Nat. Volatiles & Essent. Oils, 2021; 8(6): 317-325

NVEO
Natural Volatiles & Essential Oils

Classification And Prediction Of Essential Oils Using Mobile Nets

Dr. Machhindranath Manikrao Dhane[1], K.Praveen Kumar[2], Dr. Sandeep Rout[3], Subbulakshmi Packirisamy[4], Dr. Syed Mohd Fazal Ul Haque[5], Dr. Bonthu Kotaiah[6]

[1]Assistant Professor of Mathematics Government First Grade College, Yelahanka, Bengaluru-560064,

[2]Assistant Professor, Department of Information Technology, Kakatiya Institute of Technology and Science, warangal, Telangana

[3]Assistant Professor, Faculty of Agriculture, Sri Sri University Cuttack Odisha-754006

[4]Department of pharmacology, Meenakshi ammal dental college and hospital, Alapakkam main road, maduravoyal, Chennai 95

[5] Assistant Professor, Department of Polytechnic, Maulana Azad National Urdu University, Gachibowli Hyderabad Telangana

[6] Assistant Professor, Department of Computer Science and Information Technology, Maulana Azad National Urdu (A Central) University, Gachibowli, Hyderabad, Telangana – 500032

Abstract:

Due to its demonstrated effectiveness in handling massive data, deep learning techniques have become a significant development in various research fields. They can conduct non-linear processes and complex interactions, including biological science. Convolutional Neural Networks (CNN) were employed in this study to detect and predict the biological activities of essential oil-producing plants. The same Essential Oils (EO) dataset has also been subjected to a different class of machine learning methods, the Multiclass Neural Network. This is being done to assess the suggested CNN model's performance objectively. According to CNN and MNN's test results, the accuracy of the testing method is 98% and 81%, respectively. As a result, the EO-containing plants' bioactivities may now be classified and predicted using CNN. According to the final prediction method, the total accuracy is claimed to be around 97 per cent. A valuable model for predicting the bioactivities of EO-producing plants, at least in Egypt, is the deep learning model described here.

Keywords: CNN, MNN,Essential Oils, Deep Learning.

Introduction

A broad variety of modern life challenges, including chemical engineering, water treatment, and biological domains like genomic and proteomic investigations, have been penetrated by artificial intelligence (AI) in recent years, which is characterized by complex and nonlinear processes. With its shown capacity to take raw characteristics derived from large data sets, such as those generated by genomics, chemistry, and pharmaceutical labs. It is widely agreed that deep learning is a very promising field within artificial intelligence (AI). Deep learning is also one of the most promising sciences, since it has been proved to accept input features acquired from other challenges. Analyzing this data yields patterns that may be deduced and training process-based models that can be used to make forecasts.

Flowers, leaves, and bark, to mention a few elements of the aromatic plants that contain essential oils (EOs), are physiologically beneficial organic chemicals that are isolated from the plants' various sections. These natural materials are commonly employed in complementary medicine because of the vast variety of biological activity that they possess (CAM). Using natural alternatives to replace inorganic chemistry has been a popular subject of discussion in the contemporary biological field of study. This is due to the fact that artificial chemical compounds, when employed in the medical, pharmaceutical, cosmetic, food, and beverage industries, may have harmful

Bioactivity class Classifiers	True positive		True negative		False positive		False negative	
	MNN	CNN	MNN	CNN	MNN	CNN	MNN	CNN
Antiviral	50%	97.7%	94.4%	98.2%	5.6%	1.8%	50%	2.3%
Antiwormal	100%	100%	98.3%	98.5%	1.7%	1.5%	0	0%
Anti-inflammatory	90%	100%	80%	100%	20%	0%	10%	0%
Anticancer	85.7%	94%	96.2%	93.3%	3.8%	6.7%	14.3%	6%
Antioxidant	75.9%	98.4%	80.6%	98.6%	19.4%	1.4%	24.1%	1.6%
Antimicrobial	78.8%	97.4%	66.7%	97.3%	33.3%	2.7%	21.2%	2.6%
Antifungal	81.3%	98%	86.4%	99.2%	13.6%	0.8%	18.8%	2%
Cytotoxic activity	58.3%	96.6%	93.8%	98.7%	6.3%	1.3%	41.7%	3.4%

Table 1. In the testing phase, the MNN and CNN algorithms' confusion matrix values.

impacts on human health and the environment As a result, contemporary research focuses on finding alternative natural goods, such as essential oils that have a higher potential to be more biocompatible with live human organs while inflicting less harm than existing items.

