LETTER Communicated by Tim C. Pearce

Multilayer Perceptron Classification of Unknown Volatile Chemicals from the Firing Rates of Insect Olfactory Sensory Neurons and Its Application to Biosensor Design

Luqman R. Bachtiar
lbac004@aucklanduni.ac.nz
Charles P. Unsworth
c.unsworth@auckland.ac.nz
Department of Engineering Science, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

Richard D. Newcomb
Richard.Newcomb@plantandfood.co.nz
The New Zealand Institute for Plant and Food Research Limited, Private Bag 92169, Auckland 1142, New Zealand, and School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

Edmund J. Crampin
e.crampin@auckland.ac.nz
Department of Engineering Science and Auckland Bioengineering Institute, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

In this letter, we use the firing rates from an array of olfactory sensory neurons (OSNs) of the fruit fly, *Drosophila melanogaster*, to train an artificial neural network (ANN) to distinguish different chemical classes of volatile odorants. Bootstrapping is implemented for the optimized networks, providing an accurate estimate of a network's predicted values. Initially a simple linear predictor was used to assess the complexity of the data and was found to provide low prediction performance. A nonlinear ANN in the form of a single multilayer perceptron (MLP) was also used, providing a significant increase in prediction performance. The effect of the number of hidden layers and hidden neurons of the MLP was investigated and found to be effective in enhancing network performance with both a single and a double hidden layer investigated separately. A hybrid array of MLPs was investigated and compared against the single MLP architecture. The hybrid MLPs were found to classify all vectors of the validation set, presenting the highest degree of prediction accuracy. Adjustment of the number of hidden neurons was investigated, providing further performance gain. In addition, noise injection was investigated, proving successful for certain network designs. It was found that the best-performing MLP was that of the double-hidden-layer

hybrid MLP network without the use of noise injection. Furthermore, the level of performance was examined when different numbers of OSNs used were varied from the maximum of 24 to only 5 OSNs. Finally, the ideal OSNs were identified that optimized network performance. The results obtained from this study provide strong evidence of the usefulness of ANNs in the field of olfaction for the future realization of a signal processing back end for an artificial olfactory biosensor.

1 Introduction

In various industries, the human nose is still used as an instrument for evaluating the smell or flavor of products such as perfumes, foodstuffs, and beverages (Gardner & Bartlett, 1994). This process of olfactory testing is costly, and the labor output associated with trained experts is limited. Existing analytical methods for detecting and analyzing compounds, such as gas chromatography (Gardner & Bartlett, 1994) and gas chromatography-mass spectrometry (Haugen & Kvaal, 1998), were seen as inadequate in the past because they showed poor sensitivity and exhibit instability. However, advances have been made such that these methods have become widely used and more stable (Pearce, Schiffman, Nagle, & Gardner, 2006). Recently artificial biosensors have been developed in an attempt to emulate the workings of a biological olfactory system (Kermani, Schiffman, & Nagle 1999; Ko & Park, 2005; Lee, Jun, Ko, Kim, & Park, 2009) with considerable success (Lewis, 2004). In this letter, an artificial neural network (ANN) is used to process the data obtained from the sensory array of the proposed biosensor. A linear predictor is used to evaluate the nonlinearities present in the data and the suitability of using nonlinear classifiers for processing. A single multilayer perceptron (MLP) is the nonlinear ANN employed for the classification of the odorants of interest. Both single and double hidden-layer MLP architectures were used to extract higher-order information in the data. A hybrid MLP system, in which a defined number of MLPs is arranged in parallel, was used in contrast to the single MLP to promote network prediction. Optimal MLP length was investigated by incrementally altering the number of hidden neurons in the network. Additionally, noise injection was used to improve network training and prediction. Finally, in line with the aim of developing a biosensor, the number of OSNs that optimized network performance was identified.

1.1 The Olfactory System: *Drosophila melanogaster*.

Similar to humans, the olfactory system of insects is also capable of detecting and distinguishing a diverse array of odors (Hildebrand & Shepherd, 1997). The olfactory sensory neurons (OSNs) that detect odors are located in the antennae and maxillary palps of *Drosophila melanogaster*, as depicted in Figure 1. These olfactory organs are covered by sensilla (sensory hairs) that

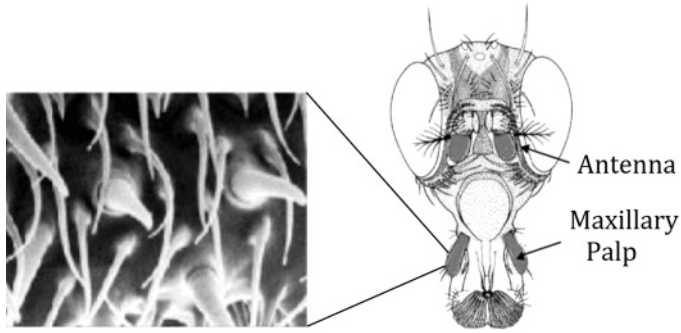


Figure 1: The antennae and maxillary palps of *Drosophila melanogaster*. The magnification of the maxillary palps shows how their surface is covered with hairlike sensilla. (Figure modified from Jefferis and Hummel, 2006, and Hallem and Carlson, 2004.)

house the OSNs. The sensilla of *D. melanogaster* contain one to four OSNs that respond to different odorants. A diverse range of sensilla contains distinct sets of OSNs. Tiny wax-filled pores in the walls of the sensilla allow the entry of odor molecules. Within the sensilla, high concentrations of odorant binding proteins solubilize and transport the odorants to the olfactory receptors expressed on the surface of the OSN dendrites. Electrophysiological techniques have been developed that record the firing rates of individual OSNs wherein it has been revealed that each OSN has a distinguishing spontaneous firing rate (Hallem & Carlson, 2006). Activation of the OSN by a relevant odorant exhibits a change in the firing rate, whether an increase or decrease in response to the binding of an odorant molecule to the olfactory receptors on the OSN. Further activation is observed in the second-order neurons located in the higher processing system or brain of the fly.

1.2 Classifying Odorants. The neuronal firing rates of olfactory neurons can in principle be analyzed to interpret and thus classify the presented odorant molecules. Numerous techniques are available for classification: statistical methods such as regression analysis, clustering methods, and multivariate data analysis (Wilson & Baietto, 2009). In this work, we employ an ANN for classification (Haugen & Kvaal, 1998). ANNs can be seen as a mathematical interpretation of biological neurons (Churchland & Sejnowski, 1992). The wires and interconnections of an artificial network model the axons and dendrites of a biological neural network, while the weighting functions of these connections represent strengths of synaptic connections (Jain, Mao, & Mohiuddin, 1996). The activation function approximates the electrical activity of the cell membrane potential upon experiencing an action potential, although this approximation is a simplification of the complex temporal properties of biological

neurons (Churchland & Sejnowski, 1992). However, although ANNs are a simplification of the workings of biological neurons, they are recognized as a viable and commonly used method in classification problems such as that of odorant prediction (Ferreira, Ludermir, & de Aquino, 2005; Fu, Li, Qin, & Freeman, 2007; Zhang, Balaban, & Principe, 2003).

The real value of ANNs lies in their unique pattern recognition ability. Namely, they are able to extract the essential features from a training data set and make use of this information to predict values. This is particularly useful for cases where there is a highly complex relationship between the data and the classes to which they belong. However, despite their unique advantages such as their generalization capabilities, arbitrary decision boundaries, and flexibility to different types of data, ANNs possess some inherent limits (Kavzoglu, 1999). These limitations can be separated into external factors, such as the characteristics and scale of the data available for training, while internal factors include the appropriate network structure, number of iterations, and transfer function employed (Kavzoglu, 1999). Recognizing these factors and deciding the appropriate parameter values is the key to successful implementation of the ANN method.

2 Motivation

We have performed preliminary studies with MLPs for olfactory recognition, using data from the firing rate of OSNs (Bachtiar, Unsworth, Newcomb, & Crampin, 2011b; Unsworth, Bachtiar, Newcomb, & Crampin, 2011) and data from chemical descriptor values (Bachtiar, Unsworth, Newcomb, & Crampin, 2011a). The motivation of this work is to provide support for the development of an olfactory biosensor, focusing on the classification performance of a network as changes are made to the structure and architecture of the network. We investigate in detail the MLPs that produce the best network performance from OSN firing rates.

3 Data

Through the use of their olfactory neurons, *D. melanogaster* are able to detect a diverse range of volatile compounds. Hallem and Carlson (2006) characterized the response of an array of olfactory receptor-expressing neurons to a standard set of odorants *in vivo*. A mutant strain that provided an empty sensory neuron was used to express each of these olfactory receptors. Spontaneous activity is quantified from the number of recorded spikes in 1 second of spontaneous activity (Hallem, Ho, & Carlson, 2004). Previous studies have reported spontaneous activity of odorant receptors ranging from approximately 1 spike/s to 30 spike/s (de Bruyne, Clyne, & Carlson, 1999; de Bruyne, Foster, & Carlson, 2001), dependent on the receptor that is expressed. Neuron firing rates for neurons expressing specific receptors were measured for a range of odorants. This work uses the quantitative

Table 1: Recorded Receptor Firing Rate Values in Response to Various Chemicals of Different Classes.

