IOSUD – „Dunărea de Jos" University of Galați

Doctoral School of Mechanical and Industrial Engineering

PhD THESIS

SUMMARY

MULTIVARIATE MATHEMATICAL METHODS APPLIED FOR ILLICIT DRUG DETECTION

PhD student,

Iulia-Florentina DARIE

Scientific adviser,

Prof univ.dr. Mirela PRAISLER

Seria I4: Inginerie industrială Nr. 95

GALAȚI

2024

IOSUD – „Dunărea de Jos" University of Galați

Doctoral School of Mechanical and Industrial Engineering

PhD THESIS

SUMMARY

MULTIVARIATE MATHEMATICAL METHODS APPLIED FOR ILLICIT DRUG DETECTION

PhD student,

Iulia-Florentina DARIE

Chairman of the scientific committee	Professor, PhD, Eng. Cătălin FETECĂU
Scientific adviser	Professor, PhD Mirela PRAISLER
Scientific referees	Professor, PhD, Eng. Remus ZĂGAN
	Professor, PhD, Eng. Gheorghe NAGÎȚ
	Associate professor, PhD Steluța GOSAV

Series I4: Industrial Engineering No. 95

GALAȚI

2024

The series of doctoral theses publicly defended in UDJG starting with October 1, 2013 are:

Fundamental field ENGINEERING SCIENCES

Series I 1**:** **Biotechnology**
Series I 2**:** **Computers and Information Technology**
Series I 3: **Electrical Engineering**
Series I 4**:** **Industrial Engineering**
Series I 5: **Materials Engineering**
Series I 6**:** **Mechanical Engineering**
Series I 7**:** **Food Engineering**
Series I 8: **Systems Engineering**
Series I 9**:** **Engineering and Management in Agriculture and Rural Development**

Fundamental field SOCIAL SCIENCES

Series E 1: **Economies**
Series E 2: **Management**
Seria E 3: **Marketing**
SSEF Series: **Science of Sport and Physical Education**
SJ Series: **Law**

Fundamental field HUMANITIES

Series U 1: **Philology - English**
Series U 2: **Philology - Romanian**
Series U 3**:** **History**
Series U 4: **Philology - French**

Fundamental field MATHEMATICS AND NATURAL SCIENCES

Series C: **Chemistry**

Fundamental field BIOMEDICAL SCIENCES

Series M: **Medicine**
Series F: **Pharmacy**

Table of contents

Introduction

The contemporary world is increasingly facing a wide variety of organised crime, a special attention being paid lately to the drug phenomenon, taking into account everything that it entails: manufacturing, transportation, distribution and consumption. Drug trafficking is a well-known danger for the population, especially for some age groups - teenagers and youth in general, which represents a huge threat regarding public health, a fact pertaining to national security.

Romania has quite easily shifted from a status of transition country to a country of drug consumers. At present, there is a variety of ways by means of which illicit drugs can enter the country (by naval, air or terrestrial transportation). Aside from money laundering, which trafficking drugs entails, the health system is also under pressure as most of the consumers develop various medical conditions.

The increasing danger of drug consumption is mentioned more and more in global reports that are trying to alert, caution and mobilise the fight against this terrible calamity. The 2023 European Drug Report made by the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) points out the fact that record quantities of illegal drugs are being seized. It emphasises an immediate need for data collecting pertaining to toxicology, which would concretely reveal the effects of the new synthetic substances [1].

The increase of drug trafficking is due not only to the use of modern technology in their manufacturing, but also by the innovative ways of smuggling. Thus, the fight against drug trafficking means using modern detection methods.

Therefore, it is imperative to detect not only the well-known drugs of abuse, but also any new substance that has a similar molecular structure. This is very useful: once a new substance is included on the list of illicit substances pertaining to the adjacent law, the clandestine laboratories immediately release new compounds on the black market in order to avoid legal consequences. These new substances are not among the banned ones. However, having a molecular structure similar to a drug of abuse, they may have a similar biological (pharmacological and/or toxicological) activity at given concentrations.

Summary and structure of the scientific work

This PhD thesis, entitled “Multivariate mathematical methods applied for illicit drug detection”, comprises five chapters alongside an introduction to the studied topic.

The paper starts with an introduction, which presents the outlines on drug trafficking and emphasises the necessity of using modern detection methods within the context of an ongoing expansion of the black market.

Chapter I comprises an outline of the analysed substances of abuse belonging to 2C-x, DOx and NBOMe classes of drugs of abuse. This chapter highlights the results published in different studies and reports on the targeted substances, their IUPAC and other names, alongside their corresponding molecular structures and formulas.

Chapter II describes the use of the ATR-FTIR spectral method for the identification and the characterisation of the targeted substances. Within this chapter, the basics of the ATR-FTIR method, the content of their ATR-FTIR spectra and their vibrational analysis are presented.

Chapter III, entitled “Mathematical methods and approaches applied for the analysis and the identification of class membership of the targeted drugs of abuse”, presents the fundamental mathematical concepts that lie at the basis of the molecular descriptors, the statistical measures, the exploratory analysis methods (PCA, ICA, Autoencoders), the machine learning methods (SVM, RF, GB, XGBoost, KNN), as well as the performance indicators used within this work.

Chapter IV presents the author’s personal scientific contributions regarding the analysis and the recognition of the class identity of representative chemical compounds belonging to the 2C-x, DOx și NBOMe classes of drugs of abuse, based on specific analytical techniques and the application of a variety of multivariate mathematical methods. Within this chapter, we are presenting not only the results obtained based on the ATR-FTIR spectra of these compounds, but also those obtained based on the molecular descriptors and the toxicity parameters calculated for these drugs of abuse.

Chapter V comprises the general conclusions that sum up the research presented within this work, by emphasising the original results obtained, accompanied by future potential directions for research and development.

The thesis ends with the list of published scientific articles, the list of scientific papers presented at national and international conferences, the participation within projects and also the won awards.

Motivation of choosing the research theme

Nowadays, the challenge of drug consumption is undeniable. This poses major issues to the current society and, unfortunately, their negative influence is reflected to a large extent on increasingly younger age groups. High school students and university students have become an easy target, their psycho-social development being seriously threatened. the present scientific work in the fast detection of drugs of abuse in order to annihilate the devastating effects which the consumption of these substances can lead to is justifiable.

The identification and the characterisation of the new synthetic drugs of abuse by means of the classic laboratory equipment takes time and implies large analysis costs. On the other hand, in the case of the identification of a new illegal substance, it is important that the decision whether to confiscate it (or not) to be taken as fast as possible, for example in order to avoid the detention of innocent

people. That is why it is necessary to develop computerised applications in order to allow a fast, *in situ*, detection, therefore performed with portable devices. For these devices to be used by persons with no advanced training in analytical chemistry and toxicology, such as policemen or customs workers, it is imperative to develop computerised applications that would automatically process the analytical information and indicate which decision must be taken regarding confiscation.

These computerised applications presuppose the implementation of certain specific techniques which rely on a complex mathematical foundation. Thus, one can deduce the important role of mathematics as a basis for the development of automatic drug detection.

The starting point of this thesis is the possibility of automatically identifying the 2C-x, DOx and NBOMe drugs of abuse by means of a variety of multivariate mathematical methods. More precisely, the research focused on creating models enabling the rapid detection of the targeted drugs of abuse, in this way helping the fight against drug trafficking. Also, the characterisation and the evaluation of the grouping tendency of the 2C-x, DOx and NBOMe drugs of abuse was pursued, in order to develop models predicting their biological activity.

Pursued research objectives

Within the present thesis, the main research objectives pursued were:

- carrying out of a literature research regarding the current state of the art, materialised in categorising the existing data in the literature pertaining to the field of study regarding the main drugs of abuse belonging to the 2C-x, DOx and NBOMe classes;
- describing the ATR-FTIR spectral method used for the characterisation and the identification of drugs of abuse;
- presenting the mathematical methods and approaches that were applied for the analysis and the identification of the class membership of the targeted compounds;
- performing the analysis and the recognition of the drug class identity based on specific multivariate techniques and the application of various mathematical methods through:
  - developing machine learning systems able to detect drugs of abuse belonging to the targeted classes, based on their ATR-FTIR spectra;
  - performing the vibrational analysis of some drugs of abuse belonging to the targeted classes based on their ATR-FTIR spectra;
  - characterising and evaluating the clustering tendency of representative drugs of abuse from the 2C-x, DOx și NBOMe classes based on their molecular descriptors and toxicity parameters.

Chapter I. Presentation of the analysed 2C-x, DOx and NBOMe compounds

1.1. Introduction

The drug market has seen a significant expansion worldwide in recent years. Over the years, more and more substances have been placed in certain risk groups, their consumption being prohibited [2]. However, in recent years, the number of drugs of abuse present on the black market has increased at a rapid pace. As certain drugs are banned, clandestine laboratories produce alternative structures for them in order to avoid legal consequences [3].

Phenylethylamines are organic compounds that stimulate the human central nervous system and are consumed globally as recreational drugs. By replacing one or more hydrogen atoms in the molecular structure of phenylethylamine, new derived compounds are obtained. They form the class of substituted phenylethylamines.

Synthetic compounds belonging to the class of substituted phenylethylamines have experienced rapid spread in the drug abuse market. Among them, very dangerous for human health are those analogues that have hallucinogenic effects (in addition to the stimulant pharmacological activity), such as those belonging to the 2C-x, DOx and NBOMe classes of drugs of abuse.

The 2C-x compounds (substituted Dimethoxyphenylethylamines) are phenylethylamines containing methoxy groups at the 2 and 5 positions of the benzene ring. Most 2C-x compounds contain various lipophilic substitutes at position 4 of the benzene ring. The name of "2C" was introduced by Alexander Shulgin and refers to the two carbon atoms present between the amino group and the benzene ring [4,5]. The general chemical structure of the 2C-x compounds is shown in Figure 1.1.

Many compounds belonging to 2C-x class were first synthesized by Alexander Shulgin in the 1970s–1980s [6]. In 1991, he published with his wife, Ann Shulgin, the book *PiHKAL (Phenethylamines I Have Known And Loved),* which presents the results of his research. The book presents both the synthesis and the side effects associated with the consumption of various compounds belonging to the 2C-x class [5]. After the publication of the book, the interest concerning the 2C-x class increased significantly [4]. Thus, many substances from the 2C-x class became recreational drugs in the late 1990s and early 2000s [6,7]. By the time they appeared on the illicit drug market, some of these substances were already classified in some countries as risk drugs and placed in

one of the known risk classes, while other substances from the 2C-x class came to be classified much later [6].

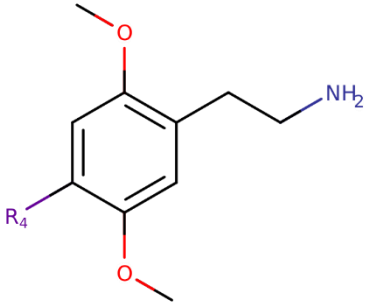

Figure 1.1 General chemical structure of the 2C-x compounds

Various studies have indicated that these substances are dangerous and addictive, leading to the gradual inclusion of the 2C-x compounds in the list of high-risk drugs at various times in most countries. For example, in the United States, many 2C-x substances, such as 2C-E (2,5-dimethoxy-4-ethylphenylethylamine), were listed as most dangerous drugs (Schedule I) in 2012, while others, such as 2C-B (2,5-dimethoxy-4-bromophenylethylamine), were introduced as early as the 1990s [4,8,9]. The European Council instituted control measures and criminal charges for the use and distribution of 2C-I (2,5-dimethoxy-4-iodophenylethylamine) in 2003, whereas in the US this was only the case in 2012 [9,10]. This classification is made based on their approved (or not) medical use, side effects, and their potential to be addictive [11].

As compounds from the 2C-x family are placed in the category of risk drugs, suppliers offer legal alternatives to these substances [6]. Drugs belonging to the 2C-x family can be administered in various forms, such as orally, buccally/sublingually or nasally, and can be found in powder, liquid, tablet, or capsule form, or even in blotter papers [4,6,12,13].

As the use of 2C-x compounds in legal pharmaceutical preparations is prohibited, the effects of these substances on humans have not been widely scientifically documented or analysed. However, the 2C-x compounds are known to induce hallucinogenic effects. The 2C-x substances act as agonists or antagonists of serotonin 5-HT_{2A} receptor and alpha adrenergic receptors [6].

The effects of these compounds vary in intensity. In small doses, the 2C-x substances can play an important role in increasing the intensity of sensory perceptions. In large doses, the effects of these drugs can be especially dangerous, triggering hallucinations, tachycardia, convulsions, high blood pressure or even death [4,6].

DOx compounds (substituted dimethoxyamphetamines) are amphetamine derivatives containing methoxy groups in positions 2 and 5 of the aromatic ring Most DOx compounds contain various lipophilic substitutes at position 4 of the aromatic ring. Amphetamines are an important class of substituted phenylethylamines. The basic chemical structure of the compounds belonging to the DOx class is shown in Figure 1.2.

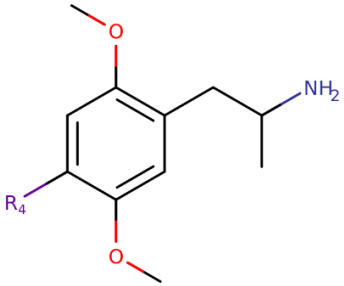

Figure 1.2 General chemical structure of the DOx compounds

DOx compounds are psychedelic drugs that have a long-lasting effect. These compounds act as partial, selective agonists of the 5-HT_{2A}, 5-HT_{2B} and 5-HT_{2C} receptors [6]. The first to synthesise these compounds is Alexander Shulgin. The synthesis and side effects associated with the DOx substances were also presented in the *PiHKAL book* [5].

Alexander Shulgin first synthesized the DOM compound (2,5-dimethoxy-4-methylamphetamine) in 1964. Other analogues such as DOB (2,5-dimethoxy-4-bromoamphetamine), DOC (2,5-dimethoxy-4-chloroamphetamine) or DOI (2,5-dimethoxy-4-iodoamphetamine) followed [6]. Although some DOx compounds have been found on the illegal drug market for several decades, they have only become extremely popular among users in the twenty-first century [6].

In addition to the hallucinogenic effects produced, taking substances from the DOx family can have negative effects on health. Consumed in large doses, they can affect the cardiovascular system [6,14]. The DOx compounds can be found in liquid, powder, tablet, capsule form or in blotting papers and are most often taken orally [15–17].

