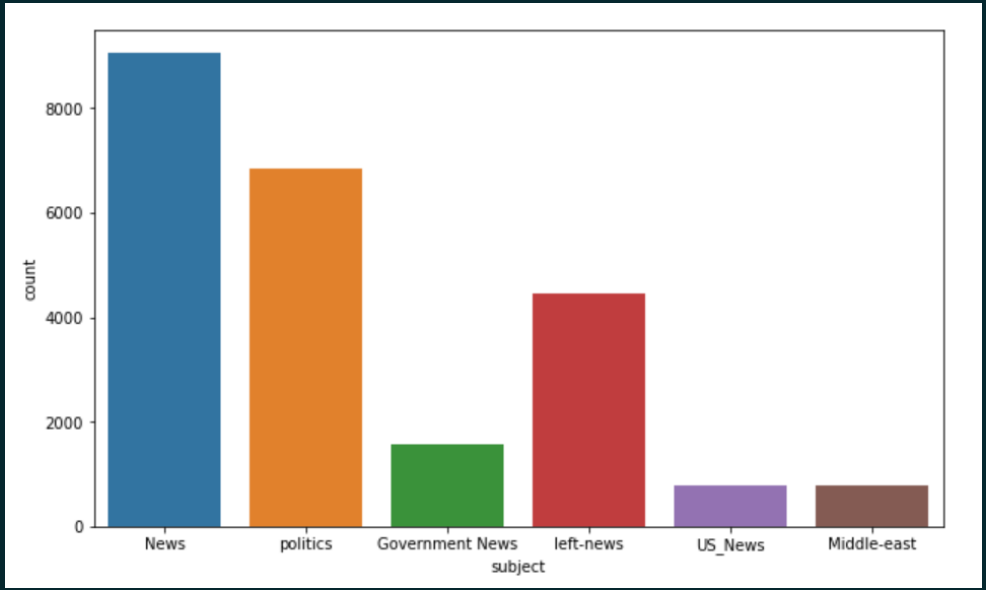
2- text

3- subject

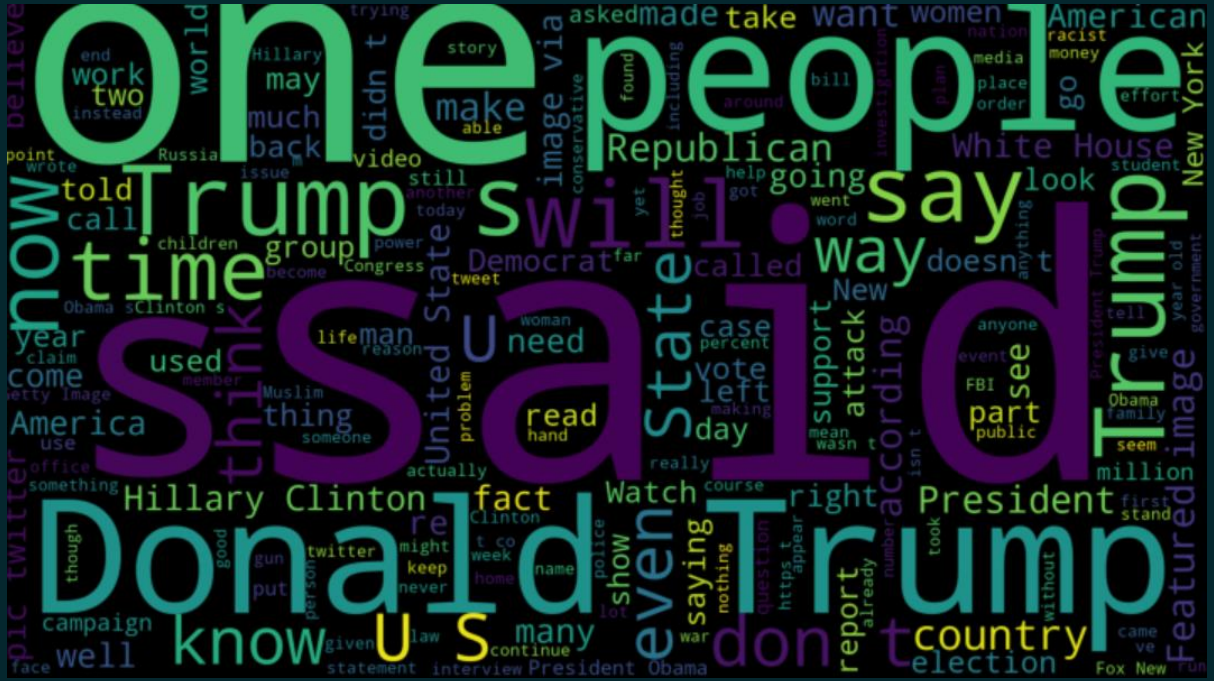
4- date

title	text	subject	date
As U.S. bu	WASHINGTON (Reuters) - The head of a conserva	politicsNe	December 31, 2017
U.S. milita	WASHINGTON (Reuters) - Transgender people wil	politicsNe	December 29, 2017
Senior U.S	WASHINGTON (Reuters) - The special counsel inve	politicsNe	December 31, 2017
FBI Russia	WASHINGTON (Reuters) - Trump campaign advise	politicsNe	December 30, 2017
Trump wa	SEATTLE/WASHINGTON (Reuters) - President Don	politicsNe	December 29, 2017
White Hou	WEST PALM BEACH, Fla./WASHINGTON (Reuters)	politicsNe	December 29, 2017
Trump say	WEST PALM BEACH, Fla (Reuters) - President Don	politicsNe	December 29, 2017
Factbox: T	The following statementsÂ were posted to the ve	politicsNe	December 29, 2017
Trump on	The following statementsÂ were posted to the ve	politicsNe	December 29, 2017
Alabama c	WASHINGTON (Reuters) - Alabama Secretary of St	politicsNe	December 28, 2017
Jones cert	(Reuters) - Alabama officials on Thursday certifie	politicsNe	December 28, 2017
New York	NEW YORK/WASHINGTON (Reuters) - The new U.S	politicsNe	December 28, 2017
Factbox: T	The following statementsÂ were posted to the ve	politicsNe	December 28, 2017
Trump on	The following statementsÂ were posted to the ve	politicsNe	December 28, 2017
Man says	(In Dec. 25 story, in second paragraph, corrects n	politicsNe	December 25, 2017
Virginia of	(Reuters) - A lottery drawing to settle a tied Virgin	politicsNe	December 27, 2017
