

**CAPSTONE PROJECT REPORT**

**ELM 5523**

**MODELING OLFACTORY FOR VIRTUAL REALITY SYSTEM**

***Submitted on:***

27th July 2025

***Prepared by***

Dawson McClary

Chris McReynolds

Candice Reyes

**Engineering Leadership and Management M.S**

**University of Oklahoma**

**Summer 2025**

# abstract

This project addresses the critical gap in virtual reality (VR) therapy for post-traumatic stress disorder (PTSD) by developing an integrated olfactory feedback system that enhances multisensory exposure through precise scent delivery. The system combines a machine-learned classification model, which predicts clinically relevant odor profiles from mass spectral data of essential oils, with a modular, wearable scent-emitting device designed for seamless integration with VR headsets. Using semantic clustering and advanced neural network architectures, the model effectively categorizes trauma-associated scents such as smoke, gunpowder, diesel, and burnt rubber, achieving high accuracy and robustness. Concurrently, the hardware design incorporates systems engineering principles to ensure safety, reliability, and clinical applicability, featuring micropump-driven scent delivery, sterilization mechanisms, and modular components. This multidisciplinary approach integrates chemical informatics, machine learning, and engineering design to overcome existing technological limitations in olfactory VR therapy. The project’s outcomes demonstrate a scalable, user-centered platform that supports enhanced therapeutic immersion and improved PTSD treatment efficacy. Beyond therapy, the technology enables richer sensory experiences in training, education, and entertainment. Designed ethically with stakeholder input, this platform advances scent-based immersion by leveraging scent’s powerful link to memory and emotion.

[abstract 3](#_Toc204539790)

[LIST OF TABLES 6](#_Toc204539791)

[LIST OF FIGURES 7](#_Toc204539792)

[LIST OF Appendices 8](#_Toc204539793)

[Problem Statement 9](#_Toc204539794)

[Introduction 10](#_Toc204539795)

[LITERATURE REVIEW 11](#_Toc204539796)

[Evaluating Virtual Reality Experiences with Olfactory Integration: A Preliminary Review 12](#_Toc204539797)

[Extension of Wearable Olfactory Display of Multisensory VR Experience 13](#_Toc204539798)

[Automatic Scent Creation by Cheminformatics Method 14](#_Toc204539799)

[Predicting Human Odor Perception Represented by Continuous Values from Mass Spectra of Essential Oils Resembling Chemical Mixtures 15](#_Toc204539800)

[From Meaning to Perception – Exploring the Space between Word and Odor Perception Embeddings 17](#_Toc204539801)

[Olfactory Perception of Chemically Diverse Molecules 19](#_Toc204539802)

[Summary and Gaps in Literature 20](#_Toc204539803)

[Methodology 21](#_Toc204539804)

[Research Background 22](#_Toc204539805)

[Scent Model Designs 22](#_Toc204539806)

[Semantic Linguistic Classification 24](#_Toc204539807)

[Scent Emitter Designs 25](#_Toc204539808)

[Project Scope Considerations 26](#_Toc204539809)

[Alignment with Project Goals 27](#_Toc204539810)

[Analysis 27](#_Toc204539811)

[Data Preparation and Pre-Processing 28](#_Toc204539812)

[Essential Oil and Mass Spectra Analysis 28](#_Toc204539813)

[VORA Device Requirement Analysis 32](#_Toc204539814)

[Applications of Modeling Methodologies 33](#_Toc204539815)

[Predictive Modeling with CNN and MLP 33](#_Toc204539816)

[SysML Architecture and Design Modeling 35](#_Toc204539817)

[Key Findings and Result Validation 36](#_Toc204539818)

[Engineering Design Results 39](#_Toc204539819)

[Support of Capstone Objectives 41](#_Toc204539820)

[Multidisciplinary Leadership and Ethical Design Considerations 41](#_Toc204539821)

[Project Management Analysis 42](#_Toc204539822)

[Conclusion 50](#_Toc204539823)

[Citation Sources 53](#_Toc204539824)

[Appendices 56](#_Toc204539825)

# LIST OF TABLES

[Table 1: Twelve Scent Profiles 29](#_Toc204539747)

[Table 2: Modeling Results 36](#_Toc204539748)

[Table 3: Risk Register Analysis 44](#_Toc204539749)

[Table 4: Project Stakeholder Analysis 49](#_Toc204539750)

# LIST OF FIGURES

[Figure 1: Faceted Line Plots of Mass Spectra Intensities of All Essential Oils 30](#_Toc204539765)

[Figure 2: Line Plot of Mass Spectra Intensities of Correlated Essential Oils 31](#_Toc204539766)

[Figure 3: Fragrance Pyramid46 32](#_Toc204539767)

[Figure 4: Smoke Scent Notes 32](#_Toc204539768)

[Figure 5: Proportion of Variance along Principal Components of Mass Spectra 34](#_Toc204539769)

[Figure 6: Accuracy Per-Cluster Results from CNN with PCA 37](#_Toc204539770)

[Figure 7: Confusion Matrix of Actual versus Predicted Results from CNN with PCA 38](#_Toc204539771)

[Figure 8: VORA Device Concept Design 41](#_Toc204539772)

[Figure 9: Heatmap of Impact and Likelihood of Project Risks 45](#_Toc204539773)

[Figure 10: Monte Carlo Simulation of Project Risk Exposures 46](#_Toc204539774)

[Figure 11: Work Breakdown Structure of Project 47](#_Toc204539775)

[Figure 12: Project Stakeholder Power-Interest Matrix 49](#_Toc204539776)

# LIST OF Appendices

[APPENDIX A: MODEL AND VISUAL GENERATION CODE 57](#_Toc204442509)

[APPENDIX B: SCENT NOTE PROFILES 83](#_Toc204442510)

[APPENDIX C: SYSML DESIGN – VORA REQUIREMENTS – VORA BOM – VORA SYSML MARKUP 83](#_Toc204442511)

[APPENDIX D: DESCRIPTOR VECTORIZATION JUSTIFICATIONS 84](#_Toc204442512)

BACKGROUND

## Problem Statement

Despite the increasing potential of virtual reality therapy for treating post-traumatic stress disorder, current virtual reality systems largely overlook the sense of smell, which is a vital sensory input for emotional processing and memory formation. Existing approaches to olfactory virtual reality face considerable challenges, including a limited variety of scents available, inaccurate odor delivery, and a lack of well-developed scent selection models specifically designed for therapeutic use. These shortcomings reduce the effectiveness of exposure therapy by restricting the patient’s ability to safely and effectively reframe trauma-associated odors within controlled environments.

This project seeks to overcome these challenges by developing a data-driven model that classifies and predicts clinically relevant scent profiles using mass spectral data of essential oils. Additionally, it involves designing a modular, wearable scent-emitting device optimized for seamless integration with virtual reality systems. The solution targets critical trigger odors associated with post-traumatic stress disorder—such as smoke, gunpowder, diesel, and burnt rubber—which remain largely absent from existing scent databases and virtual reality platforms.

In alignment with the principles of engineering leadership and management, this project employs thorough systems engineering methods, including clear requirements tracking, risk assessment, and active stakeholder involvement. These approaches ensure that the resulting device is reliable, user-friendly in clinical settings, and deployed with due consideration for ethical implications in mind. The primary stakeholders in this effort include individuals with post-traumatic stress disorder, healthcare providers, technology developers, and regulatory agencies. Each stands to benefit from a virtual reality therapy platform that is safer, more immersive, and grounded in scientific research.

By integrating expertise from machine learning, chemical engineering, and systems design, this work aims to push forward the development of multisensory therapeutic technologies. It strives to enhance treatment outcomes for post-traumatic stress disorder by creating virtual environments that incorporate realistic and targeted olfactory experiences, thus deepening immersion and therapeutic effectiveness.

## Introduction

This project aims to develop a comprehensive olfactory feedback system designed to enhance virtual reality therapy for individuals who have post-traumatic stress disorder. The system integrates two core elements: a machine learning-based scent classification model capable of predicting clinically relevant odor profiles from mass spectral data, and a modular wearable scent emitter device that integrates seamlessly with virtual reality headsets. These components work together to deliver trauma-associated odors such as smoke, gunpowder, diesel, and burnt rubber with high precision and control. These particular scents are essential for effective multisensory exposure therapy but are largely absent in current virtual reality applications.

The project scope encompasses several critical areas, beginning with data collection and preparation, including the semantic clustering of scent descriptors to construct meaningful profiles for therapeutic use. The predictive modeling component leverages convolutional neural networks to accurately classify these scent profiles from mass spectral data. On the hardware side, the project employs systems engineering principles using the systems modeling language to design a scent emitter device. This design emphasizes modularity, safety, and seamless integration with existing virtual reality platforms. Additionally, comprehensive project management strategies are incorporated, including the use of risk registers, stakeholder engagement plans, and validation procedures, to ensure the system meets clinical requirements and adheres to ethical standards.

The significance of this work lies in addressing a critical gap in the field of virtual reality therapy. Olfactory stimuli play a fundamental role in human perception, emotion, and memory formation. In the context of post-traumatic stress disorder therapy, controlled exposure to trigger scents supports the reappraisal and gradual desensitization of traumatic memories, aiding in emotional regulation and recovery. Despite this, most virtual reality systems do not incorporate the sense of smell due to several technological challenges. These challenges include a limited selection of scent options, physical constraints related to device size and ergonomics, as well as difficulties in delivering scents quickly and accurately within immersive environments. Overcoming these obstacles is vital to unlocking the full potential of virtual reality as a therapeutic tool.

By integrating advanced machine learning techniques, detailed chemical analysis, and robust systems engineering, this project advances the emerging field of multisensory virtual reality therapy. It establishes a scalable and clinically relevant framework for incorporating olfactory stimuli, thereby enhancing user immersion and improving therapeutic outcomes. Beyond its application in healthcare, the innovations developed through this project have broad implications for industries seeking to create more realistic and emotionally engaging virtual experiences.

# LITERATURE REVIEW

Recent research emphasizes the value of multi-modal models that combine visuals, hearing, and smell. Despite limits in scent variety, device size, and delivery speed, OVR (Olfactory Virtual Reality) shows strong potential for treating post-traumatic stress disorder (PTSD), easing anxiety, and expanding multisensory therapy.1 Within virtual reality (VR) environments, patients are exposed to a variety of odors to identify those that serve as specific trauma-related trigger scents.2 Over time, patients develop the ability to remain calm during simulations that include identified trigger scents. These odors are later reintroduced to help reframe their emotional associations, shifting them toward a neutral or less distressing response.2

This literature review found little to no data on odor parameters needed for OVR therapy targeting specific PTSD cases. Key odors—such as diesel, black powder, smoke, and burnt rubber—are largely absent from scent registers and academic research on odor combinations.2 Expanding the range of scents is crucial to developing more immersive and engaging virtual experiences. This literature review highlights the significance of OVR in therapeutic and training contexts, while exploring recent developments in assessing olfactory effectiveness.1 Concepts from each article played a critical part in the analysis performed to outline new pathways of OVR success for PTSD patients in the VR therapy.

## Evaluating Virtual Reality Experiences with Olfactory Integration: A Preliminary Review

VR technology is widely known for delivering immersive audio-visual experiences, yet it often overlooks the olfactory sense, which plays a vital role in human perception, cognition, and emotion.3 Traditional exposure therapy has several limitations, including, but not restricted to, access to stimuli, time constraints, and the inability to replicate certain situations. Olfaction plays a crucial role in creating immersive and realistic virtual environments. As a core part of lived experience, odors evoke strong emotional and visceral responses.4 To integrate olfaction, various odor delivery methods have been developed, typically involving airflow and scent-dispensing mechanisms embedded in the VR headset or a dedicated nosepiece.5

Smell enhances object recognition, spatial attention, and memory and connects directly to the brain’s limbic system, influencing emotions and physiological responses such as fear, pain, and vocal expression. Despite its critical role, olfaction is often overlooked in virtual reality applications, even though it offers clear benefits in therapy, education, and training. Emerging research in olfactory virtual reality (OVR) emphasizes the value of integrating scent into immersive experiences. Supported by artificial intelligence (AI), this integration aligns odors with visual and narrative elements, enhancing realism, emotional engagement, and memory retention.1

OVR shows promising applications in treating PTSD, reducing stress, and enhancing user presence. For instance, to integrate reappraisal training into OVR-based PTSD treatment, olfactory counterconditioning can be used by presenting a trauma-associated odor within a controlled virtual environment. Throughout this process, individuals are guided to reinterpret the emotional significance of the odor, gradually diminishing its negative impact. With repeated reappraisal sessions, future encounters with the scent are more likely to elicit a neutral response, making related memories less distressing. While challenges persist—including limited scent libraries, bulky hardware, and delays in scent delivery—ongoing innovations are steadily advancing the capabilities of multisensory VR, establishing olfaction as a vital frontier in immersive experience design.1

## Extension of Wearable Olfactory Display of Multisensory VR Experience

The integration of olfactory stimuli into VR environments can significantly enhance user experience and immersion.6 Existing prototypes utilize pumps and acoustic wave devices to dispense specific odors consistently, but not accurately. These devices struggle in immersive environments with lingering odors and odor intensity, which can lead to unintended mixing and user confusion. Heat buildup is another critical concern, as inadequate dissipation can significantly shorten device lifespan, especially in wearable VR systems.  Solutions such as deodorant filters, forced air cooling, and an exhaust chamber that integrates the airflow system for odor removal with the cooling system help maintain the system temperature within operational parameters. Factors that require consideration for future improvement are understanding the perceived variation of olfactory thresholds. To improve overall perceptibility, odor intensity levels require further calibration. Selection of odor components is crucial for covering a wide range of target scents and maintaining accuracy in odor generation. The challenge of wearable VR devices is and the lack of trials to determine effectiveness.7

## Automatic Scent Creation by Cheminformatics Method

While visual and auditory stimuli can be readily digitized and integrated into cyber-physical systems, olfaction remains a significant challenge. A critical step toward digitizing the sense of smell is the algorithmic design of scents aligned with specific odor descriptors. Traditional scent creation requires extensive time, experience, and iterative blending to achieve a desired result. Although prior research has explored designing novel odorous molecules with targeted descriptors, these efforts often lacked experimental validation and failed to produce actual scents.8

Recent breakthroughs in digital olfaction technologies have made the automation of scent generation more feasible.4 Two core cheminformatics technologies enable this advancement. The first is *odor reproduction*, in which complex multidimensional sensing data—such as mass spectra—are decomposed into odor components.9,10 These components are then recombined to replicate specific scents. Previous studies focused on reconstructing known mass spectra; in contrast, the present work employs these components to generate entirely new scents.

The second technology involves predicting odor profiles,11-15 either in the form of qualitative descriptors or quantitative ratings. While molecular descriptors have traditionally been used for scent prediction, they are limited to individual compounds and rely on prior knowledge of mixture composition, rendering them ineffective for complex, multi-component substances such as essential oils.

To overcome this limitation, a method was developed to accurately predict odor profiles directly from mass spectral data16-18, even for complex compound mixtures. While earlier studies confirmed that computational models could infer odor profiles, they stopped short of producing physical scents. This study advances that work by generating actual scents that correspond to targeted odor descriptors.19

Key innovations include the focus on essential oils, which offer more complex and realistic olfactory data than single molecules. This study shifts from using continuous odor ratings19 to binary descriptors, which align more closely with human sensory language and are easier to standardize across diverse scent categories. Additionally, odor components are reintroduced—not merely to reproduce existing scents but as intermediaries in the generative design of new ones.

This study demonstrates, for the first time, the successful creation of real scents with intended descriptors by integrating odor profile prediction with odor reproduction.20 The results were validated through sensory evaluation, confirming the viability of this approach for digital scent generation.

This research advances scent digitalization by demonstrating the potential of deep neural networks for odor design. It contributes to the growing field of olfactory digitization, paving the way for future innovations in digital scent synthesis.8

## Predicting Human Odor Perception Represented by Continuous Values from Mass Spectra of Essential Oils Resembling Chemical Mixtures

This study examines the science of smell. Odor perception begins when airborne molecules, known as odorants, bind to and activate olfactory receptors (ORs) located on olfactory receptor neurons (ORNs).21 These neurons transmit signals to the olfactory bulb, where unique patterns of neural activity are formed and processed by the brain to interpret specific odors. The discovery of odorant receptors has led to major advances in the biological understanding of how olfaction functions at the molecular and neural levels.22 Fragment peaks in mass spectra provide valuable insights into molecular structure and serve as effective inputs for neural network models. Unlike molecular structure parameters, which apply only to individual compounds, mass spectra capture the composite nature of mixtures—enabling the prediction of odor perception for complex substances like essential oils.23

Essential oils are composed of complex mixtures of natural, volatile, and aromatic compounds derived from plants, and as such, can be characterized using a diverse array of odor descriptors. The objective of this study is to cluster semantically and perceptually similar odor descriptors within a shared olfactory space and subsequently predict human odor perception using machine learning techniques. Specifically, the model aims to infer the presence or absence of odor descriptor groups based on the mass spectra of chemical mixtures, such as essential oils.24

Previous research has demonstrated methods for predicting odor characteristics from the mass spectra of the NIST dataset and chemical compounds listed in the Sigma-Aldrich catalog.25 However, these prior studies relied on binary odor descriptor labels for individual odorants, which introduces limitations. Binary representations obscure meaningful relationships between similar descriptors—such as ‘floral’ and ‘rose’—by treating them as mutually exclusive, despite their significant semantic and perceptual overlap.24

To address this issue, this study converts binary odor descriptor data into continuous values, allowing for the construction of more nuanced odor descriptor groups. These groups are created through correlation coefficient–based clustering of continuous descriptor values, enhanced by natural language processing techniques that evaluate the semantic similarity between terms. This combined approach enhances the model's capacity to represent and predict complex odor perceptions from chemical mixtures.24

The relationship between mass spectra and human-perceived odor impressions is inherently complex and exhibits strong nonlinearity. To address this, this study employs an autoencoder—a type of neural network designed to extract lower-dimensional feature representations from high-dimensional, nonlinear data. Specifically, the autoencoder compresses the original mass spectral data into a latent bottleneck space, generating feature vectors that retain essential structural information. The extracted feature vectors are then fed into a neural network model designed to predict corresponding odor descriptor groups.24

Given the inherent class imbalance in the dataset—where positive and negative samples are unevenly distributed—the Synthetic Minority Oversampling Technique (SMOTE) is applied to improve model performance. SMOTE is used to balance the sample distribution for both correlation coefficient–based and semantic word similarity–based odor descriptor group predictions, thereby enhancing the reliability and robustness of the classification results.24

## From Meaning to Perception – Exploring the Space between Word and Odor Perception Embeddings

This work aims to generate vector-based representations—or embeddings—of perfume notes that reflect meaningful sensory and semantic relationships. Using perfume descriptions, where a list of notes describes each fragrance, the authors trained embeddings that successfully cluster related scents (e.g., lemon and mandarin among citrus notes).26

To evaluate these embeddings, the study compares distances between notes in the perfume embedding space with distances in pre-trained word embeddings. Results show a correlation between the two, particularly for words more strongly associated with smell. The researchers further trained a regressor to map word embeddings to odor embeddings, enabling the model to infer a scent-related representation for any word, thus linking language and olfaction.26

Perfume compositions are organized into three sequential layers: top, heart, and base notes. Top notes consist of light, highly volatile molecules that evaporate quickly and form the initial impression. Heart notes follow shortly after, lasting from several minutes to about an hour, and represent the core character of the fragrance. Base notes emerge last, composed of heavier, longer-lasting molecules that can linger for up to 24 hours, often acting as fixatives to anchor and stabilize the scent.26

The study also explores the connection between smell and language, noting that while English has limited vocabulary for odors, this constraint may not exist in all languages.27-29 Some research indicates that source-based descriptors predominate olfactory language across cultures.30-32 Despite these challenges, distributional semantics—like Word2Vec—can be used to quantify odor-word associations.

Using the CBOW model26 of Word2Vec,33-34 this analysis generated word embeddings and compared them with custom odor perception embeddings derived from perfume data. The team measured the alignment between these embeddings by ranking the similarity between notes (e.g., lemon and orange) and using the Rank Biased Overlap (RBO) metric to quantify the overlap between linguistic and olfactory rankings.26

To further bridge semantic and perceptual representations, the researchers trained regressors to map word embeddings onto odor embeddings. Three models were evaluated—Linear Regression, MLP Regressor, and K-Nearest Neighbors—using embeddings reduced from 300 to 20 dimensions via PCA. While the MLP Regressor yielded the highest performance, the Linear Regressor was selected for final analysis due to its simplicity and interpretability. However, all models only marginally outperformed a baseline dummy regressor, indicating that the mapping may be inherently complex or that the models require further tuning and optimization.26

## Olfactory Perception of Chemically Diverse Molecules

Research linking molecular features to olfactory perception reveals several key insights and complexities. Molecular properties such as molecular weight, vapor pressure, and the presence of sulfur atoms strongly correlate with perceived intensity and pleasantness. Heavier molecules and those with lower vapor pressure tend to be perceived as more intense, while the presence of sulfur often leads to unpleasant perceptions.35 Conversely, molecular complexity is associated with greater pleasantness, and models combining multiple molecular features predicted pleasantness moderately well (r ≈ 0.5).36,37 Two significant challenges complicate olfactory prediction:

1. Individual variation in olfactory receptor expression means people can perceive the same molecule differently.38-42
2. Cognitive and cultural factors, such as familiarity, prior experience, language, and motivation, influence how odors are described and rated.35

Familiar odors are more likely to receive semantic labels and be rated as distinctly pleasant or unpleasant. Unfamiliar odors, in contrast, are often left undescribed and rated as neutral.

A significant finding is the nonlinear nature of intensity perception: some molecules exhibit a reversal in perceived strength at different dilutions due to unique concentration–response curves. This highlights the importance of modeling complete concentration–response functions rather than relying solely on intensity measurements at a single dilution.

Additionally, the study compares population-level ratings with individual responses, revealing that perception is not only molecule-driven but also shaped by each subject’s interpretive framework. For instance, “musky” is associated with body odor by laypeople but with pleasant base notes by fragrance experts.

Ultimately, this study suggests that assigning semantic descriptors to odors is a learned process influenced by experience, which helps explain why some descriptors are more easily predicted from molecular data than others.35

## Summary and Gaps in Literature

The integration of olfaction into Virtual Reality (OVR) is quickly becoming a pivotal advancement in the pursuit of fully immersive and emotionally engaging digital experiences. Although traditional VR has focused primarily on visual and auditory elements, the powerful influence of scent on perception, cognition, and emotion has long been underestimated. Recent technological progress is now beginning to close this sensory gap.

Current OVR prototypes, often employing airflow and scent-dispensing mechanisms, demonstrate technical feasibility but face significant challenges, including lingering odors, inconsistent intensity, heat dissipation, and limited scent libraries. Despite these hurdles, the potential applications are vast and transformative, particularly in therapeutic contexts like PTSD treatment through olfactory counterconditioning and in enhancing education, training, and overall user presence.

Crucially, breakthroughs in digital olfaction technologies, underpinned by cheminformatics and artificial intelligence, are revolutionizing the generation of scents. The ability to algorithmically design scents, predict odor profiles from complex mass spectra, and even generate entirely new physical scents with intended descriptors marks a significant leap forward. This advancement marks a shift from simple odor reproduction to true digital scent synthesis—where scents are algorithmically generated and validated through sensory evaluation.

While the complexities of olfaction—including individual variations in perception, cognitive influences, and the nonlinear nature of intensity—remain areas of ongoing research, the convergence of neuroscience, chemistry, and AI is paving the way for increasingly sophisticated and personalized olfactory VR experiences. As these technologies mature, overcoming current limitations in device size, delivery speed, and scent variety, OVR is poised to become an indispensable component of virtual environments, unlocking new dimensions of realism, emotional engagement, and therapeutic efficacy.

Gaps in current OVR literature exist regarding the achievement of seamless and perceptually accurate scent delivery using VR wearable technology. Solutioning the issues with heat buildup, lingering odors, and the science to maintain immersion present challenges. Next, calibrating odor intensity to account for vast individual variations in olfactory thresholds demands something more than a universal approach. Limited scent libraries are a factor. Most existing scent libraries are designed around pleasant, floral fragrances, limiting their relevance for therapeutic applications. Targeted trials are essential to calibrate OVR wearable devices and create richer, more effective user experiences for patients in therapeutic settings.

# Methodology

This project will develop a scent selection model and a scent-emitting device for virtual reality (VR) therapy. The approach combines data-driven scent prediction with engineering design to enable olfactory feedback in VR. Work will focus on two core components: Scent Model Development and Scent Emitter Device Design. The process will also address risk, lifecycle management, budget constraints, and stakeholder needs.

## Research Background

This research uses single-class classification modeling, semantic clustering, and systems engineering to improve olfactory feedback in VR environments. Existing models fail to capture the complexity of scent data due to limited exploration in scent profile labeling. Using and adapting a dataset from a publication in 2024 in Scientific Reports,8 our team will build upon their findings to create and develop a new method of scent profiling and handling scent nuances. Cluster-based classification better addresses this nuance6 by using semantic clustering to help group scents into meaningful profiles. These scent profiles can be customized to enhance a wide range of therapeutic VR applications by delivering targeted and emotionally meaningful olfactory stimuli.2

## Scent Model Designs

The Scent Selection Model will use mass spectra and odor ID data to predict the scent profiles of essential oils. The raw mass spectra will be standardized and normalized across various intensities to ensure their comparability across all essential oils. Scent clusters will be made using Word2Vec26 vector similarity with K-means clustering and LLM-supported reasoning to group and generate meaningful arrays of scent profiles based on descriptor linguistic semantics. The LLM, OpenAI’s ChatGPT (GPT-4.0), will be used to enhance the interpretability and relevance of cluster creation from the initial vectorization47. The model will draw on cross-domain olfactory research, fragrance taxonomy, and scent literature to categorize the descriptors. This LLM-assisted refinement will ensure an even and accurate representation of descriptors, as well as the applicability of scent clusters to therapeutic applications. The resulting clusters will then be manually reviewed and validated against literature by the team to ensure proper linguistic scent linkages and applicability to trigger scents in an organized clustering system. The data will first undergo preprocessing to identify and resolve inconsistencies or gaps. Essential oils will then be assigned to meaningful clusters based on dominant descriptor selection. To incorporate key trigger scents—such as smoke, burning rubber, gunpowder, and diesel—the initial descriptor matrix will be expanded and refined, ensuring a more accurate representation of semantic overlaps among oils. A large language model will be employed to evaluate and validate descriptor groupings, maintaining consistency and traceability across oil profiles. Additionally, the model will be used to infer relationships between oils with similar naming conventions—such as scientific versus common names—and to determine extraction methods (e.g., steam distillation vs. cold pressing), which influence the resulting scent profiles. To support systematic assignment and justification, OpenAI’s ChatGPT (GPT-4.0) will serve as a research assistant. The model will be prompted iteratively with oil-specific queries and contextual information to generate descriptor justifications, ensuring transparent and well-reasoned classification across the dataset47. The final outputs will be reviewed and validated by the team to ensure consistency and interpretability. This process will ground the cluster design in both linguistic reasoning and olfactory literature. Once the clusters are established, SMOTE24 will be applied to oversample and synthesize the dataset, addressing class imbalance and improving training accuracy. Two neural network models will be tested to validate if fully connected dense neural networks provide stronger model development compared to convolutional layering. Both models will also interpret the data using PCA to test whether reducing the dimensionality of the data will aid in stronger model generalization or lead to a loss of meaningful spectral detail. All models will be trained using a stratified 80/20 train-test split to maintain balanced class distributions across folds. Performance metrics will be calculated exclusively on the test set to prevent data leakage and ensure valid model evaluation. The final model will be selected based on key performance indicators, including accuracy, precision, and F1 score, to ensure a balanced evaluation of predictive power and reliability. Additionally, diagnostics such as confidence score distributions, heatmaps, and confusion matrices will enable more accurate assessments of the model's reliability and concerns regarding overconfidence. Once the model is finalized, each essential oil will be tested, mapped to a corresponding cluster, and assigned a confidence score. This ensures the model’s outcomes are both practical and actionable, supporting informed oil selection for target therapeutic scent profiles. A final CSV report will show these mappings. It will be used to guide oil selections for therapeutic VR scent blending, particularly targeting clinically relevant clusters such as “Smoke” and “Gunpowder” for PTSD exposure therapy.2 Post evaluation will also include visuals for prediction confidence distributions and analyzing misclassified oils to identify inconsistencies and flag weakly represented clusters for future dataset expansion.

