In [1]:

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import OneHotEncoder
from sklearn.model_selection import train_test_split
from sklearn.neural_network import MLPClassifier
from sklearn.metrics import confusion_matrix
import matplotlib.pylab as plt
import seaborn as sns
from sklearn.decomposition import PCA
from sklearn.preprocessing import OneHotEncoder
import warnings
warnings.filterwarnings("ignore")
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis as LDA
```

Procedure

The data is explored for statistics and datatypes.

The numeric columns are extracted and binned into min-25%, 25%-50,50%-75% and 75%-max ranges and replaced with the range mean.

The job roles are clubbed using similarities.

All non-numeric columns are one-hot encoded

In [2]:

```
def binRanges(x,stats):
    if x<=stats[4]:
        return (stats[3]+stats[4])/2
    elif stats[4]<x<=stats[5]:
        return (stats[4]+stats[5])/2
    elif stats[5]<x<=stats[6]:
        return (stats[5]+stats[6])/2
    elif x>stats[6]:
        return (stats[6]+stats[7])/2

def club(role):
    if role == 'Developer' or role=='Engineer' or role == 'Testing' or role == 'Assoc
        return 'Engineer'
    elif role == 'Administrator' or role == 'Support':
        return 'system admin'
    elif role == 'UX' or role=='Architect' or role=='Designer':
        return 'Designer'
    elif role=='Auditor' or role=='Manager':
        return 'Manager'
    else:
        return role
def grid(x_,test_size,alphas):
    for t in test_size:
        X_train, X_test, y_train, y_test = train_test_split(x_, y, stratify=y,random_
        for a in alphas:
            clf = MLPClassifier(hidden_layer_sizes=(300,200,100),alpha=a ,max_iter=20
            print(f"""score with test size: {t} and alpha:{a}:    {clf.score(X_test,y_
            print(f"""confusion-matrix with test size: {t} and alpha:{a}:""")
            print(confusion_matrix(clf.predict(X_test),y_test))
```

In [3]:

```
df = pd.read_csv('roo_data.csv')
df.head()
```

Out[3]:

	Acedamic percentage in Operating Systems	percentage in Algorithms	Percentage in Programming Concepts	Percentage in Software Engineering	Percentage in Computer Networks	Percentage in Electronics Subjects	Percentage in Computer Architecture	Percenta
0	69	63	78	87	94	94	87	
1	78	62	73	60	71	70	73	
2	71	86	91	87	61	81	72	
3	76	87	60	84	89	73	62	
4	92	62	90	67	71	89	73	

5 rows × 39 columns

In [4]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20000 entries, 0 to 19999
Data columns (total 39 columns):
 #   Column                                    Non-Null Count  Dtype
---  ------                                    --------------  -----
 0   Acedamic percentage in Operating Systems  20000 non-null  int64
 1   percentage in Algorithms                  20000 non-null  int64
 2   Percentage in Programming Concepts        20000 non-null  int64
 3   Percentage in Software Engineering        20000 non-null  int64
 4   Percentage in Computer Networks           20000 non-null  int64
 5   Percentage in Electronics Subjects        20000 non-null  int64
 6   Percentage in Computer Architecture       20000 non-null  int64
 7   Percentage in Mathematics                 20000 non-null  int64
 8   Percentage in Communication skills        20000 non-null  int64
 9   Hours working per day                     20000 non-null  int64
 10  Logical quotient rating                   20000 non-null  int64
 11  hackathons                                20000 non-null  int64
 12  coding skills rating                      20000 non-null  int64
 13  public speaking points                    20000 non-null  int64
 14  can work long time before system?         20000 non-null  object
 15  self-learning capability?                 20000 non-null  object
 16  Extra-courses did                         20000 non-null  object
 17  certifications                            20000 non-null  object
 18  workshops                                 20000 non-null  object
 19  talenttests taken?                        20000 non-null  object
 20  olympiads                                 20000 non-null  object
 21  reading and writing skills                20000 non-null  object
 22  memory capability score                   20000 non-null  object
 23  Interested subjects                       20000 non-null  object
 24  interested career area                    20000 non-null  object
 25  Job/Higher Studies?                       20000 non-null  object
 26  Type of company want to settle in?        20000 non-null  object
 27  Taken inputs from seniors or elders       20000 non-null  object
 28  interested in games                       20000 non-null  object
 29  Interested Type of Books                  20000 non-null  object
 30  Salary Range Expected                     20000 non-null  object
 31  In a Realtionship?                        20000 non-null  object
 32  Gentle or Tuff behaviour?                 20000 non-null  object
 33  Management or Technical                   20000 non-null  object
 34  Salary/work                               20000 non-null  object
 35  hard/smart worker                         20000 non-null  object
 36  worked in teams ever?                     20000 non-null  object
 37  Introvert                                 20000 non-null  object
 38  Suggested Job Role                        20000 non-null  object
dtypes: int64(14), object(25)
memory usage: 6.0+ MB
```

