Advanced Deep Regression Models for Smart Operation of the Oil and Gas Industry

Siavash Hosseini

Electrical and Computer Engineering

Lakehead University, Thunder Bay, Ontario

A thesis submitted to Lakehead University in partial fulfillment
of the requirements for the Master of Science degree
in the Electrical and Computer Engineering

Examining Committee Membership

The following served on the Examining Committee for this thesis:

Supervisor:	Dr. Thangarajah Akilan Assistant Professor, Department of Software Engineering.
Internal Committee Member:	Dr. Abdulsalam Yassine Assistant Professor, Department of Software Engineering.
External Committee Member:	Dr. Garima Bajwa Assistant Professor, Department of Computer Science.
Session Chair:	Dr. Yushi Zhou Assistant Professor, Department of Electrical and Computer Engineering.

Declaration of Co-Authorship / Previous Publications

I. Co-Authorship Declaration

I hereby declare that this dissertation incorporates material that is the result of joint research, as follows: This dissertation also incorporates the outcome of research under the supervision of Dr. T. Akilan and collaboration with Dr. G.H. Roshani and Dr. E. Nazemi (Chapter 2). In all cases, the key ideas, primary contributions, experimental designs, data analysis, interpretation, and writing were performed by the author, and the contribution of the coauthors was primarily through the provision of proofreading and reviewing the research papers regarding the technical content.

I am aware of the Lakehead University Policy on Authorship, and I certify that I have properly acknowledged the contribution of other researchers to my dissertation and have obtained permission from each co-author to include the above materials in my dissertation.

I certify that, with the above qualification, this dissertation, and the research to which it refers, is the product of my work.

II. Declaration of Previous Publications

This thesis includes 4 original papers that have been previously published/submitted to journals for publication, as follows:

Thesis chapter	Publication title/full citation	Status
Chapter 3	S. Hosseini and T. Akilan et al., Application of Wavelet Feature Extraction and Artificial Neural Networks for Improving the Performance of Gas–Liquid Two-Phase Flow Meters Used in Oil and Petro- chemical Industries, Polymers, 13, 3647, 2021	Published
	S. Hosseini and T. Akilan et al., Accurate Flow Regime Classification and Void Fraction Measurement in Two-phase Flowmeters using frequency-domain feature extraction and neural networks, Separations 9.7: 160, 2022.	Published
Chapter 4	S. Hosseini, A. Shahbandegan, and T. Akilan, Deep Neural Network Modeling for Accurate Electric Motor Temperature Prediction, 2022 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE). IEEE, 2022.	Published
Chapter 5	S. Hosseini and T. Akilan, Advanced Deep Regression Models for Forecasting Time Series Oil Production" IEEE Transactions on Industrial Informatics.	Submitted

III. General

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act. I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office. This thesis has not been submitted for a higher degree to any other University or Institution.

Acknowledgements

I would like to express my deepest gratitude to my thesis supervisor, Dr. Thangarajah Akilan, whose guidance, encouragement, and support were invaluable throughout my research journey. Their expertise, insights and constructive comments helped me refine my ideas and improve the quality of my work.

I am also grateful to the members of my thesis committee, Dr. Abdulsalam Yassine, Dr. Garima Bajwa, and Dr. Yushi Zhou, for their valuable feedback and suggestions.

I am indebted to my family for their unwavering love, support, and encouragement. Their belief in me gave me the strength to persevere and overcome the challenges that I faced.

Dedication

This thesis is dedicated to my dear parents.

Abstract

The first industrial revolution in the early 18th century largely exploited steam power to replace animal labor. Since then, there has been rapid development in industrial operations. Now, the world has come to the brink of the fifth industrial revolution, a.k.a. industry 5.0, where industries invest in building intelligent systems to perform complex actions more efficiently by leveraging technological advancements, including big data, and high-performance computing (HPC) platforms. Thus, modern artificial intelligence (AI), particularly deep neural networks (DNNs) has emerged as a powerful tool in industries for informed operational control, real-time fault and anomaly detection, and maintenance.

In this direction, this research aims to develop advanced regression models using artificial neural network (ANN), 1-D convolutional neural network (CNN), and long short-term memory (LSTM) units for key operations in the oil and gas industries. More specifically, this study focuses on three stages, namely drilling, transportation, and production, and proposes robust regressors for accurate prediction of void fraction, the temperature of internal components of electric motors, and the production level of hydrocarbon extracts. A precise prediction of these factors will increase resource efficiency, energy saving, and product quality, and decrease environmental pollution. An exhaustive experimental study conducted on benchmark datasets demonstrates the practicability of the proposed solutions and their robustness. It is worth mentioning that Canada is the world's fifth-largest oil producer and has one of the world's largest oil reserves and the world's third-largest proven oil reserves.

Table of Contents

List of Figures

List of Tables

List of Acronyms

Acronym	Description
ANN	Artificial Neural Network
AI	Artificial Intelligence
ML	Machine Learning
NN	Neural Network
MLP	Multi-layer Perceptrons
DL	Deep Learning
DNN	Deep Neural Network
SVM	Support Vector Machine
KNN	K-nearest Neighbours
GA	Genetic Algorithm
CNN	Convolutional Neural Network
1-D CNN	One Dimentional Convolutional Neural Network
FCN	Fully Convolutional Network
FNN	Feed-forward Neural Network
LSTM	Long Short-term Memory
RNN	Recurrent Neural Network
TCN	Temporal Convolutional Network
ReLU	The Rectified Linear Unit
Tanh	Hyperbolic Tangent

Acronym	Description
MCNP	Mont Carlo N-Particle
FFT	Fast Fourier Transform
STD	Standard deviation
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
MSE	Mean Squared Error
PMSM	Permanent Magnet Synchronous Motor
LPTN	Lumped-Parameter Thermal Networks
EWMA	Exponentially Weighted Moving Avarage
EWMS	Exponentially Weighted Moving Standard Deviation
mAP	Mean Average Performance
NARX	Nonlinear Auto-regressive Exogenous
LR	Linear Regression
HPC	High-Performance Computing
IAT	Intelligent Assistive tool
IoT	Internet of Things
MA	Meta-heuristic Algorithm
LCC	Life-cycle Cost
EnKf	Ensemble Kalman Filter
GRU	Gated Recurrent Unit
DCA	Declined Curve Analysis
PSO	Particle Swarm Optimization
HM	Historical Matching
FLOP	Floating Point Operation
DNR	Dendritic Neural Regression
WOA	Whale Optimization Algorithm

Acronym	Description
SCA	Sine-cosine Algorithm
DE	Differential Evolution
HS	Harmony Search Algorithm

Chapter 1

Introduction

1.1 Overall description of the problems and challenges

Canada is one of the world's largest oil producers and exporters. The oil and gas industry is a significant contributor to the country's economy. The integration of AI at various stages of oil and gas operations can enhance efficiency, reduce costs, and optimize drilling, transportation, and production processes (Sadjadi and Sadi-Nezhad, 2017). The oil and gas industry is comprised of three main stages: drilling, transportation, and production, each presenting unique challenges and opportunities (MacKay et al., 2021). During the drilling and extraction stage, accurate prediction of the volume of extracted materials is vital for assessing the profitability of the drilling process. Flow meter devices can provide this information (Hosseini et al., 2020). Electric motors are critical for pumping and transferring extracted materials to reservoirs in the transportation stage. Monitoring the motors' temperature can ensure safe operation and prevent disruptions caused by equipment failure (Hosseini et al., 2022b). Finally, in the production stage, production forecasting utilizes various parameters such as temperature and pressure to determine the rate of production. With the aid of AI, future production levels can be predicted and adjusted to meet market demand and price fluctuations, thereby increasing profitability. Fig. 1.1 shows the respective main stages in the oil and gas industry studied in this work.
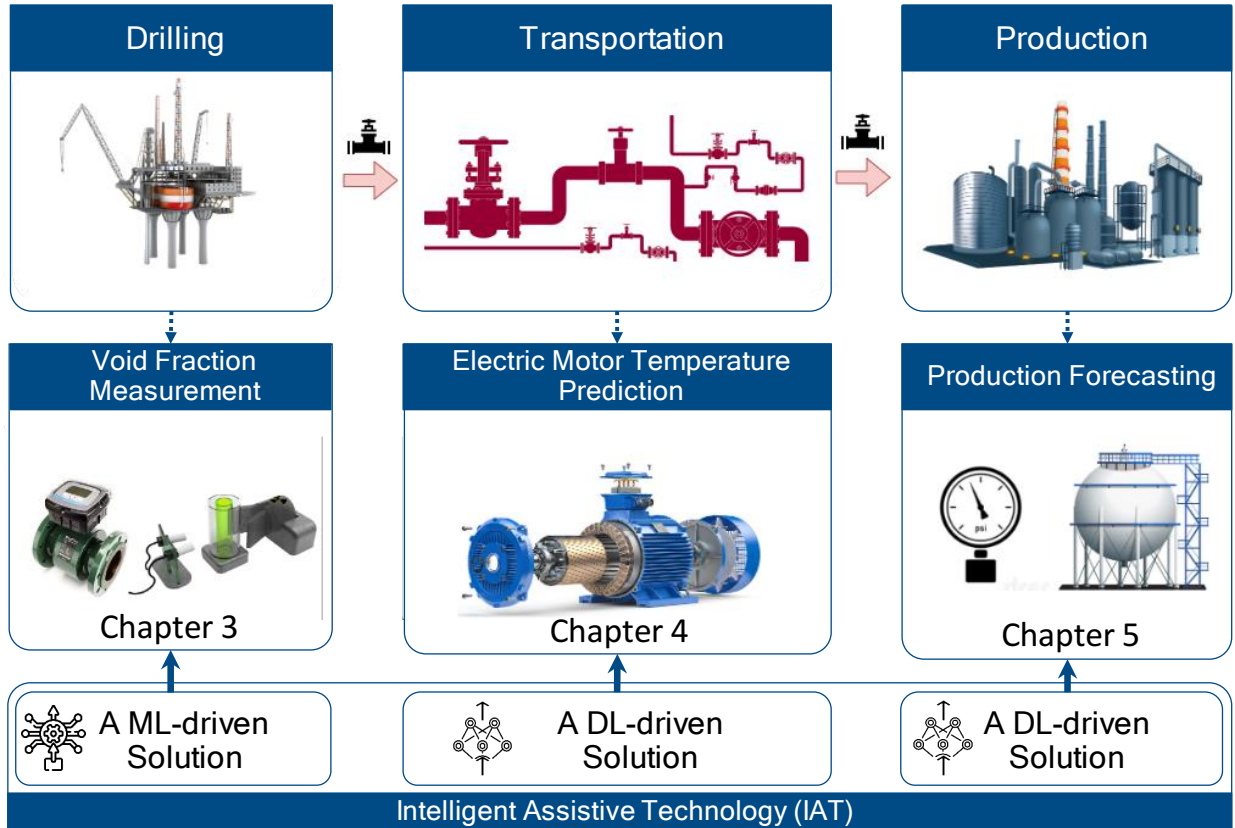
Figure 1.1: Overall illustration of three main stages in the oil and gas industry. It starts with drilling, followed by transportation and then end up with production. Chapters 2, 3 and 4 are allocated to discuss main problems and proposing solutions for each stage separately

1.2 Motivation

The oil and gas industry plays a vital role in the global economy, and its efficient operation is crucial for meeting the ever-increasing demand for energy (Hosseini et al., 2022a). The industry faces various challenges, such as the accurate prediction of key operational parameters, maintenance of equipment, and environmental concerns. The application of machine learning and deep learning techniques can significantly improve the efficiency of oil and gas operations by providing accurate predictions of critical parameters, reducing the downtime of equipment, and minimizing the environmental impact.

1.3 Technical approach

The proposed approach involves developing advanced regression models for key operations in the oil and gas industries. Three stages, namely drilling, transportation, and production, are considered in this study. The research proposes robust regressors for accurate prediction of void fraction, the temperature of internal components of electric motors, and the production level of hydrocarbon products. The proposed solutions are based on extensive experimental studies conducted on benchmark data sets. The models are trained on a large amount of historical data to predict the critical parameters with high accuracy. The results of the proposed models are compared with the existing state-of-the-art methods to demonstrate their superiority.

1.4 AI paradigm evolution

The field of Artificial Intelligence has come a long way since its inception in the 1950s. The early days of AI were marked by enthusiasm and optimism, as researchers believed that machines could be designed to mimic human intelligence in all its complexity. However, the field soon hit a roadblock as it became clear that the existing techniques were inadequate to handle the intricacies of real-world problems. This led to what is known as the "AI Winter," a period of disillusionment and lack of progress in the field. The AI paradigm underwent a major evolution in the 1980s with the advent of expert systems. These were computer programs that could reason and make decisions based on a set of rules and a knowledge base. Expert systems were widely used in various domains, such as medicine, finance, and engineering, and were considered a major breakthrough in AI. However, their limitations soon became apparent, as they were unable to handle problems that were too complex or that required a high degree of flexibility. Fig. 1.2 shows a brief overview of the AI evolution.

The 1990s saw the emergence of machine learning, which marked another significant paradigm shift in AI. Machine learning is a technique that allows machines to learn from data without being explicitly programmed. This was made possible by the availability of large data sets and powerful computational resources. Machine learning algorithms such as neural networks, decision trees,
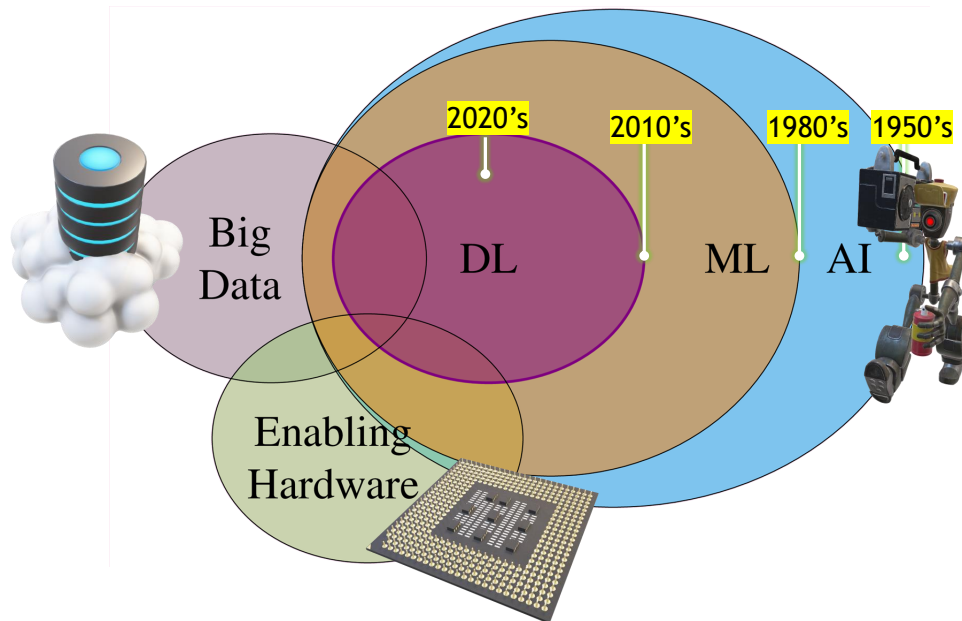
Figure 1.2: A concise overview of the evolution of AI from expert systems to the current paradigm of machine learning integrated with other technologies, with the potential to revolutionize industries.

and support vector machines became popular and were used in a wide range of applications, such as speech recognition, image classification, and natural language processing.

The 2000s saw the rise of Big Data and the Internet of Things (IoT), which generated an unprecedented amount of data from various sources. This data explosion led to the development of new machine learning techniques such as deep learning, which involves training neural networks with many layers. Deep learning has been instrumental in achieving breakthroughs in various domains, such as computer vision, speech recognition, and natural language processing. The current paradigm in AI is characterized by the integration of machine learning with other technologies such as robotics, natural language processing, and computer vision. This has led to the development of intelligent systems that can interact with humans and perform complex tasks. For example, autonomous vehicles, smart homes, and virtual assistants are all examples of intelligent systems that are based on AI. The evolution of the AI paradigm has been marked by various breakthroughs and setbacks. However, with the current advances in machine learning and the integration of other

technologies, AI has the potential to revolutionize various industries and change the way we live and work.

1.5 Artificial neural network

1.5.1 Overview

A biological neuron, also known as a nerve cell, is the fundamental unit of the nervous system. It receives, processes, and transmits information through electrochemical signals. The structure of a biological neuron consists of three parts: the cell body, dendrites, and axon. The cell body contains the nucleus and other essential organelles. Dendrites are branching structures that receive signals from other neurons or sensory receptors. The axon is a long, slender projection that sends signals to other neurons or muscles. When a signal arrives at the dendrites, it triggers an electrical charge that travels down the axon and releases chemicals called neurotransmitters at the end of the axon. These neurotransmitters then bind to receptors on the dendrites of other neurons, causing a new electrical signal to be generated, and the process continues. Artificial neural networks are computer systems inspired by the biological neural networks found in the nervous system. Fig. 1.3 shows
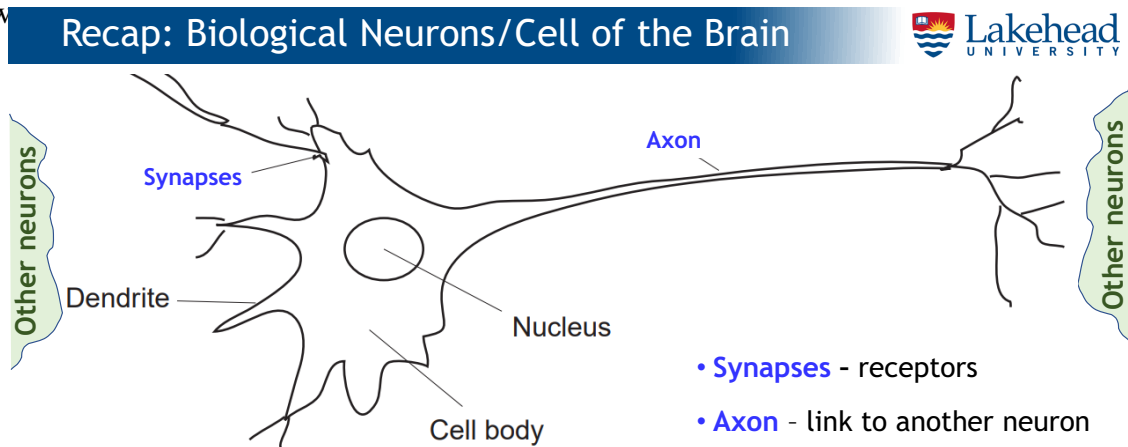
Figure 1.3: A simple representation of a biological neuron.

ANNs are a class of machine learning algorithms that are inspired by the structure and function of the human brain. They are composed of interconnected units, known as artificial neurons,

which process and transmit information. The history of ANNs can be traced back to the 1940s, when Warren McCulloch and Walter Pitts proposed the first mathematical model of a neuron. They demonstrated that a simple network of these neurons could perform logical operations and suggested that such networks could be used for computational purposes.

1.5.2 Multi-layer perceptron

In the 1950s and 1960s, researchers such as Frank Rosenblatt (Fig. 1.4) and Bernard Widrow devel-

blatt (1928 – 1971): In 1957, he introduced the "percep
cribed as the first machine "capable of having an origina

Σ ⟶ *z*

An IBM 704 comput

Image Ack: https://news.cornell.edu/stories/2019/09/professors-perceptron-paved-

(https://en.w
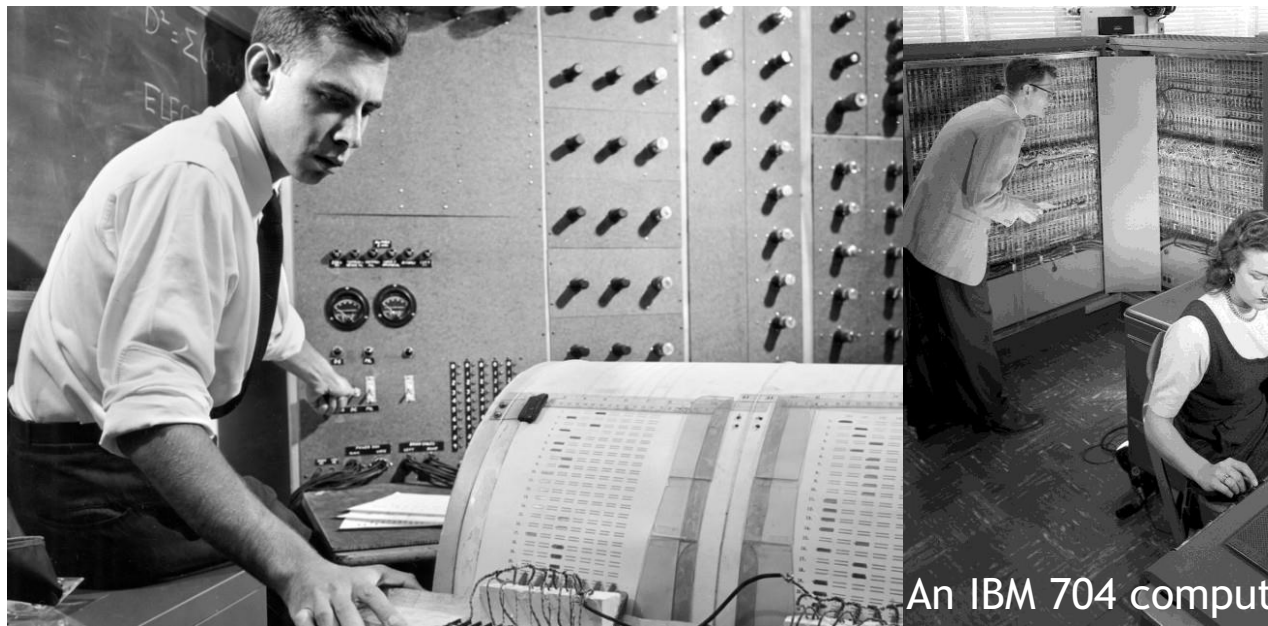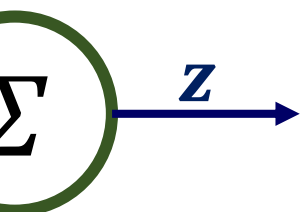
Figure 1.4: Frank Rosenblatt (1928 1971). In 1957, he introduced the " perceptron" (cor).

den) layer NN that classifies input into two possible cate

ANNs found applications in a wide range of fields, including computer vision, speech recognition,

or dog or cat – and if it's wrong, tweaks itself to make a

and natural language processing. This led to the development of more advanced neural network ar-

diction next time.

9

chitectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), which are still widely used today.

The basic building block of an ANN is the artificial neuron, which is a mathematical function that receives a number of inputs and produces a single output. The inputs are weighted, and the weights are adjusted during the training process to optimize the network's performance. An artificial neuron can be represented mathematically as follows:

$$y = f(w_1 \times x_1 + w_2 \times x_2 + ... + w_n \times x_n), \tag{1.1}$$

where y is the output of the neuron, x_1, x_2, ..., x_n are the inputs, and w_1, w_2, ..., w_n are the weights. The function f is known as the activation function, and it determines the output of the neuron based on the input and weights. Figs. 1.5 and 1.6 show the structure of a single neuron and general

Recap: Artificial Neuron – A Simple Computing Element   Lakehead UNIVERSITY

Neuron i in hidden layer l

→ Activation Functions   © Dr. T. Akilan   7
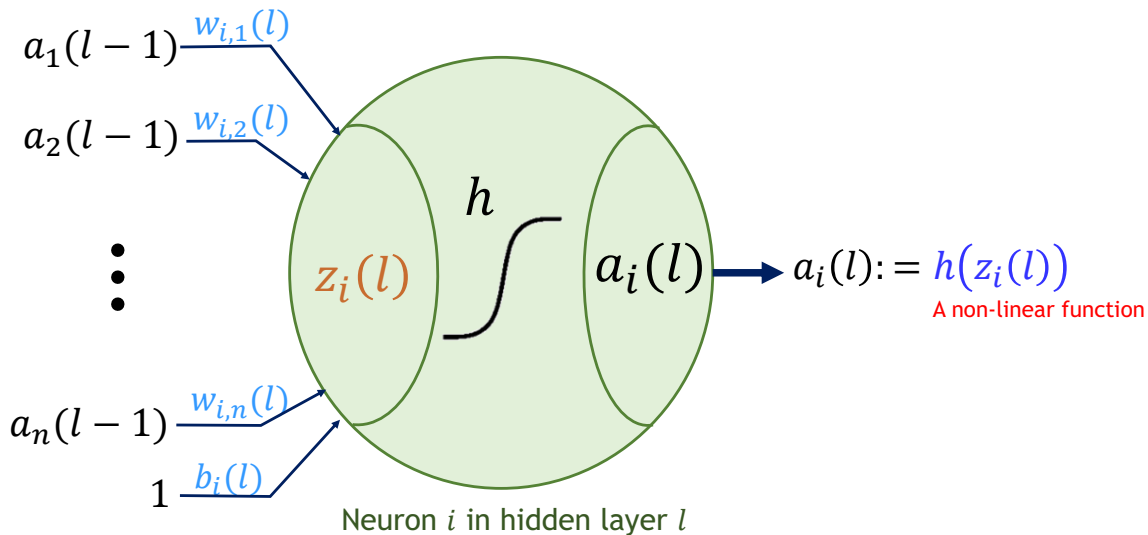
Figure 1.5: A simple illustration of a single computational unit (neuron). Here l, l-1, l+1 and i refer to current layer, previous layer, next layer and neuron index, respectively. It should be noted that $z_i(l)$ is a linear function that can be defined as $z_i(l) = \sum_j w_{i,j} \times a_j$
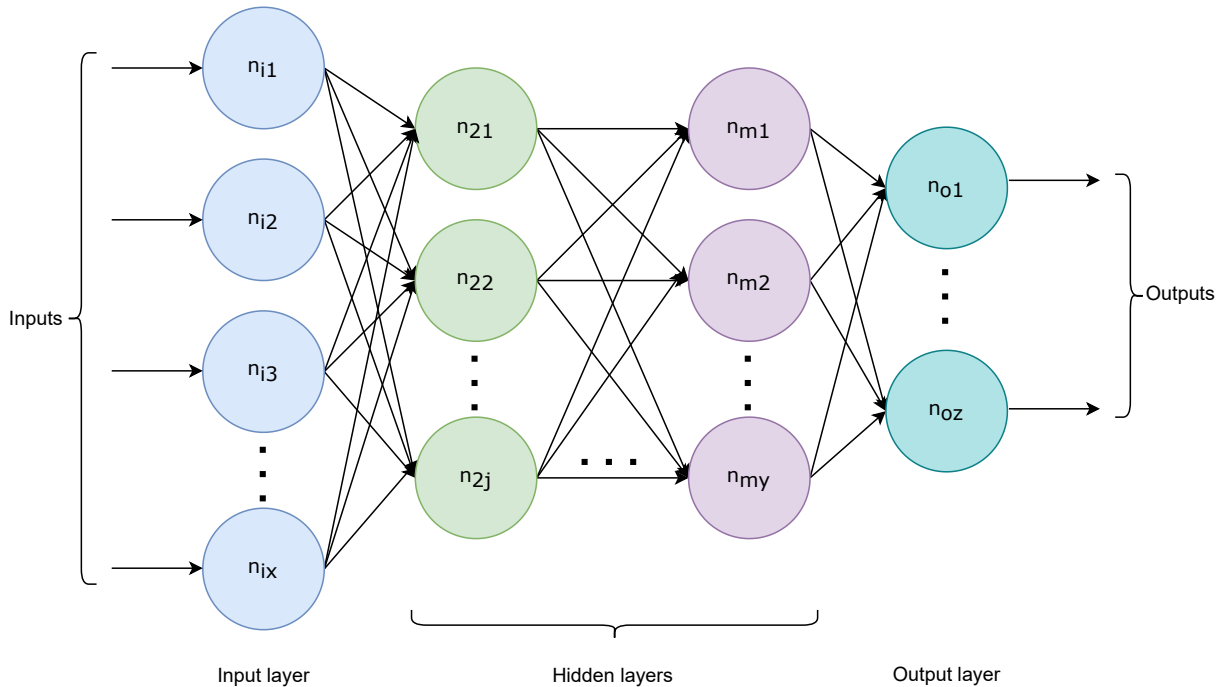
Figure 1.6: General illustration of a simple neural network where n, i, o stand for neuron, input and output layers. Moreover, x, j, y and z denote to the number of neurons in the input layer, hidden layers and output layer, respectively.

