Unsupervised feature learning for electronic nose data applied to Bacteria Identification in Blood

Martin Längkvist
Applied Autonomous Sensor Systems
School of Science and Technology
Örebro University
SE-701 82, Örebro, Sweden
martin.langkvist@oru.se

Amy Loutfi
Applied Autonomous Sensor Systems
School of Science and Technology
Örebro University
SE-701 82, Örebro, Sweden
amy.loutfi@oru.se

Abstract

Electronic nose (e-nose) data represents multivariate time-series from an array of chemical gas sensors exposed to a gas. This data is a new data set for use with deep learning methods, and is highly suitable since e-nose data is complex and difficult to interpret for human experts. Furthermore, this data set presents a number of interesting challenges for deep learning architectures per se. In this work we present a first study of e-nose data classification using deep learning when testing for the presence of bacteria in blood and agar solutions. We show in this study that deep learning outperforms hand-selected strategy based methods which has been previously tried with the same data set.

1 Introduction

Deep learning [1], [2], [3] has been applied to a variety of different data ranging from character recognition [1, 4], audio recognition [5, 6], motion capture [7], EEG event detection [8], object recognition [9] and activity recognition [10]. Unsupervised feature learning has a number of advantages, and is specifically attractive to data that is unintuitive and difficult to interpret. As we see increase in sensor technology and new sensors emerging in a variety of contexts, feature generation from data will become more and more relevant for such applications. In this paper, we present the first application of unsupervised feature learning to data which is collected from an electronic nose (e-nose). An e-nose is an instrument that consist of an array of gas sensors, typically 4-32 sensors with a pattern recognition software that quantifies and classifies the gas. The inclusion of multiple sensors with different properties increases gas selectivity but also introduce a high redundancy. These sensors have been used in experiments of food, beverage, and air quality monitoring, environmental monitoring, specific gas detection, and medical applications. Advantages of using an electronic nose include detection of difficult, odour-less, dangerous and/or unpleasant gases.

In this work, we focus on a medical application where the electronic nose is used to detect the presence of bacteria in two different media, namely blood and agar solution. The data collected can be describe as a multivariate time-series signal. Each sample represents one "‘sniff"’ where a sniff last approximately 2-3 minutes. Data is normally sampled at a frequency of 2 Hz. A snapshot of two sniffs coming from different bacteria can be seen in Figure 1. The tradition in the e-nose community has been to use hand-pick features. However, from previous work in the literature, where a plethora of features have been used, it can be concluded that an ideal feature set has not been found. Due to the high redundancy in the signals, dimensionality reduction techniques is performed before applying a classifier such as ANN or SVM. While this process has given some reasonable classification results it has been found that there is a lack of universal features across various applications in machine olfaction that give optimal performance. For this reason, this work advocates the use of unsupervised feature learning. As this is the first attempt to apply such methods to this kind of data, we use two types of RBM-based methods. The first is a DBN, which is suitable for this data as the data is not only highly redundant but may contain more complex structures than we are aware of. The DBN allows us to work directly on the raw data which is novel for the machine olfaction community. For comparison, a cRBM is also applied as it is able to capture temporal dependencies, which appear particularly in the dynamic phases of the signal (transient). Finally, both methods are compared to a traditional feature extraction solution on the same data sets. The contribution of this paper is two-fold: the first is an advocation of the use of unsupervised feature learning methods for this new type of data. Secondly, we provide a discussion on the short-coming of deep learning, and that the solutions to these short-comings could in fact be relevant for other types of data.