Preprocessing

In [5]:

```
numericCols=df.select_dtypes("int64").columns
```

In [6]:

```
for col in numericCols:
    stats = df.describe()[col]
    df[col] = df[col].apply(lambda x: binRanges(x,stats))
```

In [7]:

```
jobs = pd.DataFrame(df['Suggested Job Role'])
df_ = df.drop('Suggested Job Role',axis=1)
x = pd.get_dummies(df_,columns=df_.select_dtypes('Object').columns)
jobs['new roles']=jobs['Suggested Job Role'].apply(lambda x: x.split(' ')[-1] )
jobs['new roles clubbed']=jobs['new roles'].apply(lambda x: club(x))
y =jobs['new roles clubbed']
# X_train, X_test, y_train, y_test = train_test_split(x, y, stratify=y,random_state=1

df_['Y']=y
```

PCA and LDA

For dimensionality reduction as a part of the preprocessing.

2-D PCA plot additionally helps to visualise the distribution with respect to the original classes.

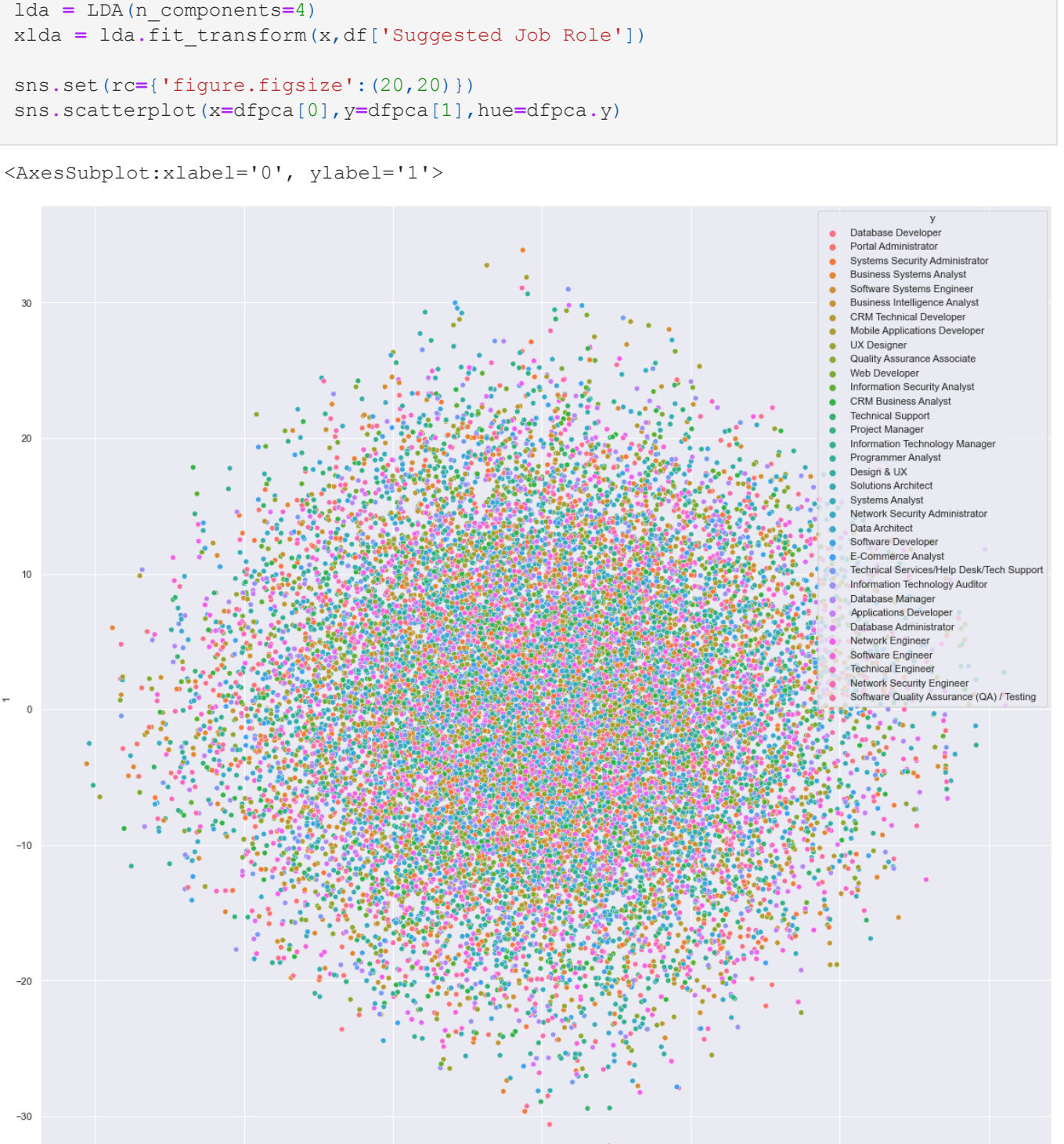
In [8]:

```
pca = PCA(n_components=2)
xpca = pca.fit_transform(x)
dfpca = pd.DataFrame(xpca)
dfpca['y']=df['Suggested Job Role']

lda = LDA(n_components=4)
xlda = lda.fit_transform(x,df['Suggested Job Role'])

sns.set(rc={'figure.figsize':(20,20)})
sns.scatterplot(x=dfpca[0],y=dfpca[1],hue=dfpca.y)
```

