Research in VGIS - Miniproject

Niclas Hjorth Stjernholm

November 30, 2018

1 Tensorflow Playground Tasks

1.1 Default settings

Circles data

The model converges very fast, within just around 150 iterations. The data is divided in a simple pattern, and with the given two hidden layers with four and two neurons, the pattern created from the data is learned in a fast manner. With this low a complexity of pattern only a few neurons a used to learn the pattern and to classify the data.

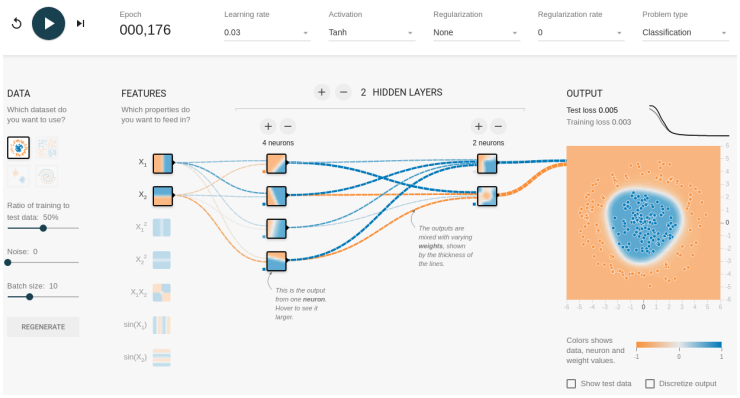


Figure 1: Default settings with a disc and a circle as data

Spiral data

The model has a training loss of just around 0.3 and a test loss even higher. This points to an underfitting which means the model is not able to properly fit to the data with the given capabilities.

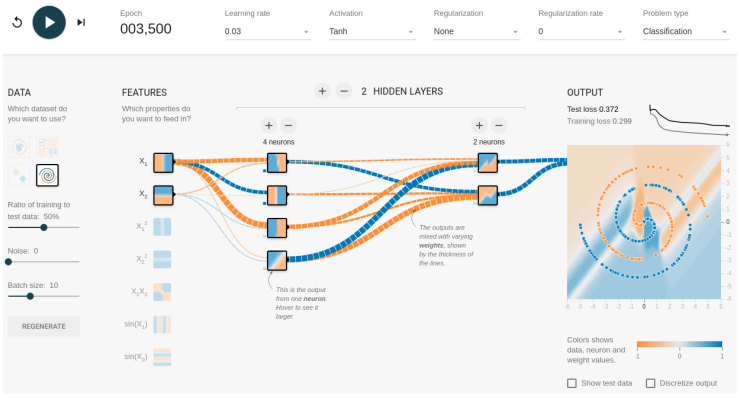


Figure 2: Default settings with the spiral data set

1.2 altered settings

By increasing the amount of neurons to eight and adding another layer with six neurons, the model is able to converge to the data and attain a fitting shape. The only other thing changed is the learning rate, changed from 0.03 to 0.01.

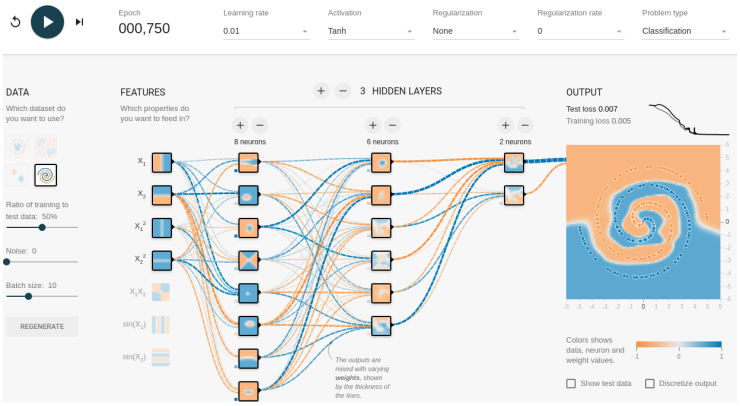
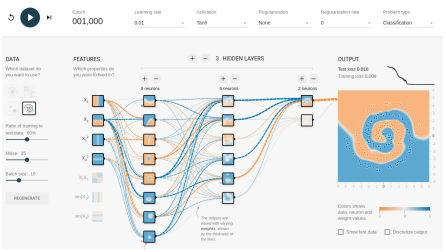


Figure 3: Altered settings with the spiral data

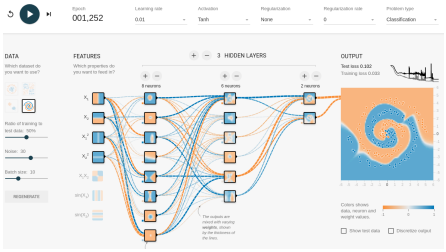
As seen in Figure 3, by adding the features X_1^2 and X_2^2 the model is able to form more complex outputs in the neurons.
With more neurons in the first hidden layer the model is once again able to generate higher complexity and by that weighting of the outputs to the next layer. With the extra layer with six neurons we are able to create more complex outputs, as seen in the figure. It is clear how the bottom neuron in the second hidden layer is already closing in on making a spiral. In the third hidden layer with just two neurons, it is again clear that a spiral is created, and the one closer to the spiral has a much higher weight than the other neuron output.

1.2.1 Adding Noise

When adding noise the model starts to break and is unstable at a noise level of 30. Figure 4 shows the the stable model at noise level 25 and the unstable model at noise level 30.



(a) Altered settings with the spiral data, adding noise at a level of 25



(b) Altered settings with the spiral data, adding noise at a level of 30

Figure 4: The two models with added noise

2 Quick, Draw! Doodle Recognition Challenge

kernel (1)

November 30, 2018

Firstly we import the nedded packages and libraries to run the python script. We then show the directories in which the data we need are in.

```
In [1]: import numpy as np # linear algebra
        import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

        from IPython.core.interactiveshell import InteractiveShell
        InteractiveShell.ast_node_interactivity = "all"
        import datetime as dt
        import matplotlib.pyplot as plt
        import seaborn as sns
        import ast
        from datetime import date, timedelta




        import pickle # Read/Write with Serialization
        import requests # Makes HTTP requests
        from io import BytesIO # Use When expecting bytes-like objects
        start = dt.datetime.now()


        import os
        print(os.listdir("../input"))

        # Any results you write to the current directory are saved as output.
```

```
['test_simplified.csv', 'sample_submission.csv', 'train_simplified', 'test_raw.csv']
```

The number of classes are easily found, looking at the amount of files in the *train_simplified* folder. Afterwards the amount of images in each class is found in a *for loop* by reading each csv file. The numbers are also used for counting the total amount of images.

```
In [2]: TRAIN_FILES_PATH = '../input/train_simplified/'
```

```
        trainingFileNameArr = os.listdir(TRAIN_FILES_PATH)
        totalClassesCount = len(trainingFileNameArr)
        totalImagesCount = 0

        print('classes:', totalClassesCount)

        #for trainingFileName in trainingFileNameArr:
        #  print(trainingFileName)
        #  dataset = pd.read_csv(TRAIN_FILES_PATH + trainingFileName, header=0).values
        #  print(dataset.shape[0])
        #  totalImagesCount += dataset.shape[0]

        print(trainingFileNameArr)
        print('images: 49707579')
        print('Average in each class: 146198.76')
```

```
classes: 340
['sleeping bag.csv', 'house plant.csv', 'bathtub.csv', 'key.csv', 'triangle.csv', 'grapes.csv',
images: 49707579
Average in each class: 146198.76
```

There are 340 classes and 49,707,579 images in total. The amount of images in each class is printed with the class name in the ouput above averaging at 146198.76 per class. ### Printing images from 8 classes. The first 8 classes has been selected for printing.

```
In [3]: for i in range(8):
            print(trainingFileNameArr[i]);
            path = os.path.join(TRAIN_FILES_PATH,trainingFileNameArr[i]);
            item = pd.read_csv(path);
            item['timestamp'] = pd.to_datetime(item.timestamp);
            item = item.sort_values(by='timestamp',ascending=False)[-20:]
            item['drawing'] = item['drawing'].apply(ast.literal_eval);

            item.head()

            n = 10
            fig, axs = plt.subplots(2, 10, figsize=(16, 3));
            for i, drawing in enumerate(item.drawing):
                ax = axs[i // n, i % n];
                for x, y in drawing:
                    ax.plot(x, -np.array(y), lw=3);
                ax.axis('off');

            fig.savefig('item.jpg', dpi=200)
            fig = plt.show();
```

```
sleeping bag.csv
```

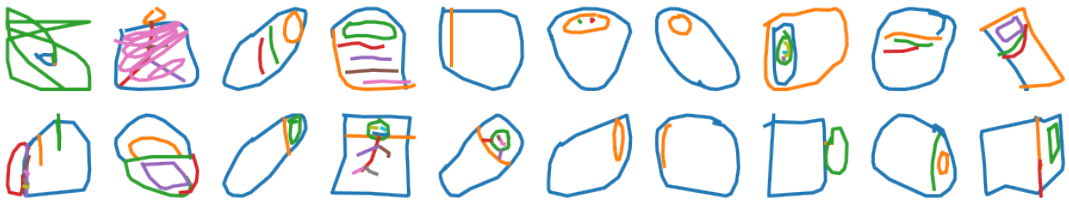
```
Out[3]:        countrycode      ...               word
        95303           US      ...       sleeping bag
        100359          US      ...       sleeping bag
        53606           JP      ...       sleeping bag
        38046           US      ...       sleeping bag
        69918           US      ...       sleeping bag

        [5 rows x 6 columns]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d7a518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d7aa58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d7ae48>]

Out[3]: (-12.75, 267.75, -129.15, 6.15)

Out[3]: [<matplotlib.lines.Line2D at 0x7f27555b2b00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d7a9e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d92780>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d92a20>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d92e48>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d9a2e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d9a748>]

Out[3]: (-6.95, 167.95, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8dad048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d9aba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8dad860>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8dad9e8>]

Out[3]: (-12.75, 267.75, -121.8, 5.800000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8db02e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8dade48>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8db0b00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8db0c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d35128>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d35588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d359e8>]

Out[3]: (-9.700000000000001, 203.7, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d3f2e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d35e48>]

Out[3]: (-12.75, 267.75, -71.4, 3.4000000000000004)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d3fb00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d3fc88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d504e0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d50668>]

Out[3]: (-7.7, 161.7, -267.7, 11.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d50ac8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d50f28>]

Out[3]: (-10.65, 223.65, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d58be0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d58908>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d5e160>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d5e2e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d5e748>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d5eba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d69048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d694a8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d69908>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d69d68>]

Out[3]: (-12.75, 267.75, -113.4, 5.4)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d70940>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d70668>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d70e80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cfa048>]

Out[3]: (-5.050000000000001, 106.05, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cfabe0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cfa908>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d00160>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d002e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d00748>]

Out[3]: (-12.25, 257.25, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d02048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d00ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d02860>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d029e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d02e48>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d122e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d12748>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d12ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d1b048>]

Out[3]: (-12.75, 267.75, -69.3, 3.3000000000000003)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d1bbe0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d1b908>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d22160>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d222e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d22748>]

Out[3]: (-6.45, 135.45, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d25048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d22ba8>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d25860>]

Out[3]: (-12.700000000000001, 266.7, -144.9, 6.9)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d259e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d25e48>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cb16a0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cb1828>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cb1c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cbe128>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cbe588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cbe9e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cbee48>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cc42e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ed15c0>]

Out[3]: (-6.25, 131.25, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cc4ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cc4fd0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cce828>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cce9b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ccee10>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ccd2b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ccd710>]

Out[3]: (-12.75, 267.75, -227.85, 10.850000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ccdb70>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ccdfd0>]

Out[3]: (-12.75, 267.75, -63.0, 3.0)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cddc88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cdd9b0>]
```

```
Out[3]: (-12.75, 267.75, -74.55, 3.5500000000000003)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ce7668>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ce7390>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ce7ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ce7d30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ce51d0>]

Out[3]: (-12.700000000000001, 266.7, -78.75, 3.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ce5d68>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ce5a90>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8c762e8>]

Out[3]: (-12.700000000000001, 266.7, -86.1, 4.1000000000000005)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8c76ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8c768d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8c7e128>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8c7e2b0>]

Out[3]: (-12.75, 267.75, -69.25, 2.25)
```



house plant.csv

```
Out[3]:        countrycode     ...              word
        120777          US     ...       house plant
        95668           US     ...       house plant
        34037           US     ...       house plant
        25792           GB     ...       house plant
        60900           US     ...       house plant

        [5 rows x 6 columns]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a91ad828>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a91ad9b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9194278>]

Out[3]: (-10.0, 210.0, -267.7, 11.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b6ecf8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2706103f60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9194f60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57a37b8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a939a828>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a939acc0>]

Out[3]: (-7.4, 155.4, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a939afd0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a93852b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9385ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a93857b8>]

Out[3]: (-5.7, 119.7, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a937c0b8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a937ca20>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a937cda0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a93ae320>]

Out[3]: (-5.7, 119.7, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a91adf60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a93ae828>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9398160>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a93aecf8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9398a90>]

Out[3]: (-6.45, 135.45, -267.75, 12.75)
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9398f98>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9398a20>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a93923c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a93927f0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9392c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9393470>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a93930f0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9393a58>]

Out[3]: (-12.75, 267.75, -227.85, 10.850000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e39438>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e395f8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9393e80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e39cf8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e52160>]

Out[3]: (-10.05, 211.05, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e52588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e52da0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e52898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e324a8>]

Out[3]: (-9.0, 189.0, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e32eb8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e328d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e41208>]

Out[3]: (-9.35, 196.35, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e41c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e41f28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e40470>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e41630>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e409b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e652b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e40f28>]

Out[3]: (-7.45, 156.45, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e65cc0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e657f0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e35390>]

Out[3]: (-8.25, 173.25, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e35710>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e35e80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e35e10>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e2d6a0>]

Out[3]: (-6.0, 126.0, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e2db70>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e2d400>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a25160>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a256d8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a25f60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a16128>]

Out[3]: (-5.15, 130.15, -267.65, 10.65)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a16978>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a25f28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a16c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a1c240>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a1c518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a1cb38>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a1c9e8>]

Out[3]: (-8.200000000000001, 172.2, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a26978>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a26358>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a26c88>]

Out[3]: (-8.700000000000001, 182.7, -267.7, 11.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a69f2198>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a69f2a58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a69f24a8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a69f7438>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a69f7048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a69f7908>]

Out[3]: (-6.6000000000000005, 138.6, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a69f7eb8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a12588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a120b8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a12f98>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60e03c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6a12ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60e0908>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60e0d30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d6668>]

Out[3]: (-11.65, 266.65, -263.55, 12.55)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d6c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d6198>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60e1390>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60e1780>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60e1c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60e1c50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d6f60>]

Out[3]: (-9.200000000000001, 193.2, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60dbb00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60dbc50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60dbf28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d27b8>]

Out[3]: (-4.800000000000001, 122.8, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d2a90>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d24a8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60c4668>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60c4240>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60c4b38>]

Out[3]: (-6.1000000000000005, 128.1, -267.75, 12.75)
```

bathtub.csv

```
Out[3]:        countrycode   ...        word
        145115          US   ...     bathtub
        51417           US   ...     bathtub
        80684           US   ...     bathtub
        75773           US   ...     bathtub
        22887           US   ...     bathtub

        [5 rows x 6 columns]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a0dee128>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90d3d30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90cc0b8>]

Out[3]: (-12.75, 267.75, -185.85, 8.85)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60c30f0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a0dee0f0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90ee390>]

Out[3]: (-12.200000000000001, 256.2, -100.75, 3.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90ee898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90ee630>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b9320>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b9208>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b9b00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b9e80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b9710>]

Out[3]: (-12.700000000000001, 266.7, -168.0, 8.0)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b8358>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b8c50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b8ba8>]

Out[3]: (-12.75, 267.75, -113.4, 5.4)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8eb0400>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8eb0940>]

Out[3]: (-12.75, 267.75, -139.65, 6.65)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b8780>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ea10b8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ea15c0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ea1dd8>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ea1ac8>]

Out[3]: (-12.75, 267.75, -128.1, 6.1000000000000005)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e78860>]

Out[3]: (-12.75, 267.75, -114.45, 5.45)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e78dd8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e789e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e787f0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e9f6d8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e9ffd0>]

Out[3]: (-12.700000000000001, 266.7, -143.85, 6.8500000000000005)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e78ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8eaa828>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8eaa400>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8eaae48>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7e2b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7e6d8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7ee80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e931d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e934a8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e938d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269a75eba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e93cf8>]

Out[3]: (-12.75, 267.75, -260.4, 12.4)

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b7fa58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b7f630>]

Out[3]: (-12.75, 267.75, -240.45, 11.450000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b7ff28>]
```

```
Out[3]: (-12.75, 267.75, -73.5, 3.5)

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b6c908>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b6ccf8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b6c128>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b62390>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b627b8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b62be0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b640f0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b64550>]

Out[3]: (-12.75, 267.75, -233.1, 11.100000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b64d68>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b64940>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b64f28>]

Out[3]: (-12.75, 267.75, -156.45, 7.45)

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b53da0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b53710>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b57240>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b57908>]

Out[3]: (-12.75, 267.75, -104.95, 3.95)

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b57ef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b4b198>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b4b898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b4bc50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b57518>]

Out[3]: (-11.700000000000001, 267.7, -112.35, 5.3500000000000005)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84e81d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84e84e0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84e8f28>]

Out[3]: (-11.700000000000001, 267.7, -124.9, 4.9)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84c64e0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699b53d30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84e8b70>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84c6ef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84c7588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84c6da0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84c7f60>]

Out[3]: (-11.15, 234.15, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84ea4a8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84ea828>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84ea8d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84eaef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84eada0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84c86a0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84c8f28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84c8b00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84d1780>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84d1470>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9967b8>]

Out[3]: (-12.75, 267.75, -184.8, 8.8)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84b95f8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84b99e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84b91d0>]

Out[3]: (-12.75, 267.75, -121.8, 5.800000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a84b9f28>]
```

Out[3]: [<matplotlib.lines.Line2D at 0x7f269981c780>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269981c438>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269981cd68>]

Out[3]: (-12.75, 267.75, -115.5, 5.5)

key.csv

Out[3]:

	countrycode	...	word
37389	NL	...	key
139408	US	...	key
27493	US	...	key
3925	US	...	key
75394	DE	...	key

[5 rows x 6 columns]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e39f60>]

Out[3]: (-11.450000000000001, 240.45, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e3f278>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bc4e320>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699811908>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e3fdd8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e3fc88>]

Out[3]: (-12.700000000000001, 266.7, -96.6, 4.6000000000000005)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e4ac88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e4a668>]

```
Out[3]: (-11.700000000000001, 267.7, -127.05, 6.050000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e4e438>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e4e3c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e4ec50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e4e908>]

Out[3]: (-6.550000000000001, 137.55, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e4f710>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e4f2e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e4fda0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e4fbe0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e5e6d8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e5ed30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5e5e588>]

Out[3]: (-12.75, 267.75, -81.9, 3.9000000000000004)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96cf438>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96cf908>]

Out[3]: (-11.05, 232.05, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96cf2e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96ca470>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96cff60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96cadd8>]

Out[3]: (-12.75, 267.75, -82.95, 3.95)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96caac8>]

Out[3]: (-5.6000000000000005, 117.6, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96b3a58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96b30f0>]

Out[3]: (-4.55, 95.55, -267.75, 12.75)
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96c62b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96b3ef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96c6898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96c6b70>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96ec400>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96ec0b8>]

Out[3]: (-11.700000000000001, 267.7, -81.9, 3.9000000000000004)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96ece10>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96bb208>]

Out[3]: (-12.700000000000001, 266.7, -82.95, 3.95)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96ece80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96bbb00>]

Out[3]: (-12.75, 267.75, -95.55, 4.55)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96c93c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96bbef0>]

Out[3]: (-10.75, 225.75, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96c9b38>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96c9208>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57d4518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57d4080>]

Out[3]: (-12.700000000000001, 266.7, -233.05, 10.05)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57d49e8>]

Out[3]: (-11.700000000000001, 267.7, -85.0, 3.0)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96bb898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57d4e10>]

Out[3]: (-6.800000000000001, 142.8, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57d7dd8>]
```

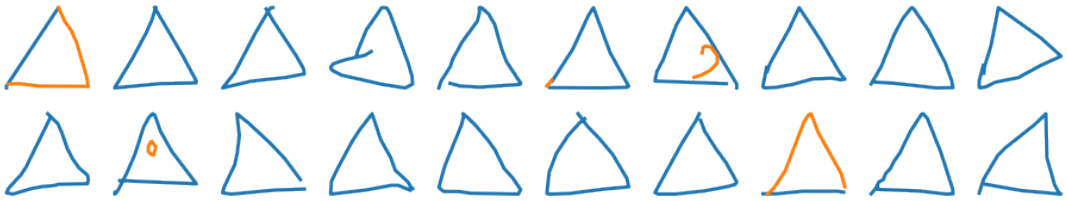
```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57d4ef0>]

Out[3]: (-12.75, 267.75, -118.6, 4.6000000000000005)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57d28d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57d79e8>]

Out[3]: (-6.25, 131.25, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57e4160>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57d2cf8>]

Out[3]: (-4.0, 84.0, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57e4080>]

Out[3]: (-5.7, 119.7, -267.75, 12.75)
```

triangle.csv

```
Out[3]:        countrycode    ...         word
        104398          RU    ...     triangle
        31368           US    ...     triangle
        113292          US    ...     triangle
        105446          US    ...     triangle
        106496          VN    ...     triangle

        [5 rows x 6 columns]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a91ce048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a91ce5f8>]

Out[3]: (-8.65, 181.65, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfa4f98>]
```

```
Out[3]: (-12.75, 267.75, -263.55, 12.55)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a91cedd8>]

Out[3]: (-11.700000000000001, 267.7, -263.55, 12.55)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfbf400>]

Out[3]: (-12.75, 267.75, -261.45, 12.450000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfbf240>]

Out[3]: (-12.75, 267.75, -213.1, 9.100000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfbf8d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfbfdd8>]

Out[3]: (-12.75, 267.75, -250.95, 11.950000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9d19b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9d1ba8>]

Out[3]: (-7.6, 181.6, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9d1e48>]

Out[3]: (-11.450000000000001, 240.45, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9e7048>]

Out[3]: (-12.350000000000001, 259.35, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9e7e48>]

Out[3]: (-10.25, 215.25, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9e7f60>]

Out[3]: (-10.3, 216.3, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9cb048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9cbb70>]

Out[3]: (-12.75, 267.75, -155.4, 7.4)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9cbf98>]

Out[3]: (-10.05, 211.05, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9c1198>]
```

```
Out[3]: (-12.75, 267.75, -213.15, 10.15)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9c1e80>]

Out[3]: (-10.5, 220.5, -267.7, 11.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9c1f60>]

Out[3]: (-8.65, 181.65, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9fa240>]

Out[3]: (-12.75, 267.75, -216.3, 10.3)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9fae80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9c1518>]

Out[3]: (-10.4, 218.4, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9da7f0>]

Out[3]: (-10.850000000000001, 227.85, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9dae10>]

Out[3]: (-8.9, 186.9, -267.75, 12.75)
```



grapes.csv

```
Out[3]:       countrycode   ...      word
        23112          US   ...    grapes
        61144          US   ...    grapes
        60593          US   ...    grapes
        30364          US   ...    grapes
        36352          US   ...    grapes

        [5 rows x 6 columns]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a61d0c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8de8c18>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57fdda0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bc72cc0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57fd9e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a580d278>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57fd400>]

Out[3]: (-7.300000000000001, 153.3, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a580d860>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a580de80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5826588>]

Out[3]: (-6.15, 129.15, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a61d0c18>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a58269e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5827358>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5827048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5827a90>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5827d68>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a58116d8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5811208>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5811f28>]

Out[3]: (-7.75, 162.75, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a58045c0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5804748>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5804e48>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5804fd0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a58048d0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a5804e10>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8771f28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8771b00>]

Out[3]: (-8.3, 174.3, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8797550>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8797c18>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8797eb8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8797cc0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8773b00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8773c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a87883c8>]

Out[3]: (-10.200000000000001, 214.2, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8773710>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a87885f8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8788d68>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a87a55f8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a87a51d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a87a5c50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a87a5b00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a87aa898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a87aa400>]

Out[3]: (-12.700000000000001, 266.7, -147.0, 7.0)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60cb240>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a87aae48>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a87aa828>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60cbc50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60cbb00>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60b4400>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60b4550>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60b4dd8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60b4c50>]

Out[3]: (-11.950000000000001, 250.95, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60bea58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60bedd8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60ba358>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60be630>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60ba898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60bafd0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60e7630>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60bacc0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60e7ef0>]

Out[3]: (-12.75, 267.75, -193.2, 9.200000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60e75c0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d10b8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d1978>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d1e48>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d14518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d14978>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a60d1c50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d13320>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d13668>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d14ef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269a77bbe0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d11390>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d110b8>]

Out[3]: (-12.75, 267.75, -152.25, 7.25)

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d11a58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d11da0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d0b390>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d0b780>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d0beb8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d2d1d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d0bba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d2dc50>]

Out[3]: (-5.65, 118.65, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d2def0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d32518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d320f0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d32f28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d32c18>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6790518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6790c50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67a3198>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67a3160>]

Out[3]: (-5.2, 131.2, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6790940>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6790f60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a677b320>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a677b048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a677ba58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a677bef0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67785f8>]

Out[3]: (-10.55, 221.55, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a677bd30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6778e10>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6789588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6789518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67899b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6789ef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a679a5f8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6789d68>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a679ab00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6320390>]

Out[3]: (-12.75, 267.75, -243.55, 10.55)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6320a20>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a679af28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2698d11d30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6303630>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a63035f8>]

Out[3]: (-7.15, 150.15, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6303c50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6303f28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6322550>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6322940>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6322d68>]

Out[3]: (-8.5, 178.5, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6310358>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a63106a0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6310b70>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a62f4400>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a62f4048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a62f4da0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a62f4d68>]

Out[3]: (-8.5, 178.5, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a631a748>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a6310630>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a62f4860>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269972c438>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269972c978>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269972c048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269972c908>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699709208>]

Out[3]: (-10.700000000000001, 224.7, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699709978>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699709630>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699709f60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269972e438>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269972eba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269972e898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699716208>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699716780>]

Out[3]: (-9.75, 204.75, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699716f98>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699710240>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699716a58>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f2699710908>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699731240>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699710e80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699731630>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699731be0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1df390>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699731eb8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9175e80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1dfdd8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1dfd30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1ed550>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1edd30>]

Out[3]: (-10.55, 221.55, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1eb4a8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1eb3c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a631a390>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1edf98>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1f64e0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1f60f0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1f6eb8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1f6a20>]

Out[3]: (-7.800000000000001, 163.8, -266.7, 12.700000000000001)
```

```
monkey.csv
```

```
Out[3]:        countrycode   ...      word
        22343           GB   ...    monkey
        68156           SE   ...    monkey
        35180           US   ...    monkey
        106105          GB   ...    monkey
        13416           US   ...    monkey

        [5 rows x 6 columns]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9e1080>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9e11d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9e1940>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a9398160>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9e1518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9c7780>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9c7390>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9c7dd8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9ed550>]

Out[3]: (-10.200000000000001, 214.2, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9c8fd0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9edeb8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9ed9e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9ff6d8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9ffb70>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9ff3c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699820518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699820940>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699820128>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f2699820f28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699811dd8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269982ea58>]

Out[3]: (-12.75, 267.75, -238.35, 11.350000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269982ecc0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b9c7c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699820dd8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269983fa90>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269983ff28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269981d588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269981da90>]

Out[3]: (-9.100000000000001, 213.1, -265.65, 12.65)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269981da58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269983f940>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699824470>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699824a58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f2699824d30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269981f240>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269981f668>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269981fe10>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b972e8>]

Out[3]: (-12.75, 267.75, -260.4, 12.4)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b979b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269982eeb8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b97f28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b98588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b98be0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f269981fb00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b71358>]

Out[3]: (-9.5, 199.5, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b71a90>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b97940>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b71f28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b7f710>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b7f6a0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b7ff98>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b7fac8>]

Out[3]: (-9.65, 224.65, -265.6, 11.600000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b9e588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b7fef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b9ecf8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b9ee48>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ba2940>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ba2630>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8ba2f60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e906d8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e903c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e90e10>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a09af128>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e79128>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e79a58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e79e80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e995f8>]

Out[3]: (-12.75, 267.75, -265.65, 12.65)
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e99ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e99d68>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7a518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7a4e0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7ad68>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7ad30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e78710>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e992e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e78ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7c048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a02f0390>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7cd30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7cd68>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e7c438>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b5ac8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b5c50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90e70b8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90e7390>]

Out[3]: (-12.75, 267.75, -165.9, 7.9)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90e7c50>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90e7be0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b7160>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b78d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b7588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90d7320>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90d7518>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90d7b38>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90d7dd8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90b7ef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bc72dd8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90e0978>]

Out[3]: (-12.75, 267.75, -186.9, 8.9)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90e0eb8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90cd860>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90cd1d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90cde10>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bef550>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8befa58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90cd7b8>]

Out[3]: (-10.700000000000001, 224.7, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a90e7c88>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bef710>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bf0b70>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bf07b8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8be2550>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8be24e0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8be2e10>]

Out[3]: (-12.75, 267.75, -260.35, 11.350000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8be6470>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8be65f8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8beff60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8be2940>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bdb400>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bdb0b8>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bdb978>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bdbf28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bed748>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bdbdd8>]

Out[3]: (-12.75, 267.75, -218.4, 10.4)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bedcc0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfba4a8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfba860>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfbada0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfbaf60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf885c0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8bedf98>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf88b38>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfae048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfae748>]

Out[3]: (-12.700000000000001, 266.7, -219.45, 10.450000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfae438>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfaea20>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf9f6d8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf9f6a0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf9fef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf83470>]

Out[3]: (-12.75, 267.75, -227.85, 10.850000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf9f2b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf83be0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf83f60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf9d630>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf9da90>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bf9df98>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d05390>]

Out[3]: (-12.75, 267.75, -158.45, 5.45)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d058d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269bfaecf8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d05f28>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d255c0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d25a90>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d05dd8>]

Out[3]: (-12.75, 267.75, -212.1, 10.100000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cf6358>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cf6048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cf6da0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cf6fd0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d0c710>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d0ca20>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cf69b0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cf1048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cf1470>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8cf1a20>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8b31048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d2e208>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d2e908>]

Out[3]: (-12.75, 267.75, -242.55, 11.55)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8d2ed68>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67c1048>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67c1908>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67c13c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67c18d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67d60b8>]

Out[3]: (-12.75, 267.75, -201.55, 8.55)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67d6ac8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67d6a20>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67d65f8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67cd710>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67cd400>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67bc128>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67cdcc0>]

Out[3]: (-6.4, 156.4, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67bcbe0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67bc5c0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67d73c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67d7358>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67d7860>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67d7cc0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a67e24a8>]

Out[3]: (-10.200000000000001, 214.2, -266.7, 12.700000000000001)
```

```
television.csv
```

```
Out[3]:       countrycode     ...            word
        3022           US     ...      television
        1377           US     ...      television
        21348          US     ...      television
        75236          CA     ...      television
        74407          AR     ...      television

        [5 rows x 6 columns]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8fb0978>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f269b99f898>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8fb0e10>]
```

```
Out[3]: (-12.75, 267.75, -199.5, 9.5)
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8f96828>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8e75d30>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8f96e10>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8fab518>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8f96cc0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8fabb00>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8fabf28>]
```

```
Out[3]: (-12.75, 267.75, -212.1, 10.100000000000001)
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8f9e3c8>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8f9e7f0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a8f9ecc0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26980d3128>]
```

```
Out[3]: (-12.3, 258.3, -267.75, 12.75)
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26980d3c50>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26980d35c0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26980fe4a8>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26980fe3c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980fe898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980fefd0>]

Out[3]: (-11.700000000000001, 245.7, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980fecc0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980c63c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980c6b38>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980f5470>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980f5128>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980f5400>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980f5e80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980cd6a0>]

Out[3]: (-12.75, 267.75, -184.8, 8.8)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980cdb70>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980d3be0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26980cdf60>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1c7748>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1c73c8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1db198>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1c7d30>]

Out[3]: (-10.3, 216.3, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1db780>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1db630>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1e84a8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1e8438>]

Out[3]: (-12.75, 267.75, -227.85, 10.850000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1e8ef0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1e8898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1cc748>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1cc668>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1ccfd0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1ccb38>]

Out[3]: (-12.700000000000001, 266.7, -242.55, 11.55)

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1f1898>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1f1470>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1f92e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1f1ef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1f9828>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f269b1f9b00>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f6048>]

Out[3]: (-9.35, 196.35, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f6860>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f6438>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a581e208>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f6d68>]

Out[3]: (-8.8, 184.8, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a581ed30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a581e710>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f5588>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f54a8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f5978>]

Out[3]: (-10.75, 225.75, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f3470>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f5da0>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f3780>]

Out[3]: (-12.75, 267.75, -235.2, 11.200000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f3d30>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57f3fd0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57fc8d0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57fc5c0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57fcef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57c8358>]

Out[3]: (-12.75, 267.75, -219.45, 10.450000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57c8cf8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57c8ba8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57ba320>]

Out[3]: (-10.4, 218.4, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57bacf8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57ba978>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57b9048>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57b94e0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57b9908>]

Out[3]: (-11.700000000000001, 267.7, -219.4, 9.4)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57ac390>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57ac320>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57ac6a0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57aceb8>]

Out[3]: (-12.75, 267.75, -201.6, 9.600000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57acb70>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57be7f0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57be208>]
```

```
Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57bee80>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57c05f8>]

Out[3]: (-7.65, 160.65, -266.7, 12.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57c0be0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57c0ef0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96d85c0>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a57c02e8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96d8d30>]

Out[3]: (-9.850000000000001, 206.85, -267.75, 12.75)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96c54a8>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96d8748>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96c5c18>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96c5ba8>]

Out[3]: (-9.950000000000001, 208.95, -267.7, 11.700000000000001)

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96d4550>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96d4278>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96d4d68>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96d4a58>]

Out[3]: [<matplotlib.lines.Line2D at 0x7f26a96d56d8>]

Out[3]: (-12.75, 267.75, -243.6, 11.600000000000001)
```

0.0.1 Preparing data for network

In [4]:
```
%reset -f
```

In [5]:
```
#%% import
import os
from glob import glob
import re
import ast
import numpy as np
import pandas as pd
from PIL import Image, ImageDraw
from tqdm import tqdm
from dask import bag

import tensorflow as tf
from tensorflow import keras
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Dropout, Flatten
from tensorflow.keras.layers import Conv2D, MaxPooling2D
from tensorflow.keras.metrics import top_k_categorical_accuracy
from tensorflow.keras.callbacks import ModelCheckpoint, ReduceLROnPlateau, EarlyStopping
```

In [6]:
```
#%% set label dictionary and params
classfiles = os.listdir('../input/train_simplified/')
numstonames = {i: v[:-4].replace(" ", "_") for i, v in enumerate(classfiles)} #adds unde

num_classes = 340    #340 max
imheight, imwidth = 32, 32
ims_per_class = 2000  #max?
```

In [7]:
```
# faster conversion function
def draw_it(strokes):
    image = Image.new("P", (256,256), color=255)
    image_draw = ImageDraw.Draw(image)
    for stroke in ast.literal_eval(strokes):
        for i in range(len(stroke[0])-1):
            image_draw.line([stroke[0][i],
                             stroke[1][i],
                             stroke[0][i+1],
                             stroke[1][i+1]],
                            fill=0, width=5)
    image = image.resize((imheight, imwidth))
    return np.array(image)/255.

#%% get train arrays
train_grand = []
class_paths = glob('../input/train_simplified/*.csv')
for i,c in enumerate(tqdm(class_paths[0: num_classes])):
```

```
            train = pd.read_csv(c, usecols=['drawing', 'recognized'], nrows=ims_per_class*5//4)
            train = train[train.recognized == True].head(ims_per_class)
            imagebag = bag.from_sequence(train.drawing.values).map(draw_it)
            trainarray = np.array(imagebag.compute())  # PARALLELIZE
            trainarray = np.reshape(trainarray, (ims_per_class, -1))
            labelarray = np.full((train.shape[0], 1), i)
            trainarray = np.concatenate((labelarray, trainarray), axis=1)
            train_grand.append(trainarray)

        train_grand = np.array([train_grand.pop() for i in np.arange(num_classes)]) #less memory
        train_grand = train_grand.reshape((-1, (imheight*imwidth+1)))

        del trainarray
        del train
```

```
100%|| 340/340 [11:15<00:00,  2.14s/it]
```

```
In [8]: # memory-friendly alternative to train_test_split?
        valfrac = 0.1
        cutpt = int(valfrac * train_grand.shape[0])

        np.random.shuffle(train_grand)
        y_train, X_train = train_grand[cutpt: , 0], train_grand[cutpt: , 1:]
        y_val, X_val = train_grand[0:cutpt, 0], train_grand[0:cutpt, 1:] #validation set is reco

        del train_grand

        y_train = keras.utils.to_categorical(y_train, num_classes)
        X_train = X_train.reshape(X_train.shape[0], imheight, imwidth, 1)
        y_val = keras.utils.to_categorical(y_val, num_classes)
        X_val = X_val.reshape(X_val.shape[0], imheight, imwidth, 1)

        print(y_train.shape, "\n",
              X_train.shape, "\n",
              y_val.shape, "\n",
              X_val.shape)
```

```
(612000, 340)
 (612000, 32, 32, 1)
 (68000, 340)
 (68000, 32, 32, 1)
```

```
In [9]: model = Sequential()
        model.add(Conv2D(32, kernel_size=(3, 3), padding='same', activation='relu', input_shape=
        model.add(MaxPooling2D(pool_size=(2, 2)))
        model.add(Conv2D(64, kernel_size=(3, 3), padding='same', activation='relu'))
        model.add(MaxPooling2D(pool_size=(2, 2)))
```

```
        model.add(Dropout(0.2))
        model.add(Flatten())
        model.add(Dense(680, activation='relu'))
        model.add(Dropout(0.5))
        model.add(Dense(num_classes, activation='softmax'))
        model.summary()
```

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 32, 32, 32)        320
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 16, 16, 32)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 16, 16, 64)        18496
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 8, 8, 64)          0
_________________________________________________________________
dropout (Dropout)            (None, 8, 8, 64)          0
_________________________________________________________________
flatten (Flatten)            (None, 4096)              0
_________________________________________________________________
dense (Dense)                (None, 680)               2785960
_________________________________________________________________
dropout_1 (Dropout)          (None, 680)               0
_________________________________________________________________
dense_1 (Dense)              (None, 340)               231540
=================================================================
Total params: 3,036,316
Trainable params: 3,036,316
Non-trainable params: 0
_________________________________________________________________
```

```
In [10]: def top_3_accuracy(x,y):
             t3 = top_k_categorical_accuracy(x,y, 3)
             return t3

         reduceLROnPlat = ReduceLROnPlateau(monitor='val_loss', factor=0.5, patience=3,
                                            verbose=1, mode='auto', min_delta=0.005, cooldown=5,
         earlystop = EarlyStopping(monitor='val_top_3_accuracy', mode='max', patience=5)
         callbacks = [reduceLROnPlat, earlystop]

         model.compile(loss='categorical_crossentropy',
                       optimizer='adam',
                       metrics=['accuracy', top_3_accuracy])

         model.fit(x=X_train, y=y_train,
```

```
                    batch_size = 32,
                    epochs = 22,
                    validation_data = (X_val, y_val),
                    callbacks = callbacks,
                    verbose = 1)

Train on 612000 samples, validate on 68000 samples
Epoch 1/22
612000/612000 [==============================] - 190s 310us/step - loss: 3.0555 - acc: 0.3142 -
Epoch 2/22
612000/612000 [==============================] - 189s 308us/step - loss: 2.4344 - acc: 0.4169 -
Epoch 3/22
612000/612000 [==============================] - 189s 308us/step - loss: 2.2981 - acc: 0.4427 -
Epoch 4/22
612000/612000 [==============================] - 188s 307us/step - loss: 2.2229 - acc: 0.4573 -
Epoch 5/22
612000/612000 [==============================] - 187s 305us/step - loss: 2.1768 - acc: 0.4678 -
Epoch 6/22
612000/612000 [==============================] - 187s 306us/step - loss: 2.1471 - acc: 0.4731 -
Epoch 7/22
612000/612000 [==============================] - 188s 308us/step - loss: 2.1263 - acc: 0.4768 -
Epoch 8/22
612000/612000 [==============================] - 188s 307us/step - loss: 2.1104 - acc: 0.4809 -
Epoch 9/22
612000/612000 [==============================] - 187s 306us/step - loss: 2.1009 - acc: 0.4828 -
Epoch 10/22
612000/612000 [==============================] - 186s 305us/step - loss: 2.0918 - acc: 0.4846 -
Epoch 11/22
611968/612000 [==============================>.] - ETA: 0s - loss: 2.0864 - acc: 0.4858 - top_3_ac
Epoch 00011: ReduceLROnPlateau reducing learning rate to 0.0005000000237487257.
612000/612000 [==============================] - 188s 307us/step - loss: 2.0864 - acc: 0.4858 -
Epoch 12/22
612000/612000 [==============================] - 188s 307us/step - loss: 1.9707 - acc: 0.5080 -
Epoch 13/22
612000/612000 [==============================] - 188s 307us/step - loss: 1.9391 - acc: 0.5150 -
Epoch 14/22
612000/612000 [==============================] - 187s 306us/step - loss: 1.9250 - acc: 0.5175 -
Epoch 15/22
612000/612000 [==============================] - 187s 305us/step - loss: 1.9133 - acc: 0.5191 -
Epoch 16/22
612000/612000 [==============================] - 187s 305us/step - loss: 1.9051 - acc: 0.5205 -
Epoch 17/22
612000/612000 [==============================] - 187s 306us/step - loss: 1.9016 - acc: 0.5223 -
Epoch 18/22
612000/612000 [==============================] - 189s 308us/step - loss: 1.8924 - acc: 0.5238 -
Epoch 19/22
612000/612000 [==============================] - 188s 307us/step - loss: 1.8888 - acc: 0.5247 -
Epoch 20/22
```

```
612000/612000 [==============================] - 187s 305us/step - loss: 1.8822 - acc: 0.5253 -
Epoch 21/22
611872/612000 [============================>.] - ETA: 0s - loss: 1.8803 - acc: 0.5261 - top_3_ac
Epoch 00021: ReduceLROnPlateau reducing learning rate to 0.00025000001187436283.
612000/612000 [==============================] - 186s 305us/step - loss: 1.8803 - acc: 0.5261 -
Epoch 22/22
612000/612000 [==============================] - 187s 306us/step - loss: 1.8126 - acc: 0.5389 -
```

```
Out[10]: <tensorflow.python.keras.callbacks.History at 0x7f26a658ce80>
```

0.1 Predicting on the Test data

The CNN does OK on the validation data, even with a basic model and limited training data. Let's generate predictions on the test set and submit.

```
In [11]: #%% get test set
         ttvlist = []
         reader = pd.read_csv('../input/test_simplified.csv', index_col=['key_id'],
             chunksize=2048)
         for chunk in tqdm(reader, total=55):
             imagebag = bag.from_sequence(chunk.drawing.values).map(draw_it)
             testarray = np.array(imagebag.compute())
             testarray = np.reshape(testarray, (testarray.shape[0], imheight, imwidth, 1))
             testpreds = model.predict(testarray, verbose=0)
             ttvs = np.argsort(-testpreds)[:, 0:3]  # top 3
             ttvlist.append(ttvs)

         ttvarray = np.concatenate(ttvlist)
```

```
100%|| 55/55 [02:14<00:00,  2.34s/it]
```

```
In [12]: preds_df = pd.DataFrame({'first': ttvarray[:,0], 'second': ttvarray[:,1], 'third': ttva
         preds_df = preds_df.replace(numstonames)
         preds_df['words'] = preds_df['first'] + " " + preds_df['second'] + " " + preds_df['thir

         sub = pd.read_csv('../input/sample_submission.csv', index_col=['key_id'])
         sub['word'] = preds_df.words.values
         sub.to_csv('subcnn_small.csv')
         sub.head()
```

```
Out[12]:                                                  word
         key_id
         9000003627287624                radio motorbike stereo
         9000010688666847              hockey_puck pool sandwich
         9000023642890129  The_Great_Wall_of_China camel castle
         9000038588854897                mountain tent triangle
         9000052667981386            campfire fireplace feather
```

```
In [13]: import sys

         # These are the usual ipython objects, including this one you are creating
         ipython_vars = ['In', 'Out', 'exit', 'quit', 'get_ipython', 'ipython_vars']

         # Get a sorted list of the objects and their sizes
         sorted([(x, sys.getsizeof(globals().get(x))) for x in dir() if not
             x.startswith('_') and x not in sys.modules and x
             not in ipython_vars], key=lambda x: x[1], reverse=True)

Out[13]: [('preds_df', 30572830),
          ('sub', 9900835),
          ('ttvarray', 2692888),
          ('testpreds', 2185632),
          ('chunk', 1053432),
          ('labelarray', 16112),
          ('numstonames', 9320),
          ('class_paths', 3104),
          ('Sequential', 3096),
          ('classfiles', 2896),
          ('Dense', 2000),
          ('Dropout', 2000),
          ('Flatten', 2000),
          ('MaxPooling2D', 2000),
          ('EarlyStopping', 1464),
          ('ReduceLROnPlateau', 1464),
          ('ModelCheckpoint', 1056),
          ('Conv2D', 888),
          ('ttvlist', 528),
          ('X_train', 144),
          ('X_val', 144),
          ('testarray', 144),
          ('draw_it', 136),
          ('top_3_accuracy', 136),
          ('top_k_categorical_accuracy', 136),
          ('ttvs', 112),
          ('y_train', 112),
          ('y_val', 112),
          ('c', 86),
          ('Image', 80),
          ('ImageDraw', 80),
          ('bag', 80),
          ('callbacks', 80),
          ('keras', 80),
          ('np', 80),
          ('pd', 80),
          ('tf', 80),
          ('earlystop', 56),
```

```
('imagebag', 56),
('model', 56),
('reader', 56),
('reduceLROnPlat', 56),
('cutpt', 28),
('i', 28),
('imheight', 28),
('ims_per_class', 28),
('imwidth', 28),
('num_classes', 28),
('valfrac', 24)]
```