## Semantic Linguistic Classification

This classification maps scent clusters to trigger scents, enabling the model to align scent profiles with specific VR visual stimuli. It links clustered scents to targeted responses. By anchoring scent clusters in both semantic proximity and therapeutic applicability, the model ensures that each predicted cluster reflects a generalized olfactory experience rather than a single rigid label. This flexibility enhances the system's ability to generalize odor profiles across all oils and supports consistent emotional and behavioral effects when used in VR-based therapeutic applications. Eller and Vosshall found that pairing scent with visual cues can shape emotional perception, allowing one scent to represent a range of related scents within a profile.45 Ultimately, by creating meaningful scent profiles, the model can help achieve targeted psychological and clinical outcomes by providing a predefined trigger spectrum that aligns with and functions in various VR visual environments.

## Scent Emitter Designs

Building on the model's outputs, the team will design a scent-emitting device capable of delivering targeted trigger scents in VR environments. This device, referred to as the VORA Scent Emitter, will be engineered to mount directly onto pre-existing VR headsets, providing seamless sensory integration without compromising user comfort or requiring new investments in VR headsets. The design process will begin with gathering and forming a functional requirements list for the emitter, taking into account the needs of users and clinicians. These will include criteria such as portability, device accuracy, device modularity, safety, and subsystem management. Requirements are derived from system constraints, literature, and stakeholder needs and organized using SysML architecture to show requirement definitions and requirement instances. These will be formally mapped to logical and behavioral components and visualized to show system flow. Once the requirements are mapped and well-defined, the hardware design will build upon previous research to create a device that is more accurate and customized to handle the wide variety of scent profiles targeted in the model. The design will improve upon the previous SAW atomization approach, address concerns of safety and usability, and seamless integration for hygienic clinical applications. Another key component will be the central mixing manifold, which supports the blending flexibility of various oil components. Future iterations of the system will include an embedded control system that functions with an open API, allowing integration with various VR platforms, such as Unity or Unreal Engine. This software layer will support the need for real-time synchronization between VR stimuli and the user's olfactory output.

The final system will be validated against the initial requirements using SysML satisfaction links to show traceability from user needs to system performance. The complete architecture will be comprised of four layers in the SysML model: logical and behavioral mappings and requirement definitions and constrained attributes. A comprehensive Bill of Materials (BOM) will be created once the system design is fully defined, supporting efficient part procurement and ensuring traceability of all components throughout development and deployment. The BOM will list all mechanical, electronic, and consumable parts required for the VORA assembly. Using this structured approach will support a reliable, scalable, and clinically applicable scent delivery system that will assist in the outcome of therapeutic applications with PTSD exposure therapy, cognitive rehabilitation, and anxiety reduction.

## Project Scope Considerations

This project will encompass both technical development and strategic planning across a variety of domains, like data modeling, clinical application integration, hardware system design, and cross-functional collaboration. The scope of this project will include the development of a scent classification model, the design of a VR-compatible scent emitting device, and the formation of supporting system documents. Key project management elements will be integrated throughout the project to ensure accountability, maintain feasibility, and align objectives with stakeholder expectations. These elements will include:

* A hierarchical Work Breakdown Structure that will organize the project into manageable tasks and show their dependencies.
* A mapped stakeholder analysis that will address key stakeholders' needs with a power-interest matrix.
* A formal risk analysis that addresses the likelihood and impact scores of potential risks. To quantify uncertainty, a Monte Carlo simulation will be performed to assess the overall severity of risk exposure and variance across a range of potential scenarios.

## Alignment with Project Goals

The methodology follows a clear set of steps, including the use of cluster-based classification for scent profiling, incorporation of device design, and application of modeling languages. Machine learning manages technical tasks, while systems engineering supports project management. Together, these elements create a comprehensive solution to enhance VR therapy with olfactory feedback.

# Analysis

The analysis of the capstone project, VORA (Virtual Olfactory Reconfigurable Architecture), involved the development of two integral components. The first component was a predictive machine-learned classification model that utilized mass spectra of essential oils, and the second was a modular SysML-modeled scent-emitting device designed for applications in clinical VR therapy. These efforts both support the goal of improving and aiding PTSD treatment with the controlled delivery of olfactory stimuli alongside visual stimuli. The analysis outlines the outcomes of the methodologies applied, the overall insights gained, and how the findings validate and push the boundaries of the project's multidisciplinary goals.

Along with the technical modeling analysis, this project utilized structured project management tools to ensure that the VORA system could be feasibly created, thoroughly validated, and deployed with traceability to the project requirements. Viewing the project’s stakeholder alignment and ethical safeguards is a key principle to ensure that the project's foundation is well thought out and documented. Using risk registers with quantified simulations, a timeline of events in a work breakdown structure, and a stakeholder engagement plan reinforces the capstone’s alignment with systems engineering and leadership best practices, providing long-term project sustainability.

## Data Preparation and Pre-Processing

### Essential Oil and Mass Spectra Analysis

The project aimed to improve a predictive model using two datasets: a high-resolution mass spectra file, which provided the m/z intensities for essential oils, and a binary descriptor matrix that classified odor descriptors for each oil.8 hrough semantic-linguistic clustering, we grouped these descriptors into twelve clinically and chemically meaningful scent profiles below in Table 1. Through techniques such as Word2Vec, we incorporated a natural language processing model, as outlined in Appendix A, to help form the initial cluster profiles based on context-driven learning of the descriptor labels. The original data set contained39 unique scent descriptors with some having similar overlap, such as “Light” and “Pleasant”. Initial modeling and data verification revealed that five oils had similar names but were distinguishable based on differences in their mass spectra. Upon further analysis, the observed differences were attributed to the extraction methods—steam distillation versus cold pressing—which significantly impacted the mass spectral profiles and their correlation with odor descriptors, as each method emphasizes distinct aromatic notes.44 Only one oil, Origanum Compactum, was removed due to little differences in spectral data and lack of meaningful differences for the model to train on. Initial models revealed that the data lacked sufficient descriptor consistency and lacked adequate descriptor labels for us to apply them effectively to a model for the target trigger scent profiles. Through further research, we added 14 more descriptor labels to enhance the model's learning capabilities and provide a deeper range of scent profile descriptors.45 With the addition of these labels, all descriptors were systematically reviewed to ensure alignment with established research on semantic label classification. A large language model was employed to refine and validate the descriptor set, grounding the revised framework in relevant literature. Although the initial descriptor vectorization relied on expert judgment rather than a comprehensive empirical study, the resulting variations in oil classification were determined to have minimal impact on overall user experience. Justifications for these decisions are documented and evaluated in Appendix D. The first four clusters were designed to capture key trigger scent profiles, while the remaining clusters were refined using a combination of Word2Vec vector analysis and domain expertise informed by current research.

|  |  |  |
| --- | --- | --- |
| Cluster | Cluster Name | Descriptors |
| 1 | Smoke | Smoky, Phenolic, Burnt, Medicinal |
| 2 | Burning Rubber | Rubber, Sulfurous, Acrid, Metallic |
| 3 | Gunpowder | Sharp, Dry, Powdery, Leathery |
| 4 | Diesel | Musky, Animalic, Pungent, Oily |
| 5 | Herbal Green | Green, Grassy, Herbaceous, Leafy, Grass |
| 6 | Citrus Zest | Citrus, Lemon, Orange, Citronellal, Pleasant |
| 7 | Floral Soft | Floral, Rose, Tea, Light |
| 8 | Spiced Sweet | Spicy, Sweet, Fatty, Anisic, Balsamic |
| 9 | Wood Resin | Woody, Root, Earthy, Resinous, Pine |
| 10 | Cooling Fresh | Fresh, Minty, Cooling, Eucalyptol |
| 11 | Aromatic Herbal | Terpene, Aromatic, Peppery, Anethole |
| 12 | Warm Rich | Warm, Rich, Bitter, Fruity |

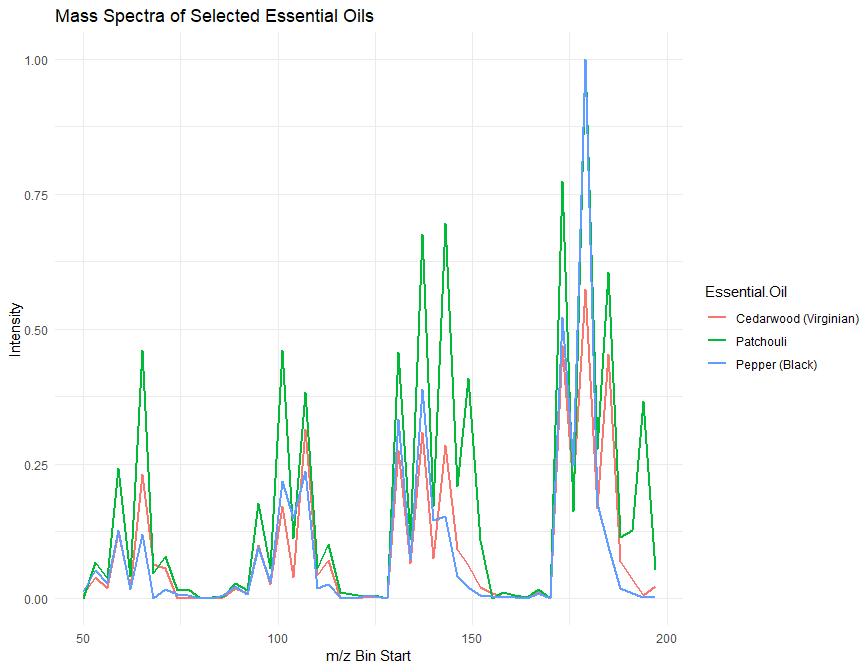
Table 1: Twelve Scent Profiles

Based on this framework, each oil was assigned a dominant cluster label to reduce model complexity during training and to enhance interpretability in practical therapeutic contexts. Initial insights into the individual data sets revealed sufficient uniqueness in the intensities of each oil to enable the model to learn unsupervised. Examining the faceted view of all the oils in Figure 1 below, the intensities and their corresponding peak frequency intervals show similar overlap in certain oils. Taking a smaller sample of three specific oils with common descriptors that fit into the smoke scent profile reveals a common occurrence of intensities and even similar levels of mass spectral intensities, suggesting that these oils closely fit within that scent profile. Figure 2 below shows cedarwood, patchouli, and black pepper peaking at similar sampling frequencies and having overlap in intensity values at those sampling frequencies.

Figure 1: Faceted Line Plots of Mass Spectra Intensities of All Essential Oils



Figure 2: Line Plot of Mass Spectra Intensities of Correlated Essential Oils



Looking deeper into these specific oils highlights the purpose and value that a scent note has in a person's perception. Scents are composed of a blend of notes that unfold over time, typically categorized into three layers: top notes, heart notes, and base notes. These elements unfold gradually over time. Top notes consist of light, small, and highly volatile molecules that evaporate quickly. Heart notes emerge as the top notes fade, typically within two minutes to an hour after application. Base notes appear once the heart notes dissipate, comprising heavier, larger, and less volatile molecules. Often serving as fixatives, base notes become noticeable after about 30 minutes and can linger for up to 24 hours.26 This study examined specific notes that determine "smoke", "burnt rubber", "gunpowder", and "diesel" scents. The results of our findings, showing the most similar notes for each scent, are presented in Appendix B. Figure 3 illustrates the Fragrance Pyramid, indicating the typical duration a scent remains with a person. Figure 4 shows the specific notes associated with the "Smoke" scent.

Figure 3: Fragrance Pyramid46

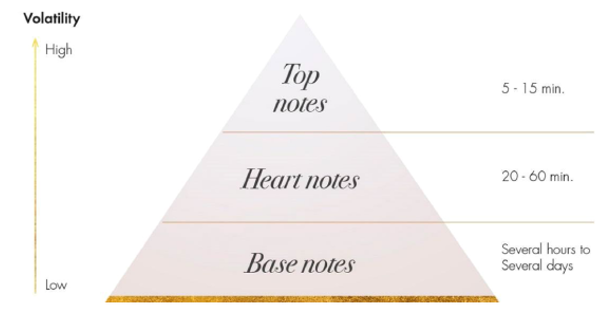
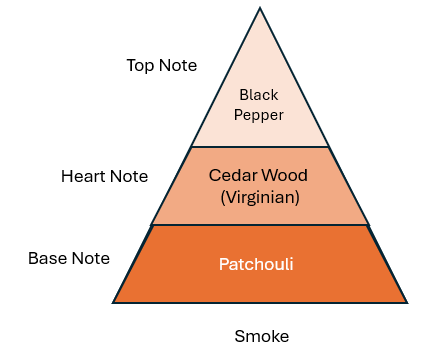


Figure 4: Smoke Scent Notes



### VORA Device Requirement Analysis

The hardware system preprocessing began with gathering stakeholder requirements derived from both clinical objectives and end-user needs. These were translated into 69 formal requirements, detailed in Appendix C, and organized into system-level and subsystem-level categories. The requirements were modeled using a SysML kernel, incorporating Boolean and integer attributes to ensure alignment with the constraints and functionality of real-world embedded devices. This process resulted in the creation of a Bill of Materials, detailed in Appendix C, which outlined the total unit costs and all the components required for assembly.

## Applications of Modeling Methodologies

### Predictive Modeling with CNN and MLP

To evaluate the most effective approach for modeling mass spectral data about odor clusters, two neural network architectures were tested: convolutional neural networks (CNNs) and multilayer perceptrons (MLPs). CNNs were selected for their ability to capture local spectral patterns, while MLPs demonstrated effectiveness in processing dense vector representations. Both models were developed using the TensorFlow framework, incorporating early stopping and reproducibility constraints to ensure training stability and consistency. All four models, as well as the master script, can be found in Appendix A. To address class imbalance among odor clusters, the Synthetic Minority Oversampling Technique (SMOTE) was employed. SMOTE mitigates model bias by generating synthetic feature vectors through interpolation with the k-nearest neighbors of minority samples, effectively augmenting underrepresented clusters. For a minority sample xᵢ, a synthetic sample xnew is generated by:24

xnew = xᵢ + δ · (xₙₙ − xᵢ)

where xₙₙ is a randomly selected nearest neighbor, and δ ∈ [0, 1] is a random number. All spectral data were standardized using Z-score normalization to prevent model distortion caused by significant disparities in value ranges. For a normalization of the sample, μ is the mean, σ is the standard deviation, and x ' is generated by: 24

x' = (x − μ) / σ

The PCA analysis was crucial to test whether the data could be better understood and trained without reducing dimensionality. Initial data preprocessing revealed that approximately 64% of the variance could be accounted for by the first four principal components of the Mass Spectra. Figure 5 below shows the proportion of variance along each component, tapering off heavily after the fourth component.

Figure 5: Proportion of Variance along Principal Components of Mass Spectra

A graph of a function

Description automatically generated

This result is a significant finding because it suggests that there is an underlying structure in the data surrounding the relationships between spectral features alone, indicating to the team that there is room to reduce high-dimensional data without losing essential information. The relatively steep drop-off in variance also suggests that many spectral features are correlated or redundant, supporting the notion that the essential oils may have recurring chemical scent patterns or dominant peaks across multiple oils. Moving forward with the development of the four models with a justified reasoning for dimensionality reduction, each model provides a series of metrics, including accuracy, F-1 score, confidence distributions, cluster-specific heatmaps, and confusion matrices, to ensure the results are transparent and traceable to the model’s performance.

### SysML Architecture and Design Modeling

Initial design prototypes were based on a preliminary academic research model that pursued similar objectives.7 Recognizing the limitations and constraints of that approach, we enhanced the design by introducing quantifiable criteria for component selection and adaptability to enhance the functionality and value of virtual reality therapeutic applications.

The VORA scent emitter transitioned from earlier surface acoustic wave (SAW) atomization to a micropump-driven architecture. This shift was crucial to the device's overall functionality in handling multiple oils and blending to meet the demands of a wide range of scent profiles. The SAW device had poor precision of scent volume, limited scalability, and issues with overheating. This revised system will utilize Bartels MP6 micropumps with pulse width modulation, like the previous device, to provide improved control of oil delivery and enhanced thermal safety. Our model will double the oil pod capacity and maintain ergonomics and space, thanks to the design of VORA's housing around the front of the user's face. The team incorporates several safety and reliability features, including RFID chips on oil pods to track quality and usage, a UV-C LED as an internal sterilizer between uses, and disposable silicone nozzle tips to ensure accurate and clean vapor delivery for each user. A Bill of Materials, found in Appendix C, was compiled based on this updated structure, which provides transparency and ease of tracking when purchasing components. From this foundation, an architecture was developed to model and illustrate the interrelated components of the VORA device.

The architecture of the SysML Design was created using:

* Requirement Definitions: Using specific “require constraint” clauses tied to unique attributes like “power draw” or “battery life minutes”.
* Logical Architecture: Structured using “part”, “port”, and “connect” statements to provide overall system and subsystem flow with power mapping between subsystems.
* Behavioral Architecture: Mirrored the logical system and subsystem hierarchy to display future device sequencing and processing of information.
* Validation Layer: Used “satisfy” statements to ensure traceability from the logical side to the requirement definitions.

All modeling pieces were reviewed and validated to ensure they met all system requirements, thereby providing diagrams of traceable blueprints which are contained in Appendix C.

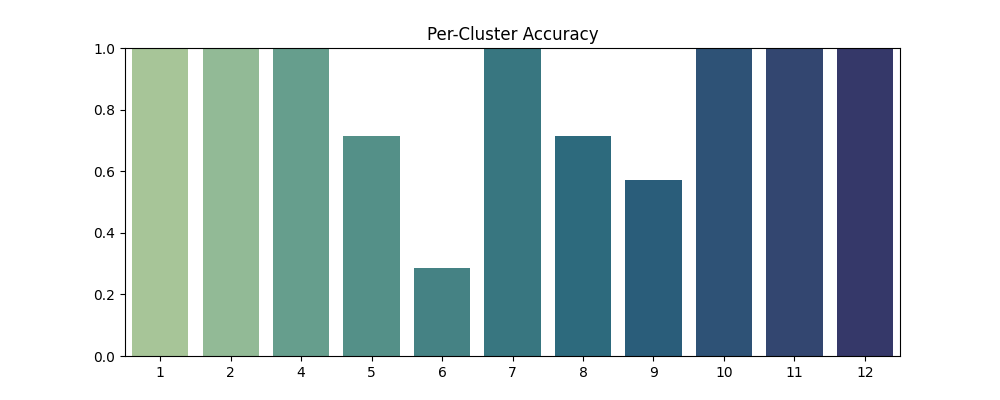
### Key Findings and Result Validation

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Accuracy | F1 Score | Mean Confidence |
| CNN with PCA | 0.8421 | 0.8421 | 0.9097 |
| CNN without PCA | 0.7957 | 0.7957 | 0.9267 |
| MLP with PCA | 0.7632 | 0.7632 | 0.8218 |
| MLP without PCA | 0.7527 | 0.7527 | 0.9099 |

Table 2: Modeling Results

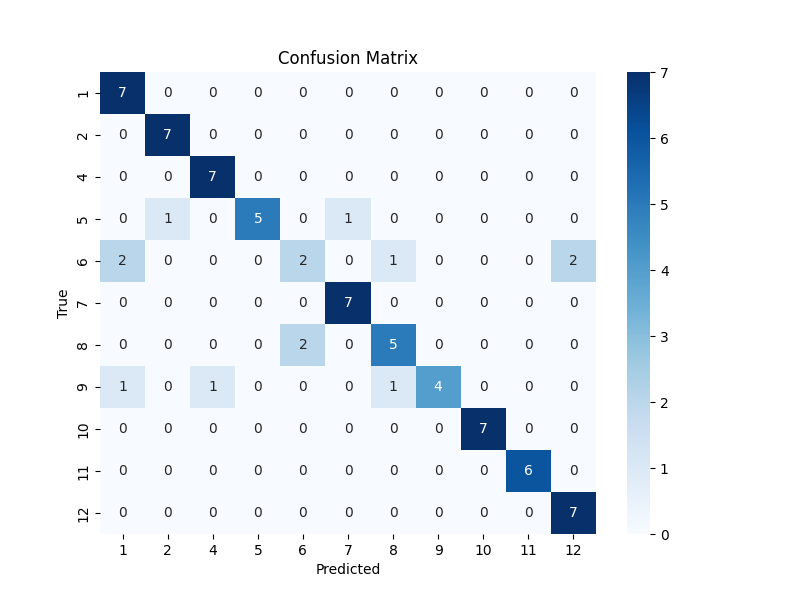
As illustrated in Table 2, the CNN models outperformed the MLP’s significantly, especially with the use of PCA. These results suggest that the initial theory of capturing variance through dimensional reduction helped the model learn local features in the mass spectra more effectively. Total mean confidence stayed high across the board, and overconfidence in misclassifications remained low. Although all models struggled to train on one of the clusters due to a lack of oil assignments, eleven of the twelve clusters were successful. The cluster not included was the third target cluster, Gunpowder, which was particularly difficult to assign oils to due to a lack of dominant descriptor values. Because of how SMOTE works, there needs to be at least one oil assignment per cluster for it to synthesize meaningful data. This finding suggests there is room for additional oils, particularly catered to typical gunpowder descriptors, to be added to the dataset to help the model grow. For the remaining clusters, Figure 6 shows the accuracy of each cluster, presenting strong results from the model.

Figure 6: Accuracy Per-Cluster Results from CNN with PCA



Looking deeper into some of the reasons for the accuracy errors, the team evaluates a confusion matrix of the prediction results to look for any overlaps in misclassification. Figure 7 below shows a strong diagonal trend, indicating a high level of agreement between actual and predicted cluster assignments. However, prominent misclassifications reveal some systematic patterns that could stem from either semantic or chemical overlaps between certain oils and various clusters.

Figure 7: Confusion Matrix of Actual versus Predicted Results from CNN with PCA



For example, Cluster 6 (“Citrus Zest”) frequently misclassified samples from Cluster 1 (“Smoke”) and Cluster 12 (“Warm Rich”), which suggests potential overlap in their underlying volatile compounds or ambiguity in descriptor labels such as “Pleasant” and “Warm” or “Medicinal” and “Citronella.” These descriptors are semantically broad and may contribute to perceptual similarity, thus influencing the model's confusion. Clusters that exhibit near-zero recall or precision likely reflect the model’s inability to learn distinct, generalizable representations of those scent clusters, which could be due to issues of low sample sizes or overlapping descriptor semantics. While Figure 7 demonstrates high overall classification accuracy across most clusters, it is essential to acknowledge that raw accuracy alone can mask underlying prediction biases. For instance, a model that overpredicts a dominant cluster may achieve high accuracy while misclassifying less-represented classes. Therefore, the use of the F1 score becomes crucial in evaluating model performance, as it balances both precision and recall, offering a more reliable assessment of classification efficacy in imbalanced and semantically nuanced datasets like this one. Using F-1 metrics to justify the precision and accuracy trade-off, the CNN with PCA model was selected as the final choice due to its demonstrated generalizability, yielding results that were both meaningful and interpretable for oil selection in clinical trials.

This model not only provides a method for future classification of other essential oils into scent profiles but also lays a foundation for further innovation in scent profile selection and descriptor label selection. Using the high confidence score from this model, we can then move to select the oils we aim to use to target these specific scent clusters. Using research on these specific oils, we discover different top note, middle note, and base note profiles that emit a range of scents as the user absorbs them.26 Future research is needed to fine-tune these essential oil notes, thereby providing high confidence in determining which specific blends and concentrations should be used in the pods of the VORA device.

### Engineering Design Results

The design component successfully has:

* A complete mapping of 69 stakeholder-backed requirements
* A logical and behavioral architectures fully synchronized in SysML
* A verified validation mapping to all blocks of architectures
* A traceability matrix embedded across all packages
* A thorough Bill of Materials linked to entire device assembly

Based on VORA’s improvement over preliminary designs, the device exceeds standard research availability, offering options to blend multiple different oils through a vortex system to aid in scent immersion between different scent profiles. The design provides a robust foundation for future iterations of broader clinical use, offering opportunities for growth in other industries. The next step of this component involves creating iterations of computer-aided designs (CADs) that display the interfaces and interworking details of the device to build and create the system and subsystems. Using the information from the BOM and the overall design of the system and function, a future visualization was created using OpenAI’s ChatGPT (GPT-4.0) to help present a proof-of-concept design47. This design, shown in Figure 8, presents the functional view of the VORA device, including its housing's appearance, its location, and its placement on the user. The design shows inlet ports and exhaust ports for routing fresh air in and removing residual air out. There is room for a band or strap to connect from the base of the device around the user's neck, ensuring stability during operation. The front panel display functions to show color-mapped frequencies that correlate with the RFID reading from which pods are active. The interface will also indicate whether the device is on or off, as well as when it is in a cleaning cycle. While this is the ultimate goal of device creation and design, prioritizing the meeting of stakeholder requirements and ensuring adherence to clinical hygiene practices will take precedence when considering the housing shape and functionality.

Figure 8: VORA Device Concept Design



## Support of Capstone Objectives

### Multidisciplinary Leadership and Ethical Design Considerations

This project required several cross-functional integrations of:

* Systems Engineering: Building SysML architectures linked to requirements
* Chemical Engineering: Interpreting spectral data and building maps to descriptor profiles
* Software Engineering: Constructing and validating neural machine learned models
* Healthcare UX: Forming requirements to align with clinical and user protocols

Creating a diverse system architecture that meets user-centered design and ensures software predictability addresses the challenges of this capstone by incorporating well-defined device logic and a chemical model that targets specific scent profiles. Merging these roles across the team highlights the value of engineering leadership and its application in integrating technology with VR therapy.

