Highlights

Improvement the classification of a nanocomposite using nanoparticules based on a meta-analysis study, RNN and RNN-Monte Carlo algorithms

Rania LOUKIL,Wejdene GAZEHI,Mongi BESBES

- Proposition of new techniques of classification based RNN and RNN-Monte Carlo for a nanocomposite using nanoparticules.
- A structured meta-analysis task for synthesizing and analyzing the results of multiple recent studies.

Improvement the classification of a nanocomposite using nanoparticules based on a meta-analysis study, RNN and RNN-Monte Carlo algorithms

Rania LOUKIL[a,b], Wejdene GAZEHI[a,b] and Mongi BESBES[a,b]

[a]*Higher institue of Information and Communication Technologies, University of Carthage, Tunisia*
[b]*Laboratory of Robotics, Informatics, and Complex Systems ,ENIT, UTM, Tunisia*

ARTICLE INFO

Keywords:
Deep Learning
Polymeric nanoparticles
Meta-analytic study
physical properties
Classification
Regression
RNN
LSTM

ABSTRACT

This paper may be the first meta-analysis that presents a comprehensive synthesis of scientific works spanning the last five years, focusing on methodologies and results related to the analysis of nanocomposite using nanoparticules. The primary objective is to identify the optimal algorithm using software information and leading to better classification methodology. Specifically, this study come up with the advantages and the drawbacks of the most used algorithms and proposes an enhancement and performance of RNN based LSTM neurons. Besides, a comparaison of Deep Learning methods for the classification of polymeric nanoparticles, with polypropylene serving as a case study will be implemented. Experiment comparison were conducted to assess with one physical property, later expanded to four properties and finally to eight properties. Neural networks, including Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), and RNN-Monte Carlo, were employed for simulations. The evaluation criteria encompassed accuracy, calculation time, mean square error (MSE) and other metrics. The findings contribute to the selection of an optimal algorithm for the analysis of polymeric nanoparticles, emphasizing the potential of Deep Learning methodologies, particularly RNN-Monte Carlo, in advancing classification accuracy and efficiency.

1. Introduction

In recent years, the integration of advanced computational techniques, particularly those falling under the umbrella of artificial intelligence and machine learning, has significantly impacted materials science and engineering Banga, Gehani, Bhilare, Patel and Kara (2018).This influence is particularly pronounced in the evaluation and design of novel materials with enhanced properties. This overview introduces several research studies at the forefront of this intersection, highlighting their contributions to the understanding and advancement of materials science. In fact, The rapid evolution of Deep Learning techniques has catalyzed a profound transformation across various domains, ranging from image recognition to natural language processing and medical diagnostics. In light of this, a meta-analytic study becomes indispensable to comprehensively assess the landscape of Deep Learning approaches, understand their effectiveness, and identify overarching trends and patterns. In addition, the Deep Learning approaches have found applications in diverse fields, including computer vision Voulodimos, Doulamis, Doulamis, Protopapadakis et al. (2018) Ioannidou, Chatzilari, Nikolopoulos and Kompatsiaris (2017), speech recognition Deng and Platt (2014)Zhang, Geiger, Pohjalainen, Mousa, Jin and Schuller (2018), natural language understanding Otter, Medina and Kalita (2020) Goyal, Pandey and Jain (2018), and healthcare Miotto, Wang, Wang, Jiang and Dudley (2018) Faust, Hagiwara, Hong, Lih and Acharya (2018). As a consequence, a multitude of studies and implementations have emerged, each tailored to address specific challenges within these domains. With the growing number of deep learning studies, a meta-analytic approach becomes crucial for consolidating and synthesizing the knowledge dispersed across a myriad of research papers.

Nanomaterials are materials composed of particles where at least one dimension falls within the range of 1 to 100 nanometers, with at least 1% of these particles falling within this dimensional range. Their specific surface area exceeds 60 m²/cm³, excluding materials consisting of particles smaller than the nanometer Gaffet (2011). There are various categories of nanomaterials, each characterized by distinct composition, structure, and propertiesSingh, Rehman and Pandey (2023). Nanotechnology encompasses the design, characterization, production, and application of structures, devices, and systems by controlling shape and size at the nanoscale. The terms nanosciences and nanotechnologies encompass a body of knowledge and techniques shared across various traditional scientific disciplines, including chemistry, physics, materials science, technology, biological sciences, medicine, and environmental sciences. The unifying factor is the nanoscale size of the studied object and/or the tools used to manipulate it. It can now be considered that the emergence of nanotechnologies represents a major turning point in the development of industry in the 21st century Sanchez and Sobolev (2010). Nanomaterials are present in various sectors and play a crucial role in many industries, thanks to their unique properties and ability to

rania.loukil@istic.ucar.tn (R. LOUKIL); wejden.gazehi@istic.ucar.tn (W. GAZEHI); mongi.besbes@ucar.tn (M. BESBES)
ORCID(s): 0009-0000-2817-2726 (R. LOUKIL); 0009-0005-9230-3264 (W. GAZEHI); 0000-0003-0953-356X (M. BESBES)

revolutionize existing applications. They form the foundation of numerous technological and scientific advancements, contributing significantly to innovation and the improvement of products and processes in diverse fields such as electronics, medicine, energy, chemical industry, aerospace, and many others Lartigue (2010). Nanotechnologies have a revolutionary potential in the medical field, offering significant advancements in early diagnostics, personalized treatments, and notable improvements in the management of various conditions, especially in oncology and the treatment of disabilities Chouard, Cabanis, Chambron, Milgrom, Adolphe, Ardaillou, Aurengo, Bach, Battin, Baulieu et al. (2008). Nanotechnology emerges as a transformative force promising to profoundly revolutionize the fields of medicine and dentistry, providing significant opportunities to enhance human health and push the boundaries of traditional medical and dental treatments. Mallanagouda Patil et al. Patil, Mehta and Guvva (2008) underscore in their article the revolutionary importance of nanotechnologies in the realms of medicine and dentistry. These advancements offer substantial potential in the diagnosis, treatment, and prevention of diseases, propelling nanomedicine to the forefront of future developments. Progress in nanotechnologies allows envisioning a future where physicians will have access to tools such as medical nanorobots, capable of conducting therapeutic and reconstructive interventions at the cellular and molecular levels. The development of new nanoscale materials for the dental industry presents innovative prospects for periodontal management and oral healthcare. In the field of the environment, nanotechnologies are employed for the preparation and shaping of nanomaterials. Innovative synthesis methods enable the creation of nanostructured materials with specific properties Roualdès (2009).
In the energy domain, nanotechnologies are applied to membrane synthesis, contributing to the production of membranes with enhanced selectivity for the treatment of water or effluents through nanofiltration Roualdès (2009).

In 2019, Bingcheng Luo et al. Luo, Wang, Tian, Cai and Li (2019) contributed to the understanding of nanoparticle dispersion within composite materials. Employing machine learning algorithms, the study introduces quantifiable tools such as the coefficient of variation of K-nearest neighbor distances and information entropy. The research analyzes 230 different morphologies of nano-composite dispersion, revealing that lower values of these metrics correspond to a more homogeneous dispersion of nanoparticles.

In 2020, Mirza Sahaluddin and collaborators Sahaluddin, Alade, Oyedeji and Aliyu (2020) presented a machine learning-based model for estimating the density of nitride nano-fluids dispersed in ethylene glycol. Utilizing support vector regression (SVR), the study demonstrates high precision in predicting nano-fluid densities. The proposed model, trained on experimental data of aluminum, titanium, and silicon nitride nanoparticles, outperforms traditional mixing theories, showcasing the efficacy of machine learning in predicting complex fluid properties.
In 2021, Masoud Fetanat et al. Fetanat, Keshtiara, Keyikoglu, Khataee, Daiyan and Razmjou (2021) present the machine learning for the design of thin-film nano-composite membranes. This research employs artificial neural networks (ANN) and convolutional graph networks (GCN) to establish relationships between membrane microstructure and macro-properties. The proposed method showcases superior accuracy in predicting penetration and rejection of pollutants, emphasizing the potential of machine learning in optimizing membrane design for diverse applications.

In 2022, The study conducted by Kyungmin Baek et al.Baek, Hwang, Lee, Chung and Cho (2022) explores the application of deep learning techniques in the evaluation of electromechanical properties in intricately structured polymer nanocomposites. Leveraging two types of deep neural networks (DNN), trained on multiscale analysis results, the research demonstrates the superiority of deep learning over traditional models. Specifically, convolutional graph networks (CNN) exhibit enhanced capabilities in assessing the impact of nanoparticle distribution and agglomeration on the electrical and mechanical properties of polypropylene matrix composites.

In 2023, the investigation led by Kyungmin Baek et al. Baek, Kim, Shin, Park and Cho (2023) focuses on a multiscale study to understand the agglomeration effects of nano-SiC particles in polypropylene matrix composites. The study aims to elucidate the impact of nanoparticle distribution on electrical conductivity. By employing advanced multiscale techniques, the research seeks to unravel the intricacies of nanoparticle agglomeration and its correlation with the electrical properties of the composite material.

In summary, these studies showcase the transformative potential of advanced computational techniques, including Deep Learning and machine learning, in materials science research. The integration of these methodologies allows for more accurate predictions, optimization of material design, and a deeper understanding of the complex relationships within composite materials. Our study aims to distill insights from a broad spectrum of research endeavors, providing a comprehensive overview of the current state of Deep Learning methodologies. As deep learning continues to reshape the technological landscape, this meta-analytic study endeavors to provide a panoramic view of the current status and trends within the field. By distilling insights from an extensive body of research, it aspires to contribute to the ongoing discourse surrounding the optimization and application of Deep Learning techniques across diverse disciplines.

This paper is orgonized of four parts. It is structured as follow: in Section 2, we summarise the problem classification and employed solution for analyzing nanocomposites. Initially, we would like to delve into the issue surrounding the classification of nanocomposites. This crucial question arises from the inherent complexity in characterizing these materials at the nanoscale. To address this challenge, a previously employed solution will be examined in detail in a meta analysis job for previous jobs during the last five years.

In Section 3, we will propose new techniques to improve nanocomposite classification. In fact, we will describe the used Dataset, then we will introduce the ANN model and its results , the RNN model and the RNN-Monte Carlo. In each part, we will intorduce the architecture of the network then our simulation results using the Matlab platform.
In Section 4, we present comparative tables between the previous works and our findings. We will discuss the results in term of IA methods and changing the number of properties in order to involve a thorough comparison and detailed discussion between the results obtained in our research and those stemming from previous studies.
Finally, the section 5 will be the conclusion and the perspectives of our methodology.

2. Problem Classification and employed solution for analyzing nanocomposites

Our objective is to reliably detect the properties of nanoparticles, especially in the face of challenges associated with nanoparticle measurement that can be difficult or even unattainable due to their extremely small scale and associated technological limitations. The issue becomes even more complex when certain physical properties are not directly accessible, posing an additional challenge in the precise characterization of these materials at the nanoscale. How to overcome these obstacles and implement effective methods, perhaps by exploring innovative approaches such as the use of advanced measurement techniques or the application of machine learning models, to achieve a reliable assessment of nanoparticle properties? This research aims to address these issues and propose solutions to enhance the understanding, classification, and detection of nanoparticle characteristics, despite the inherent challenges of their nature and scale.
The challenge lies in the laborious and costly process researchers face when attempting to characterize the composition of nanoparticles. Frequently, determining the number of nanoparticles in a composition takes time, and interactions between nanoparticles can complicate their classification.

2.1. Studied properties

We will focus in this section on the most popular nanomaterial properties which has been already studied from P1 to P8, the other properties from P9 to P36 are defined in the appendix.

1. P1: Electrical conductivity (EC) : In the context of nanocomposites refers to the material's ability to allow the passage of electric current through its structure. In the specific case of nanocomposites, which are materials composed of a polymer matrix and dispersed nanoparticles, electrical conductivity is influenced by the distribution, dispersion, and agglomeration of nanoparticles within the polymer matrix Baek et al. (2022). The unit of measurement for electrical conductivity is siemens per meter, often abbreviated as S/m.
2. P2: Elastic modulus (M): The elastic modulus of polymer nanomaterials, also known as the elastic modulus of polymer nanocomposites, is a mechanical property that measures the stiffness of a polymer material on the nanoscale when reinforced or modified by the incorporation of nanoparticles. It is often abbreviated as "nanoscale elastic modulus" or "nanomechanical elastic modulus." This elastic modulus is a measure of the material's ability to resist elastic (reversible) deformation when stress is applied. It is expressed in pascals (Pa) or gigapascals (GPa) and can be used to assess the mechanical performance of polymer nanocomposites in applications where deformation resistance and dimensional stability are crucial Shady and Gowayed (2010).
3. P3: Tensile strength (T): Tensile strength is a mechanical property of materials that measures their ability to withstand a tensile force, i.e., a force that tends to stretch the material. It is usually expressed in stress units (pascals, megapascals, or newtons per square millimeter) and represents the maximum force a material can endure before breaking or undergoing significant plastic deformation.
4. P4: Concentration: The concentration refers to the number concentration which is the quantity of nanoscale entities per unit volume. It is commonly expressed in terms of particles per unit volume and is an important property when characterizing nanomaterial dispersions Haiss, Thanh, Aveyard and Fernig (2007).
5. P5: molecular weight: The molecular weight of a nanomaterial refers to the sum of the atomic weights of all atoms in a single molecule of that nanomaterial. It is typically expressed in atomic mass units (u) or daltons (Da) Legrand and Gaines Jr (1969).
6. P6: Temperature: The temperature (T) is a measure of the average kinetic energy of particles within a substance, including nanomaterials. It reflects the degree of hotness or coldness of a material and determines the direction of heat flow. It is typically measured in degrees Celsius (°C) or Kelvin (K)Rittigstein and Torkelson (2006).
7. P7: The size of nanomaterials (S): Nanomaterials exhibit unique characteristics due to their extreme smallness on the nanoscale. Specifically, a nanomaterial can be defined as a material with at least one dimension ranging from 1 nm to 100 nm Bégin-Colin, Le Caër and Girot (2002) Azouani (2009).
8. P8: Flexural strength (F): The flexural strength of polymer nanoparticles refers to the ability of a composite material containing polymer nanoparticles to resist an applied force that tends to bend or flex it. This strength measures the material's ability to maintain its structural integrity when subjected to bending stresses. Flexural strength is measured in pascals (Pa) or megapascals (MPa), equivalent to one million pascals.

