SOM & KMeans Comparison

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Methods Tags Batch no. | Algo | Initial Centroids/Neurons | Results | | |
| Total Iterations | Clusters | Outliers |
| 5 | KMeans | Farthest : [0, 3] | 2 | C0: [0, 1, 6, 7, 8, 10, 12]  C1: [2, 3, 4, 5, 9, 11] | No |
| Random | 4 | C0: [2, 3, 4, 5, 9, 11, 12] C1: [0, 1, 6, 7, 8, 10] | 12 |
| SOM | Original tags Farthest : [0, 2] | 39 | C0: [0, 1, 6, 7, 8, 10, 12]  C1: [2, 3, 4, 5, 9, 11] | No |
| Base tags  Farthest : [0, 2] | 39 | C0: [0, 1, 6, 7, 8, 10, 12]  C1: [2, 3, 4, 5, 9, 11] | No |
| 4 | KMeans | Farthest : [0, 2] | 2 | C0: [0, 1, 6, 7, 8, 9, 10, 11, 12]  C1: [2, 3, 4, 5] | 9, 11 |
| Random | 2 | C0: [2, 3, 6, 8 10, 12]  C1: [0, 1, 4, 5, 7, 9, 11] | 2, 3, 0, 1, 7, 9, 11 |
| SOM | Farthest : [0, 2] | 39 | C0: [0, 1, 6, 7, 8]  C1: [2, 3, 4, 5, 9, 10, 11, 12] | 10, 12 |
| Random |  | Doesn’t give appropriate results |  |
| 10 | KMeans with original tags |  |  |  |  |
| KMeans with base words |  |  |  |  |
|  | Random | 3 | C0: [5, 7, 8, 10, 14]  C1: [0, 1, 2, 3, 4, 6, 9, 11, 12, 13, 15]  C2: [16] | Too many |
| SOM with base words | Farthest: [0, 4, 6] | 39 | C0: [0, 1, 3, 9, 14]  C1: [6, 13, 15, 16]  C2: [2, 4, 5, 7, 8, 10, 11, 12] | 2, 12 |
| SOM with original tags | Farthest: [0, 2, 4, 6, 13] | 39 | C0: [0, 1, 3, 9, 14]  C1: [2, 12]  C2: [4, 5, 7, 8, 10, 11]  C3: [6, 16]  C4: [13, 15] | C0 & C1 to be merged. C3 & C4 to be merged |
| 7 | KMeans | Farthest : [0, 3] | 2 | C0: [0, 1, 6, 7, 8, 10, 12]  C1: [2, 3, 4, 5, 9, 11] | No |
| Random | 4 | C0: [2, 3, 4, 5, 9, 11, 12] C1: [0, 1, 6, 7, 8, 10] | 12 |
| SOM | Original tags Farthest : [0, 7, 13, 14] | 39 | C0: [0, 1, 2, 3, 4, 5, 6]  C1: [7,8, 9, 10, 11, 12]  C2: [13, 15]  C3: [14, 16] | 14, 16 were better to be the part of C2 |
| Base tags  Farthest : [0, 7, 13] | 39 | C0: [0, 1, 2, 3, 4, 5, 6]  C1: [7, 8, 9, 10, 11, 12]  C2: [13, 14, 15, 16] | No |
| 6 | KMeans |  |  |  |  |
|  |  |  |  |
| SOM | Original tags Farthest : [0, 7] | 39 | C0: [0, 1, 2, 3, 4, 5, 6]  C1: [7, 8, 9, 10, 11, 12] | No |
| Base tags  Farthest : [0, 7] | 39 | C0: [0, 1, 2, 3, 4, 5, 6]  C1: [7, 8, 9, 10, 11, 12] | No |
| 8 | KMeans |  |  |  |  |
|  |  |  |  |
| SOM | Original tags Farthest : [0, 1, 4] | 39 | C0: [0, 3]  C1: [1,2, 6]  C3: [4, 5, 7] | No |
| Base tags  Farthest : [0, 1, 4] | 39 | C0: [0, 3]  C1: [1,2, 6]  C3: [4, 5, 7] | No |
| 9 | KMeans |  |  |  |  |
|  |  |  |  |
| SOM | Original tags Farthest : [0, 4, 6, 13] | 39 | C0: [0, 1, 2, 3, 9, 12, 14]  C1: [4, 5, 7, 8, 10, 11]  C2: [6, 16]  C3: [13, 15] | C2, C3 better to be merged |
| Base tags  Farthest [0, 4, 6] | 39 | C0: [0, 1, 3, 9, 14]  C1: [2, 4, 5, 7, 8, 10, 11, 12]  C2: [6, 13, 15, 16] | 2, 12 |
| Data provided by Shamsa | KMeans |  |  |  |  |
|  |  |  |  |
| SOM | Original tags Farthest : [0, 1, 4, 5, 7] | 39 | C0: [0]  C1: [1, 2, 3, 6]  C2: [4]  C3: [5]  C4: [7] | None |
| Base tags  Farthest [ | 39 |  |  |

* Reason of outliers in batch 4: compound tags