U.S. lawm	WASHINGTON (Reuters) - A Georgian-American b	politicsNe	December 27, 2017
Trump on	The following statementsÂ were posted to the ve	politicsNe	December 26, 2017
U.S. appea	(Reuters) - A U.S. appeals court in Washington on	politicsNe	December 26, 2017
Treasury S	(Reuters) - A gift-wrapped package addressed to t	politicsNe	December 24, 2017
Federal ju	WASHINGTON (Reuters) - A federal judge in Seattl	politicsNe	December 24, 2017
Exclusive:	NEW YORK (Reuters) - The U.S. Justice Departmen	politicsNe	December 23, 2017
Trump tra	(Reuters) - A U.S. appeals court on Friday said Pre	politicsNe	December 23, 2017
Second co	WASHINGTON (Reuters) - A federal appeals court	politicsNe	December 23, 2017
Failed vote	LIMA (Reuters) - Peruâ€™s President Pedro Pablo	politicsNe	December 23, 2017
Trump sig	WASHINGTON (Reuters) - U.S. President Donald T	politicsNe	December 22, 2017
Companie	WASHINGTON (Reuters) - U.S. financial regulators	politicsNe	December 23, 2017
Trump on	The following statementsÂ were posted to the ve	politicsNe	December 22, 2017
Mexico to	MEXICO CITY (Reuters) - Mexicoâ€™s finance min	politicsNe	December 22, 2017
Senate lea	WASHINGTON (Reuters) - U.S. Senate Majority Le	politicsNe	December 22, 2017
Alabama t	(Reuters) - Democrat Doug Jonesâ€™ surprise vict	politicsNe	December 22, 2017
McConnel	WASHINGTON (Reuters) - A summer spat betwee	politicsNe	December 22, 2017

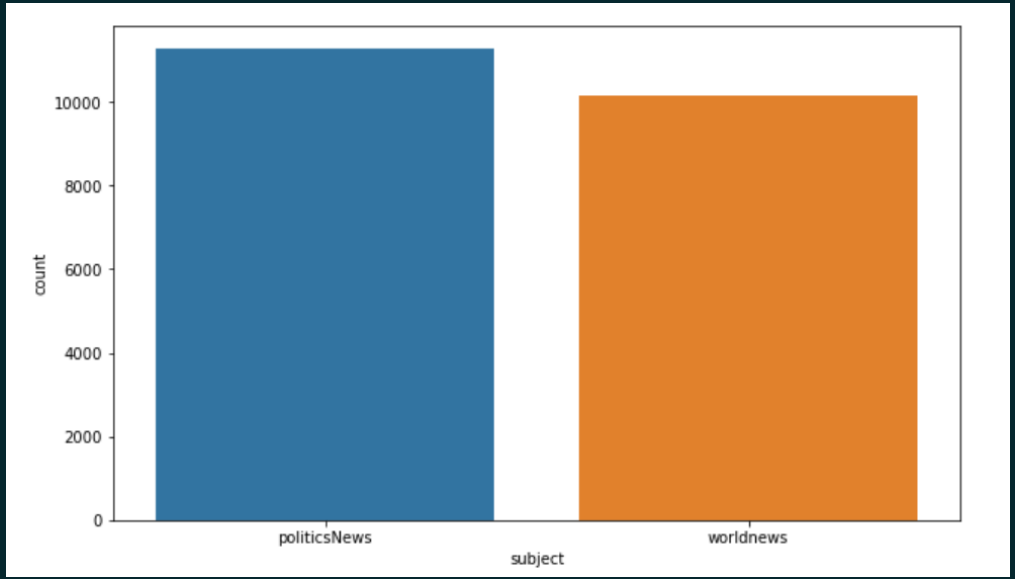
What is the most frequent subject in fake dataset?