Table 1
Properties and IA methods

Researchers	Algorithms	Properties	Regression
Binh et al.(2019)	SVM	P14..P27	Linear
Binh et al.(2019)	ANN	P14..P27	Linear
Matos et al.(2019a)	ANN	P1,P7,P9,P24	Linear
Mirza et al.(2020)	SVR	P7,P12..P14	Linear
Divya et al.(2021)	ANN	P4..P7	Linear
Baek et al.(2022)	ANN	P1	Linear
Baek et al.(2022)	CNN	P1	Linear
Champa,E. et al.(2024)	RNN	P5	Linear
Champa,E. et al.(2024)	RNN-LV	P5	Linear

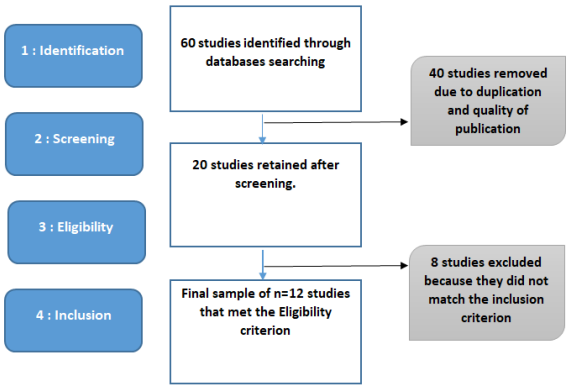

Figure 1: The Prisma process diagram

We have already presenting the most used properties in litterature but there are other properties which are explained in the appendix. The following table (1) provides a comprehensive overview of researchers in terms of some properties such as conductivity distance and distribution, density, flux, tensile strength, and compression coefficient. Besides, IA methods used in the last five years of some studied researchers will be explained in table 1.

We find that the most of the researchers used the SVM, ANN and CNN algorithms. All of them present the Linear regression. In fact, it is used for predicting a continuous outcome. The output is a continuous value, and the algorithm tries to fit a straight line that best represents the relationship between the independent variables and the dependent variable.

2.2. Methodology and Analysis

For our investigation, we employed the PRISMA method (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) to construct our sample. Illustrated in Fig. 1, this method encompasses the four stages of identification, screening, eligibility, and inclusion. Subsequently, we refined our sample by applying quality criteria in the second step. To ensure the study's quality, only peer-reviewed journal articles were considered, while those published in books, meetings, or proceedings were excluded Gedda (2015). Consequently, our sample comprises 20 peer-reviewed journal papers. In the third step, we conducted a more detailed examination of the papers using specific qualifying criteria. (i) The primary subject of the studies should be related to our investigation. (ii) Studies with unclear concept definitions or those focusing on mathematical modeling were omitted. (iii) The research must be quantitative and report Pearson's correlation or include values such as p-values and t-values that can be converted into Pearson correlation. At this juncture, a total of 8 studies were excluded as they did not meet all the requirements. The remaining sample consists of 12 distinct studies.

For the work of Bingcheng Luo et al. (2019) Luo et al. (2019); a quantitative assessment of nanoparticle dispersion was conducted using machine learning algorithms. The morphologies of nano-composites were employed as the input training dataset, and both distance and distribution were extracted to be used as a fingerprint. The correlation coefficient and information entropy derived from distance and distribution were adopted as indices for evaluating dispersed information. Utilizing the K-nearest neighbors algorithms to bridge the relationship between correlation and dispersion sets the standard for determining what constitutes "good" or "bad" nanoparticle dispersion. The information entropy (also known as Shannon entropy) measures and quantifies the disorder or uncertainty associated with a random variable, with each distribution having a corresponding entropy.

ö,,Miguel A.S. Matos and al. Matos, Pinho and Tagarielli (2019b) present in their article predictive multiscale models of the response to multiaxial deformation in conductive carbon nanotube (CNT) and polymer-based composites. The authors use detailed finite element (FE) models at the micron scale to generate training data for an artificial neural network. This neural network is then employed at a macroscopic scale to predict the electromechanical response of components with arbitrary shapes subjected to a non-uniform, multiaxial strain field, allowing computational time savings of six orders of magnitude. In other words, the researchers have established a multi-scale simulation methodology based on physical models to predict how CNT-polymer composites react to non-uniform multiaxial deformation fields. They utilized detailed simulations at the micron scale to train an artificial neural network, which can subsequently predict the electromechanical response at the macroscopic scale much more efficiently in terms of computational time. This approach has been applied to explore the use of CNT-polymer composites in the construction of various types of sensors and for damage detection. In summary, the article proposes a modeling method that combines the accuracy of physically-based simulations at a small scale with the efficiency of large-scale machine learning to predict the

electromechanical behavior of composites under complex deformation conditions .

The work of M.A.S. Matos et al.(2019) Matos, Pinho and Tagarielli (2019a); addresses the prediction of electrical conductivity in polymer and carbon nanotube composites through the application of an artificial neural network. Industrial applications of conductive polymer composites with carbon nanotubes require precise tuning of their electrical properties. Existing theoretical methods for predicting overall conductivity often rely on experimental adjustments and frequently employ power laws that are only valid in the vicinity of the percolation threshold. Although numerical methods offer high precision, their accuracy comes with substantial computational efforts. In this article, the authors utilize recently developed physics-based finite element analyses to successfully train an artificial neural network, enabling predictions of the overall conductivity of carbon nanotube-polymer composites at negligible computational cost .

Binh Thai Pham et al. Pham, Nguyen, Van Dao, Prakash, Ly, Le, Ho, Nguyen, Ngo, Hoang et al. (2019) focus in their article on the development and application of artificial intelligence (AI) models for predicting the Compression Coefficient of soil (Cc), a crucial geotechnical parameter. The study employs three different AI models: Artificial Neural Network (ANN), Adaptive Network based Fuzzy Inference System (ANFIS), and Support Vector Machine (SVM). The researchers utilized a Monte Carlo approach for sensitivity analysis of the AI models and their input parameters. The study involved the analysis of 189 clayey soil samples to construct and validate the models. Thirteen input parameters were considered, including sample depth, bulk density, plasticity index, moisture content, clay content, specific gravity, void ratio, liquid limit, dry density, porosity, plastic limit, degree of saturation, and liquidity index, to predict the Compression Coefficient (Cc) as the output parameter. Model validation was performed using statistical methods such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Determination (R2). The results indicated that all three models performed well, with the SVM model being the most effective in predicting Cc. Monte Carlo sensitivity analysis revealed that four input parameters—clay content, degree of saturation, specific gravity, and sample depth—were the most relevant in predicting Cc. Conversely, parameters such as bulk density, dry density, void ratio, and porosity were found to be less significant. The removal of these less significant parameters improved the models' performance by reducing the dimension of the input space and decreasing the model's runtime.

Merve Özkan and al. Özkan, Karakoç, Borghei, Wiklund, Rojas and Paltakari (2019) address the application of machine learning methods in the design of nanocellulose films in their article. Nanocellulose is a nanomaterial derived from cellulose fibers, possessing unique mechanical and structural properties. The researchers aimed to optimize the mechanical properties of nanocellulose films by incorporating commonly used additives in the paper industry, namely polyvinyl alcohol (PVA), glyoxal (Gx), and ammonium zirconium carbonate (AZC). The researchers employed artificial neural network (ANN), random forest (RF), and multiple linear regression (MLR) methods to predict the mechanical properties of three-component nanocomposite films. These films consisted of polyvinyl alcohol (PVA),crosslinked 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) oxidized cellulose nanofibers (TOCNFs), and either ammonium zirconium carbonate (AZC) or glyoxal (Gx). The prediction models were trained using the mechanical properties of mono-component TOCNF films and two-component nanocomposites as inputs. The results showed that machine learning methods, especially artificial neural networks (ANN), were effective in predicting the mechanical properties of the three-component nanocomposite films. ANN displayed the lowest error rates, outperforming random forest (RF) and multiple linear regression (MLR). The study demonstrated that the mechanical properties of TOCNF films could be adjusted by crosslinking the cellulose network with the selected additives. Furthermore, physically or chemically crosslinked hybrid films, with optimized amounts of crosslinkers, exhibited significantly higher tensile strength and Young's modulus compared to pure nanocellulose films. The use of machine learning in this context potentially allows for a more efficient prediction of the interaction between different components in the nanocomposite, leading to cost and time savings in the experimental design process.

In 2020, Steph-Yves Louis and al. Louis, Zhao, Nasiri, Wang, Song, Liu and Hu (2020); focused on the development of a new graph neural network model, named GATGNN, aimed at improving the prediction of inorganic materials' properties. GATGNN stands out through the utilization of augmented graph attention layers (AGAT) and a global attention layer. These AGAT and global attention layers enable the model to learn both the local relationships between neighboring atoms and the overall contribution of each atom to the material's property. The goal is to achieve significantly enhanced prediction performance across various tested properties. The conducted experiments demonstrate that the proposed method outperforms state-of-the-art existing graph neural network (GNN) models. Additionally, the method provides a measurable insight into the correlation between atoms and their material properties. In summary, the article explores an innovative approach using graph neural networks with global attention to enhance the prediction of material properties .

In 2020, Mirza Sahaluddin and collaborators Sahaluddin et al. (2020) have already used the Support Vector Regression (SVR) model for accurately estimating the density of nanofluids composed of nitrides in ethylene glycol presents

a promising avenue. The study of the thermo-physical properties of these nanofluids offers considerable advantages. Although some studies have examined the thermo-physical properties of nitride-based nanofluids, the application of machine learning techniques to estimate the density of nitride-based nanofluids has not been explored. The proposed model is built on experimental data derived from aluminum nitride, titanium nitride, and silicon nitride nanoparticles. The merit of the proposed model lies in the ease of determining its inputs, such as mass fractions, temperature, nanoparticle size, and molecular mass. To highlight the performance of the proposed model, its results were compared with those of mixture theory. The proposed model produces significantly more accurate results than mixture theory, recommending it as a viable method for estimating nanofluid density. The proposed method was evaluated on two datasets through training, validation, and test sets, as well as an unseen dataset. A total of 2250 different initial weights and the number of neurons in the hidden layer for the proposed Artificial Neural Network (ANN) models were considered and compared to identify the optimized ANN models. Mean Squared Error (MSE) and the Coefficient of Determination (R2) were employed to select the top 20 ANN models for further analysis.

In 2021, Tien-Thinh Le Le (2021) dedicated himself to the development and creation of a practical Machine Learning (ML)-based model for predicting the tensile strength of polymer/carbon nanotube (CNT) composites. A database was compiled from available literature, consisting of 11 input variables. These variables were chosen due to their significance in predicting the tensile strength of nanocomposites, including the category of polymer matrix, the mechanical attributes of the polymer matrix, the physical traits of CNTs, the mechanical attributes of CNTs, and integration factors like CNT weight ratio, the method of CNT surface modification, and processing techniques. The developed model utilizes Gaussian Process Regression (GPR) for prediction and was optimized through a parametric study. Error measurement criteria such as the correlation coefficient (R),during the GPR model training process, the utilization of Willmott's index of agreement (IA), regression slope, Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) was incorporated. The GPR model demonstrated good performance in both the training and testing phases. The article also emphasizes employing uncertainty analysis to calculate confidence intervals for predictions. Additionally, it investigates and discusses the predictive performance of the GPR model across various input variable ranges. To facilitate practical use, a Graphical User Interface (GUI) was created using Matlab to forecast the tensile strength of nanocomposites.

Masoud Fetanat et al.(2021) Fetanat et al. (2021); proposed a method which can offer valuable insights into formulating the penetration flow and pollutant rejection, considering the effects of each experimental condition on nano-composite filtration membranes, without conducting actual experiments, which are both time-consuming and costly. This concerns Machine Learning applied to Thin-Film Nanocomposites (TFN). Nine independent variables are provided to Artificial Neural Networks (ANN), and the penetration flow as well as pollutant rejection are derived as outputs from the ANN .

