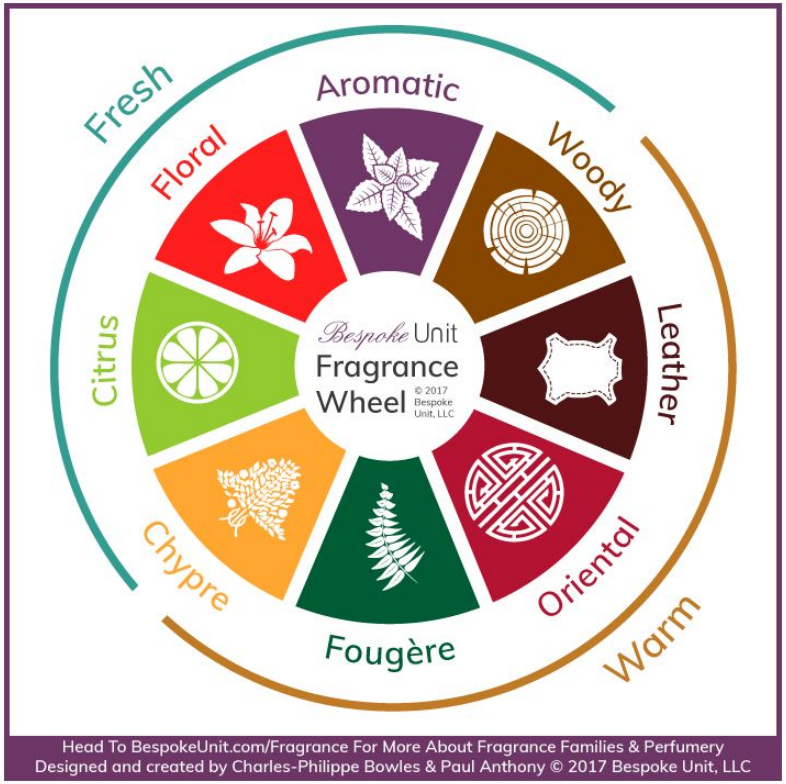
Collection of Gas Chromatography (GC) Signals for the Development of a Scent Categorization and Description Algorithm

Alexandra Driehaus, Dr. Dobrokhotov, Dr. Novikov, and Mr. Matthew Pimienta

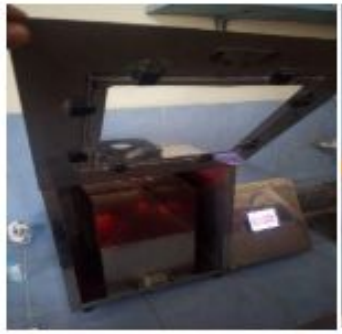
Western Kentucky University
Department of Physics and Astronomy

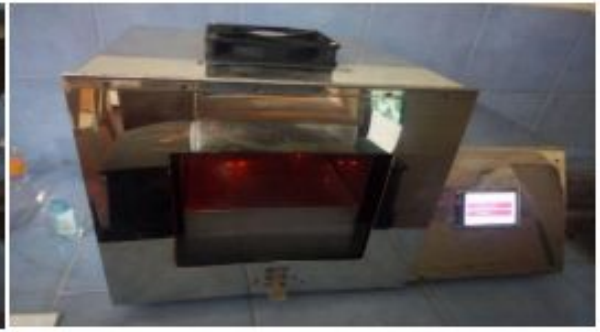
Introduction

- Scents are grouped into families based on their verbal descriptions.
  - These families include citrus, aromatic, woody, oriental, green, leather, and floral.
- In his *Atlas of Odor Character Profiles* (1985), Andrew Dravnieks found the verbal descriptions of 160 chemicals
  - He surveyed about 140 people with 146 descriptors for each chemical
    - Cinnamon was found to have 76% applicability for Cinnamon Bark Oil
- These descriptors can be used to group the chemicals into their scent families.



Past Work





Many are working to connect quantifiable data to scent.

- *Development of an Electronic Nose for Olfactory System Modelling using Artificial Neural Network* (2018) Roa and Fernandez
  - Applied an e-nose to general odor classification
- *Analysis of Fragrances in Cosmetics by Gas* (1995) Spectrometry and Rastogi
  - Applied gas chromatography techniques to analyze different cosmetics, including lotions and perfumes

