CHUKWUEBUKA IKPEAMA C

THE THESIS TITLE

MASTER/PhD THESIS

YEAR

NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF CHEMISTRY

ELECTROCHEMICAL DETERMINATION OF DOPAMINE IN THE PRESENCE OF ASCORBIC ACID USING MACHINE LEARNING TECHNIQUE

M.Sc. THESIS

CHUKWUEBUKA IKPEAMA C

Nicosia

September, 2023

NEAR EAST UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
DEPARTMENT OF CHEMISTRY

ELECTROCHEMICAL DETERMINATION OF DOPAMINE IN THE PRESENCE OF ASCORBIC ACID USING MACHINE LEARNING TECHNIQUE

M.Sc. THESIS

CHUKWUEBUKA IKPEAMA C

Supervisor
Prof. Dr. Süleyman ASIR
Co-supervisor
Prof. Kamıl Dımılıler

Nicosia
September, 2023

Approval

We certify that we have read the thesis submitted by Ikpeama Chukwuebuka C titled **"Electrochemical determination of dopamine in the presence of ascorbic acid using machine learning technique"** and that in our combined opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Examining Committee	Name-Surname	Signature
Head of the Committee:	Assoc. Prof. Dr. Boran Sekeroglu	.....
Committee Member*:	Assoc. Prof. Dr. Elbrus Imanov	.....
Committee Member*:	Assoc. Prof. Dr. Aydin Hassani	.....
Supervisor:	Assoc. Prof. Dr. Suleyman Asir	.....
Co-supervisor	Assoc. Prof. Dr. Kamil Dimililer	.....

Approved by the Head of the Department

24/07/2023

.....

Assoc Prof. Dr. Süleyman Aşır

Title, Name-Surname

Head of Department

Approved by the Institute of Graduate Studies

...../...../20...

Prof. Dr. Kemal Hüsnü Can Başer

Head of the Institute

Declaration

I hereby declare that all information, documents, analysis and results in this thesis have been collected and presented according to the academic rules and ethical guidelines of Institute of Graduate Studies, Near East University. I also declare that as required by these rules and conduct, I have fully cited and referenced information and data that are not original to this study.

Name and Surname of the Student

…../…../…..

Day/Month/Year

Acknowledgments

I would like to express my sincere gratitude to my thesis advisor, Prof. Dr. Süleyman ASIR and Prof. Kamıl Dımılıler for theır guidance, support, and encouragement throughout my research. I would also like to thank the members of my thesis committee for the valuable comments and suggestions, I am grateful to **Near East University** for providing me with the necessary resources to complete my research. I would like to thank my fiancée, family, and friends for their unwavering support and encouragement throughout my academic journey. Finally, I would like to thank God, for letting me get through all the difficulties. I have experienced your guidance day by day. You are the one who let me finish my degree. I will keep on trusting you for my future.

Chukwuebuka Ikpeama C

ABSTRACT

The accurate and reliable detection of dopamine in biological samples is of great importance for the diagnosis and treatment of neurological disorders. In this study, we created a way to use a pencil graphite electrode and machine learning to detect dopamine while ascorbic acid is present. We used cyclic voltammetry (CV) to analyze the electrochemical behavior of dopamine and ascorbic acid, and found that the PGE was able to selectively detect dopamine in the presence of ascorbic acid. We also used four machine learning algorithms, namely ADABOOSTCLASS, BPNNCLASS, SVM, and GBA, to improve the accuracy of the detection and reduce the interference from ascorbic acid. The results indicated that the limit of detection (LOD) and limit of quantification (LOQ) for dopamine were 4.76 mM and 15.87 mM, respectively. The electrochemical method and machine learning algorithms developed in this study have important implications for the development of more accurate and reliable methods for detecting neurotransmitters and other biomolecules in biological samples.

Keywords**:** Dopamine, Ascorbic acid, PGE, CV, Machine learning.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

(Fe^{+2})	Ferrous ions
($K_4Fe(CN)_6$)	Potassium ferrocyanide
AA	Ascorbic acid
ADABOOSTCLASS	Adaptive Boosting Classifier
Ag/AgCl	Silver silver chloride
AI	Artificial intelligence
BPNN	Back Propagation Neural network
CE	Counter electrode
C-N	Cabon-nitrogen
CNN	Convolutional neural network
CV	Cyclic voltammetry
DA	Dopamine
DNA	Deoxyribonucleic acid
DPV	differential pulse voltammetry
DT	Decision tree
GBA	Gradient Boosting Algorithm
H_2O_2	Hydrogen peroxide
IUPAC	International union of pure and applied science
K_2HPO_4	Di-potassium hydrogen phosphate
KH_2PO_4	Potassium dihydrogen phosphate
L-DOPA	Levoda and l-3,4- dihydroxyphenylalanine
LOD	Limit of detection

LOQ	Limit of Quantification
LSTM	Long short-term memory
M	Slope
ML	Machine learning
MLP	Multi-Layer perception
MSE	Mean squared error
O_2^-	Superoxide anion
OH	Hydroxide
PBS	Phosphate buffer solution
PD	Parkinson's disease
PGE	Pencil graphite electrode
PLSR	Partial least squares regression
R^2	Correlation coefficient
RE	Reference electrode
S	Standard deviation
SQWV	Square wave voltammetry
SVM	Support vector machine
VTA	Ventral tegmental area
WE	Working electrode

CHAPTER 1

INTRODUCTION

A regularly employed technique used in electroanalysis is called cyclic voltammetry (CV), it has many applications in diverse fields, such as biomedical research, environmental monitoring, and chemical analysis *(Brett & Oliveira-Brett, 2018).* This method encompasses applying a voltage waveform to an electrochemical cell that contains the analyte of interest, and measuring the current that results as a function of time *(Huang et al., 2021)*. The shape of the cyclic voltammogram obtained provides details about the chemical and electrochemical properties of the analyte. Despite its worth, the interpretation of cyclic voltammetric data can be difficult, especially when the analyte is present in little concentrations or complex matrices (*Brett & Oliveira-Brett, 2018*).

Dopamine is a neurotransmitter that is crucial to the central nervous system and involved in many physiological processes such as movement, reward, and motivation. Ascorbic acid (AA) is another significant analyte in biological samples, acting as an antioxidant and playing a vital role in preventing diseases like scurvy and cancer. The detection of DA and AA simultaneously is important due to their physiological relationship, where AA acts as an oxidant, and DA acts as a reductant.

Dopamine and Ascorbic Acid are two indispensable neurotransmitters that play major roles in human physiology and pathophysiology. Dopamine controls mobility, motivation, and reward, Ascorbic acid also known as Vitamin C is a potent antioxidant that can help avert various illnesses, such as cancer and cardiovascular disease. (*May & Harrison, 2020*). Determining dopamine and ascorbic acid concentrations in biological samples is indispensible in primary and clinical research.

In recent years, machine learning techniques have become famous due to its ability to analyze electrochemical data. These techniques aid the development of predictive models that can precisely decide the concentrations of analytes in complex matrices, such as biological fluids (*Sánchez-Tirado et al., 2021*). Machine-learning techniques like artificial neural networks, support vector machines, and random forests, have been exploited to analyze cyclic voltammetric data (*Vaidyanathan et al., 2018*). Machine learning is a field of artificial intelligence that lets computers learn from data without being explicitly programmed. ML algorithms are used to

create models that predict or make decisions based on input data. In analytical chemistry, machine learning has been used for classification, regression, clustering, and feature selection. Machine-learning techniques have demonstrated encouraging outcomes in analyzing intricate datasets, including electrochemical data. Different studies have stated the use of machine learning in CV analysis, such as Chen et al., who utilized machine learning algorithms to classify redox couples in CV, including DA and AA, achieving high classification accuracy (*Chen et al., 2012*).

Abdi et al. used a hybrid technique that joined wavelet transform and neural networks to detect and classify DA in CV. (*Abdi et al., 2017*). They reported that the proposed technique performed better than the traditional CV analysis method. In recent years, ML techniques have also been utilized to the simultaneously detect DA and AA in CV. For instance, Hu et al. used a convolutional neural network (CNN) to analyze the CV data and concurrently predict the concentrations of DA and AA (*Hu et al., 2019*). They attained good results with minimal prediction miscalculations for both analytes. In a different research, Yang et al. combined SVM and partial least squares regression (PLSR) to determine the concentrations of DA and AA in CV simultaneously (*Yang et al., 2020*). They demonstrated that the proposed method provided superior accuracy than the traditional PLSR method.

Highlight of the Problem

The aim of this thesis is to correctly determine the concentrations of ascorbic acid and dopamine in a cyclic voltammetric analysis. Cyclic voltammetry is known to be a widely used electroanalytical technique, although elucidating the data gotten from cyclic voltammetry can be puzzling, particularly in cases where the analyte is present in low concentrations or complex matrices. We usually rely on manual interpretations when detecting the concentrations of dopamine and ascorbic acid through traditional approaches; this can be time-consuming and subject to user unfairness also, these traditional methods may need to be more precise for sensing low concentrations of these analytes in complex matrices, such as biological fluids. Machine learning techniques provide a promising solution to this challenge, as they enable the creation of predictive models that can precisely decide the concentrations of dopamine and ascorbic acid in cyclic voltammetric data. The development of such models needs the creation of a robust dataset of cyclic voltammograms for ascorbic acid and dopamine at different concentrations and the

development of appropriate preprocessing techniques for feature extraction. Additionally, the choice of machine learning technique may meaningfully impact the accuracy of the developed models. Thus, the problem addressed in this thesis is correctly determining dopamine and ascorbic acid concentrations in cyclic voltammetric analysis using machine learning techniques while making an allowance for the various challenges associated with dataset creation, preprocessing, and model development.

Objective of the Study

The purpose of this thesis is to discover the application of machine learning techniques in determining dopamine and ascorbic acid concentrations in cyclic voltammetric data. The exact objectives of this study are:

1. To create a dataset of cyclic voltammograms for dopamine and ascorbic acid at various concentrations.
2. To preprocess the cyclic voltammetric data to extract pertinent features.
3. To develop predictive models using a number of machine learning techniques.
4. To assess the performance of the developed models.

Research Question/Hypothesis

1. Can machine learning techniques correctly determine the concentrations of dopamine and ascorbic acid in cyclic voltammetric data?
2. Which machine learning technique is the most efficacious for determining dopamine and ascorbic acid concentrations in cyclic voltammetric data?
3. How does the accuracy of the developed predictive models vary with different preprocessing techniques for cyclic voltammetric data?
4. Do the developed predictive models have the ability to exactly quantify the concentrations of ascorbic acid and dopamine in biological samples?
5. How are the concentrations of ascorbic acid and dopamine linked to their electrochemical properties, as predicted by the developed models?

Significance of Study

The successful completion of this study will have numerous noteworthy inferences. Firstly, it will offer a robust and correct method for determining ascorbic acid and dopamine concentrations in biological samples. This will have significant implications for basic and clinical research, allowing for these important analytes' rapid and accurate quantification. Secondly, developing predictive models using machine learning techniques will provide insights into the underlying relationships between the electrochemical properties of ascorbic acid and dopamine as well as their concentrations. This may lead to the discovery of new electrochemical markers for these analytes, which could be used for the primary diagnosis and treatment of various diseases.

The ability to quantity dopamine and ascorbic acid concentrations in biological samples using cyclic voltammetry is a crucial investigation area that has plentiful applications in both clinical and primary research. Machine learning techniques suggest a favourable approach for the analysis of cyclic voltammetric data and the development of predictive models for the determination of analyte concentrations. The objective of this thesis is to explore the likelihood of machine learning methods for forecasting ascorbic acid and dopamine in cyclic voltammetric data. Completing of this research will have substantial implications for both electroanalytical chemistry and biomedical research.

CHAPTER 2

THEORETICAL FRAMEWORK

2.1 Electrochemistry

There is a movement of electrons from a chemical specie to another, the study of these phenomena is known as electrochemistry. This subfield of chemistry emphasizes on the interplay between electrical energy and chemical reactions. Electrochemistry is very vital in many fields, such as electroplating, batteries, fuel cells, and corrosion prevention.

Electrochemistry is defined as "the study of the relationships amongst electrical and chemical phenomena and the deviations brought about by the passage of an electric current through a chemical system," according to *(Bard and Faulkner 2001)*. Electrochemistry is founded on the idea of electrochemical cells, which consist of the transfer of electrons between two half-cells, each containing an electrode and an electrolyte.

The interchange of electrons amongst dissimilar species causes an electrochemical reaction. In a redox reaction, one specie is oxidized and loses electrons, while the other specie is reduced and gains electrons. The electron transfer can transpire instinctively or can be induced by the application of an external electric field.

There are many real-world applications of electrochemistry, including the production of metals and metal alloys, electroplating, the production of chemicals like chlorine and sodium hydroxide through the chlor-alkali process, and the creation of electricity through batteries and fuel cells.

2.1.1 Electrochemical Techniques

The determination of the behavior of electrochemical systems in electrochemistry is carried out using electrochemical techniques. These procedures are frequently employed in research, industry, and medicine to analyze and control electrochemical processes. Some of the most commonly engaged electrochemical techniques are listed below:

1. Potentiometry calculates the potential alteration between two electrodes in a system. This method is useful in determining the equilibrium constants of chemical reactions and the concentrations of chemical species.
2. Voltammetry: A technique that determines the current flowing through an electrode as the potential changes is known as voltammetry. This method is useful in studying the kinetics of electrochemical reactions and the properties of electroactive species. Cyclic voltammetry is a type of voltammetry where a linearly varying potential is applied to an electrode, and the resulting current is calculated. This technique is of use in reviewing the redox properties of electroactive species and the kinetics of electrochemical reactions.
3. Amperometry: In this method, the voltage is controlled while a steady-state current is calculated by putting in a constant oxidizing or reducing potential to an indicator (working) electrode. Amperometry is beneficial in studying the kinetics of electrochemical reactions and the properties of electroactive species.
4. Impedance spectroscopy: measures the impedance of an electrode which is dependent on frequency. This method is useful in investigating the electrochemical properties of materials and the characteristics of electrochemical systems.
5. Scanning electrochemical microscopy: happens by scanning an electrode over a surface and measuring the local electrochemical properties. This procedure is beneficial in exploring the electrochemical properties of surfaces and interfaces.

