Investigating the Role of Stimulus Concentration in the Way Glomerular Activity Distributes Chemical Groups in the Surface of the Mouse Olfactory Bulb Using Hierarchical Clustering

Senior Thesis

Presented to

The Faculty and Department of Neuroscience

Bates College

In partial fulfillment of the requirements

For the Degree Bachelor of Arts

By Joel Hallkaj

Lewiston, Maine

Fall, 2024

Acknowledgements:

I would like to thank my professor and advisor, Dr. Jason Castro, for introducing me to the world of olfaction and for allowing me to work on this unique project that taught me a lot about research and myself. Thank you for the motivation and technical support throughout this project. Special thanks to Dr. Matt Wachowiak in the University of Utah School of Medicine and his team for giving me the opportunity to work with their unique and impressively collected data.

I would like to thank all of my friends who have been a constant source of joy, support and have challenged my curiosity and intelligence constantly throughout my career here at Bates. Special thanks to Giancarlo and Pico who have been close to me during writing this thesis project both tolerating my frustrations with Python as I lay in their couches and for actively listening to me as I tried to explain what I was doing. Furthermore, I would like to thank everyone who has shared some laughter with me during this stressful time and shown their support.

I would like to thank my extended family and whoever I count as family for their constant support and encouragement. Moreover, I would especially like to thank my mother, Elda, and my sister, Joana, for their constant support and unconditional love throughout this project and for listening to what I was doing despite not being fully able to understand what was going on. I wouldn't have been able to be here at Bates without them.

Table of Content:

.

INTRODUCTION

All sensory systems face a trade-off between selectivity and sensitivity as they encounter stimuli and have to balance different goals while having to distill noisy information. Improving one's performance on one will always come at the detriment of the other due to limited biological resources. Put more colloquially it is a question of whether the system prioritizes range, i.e. how wide would the system cast the net of what it can process, or sensitivity, i.e. how granular does the information have to be. Sensory systems have evolved across modalities and species their own balances of this trade-off, making it a crucial quality of every system. Therefore, understanding this balance better at multiple stagepoints, can shed light into the unique way the system is organised as a whole.

The olfactory system, like all other systems, has to parse out information across many combinations of odorants that range across magnitudes of intensity and concentration. Hence, this tradeoff between selectivity and sensitivity is especially salient for the olfactory system, which has to discriminate between elements of a large, heterogeneous, and high dimensional space of stimuli over many orders of magnitude . For example, a sensitive olfactory system would respond to the low and minute details but it would inundate its differentiation at a relatively lower intensity of stimuli. Understanding the logic and mechanisms that the olfactory system employs with regards to stimuli in this respect is a key issue in understanding and perhaps further decoding the behavior of the olfactory system. In this study, we looked at how the mouse olfactory bulb (OB) represents stimuli varying in magnitude of concentration by analyzing data from odor-evoked activity in the cortex of the mouse olfactory bulb.

The OB lies in the orbitofrontal cortex sitting right above the nasal cavity (see Figure 1) and is the first central nervous system that is involved in olfactory processing . It sits right at the top of Cribriform plate from where it directly extends downstream single neuronal connections into the olfactory epithelium by employing bipolar cells called olfactory sensory neurons (OSNs). In the epithelium, they extend organelles called cilia that detect airborne molecules and transduce such chemical signals into electrical signals that are passed along upstream. The OSN are expressed by about 400 genes in humans, and about 1000 in mice, and are expressed by one gene per type of OSN (Ebrahimi & Chess, 1988). The specificity of these OSNs also continues as we see the glomeruli responding more strongly to single odorants (Wachowiak et al, 2022). The OSNs are modulated by g-protein coupled receptors that can hold complex configurations to accommodate ranges of odorants (Wu et al, 2024). The body of the OSN crosses the layers of the epithelium and clusters in the OB into bundles of cells called glomeruli, where they interact with excitatory Mitral/Tufted cells that send information along the olfactory tract into the brain for higher order processing. Their spatial arrangement corresponds to some degree also in the spatial arrangement of glomeruli in the OB, as dorsal OSNs can be represented on the dorsal OB. The OSN projects in two places in the OB once on the surface, and once more within the bulb. The continued similarity between the OSN and the glomeruli shows an affinity between chemical attributes and the activation in the OB. The glomeruli seem to serve a key role in olfactory processing due to their patterned activation encoding different stimuli attributes.

When the glomeruli connect to the Mitral/Tufted Cells, excitatory neurons, there are inhibitory interneurons connections between these glomeruli that contribute to an improved noise-to-signal ratio and the sparseness level of activation in the OB. This processing in the OB is important in normalizing signal processing from the ORN to the Mitral/Tufted Cells. This

lateral inhibition is similar to other sensory systems such as in vision. In the retina, there are also lateral inhibitory neurons that function between the photoreceptor cells and bipolar cells that trim and sharpen the excitatory activity as the signal travels towards the optic nerve; in order to achieve a similar goal of improvement in sensory acuity (Cree and Weimer, 2003). The mediator role that glomeruli serve through the inhibitory interconnections that they exhibit and their connections to excitatory cells that project upstream for sensory processing show that in this stage they play a key role in the trade-off between sensitivity, which in our study shows the concentration level at which the glomerulus will respond to the odorant and selectivity, the sparseness displayed on the OB for certain odorants. Therefore understanding the mapping of odorants and its clustering in the bulb is a key area of inquiry for understanding olfaction.

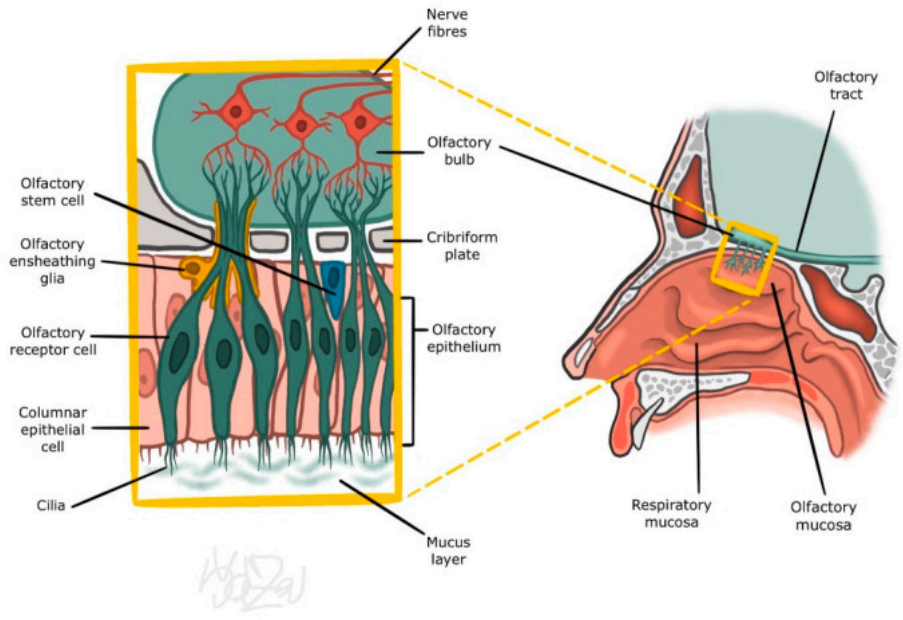

Figure 1. Diagram visualising the early olfactory system and the OB image from (Červený et al, 2022).

One important way in understanding olfaction is thinking about the topography of the OB as a space where we can map different odorant stimuli and understanding how different odorant

features are clustered together in this surface. Johnson and Leon (2007) highlighted that the organisation of glomerular responses happens in a spatial manner, where clustering of odorant attributes, such as functional groups, hydrocarbon structures, and other attributes such as water-solubility or volatility, seem to stem from OSN and Olfactory Receptor (OR) projections into the glomeruli. Olfactory scientists have attempted to create maps of bulbar activity based on odorant stimuli. The understanding is that neighboring glomeruli form clusters that respond to similar chemistry of odorants. For example, in the bulb we see that the clusters representing aliphatic and carboxylic acid odorants seem to extend rostrally as hydrocarbon chain length increases (Johnson and Leon, 2007), although this progression across the bulb is not present in all groups. The olfactory system has been thought to be activated by combinatorial coding of the odorant features, meaning that while certain odorants can activate glomeruli that are usually activated by a single odorant, there is also possible stimulation in other clusters in different places of the bulb. Aliphatic odorants, for example, can activate domains related to both their functional group and hydrocarbon chain length (Johnson et al, 2004). This example and other substitutions borrowed from (Johnson and Leon, 2007), show support for the theory that the brain relies on combinatorial coding to make sense of the odor space. Which, intuitively, also is plausible since there are more odorants than specialised types of ORs. This line of thinking has led to attempts to manually decode individual odorants.

