Rogoz Ana-Cristina, 232

DOCUMENTATIE PROIECT
Inteligenta Artificiala

Arhitectura retelei antrenate

Pentru acest proiect, am utilizat doua retele.

Prima retea a folosit tot setul de antrenare dat, a avut 2 straturi ascunse, unul cu 64 de perceptroni, unul cu 16 perceptroni si un strat de iesire cu 10 perceptroni, iar functiile de transfer de pe aceste straturi au fost logsig pentru straturile ascunse si softmax pentru stratul de iesire.

Cea de-a doua retea a folosit un numar aproximativ egal de exemple din fiecare clasa a setului de antrenare (in jur de 750 exemple din fiecare clasa de la 0-9) si a pastrat arhitectura primei retele. Pe aceasta am folosit-o exclusiv pentru imbunatatirea outputului dat de prima retea, astfel daca prima retea considera ca un exemplu face parte dintr-o clasa ce nu este 9, utilizam raspunsul dat de aceasta retea echilibrata.

Algoritmul de antrenare

Atat pentru prima retea, cat si pentru cea de-a doua retea am pastrat functia de antrenare a patternnetului si anume trainscg, antrenare pe baza gradientului conjugat scalat, am pastrat si preprocesarile aferente pentru input si output, dar am inlocuit valorile de NaN cu media pe linie.

Pentru prima retea am setat rata de invatare ca fiind 0.007, numarul maxim de iteratii pentru care sa nu existe nicio imbunatatire pe multimea de validare l-am pus ca fiind 30 (trainParam.max_fail), am pastrat raportul pentru antrenare/validare/testare, iar modul in care au fost impartite exemplele a fost random. Pentru cea de-a doua retea, cea echilibrata am setat rata de invatare ca fiind 0.008, am folosit toate exemplele pentru antrenare si am

oprit antrenarea in momentul in care performanta a ajuns la 0.05, intrucat in urma mai multor incercari, a parut sa nu faca overfitting daca o opresc in jurul momentului aceluia.

Performanta de clasificare medie

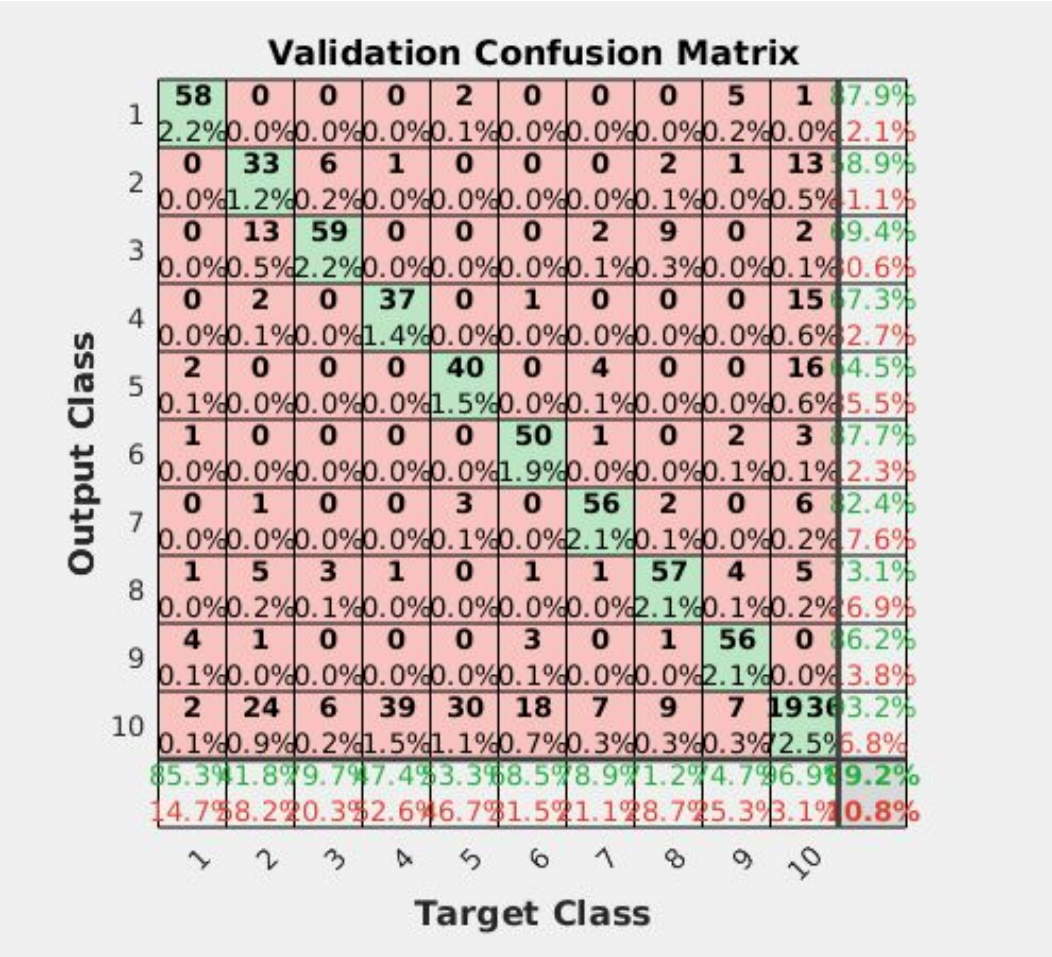
Am calculat performanta medie de clasificare a primei retele generand o permutare a indicilor multimii de antrenare, pe care i-am impartit in 10 submultimi disjuncte. Astfel, am obtinut urmatoarele rezultate:

Fold	Performanta
1	0.026496
2	0.026662
3	0.026744
4	0.026716
5	0.026763
6	0.026541
7	0.026461
8	0.026895
9	0.026734
10	0.02678