According to basic rules of thumb, the bioactivity of an EO-producing plant is defined by its chemical structure and its EOs content. Both factors impact the overall bioactivity of the plant in question. EOs,

on the other hand, are composed of a variety of low nuclear weight natural blends that are capable of complete organic motility and may be combined in a variety of ways. Based on their structural features, these dynamic mixtures may be divided into several distinct groups (for example, hydrocarbons, oxygenated mixture &sulphur or perhaps nitrogen). We can now explain the biological role of each EO3 thanks to the pools identified in these combinations.

The essential oils' many biological effects include antiseptic & antimicrobial, antifungal, anticancer, antiviral, and anti-inflammatory properties. A plant's chemical composition may differ from one variety to the next based on its geographical area, agricultural conditions, climate or seasonal changes. An EO's bioactivity cannot be reliably assigned to anyone one molecule in the EO's mixture, as should be noted. As this study's findings demonstrate, the precise connections between the EO's cocktail and its biological effects are pretty non-linear when investigated across diverse pools of chemical structures. Predicting biological effects from data with such a wide variety of structural properties may be difficult using typical methodologies. This is why. Thus, they were developing an artificial neural network model that could identify and predict the biological activities of EO-producing plants

using variations in their chemical composition. This would save both time and money because in-vitro experiments, which would otherwise be required, would not be necessary because of this.
It has been claimed that machine learning (ML) approaches, notably Artificial Neural Networks (ANN), may help solve a number of biological challenges in recent decades. Nodes and neurons (i.e. layers) are the basic building blocks of an artificial neural network (ANN), which is a numerical representation of a given structure. Hidden neurons in each layer change the values of the inputs and send the resulting outputs to the appropriate neurons. Many factors in a fully connected neural network need extensive processing inside the network layers to produce the desired output, as biological information has grown in recent years. Large data sets with a high degree of complexity and structure have shown the effectiveness of deep learning methods. With millions of variables (pixels) grouped into clearly defined objects in an image, deep learning methods are often used in image processing6. Some text classification problems, like those requiring DNA sequence categorization, may be solved using deep learning approaches, which are no longer exclusive to the field of image processing. It seems that deep learning might be an efficient strategy for dealing with the complex interactions between different chemicals and biological processes that make up substances in the world around us.
Convolutional Neural Networks are a powerful deep learning model (CNN). Convolutional and pooling layers differentiate the CNN. These layers utilize filters to restrict the input data range, allowing for discovering significant or memorable locations within it To build a CNN you need three layers: an initial input layer, an

Bioactivity class Classifiers	True positive		True negative		False positive		False negative	
	MNN	CNN	MNN	CNN	MNN	CNN	MNN	CNN
Antiviral	50%	97.7%	94.4%	98.2%	5.6%	1.8%	50%	2.3%
Antiwormal	100%	100%	98.3%	98.5%	1.7%	1.5%	0	0%
Anti-inflammatory	90%	100%	80%	100%	20%	0%	10%	0%
Anticancer	85.7%	94%	96.2%	93.3%	3.8%	6.7%	14.3%	6%
Antioxidant	75.9%	98.4%	80.6%	98.6%	19.4%	1.4%	24.1%	1.6%
Antimicrobial	78.8%	97.4%	66.7%	97.3%	33.3%	2.7%	21.2%	2.6%
Antifungal	81.3%	98%	86.4%	99.2%	13.6%	0.8%	18.8%	2%
Cytotoxic activity	58.3%	96.6%	93.8%	98.7%	6.3%	1.3%	41.7%	3.4%

Table 2. The MNN and CNN algorithms' accuracy and relevance measures are now being evaluated.

Bioactivity class Classifiers	True positive	True negative	False positive	False negative
Antiviral	98%	96.3%	3.7%	2%
Antiwormal	100%	98.3%	1.7%	0%
Anti-inflammatory	100%	100%	0%	0%
Anticancer	97%	94.7%	5.3%	3%
Antioxidant	100%	95%	5%	0%
Antimicrobial	94%	93.5%	6.5%	6%
Antifungal	100%	96.6%	3.4%	0%
Cytotoxic activity	94%	93.1%	6.9%	6%

Table 3. The confusion matrix values for the predicting model.

intermediate convolution layer, and a final ReLU (or Rectified Linear Unit) layer. There are three steps: max(0, x), max pooling, and the final output layer. Mathematically it is represented as max(0, x) (any value less than 0 is transformed to 0, while any positive value passes intact in other words (i.e. a fully-connected neural network layer, which makes the output based on the activation function).