Deciding on a predictive model for scent therapy must have ethical considerations in all facets. The model should have:

* A prediction confidence that is included as an export to prevent overreliance
* Cluster mappings that reflect real clinical applications
* A modular design for ease of accessibility and flexibility for open sour extensions
* Safety features to limit potential data breaches or user harm

Ensuring that the VORA device is both responsibly used and flexible for the intended user's clinical usage helps promote patient safety and data transparency without compromising new, innovative scent selections and delivery methods..

### Project Management Analysis

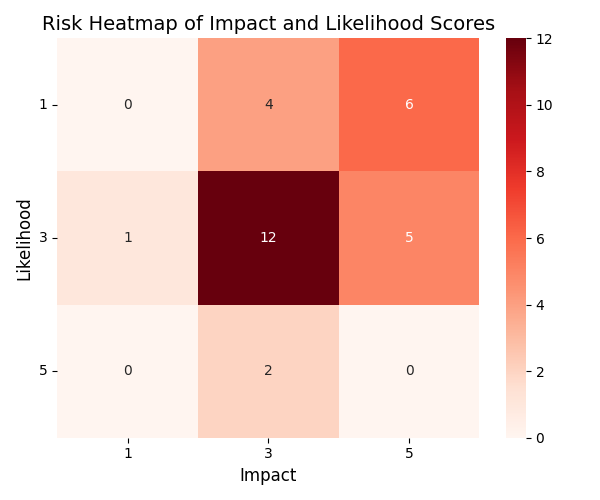
The capstone project was structured to deliver both technical and analytical modeling, guided by structured project management strategies to ensure stakeholder alignment and ethically informed decision-making. A detailed risk analysis, presented in Table 3, identifies potential risks, assigns likelihood and impact scores, and outlines corresponding mitigation strategies.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ID** | **Risk Description** | **Category** | **Likelihood** | **Impact** | **Severity** | **Mitigation Strategy** | **Likelihood Score** | **Impact Score** | **Severity Score** | **Residual Likelihood** | **Residual Impact** | **Residual Severity** |
| R1 | Scent overlap causes unintended blends during operation | Hardware Design | Medium | High | High | Implement precise micro-dosing; purge/flush cycle between scents | 3 | 5 | 15 | 2 | 4 | 8 |
| R2 | Micropump or valve failure interrupts scent delivery | Mechanical | Medium | Medium | Medium | Redundant array design; diagnostic sensors for real-time failure detection | 3 | 3 | 9 | 2 | 2 | 4 |
| R3 | Odor triggers unintended psychological response in PTSD patients | Clinical Safety | Low | High | High | Clinical validation of scent clusters; allow emergency stop; allow scent opt-out | 1 | 5 | 5 | 1 | 4 | 4 |
| R4 | Thermal module exceeds safe temperature threshold | Hardware Safety | Low | Medium | Medium | Limit via power constraint (<25W); thermal fuse; sensor-controlled shutoff | 1 | 3 | 3 | 1 | 2 | 2 |
| R5 | Prediction model is overconfident on incorrect cluster assignments | Software Model | High | Medium | High | Include prediction confidence output; retrain on updated feedback; soft alerts | 5 | 3 | 15 | 4 | 2 | 8 |
| R6 | Device is not compatible with various VR platforms | Integration | Medium | High | High | Use modular VR API interface; test across hardware platforms | 3 | 5 | 15 | 2 | 4 | 8 |
| R7 | Cleaning system fails to remove residual oils leading to scent contamination | Maintenance | Medium | Medium | Medium | Auto-cleaning cycle after use; pressure-based flushing detection | 3 | 3 | 9 | 2 | 2 | 4 |
| R8 | Cartridges inserted in wrong orientation or slot | User Error | Medium | Medium | Medium | RFID validation before activation; keyed cartridge shapes | 3 | 3 | 9 | 2 | 2 | 4 |
| R9 | Descriptor clustering fails to represent user-perceived scent groupings | Data Modeling | Medium | High | High | Incorporate user validation of clusters; allow mapping updates | 3 | 5 | 15 | 2 | 4 | 8 |
| R10 | Low battery life limits use in long clinical sessions | Usability | Low | Medium | Medium | Optimize energy use; modular battery or wired option | 1 | 3 | 3 | 1 | 2 | 2 |
| R11 | SMOTE oversampling introduces noise in minority class clusters | Modeling Bias | Medium | Medium | Medium | Combine SMOTE with stratified validation and confidence filtering | 3 | 3 | 9 | 2 | 2 | 4 |
| R12 | Requirement traceability breaks due to unlinked logical or behavioral elements | Systems Modeling | Low | High | Medium | Regular model validation with traceability matrix | 1 | 5 | 5 | 1 | 4 | 4 |
| R13 | Float values introduced in SysML cause validation errors | Modeling Syntax | Low | Medium | Medium | Enforce integer-only kernel; convert decimals to scaled integers | 1 | 3 | 3 | 1 | 2 | 2 |
| R14 | Confusion matrix shows persistent misclassification in similar scent clusters | Model Accuracy | High | Medium | High | Re-tune cluster mapping; introduce hierarchical model refinement | 5 | 3 | 15 | 4 | 2 | 8 |
| R15 | Failure to meet VR frame rate performance causes latency in scent delivery | Integration | Low | High | Medium | Async scent dispatch logic; prioritize low-latency microcontroller loop | 1 | 5 | 5 | 1 | 4 | 4 |
| R16 | Patient discomfort due to device weight or heat buildup | Ergonomics | Medium | Medium | Medium | Reduce component weight; improve airflow; balanced head strap design | 3 | 3 | 9 | 2 | 2 | 4 |
| R17 | Nozzle clogs from dried oil residue | Mechanical | Medium | Medium | Medium | Self-cleaning tip; filter screens; maintenance alerts | 3 | 3 | 9 | 2 | 2 | 4 |
| R18 | Intellectual property issues when open-sourcing model or emitter design | Legal/Ethical | Low | Medium | Medium | Document sources; license under CC/MIT | 1 | 3 | 3 | 1 | 2 | 2 |
| R19 | Team transition leads to knowledge loss in future development | Project Handoff | Medium | Medium | Medium | Detailed documentation, diagrams, and SysML exports | 3 | 3 | 9 | 2 | 2 | 4 |
| R20 | Scent availability becomes limited due to supply chain disruptions | Operational | Medium | Low | Medium | Allow cartridge substitution; multi-vendor sourcing strategy | 3 | 1 | 3 | 2 | 1 | 2 |
| R21 | Surface temperature exceeds safe skin contact threshold (42°C) | Safety | Low | High | Medium | Thermal insulation; real-time temperature sensors | 1 | 5 | 5 | 1 | 4 | 4 |
| R22 | Airflow bypasses filtration before reaching the user | Hygiene | Medium | High | High | Dual HEPA filters; flow-blocking failsafe | 3 | 5 | 15 | 2 | 4 | 8 |
| R23 | Physical contact with face spreads contaminants or causes discomfort | Hygiene | Medium | Medium | Medium | 1–2 cm standoff design; suspended mount | 3 | 3 | 9 | 2 | 2 | 4 |
| R24 | UV light exposure from sanitation module leaks or reflects toward user | Clinical Safety | Low | High | Medium | UV shielding; exposure lockout; IEC 62471 compliance | 1 | 5 | 5 | 1 | 4 | 4 |
| R25 | Device fails to shut off automatically when dislodged from face | Safety | Low | High | Medium | Accelerometer-triggered scent disable | 1 | 5 | 5 | 1 | 4 | 4 |
| R26 | Hygiene routine skipped, causing scent residue accumulation | Maintenance | Medium | Medium | Medium | Auto-cleaning; confirm clean state before use | 3 | 3 | 9 | 2 | 2 | 4 |
| R27 | Filter lifespan not tracked, leading to unnoticed performance degradation | Reproducibility | Medium | Medium | Medium | Filter load sensors; replacement alerts | 3 | 3 | 9 | 2 | 2 | 4 |
| R28 | Variation in scent delivery timing across sessions | Reproducibility | Medium | Medium | Medium | Warm-up readiness check; calibrated fan/pump timing | 3 | 3 | 9 | 2 | 2 | 4 |
| R29 | Failed integration with game engine control API disrupts scent synchronization | Usability | Medium | High | High | Pre-release Unity/Unreal testing; diagnostics | 3 | 5 | 15 | 2 | 4 | 8 |
| R30 | VR frame latency creates desynchronization between scene and scent delivery | Integration | Medium | Medium | Medium | Asynchronous scent queuing and delay tuning | 3 | 3 | 9 | 2 | 2 | 4 |

Table 3: Risk Register Analysis

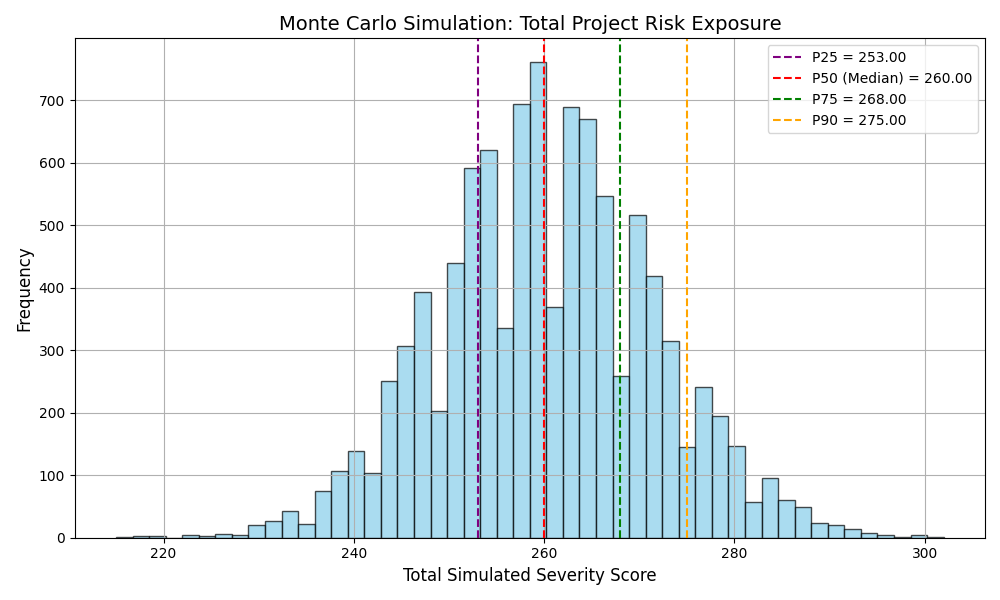
A heatmap, shown in Figure 9, helps visualize these risks across the specified dimensions, revealing a concentration of high-severity risk in the upper-right corner. The risk heatmap highlights the complexity of integrating predictive scent models and scent blending hardware with VR platforms

Figure 9: Heatmap of Impact and Likelihood of Project Risks



Noting risks such as scent blending errors (R1), platform incompatibilities (R6), and model prediction errors (R5), which show a high likelihood and high impact, suggest that these risks demand robust mitigation strategies and iterative validation. On the other hand, a large portion of the remaining risks falls within the moderate zone, where systems-engineered designs, such as "self-cleaning nozzles" and "RFID-locked cartridges," and documentation-based strategies help reduce the residual severity of these moderate risks. There are a few low-impact risks, such as low battery life (R10) and chain supply disruptions (R20), which have been sufficiently mitigated through specific design choices and noted mitigation strategies. Every noted risk has an associated strategy to mitigate its impact, which means that each of the 30 labeled risks can be successfully addressed. The distribution of impact and likelihood validates the layered approach to combining human-factor risks and technical risks, ensuring that the system remains resilient to change without overwhelming the project's development. To account for uncertainty, Monte Carlo simulations were employed, enabling the team to estimate total risk exposure with a potential variance using a triangular distribution method centered around its assessed risk values. 10,000 randomly generated scenarios reflect the small variations to assess uncertainty of impact and likelihood. The total severity for each scenario was calculated by combining likelihood and impact scores, resulting in an overall severity score to facilitate comprehensive risk assessment. The results in Figure 10 show that the majority of simulated risk scenarios have a total severity score between 253 and 268, indicating a tight distribution of outcomes.

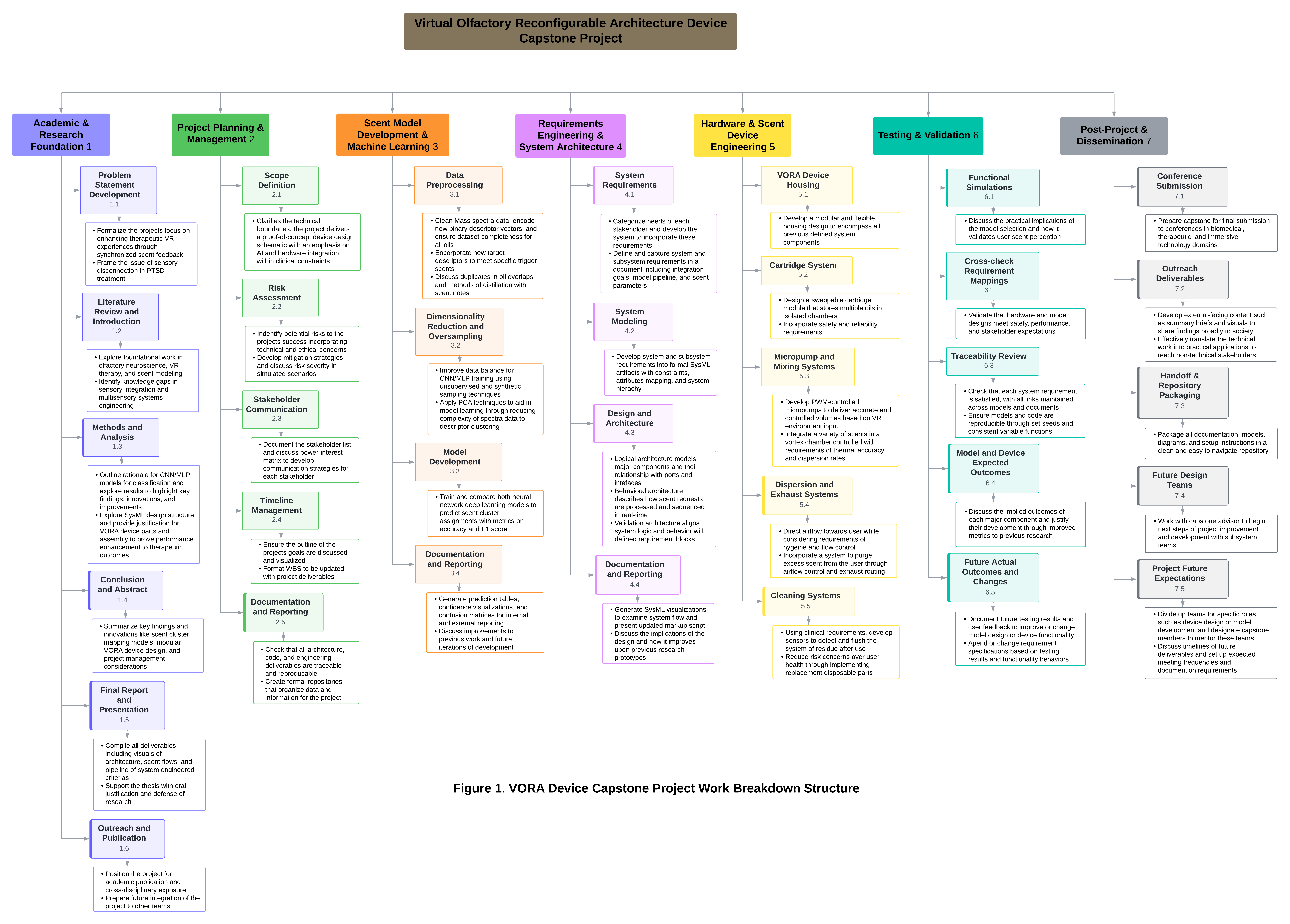
Figure 10: Monte Carlo Simulation of Project Risk Exposures



The P90 value indicates a high-end scenario, where 90% of the scenarios did not exceed the threshold of 275, whereas the median score currently stands at 260. Given the tight clustering near this median value of severity scores, small efforts to mitigate high severity risks, such as scent overlaps (R1) or failure in cluster representations in user-perceived scents (R9), could drastically reduce the overall project exposure. Using this data, the team can set risk tolerance thresholds, identify outlier risk, and allocate resources to mitigation proportionally. This overall analysis validated the effectiveness of key risk controls, such as sensor-based shutoff mechanisms and SMOTE techniques, both of which significantly reduced the probability and severity of data-related errors and high project severity scores.

A work breakdown structure, Figure 11, helps encompass all aspects of the capstone, including future development and potential full-scale models.

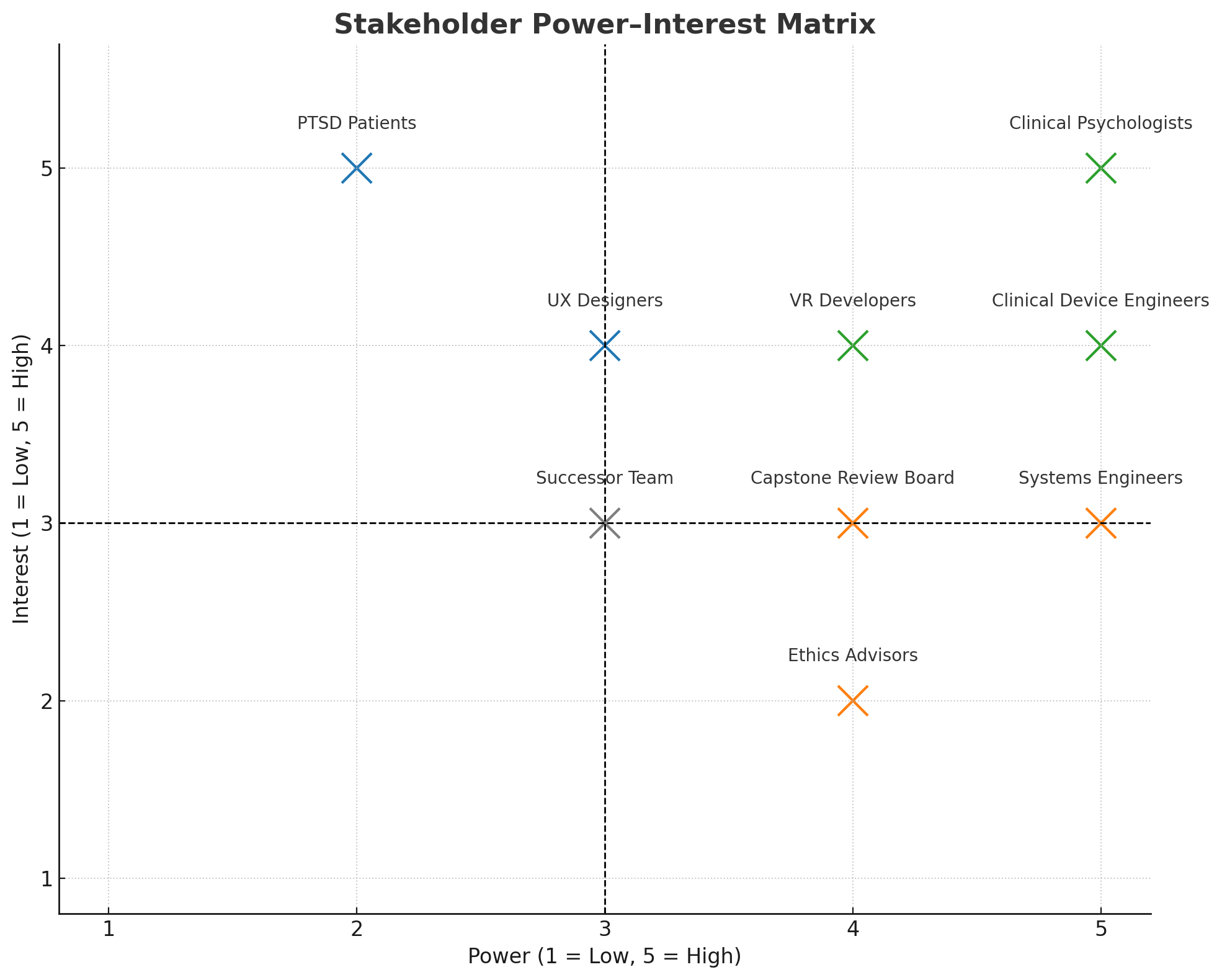
Figure 11: Work Breakdown Structure of Project



It includes deliverables for each structure and provides a chronological flow of events to ensure timeline alignment and clarified ownership. This structure facilitates the project's scope throughout its lifecycle, allowing the team to make adjustments and follow through on future objectives. The structure spans academic preparation, system architecture, scent modeling, hardware design, validation, and final dissemination. The foundation of the project is academic research and problem framing, which helps to ensure that the project is rooted in relevant literature, a thorough methodology, and sufficiently meets all the capstone requirements. The subsequent phases of the project focus on project planning, stakeholder communication, and systems engineering, emphasizing traceability and consistency. System design is structured around the architectures of the SysML blocks, which enable modular definition and flow control throughout the components. The device engineering and modeling sections highlight the implementation of subsystems that trace various requirements and outline clear steps taken to understand, process, and develop the two core components. The final stages include thorough testing and validation of all components, as well as the commencement of the next phase of the project, which involves ensuring that all necessary documentation and outreach channels are established. Future iterations of the project will involve forming subsystem teams and procuring and testing devices. The hierarchical structure of the WBS not only enables the team to manage tasks and assign ownership effectively, but also ensures that project goals and objectives are aligned and delivered systematically with clear justifications.

The stakeholder evaluation utilized a power-interest matrix, as shown in Figure 12, which mapped to different engagement strategies for each stakeholder, as outlined in Table 4. Understanding where the stakeholder is and how much information or power they require allows the team to make better-informed decisions when discussing what information to share.

Figure 12: Project Stakeholder Power-Interest Matrix



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Stakeholder** | **Role** | **Power (1–5)** | **Interest (1–5)** | **Engagement Strategy** | **Engagement Risk** |
| Clinical Psychologists | Therapy Integration Experts | 5 | 5 | Engage Closely | Unvalidated scents could trigger patients |
| PTSD Patients | End Users | 2 | 5 | Empower via Feedback | Poor UX or harmful scent exposure |
| VR Developers | Integration Engineers | 4 | 4 | Engage | Incompatibility with VR platforms |
| Systems Engineers | SysML Modelers | 5 | 3 | Consult and Support | Model traceability could fail |
| Clinical Device Engineers | Hardware Designers | 5 | 4 | Engage Closely | BOM mismatches or design flaws |
| Capstone Review Board | Academic Sponsors | 4 | 3 | Keep Satisfied | Project may not meet academic rigor |
| Successor Team | Future Developers | 3 | 3 | Inform and Prepare | Knowledge loss during handoff |
| UX Designers | Interface & Comfort Experts | 3 | 4 | Consult | Ergonomic or comfort issues not resolved |
| Ethics Advisors | Oversight and Compliance | 4 | 2 | Inform | Failure to meet safety or licensing guidelines |

Table 4: Project Stakeholder Analysis

Stakeholders, including clinicians, patients, developers, and review boards, were involved to ensure that both systems could be ethically and thoroughly evaluated from concept to deployment. Prioritizing engagement strategies for these stakeholders will enable the team to understand the level of communication better and provide the necessary information. Identifying stakeholders, such as clinical psychologists and device engineers, who occupy the high-power and high-interest quadrant, indicates that they require close, continuous involvement due to their central role in the integration of the device and the hardware's reliability. The patients and the UX designers fall into another quadrant, expressing their high interest in the project but low organizational power, which indicates to the team that they require proactive communication and iterative feedback loops. In contrast, stakeholders such as ethics advisors and review boards hold significant power over the project's success, but require fewer day-to-day updates, suggesting that they are informed on a milestone basis to maintain satisfaction and avoid excessive communication. Three of the four quadrants are filled except for the low-interest and low-power quadrant, which reflects the complexity and interdisciplinary nature of the project, where nearly all stakeholders have a critical role in ethical, design, and experiential values. Understanding each stakeholder's role and knowing the difference in communication strategies ensures that there is a communication path that properly engages with each stakeholder at an appropriate level of depth and frequency.

# Conclusion

This project successfully developed an integrated system that advances virtual reality therapy for PTSD through controlled olfactory feedback. At its core, the system comprises two key innovations: a predictive, machine-learned scent classification model and a modular, wearable scent-emitting device.

Using mass spectral data from essential oils and semantic clustering of odor descriptors, the model identified twelve clinically relevant scent profiles, including trauma-associated odors such as smoke, gunpowder, diesel, and burnt rubber. Among the modeling techniques evaluated, convolutional neural networks combined with principal component analysis demonstrated the highest accuracy and robustness, effectively capturing complex spectral features while mitigating class imbalance through synthetic oversampling techniques. These results validate the model’s ability to support precise and clinically relevant scent selection, addressing a critical technological gap in virtual reality exposure therapy platforms.

The hardware component was developed using rigorous systems engineering practices, including traceability, stakeholder engagement, and risk management. Transitioning from atomization-based delivery to micropump-driven dispersion improved scent precision, thermal safety, and scalability. Key features, including RFID oil pod tracking, UV sterilization, and disposable nozzles, ensure hygiene and device reliability in clinical environments. The modular design enables integration with existing VR headsets and allows for future expansion of scents. Together, these technical contributions represent a robust, user-centered platform that brings engineering leadership to therapeutic innovation.

Beyond its technical merit, this platform makes meaningful contributions to society by enhancing the effectiveness of virtual reality therapy for PTSD, supporting emotional processing through the deep neurological connection between scent, memory, and emotion. The system aligns with clinical protocols, prioritizes patient safety, and responds directly to the needs of stakeholders in mental health treatment. Additionally, this innovation has wide-ranging applications in education, professional training, and entertainment, where realistic multisensory immersion can improve engagement and learning outcomes.

By bridging chemical informatics, machine learning, and systems engineering, the project establishes a scalable framework for digitizing scent, a historically elusive sensory modality in virtual environments. The documented design approach offers a transferable blueprint not only for the deployment of VORA but also for future innovations in multisensory therapeutic systems. This work lays a strong foundation for continued research and development in olfactory VR, opening new possibilities for advancing immersive technologies in healthcare and beyond.