Out[8]: <AxesSubplot:xlabel='0', ylabel='1'>



Fitting on Grid

With Test sizes of 0.2,0.25,0.3,0.35,0.4 respectively

alpha values of .0000001,0.0001,0.001,0.01,0,1 respectively

In [10]:

```
ts= [0.2,0.25,0.3,0.35,0.4]
a=[0.0000001,0.0001,0.001,0.01,0,1]
```

In [11]:

```
grid(xpca,ts,a)
```

```
score with test size: 0.2 and alpha:1e-07:    0.369
confusion-matrix with test size: 0.2 and alpha:1e-07:
[[   9    3   13    4    8]
 [   3    1    4    2    3]
 [ 748  449 1456  447  776]
 [   1    0    3    2    1]
 [  14   11   25    9    8]]
score with test size: 0.2 and alpha:0.0001:    0.3645
confusion-matrix with test size: 0.2 and alpha:0.0001:
[[   3    8   22    2    5]
 [   5    2    6    1    2]
 [ 758  447 1440  453  775]
 [   0    1    0    1    2]
 [   9    6   33    7   12]]
score with test size: 0.2 and alpha:0.001:    0.37225
confusion-matrix with test size: 0.2 and alpha:0.001:
[[   3    2   12    0    2]
 [   4    0    2    4    4]
 [ 761  457 1472  456  776]
 [   0    1    1    1    1]
 [   7    4   14    3   13]]
score with test size: 0.2 and alpha:0.01:    0.374
confusion-matrix with test size: 0.2 and alpha:0.01:
[[   2    2   10    0    5]
 [   0    0    0    0    0]
 [ 770  460 1488  462  785]
 [   0    0    0    0    0]
 [   3    2    3    2    6]]
score with test size: 0.2 and alpha:0:    0.36625
confusion-matrix with test size: 0.2 and alpha:0:
[[  10    9   30    4    8]
 [   1    1    0    1    1]
 [ 746  443 1433  444  765]
 [   1    0    2    0    1]
 [  17   11   36   15   21]]
score with test size: 0.2 and alpha:1:    0.37525
confusion-matrix with test size: 0.2 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [ 775  464 1501  464  796]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.25 and alpha:1e-07:    0.3696
confusion-matrix with test size: 0.25 and alpha:1e-07:
[[   4    4   17    2    5]
 [  10    4    9    5    3]
 [ 934  558 1822  559  969]
 [   6    2    4    2    3]
 [  14   12   24   12   16]]
score with test size: 0.25 and alpha:0.0001:    0.3684
confusion-matrix with test size: 0.25 and alpha:0.0001:
[[   2    5   14    3    5]
 [   1    0    0    1    0]
 [ 954  566 1819  566  968]
 [   0    1    1    1    3]
 [  11    8   42    9   20]]
score with test size: 0.25 and alpha:0.001:    0.3758
confusion-matrix with test size: 0.25 and alpha:0.001:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [ 966  579 1874  578  991]
 [   0    0    0    0    0]
 [   2    1    2    2    5]]
score with test size: 0.25 and alpha:0.01:    0.374
confusion-matrix with test size: 0.25 and alpha:0.01:
[[   0    1    4    0    1]
 [   0    0    0    0    0]
 [ 963  577 1863  575  988]
 [   0    0    0    0    0]
 [   5    2    9    5    7]]
score with test size: 0.25 and alpha:0:    0.3748
confusion-matrix with test size: 0.25 and alpha:0:
[[   0    1    5    0    1]
 [   0    0    0    0    0]
 [ 964  577 1867  577  988]
 [   0    0    0    0    0]
 [   4    2    4    3    7]]
score with test size: 0.25 and alpha:1:    0.3752
confusion-matrix with test size: 0.25 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [ 968  580 1876  580  996]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.3 and alpha:1e-07:    0.36616666666666664
confusion-matrix with test size: 0.3 and alpha:1e-07:
[[   7    5   27    6    8]
 [   6    4    8    4    4]
 [1117  664 2152  670 1148]
 [   4    2    4    0    1]
 [  28   21   60   16   34]]
score with test size: 0.3 and alpha:0.0001:    0.366
confusion-matrix with test size: 0.3 and alpha:0.0001:
[[  22    8   38    9   20]
 [   1    2    4    2    0]
 [1106  666 2140  661 1141]
 [  10    6   25    9   11]
 [  23   14   44   15   23]]
score with test size: 0.3 and alpha:0.001:    0.37066666666666664
confusion-matrix with test size: 0.3 and alpha:0.001:
[[   7    5   18    5   10]
 [  10    4   16    6    6]
 [1127  672 2189  675 1155]
 [   4    1    4    1    1]
 [  14   14   24    9   23]]
score with test size: 0.3 and alpha:0.01:    0.37483333333333335
confusion-matrix with test size: 0.3 and alpha:0.01:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [1156  693 2240  692 1186]
 [   0    0    0    0    0]
 [   6    3   11    4    9]]
score with test size: 0.3 and alpha:0:    0.36433333333333334
confusion-matrix with test size: 0.3 and alpha:0:
[[  16   18   55   14   19]
 [   5    3   14    3    5]
 [1112  658 2140  661 1144]
 [   3    2    4    1    1]
 [  26   15   38   17   26]]
score with test size: 0.3 and alpha:1:    0.37516666666666665
confusion-matrix with test size: 0.3 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [1162  696 2251  696 1195]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.35 and alpha:1e-07:    0.3707142857142857
confusion-matrix with test size: 0.35 and alpha:1e-07:
[[  13    9   26    9   13]
 [   4    5    5    2    1]
 [1301  773 2514  776 1317]
 [   0    0    0    0    0]
 [  38   25   81   25   63]]
score with test size: 0.35 and alpha:0.0001:    0.367
confusion-matrix with test size: 0.35 and alpha:0.0001:
[[  29   27   67   16   29]
 [   6    3    9    2    7]
 [1292  758 2485  774 1304]
 [   2    4    8    4    6]
 [  27   20   57   16   48]]
score with test size: 0.35 and alpha:0.001:    0.36414285714285716
confusion-matrix with test size: 0.35 and alpha:0.001:
[[  24   16   57   13   17]
 [   5    2   11    3    5]
 [1290  766 2475  766 1324]
 [   1    2    1    0    0]
 [  36   26   82   30   48]]
score with test size: 0.35 and alpha:0.01:    0.37514285714285717
confusion-matrix with test size: 0.35 and alpha:0.01:
[[   1    1    3    3    2]
 [   0    0    0    0    0]
 [1353  811 2620  808 1387]
 [   0    0    0    0    0]
 [   2    0    3    1    5]]
score with test size: 0.35 and alpha:0:    0.365
confusion-matrix with test size: 0.35 and alpha:0:
[[  26   13   69   15   36]
 [   7    7   20    6    6]
 [1297  773 2494  780 1322]
 [   5    3    3    1    3]
 [  21   16   40   10   27]]
score with test size: 0.35 and alpha:1:    0.37514285714285717
confusion-matrix with test size: 0.35 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [1356  812 2626  812 1394]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.4 and alpha:1e-07:    0.367375
confusion-matrix with test size: 0.4 and alpha:1e-07:
[[  25   19   64   16   31]
 [   0    0    1    1    1]
 [1490  889 2877  892 1523]
 [   1    0    2    0    1]
 [  34   20   57   19   37]]
score with test size: 0.4 and alpha:0.0001:    0.36975
confusion-matrix with test size: 0.4 and alpha:0.0001:
[[  25   14   43   14   22]
 [   5    1   13    3    6]
 [1505  900 2909  895 1541]
 [   2    3    6    2    3]
 [  13   10   30   14   21]]
score with test size: 0.4 and alpha:0.001:    0.365
confusion-matrix with test size: 0.4 and alpha:0.001:
[[  20   16   56   13   19]
 [   2    2   10    0    2]
 [1482  884 2855  889 1526]
 [   6    2    6    2    5]
 [  40   24   74   24   41]]
score with test size: 0.4 and alpha:0.01:    0.371875
confusion-matrix with test size: 0.4 and alpha:0.01:
[[   7    6   22    6    5]
 [   0    0    3    0    3]
 [1532  914 2953  914 1570]
 [   0    0    0    0    0]
 [  11    8   23    8   15]]
score with test size: 0.4 and alpha:0:    0.369375
confusion-matrix with test size: 0.4 and alpha:0:
[[  22   11   32    9   16]
 [   8    4   13    3    8]
 [1498  890 2902  897 1539]
 [   2    2    9    1    4]
 [  20   21   45   18   26]]
score with test size: 0.4 and alpha:1:    0.375125
confusion-matrix with test size: 0.4 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [1550  928 3001  928 1593]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
```