NBOMe compounds (N-(2-methoxybenzyl)phenylethylamines) are analogues of the compounds belonging to the 2C-x family, containing an N-(2-methoxy)benzyl substitute. The general chemical structure of the NBOMe compounds is represented in Figure 1.3.

Like the 2C-x drugs, the compounds from the NBOMe class have been shown to be serotonin 5-HT_{2A} receptor agonists, thus endowed with

halluciongenic properties [18–20]. Also, due to the interaction with alpha-adrenergic receptors, NBOMe compounds can produce stimulant effects [18].

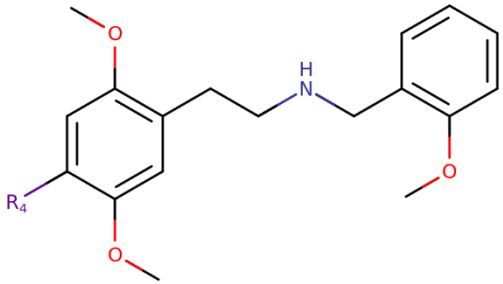

Figure 1.3 General chemical structure of the NBOMe compounds

Research by Richard Glennon et al., followed by studies conducted by Ralph Heim and Martin Hansel, led to the synthesis and characterisation of the compounds belonging to the NBOMe class [21–24]. In 2010, the compounds from the NBOMe class appeared on the illicit drug market [25,26].

The NBOMe drugs of abuse can be administered by various routes, such as sublingual/oral route, intravenously, nasal insufflations or smoking [25]. They can be found in samples in many forms such as powder, liquid, pill form or in blotting papers [25,27]. Among them, most often, the NBOMe compounds are found in in blotter papers inscribed with various colored patterns, works of art, illustrations of cartoon characters or music and film posters [25,28,29].

The choice of blotting papers for drug dispensing is common for drugs for which the usual dose administered is low [30]. In the case of a liquid or powder sample, limiting the dose administered is more difficult and the consequences of administering an overdose can be very serious [31]. The NBOMe compounds can produce effects such as visual and auditory hallucinations, convulsions, agitation, panic, hypertension, tachycardia, mydriasis (pupil dilation) and so on [25,29,32].

According to an analysis of drugs owned by participants in a festival of psychedelic music and culture in Portugal, the presence of several DOx and NBOMe compounds was identified in samples sold as LSD (lysergic acid diethylamide) [33]. Involuntary use of DOx and NBOMe drugs of abuse can increase the risk of poisoning with these substances, thus having dangerous effects on the body [33]. An analysis of commercial blotting papers impregnated with various NBOMe substances revealed that often these substances are found in significantly higher concentrations than those stated by the supplier, which can lead to adverse consequences [30].

Chapter II. Spectral characterisation and identification of the analysed compounds by ATR-FTIR

2.2. Infrared spectroscopy

2.2.1. Fourier Transform Infrared Spectroscopy (FTIR)

The first infrared spectrometers were of the dispersive type. They use scattering elements, such as prisms or diffraction networks, to separate individual signals emitted from the source. Each infrared spectrometer of dispersive type is equipped with a detector that measures the amount of energy at each frequency separately, resulting in the final spectrum of the analysed sample [34].

In recent years, a new type of infrared spectrometers, called Fourier Transform infrared spectrometers, have gained popularity due to the advantages that their use brings. Unlike dispersive infrared spectrometers, these new spectrometers analyse all infrared frequencies simultaneously, making them faster and more accurate analysis tools [34].

Fourier Transform spectrometers consist of three basic elements: a radiation source, an interferometer and a detector. Their main component is the interferometer, the most common being the Michelson interferometer [34,35].

The Michelson interferometer consists of two plane mirrors perpendicular to each other, one of which is fixed and the other movable, and of a beam divider (semitransparent instrument). The beam divider divides the radiation coming from the source into two identical beams so that each beam has a different path. One of the beams is transmitted, while the other is reflected. Thus, one of the beams reaches the movable mirror, and the other, to the fixed mirror. When the two beams are reflected back by the mirrors, they are recombined at the radiation divider, after an optical path difference has been introduced between them by means of the movable mirror (see Figure 2.1). The signal is then transmitted to the detector in the form of an interferogram, which is then processed by a mathematical process, called the Fourier Transform, which ensures the passage into the frequency range, resulting in a spectrum [34,35].

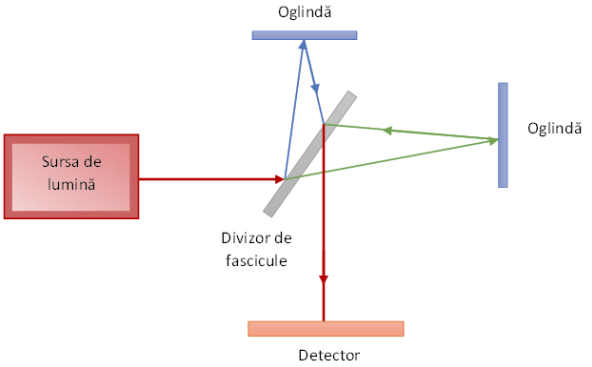

Figure 2.1 Schematic diagram of the Michelson interferometer as part of an FTIR spectrometer

2.2.2. Attenuated Total Reflectance (ATR)

The ATR accessory is an instrument that can be used in FTIR spectrometers. It is designed to measure changes that occur in an internally reflected infrared beam when it comes into contact with a sample. The beam is directed to a dense optical crystal, which has a high refractive index, at a certain angle. This creates an evanescent wave that extends beyond the crystal surface in the sample kept in contact with it. This evanescent wave is attenuated in areas of the infrared spectrum where energy is absorbed by the sample. The attenuated beam exits through the opposite end of the crystal and is directed to the detector in the spectrometer. The main advantage of using the ATR accessory is that it allows samples to be analysed in their natural state, without requiring complex prior preparation [36].

2.4. Vibrational analysis of chemical compounds based on their ATR-FTIR spectra

Vibrational spectroscopy is a non-destructive method of identification and characterisation, by which specific chemical bonds of atoms, chemical compounds or functional groups present in a particular sample can be identified. This technique is based on the vibrational movements of molecules and allows analysis of liquid, solid or gaseous samples [37,38].

An important type of vibrational spectroscopy refers to infrared spectroscopy. In general, almost any substance with covalent bonds absorbs electromagnetic radiation from the infrared region of the electromagnetic spectrum at different frequencies. By absorbing infrared radiation by molecules, the vibrational energy of binders between atoms changes. Thus, it is possible to identify specific chemical bonds of atoms, chemical compounds or functional groups present in the chemical structure of the analysed substances [37]. There are two important groups of vibrations: stretching vibrations and bending

vibrations. The stretching vibrations correspond to a continuous change in the interatomic distance along the axis of the bond between two atoms. Bending vibrations correspond to a change in the angle formed between two atomic bonds [39].

Chapter III. Mathematical methods and approaches applied for the analysis and the identification of class membership of the targeted drugs of abuse

3.1. Molecular descriptors and toxicity parameters

Molecular descriptors are mathematical representations of the properties of a molecule and have been widely used in recent decades. They convert the physico-chemical and structural information of molecules into numerical format, thus becoming a useful tool in applying various methods of computational analysis [40–43]. An important role of this mathematical approach is to characterise and provide easier identification of related molecular compounds. There are currently many categories of molecular descriptors that can be used for this purpose, such as charge descriptors, radial distribution function (RDF) descriptors, delocalisation-degree indices, topological descriptors (indices), functional group counts descriptors, quantum molecular descriptors and toxicity molecular descriptors. Toxicity molecular descriptors are estimates of the toxicological character of molecules. Toxicity can also be assessed on the basis of estimates of toxicity parameters such as lethal dose 50 (LD50), carcinogenicity and immunotoxicity. In recent years, various software tools have been developed, by which molecular descriptors and toxicity parameters can be quickly determined.

3.3. Exploratory analysis methods applied for the analysis and the identification of class membership of the targeted drugs of abuse

3.3.1. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a multivariate mathematical method used to reduce the dimension of a data set by transforming the initial variables into new orthogonal variables, called principal components. They represent linear combinations of the initial variables and are determined so as to encapsulate the maximum variance that can be preserved from the initial data set [44]. In other words, PCA aims to reduce the dimension of a data set while retaining its maximum amount of information. PCA is also successfully used as an exploratory analysis method, being a useful tool for highlighting the grouping tendency of the analysed drugs and for identifying the variables that contribute most to the modeling of the targeted classes [45,46].

3.4. Machine learning methods applied for the analysis and the identification of class membership of the targeted drugs of abuse

3.4.1. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a machine learning method that can be successfully used in classification problems. Initially, this method was designed to deal with binary classification problems (involving two distinct classes), but it can now deal with problems involving more than two classes. SVM is effective for analyzing both simple linear problems and complex nonlinear problems [44].

3.4.2. Random Forest (RF)

Random Forest (RF) is a machine learning method applicable in solving classification problems. To determine whether a sample belongs to a particular class, the method uses a collection of decision trees, each of which provides a prediction. Finally, the sample is assigned to the class that was predicted by most decision trees, with each tree voting for only one class. Alternatively, predicting whether a sample belongs to a particular class can be accomplished by averaging probabilistic predictions, calculated based on each decision tree [47].

3.4.3. Gradient Boosting (GB)

Gradient Boosting (GB) is a machine learning method that can be successfully implemented for solving classification problems. The method combines several weak learners (trees, in general) to create a powerful predictive model, this techniqe being known as *boosting*. This approach involves sequentially training the models, each new model trying to correct errors made by previous models [48,49].

3.4.4. eXtreme Gradient Boosting (XGBoost)

The eXtreme Gradient Boosting (XGBoost) method is a machine learning method that stands out for its good results in solving both binary classification problems and multiclass classification problems. The algorithm associated with the XGBoost method is based on the algorithm associated with the GB method and is known as a powerful, optimised and scalable algorithm, preferred for developing predictive models, due to its high degree of efficiency, accuracy and adaptability [50,51].

3.4.5. K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) is a machine learning method that solves classification problems based on a relatively simple algorithm. It assigns the class label of a test object based on the predominant class in the group of objects (neighbors) in the training dataset which are closest to the test object K[52].

3.4.6. Evaluation of the performance of the classification models associated with the utilised machine learning methods

As discussed in previous sub-chapters, there are various machine learning methods that can be applied to develop classification models. Assessment of the quality of the models obtained can be carried out using various performance metrics. Their analysis allows comparing the performance of the obtained classification models, thus leading to the choice of the best model in the context of the given problem.

Some of the most commonly used performance metrics are confusion matrix, sensitivity, specificity, accuracy, balanced accuracy, Matthews correlation coefficient (MCC) and area under the ROC curve (Area Under the Receiver Operating Characteristic Curve, ROC AUC).

Chapter IV. Personal contributions regarding the analysis and the recognition of class identity of the 2C-x, DOx and NBOMe drugs of abuse based on specific techniques and the application of multivariate mathematical methods

4.1. Analysis and class identity recognition of drugs of abuse belonging to the 2C-x, DOx and NBOMe classes based on their ATR-FTIR spectra

Considering that the number of new drugs, derivatives of the existing drugs, is constantly increasing on the black market, it is necessary to find tools which could be used to achieve the attribution of their class identity. Thus, it becomes essential to urgently develop tools to effectively detect both already known drugs and new drugs, which are often derivatives of the first mentioned substances. A useful approach to this is to use the ATR-FTIR spectra of targeted drugs. The results obtained regarding the analysis and the class identity recognition of the drugs belonging to the 2C-x, DOx and NBOMe classses based on their ATR-FTIR spectra are presented below.

4.1.1. KNN-type machine learning system designed for the detection of illicit drugs based on ATR-FTIR spectra [53]

Machine learning methods have been successfully applied in recent years to deal with various classification problems in various fields. In this subchapter are presented the results obtained based on the KNN algorithm, which was applied to detect substances belonging to the NBOMe and opioids classes, which are abused for recreational purposes [53]. Opioids represent a class of drugs of abuse with an important role in treating pain, used for medical but also for illicit purposes [54,55].

For the detection of the illicit drugs belonging to these two classes, a KNN classification model was built, using as database 68 ATR-FTIR spectra selected from the SWGDRUG public spectral library [56]. The spectral database was divided into three different classes: class 1 – NBOMe, class 2 – opioids and class 3 – negatives. The first class comprises the spectra of 12 NBOMe compounds, the second class includes the spectra of 30 opioids, and the third class consists of the spectra of 26 different substances that do not belong to the other two modelled classes of substances.

The KNN model was built by using the *Python* software. The database was randomly divided into two parts, with 70% of the spectra forming the training set and 30% forming the test set. For a better estimate of accuracy, the training session was repeated 10 times. An average accuracy of 79.99% was achieved for the training set and 70.17% for the test set, thus highlighting the substantial potential of the model.

The confusion matrix corresponding to accuracy having the closest value to mean accuracy is shown in Figure 4.1. This indicates that the KNN model correctly classified all NBOMe compounds (100%), while only 70% of the opioids were correctly classified. The rest of the opioids were misclassified as NBOMe (10%) or negatives (20%).

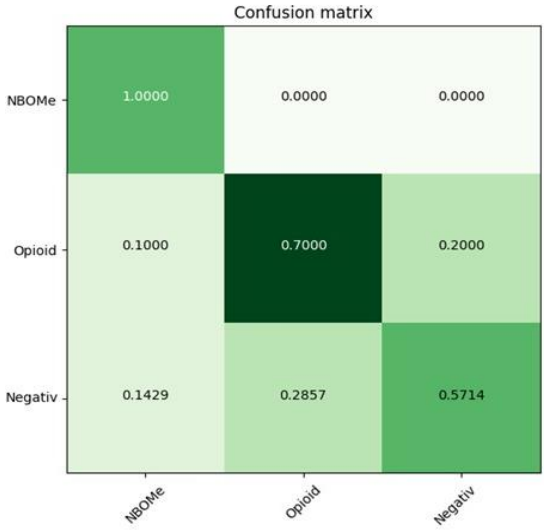

Figure 4.1 Confusion matrix for the KNN model built to detect substances beloinging to the NBOMe and opioids classes [53]

The KNN algorithm analysed in this study provided promising results for classifying and detecting substances beloinging to the NBOMe and opioids classes based on their ATR-FTIR spectral data. One notable result of this research is the model's ability to accurately identify 100% of NBOMe compounds, which are well known for their high potency and abuse potential. Opioids, on the other hand, were correctly classified in 70% of cases. The remaining misclassifications occurred mainly between the opioids and the negative classes, indicating the need for further refinement of the model to improve the differentiation between these classes.