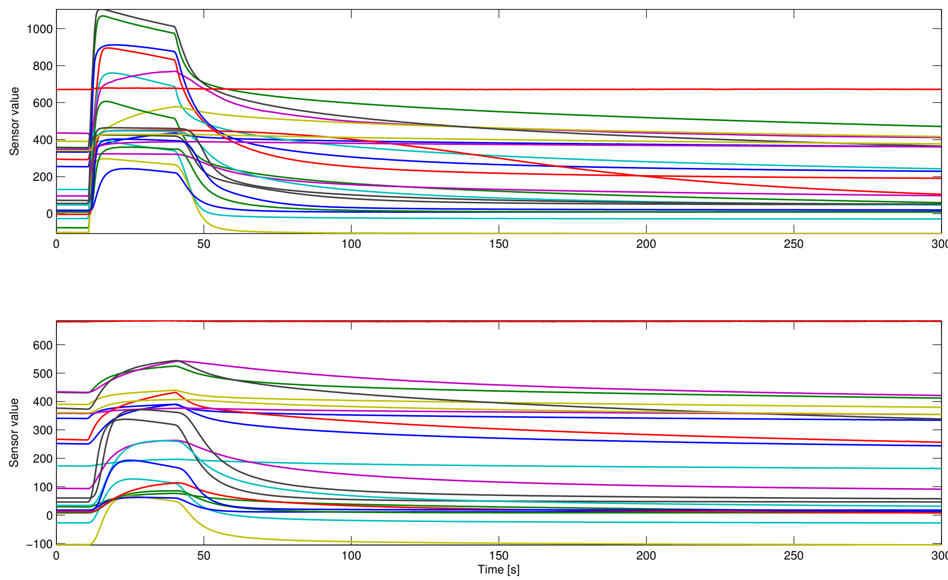


Figure 1: Raw sensor data from two bacteria in blood. The three phases are clearly visible with the baseline at 10 seconds and recovery phase at 40 seconds.

The paper begins with a short introduction to electronic noses and a description of the specific data set used in this work. Following is an outline of the algorithms used and the results achieved. All data sets and part of the code are available at `http://aass.oru.se/~mlt`

2 Electronic noses

The sense of smell is important for biological beings in order to localize food sources, detects hazards and in some species, finding a mate. Machine olfaction is a growing field that has emerged in order to quantify and objectively analyse odours in industrial applications. An electronic nose requires two components: sensors with varying selectivity, and a pattern recognition system [11], and therefore interests both sensor communities and pattern recognition communities.

The general principle behind the chemical gas sensor is based on the fact that analyte molecules come into contact with a chemically sensitive material which causes a change in the properties of the material, resulting in a change in the electrical signal. The most common material used for gas sensors is the tin-dioxide semiconducter, which is doped in order to provide selectivity. Examples of selectivity include selectivity towards volatile organic compounds (VOC), alcohols, sulphurs, carbon-based molecules e.t.c.

Typically the sensors are contained in an instrument that regulates the flow of air. Sampling is done in three phases: baseline phase, sampling phase, and recovery phase. The gas to be analysed is exposed to the sensor array in the sampling phase while a reference gas is used during baseline and recovery phase in order to return the sensor values to the initial state. Valuable information is not only obtained in the transient and static sensor values in the sampling phase but also the dynamic appearance in the recovery phase.

As previously mentioned, signal processing in the area of machine olfaction has traditionally been based on feature extraction using various properties of the signals such as maximum value, area under the curve, transient derivatives [12], Dynamic Time Warping [13] e.t.c. This step is followed by dimensionality reduction, either by PCA, feature selection and/or sensor selection. Finally, the reduced data is sent to a shallow supervised classifier. An overview of previous methods and features can be seen in [14].

2.1 Mednose data sets

The Mednose project aims to use an e-nose in order to discriminate between different type of bacteria that is typically found in blood and can lead to septisimia. Identifying bacteria in blood using an electronic nose has been done before with a 22 sensor array [12], as well as with a single sensor [13]. The two data sets used in this work are outlined below. In these data sets 1 of 10 different bacteria types is present in each sample.

Bacteria in blood The sampling system used for this data set is NST 3220 Emission Analyzer from Applied Sensors, Linköping, Sweden, which is composed of 10 MOS and 12 MOSFET sensors, for a total of 22 sensors. The odour sampling phases and recovery phase is 30 and 260 seconds long respectively. With a baseline phase of 10 seconds, the total length of one sampling cycle is 5 minutes. If one where to consider the sampling in Figure 1 as one sniff, this data set consists of a total of 800 sniffs. For more detailed description on how the samples were prepared see [12].

Bacteria in agar The device used for data collection in this data set is the Cyranose 320, which is a commercial generic e-nose system consisting of a sensor array of 32 conducting polymer sensors. The odour sampling phases and recovery phase is 20 and 80 seconds long, respectively. This data set contains a total of 740 sniffs. No publication with more detailed description for this data set is available.

Both data sets have the same sampling frequency of 2 Hz. The complete list of bacteria with the number of samples is given in Table 1.

3 Methods

We demonstrate the use of two deep learning architectures, namely DBN and cRBM.

