

Topic	DEFINE MODEL FOR SENTIMENT ANAL	YSIS
Class Description	Analyzing the sentiment related to text by defining a machine learning model.	
Class	PRO C115	
Class time	45 mins	
Goal	 Introduction to the CNN model for text analysis Introduction to Embedding, Conv1D and LSTM 	
Resources Required	 Teacher Resources: Laptop with internet connectivity Earphones with mic Notebook and pen Smartphone Student Resources: Laptop with internet connectivity Earphones with mic Notebook and pen 	
Class structure	Warm-Up Teacher-Led Activity 1 Student-Led Activity 1 Wrap-Up	10 mins 10 mins 20 mins 05 mins
Credit & Permissions:	Tensorflow by Google Brain team Keras by Google Engineer Francois Chollet Numpy by The NumPy community Kaggle: Emotion Detection from Text	
WARM-UP SESSION - 10 mins		
Teacher Starts Slideshow		

© 2021 - WhiteHat Education Technology Private Limited.

Note: This document is the original copyright of WhiteHat Education Technology Private Limited.

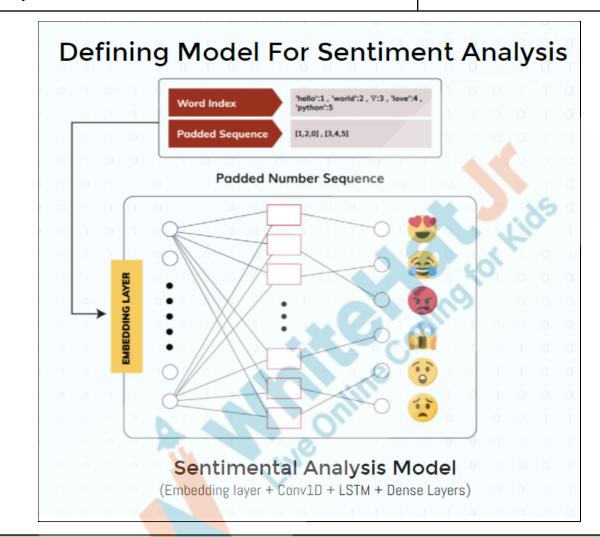


Slide 1 to 5 Refer to speaker notes and follow the instructions on each slide.		
Teacher Action	Student Action	
Hey <student's name="">. How are you? It's great to see you! Are you excited to learn something new today?</student's>	ESR: Hi, thanks! Yes, I am excited about it!	
Can you tell me what we learned in the previous class?	Click on the slide show tab and present the slides	
Note : Encourage the student to give answers and be more involved in the discussion.	* 1.95	
Amazing!	O toll	
Till now we have preprocessed the text data to visualize it using Pandas.	ding	
We also learned how we can convert text into a number sequence.		
WARM-UP QUIZ Click on In-Class Quiz		
Continue WARM-UP Session Slide 6 to 12		
Now that we have our text in the form of numbers, we'll now define a CNN model for sentiment prediction.		
As text and images are different, we need different layers for text feature extraction and classification.		
Today we are going to learn about new layers of the CNN model. i.e Embedding layer , Conv1D and LSTM layer .		
	ESR: Yes	

Note: This document is the original copyright of WhiteHat Education Technology Private Limited.



Are you excited?



Teacher Ends Slideshow



TEACHER-LED ACTIVITY - 10 mins

Teacher Initiates Screen Share

ACTIVITY

- Introduction to embedding layer, Conv1D layer and LSTM layer.
- Define the CNN model for text analysis.

© 2021 - WhiteHat Education Technology Private Limited.

Note: This document is the original copyright of WhiteHat Education Technology Private Limited. Please don't share, download or copy this file without permission.



Teacher Action Student Action Note: Open Teacher activity 2 for the boilerplate code and run all the cells. We have already converted our text dataset into number sequences. These padded sequences will be used to train the CNN model. The next step would be to convert these lists of padded sequences and labels into **Numpy** arrays to give it to the **Embedding** layer as shown below: import numpy as np training_padded = np.array(training_padded) training_labels = np.array(training_labels) From tensorflow.keras import the Sequential class to define a **Sequential CNN** model. Also, to extract features and for classification, different layers are imported such as Embedding, LSTM, Dense, Conv1D, Dropout, Maxpooling1D. Conv1D Max Pooling Word What were the two important steps to pre-process images? **ESR:** Extract features and classification Great! To extract the features from text data we will be using the following layers:

© 2021 - WhiteHat Education Technology Private Limited.

