

# Emojify!

Welcome to the second assignment of Week 2. You are going to use word vector representations to build an Emojifier.

Have you ever wanted to make your text messages more expressive? Your emojifier app will help you do that. So rather than writing:

"Congratulations on the promotion! Let's get coffee and talk. Love you!"

The emojifier can automatically turn this into:

"Congratulations on the promotion! 🙌 Let's get coffee and talk. 🍵 Love you! ❤️"

- You will implement a model which inputs a sentence (such as "Let's go see the baseball game tonight!") and finds the most appropriate emoji to be used with this sentence (🚫).

## Using word vectors to improve emoji lookups

- In many emoji interfaces, you need to remember that ❤️ is the "heart" symbol rather than the "love" symbol.
  - In other words, you'll have to remember to type "heart" to find the desired emoji, and typing "love" won't bring up that symbol.
- We can make a more flexible emoji interface by using word vectors!
- When using word vectors, you'll see that even if your training set explicitly relates only a few words to a particular emoji, your algorithm will be able to generalize and associate additional words in the test set to the same emoji.
  - This works even if those additional words don't even appear in the training set.
  - This allows you to build an accurate classifier mapping from sentences to emojis, even using a small training set.

## What you'll build

1. In this exercise, you'll start with a baseline model (Emojifier-V1) using word embeddings.
2. Then you will build a more sophisticated model (Emojifier-V2) that further incorporates an LSTM.

## Updates

If you were working on the notebook before this update...

- The current notebook is version "2a".
- You can find your original work saved in the notebook with the previous version name ("v2")
- To view the file directory, go to the menu "File->Open", and this will open a new tab that shows the file directory.

### List of updates

- sentence\_to\_avg
  - Updated instructions.
  - Use separate variables to store the total and the average (instead of just `avg`).
  - Additional hint about how to initialize the shape of `avg` vector.
- sentences\_to\_indices
  - Updated preceding text and instructions, added additional hints.
- pretrained\_embedding\_layer
  - Additional instructions to explain how to implement each step.
- Emoify\_V2
  - Modifies instructions to specify which parameters are needed for each Keras layer.
  - Remind users of Keras syntax.
  - Explanation of how to use the layer object that is returned by `pretrained_embedding_layer`.
  - Provides sample Keras code.
- Spelling, grammar and wording corrections.

Let's get started! Run the following cell to load the package you are going to use.

```
In [1]: import numpy as np
        from emo_utils import *
        import emoji
        import matplotlib.pyplot as plt
```

# 1 - Baseline model: Emojifier-V1






## 1.1 - Dataset EMOJISSET

Let's start by building a simple baseline classifier.

You have a tiny dataset (X, Y) where:

- X contains 127 sentences (strings).
- Y contains an integer label between 0 and 4 corresponding to an emoji for each sentence.

X (sentences)	Y (labels)
I love you	0
Congrats on the new job	2
I think I will end up alone	3
I want to have sushi for dinner!	4
It was funny lol	2
she did not answer my text	3
Happy new year	2
my algorithm performs poorly	3
he can pitch really well	1
you are failing this exercise	3
you did well on your exam.	2
What you did was awesome	2
I am frustrated	3

code	emoji	label
:heart:		0
:baseball:		1
:smile:		2
:disappointed:		3
:fork_and_knife:		4

**Figure 1:** EMOJISSET - a classification problem with 5 classes. A few examples of sentences are given here.

Let's load the dataset using the code below. We split the dataset between training (127 examples) and testing (56 examples).

```
In [2]: X_train, Y_train = read_csv('data/train_emoji.csv')
```

```
In [3]:
```

Run the following cell to print sentences from X\_train and corresponding labels from Y\_train.

- Change `idx` to see different examples.
- Note that due to the font used by iPython notebook, the heart emoji may be colored black rather than red.

```
In [4]: for idx in range(10):
```

```

never talk to me again 😞
I am proud of your achievements 😊
It is the worst day in my life 😞
Miss you so much ❤️
food is life 🍴
I love you mum ❤️
Stop saying bullshit 😞
congratulations on your acceptance 😊
The assignment is too long 😞
I want to go play 🎮
```

## 1.2 - Overview of the Emojifier-V1

In this part, you are going to implement a baseline model called "Emojifier-v1".

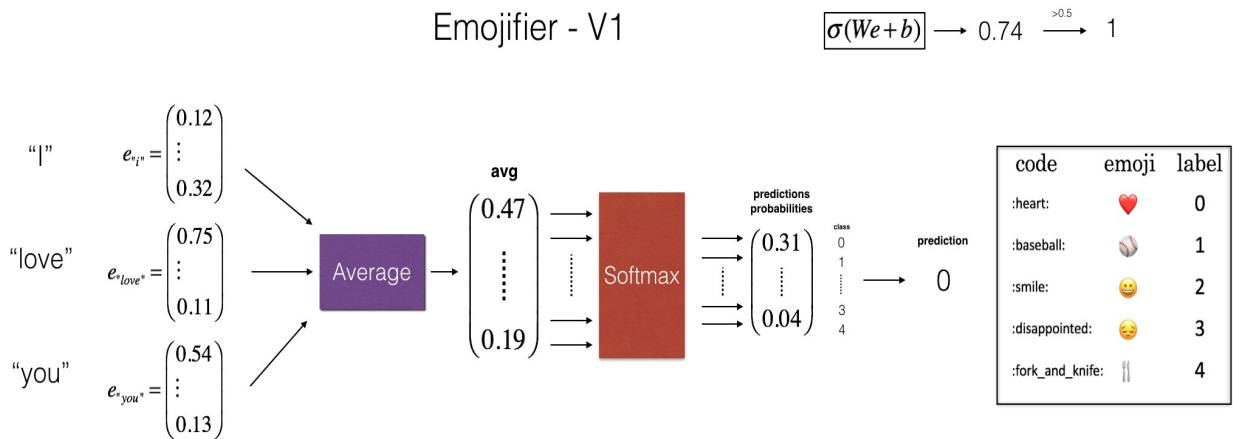


Figure 2: Baseline model (Emojifier-V1).