Despite its limitations, this study shows the potential of the KNN method to detect and classify drugs belonging to the two analysed classes. These results form a basis that can ultimately contribute to the development of rapid, accurate and noninvasive drug detection methods, crucial for managing the global problem of drug abuse.

4.1.2. Machine learning systems designed for the detection of illicit drugs based on ATR-FTIR spectra [46]

The results obtained in the previous sub-chapter determined the extension of the research by analysing more classes of illicit drugs and diversifying the machine learning methods applied for the detection of the illicit drugs belonging to those classes.

This sub-chapter presents a comparative study aiming to determine the most effective multivariate model that can be used to detect important classes of drugs of abuse based on their ATR-FTIR spectra. The substances covered by the analysis were divided into four classes: the hallucinogenic phenylethylamines class (composed mainly of the substances belonging to the 2C-x, DOx and NBOMe classes), the cannabinoids class, the opioids class and the negatives class, comprising various other substances of forensic interest [46]. Cannabinoids are a class of drugs comprising natural, synthetic and semisynthetic compounds that interact with cannabinoid receptors [57]. With the expansion of the accessibility of cannabinoids, especially the synthetic ones, public interest in these compounds has increased significantly [58].

Such illicit drugs constantly appearing on the black market are a real problem today. From this point of view, it is important to develop models that can automatically detect the class membership of these new compounds.

The aim of this study is to develop a machine learning system that can be used to detect hallucinogenic phenyethylamines (mainly 2C-x, DOx and NBOMe compounds), opioids and cannabinoids, based on their ATR-FTIR spectra.

ATR-FTIR spectrometers are increasingly used for field screening of illicit drugs, as they are portable instruments and do not require sample preparation [59]. The ATR-FTIR spectra used in this study were extracted from the SWGDRUG public spectral library [56]. The spectral database comprises 95

spectra corresponding to targeted illicit drugs and negatives (different randomly selected substances of forensic interest).

The spectral database was divided into four classes: class 1 – hallucinogenic phenylethylamines (mainly including spectra of compounds belonging to the 2C-x, DOx and NBOMe classes), class 2 – opioids, class 3 – cannabinoids and class 4 – negatives. The hallucinogenic phenylethylamines class comprises 25 spectra, the opioids class consists of 34 spectra, the cannabinoids class contains 18 spectra, and the negatives class consists of 18 spectra corresponding to different randomly selected compounds. The machine learning models were built by using the *Python* software.

In order to analyse the grouping tendency of data, PCA was initially applied considering two principal components. Figure 4.2 displays the score plot obtained for the two PCs. It shows that the hallucinogenic phenylethylamines form the most compact cluster. The points associated with the opioids and the cannabinoids are much more spread out than those corresponding to the substances belonging to the class of hallucinogenic phenylethylamines. Many of the points associated with the negatives overlap the clusters formed by the positives, especially the cluster of the opioids.

Next, five machine learning models were developed to detect the targeted illicit drugs, namely SVM, XGBoost, RF, GB and KNN models [50,60–63]. These models were chosen due to their efficiency, simplicity and quick implementation.

For the development of each model, the dataset was randomly divided into two parts: 60% of the total spectra were used for training and the remaining 40% were used for testing. Each model was then trained on the training set and evaluated on the test set. The model, training and test datasets were then deleted. We define this process as a training session. Although the initial dataset for each session was the same, the training and test sets were different at each iteration because the entries were randomly selected each time. In other words, models were trained and evaluated each time on different selections of the same dataset. Each training session was repeated 10 times.

In order to determine and compare the performance of the obtained models, their confusion matrices were determined, but also a series of performance metrics, namely the balanced accuracy, the sensitivity, the specificity, the Matthews correlation coefficient and the area under the ROC curve.

The confusion matrices determined for each model are shown in Figure 4.3, Figure 4.4, Figure 4.5, Figure 4.6 and Figure 4.7. Table 4.1 presents the average values of the performance metrics used, obtained after 10 runs, corresponding to the five models developed.

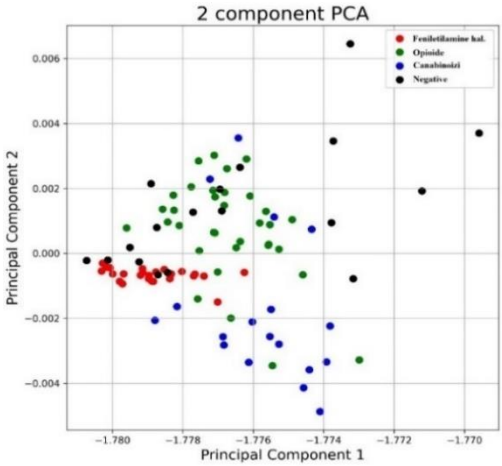

Figure 4.2 Score plot of the first two principal components of a two-component PCA displaying the clusters formed by hallucinogenic phenylethylamines (red), opioids (green), cannabinoids (blue), and negatives (black) [46]

Table 4.1 indicates that the SVM and XGBoost models have the highest balanced accuracy. At the same time, the SVM model has the highest specificity, while the XGBoost is the most sensitive model. The SVM and XGBoost models have the best (and comparable) Matthews correlation coefficient, while the coefficient determined for the other models is significantly lower. The value of this coefficient is positive for all models, which indicates positive correlations in all cases. The SVM and XGBoost models also have the highest ROC AUC score, having almost the same value for both models. As the ROC AUC score is very high (close to 1), we may conclude that these two models have a very good prediction rate.

If we consider that the models tested are tree-based models (XGBoost, RF and GB), decision boundary models (SVM) and non-parametric models (KNN), we may conclude that the decision boundary models had the best results, followed by the tree-based models and the non-parametric models.

The confusion matrices (Figure 4.3, Figure 4.4, Figure 4.5, Figure 4.6 and Figure 4.7) indicates that, except for the GB model, all models classify the hallucinogenic phenylethylamines with 100% accuracy. The GB model is not much less efficient, with an accuracy of 80% concerning the classification of the hallucinogenic phenylethylamines. The main difference between the models regarding the hallucinogenic phenylethylamine class is related to the false

positive rate, which is 20% for the GB model, 25% for the RF model, 32.14% for the SVM model, 33.33% for the XGBoost model and 53.33% for the KNN model.

Table 4.1 Average values of the performance metrics corresponding to the classification models [46]

Modelul	**Balanced accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Matthews correlation coefficient**	**ROC AUC**
Support Vector Machines (SVM)	93,1 ± 5,54	89,41 ± 4,16	96,79 ± 3,66	0,89 ±0,04	0,92
eXtreme Gradient Boosting (XGBoost)	89,66 ± 8,42	94,66 ± 7,95	84,66 ± 6,74	0,86 ± 0,5	0,91
Random Forest (RF)	82,76 ± 8,14	71,86 ± 6,93	93,66 ± 7,62	0,69 ± 0,09	0,82
Gradient Boosting (GB)	75,86 ± 5,21	65,62 ± 5,27	86,10 ± 4,98	0,61 ± 0,05	0,76
K-Nearest Neighbors (KNN)	65,52 ± 9,23	61,25 ± 10,35	91,12 ± 8,33	0,56 ± 0,09	0,75

The opioids are 100% correctly classified by the XGBoost model. The second best correct classification rate (92.86%) is recorded for the SVM model, with 7.14% of the opioids being misclassified as hallucinogenic phenylethylamines. It is worth noting, however, that the false positive rate is 66.67% in the case of the XGBoost model and 0% in the case of the SVM model. The other models perform worse at correctly attributing the class identity to the opioids.

Cannabinoids are recognised as such with 100% accuracy by the SVM and the XGBoost models, both of which have a false positive rate of 0%. The rest of the models have significantly lower performance in this regard.

Taking into account both the accuracy and the misclassification rates, the negatives appear to be the hardest to correctly classify for all the models, most likely due to the wide variety of substances that make up this class.

The availability of screening tools capable of detecting illicit substances harmful to humans quickly and reliably is essential for public safety. The models presented in this thesis can work in harmony with the currently recommended methodology for designer drug detection.

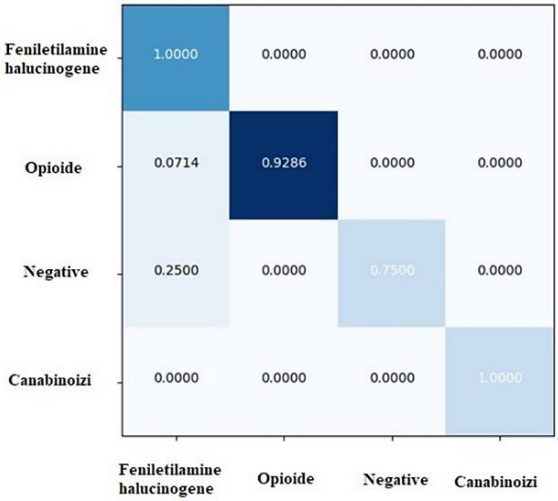

Figure 4.3 Confusion matrix corresponding to the SVM model

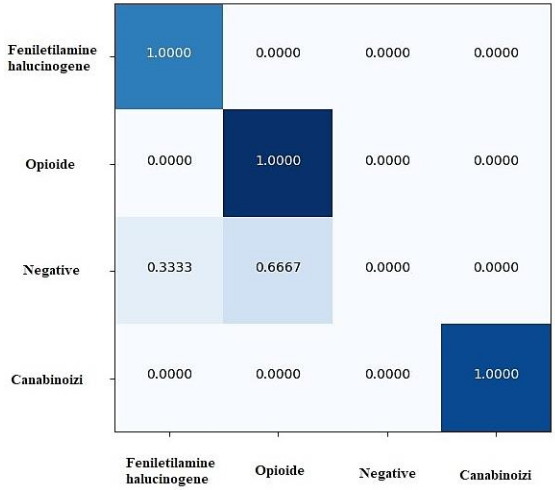

Figure 4.4 Confusion matrix corresponding to the XGBoost model

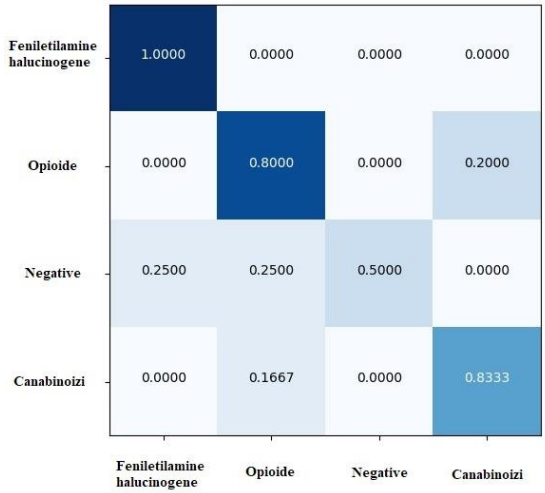

Figure 4.5 Confusion matrix corresponding to the RF model

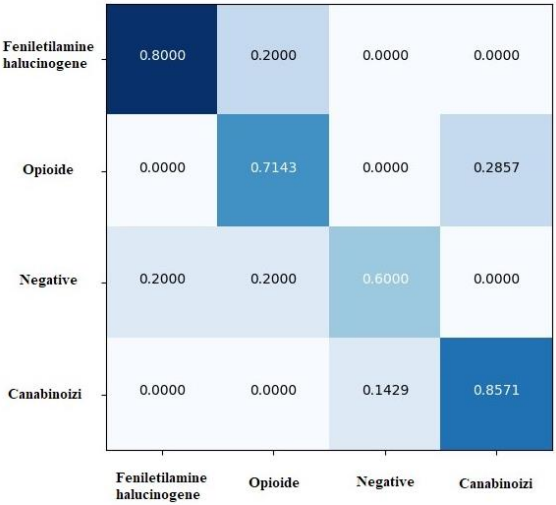

Figure 4.6 Confusion matrix corresponding to the GB model

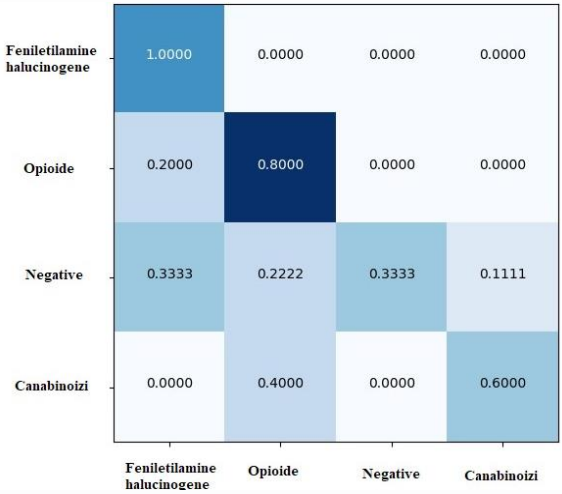

Figure 4.7 Confusion matrix corresponding to the KNN model

In terms of overall accuracy, the best performing model is SVM. As screening forensic systems designed to operate ATR-FTIR field (portable) analytical tools, the models developed should be able to perform cost-effective, non-destructive, direct, real-time, on-site testing. However, the main objective of these models is to restrict the number of samples subjected to further in-depth analysis with more sophisticated stationary analytical instruments in the laboratory. Only the samples tested on site and assigned a positive class identity (hallucinogenic phenylethylamines, cannabinoids and opioids) will be analysed in the laboratory to determine their individual identity (not only their class membership).

Therefore, an essential feature of such a screening system is its effectiveness in detecting positives. In this case, no compounds such as hallucinogenic phenylethylamine, cannabinoid or opioid should be misclassified as falsely negative. For this reason, XGBoost is at least as suitable for this purpose as SVM, since XGBoost does not produce false negatives. While 7.14% of the opioids are erroneously classified as hallucinogenic phenylethylamines by SVM, no hallucinogenic phenylethylamine, opioid, or cannabinoid is misclassified by XGBoost.

It is true that XGBoost has a higher erroneous classification rate than the SVM model. XGBoost fails to correctly classify the negatives, while SVM misclassifies only 7.14% of the opioids as hallucinogenic phenylethylamines and 25% of the negatives as hallucinogenic phenylethylamines. However, the false

positives (false hallucinogenic phenylethylamines or opioids), although not desirable, are less important. As mentioned earlier, their individual identity (molecular structure) will be determined during further laboratory testing, based on a series of analytical methods that are recommended for each class of drug abuse by specialized international agencies such as the United Nations Office on Drugs and Crime [64,65]. In conclusion, SVM works best of all models tested, but XGBoost is a choice at least as good as SVM from a forensic point of view.