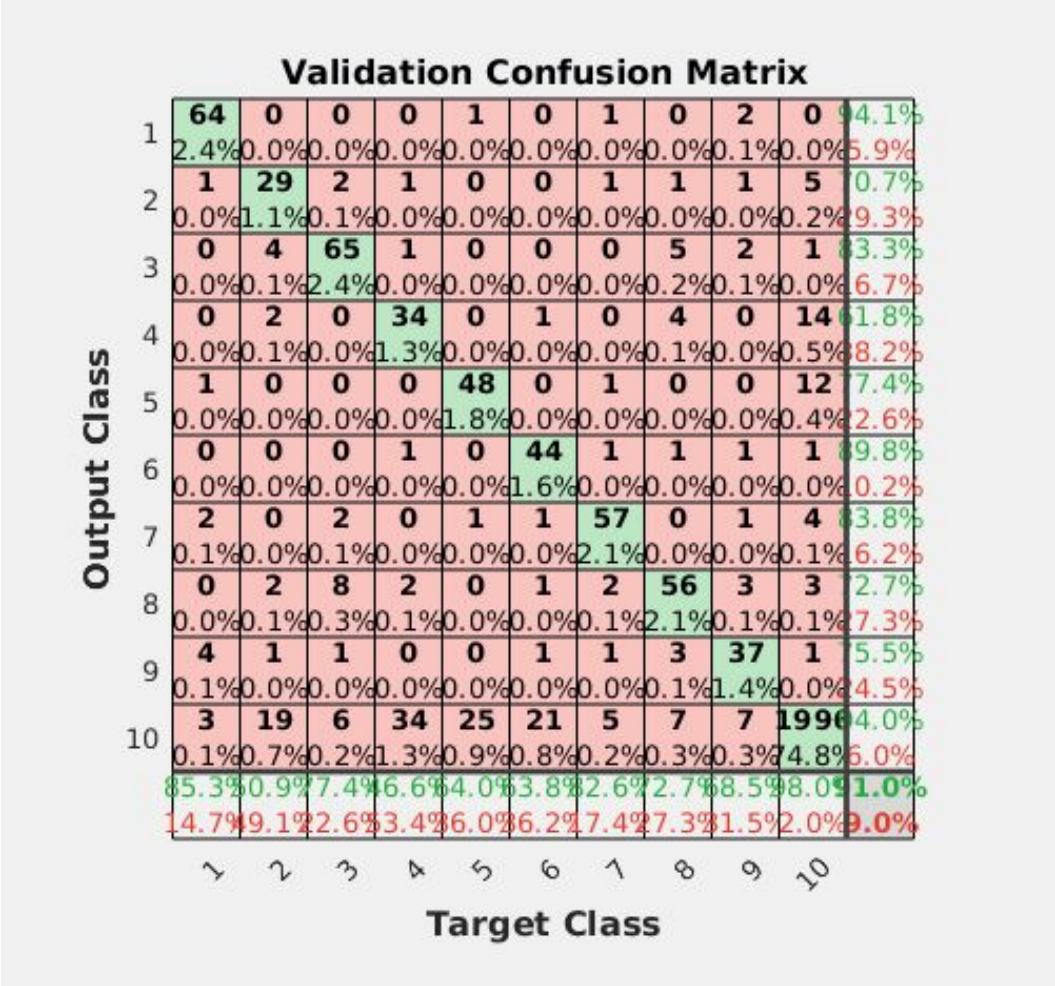
Asadar, performanta medie de clasficare este 0.0266792.

Matricea de confuzie pentru cele 10 seturi de validare:

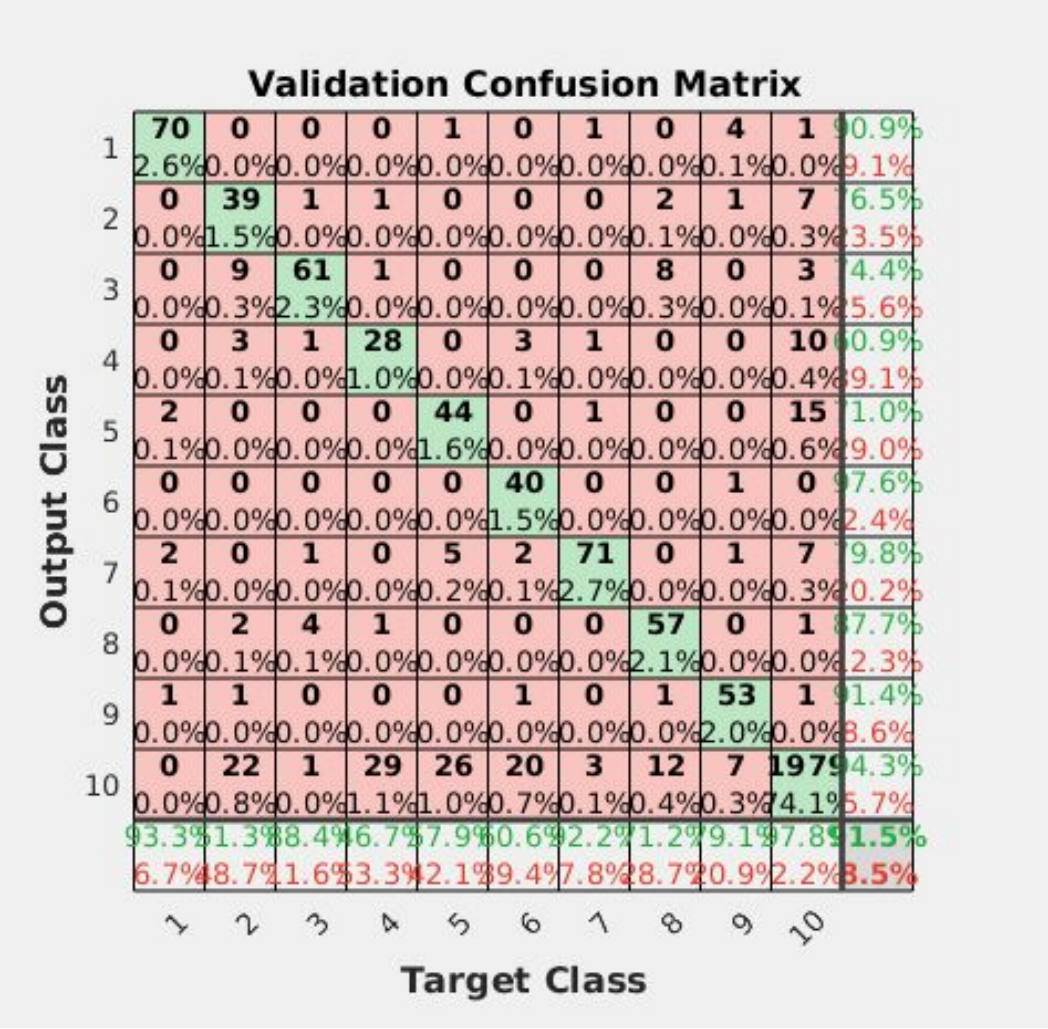
Fold 1



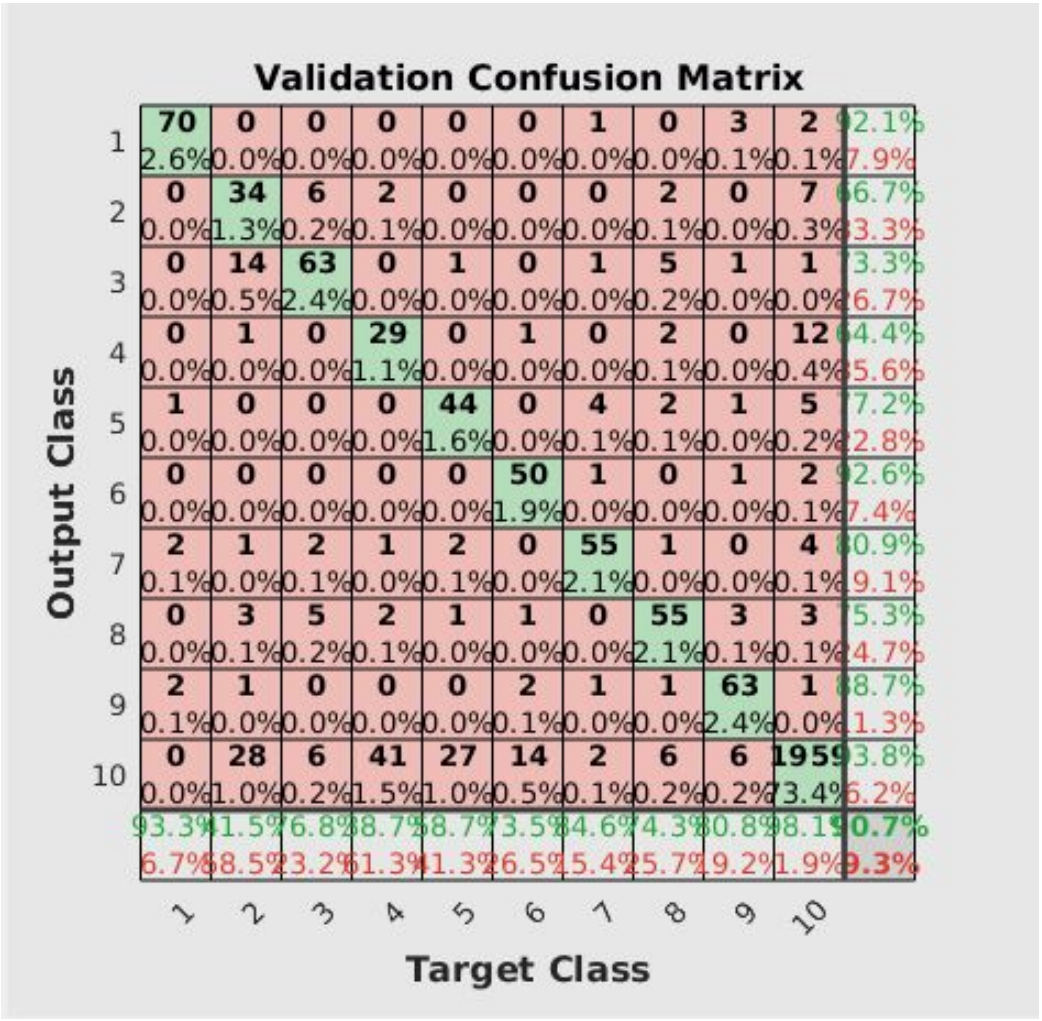
Fold 2


Validation Confusion Matrix
Output Class
Target Class
1 2 3 4 5 6 7 8 9 10
64 0 0 0 1 0 1 0 2 0 94.1% 5.9%
2.4% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.1% 0.0%
1 29 2 1 0 0 1 1 1 5 70.7% 29.3%
0.0% 1.1% 0.1% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.2%
0 4 65 1 0 0 0 5 2 1 83.3% 16.7%
0.0% 0.1% 2.4% 0.0% 0.0% 0.0% 0.0% 0.2% 0.1% 0.0%
0 2 0 34 0 1 0 4 0 14 61.8% 38.2%
0.0% 0.1% 0.0% 1.3% 0.0% 0.0% 0.0% 0.1% 0.0% 0.5%
1 0 0 0 48 0 1 0 0 12 77.4% 22.6%
0.0% 0.0% 0.0% 0.0% 1.8% 0.0% 0.0% 0.0% 0.0% 0.4%
0 0 0 1 0 44 1 1 1 1 89.8% 10.2%
0.0% 0.0% 0.0% 0.0% 0.0% 1.6% 0.0% 0.0% 0.0% 0.0%
2 0 2 0 1 1 57 0 1 4 83.8% 16.2%
0.1% 0.0% 0.1% 0.0% 0.0% 0.0% 2.1% 0.0% 0.0% 0.1%
0 2 8 2 0 1 2 56 3 3 72.7% 27.3%
0.0% 0.1% 0.3% 0.1% 0.0% 0.0% 0.1% 2.1% 0.1% 0.1%
4 1 1 0 0 1 1 3 37 1 75.5% 24.5%
0.1% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.1% 1.4% 0.0%
3 19 6 34 25 21 5 7 7 1996 94.0% 6.0%
0.1% 0.7% 0.2% 1.3% 0.9% 0.8% 0.2% 0.3% 0.3% 74.8%
85.3% 50.9% 77.4% 46.6% 64.0% 53.8% 62.6% 72.7% 68.5% 98.0% 91.0%
14.7% 49.1% 22.6% 53.4% 36.0% 46.2% 37.4% 27.3% 31.5% 2.0% 9.0%

