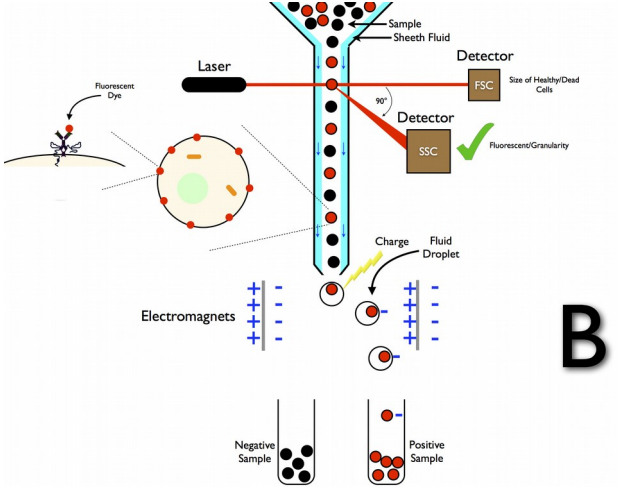
Klastrovanie buniek podľa cytometrických dát

Marek Behún

Popis úlohy

Input: flow cytometry dáta (100k+, 13 až 50 dimenzií)
(napríklad bunky z kostnej drene od zdravého darcu)
Chceme: rozoznať populácie rovnakých buniek



Problém

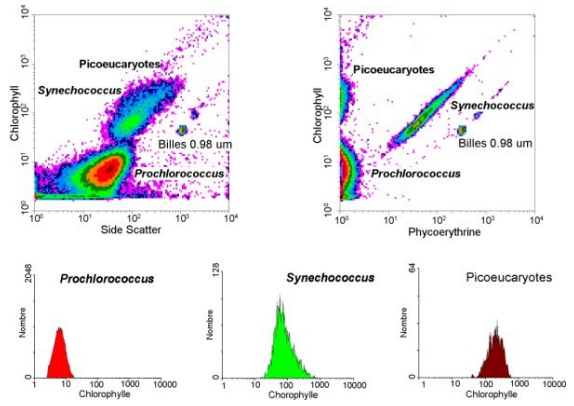
Pri veľkom počte buniek štandardné klastrovacie algoritmy nefungujú dobre:

- šum
- evolučná kontinuita
- ...

Lepšie algoritmy sú zase pomalé (rádovo 100 tisíce buniek).

Hodnotenie presnosti

- časti datasetov sú olabelované tzv. “manuálnym gatingom”
  - ide o vizuálne pozorovanie scatterplotov, histogramov, …
  - limitácie: subjektivita, chyba operátora, zlá reproducibilita, ťažkosti detekovať nové typy buniek
- používame F1 skóre
- hľadáme priradenie clusterov s najlepším súčtom F1 (cez všetky permutácie)

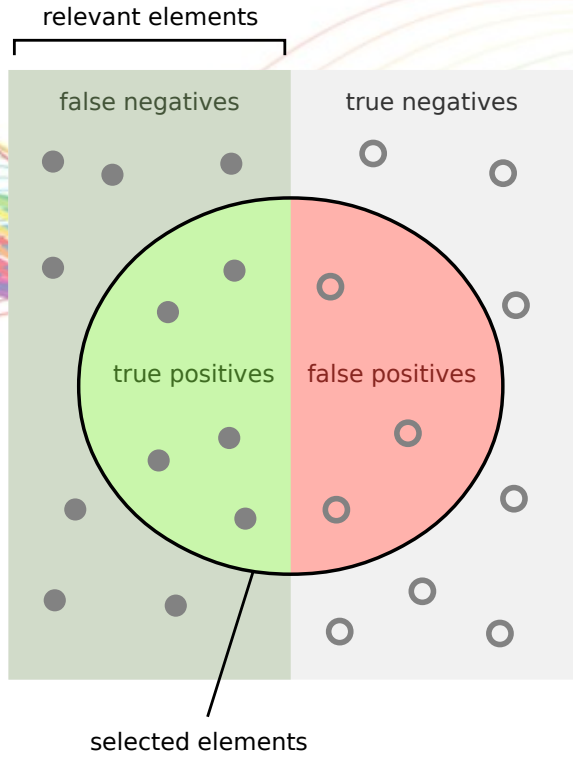


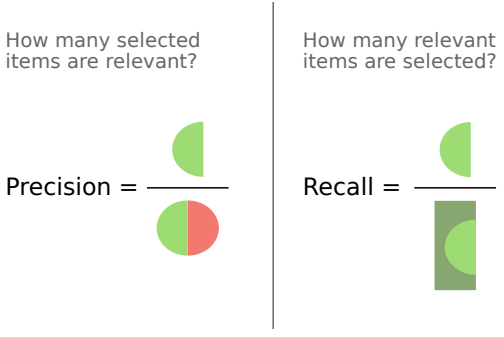
F1 skóre

- harmonický priemer *precision* a *recall*

$$F_1 = \left(\frac{precision^{-1} + recall^{-1}}{2} \right)^{-1}$$

- F_1 je v [0, 1]
- F_1 = 1 znamená perfektný *precision* a *recall*
- F_1 = 0 najhorší