The high classification accuracy of the presented models indicates that AI-based strategies represent an important way forward in the context of automatising the processing of the ATR-FTIR spectra during field operations. The model that works best within the classification strategy that only considers the overall accuracy is SVM. However, since these are forensic tools, the classification strategy should also consider the false negative rate. For this reason, XGBoost is a choice at least as good as SVM, and its overall accuracy is comparable to that of SVM.

4.1.3. Vibrational analysis of some hallucinogenic phenylethylamines belonging to the 2C-x and DOx classes of drugs of abuse [38]

Vibrational spectroscopy is a useful, non-destructive technique that can be used to investigate and identify chemical compounds by determining their functional groups. This subchapter presents the vibrational analysis of three substances belonging to the 2C-x and DOx classes, namely 2-(4-iodo-2,5-dimethoxyphenyl)ethanamine (2C-I), 2-(4-bromo-2,5-dimethoxyphenyl)ethanamine (2C-B) and 1-(4-bromo-2,5-dimethoxyphenyl)propan-2-amine (DOB) based on their ATR-FTIR spectra [38]. The spectra of the analysed substances (depending on the wavenumber and absorbance), obtained from the spectral data provided by SWGDRUG, are presented in Figure 4.8, Figure 4.9 and Figure 4.10 [56].

The functional groups and the modes of vibration identificate identified for the most important peaks present in the spectra of the targeted 2C-x and DOx substances are presented in Table 4.2, Table 4.3 and Tabelul 4.4 [66,67].

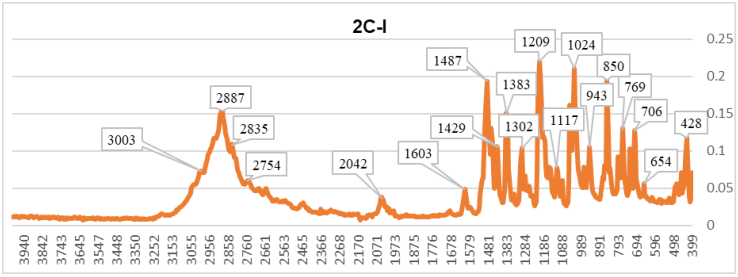

Figure 4.8 ATR-FTIR spectrum of 2-(4-iodo-2,5-dimethoxyphenyl)ethanamine (2C-I) [38]

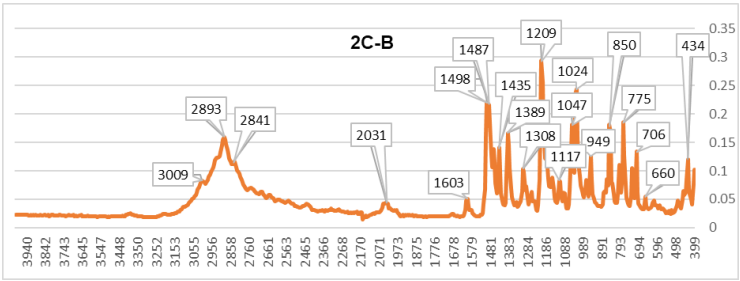

Figure 4.9 ATR-FTIR spectrum of 2-(4-bromo-2,5-dimethoxyphenyl)ethanamine (2C-B) [38]

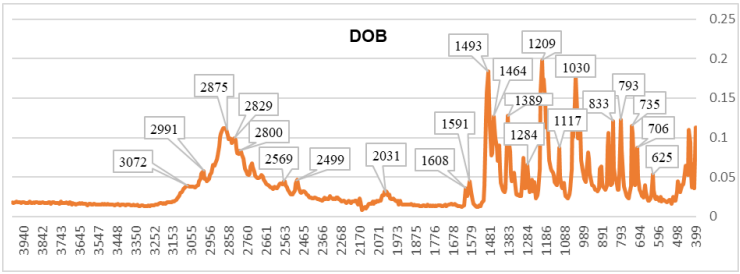

Figure 4.10 ATR-FTIR spectrum of 1-(4-bromo-2,5-dimethoxyphenyl)propane-2-amine (DOB) [38]

Table 4.2 Vibrational analysis of 2-(4-iodo-2,5-dimethoxyphenyl)ethanamine [38]

2C-I	Functional group	Mode of vibration
3003	C-H (in the aromatic ring)	stretching
2887	C-H (alkane)	stretching
	N-H (amine)	stretching
2835	N-H (amine)	stretching
2754		
2042		
1603	C-C (in the aromatic ring)	stretching

	N-H (amine)	bending
1487	C-C (in the aromatic ring)	stretching
1429	C-C (in the aromatic ring)	stretching
	C-H (alkane - methyl group)	bending
1383	C-H (alkane - methyl group)	bending
1302		
1209	C-O (alkyl aryl ether)	stretching
1117	C-N (amine)	stretching
1024	C-N (amine)	stretching
943		
850	C–H	bending
769	C–H	bending
706	C–H	bending
654	C-I	stretching
428		

Table 4.3 Vibrational analysis of 2-(4-bromo-2,5-dimethoxyphenyl)ethanamine [38]

2C-B	**Functional group**	**Mode of vibration**
3009	C-H (in the aromatic ring)	stretching
2893	C-H (alcan)	stretching
	N-H (amine)	stretching
2841	N-H (amine)	stretching
2031		
1603	C-C (in the aromatic ring)	stretching

	N-H (amine)	bending
1498	C-C (in the aromatic ring)	stretching
1487	C-C (in the aromatic ring)	stretching
1435	C-C (in the aromatic ring)	stretching
	C-H (alkane - methyl group)	bending
1389	C-H (alkane - methyl group)	bending
1308		
1209	C-O (alkyl aryl ether)	stretching
1117	C-N (amine)	stretching
1047	C-N (amine)	stretching
1024	C-N (amine)	stretching
949		
850	C–H	bending
798	C–H	bending
775	C–H	bending
706	C–H	bending
660	C-Br	stretching
434		

Tabelul 4.4 Vibrational analysis of 1-(4-bromo-2,5-dimethoxyphenyl)propane-2-amine [38]

DOB	Functional group	Mode of vibration
3072	C-H (in the aromatic ring)	stretching
2991	C-H (alcan)	stretching
	N-H (amine)	stretching
2875	C-H (alcan)	stretching
	N-H (amine)	stretching
2829	C-H (alcan)	stretching
	N-H (amine)	stretching
2800	N-H (amine)	stretching
2737		
2569		
2499		
2031		
1608	N-H (amine)	bending
1591	N-H (amine)	bending
	C-C (in the aromatic ring)	stretching
1493	C-C (in the aromatic ring)	stretching
1464	C-C (in the aromatic ring)	stretching
	C-H (alkane - methyl group)	bending
1389	C-H (alkane - methyl group)	bending
1354		
1308		

1284		
1209	C-O (alkyl aryl ether)	stretching
1198	C-N (amine)	stretching
1117	C-N (amine)	stretching
1030	C-N (amine)	stretching
966		
897	C–H	bending
856	C–H	bending
833	C–H	bending
793	C–H	bending
735	C–H	bending
706	C–H	bending
625	C-Br	stretching
492		
451		
434		
399		

The vibrational analysis of the ATR-FTIR spectra indicates that the three substances have the most important peaks in approximately the same regions, namely in the spectral regions 3010-2500 cm^{-1}, 1600-700 cm^{-1} and 690-500 cm^{-1}. Thus, the existence of three main spectral domains is noted. The first spectral range corresponds to 3010-2500 cm^{-1} and reveals the presence of the N-H and C-H groups, suggesting also the presence of the aromatic ring. The second spectral range, corresponding to 1600-700 cm^{-1}, comprises the largest number of peaks of the spectra and shows the presence of the C-H, N-H, C-C, C-O and C-N groups. The third spectral region ranges between 690 and 500 cm^{-1} and suggests the presence of the C-I and C-Br groups, specific to the analysed compounds. The vibrational analysis of the spectra highlights the similarities between the analysed compounds, being an efficient, selective method of characterisation.

4.2. Analysis and class identity recognition of drugs of abuse belonging to the 2C-x, DOx and NBOMe classes based on molecular descriptors and toxicity parameters

Molecular descriptors convert the chemical and structural information of the substances into mathematical language. With their help, the physical and chemical properties of substances can be expressed in numerical format. An important role of this transformation is to characterise and provide easier identification of related substances [40–43]. In this respect, various molecular descriptors calculated for some of the most representative substituted phenylethylamines belonging to the 2C-x DOx and NBOMe classes were analysed.

4.2.1. Use of geometrical, topological and functional group counts descriptors for the characterisation of some compounds belonging to the 2C-x and DOx classes of drugs of abuse [43]

Geometrical descriptors, topological descriptors (indices) and functional group counts descriptors are three important categories of molecular descriptors that can be successfully used to characterise the targeted drugs [43]. A significant number of descriptors from these categories were determined for six representative illicit drugs belonging to the 2C-x and DOx classes. The list of the analysed compounds is given in Tabelul 4.5.

Tabelul 4.5 List of the analysed 2C-x and DOx compounds [43]

Compound name	Class
2,5-dimethoxy-4-ethylthiophenylethylamine (2C-T-2)	2C-x
2,5-dimethoxy-4-propylthiophenylamine (2C-T-7)	2C-x
2,5-dimethoxy-4-bromophenylethylamine (2C-B)	2C-x
2,5-dimethoxy-4-ethylamphetamine (DOET)	DOx
2,5-dimethoxy-4-bromoamphetamine (DOB)	DOx
2,5-dimethoxy-4-methylamphetamine (DOM)	DOx

From the geometric descriptors category, three delocalisation-degree indices were calculated for each compound: the Harmonic Oscillator Model of Aromaticity index ($HOMA$), the aromaticity index ($AROM$) and the $HOMA$ total ($HOMT$). From the topological descriptors category we have determined the first Zagreb index ($ZM1$), the second Zagreb index ($ZM2$), the quadratic index ($Qindex$), the simple topological index Narumi ($SNar$) and the geometric topological index Narumi ($Gnar$).

From the functional group counts descriptors category the following descriptors were selected for analysis: the number of aromatic C($sp2$) ($nCar$), number of unsubstituted benzene C($sp2$) ($nCbH$), the number of substituted benzene C(sp2) ($nCb-$) and the number of acceptor atoms for H-bonds (N, O, F) ($nHAcc$).

The molecular descriptors in the three categories were calculated using the *Dragon 5.5 software* [68]. The results (rounded to three decimal places) are presented inTable 4.6, Table 4.7 and Table 4.8.

Table 4.6 Values of the topological indices calculated for the analysed 2C-x and DOx compounds

	Topological indices				
	$ZM1$	$ZM2$	$Qindex$	$SNar$	$GNar$
2C-T-2	72	82	7	9,94	1,852
2C-T-7	76	86	7	10,633	1,869
2C-B	64	73	7	8,553	1,842
DOET	74	84	8	9,652	1,828
DOB	70	79	8	8,959	1,817
DOM	70	79	8	8,959	1,817

Table 4.7 Values of the geometrical descriptors calculated for the analysed 2C-x and DOx compounds

	Geometrical descriptors		
	$HOMA$	$AROM$	$HOMT$
2C-T-2	0,935	0,986	5,611
2C-T-7	0,935	0,986	5,611
2C-B	0,938	0,984	5,628
DOET	0,94	0,987	5,643
DOB	0,94	0,987	5,642
DOM	0,94	0,987	5,643

Table 4.8 Values of the functional group counts descriptors calculated for the analysed 2C-x and DOx compounds

	Functional group counts descriptors			
	$nCar$	$nCbH$	$nCb-$	$nHAcc$
2C-T-2	6	2	4	3
2C-T-7	6	2	4	3
2C-B	6	2	4	3
DOET	6	2	4	3
DOB	6	2	4	3
DOM	6	2	4	3

The results obtained highlight the similarities and differences between the analysed compounds. Among the molecular descriptors included in the analysis, the most relevant for achieving a thorough discrimination between the two classes concerned (2C-x and DOx) proved to be those belonging to the category of geometrical descriptors, namely $HOMA$, $AROM$ and $HOMT$, but also the topological indices $Qindex$ and $GNar$. On the other hand, the functional group counts descriptors could be relevant for the discrimination between the analysed compounds and the compounds belonging to other classes of drugs. Thus, the molecular descriptors calculated in this subchapter can be considered useful tools for identifying he class membership of the of the drugs concerned.

4.2.2. PCA applied for the characterisation and the evaluation of the clustering tendency of some drugs of abuse belonging to the 2C-x class based on topological indices [69]

This subchapter presents the results obtained based on the application of the PCA method, with the main purpose of assessing the grouping tendency of the compounds belonging to the 2C-x class based on the topological indices [69]. The database involved in the analysis included 6 representative compounds belonging to the 2C-x class (called the class of positives) and 10 different compounds of forensic interest (generically called the class of negatives). Table 4.9 presents the complete list of these substances. For each of these compounds, 79 topological indices were calculated, representing the current total number of descriptors in this category that can be determined using *alvaDesc,* the software tool which was used to calculate them [70]. The *Unscrambler X 10.4* software was then used to perform PCA [71].

Table 4.9 List of the 2C-x compounds and those belonging to the class of negatives included in the analysis [69]

Compound name	Compound type
2C-B	2C-x
2C-E	2C-x
2C-H	2C-x
2C-I	2C-x
2C-T-2	2C-x
2C-T-7	2C-x
Butylone	Negative
Cathine	Negative
Ciprofloxacin	Negative
Januvia	Negative
JWH-018 adamantyl-carboxamide	Negative
JWH 019	Negative
JWH-200	Negative
JWH-250	Negative
LSD	Negative
Nitroaspirin	Negative

Figure 4.11 presents 2D score plot relative to the first two principal components. It can be noted that substances belonging to the 2C-x class form a dense cluster, while the substances belonging to the class of negatives are scattered over the entire surface of the plot. Two false positives have also been identified, namely butylone and cathine. They were incorrectly classified probably because of their molecular structures, which are relatively similar to those of the 2C-x substances.