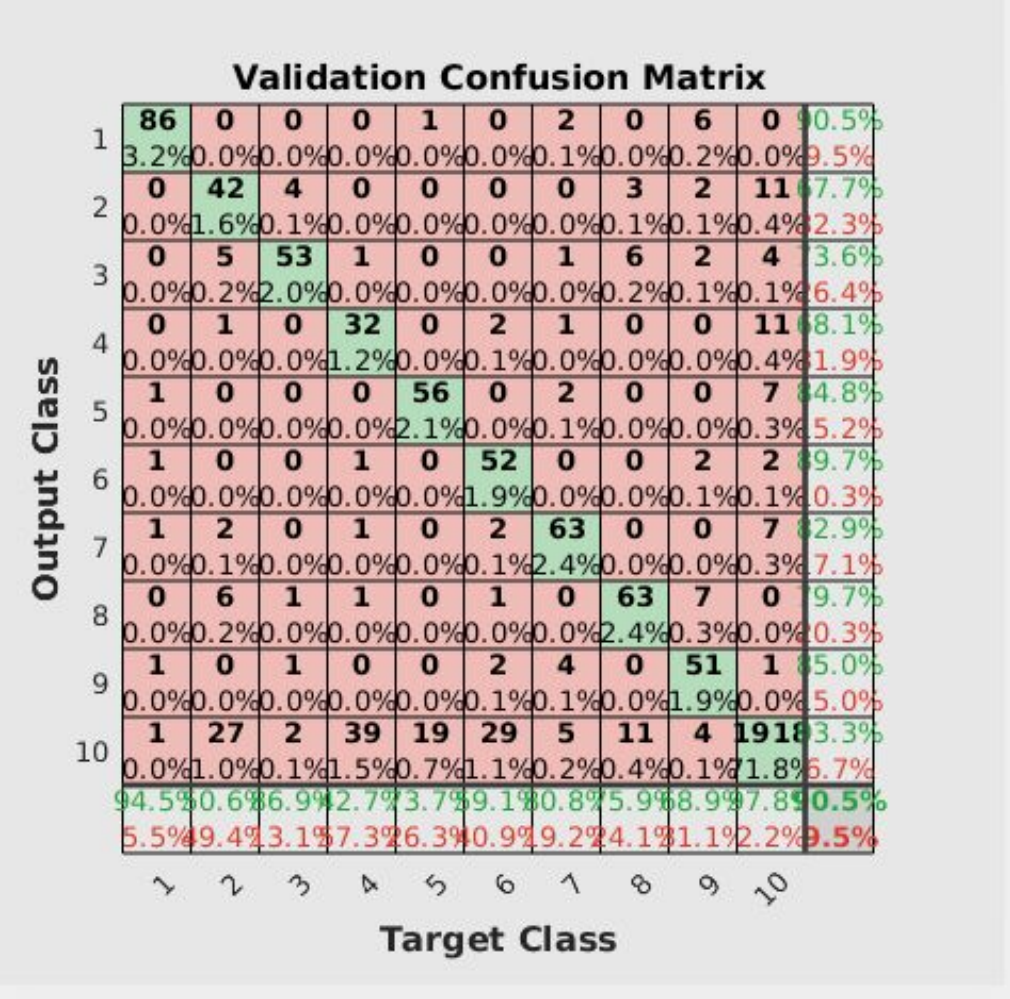
Fold 3


Validation Confusion Matrix
Output Class
Target Class
1 2 3 4 5 6 7 8 9 10
70 0 0 0 1 0 1 0 4 1 90.9% 9.1%
0 39 1 1 0 0 0 2 1 7 76.5% 23.5%
0 9 61 1 0 0 0 8 0 3 74.4% 25.6%
0 3 1 28 0 3 1 0 0 10 60.9% 39.1%
2 0 0 0 44 0 1 0 0 15 71.0% 29.0%
0 0 0 0 0 40 0 0 1 0 97.6% 2.4%
2 0 1 0 5 2 71 0 1 7 79.8% 20.2%
0 2 4 1 0 0 0 57 0 1 87.7% 12.3%
1 1 0 0 0 1 0 1 53 1 91.4% 8.6%
0 22 1 29 26 20 3 12 7 1979 94.3% 5.7%
93.3% 71.3% 88.4% 66.7% 57.9% 60.6% 92.2% 71.2% 79.1% 97.8% 91.5%
6.7% 28.7% 11.6% 33.3% 42.1% 39.4% 7.8% 28.7% 20.9% 2.2% 8.5%

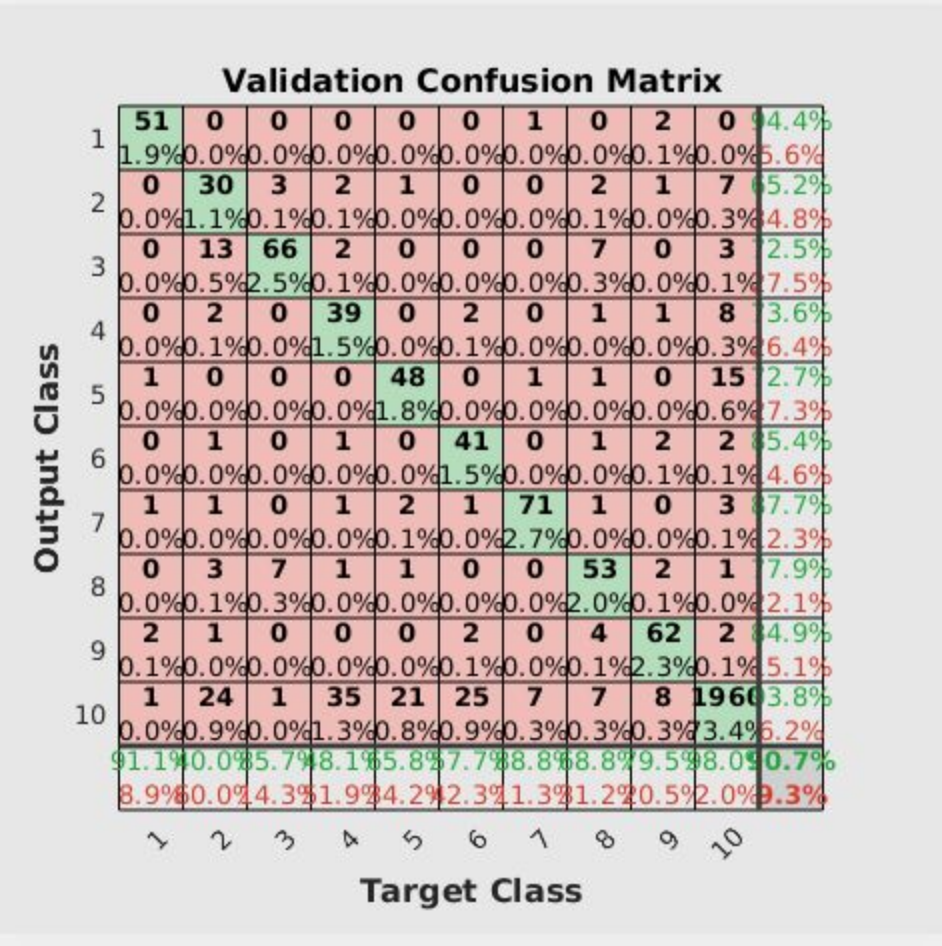
Fold 4


Validation Confusion Matrix
Output Class
Target Class
1 2 3 4 5 6 7 8 9 10
70 0 0 0 0 0 1 0 3 2 92.1%
2.6% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.1% 0.1% 7.9%
0 34 6 2 0 0 0 2 0 7 66.7%
0.0% 1.3% 0.2% 0.1% 0.0% 0.0% 0.0% 0.1% 0.0% 0.3% 33.3%
0 14 63 0 1 0 1 5 1 1 73.3%
0.0% 0.5% 2.4% 0.0% 0.0% 0.0% 0.0% 0.2% 0.0% 0.0% 26.7%
0 1 0 29 0 1 0 2 0 12 64.4%
0.0% 0.0% 0.0% 1.1% 0.0% 0.0% 0.0% 0.1% 0.0% 0.4% 35.6%
1 0 0 0 44 0 4 2 1 5 77.2%
0.0% 0.0% 0.0% 0.0% 1.6% 0.0% 0.1% 0.1% 0.0% 0.2% 22.8%
0 0 0 0 0 50 1 0 1 2 92.6%
0.0% 0.0% 0.0% 0.0% 0.0% 1.9% 0.0% 0.0% 0.0% 0.1% 7.4%
2 1 2 1 2 0 55 1 0 4 80.9%
0.1% 0.0% 0.1% 0.0% 0.1% 0.0% 2.1% 0.0% 0.0% 0.1% 19.1%
0 3 5 2 1 1 0 55 3 3 75.3%
0.0% 0.1% 0.2% 0.1% 0.0% 0.0% 0.0% 2.1% 0.1% 0.1% 24.7%
2 1 0 0 0 2 1 1 63 1 88.7%
0.1% 0.0% 0.0% 0.0% 0.0% 0.1% 0.0% 0.0% 2.4% 0.0% 11.3%
0 28 6 41 27 14 2 6 6 1959 93.8%
0.0% 1.0% 0.2% 1.5% 1.0% 0.5% 0.1% 0.2% 0.2% 73.4% 6.2%
93.3% 41.5% 76.8% 38.7% 38.7% 43.5% 84.6% 84.3% 80.8% 98.1% 90.7%
6.7% 58.5% 23.2% 61.3% 61.3% 56.5% 15.4% 15.7% 19.2% 1.9% 9.3%