### Inputs and outputs

- The input of the model is a string corresponding to a sentence (e.g. "I love you").
- The output will be a probability vector of shape (1,5), (there are 5 emojis to choose from).
- The (1,5) probability vector is passed to an argmax layer, which extracts the index of the emoji with the highest probability.

### One-hot encoding

- To get our labels into a format suitable for training a softmax classifier, let's convert  $Y$  from its current shape  $(m, 1)$  into a "one-hot representation"  $(m, 5)$ ,
  - Each row is a one-hot vector giving the label of one example.
  - Here,  $Y_{oh}$  stands for "Y-one-hot" in the variable names  $Y_{oh\_train}$  and  $Y_{oh\_test}$ :

```
In [5]: Y_oh_train = convert_to_one_hot(Y_train, C = 5)
```

Let's see what `convert_to_one_hot()` did. Feel free to change `index` to print out different values.

```
In [6]: idx = 50
print(f"Sentence '{X_train[idx]}' has label index {Y_train[idx]}, which is emoji {label
Sentence 'I missed you' has label index 0, which is emoji ❤️
Label index 0 in one-hot encoding format is [ 1.  0.  0.  0.  0.]
```

All the data is now ready to be fed into the Emojify-V1 model. Let's implement the model!

### 1.3 - Implementing Emojifier-V1

As shown in Figure 2 (above), the first step is to:

- Convert each word in the input sentence into their word vector representations.
- Then take an average of the word vectors.
- Similar to the previous exercise, we will use pre-trained 50-dimensional GloVe embeddings.

Run the following cell to load the `word_to_vec_map`, which contains all the vector representations.

In [7]:

You've loaded:

- `word_to_index`: dictionary mapping from words to their indices in the vocabulary
  - (400,001 words, with the valid indices ranging from 0 to 400,000)
- `index_to_word`: dictionary mapping from indices to their corresponding words in the vocabulary
- `word_to_vec_map`: dictionary mapping words to their GloVe vector representation.

Run the following cell to check if it works.

In [8]:

```
word = "cucumber"
idx = 289846
print("the index of", word, "in the vocabulary is", word_to_index[word])
```

```
the index of cucumber in the vocabulary is 113317
the 289846th word in the vocabulary is potatos
```

**Exercise:** Implement `sentence_to_avg()`. You will need to carry out two steps:

1. Convert every sentence to lower-case, then split the sentence into a list of words.
  - `X.lower()` and `X.split()` might be useful.
2. For each word in the sentence, access its GloVe representation.
  - Then take the average of all of these word vectors.
  - You might use `numpy.zeros()`.

#### Additional Hints

- When creating the `avg` array of zeros, you'll want it to be a vector of the same shape as the other word vectors in the `word_to_vec_map`.
  - You can choose a word that exists in the `word_to_vec_map` and access its `.shape` field.
  - Be careful not to hard code the word that you access. In other words, don't assume that if you see the word 'the' in the `word_to_vec_map` within this notebook, that this word will be in the `word_to_vec_map` when the function is being called by the automatic grader.
  - Hint: you can use any one of the word vectors that you retrieved from the input `sentence` to find the shape of a word vector.

```
In [9]: # GRADED FUNCTION: sentence_to_avg

def sentence_to_avg(sentence, word_to_vec_map):
    """
    Converts a sentence (string) into a list of words (strings). Extracts the GloVe re
    and averages its value into a single vector encoding the meaning of the sentence.

    Arguments:
    sentence -- string, one training example from X
    word_to_vec_map -- dictionary mapping every word in a vocabulary into its 50-dimer

    Returns:
    avg -- average vector encoding information about the sentence, numpy-array of shap
    """

    ### START CODE HERE ###
    # Step 1: Split sentence into list of lower case words (~ 1 line)
    words = [i.lower() for i in sentence.split()]

    # Initialize the average word vector, should have the same shape as your word vect
    avg = np.zeros((50,))

    # Step 2: average the word vectors. You can loop over the words in the list "words
    total = 0
    for w in words:
        total += word_to_vec_map[w]
    avg = total/len(words)

    ### END CODE HERE ###
```

```
In [10]: avg = sentence_to_avg("Morrocan couscous is my favorite dish", word_to_vec_map)

avg =
[-0.008005    0.56370833 -0.50427333  0.258865    0.55131103  0.03104983
 -0.21013718  0.16893933 -0.09590267  0.141784   -0.15708967  0.18525867
  0.6495785   0.38371117  0.21102167  0.11301667  0.02613967  0.26037767
  0.05820667 -0.01578167 -0.12078833 -0.02471267  0.4128455   0.5152061
  0.38756167 -0.898661   -0.535145   0.33501167  0.68806933 -0.2156265
  1.797155    0.10476933 -0.36775333  0.750785    0.10282583  0.348925
 -0.27262833  0.66768    -0.10706167 -0.283635    0.59580117  0.28747333
 -0.3366635   0.23393817  0.34349183  0.178405    0.1166155   -0.076433
  0.1445417   0.09808667]
```

### Expected Output:

```
avg =
[-0.008005    0.56370833 -0.50427333  0.258865    0.55131103  0.03104983
 -0.21013718  0.16893933 -0.09590267  0.141784   -0.15708967  0.18525867
  0.6495785   0.38371117  0.21102167  0.11301667  0.02613967  0.26037767
  0.05820667 -0.01578167 -0.12078833 -0.02471267  0.4128455   0.5152061
  0.38756167 -0.898661   -0.535145   0.33501167  0.68806933 -0.2156265
  1.797155    0.10476933 -0.36775333  0.750785    0.10282583  0.348925
 -0.27262833  0.66768    -0.10706167 -0.283635    0.59580117  0.28747333
 -0.3366635   0.23393817  0.34349183  0.178405    0.1166155   -0.076433
  0.1445417   0.09808667]
```

## Model

You now have all the pieces to finish implementing the `model()` function. After using `sentence_to_avg()` you need to:

- Pass the average through forward propagation
- Compute the cost
- Backpropagate to update the softmax parameters

**Exercise:** Implement the `model()` function described in Figure (2).

- The equations you need to implement in the forward pass and to compute the cross-entropy cost are below:
- The variable  $Y_{oh}$  ("Y one hot") is the one-hot encoding of the output labels.

$$z^{(i)} = W \cdot avg^{(i)} + b$$

$$a^{(i)} = softmax(z^{(i)})$$

$$\mathcal{L}^{(i)} = - \sum_{k=0}^{n_y-1} Y_{oh,k}^{(i)} * \log(a_k^{(i)})$$

**Note** It is possible to come up with a more efficient vectorized implementation. For now, let's use nested for loops to better understand the algorithm, and for easier debugging.