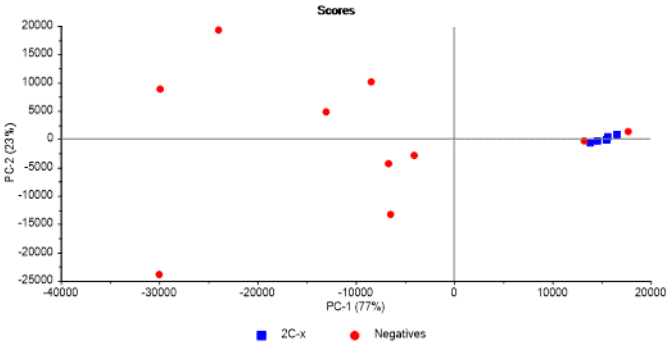

Figure 4.11 2D score plot showing the cluster formed by the 2C-x compounds (blue), as compared to substances from the negatives class (red) [69]

Figure 4.12 presents the loading plot, relative to the first two principal components, which highlights the contribution of each variable to the formation and discrimination of the various formed clusters. Analysing this graph it can be seen that the input variables with the highest discriminating power are: the all-path Wiener index Wap, the Gutman molecular topological index $GMTI$, the Schultz molecular topological index $SMTI$, the Gutman molecular topological index by valence vertex degrees $GMTIV$ and the Schultz molecular topological index by valence vertex degrees $SMTIV$.

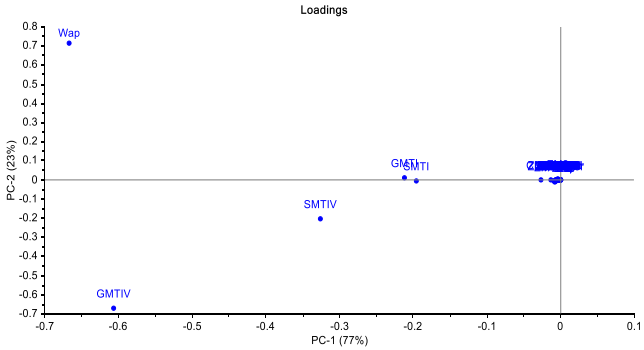

Figure 4.12 Loading plot illustrating the contribution of the topological indices to the formation of the clusters determined by the 2C-x compounds [69]

The results obtained in the analysis indicates that the topological indices can be used to assign the class identity of the 2C-x drugs and can be used as input data for the development of advanced systems built to classify them and to predict their biological activity.

4.2.3. PCA applied for the characterisation and the evaluation the clustering tendency of some drugs of abuse belonging to the 2C-x, DOx and NBOMe classes based on topological indices, charge and RDF descriptors [45]

The results presented in the 4.2.2. subchapter encouraged the expansion of the analysis by adding more compounds from the DOx and NBOMe classes to the database, but also by expanding the negatives class. Furthermore, for the application of the PCA method, other descriptors, representing charge or RDF descriptors, were included in the analysis in addition to the topological indices [45]. Thus, 304 molecular descriptors were calculated during the analysis, 79 of them being topological indices, 15 of them being charge descriptors, and the remaining 210 descriptors being RDF descriptors, representing the current total number of descriptors in these categories that can be determined using *alvaDesc*, the software tool which was used to calculate them [70].

Three classes were considered for the analysis. The first class consisted of some drugs belonging to the 2C-x and DOx clasees, which were treated as a single class due to the similarities existing between their members. The second class comprised some NBOMe drugs, and the third class was formed with different substances of forensic interest and was referred to as the negatives class. Specifically, the database consisted of 10 compounds belonging to the 2C-x or DOx classes (class 1), 10 compounds belonging to the NBOMe class (class 2) and 14 compounds with different molecular structures (class 3). The list of the compounds included in the database is presented in Table 4.10.

The purpose of the analysis was to assess to what extent the molecular descriptors from the three groups (topological, charge and RDF) can be used for a successful assignment of the class identity of the 2C-x, DOx and NBOMe compounds and therefore for predicting the psychedelic activity of an unknown.

The evaluation was carried out using PCA, which aimed to:

a) assess the natural (unsupervised) formation of the 2C-x / DOx and NBOMe clusters;
b) asses the relevance of the molecular descriptors for modelling these classes of drugs of abuse and discriminate them from negatives (any other compound);
c) identify the descriptors that are the most important for modeling / discriminating each of these three groups of substances.

Table 4.10 List of the 2C-x/DOx, NBOMe compounds and those belonging to the class of negatives included in the analysis (© 2022 IEEE) [45]

2C-x/DOx compounds	**NBOMe compounds**	**Negatives**
2C-B	25C-NBOMe	Buphedrone
2C-E	25D-NBOMe	Butylone
2C-H	25E-NBOMe	Ciprofloxacin
2C-I	25H-NBOMe	EGFR/ErbB-2 Inhibitor
2C-T-2	25I-N3BOMe	Januvia
2C-T-7	25I-N4BOMe	JWH-018 adamantyl-carboxamide
DOB	25I-NBOMe	JWH-019
DOC	25B-NBOMe	JWH-200
DOET	25C-NB3OMe	JWH-250
2,5-DMA (DOH)	25C-NB4OMe	L_amoxicilin
		LSD
		Nitroaspirin
		Penicilin_v
		Xanax

The *Unscrambler X 10.4* software was used to perform *PCA* [71]. For this purpose, the data was mean centered. The blocks were divided based on the standard deviation. SDV was used as computational algorithm.

The explained variance plot, shown in Figure 4.13, indicates that the first two principal components are responsible for most of the explained variance of the data. PC1 și PC2 account for 78% and 21% of the explained variance, respectively. Therefore, the initial 304 variables can be reduced to just two variables (the first two principal components). Reducing the number of variables allows easier interpretation of the data clustering without losing much information.

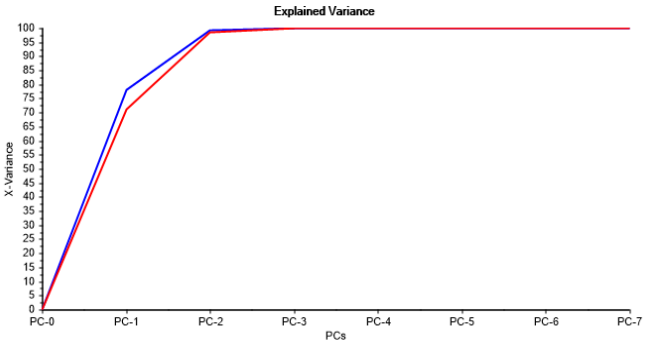

Figure 4.13 Explained variance plot corresponding to the calibration set (blue) and to the validation set (red) (© 2022 IEEE) [45]

The 2D score plot relative to the first two principal components is shown in Figura 4.14. It indicates that the first two modeled classes of drugs of abuse (2C-x/DOx class and NBOMe class) form well-defined clusters. The 2C-x and DOx compounds form the densest cluster, which can be identified by high positive PC1 scores and close to zero negative PC2 scores. Two false positives have been identified, namely butylone and buphedrone, which have been misclassified, probably because their chemical structures are relatively similar to those of the 2C-x or DOx compounds. The NBOMe compounds form a cluster characterised by relatively close to zero PC1 scores and positive PC2 scores. Since their molecular structures are very different, the negatives form a diffuse cloud, which is located in quadrants II and III. The negatives can be easily distinguished from the positives due to the PC1 scores.

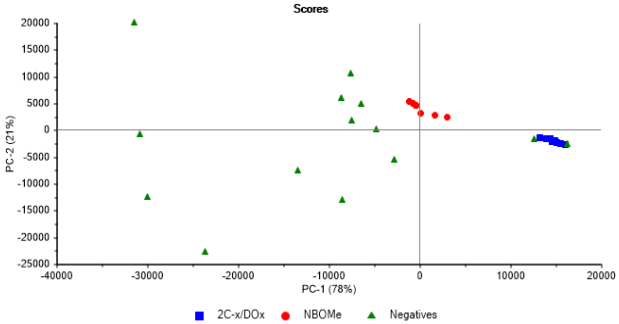

Figure 4.14 2D score plot highlighting the clusters formed by the 2C-x/DOx (blue) and the NBOMe (red) drugs of abuse compared to the negatives (green) (© 2022 IEEE) [45]

Although the variance encapsulated by PC3 is significantly lower than that corresponding to the PC1 and PC2 components, the 3D PCA score plot can be useful for a better visualization, as shown in Figure 4.15.

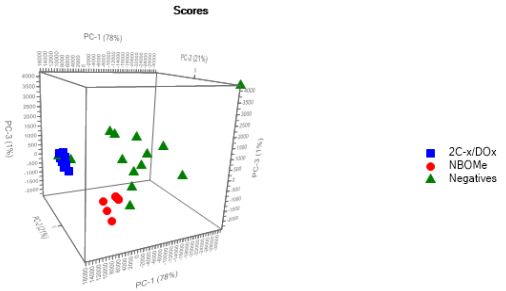

Figure 4.15 3D score plot illustrating the clusters formed formed by the 2C-x/DOx (blue) and the NBOMe (red) drugs of abuse compared to the negatives (green) (© 2022 IEEE) [45]

The loading plot, presented in the all-path Wiener index Wap, the Gutman molecular topological index $GMTI$, the Schultz molecular topological index $SMTI$, the Gutman molecular topological index by valence vertex degrees $GMTIV$ and the Schultz molecular topological index by valence vertex degrees $SMTIV$.

Figure 4.16, shows the contribution of each variable to the modeling and the discrimination of each cluster considered. Based on the analysis of this plot we can identify the input variables with the highest discrimination power, namely the all-path Wiener index Wap, the Gutman molecular topological index $GMTI$,

the Schultz molecular topological index $SMTI$, the Gutman molecular topological index by valence vertex degrees $GMTIV$ and the Schultz molecular topological index by valence vertex degrees $SMTIV$.

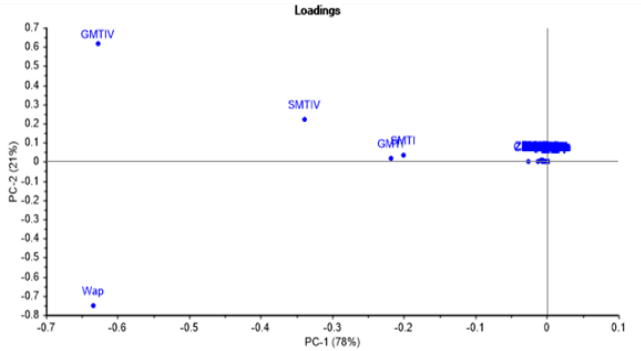

Figure 4.16 Loading plot showing the contribution of the molecular descriptors to the formation of the 2C-x / DOx and NBOMe clusters (© 2022 IEEE) [45]

The analysis carried out based on the topological indices, the charge descriptors and the RDF descriptors shows that they may be successfully used to assign the class identity of the drugs belonging to the 2C-x/DOx and NBOMe classes. The first two modeled classes of drugs (2C-x / DOx and NBOMe classes) form well-defined clusters. However, the results indicate that some misclassifications can be expected, i.e. negatives mistakenly classified as 2C-x/DOx. But from a forensic point of view the most important aspect is that no positive (2C-x/DOx or NBOMe substances) have been misclassified as negative.

In conclusion, the selected molecular descriptors could be successfully used as input into advanced systems used to classify drugs belonging to the targeted classes, including new compounds, and to predict their biological activity.

4.2.4. Use of quantum molecular descriptors and molecular electrostatic potential diagrams for the characterisation of some drugs of abuse belonging to the 2C-x and NBOMe classes [72]

Quantum molecular descriptors and molecular electrostatic potential (MEP) diagrams are useful tools for the physico-chemical characterisation of the compounds. In order to characterise some representative compounds belonging to the 2C-x and NBOMe classes of illicit drugs, their MEP diagrams were determined, along with ten quantum molecular descriptors: the dipole moment (DM), the minimum energy (E_{min}), the energy of the highest occupied molecular orbital (E_{HOMO}), the energy of the lowest unoccupied molecular orbital (E_{LUMO}), the gap energy (E_{gap}), the chemical hardness (η), the chemical softness (σ), the

electronegativity (χ), the chemical potential (μ) and the electrophilicity index (ω) [72].

The targeted compounds are 2-(2,5-dimethoxyphenyl)ethanamine (2C-H), 2-(4-bromo-2,5-dimethoxyphenyl)ethanamine (2C-B), 2-(4-ethyl-2,5-dimethoxyphenyl)ethanamine (2C-E), 2-(2,5-dimethoxyphenyl)-N-[(2-methoxyphenyl)methyl]ethanamine (25H-NBOMe), 2-(4-bromo-2,5-dimethoxyphenyl)-N-[(2-methoxyphenyl)methyl]ethanamine (25B-NBOMe), and 2-(4-ethyl-2,5-dimethoxyphenyl)-N-[(2-methoxyphenyl)methyl]ethanamine (25E-NBOMe).

Some of the quantum molecular descriptors were determined by using the *HyperChem 8.0* software, namely the dipole moment (DM), the minimum energy (E_{min}), the energy of the highest occupied molecular orbital (E_{HOMO}) and the energy of the lowest unoccupied molecular orbital (E_{LUMO}) [73]. The other quantum molecular descriptors were determined based on the formulas presented in the sub-chapter 3.1. The values of the quantum molecular descriptors (rounded to two decimal places) corresponding to the analysed compounds are presented in Table 4.11 and Table 4.12.

Table 4.11 Values of the quantum molecular descriptors corresponding to the targeted compounds [72]

	E_{min} (kcal/mol)	E_{HOMO} (eV)	E_{LUMO} (eV)	E_{gap} (eV)	η (eV)
2C-H	-2783,23	-8,86	0,29	9,15	4,57
2C-B	-2752,17	-9,00	-0,16	8,85	4,42
2C-E	-3346,64	-8,46	0,31	8,77	4,39
25H-NBOMe	-4634,84	-8,51	0,35	8,86	4,43
25B-NBOMe	-4601,15	-8,97	-0,07	8,90	4,45
25E-NBOMe	-5197,14	-8,41	0,34	8,75	4,38

Table 4.12 Values of the quantum molecular descriptors corresponding to the targeted compounds [72]

	σ $(eV)^{-1}$	χ (eV)	μ (eV)	ω (eV)	*DM* (debye)
2C-H	0,11	4,29	-4,29	2,01	2,17
2C-B	0,11	4,58	-4,58	2,37	3,06
2C-E	0,11	4,07	-4,07	1,89	1,46
25H-NBOMe	0,11	4,08	-4,08	1,88	1,70
25B-NBOMe	0,11	4,52	-4,52	2,30	3,26
25E-NBOMe	0,11	4,03	-4,03	1,86	0,73

The chemical hardness ranges from 4.38 to 4.57 eV, while the chemical softness has an approximate value of 0.11 $(eV)^{-1}$ for all the compounds, indicating their good chemical stability. This property is also highlighted by the values of the

energy gap, which range from 8.75 to 9.15 eV. The values of the electrophilicity index are between 1.86 and 2.37 eV, which indicates that all analysed compounds are strong electrophiles [74].