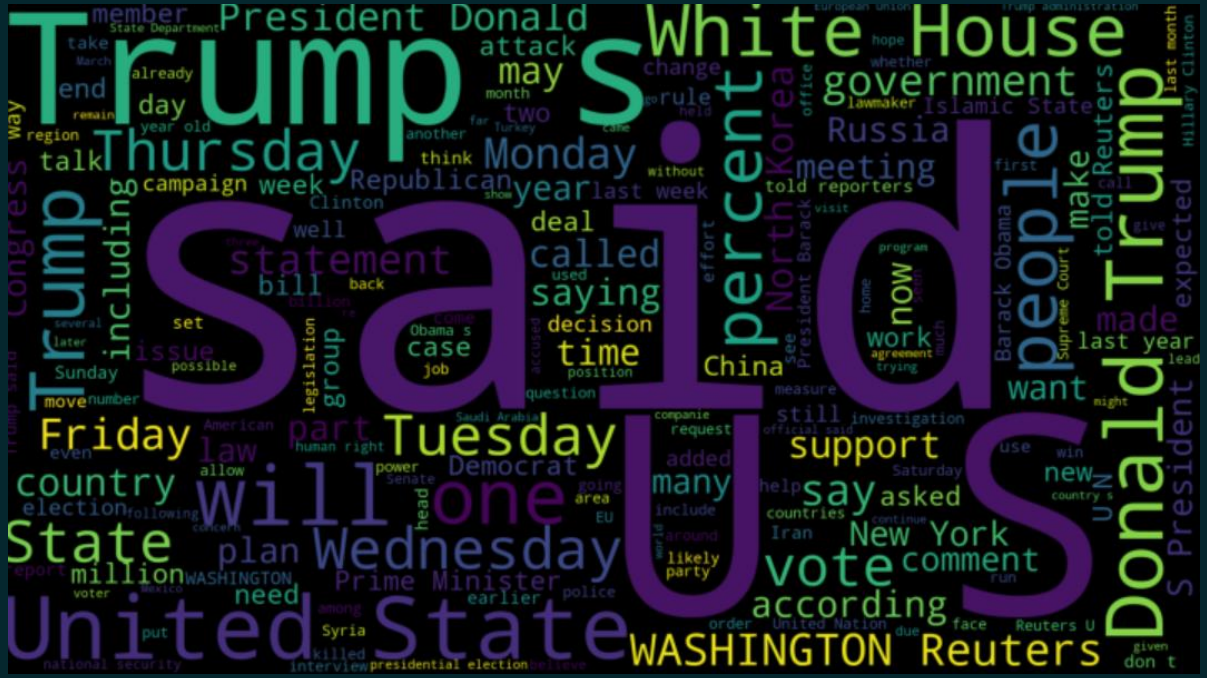
Word cloud visualizing the most frequent subject in fake dataset



What is the most frequent subject in real dataset?



Word cloud visualizing the most frequent subject in fake dataset



EDA

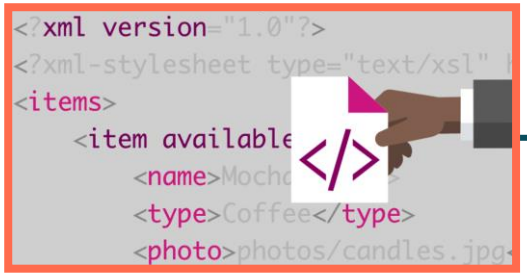
Unknow / Known Publishers

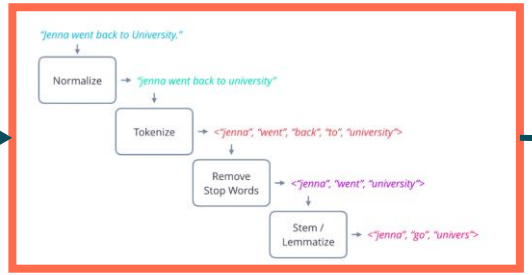
```
# creating a list of index that don't have the publiction part
unknow_publishers = []
for index, row in enumerate(real.text.values):
    try:
# if it doesn't have - it will create an error and if it's longer than 260 it will create an error, in both situations the
# error means that there's no publisher that's why we appended it it the exception to no_publishers list
        record = row.split("-", maxsplit = 1)
        record[1]
        # is it twitter news ?
        assert(len(record[0]) < 120)

    except:
        unknow_publishers.append(index)

len(unknow_publishers)
```

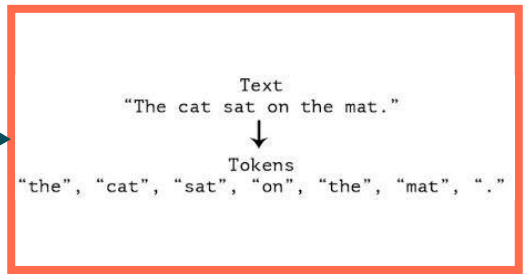
```
222
```

```
# now we create an array with all the publishers names
publishers = []
tmp_text = []

for index, row in enumerate (real.text.values):
    if index in unknow_publishers:
        tmp_text.append(row)
        publishers.append('unknown')
    else:
        record= row.split("-", maxsplit=1)
        publishers.append(record[0].strip())
        tmp_text.append(record[1].strip)
```

Data Pre-processing







Noise Removal

Deletes or transform things in text that degrade the NLP task model:

- Remove text file headers, footers
- Remove HTML, XML, etc. markup and metadata
- Extract valuable data from other formats: JSON, within the database

Normalization

Normalization refers to series of related tasks meant to put all text in same playing field it consist of:

1- Stemming: running → run
2- Lemmatization: better → good
3- Everything else

Tokenization

Tokenization is step which splits longer strings of text into smaller pieces.

Crayon

DATA PRE-PROCESSING

Normalization

1.