Použité datasety

Levine_13dim
- 13 dimenzií
- 167 034 záznamov
- 49% manuálne gatovaných

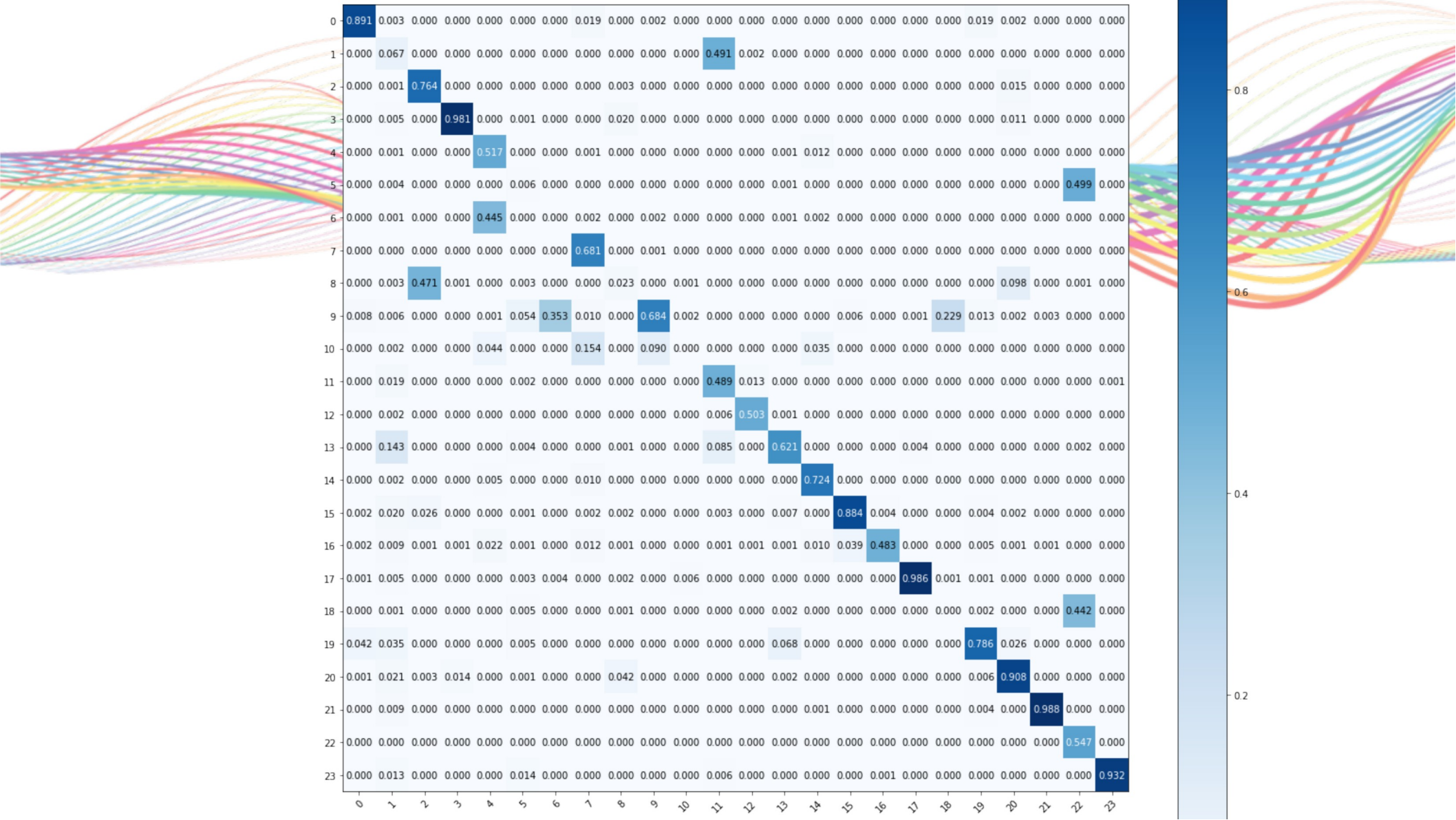
Samusik_all
- 39 dimenzií
- 841 643 záznamov
- 61% manuálne gatovaných

V oboch prípadoch surové dáta prešli transformáciou arcsinh(x/5)