A Restricted Boltzmann Machine (RBM) has visible units, **v**, with bias vector, **c**, and hidden units, **h**, with bias vector, **b**. A weight matrix, **W**, connects the hidden and visible units. For binary visible and hidden units (Bernoulli-Bernoulli), the probability that hidden unit h_j is activated given visible vector, **v** and the probability that visible unit v_i is activated given hidden vector, **h** are given by

$$P(h_j|\mathbf{v}) = \sigma(b_j + \sum_i W_{ij} v_i)$$
$$P(v_i|\mathbf{h}) = \sigma(c_i + \sum_j W_{ij} h_j)$$

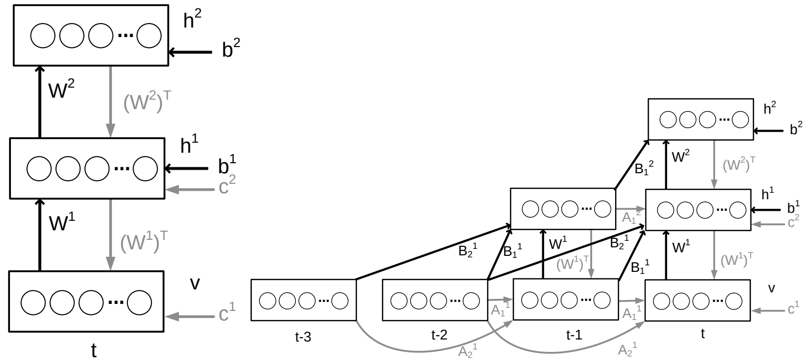
where $\sigma()$ is the sigmoid activation function. The parameters **W**, **b**, and **c**, are trained with contrastive divergence [15], which, in a similar fashion as auto-encoders, trains the model under the constraint to minimize the reconstruction error. For our data set the input vector to a DBN **v** is given by

$$\mathbf{v} = [S_{1,1}, \ldots, S_{w,1}, \ldots, S_{i,s}, \ldots, S_{w,s}] \quad (1)$$

where $S_{w,s}$ is the reading from sensor s from 1 to window size w for one sniff.

For the cRBM the input vector is given by concatenating sensor values across all sensors by the following:

$$\mathbf{v}(t) = [S_{t,1}, \ldots, S_{t,s}] \quad (2)$$



(a) (b)

Figure 2: Graphical depiction of (a) deep RBM and (b) deep cRBM with model order 2 in first layer and model order 1 in second layer.

The model for a Conditional Restricted Boltzmann Machine (cRBM) looks very similar to a RBM, except visible units are dependent on previous visible units, and hidden units are dependent on previous hidden as well as visible units. The probabilities for going up or down a layer is now

$$P(h_j|\mathbf{v}) = \sigma(b_j + \sum_i W_{ij} v_i + \sum_k \sum_i B_{ijk} v_i(t-k))$$
$$P(v_i|\mathbf{h}) = \sigma(c_i + \sum_j W_{ij} h_j) + \sum_k \sum_i A_{ijk} v_i(t-k)$$

The parameters **W**, **b**, **c**, **A**, and **B**, are trained using contrastive divergence.

A deep network of the two models can be formed by stacking them on top of each other where the lower-level hidden layer becomes the visible layer for the layer above, see Figure 2. Classification is achieved by attaching a single set of "‘softmax"’ units to the top hidden layer. The probability that predicted class y belongs to class j given input vector x and weight matrix θ is given by

$$P(y_i = j|x_i;\theta) = \frac{\exp^{\theta_j^T x_i}}{\sum_{l=1}^{k} \exp^{\theta_l^T x_i}}$$

The weight matrix θ is trained by minimizing the cost function

$$\begin{aligned} J(\theta) &= -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{k} 1\left\{y_i = j\right\}\log P(y_i = j)\right] + \\ &\quad + \frac{\lambda}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\theta_{ij}^2 \end{aligned}$$

where $1\left\{\cdot\right\}$ is the indicator function. The classes label j is one of the possible 10 bacteria that can be identified.

With exception of the top layer, each layer is only suboptimal, therefore, a fine tuning step is performed [16] for both methods. Training for the DBN is done by first (1) unsupervised pre-training of each layer, then (2) unsupervised fine-tuning of all layers with backpropagation, and finally (3) supervised fine-tuning of all layers with backpropagation. No experiments with fine-tuning a DBN with CD was done. Training for the deep cRBM is done by first (1) unsupervised pre-training of each layer, and then (2) supervised fine-tuning of all layers with backpropagation. We noted that fine-tuning with CD for cRBM did not improve our classification accuracy. In our implementation of supervised backpropagation for cRBM, we only update the bottom-up weights (W, b and B), which lower the reconstruction capabilities. However, since the task is to do classification and not reconstruction tasks, e.g., generating casual input data or noise reduction, we do not need the reconstruction parameters (A and c) once initial greedy layer-wise training is done.