We provided the function `softmax()`, which was imported earlier.

```

In [11]: # GRADED FUNCTION: model

def model(X, Y, word_to_vec_map, learning_rate = 0.01, num_iterations = 400):
    """
    Model to train word vector representations in numpy.

    Arguments:
    X -- input data, numpy array of sentences as strings, of shape (m, 1)
    Y -- labels, numpy array of integers between 0 and 7, numpy-array of shape (m, 1)
    word_to_vec_map -- dictionary mapping every word in a vocabulary into its 50-dimer
    learning_rate -- learning_rate for the stochastic gradient descent algorithm
    num_iterations -- number of iterations

    Returns:
    pred -- vector of predictions, numpy-array of shape (m, 1)
    W -- weight matrix of the softmax layer, of shape (n_y, n_h)
    b -- bias of the softmax layer, of shape (n_y,)
    """

    np.random.seed(1)

    # Define number of training examples
    m = Y.shape[0]                    # number of training examples
    n_y = 5                          # number of classes
    n_h = 50                         # dimensions of the GloVe vectors

    # Initialize parameters using Xavier initialization
    W = np.random.randn(n_y, n_h) / np.sqrt(n_h)
    b = np.zeros((n_y,))

    # Convert Y to Y_onehot with n_y classes
    Y_oh = convert_to_one_hot(Y, C = n_y)

    # Optimization loop
    for t in range(num_iterations): # Loop over the number of iterations
        for i in range(m):         # Loop over the training examples

            ### START CODE HERE ### (~ 4 lines of code)
            # Average the word vectors of the words from the i'th training example
            avg = sentence_to_avg(X[i], word_to_vec_map)

            # Forward propagate the avg through the softmax layer
            z = np.dot(W, avg) + b
            a = softmax(z)

            # Compute cost using the i'th training label's one hot representation and
            cost = -1*np.sum(np.multiply(Y_oh[i], np.log(a)))
            ### END CODE HERE ###

            # Compute gradients
            dz = a - Y_oh[i]
            dW = np.dot(dz.reshape(n_y,1), avg.reshape(1, n_h))
            db = dz

            # Update parameters with Stochastic Gradient Descent
            W = W - learning_rate * dW
            b = b - learning_rate * db

        if t % 100 == 0:
            print("Epoch: " + str(t) + " --- cost = " + str(cost))
            pred = predict(X, Y, W, b, word_to_vec_map) #predict is defined in emo_utili

```



Run the next cell to train your model and learn the softmax parameters ( $W, b$ ).

```
Epoch: 0 --- cost = 1.95204988128
Accuracy: 0.34848484848485
Epoch: 100 --- cost = 0.0797181872601
Accuracy: 0.93181818181818
Epoch: 200 --- cost = 0.0445636924368
Accuracy: 0.95454545454545
Epoch: 300 --- cost = 0.0343226737879
Accuracy: 0.969696969697
[[ 3.]
 [ 2.]
 [ 3.]
 [ 0.]
 [ 4.]
 [ 0.]
 [ 3.]
 [ 2.]
 [ 3.]
 [ 1.]
 [ 3.]
 [ 2.]
```

**Expected Output** (on a subset of iterations):

```
Epoch: 0    cost = 1.95204988128  Accuracy: 0.348484848485
Epoch: 100  cost = 0.0797181872601  Accuracy: 0.931818181818
Epoch: 200  cost = 0.0445636924368  Accuracy: 0.954545454545
Epoch: 300  cost = 0.0343226737879  Accuracy: 0.969696969697
```

Great! Your model has pretty high accuracy on the training set. Lets now see how it does on the test set.

## 1.4 - Examining test set performance

- Note that the `predict` function used here is defined in `emo_util.spy`.

```
In [14]: print("Training set:")
pred_train = predict(X_train, Y_train, W, b, word_to_vec_map)
print('Test set:')

Training set:
Accuracy: 0.977272727273
Test set:
Accuracy: 0.857142857143
```

**Expected Output:**


```
Train set accuracy  97.7
Test set accuracy   85.7
```

- Random guessing would have had 20% accuracy given that there are 5 classes. ( $1/5 = 20\%$ ).
- This is pretty good performance after training on only 127 examples.

### The model matches emojis to relevant words

In the training set, the algorithm saw the sentence

*"I love you"*

with the label .

- You can check that the word "adore" does not appear in the training set.
- Nonetheless, lets see what happens if you write *"I adore you."*

```
In [15]: X_my_sentences = np.array(["i adore you", "i love you", "funny lol", "lets play with a ball", "food is ready", "not feeling happy"])
Y_my_labels = np.array([[0], [0], [2], [1], [4], [3]])

pred = predict(X_my_sentences, Y_my_labels , W, b, word_to_vec_map)
print_predictions(X_my_sentences, pred)
```

Accuracy: 0.833333333333

i adore you ❤️  
i love you ❤️  
funny lol 😄  
lets play with a ball 🏀  
food is ready 🍴  
not feeling happy 😞

Amazing!

- Because *adore* has a similar embedding as *love*, the algorithm has generalized correctly even to a word it has never seen before.
- Words such as *heart*, *dear*, *beloved* or *adore* have embedding vectors similar to *love*.
  - Feel free to modify the inputs above and try out a variety of input sentences.
  - How well does it work?

### Word ordering isn't considered in this model

- Note that the model doesn't get the following sentence correct:

"not feeling happy"

- This algorithm ignores word ordering, so is not good at understanding phrases like "not happy."