In [12]:

```
grid(xlda,ts,a)
```

```
score with test size: 0.2 and alpha:1e-07:    0.28425
confusion-matrix with test size: 0.2 and alpha:1e-07:
[[132  84 255  78 125]
 [ 47  27  84  19  40]
 [428 239 820 267 447]
 [ 50  37 126  36  62]
 [118  77 216  64 122]]
score with test size: 0.2 and alpha:0.0001:    0.2945
confusion-matrix with test size: 0.2 and alpha:0.0001:
[[ 93  52 166  53  65]
 [ 55  34 116  38  61]
 [457 262 897 265 497]
 [ 43  33 111  28  47]
 [127  83 211  80 126]]
score with test size: 0.2 and alpha:0.001:    0.292
confusion-matrix with test size: 0.2 and alpha:0.001:
[[107  57 234  56 107]
 [ 49  33  73  30  43]
 [449 264 867 278 475]
 [ 29  24  68  20  30]
 [141  86 259  80 141]]
score with test size: 0.2 and alpha:0.01:    0.30175
confusion-matrix with test size: 0.2 and alpha:0.01:
[[ 97  50 208  51  69]
 [ 42  30  65  27  37]
 [498 294 926 308 516]
 [ 44  27  81  25  45]
 [ 94  63 221  53 129]]
score with test size: 0.2 and alpha:0:    0.27575
confusion-matrix with test size: 0.2 and alpha:0:
[[105  79 236  68 105]
 [ 55  30 117  37  65]
 [435 235 797 250 435]
 [ 56  40 125  36  56]
 [124  80 226  73 135]]
score with test size: 0.2 and alpha:1:    0.37525
confusion-matrix with test size: 0.2 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [ 775  464 1501  464  796]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.25 and alpha:1e-07:    0.2972
confusion-matrix with test size: 0.25 and alpha:1e-07:
[[ 108   75  264   63  127]
 [  66   43  112   38   71]
 [ 589  355 1136  354  587]
 [  54   27   98   37   49]
 [ 151   80  266   88  162]]
score with test size: 0.25 and alpha:0.0001:    0.2706
confusion-matrix with test size: 0.25 and alpha:0.0001:
[[126  99 290  96 131]
 [ 76  38 133  34  85]
 [505 265 907 288 483]
 [ 66  42 137  47  62]
 [195 136 409 115 235]]
score with test size: 0.25 and alpha:0.001:    0.2942
confusion-matrix with test size: 0.25 and alpha:0.001:
[[ 152   94  290   86  132]
 [  59   34  104   25   53]
 [ 579  336 1109  339  604]
 [  52   36  115   35   66]
 [ 126   80  258   95  141]]
score with test size: 0.25 and alpha:0.01:    0.297
confusion-matrix with test size: 0.25 and alpha:0.01:
[[ 104   76  236   67  107]
 [  50   31   97   27   51]
 [ 631  360 1182  377  642]
 [  57   33  107   30   58]
 [ 126   80  254   79  138]]
score with test size: 0.25 and alpha:0:    0.2978
confusion-matrix with test size: 0.25 and alpha:0:
[[ 145   92  298   87  135]
 [  40   37   77   29   41]
 [ 574  347 1113  342  602]
 [  66   27  126   36   60]
 [ 143   77  262   86  158]]
score with test size: 0.25 and alpha:1:    0.3752
confusion-matrix with test size: 0.25 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [ 968  580 1876  580  996]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.3 and alpha:1e-07:    0.264
confusion-matrix with test size: 0.3 and alpha:1e-07:
[[ 147  106  340   97  160]
 [  51   26   87   28   53]
 [ 621  376 1096  340  625]
 [ 124   56  219   73  115]
 [ 219  132  509  158  242]]
score with test size: 0.3 and alpha:0.0001:    0.265
confusion-matrix with test size: 0.3 and alpha:0.0001:
[[ 191  115  410  126  187]
 [  92   70  178   58   91]
 [ 594  343 1052  337  595]
 [ 106   53  209   56  101]
 [ 179  115  402  119  221]]
score with test size: 0.3 and alpha:0.001:    0.2925
confusion-matrix with test size: 0.3 and alpha:0.001:
[[ 146   90  310   91  125]
 [  69   52  147   43   95]
 [ 715  410 1360  428  747]
 [  90   52  162   50   81]
 [ 142   92  272   84  147]]
score with test size: 0.3 and alpha:0.01:    0.2855
confusion-matrix with test size: 0.3 and alpha:0.01:
[[ 186  131  395  111  184]
 [  48   32  103   26   48]
 [ 691  381 1268  422  711]
 [  76   48  134   46   71]
 [ 161  104  351   91  181]]
score with test size: 0.3 and alpha:0:    0.2755
confusion-matrix with test size: 0.3 and alpha:0:
[[ 174  110  402  130  187]
 [  90   73  191   50  100]
 [ 623  344 1148  367  590]
 [ 114   69  181   57  117]
 [ 161  100  329   92  201]]
score with test size: 0.3 and alpha:1:    0.37516666666666665
confusion-matrix with test size: 0.3 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [1162  696 2251  696 1195]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.35 and alpha:1e-07:    0.26571428571428574
confusion-matrix with test size: 0.35 and alpha:1e-07:
[[ 219  141  444  119  221]
 [  91   70  191   60  109]
 [ 667  386 1226  416  662]
 [  99   45  179   55  112]
 [ 280  170  586  162  290]]
score with test size: 0.35 and alpha:0.0001:    0.28585714285714287
confusion-matrix with test size: 0.35 and alpha:0.0001:
[[ 275  164  552  164  276]
 [  70   51  163   49   82]
 [ 763  430 1407  422  750]
 [  65   40  131   49   67]
 [ 183  127  373  128  219]]
score with test size: 0.35 and alpha:0.001:    0.29228571428571426
confusion-matrix with test size: 0.35 and alpha:0.001:
[[ 177  109  364  112  173]
 [  93   60  202   47  113]
 [ 809  489 1570  499  854]
 [  75   30  114   37   52]
 [ 202  124  376  117  202]]
score with test size: 0.35 and alpha:0.01:    0.30557142857142855
confusion-matrix with test size: 0.35 and alpha:0.01:
[[ 142   84  274   99  140]
 [  94   65  189   66  109]
 [ 885  521 1699  505  903]
 [  57   27  102   34   43]
 [ 178  115  362  108  199]]
score with test size: 0.35 and alpha:0:    0.2815714285714286
confusion-matrix with test size: 0.35 and alpha:0:
[[ 141   93  302  119  148]
 [  56   40  154   44   80]
 [ 819  462 1486  451  826]
 [ 135   79  245   69  105]
 [ 205  138  439  129  235]]
score with test size: 0.35 and alpha:1:    0.37514285714285717
confusion-matrix with test size: 0.35 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [1356  812 2626  812 1394]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.4 and alpha:1e-07:    0.2845
confusion-matrix with test size: 0.4 and alpha:1e-07:
[[ 229  134  455  145  218]
 [ 142   78  270   80  136]
 [ 876  509 1683  502  902]
 [ 122   68  184   67  118]
 [ 181  139  409  134  219]]
score with test size: 0.4 and alpha:0.0001:    0.294375
confusion-matrix with test size: 0.4 and alpha:0.0001:
[[ 179  125  391  119  193]
 [  84   52  142   60   92]
 [ 923  521 1765  518  897]
 [ 123   75  222   81  133]
 [ 241  155  481  150  278]]
score with test size: 0.4 and alpha:0.001:    0.28775
confusion-matrix with test size: 0.4 and alpha:0.001:
[[ 224  143  438  137  212]
 [ 116   57  220   54   96]
 [ 884  532 1718  515  918]
 [ 138   70  257   92  156]
 [ 188  126  368  130  211]]
score with test size: 0.4 and alpha:0.01:    0.302
confusion-matrix with test size: 0.4 and alpha:0.01:
[[ 141   76  285   72  136]
 [  72   34  123   44   68]
 [ 987  579 1847  589  975]
 [  96   52  164   57   77]
 [ 254  187  582  166  337]]
score with test size: 0.4 and alpha:0:    0.27575
confusion-matrix with test size: 0.4 and alpha:0:
[[ 230  130  461  142  211]
 [ 143   79  267   75  152]
 [ 838  483 1582  509  837]
 [ 158   96  298   89  167]
 [ 181  140  393  113  226]]
score with test size: 0.4 and alpha:1:    0.375125
confusion-matrix with test size: 0.4 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [1550  928 3001  928 1593]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
```