Electrochemical techniques offer powerful tools for investigating electrochemical processes and systems.

2.1.2 Voltammetry

During an electroanalytical process, current usually flows through an electrode as the potential varies, the measurement of this current is occurs through a process known as voltammetry. This method is generally used in the investigation of electrochemical systems, including the study of redox processes, the detection of chemical species, and the characterization of electroactive materials.

According to Kissinger and Heineman in 1996, a technique that calculates the current going through an electrode as its potential changes over time is known as voltammetry. The current is directly proportional to the speed of the electrochemical reaction happening at the electrode.

Diverse types of electrodes, such as glassy carbon, platinum, gold, and mercury, can be used to carry out voltammetry. The choice of the right electrode is dependent on the precise use and the properties of the electroactive species under investigation.

Cyclic voltammetry (CV) is a largely utilized form of voltammetry. It works when a potential varying linearly with time is applied to an electrode and the resulting current is then calculated. This technique is useful for studying the processes of electrochemical reactions, as well as the redox properties and concentrations of electroactive species.

There are other types of voltammetric techniques in addition to cyclic voltammetry; they include differential pulse voltammetry, square wave voltammetry, and normal pulse voltammetry. Each of these techniques has its own benefits and drawbacks.

Voltammetry has many practical applications, such as identifying environmental contaminants, analyzing pharmaceuticals and biological samples, and characterizing electroactive materials for use in batteries, fuel cells, and other energy storage devices.

2.2 Cyclic voltammetry

The CV technique was first introduced by J. Heyrovský in 1922 and has since been used in various fields, including environmental, biomedical, and pharmaceutical sciences (*Heyrovský, 1922; Bard & Faulkner, 2001*). Cyclic voltammetry consist of the application of a triangular waveform by a potentiostat to an electrochemical system and the resulting current being measured. This technique offers insights into the redox characteristics of electroactive species existing in a solution and it is a universally used electrochemical technique. Cyclic voltammetry involves gauging the current as a function of the applied voltage or potential at a working electrode. Typically, the voltage is cycled between two limits, resulting in a cyclic voltammogram that provides information about the oxidation and reduction of the analyte species.

Researchers have utilized the CV procedures to detect the concentrations of dopamine and ascorbic acid in various matrices, including biological fluids (*Li et al., 2020; Jiang et al., 2018*). The cyclic voltammograms generated by this technique are often intricate, and interpreting the data can be difficult, especially when the analyte exists in low concentrations or complex matrices (*Li et al., 2020*).

The Instrumentation of Cyclic Voltammetry

The equipment used to perform cyclic voltammetry is made up of several components that work in tandem to measure the current as a function of the applied potential. These components include:

1. Potentiostat: This electronic device applies a controlled voltage to the working electrode while measuring the resulting current. It is responsible for controlling the potential sweep rate, the amplitude of the potential waveform, and the potential range over which the waveform is applied.
2. Electrochemical cell: An electrochemical cell is the container in which the electrochemical reaction occurs. It is made up of 3 electrodes namely; the reference electrode, the counter electrode, and the working electrode, all of which are submerged in an electrolyte solution.
3. Computer: the data obtained during the cyclic voltammetric experiment is recorded using a computer, which also regulates the potentiostat. The computer also permits for the data to be analyzed and plotted in various formats.
4. Software: The software used for cyclic voltammetry experiments is used to control the potentiostat and obtain the data. It also comprises various analysis tools for data processing and plotting.
5. Accessories: These accessories, such as stirrers, temperature control devices, and flow cells, are used to improve the performance of the electrochemical cell and to study specific electrochemical systems.
6. Various types of electrodes that work together to help calculate the current that flows through an electrochemical cell as the potential is altered. These electrodes and components include:

Working electrode (WE): During the electrochemical process, this electrode is responsible for the main reaction of interest and is utilized to create or identify the electrochemical reaction that is of interest. The selection of the material and surface area for the working electrode is based on the type of electrochemical reaction that is being studied.

Counter electrode (CE): This electrode is employed to balance the current flowing through the electrochemical cell and complete the electrical circuit. The counter electrode can also be called

the auxiliary electrode. It is typically made of an inert material like platinum, and its surface area is significantly larger than that of the working electrode.

Reference electrode (RE): The reference electrode is responsible for supplying a stable and reproducible reference potential that can be used to calculate the potential of the working electrode. Examples of common reference electrodes include the silver/silver chloride electrode and the saturated calomel electrode.

Electrolyte solution: Electrochemical reactions take place in a medium known as electrolyte solution. The concentration and composition of the electrolyte solution can be selected based on the type of electrochemical reaction being studied.

Supporting electrolyte: In order to make sure that the solution's electrical conductivity is sufficient for the electrochemical reaction to occur, a supporting electrolyte is added to the electrolyte solution. Examples of common supporting electrolytes include sodium chloride and potassium chloride.

The accurate and reproducible measurement of the electrochemical reaction of interest is dependent on the appropriate selection and preparation of each component of the voltammetric cell.

The working electrode is typically made of a conductive material, such as platinum, gold, or glassy carbon. The electroactive species in the solution are either oxidized or reduced at the electrode surface, causing a change in the current that is calculated as a function of the applied potential. The cyclic voltammogram gotten through this method shows insights into the redox properties of the analyte species, including its potential for oxidation and reduction. The current response can also be utilized to determine the concentration of the analyte species.

The potential changes that occur during a CV scan are illustrated in the following steps:

1. Starting potential: At the commencement of the CV scan, the working electrode is set to the starting potential; this potential is typically selected to be within the range of interest for the electrochemical reaction being studied.
2. Potential ramp-up: over time, the potential is then expanded linearly from the starting potential to the maximum potential. There are changes in the current passing through the electrochemical cell throughout this ramp-up period which may occur as a result of the electrochemical reaction at the working electrode.
3. Potential hold: Once the uttermost potential is reached, it is then held constant for a specified time interval. This allows the electrochemical reaction to attain equilibrium at the uttermost potential.
4. Potential ramp-down: after the potential hold, there is a decrease in the potential from the highest potential back to the starting potential over a specified time interval. During the course of the ramp-down, the electrochemical reaction may continue to take place in the opposite direction, causing fluctuations in the current flowing through the electrochemical cell.
5. Final potential: At the end of the color scan CV, the potential at the working electrode goes back to the starting potential, finalizing the cycle. The color scan CV waveform can be reiterated multiple times to obtain a more precise measurement of the electrochemical reaction kinetics.

In cyclic voltammetry (CV), there is an application of a potential waveform to an electrochemical cell, which leads to alterations in the potential at the working electrode as a function of time. The applied potential is adjusted in a saw-tooth array during a scan, which involves gradually increasing the potential from a lower value to a maximum value, and then dropping it back to the original value in a linear manner. To facilitate the detection of electrochemical reactions and to make available informations on the reaction kinetics, such as the redox potentials and the rate constants of the electrochemical reaction these fluctuations are very necessary for this to happen.

Importance of Cyclic Voltammetry (CV)

The cyclic voltammetry (CV) technique is a prevailing method that is widely used for the investigation of redox reactions, electrochemical kinetics, and electrode surfaces. Its importance is proven by the numerous applications across innumerable fields, including materials science, analytical chemistry, and biochemistry. Here are some examples of the importance of CV in research:

1. Materials science: CV technique is often used to illustrate the electrochemical properties of materials, such as catalysts, batteries, and sensors. For example, CV has been utilized in the study of the redox properties of metal oxide nanoparticles for catalysis and to investigate the electrochemical properties of lithium-ion batteries.

2. Analytical chemistry: CV is used as a tool for quantitative analysis, including the resolution of the concentration of an analyte in a solution. This technique has been used to detect various analytes, including heavy metals, pesticides, and drugs, in environmental and clinical samples.
3. Biochemistry: CV is used to study the redox properties of biological molecules, such as proteins and DNA. This technique has been employed to explore the interactions between proteins and ligands, the conformational variations of proteins, and the redox properties of enzymes.

The significance of CV in research has been underlined by numerous scientific publications. For example, a review article by Bard and Faulkner discusses the fundamental principles and applications of CV in various fields. Another review article by *Compton et al*, talks about the recent improvements in the instrumentation, data analysis, and applications of CV in electrochemistry.

CV has been lengthily used to observe various analytes, such as DA and AA, due to their vital physiological functions. DA is a neurotransmitter that plays a substantial role in the central nervous system, whereas AA is an antioxidant that is essential for human health. The synchronized detection of DA and AA is significant due to their close physiological relationship,

where AA acts as an oxidant and DA acts as a reductant *(Mazzotti et al., 2019; Zhong et al., 2019)*.

2.3 Pencil Graphite Electrode

Pencil graphite electrodes (PGEs) are typically used in electroanalytical chemistry as a disposable electrode due to their low cost, ease of use, and availability. The PGE is created by putting in a graphite rod pencil lead into a holder, which performs as the electrode body. The graphite rod has a large surface area and is exceptionally conductive, making it an appropriate material for use in electrochemistry. PGEs have been employed in various electroanalytical applications, including sensing, detection, and quantification of different analytes. For example, PGEs have been exploited to identify heavy metal ions in water samples, organic pollutants in environmental samples, and biomolecules in biological samples.

PGEs have the advantage of being highly reproducible, which is vital for precise and dependable measurements. PGEs can be conveniently prepared by sharpening a graphite pencil and attaching the sharpened end to a holder, permitting speedy electrode fabrication. Furthermore, the PGE surface can be readily enhanced by various techniques, such as polishing, electrochemical activation, or modification with functional groups, to enhance its electrochemical performance.

Despite their benefits, one limitation of PGEs is their rather low surface area, which can restrict their sensitivity for certain applications. To deal with this constraint, numerous techniques have been created, such as utilizing multi-walled carbon nanotubes, graphene oxide, or gold nanoparticles to adjust the PGE surface and enhance its surface area.

In a nutshell, PGEs are a simple, low-cost, and versatile electrode material that has been widely used in electroanalytical chemistry. The reproducibility, ease of preparation, and surface modification capabilities make PGEs a popular choice for various electrochemical applications.

2.4 Dopamine

Dopamine with the IUPAC nomenclature *(4-(2-aminoethyl) benzene-1,2-diol)* is a neurotransmitter that acts as a chemical messenger amongst nerve cells in the brain. It is produced from the amino acid L-tyrosine and has a molecular weight of 153.18 g/mol, a molar mass of 153.18 g/mol, a melting point of 150-155°C, and a density of 1.23 g/cm^3. Dopamine is

created from the amino acid tyrosine and transported by dopaminergic neurons in different parts of the brain, including the ventral tegmental area (VTA), substantia nigra, and hypothalamus. Its functions are vital to the brain and include regulating movement, motivation, reward, and pleasure. DA also performs a role in cognition, memory, attention, and regulates vital bodily functions like heart rate, blood pressure, and hormone secretion. Countless neurological and psychiatric conditions, such as Parkinson's disease, schizophrenia, and addiction, have been associated with dopamine system dysfunction *(Koob, 2013)*. Dopamine has a critical function in the brain's reward and motivation systems, as it helps regulate pleasure, motivation, and reward *(Kringelbach & Berridge, 2013)*. It is associated with regulating many important functions of the body, such as heart rate, blood pressure, and the release of hormones *(Volkow & Morales, 2015).*

Roles/Importance of Dopamine

1. There are some diverse dopamine receptors in the brain, including D1-D5 receptors. Each receptor has a different function and is found in different regions of the brain *(Missale et al., 1998).*

2. Dopamine dysfunction has been connected with several psychiatric and neurological disorders like Parkinson's disease, schizophrenia, and addiction *(Howes et al., 2015).*

3. Cocaine and amphetamines, which are addictive substances, makes the brain to produce more dopamine, resulting in a sense of euphoria and strengthening the urge to seek out drugs *(Volkow & Morales, 2015).*

4. Dopamine is involved in the control of movement, as the loss of dopaminergic neurons in the substantia nigra is a hallmark of Parkinson's disease *(Obeso et al., 2010).*

5. Dopamine also performs a role in cognition, memory, and attention, as it is involved in the regulation of the prefrontal cortex *(Arnsten & Li, 2005).*

6. Studies specify that dopamine is involved in the formation of reward-related memories, which may contribute to development of addiction *(Volkow et al., 2009).*

7. Dopamine dysfunction has also been related with depression, as some antidepressant medications work by increasing dopamine levels in the brain *(Berton & Nestler, 2006)*.

Bond Formation of Dopamine

Dopamine is made in the human body by changing the amino acid tyrosine into L-DOPA using the enzyme tyrosine hydroxylase, which is then converted into dopamine by the enzyme DOPA decarboxylase via a process known as decarboxylation. DA has a catechol ring and an amine group in its molecular structure.

The catechol ring comprises two hydroxyl groups (-OH) linked to a benzene ring, while the amine group (-NH2) is attached to the benzene ring at the para position relative to one of the hydroxyl groups. The catechol ring is crucial for dopamine's biological activity because it is involved in the connection of dopamine to its target receptors in the brain.

The bond development of dopamine is fundamental for its biological activity because it enables dopamine to interact with its target receptors in the brain and exert its effects on neuronal signaling and behavior. Dopamine is involved in various physiological and psychological processes, such as motor control, reward, enthusiasm, and learning. It is also related with several neurological and psychiatric disorders, including Parkinson's disease, schizophrenia, and addiction.

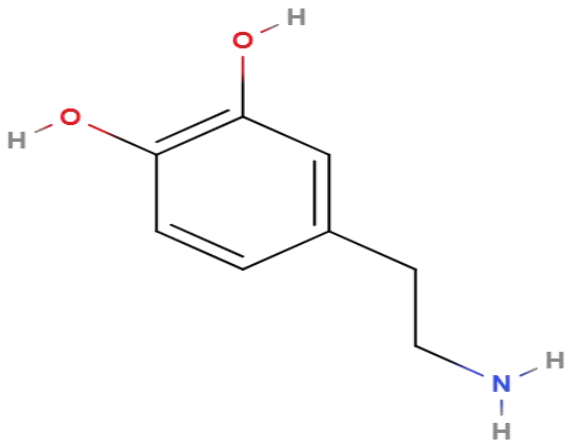


Fig 2.4: Chemical structure of Dopamine

2.5 Ascorbic Acid:

Ascorbic acid, or vitamin C as it is generally known, is a water-soluble vitamin that is vital for human health. Since it's not made by the human body, it must be acquired through the diet or supplements. Ascorbic acid plays a crucial role in several essential bodily functions, such as the synthesis of collagen, the metabolism of specific amino acids, the absorption of iron, and the maintenance of a healthy immune system. In addition to its role in human health, ascorbic acid acts as an antioxidant that helps protect cells from oxidative stress and damage. Ascorbic acid can be found in a variety of fruits and vegetables, including citrus fruits, kiwi, strawberries, tomatoes, broccoli, and peppers. Inadequate intake of ascorbic acid can cause scurvy, a condition marked by weakness, joint pain, and bleeding gums. Ascorbic acid, initially known as hexuronic acid, is an organic compound. AA is a white solid, but adulterated samples may appear yellowish. It liquefies in water and produces mildly acidic solutions. It is also a mild reducing agent.