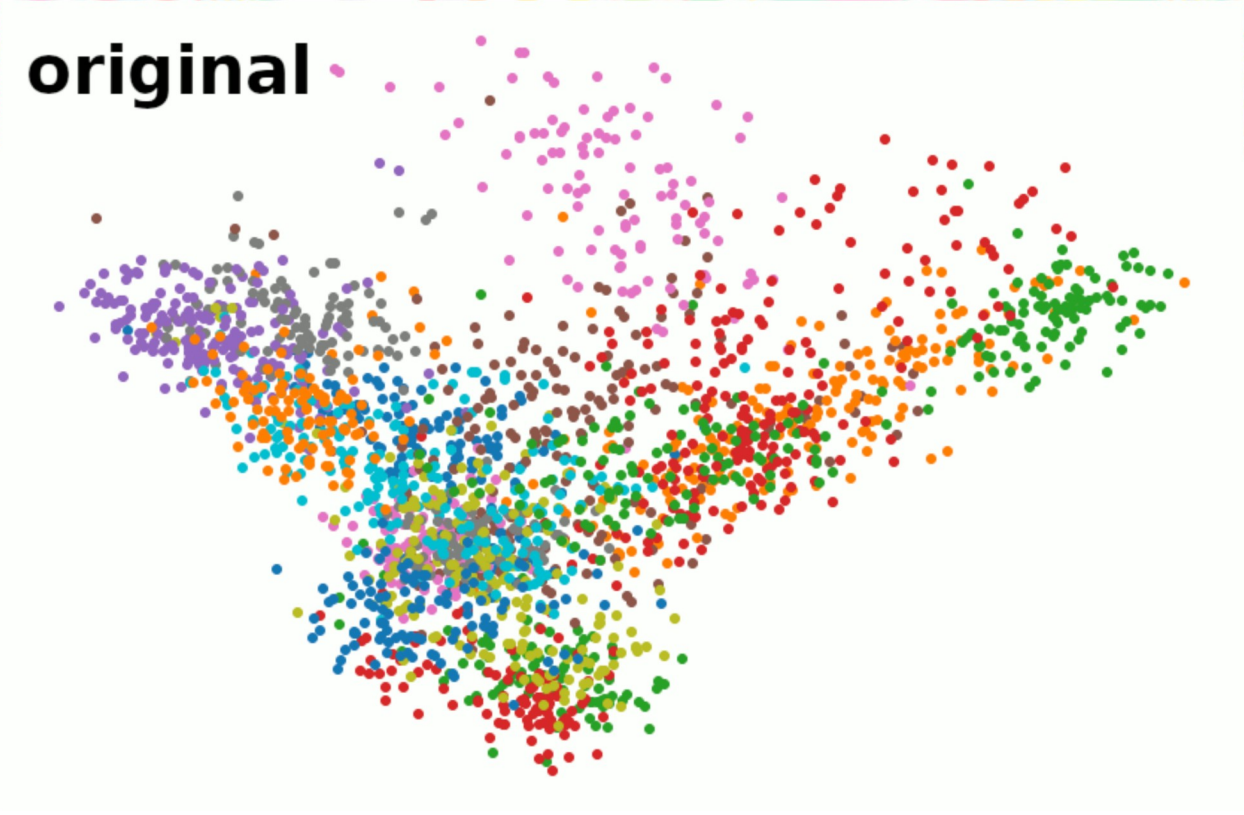
KMeans: Samusik

F_1 skóre: 0.547035

Čas: 97s

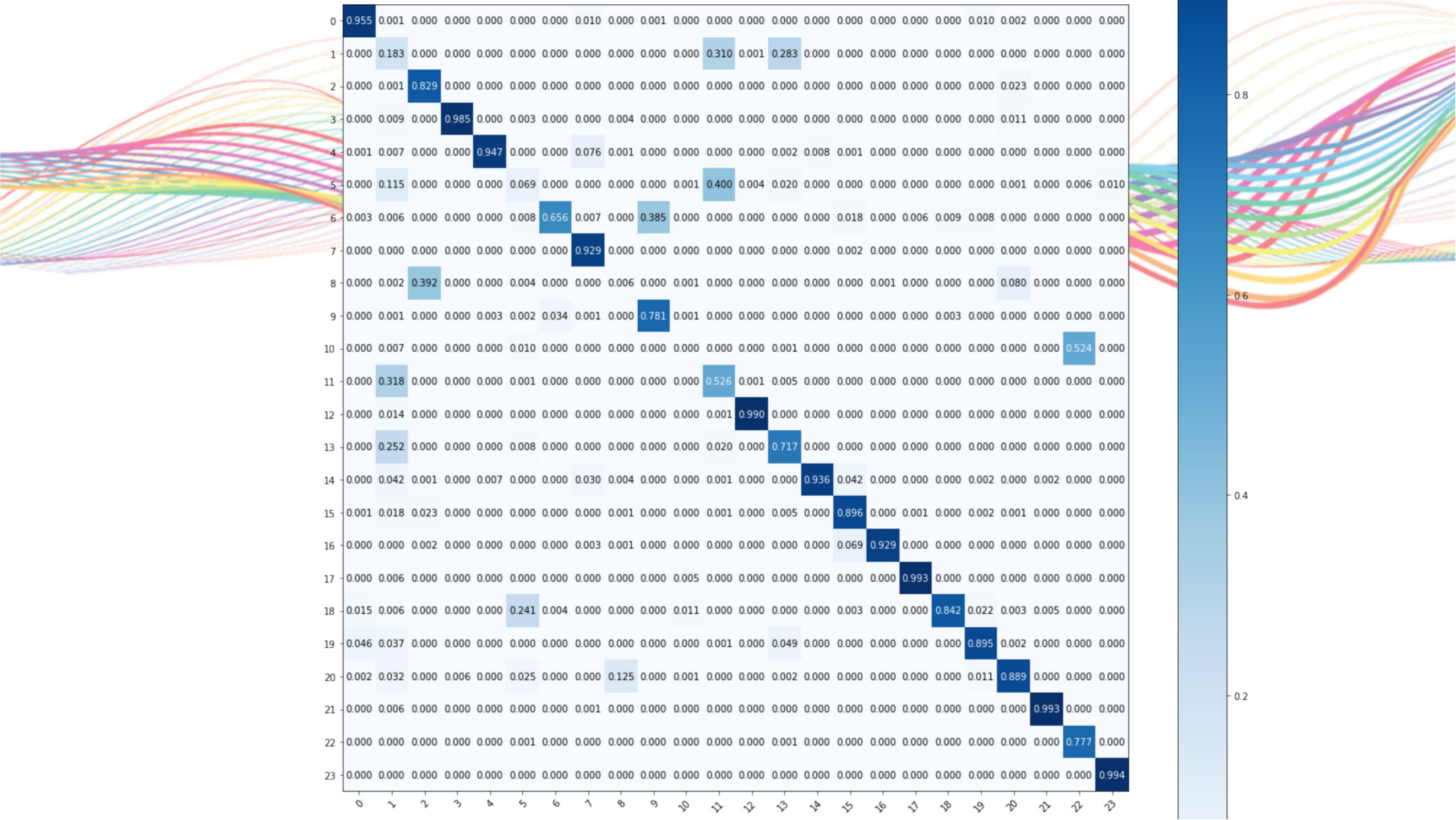

0 0.891 0.003 0.000 0.000 0.000 0.000 0.000 0.019 0.000 0.002 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.019 0.002 0.000 0.000 0.000
1 0.000 0.067 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.491 0.002 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
2 0.000 0.001 0.764 0.000 0.000 0.000 0.000 0.000 0.003 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.015 0.000 0.000 0.000
3 0.000 0.005 0.000 0.981 0.000 0.001 0.000 0.000 0.020 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.011 0.000 0.000 0.000
4 0.000 0.001 0.000 0.000 0.517 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.001 0.012 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
5 0.000 0.004 0.000 0.000 0.000 0.006 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.499 0.000
6 0.000 0.001 0.000 0.000 0.445 0.000 0.000 0.002 0.000 0.002 0.000 0.000 0.000 0.001 0.002 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
7 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.681 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
8 0.000 0.003 0.471 0.001 0.000 0.003 0.000 0.000 0.023 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.098 0.000 0.001 0.000
9 0.008 0.006 0.000 0.000 0.001 0.054 0.353 0.010 0.000 0.684 0.002 0.000 0.000 0.000 0.000 0.006 0.000 0.001 0.229 0.013 0.002 0.003 0.000 0.000
10 0.000 0.002 0.000 0.000 0.044 0.000 0.000 0.154 0.000 0.090 0.000 0.000 0.000 0.000 0.035 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
11 0.000 0.019 0.000 0.000 0.000 0.002 0.000 0.000 0.000 0.000 0.000 0.489 0.013 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001