Pashkovski et al. (2020), also examined this combinatorial code by correlating glomerular activity and the interconnectivity of lateral neurons with mitral to the activity of the Piriform Cortex, responsible for organising the information sensory. They found that the cortex reorganizes the glomerular information in pairwise relationships and is also transformed in the cortex based on statistics of the odor environment to single tuned neuron curves. They do this by

statistically sampling the activity in glomeruli to generate a space of odor representation. While the glomerulus organizes and normalizes the inputs in the bulb, akin to Z-normalisation (Leon and Johnon, 2007; Pashkovski et al., 2020) the cortex reorganizes said relationships emphasizing some and de-emphasizing others in perception. This shows that the combinatorial nature of the cortex is strong and important in the whole olfactory system.

Wachowiak et al. (2022), highlighted that this combinatorial evidence is true in higher concentrations but in lower concentrations we see more single neuron relationships in the OBs of mice. They mapped 186 odorants in 8 different bulbs to show that the relationships were sparse. They identified that there are limitations to this model at large due to the strategy that the OB employs in coding information and the nature of bulbar response remains yet unclear. Highlighting the need for further understanding. They employed a paradigm where they, through a series of pilot studies, determined low detectable concentrations of odorants to be used on two concentration regimes as a way of investigating a more granular relationship between odorants and glomerular activity. While absolute levels of glomerular activity increase with odorant concentration (Leon and Johnston, 207), there is a possibility that the observed increased overall activation could be a noisier or more ‘corrupted’ version of the activation we see at the lower activation level. Indeed, the uniqueness of the Wachowiak et al. study is the fact that they were able to obtain data at the lower concentrations (ranging from the nanomolar to the picomolar scale). They measured lifetime sparsity, an indicator of selectivity, determining that these sparse interactions and selectivity for sparse glomeruli is not due to artifacts of signal-to-noise ratios of weak response glomeruli but due to an organisational effect by the OB.

They indicated that odorant coding is not a lower dimensional space, as some arguments mentioned in (Johnson and Leon, 2007), but it is in fact a high dimensional space. With tuning of

glomeruli to certain deodorants only becoming slightly less sparse at the higher concentrations. Through the analysis of covariance in odorant pattern activations in the OB they proposed that the dimensionality was lower for the higher concentration, perhaps indicating that there is a significant difference between how odorant concentrations are presented in the two different concentration levels.

The attempts have been done manually and serious mapping has been achieved of only bulbar responses through glomeruli responsible for certain odorants or stimuli (for example, Zhang et al, 2013); however, the odorant space remains vast although still unknown (Gerkin and Castro, 2015). This approach also requires careful tuning for each individual model and is a cumbersome process for finding the perfect tuning of physicochemical features, etc. The findings of Wachowiak et al. indicate a new way to think about the link of ORs to the odorants by proposing a new paradigm of investigating the tuning of the OB in lower concentration regimes.

In this study, we will investigate data provided by the Wachowiak lab that shows the activation of eight olfactory bulbs tested against 59 odorants at threshold level concentration and 25x higher concentration for each odorant. The data has not yet been published in another study and it is unique for offering a level of granularity by widefield and 2-photon imaging for a low concentration that is extremely sensitive at these lower concentrations, barely detectable by the OB. The field of olfaction is transitioning its methods of investigating the olfactory system by employing more and more machine learning and computational methods in its arsenal as a way to combat the inherent complexity and dimensionality present in this system as seen in Pashkovski et al. (2020). Keeping on with the trends in the field, we investigated this dataset through the use of unsupervised learning and clustering methods to model the way that the OB would seem to cluster the odorants and employed a predictive model, using a random forest

classifier, to see the predictive accuracy that the different algorithms trained in one concentration regime would have on the other. We explored different clustering methods and graphical representations of this dataset in order to explore its properties. We hoped to employ this approach intending to show both that the clustering is different in different concentrations and also that the accuracy would highlight differences in these activations. Thereby, attempting to lend support to the theory that the selectivity of the bulb is different for different concentration regimes as well as shows divergent selectivity in its responses. The two main goals are A) investigating the hierarchical relationship and how to present this relationship of odorant co-clustering and B) seeing whether stimulus concentration affects the way that the OB distributes chemical attributes.

Methods:

Data Collection:

In this study we used eight different maps of mice olfactory bulbs that recorded glomerular activation. There were 59 odorants tested, which were grouped into 10 groups of varying lengths. This grouping was done subjectively by the Wachowiak lab and it considered progressive hydrocarbon chains, functional groups as well as informed judgements about the structure. These were tested on 4 mice and 8 bulbs. There is no significant difference between the bulbs and the lateralisation in terms of within and between mice variability (Pashkovski et al, 2020?). The results are stored in text files as matrices of varying dimensions for each mouse: [59 x 126], [59 x 119], [59 x 159], and [59 x 163]. The low concentration being threshold detection level and the high concentration being 25x that concentration. The asymmetry between matrix length being due to the sharpness and acuity of measurement required for each mouse. Wachowiak et al. (2022), showed that the activations are very sparse and this was confirmed with the dataset that we got by testing whether the number of zeros was larger than half of the elements of the matrix.

The data that we used were collected by our collaborators in the Wachowiak Lab, and were collected in a similar procedure to the measurements done in the study of Wachowiak et al. (2022). The experiments were performed on transgenic mice obtained from heterozygous crosses of OMP-IRES-tTA (Jackson Laboratory stock #017754) (Yu et al., 2004) and tetO-GCaMP6s (Jackson Labo-ratory stock #024742) (Wekselblatt et al., 2016). The mice were aged between 2-6 months and included both male and female. They were caged with up to 5 conspecifics and in a 12h light/dark cycle. The odorants were diluted using serial dilution until the desired levels and were presented to the mice for 2 seconds in 3-5 trials per odorant, then leaving time for rest

and an opportunity to clean the air before administering another odorant in the olfactometer. The olfactometer was custom built by the lab and equipped with an eductor for odor delivery through steam.

Imaging of the glomeruli was performed using in vivo widefield and two-photon calcium imaging on the dorsal OBs of the mice. The mice were kept under anesthesia where kept under anesthesia for the tracheotomy where a tube supporting artificial inhalation (150 ms duration, 300 mL/min flow rate), with the goal of decoupling olfaction from the confounds of breathing (Eiting and Wachowiak, 2018). The mice were head-fixed in order to be administered a thinning of the bone over the dorsal OB, forming a well, afterwards it was covered with dental cement and then rendered transparent with Ringer's solution for the purpose of having a window for imaging. Epifluorescence was collected through a 4×, 0.28 N.A. air objective (Olympus) at 256×256-pixel resolution and 25 Hz frame rate using a back-illuminated CCD camera (NeuroCCD-SM256; RedShirt Imaging) and Neuroplex software, with illumination provided by a 470 nm LED (M470L2, Thorlabs) and green fluorescent protein filter set (GFP-1828A-000, Semrock).

From these readings glomerular response maps were generated where the numbers of presentations were averaged over. The odorants were selected and made sure to be recorded with the most direct and specific glomerular activation. They were presented as the matrices we used in our data analysis.

Data Analysis:

Data analysis was performed by us using Python and R programming languages. They are preferred for their abilities to perform statistical and machine learning programs with

confidence and are commonly used for such tasks and the available libraries for both machine learning and hierarchical clustering, such as Scipy and Scikit-Learn. The code used in this study is provided in the Code Appendix page 35.

Data processing:

In order to perform hierarchical clustering we need to form distance matrices, that are square, symmetric measurements of pairwise distances along the different sequence of the matrix and its transpose. Considering our data was sparse and highly dimensional, we decided to use cosine distance as the metric for our analysis. Cosine distance is beneficial in our case because it measures the dissimilarity based on the cosine of the angle which can circumvent noise generated by sparsity since it only considers non-zero values. By calculating the relationship between the distances as a function of the angle, which also scales the data for easier processing. The formula for a matrix $\boldsymbol{X}$ would be:

$$Cosine\ Distance(X,\ X.T)\ =\ 1\ -\ X \cdot X.T\ /\ ||\mathrm{X}||\ ||\mathrm{X.T}||$$

Other distance metrics like euclidean, manhattan, and minkowski were tried throughout the process for exploratory purposes but performed subpar, as expected, to cosine distance. In our subsequent analyses in Python, we encountered problems since non-orthogonal data is stored as zero which creates problems for calculations that require divisions, producing infinities. R calculates the distance differently reducing the non-zeros and we also replicated the method in Python together with choosing to go with adding an epsilon values of $1 * 10^{10}$ to alleviate that problem and adding a normally distributed value of range 1e-8 to alleviate some errors raised by our function.. We confirmed high correlations between these matrices with a pearson correlation in order to continue with the rest of the analysis.