The purpose of this work is to categorise and predict the biological activities of essential oil-producing plants based on their EOs content using an experimental case study. Classification is conducted using different algorithms to assess the utility of machine and deep learning approaches: Multiclass Neural Network (MNN) and Convolutional Neural Network (CNN). The practical method is used to construct a

model for forecasting the biological activity of EOs-producing examples planted in various countries. The sections that follow outline the research skills employed in this study: To begin, we will discuss the findings of two algorithms applied to the essential oils dataset; MNN is a machine learning method, while CNN is a deep learning technique. These findings were compared to both existing and newly created untested datasets. Finally, the report outlines the methodologies used to conduct the research. Confusion matrix validation is the best method for evaluating classification performance. For CNN and MNN, the classification matrix is shown in Table 1. True Positive, False Positive, True Negative, and False Negative are all binary categories. Different measures for measuring classification performance based on confusion matrix values10 include accuracy, precision, recall, and F1 score.

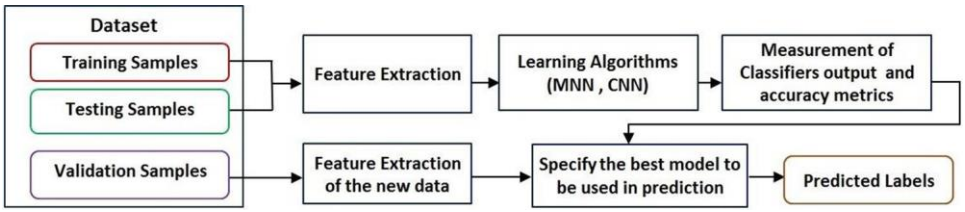

Fig 1. The suggested methodology phases are shown in a detailed flowchart.

Methodology

Using CNN algorithms to develop a prediction model for the bioactivity of EOs.

The classification technique determined that the CNN algorithm maintains a high accuracy throughout the training and testing phases. This is because it can process massive amounts of data while focusing on the most critical components of the dataset. A model for anticipating biological activity has been developed due to the CNN-suggested technique. Forecasting the bioactivity of previously unknown essential oils has been predicted to be around 97 per cent accurate. The suggested CNN prediction model's final metric values and discriminating results are shown in Tables 2 and 3 and Fig. 2.

In order to solve medical and biological issues, it is vital that artificial intelligence, namely machine learning and deep learning algorithms, be used. It is generally known that essential oils (EOs) have therapeutic effects and biological processes that protect against a number of pathogens and cancers. Due to the absence of complete understanding on their underlying chemical variety and function, it is difficult to demonstrate their consistent activity.

To speed up in-vitro operations and create reliable predictions regarding experimental results, several EO researchers have used AI. Classification of EOs and identification of those with a high chance of preventing the growth of all bacterial strains was accomplished using an unsupervised machine learning technique. Additionally, a machine learning-based binary classification technique was used to show the antibacterial activity of essential oils isolated from a range of Mediterranean plants. The

antibacterial activity of 49 EOs against four distinct illnesses was predicted using a fast artificial neural network (FNN). The FNN algorithm accurately predicted more than 70% of antibacterial activity within a 10 mm error range.

Due to the variety and immense size of aromatic plant species as a consequence of their climatic environment, it may be possible to enhance the synthesis of highly concentrated secondary metabolites. As a consequence, there is an urgent need to strengthen the scientific basis for aromatic species via the use of cutting-edge artificial intelligence techniques such as machine and deep learning. Essential oils are used in this particular experiment. There were peer-reviewed publications and in vitro tests used to get the data. To treat a particular condition, it may be necessary to determine the essential oil's chemical makeup or to provide one or more EOs (e.g. bacteria, viruses, or cancer). Because of this, researchers have turned to machine learning and deep learning technologies to explore the biological activities of certain plants that produce essential oils. The training and testing methodologies make use of a unique collection of plants that generate essential oils that the authors compiled and updated from peer-reviewed scientific papers. A random sample of twelve hundred (120) plants was used to construct the dataset. This collection has 573 chemical compounds in total. In the recommended experimental setting, the biological actions of essential oils are classified into eight categories (anticancer, antioxidant, antimicrobial, antifungal, antiviral, anti-wormal, anti-inflammatory, and cytotoxic activities). EOs may be identified via the use of two supervised learning-based classification techniques. The Multiclass Neural Network (MNN) architecture is used to construct the initial model (i.e., the Multiclass NN module in Azure). Open-source Azure Machine Learning studio56 is used to run the second model based on Convolutional Neural Networks (CNNs). An all-encompassing flowchart of the suggested method is shown in Figure 3. About 68,760 chemical component % concentration values correspond to the two recommended supervised learning techniques in the EO dataset, trained using the two algorithms. Approximately 60% of the information is utilised to train MNN and CNN algorithms, with the other 40% for testing and foretelling purposes. This implies that there are 45,840, 11,460, and 11,460 people in the training, testing, and validation processes, respectively. EOs dataset is shown in Azure format in Figure 1.

