Dany Raihan (53341608)

Artificial intelligence applications offer higher accuracy and efficiency in the molecule odor-prediction models.

The sense of smell is a complex phenomenon that is critical in many aspects of life. For most organisms, the ability to perceive odor is associated with various survival mechanisms, such as avoiding predators, finding mates, and hunting food (Choi et al., 2014). Apex predators like tigers can hunt prey by detecting the scent of blood from thousands of meters away, while insects like mosquitoes sense carbon dioxide in order to find mammals (Raji & DeGennaro, 2017). According to Buck & Axel (1991), smell is defined as a chemical stimulation followed by a physiological response that can be caused by a single or a combination of odorant molecules. These molecules have odors due to their relatively high vapor pressure and hydrophobicity, which are interconnected with their molecular structure (Herrmann & Wiley, 2010). While studies have found that molecules in a functional group tend to share common odors, there are only a few odors that can be determined that way. Contrarily, some odorant molecules have substantially different odors despite being in the same functional group (Nara et al., 2011). Understanding the role of molecular structure in its correlated smell has always been considered hard due to the diverse origins of different odorant molecules. However, the development of artificial intelligence may allow the analysis of complex molecular structures more reliable. In this paper, I argue that artificial intelligence applications offer higher accuracy and efficiency in the molecule odor-prediction model. An encouraging finding by Sharma et al. (2021) revealed strong support for the idea that high accuracy prediction can be achieved using a combination of neural network algorithms. Accordingly, Shang et al. (2021) also found that machine learning models, particularly support vector machine, have promising prospects for accurate and rapid prediction of the smell of a molecule.

A major advantage of using artificial intelligence in the prediction of molecular odors is the high level of accuracy that can be achieved. Various limitations in identifying the relationship between odor and its molecular shape using conventional methods lead to the opportunity for new olfaction research using new technology such as e-nose, neural networks, and machine learning. Machine learning algorithms were trained using large datasets of molecular odor profiles. These models are able to comprehend complicated input, which allows them to predict the smell of a specific molecule with extreme accuracy. The result of the study done by Sharma et al. (2021) revealed supportive evidence for the application of artificial intelligence, notably neural networks. They found that the combined architecture of the neural network models predicted 64 odor samples with perfect precision. These studies show a strong support for artificial intelligence in finding the precise correlation between molecule properties and their associated odor.

Artificial intelligence prediction models also offer a remarkable level of efficiency. Compared to conventional, time-consuming methods that require extensive chemical analysis, artificial intelligence models can make predictions considerably faster. This allows researchers to screen more molecules for their odor properties. For the last two decades, machine learning models have become more prevalent in the field of olfactory studies, such as the physiology and identification of odors (Genva et al., 2018). Based on the result of a study by Shang et al. (2021), trained artificial neural network prediction models can predict the odor of a molecule within seconds. These models also established an average accuracy of R^2 of 0.8807. This result implies that very efficient artificial intelligence models can potentially revolutionize the field of odor prediction.

Another benefit of using artificial intelligence models is the capability of identifying complicated patterns in molecular shape. One of the well-known chemical notations, the simplified molecular input line entry system (SMILES), provides a way to represent chemical

structures in an ASCII-string, a computer-readable format. This notation allows scientists to analyze complex molecular structures using more effective computational power. Shang et al. (2021) gathered SMILES of odorant molecules from NCBI (National Library of Medicine) in accordance with their CAS (Chemical Abstracts Service) number. After computing the physicochemical parameters of SMILES strings, they found that out of 5270 molecular parameters (MP), only 1006 MPs are related to the odor features of a molecule. Due to an imbalanced distribution of odor samples where a particular odor dominates other classes quantitatively, synthetic minority oversampling techniques (SMOTE) were performed. It was discovered that to balance the majority class, the minority class needed to be oversampled by a 3:1 ratio. Support vector machine classification algorithm models along with extreme learning machine models were then implemented to these parameters and the SMILES data. After validating the models, Shang et al. (2021) discovered that odor prediction models calibrated by SVM demonstrated higher accuracy relative to other models, with an accuracy of 97.19% ± 0.93%. This significant result in this study reflects the capability of artificial intelligence models in processing complex molecular parameter data effectively. By accounting for more factors, artificial intelligence can provide more efficient and comprehensive predictions of molecule odor, improving the reliability of the prediction model.