Mantel Test:

We wanted to explore the correlations between all pairwise combinations of the OBs using a mantel test and calculating spearman's rho, in R. The Mantel test calculates two matrices of the same shape and calculates the correlation through euclidean distance between the points. We calculated the cosine distance matrices for each OB to determine a preliminary sense of similarity between the OBs.

Hierarchical Clustering:

Hierarchical clustering is a machine learning algorithm that represents distance matrices and plots the different linkages between clusters of points. We represented and clustered glomerular activity using this method.

Firstly, we wrote a helper function that would help us decide the best linkages for clustering our data by maximizing the cophenetic correlation coefficient, that measures the cophenetic distances in the dendrogram through seeing how the dendrogram preserves the original distances of the data. We tested the linkages compatible with cosine distance: Single, Complete, Average, and Weighted. This was performed using the Scipy package in Python.

Secondly, we constructed Elbow graphs to determine the optimal number of clusters to extract from the data using three different methods: a custom written python function for Within Cluster Sum of Squares, a custom-written python function that calculates the acceleration of distance growth between clusters and thirdly using clustering error measurements like silhouette scores. We also calculated the gap statistic as a way to choose the optimal amount of clusters in Python.

Predictive Algorithm:

We wanted to design a test to show the independence of the two concentration conditions based on how well a classification algorithm predicts the clustering labels of one condition being

trained on the other. To that end, we performed a grid search of the following classifiers from the sci-kit learn package in Python: SVCs, Naive-Bayes Classifier, Ridge Classifier, K Neighbors Classifier, Random Forest Classifier, MLPC, and Gaussian NB. Opting in the end, to conduct the analysis with a Random Forest Classifier with the following specifications: ._estimators=1000, max_depth=8, min_samples_split=3, random_state=42, n_jobs=-1, criterion='gini', max_features='sqrt', bootstrap=True, class_weight='balanced_subsample', min_samples_leaf=3) Furthermore, we performed another grid search for hyperparameters of Random Forest Classifier using leave one out cross-validation that helps with our relatively small dataset to validate our scores. We also reran the analysis with the unclustered data to check for differences in the trend of the classifier to highlight possible confounds inherent in the classification through the use of hierarchical clustering labels as the true labels in the algorithm.

We concatenated all the glomeruli together and all the respective clustering labels as true labels in order to use in our grid search and classification. Due to unequal sizes there were many NaN values and we opted to cut the dataframe to the last column without NaNs which then gave us 117 glomeruli columns.

Results:

Hierarchical Clustering:

We wanted to represent clustering relationships using linkage dendrograms and to that end investigated the best way to implement this clustering approach. First, to confirm the nature of the data we ran a simple sparsity test that revealed that all OBs showed high levels of sparsity, especially the OBs treated with the low concentrations (see Table 1). This confirms the finding done in the previous experiments by Wachowiak et al. (2022) that showed that sparsity is high for all conditions and higher for lower concentration conditions. This can also be confirmed visually by looking at the activity of glomeruli for each tested odorant (see Figure 2). We expected sparsity scores of this kind through the visual representation in Figure 2, which confirms our choice for investigating cosine distances due to their efficiency in sparse and highly-dimensional data.

Names	OB 1 High	OB 2 High	OB 3 High	OB 4 High	OB 1 Low	OB 2 Low	OB 3 Low	OB 4 Low
Sparsity	0.892	0.888	0.890	0.880	0.956	0.955	0.996	0.962

Table 1. Simple sparsity measurements for each imaged OB.

To investigate the nature of the data further we calculated the cosine distance matrices in R, and plotted the mantel test between the cosine distance matrices using spearman correlation for each combination of all OBs (see Figure 2). We discovered moderate correlations between bulbar activity in different concentration regimes and saw moderate-to-strong correlations between some in the same concentration regime. Confirming our intuition for how the data behaves and confirming a relationship between the activity of each OB that forms an intuitive bases for applying our predictive algorithm tests.

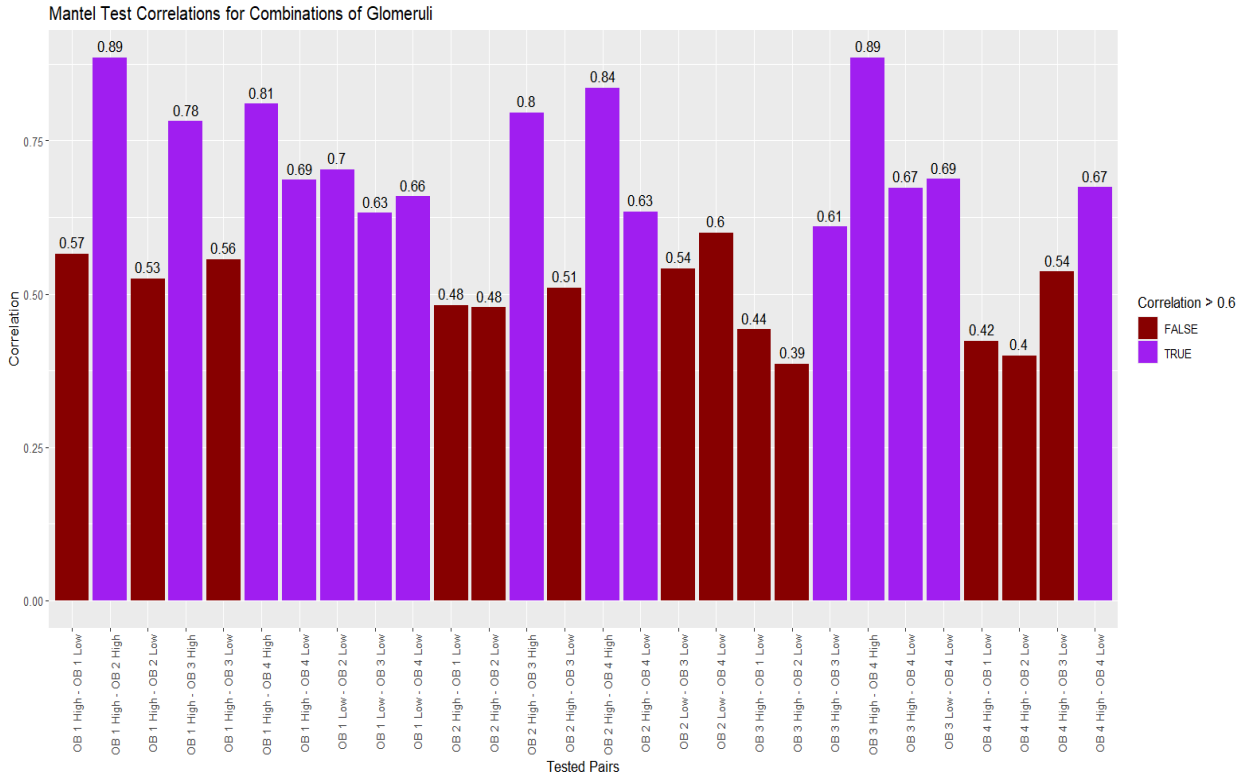

Figure 2. Mantel tests for all combinations of OBs within each other, highlighting with purple all the combinations that yielded a moderate to strong correlation as defined by rho > 0.6, all correlations were statistically significant with $p < 0.001$ for each combination tested.

Secondly, in order to show the hierarchical properties of our data through a dendrogram we tested different linkage methods that were compatible with the cosine distance metric. This was achieved by testing the different linkages that are built in the Scipy packages: 'average', 'weighted', 'single', and 'complete'. We determined the cophenetic correlation coefficient to investigate which linkages preserved the original data structures, i.e., which linkage method causes less distortions of the original distance matrix calculated for each OB. The results show that average linkage scores the highest cophenetic relationship across all the OBs (mean cophenetic correlation of 0.802; see Figure 4). We expected this method to be beneficial as it