Eos name Score probabilities	Antiviral	Antivormal	Anti-inflammatory	Anticancer	Antioxidant	Antimicrobial	Antifungal	Cytotoxic activity
Citrus deliciosa var. *tangarina*—fina clementine—peels	0.9888	0.0014	0.0032	0.0001	0.9854	0.9981	0.9983	0.9966
Citrus deliciosa var. *tangarina*—fina clementine—leaves	0.9999	0.9986	0.0005	0.0003	0.9992	0.9984	0.9983	0.0013
Citrus deliciosa var. *tangarina*—Nour Clementine—peels	0.9465	0.0022	0.3243	0.0002	0.9894	0.9854	0.9684	0.7711
Citrus deliciosa var. *tangarina*—Nour Clementine—leaves	0.9999	0.0013	0.0641	0.0002	0.9995	0.9997	0.9996	0.0003
Citrus deliciosa var. *tangarina*—Spinosa Clementine—peels	0.8294	0.0795	0.0001	0.0003	0.8291	0.9294	0.9993	0.6654
Citrus deliciosa var. *tangarina*—Spinosa Clementine—leaves	0.9977	0.0015	0.0320	0.0084	0.9968	0.9982	0.9961	0.0049
Citrus deliciosa var. *tangarina*—Thornless Clementine—peels	0.9991	0.0007	0.0931	0.0034	0.9987	0.9999	0.998	0.2611
Citrus deliciosa var. *tangarina*—Thornless Clementine—leaves	0.9999	0.0020	0.0134	0.0082	0.9999	0.9998	0.9996	0.0007
Callistemon comboynensis	0.9975	0.0037	0.1375	0.0011	0.9947	0.9955	0.9945	0.0031
Cupressus sempervirens L *Semperuirens* L	0.8982	0.0011	0.0004	0.0012	0.9947	0.9967	0.0002	0.0035
Cuminum cyminum	0.8923	0.9747	0.9384	0.0005	0.9297	0.9997	0.7909	0.0001
Ocimum basilicum L. (sinai)	0.0004	0.0005	0.0883	0.9977	0.9991	0.9990	0.9089	0.0003
Tagetes minuta L	0.4107	0.0011	0.0034	0.0253	0.5146	0.7167	0.6179	0.1418
Achillea fragrantissima	0.0004	0.0003	0.0003	0.0007	0.0008	0.0013	0.0001	0.9985
Pluchea dioscoridis	0.0052	0.7622	0.0054	0.9863	0.9931	0.9991	0.8932	0.9915
Myrtus communis—leaves	0.0022	0.9887	0.0554	0.0018	0.0003	0.0021	0.0023	0.0028
Myrtus communis—fruits	0.0007	0.6754	0.0001	0.0005	0.0013	0.0023	0.0011	0.0004
Eugenia supraxillaris—leaves	0.001	0.9998	0.3451	0.0075	0.0050	0.0075	0.0065	0.0032
Eugenia supraxillaris—fruits	0.9955	0.7865	0.0003	0.0003	0.9995	0.9934	0.0273	0.8539

According to **Fig 2** the CNN model used in this experiment consists of two convolutional layers, two pooling layers, and one hidden layer. [120 573] denotes "the number of essential oils and their chemical components, respectively," which serves as input for the suggested CNN. A convolutional layer convexes the input matrix in the subsequent layers. Each output node is the filter's convolution operation resulting in the convolution layer. The pooling layer is used to minimise overfitting by compressing and generalising data. This experiment employs tiny split sections of the input matrix to extract the maximum value using local maxpooling.

