UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Câmpus de Ilha Solteira

Faculdade de Engenharia

Neural Networks: Enhancing Intelligent Systems with Deep Learning

Gabriel D. Silva
Student
Douglas D. Bueno
Advisor

UNESP
Ilha Solteira – SP
2023

Research Report

The present report approaches a way to improve smart systems. Through artificial intelligence applied in the mechanical engineering field, it provides a consistent algorithm that can reads data, trains the machine and provides results about the situation and what to do with it. It will be studied two cases, one of them using machine learning classical techniques to determine the forces applied to a unnamed aerial vehicle and other using deep learning techniques like neural networks in the structural health monitoring area.

Complete after the research is done.

Keywords: machine learning, structural health monitoring, unnamed aerial vehicle

LIST OF FIGURES

LIST OF ACRONYM

AI Artificial Intelligence
ANN Artificial Neural Network

BD Big Data

CNN Convolutional Neural Network

DL Deep Learning

FEM Finite Element Method

IoT Internet of Things

MEMS Micro Electromechanical Systems
ML Machine Learning
MLP Multilayer Perceptron
MSE Mean Squared Error

NN Neural Network

RNN Recurrent Neural Network

SGD Stochastic Gradient Descent
SHM Structural Health Monitoring

UAV Unnamed Aerial Vehicle

CONTENTS

1 INTRODUCTION

The use of Artificial Intelligence (AI) is very present nowadays [38, 51, 52]. This area of statistics neither is new nor started just now with the autonomous cars and voice assistants [47], but it is clear that in the last years it has been increasingly gaining more popularity. This happens mainly because of the advances that the World Wide Web has been had over the years [39, 14], since dial-up internet connection, back in the eighties, until now, with broadband internet and smartphones equipped with 5G connection. Another factor is that in the past, the cost to get a large capacity of storage memory was significantly more expensive than it is now, what makes today cheaper and easy to get memory to store information [27]. With the amount of data available, the evolution of internet and storage capacity, now it is not difficult to obtain, keep and analyze databases to make decisions [20].

AI application is everywhere and today, more than ever, it is easy to realize that. Either to get multimedia recommendations on streaming platforms, like occurs at Netflix, YouTube, Spotify, and so many others platforms [13], or to make predictions on the financial market and sports betting [46, 35, 30], AI is there behind the scenes making all the magic happen. Evidently there is nothing really magical about them, it is pure mathematics combined with a programming language that produces the algorithm capable of doing those things [28, 3, 54, 55]. The launch of ChatGPT–3, and shortly thereafter ChatGPT–4, has shown the power of those technologies and how they can change the way people do things [10, 9, 41, 6].

Getting into the smart systems application, the use of AI is widely used to Structural Health Monitoring (SHM), which is heavily used in the aerospace and civil fields, [5, 72]. The level and the complexity of the AI to be applied to monitor the structure, whether is going to use Deep Learning (DL) and Neural Network (NN) or simpler methods of Machine Learning (ML) like regressions, is determined by the problem itself and the results desired [22]. In some cases, the standards methods use numerical techniques and they may not be feasible, especially when there is a huge data to be analyzed. Thus, taking the AI road is an alternative to get the needed results for the monitoring in a more practical way [62, 65].

Still in this context, but in the field of Unnamed Aerial Vehicle (UAV), the use of AI can be combined to integrate UAV through wireless communication networks [37] what can be useful in the agriculture sphere [1] with technologies like Internet of Things (IoT) [69, 68]. Also, the use of the AI can be subtle, such as the use of a built-in MATLAB function to make a simple NN to determine the final pose of a UAV based on the initial pose and the forces applied on it [25], or can be more sophisticated, like the use of ML and DL algorithms to predict materials properties, design new materials, discover new mechanisms and control real dynamic systems [29, 2].

It is clear, therefore, that AI can transit into different fields, such as entertainment, business, health care, marketing, financial, agriculture, engineering, among others [56, 73, 18, 70, 45, 50, 26]. The use of the Big Data (BD) can not only make it clear the scenario to be studied, but also to support making strategical decisions [31, 36]. The internet and hardware improvement [7], alongside the facility to storage data with accessible costs, encourages the AI use due to the benefits it can provide.

1.1 Motivation

1.2 Objective

To develop an AI algorithm based on NN to apply in smart systems. The main goals are:

- to determine the forces used to move an UAV based on its initial and final pose;
- to detect railways cracks through piezoelectric signal for SHM.

2 LITERATURE REVIEW

This chapter deals with the history, the main concepts and some practical cases of SHM inside the industry and academic area, besides showing how it may be used in the railway crack detection context. Next, in the dynamic field, it will be studied the main mechanical concepts to get the necessary understanding to an UAV motion as well the basics to know how an UAV can be controlled. Then, it will be shown the mathematics behind the algorithms of deep learning that will be implemented in the Chapter 4. Finally, the way how the algorithms are going to be implemented and the tools necessary to achieve the desired neural network.

2.1 Structural Health Monitoring

2.1.1 Definition

According to Balageas et al. [8], the SHM main purpose is to provide, during the life of a structure, a diagnosis of: the state of the constituent material; the different parts of the structure; and the full assembly of each part that makes the structure as a whole. It is an improved way to make non-destructive evaluation. It can be applied in several areas such as civil infrastructure, like bridges and buildings; aerospace, like airplanes and spaceships; and mechanical, like machines.

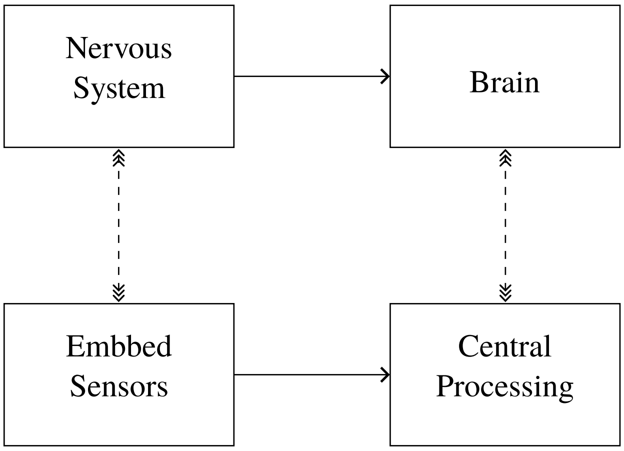


Figure 2.1.1. SHM and Human Nervous System Analogy. The nerve endings are responsible to pass the information to the brain; as the sensors embedded in the structure play the role of the nerve endings, the central processing does the brain role.

Furthermore, it also can be associated as an analogy to the human nervous system. Just like the sensors send a signal to the central processor, the human senses send a signal to the brain to make the recognition of what is happening, as shown in the Fig. 2.1.1.

2.1.2 Brief History

The SHM development began back in the 20th century and it has been coupled with the evolution of the digital computing hardware, what allowed the costs of the applied techniques less expensive over time.

It all starts back in the early 1970s and 1980s. The oil industry tried to develop vibration-based damage identification methods for offshore platforms by simulating damage scenarios, examining the changes in the resonant frequencies and correlating them with those measured on a platform. In the same period, the aerospace community studied vibration-based damage identification along with the development of the space shuttle. From that, it was developed the shuttle modal inspection system which aimed to identify fatigue damage in components like fuselage panels and control surfaces. The system was so successful that all orbiter vehicles had been periodically subjected to this test. Also, the civil engineering community studied vibration-based damage evaluation of bridge structures and buildings in the late 1980s [21].

From the late 1990s to the early 2000s, Sohn et al. [63] showed the evolution of the techniques used in SHM, analyzing mainly the following factors: the operational evaluation; data acquisition and cleansing; feature extraction; and statistical modeling for feature discrimination. He also verified that the statistical patter recognition had not been embraced by the researchers to be more often used in such matter.

Nowadays, in order to contour inherent issues of SHM methods, as large computational effort and hand-crafted work that results in poor classification performance, many deep learning techniques have been used, such as Convolutional Neural Network (CNN) [4].

2.1.3 Main Techniques

Accelerometers

The use of accelerometers is consolidated in the engineering community to be used in several areas and it is present in the people daily life in things like game consoles, smartphones, and tablets.

Micro Electromechanical Systems (MEMS) sensor have several applications in measuring linear acceleration or angular motion along axis as an input to control a system. MEMS accelerometer sensors often measure the movement of a mass with a position measuring interface circuit that is converted into a digital electrical signal by an analog-to-digital converter for digital processing [16].

In SHM situation, the accelerometers are in the MEMS. The MEMS are, then, embedded in the structure and can provide information about the structure by detecting low-amplitude and low-frequency vibrations that are not always viable with the conventional low-cost sensor boards [57].

There are many others sensors used in vibration-base techniques like velocity and displacement sensors, however the accelerometers are widely used for this purpose.

Graphical Inspection

The use of digital cameras to detect any kind of irregularity in the surface is also a way to monitor the structure, mainly in the surface areas. The camera itself may be static in a strategical position that allow it to provide good images to be analyzed or can be embedded in the structure itself or in an UAV that will surround it.

To automate and improve the accuracy of the damage detection, image processing techniques are employed, that being a non-conventional approach [60]. In the civil engineering context, it is commonly utilized computer vision to damage detection [23] and also UAV integrated in the same local as the structures for SHM [59].

Many of the images obtained can have their not only the images improved by AI, but also the analyses can take advantage of it.

Piezoelectric Materials

When dealing with acoustic-based techniques, the use of piezoelectric materials as sensor is a great choice due to its ability to respond to stimuli, incorporation, and compatibility with construction materials. Beyond that, these materials are relatively cost-efficient and can sense vibrations in the structures they are installed [32].

Piezoelectric materials and their main property were discovered back in 1880 [15]. The phenomenon is that by the application of pressure in those kinds of materials in the correct direction, it is observed the production of a potential difference and consequently an electrical charge. Examples of materials that are piezoelectric are quartz, zinc, sodium chlorate, tourmaline, calamine, topaz, tartaric acid, cane sugar, and others [11].

The application of these materials in SHM is basically to install the piezoelectric sensor in the structure intend to be monitored and through the tension or compression in it done, a sign will be sent to the central system by the potential differential. The signal indicates that something not usual is happening in the structure. Of course there are levels of the signals and each case must be evaluated in its context. In the last years the use of piezoelectric materials has been capable to identify failures, like the presence of delamination damage, as long as the piezoelectric sensors are close to the damage [42].

2.1.4 Railway Cracks

Train is one of the most used means of transportation around the world, either to transport people or groceries, therefore, there are inherent problems in the attached to it. One of the most common problems is the crack on the railway track, mainly due to the expansion and contraction caused by the heat and to constant pressure because of the wagon.

The crack in a railway is considerable problem because it may cause fatal accidents since the wagon is able to leave the railway. In this scenario, many methods are used to detect the crack or to foresee it before any misfortune happen. Karthick and Ramalingam [33] proposed a system to identify the cracks and prevent the accidents. One of its advantages is that if some crack is detected on the track, the train starts to slow and stop before it passes by there. Other method includes the use of sensor coupled in the track that allow to detect the crack and send a signal to the command center through IoT [58].

The use of piezoelectric materials for SHM is very common, as seen in the Section 2.1.3. Loveday [40] presents a system where piezoelectric transducer are installed along the railway track. They receive an electrical wave and send it, then, a signal to the receiver, making it possible, also through Finite Element Method (FEM), to detect any inconsistency that should not be there.

There are, hence, lots of methods that can be used to detect and prevent accidents in railway tracks. Putting they together and optimize them with SHM techniques are an efficient way of improve the railway ecosystem.

2.2 Unnamed Aerial Vehicle

2.2.1 Usage

An UAV has several applications, going from the simplest to the most sophisticated. It can be used since for entertainment, like toys; commercially, to record big shows in arenas; surveillance, to monitor places; and also in engineering, aiding in various context to improve some processing.

Due to its portability and autonomy, it can be used to facilitate the delivery o medicines. In this sense, UAV can be used for transportation of medical goods in critical times, where other means of transportation may not be feasible. In the final of 2019, COVID-19 pandemics spread throughout the world, making it difficult to deliver patients their needed medicines [53, 44]. Besides, risks are inherent to the transportation and in come countries, like the USA, UAV usage may be restricted [67]. A strategical way to use them is also welcome.

In the agriculture context, in order to boost the productivity, UAV can be used to remotely sense the farming, obtaining information on the state of the fields with non-contact procedures, like nutrient evaluation and soil monitoring; or even for aerial spraying, using pesticide to prevent damages in the plantation [19].

The main reason for its adoptions is the mobility, low maintenance costs, hovering capacity, ease of deployment, etc. It is widely used for the civil infrastructure, gathering photographs faster than satellite imagery and with better quality. Combining those benefits with AI can be a powerful tool for the future. Sivakumar and Tyj [61].

2.2.2 Motion Equation

Considering the UAV a quadcopter, as the Fig. 2.2.1 shows, it is assumed a defined coordinate system fixed in the body (Z_M). Fossen [24] described the equation of motion as being

$$\mathbf{M}_\eta(\boldsymbol{\eta})\ddot{\boldsymbol{\eta}} + \mathbf{C}_\eta(\boldsymbol{\nu}, \boldsymbol{\eta})\dot{\boldsymbol{\eta}} + \mathbf{g}_\eta(\boldsymbol{\eta}) = \boldsymbol{\tau}_\eta(\boldsymbol{\eta}) + \mathbf{F}_d \tag{2.2.1}$$

where $\mathbf{M}_\eta(\boldsymbol{\eta})$ is the inertial matrix; $\mathbf{C}_\eta(\mathbf{v}, \boldsymbol{\eta})$ is the Coriolis matrix; $\mathbf{g}_\eta(\boldsymbol{\eta})$ is the gravitational vector; $\boldsymbol{\tau}_\eta(\boldsymbol{\eta})$ is the control torque; $\mathbf{F}_d$ is the gust vector; and $\boldsymbol{\nu}$ is the velocity generalized coordinate in the body-frame. The generalized coordinate's vector of the quadrotor in the inertial

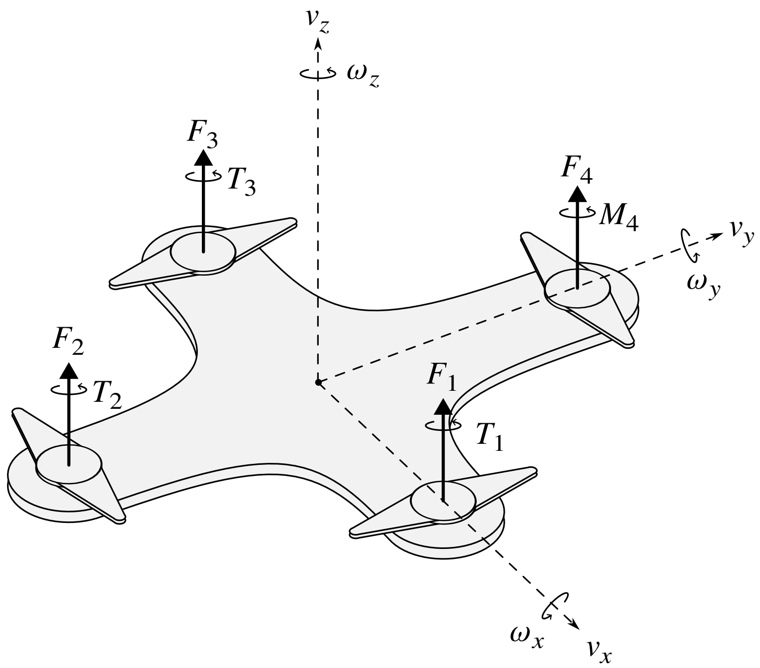


Figure 2.2.1. Quadcopter Dynamic Scheme. F_i and T_i, for $i = 1, 2, 3, 4$, are the forces and the torque applied in the propeller, respectively. ω_j and v_j, for $j = x, y, z$, are the momentum and the velocities applied in the UAV, respectively.

reference frame is

$$\boldsymbol{\eta} = \begin{Bmatrix} x & y & z & \theta & \phi & \psi \end{Bmatrix}^\top \tag{2.2.2}$$

where $x,\ y,\ z$ and $\phi,\ \theta,\ \psi$ are the positions and orientation angles of the quadrotor. As detailed by Geronel et al. [25], the matrices of the Eq. (2.2.1) are

$$\mathbf{M}_\eta(\boldsymbol{\eta}) = \mathbf{I}_{6\times 6} \tag{2.2.3}$$

$$\mathbf{C}_\eta(\boldsymbol{\nu}, \boldsymbol{\eta}) = \mathbf{J}\mathbf{M}^{-1}\mathbf{C}\mathbf{J}^{-1} - \dot{\mathbf{J}}\mathbf{J}^{-1} \tag{2.2.4}$$

$$\mathbf{g}_\eta(\boldsymbol{\eta}) = \mathbf{J}\mathbf{M}^{-1}\mathbf{g}_0 \tag{2.2.5}$$

$$\boldsymbol{\tau}_\eta(\boldsymbol{\eta}) = \mathbf{J}\mathbf{M}^{-1}\boldsymbol{\tau} \tag{2.2.6}$$

where $\mathbf{J}$ is the derivative of the transformation matrix with respect to time. The main matrices of the Eq. (2.2.6) are

$$\mathbf{M} = \mathrm{diag}(m, m, m, I_{xx}, I_{yy}, I_{zz}) \tag{2.2.7}$$

$$\mathbf{C} = \begin{bmatrix} 0 & -m\omega_z & m\omega_y & 0 & 0 & 0 \\ m\omega_z & 0 & -m\omega_x 0 & 0 & 0 & 0 \\ -m\omega_y & m\omega_x & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_1\omega_y \\ 0 & 0 & 0 & I_2\omega_z & 0 & 0 \\ 0 & 0 & 0 & 0 & I_3\omega_x & 0 \end{bmatrix} \tag{2.2.8}$$

where $I_1 = I_{zz} - I_{yy}$ is the moment of inertia along the x axis; $I_2 = I_{xx} - I_{zz}$ is the moment of inertia along the y axis;$I_3 = I_{yy} - I_{xx}$ is the moment of inertia along the z axis; $\mathbf{g}_0 = \begin{bmatrix} -mgs\theta & mg\theta s\phi & mg\theta\phi & 0 & 0 & 0 \end{bmatrix}^\top$ is the gravitational vector.

A new transformation matrix $\mathbf{J}$ is considered by using the relationship $\boldsymbol{\eta} = \mathbf{J}\boldsymbol{\nu}$, where $\mathbf{J}$ is defined as

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}(\boldsymbol{\eta}) & \mathbf{0}_{6\times 1} \\ \mathbf{j}_n & \theta\phi \end{bmatrix} \tag{2.2.9}$$

and $\mathbf{J}(\boldsymbol{\eta})$ is defined as

$$\mathbf{J}(\boldsymbol{\eta}) = \begin{bmatrix} \psi\theta & s\phi s\theta\psi - \phi s\psi & \phi s\theta\psi + s\psi s\phi & 0 & 0 & 0 \\ s\psi\theta & s\phi s\theta s\psi + \phi\psi & \phi s\theta s\psi - s\phi\psi & 0 & 0 & 0 \\ -s\theta & s\phi\theta & \phi\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & s\phi t\theta & \phi t\theta \\ 0 & 0 & 0 & 0 & \phi & -s\phi \\ 0 & 0 & 0 & 0 & \frac{s\phi}{\theta} & \frac{\phi}{\theta} \end{bmatrix} \tag{2.2.10}$$

where $\mathbf{j}_n = \begin{bmatrix} -s\theta & \theta s\phi & 0 & 0 & 0 & 0 \end{bmatrix}$.

2.2.3 Control Algorithm

Geronel et al. [25] also developed a MATLAB algorithm to control the quadrotor. It can control it in three different trajectories: rectangular, circular and linear. Given the initial pose and angles, the trajectory wanted, and the control required, it is able to give a complete overview of the quadrotor's motion.

2.3 Artificial Neural Networks

2.3.1 Deep Learning

The concepts of deep learning studied in this section is going to be based on the work of Goodfellow et al. [28] and the documentation of PyTorch[1] and MATLAB[2].

There are several definitions of AI [71], but the computer scientist McCarthy [43] defines it as “the science and engineering of making intelligent machines, especially intelligent computer programs.”. He also states that “it is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.”.

The big area of study is the AI and it includes several branches like fuzzy logics, robotics, machine learning and so on. The later one, in turn, is another field with also some branches and one of them is the deep learning. This can be represented in a Venn diagram, as the Fig. 2.3.1 shows. However, all the three terms can be interchangeable in the major context.

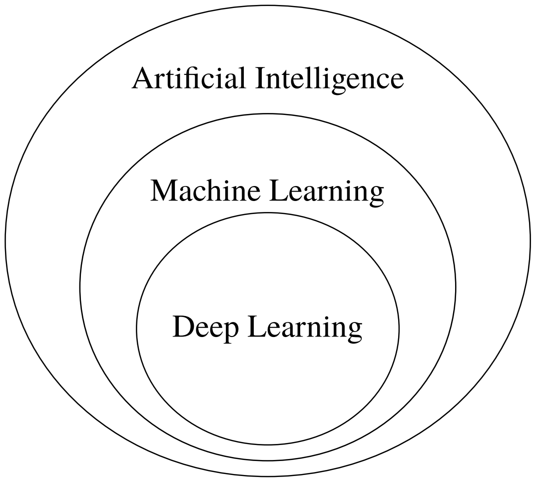


Figure 2.3.1. Subareas of Artificial Intelligence

The deep learning history goes back to the 1940s and it had several names over the years. It was called by *cybernetics* (1940s–1960s), *connectionism* (1980s–1990s), and from 2006 until now is known as *deep learning*. The DL models were engineered systems inspired by the biological brain and they were denominated Artificial Neural Network (ANN). One of the motivations of the neural perspective was to understand that the brain provides a proof by example that intelligent behavior is possible and try to reverse engineer the computation principals behind the brain, duplicating its functionality. Today it goes beyond the neuroscientist perspective and it is more of general principle of learning multiple levels of composition.

DL dwells in the programming sphere. The approach, however, it is not like the traditional programming scripts and models. To automate stuff, there are three main parts: (i) the input data, (ii) the rule (function) and (iii) the output data. In both types there are two of three parts

[1] https://pytorch.org/docs/stable/index.html
[2] https://www.mathworks.com/help/matlab/

available, but different ones for each other. In the traditional programming, there is the input data and the rule, for the algorithm output the data. In deep learning, there is the input data and the output data, for the algorithm provides the rule. A good analogy is cooking: in the traditional programming context, one has the ingredients and the recipe to make the main course; in the deep learning context, one has the ingredients and the main course to discover the recipe.

2.3.2 Neural Networks Models

A ANN is machine learning a model that simulate a biological NN to make a machine learns as the human being learns. ANN are the heart of DL and there are several models of them, each one most suitable for different kind of problems. Some of them are Multilayer Perceptron (MLP), CNN, Recurrent Neural Network (RNN), among others.

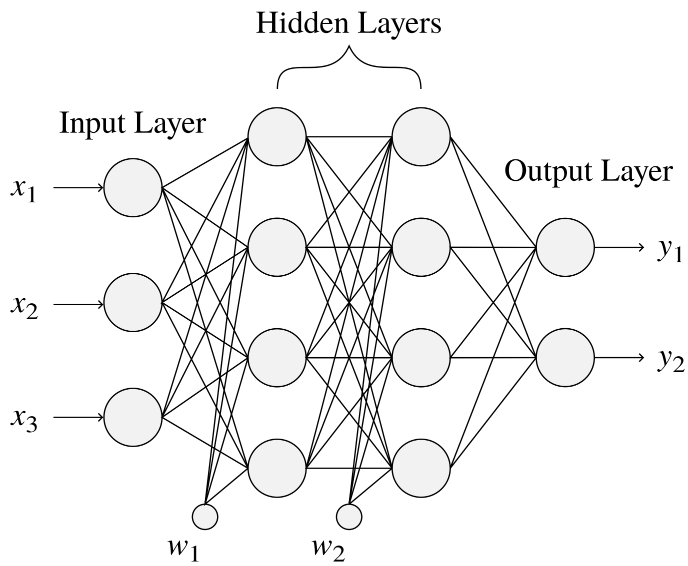


Figure 2.3.2. Visual Representation of a Artificial Neural Network. The input data x_i is given in the input layer and i can assume any integer, just as the output data y_i. The weights w_i are there to not allow the data to be biased.

Multi-layer Perceptron

Convulutional Neural Network

2.3.3 Loss Function

The *loss function*, also called *cost function* or *error function*, is the one used measure the error between the predicted output of an algorithm and the real target output. There are several loss functions suitable to different kind of situation. For each distributed data there is one that fits better. Many kinds of them are available and must be analyzed the most proper one to each case. The choice of what loss function should be picked will depend on not only the data and its pattern, but also the computational processing and the cost attached to it.

Regression

For regression problems, a common loss functions adopted is the Mean Squared Error (MSE) [12].

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2.3.1}$$

where n is the sample size; y_i is the predicted output; and $\hat{y}_i$ is the real target.

The Fig. 2.3.3 shows a linear data and how the domain of the loss function is obtained for a linear regression.

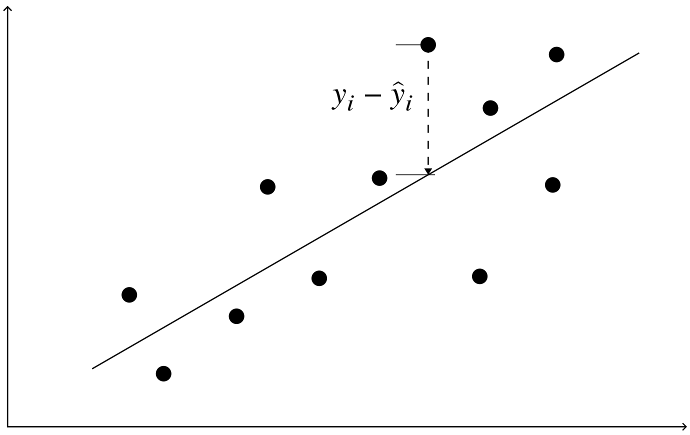


Figure 2.3.3. Loss Function for Linear Regression. The loss function take all the distances between the predicted and the target value to verify if the model is in the right path. The lower the distance, the better the model. The y-axis represents the output data and the x-axis represents the input data.

Classification

For classification problems, a common loss function adopted is the cross-entropy. It measures the “distance” between the probability distribution and the true probability, being ideal for this kind of problem. For multi-classification, i.e., when there are more than two outputs for the model, the cross-entropy is defined as

$$H(P,Q) = -\sum_{i=1}^{n} P(x)\ \log Q(x) \tag{2.3.2}$$

where Q is a discrete distribution relative P distribution (both discrete); and n is the size of the sample.

For binary classification, i.e., when there is only two outputs for the model (true or false), the cross-entropy is defined as

$$H(P,Q) = -\frac{1}{n}\sum_{i=1}^{n} x\ \log P(x) + (1-x)\ \log\left(1 - Q(x)\right) \tag{2.3.3}$$

2.3.4 Optimizer

The optimizer is an algorithm that updates the model in response to the output of the loss function, that is, it aids to minimize the loss function. As the loss function minimizes, the model is getting closer to the target values and, hence, closer to the real pattern.

Gradient Descent

The *gradient descent* is one of the main algorithm [48] that optimizes the model and many important ones are based on it, like the Stochastic Gradient Descent (SGD). The goal is to get the minimum, as the error (loss) between the predicted and the target data is null. This would mean that the model fits to the pattern of the data.

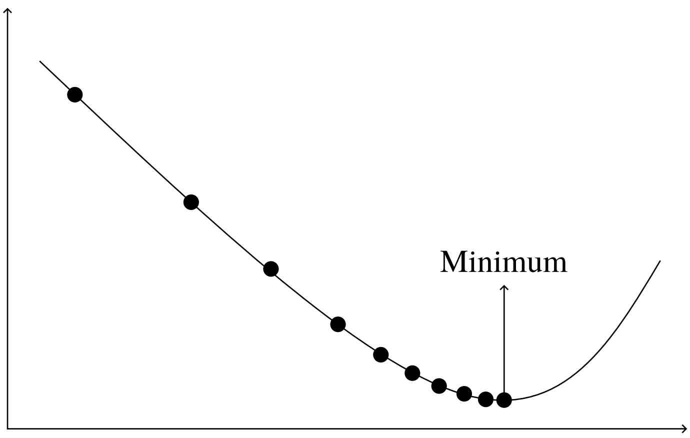


Figure 2.3.4. Gradient Descent Process. In this case, the loss function (yellow curve) can be represented in a two-axes plan. Depending on the data, it is not possible to represent graphically due to its multi dimension. Each point represents the learning step. When the gradient descent reaches the minimum of the loss function, it means that the model may be accurate. Note that the gradient descent can reach a local minimum of the function and not the global minimum necessarily. The y-axis represents the loss function and the x-axis represents the weight values.

The gradient descent is a powerful algorithm that reduces the loss function, minimizing the error between the predicted value and the target value.

Since the gradient of a function gives the direction of the steepest ascent of a function and it is orthogonal to the surface at a determined point, it seems reasonable that moving in the perpendicular direction gives the maximum increase of the function [64]. On the other hand, the negative of the gradient may be used to find the opposite, that is, the steepest descent of the function, or the minimum decrease. If the steps given to the direction of the negative gradient of the function are small, there is a good chance to get minimum value of the function. However, if the steps are too long, the chance to pass by the minimum value is high [49]. These steps are called *learning rate* and should be chosen wisely.

This way, let $\boldsymbol{x}$ be the entry vector with the predicted data and L the loss function adopted for some deep learning model, and ϵ the learning rate, the gradient descent is:

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t - \epsilon \nabla L(\boldsymbol{x}_t) \tag{2.3.4}$$

In determined cases, it is possible to avoid running the iterative algorithm and just go directly to the critical point by solving $\nabla L(\boldsymbol{x}_t) = 0$ for $\boldsymbol{x}$.

Stochastic Gradient Descent

As seen, gradient descent is a powerful tool to minimize the loss function, however, for large data, the cost of operation is very high and its use is not feasible. The main ideia of SGD is that the gradient is an expectation. Later, the data is divided in subsets, also called *mini-batch* and then the gradient is performed over them. The mini-batche size is chosen to be a realatively small numbers of examples. The data inside each subset may be considered redundant, that is why it uses one single value of the subset to compute the gradient descent. This way, the process is considerable better for computational resources.

The SGD can be written as:

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t - \frac{\epsilon}{m} \sum_{i=1}^{m} \nabla L(\boldsymbol{x}_t; p^{(i)}, q^{(i)}) \tag{2.3.5}$$

where m is the mini-batch size; and $\nabla L(\boldsymbol{x}; p^{(i)}, q^{(i)})$ is the gradient of the loss function with respect to the parameter vector $\boldsymbol{x}$ for the i^{th} example $(p^{(i)}, q^{(i)})$ in the mini-batch.

Yet, nowadays, with the amount of data, many techniques are still applied in SGD as creating an automatic adaptive learning rates which achieve the optimal rate of convergence [17] and the momentum technique to improve it [66].

Adam

Adam is an algorithm for first-order gradient-based optimization of stochastic objective functions, like the loss function, as seen in the Section 2.3.3. It is based on adaptive estimates of low-order moments and computationally efficient, requiring little computational memory. Adam is a strategical choice when using large data or parameters and with very noisy/sparse gradients [34].

The algorithm can be implemented as it follows

Adam Algorithm

Require: α: stepsize
Require: $\beta_1,\ \beta_2\ \in [0,1)$: exponential decay rates for the moment estimates
Require: $f(\theta)$: loss function
Require: θ_0: initial parameter

```
m_0 ← 0                                        ▷ Initialize first moment vector
v_0 ← 0                                        ▷ Initialize second moment vector
t ← 0                                          ▷ Initialize timesep
while θ_t not converged do
    t ← t + 1
    g_t ← ∇_θ f_t(θ_{t−1})                     ▷ Get gradients w.r.t. stochastic objective at time step t
    m_t ← b_1 · m_{t−1} + (1 − β_1) · g_t      ▷ Update biased first moment estimate
    v_t ← β_2 · v_{t−1} + (1 − β_2) · g_t^2    ▷ Update biased second raw moment estimate
    m̂_t ← m_t/(1 − β_1^t)                     ▷ Compute bias-corrected first moment estimate
    v̂_t ← v_t/(1 − β_2^t)                     ▷ Compute bias-corrected second raw moment estimate
    θ_t ← θ_{t−1} − α · m̂_t/(√(v̂_t + ε))      ▷ Update parameters
end while
return θ_t                                     ▷ Resulting parameters
```

Good default setting are $\alpha = 0.001,\ \beta_1 = 0.9,\ \beta_2 = 0.999$ and $\epsilon = 10^{-8}$. Operations on vectors are element-wise.

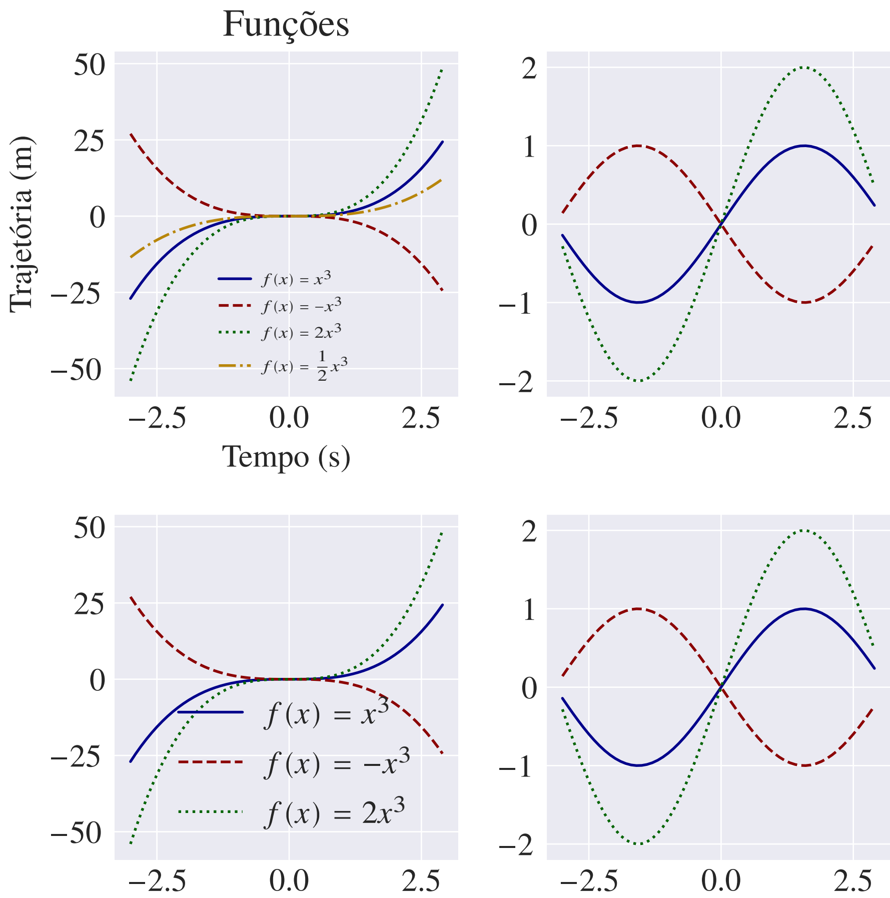
3 METHODOLOGY

The code implementation will be pragmatical and the lines of the code will not be fully explained. The frameworks methods will not be explained either, but their documentation are reasonably comprehensive with previous programming knowledge, especially in Python and MATLAB, and they are going to be linked whenever possible.

While the engineering goal of AI is to solve real-world problems using it as an equipment, the scientific goal is to determine which ideas explain the various sorts of intelligence [71] and the current objective is to use AI from the engineering perspective.

$$I = \int_0^1 f(x)\, dx \tag{3.0.1}$$

4 RESULTS AND DISCUSSION



Funções
Trajetória (m)
Tempo (s)
$f(x) = x^3$
$f(x) = -x^3$
$f(x) = 2x^3$
$f(x) = \frac{1}{2}x^3$
50
25
0
−25
−50
−2.5
0.0
2.5
2
1
0
−1
−2

5 CONCLUSION

BIBLIOGRAPHY

[1] Ahirwar, S., R. Swarnkar, S. Bhukya, and G. Namwade (2019, January). Application of Drone in Agriculture. *International Journal of Current Microbiology and Applied Sciences 8*(01), 2500–2505.

[2] Assilian, S. (1974). *Artificial Intelligence in the Controle of Real Dynamic Systems*. Ph. D. thesis, Queen Mary University of London.

[3] Aurélien, G. (2022). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc.

[4] Avci, O., O. Abdeljaber, S. Kiranyaz, and D. Inman (2017). Structural Damage Detection in Real Time: Implementation of 1D Convolutional Neural Networks for SHM Applications. In C. Niezrecki (Ed.), *Structural Health Monitoring & Damage Detection, Volume 7*, pp. 49–54. Cham: Springer International Publishing.

[5] Azimi, M., A. Eslamlou, and G. Pekcan (2020, May). Data-Driven Structural Health Monitoring and Damage Detection through Deep Learning: State-of-the-Art Review. *Sensors 20*(10), 2778.

[6] Baidoo-Anu, D. and L. Owusu Ansah (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *SSRN Electronic Journal*.

[7] Baji, T. (2018, March). Evolution of the GPU Device widely used in AI and Massive Parallel Processing. In *2018 IEEE 2nd Electron Devices Technology and Manufacturing Conference (EDTM)*, Kobe, pp. 7–9. IEEE.

[8] Balageas, D., C.-P. Fritzen, and A. Güemes (2010). *Structural Health Monitoring*, Volume 90. John Wiley & Sons.

[9] Biswas, S. S. (2023a, March). Potential Use of Chat GPT in Global Warming. *Annals of Biomedical Engineering*.

[10] Biswas, S. S. (2023b, March). Role of Chat GPT in Public Health. *Annals of Biomedical Engineering*.

[11] Brown, C., R. Kell, R. Taylor, and L. Thomas (1962, December). Piezoelectric Materials, A Review of Progress. *IRE Transactions on Component Parts 9*(4), 193–211.

[12] Bussab, W. d. O. and P. A. Morettin (2017). *Estatística Básica*. Saraiva Uni.

[13] Chan-Olmsted, S. M. (2019, October). A Review of Artificial Intelligence Adoptions in the Media Industry. *International Journal on Media Management 21*(3-4), 193–215.

[14] Cohen-Almagor, R. (2011, April). Internet History:. *International Journal of Technoethics 2*(2), 45–64.

[15] Curie, J. and P. Curie (1880). Développement par compression de l'électricité polaire dans les cristaux hémièdres à faces inclinées. *Bulletin de la Société minéralogique de France 3*(4), 90–93.

[16] Dadafshar, M. (2014). Accelerometer and Gyroscopes Sensors: Operation, Sensing, and Applications. *Maxim Integrated [online].*

[17] Darken, C. and J. E. Moody (1991). Towards Faster Stochastic Gradient Search. *4*.

[18] Davenport, T. and R. Kalakota (2019, June). The potential for artificial intelligence in healthcare. *Future Healthcare Journal 6*(2), 94–98.

[19] del Cerro, J., C. Cruz Ulloa, A. Barrientos, and J. de León Rivas (2021, January). Unmanned Aerial Vehicles in Agriculture: A Survey. *Agronomy 11*(2), 203.

[20] Duan, Y., J. S. Edwards, and Y. K. Dwivedi (2019, October). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management 48*, 63–71.

[21] Farrar, C. R. and K. Worden (2007, February). An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365*(1851), 303–315.

[22] Farrar, C. R. and K. Worden (2012). *Structural Health Monitoring: A Machine Learning Perspective*. John Wiley & Sons.

[23] Feng, D. and M. Q. Feng (2018, February). Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection – A review. *Engineering Structures 156*, 105–117.

[24] Fossen, T. I. (1994). *Guidance and Control of Ocean Vehicles*. Chichester ; New York: Wiley.

[25] Geronel, R. S., R. M. Botez, and D. D. Bueno (2023, January). Dynamic responses due to the Dryden gust of an autonomous quadrotor UAV carrying a payload. *The Aeronautical Journal 127*(1307), 116–138.

[26] Ghatrehsamani, S., G. Jha, W. Dutta, F. Molaei, F. Nazrul, M. Fortin, S. Bansal, U. Debangshi, and J. Neupane (2023, January). Artificial Intelligence Tools and Techniques to Combat Herbicide Resistant Weeds—A Review. *Sustainability 15*(3), 1843.

[27] Goda, K. and M. Kitsuregawa (2012, May). The History of Storage Systems. *Proceedings of the IEEE 100*(Special Centennial Issue), 1433–1440.

[28] Goodfellow, I., Y. Bengio, and A. Courville (2016). *Deep Learning*. MIT Press.

[29] Guo, K., Z. Yang, C.-H. Yu, and M. J. Buehler (2021). Artificial intelligence and machine learning in design of mechanical materials. *Materials Horizons 8*(4), 1153–1172.

[30] Hubáček, O., G. Šourek, and F. Železný (2019, April). Exploiting sports-betting market using machine learning. *International Journal of Forecasting 35*(2), 783–796.

[31] Jeble, S., S. Kumari, and Y. Patil (2018, January). Role of Big Data in Decision Making. *Operations and Supply Chain Management: An International Journal*, 36–44.

[32] Jiao, P., K.-J. I. Egbe, Y. Xie, A. Matin Nazar, and A. H. Alavi (2020, July). Piezoelectric Sensing Techniques in Structural Health Monitoring: A State-of-the-Art Review. *Sensors 20*(13), 3730.

[33] Karthick, N. and N. Ramalingam (2017, June). Implementation of Railway Track Crack Detection and Protection. *International Journal Of Engineering And Computer Science*.

[34] Kingma, D. P. and J. Ba (2017, January). Adam: A Method for Stochastic Optimization.

[35] Kollár, A. (2021, March). Betting models using AI: A review on ANN, SVM, and Markov Chain. Preprint, Open Science Framework.

[36] Kościelniak, H. and A. Puto (2015). BIG DATA in Decision Making Processes of Enterprises. *Procedia Computer Science 65*, 1052–1058.

[37] Lahmeri, M.-A., M. A. Kishk, and M.-S. Alouini (2021). Artificial Intelligence for UAV-Enabled Wireless Networks: A Survey. *IEEE Open Journal of the Communications Society 2*, 1015–1040.

[38] Lee, R. S. T. (2020). *Artificial Intelligence in Daily Life*. Singapore: Springer Singapore.

[39] Leiner, B. M., V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. G. Roberts, and S. Wolff (2009, October). A brief history of the internet. *ACM SIGCOMM Computer Communication Review 39*(5), 22–31.

[40] Loveday, P. W. (2000, April). Development of piezoelectric transducers for a railway integrity monitoring system. In S.-C. Liu (Ed.), *SPIE's 7th Annual International Symposium on Smart Structures and Materials*, Newport Beach, CA, pp. 330–338.

[41] Lund, B. D. and T. Wang (2023, February). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News*.

[42] Maio, C. E. B. (2011, March). *Técnicas para monitoramento de integridade estrutural usando sensores e atuadores piezoelétricos*. Mestrado em Dinâmica das Máquinas e Sistemas, Universidade de São Paulo, São Carlos.

[43] McCarthy, J. (2007). What Is Artificial Intelligence? *Stanford University*.

[44] McPhillips, D. (2022, December). Home delivery of medications can help improve access, especially when time is tight. *CNN Health*.

[45] Mhlanga, D. (2020, July). Industry 4.0 in Finance: The Impact of Artificial Intelligence (AI) on Digital Financial Inclusion. *International Journal of Financial Studies 8*(3), 45.

[46] Milana, C. and A. Ashta (2021, May). Artificial intelligence techniques in finance and financial markets: A survey of the literature. *Strategic Change 30*(3), 189–209.

[47] Muthukrishnan, N., F. Maleki, K. Ovens, C. Reinhold, B. Forghani, and R. Forghani (2020, November). Brief History of Artificial Intelligence. *Neuroimaging Clinics of North America 30*(4), 393–399.

[48] Nesterov, I. E. (2004). *Introductory Lectures on Convex Optimization: A Basic Course*. Number v. 87 in Applied Optimization. Boston: Kluwer Academic Publishers.

[49] Nielsen, M. (2015). *Neural Networks and Deep Learning*, Volume 25. Determination press San Francisco, CA, USA.

[50] Pannu, A. (2015). Artificial Intelligence and its Application in Different Areas. *4*(10).

[51] Poola, I. (2017). How Artificial Intelligence in Impacting Real life Everyday. *International Journal for Advance Research and Development 2*(10), 96–100.

[52] Rabunal, J. R. and J. Dorado (Eds.) (2006). *Artificial Neural Networks in Real-Life Applications*. Hershey PA: Idea Group Pub.

[53] Ramakrishnan, M., P. G. Poojari, M. Rashid, S. Nair, V. Pulikkel Chandran, and G. Thunga (2023, March). Impact of COVID-19 pandemic on medicine supply chain for patients with chronic diseases: Experiences of the community pharmacists. *Clinical Epidemiology and Global Health 20*, 101243.

[54] Raschka, S. (2015). *Python Machine Learning*. Packt Publishing Ltd.

[55] Raschka, S., Y. H. Liu, V. Mirjalili, and D. Dzhulgakov (2022). *Machine Learning with PyTorch and Scikit-Learn: Develop Machine Learning and Deep Learning Models with Python*. Packt Publishing Ltd.

[56] Ruiz-Real, J. L., J. Uribe-Toril, J. A. Torres, and J. De Pablo (2020, October). Artificial Intelligence in Business and Economics Research: Trends and Future. *Journal of Business Economics and Management 22*(1), 98–117.

[57] Sabato, A., C. Niezrecki, and G. Fortino (2017, January). Wireless MEMS-Based Accelerometer Sensor Boards for Structural Vibration Monitoring: A Review. *IEEE Sensors Journal 17*(2), 226–235.

[58] Sakena Benazer, S., M. Sheik Dawood, S. Karthick Ramanathan, and G. Saranya (2021). Efficient model for IoT based railway crack detection system. *Materials Today: Proceedings 45*, 2789–2792.

[59] Sankarasrinivasan, S., E. Balasubramanian, K. Karthik, U. Chandrasekar, and R. Gupta (2015). Health Monitoring of Civil Structures with Integrated UAV and Image Processing System. *Procedia Computer Science 54*, 508–515.

[60] Sharma, A. and N. Mehta (2016, August). Structural Health Monitoring Using Image Processing Techniques-A Review.

[61] Sivakumar, M. and N. M. Tyj (2021). A Literature Survey of Unmanned Aerial Vehicle Usage for Civil Applications. *Journal of Aerospace Technology and Management 13*, e4021.

[62] Smarsly, K., K. Lehner, and D. Hartmann (2007). Structural Health Monitoring based on Artificial Intelligence Techniques. In *Computing in Civil Engineering (2007)*, pp. 111–118.

[63] Sohn, H., C. R. Farrar, F. Hemez, and J. Czarnecki (2003). A Review of Structural Health Monitoring Literature 1996 – 200. *Los Alamos National Laboratory, USA 1*, 16.

[64] Stewart, J. (2016). *Calculus* (EIghth edition ed.). Boston, MA, USA: Cengage Learning.

[65] Sun, L., Z. Shang, Y. Xia, S. Bhowmick, and S. Nagarajaiah (2020). Review of Bridge Structural Health Monitoring Aided by Big Data and Artificial Intelligence: From Condition Assessment to Damage Detection. *Journal of Structural Engineering 146*.

[66] Sutskever, I., J. Martens, G. Dahl, and G. Hinton (2013). On the Importance of Initialization and Momentum in Deep Learning. pp. 1139–1147.

[67] Thiels, C. A., J. M. Aho, S. P. Zietlow, and D. H. Jenkins (2015, March). Use of Unmanned Aerial Vehicles for Medical Product Transport. *Air Medical Journal 34*(2), 104–108.

[68] Tzounis, A., N. Katsoulas, T. Bartzanas, and C. Kittas (2017, December). Internet of Things in agriculture, recent advances and future challenges. *Biosystems Engineering 164*, 31–48.

[69] Verdouw, C., S. Wolfert, and B. Tekinerdogan (2016, January). Internet of Things in agriculture. *CABI Reviews 2016*, 1–12.

[70] Verma, S., R. Sharma, S. Deb, and D. Maitra (2021, April). Artificial intelligence in marketing: Systematic review and future research direction. *International Journal of Information Management Data Insights 1*(1), 100002.

[71] Winston, P. H. (1992). *Artificial Intelligence* (3rd ed ed.). Reading, Mass: Addison-Wesley Pub. Co.

[72] Ye, X., T. Jin, and C. Yun (2019, November). A review on deep learning-based structural health monitoring of civil infrastructures. *Smart Structures and Systems 24*(5), 567–585.

[73] Yu, K.-H., A. L. Beam, and I. S. Kohane (2018, October). Artificial intelligence in healthcare. *Nature Biomedical Engineering 2*(10), 719–731.