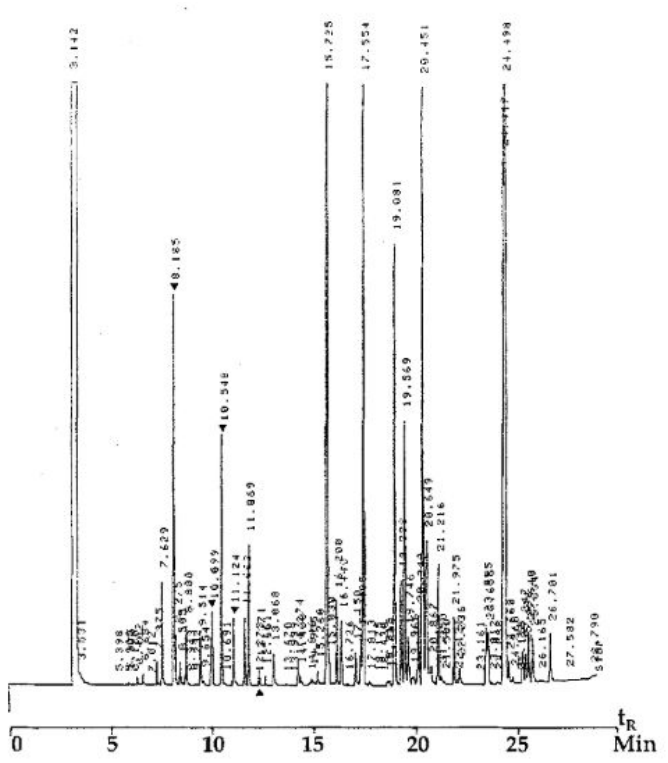
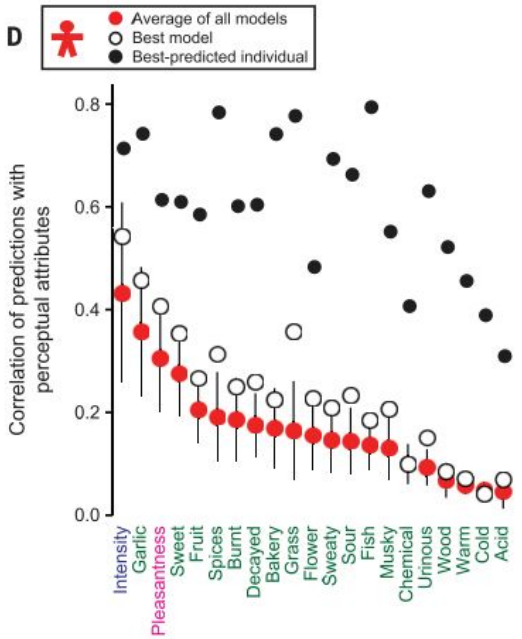


Figure 3. Analysis of the target fragrance substances in an eau de toilette (3-1714) by GC-FID. t_R 8.815 – internal standard (citronellal), 10.099 – geraniol, 10.548 – hydroxycitronellal, 11.124 – cinnamic alcohol, 12.371 – eugenol.

Past Work (Cont.)

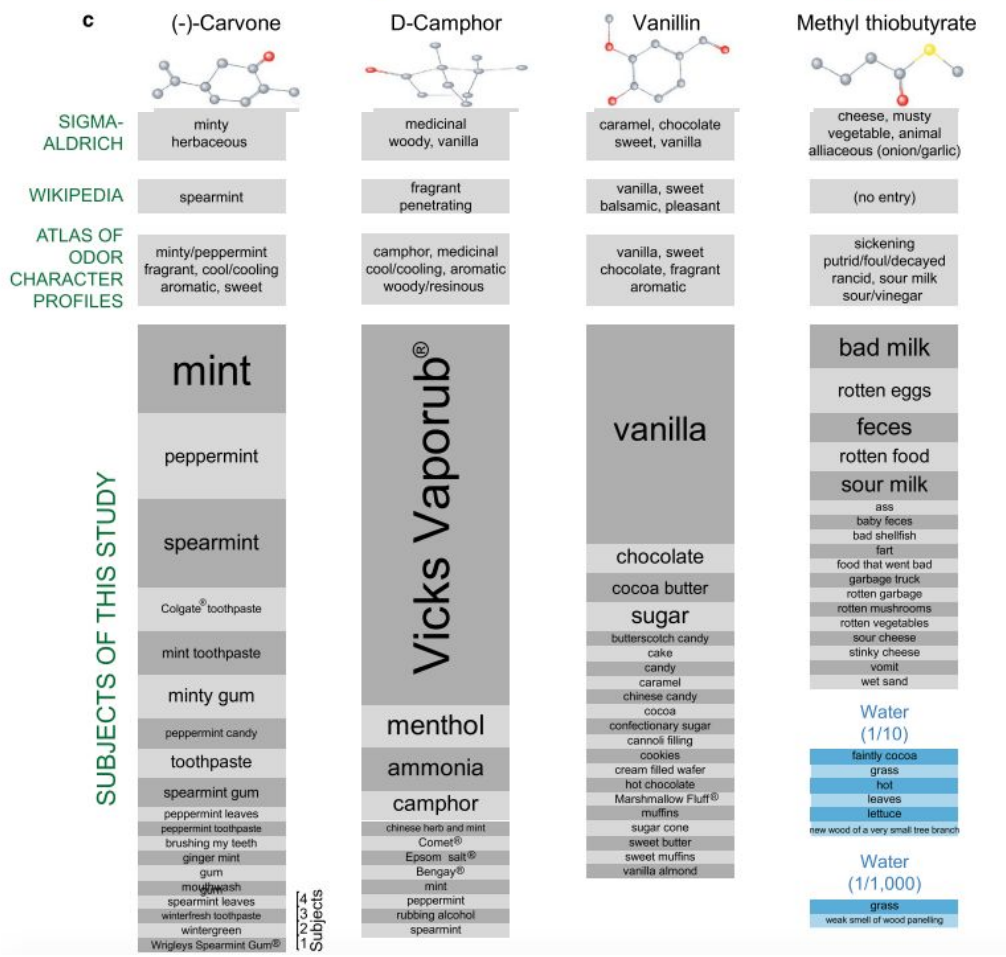
Keller, Gerkin, Guan, et.al.'s "Predicting Human Olfactory Perception from Chemical Features of Odor Molecules"