1.5.3 Activation function

In ANNs, an activation function is used to introduce nonlinearity into the output of a neuron. The choice of activation function can greatly affect the performance of the network, and several different functions have been proposed over the years. Some of the most commonly used activation functions are as follows: sigmoid, ReLU, tanh, softmax and leaky ReLU. Fig. 1.7 shows several common activation functions.

Sigmoid: The sigmoid function is a classic activation function that maps any input value to a value between 0 and 1. The sigmoid function is defined as (1.2):

$$h_\theta(x) = \frac{1}{1 + e^{-x}}, \tag{1.2}$$

Where x is the input to the function. The sigmoid function is commonly used in binary classification problems, where the goal is to predict whether an input belongs to one of two classes.

Step	Sign	Sigmoid	Identity
$h(z) = \begin{cases} 1, z \geq 0 \\ 0, z < 0 \end{cases}$	$h(z) = \begin{cases} +1, z \geq 0 \\ -1, z < 0 \end{cases}$	$h(z) = \dfrac{1}{1+e^{-z}}$	$h(z) = z$
$[0, 1]$	$[-1, 1]$	$(0, 1)$	$(-\infty, \infty)$
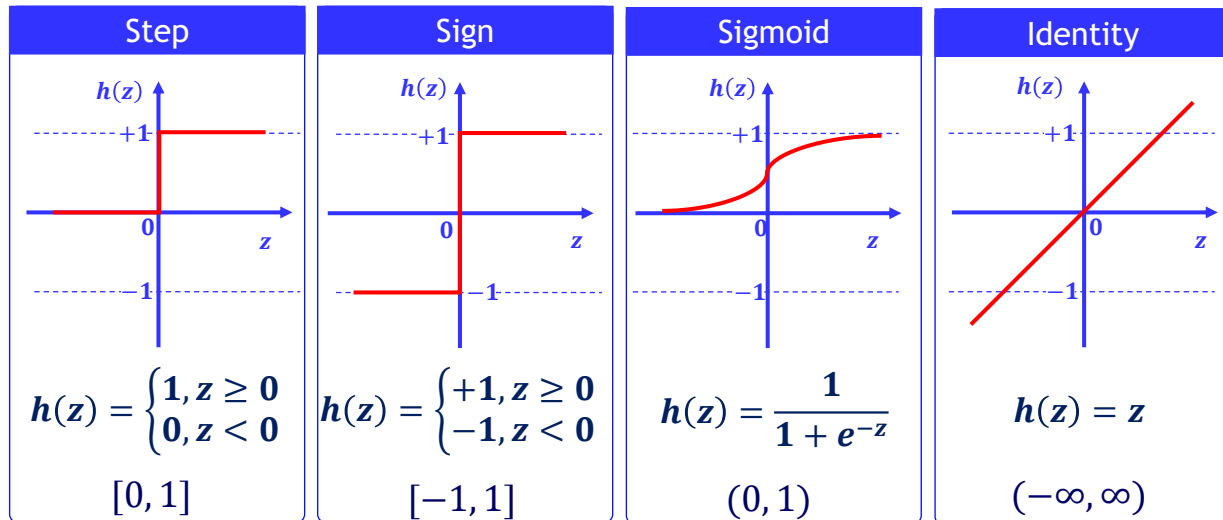
- What is the main purpose of using activation functions?

Figure 1.7: Common types of activation functions.

© Dr. T. Akilan   8

ReLU: The Rectified Linear Unit (ReLU) function is a simple and popular activation function that sets any negative input to zero and leaves positive inputs unchanged. The ReLU function is defined as (1.3):

$$F(x) = max(0, x), \tag{1.3}$$

Where x is the input to the function. The ReLU function is commonly used in deep learning architectures because of its simplicity and effectiveness.

Tanh: The hyperbolic tangent (tanh) function is similar to the sigmoid function, but it maps input values to a range between -1 and 1. The tanh function is defined as (1.4):

$$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{1 - e^{-2x}}{1 + e^{-2x}}, \tag{1.4}$$

Where x is the input to the function. The tanh function is often used in image processing and speech recognition applications.

Softmax: The softmax function is commonly used in multiclass classification problems, where the goal is to predict which of several classes an input belongs to. The softmax function maps the outputs of multiple neurons to a probability distribution over the possible classes. The softmax function is defined as (1.5):

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad for \ i = 1, 2, \ldots, K \tag{1.5}$$

Where z_i is the output of the $i - th$ neuron, and j ranges over all the neurons in the output layer. The softmax function ensures that the probabilities of all the possible classes sum to 1.

Leaky ReLU: The Leaky ReLU function is a modification of the ReLU function that allows a small gradient when the input is negative. This can help prevent the "dying ReLU" problem, where neurons that are stuck in the negative side of the ReLU function never activate. The Leaky ReLU function is defined as (1.6):

$$F(x) = max(ax, x), \tag{1.6}$$

Where a is a small positive constant, usually 0.01 or less.

Activation functions play a crucial role in determining the effectiveness of artificial neural networks. The choice of activation function should be based on the specific problem being solved, and experimentation is often required to find the best function for a particular application.

1.5.4 Training strategy

ANNs are organized into layers, with the input layer receiving the raw data and the output layer producing the final prediction. Between the input and output layers, there may be one or more hidden layers, which contain additional artificial neurons that process the data and transmit it to the next layer. The training process involves adjusting the weights of the network in order to minimize the error between the predicted output and the desired output. This is typically done using an optimization algorithm, such as stochastic gradient descent, which iteratively adjusts the weights in a way that reduces the error. Forward propagation and backpropagation are two fundamental concepts in the training of the ANNs. Forward propagation refers to the process of passing the input data through the network, layer by layer until it reaches the output layer. Backpropagation, on the other hand, is the process of adjusting the weights of the network in order to minimize the error between the predicted output and the desired output. Once the prediction has been made, the back propagation algorithm is used to update the weights and biases of the network in order

to minimize the cost function. The cost function measures the difference between the predicted output and the actual output.

Gradient descent is an iterative optimization algorithm used to update the weights of a neural network during training. The basic idea behind gradient descent is to compute the gradient of the loss function with respect to the weights, and then update the weights in the direction that minimizes the loss. The algorithm works as follows:

- Initialize the weights of the neural network with random values.

- Feed a batch of training data through the network and compute the output.

- Compute the loss between the predicted output and the true output.

- Compute the gradient of the loss with respect to each weight in the network using backpropagation.

- Update the weights in the direction of the negative gradient, multiplied by a learning rate hyperparameter using Eq. (1.7), which controls the size of the updates.

$$w_{i+1} = w_i - \alpha \nabla_w L(w_i) \tag{1.7}$$

- Repeat the above steps for a fixed number of iterations or until the loss converges.

The update rule above is called the gradient descent update rule. The negative gradient tells us the direction of steepest descent, and multiplying by the learning rate controls the step size. The learning rate is a hyperparameter that can be tuned to achieve faster convergence without overshooting the minimum.

Here, w_i is the weight vector at iteration i, $L(w_i)$ is the loss function evaluated at w_i, $\nabla_w L(w_i)$ is the gradient of the loss function with respect to w_i, and α is the learning rate. The gradient $\nabla_w L(w_i)$ is a vector that contains the partial derivatives of the loss function with respect to each weight in the network. Intuitively, the gradient tells us how much the loss would change if we were to nudge each weight up or down slightly. By moving the weights in the direction of the negative

gradient, we can decrease the loss and improve the performance of the network. The learning rate α is a hyperparameter that controls the size of the weight updates. If the learning rate is too high, the weight updates can be too large and cause the optimization process to overshoot the minimum of the loss function. On the other hand, if the learning rate is too low, the optimization process can be very slow and take a long time to converge. The update rule $w_{i+1} = w_i - \alpha \nabla_w L(w_i)$ is repeated for a fixed number of iterations or until the loss converges. Each iteration of the algorithm involves computing the gradient of the loss with respect to the weights, and then updating the weights accordingly. As the weights are updated, the loss generally decreases, and the network becomes better at making predictions.

Overall, the equation for gradient descent is a fundamental component of neural network optimization. By iteratively updating the weights in the direction of the negative gradient, we can optimize the network to minimize the loss and make accurate predictions.

1.6 Deep neural network

1.6.1 Overview

Simple neural networks, also known as shallow neural networks, have just a few layers of neurons and are not able to learn and represent very complex patterns in data. These networks were developed in the 1980s and were initially used for tasks such as image and speech recognition. However, their performance was not as good as desired, and they were not able to achieve the same level of accuracy as humans. Deep neural networks, on the other hand, have many layers of neurons and are able to learn and represent very complex patterns in data. The additional layers and increased learning capacity of deep neural networks allow them to learn and represent patterns that are much more complex than those that can be learned by simple neural networks. Enhancements in deep learning have led to significant progress in a wide range of applications, including image and speech recognition, natural language processing, and machine translation. The development of deep learning can be traced back to the 1980s and 1990s when researchers started to explore the use of multiple layers in neural networks. However, it was not until the early 2010s,

with the availability of large data sets and the development of more powerful computing hardware, that deep learning began to achieve significant performance improvements and gained widespread attention. Today, deep learning is an active area of research and is being applied in a wide range of fields.

The oil and gas industry has recently seen a surge in the application of deep learning techniques, particularly in the areas of exploration and production. Deep learning involves the use of artificial neural networks to analyze large amounts of data and make predictions or decisions. This technology has the ability to learn and adapt to new data, making it particularly well-suited for the complex and constantly changing nature of the oil and gas industry. One key application of deep learning in the oil and gas industry is in the exploration and drilling process. Traditionally, this process has been heavily reliant on geologists and geophysicists to interpret data and identify potential drilling locations. However, with the use of deep learning techniques, this process can be greatly enhanced and made more efficient. By analyzing data from various sources such as seismic surveys, well logs, and geological maps, deep learning algorithms can identify patterns and trends that may indicate the presence of oil or gas deposits. This can greatly reduce the time and cost of the exploration process, as well as increase the chances of success. Another area where deep learning has been applied in the oil and gas industry is in the prediction and prevention of equipment failures. The maintenance and repair of oil and gas production equipment is a critical aspect of the industry, as any downtime can result in significant losses. By analyzing data from sensors and other sources, deep learning algorithms can identify patterns and trends that may indicate an impending failure. This can allow for proactive maintenance and repair, reducing the risk of unplanned downtime and increasing overall efficiency.

Deep learning has also been applied in the prediction of reservoir performance. In the oil and gas industry, it is important to accurately predict the yield of a reservoir in order to optimize production. By analyzing data from sources such as well logs and production data, deep learning algorithms can make more accurate predictions of reservoir performance. This can help companies make more informed decisions about their operations, leading to increased efficiency and profitability. Another area where deep learning has been applied in the oil and gas industry is in the

optimization of drilling and completion operations. By analyzing data from various sources such as drilling rigs and well logs, deep learning algorithms can identify patterns and trends that can help optimize drilling and completion operations. This can result in reduced costs and increased efficiency, leading to improved profitability.

Deep models have also been applied in the area of environmental monitoring and compliance. The oil and gas industry is heavily regulated, with strict requirements for monitoring and reporting environmental impacts. By analyzing data from sources such as sensors and monitoring equipment, deep learning algorithms can help companies identify and address potential environmental issues in real time, reducing the risk of non-compliance and fines. Generally, the oil and gas industry has seen a significant increase in the application of deep learning techniques in recent years. This technology has the potential to greatly enhance and improve various aspects of the industry, including exploration and drilling, equipment maintenance and repair, reservoir performance prediction, drilling and completion optimization, and environmental monitoring and compliance. By analyzing large amounts of data and identifying patterns and trends, deep learning algorithms can help companies make more informed decisions and operate more efficiently, leading to increased profitability and sustainability.

1.6.2   Types of DNNs

Deep neural networks are a subset of machine learning algorithms that use multiple layers of artificial neurons to extract features and patterns from data. There are various types of deep neural networks, each suited for specific tasks. Two popular types are Convolutional Neural Networks and Recurrent Neural Networks.

1D CNNs are often used for processing time series data, such as sensor readings or stock prices. They typically use one-dimensional convolutional filters to extract features from the input data, followed by pooling layers that reduce the spatial dimensionality of the output. 2D CNNs, on the other hand, are primarily used for image and video processing tasks. They use two-dimensional convolutional filters to extract spatial features from the input images, followed by pooling layers that reduce the spatial resolution of the output.

In addition to CNNs, another popular type of deep neural network are RNNs, that are designed to handle sequential data that exhibit temporal dependencies. They are commonly used in speech recognition, natural language processing, and time-series analysis applications. RNNs work by propagating information from one time step to the next, allowing them to capture long-term dependencies in the input sequence.

Long Short-Term Memory (LSTM) is a type of RNN that overcomes the vanishing gradient problem by using a gating mechanism to selectively forget or remember information from previous time steps. This makes it well-suited for tasks that require modeling long-term dependencies, such as language translation and speech recognition. Deep learning techniques, such as LSTM and one-dimensional convolutional neural networks, have become increasingly popular in recent years for a variety of applications, including those in the oil and gas industry. These techniques are particularly useful for analyzing sequential data, such as time series data, and can be used to make predictions or decisions based on past and current data.

LSTM-based model

LSTM networks have shown great promise in sequential data analysis, such as moving object detection in videos (Akilan et al., 2020), speech recognition (Graves and Schmidhuber, 2005), and stock market prediction (Chen et al., 2015). LSTM models are designed to remember and process long-term dependencies in data. They are particularly well-suited for analyzing time series data, as they can take into account both past and present data when making predictions. One key application of LSTM models in the oil and gas industry is in the prediction and prevention of equipment failures. The maintenance and repair of oil and gas production equipment is a critical aspect of the industry, as any downtime can result in significant losses. By analyzing data from sensors and other sources, LSTM models can identify patterns and trends that may indicate an impending failure. This can allow for proactive maintenance and repair, reducing the risk of unplanned downtime and increasing overall efficiency.

LSTM models have also been used in the oil and gas industry to predict the production of oil and gas wells. By analyzing data from sources such as well logs and production data, LSTM

models can make more accurate predictions of well production. This can help companies make more informed decisions about their operations, leading to increased efficiency and profitability. In addition to these applications, LSTM models have also been used in the oil and gas industry to optimize the design of drilling and completion operations. By analyzing data from previous operations, LSTM models can identify patterns and trends that can be used to optimize the design of future operations. This can result in reduced costs and increased efficiency, leading to improved profitability.

Due to their recurrent connections and memory mechanism, the LSTM networks can preserve valuable past observations. Through three gating functions: input gate, output gate, and forget gate defined by (1.8) - (1.12), these networks are capable of adding and removing information to the memory. Fig. 1.8 illustrates the interconnection of these gates in a standard LSTM unit.
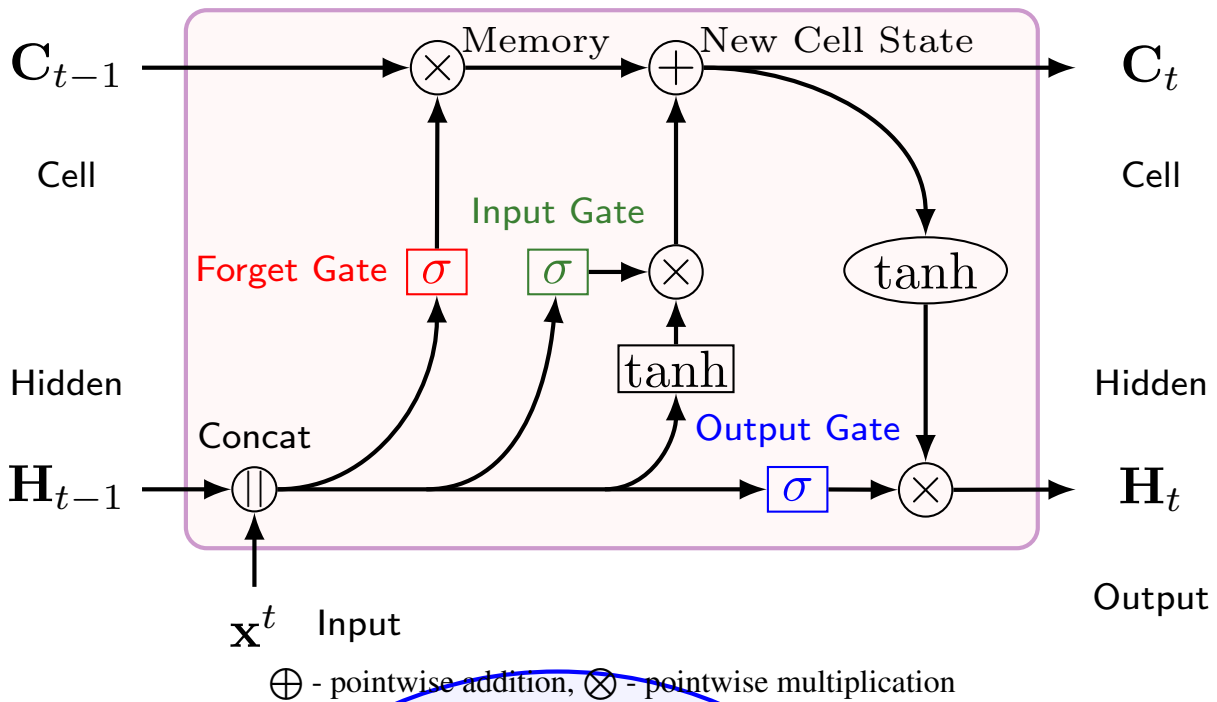
Figure 1.8: An illustration of a standard LSTM cell with three gates that control information flow, where $\mathbf{X}_t$, $\mathbf{C}_t$, and $\mathbf{H}_t$, are the input quantity from a time-series data, cell state, and hidden state, respectively, at timestamp t.
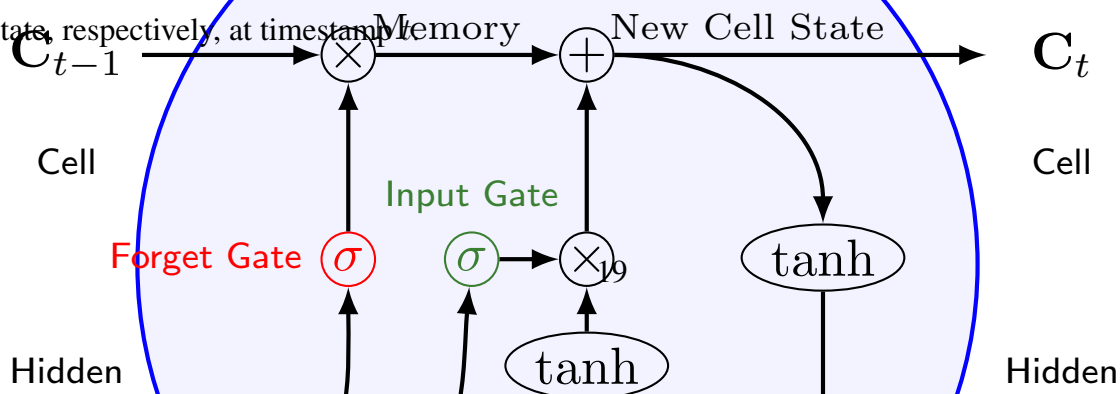
$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + b_i), \tag{1.8}$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + b_f), \tag{1.9}$$

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + b_o), \tag{1.10}$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c), \tag{1.11}$$

$$H_t = o_t \circ \tanh(C_t), \tag{1.12}$$

where X_t is an input quantity from a time-series data, C_t is the cell state, H_t is the hidden state, and i_t, f_t, and o_t are the gates of the LSTM block at timestamp t. Hence, W, '$*$', and '$\circ$' denote conv kernels specific to the gates and internal states, the conv operator, and the Hadamard product. The σ is a `hard sigmoid` function.

1-D CNN-based model

1-D CNNs are another type of deep learning technique that can be used to analyze sequential data. CNNs are widely used in visual recognition-based applications, viz. image classification, object detection, tracking, and segmentation (Akilan et al., 2017, 2020; Yang et al., 2020). These models use convolutional filters to extract features from data, and are particularly suitable for analyzing data with a spatial component, such as audio or video data. In the oil and gas industry, 1-D CNNs could be used to analyze data from sensors or monitoring equipment to identify patterns or trends that may indicate potential environmental issues.

One of the most important applications of 1-D CNNs in the oil and gas industry is in the prediction of reservoir performance. In the oil and gas industry, it is important to accurately predict the yield of a reservoir in order to optimize production. By analyzing data from sources such as well logs and production data, 1-D CNNs can make more accurate predictions of reservoir performance. One of the outstanding characteristics of CNNs is the spatio-local connectivity, through which layers share the parameters, making them efficient learning models. It has been realized that not only do these networks have superior performance, but also have dominant performance in sequential data analytical problems. In CNNs, the convolution (Conv) layer plays a vital role in

feature extraction. The data that pass through the Conv layers convolve with the respective kernels in each layer. The Conv operation (i.e., dot product) between the whole input data and the kernels will generate a volume of feature maps. In this case, as the data is a 1-D sequential data, each Conv layer receives a 1-D input data, $x(n)$. Then, a 1-D kernel $w(n)$, convolve with the input will generate a feature map, $z(n)$ as defined in (1.13) (Zhao et al., 2019).

$$z(n) = x(n) * w(n) = \sum_{m=-l}^{l} x(m) \cdot w(n-m), \tag{1.13}$$

where l stands for the kernel size.

Fig. 1.9 shows the general illustration of the consecutive hidden layers in 1-D CNN models for sequence learning.
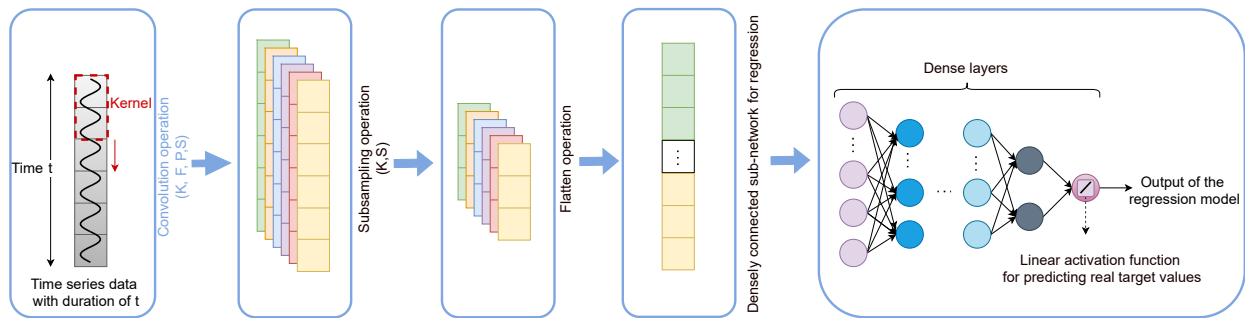
Figure 1.9: General illustration of applying 1-D CNN for sequence learning: In this illustration, a sequence of 6 is considered. K, F, P and S stand for kernel size, number of filters, padding and stride rate, respectively. Then, a sub-sampling operation with a kernel size of 2 and stride rate of 1 has been applied to reduce the dimension of the feature map.

As the introduction has outlined, the use of machine learning and deep learning techniques has the potential to greatly benefit the oil and gas industry. In the following chapters, we will delve deeper into this topic by examining specific case studies of how these techniques have been successfully implemented in three main stages in the oil and gas industry. This will involve a detailed examination of the various methods and approaches that have been used, as well as a discussion of the limitations and challenges that have been encountered in previous research. By doing so, we will be able to establish a clear foundation for our own research and provide a framework for un-

derstanding the gained results. Through this analysis, we aim to provide a comprehensive overview of the capabilities of machine learning and deep learning for solving problems in this sector.

1.7   Application of regression in the oil and gas industry

Regression is a useful tool for facilitating prediction in all of these stages. Regression analysis is one of the most important methods for analyzing multi-factor data, which is used for predicting numerical values (i.e. quantity prediction). The widespread appeal and usefulness of this concept can be attributed to the logical process of representing the relationship between a dependent variable and a set of related independent variables through an equation. Using an equation allows for a clear and concise representation of the relationship between the variables, making it an effective and efficient tool for analysis. This approach is particularly useful in fields such as statistics, economics, and social sciences, where understanding the relationship between different variables is essential for making informed decisions and predictions. Its popularity arises from its conceptually logical approach for using equations to express relationships between input and target variables. In addition, the importance of regression comes from its mathematical and well-developed statistical theory. It is crucial to have an understanding of both the theory and problem domain when the technique is adopted with real-world data. As a result of the wide applicability of regression in an enormous range of problems, regression analysis can be considered the most broadly used statistical technique.

Informed-decision making is an essential element of Industry 4.0 and emerging Industry 5.0 because it enables businesses to make informed decisions about their operations and processes that can improve productivity, efficiency, and profitability. It involves using data and analytics to make logical decisions and help businesses to identify and address potential issues before they become problems. One key benefit of informed-decision making is that it can be used as a preventive measure to prevent machine failures, maintenance, and production-level issues. By analyzing data and identifying patterns and trends, businesses can take proactive steps to prevent problems from occurring, which can help to reduce downtime and minimize the impact on production and

profits. In addition, to prevent problems, informed-decision making can also be used to reduce downtime and environmental impact. By identifying and addressing potential issues before they cause disruptions, businesses can minimize the amount of time that their operations are stopped or slowed, which can help to reduce the environmental impact of their operations.

Finally, informed-decision making can be used to regulate operational conditions and increase work efficiency and profits. By analyzing data and identifying areas for improvement, businesses can optimize their processes and operations to increase efficiency and productivity, which can help to boost profits and improve overall competitiveness. Regression models have reached tremendous progress in recent years. Nowadays, regression models are an inseparable part of diagnostic analyses, model selection, and time series forecasting. Regression models provide a solution for interpreting data that usual analytical approaches cannot describe. As a result, these models have been the centre of global attention for researchers in technical and industry aspects. A selection criterion plays an important role in regression analysis when determining how many model parameters to include and which distribution best describes the data. Regression is a type of machine learning algorithm that is used to predict continuous values, such as prices, temperatures, or stock values. In recent years, deep learning techniques, which involve using neural networks with multiple layers, have become increasingly popular for regression tasks. These deep regression models, which are also known as deep neural networks, are able to learn and model more complex relationships between the input and output data than traditional machine learning algorithms. Deep architectures for regression tasks often include multiple convolutional layers, as well as LSTM cells, which are specialized units that are able to retain information for longer periods of time and are often used to analyze sequential data. These layers are typically followed by one or more fully connected layers, which are responsible for making the final prediction based on the information processed by the previous layers. Overall, deep regression models have shown impressive performance on a wide range of oil and gas-related tasks and have become the state-of-the-art choice for many predictive modelling problems.

