NMF in Screening Some Spirometric Data, an Insight into 12-Dimensional Data Space

Anna M. Bartkowiak[1(✉)] and Jerzy Liebhart[2]

[1] Institute of Computer Science, Wroclaw University, 50-383 Wroclaw, Poland
aba@cs.uni.edu.pl
[2] Department and Clinic of Internal Diseases and Allergology,
Wroclaw Medical University, Wroclaw, Poland

Abstract. We present the usage of the Non-negative Matrix Factorization (NMF), an unsupervised machine learning method, which learns normal and abnormal state of patient's ventilatory systems. This is done using samples of patients having defects of obturative and restrictive kind and a control group.

We show that the NMF method can identify patients being in the normal state and screen them off from the remaining patients; however the kind of the ventilatory disorder for the remaining patients is not recognized. This is confronted with clustering provided by the k-means method and visualization of the 12-dimensional data using heatmaps and Kohonen's self-organizing maps.

The data set can be reconstructed with a 0.9746 accuracy (fraction of explained variance) from 6 base vectors provided by the NMF and using appropriate encoders provided also by the NMF; while 3 factors yield an 0.8573 fraction of explained variance.

Keywords: Healthy state · Abnormal state · Non-negative matrix factorization (NMF) · Heatmap · Self-organizing map (SOM) · Inner factors · Reconstruction of data

1 Introduction

Non-linear Matrix Factorization launched in [9] is a method for *approximating* a given real data matrix by derived factor matrices of lower rank. It may by aligned with two others methods serving for this purpose for a centenary of years: Principal Components (PCA) and Singular Value Decomposition (SVD) [6] based on spectral decomposition of a (preprocessed) data matrix. There are essential differences in the criteria of optimality of these methods, which lead to different factorization of the given data matrix. Generally speaking, NMF works on the assumption that both the analyzed data matrix and the derived factor matrices should be non-negative, that is to mean, their elements are allowed to show only non-negative values. It appears that this leads to better interpretability of

K. Saeed et al. (Eds.): CISIM 2017, LNCS 10244, pp. 155–166, 2017.
DOI: 10.1007/978-3-319-59105-6_14

the derived factors. The NMF method gains more and more popularity, especially that feasible algorithms working under the constraints of non-negativity of all the derived elements have been elaborated (see, e.g., [4]). There are also algorithms for big data and data streams [14].

All the three mentioned data analysis methods (PCA, SVD and NMF) are frequently viewed also as dimensionality reducing methods. A satisfactory ('good') reduction to dimensionality equal 2 or 3 makes, that the data may be visualized in a 2D or 3D space, and usually we have some additional information on the fidelity of this visualization. By looking at the 2D or 3D representation, we may get insight into the multivariate data space. Such a visualization may also be the basis for finding some structure in the data, e.g. some clusters. This is difficult to achieve when working with PCA or SVD. However NMF, thanks to the non-negative constraints, *may* provide directly some information on the group membership of the data vectors contained in the analyzed data set. In the following we will consider *data vectors* that denote some *subjects* (e.g. patients), each of them characterized by a number of attributes, called in statistical data analysis *variables*. The clustering information is not always meaningful, and for some data sets we may get the clustering of the subjects, and for other not [1,15]. Moreover, sometimes we may get even a *bi-clustering*, that is a simultaneous clustering both of subjects and of variables [7,10]. Of course, the NMF – as *a priori* un-supervised machine learning method – gets only the data for analysis, however it gets no *a priori* information on the putative groups contained in the analyzed data.

The main goal of this paper is to explore the clustering ability of NMF when applied to a real medical data set concerned with diagnosing normal and abnormal state of a sample of patients, part of them suffering from pulmonary diseases, and the other part being in normal state. The data set is not big; it contains 77 patients, each characterized by 12 variables. We will show in detail that NMF is able to screen off the normal state patients, and even find some outliers in the data. We will confirm our results by using alternate visualization methods dealing with multivariate data. This statement ends the first Section of the paper.