- Keller's group worked to predict how humans would perceive scents based off the chemical features
- There is still no way to fully determine based of appearance alone how a human olfactory system will perceive a given chemical scent
- In order to find similarities in scent, known fragrant chemicals must be found to exist in multiple samples.



Past Work (Cont.)

- *Olfactory perception of chemically diverse molecules* (2016) Keller and Vosshall
  - Tested 480 chemicals with 55 humans to show the correlation between chemical structure and verbal description

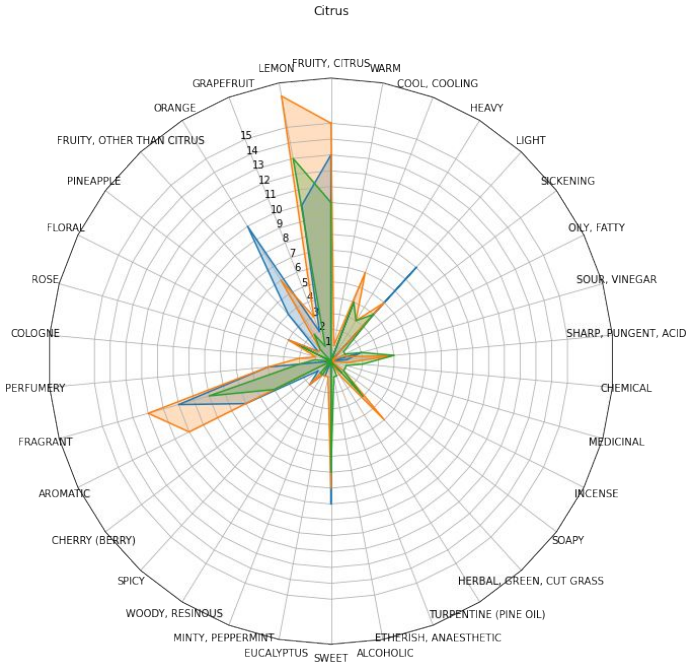


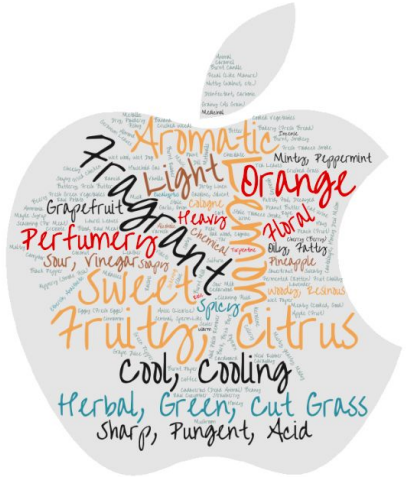
The Problem

- How to generate a verbal description of scent without human interaction, to show the relationship between chemical structure and human perception
- Approach:
  - Take the scents from the Atlas and categorize them into their scent families
  - Generate their respective chromatograms through Multisensory Gas Chromatography (MGC)
  - These will be the base for the convolutional neural network
  - Generate chromatograms for unknown scents
  - Use these to determine the scents of unknown scents
- We seek to measure chromatograms and provide verbal description of scents.