Aowabin Rahman and al. Rahman, Deshpande, Radue, Odegard, Gowtham, Ghosh and Spear (2021) discuss in their article the use of a machine learning framework to predict the shear strength of interfaces between carbon nanotubes (CNTs) and polymers, based on molecular dynamics simulation data. In modern aerospace applications, there is a growing need for lightweight materials possessing exceptional strength and stiffness. Carbon nanotube (CNT)-reinforced composites hold promise in meeting these demands. However, a key limitation hindering the potential of these composites lies in the restricted capacity to transfer load between CNTs within a polymer matrix, primarily due to the low shear strength observed at the molecular scale of the CNT-polymer interface. While molecular dynamics (MD) simulations offer a means to investigate this interface, their computational demands are considerable. Consequently, exploring a sufficiently extensive design space for interface modifications and optimization solely through MD simulations presents a significant challenge. This encourages the adoption of surrogate models as a means to efficiently correlate CNT-polymer interfacial configurations with interfacial strength. In this study, the researchers introduce a machine learning surrogate model trained using MD simulations of functionalized CNT-epoxy systems and their corresponding interfacial shear strengths. The proposed machine learning framework comprises (i) a feature representation approach based on radial distribution functions for CNT-epoxy systems; (ii) a convolutional neural network architecture to correlate these feature representations with a measure of interfacial shear strength; and (iii) a data augmentation technique to optimize the utilization of limited training data. Despite significant variability in the pullout force, the proposed machine learning model demonstrates notable accuracy in predicting CNT pullout forces.

Also, in 2021 Tien-Thinh Le and Minh Vuong Le Le and Le (2021) propose the development of a prediction model based on a neural network (NN) to estimate the tensile modulus of carbon nanotube (CN) and polymer composites.The methodology relies on the utilization of an experimental database comprising 282 configurations, gathered from various available sources. The incorporated input variables in the dataset encompass multiple aspects, including the mechanical properties of separate phases, the density of the polymer matrix, the processing method, the geometry of CN, the modification method at the CN surface, etc. The model aims to predict the tensile modulus of the nanocomposite.The article also emphasizes the conduct of parametric studies to determine the optimal architecture of

the proposed NN model. These studies involved exploring various configurations and parameters of the neural network, aiming to identify the most effective structure for accurately predicting the tensile modulus of CN/polymer nanocomposites. This work thus provides a significant contribution to the understanding and modeling of the mechanical properties of these composite materials.

In 2021, Itamar A. Shabtai, Laurel M. Lynch, and Yael G. Mishael Shabtai, Lynch and Mishael (2021) concentrated on the application of clay-polymer nanocomposites (CPNs) as sorbents in water treatment. Their main goal was to present recent developments in CPN research using a meta-analysis approach, with a focus on outlining key steps to bridge the gap between fundamental research and the practical implementation of these materials. The meta-analysis drew insights from 99 research articles on CPNs and 8 review articles covering other commonly studied sorbents. CPNs exhibited notable adsorption capacities across various inorganic and organic pollutants, encompassing heavy metals, oxyanions, and dyes. Employing statistical techniques such as principal component analysis, analysis of variance, and multiple linear regressions, their study delved into understanding how CPN properties and pollutant characteristics influence adsorption mechanisms. The findings underscored the significant impact of factors like CPN fabrication method, polymer functional groups, and pollutant properties on adsorption performance. The review emphasized the necessity to shift from fundamental research to practical application, highlighting existing limitations, such as the absence of comparisons with commercial sorbents, real-world trials, regeneration, and cost assessment. The authors advocated for the development of intelligent and functional CPNs, emphasizing the importance of in-depth characterization of material properties to optimize their overall performance.

Divya P. Barai and colleagues Barai, Bhanvase and Pandharipande (2022) article introduces an advanced artificial neural network (ANN) model designed to predict the thermal conductivity of nanofluids incorporating reduced graphene oxide (rGO) and metal oxide nanocomposites. The model relies on data from five distinct water-based nanofluids containing rGO–metal oxide nanocomposite particles sourced from existing literature. Four key input variables—nanocomposite molecular weight, average particle size, nanofluid concentration, and temperature—are considered to forecast the thermal conductivity of the nanofluids. Two ANN models were developed within the same framework: one using a comprehensive dataset of 185 data points and the other segmenting the data into training and testing subsets. Both models demonstrated a strong alignment with experimental data, achieving an impressive R2 value of 0.956 for the testing subset. Additionally, the difference between predicted and actual thermal conductivity values across all data points was minimal, with an average residual of $\pm$0.048 W/mK. The article highlights the crucial role of heat transfer in various industrial and domestic applications, exploring the potential of nanofluids—fluids infused with solid nanoparticles—to enhance heat transfer capabilities due to their elevated thermal conductivity. The focus of the study is on nanofluids featuring rGO–metal oxide nanocomposites, which exhibit promising attributes for thermal transport. By examining factors such as particle size, concentration, and temperature, the article provides insights into the nuanced determinants of nanofluid thermal conductivity, emphasizing the need for precise modeling to accurately anticipate these properties.

Hossein Adel et al. Adel, Palizban, Sharifi, Ilchi Ghazaan and Habibnejad Korayem (2022) focuses on the prediction of mechanical properties in carbon nanotube (CNT)-reinforced cementitious nanocomposites using interpretable machine learning models. The critical mechanical properties under examination are compressive strength and flexural strength, pivotal for assessing the effectiveness of CNT incorporation in cementitious nanocomposites. The article initiates by underscoring the importance of compressive and flexural strength in evaluating construction materials, highlighting the associated costs and time constraints in conducting numerous mechanical tests on samples at various ages. Given the complexity of contemporary construction materials, traditional experimental and statistical models prove inadequate for such intricate materials, necessitating the adoption of machine learning techniques. Subsequently, the article outlines the data collection process from literature, encompassing compressive and flexural strength properties of cementitious nanocomposites, along with pertinent input variables such as water-to-cement ratio, its type, content,length and diameter. The overarching objective of the article is to propose an approach rooted in interpretable machine learning for predicting the mechanical properties of CNT-reinforced cementitious nanocomposites. Additionally, the study aims to assess the significance of input variables, providing insights into the factors influencing these mechanical properties..

Meng Shi and al.Shi, Feng, Li and Guo (2022) uThe study focuses on leveraging machine learning to optimize nanocomposite materials for electromagnetic interference (EMI) shielding. Carbon-based nanocomposites, combining conductive fillers with polymers, are explored due to their superior electrical conductivity and lightweight characteristics. However, the traditional development of these composites often relies on researchers' experience and repetitive experiments, leading to prolonged development cycles and increased costs.

In this research, the scientists adopt a machine learning approach to establish a swift prediction model for shielding effectiveness and analyze crucial factors and guidelines in material design. They assemble a dataset containing features of carbon based conductive particles and polymer nanocomposites intended for EMI applications. Employing a technique termed "Weighted Average Ensemble," the

researchers merge five distinct base models on the dataset, revealing that the final prediction model surpasses the performance of all individual base models. Additionally, the article delves into the analysis of feature importance through variable importance rankings. Rules influencing EMI, based on critical features, are explored using model-agnostic techniques such as partial dependence plots and individual conditional expectation plots. The resulting prediction model is deemed a valuable tool for swiftly estimating shielding performance. Alongside the identified rules, the model can guide material development, shorten development cycles, and reduce costs.

In their study, Liu et al. (2022) Liu, Vu-Bac, Zhuang, Fu and Rabczuk (2022) explore a comprehensive stochastic methodology integrating machine learning techniques across various scales to forecast the macroscopic thermal conductivity in carbon nanotube-reinforced polymeric composites (CNT-PCs). They introduce seven distinct machine learning models, including Multivariate Adaptive Regression Splines (MARS), Support Vector Machine (SVM), Regression Tree (RT), Bagging Tree (Bag), Random Forest (RF), Gradient Boosting Machine (GBM), and Cubist. These models form the basis of a stochastic modeling framework designed to establish the connection between uncertain input parameters and the target output, which is the composite's macroscopic thermal conductivity. The study incorporates Particle Swarm Optimization (PSO) for hyperparameter optimization, aiming to identify global optimal values and thereby reduce computational expenses significantly. Additionally, the article evaluates the computational costs and model complexities of the different methodologies, highlighting their respective advantages and limitations. The authors contend that their proposed stochastic integrated machine learning approach, which accounts for uncertainties, constitutes a significant advancement in the computational design of novel composites for applications pertaining to thermal management. In summary, the study proposes a methodology that utilizes machine learning to predict the thermal conductivity of CNT-PCs while considering associated uncertainties.

Elizabeth Champa-Bujaico et al. Champa-Bujaico, García-Díaz and Díez-Pascual (2022) address in their article the application of machine learning (ML) for the prediction and optimization of polymeric nanocomposites, providing a state-of-the-art overview.In recent times, there has been a notable surge in scientific and industrial interest towards polymeric nanocomposites. This heightened attention stems from the remarkable enhancements achieved through the synergistic integration of properties derived from both a polymeric matrix and organic or inorganic nanomaterials. These materials demonstrate exceptional mechanical strength, resilience, stiffness, electrical and thermal conductivity, improved flame resistance, as well as heightened barrier properties against moisture and gases. The unique design possibilities of nanocomposites offer advantages for developing multifunctional materials tailored for specific applications. The article explores how machine learning serves as a powerful predictive tool for data-driven multiphysics modeling, providing insights beyond traditional computational and experimental analyses. It covers prediction, optimization, feature identification, and uncertainty quantification using various ML algorithms in the context of polymeric nanocomposites. The discussion includes examples of ML applications in predicting properties. Polymeric nanocomposites with multifunctional properties find applications in sectors such as aerospace, automotive, civil engineering, marine, and other technologically demanding industries. However, investigating their physical, chemical, and mechanical behavior under diverse environmental conditions is complex and time-consuming. The article emphasizes that machine learning, trained on extensive data, proves to be a powerful tool for predicting properties and accelerating the design process. Various ML algorithms such as artificial neural networks (ANN), ANFIS, MLP, CNN, GA, GP have successfully established a mapping between input features and target properties, demonstrating a strong correlation with experimental values. The article notes a recent increase in ML-based studies for polymeric nanocomposites, particularly in predicting mechanical properties. However, the accuracy and generalization of ML models depend on the quantity and quality of available data, which remains limited for polymeric nanocomposites. Efforts are underway to address this limitation through the creation of specific databases and advanced methods, including natural language processing (NLP) techniques.

In 2022, the study conducted by Kyungmin Baek et al.Baek et al. (2022) is based on Deep Neural Networks (DNNs): These methods not only surpass the limitations of the traditional cluster density-based approach, which offers only a singular property value for a given cluster density but also excel over conventional methods in assessing the electromechanical characteristics of nanocomposites. Even with fewer features considered, a basic Graph Convolutional Network (GCN) has demonstrated superior capability compared to a complex Artificial Neural Network (ANN). The GCN proves more adept and user-friendly in evaluating the impact of nanoparticle distribution and agglomeration. Moreover, the GCN efficiently provides mechanical and electrical properties for a large Representative Volume Element (RVE) without sacrificing precision.

NX Ho, TT Le, and MV Le Ho, Le and Le (2022) presented a paper focusing on the development of a model based on Artificial Intelligence (AI) is introduced to forecast the Young's modulus of polymer/carbon-nanotube (CNT) composites. AI is suggested as a remedy to address the challenges associated with investigating the properties of novel composite materials, such as the time-consuming nature of experimental studies or the resource-intensive requirements of numerical methods. An Artificial Neural Network (ANN) model is chosen, and its architecture is

fine-tuned through a parametric analysis. The primary aim of this research is to validate the efficacy of the proposed AI approach for nanocomposites while laying the groundwork for further enhancements in terms of computational efficiency and resource utilization. The findings indicate that the proposed model demonstrates exceptional performance in both training and testing phases, achieving correlation coefficients of 0.986 for the training phase and 0.978 for the testing phase.

Recently, Kyungmin Baek and all. (2023) Baek et al. (2023) have done a multiscale modeling involves calculations rooted in the principles of density functional theory (clustering density) and numerical homogenization based on the Finite Element (FE) method.

F. Chen andal.Chen, Weng, Wang, Wu, Ma, Pan and Ding (2023) address in their article the development of an adaptive framework that merges domain-specific expertise with machine learning to expedite the optimization of flame-retardant high-composite materials. Research in the polymer field has widely adopted machine learning methods, encompassing both linear and nonlinear algorithms, which have brought about significant changes in areas such as metals, catalysts, polymers, etc. However, the majority of these polymer-related studies often focus on the molecular design of the polymers themselves or simulations, rather than exploring the composition of functional polymer-based composites.The article proposes an adaptive framework that integrates domain-specific knowledge with machine learning techniques to expedite the optimization of composites with high flame resistance. The adaptive framework utilizes data resources from three distinct approaches: experiments, handbooks, and published papers. These sources are cleverly employed for training, feedback, or prediction purposes. A comprehensive feature engineering of flame-retardant polymer-based composites is presented and classified in detail. Four machine learning methods, comprising conventional linear regression (Lasso and Ridge), nonlinear artificial neural networks (ANN), and their combination (L-ANN), are compared within the adaptive framework. Models based on the L-ANN method for the Limiting Oxygen Index (LOI) demonstrate higher accuracy in two runs, thereby guiding new experiments with high flame resistance and effectively predicting across different flame retardants to address the intuitive trial-and-error problem. The ultimately optimized models from this adaptive framework could also prove beneficial for artificial intelligence applications in the field of engineering flame-retardant polymer-based composites.

The article Champa-Bujaico, Díez-Pascual, Redondo and Garcia-Diaz (2024) addresses the optimization of mechanical properties in multiscale hybrid polymer nanocomposites. The authors employ a combination of experimental techniques and machine learning (ML) for predicting the mechanical characteristics of nanocomposites composed of poly(3-hydroxybutyrate) (P3HB) reinforced with different concentrations of multi-walled carbon nanotubes (MWCNTs), WS2 nanosheets, and sepiolite (SEP) nanoclay. The nanocomposites were fabricated through solution casting, and scanning electron microscopy (SEM) images revealed a uniform and random dispersion of the three nanofillers within the matrix. A substantial improvement in stiffness, reaching 132% with the addition of 1:2:2 wt% SEP:MWCNTs:WS2, was achieved through a synergistic reinforcement effect. However, the increases in strength were moderate, up to 13.4%. Various ML models, including Recurrent Neural Network (RNN), RNN with Levenberg's algorithm (RNN-LV), decision tree (DT), and Random Forest (RF), were applied to predict these properties.Results indicate that the RNN-LV model with 3 hidden layers and 50 neurons in each layer performed best in predicting Young's modulus, while the RF model with 100 estimators and a maximum depth of 100 was optimal for predicting tensile strength. The models exhibited strong correlations and minimal errors, underscoring their precision. The authors conclude that these ML models serve as potent tools for optimizing the mechanical properties of multiscale hybrid polymer nanocomposites, offering time and resource savings in the experimental characterization process. This is particularly beneficial in the development of new nanocomposites, providing increased efficiency in the design and virtual assessment of diverse material combinations.

3. Proposal of new techniques to improve nanocomposite classification

Our methodology relies on the exploitation of a specific dataset associated with polymer-based nanoparticles. First, we implement artifical neural network models, to classify the various variants of nanoparticles based on their characteristics.

3.1. Dataset

We normalized the database used in our study, which is based on 8 properties of nanocomposites: electrical conductivity, elastic modulus, tensile strength, Concentration, molecular weight, temperature, nanomaterial size, and flexural strength . Indeed, normalization involves transforming measurements values to make them more suitable for a specific analysis or processing. It typically entails adjusting the data to follow a normal distribution, where mean and standard deviation are key parameters. This step is crucial in machine learning as it makes data more comparable and simplifies subsequent processing (see Appendix).

3.2. Artificial Neural Networks (ANN)

An artificial neural network (ANN) represents a common form of supervised Deep Neural Network employed for regression or classification tasks. Due to its inherent clarity in elucidating the learning process and principles, the ANN stands out as the favored deep learning model, particularly for researchers who are newcomers to the realms of machine learning or Deep Learning Baek et al. (2022). Figure 2

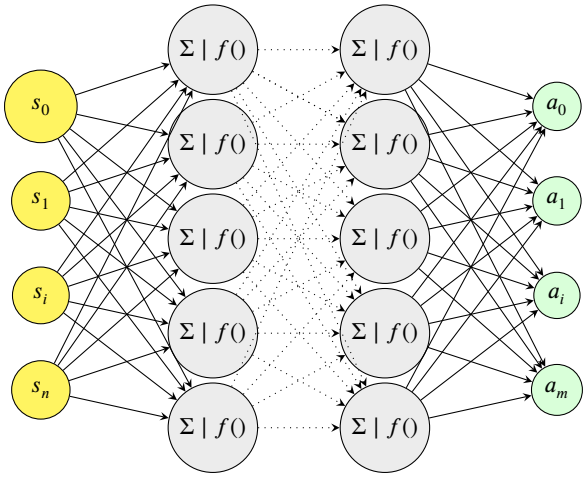

Figure 2: The overall architecture of the artificial neural network (ANN)

illustrates the overall architecture of the artificial neural network (ANN), comprising an input layer, hidden layers, and an output layer. Within the input layer, a user-selected set of features is arranged in a column to effectively classify or predict targets. The activation function introduces the nonlinearity during forward or backward propagationsBaek et al. (2022).

3.2.1. ANN classification with conductivity property (P1)

Figure 3 displays the classification results obtained using the *ANN* model for classification.

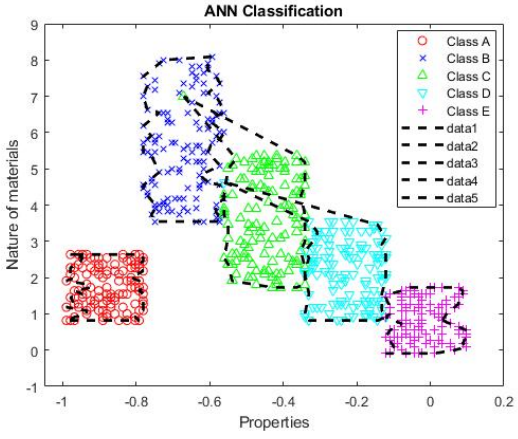

Figure 3: ANN classification

In fact, the x-axis represents the different properties and the y-axis represents the nature of materials. We find five classes from A to E which are imbalanced. This emphasizes how the used model handles minority or majority classes. Then, the figure 4 shows the evolution of the Mean Square Error in phase of train (blue), validation (green), test (red) and the best validation performance is 0.073288 at epoch 6.

These results are more explained through figure 4.

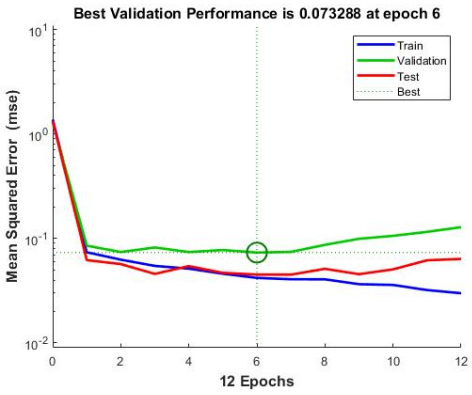

Figure 4: MSE with conductivity

However , Baek et al. Baek et al. (2022) obtained for the same property a value of MSE equal to 0.004.

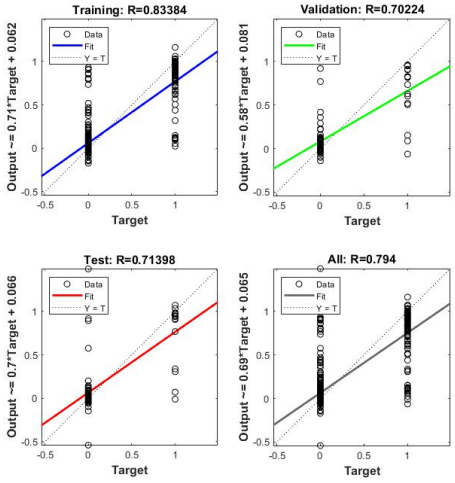

Figure 5: Regression with conductivity

Analyzing the regression of the Artificial Neural Network (ANN) model for classification through the variation of the coefficient of correlation (R) provides valuable insights into the relationship between the predicted probabilities and the actual class labels. The coefficient of correlation, often denoted as R or R-squared, quantifies the degree of linear dependence between the predicted and observed values (figure 5)

3.2.2. ANN classification with four properties (P4..P7)

We simulated for establishing the difference between our method and Divya et al(2022)Barai et al. (2022). Fixing four properties from P4 to P7, we obtained the Mean square error (figure 6) which has the best validation performance at epoch 3 and it is equal to 0.025242. We use some metric values such as accuracy , R^2 and RMSE which are explained in table 2.

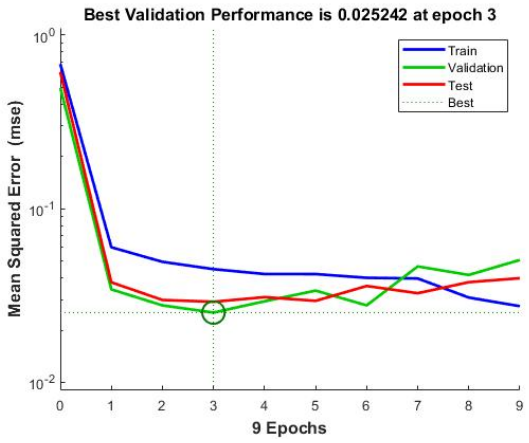

Figure 6: MSE with P4..P7

Table 2
Table of Simulation Results for *ANN* for 4 properties

	Accuracy	R^2	RMSE
Divya et al (2022)	95% ±5	0.956	0.077
ANN (Our work)	90.63% ±4	0.8680	0.1038

3.2.3. ANN classification with eight properties

Figure 7 shows the Mean Squared Error (*MSE*) of the *ANN* for classification. In this particular scenario, the lowest Mean Squared Error (MSE) is observed and reached its peak at 0.00036021 at 16 epochs. This value is affecting the precision of the model's predictions during training, testing, and validation.

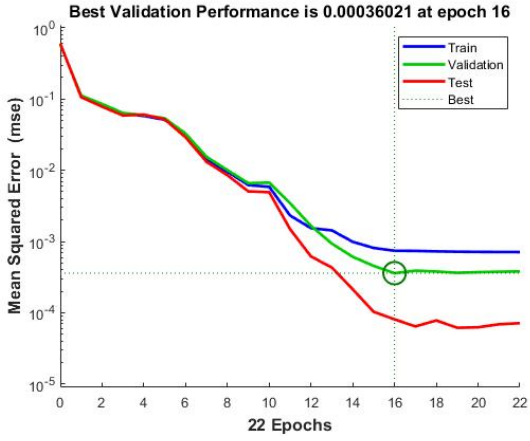

Figure 7: MSE with eight properties

The histogram is a visual representation of numerical data, where the data is divided into equal-width bins. Each bin is represented by a bar, with the height of the bar indicating the frequency of data points within that bin. Additionally, bins may also be referred to as "intervals," "classes," or "buckets." Moreover, we depict the histogram of errors generated by the *ANN* model during classification (figure 8). It shows the histogram of the Artificial Neural Network (ANN) error provides a visual representation of the distribution of errors generated by the network. This offers valuable insights into the model's performance by revealing the frequency of different error values. A well-interpreted histogram can highlight trends, concentrations, or areas where the model may encounter difficulties. For instance, an asymmetric distribution may indicate systematic errors, while a more uniform distribution might suggest a more balanced performance. Analyzing this histogram, we find that the error histogram with 20 Bins have already been allowed for adjustments and improvements to the model to achieve more accurate and reliable results.

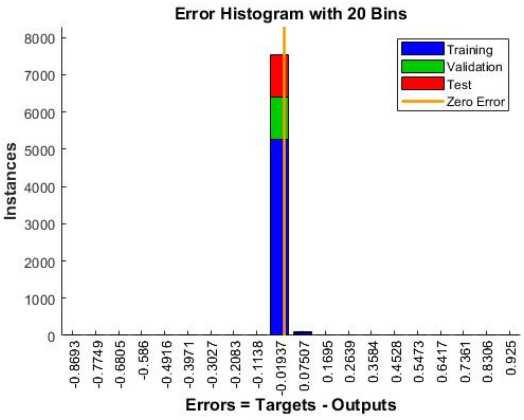

Figure 8: The histogram of the ANN network's error

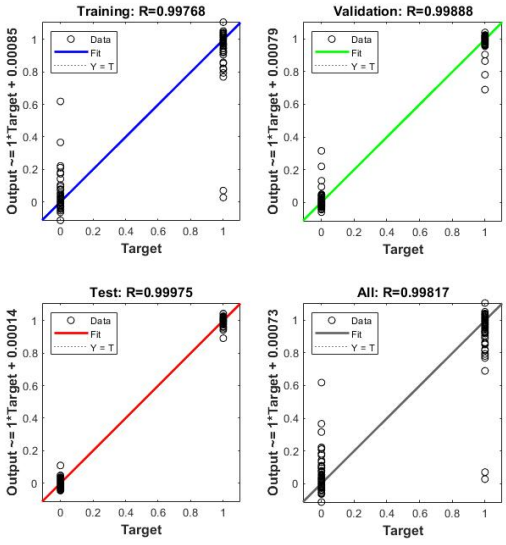

Figure 9: The regression of the ANN model

The regression of the *ANN* model for classification is as follows(figure 9). As the coefficient of correlation varies, it reflects how well the model captures the underlying patterns and associations within the classification data. A higher R

indicates a stronger linear relationship, signifying that the predicted probabilities align well with the true class labels. Understanding the differences in R coefficients during different phases of training R=0.99768, testing R=0.99975, validation R=0.99888, and overall R=0.99817 can provide insights into the performance and generalization ability of the model. After simulating the ANN with one property, four property and eight property, we try to generalize the results and why not to improve the performances by using the recurrent neural network (RNN).

3.3. Recurrent Neural Networks (RNN)

Recurrent Neural Networks (RNN) are a specific type of artificial neural networks designed to work with data that evolves over time or has a particular sequence.(RNNs) have the particularity of having special connections that allow them to remember previous steps in the sequence. This ability to retain internal memory makes them particularly well-suited for tasks such as automatic translation, text generation, analysis of time-series data, speech recognition, and other tasks involving the manipulation of sequential dataLe Bolzer, Lambert and Schnitzler (2020)Tellier (2016).

Here is an example of the general architecture of a Recurrent Neural Network (RNN).

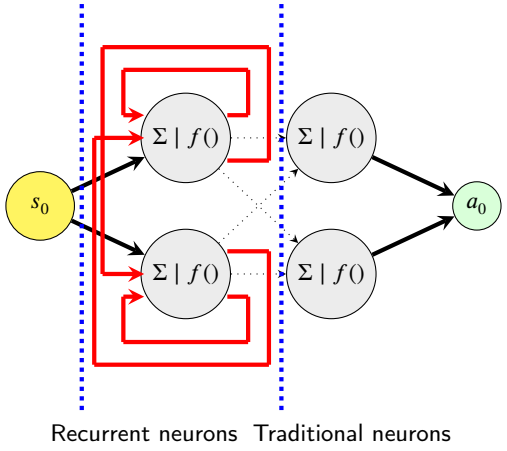

Figure 10: Recurrent neurons vs. traditional neurons

Recurrent neurons are distinguished from traditional neurons by their ability to process sequential data through recurrent connections, enabling them to maintain an internal state and consider sequential context. This characteristic makes them particularly well-suited for tasks such as natural language modeling, sequence prediction, and text generation, where data sequentiality is crucial. In contrast, traditional neurons, also known as feedforward neurons, are better suited for tasks where sequentiality is not a key element, such as image classification or pattern recognition (figure 10).

We have chosen the Recurrent Neural Network (RNN) as a model to consider. In fact, the Recurrent Neural Networks (RNNs) present distinct advantages in classification tasks when compared to conventional Artificial Neural Networks (ANNs). A key asset of RNNs is their adeptness at processing sequential data, making them well-suited for tasks where the order of input information holds significance. This sequential processing capability equips RNNs to discern dependencies and patterns in temporal data, a feature critical in numerous practical applications. In contrast to ANNs, RNNs excel in handling variable-length input sequences, showcasing adaptability in managing data of diverse lengths. Furthermore, the architecture of RNNs, particularly variants like Long Short-Term Memory (LSTM), effectively mitigates the vanishing gradient problem, enabling them to capture long-range dependencies in data. This inherent ability to retain memory of previous information positions RNNs as advantageous in tasks involving temporal relationships, providing heightened accuracy, particularly in scenarios with substantial data volume or temporal patterns. We conduct the simulation now by fixing the same property which is the weight(P5) and we use the Recurrent Neural Network with the *LSTM* method.

3.3.1. RNN classification with weight property (P5)

The result of our Mean Squared Error (MSE) test using the Recurrent Neural Network (RNN) with one property, which is the weight (P5), is approximately equal to 0.4270. Then, we elaborate a small comparison based on metric values such as MSE, R^2, and MAE. In fact, for the MSE, Champa et al.(2024) found a value of MSE equal to 0.435 which is so near to our value 0.4270. Then, Champa et al.(2024) found an $R^2 = 0.6440$ which is so near to our value R^2=0.6497. Finally, we calculate the MAE and our result is 0.3820 however, Champa et al.(2024) have got an MAE = 0.5444 which is bigger than ours.

3.3.2. RNN classification with (P4..P7) properties

We repeat the same tasks but with four properties and we have used the RNN in order to compare our work and finding results to the reserchers Diviya et al.(2022)Barai et al. (2022). The simulation of the RNN classification with these properties is explained through figure 11. Besides, we put all the comparison results into table 3.

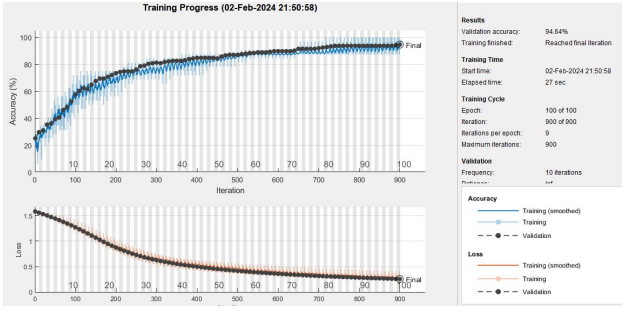

Figure 11: Precision Curve RNN with four properties

Table 3
Table of Simulation Results for RNN with four properties

	Accuracy	R^2	RMSE
Divya et al (2022)	95% ± 5	0.956	0.077
RNN (Our work)	94.64% ± 2	0.9701	0.0610

3.3.3. RNN classification with eight properties

We repeat the simulation, now we are fixing the eight properties, and employ the recurrent neural network with the *LSTM* method. Figure 12 describes the material classification into 5 categories: classes A, B, C, D, and E.

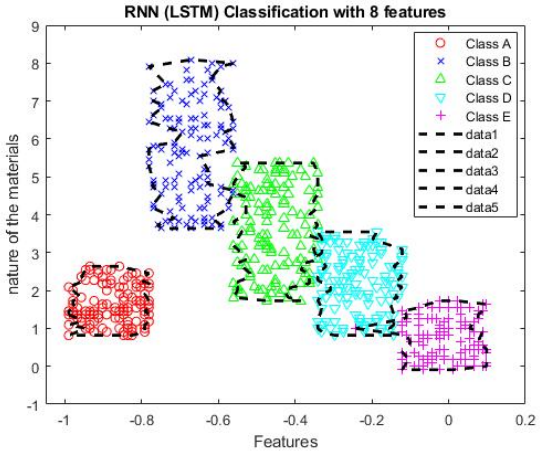

Figure 12: RNN classification with eight properties

Then, we put a curve presenting the accuracy of the recurrent neural network with eight properties P1..P8 (figure 13). The accuracy has been well improved when we have changed the algorithm (table 4). It's around 97%.

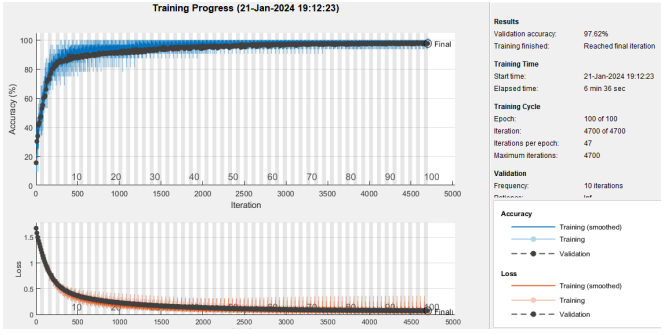

Figure 13: Precision Curve RNN with eight properties

The result of the Mean Squared Error (*MSE*) of the RNN network for eight properties is displayed. Indeed, as we increase the number of properties, we assess the loss of information. However, this value is now equal to 0.0198.

Table 4
Table of Simulation Results for RNN with eight properties

	Results for *RNN*
Accuracy	97.62% ± 1
MSE	0.0198
Duration	6m and 36s

Based on the outcomes of recurrent neural networks, we have noted that an increase in the number of features leads to an enhancement in result accuracy. Although the simulation requires more time, the mean squared error (MSE) is reduced.

3.4. Monte Carlo Method

Monte Carlo is a numerical method that employs random simulations to estimate unknown quantities or solve complex mathematical problems. It is often used when analytical calculations are difficult or impossible. This method is named after the famous Monte Carlo casino in Monaco, renowned for its games of chance. In general, what is referred to as a Monte Carlo method is a technique aimed at calculating a deterministic quantity through a random process Burch, Najm, Yang and Trick (1993) Kantas, Doucet, Singh and Maciejowski (2009).
Then, We propose to incorporating the previously defined statistical tool (Monte Carlo). The latter will be associated with the recurrent network to observe improvements in the estimated classification.

3.4.1. RNN- Monte Carlo classification with eight properties

Indeed, the classification with the RNN-Monte Carlo with eight properties is determined in the figure 13. Then, The figure 14 depicts the evolution of accuracy. We aim to increase this value compared to the previous simulations.

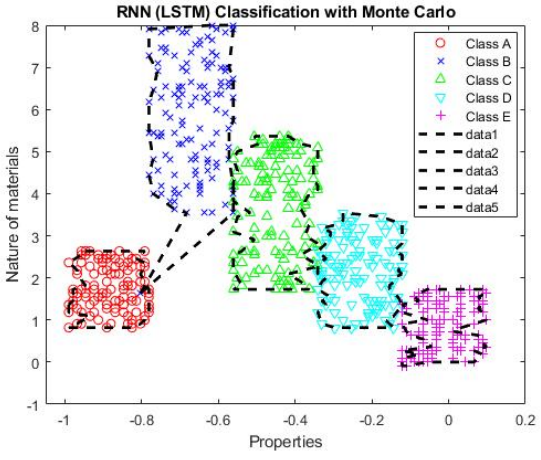

Figure 14: Classification with RNN-Monte Carlo

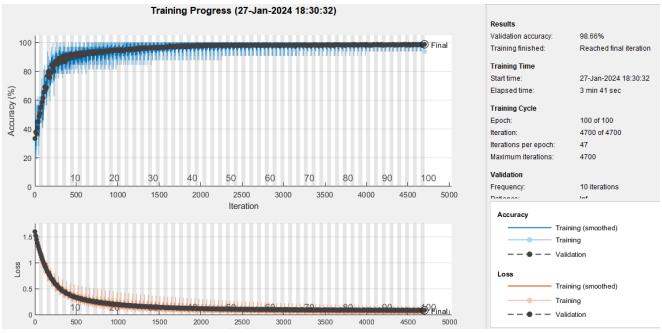

Figure 15: Precision curve RNN-Monte Carlo

Table 5
Table of Simulation Results for RNN- Monte Carlo

	Results for RNN- Monte Carlo
Accuracy	98.66 % ± 1
MSE	0.0321
Duration	3m and 41s

We observe that despite the incorporation of multiple data, the performance in terms of accuracy remains unchanged, if not improved, for the RNN algorithm with eight properties.

We still have a trade-off between performance during the classification phase. In fact, the introduction of the Monte Carlo method to RNN leads to a substantial reduction in execution time, establishing itself at 3 minutes and 41 seconds. This result suggests that the increased efficiency of the Monte Carlo method contributes to optimizing the trade-off between execution speed and accuracy of RNN in the context of the studied classification (table 5). On the contrary, RNN requires 6 minutes and 36 seconds to accomplish the same task, highlighting a significant increase in computational time compared to ANN, yet compensated by a considerable improvement in accuracy. Indeed, we have gained in terms of accuracy, but the value of the mean squared error has increased slightly. However, it remains quite interesting.

4. Comparison and Discussion

To validate our IA models, we must calculate different kind of metrics which are more explicated in the table 6.

It's certainly known that we have a compromise between the different performances which are already calculated and then evaluated based on metric values. For example, when we change the IA algorithm from ANN to RNN and RNN-Monte Carlo, we have improved the accuracy but not the duration. This comparison is more explained in the next paragraph.

Table 6
Results between our three approaches with eight properties

	ANN	RNN	RNN-Monte Carlo
Accuracy	-	97.62% ± 1	98.66 % ± 1
Duration	6s	6m and 36s	3m and 41s
MSE	0.00036	0.0198	0.0321
MAPE	0.20%	0.89%	0.48%
MASE	2.4962	2.00	2.25
RMSPE	19.43%	3.73%	4.93%
RMSE	0.0258	0.03522	-
RMSSE	0.45%	0.03%	0.02%
R	-	0.9955	0.9970
R^2	0.1363	0.9885	0.9901
CP	-	0.8910	0.6887

4.1. Comparison of Accuracy results

Following our meticulous analysis, the obtained accuracy results reveal distinct performances among the various studied approaches (figure 16). We take the properties from P4 to P7, so we have four to test with; especially when we compare with Divya et al.Barai et al. (2022). This last has an amount of accuracy equal to 95% ±5. We integrate first the ANN that is achieving an accuracy of 90.6% ±4. Besides, The Recurrent Neural Networks (RNN) exhibit a remarkable accuracy of 94.64% ±2 , emphasizing its ability to significantly outperform, see table 7.

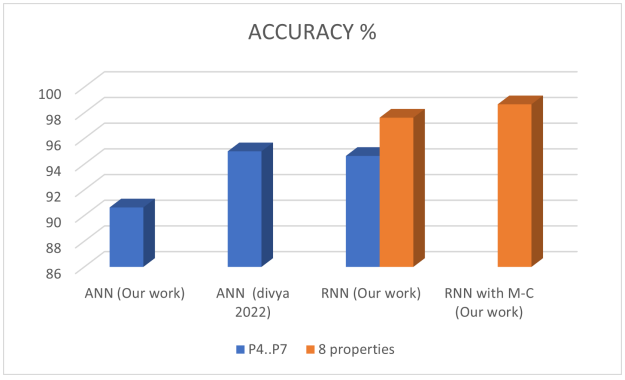

Figure 16: Accuarcy comparison

Table 7
Table of accuracy results with four properties

	Accuracy
Divya et al (2022)	95% ±5
ANN (Our work)	90.63% ±4
RNN (Our work)	94.64% ±2

4.2. Comparison of MSE results

Fixing eight properties, we add the RNN-Monte Carlo. In the evaluation of performance, it becomes essential to scrutinize the results of the Mean Squared Error (MSE). The MSE values obtained for artificial neural networks (ANNs), recurrent neural networks (RNNs), and RNN- Monte Carlo method are recorded as 0.00036, 0.0198, and 0.0321, respectively (figure 17). It is noteworthy that the ANN exhibits the lowest MSE at 0.00036, thus demonstrating the superiority of the ANN in terms of MSE. This indicates a better fitting of the model to the data compared to other approaches. After the meta analysis study, for one property P1 which is the conductivity we note that the work of Baek et al.Baek et al. (2022) has an MSE value equal to 0.004 for ANN and 0.036 for CNN despite of our work with the same property, we notice a value of 0.0732. Besides, we compare the MSE results with an other property P5 which is the weight with champa et al.Champa-Bujaico et al. (2024). We find in her work two MSE values for each IA method. So, for RNN, MSE is equal to 0.435 and for RNN-LV it's 0.3932, see table 8.

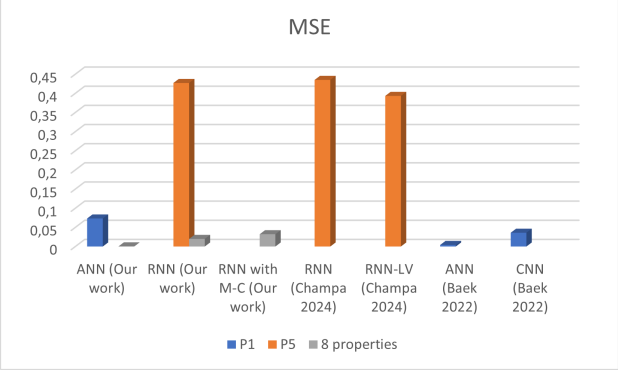

Figure 17: MSE comparison

Table 8
Results of MSE

	P1	P5	P1..P8
ANN (Our work)	0.0732	-	0.00036
RNN (Our work)	-	0.427	0.0198
RNN M-C (Our work)	-	-	0.0321
RNN (Champa 2024)	-	0.435	-
RNN L-V (Champa 2024)	-	0.3932	-
ANN (Baek 2022)	0.004	-	-
CNN (Baek 2022)	0.036	-	-

4.3. Comparison of R^2 results

Within our comparative study, the values of the coefficient of determination (R^2) shed instructive light on the performance of recurrent neural networks (RNNs) with eight properties P1..P8. The RNN demonstrates an impressive coefficient of determination of 0.9885. However, the integration of the Monte Carlo method with the RNN leads to a notable improvement, propelling the coefficient of determination (R^2) to 0.9901. In contrast, artificial neural networks (ANNs) exhibit a significantly more modest R^2 results, standing at 0.1363 (figure 18). This disparity further underscores the enhanced efficiency of the approach we advocate, highlighting the improved capacity of the model to explain variance in the data. Besides, we simulate with four properties P4..P7.
We find R^2= 0.868 using ANN and R^2=0.9701 using RNN; comparing these results with Divya et al. (2022)Barai et al. (2022), R^2=0.956.
Finally, we fix one property P5 which is the weight, we simulate using the RNN and we find R^2= 0.6497.
However, champa (2024)**?** has two values equal to 0.6640 using RNN and 0.6784 using RNN L-V, see table 9.

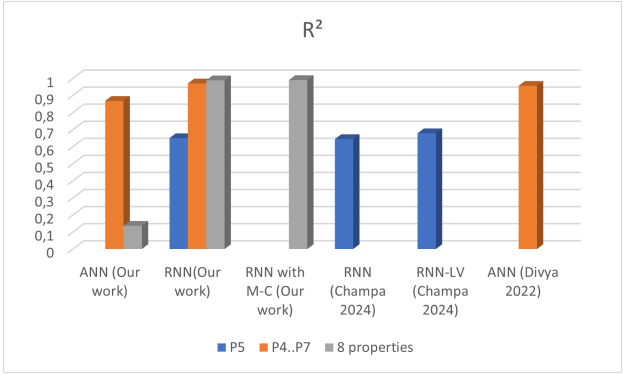

Figure 18: R^2 comparison

Table 9
Results of R^2

	P5	P4..P7	P1..P8
ANN (Our work)	-	0.868	0.1363
RNN (Our work)	0.6497	0.9701	0.9885
RNN M-C (Our work)	-	-	0.9901
RNN (Champa 2024)	0.6440	-	-
RNN L-V (Champa 2024)	0.6784	-	-
ANN (Divya 2022)	-	0.956	-

4.4. Discussions

We summarize all the results in the table 10 resuming the previous researchers' results especially the metric findings such as the MSE,the ACCURACY, R, R^2, MAE and RMSE. The number and nature of properties are introduced also and they are added to the used IA algorithm. Our principal target is to compare our actual work with others. The simulation results will be so interesting in this case. For us, we begin our simulation with the same property P1, then P5 and compare them. Besides, we simulate with P4..P7 and finally we extend to eight properties.

Besides, we integrate the Monte Carlo to our RNN in order to improve especially the duration and a little bit the accuracy. This fact was already been explained previously. In addition, the strength point of our work is the use of three approches ANN, RNN and RNN-Monte Carlo with different combination of properties in order to have the same condition with other researchers and so we can compare our work with them. We noted that the use of RNN has many benefits especially when it is added to Monte Carlo. In fact,we have chosen this kind of IA model due to its benefits such as the sequential Data Processing, the variable-Length Input, the memory and Context Preservation and the Flexibility in Architecture.

We found that all previous studied works have used the linear regression as shown previously in the meta analysis study for example in the paragraph 2.1. The use of nonlinear regression is more significant because the borders are presented as curves and they aren't straight.It's known, that the various RNN architectures, such as Long Short-Term Memory (LSTM) addresses the vanishing gradient problem and enable RNNs to better capture long-range dependencies. In fact, when the amount of data increases for artificial neural networks (ANN), issues of accuracy may arise. Conversely, recurrent neural networks (RNN) stand out for their superior accuracy in such situations. So, we propose already to integrate the logistic non linear regression in order to improve some performances concerning the classification target as a perspective in future studies.

Comparing our results with those of Divya et al. (2022), who utilized an artificial neural network (ANN) with four distinct properties (P4 to P7), we observe significant differences in terms of accuracy, coefficient of determination (R^2), and root mean square error (RMSE). Divya et al. (2022) reported results of 95% $\pm$5 for accuracy, an R^2 of 0.956, and an RMSE of 0.077. In contrast, our results were close, with an accuracy of 90.63% $\pm$4, an R^2 of 0.8680, and an RMSE of 0.1038. However, when we employed the RNN with four properties, our performance significantly improved, showing an accuracy of 94.64% $\pm$2, an R^2 of 0.9701, and an RMSE of 0.0610. This increase in accuracy, R^2 value, and the reduction of RMSE attest to the enhanced efficiency of our model.

Consequently, we decided to expand our analysis by increasing the number of properties to eight, resulting in an accuracy of , 97.62% $\pm$1an R^2 of 0.9885, and an RMSE of 0.0352. Finally, the incorporation of the Monte Carlo method further improved our performance, bringing the accuracy to 98.66% $\pm$1 and the R^2 to 0.9901. These results demonstrate the robustness and efficiency of our approach, reinforcing the relevance of using neural networks for this type of modeling.

To recapitulate the research article methodology and findings,we are presenting a summary in Table 10 that focuses on metrics from previous studies. The primary objective is to compare the current work with previous research, starting with simulations on specific properties and gradually extending to eight. The integration of Monte Carlo into the Recurrent Neural Network (RNN) is highlighted for improved duration and accuracy. The study employs three approaches (ANN, RNN, and RNN-Monte Carlo) with various property combinations for meaningful comparisons. The benefits of RNN, especially when combined with Monte Carlo, are emphasized, addressing sequential data processing and architecture flexibility.

Indeed, the other researchers employed one to four properties, and they never reached the same level of accuracy as we did with eight. The use of nonlinear regression in RNN architectures like Long Short-Term Memory (LSTM) is justified, contrasting with previous linear regression approaches. Challenges in Artificial Neural Networks (ANN) as data increases are noted, with RNNs standing out for superior accuracy. To conclude, we have proposed the integration of logistic nonlinear regression for enhanced classification in future studies, demonstrating a forward-looking perspective.

Table 10
Metric comparisons based on meta analysis study

RESEACHERS	ALGORITHM	Properties	MSE	ACCURACY	R	R^2	MAE	RMSE
Binh et al. (2019)	SVM	P14..P27	-	-	-	0.909	0.021	0.034
Binh et al. (2019)	ANN	P14..P27	-	-	-	0.875	0.032	0.041
Mirza et al.(2020)	SVR	P7,P12..P14	-	-	0.09987	-	-	1.0105
Divya et al. (2022)	ANN	P4..P7	-	95% ±5%	-	0.956	-	0.077
Baek et al.(2022)	ANN	P1	0.004	-	-	-	-	-
Baek et al.(2022)	CNN	P1	0.036	-	-	-	-	-
Adel et al.(2022)	SVR	P8,P29..P36	32.99	-	-	0.988	-	-
Adel et al.(2022)	ANN	P8,P29..P36	34.428	-	-	0.904	-	-
Champa,E. et al.(2024)	RNN	P5	0.435	-	-	0.6440	0.5444	-
Champa,E. et al.(2024)	RNN-LV	P5	0.3932	-	-	0.6784	0.5517	-
Our work	ANN	P1	0.0732	-	0.4210	0.0329	0.3469	0.1712
Our work	RNN	P5	0.4270	-	0.8103	0.6497	0.3820	0.1633
Our work	ANN	P4..P7	0.0252	90.63% ±4%	0.9579	0.8680	0.1724	0.1038
Our work	RNN	P4..P7	0.0587	94.64% ±2%	0.9850	0.9701	0.0597	0.0610
Our work	ANN	P1,P2,..P8	0.00036	-	-	0.1363	0.0061	0.0258
Our work	RNN	P1,P2,..P8	0.0198	97.62% ±1%	0.9955	0.9885	0.0229	0.0352
Our work	RNN M-C	P1,P2,..P8	0.0321	98.66% ±1%	0.997	0.9901	-	-

5. Conclusion

In conclusion, our meta-analysis represents a significant contribution to research on nanocomposites, shedding light on the advancements made in the methodologies and results related to the analysis of nanocomposites using nanoparticles over the past five years. Our primary objective was to identify the optimal algorithm using software information, leading to the development of a meta-analytic approach. We extensively examined the advantages and disadvantages of the most commonly used algorithms, proposing an enhancement of the performance of LSTM-based RNN neurons.

The study also included a comparison of Deep Learning methods for the classification of polymer nanoparticles, with polypropylene as a case study.
Our experiments focused on evaluating seven physical properties, utilizing neural networks such as artificial neural networks (ANN), recurrent neural networks (RNN), and RNN- Monte Carlo for simulations. The results underscore the remarkable efficiency of the last algorithm in achieving superior classification performance.

The findings obtained from the meta-analysis reinforce the justification for selecting an optimal algorithm for the analysis of polymer nanoparticles. This study highlights the significant potential of Deep Learning methodologies, particularly recurrent neural networks RNN- Monte Carlo method, to enhance the accuracy and efficiency of classification.
The evaluation of our methods(table 6) was conducted using several metrics, including Mean Absolute Percentage Error (MAPE), Mean Absolute Scaled Error (MASE), Root Mean Squared Percentage Error (RMSPE), Root Mean Square Error (RMSE), Root Mean Squared Scaled Error (RMSSE), correlation coefficient (R) and the determination coefficient (R^2). These measures provide a comprehensive assessment of the performance of the tested algorithms in this specific context.

In consideration of previous studies, our article aligns with the continuity of innovative work conducted by researchers such as Binh et al. , Mirza et al. Diviya et al., Baek et al., Adel et al., Champa, E et al.(see table 10). They have contributed to the understanding and optimization of nanomaterials through advanced machine learning techniques using different IA algorithm (SVM, SVR, ANN, CNN, RNN...) The task of comparison was elaborated in section 4 and we simulate with the same commun properties for example P1 or P5 even with four properties P4..P7. We use also the same IA algorithm. Our study, comparing results with Divya et al. (2022), highlights significant differences in accuracy, R^2, and RMSE.

Our results are so interesting especially after a rigourous meta-analysis task.In fact, the other reserachers used one or

four properties, and they never achieve the same accuracy with eight , as we have achieved. By distilling ideas from a broad set of research, we hope to contribute to the optimization and the application of Deep Learning techniques in diverse areas.
The future of nanocomposite analysis looks promising, with prospects for continuous improvement in accuracy and efficiency through the advances in Deep Learning that our study has highlighted.

A. My Appendix

A.1. RNN

The general equation for an RNN at a given time step is as follows:

$$h(t) = f(W_{hh}h(t-1) + W_{hx}x(t) + b_h) \tag{1}$$

Where:

- $h(t)$: Hidden state at time t.
- f : Activation function applied to the linear operation.
- W_{hh} : Weight matrix associated with the hidden state.
- $h(t-1)$: Hidden state at the previous time step.
- W_{hx} : Weight matrix associated with the input $x(t)$.
- $x(t)$:Input at time t.
- b_h : Bias term for the hidden state.

We used an Intel(R), Core(TM), i7-8550U CPU, @1.80GHz 1.99 GHz for the simulation.

A.2. The metrics

Mean Absolute Percentage Error :

This formula computes the average of the absolute percentage errors between actual and forecasted values, and it is expressed as a percentage Chicco, Warrens and Jurman (2021).

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - F_i}{A_i}\right| \times 100 \tag{2}$$

Where:

- $MAPE$ is the Mean Absolute Percentage Error.
- n is the number of observations.
- A_i is the actual value for observation i.
- F_i is the forecasted value for observation i.

A low MAPE indicates that the model's predictions are close to the actual values in terms of percentage error. The lower the MAPE, the better the model's performance in terms of forecast precision.

Mean Absolute Scaled Error :

The MASE compares the model's mean absolute errors with the mean absolute errors of a naive model (e.g., random walk) on the training set. A MASE less than 1 generally indicates better performance compared to a naive model. If the MASE is equal to 1, the model performs similarly to a naive model.

$$\text{MASE} = \frac{\frac{1}{h}\sum_{t=1}^{h}|Y_t - \hat{Y}_t|}{\frac{1}{n-1}\sum_{t=2}^{n}|Y_t - Y_{t-1}|} \tag{3}$$

Where:

- $MASE$ is the Mean Absolute Scaled Error.
- n is the number of observations in the time series.
- h is the number of forecast periods.
- Y_t is the actual value at time t.
- $\hat{Y}_t$ is the forecasted value at time t.

Root Mean Squared Percentage Error : The Root Mean Squared Percentage Error (RMSPE) is a metric used to assess the accuracy of predictions in forecasting and time series analysis.

$$\text{RMSPE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{A_i - F_i}{A_i}\right)^2} \times 100 \tag{4}$$

Where :

- $RMSPE$ is the Root Mean Squared Percentage Error.
- n is the number of observations.
- A_i is the actual value for observation i.
- F_i is the forecasted value for observation i

Similar to other error metrics, a lower RMSPE indicates better model accuracy. However, it is important to note that RMSPE is sensitive to values close to zero, much like MAPE.

https://www.overleaf.com/project/658dda18892054c2aad5cc39

Root Mean Squared Scaled Error :

The Root Mean Squared Scaled Error (RMSSE) is an evaluation metric used to measure the accuracy of forecasts in the field of time series modeling Kaur, Singh, Kahlon and Bassi (2010).

$$\text{RMSSE} = \sqrt{\frac{1}{h}\sum_{t=n+1}^{n+h}\left(\frac{Y_t - \hat{Y}_t}{S_{t-n}}\right)^2} \tag{5}$$

Where :

- $RMSSE$ is the Root Mean Squared Scaled Error.

- n is the number of observations in the historical data.
- h is the number of forecast periods.
- Y_t is the actual value at time t.
- $\hat{Y}_t$ is the forecasted value at time t.
- S_{t-n} is the seasonal factor for time $t - n$.

The coefficient of correlation R

$$R = \frac{\sum (X_i - \bar{X})^2}{\sum (Y_i - \bar{Y})^2} \sum (X_i - \bar{X})(Y_i - \bar{Y}) \tag{6}$$

Where:

- X_i and Y_i are the individual observations of X and Y,
- $\bar{X}$ and $\bar{Y}$ are the means of X and Y, respectively.

The correlation coefficient varies from -1 to 1 Ratner (2009), where:

- 1 indicates a perfect positive correlation (when one variable increases, the other also increases proportionally).
- - 1 indicates a perfect negative correlation (when one variable increases, the other decreases proportionally).
- 0 indicates an absence of linear correlation.

Coefficient of Determination R^2

R^2 is calculated by squaring the Pearson correlation coefficient between the dependent variable (Y) and the variable predicted by the regression model ($\hat{Y}$) Hahn (1973) Chicco et al. (2021).

$$R^2 = \left(\frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \right) \tag{7}$$

Where:

- $\hat{Y}_i$ are the values predicted by the model,
- Y_i are the observed values of the dependent variable,
- $\bar{Y}$ is the mean of the observed values.

The value of R^2 ranges from 0 to 1. An R^2 of 1 indicates that the model perfectly explains the variance of the dependent variable, while an R^2 of 0 indicates that the model explains no variance. In practice, intermediate values of R^2 are more common.

The coefficient of Pearson correlation (CP)

The Pearson coefficient measures the covariance between X and Y (numerator) normalized by the product of the standard deviations of X and Y (denominator).

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \tag{8}$$

Where :

- r est le coefficient de correlation de Pearson.
- X_i et Y_i représentent les observations individuelles dans les séries de données X et Y, respectivement.
- $\bar{X}$ et $\bar{Y}$ représentent les moyennes des séries de données X et Y, respectivement.

This renders it independent of the scales of the variables. The coefficient can vary from -1 to 1. A coefficient of 1 indicates a perfect positive linear correlation, -1 indicates a perfect negative linear correlation, and 0 indicates no linear correlation between the variables Xu and Deng (2017).

A.3. Normalization

Esperance (or mean) of a random variable is a measure of the central tendency of that variable. It represents the average value one can expect from this variable over a large number of observations or trials Boyé, Comairas and de la Loire (2002) Bru and Heinich (1980).It is denoted by μ and is calculated by summing all the values in a dataset and dividing by the total number of observations (N).

$$\mu = \frac{\sum_{i=1}^{N} L_i}{N} \tag{9}$$

where:

- μ represents the mean.
- L_i represents each value in the dataset.
- N is the total number of observations.

The standard deviation σ is a measure of the dispersion of observations relative to the mean. It reflects how values are distributed around the average value within a dataset Vermette (2016) Boyé et al. (2002).

The variance, denoted as V, is calculated by taking the average of the squared differences between each value and the mean ZENDRERA and MARION.

$$V = \frac{\sum_{i=1}^{N} (L_i - \mu)^2}{N} \tag{10}$$

where:

- V represents the variance.
- μ represents the mean.
- L_i represents each value in the dataset.
- N is the total number of observations.

$$\sigma = \sqrt{V} \tag{11}$$

These professional equations allow for the precise and formal calculation of the mean, variance, and standard deviation, which is essential in statistics for quantifying the central tendency and dispersion of data.

A.4. The Other used properties

we begin from P9 to P36.

P9: Distance and distribution (DD) : The distance refers to the spatial separation or spacing between nanoparticles within the matrix of the nanocomposite.The distribution pertains to how nanoparticles are arranged or spread throughout the matrix of the nanocomposite Luo et al. (2019).

P10: Coefficient of thermal expansion (C): The coefficient of thermal expansion, also known as the thermal expansion coefficient, is a physical quantity that measures the change in dimensions of a material in response to a temperature change. Specifically, the thermal expansion coefficient expresses the relative variation in length, area, or volume of a material per degree of temperature change Yoon, Fornes and Paul (2002).

P11: Volume fraction (V): The volume fraction of polymer nanoparticles is often expressed as a percentage for clarity, where *ro* represents the percentage proportion of the volume of polymer nanoparticles relative to the total volume of the material. It is important for characterizing the distribution and concentration of polymer nanoparticles in a polymer matrix, which can have significant implications for the properties and performance of materials in various applications.

P12: Compression coefficient Cc (CC) :The Compression Coefficient (Cc) is a physico-mechanical parameter of soils that reflects their compressibility during the consolidation process Pham et al. (2019).

P13: Density (D) : The density of a substance is a physical characteristic that indicates how much matter the substance contains in a given space. The Greek symbol rho ρ is commonly used to represent density, and it is measured in units such as kilograms per cubic meter kg/m^3 in the International System. In other words, density provides insight into the concentration of mass within a specific volume, which is valuable for understanding the behavior of materials in various fields such as physics and chemistry.

P14: Mass Fraction which is a Mass fraction refers to the ratio of the mass of a specific component to the total mass of the mixture or material.

P15: Depth of sample represents the vertical distance from the surface to a specified point within a material or substance.

P16 : Bulk density is the mass of a material per unit volume, including both solid and void spaces. It is often used to characterize soil or powders Blake (1965).

P17: Plasticity index is a measure of the plasticity of a soil, indicating the range of moisture content within which the soil exhibits plastic behavior.

P18: Moisture content is the amount of water present in a material, usually expressed as a percentage of the total mass Gardner (1986).

P19: Clay content refers to the proportion of soil particles that are classified as clay-sized, typically with diameters less than 0.002 mm Soinne, Keskinen, Tähtikarhu, Kuva and Hyväluoma (2023).

P20: Specific gravity is the ratio of the density of a substance to the density of a reference substance (usually water for liquids and air for gases) Figura and Teixeira (2023).

P21: Void ratio is the volume of voids (spaces) to the volume of solids in a soil or other granular material Sarkar, König and Goudarzy (2019).

P22: Liquid limit is the moisture content at which a soil changes from a plastic to a liquid state under specified conditions.

P23: Dry density is the density of a material when all the moisture is removed, typically expressed in mass per unit volume.

P24: Porosity is the ratio of the volume of voids to the total volume of a material, often expressed as a percentage.

P25: Plastic limit is the moisture content at which a soil ceases to behave plastically and begins to crumble when rolled into a thread.

P26: Degree of saturation is the ratio of the volume of water present in a material to the volume of voids, expressed as a percentage.

P27: Liquidity index is a measure used in soil mechanics to quantify the consistency of a soil, indicating whether it is in a liquid, plastic, or solid state.

P28: Flux (F) is a measure of the flow of a property through a surface or across a boundary. In various contexts, it can represent the flow of particles, energy, or other quantities. In the article by Masoud Fetanat et al. Fetanat et al. (2021), the term "flux" refers to the transfer rate of a substance, specifically the permeate flux in nanocomposite filtration membranes. Permeate flux denotes the quantity of fluid passing through the membrane per unit area and time, serving as a crucial parameter for assessing the efficiency and performance of filtration membranes.

P29: Water to cement ratio (W/C) The ratio of the mass of water to the mass of cement in the mortar mixture Singh,

Munjal and Thammishetti (2015).

P30: Type : Nanomaterials exhibit distinct properties or characteristics at the nanometer scale (typically, 1 to 100 nanometers). They can be categorized based on their composition, structure, or attributes, including nanoparticles, nanotubes, nanowires, and nanocomposites.

P31 : Content : The composition of a nanomaterial refers to the elemental, compound, or structural makeup of the material at the nanoscale. This encompasses the specific atoms, molecules, or configurations present within the material.

P32 : Length : The extent of a nanomaterial denotes the measurement of its longest dimension in a particular direction. This measurement varies depending on the nanomaterial type; for example, in nanotubes, it signifies the length of the tubular structure.

P33 : Diameter : The dimensions of a nanomaterial include the measurement of its width or thickness. This can differ based on the specific type of nanomaterial; for instance, in nanoparticles, it pertains to the size of the particle across its broadest dimension.

P34: Sand to cement ratio for mortars (S/C) is the ratio of the mass of sand to the mass of cement in the mortar mixture Plawsky, Jovanovic, Littman, Hover, Gerolimatos and Douglas (2003).

P35 : Surfactant type : the type of surfactant used in the mortar mixture.

P36: Curing days (age), the duration for which the mortar samples were cured before testing, often expressed in days Çakır and Aköz (2008).

References

Adel, H., Palizban, S.M.M., Sharifi, S.S., Ilchi Ghazaan, M., Habibnejad Korayem, A., 2022. Predicting mechanical properties of carbon nanotube-reinforced cementitious nanocomposites using interpretable ensemble learning models. Construction and Building Materials 354, 129209. URL: https://www.sciencedirect.com/science/article/pii/S0950061822028653, doi:https://doi.org/10.1016/j.conbuildmat.2022.129209.

Azouani, R., 2009. Elaboration de nouveaux nanomatériaux photocatalytiques actifs sous rayonnement visible. Ph.D. thesis. Université Paris-Nord-Paris XIII.

Baek, K., Hwang, T., Lee, W., Chung, H., Cho, M., 2022. Deep learning aided evaluation for electromechanical properties of complexly structured polymer nanocomposites. Composites Science and Technology 228, 109661.

Baek, K., Kim, H., Shin, H., Park, H., Cho, M., 2023. Multiscale study to investigate nanoparticle agglomeration effect on electrical conductivity of nano-sic reinforced polypropylene matrix composites. Mechanics of Advanced Materials and Structures 30, 2442–2452.

Banga, S., Gehani, H., Bhilare, S., Patel, S., Kara, L., 2018. 3d topology optimization using convolutional neural networks. arXiv preprint arXiv:1808.07440 .

Barai, D.P., Bhanvase, B.A., Pandharipande, S.L., 2022. Artificial neural network for prediction of thermal conductivity of rgo–metal oxide nanocomposite-based nanofluids. Neural Computing and Applications 34, 271–282.

Bégin-Colin, S., Le Caër, G., Girot, T., 2002. Les nanomatériaux: une caractérisation difficile..., in: Journal de Physique IV (Proceedings), EDP sciences. pp. 441–454.

Blake, G., 1965. Bulk density. Methods of soil analysis: Part 1 physical and mineralogical properties, including statistics of measurement and sampling 9, 374–390.

Boyé, A., Comairas, M.C., de la Loire, I.d.P., 2002. Moyenne, mediane, écart-type. Quelques regards sur l'histoire .

Bru, B., Heinich, H., 1980. Sur l'espérance des variables aléatoires vectorielles, in: Annales de l'IHP Probabilités et statistiques, pp. 177–196.

Burch, R., Najm, F.N., Yang, P., Trick, T.N., 1993. A monte carlo approach for power estimation. IEEE Transactions on Very Large Scale Integration (VLSI) Systems 1, 63–71.

Çakır, Ö., Aköz, F., 2008. Effect of curing conditions on the mortars with and without ggbfs. Construction and Building Materials 22, 308–314.

Champa-Bujaico, E., Díez-Pascual, A.M., Redondo, A.L., Garcia-Diaz, P., 2024. Optimization of mechanical properties of multiscale hybrid polymer nanocomposites: A combination of experimental and machine learning techniques. Composites Part B: Engineering 269, 111099.

Champa-Bujaico, E., García-Díaz, P., Díez-Pascual, A.M., 2022. Machine learning for property prediction and optimization of polymeric nanocomposites: a state-of-the-art. International Journal of Molecular Sciences 23, 10712.

Chen, F., Weng, L., Wang, J., Wu, P., Ma, D., Pan, F., Ding, P., 2023. An adaptive framework to accelerate optimization of high flame retardant composites using machine learning. Composites Science and Technology 231, 109818.

Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. PeerJ Computer Science 7, e623.

Chouard, C.H., Cabanis, E.A., Chambron, J., Milgrom, E., Adolphe, M., Ardaillou, R., Aurengo, A., Bach, J.F., Battin, J., Baulieu, É.É., et al., 2008. Nanosciences et médecine. Bulletin de l'Académie Nationale de Médecine 192, 1253–1259.

Deng, L., Platt, J., 2014. Ensemble deep learning for speech recognition, in: Proc. interspeech.

Faust, O., Hagiwara, Y., Hong, T.J., Lih, O.S., Acharya, U.R., 2018. Deep learning for healthcare applications based on physiological signals: A review. Computer methods and programs in biomedicine 161, 1–13.

Fetanat, M., Keshtiara, M., Keyikoglu, R., Khataee, A., Daiyan, R., Razmjou, A., 2021. Machine learning for design of thin-film nanocomposite membranes. Separation and Purification Technology 270, 118383.

Figura, L.O., Teixeira, A.A., 2023. Mass density, in: Food Physics: Physical Properties-Measurement and Applications. Springer, pp. 59–100.

Gaffet, E., 2011. Nanomatériaux: Une revue des définitions, des applications et des effets sur la santé. comment implémenter un développement sûr. Comptes Rendus Physique 12, 648–658.

Gardner, W.H., 1986. Water content. Methods of Soil Analysis: Part 1 Physical and Mineralogical Methods 5, 493–544.

Gedda, M., 2015. Traduction française des lignes directrices prisma pour l'écriture et la lecture des revues systématiques et des méta-analyses. Kinésithérapie, la Revue 15, 39–44.

Goyal, P., Pandey, S., Jain, K., 2018. Deep learning for natural language processing. New York: Apress .

Hahn, G.J., 1973. The coefficient of determination exposed. Chemtech 3, 609–612.

Haiss, W., Thanh, N.T., Aveyard, J., Fernig, D.G., 2007. Determination of size and concentration of gold nanoparticles from uv- vis spectra. Analytical chemistry 79, 4215–4221.

Ho, N.X., Le, T.T., Le, M.V., 2022. Development of artificial intelligence based model for the prediction of young's modulus of polymer/carbon-nanotubes composites. Mechanics of Advanced Materials and Structures 29, 5965–5978.

Ioannidou, A., Chatzilari, E., Nikolopoulos, S., Kompatsiaris, I., 2017. Deep learning advances in computer vision with 3d data: A survey. ACM computing surveys (CSUR) 50, 1–38.

Kantas, N., Doucet, A., Singh, S.S., Maciejowski, J.M., 2009. An overview of sequential monte carlo methods for parameter estimation in general state-space models. IFAC Proceedings Volumes 42, 774–785.

Kaur, J., Singh, S., Kahlon, K.S., Bassi, P., 2010. Neural network-a novel technique for software effort estimation. International Journal of Computer Theory and Engineering 2, 17.

Lartigue, L., 2010. Synthesis, Characterization, and Theranostic Application of Iron Based Magnetic Nanoparticles. Ph.D. thesis. Montpellier 2.

Le, T.T., 2021. Prediction of tensile strength of polymer carbon nanotube composites using practical machine learning method. Journal of Composite Materials 55, 787–811.

Le, T.T., Le, M.V., 2021. Prediction model for tensile modulus of carbon nanotube–polymer composites, in: International Conference on Engineering Research and Applications, Springer. pp. 786–792.

Le Bolzer, F., Lambert, A., Schnitzler, F., 2020. Vers des réseaux de neurones récurrents flexibles .

Legrand, D., Gaines Jr, G., 1969. The molecular weight dependence of polymer surface tension. Journal of Colloid and Interface Science 31, 162–167.

Liu, B., Vu-Bac, N., Zhuang, X., Fu, X., Rabczuk, T., 2022. Stochastic integrated machine learning based multiscale approach for the prediction of the thermal conductivity in carbon nanotube reinforced polymeric composites. Composites Science and Technology 224, 109425. URL: https://www.sciencedirect.com/science/article/pii/S0266353822001671, doi:https://doi.org/10.1016/j.compscitech.2022.109425.

Louis, S.Y., Zhao, Y., Nasiri, A., Wang, X., Song, Y., Liu, F., Hu, J., 2020. Graph convolutional neural networks with global attention for improved materials property prediction. Physical Chemistry Chemical Physics 22, 18141–18148.

Luo, B., Wang, X., Tian, M., Cai, Z., Li, L., 2019. Homogeneity quantification of nanoparticles dispersion in composite materials. Polymer Composites 40, 1000–1005.

Matos, M., Pinho, S., Tagarielli, V., 2019a. Predictions of the electrical conductivity of composites of polymers and carbon nanotubes by an artificial neural network. Scripta Materialia 166, 117–121. URL: https://www.sciencedirect.com/science/article/pii/S1359646219301368, doi:https://doi.org/10.1016/j.scriptamat.2019.03.003.

Matos, M.A., Pinho, S.T., Tagarielli, V.L., 2019b. Application of machine learning to predict the multiaxial strain-sensing response of cnt-polymer composites. Carbon 146, 265–275.

Miotto, R., Wang, F., Wang, S., Jiang, X., Dudley, J.T., 2018. Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics 19, 1236–1246.

Otter, D.W., Medina, J.R., Kalita, J.K., 2020. A survey of the usages of deep learning for natural language processing. IEEE transactions on neural networks and learning systems 32, 604–624.

Özkan, M., Karakoç, A., Borghei, M., Wiklund, J., Rojas, O.J., Paltakari, J., 2019. Machine learning assisted design of tailor-made nanocellulose films: A combination of experimental and computational studies. Polymer Composites 40, 4013–4022.

Patil, M., Mehta, D.S., Guvva, S., 2008. Future impact of nanotechnology on medicine and dentistry. Journal of Indian society of periodontology 12, 34.

Pham, B.T., Nguyen, M.D., Van Dao, D., Prakash, I., Ly, H.B., Le, T.T., Ho, L.S., Nguyen, K.T., Ngo, T.Q., Hoang, V., et al., 2019. Development of artificial intelligence models for the prediction of compression coefficient of soil: An application of monte carlo sensitivity analysis. Science of The Total Environment 679, 172–184.

Plawsky, J., Jovanovic, S., Littman, H., Hover, K., Gerolimatos, S., Douglas, K., 2003. Exploring the effect of dry premixing of sand and cement on the mechanical properties of mortar. Cement and concrete research 33, 255–264.

Rahman, A., Deshpande, P., Radue, M.S., Odegard, G.M., Gowtham, S., Ghosh, S., Spear, A.D., 2021. A machine learning framework for predicting the shear strength of carbon nanotube-polymer interfaces based on molecular dynamics simulation data. Composites Science and Technology 207, 108627.

Ratner, B., 2009. The correlation coefficient: Its values range between+ 1/- 1, or do they? Journal of targeting, measurement and analysis for marketing 17, 139–142.

Rittigstein, P., Torkelson, J.M., 2006. Polymer–nanoparticle interfacial interactions in polymer nanocomposites: confinement effects on glass transition temperature and suppression of physical aging. Journal of Polymer Science Part B: Polymer Physics 44, 2935–2943.

Roualdès, S., 2009. Les nanomatériaux pour l'environnement et l'énergie. l'actualité chimique , 19.

Sahaluddin, M., Alade, I.O., Oyedeji, M.O., Aliyu, U.S., 2020. A machine learning-based model to estimate the density of nanofluids of nitrides in ethylene glycol. Journal of Applied Physics 127.

Sanchez, F., Sobolev, K., 2010. Nanotechnology in concrete–a review. Construction and building materials 24, 2060–2071.

Sarkar, D., König, D., Goudarzy, M., 2019. The influence of particle characteristics on the index void ratios in granular materials. Particuology 46, 1–13.

Shabtai, I.A., Lynch, L.M., Mishael, Y.G., 2021. Designing clay-polymer nanocomposite sorbents for water treatment: A review and meta-analysis of the past decade. Water Research 188, 116571.

Shady, E., Gowayed, Y., 2010. Effect of nanotube geometry on the elastic properties of nanocomposites. Composites Science and Technology 70, 1476–1481.

Shi, M., Feng, C.P., Li, J., Guo, S.Y., 2022. Machine learning to optimize nanocomposite materials for electromagnetic interference shielding. Composites Science and Technology 223, 109414.

Singh, D., Rehman, N., Pandey, A., 2023. Nanotechnology: the alternative and efficient solution to biofouling in the aquaculture industry. Applied Biochemistry and Biotechnology , 1–16.

Singh, S., Munjal, P., Thammishetti, N., 2015. Role of water/cement ratio on strength development of cement mortar. Journal of Building Engineering 4, 94–100.

Soinne, H., Keskinen, R., Tähtikarhu, M., Kuva, J., Hyväluoma, J., 2023. Effects of organic carbon and clay contents on structure-related properties of arable soils with high clay content. European Journal of Soil Science 74, e13424.

Tellier, M.D.I., 2016. Étude de réseaux de neurones récurrents pour l'étiquetage de séquences. PARIS Inalco du 4 au 8 juillet 2016 , 98.

Vermette, S., 2016. L'écart-type: au-delà de l'algorithme. Bulletin AMQ 56.

Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E., et al., 2018. Deep learning for computer vision: A brief review. Computational intelligence and neuroscience 2018.

Xu, H., Deng, Y., 2017. Dependent evidence combination based on shearman coefficient and pearson coefficient. Ieee Access 6, 11634–11640.

Yoon, P., Fornes, T., Paul, D., 2002. Thermal expansion behavior of nylon 6 nanocomposites. Polymer 43, 6727–6741.

ZENDRERA, N., MARION, J.M., . Evolution des conceptions au regard de l'ecart-type chez les etudiants en shs .

Zhang, Z., Geiger, J., Pohjalainen, J., Mousa, A.E.D., Jin, W., Schuller, B., 2018. Deep learning for environmentally robust speech recognition: An overview of recent developments. ACM Transactions on Intelligent Systems and Technology (TIST) 9, 1–28.

Rania Loukil, Assistant professor Automation and Industrial Computing from 2015 until today. She holds a Doctorate in Automation and Industrial Computing in 2014 about the diagnostic of complex systems based on nonlinear observers. Her a master's degree in electrical engineering in 2009, was Conducted in collaboration with INRIA (Paris) on the topic of visual odometry for stereoscopic cameras mounted on electric vehicles. She has the Engineering Diploma in Electronics from the National School of Engineers from Sfax, Tunisia in 2008. She was the Director of the department of Electrical Engineering at Private Central University from 2014 to 2015. She was a coach senor at BIAT foundation. Her research focuses on Diagnosis and Control, nonlinear systems and observers, neural networks, sliding mode, Deep Learning, IoT and Smart grid.

Wejdene Gazehi,in June 2018 received her Bachelor's degree and her License Degree in Industrial Computer Science in 2021 from the Institut supérieur des technologies de l'information et de la communication (ISTIC) at the University of Carthage, Tunisia. Her academic journey continued with a Master's degree in Data Science and Intelligent Services (D3S) in November 2023, also at ISTIC. Currently, she is a PhD student at National Engineering School of Tunis. Her research interests encompass nanotechnology and artificial intelligence, with a particular focus on nanoparticle classification, deep learning, and recurrent neural networks.

Mongi Besbes, professor of Automation and Industrial Computing, holds a Doctorate in Automation and Industrial Computing, a master's degree in electrical engineering, and an Engineering Diploma in Electronics from the National School of Engineers from Tunis. He held several academic positions at the Tunisian University as Director of the Department of Electrical Engineering at the National School of Engineers of Carthage (from 2004 to 2011) and Director of the Department of e-Learning at the University of Carthage (from 2011 to 2014). Director of the Higher Institute of Information and Communication Technologies (2014-2020). From 2020, he is Vice Rector at the University of Carthage. His research focuses on Robust Control, Robotics, LPV nonlinear systems and observers, fuzzy logic, multi-level inverters, and renewable energies.