In the oil and gas industry, regression models can be utilized to predict a range of outcomes, such as forecasting oil and gas prices, estimating reserves, predicting production, optimizing

drilling and completion operations, and evaluating investments. For instance, regression models can be employed to analyze historical data on oil and gas prices and identify trends and patterns that can be utilized to predict future prices. Additionally, regression models can be used to estimate the size of oil and gas reserves in a particular area by analyzing data on factors such as geologic features, drilling results, and production data. Regression models can also be utilized to predict future oil and gas production from a specific field or well by analyzing data on factors such as reservoir characteristics, well characteristics, and past production data. Furthermore, they can be employed to optimize drilling and completion operations by predicting factors such as drilling time, costs, and success rate based on variables such as well depth, geology, and drilling equipment. Finally, regression models can be utilized to evaluate the potential return on investment for oil and gas projects. By analyzing historical data on production, prices, and other relevant factors, regression models can help companies forecast future production and make informed decisions about investment and resource allocation. This is particularly important in the oil and gas industry, where production forecasting is critical for managing supply and demand, setting prices, and optimizing operations. The accuracy and reliability of these forecasts can have a significant impact on the financial performance and success of a company, and thus, the use of regression analysis in production forecasting is an important consideration in the industry. This thesis particularly focuses on applying advanced deep regression models for optimizing drilling, transportation and production stages in oil and gas industry.

1.8 Thesis contribution

The main contribution of this research is the development of advanced regression models for key operations in the oil and gas industries. The proposed models can be easily integrated into the existing oil and gas industry infrastructure, making them highly scalable and practical. The research contributes to the emerging field of applying machine learning and deep learning techniques for solving oil and gas-related problems, paving the way for more efficient and sustainable oil and gas operations. this study's primary contributions are as follows:

- This thesis offers a thorough examination of data pre-processing techniques, including how to handle missing values, scale data, and choose features based on knowledge from the oil and gas sector.

- Optimize model performance by investigating the most effective hyper-parameters.

- The study's models are designed to be applicable across different wells in the Volve oil field.

- The proposed models' efficacy is extensively tested through an ablation study and comparative analysis.

- Improving efficiency and reducing error ratio in the three mentioned problems. It is to be noted that, proposed models have significant improvements compared to existing works.

Chapter 2

Related Works

The related work section of this thesis provides an overview of existing research and literature relevant to the study. This section also highlights the existing knowledge gap and explains how former works contribute to the field and explain how their study fill those gaps. This chapter will focus on three key areas of the oil and gas industry: drilling, transportation, and production forecasting. Related work section aims to provide a comprehensive understanding of the current state of research in the field and set the foundation for the study's contributions.

2.1   Void fraction measurement

Accurate void fraction measurement is crucial in the oil and gas industry for determining volumetric percentages of hydrocarbon extracts. To enhance the precision of this task, numerous research studies have been conducted in recent years, utilizing machine learning and deep learning techniques. In this case Nazemi *et al.* improved the accuracy of estimation by applying two features of registered signals in Radial Basis Function (RBF) neural network for determining void fraction. By using the proposed method, the percentages of volume fraction were determined autonomous of density alterations in the liquid phase of the stratified regime (Nazemi et al., 2015). Utilizing fewer detectors in structure is a matter of key importance in industries, not only does it lessen expenditures, but also makes it easier to work with these systems. Roshani *et al.* analyzed a simple setup

with a single NaI detector as well as a Co60 source, but authors found that it was impossible to classify all the flow regimes using one detector in the structure and only 2 of the regimes were identified (Roshani et al., 2016). Different features in the frequency domain were presented by Hanus and co-workers in order to identify the flow regimes in dynamic condition (Hanus et al., 2016). In this work, three various structures of two-phase flows (air-water), including plug flow, bubble flow and transitional plug-bubble flow, were studied. Salgado *et al.* have done several works which aim to distinguish flow regimes and determine void fractions using ANNs (de Freitas Dam et al., 2021). Sattari *et al.* carried out research work by taking advantage of time-domain feature extraction for regime classification and void fraction prediction. By adopting time-domain techniques, volumetric percentages were estimated with RMSE of 5.32 (Sattari et al., 2021) Veisy *et al.* discussed the challenge of accurately determining phase fractions in two-phase flows in power plants, petroleum, and petrochemical industries. The authors validate finite element simulations for water-air two-phase flow in an annular pattern using experimental investigations with a concave electrode shape. They also design an Artificial Neural Network (ANN) model to predict void fractions in two-phase flows consisting of petroleum products as the liquid phase, achieving a prediction error of less than 2%. The presented metering system can precisely measure the void fraction of any annular two-phase flow with different liquids (Veisi et al., 2023).

Mayet *et al.* presented a novel technique for controlling liquid petrochemical and petroleum products in a transmitting pipe using a feature extraction system and neural networks. A simulation setup was conducted using the MCNPX code to examine a mixture of four different petroleum products in various volumetric ratios. Twelve time characteristics were extracted from the received signal, and the most effective ones were selected for neural network training. Three MLP neural networks were applied to indicate the volume ratio of three kinds of petroleum products, and the volume ratio of the fourth product can be feasibly achieved through the results of the three aforementioned networks. The proposed system achieved high accuracy with an RMSE of 1.21, which is attributed to the use of appropriate characteristics as the neural network input. The study's innovation lies in increasing the accuracy of predicting the volume ratio, making it an efficient method for the oil industry (Mayet et al., 2022a).

Mayet *et al.* proposed an artificial intelligence method to determine the flow regime and volume percentage of a two-phase flow in the presence of scale inside oil and gas pipelines. The non-invasive method involves using barium-133 and cesium-137 isotopes to irradiate the pipe and detecting the photons that pass through with a single detector. The Monte Carlo N Particle Code is used to simulate frequency features from the recorded data, which are then used as inputs for radial basis function neural networks. These networks accurately determine the volume percentage and flow pattern classification, independent of scale thickness. The proposed system has a higher precision and simpler structure compared to conventional systems that use two detectors. Overall, the paper offers a promising solution to address the issue of scale formation and its impact on pipeline efficiency and equipment depreciation (Mayet et al., 2022b).

Alanazi *et al.* proposed an innovative and non-invasive system to determine the flow pattern and volume percentage in a two-phase flow, taking into consideration the thickness of the scale in the tested pipeline. The system included a dual-energy gamma source, two sodium iodide detectors, and artificial intelligence techniques. The dual-energy gamma source emitted photons, and two detectors recorded transmitted and scattered photons, respectively. Time characteristics were extracted from the recorded data and used as inputs for the multi-layer perceptron neural network. The system was able to accurately determine volume percentages and classify all flow regimes, which can be a valuable tool for the oil and gas industry in detecting and managing scales in pipelines (Alanazi et al., 2022).

Alamoudi *et al.* conducted a novel approach to detect scale deposits in petroleum pipelines using a dual-energy gamma attenuation technique with a radial basis function neural network. The study aimed to determine scale thickness in petroleum pipelines where two-phase flows with different symmetrical flow regimes and void fractions exist. The system consisted of a dual-energy gamma source with Ba-133 and Cs-137 radioisotopes, and two sodium iodide detectors to record photons. The first detector recorded transmitted photons, and the second recorded scattered photons. The authors used radial basis function neural network to process the scattering detector's total counts as inputs and assigned the scale thickness as the output. The proposed approach is

non-invasive and effective and can help improve equipment efficiency in the oil and petrochemical industry (Alamoudi et al., 2021).

Amiri *et al.* described a system for determining the type of flow regime and void fraction percentage of a two-phase flow using an X-ray tube and a sodium iodide crystal detector. The proposed system utilizes the MCNP-X code for physical modeling and implements radial basis function (RBF) for analyzing and classifying the obtained data. The system can recognize all flow regimes and predict the void fraction percentage of a modeled liquid-gas two-phase flow with an acceptable error. The proposed methodology has three main novelties over former studies: the use of an X-ray tube instead of radioisotope sources as the radiation source, the utilization of only one detector instead of at least two, and the use of just one neural network instead of more than one. The findings of this study provide a significant contribution to the research field of two-phase flows and have the potential to be applied in various industrial applications (Amiri et al., 2021).

Basahel *et al.* presented a study on enhancing the accuracy of X-ray radiation-based two-phase flowmeters in the oil industry using artificial neural networks (ANN) and feature extractions. The proposed detection system includes an X-ray tube, a NaI detector, and a Pyrex-glass pipe for modeling the system's geometry using the Monte Carlo MCNP-X code. Five features in the time domain were extracted from the collected data, which were then used as input to a Multi-Layer Perceptron (MLP) to approximate the function related to the input-output relationship. The introduced approach was successful in accurately identifying flow patterns and predicting the volume fraction of two-phase flow components, achieving a high level of precision with root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) of less than 0.51, 0.4 and 1.16%, respectively. The findings of this study demonstrate an improvement in precision compared to previous research and highlight the potential of using ANN and feature extractions in enhancing flow measurement accuracy in the oil industry (Basahel et al., 2021).

In recent years, many researchers put a great deal of effort into oil and gas field for flow regime identification and void fraction measurement by utilizing different methods such as GMDH and wavelet feature extraction (Roshani et al., 2021a,b).

2.2 Electric motor temperature prediction

Kirchgassner *et al.* (Kirchgässner et al., 2021) studied multiple machine learning models for accurate estimation of motors' thermal stress. For this, they analyzed the applicability of different approaches, such as support vector machine (SVM), K-nearest neighbours (KNN), randomized trees, and neural networks (NN). They argued that classical supervised learning models can predict temperature profiles with acceptable accuracy even in high dynamic drive cycles. At the same time, some researchers attempted to exploit the capability of DNNs in handling sequential data, in this case, the temperature of PMSM's internal elements. For instance, Kasburg *et al.* (Kasburg and Stefenon, 2019) utilized LSTM neural network for photovoltaic power generation estimation that achieved a significant improvement over former studies. Gui *et al.* (Gui et al., 2019) implemented a DNN-based model, for multi-step feature selection, and a genetic algorithm (GA)-based solution, to predict the temperature of a reheater system. Similarly, Wallscheid *et al.* (Wallscheid et al., 2017a) investigated the application of recurrent neural networks (RNNs) to accomplish precise temperature prediction of PMSMs. They used particle swarm optimization for finding suitable hyper-parameters, for example, the number of hidden layers and neurons of their model.

Kirchgässner *et al.* (Kirchgässner et al., 2020) employed temporal convolutional network (TCN) for electric motors' temperature prediction. Their model achieves a mean average MSE (MA-MSE) of 3.04 $°C^2$. Similar to the canonical RNN, the TCN is a causal system such that response at timestamp t is the result of 1D-kernels convolved only with input quantities from the time stamp t and before in the earlier time stamps. Thus, the TCN can be viewed as 1-D fully convolutional networks (FCN) with causal convolutions. It is found that the TCN outperforms generic RNNs in certain applications, like polyphonic music modelling, character/word-level language modelling, and IoT anomaly detection (Wang et al., 2021b). On the other hand, Lee *et al.* (Lee and Ha, 2020) proposed a feed-forward neural network (FNN) based on nonlinear auto-regressive exogenous (NARX) to estimate the temperature of the permanent magnet and the stator winding in PMSMs. The authors indicated that their NARX structure would produce better results than the RNN and LSTM-based models if the relationship between the input and the output is known.

Kirchgässner *et al.* evaluate the effectiveness of machine learning models in predicting the temperature of the magnets in permanent magnet synchronous motors (PMSMs) used in automotive applications. Due to the difficulty of signal injection or sensor-based methods, temperature monitoring has been a long-standing challenge, and inaccurate estimations can lead to reduced device utilization and higher material cost. The paper compares the accuracy of several ML models, including ordinary least squares, support vector regression, k-nearest neighbors, randomized trees, and neural networks, with classical thermal models built on thermodynamic theory. The results show that simpler ML models can achieve similar estimation performance as deep neural networks, but with lower model sizes and data demand. Linear regression and simple feed-forward neural networks with optimized hyper parameters are identified as strong predictors of magnet temperature. This study highlights the potential of ML models in accurately predicting magnet temperature and could potentially improve the design and control strategies of PMSMs for automotive applications (Kirchgässner et al., 2021).

Schenke *et al.* developed a model-free deep reinforcement learning torque controller for electric drives, which is a challenging task due to high dynamics, input and state constraints, and secondary objectives such as power efficiency. The paper utilizes the deep Q-learning algorithm with a finite action set, which fits well with the considered system, a permanent magnet synchronous motor supplied by a two-level voltage source inverter. The proposed algorithm incorporates a reward function that considers torque tracking, efficiency maximization, and compliance with operation limits, and a comprehensive hyper parameter optimization is performed to achieve an optimal controller configuration. The paper presents advantages and remaining challenges of the proposed algorithm through an extensive validation, including a direct comparison with a state-of-the-art model predictive direct torque controller. Overall, the paper provides a promising solution for torque control of electric drives without requiring thorough knowledge about the system model (Schenke and Wallscheid, 2021).

Buettner *et al.* discussed a novel approach for Optimal Feedforward Torque Control (OFTC) strategy for Interior Permanent Magnet Synchronous Machines (IPMSMs) by utilizing Artificial Neural Networks (ANN). Unlike conventional OFTC strategies which use large look-up tables or

iterative numerical or analytical computations to find the optimal reference currents, the proposed ANN-based OFTC strategy can analytically compute the optimal reference currents by training and validating the ANN. The proposed approach is faster and easier to implement, while still accounting for nonlinearities and nonidealities such as magnetic cross-coupling and saturation and speed-dependent iron losses, resulting in more accurate optimal reference currents. Simulation results for a real and highly nonlinear IPMSM demonstrate the benefits of the proposed ANN-based OFTC approach compared to conventional OFTC strategies (Buettner et al., 2022).

Kirchgässner *et al.* aimed to improve the accuracy and efficiency of temperature monitoring in permanent magnet synchronous machines (PMSMs), which are widely used in traction drive applications due to their high energy and power density and moderate assembly costs. The lack of accurate temperature monitoring capabilities can harm the thermal robustness of electric motors, leading to oversized materials and reduced effective utilization. To address this issue, the paper proposes using linear regression as an alternative to lumped-parameter thermal networks (LPTNs) for thermal modeling. The paper shows that linear regression can achieve similar predictive performance with low computational complexity, as long as input representations are preprocessed with exponentially weighted moving averages. The proposed method reduces the dependence on domain knowledge and relies entirely on collected data and considered input representations. The paper also examines the impact of data quantity and diversification to determine the minimal amount of test bench measurements required for accurate temperature monitoring. Overall, this work provides a valuable contribution to the field of thermal monitoring in PMSMs and offers an alternative approach that could improve the effective utilization of these machines (Kirchgässner et al., 2019).

Jung *et al.* proposed a method to estimate the temperature of an NdFeB magnet in an interior permanent magnet synchronous motor (IPMSM) operating at medium and high speeds. The proposed method uses the fundamental reactive energy to estimate the magnet temperature, resulting in higher sensitivity compared to conventional methods. The method also eliminates the stator resistance error effect caused by stator temperature without the use of any additional temperature sensor on stator winding. Additionally, the proposed method is robust to inverter nonlinearity ef-

fect and ac resistance effect. The article demonstrates that the proposed method can accurately estimate the magnet temperature in real-time, even in load variation and stator temperature variation. The error in magnet temperature estimation was less than 3.7 °C in online estimation for 10,000 seconds under wide speed and torque variations (Jung et al., 2021).

Although there are some works that focus on the application of DNNs for PMSMs' internal components' temperature prediction, there is still a huge scope for further investigation and development of such models to be entrusted by the industrial community. Thus, in this work, two different DNNs have been introduced aiming to enhance the precision of temperature prediction of the permanent magnet, stator tooth, stator winding, and stator yoke of PMSMs, simultaneously.

2.3   Oil production forecasting

In the modern era, industries want to explore more pathways for cost saving, increasing productivity, and enhancing safety in the operational environment. Thus, engineers and scientists develop various IATs to facilitate industries in achieving their desired goals. For example, in recent years, for solving several problems related to the oil and gas industries, researchers have explored a combination of the nature-inspired meta-heuristic algorithm (MA) and ML techniques. These algorithms are found to have robust performances and converge to the global optimum solution (Bahiraei et al., 2021a,b; Ezugwu et al., 2020).

On the other hand, Alakeely *et al.* implemented a recurrent neural network model along with convolutional neural networks to simulate reservoir behavior (Alakeely and Horne, 2020). In addition to reservoir engineering, some research works focused on applying ML for drilling and construction engineering in the petrochemical industry. For instance, Syed *et al.* (Syed et al., 2020) investigated ML models to predict lift selection, assess the wells' performance, and to classify them as "Good" or "Bad" wells based on their life-cycle cost (LCC).

Similarly, Adedigba *et al.* (Adedigba et al., 2018) conducted research touching upon risk assessment of drilling operations using a Bayesian tree augmented Naïve Bayes (TAN). They developed this model to predict time-dependent blowout risk based on the current status of the key

drilling parameters in real-time. Hence, it is intended for informed decision-making to avoid preventable workplace accidents and enhance the safety of drilling operations. Furthermore, Ozbayoglu *et al.* proposed an artificial neural network (ANN)-based model to estimate flow rate and velocity of pipe rotation for real-time drilling optimization and automation (Ozbayoglu et al., 2021).

ML-driven data analytics– a branch of science taking advantage of advanced statistical and neural network techniques– is used to realize and unearth insights and trends in large-scale datasets. It can be potentially exploited to drive meaningful information from hydrocarbon well's raw data aiming at increasing production efficiency and maximizing the profit of petrochemical industries. For instance, Bao *et al.* (Bao et al., 2020) investigated the performance of RNN combined with an ensemble Kalman filter (EnKf) for predicting production to assist reservoir characterization and development. They verified their model on synthetic historical production data rather than real data collected from an oil field.

On the contrary, some researchers attempted to develop production forecasting models using actual data. For example, Zanjani *et al.* (Zanjani et al., 2020) developed multiple algorithms based on ANNs, SVM and linear regression (LR) for production forecasting using well-specific information. Their results on well NO 159 F-1 C in the Volve oil field show that the ANN-based model performs better than the other two algorithms. Since they focus on single well-specific model development, their approach is not scalable.

Wang *et al.* (Wang et al., 2021a) conducted a study using machine learning to predict future production. In this research, a machine learning algorithm called the random forest ensemble was implemented to predict time-lapse oil saturation profiles. The algorithm was optimized using feature selection based on feature importance scores and Pearson correlation coefficients in combination with geophysical domain knowledge. The workflow was demonstrated using data from a structurally complex, heterogeneous, and heavily faulted offshore reservoir and was able to predict future time-lapse oil saturation profiles with high accuracy, as measured by over 90% R-square. This approach is notable because it does not require input parameters derived from cores, petrophysical logs, or seismic data and incorporates production data, which is an essential reflection of

dynamic reservoir properties and is typically the most frequently and reliably measured quantity throughout the life of a field.

Li *et al.* conducted a study for pressure prediction by combining the physics of well's behaviour and DL models. Gated Recurrent Unit (GRU) and LSTM models were implemented to compare their results with the RNN model. Results showed that GRU and LSTM performed better compared to RNN. Well NO 15/9-F-1 C from April 2014 to April 2016 was used as a testing profile (Li et al., 2019).

Masina *et al.* studied automated declined curve analysis (DCA) using AI to predict production rate. Their results showed that the DCA method is able to predict the desired output with a goodness of fit of 0.82 on the test set (Masini et al., 2019).

Zhang *et al.* (Zhang et al., 2018) proposed a method for detecting and locating leaks in liquid pipelines, which combines inverse hydraulic-thermodynamic transient analysis with an improved version of the particle swarm optimization (PSO) algorithm. The finite volume method is used to solve the continuity, momentum, and energy equations numerically. Four different algorithms were tested to determine the best-performing version of the improved PSO algorithm, and the results were evaluated based on accuracy, stability, robustness, and false alarm rate. The SIPSO algorithm was found to be the most effective. The proposed method was applied to two oil pipelines in real-world scenarios, one during a field-opening experiment and the other during a leak incident. The method was able to accurately estimate the location, coefficient, and starting time of the leaks with low relative errors.

Noshi *et al.* explored the potential application of Machine Learning algorithms in production prediction. They took advantage of the AdaBoost technique for production prediction. Mean absolute error (MAE) was used as an error metric to show the method's performance. Six features that affect production prediction, including: on stream hours, average choke size, bore oil volume, bore gas volume, bore water volume, and finally, average wellhead pressure were used as input parameters (Noshi et al., 2019).

Panja *et al.* carried out a study to predict hydrocarbon production from hydraulically fractured wells. Two common types of ML models, namely the Least Square Support Vector Machine

(LSSVM) and the ANN were analyzed and compared to the traditional curve fitting method known as Response Surface Model (RSM) using second-order polynomial equations to determine production rate (Panja et al., 2018).

Wui Ng *et al.* studied the LSTM model for Volve oilfield production forecasting. In the mentioned work, only well NO 15/9-F-14 H were used for both the training and testing process (Ng et al., 2022). This paper used an incorrect strategy in methodology. Specifically, the correlation between the production of oil and gas was found to be equal to 1, but the authors used gas as the input in their network and oil as the output. A more appropriate approach would have been to remove gas from the input variables in order to avoid this issue, as it is highly correlated with the output and will result in a bias in the predictions made by the network.

One of the classical methods which were used for hydrocarbon production forecasting in recent decades is decline curve analysis (DCA). This method was initiated by Arps *et al.* (1945), and then oil and gas companies adopted this method and its specific applications in related industries (Arps, 1945). As a result of its simple development, it has been broadly used in various situations (Hong et al., 2019). For forecasting hydrocarbon production, numerical reservoir simulation (NRS) can be used as an alternative to DCA. Yet, the performance of the NRS method depends on how historical matching (HM) has been done (Liu et al., 2019). Also, NRS requires several features, comprising well locations, fluid properties, geological data, etc. To forecast production more accurately, the simulation model should be updated via HM when new real-time data emerges. As a result, this method has apparent limitations (Ng et al., 2022).

Data-driven modelling has become a viable option for hydrocarbon production forecasting with the advancement of data analytics and computing technology. Not only is this method easy to implement, but it also captures the intricate relationships between inputs and outputs. Utilizing ML has led to notable advancement in the oil and gas industry, especially in reservoir engineering (Ng et al., 2022).

The production of oil from horizontal well steam stimulation in heavy oil reservoirs needs to be accurately predicted. To address the limitations of traditional forecasting methods, a new method called PCA-GBDT has been developed. This method involves combining and processing data from

geological sources and field production, using grey correlation analysis to identify the main factors that influence oil production, applying principal component analysis to eliminate correlations among these factors, building a sample library for machine learning, and training a Gradient Boosting Decision Tree model using this sample library and optimizing the model's parameters through grid search. The PCA-GBDT model was found to be effective in predicting oil production, with an average absolute error of less than 10%. This method is useful for optimizing the production and development of heavy oil horizontal well steam huff and puff and for the application of machine learning in unconventional resources (Liu et al., 2021a).

Ning *et al.* (Ning et al., 2022) presented a machine learning-based time series forecasting method for predicting the production performance of unconventional reservoirs. The traditional methods of production forecasting are subjective and do not account for the complexity of unconventional reservoirs. The study compared three algorithms: ARIMA, LSTM, and Prophet, and evaluated their performance against decline curve analysis and reservoir simulation modeling forecasting. The results show that the machine learning models outperform traditional methods in terms of reliability, speed, and simplicity of the workflow. The Prophet model, in particular, captured seasonal influences on oil production, which is often overlooked by previous studies. The study also demonstrated the transferability of the model parameters across wells in the same area.

Al Qaness *et al.* (Al-Qaness et al., 2022) discussed the use of artificial neural networks in the oil industry for forecasting and predicting crude oil production. Specifically, the authors focus on the dendritic neural regression (DNR) model, which has shown promising results in time-series prediction, but has limitations in training and parameter configuration. To address these limitations, the authors utilize metaheuristic optimization algorithms, including the whale optimization algorithm (WOA), particle swarm optimization algorithm (PSO), genetic algorithm (GA), sine–cosine algorithm (SCA), differential evolution (DE), and harmony search algorithm (HS). The authors evaluated the performance of these algorithms using oil-production datasets from seven real-world oilfields in China. The findings suggest that the application of metaheuristic methods improves the performance of the DNR model, with PSO and WOA achieving the best results. Overall, this

study confirms the applicability of metaheuristic optimization algorithms with DNR in time-series forecasting applications for crude oil production.

However, the hydrocarbon industries face challenges in accomplishing this, as the existing conventional tools are not generalized and robust enough. In recent decades, AI and DL-based innovative solutions have emerged to improve the efficiency of operations in industries.

In recent years, DNNs and ML methods have been practically used in the industry for problem-solving. This research has been carried out, which aims to propose two models based on convolutional neural network and long short-term memory for hydrocarbon production.

Chapter 3

Void Fraction Measurement

3.1   Overview

Two-phase flow is very important in many areas of science, engineering, and industry. Two-phase flow comprising gas and liquid phases is a common occurrence in oil and gas-related industries. This study considers three flow regimes, including homogeneous, annular, and stratified regimes ranging from 5–90% of void fractions simulated via the Mont Carlo N-Particle (MCNP) Code. In the proposed model, two NaI detectors were used for recording the emitted photons of a cesium 137 source that passes through the pipe. Following that, fast Fourier transform (FFT), which aims to transfer recorded signals to the frequency domain, was adopted. By analyzing signals in the frequency domain, it is possible to extract some hidden features that are not visible in the time domain analysis. Four distinctive features of registered signals, including average value, the amplitude of dominant frequency, standard deviation (STD), and skewness were extracted. These features were compared to each other to determine the best feature that can offer the best separation. Furthermore, artificial neural network was utilized to increase the efficiency of two-phase flowmeters. Additionally, two multi-layer perceptrons (MLP) neural networks were adopted for classifying the considered regimes and estimating the volumetric percentages. Applying the proposed model, the outlined flow regimes were accurately classified, resulting in volumetric percentages with a low root mean square error (RMSE) of 1.1%.

3.2 Introduction

One of the most costly processes within the oil and gas industry is the measurement of volumetric percentages of extracted hydrocarbon materials. This type of measurement is a complex task that requires a high level of precision and expertise. To accurately measure the volumetric percentages and flow rates of oil and gas, highly sensitive flow meters are utilized. Maintaining a high level of precision is essential to meet industry quality standards.

