Université Mohamed Khider de Biskra
Faculté des sciences et de la technologie
Département de chimie industrielle

MÉMOIRE DE MASTER

Domaine : Sciences et Technologies
Filière : Génie des procédés
Spécialité : Génie chimique

Réf. : Entrez la référence du document

Présenté et soutenu par :
Ibtissam LAKHDARI
Siham CHAMEKH
Le : 2/juin /2025

Prediction of Liquid-Liquid Equilibrium Using Artificial Neural Networks

---

Jury :

Dr	Khaled ATHMANI	MCB	Université de Biskra	Encadreur
Dr	Elhachmi GUETTAF TAMAM	MCA	Université de Biskra	Examinateur
Dr	Kalthoum ADAIKA	MCA	Université de Biskra	Président

Année universitaire : 2024- 2025

ACKNOWLEDGMENT

First of all, we want to thank Allah for giving us enough power, guidance and precious opportunity to complete this work.

My sincere thanks go to our advisor ***Dr. Khaled ATHMANI*** *for his patience, noble guidance and support to accomplish this work. He has been giving us encouragement and valuable suggestions for this whole year. Thank you very much* ***Dr. ATHMANI****.*

We would like to express our deep gratitude to ***Dr. Elhachmi Guettaf Tamam*** *and* ***Dr. Kalthoum Adaika****, who have good experience, for agreeing to judge this work and enriching it with these constructive comments and criticisms.*

Thank you to all the teachers in the Department of Industrial Chemistry.

Last but not least, we would like to thank and dedicate this research paper to our loved parents, our whole family, our friends and our colleagues inuniversity of Biskra.

Thank you all.

Dediсation

{فَرِحِينَ بِمَآ ءَاتَىٰهُمُ ٱللَّهُ مِن فَضۡلِهِۦ}

Praise be to God, by whose grace good deeds are accomplished, and by whose guidance goals are achieved.

To my beloved mother, my refuge in times of weakness, and my companion on my journey. To you, who have always been the heart that beats with prayers for me, to you who walked and toiled so that this day might come true. You are the light that illuminated my path and the bridge that I crossed to success.

To the soul of my dear departed heart, who departed this world but never departed my heart. To the one who left me, but whose soul still flutters in the sky of my life. To that pure soul.

I ask God to make your abode in Paradise and to bless the reward of this work so that it may bring you light and mercy.

To my brother my hero, "Lakhdari Zouhir", and all my brothers, "Mohammed", "Karim", and sisters, "Chaima", "Iman", "Houda", and "Samia", my life partners and support, who have always been with me every step of the way.

To my esteemed professor, "Athmani Khaled", I extend my sincere thanks and appreciation for all your efforts.

To all my family and friends, thank you for your support and encouragement, especially to my favorite person. I hope that together we will achieve more successes.

To all of you, I dedicate this humble endeavor, filled with pride and gratitude, the fruits of what God has guided me to.

***Ibtissam Lakhdari*.**

Dedication

Praise be to Allah, by whose grace good deeds are completed and goals are achieved.

I extend my heartfelt thanks and gratitude to those who, after Allah, played a significant role in helping me reach this stage of my academic journey:

To my beloved mother, the source of love and compassion, whose prayers have always been my support, and whose constant presence has been the secret behind my success.

To my dear father, my role model and first teacher, who instilled in me a love for knowledge and perseverance, and who has always been a pillar of strength and encouragement.

To my dear siblings, who stood by my side, shared both the hardships and the joys, and to whom I owe all my love and appreciation.

To my loyal friends, my companions along this path, who lightened the burdens of the journey and brought joy and hope into my academic life.

I would also like to express my sincere appreciation to all my esteemed professors, who generously shared their knowledge and guidance and were true role models for us.

A special thanks goes to my supervisor, Khaled ATHMANI , for their dedication, guidance, and patience throughout the preparation of this thesis. I am truly grateful for their support.

To everyone who contributed to my success, even with a kind word, sincere advice, or a supportive smile—I thank you from the bottom of my heart.

This achievement would not have been possible without the grace of Allah and your presence in my life.

Siham Chamekh .

Abstract:

This study explores the use of Artificial Neural Networks (ANNs) to model phase compositions in a liquid-liquid extraction system based on experimental data. Several ANN architectures were developed and tested by varying the number of hidden layers (0, 1, and 2) and the number of neurons in each layer. The performance of each configuration was evaluated using mean squared error (MSE) and the coefficient of determination (R^2). The results showed that increasing the network complexity led to improved prediction accuracy. The most effective architecture, with a 6/12/12/6 configuration (input layer/hidden layer 1/hidden layer 2/output layer), achieved a low MSE of 0.00023. These results demonstrate that ANNs are powerful and reliable tools for modeling complex separation processes such as liquid-liquid extraction.

Key words: Modeling, ANN, L-L extraction, phase equilibrium.

ملخص:

تستكشف هذه الدراسة استخدام الشبكات العصبية الاصطناعية (ANNs) لنمذجة تركيبات الطور في نظام استخلاص سائل-سائل بناءً على بيانات تجريبية. طُوّرت واختُبرت العديد من هياكل الشبكات العصبية الاصطناعية (ANNs) من خلال تغيير عدد الطبقات المخفية (0، 1، و2) وعدد الخلايا العصبية في كل طبقة. قُيّم أداء كل تكوين باستخدام متوسط مربع الخطأ (MSE) ومعامل التحديد (R^2). أظهرت النتائج أن زيادة تعقيد الشبكة أدى إلى تحسين دقة التنبؤ. حققت البنية الأكثر فعالية، ذات التكوين 6/12/12/6 (طبقة الإدخال/الطبقة المخفية 1/الطبقة المخفية 2/طبقة الإخراج)، قيمة MSE منخفضة قدرها 0.00023. تُظهر هذه النتائج أن الشبكات العصبية الاصطناعية (ANNs) أدوات قوية وموثوقة لنمذجة عمليات الفصل المعقدة مثل استخلاص سائل-سائل.

الكلمات المفتاحية : النمذجة، الشبكات العصبية الاصطناعية، استخراج سائل-سائل ، توازن الطور.

Résumé :

Cette étude explore l'utilisation de réseaux de neurones artificiels (RNA) pour modéliser les compositions de phases dans un système d'extraction liquide-liquide, à partir de données expérimentales. Plusieurs architectures de RNA ont été développées et testées en faisant varier le nombre de couches cachées (0, 1 et 2) et le nombre de neurones dans chaque couche. Les performances de chaque configuration ont été évaluées à l'aide de l'erreur quadratique moyenne (EQM) et du coefficient de détermination (R^2). Les résultats ont montré qu'une augmentation de la complexité du réseau permettait d'améliorer la précision des prédictions. L'architecture la plus efficace, avec une configuration 6/12/12/6 (couche d'entrée/couche cachée 1/couche cachée 2/couche de sortie), a atteint une EQM faible de 0,00023. Ces résultats démontrent que les RNA sont des outils puissants et fiables pour la modélisation de processus de séparation complexes tels que l'extraction liquide-liquide.

Mots-clés : Modélisation, ANN, extraction L-L, équilibre de phase.

Table of contents :

List of figures:

List of tables:

General Introduction:

Liquid–liquid extraction (LLE) is an important separation process in chemical industries, such as petrochemicals, pharmaceuticals, and environmental engineering. This technique relies on distributing a solute between two immiscible liquid phases, allowing for efficient separation without the need for harsh operating conditions, such as distillation, especially when dealing with thermally sensitive or non-volatile compounds. This method is used in multiple applications, including industrial wastewater treatment and the extraction of bioactive compounds from biomass **[1]**.

Activity parameter models, such as NRTL and UNIQUAC, are used to characterize the distribution of components between phases in liquid–liquid extraction systems. These models are based on thermodynamic equations that require precise knowledge of the interaction parameters between components, making their use difficult in multicomponent systems or under suboptimal conditions **[2]**.

In recent years, artificial neural networks (ANNs) have emerged as an effective tool for modeling nonlinear and complex systems, such as liquid–liquid extraction systems. ANNs are characterized by their ability to learn from experimental data and discover hidden relationships between variables without the need for explicit physical models, making them suitable for modeling problems where data are available but theoretical knowledge is limited **[2].**

This study aims to develop a model using artificial neural networks to predict the component ratios in the raffinate and extract phases in a liquid-liquid extraction system, based on the composition of the feed phase. This research focuses on building and optimizing the neural network architecture to obtain accurate and reliable predictions, while identifying the best architecture that achieves the lowest mean squared error (MSE). The model's performance was evaluated by comparing it with experimental results reported in previous studies **[3].**

This thesis is divided into three main chapters. The first chapter covers the basic principles of the liquid-liquid extraction process. The second chapter reviews the general concepts of artificial neural networks, the model development methodology, and the technical details of the training and testing process. The third chapter analyzes and compares the results, as well as conclusions and future recommendations.

Chapter I:

Liquid-liquid extraction

I.1 Introduction:

Liquid-liquid extraction is a physicochemical separation process that has seen significant development, particularly in the nuclear, pharmaceutical, petrochemical, and oil industries.

It is based on the principle of transferring a substance or substances, known as the “solute,” between two immiscible liquid phases. **[4]**

The system consists of:

1. An aqueous solution of metallic cations, which, depending on the nature and composition of the aqueous phase, can exist as free ions, charged compounds, or neutral complexes.
2. An organic phase composed of an extractant or an extractant dissolved in a diluent.

To perform liquid-liquid extraction, two distinct operations must be carried out:

• The intimate mixing of the two phases by agitation.

• The separation of the two phases by decantation.

The duration of agitation is governed by the kinetics of solute transfer to reach equilibrium concentration, while the decantation time is determined by the separation time of the immiscible phases.

At equilibrium, the aqueous phase is referred to as the “raffinate” and the organic phase as the “extract.”

I.2 Definitions:

I.2.1 Solvent:

A solvent is an organic compound capable of forming combinations with the metallic solute that is soluble in the organic phase. It has physico-chemical properties that allow it to form a continuous organic phase, immiscible with the aqueous phase.

I.2.2 Extractant:

The extractant has the ability to form an organometallic compound with the metallic solute from the aqueous phase, which is soluble in the organic phase **[5,6].**

The criteria for using a good extractant are as follows:

• It is relatively inexpensive.

• It has low solubility in the aqueous phase.

• It does not form stable emulsions when mixed with the aqueous phase.

• It is highly stable, especially during continuous recycling.

• It has a high metal loading capacity.

• Its purification is easy after extraction.

• It has high solubility in aliphatic and aromatic diluents.

This product is inflammable, non-volatile, and non-toxic, and is dissolved or diluted in a hydrocarbon with high chemical inertness. Its physical properties are favorable for liquid flow and decantation.

I.2.3 Diluent:

A diluent is a compound that has no affinity for the solute to be extracted and has the major advantage of forming a continuous organic phase that is immiscible with an aqueous solution. It is generally used to solubilize extractants, dilute solvents, and, most importantly, to stabilize the physico-chemical properties of the organic phase.

I.3 Principle of Liquid-Liquid Extraction:

Liquid-liquid extraction is carried out by the intimate contact of the organic solvent with the aqueous solution in devices designed to mix the two phases, such as flasks, columns, and mixers. The separation of phases is achieved through gravitational decantation or centrifugation **(Figure I.1)**.

The transfer of the solute into the organic solvent theoretically leads to an equilibrium in the composition of the phases. Achieving this equilibrium is related to the diffusion rate of the solute from one phase to another.

The extraction rates are higher when the concentration differences of the solute between the two contacting phases are large, and when the surface area of exchange between the two phases is large. Agitation of the medium increases the contact surface between the phases and promotes the diffusion of the solute within each phase

The operational conditions for liquid-liquid extraction are chosen based on numerous parameters: temperature, concentration, pH, redox potential, and various solvents, which explains the wide range of applications for this process **[7].**

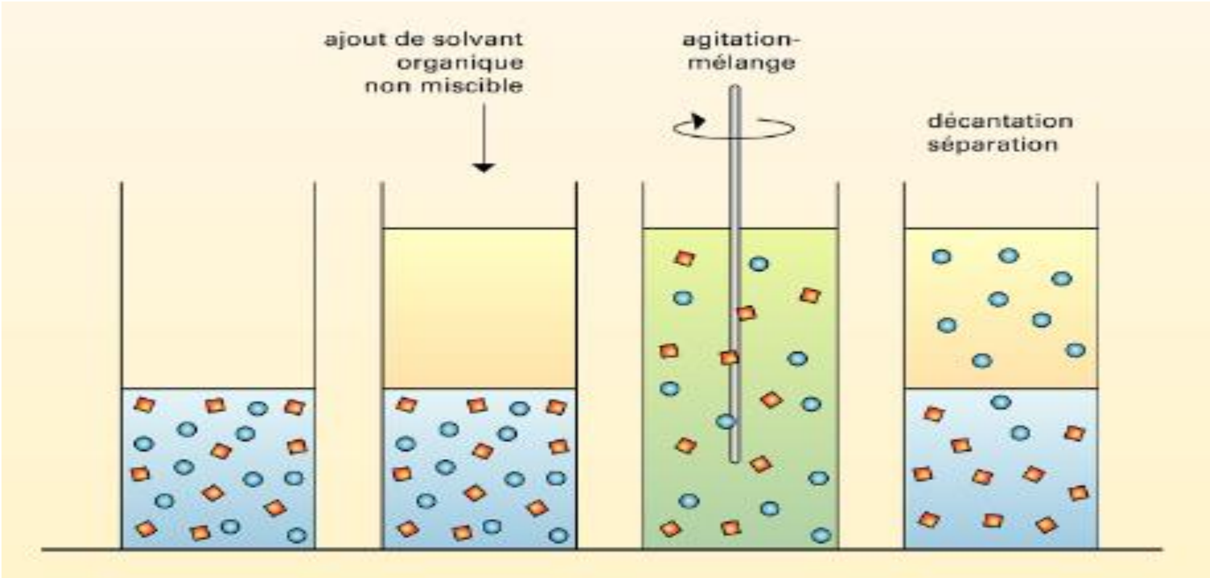

Figure I.1: Presentation of a liquid-liquid extraction.

To prform this extraction, two main steps are involved:

1. Intimate contact of the two liquid phases for a sufficient amount of time to reach equilibrium or a state close to equilibrium, during which the solute(s) are transferred from the aqueous phase (feed solution) to the organic solvent **(Figure I.2)**.
2. Subsequent separation of the two liquids (extract and raffinate) under the effect of natural gravity, to which additional forces may be applied in some cases, such as centrifugal force, electric field, etc.

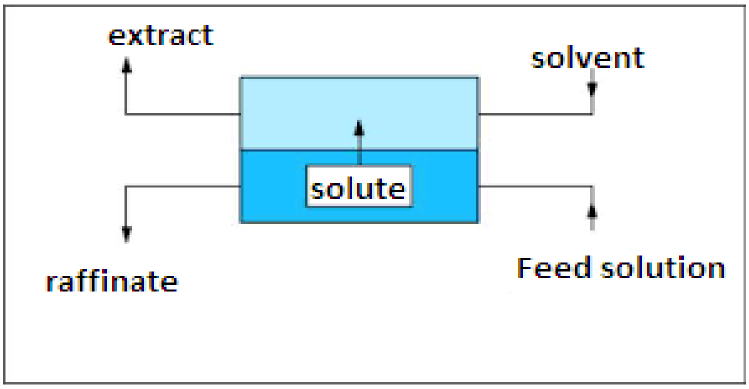

Figure I.2: Schematic diagram of liquid-liquid extraction.

All of these operations constitute one extraction stage. Generally, the previous operations are followed by a third operation (re-extraction or de-extraction), which involves regenerating the solvent and the extractant in order to recycle them.

I.4. The triangular diagram:

The triangular diagram **(Figure I.3)** is widely used for representing ternary systems. Since the effects of pressure and the vapor phase are neglected, the only factors to consider are temperature and composition.

The three vertices of the triangle represent the three components. The temperature axis is perpendicular to the plane of the triangle, resulting in a triangular prism shape, with the sides representing the binary systems A-B, B-C, and A-C. Thus, the triangle illustrates an isothermal section of the system. By drawing a line through point M parallel to AC, the percentage of component B remains constant at 70% along the entire line MN. Similarly, for component A, its percentage stays constant along a line parallel to BC. Therefore, point M corresponds to 10% of component A. As a result, the composition of the phase at point M is: 70% of B, 10% of A, and 20% of C **[8] .**

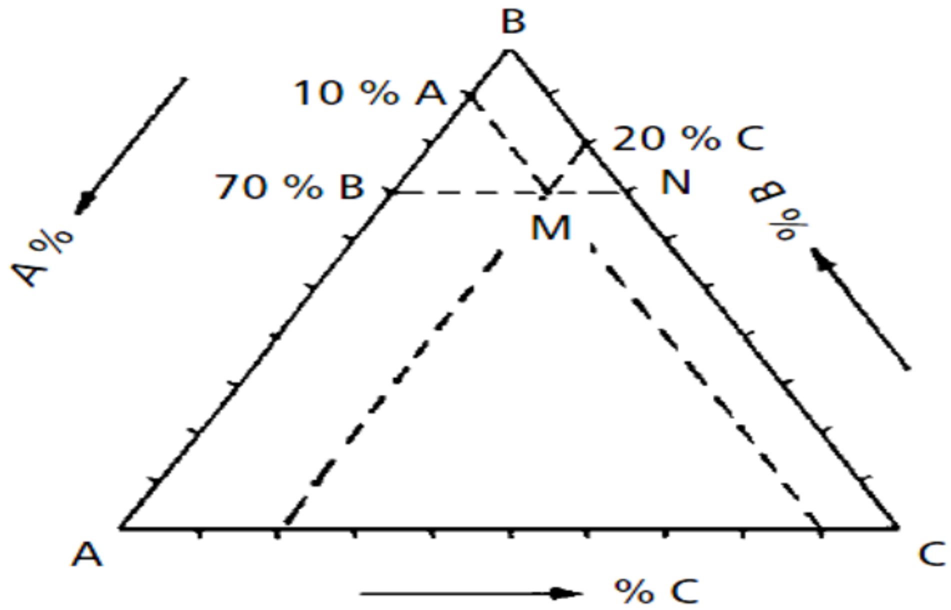

Figure I.3: Ternary system: triangular diagram **[8].**

I.4.1 Triangle diagram type:

At specific temperature and pressure conditions, the shape of the phase diagrams in liquid-liquid extraction depends on how the components of a ternary mixture dissolve in one another when considered pair by pair. The possible types of phase diagrams are as follows **(Figure I.4)**

All three components form a single homogeneous phase regardless of their proportions. One of the binary mixtures shows partial miscibility, creating a gap in solubility (Type I diagram).

Two of the binary mixtures exhibit a miscibility gap (Type II diagram). All three binary mixtures exhibit a miscibility gap (Type III diagram).

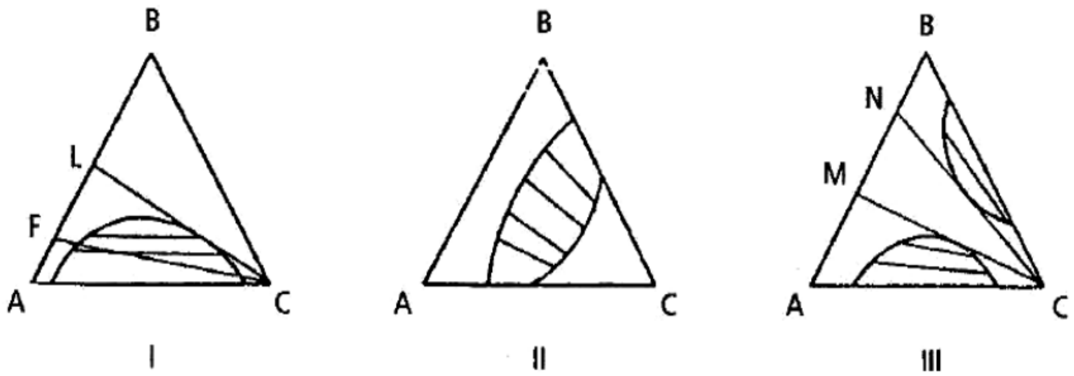

Figure I.4: Type of ternary systems.

When the three components of a mixture are miscible in all proportions, the system becomes unusable in liquid-liquid extraction. The most common case, represented by a Type I diagram, will be discussed here. To represent the system, an isothermal diagram is used, as solubility varies mainly with temperature and only slightly with pressure.

In the diagram shown in **Figure (I.5)**, component B is miscible in all proportions with both A and C, while A and C are not fully miscible with each other **[9].**

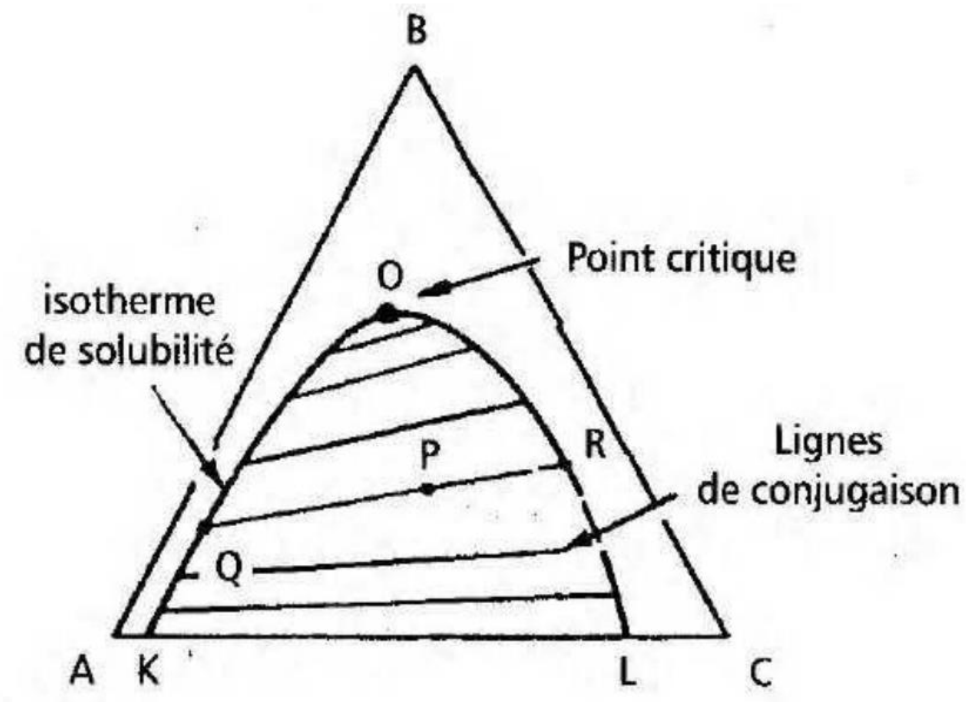

Figure I.5: Isothermal diagram of one of the binary mixtures showing a miscibility gap**[10].**

The bimodal curve, or solubility isotherm, divides the diagram into two regions: above the curve, the system consists of three phases, while below the curve, the system is homogeneous. In the immiscibility zone, if point P represents a specific state of the system, points Q and R represent the compositions of the two distinct phases, and these points are fixed. The line connecting Q and R is referred to as conjugate points, and there is a special point, O, called the critical point, which is its own conjugate.

The segment KO, corresponding to phases rich in the diluent, defines the composition of raffinate in liquid-liquid extraction, while the segment OL, related to phases rich in solvent, defines the composition of the extract.

Representing a system with four components requires a spatial diagram, such as a tetrahedral plot. For systems with more than four components, creating a complete graphical representation becomes highly complex. In practice, however, in many cases, it is possible to group these components based on their chemical properties, thus treating the complex mixture as a blend of a limited number of groups, simplifying the system to a ternary system**[10].**

I.4.2. Effect of Temperature on Diagrams:

In general, pressure has little impact on the properties of liquid phases, including the chemical potential or fugacity of the components (as long as we are far from the liquid-vapor critical point). Therefore, pressure is not a significant state variable for studying liquid-liquid equilibria, as long as it remains sufficient to ensure that no vapor phase is present. When dealing with liquid-liquid equilibria, the variance is often calculated as: **$v = c + 1 - \varphi$**, meaning it is reduced by one unit to account for the lack of pressure influence.

Temperature, on the other hand, has a moderate effect on phase diagrams: generally, it can be assumed that an increase in temperature tends to reduce the extent of the two-phase region. The rise in temperature promotes.

Certain parameters are often used to characterize the effectiveness of extraction **[11]**.

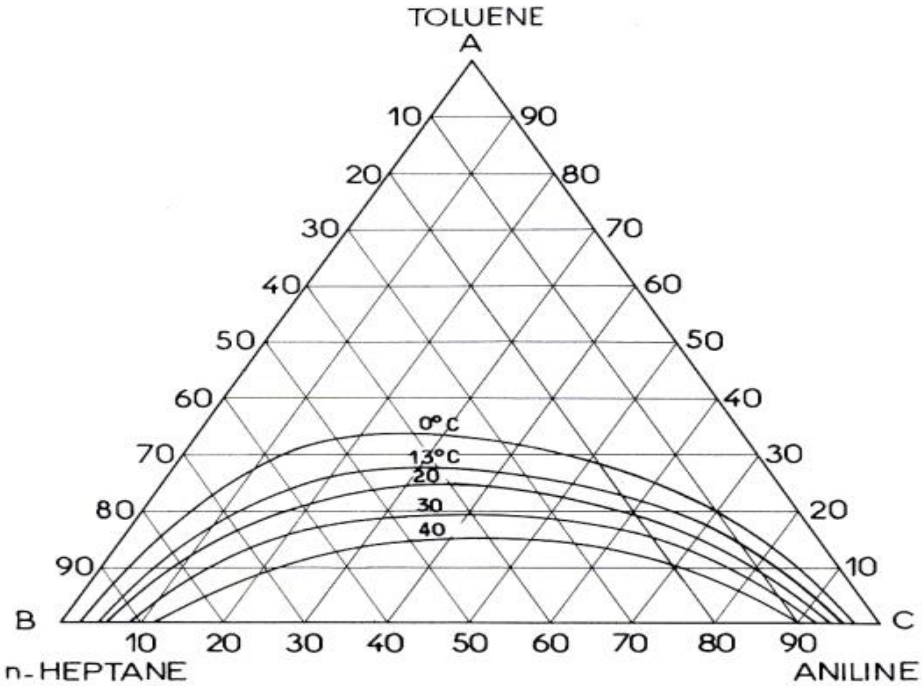

Figure I.6: The influence of temperature on the miscibility zone.

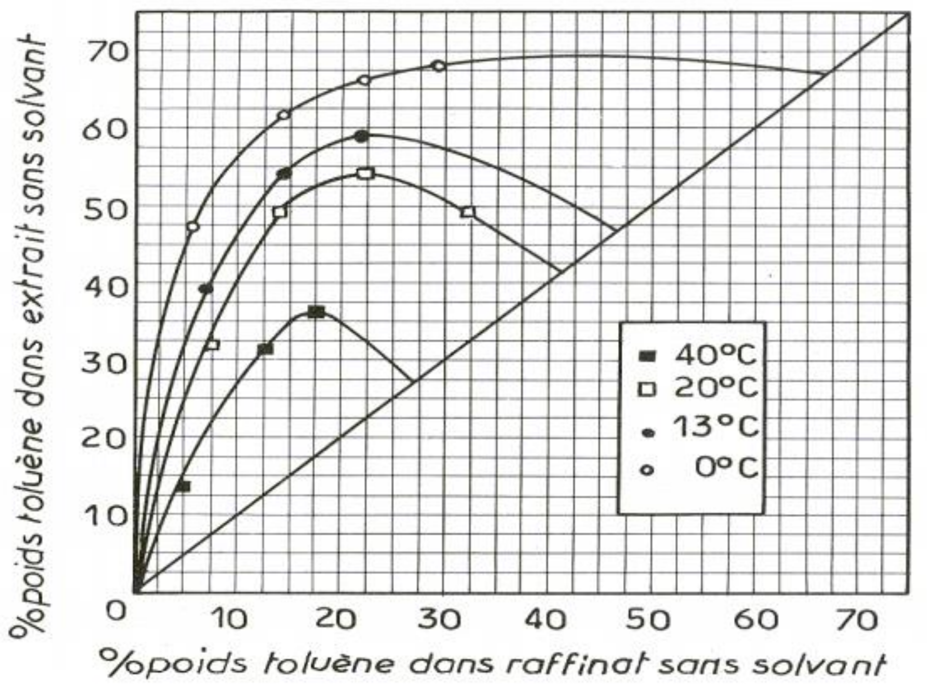

Figure I.7: The influence of temperature on the selectivity curve **[11]**.

I.5 Parameter followed:

Often certain parameters are used to determine the extraction efficiency.

I.5.1 Distribution or Partition Coefficient:

The distribution or partition coefficient "m" represents the ratio of the total solute concentration in the extract to the total solute concentration in the raffinate. The distribution coefficient is defined by the following equation:

$$m = \frac{c\prime}{c} = \frac{c_{in}v_{in} - cv}{cv\prime} \quad \textbf{(I.1)}$$

Where:

C' : the total solute concentration in the extract.

C : the total solute concentration in the raffinate.

V : total volume of the extract phase.

V' : total volume of the raffinate phase.

The distribution or partition curve is obtained by plotting, in rectangular coordinates for both phases (extract and raffinate), the variation in solute concentration in the solvent-rich phase as a function of the solute concentration in the solvent-poor phase. It is a curve that represents the equilibrium distribution of the solute between the two phases, extract and raffinate

1.5.2 Selectivity:

The selectivity curve is plotted in rectangular coordinates and represents the variation of the mass fraction of the solute in the extract relative to the mass fraction of the solute dissolved in the raffinate after removal of the solvent. The degree of solvent selectivity is defined by:

$$Sélectivité = \frac{Rapport\ entre\ les\ fractions\ massiques\ de\ solute\ et\ d'inerte\ dans\ l'éxtrait}{Rapport\ entre\ les\ fractions\ massiques\ de\ solute\ et\ d'inerte\ dans\ le\ raffinaat} \quad \textbf{(I.2)}$$

I.5.3 Separation Factor:

In cases where two solutes, 1 and 2, are to be separated, the selectivity of the solvent for solute 2 relative to solute 1 must also be taken into account. This selectivity is represented by the separation factor β*:

$$\boldsymbol{\beta}^* = \frac{c'_1/c'_2}{c_1/c_2} = \frac{m_1}{m_2} \quad \textbf{(I.1)}$$

Two solutes, 1 and 2, are more easily separable when their separation factor is high. Therefore, the choice of a solvent for liquid-liquid extraction is influenced, among other factors, by the need for satisfactory separation factors**[8]**.

I.6 Calculation of Liquid-Liquid Equilibrium:

Typically, the calculation of fugacities requires a heterogeneous method because the mixture laws applied to the equations of state cannot, at least in their classical form, be used for systems involved in liquid-liquid equilibria due to their limited flexibility and the polarity of certain components.

The liquid-liquid equilibrium condition can be written as follows:

$$x_i^E \gamma_i^{L,E} = x_i^R \gamma_i^{L,R} \quad \textbf{(I.3)}$$

In this way, it is possible to relate the "solvent power," "partition coefficient," and "selectivity" to the activity coefficients γiL.

By "solvent power," we refer to the solubility of a representative compound from a given family in the solvents studied, for example, hexane in polar solvents. The following equation, using the index 1 for hexane, the exponent R for the hexane-rich phase, and the exponent E for the solvent-rich phase, shows that the solubility of hexane is expressed by the relation:

$$x_1^E = \frac{x_1^R \gamma_1^{L,R}}{\gamma_1^{L,E}} \quad \textbf{(I.4)}$$

Moreover, if the reciprocal solubilities of hexane and the solvent are low, then the activity of hexane (equal to the product $x_1^R\ \gamma_1^{L,R}$) in the hexane-rich phase will be very close to 1, (as will the activity of the solvent in the solvent-rich phase), and we can write:

$$x_1^E \approx \frac{1}{\gamma_1^{L,E}} \quad \textbf{(I.6)}$$

There are many data in the literature on activity coefficients at infinite dilution, which, when they have a high value, correspond to the inverse of the solvent power.

The "partition coefficient" of a compound is the ratio of the concentrations in each phase; therefore, it is equal to the ratio of the activity coefficients:

$$\frac{x_i^E}{x_i^R} = \frac{\gamma_i^{L,R}}{\gamma_i^{L,E}} \quad \textbf{(I.7)}$$

Finally, the "selectivity" of a solvent with respect to two solutes is simply the ratio of the partition coefficients or, in other words, the ratio of the concentrations of the two solutes in the extract phase and the raffinate phase.

$$\alpha_{ij} = \frac{x_i^E / x_i^R}{x_j^E / x_j^R} = \frac{x_i^E / x_j^E}{x_i^R / x_j^R} = \frac{\gamma_j^{L,E} / \gamma_i^{L,E}}{\gamma_j^{L,R} / \gamma_i^{L,R}} \quad \textbf{(I.8)}$$

In the case where the raffinate phase is primarily composed of the two solutes, and their activity coefficients in this phase are close to unity, we can then write:

$$\alpha_{ij} = \frac{\gamma_j^{L,E}}{\gamma_i^{L,E}} \quad \textbf{(I.9)}$$

However, this is only a convenient approximation. For example, in the separation of aromatics and paraffins, hydrocarbon mixtures are far from ideal. Nevertheless, this approximation, using the activity coefficients at infinite dilution for a paraffin and an aromatic in polar solvents, allows for the establishment of a selectivity scale that can assist in the selection of appropriate solvents**[12]**.

I.7. Activity Coefficient Models:

The ability to correlate and predict equilibrium data is crucial in the design of separation equipment in the chemical industry. Many studies have been conducted to develop thermodynamic models that provide reliable methods for handling non-electrolyte mixtures under normal temperature and pressure conditions

.One of the key parameters in phase equilibrium calculations is the activity coefficient, and many models have been proposed for its calculation, including the NRTL model, the UNIQUAC model, and the UNIFAC model **[13]**.

I.7.1. NRTL Model:

The NRTL (Non-Random Two Liquids) model was introduced in 1968 by RENON and PRAUSNITZ. This model is based on expressing the internal energy of the mixture through the concept of local compositions. According to this concept, in a mixture, around each molecule i, there are molecules of the same type i as well as molecules of different types j. The arrangement of these molecules is not necessarily uniform; for instance, polar molecules tend to cluster together, excluding non-polar molecules. Therefore, the local composition is represented by xij, where j is the center of attraction.

The activity coefficient is given by:

$$\ln\gamma_i = \frac{\sum_j \tau_{ji} G_{ji} x_j}{\sum_k G_{ki} x_k} + \sum_i \frac{x_j G_{ij}}{\sum_k G_{ki} x_k}\left[\tau_{ij} - \frac{\sum_r x_r \tau_{rj} G_{rj}}{\sum_k G_{kj} x_k}\right] \qquad \textbf{(I.10)}$$

With :

$$\tau_{ij} = \left(\frac{g_{ji} - g_{ii}}{RT}\right) = \left(\frac{\Delta g_{ij}}{RT}\right) \qquad \textbf{(I.11)}$$

Or :

$$g_{ij} = g_{ji} \text{ donc } \tau_{ij} \neq \tau_{ji} \qquad \textbf{(I.12)}$$

And :

$$G_{ji} = exp(-\alpha_{ji} . \tau_{ji}) \qquad \textbf{(I.13)}$$

Or :

$$\alpha_{ji} = \alpha_{ij} = \alpha \qquad \textbf{(I.14)}$$

Δgij is a variable parameter that represents the interaction energy between components i and j, while xi denotes the molar fraction of component i. Additionally, αij is an empirical parameter that describes how species i and j are likely to avoid random association.

Typically, the value of the α parameter is set to a fixed value, often 0.2, 0.3, or 0.5, depending on the specific components in the mixture.

I.7.2. UNIQUAC Model:

The UNIQUAC (Universal Quasi-Chemical) model was introduced by Abrams and Prausnitz in 1975. This model calculates the excess Gibbs free energy by combining two terms: the combinatorial term, which takes into account the molecular size and shape in the mixture, and the residual term, which reflects the interactions between the molecules.

The expression is as follows:

$$g^E = g^{E,comb} + g^{E,res}\ \alpha \quad \text{(I.15)}$$

With :

$$g^{E,comb} = \sum_i x_i ln\left(\frac{\varphi_i}{x_i}\right) + \frac{Z}{2} ln\left(\frac{\varphi_i}{x_i}\right) \quad \text{(I.16)}$$

$$g^{E,res} = -\sum_i q_i x_i\, ln\left(\sum_j \theta_j \tau_{ji}\right) \quad \text{(I.17)}$$

Φi and θi denote the volume and surface fractions of molecule i, which are determined based on the structural parameters ri and qi. These parameters depend on the volume and surface area of the molecules, respectively, and are expressed as:

$$\varphi_i = \frac{r_i x_i}{\sum_i r_i x_i} \quad \text{(I.18)}$$

$$\theta_i = \frac{q_i x_i}{\sum_i q_i x_i} \quad \text{(I.19)}$$

Z is the coordination number, equal to 10 in the UNIQUAC model, and l_i is a parameter specific to component i, given by the following expression:

$$l_i = \frac{Z}{2}(r_i - q_i) - (r_i - 1) \quad \text{(I.20)}$$

The parameters r_i and q_i are related to the molecular structure. They are obtained by summing the surface R_k and volume Q_k parameters of each functional group in the molecule:

$$r_i = \sum_k v_k^i R_k \quad \text{(I.21)}$$

$$q_i = \sum_k v_k^i Q_k \quad \text{(I.22)}$$

Where v_k^i is the number of functional groups of type k in molecule i. The parameters R_k and Q_k are obtained from the volumes V_k and the surface R_k areas of the Van der Waals radii (Bondi, 1964),

$$\left(R^k = {}^{V_k}\!/_{15.17}\,;\, Q^k = {}^{A_k}\!/_{2.5}\right) \quad \text{(I.23)}$$

given by the following relations:

$$\tau_{ij} = exp\left(-\frac{u_{ij} - u_{ii}}{RT}\right) = exp\left(-\frac{A_{ij}}{RT}\right) \quad \text{(I.24)}$$

$$u_{ij} - u_{ii} = \Delta u_{ij} = a_{ij} + b_{ij} T + c_{ij} T^2 \quad \text{(I.25)}$$

Where *a* ,*b*, *c* and *u* are the parameters of UNIQUAC Model**[13]** .

Chapter II:

Artificial Neural Networks (ANN)

II.1 Introduction:

We begin this chapter with some basic definitions of neural networks, and then present the most commonly used learning methods.

Neural networks are a data processing technique that will soon be part of the toolbox of any engineer concerned with extracting the most relevant information from their data: developing models, recognizing patterns or signals, controlling systems, making predictions, etc **[14].**

The concept of artificial neural networks was invented in 1943 by two researchers at the University of Chicago: neurophysicist Warren McCullough and mathematician Walter Pitts. In an article published in the journal Brain Theory, the two researchers presented their theory that neural activation is the basic unit of brain activity. In 1957, the Perceptron was invented. It is the oldest machine learning algorithm, designed to perform complex pattern recognition tasks. It was this algorithm that would later enable machines to learn to recognize objects in images. It wasn't until the early 2010s, with the rise of big data and massively parallel processing, that data scientists had the data and computing power necessary to run complex neural networks. In 2012, during a competition organized by ImageNet, a neural network managed for the first time to outperform a human in image recognition. This is why this technology is once again a central concern for scientists. Today, artificial neural networks continue to improve and evolve every day **[15].**

II.2. Artificial Intelligence Concepts:

II.2.1 Definition of Artificial Intelligence:

Today, people mostly learn about artificial intelligence (AI) through the media, films, and its growing presence in everyday life, as illustrated in **Figure II.1.**

A widely recognized, though relatively early, definition of AI was proposed by John McCarthy during the 1956 Dartmouth Conference. He described artificial intelligence as the attempt to enable machines to simulate human intelligence as accurately as possible. However, this definition seems to overlook the concept of strong AI, which refers to machines that possess the ability to reason and solve problems independently.

Before we explore what “artificial intelligence” truly means, it’s helpful to first understand the concept of “intelligence.”

According to the theory of multiple intelligences, human intelligence can be divided into seven types: linguistic, logical-mathematical, spatial, bodily-kinesthetic, musical, interpersonal, and intrapersonal intelligence.

- **Linguistic Intelligence:**

  Linguistic intelligence involves the ability to clearly express ideas through spoken or written language, to understand the language of others, and to skillfully use the phonology, semantics, and grammar of a language **[16]**.

Table.II.1: Social cognition of AI **[17]**.

Haidian park , the World is first AI park! Al program Defented Top Human Pmayers at StarCraft. AlphaStar Gained fame! Portrait by AI Program Portrait of Edmond Belamy Sells for 430,000$ al programmer Demand Skyrocketed 35 Times !Salary Ranked 0No.1! 50%od the Jobs Will be Replaced by AI in the Future The Winter is Coming ?AI Faces Big Challenges ……	The Terminator 2001:A Space Odyssey The Matrix I,Robots Blade Runner Her Bicentennial Man …..	Self-service security screening Speaking skills assessment Movie and misic recommendation Smart loudspeaker Robot vacuums Bank self-service terminal Smart service Siri ……
News	Movies	Daily Application
Application of AI Industry trends and outlook for AI Challenges of AI …..	AI controls humans Fall in live with AI Self-consciousness of AI …..	Security & protection Entertainment Smart home Finance …..

Verbal thinking refers to the ability to process and express ideas through language. It also involves interpreting the deeper meaning or connotation behind spoken or written expressions. Individuals with strong linguistic intelligence often excel in roles that require effective communication. Ideal career paths include politician, activist, TV host, lawyer, public speaker, editor, writer, journalist, and teacher.

- **Logical-Mathematical Intelligence:**

Logical-mathematical intelligence refers to the ability to calculate, quantify, reason logically, analyze information, and solve complex mathematical problems. This type of intelligence involves a strong sensitivity to abstract concepts such as logical patterns, relationships, statements, and functions. People who excel in this area are often suited to careers like scientist, accountant, statistician, engineer, or software developer.

- **Spatial Intelligence:**

Spatial intelligence is the ability to accurately perceive and interpret the visual-spatial world and to mentally manipulate or recreate those perceptions through drawings or models. Those with high spatial intelligence are particularly sensitive to elements like color, shape, line, form, and spatial relationships. Careers suited to these individuals include architect, interior designer, photographer, painter, pilot, and other visual-based professions.

- **Bodily-Kinesthetic Intelligence:**

Bodily-kinesthetic intelligence involves the use of one's body to express ideas and emotions or to create and manipulate objects skillfully. This includes fine and gross motor skills such as balance, coordination, strength, agility, and dexterity. Individuals with strong bodily-kinesthetic intelligence often thrive in professions like athlete, actor, dancer, surgeon, jeweler, or mechanic.

- **Musical Intelligence:**

Musical intelligence is the ability to recognize and create musical elements such as pitch, tone, melody, rhythm, and timbre. People with high musical intelligence have a heightened sensitivity to sound patterns and musical structures, allowing them to excel in activities such as composing, performing, or analyzing music. Careers in this area include musician, composer, music teacher, conductor, or sound engineer **[17]**.

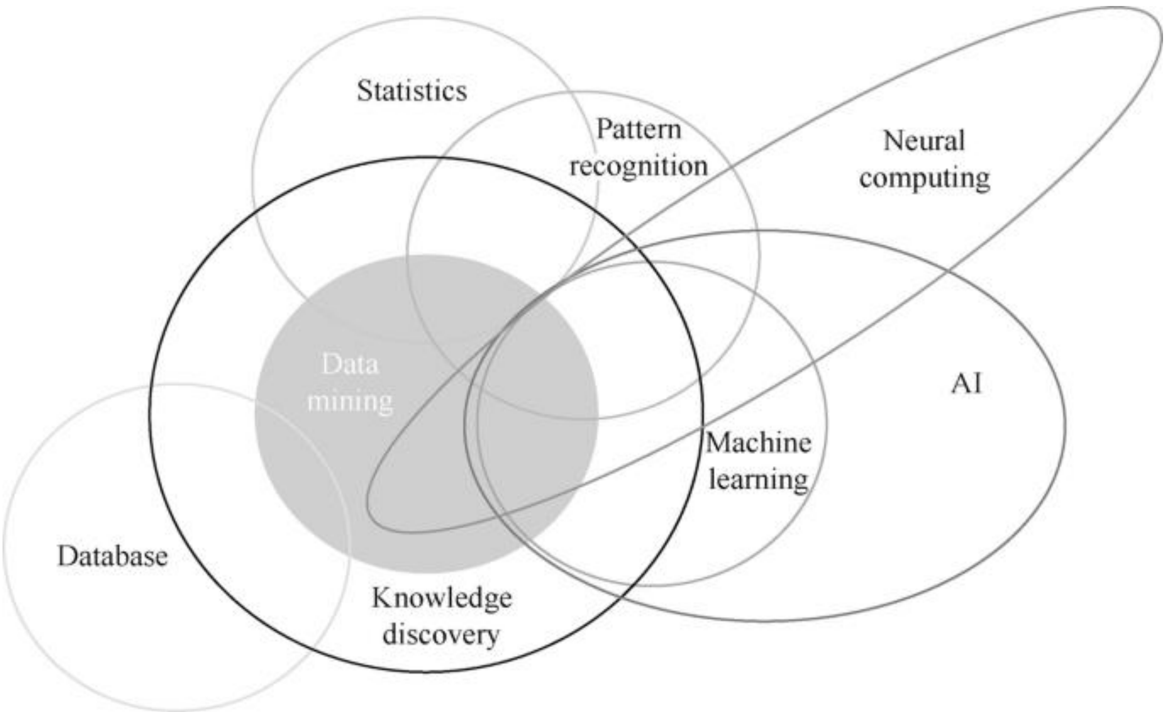

Figure.II.1: Fields covered by artificial intelligence **[17]**.

They are skilled at performing, composing, and analyzing music. Recommended careers for individuals with strong musical intelligence include singer, composer, conductor, music critic, piano tuner, and similar roles.

- **Interpersonal Intelligence:**

Interpersonal intelligence is the ability to understand and interact effectively with other people. Individuals who possess this form of intelligence are highly attuned to others' emotions, moods, and motivations. They can empathize with others, pick up on subtle social cues, and navigate relationships with sensitivity and awareness. Careers suited to people with strong interpersonal intelligence include politician, diplomat, leader, psychologist, public relations officer, salesperson, and similar professions.

- **Intrapersonal Intelligence:**

Intrapersonal intelligence is the capacity for self-awareness and self-reflection. It involves understanding one's own emotions, strengths, weaknesses, values, and motivations, and using this insight to guide behavior and decision-making. People with strong intrapersonal intelligence tend to enjoy deep thinking and self-exploration. Ideal careers include philosopher, politician, thinker, psychologist, and other roles that require introspection and independent thought.

- **Naturalist Intelligence**

Naturalist intelligence is the ability to recognize, categorize, and draw connections between different elements in nature. It involves keen observation of the natural world and the skill to identify and differentiate between flora, fauna, weather patterns, and other natural phenomena. It also includes an understanding of the interaction between natural and artificial systems. Careers in biology, environmental science, farming, botany, and wildlife conservation are well-suited for those with strong naturalist intelligence **[16]**.

II.2.2 Artificial Intelligence as a Technological Discipline :

Artificial Intelligence is a modern technological science focused on the study and development of theories, methods, technologies, and application systems that simulate, enhance, and extend human intelligence. AI aims to equip machines with the ability to reason, learn, and solve problems in a manner similar to humans.

Over time, the definition and scope of AI have expanded significantly. It has evolved into a highly interdisciplinary field, drawing from computer science, cognitive science, neuroscience, mathematics, linguistics, psychology, and more. This integration is illustrated in **Figure. II.1** high lighting AI's broad and interconnected nature.

- **Machine Learning (ML):**

Machine learning (ML) is widely recognized as one of the core areas within the broader interdisciplinary field of artificial intelligence.

According to Tom Mitchell—often referred to as "the godfather of global ML"—machine learning is defined as follows:

A computer program is said to learn from experience (E) with respect to some class of tasks (T) and performance measure (P), if its performance at tasks in T, as measured by P, improves with experience E.

Though relatively simple and abstract, this definition captures the core idea of machine learning.

As our understanding of machine learning evolves, both its connotation (implied meaning) and denotation (literal definition) continue to develop. Defining machine learning in just one or two sentences proves difficult—not only because it spans a wide range of theories and applications, but also because it is a rapidly evolving field.

In general, machine learning systems and algorithms work by analyzing data and identifying hidden patterns in order to make predictions or decisions. It is a vital subfield of artificial intelligence and is closely linked with other areas such as data mining (DM) and knowledge discovery in databases (KDD), forming a broader ecosystem of intelligent data-driven technologies **[17]**.

II.3. The Relationship Between AI, Machine Learning, and Deep Learning:

The field of machine learning focuses on enabling computers to replicate human learning abilities—acquiring new knowledge and developing new skills through data-driven experiences.

Deep learning (DL), a more recent and advanced branch of machine learning, originates from the study of artificial neural networks (ANNs). It is specifically designed to imitate the way the human brain processes information, allowing machines to interpret complex data such as images, audio, and text more effectively.

The hierarchical relationship among artificial intelligence (AI), machine learning (ML), and deep learning (DL) is typically visualized as nested subsets—where deep learning is a part of machine learning, which in turn is a subset of the broader field of AI. This relationship is illustrated in **Figure. II.2**.

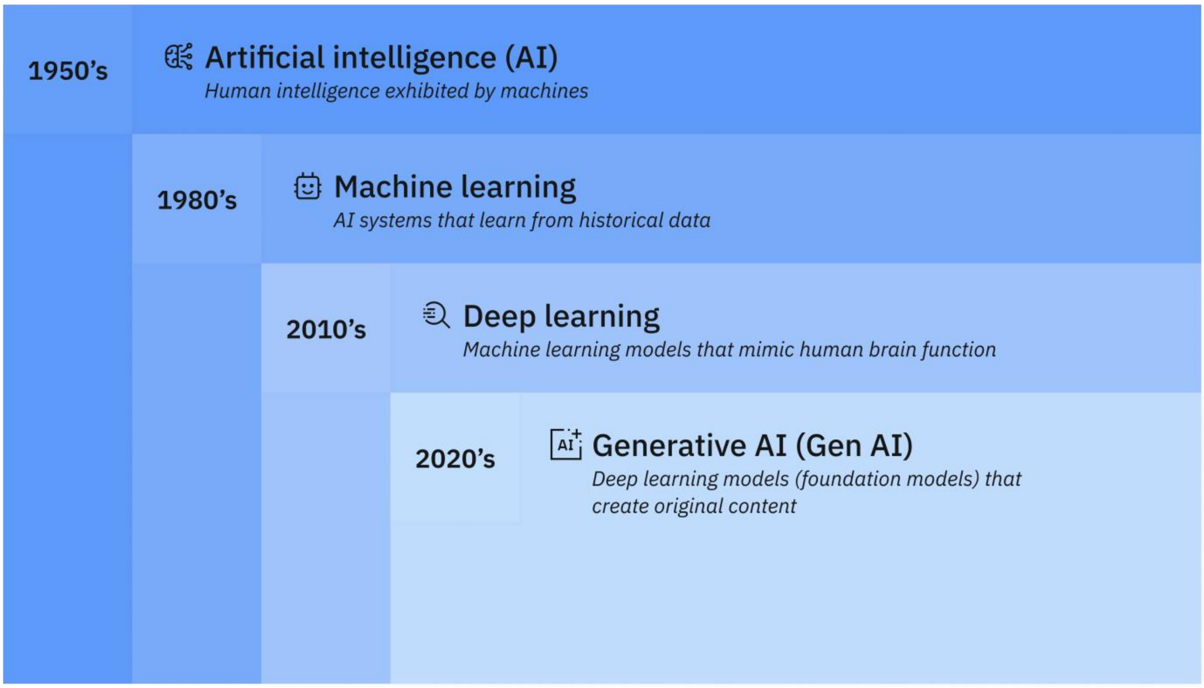

Figure.II.2: The relationship between artificial intelligence, machine learning, deep learning and generative AI **[18].**

Among the three concepts, machine learning is an approach within the broader field of artificial intelligence, while deep learning is a specialized subset of machine learning.

To put it metaphorically: if AI represents the brain, then machine learning is the process of acquiring cognitive abilities, and deep learning is the highly efficient, self-training system that powers and refines this learning process.

In essence, artificial intelligence is the ultimate goal and outcome, while machine learning and deep learning are the methods and tools used to achieve it **[18]**.

Generative AI is a type of artificial intelligence that focuses on creating new content, such as text, images, music, and videos, rather than just analyzing or recognizing existing data. It leverages large AI models to learn patterns and structures from vast datasets and then generates new content based on that learning.

II.3.1. Types of AI:

Artificial intelligence can generally be classified into two types: strong AI and weak AI. Strong artificial intelligence refers to the theoretical possibility of creating intelligent machines capable of true reasoning and problem-solving. These machines are believed to possess consciousness, self-awareness, and the ability to think independently, generating optimal solutions to complex problems. Strong AI is envisioned as having its own set of values and worldview, as well as instinctual drives—such as the need for survival and safety—similar to those of living beings. In a broader sense, strong AI could be seen as the emergence of an entirely new form of civilization.

Weak artificial intelligence, on the other hand, describes systems that do not truly possess intelligence or self-awareness. These machines may appear intelligent, but they lack the ability to reason or understand. Instead, they are designed to perform specific tasks efficiently, without general cognitive capabilities.

At present, we are still in the era of weak AI. This type of AI supports and enhances intellectual work by mimicking certain human functions—much like advanced bionic systems. Examples of weak AI include AlphaGo, or AI programs capable of writing news articles or novels. These systems can surpass human performance in narrowly defined areas, but they do not possess general intelligence.

In the current age of weak AI, data and computing power are critical drivers that accelerate its development and commercialization. Looking ahead, these two factors will remain essential in the eventual transition to strong AI. Moreover, ongoing research in quantum computing by companies such as Google and IBM is laying the groundwork for the potential emergence of strong AI in the future**[18]**.

II.4. Machine Learning:

Machine learning (including its subfield, deep learning) is the study of learning algorithms. In this context, "learning" means that if the performance of a computer program on a specific task T, as measured by a performance metric P, improves with experience E, then we say that the program is learning from that experience.

For example, consider the task of identifying spam emails this is our task T. As humans, we can perform this task quite well because we have plenty of relevant experience from everyday life, such as reading emails, encountering spam messages, or seeing misleading advertisements. Based on this experience, we can infer that emails from unknown senders containing words like "discount" or "zero risk" are more likely to be spam.

Using this learned knowledge, we can often determine whether a new, unread email is spam (as illustrated in **Figure.II.3.a**). But can we design a computer program to simulate this process?

As shown in (**Figure. II.3.b)**, we can start by collecting a large number of emails and manually labeling which ones are spam. This labeled data serves as the experience E for the program. However, unlike humans, the program cannot automatically extract meaningful

patterns from the data on its own. That's where machine learning algorithms come in—they are used to train the program to learn from the labeled data.
Once trained, the program becomes a model. Generally, the more data we use for training, the better the model performs. As a result, the value of the performance metric P increases, meaning the program is better at accurately identifying spam emails.

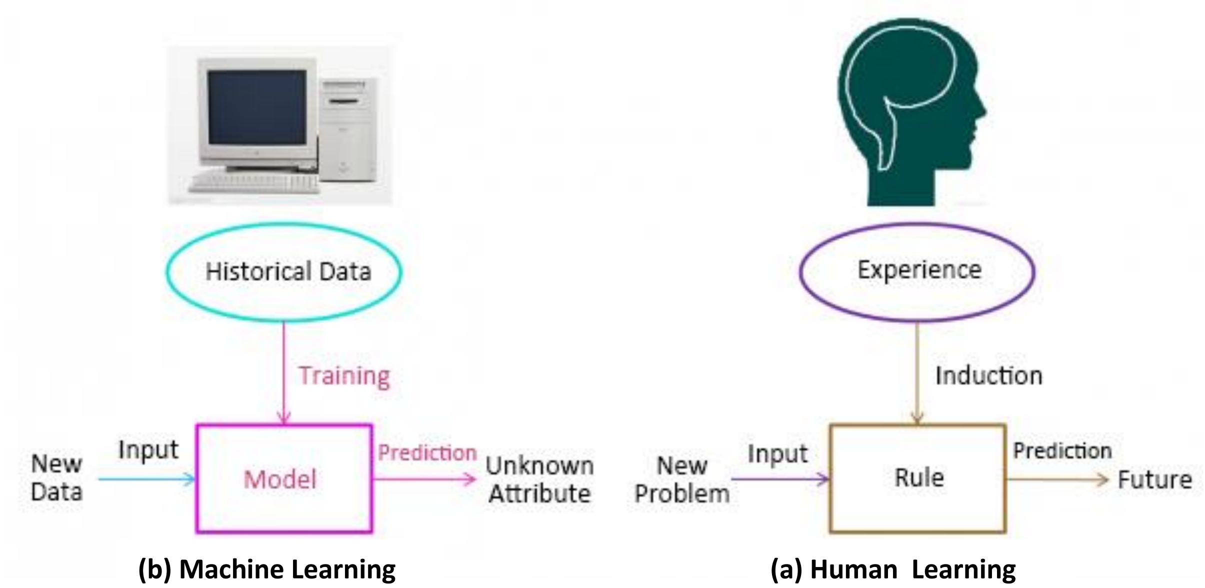

Figure.II.3: Learning mode**[19]**.

II.4.1. Challenges in Spam Identification and the Role of Machine Learning:

Identifying spam emails is extremely difficult using traditional programming methods. In theory, one could attempt to create a set of explicit rules that distinguish spam emails from regular ones. This is known as a rule-based approach. However, in practice, it is nearly impossible to formulate a complete and consistent set of rules that accurately captures all the features of spam emails.
To address this limitation, machine learning adopts a statistical-based approach. Instead of explicitly programming rules, machine learning enables computers to automatically learn patterns and rules from a collection of labeled samples (data). In essence, machine learning allows machines to learn from experience and discover complex rules that would be difficult or even impossible to define manually.
Compared to rule-based methods, machine learning can handle more complex, flexible, and subtle patterns in data, making it suitable for a wider range of problems **[17]**.

II.4.2. When to Use Machine Learning :

Although machine learning is a powerful and versatile tool in the AI field, it is not always the most efficient solution. As illustrated in **Figure.II.3**, machine learning is best suited for problems that:

- Require complex solutions.
- Involve large volumes of data.
- Have an unknown or changing probability distribution.

In contrast, for simpler problems especially those where effective rules can be easily written traditional rule-based methods may be more cost-effective and efficient. For instance, in scenarios that fall into the second quadrant of **Figure.II.3** (i.e., small-scale problems solvable by hand-coded rules), using machine learning may introduce unnecessary complexity and cost.

II.4.3. Common Application Scenarios for Machine Learning:

There are two primary situations where machine learning is especially useful:

1. The rules are too complex or cannot be clearly defined, such as in:
   - Face recognition
   - Speech recognition
2. The data distribution changes over time, and the system needs to continuously adapt, as in:
   - Predicting commodity sales trends
   - Stock market forecasting
   - User behavior modeling

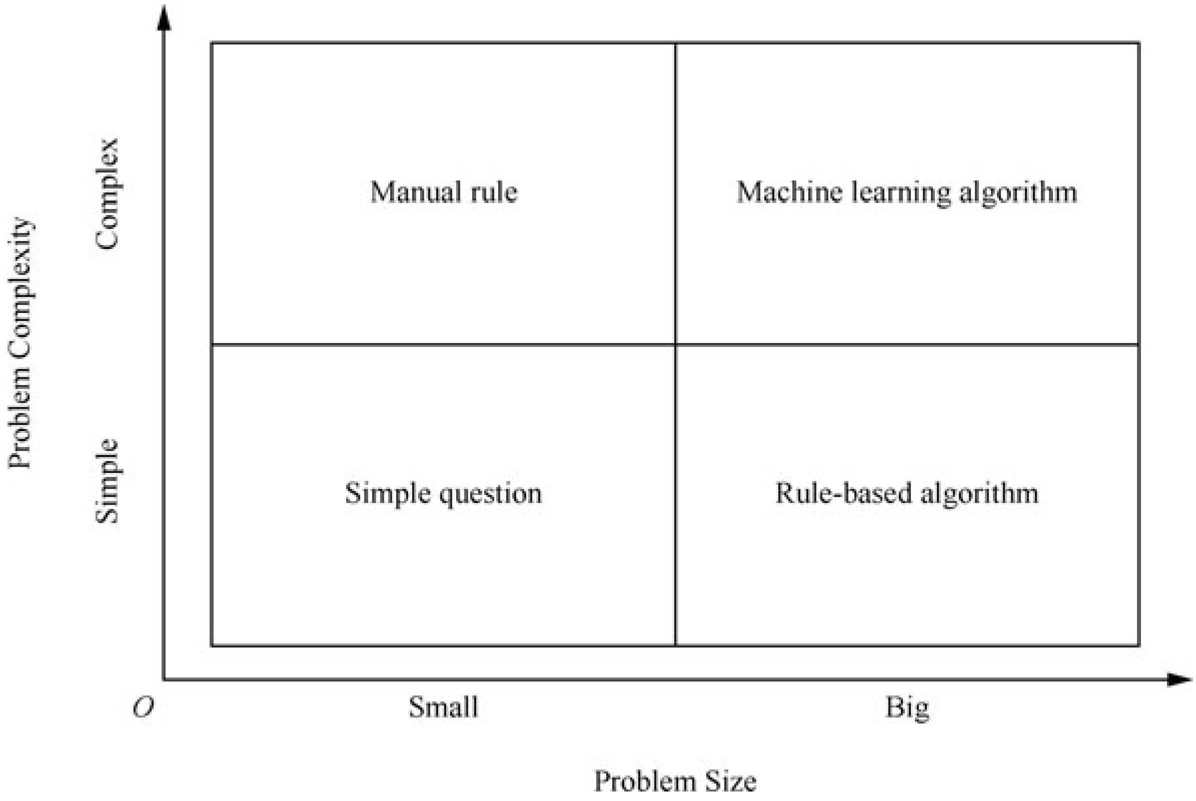

Figure.II.4: Application scenarios of machine learning **[17]**..

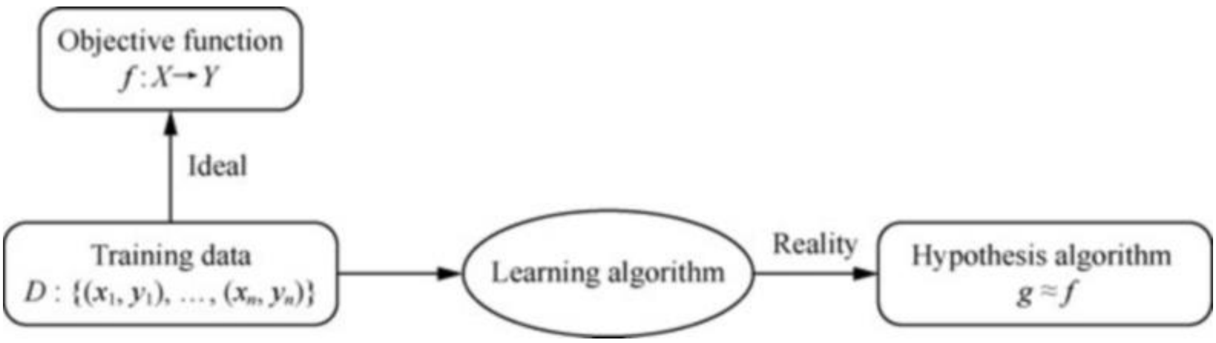

Figure.II.5: The relationship between the hypothesis function and the objective function **[17]**.

II.4.4. Major Problems Solved by Machine Learning:

Machine learning is capable of addressing various types of problems, with the most common being classification, regression, and clustering. Among these, classification and regression are the two main types of prediction problems, accounting for about 80–90% of all machine learning applications.

- **Classification:**

In a classification problem, the goal is to determine which of the predefined categories (classes) a given input belongs to. The output is discrete, typically represented as labels such as {1, 2, ..., k}.

Machine learning algorithms designed for classification tasks generate a mapping from an input domain D to a set of category labels. A classic example is image classification, where the program assigns an image to one of several categories (e.g., "cat", "dog", "car").

- **Regression:**

Regression involves predicting a continuous value based on input data. In this case, the algorithm learns a mapping from domain D to the real number domain R.

Examples of regression problems include:

- Predicting insurance claim amounts to determine premiums.
- Forecasting stock prices or commodity values.

Interestingly, classification problems can sometimes be reframed as regression problems—for instance, by predicting the probability that an image belongs to each class, then choosing the class with the highest probability.

- **Clustering:**

Clustering is an unsupervised learning task where the goal is to group data into clusters based on inherent similarities within the data—without using pre-assigned labels.

Unlike classification, clustering does not rely on labeled datasets. Instead, clustering algorithms aim to:

- Maximize similarity within the same group (cluster),
- Minimize similarity between different groups.

Common use cases for clustering include:

- Image retrieval (e.g., grouping visually similar images),
- User profiling or audience segmentation (e.g., generating “user portraits” for personalized recommendations) **[17]**.

II.5. Definition Neural Network:

An Artificial Neural Network (ANN) is a signal or information processing system made up of numerous simple processing units called neurons. These units are interconnected via direct links and work together to perform parallel and distributed processing to solve complex computational tasks.

Neural networks are designed to process information in a manner similar to the human brain. Inspired by biological neural systems, such as those found in the brain, ANNs are capable of learning by example—that is, they improve their performance by being exposed to data and adjusting their internal connections accordingly **[21]**.

II.5.1. Applications:

Neural networks are now being applied across a wide range of fields. For instance, Google uses these networks to classify images and automatically assign keywords, while Microsoft has developed technologies to convert speech from English to Chinese. These applications are expanding due to the ability of neural networks to solve problems in many areas, including engineering, medicine, business, finance, and science. Additionally, advancements in computers and faster algorithms have made it possible to use neural networks to tackle complex industrial problems that previously required extensive computation **[21]**.

II.5.2. Biological Inspiration:

In this section we will briefly describe the characteristics of brain function that have inspired the development of artificial neural networks.
The brain consists of a large number of highly connected elements called neurons.
These neurons have three principal components: the dendrites, the cell body and the axon as shown in **Figure.II.6**.

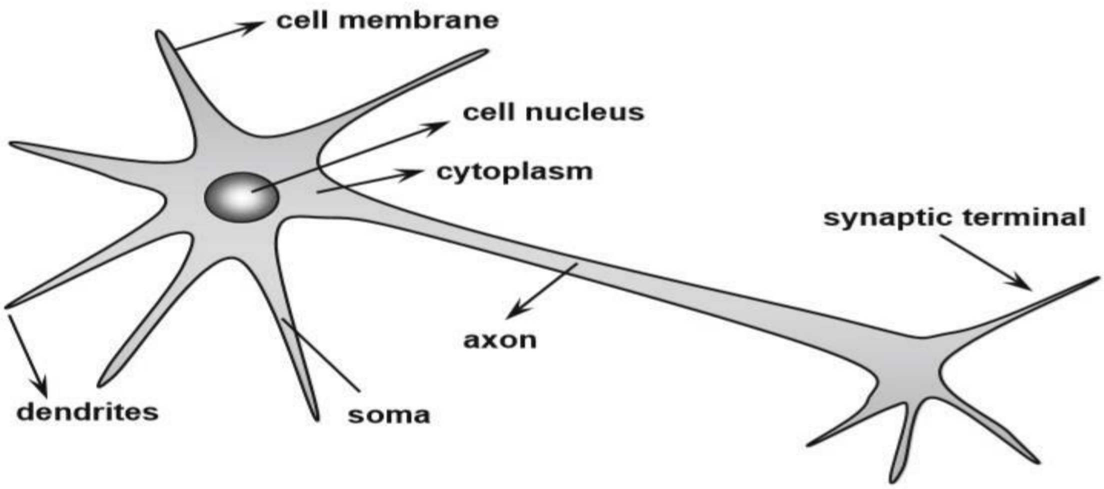

Figure.II.6: Biological Neuron**[21]**.

The dendrites are tree-like receptive networks of nerve fibers that carry electrical signals into the cell body.

The cell body effectively sums and thresholds these incoming signals.

The axon is a single long fiber that carries the signal from the cell body out to other neurons.

The point of contact between an axon of one cell and a dendrite of another cell is called a synapse. It is the arrangement of neurons and the strengths of the individual synapses, determined by a complex chemical process, that establishes the function of the neural network.

The synapses are the connections which enable the transfer of electric axon impulses from a particular neuron to dendrites of other neurons, as illustrated in **Figur.II.7 [21]**.

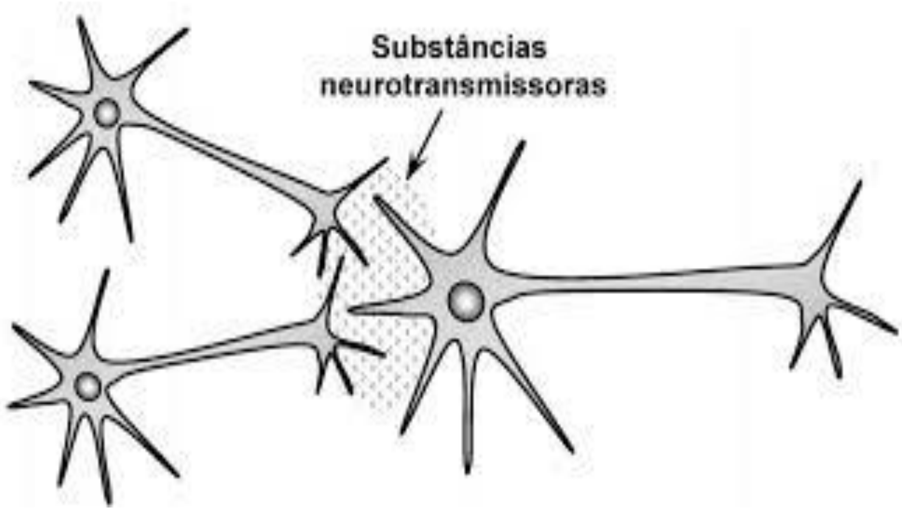

Figure.II.7: The synaptic connection between neurons **[21]**.

II.5.3. Concept of Artificial Neural Networks:

An artificial neuron can be viewed as a unit comprising inputs, an output, and a transfer function. Neurons are interconnected, allowing for the exchange of information through these connections. These neurons are organized into layers, with each layer forming a group that performs its own specific operation independently of the other groups. Typically, each neuron is connected to every neuron in the previous and next layers, except for the input and output layers of the network. Information is passed through the network from the input to the output, layer by layer. The network is structured using a learning algorithm, which calculates the parameters (such as weights associated with the various connections).

Each neuron, apart from those in the input layer, consists of two main components:
• A summation unit that adds the input signals after multiplying them by the corresponding weight matrix Wi.
• An activation function that receives two signals: the output of the summation unit and the “bias” value.

Mathematically, the input to neuron i is expressed as:

$$e_i = \sum_{j=1}^{n} W_{ij}.X_i \tag{II.1}$$

Where :

e_i: Global input to neuron i

X_i: Input signals to neuron i

W_{ij}: Weight associated with the signal connecting neuron i to neuron j in the previous layer.

And its output after transformation is written as:

$$y_i = f(e_i - \theta) \tag{II.2}$$

θ: Threshold or activation level of the neuron.

f: Activation function (transfer function).

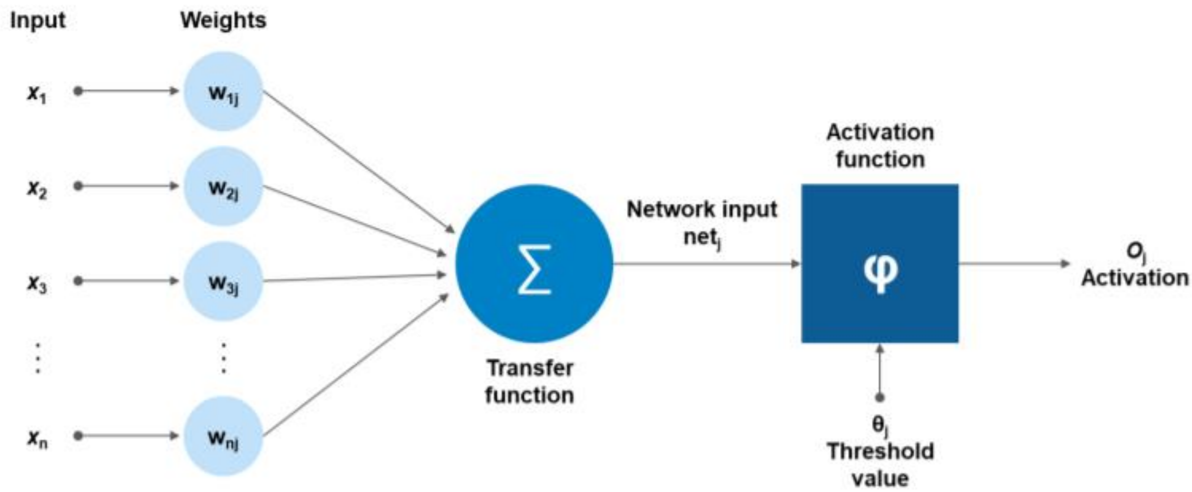

Figure.II.8: the operation of the artificial neuron.

The choice of an activation function turns out to be an important building block of neural networks.

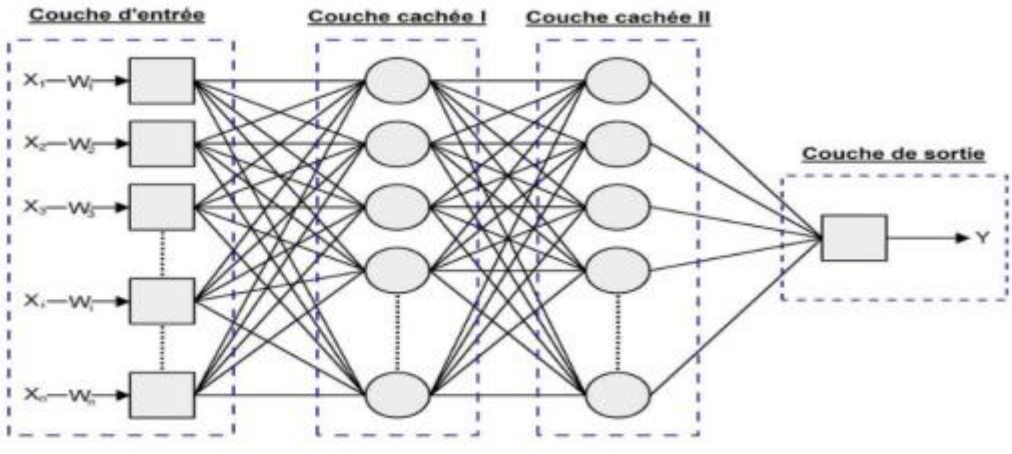

Figure.II.9: Layered neural network

We distinguish two network structures, based on the graph of their connections, that is, the graph whose nodes are the neurons and whose edges are the "connections" between them:

II.5.3.1. Static (or acyclic, or non-looping) neural networks:

A non-recurrent neural network performs one or more algebraic functions on its inputs by composing the functions executed by each neuron. In such a network, the flow of information moves from the inputs to the outputs without any feedback loops. This type of network is primarily used for tasks like nonlinear function approximation and modeling of static nonlinear processes.

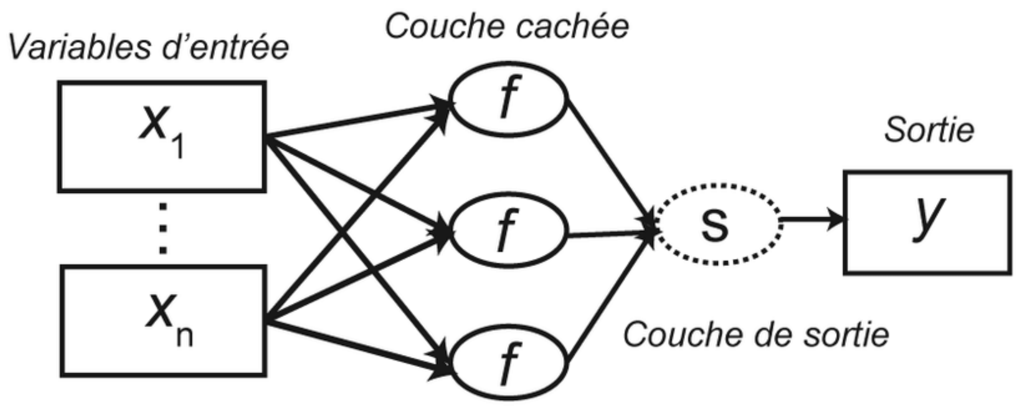

Figure.II.10: Example of a non-loopable neural network **[22]**.

II.5.3.2. Dynamic (or Recurrent or Feedback) Neural Networks:

Feedback neural networks are dynamic systems operating in discrete time, governed by one or more nonlinear difference equations. These equations arise from the combination of the functions performed by each neuron and the delays associated with each connection. They are used to model dynamic systems, process control, system regulation, or filtering tasks.

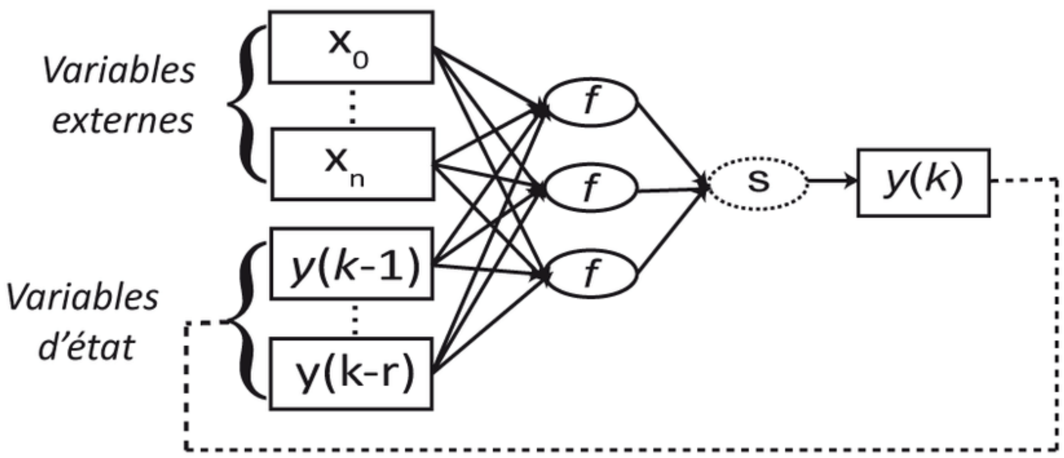

Figure.II.11: Example of a looped or recurrent neural network **[22]**.

II.5.4.FORMAL NEURAL NETWORKS:

There are two main types of neural network architectures: feedforward neural networks and recurrent neural networks.

II.5.4.1. Feedforward Neural Networks:

A feedforward neural network performs one or more algebraic functions of its inputs, achieved through the composition of the functions executed by each of its neurons. A feedforward neural network is graphically represented as a set of neurons "connected" to each other, with information flowing from the inputs to the outputs without any "feedback." If the network is depicted as a graph, where the nodes represent the neurons and the edges represent the "connections" between them, the graph of a feedforward network is acyclic. The term "connections" is a metaphor: in most applications, neural networks are algebraic formulas whose numerical values are computed by computer programs, not physical objects (specialized electronic circuits); however, the term "connection," which originates from the biological roots of neural networks, has become commonplace, despite being somewhat misleading. It even led to the term "connectionism."

(Figure.II.12)shows a feedforward neural network with a particular structure, commonly used in practice: it consists of input neurons, a hidden layer of neurons, and output neurons. The neurons in the hidden layer are not connected to each other. This structure is referred to as a Multilayer Perceptron

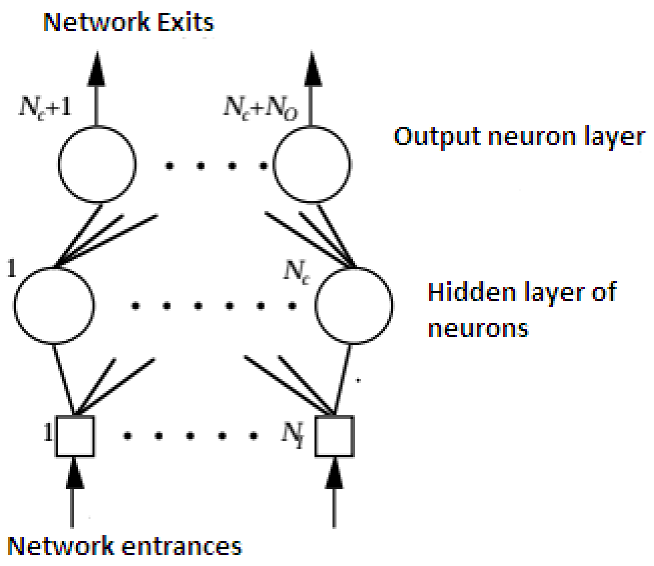

Figure.II.12: A multi-layer perceptron **[22]**.

Non-recurrent neural networks are static objects: if the inputs are independent of time, the outputs will be as well. They are primarily used for tasks such as nonlinear function approximation, classification, or modeling of static nonlinear processes.

II.5.4.2.Recurrent Neural Networks:

In contrast to non-recurrent neural networks, which have an acyclic connection graph, recurrent neural networks can have any kind of connection topology, including loops that feed outputs back to the inputs. For such a system to be causal, a delay must be associated with every loop: therefore, a recurrent neural network is a dynamic system governed by differential equations. Since most applications are implemented through computer programs, we typically work within the framework of discrete-time systems, where differential equations are replaced by difference equations.

A discrete-time recurrent neural network is governed by one (or more) nonlinear difference equations, resulting from the composition of functions performed by each neuron and the delays associated with each connection. The most general form of the equations governing a recurrent neural network is called the canonical form:

$$x(k+1) = \varphi\,[x(k), u(k)] \qquad \textbf{(II.3)}$$

$$y(k) = \psi\,[x(k), u(k)] \qquad \textbf{(II.4)}$$

where φ and ψ are nonlinear functions performed by a non-recurrent neural network (such as, but not limited to, a multilayer perceptron), and where k represents time (discrete). The canonical form is depicted in Figure 3. Any neural network, no matter how complex, can be automatically transformed into this canonical form **[22]**. Thus, the network shown in Figure.II.12. can be converted into an equivalent but more manageable canonical form, as shown in Figure.II.13.

Recurrent neural networks are used for tasks such as dynamic system modeling, process control, and filtering.

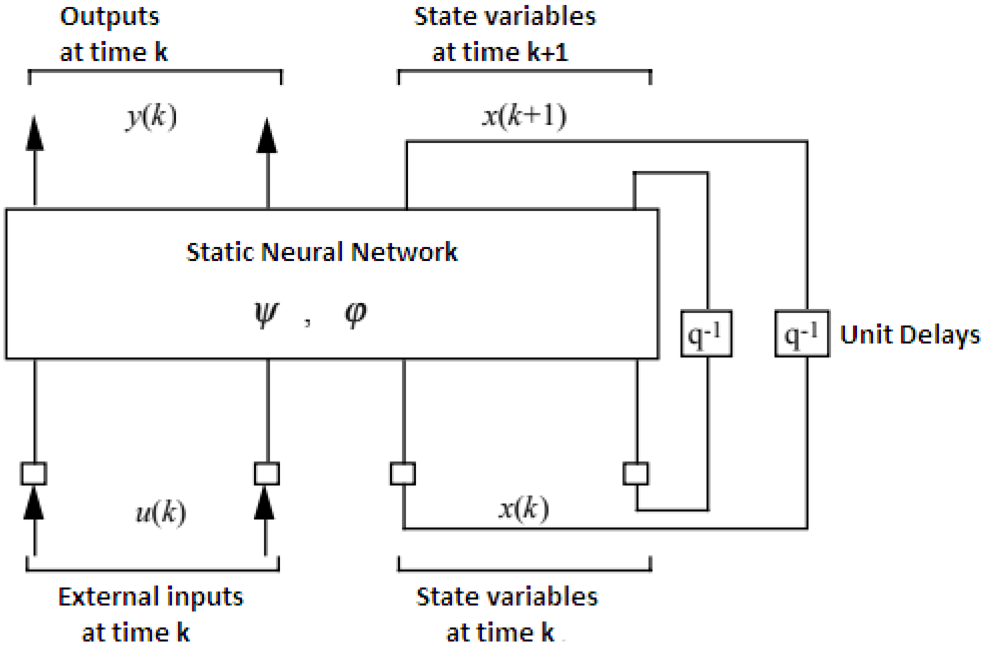

Figure.II.13: Canonical Form of a Loop-Backed Neural Network**[22]**.

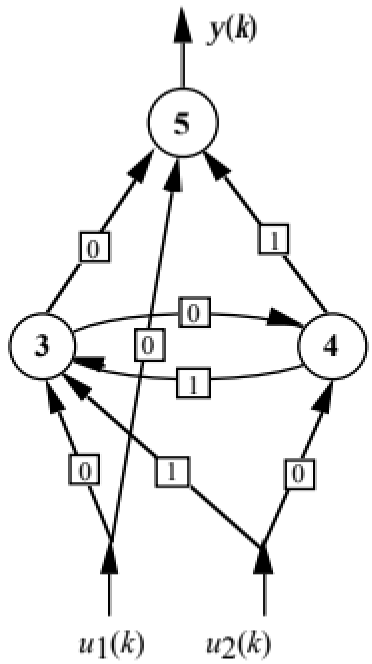

Figure.II.14: A looped neural network The numbers in the squares are the delays (expressed as the number of sampling periods) associated with each connection**[23]**.

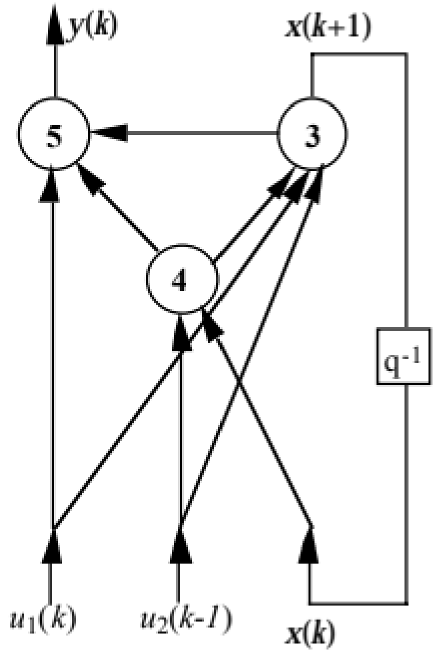

Figure.II.15: The canonical form of the network in **(Figure.II.14) [23]**.

Chapter III :

Results and Discussion

III.1 Introduction:

This chapter presents the results of applying Artificial Neural Networks (ANNs) to predict the mass fractions in the raffinate and extract phases of a liquid-liquid extraction system. The focus is on evaluating how the number of hidden layers (0, 1, and 2) and the number of neurons in each layer affect model performance. Several ANN architectures were trained and tested using experimental data. Their accuracy was assessed using metrics such as mean squared error (MSE) and R^2. The results are compared to identify the most effective structure for accurate and reliable prediction.

III.2 Application of Artificial Neural Networks (ANN):

In this study, Artificial Neural Networks (ANNs) were applied to model the complex relationship between process input variables and the mass fractions of components in both the raffinate and extract phases of a liquid-liquid extraction system. The model was developed and implemented using MATLAB, and it was trained on experimental data to capture nonlinear correlations between variables. Input and Output Variables To train the ANN model, six input variables were considered [24,25]:

Feed composition:

- X1
- X2
- X3

Salt concentration (C_salt):

- Values: 0, 0.005, 0.10, 0.15

Solubility:

- Values: 0.36, 0.34

Carbon concentration (C_carbon):

- Values: 1, 3

These inputs were selected based on their significant influence on the phase behavior in the liquid-liquid extraction process. The outputs of the model were the mass fractions of the same three components in both **the raffinate** and **extract** phases:

Raffinate phase:

- X1 , X2 , X3

Extract phase:

- X1, X2 , X3

The input and target data used to train and validate the network are presented in **Table III.1** and **Figure III.1**

Network Architecture:

Several ANN architectures were tested to evaluate the effect of network complexity onpredictive accuracy. The architecture was defined by:

Input layer: 6 neurons (corresponding to the six input variables)

Hidden layers: The number of hidden layers (0, 1, or 2) and the number of neurons ineach were varied systematically

Output layer: 6 neurons (corresponding to the six output variables)

Training Algorithm:

The network was trained using the **Levenberg–Marquardt algorithm**, which is known for its fast convergence and efficiency in nonlinear optimization problems. This algorithm provides a balance between the Gauss-Newton method and gradient descent.

Activation Function:

The tansig activation function (hyperbolic tangent sigmoid transfer function) was used in the hidden layers. It is defined as:

$$\text{tansig}(x) = \frac{2}{1+e^{-2x}} - 1 \qquad \textbf{(III.1)}$$

This function introduces nonlinearity and maps the input values to a range between -1 and 1, which facilitates faster learning and better performance in many cases.

Data Division

To obtain a reliable and generalizable ANN model, the dataset was divided into three subsets:

- **70% for training:** Used to teach the network by adjusting weights and biases.
- **15% for validation:** Used to tune model parameters and monitor overfitting during training.
- **15% for testing: Used** to evaluate the model's final performance on unseen data.

This division strategy ensures that the model is trained effectively while maintaining its ability to generalize to new inputs.

Performance Evaluation

The performance of each ANN architecture was assessed using the following metrics:

- **Mean Squared Error (MSE**): The primary performance function used to evaluate the average squared difference between predicted and actual values.

$$MSE = \frac{\sum(y_i - p_i)^2}{n} \quad \textbf{(III.2)}$$

- **Coefficient of determination (R^2):** Used to assess the proportion of variance in the output that can be predicted from the inputs.

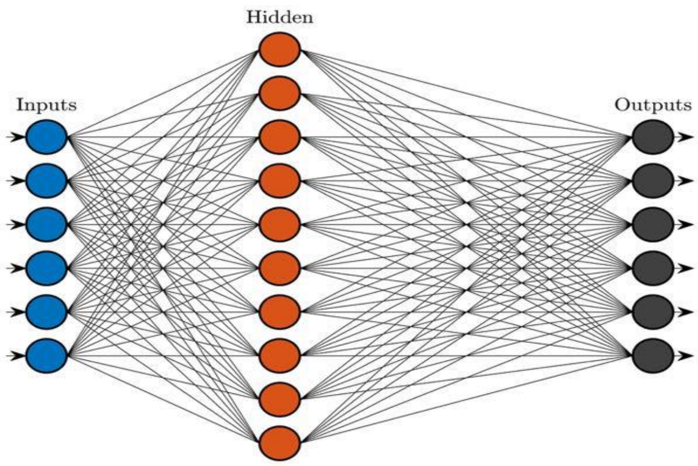

Figure.III.1 : Input-output architecture and variables in ANN model.

Table .III.1: The input and target data used to train and validate the network.

Input		**Output**	
Feed	**X_1**	**Raffinate**	**X1**
	X_2		**X2**
	X3		**X3**
C of salt	**(0, 0.005,0.1,0.15)**	**Extract**	**X1**
Solubility	**(0.36,0.34)**		**X2**
C of carbon	**(1, 3)**		**X3**

The table represents the results of applying artificial neural networks (ANNs). It contains three columns:

1. Number of hidden layers
2. Architecture
3. Mean square error (MSE)

Table III. 2: Results of applying Artificial Neural Networks (ANNs)

Number of hidden layers	Architecture	MSE
0	6/6	0.00164
1	6/2/6	0.00110
	6/4/6	0.00069
	6/6/6	0.00036
	6/8/6	0.00025
	6/10/6	0.00063
	6/12/6	0.00077
2	6/2/2/6	0.00097
	6/4/4/6	0.00051
	6/6/6/6	0.00040
	6/8/8/6	0.00042
	6/10/10/6	0.00042
	6/12/12/6	0.00023

Artificial neural networks were tested with different numbers of hidden layers (0, 1, or 2), and the number of nodes (neurons) in each layer was varied, as shown in the "Architecture" column.

Architecture, for example, 6/6 means the number of nodes in each layer (in this example: 6 inputs, 6 outputs, and no hidden layers).

MSE is a measure of accuracy. The lower the number, the more accurate the neural network.

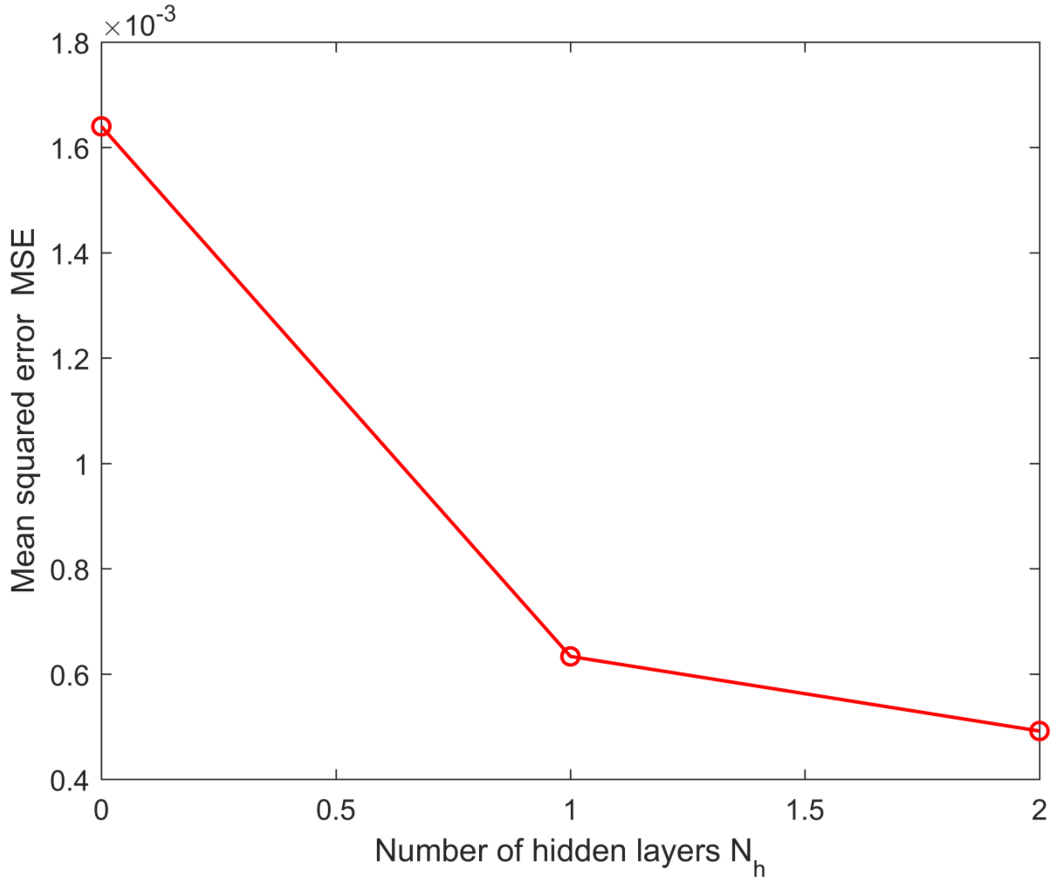

Figure.III.2: The effect of the number of hidden layers (Nh) on ANN.

This figure shows the effect of the number of hidden layers (Nh) on ANN . We notice that the more the number of hidden layers increases, the Mean Squared Error (MSE) decreases. From this, we conclude that the number of hidden layers should be increased for more accurate results.

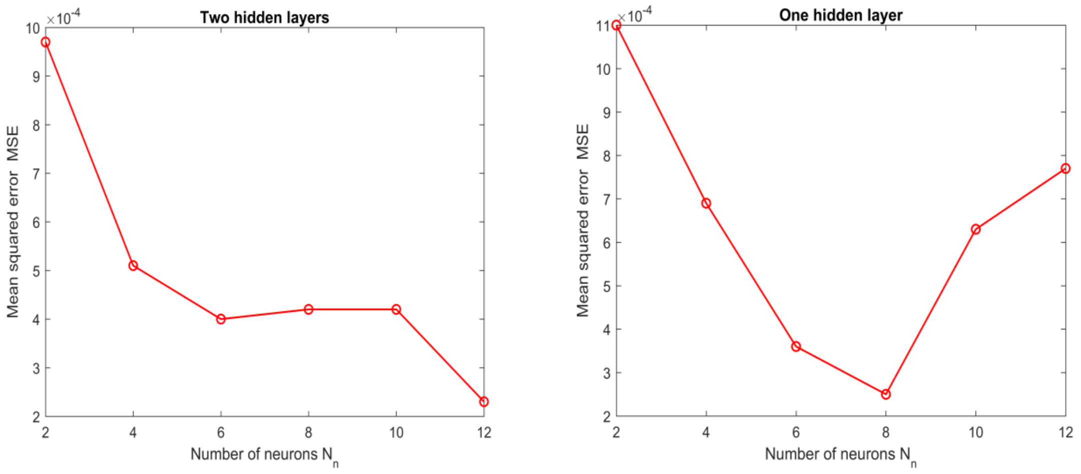

Figure.III.3: The effect of the number of neurons (Nn) on ANN.

This figure shows the effect of the number of neurons (Nn) on ANN . We notice that from 0 to 8 neurons the Mean Squared Error decreases, and from 8 to 12 it rises again . From this we conclude that 8 neurons is the most appropriate number.

III.3 Network Architectures and Results:

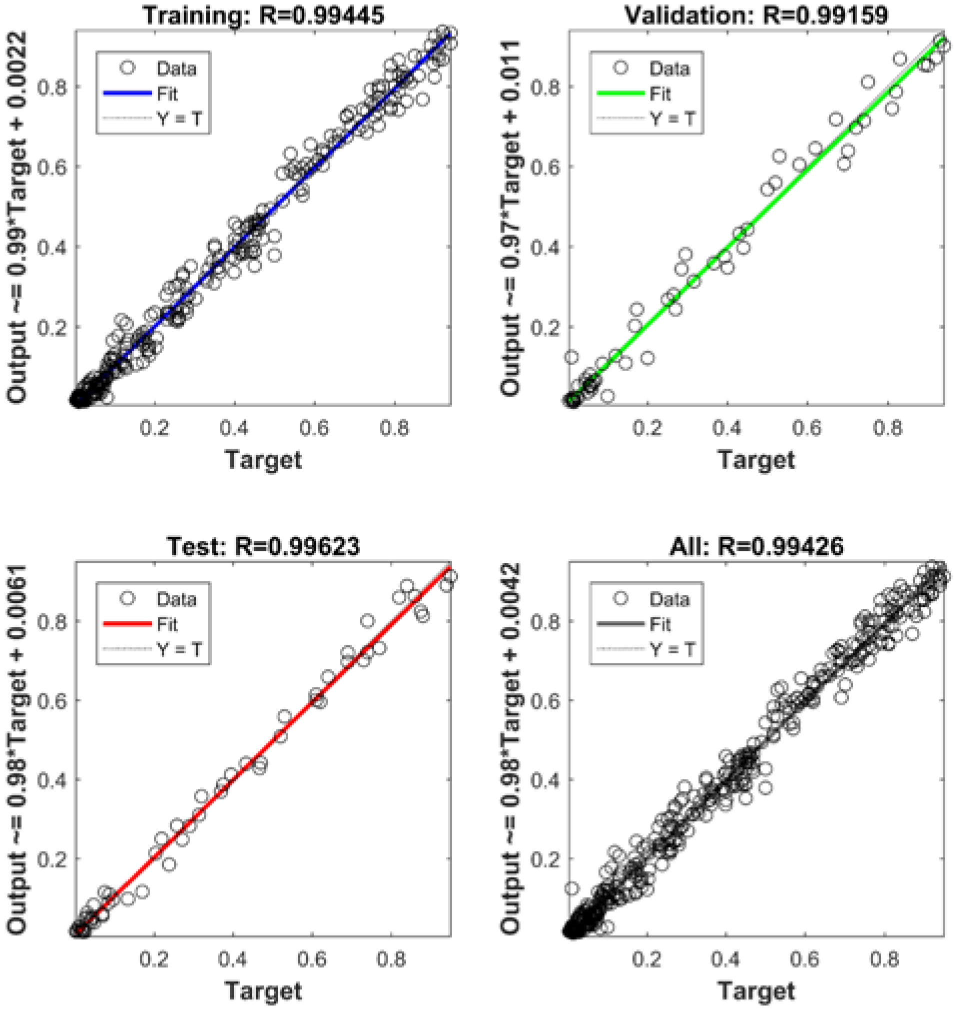

Figure.III.4 : The results of Architecture 6/6

This figure shows the results of Architecture 6/6 the graphs illustrate the relationship between the actual values (target) and predicted values (output) for three datasets: training, validation, and testing, in addition to the overall performance. The line in the middle is the linear regression line, which represents the extent to which the predictions match the actual values, and R represents the correlation coefficient. We notice that R values are close to 1, and all points are close to the middle line, indicating that the model shows excellent performance, demonstrating the accuracy of prediction and the stability of performance.

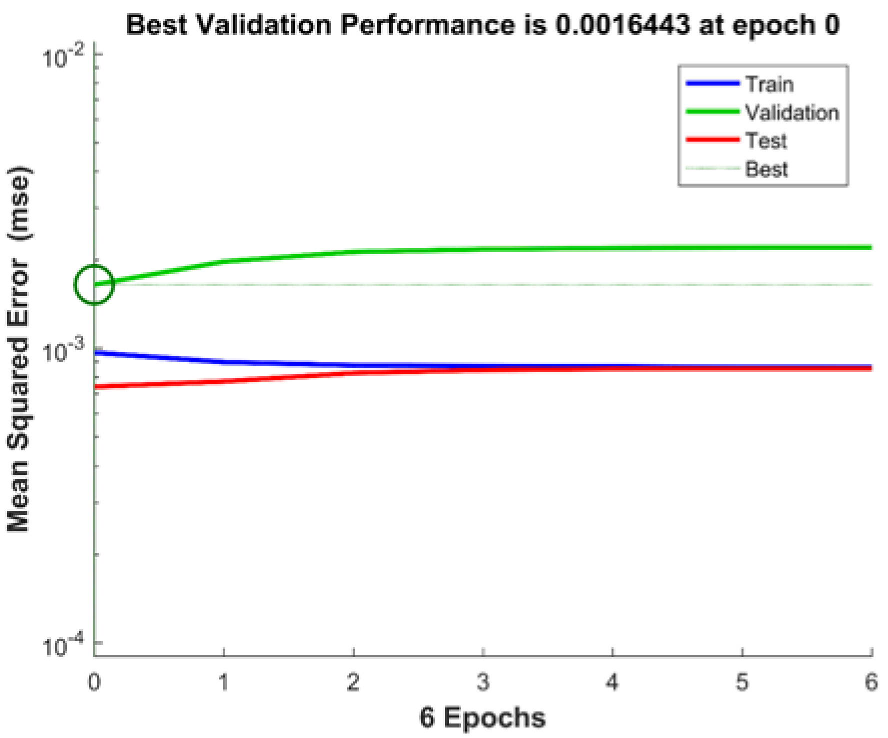

Figure.III.5: MSE curve for architecture 6/6.

This figure shows the mean squared error (MSE) curve and how the model's performance changes across a number of epochs on three sets (training , validation, and testing). The x-axis represents the number of epochs the number of times all the data was passed through the model. The y-axis represents the mean squared error (MSE), a measure of how far the model's predictions are from the true values. The lower the MSE, the better the model's performance. The blue (Train) line is relatively low and almost constant. The green (Validation) line is above the training data and begins to rise slightly. The red line is also low and constant.

The green circle represents the best validation performance, which occurred in the first epoch (epoch 0) with a value of 0.0016443.

We conclude that the error is very low across all sets, indicating excellent model performance.

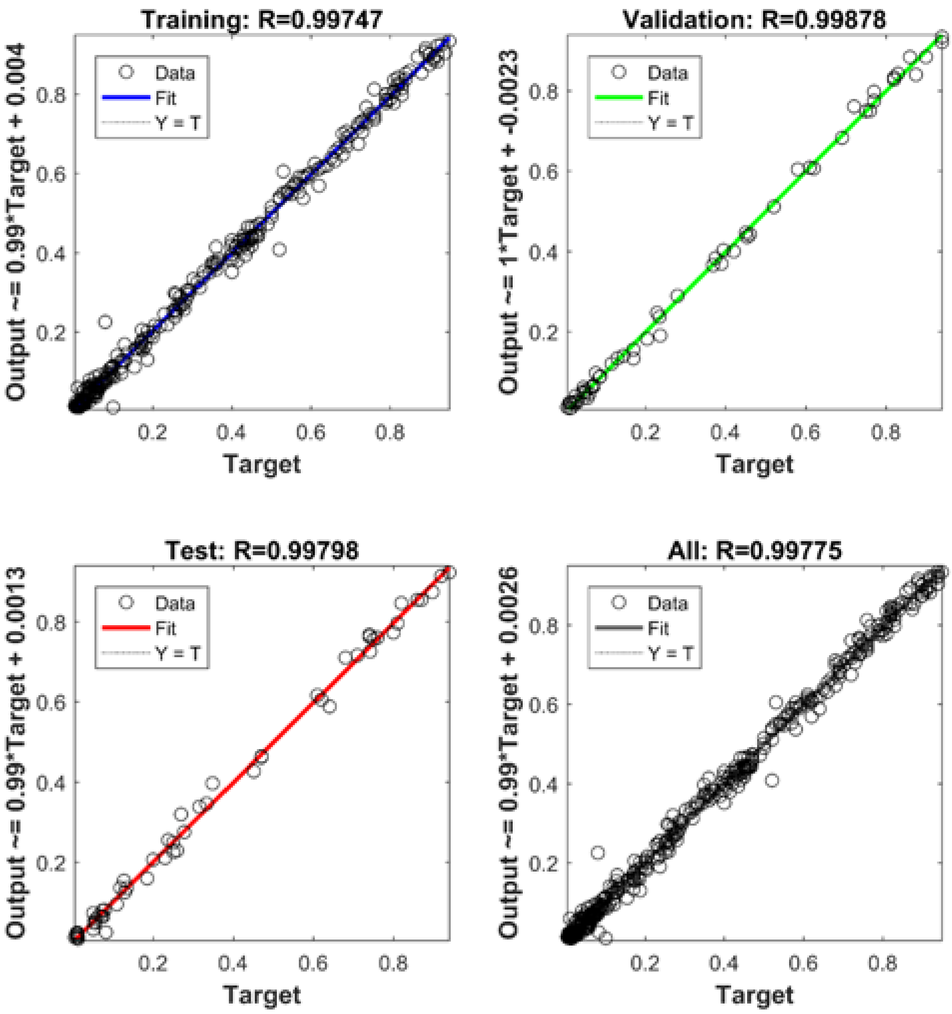

Figure.III.6 : ANN results for architecture 6/8/6

This figure shows the results of Architecture 6/8/6 .We notice that R values are very close to 1, and all points are very close to the middle line, indicating that the model shows excellent performance, demonstrating the accuracy of prediction and the stability of performance.

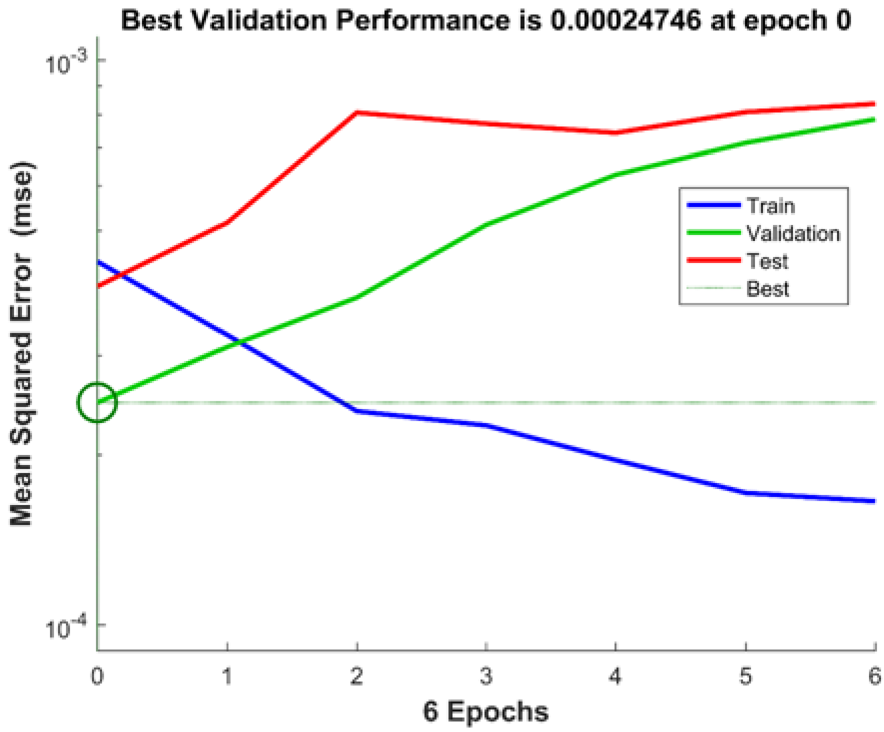

Figure.III.7 : MSE curve for architecture 6/8/6.

This figure shows the mean squared error (MSE) curve and how the model's performance changes across a number of iterations (epochs) in three sets: (Train, Validation, Test). The x-axis represents the number of epochs, and the y-axis represents the mean squared error (MSE). We notice that the blue (Train) line is steadily declining, indicating that the model is learning well from the training data. In contrast, the green (Validation) line starts to rise after the first iteration, indicating that the model's performance on the validation data is gradually deteriorating, an indication of over fitting. The red (Test) line is also rising, indicating that the model is not generalizing well to new data. The green circle represents the best validation performance with an approximate value of 0.00024746. This occurred in the first cycle (epoch 0), indicating that the model was at its best at the beginning of training. We conclude that: The error is very low at the beginning, indicating excellent performance of the model in the first cycle, but continued training after that led to poor performance of the model on the validation and test data due to overfitting of the training data.

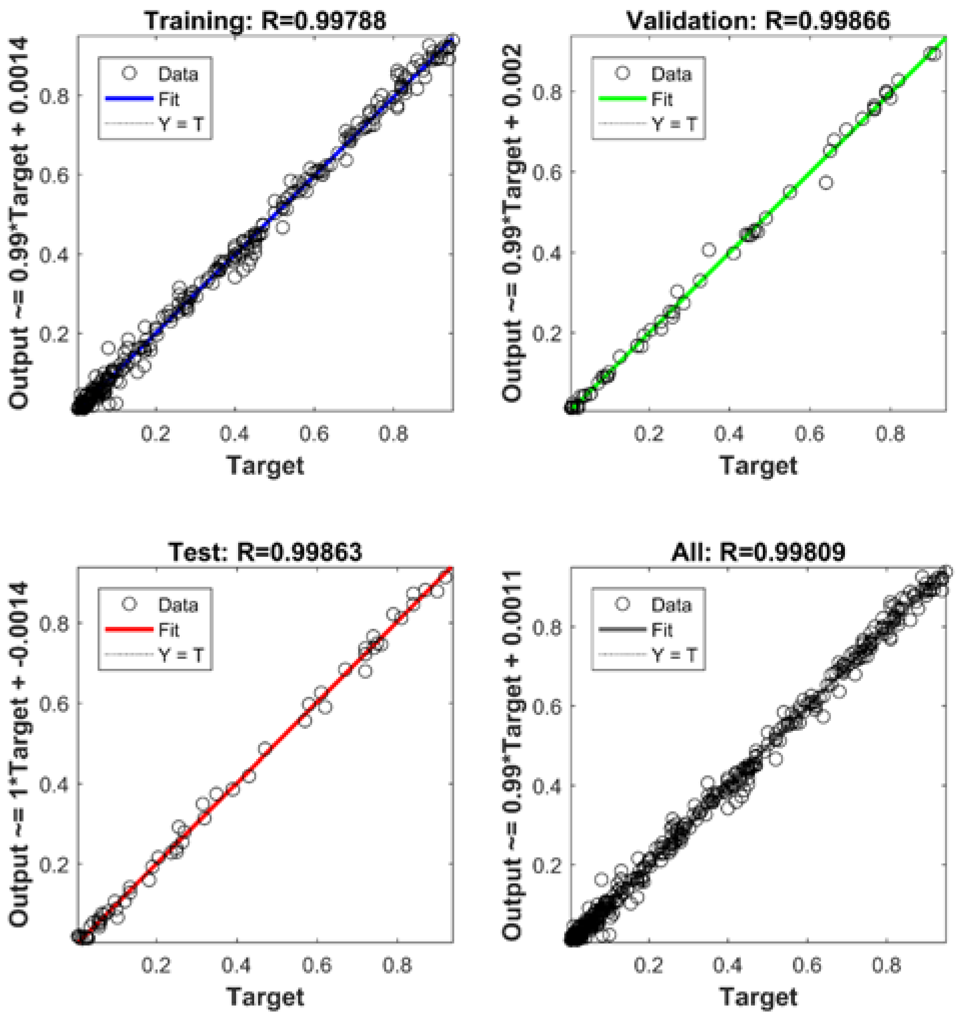

Figure.III.8 : ANN results for architecture 6/12/12/6.

This figure shows the results of a neural network model using a 6/12/12/6 architecture to predict target values.We can see that the correlation coefficient (R) values are very close to 1 (closer than previous cases), indicating a strong positive correlation and a high level of prediction accuracy. Furthermore, the data points are closely distributed around the middle line in all graphs reflecting the model's stability and reliability across different datasets. This confirms that the model has excellent generalization ability and consistent performance.

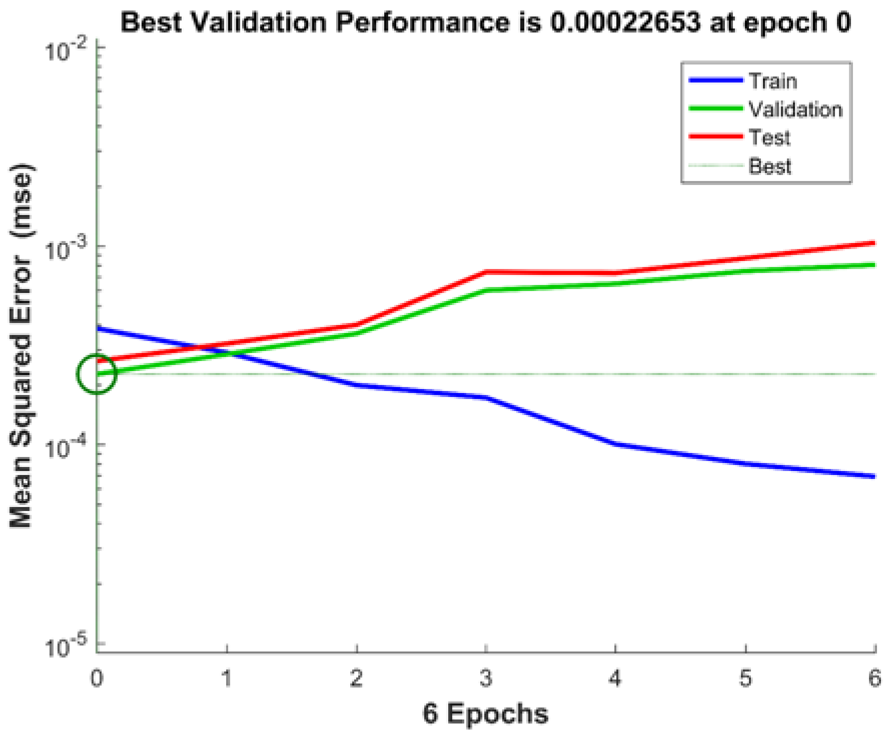

Figure.III.9 : MSE curve for architecture 6/12/12/6.

This figure shows the mean squared error (MSE) curve and how the model's performance changes across a number of iterations (epochs) in three sets: (Train, Validation, Test). The x-axis represents the number of epochs, and the y-axis represents the mean squared error (MSE). We notice that the blue (Train) line is steadily declining, indicating that the model is learning well from the training data. In contrast, the green (Validation) line starts to rise after the first iteration, indicating that the model's performance on the validation data is gradually deteriorating, an indication of over fitting. The red (Test) line is also rising, indicating that the model is not generalizing well to new data. The green circle represents the best validation performance with an approximate value of 0.00024746. This occurred in the first cycle (epoch 0), indicating that the model was at its best at the beginning of training. We conclude that: The error is very low at the beginning, indicating excellent performance of the model in the first cycle, but continued training after that led to poor performance of the model on the validation and test data due to over fitting of the training data.

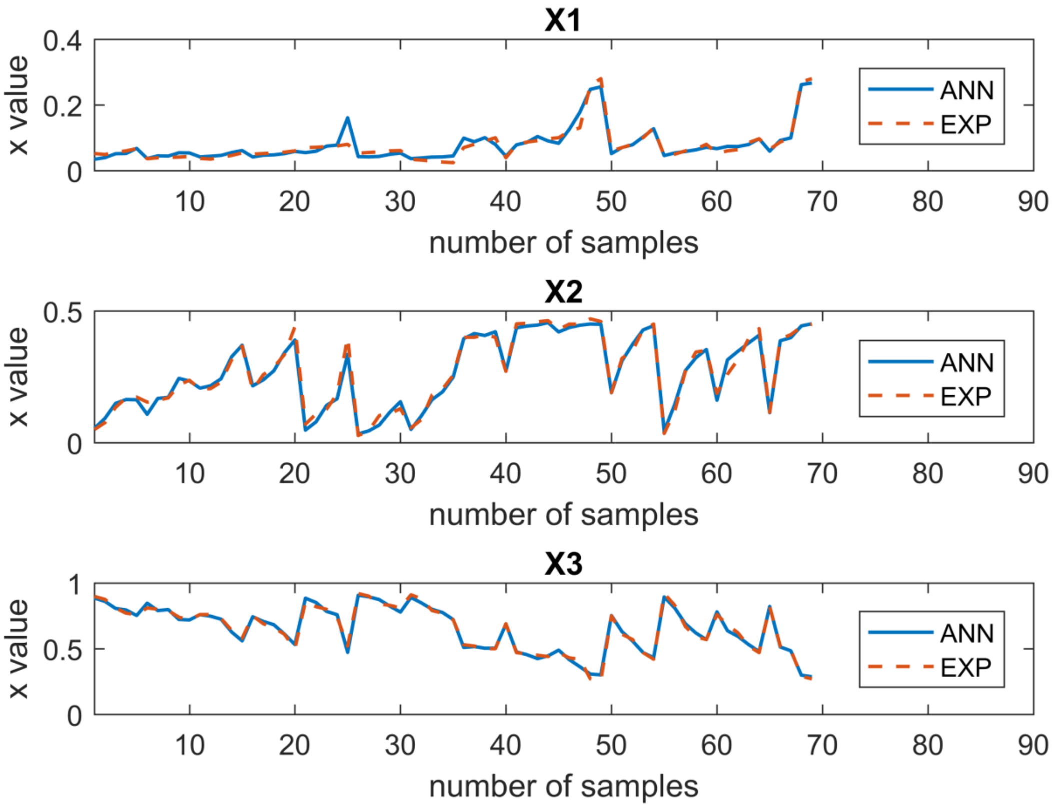

Figure.III.10: Comparison between the predicted values by (ANN) with (EXP) for structure 6/12/12/6.

These graphs show the results of structure 6/12/12/6. For raffinate compare the predicted values of the artificial neural network (ANN) with the experimental values (EXP) for three variables: X1, X2, and X3 are identical. The model performs best on X2, followed by X3, and slightly weaker on X1, but the differences are still minimal.

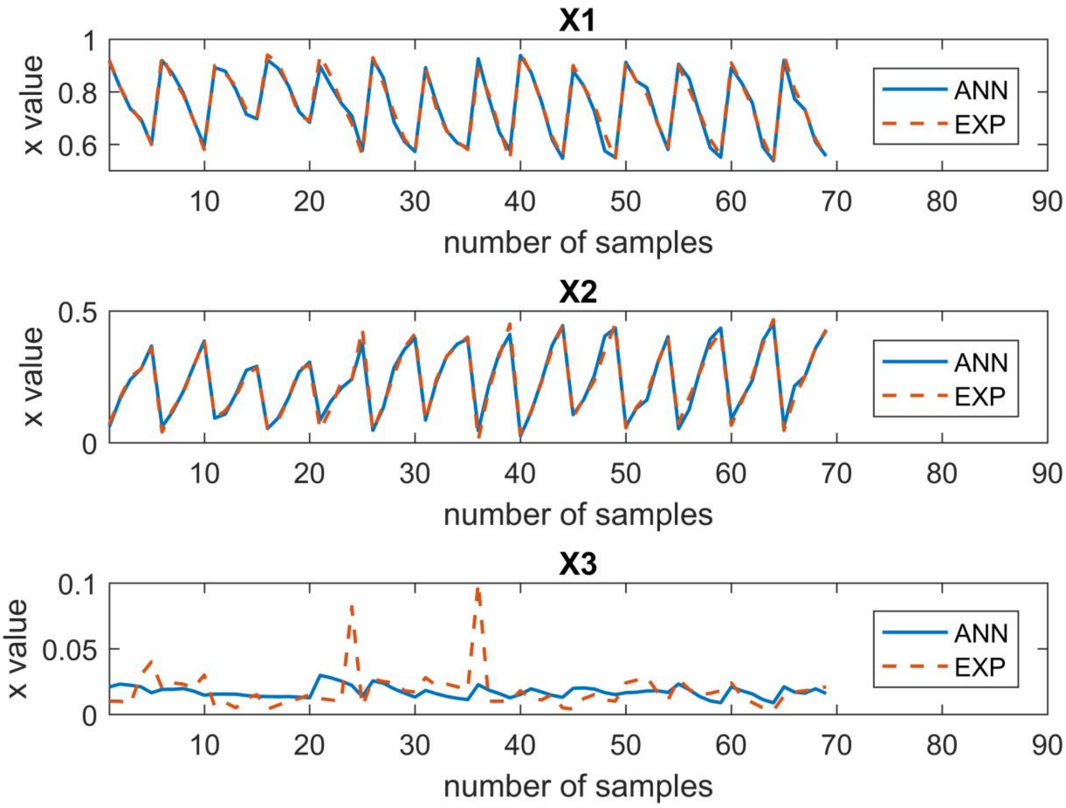

Figure.III.11: Comparison between the predicted values by (ANN) with (EXP) for structure 6/12/12/6.

.

These graphs show the results of structure 6/12/12/6. For extract compare the predicted values of the artificial neural network (ANN) with the experimental values (EXP) for three variables: X1, X2, and X3. X1 and X2 show high performance and excellent agreement. There are clear differences between X3;

III.4. Conclusion:

In this study, The network architecture was systematically evaluated by varying the number of hidden layers (0, 1, and 2) and testing multiple neuron configurations within each layer. The results clearly demonstrated that increasing the number of hidden layers led to improved predictive accuracy, with the two-hidden-layer architecture outperforming the simpler models. Among the various neuron arrangements tested, the configuration of 6/12/12/6 (input layer/hidden layer 1/hidden layer 2/output layer) was identified as the most effective, offering the highest precision in predicting phase compositions. These findings highlight the importance of network architecture optimization in enhancing the performance of ANN models for complex separation processes such as liquid-liquid extraction.

General Conclusion

In this study, the architecture of the artificial neural network was systematically evaluated by varying the number of hidden layers (0, 1, and 2) and testing several neuron configurations within each layer. The results clearly demonstrated that increasing the number of hidden layers improved the predictive performance, with the two-hidden-layer model outperforming simpler configurations. Among the tested structures, the architecture with a configuration of 6/12/12/6 (input layer/hidden layer 1/hidden layer 2/output layer) proved to be the most effective, offering the highest accuracy in predicting the mass fractions in the raffinate and extract phases. This architecture achieved a mean squared error (MSE) of 0.00023, indicating excellent predictive capability. Overall, the findings confirm that artificial neural networks are effective tools for modeling liquid-liquid extraction systems. These results underscore the importance of optimizing network architecture to enhance the performance of ANN models, particularly for complex separation processes.

Reference

[1] Poole, C. F. (2019). Liquid–liquid extraction. In *Liquid-Phase Extraction* (pp. 35–67). Elsevier. https://doi.org/10.1002/0471238961.koe00041

[2] International Journal of Scientific and Research Publications, Volume 2, Issue 6, June 2012 1

[3] Mjalli, F. S., & Al-Asheh, S. (2007). Artificial neural network based modeling of liquid–liquid extraction of phenol using soybean oil. *Chemical Engineering and Processing: Process Intensification*, 46(4), 387–392.

[4] Rydberg, J., Cox, M., Musikas, C., & Choppin, G.R. (2004). *Solvent Extraction Principles and Practice* (2nd ed.). CRC Press.

[5] A. Hamdi, Thèse de magister. Université Houari Boumediene, Alger, (1985).

[6] D. S Flett, D .R. Spink. Hydrométallurgies, 1, 207, (1974).

[7] E.KOLLER, Aide-mémoire de Génie chimique-4e éd, Paris (2001), p 433- 440.

[8] Pierre Wuithier. Raffinage et Génie Chimique, Tome II. Edition TECHNIP, Paris, 1972.p 939

[9] Azeddine Bouraqad Idrissi «Extraction par solvant ; Etude et modélisation du système Tributylphosphate- acides monocarboxiliques». Thèse de doctorat, institut national polytechnique de TOULOUSE. France, 2006.p 45.

[10] Emilian Koller; Aide mémoire Génie chimique, Edition DUNOD, 2eme Édition, 2004. p 202.

[11] Jacques Schwartzentruber. Équilibres liquide-liquide de mélanges ternaires Éditeur, EMAC. Couverture géographique, France, 26 juillet 2010.

[12] J.P.WAUQUIER; Le raffinage du pétrole, procédé de séparation; Edition Technip, Paris, 1998.p 65.

[13] Timedjeghdine. M, « Etude expérimentale et modélisation des systèmes (eau + acide carboxylique +solvants) », thèse de Doctorat en chimie industrielle. Université Mohamed Khider – Biskra, 2016.

[14] MERABTI.Y, « Optimisation des réseaux de neurones MLP par l'algorithme hybride AG-RT pour le contrôle d'un système non linéaire», mémoire de master, université Oum el bouaghi, 2015.

[15] https://www.lebigdata.fr/reseau-de-neurones-artificiels-definition. (25 avril 2019)

[16] Huawei Technologies Co., Ltd., Artificial Intelligence Technology, https://doi.org/10.1007/978-981-19-2879-6_1

[17] The Author(s) 2023 Huawei Technologies Co., Ltd., Artifi cial Intelligence Technology, https://doi.org/10.1007/978-981-19-2879-6_2

[18] Artificial intelligence,machine learning and deep learning Nadia BERCHANE (M2 IESCl,2018)

[19] Learning Machine Learning, Part 1: An Introduction Alibaba Clouder December 26, 2016

[20] Published by Springer International Publishing Switzerland 2016. All Rights Reserved K. Saeed and W. Homenda (Eds.): CISIM 2016, LNCS 9842, pp. 3–14, 2016.

[21] Artificial Neural Networks (ANN) 2020-2021 Dr. Najlaa M. Hussein

[22] C. BISHOP, Neural Networks for Pattern Recognition, Oxford University Press, 1995.

[23] G. DREYFUS, Y. IDAN, "The Canonical Form of Nonlinear Discrete-Time Models", Neural Computation Vol. 10, n°1, 1998.

[24] Laiadi, C., Lami, N., Merzougui, A., & Laouini, S. E. (2021). The influence of cation and temperature on the liquid–liquid equilibrium of the water + 2-butanone system and its simulation using ANN and UNIQUAC models. *Desalination and Water Treatment, 221*, 136–145. https://doi.org/10.5004/dwt.2021.27095

[25] Athmani, K., Almeasar, K. S., Elhouichet, H., & Guettala, A. (2025). Machine learning methods for predicting durability of earth mortars with date palm ash. *Asian Journal of Civil Engineering*. https://doi.org/10.1007/s42107-025-01365-0

ملحق القرار رقم: 933 المؤرخ في: 28 جويلية 2016

الذي يحدد القواعد المتعلقة بالوقاية من السرقة العلمية ومكافحتها

الجمهورية الجزائرية الديمقراطية الشعبية

مؤسسة التعليم العالي:

نموذج التصريح الشرفي

خاص بالالتزام بقواعد النزاهة العلمية لانجاز بحث

أنا الممضي أدناه،

السيد: شامخ سهام ..... الصفة: طالب، أستاذ باحث، باحث دائم: طالبة ....

الحامل لبطاقة التعريف الوطنية رقم: 205835915 والصادرة بتاريخ 2020/03/16

المسجل بكلية العلوم والتكنولوجيا قسم الكيمياء الصناعية ....

و المكلف بإنجاز أعمال بحث (مذكرة التخرج ، مذكرة ماستر ، مذكرة ماجستير ، أطروحة دكتوراه)، عنوانها: Prediction of liquids – liquid Equilibrium Using Artificial Neural Networks ....

أصرح بشرفي أني ألتزم بمراعاة المعايير العلمية والمنهجية ومعايير الأخلاقيات المهنية والنزاهة الأكاديمية المطلوبة في انجاز البحث المذكور أعلاه.

التاريخ: 15/06/2025

إمضاء المعني

ملحق القرار رقم: 933 المؤرخ في: 28 جويلية 2016

الذي يحدد القواعد المتعلقة بالوقاية من السرقة العلمية ومكافحتها

الجمهورية الجزائرية الديمقراطية الشعبية

مؤسسة التعليم العالي:

نموذج التصريح الشرفي

خاص بالالتزام بقواعد النزاهة العلمية لانجاز بحث

أنا الممضي أدناه،

السيد: لجزاري ابتسام الصفة: طالب، أستاذ باحث، باحث دائم: طالبة

الحامل لبطاقة التعريف الوطنية رقم: 205479386 والصادرة بتاريخ: 2019/12/15

المسجل بكلية: العلوم والتكنولوجيا قسم: كيمياء صناعية

و المكلف بإنجاز أعمال بحث (مذكرة التخرج ، مذكرة ماستر، مذكرة ماجستير ، أطروحة دكتوراه)، عنوانها: Prediction of liquid-liquid Equilibrium using Artificial Neural Networks

أصرح بشرفي أني ألتزم بمراعاة المعايير العلمية والمنهجية ومعايير الأخلاقيات المهنية والنزاهة الأكاديمية المطلوبة في انجاز البحث المذكور أعلاه.

التاريخ: 15/6/2025

إمضاء المعني

الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة محمد خيضر-بسكرة

كلية العلوم والتكنولوجيا

قسم الكيمياء الصناعية

بسكرة في: 15/06/2025

إذن بإيداع مذكرة الماستر بعد التصحيحات

أنا الممضي أسفله الأستاذ: عثماني خالد

الرتبة: أستاذ MCB

أستاذ مشرف على مذكرة ماستر-للطالب (ة): عثماني خالد

الشعبة: هندسة الطرائق

التخصص: هندسة كيميائية

بعنوان: Prediction of liquids - liquid Equilibrium Using Artificial Neural Networks

أرخص بإيداع المذكرة المذكورة.

الأستاذ المشرف

رئيس لجنة المناقشة