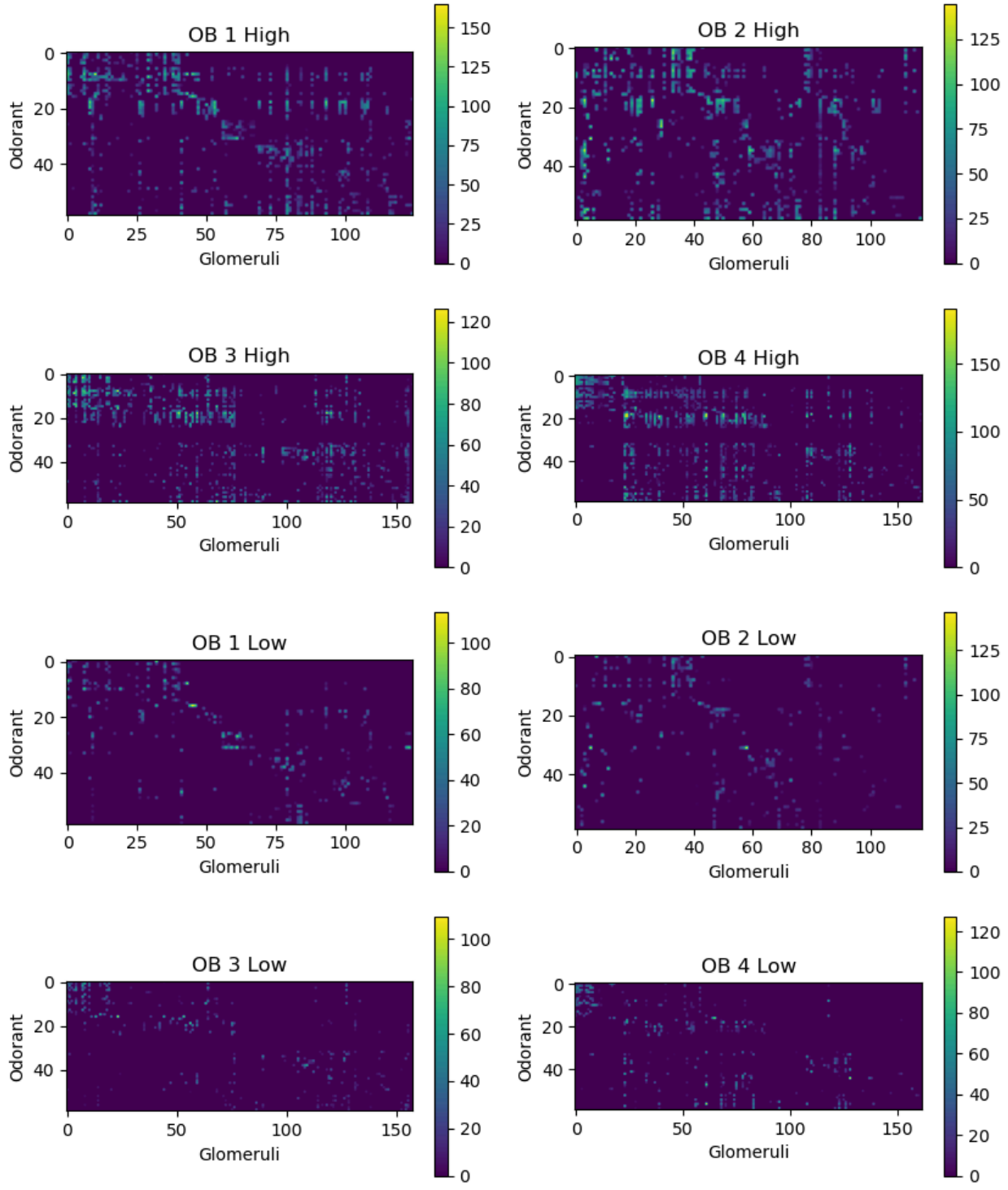

Figure 3. Activation shown in all OBs plotting the respective individual glomeruli activations per each of the 59 odorants coming from the raw data collected.

finds a balance between both single and complete linkage making it a more versatile approach to investigate within both groups and could indicate a standard for looking at glomerular activations on the OB in future similar experiments.

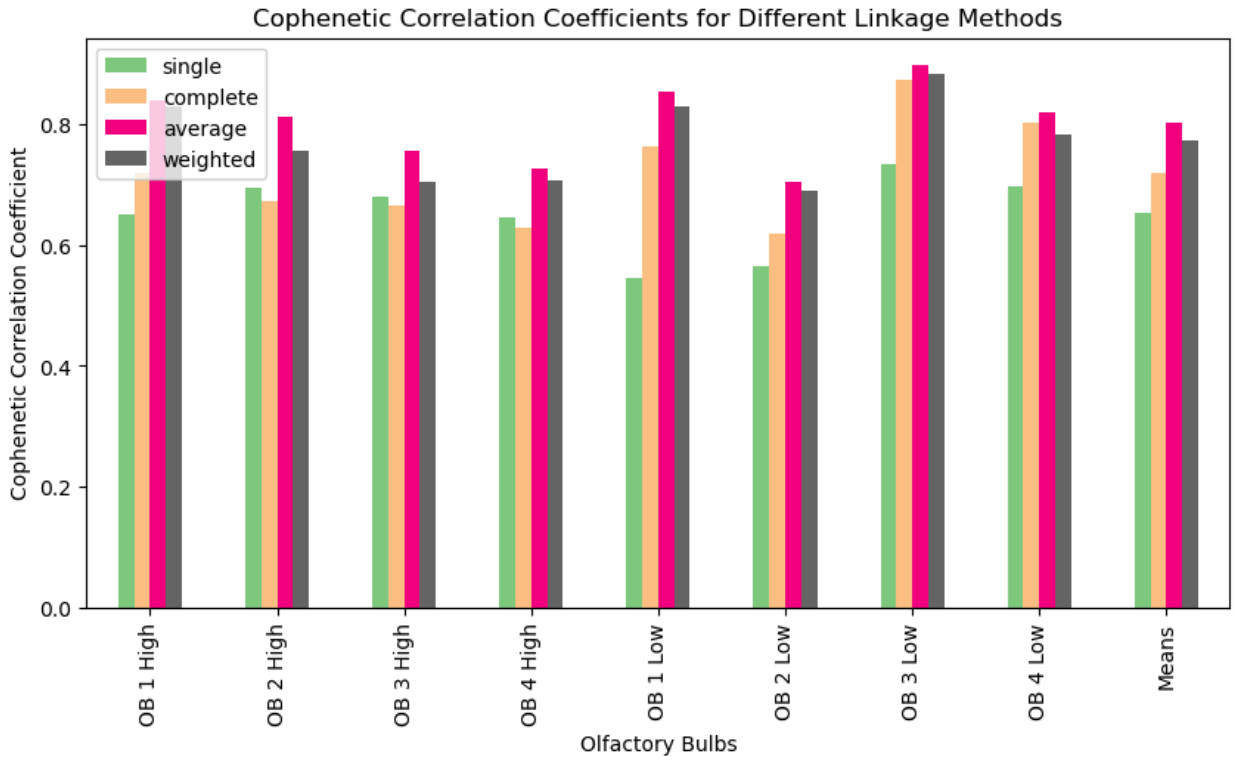

Figure 4. Cophenetic Correlations for the different linkage methods (see legend) that were tested on each OB, with the mean for all OBs appended at the right of the graph.

Choosing the Optimal Number of Clusters:

We determined the optimal number of clusters for the hierarchical clustering algorithm through visualising the information in each OB by plotting the within cluster sum of squares for each number of clusters (k) to be calculated (see Figure 5). We determined that across the elbow graphs the optimal way k hovers around 5. Other methods such as calculating the Gap Statistic yielded ranges hovering around 4-6 k, and also calculating a Davis Bouldin error score yielded results, although more inconsistent around 4 k and were excluded from this analysis.