# Citation Sources

1. da Silva, M. F. B., Sanches, I. H., Borba, J. V. V. B., Barros, A. C. de A., Feitosa, F. L., de Carvalho, R. M., Filho, A. R. G., & Andrade, C. H. (2024). Elevating Virtual Reality Experiences with Olfactory Integration: A Preliminary Review. *Journal of the Brazilian Computer Society*, *30*(1), 639–652. https://doi.org/10.5753/jbcs.2024.4632
2. S Herz R. Olfactory Virtual Reality: A New Frontier in the Treatment and Prevention of Posttraumatic Stress Disorder. Brain Sci. 2021 Aug 16;11(8):1070. doi: 10.3390/brainsci11081070. PMID: 34439689; PMCID: PMC8391796.
3. Barfield, W.; Danas, E. Comments on the use of olfactory displays for virtual environments. Presence Teleoperators Virtual Environ. 1996, 5, 109–121.
4. Munyan, B.G., III; Neer, S.M.; Beidel, D.C.; Jentsch, F. Olfactory stimuli increase presence in virtual environments. PLoS ONE 2016, 11, e0157568.
5. Ischer, M.; Baron, N.; Mermoud, C.; Cayeux, I.; Porcherot, C.; Sander, D.; Delplanque, S. How incorporation of scents could enhance immersive virtual experiences. Front. Psychol. 2014, 5, 736.
6. Debnath T, Nakamoto T (2020) Predicting human odor perception represented by continuous values from mass spectra of essential oils resembling chemical mixtures. PLoS ONE 15(6): e0234688. https://doi.org/10.1371/journal.pone.0234688
7. Zou, Zhe & Prasetyawan, Dani & Wu, Hsueh & Cheng, Kelvin & Nakamoto, Takamichi. (2024). Extension of Wearable Olfactory Display for Multisensory VR Experience. 10.2312/egve.20241374.
8. Aleixandre, M., Prasetyawan, D. & Nakamoto, T. Automatic scent creation by cheminformatics method. *Sci Rep* 14, 31284 (2024). https://doi.org/10.1038/s41598-024-82654-7
9. Nakamoto, T., Ohno, M. & Nihei, Y. Odor approximation using mass spectrometry. *IEEE Sens. J.* **12**, 3225–3231. h t t p s : / / d o i . o r g / 1 0 . 1 1 0 9 / J S E N . 2 0 1 2 . 2 1 9 0 5 0 6 (2012).
10. Prasetyawan, D. & Nakamoto, T. Odor reproduction technology using a small set of odor components. *IEEJ Trans. Elec Electron. Eng.* **19**, 4–14. https://doi.org/10.1002/tee.23915 (2024).
11. Gutiérrez, E. D. et al. Predicting natural language descriptions of mono-molecular odorants. *Nat. Commun.* **9**, 4979. h t t p s : / / d o i . o r g / 1 0 . 1 0 3 8 / s 4 1 4 6 7 - 0 1 8 - 0 7 4 3 9 - 9 (2018).
12. Keller, A. et al. Predicting human olfactory perception from chemical features of odor molecules. *Science* **355**, 820–826. h t t p s : / / d o i . o r g / 1 0 . 1 1 2 6 / s c i e n c e . a a l 2 0 1 4 (2017).
13. Khan, R. M. et al. Predicting odor pleasantness from odorant structure: Pleasantness as a reflection of the physical world. *J. Neurosci.* **27**, 10015–10023. https://doi.org/10.1523/JNEUROSCI.1158-07.2007 (2007).
14. Sanchez-Lengeling et al. Machine learning for scent: Learning generalizable perceptual representations of small molecules. *arXiv preprint* arXiv:1910.10685 (2019). https://doi.org/10.48550/arXiv.1910.10685
15. Lee, B. K. et al. A principal odor map unifies diverse tasks in olfactory perception. *Science* **381**, 999–1006. h t t p s : / / d o i . o r g / 1 0 . 1 1 2 6 / s c i e n c e . a d e 4 4 0 1 (2023).
16. Nozaki, Y. & Nakamoto, T. Odor impression prediction from mass spectra. *PLoS One*. **11**, e0157030. h t t p s : / / d o i . o r g / 1 0 . 1 3 7 1 / j o u r n a l . p o n e . 0 1 5 7 0 3 0 (2016)
17. Nozaki, Y. & Nakamoto, T. Predictive modeling for odor character of a chemical using machine learning combined with natural language processing. *PLoS ONE*. **13**, e0198475. https://doi.org/10.1371/journal.pone.0198475 (2018).
18. Debnath, T., Prasetyawan, D. & Nakamoto, T. Predicting odor perception of mixed scent from mass spectrometry. *J. Electrochem. Soc.* **168**, 117505. https://doi.org/10.1149/1945-7111/ac33e0 (2021).
19. Hasebe, D., Aleixandre, M. & Nakamoto, T. Exploration of sensing data to realize intended odor impression using mass spectrum of odor mixture. *PLoS ONE*. **17**, e0273011. https://doi.org/10.1371/journal.pone.0273011 (2022).
20. Meilgaard, M. C., Carr, B. T. & Carr, B. T. *Sensory Evaluation Techniques*. ISBN: 9780429195143. https://doi.org/10.1201/b16452 (CRC PressLondon , 2016).
21. Nakamoto T., Essentials of Machine Olfaction and Taste, vol. 1. wiley, 2016.
22. Buck L. and Axel R., “A novel multigene family may encode odorant receptors: A molecular basis for odor recognition,” Cell, vol. 65, no. 1, pp. 175–187, Apr. 1991. https://doi.org/10.1016/0092-8674(91)
23. T. Debnath, D. Prasetyawan, and T. Nakamoto, “Prediction of Odor Descriptor Group of Essential Oils from Mass Spectra using Machine Learning,” in 2019 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), 2019, pp. 1–3. https://doi.org/10.1109/ISOEN.2019.8823226
24. Debnath T, Nakamoto T (2020) Predicting human odor perception represented by continuous values from mass spectra of essential oils resembling chemical mixtures. PLoS ONE 15(6):e0234688. https://doi.org/10.1371/journal.pone.0234688
25. Nozaki Y, Nakamoto T (2018) Correction: Predictive modeling for odor character of a chemical using machine learning combined with natural language processing. PLOS ONE 13(12): e0208962. https:// doi.org/10.1371/journal.pone.0208962
26. Amann, Janek & Agirrezabal, Manex. (2022). From meaning to perception -- exploring the space between word and odor perception embeddings. 10.48550/arXiv.2203.10294.
27. Asifa Majid, Seán G Roberts, Ludy Cilissen, Karen Emmorey, Brenda Nicodemus, Lucinda O’grady, Bencie Woll, Barbara LeLan, Hilário De Sousa, Brian L Cansler, et al. Differential coding of perception in the world’s languages. Proceedings of the National Academy of Sciences, 115(45):11369–11376, 2018.
28. Jonas K Olofsson and Jay A Gottfried. The muted sense: neurocognitive limitations of

olfactory language. Trends in cognitive sciences, 19(6):314–321, 2015.

1. Yaara Yeshurun and Noam Sobel. An odor is not worth a thousand words: from multidimensional odors to unidimensional odor objects. Annual review of psychology, 61:219–241, 2010.
2. Rachel S Herz. 17 perfume. Neurobiology of Sensation and Reward, page 371, 2011.
3. Asifa Majid. Human olfaction at the intersection of language, culture, and biology. Trends in Cognitive Sciences, 2020.
4. Asifa Majid and Niclas Burenhult. Odors are expressible in language, as long as you speak the right language. Cognition, 130(2):266–270, 2014.
5. T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations.
6. T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean. Distributed representations of.odors to unidimensional odor objects. Annual review of psychology, 61:219–241, 2010.
7. Keller, A., Vosshall, L.B. Olfactory perception of chemically diverse molecules. *BMC Neurosci* **17**, 55 (2016). https://doi.org/10.1186/s12868-016-0287-2
8. Todeschini R, Consonni V. Molecular descriptors for chemoinformatics, alphabetical listing, vol. 1. New York: Wiley; 2009.
9. Todeschini R, Consonni V. Molecular descriptors for chemoinformatics. appendices, bibliography, vol. 2. New York: Wiley; 2009.
10. Menashe I, Man O, Lancet D, Gilad Y. Different noses for different people. Nat Genet. 2003;34(2):143–4.
11. Keller A, Zhuang H, Chi Q, Vosshall LB, Matsunami H. Genetic variation in a human odorant receptor alters odour perception. Nature. 2007;449(7161):468–72.
12. Olender T, Waszak SM, Viavant M, Khen M, Ben-Asher E, Reyes A, Nativ N, Wysocki CJ, Ge DL, Lancet D. Personal receptor repertoires: olfaction as a model. BMC Genom. 2012;13:414.
13. Keydar I, Ben-Asher E, Feldmesser E, Nativ N, Oshimoto A, Restrepo D, Matsunami H, Chien M-S, Pinto JM, Gilad Y, et al. General olfactory sensitivity database (GOSdb): candidate genes and their genomic variations. Hum Mutat. 2013;34(1):32–41.
14. Mainland JD, Keller A, Li YR, Zhou T, Trimmer C, Snyder LL, Moberly AH, Adipietro KA, Liu WLL, Zhuang H, et al. The missense of smell: functional variability in the human odorant receptor repertoire. Nat Neurosci. 2014;17(1):114–20.
15. Keller, A., & Vosshall, L. B. (2016). The sensory perception of odor. *BMC Neuroscience, 17*(55). https://doi.org/10.1186/s12868-016-0287-2
16. Ndablog. (2025, May 2). *UNTAPPING THE POWER OF NATURE: ESSENTIAL OIL EXTRACTION METHODS*. Blog - New Directions Aromatics Inc. https://www.newdirectionsaromatics.com/blog/untapping-the-power-of-nature-essential-oil-extraction-methods/
17. Cortese BM, Leslie K, Uhde TW. Differential odor sensitivity in PTSD: Implications for treatment and future research. J Affect Disord. 2015 Jul 1;179:23-30. doi: 10.1016/j.jad.2015.03.026. Epub 2015 Mar 31. PMID: 25845746; PMCID: PMC4437877.
18. Asma Iqbal, Mohammad Amil Bhat, Qazi Muneeb, Muazam Javid,Revolutionizing perfume creation: PTD's innovative approach, Digital Chemical Engineering, Volume 15, 2025,100223, ISSN 2772-5081,https://doi.org/10.1016/j.dche.2025.100223.
19. OpenAI. (2025). *ChatGPT (July 2025 version) [Large language model]*. https://chat.openai.com/

# Appendices

APPENDIX A: MODEL AND VISUAL GENERATION CODE

|  |
| --- |
| **MASTER MODEL Comparison python script** |
| import pandas as pd  import os  # === Define paths ===  base\_dir = "C:/Users/Capstone - VR Sense/final model"  cnn\_with\_pca\_dir = os.path.join(base\_dir, "cnn with pca")  cnn\_without\_pca\_dir = os.path.join(base\_dir, "cnn without pca")  mlp\_with\_pca\_dir = os.path.join(base\_dir, "mlp with pca")  mlp\_without\_pca\_dir = os.path.join(base\_dir, "mlp without pca")  model\_paths = {  "cnn\_with\_pca": os.path.join(cnn\_with\_pca\_dir, "cnn\_pca\_predictions\_with\_confidence.csv"),  "cnn\_without\_pca": os.path.join(cnn\_without\_pca\_dir, "cnn\_predictions\_with\_confidence.csv"),  "mlp\_with\_pca": os.path.join(mlp\_with\_pca\_dir, "mlp\_pca\_predictions\_with\_confidence.csv"),  "mlp\_without\_pca": os.path.join(mlp\_without\_pca\_dir, "mlp\_predictions\_with\_confidence.csv"),  }  results = []  for model\_name, path in model\_paths.items():  if not os.path.exists(path):  print(f"⚠️ Missing file for {model\_name}: {path}")  continue  df = pd.read\_csv(path)  if {"True Cluster", "Predicted Cluster", "Confidence"}.issubset(df.columns):  accuracy = (df["True Cluster"] == df["Predicted Cluster"]).mean()  f1 = 2 \* (accuracy \* accuracy) / (accuracy + accuracy) if accuracy > 0 else 0  confidence = df["Confidence"].mean()  results.append({  "Model": model\_name,  "Accuracy": round(accuracy, 4),  "F1 Score": round(f1, 4),  "Mean Confidence": round(confidence, 4)  })  print(f"📊 {model\_name}")  print(f" - Accuracy: {accuracy:.4f}")  print(f" - F1 Score: {f1:.4f}")  print(f" - Mean Confidence: {confidence:.4f}")  else:  print(f"❌ File for {model\_name} is missing required columns.")  # === Save comparison summary ===  summary\_df = pd.DataFrame(results)  summary\_csv = os.path.join(base\_dir, "model\_comparison\_summary.csv")  summary\_df.to\_csv(summary\_csv, index=False)  print(f"\n✅ Comparison summary saved to: {summary\_csv}") |

|  |
| --- |
| **CNN model with pca python script** |
| import pandas as pd  import numpy as np  import os  import random  import tensorflow as tf  from sklearn.preprocessing import StandardScaler  from sklearn.model\_selection import StratifiedShuffleSplit  from sklearn.decomposition import PCA  from sklearn.metrics import accuracy\_score, classification\_report, f1\_score, confusion\_matrix  from tensorflow.keras.models import Sequential  from tensorflow.keras.layers import Conv1D, MaxPooling1D, Flatten, Dense, Dropout  from tensorflow.keras.utils import to\_categorical  from imblearn.over\_sampling import SMOTE  import seaborn as sns  import matplotlib.pyplot as plt  # === Set seeds for reproducibility ===  SEED = 777  random.seed(SEED)  np.random.seed(SEED)  tf.random.set\_seed(SEED)  # === Set paths ===  base\_dir = "C:/Users/Capstone - VR Sense"  xlsx\_path = os.path.join(base\_dir, "EODataSets.xlsx")  output\_dir = os.path.join(base\_dir, "final model", "cnn with pca")  os.makedirs(output\_dir, exist\_ok=True)  # === Load Excel data ===  spectra\_df = pd.read\_excel(xlsx\_path, sheet\_name="MassSpectraDataSet")  odor\_df = pd.read\_excel(xlsx\_path, sheet\_name="OdorDataSet")  spectra\_df["Essential Oil"] = spectra\_df["Essential Oil"].str.strip()  odor\_df["Essential Oil"] = odor\_df["Essential Oil"].str.strip()  odor\_df.set\_index("Essential Oil", inplace=True)  # === Descriptor to cluster mapping ===  descriptor\_to\_cluster = {  'Smoky': 1, 'Phenolic': 1, 'Burnt': 1, 'Medicinal': 1,  'Rubber': 2, 'Sulfurous': 2, 'Acrid': 2, 'Metallic': 2,  'Sharp': 3, 'Dry': 3, 'Powdery': 3, 'Leathery': 3,  'Musky': 4, 'Animalic': 4, 'Pungent': 4, 'Oily': 4,  'Green': 5, 'Grassy': 5, 'Herbaceous': 5, 'Leafy': 5, 'Grass': 5,  'Citrus': 6, 'Lemon': 6, 'Orange': 6, 'Citronellal': 6, 'Pleasant': 6,  'Floral': 7, 'Rose': 7, 'Tea': 7, 'Light': 7,  'Spicy': 8, 'Sweet': 8, 'Fatty': 8, 'Anisic': 8, 'Balsamic': 8,  'Woody': 9, 'Root': 9, 'Earthy': 9, 'Resinous': 9, 'Pine': 9,  'Fresh': 10, 'Minty': 10, 'Cooling': 10, 'Eucalyptol': 10,  'Terpene': 11, 'Aromatic': 11, 'Peppery': 11, 'Anethole': 11,  'Warm': 12, 'Rich': 12, 'Bitter': 12, 'Fruity': 12  }  # === Assign dominant cluster ===  oil\_to\_cluster = {}  for oil, row in odor\_df.iterrows():  cluster\_counts = {}  for descriptor, present in row.items():  if present == 1 and descriptor in descriptor\_to\_cluster:  cluster = descriptor\_to\_cluster[descriptor]  cluster\_counts[cluster] = cluster\_counts.get(cluster, 0) + 1  if cluster\_counts:  dominant\_cluster = max(cluster\_counts.items(), key=lambda x: (x[1], -x[0]))[0]  oil\_to\_cluster[oil] = dominant\_cluster  # === Filter and align ===  spectra\_df = spectra\_df[spectra\_df["Essential Oil"].isin(oil\_to\_cluster.keys())]  spectra\_df["Cluster"] = spectra\_df["Essential Oil"].map(oil\_to\_cluster)  X = spectra\_df.drop(columns=["Essential Oil", "Cluster"]).values  y = spectra\_df["Cluster"].values  oil\_names = spectra\_df["Essential Oil"].values  # === PCA for 2D visualization ===  pca\_vis = PCA(n\_components=2)  X\_pca = pca\_vis.fit\_transform(X)  plt.figure(figsize=(6,5))  sns.scatterplot(x=X\_pca[:,0], y=X\_pca[:,1], hue=y, palette="tab10")  plt.title("PCA 2D View of Oils by Cluster")  plt.savefig(os.path.join(output\_dir, "pca\_2d\_visualization.png"))  plt.close()  # === PCA for modeling ===  pca = PCA(n\_components=20)  X\_pca\_model = pca.fit\_transform(X)  # === SMOTE ===  scaler = StandardScaler()  X\_scaled = scaler.fit\_transform(X\_pca\_model)  sm = SMOTE(random\_state=SEED, k\_neighbors=1)  X\_resampled, y\_resampled = sm.fit\_resample(X\_scaled, y)  oil\_names\_resampled = np.array(oil\_names)[sm.fit\_resample(np.arange(len(oil\_names)).reshape(-1, 1), y)[0].flatten()]  # === Plot class balance ===  plt.figure(figsize=(8,4))  sns.countplot(x=y\_resampled)  plt.title("Cluster Counts After SMOTE")  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "post\_smote\_cluster\_counts.png"))  plt.close()  # === Stratified split ===  splitter = StratifiedShuffleSplit(n\_splits=1, test\_size=0.3, random\_state=SEED)  for train\_idx, test\_idx in splitter.split(X\_resampled, y\_resampled):  X\_train, X\_test = X\_resampled[train\_idx], X\_resampled[test\_idx]  y\_train\_raw, y\_test\_raw = y\_resampled[train\_idx], y\_resampled[test\_idx]  oil\_names\_test = oil\_names\_resampled[test\_idx]  y\_train = to\_categorical(y\_train\_raw - 1, num\_classes=12)  y\_test = to\_categorical(y\_test\_raw - 1, num\_classes=12)  # === Reshape for CNN ===  X\_train = X\_train[..., np.newaxis]  X\_test = X\_test[..., np.newaxis]  # === CNN model ===  model = Sequential([  Conv1D(32, 3, activation='relu', padding='same', input\_shape=(X\_train.shape[1], 1)),  MaxPooling1D(2),  Dropout(0.3),  Flatten(),  Dense(64, activation='relu'),  Dense(12, activation='softmax')  ])  model.compile(optimizer="adam", loss="categorical\_crossentropy", metrics=["accuracy"])  history = model.fit(X\_train, y\_train, epochs=50, batch\_size=8, validation\_split=0.2)  # === Training Curves ===  plt.figure(figsize=(8,4))  plt.plot(history.history['loss'], label='Train Loss')  plt.plot(history.history['val\_loss'], label='Val Loss')  plt.title("CNN Loss Curve")  plt.legend()  plt.savefig(os.path.join(output\_dir, "loss\_curve.png"))  plt.close()  plt.figure(figsize=(8,4))  plt.plot(history.history['accuracy'], label='Train Acc')  plt.plot(history.history['val\_accuracy'], label='Val Acc')  plt.title("CNN Accuracy Curve")  plt.legend()  plt.savefig(os.path.join(output\_dir, "accuracy\_curve.png"))  plt.close()  # === Evaluation ===  y\_pred\_probs = model.predict(X\_test)  y\_pred = np.argmax(y\_pred\_probs, axis=1) + 1  y\_true = np.argmax(y\_test, axis=1) + 1  # === Confusion Matrix ===  conf = confusion\_matrix(y\_true, y\_pred)  plt.figure(figsize=(8,6))  sns.heatmap(conf, annot=True, fmt='d', cmap="Blues", xticklabels=range(1,13), yticklabels=range(1,13))  plt.title("Confusion Matrix")  plt.xlabel("Predicted")  plt.ylabel("True")  plt.savefig(os.path.join(output\_dir, "confusion\_matrix.png"))  plt.close()  # === Per-Cluster Accuracy ===  cluster\_acc = {}  for i in range(1, 13):  mask = y\_true == i  if mask.sum() > 0:  acc = (y\_pred[mask] == i).sum() / mask.sum()  cluster\_acc[i] = acc  plt.figure(figsize=(10,4))  sns.barplot(x=list(cluster\_acc.keys()), y=list(cluster\_acc.values()), palette="crest", hue=list(cluster\_acc.keys()), legend=False)  plt.title("Per-Cluster Accuracy")  plt.ylim(0, 1)  plt.savefig(os.path.join(output\_dir, "cluster\_accuracy.png"))  plt.close()  # === Confidence Histogram ===  confidences = np.max(y\_pred\_probs, axis=1)  plt.figure(figsize=(6,4))  sns.histplot(confidences, bins=10, kde=True)  plt.title("Prediction Confidence Distribution")  plt.savefig(os.path.join(output\_dir, "confidence\_distribution.png"))  plt.close()  # === Metrics and Report ===  print(f"✅ Accuracy: {accuracy\_score(y\_true, y\_pred):.4f}")  print(f"🎯 F1 Score: {f1\_score(y\_true, y\_pred, average='weighted'):.4f}")  print("\n📋 Classification Report:")  print(classification\_report(y\_true, y\_pred))  # === Save Predictions with Confidence ===  pred\_df = pd.DataFrame({  "Oil": oil\_names\_test,  "True Cluster": y\_true,  "Predicted Cluster": y\_pred,  "Confidence": confidences  })  pred\_df.to\_csv(os.path.join(output\_dir, "cnn\_pca\_predictions\_with\_confidence.csv"), index=False) |

|  |
| --- |
| **CNN model without pca python script** |
| import pandas as pd  import numpy as np  import os  import random  import tensorflow as tf  from sklearn.preprocessing import StandardScaler  from sklearn.model\_selection import StratifiedShuffleSplit  from sklearn.decomposition import PCA  from sklearn.metrics import accuracy\_score, classification\_report, f1\_score, confusion\_matrix  from tensorflow.keras.models import Sequential  from tensorflow.keras.layers import Conv1D, MaxPooling1D, Flatten, Dense, Dropout  from tensorflow.keras.utils import to\_categorical  from imblearn.over\_sampling import SMOTE  import seaborn as sns  import matplotlib.pyplot as plt  # === Set seeds for reproducibility ===  SEED = 777  random.seed(SEED)  np.random.seed(SEED)  tf.random.set\_seed(SEED)  # === Set paths ===  base\_dir = "C:/Users/Capstone - VR Sense"  xlsx\_path = os.path.join(base\_dir, "EODataSets.xlsx")  output\_dir = os.path.join(base\_dir, "final model", "cnn without pca")  os.makedirs(output\_dir, exist\_ok=True)  # === Load Excel data ===  spectra\_df = pd.read\_excel(xlsx\_path, sheet\_name="MassSpectraDataSet")  odor\_df = pd.read\_excel(xlsx\_path, sheet\_name="OdorDataSet")  spectra\_df["Essential Oil"] = spectra\_df["Essential Oil"].str.strip()  odor\_df["Essential Oil"] = odor\_df["Essential Oil"].str.strip()  odor\_df.set\_index("Essential Oil", inplace=True)  # === Descriptor to cluster mapping ===  descriptor\_to\_cluster = {  'Smoky': 1, 'Phenolic': 1, 'Burnt': 1, 'Medicinal': 1,  'Rubber': 2, 'Sulfurous': 2, 'Acrid': 2, 'Metallic': 2,  'Sharp': 3, 'Dry': 3, 'Powdery': 3, 'Leathery': 3,  'Musky': 4, 'Animalic': 4, 'Pungent': 4, 'Oily': 4,  'Green': 5, 'Grassy': 5, 'Herbaceous': 5, 'Leafy': 5, 'Grass': 5,  'Citrus': 6, 'Lemon': 6, 'Orange': 6, 'Citronellal': 6, 'Pleasant': 6,  'Floral': 7, 'Rose': 7, 'Tea': 7, 'Light': 7,  'Spicy': 8, 'Sweet': 8, 'Fatty': 8, 'Anisic': 8, 'Balsamic': 8,  'Woody': 9, 'Root': 9, 'Earthy': 9, 'Resinous': 9, 'Pine': 9,  'Fresh': 10, 'Minty': 10, 'Cooling': 10, 'Eucalyptol': 10,  'Terpene': 11, 'Aromatic': 11, 'Peppery': 11, 'Anethole': 11,  'Warm': 12, 'Rich': 12, 'Bitter': 12, 'Fruity': 12  }  # === Assign dominant cluster ===  oil\_to\_cluster = {}  for oil, row in odor\_df.iterrows():  cluster\_counts = {}  for descriptor, present in row.items():  if present == 1 and descriptor in descriptor\_to\_cluster:  cluster = descriptor\_to\_cluster[descriptor]  cluster\_counts[cluster] = cluster\_counts.get(cluster, 0) + 1  if cluster\_counts:  dominant\_cluster = max(cluster\_counts.items(), key=lambda x: (x[1], -x[0]))[0]  oil\_to\_cluster[oil] = dominant\_cluster  # === Filter and align ===  spectra\_df = spectra\_df[spectra\_df["Essential Oil"].isin(oil\_to\_cluster.keys())]  spectra\_df["Cluster"] = spectra\_df["Essential Oil"].map(oil\_to\_cluster)  X = spectra\_df.drop(columns=["Essential Oil", "Cluster"]).values  y = spectra\_df["Cluster"].values  oil\_names = spectra\_df["Essential Oil"].values # aligned with X and y  # === PCA for 2D visualization ===  pca = PCA(n\_components=2)  X\_pca = pca.fit\_transform(X)  plt.figure(figsize=(6,5))  sns.scatterplot(x=X\_pca[:,0], y=X\_pca[:,1], hue=y, palette="tab10")  plt.title("PCA 2D View of Oils by Cluster")  plt.savefig(os.path.join(output\_dir, "pca\_2d\_visualization.png"))  plt.close()  # === SMOTE ===  scaler = StandardScaler()  X\_scaled = scaler.fit\_transform(X)  sm = SMOTE(random\_state=SEED, k\_neighbors=1)  X\_resampled, y\_resampled = sm.fit\_resample(X\_scaled, y)  # === Track real oils only ===  real\_indices = np.arange(len(y)) # only original samples  oil\_names\_resampled = np.concatenate([  oil\_names,  np.array(["(synthetic)"] \* (len(X\_resampled) - len(oil\_names)))  ])  # === Stratified split ===  splitter = StratifiedShuffleSplit(n\_splits=1, test\_size=0.3, random\_state=SEED)  for train\_idx, test\_idx in splitter.split(X\_resampled, y\_resampled):  X\_train, X\_test = X\_resampled[train\_idx], X\_resampled[test\_idx]  y\_train\_raw, y\_test\_raw = y\_resampled[train\_idx], y\_resampled[test\_idx]  oil\_names\_train = oil\_names\_resampled[train\_idx]  oil\_names\_test = oil\_names\_resampled[test\_idx]  y\_train = to\_categorical(y\_train\_raw - 1, num\_classes=12)  y\_test = to\_categorical(y\_test\_raw - 1, num\_classes=12)  # === CNN Input Reshape ===  X\_train = X\_train[..., np.newaxis]  X\_test = X\_test[..., np.newaxis]  # === CNN model ===  model = Sequential([  Conv1D(32, 3, activation='relu', padding='same', input\_shape=(X\_train.shape[1], 1)),  MaxPooling1D(2),  Dropout(0.3),  Flatten(),  Dense(64, activation='relu'),  Dense(12, activation='softmax')  ])  model.compile(optimizer="adam", loss="categorical\_crossentropy", metrics=["accuracy"])  history = model.fit(X\_train, y\_train, epochs=50, batch\_size=8, validation\_split=0.2)  # === Plot Loss/Accuracy ===  plt.figure(figsize=(8,4))  plt.plot(history.history['loss'], label='Train Loss')  plt.plot(history.history['val\_loss'], label='Val Loss')  plt.title("CNN Loss Curve")  plt.legend()  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "loss\_curve.png"))  plt.close()  plt.figure(figsize=(8,4))  plt.plot(history.history['accuracy'], label='Train Acc')  plt.plot(history.history['val\_accuracy'], label='Val Acc')  plt.title("CNN Accuracy Curve")  plt.legend()  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "accuracy\_curve.png"))  plt.close()  # === Evaluation ===  y\_pred\_probs = model.predict(X\_test)  y\_pred = np.argmax(y\_pred\_probs, axis=1) + 1  y\_true = np.argmax(y\_test, axis=1) + 1  confidences = np.max(y\_pred\_probs, axis=1)  # === Confusion Matrix ===  conf = confusion\_matrix(y\_true, y\_pred)  plt.figure(figsize=(8,6))  sns.heatmap(conf, annot=True, fmt='d', cmap="Blues", xticklabels=range(1,13), yticklabels=range(1,13))  plt.title("Confusion Matrix")  plt.xlabel("Predicted")  plt.ylabel("True")  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "confusion\_matrix.png"))  plt.close()  # === Per-Cluster Accuracy ===  cluster\_acc = {}  for i in range(1, 13):  mask = y\_true == i  if mask.sum() > 0:  acc = (y\_pred[mask] == i).sum() / mask.sum()  cluster\_acc[i] = acc  plt.figure(figsize=(10,4))  sns.barplot(x=list(cluster\_acc.keys()), y=list(cluster\_acc.values()), palette="crest", hue=list(cluster\_acc.keys()), legend=False)  plt.title("Per-Cluster Accuracy")  plt.ylim(0, 1)  plt.savefig(os.path.join(output\_dir, "cluster\_accuracy.png"))  plt.close()  # === Confidence Histogram ===  plt.figure(figsize=(6,4))  sns.histplot(confidences, bins=10, kde=True)  plt.title("Prediction Confidence Distribution")  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "confidence\_distribution.png"))  plt.close()  # === Evaluation Printout ===  print(f"✅ Accuracy: {accuracy\_score(y\_true, y\_pred):.4f}")  print(f"🎯 F1 Score: {f1\_score(y\_true, y\_pred, average='weighted'):.4f}")  print("\n📋 Classification Report:")  print(classification\_report(y\_true, y\_pred))  # === Predict on original oils (before SMOTE) ===  X\_real\_scaled = scaler.transform(X) # scale original oils  X\_real\_input = X\_real\_scaled[..., np.newaxis]  real\_probs = model.predict(X\_real\_input)  real\_preds = np.argmax(real\_probs, axis=1) + 1  real\_confidences = np.max(real\_probs, axis=1)  # === Output CSV: All oils mapped to predicted clusters with confidence ===  full\_output = pd.DataFrame({  "Oil": oil\_names,  "True Cluster": y,  "Predicted Cluster": real\_preds,  "Confidence": real\_confidences  })  full\_output.to\_csv(os.path.join(output\_dir, "cnn\_predictions\_with\_confidence.csv"), index=False) |