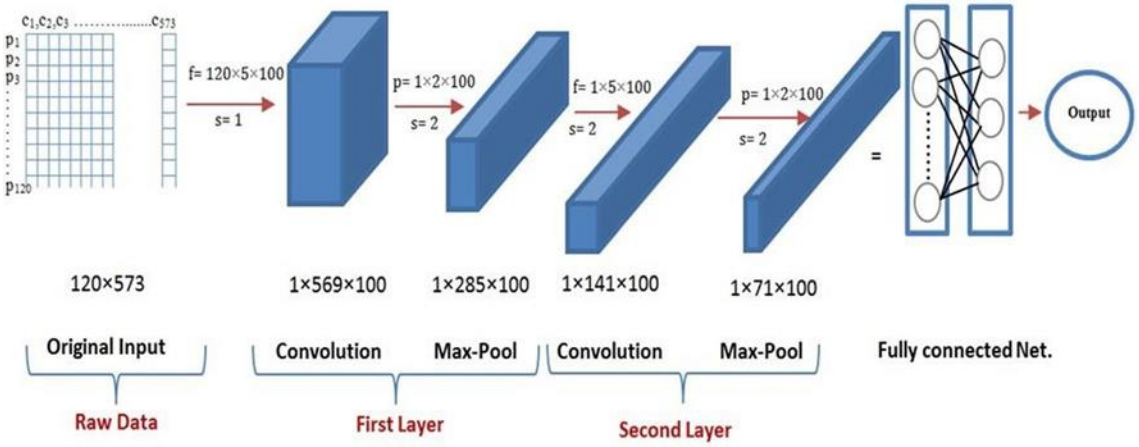

Fig3. The New Conceptualization Proposed Analysis of essential oil data begins with CNN using an original input matrix. The network's underlying structure consists of two convolution and pooling layers. Sigmoid activation is used in a fully connected feed forward neural network to select the correct output from these layers.

Layers	Filter size	Stride	Output size	Output Dense
Input	–	–	120 × 573	68,760
1st convolution layer	120 × 5 × 100	1	1 × 569 × 100	56,900
1st max pooling layer	1 × 2 × 100	2	1 × 285 × 100	28,500
2nd convolution layer	1 × 5 × 100	2	1 × 141 × 100	14,100
2nd max pooling layer	1 × 2 × 100	2	1 × 71 × 100	7100

Table.4 Hyper-parameter values for CNN layers are shown in the following table.

Last is a "sigmoid" active layer that comes after the convolution and pooling layers. Since each node's score is translated to a probability value between 0 and 1, sigmoid functions are often utilised while developing a CNN. It is possible to categorise the input into several subcategories (Suppl Information).

Conclusion

With the use of deep learning, researchers want to construct a computer model able to categorize and predict EOs' bioactive properties without the need for in vitro testing. Because of their efficacy as anti-septics and anti-microbials, antifungals, antioxidants, anti-tumours, antivirals & anti-inflammatories, EOs have drawn a lot of attention in the health care business.. However, a significant challenge is tracing the chemical components of the EO to their biological effects using conventional in-vitro research methods. This experiment categorized and analyzed 120 Egyptian essential oils using two classification methods. Both machine learning and deep learning were tested in this experiment multiclass Neural Networks and Artificial Neural Networks are used for supervised learning. CNN has a two-to-one accuracy advantage over MNN (98.13 percent) (81.88 percent).

References

1. Dhifi, W ., B ellili, S., Sabrine Jazi, N. B. W. M. Essential oils ’ chemical characterization and investigation of some biological activi- ties . Medicines 3, 1–16 (2020).

2. Jones, D. T. Setting the standards for machine learning in in biology. Nat. Rev. Mol. Cell Biol. 20, 659–660 (2019).

3. Tharwat,A. Applied computing and information classification assessment methods. Appl. Comput. Inform. 1–13. https://doi. org/10.1016/j.aci.2018.08.003 (2018).

4. Ibrahim, T.A., ., El-hela, A. A., El-hefnawy, H. M. & Al-taweel, A. M. Chemical composition and antimicrobial activities of essential oils oils of some coniferous plants cultivated in Egypt. Iran. J. Pharm. Res.16,328-337(2017)

5. Muniz, I. et al. Classification of food vegetable oils by fluorimetry and artificial neural networks. Food Control 47, 86–91 (2015).

6. El-ghorab, A. H., Ramadan, M. M., El-moez, S. I. A., April, M. & April, M. Research journal of pharmaceutical, biological and chemical sciences essential oil, antioxidant, antimicrobial and anticancer activities of Egyptian Pluchea dioscoridis extract. .Res. J. Pharm. Biol. Chem. Sci. 6, 1255–1265 (2015).

7. Daynac, M., Cortes-cabrera, A. & Prieto, J. M. Application of artificial intelligence to the prediction of the antimicrobial activity of essential oils. Evidence-Based Complement. Altern. Med. 2015 (2015).

8. Badwy, M. E. I. & Abdelgaleil, S. A. M. Composition and antimicrobial activity of essential oils isolated from Egyptian plants against plant pathogenic bacteria and fungi Ind. Crop. Prod. 52, 776–782 (2014).

9. Abdelhady, M. I. & Aly, H. A. H. Original article antioxidant and antimicrobial activities of Callistemon comboynensis essential oil. Free Radic. Antioxid. 2, 37–41 (2012).