			Olfactory Receptor			
Chemical Class	Total Odorants	Chemical Odorant	Or2a (X_1)	Or7a (X_2)	Or88a (X_{23})	Or98a (X_{24})
Amines	**2**	**Putrescine**	**6**	**−36**	**−6**	**17**
Lactones	5	G-butyrolactone	9	34	2	29
Acids	15	Propionic Acid	1	5	−7	11
Sulfur compounds	**2**	**Dimethyl sulfide**	**0**	**−8**	**−6**	**−11**
Terpenes	16	Terpinolene	−6	6	6	81
Aldehydes	8	Butanal	43	54	−3	26
Ketones	6	Acetone	1	−10	−9	−1
Arormatics	13	Benyl alcohol	−4	5	1	1
Alcohols	17	Methanol	10	22	−8	5
Esters	24	Methyl acetate	5	−6	14	4

Notes: The "Chemical Odorant" column presents a member of the chemical class and the corresponding firing rates of four receptors on exposure to said chemical. The odorants in the amine and sulfur compound chemical classes, highlighted in bold, were not used in this study.

firing rate values recorded on exposure to the volatile chemical odorants where different receptors gave rise to different firing rates for the same odorant and different odorants produce different firing rates for the same expressed receptor.

Using transformations to normalize input data has been found to improve network performance (Garson, 1998; Sola & Sevilla, 1997). Normalization allows for a better-behaved training data set, which provides a more efficient training process in the aim of minimizing the network error. Prior to ANN testing, the data are preprocessed via statistical normalization, which involves subtracting the mean and dividing by the standard deviation as follows:

$$x_t = \frac{X_t - \mu}{\sigma},$$

where X_t is the raw value of the tth training case, μ the mean, μ the standard deviation, and x_t the calculated normalized value. The processed data has a zero mean and a standard deviation of 1; all values center about zero. Statistical normalization ensures that the initially large raw values now cluster in an appropriate area, further enhancing the learning algorithm implemented (Sola & Sevilla, 1997).

Table 1 represents an abridged version of the data used in this research. Initially odorants from a set of 10 chemical classes were available for processing: amines (2 odorant chemicals), lactones (5), acids (15), sulphur

compounds (2), terpenes (16), aldehydes (8), ketones (6), aromatics (13), alcohols (17), and esters (24). For the amine and sulfur compounds highlighted in bold, only two unique chemicals were available in the data set thus, to prevent the network from overtraining and overgeneralizing the data, it was deemed appropriate to remove these two chemical classes from our analysis (Foody, 1995). In conjunction with the firing rates of the 24 OSNs, the eight chemical classes that the 104 different chemical volatile compounds belonged to were used for the neural network analysis in this study.

Hallem and Carlson (2006) used a small subset of 10 odorants to observe the effects of varying odorant concentration. Only 10 unique chemicals, of the ester (5 odorant chemicals), alcohol (2), ketone (2) and aldehyde (1), were used for this experiment. As we stated earlier, such limited raw data may result in overtraining and overgeneralization by the network (Foody, 1995). Because this subset proved limited, the effect of odorant concentration on network prediction will not be analyzed. Instead, this work uses odorants that are of a constant 10^{-2} dilution, a concentration set deliberately high as to ensure an olfactory receptor response (Hallem & Carlson, 2006).

Of the 104 chemical compounds in the data set used in this research, approximately 10% of each chemical class was randomly selected to be included in the validation set, used to measure the degree of network learning. Due to the limited number of chemicals available in each class, only 10% of the chemicals from each class were used. Thus, the numbers of chemicals in each chemical class of the validation set were lactones (1 chemical odorant), acids (2), terpenes (2), aldehydes (1), ketones (2), aromatics (2), alcohols (2), and esters (3). The training set of the neural networks is composed of the 89 remaining chemicals.

Bootstrapping is a method in which training sets are sampled from the raw data set for network processing and then replaced by a different training set, in repeated fashion (Dreiseitl & Ohno-Machado, 2002; Efron & Tibshirani, 1993). As Tibshirani (1996) found, bootstrapping methods elicited the most accurate estimates of a network's predicted values and the corresponding standard errors, outperforming cross-validation methods (Efron, 1983; Harrell, 2001). Bootstrapping will be implemented on optimized networks of 10,000 different training and validation sets. The use of bootstrapping will be presented as a bar graph of the mean and standard error of the network prediction. A fixed validation set is used in cases that demand longer processing time, in which case bootstrapping will not be used. It should be noted that bootstrapping is not performed in the final biosensor design. Instead, a catalogue of known odorants is used by the biosensor to predict unknown odorants.

4 MLP Architecture

The main ANN presented in this letter is that of a simple feedforward MLP. This section documents the different MLP architectures investigated and

their subsequent prediction performance using olfactory sensory data. Each input vector represents the firing rate of the 24 olfactory neurons to a single chemical odorant. Thus, the input layer of the MLPs consisted of 24 neurons, which aligned with the 24 OSNs. The lengths of the following layers of hidden and output layers were altered and their effects investigated in the later sections.

The weights to the network were drawn from a symmetric gaussian distribution with a zero mean and variance of unity providing weight values between -1 and 1. These weight values were chosen to be small so as to optimize the weight decay regularizer, which has been shown to improve network generalization and avoid overfitting for an MLP, Bishop (1995). An example of the weights and their evolution for the optimal hybrid network is shown in Figures 7D and 7F for the different layers of the network. As Bishop (1995) suggested, the stopping criteria were chosen after an exhaustive series of preliminary runs that served to reduce computation CPU time, were for a number of fixed iterations, and stopped when the error of prediction fell below a certain prescribed limit. In our case, we found that 200 epochs provided a good CPU single run time of about 3 minutes, which provided a reasonable experimentation time of approximately 2 weeks when bootstrapping was employed. Our error prediction cut-off was set to 0.01, which was also found to occur around 200 epochs for all of the optimal network architectures, highlighted in Figure 7G for the optimal hybrid network. Under bootstrapping, the weighting functions are defined and fixed for the subsequent 10,000 different validation and training sets tested. The stopping criterion of network training is based on the number of epochs performed by the network, and 200 epochs are used. The outputs of the network are continuous, and the prediction percentage values are obtained by dividing the validation vector output by the absolute sum of the output signal as follows:

$$Ester^1\ Prediction(\%) = \frac{output\ \ signal_{Ester^1}}{\left|\sum_{i=1}^{n} output\ \ signal_{Class}\right|}.$$

Unless stated otherwise, the results of prediction represent the percentage accuracy for which the unseen chemicals were placed into their correct chemical class; a value closer to 100% indicates a high probability that the chemical has been classified in its correct class.

4.1 Linear Prediction of Chemical Class. A linear predictor is equivalent to a single artificial node with a linear activation function. This alternative to the ANN system is helpful to address the complexity of the data and relevance of the nonlinearities that appear for certain chemical classes. Figure 2 shows the prediction performance of the linear predictor across the validation set. Because the odorant prediction exhibited a wide

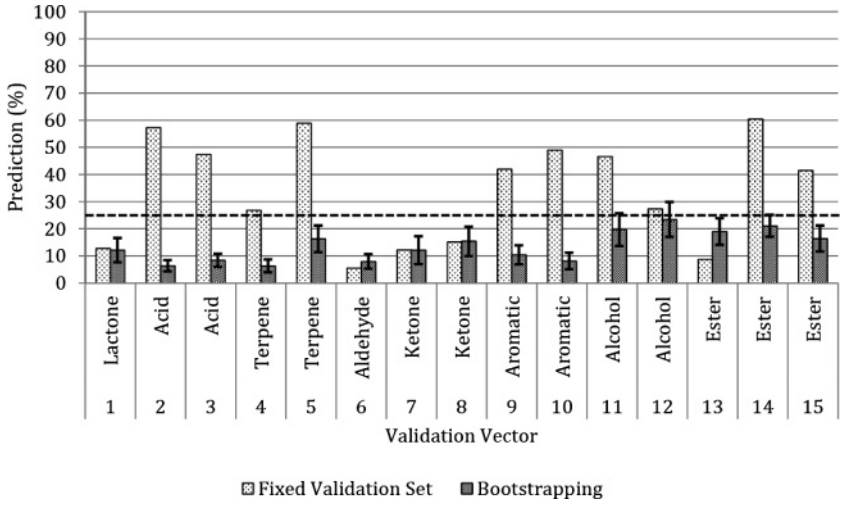


Figure 2: Performance of the linear predictor using the fixed validation set (left bar) and under bootstrapping (right bar). The horizontal dashed lines represent the 25% threshold value that defines correct classification of an odorant of the validation set.

range, a threshold value was defined to characterize the successful classification of odorants. The largest chemical class present in the data is the ester group; therefore, it is represented by the most chemicals in the validation set: out of the 15 validating odorants, 3 were known to be esters. The probability of choosing an ester randomly was 3/15, or 20%. This value of 20% was used to identify the minimum threshold level of detection of the odorant validation set. A margin of 5% was included with this value as an added safeguard. Thus, a threshold level of 25% accuracy was employed to the network prediction results. The purpose of the 25% threshold value, depicted by horizontal dashed lines in Figure 2, is that it provides a qualitative indicator to the number of validation vectors that are correctly classified. If the prediction percentage value of a validation vector exceeds the 25% threshold value, it is regarded as having been correctly classed by the network.