|  |
| --- |
| **MLP MODEL WITH pca python script** |
| import pandas as pd  import numpy as np  import os  import random  import tensorflow as tf  from sklearn.preprocessing import StandardScaler  from sklearn.decomposition import PCA  from sklearn.model\_selection import StratifiedShuffleSplit  from sklearn.metrics import accuracy\_score, classification\_report, f1\_score, confusion\_matrix  from tensorflow.keras.models import Sequential  from tensorflow.keras.layers import Dense, Dropout  from tensorflow.keras.utils import to\_categorical  from imblearn.over\_sampling import SMOTE  import seaborn as sns  import matplotlib.pyplot as plt  # === Set seeds for reproducibility ===  SEED = 777  random.seed(SEED)  np.random.seed(SEED)  tf.random.set\_seed(SEED)  # === Paths ===  base\_dir = "C:/Users/Capstone - VR Sense"  xlsx\_path = os.path.join(base\_dir, "EODataSets.xlsx")  output\_dir = os.path.join(base\_dir, "final model", "mlp with pca")  os.makedirs(output\_dir, exist\_ok=True)  # === Load data ===  spectra\_df = pd.read\_excel(xlsx\_path, sheet\_name="MassSpectraDataSet")  odor\_df = pd.read\_excel(xlsx\_path, sheet\_name="OdorDataSet")  spectra\_df["Essential Oil"] = spectra\_df["Essential Oil"].str.strip()  odor\_df["Essential Oil"] = odor\_df["Essential Oil"].str.strip()  odor\_df.set\_index("Essential Oil", inplace=True)  # === Descriptor to cluster mapping ===  descriptor\_to\_cluster = {  'Smoky': 1, 'Phenolic': 1, 'Burnt': 1, 'Medicinal': 1,  'Rubber': 2, 'Sulfurous': 2, 'Acrid': 2, 'Metallic': 2,  'Sharp': 3, 'Dry': 3, 'Powdery': 3, 'Leathery': 3,  'Musky': 4, 'Animalic': 4, 'Pungent': 4, 'Oily': 4,  'Green': 5, 'Grassy': 5, 'Herbaceous': 5, 'Leafy': 5, 'Grass': 5,  'Citrus': 6, 'Lemon': 6, 'Orange': 6, 'Citronellal': 6, 'Pleasant': 6,  'Floral': 7, 'Rose': 7, 'Tea': 7, 'Light': 7,  'Spicy': 8, 'Sweet': 8, 'Fatty': 8, 'Anisic': 8, 'Balsamic': 8,  'Woody': 9, 'Root': 9, 'Earthy': 9, 'Resinous': 9, 'Pine': 9,  'Fresh': 10, 'Minty': 10, 'Cooling': 10, 'Eucalyptol': 10,  'Terpene': 11, 'Aromatic': 11, 'Peppery': 11, 'Anethole': 11,  'Warm': 12, 'Rich': 12, 'Bitter': 12, 'Fruity': 12  }  # === Assign dominant cluster ===  oil\_to\_cluster = {}  for oil, row in odor\_df.iterrows():  cluster\_counts = {}  for descriptor, present in row.items():  if present == 1 and descriptor in descriptor\_to\_cluster:  cluster = descriptor\_to\_cluster[descriptor]  cluster\_counts[cluster] = cluster\_counts.get(cluster, 0) + 1  if cluster\_counts:  dominant\_cluster = max(cluster\_counts.items(), key=lambda x: (x[1], -x[0]))[0]  oil\_to\_cluster[oil] = dominant\_cluster  # === Filter and align ===  spectra\_df = spectra\_df[spectra\_df["Essential Oil"].isin(oil\_to\_cluster.keys())]  spectra\_df["Cluster"] = spectra\_df["Essential Oil"].map(oil\_to\_cluster)  X = spectra\_df.drop(columns=["Essential Oil", "Cluster"]).values  y = spectra\_df["Cluster"].values  oil\_names = spectra\_df["Essential Oil"].values  # === PCA reduction ===  scaler = StandardScaler()  X\_scaled = scaler.fit\_transform(X)  pca = PCA(n\_components=15)  X\_pca = pca.fit\_transform(X\_scaled)  # === Pre-SMOTE cluster distribution ===  plt.figure(figsize=(8, 4))  sns.countplot(x=y)  plt.title("Cluster Distribution Before SMOTE")  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "pre\_smote\_cluster\_counts.png"))  plt.close()  # === SMOTE ===  sm = SMOTE(random\_state=SEED, k\_neighbors=1)  X\_resampled, y\_resampled = sm.fit\_resample(X\_pca, y)  # === Oil name tracking for predictions ===  # Track original oils and fill synthetic ones randomly  oil\_names\_resampled = np.concatenate([  oil\_names,  np.random.choice(oil\_names, size=len(X\_resampled) - len(X), replace=True)  ])  # === Stratified Split ===  splitter = StratifiedShuffleSplit(n\_splits=1, test\_size=0.3, random\_state=SEED)  for train\_idx, test\_idx in splitter.split(X\_resampled, y\_resampled):  X\_train, X\_test = X\_resampled[train\_idx], X\_resampled[test\_idx]  y\_train\_raw, y\_test\_raw = y\_resampled[train\_idx], y\_resampled[test\_idx]  test\_oils = oil\_names\_resampled[test\_idx]  y\_train = to\_categorical(y\_train\_raw - 1, num\_classes=12)  y\_test = to\_categorical(y\_test\_raw - 1, num\_classes=12)  # === MLP Model ===  model = Sequential([  Dense(64, activation='relu', input\_shape=(X\_train.shape[1],)),  Dropout(0.3),  Dense(64, activation='relu'),  Dense(12, activation='softmax')  ])  model.compile(optimizer='adam', loss='categorical\_crossentropy', metrics=['accuracy'])  history = model.fit(X\_train, y\_train, epochs=50, batch\_size=8, validation\_split=0.2)  # === Loss/Accuracy Curves ===  plt.figure(figsize=(8, 4))  plt.plot(history.history["loss"], label="Train Loss")  plt.plot(history.history["val\_loss"], label="Val Loss")  plt.title("MLP Loss Curve")  plt.legend()  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "loss\_curve.png"))  plt.close()  plt.figure(figsize=(8, 4))  plt.plot(history.history["accuracy"], label="Train Acc")  plt.plot(history.history["val\_accuracy"], label="Val Acc")  plt.title("MLP Accuracy Curve")  plt.legend()  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "accuracy\_curve.png"))  plt.close()  # === Evaluate ===  y\_pred\_probs = model.predict(X\_test)  y\_pred = np.argmax(y\_pred\_probs, axis=1) + 1  y\_true = np.argmax(y\_test, axis=1) + 1  confidences = np.max(y\_pred\_probs, axis=1)  # === Classification Outputs ===  accuracy = accuracy\_score(y\_true, y\_pred)  f1 = f1\_score(y\_true, y\_pred, average='weighted')  print(f"✅ Accuracy: {accuracy:.4f}")  print(f"🎯 F1 Score: {f1:.4f}")  print("\n📋 Classification Report:")  print(classification\_report(y\_true, y\_pred))  # === Confusion Matrix ===  conf = confusion\_matrix(y\_true, y\_pred)  plt.figure(figsize=(8, 6))  sns.heatmap(conf, annot=True, fmt='d', cmap="Blues", xticklabels=range(1, 13), yticklabels=range(1, 13))  plt.title("Confusion Matrix")  plt.xlabel("Predicted")  plt.ylabel("True")  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "confusion\_matrix.png"))  plt.close()  # === Per-Cluster Accuracy ===  cluster\_acc = {}  for i in range(1, 13):  mask = y\_true == i  if mask.sum() > 0:  acc = (y\_pred[mask] == i).sum() / mask.sum()  cluster\_acc[i] = acc  plt.figure(figsize=(10, 4))  sns.barplot(x=list(cluster\_acc.keys()), y=list(cluster\_acc.values()), palette="crest")  plt.title("Per-Cluster Accuracy")  plt.ylim(0, 1)  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "cluster\_accuracy.png"))  plt.close()  # === Confidence Distribution ===  plt.figure(figsize=(6, 4))  sns.histplot(confidences, bins=10, kde=True)  plt.title("Prediction Confidence Distribution")  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "confidence\_distribution.png"))  plt.close()  # === Save Predictions with Confidence ===  pd.DataFrame({  "Oil": test\_oils,  "True Cluster": y\_true,  "Predicted Cluster": y\_pred,  "Confidence": confidences  }).to\_csv(os.path.join(output\_dir, "mlp\_pca\_predictions\_with\_confidence.csv"), index=False) |

|  |
| --- |
| **MLP Model without pca python script** |
| import pandas as pd  import numpy as np  import os  import random  import tensorflow as tf  from sklearn.preprocessing import StandardScaler  from sklearn.model\_selection import StratifiedShuffleSplit  from sklearn.decomposition import PCA  from sklearn.metrics import accuracy\_score, classification\_report, f1\_score, confusion\_matrix  from tensorflow.keras.models import Sequential  from tensorflow.keras.layers import Dense, Dropout  from tensorflow.keras.utils import to\_categorical  from imblearn.over\_sampling import SMOTE  import seaborn as sns  import matplotlib.pyplot as plt  # === Set seeds for reproducibility ===  SEED = 777  random.seed(SEED)  np.random.seed(SEED)  tf.random.set\_seed(SEED)  # === Set paths ===  base\_dir = "C:/Users/Capstone - VR Sense"  xlsx\_path = os.path.join(base\_dir, "EODataSets.xlsx")  output\_dir = os.path.join(base\_dir, "final model", "mlp without pca")  os.makedirs(output\_dir, exist\_ok=True)  # === Load Excel data ===  spectra\_df = pd.read\_excel(xlsx\_path, sheet\_name="MassSpectraDataSet")  odor\_df = pd.read\_excel(xlsx\_path, sheet\_name="OdorDataSet")  spectra\_df["Essential Oil"] = spectra\_df["Essential Oil"].str.strip()  odor\_df["Essential Oil"] = odor\_df["Essential Oil"].str.strip()  odor\_df.set\_index("Essential Oil", inplace=True)  # === Descriptor to cluster mapping ===  descriptor\_to\_cluster = {  'Smoky': 1, 'Phenolic': 1, 'Burnt': 1, 'Medicinal': 1,  'Rubber': 2, 'Sulfurous': 2, 'Acrid': 2, 'Metallic': 2,  'Sharp': 3, 'Dry': 3, 'Powdery': 3, 'Leathery': 3,  'Musky': 4, 'Animalic': 4, 'Pungent': 4, 'Oily': 4,  'Green': 5, 'Grassy': 5, 'Herbaceous': 5, 'Leafy': 5, 'Grass': 5,  'Citrus': 6, 'Lemon': 6, 'Orange': 6, 'Citronellal': 6, 'Pleasant': 6,  'Floral': 7, 'Rose': 7, 'Tea': 7, 'Light': 7,  'Spicy': 8, 'Sweet': 8, 'Fatty': 8, 'Anisic': 8, 'Balsamic': 8,  'Woody': 9, 'Root': 9, 'Earthy': 9, 'Resinous': 9, 'Pine': 9,  'Fresh': 10, 'Minty': 10, 'Cooling': 10, 'Eucalyptol': 10,  'Terpene': 11, 'Aromatic': 11, 'Peppery': 11, 'Anethole': 11,  'Warm': 12, 'Rich': 12, 'Bitter': 12, 'Fruity': 12  }  # === Assign dominant cluster ===  oil\_to\_cluster = {}  for oil, row in odor\_df.iterrows():  cluster\_counts = {}  for descriptor, present in row.items():  if present == 1 and descriptor in descriptor\_to\_cluster:  cluster = descriptor\_to\_cluster[descriptor]  cluster\_counts[cluster] = cluster\_counts.get(cluster, 0) + 1  if cluster\_counts:  dominant\_cluster = max(cluster\_counts.items(), key=lambda x: (x[1], -x[0]))[0]  oil\_to\_cluster[oil] = dominant\_cluster  # === Filter and align ===  spectra\_df = spectra\_df[spectra\_df["Essential Oil"].isin(oil\_to\_cluster.keys())]  spectra\_df["Cluster"] = spectra\_df["Essential Oil"].map(oil\_to\_cluster)  X = spectra\_df.drop(columns=["Essential Oil", "Cluster"]).values  y = spectra\_df["Cluster"].values  oil\_names = spectra\_df["Essential Oil"].values  # === PCA for 2D visualization only ===  pca = PCA(n\_components=2)  X\_pca = pca.fit\_transform(X)  plt.figure(figsize=(6,5))  sns.scatterplot(x=X\_pca[:,0], y=X\_pca[:,1], hue=y, palette="tab10")  plt.title("PCA 2D View of Oils by Cluster")  plt.savefig(os.path.join(output\_dir, "pca\_2d\_visualization.png"))  plt.close()  # === SMOTE oversampling ===  scaler = StandardScaler()  X\_scaled = scaler.fit\_transform(X)  sm = SMOTE(random\_state=SEED, k\_neighbors=1)  X\_resampled, y\_resampled = sm.fit\_resample(X\_scaled, y)  # === Real oil tracking ===  oil\_names\_resampled = np.concatenate([  oil\_names,  np.array(["(synthetic)"] \* (len(X\_resampled) - len(oil\_names)))  ])  # === Stratified train-test split ===  splitter = StratifiedShuffleSplit(n\_splits=1, test\_size=0.3, random\_state=SEED)  for train\_idx, test\_idx in splitter.split(X\_resampled, y\_resampled):  X\_train, X\_test = X\_resampled[train\_idx], X\_resampled[test\_idx]  y\_train\_raw, y\_test\_raw = y\_resampled[train\_idx], y\_resampled[test\_idx]  oil\_names\_train = oil\_names\_resampled[train\_idx]  oil\_names\_test = oil\_names\_resampled[test\_idx]  y\_train = to\_categorical(y\_train\_raw - 1, num\_classes=12)  y\_test = to\_categorical(y\_test\_raw - 1, num\_classes=12)  # === MLP Model ===  model = Sequential([  Dense(128, activation='relu', input\_shape=(X\_train.shape[1],)),  Dropout(0.3),  Dense(64, activation='relu'),  Dense(12, activation='softmax')  ])  model.compile(optimizer='adam', loss='categorical\_crossentropy', metrics=['accuracy'])  history = model.fit(X\_train, y\_train, epochs=50, batch\_size=8, validation\_split=0.2)  # === Loss/Accuracy Plots ===  plt.figure(figsize=(8,4))  plt.plot(history.history['loss'], label='Train Loss')  plt.plot(history.history['val\_loss'], label='Val Loss')  plt.title("MLP Loss Curve")  plt.legend()  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "loss\_curve.png"))  plt.close()  plt.figure(figsize=(8,4))  plt.plot(history.history['accuracy'], label='Train Accuracy')  plt.plot(history.history['val\_accuracy'], label='Val Accuracy')  plt.title("MLP Accuracy Curve")  plt.legend()  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "accuracy\_curve.png"))  plt.close()  # === Evaluation ===  y\_pred\_probs = model.predict(X\_test)  y\_pred = np.argmax(y\_pred\_probs, axis=1) + 1  y\_true = np.argmax(y\_test, axis=1) + 1  confidences = np.max(y\_pred\_probs, axis=1)  # === Confusion Matrix ===  conf = confusion\_matrix(y\_true, y\_pred)  plt.figure(figsize=(8,6))  sns.heatmap(conf, annot=True, fmt='d', cmap="Blues", xticklabels=range(1,13), yticklabels=range(1,13))  plt.title("Confusion Matrix")  plt.xlabel("Predicted")  plt.ylabel("True")  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "confusion\_matrix.png"))  plt.close()  # === Per-cluster accuracy ===  cluster\_acc = {}  for i in range(1, 13):  mask = y\_true == i  if mask.sum() > 0:  acc = (y\_pred[mask] == i).sum() / mask.sum()  cluster\_acc[i] = acc  plt.figure(figsize=(10,4))  sns.barplot(x=list(cluster\_acc.keys()), y=list(cluster\_acc.values()), palette="crest", hue=list(cluster\_acc.keys()), legend=False)  plt.title("Per-Cluster Accuracy")  plt.ylim(0, 1)  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "cluster\_accuracy.png"))  plt.close()  # === Confidence Histogram ===  plt.figure(figsize=(6,4))  sns.histplot(confidences, bins=10, kde=True)  plt.title("Prediction Confidence Distribution")  plt.tight\_layout()  plt.savefig(os.path.join(output\_dir, "confidence\_distribution.png"))  plt.close()  # === Print Evaluation Metrics ===  print(f"✅ Accuracy: {accuracy\_score(y\_true, y\_pred):.4f}")  print(f"🎯 F1 Score: {f1\_score(y\_true, y\_pred, average='weighted'):.4f}")  print("\n📋 Classification Report:")  print(classification\_report(y\_true, y\_pred))  # === Predict on original oils (before SMOTE) ===  X\_real\_scaled = scaler.transform(X)  real\_probs = model.predict(X\_real\_scaled)  real\_preds = np.argmax(real\_probs, axis=1) + 1  real\_confidences = np.max(real\_probs, axis=1)  # === Export Final Oil Predictions ===  pd.DataFrame({  "Oil": oil\_names,  "True Cluster": y,  "Predicted Cluster": real\_preds,  "Confidence": real\_confidences  }).to\_csv(os.path.join(output\_dir, "mlp\_predictions\_with\_confidence.csv"), index=False) |
| **WORD2VEC CLUSTER FORMATION python Script** |
| import pandas as pd  from sklearn.cluster import KMeans  from gensim.models import Word2Vec  import numpy as np  # === Step 1: Full list of 52 descriptors ===  descriptors = [  'Smoky', 'Phenolic', 'Burnt', 'Medicinal',  'Rubber', 'Sulfurous', 'Acrid', 'Metallic',  'Sharp', 'Dry', 'Powdery', 'Leathery',  'Musky', 'Animalic', 'Pungent', 'Oily',  'Green', 'Grassy', 'Herbaceous', 'Leafy', 'Grass',  'Citrus', 'Lemon', 'Orange', 'Citronellal', 'Pleasant',  'Floral', 'Rose', 'Tea', 'Light',  'Spicy', 'Sweet', 'Fatty', 'Anisic', 'Balsamic',  'Woody', 'Root', 'Earthy', 'Resinous', 'Pine',  'Fresh', 'Minty', 'Cooling', 'Eucalyptol',  'Terpene', 'Aromatic', 'Peppery', 'Anethole',  'Warm', 'Rich', 'Bitter', 'Fruity'  ]  # === Step 2: Generate simple context sentences for training ===  sentences = [[desc1, desc2] for i, desc1 in enumerate(descriptors) for j, desc2 in enumerate(descriptors) if i != j]  # === Step 3: Train Word2Vec model ===  model = Word2Vec(sentences, vector\_size=50, window=3, min\_count=1, workers=1, seed=42)  # === Step 4: Vectorize descriptors ===  vectors = np.array([model.wv[word] for word in descriptors])  # === Step 5: Apply KMeans clustering ===  kmeans = KMeans(n\_clusters=12, random\_state=777, n\_init=10)  labels = kmeans.fit\_predict(vectors)  # === Step 6: Build output cluster dictionary ===  cluster\_dict = {}  for i, label in enumerate(labels):  cluster\_dict.setdefault(f"Cluster {label + 1}", []).append(descriptors[i])  # === Step 7: Format for display ===  cluster\_df = pd.DataFrame({  "Cluster": cluster\_dict.keys(),  "Descriptors": ['; '.join(descs) for descs in cluster\_dict.values()]  })  # === Save the clustered descriptors to CSV ===  output\_df = pd.DataFrame({  "Cluster": cluster\_dict.keys(),  "Descriptors": ['; '.join(descs) for descs in cluster\_dict.values()]  })  output\_df.to\_csv("word2vec\_descriptor\_clusters.csv", index=False)  print("✅ Clusters saved to 'word2vec\_descriptor\_clusters.csv'")  print("\n📦 Word2Vec + KMeans Clusters:\n")  for cluster\_label, descriptors in cluster\_dict.items():  print(f"{cluster\_label}:")  print(" " + ", ".join(descriptors))  print("-" \* 40) |

|  |
| --- |
| **INITIAL DATA ANALYSIS AND verification R script** |
| #Scent Data Sets from previous Mass Spectra data and Essential Oil odor comparrisons #Author: Dawson McClary, Candice Reyes, Chris McReynolds #Last Revision: 6/26/2025  #library for functions, themes, and plots library(hdf5r) library(R.matlab) library(tidyverse) library(dplyr) library(datasets) library(ggbiplot) library(ggplot2) library(conflicted) library(readr) library(readxl) library(reshape2) library(knitr) library(AppliedPredictiveModeling) library(caret) library(tidyverse) library(mlbench) library(glmnet) library(earth) library(rpart) library(party) library(partykit) library(ROCR) library(rattle) library(rpart.plot) library(RColorBrewer) library(adabag) library(ipred) library(randomForest) library(patchwork) library(tidyr) library(ggrepel) library(dendextend) library(pheatmap) library(UpSetR) library(grid) library(CCA) library(matrixStats)  #View the Matlab files in R EOwsd <- readMat('C:/Users/alexm/Documents/Capstone - VR Sense/EOwsd.mat') EOwsd\_MS <- readMat('C:/Users/alexm/Documents/Capstone - VR Sense/EOwsd\_MS.mat') EOwsd\_Names <- readMat('C:/Users/alexm/Documents/Capstone - VR Sense/EOwsd\_Names.mat') EOwsd\_OdorsList <- readMat('C:/Users/alexm/Documents/Capstone - VR Sense/EOwsd\_OdorsList.mat') EOwsd\_SDOHen <- readMat('C:/Users/alexm/Documents/Capstone - VR Sense/EOwsd\_SDOHen.mat')  #Converted the tables from R into an excel sheet with tables easily formatted for importing  #Reads excel workbook and assigns sheets with feature data to a variable OdorDataSet <- read\_excel("C:/Users/alexm/Documents/Capstone - VR Sense/EODataSets.xlsx", sheet = "OdorDataSet") MassSpecDataSet <- read\_excel("C:/Users/alexm/Documents/Capstone - VR Sense/EODataSets.xlsx", sheet = "MassSpectraDataSet")  #view(OdorDataSet) #view(MassSpecDataSet)  OdorDataSet\_df <- data.frame(OdorDataSet) MassSpecDataSet\_df <- data.frame(MassSpecDataSet)  #Line plot of mass spectra  Assuming MassSpecDataSet\_df is your data.frame (94 oils × 50 bins)  ###################### Melt your data to long format ################################################## MSdf\_long <- melt(MassSpecDataSet\_df, id.vars = "Essential.Oil", variable.name = "Bin", value.name = "Intensity") ###################### Replace all non-breaking spaces (U+00A0) with a normal space MSdf\_long$Essential.Oil <- gsub("\u00A0", " ", MSdf\_long$Essential.Oil) MSdf\_long$Essential.Oil <- gsub("[[:space:][:cntrl:]]+", " ", MSdf\_long$Essential.Oil) MSdf\_long$Essential.Oil <- trimws(MSdf\_long$Essential.Oil)  ###################### Convert factor to character first, then extract numeric start ################### MSdf\_long$BinStart <- as.numeric(sub(".?(\d+).", "\1", as.character(MSdf\_long$Bin)))  Choose a few oils you know exist in your dataset  selected\_oils <- c("Pepper (Black)", 'Patchouli', "Cedarwood (Virginian)")  ###################### Line plot of select oils based on "selected\_oils" variable ###################### MSplot\_LineSelect <- ggplot(subset(MSdf\_long, Essential.Oil %in% selected\_oils), aes(x = BinStart, y = Intensity, color = Essential.Oil)) + geom\_line(size = 1) + labs(title = "Mass Spectra of Selected Essential Oils", x = "m/z Bin Start", y = "Intensity") + theme\_minimal()  print(MSplot\_LineSelect)  ###################### Line plot of all oils ############################################################ MSplot\_LineAll <- ggplot(MSdf\_long, aes(x = BinStart, y = Intensity, color = Essential.Oil)) + geom\_line(alpha = 0.7, size = 0.8) + labs(title = "Mass Spectra of All Essential Oils", x = "m/z Bin Start", y = "Intensity") + theme\_minimal() + theme(strip.text = element\_text(size = 6), axis.text = element\_text(size = 5), axis.title = element\_text(size = 8))  print(MSplot\_LineAll)  ###################### Faceted line plots of all oils #################################################### MSplot\_FacetAll <- ggplot(MSdf\_long, aes(x = BinStart, y = Intensity)) + geom\_line(color = "steelblue", size = 0.6) + facet\_wrap(~ Essential.Oil, scales = "free\_y", ncol = 6) + labs(title = "Mass Spectra of All Essential Oils", x = "m/z Bin Start", y = "Intensity") + theme\_minimal(base\_size = 6) + theme(strip.text = element\_text(size = 6))  print(MSplot\_FacetAll)  ###################### Make heatmap-style data ############################################################ MSheatmap\_df <- MSdf\_long %>% mutate(Bin = factor(Bin, levels = unique(Bin))) %>% select(Essential.Oil, Bin, Intensity)  MSplot\_HeatMap <- ggplot(MSheatmap\_df, aes(x = Bin, y = Essential.Oil, fill = Intensity)) + geom\_tile() + scale\_fill\_viridis\_c() + labs(title = "Mass Spectra Heatmap of Essential Oils", x = "m/z Bin", y = "Essential Oil") + theme\_minimal(base\_size = 7) + theme(axis.text.x = element\_text(angle = 90, hjust = 1))  print(MSplot\_HeatMap)  ###################### PCA Prep ############################################################################  Extract numeric data only for PCA  spectra\_matrix <- MassSpecDataSet\_df %>% select(-Essential.Oil) %>% mutate(across(everything(), as.numeric)) %>% as.matrix()  Run PCA  MSpca\_result <- prcomp(spectra\_matrix, scale. = TRUE)  Build PCA plot  MSpca\_df <- as.data.frame(pca\_result$x) MSpca\_df$Essential.Oil <- MSdf\_wide$Essential.Oil  MSplot\_PCA <- ggplot(MSpca\_df, aes(x = PC1, y = PC2, label = Essential.Oil)) + geom\_point(aes(color = Essential.Oil), size = 2) + geom\_text(size = 2.5, vjust = 1, hjust = 1, check\_overlap = TRUE) + labs(title = "PCA of Essential Oil Mass Spectra") + theme\_minimal() + theme(legend.position = "none")  print(MSplot\_PCA)  PCA results and look at proportion of variance captured per PC  summary(MSpca\_result) var\_exp <- (MSpca\_result$sdev)^2 / sum((MSpca\_result$sdev)^2) plot(var\_exp, type = "b", xlab = "Principal Component", ylab = "Proportion of Variance Explained") |