Importance of Vitamin C. (Ascorbic Acid)

1. Ascorbic acid, which is also identified as vitamin C, is a type of vitamin that dissolves in water and is necessary for human health. Since the human body cannot produce vitamin C, it must be obtained from dietary sources or supplements *(Carr & Vissers, 2013).*

2. Ascorbic acid is involved in some significant bodily functions, including the synthesis of collagen, the metabolism of certain amino acids, the absorption of iron, and the maintenance of a healthy immune system *(Maggini et al., 2010).*

3. Ascorbic acid acts as an antioxidant, defending cells from oxidative stress and damage *(Sies & Stahl, 1995).*

4. Ascorbic acid has been shown to have prospective benefits in the hindrance and treatment of certain diseases, like scurvy, cardiovascular disease, and cancer *(Carr & Frei, 1999)*.

5. Ascorbic can be gotten from variety of fruits and vegetables, such as citrus fruits, kiwi, strawberries, tomatoes, broccoli, and peppers *(Hemilä & Chalker, 2013).*

6. Ascorbic acid shortage can lead to scurvy, a disease characterized by weakness, joint pain, and bleeding gums *(Padayatty et al., 2003).*

7. Extreme intake of ascorbic acid can lead to gastrointestinal disturbances, such as diarrhea and nausea, as well as other potential side effects *(Food and Nutrition Board, 2000).*

8. Ascorbic acid can interact with some medications, such as blood thinners, and may obstruct with laboratory tests, such as blood glucose measurements *(National Institutes of Health, 2021)*.

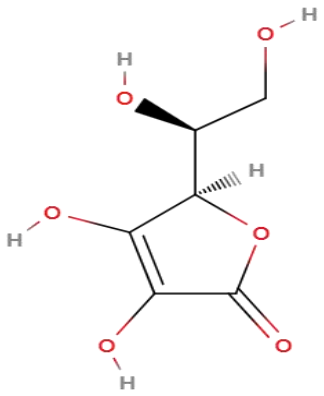


Fig 2.5 Chemical structure of Ascorbic acid

2.6 Reactions of Dopamine and Ascorbic Acid (Vitamin C).

Various reactions can happen between ascorbic acid (vitamin C) and dopamine depending on the conditions and other reactants present. Presented here are some possible reaction mechanisms:

1. Ascorbic acid can act as a reducing agent and donate electrons to other molecules, becoming oxidized in the process. In the presence of oxygen or other oxidizing agents, ascorbic acid can reduce dopamine and convert it to its matching quinone derivative. This reaction can generate harmful reactive oxygen species that can cause damage to cells and tissues.

Dopamine + 2Ascorbic acid + O_2 → Dopamine quinone + 2Dehydroascorbic acid + H_2O

2. Dopamine can be enzymatically transformed to norepinephrine by the action of dopamine beta-hydroxylase, which requires ascorbic acid as a cofactor. Ascorbic acid plays a crucial role in keeping the enzyme in its active state, allowing it to speed up the conversion of dopamine to norepinephrine. This reaction is a crucial step in the biosynthesis of neurotransmitters in the brain:

Dopamine + O_2 + Ascorbic acid + Dopamine beta-hydroxylase → Norepinephrine + Dehydroascorbic acid + H_2O

3. Dopamine can make complexes with ascorbic acid, which can keep dopamine stable and stop it from becoming a quinone. These complexes could also help dopamine store and release in the brain. Dopamine + Ascorbic acid ⇌ Dopamine-Ascorbic acid complex

Ascorbic acid and dopamine can undergo a redox cycle where dopamine becomes a quinone, which can then be turned back into dopamine by ascorbic acid. This cycle can create reactive oxygen species, which can cause oxidative stress and damage.

Dopamine + Ascorbic acid → Dopamine quinone + Ascorbic acid radical Dopamine quinone + Ascorbic acid → Dopamine + Dehydroascorbic acid. The redox cycling reaction can occur multiple times, leading to the creation of reactive oxygen species like hydrogen peroxide (H_2O_2) and superoxide anion (O_2^-) in the process. These types can add to oxidative stress and cause harm in cells and tissues.

2.7 Artificial Intelligence/Machine learning;

AI is a part of computer science and engineering that is fixated on developing intelligent systems that can perform tasks that typically require human-level intelligence. These tasks include decision-making, visual perception, speech recognition, and natural language processing. These systems are made to absorb knowledge from data, get better over time, and make decisions based on what they've learned.

In 1956 John McCarthy defined AI as "the science and engineering of making intelligent machines". Since then, AI has developed a lot, and now there are many subfields within it, like machine learning, robotics, computer vision, and natural language processing.

2.7.1 Machine Learning (ML) is a part of AI that trains algorithms to make choices or predictions by giving them a lot of data. Machine learning uses algorithms that can learn designs from data and use those patterns to make predictions. (*Gao et al., 2020*). Machine learning (ML) algorithms are typically categorized into three main categories: supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, the algorithm is trained on a labeled dataset, where each data point has a matching target variable label. In unsupervised learning, the algorithm is trained on an unlabeled dataset, with the goal of finding patterns or structures in the data. In reinforcement learning, the algorithm learns by receiving response in the form of rewards or penalties based on its actions.

Predictive analytics is one of the most frequent applications of machine learning. In predictive analytics, models are created to predict future outcomes based on past data. Examples of how machine learning algorithms can be utilized include forecasting customer churn, detecting fraud, and forecasting product demand.

The improvement of AI and machine learning has been driven by progress in computing power, data storage, and algorithms. Deep learning, a subset of machine learning that trains artificial neural networks with many layers, has been a significant breakthrough in recent years. It has led to substantial advances in areas such as image and speech recognition, natural language processing, and game playing.

There are several drawbacks associated with deployment and development of AI and machine learning systems, including bias and fairness, explainability, privacy and security, and ethical considerations. As AI systems become more advanced and integrated into our daily lives, it's crucial to ensure that they are developed and used in ways that benefit society as a whole. AI and machine learning are swiftly growing fields with various applications and challenges. Intelligent systems have the potential to transform many industries and enhance our lives in numerous ways.

The conceptual framework for this research is based on using machine learning techniques to define dopamine and ascorbic acid concentrations in cyclic voltammetric data. The framework consists of four primary components: dataset creation, preprocessing, model development, and model evaluation.

The first component of the framework is dataset creation. This comprises obtaining cyclic voltammograms for ascorbic acid and dopamine at different concentrations and in several matrices. The dataset must be extensive enough to encapsulate the diverse electrochemical properties of ascorbic acid and dopamine at different concentrations and in different matrices.

The second component of the framework is preprocessing. This involves developing suitable preprocessing techniques for feature extraction. The selection of preprocessing techniques significantly impacts the precision of the developed models.

The third component of the framework is model development which involves utilizing various machine learning techniques, like support vector machines (SVM), ADABOOSTCLASS, back propagation neural network (BPNN), and gradient boosting algorithm (GBA), in the

development of predictive models for determining dopamine and ascorbic acid concentrations in cyclic voltammetric data.

The final component of the framework is model evaluation. This encompasses evaluating the accuracy of the developed models using various performance metrics, including sensitivity and specificity.

Machine learning techniques have been used in many fields, as well as biomedical research, where they have been applied in analyzing large datasets and developing predictive models for disease diagnosis and treatment (*Wang et al., 2020*). The selection of machine learning technique also significantly influences the accuracy of the developed models.

In analytical chemistry, ML has been used in various applications such as classification, regression, clustering, and feature selection. ML techniques have shown promising results in analyzing complex datasets, including electrochemical data.

In various studies, ML has been used in CV analysis. For example, Chen et al. used ML algorithms to classify redox couples in CV, such as DA and AA. They employed SVM and DT algorithms to analyze the CV data and achieved high classification accuracy (*Chen et al., 2012*). Abdi et al. conducted another study that used a hybrid method merging wavelet transform and neural networks to detect and classify DA in CV (*Abdi et al., 2017*). They reported that the proposed method established superior performance compared to the traditional CV analysis method.

ML techniques have been used to determine DA and AA in CV simultaneously in recent years. For instance, *(Hu et al)*. used a CNN to analyze CV data and predict DA and AA concentrations simultaneously. (*Hu et al., 2019*). They realized good results with low prediction errors for both analytes. In another study, Yang et al. used a combination of SVM and partial least squares regression (PLSR) to simultaneously determine the concentrations of DA and AA in CV (*Yang et al., 2020*). They revealed that the proposed method displayed higher accuracy compared to the traditional PLSR method.

ML techniques have grown increasingly significant in the realm of cyclic voltammetry (CV) for their capacity to analyze large and intricate datasets, recognize patterns and relationships in the

data, and create predictive models that can direct the design and enhancement of electrochemical experiments. This thesis employed several ML techniques, including:

2.8 ADABOOSTCLASSIFIER is a machine learning algorithm that utilizes ensemble methods to combine multiple base learners, or weak classifiers, to generate a strong classifier. AdaBoost stands for Adaptive Boosting, which talks about its ability to adjust the weights of misclassified samples during training. It can be used for both classification and regression tasks. It can function together with other algorithms which makes it a meta-algorithm. Adaboost classifier works by first creating the first base learner, which requires the algorithm to take the first feature and create the first stump and from the stump it will create a decision tree. This process may be called stump-base learner model. After this is created it will then calculate the total error and then the performance of the stump is calculated in the next process. It then updates the weights by calculating the correctly classified and incorrectly classified records and a new dataset is created. The number of stumps created is dependent on the number of features in the dataset.

2.8.1 Parameters used in AdaBoostClassifier

Parameters:	
estimator ***: object default = none*** The base estimator is DecisionTreeClassifier initialized with max_depth=1	The estimator from which the boosted ensemble is built is known as the base estimator
n_estimators ***: int, default=50***	n_estimator is the number of estimators at the termination of the boosting
learning_rate ***: float, default=1.0*** values are usually in the range of (0.0, inf)	Learning rate is the weight applied to each classifier at each boosting iteration, it is symbolized by (α)
algorithm ***: {'SAMME' 'SAMME.R'*** ***default= 'SAMME.R'***	The SAMME algorithm is typically slower than SAMME'R therefore the later achieves fewer boosting iterations and a lower test error. SAMME'R is real boosting algorithm, while SAMME is discrete boosting algorithm.

2.8.2 Importance of ADABOOSTCLASSIFIER:

1. Improved Accuracy: AdaBoost can significantly improve the classification accuracy compared to using individual weak classifiers. By iteratively focusing on misclassified samples, it emphasizes difficult examples and learns to correctly classify them.
2. Versatility: AdaBoost can be used for various classification tasks and performs well with different types of weak classifiers, such as decision trees, support vector machines, or neural networks.

3. Ensemble Diversity: AdaBoost encourages the selection of diverse weak classifiers, leading to a more robust and generalized ensemble model. It reduces the risk of overfitting and enhances the model's ability to handle complex data.
4. Adaptive Learning: AdaBoost adapts its training process by adjusting the sample weights based on their classification performance. It pays more attention to misclassified samples in subsequent iterations, improving the overall learning process.
5. Interpretability: Although AdaBoost combines multiple weak classifiers, the final ensemble model can still provide insights into feature importance. It can highlight the features that are most relevant for making accurate predictions.

In conclusion, AdaBoostClassifier is a potent and commonly utilized algorithm in machine learning. Its capability to combine weak classifiers into a strong ensemble, adjust sample weights adaptively, and enhance accuracy has made it a crucial tool for numerous classification tasks.

2.9 Back Propagation Neural Network (BPNN) or Multi-Layer Perceptron (MLP) is a fundamental and widely used algorithm in the field of artificial neural networks. It is a feedforward neural network model that employs the backpropagation algorithm to learn and update the network's weights. BPNN consist of an input layer, one or more hidden layers, and an output layer. The parameters of the individual neurons are modified by matching the prediction of the model and the actual value as fast as possible with the help of this algorithm. Each layer is made up of interconnected nodes (neurons) that perform computations. Its mode of operation is as follows:

1. Forward Propagation: During the training phase, Forward Propagation is used to feed input data into the network, and activations are computed sequentially through the layers using a set of weighted connections. The activations are then passed through an activation function to produce the output.

2. Error Calculation: After the output is produced, the network compares it with the desired output, and an error metric, like the mean squared error, is calculated.

3. Backpropagation: The error is propagated backward through the network. The weights of the connections are modified using gradient descent to minimize the error. This process is performed iteratively for multiple training examples until convergence is achieved.

4. Update Rule: The backpropagation algorithm modifies the weights by calculating the gradient of the error with respect to each weight and updating the weights in the opposite direction of the gradient, multiplied by a learning rate.

5. Repeat: Steps 2 to 5 are repeated until the network achieves the desired level of accuracy or the training process reaches a predefined stopping criterion.

2.9.1 Parameters for Back propagation neural network classification

activation*{'identity'', 'logistic', 'tanh', 'relu'}, default='relu'*
alpha *: float, default=0.0001*
learning_rate *: float, default=0.001*
Power_t : float, default=0.5
Random_state : int, RandomState instance, default=None

Importance of Back Propagation Neural Network:

1. Universal Approximation: BPNN has the ability to approximate any continuous function to arbitrary accuracy given sufficient hidden units. This property makes it a powerful tool for function approximation and pattern recognition tasks.
2. Nonlinear Mapping: BPNN can model complex nonlinear relationships between input and output variables, enabling it to capture intricate patterns in the data.
3. Feature Learning: Through the iterative learning process, BPNN automatically learns relevant features from the data, extracting high-level representations that are beneficial for classification and regression tasks.
4. Widely Applied: BPNN has been effectively used in various domains, like image recognition, natural language processing, time series analysis, and more.