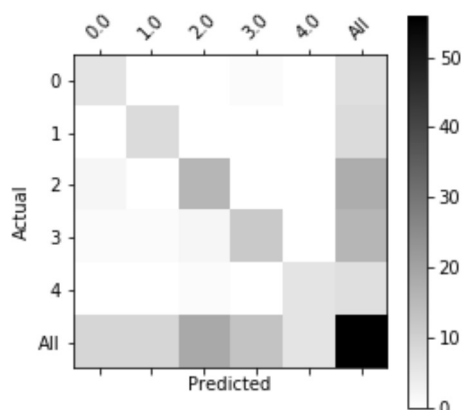
### Confusion matrix

- Printing the confusion matrix can also help understand which classes are more difficult for your model.
- A confusion matrix shows how often an example whose label is one class ("actual" class) is mislabeled by the algorithm with a different class ("predicted" class).

```
In [16]: print(Y_test.shape)
print(' ' + label_to_emoji(0) + ' ' + label_to_emoji(1) + ' ' + label_to_emoji(2) + ' ' + label_to_emoji(3) + ' ' + label_to_emoji(4) + ' ')
print(pd.crosstab(Y_test, pred_test.reshape(56,), rownames=['Actual'], colnames=['Predicted'], margins=True))
plot_confusion_matrix(Y_test, pred_test)
```

(56,)

	0.0	1.0	2.0	3.0	4.0	All
Actual						
0	6	0	0	1	0	7
1	0	8	0	0	0	8
2	2	0	16	0	0	18
3	1	1	2	12	0	16
4	0	0	1	0	6	7
All	9	9	19	13	6	56



## What you should remember from this section

- Even with a 127 training examples, you can get a reasonably good model for Emojifying.
  - This is due to the generalization power word vectors gives you.
- Emojify-V1 will perform poorly on sentences such as *"This movie is not good and not enjoyable"*
  - It doesn't understand combinations of words.
  - It just averages all the words' embedding vectors together, without considering the ordering of words.

You will build a better algorithm in the next section!

## 2 - Emojifier-V2: Using LSTMs in Keras:

Let's build an LSTM model that takes word **sequences** as input!

- This model will be able to account for the word ordering.
- Emojifier-V2 will continue to use pre-trained word embeddings to represent words.
- We will feed word embeddings into an LSTM.
- The LSTM will learn to predict the most appropriate emoji.

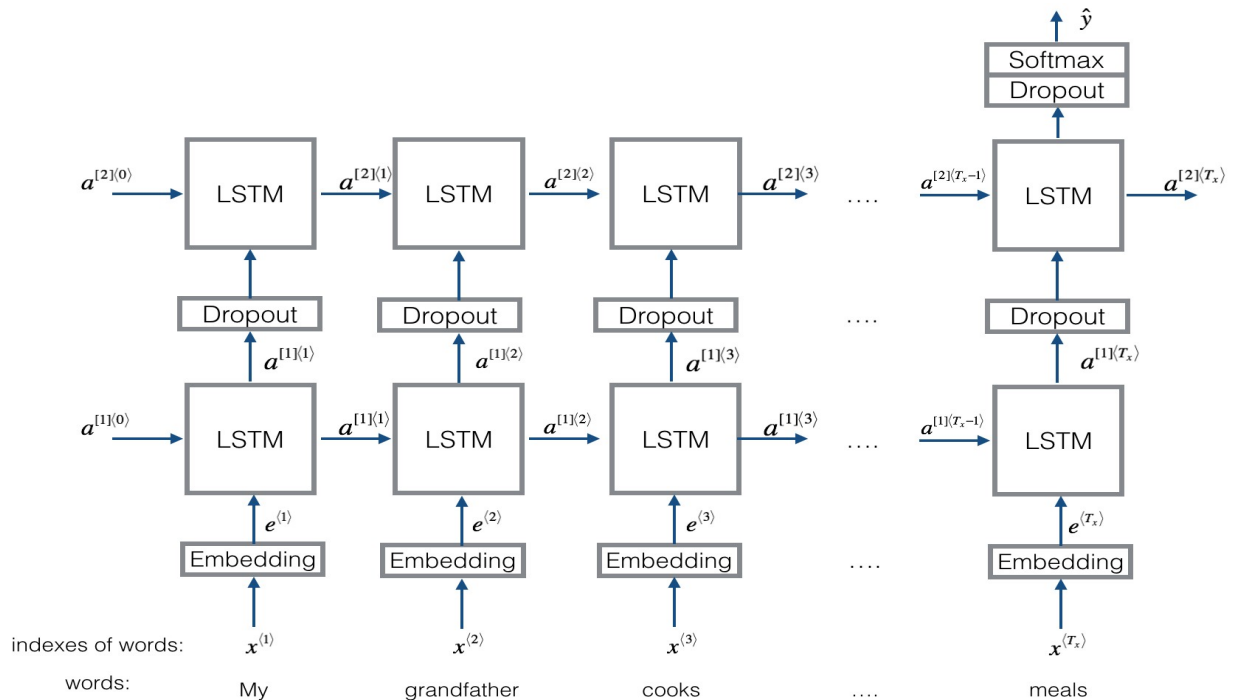
Run the following cell to load the Keras packages.

```
In [17]: import numpy as np
np.random.seed(0)
from keras.models import Model
from keras.layers import Dense, Input, Dropout, LSTM, Activation
from keras.layers.embeddings import Embedding
from keras.preprocessing import sequence
from keras.initializers import glorot_uniform
```

Using TensorFlow backend.

## 2.1 - Overview of the model

Here is the Emojifier-v2 you will implement:



**Figure 3:** Emojifier-V2. A 2-layer LSTM sequence classifier.

## 2.2 Keras and mini-batching

- In this exercise, we want to train Keras using mini-batches.
- However, most deep learning frameworks require that all sequences in the same mini-batch have the **same length**.
  - This is what allows vectorization to work: If you had a 3-word sentence and a 4-word sentence, then the computations needed for them are different (one takes 3 steps of an LSTM, one takes 4 steps) so it's just not possible to do them both at the same time.

### Padding handles sequences of varying length

- The common solution to handling sequences of **different length** is to use padding. Specifically:
  - Set a maximum sequence length
  - Pad all sequences to have the same length.