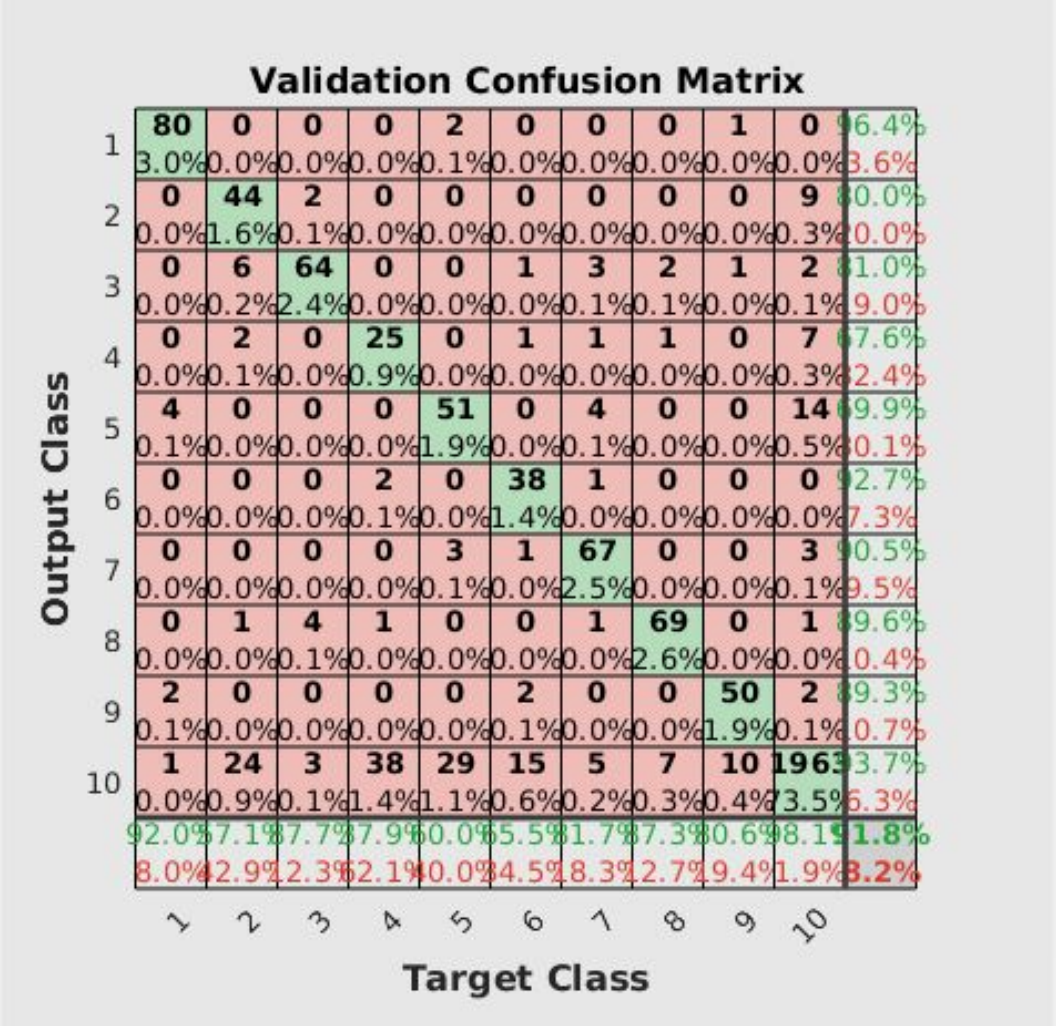
Fold 5


Validation Confusion Matrix
Output Class
Target Class
1 2 3 4 5 6 7 8 9 10
86 0 0 0 1 0 2 0 6 0 90.5%
3.2% 0.0% 0.0% 0.0% 0.0% 0.0% 0.1% 0.0% 0.2% 0.0% 9.5%
0 42 4 0 0 0 0 3 2 11 67.7%
0.0% 1.6% 0.1% 0.0% 0.0% 0.0% 0.0% 0.1% 0.1% 0.4% 32.3%
0 5 53 1 0 0 1 6 2 4 73.6%
0.0% 0.2% 2.0% 0.0% 0.0% 0.0% 0.0% 0.2% 0.1% 0.1% 26.4%
0 1 0 32 0 2 1 0 0 11 68.1%
0.0% 0.0% 0.0% 1.2% 0.0% 0.1% 0.0% 0.0% 0.0% 0.4% 31.9%
1 0 0 0 56 0 2 0 0 7 84.8%
0.0% 0.0% 0.0% 0.0% 2.1% 0.0% 0.1% 0.0% 0.0% 0.3% 15.2%
1 0 0 1 0 52 0 0 2 2 89.7%
0.0% 0.0% 0.0% 0.0% 0.0% 1.9% 0.0% 0.0% 0.1% 0.1% 10.3%
1 2 0 1 0 2 63 0 0 7 82.9%
0.0% 0.1% 0.0% 0.0% 0.0% 0.1% 2.4% 0.0% 0.0% 0.3% 17.1%
0 6 1 1 0 1 0 63 7 0 79.7%
0.0% 0.2% 0.0% 0.0% 0.0% 0.0% 0.0% 2.4% 0.3% 0.0% 20.3%
1 0 1 0 0 2 4 0 51 1 85.0%
0.0% 0.0% 0.0% 0.0% 0.0% 0.1% 0.1% 0.0% 1.9% 0.0% 15.0%
1 27 2 39 19 29 5 11 4 1918 3.3%
0.0% 1.0% 0.1% 1.5% 0.7% 1.1% 0.2% 0.4% 0.1% 71.8% 96.7%
94.5% 50.6% 86.9% 42.7% 73.7% 59.1% 80.8% 75.9% 68.9% 97.8% 90.5%
5.5% 49.4% 13.1% 57.3% 26.3% 40.9% 19.2% 24.1% 31.1% 2.2% 9.5%