For the fixed validation set, the linear predictor presented 10 odorants that exceeded the threshold value, while 5 others failed. Acids, terpenes, aromatics, alcohols, and esters were predicted well; however the lactone, aldehyde, and ketone chemical class did not perform well. The failure of some validation vectors is due to the MLP classifying the vectors into an incorrect class. The known lactone vector was classed as an acid and an aldehyde as an acid; ketones were recognized as alcohols and an ester as an alcohol. This could be due to the limited data available. A larger training set

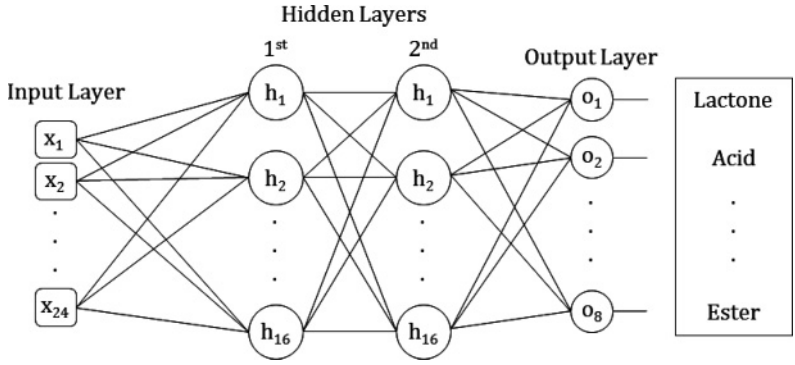


Figure 3: Schematic of a double-hidden-layer MLP with a multineuron output layer

could potentially yield improvements for outliers. Bootstrapping resulted in no odorant mean values that surpassed, with a possible range of $0 \leq n$ odorants superseding threshold ≤ 3. These results suggest that the nature of the data does not favor a linear model. Thus, it would be intuitive to investigate a more complex nonlinear ANN to model the data, as described in the following section.

4.2 A Single-Layer MLP. An initial MLP architecture based on Blum (1992), with an input size of 24 neurons, 16 hidden neurons, and an output size of 8 neurons was obtained to create a 24-16-8 MLP structure. A standard backpropagation method under supervised learning with a hard-limiting binary activation function was used. A momentum function was included to enhance the gradient-descent algorithm, allowing a more stabilized network, preventing it from stagnating at local minima (Huang, Tan, & Tang, 2004). An epoch was used as a measure of network training, where a single epoch is defined as when one whole training set has passed through the network. Thus, the stopping criteria of network training are based on the number of epochs performed by the network; in this work, 200 epochs were to be found sufficient as a stopping criteria.

4.3 A Double-Layer MLP. One hidden layer has been found to be sufficient for the universal approximation property (Hornik, Stinchcombe, & White, 1989). However, an additional layer was introduced to the existing basic MLP to extract any higher-order relationships that were present in the data (De Villiers & Barnard, 1993; Longstaff & Cross, 1987). The same number of neurons as the first hidden layer was used to produce an MLP structure of 24-16-16-8, illustrated by the simplified model in Figure 3.

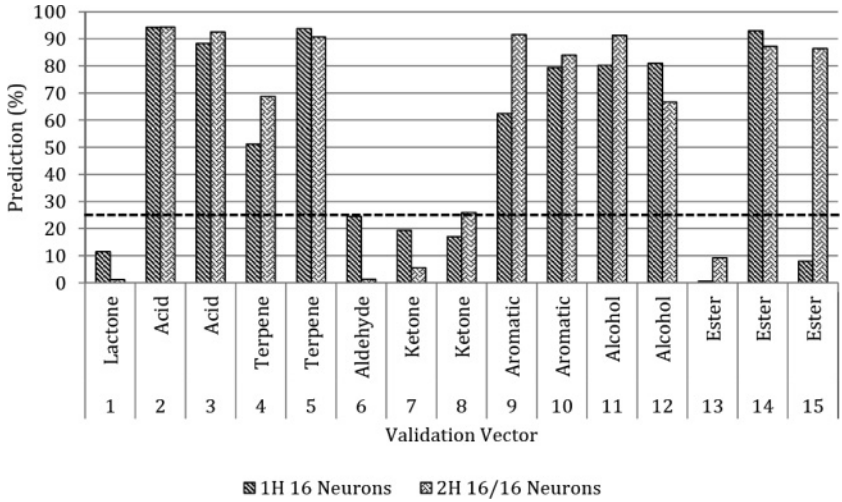


Figure 4: Performance of the MLP using a single (left bar) and double (right bar) hidden layer. The horizontal broken lines represent the 25% threshold value that defines correct classification of an odorant of the validation set.

As presented in Figure 4, the double-hidden-layer MLP classified 11 odorants compared to the 9 odorants of the single-hidden-layer MLP (where the left bar in each chemical class represents a single-layer MLP and the right bar a double-layer MLP). The most significant change is seen in odorant 15 of the ester class, which showed an 80% increase in prediction accuracy through the use of the double-hidden-layer MLP. The double-hidden-layer MLP system has proven to increase network prediction across the validation set using this initial architecture configuration. However, a number of odorants decreased in the accuracy of prediction, such as that of odorant 6, which shrank by nearly 30%, thus ruling it out of the successful classification. This is due to the MLP's classifying the odorant into the incorrect class. Similar to the linear predictor, the MLP system classed the known lactone vector as an acid and an aldehyde as an acid, while ketones were recognized as alcohols and an ester as an alcohol. A larger training set could potentially yield improvements for outliers. The spread in prediction values of the MLP system appears more precise, with a number of odorants showing high prediction values. On the other hand, results from the linear predictor as seen in section 4.1 are quite close to the defined threshold value. This result emphasizes the suitability of a nonlinear classifier over a linear classifier.

4.4 Incrementally Altering the Number of Neurons in the Hidden Layers. As network size affects network complexity, learning time, and the

generalization capabilities of the network, the quality of a solution thus relies heavily on the network size used (Baum & Haussler, 1989). The optimum number of hidden neurons producing the best network performance for the MLPs was found by varying the number of neurons of the hidden layers. With a defined input length of 24, corresponding to the olfactory receptor sensors, a maximum hidden-layer length of 48 neurons was established. This is because a hidden layer more than double that of the input layer has been found to be ineffective: it requires more time to train the network and more training samples are needed to achieve good generalization (Berry & Linoff, 2009; Churchland & Sejnowski, 1992; Kavzoglu, 1999; Swingler, 1996).

This test was performed using both a single and a double hidden layer. Assembling these series of tests proved laborious, as many MLPs were constructed to ensure all conditions were tested; 16,464 MLPs were used to determine the optimum architectures. Selection of the best structure was based on the highest accuracy of the odorant with the lowest prediction accuracy across the validation set. Bootstrapping was applied only to the optimized networks.

Based on the defined threshold, the single-hidden-layer design correctly classified the entire fixed validation set, while the double hidden layer correctly classified 14 of the 15 odorants. The lowest-performing vector of the single- and double-hidden-layer design presented an accuracy of 28% and 23%, respectively. As presented in Figure 5, both of these odorants bordered the defined threshold of 25%, which suggests that with further refinement, the double hidden layer also has the potential of correctly predicting the complete validation set.

With the implementation of bootstrapping, an odorant vector of the ketone chemical group failed to supersede the threshold value, Figure 5B. Aside from the failure of the ketone group, the performance of this optimal hidden-layer multiple-output MLP on average classified 14 odorant vectors with a range of $12 \leq n$ odorants, superseding threshold ≤ 15 for the single-hidden-layer structure, while the double hidden layer on average classed 14 odorant vectors with a range of $11 \leq n$ odorants superseding threshold ≤ 15. The single-hidden-layer structure of the multiple-output MLP outperformed the double-hidden-layer equivalent. However, the overall mean across the validation set after bootstrapping is in favor of the double-hidden-layer structure, with a mean of 58.9% over the single-hidden-layer structure output prediction of 57.5%. With the implementation of bootstrapping, the performance of these two structures is very close. Thus, the network architectures that will be tested in the following sections include analysis of both single and double hidden layers in the MLP system.