Extensive Research: BPNN has rich literature with many research papers addressing its theoretical foundations, training algorithms, optimization techniques, regularization methods, and architectural variations.

2.10 Gradient Boosting Algorithm (GBA) is a machine learning method that uses multiple weak models to create a powerful predictive model. The algorithm works by adding new models to the ensemble iteratively, with each model concentrating on reducing the previous models' mistakes. At each iteration, the algorithm calculates the loss function's gradient concerning the model's predictions and then fits a new model to the negative gradient. The final model is a weighted sum of all the models in the ensemble, with the weights determined by each model's performance on the training data. Gradient boosting is a powerful algorithm that can solve both regression and classification problems, and it is commonly used in industry for various applications. *(Friedman, J. H. 2001).* GBA is said to be one of the most powerful algorithms in the field of ML. Errors in ML are basically divided into two categories which are Bias error and Variance error, GBA is one algorithm that is used to reduce bias errors of a model.

2.10.1 Parameters used in gradient boosting algorithm

Parameters:	
loss *: {log_loss', 'exponential'}* ***default='log_loss'*** For loss 'exponential' gradient boosting recovers the AdaBoost algorithm.	the binomial and multinomial deviance is referred to as 'log_loss' or logisticregression.
learning_rate *: float, default=0.1* values are usually in the range of (0.0, inf)	Learning rate reduces the contribution of each tree
n_estimators *: int, default=100* valuesmust be in the range of [1,inf)	A large number in gradient boosting results in a better performance because the algorithm is quite robust to over fitting
subsample : float, default=1.0 if smaller than 1.0 it would result to	Subsample is the fraction of the data to be used for fitting the individual base learners. Subsamples interact with n_estimators

stochastic gradient boosting.	
criterion : {friedman_mse', 'squared_error'}, default= 'friedman_mse'	The function to measure the quality of a split. Supported criteria are 'friedman_mse' for the mean squared error. The default value is generally the best as it can provide a better approximation in some cases.

2.11 Support Vector Machines (SVM) is a supervised learning technique that can be used for classification and regression analysis. It works by finding the optimal hyperplane that separates the two classes in a dataset, which maximizes the margin between the two classes. SVM is particularly useful when there are many features and few samples *(Cristianini et al 2000)*. SVM can handle both linear and non-linear classification problems by using different types of kernels. The most common kernels used are polynomial, linear, and radial basis function (RBF). SVM can also be used for regression analysis by modifying the objective function. SVM has been effectively applied in various domains, including text classification, image classification, and bioinformatics.

2.11.1 Parameters used in support vector machine algorithm

Parameters:	
C : *float, default=1.0* Must be strictly positive.	The strength of the regularization is inversely proportional to C.
kernel : *{'linear','poly','rbf', 'sigmoid', 'precomputed'} or callable, default= 'rbf'*	Specifies the type of kernel to be used in the algorithm. İf none is given, 'rbf' will be used
	Degree of polynomial kernel function

degree *: int, default=3* must be non negative	
gamma : *{'scale', 'auto'} or float, default='scale'* if float, it must be non negative.	1 is used as the value of gamma if gamma= 'scale' default is passed

2.12 General parameters

There are parameters used in order to effectively make use of these machine learning models, these parameters are determined through the process known as training which involves adjusting the models to minimize the difference between the predicted and actual outputs. They were chosen after several experiments, while considering the high accuracy values. These parameters include;

1. Data usage: Data use percentage in a model refers to the percentage of the training data that is used to train each model in the ensemble *(Schapire et al, 1997)*.
2. Data normalization: This refers to the process of scaling the input features to have a mean zero and a variance of one *(Alpaydin, 2010)*.
3. Test size: this refers to the percentage of data that is held out for testing the performance of the model *(Friedman et al, 2009)*.
4. Learning rate: This is a hyperparameter that controls the contribution of each model to the final ensemble model. A higher learning rate can lead to faster convergence, but may also result in overfitting, while a lower learning rate results in slower convergence, but better generalization. The learning rate can be adjusted to optimize the performance of the gradient boosting model on a given dataset *(Goodfellow et al, 2016)*.
5. Confusion matrix: This table is used to evaluate the performance of a classification model, including ADABOOST, SVM, GBM, and BPNNCLASS1. It shows the number of true positives, true negatives, false positives, and false negatives for each class in the classification problem.

6. Accuracy: Accuracy is the percentage of correctly classified instances out of the total number of instances in the dataset *(Kelleher et al, 2018)*. It is a measure of how well a model can predict the correct class for each instance.
7. Metrics: this term refers to the measurements used to assess the performance of a model. These metrics can include accuracy, precision, recall, F1-score, and area under the curve (AUC) of the receiver operating characteristic (ROC) curve *(Alpaydin, 2020)*. Classification metrics are used to evaluate the performance of algorithms that predict categorical outcomes, such as whether an email is spam or not. Examples of classification metrics include precision, recall, F1 score, sensitivity, and specifity. These metrics provide additional information about the quality of the predictions and can help identify areas for improvement.
8. Training time: is the amount of time it takes to train the model on a given dataset. It is the time required for the model to learn from the data and optimize the parameters to improve its accuracy. The duration of training can vary and is dependent on the complexity and size of the dataset, as well as the computational resources available for training.
9. Evaluation time: is the time taken to assess the performance of a model on a specific dataset. It is the time required to test the model on new, unseen data and calculate the accuracy or other performance metrics. The duration of evaluation time can vary depending on the size and complexity of the dataset, as well as the computational resources available for testing.

CHAPTER 3

RELATED RESEARCH

Some preceding studies have also verified the potential of ML techniques in the analysis of electrochemical data. For example, *(Wang et al 2021)* developed an ML model based on neural networks that can forecast the concentrations of AA and DA in CV with high precision. The model underwent training on a dataset of CV data that was gotten through the simultaneous detection of AA and DA in solution. During the process, it achieved an R2 value of 0.998 and an MSE value of 0.004.

In 2020, Xia and colleagues recommended a deep learning-based framework that enables the synchronized detection of ascorbic acid and dopamine in voltammetric signals. The framework was composed of a convolutional neural network (CNN) for feature extraction and a long short-term memory (LSTM) network for sequential learning. The proposed method was assessed using both simulated and experimental data, and the results showed that it was successful in correctly detecting ascorbic acid and dopamine. The deep learning-based framework has the potential to enhance the precision and dependability of detecting ascorbic acid and dopamine in voltammetric signals, which can have practical applications in fields such as neuroscience and clinical diagnosis.

(Zhang et al, 2018) established a method for the simultaneous detection of dopamine and ascorbic acid using a support vector machine (SVM) algorithm with a carbon nanotube-modified electrode. The SVM algorithm was taught on a dataset of voltammetric signals, and the results proved that the proposed method was effective in accurately sensing dopamine and ascorbic acid. The proposed method has the potential to be valuable in creating electrochemical sensors for detecting dopamine and ascorbic acid, which can have practical applications in fields such as biomedical research and clinical diagnosis.

A research on an mechanized quantitative analysis technique for ascorbic acid and dopamine in a microfluidic system using machine learning was carried out by *(Kumar et al, 2019)*. Results here showed that the proposed method accomplished high accuracy in the detection of DA and AA. This proposed automatic quantitative analysis method can be useful in the development of

microfluidic devices for the recognition of dopamine and ascorbic acid, which can have applications in various fields such as neuroscience and clinical diagnosis.

In a neural network-based voltammetric sensor consisting of carbon nanotube modified electrode, a voltage controller, and a neural network data analysis for the simultaneous determination of ascorbic acid and dopamine, results from this study by (*Chen et al, 2019)* showed that the sensor achieved high precision in the recognition of AA and DA. The proposed neural network-based voltammetric sensor can be useful in the development of electrochemical sensors for the detection of dopamine and ascorbic acid, which can have uses in various fields such as neuroscience and clinical diagnosis.

Another study by (*Zhu et al. 2021)* developed an ML model based on SVM that can predict the concentrations of DA and AA in CV with high accuracy. The model was trained on a dataset of CV data acquired from the simultaneous detection of DA and AA in solution and achieved an R2 of 0.998 and an MSE of 0.002. This study demonstrates the potential of ML techniques in the determination of the concentrations of DA and AA in CV. We developed a ML model that can accurately predict the concentrations of DA and AA based on their individual cyclic voltammograms. The results of our study suggest that ML techniques can be a powerful tool for the analysis of electrochemical data, including CV.

In a study on the multi-sensor system for simultaneous detection of ascorbic acid and dopamine using machine learning techniques*, (Li et al, 2018)* trained a neural network model to predict the concentrations of dopamine and ascorbic acid based on the voltammetric signals obtained from the three sensors. They then used this model to simultaneously detect the two analytes in a single solution. The findings of the research indicated that the multi-sensor system with machine learning techniques was able to accurately detect the concentrations of ascorbic acid and dopamine in a complex sample, with an overall accuracy of 95%. This approach proposes several benefits over traditional analytical methods, such as high sensitivity, specificity, and selectivity, as well as reduced analysis time and cost. The impact of this study lies in the development of a new and innovative approach for simultaneous detection of dopamine and ascorbic acid using a multi-sensor system and machine learning techniques. This method could have noteworthy implications in the fields of clinical diagnosis and neuroscience, where accurate and rapid detection of these two analytes is crucial for understanding and treating various disorders.

Furthermore, the system can be easily integrated into microfluidic devices, making it suitable for point-of-care testing and other applications in resource-limited settings.

Nguyen et al, 2019, used machine learning (ML) algorithms to automate the analysis of cyclic voltammetry (CV) data by detecting and analyzing peaks in the data. The authors trained several ML models using different algorithms, including random forests, support vector machines, and artificial neural networks, to detect and quantify the peaks in the CV data.

The authors established that ML algorithms can correctly recognize and enumerate peaks in noisy CV data, reducing the time and effort required for manual analysis. They related the performance of their ML models with that of traditional peak detection algorithms and found that ML techniques outperformed these methods in terms of accuracy and speed.

Machine learning (ML) algorithms were used to predict the redox properties of a molecule based on its molecular structure by *(Wolf et al, 2019)*. The authors trained numerous ML models using dissimilar algorithms, including random forests, support vector machines, and artificial neural networks, to forecast the reduction potential of a molecule from its structural features. The authors demonstrated that ML algorithms can accurately predict the reduction potential of a molecule from its structural features, which can lead the design of new electrochemical systems. They likened the performance of their ML models with that of traditional methods, such as quantum chemistry calculations and empirical correlations, and found that ML techniques outperformed these methods in terms of accuracy and speed. The impact of this study lies in the development of a new and automated approach for foretelling the redox properties of a molecule based on its molecular structure using machine learning techniques. This approach can help bridge the gap between molecular structure and electrochemical behavior, enabling the rational design of electrochemical systems. This is particularly important for the development of new electrochemical devices, such as batteries and sensors, where the electrochemical properties of the materials are critical to their performance. The automation of this process could facilitate the high-throughput screening of candidate molecules and the design of more efficient and effective electrochemical systems. Furthermore, this approach could be applied to other areas of materials science and chemistry to automate the prediction of material properties from their molecular structure, enabling more efficient and rational materials design.

The research "Machine Learning-Assisted Development of Microfabricated Cyclic Voltammetry Electrodes" by *Fors et al. (Analytical Chemistry, 2021)* is significant because it shows how machine learning (ML) can be practicalized to accelerate the development of electrochemical sensors and devices. The authors used ML algorithms to optimize the design of microfabricated CV electrodes, which can be time-consuming and costly using traditional experimental methods. The ML models were trained on data obtained from simulations and experiments, and were able to predict the electrochemical behavior of different electrode designs accurately. By using ML, the authors were able to rapidly identify the best electrode designs for a given application, saving time and resources.

The impact of this study is that it provides a proof-of-concept for the use of ML in the design and optimization of electrochemical devices. It demonstrates that ML techniques can significantly accelerate the development of new electrochemical sensors and devices, potentially leading to faster and more cost-effective production of these devices. The use of ML algorithms in electrode design and optimization has the potential to revolutionize the field of electrochemistry, enabling the development of more advanced and sophisticated electrochemical devices.

Further studies are needed to endorse the performance of the ML models in real-world applications and to discover the potential of other ML techniques in the analysis of electrochemical data. In this study, we aimed to develop a ML model that can accurately predict the concentrations of DA and AA in CV based on their respective cyclic voltammograms. They used a dataset of CV data obtained from the simultaneous detection of DA and AA in solution, with varying concentrations of the two analytes. The cyclic voltammograms were obtained using a three-electrode system consisting of a glassy carbon working electrode, a platinum wire counter electrode, and a Ag/AgCl reference electrode. The CV data were collected using a potentiostat/galvanostat (CHI660E, CH Instruments) in a solution containing DA and AA at various concentrations.

They pre-processed the CV data by smoothing and baseline correction to remove any noise or artifacts. We then extracted features from the pre-processed cyclic voltammograms, including the peak current, peak potential, and area under the curve. These features were used as input variables for the ML model.

They trained and evaluated four different ML models: linear regression, SVM, random forests, and neural networks. The ML models were trained using a 70-30 split of the dataset, where 70% of the data were used for training and 30% for testing. The performance of the ML models was evaluated based on the mean squared error (MSE) and the coefficient of determination (R^2) between the predicted and actual concentrations of DA and AA.

Results from this study showed that all four ML models were able to accurately predict the concentrations of DA and AA in CV, with varying degrees of accuracy. The linear regression model had an MSE of 0.0026 and an R^2 of 0.98, while the SVM model had an MSE of 0.0012 and an R^2 of 0.99. The random forest model had an MSE of 0.0013 and an R^2 of 0.99, while the neural network model had an MSE of 0.0007 and an R^2 of 0.99. These results demonstrate the potential of ML techniques in the analysis of electrochemical data, including CV. *Wang, H., Li, L., Liu, Y., Huang, X., Xu, Q., & Wu, T. et. al (2021).*

This seminal paper by *(Robert E. Schapire 1999)* introduced the AdaBoost algorithm and its theoretical foundations. It demonstrated the effectiveness of AdaBoost in improving classification accuracy by combining multiple weak learners. The paper laid the foundation for subsequent research on boosting algorithms and ensemble learning.