Next, the *HyperChem 8.0* software was used to generate the MEP diagrams for the targeted substances. The MEP diagram of the 2C-B compound is shown in Figure 4.17 [73]. By analysing the MEP diagrams, possible active sites for the electrophilic attack can be identified, corresponding to regions with negative electrostatic potential.

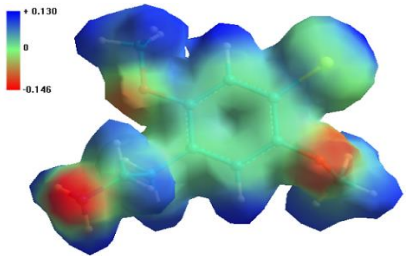

Figure 4.17 MEP diagram of 2-(4-bromo-2,5-dimethoxyphenyl)ethanamine (2C-B) [72]

In conclusion, the quantum molecular descriptors and the molecular electrostatic potential (MEP) diagrams provide relevant information on the physico-chemical properties of the targeted compounds belonging to the 2C-x and NBOMe classes of illicit drugs, thus representing useful tools for their characterisation.

4.2.5. PCA applied for the characterisation and the evaluation of the clustering tendency of some drugs of abuse belonging to the 2C-x, DOx and NBOMe classes based on some molecular descriptors and toxicity parameters [75]

This subchapter presents important toxicological aspects corresponding to some representative drugs belonging to the DOx and NBOMe classes of illicit drugs. More specifically, three molecular toxicity descriptors were determined, namely the Verhaar Fish baseline toxicity ($BLTF96$), the Verhaar Daphnia baseline toxicity ($BLTD48$), the Verhaar Algae baseline toxicity ($BLTA96$), together with three toxicity parameters, namely the median lethal dose ($LD50$), the immunotoxicity and the carcinogenicity. These toxicity indicators were then subjected to PCA to assess the clustering tendency and the potential discrimination between the analysed classes of compounds [75]. The list of the targeted compounds is presented in Table 4.13.

The first three toxicity indicators, namely the $BLTF96$, $BLTD48$ and $BLTA96$ descriptors were determined using the *alvaDesc* software [70]. The values of the $BLTF96$, $BLTD48$ and $BLTA96$ determined for the targeted compounds are given in Table 4.14.

Table 4.13 List of the targeted DOx and NBOMe compounds [75]

Compound name	Compound class
2,5-dimethoxyamphetamine (2,5-DMA)	DOx
2,5-dimethoxy-4-bromoamphetamine (DOB)	DOx
2,5-dimethoxy-4-chloroamphetamine (DOC)	DOx
2,5-dimethoxy-4-ethylamphetamine (DOET)	DOx
2,5-dimethoxy-4-methylamphetamine (DOM)	DOx
25B-NBOMe	NBOMe
25C-NBOMe	NBOMe
25C-NB3OMe	NBOMe
25C-NB4OMe	NBOMe
25D-NBOMe	NBOMe

Table 4.14 Values of the $BLTF96$, $BLTD48$ and $BLTA96$ descriptors corresponding to the targeted compounds [75]

	Compound name	BLTF96	BLTD48	BLTA96
1	2,5-DMA	-2,8	-2,89	-2,89
2	DOC	-3,26	-3,41	-3,44
3	DOET	-3,27	-3,42	-3,44
4	DOM	-3,04	-3,16	-3,17
5	DOB	-3,38	-3,55	-3,57
6	25B-NBOMe	-4,17	-4,43	-4,5
7	25C-NBOMe	-4,08	-4,32	-4,39
8	25C-NB3OMe	-4,08	-4,32	-4,39
9	25C-NB4OMe	-4,08	-4,32	-4,39
10	25D-NBOMe	-3,85	-4,07	-4,12
11	25E-NBOMe	-4,04	-4,28	-4,35
12	25H-NBOMe	-3,65	-3,85	-3,89

The following toxicity indicators, namely the median lethal dose ($LD50$), the immunotoxicity and the carcinogenicity were predicted using *ProTox-II* software [76]. The canonical SMILES (*Simplified Molecular Input Line Entry System*), by which the molecular structures of compounds are represented as strings, were collected from the *PubChem* website and then used as input data for *the ProTox-II* software [76,77]. The results obtained are presented in Table 4.15. For the immunotoxicity and the carcinogenicity, "inactive" was coded as "0", while "active" was assigned the code "1".

Table 4.15 Predicted values of the LD50, immunotoxicity and carcinogenicity corresponding to the targeted compounds [75]

	Comppund name	**LD50 (mg/kg)**	**Immunotoxicity**	**Carcinogenicity**
1	2,5-DMA	171	0	1
2	DOC	330	0	0
3	DOET	330	0	1
4	DOM	330	0	1
5	DOB	400	0	1
6	25B-NBOMe	400	1	0
7	25C-NBOMe	300	1	0
8	25C-NB3OMe	800	1	0
9	25C-NB4OMe	940	1	0
10	25D-NBOMe	300	1	0
11	25E-NBOMe	300	1	0
12	25H-NBOMe	300	1	0

Next, based on the determined toxicity indicators ($BLTF96$, $BLTD48$, $BLTA96$, $LD50$, immunotoxicity and carcinogenicity), PCA was parformed using *Minitab 19* software with its default settings [78]. The scree plot obtained is presented in Figure 4.18. This indicates that the first two principal components (PCs) are enough for future analysis, as they encapsulate a cumulated explained variance of 94% (out of which 81.10% corresponds to PC1).

The score plot, shown in Figure 4.19, indicates that the NBOMe drugs of abuse are characterised by negative PC1 scores, while the DOx illicit compounds have positive PC1 scores. Therefore, these two classes of drugs can be clearly discriminated, and this only based on their PC1 scores. The score plot also indicates that PC2 plays an important role in the discrimination between the NBOMe compounds. Most NBOMe compounds have positive PC2 scores, but there are also two exceptions, namely 25C-NB3OMe and 25C-NB4OMe, which are characterised by negative PC2 scores.

The loading plot, presented in Figure 4.20, indicates the most important variables contributing to the discrimination between the DOx and the NBOMe drugs of abuse. The DOx compounds cluster mainly due to their carcinogenicity resoponses and the values of the BLT descriptors. On the other hand, the NBOMe compounds may be clearly distinguished from DOx compounds,

especially due to their negative PC1 scores generated by their immunotoxicity responses and their $LD50$ values. What distinguishes the NBOMe substances characterised by negative PC2 scores from those that form the cluster located in the quadrant II are their $LD50$ values. The $LD50$ values corresponding to 25C-NB3OMe and 25C-NB4OMe are much (almost three times) higher than the $LD50$ values obtained for the NBOMe compounds grouped in the quadrant II.

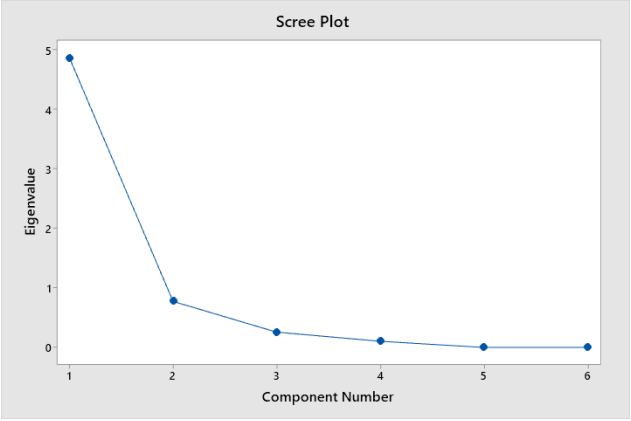

Figure 4.18 Scree plot obtained based on the calculated toxicity indicators [75]

Figure 4.19 2D score plot highlighting the discrimination between the DOx and the NBOMe drugs of abuse based on the determined toxicity indicators [75]

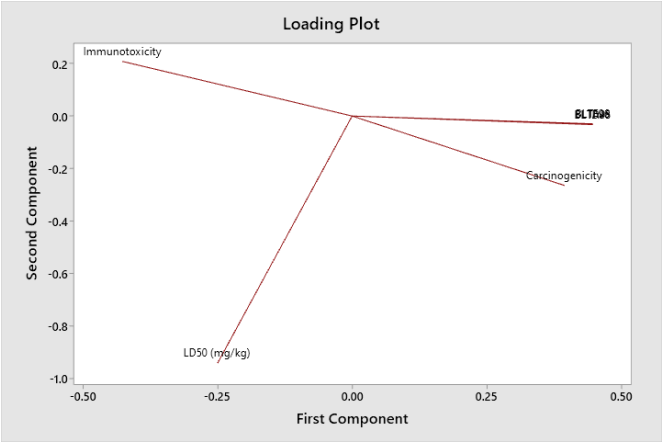

Figure 4.20 Loading plot illustrating the contribution of the toxicity indicators to the formation of the DOx and NBOMe clusters [75]

The computational assessment of the toxicity of the compounds is very important because it indicates in a rapid and cost-effective way which new drugs of abuse pose a higher threat from the public health point of view. The results show that the toxicity indicators chosen are relevant for a clear discrimination of the DOx and the NBOMe illicit drugs. Performing PCA was also useful for assessing which toxicity indicators (and associated medical conditions) ensure this discrimination.

Chapter V. General conclusions and future directions for research and development

General conclusions

The research activity within this doctoral thesis was carried out in two directions:

- research conducted based on the ATR-FTIR spectra of the analysed 2C-x, DOx and NBOMe drugs
- research conducted based on some molecular descriptors and toxicity parameters determined for the analysed 2C-x, DOx and NBOMe drugs

The research conducted based on the ATR-FTIR spectra has primarily resulted in the development of some machine learning systems aimed at detecting 2C-x, DOx and NBOMe illicit drugs, as well as at detecting drugs belonging to other classes of forensic interest, such as opioids and cannabinoids. For the development of these machine learning systems, the following methods

were used: SVM, XGBoost, RF, GB and KNN. Of these, the best results were obtained by SMV and XGBoost. Also, within this research direction, the vibrational analysis of the ATR-FTIR spectra of some drugs belonging to the targeted classes was carried out.

The research based on the molecular descriptors and the toxicity parameters consisted in characterising and assessing the grouping tendency of the representative drugs belonging to the 2C-x, DOx and NBOMe classes. The results showed that the molecular descriptors and the toxicity parameters used have a high discrimination power for the compounds belonging to these three classes of drugs concerned and may be used as input data for future research.

Future directions for research and development

With regard to the future directions for research and development, the following are envisaged:

- the development of the current spectral database by adding the ATR-FTIR spectra of other known 2C-x, DOx and NBOMe drugs of abuse, as they are provided by specialised laboratories;
- the development of the current spectral database by adding the ATR-FTIR spectra of new 2C-x, DOx and NBOMe compounds, as they are produced in clandestine laboratories and subsequently identified, and their spectra are provided by specialised laboratories;
- the development of the database built based on the molecular descriptors and the toxicity parameters by calculating them for other 2C-x, DOx and NBOMe compounds, both known and new compounds produced in clandestine laboratories and subsequently identified;
- the application of other multivariate mathematical methods for the detection of the 2C-x, DOx and NBOMe drugs of abuse, such as ANN or Logistic Regression;
- the development of some QSAR models based on the analysed molecular descriptors in order to predict the pharmacological and/or toxicological properties of the 2C-x, DOx and NBOMe substances;
- the expansion of the research by developing new classification models to detect other types of drugs of abuse.

List of publications

ISI publications

- **ISI (Web of Science) indexed/quoted papers**

1. **Darie, I.-F.**, Anton S. R., Praisler, M., Machine learning systems detecting illicit drugs based on their ATR-FTIR spectra, Inventions 8 (2023) Art. no. 56. DOI: 10.3390/inventions8020056 **WOS:000978413800001 Impact Factor: 3,4**
https://www.mdpi.com/2411-5134/8/2/56

- **ISI (Web of Science) in progress indexing papers**

1. **Darie, I.-F.**, Gosav, S., Praisler, M., Characterisation of novel illicit drugs based on computational toxicology, 2023 International Conference on E-Health and Bioengineering (EHB 2023), 11th Edition, 2023, SpringerLink, *in press* http://www.ehbconference.ro/Home.aspx

2. **Darie, I.-F.**, Praisler, M., Principal Component Analysis Assessing the Potential Clustering of 2C-x and DOx amphetamines, 2022 International Conference on E-Health and Bioengineering (EHB 2022), 10th Edition, 2022, IEEE, pp. 1-4, DOI: 10.1109/EHB55594.2022.9991592, Electronic ISBN:978-1-6654-8557-9 Print on Demand(PoD) ISBN:978-1-6654-8558-6, Electronic ISSN: 2575-5145, Print on Demand(PoD) ISSN: 2575-5137
https://ieeexplore.ieee.org/abstract/document/9991592
http://www.ehbconference.ro/2022/Home.aspx

BDI publications

1. **Darie, I.-F.**, Gosav, S., Praisler, M., Evaluation of physico-chemical parameters of some psychoactive compounds using molecular modeling Annals of "Dunarea de Jos" University of Galati, Mathematics, Physics, Theoretical Mechanics, Fascicle II, Year XV (XLVI) 2023, No. 1, pp. 1 - 4. DOI: 35219/ann-ugal-math-phys-mec.2023.1.01 ISSN 2067-2071
https://www.gup.ugal.ro/ugaljournals/index.php/math/article/view/6232

2. **Darie, I.-F.**, Praisler, M., Vibrational analysis of 2C-x and DOx hallucinogenic amphetamines, Annals of "Dunarea de Jos" University of Galati, Mathematics, Physics, Theoretical Mechanics, Fascicle II, Year XIV (XLV) 2022, 45(1), pp. 5 – 10. DOI: 10.35219/ann-ugal-math-phys-mec.2022.1.02 ISSN 2067-2071
https://www.gup.ugal.ro/ugaljournals/index.php/math/article/view/5508