One of the criticisms of artificial intelligence models is their lack of transparency. It can be difficult to understand how a model arrives at a particular prediction, which can be problematic in fields where transparency is important. Monett et al. (2020) stated that artificial intelligence is indeed a contested concept that lacks clear consensus. Additionally, artificial intelligence models can be highly complex and difficult to understand, making transparency challenging. It may not always be possible to provide a complete explanation of how a model makes decisions, particularly when using more advanced deep learning models. However, a systemic standardization of explainable artificial intelligence (XAI) has emerged to address this transparency issue. XAI can be used to make artificial intelligence models more transparent and explainable. XAI examines each feature and its impact on the

result. This analysis helps to develop new scenarios and comprehend how altering the input values affects the result. This analysis also enables researchers to understand and trust the output produced by machine learning algorithms. Gunning & Aha (2019) stated that XAI methods address two relevant challenges, which are data analytics and trust. By using XAI methods, it might be possible to understand how an artificial intelligence model arrives at a particular prediction, especially in an artificial intelligence-powered odor prediction model.

In conclusion, artificial intelligence can be used to predict the odor of a molecule based on its molecular structure. Artificial intelligence can process complex molecular data more effectively. Although there might be some transparency issues, the development of XAI may gradually make artificial intelligence models more transparent and explainable. Artificial intelligence also allows data-driven approaches to explore a deeper understanding of the relationships between the chemical structure and its related odor. In order to more effectively evaluate the connection between a molecule and its structure, future research should focus on building more advanced and transparent machine learning algorithms for odor prediction models.

References Cited

Block, E. (2018). Molecular basis of mammalian odor discrimination: A status report. *Journal of Agricultural and Food Chemistry*, 66(51), 13346-13366. https://doi.org/10.1021/acs.jafc.8b04471

Buck, L., & Axel, R. (1991). A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell*, 65(1), 175-187. https://doi.org/10.1016/0092-8674(91)90418-X

Choi, N., Han, J. H., & Wiley Online Library. (2014). *How flavor works: The science of taste and aroma*. John Wiley & Sons Inc. https://doi.org/10.1002/9781118865439

Genva, M., Kenne Kemene, T., Deleu, M., Lins, L., & Fauconnier, M. (2019). Is it possible to predict the odor of a molecule on the basis of its structure? *International Journal of Molecular Sciences*, 20(12), 3018. https://doi.org/10.3390/ijms20123018

Grabe, V., & Sachse, S. (2018). Fundamental principles of the olfactory code. *Biosystems*, 164, 94-101. https://doi.org/10.1016/j.biosystems.2017.10.010

Gunning, D., & Aha, D. W. (2019). DARPA's explainable artificial intelligence (XAI) program. *The AI Magazine*, 40(2), 44-58. https://doi.org/10.1609/aimag.v40i2.2850

Herrmann, A., & Wiley Online Library. (2010). *The chemistry and biology of volatiles*. Wiley. https://doi.org/10.1002/9780470669532

Monett, D., Lewis, C. W., & Thórisson, K. R. (2020). Introduction to the JAGI Special Issue "On Defining Artificial Intelligence"—Commentaries and Author's Response. *Journal of Artificial General Intelligence*, 11(2), 1–100. https://doi.org/10.2478/jagi-2020-0003

Nara, K., Saraiva, L. R., Ye, X., & Buck, L. B. (2011). A large-scale analysis of odor coding in the olfactory epithelium. *The Journal of Neuroscience*, 31(25), 9179-9191. https://doi.org/10.1523/JNEUROSCI.1282-11.2011

Raji, J. I., & DeGennaro, M. (2017). Genetic Analysis of Mosquito Detection of Humans. *Current opinion in insect science*, 20, 34–38. https://doi.org/10.1016/j.cois.2017.03.003

Shang, L., Liu, C., Tomiura, Y., & Hayashi, K. (2017). Machine-learning-based olfactometer: Prediction of odor perception from physicochemical features of odorant molecules. *Analytical Chemistry (Washington), 89*(22), 11999-12005. https://doi.org/10.1021/acs.analchem.7b02389

Sharma, A., Kumar, R., Ranjta, S., & Varadwaj, P. K. (2021). SMILES to smell: Decoding the Structure–Odor relationship of chemical compounds using the deep neural network approach. *Journal of Chemical Information and Modeling, 61*(2), 676-688. https://doi.org/10.1021/acs.jcim.0c01288

Zhang, L., Mao, H., Zhuang, Y., Wang, L., Liu, L., Dong, Y., Du, J., Xie, W., & Yuan, Z. (2021). Odor prediction and aroma mixture design using machine learning model and molecular surface charge density profiles. Chemical Engineering Science, 245, 116947. https://doi.org/10.1016/j.ces.2021.116947