Note: This document is the original copyright of WhiteHat Education Technology Private Limited.



1. Embedding layer:

The embedding layer is the very first layer used for creating ML models for NLP. It requires that the input data be tokenized. It converts each word into a fixed size array depending upon different features.

2. Conv1D:

The **Conv1D** layer is used to extract features from the text data. A kernel is used to run over word embeddings to give a **Feature array**.

3. Max Pooling layer:

The **Max Pooling** layer is used to extract the most prominent feature from the feature array.

The **Embedding** layer uses the **Word Embedding** approach for representing words using floating-point numbers depending upon different features; it gives an array for feature learning.

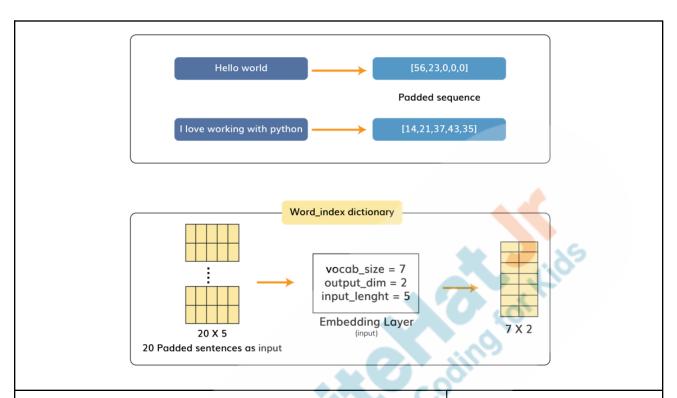
Do you remember how the **word_index** dictionary is created?

Great!

So now using these values we had created padded sequences. In these sequences, the words are identified by unique numbers.

ESR: Yes. Each word was assigned with a numerical value





The **Embedding layer** is defined as the first hidden layer of a neural network. It must specify three arguments:

Teacher activity 3: Embedding layer

- vocab_size: It is the number of unique words in training data. For example, we have 7 unique words in 2 sentences, then the size of the vocabulary would be 7 unique words.
- output_dim: is the size of the output array in which words will be embedded.
- **input_length:** It is the length of input sequence. It is the same as max_len used in padding the sequences. Here the input_length would be 5.

If we have 20 sentences with 7 unique words where the padding is done with max_length of 5, this gives us a reduced array of 7 x 2.



Dropout may be implemented on any or all hidden layers in the network as well as the visible or input layer. It is used to avoid errors while training the model.

Conv1D layer:

In the Conv1D layer, the convolution kernel only moves along a single axis, thus it is applied to linear data like text. A convolution kernel slides over an embedding array to give **a feature array**.

In case of **text**, a convolutional **kernel** is applied at embeddings for multiple words. The kernel will no longer be square, instead, it will be a rectangle. Here we are considering four features so the kernel used has 3X4 size.

Teacher activity 4: Convolution 1D layer

Convolution over Word Sequences:

To process an entire sequence of words, these kernels will slide down a list of word embeddings. This is called a **1D** convolution because the kernel is moving in only one dimension.

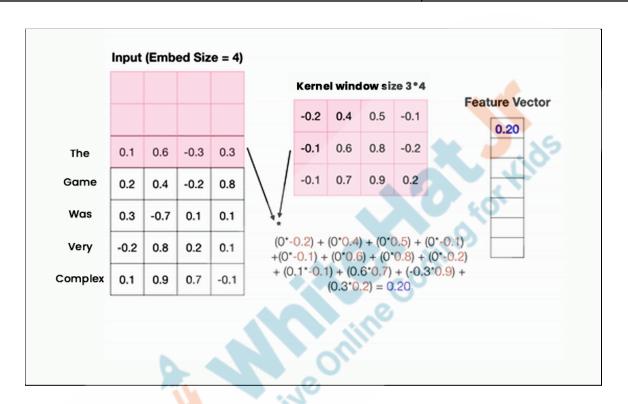
A single kernel will move one by one down a list of input embeddings. The resultant output will be a **feature array** that contains about as many values as there were input embeddings.