Scent Family: Citrus

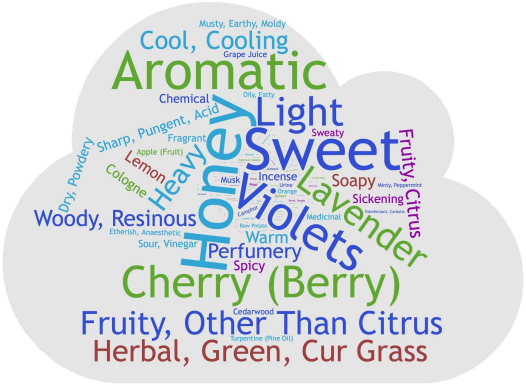
- Three chemicals from the Atlas were found to be consistent with the citrus family description.
  - These were:Limonene, Citral, and Citralva
- Citrus scents, for example, have notes of lemon, orange, and grapefruit.
- These are plots of the applicability of the family.

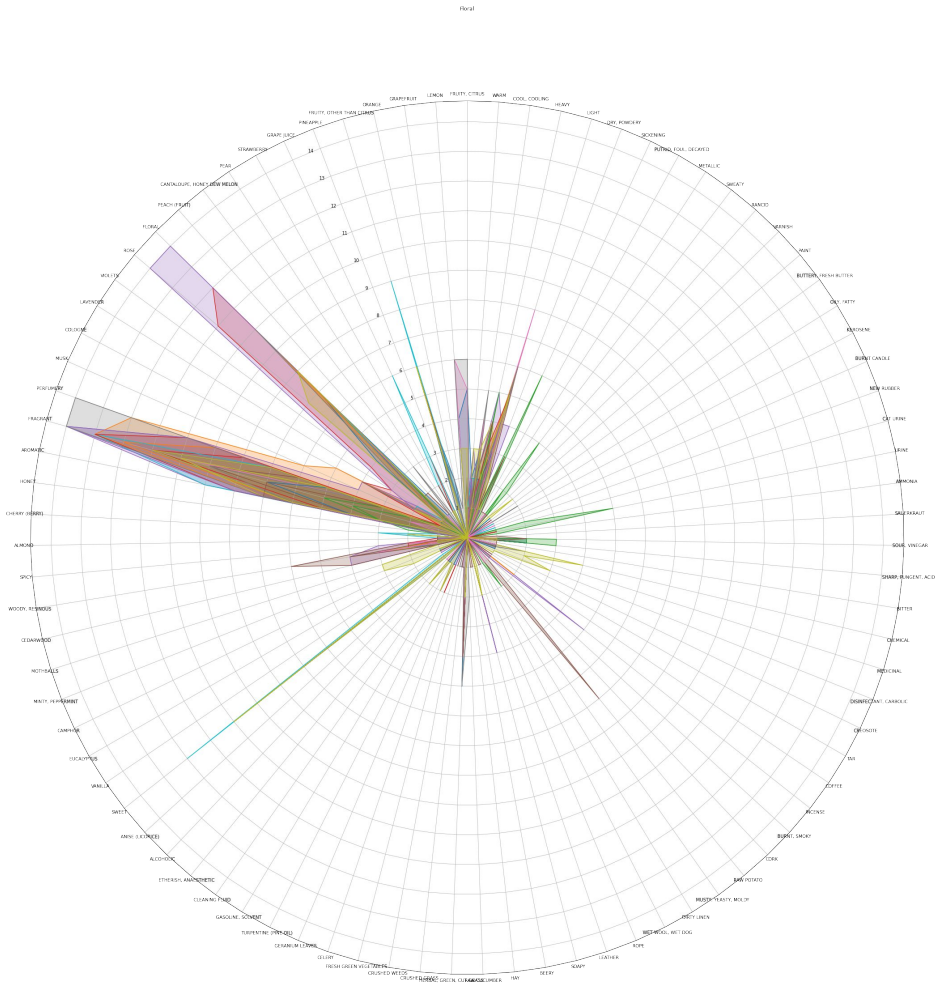




Scent Family: Floral

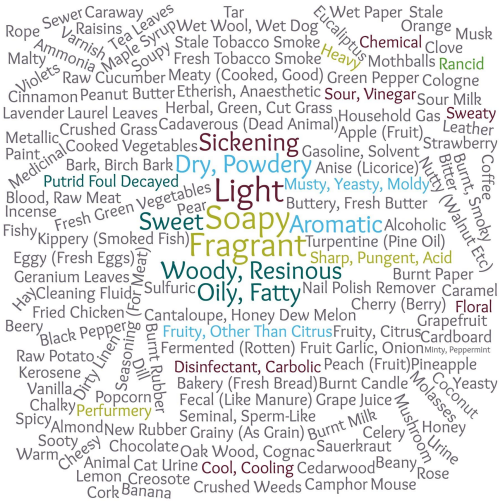
- Eighteen chemicals from the Atlas were found to be consistent with the floral family description.
- Floral scents, for example, have notes of roses and lavender.
- These are plots of the applicability of the family.