4.5 A Hybrid Network of MLPs. The MLPs constructed in previous sections had an output layer length of eight neurons, corresponding to the different chemical classes of the odorant data. This brings about additional

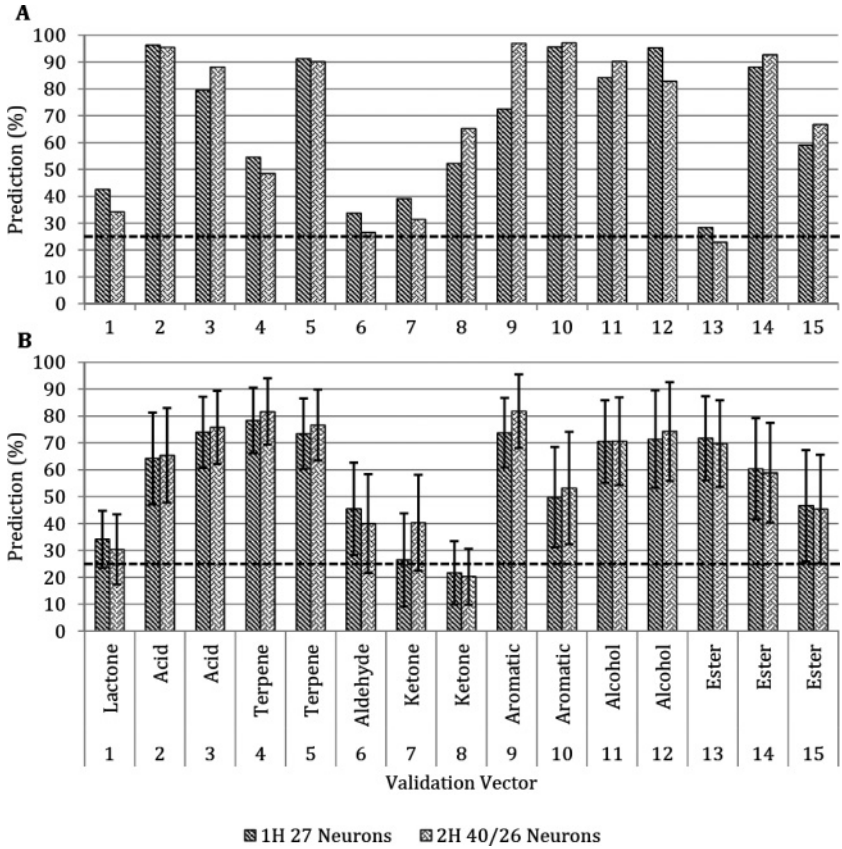


Figure 5: Performance of single MLP prediction using optimum neuron lengths for single (left bar) and double (right bar) hidden layers with (A) a fixed validation set and (B) implementation of bootstrapping. The horizontal dashed lines represent the 25% threshold value that defines correct classification of an odorant of the validation set.

weights between the hidden layers and the output layer of the MLP, which is somewhat undesirable as it escalates the complexity of the network, decreasing the network's rate of convergence (Kavzoglu, 1999). To overcome the complexity of a multiclass data set, we reduced the output layer to a single neuron and produced eight different MLPs, where each MLP corresponded to one of the chemical classes, as depicted in Figure 6. Each of the eight MLPs was trained to give a desired output close to 1 for its corresponding chemicals, while a desired output near 0 was put in place for all other chemical classes. Combining individual networks in parallel is sometimes referred to as a hybrid network (Gardner, Hines, & Wilkinson,

1990). This system is said to be more suitable for multiclass problems as the association of one class is not influenced by any decisions regarding other classes (Longstaff & Cross, 1987). As Longstaff and Cross found, this MLP structure is guaranteed to be able to represent all multiple-class problems, however complex.

As discussed in section 4.3, the double-hidden-layer system was found to enhance network performance. With the use of the hybrid network, it was of interest also to investigate the performance difference when a double hidden layer was used. Two hybrid network systems were constructed: 24-16-1 and 24-16-16-1. A simplified model of double-hidden-layer system is depicted in Figure 6.

Figure 7A presents the prediction performance of the hybrid network system of MLPs when using a single- and double-hidden-layer design. The single hidden-layer hybrid design showed an improvement with the classification of 11 odorants, an addition of 2 odorants compared to using the single-MLP system. An improvement is also seen in the odorants that did not surpass the threshold level, where two vectors are predicted just under the 25% requirement. Using a double hidden layer also showed improvement in the odorants that had previously scarcely passed the threshold; however, no additional chemical odorants exceeded the threshold when compared to the single MLP system.

Applying the optimum number of hidden neurons for the hybrid system proved valuable. As presented in Figure 7B, the magnitude of the prediction values has increased. The single hidden layer successfully classified 14 of the 15 odorants. The unclassified odorant presented a prediction accuracy value of 24.4%, failing classification by less than 1%. The minimum prediction accuracy of the double-layer hybrid network was 45.9%, exceeding the defined threshold by more than 20%. These results from the array of the hybrid network contrast with that of the single MLP, in which the performances of the single and double hidden layer are now reversed, with the double-hidden-layer structure outperforming the single-hidden-layer structure.

Applying bootstrapping on this optimum hidden neuron hybrid system showed a reduction in the number of validation vectors correctly classed. The single-hidden-layer system on average classified 14 odorant vectors with a range of $12 \leq n$ odorants superseding threshold ≤ 14, while the double-hidden-layer system presented a result of 13 odorant vectors classed on average with a range of $9 \leq n$ odorants superseding threshold ≤ 14. As noted earlier, bootstrapping allows a more accurate presentation of the network prediction performance. As seen from Figure 7A, a few of the validation odorants were classed into the incorrect chemical class by the predictor. This result was also observed in the linear predictor (see section 4.1) and in the basic single- and double-hidden-layer MLP (see section 4.4) through the use of the fixed validation set. Using 10,000 different training and validation sets removed, the effect of the low-performing

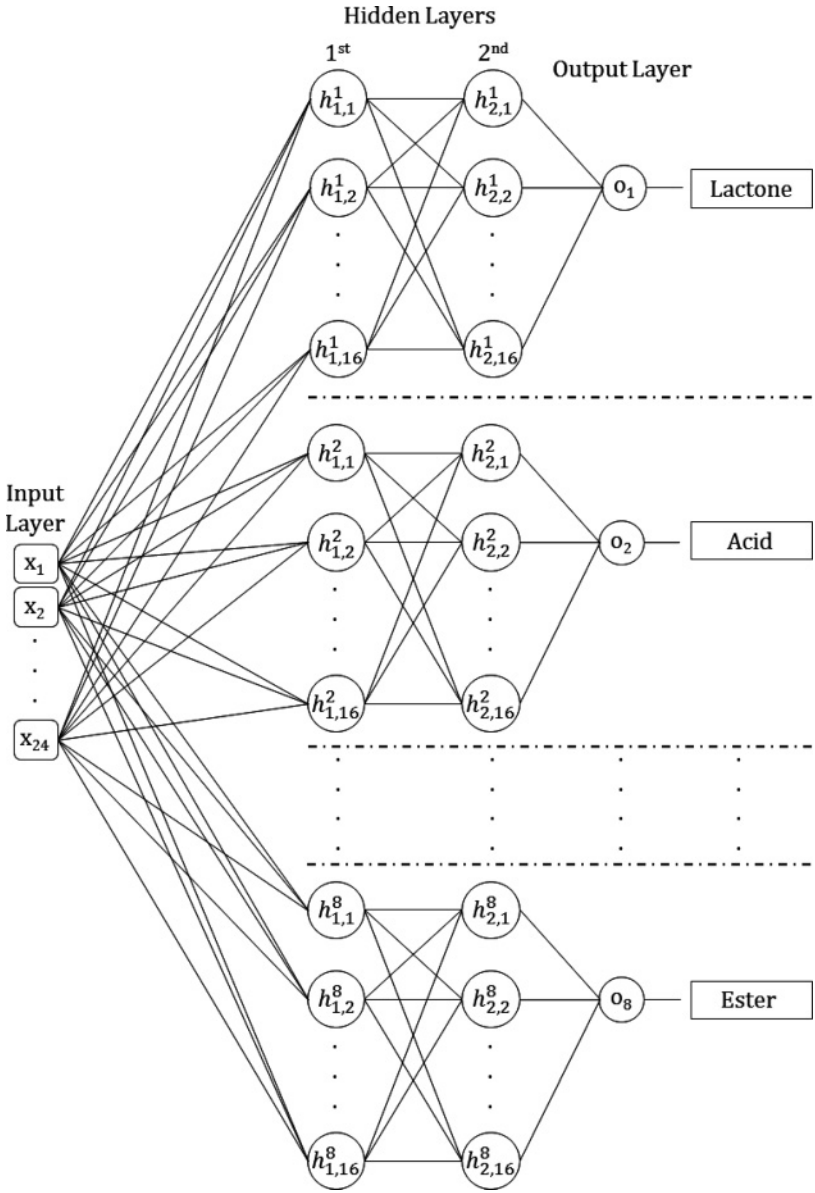


Figure 6: Schematic of a double-hidden-layer hybrid MLP with a single-neuron output layer corresponding to the chemical classes of the odorant molecules.

odorants of the fixed validation set. A larger database of odorant receptor responses is required to yield improvements in the outliers observed.