### Example of padding

- Given a maximum sequence length of 20, we could pad every sentence with "0"s so that each input sentence is of length 20.
- Thus, the sentence "I love you" would be represented as  $(e_I, e_{love}, e_{you}, \vec{0}, \vec{0}, \dots, \vec{0})$ .
- In this example, any sentences longer than 20 words would have to be truncated.
- One way to choose the maximum sequence length is to just pick the length of the longest sentence in the training set.

## 2.3 - The Embedding layer

- In Keras, the embedding matrix is represented as a "layer".
- The embedding matrix maps word indices to embedding vectors.
  - The word indices are positive integers.
  - The embedding vectors are dense vectors of fixed size.
  - When we say a vector is "dense", in this context, it means that most of the values are non-zero. As a counter-example, a one-hot encoded vector is not "dense."
- The embedding matrix can be derived in two ways:
  - Training a model to derive the embeddings from scratch.
  - Using a pretrained embedding

### Using and updating pre-trained embeddings

- In this part, you will learn how to create an `Embedding()` (<https://keras.io/layers/embeddings/>) layer in Keras
- You will initialize the Embedding layer with the GloVe 50-dimensional vectors.
- In the code below, we'll show you how Keras allows you to either train or leave fixed this layer.
- Because our training set is quite small, we will leave the GloVe embeddings fixed instead of updating them.

## Inputs and outputs to the embedding layer

- The `Embedding()` layer's input is an integer matrix of size **(batch size, max input length)**.
  - This input corresponds to sentences converted into lists of indices (integers).
  - The largest integer (the highest word index) in the input should be no larger than the vocabulary size.
- The embedding layer outputs an array of shape (batch size, max input length, dimension of word vectors).
- The figure shows the propagation of two example sentences through the embedding layer.
  - Both examples have been zero-padded to a length of `max_len=5`.
  - The word embeddings are 50 units in length.
  - The final dimension of the representation is `(2, max_len, 50)`.

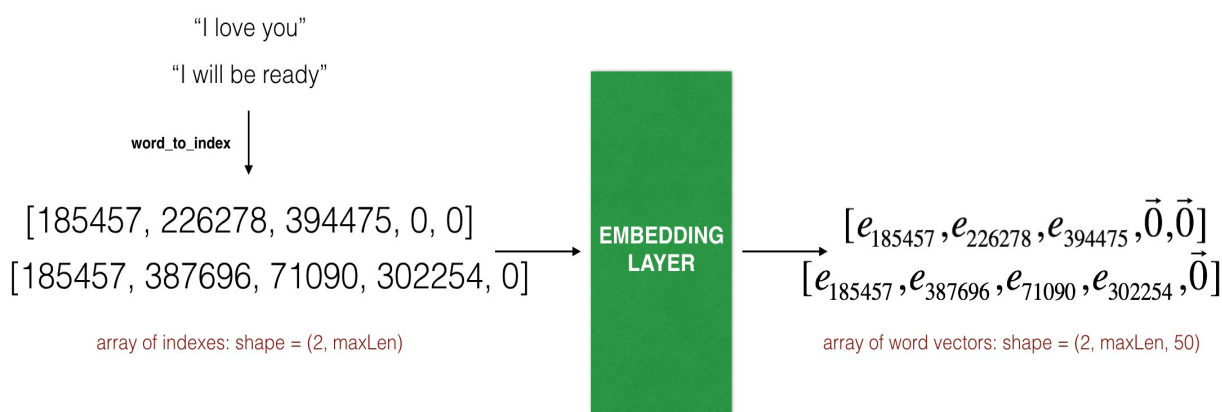


Figure 4: Embedding layer

## Prepare the input sentences

### Exercise:

- Implement `sentences_to_indices`, which processes an array of sentences (X) and returns inputs to the embedding layer:
  - Convert each training sentences into a list of indices (the indices correspond to each word in the sentence)
  - Zero-pad all these lists so that their length is the length of the longest sentence.

### Additional Hints

- Note that you may have considered using the `enumerate()` function in the for loop, but for the purposes of passing the autograder, please follow the starter code by initializing and incrementing `j` explicitly.

```
In [18]: for idx, val in enumerate(["I", "like", "learning"]):
          0 I
          1 like
          2 learning
```

```
In [23]: # GRADED FUNCTION: sentences_to_indices

def sentences_to_indices(X, word_to_index, max_len):
    """
    Converts an array of sentences (strings) into an array of indices corresponding to words.
    The output shape should be such that it can be given to `Embedding()` (described in
    Embedding layer).

    Arguments:
    X -- array of sentences (strings), of shape (m, 1)
    word_to_index -- a dictionary containing the each word mapped to its index
    max_len -- maximum number of words in a sentence. You can assume every sentence in
    X has a length less than or equal to max_len.

    Returns:
    X_indices -- array of indices corresponding to words in the sentences from X, of shape (m, max_len)
    """

    m = X.shape[0] # number of training examples

    ### START CODE HERE ###
    # Initialize X_indices as a numpy matrix of zeros and the correct shape (~ 1 line)
    X_indices = np.zeros((m,max_len))

    for i in range(m): # loop over training examples
        # Convert the ith training sentence in lower case and split it into words. You
        # can assume every sentence in X has a length less than or equal to max_len.
        sentence_words = [i.lower() for i in X[i].split()]

        # Initialize j to 0
        j = 0

        # Loop over the words of sentence_words
        for w in sentence_words:
            # Set the (i,j)th entry of X_indices to the index of the correct word.
            X_indices[i, j] = word_to_index[w]
            # Increment j to j + 1
            j = j+1

    ### END CODE HERE ###
```

Run the following cell to check what `sentences_to_indices()` does, and check your results.

```
In [24]: X1 = np.array(["funny lol", "lets play baseball", "food is ready for you"])
X1_indices = sentences_to_indices(X1,word_to_index, max_len = 5)
print("X1 =", X1)

X1 = ['funny lol' 'lets play baseball' 'food is ready for you']
X1_indices =
[[ 155345.  225122.     0.     0.     0.]
 [ 220930.  286375.  69714.     0.     0.]
 [ 151204.  192973.  302254.  151349.  394475.]]
```

**Expected Output:**

```
X1 = ['funny lol' 'lets play baseball' 'food is ready for you']
X1_indices =
[[ 155345.  225122.     0.     0.     0.]
 [ 220930.  286375.  69714.     0.     0.]
 [ 151204.  192973.  302254.  151349.  394475.]]
```



## Build embedding layer

- Let's build the `Embedding()` layer in Keras, using pre-trained word vectors.
- The embedding layer takes as input a list of word indices.
  - `sentences_to_indices()` creates these word indices.
- The embedding layer will return the word embeddings for a sentence.