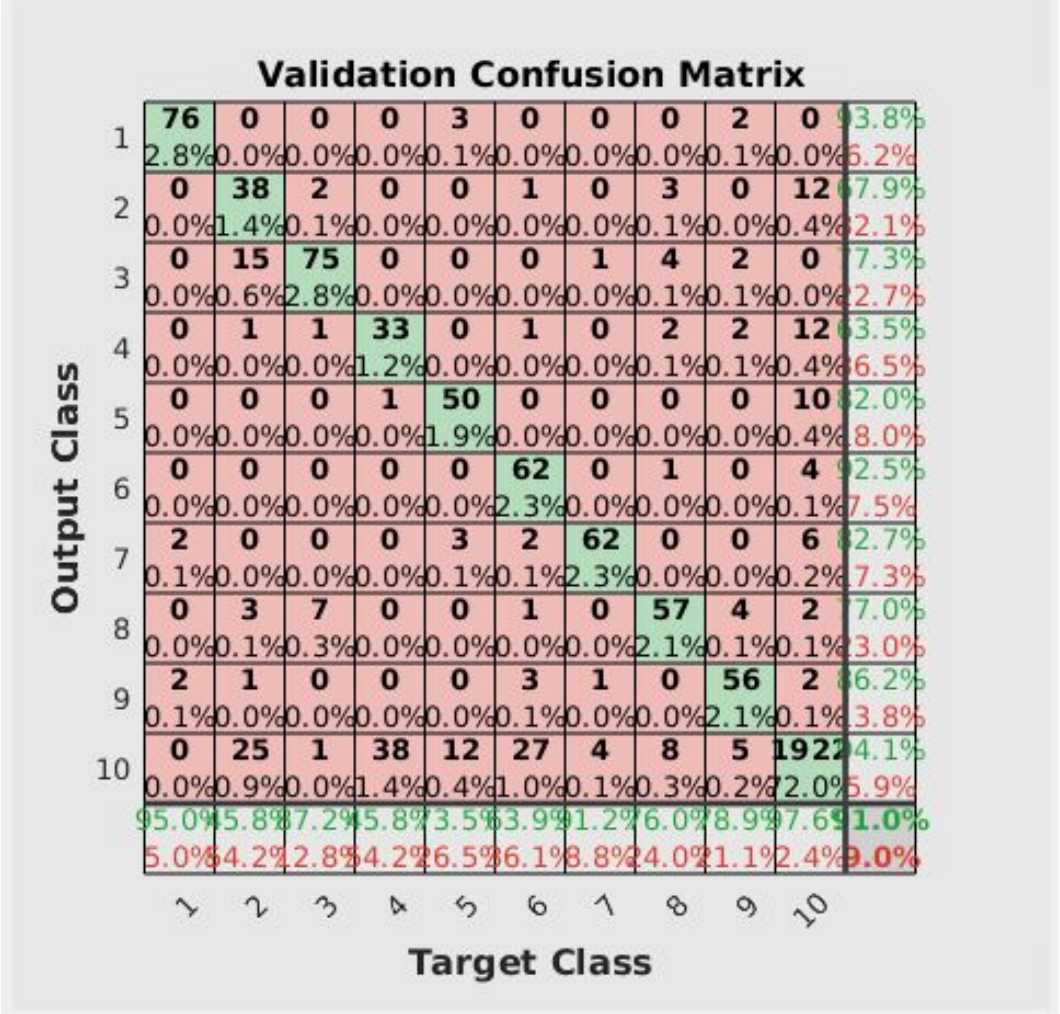
Fold 6


Validation Confusion Matrix
Output Class
Target Class
1 2 3 4 5 6 7 8 9 10
51 0 0 0 0 0 1 0 2 0 94.4% 5.6%
1.9% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.1% 0.0%
0 30 3 2 1 0 0 2 1 7 65.2% 34.8%
0.0% 1.1% 0.1% 0.1% 0.0% 0.0% 0.0% 0.1% 0.0% 0.3%
0 13 66 2 0 0 0 7 0 3 72.5% 27.5%
0.0% 0.5% 2.5% 0.1% 0.0% 0.0% 0.0% 0.3% 0.0% 0.1%
0 2 0 39 0 2 0 1 1 8 73.6% 26.4%
0.0% 0.1% 0.0% 1.5% 0.0% 0.1% 0.0% 0.0% 0.0% 0.3%
1 0 0 0 48 0 1 1 0 15 72.7% 27.3%
0.0% 0.0% 0.0% 0.0% 1.8% 0.0% 0.0% 0.0% 0.0% 0.6%
0 1 0 1 0 41 0 1 2 2 85.4% 14.6%
0.0% 0.0% 0.0% 0.0% 0.0% 1.5% 0.0% 0.0% 0.1% 0.1%
1 1 0 1 2 1 71 1 0 3 87.7% 12.3%
0.0% 0.0% 0.0% 0.0% 0.1% 0.0% 2.7% 0.0% 0.0% 0.1%
0 3 7 1 1 0 0 53 2 1 77.9% 22.1%
0.0% 0.1% 0.3% 0.0% 0.0% 0.0% 0.0% 2.0% 0.1% 0.0%
2 1 0 0 0 2 0 4 62 2 84.9% 15.1%
0.1% 0.0% 0.0% 0.0% 0.0% 0.1% 0.0% 0.1% 2.3% 0.1%
1 24 1 35 21 25 7 7 8 1960 93.8% 6.2%
0.0% 0.9% 0.0% 1.3% 0.8% 0.9% 0.3% 0.3% 0.3% 73.4%
91.1% 0.0% 85.7% 48.1% 65.8% 57.7% 88.8% 68.8% 79.5% 98.0% 90.7%
8.9% 60.0% 14.3% 51.9% 34.2% 42.3% 11.3% 31.2% 20.5% 2.0% 9.3%