When the kernel is run over the array received from the embedding layer, we get a single output value as shown below.



Note: Refer to additional activities of C112 for matrix multiplication.

The image below is for Teacher reference. It shows how a kernel works in case of Conv1D layer to give feature array.



Do you remember the parameters we needed for the Conv2D layer?

Convolution 1D layer too must have these arguments:

- 1. **filters**: Defines the number of filters required to perform the convolution operation.
- 2. **kernel_size**: Defines the size of the kernel window.
- 3. **activation:** Defines the activation function for the Conv1D layer.

ESR: filters, kernel size and activation function.



Max Pooling layer:

A convolutional operation produces a **feature array** that can represent local features in sequences of word embeddings.

Identify the significant words from the following sentence:

'It was a tough fight, I managed to win'.

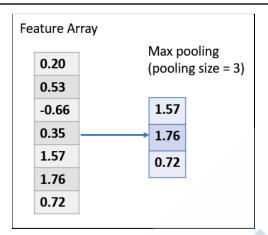
Yes!

These words may appear anywhere in the sentence, it signifies fight and victory. These are considered high-level features that are extracted by the **Max Pooling** layer.

Max Pooling layer retains only maximum value in the feature array which is the most useful, local feature. In the above feature array, three maximum values are considered as shown for pool size 3.

ESR: tough, win





Pooling layers **subsample** their input. Thus, the output after the max-pooling layer would be a **feature map** containing the **most prominent features** of the previous feature map.

Max pooling layer takes parameter **pool_size** to select the prominent features.

Teacher Activity 5: Max Pooling layer

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Embedding, LSTM, Dense
from tensorflow.keras.layers import Conv1D, Dropout, MaxPooling1D

model = tf.keras.Sequential([
    Embedding(vocab_size, embedding_dim, input_length=max_length),
    Dropout(0.2),

Conv1D(filters = 256, kernel_size = 3, activation = "relu"),
    MaxPooling1D(pool_size = 3),