|  |
| --- |
| **PROJECT MANAGEMENT VISUALs python Script** |
| import numpy as np import matplotlib.pyplot as plt import pandas as pd import seaborn as sns  === Risk Table Scores ===  likelihoods = [3,3,1,1,5,3,3,3,3,1,3,1,1,5,1,3,3,1,3,3,1,3,3,1,1,3,3,3,3,3] impacts = [5,3,5,3,3,5,3,3,5,3,3,5,3,3,5,3,3,3,3,1,5,5,3,5,5,3,3,3,5,3]  === Simulation Settings ===  n\_simulations = 10000 simulated\_totals = []  Set the seed for reproducibility  np.random.seed(777)  === Monte Carlo Simulation ===  for \_ in range(n\_simulations): total\_severity = 0 for l\_base, i\_base in zip(likelihoods, impacts): # Triangular sampling with clamp between 1 and 5 l\_sim = int(np.round(np.clip(np.random.triangular(l\_base-1, l\_base, l\_base+1), 1, 5))) i\_sim = int(np.round(np.clip(np.random.triangular(i\_base-1, i\_base, i\_base+1), 1, 5))) total\_severity += l\_sim \* i\_sim simulated\_totals.append(total\_severity)  simulated\_totals = np.array(simulated\_totals)  === Percentile Statistics ===  mean\_val = np.mean(simulated\_totals) p25 = np.percentile(simulated\_totals, 25) p50 = np.percentile(simulated\_totals, 50) p75 = np.percentile(simulated\_totals, 75) p90 = np.percentile(simulated\_totals, 90)  === Plotting Histogram with Percentile Lines ===  plt.figure(figsize=(10, 6)) plt.hist(simulated\_totals, bins=50, color="skyblue", edgecolor="black", alpha=0.7) plt.axvline(p25, color='purple', linestyle='--', label=f'P25 = {p25:.2f}') plt.axvline(p50, color='red', linestyle='--', label=f'P50 (Median) = {p50:.2f}') plt.axvline(p75, color='green', linestyle='--', label=f'P75 = {p75:.2f}') plt.axvline(p90, color='orange', linestyle='--', label=f'P90 = {p90:.2f}') plt.title("Monte Carlo Simulation: Total Project Risk Exposure", fontsize=14) plt.xlabel("Total Simulated Severity Score", fontsize=12) plt.ylabel("Frequency", fontsize=12) plt.legend() plt.grid(True) plt.tight\_layout() plt.show()  === Print Numeric Results ===  print(f"Mean Total Risk Exposure: {mean\_val:.2f}") print(f"P25: {p25:.2f}") print(f"P50 (Median): {p50:.2f}") print(f"P75: {p75:.2f}") print(f"P90: {p90:.2f}")  Risk Heatmap  === Create Frequency Matrix (Impact = columns, Likelihood = rows) ===  heatmap\_data = pd.DataFrame(0, index=range(1, 6), columns=range(1, 6)) for l, i in zip(likelihoods, impacts): heatmap\_data.loc[l, i] += 1 # Now Likelihood is the row, Impact is the column  === Filter for only 1, 3, 5 scores ===  filtered\_data = heatmap\_data.loc[[1, 3, 5], [1, 3, 5]]  === Plot the Heatmap ===  plt.figure(figsize=(6, 5)) sns.heatmap(filtered\_data, annot=True, fmt="d", cmap="Reds", cbar=True)  === Labels and Formatting ===  plt.title("Risk Heatmap of Impact and Likelihood Scores", fontsize=14) plt.xlabel("Impact", fontsize=12) plt.ylabel("Likelihood", fontsize=12) plt.xticks(ticks=[0.5, 1.5, 2.5], labels=[1, 3, 5]) plt.yticks(ticks=[0.5, 1.5, 2.5], labels=[1, 3, 5], rotation=0) plt.tight\_layout() plt.show()  Stakeholder Power-Interest Matrix  === Stakeholder Data ===  stakeholders = [ "Clinical Psychologists", "PTSD Patients", "VR Developers", "Systems Engineers", "Clinical Device Engineers", "Capstone Review Board", "Successor Team", "UX Designers", "Ethics Advisors" ] power = [5, 2, 4, 5, 5, 4, 3, 3, 4] interest = [5, 5, 4, 3, 4, 3, 3, 4, 2]  === Color Coding by Engagement Strategy ===  colors = [] for p, i in zip(power, interest): if p >= 4 and i >= 4: colors.append("#2ca02c") # green: Engage Closely elif p >= 4 and i < 4: colors.append("#ff7f0e") # orange: Keep Satisfied elif p < 4 and i >= 4: colors.append("#1f77b4") # blue: Keep Informed else: colors.append("#7f7f7f") # gray: Monitor (none currently in this case)  === Plotting the Matrix ===  plt.figure(figsize=(10, 8)) plt.scatter(power, interest, c=colors, s=300, edgecolors='black')  === Add Labels ===  for i, label in enumerate(stakeholders): plt.text(power[i], interest[i] + 0.1, label, ha='center', va='bottom')  === Add Quadrant Guidelines ===  plt.axhline(y=3, color='black', linestyle='--', linewidth=1) plt.axvline(x=3, color='black', linestyle='--', linewidth=1)  === Axes and Styling ===  plt.title("Stakeholder Power–Interest Matrix", fontsize=16, fontweight='bold') plt.xlabel("Power (1 = Low, 5 = High)", fontsize=12) plt.ylabel("Interest (1 = Low, 5 = High)", fontsize=12) plt.xticks(range(1, 6)) plt.yticks(range(1, 6)) plt.grid(True, linestyle=':', linewidth=0.7) plt.xlim(0.8, 5.2) plt.ylim(0.8, 5.2) plt.tight\_layout() plt.show() |

APPENDIX B: SCENT NOTE PROFILES

A screen shot of a computer

Description automatically generated

APPENDIX C: SYSML DESIGN – VORA REQUIREMENTS – VORA BOM – VORA SYSML MARKUP (click each PDF file to view document)

****

****

APPENDIX D: DESCRIPTOR VECTORIZATION JUSTIFICATIONS

Appendix D 1: Lavandin (Super)

Likely Extraction Method: Steam Distillation  
 Overview: Lavandin (Super) is a hybrid of Lavandula angustifolia and Lavandula latifolia. It has a sharper, more camphoraceous aroma than true lavender, with slightly green and herbaceous notes, and is commonly used in perfumery and functional fragrance applications.

##### Descriptors and Justifications:

Fresh: Lavandin carries a light and cleansing top note often associated with “fresh” scents in soaps and household products. This brightness is attributed to its linalool and linalyl acetate content.

Cooling: Due to the presence of camphor and 1,8-cineole (eucalyptol), Lavandin emits a brisk, menthol-like aroma that evokes a “cooling” sensation upon inhalation.

Herbaceous: This descriptor is fitting as Lavandin retains the leafy, green character from its parent L. angustifolia, presenting a botanical freshness tied to terpene derivatives.

Floral: Despite its sharper tone, Lavandin still exhibits a floral core typical of its lavender lineage. The linalool-linalyl acetate balance contributes to this aspect.

Pungent: Camphor and cineole impart a nasal-tingling intensity that can be described as pungent, more aggressive than soft floral oils.

Aromatic: A broad term, “aromatic” is applicable to Lavandin for its volatile compound complexity, encompassing floral, green, and slightly medicinal tonalities.

Sharp: The presence of terpenes and camphor results in a stinging or piercing top note, aligning with the "sharp" descriptor.

Appendix D 2: Peppermint (English)

Likely Extraction Method: Steam Distillation  
 Overview: Peppermint oil is obtained from Mentha × piperita, known for its menthol-rich profile. It is widely used in both flavoring and aromatherapy.

##### Descriptors and Justifications:

Fresh: The oil’s crispness is largely due to menthol and menthone, which impart a refreshing, uplifting sensory response.

Minty: A primary scent note of peppermint, “minty” is chemically driven by menthol, which is cooling and piercing.

Cooling: Menthol stimulates thermoreceptors in the nasal passage, resulting in a perceived decrease in temperature.

Herbaceous: As a member of the mint family, the oil retains a green, leafy character beneath its menthol bite.

Pungent: High menthol content causes a sharp, burning sensation that fits the definition of “pungent.”

Aromatic: Volatile menthyl esters and monoterpenes give it a strong olfactory footprint, earning the aromatic label.

Sharp: The bite of menthone and menthyl acetate gives a crisp, nearly metallic edge to the scent.

Eucalyptol: While present in smaller amounts than in eucalyptus oil, some English peppermint varieties do contain trace amounts of 1,8-cineole, contributing to its brightness.

Appendix D 3: Petitgrain (Citrus aurantium)

Likely Extraction Method: Steam Distillation  
 Overview: Petitgrain is derived from the leaves and twigs of the bitter orange tree, rather than the fruit, and exhibits floral-citrus-woody nuances.

##### Descriptors and Justifications:

Woody: Petitgrain carries a dry woodiness from the linalool-terpene backbone, reminiscent of twigs and stems.

Green: This arises from its leafy origin, producing vegetal and chlorophyll-like overtones.

Bitter: Linalyl acetate and other terpenoids create a bitter, slightly medicinal nuance.

Fresh: Petitgrain’s brisk profile is commonly used in colognes due to its sparkling brightness.

Herbaceous: The steam-distilled leaf matter yields a lightly savory, green-floral undertone.

Floral: Though less than neroli, petitgrain has floral softness from its linalool component.

Grassy: This emerges from the oil’s slight sharp-green profile, more plant-like than herbaceous.

Tea: The balance of soft floralcy and green aromatic depth evokes a faint, delicate tea-leaf note.

Aromatic: A general descriptor covering its diverse bouquet.

Pleasant: Petitgrain’s use in high-end colognes and soaps shows its wide acceptability and low olfactory fatigue.

Appendix D 4: Juniperberry

Likely Extraction Method: Steam Distillation  
 Overview: Distilled from the ripe berries of Juniperus communis, the oil is bright, piney, and slightly sweet.

##### Descriptors and Justifications:

Woody: Juniper is known for its dry, pinewood base due to high concentrations of α-pinene and β-pinene.

Fresh: The crisp, forest-like character evokes freshness, aided by monoterpene volatility.

Pine: This association is nearly direct, as juniper shares many of the same terpenes with pine species.

Resinous: Underlying balsamic notes from terpinene-4-ol give it a viscous, sappy edge.

Smoky: When oxidized or aged, juniper oil develops a faintly smoky nuance, similar to dried wood embers.

Sharp: Its brisk top note lends a clean-cut astringency to its overall profile.

Aromatic: A highly complex scent from multiple terpene groups.

Appendix D 5: Cedarwood (Virginian)

Likely Extraction Method: Steam Distillation  
 Overview: Extracted from Juniperus virginiana, this is technically a juniper species and features a dry, pencil-shavings-like profile.

##### Descriptors and Justifications:

Woody: Dominated by cedrol and thujopsene, the scent is strongly wood-based, reminiscent of cedar chips.

Floral: Though minimal, some samples contain methyl octine carbonate-like soft floral undertones.

Earthy: A grounding scent that mimics the aroma of dry soil or bark.

Resinous: A slightly sticky, balsamic character emerges from sesquiterpenes.

Pine: Similar to juniper, this wood contains terpene analogs of pinewood species.

Smoky: The dry distillation can result in light smoky notes resembling charred cedar.

Phenolic: Though milder than tar oils, aged samples may develop phenolic depth akin to smoke or leather.

Sharp: A faint pointedness appears in the top note phase.

Aromatic: The volatile blend projects well and retains its signature in compositions.

Appendix D 6: Petitgrain (Citrus aurantium)

**Overview:**  
 **Petitgrain essential oil is derived from the leaves and green twigs of the bitter orange tree (Citrus aurantium), distinct from neroli (flowers) and bitter orange (peel). It has a fresh, green, and mildly woody aroma with a lightly floral undertone. Commonly used in perfumery and aromatherapy for its calming effects.**

**Inferred Extraction Method:**  
 **Steam Distillation – Petitgrain is traditionally steam distilled from the leaves and twigs, as opposed to the cold pressing method used for citrus peels.**

**Descriptors and Justifications:**

Woody: Petitgrain contains linalool and linalyl acetate, both of which contribute to the subtle woody undertones common in green twig-based oils .

Green: Its profile includes high levels of green aldehydes and monoterpenes like alpha-terpineol and myrcene, responsible for the sharp, green scent .

Bitter: Petitgrain’s slightly bitter edge is attributed to terpenoid compounds and its relation to the bitter orange tree, which imparts a citrus bitterness in the top note.

Fresh: The combination of linalool, terpineol, and citronellol gives Petitgrain a clean, brisk sensation typical of fresh notes .

Herbaceous: Steam-distilled leaves yield an herbaceous nuance due to alcohols like linalool, which are aromatic and slightly sweet-herbal.

Tea: While subtle, its light floral and green blend is often used as a tea-like note in perfumery, reminiscent of Earl Grey-like profiles.

Floral: Linalyl acetate and nerolidol impart delicate floral characteristics similar to neroli but milder.

Aromatic: Monoterpenes and oxygenated compounds such as nerol and geraniol give Petitgrain a mildly aromatic, cologne-like base.

Appendix D 7: Juniperberry

**Overview:**  
 **Juniperberry essential oil is extracted from the berries of Juniperus communis. It has a sharp, clean, and pine-like aroma with woody and peppery undertones, often used in cleansing blends and detox treatments.**

**Inferred Extraction Method:**  
 **Steam Distillation – Juniperberries are typically dried and then steam distilled to produce a crisp, resinous essential oil.**

**Descriptors and Justifications:**

Woody: Its high content of alpha-pinene and other terpenes imparts a distinctly woody and balsamic aroma, similar to other conifer-derived oils .

Fresh: Alpha-pinene and myrcene produce a bright, invigorating freshness, contributing to Juniperberry’s role in uplifting aromatic blends.

Peppery: The oil contains sabinene and beta-caryophyllene, which are associated with warm, slightly spicy peppercorn-like notes.

Aromatic: With a composition rich in monoterpenes (alpha- and beta-pinene), it has a sharply aromatic scent widely used in colognes.

Resinous: The sticky, balsamic character—especially when aged—links Juniperberry to a slightly resinous quality, akin to frankincense and pine.

Pine: Shares major chemical compounds like alpha-pinene with pine oils, explaining the olfactory similarity.

Sharp: Its intense monoterpene structure gives rise to a stinging, clean sharpness at first inhalation.

Appendix D 8: Cedarwood (Virginian)

**Overview:**  
 **Cedarwood Virginian (Juniperus virginiana) has a dry, woody, and smoky scent profile. Unlike true cedars (genus Cedrus), it’s technically a juniper, though still widely referred to as cedarwood in aromatherapy.**

**Inferred Extraction Method:**  
 **Steam Distillation – The wood chips are steam distilled to yield a thick, deep-hued oil.**

**Descriptors and Justifications:**

Woody: Rich in sesquiterpenes like cedrene, Cedarwood Virginian is one of the benchmark woody oils .

Earthy: Terpenic derivatives lend the oil a base note reminiscent of damp soil and dry bark.

Resinous: Contains thujopsene and alpha-cedrene, both of which contribute to its balsamic, resin-like base.

Pine: Shared origin from the Cupressaceae family and alpha- and beta-pinene content reinforce its pine-needle characteristic.

Smoky: Often described as having a smoky nuance due to the distillation process and terpenic degradation.

Phenolic: The sharpness in the base note may also be linked to mild phenolic tones similar to creosote wood.

Sharp: Initial vapor has a pointed, stinging note due to the interaction between terpenes and oxygenated hydrocarbons.

Appendix D 9: Lemon

**Overview:**  
 **Lemon oil is known for its sharp, bright, and uplifting scent and is derived from the peel of the lemon fruit (Citrus limonum).**

**Inferred Extraction Method:**  
 **Cold Pressed – Citrus peels are traditionally cold pressed, which retains more of the volatile top-note compounds such as limonene.**

**Descriptors and Justifications:**

Fresh: Dominated by limonene (up to 70%), which is responsible for the fresh, zesty aroma of citrus oils.

Lemon: The source of this oil directly aligns with this descriptor—lemons.

Citrus: Includes a wide range of citrus aldehydes and esters, including citral and citronellal.

Light: As a top-note oil with a low molecular weight, it evaporates quickly and imparts an airy brightness.

Orange: While lemon-forward, there is overlap in scent chemistry with orange (d-limonene), producing a rounded citrus note.

Pleasant: The universally pleasant, clean aroma is widely used in personal care and cleaning products.

Dry: The astringent, slightly acidic nature of lemon peel distillate evokes a dry, effervescent mouthfeel in perfumery.

Appendix D 10: Citronella

**Overview:**  
 **Citronella oil is extracted from the stems and leaves of Cymbopogon nardus or Cymbopogon winterianus. It has a sharp, lemony scent with grassy and slightly woody undertones.**

**Inferred Extraction Method:**  
 **Steam Distillation – As a grass-derived oil, it is always steam distilled from leaves and stems.**

**Descriptors and Justifications:**

Fresh: Citronellol and geraniol contribute to its invigorating freshness.

Herbaceous: The leafy source material imparts a dry, grassy herbaceous quality.

Lemon: Named for its lemon-like scent, due to high citronellal content.

Citrus: Though not a true citrus, the aldehydes and monoterpenoids simulate citrus characteristics.

Citronellal: Dominated by citronellal, which gives it insect-repellent power and a strong lemony odor.

Pungent: The intensity and volatility of citronella oil create a sharp, slightly overwhelming note in high concentrations.

Dry: Some versions exhibit a dry, grassy fade-out due to geranyl acetate content.

Appendix D 11: Pepper (Black)

**Inferred Extraction Method: Steam Distillation**  
 **Overview: Black pepper essential oil is derived from the dried unripe fruits (peppercorns) of Piper nigrum and is rich in monoterpenes such as α-pinene and limonene, along with sesquiterpenes like caryophyllene. It is known for its sharp, dry, and spicy aroma.**

**Descriptors and Justifications:**

Spicy: The presence of β-caryophyllene and limonene imparts a classic spicy character, evoking the pungency associated with freshly ground black pepper.

Woody: Terpenes like α-humulene contribute warm, woody undertones often observed in steam-distilled spice oils.

Pungent: Piperine, while not volatile itself, is accompanied by terpenoid volatiles that give the oil its aggressive, sinus-clearing sharpness.

Peppery: This descriptor directly aligns with the oil’s botanical origin and profile—sharp, nose-tingling, and resonant of the black peppercorn spice.

Aromatic: Monoterpenes such as limonene and α-pinene lend a radiant aromatic quality to the scent.

Metallic: Though not common in most spice oils, the high sharpness and nasal penetration can translate into a metallic sensation during olfactory processing.

Leathery: Some aged samples or oxidized components may develop warm, dense leathery notes, resembling well-worn leather.

Animalic: In certain cases, black pepper oil has undertones reminiscent of body warmth or musk, often due to β-caryophyllene oxidation.

Sharp: A defining characteristic; the oil’s immediate nasal impact is due to terpenes stimulating trigeminal nerves.

Appendix D 12: Vetiver

**Inferred Extraction Method: Steam Distillation from dried roots**  
 **Overview: Extracted from the roots of Vetiveria zizanoides, vetiver oil is complex, deep, and earthy. It contains khusimol, vetiselinenol, and isovalencenol among other sesquiterpenoids.**

**Descriptors and Justifications:**

Woody: The grounding scent of vetiver is woody by nature, derived from sesquiterpene alcohols in the roots.

Root: As a root oil, this descriptor is both literal and olfactive, reflecting its earthy, dry, and grounding base notes.

Earthy: Vetiver is one of the most archetypal earthy essential oils, prized in perfumery for its soil-like and mineral depth.

Floral: Though not floral in the conventional sense, vetiver may contain faintly sweet notes in the drydown that evoke the backdrop of dried petals.

Resinous: Some vetiver samples develop a sweet-resinous or balsamic quality due to oxidation and root aging during drying.

Pine: While pine isn't dominant, vetiver can exhibit a coniferous sharpness due to its terpenoid mix, especially in aged material.

Smoky: A key characteristic; smoky tones in vetiver stem from deep-rooted sesquiterpenes and their degradation products.

Rubber: Certain distillation conditions or root aging can yield rubbery, burnt-tire-like notes, commonly found in darker vetiver distillations.

Phenolic: Similar to smoky, the heavy phenolic tones evoke charred wood or burnt soil.

Medicinal: While not overtly camphoraceous, the bitter, resinous quality of some vetiver batches can resemble old-world medicinal roots.

Dry: A classic base note; dry, powdery-earth hints pervade vetiver as it settles on skin.

Appendix D 13: Thyme (Sweet)

**Inferred Extraction Method: Steam Distillation**  
 **Overview: Sweet thyme oil (Thymus vulgaris chemotype linalool) is more gentle and floral than thymol-type thyme. It contains linalool, terpineol, and small amounts of thymol.**

**Descriptors and Justifications:**

Spicy: The warming effect of thymol and terpineol lends a peppery, kitchen-spice warmth.

Herbaceous: Thyme’s core aromatic nature is green and herbal, derived from its close family of culinary herbs.

Pungent: While “sweet” thyme is milder, it still retains the phenolic punch found in all thyme varieties.

Aromatic: Linalool and terpineol give it bright, floral-aromatic highs.

Sharp: Present even in sweet variants; high monoterpene content gives thyme a penetrating bite.

Appendix D 14: Carrot Seed

**Inferred Extraction Method: Steam Distillation of dried seeds**  
 **Overview: Carrot seed oil is distilled from Daucus carota seeds and contains carotol, daucol, and beta-caryophyllene. It has a dry, earthy, slightly musky scent.**

**Descriptors and Justifications:**

Woody: Though not wood-derived, the dry, fibrous odor of carrot seed resembles cedar shavings or pencil shavings.

Warm: The oil’s soft, musky undertones feel enveloping and warm, especially in blends.

Root: The scent bears an unmistakable “earthy root” character, like parsnip or dried tuber.

Herbaceous: Reflects the plant’s family—Apiaceae—and the green, dry tone of the seeds.

Floral: In the drydown, a faint musky-floral note may emerge, contributing to its use in vintage perfumes.

Dry: The top note is typically dry and dusty, like aged herbs or fields after harvest.

Appendix D 15: Myrtle

**Inferred Extraction Method: Steam Distillation of leaves**  
 **Overview: Myrtle (Myrtus communis) essential oil has a clean, slightly medicinal profile with a floral-green overtone. Major constituents include myrtenyl acetate, 1,8-cineole (eucalyptol), and linalool.**

**Descriptors and Justifications:**

Sweet: Myrtenyl acetate gives myrtle its delicate, slightly sweet quality, used in perfumery for its softness.

Fresh: The oil opens with a fresh, uplifting green note from eucalyptol and linalool.

Herbaceous: Derived from green leaves, its base character is herbaceous, like a milder rosemary.

Tea: Myrtle belongs to the myrtle family (Myrtaceae) and shares certain tea-like, astringent nuances.

Floral: In some chemotypes, the ester content lends soft floral notes akin to lavender or rosewood.

Light: The overall impression is airy, soft, and non-cloying—making it ideal for gentle diffusion.

Appendix D 16: Cedarwood (Himalayan)

**Inferred Extraction Method: Steam Distillation of wood**  
 **Overview: Derived from the wood of Cedrus deodara, this oil is dry, woody, and subtly smoky with grounding undertones. It is high in sesquiterpenes such as himachalol, atlantone, and cedrol.**

**Descriptors and Justifications:**

Woody: As a wood-derived oil, this is its primary olfactory feature—dry, deep, and pencil-shaving-like.

Floral: Surprisingly, cedarwood Himalayan contains faint notes that can be interpreted as floral, especially in the drydown, possibly due to balance with cedrol.

Leafy: A soft, green nuance is present due to unextracted leaf matter during wood steam distillation, adding to the forest-floor character.

Resinous: Contains balsamic notes reminiscent of tree sap or pine resin, common to cedar oils.

Pine: The high sesquiterpene content gives it a coniferous edge, making it similar in feel to pine needle oils.

Dry: The astringency and arid, incense-like trail of this oil makes "dry" an appropriate descriptor.

Appendix D 17: Fennel (Sweet)

**Inferred Extraction Method: Steam Distillation of crushed seeds**  
 **Overview: Obtained from Foeniculum vulgare, sweet fennel oil is rich in trans-anethole, giving it a candy-like, licorice scent.**

**Descriptors and Justifications:**

Spicy: While not hot-spicy, the warm, aromatic spice character is typical of seed-based oils.