Converting to lower case characters

```
In [22]:
real['text'] = real['text'].apply(lambda x : str(x).lower())
fake['text'] = fake['text'].apply(lambda x : str(x).lower())
```

2.

```
real['class']=1
fake['class']=0
real
```

	title	text	subject	date	publisher	class
0	As U.S. budget fight looms, Republicans flip t...	as u.s. budget fight looms, republicans flip t...	politicsNews	December 31, 2017	WASHINGTON (Reuters)	1
1	U.S. military to accept transgender recruits o...	u.s. military to accept transgender recruits o...	politicsNews	December 29, 2017	WASHINGTON (Reuters)	1

3.

Remove special characters -- preprocess_kgptalkie library

```
data['text'] = data['text'].apply(lambda x: ps.remove_special_chars(x))
```

DATA PRE-PROCESSING

split() / tolist()

As U.S budget fight looms, Republicans flip their first fiscal script →

```
[['as',
  'us',
  'budget',
  'fight',
  'looms',
  'republicans',
  'flip',
  'their',
  'fiscal',
  'script',
  'builtin',
  'method',
  'strip',
  'of',
  'str',
  'object',
  'at',
  '0x0000021c038edfb0'],
 ['us',
```

DATA PRE-PROCESSING

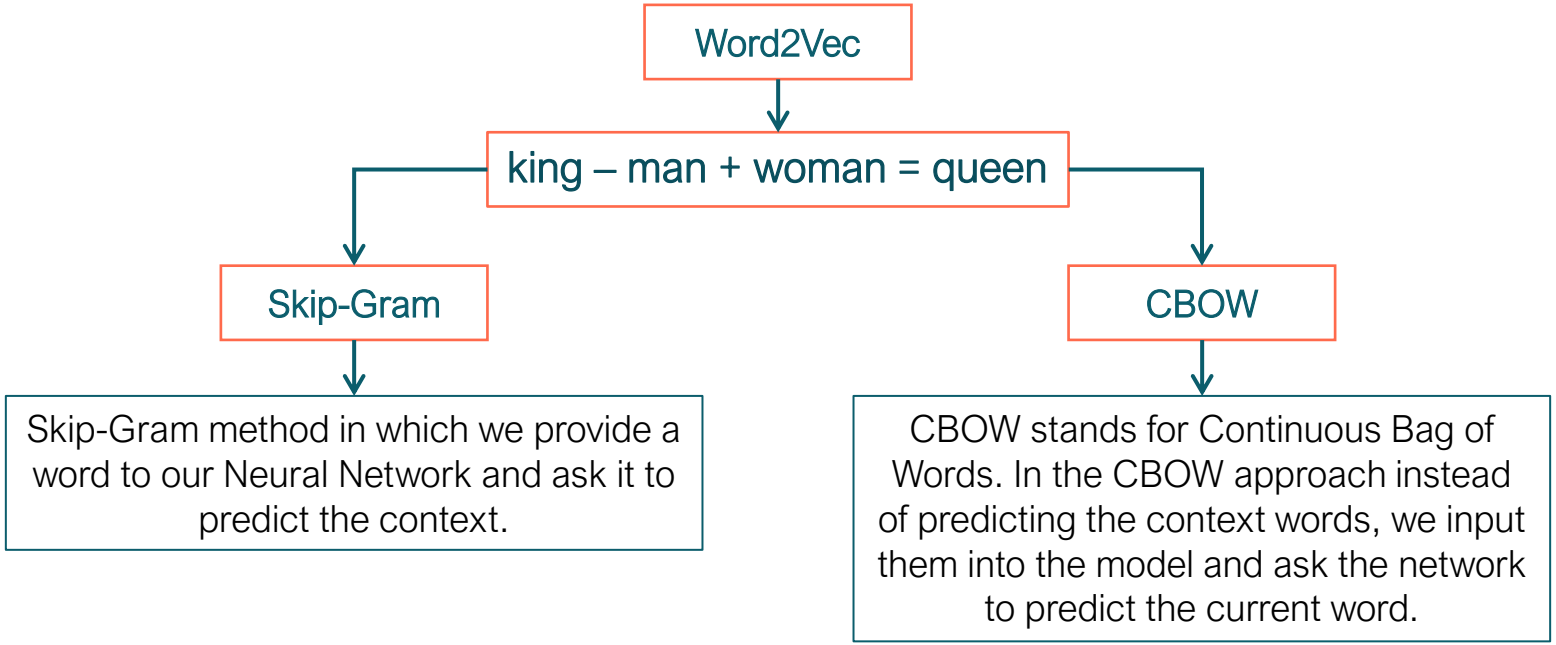
Vectorization

Machine learning and deep learning needs data in the numeric form. We basically used encoding technique (BagOfWord, TF-IDF, Word2Vec) to encode text into numeric vector.

Word2Vec:

We choose Word2Vec because in BagOfWord and TF-IDF every word was treated as individual entity, and semantics were completely ignored. With the introduction of Word2Vec, the vector representation of words was said to be contextually aware.