12 0.000 0.002 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.006 0.503 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
13 0.000 0.143 0.000 0.000 0.000 0.004 0.000 0.000 0.001 0.000 0.000 0.085 0.000 0.621 0.000 0.000 0.000 0.004 0.000 0.000 0.000 0.000 0.002 0.000
14 0.000 0.002 0.000 0.000 0.005 0.000 0.000 0.010 0.000 0.000 0.000 0.000 0.000 0.000 0.724 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
15 0.002 0.020 0.026 0.000 0.000 0.001 0.000 0.002 0.002 0.000 0.000 0.003 0.000 0.007 0.000 0.884 0.004 0.000 0.000 0.004 0.002 0.000 0.000 0.000
16 0.002 0.009 0.001 0.001 0.022 0.001 0.000 0.012 0.001 0.000 0.000 0.001 0.001 0.001 0.010 0.039 0.483 0.000 0.000 0.005 0.001 0.001 0.000 0.000
17 0.001 0.005 0.000 0.000 0.000 0.003 0.004 0.000 0.002 0.000 0.006 0.000 0.000 0.000 0.000 0.000 0.000 0.986 0.001 0.001 0.000 0.000 0.000 0.000
18 0.000 0.001 0.000 0.000 0.000 0.005 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.002 0.000 0.000 0.000 0.000 0.000 0.002 0.000 0.000 0.442 0.000
19 0.042 0.035 0.000 0.000 0.000 0.005 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.068 0.000 0.000 0.000 0.000 0.000 0.786 0.026 0.000 0.000 0.000
20 0.001 0.021 0.003 0.014 0.000 0.001 0.000 0.000 0.042 0.000 0.000 0.000 0.000 0.002 0.000 0.000 0.000 0.000 0.000 0.006 0.908 0.000 0.000 0.000
21 0.000 0.009 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.004 0.000 0.988 0.000 0.000
22 0.000 0.547 0.000
23 0.000 0.013 0.000 0.000 0.000 0.014 0.000 0.000 0.000 0.000 0.000 0.006 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.932
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
0.8
0.6
0.4
0.2

HCA: Samusik

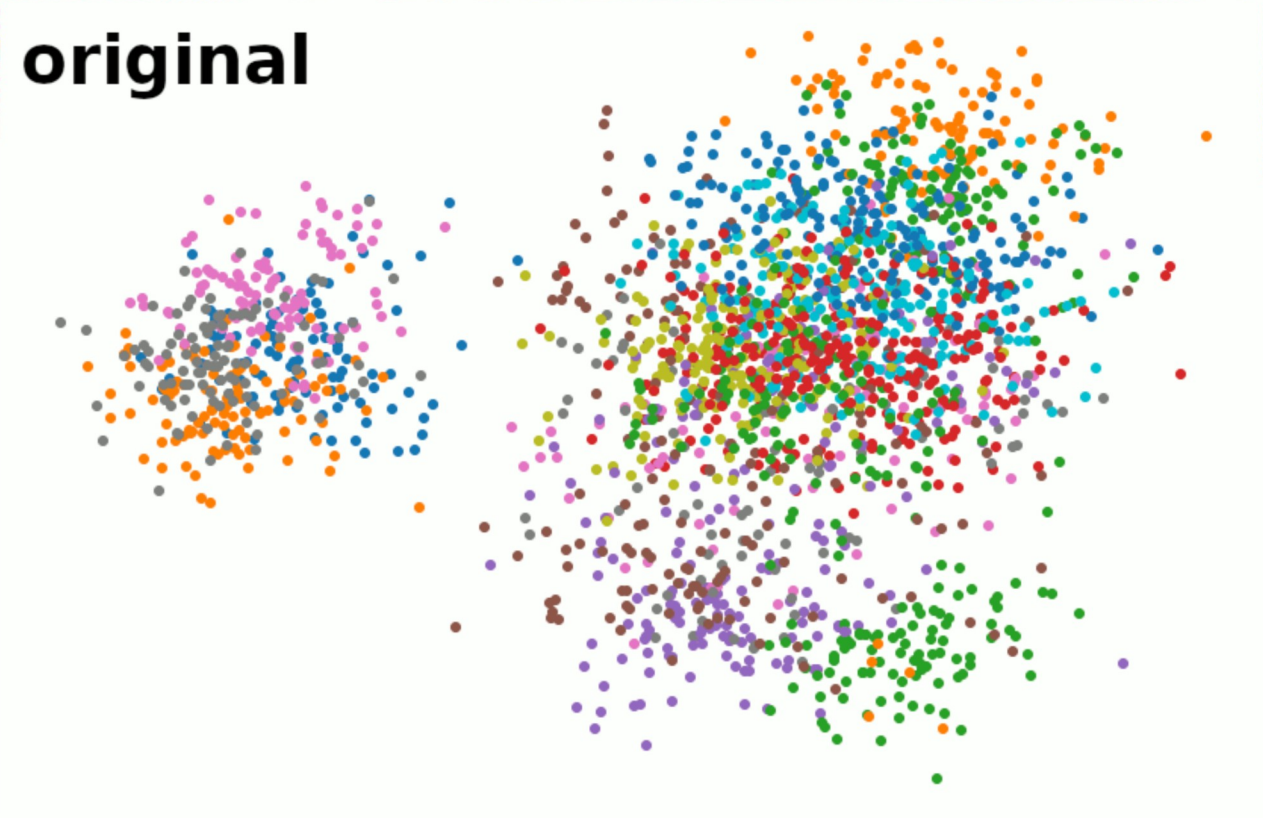


F_1 skóre: 0.738262

Čas: 434s (86 468 samples)