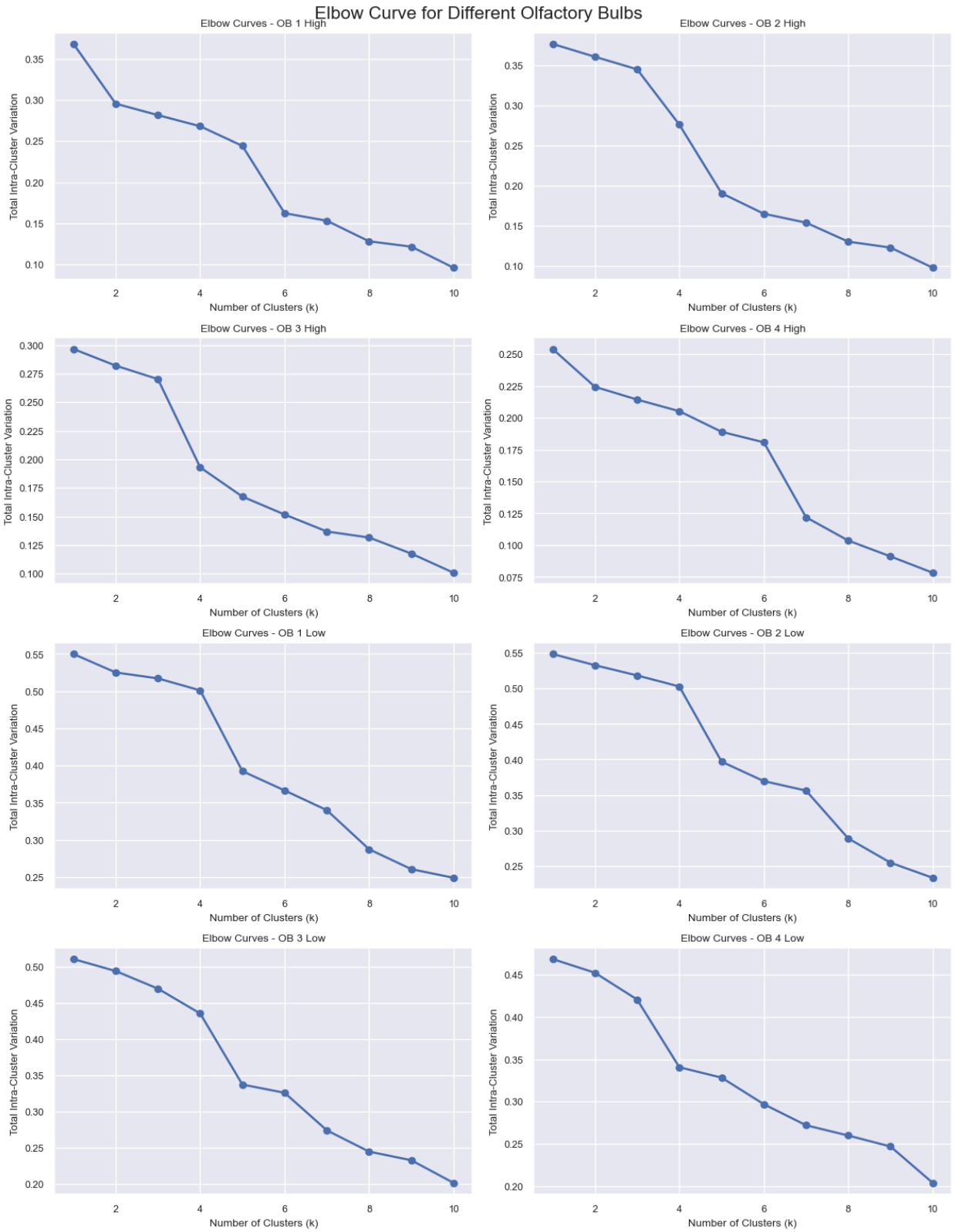

Figure 5. Elbow curves for each OB calculating the within cluster sum of squares using cosine distance metrics.

The number of optimal clusters was not surprising as we considered that two of our groups, Lactone and Hexanol/Hexanal, comprised two elements each which could be classified by activation perhaps to esters/carboxylic acids or alcohol respectively. The clustering algorithm calculates the relationships from the bottom up, and perhaps some of the activated glomeruli could respond to any of the shared features categorising them with other similar groups. Therefore, we expected less groups, and looking at the looser pre-experimental classification we can expect strong and larger groups like alcohols or aldehydes to cluster first and incorporate more elements.

Having determined the optimal number of clusters for the dendrogram we then plotted the relationships of the clustering of odorants into the five different groups. We plotted the information in different dendrograms as seen in (Figure 6). We extracted the cluster labels for every odorant through a custom function routine and then reorganized odorant information by odorant, cluster, and group name. This helped us preprocess our data for the supervised learning part of the study, since also cosine distances normalise the data eliminating the need for further preprocessing.

Another important result of our analysis was to plot the relationships between the hierarchical clusters and the original grouping of the data. We plotted the heatmap of the original grouping and the subsequent clusters in Figure 7. We saw that the clustering that Ketones, Thiazoles, Pyrazines, Lactones tend to cluster the same in both conditions, and on some of the clusters we see Carboxylic acids, Aldehydes and alcohols, which show a somewhat consistent co-clustering. We also see a difference in co-clustering between said groups in the different concentrations as well. Seeing the groups overlapping between clusters in clusters 0 and 1 low but largely separated still. Since Hierarchical clustering is calculated from the bottom up, it is

difficult to make further analysis since the clusters are arbitrary within the OBs and statistical comparisons are more difficult to make using this type of clustering.

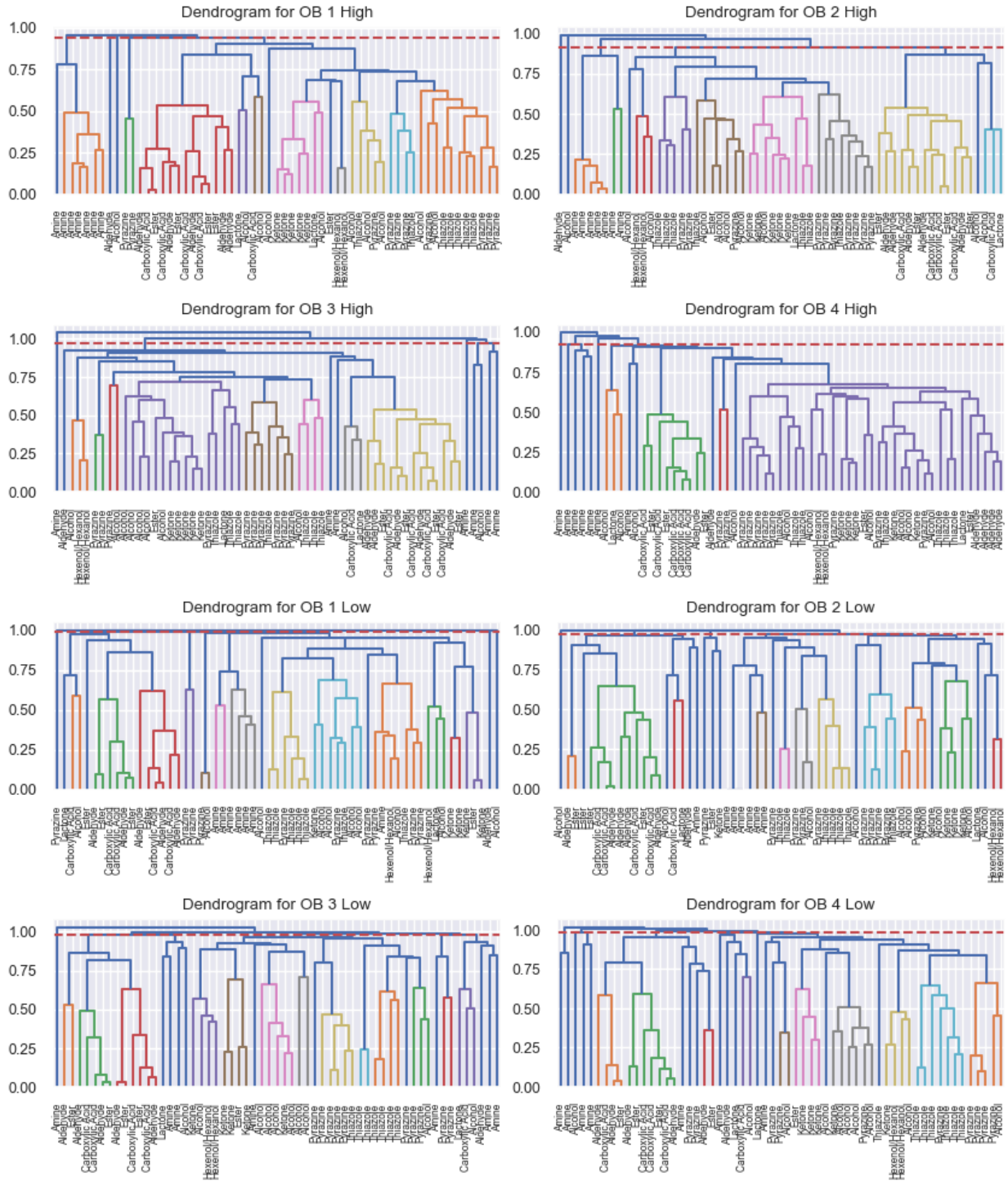

Figure 6. Cosine Distance Dendrograms for each OB in the respective condition.