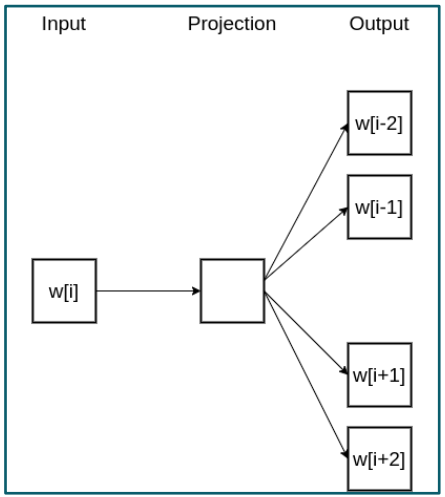
Vectorization – Word2Vec



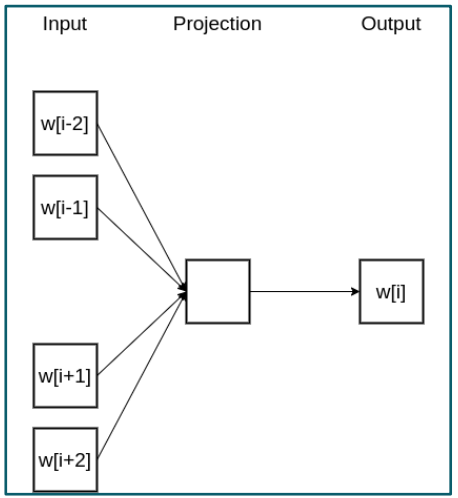
DATA PRE-PROCESSING

Vectorization – Word2Vec

Skip-Gram



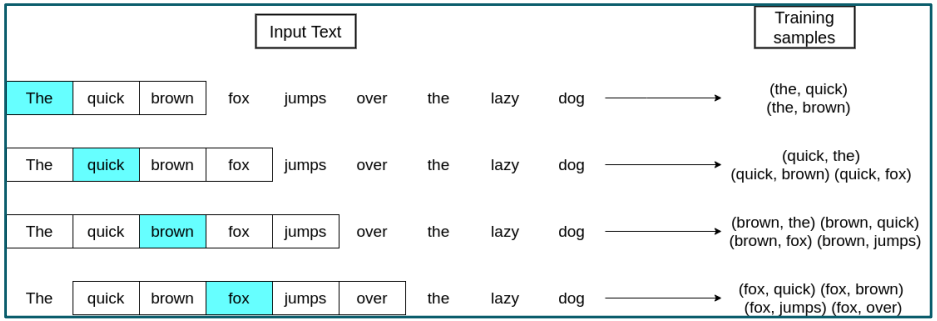
CBOW



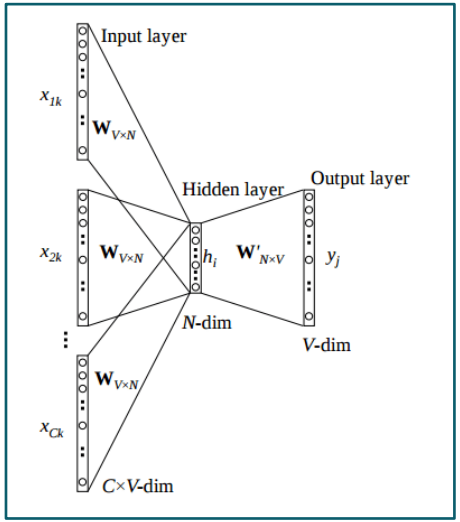
DATA PRE-PROCESSING

Vectorization – Word2Vec

Skip-Gram



CBOW



DATA PRE-PROCESSING

Vectorization

```
DIM =100
w2v_model = gensim.models.Word2Vec(X, vector_size  = DIM, window = 10, min_count=1)
```

```
# number of unique words
len(w2v_model.wv)
```

```
214703
```

```
w2v_model
```

```
w2v_model.wv.key_to_index
#Use KeyedVector's .key_to_index dict, .index_to_key list, and methods
#.get_vecattr(key, attr) and .set_vecattr(key, attr, new_val) instead.
```

```
{'the': 0,
 'to': 1,
 'of': 2,
 'and': 3,
 'a': 4,
 'in': 5,
 'that': 6,
 's': 7,
 'is': 8,
 'for': 9,
 'on': 10,
 'trump': 11,
 'it': 12
```

DATA PRE-PROCESSING

Vectorization

```
# getting the vector of a particular word
w2v_model.wv['china']
```

```
array([ 1.9093356 ,  2.2974195 ,  0.28128278,  1.549105  ,  0.77254045,
       -2.9315991 , -4.98908   ,  0.21570095,  3.195629  , -1.3674172 ,
        3.5499957 , -0.8564458 , -1.7446386 ,  1.1456767 , -0.03049176,
       -2.6928725 , -0.4091974 , -3.7493014 ,  2.0395038 ,  1.2551712 ,
       -2.5972311 , -0.00731431,  0.65777314, -3.1004255 , -2.128144  ,
        0.34549403,  0.32970116,  1.137637  ,  3.21343   , -2.8638592 ,
       -0.41603228,  0.70561355,  0.13587774,  0.14104858,  0.41028628,
       -1.5703479 ,  1.33233   ,  1.5106847 , -1.8140994 , -0.62216204,
       -0.01351932, -2.1443639 , -1.5058217 ,  2.416285  , -2.6939218 ,
        3.063219  ,  0.3016466 , -2.1021864 , -1.0553753 ,  1.567231  ,
       -1.0580026 , -0.1777452 ,  1.1575197 ,  2.9422908 , -0.58539903,
        1.4489905 , -0.8965168 ,  1.0115447 , -1.4507017 ,  0.80967075,
        0.6183848 ,  3.0762324 ,  0.14367294, -2.072617  , -5.020803  ,
       -0.07560263,  0.9498661 ,  0.58744496, -0.771867  ,  3.4488702 ,
        0.56288576, -0.67099655,  3.4534557 ,  1.781623  ,  1.8199677 ,
       -0.7096106 , -0.16626604, -0.36244583,  0.6749105 ,  1.1802186 ,
       -1.7399623 ,  3.7148132 , -1.3802011 ,  2.2915728 , -2.0361164 ,
        2.1238163 ,  1.5755166 , -2.0036867 ,  1.5545765 ,  4.9804325 ,
       -1.2157688 , -1.4319847 , -2.6330934 , -3.2006829 ,  2.003402  ,
        1.8683401 ,  1.1045316 , -2.0308616 , -2.153113  , -3.030711  ],
      dtype=float32)
```

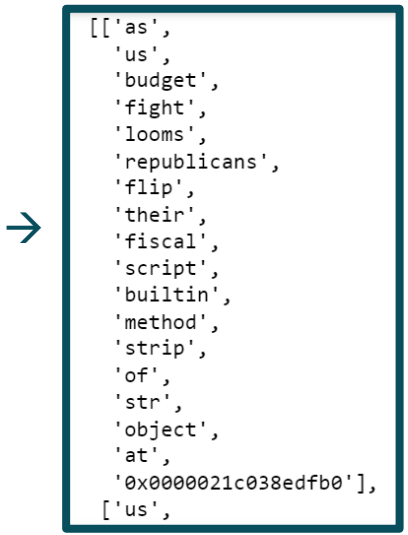
```
w2v_model.wv.most_similar('china')
```

```
[('japan', 0.7840046882629395),
 ('nafta', 0.7474444508552551),
 ('nato', 0.7117760181427002),
 ('pyongyang', 0.7061722278594971),
 ('taiwan', 0.6976385712623596),
 ('philippines', 0.6917264461517334),
 ('asia', 0.6822218894958496),
 ('iran', 0.6785287261009216),
 ('diplomatic', 0.6755640506744385),
 ('chinas', 0.6744887828826904)]
```

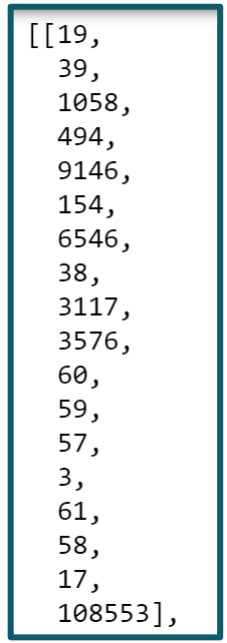
DATA PRE-PROCESSING

Tokenization

```
tokenizer = Tokenizer()
tokenizer.fit_on_texts(X)
X
```



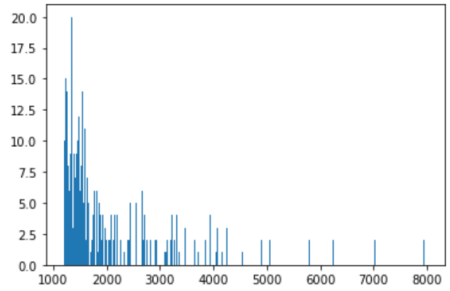

```
X = tokenizer.texts_to_sequences(X)
X
```



MODEL BUILDING

Length of x The length of words in the news

```
# we are calculating the length of x, the length of words in a news
plt.hist([len(x) for x in X], bins = 700, range=[1200 , 8000])
plt.show()
```




```
# number of news that has greater than 1000 words
nos = np.array([len(x) for x in X])
len(nos[nos > 1000])
```

```
1054
```

```
# When the sequence is more than 1000 then it is truncated
maxlen = 1000
X = pad_sequences(X, maxlen = maxlen)
X
```

```
array([[     0,      0,      0, ...,     58,     17, 108553],
       [     0,      0,      0, ...,     58,     17, 108554],
       [     0,      0,      0, ...,     58,     17, 108555],
       ...,
       [     2,   1265,     10, ...,    531,    357,   1594],
       [     0,      0,      0, ...,    531,   4463,    198],
       [     0,      0,      0, ...,    531,    491,   1594]])
```

```
# when you check the length of any index of x it is always 1000
len(X[0])
```

```
1000
```

The above code will pad every sequence. Following is the output after applying the pad_sqeuences function. We can see that every sequence has a length of 1000 numbers.

MODEL BUILDING

Weight Matrix

```
vocab_size = len(tokenizer.word_index) + 1
vocab = tokenizer.word_index
```

```
def get_weight_matrix(model):

    weight_matrix = np.zeros((vocab_size, DIM))

    for word, i in vocab.items():
        weight_matrix[i] = model.wv[word]

    return weight_matrix
```

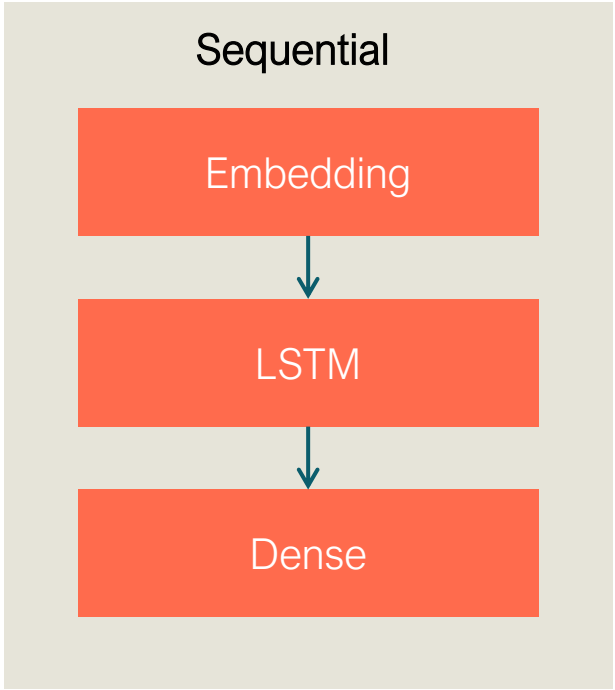
```
embedding_vectors = get_weight_matrix(w2v_model)

embedding_vectors.shape
```

```
(214704, 100)
```

MODEL BUILDING

Sequential

The Sequential model is a linear stack of layers.



MODEL BUILDING

Embedding

What is word embedding?

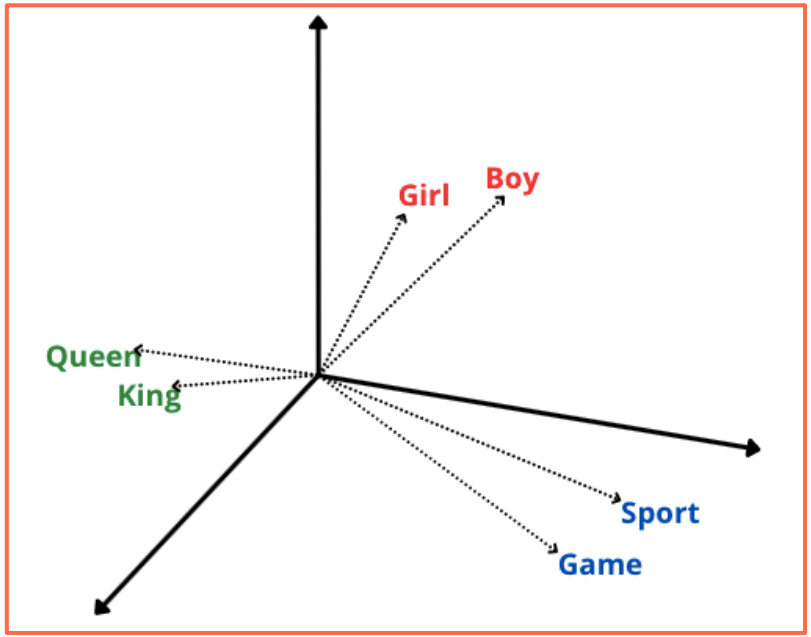
Word embedding is a technique to represent the word with the vector of numbers.

Word embedding means text as numbers.

Two vectors in the vector space are closer to each other

=

the two words are expected to have a similar meaning.



MODEL BUILDING

Embedding

What is the embedding layer in Keras?

Keras provides an embedding layer that converts each word into a fixed-length vector of defined size. in embedding layers, every word has a real-valued vector of fixed length.

The main advantage: of this technique. Due to the reduction in dimensions, the representation of a word takes minimum space compared to one-hot-encoding.

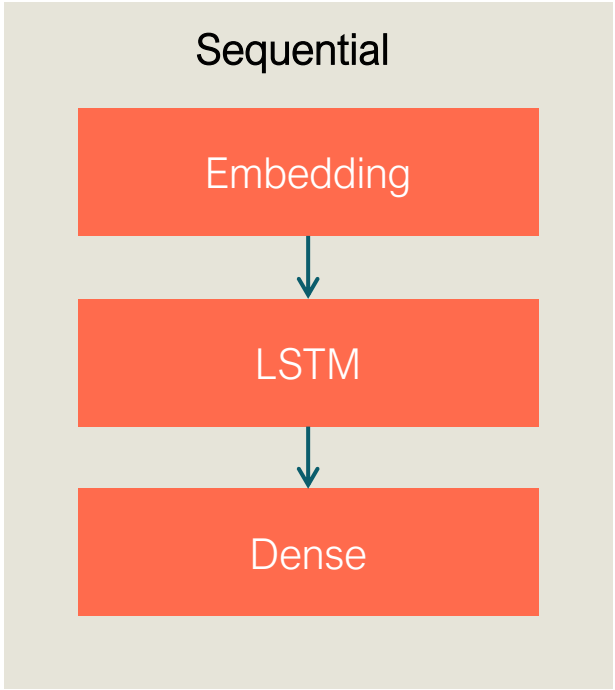
The embedding layer requires integer encoded input data to represent each word uniquely.

The output: of the embedding layer is a 2D vector where every vector represents a single word from the vocabulary.

MODEL BUILDING

Sequential

The Sequential model is a linear stack of layers.



MODEL BUILDING

LSTM

RNN

David, a 36-year-old man lives in San Francisco. He has a female friend Maria. Maria works as a cook in a famous restaurant in New York whom he met recently in a school alumni meet. Maria told him that she always had a passion for _____ .

we want our network to learn from dependency ‘cook’ to predict ‘cooking. There is a gap between the information what we want to predict and from where we want it to get predicted . This is called long-term dependency. We can say that anything larger than trigram as a long-term dependency. Unfortunately, RNN does not work practically in this situation.

MODEL BUILDING

LSTM

What is LSTM?

is a variation of a recurrent neural network (RNN) that is quite effective in predicting the long sequences of data like sentences and stock prices over a period.

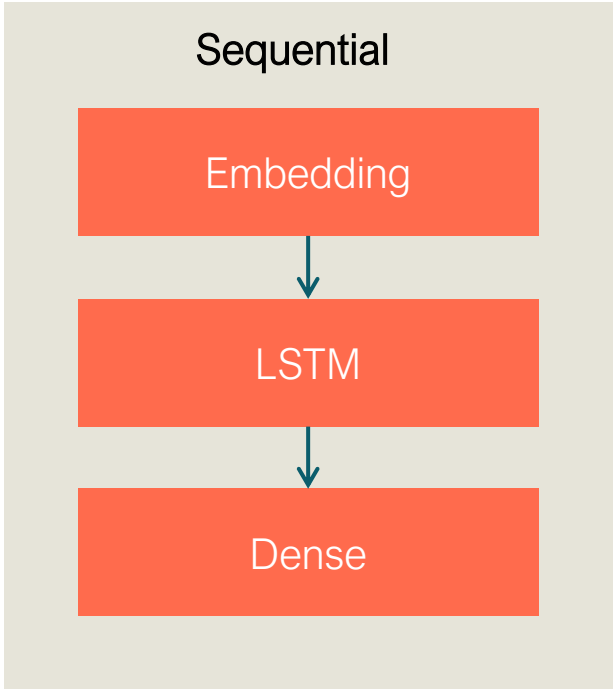
It differs from a normal feedforward network because there is a feedback loop in its architecture. It also includes a special unit known as a memory cell to withhold the past information for a longer time for making an effective prediction.

In fact, LSTM with its memory cells is an improved version of traditional RNNs which cannot predict using such a long sequence of data and run into the problem of vanishing gradient.

MODEL BUILDING

Sequential

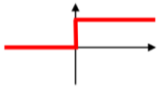
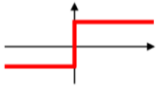
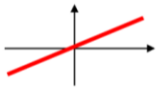
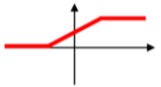
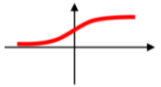
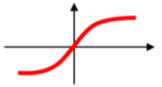
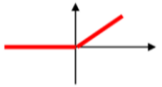
The Sequential model is a linear stack of layers.



Dense Layer

A dense layer or a fully connected layer is a layer where every neuron from the previous layer connects with every neuron of the next layer

The dense layer is one of the most used layers in designing deep neural networks. The figure below shows the various options of the Dense layer provided by the Keras library.

Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$	Multi-layer Neural Networks	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = max(0, z)$	Multi-layer Neural Networks	
Rectifier, softplus	$\phi(z) = \ln(1 + e^{z})$	Multi-layer Neural Networks	

MODEL BUILDING

Sequential

In [48]:
```
model = Sequential()
model.add(Embedding(vocab_size, output_dim=DIM, weights = [embedding_vectors], input_length = maxlen , trainable = False))
model.add(LSTM(units=128))
model.add(Dense(1, activation = 'sigmoid'))
model.compile(optimizer = 'adam', loss='binary_crossentropy', metrics=['acc'])
```

In [49]:
```
model.summary()
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 1000, 100)         21470400

 lstm (LSTM)                 (None, 128)               117248

 dense (Dense)               (None, 1)                 129

=================================================================
Total params: 21,587,777
Trainable params: 117,377
Non-trainable params: 21,470,400
_________________________________________________________________
```

MODEL BUILDING

Train and test split

Split to train and test

```
X_train, X_test, y_train, y_test = train_test_split(X,y)
```

```
model.fit(X_train, y_train, validation_split =0.3, epochs=6)
```

```
Epoch 1/6
737/737 [==============================] - 719s 973ms/step - loss: 0.0195 - acc: 0.9952 - val_loss: 0.0086 - val_acc: 0.9980
Epoch 2/6
737/737 [==============================] - 739s 1s/step - loss: 0.0091 - acc: 0.9970 - val_loss: 0.0076 - val_acc: 0.9978
Epoch 3/6
737/737 [==============================] - 2099s 3s/step - loss: 0.0050 - acc: 0.9983 - val_loss: 0.0085 - val_acc: 0.9984
Epoch 4/6
737/737 [==============================] - 3953s 5s/step - loss: 0.0022 - acc: 0.9992 - val_loss: 0.0081 - val_acc: 0.9982
Epoch 5/6
737/737 [==============================] - 989s 1s/step - loss: 0.0011 - acc: 0.9997 - val_loss: 0.0082 - val_acc: 0.9984
Epoch 6/6
737/737 [==============================] - 1000s 1s/step - loss: 3.8560e-04 - acc: 1.0000 - val_loss: 0.0100 - val_acc: 0.99
84
```

```
3]: <keras.callbacks.History at 0x17f8536ac40>
```

```
y_pred = (model.predict(X_test) >= 0.5).astype(int)
```

```
351/351 [==============================] - 136s 384ms/step
```

Model Evaluation

```
In [56]: accuracy_score(y_test, y_pred)

Out[56]: 0.9974164810690423

In [57]: print(classification_report(y_test, y_pred))
```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	5796
1	1.00	1.00	1.00	5429
accuracy			1.00	11225
macro avg	1.00	1.00	1.00	11225
weighted avg	1.00	1.00	1.00	11225

Testing With External Text Data

```
x = ['Drone attack on Russian bomber base leaves three dead']
x = tokenizer.texts_to_sequences(x)
x = pad_sequences(x, maxlen=maxlen)
```

```
# 0 == fake news, 1== real news
( model.predict(x) >=0.5 ).astype(int)
```

```
1/1 [==============================] - 0s 116ms/step

array([[0]])
```

Thank you

Crayon