**Exercise:** Implement `pretrained_embedding_layer()` with these steps:

1. Initialize the embedding matrix as a numpy array of zeros.
  - The embedding matrix has a row for each unique word in the vocabulary.
    - There is one additional row to handle "unknown" words.
    - So `vocab_len` is the number of unique words plus one.
  - Each row will store the vector representation of one word.
    - For example, one row may be 50 positions long if using GloVe word vectors.
  - In the code below, `emb_dim` represents the length of a word embedding.
2. Fill in each row of the embedding matrix with the vector representation of a word
  - Each word in `word_to_index` is a string.
  - `word_to_vec_map` is a dictionary where the keys are strings and the values are the word vectors.
3. Define the Keras embedding layer.
  - Use `Embedding()` (<https://keras.io/layers/embeddings/>).
  - The input dimension is equal to the vocabulary length (number of unique words plus one).
  - The output dimension is equal to the number of positions in a word embedding.
  - Make this layer's embeddings fixed.
    - If you were to set `trainable = True`, then it will allow the optimization algorithm to modify the values of the word embeddings.
    - In this case, we don't want the model to modify the word embeddings.
4. Set the embedding weights to be equal to the embedding matrix.
  - Note that this is part of the code is already completed for you and does not need to be modified.

```
In [25]: # GRADED FUNCTION: pretrained_embedding_layer

def pretrained_embedding_layer(word_to_vec_map, word_to_index):
    """
    Creates a Keras Embedding() layer and loads in pre-trained GloVe 50-dimensional ve

    Arguments:
    word_to_vec_map -- dictionary mapping words to their GloVe vector representation.
    word_to_index -- dictionary mapping from words to their indices in the vocabulary

    Returns:
    embedding_layer -- pretrained layer Keras instance
    """

    vocab_len = len(word_to_index) + 1 # adding 1 to fit Keras embedd
    emb_dim = word_to_vec_map["cucumber"].shape[0] # define dimensionality of you

    ### START CODE HERE ###
    # Step 1
    # Initialize the embedding matrix as a numpy array of zeros.
    # See instructions above to choose the correct shape.
    emb_matrix = np.zeros((vocab_len, emb_dim))

    # Step 2
    # Set each row "idx" of the embedding matrix to be
    # the word vector representation of the idx'th word of the vocabulary
    for word, idx in word_to_index.items():
        emb_matrix[idx, :] = word_to_vec_map[word]

    # Step 3
    # Define Keras embedding layer with the correct input and output sizes
    # Make it non-trainable.
    embedding_layer = Embedding(vocab_len, emb_dim, trainable=False)
    ### END CODE HERE ###

    # Step 4 (already done for you; please do not modify)
    # Build the embedding layer, it is required before setting the weights of the embe
    embedding_layer.build((None,)) # Do not modify the "None". This line of code is c

    # Set the weights of the embedding layer to the embedding matrix. Your layer is no
    embedding_layer.set_weights([emb_matrix])
```

```
In [26]: embedding_layer = pretrained_embedding_layer(word_to_vec_map, word_to_index)
weights[0][1][3] = -0.3403
```

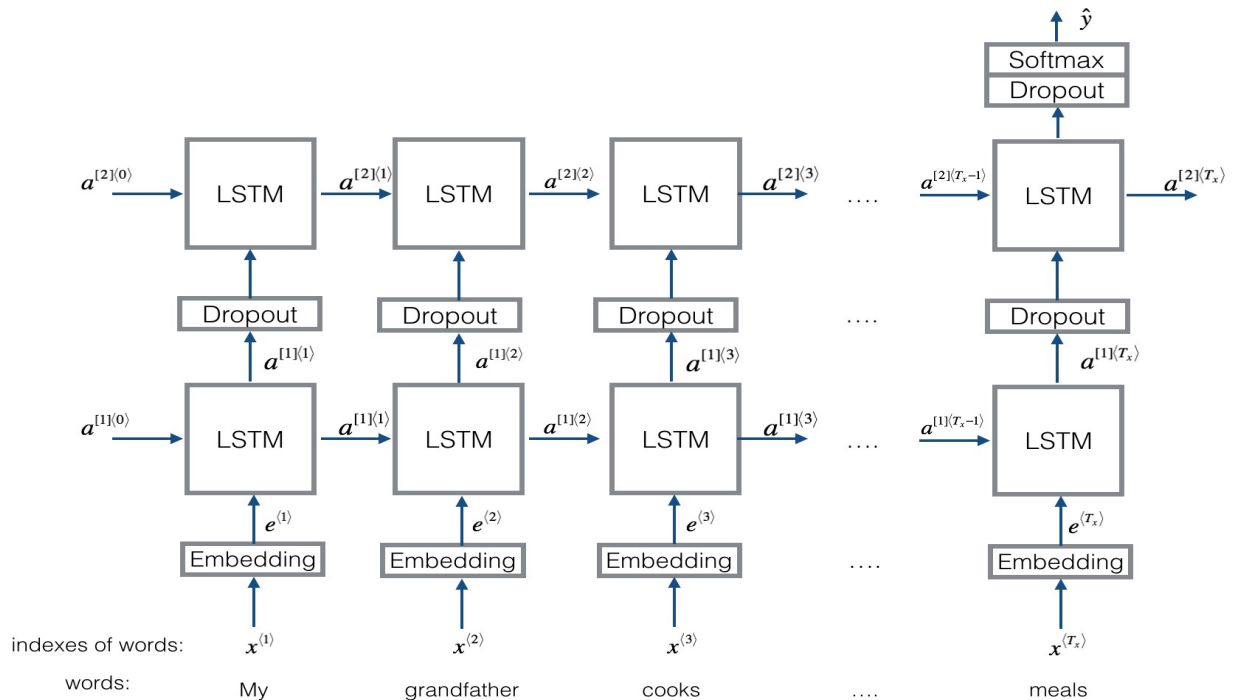
### Expected Output:

```
weights[0][1][3] = -0.3403
```

## 2.3 Building the Emojifier-V2

Lets now build the Emojifier-V2 model.

