IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Scent Synthesis: An AI-Driven Framework For Creating Novel Fragrance Profiles

Palash Halder

Department of Computer Science & Engineering, IIMT College of Engineering

Role: Independent Researcher, Founder – Valoure Product PVT. LTD.

Abstract

The world of perfumery – a classically handmade industry – is discreetly slinking into the 21st century with artificial intelligence (AI). We introduce in this paper "ScentSynth", a new AI-based framework targeting the generation of fragrances. By employing olfactory databases, molecular descriptors and consumer taste patterns, the model combines supervised and generative machine learning methods. The goal is to drastically save on time and cost to formulate scent as well as to improve creativity and customization. We advance a naturally evolving technical map of AI in perfumery, build a detailed breakdown of the system architecture, as well as validate the feasibility of the system by deployment and evaluation.

Keywords: Artificial Intelligence, Perfumery, Machine Learning, Fragrance Formulation, Scent Prediction, Molecular Descriptors

1. Introduction

Human beings have created perfume as an art form for thousands of years, based on experience, creativity, and the perfumer's gut instincts. But, thanks to breakthroughs in AI and big data, technology is now on the verge of complementing — and sometimes even co-creating with — noses. This paper presents a novel machine learning framework, ScentSynth, that is capable of the analysis of molecular data as well as autonomously create new fragrance blends, utilizing AI models including XGBoost, Graph Neural Networks(GNNs) and Variational Autoencoders(VAEs). ScentSynth seeks to enhance the creativity of perfume creation using computational tools based on perceptions.

2. Related Work

There have been recent collaboration like IBM and Firmenich [4] which have been leading the way in this AI assisted scent design. Keller et al. [6] showed that DL based models can deliver better olfactory predictions than traditional QSAR ones. Tools like RDKit [2] for cheminformatics and The GoodScents Company [3] database have been used widely for dataset pre-processing in scent-based machine learning studies. These first efforts pave the way for the incorporation of the molecular aspects in predictive odor modeling.

3. Methodology

ScentSynth framework is based on a 5-step routine:

1. Data Collection: SMILES strings and molecular descriptors were collected from public databases such as GoodScents [3].
2. Label Normalization: The inconsistent scent labels were normalized using BERT (Bidirectional Encoder Representation from Transformers).
3. Feature Engineering: Selected molecular fingerprints using RDKit [2] such as topological polar surface area, LogP, and molecular weight.
4. Model Training:
    - XGBoost: A powerful gradient-boosted tree algorithm on (source of DNA sequences) tabular molecular data.
    - GNN (Graph Neural Network) - A type of neural network able to learn the molecular graph structure directly.
    - VAE (Variational Autoencoder): Employed in the generation of novel molecules.
5. Evaluation: Models were evaluated based on the metrics of Tanimoto similarity and Synthetic Accessibility.

4. Results and Discussion

Three models were trained and tested with a curated dataset. The performance is tabulated in Table 1.

Table 1 : Model Performance

Model	Accuracy	F1-Score	Training Time
Random Forest	81%	0.78	30s
XGBoost	85%	0.81	12s
GNN	88%	0.84	55s

A VAE-based generator was also evaluated on molecule generation, with TD score of 0.81 and SA score of 0.72. Further, in a double double-blind test with 5 users 4 judged the AI-generated composition more favorably than the baseline composition.

5. System Architecture

The following steps represent the architecture of the ScentSynth system:

Data Ingestion (SMILES + Descriptors)

↓

NLP Normalization **(**using BERT for label standardization**)**

↓

Feature Extraction **(**with RDKit for molecular descriptors**)**

↓

Model Training **(**XGBoost, GNN, and VAE models**)**

↓

Prediction Interface (Flask + Streamlit-based UI)

6. Case Study

A prototype app was also developed using Flask and Python Streamlit. And the tool allows users to choose end descriptors (think "fresh" or "woody" or "ambery") and get on-the-fly formula suggestions.

Example Formulation:

Base Notes: Iso E Super (40%) Amberlyn (20%)

Heart Note: Hedione (15%), Linalool (10%)

Top Note: Limonene (10%) 2nd: Aldehyde C12 (5%)

7. Sample Code

```
from rdkit import Chem
from rdkit.Chem import Descriptors, MACCSkeys
from xgboost import XGBClassifier
import pandas as pd

smile = "CC(=O)OC1=CC=CC=C1C(=O)O"
mol = Chem.MolFromSmiles(smile)
feature = [Descriptors.MolWt(mol), Descriptors.TPSA(mol)]
model = xgb.Booster()
model.load_model("scent_model.json")
prediction = model.predict(xgb.DMatrix([features]))
print("Predicted scent class:", prediction)
```

8. Challenges and Limitations

- Scent Descriptor Inconsistencies: Descriptions of fragrance differ between datasets.
- Data disadvantage: The public fragrance data have been very limited and not standardized.
- Real-Time Scent Rendering: The hardware support for digital consumption of scents has not fully developed.

9. Applications and Societal Impact

The fields where AI-led perfumery could transform various industries:

- E-commerce: Personalized scent recommendations.
- Aromatherapy: AI-generated mix for mood or sleep.
- Applications: culture-oriented machine of smart attar blending for India.
- Education: Teaching the science of smell with immersive simulations.

10. Future Work

Future developments will include:

- Explainable AI applied to the interpretation of scent predictions.
- Multimodal learning with text, images, and chemistry.
- Biometric feedback integration.
- Collaborative AI dashboards for perfumers.

11. Author Contribution and Personalization

The authors implemented the ScentSynth pipeline, developed the prototype application, carried out a small-scale blind study, and trained the machine learning models. Indian-specific words associated with fragrance such as “attar” and “gulab” were incorporated in the data to be more in-line with cultural context.

12. Conclusion

ScentSynth is proof of concept that AI can help, as an aid, perfumery as a creative craft. With better quality data and better algorithms, AI can act as a true partner in fragrance innovation, the harbinger of speed, the friend of cost efficiency and the deliverer of personalization at scale.

13. Acknowledgment

I would like to express my deepest gratitude to my father and mother for being my first teachers in perfumery. Their encouragement, support and unfailing guidance have helped me to form this path. From reviling me of knowledge of fragrance basics to encouraging me at each and every obstacles, their involvement has been the cornerstone of this work. I am thankful to the faculty at IIMT College of Engineering, Grater Noida for the academic support, and to all those who contributed to the construction and the validation of this work.

14. Plagiarism Declaration

I Palash Halder, hereby declare that the above stated project work and findings are original to the best of my knowledge and belief and have not been submitted for the award of any degree, diploma or any other publication since any date Someone AI or so. I acknowledge that external sources must be appropriately cited and referenced.

15. References

[1] A. Aspuru-Guzik et al., "The Matter Simulation (MolGAN): Transfer Learning for Multimodal Generative Model," arXiv preprint, arXiv:1805.11973, 2018.

[2] RDKit, "Open-source cheminformatics," [Online]. Available: https://www.rdkit.org.

[3] The GoodScents Company, [Online]. Available: https://www.thegoodscentscompany.com.

[4] IBM and Firmenich, "The AI Scent Designer Project," IBM Research Blog, Jan. 2020.

[5] E. Gifford et al., "Machine learning for olfaction," Journal of Sensory Studies, Feb. 2022.

[6] A. Keller, L. B. B. Gerkin, T. Paschke, and F. K. Pyrkosz, "Predicting human olfactory perception from chemical features of odor molecules," Science, vol. 355, pp. 820–826, 2017.