Flow meters are devices used to measure the flow rate, or volume, of a fluid as it passes through a pipeline. These devices are commonly used in the oil and gas, nuclear, and petrochemical industries to measure the flow of materials such as oil, gas, and chemicals through pipelines. Accurate measurement of the flow of these materials is important for various reasons, including monitoring the efficiency of the drilling process and ensuring that the right amount of materials is being extracted and transported. Flow meters can be used to measure the flow of fluids in different flow regimes, including homogeneous, annular, and stratified. The precise measurement of the flow of oil and gas is also critical for the economy of a country, as it can impact the growth and development of these industries. Fig. 3.1 is a simple illustration of the mentioned flow regimes.
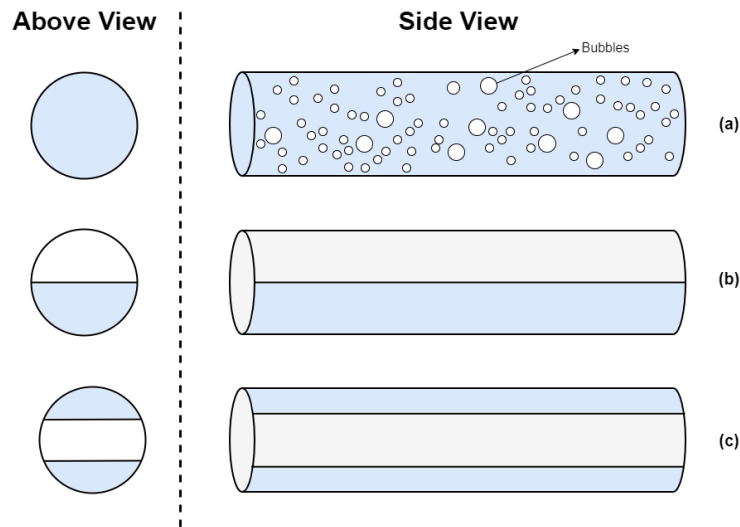
Figure 3.1: Types of flow regimes: a) Homogeneous, b) Stratified, c) Annular.

There are specific definitions for each flow regime as follows:

- **Homogeneous:** In this flow regime, the liquid and gas phases are uniformly distributed throughout the pipeline. This can occur when the gas and liquid flow rates are well balanced, resulting in a homogeneous mixture of the two phases.

- **Annular:** In the annular flow regime, the liquid phase is situated on both sides of the pipe, while the gas phase is in the middle. This can occur when the gas flow rate is much higher than the liquid flow rate, resulting in the gas phase being surrounded by the liquid phase.

- **Stratified:** In the stratified flow regime, the liquid phase is located at the bottom of the pipe, while the gas phase is at the top. This can occur when the liquid flow rate is much higher than the gas flow rate, resulting in the liquid phase being more densely packed at the bottom of the pipe.

These flow regimes can have significant impacts on the design and operation of pipelines in the oil and gas industry, as they can affect the pressure drop, flow rate, and stability of the flow. It is important to accurately predict and control the flow regime in order to optimize the operation of a pipeline. This can be achieved through the use of appropriate flow measurement and control devices , as well as careful selection of pipe diameter and material to ensure that the desired flow regime is maintained.

Nowadays, analyzing flow regimes and volumetric percentages of multiphase flows is a significant and notable topic in many industries (Nazemi et al., 2016b). Liquids and gases are the most important elements of oil and gas storage. For better comprehension, of whether the drilling process is sensible or not, it is essential to measure each parameter (Salgado et al., 2020). Also, the separation procedure will be developed much better with adequate information about the regime types and volume of each phase. In this regard, different kinds of methods have been studied for determining the mentioned parameters. According to former studies, the gamma-ray attenuation technique was the utmost accurate method. Abro and his colleagues have investigated the efficiency of single and multi-beam gamma-ray densitometry to estimate the volumetric percentages in two-phase flows consisting of gas and liquid (Åbro and Johansen, 1999). According to their acquired results, the multi-beam gamma ray method had higher accuracy compared to the single-

beam technique. Chunguo and co-workers investigated dual modality densitometry to classify the flow regimes in a vertical pipe (Chunguo and Qiuguo, 2009). In 2014, three flow regimes (homogeneous, annular, and stratified) were simulated by MCNP code (Roshani et al., 2015). One 137Cs source, and one transmitted and scattered detector were utilized as the proposed structure. For classifying the flow regimes and predicting volumetric percentages, three attributes of signals were extracted and used as the ANN inputs. Faghihi et al. studied stratified, homogeneous, and annular regimes in a pipe with a vertical position for three different flow regimes (Faghihi et al., 2015). Nazemi et al. investigated the gamma-ray attenuation technique in annular, bubbly, and stratified flow regimes in a 2-phase flow structure. In this article, Volume fraction was determined independent of regime type (Nazemi et al., 2016a). Two transmitted detectors were applied to register the transmitted photons. In this situation, void fraction percentages were calculated using the MLP neural network.

3.3 Methodology

3.3.1 Simulation procedure

The data collection process in this work includes two separated stages. The data set comprises 14,892 data samples for each regime (44,676 data samples for the three flow regimes). Firstly, three principal regimes including homogeneous, annular, and stratified, were simulated using MCNP code based on the illustrated structure in Fig. 3.2. Simulations were accomplished for 5%-90% void fraction. Gasoil with the density of 835 kg/m3 as well as the air with the density of $1kg/m^3$ were considered as the liquid and gas phases, respectively. A 137 Cs radioactive source (emitter energy: 0.662 MeV) and two 25.4 mm NaI detectors were utilized in order to register photons that passed through the pipe with an inner diameter of 95 mm and thickness of 2.5 mm. Secondly, for evaluating the accuracy of the simulated structure in MCNP code, simulated geometry, was assessed for validity with multiple experiments in (Nazemi et al., 2016b). Fig 3.3 shows different stages of The proposed void fraction regressor.
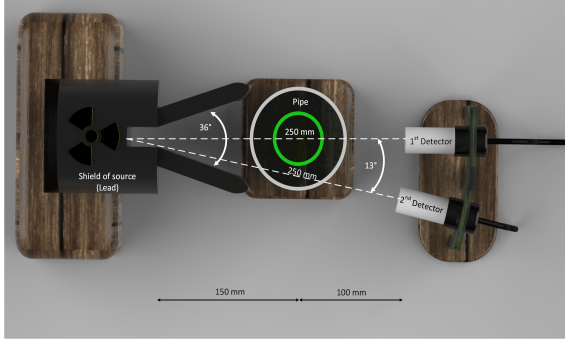
(a) Top View

(b) 3-D View

Figure 3.2: The data gathering apparatuses with MCNP (Hosseini et al., 2022a).

Fig. 3.4 shows registered photon energy spectra for the three flow regimes and void fraction of 5%.

3.3.2 Data conditioning and feature extraction

In this study, fast Fourier transform was used to transform registered signals into frequency domain as defined in Eq. (3.1):

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)e^{-j\omega t}dt, \tag{3.1}$$

where $F(\omega)$ and $f(t)$, stand for signal in the frequency-domain and time-domain. ω and t are angular frequency and time frequency, respectively.

Fig. 3.5 and Fig. 3.6 show a simple illustration of applying FFT on the time domain and an example of an annular signal in the frequency domain, respectively.

As is observed in Fig. 3.7, in all three flow regimes, there is a definite link between the air percentages in the pipe and the amplitude of the dominant frequency. In other words, by increasing the volume of air in the pipe, the amplitude of the dominant frequency gradually increases.

The adopted features are as follows:

- The signal average:

$$m = \frac{1}{N} \sum_{n=1}^{N} x[n] \tag{3.2}$$
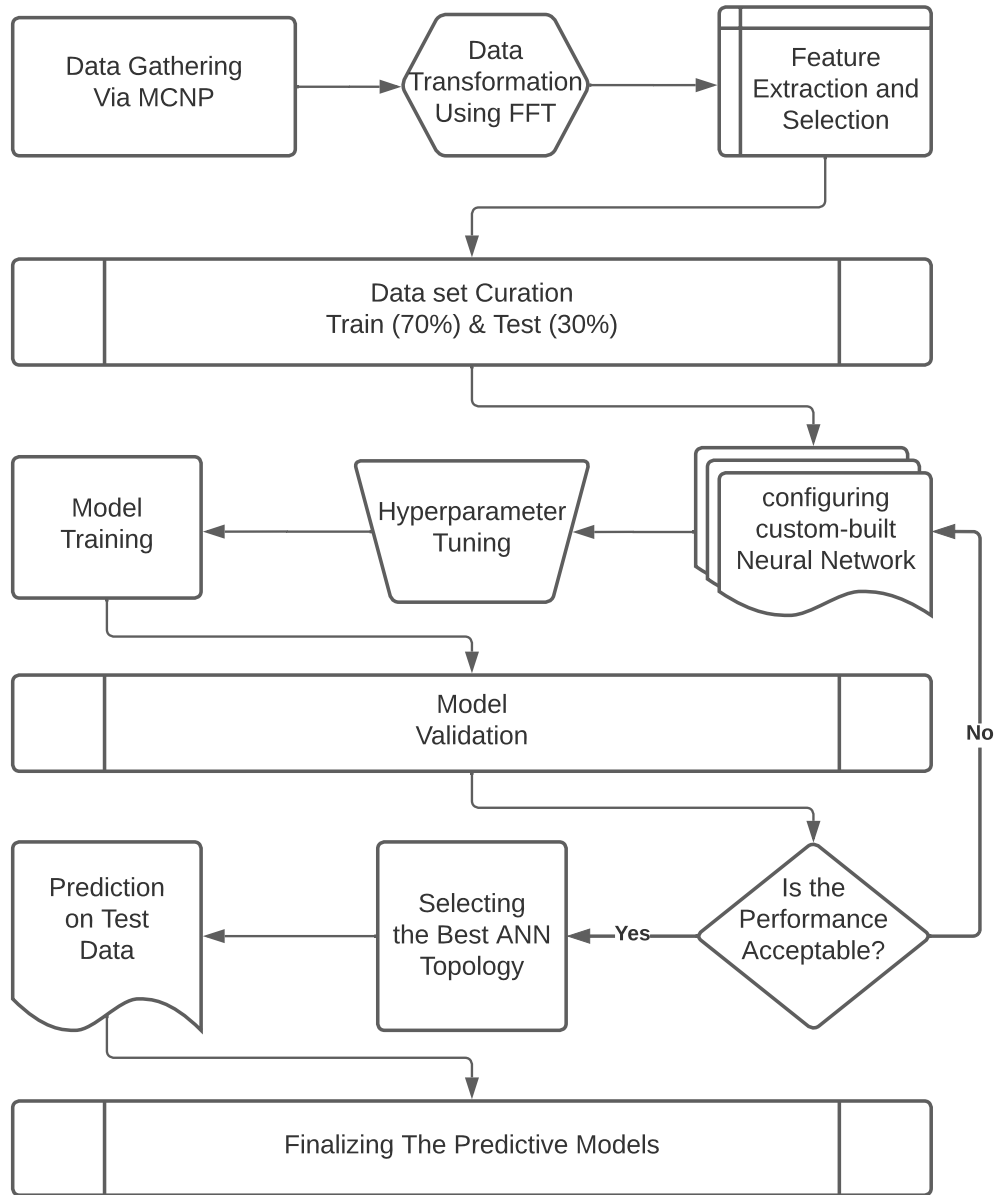
Figure 3.3: Stages of the proposed solution from data gathering to model building, training and evaluation.
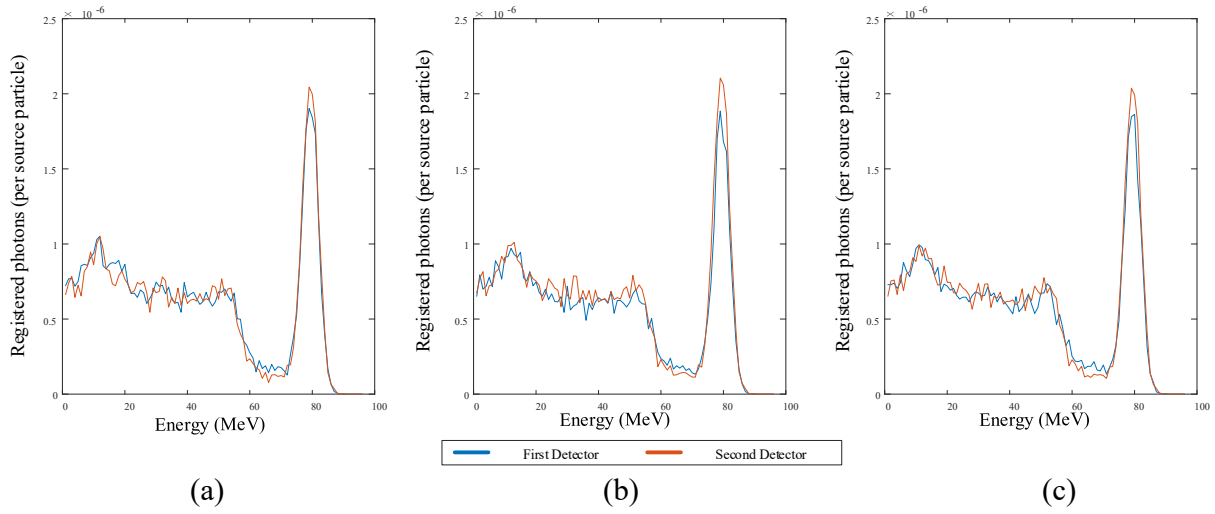
Figure 3.4: Registered photon energy spectra in the 1st & 2nd detectors (void frac.=5%): a) Annular, b) Homogeneous, and c) Stratified (Hosseini et al., 2022a).
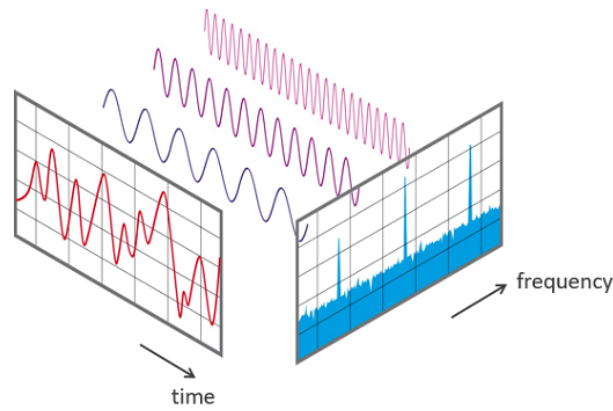
Figure 3.5: A simple illustration of the fast fourier transform.

- The amplitude of the dominant frequency

- Standard deviation (STD):

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} (x[n] - m)^2}, \tag{3.3}$$

- Skewness:

$$s = \frac{m_3}{\sigma^3}, m_3 = \frac{1}{N} \sum_{n=1}^{n=N} (x[n] - m)^3, \tag{3.4}$$

45
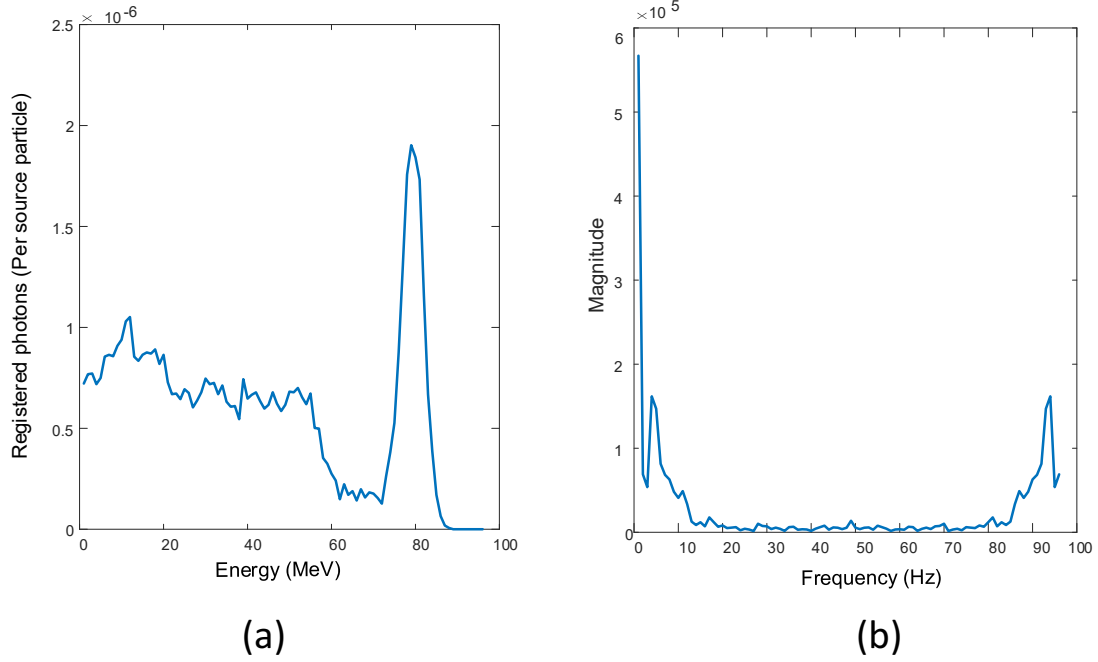
Figure 3.6: An example for FFT of the annular regime with void fraction = 5%): Output signal of the 1st detector in (a) Time-domain and (b) Frequency domain (Hosseini et al., 2022a).
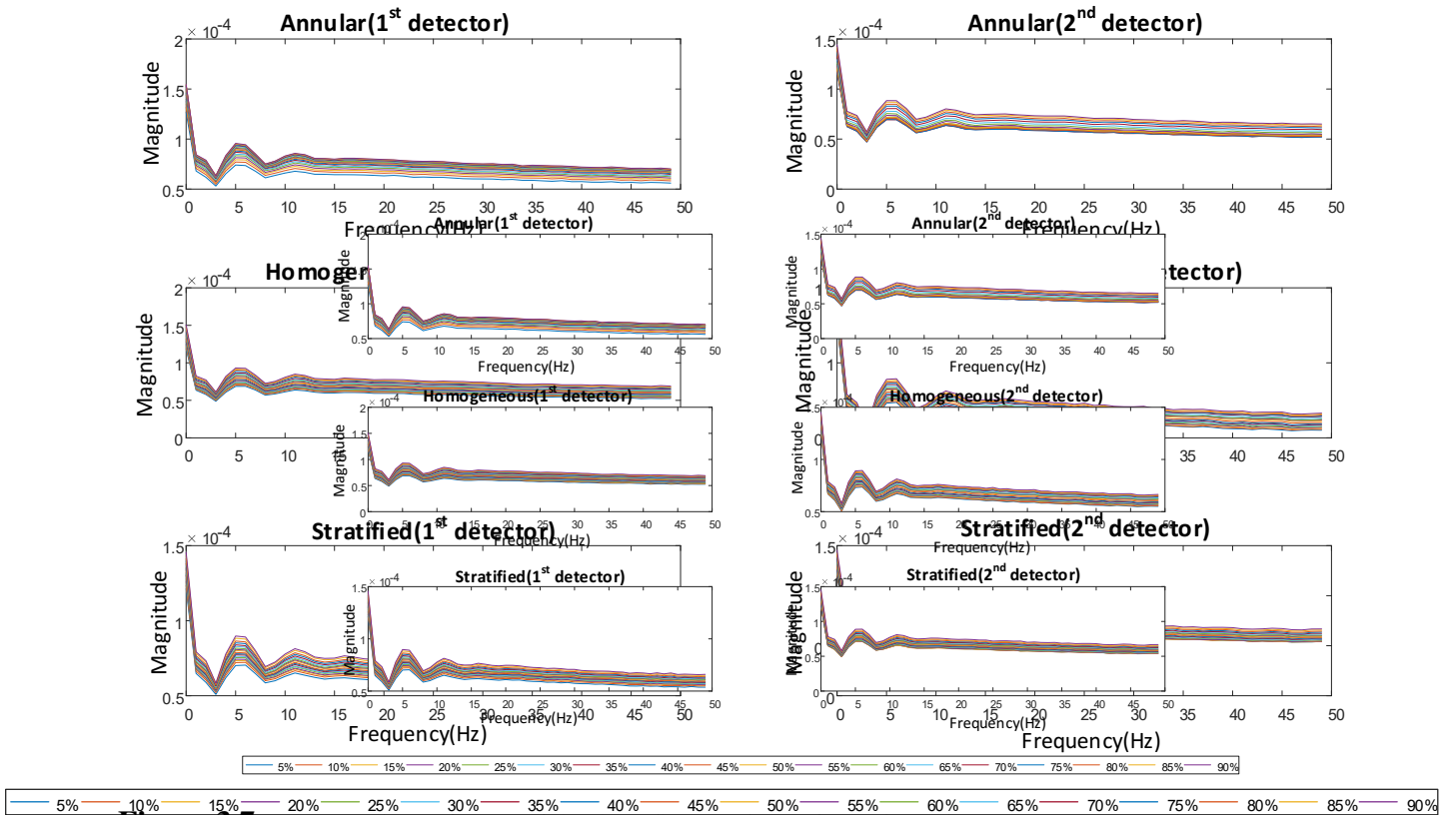
Figure 3.7: A comparison of FFTs of the three flow Regimes' for various air percentages (Hosseini et al., 2022a).

Where x and N stand for the corresponding registered data and total number of data samples, respectively.

Fig. 3.8 shows the diagram of extracted features (the first detector versus the second detector), which indicates the separation ability for every feature. As is clear, the classification procedure of flow regimes is possible only with one feature (standard deviation), and the other three extracted features are not able to classify the mentioned flow regimes because of overlap in their diagrams. According to these obtained results, it can be concluded that the standard deviation is the best feature. Also, the indicated points in each graph show the different void fraction percentages.

3.3.3 Model building

ANNs have a wide range of applications in several engineering studies. There are different advantages, such as flexibility, availability and simplicity, which lead to great attention by researchers for analyzing various processes in many fields and systems. ANNs can be expressed as mathematical systems which consist of several neurons performing in a parallel way, and it is produced in single or several layers (Hosseini et al., 2020). These networks have been inspired by biological neural networks (Bahiraei et al., 2020). MLP is a common type of neural network (Bahiraei et al., 2021a). ANN is a suitable technique for handling the models and classification as well as prediction (Bahiraei et al., 2021b). ANNs are widely adopted in many different fields because of their high potential applications. ANNs are almost like human brain data processing technology. The neurons are a basic components of biological networks, and other elements are synapses, soma or cell body and dendrites. Input information or signals come to the dendrites, and soma acts like a processor, synaptic perform like a link, and the axon is responsible for transmitting output signals to the other neurons and also with non-linear performance [31]. ANN model comprises several processing components, called neurons, which act parallel. Inputs are transmitted to each neuron in proportion to their allocated weights, and a non-linear activation function affects and generates specific data which may transmit to others.

The parameters of the proposed neural network-based solution for predicting numerical values of a void fraction are as follows:
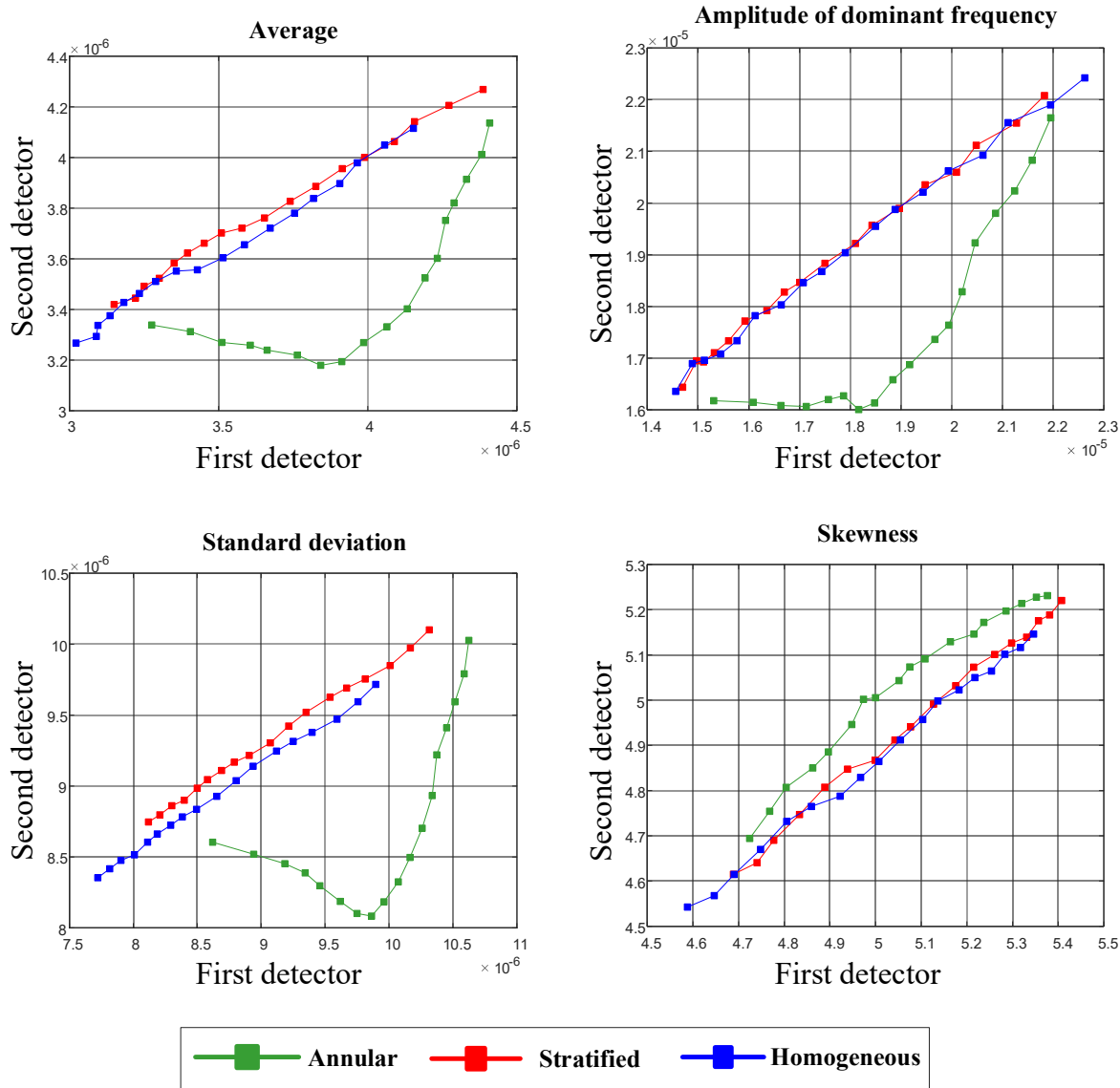
Figure 3.8: Extracted features from both detectors (Hosseini et al., 2022a).

- Input layer: 2 neurons

- 1st hidden layer: 3 neurons

- 2nd hidden layer: 3 neurons

- Output layer: 1 neuron

- Number of epochs: 250

Fig. 3.9 shows the architecture employed for void fraction measurement.
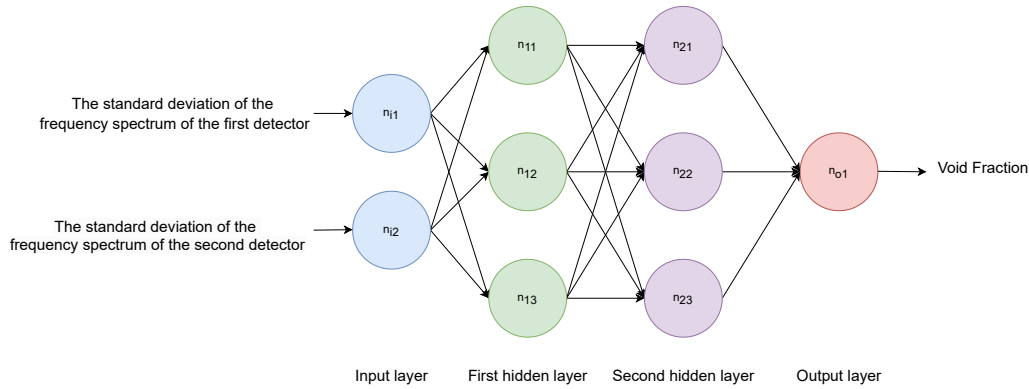
Figure 3.9: An illustration of the proposed neural network-based void fraction regressor.