- You feed the embedding layer's output to an LSTM network.



**Figure 3:** Emojifier-v2. A 2-layer LSTM sequence classifier.

**Exercise:** Implement `Emojify_V2()`, which builds a Keras graph of the architecture shown in Figure 3.

- The model takes as input an array of sentences of shape  $(m, \text{max\_len}, )$  defined by `input_shape`.
- The model outputs a softmax probability vector of shape  $(m, C = 5)$ .
- You may need to use the following Keras layers:
  - `Input()` (<https://keras.io/layers/core/#input>)
    - Set the `shape` and `dtype` parameters.
    - The inputs are integers, so you can specify the data type as a string, 'int32'.
  - `LSTM()` (<https://keras.io/layers/recurrent/#lstm>)
    - Set the `units` and `return_sequences` parameters.
  - `Dropout()` (<https://keras.io/layers/core/#dropout>)
    - Set the `rate` parameter.
  - `Dense()` (<https://keras.io/layers/core/#dense>)
    - Set the `units`,
    - Note that `Dense()` has an `activation` parameter. For the purposes of passing the autograder, please do not set the activation within `Dense()`. Use the separate `Activation` layer to do so.
  - `Activation()` (<https://keras.io/activations/>).
    - You can pass in the activation of your choice as a lowercase string.
  - `Model()` (<https://keras.io/models/model/>) Set inputs and outputs.

**Additional Hints**

```
In [29]: # GRADED FUNCTION: Emojify_V2

def Emojify_V2(input_shape, word_to_vec_map, word_to_index):
    """
    Function creating the Emojify-v2 model's graph.

    Arguments:
    input_shape -- shape of the input, usually (max_len,)
    word_to_vec_map -- dictionary mapping every word in a vocabulary into its 50-dimer
    word_to_index -- dictionary mapping from words to their indices in the vocabulary

    Returns:
    model -- a model instance in Keras
    """

    ### START CODE HERE ###
    # Define sentence_indices as the input of the graph.
    # It should be of shape input_shape and dtype 'int32' (as it contains indices, whi
    sentence_indices = Input(input_shape, dtype='int32')

    # Create the embedding layer pretrained with GloVe Vectors (~1 line)
    embedding_layer = pretrained_embedding_layer(word_to_vec_map, word_to_index)

    # Propagate sentence_indices through your embedding layer
    # (See additional hints in the instructions).
    embeddings = embedding_layer(sentence_indices)

    # Propagate the embeddings through an LSTM layer with 128-dimensional hidden state
    # The returned output should be a batch of sequences.
    X = LSTM(128, return_sequences=True)(embeddings)
    # Add dropout with a probability of 0.5
    X = Dropout(0.5)(X)
    # Propagate X trough another LSTM layer with 128-dimensional hidden state
    # The returned output should be a single hidden state, not a batch of sequences.
    X = LSTM(128, return_sequences=False)(X)
    # Add dropout with a probability of 0.5
    X = Dropout(0.5)(X)
    # Propagate X through a Dense layer with 5 units
    X = Dense(5)(X)
    # Add a softmax activation
    X = Activation('softmax')(X)

    # Create Model instance which converts sentence_indices into X.
    model = Model(inputs=sentence_indices, outputs=X)

    ### END CODE HERE ###
```

Run the following cell to create your model and check its summary. Because all sentences in the dataset are less than 10 words, we chose `max_len = 10`. You should see your architecture, it uses "20,223,927" parameters, of which 20,000,050 (the word embeddings) are non-trainable, and the remaining 223,877 are. Because our vocabulary size has 400,001 words (with valid indices from 0 to 400,000) there are  $400,001 \times 50 = 20,000,050$  non-trainable parameters.

```
In [30]: model = Emojify_V2((maxLen,), word_to_vec_map, word_to_index)
```

Layer (type)	Output Shape	Param #
=====	=====	=====
input_2 (InputLayer)	(None, 10)	0
embedding_3 (Embedding)	(None, 10, 50)	20000050
lstm_3 (LSTM)	(None, 10, 128)	91648
dropout_3 (Dropout)	(None, 10, 128)	0
lstm_4 (LSTM)	(None, 128)	131584
dropout_4 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 5)	645
activation_2 (Activation)	(None, 5)	0
=====	=====	=====
Total params: 20,223,927		
Trainable params: 223,877		
Non-trainable params: 20,000,050		

As usual, after creating your model in Keras, you need to compile it and define what loss, optimizer and metrics you are want to use. Compile your model using `categorical_crossentropy` loss, `adam` optimizer and `['accuracy']` metrics:

```
In [31]:
```

It's time to train your model. Your `Emojifier-V2` model takes as input an array of shape `(m, max_len)` and outputs probability vectors of shape `(m, number of classes)`. We thus have to convert `X_train` (array of sentences as strings) to `X_train_indices` (array of sentences as list of word indices), and `Y_train` (labels as indices) to `Y_train_oh` (labels as one-hot vectors).

```
In [32]: X_train_indices = sentences_to_indices(X_train, word_to_index, maxLen)
```

Fit the Keras model on `X_train_indices` and `Y_train_oh`. We will use `epochs = 50` and `batch_size = 32`.