Each sensor signal was scaled to have values between 0 and 1. This was done by subtracting each sensor signal with a baseline approximation and then dividing by the maximum value over all subtracted training examples. This also eliminates any possible sensor drift that might have occurred. The only form of dimensionality reduction is down sampling by a factor of 2, not by any feature or sensor selection.

4 Experimental Results

Table 1 lists the 10 bacteria that was inserted in the different media of blood and agar and the number of training, testing and validation samples for each class. Each class was randomly divided into training and testing sets. The number of samples increased when using a moving window. The same set of softmax units are connected to the units of the last hidden layer for all window positions.

Table 1: Bacteria names and number of train, test and validation samples

	Bacteria in blood			Bacteria in agar		
Bacteria	Train	Test	Validation	Train	Test	Validation
E.coli (ECOLI)	62	6	12	56	10	8
Pseudomonas aeruginosa (PSAER)	66	8	6	60	6	8
Staphylococcus aureus (STA)	64	5	11	63	7	5
Klebsiella oxytoca (KLOXY)	64	10	6	61	7	7
Proteus mirabilis (PRMIR)	66	7	7	60	9	9
Enterococcus faecalis (SRFCL)	63	9	8	64	4	8
Staphylococcus lugdunensis (STLUG)	63	10	7	57	9	8
Pasteurella multocida (PASMU)	63	8	9	57	7	10
Steptococcus pyogenes (HSA)	64	6	10	55	6	4
Hemophilus influenzae (HINFL)	65	11	4	59	9	7
Total	640	80	80	592	74	74

Three different experiments for both data sets were performed: classification using only raw data, a DBN, and a cRBM. Each experiment included tests of using either 1 and 2 layers as well as using either the full sample as input or only a 20 sample moving window as input. Step size of the window was set to 20 for the blood dataset and 5 for the agar dataset. An average over all individual windows were calculated to predict the class for the whole sample. Each layer had 200 hidden units and was trained with 200 epochs. Initial biases of the hidden units were set to -4 to encourage sparsity for the DBN [17].

The classification accuracy for all experiments can be seen in Table 2. Raw data was presented to the softmax classifier in the first two experiments, where it can be noted that comparable results with a feature-based method for Bacteria in blood is almost already achieved. A RBM with full sample as input did not perform well for either of the data sets. Increasing the number of levels improved the results only slightly. Using a 20 sample moving average window on the bacteria in blood dataset did, however, give satisfactory results. No reasonable results with a RBM could be obtained with the bacteria in agar dataset. We believe this is due to a much shorter recovery time for the agar dataset, resulting in a difficulty in properly aligning the sensor data. However, using a cRBM on the agar dataset gave much better results. In particular when a smaller model order was used. The number of levels for cRBM did not seem to have a significant effect on the classification accuracy for either of the data sets.

5 Discussion

In this work we have shown that unsupervised feature learning can be applied to electronic nose data, thus removing the task of designing hand-made features. We see the need to look into methods that are less influenced by human engineers in order to discover relevant patterns in complex data such as e-nose data. The e-nose community is a growing one with a number of emerging applications. However, for the data from an e-nose to be interpreted in a consistent manner, it is not tractable for each application to have its own hand-picked features for classification. A generalized

Table 2: Classification accuracy for full sample window

Setup	Bacteria in blood	Bacteria in agar
Features + SVM [12]	93.7	84.0
Raw data (full sample window)	80.0	52.7
Raw data (20 sample window)	84.1	39.2
RBM (1-layer, full sample window)	38.8	36.5
RBM (2-layer, full sample window)	43.8	44.6
RBM (1-layer, 20 sample window)	93.8	41.9
RBM (2-layer, 20 sample window)	**96.2**	47.3
cRBM (1-layer, model order 5)	85.0	**96.1**
cRBM (2-layer, model order 10-10)	75.0	90.0
cRBM (2-layer, model order 5-5)	85.0	96.0