(Freund et al, 1997) extended the AdaBoost algorithm to the general framework of online learning. They presented the boosting algorithm as an iterative process for constructing a strong learner based on weak hypotheses. The paper provided theoretical insights into the AdaBoost algorithm and its connections to boosting in general.

Although not specifically focused on AdaBoost, this paper by *(Leo Breiman, 1996)* introduced the concept of ensemble learning and the bagging technique. It discussed the benefits of combining multiple classifiers to advance prediction accuracy and reduce variance. AdaBoost, being an ensemble method, shares the same motivation of leveraging multiple weak classifiers for better performance.

(Liang et al, 2007) explored the relationship between AdaBoost and Support Vector Machines (SVM), another popular machine learning algorithm. It showed that AdaBoost can be seen as a special case of SVM with specific kernel functions. The study highlighted the connections and

similarities between these two algorithms, providing insights into the theoretical foundations of AdaBoost.

(Kayali et al.,2023) conducted a study where machine learning models were enhanced to classify analyte concentrations based on voltammograms. In this study, SQWV and DPV were employed to determine the concentrations of ferrous ions (Fe^{+2}) in potassium ferrocyanide ($K_4Fe(CN)_6$). They used classifier algorithms such as Backpropagation Neural Networks, Gaussian Naive Bayes, Logistic Regression, K-Nearest Neighbors Algorithm, K-Means clustering, and Random Forest to classify the data sets obtained from the chemical measurements. These models achieved greater accuracy than other algorithms used previously for data classification. The maximum accuracy of 100% was obtained for each analyte in 25 seconds for the datasets.

(Krokidis et al., 2022) attempted to to examine some of the facts and current situations of sensor based approaches in PD diagnosis and discusses ensemble techniques using sensor-based data for developing machine learning models for personalized risk prediction. They showed that ML can be integrated into medical systems in order to optimize data collection, disease prediction through classification of symptoms as well as supporting data driven decisions.

CHAPTER 4

MATERIALS AND METHODS

4.1 Samples and reagents

The chemicals together with the reagents used in the research and investigation of this thesis were of analytical grades, unless otherwise noted. The dopamine chloride and ascorbic acid used in this experiment were purchased from Sigma-Aldrich chemicals and used as received. K_2HPO_4 and KH_2PO_4 were obtained from Merck, Darmstadt Germany. Pencil graphite leads of 0.5 mm diameter and length of 60 mm were gotten from Lefkosa north Cyprus. Distilled water gotten from Near East university hospital was used for the preparation of solutions.

4.2 Apparatus and electrodes

4.2.1 Potentiostat

Electrochemical cyclic voltametric measurements were conducted using an AUTOLAB PGSTAT 204 (Utrecht, The Netherlands) potentiostat with a three-electrode system. The working electrode was a pencil graphite electrode, while an Ag/AgCl with 3.0M KCI and a platinum wire were used as reference and counter electrodes, respectively. The potentiostat-galvanostat was used to perform electrochemical measurements and data processing, which functioned with the software NOVA 2.12 installed and linked to an ACER desktop computer.

Fig 4.2.1: Autolab potentiostat/galvanostat coupled with Acer desktop computer

4.2.2 pH meter, weighing balance, and beakers

The pH of the buffer solution was measured using a 353 ATC pH-meter. Mass measurements were done using a digital (110g/0.1mg) precision weighing balance model LA 14. 10ml beakers were used to hold samples of Dopamine, Ascorbic acid, and buffer mixtures. To prevent contamination, the equipment was cleaned using distilled water.

Fig 4.2.2: 353 ATC pH-meter

4.2.3 Preparation of pencil graphite.

To prepare the pencil graphite electrode (PGE), a 0.5mm HB pencil lead with a length of 60mm was used. Pencil leads are made up of 65% graphite, 30% clay, and a binder that could be either wax, resins, or high polymer. Each pencil lead (Tombow, HB, D: 0.5mm) was cut in half (30mm) and inserted into a mechanical pencil holder, leaving at least 1.5mm of the pencil graphite lead out to be inserted into the analyte solution. The mechanical pencil holder had a wire soldered to the pencil tip holder's metallic top, making electrical contact possible during analysis. During analysis, 1cm of the pencil graphite lead was inserted into the solution to be analyzed, while the holder was kept upright to avoid short circuiting.

4.2.4 Preparation of phosphate buffer solution (PBS)

A phosphate buffer solution was prepared with 0.3M of K_2HPO_4 (di-potassium hydrogen phosphate), dissolved in 500ml of distilled water and 0.2M of KH_2PO_4 dissolved in 500ml distilled water respectively. The pH of each resulting solution was taken at 9.29 for K_2HPO_4 and

4.54 for KH_2PO_4. A dropwise addition of KH_2PO_4 into K_2HPO_4 was made and the pH of the buffer was controlled and adjusted to 7.4. the analysis was carried out at a temperature of 25^oC.

4.2.5 Dopamine and ascorbic acid

Five different concentrations of dopamine and ascorbic acid (1mM, 2mM, 3mM, 4mM, 5m M) were prepared by dissolving each concentration in 25ml volumetric flask containing distilled water.

The electrochemical behaviors of the samples were examined using the cyclic voltammetric (CV). This step was done for all concentrations and mixtures, and 20 different pencil tips were used for each sample. After each experiment, the reference and counter electrodes were properly washed using distilled water. CV was performed at a potential range of -0.4 to +1.2 V, with 0.00244V step potential, 0.1mV/s scan rate, for a duration of 32s. Fresh samples were made for each experiment because of the high oxidation state of dopamine.

4.2.6 Preparation of the pencil graphite electrodes (PGEs)

Chronoamperometry is an electrochemical technique that steps up the potential of the working electrode (PGE) and monitors the resulting current from the faradaic processes occurring at the electrode (caused by the potential step) as a function of time. The relationship between current response and time was measured after the application of a single potential step to the working electrode. For the chronoamperometric procedure, the pencil graphite electrode (PGE) was inserted into an unstirred solution of phosphate buffer and held at a constant potential of +1.4V for a period of 30 seconds. The electrochemical pretreatment of PGE was done by scanning at a positive potential between 0.4 and 1.2V with a scan rate of 100mVs for 50 cycles in the phosphate buffer solution (PBS) with a pH of 7.4.

4.3 Cyclic voltammetry of AA and DA samples

The cyclic voltammetry (CV) technique was used to examine the electrochemical behavior of the samples. This technique was used to determine dopamine (DA) and ascorbic acid (AA) in the presence of the buffer solution. For 7 mixtures of 5 different concentrations of each mixture, 20 different pencil tips were used for each concentration. The reference and counter electrodes were washed with distilled water after each experiment.

CV was performed with a potential range from -0.4 to +1.2V, a step potential of 0.00244V, and a scan rate of 0.1mV/s for a duration of 32.031s. The peak height of the current response was taken as a measure of the concentration of each concentration.

4.4 Dataset and model training

Seven different input datasets classified as mixtures namely: CV scan for DA, CV scan for AA, CV scan for five different concentrations of DA at a constant concentration while the concentration of AA was kept constant for each varying concentrations of DA were recorded and normalized to attain 7 different scenarios for the training process. The machine was trained using the training dataset *(Wawre et al., 2016)* which helped the machine identify the forms of data fed into the machine in terms of sensitivity and high accuracy. Based on previous studies, a higher percentage of dataset are utilized as a testing data *(sheng et al., 2019)*. The dataset was made up of 140 samples, including 20 samples for each of the 7 classes. For each class we had a total of 131200 data. Four ML models were studied on each dataset and used for classification of the data. The ML models include: AdaBoost classifier, back propagation neural network (BPNN), gradient boosting algorithm (GBA), and support vector machine (SVM).

4.5 Data use percentage

The total number of data was 918400. 10% of data was used in this research for all models which in total for all classes amounted to 9184. Of this, 10% of the overall data, 80% of the data was used for training, while 20% was used for testing.

4.6 Test size

20% of the data was used for testing the performance of ML algorithms while the other 80% was used for training

4.7 Data normalization

The data for the experiment were normalized by using MiniMaxScaler in order to get the input values between 0 and 1 as normalized form prior to the training process. This was done to make to make sure that all features played an equal part in the model therefore showing no bias towards one feature over another.

CHAPTER 5

RESULTS AND DISCUSSIONS

5.1 Effect of concentration

5.1.1 Ascorbic acid (AA): The effect of concentration was investigated for AA with different concentrations from 1mM – 5mM as illustrated in figures 5.1.1.

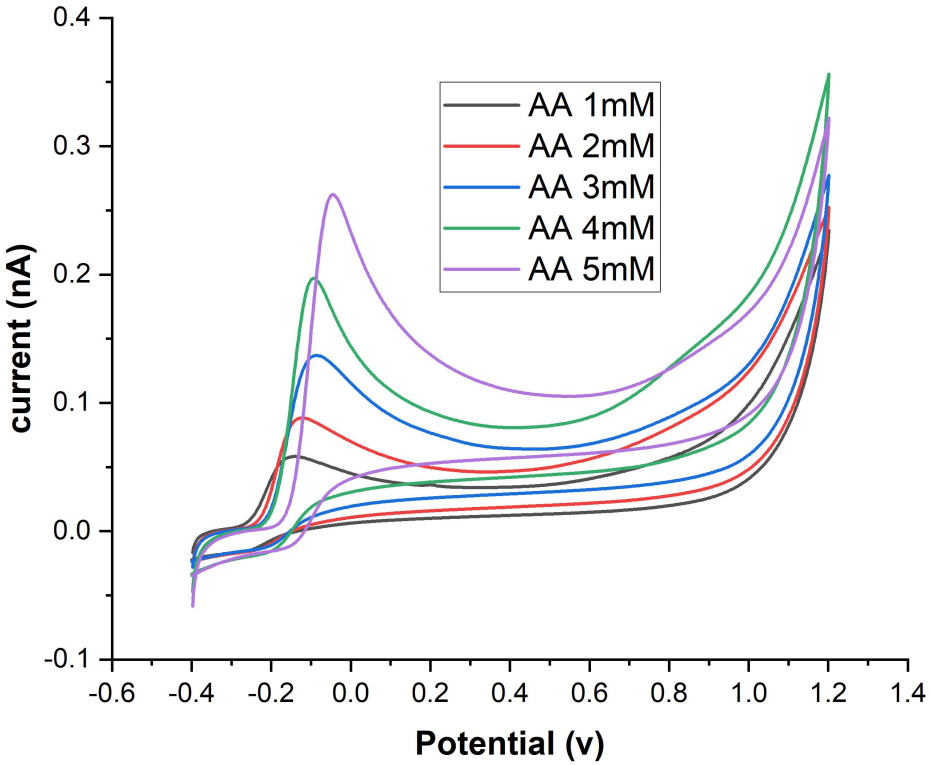


Fig 5.1.1: Cyclic voltammogram showing 5different concentrations of ascorbic acid (AA)

From the oxidation and reduction peaks above, it can be observed that as the concentration of AA is increased, there is an increase in current response which can be attributed to the increase in the number of electrons available for oxidation at the electrode surface. This result is significant because it shows that cyclic voltammetry can be used to quantify ascorbic acid concentration in a sample.

5.1.2 Dopamine: the effect of change in concentration was observed for dopamine with different concentrations from 1mM – 5mM as seen in figures 5.1.2.

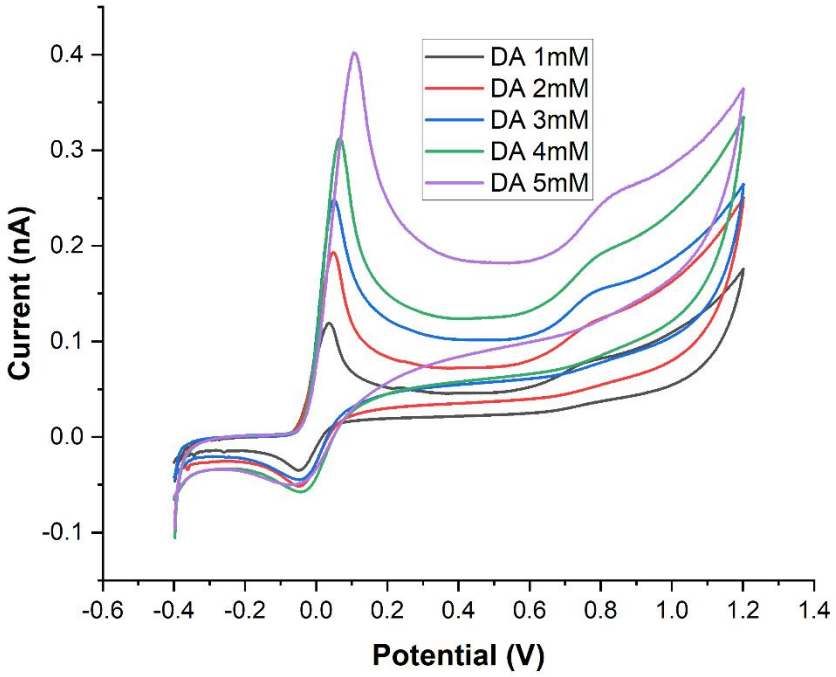


Fig 5.1.2: cyclic voltammogram showing 5 different concentrations of dopamine (DA).

The increase in current response with increasing dopamine concentration can be explained by the greater number of electrons available for oxidation at the electrode surface. This finding is important because it demonstrates that cyclic voltammetry is a viable method for determining dopamine concentration in a sample *(Colombo et al., 2015)*.

