

The background is a solid blue color with abstract geometric shapes and network diagrams. In the top-left and bottom-right corners, there are white line graphs with nodes connected by lines, resembling a network or data structure. The main text is centered in the upper half of the image.

Document Image Classifier

Developing a document Image Classifier to accurately **classify** different types of documents

Prediction: ADVE | Real class: ADVE | Result: True



Given a document, the task is to predict which type of document it is.

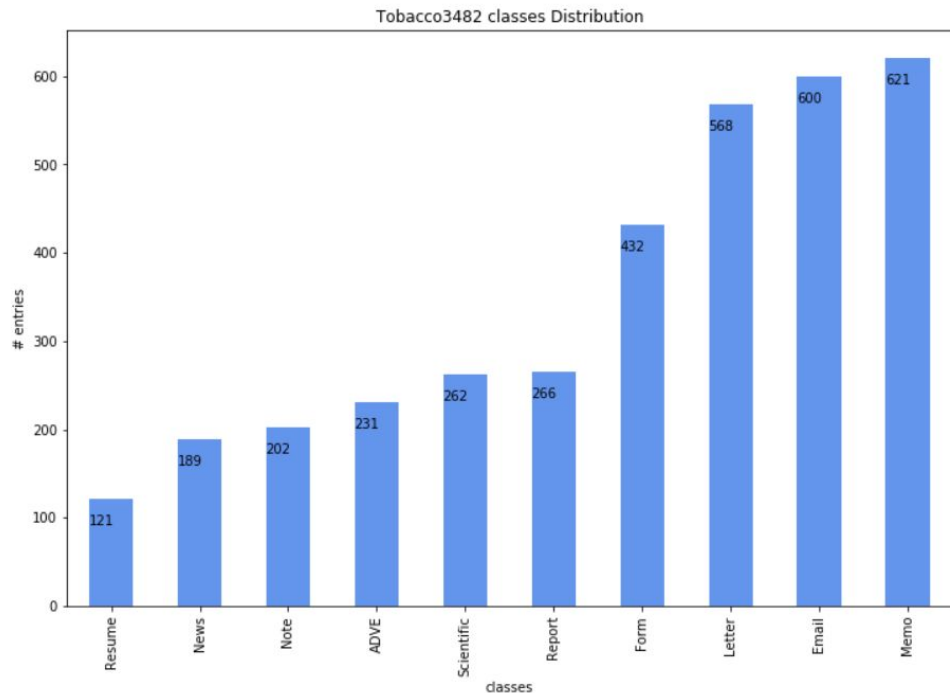
Overview of what the **dataset** looks like

Data Collection

The data is Tobacco3482 dataset .

There are 10 classes of documents

'Letter', 'Resume', 'Scientific', 'ADVE',
'Email', 'Report', 'News', 'Memo', 'Form',
'Note'



Implementation details

Preprocessing Images

Resize and transform to Numpy array

Label Binarizer

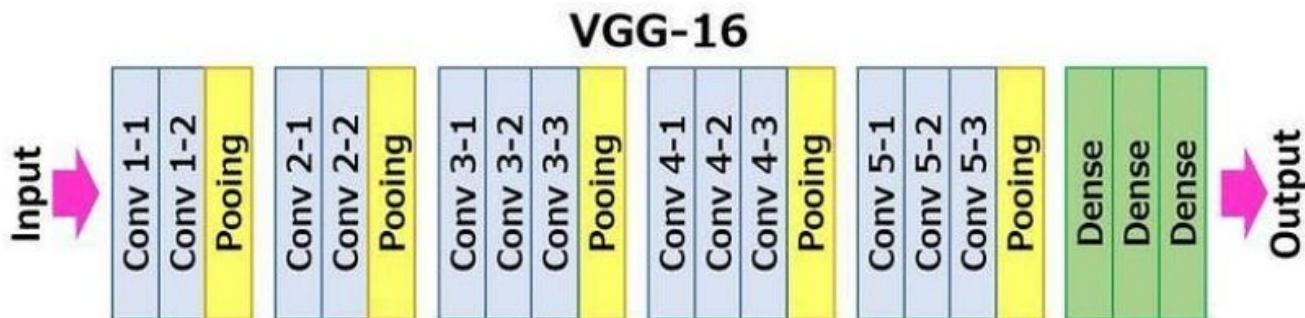
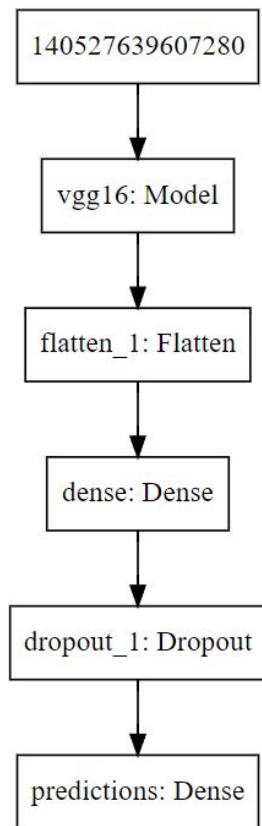
```
train image shape : (2437, 299, 299, 3)
train label shape : (2437, 10)
test image shape : (347, 299, 299, 3)
test label shape : (347, 10)
valdiation image shape : (696, 299, 299, 3)
Validation label shape :(696, 10)
```

ADVE

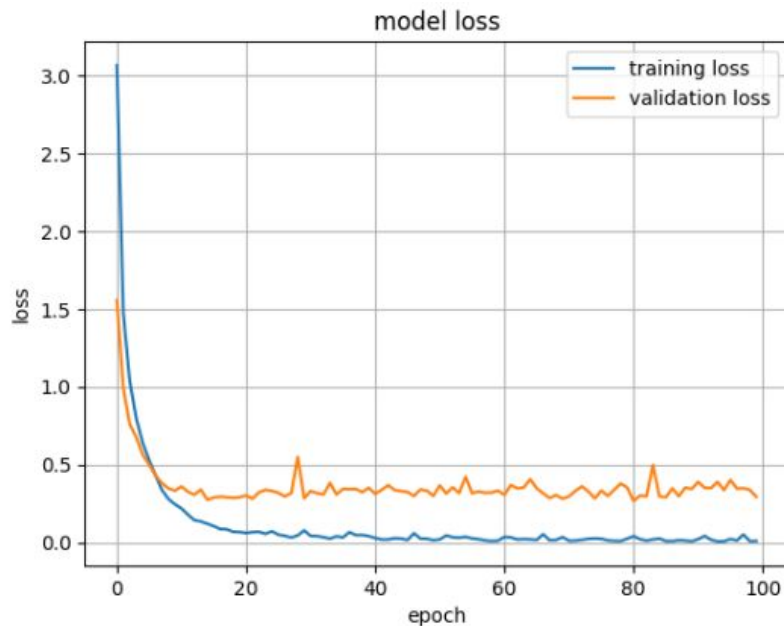
```
[1 0 0 0 0 0 0 0 0 0]
```

```
['ADVE' 'Email' 'Form' 'Letter' 'Memo' 'News' 'Note' 'Report' 'Resume' 'Scientific']
```

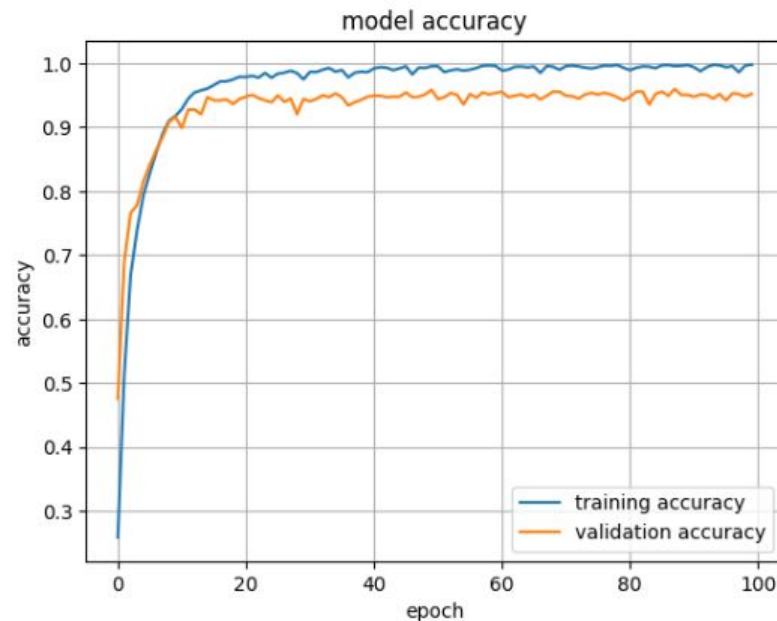
Implementation details: **Base architecture**



Model Evaluation



Adam optimizer
Categorical crossentropy loss
100 epochs



10 epoch
accuracy: 0.9175
val_accuracy: 0.9029

100 epoch
accuracy: 0.9977
val_accuracy: 0.9519

Results

Prediction: Email | Real class: Email | Result: True
Prediction: Email | Real class: Email | Result: True
Prediction: Memo | Real class: Memo | Result: True
Prediction: ADVE | Real class: ADVE | Result: True
Prediction: Form | Real class: Form | Result: True
Prediction: Report | Real class: Report | Result: True
Prediction: Memo | Real class: Note | Result: False
Prediction: Report | Real class: Report | Result: True
Prediction: Report | Real class: Resume | Result: False
Prediction: Letter | Real class: Letter | Result: True
Prediction: Email | Real class: Email | Result: True
Prediction: Form | Real class: Form | Result: True
Prediction: Form | Real class: Form | Result: True
Prediction: Form | Real class: Form | Result: True
Prediction: Note | Real class: Scientific | Result: False
Prediction: Memo | Real class: Memo | Result: True
Prediction: Scientific | Real class: Form | Result: False

Prediction: ADVE | Real class: ADVE | Result: True





THANK YOU!