3. **Darie, I.-F.**, Praisler, M., Molecular descriptors - a mathematical approach to characterize psychedelic substances, Annals of "Dunarea de Jos" University of Galati, Mathematics, Physics, Theoretical Mechanics, Fascicle II, Year XIV (XLV) 2022, 45(1), pp. 1 – 4. DOI: 10.35219/ann-ugal-math-phys-mec.2022.1.01 ISSN 2067-2071
https://www.gup.ugal.ro/ugaljournals/index.php/math/article/view/5507

4. Negoiță, C., Praisler, M., **Darie, I.-F.**, Automatic identification of haluginogenic amphetamines based on their ATR-FTIR spectra processed with Convolutional Neural Networks, 9th International Multidisciplinary Symposium "UNIVERSITARIA SIMPRO", 27 – 28 May 2021, Petrosani, Romania, MATEC Web of Conferences 342 (2021) 05003. DOI: 10.1051/matecconf/202134205003
https://www.matec-conferences.org/articles/matecconf/abs/2021/11/matecconf_simpro21_05003/matecconf_simpro21_05003.html
https://www.upet.ro/simpro/2021/

5. **Darie, I.-F.**, Praisler, M., Negoiță, C., 2C-x and DOx hallucinogens: a systematic review, Annals of "Dunarea de Jos" University of Galati, Mathematics, Physics, Theoretical Mechanics, Fascicle II, Year XIII (XLIV) 2021, No. 1, pp. 46-52. DOI: 10.35219/ann-ugal-math-phys-mec.2021.1.07 ISSN 2067-2071
https://gup.ugal.ro/ugaljournals/index.php/math/article/view/4925

Papers communicated at international conferences

1. **Darie, I.-F.**, Gosav, S., Praisler, M., Characterisation of novel illicit drugs based on computational toxicology, 2023 International Conference on E-Health and Bioengineering (EHB 2023), 11th Edition, 9-10 November 2023, Bucharest
http://www.ehbconference.ro/Home.aspx

2. **Darie, I.-F.**, Gosav, S., Praisler, M., Computational assessment of the toxicity of new psychoactive controlled substances, International Conference and Workshop "Interdisciplinary applications of advanced analytical and control techniques in environment, health and materials science - INTERVENT", 19-20 October 2023, Galati
https://www.researchgate.net/publication/375548506_Abstract_Book_-_International_Conference_and_Workshop_Interdisciplinary_applications_of_advanced_analytical_and_control_techniques_in_environment_health_and_materials_science_-_INTERVENT_19-20_October
https://www.researchgate.net/publication/373217117_Interdisciplinary_applications_of_advanced_analytical_and_control_techniques_in_environment_health_and_materials_science_-_INTERVENT

3. **Darie, I.-F.**, Anton, S. R., Praisler, M., Drug detection using K-Nearest Neighbors algorithm, International Symposium & International Student Workshop on Interdisciplinary Mathematics in the CiTi areas, *ISIM & ISWIM 2023*, 26-29 June, 2023, Bucharest
https://www.researchgate.net/publication/377160085_isim2023_Book_of_Abstracts
https://www.isimconference.eu/

4. **Darie, I.-F.**, Praisler, M., Principal Component Analysis Assessing the Potential Clustering of 2C-x and DOx amphetamines, 2022 International Conference on E-Health and Bioengineering (EHB 2022), 10th Edition, 17-19 November 2022, Iasi
https://ieeexplore.ieee.org/abstract/document/**9991592**
http://www.ehbconference.ro/2022/Home.aspx

5. **Darie, I.-F.**, Praisler, M., S. Gosav, Topological indices – application of graph theory in chemistry, International Symposium & International Student Workshop on Interdisciplinary Mathematics in the CiTi areas, *ISIM & ISWIM 2022,* 26-29 June 2022, Bucharest
https://www.researchgate.net/publication/367008968_1st_International_Symposium_International_Student_Workshop_on_Interdisciplinary_Mathematics_in_the_CiTi_areas_ISIM_ISWIM_2022
https://www.isimconference.eu/index.php

6. Negoiță, C., Praisler, M., **Darie, I.-F.**, Automatic identification of haluginogenic amphetamines based on their ATR-FTIR spectra processed with Convolutional Neural Networks, 9th International Multidisciplinary Symposium "UNIVERSITARIA SIMPRO", 27 – 28 May 2021, Petrosani
https://www.matec-conferences.org/articles/matecconf/abs/2021/11/matecconf_simpro21_05003/matecconf_simpro21_05003.html
https://www.upet.ro/simpro/2021/

Papers communicated at national conferences

1. **Darie, I.-F.**, Gosav, S., Praisler, M., Evaluation of physico-chemical parameters of some psychoactive compounds using molecular modeling, *Book of abstracts SCDS-UDJG 2023, 11th Edition*, 8 - 9 June 2023, Galati
http://www.cssd-udjg.ugal.ro/

2. **Darie, I.-F.**, Praisler, M., Toxicity assessment of some DOx, 2C-x and NBOMe psychoactive compounds using dedicated models, *Book of abstracts SCDS-UDJG 2023, 11h Edition*, 8 - 9 June 2023, Galati
http://www.cssd-udjg.ugal.ro/

3. **Darie, I.-F.**, Praisler, M., Vibrational analysis of 2C-x and DOx hallucinogenic amphetamines, *Book of abstracts SCDS-UDJG 2022, 10th Edition*, 9 - 10 June 2022, Galati
http://www.cssd-udjg.ugal.ro/

4. **Darie, I.-F.**, Praisler, M., Molecular descriptors - a mathematical approach to characterize psychedelic substances, *Book of abstracts SCDS-UDJG 2022, 10th Edition*, 9 - 10 June 2022, Galati
http://www.cssd-udjg.ugal.ro/

5. **Darie, I.-F.**, Praisler, M., Negoiță, C., 2C-x and DOx hallucinogens: a systematic review, *Book of abstracts SCDS-UDJG 2021, 9th Edition*, 10 - 11 June 2021, Galati
http://www.cssd-udjg.ugal.ro/

6. **Darie, I.-F.**, Gosav, S., Praisler, M., Exploratory analysis of hallucinogenic amphetamines based on their ATR-FTIR spectra and Principal Component Analysis, *Book of abstracts SCDS-UDJG 2021, 9th Edition*, 10 - 11 June 2021, Galati
http://www.cssd-udjg.ugal.ro/

Awards

1. First prize obtained at the *Excellence Research Gala - CEREX UDJG* for excellent results in research activity within IOSUD-UDJG, 6 December 2023, Galati
https://www.ugal.ro/anunturi/stiri-si-evenimente/12915-gala-cerex-a-ajuns-la-a-patra-editie

2. First prize at the International Competition *Ma thèse en 180 secondes*, Polytechnic University of Bucharest, for the presentation in 3 minutes of the doctoral thesis entitled "Les méthodes mathématiques multivariées pour dépister les drogues", 3 June 2022, Bucharest
https://www.auf.org/europe-centrale-orientale/nouvelles/actualites/iulia-darie-laureate-de-la-finale-nationale-de-ma-en-180-secondes-roumanie-edition-2022/

3. Third prize obtained within the Scientific Conference of Doctoral Schools of University "Dunarea de Jos" Galati (SCDS-UDJG), **Darie, I.-F.**, Gosav, S., Praisler, M., Evaluation of physico-chemical parameters of some psychoactive compounds using molecular modeling, *Book of abstracts SCDS-UDJG 2023, 11th Edition*, 8 - 9 June 2023, Galati
http://www.cssd-udjg.ugal.ro/

Project participations

1. Participation in the project *Academic and social internationalization of students at "Dunarea de Jos" University of Galati - IDEI (UGAL)22*, 2022.

Selective references

[1] European Monitoring Centre for Drugs and Drug Addiction, Situația drogurilor în Europa până în 2023, Publications Office, LU, 2023. https://data.europa.eu/doi/10.2810/8277 (accessed 3 February 2024).

[2] I.-F. Darie, M. Praisler, C. Negoita, 2C-x and DOx hallucinogens: a systematic review, 44 (2021) 46–52. https://doi.org/10.35219/ann-ugal-math-phys-mec.2021.1.07.

[3] S.L. Hill, S.H.L. Thomas, Clinical toxicology of newer recreational drugs, Clin. Toxicol. (Phila) 49 (2011) 705–719. https://doi.org/10.3109/15563650.2011.615318.

[4] B.V. Dean, S.J. Stellpflug, A.M. Burnett, K.M. Engebretsen, 2C or not 2C: phenethylamine designer drug review, J. Med. Toxicol. 9 (2013) 172–178. https://doi.org/10.1007/s13181-013-0295-x.

[5] A. Shulgin, A. Shulgin, PIHKAL: A Chemical Love Story, First Edition, Transform Press, Berkeley, 1990.

[6] E.S. Herrmann, P.S. Johnson, M.W. Johnson, R. Vandrey, Novel Drugs of Abuse, in: Neuropathology of Drug Addictions and Substance Misuse, Elsevier, 2016: pp. 893–902. https://doi.org/10.1016/B978-0-12-800634-4.00088-3.

[7] D. de Boer, I. Bosman, A new trend in drugs-of-abuse; the 2C-series of phenethylamine designer drugs, Pharm. World Sci. 26 (2004) 110–113. https://doi.org/10.1023/b:phar.0000018600.03664.36.

[8] National Drug Intelligence Center. Department of Justice. Washington, D.C., Information Bulletin 2C-B (Nexus) Reappears on the Club Drug Scene, (2001). https://www.justice.gov/archive/ndic/pubs0/665/index.htm (accessed 24 September 2023).

[9] Sen T. Harkin, Text - S.3187 - 112th Congress (2011-2012): Food and Drug Administration Safety and Innovation Act, (2012). https://www.congress.gov/bill/112th-congress/senate-bill/3187/text (accessed 25 October 2023).

[10] Council Decision on 2C-I, 2C-T-2, 2C-T-7 and TMA-2 | www.emcdda.europa.eu, (n.d.). https://www.emcdda.europa.eu/drugs-library/2003847jha-council-decision-2003847jha-27-november-2003-concerning-control-measures-and-criminal-sanctions-respect-new-synthetic-drugs-2c-i-2c-t-2-2c-t-7-and-tma-2_en (accessed 25 October 2023).

[11] Drug Scheduling, (n.d.). https://www.dea.gov/drug-information/drug-scheduling (accessed 25 October 2023).

[12] M. McElwain, #TestIt Alert: Misrepresented 'Beatles Blotter' Sold as 2C-B in Saskatchewan, Canada | DanceSafe, (2019). https://dancesafe.org/testit-alert-misrepresented-beatles-blotter-sold-as-2c-b-in-saskatchewan-canada/ (accessed 25 October 2023).

[13] X.W.S. Chia, M.C. Ong, Y.Y.C. Yeo, Y.J. Ho, E.I. Binte Ahmad Nasir, L.-L.J. Tan, P.Y. Chua, T.W.A. Yap, J.L.W. Lim, Simultaneous analysis of 2Cs, 25-NBOHs, 25-NBOMes and LSD in seized exhibits using liquid chromatography–tandem mass spectrometry: A targeted approach, Forensic Sci. Int. 301 (2019) 394–401. https://doi.org/10.1016/j.forsciint.2019.05.036.

[14] L.E. Hollister, M.F. Macnicol, H.K. Gillespie, An hallucinogenic amphetamine analog (DOM) in man, Psychopharmacologia 14 (1969) 62–73. https://doi.org/10.1007/BF00401535.

[15] C. Furnari, V. Ottaviano, F. Rosati, [Determination of 4-bromo-2,5-dimethoxyamphetamine (DOB) found in illicit tablets seized in Italy], Ann. Ist. Super. Sanita 37 (2001) 297–300.

[16] L. Burrai, M. Nieddu, M. Palomba, M.A. Pirisi, Identification and quantitation of 4-bromo-2,5-dimethoxyamphetamine in seized blotters, Leg. Med. 17 (2015) 56– 59. https://doi.org/10.1016/j.legalmed.2014.09.001.

[17] 4-Chloro-2,5-dimethoxyamphetamine, (n.d.). https://www.chemeurope.com/en/encyclopedia/4-Chloro-2%2C5-dimethoxyamphetamine.html (accessed 25 October 2023).

[18] A. Rickli, D. Luethi, J. Reinisch, D. Buchy, M.C. Hoener, M.E. Liechti, Receptor interaction profiles of novel N-2-methoxybenzyl (NBOMe) derivatives of 2,5-dimethoxy-substituted phenethylamines (2C drugs), Neuropharmacology 99 (2015) 546–553. https://doi.org/10.1016/j.neuropharm.2015.08.034.

[19] M.H.Y. Tang, C.K. Ching, M.S.H. Tsui, F.K.C. Chu, T.W.L. Mak, Two cases of severe intoxication associated with analytically confirmed use of the novel psychoactive substances 25B-NBOMe and 25C-NBOMe, Clin. Toxicol. (Phila) 52 (2014) 561–565. https://doi.org/10.3109/15563650.2014.909932.

[20] M.R. Braden, J.C. Parrish, J.C. Naylor, D.E. Nichols, Molecular Interaction of Serotonin 5-HT 2A Receptor Residues Phe339 (6.51) and Phe340 (6.52) with Superpotent N -Benzyl Phenethylamine Agonists, Mol. Pharmacol. 70 (2006) 1956–1964. https://doi.org/10.1124/mol.106.028720.

[21] D. Pasin, A. Cawley, S. Bidny, S. Fu, Characterization of hallucinogenic phenethylamines using high-resolution mass spectrometry for non-targeted screening purposes, Drug Test. Anal. 9 (2017) 1620–1629. https://doi.org/10.1002/dta.2171.

[22] M. Hansen, Design and Synthesis of Selective Serotonin Receptor Agonists for Positron Emission Tomography Imaging of the Brain (PhD), 2010. https://doi.org/10.13140/RG.2.2.33671.14245.

[23] R. Heim, Synthese und Pharmakologie potenter 5-HT2A-Rezeptoragonisten mit N-2 -Methoxybenzyl-Partialstruktur: Entwicklung eines neuen Struktur-Wirkungskonzepts, (2004). https://doi.org/10.17169/refubium-16193.

[24] R.A. Glennon, M. Dukat, M. el-Bermawy, H. Law, J. De los Angeles, M. Teitler, A. King, K. Herrick-Davis, Influence of amine substituents on 5-HT2A versus 5-HT2C binding of phenylalkyl- and indolylalkylamines, J. Med. Chem. 37 (1994) 1929–1935. https://doi.org/10.1021/jm00039a004.