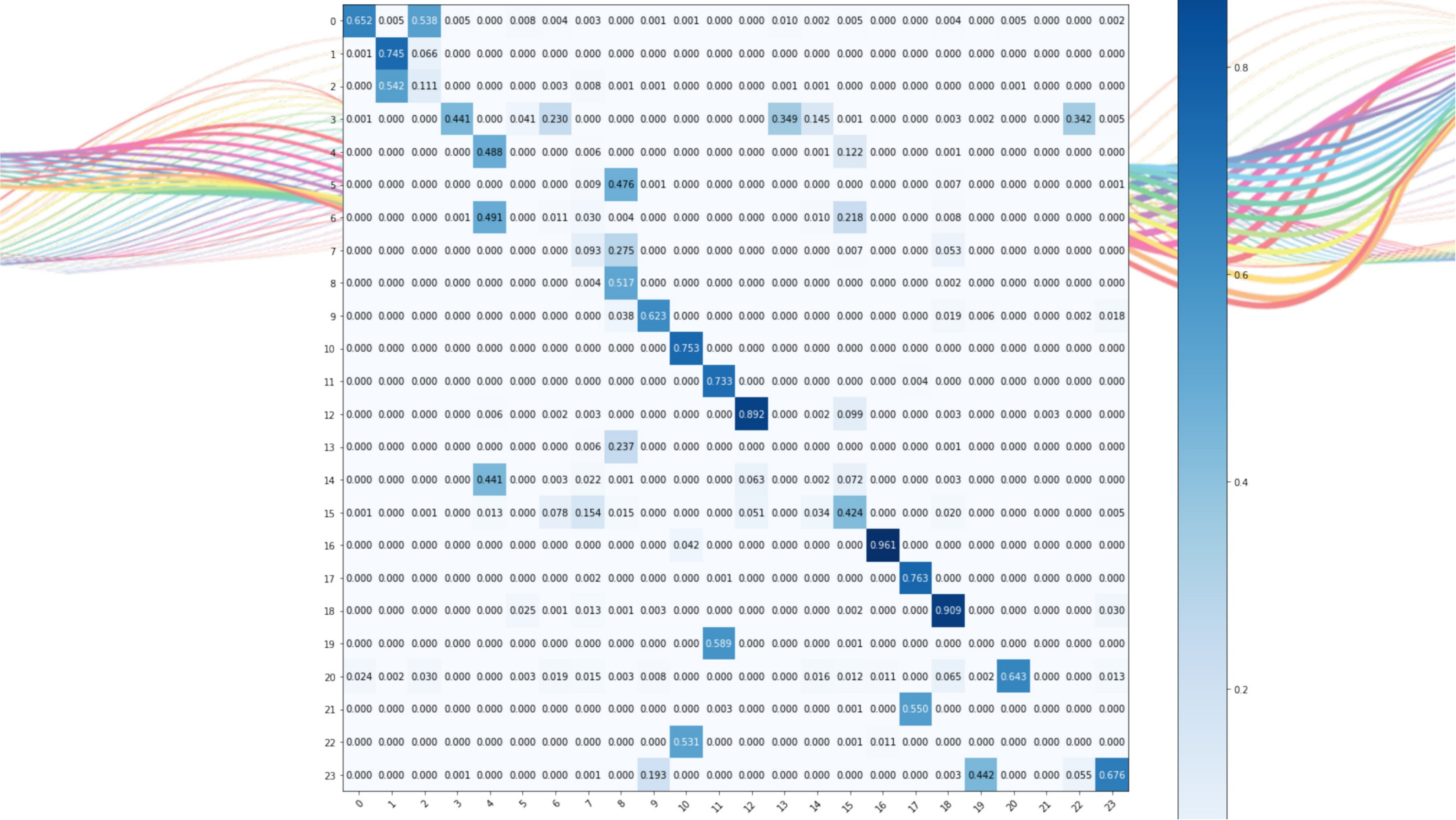

0 0.955 0.001 0.000 0.000 0.000 0.000 0.000 0.010 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.010 0.002 0.000 0.000 0.000
1 0.000 0.183 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.310 0.001 0.283 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
2 0.000 0.001 0.829 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.023 0.000 0.000 0.000
3 0.000 0.009 0.000 0.985 0.000 0.003 0.000 0.000 0.004 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.011 0.000 0.000 0.000
4 0.001 0.007 0.000 0.000 0.947 0.000 0.000 0.076 0.001 0.000 0.000 0.000 0.000 0.002 0.008 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
5 0.000 0.115 0.000 0.000 0.000 0.069 0.000 0.000 0.000 0.000 0.001 0.400 0.004 0.020 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.006 0.010
6 0.003 0.006 0.000 0.000 0.000 0.008 0.656 0.007 0.000 0.385 0.000 0.000 0.000 0.000 0.000 0.018 0.000 0.006 0.009 0.008 0.000 0.000 0.000 0.000 0.000
7 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.929 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.002 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
8 0.000 0.002 0.392 0.000 0.000 0.004 0.000 0.000 0.006 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.080 0.000 0.000 0.000 0.000
9 0.000 0.001 0.000 0.000 0.003 0.002 0.034 0.001 0.000 0.781 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.003 0.000 0.000 0.000 0.000 0.000 0.000
10 0.000 0.007 0.000 0.000 0.000 0.010 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.524 0.000
11 0.000 0.318 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.526 0.001 0.005 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
12 0.000 0.014 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.990 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
13 0.000 0.252 0.000 0.000 0.000 0.008 0.000 0.000 0.000 0.000 0.000 0.020 0.000 0.717 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
14 0.000 0.042 0.001 0.000 0.007 0.000 0.000 0.030 0.004 0.000 0.000 0.001 0.000 0.000 0.936 0.042 0.000 0.000 0.000 0.002 0.000 0.002 0.000 0.000 0.000
15 0.001 0.018 0.023 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.001 0.000 0.005 0.000 0.896 0.000 0.001 0.000 0.002 0.001 0.000 0.000 0.000 0.000
16 0.000 0.000 0.002 0.000 0.000 0.000 0.000 0.003 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.069 0.929 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
17 0.000 0.006 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.005 0.000 0.000 0.000 0.000 0.000 0.000 0.993 0.000 0.000 0.000 0.000 0.000 0.000 0.000
18 0.015 0.006 0.000 0.000 0.000 0.241 0.004 0.000 0.000 0.000 0.011 0.000 0.000 0.000 0.000 0.003 0.000 0.000 0.842 0.022 0.003 0.005 0.000 0.000 0.000
19 0.046 0.037 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.049 0.000 0.000 0.000 0.000 0.000 0.895 0.002 0.000 0.000 0.000 0.000
20 0.002 0.032 0.000 0.006 0.000 0.025 0.000 0.000 0.125 0.000 0.001 0.000 0.000 0.002 0.000 0.000 0.000 0.000 0.000 0.011 0.889 0.000 0.000 0.000 0.000
21 0.000 0.006 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.993 0.000 0.000 0.000
22 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.777 0.000 0.000
23 0.000 0.994 0.000
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
0.8
0.6
0.4
0.2