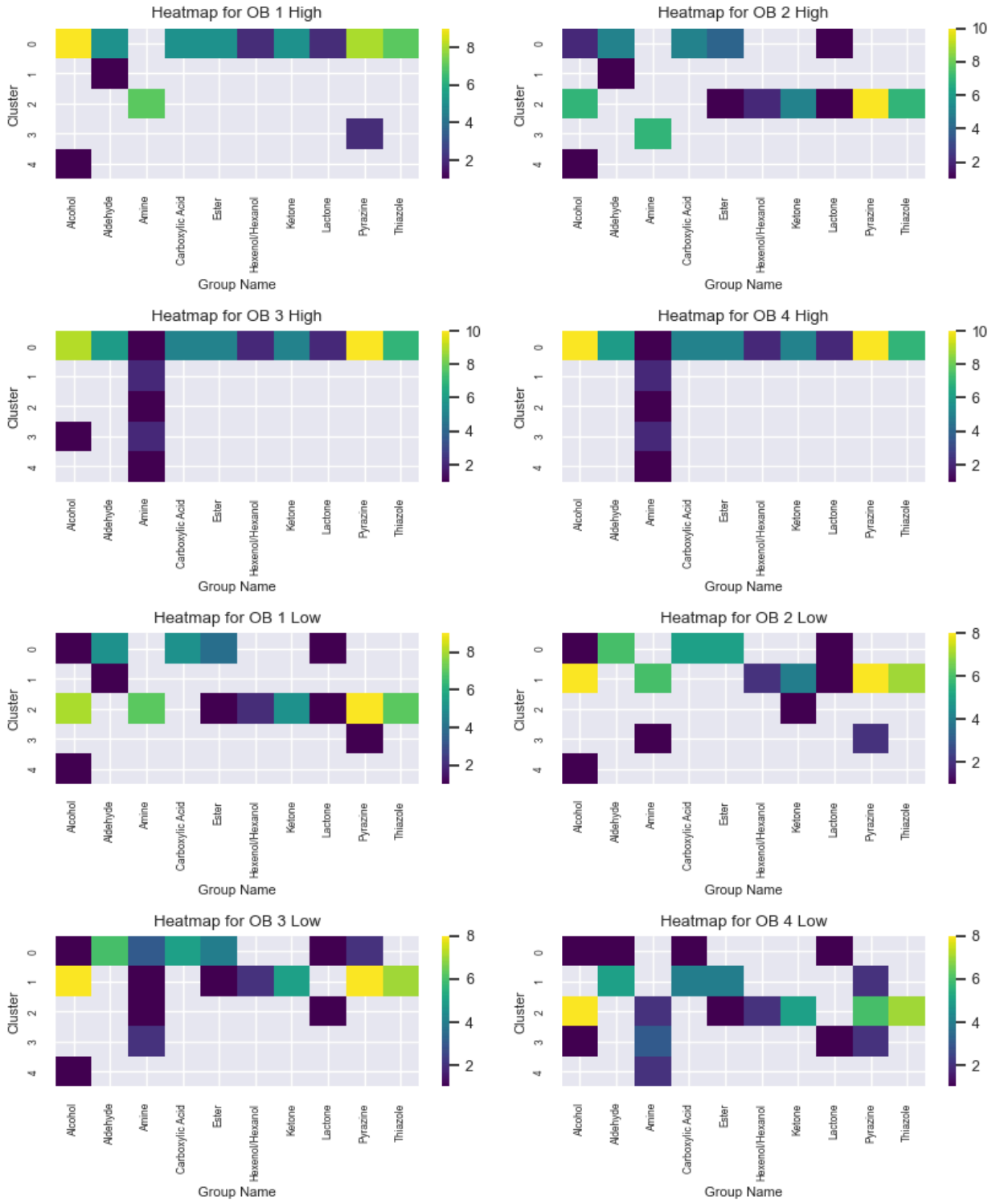

Figure 7. A heatmap showing pivot table for the predefined group and the cluster label assigned to it.

Predictive Algorithm:

We found that a Random Forest Classifier was the best at predicting the cluster labels at the OBs separated by concentration. We divided the data into two matrices one where we concatenated the high concentration data and one with the low concentration and their respective labels that were shown in Figures 6 and 7. We created four custom tests where we mixed and matched the trained classifier to the test data of the other concentration separated through stratified shuffle separation to balance the different cluster sizes in the data and found that the predicted accuracies seen were for High-High 0.813, High predicting Low 0.291, Low predicting High 0.375 and Low predicting Low with 0.563 accuracy (see figure 8), with chance level being 0.2 since we have 5 labels. We did expect differences by virtue of the fact that the datasets are in

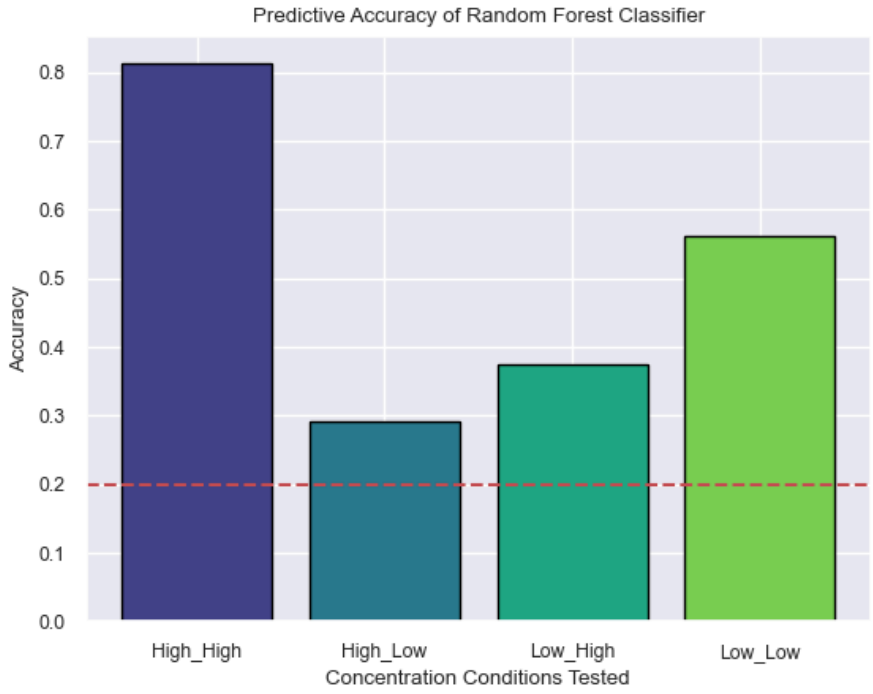

Figure 8. Predictive accuracy score of the Random Forest Classifier for different tests.

fact different datasets. The higher accuracy in the first test seems interesting since the higher concentration OBs due have more nonzero values and could improve accuracy when compared

to the lower concentrations. However, the lower accuracy in the mixed concentration tests, far below the lower concentration to a hierarchical difference that seems to somewhat agree to tthe observed co-clustering obtained by the linkage matrix and seen in the heatmap in figure 7. Moreover, looking at the confusion matrices can elucidate more information about the performance of the algorithm.

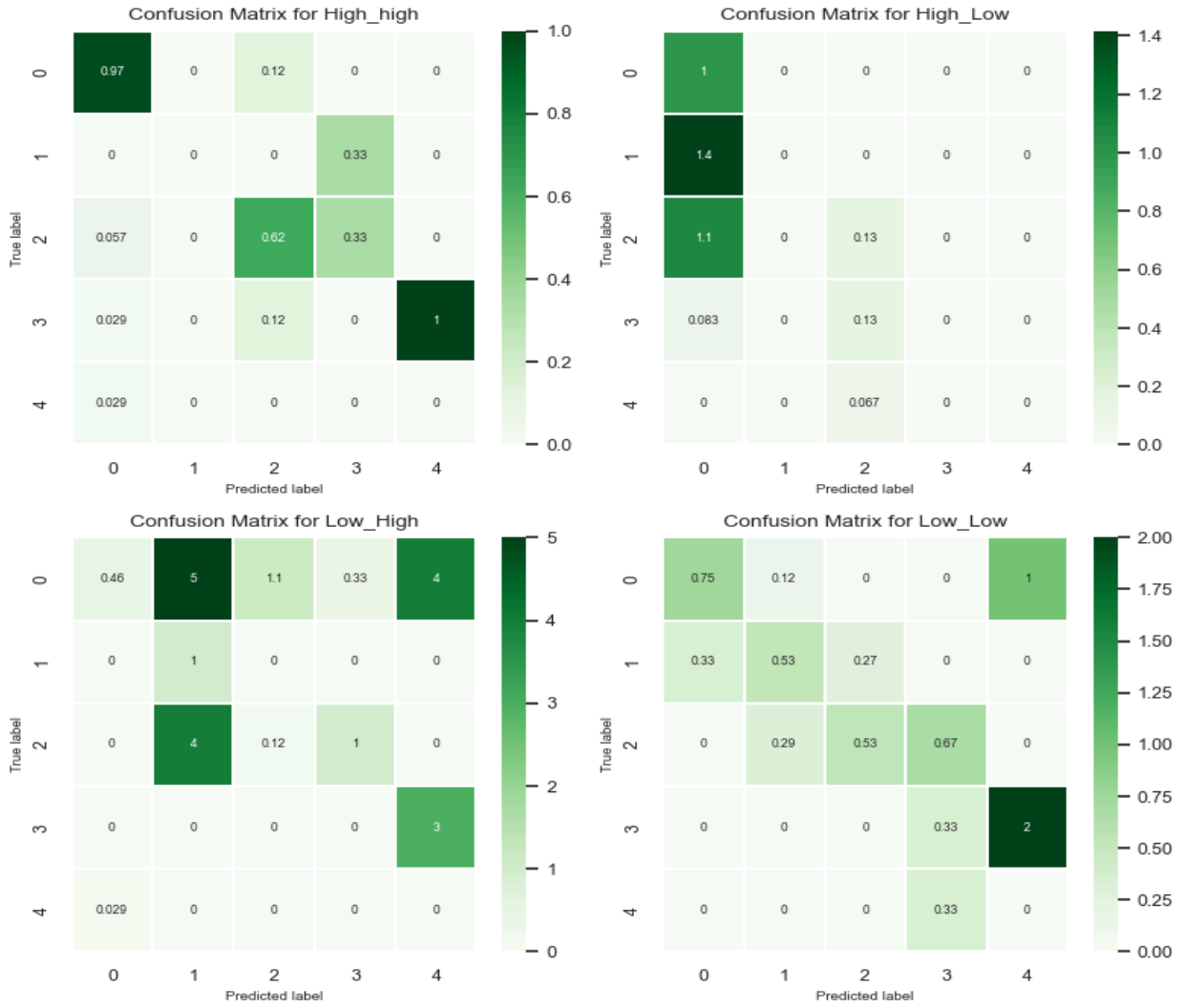

Figure 9. Confusion matrix for the combined tests from the Random Forest Classification.