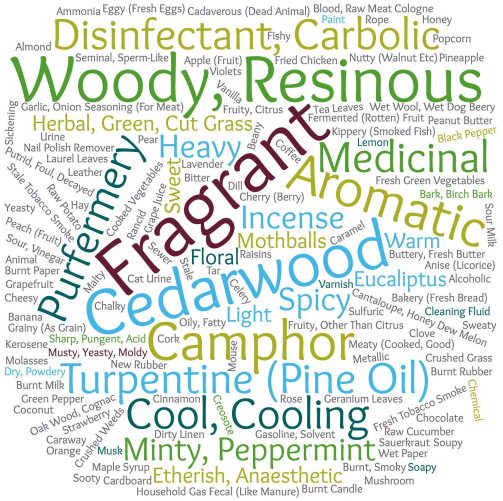




Scents from the Same Family

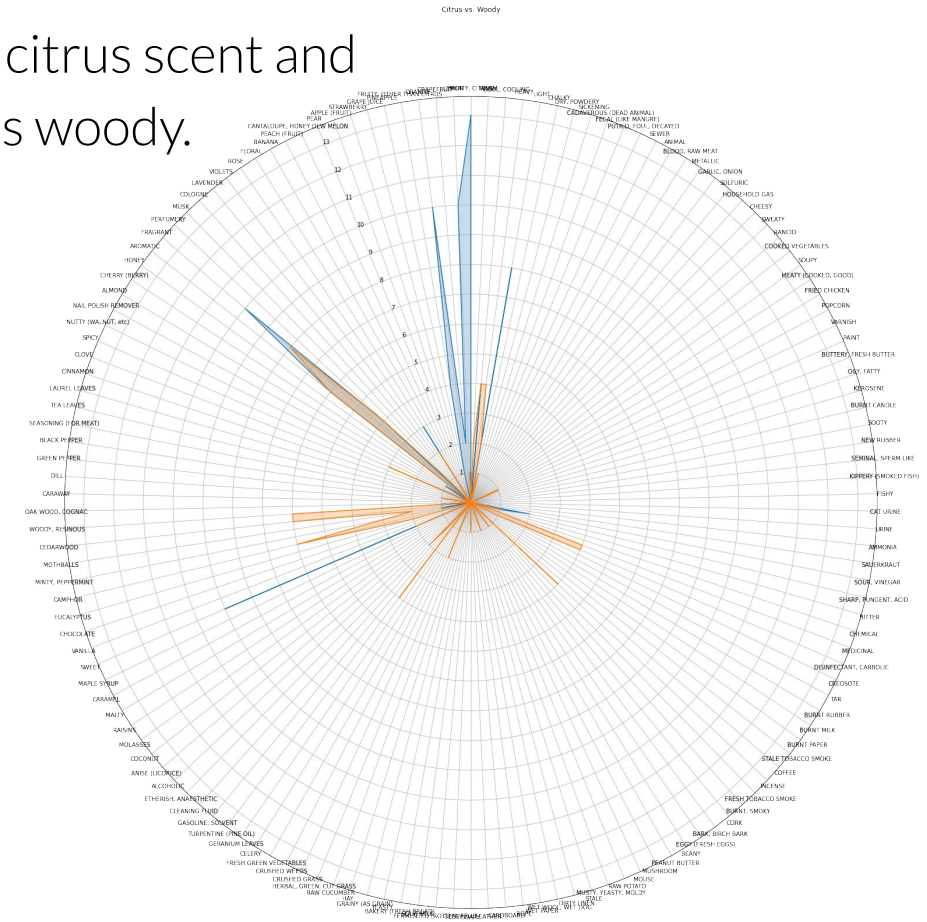
- The verbal descriptions of scents from the same family are similar. These are both chemicals from the woody family: Oneatic Ether (Top) and Patchouli Oil (Bottom)
- The woody family often has notes of cedarwood and patchouli.