In [33]:

```
Epoch 1/50
132/132 [=====] - 0s - loss: 1.5989 - acc: 0.2652
Epoch 2/50
132/132 [=====] - 0s - loss: 1.5249 - acc: 0.3636
Epoch 3/50
132/132 [=====] - 0s - loss: 1.4816 - acc: 0.3485 - E
TA: 0s - loss: 1.5158 - acc: 0.2
Epoch 4/50
132/132 [=====] - 0s - loss: 1.4408 - acc: 0.4318
Epoch 5/50
132/132 [=====] - 0s - loss: 1.3312 - acc: 0.5000
Epoch 6/50
132/132 [=====] - 0s - loss: 1.2075 - acc: 0.5455
Epoch 7/50
132/132 [=====] - 0s - loss: 1.1065 - acc: 0.5833
Epoch 8/50
132/132 [=====] - 0s - loss: 1.0038 - acc: 0.5758
Epoch 9/50
132/132 [=====] - 0s - loss: 0.9677 - acc: 0.6364 - E
TA: 0s - loss: 0.9520 - acc: 0.640
```

Your model should perform around **90% to 100% accuracy** on the training set. The exact accuracy you get may be a little different. Run the following cell to evaluate your model on the test set.

In [34]:

```
X_test_indices = sentences_to_indices(X_test, word_to_index, max_len = maxLen)
Y_test_oh = convert_to_one_hot(Y_test, C = 5)
loss, acc = model.evaluate(X_test_indices, Y_test_oh)
print()

32/56 [=====>.....] - ETA: 0s
Test accuracy = 0.803571437086
```

You should get a test accuracy between 80% and 95%. Run the cell below to see the mislabelled examples.

In [35]:

```
# This code allows you to see the mislabelled examples
C = 5
y_test_oh = np.eye(C)[Y_test.reshape(-1)]
X_test_indices = sentences_to_indices(X_test, word_to_index, maxLen)
pred = model.predict(X_test_indices)
for i in range(len(X_test)):
    x = X_test_indices
    num = np.argmax(pred[i])
    if(num != Y_test[i]):
        Expected emoji: 😞 prediction: work is hard 😊
        Expected emoji: 😞 prediction: This girl is messing with me ❤️
        Expected emoji: 😞 prediction: work is horrible 😞
        Expected emoji: 😞 prediction: you brighten my day ❤️
        Expected emoji: 😞 prediction: she is a bully ❤️
        Expected emoji: 🙄 prediction: enjoy your game 😊
        Expected emoji: 😞 prediction: My life is so boring ❤️
        Expected emoji: 😞 prediction: will you be my valentine ❤️
        Expected emoji: 🙄 prediction: what is your favorite baseball game 😊
        Expected emoji: ❤️ prediction: I love you to the stars and back 😊
        Expected emoji: 😞 prediction: go away 🙄
```

Now you can try it on your own example. Write your own sentence below.

```
In [36]: # Change the sentence below to see your prediction. Make sure all the words are in the
x_test = np.array(['not feeling happy'])
X_test_indices = sentences_to_indices(x_test, word_to_index, maxlen)

not feeling happy 😊
```

## LSTM version accounts for word order

- Previously, Emojify-V1 model did not correctly label "not feeling happy," but our implementation of Emojify-V2 got it right.
  - (Keras' outputs are slightly random each time, so you may not have obtained the same result.)
- The current model still isn't very robust at understanding negation (such as "not happy")
  - This is because the training set is small and doesn't have a lot of examples of negation.
  - But if the training set were larger, the LSTM model would be much better than the Emojify-V1 model at understanding such complex sentences.

## Congratulations!

You have completed this notebook! ❤️❤️❤️

## What you should remember

- If you have an NLP task where the training set is small, using word embeddings can help your algorithm significantly.
- Word embeddings allow your model to work on words in the test set that may not even appear in the training set.
- Training sequence models in Keras (and in most other deep learning frameworks) requires a few important details:
  - To use mini-batches, the sequences need to be **padded** so that all the examples in a mini-batch have the **same length**.
  - An `Embedding()` layer can be initialized with pretrained values.
    - These values can be either fixed or trained further on your dataset.
    - If however your labeled dataset is small, it's usually not worth trying to train a large pre-trained set of embeddings.
  - `LSTM()` has a flag called `return_sequences` to decide if you would like to return every hidden states or only the last one.
  - You can use `Dropout()` right after `LSTM()` to regularize your network.

### Input sentences:

```
"Congratulations on finishing this assignment and building an Emojifier."
"We hope you're happy with what you've accomplished in this notebook!"
```

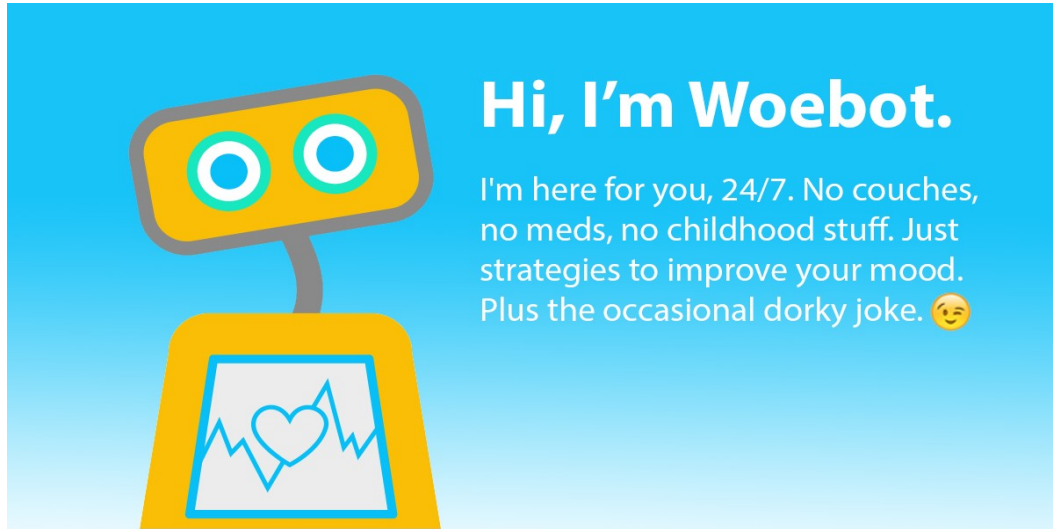
### Output emojis:



## Acknowledgments

Thanks to Alison Darcy and the Woebot team for their advice on the creation of this assignment.

- Woebot is a chatbot friend that is ready to speak with you 24/7.
- Part of Woebot's technology uses word embeddings to understand the emotions of what you say.
- You can chat with Woebot by going to <http://woebot.io> (<http://woebot.io>)



In [ ]: