MUSTAFA IBRAHIM WAFA
DEVELOPING AN ALTERNATIVE HIGH-PERFORMANCE HYBRID MULTILINEAR REGRESSION-BASED INTELLIGENCE MODEL FOR QUANTITATIVE STRUCTURE-PROPERTY RELATIONSHIP (QSPR) STUDIES OF MYCOTOXINS IN FOODS
NEU
2020

DEVELOPING AN ALTERNATIVE HIGH-PERFORMANCE HYBRID MULTILINEAR REGRESSION-BASED INTELLIGENCE MODEL FOR QUANTITATIVE STRUCTURE-PROPERTY RELATIONSHIP (QSPR) STUDIES OF MYCOTOXINS IN FOODS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES OF NEAR EAST UNIVERSITY

By
MUSTAFA IBRAHIM WAFA

In Partial Fulfillment of the Requirements for the Degree of Master of Science In Food Engineering

NICOSIA, 2020

DEVELOPING AN ALTERNATIVE HIGH-PERFORMANCE HYBRID MULTILINEAR REGRESSION-BASED INTELLIGENCE MODEL FOR QUANTITATIVE STRUCTURE-PROPERTY RELATIONSHIP (QSPR) STUDIES OF MYCOTOXINS IN FOODS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES OF NEAR EAST UNIVERSITY

By
MUSTAFA IBRAHIM WAFA

In Partial Fulfillment of the Requirements for the Degree of Master of Science In Food Engineering

NICOSIA, 2020

Mustafa Ibrahim WAFA: DEVELOPING AN ALTERNATIVE HIGH-PERFORMANCE HYBRID MULTILINEAR REGRESSION-BASED INTELLIGENCE MODEL FOR QUANTITATIVE STRUCTURE-PROPERTY RELATIONSHIP (QSPR) STUDIES OF MYCOTOXINS IN FOODS

Approval of Director of Graduate School of

Applied Sciences

Prof. Dr. Nadire CAVUS

We certify this thesis is satisfactory for the award of the degree of Masters of Science in

Food Engineering

Examining Committee in Charge:

Assist. Prof. Dr. Perihan Adun 	Committee Chairman, Department of Food Engineering, NEU
Assist. Prof. Dr. Kaan Uyar 	Department of Computer Engineering, NEU
Assist. Prof. Dr. Kehinde Adesina 	Supervisor, Department of Food Engineering, NEU

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have finally cited and referenced all materials and results that are not original to this work.

Name, Last Name: Mustafa Ibrahim Wafa

Signature 

Date: 05.12.2020

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my parents, wife, daughter, brothers, sisters and friends who are always there for me. Your support and trust in me inspired and encouraged me to do my best in life. Thank you for everything you have done for me. Special thanks to my Supervisor Assist. Prof. Dr. Kehinde Adewale Adesina, I would like to thank you for supporting me, encouraging my thesis. You were always ready to help and advise me. I'm really glad to have such a Supervisor for my thesis like you. I am truly thankful to the Food Engineering Chairperson, Assist. Prof. Dr. Perihan Aysal Adun and to all my instructors at the Near East University. I appreciate them for their time, patience and consistent help.

To my parents, wife and daughter…

ABSTRACT

Mycotoxins are group of fungi that are of general health concern especially for humans and other animals. The use of data-intelligent oriented approach such as ANN, MLR, SVM, and ANFIS has been established for the quantitative structure–property relationship (QSPR) studies of mycotoxins. In this thesis, these four data-driven model approaches are applied for QSPR examination of the retention time (tR) of various classes of mycotoxins using the quantification and elicitation data obtained by HPLC chromatographic technique. This is with view of developing a new hybrid MLR based intelligence model. For the QSPR modeling, four input variables - retention index, peak symmetry, mono isotopic mass and relative sensitivity factor are used. Then Spearman Pearson correlation is employed to measure the relationships between these input variables with the retention time, purposely for the development various parameter settings for the models.

Determination of coefficient, root error mean square, error mean square, correlation coefficient is used to check the fitness, performance and adequacy of the models. The hybrid MLR-based intelligence models MLR-SVM, MLR-ANN and MLR-ANFIS are examined to cope with non-linearity in the input data. The results show that ANFIS outperformed the other three single models (MLR, SVM and ANN). The new hybrid MLR-based intelligence approach proposed shows that hybrid models outperformed the single models with hybrid model MLR-ANFIS as the most adequate model. Thus hybrid MLR-based intelligence can be employed as alternative high-performance models for QSPR simulation.

Keywords: Mycotoxins; fungi; HPLC; retention time; AI-based models; multilinear regression; hybrid techniques

ÖZET

Mikotoksinler, özellikle insanlar ve diğer hayvanlar için genel sağlık sorunu olan mantar grubudur. Mikotoksinlerin kantitatif yapı-özellik ilişkisi (QSPR) çalışmaları için YSA, MLR, SVM ve ANFIS gibi veri-akıllı odaklı yaklaşımın kullanımı oluşturulmuştur. Bu tezde, bu dört veriye dayalı model yaklaşımı, HPLC kromatografik tekniği ile elde edilen kantifikasyon ve belirleme verileri kullanılarak çeşitli mikotoksin sınıflarının alıkonma süresinin (tR) QSPR incelemesi için uygulanır. Bu, yeni bir hibrit MLR tabanlı istihbarat modeli geliştirme görüşüyle. QSPR modellemesi için dört girdi değişkeni - tutma indeksi, tepe simetri, mono izotopik kütle ve bağıl duyarlılık faktörü kullanılır. Daha sonra, modeller için çeşitli parametre ayarlarının geliştirilmesi amacıyla, bu girdi değişkenleri ile tutma süresi arasındaki ilişkileri ölçmek için Spearman Pearson korelasyonu kullanılır.

Modellerin uygunluğunu, performansını ve yeterliliğini kontrol etmek için katsayı, kök hata ortalama karesi, hata ortalama karesi, korelasyon katsayısının belirlenmesi kullanılır. Hibrit MLR tabanlı zeka modelleri MLR-SVM, MLR-ANN ve MLR-ANFIS, giriş verilerindeki doğrusal olmayanlıkla baş etmek için incelenir. Sonuçlar, ANFIS'in diğer üç modelden (MLR, SVM ve YSA) daha iyi performans gösterdiğini göstermektedir. Önerilen yeni hibrit MLR tabanlı zeka yaklaşımı, hibrit modellerin en uygun model olarak hibrit model MLR-ANFIS ile tekli modellerden daha iyi performans gösterdiğini göstermektedir. Bu nedenle hibrit MLR tabanlı zeka, QSPR simülasyonu için alternatif yüksek performanslı modeller olarak kullanılabilir.

Anahtar Kelimeler: Mikotoksinler; mantarlar; HPLC; saklama süresi; yapay zeka Tabanlı modeller; Çok çizgili regresyon; hibrit teknikler

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFT: Aflatoxin

AI: Artificial Intelligence

ANFIS: Adaptive Neuro-Fuzzy Inference Systems

ANN: Artificial Neural Network

BPNN: Back Propagation Neural Network

CNS: Central Nervous System

ELM: Extreme Learning Machine

FFNN: Feed Forward Neural Network

FUM: Fumonisins

GC: Gas Chromatography

HPTLC: High Performance Thin Layer Chromatography

HPLC: High Performance Liquid Chromatography

HW: Hammerstein Wiener

IARC: International Agency for Research on Cancer

LC-MS: Liquid Chromatography-Mass Spectrometry

MFs: Membership Functions

MLP: Multilayer Perceptron

MLR: Multiple Linear Regression

MS: Mass Spectrometry

MSE: Mean Square Error

m/z: Mass-to-charge ratio

PCA: Principal Component Analysis

PLS:	Partial Least Squares
Ppb:	Part Per Billion
QSPR:	Quantitative Structure-Property Relationship
R:	Correlation Co-efficient
R^2:	Determination Co-efficient
RMSE:	Root Mean Square Error
SFC:	Supercritical Fluid Chromatography
SVM:	Support Vector Machine
SWLR:	Step-wise Linear Regression
THZ:	Terahertz
tR:	Retention Time
UPLC:	**U**ltra-Performance Liquid Chromatography

CHAPTER 1
INTRODUCTION

1.1 Background of study

Generally, fungi are considered as one of the major disease causative agents of insects and plants (Khosrokhavar, Ghasemi, and Shiri, 2010). Even though, they are not as vital as agents that causes diseases in vertebrates (meaning the number of fungi that are of medical significance is relatively small as compared with other organism such as virus and bacteria). The growth of Fungi on animal hosts lead to the production of diseases and disorders. These are collectively called mycoses; while dermal, respiratory, and dietary and others exposure to harmful fundal metabolites and precursors lead to production of diseases collectively known as mycotoxicoses (Magan and Olsen, 2004).

These mycotoxicoses are generally analogous to pathologies that are caused through natural poisoning, produced through exposure to heavy metals scums and pesticides (El-Nezami et al., 1998). The mycotoxicosis symptoms rely on the kind of mycotoxin, its interaction with other harmful substances, dietary status, genetics, the exposed person's gender, health status, age, and the time taken by the exposure (Assunção, Silva and Alvito 2016). Therefore, the intensity of mycotoxin poisoning may be attributed to factors like deficiency of a certain vitamin, health status, abuse of chemical such as alcohol, drugs as well as caloric deprivation. Consecutively, mycotoxicoses may lead to microbial diseases, increase the synergistic effects of other harmful and toxic substances and increase the negative impact of malnutrition (Tsakiris et al., 2013).

Recent studies have described that various mycotoxins composed of some carcinogenic characteristics . Most of them are DNA-reactive while some are not. But when the endpoint is considered to be carcinogenic, in-vivo or in-vitro studies may be needed in order to construct a way of determining possible molecular modifications related with gene expression, which are related to tumor suppressor genes or proto-oncogenes as well as instability that has to do with the genomic make-up. In return, this will aid in knowing the mode of the cancer as well as in finding possible and suitable solution (Bull, Ward, and Goodfellow, 2000).

Mycotoxins may lead to developmental defects such as birth effects, affecting the human reproductive systems, the immune systems, can be neurotoxic that will affect various target organs as well as alter the hormonal activity in the body. Moreover, it may lead to certain damages to the gastrointestinal systems; lead to growth reduction, causes hematological defects and can lead to skin irritation (Foerster et al., 2020).

Generally, mycotoxins enters the body through ingestion of highly contaminated food substance, even though other routes such as inhalations of toxins, dermal contact with the mycotoxins especially directly are equally considered as vital routes for mycotoxins to be found in the body (Schlosser, Robert and Noyon, 2020). The Mycotoxins that exists in food materials are usually secondary metabolites of various filamentous fungi that have the ability of contaminating the food as well as the food crops via food chain. Even though, different classes of fungal toxins were known, but limited number of them are considered vital in playing role towards the area of food safety. Therefore, based on these a wide range of analytical tools has been used to designed methods for their elucidations and determination (Györgyi, 2010).

However, microfungi are regarded as rich sources of chemical diversity. Microfungi and actinomytes are the major sources of metabolites that can be used in the pharmaceutical application used in their raw of derivatized form. Only a little percentage of mycota has been known until today and fungi lead to the production of numerous unknown metabolites. Therefore, these fungi are considered as one of the promising microbiotic source of signing theoretical models that can be used in the simulation (for example retention behavior) of mycotoxins is of paramount importance, due to their toxicity to animals and humans. Since, the diversity of the chemical nature if high within micro-fungi, therefore different compounds can be found in the sample extract (Martins et al., 2020). For example, cyclic peptides, antraquinones, alkaloids, ketones, alcohols and small acids. In order to cope for this wide range of chemical behavior, a reversed phase liquid chromatography coupled with diode array detector is used to determine these structures (Cunha & Fernandes, 2018).

The method equally involves the use of water-acetonitrile gradient elution system using carbon-18 as the stationary phase. Besides, various reports in the literature have reported

the quantitative relationship between the retention behavior and molecular properties of mycotoxins (Nielsen & Smedsgaard, 2003). Generally, the computational approaches involve in predicting retention behavior can be categorized into two classes (Wang et al., 2014). One of the methods involves the application of mathematical equation to calculate the retention behavior using the molecular properties as the corresponding variables. The other method involve the use of advance Quantitative Structure-Property Relationship (QSPR) methods such as multiple linear regression (MLR), Artificial neural network (ANN), Adaptive neuro-fuzzy inference systems (ANFIS) and support vector machine (SVM) amongst others. These computational methods offer great advantage over the trivial mathematical equation method through simple implementation mechanism, dimensional independence, less tedious and more time efficient (Abba, Usman, and Isik, 2020).

1.2 Problem Statement

To understand the mechanism of various mycotoxins, their chemical behavior needs to be understood. These can be achieve using various analytical instruments such as high performance liquid chromatography (HPLC), ultra-performance liquid chromatography (UPLC), gas chromatography (GC), high performance thin layer chromatography (HPTLC) and electrophoresis. Due to the high costs of the instruments, reagents as well as failure during the laboratory experiment leading to high cost. This has led to the low number of studies reported in the technical literature indicating the determination of mycotoxins using these chromatographic methods. It can be observed that the area of determination and elucidation of performance of mycotoxins using chromatographic methods still needs relatively higher number of studies to fill the gap.

Consequently, most of the studies on mycotoxins using chromatographic techniques reported in the literature that focused on data-driven methods and machine learning such as artificial intelligence (AI) methods and hybrid data-intelligence approaches are still limited. Therefore, this study employs the application of AI-based models using various molecular properties of mycotoxins in simulating their retention behavior in HPLC technique during the first stage of the research. Based on the literature, despite the rising implementation of these AI-based models, still it is associated with some drawbacks and weaknesses, which has led to lower accuracy production due to over fitting especially for

ANN. Therefore, this work introduces a novel hybrid approach MLR-intelligence. This will enhance the performance of the model and cope with inadequacies of both the classical linear model MLR and the AI-based models (ANN, ANFIS and SVM) orchestrated by the linear and non-linear properties of the models.

1.3 Aim and specific objectives

The aim of the study is to determine an adequate new hybrid MLR-based intelligence model for the quantitative structure–property relationship (QSPR) studies of mycotoxins.
The specific objectives of this work are as follows:

- To comparatively study the single applications of multiple linear regressions (MLR), Support vector machine (SVM), Artificial neural network (ANN) and Adaptive neuro-fuzzy inference systems (ANFIS) for quantitative structure–property relationship (QSPR) studies of mycotoxins.
- To examine the hybrid MLR-based intelligence model for the quantitative structure–property relationship (QSPR) studies of mycotoxins.
- To establish an adequate hybrid MLR-based intelligence for QSPR studies of mycotoxin.

1.4 Significant of Study

According to our literature search, limited studies have been conducted using reversed phase HPLC for the elucidation of mycotoxins using experimental design. Therefore, the results obtained from this studies will be vital to the area of Food engineering and technology through determining the retention behaviour of these mycotoxins using multiple linear regression (MLR), Support vector machine (SVM), Artificial neural network (ANN) and Adaptive neuro-fuzzy inference systems (ANFIS). Moreover, the literature demonstrated that there is no much study conducted showing the applications of the hybrid-intelligence for modelling the qualitative properties of mycotoxins. To cope with the non-linearity in the descriptor input data used, which is one of the shortages of the single model approaches, hybrid MLR-based intelligence approach is proposed and examined. This is also gives credence to this work.

1.5 Scope of Study

The data of mycotoxins reported in this current study was taken from previous experimental data reported by Nielsen and Smedsgaard, 2003. More also, the modelling was conducted based on secondary data collected using HPLC method for quantification of mycotoxin in food samples.

1.6 Thesis Organization

- Chapter 1: It involves the explanation of the topic introduction, describing the objective and the problem statement of the study.
- Chapter 2: It involves the explanation and discussion about previous works conducted in the literature that are related to this work.
- Chapter 3: It is describing detailed procedures and methods involved in achieving the objectives of the research.
- Chapter 4: It involves detailed results and discussion.
- Chapter 5: Involves the conclusion, recommendations, and direction for future work.

CHAPTER 2
LITERATURE REVIEW

The basic idea of designing this chapter is to discuss on previous studies conducted in the literature, which are related to the application of data driven algorithms used in prediction of food-induced mycotoxins monitoring and chromatographic applications on mycotoxins. However, a comprehensive review of different task that are related with mycotoxins, High performance liquid chromatography (HPLC) and Gas chromatography (GC) will be given more emphasis in order to understand the mechanisms and approaches involved in the study, which will be discussed in this chapter. Chromatography is considered as a method used in separating mixture through distributing its components between the stationary and mobile phase. The elucidation of mycotoxins can be done using chromatographic methods more especially HPLC, GC and thin layer chromatography in order to understand the chemical behavior of the mycotoxins so as to provide an sight for its mechanism and give a preventive measures against its adverse effects.

2.1 Mycotoxins

The word Mycotoxin combines the Greek word ‘mykes’ which means fungus and a Latin word ‘toxicum’ which means poison. Mycotoxins are formed as secondary metabolites of various fungi that can easily colonise food produce in the farm of in stores after harvesting (Afsah-Hejri et al., 2020). Therefore, it results in potential threat to animals and human health via ingestion of the food product that is prepared from such commodities. The contamination can occur before or after the harvest. For example aflatoxins, ochratoxins, T-2 toxin and deoxynivalenol. Mostly, plant produce that are kept for a long period of time are more prone for mycotoxin growth and mould formation. Mycotoxins have the ability of occurring in both tropical and temperate regions worldwide, which depends on the type of the fungi species involved (Ayofemi, 2020).

The main food crops affected are beans, dried peas, oil seeds, spices, cocoa, coffee, dried fruit, nuts and cereals. These toxic compounds may equally be found in wine and beer, resulted from the usage of barley that was contaminated as well as grapes and other cereals involved in their synthesis (Foerster et al., 2020). They have the ability of entering human

food chain through meat as well as other animal products like cheese, milk and eggs due to the fact that the livestock have ingested a food contaminated with the mycotoxins. Sometimes, they are genotypically specific, even though they may be produced from one or numerous fungal species. For instance, Ochratoxins can be produced by various species like A. Ochraceus, generally in the tropical parts of the world, while *P. verrucosum* generally produced it during storage period in temperate regions and in special cases a single species may produce multiple mycotoxins (Colombo and Papetti, 2020). Usually mycotoxins are structurally and chemically diverse (see Figure 2.1). Since most of the secondary metabolites are produced through simple biosynthetic reaction involving smaller molecules such as pyruvates and acetates etc. This can be surprising, thus, this may results in the compounds to compose of wide range of toxic effects, both chronic and acute (Bauchet et al., 2020).

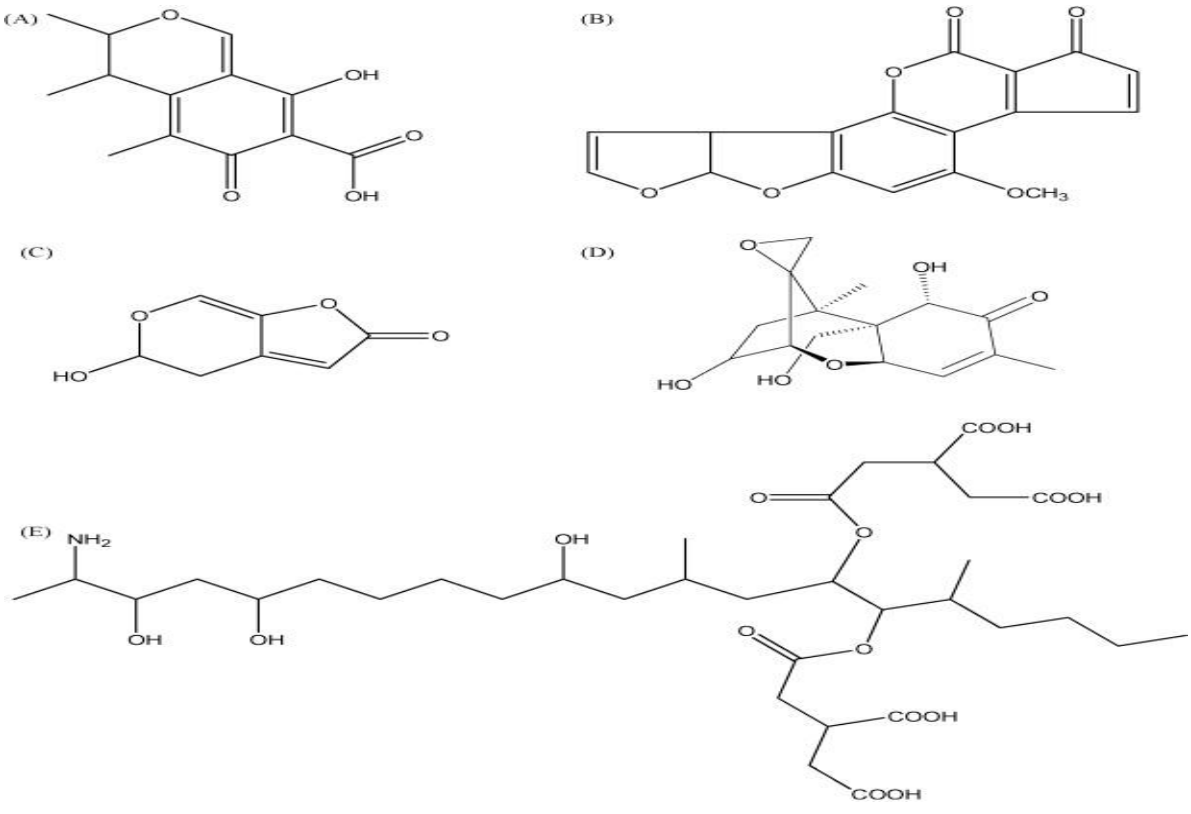


Figure 2.1: Examples of common mycotoxins demonstrating structural diversity of these compounds. (A) Citrinin, (B) Aflatoxin B1, (C) patulin, (D) deoxynivalenol, and (E) fumonsin B1.

Table 2.1: Classes of major mycotoxins and their physiological effects

Toxin	Major effects on mammalian systems
Aflatoxins	Impaired immune system, acute hepatitis, carcinogenic
Citrinin	Nephrotoxic
Citreoviridin	Neurotoxicity
Cyclopiazonic acid	Weight loss, depression hyperesthesia, kodua poisoning, kidney lesions, lactation loss, liver leisions, diarrhoea, carcinogenic
Fumonisins	Causing leukoencephalomalacia to certain animals, hepatotoxic, carcinogenic
Moniliformin	Keshan disease, intestitinal haemorrhage, carcinogenic weight loss
Ochratoxins	Teratogenic, hepatotoxic, nephrotoxic, carcinogenic
Patulin	Brain and lung haemorrhage
Sterigmatocystin	Kidney lesions, lactation loss, liver lesions, diarrhoea, carcinogenic
Trichothecenes	gastrointestinal haemorrhage and Immuno-depressants,
Zearalenone	Estrogenic activity

2.1.1. Aflatoxins

Aflatoxins are considered as widely related and closed group of mycotoxins, which are produced from fungi *A. parasiticus* and *A. flavus*. This group of mycotoxins were first

understood following the death of almost 100,000 Turks in England on a poultry farm, which are related to the ingestion of various Brazilian peanut meal. Aflatoxins are said to be regarded as difuranocoumarin derivatives, which are synthesized via polyketide mechanism through various strains of *A. parasiticus* and *A. flavus* (Bauchet et al., 2020). Even though *A. flavus* was reported and considered as a common contaminant to agricultural crops. Moreover, *A. pseudotamari, A. nomius, A. ochraceoroseus* and *A. bombycis* are equally reported as aflatoxin producing agents, which are generally encountered not frequently.

Based on mycological points of view, there are quantitative and qualitative differences in the toxigenic capabilities shown by various strains found in each aflatoxin specie (Abdelhaliem sand Al-Otaibi, 2020). For instance, about 50% of *A. flavus* are involved in the production of aflatoxins. The four major classes of aflatoxins are G1, G2, B1, and B2, which are generally classified based on their UV-properties as well as chromatographic behaviour (Roohi et al., 2020). B1 is one of the potent natural carcinogenic chemical compounds, which is known so far. The toxicology of various aflatoxins are both complex and challenging. Their variations in susceptibility between individuals lies greatly on the fraction and amount of dosage, which is directed towards different pathways in the human body, having toxic 'biological' penetration being the results of the activation of the epoxide with the DNA and proteins (Javanmardi et al., 2020).

2.1.2. Ochratoxins

This is equally produced by various *Penicillium* and *Aspergillus* genera that are natural opportunistic bio-deterioration agent. These species are widely spread, since the species are grown in different environmental and biological conditions such as temperature, moisture, pH and substrate (El-Shahir et al., 2020). Ochratoxins was discovered in the year 1965, which is considered as a fungal metabolite, which showed harmful behaviour to animals. The structure of these mycotoxins as well as its analogues are indicated in Figure 2.2 Mellein, OTAβ and OTAα are considered as dihydroisocoumarins which are generally synthesized by the same chemical group, which is equally linked to the biosynthesis of OTA (Abdelhaliem and Al-Otaibi, 2020).

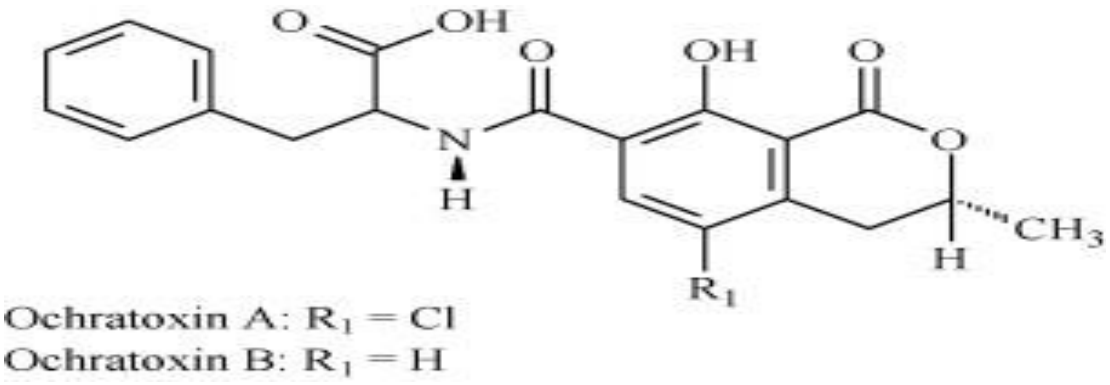


Figure 2.2: Structural representations of ochratoxin A and B. Ochratoxin A: R1 = Cl and ochratoxin B: R1 = H.

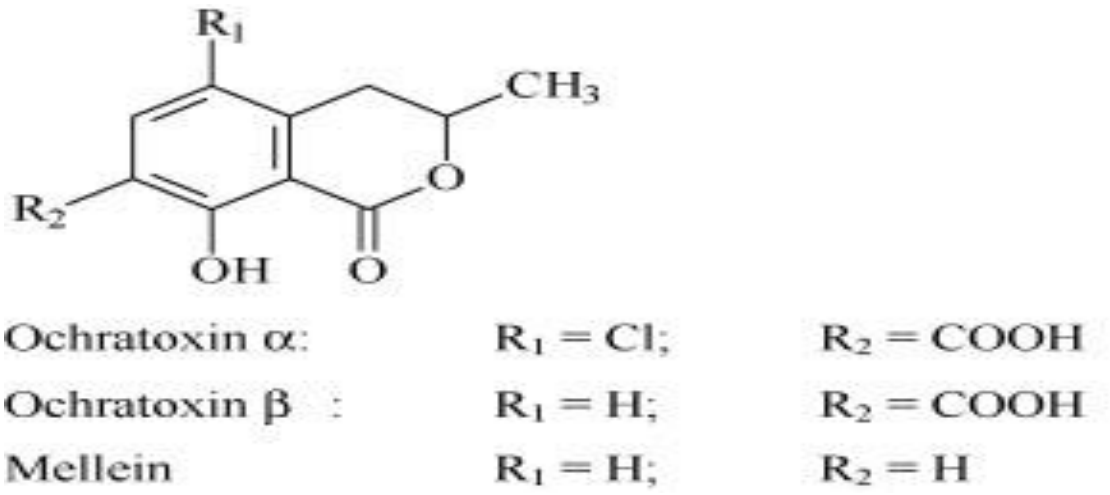


Figure 2.3: Structural representations of the dihydroisocoumarins, ochratoxin α, ochratoxin β and mellein. Ochratoxin α: R1 = Cl; R2 = COOH, ochratoxin β: R1 = H;R2 = COOH, and Mellein R1 = H; R2 = H.

Ochratoxins is generally found in cereals, other starch source foods as well as coffee, dried fruits and spices. Mostly the concentration of Ochratoxins is not up to some part per billion (ppb). The biological properties of this mycotoxin are well reported and documented in the technical literature (Veenaas, Linusson, and Haglund, 2018). Some of its effects such as carcinogenic effects, mutagenic, fertility inhibition, teratogenic as well as immune suppressive nature has been discussed in the technical literature and many reviews have indicated this subject matter (Dankovich, 2019).

The International Agency for Research on Cancer (IARC) has classified Ochratoxins as a potential carcinogenic agent. Its toxicity have shown to be related to its capability in inhibiting protein synthesis through competition with other chemical compounds such as phenylalanine (Yassein et al.,2020).

2.2 Determination of mycotoxins

Majority of mycotoxins are composed of chemically stable compounds that survive during processing and storage, some even during cooking at higher temperatures while others even up to baking or varieties of breakfast production cereals. Therefore, this makes it vital to avoid conditions, which might lead to the synthesis of mycotoxins. Moreover, the understanding of this conditions and processes is tedious, which is not easily achievable and even impossible in practice (Evtugyn and Hianik, 2019). Generally, mycotoxins are notorious and uneasy organisms to be removed as the best technique for its control is to through prevention. The present of the recognised properties of the toxin-producing Fungi, do not necessitate the presence of the mycotoxins, due to the fact that numerous factors play a role in the formation of the mycotoxins (El-Nezami et al., 1998).

Moreover, the absence of any mould is not a prove that a certain mycotoxin is not present as the mould might have already become extinct and still leave the toxin properties intact. Mostly, a fungus develops in an isolated place, which is unevenly distributed in various stored and processed commodities. Hence, it is vital to create a protocol in order to make sure that when taking a sample, it should maintain its own properties prior to analysis using various analytical tools (Alsharif et al., 2019). Indeed, almost 80% of errors related to mycotoxin analysis are related to how the sample collection was done. Taking an accurate sample is very difficult, since the toxins are unevenly distributed in feeds and grain. More also, these toxins are generally found in relatively lower concentrations needs a highly sensitive, reliable and specific methods for its elucidation. Sampling and analysis are of paramount importance, due to the fact that failure in achieving a reliable and satisfactory analysis may aggravate to consignments that are not acceptable (Fish, Fernandes and Ivanova, 2019).

Because of this complexity in the chemical structures of mycotoxins, it is impossible to apply a single unique method in detecting most the mycotoxins, due to the fact that each

mycotoxin need a different technique (Sobral et al., 2019). The working principle of some compounds may not be applicable to other compounds even if they have similar physical and chemical properties. However, depending on the chemical and physical properties of the compounds, different standard analytical methods have been developed using various analytical tools, which gives broad-based and flexible procedures that be applied in detecting various mycotoxins (Alsharif et al., 2019). It could be appropriate to have simple detection techniques that can be used even by non-scientific individuals that are inexpensive and fast.

The use of simple, cost effective and cheaper solutions for the elucidation of mycotoxins is becoming increasingly required, due to their importance, which considers their toxicity and regulating their amounts in food products (Carballo et al., 2019). A desired detection technique should have higher ability for flexibility, should be robust, and should have higher sensitivity and selectivity over a wider range of chemical species. The method should equally be reproducible at a higher level and the obtained results should be easy to comprehend and relevant to the study under practice. The technique should be portable and rapid especially when the analysis involves fieldwork. Many methods have been applied in the analysis of mycotoxins, whereby majority of them are lab-based (Dada et al.,2020)

2.3 Mycotoxin evaluation and quantification

Generally, different methods can be used in the determination, evaluation and quantification of mycotoxins to understand the behavior, structure and how much of the mycotoxin is present in a certain sample matric. Analytical methods such as spectroscopy (e.g Nuclear magnetic resonance, Raman spectroscopy and UV-spectroscopy), spectrometry method (e.g Mass spectrometry) and Chromatographic method (HPLC, GC-MS and GC-FID). Chromatographic method is a combination of analytical processes that involve the separation of mixture through distributing their components into both mobile and stationary phases (Xiong et al., 2020). Whereby, the mobile phase carries the components of the mixture via the medium being used while the stationary phase remain fixed throughout the analysis (Encarnação, 2020).

Generally, we have various modes of classifying chromatographic methods, whereby the popular one is based on the physical state of the mobile phase e.g Liquid chromatography

such as high performance liquid chromatography (HPLC), high performance thin layer chromatography (HPTLC) and ultra- performance liquid chromatography (UPLC), Gas chromatography such as GC-MS and GC-FID and Supercritical Fluid Chromatography (SFC) (Kotapati and Bates, 2020). Chromatographic analysis are employed due to its vital application in increasing the specificity and complexity of the chemical compounds in order to remove interferences, in order to pre-concentrate the target analyte due to lower sensitivity of the target analytical tool in order to enrich the target compound's concentration as well as when their this non-compatibility between the sample matrix and the targeted analytical tools (Abdelkhalek et al., 2020).

2.3.1 High performance liquid chromatography (HPLC)

The high performance liquid chromatography (HPLC) or sometimes-called high-pressure liquid chromatography (HPLC) is one the classes of liquid chromatography method that involves the use of a liquid mobile phase (Ostertag et al., 2020). The idea of HPLC development was designed based on the need for providing an efficient separation method that can be achieved by the aid of a well-packed column, which will subsequently minimized the analysis time (Sparkman, 2008). This can be achieved through the use of a pump, which will deliver the mobile phase (carrying the analyte) into the column, which can only be done using special tools, hence this lead to the development and creation of HPLC instrument (Adamson et al., 1999).

Liquid chromatography can be classified into four different classes; which are, size-exclusion chromatography, ion-exchange chromatography, adsorption chromatography and partition chromatography (Biswas et al., 2007). The partition chromatography is the common chromatographic method employed nowadays, which is equally further subdivided into normal phase and reversed phase chromatography based on the polarity of their mobile phases (Zapata, Rodríguez and Garrido, 2000).

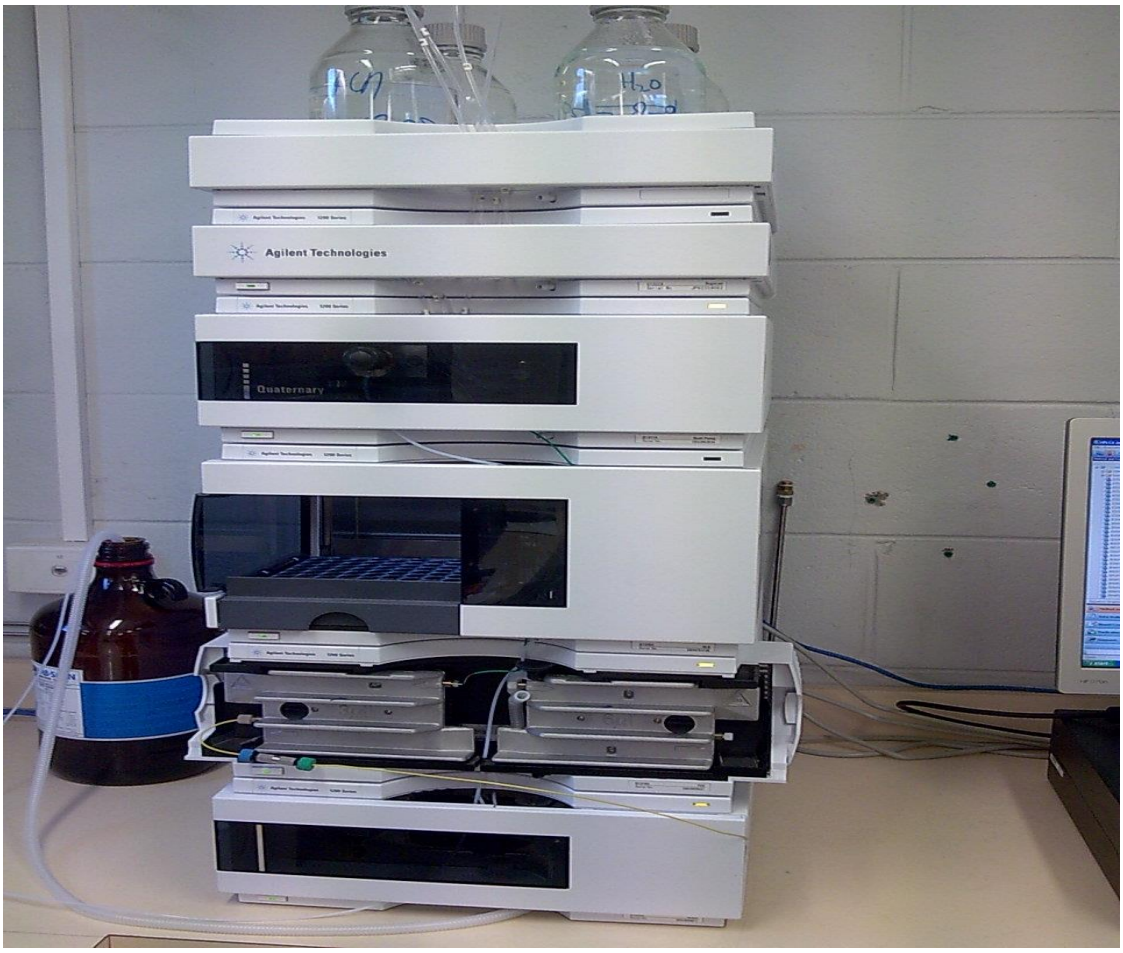


Figure 2.4: The High performance liquid Chromatography (HPLC) instrument

2.3.2 Mass Spectrometry (MS)

Mass spectrometry is considered as one of the most informative analytical instrument, owing to its wide range of application (Mingxun Wang et al., 2020). Its application composed of determination of isotopic elements, elucidation of quantitative and qualitative composition of complex mixtures, various elemental determination of samples and structural determinations of biological, organic and inorganic samples (Cui et al., 2020). MS can be classified into two main categories 1) The atomic MS, which can be used in determining elemental species present in the samples as well as their corresponding concentrations 2) The molecular MS instrument, which are employed for the quantitative and qualitative determination of various molecular compounds, which are present in the sample matrix. In this research the molecular MS was used, which will be further discussed in detail in chapter three (Cui et al., 2020).

1. Mass Spectrometer

This is the analytical tool used in MS analysis, which involves the separation and production of ions based on their mass-to-charge ratio (m/z). Mostly the ions are produced as single charges (either positive or negative charges) (Liu, Zhang and Gross, 2020). The working principle of molecular MS composed of the analyte bombardment using stream of electrons that leads to a loss of an electron, which results in the formation of molecular ion M+. The fragmentation is very vital used in determining various compounds (Aron et al., 2020).

Figure 2.5: The Mass Spectrometer instrument

2. Liquid Chromatography-Mass Spectrometry (LC-MS)

The coupling of MS and liquid chromatography is a strong hyphenation, which combined both the advantages of the separation efficiency of the liquid chromatography as well as the selectivity and sensitivity of the MS (Xuan et al., 2020). Even though one of the major drawbacks of this hyphenation is the need for gaseous sample for the MS, while the outputs of the liquid chromatography are liquid as the mobile phase and solute analyte, which is dissolved in a suitable solvent (Xuan et al., 2020). Therefore, a vaporized solvent is needed. The vapor that can be produce by liquid chromatography solvent is much higher than the carrier gasses in gas chromatography (GC) (Yoshikawa et al., 2020). Therefore,

excess of the solvent need to be removed using a suitable method. The major application of LC-MS method is that it gives the molecular fingerprint of a certain eluate without much relying on the retention time, which is contrary in the case of the classical HPLC instrument. This combination equally gives information on the molecular mass, exert quantitative result as well as structural information (Yoshikawa et al., 2020).

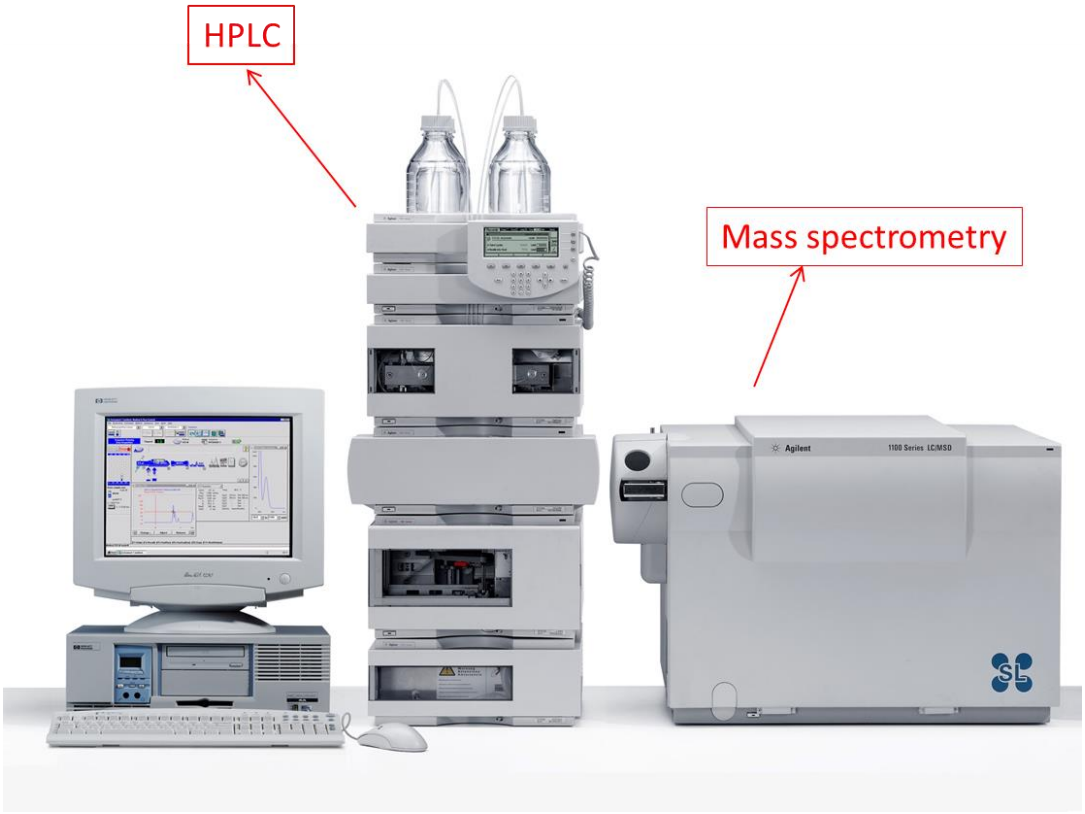


Figure 2.6; The liquid chromatography-Mass spectrometer (LC-MS) coupling

2.4 Sample pre-treatment methods

Majority of the methods involve in determining mycotoxin relies on accurate extraction technique as well as sample clean up. These pre-analysis stages are important steps for a successful and good result. Even though they are time consuming and sometimes expensive, but the delay is worth it as it influence the final choice for the detection technique to be used. These stages are important for an effective protocol, as they are generally considered as time-consuming step (which generally consume almost two-third of the total time), and this influence the final choice of the analytical tool to be used. The kind of extraction technique to be used in extracting a certain mycotoxin from the

biological matter solely depends on the structure of the toxin (Phwan et al., 2019). Polar analytes, like Fumonisins (FUM), needs polar solvents or extractants such as water and other organic solvents like methanol, ethanol and acetonitrile.

Non-polar metabolites, such as AFT solely can be extracted using hydrophobic organic solvents such as N-hexane, chloroform and certain alcohols, which can be through direct extraction or derivatization using another kind of solvent that can be used in partial cleaning up of the matrix in order to get rid of the excessive component of the biological material (Tessini et al., 2010). The matrix equally plays a vital role in choosing the extraction solvent, which is affected by various chemical compounds, from which the target analyte will be taken from prior to the analysis. The application of chlorinated solvents for extraction is dramatically minimized, as they are proven hazardous to the environment and to the human life. The sample clean-up is one of the most important steps used in developing an experimental method, due to the fact that the sample purity affects the selectivity as well as the sensitivity of the obtained result.

Trace amount of the target mycotoxin can be masked by other interfering chemical compounds, which can be found in the matrix as well as in other materials, solvents and chemical applied during the analysis. Glassware should equally be contamination free, like alkaline detergents has the ability of forming salts with other chemical materials, which can subsequently minimized the detection ability (Arroyo et al., 2012).

2.5 Method of mycotoxin prevention and control

One of the major prevention technique of mycotoxins, which is well accepted is to prevent crops from been contaminated from the toxins. In the farm, it composed of good agronomical practice, which will increase the growth as well as prevents its infection by the toxic fungi (Cvjetko et al., 2018). Such practice consists of the application of resistant varieties, crop rotation, shifting cultivation and drought stress prevention. Since mycotoxins do not grow in dried foods, effective drying of the food products as well as maintenance of the dry materials is one of the major measures that can be taken against the production of mycotoxins. In order to prevent post-harvest mycotoxin production, drying should be conducted immediately after harvest (Arroyo-Manzanares et al., 2012). The water content of the farm produce that is safe for storage should correspond to the water

activity, which is an efficient method employed worldwide in controlling of fungal spoilage as well as production of mycotoxins in food and food products (El-Nezami et al., 1998).

Even though, it is feasible to control the growth of mycotoxins in stored food products through controlling the atmosphere of via the applications of natural inhibitors and preservatives, these methods are usually very expensive as compared with the efficient drying, which are equally not feasible especially in the developing nations (Yassein et al., 2020). Another method that can be applied in preventing mycotoxin growth in farm produce after harvest is to reduce the moisture content of the foodstuffs down in order to create an unfavourable situation for fungal growth. Moreover, antifungal agents can also be applied in controlling the growth of mycotoxins and other fungi (Javanmardi et al., 2020).

2.6 Artificial neural networks (ANNs)

This non-linear model consists of information processing unit known as neurons incorporated with advance tools applied in simulation and modelling, which mimicked the human and biological neurons (Ma et al., 2014). Due to this fact, ANN acts just like the CNS, which has the ability of solving complicated and complex issues (Barmpalexis et al., 2018). This model work through combining various characteristics such as decision making, learning power, generalization ability and parallel processing. The structure of ANN generally composed of three layers, whereby each layer perform different and special tasks; The layer is the one responsible for processing the information, the input layer distributes the data set into the network, while the output layer processes each of the input vector (D'Archivio, Giannitto, and Maggi, 2013). The neurons are considered as the smallest units which aid in processing the networks (Veenaas et al., 2018).

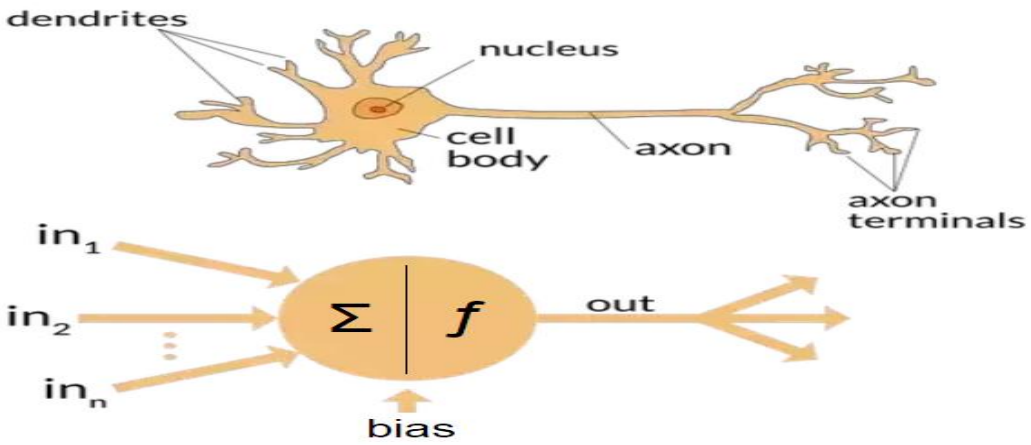


Figure 2.7: The structure of the three-layer artificial neural network

ANN have different classes such as Feed forward neural network (FFNN), Back propagation neural network (BPNN) and multi-layer perceptron. In this study the multilayer perceptron (MLP) was used, which is one of the commonest ANNs applied nowadays (Lotfi and Akbarzadeh-T., 2014). MLP is a vital technique that has the capacity of simulating the non-linear relation between dependent and independent variables (Lotfi and Akbarzadeh-T, 2013).One of the fundamental properties of MLP composed of the completing the message processing unit based on the interactive linkage denoted between neurons and without needing any advanced model design (Choubin, Khalighi-Sigaroodi, Malekian, and Kişi, 2016). Just like the normal ANN, MLP also composed of input layer, a minimum of one hidden layer as well as an output layer. Below is the structure of MLP (G. Elkiran et al., 2018).

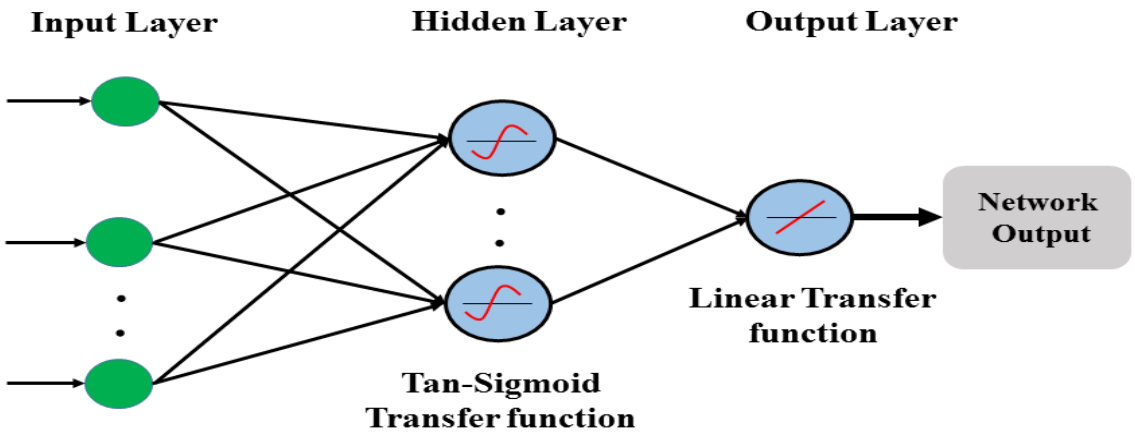


Figure 2.8: The architecture of Multilayer perceptron (MLP)

2.7 Adaptive-Neuro Fuzzy Inference System (ANFIS)

This is also another kind of non-linear data driven approach machine learning, which combines the learning ability of both fuzzy logic and neural networks (Ahmed, Mustakim, and Shah, 2017). It is equally considered as a real estimator due to the fact it is capable in approximating real function (Ahmed et al., 2017). In general, we have three kinds of ANFIS, which consists of Mamdani, Sugeno and Tsumoto. Modellers generally employers the used of Sugeno just as in the case of this current study. Defuzzifier and Fuzzier are main componets of the fuzzy logic systems (Tao et al., 2017). The fuzzy logic works through converting the input data into fuzzy values by using the membership functions (MFs).

The nodes here act as the membership function which converts the data into values from 0 to 1. Generally we have four classes of MFs, namely; trapezoidal, Gaussian, Sigmoid and triangular (Saini and Kumar, 2016). If we make an assumption that the FIS is composed of 2 inputs namely ‘x’ and ‘y’ having a single output ‘f’. The Sugeno fuzzy first order will then have these rules.

Rule 1: if $\mu(x)$is A_1 and $\mu(y)$ B_1then $f_1 = p_1x + q_1y + r_1$ (2.1)

Rule 2: if $\mu(x)$is A_2 and $\mu(y)$is B_2 then $f_2 = p_2x + q_2y + r_2$ (2.2)

A_1, B_1, A_2, B_2 parameters are membership functions for x and y inputs;

$p_1, q_1, r_1, p_2, q_2, r_2$ are outlet function parameters.

The structure and formulation of ANFIS follows a five-layer neural network arrangement. Refer to [9] for more information about ANFIS.

Generally, in modelling ANFIS different kind of MFs as well as epochs iterations were applied through trial by error approach in order to get the best architecture (Gaya et al., 2014).

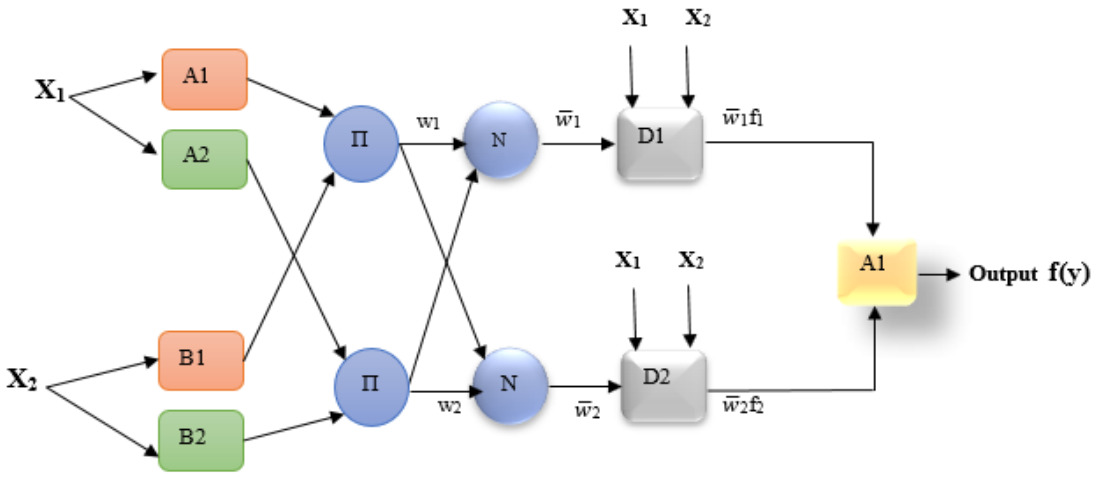


Figure 2.9 The architecture ANFIS model

2.8 Support Vector Machine (SVM)

The first idea of SVM was introduced and developed by in 1995, which gives a reliable approach towards pattern recognition, regression, classification and problem prediction (Vapnik, 1995). SVM follows the concept of machine learning and modelling that constitutes the use of data driven technique. Basically, SVM has numerous applications, whereby; the minimization of structural risk and statistical learning theory are the major applications of this approach (Haghiabi, Azamathulla, and Parsaie, 2017). Therefore this makes it differs from all other machine learning approaches such as ANFIS and ANN due to its capability in minimizing errors, complexity as well the ability in increasing the performance generalization ability of the network (Sharghi etal., 2018). SVM is generally classified into two classes; the linear as well as the non-linear SVM. Moreover, the

regression SVM is also referred as SVR (Gozen Elkiran, Nourani, and Abba, 2019). In SVR the weights are fitted into the data, whereby subsequently the outputs then move to the non-linear kernel, so as to check the no-linear behavior of the data sets. The general equation of SVM is demonstrated in equation 8 below (Su et al., 2019).

$$y = f(x) = w\phi(x_i) + b \tag{2.3}$$

where $\varphi(x_i)$ indicates feature spaces, nonlinearly mapped from input vector x.

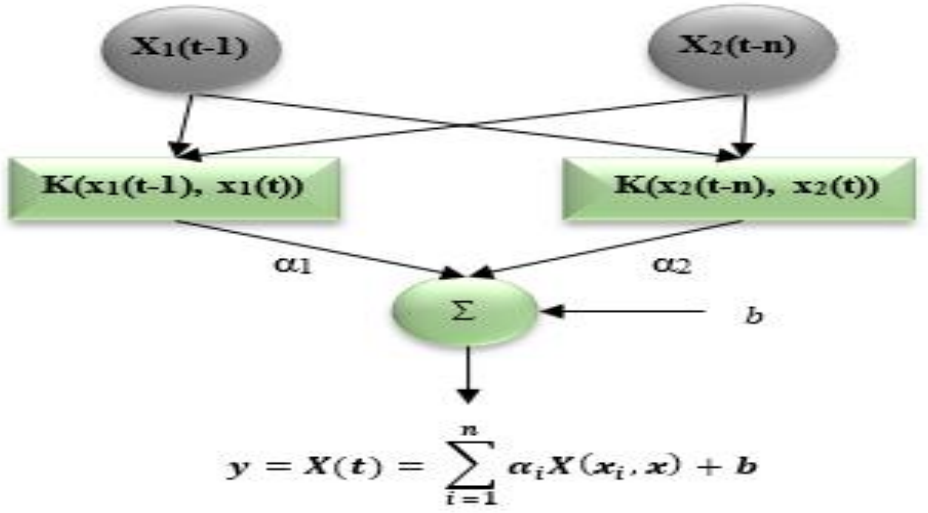


Figure 2.10: The Structure of Support Vector Machine Algorithms

2.9 Multi-linear regression (MLR)

In general, the regression models are used in predicting the correlation that exists between the dependent and the independent parameters. The mechanism of linear regression is based on the least square method (Lee et al., 2017). The linear regression is classified into 2 main groups; the simple and multiple linear regressions. Simple linear regression (SLR) involved the modelling of a single dependent variable using a single independent variable (Karimi et al., 2013). Whereby, a multiple linear regression (MLR) model is the one that involves the use of two or more independent variables toward the prediction and modelling of a single dependent variable (Khademi and Behfarnia, 2016).

MLR is the commonest applied linear model used nowadays in various fields of study such as engineering, science and social science (Kazemi et al., 2016). It is very important to note that MLR displays its correlation based on a straight line (linear) pattern that can be used in estimating the entire data points of both the dependent and independent variables (D'Archivio et al., 2013). MLR can be best described using equation 2.4 below

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots b_i x_i \quad (2.4)$$

Where by x_1,is the value of the i^{th} output, b_0 is the regression constant, and $b_i$is the coefficient of the i^{th} output.

2.10 Hybrid techniques development

The performance accuracy of the simple models is mostly related to various conditions such as the time scale, input determination and configuration of the model (Ghaedi, Hosaininia, Ghaedi, Vafaei, and Taghizadeh, 2014). These parameters were sum-up, which will subsequently have impact on the chromatographic conditions (Marrero-Ponce et al., 2018). However, some of the problems of these single models may be solved through the application of novel techniques that has the capability of measuring both the non-linear and linear co-relations that exists between the independent and dependent variables (Park et al., 2017). As discussed previously, the non-linear models (such as SVM, ANFIS and ANN) have shown good potentials in both regression and classification (Chandwani et al.,2015).

According to Marrero-Ponce et al. 2018, both linear and non-linear models have demonstrated success in classification and regression, especially when we focused more on enjoying the advantages that will exists if we coupled the two data driven approaches. If we consider the "no free lunch" theorem", we will understand that, there is no single existing model so far that can be used in all dataset (Yaseen et al., 2019). The properties of data such as size, normality and linearity influence so much on the performance of various learning algorithms (Yaseen et al., 2018). Moreover, different works have illustrated that

even when the same data set is used in modelling the performance of various properties, the performance of various models differs (Pham et al.,2019).

However, hybrid intelligence approaches have been substantiated to be effective in various forms of problems (Kazienko et al., 2013). Thereby, in this work, four different simple models namely; SVM, ANFIS, ANN and MLR are applied to determine the retention behaviour of the mycotoxins. Afterwards, the hybrid approach was proposed by coupling the linear and the non-linear models (MLR- (SVM, ANFIS and ANN)) in order to enjoy the prime features and properties of both the non-linear models as well as the linear MLR model for the data pattern prediction. Intrinsically, various researches in the literature have proved that it is highly needed to couple the simple models so as to boost and increase the overall performance of the modelling (Ghorbani et al., 2018). The hybrid technique involved two steps; the first stage involved the training of the MLR model in order to have the best linear fitted values. Subsequently, since MLR is a single model that can capture only the linear behaviour of the dataset, therefore the residuals of MLR result will be taken, which contains the information regarding the non-linearity properties of the datasets (Solgi, Pourhaghi, Bahmani, and Zarei, 2017).

This message is equally merge with that of the fitted values of the non-linear models, whereby this step is considered as the second step of the modelling process (Nourani et al., 2019).

$$f(y_t) = q_t + r_t \tag{2.5}$$

where q_t represents the linear phase and r_t represents the non-linear phase. These two phases have to be estimated from the data. Let ϵ denote the residual at time t from the linear model, then:

$$\epsilon_t = y_t - \hat{q}_t \tag{2.6}$$

Where $\hat{q}_t$ is the forecast value for time t from the estimated correlation, by modelling residuals using AI-based models, nonlinear relationships can be discovered. With n input nodes, the overall AI model for the residuals will be:

$$\epsilon_t = f(\epsilon_{t-1}, \epsilon_{t-2}, \ldots., \epsilon_{t-n}) + \varepsilon_t \quad (2.7)$$

where f is a nonlinear function determined by the AI models (FFNN and SVM) and ε_t is the random error.

2.11 Chromatographic application in Mycotoxins determination and Quantification

Various analytical methods hasve been applied in the determination and analysis of mycotoxins and other food samples, these analysis involves; the quantification, identification and classification of the mycotoxins, which can be done using different analytical methods such as spectroscopy (e.g Raman spectroscopy, UV-spectroscopy and atomic absorption spectroscopy), spectrometry (e.g mass spectrometry and atomic emission spectrometry) and chromatographic methods (e.g thin layer chromatography and high performance liquid chromatography). The need for the qualitative as well as quantitative evaluation and elucidation of mycotoxins is vital.

24 samples were investigated, which includes 14 functional foods as well as 10 spices, which were obtained from Chinese markets. The samples mould profile was subsequently examined using HPLC-FLC technique. The obtained results showed that a considerable amount of various classes of mycotoxins and fungi were identified including penicilium, Aspergilllus, Ochratoxin and aflatoxin B1(Kong et al., 2014).

It is equally reported that simultaneous evaluation of different aflatoxin G1,G2, B1, B2 T-2 toxin, Zearalanone and Ochratoxin A using the HPLC hyphenated with tandem MS using a locally made column. The method proposed meets all the necessary requirements considered for quick sample preparation as well as higher sensitivity for multiple detection (Zhang et al., 2016). Various classes of Ochratoxin and aflatoxin were examined from tea samples using HPLC method development with fluorescence detector (FD). The samples treated and the fungi were cultured, which were subsequently identified. The authors

recommended that due to the higher contamination of tea samples from mycotoxins, the need for regular and frequent evaluation in processing the tea can be used to improve the quality of the tea (Pakshir et al., 2020).

The analysis of various beer brewing samples for the determination of Ochratoxin and aflatoxin were also reported. Ochratoxin and aflatoxin are mycotoxins, synthesized from opaque beer sorghum malt, maize and fungi samples. The analysis was conducted using HPLC technique. The result proved the need for the evaluation of mycotoxins in beer samples (Marume et al., 2020).

2.12 Previous applications of MLR, ANN, SVM and ANFIS models

Chromatographic as well as other instrumental methods of analysis have proved to be effective in the evaluation of mycotoxins. The fact that, most of these methods are expensive and time consuming can hinder the rapidness of these fungi elucidation. Therefore, based on these problems various researchers have adopted the application of various data-driven approaches, for modelling the properties of mycotoxins using different input and output variables.

The application of MLR and SVM methods were employed for the prediction of retention behaviour of different mycotoxins using data obtained through the liquid chromatography with Uv and MS detector. The presented results showed strong agreement between the experimental and the simulated values using the R^2 and Q^2 (Khosrokhavar et al., 2010). Moreover, the use of image processing, SVM and discriminant analysis (DA) algorithms in order to classify and discriminate various wheat grains, which are being infected with a form of mycotoxin called *F. graminearum* was considered. The obtained result indicated that classification using the linear SVM technique gives higher performance as compared with its non-linear counterpart (Abbaspour-Gilandeh, Ghadakchi-Bazaz and Davari, 2020).

The comparative performance study of both MLR and ANN models for the prediction of the overall quality of cheese was reported. The presented result showed the robustness of ANN over MLR through the higher R^2 and lower RMSE values (Stangierski, Weiss, and Kaczmarek, 2019). In one of the studies, they investigated 160 samples in order to detect the presence of aflatoxin B1 using terahertz (THZ) spectral method with different range of concentration. Furthermore, various data driven algorithms such as SVM, principal

component regression, partial least squares (PLS) and PCA-SVM technique were employed. The results revealed the reliability of the non-linear SVM and PCA-SVM models. Higher performance accuracy were observes as opposed the other two techniques (Ge et al., 2016).

Another study, applies the hybrid models in order to boost the accuracy performance of the classical MLR model for the simulation of Chlorophyll. The chlorophyll was predicted using various physical and biological properties in order to determine the mechanism of ocean water system. The use of a hybrid MLR-ANN was observed to enhance the performance accuracy of the MLR result obtained. The results indicated that the proposed techniques have the ability of improving the performance by reducing the errors generated as well as increasing the correlation of the variables. Therefore, this proves that the hybrid technique is effective, accurate, efficient and with higher performance efficiency as compared with the simple models (Lola et al., 2016).

CHAPTER 3
METHODOLOGY

3.1 Model conceptualization

In this work, we employ the application of data driven approach in two scenarios. First, the prediction of the mycotoxins using HPLC technique by applying single models inform of SVM, ANFIS, ANN and MLR, for modelling their retention time as the dependent variables. Second, the application of the hybrid models in order to capture both the non-linear and linear features of the data using MLR-ANN, MLR-ANFIS and MLR-SVM.

The retention factor as the dependent variable is predicted using retention index, mono isotopic mass, and relative sensitivity factor and peak symmetry as the independent variable for each of the corresponding mycotoxin.

3.2 Phase 1: Data acquisition

A set of 150 data points from seven different mycotoxins group is employed in modelling the retention time of the corresponding mycotoxins. The complete data set used in this research is given in the Appendix 1. More so, 100 (67%) of the data set are considered as the training data; 50 (33%) of the data set are employed in order to test the models. For modelling SVM, ANFIS, ANN and MLR, 150 experimental data are taken from historical data. Five different model parameter settings are developed as discussed above (M1-M5). One of the major reasons for checking the data is to ensure the fitness of the data as well as to avoid the over-fitting and under fitting of the calibration data set. The data are therefore, validated in order to control and check potential modelling issues such as local minima (Gozen, Nourani, and Abba, 2019).

Table: 3.1: The historical data used for simulation of mycotoxins in the study

Class	Metabolites	Retention Index	Peak Symmetry	Mono Isotopic Mass	Relative Sensitivity Factor	Retention Time (tR) (Min)
	Aflatoxicol I	880	1.1	314.08	0.1	12.45
	Aflatoxin B1	859	1.1	312.06	0.2	11.50
	Aflatoxin B2	834	1.1	314.08	0.2	10.33
	Aflatoxin B2α	753	1.4	330.074	0	6.6
	Aflatoxin G1	830	1.1	328.06	0.2	10.16
	Aflatoxin G2	804	1.1	330.07	0.2	8.97
Aflatoxins and their Precursors	Aflatoxin G2α	718	1.4	346.07	0	5
	Aflatoxin M1	781	1.3	328.06	0.3	7.21
	Austocystin A	1140	1.3	372.04	0.2	21.57
	Averufin	1289	1.5	368.09	0.07	25.65
	5-Methoxysterigmatocystin	1028	0.9	354.07	0.2	18.02
	Dihydroxysterigmatocystin	1018	0.9	326.08	0.1	17.7
	Methoxysterigmatocystin	944	1.1	338.08	0.05	15.03
	Sterigmatocystin	1055	0.9	324.06	0.7	18.91
	Norsolorinic acid	1514	2.3	370.105	0.0002	31.08
	parasiticol	842	1.3	302.08	0.1	10.73
	Nivalenol	638	0.9	312.12	0.002	1.27
	Fusarenone X	661	1.5	354.13	0.005	2.35
	Deoxynivalenol	644	1.2	296.13	0.003	1.54
	3-Acetyldeoxynivalenol	723	1.3	338.14	0.011	5.21
	15-O-Acetyl-4-deoxynivalenol	721	1.4	338.14	0.02	5.1
	Scirpentriol	650	1.2	282.1467	0.03	1.82
	15-Acetoxyscirpenol	770	1.4	324.16	0.01	7.4
	Diacetoxyscirpenol	854	1.6	366.17	0.015	11.28
	3α-Acetyldiacetoxyscirpenol	958	1.5	408.18	0.025	15.56
	Neosolaniol	679	1.4	382.2	0.02	3.19
Trichothecenes	T-2 Triol	841	1.4	382.199	0.005	10.66
	HT-2 Toxin	908	1.1	424.2097	0.01	13.69
	T-2 Toxin	999	1.4	466.2203	0.001	17.06
	Iso-T-2 toxin	1015	1.8	466.2203	0.02	17.61
	Acetyl-T- 2toxin	1125	1.6	508.23	0.04	21.12
	Trichodermin	974	1.3	292.17	0.04	16.13
	Trichodermol	820	1	250.16	0.005	9.69
	7-α-Hydroxytrichodermol	666	1.4	266.16	0.03	2.59
	Verrucarol	673	1.5	266.16	0.003	2.89
	4,15 Diacetylverrucarol	920	1.2	350.17	0.015	14.15
	Trichothecin	978	1.5	332.1624	0.001	16.29
	Trichothecolone	689	1.6	264.1362	0.005	3.63
	Isosatratoxin F	960	1.4	542.23	0.2	15.63
	Roridin A	978	1.4	532.27	0.003	16.29
	RoridinE	1107	1.7	514.27	0.3	20.6
	Roridin H	1238	1.8	512.24	0.07	24.33
	Roridin L-2	900	1.4	530.25	0.05	13.37
Macrocyclic	Satratoxin G	901	1.5	544.2	0.08	13.43

Trichothecens and Precursors	Satratoxin H	920	1.4	528.24	0.02	14.14
	Trichoverrin A	922	1.3	532.27	0.2	14.2
	Trichoverin B	919	1.3	532.26	0.1	14.1
	Trichoverrol A	830	1.4	420.22	0.1	10.16
	Trichoverrol B	823	1.4	420.22	0.2	9.85
	Verrucarin A	981	1.4	502.22	0.005	16.4

Source: (Nielsen and Smedsgaard 2003)

3.3 Phase 2: Data Normalization

The data set employed in this work is normalized into a range of 0-1 based on equation 3.1. One of the main applications of data normalization prior to AI modelling is to minimize the redundancy of the data and to reduce the larger numerical errors (Pfeifer et al., 2007).

$$y = \left(\left(\frac{x - x_{min}}{x_{max} - x_{min}}\right)\right) \quad (3.1)$$

3.4 Phase 3: Correlation and statistical analysis

The Spearman correlation analysis is conducted using the excel 2016. The correlation analysis was done with 95% confidence limit. The statistical analysis is equally done on excel 2016. The data is then drag into the data analysis to determine the descriptive statistics.

3.5: Phase 4: Model development using correlation results

Five different models are developed based on the correlation analysis results. These are as follows;

M1: Mono isototopic mass, relative sensitivity factor

M2: Mono isotopic mass, peak symmetry, relative sensitivity factor

M3: Retention index, peak symmetry, relative sensitivity factor, Mono isotopic mass

M4: Retention index, peak symmetry, relative sensitivity factor

M5: Retention index, peak symmetry, Mono isotopic mass

3.6 Phase 5: Model development and Simulation

3.6.1 Simulation using single models

This work is a data driven approach, which presented different computational algorithms such as; SVM, ANFIS, ANN and MLR for the modelling of the qualitative properties of various mycotoxins in form of their retention behaviour using the HPLC technique.

Simulation 1: MLR

This method is conducted using the MS Excel 2016 using the normalized data. The regression is done using single output variable retention time and input variables parameters as presented in Phase 4. The MLR model follows, the least square method, which follows the following equation as discussed in chapter 2.

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots b_i x_i \qquad (2.4)$$

Simulation 2: ANN

This simulation technique was conducted using MATLAB 9.3 (R2019a). The ANN topology adopted consists of three layers viz; input layer, out layer and hidden layer are used during the analysis. The numbers of neurons at the hidden layer for the settings are: for M1, 3 neurones were used; M2 involved the use of 4 neurons; M3 5 neurons while in M4 and M5 we use 4 neurons. The Tan-sigmoid transfer function was used also during the simulation. The Multilayer perceptron (MLP) is equally used during the prediction.

Simulation 3: ANFIS

This method was also done in MATLAB 9.3 (R2019a). The first stage involved the importation of the normalized data into the MATLAB. The data is then classified into train and testing prior to the modelling. The sub-clustering Membership function (MP) is used throughout the analysis by changing the number of the input variables. The grid-partitioning method was used for optimizing the technique as the training FIS. 0.0005 error tolerance and 60 epochs were used throughout the analysis.

Simulation 4: SVM

During SVM simulation, we used the MATLAB 9.3 (R2019a) software. The normalized data are imported into the working space and then transferred to the regression learner. The cubic SVM is employed; data are trained and tested in order to get the result of the analysis using various parameter settings.

3.6.2 Hybrid MLR-based intelligence models

The hybrid data-driven approach involves the integration of both the linear (MLR) and non-linear models (SVM, ANFIS and ANN) models. In this work, the best performing model of the single models inform of M4 was employed to evaluate the performance the hybrid MLR-based intelligence over single models.

Hybrid I: MLR-ANN

Firstly, the residuals of MLR (difference between the values of the MLR model and the experimental values) are taken. Secondly, the residuals are subsequently added to the best performing result of the ANN model. The result is considered as the hybrid of the two models, which aid in capturing both the linear and non-linear properties of the data as well as can boosting the performance of the single models.

Hybrid II: MLR-ANFIS

First, the residuals of MLR (difference between the values of the MLR model and the experimental values) are taken. Secondly, the residuals are further added to the best performing result of the ANFIS (M4) model. The result is considered as the hybrid of the two models (MLR-ANFIS), which help in capturing both the linear and non-linear properties of the data.

Hybrid III: MLR-SVM

First, the residuals of MLR (difference between the values of the MLR model and the experimental values) were taken. Secondly, the residuals were further added to the best performing result of the SVM (M4) model. The result is considered as the hybrid of the two models (MLR-SVM), which help in capturing both the linear and non-linear properties of the data as well, can boost the performance of the single models.

3.7 Phase 6: Performance metrics

In any data driven approach, the performance criteria is checked using different indices through comparing the experimental result with the predicted results. The data driven algorithms (SVM, MLR, ANFIS and ANN) developed in both the training and testing phases are checked for their prediction ability using four performance indicator metrics. These are as follows; mean square error (MSE), root mean square error (RMSE), correlation co-efficient (R) and determination co-efficient (R^2) given respectively as Equations below.

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(Y_{obsi} - Y_{comi})^2 \tag{3.2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Y_{obsi}-Y_{comi})^2}{N}} \tag{3.3}$$

$$R = \frac{\sum_{i=1}^{N}(Y_{obs}-\bar{Y}_{obs})(Y_{com}-\bar{Y}_{com})}{\sqrt{\sum_{i=1}^{N}(Y_{obs}-\bar{Y}_{obs})^2 \sum_{i=1}^{N}(Y_{com}-\bar{Y}_{com})^2}} \tag{3.4}$$

$$R^2 = 1 - \frac{\sum_{j=1}^{N}\left[(Y)_{obs,j}-(Y)_{com,j}\right]^2}{\sum_{j=1}^{N}\left[(Y)_{obs,j}-\overline{(Y)}_{obs,j}\right]^2} \tag{3.5}$$

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Statistical and correlation analysis results

Based on the processing and pre-analysis test, Table 4.1 describes an accurate investigation of the input variables together with the output parameter for modelling the chemical and physical behaviour of various mycotoxins. Based on Elkiran et al., 2019 higher skewness has the ability of influencing extensively on the ANN model, whereby lower skewness lowers the prediction ability of the model (Elkiran et al., 2019). From Table 4.1 it can be observed that the retention index as the dominant input variable has the highest skewness value and hence has the highest correlation as seen in table 4.2. Moreover, all the other input variables showed higher and good skewness values. Whereby, this affects the correlation analysis of both the input and output parameters, which subsequently influence the modelling as shown in Table 4.2.

Table 4.1: The statistical analysis of the experimental variables

Variables	Retention index	Peak symmetry [a]	Mono isotopic Mass	Relative sensitivity factor [b]	Retention time (tR) (min)
Mean	1066.8	2.7	432.2	0.3	16.3
Median	981.0	1.4	416.7	0.1	16.5
Mode	1055.0	1.4	481.3	0.2	6.6
Standard Deviation	769.7	3.1	189.8	0.6	7.6
Kurtosis	124.3	6.3	29.2	21.9	0.1
Skewness	10.7	2.6	4.1	4.4	0.2
Range	9371.0	17.0	1963.1	4.0	37.4
Minimum	638.0	0.0	0.0	0.0	1.3
Maximum	10009.0	17.0	1963.1	4.0	38.6

Further evaluation was done using the correlation analysis method in order to ascertain the bearing and quantity relationship among the variables, which aid in the developments of five different classes (inform of models M1-M5). To improve the methodologies of the

models, assurance of the correlation co-efficient conducted are presented in Table 4.2 Where the guiding symbols (+ or -) shows the connection among the variables.

Table 4.2: The Correlation analysis matrix

Variables	Retention index	Peak symmetry	Mono isotopic Mass	Relative sensitivity factor	Retention time (tR) (min.)
Retention index	1.0000				
Peak symmetry	-0.0947	1.0000			
Mono isotopic Mass	0.4669	-0.1329	1.0000		
Relative sensitivity factor	0.1029	0.3690	0.0715	1.0000	
Retention time (tR)	0.7810	0.6105	-0.1454	0.1336	1.0000

Based on the correlation results it can be seen that retention index and peak symmetry has the highest correlation with the target output variable (retention time) with R-values 0.7810 and 0.6105 respectively. The correlation analysis helps us in determining the highest and lowest parameters that are related with one another. Moreover, it aid in knowing the science and the mechanism of the data prior to the dwelling into the modelling method by showing the parameter with highest correlation with the output variable. This can equally help during experimental analysis especially during the optimization method. Hence, based on the correlation analysis various models were formed to practically observe the influence of the method towards the simulation process. These models are shown in Table 4.3.

Table 4.3: Input variable parameter settings

Model Type	Model input combination
M1	Mono isotopic mass + relative sensitivity factor
M2	Mono isotopic mass + peak symmetry + relative sensitivity factor
M3	Retention index + peak symmetry + relative sensitivity factor + Mono isotopic mass
M4	Retention index + peak symmetry + relative sensitivity factor
M5	Retention index + peak symmetry + Mono isotopic mass

4.2: Results of the single models (SVM, ANN, ANFIS and MLR)

The performance ability of the single models of SVM, ANN, ANFIS and MLR is presented in Table 4.4. According to the comparative study of the techniques, it can be observed that the M3, M4 and M5) modeled the performance of the different classes of mycotoxins with the minimum accepted performance results. According to Nourani et al. 2019, the minimum determination co-efficient (R^2) required for any data driven approach to be acceptable should be 80%. of In another word, the model should have a minimum of 0.8 R^2-value for it to be acceptable (Nourani et al., 2019). In line with the aforesaid, M2 of ANFIS and ANN are considered acceptable results compared with the SVM and MLR models. This can be attributed to the ability of the models in capturing highly non-linearity of a data.

Table 4.4: The single models results

	Training				Testing			
Models	**R^2**	**RMSE**	**MSE**	**CC**	**R^2**	**RMSE**	**MSE**	**CC**
MLR-M1	0.4126183	0.126091	0.015899	0.642354	0.459973	0.11278	0.012719	0.678213
MLR-M2	0.5245308	0.113445	0.01287	0.724245	0.607718	0.096122	0.00924	0.779563
MLR-M3	0.8137542	0.100138	0.010028	0.902083	0.815155	0.065983	0.004354	0.902859
MLR-M4	**0.8727467**	**0.058689**	**0.003444**	**0.934209**	**0.876753**	**0.053878**	**0.002903**	**0.936351**
MLR-M5	0.8460088	0.064561	0.004168	0.919787	0.812934	0.066378	0.004406	0.901628
ANN-M1	0.7854856	0.076199	0.005806	0.886276	0.559203	0.101893	0.010382	0.747799
ANN-M2	0.8462503	0.064511	0.004162	0.919919	0.725962	0.08034	0.006454	0.852034
ANN-M3	0.9342788	0.042177	0.001779	0.966581	0.909113	0.046268	0.002141	0.953474
ANN-M4	**0.9750963**	**0.025963**	**0.000674**	**0.98747**	**0.978962**	**0.02226**	**0.000496**	**0.989425**
ANN-M5	0.9648481	0.030846	0.000951	0.982267	0.926096	0.041722	0.001741	0.962339
SVM-M1	0.5722622	0.1076	0.011578	0.75648	0.562944	0.10146	0.010294	0.750296
SVM-M2	0.6378167	0.099012	0.009803	0.798634	0.6792	0.12293	0.015112	0.824136
SVM-M3	0.9332619	0.042502	0.001806	0.966055	0.910133	0.046007	0.002117	0.954009
SVM-M4	**0.9463757**	**0.038098**	**0.001451**	**0.972818**	**0.928889**	**0.040926**	**0.001675**	**0.963789**
SVM-M5	0.9267449	0.044529	0.001983	0.962676	0.929217	0.040831	0.001667	0.963959
ANFIS-M1	0.667417	0.09488	0.009002	0.816956	0.67565	0.087404	0.00764	0.821979
ANFIS-M2	0.8230232	0.069212	0.00479	0.907206	0.822984	0.06457	0.004169	0.907185
ANFIS-M3	0.9878777	0.018114	0.000328	0.99392	0.958831	0.03114	0.00097	0.979199
ANFIS-M4	**0.9893603**	**0.01697**	**0.000288**	**0.994666**	**0.995069**	**0.010777**	**0.000116**	**0.997531**
ANFIS-M5	0.9605229	0.032689	0.001069	0.980063	0.930701	0.040401	0.001632	0.964728

Furthermore, the results showed that M1 of all the single models could not attain the minimum acceptable determination co-efficient (R^2) for the modelling of various classes of the mycotoxins using the chromatographic method. Further comparison of the results proved that the M4 of all the models demonstrated the highest performance in the modelling process. This might be attributed to the class combination as seen in the correlation analysis Table 4.2. Therefore, the result of the single models further corroborates the observations in Table 4.2.

Further comparative analysis of the single best models (M4) indicated that ANFIS-M4 showed higher superiority over the other single models SVM-M4, ANN-M4 and MLR-M4 in both the calibration and verification stages using the evaluation performance criteria R^2, CC, MSE and RMSE. Moreover, the performance efficiency of the single models based on the R^2-values demonstrates that ANFIS-M4 outperformed SVM-M4, ANN-M4 and MLR-M4 with the performance efficiency of 6.618%, 1.6107% and 11.8301% respectively.

The comparative performance analyses of the single models are shown with the scatter plots Figure 4.1- 4.4. The scatter plot is employed to the adequate models of each of the single Model. Based on the plot, it can be observed that all the four models showed good agreement between the predicted and the experimental results; this is in line with Table 4.4. These scatter plots revealed ANFIS-M4 as the model with the highest prediction ability. Hence, it is the adequate single model in the simulation of the retention time of the various classes of the mycotoxins in the HPLC method development.

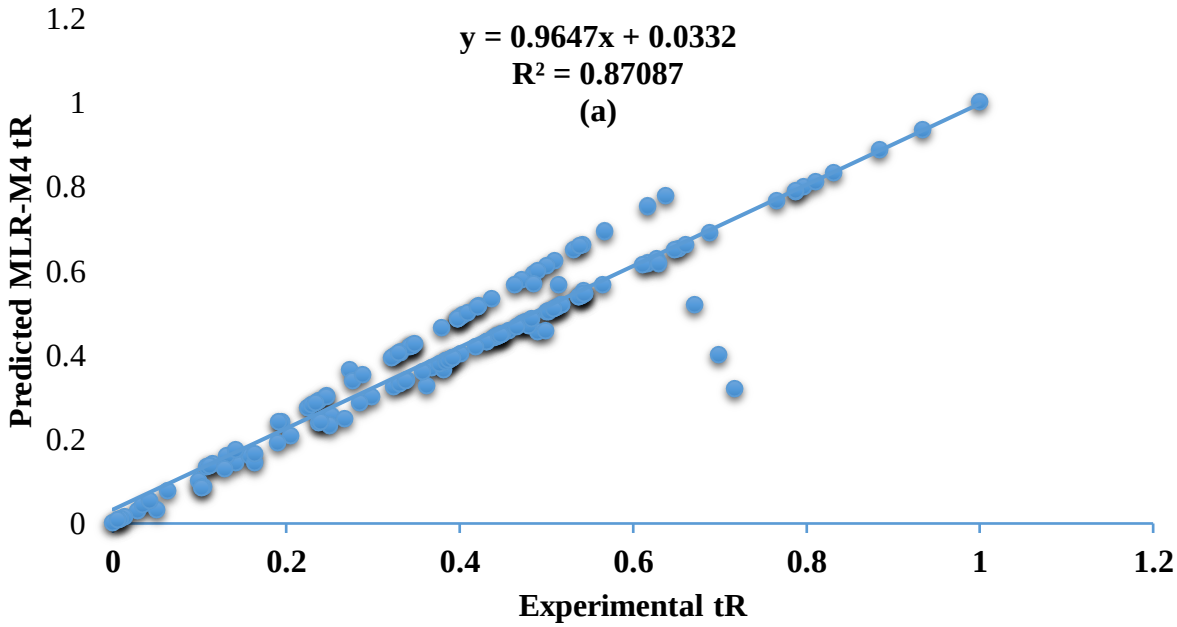


Figure 4.1: Graphical representations of the MLR-M4 results for modelling the retention time (tR) of various Mycotoxins using HPLC method

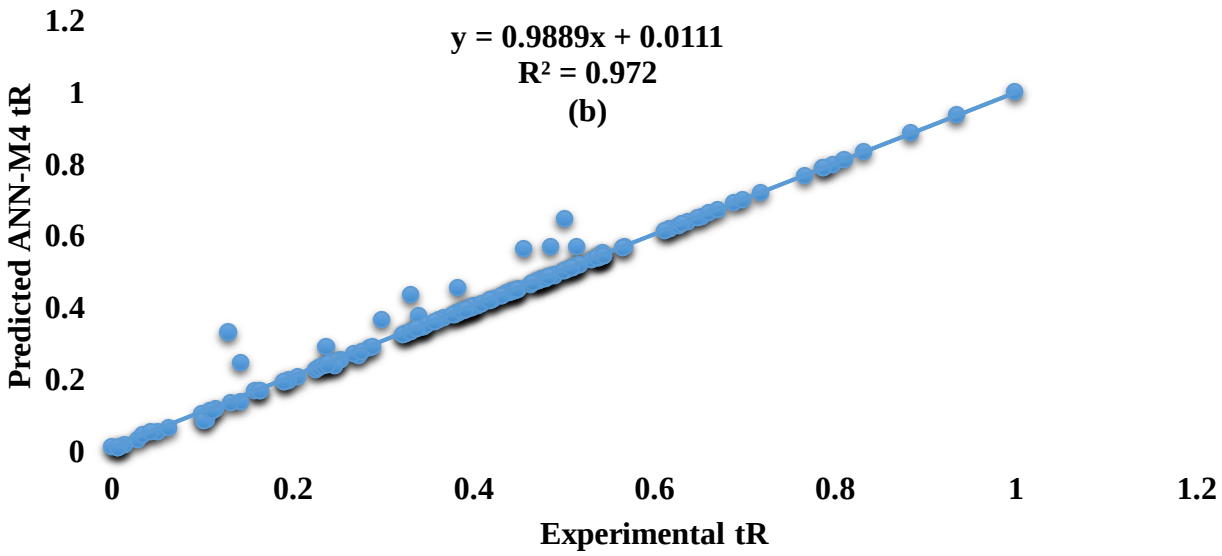


Figure 4.2: Graphical representations of the ANN-M4 results for modelling the retention time (tR) of various Mycotoxins using HPLC method

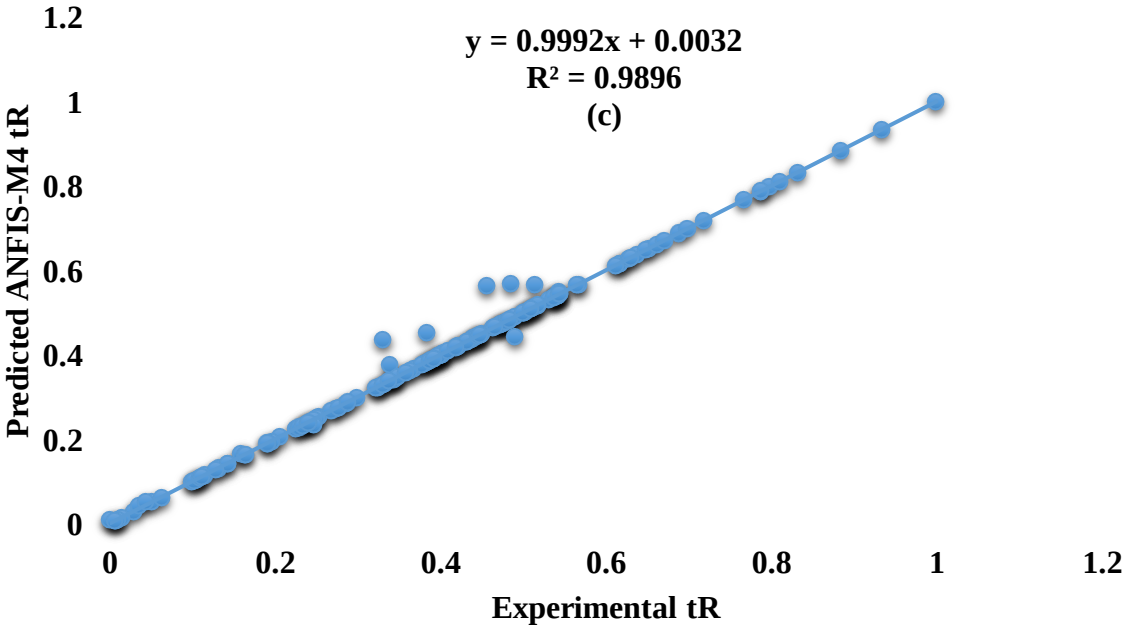


Figure 4.3: Graphical representations of the ANFIS-M4 results for modelling the retention time (tR) of various Mycotoxins using HPLC method

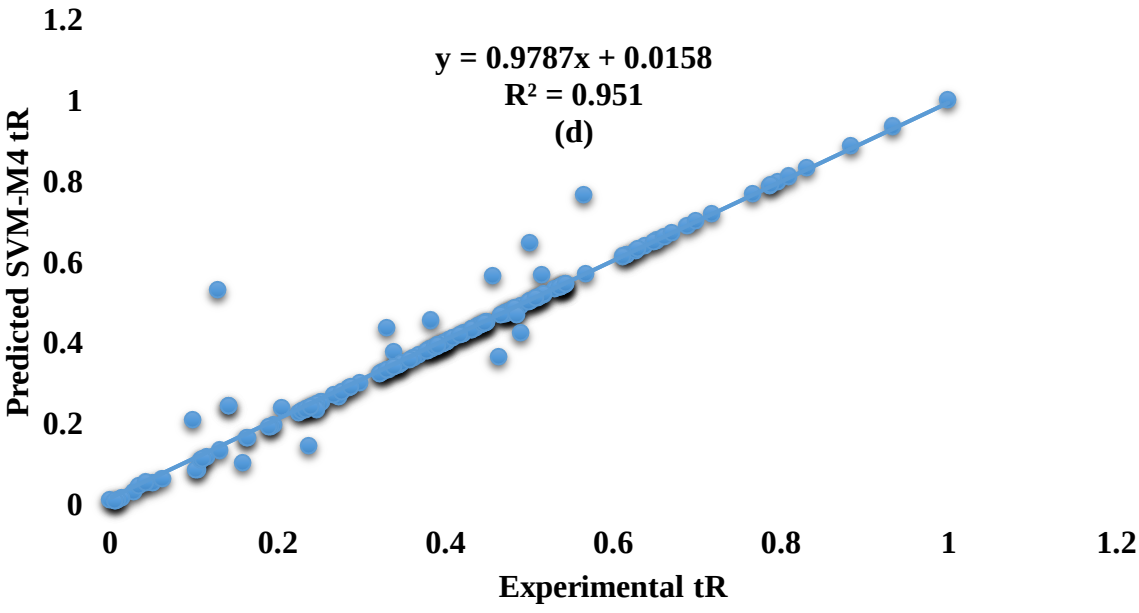


Figure 4.4: Graphical representations of the MLR-M4 results for modelling the retention time (tR) of various Mycotoxins using HPLC method

The comparative predictive performance analysis of the single models can equally be demonstrated using the performance error of the models, whereby the higher the error the lower the performance of the model and vice-versa. Figure 4.5 demonstrates the performance error using RMSE-values as described in Table 4.4. The essence of depicting this graph is to clearly demonstrate the performance of each model using a bar chart, which can be easier for understanding and assimilation. It is vital to note that ANFIS-M4 showed higher performance ability based on this illustration in both the verification and calibration phases. This assertion further shows that the model is adequate single model for the performance modeling of various classes of mycotoxin through the inputs and output parameter used in the study .

The capability of this model in capturing the chaotic and highly complex properties of the data is also revealed. These findings of this study are in agreement with those of Ghali et al., 2020 and D'Archivio, 2019. In which Ghali et al., 2020 demonstrated the applications of three different single AI-based models for modelling the performance of an anti-Alzheimer agent. The results of their work shows higher performance of these models (Ghali et al., 2020). D'Archivio, 2019 also reported the applications of MLP-ANN in the modelling of 16 amino acids in HPLC method developments. The results showed the effective application of ANN in modelling the amino acids (D'Archivio, 2019).

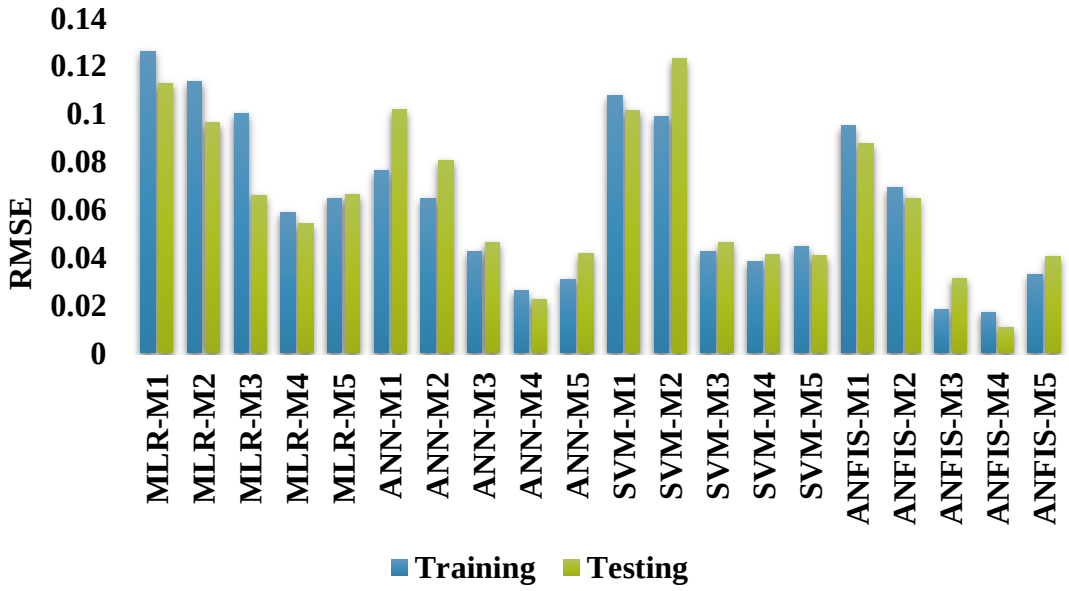


Figure 4.5: RMSE performance comparison of the single models

4.3: The advance hybrid-data intelligence approaches results

The hybrids data intelligence models are conceived in the attempt of strengthening methods used in the prediction and elucidation of the mechanisms and properties of mycotoxins with view of developing methods of prevention and control. Table 4.5 describes the performance efficiency of the hybrid models MLR-ANN, MLR-ANFIS and MLR-SVM. The hybrid involves the integration of the adequate models as revealed by the results of the single non-linear models (SVM-M4, ANFIS-M4 and ANN-M4) with that of the linear model (MLR-M4) to cope with the non-linearity in the experimental data. This integration is also aimed at dealing with the variability that may influence the accuracy of the estimated performance of the models.

Table 4.4 presents the effectiveness and efficiency of the hybrid-learning models as ascertained over the classical single model approaches currently being used for the prediction. The hybrid-intelligence MLR-ANFIS model showed the highest performance among other hybrid and single models having the highest values of CC and R^2 and the lower values of MSE and RMSE. Moreover, a research conducted by Abba et al., showed the application of various single models (MLR, ANN and ANFIS) used in the simulation of Amiloride and Methyclothiazide in HPLC method development. The result obtained is in agreement with the result demonstrated by the single models of this current research (Abba, Usman and Is, 2020). Abdullahi et al,. 2020 equally reported the application of ANN model in the prediction of isoqurcetin using HPLC method development.

The result in terms of R^2, RMSE and MSE is in agreement with the current result (Abdullahi et al., 2020). For example, to compare both the single and hybrid models in order to demonstrates the advantage of the hybrid models over the single models. From table 4.4 it can be observed that ANFIS single model demonstrates the highest performance with R^2-value of 0.9893603 in the training phases, while in for table 4.5 MLR-ANFIS hybrid model showed the highest performance accuracy with R2-value of 0.99996 in the training phase and this clearly indicates the higher performance of the hybrid models as compared with the single models.

Table 4.5: The hybrid models results

	Training				Testing			
Models	**R2**	**RMSE**	**MSE**	**CC**	**R2**	**RMSE**	**MSE**	**CC**
MLR-ANN	0.9823	0.0322	0.0010	0.9911	0.9828	0.0201	0.0004	0.9914
MLR-SVM	0.9736	0.0394	0.0016	0.9867	0.9741	0.0247	0.0006	0.9870
MLR-ANFIS	**0.99996**	**0.0015**	**2E-06**	**0.9999**	**0.9899**	**0.0154**	**0.0002**	**0.995**

The prediction accuracy of the hybrid machine learning method is equally shown using the radar plots. Nevertheless, the correlation co-efficient (CC) demonstrated in Table 4.5 indicated the applications of the hybrid-intelligence models in modelling the performance efficiency of the various mycotoxins classes in the HPLC technique. Figure 4.6 indicates the performance of the mycotoxins simulation using a radar plots in the testing and training phases. Moreover, it has been reported that the radar plots ranges between 0 to 1, with the low performing model closer to zero and the adequate model closer to one.

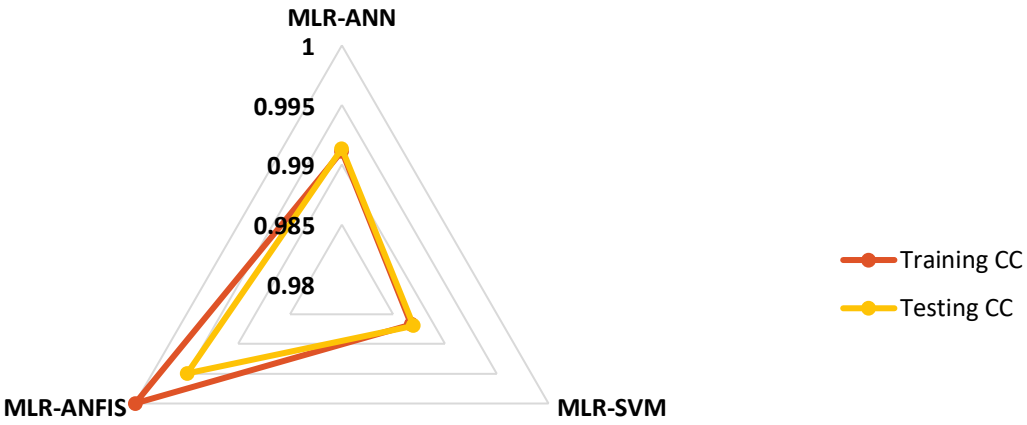
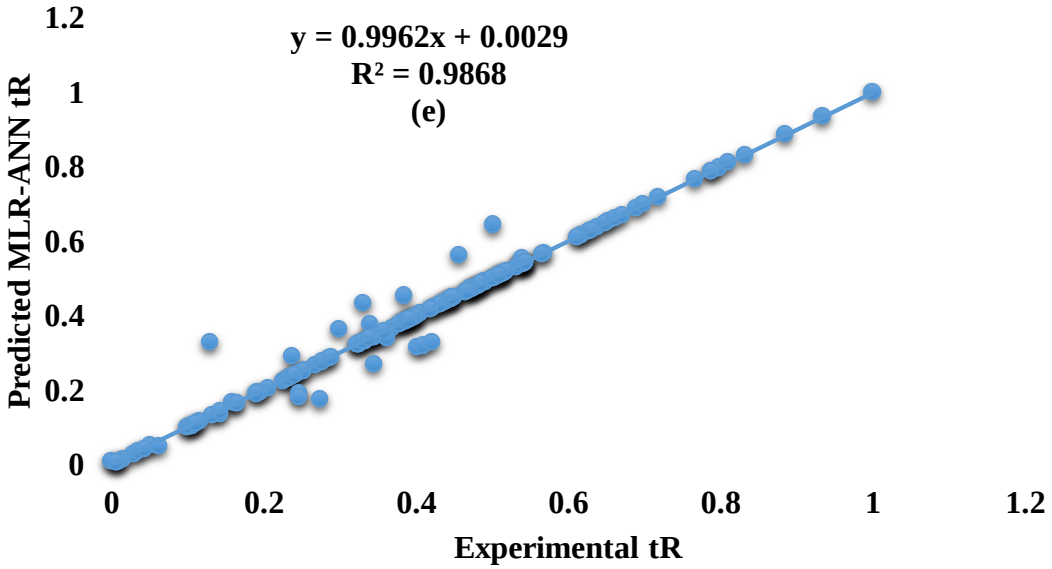
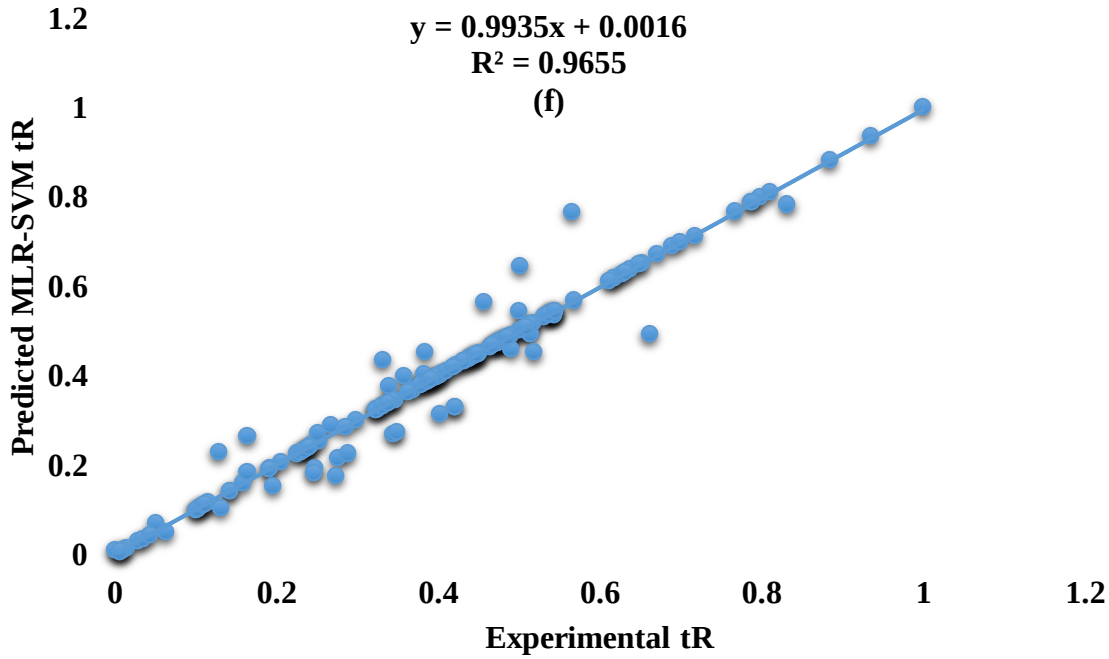


Figure 4.6 Comparative performance of correlation co-efficient (CC) in terms Radar chart for the hybrid learning approaches for mycotoxins modelling performance

Based on the radar chart, it can be observed that MLR-ANFIS have the highest performance accuracy among the other hybrid models.

The comparative performance of the hybrid-intelligence models are presented with the aid of the scatter plots to show the best fit the experimental and predicted values. Figure 4.7 indicates the performance of MLR-ANN, MLR-ANFIS and MLR-SVM through the application of the scatter plot.


y = 0.9962x + 0.0029
R² = 0.9868
(e)
Predicted MLR-ANN tR
Experimental tR
0
0.2
0.4
0.6
0.8
1
1.2


y = 0.9935x + 0.0016
R² = 0.9655
(f)
Predicted MLR-SVM tR
Experimental tR
0
0.2
0.4
0.6
0.8
1
1.2

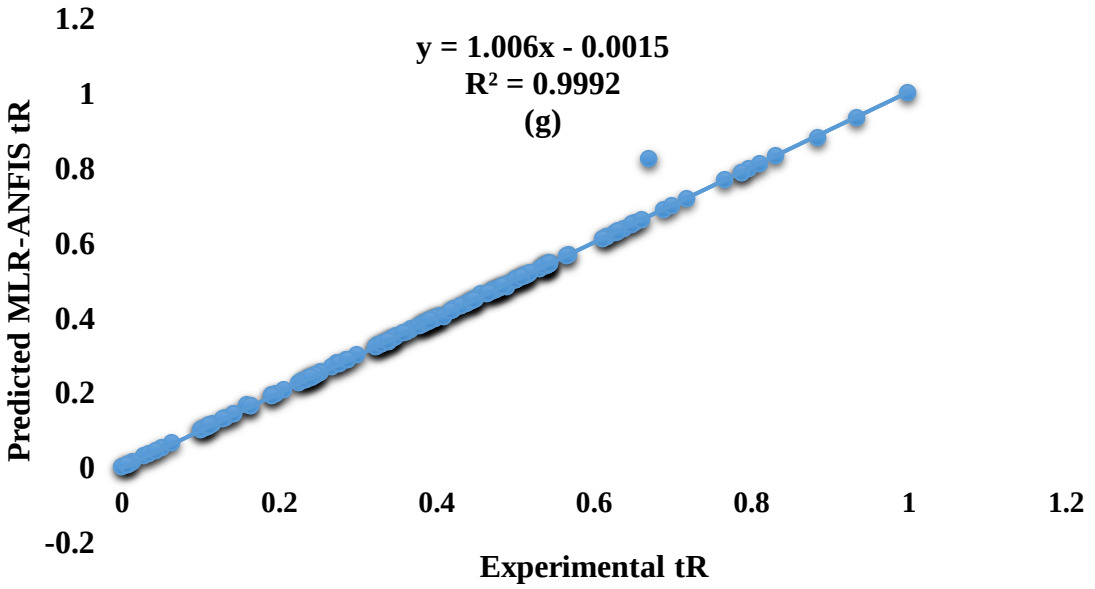


Figure 4.7: Graphical representation of the hybrid models (e) MLR-ANN (f) MLR-ANFIS (g) MLR-SVM

The performance of the hybrid models are checked based on their respective MSE in both the calibration and verification phases. Therefore, a graphical illustration informs through the bar chart shown in Figure 4.8. It can be observed that all the MSE values of the models are within the acceptable range. More so, this presentation as revealed is in tandem with the models presented on Table 4.5. Alsharksi et al., 2020 reported the application of hybrid ANFIS model in the simulation of clostridium from various patients. The results equally is in concordance with the assertions made in the current research (Alsharksi et al., 2020). Pham et al., 2019 also reported the application of the potential of hybrid models in modelling rainfall. The result demonstrated is in agreement with that of this current research (Pham et al., 2019).

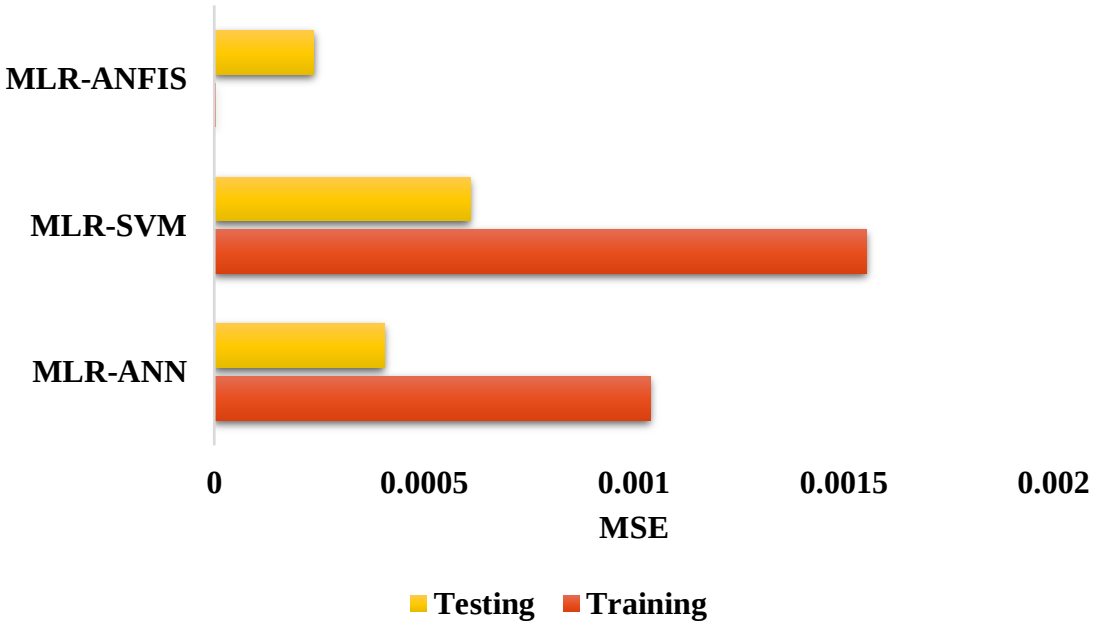


Figure 4.8: Performance comparison of the hybrid learning approach based on MSE for modelling the chromatographic performance of various mycotoxins classes

CHAPTER 5
CONCLUSION AND RECOMMENDATION

5.1 Conclusions

The major objective of this research is to assess the potential applications of various data driven approaches consisting of both single and hybrid modelling methods. These techniques are employed in order to predict the performance properties of various classes of mycotoxins using the HPLC technique. This research involved the modelling of various mycotoxins through their retention behaviour in HPLC method, in which the data used in the research were collected from previous research conducted in the literature. The retention behaviour is simulated using four different single models (SVM, ANFIS, ANN and MLR) with retention index, peak symmetry, mono isotopic mass and relative sensitivity factor as the input parameters, while the retention time is taken as the output parameter.

Based on the evaluation metrics, RMSE, CC, MSE and R^2, the result obtained showed that ANFIS-M4 has the highest performance, which is attributed to the non-linear ability of the model as well as the input parameter settings. Nonetheless, various results of the single models at different time interval account for the necessity to cope both the linear and linear models. Therefore, three different hybrid models (MLR-ANN. MLR-ANFIS and MLR-SVM) are developed to improve the performance of the single models. To conclude, the general comparative of both the single and hybrid techniques indicates that MLR-ANFIS showed higher performance than all the models employed the current research.

5.2 Recommendations

Moreover, as means for further research on the proposed hybrid MLR-based intelligence model, other machine learning methods in form of single models such as Hammerstein Wiener (HW), step-wise linear regression (SWLR), principal component analysis (PCA) and extreme learning machine (ELM), ensemble machine learning, other hybrid techniques (SWLR-HW, SWLR-ELM and SWLR-PCA etc) should be employed in the prediction of various mycotoxins using the HPLC method development. More so, it will be of paramount importance to employ the classification learning method in order to categorize

various classes of the mycotoxins using data from the HPLC and other analytical approaches. Furthermore, there is equally the need to carry out sensitivity analysis. Due to the fact that many criteria are going to be influencing the determination and thus variation due to uncertainty could be a thing of concern. Integrated multi-criteria approaches that are insensitive to the effects of variations such Taguchi-data envelopment could be tested to examine the effects of variations on the predictive tendency of the hybrid MLR-based intelligence model proposed.

REFERENCES

Abba, S. I., Usman, A. G., & Is, S. (2020). *Chemometrics and Intelligent Laboratory Systems Simulation for response surface in the HPLC optimization method development using arti fi cial intelligence models : A data-driven approach*. *201*(August 2019). https://doi.org/10.1016/j.chemolab.2020.104007

Abbaspour-Gilandeh, Y., Ghadakchi-Bazaz, H., & Davari, M. (2020). Discriminating Healthy Wheat Grains from Grains Infected with Fusarium graminearum Using Texture Characteristics of Image-Processing Technique, Discriminant Analysis, and Support Vector Machine Methods. *Journal of Intelligent Systems*, *29*(1), 1576–1586. https://doi.org/10.1515/jisys-2018-0430

Abdelhaliem, E., & Al-Otaibi, H. (2020). Modulation of ochratoxins a-induced genotoxicity in phaseolus vulgaris by multi-walled carbon nanotubes evaluated by phenotypic, cellular ultrastructure, and nuclear dna bioassays. *Genetics and Molecular Research*, *19*(1), 1–33. https://doi.org/10.4238/gmr18509

Abdelkhalek, A., Salem, M. Z. M., Kordy, A. M., Salem, A. Z. M., & Behiry, S. I. (2020). Antiviral, antifungal, and insecticidal activities of Eucalyptus bark extract: HPLC analysis of polyphenolic compounds. *Microbial Pathogenesis*, *147*(May), 104383. https://doi.org/10.1016/j.micpath.2020.104383

Abdullahi, H. U., Usman, A. G., & Abba, S. I. (2020). *Modelling the Absorbance of a Bioactive Compound in HPLC Method using Artificial Neural Network and Multilinear Regression Methods*. *6*(2), 362–371.

Adamson, G. E., Lazarus, S. A., Mitchell, A. E., Prior, R. L., Cao, G., Jacobs, P. H., … Schmitz, H. H. (1999). HPLC method for the quantification of procyanidins in cocoa and chocolate samples and correlation to total antioxidant capacity. *Journal of Agricultural and Food Chemistry*, *47*(10), 4184–4188. https://doi.org/10.1021/jf990317m

Afsah-Hejri, L., Hajeb, P., & Ehsani, R. J. (2020). Application of ozone for degradation of mycotoxins in food: A review. *Comprehensive Reviews in Food Science and Food*

Safety, *19*(4), 1777–1808. https://doi.org/10.1111/1541-4337.12594

Ahmed, A. A. M., Mustakim, S., & Shah, A. (2017). Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River. *Journal of King Saud University - Engineering Sciences*, *29*(3), 237–243. https://doi.org/10.1016/j.jksues.2015.02.001

Alsharif, A. M. A., Choo, Y. M., & Tan, G. H. (2019). Detection of five mycotoxins in different food matrices in the malaysian market by using validated liquid chromatography electrospray ionization triple quadrupole mass spectrometry. *Toxins*, *11*(4). https://doi.org/10.3390/toxins11040196

Alsharif, A. M. A., Choo, Y. M., Tan, G. H., & Abdulra'uf, L. B. (2019). Determination of Mycotoxins Using Hollow Fiber Dispersive Liquid–Liquid–Microextraction (HF-DLLME) Prior to High-Performance Liquid Chromatography–Tandem Mass Spectrometry (HPLC - MS/MS). *Analytical Letters*, *52*(12), 1976–1990. https://doi.org/10.1080/00032719.2019.1587766

Alsharksi, A. N., Danmaraya, Y. A., Abdullahi, H. U., Ghali, U. M., & Usman, A. G. (2020). *Potential of Hybrid Adaptive Neuro Fuzzy Model in Simulating Clostridium Difficile Infection Status*. (1), 1–6. https://doi.org/10.35940/ijbsac.A0191.073120

Aron, A. T., Gentry, E. C., McPhail, K. L., Nothias, L. F., Nothias-Esposito, M., Bouslimani, A., … Dorrestein, P. C. (2020). Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nature Protocols*, *15*(6), 1954–1991. https://doi.org/10.1038/s41596-020-0317-5

Arroyo-Manzanares, N., Gámiz-Gracia, L., & García-Campaña, A. M. (2012). Determination of ochratoxin A in wines by capillary liquid chromatography with laser induced fluorescence detection using dispersive liquid-liquid microextraction. *Food Chemistry*, *135*(2), 368–372. https://doi.org/10.1016/j.foodchem.2012.05.009

Assunção, R., Silva, M. J., & Alvito, P. (2016). Challenges in risk assessment of multiple mycotoxins in food. *World Mycotoxin Journal*, *9*(5), 791–811. https://doi.org/10.3920/WMJ2016.2039

Ayofemi Olalekan Adeyeye, S. (2020). Aflatoxigenic fungi and mycotoxins in food: a review. *Critical Reviews in Food Science and Nutrition*, *60*(5), 709–721. https://doi.org/10.1080/10408398.2018.1548429

Barmpalexis, P., Karagianni, A., Karasavvaides, G., & Kachrimanis, K. (2018). Comparison of multi-linear regression, particle swarm optimization artificial neural networks and genetic programming in the development of mini-tablets. *International Journal of Pharmaceutics*, *551*(1–2), 166–176. https://doi.org/10.1016/j.ijpharm.2018.09.026

Bauchet, J., Prieto, S., & Ricker-Gilbert, J. (2020). Improved Drying and Storage Practices that Reduce Aflatoxins in Stored Maize: Experimental Evidence from Smallholders in Senegal. *American Journal of Agricultural Economics*, *00*(00), 1–21. https://doi.org/10.1111/ajae.12106

Biswas, A. K., Rao, G. S., Kondaiah, N., Anjaneyulu, A. S. R., Mendiratta, S. K., Prasad, R., & Malik, J. K. (2007). A simple multi-residue method for determination of oxytetracycline, tetracycline and chlortetracycline in export buffalo meat by HPLC-photodiode array detector. *Journal of Food and Drug Analysis*, *15*(3), 278–284. https://doi.org/10.38212/2224-6614.2419

Bull, A. T., Ward, A. C., & Goodfellow, M. (2000). Search and Discovery Strategies for Biotechnology: the Paradigm Shift. *Microbiology and Molecular Biology Reviews*, *64*(3), 573–606. https://doi.org/10.1128/mmbr.64.3.573-606.2000

Carballo, D., Tolosa, J., Ferrer, E., & Berrada, H. (2019). Dietary exposure assessment to mycotoxins through total diet studies. A review. *Food and Chemical Toxicology*, *128*(March), 8–20. https://doi.org/10.1016/j.fct.2019.03.033

Chandwani, V., Vyas, S. K., Agrawal, V., & Sharma, G. (2015). Soft Computing Approach for Rainfall-runoff Modelling: A Review. *Aquatic Procedia*, *4*(April), 1054–1061. https://doi.org/10.1016/j.aqpro.2015.02.133

Choubin, B., Khalighi-Sigaroodi, S., Malekian, A., & Kişi, Ö. (2016). Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system

for forecasting precipitation based on large-scale climate signals. *Hydrological Sciences Journal*, *61*(6), 1001–1009. https://doi.org/10.1080/02626667.2014.966721

Colombo, R., & Papetti, A. (2020). Pre-concentration and analysis of mycotoxins in food samples by capillary electrophoresis. *Molecules*, *25*(15), 1–19. https://doi.org/10.3390/molecules25153441

Cui, J. J., Wang, L. Y., Tan, Z. R., Zhou, H. H., Zhan, X., & Yin, J. Y. (2020). Mass Spectrometry-Based Personalized Drug Therapy. *Mass Spectrometry Reviews*, *39*(5–6), 523–552. https://doi.org/10.1002/mas.21620

Cunha, S. C., & Fernandes, J. O. (2018). Extraction techniques with deep eutectic solvents. *TrAC - Trends in Analytical Chemistry*, *105*, 225–239. https://doi.org/10.1016/j.trac.2018.05.001

Cvjetko Bubalo, M., Vidović, S., Radojčić Redovniković, I., & Jokić, S. (2018). New perspective in extraction of plant biologically active compounds by green solvents. *Food and Bioproducts Processing*, *109*, 52–73. https://doi.org/10.1016/j.fbp.2018.03.001

D'Archivio, A. A. (2019). Artificial neural network prediction of retention of amino acids in reversed-phase HPLC under application of linear organic modifier gradients and/or pH gradients. *Molecules*, *24*(3). https://doi.org/10.3390/molecules24030632

D'Archivio, A. A., Giannitto, A., & Maggi, M. A. (2013). Cross-column prediction of gas-chromatographic retention of polybrominated diphenyl ethers. *Journal of Chromatography A*, *1298*, 118–131. https://doi.org/10.1016/j.chroma.2013.05.018

Dada, T. A., Ekwomadu, T. I., & Mwanza, M. (2020). Multi mycotoxin determination in dried beef using liquid chromatography coupled with triple quadrupole mass spectrometry (LC-MS/MS). *Toxins*, *12*(6). https://doi.org/10.3390/toxins12060357

Dankovich, R. S. (2019). Clinical sings and pathomorphologic changes of pigs by the spontaneous ochratoxicosis. *Scientific Messenger of LNU of Veterinary Medicine and Biotechnologies*, *21*(96), 75–80. https://doi.org/10.32718/nvlvet9613

El-Nezami, H., Kankaanpaa, P., Salminen, S., & Ahokas, J. (1998). Ability of dairy strains of lactic acid bacteria to bind a common food carcinogen, aflatoxin B1. *Food and Chemical Toxicology*, *36*(4), 321–326. https://doi.org/10.1016/S0278-6915(97)00160-9

El-Shahir, A. A., Abdel-Sater, M. A., & Yassein, A. S. (2020). Lipolysis Activity of Aflatoxins and Ochratoxins Producing Fungal Strains Contaminating Fresh Cow Meat. *Biosciences, Biotechnology Research Asia*, *17*(01), 163–172. https://doi.org/10.13005/bbra/2821

Elkiran, G., Nourani, V., & Abba, S. I. (2019). Multi-step ahead modelling of river water quality parameters using ensemble artificial intelligence-based approach. *Journal of Hydrology*, *577*(July), 123962. https://doi.org/10.1016/j.jhydrol.2019.123962

Elkiran, G., Nourani, V., Abba, S. I., & Abdullahi, J. (2018). Artificial intelligence-based approaches for multi-station modelling of dissolve oxygen in river. *Global Journal of Environmental Science and Management*, *4*(4), 439–450. https://doi.org/10.22034/gjesm.2018.04.005

Elkiran, Gozen, Nourani, V., & Abba, S. I. (2019). Multi-step ahead modelling of river water quality parameters using ensemble artificial intelligence-based approach. *Journal of Hydrology*, *577*(April), 123962. https://doi.org/10.1016/j.jhydrol.2019.123962

Encarnação, T. et al. (2020). Development and validation of a RP-HPLC method for the simultaneous analysis of paracetamol, ibuprofen, olanzapine, and simvastatin during microalgae bioremediation. Submitted for publication inEnvironmental. *Methods X*, 101083. https://doi.org/10.1016/j.mex.2020.101083

Evtugyn, G., & Hianik, T. (2019). Electrochemical immuno- and aptasensors for mycotoxin determination. *Chemosensors*, *7*(1). https://doi.org/10.3390/CHEMOSENSORS7010010

Fish, P. P., Fernandes, J. M. O., & Ivanova, L. (2019). Chromatography High-Resolution Mass Spectrometry. *Toxins*, *11*.

Foerster, C., Muñoz, K., Delgado-Rivera, L., Rivera, A., Cortés, S., Müller, A., … Rios, G. (2020). Occurrence of relevant mycotoxins in food commodities consumed in Chile. *Mycotoxin Research*, *36*(1), 63–72. https://doi.org/10.1007/s12550-019-00369-5

Gaya, M. S., Abdul Wahab, N., Sam, Y. M., & Samsudin, S. I. (2014). ANFIS Modelling of Carbon and Nitrogen Removal in Domestic Wastewater Treatment Plant. *Jurnal Teknologi*, *67*(5). https://doi.org/10.11113/jt.v67.2839

Ge, H., Jiang, Y., Lian, F., Zhang, Y., & Xia, S. (2016). Quantitative determination of aflatoxin B1 concentration in acetonitrile by chemometric methods using terahertz spectroscopy. *Food Chemistry*, *209*, 286–292. https://doi.org/10.1016/j.foodchem.2016.04.070

Ghaedi, M., Hosaininia, R., Ghaedi, A. M., Vafaei, A., & Taghizadeh, F. (2014). Adaptive neuro-fuzzy inference system model for adsorption of 1,3,4-thiadiazole-2,5-dithiol onto gold nanoparticales-activated carbon. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *131*, 606–614. https://doi.org/10.1016/j.saa.2014.03.055

Ghali, U. M., Alhosen, M., Degm, A., Alsharksi, A. N., Hoti, Q., & Usman, A. G. (2020). *Development Of Computational Intelligence Algorithms For Modelling The Performance Of Humanin And Its Derivatives In HPLC Optimization Method Development*. *9*(08), 110–117.

Ghorbani, M. A., Deo, R. C., Yaseen, Z. M., H. Kashani, M., & Mohammadi, B. (2018). Pan evaporation prediction using a hybrid multilayer perceptron-firefly algorithm (MLP-FFA) model: case study in North Iran. *Theoretical and Applied Climatology*, *133*(3–4), 1119–1131. https://doi.org/10.1007/s00704-017-2244-0

Györgyi, B. (2010). *Risk effects of the spread route of mycotoxins Risk effects of the spread route of mycotoxins*. (January).

Haghiabi, A. H., Azamathulla, H. M., & Parsaie, A. (2017). Prediction of head loss on cascade weir using ANN and SVM. *ISH Journal of Hydraulic Engineering*, *23*(1), 102–110. https://doi.org/10.1080/09715010.2016.1241724

Javanmardi, F., Khodaei, D., Sheidaei, Z., Bashiry, M., Nayebzadeh, K., Vasseghian, Y., & Mousavi Khaneghah, A. (2020). Decontamination of Aflatoxins in Edible Oils: A Comprehensive Review. *Food Reviews International*, *00*(00), 1–17. https://doi.org/10.1080/87559129.2020.1812635

Karimi, S., Kisi, O., Shiri, J., & Makarynskyy, O. (2013). Neuro-fuzzy and neural network techniques for forecasting sea level in Darwin Harbor, Australia. *Computers and Geosciences*, *52*, 50–59. https://doi.org/10.1016/j.cageo.2012.09.015

Kazemi, P., Khalid, M. H., Szlek, J., Mirtič, A., Reynolds, G. K., Jachowicz, R., & Mendyk, A. (2016). Computational intelligence modeling of granule size distribution for oscillating milling. *Powder Technology*, *301*, 1252–1258. https://doi.org/10.1016/j.powtec.2016.07.046

Kazienko, P., Lughofer, E., & Trawiński, B. (2013). Hybrid and ensemble methods in machine learning J.UCS special issue. *Journal of Universal Computer Science*, *19*(4), 457–461.

Khademi, F., & Behfarnia, K. (2016). Evaluation of Concrete Compressive Strength Using Artificial Neural Network and Multiple Linear Regression Models. *Iust*, *6*(3), 423–432. Retrieved from http://ijoce.iust.ac.ir/article-1-260-en.html

Khosrokhavar, R., Ghasemi, J. B., & Shiri, F. (2010). 2D Quantitative Structure-property relationship study of mycotoxins by multiple linear regression and support vector machine. *International Journal of Molecular Sciences*, *11*(9), 3052–3068. https://doi.org/10.3390/ijms11093052

Kong, W., Wei, R., Logrieco, A. F., Wei, J., Wen, J., Xiao, X., & Yang, M. (2014). Occurrence of toxigenic fungi and determination of mycotoxins by HPLC-FLD in functional foods and spices in China markets. *Food Chemistry*, *146*, 320–326. https://doi.org/10.1016/j.foodchem.2013.09.005

Kotapati, H. K., & Bates, P. D. (2020). Normal phase HPLC method for combined separation of both polar and neutral lipid classes with application to lipid metabolic flux. *Journal of Chromatography B: Analytical Technologies in the Biomedical and*

Life Sciences, *1145*(March), 122099. https://doi.org/10.1016/j.jchromb.2020.122099

Lee, J. K., Han, W. S., Lee, J. S., & Yoon, C. N. (2017). A Novel Computational Method for Biomedical Binary Data Analysis: Development of a Thyroid Disease Index Using a Brute-Force Search with MLR Analysis. *Bulletin of the Korean Chemical Society*, *38*(12), 1392–1397. https://doi.org/10.1002/bkcs.11308

Liu, X. R., Zhang, M. M., & Gross, M. L. (2020). Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications. *Chemical Reviews*. https://doi.org/10.1021/acs.chemrev.9b00815

Lola, M. S., Noor, M., Ramlee, A., Gunalan, G. S., Zainuddin, N. H., Zakariya, R., … Khalil, I. (2016). *Improved the Prediction of Multiple Linear Regression Model Performance Using the Hybrid Approach : A Case Study of Chlorophyll-a at the Offshore Kuala Terengganu , Terengganu*. (October). https://doi.org/10.4236/ojs.2016.65065

Lotfi, E., & Akbarzadeh-T., M. R. (2014). Adaptive brain emotional decayed learning for online prediction of geomagnetic activity indices. *Neurocomputing*, *126*, 188–196. https://doi.org/10.1016/j.neucom.2013.02.040

Lotfi, E., & Akbarzadeh-T, M. R. (2013). Brain emotional learning-based pattern recognizer. *Cybernetics and Systems*, *44*(5), 402–421. https://doi.org/10.1080/01969722.2013.789652

Ma, J., Cai, J., Lin, G., Chen, H., Wang, X., Wang, X., & Hu, L. (2014). Development of LC-MS determination method and back-propagation ANN pharmacokinetic model of corynoxeine in rat. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, *959*, 10–15. https://doi.org/10.1016/j.jchromb.2014.03.024

Magan, N., & Olsen, M. (2004). Mycotoxins in Food: Detection and Control. In *Mycotoxins in Food: Detection and Control*. https://doi.org/10.1533/9781855739086

Marrero-Ponce, Y., Barigye, S. J., Jorge-Rodríguez, M. E., & Tran-Thi-Thu, T. (2018). QSRR prediction of gas chromatography retention indices of essential oil

components. *Chemical Papers*, *72*(1), 57–69. https://doi.org/10.1007/s11696-017-0257-x

Martins, C., Assunção, R., Nunes, C., Torres, D., & Alvito, P. (2020). Are Data from Mycotoxins' Urinary Biomarkers and Food Surveys Linked? A Review Underneath Risk Assessment. *Food Reviews International*, *00*(00), 1–26. https://doi.org/10.1080/87559129.2019.1709200

Marume, P., Ndlovu, N., Manhokwe, S., & Ndlovu, C. (2020). *African Journal of Biological Sciences Detection of mycotoxins in opaque beer production*. *2*(4), 84–91. https://doi.org/10.33472/AFJBS.2.4.2020.84-91

Nielsen, K. F., & Smedsgaard, J. (2003). Fungal metabolite screening: Database of 474 mycotoxins and fungal metabolites for dereplication by standardised liquid chromatography-UV-mass spectrometry methodology. *Journal of Chromatography A*, *1002*(1–2), 111–136. https://doi.org/10.1016/S0021-9673(03)00490-4

Nourani, V., Molajou, A., Uzelaltinbulat, S., & Sadikoglu, F. (2019). Emotional artificial neural networks (EANNs) for multi-step ahead prediction of monthly precipitation; case study: northern Cyprus. *Theoretical and Applied Climatology*, *138*(3–4), 1419–1434. https://doi.org/10.1007/s00704-019-02904-x

Ostertag, F., Schmidt, C. M., Berensmeier, S., & Hinrichs, J. (2020). Development and validation of an RP-HPLC DAD method for the simultaneous quantification of minor and major whey proteins. *Food Chemistry*, (September), 128176. https://doi.org/10.1016/j.foodchem.2020.128176

Pakshir, K., Mirshekari, Z., Nouraei, H., Zareshahrabadi, Z., Zomorodian, K., Khodadadi, H., & Hadaegh, A. (2020). Mycotoxins Detection and Fungal Contamination in Black and Green Tea by HPLC-Based Method. *Journal of Toxicology*, *2020*. https://doi.org/10.1155/2020/2456210

Park, S. H., Haddad, P. R., Talebi, M., Tyteca, E., Amos, R. I. J., Szucs, R., … Pohl, C. A. (2017). Retention prediction of low molecular weight anions in ion chromatography based on quantitative structure-retention relationships applied to the linear solvent

strength model. *Journal of Chromatography A*, *1486*, 68–75. https://doi.org/10.1016/j.chroma.2016.12.048

Pfeifer, N., Leinenbach, A., Huber, C. G., & Kohlbacher, O. (2007). Statistical learning of peptide retention behavior in chromatographic separations: A new kernel-based approach for computational proteomics. *BMC Bioinformatics*, *8*, 1–14. https://doi.org/10.1186/1471-2105-8-468

Pham, Q. B., Abba, S. I., Usman, A. G., Thi, N., & Linh, T. (2019). *Potential of Hybrid Data-Intelligence Algorithms for Multi-Station Modelling of Rainfall*.

Phwan, C. K., Chew, K. W., Sebayang, A. H., Ong, H. C., Ling, T. C., Malek, M. A., … Show, P. L. (2019). Effects of acids pre-treatment on the microbial fermentation process for bioethanol production from microalgae. *Biotechnology for Biofuels*, *12*(1), 1–9. https://doi.org/10.1186/s13068-019-1533-5

Roohi, R., Hashemi, S. M. B., & Mousavi Khaneghah, A. (2020). Kinetics and thermodynamic modelling of the aflatoxins decontamination: a review. *International Journal of Food Science and Technology*, 1–8. https://doi.org/10.1111/ijfs.14689

Saini, R., & Kumar, P. (2016). Optimization of chlorpyrifos degradation by Fenton oxidation using CCD and ANFIS computing technique. *Journal of Environmental Chemical Engineering*, *4*(3), 2952–2963. https://doi.org/10.1016/j.jece.2016.06.003

Schlosser, O., Robert, S., & Noyon, N. (2020). Airborne mycotoxins in waste recycling and recovery facilities: Occupational exposure and health risk assessment. *Waste Management*, *105*, 395–404. https://doi.org/10.1016/j.wasman.2020.02.031

Sharghi, E., Nourani, V., & Behfar, N. (2018). Earthfill dam seepage analysis using ensemble artificial intelligence based modeling. *Journal of Hydroinformatics*, jh2018151. https://doi.org/10.2166/hydro.2018.151

Sobral, M. M. C., Faria, M. A., Cunha, S. C., Miladinovic, B., & Mplvo, I. (2019). Transport of mycotoxins across human gastric NCI – N87 and intestinal Caco-2 cell models. *Food and Chemical Toxicology*, *131*(May), 110595. https://doi.org/10.1016/j.fct.2019.110595

Solgi, A., Pourhaghi, A., Bahmani, R., & Zarei, H. (2017). Improving SVR and ANFIS performance using wavelet transform and PCA algorithm for modeling and predicting biochemical oxygen demand (BOD). *Ecohydrology and Hydrobiology*, *17*(2), 164–175. https://doi.org/10.1016/j.ecohyd.2017.02.002

Sparkman, O. D. (2008). GC/MS: A practical user's guide, 2nd edition Marvin McMaster. *Journal of the American Society for Mass Spectrometry*, *19*(8), R1–R5. https://doi.org/10.1016/j.jasms.2008.05.001

Stangierski, J., Weiss, D., & Kaczmarek, A. (2019). Multiple regression models and Artificial Neural Network (ANN) as prediction tools of changes in overall quality during the storage of spreadable processed Gouda cheese. *European Food Research and Technology*, *245*(11), 2539–2547. https://doi.org/10.1007/s00217-019-03369-y

Su, M., Zhang, Z., Zhu, Y., & Zha, D. (2019). Data-driven natural gas spot price forecasting with least squares regression boosting algorithm. *Energies*, *12*(6). https://doi.org/10.3390/en12061094

Tao, Y., Wang, Y., Pan, M., Zhong, S., Wu, Y., Yang, R., … Zhou, J. (2017). Combined ANFIS and numerical methods to simulate ultrasound-assisted extraction of phenolics from chokeberry cultivated in China and analysis of phenolic composition. *Separation and Purification Technology*, *178*, 178–188. https://doi.org/10.1016/j.seppur.2017.01.012

Tessini, C., Mardones, C., von Baer, D., Vega, M., Herlitz, E., Saelzer, R., … Torres, O. (2010). Alternatives for sample pre-treatment and HPLC determination of Ochratoxin A in red wine using fluorescence detection. *Analytica Chimica Acta*, *660*(1–2), 119–126. https://doi.org/10.1016/j.aca.2009.11.011

Tsakiris, I. N., Tzatzarakis, M. N., Alegakis, A. K., Vlachou, M. I., Renieri, E. A., & Tsatsakis, A. M. (2013). Risk assessment scenarios of children's exposure to aflatoxin M1 residues in different milk types from the Greek market. *Food and Chemical Toxicology*, *56*, 261–265. https://doi.org/10.1016/j.fct.2013.02.024

Vapnik, V. (1995). *The nature of statistical learning theory* (p. 188). p. 188.

https://doi.org/10.1007/978-1-4757-2440-0

Veenaas, C., Linusson, A., & Haglund, P. (2018). Retention-time prediction in comprehensive two-dimensional gas chromatography to aid identification of unknown contaminants. *Analytical and Bioanalytical Chemistry*, *410*(30), 7931–7941. https://doi.org/10.1007/s00216-018-1415-x

Wang, Mei, Avula, B., Wang, Y. H., Zhao, J., Avonto, C., Parcher, J. F., … Khan, I. A. (2014). An integrated approach utilising chemometrics and GC/MS for classification of chamomile flowers, essential oils and commercial products. *Food Chemistry*, *152*, 391–398. https://doi.org/10.1016/j.foodchem.2013.11.118

Wang, Mingxun, Jarmusch, A. K., Vargas, F., Aksenov, A. A., Gauglitz, M., Weldon, K., … Gentry, E. C. (2020). *HHS Public Access*. *38*(1), 23–26. https://doi.org/10.1038/s41587-019-0375-9.Mass

Xiong, Y., Zhang, P., Warner, R. D., Shen, S., Johnson, S., & Fang, Z. (2020). Comprehensive profiling of phenolic compounds by HPLC-DAD-ESI-QTOF-MS/MS to reveal their location and form of presence in different sorghum grain genotypes. *Food Research International*, *137*(September), 109671. https://doi.org/10.1016/j.foodres.2020.109671

Xuan, Q., Zheng, F., Yu, D., Ouyang, Y., Zhao, X., Hu, C., & Xu, G. (2020). Rapid lipidomic profiling based on ultra-high performance liquid chromatography–mass spectrometry and its application in diabetic retinopathy. *Analytical and Bioanalytical Chemistry*, *412*(15), 3585–3594. https://doi.org/10.1007/s00216-020-02632-6

Yaseen, Z. M., Ehteram, M., Hossain, M. S., Fai, C. M., Koting, S. B., Mohd, N. S., … El-Shafie, A. (2019). A novel hybrid evolutionary data-intelligence algorithm for irrigation and power production management: Application to multi-purpose reservoir systems. *Sustainability (Switzerland)*, *11*(7). https://doi.org/10.3390/su11071953

Yaseen, Z. M., Ghareb, M. I., Ebtehaj, I., Bonakdari, H., Siddique, R., Heddam, S., … Deo, R. (2018). Rainfall Pattern Forecasting Using Novel Hybrid Intelligent Model Based ANFIS-FFA. *Water Resources Management*, *32*(1), 105–122.

https://doi.org/10.1007/s11269-017-1797-0

Yassein, A. S., El-Said, A. H. M., & El-Dawy, E. G. A. (2020). Biocontrol of toxigenic Aspergillus strains isolated from baby foods by essential oils. *Flavour and Fragrance Journal*, *35*(2), 182–189. https://doi.org/10.1002/ffj.3551

Yoshikawa, K., Furuno, M., Tanaka, N., & Fukusaki, E. (2020). Fast enantiomeric separation of amino acids using liquid chromatography/mass spectrometry on a chiral crown ether stationary phase. *Journal of Bioscience and Bioengineering*, *130*(4), 437–442. https://doi.org/10.1016/j.jbiosc.2020.05.007

Zapata, M., Rodríguez, F., & Garrido, J. L. (2000). Separation of chlorophylls and carotenoids from marine phytoplankton: A new HPLC method using a reversed phase C8 column and pyridine-containing mobile phases. *Marine Ecology Progress Series*, *195*, 29–45. https://doi.org/10.3354/meps195029

Zhang, Z., Hu, X., Zhang, Q., & Li, P. (2016). Determination for multiple mycotoxins in agricultural products using HPLC-MS/MS via a multiple antibody immunoaffinity column. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, *1021*, 145–152. https://doi.org/10.1016/j.jchromb.2016.02.035

APPENDICES

APPENDIX I: ETHICAL APPROVAL DOCUMENT

NEAR EAST UNIVERSITY

ETHICAL APPROVAL DOCUMENT

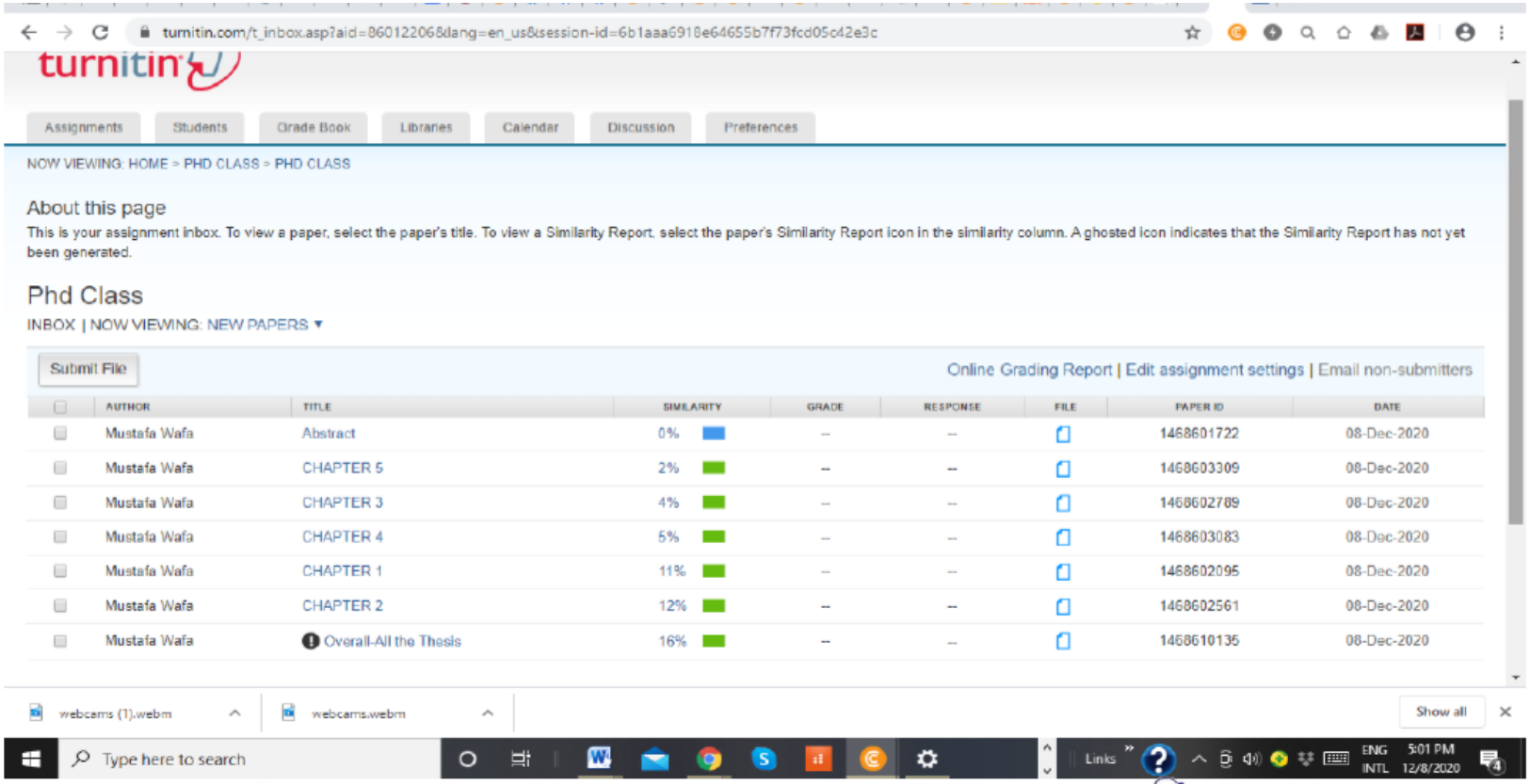
Date: 10/12/2020

To the Graduate School of Applied Sciences

For the thesis project entitled as "**DEVELOPING AN ALTERNATIVE HIGH PERFORMANCE HYBRID MULTILINEAR REGRESSION-BASED INTELLIGENCE MODEL FOR QUANTITATIVE STRUCTURE-PROPERTY RELATIONSHIP (QSPR) STUDIES OF MYCOTOXINS IN FOODS**", the researchers declare that they did not collect any data from human/animal or any other subjects.
Therefore, this project does not need to go through the ethics committee evaluation.

Title: Assist. Prof. Dr.
Name Surname: Kehinde Adewale ADESINA
Signature:
Role in the Research Project: Supervisor

APPENDIX II: THESIS PLAGIARISM INDEX



Mustafa Ibrahim Wafa ..... Assist. Prof. Dr. Kehinde Adewale Adesina.....

(Student) (Supervisor)