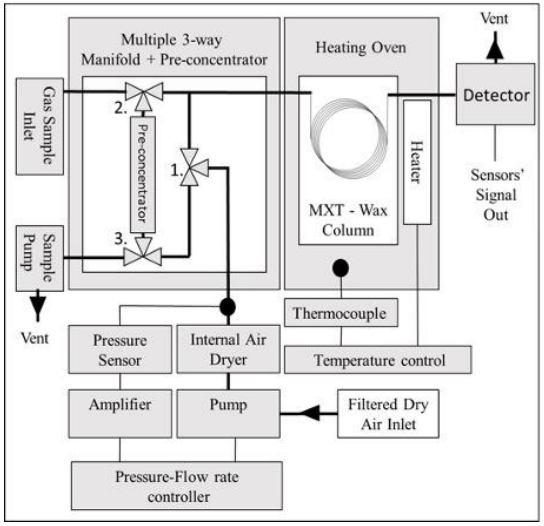
Limonene and Patchouli Oil

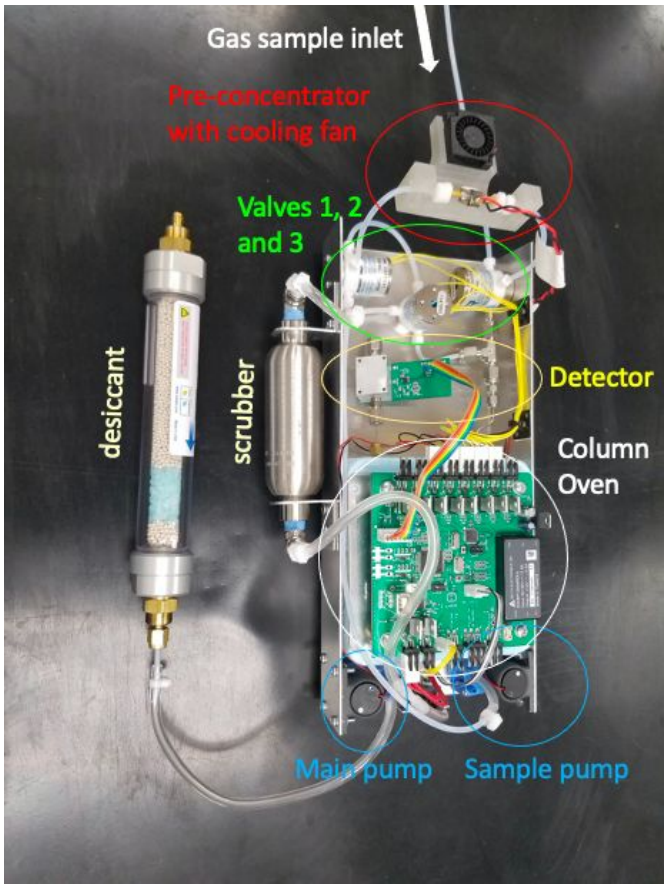
- Limonene is a citrus scent and Patchouli Oil is woody.



Gas Column Chromatography

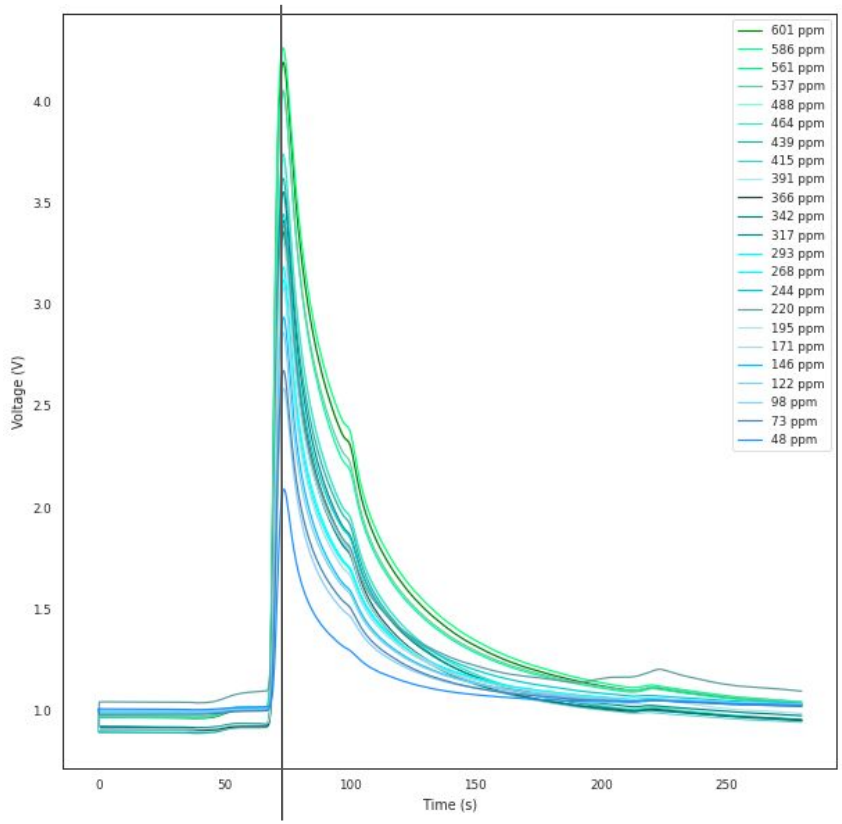
- GCC collects data by using the apparatus to analyze scented chemicals
  - Creates a chromatogram for each scent





Preliminary Experimental Results

- Chromatographic data can be used to determine chemical structure.
  - The higher the concentration, the higher the peak.
  - The larger the molecule, the wider the peak.
  - The heavier the molecule, the higher the retention time
    - Retention time defines the position of the peak on the time axis.
  - Some characteristic peaks are distinguishable at concentrations higher than the detection limit.