In next Sect. 2, we will show in more detail the pulmonary data used in our elaboration. They will be visualized by a heatmap. Then, Sect. 3 shows the details of the NMF model and its work when reducing the dimensions of the data to rank 2 and rank 3. The indication of the clusters will be discussed. Section 4 shows analogous results obtained by the k-means method, when splitting directly the data into $k = 2$ and $k = 3$ clusters. Section 5 shows an alternate multivariate visualization of the of the data using Kohonen's self-organizing maps. Also the placement of the true group structure obtained from an independent medical diagnosis is depicted. Finally, Sect. 6 contains a global summary and final conclusions.

2 The Analyzed Data Set

The data were gathered in the Department and Clinic of Internal Medicine and Allergology, Wroclaw Medical University. We consider a multivariate data sample containing patients being either in the normal or the abnormal state of their ventilatory system. The abnormal state was diagnosed as *obturation* or *restriction* in ventilatory disorders. In the elaborated sample counting $n = 77$ patients we have:

- Group 1, $n_1 = 28$ patients with obturative disorders,
- Group 2, $n_2 = 21$ patients with restrictive disorders,
- Group 3, $n_2 = 28$ patients with normal functioning, serving as control.

Each patient has its record (data vector) composed of 15 values: Patient's ID no., residual value (RV), total lung capacity (TLC), weight, height and 10 various spirometric variables derived from the patient's spirogram. The data were gathered with the aim to predict the RV and TLC variables (they need measurements using a different device) just from the spirometric variables considered as predictors [3,11]. In the following we will use from the these records only 12 variables defined in Table 1.

The twelve variables displayed in Table 1 will be hereafter called 'spirometric variables'. The first two of them are weight and height, the remaining ten – as obtained from a spirogram – are *sensu stricto* spirometric variables.

The recorded data are memorized as the $n \times m$ matrix $\mathbf{X}$ (with $n = 77$ and $m = 12$) containing in its rows the patients and in its columns the values of the variables characterizing the patients. The rows of $\mathbf{X}$ contain firstly the $n_1 = 28$ patients of group 1 (obturation), next the $n_2 = 21$ patients from group 2 (restriction), and finally the $n_3 = 28$ patients from group 3 (normal, control).

Table 1. Labels and definitions of the variables used in our analysis

X1:	Age [years],
X2:	Height [cm],
X3:	FVC, Forced Vital Capacity $[cm^3]$,
X4:	$FVC\%$ evaluated as $FVC/FVC_{predicted} \times 100$ [%],
X5:	FEV_1, Forced Expiratory Volume in one second $[cm^3]$,
X6:	Ratio $FEV_1/FVC \times 100$ [%],
X7:	$FEF_{0.2-1.2}$, Forced Expiratory Flow at the level of $0.2 - 1.2\,dm^3$ of FVC $[dm^3/min]$,
X8:	$MMFR$, Maximal Mid–expiratory Flow Rate $[dm^3/min]$,
X9:	$MMFT$, Maximal Mid–expiratory Flow Time [sec],
X10:	Ratio $MMFR/MMFT$ $[dm^3/min^2]$,
X11:	Ratio $FEF_{0.2-1.2}/FVC$ calculated as $1000\times$ X7/X3 $[min^{-1}]$,
X12:	Ratio $FVC/MMFT$ $[cm^3/sec]$

All elements of the data matrix **X** were strictly positive, however its columns were expressed in different scales. To annihilate the effects of largely differentiated scales of variables, the columns of the data matrix were standardized by 'range' to take values from the [0,1] interval. The data matrix with standardized elements will be referred to as **Xs**.

A holistic view of a not very big data matrix with non-negative elements may be obtained by displaying its values in the form of a heatmap. As such, the data matrix is considered as an image in the scale of an assumed color palette defined in computer graphics as colormap. Figure 1 shows such a heatmap of the transposed matrix $(\mathbf{Xs})^T$ of our data, when using the colormap *JET* (the default in Matlab).

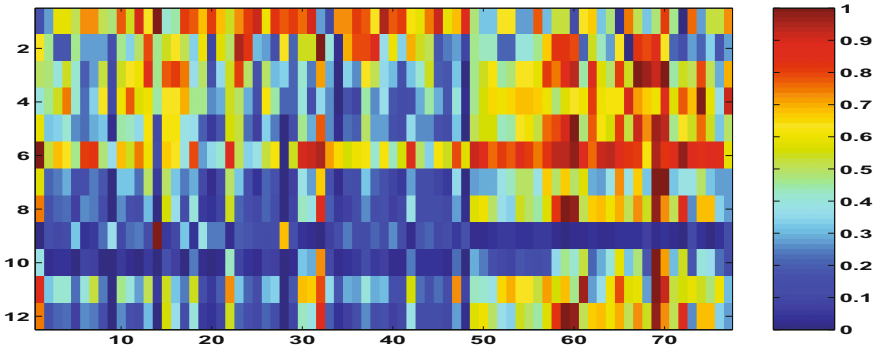

Fig. 1. Heatmap for $(\mathbf{Xs})^T$, the transposed standardized data matrix

The vertical constituents (columns) of the image are 77 columns numbered $1, 2, \ldots, 77$. They correspond to the subsequent patients constituting our data set. The no.s in the x-axis denote patients. Each column j visualizes the 12-dimensional data vector recorded for the j-th patient. In particular, no.s 1–28 indicate patients from group 1 (obturation), no.s 29–49 from group 2 (restriction), and no.s 50–77 from group 3 (normal, control)

Looking at Fig. 1 one may state that the patient group 3 (no.s 50–77) is clearly different as the groups 1 and 2 with no.s 1–49. The conclusion is: the group 3 of patients differs much in its inner structure from the the remaining two groups which look rather similar. Therefore, the 'normal' group should be rather easily identifiable in the data, while for the disease groups (obturation and restriction) this is doubtful.

3 The NMF Method

The NMF (non-negative matrix approximation) method was launched in [9] with the idea to find inner (hidden, unobservable) structure of the data which could be parameterized by lower rank factor matrices permitting to reconstruct

the observed data matrix $\mathbf{V}$. It was assumed that the approximation (*alias* reconstruction) will be of the form $\mathbf{V} \approx \mathbf{WH}$, with $\mathbf{W}$ and $\mathbf{H}$ denoting the inner structure factors found using an appropriate goodness of fit criterion. The methodology proved to yield very interesting applications and was subsequently developed using other variants of approach (see, e.g., [4,14–17]) and exploiting its properties (see, e.g., [1,2,10,12]). In the following we use matrix denotation $\mathbf{V}, \mathbf{W}, \mathbf{H}$ introduced in [9].

Suppose, we have data observed for n subjects characterized by m variables denoting attributes of the subjects. If we are more interested in the subjects (in their relationship, clustering) then we should put them in the columns of the matrix $\mathbf{V}$ to be dealt with by NMF. In such a case the model NMF is given as

$$\mathbf{V}_{m\times n} \approx \mathbf{W}_{m\times r}\mathbf{H}_{r\times n}, \tag{1}$$

where r is usually much smaller as $min(n, m)$ and is called *rank* of approximation.

The column vectors in $\mathbf{W}$ are called *base vectors*, they play the role of a kind of prototypes (representatives) of the subjects included into the data. If the data set contains r groups (clusters), then the derived base vectors are expected to be representatives of these groups [5].

The column vectors in $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_n]$ are called *encoders* or *coefficient vectors*, they play a crucial role in reconstructing the matrix $\mathbf{V}$ from the base vectors $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_r]$.

We will carry out the NMF approximation of the observed (visible) data matrix $\mathbf{V}_{12\times 77}$ by rank $r = 2$ and $r = 3$ factors. Computations will be done using the Matlab function snmf available at http://mikkelschmidt.dk/code.html. The criterium of goodness of fit of the model is the *explained variance* defined as follows

$$explained\ variance = (sst - sse)/sst, \tag{2}$$

where *sst* denotes the total sum of squares of deviations of elements v_{ij} from their overall mean, and *sse* is the sum of squared terms of errors $e_{ij} = v_{ij} - (\mathbf{WH})_{ij}$.

We start computing the NMF factorization for rank $r = 2$, obtaining the approximation $\mathbf{V}_{12\times 77} \approx \mathbf{W}_{12\times 2}\mathbf{H}_{2\times 77}$. In such a case there are only two base vectors $\mathbf{w}_1, \mathbf{w}_2$. Simple linear algebra shows, that these two vectors, combined in various proportions (each proportions taken as one column vector from the set $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_n]$), permit to reconstruct in sequence the subjects vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ contained in $\mathbf{V}$. For example, the 1st and 2nd subjects from $\mathbf{V}$, that is the subject vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are reconstructed as

$$\mathbf{v}_1 = \mathbf{w}_1 h_{11} + \mathbf{w}_2 h_{21}, \quad \mathbf{v}_2 = \mathbf{w}_1 h_{12} + \mathbf{w}_2 h_{22}, \quad \ldots \tag{3}$$

and so on. The elements of the coefficient vector $\mathbf{h}_j$ say how important are subsequent base vectors $w_1, \ldots, w_r$ in reproducing just the j-th subject vector. This may be seen globally when inspecting the row profiles of $\mathbf{H}$ depicted in Fig. 2 (for $r = 2$) and in Fig. 3 (for $r = 2$).

We formulate the following *Principle of group assignment in NMF*:
Find for given j, $(j = 1, 2, \ldots n)$, which of the coefficients h_{ij}, $(i = 1, \ldots, r)$ is

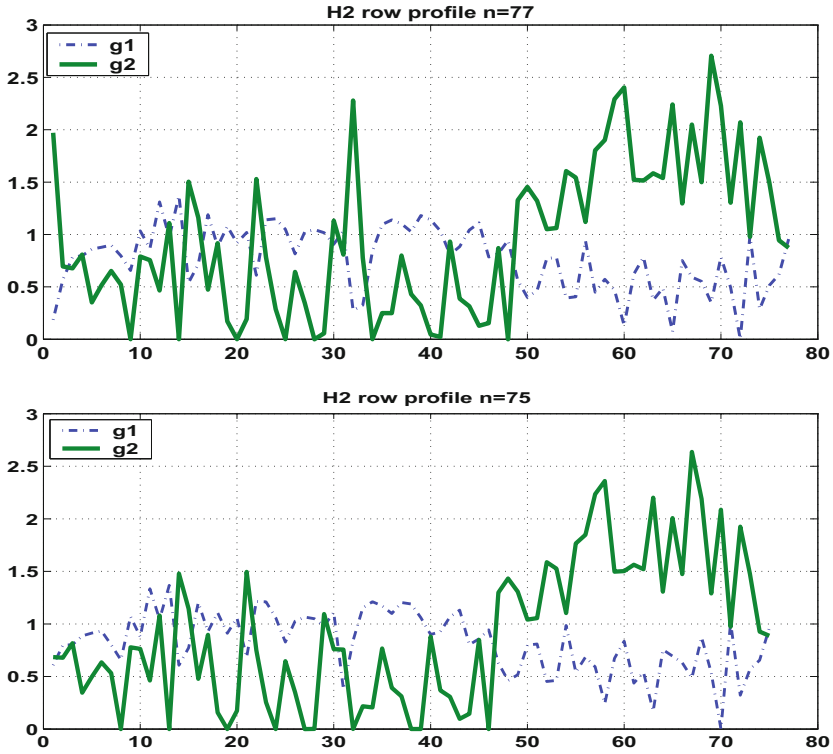

Fig. 2. NMF for $r = 2$. Top: Profiles of the coefficient matrix $H_{2\times 77}$ obtained when using the entire data matrix $\mathbf{V}$. Notice the outstanding green pics for *no.* 1 and *no.* 32. Bottom: Profiles of the coefficient matrix $H_{2\times 75}$ obtained when using the data matrix $\mathbf{V}$ from which the outlying vectors *no.* 1 and *no.* 32 were removed. Notice in both exhibits the discrepancy of the green and blue curves for the last 28 indices in the x-axis indicating patients from the control group. (Color figure online)

the largest. Say this is the i_{max} coefficient. Then assign the j-th data vector to the group no. i_{max}.

Figure 2, top exhibit, shows the row profiles plot of the coefficient matrix $\mathbf{H}$ obtained for $r = 2$. We see the values h_{1j} and h_{2j}, for the indices $j = 1, 2, \ldots, 77$, connected into two curves colored in blue and green respectively. The blue and green curves show - for the index j - the prevalence of the subject coefficient $\mathbf{h}_j$ to the first or second basic vector. This prevalence means that the respective base vector has a larger contribution in reconstructing the object no. j. In such a way the subjects are assigned to one of the two groups represented by the two base vectors.

Looking at the top exhibit in Fig. 2 we see there a particular intertwining of the two curves. For $j = 1$ to 49 (patients with obturation or restriction) the two curves are intermingled with a slight prevalence of the blue curve. Moreover, for $j = 1$ and $j = 32$ there are two big eruptions of the green curve. For $j = 50$ to

77 (patients from the normal group) the green curve connected with the base vector $\mathbf{w}_2$ is decidedly dominating.

The analysis was repeated with the data set reduced to 75 patients (patients *no.* 1 and *no.* 32 notified as outliers were removed). Results of the new analysis, without the two outliers, are shown in bottom exhibit of Fig. 2 and look more consistent. The essential pattern noticed in the top panel is present also here: the control group is decidedly connected with the base vector $\mathbf{w}_2$; the obturation and restriction groups show the same intermingled pattern with a mild preference of the base vector $\mathbf{w}_1$.

When looking at Fig. 2 one may come to the following conclusion: Similarly as in the heatmap, one sees here clearly a subdivision of the entire data into two

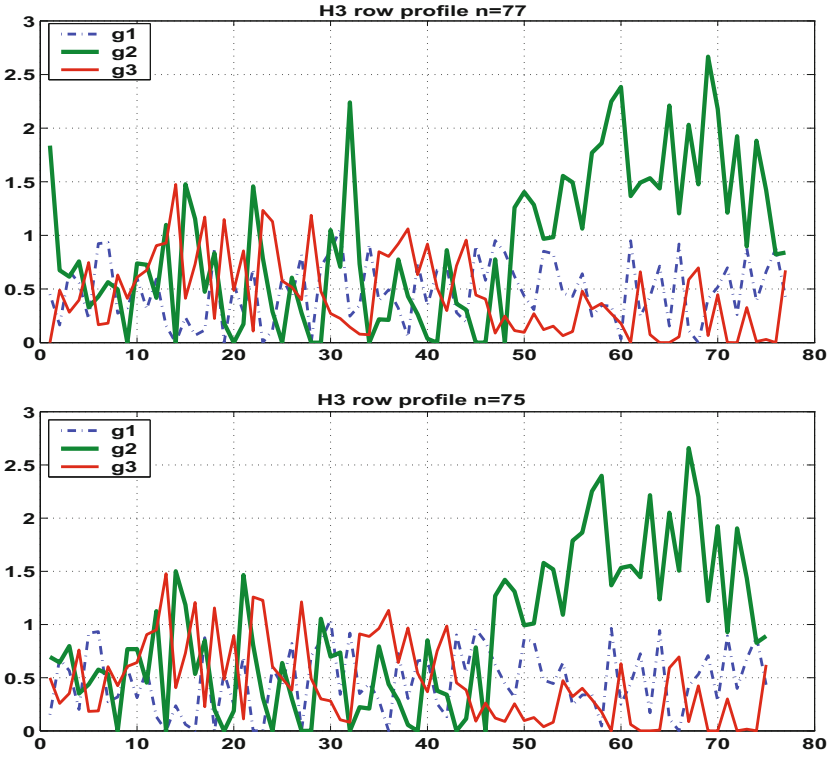

Fig. 3. NMF for rank $r = 3$. Top: Profiles of the coefficient matrix $\mathbf{H}_{3\times 77}$ obtained when using the entire data matrix $\mathbf{V}$. Bottom: Profiles of the coefficient matrix $\mathbf{H}_{3\times 75}$ obtained when using the data matrix $\mathbf{V}$ from which the subjects *no.* 1 and *no.* 32 were removed. Notice the big discrepancy of the red curve from both the green and the blue curve when observed for the last 28 indices in the x-axis – the red curve indicating the control group - is for no.s starting from *no.* 49 decidedly dominating. Notice also the appearance of two new pics of the red curve for no.s covering the first group (obturation). (Color figure online)

groups: the 'normal' state group (no.s 50 through 77) and the abnormal group designated by no.s 1–49. The group assignment for the normal group is good: 26 out from 28 patients are correctly located. Thus the first kind error is small. However the second kind error is considerable: several patients from the disease group are assigned to the normal, i.e. healthy group. Removing some putative outliers reduces the erroneous assignments in the disease group.

The analysis of the results obtained when assuming 3 groups of data (see Fig. 3) shows clearly that the assumption on 3 groups of data cannot be sustained.

Concerning the efficacy of the approximation measured by the fraction of explained variance (2): The models using $r = 2$ and $r = 3$ explain 79.18% and 87.43% of total variance. Model with $r = 6$ yields 97.46% of explained variance.

4 Clusters by Classic k-Means

For comparison with the above results by NMF, search for $k = 2$ and $k = 3$ clusters using the classical k-means algorithm was carried out. Below the group classification indices obtained from this method. The calculations were performed twice:

- For the full data matrix **Xs** size 77×12,
- For the reduced data matrix **Zs** size 77×12 (without two outliers discovered by the NMF).

The results are shown in Table 2 (for $k = 2$) and Table 3 (for $k = 3$) below. Notice that the k-means algorithm has established its own cluster denotations which is different from our group denotation linked with the diagnostic of obturative ("O") or restrictive ("R") ventilatory disorders, or normal ("N") state of the patient.

Working with $k = 2$ clusters the k-means yielded reasonable results similar to those obtained by NMF: the control group was diagnosed with 5 and 2 false assignments. They appear in Table 2 under the heading 'frequencies'.

Table 2. Results of k-means for $k = 2$

Group ↓	Subdividing data Xs into k = 2 subgroups	Frequencies
	k-means indicators	
g1 "O" n = 28	12222222222212112222221222222	1(5), 2(23)
g2 "R" n = 21	212122222222222222221	1(3), 2(18)
g3 "N" n = 28	11221111111111111111111121122	1(23), 2(5)
	Zs, i.e. Xs with two outliers removed	
g1 "O" n = 27	22222222222121122222212222222	1(4), 2(23)
g2 "R" n = 20	21222222222222222221	1(2), 2(18)
g3 "N" n = 28	11111111111111111111111122	1(26), 2(2)

Table 3. Results of k-means for $k = 3$

Group ↓	Subdividing data Xs into k = 3 subgroups	Frequencies
	k-means indicators	
g1 "O" n = 28	12223222322223122233312332 33	1(3), 2(15) 3(10)
g2 "R" n = 21	322123332333323333232	1(1), 2(7) 3(13)
g3 "N" n = 28	12221121111111111211112121122	1(19), 2(9) 3(0)
	Zs, i.e. Xs with two outliers removed	
g1 "O" n = 27	33323332333321333222132232 2	1(2) 2(10) 3(15)
g2 "R" n = 20	23332223222232222323	1(0) 2(13) 3(7)
g3 "N" n = 28	13331131111111111311113131133	1(19) 2(0) 3(9)

For $k = 3$ clusters the results from k-means were much worse as those by NMF: the control group was diagnosed with 9 false assignments both in **Xs** and **Zs**. The assignment to groups 2 and 3 was a mess of '2' and '3' assignments.

5 Visualization of the Data Using Kohonen's SOM

Now we visualize the data matrix **Xs** using Kohonen's SOMs. The SOM (self-organizing map) is *de facto* a neural network which reflects specifically the positions of data points in the multi-dimensional data space with preserving the data neighborhood topology. The principles of the method and the tricky algorithm are described in [8]. Maps obtained for our data - obtained using the free Matlab SOM toolbox [13] - are shown in Fig. 4. One may notice there that the maps have rectangular shape and are composed from $M = 11 \times 4 = 44$ equal size hexagons.

The SOM methodology performs a *vector quantization* and subdivides the data space into M so called Voronoi regions. The regions contain all the data points and reflect their density in the data space; some regions may be empty. There is a strict correspondence between the hexagons in the map and the Voronoi regions in the multivariate data space. We can find which data points are very near each other, and which are distant. Each Voronoi region has a representative data point (called codebook vector), which is in one-to-one correspondence to the center of hexagon associated with the respective Voronoi Region. There are two fitness indices: the quantization error *vqe* (means the average distance of data points in one Voronoi region to their representative codebook vector located in that region; for our data $vqe = 0.304$) and the topological error *tpe* (says, how many hexagons being neighboring hexagons in the map are not neighboring Voronoi regions in the data space; for our map we got $tpe = 0.0000$).

Now, where are our patient vectors from the categories "O", "R", and "N"? We found and identified all of them using the hit utility of the toolbox. The respective information (how many items from each group are represented by

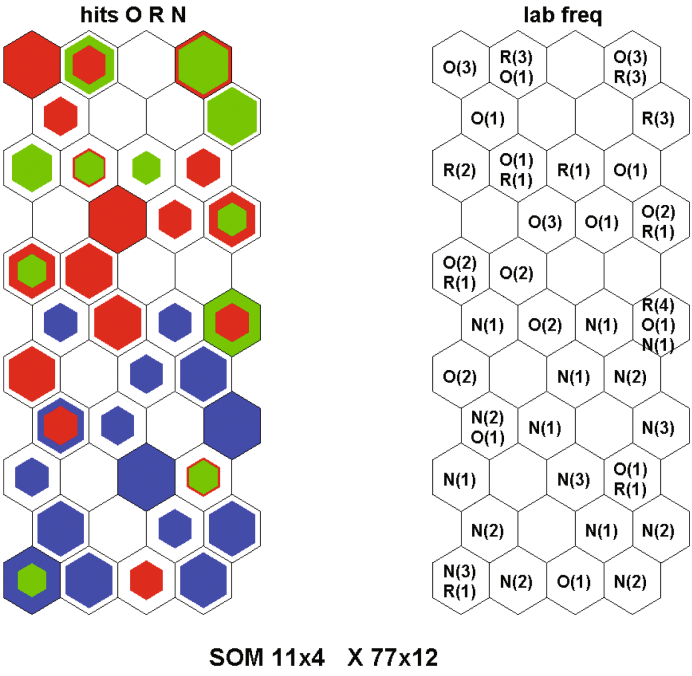

Fig. 4. Data groups assignment in Kohonen SOM. Letters O, R, N denote data points identified with patients belonging to the Obturation, Restriction or Normal group according the definition in Sect. 2. Colors red, green, blue are linked with groups "O", "R", "N" appropriately. Notice, that the "N"s appear in the southern part of the maps, while the "O"s and the "R"s reside in the northern part of the maps. (Color figure online)

each hexagon) is inscribed as ordinary text onto the map located in the right exhibit of Fig. 4. E.g., one may notice in the displayed map, that the utmost south-west hexagon (coordinates 11,1) contains three "Ns" and one "R" items (this "R" item happens to be the outlier no. 32 found by NMF and visible in Fig. 2, top exhibit).

Having the allocation of each data vector, one may display them group-wise, assigning to them specific colors and drawing in each map hexagon a smaller colored one, with radius proportional to the square root frequency of items linked to the given hexagon. The result is shown in the map appearing in Fig. 4, left exhibit. Looking at that exhibit, it is interesting to find that items of the 3 groups of our data are spread out over the entire map. Blue color (denoting "N" items) concentrates rather in the south of the map, however the region is not pure. In particular, there are two outliers at the south edge (no.s 1 and 32 notified previously and removed from part of our analyzes shown in Sect. 3). There is also one other "O" and one other "R" outlier located at the south-west edge. The 'disease' items (linked with red and green) appear intermixed in the upper

part of the same map, however some of them invade the blue territory of the "N" items. This observation is in agreement with the former comment on the results of NMF, that the results displayed in Figs. 2 and 3 are supporting the assumption on only $r = 2$ discernible clusters in the analyzed data.

6 General Summary and Concluding Remarks

The aim of the investigation was to explore the role of the NMF (non-negative matrix factorization) method when applied to real medical data. We have considered for this purpose a medical data set obtained from spirographic investigation of 3 groups of patients diagnosed as having two kinds of ventilatory disorders (g1, obturation, n = 28, and g2, restriction, n = 21) and additionally a control group (g3, normal, n = 28). Each patient was characterized by 12 variables, 10 of them being spirogram characteristics.

The NMF is a model based approach which can find in data some hidden factors (unobservable inner structure) permitting to reconstruct the entire data set with a hopefully small error. It was reported, that this inner structure might be connected with some groups into which the investigated data samples can be subdivided. In medical context this might mean subdivision into subgroups of various subtypes, e.g. various types of a disease. Our question was: Can the various types (groups of patients appearing in our data) be recognized – in a unsupervised way – by the NMF algorithm?