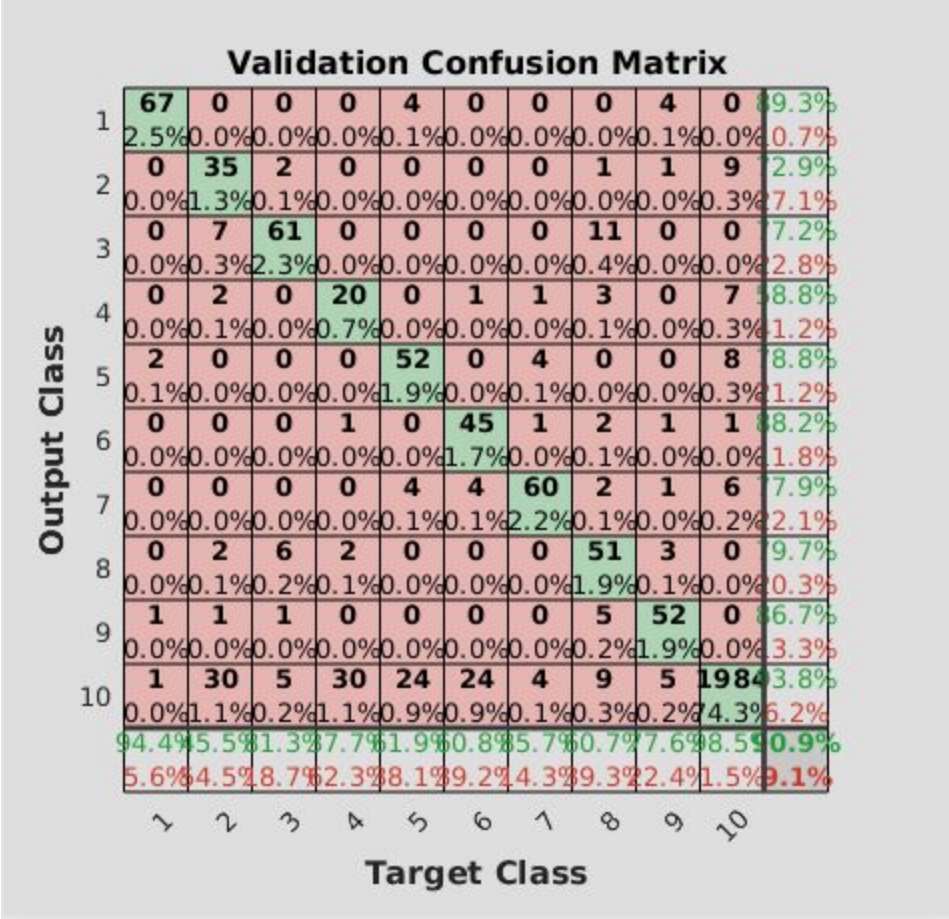
Fold 7


Validation Confusion Matrix
Output Class
Target Class
1 2 3 4 5 6 7 8 9 10
80 0 0 0 2 0 0 0 1 0 96.4%
3.0% 0.0% 0.0% 0.0% 0.1% 0.0% 0.0% 0.0% 0.0% 0.0% 3.6%
0 44 2 0 0 0 0 0 0 9 80.0%
0.0% 1.6% 0.1% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.3% 20.0%
0 6 64 0 0 1 3 2 1 2 81.0%
0.0% 0.2% 2.4% 0.0% 0.0% 0.0% 0.1% 0.1% 0.0% 0.1% 19.0%
0 2 0 25 0 1 1 1 0 7 67.6%
0.0% 0.1% 0.0% 0.9% 0.0% 0.0% 0.0% 0.0% 0.0% 0.3% 32.4%
4 0 0 0 51 0 4 0 0 14 69.9%
0.1% 0.0% 0.0% 0.0% 1.9% 0.0% 0.1% 0.0% 0.0% 0.5% 30.1%
0 0 0 2 0 38 1 0 0 0 92.7%
0.0% 0.0% 0.0% 0.1% 0.0% 1.4% 0.0% 0.0% 0.0% 0.0% 7.3%
0 0 0 0 3 1 67 0 0 3 90.5%
0.0% 0.0% 0.0% 0.0% 0.1% 0.0% 2.5% 0.0% 0.0% 0.1% 9.5%
0 1 4 1 0 0 1 69 0 1 89.6%
0.0% 0.0% 0.1% 0.0% 0.0% 0.0% 0.0% 2.6% 0.0% 0.0% 10.4%
2 0 0 0 0 2 0 0 50 2 89.3%
0.1% 0.0% 0.0% 0.0% 0.0% 0.1% 0.0% 0.0% 1.9% 0.1% 10.7%
1 24 3 38 29 15 5 7 10 1963 93.7%
0.0% 0.9% 0.1% 1.4% 1.1% 0.6% 0.2% 0.3% 0.4% 73.5% 6.3%
92.0% 57.1% 87.7% 37.9% 60.0% 65.5% 81.7% 87.3% 80.6% 98.1% 91.8%
8.0% 42.9% 12.3% 62.1% 40.0% 34.5% 18.3% 12.7% 19.4% 1.9% 8.2%