Retention Time

Chromatography and Scent

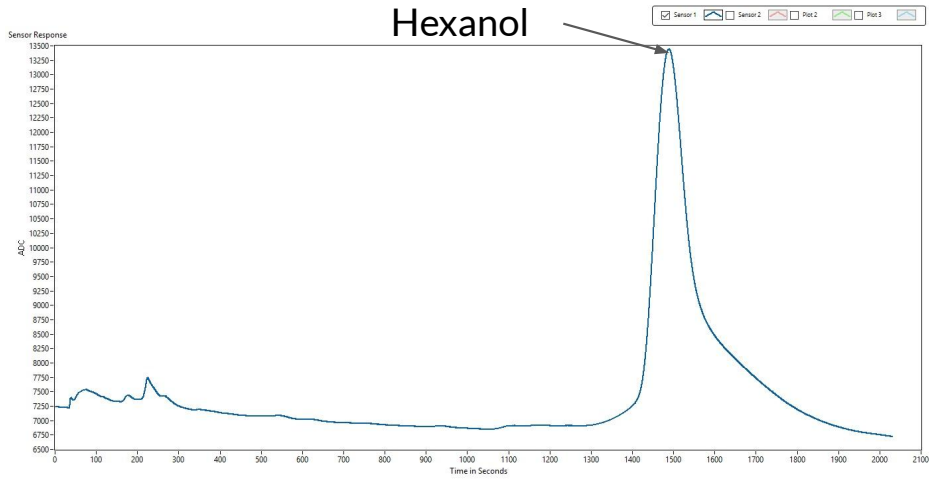
Three chemicals were chosen from each of three scent families for the gathering of known data:

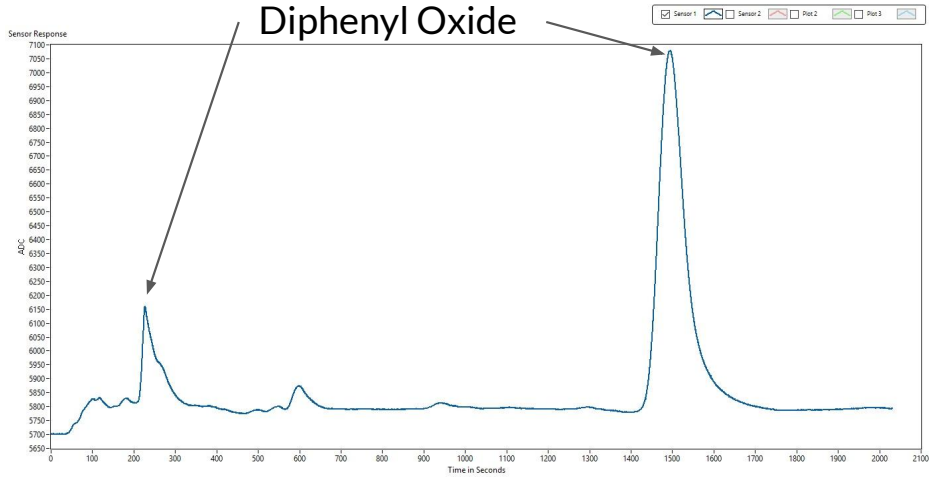
- Green
  - 1-Hexanal, Diphenyl Oxide, and Hexanol
- Aromatic
  - (-) Menthol, Eucalyptol, and Methyl Salicylate
- Floral
  - Acetophenone, Hydroxy Citronellol, and Phenyl Ethanol

Experimental Data - Same Scent Family

These are signals from the green family.

- As shown, the peaks have different retention times and widths.
    - Unlabeled “peaks” fall below the detection limit, and thus are just noise.
- Within the same scent family, our data shows similar retention times.
    - In order to solidify this statement, more tests would need to be run with a wider range of samples.