Our results. We have applied the NMF method seeking for k = 2 and k = 3 inner factors. The results from NMF were confronted with the *classic k-means* and two graphical methods: *Kohonen SOM* and *heatmaps* permitting to get a holistic vision of the 12-dimensional data vectors contained in the analyzed data set.

Conclusions from the investigation:

1. The NMF is able to recognize the normal group (92.86 % accuracy), however it can not recognize properly the abnormal state groups.
2. The NMF has an ability to recognize outliers hidden in the data.
3. Comparing to clustering abilities of the classic k-means, the NMF is superior.
4. The overall inspection of the results leads to the conclusion that the NMF has some abilities to recognize the group structure and can be considered as a 'weak learner', that - along with others weak learners - might be useful in *ensemble learning*.

References

1. Bartkowiak, A.M.: Classic and convex non-negative matrix visualization in clustering two benchmark data. Przeglad Elektrotechniczny **R93**(1), 53–59 (2017)
2. Bartkowiak, A.M., Zimroz, R.: NMF and PCA as applied to gearbox fault data. In: Jackowski, K., Burduk, R., Walkowiak, K., Woźniak, M., Yin, H. (eds.) IDEAL 2015. LNCS, vol. 9375, pp. 199–206. Springer, Cham (2015). doi:10.1007/978-3-319-24834-9_24

3. Bartkowiak, A., Liebhart, E.: Estimation of the spirometric residual volume (RV) by a regression built from Gower distances. Biometrical J. **37**(2), 131–149 (1995)
4. Cichocki, A., Zdunek, R., Phan, A.H., Amari, S.: Nonnegative Matrix and Tensor Factorizations. Applications to Exploratory Multi-way Data Analysis and Blind Source Separation. Wiley, Chichester (2009)
5. Ding, C., Li, T., Jordan, M.: Convex and semi-nonnegative matrix factorizations. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI) **32**, 45–55 (2010)
6. Jolliffe, I.T.: Principal Component Analysis, 2nd edn. Springer, New York (2002)
7. Kasim, A., Shkedy, Z., Kaiser, S., Hochreiter, S., Talloen, S. (eds.): Applied Biclustering Methods for Big and High-Dimensional Data Using R. CRC Press, Taylor & Francis Group, A Chapman & Hall Book, Boca Raton (2017)
8. Kohonen, T.: Self-Organizing Maps, Third Extended Edition. Springer, Heidelberg (2001)
9. Lee, D.D., Seung, H.S.: Learning the parts of objects by nonnegative matrix factorization. Nature **401**, 788–791 (1999)
10. Li, Y., Ngom, A.: The non-negative matrix factorization toolbox for biological data mining. BMC Source Code Biol. Med. **8**(10), 1–15 (2013)
11. Liebhart, J., Bartkowiak, A., Liebhart, E.: The impact of outliers in in the regression estimating TLC from age and some spirometric observations. Model. Simul. Control C **15**, 1–19 (1989). AMSE Press
12. Schmidt, M.N., Larsen, J., Hsiao, F.-T.: Wind noise reduction using non-negative sparse coding. In: IEEE International Workshop on Machine Learning for Signal Processing, (MLSP), pp. 431–436, August 2007
13. Vesanto, J., et al.: SOM Toolbox for Matlab 5, Som Toolbox Team, HUT, Finland. Libella Oy, Espoo, Version 0beta 2.0, pp. 1–54, November 2001
14. Zdunek, R.: Extraction of nonnegative features from multidimensional nonstationary signals. In: Tan, Y., Shi, Y. (eds.) DMBD 2016. LNCS, pp. 557–566. Springer, Heidelberg (2016)
15. Zdunek, R.: Convex nonnegative matrix factorization with Rank-1 update for clustering. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2015. LNCS, vol. 9120, pp. 59–68. Springer, Cham (2015). doi:10.1007/978-3-319-19369-4_6
16. Zdunek, R.: Data clustering with semi-binary nonnegative matrix factorization. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2008. LNCS, vol. 5097, pp. 705–716. Springer, Heidelberg (2008). doi:10.1007/978-3-540-69731-2_68
17. Zurada, J.M., Ensari, T., Asi, E.H., Chorowski, J.: Nonnegative matrix factorization and its application to pattern recognition and text mining. In: Proceedings of the 13th Federated Conference on Computer Science and Information Systems, Cracow, pp. 11–16 (2013)