method is required and unsupervised feature learning provides a framework towards this direction. There are some aspect about e-nose data that are not addressed within the methods that was used here. Namely, the high-level knowledge about which phase of the e-nose signal is most informative might suggest that a weighted window approach is more suitable, preferably with a separately trained model and classifier for each window. Further, the type of fine-tuning that is required to run is still not clear. It is especially difficult for e-nose data to verify what the model has learned by examining the reconstruction or visualizing the model parameters, as e-nose data is inherently difficult to understand. Finally, we see that there is an advocacy for non-RBM based approaches, such as stacked auto-encoders, as they seem easier to train. However, RBM-based methods are particularly suited for time-series multivariate data. Future work will focus on further examination of the suitability of RBM techniques for processing e-nose data, specifically continuous e-nose data that is done in open sampling systems where a three-phase sampling is not necessarily present, and when the gas character changes throughout the signal.

Acknowledgements

The authors would like to thank Lena Barkman, Bo Söderquist, and Per Thunberg at Örebro University Hospital for their assistance in the sample preparation, as well as Marco Trincavelli for assisting in data extraction. We also thanks Silvia Coradeschi for her support and suggestions. And finally, a special thanks to D F Wulsin for sharing his DBN implementation[1], as well as Graham Taylor for sharing his cRBM implementation[2].

This work was funded by NovaMedTech.

References

[1] G. E. Hinton, O. S., T. Y., A fast learning algorithm for deep belief nets, Neural Computation 18 (2006) 1527–1554.

[2] Y. Bengio, P. Lamblin, D. Popovici, H. Larochelle, Greedy layer-wise training of deep networks, Advances in Neural Information Processing Systems 19 (NIPS 2006) (2006) pp. 153–160.

[3] M. Ranzato, C. Poultney, S. Chopra, Y. LeCun, Efficient learning of sparse representations with an energy-based model, in: J. P. et al. (Ed.), Advances in Neural Information Processing Systems (NIPS 2006), MIT Press, 2006.

[4] A. Coates, B. Carpenter, C. Case, S. Satheesh, B. Suresh, T. Wang, A. Y. Ng, Text detection and character recognition in scene images with unsupervised feature learning, in: In Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR 2011), 2011.

[1] Download link can be found in [8]
[2] Project website can be found in [7]

[5] N. Jaitly, G. E. Hinton, Learning a better representation of speech sound waves using restricted boltzmann machines., in: ICASSP-2011, 2011.

[6] H. Lee, Y. Largman, P. Pham, A. Y. Ng, Unsupervised feature learning for audio classification using convolutional deep belief networks, in: NIPS, 2009.

[7] G. Taylor, G. E. Hinton, S. Roweis, Modeling human motion using binary latent variables, in: Advances in Neural Information Processing Systems, 2007.

[8] D. Wulsin, J. Gupta, R. Mani, J. Blanco, B. Litt, Modeling electroencephalography waveforms with semi-supervised deep belief nets: faster classification and anomaly measurement, Journal of Neural Engineering 8 (2011) 1741 – 2552.

[9] V. Nair, G. E. Hinton, 3d object recognition with deep belief nets, in: NIPS, 2006.

[10] Q. V. Le, W. Y. Zou, S. Y. Yeung, A. Y. Ng, Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis, in: Computer Vision and Pattern Recognition (CVPR), 2011.

[11] J. Gardner, P. Bartlett, Electronic Noses, Principles and Applications, Oxford University Press, New York, NY, USA., 1999.

[12] M. Trincavelli, S. Coradeschi, A. Loutfi, B. Söderquist, P. Thunberg, Direct identification of bacteria in blood culture samples using an electronic nose, IEEE Trans Biomedical Engineering 57 (Issue 12) (2010) 2884–2890.

[13] M. Bruins, A. Bos, P. L. Petit, K. Eadie, A. Rog, R. Bos, G. H. van Ramshorst, A. van Belkum, Device-independent, real-time identification of bacterial pathogens with a metal oxide-based olfactory sensor, Eur. J. Clin. Microbiol. Infect. Dis. 28 (2009) 775–780.

[14] R. Gutierrez-Osuna, Pattern analysis for machine olfaction: A review, IEEE Sensors Journal 2(3) (2002) 189–202.

[15] G. E. Hinton, Training products of experts by minimizing contrastive divergence, Neural Computation 14 (2002) 1771 – 1800.

[16] G. W. Taylor, Composable, distributed-state models for high-dimensional time series, Ph.D. thesis, Departmet of Computer Science University of Toronto (2009).

[17] G. E. Hinton, A Practical Guide to Training Restricted Boltzmann Machines (Aug 2010).