While the labels slightly shifted during hierarchical clustering they were fairly more consistent within the concentration groups and the confusion values around the diagonal may

reflect the labelling problem. Also we see through the classification reports that the support for clusters the larger clusters are much higher and they will always skew the classifier in their direction. However in the lower concentration test we see that the diagonal is more consistent that can reflect the sparser nature of the lower concentration glomerular activity in the OB.

We ran another classifier but with the original group provided by the Wachowiak Lab, as a control to the possible confound of clustering. We saw similar trends in predictive accuracy as well as in the confusion matrix classification (see figure 10 and 11). Indicating that the trend is present in the data to similar magnitudes despite the possible confounds of the data.

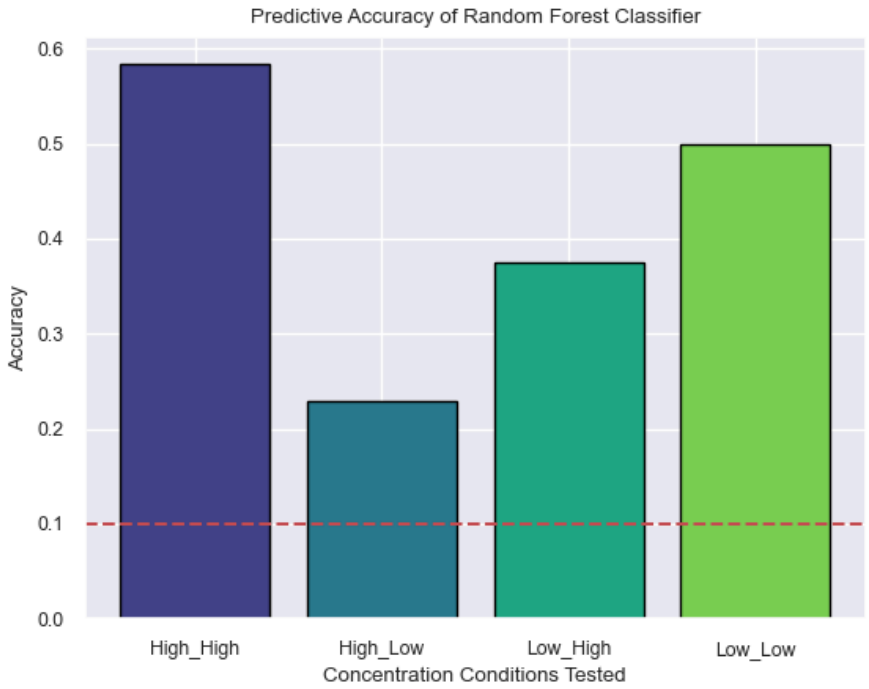

Figure 10. Predictive accuracy for the original unclustered data using the labels provided by the group names, obtained from the Random Forest Classifier

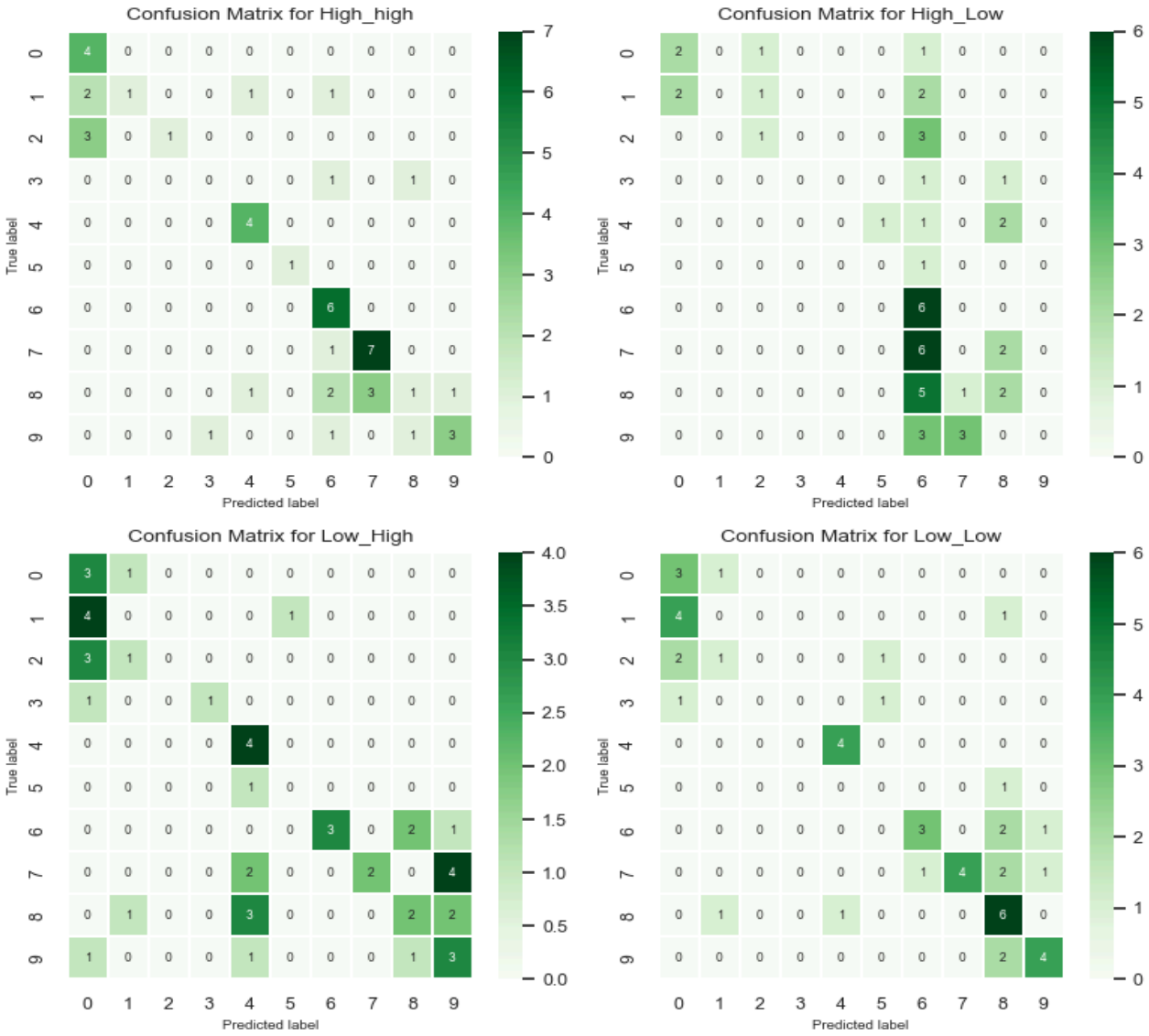

Figure 11. Confusion matrix for the original unclustered data using the labels provided by the group names obtained from the Random Forest Classifier.

Ultimately, the difference in predictive accuracy could still inform a possible inherent difference in the structure and spread of chemical groups of odorants on the OB surface across the different concentrations. It is important to note that these findings are more exploratory in nature and do not precisely prove a true distinction between the two groups due to the limitations of the classifier as a tool to compare these two different conditions.

Discussion:

In this project we had two goals from the beginning: firstly, to find an optimal way to represent the data from the OB and the glomerular activation in each glomerulus therein in a hierarchical way to demonstrate the relationships within the bulb; and secondly, to test how the relationships between the OBs in different concentrations change from each other building on the research done by the Wachowiak lab.

In the field of olfaction, we see an increasing trend of more sophisticated machine learning techniques implemented to combat the inherent complexity that is a feature of the Olfactory system. From the interactions of OSNs with the odorants, we already have a multidimensional space since they are mediated through g-protein coupled receptors making the calculations of odorant features and representation a complicated and onerous task from its genesis and moving upstream to the OB we have a vast and sparse space that responds to thousands of odorants. In this study, we are only looking at the cortex of the OB, but we know that OSNs project both in the cortex and medially into the bulb (Nagao et al., 200) marking an unexplored space yet to be added into our information. Moreover, the lateral communications within glomeruli are another level of complexity still in need to be parsed out in our context. Therefore, unsupervised tools such as hierarchical clustering offer valuable advantages in understanding such complicated cases.