[25] J.B. Zawilska, M. Kacela, P. Adamowicz, NBOMes–Highly Potent and Toxic Alternatives of LSD, Front. Neurosci. 14 (2020). https://doi.org/10.3389/fnins.2020.00078
[26] M. Herian, A. Wojtas, M. Maćkowiak, A. Wawrzczak-Bargiela, A. Solarz, A. Bysiek, K. Madej, K. Gołembiowska, Neurotoxicological profile of the hallucinogenic compound 25I-NBOMe, Sci. Rep. 12 (2022) 2939. https://doi.org/10.1038/s41598-022-07069-8.
[27] I. Papoutsis, P. Nikolaou, M. Stefanidou, C. Spiliopoulou, S. Athanaselis, 25B-NBOMe and its precursor 2C-B: modern trends and hidden dangers, Forensic Toxicol. 33 (2015) 1–11. https://doi.org/10.1007/s11419-014-0242-9.
[28] B.P. dos Santos, L. Birk, P. de Souza Schwarz, S. Eller, T.F. de Oliveira, M.D. Arbo, A Comprehensive Analysis of Legislative Strategies for New Psychoactive Substances: The Brazilian Panorama, Psychoactives 2 (2023) 242–255. https://doi.org/10.3390/psychoactives2030016.
[29] J. Suzuki, M.A. Dekker, E.S. Valenti, F.A. Arbelo Cruz, A.M. Correa, J.L. Poklis, A. Poklis, Toxicities associated with NBOMe ingestion-a novel class of potent hallucinogens: a review of the literature, Psychosomatics 56 (2015) 129–139. https://doi.org/10.1016/j.psym.2014.11.002.
[30] J.L. Poklis, S.A. Raso, K.N. Alford, A. Poklis, M.R. Peace, Analysis of 25I-NBOMe, 25B-NBOMe, 25C-NBOMe and Other Dimethoxyphenyl- N -[(2-Methoxyphenyl) Methyl]Ethanamine Derivatives on Blotter Paper, J. Anal. Toxicol. 39 (2015) 617–623. https://doi.org/10.1093/jat/bkv073.
[31] European Monitoring Centre for Drugs and Drug Addiction., 25I-NBOMe: report on the risk assessment of 2 (4 iodo 2,5 dimethoxyphenyl) N (2 methoxybenzyl) ethanamine (25I NBOMe) in the framework of the Council Decision on new psychoactive substances., Publications Office, LU, 2014. https://data.europa.eu/doi/10.2810/5725 (accessed 23 October 2023).
[32] M.B. Forrester, NBOMe Designer Drug Exposures Reported to Texas Poison Centers, J. Addict. Dis. 33 (2014) 196–201. https://doi.org/10.1080/10550887.2014.950027.
[33] D. Martins, M.J. Barratt, C.V. Pires, H. Carvalho, M.V. Vilamala, I.F. Espinosa, H. Valente, The detection and prevention of unintentional consumption of DOx and 25x-NBOMe at Portugal's Boom Festival, Hum. Psychopharm. Clin. 32 (2017) e2608. https://doi.org/10.1002/hup.2608.
[34] A. Dutta, Fourier Transform Infrared Spectroscopy, in: Spectroscopic Methods for Nanomaterials Characterization, Elsevier, 2017: pp. 73–93. https://doi.org/10.1016/B978-0-323-46140-5.00004-2.
[35] IR spectroscopy Hsu.pdf, (n.d.). https://mmrc.caltech.edu/FTIR/Literature/General/IR%20spectroscopy%20Hsu.pdf (accessed 6 November 2023).
[36] FTIR Sample Techniques: Attenuated Total Reflection (ATR) - RO, (n.d.). https://www.thermofisher.com/tr/en/home/industrial/spectroscopy-elemental-isotope-analysis/spectroscopy- elemental- isotope- analysis-learning- center/molecular- spectroscopy-information/ftir-information/ftir-sample-handling-techniques/ftir-sample-handling-techniques-attenuated-total-reflection-atr.html (accessed 18 November 2023).

[37] A.A. Bunaciu, H. Aboul-Enein, Vibrational Spectroscopy Applications in Drugs Analysis, in: Reference Module in Chemistry, Molecular Sciences and Chemical Engineering, 2016. https://doi.org/10.1016/B978-0-12-409547-2.12214-0.

[38] I.-F. Darie, M. Praisler, Vibrational analysis of hallucinogenic amphetamines, Analele Universității ”Dunarea de Jos” Din Galați. Fascicula II, Matematică, Fizică, Mecanică Teoretică / Annals of the ”Dunarea de Jos” University of Galati. Fascicle II, Mathematics, Physics, Theoretical Mechanics 45 (2022) 5–10. https://doi.org/10.35219/ann-ugal-math-phys-mec.2022.1.02.

[39] Infrared Spectroscopy, Chemistry LibreTexts (2013). https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Spectroscopy/Vibrational_Spectroscopy/Infrared_Spectroscopy/Infrared_Spectroscopy (accessed 4 November 2023).

[40] R. Parthasarathi, A. Dhawan, Chapter 5 - In Silico Approaches for Predictive Toxicology, in: A. Dhawan, S. Kwon (Eds.), In Vitro Toxicology, Academic Press, 2018: pp. 91–109. https://doi.org/10.1016/B978-0-12-804667-8.00005-5.

[41] J.C. Copeland, L.J. Zehr, R.L. Cerny, R. Powers, The Applicability of Molecular Descriptors for Designing an Electrospray Ionization Mass Spectrometry Compatible Library for Drug Discovery, Comb. Chem. High Throughput Screen. 15 (2012) 806–815.

[42] B. Chandrasekaran, S.N. Abed, O. Al-Attraqchi, K. Kuche, R.K. Tekade, Chapter 21 - Computer-Aided Prediction of Pharmacokinetic (ADMET) Properties, in: R.K. Tekade (Ed.), Dosage Form Design Parameters, Academic Press, 2018: pp. 731–755. https://doi.org/10.1016/B978-0-12-814421-3.00021-X.

[43] I.-F. Darie, M. Praisler, S. Gosav, Molecular descriptors - a mathematical approach to characterize psychedelic substances, 45 (2022) 1–4. https://doi.org/10.35219/ann-ugal-math-phys-mec.2022.1.01.

[44] E. Deconinck, C. Duchateau, M. Balcaen, L. Gremeaux, P. Courselle, Chemometrics and infrared spectroscopy – A winning team for the analysis of illicit drug products, Rev. Anal. Chem. 41 (2022) 228–255. https://doi.org/10.1515/revac-2022-0046.

[45] I.-F. Darie, M. Praisler, Principal Component Analysis Assessing the Potential Clustering of 2C-x and DOx Amphetamines, in: 2022 E-Health and Bioengineering Conference (EHB), 2022: pp. 01–04. https://doi.org/10.1109/EHB55594.2022.9991592.

[46] I.-F. Darie, S.-R. Anton, M. Praisler, Machine Learning Systems Detecting Illicit Drugs Based on Their ATR-FTIR Spectra, Inventions 8 (2023) 56. https://doi.org/10.3390/inventions8020056.

[47] 1.11. Ensembles: Gradient boosting, random forests, bagging, voting, stacking, Scikit-Learn (n.d.). https://scikit-learn.org/stable/modules/ensemble.html (accessed 18 December 2023).

[48] A. Natekin, A. Knoll, Gradient boosting machines, a tutorial, Front. neurorobot. 7 (2013). https://doi.org/10.3389/fnbot.2013.00021.

[49] R. Raut, S. Krit, P. Chatterjee, Machine Vision for Industry 4.0: Applications and Case Studies, CRC Press, 2022.

[50] C. Bentéjac, A. Csörgő, G. Martínez-Muñoz, A comparative analysis of gradient boosting algorithms, Artif. Intell. Rev. 54 (2021) 1937–1967. https://doi.org/10.1007/s10462-020-09896-5.
[51] D. Tarwidi, S.R. Pudjaprasetya, D. Adytia, M. Apri, An optimized XGBoost-based machine learning method for predicting wave run-up on a sloping beach, MethodsX 10 (2023) 102119. https://doi.org/10.1016/j.mex.2023.102119.
[52] X. Wu, V. Kumar, The Top Ten Algorithms in Data Mining, CRC Press, 2009.
[53] I.-F. Darie, S. R. Anton, M. Praisler, Drug detection using K-Nearest Neighbors algorithm, in: International Symposium & International Student Workshop on Interdisciplinary Mathematics in the CiTi areas ISIM & ISWIM 2023, 2023
[54] S. Kerrigan, B.A. Goldberger, Opioids, in: B.S. Levine, S. KERRIGAN (Eds.), Principles of Forensic Toxicology, Springer International Publishing, Cham, 2020: pp. 347–369. https://doi.org/10.1007/978-3-030-42917-1_22.
[55] C. Zöllner, C. Stein, Opioids, Handb. Exp. Pharmacol. (2007) 31–63. https://doi.org/10.1007/978-3-540-33823-9_2.
[56] SWGDRUG Infrared Library, (n.d.). https://swgdrug.org/ir.htm (accessed 6 February 2024).
[57] G.J. Harry, Chapter One - Cannabinoids, in: W. Slikker, M. Aschner, L.G. Costa (Eds.), Advances in Neurotoxicology, Academic Press, 2022: pp. 1–48. https://doi.org/10.1016/bs.ant.2022.06.001.
[58] V.Y. Shi, J.L. Hsiao, M.A. Lowes, I.H. Hamzavi, A Comprehensive Guide to Hidradenitis Suppurativa, Elsevier Health Sciences, 2021.
[59] K. Piorunska-Sedlak, K. Stypulkowska, Strategy for identification of new psychoactive substances in illicit samples using attenuated total reflectance infrared spectroscopy, Forensic Sci. Int. 312 (2020) 110262. https://doi.org/10.1016/j.forsciint.2020.110262.
[60] R. Rodríguez-Pérez, J. Bajorath, Evolution of Support Vector Machine and Regression Modeling in Chemoinformatics and Drug Discovery, J. Comput. Aided Mol. Des. 36 (2022) 355–362. https://doi.org/10.1007/s10822-022-00442-9.
[61] C. Wade, K. Glynn, Hands-On Gradient Boosting with XGBoost and scikit-learn: Perform accessible machine learning and extreme gradient boosting with Python, Packt Publishing Ltd, 2020.
[62] C. Smith, Decision Trees and Random Forests: A Visual Introduction for Beginners, Blue Windmill Media, 2017.
[63] J. Vitola, F. Pozo, D.A. Tibaduiza, M. Anaya, A Sensor Data Fusion System Based on k-Nearest Neighbor Pattern Classification for Structural Health Monitoring Applications, Sensors 17 (2017) 417. https://doi.org/10.3390/s17020417.
[64] Recommended methods for the identification and analysis of amphetamine, methamphetamine and their ring-substituted analogues in seized materials, (n.d.). https://www.unodc.org/unodc/en/scientists/recommended-methods-for-the-identification-and-analysis-of-amphetamine--methamphetamine-and-their-ring-substituted-analogues-in-seized-materials.html (accessed 8 February 2024).

[65] Recommended Methods for the Identification and Analysis of Synthetic Cannabinoid Receptor Agonists in Seized Materials, (n.d.). https://www.unodc.org/unodc/en/scientists/recommended-methods-for-the-identification-and-analysis-of-synthetic-cannabinoid-receptor-agonists-in-seized-materials.html (accessed 8 February 2024).
[66] Infrared Spectroscopy Absorption Table, Chemistry LibreTexts (2014). https://chem.libretexts.org/Ancillary_Materials/Reference/Reference_Tables/Spectroscopic_Reference_Tables/Infrared_Spectroscopy_Absorption_Table (accessed 14 February 2024).
[67] IR: aromatics, (n.d.). https://orgchemboulder.com/Spectroscopy/irtutor/aromaticsir.shtml (accessed 14 February 2024).
[68] Dragon Professional Version 5.5, Talete srl (2007) (http://www.talete.mi.it/dragon.htm)., (n.d.).
[69] I.-F. Darie, M. Praisler, S. Gosav, Topological indices - application of graph theory in chemistry, in: 1st International Symposium & International Student Workshop on Interdisciplinary Mathematics in the CiTi Areas ISIM & ISWIM 2022, 2022.
[70] A. Mauri, alvaDesc: A Tool to Calculate and Analyze Molecular Descriptors and Fingerprints, in: K. Roy (Ed.), Ecotoxicological QSARs, Springer US, New York, NY, 2020: pp. 801–820. https://doi.org/10.1007/978-1-0716-0150-1_32.
[71] The Unscrambler X, Version 10.4, Camo software., (n.d.).
[72] I.-F. Darie, S. Gosav, M. Praisler, Evaluation of physico-chemical parameters of some psychoactive compounds based on molecular modeling, Analele Universității „Dunărea de Jos” Din Galați. Fascicula II, Matematică, Fizică, Mecanică Teoretică / Annals of the ”Dunarea de Jos” University of Galati. Fascicle II, Mathematics, Physics, Theoretical Mechanics 46 (2023) 1–4. https://doi.org/10.35219/ann-ugal-math-phys-mec.2023.1.01.
[73] The Hyperchem software, Version 8.0.3., Hypercube, Inc, USA., (n.d.).
[74] P. Pérez, L. Domingo, A. Aizman, R. Contreras, The Electrophilicity Index in Organic Chemistry, Theor. Comput. Chem.19 (2007) 139–201. https://doi.org/10.1016/S1380-7323(07)80010-0.
[75] I.-F. Darie, S. Gosav, M. Praisler, Characterisation of novel illicit drugs based on computational toxicology, in: 2023 E-Health and Bioengineering Conference (EHB), SpringerLink, in press, 2023.
[76] P. Banerjee, A.O. Eckert, A.K. Schrey, R. Preissner, ProTox-II: a webserver for the prediction of toxicity of chemicals, Nucleic Acids Res. 46 (2018) W257–W263. https://doi.org/10.1093/nar/gky318.
[77] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B.A. Shoemaker, P.A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, E.E. Bolton, PubChem 2023 update, Nucleic Acids Res. 51 (2023) D1373–D1380. https://doi.org/10.1093/nar/gkac956.
[78] S.A. Lesik, Applied Statistical Inference with MINITAB®, Second Edition, CRC Press, 2018.