Sweet: Trans-anethole gives this oil a markedly sweet, sugary profile.

Anisic: Almost synonymous with the term "aniseed"—fennel is a major natural source of this compound.

Floral: Some users detect a light floral veil behind the dominating sweet-spice note, akin to neroli or lavender in trace levels.

Peppery: Though mild, a fresh peppery edge exists in early top notes.

Anethole: This descriptor is directly tied to its main compound—anethole—found in high concentrations.

Appendix D 18: Tea Tree (Melaleuca alternifolia)

**Inferred Extraction Method: Steam Distillation of leaves**  
 **Overview: Tea tree oil is best known for its intense medicinal, camphoraceous scent. It contains terpinen-4-ol, γ-terpinene, and α-terpineol.**

**Descriptors and Justifications:**

Fresh: The top notes are clean and brisk, often compared to antiseptic wipes or eucalyptus.

Minty: Not mentholic, but the camphor-like bite gives a fresh-cooling impression.

Cooling: This sensation stems from its interaction with trigeminal receptors, even without menthol.

Pungent: Very strong and forceful; the sharpness is unmistakable.

Terpene: High in monoterpenes, this descriptor is chemically and olfactively justified.

Sharp: The oil is often described as "piercing" or "abrasive" in scent—hence the term.

Appendix D 19: Clove Bud

**Inferred Extraction Method: Steam Distillation of dried flower buds**  
 **Overview: Rich in eugenol, clove bud oil is powerfully spicy, warm, and phenolic. Often used in dental care.**

**Descriptors and Justifications:**

Spicy: This is its primary and most intense note—hot, dry spice with immediate warmth.

Sweet: Eugenol has a sweet, almost vanilla-adjacent drydown.

Warm: The warming sensation is both chemical (warming rubs) and emotional (comforting holiday spice).

Balsamic: Some batches have syrupy, thick undercurrents, especially in aged oils.

Phenolic: The smoky, tar-like edge of eugenol gives clove a phenolic, almost tarry dimension.

Medicinal: Used traditionally for toothaches and antiseptic applications, the scent is tightly associated with medicine.

Sharp: Especially strong in undiluted form, the aroma cuts through and lingers.

Appendix D 20: Ylang Ylang (Extra)

**Inferred Extraction Method: Steam Distillation (first fraction) of fresh flowers**  
 **Overview: The "Extra" grade is the most fragrant fraction of ylang ylang distillation. It’s intensely floral, rich, and exotic, high in benzyl acetate, linalool, and p-cresyl methyl ether.**

**Descriptors and Justifications:**

Sweet: Intensely sweet, often compared to banana or candy.

Bitter: The slight metallic astringency in the heart note can evoke bitterness.

Floral: One of the most floral essential oils—a cornerstone of many perfumes.

Pleasant: Subjective, but universally regarded as smooth, calming, and mood-lifting.

Fruity: Hints of ripe banana or melon, especially in the top note due to esters.

Balsamic: Soft, creamy undertones ground the top florals, adding weight and warmth.

Appendix D 21: Spearmint (Mentha spicata)

**Inferred Extraction Method: Steam Distillation of fresh leaves**  
 **Overview: Spearmint oil is rich in carvone, a ketone responsible for its sweet minty flavor and cooling quality.**

**Descriptors and Justifications:**

Sweet: Carvone gives it a sweet, refreshing candy-like edge.

Green: Bright and leafy, spearmint evokes cut grass and herbs.

Minty: Self-explanatory; spearmint defines this category.

Herbaceous: Comes from the herbaceous plant family Lamiaceae.

Eucalyptol: Small concentrations of 1,8-cineole provide a fresh, crisp edge.

Appendix D 22: Frankincense

**Inferred Extraction Method: Steam Distillation of resin from Boswellia species**  
 **Overview: Frankincense (typically Boswellia carterii) is citrusy, resinous, and piney. Contains alpha-pinene, limonene, and incensole.**

**Descriptors and Justifications:**

Green: A sharp green top note is sometimes present from pinene.

Floral: The drydown is often described as soft and powdery, especially in fresh distillations.

Root: Earthy and grounded—some frankincense oils have a subtle carrot/root note.

Resinous: Being distilled from resin, this descriptor is essential.

Pine: α-pinene gives a pine-like freshness.

Smoky: Often associated with church incense, which includes frankincense resin burned over charcoal.

Phenolic: The smoky, tarry components lean phenolic.

Medicinal: Used historically in wound care and for spiritual cleansing.

Appendix D 23: Pine Scotch

**Inferred Extraction Method: Steam Distillation of needles and twigs**  
 **Overview: Pine Scotch (Pinus sylvestris) oil has a bright, clean, and invigorating scent due to high levels of α-pinene and β-pinene.**

**Descriptors and Justifications:**

Pungent: The oil’s fresh sharpness makes it stimulating and bracing.

Terpene: Rich in α-pinene and limonene—hallmark terpene content.

Appendix D 24: Elemi

**Inferred Extraction Method: Steam Distillation of resin**  
 **Overview: Elemi (Canarium luzonicum) is lemony-spicy and resinous, often considered a bridge between citrus and frankincense in perfumery.**

**Descriptors and Justifications:**

Spicy: Elemi has a peppery-citrus top note common in resin oils.

Green: Sharp leafy hints due to limonene and elemicin.

Fatty: A waxy, dense body sometimes emerges in aged samples.

Bitter: Touches of sharp citrus rind bitterness linger in the top.

Lemon: Rich in limonene, giving it a lemony bite.

Light: Uplifting and luminous; used in top notes for its brightness.

Resinous: Extracted from resin; balsamic and incense-like.

Powdery: The drydown includes a chalky, powdery skin scent.

Phenolic: Some samples carry a phenolic backbone akin to smoky pine.

Appendix D 25: Basil

**Inferred Extraction Method: Steam Distillation of fresh leaves**  
 **Overview: Basil essential oil (typically Ocimum basilicum) is rich in linalool and methyl chavicol (estragole), giving it a sweet-herbal and slightly spicy scent.**

**Descriptors and Justifications:**

Spicy: Estragole provides a warm, spicy nuance.

Sweet: Linalool lends a sweet, gentle floral note.

Anisic: Basil chemotypes high in estragole have a faint anise character.

Fruity: Linalool and methyl cinnamate give off soft fruit-like hints.

Peppery: A light peppery top note is present in some batches.

Dry: As it oxidizes, it can become woody and dry, especially in the base.

Appendix D 26: Cinnamon Bark

**Inferred Extraction Method: Steam Distillation of bark**  
 **Overview: Extracted from the inner bark of Cinnamomum verum or C. cassia, cinnamon bark oil is rich in cinnamaldehyde and eugenol. It is warmer, sweeter, and spicier than cinnamon leaf oil and is frequently used in perfumery and natural remedies.**

**Descriptors and Justifications:**

Spicy: Cinnamaldehyde, the dominant compound, gives cinnamon its classic sharp and warm spice aroma.

Sweet: Beneath the spice, eugenol and benzaldehyde contribute sugary, baked-good nuances.

Balsamic: The warm, dense base note with a syrup-like depth is typical of oils rich in aromatic aldehydes and alcohols.

Phenolic: While more dominant in cinnamon leaf, trace eugenol still imparts a slightly smoky-phenolic edge.

Medicinal: Cinnamon's traditional uses in oral care and digestion link it strongly to medicinal olfactory cues.

Warm: Emotionally and chemically warming—cinnamaldehyde is a rubefacient that triggers warmth when applied topically.

Appendix D 27: Clary Sage

**Inferred Extraction Method: Steam Distillation of flowering tops**  
 **Overview: Clary Sage (Salvia sclarea) is soft, herbaceous, and sweet with nutty and slightly amber-like undertones. Its key components include linalyl acetate, linalool, and sclareol.**

**Descriptors and Justifications:**

Floral: Linalyl acetate gives clary sage a rounded, musky-floral aroma similar to lavender.

Sweet: Mildly sweet due to its ester content and soft fruity backbone.

Fatty: The viscous, smooth quality of the oil imparts a fatty note in the mid and base.

Herbaceous: As a member of the mint family, it retains the green, leafy depth typical of Lamiaceae oils.

Minty: Mild top notes of linalool and terpinene offer a gentle, mint-adjacent freshness.

Aromatic: Complex aromatic presence due to its high ester content and minor diterpenes.

Appendix D 28: Cardamom

**Inferred Extraction Method: Steam Distillation of dried seeds**  
 **Overview: Extracted from Elettaria cardamomum, cardamom oil is sweet, spicy, and uplifting with notes of cineole, α-terpinyl acetate, and limonene.**

**Descriptors and Justifications:**

Spicy: Warm and clean spice derived from cineole and terpinyl acetate—characteristic of cardamom pods.

Sweet: Known for its bright, sugary top note—commonly used in confections and perfumes.

Floral: The middle note is gently floral, especially in perfumery use, due to ester-like smoothness.

Aromatic: Rich aromatic complexity from its balanced profile of esters and oxides.

Pleasant: Universally described as smooth and agreeable—rarely sharp or aggressive.

Terpene: Limonene and other monoterpenes contribute to the vibrant, volatile lift.

Appendix D 29: Chamomile (Roman)

**Inferred Extraction Method: Steam Distillation of fresh flowers**  
 **Overview: Derived from Chamaemelum nobile, Roman chamomile has a sweet, herbaceous scent with apple-like undertones due to angelic acid esters and isobutyl angelate.**

**Descriptors and Justifications:**

Sweet: The ester-rich profile gives Roman chamomile its soothing, apple-honey sweetness.

Herbaceous: Flowering plant oils, especially Asteraceae family members, contain green, leafy backnotes.

Tea: A signature chamomile feature—used for centuries in calming herbal teas.

Floral: A light, daisy-like floral note sits atop its base.

Warm: Known for its comforting and emotionally grounding effects.

Pleasant: Among the gentlest and most soothing oils; often used for children and sensitive users.

Appendix D 30: Palmarosa

**Inferred Extraction Method: Steam Distillation of grass**  
 **Overview: Extracted from Cymbopogon martini, palmarosa oil smells like a cross between rose and lemongrass. Its main component is geraniol.**

**Descriptors and Justifications:**

Floral: Geraniol gives palmarosa its distinctly rosy floral signature.

Sweet: Smooth and sugary in drydown, common in perfume bases.

Pleasant: Widely considered relaxing and agreeable—used for mood lifting.

Herbaceous: Grassy and green aspects are present in top notes due to its Cymbopogon origin.

Citrus: Has a citrus-leaning freshness due to linalool and limonene traces.

Dry: The finish is slightly powdery and dry as top notes fade.

Green: Its botanical origin as a grass yields a chlorophyllic sharpness early in the vapor trail.

Appendix D 31: Lime (Distilled)

**Inferred Extraction Method: Steam Distillation (unlike expressed lime, this version avoids phototoxicity)**  
 **Overview: Lime oil has a sharp, bright, and effervescent citrus scent rich in limonene and γ-terpinene.**

**Descriptors and Justifications:**

Citrus: The dominant character—limonene content is often over 50%.

Lemon: Close in chemical profile to lemon oil, with more tartness.

Pleasant: Zesty, uplifting, and commonly associated with cleanliness and freshness.

Orange: Some oils contain minor amounts of α-pinene and myrcene, adding soft orange-like notes.

Fresh: Crisp and invigorating top notes; among the brightest of citrus oils.

Dry: Steam-distilled lime often lacks the juicy aspect of cold-pressed oils, feeling more astringent.

Light: Rapidly evaporates, leaving a faint airy sweetness.

Appendix D 32: Marjoram (Sweet)

**Inferred Extraction Method: Steam Distillation of flowering tops**  
 **Overview: Origanum majorana oil is gentle, warm, and herbaceous, used traditionally in muscle relaxation blends. Key constituents include terpinen-4-ol and sabinene hydrate.**

**Descriptors and Justifications:**

Warm: Subtle warming qualities both in scent and topical effect.

Herbaceous: As a classic Mediterranean herb, this descriptor fits its green, savory signature.

Sweet: Softly sweet top notes, not as sugary as floral oils, but gently rounded.

Tea: Its warm herbaceousness resembles soft tea leaves in vapor.

Floral: Linalool and terpinolene contribute floral hints in drydown.

Pleasant: Known for emotional relaxation and soft body notes.

Appendix D 33: Hyssopus Officinalis

**Inferred Extraction Method: Steam Distillation of flowering tops**  
 **Overview: Hyssop oil is a pungent, medicinal-smelling oil with camphoraceous and herbal tones. Contains pinocamphone, isopinocamphone, and beta-pinene.**

**Descriptors and Justifications:**

Fresh: Camphor-rich top note opens with a sharp, clean hit.

Herbaceous: Traditional herbal tone with strong green complexity.

Medicinal: Classic medicinal scent due to camphor and terpene ketones.

Minty: Not sweet-mint, but a cooling mentholic sharpness is noticeable.

Sharp: Vaporous intensity immediately apparent in top notes.

Dry: The end note feels brittle and arid, characteristic of high ketone oils.

Pine: β-pinene gives a slightly foresty impression in the background.

Appendix D 34: Patchouli

**Inferred Extraction Method: Steam Distillation of dried leaves**  
 **Overview: Known for its deep, earthy, and musky scent, patchouli (Pogostemon cablin) is rich in patchoulol and other sesquiterpenes.**

**Descriptors and Justifications:**

Woody: Primary descriptor; dark wood-like base note dominates.

Earthy: Soil-heavy tones reminiscent of damp earth or composted bark.

Root: Deep, anchoring olfactory footprint similar to vetiver.

Musky: Natural muskiness due to sesquiterpenoid interactions with skin.

Resinous: Aged patchouli takes on thick, almost incense-like balsamic layers.

Warm: The base warms as it diffuses, often used as a fixative in perfume.

Appendix D 35: Cistus (Labdanum)

**Inferred Extraction Method: Steam Distillation of resin or oleoresin**  
 **Overview: Also known as labdanum oil, cistus comes from Cistus ladanifer and is resinous, amber-like, and leathery. Key molecules include labdane derivatives and pinene.**

**Descriptors and Justifications:**

Woody: Dense wood-resin notes saturate the profile.

Leathery: A defining feature—evokes tanned leather or antique books.

Resinous: Extracted from plant gum; incense-like and tenacious.

Earthy: Labdanum is rooted, dark, and reminiscent of forest ground.

Warm: Comforting, ambery warmth unfolds in the drydown.

Balsamic: Syrupy, sticky sweetness lies beneath the wood and leather.

Appendix D 36: Lemongrass (Cymbopogon Citratus)

**Inferred Extraction Method: Steam Distillation of fresh or dried grass**  
 **Overview: Lemongrass oil, rich in citral (a mix of geranial and neral), has a powerful lemony, grassy, and slightly spicy profile. Commonly used in insect repellents, perfumery, and cuisine.**

**Descriptors and Justifications:**

Citrus: Citral contributes directly to its lemon-citrus identity.

Lemon: The sharp citral aroma mimics fresh lemon zest, especially in top notes.

Grassy: Extracted from a grass species; its chlorophyllic freshness persists throughout.

Pleasant: Known for its energizing and uplifting effects, widely used in aromatherapy.

Herbaceous: Slightly vegetal undertone reflects its herbal culinary uses.

Dry: Astringent drydown, with aldehydic crispness as citrus fades.

Appendix D 37: Helichrysum (Everlasting)

**Inferred Extraction Method: Steam Distillation of flowering tops**  
 **Overview: Also called immortelle, this oil from Helichrysum italicum carries sweet, tea-like, and hay-floral notes. It contains neryl acetate, curcumene, and italidione.**

**Descriptors and Justifications:**

Sweet: Neryl acetate lends a smooth, honeyed sweetness.

Tea: Its herbaceous-floral top mimics chamomile or green tea.

Warm: The overall scent profile is cozy, slightly musky, and healing.

Pleasant: Mild and emotionally soothing—used in trauma and scar blends.

Earthy: Drydown leans toward hay, soil, and moss due to sesquiterpenes.

Fatty: The thick mouthfeel and texture of the oil create a "fatty" base sensation.

Appendix D 38: Cypress

**Inferred Extraction Method: Steam Distillation of needles and twigs**  
 **Overview: From Cupressus sempervirens, this oil is a dry, woody conifer scent with green top notes and a musky finish. α-pinene and cedrol dominate its chemical structure.**

**Descriptors and Justifications:**

Woody: Derived from coniferous trees; backbone scent is dry wood.

Root: Slightly grounding, earthy base with root-like vetiver tones.

Earthy: Damp forest floor nuance—linked to terpenoid content.

Fresh: Camphoraceous freshness hits early in the diffusion.

Pine: Pinene content gives crisp conifer clarity.

Dry: Overall scent has an arid, astringent feel.

Appendix D 39: Myrrh

**Inferred Extraction Method: Steam Distillation of resin**  
 **Overview: Derived from the gum of Commiphora myrrha, this oil is smoky, earthy, and slightly bitter. Main compounds include furanoeudesma-1,3-diene and curzerene.**

**Descriptors and Justifications:**

Smoky: Burnt incense quality common to resin distillates.

Root: Deep and bitter—connects olfactorily to root vegetables or vetiver.

Earthy: Heavy and mineral-rich, evoking arid soil.

Resinous: Extracted from resin—sticky, balsamic, thick.

Balsamic: Sweet and warm base note typical of ancient incense materials.

Fatty: Dense, waxy base texture, especially in aged oils.

Appendix D 40: Geranium (Egyptian)

**Inferred Extraction Method: Steam Distillation of leaves and stems**  
 **Overview: This floral-green oil from Pelargonium graveolens contains citronellol, geraniol, and linalool, giving it a complex floral-herbal scent.**

**Descriptors and Justifications:**

Floral: High geraniol content makes it a rose-like top note.

Green: Sharp leaf and stem notes accompany the floral sweetness.

Citronellal: Direct presence in the chemical profile—adds citrusy brightness.

Pleasant: Balanced, rounded scent used heavily in perfumery.

Sweet: Rosey-sweet drydown.

Light: Not as powdery or heavy as true rose; floats gently.

Appendix D 41: Lemon (Cold-Pressed)

**Inferred Extraction Method: Cold Pressed from rinds**  
 **Overview: Lemon oil is a top-note citrus oil with high limonene content. The cold-pressed version retains minor non-volatile compounds and oxidizes faster.**

**Descriptors and Justifications:**

Citrus: Primary profile—fresh zest.

Lemon: Limonene, β-pinene, and citral create bright lemon scent.

Orange: Trace compounds like γ-terpinene and α-pinene offer orange-like warmth.

Pleasant: Energetic and uplifting; commonly linked with cleanliness.

Fresh: Crisp, volatile, evaporates quickly.

Light: High in top notes—fleeting and delicate.

Appendix D 42: Eucalyptus (Citriodora)

**Inferred Extraction Method: Steam Distillation of leaves**  
 **Overview: This variant of eucalyptus contains high levels of citronellal, making it citrusy and less camphorous than E. globulus.**

**Descriptors and Justifications:**

Citronellal: Named for it; this is the dominant compound.

Fresh: Cooling top notes akin to clean linen or disinfectant.

Minty: Mild secondary note from cineole, if present.

Cooling: Opens airways and produces trigeminal cooling effects.

Pleasant: Considered more palatable than traditional eucalyptus.

Green: Leafy freshness overlays the citrus-like sharpness.

Appendix D 43: Orange (Sweet)

**Inferred Extraction Method: Cold Pressed from rinds**  
 **Overview: Sweet orange oil is cheerful, fruity, and sugary, dominated by limonene (often 90%+).**

**Descriptors and Justifications:**

Citrus: Archetypal citrus note, juicy and radiant.

Orange: Most direct representation of this fruit in essential oil form.

Pleasant: Universally liked, calming and joyful.

Sweet: Rich in sugars and esters, like fruit juice.

Fresh: Bursting top note when diffused.

Fruity: Obvious fruit resemblance in aroma.

Appendix D 44: Peppermint (English)

**Inferred Extraction Method: Steam Distillation of leaves**  
 **Overview: Rich in menthol and menthone, peppermint oil is strong, cooling, and minty—often used for alertness and pain relief.**

**Descriptors and Justifications:**

Minty: Dominated by menthol; defines mint category.

Cooling: Topical and nasal cooling effects are immediate.

Fresh: Sharp, awakening aroma with high vapor tension.

Sharp: Edgy and piercing—menthol stimulates trigeminal nerves.

Pleasant: Commonly used in therapeutic and consumer applications.

Herbaceous: Originates from mint family; green and plant-like tones in base.

Appendix D 45: Thyme (Sweet)

**Inferred Extraction Method: Steam Distillation of flowering tops**  
 **Overview: Thymus vulgaris (sweet chemotype) contains linalool and borneol, offering a milder, sweeter version of typical thyme.**

**Descriptors and Justifications:**

Sweet: Rich in linalool, softens the spicy harshness.

Spicy: Retains a degree of warm, savory spice from thymol/borneol.

Herbaceous: Thyme's essential character—dry, Mediterranean herb tone.

Peppery: Slight piquancy on the nose in opening notes.

Minty: Subtle cooling backnote from p-cymene or cineole content.

Light: The chemotype gives a delicate, airy impression vs. traditional thyme.

Appendix D 46: Sandalwood

**Inferred Extraction Method: Steam Distillation of heartwood**  
 **Overview: Sandalwood oil, primarily from Santalum album, is revered for its rich, smooth, woody scent. Dominated by santalol (α- and β-santalol), this oil has a persistent, grounding aroma used in perfumery and spiritual rituals.**

**Descriptors and Justifications:**

Woody: The primary olfactory profile; derived from santalols that give a creamy, woody depth.

Sweet: A soft, rounded sweetness balances the dry wood base—distinct from sugary sweet but warm and enveloping.

Balsamic: Its fixative quality and resin-like drydown resemble that of aged balsam or labdanum.

Resinous: Though not a true resin, its viscosity and smooth base note echo this quality.

Appendix D 47: Myrrh

**Inferred Extraction Method: Steam Distillation of resin**  
 **Overview: Myrrh oil is distilled from the gum resin of Commiphora myrrha. Its earthy, smoky aroma is used for grounding and calming in both aromatherapy and perfumery. Key compounds include curzerene, furanoeudesma-1,3-diene, and limonene.**

**Descriptors and Justifications:**

Smoky: Characteristic burnt-incense nuance; common to dried resins upon distillation.

Root: The earthy depth is reminiscent of root oils like vetiver or angelica root.

Earthy: Strong mineral and dust-like undertones from sesquiterpenes.

Resinous: Naturally derived from hardened resin—thick, tenacious scent.

Balsamic: Warm, meditative note like benzoin or frankincense.

Fatty: Slightly waxy density in aged oils or thicker cuts; linked to lipid-like fixative nature.

Appendix D 48: Rose Absolute

**Inferred Extraction Method: Solvent Extraction of petals**  
 **Overview: Rose absolute is obtained through solvent extraction of Rosa damascena or Rosa centifolia petals. It is richer and heavier than steam-distilled rose otto, containing phenylethyl alcohol, citronellol, geraniol, and eugenol.**

**Descriptors and Justifications:**

Floral: The defining characteristic—delicate, lush, and complex.

Sweet: Strong sugar-like opening from phenylethyl alcohol.

Pleasant: Universally recognized as comforting and luxurious.

Light: Despite richness, top notes remain ethereal and uplifting.

Green: Leafy nuance from citronellol and trace aldehydes gives it a realistic bloom character.

Appendix D 49: Chamomile (Maroc)

**Inferred Extraction Method: Steam Distillation of flowers (Ormenis multicaulis)**  
 **Overview: Known as Moroccan chamomile, this oil is chemically different from German and Roman varieties. It has a delicate, herbaceous floral profile rich in pinocarvone and trans-pinocarveol.**

**Descriptors and Justifications:**

Herbaceous: Slightly sharp, grassy-green from monoterpenes.

Fresh: Clean, airy quality—lighter than Roman or German types.

Floral: Mild floral backdrop from aromatic esters and lactones.

Pleasant: Calming and balancing effect; popular in gentle aromatherapy blends.

Appendix D 50: Melissa

**Inferred Extraction Method: Steam Distillation of aerial parts**  
 **Overview: Melissa (lemon balm) oil is rare and expensive due to its low yield. It features citronellal, geranial, and neral, offering a lemony-herbal scent with a soft floral drydown.**

**Descriptors and Justifications:**

Lemon: Citral compounds give this a citrus identity without being sharp like lemon oil.

Citronellal: Secondary compound lending a clean, lemony-metallic tone.

Floral: Slight floral nuance from geraniol traces in the drydown.

Pleasant: Widely regarded as calming and soft, often used for stress relief.

Sweet: Round, honeyed finish—distinct from tart lemon.

Light: The volatility and smooth tone contribute to a floating, gentle aroma.

Appendix D 51: Patchouli

**Inferred Extraction Method: Steam Distillation of dried leaves**  
 **Overview: Extracted from Pogostemon cablin, patchouli oil is earthy, dark, and musky, often used as a fixative. Its signature profile comes from patchoulol and other sesquiterpenes.**

**Descriptors and Justifications:**

Earthy: The dominant characteristic—rich, soil-like base notes due to high patchoulol content.

Woody: Dry down reveals aged wood facets; enhanced by α-bulnesene and guaiene isomers.

Warm: Deep, grounding warmth in the base from oxidized sesquiterpenes.

Root: Aged patchouli smells like roots or damp forest floor—analogous to vetiver.

Fatty: Dense, waxy feel in texture and smell—often used in base-heavy perfume blends.

Musky: Animalic tone in the drydown; often described as primal or sensual.

Appendix D 52: Clary Sage

**Inferred Extraction Method: Steam Distillation of flowering tops and leaves**  
 **Overview: Salvia sclarea yields a soft, herbaceous oil known for its relaxing and hormone-balancing properties. Dominated by linalyl acetate and linalool.**

**Descriptors and Justifications:**

Herbaceous: Sage-like green, dry top notes from sclareol and monoterpenes.

Sweet: Linalool adds a soft floral sweetness.

Green: Crisp leafy tone prominent in early diffusion.

Pleasant: Universally appreciated as calming and non-irritating.

Minty: Subtle menthol edge—possibly from α-terpineol or trace 1,8-cineole.

Appendix D 53: Neroli

**Inferred Extraction Method: Steam Distillation of Citrus aurantium blossoms**  
 **Overview: Neroli is derived from the bitter orange tree’s flowers, distinct from orange blossom absolute. It has a delicate, floral, and citrus profile due to linalool, nerolidol, and limonene.**

**Descriptors and Justifications:**

Floral: White floral elegance from nerolidol and linalool.

Citrus: Top note zest from limonene gives brightness.

Pleasant: Often used in perfumery for calming and universally liked nature.

Fresh: Crisp, aldehydic opening adds effervescence.

Light: Wispy, ephemeral top note quality characteristic of blossom oils.

Appendix D 54: Galbanum

**Inferred Extraction Method: Steam Distillation of resin**  
 **Overview: Galbanum oil, obtained from Ferula galbaniflua resin, is intensely green and bitter, with sharp aldehydes and terpenes. Used in vintage perfumery and grounding blends.**

**Descriptors and Justifications:**

Green: One of the greenest scents—hexenyl compounds dominate.