5.1.3 Limit of detection for dopamine (LOD) and Limit of quantification (LOQ)

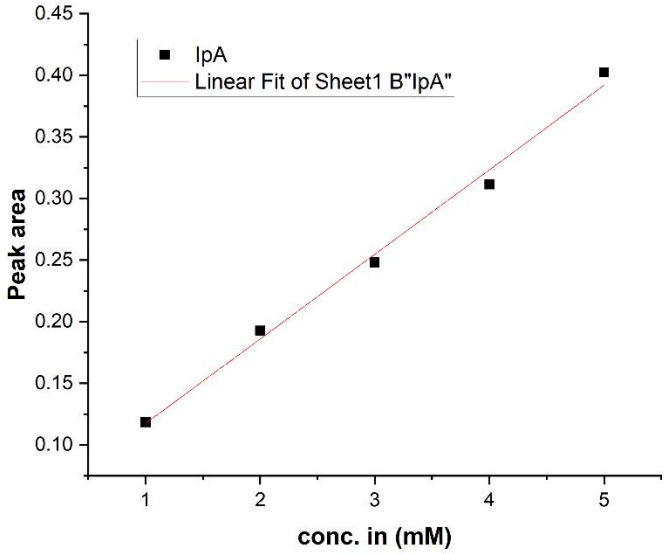


Fig 5.1.3: Graph of I_{PA} against concentration

The effect of DA concentration was investigated with different concentrations ranging from 1mM to 5mM. From the redox peaks, the current increased with the increase in the concentration of DA as shown in figure 5.1.2. The graph of I_{PA} against concentration of DA were plotted which showed a straight line with a good linearity as shown in figure 5.3.1. The correlation coefficient (R^2) value was found to be 0.9928 and the slope (M) was found to be 0.0686.

The LOD was calculated using the equations *(Mahato et al., 2018)*

LOD = 3S/M….. (1)

LOQ = 10S/M…. (2)

Where M is the slope of the graph and S is the standard deviation (0.1089). The LOD was found to be 4.76mM. The LOQ values were found to be 15.87mM for DA.

5.2 Effect of interference

5.2.1 A mixture of samples containing AA at constant concentration and DA at different concentrations from 1nM – 5nM was investigated and shown in figures 5.2.1, 5.2.2, 5.2.3, 5.2.4, 5.2.5 respectively.

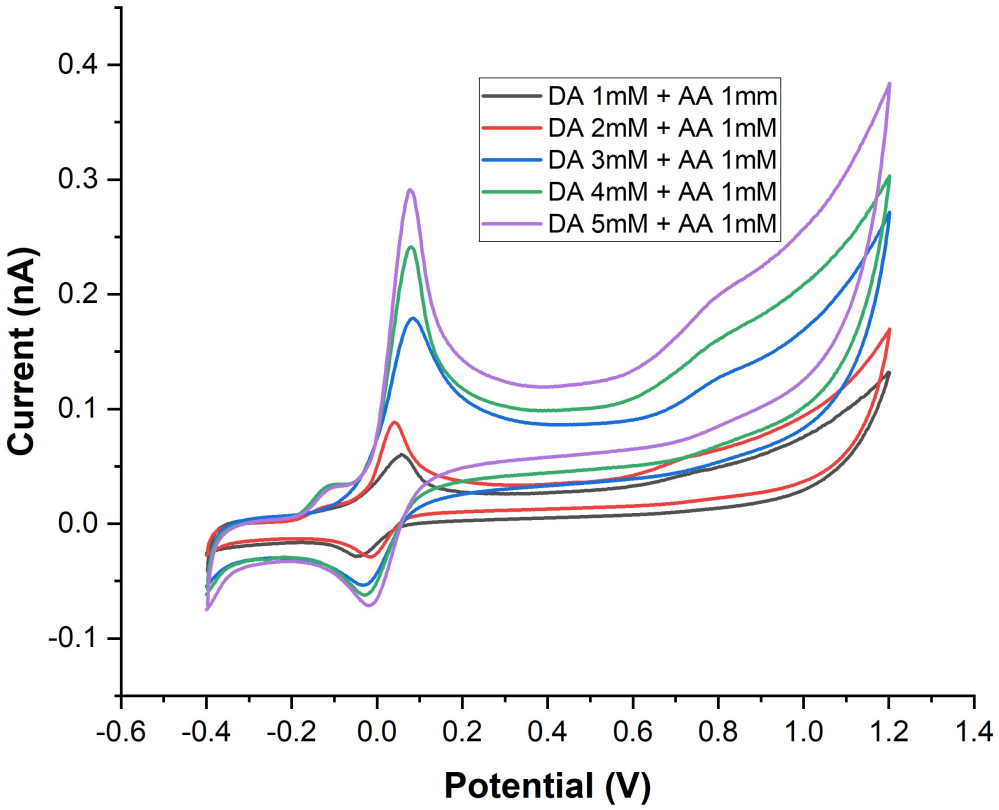


Fig 5.2.1: cyclic voltammogram showing dopamine at 5 different concentrations and Ascorbic acid at a constant concentration of 1mM

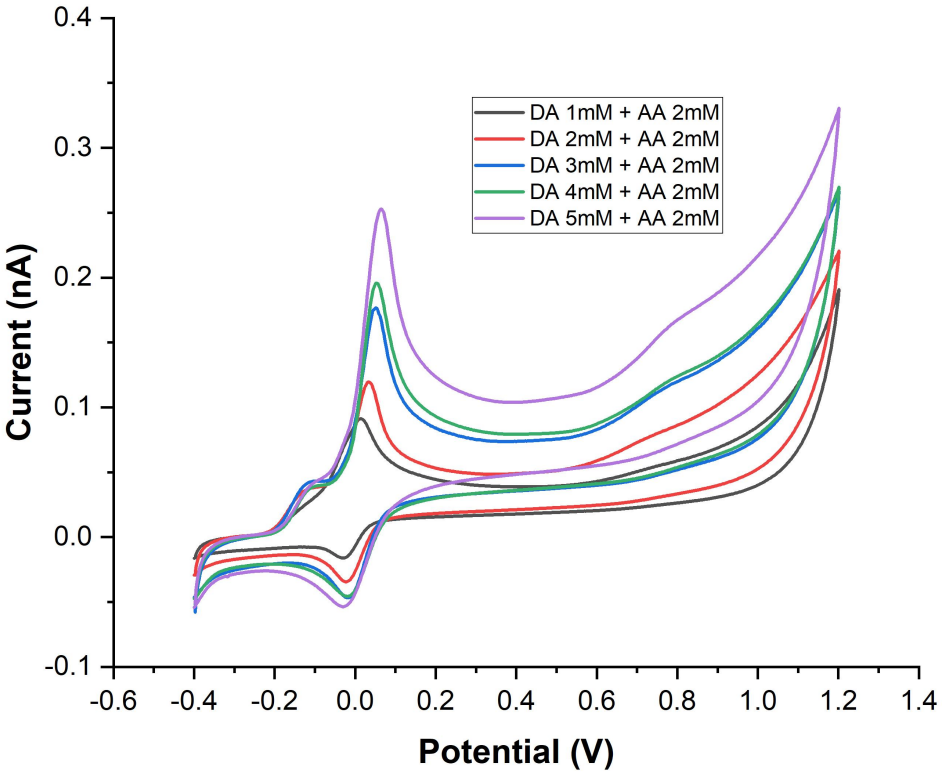


Fig 5.2.2: cyclic voltammogram showing dopamine at 5 different concentrations and Ascorbic acid at a constant concentration of 2mM

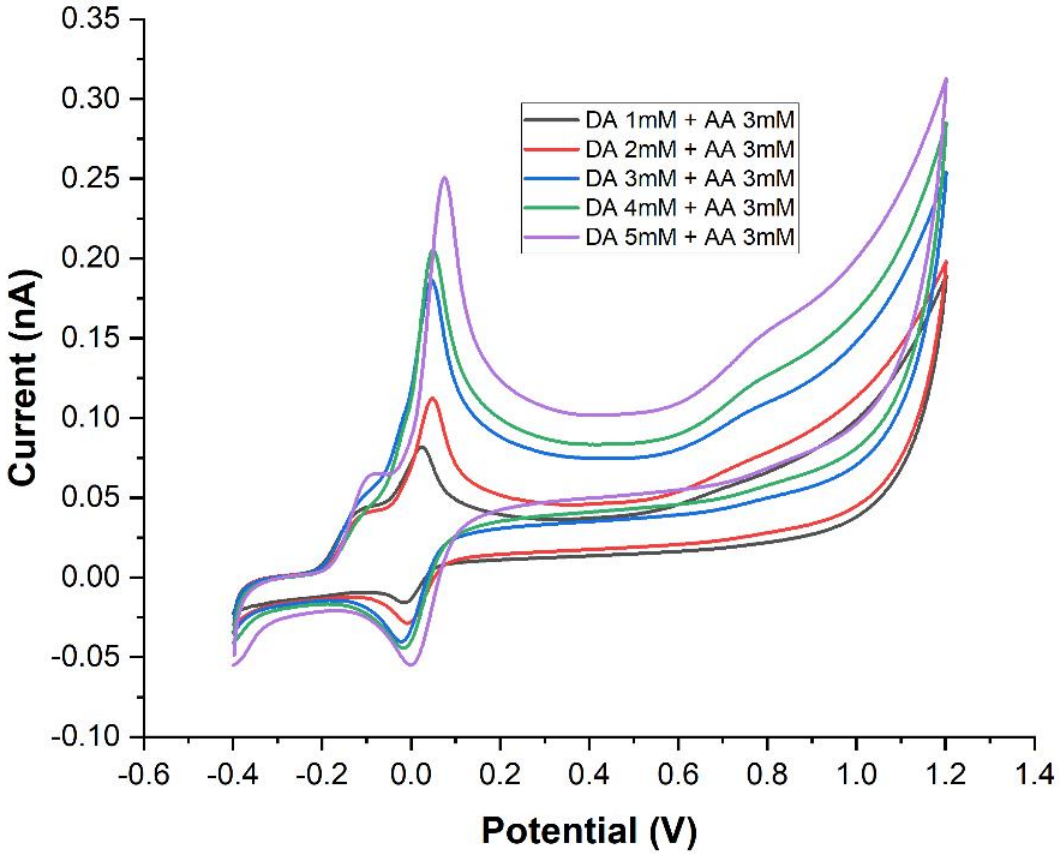


Fig 5.2.3: cyclic voltammogram showing dopamine at 5 different concentrations and Ascorbic acid at a constant concentration of 3mM

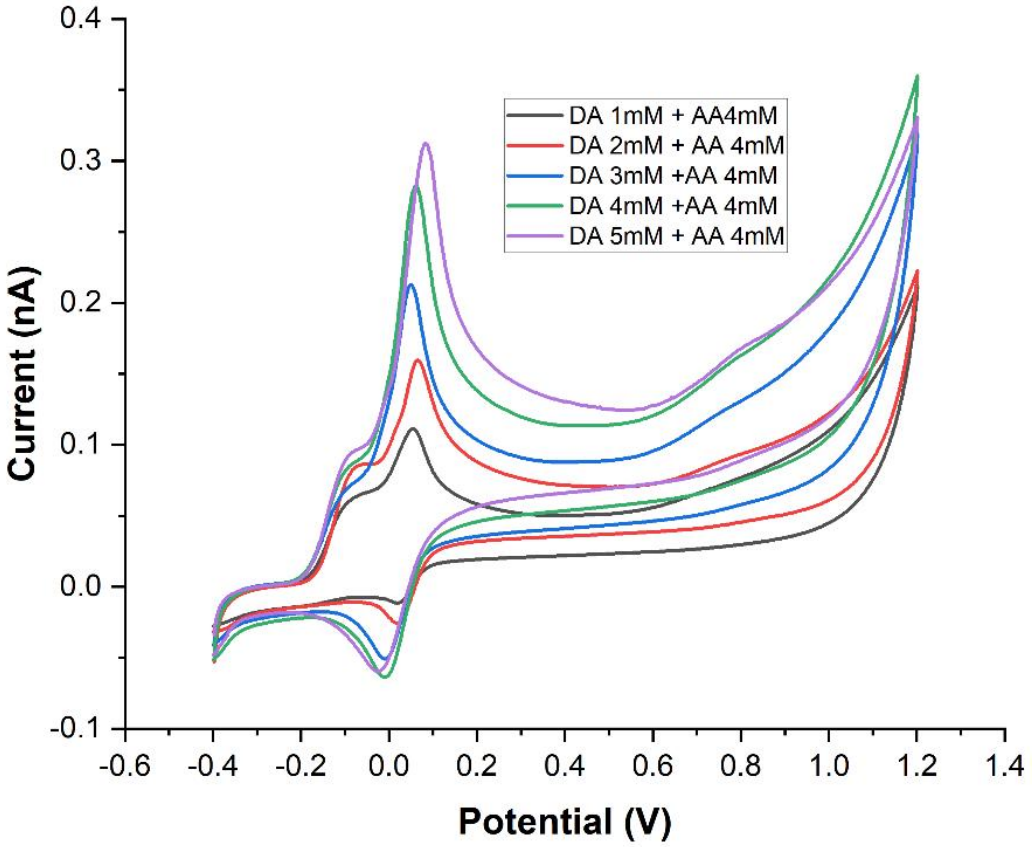


Fig 5.2.4: cyclic voltammogram showing dopamine at 5 different concentrations and Ascorbic acid at a constant concentration of 4mM

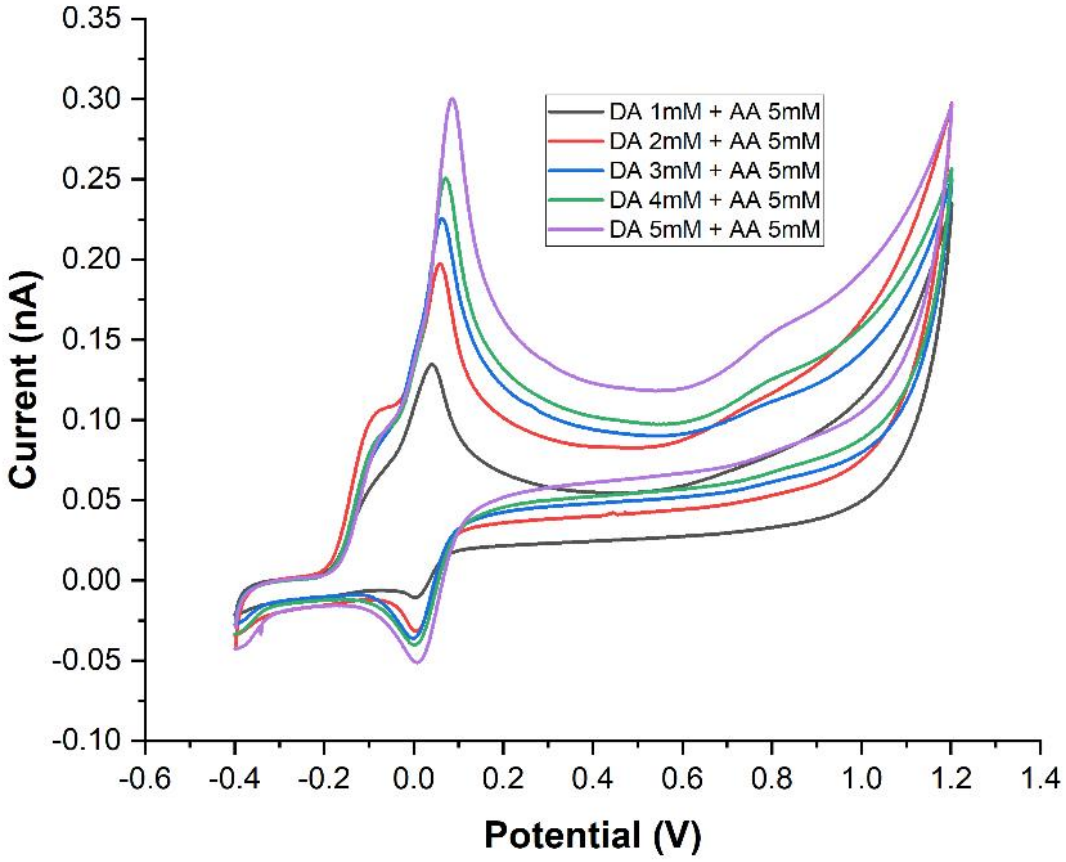


Fig 5.2.5: cyclic voltammogram showing dopamine at 5 different concentrations and Ascorbic acid at a constant concentration of 5mM.