The network performance for void fraction measurement for training and testing data samples is indicated in Figs. 3.10 and 3.11, respectively.
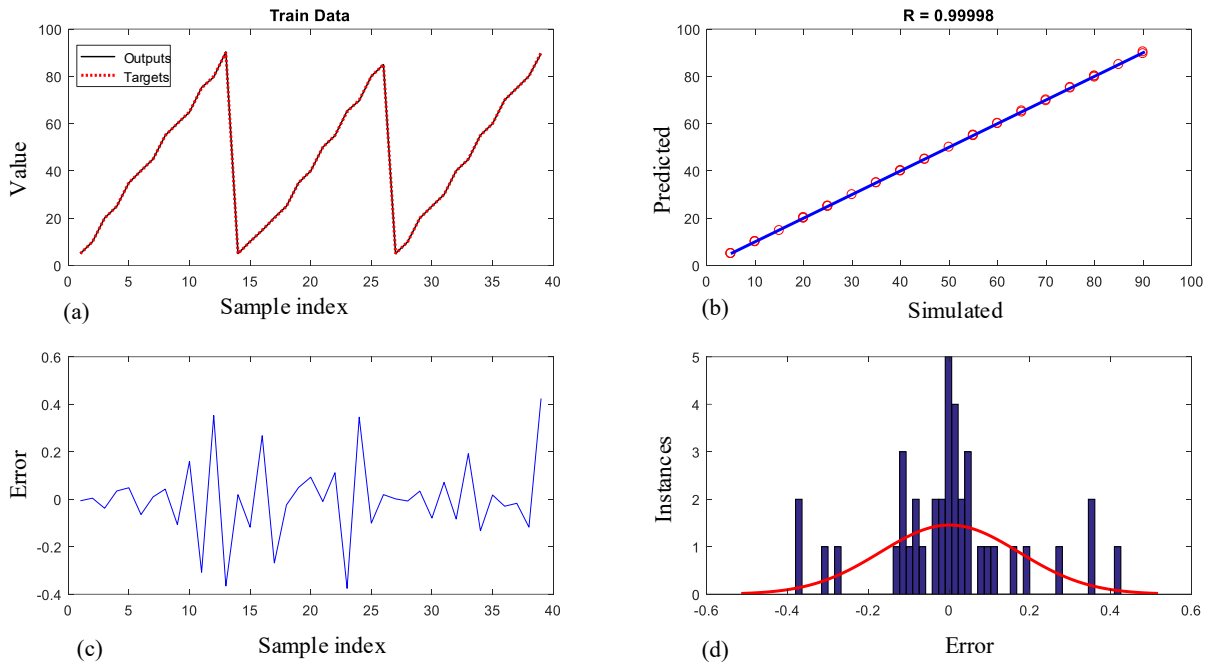
Figure 3.10: Training performance of the proposed neural network-based regressor: a) Fitting, b) Regression, c) Error, d) Error histogram diagram Hosseini et al. (2022a).

To evaluate the Adopted ANN, the root mean square error percentage (RMSE %) and coefficient of determination (R-squared) were computed by Eqs. (3.5) and (3.6), respectively. It is highly
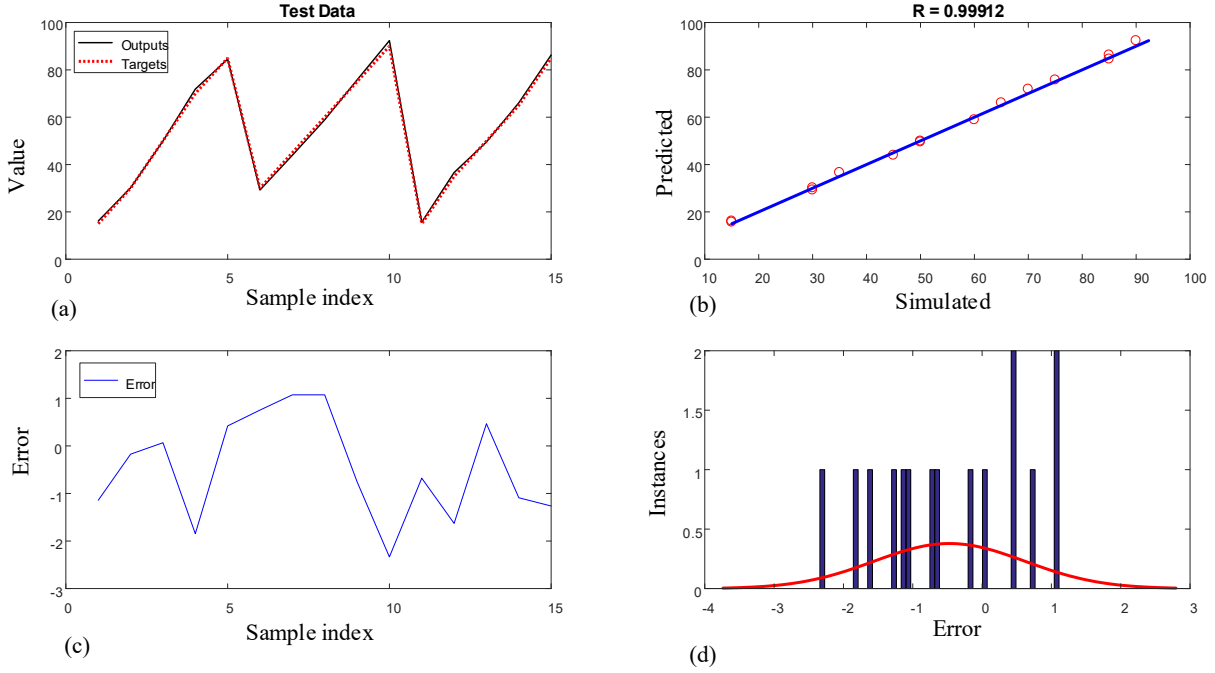
Figure 3.11: Performances of the proposed model in test set: a) Fitting, b) Regression, c) Error, d) Error Histogram Hosseini et al. (2022a).

noted that the coefficient of determination (R-squared) is applied as the goodness of fit index is described as the proportion of variation explained by the best line model. The achieved errors are indicated in Table 3.1.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{3.5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}, \tag{3.6}$$

where y_i, $\hat{y}_i$ and $\bar{y}$ stand for actual, predicted and average values of target attributes, respectively.

Table 3.1: Comparison between training and testing performances.

Data	RMSE $\%$	R^2
Training	0.86	0.99
Testing	1.15	0.99

The comparison between this study and several research items in this field can be found in Table 3.2.

Table 3.2: Comparison between this work and previous studies based on RMSE for volumetric percentages (The best performance is inked in Blue)

Ref.	Technique	RMSE	Improvement
Nazemi et al. (2019)	Total count	2.12	-10%
Sattari et al. (2021)	Time-domain	5.32	-177%
Hosseini et al. (2021b)	Wavelet feature extraction	1.92	Baseline
Current research Hosseini et al. (2022a)	Frequency-domain	1.15	+40%

3.4 Conclusion

This chapter proposed the use of fast Fourier transform to transform and analyze the frequency domain signals of three flow regimes simulated using MCNP code. Three attributes were extracted in the frequency domain, and the standard deviation was recognized as the best feature for void fraction measurement.Moreover, by using the feature extraction technique and applying neural networks, void fraction percentages were predicted with a low root mean square error of 1.1%, which is indicative of the utility of the proposed model. This thesis did not investigate the impact of different regression models on prediction precision. However, this aspect remains a subject for future research.

Chapter 4

Electric Motor Temperature Prediction

4.1 Overview

Electric motors are becoming widely used in many different applications, such as electric cars and turbines. Measuring the temperature of internal components of an electric motor, like a permanent magnet synchronous motor (PMSM) is vital to maintain its safe operation. However, measuring the temperature of the permanent magnet and stator directly comes at the expense of higher cost and additional hardware requirements, for instance, a sensor network. To overcome these limitations, machine learning (ML) techniques can be employed to model the mentioned parameters without the need for specialized sensors and design ideas for housing them inside the motors. Classical methods, like lumped-parameter thermal networks (LPTNs), are capable of calculating the temperature of internal elements of PMSMs. But, these methods require expertise and may lack acceptable accuracy. In this study, two deep neural networks were modelled using CNN and long LSTM units to predict the temperature of four target values of PMSMs: stator tooth, stator yoke, stator winding, and permanent magnet. For attribute conditioning, exponentially weighted moving average (EWMA) and exponentially weighted moving standard deviation (EWMS) were applied. A thorough ablation analysis shows that the CNN-based model predicts the targets better than the LSTM model with an average mean squared error (MSE) of 2.64 °C^2 and an average R^2

of 0.9924. It is also found that the proposed CNN-based model achieves a 13% mean average performance (mAP) improvement compared to the existing state-of-the-art solution.

4.2   Introduction

The oil and gas industry is a vital part of the global economy, providing energy for a wide range of purposes including transportation, manufacturing, and electricity generation. One of the key components of the oil and gas industry is drilling rig equipment, which is used to extract resources. These rigs rely on electric motors as their power source, and these motors must be able to withstand harsh conditions including vibration, extreme temperatures, and corrosive environments. The electric motors used in the oil and gas industry play a crucial role in the drilling process as well as transportation, and it is important to have a proper preventative maintenance plan in place to ensure optimal performance, safety, and efficiency.

The power generated by these motors is used to support various stages of the oil and gas industry, including extraction, processing, storage, and transportation. For example, motors may be used to operate pumps and compressors in order to extract and transport oil and natural gas from the ground. They may also be used to power equipment used in the refining process, such as distillation towers and heat exchangers. One common issue that can arise with electric motors in the oil and gas industry is an unexpected increase in temperature. This can lead to problems with performance and efficiency, and can also pose a safety hazard if not addressed promptly. There are several potential causes of overheating in electric motors, including faulty wiring, overloading, and a lack of proper lubrication. It is important to regularly monitor the temperature of electric motors and take steps to address any issues that may arise.

In order to guarantee the lifespan and proper functionality of the PMSMs, there is a focus in the research community to measure the temperature of PMSM's internal components (cf. Fig. 4.1) using mathematical modelling to modern data-driven approaches. For instance, Wallscheid *et al.* (Wallscheid et al., 2017b) utilized precise flux observer to measure the magnet's temperature indirectly in the fundamental wave domain without any signal injectors and sensors. However, the
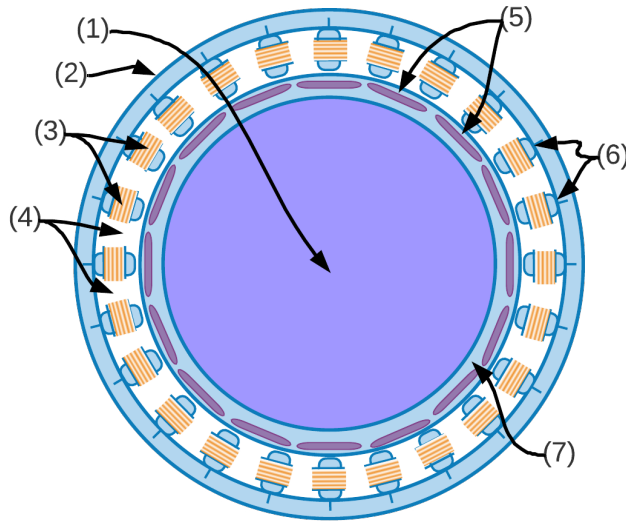
Figure 4.1: PMSM's cross-section: (1) rotor shaft, (2) stator yoke, (3) stator winding, (4) stator slot, (5) permanent magnet, (6) stator tooth, (7) rotor ion.

methods that have been used in this case comes with huge cost and preparation of apparatus. This is why researchers have started applying data-driven models for electric motors' temperature estimation after gathering enough data. In this line, the traditional methods deal with finite element methods for temperature prediction (Grobler et al., 2015). It is mainly derived by simulating transient temperature by utilizing finite elements, establishing the temperature field, and finally estimating the motor temperature. This technique is capable to measure the current linear data and it is not applicable to large-scale and nonlinear historical data. On the other hand, ML techniques can efficiently solve the aforementioned problem and improve the prediction process in comparison with the former methods. For instance, Hyeontae *et al.* (Jo et al., 2019) used an ML-based approach, particularly, the decision tree, for temperature estimation of Linz–Donawiz converter for improving the quality of melted pig iron. Similarly, Wang (Wang, 2016) built a random forest model to predict the temperature of a ladle furnace. Su *et al.* (Su et al., 2019) proposed an extreme learning machine-based implementation to estimate the iron temperature in a blast furnace. Hence, Zhukov *et al.* (Zhukov et al., 2019) studied different methods based on the random forest for assessing the security level in power systems. The feature extraction, in most of the ML techniques, rely on human involvement (Akilan et al., 2017; Hosseini et al., 2020, 2021b). For elementary tasks, this way of feature extraction is efficient and simple but for a complex problem,

manual feature extraction is a challenging procedure (Bahiraei et al., 2020, 2021b; Hosseini et al., 2021a, 2022a). Meanwhile, the DNNs are capable of handling the feature extraction automatically within themselves via an end-to-end learning procedure (Li, 2017; Yang et al., 2020). In the recent decades, DNNs have been widely adopted for solving practical industrial problems (Bahiraei et al., 2021a). Thus, by exploiting the advantage of DNNs, this work proposes two models based on CNNs and LSTM modules. The rest of this paper is organized as follows. Section **??** reviews important relevant works. Section 5.3 elaborates on the proposed methodology, and Section 5.4 discusses the experimental findings. Finally, Section 5.6 concludes the paper with future direction.

4.3   Methodology

4.3.1   The proposed models

DNNs have been proven to be strong tools for dealing with regression, classification, and optimization problems in industrial environments (Bahiraei et al., 2021b; Hosseini et al., 2021b; Li and Akilan, 2022). Thus, this work extends upon the existing DNN technologies, such as 1-D convolution and LSTM cells to improve the temperature prediction accuracy of PMSMs. The following subsections highlight the key details of the two proposed models.

1-D CNN-based Model

The proposed 1-D CNN-based model consists of six layers subsuming a total of 204, 196 trainable parameters. Table 5.1 summarizes the proposed CNN-based model's connectivity pattern with respective layer details.

LSTM-based Model

The proposed LSTM-based model consists of five layers subsuming a total of 936, 228 trainable parameters. Table 5.2 summarizes the proposed LSTM-based model's connectivity pattern with respective layer details.

Table 4.1: Architecture Detail of the 1-D CNN-based Model

Layer ID	Layer Type	Output Dimension
Input	Input layer	(1024, 128, 108)
L1	Conv 1D	(1024, 122, 64)
L2	Conv 1D	(1024, 116, 64)
L3	MaxPooling 1D	(1024, 58, 64)
L4	Flatten	(1024, 3712)
L5	Dense	(1024, 32)
L6	Dense	(1024, 4)

Total number of trainable parameters: 204,196
Number of filters in L1 and L2: 64
Kernel size in L1 and L2: 7, and in L3: 2
Activation function: L1-L5: ReLu; L6: linear

Table 4.2: Architecture Detail of the LSTM-based Model

Layer ID	Layer Type	Output Dimension
Input	Input layer	(1024, 128, 108)
L1	LSTM	(1024, 128, 256)
L2	LSTM	(1024, 256)
L3	Dense	(1024, 128)
L4	Dense	(1024, 32)
L5	Dense	(1024, 4)

Total number of trainable parameters: 936,228
Number of hidden units in L1 and L2: 256
Activation function: L1, L2: tanh; L3, L4: Relu; L5: linear

4.3.2 Data pre-processing

Data scaling

Data scaling is an essential pre-processing step in time-series prediction problems and it has shown great performance gain in DNNs (Bhanja and Das, 2018). In this work, the standard scalar was used to normalize the values by subtracting the mean and dividing by the standard deviation of the values considering all the samples as given in (5.3). It is worth mentioning that during inference time on the test set, the predicted target values were re-scaled back to the original range.

$$x_{scaled} = \frac{x - \mu(x)}{\sigma(x)}, \tag{4.1}$$

Table 4.3: Input and Target Variables (Kirchgässner et al., 2020)

Variable name	Symbol
Measured Input Variables	
Ambient temperature	θ_a
Liquid coolant temperature	θ_c
Actual voltage d-axis component	u_d
Actual voltage q-axis component	u_q
Actual current d-axis component	i_d
Actual current q-axis component	i_q
Motor speed	n_{mech}
Torque	τ
Derived Input Features	
Voltage magnitude: $\sqrt{u_d^2 + u_q^2}$	u_s
Current magnitude: $\sqrt{i_d^2 + i_q^2}$	i_s
Electric apparent power: $1.5 * u_s * i_s$	S_{el}
Joint Interaction-A: $i_s * n_{mech}$	J_1
Joint Interaction-B: $S_{el} * n_{mech}$	J_2
Target Variables	
Permanent magnet temperature	θ_{PM}
Stator teeth temperature	θ_{ST}
Stator winding temperature	θ_{SW}
Stator yoke temperature	θ_{SY}

where x is the raw input, and its sample mean and standard deviation are denoted by $\mu(x)$ and $\sigma(x)$, respectively.

Derived attributes

According to the physical representations of quantities, like the total current, the total voltage and the apparent power, and the joint interaction between two inputs, a total of five features are derived as tabulated in Table 4.3.

EWMA and EWMS

New features are computed based using EWMA and EWMS with different span values defined by (4.2) and (4.3), where $\alpha = 2/(s + 1)$ and s is an arbitrary span or a lookback period defined by the user. As a result, at each time step t the EWMA and EWMS of each input (cf. Table 4.3)
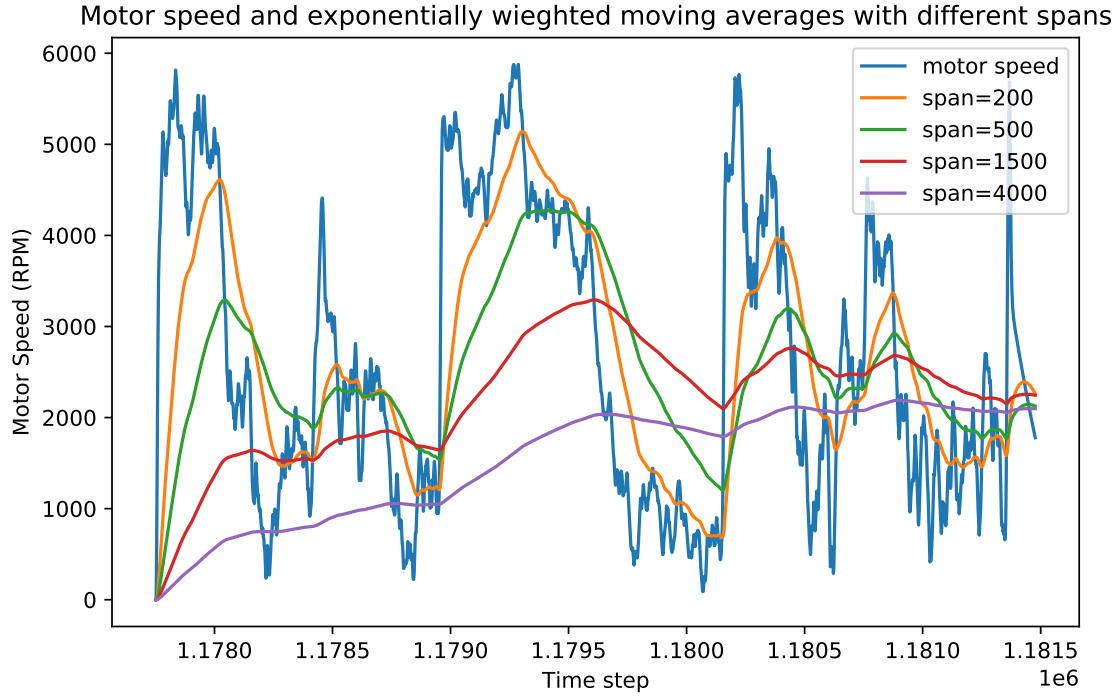
Figure 4.2: Raw motor speed and its EWMA features with different spans.

with four different spans are fed to the proposed models as additional inputs. As the input series is extremely dynamic with high fluctuation, the exponentially weighted features can capture some historical information of the inputs and boost the models' performances. A sample feature and its EWMA and EWMS with different spans are depicted in Fig. 4.2 and Fig. 4.3.

$$\mu_x = \frac{\sum_{i=0}^{t}(1-\alpha)^i x_{t-i}}{\sum_{i=0}^{t}(1-\alpha)^i}, \tag{4.2}$$

$$\sigma_x = \frac{\sum_{i=0}^{t}(1-\alpha)^i (x_i - \mu_t)}{\sum_{i=0}^{t}(1-\alpha)^i}. \tag{4.3}$$

4.3.3 Model implementation and training strategy

Before finalizing the models described in Tables 5.1 and 5.2, several CNN and LSTM networks with different numbers of layers and various architectures were implemented and their perfor-
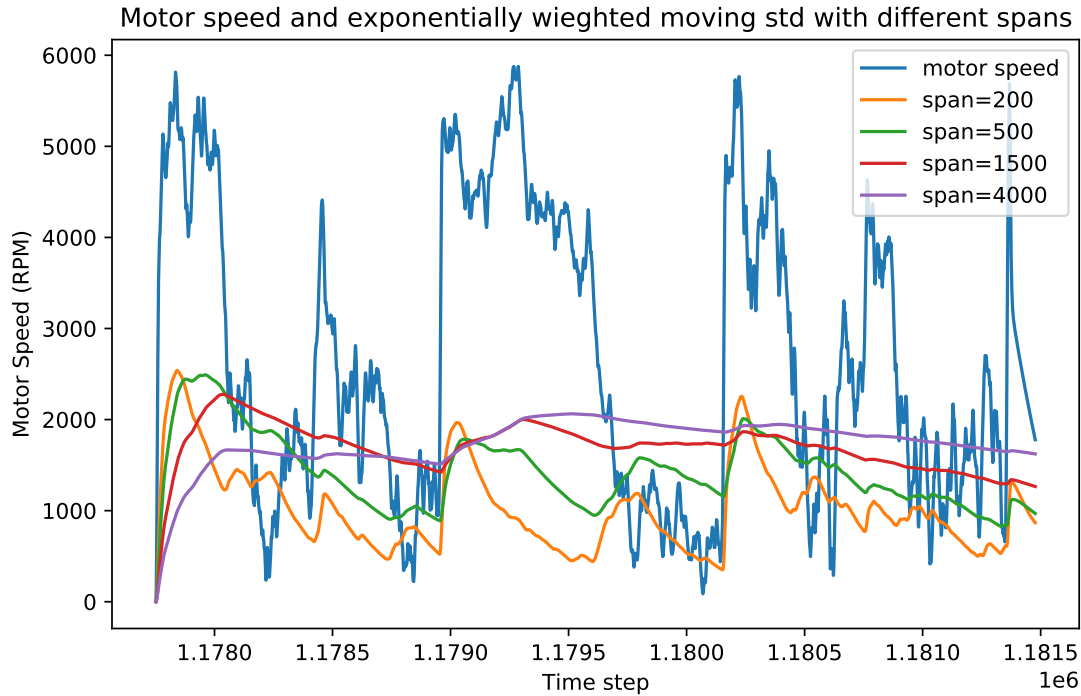
Figure 4.3: Raw motor speed and its EWMS with different spans.

mances were recorded in the test set. Those primitive results are omitted due to space limitations. The finalized models are trained using `Adam` optimizer with an initial learning rate of 0.001, and an input batch size of 1024. Moreover, the early stopping technique was used to handle overfitting issues and both models reached to the optimal condition between 8 and 17 epochs.

4.4 Experimental analysis

4.4.1 Data set

The data set comprises 185 hours of temperature recording of a PMSM sampled at 2 Hz frequency (Kirchgässner et al., 2020). The variables of the data set along with the derived attributes inspired by (Kirchgässner et al., 2020) and target values are listed in Table 4.3. It consists of sixty-nine profiles under different load conditions with the duration of each profile ranging from 0.3 to 6.1 hours. For a fair comparison, following Kirchgässner *et al.* (Kirchgässner et al., 2020), the profile ID 65 is chosen as the testing set. The rest is used for training, except for one profile (about

4.6 hours), which is used for model validation. Fig. 4.4 shows the sample size of each profile while Fig. 4.5 represents sample distribution across recording length.
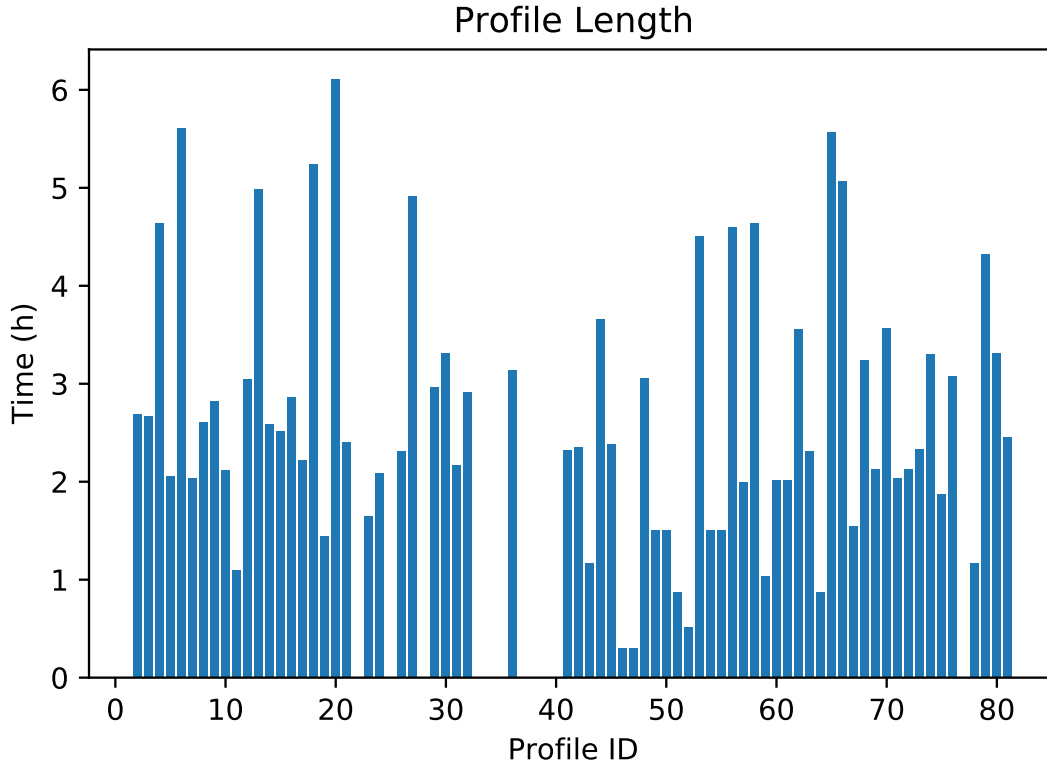
Figure 4.4: Profile length based on recorded hours.

A sample from the training set, profile ID 52, is visualized through a trend plot in Fig. 4.6.

To better understand the nature of the data, the correlation between each of the attributes is calculated and visualized via a heat map as in Fig. 4.7. Notably, the correlation between two attributes, namely, `i_q` and `torque`, is equal to 1. Thus, it is safe to remove `torque` from the input attribute list of the proposed models. In addition, due to some elements' close proximity in PMSM's internal structure, they experience similar thermal stress, so they have high correlation values. For example, the correlation between θ_{ST}, θ_{SY} and θ_{SW} is close to 1 because they are close to each other (cf. Fig. 4.1) and receive approximately the same amount of heat compared to center element – the rotor shaft.
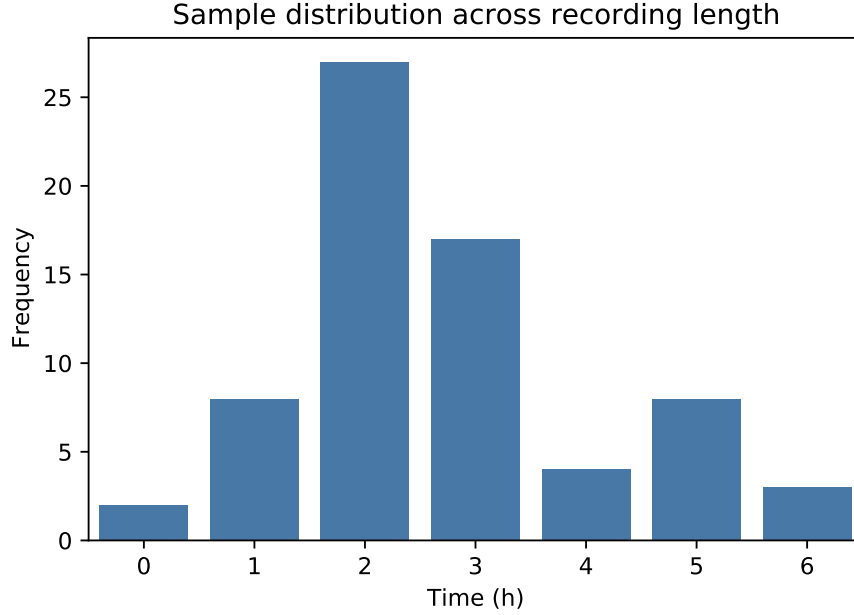
Figure 4.5: Sample distribution across recording length.

4.4.2 Evaluation metrics

Since this work handles regression problems (i.e., real value prediction), the mean squared error defined in (4.4) is used as the primary evaluation metric. The smaller MSE value indicates better performance of the model under study.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2, i = 1, 2, \cdots n, \tag{4.4}$$

where n, y_i, $\hat{y}_i$, and i are the total number of data points predicted, true value, the predicted value, and sample index, respectively.

4.4.3 Overall analysis

The data set used in this work is an emerging one; hence, not many works have used the data set for experimental studies. Thus, this work can perform a comparative analysis with only one recent work - Kirchgässner *et al.* (Kirchgässner et al., 2020) in Table 4.4. The analysis reveals that the 1-D CNN-based model has performed better than the LSTM-based model and the existing solution.
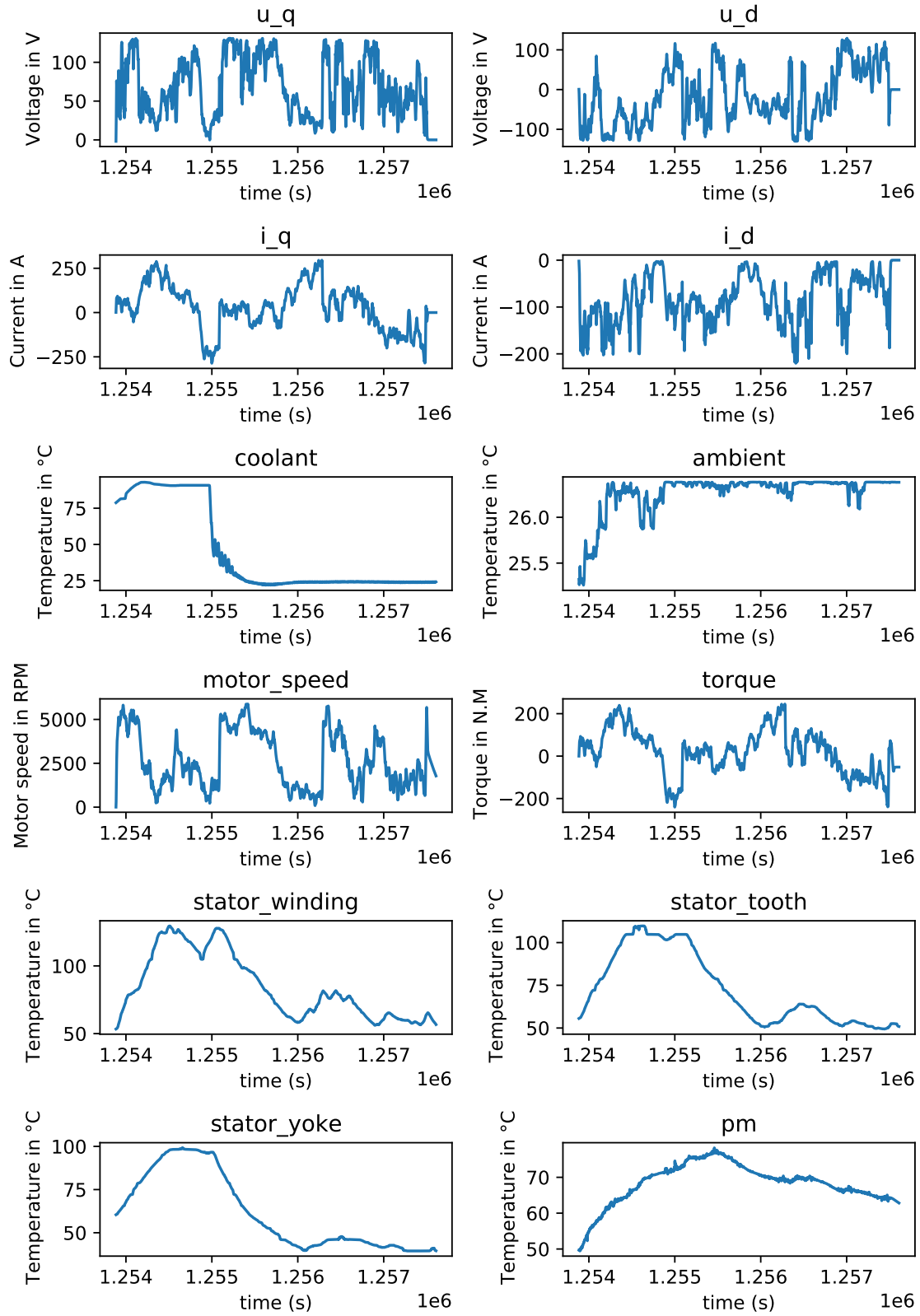
Figure 4.6: Profile ID 52. A sample load profile from the training set.

Figure 4.7: Heatmap of linear correlations between different data attributes.

Compared to the LSTM-based model, the CNN-based model records $> 35\%$ better overall performance. The intuition for this improvement is the excellent feature learning capability of CNNs. As EWMA and EWMS capture historical information, the CNN model unearths some intricate features that can not be learnt by pure long-term memory-based systems. At the same time, compared to the existing work, the proposed CNN-based model also achieves a 13% enhancement in average overall performance across all the target variables. Hence, the model's R^2 score for each target variable: θ_{ST}, θ_{SW}, θ_{SY}, and θ_{PM} is found to be 0.9950, 0.9954, 0.9948 and 0.9846, respectively. However, the R^2 scores are omitted in Table 4.4 since non of the existing works present this

measure. The performances of the proposed models on the test profile are compared to the actual target values in Fig. 4.8 and 4.9.
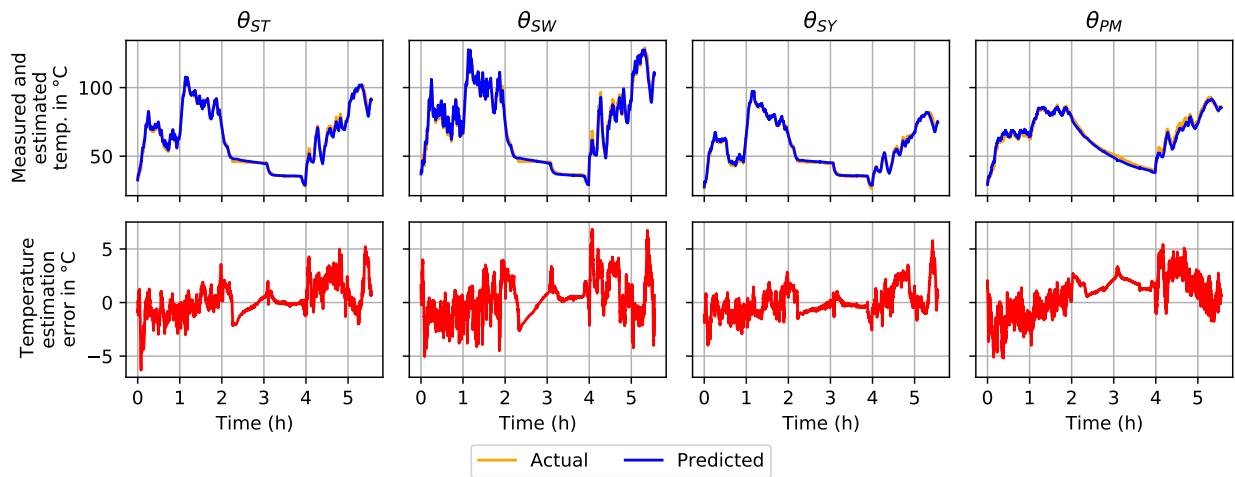
Figure 4.8: Performance of the 1-D CNN model on Profile ID: 65.
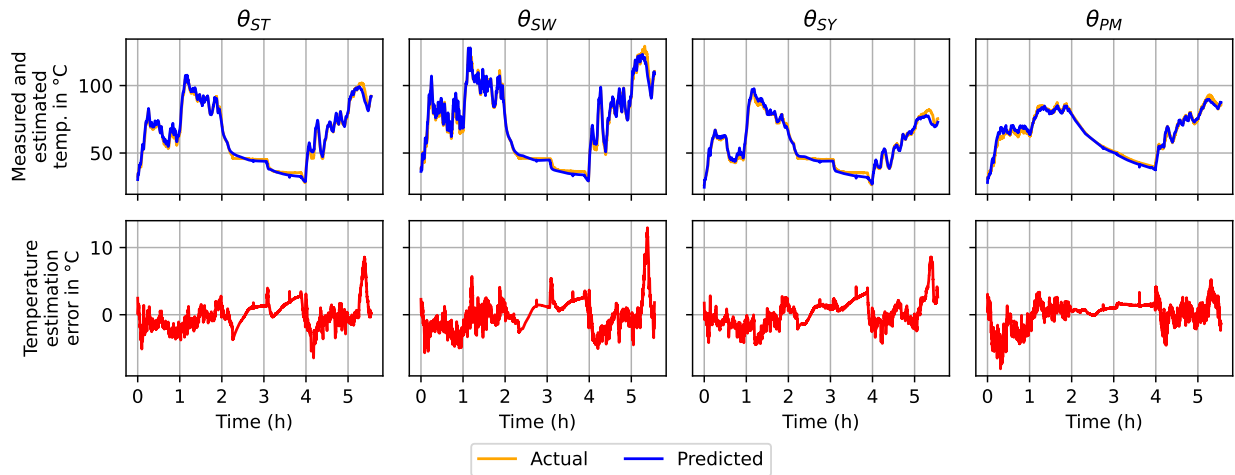
Figure 4.9: Performance of the LSTM model on Profile ID: 65.

4.5 Conclusion

This chapter proposed a 1-D CNN-based model and LSTM-based architectures to undertake temperature prediction of the internal elements of permanent magnet synchronous motors. In the initial stage, EWMA and EWMS were utilized for attribute conditioning and generating sufficient

Table 4.4: Comparison between different models wrt MSE: MA-MSE - Mean Average MSE Across All for targets, mAP - Mean Average Performance Compared to the Baseline Model, TCN (Kirchgässner et al., 2020). Note: The Best Performances Are Inked in Blue

Model	θ_{ST}	θ_{SW}	θ_{SY}	θ_{PM}	MA-MSE	mAP Improvement
Proposed 1-D CNN Model	2.21	3.34	1.52	3.50	2.64	+13%
Proposed LSTM Model	3.75	5.62	3.98	3.14	4.12	-35%
TCN (Kirchgässner et al., 2020)	2.84	6.90	1.80	0.65	3.04	Baseline
LTPN (Gedlu et al., 2020)	-	-	-	-	5.73	-88%
RNN (Kirchgässner et al., 2020)	-	-	-	-	8.70	-186%

features. Experimental results show that the CNN model can predict the desired target values with high precision and an average MSE within 2.64 °C^2. It is a significant improvement in comparison with former studies. The introduced data-driven temperature estimation method shows great promise that it can be a viable replacement for classical methods. From an application perspective, the proposed models can be generalized to other electric motors and even system components where thermal stress prediction is a paramount safety concern. However, this paper has not investigated the generalizability of the model and it is left for future work. It is also worth analyzing the effect of individual predictors and their respective exponentially weighted history on the performance of the proposed models.

Chapter 5

Oil Production Forecasting

5.1 Overview

Global oil demand is rapidly increasing and is expected to reach 106.3 million barrels per day by 2040. Thus, it is vital for hydrocarbon extraction industries to forecast their production to optimize their operations and avoid losses. Big companies have realized that exploiting the power of deep learning and the massive amount of data from various oil wells for this purpose can save a lot of operational costs and reduce unwanted environmental impacts. In this direction, researchers have proposed models using conventional machine learning techniques for oil production forecasting. However, these techniques are inappropriate for this problem as they can not capture historical patterns found in time series data, resulting in inaccurate predictions. This research aims to overcome these issues by developing advanced data-driven regression models using sequential convolutions and LSTM units. Exhaustive analyses are conducted to select the optimal sequence length, model hyperparameters, and cross-well dataset formation to build highly generalized robust models. A comprehensive experimental study on Volve oilfield data validates the proposed models. It reveals that the LSTM-based sequence learning model can predict oil production better than the 1-D CNN with mean absolute error (MAE) and R^2 score of 111.16 and 0.98, respectively. It is also found that the LSTM-based model performs better than all the existing state-of-the-art solutions and achieves

a 37% improvement compared to a standard linear regression, which is considered the baseline model in this work.

5.2   Introduction

The 18th century marked the first profound industrial revolution that predominantly exploited steam power replacing animal labour. Since then, there has been rapid development in industrial operations Xu et al. (2014). Now, the world has come to the brink of the fifth industrial revolution, a.k.a. industry 5.0, where smart systems are built to perform complex tasks more efficiently by leveraging advanced technologies, such as big data, high-performance computing (HPC) platforms, and data-driven analytics Ahmed et al. (2022); Liu et al. (2021b). Thus, industries are increasingly striving to create new and efficient production methods by utilizing the capabilities of AI. These advanced technologies offer a wide range of potential benefits, including increased automation, improved decision-making, and enhanced ability to process and analyze large amounts of data. Hence, the DNNs have become a cornerstone of several industrial operations, including accurate prediction or concept classification of operational conditions, aiming at smart control, real-time fault detection, and maintenance. For instance, in the oil and gas industry, intelligent assistive tools (IATs) for production forecasting based on readily accessible parameters is crucial for economic assessment and gain. Nevertheless, it is a challenging task due to (i) the complexity of the environmental and geographical subsurface conditions, (ii) the non-linear relationship between production volume and petro-physical parameters, such as permeability and density, and (iii) the shortage of curated data availability. Therefore, despite the technological advancement, hydrocarbon production analysis remains an active research field. It urges the research community to develop reliable and precise predictive models. Such models should provide more comprehension of the ongoing production, resulting in efficient operation, and informed decision-making and management.

In this direction, this work pragmatically develops two deep-learning models using 1-D CNN and LSTM to forecasting oil production. The main contributions of this work are summarized as follows.

- Comprehensive analysis of data pre-processing, including handling missing values, data scaling, and feature selection based on domain knowledge gained from the oil and gas industry.

- Applying systematic time series analysis that provides an in-depth analysis of sequence generation for model building.

- Hyper-parameter optimization by investigating the most effective model parameters.

- Generalized model development using data from various wells in the Volve oil field.

- Exhaustive ablation study and comparative analysis to validate the proposed models' performances.

It is worth mentioning that these highlighted contributions are fully or partially missing in the existing works conducted by other researchers on the same data collected from the Volve oil field. Thus, this study aims to bridge the main research gaps and propose new strategies to improve production forecasting in the hydrocarbon industries.

The rest of this paper is organized as follows. Section **??** reviews relevant works, Section 5.3 elaborates on the proposed models, and Section 5.4 presents the environmental setup, explanatory data analysis, model training details, and discusses the experimental results. Finally, Section 5.5 and Section 5.6 provides an overall summary and conclusion of the paper with future directions, respectively.

5.3 Methodology

In the last few decades, there has been a significant amount of focus on improving the architecture of DNNs. This attention is due to the fact that DNNs have been effective in solving a wide range

Standard scalar was applied as the normalization method

$$x_{scaled} = \frac{x - \mu(x)}{\sigma(x)} \quad (9)$$

1-D CNN and LSTM models were implemented to model the time series behavior of oil production

Each well's data was divided into training and testing (70% Training - 30% Testing) and in model training and testing, data from all of the wells was used

Hyperparameter setting, Model training and Model validation

R^2

MAE

NO

Sequence generation

the data was split sequentially throughout different stages of the well's production life

Selecting the best model topology and evaluating model on test data

Yes

February 2008 till September 2016

In model training and testing, data from all of the wells were used. This strategy makes the model more robust and generalizable that can be used for other wells.

Figure 5.1: An illustration of the phases involved in building the advanced deep learning-based oil production forecasting models. It subsumes several operations, including data gathering, data curation, model configuration, model training, and model evaluation.

of practical problems that arise in various industries. One of the most significant advantages of using these networks is their high capability to learn non-linear relationships regardless of the type of data (Akilan et al., 2017; Bahiraei et al., 2021a; Hosseini et al., 2020, 2021a,b, 2022a; Li and Akilan, 2022; Rad and Shahbandegan, 2020). As a result, there is a growing interest among researchers in optimizing the structure of DNNs. This has involved the exploitation of different topologies, such as skip connections, the application of various techniques to reduce the number of trainable parameters, and novel fast retraining to fine-tune pre-trained DNNs Yang et al. (2020).

In this direction, this work advances the regression models for hydrocarbon production forecasting by exploiting the power of the learning capability of the DNNs, particularly in time-series data analysis. This work adopts a modelling technique that progressively designs and develops predictive models from a baseline least squares-based linear regression to advanced deep learning-based regression models. Fig. 5.1 illustrates the phases involved in the building of the advanced deep learning-based solutions. The following subsections elaborate on the proposed solutions in a step-by-step manner.

5.3.1 Base model: A standard linear regressor

Linear regression (LR) estimates the linear relationship between different explanatory attributes and a dependant variable of given data samples (cf. 5.1) by minimizing an objective function, say the sum of the squares, i.e., the distance between each predicted and actual value of the dependent variable is squared and then summed up for all training samples. Due to its well-established mathematical foundation and easy training procedure, LR is widely accepted as a baseline model for various regression problems in several fields, including engineering, biomedical, behavioural, and social sciences, and business. The standard linear regression model can be defined as in (5.1).

$$y = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n + \beta, \tag{5.1}$$

where y, α_i, x_i, and β stand for the dependent variable (output), coefficient of the ith input attribute, ith input attribute, and bias, respectively. The coefficients are optimized by minimizing the total

Table 5.1: Architecture detail of the 1-D CNN-based Regressor

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	Conv 1D	(32, 4, 64)
L2	Conv 1D	(32, 3, 64)
L3	Conv 1D	(32, 2, 64)
L4	MaxPooling 1D	(32, 1, 64)
L5	Flatten	(32, 64)
L6	Dense	(32, 1)

Total number of trainable parameters: 18,049
Number of filters (F) in L1, L2, and L3: 64
Kernel size (K) in L1, L2, L3, and L4: 2
Activation function: L1-L5: ReLu; L6: linear
Stride rate (S) was equal to 1 in all layers
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

sum of squares (SST) defined in (5.2), which is the aggregation of the sum of squares (SSE), $\sum_{i=1}^{n}(y_i - \widehat{y_i})^2$ and the sum of squares (SSR), $\sum_{i=1}^{n}(\widehat{y_i} - \overline{y})^2$.

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}(y_i - \widehat{y_i})^2 + \sum_{i=1}^{n}(\widehat{y_i} - \overline{y})^2, \tag{5.2}$$

where y_i stands for the real value in observation i, $\overline{y}$ is the mean value of the dependant variable, y from all n observations, and $\widehat{y_i}$ denotes the predicted value of the dependant variable for the given ith observation's input attributes.

5.3.2 1-D CNN-based regressor

The hyper-parameters, such as the number of hidden layers, kernel size (K), number of filters (F), sub-sampling factor, and type of activation function used in each layer determine the structure of the 1-D CNN model. In this work, the proposed 1-D CNN regressor's structure, its layer connectivity, and the hyperparameter setting are given in Table 5.1.
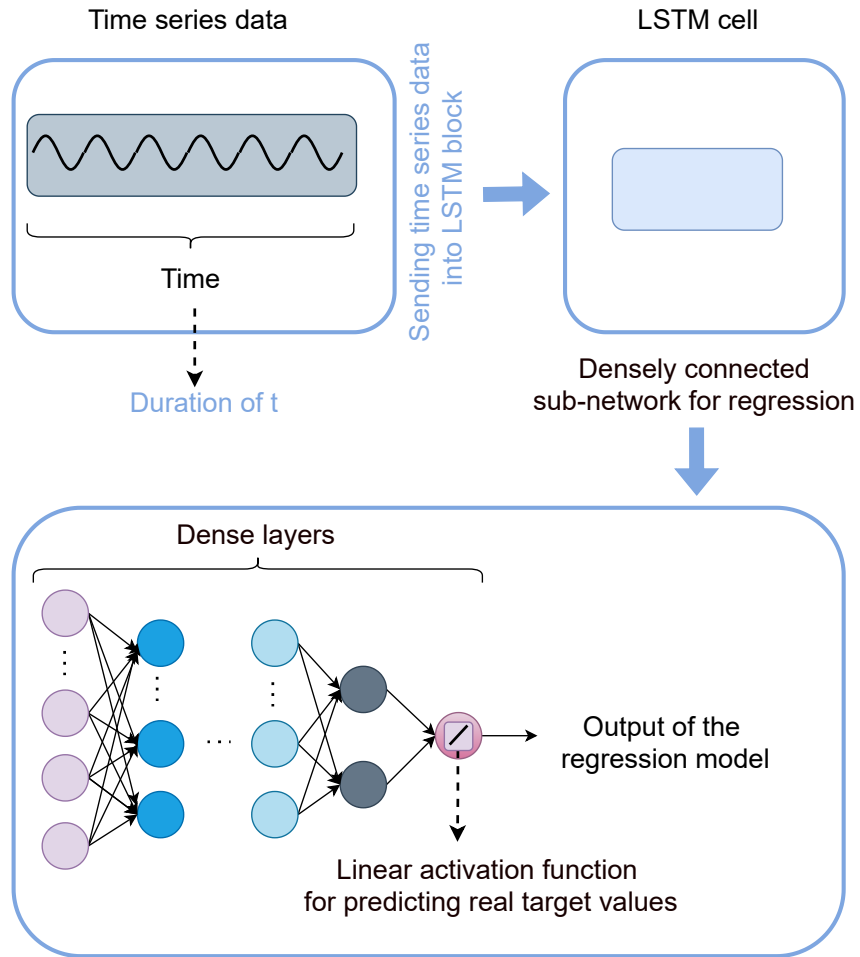
Figure 5.2: An overview of applying LSTM for sequence learning. Here, time series data with a sequence length of t is input to the LSTM subnetwork. The learnt representation by the LSTM subnetwork is then forwarded to the densely connected regression sub-network that generates the final output.

5.3.3 LSTM-based regressor

While Fig. 5.2 illustrates the general idea of how LSTMs can be applied for a sequence-based regression problem, Table 5.2 provides the layer connectivity details of the proposed LSTM-based regressor.

Table 5.2: Architecture detail of the LSTM-based Regressor

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	LSTM	(32, 5, 30)
L2	LSTM	(32, 20)
L3	Dense	(32, 20)
L4	Dense	(32, 16)
L5	Dense	(1)

Total number of trainable parameters: 9,893
Number of hidden units in L1 and L2: 30 and 20
Activation function: L1, L2: tanh; L3, L4: Relu; L5: linear
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

5.4 Experimental analysis

5.4.1 Environment

This work exploits Google Colaboratory cloud resources with one Tesla T4 graphical processing unit, 12 Gigabytes of RAM, and 2 CPU cores for training and evaluation. The models are built using Python programming language and the pre-built libraries, like Numpy, Pandas, Matplotlib, Seaborn, and the open-source deep learning library TensorFlow library.

5.4.2 Explanatory data analysis

Data source

This work uses the Volve oil field database for the experimental study. The Volve oil field is located in the central part of the North Sea at 2750 - 3120 m depth. The field was revealed in 1993, and the drilling process started in May 2007 Ng et al. (2022). After 8.5 years, the Volve oil field was decommissioned in 2016 Noshi et al. (2019).

In May 2018, Equinor released the Volve database publicly for research and development purposes Equinor (2018a,b,c). The database includes different categories of data collected from various operations, but this study focuses on real-field production data. The production data subsumes information gathered from seven wells (five producers and two injectors) as summarized

Table 5.3: Historical recordings of different wells in the Volve oil field. Note that data were recorded daily from each well.

Well code	Type	Start of recording	End of recording
NO 15/9-F-1 C	Producer	07-Apr-14	21-Apr-16
NO 15/9-F-11 H	Producer	08-Jul-13	17-Sep-16
NO 15/9-F-12 H	Producer	12-Feb-08	17-Sep-16
NO 15/9-F-14 H	Producer	12-Feb-08	17-Sep-16
NO 15/9-F-15 D	Producer	12-Jan-14	17-Sep-16
NO 15/9-F-4 AH	Injector	01-Sep-07	01-Dec-16
NO 15/9-F-5 AH	Injector	01-Sep-07	18-Sep-16

in Table 5.3, namely NO 15/9-F-1 C, NO 15/9-F-11H, NO 15/9-F-12 H, NO 159–F–14 H, NO 15/9-F-15 D, NO15/9-F-4 AH, and NO 15/9-F-5 AH, where 15/9-F-4 and 15/9-F-5 are the two injectors.

Attributes

Table 5.4 lists the attributes of the collected data from seven wells stated in Table 5.3. To understand the nature of each attribute, they are visualized using trend plots wrt duration. For example, Fig. 5.3 visualizes the attributes of well no. 14 (NO 15/9-F-14 H). Similarly, the attributes' basic statistical information are also analyzed for each well. For instance, Table 5.5 provides the statistical information: mean, standard deviation, minimum value, 1st quartile, median, 3rd quartile, and maximum value of each attribute of well no. 14.