KMeans: Levine13

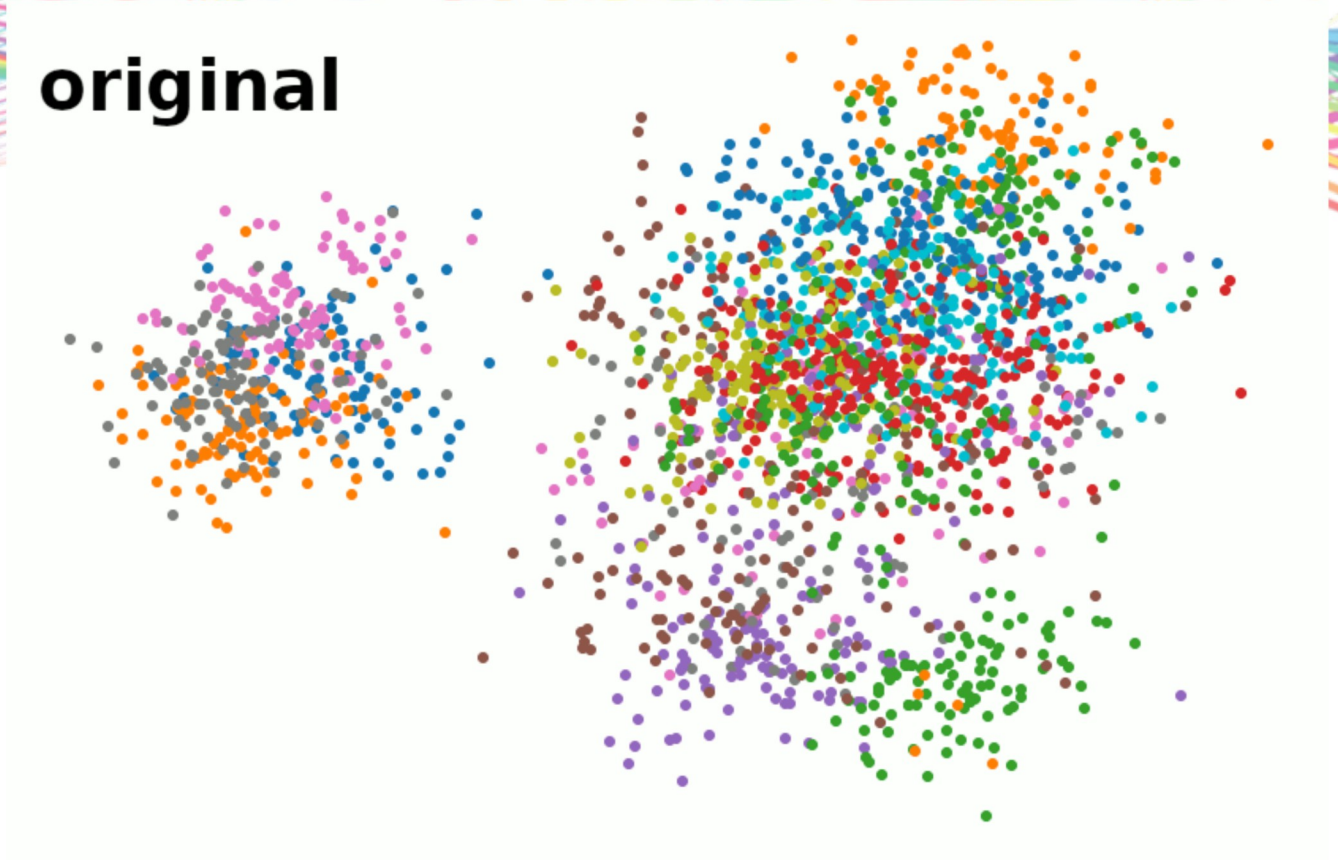


F_1 skóre: 0.434779

Čas: 11s

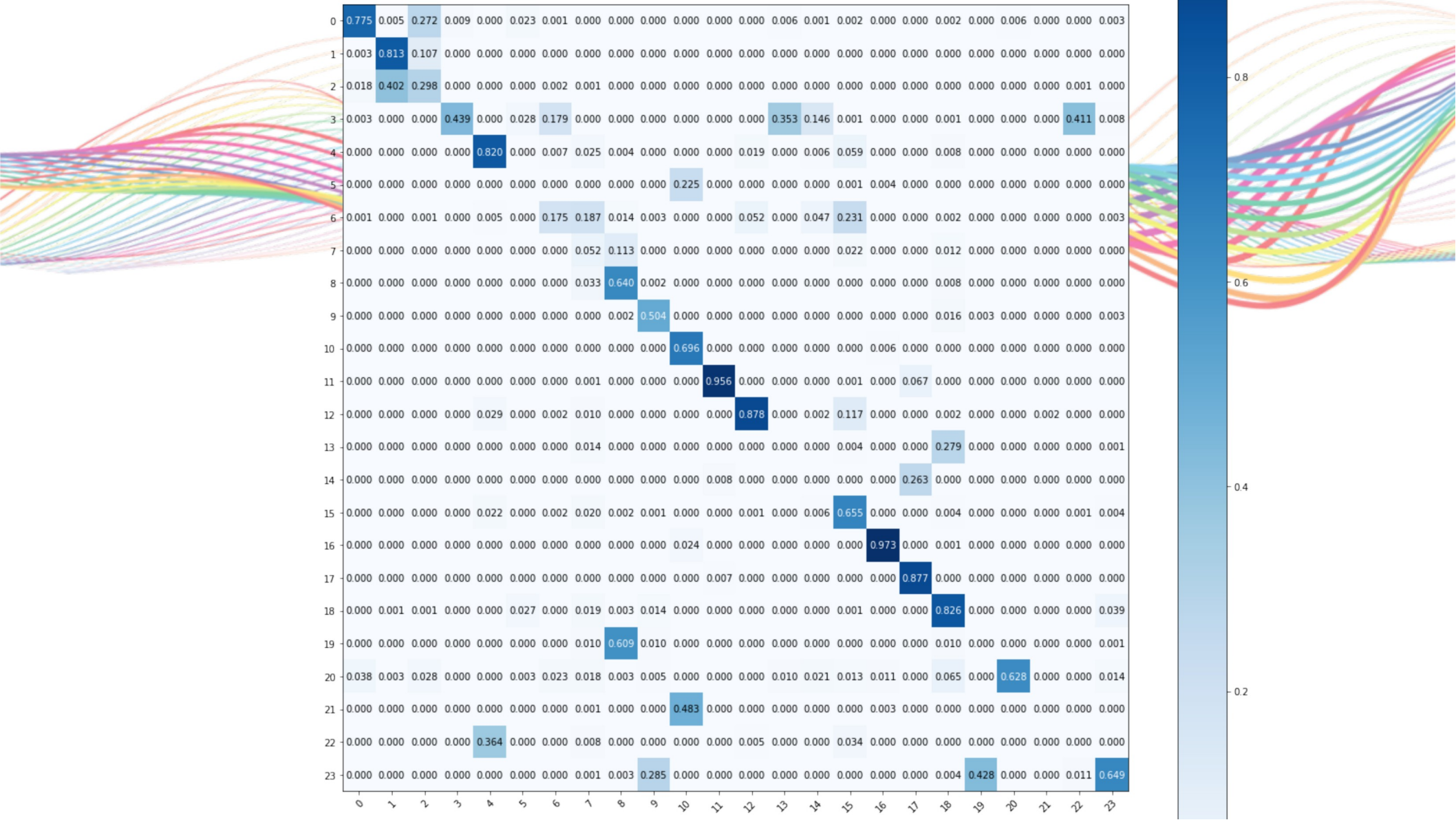

0 0.652 0.005 0.538 0.005 0.000 0.008 0.004 0.003 0.000 0.001 0.001 0.000 0.000 0.010 0.002 0.005 0.000 0.000 0.004 0.000 0.005 0.000 0.000 0.002
1 0.001 0.745 0.066 0.000
2 0.000 0.542 0.111 0.000 0.000 0.000 0.003 0.008 0.001 0.001 0.000 0.000 0.000 0.001 0.001 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000
3 0.001 0.000 0.000 0.441 0.000 0.041 0.230 0.000 0.000 0.000 0.000 0.000 0.000 0.349 0.145 0.001 0.000 0.000 0.003 0.002 0.000 0.000 0.342 0.005
4 0.000 0.000 0.000 0.000 0.488 0.000 0.000 0.006 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.122 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000
5 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.009 0.476 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.007 0.000 0.000 0.000 0.000 0.001
6 0.000 0.000 0.000 0.001 0.491 0.000 0.011 0.030 0.004 0.000 0.000 0.000 0.000 0.000 0.010 0.218 0.000 0.000 0.008 0.000 0.000 0.000 0.000 0.000
7 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.093 0.275 0.000 0.000 0.000 0.000 0.000 0.000 0.007 0.000 0.000 0.053 0.000 0.000 0.000 0.000 0.000
8 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.004 0.517 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.002 0.000 0.000 0.000 0.000 0.000
9 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.038 0.623 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.019 0.006 0.000 0.000 0.002 0.018
10 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.753 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
11 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.733 0.000 0.000 0.000 0.000 0.000 0.004 0.000 0.000 0.000 0.000 0.000 0.000