We planned to use the labels obtained by hierarchical clustering in order to redefine the data and chemical groups represented in the OB, as opposed to the groups given to us by the Wachowiak lab. These labels would inherit the advantages of the clustering algorithm and we test for a difference between labeling in both high and low conditions using a Random Forest

Classifier to see that the categorisation is different for both concentrations. However, these techniques come with their own technical limitations as well.

In the beginning, we tested different linkage methods based on euclidean distance. We tried different methods like Ward linkage, and Centroid linkage, as well as the linkages we used with the cosine distance metrics. While we did get higher cophenetic correlations between these distances when we plotted the graph we could see a lot of hierarchical relationships that are quite progressive and simple. The correlation coefficients did look promising (sometimes around 0.9), we judged that because of the dendrograms exhibiting this relationship and the fact that cosine distances between clusters offer more advantage in sparse and higher dimensional data we decided to go forward with cosine distance metrics for the rest of our linkage methods. Unsupervised learning methods sometimes require human judgment to watch out for cases of overfitting the linkages, since hierarchical clustering is a greedy algorithm and will always bring about a relationship in the data. Therefore, we decided to implement cosine distance due to its advantages with data of our nature and obtained although slightly lower cophenetic correlations, the linkages seen in figure 6, do show a more nuanced nature of our data, reflecting the multiple inputs of chemical classes that we saw with our raw data from the Wachowiak lab.

Cosine distance calculation required some amount of troubleshooting due to the nature that the Scipy library's *pdist* function in python calculates such distances. All distance matrices, except for OBs 1 and 2 high, contained negative or very small values, where zeros should have been; while in R the method from Proxy library is different as it calculates cosine distance differently eliminating the low values. We tried different methods of calculating cosine distance using cosine similarity in scikit-learn and even manually input a function using the cosine distance formula, ultimately we decided to import the distance matrices from R, since we ran the

mantel tests with them for continuity and consistency. These distance matrices were used in calculating the cophenetic correlation coefficient for different linkage methods as the distance matrix to compare the clustered data. For the rest of the analyses, we decided to add a small epsilon value of 1e-10, plus a random normal number of magnitude of 1e-8 to be able to work with our data and troubleshoot the tests that we carried out. A future study will have to consider the implementation of cosine distances when dealing with similar datasets.

In order to determine the optimal number of clusters we determined that it was best to look at the within-cluster sum of squares as a measure of internal cluster variation. The idea being that picking a point at the elbow of the curve would allow us to find a number of clusters where picking any more clusters would not give us much more meaningful information within clusters. Theoretically, we could pick up to 59 clusters and that would give us a cluster per odorant but that would defeat the purpose. We looked at an amount of clusters smaller and greater than the original 10 groups in which the data was organised into but found in most methods no meaningful information for clusters greater than 10. Considering the nature of the data we expected less clusters. Due to the amount of information and the general size of the data it is hard to support all groups as stand-alone clusters. We would require more OBs imaged to increase the power of the clustering algorithm in order to do so. We attempted multiple tests in both Python and R by looking at clustering scores such as Silhouette Scores, Davies Bouldin Score, Calinski-Harabasz and also the Gap Statistic in order to complement our Elbow plot where we looked at cosine distance clustering. However, the graphs obtained would behave abnormally when dealing with our data, making us interpret their results with more caution. They showed fluctuations between clusters as well as between datasets making them unreliable in the

context of this dataset. Further work would be required to infer an optimal amount of clusters that way.

In the second part of our project, we wanted to inform our understanding of the distribution of chemical classes as pertaining to the two different stimulus concentrations by employing a classifier trained and tested on two different sets of data separated by concentration. We discovered that: when the high data predicts high and low data predicts low, the accuracy is higher than when we tested low predicts high and vice versa. In the scope of this project, it was difficult to determine a test that could separate the magnitude of the different organisation of the bulb from noise/error that arises by virtue of the classifier predicting different data. As the sample size was smaller than what the classifiers prefer for a powerful classification we refrained from further splitting the data. Another limitation we found was that since the original matrices constituted of different numbers of columns (i.e. glomeruli measured) we had to trim the original matrices for processing before the train-test split due to having NaN values. Since the data is smaller and sparse we refrained at this point from imputing data, but future research could implement a method such as a K nearest neighbor imputer or work with more datasets to avoid this type of data cleaning.

We suspect that part of the difference between the High-Low and Low-High tests come from the fact that clustering as seen in figure 6, slightly rearranged the labels, in the Low concentration we see that what would be cluster 0 is split in half between two clusters, which would lower the predictive accuracy of the classification. We decided to run the classification using the original groups of the dataset provided by the Wachowiak lab to possibly circumvent the classification differences when obtaining cluster labels from the hierarchical clustering task. The classification had similar trends, with the main difference being the High-High and

Low-Low groups having much closer, but overall lower, predictive accuracies. The confusion matrices for this task showed the same trends as the first classification task, even having a more outlined diagonal in the Low trained tests as opposed to the high. Perhaps the noisier behavior seen with the high datasets, such as having a large and greedy cluster 0, could indicate that in fact the activation of glomeruli in the cortex of the OB is biologically noisy as well. More specifically a noisier version of the low concentration odorants, which were better at predicting higher concentration data than the other way around. These conclusions are more speculative in this phase as there is still a need to understand the difference between the statistical behavior of the classifier versus the biological behavior of the glomeruli in the OB.

In this project, we were able to visually represent the different concentrations and their glomerular activations in the OB by clustering the chemical families together through unsupervised learning tasks and testing said relationship with a supervised learning task. We observed that there are different behaviors for the different concentrations tested that would require in the future further testing. We think that the combination of unsupervised and supervised learning routines in examining the surface of the OB are useful for understanding a very complex and high-dimensional such as the OB and glomerular activation could help further decode a lot of the relationships between odorant co-clustering and the topography of chemical class distribution across the OB. Further refined methods for machine learning tasks on more data could be able to inform us of the inner workings of the olfactory system in a way that is more and more accessible to us.

REFERENCES:

Burton, S. D., Brown, A., Eiting, T. P., Youngstrom, I. A., Rust, T. C., Schmuker, M., & Wachowiak, M. (2022). Mapping odorant sensitivities reveals a sparse but structured representation of olfactory chemical space by sensory input to the mouse olfactory bulb. *ELife*, *11*. https://doi.org/10.7554/eLife.80470

Červený, K., Janoušková, K., Vaněčková, K., Zavázalová, Š., Funda, D., Astl, J., & Holy, R. (2022). Olfactory Evaluation in Clinical Medical Practice. *Journal of Clinical Medicine*, *11*(22), 6628. https://doi.org/10.3390/jcm11226628

Cree, B. A., & Weimer, L. H. (2003). Sensory System, Overview. *Encyclopedia of the Neurological Sciences: Volumes 1-4*, *4*, V4-234-V4-241. https://doi.org/10.1016/B0-12-226870-9/00897-2

Ebrahimi, F. A. W., & Chess, A. (1998). Olfactory G proteins: Simple and complex signal transduction. *Current Biology*, *8*(12), R431–R433. https://doi.org/10.1016/S0960-9822(98)70271-4

Eiting, T., & Wachowiak, M. (2018). Artificial Inhalation Protocol in Adult Mice. *BIO-PROTOCOL*, *8*(18). https://doi.org/10.21769/BioProtoc.3024

Gerkin, R. C., & Castro, J. B. (2015). The number of olfactory stimuli that humans can discriminate is still unknown. *ELife*, *4*. https://doi.org/10.7554/eLife.08127

Nagao, H., Yoshihara, Y., Mitsui, S., Fujisawa, H., & Mori, K. (2000). Two mirror-image sensory maps with domain organization in the mouse main olfactory bulb. *NeuroReport*, *11*(13), 3023–3027. https://doi.org/10.1097/00001756-200009110-00039

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D.,

Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wekselblatt, J. B., Flister, E. D., Piscopo, D. M., & Niell, C. M. (2016). Large-scale imaging of cortical dynamics during sensory perception and behavior. *Journal of Neurophysiology*, *115*(6), 2852–2866. https://doi.org/10.1152/jn.01056.2015

Zhang, J., Pacifico, R., Cawley, D., Feinstein, P., & Bozza, T. (2013). Ultrasensitive Detection of Amines by a Trace Amine-Associated Receptor. *The Journal of Neuroscience*, *33*(7), 3228–3239. https://doi.org/10.1523/JNEUROSCI.4299-12.2013