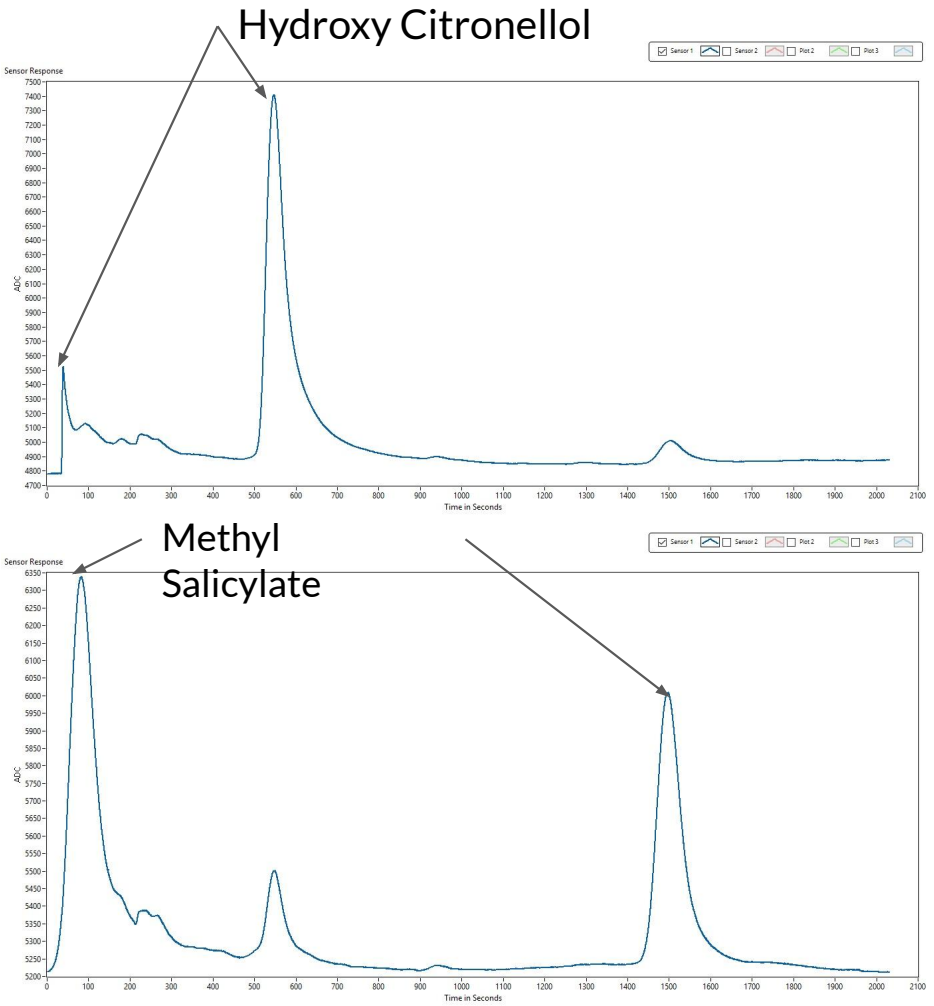


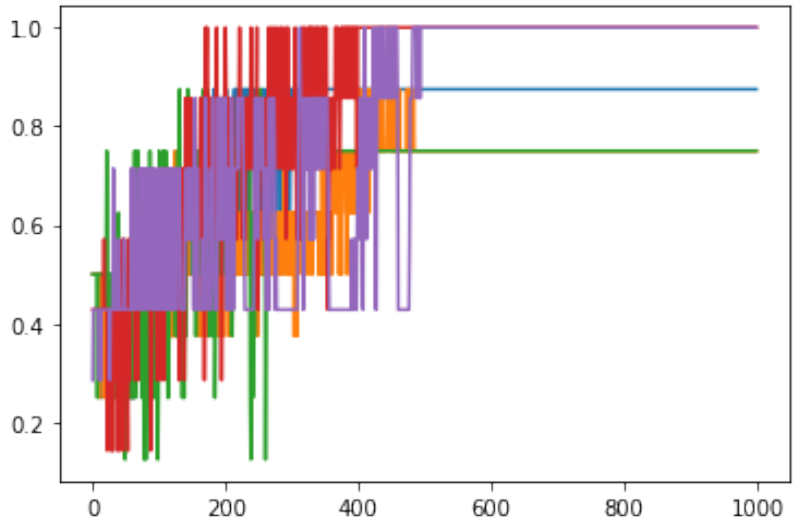


Experimental Data - Different Scent Families

These are signals from the floral (top) and aromatic (bottom) families.

- As shown, the peaks have different retention times and widths.
    - Unlabeled “peaks” fall below the detection limit, and thus are just noise.
- There are multiple dominant peaks for the aromatic scent.
- Within the different scent families, our data shows ranges of retention times for each family.
    - In order to solidify this statement, more tests would need to be run with a wider range of samples.





Validation run 5 times

- On each validation run, 1000 epochs (trains on the same data 1000 times)

Average accuracy of 87.5%

Input is voltage curve

- 1000 evenly-spaced points taken from that curve
- Deep learning model performs a series of transformations on the input to learn high-level patterns and ultimately give the prediction for scent family
    - Model based on 1 dimensional convolutions

Future Work

- Formulate the algorithm
- Perform the experiment (again)
  - With unknown scents
  - Generates a verbal description
- Test the validity of the experiment
  - Check that the generated description matches the manufacturer's description

Acknowledgements

This project was funded by KY EPSCoR.