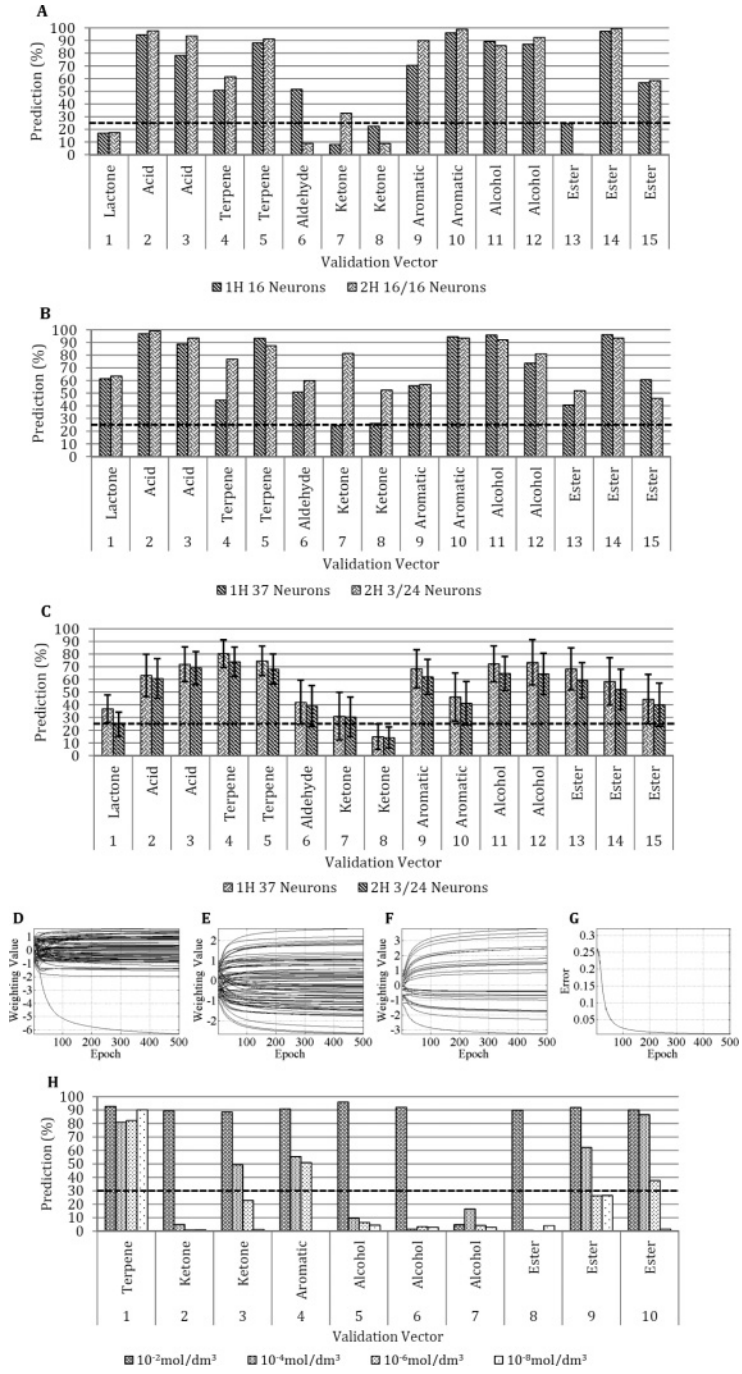
These results suggest that the hybrid network system improves on the basic single MLP design. However, the computational power involved when preparing and establishing the array of MLPs must be considered. Because the data set in this project contains eight unique classes for categorization, it can be said that using a single MLP for analysis requires one-eighth of the processing power needed for the array of MLPs.

5 Applying Noise Injection to the Data Set

Adding jitter or noise to the training data set can increase the generalization performance of the network (An, 1996). Sietsma and Dow (1991) demonstrated that noise injection leads to an improvement of the performance of the ANN as it requires the ANN to have more weights or hidden units. Noise injection has also been found to be independent to the type of noise used (Levin, Lieven, & Lowenberg, 2000). While a magnitude of approximately 10% of noise was discovered to be optimal for improving generalization of the network (Levin et al., 2000), 25% noise was successfully used to recognize patterns through obscuration (Unsworth & Coghill, 2006). In accordance with An (1996), a uniform additive noise of 10% was employed and applied to the training sets in this study.

Figures 8A and 8B depict the training process of the raw data and the added noise data when using a single MLP (see Figure 8A) and a hybrid MLP system (see Figure 8B). The single MLP system showed a faster rate of convergence with added noise. The MLP shown for the hybrid system is that of the lactone class. The training of raw data presents a fluctuation in the error. However, like the single MLP system, with added noise, the MLP shows a faster rate of convergence. This trend in improvement of training error is also seen in the MLPs of the seven other chemical classes of the hybrid system.

Noise injection was shown to be valuable in certain network designs. As presented in Figure 8C, the noise-injected training data induce a further increase in predicting odorants that were successfully classified when using the raw data. However, the odorants that initially bordered the threshold value of 25% when using the raw data showed a reduction in prediction when noise injection was applied. Although an overall improvement in magnitude of prediction accuracy was observed, the performance of an MLP in this research is ultimately based on the number of correctly classed odorants, determined by the threshold of 25% prediction. The raw data classified 14 odorants, while the noise-injected data classified 12. The use of noise injection had a negative impact for odorant prediction. Thus, because it was found that MLP performance was better for networks where noise injection was not used, bootstrapping was applied only to the MLPs with no noise injection.


A
Prediction (%)
Lactone Acid Acid Terpene Terpene Aldehyde Ketone Ketone Aromatic Aromatic Alcohol Alcohol Ester Ester Ester
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Validation Vector
1H 16 Neurons
2H 16/16 Neurons
B
Prediction (%)
Lactone Acid Acid Terpene Terpene Aldehyde Ketone Ketone Aromatic Aromatic Alcohol Alcohol Ester Ester Ester
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Validation Vector
1H 37 Neurons
2H 3/24 Neurons
C
Prediction (%)
Lactone Acid Acid Terpene Terpene Aldehyde Ketone Ketone Aromatic Aromatic Alcohol Alcohol Ester Ester Ester
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Validation Vector
1H 37 Neurons
2H 3/24 Neurons
D
Weighting Value
Epoch
E
Weighting Value
Epoch
F
Weighting Value
Epoch
G
Error
Epoch
H
Prediction (%)
Terpene Ketone Ketone Aromatic Alcohol Alcohol Alcohol Ester Ester Ester
1 2 3 4 5 6 7 8 9 10
Validation Vector
10-2mol/dm3
10-4mol/dm3
10-6mol/dm3
10-8mol/dm3

6 Reducing the Number of Olfactory Receptors Required for Classification of Odorant Chemical

Our motivation for this work is the development of a biosensor device based on insect olfactory receptor signaling. Such a device must be able to differentiate odor classes. However, for practical reasons, it would be advantageous to minimize the number of different receptors required in the device. Therefore, we have considered whether a reduced number of receptors could be chosen from the 24 included in this study while still maintaining discriminatory power. A bank of possible receptor input patterns for the neural network was formed, as shown in Figure 9A, where a different ordering of the same combinations was not permitted and receptors chosen did not occur twice in a combination. As the number of receptors is decreased from 24, the number of possible receptor network combinations increases, as illustrated in Figure 9B. Figure 9B highlights the difficulty in processing time of vast combinations. Thus, we elected to use the full 42,504 combination bank for the five-receptor ANN because this was of particular importance, and then used a sample of 10,000 combinations for all the other cases. Once the best-performing combination of receptors has been found, the resulting network is then put through bootstrapping where 10,000 different training and validation sets are used.

Because the hybrid MLP system proved the best MLP architecture in previous sections, it is used for the analysis when using only five receptors. We compared single- and double-hidden-layer architecture where the number of hidden neurons was based on the trend observed when using the complete data set (see section 4.4). The single hidden layer used was a 5-5-1 MLP, and the double-hidden-layer MLP had a 5-5-7-1 structure. Each of these designs was presented with the complete bank of patterns, so 85,008 simulations were performed, yielding 680,064 MLPs, which analyzed the data over two weeks.

The best-performing sequence of the input pattern bank was based on the highest accuracy of the worst-performing odorant of the validation set.

Figure 7: Performance of the hybrid MLP system using a single (left bar) and double (right bar) hidden layer, with the (A) basic number of hidden neurons and (B) optimized number of hidden neurons for the particular network architecture. (C) Bootstrapping is performed on the optimized hidden neuron hybrid network architecture. Evolution of weighting values of the (D) input layer, (E) first hidden layer, and (F) second hidden layer over 500 network epochs. (G) Evolution in network error over 500 network epochs. (H) Performance of the hybrid double-hidden-layer MLP system in predicting validation odorants of various concentration levels (10^{-2}, 10^{-4}, 10^{-6}, and 10^{-8} mol/dm^3). The horizontal dashed lines represent the 25% threshold value that defines correct classification of an odorant of the validation set.