Bitter: High concentration of terpenoids like cadinene give acrid, sharp bite.

Root: Deep earthy base, reminiscent of celery root or parsnip.

Resinous: Extracted from resin; sticky, balsamic base remains throughout.

Sharp: Top note has piercing, almost metallic quality.

Appendix D 55: Jasmine Absolute

**Inferred Extraction Method: Solvent extraction of fresh blossoms (Jasminum grandiflorum)**  
 **Overview: Jasmine absolute is a rich, intensely floral oil containing benzyl acetate, indole, and linalool. It is widely used in perfumery and sensual blends.**

**Descriptors and Justifications:**

Floral: Iconic floral note—indole and benzyl acetate define this class.

Sweet: Sugar-like warmth in the mid and base notes.

Pleasant: Evocative, euphoric—common in calming and aphrodisiac formulas.

Light: Initial top notes float before deepening to heady richness.

Aromatic: Complex blend of aldehydes, esters, and alcohols give rich olfactory complexity.

Appendix D 56: Rosemary

**Inferred Extraction Method: Steam Distillation of leaves and flowering tops**  
 **Overview: Rosmarinus officinalis oil is sharp, herbaceous, and stimulating. Rich in 1,8-cineole, camphor, and α-pinene.**

**Descriptors and Justifications:**

Herbaceous: Classic sharp green note, typical of Mediterranean shrubs.

Fresh: Eucalyptol (1,8-cineole) contributes to fresh, invigorating quality.

Minty: Not as dominant as peppermint, but present from trace menthone-like molecules.

Sharp: Penetrating top notes from α-pinene and cineole stimulate focus.

Pleasant: Uplifting and commonly used in wellness products.

Appendix D 57: Cardamom

**Inferred Extraction Method: Steam Distillation of seeds**  
 **Overview: Elettaria cardamomum yields a bright, spicy-sweet oil dominated by cineole, terpinyl acetate, and linalyl acetate.**

**Descriptors and Justifications:**

Spicy: Warming, aromatic top note due to eucalyptol and other terpenes.

Sweet: Terpinyl acetate contributes a sugary green sweetness.

Aromatic: Complex scent often likened to incense and sweet herbs.

Pleasant: Bright and well-tolerated; used to soothe nausea and anxiety.

Light: Not overpowering; airy and clean spice.

Appendix D 58: Bergamot

**Inferred Extraction Method: Cold Pressed from rind**  
 **Overview: From the peel of Citrus bergamia, this oil combines citrus brightness with green and floral undertones. Main compounds include limonene, linalyl acetate, and linalool.**

**Descriptors and Justifications:**

Citrus: Primary profile—zesty top note from limonene.

Lemon: Bright lemon-lime scent from α-pinene and citral.

Pleasant: Common in perfumery and Earl Grey tea.

Sweet: Linalool and linalyl acetate balance tartness.

Fresh: Opening is bright, volatile, and clean.

Appendix D 59: Spikenard

**Inferred Extraction Method: Steam Distillation of roots**  
 **Overview: Spikenard (Nardostachys jatamansi) is musky, earthy, and deep. It has a history in ancient rituals and Ayurvedic medicine. Composed of valerianol, jatamansone, and calarene.**

**Descriptors and Justifications:**

Earthy: Dominant scent—reminiscent of moss, soil, or decay.

Root: True root oil; dense and grounding.

Musky: Animalic tones that echo civet or aged leather.

Woody: Drydown becomes more sandalwood-like and aged.

Resinous: Sticky, ambered base can resemble dried sap.

Appendix D 60: Litsea Cubeba

**Inferred Extraction Method: Steam Distillation of fruit**  
 **Overview: Also known as May Chang, this oil has a lemony, fresh, and slightly spicy profile. Dominated by citral, limonene, and geranial.**

**Descriptors and Justifications:**

Lemon: Citral is 70–85% of the oil—defining the lemon scent.

Citrus: Sparkling top notes that mimic lemon peel.

Pleasant: Used for calming and mood-brightening.

Fresh: Extremely volatile and airy on first impression.

Spicy: Slight backnote from terpenes adds zing to the citrus.

Appendix D 61: Ylang Ylang (Extra)

**Inferred Extraction Method: Steam Distillation (fractional, early cut) of flowers**  
 **Overview: The “Extra” grade captures the top floral and fruity esters of Cananga odorata. Rich in benzyl acetate, linalool, and methyl benzoate.**

**Descriptors and Justifications:**

Floral: Heady, exotic floral from esters and alcohols.

Sweet: Rich, sensual sweetness—sometimes overwhelming.

Fruity: Banana-like, due to isoamyl acetate and related compounds.

Pleasant: Used in calming, romantic blends.

Light: “Extra” grade has more volatility than base-grade ylang.

Appendix D 62: Spearmint

**Inferred Extraction Method: Steam Distillation of leaves**  
 **Overview: Milder than peppermint, spearmint (Mentha spicata) contains carvone, limonene, and menthone derivatives.**

**Descriptors and Justifications:**

Minty: Carvone provides classic mint candy profile.

Cooling: Gentle trigeminal effect—softer than menthol.

Fresh: Clean, bright top note from limonene.

Sweet: Spearmint has a naturally sweet base from esters.

Pleasant: Popular in children’s products for its soft character.

Appendix D 63: Juniper Berry

**Inferred Extraction Method: Steam Distillation of dried berries**  
 **Overview: Extracted from Juniperus communis, this oil is bright, dry, and pine-like with a clean, coniferous aroma. Composed of α-pinene, sabinene, and myrcene.**

**Descriptors and Justifications:**

Woody: Dry cedar-like base from sesquiterpenes.

Pine: High α-pinene content gives forest-fresh tone.

Fresh: Sparkling top note from monoterpenes.

Sharp: Top-heavy oil that stings gently at high doses.

Pleasant: Popular in colognes and detox blends.

Appendix D 64: Helichrysum

**Inferred Extraction Method: Steam Distillation of flowers (Helichrysum italicum)**  
 **Overview: Also called immortelle, this oil has a medicinal, honeyed, and herbal scent. Rich in neryl acetate, diketones, and curcumenes.**

**Descriptors and Justifications:**

Resinous: Thick, rich base note resembling aged sap or glue.

Herbaceous: Bitter green top notes from artemisia-like components.

Sweet: Dried honey-flower sweetness in drydown.

Medicinal: Used in wound care; scent resembles bandages or aged herbal tinctures.

Root: Earthy base akin to turmeric or dried roots.

Appendix D 65: Lemongrass

**Inferred Extraction Method: Steam Distillation of grass blades**  
 **Overview: Cymbopogon citratus produces a bold lemon-scented oil high in citral and myrcene.**

**Descriptors and Justifications:**

Lemon: Citral dominates—up to 80% of content.

Grassy: Leafy, dry-grass tone from myrcene and aldehydes.

Citrus: Sparkling citrus bite—harsher than lemon oil.

Pleasant: Used widely in soaps and massage oils.

Sharp: Pungent top note—stimulates with aldehydic impact.

Appendix D 66: Cinnamon (Bark)

**Inferred Extraction Method: Steam Distillation of inner bark (Cinnamomum verum)**  
 **Overview: Cinnamon bark oil is hot, spicy, and intense. Rich in cinnamaldehyde, it is more pungent than cinnamon leaf oil and often used in stimulating or festive blends.**

**Descriptors and Justifications:**

Spicy: Dominated by cinnamaldehyde—warm, piquant aroma.

Sweet: Sugary undertone beneath the spice, especially as it dries down.

Warm: Deep internal heat—used in blends for circulation and seasonal warmth.

Pungent: Nose-stinging sharpness; one of the strongest oils aromatically.

Pleasant: Widely liked in low doses; festive, nostalgic aroma.

Appendix D 67: Lime (Distilled)

**Inferred Extraction Method: Steam Distillation of rind**  
 **Overview: Unlike cold-pressed lime oil, distilled lime is clear and phototoxic-free. It smells more tart and less sweet due to the removal of heavier compounds.**

**Descriptors and Justifications:**

Citrus: Pure lime aroma with minimal bitterness.

Lemon: Has citral and limonene crossover with lemon oils.

Fresh: Zesty and clean; more volatile than pressed variants.

Pleasant: Bright and clean; ideal for uplifting environments.

Sharp: Tart aldehydes give a piercing freshness on the nose.

Appendix D 68: Fennel (Sweet)

**Inferred Extraction Method: Steam Distillation of seeds**  
 **Overview: Sweet fennel oil is mildly sweet, herbaceous, and spicy. High in trans-anethole, it’s often compared to licorice or anise.**

**Descriptors and Justifications:**

Sweet: Anethole imparts a sugar-like licorice note.

Spicy: Warming character common to seed oils.

Herbaceous: Undercurrent of dry green notes from fenchone.

Pleasant: Calming and well-liked in digestive or hormonal blends.

Anisic: Smells distinctly like anise candy or absinthe.

Appendix D 69: Cypress

**Inferred Extraction Method: Steam Distillation of needles and twigs**  
 **Overview: Cypress oil has a clean, woody, and slightly smoky profile. Its primary components are α-pinene, δ-3-carene, and cedrol.**

**Descriptors and Justifications:**

Woody: Strong coniferous profile like pine or fir.

Earthy: Dry and grounding base.

Fresh: Bright, forest-like top note.

Pleasant: Often used in emotional support and detox blends.

Resinous: Sticky, pine-sap undertone from sesquiterpenes.

Appendix D 70: Tarragon

**Inferred Extraction Method: Steam Distillation of aerial parts**  
 **Overview: Artemisia dracunculus produces a sweet, green oil high in estragole (methyl chavicol), which smells similar to anise but more herbal.**

**Descriptors and Justifications:**

Anisic: Main compound estragole gives licorice-like note.

Green: Bright green top note—leafy, grassy feel.

Herbaceous: Classic Mediterranean herb quality.

Pleasant: Gentle sweetness makes it widely liked.

Light: Fleeting aromatic profile with quick evaporation.

Appendix D 71: Valerian

**Inferred Extraction Method: Steam Distillation of roots**  
 **Overview: Valerian root oil is deeply earthy, musky, and pungent. Key molecules include valerianol, bornyl acetate, and isovaleric acid.**

**Descriptors and Justifications:**

Earthy: Dominates the scent—deep, wet soil aroma.

Root: Strong underground vegetal profile.

Musky: Animalic quality from isovaleric compounds.

Bitter: Harsh, acrid note, especially in high concentrations.

Pungent: Piercing, unpleasant to some—very potent.

Appendix D 72: Palmarosa

**Inferred Extraction Method: Steam Distillation of grass blades**  
 **Overview: Cymbopogon martinii oil has a rosy-floral scent despite being a grass. Rich in geraniol and geranyl acetate.**

**Descriptors and Justifications:**

Floral: Geraniol mimics rose-like sweetness.

Sweet: Honeyed and slightly fruity undertone.

Pleasant: Used as a rose substitute in perfumery.

Fresh: Bright and dewy initial top note.

Green: Grassy and herbal backdrop balances florals.

Appendix D 73: Ginger

**Inferred Extraction Method: Steam Distillation of dried rhizome**  
 **Overview: Ginger oil is warm, spicy, and slightly citrusy. It contains zingiberene, β-sesquiphellandrene, and ar-curcumene.**

**Descriptors and Justifications:**

Spicy: Defining characteristic—peppery heat from gingerols.

Warm: Deep internal warmth used in circulation blends.

Pleasant: Soothing when diffused at low levels.

Sharp: Biting opening note from fresh terpenes.

Root: Earthy tone remains in drydown—reminds of fresh ginger.

Appendix D 74: Black Spruce

**Inferred Extraction Method: Steam Distillation of needles**  
 **Overview: Picea mariana produces a fresh, piney oil with deep balsamic notes. Major components include bornyl acetate, camphene, and α-pinene.**

**Descriptors and Justifications:**

Pine: Main scent—classic evergreen note.

Woody: Dry woody base after diffusion.

Fresh: Cooling top notes from camphene and limonene.

Resinous: Sticky, sap-like nuance from sesquiterpenes.

Pleasant: Uplifting and grounding—used in emotional blends.

Appendix D 75: Basil (Sweet)

**Inferred Extraction Method: Steam Distillation of flowering tops**  
 **Overview: This chemotype (Ocimum basilicum) is rich in linalool and methyl chavicol. It’s fresh, herbaceous, and slightly spicy.**

**Descriptors and Justifications:**

Herbaceous: Classic green culinary tone.

Spicy: Mild clove-like spice from eugenol and estragole.

Fresh: Sparkling top note from linalool and cineole.

Pleasant: Energizing and familiar to most users.

Minty: Faint menthol note from trace camphor or eucalyptol.

Appendix D 76: Cajeput

**Inferred Extraction Method: Steam Distillation of leaves and twigs (Melaleuca cajuputi)**  
 **Overview: Often compared to tea tree and eucalyptus, cajeput oil is piercingly fresh, camphorous, and herbaceous. It contains high levels of 1,8-cineole, α-terpineol, and limonene.**

**Descriptors and Justifications:**

Fresh: Immediate vapor tension upon diffusion—stimulates the nasal passages with sharp clarity.

Cooling: The eucalyptol content triggers a trigeminal cooling effect, opening airways and offering a medicinal sensation.

Eucalyptol: A primary compound—1,8-cineole is the backbone of the oil’s scent and functionality.

Herbaceous: Leafy-green background stemming from the Melaleuca family origin and its secondary terpene content.

Pleasant: Despite its intensity, cajeput is favored in respiratory and purifying blends for its cleansing and revitalizing feel.

Appendix D 77: Anise

**Inferred Extraction Method: Steam Distillation of seeds (Pimpinella anisum)**  
 **Overview: Anise seed oil is intensely sweet and licorice-like, with trans-anethole as the dominant compound, giving it its characteristic flavor and aroma.**

**Descriptors and Justifications:**

Anisic: Defined by trans-anethole, which provides the unmistakable scent of licorice and anisette liqueurs.

Sweet: One of the sweetest essential oils—anethole mimics sugar on the olfactory palette.

Pleasant: Strong but often nostalgic; used in traditional medicine and flavoring.

Aromatic: Complex mixture of phenolic and terpenoid compounds adds depth beyond its sweetness.

Spicy: While not spicy-hot, the oil has warm intensity and sharpness from phenylpropanoids.

Appendix D 78: Frankincense (Carterii)

**Inferred Extraction Method: Steam Distillation of resin (Boswellia carterii)**  
 **Overview: Known for its meditative and sacred use, frankincense oil is rich in α-pinene, limonene, and incensole acetate. It has a warm, balsamic, and resinous profile.**

**Descriptors and Justifications:**

Resinous: Extracted from tree resin—thick, balsamic core is omnipresent.

Woody: Drydown resembles aged wood and incense, anchored by sesquiterpenes.

Smoky: Subtle incense-like smoke tone—especially in aged or oxidized oils.

Balsamic: Smooth, warm, slightly sweet and rounded aroma from ester components.

Pleasant: Widely used in calming, grounding applications for emotional balance.

Appendix D 79: Cedarwood (Virginian)

**Inferred Extraction Method: Steam Distillation of wood (Juniperus virginiana)**  
 **Overview: Cedarwood Virginian is dry, woody, and pencil-shaving-like. It has high levels of cedrol and thujopsene and is technically a type of juniper.**

**Descriptors and Justifications:**

Woody: Archetypal wood scent—evokes cedar chests and freshly cut timber.

Earthy: Grounding tone present in the base from aged wood and resinous lignans.

Resinous: Waxy, balsamic depth—although lighter than Atlas or Himalayan cedarwoods.

Smoky: A faint smoky veil may emerge, especially if oxidized.

Root: Anchoring, subterranean character likened to vetiver or dry soil.

Appendix D 80: Vetiver

**Inferred Extraction Method: Steam Distillation of roots (Chrysopogon zizanioides)**  
 **Overview: Vetiver oil is profoundly earthy, smoky, and grounding. It contains khusimol, vetivenol, and other complex sesquiterpenes that contribute to its viscosity and depth.**

**Descriptors and Justifications:**

Earthy: A defining trait—like damp, mineral-rich soil.

Root: Extracted from roots and smells as such—deep, anchored, with moist depth.

Woody: Drydown is akin to aged wood or incense resin.

Smoky: Naturally smoky due to slow distillation and chemical profile.

Resinous: Viscous and balsamic undercurrent—acts as a fixative in perfumery.

Appendix D 81: Chamomile (Roman)

**Inferred Extraction Method: Steam Distillation of flowers (Anthemis nobilis)**  
 **Overview: Roman chamomile has a soft, apple-like aroma with calming, sweet, and herbal tones. It contains esters such as isobutyl angelate and methallyl angelate.**

**Descriptors and Justifications:**

Floral: Gentle floral top note—like wildflowers, contributed by ester content.

Sweet: Honey-apple scent is soft, syrupy, and comforting.

Pleasant: Universally associated with calm and safety—used widely for children and anxiety relief.

Light: Evaporates easily; has a wispy, ephemeral nature.

Tea: Soft infusion note similar to warm chamomile herbal tea.

Appendix D 82: Marjoram (Sweet)

**Inferred Extraction Method: Steam Distillation of aerial parts (Origanum majorana)**  
 **Overview: Sweet marjoram has a warm, woody-herbal scent. It contains terpinen-4-ol, sabinene hydrate, and linalool.**

**Descriptors and Justifications:**

Herbaceous: Classic Mediterranean herbal character from monoterpenols.

Spicy: Warming character reminiscent of culinary herbs like oregano.

Pleasant: Soothing and well-tolerated—used for stress and sleep support.

Tea: Brew-like softness emerges in drydown, especially when blended.

Light: More delicate than oregano or thyme—subtle and airy.

Appendix D 83: Myrtle

**Inferred Extraction Method: Steam Distillation of leaves and twigs (Myrtus communis)**  
 **Overview: Myrtle oil has a fresh, camphorous profile with green and floral nuances. It is rich in 1,8-cineole and myrtenyl acetate.**

**Descriptors and Justifications:**

Fresh: Invigorating eucalyptus-like top note from cineole.

Green: Crisp leafiness layered over the camphoraceous burst.

Herbaceous: Mediterranean shrub tone from linalool and myrtenol.

Tea: Slight tea-leaf warmth in drydown—subtle and pleasant.

Pleasant: Calming and clean—used in gentle respiratory blends.

Appendix D 84: Elemi

**Inferred Extraction Method: Steam Distillation of resin (Canarium luzonicum)**  
 **Overview: Elemi oil is citrusy-resinous and slightly peppery. It is a close relative to frankincense and myrrh and is often used in perfumery.**

**Descriptors and Justifications:**

Resinous: Extracted from oleoresin—sticky, balsamic aroma.

Citrus: Limonene and phellandrene give sparkling citrus lift.

Pleasant: Balancing and harmonizing—middle note used in sacred blends.

Sweet: Soft and honeyed undertones develop over time.

Woody: Drydown reveals dry, incense-like character.

Appendix D 85: Cinnamon (Leaf)

**Inferred Extraction Method: Steam Distillation of leaves (Cinnamomum verum)**  
 **Overview: Cinnamon leaf oil is more herbaceous and less pungent than bark oil. It is high in eugenol and often used for muscular or antimicrobial blends.**

**Descriptors and Justifications:**

Spicy: Warm clove-like spice due to eugenol dominance.

Herbaceous: Leafy quality overlays the spice, making it less fiery.

Warm: Cozy, heating scent common in massage blends.

Pleasant: Well tolerated in dilution—used in fall/winter blends.

Bitter: Slight astringency from phenols in the top notes.

Appendix D 86: Cardamom

**Inferred Extraction Method: Steam Distillation of dried seeds (Elettaria cardamomum)**  
 **Overview: Cardamom oil has a warm, spicy, and slightly sweet aroma. Rich in α-terpinyl acetate, cineole, and limonene, it's used in both perfumery and digestion blends.**

**Descriptors and Justifications:**

Spicy: Peppery warmth from cineole and terpenes.

Sweet: Smooth, soft background sweetness from esters.

Pleasant: Often used for relaxation and clarity; appealing across cultures.

Aromatic: Intensely fragrant—common in Indian and Middle Eastern perfumes.

Cooling: Cineole content gives it a fresh yet warming duality.

Appendix D 87: Bay Laurel

**Inferred Extraction Method: Steam Distillation of leaves (Laurus nobilis)**  
 **Overview: Bay laurel oil is herbaceous, spicy, and slightly floral. Contains 1,8-cineole, linalool, and eugenol.**

**Descriptors and Justifications:**

Spicy: Warm, clove-like sharpness from eugenol.

Herbaceous: Leafy, aromatic profile—classic Mediterranean cooking scent.

Pleasant: Invigorating and balancing; used for confidence and clarity.

Aromatic: Rich in volatile oils—potent and layered.

Minty: Slight mentholated freshness from cineole backbone.

Appendix D 88: Helichrysum (Italicum)

**Inferred Extraction Method: Steam Distillation of flowering tops**  
 **Overview: Known as immortelle, helichrysum oil has a sweet, earthy, and honey-like scent. High in neryl acetate and diketones, it's revered for healing properties.**

**Descriptors and Justifications:**

Sweet: Warm honey note—a signature trait of neryl esters.

Earthy: Base note depth—rich in grounding, dry soil-like aroma.

Balsamic: Syrupy warmth in drydown—almost medicinal in tone.

Pleasant: Comforting and calming—used in trauma recovery blends.

Woody: Slight woody drydown from sesquiterpenes.

Appendix D 89: Yarrow

**Inferred Extraction Method: Steam Distillation of flowering tops**  
 **Overview: Yarrow oil is blue in color due to chamazulene and has a herbaceous, earthy scent. It’s a cousin of chamomile and is used in skin and wound care.**

**Descriptors and Justifications:**

Earthy: Soil-heavy aroma typical of wild flowering herbs.

Herbaceous: Dry green sharpness from monoterpenes.

Pleasant: Used for calming emotions and treating skin; regarded positively.

Resinous: Chamazulene gives viscous, inky, balsamic tone.

Smoky: A soft smokiness from the azulene component during drydown.

Appendix D 90: Benzoin

**Inferred Extraction Method: Solvent extraction of resin (Styrax benzoin)**  
 **Overview: A balsamic resin with a vanilla-like sweetness, benzoin is used as a fixative and in grounding blends.**

**Descriptors and Justifications:**

Sweet: Vanilla-like profile from benzoic and cinnamic acid esters.

Resinous: Gum resin origin defines the core of its scent.

Balsamic: Rich, syrupy, and lingering—common base note in perfumery.

Warm: Deep, comforting aroma like spiced amber.

Pleasant: Universally liked—used in spiritual and meditative traditions.

Appendix D 91: Spikenard

**Inferred Extraction Method: Steam Distillation of rhizomes (Nardostachys jatamansi)**  
 **Overview: Spikenard is intensely earthy, woody, and musky. It contains jatamansone and is traditionally used in spiritual rituals.**

**Descriptors and Justifications:**

Earthy: Damp forest floor scent—deeply grounding.

Root: Rhizome origin is apparent—pungent and subterranean.

Woody: Aged wood and incense-like notes as it dries down.

Musky: Animalic depth—evokes sacred incense.

Bitter: Harsh opening notes—medicinal, almost acrid in pure form.

Appendix D 92: Peru Balsam

**Inferred Extraction Method: Solvent extraction of tree resin (Myroxylon balsamum)**  
 **Overview: Peru balsam smells sweet, vanilla-like, and resinous. Contains cinnamic acid and benzyl benzoate.**

**Descriptors and Justifications:**

Sweet: Syrupy, caramel warmth—used in natural vanilla alternatives.

Balsamic: Sticky, rounded warmth—typical of resin distillates.

Resinous: Gum exudate origin provides lasting base note.

Pleasant: Soothing and nostalgic scent profile.

Warm: Cozy and sensual—often used in perfumery.

Appendix D 93: Wintergreen

**Inferred Extraction Method: Steam Distillation of leaves (Gaultheria procumbens)**  
 **Overview: Known for its minty, medicinal scent, wintergreen is high in methyl salicylate—chemically similar to aspirin.**

**Descriptors and Justifications:**

Minty: Signature minty sharpness—identical to root beer or muscle rubs.

Cooling: Evaporates with a trigeminal cooling sensation.

Sharp: Penetrating scent—stimulates the nose strongly.

Pleasant: While intense, it’s liked in moderation—used for pain relief.

Pungent: Almost harsh—needs dilution; high olfactory intensity.

Appendix D 94: Cumin

**Inferred Extraction Method: Steam Distillation of dried seeds (Cuminum cyminum)**  
 **Overview: Cumin oil is deeply spicy, musky, and warming. Rich in cuminaldehyde and cymene.**

**Descriptors and Justifications:**

Spicy: Main aromatic category—sharp, warm intensity.

Pungent: Bold and assertive—used in low concentrations in perfumery.

Herbaceous: Underlying dry grassiness from cymene.

Pleasant: Familiar in culinary use, though polarizing in raw form.

Warm: Deep base character, enhances blends for energy and warmth.

Appendix D 95: Bay (West Indian)

**Inferred Extraction Method: Steam Distillation of leaves (Pimenta racemosa)**  
 **Overview: West Indian bay is spicy, woody, and slightly sweet. Contains eugenol, chavicol, and myrcene.**

**Descriptors and Justifications:**

Spicy: Dominated by eugenol—warm and sharp like clove.

Woody: Resinous, drydown mimics aged wood and bark.

Sweet: Mild sweetness that emerges beneath spicy top.

Pleasant: Often used in barbershop scents and tonics.

Aromatic: Rich and layered with complex phenolic components.

**References for Listed Descriptors:**

Sellar, W. The Directory of Essential Oils (Vermilion, 2007) – foundational for oil sourcing and general aroma profiles.

Baser, K.H.C. & Buchbauer, G. Handbook of Essential Oils: Science, Technology, and Applications (CRC Press, 2010).

Edris, A.E. “Pharmaceutical and therapeutic potentials of essential oils and their individual volatile constituents: a review.” Phytotherapy Research, 21(4), 2007.

Ramos et al. “Chemical Composition and Biological Activities of Essential Oils: A Review.” Molecules, 2021.

Ferhat et al. “Comparative study of different extraction methods of essential oil from citrus fruits.” Journal of Food Engineering, 2007.

Lis-Balchin, M. Aromatherapy Science: A Guide for Healthcare Professionals (Pharmaceutical Press, 2006).

Orav et al. “Essential oil composition of various Mentha species.” Proceedings of the Estonian Academy of Sciences, Chemistry, 2004.

Lawless, J. The Encyclopedia of Essential Oils (HarperCollins, 2014).

PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) – used for verifying major compound identities and properties.

The Good Scents Company (<http://www.thegoodscentscompany.com>) – industry database for aromatic compound profiles.

Essential Oil Safety – Tisserand & Young (2nd Edition, 2014) – to cross-check safety and compound-specific profiles.

Sigma-Aldrich Essential Oil MSDS Sheets – used for composition patterns.

AromaWeb (<https://www.aromaweb.com>) – general aroma descriptions and consumer-level descriptor matching.

Plant Therapy Blog & Edens Garden Blog – used for distillation methods and comparative insights.

Robertet Group & IFF Scent Ingredient Catalogs – industry-grade scent notes and usage classification.