Conv1D(filters = 128, kernel_size = 3, activation = "relu"),
    MaxPooling1D(pool_size = 3),
```

Long Short Term Memory (LSTM):

Long Short Term Memory Network is another type of Neural Network that allows relevant information to persist.



Let me check how much you can remember things. I am going to tell you a story in a few sentences. I will be asking you questions related to the story.

There were two friends who were walking across a desert. They kept on walking until they found an oasis, where they decided to take a bath. One of them got stuck in the mire and started drowning, but his friend saved him. After he recovered from the near drowning, he wrote on a stone "Today my best friend saved my life".

Tell me some of the words I used in the story?

Great!

So you were able to recollect the words which were related or relevant to the story.

Similarly, LSTM remembers the previous information and uses it for processing the current input.

The LSTM consists of three parts, as shown in the image below and each part performs an individual function.

The **first part** or **input gate** chooses whether the information coming from the previous input is to be remembered. If it is irrelevant, it can be forgotten.

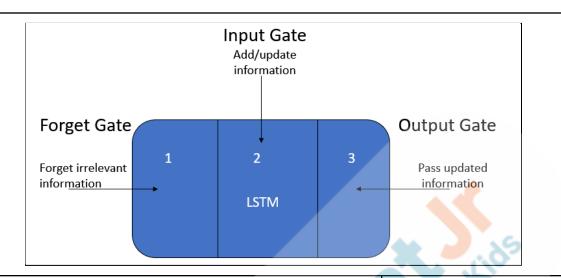
The **second part** or **output gate** tries to learn new information from the input to this cell.

At last, in the **third part** or **forget** gate passes the updated information from the current timestamp to the next input.

A common LSTM unit also called a **cell** contains an **input gate**, an **output gate** and a **forget gate**.

ESR: Varied





The **cell** remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell.

The input propagates forwards and backward through the LSTM layer and then concatenates with the outputs.

This makes LSTM an efficient network as irrelevant data is discarded and only keeps information relevant for training the model. We then fit it into a Dense neural network to do classification.

Teacher activity 6: Reference: LSTM layer

Define LSTM with 128 output units that return 128 features depending on the three timestamps by three gates of cells.

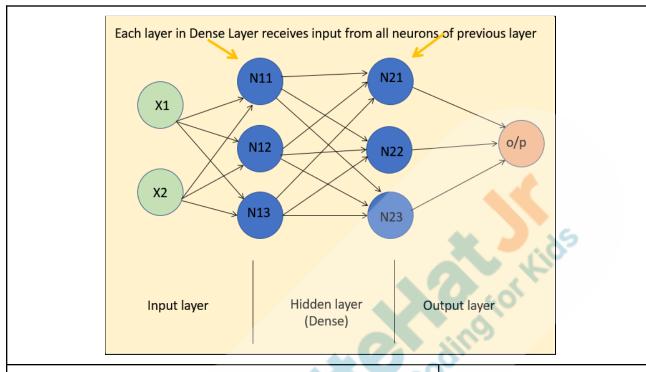


Dense layer:

Now, in the dense layer, each neuron is connected with the neuron in its preceding layer.

This layer is used at the final stage of CNN after LSTM to perform classification.





So, our first Dense layer would have **128** neurons and a **relu** activation function.

We have a Dropout layer to avoid errors.

Next, we are defining two more dense layers with **relu** and **softmax** activations.



Now that the model is created we need to compile, fit and save the model.

That was interesting!

Are you excited to define and test the model?

ESR: Yes

Teacher Stops Screen Share

So now it's your turn.

Please share your screen with me.



Teacher Starts Slideshow Slide 13 to 14

Refer to speaker notes and follow the instructions on each slide.

We have one more class challenge for you.

Remember in the previous class we trained the model in

© 2021 - WhiteHat Education Technology Private Limited.

Note: This document is the original copyright of WhiteHat Education Technology Private Limited.



teachable machines and then we had downloaded and tested the model locally?

After training your model on **Google Colab** we will test it on our system.

Let's try. I will guide you through it.

Teacher Ends Slideshow

STUDENT-LED ACTIVITY - 20 mins

- Ask the student to press the ESC key to come back to the panel.
- Guide the student to start Screen Share.
- The teacher gets into Full Screen.

Student Initiates Screen Share

ACTIVITY

- Create a CNN model for text sentiment analysis.
- Compile, fit and save the model.
- Predict the sentiment from any random text.
- Download the .h5 file and run it locally in your system to check the performance of your model.

Teacher Action	Student Action
Guide the student to open the boilerplate code <u>Student</u> <u>Activity 1</u> and run all the code.	
Note : The student will perform the same activities as done by the teacher.	
Note : Please refer to Teacher Activity 2 to follow through and guide the student to do the activities.	
Guide the student to perform the activities as mentioned below:	



- 1. Define the model by importing layers from tensorflow.keras.
- 2. Define Embedding layer, Conv1d, LSTM and Dense layers.
- 3. After defining layers, compile the model as performed in class 113.

Do you remember the parameters taken by the **compile()** function?

ESR: loss, optimizer and metrics

4. Similarly, compile the model using **loss**, **optimizer** and **metrics** parameters first.

model.compile(loss='sparse_categorical_crossentropy',optimizer='adam',metrics=['accuracy'])

Model: "sequential"		0
Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 100, 16)	16000 0
dropout (Dropout)	(None, 100, 16)	0
conv1d (Conv1D)	(None, 98, 256)	12544
<pre>max_pooling1d (MaxPooling1D)</pre>	(None, 32, 256)	0
conv1d_1 (Conv1D)	(None, 30, 128)	98432
<pre>max_pooling1d_1 (MaxPooling 1D)</pre>	(None, 10, 128)	0
lstm (LSTM)	(None, 128)	131584
dense (Dens <mark>e</mark>)	(None, 128)	16512
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 64)	8256
dense_2 (Dense)	(None, 6)	390
Total params: 427,718 Trainable params: 427,718 Non-trainable params: 0		



Thus the model is being compiled we got the total parameters and the trainable parameters as output from these layers.

Output data from **CNN** is also a 3D array consisting of **batch_size**, **height**, **width**.

So here each layer returns an array. As the output is not received in batches, **batch_size='None'**.

Also, each layer gives **parameters** that are used to train the model for prediction.

5. After compiling we have to train the model. Can you suggest what will be the next step?

ESR: Using the **fit()** function



Great!!!

6. Define the number of epochs to 30 and then train the model using the **fit()** function.