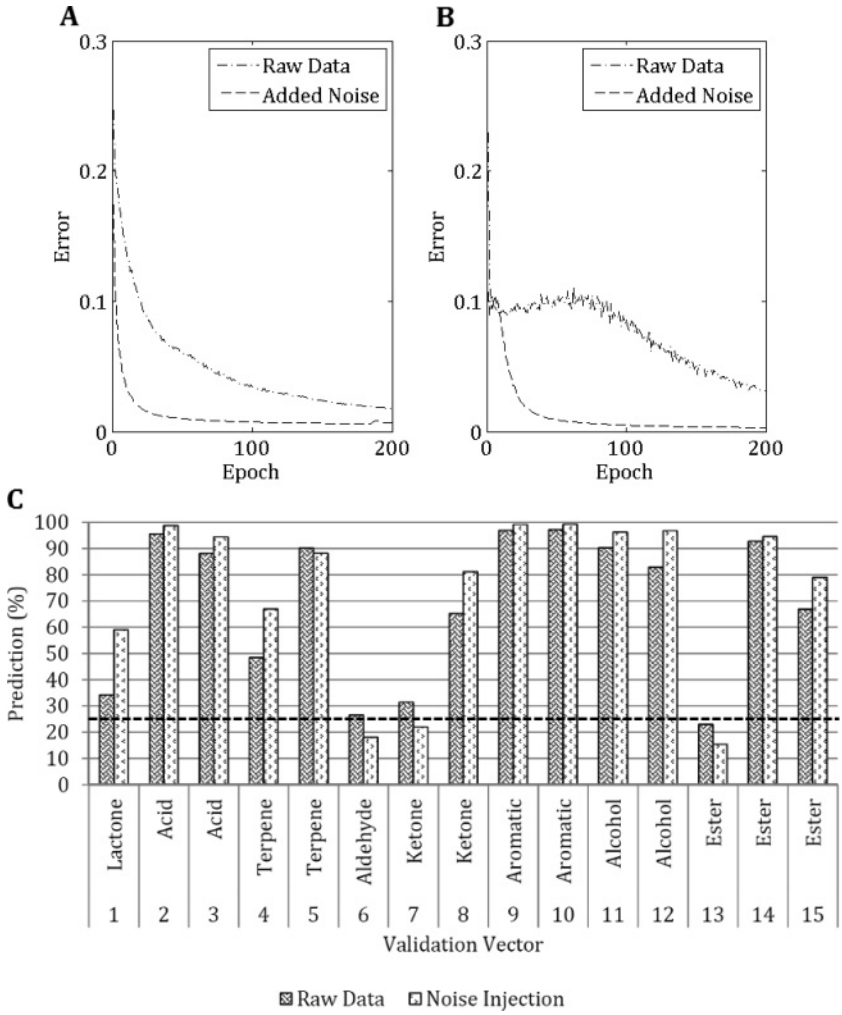
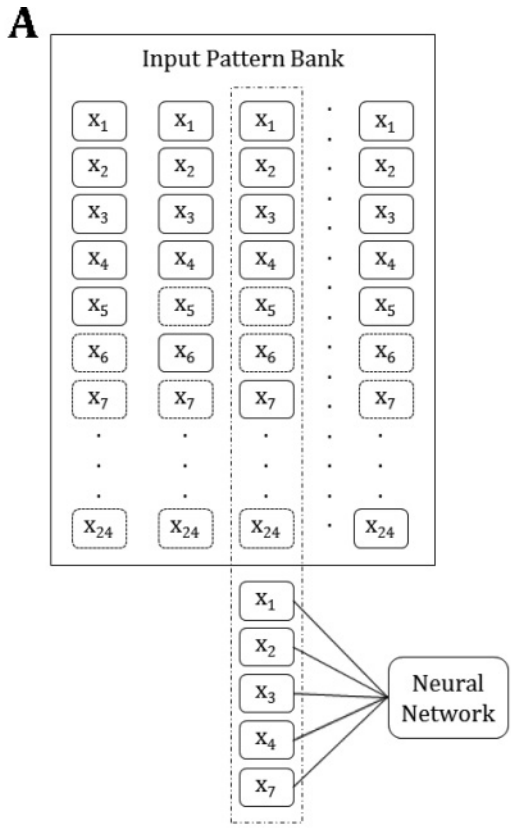


Figure 8: Training for 200 epochs with raw and noise-injected data of the (A) single MLP system and (B) lactone class MLP of the hybrid MLP system. (C) Performance of the optimum double hidden layer, single MLP system using raw (left bar) and noise-injected (right bar) training data. The horizontal dashed lines represent the 25% threshold value that defines correct classification of an odorant of the validation set.

The sequence of five receptors that generated the best result when using a single hidden layer were receptors: Or7a, Or47a, Or49b, Or59b and Or65a. When a double hidden layer was used, the following receptor combinations proved to be the best: Or2a, Or7a, Or49b, Or59b, and Or65a. As depicted

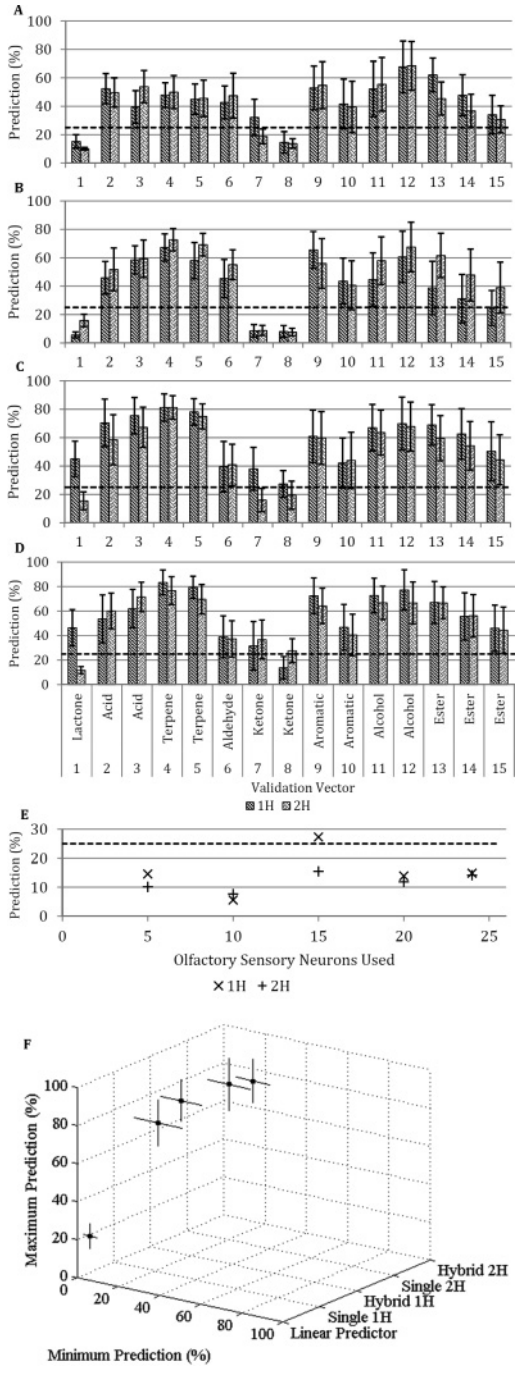


B

	Olfactory Sensory Neurons Used			
	5	10	15	20
Combinations for Processing	42,504	1,961,256	1,307,504	10,600

Figure 9: (A) Schematic of the procedure to obtain a set of five receptors from the available bank of possible receptor input combinations. (B) Number of combinations available corresponding to the varying numbers of OSNs used for processing with the hybrid MLP system.

in Figure 10A in which the combinations were put through bootstrapping, the single-hidden-layer MLP on average presented a result of 13 classified odorant vectors with a range of $11 \leq n$ odorants superseding threshold ≤ 13, while the double hidden layer on average classified 12 odorant vectors with a range of $10 \leq n$ odorants superseding threshold ≤ 12.


A
B
C
D
Prediction (%)
Lactone
Acid
Acid
Terpene
Terpene
Aldehyde
Ketone
Ketone
Aromatic
Aromatic
Alcohol
Alcohol
Ester
Ester
Ester
Validation Vector
1H 2H
E
Olfactory Sensory Neurons Used
×1H +2H
F
Maximum Prediction (%)
Minimum Prediction (%)
Hybrid 2H
Single 2H
Hybrid 1H
Single 1H
Linear Predictor

The number of OSNs used was also tested. The corresponding combinations for processing are included in Figure 9B. Similar to the five-receptor analysis, the performance of a single and double hidden layer was compared where the number of hidden neurons of the MLPs was also based on the trend observed in section 4.4.

The best-performing combination of each case is put through bootstrapping, as presented in Figures 10B, 10C, and 10D. Using 10 receptors with a single hidden layer on average classified 12 odorant vectors with a range of $9 \leq n$ odorants, superseding threshold $\leq$12, while the double hidden layer classified on average 12 odorant vectors with a range of $10 \leq n$ odorants, superseding threshold $\leq$12 (see Figure 10B). Increasing the receptors for analysis to 15 yielded an improvement, where the single hidden layer classified on average 15 odorant vectors with a range of $12 \leq n$ odorants, superseding threshold $\leq$15, while the double hidden layer on average classified 12 odorant vectors with a range of $12 \leq n$ odorants, superseding threshold $\leq$13 (see Figure 10C). With 20 receptors, the system on average classified 14 odorant vectors with a range of $12 \leq n$ odorants, superseding threshold $\leq$14, using a single hidden layer, while the double hidden layer classified on average 14 odorant vectors with a range of $10 \leq n$ odorants, superseding threshold $\leq$14 (see Figure 10D).

Using 15 receptors presented the highest minimum prediction, with a mean prediction value of 27.33% for the ketone odorant vector and a mean prediction of 58.46% across the validation set. The effect of changing from 10 or 20 receptors to 15 receptors is clearly seen in the increase in mean prediction magnitude of the lactone, terpene, and ester chemical classes. This result is illustrated by the higher magnitude of prediction of Figure 10C compared to Figures 10B and 10D, with prediction differences ranging from 15% to 40%.