Handling missing values

Handling missing values plays a crucial role in data-driven model building. The Volve oil field database contains missing values in certain attributes. For example, one can find from the statistical information of well no. 15/9-F-14 H summarized in Table 5.5 that there are data samples with missing values in attributes ADP, ADT, ADPT, ACP, AAP, AWP, AWT, and DPC. It's worth noting that around 7 percentage of the dataset contains missing values. To resolve this, it is important to understand the information distribution of attributes across all the samples. In this case, the distribution is observed using boxplots as shown in Fig. 5.4. From these plots, it is clear that the

Table 5.4: Description of wells' attributes recorded from the Volve oil field Ng et al. (2022)

Attributes	Description	Abbreviations
DATEPRD	Date of Record	DOR
ON_STREAM_HRS	On stream hours	OSH
AVG_DOWNHOLE_PRESSURE	Average Downhole Pressure	ADP
AVG_DOWNHOLE_TEMPERATURE	Average Downhole Temperature	ADT
AVG_DP_TUBING	Average Differential Pressure (Tubing)	ADPT
AVG_ANNULUS_PRESS	Average Annular Pressure	AAP
AVG_CHOKE_SIZE_P	Average Choke Size Percentage	ACP
AVG_WHP_P	Average Wellhead Pressure	AWP
AVG_WHT_P	Average Wellhead Temperature	AWT
DP_CHOKE_SIZE	Differential Pressure of Choke Size	DPC
BORE_OIL_VOL	Oil Volume from Well	O
BORE_WAT_VOL	Water Volume from Well	W
BORE_GAS_VOL	Gas Volume from Well	G
BORE_WI_VOL	Water Volume Injected	WI
FLOW_KIND	Type of Flow (production or injection)	TOF
WELL_TYPE	Type of Well (Producer or Injector)	TOW

Table 5.5: Fundamental statistical information of different features of well No 15/9-F-14 H from the Volve database

	OSH	ADP	ADT	ADPT	ACP	AAP	AWP	AWT	DPC	G	O
Count	9001	8980	8980	8980	8759	7730	8995	8995	8995	9001	9001
Mean	20.1969	181.8039	77.16297	154.0288	54.91872	14.79186	45.80569	68.14038	19.4882	163182.8	1110.534
STD	8.266102	109.7124	45.65795	76.75237	36.67992	8.411588	24.72463	27.67885	22.62904	189092.9	1330.366
Min	0	0	0	0	0	0	0	0	0	0	0
1st quartile	24	0	0	83.66536	18.59543	10.7977	31.42122	58.96197	3.480234	29381.58	190.4
Median	24	232.8969	103.1867	175.5889	51.6073	16.18796	38.3469	80.2844	9.888581	89678.31	574.57
3rd quartile	24	255.4015	106.2766	204.32	99.9372	21.26906	57.80183	88.11501	27.39888	208411.8	1377.51
Max	25	397.5886	108.5022	345.9068	100	30.01983	137.311	93.50958	125.7186	851131.5	5901.84

data is highly skewed, so the best approach for missing value imputation is replacing the missing value with the median value of the respective attribute.

Feature selection

In order to gain a better comprehension of the data, correlations among all of the attributes are calculated and visualized using a heat map shown in Fig. 5.5. It is worth mentioning that the correlation between ADP, ADT, and ADPT is equal to 0.97 and 0.95, respectively. Therefore,
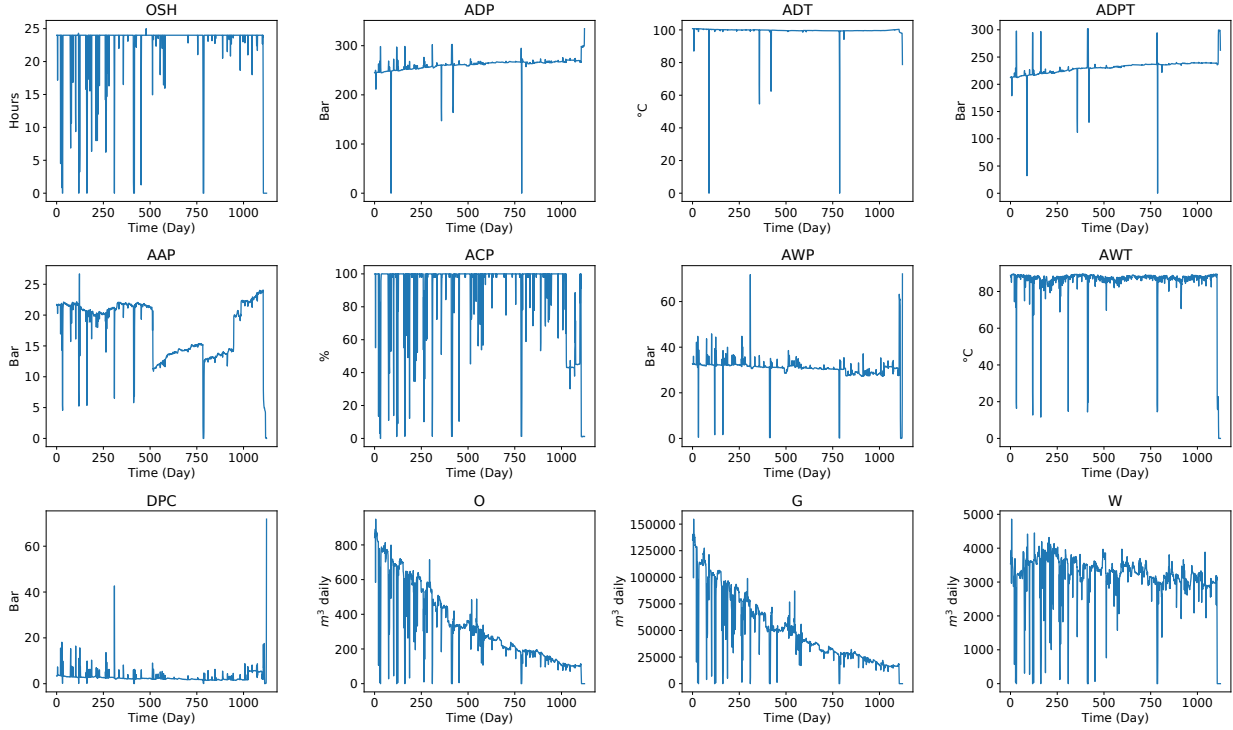
Figure 5.3: Trend plot visualization of the attributes: OSH, ADP, ADT, ADPT, AAP, ACP, AWP, AWT, DPC, O, G and W for the well no 15/9-F-14 H.

ADP is removed from the input variable list. Moreover, the correlation between produced gas and oil is equal to 1, which complies that oil flow produces gas.

5.4.3 Data preprocessing

Data scaling

Data scaling is a critical preprocessing step, which shows substantial performance gain in time-series analysis Akilan et al. (2018); Bahiraei et al. (2020); Hosseini et al. (2021a). In this work, the standard scalar was applied as the normalization method. This method normalizes the data by subtracting the mean and dividing it by the standard deviation as given in (5.3). It should be noted that during the process of analyzing the test set, predicted target values are re-scaled to the original range.

$$x_{scaled} = \frac{x - \mu(x)}{\sigma(x)}, \tag{5.3}$$
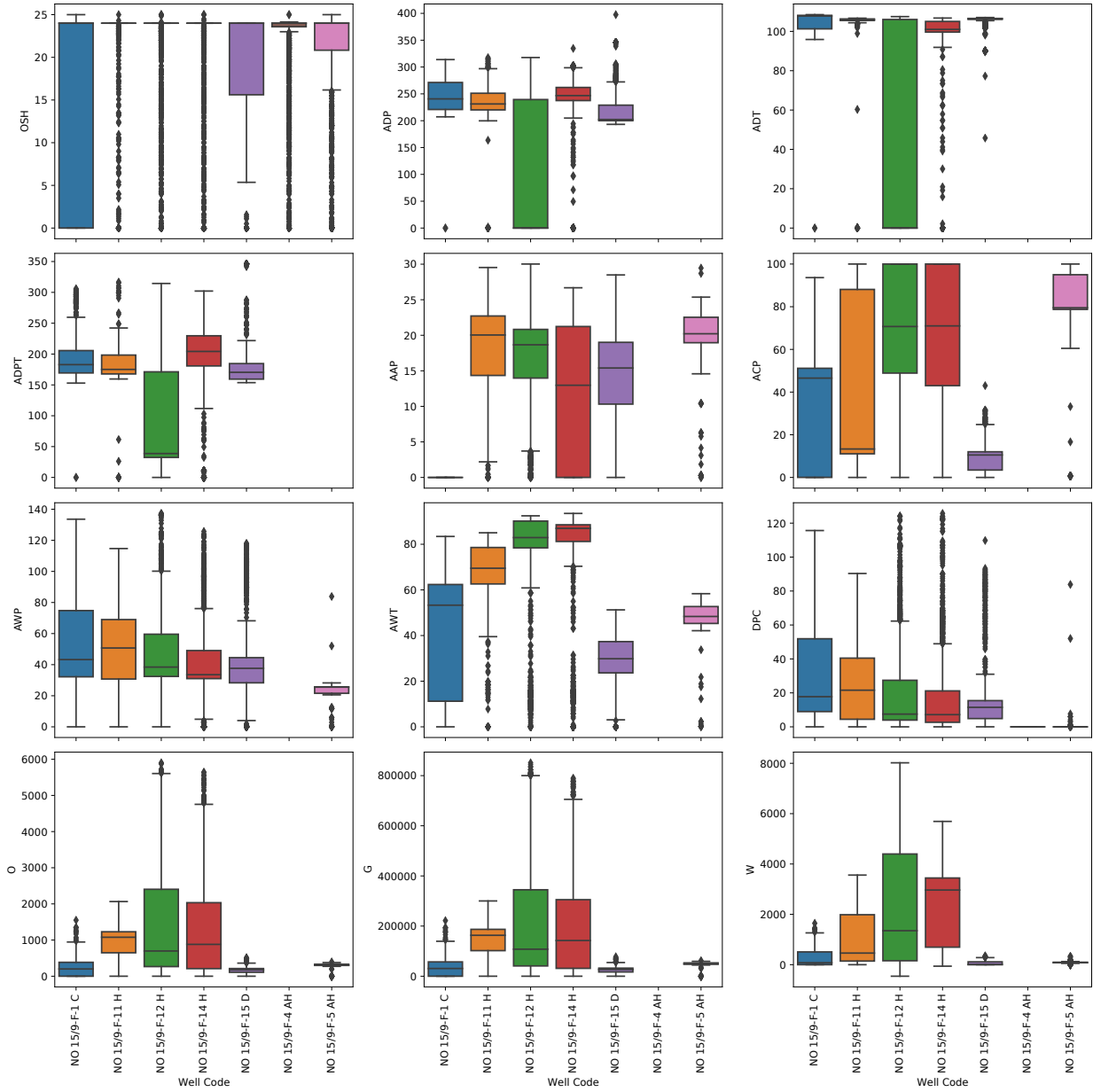
Figure 5.4: Visualizing the data distribution of the key attributes with respect to each well listed in Table 5.3.

where x, $\mu(x)$, and $\sigma(x)$ are the raw input, the sample mean, and the standard deviation, respectively.

Dataset curation

It is found that the existing works on the Volve oil field production database are well-specific models and they are not generalized solutions to all the wells. To address this, we organized the

Figure 5.5: Heatmap of linear correlations between different Features.

data set for training and testing that comprises sequential data samples from the five production-related wells listed in Table 5.3.

In this regard, after generating sequences, each well's data was divided into training and testing (70% Training - 30% Testing) to make a train and test set named "global train set" and "global test set". The global train set includes 70% of each well's data to use in the model training process, and the global test set comprises 30% of each well's data for the testing procedure. This strategy makes the model more robust and generalized that can be used for other wells in different oilfields.

5.4.4 Evaluation metrics

This work uses MAE and R-squared (R^2 score) defined in (5.4) and (5.5) to measure the precision of the oil production forecasts by the proposed models.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n}, \tag{5.4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}, \tag{5.5}$$

where y_i, $\hat{y}_i$ and $\bar{y}$ stand for actual, predicted and average values of target attributes, respectively.

5.4.5 Hyper-parameter tuning

Hyperparameter tuning is a meta-optimization task, which plays a vital role in the realm of DL. Effective hyperparameter adjustments have proved to improve predictive models' performances. In this case, two important aspects, such as the input sequence length and the layer configuration of the proposed model are considered for hyper-parameter tuning.

Optimizing sequence length

In time-series analysis, selecting the optimal sequence length highly influences the prediction's precision. It requires a strong domain knowledge to comprehend the impact of previous data points in the data sequence. In this work, for the proposed 1-D CNN-based and LSTM-based oil production forecasting models, various sequence lengths (ranging from three to eight) are considered in selecting the optimal sequence length. Fig. 5.6 shows the influence of sequence length in the proposed models' performance in terms of MAE. From these analyses, it is evident that when the sequence length is increased from three to five, the error in the forecast gradually decreases, but beyond that range, the error increases. As a result, a sequence length of five is chosen as the optimal value.
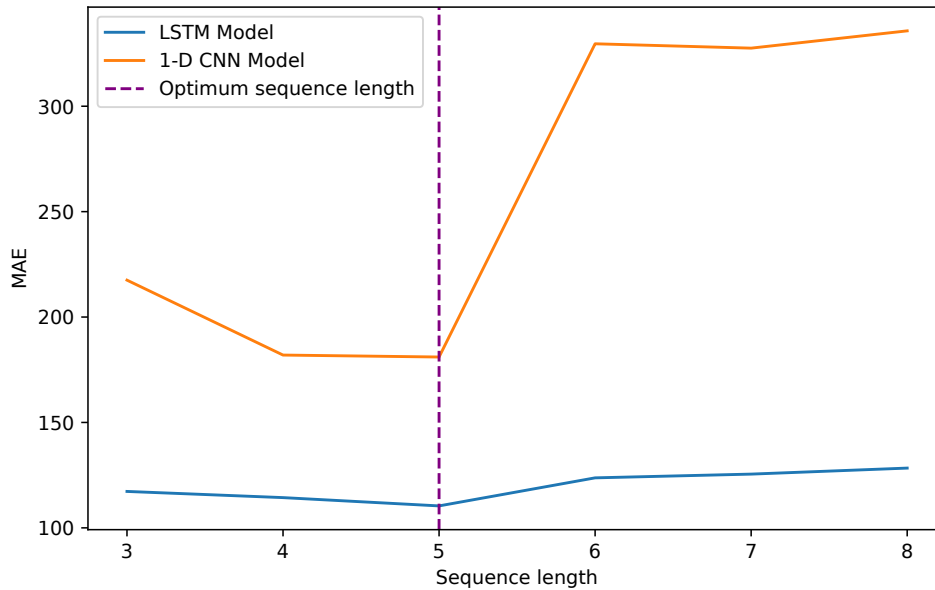
Figure 5.6: Analysing the impact of input sequence length in forecasting the oil production with respect to evaluation metric (MAE).

Optimizing layer configuration

To finalize an optimal model wrt the number of hidden layers and neurons several sanity analyses are conducted. For example, Table 5.6 summarizes the performances of ten different architectural configurations that led us to finalize the optimal model elaborated in Section 5.3.3 and Table 5.2.

It is clear from this table, the best performant models (i.e., with the lowest MAE and the highest R-squared) are model 3 and model 4, respectively for LSTM-based and 1-D-based regressors.

For more information about the optimized LSTM network, Table 5.2 was provided to summarize the model's connectivity pattern with respective layer details. Furthermore, the best performance for 1-D CNN network was gained by model number 4, which includes three convolutions, one max pooling, one flatten and one dense layer. A table was provided to conclude the details of the proposed CNN-based model.

Additional architectural information on how we finalize the best deep regression model for oil production forecasting was provided in the following Tables. It summarizes all the models' (i.e., LSTM models 1 - 5 and 1-D CNN models 1 - 5) layer-wise connectivity patterns along with their hyper-parameter details.

Table 5.6: Sanity Analysis for Finalizing the Best Architecture for Oil Production Forecasting.

Type	Model ID	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6	Layer 7	FLOPs	MAE	R^2
	Model 1	LSTM	Dense	Dense	∅	∅	∅	∅	27552	136.37	0.97
	Model 2	LSTM	LSTM	Dense	Dense	∅	∅	∅	30272	113.39	0.98
LSTM	Model 3	LSTM	LSTM	Dense	Dense	Dense	∅	∅	33312	111.16	0.98
	Model 4	LSTM	LSTM	LSTM	Dense	Dense	Dense	∅	40352	114.58	0.97
	Model 5	LSTM	LSTM	LSTM	Dense	Dense	Dense	Dense	48288	117.40	0.97
	Model 1	Conv	MaxPooling	Flatten	Dense	∅	∅	∅	385056	199.88	0.94
	Model 2	Conv	Conv	MaxPooling	Flatten	Dense	∅	∅	1955872	322.88	0.84
1-D CNN	Model 3	Conv	Conv	MaxPooling	Flatten	Dense	Dense	∅	1993696	293.49	0.85
	Model 4	Conv	Conv	Conv	MaxPooling	Flatten	Dense	∅	3008544	151.64	0.96
	Model 5	Conv	Conv	Conv	MaxPooling	Flatten	Dense	Dense	3046368	161.34	0.95

Note: ∅ means that model does not include that layer.

We derive the best model starting from a vanilla architecture (LSTM-model#1, and 1-D CNN-model#1), then gradually increase their complexity by adding more hidden layers as shown in Table 5.6. Tables 5.7-5.16 are provided to show more information about the architecture of the considered models in this thesis.

Table 5.7: Architecture detail of the 1-D CNN-based Regressor (Model 1)

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	Conv 1D	(32, 4, 64)
L2	MaxPooling 1D	(32, 2, 64)
L3	Flatten	(32, 128)
L4	Dense	(32, 1)

Total number of trainable parameters: 1,601
Number of filters (F) in L1: 64
Kernel size (K) in L1, L2: 2
Activation function: L1-L3: ReLu; L4: linear
Stride rate (S) was equal to 1 in all layers
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

Table 5.8: Architecture detail of the 1-D CNN-based Regressor (Model 2)

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	Conv 1D	(32, 4, 64)
L2	Conv 1D	(32, 3, 64)
L3	MaxPooling 1D	(32, 1, 64)
L4	Flatten	(32, 64)
L5	Dense	(32, 1)

Total number of trainable parameters: 9,793
Number of filters (F) in L1 and L2: 64
Kernel size (K) in L1, L2 and L3: 2
Activation function: L1-L4: ReLu; L5: linear
Stride rate (S) was equal to 1 in all layers
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

Table 5.9: Architecture detail of the 1-D CNN-based Regressor (Model 3)

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	Conv 1D	(32, 4, 64)
L2	Conv 1D	(32, 3, 64)
L3	MaxPooling 1D	(32, 1, 64)
L4	Flatten	(32, 64)
L5	Dense	(32, 10)
L6	Dense	(32, 1)

Total number of trainable parameters: 10,389
Number of filters (F) in L1 and L2: 64
Kernel size (K) in L1, L2 and L3: 2
Activation function: L1-L5: ReLu; L6: linear
Stride rate (S) was equal to 1 in all layers
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

Table 5.10: Architecture detail of the 1-D CNN-based Regressor (Model 4)

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	Conv 1D	(32, 4, 64)
L2	Conv 1D	(32, 3, 64)
L3	Conv 1D	(32, 2, 64)
L4	MaxPooling 1D	(32, 1, 64)
L5	Flatten	(32, 64)
L6	Dense	(32, 1)

Total number of trainable parameters: 18,049
Number of filters (F) in L1, L2 and L3: 64
Kernel size (K) in L1, L2, L3 and L4: 2
Activation function: L1-L5: ReLu; L6: linear
Stride rate (S) was equal to 1 in all layers
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

Table 5.11: Architecture detail of the 1-D CNN-based Regressor (Model 5)

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	Conv 1D	(32, 4, 64)
L2	Conv 1D	(32, 3, 64)
L3	Conv 1D	(32, 2, 64)
L4	MaxPooling 1D	(32, 1, 64)
L5	Flatten	(32, 64)
L6	Dense	(32, 10)
L7	Dense	(32, 1)

Total number of trainable parameters: 18,645
Number of filters (F) in L1, L2 and L3: 64
Kernel size (K) in L1, L2, L3 and L4: 2
Activation function: L1-L6: ReLu; L7: linear
Stride rate (S) was equal to 1 in all layers
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

5.4.6 Model complexity analysis

In deep learning, the process of passing a single input through a model to produce an output is called an inference. It is important to know the inference time of a model in advance because it allows researchers to design and optimize the model for better performance. To measure the

Table 5.12: Architecture detail of the LSTM-based Regressor (Model 1)

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	LSTM	(32, 30)
L2	Dense	(32, 20)
L3	Dense	(32, 1)

Total number of trainable parameters: 5,681
Number of hidden units in L1 : 30
Activation function: L1: tanh; L3: Relu; L5: linear
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

Table 5.13: Architecture detail of the LSTM-based Regressor (Model 2)

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	LSTM	(32, 5, 30)
L2	LSTM	(32, 20)
L3	Dense	(32, 20)
L4	Dense	(32, 1)

Total number of trainable parameters: 9,561
Number of hidden units in L1 and L2: 30 and 20
Activation function: L1 and L2: tanh; L3: Relu; L4: linear
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

Table 5.14: Architecture detail of the LSTM-based Regressor (Model 3)

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	LSTM	(32, 5, 30)
L2	LSTM	(32, 20)
L3	Dense	(32, 20)
L3	Dense	(32, 16)
L4	Dense	(32, 1)

Total number of trainable parameters: 9,893
Number of hidden units in L1 and L2: 30 and 20
Activation function: L1 and L2: tanh; L3 and L4: Relu; L5: linear
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

Table 5.15: Architecture detail of the LSTM-based Regressor (Model 4)

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	LSTM	(32, 5, 30)
L2	LSTM	(32, 5, 25)
L3	LSTM	(32, 20)
L4	Dense	(32, 15)
L5	Dense	(32, 10)
L6	Dense	(32, 1)

Total number of trainable parameters: 14,806
Number of hidden units in L1, L2 and L3 : 30, 25 and 20
Activation function: L1, L2 and L3: tanh; L4 and L5: Relu; L6: linear
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

Table 5.16: Architecture detail of the LSTM-based Regressor (Model 5)

Layer ID	Layer Type	Output Dimension
Input	Input layer	(32, 5, 12)
L1	LSTM	(32, 5, 30)
L2	LSTM	(32, 5, 25)
L3	LSTM	(32, 20)
L4	Dense	(32, 15)
L5	Dense	(32, 10)
L6	Dense	(32, 5)
L7	Dense	(32, 1)

Total number of trainable parameters: 14,856
Number of hidden units in L1, L2 and L3 : 30, 25 and 20
Activation function: L1, L2 and L3: tanh; L4, L5 and L6: Relu; L7: linear
Learning rate: 0.001; Number of epochs: 200
Optimizer: Adam; Batch size: 32

inference time, the total number of computations performed by the model must be calculated. One way to do this is by using a measure called Floating Point Operations (FLOPs), which counts the number of operations involving floating point values in the model. The inference time can then be calculated by dividing the number of FLOPs by the number of FLOPs that the CPU can perform in a given amount of time.

5.5 Overall analysis

5.5.1 Research gap in the existing works

In the existing works, there is a lack of experiments conducted on the Volve oil field production datasets. We believe it is because the dataset became public quite recently in 2018. In addition, all the existing works focus on well-specific model developments. Such models are not applicable to forecast the oil production in other wells. Hence, the existing works do not report their performances in all relevant evaluation metrics (MAE and R^2 score). However, this work develops a generalized model, which is applicable to all the oil production wells in the Volve oil field and the performance of the models is evaluated using MAE an R^2 score.

5.5.2 Quantitative analysis

A thorough comparative analysis is conducted in comparison to the existing works, in Table 5.17, where the standard linear regression described in Section 5.3.1 is considered as the baseline. When compared to this baseline, the best existing work, the conventional neural network-based solution proposed by Chahar *et al.* (Chahar et al., 2022) achieves 8.5% improvement, while the proposed LSTM-based and 1-D CNN-based models provide significant improvement of 37% and 14%, respectively. In addition to its superior performance, the complexity of the proposed LSTM-based regressor is $\simeq 45\%$ less than that of the CNN-based counterpart in terms of the number of trainable parameters (cf. Table 5.1, Table 5.2 and Table 5.6). Hence, it requires only about 2% of computations compared to the CNN-based counterpart in terms of FLOPs. Therefore, the proposed LSTM-based model is a more resource-friendly and efficient solution. The holistic analysis suggests that the LSTM-based oil production forecasting model is a robust, generalized, and reliable solution. As it can be seen from Table 5.6, LSTM models 2 and 3 have similar MAE and FLOPs. Model 2 is preferable for the resource-limited computational platform, while model 3 is a better solution when prediction precision is very important for the optimal operation of oil production forecasting.

Table 5.17: A Quantitative Analysis. Comparing the Performances of the Proposed Models and the Existing Works. Note: In terms of comparison, (1) FLOPS and MAE, the Lower the Better, (2) R^2, the Higher the Better, (3) "—" Means No Information Is Found in the Respective Work, and (3) ⇑ - Average Percentage of Improvement Compared to the Baseline Model.

Network	Error Metric	15/9-F-12 H	15/9-F-14 H	15/9-F-1 C	15/9-F-11 H	15/9-F-15 D	Global Test Set	FLOPs	Parameters	⇑
Proposed LSTM Model	MAE	109.42	73.77	177.64	133.49	156.74	111.16	33312	9893	+37%
	R^2	0.98	0.99	0.95	0.97	0.95	0.98			
Proposed 1-D CNN Model	MAE	157.04	106.62	201.29	186.58	196.52	151.64	3008544	18049	+14%
	R^2	0.97	0.98	0.95	0.94	0.91	0.96			
Linear Regression Model	MAE	209.65	218.76	256.10	196.01	270.40	176.36	Not defined	Not defined	Base line
	R^2	0.94	0.91	0.79	0.90	0.90	0.92			
Neural Network Chahar et al. (2022)	MAE	161.24								+8.5%
	R^2	0.94	—	—	—	—	—	—	—	
Random Forest Wang et al. (2021a)	MAE	—								-1%
	R^2	0.91	—	—	—	—	—	—	—	
Decline Curve Masini et al. (2019)	MAE	—	—							-10%
	R^2		0.82	—	—	—	—	—	—	
AdaBosst Noshi et al. (2019)	MAE	—	300							-70%
	R^2		—	—	—	—	—	—	—	
LSTM Ng et al. (2022)	MAE	—	—							+5%
	R^2		0.97	—	—	—	—	—	—	

5.5.3 Qualitative analysis

Figures 5.7 and 5.8 compare the proposed models' oil production forecasting results with the ground truths of the respective test sequences wrt the five production wells in the Volve oil field. From the plots of the prediction and actual values, one can observe that the proposed models' forecasts are very close to the actual values. It is further verified by the quantitative comparisons given in Table 5.17.
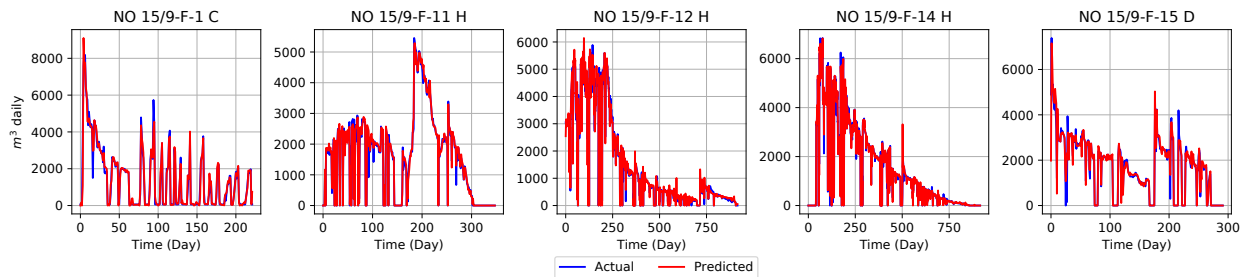

Figure 5.7: Qualitative analysis I. Plots of the proposed LSTM-based oil production forecasting along with the ground truth values for all five production wells, No 15/9-F-1 C, No 15/9-F-11 H, No 15/9-F-12 H, No 15/9-F-14 H, and No 15/9-F-15 D, in the Volve oil field.