12 0.000 0.000 0.000 0.000 0.006 0.000 0.002 0.003 0.000 0.000 0.000 0.000 0.892 0.000 0.002 0.099 0.000 0.000 0.003 0.000 0.000 0.003 0.000 0.000
13 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.006 0.237 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000
14 0.000 0.000 0.000 0.000 0.441 0.000 0.003 0.022 0.001 0.000 0.000 0.000 0.063 0.000 0.002 0.072 0.000 0.000 0.003 0.000 0.000 0.000 0.000 0.000
15 0.001 0.000 0.001 0.000 0.013 0.000 0.078 0.154 0.015 0.000 0.000 0.000 0.051 0.000 0.034 0.424 0.000 0.000 0.020 0.000 0.000 0.000 0.000 0.005
16 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.042 0.000 0.000 0.000 0.000 0.000 0.961 0.000 0.000 0.000 0.000 0.000 0.000 0.000
17 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.002 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.763 0.000 0.000 0.000 0.000 0.000 0.000
18 0.000 0.000 0.000 0.000 0.000 0.025 0.001 0.013 0.001 0.003 0.000 0.000 0.000 0.000 0.000 0.002 0.000 0.000 0.909 0.000 0.000 0.000 0.000 0.030
19 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.589 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
20 0.024 0.002 0.030 0.000 0.000 0.003 0.019 0.015 0.003 0.008 0.000 0.000 0.000 0.000 0.016 0.012 0.011 0.000 0.065 0.002 0.643 0.000 0.000 0.013
21 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.003 0.000 0.000 0.000 0.001 0.000 0.550 0.000 0.000 0.000 0.000 0.000 0.000
22 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.531 0.000 0.000 0.000 0.000 0.001 0.011 0.000 0.000 0.000 0.000 0.000 0.000 0.000
23 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.001 0.000 0.193 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.003 0.442 0.000 0.000 0.055 0.676
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
0.8
0.6
0.4
0.2