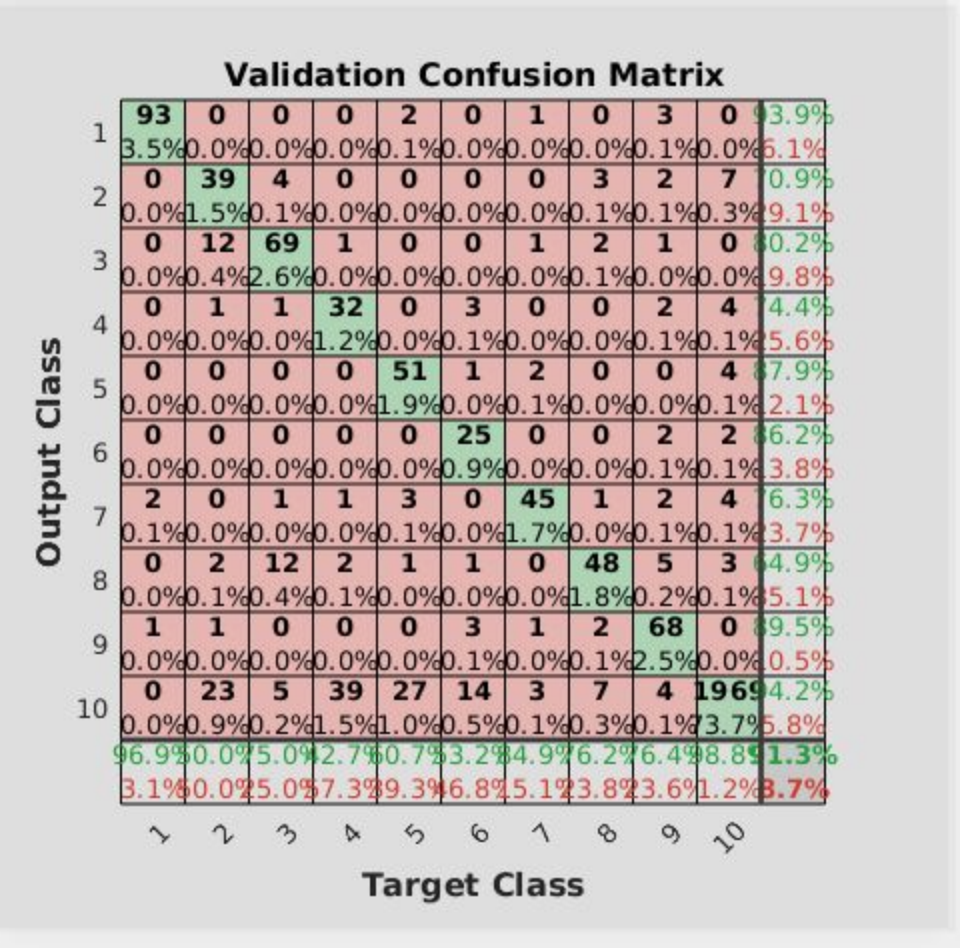
Fold 8


Validation Confusion Matrix
Output Class
Target Class
1 2 3 4 5 6 7 8 9 10
76 0 0 0 3 0 0 0 2 0 93.8%
2.8% 0.0% 0.0% 0.0% 0.1% 0.0% 0.0% 0.0% 0.1% 0.0% 6.2%
0 38 2 0 0 1 0 3 0 12 67.9%
0.0% 1.4% 0.1% 0.0% 0.0% 0.0% 0.0% 0.1% 0.0% 0.4% 32.1%
0 15 75 0 0 0 1 4 2 0 77.3%
0.0% 0.6% 2.8% 0.0% 0.0% 0.0% 0.0% 0.1% 0.1% 0.0% 22.7%
0 1 1 33 0 1 0 2 2 12 63.5%
0.0% 0.0% 0.0% 1.2% 0.0% 0.0% 0.0% 0.1% 0.1% 0.4% 36.5%
0 0 0 1 50 0 0 0 0 10 82.0%
0.0% 0.0% 0.0% 0.0% 1.9% 0.0% 0.0% 0.0% 0.0% 0.4% 18.0%
0 0 0 0 0 62 0 1 0 4 92.5%
0.0% 0.0% 0.0% 0.0% 0.0% 2.3% 0.0% 0.0% 0.0% 0.1% 7.5%
2 0 0 0 3 2 62 0 0 6 82.7%
0.1% 0.0% 0.0% 0.0% 0.1% 0.1% 2.3% 0.0% 0.0% 0.2% 17.3%
0 3 7 0 0 1 0 57 4 2 77.0%
0.0% 0.1% 0.3% 0.0% 0.0% 0.0% 0.0% 2.1% 0.1% 0.1% 23.0%
2 1 0 0 0 3 1 0 56 2 86.2%
0.1% 0.0% 0.0% 0.0% 0.0% 0.1% 0.0% 0.0% 2.1% 0.1% 13.8%
0 25 1 38 12 27 4 8 5 1922 94.1%
0.0% 0.9% 0.0% 1.4% 0.4% 1.0% 0.1% 0.3% 0.2% 72.0% 5.9%
95.0% 85.8% 87.2% 85.8% 73.5% 63.9% 91.2% 76.0% 78.9% 97.6% 91.0%
5.0% 14.2% 12.8% 14.2% 26.5% 36.1% 8.8% 24.0% 21.1% 2.4% 9.0%

Fold 9


Validation Confusion Matrix
Output Class
Target Class
1 67 0 0 0 4 0 0 0 4 0 89.3%
2.5% 0.0% 0.0% 0.0% 0.1% 0.0% 0.0% 0.0% 0.1% 0.0% 10.7%
2 0 35 2 0 0 0 0 1 1 9 72.9%
0.0% 1.3% 0.1% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.3% 27.1%
3 0 7 61 0 0 0 0 11 0 0 77.2%
0.0% 0.3% 2.3% 0.0% 0.0% 0.0% 0.0% 0.4% 0.0% 0.0% 22.8%
4 0 2 0 20 0 1 1 3 0 7 58.8%
0.0% 0.1% 0.0% 0.7% 0.0% 0.0% 0.0% 0.1% 0.0% 0.3% 41.2%
5 2 0 0 0 52 0 4 0 0 8 78.8%
0.1% 0.0% 0.0% 0.0% 1.9% 0.0% 0.1% 0.0% 0.0% 0.3% 21.2%
6 0 0 0 1 0 45 1 2 1 1 88.2%
0.0% 0.0% 0.0% 0.0% 0.0% 1.7% 0.0% 0.1% 0.0% 0.0% 11.8%
7 0 0 0 0 4 4 60 2 1 6 77.9%
0.0% 0.0% 0.0% 0.0% 0.1% 0.1% 2.2% 0.1% 0.0% 0.2% 22.1%
8 0 2 6 2 0 0 0 51 3 0 79.7%
0.0% 0.1% 0.2% 0.1% 0.0% 0.0% 0.0% 1.9% 0.1% 0.0% 20.3%
9 1 1 1 0 0 0 0 5 52 0 86.7%
0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.0% 0.2% 1.9% 0.0% 13.3%
10 1 30 5 30 24 24 4 9 5 1984 93.8%
0.0% 1.1% 0.2% 1.1% 0.9% 0.9% 0.1% 0.3% 0.2% 74.3% 6.2%
94.4% 55.5% 81.3% 37.7% 61.9% 60.8% 85.7% 60.7% 77.6% 98.5% 90.9%
5.6% 44.5% 18.7% 62.3% 38.1% 39.2% 14.3% 39.3% 22.4% 1.5% 9.1%
1 2 3 4 5 6 7 8 9 10

Fold 10



(In tabelele matricilor de confuzie indicii sunt reprezentati de la 1 la 10 si nu de la 0 la 9 precum clasele din care fac parte exemplele.)

In urma celor 10 matrici de confuzie am observat ca un numar foarte mare de exemple ce in realitate fac parte din clasa 9 sunt clasificate ca fiind in clasa 9 -- incepand de la primul fold unde avem un procent de 96.9% pana in ultimul fold unde ajungem la un procent de aproape 99%. Pe de alta parte, reteaua ajunge sa clasifice mult prea mult exemple ca fiind in clasa 9, astfel doar 94% din cele clasificate ca fiind in clasa 9 chiar fac parte din aceasta.

De asemenea, putem observa ca cele mai multe misclasificari ale retelei sunt exemplele din clasa 1, 3, 4 si 5 pe care le considera ca apartin clasei 9. Din acest motiv, am adaugat cea de-a doua retea care contine un numar aproximativ egal de exemple din fiecare clasa.

Alte variante ale retelei:

Pentru a imbunatati reteaua am mai incercat o serie de idei precum adaugarea regularizarii si a normalizarii standard, initializarea ponderile de pe fiecare strat cu valori repartizate uniform in intervalul [-3 * n^(-0.5), 3 * n^(-0.5)], unde n reprezinta numarul de intrari si 5-fold cross-validation pentru reteaua cu input echilibrat.