Figure 5.8: Qualitative analysis II. Plots of the proposed 1D CNN-based oil production forecasting along with the ground truth values for all five production wells, No 15/9-F-1 C, No 15/9-F-11 H, No 15/9-F-12 H, No 15/9-F-14 H, and No 15/9-F-15 D, in the Volve oil field.
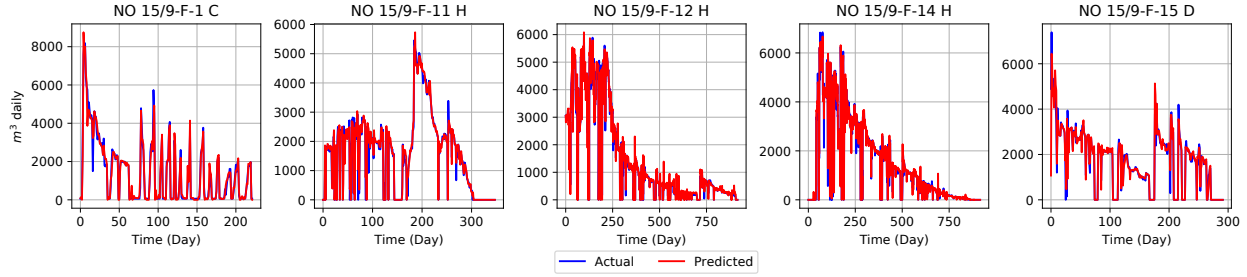
5.6   Conclusion

In this study, an LSTM-based model and a 1-D CNN-based model were proposed for the time series production forecasting of the Volve oilfield. To determine the appropriate sequence length in the time series data analysis, a comprehensive investigation was conducted at the outset. After the best model topology was chosen based on the hyper-parameter tuning procedure, it was found that the LSTM-based model had better performance than the 1-D CNN model, as demonstrated by the MAE, R-squared, and complexity metric. From an applied perspective, since data from all of the wells was used in the training and testing of the models, they can be generalized to the other existing wells in the Volve oilfield . However, it is important to note that the generalizability of the models for the other oilfields has not been investigated in this paper and is a topic for future research. In addition, another potential direction for future research in oil production forecasting using deep neural networks could be the integration of additional data sources. For example, incorporating data on well activity, drilling plans, and geological information could potentially improve the accuracy of forecasts.

Chapter 6

Conclusion

In the oil and gas industry, the drilling stage is of utmost importance as it is responsible for extracting valuable resources from beneath the earth's surface. Accurately measuring the volume of extracted materials is crucial for determining the economic viability of the operation. To address this issue, a machine learning-driven solution has been developed, which has shown a significant improvement compared to existing studies. The proposed model has an RMSE of 1.1%, making it a valuable tool for the industry.

The transportation stage is another critical aspect of the industry, as it involves moving extracted materials from the drilling site to the processing plant. Electric motors play a significant role in this stage, and their failure can cause significant disruptions and losses. To prevent such incidents, two deep learning-based solutions have been proposed, namely a 1-D CNN and an LSTM model. Through a thorough analysis, the CNN-based model was found to be more accurate than the LSTM model, achieving an average MSE of 2.64 °C^2 and an average R^2 of 0.9924. Additionally, the proposed model showed a 13% mean average performance improvement compared to existing state-of-the-art solutions.

The production stage is the final stage in the oil and gas industry, where extracted materials are processed to obtain valuable products. Accurately forecasting future production levels is crucial for planning and decision-making purposes. To achieve this, two advanced data-driven regression models using sequential convolutions and LSTM have been developed. These models have

been tested and validated through a comprehensive experimental study on Volve oilfield data. The LSTM-based model outperformed the 1-D CNN model, achieving an MAE of 111.16 and an R^2 score of 0.98. This is a significant improvement compared to the baseline model of standard linear regression.

In conclusion, the proposed solutions for the three main stages in the oil and gas industry have shown promising results and can potentially improve the efficiency and profitability of the industry. These models have the potential to reduce costs, increase accuracy, and minimize disruptions caused by equipment failure, ultimately leading to a more sustainable and efficient oil and gas industry.

Bibliography

Professor's perceptron paved the way for AI – 60 years too soon — Cornell Chronicle — news.cornell.edu, https://news.cornell.edu/stories/2019/09/ professors-perceptron-paved-way-ai-60-years-too-soon, [Accessed 07-Apr-2023].

Åbro, E. and Johansen, G.: Improved void fraction determination by means of multibeam gamma-ray attenuation measurements, Flow measurement and instrumentation, 10, 99–108, 1999.

Adedigba, S. A., Oloruntobi, O., Khan, F., and Butt, S.: Data-driven dynamic risk analysis of offshore drilling operations, Journal of petroleum science and engineering, 165, 444–452, 2018.

Ahmed, I., Jeon, G., and Piccialli, F.: From Artificial Intelligence to Explainable Artificial Intelligence in Industry 4.0: A Survey on What, How, and Where, IEEE Transactions on Industrial Informatics, 18, 5031–5042, https://doi.org/10.1109/TII.2022.3146552, 2022.

Akilan, T., Wu, Q. J., Yang, Y., and Safaei, A.: Fusion of transfer learning features and its application in image classification, in: 2017 IEEE 30th Canadian Conf. on Elec. and Compu. Engineering, pp. 1–5, https://doi.org/10.1109/CCECE.2017.7946733, 2017.

Akilan, T., Wu, Q. J., and Zhang, H.: Effect of fusing features from multiple DCNN architectures in image classification, IET Image Processing, 12, 1102–1110, 2018.

Akilan, T., Wu, Q. J., Safaei, A., Huo, J., and Yang, Y.: A 3D CNN-LSTM-Based Image-to-Image Foreground Segmentation, IEEE Trans. Intelligent Transportation Systems, 21, 959–971, https://doi.org/10.1109/TITS.2019.2900426, 2020.

Al-Qaness, M. A., Ewees, A. A., Abualigah, L., AlRassas, A. M., Thanh, H. V., and Abd Elaziz, M.: Evaluating the applications of dendritic neuron model with metaheuristic optimization algorithms for crude-oil-production forecasting, Entropy, 24, 1674, 2022.

Alakeely, A. and Horne, R. N.: Simulating the behavior of reservoirs with convolutional and recurrent neural networks, SPE Reservoir Evaluation & Engineering, 23, 0992–1005, 2020.

Alamoudi, M., Sattari, M. A., Balubaid, M., Eftekhari-Zadeh, E., Nazemi, E., Taylan, O., and Kalmoun, E. M.: Application of Gamma Attenuation Technique and Artificial Intelligence to Detect Scale Thickness in Pipelines in Which Two-Phase Flows with Different Flow Regimes and Void Fractions Exist, Symmetry, 13, 1198, 2021.

Alanazi, A. K., Alizadeh, S. M., Nurgalieva, K. S., Nesic, S., Grimaldo Guerrero, J. W., Abo-Dief, H. M., Eftekhari-Zadeh, E., Nazemi, E., and Narozhnyy, I. M.: Application of neural network and time-domain feature extraction techniques for determining volumetric percentages and the type of two phase flow regimes independent of scale layer thickness, Applied Sciences, 12, 1336, 2022.

Amiri, S., Ali, P. J. M., Mohammed, S., Hanus, R., Abdulkareem, L., Alanezi, A. A., Eftekhari-Zadeh, E., Roshani, G. H., Nazemi, E., and Kalmoun, E. M.: Proposing a nondestructive and intelligent system for simultaneous determining flow regime and void fraction percentage of gas–liquid two phase flows using polychromatic X-ray transmission spectra, Journal of Nondestructive Evaluation, 40, 47, 2021.

Arps, J. J.: Analysis of decline curves, Transactions of the AIME, 160, 228–247, 1945.

Bahiraei, M., Mazaheri, N., and Hosseini, S.: Neural network modeling of thermo-hydraulic attributes and entropy generation of an ecofriendly nanofluid flow inside tubes equipped with novel rotary coaxial double-twisted tape, Powder technology, 369, 162–175, 2020.

Bahiraei, M., Foong, L. K., Hosseini, S., and Mazaheri, N.: Neural network combined with nature-inspired algorithms to estimate overall heat transfer coefficient of a ribbed triple-tube heat exchanger operating with a hybrid nanofluid, Measurement, 174, 108 967, 2021a.

Bahiraei, M., Foong, L. K., Hosseini, S., and Mazaheri, N.: Predicting heat transfer rate of a ribbed triple-tube heat exchanger working with nanofluid using neural network enhanced by advanced optimization algorithms, Powder Technology, 381, 459–476, 2021b.

Bao, A., Gildin, E., Huang, J., and Coutinho, E. J.: Data-driven end-to-end production prediction of oil reservoirs by enkf-enhanced recurrent neural networks, in: SPE latin american and caribbean petroleum engineering conference, OnePetro, 2020.

Basahel, A., Sattari, M. A., Taylan, O., and Nazemi, E.: Application of feature extraction and artificial intelligence techniques for increasing the accuracy of X-ray radiation based two phase flow meter, Mathematics, 9, 1227, 2021.

Bhanja, S. and Das, A.: Impact of data normalization on deep neural network for time series forecasting, arXiv:1812.05519, 2018.

Buettner, M. A., Monzen, N., and Hackl, C. M.: Artificial neural network based optimal feedforward torque control of interior permanent magnet synchronous machines: A feasibility study and comparison with the state-of-the-art, Energies, 15, 1838, 2022.

Chahar, J., Verma, J., Vyas, D., and Goyal, M.: Data-driven approach for hydrocarbon production forecasting using machine learning techniques, Journal of Petroleum Science and Engineering, 217, 110 757, 2022.

Chen, K., Zhou, Y., and Dai, F.: A LSTM-based method for stock returns prediction: A case study of China stock market, in: 2015 IEEE international conference on big data (big data), pp. 2823–2824, IEEE, 2015.

Chunguo, J. and Qiuguo, B.: Flow regime identification of gas/liquid two-phase flow in vertical pipe using RBF neural networks, in: 2009 Chinese Control and Decision Conference, pp. 5143–5147, IEEE, 2009.

de Freitas Dam, R. S., Salgado, W. L., Affonso, R. R. W., Schirru, R., and Salgado, C. M.: Optimization of radioactive particle tracking methodology in a single-phase flow using MCNP6 code and artificial intelligence methods, Flow Measurement and Instrumentation, 78, 101 862, 2021.

Equinor: Volve data village, Equinor Data Portal Beta, 2018a.

Equinor: Volve data village dataset: released under a license based on CC BY 4.0., 2018b.

Equinor: Disclosing all Volve data, https://www.equinor.com/news/archive/ 14jun2018-disclosing-volve-data/, [Online; accessed 19-July-2022], 2018c.

Ezugwu, A. E., Adeleke, O. J., Akinyelu, A. A., and Viriri, S.: A conceptual comparison of several metaheuristic algorithms on continuous optimisation problems, Neural Computing and Applications, 32, 6207–6251, 2020.

Faghihi, R., Nematollahi, M., Erfaninia, A., and Adineh, M.: Void fraction measurement in modeled two-phase flow inside a vertical pipe by using polyethylene phantoms, International Journal of Hydrogen Energy, 40, 15 206–15 212, 2015.

Gedlu, E. G., Wallscheid, O., and Böcker, J.: Permanent magnet synchronous machine temperature estimation using low-order lumped-parameter thermal network with extended iron loss model, in: The 10th International Conference on Power Electronics, Machines and Drives, vol. 2020, pp. 937–942, IET, 2020.

Graves, A. and Schmidhuber, J.: Framewise phoneme classification with bidirectional LSTM networks, in: Proceedings. 2005 IEEE Intl, Joint Conf. on Neural Networks, 2005., vol. 4, pp. 2047–2052, IEEE, 2005.

Grobler, A. J., Holm, S. R., and van Schoor, G.: A two-dimensional analytic thermal model for a high-speed PMSM magnet, IEEE Trans. Industrial Electronics, 62, 6756–6764, 2015.

Gui, N., Lou, J., Qiu, Z., and Gui, W.: Temporal feature selection for multi-step ahead reheater temperature prediction, Processes, 7, 473, 2019.

Hanus, R., Zych, M., Petryka, L., Jaszczur, M., and Hanus, P.: Signals features extraction in liquid-gas flow measurements using gamma densitometry. Part 1: time domain, in: EPJ Web of Conferences, vol. 114, p. 02035, EDP Sciences, 2016.

Hong, A., Bratvold, R. B., Lake, L. W., and Ruiz Maraggi, L. M.: Integrating model uncertainty in probabilistic decline-curve analysis for unconventional-oil-production forecasting, SPE Reservoir Evaluation & Engineering, 22, 861–876, 2019.

Hosseini, S., Roshani, G., and Setayeshi, S.: Precise gamma based two-phase flow meter using frequency feature extraction and only one detector, Flow Measurement and Instrumentation, 72, 101 693, 2020.

Hosseini, S., Setayeshi, S., Roshani, G., Zahedi, A., and Shama, F.: Increasing efficiency of two-phase flowmeters using frequency-domain feature extraction and neural network in the detector output spectrum, Journal of Modeling in Engineering, 19, 47–57, https://doi.org/10.22075/jme.2021.19817.1860, 2021a.

Hosseini, S., Taylan, O., Abusurrah, M., Akilan, T., Nazemi, E., Eftekhari-Zadeh, E., Bano, F., and Roshani, G. H.: Application of Wavelet Feature Extraction and Artificial Neural Networks for Improving the Performance of Gas–Liquid Two-Phase Flow Meters Used in Oil and Petrochemical Industries, Polymers, 13, 3647, 2021b.

Hosseini, S., Iliyasu, A. M., Akilan, T., Salama, A. S., Eftekhari-Zadeh, E., and Hirota, K.: Accurate Flow Regime Classification and Void Fraction Measurement in Two-Phase Flowmeters Using Frequency-Domain Feature Extraction and Neural Networks, Separations, 9, 160, 2022a.

Hosseini, S., Shahbandegan, A., and Akilan, T.: Deep Neural Network Modeling for Accurate Electric Motor Temperature Prediction, in: 2022 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 170–175, https://doi.org/10.1109/CCECE49351.2022.9918222, 2022b.

95

Jo, H., Hwang, H. J., Phan, D., Lee, Y., and Jang, H.: Endpoint temperature prediction model for LD converters using machine-learning techniques, in: 2019 IEEE 6th International Conference on Industrial Engineering and Applications (ICIEA), pp. 22–26, IEEE, 2019.

Jung, H.-S., Kim, H., Sul, S.-K., and Berry, D. J.: Temperature Estimation of IPMSM by Using Fundamental Reactive Energy Considering Variation of Inductances, IEEE Transactions on Power Electronics, 36, 5771–5783, https://doi.org/10.1109/TPEL.2020.3028084, 2021.

Kasburg, C. and Stefenon, S. F.: Deep learning for photovoltaic generation forecast in active solar trackers, IEEE Latin America Transactions, 17, 2013–2019, 2019.

Kirchgässner, W., Wallscheid, O., and Böcker, J.: Estimating Electric Motor Temperatures with Deep Residual Machine Learning, IEEE Trans. Power Electronics, 36, 7480–7488, 2020.

Kirchgässner, W., Wallscheid, O., and Böcker, J.: Data-Driven Permanent Magnet Temperature Estimation in Synchronous Motors with Supervised Machine Learning: A Benchmark, IEEE Trans. Energy Conversion, 36, 2059–2067, 2021.

Kirchgässner, W., Wallscheid, O., and Böcker, J.: Empirical Evaluation of Exponentially Weighted Moving Averages for Simple Linear Thermal Modeling of Permanent Magnet Synchronous Machines, in: 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE), pp. 318–323, https://doi.org/10.1109/ISIE.2019.8781195, 2019.

Kirchgässner, W., Wallscheid, O., and Böcker, J.: Data-Driven Permanent Magnet Temperature Estimation in Synchronous Motors With Supervised Machine Learning: A Benchmark, IEEE Transactions on Energy Conversion, 36, 2059–2067, https://doi.org/10.1109/TEC.2021.3052546, 2021.

Lee, J. and Ha, J.-I.: Temperature Estimation of PMSM Using a Difference-Estimating Feedforward Neural Network, IEEE Access, 8, 130 855–130 865, 2020.

Li, H.: Deep learning for natural language processing: advantages and challenges, National Science Review, 2017.

Li, J. and Akilan, T.: Global Attention-based Encoder-Decoder LSTM Model for Temperature Prediction of Permanent Magnet Synchronous Motors, https://arxiv.org/abs/2208.00293, 2022.

Li, Y., Sun, R., and Horne, R.: Deep learning for well data history analysis, in: SPE Annual Technical Conference and Exhibition, OnePetro, 2019.

Liu, H., Gu, J., Wang, Y., and Wei, Z.: Prediction Method of Heavy Oil Horizontal Well Cycle Oil Production Based on PCA and Gradient Boosting Decision Tree, in: 2021 3rd International Conference on Intelligent Control, Measurement and Signal Processing and Intelligent Oil Field (ICMSP), pp. 276–280, https://doi.org/10.1109/ICMSP53480.2021.9513392, 2021a.

Liu, W., Liu, W. D., and Gu, J.: Petroleum production forecasting based on machine learning, in: Proceedings of the 2019 3rd International Conference on Advances in Image Processing, pp. 124–128, 2019.

Liu, Y., Ma, X., Shu, L., Hancke, G. P., and Abu-Mahfouz, A. M.: From Industry 4.0 to Agriculture 4.0: Current Status, Enabling Technologies, and Research Challenges, IEEE Transactions on Industrial Informatics, 17, 4322–4334, https://doi.org/10.1109/TII.2020.3003910, 2021b.

MacKay, K., Lavoie, M., Bourlon, E., Atherton, E., O'Connell, E., Baillie, J., Fougère, C., and Risk, D.: Methane emissions from upstream oil and gas production in Canada are underestimated, Scientific reports, 11, 1–8, 2021.

Masini, S. R., Goswami, S., Kumar, A., and Chennakrishnan, B.: Decline curve analysis using artificial intelligence, in: Abu Dhabi International Petroleum Exhibition & Conference, OnePetro, 2019.

Mayet, A. M., Alizadeh, S. M., Nurgalieva, K. S., Hanus, R., Nazemi, E., and Narozhnyy, I. M.: Extraction of time-domain characteristics and selection of effective features using correlation analysis to increase the accuracy of petroleum fluid monitoring systems, Energies, 15, 1986, 2022a.

Mayet, A. M., Salama, A. S., Alizadeh, S. M., Nesic, S., Guerrero, J. W. G., Eftekhari-Zadeh, E., Nazemi, E., and Iliyasu, A. M.: Applying data mining and artificial intelligence techniques for high precision measuring of the two-phase flow's characteristics independent of the pipe's scale layer, Electronics, 11, 459, 2022b.

Nazemi, E., Feghhi, S., Roshani, G., Setayeshi, S., and Peyvandi, R. G.: A radiation-based hydrocarbon two-phase flow meter for estimating of phase fraction independent of liquid phase density in stratified regime, Flow Measurement and Instrumentation, 46, 25–32, 2015.

Nazemi, E., Feghhi, S., Roshani, G., Peyvandi, R. G., and Setayeshi, S.: Precise void fraction measurement in two-phase flows independent of the flow regime using gamma-ray attenuation, Nuclear Engineering and Technology, 48, 64–71, 2016a.

Nazemi, E., Roshani, G., Feghhi, S., Setayeshi, S., Zadeh, E. E., and Fatehi, A.: Optimization of a method for identifying the flow regime and measuring void fraction in a broad beam gamma-ray attenuation technique, international journal of hydrogen energy, 41, 7438–7444, 2016b.

Nazemi, E., Feghhi, S., Roshani, G., Peyvandi, R. G., and Setayeshi, S.: Precise void fraction measurement in two-phase flows independent of the flow regime using gamma-ray attenuation, Nuclear Engineering and Technology, 48, 64–71, 2019.

Ng, C. S. W., Ghahfarokhi, A. J., and Amar, M. N.: Well production forecast in Volve field: Application of rigorous machine learning techniques and metaheuristic algorithm, Journal of Petroleum Science and Engineering, 208, 109 468, 2022.

Ning, Y., Kazemi, H., and Tahmasebi, P.: A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet, Computers & Geosciences, 164, 105 126, 2022.

Noshi, C. I., Eissa, M. R., and Abdalla, R. M.: An intelligent data driven approach for production prediction, in: Offshore Technology Conference, OnePetro, 2019.

Ozbayoglu, E., Ozbayoglu, M., Ozdilli, B. G., and Erge, O.: Optimization of flow rate and pipe rotation speed considering effective cuttings transport using data-driven models, Energies, 14, 1484, 2021.

Panja, P., Velasco, R., Pathak, M., and Deo, M.: Application of artificial intelligence to forecast hydrocarbon production from shales, Petroleum, 4, 75–89, 2018.

Rad, M. Y. and Shahbandegan, S.: An Intelligent Algorithm for Mapping of Applications on Parallel Reconfigurable Systems, in: 2020 6th Iranian Conf. on Sig. Process. and Intelli. Sys. (ICSPIS), pp. 1–6, https://doi.org/10.1109/ICSPIS51611.2020.9349558, 2020.

Roshani, G., Nazemi, E., Feghhi, S., and Setayeshi, S.: Flow regime identification and void fraction prediction in two-phase flows based on gamma ray attenuation, Measurement, 62, 25–32, 2015.

Roshani, G., Nazemi, E., and Feghhi, S.: Investigation of using 60Co source and one detector for determining the flow regime and void fraction in gas–liquid two-phase flows, Flow Measurement and Instrumentation, 50, 73–79, 2016.

Roshani, M., Phan, G., Faraj, R. H., Phan, N.-H., Roshani, G. H., Nazemi, B., Corniani, E., and Nazemi, E.: Proposing a gamma radiation based intelligent system for simultaneous analyzing and detecting type and amount of petroleum by-products, Nuclear Engineering and Technology, 53, 1277–1283, 2021a.

Roshani, M., Phan, G., Roshani, G. H., Hanus, R., Nazemi, B., Corniani, E., and Nazemi, E.: Combination of X-ray tube and GMDH neural network as a nondestructive and potential technique for measuring characteristics of gas-oil–water three phase flows, Measurement, 168, 108 427, 2021b.

Sadjadi, S. and Sadi-Nezhad, S.: Ranking Canadian oil and gas projects using TOPSIS, Journal of Project Management, 2, 87–92, 2017.

Salgado, C., Dam, R., Salgado, W., Werneck, R., Pereira, C., and Schirru, R.: The comparison of different multilayer perceptron and General Regression Neural Networks for volume fraction prediction using MCNPX code, Applied Radiation and Isotopes, 162, 109 170, 2020.

Sattari, M. A., Roshani, G. H., Hanus, R., and Nazemi, E.: Applicability of time-domain feature extraction methods and artificial intelligence in two-phase flow meters based on gamma-ray absorption technique, Measurement, 168, 108 474, 2021.

Schenke, M. and Wallscheid, O.: A Deep Q-Learning Direct Torque Controller for Permanent Magnet Synchronous Motors, IEEE Open Journal of the Industrial Electronics Society, 2, 388–400, https://doi.org/10.1109/OJIES.2021.3075521, 2021.

Su, X., Zhang, S., Yin, Y., and Xiao, W.: Prediction model of hot metal temperature for blast furnace based on improved multi-layer extreme learning machine, International Journal of Machine Learning and Cybernetics, 10, 2739–2752, 2019.

Syed, F. I., Alshamsi, M., Dahaghi, A. K., and Neghabhan, S.: Artificial lift system optimization using machine learning applications, Petroleum, 2020.

Veisi, A., Shahsavari, M. H., Roshani, G. H., Eftekhari-Zadeh, E., and Nazemi, E.: Experimental Study of Void Fraction Measurement Using a Capacitance-Based Sensor and ANN in Two-Phase Annular Regimes for Different Fluids, Axioms, 12, 66, 2023.

Wallscheid, O., Kirchgässner, W., and Böcker, J.: Investigation of long short-term memory networks to temperature prediction for permanent magnet synchronous motors, in: Intl. joint conf. on neural networks, pp. 1940–1947, IEEE, 2017a.

Wallscheid, O., Specht, A., and Böcker, J.: Observing the permanent-magnet temperature of synchronous motors based on electrical fundamental wave model quantities, IEEE Trans. Industrial Electronics, 64, 3921–3929, 2017b.

Wang, B., Sharma, J., Chen, J., and Persaud, P.: Ensemble Machine Learning Assisted Reservoir Characterization Using Field Production Data–An Offshore Field Case Study, Energies, 14, 1052, 2021a.

Wang, X.: Ladle furnace temperature prediction model based on large-scale data with random forest, IEEE/CAA Journal of Automatica Sinica, 4, 770–774, 2016.

Wang, Z., Tian, J., Fang, H., Chen, L., and Qin, J.: LightLog: A lightweight temporal convolutional network for log anomaly detection on the edge, Computer Networks, p. 108616, 2021b.

Xu, L. D., He, W., and Li, S.: Internet of Things in Industries: A Survey, IEEE Transactions on Industrial Informatics, 10, 2233–2243, https://doi.org/10.1109/TII.2014.2300753, 2014.

Yang, Y., Wu, Q. M. J., Feng, X., and Akilan, T.: Recomputation of the Dense Layers for Performance Improvement of DCNN, IEEE Trans. on Pattern Analy. and Mach. Intelli., 42, 2912–2925, https://doi.org/10.1109/TPAMI.2019.2917685, 2020.

Zanjani, M. S., Salam, M. A., and Kandara, O.: Data-driven hydrocarbon production forecasting using machine learning techniques, International Journal of Computer Science and Information Security (IJCSIS), 18, 65–72, 2020.

Zhang, H., Liang, Y., Zhang, W., Xu, N., Guo, Z., and Wu, G.: Improved PSO-Based Method for Leak Detection and Localization in Liquid Pipelines, IEEE Transactions on Industrial Informatics, 14, 3143–3154, https://doi.org/10.1109/TII.2018.2794987, 2018.

Zhao, J., Mao, X., and Chen, L.: Speech emotion recognition using deep 1D & 2D CNN LSTM networks, Biomed. sig. process. and cont., 47, 312–323, 2019.

Zhukov, A., Tomin, N., Kurbatsky, V., Sidorov, D., Panasetsky, D., and Foley, A.: Ensemble methods of classification for power systems security assessment, App. comp. & info., 15, 45–53, 2019.

Chapter 7

Appendix

7.1   Appendix A: IEEE Permission to Reprint

In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of Lakehead University's products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link. html and https://www.ieee.org/publications/rights/author-rights-responsibilities.html to learn how to obtain a License from RightsLink.

7.2 Appendix B: Elsevier Permission to Reprint

In reference to Elsevier copyrighted material which is used with permission in this thesis, the Elsevier does not endorse any of Lakehead University's products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing Elsevier copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights.

7.3   Appendix C: MDPI Permission to Reprint

In reference to MDPI copyrighted material which is used with permission in this thesis, the MDPI does not endorse any of Lakehead University's products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing MDPI copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please visit https://www.mdpi.com/authors/rights.

7.4 Appendix D: Source Code

The source codes of this thesis are available on GitHub.

For more information about the author's publications, please refer to Google Scholar and LinkedIn profiles.

Google Scholar: Google Scholar.

LinkedIn: LinkedIn.