```
num_epochs = 30
history = model.fit(training_padded, training_labels, epochs=num_epochs, verbose=2)
```

```
Epoch 1/30
500/500 - 26s - loss: 1.5039 - accuracy: 0.3449 - 26s/epoch - 52ms/step
Epoch 2/30
500/500 - 21s - loss: 0.8811 - accuracy: 0.6001 - 21s/epoch - 43ms/step
Epoch 3/30
500/500 - 21s - loss: 0.5955 - accuracy: 0.7616 - 21s/epoch - 43ms/step
Epoch 4/30
500/500 - 22s - loss: 0.4234 - accuracy: 0.8598 - 22s/epoch - 44ms/step
Epoch 5/30
```

Save the model using the save() function.

```
model.save("Text_Emotion.h5")
```

Let's test the model.

1. Create a list of sentences and write some random sentences for testing.

© 2021 - WhiteHat Education Technology Private Limited.

Note: This document is the original copyright of WhiteHat Education Technology Private Limited.



- Once this is done we need to tokenize these sentences by using the text to sequence() method. Perform padding using the pad_sequences() method.
- 3. Test the model using **predict()** method. Find the result using the **argmax()** method. The argmax() method will give the label of the class with the highest probability in the result.

It is returning the class labels for respective sentences.

As you can see the prediction of your model is almost right.

Let's test the model locally on your device:

- 1. As we saved the model, we have a .h5 file saved in our colab directory.
- 2. **Download** the file and **save** it in a folder.
- 3. Open <u>Student Activity 2</u> and download the boilerplate code. Paste it in the same folder.
- 4. Now open the boilerplate code in VS code.



5. Here we have to predict the sentiments of the given sentences written in the list called a **sentence**.

Note: The student can write any sentence.

Perform tokenization and padding on these sentences.



7. Use a .h5 file to load the model and use the **predict()** method to predict the sentiment of the text.

```
# Define the model using .h5 file
model=tensorflow.keras.models.load_model('Text_Emotion.h5')

# Test the model
result=model.predict(padded)
print(result)

# Print the result
predict_class = np.argmax(result, axis=1)
print(predict_class)

# {"anger": 0, "fear": 1, "joy": 2, "love": 3, "sadness": 4, "surprise": 5}
```

- 8. Go to the command prompt and traverse the working folder as we did in class 110.
- Create a virtual environment (Windows/Mac) using python -m venv <name_of_the _environment>for testing the model.
- 10. Activate the virtual environment using the following command:

<name_of_the_environment>\Scripts\activate

Run the command to create a virtual environment with the name "Text_Sentiment_Analysis". The environment is user-defined. We can keep the name as we want relevant to our project.

C:\whitehat jr\PRO C115\Text_Sentiment_Analysis>Scripts\activate

11. Now, run the desired python file to check the output. You can run the code stepwise and check the output one by one.

© 2021 - WhiteHat Education Technology Private Limited.

 ${\bf Note: This\ document\ is\ the\ original\ copyright\ of\ WhiteHat\ Education\ Technology\ Private\ Limited.}$



(Text_Sentiment_Analysis) C:\whitehat jr\PRO C115>python Text_Sentiment.py

Tokenized and padded sequences:

Result and prediction of the sentiment using labels:

```
[[1.26573211e-03 1.21657131e-03 2.26438437e-02 6.40540496e-02 1.04617124e-04 9.10715103e-01]
[4.97628838e-01 7.16012642e-02 3.41708571e-01 1.01922369e-02 7.38635287e-02 5.00548165e-03]]
[5 0]
```

Analyzing output:

As you can see the first array has six parameters since we have six classes.

Do you remember we had used the argmax() method?

Yes!! we had used the argmax() method for finding the maximum probability related to a particular class?

Here, you can see that after using the argmax() method we got output as [5 0].

5 corresponds to **surprise** and 0 corresponds to **anger**.

Since the ML model cannot be 100% correct we may not get the expected predictions always.

Great work!! This way we can create a Machine Learning model to predict the sentiments.

ESR: Yes, we had used it to find the class related to testing data



You can stop sharing your screen now.

Teacher Guides Student to Stop Screen Share

WRAP-UP SESSION - 05 mins

Teacher Starts Slideshow Slide 15 to 19

Activity details:

Following are the WRAP-UP session deliverables:

- Appreciate the student.
- Revise the current class activities.
- Discuss the quizzes.

WRAP-UP QUIZ

Click on In-Class Quiz



Continue WRAP-UP Session Slide 20 to 25

Activity Details

Following are the session deliverables:

- Explain the facts and trivia
- The next class challenge is to find how python is used for creating websites
- Project for the day
- Additional Activity (Optional)

FEEDBACK

- Appreciate and compliment the student for trying to learn a difficult concept.
- Get to know how they are feeling after the session.
- Review and check their understanding.

© 2021 - WhiteHat Education Technology Private Limited.

Note: This document is the original copyright of WhiteHat Education Technology Private Limited.



Teacher Action	Student Action
You get "hats-off" for your excellent work!	Make sure you have given at least 2 hats-off during the class for:
In the next class, We will be learning about Flask which is	Creatively Solved Activities +10 Great Question +10 Strong Concentration
used for creating web pages using Python.	O.
PROJECT OVERVIEW DISCUSSION Refer the document below in Activity Links Sections	
Teacher Clicks × End Class	



ACTIVITY LINKS		
Activity Name	Description	Links
Teacher Activity 1	Previous Class Code	https://colab.research.google.com/ drive/1-XnOC4oGfKqDT6HGdA1gu osGMlO6QWkT?usp=sharing
Teacher Activity 2	Boilerplate Code	https://colab.research.google.com/ drive/1sFb2RvAZJIMpdyKXkEK8M azW9HyG0gqx?usp=sharing
Teacher Activity 3	Embedding layer	https://www.tensorflow.org/api_doc s/python/tf/keras/layers/Embedding
Teacher Activity 4	Convolution 1D layer	https://www.tensorflow.org/api_doc s/python/tf/keras/layers/Conv1D
Teacher activity 5	Max Pooling layer	https://www.tensorflow.org/api_doc s/python/tf/keras/layers/MaxPool1D
Teacher Activity 6	LSTM layer	https://www.tensorflow.org/api_doc s/python/tf/keras/layers/LSTM
Teacher Reference 1	Reference Code	https://colab.research.google.com/ drive/1-kuwdZlqnczX1aSsFICzAt1 BC2g7P6H2?usp=sharing
Teacher Reference 2	Reference Code (VS Code)	https://github.com/procodingclass/P RO-C115-Reference-Code
Teacher Reference 3	Project	https://s3-whjr-curriculum-uploads. whjr.online/a2f7b916-9e9b-4476-84 70-90a4bde2627d.pdf
Teacher Reference 4	Project Solution	https://colab.research.google.com/ drive/1T3d8LQizWHHzP9wP_Rkgv TyWylDjFtCt?usp=sharing
Teacher Reference 5	Visual-Aid	https://s3-whjr-curriculum-uploads. whjr.online/7e0118ac-aba0-43e4-a 957-e628facd8167.html
Teacher Reference 6	In-Class Quiz	https://s3-whjr-curriculum-uploads.

^{© 2021 -} WhiteHat Education Technology Private Limited.

Note: This document is the original copyright of WhiteHat Education Technology Private Limited.



		whjr.online/c57b9df6-fb4f-4a2e-a9d 7-abf218c1277f.pdf
Student Activity 1	Boilerplate Code	https://colab.research.google.com/ drive/1sFb2RvAZJIMpdyKXkEK8M azW9HyG0gqx?usp=sharing
Student Activity 2	Boilerplate Code(VS Code)	https://github.com/procodingclass/P RO-C13-Student-Boilerplate-VS-co de-