Selection of the double-hidden-layer hybrid MLP as the network of choice for this final analysis is presented in Figure 10F by the maximum and minimum prediction values of the mean validation output obtained through bootstrapping the different ANN architectures employed in this work. The double-hidden-layer hybrid MLP system, hybrid 2H, in Figure 10F clearly supersedes the other ANN architectures: linear predictor, single

Figure 10: Network output of the best-performing combinations for a single and double hidden layer using a varying number of OSNs A–D: (A) 5 OSNs, (B) 10 OSNs, (C) 15 OSNs, (D) 20 OSNs, where the left bar (A–D) represents one hidden layer and right bar represents two hidden layers. (E) The minimum prediction recorded in the validation set with different numbers of OSNs used for processing. The horizontal dashed lines represent the 25% threshold value that defines correct classification of an odorant of the validation set. (F) A 3D plot of the highest and lowest prediction percentage of the mean validation set obtained through bootstrapping the different ANN architectures.

1H, hybrid 1H, and single 2H. This reinforces the suitability of this MLP for analysis of altering receptor length combinations. However, because the difference between the single- and double-hidden-layer hybrid MLP is small both of these architectures were used in the analysis of this section.

Employing the 25% threshold value described in section 4.1 allows us to quantify the degree of network learning and its response to unseen chemical odorants. The minimum prediction value of a validation set represents the weakest predicted value of the unseen odorants. Hence, a minimum prediction exceeding the threshold value implies the classification of the full validation set. There appears to be no trend observed in the minimum prediction in Figure 10E of the bootstrapping analysis of the receptor lengths. The only receptor length combination that exceeded the threshold value is that of using 15 receptors, superseding the prediction result when the full receptor array was used.

The performance of a receptor is of particular interest for the development of a biosensor based on olfactory sensory receptors. As more OSNs are used for processing, the receptors that elicited strong network performance changed. Table 2 lists the receptors that produced the best network performance. It should be noted that the receptors chosen are not necessarily the best ones that might be chosen for a biological study; rather they are the receptors that produce the best prediction for the network architecture with a fixed input space and the necessary receptors for a signal processing biosensor back end. Essential receptors that produced the best performance could now be identified when they were used by all networks, such as receptor Or7a. It is highly likely that Or2a is also a fundamental receptor. Furthermore, the importance of receptors Or9a, Or19a, and Or85b is seen by their addition to the best-performing combinations when 10 or more OSNs are used for analysis.

The difference between receptors chosen for biological study and those chosen for network performance can be highlighted by Or67a. From the study conducted by Hallem and Carlson (2006), we know that the response of individual receptors ranges along a continuum from narrowly tuned to broadly tuned. Or67a in particular was strongly excited by 31 of the 110 odorants of the lactone, acid, aldehyde, ketone, aromatic, alcohol, and ester classes. This broadly tuned receptor would appear ideal for conditions when only limited receptors are available for use; however, it was selected only for when 20 OSNs were used. Our hypothesis is that although this receptor is suitable for *in vitro* work in a biological study, it may not necessarily be suitable for distinguishing between classes with an ANN due to its broad range of response. Rather, receptors narrowly tuned to different chemical classes may provide better outcomes between separate odorant classes, as is the case in the selection of Or49b in the five-receptor-array case. The odor response of Or49b was found to be somewhat narrowly tuned and showed exceptional selectivity, with strong excitatory responses to odorants that contained a benzene ring (Hallem & Carlson, 2006). This

Table 2: Olfactory Receptors That Presented the Best Network Performance.

OSNs Used	Hidden Layers	Olfactory Receptor																							
		2a	7a	9a	10a	19a	22a	23a	33b	35a	43a	43b	47a	47b	49b	59b	65a	67a	67c	82a	85a	85b	85f	88a	98a
5	1H		✔										✔		✔	✔	✔								
	2H	✔	✔												✔	✔	✔								
10	1H	✔	✔	✔	✔	✔		✔										✔	✔			✔			✔
	2H	✔	✔	✔	✔	✔	✔		✔		✔		✔									✔			
15	1H	✔	✔	✔	✔	✔	✔	✔	✔		✔			✔		✔					✔	✔	✔		✔
	2H	✔	✔	✔	✔	✔	✔	✔	✔	✔		✔		✔	✔				✔			✔	✔		
20	1H	✔	✔	✔		✔	✔	✔	✔		✔	✔	✔	✔	✔	✔	✔	✔	✔		✔	✔	✔		✔
	2H	✔	✔	✔	✔	✔	✔	✔		✔	✔	✔		✔		✔	✔	✔	✔	✔	✔	✔	✔		✔

Note: The best hidden-layer structures for a particular olfactory receptor length chosen were, for 5 OSNs, layer 1H; for 10 OSNs, layer 2H; for 15 OSNs, layer 2H; and for 20 OSNs, layer 2H.

suggests that receptors showed selectivity for odorants within a chemical class. Thus, an important insight for biosensor design is that in building a biosensor, it is not necessarily a one-to-one correspondence for the receptors that are chosen for biological study to those that are chosen for efficient network performance, as reflected by the selection of the narrowly or nearly narrowly tuned receptors Or2a, Or49b, Or59b, and Or65a that elicited the best performance of the five-array MLP case. Only one receptor, Or7a, in this receptor combination was a near broadly tuned receptor. This shows that narrowly tuned receptors provide better prediction, as they aid in differentiating one chemical class from another, whereas a broadly tuned receptor will confuse the training of the network by firing at every odorant chemical class. The choice of receptors is application dependent. We stress here that the receptors that provide the best prediction across most networks are the most suitable for the signal processing back end for the biosensor rather than the receptors that work best in a biological setting.

7 Effects of Concentration Level on Network Performance

This section documents the findings of how the optimal hybrid network of section 4.5 performed for varying levels of odorant concentration. In this letter, the data from Hallem and Carlson (2006) used an odorant concentration level of 10^{-2} mol/dm^3. However, they also reported results from smaller data sets of concentration levels (10^{-4}, 10^{-6}, and 10^{-8} mol/dm^3).

Understanding the issue of concentration is important for a final biosensor design as it is rarely possible to accurately control stimulus concentration in reality. In Hallem and Carlson (2006), only 10 vectors were provided at each concentration level consisting of terpene (1 vector), ketone (2), aromatic (1), alcohol (3), and ester (3). Due to the few concentration data available, it was not possible to train the network on different concentrations. However, it was possible to assess how well a network trained on the original concentration level of 10^{-2} mol/dm^3 performed at classifying vectors from the other concentration levels. This would allow us to gain an empirical sense of the lower-bound (or worst-possible-case) performance of the network, and one would expect better performance if adequate concentration training data were provided. Figure 7H shows that as the concentration level is reduced from the 10^{-2} mol/dm^3 (i.e., the training concentration level) to the lowest concentration of 10^{-8} mol/dm^3, the prediction performance reduces accordingly, giving prediction performance (90% for 10^{-2} mol/dm^3, 50% for 10^{-4} mol/dm^3, 33% for 10^{-6} mol/dm^3, and 10% for 10^{-8} mol/dm^3). This simple test demonstrates that a network trained on one specific concentration level is more sensitive to concentration levels that are adjacent to the trained concentration level than those farther away.

Thus, from these results, we can suggest that a realistic working biosensor would require the network to be trained on a variety of concentrations rather than a single concentration level to maximize performance. This

result highlights the challenges faced in olfactory chemosensory processing systems: dealing with very large dynamic ranges of input that commonly occur with concentration variation. As well as training over multiple concentration vectors in the training set, additional suggestions to the concentration invariance problem that may potentially yield improvements could be by way of normalization of the input vectors.

8 Conclusion

In this letter, we have demonstrated how it is possible to use the generalization properties of an artificial neural network (ANN) for prediction in the field of chemical olfaction. This was ultimately achieved by training the firing rates recorded from an array of insect olfactory sensory neurons (OSN) on a set of parallel multilayer perceptrons (MLPs) to predict and differentiate eight distinct chemical classes.

Initial tests with a linear predictor on the fixed validation set presented 10 of the 15 possible odorants having exceeded the defined threshold value of 25%. Implementation of bootstrapping resulted in no odorant mean values that surpassed, with a possible range of $0 \leq n$ odorants superseding threshold ≤ 3. The performance of the linear predictor is insufficient for the highly complex olfactory data used in this project. Interconnections between the layers of a MLP contribute to their nonlinear characteristic, which is more suitable for the data that are used. A single-hidden-layer MLP classed 9 odorants, and a double-hidden-layer design classed 11 of the 15 fixed validation set odorants. By reducing the output layer to a single neuron and constructing eight different MLPs (corresponding to the eight chemical classes) in parallel, additional validation odorants were able to be correctly classified. Both the single and double-hidden-layer of this hybrid MLP system correctly classed 11 odorants, an improvement of two additional odorants for the single-hidden-layer design. An optimal layer length was found by incrementally altering the number of neurons of the hidden layers. Noise injection was implemented and was found to enhance network training by increasing the rate of convergence. By introducing 10% noise to the system, we were able to attain improvement for certain network designs.

The double-hidden-layer hybrid network design (24-3-24-1) correctly classified all 15 odorants of the fixed validation set. This MLP design also recorded the highest accuracy of 46% for the least-detected odorant of the set. With this, we demonstrated that with a hybrid network and optimum hidden-layer lengths, the classification of chemicals from the firing rates of insect olfactory receptors can be performed with high accuracy and success. Applying bootstrapping on this optimum hidden neuron hybrid system presented a reduction in performance of an average classification of 14 odorant vectors with a range of $12 \leq n$ odorants superseding threshold ≤ 14 for the single-hidden-layer system, and the double-hidden-layer system

on average classified 13 odorant vectors with a range of $9 \leq n$ odorants superseding threshold $\leq$14. Although it is not implemented in the final biosensor design, bootstrapping allows a more accurate presentation of the network prediction performance.