In [13]:

```
grid(x,ts,a)
```

```
score with test size: 0.2 and alpha:1e-07:    0.24125
confusion-matrix with test size: 0.2 and alpha:1e-07:
[[159  95 314  97 176]
 [ 85  49 168  62  77]
 [291 172 570 161 336]
 [ 81  40 119  40  60]
 [159 108 330 104 147]]
score with test size: 0.2 and alpha:0.0001:    0.2485
confusion-matrix with test size: 0.2 and alpha:0.0001:
[[146  93 285  91 164]
 [ 94  44 154  52  79]
 [333 187 600 195 309]
 [ 66  50 171  41  81]
 [136  90 291  85 163]]
score with test size: 0.2 and alpha:0.001:    0.2505
confusion-matrix with test size: 0.2 and alpha:0.001:
[[143 110 289  91 151]
 [ 69  42 142  47  91]
 [320 191 618 175 317]
 [ 86  43 160  44  82]
 [157  78 292 107 155]]
score with test size: 0.2 and alpha:0.01:    0.25925
confusion-matrix with test size: 0.2 and alpha:0.01:
[[128  73 241  89 139]
 [ 87  45 168  32  66]
 [357 209 665 209 369]
 [ 89  37 148  43  66]
 [114 100 279  91 156]]
score with test size: 0.2 and alpha:0:    0.25675
confusion-matrix with test size: 0.2 and alpha:0:
[[150  84 268  81 148]
 [ 86  54 169  61 105]
 [321 212 638 200 321]
 [103  40 178  50  87]
 [115  74 248  72 135]]
score with test size: 0.2 and alpha:1:    0.37525
confusion-matrix with test size: 0.2 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [ 775  464 1501  464  796]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.25 and alpha:1e-07:    0.245
confusion-matrix with test size: 0.25 and alpha:1e-07:
[[170  91 295  93 170]
 [118  63 238  53  99]
 [391 217 698 212 390]
 [120  63 236  73 116]
 [169 146 409 149 221]]
score with test size: 0.25 and alpha:0.0001:    0.255
confusion-matrix with test size: 0.25 and alpha:0.0001:
[[206 125 400 133 212]
 [ 73  43 176  51  90]
 [395 248 759 215 404]
 [ 94  51 186  56  79]
 [200 113 355 125 211]]
score with test size: 0.25 and alpha:0.001:    0.2332
confusion-matrix with test size: 0.25 and alpha:0.001:
[[197 117 392 116 233]
 [102  72 219  58  98]
 [363 208 691 248 407]
 [123  70 209  56 108]
 [183 113 365 102 150]]
score with test size: 0.25 and alpha:0.01:    0.2592
confusion-matrix with test size: 0.25 and alpha:0.01:
[[179  97 350 102 201]
 [108  63 178  60 106]
 [432 255 826 256 412]
 [ 93  55 183  48  97]
 [156 110 339 114 180]]
score with test size: 0.25 and alpha:0:    0.2496
confusion-matrix with test size: 0.25 and alpha:0:
[[185 108 360 104 183]
 [100  68 187  65 109]
 [396 229 770 243 413]
 [ 99  51 184  48 114]
 [188 124 375 120 177]]
score with test size: 0.25 and alpha:1:    0.3752
confusion-matrix with test size: 0.25 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [ 968  580 1876  580  996]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.3 and alpha:1e-07:    0.24883333333333332
confusion-matrix with test size: 0.3 and alpha:1e-07:
[[224 158 479 147 287]
 [110  69 200  65 108]
 [484 256 910 278 471]
 [126  55 218  78 117]
 [218 158 444 128 212]]
score with test size: 0.3 and alpha:0.0001:    0.24183333333333334
confusion-matrix with test size: 0.3 and alpha:0.0001:
[[231 145 469 126 229]
 [135  84 266  69 131]
 [424 228 793 263 449]
 [140  92 263  77 120]
 [232 147 460 161 266]]
score with test size: 0.3 and alpha:0.001:    0.25483333333333336
confusion-matrix with test size: 0.3 and alpha:0.001:
[[272 132 432 134 246]
 [133  65 257  74 127]
 [412 274 870 275 447]
 [116  68 234  61 114]
 [229 157 458 152 261]]
score with test size: 0.3 and alpha:0.01:    0.25683333333333336
confusion-matrix with test size: 0.3 and alpha:0.01:
[[251 123 418 134 236]
 [ 99  58 254  63 121]
 [498 312 950 302 505]
 [120  70 228  70 121]
 [194 133 401 127 212]]
score with test size: 0.3 and alpha:0:    0.24933333333333332
confusion-matrix with test size: 0.3 and alpha:0:
[[222 147 470 111 239]
 [150  82 238  74 132]
 [443 253 856 280 445]
 [134  77 247  85 128]
 [213 137 440 146 251]]
score with test size: 0.3 and alpha:1:    0.37516666666666665