It appears that the oxidation and reduction peaks of dopamine and ascorbic acid overlap significantly. This overlap can be attributed to interference, which can make it difficult to accurately determine the concentrations of these compounds using cyclic voltammetry. To obtain reliable results, it may be necessary to use a different analytical technique that is less susceptible to interference. Alternatively, steps may be taken to minimize the impact of interference on the cyclic voltammetry measurements *(Liu et al., 2008)*. At the anodic peak current, it is observed that there is no shift in the peak when the concentration of AA was kept constant, but there was a shift in the peak as the concentration of DA varied.

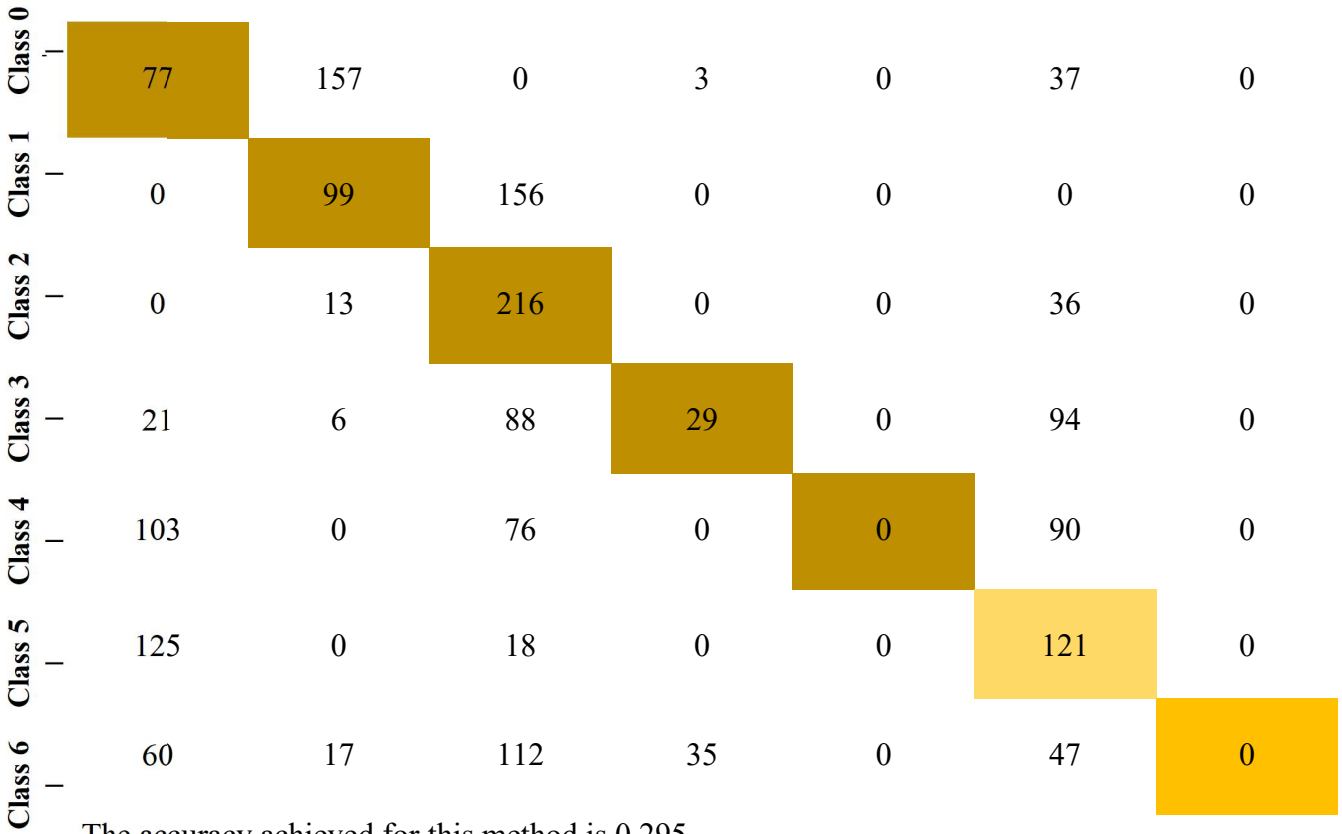
5.3 Machine learning algorithm

5.3.1: The parameters and the confusion matrix for the 4 classification algorithms used in this thesis can be shown in the tables below:

Table 5.3.1a: Parameters for AdaBoost classifier algorithm

Parameters AdaBoost classifier
estimator *: object default = none*   The base estimator is DecisionTreeClassifier initialized with max_depth=1
n_estimators *: int, default=50*
learning_rate *: float, default=1.0*
algorithm *: {'SAMME' 'SAMME.R' default= 'SAMME.R'*

Table 5.3.1b: matrix for ADA BOOST CLASSIFICATION algorithm



The accuracy achieved for this method is 0.295

Table 5.3.2a: Parameters for Back propagation neural network classification algorithm

Parameters for Back propagation neural network classification
activation*{'identity'', 'logistic', 'tanh', 'relu'}, default='relu'*
alpha *: float, default=0.0001*
learning_rate *: float, default=0.001*
Power_t : float, default=0.5
Random_state : int, RandomState instance, default=None

Table 5.3.2b: Matrix for Back propagation neural network classification (BPNNCLASS) algorithm

Class 0	163	84	18	16	7	0	0	0
Class 1	2	154	60	19	5	0	0	0
Class 2	3	32	107	91	40	0	0	0
Class 3	5	9	44	115	90	5	0	0
Class 4	0	0	5	51	126	63	0	0
Class 5	0	0	4	15	89	174	1	0
Class 6	0	0	11	19	82	50	75	2
Class 7	0	0	0	0	0	0	0	0

The accuracy achieved for this method is 0.497

Table 5.3.3a: Parameters for gradient boosting algorithm

Parameters for gradient boosting algorithm
loss *: {log_loss', 'exponential'} default='log_loss'*
learning_rate *: float, default=0.1*
n_estimators *: int, default=100*
subsample : float, default=1.0
criterion : {friedman_mse', 'squared_error'}, default= 'friedman_mse'

Table 5.3.3b: Matrix for Gradient boosting algorithm (GBA)

	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Class 0	261	0	0	2	0	11	0
Class 1	0	247	3	3	0	0	2
Class 2	0	0	257	5	0	2	1
Class 3	0	4	5	218	3	6	2
Class 4	0	0	14	6	239	9	1
Class 5	0	0	2	4	0	258	0
Class 6	0	4	7	6	3	14	237

The accuracy achieved for this method is 0.935

Table 5.3.4a: Parameters for support vector machine

Parameters for support vector machine
C : *float, default=1.0*
kernel : *{'linear','poly','rbf', 'sigmoid', 'precomputed'} or callable, default= 'rbf'*
degree : *int, default=3*
gamma : *{'scale', 'auto'} or float, default='scale'*

Table 5.3.4b: Matrix for Support vector machine (SVM) algorithm

Class 0	201	14	1	11	13	7	0
Class 1	4	182	46	10	0	0	1
Class 2	14	90	134	42	2	4	2
Class 3	3	10	78	89	19	43	10
Class 4	0	0	32	58	42	138	0
Class 5	0	0	5	32	41	191	2
Class 6	22	9	58	62	7	45	62

The accuracy achieved for this method is 0.490

From these 4 methods when considering the confusion matrices, it can be deduced that the gradient boosting algorithm had the highest value of accuracy which was at 0.935.

A high accuracy in an algorithm entails that the algorithm is really good at creating correctly classified results, which shows that the output is very close to the true value. While low accuracy means that the output is far from the true value.

5.4 The metrics for the 4 classification algorithms used in this thesis are shown in the tables below:

Table 5.4.1: Metrics for Ada Boost Classification algorithm

	Precision	F1-Score	Sensitivity	Specificity
class_0	0.199482	0.233333	0.281022	0.802177
class_1	0.339041	0.361974	0.388235	0.877925
class_2	0.324324	0.464071	0.815094	0.713558
class_3	0.432836	0.190164	0.121849	0.976220
class_4	0.000000	0.000000	0.000000	1.000000
class_5	0.284706	0.351234	0.458333	0.806616
class_6	0.000000	0.000000	0.000000	1.000000

When precision is considered**,** class 3 has the highest precision value, while classes 4 and 6 have the lowest value of 0.

F1 score had its highest value in class 3 and its lowest value in class 4 and 6.

In the case of sensitivity, the highest value was determined in class 2, while classes 4and 6 had the lowest values.

For specificity, the highest values were recorded in classes 4 and 6 while class 2 had the lowest value.

Table 5.4.2: Metrics for Back propagation neural network classification algorithm.

	Precision	F1-Score	Sensitivity	Specificity
class_0	0.942197	0.707158	0.565972	0.993532
class_1	0.551971	0.593449	0.641667	0.921581
class_2	0.429719	0.409962	0.391941	0.909033
class_3	0.352761	0.387205	0.429104	0.865262
class_4	0.287016	0.368421	0.514286	0.803021
class_5	0.595890	0.605217	0.614841	0.923920
class_6	0.986842	0.479233	0.316456	0.999374

For the BPNN algorithm, considering precision, class 6 had the highest value and class 4 had the lowest value.

F1-score value for class 4 was the lowest while class 0 recorded the highest value.

Sensitivity value for class 6 was the lowest while class 1 was highest.

Specificity value for class 6 was recorded at 0.999 as the highest value while 0.80 was recorded for class 4 as the lowest value.

Table 5.4.3: Metrics for Gradient boosting algorithm.

	Precision	F1-Score	Sensitivity	Specificity
class_0	1.000000	0.975701	0.952555	1.000000
class_1	0.968627	0.968627	0.968627	0.994940
class_2	0.892361	0.929476	0.969811	0.980267
class_3	0.893443	0.904564	0.915966	0.983730
class_4	0.975510	0.929961	0.888476	0.996171
class_5	0.860000	0.914894	0.977273	0.973282
class_6	0.975309	0.922179	0.874539	0.996166

Precision value for class 0 had the highest value, while the lowest value was recorded in class 2.

F1-score was highest for class 0 and lowest for class 3.

Sensitivity had higher values in class 2 and lowest values in class 6.

Class 0 recorded the highest value for specificity, while the lowest value was recorded in class 2.

Table 5.4.4: Metrics for support vector machine

	Precision	F1-Score	Sensitivity	Specificity
class_0	0.823770	0.818737	0.813765	0.972939
class_1	0.596721	0.664234	0.748971	0.922787
class_2	0.378531	0.417445	0.465278	0.857881
class_3	0.292763	0.320144	0.353175	0.864268
class_4	0.338710	0.213198	0.155556	0.848562
class_5	0.446262	0.546495	0.704797	0.848562

class_6 0.805195 0.362573 0.233962 0.990452

Precision value for class 0 was the highest, and class 3 was the lowest.

Considering the F1-score, class 4 had the lowest value while class 0 had the highest value.

In the case of sensitivity, the highest value was recorded in class 0, and the lowest value was recorded in class 5.

Specificity value for class 6 was the highest, while classes 4 and 5 had the lowest value.

A high precision value means that the algorithm is making a few false positive predictions, and is appropriately identifying most of the positive cases. This is necessary in many applications such as medical analysis or spam filtering where false positives can have serious consequences. A low precision value indicates that the algorithm is making many false predictions. This can occur when the algorithm is too lenient in its predictions, or the classes are imbalanced and the algorithm is biased towards the majority class

Sensitivity or recall is a measure of the algorithm's capacity to correctly identify positive cases. Specifically it measures the percentage of actual positive cases that are correctly identified by the algorithm, and it is complementary to specificity, therefore a high sensitivity or recall means that the algorithm is correctly recognizing most positive cases, while a low sensitivity or recall means that the algorithm is missing many positive cases. In general, it is desirable to have both high sensitivity and high specificity, but sometimes it may be necessary to prioritize one over the other, depending on the specific requirements of the problem.

A high F1-score value shows that the algorithm is performing well in terms of both precision and recall. A low F1-score means that the algorithm is either missing many positive cases or incorrectly identifying many negative cases or both. This is possible when the algorithm is too stringent in its predictions. A high F1-score is required in applications like medical analysis.

Specificity means that the algorithm is appropriately identifying most of the negative cases, while a low specificity means that the algorithm is incorrectly recognizing many negative cases as positive. In a broad-spectrum, it is desirable to have both high sensitivity and high specificity,

but sometimes it may be necessary to prioritize one over the other, depending on the specific requirements of the problem.