In accordance with the aim of minimizing the number of olfactory sensory receptors required to be used in a biosensor, simulations were performed to find the five receptors that yielded the best prediction of chemical class. A single- and double-hidden-layer MLP system was applied on the receptor combinations, and four common olfactory receptors yielded the greatest network prediction: Or7a, Or49b, Or59b, and Or65a. The single-hidden-layer hybrid MLP system elicited the best performance with a performance on average classifying 13 odorant vectors with a range of $11 \leq n$ odorants superseding threshold $\leq$13. The number of receptors needed for complete classification was found to be 15 olfactory receptors using a single-hidden-layer hybrid system. In addition, receptors Or2a, Or7a, Or9a, Or19a, and Or85b were identified to be the best for performance across all network architectures. However, what is interesting from this work is that the receptors chosen for the signal processing back end of a biosensor are not necessarily the same odorant structures, such as the broadly tuned receptors used for in vitro biological studies.

In this letter, we also assessed how the optimum hybrid network trained on odorants of 10^{-2} mol/dm^3 performed on odorants of lower concentration levels (10^{-4}, 10^{-6} and 10^{-8} mol/dm^3). We found that a network trained on one specific concentration level is more sensitive to concentration levels that are adjacent to the trained concentration level than those farther away. Thus, a realistic working biosensor would require an MLP to be trained on a variety of concentrations rather than a single concentration level to maximize its performance.

Acknowledgments

This work was supported by New Zealand's New Economy Research Fund C06X0701. We thank John Carlson for kindly providing the raw data.

References

An, G. (1996). The effects of adding noise during backpropagation training on a generalization performance. *Neural Computation*, *8*(3), 643–674.

Bachtiar, L. R., Unsworth, C. P., Newcomb, R. D., & Crampin, E. J. (2011a). Predicting odorant chemical class from odorant descriptor values with an assembly of mulit-layer perceptrons. In *Proceedings of the 33rd Annual International IEEE EMBS Conference* (pp. 2756–2759). Piscataway, NJ: IEEE.

Bachtiar, L. R., Unsworth, C. P., Newcomb, R. D., & Crampin, E. J. (2011b). Using artificial neural networks to classify unknown volatile chemicals from the firings

of insect olfactory sensory neurons. In *Proceedings of the 33rd Annual International IEEE EMBS Conference* (pp. 2752–2755). Piscataway, NJ: IEEE.

Baum, E., & Haussler, D. (1989). What size net gives valid generalization? *Neural Computation*, *1*, 151–160.

Berry, M.J.A., & Linoff, G. S. (2009). *Data mining techniques*. New York: Wiley.

Bishop, C. M. (1995). *Neural networks for pattern recognition*. New York: Oxford University Press.

Blum, A. (1992). *Neural networks in C++: An object-oriented framework for building connectionist systems*. New York: Wiley.

Churchland, P. S., & Sejnowski, T. J. (1992). *The computational brain*. Cambridge, MA: MIT Press.

de Bruyne, M., Clyne, P. J., & Carlson, J. R. (1999). Odor coding in a model olfactory organ: The drosophila maxillary palp. *Journal of Neuroscience*, *19*(11), 4520–4532.

de Bruyne, M., Foster, K., & Carlson, J. R. (2001). Odor coding in the Drosophila antenna. *Neuron*, *30*(2), 537–552.

De Villiers, J., & Barnard, E. (1993). Backpropagation neural nets with one and two hidden layers. *IEEE Transactions on Neural Networks*, *4*(1), 136–141.

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics*, *35*(5–6), 352–359.

Efron, B. (1983). Estimating the error rate of a prediction rule: Improvement on cross-validation. *Journal of the American Statistical Association*, 316–331.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. London: Chapman & Hall/CRC.

Ferreira, A., Ludermir, T., & de Aquino, R. (2005). A comparative study of neural network to artificial noses. In *Proceedings of the 2005 IEEE International Conference on Neural Networks* (pp. 2081–2086). Piscataway, NJ: IEEE.

Foody, G. (1995). Using prior knowledge in artificial neural network classification with a minimal training set. *International Journal of Remote Sensing*, *16*(2), 301–312.

Fu, J., Li, G., Qin, Y., & Freeman, W. J. (2007). A pattern recognition method for electronic noses based on an olfactory neural network. *Sensors and Actuators B: Chemical*, *125*(2), 489–497.

Gardner, J. W., & Bartlett, P. N. (1994). A brief history of electronic noses. *Sensors and Actuators B: Chemical*, *18*(1–3), 210–211.

Gardner, J., Hines, E., & Wilkinson, M. (1990). Application of artificial neural networks to an electronic olfactory system. *Measurement Science and Technology*, *1*, 446–451.

Garson, G. D. (1998). *Neural networks: An introductory guide for social scientists*. Thousand Oaks, CA: Sage.

Hallem, E. A., & Carlson, J. R. (2004). The odor coding system of *Drosophila. Trends in Genetics, 20*, 453–459.

Hallem, E. A., & Carlson, J. R. (2006). Coding of odors by a receptor repertoire. *Cell*, *125*(1), 143–160.

Hallem, E. A., Ho, M. G., & Carlson, J. R. (2004). The molecular basis of odor coding in the drosophila antenna. *Cell*, *117*(7), 965–979.

Harrell, F. E. (2001). *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*. New York: Springer-Verlag.

Haugen, J. E., & Kvaal, K. (1998). Electronic nose and artificial neural network. *Meat Science*, *49*, S273–S286.

Hildebrand, J. G., & Shepherd, G. M. (1997). Mechanisms of olfactory discrimination: Converging evidence for common principles across phyla. *Annu. Rev. Neurosci.*, *20*, 595–631.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366.

Huang, S., Tan, K. K., & Tang, K. Z. (2004). *Neural network control: Theory and applications*. Baldock, Hertfordshire: Research Studies Press.

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, *29*(3), 31–44.

Jefferis, G. S., & Hummel, T. (2006). Wiring specificity in the olfactory system. *Seminars in Cell Developmental Biology, 17*(1), 50–65.

Kavzoglu, T. (1999). Determining optimum structure for artificial neural networks. In *Proceedings of the 25th Annual Technical Conference and Exhibition of the Remote Sensing Society* (pp. 675–682). Nottingham, UK: Remote Sensing Society.

Kermani, B. G., Schiffman, S. S., & Nagle, H. T. (1999). Using neural networks and genetic algorithms to enhance performance in an electronic nose. *Transactions on Biomedical Engineering*, *46*(4), 429–436.

Ko, H. J., & Park, T. H. (2005). Piezoelectric olfactory biosensor: Ligand specificity and dose-dependence of an olfactory receptor expressed in a heterologous cell system. *Biosensors and Bioelectronics*, *20*(7), 1327–1332.

Lee, S. H., Jun, S. B., Ko, H. J., Kim, S. J., & Park, T. H. (2009). Cell-based olfactory biosensor using microfabricated planar electrode. *Biosensors and Bioelectronics*, *24*(8), 2659–2664.

Levin, R., Lieven, N., & Lowenberg, M. (2000). Measuring and improving neural network generalization for model updating. *Journal of Sound and Vibration*, *238*(3), 401–424.

Lewis, N. S. (2004). Comparisons between mammalian and artificial olfaction based on arrays of carbon black-polymer composite vapor detectors. *Accounts of Chemical Research*, *37*(9), 663–672.

Longstaff, I. D., & Cross, J. F. (1987). A pattern recognition approach to understanding the multi-layer perception. *Pattern Recognition Letters*, *5*(5), 315–319.

Pearce, T. C., Schiffman, S. S., Nagle, H. T., & Gardner, J. W. (2006). *Handbook of machine olfaction: Electronic nose technology*. New York: Wiley.

Sietsma, J., & Dow, R. J. F. (1991). Creating artificial neural networks that generalize. *Neural Networks*, *4*(1), 67–79.

Sola, J., & Sevilla, J. (1997). Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Transactions on Nuclear Science*, *44*(3), 1464–1468.

Swingler, K. (1996). *Applying neural networks: A practical guide*. San Mateo, CA: Morgan Kaufmann.

Tibshirani, R. (1996). A comparison of some error estimates for neural network models. *Neural Computation*, *8*(1), 152–163.

Unsworth, C. P., Bachtiar, L. R., Newcomb, R. D., & Crampin, E. J. (2011). *The Cybernose Project: Predicting the sense of smell*. Paper presented at the Annual Joint Symposium on Neural Computation, San Diego.

Unsworth, C., & Coghill, G. (2006). Excessive noise injection training of neural networks for markerless tracking in obscured and segmented environments. *Neural Computation*, *18*(9), 2122–2145.

Wilson, A. D., & Baietto, M. (2009). Applications and advances in electronic-nose technologies. *Sensors*, *9*(7), 5099–5148.

Zhang, H., Balaban, M. Ö., & Principe, J. C. (2003). Improving pattern recognition of electronic nose data with time-delay neural networks. *Sensors and Actuators B: Chemical*, *96*(1–2), 385–389.

Received February 29, 2012; accepted August 6, 2012.