HCA: Levine13



F_1 skóre: 0.485611

Čas: 207s


0 0.775 0.005 0.272 0.009 0.000 0.023 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.006 0.001 0.002 0.000 0.000 0.002 0.000 0.006 0.000 0.000 0.003
1 0.003 0.813 0.107 0.000
2 0.018 0.402 0.298 0.000 0.000 0.000 0.002 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000
3 0.003 0.000 0.000 0.439 0.000 0.028 0.179 0.000 0.000 0.000 0.000 0.000 0.000 0.353 0.146 0.001 0.000 0.000 0.001 0.000 0.000 0.000 0.411 0.008
4 0.000 0.000 0.000 0.000 0.820 0.000 0.007 0.025 0.004 0.000 0.000 0.000 0.019 0.000 0.006 0.059 0.000 0.000 0.008 0.000 0.000 0.000 0.000 0.000
5 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.225 0.000 0.000 0.000 0.000 0.001 0.004 0.000 0.000 0.000 0.000 0.000 0.000 0.000
6 0.001 0.000 0.001 0.000 0.005 0.000 0.175 0.187 0.014 0.003 0.000 0.000 0.052 0.000 0.047 0.231 0.000 0.000 0.002 0.000 0.000 0.000 0.000 0.003
7 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.052 0.113 0.000 0.000 0.000 0.000 0.000 0.000 0.022 0.000 0.000 0.012 0.000 0.000 0.000 0.000 0.000
8 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.033 0.640 0.002 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.008 0.000 0.000 0.000 0.000 0.000
9 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.002 0.504 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.016 0.003 0.000 0.000 0.000 0.003
10 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.696 0.000 0.000 0.000 0.000 0.000 0.006 0.000 0.000 0.000 0.000 0.000 0.000 0.000
11 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.956 0.000 0.000 0.000 0.001 0.000 0.067 0.000 0.000 0.000 0.000 0.000 0.000
12 0.000 0.000 0.000 0.000 0.029 0.000 0.002 0.010 0.000 0.000 0.000 0.000 0.878 0.000 0.002 0.117 0.000 0.000 0.002 0.000 0.000 0.002 0.000 0.000
13 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.014 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.004 0.000 0.000 0.279 0.000 0.000 0.000 0.000 0.001
14 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.008 0.000 0.000 0.000 0.000 0.000 0.263 0.000 0.000 0.000 0.000 0.000 0.000
15 0.000 0.000 0.000 0.000 0.022 0.000 0.002 0.020 0.002 0.001 0.000 0.000 0.001 0.000 0.006 0.655 0.000 0.000 0.004 0.000 0.000 0.000 0.001 0.004
16 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.024 0.000 0.000 0.000 0.000 0.000 0.973 0.000 0.001 0.000 0.000 0.000 0.000 0.000
17 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.007 0.000 0.000 0.000 0.000 0.000 0.877 0.000 0.000 0.000 0.000 0.000 0.000
18 0.000 0.001 0.001 0.000 0.000 0.027 0.000 0.019 0.003 0.014 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.826 0.000 0.000 0.000 0.000 0.039
19 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.010 0.609 0.010 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.010 0.000 0.000 0.000 0.000 0.001
20 0.038 0.003 0.028 0.000 0.000 0.003 0.023 0.018 0.003 0.005 0.000 0.000 0.000 0.010 0.021 0.013 0.011 0.000 0.065 0.000 0.628 0.000 0.000 0.014
21 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.483 0.000 0.000 0.000 0.000 0.000 0.003 0.000 0.000 0.000 0.000 0.000 0.000 0.000
22 0.000 0.000 0.000 0.000 0.364 0.000 0.000 0.008 0.000 0.000 0.000 0.000 0.005 0.000 0.000 0.034 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
23 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.003 0.285 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.004 0.428 0.000 0.000 0.011 0.649
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
0.8
0.6
0.4
0.2

Záver

Levine_13dim

- nepodarilo sa prekročiť F_1 skóre z originálneho článku (0.518)
- HCA s 0.485 by bol tretích

Samusik_all

- najlepšie F_1 je v originálnom článku 0.702
- HCA: 0.738
- neviem ale, či by bol rýchlejší :-(

Future: Mahalanobis Clustering?

- chcel by som vyskúšať aglomeratívny clustering s Mahalanobisovou metrikou
- predstavujem si, že zhluky sa môžu vo flow cytometry dátach tvarovať do elipsoidov
- zatiaľ to nie je naimplementované v SciPy / scikit-learn (skúšal som to dohackovať, ale nepodarilo sa)

Zdroje a odkazy

Videá: https://blackhole.sk/~kabel/doc/bioinf

Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data, https://doi.org/10.1002/cyto.a.23030