confusion-matrix with test size: 0.3 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [1162  696 2251  696 1195]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.35 and alpha:1e-07:    0.248
confusion-matrix with test size: 0.35 and alpha:1e-07:
[[ 254  167  559  172  292]
 [ 134  103  285   83  134]
 [ 522  306 1040  296  589]
 [ 160   81  265   96  136]
 [ 286  155  477  165  243]]
score with test size: 0.35 and alpha:0.0001:    0.25
confusion-matrix with test size: 0.35 and alpha:0.0001:
[[ 246  168  534  132  264]
 [ 137   94  282  101  145]
 [ 559  331 1048  341  562]
 [ 160   81  280   80  141]
 [ 254  138  482  158  282]]
score with test size: 0.35 and alpha:0.001:    0.24528571428571427
confusion-matrix with test size: 0.35 and alpha:0.001:
[[ 243  141  524  150  286]
 [ 111   74  247   64  122]
 [ 557  333 1064  336  601]
 [ 158   74  268   95  144]
 [ 287  190  523  167  241]]
score with test size: 0.35 and alpha:0.01:    0.257
confusion-matrix with test size: 0.35 and alpha:0.01:
[[ 238  154  466  125  270]
 [ 142   73  277   86  120]
 [ 612  369 1186  392  637]
 [ 172   86  273   73  138]
 [ 192  130  424  136  229]]
score with test size: 0.35 and alpha:0:    0.24528571428571427
confusion-matrix with test size: 0.35 and alpha:0:
[[ 251  171  550  164  278]
 [ 167   83  271   80  138]
 [ 565  292 1052  336  588]
 [ 123   95  253   78  137]
 [ 250  171  500  154  253]]
score with test size: 0.35 and alpha:1:    0.37514285714285717
confusion-matrix with test size: 0.35 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [1356  812 2626  812 1394]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
score with test size: 0.4 and alpha:1e-07:    0.246
confusion-matrix with test size: 0.4 and alpha:1e-07:
[[ 293  197  568  184  292]
 [ 162  101  323   94  167]
 [ 615  369 1179  383  658]
 [ 185   82  369   88  169]
 [ 295  179  562  179  307]]
score with test size: 0.4 and alpha:0.0001:    0.24475
confusion-matrix with test size: 0.4 and alpha:0.0001:
[[ 289  171  613  167  326]
 [ 164   93  310  103  143]
 [ 584  356 1168  355  646]
 [ 182  106  322   92  162]
 [ 331  202  588  211  316]]
score with test size: 0.4 and alpha:0.001:    0.2535
confusion-matrix with test size: 0.4 and alpha:0.001:
[[ 289  194  552  172  328]
 [ 160  105  336  118  160]
 [ 631  362 1235  384  629]
 [ 191   95  320   91  168]
 [ 279  172  558  163  308]]
score with test size: 0.4 and alpha:0.01:    0.24225
confusion-matrix with test size: 0.4 and alpha:0.01:
[[ 263  189  556  172  304]
 [ 177  123  360  108  221]
 [ 596  323 1112  358  594]
 [ 165   85  353  113  147]
 [ 349  208  620  177  327]]
score with test size: 0.4 and alpha:0:    0.247625
confusion-matrix with test size: 0.4 and alpha:0:
[[ 289  176  639  184  339]
 [ 185  114  312   94  155]
 [ 585  372 1171  363  604]
 [ 175   97  294   89  177]
 [ 316  169  585  198  318]]
score with test size: 0.4 and alpha:1:    0.375125
confusion-matrix with test size: 0.4 and alpha:1:
[[   0    0    0    0    0]
 [   0    0    0    0    0]
 [1550  928 3001  928 1593]
 [   0    0    0    0    0]
 [   0    0    0    0    0]]
```

confusion matrix visualisation of a reasonable outcome

In [17]:

```
# score with test size: 0.4 and alpha:0:    0.27575
# confusion-matrix with test size: 0.4 and alpha:0:


X_train, X_test, y_train, y_test = train_test_split(xlda, y, stratify=y,random_state=

clf = MLPClassifier(hidden_layer_sizes=(300,200,100),alpha=0 ,max_iter=200).fit(X_tra

cm = confusion_matrix(clf.predict(X_test),y_test)
```

```
<Figure size 720x720 with 0 Axes>
```

In [20]:

```
sns.set(rc={'figure.figsize':(10,10)})
sns.heatmap(cm)
```

Out[20]: <AxesSubplot:>

Conclusion

The data shows a reasonable class distribution in the confusion matrix for lower values of accuracies.

As we reach higher accuracy the class distribution tends to get imbalanced.

For most train test ratio, smaller value of the alpha value tends to give a more balanced confusion matrix