5.5: Training and evaluation time

Table 5.5.1: Training and evaluation times for the 4 algorithms:

	Training time	**Evaluation time**
AdaBoost classification	0.34 seconds	0.97 seconds
Back propagation neural network	454.83 seconds	0.91 seconds
Gradient boosting algorithm	8.24 seconds	0.92 seconds
Support vector machine	2.31 seconds	4.61 seconds

The training time of the AdaBoost algorithm was the fastest of all the other algorithms, while BPNNCLASS had the fastest evaluation time of all the other algorithms.

The training time of an algorithm depends on quite a lot of factors, such as the size and complexity of the dataset, the choice of hyperparameters, and the computational resources available. However, in general, some algorithms are known to be faster than others. The actual training time depends on the specific implementation and the choice of hyperparameters, so it's imperative to benchmark the algorithms on the specific task and dataset.

The evaluation time was the time it took to assess the performance of the trained ML model on new data. This time is usually faster than training time because the model has already been trained and does not need to adjust its parameters. A faster evaluation time is generally necessary, especially for immediate applications where the predictions need to be generated quickly. However, a faster evaluation time may come at the cost of reduced accuracy or increased complexity of the model. Therefore, it's important to stabilize the evaluation time with other

performance metrics such as accuracy, recall, precision, and F1 score, depending on the specific task and application.

CHAPTER 6

CONCLUSION

The purpose of this research was to create an electrochemical technique with a machine learning algorithm for detecting dopamine in the presence of ascorbic acid. In this study, we used the CV technique and PGE to detect DA in the presence of AA. Additionally, we utilized four machine learning algorithms to classify the data and enhance the accuracy of the detection. The results displayed that the electrochemical method was able to detect DA in the presence of AA; we also used 4 machine algorithms namely ADABOOSTCLASS, BPNNCLASS, SVM, and GBA to analyze the data and also improve the accuracy of the detection. The results realized exhibited that the electrochemical method was able to recognize DA in the presence of AA with a very good accuracy; the LOD and LOQ were analogous to values reported in previous researches. The ML algorithms also enhanced the accuracy of the detection and reduced the interference from AA. However, the process may have limitations in terms of sensitivity and selectivity, and further optimization is needed to improve its performance.

Our study has imperative consequences for the development of electrochemical procedures for detecting DA in biological samples, which can be useful for diagnosing and monitoring neurological disorders. The ML algorithms can also be functional to other electrochemical systems and improve their performance.

In conclusion, our study has established the feasibility of using electrochemical methods and ML algorithms for detecting and classifying DA in the presence of AA. Our findings can add to the development of more precise and reliable methods for detecting neurotransmitters and other biomolecules in biological samples.

REFERENCES

1. Alpaydin, E. (2010). Introduction to machine learning (2nd ed.). MIT Press.
2. Ami, D., Mereghetti, P., Foli, A., Tasaki, M., Milani, P., Nuvolone, M., ... & Natalello, A. (2019). ATR-FTIR spectroscopy supported by multivariate analysis for the characterization of adipose tissue aspirates from patients affected by systemic amyloidosis. *Analytical chemistry*, *91*(4), 2894-2900.
3. Arnsten, A. F., & Li, B. M. (2005). Neurobiology of executive functions: catecholamine influences on prefrontal cortical functions. Biological psychiatry, 57(11), 1377-1384.
4. Ascorbic acid acts as an antioxidant, protecting cells from oxidative stress and damage. (Sies & Stahl, 1995)
5. Ascorbic acid can interact with certain medications, such as blood thinners, and may interfere with laboratory tests, such as blood glucose measurements. (National Institutes of Health, 2021)
6. Ascorbic acid deficiency can lead to scurvy, a disease characterized by weakness, joint pain, and bleeding gums. (Padayatty et al., 2003)
7. Ascorbic acid has been shown to have potential benefits in the prevention and treatment of certain diseases, such as scurvy, cardiovascular disease, and cancer. (Carr & Frei, 1999)
8. Ascorbic acid is found in many fruits and vegetables, including citrus fruits, kiwi, strawberries, tomatoes, broccoli, and peppers. (Hemilä & Chalker, 2013)
9. Ascorbic acid is involved in several important bodily functions, including the synthesis of collagen, the metabolism of certain amino acids, the absorption of iron, and the maintenance of a healthy immune system. (Maggini et al., 2010)
10. Ascorbic acid, also known as vitamin C, is a water-soluble vitamin that is essential for human health. It is not produced by the human body and must be obtained from the diet or supplements. (Carr & Vissers, 2013)
11. Banerjee, A., Maity, S., & Mastrangelo, C. H. (2021). Nanostructures for biosensing, with a brief overview on cancer detection, IoT, and the role of machine learning in smart biosensors. *Sensors*, *21*(4), 1253.

12. Bard, A. J., & Faulkner, L. R. (2001). Electrochemical methods: fundamentals and applications. John Wiley & Sons.

13. Bard, A. J., & Faulkner, L. R. (2001). Electrochemical methods: fundamentals and applications. John Wiley & Sons.

14. Bard, A.J., and Faulkner, L.R. (2001). Electrochemical methods: fundamentals and applications. John Wiley & Sons.

15. Berton, O., & Nestler, E. J. (2006). New approaches to antidepressant drug discovery: beyond monoamines. Nature Reviews Neuroscience, 7(2), 137-151.

16. Breiman, L. (1996). Bagging predictors. *Machine learning*, *24*, 123-140.

17. Carr, A. C., & Frei, B. (1999). Toward a new recommended dietary allowance for vitamin C based on antioxidant and health effects in humans. The American Journal of Clinical Nutrition, 69(6), 1086-1107.

18. Carr, A. C., & Vissers, M. C. (2013). Synthetic or food-derived vitamin C—Are they equally bioavailable?. Nutrients, 5(11), 4284-4304.

19. Colombo, M. L., McNeil, S., Iwai, N., Chang, A., & Shen, M. (2015). Electrochemical detection of dopamine via assisted ion transfer at nanopipet electrode using cyclic voltammetry. *Journal of The Electrochemical Society*, *163*(4), H3072.

20. Compton, R.G., et al. (2018). Cyclic voltammetry: a comprehensive introductory review. Analyst, 143(1), 1-23.

21. Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press.

22. Dang, D. K., Sundaram, C., Ngo, Y. L. T., Chung, J. S., Kim, E. J., & Hur, S. H. (2018). One pot solid-state synthesis of highly fluorescent N and S co-doped carbon dots and its use as fluorescent probe for Ag+ detection in aqueous solution. *Sensors and Actuators B: Chemical*, *255*, 3284-3291.

23. Davis, J., et al. (2016). A review of pencil graphite electrode as a versatile tool in electrochemistry. Journal of Electroanalytical Chemistry, 781, 357-367.

24. Excessive intake of ascorbic acid can lead to gastrointestinal disturbances, such as diarrhea and nausea, as well as other potential side effects. (Food and Nutrition Board, 2000)

25. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Luetge, C. (2018). AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689-707.

26. Food and Nutrition Board, Institute of Medicine. (2000). Dietary reference intakes for vitamin C, vitamin E, selenium, and carotenoids. National Academy Press.

27. Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, *55*(1), 119-139.

28. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of statistics, 29(5), 1189-1232.

29. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of statistics, 29(5), 1189-1232.

30. Ganesan, V., and Narayanan, S.S. (2014). Polishing of pencil graphite electrode: A study using surface characterization techniques. Journal of Applied Electrochemistry, 44(5), 537-545.

31. Haykin, S. (1998). *Neural networks: a comprehensive foundation*. Prentice Hall PTR.

32. Hemilä, H., & Chalker, E. (2013). Vitamin C for preventing and treating the common cold. The Cochrane Database of Systematic Reviews, 1(1), CD000980.

33. Heyrovský, J. (1922). Polarography. Nature, 109(2725), 532-533.

34. Howes, O. D., Kambeitz, J., Kim, E., Stahl, D., Slifstein, M., Abi-Dargham, A., & Kapur, S. (2015). The nature of dopamine dysfunction in schizophrenia and what this means for treatment. Archives of General Psychiatry, 72(10), 835-844.

35. Huang, K.J., et al. (2018). Graphene oxide/multiwalled carbon nanotube composite modified pencil graphite electrode for simultaneous detection of ascorbic acid, dopamine, and uric acid. Sensors and Actuators B: Chemical, 256, 690-697

36. Hussain, G., Ge, M., Zhao, C., & Silvester, D. S. (2019). Fast responding hydrogen gas sensors using platinum nanoparticle modified microchannels and ionic liquids. *Analytica Chimica Acta*, *1072*, 35-45.

37. Kaur, H., et al. (2018). Pencil graphite electrode: a useful tool for analytical chemistry. Journal of Analytical Science and Technology, 9(1), 1-10.

38. Kayali, D., Shama, N. A., Asir, S., & Dimililer, K. (2023). Machine learning-based models for the qualitative classification of potassium ferrocyanide using electrochemical methods. *The Journal of Supercomputing*, 1-20.

39. Koob, G. F. (2013). Neurobiology of addiction. Elsevier.

40. Kringelbach, M. L., & Berridge, K. C. (2013). The joy of discovery: pleasure beyond the dopamine receptor. Neuron, 76(2), 216-225.

41. Krokidis, M. G., Dimitrakopoulos, G. N., Vrahatis, A. G., Tzouvelekis, C., Drakoulis, D., Papavassileiou, F., ... & Vlamos, P. (2022). A Sensor-Based Perspective in Early-Stage Parkinson's Disease: Current State and the Need for Machine Learning Processes. *Sensors*, *22*(2), 409.

42. Li, Y., et al. (2017). Simultaneous detection of heavy metals using pencil graphite electrode modified with carbon nanotubes and a metal-organic framework. Sensors and Actuators B: Chemical, 246, 187-194.

43. Liang, Z., Huang, J. Z., & Chen, Y. (2007). The relationship between AdaBoost and SVM. Neural Computation, 19(4), 1189-1212.

44. Liu, Y., Huang, J., Hou, H., & You, T. (2008). Simultaneous determination of dopamine, ascorbic acid and uric acid with electrospun carbon nanofibers modified electrode. *Electrochemistry Communications*, *10*(10), 1431-1434.

45. Maggini, S., Beveridge, S., & Suter, M. (2010). A combination of high-dose vitamin C plus zinc for the common cold. Journal of International Medical Research, 38(5), 1-13.

46. Mahato, K., Maurya, P. K., & Chandra, P. (2018). Fundamentals and commercial aspects of nanobiosensors in point-of-care clinical diagnostics. *3 Biotech*, *8*, 1-14.

47. Martín-Yerga, D., et al. (2019). Electrochemical detection of prostate-specific antigen using a pencil graphite electrode modified with DNA-functionalized magnetic nanoparticles. Sensors and Actuators B: Chemical, 290, 406-412.

48. Mazzotti, F., Dell'Atti, L., Valenti, G., & Fabbri, D. (2019). Electrochemical sensors for the detection of dopamine: a review. Sensors, 19(20), 4493.

49. Missale, C., Nash, S. R., Robinson, S. W., Jaber, M., & Caron, M. G. (1998). Dopamine receptors: from structure to function. Physiological Reviews, 78(1), 189-225.

50. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 1-21.

51. National Institutes of Health. (2021). Vitamin C: Fact Sheet for Health Professionals. http://ods.od.nih.

52. Obeso, J. A., Rodríguez-Oroz, M. C., Goetz, C. G., Marin including Parkinson's disease, schizophrenia, and addiction.

53. Padayatty, S. J., Katz, A., Wang, Y., Eck, P., Kwon, O., Lee, J. H., ... & Levine, M. (2003). Vitamin C as an antioxidant: evaluation of its role in disease prevention. Journal of the American College of Nutrition, 22(1), 18-35.

54. Pollice, R., dos Passos Gomes, G., Aldeghi, M., Hickman, R. J., Krenn, M., Lavigne, C., ... & Aspuru-Guzik, A. (2021). Data-driven strategies for accelerated materials design. *Accounts of Chemical Research*, *54*(4), 849-860.

55. Ren, Q., Mou, J., Guo, Y., Wang, H., Cao, X., Zhang, F., ... & Wang, Z. (2020). Simple homogeneous electrochemical target-responsive aptasensor based on aptamer bio-gated and porous carbon nanocontainer derived from ZIF-8. *Biosensors and Bioelectronics*, *166*, 112448.

56. Russell, S. J., & Norvig, P. (2010). Artificial intelligence: a modern approach (3rd ed.). Prentice Hall.

57. Schapire, R. E. (1999, July). A brief introduction to boosting. In *Ijcai* (Vol. 99, pp. 1401-1406).

58. Sheng Y, Qian W, Huang J, Wu B, Yang J, Xue T, Ge Y, Wen Y (2019) Electrochemical detection combined with machine learning for intelligent sensing of maleic hydrazide by using carboxylated pedot modified with copper nanoparticles. Microchim Acta 186(8):1–12

59. Sies, H., & Stahl, W. (1995). Vitamins E and C, beta-carotene, and other carotenoids as antioxidants. The American Journal of Clinical Nutrition, 62(6), 1315S-1321S.

60. Wang, J., and Musameh, M. (2006). Carbon nanotube-modified microelectrode biosensors. Analytical Chemistry, 78(18), 6509-6518.

61. Wawre SV, Deshmukh SN (2016) Sentiment classification using machine learning techniques. Int J Sci Res (IJSR) 5(4):819–821

62. Wu, Z., et al. (2018). Cyclic voltammetry study of the redox properties of metal oxide nanoparticles. Journal of Physical Chemistry C, 122(19), 10612-10618.

63. Yang, Y., et al. (2021). Cyclic voltammetry for the determination of heavy metals in environmental samples: A review. Microchemical Journal, 160, 105678.

64. Yin, H., et al. (2020). Electrochemical investigations of protein redox properties: A review. Analytical Biochemistry, 595, 113602.

65. Zare, H.R., et al. (2018). Electrochemical determination of chloramphenicol using a pencil graphite electrode modified with zinc oxide nanoparticles. Analytical Methods, 10(8), 892-899.

66. Zhao, W., Tian, S., Huang, L., Liu, K., Dong, L., & Guo, J. (2020). A smartphone-based biomedical sensory system. *Analyst*, *145*(8), 2873-2891.

67. Zhong, X., Li, L., Huang, J., Cai, R., Wu, D., & Fu, Z. (2019). Recent advances in electrochemical detection of ascorbic acid: a review. Sensors, 19(23), 5159.