ARTIFICIAL NEURAL NETWORK BASED THERMAL CONDUCTIVITY

PREDICTION OF PROPYLENE GLYCOL SOLUTIONS WITH

REAL TIME HEAT PROPAGATION APPROACH

Andrew Jarrett, B.E.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

August 2022

APPROVED:

Tae-Youl Choi, Major Professor
Huseyin Bostanci, Committee Member
Hamid Sadat, Committee Member
Herman Shen, Chair of the Department of
    Mechanical Engineering
Shengli Fu, Interim Dean of the College of
    Engineering
Victor Prybutok, Dean of the Toulouse
    Graduate School

Jarrett, Andrew. *Artificial Neural Network Based Thermal Conductivity Prediction of Propylene Glycol Solutions with Real Time Heat Propagation Approach*. Master of Science (Mechanical and Energy Engineering), August 2022, 48 pp., 7 tables, 26 figures, 46 numbered references.

Machine learning is fast growing field as it can be applied to solve a large amount of problems. One large subsection of machine learning are artificial neural networks (ANN), these work on pattern recognition and can be trained with data sets of known solutions. The objective of this thesis is to discuss the creation of an ANN capable of classifying differences in propylene glycol concentrations, up to 10%. Utilizing a micro pipette thermal sensor (MTS) it is possible to measure the heat propagation of a liquid from a laser pulse. The ANN can then be trained beforehand with simulated data and be tested in real time with temperature data from the MTS. This method could be applied to find the thermal conductivity of unknown fluids and biological samples, such as cells and tissues.

TABLE OF CONTENTS

LIST OF FIGURES

Page

NOMENCLATURE

ANN – Artificial neural network

C – Specific heat (kJ/Kg*K)

DSC – Differential scanning calorimetry

$\boldsymbol{d}_k$ – Conjugate direction

$E(\theta_k)$ – Cost function

FN – False negative

FP – False positive

h – Contact thermal conductance (W/(m^2*K))

H_k – Hessian matrix

k – Thermal conductivity (W/m*K)

L_{CE} – Cross entropy loss function

LPHT – Laser point heating thermometry

MEMS – Micro electro mechanical systems

ML – Machine learning

MTS – Micropipette thermal sensor

N – Number of samples

$q_x"$ – Heat flux (W/m)

S – Seebeck coefficient (µV/K)

SThM – Scanning thermal microscopy

t_i – Truth label

TN – True negative

TP – True positive

∇T – Temperature gradient

ΔT – Change of temperature (K)

ΔV – Change of voltage (V)

x – Length (m)

r – Radius (m)

ρ – Density (Kg/m^3)

θ_k – Weight of perception

α – Learning rate

γ_k – Step size

x_i – Observed value

$\hat{x}_i$ – Expected value

p_i – Probability

CHAPTER 1

INTRODUCTION

Measuring the thermal properties of biological samples is a rapidly growing field of study due to fact that thermal energy is related to all biological reactions inside of the cell [1] [2]. One of the most essential of these properties is thermal conductivity as it is the governing parameter in all heat conduction and convection problems in engineering. It has been shown to also be a viable way to evaluate the viability of cells [3]. Along with this thermal conductivity has been utilized as a biomarker to measure the disease state of cancer, like the proliferation index in which tumor progression is assessed [4]. Therefore, a way to measure this thermal property is needed on a micro scale that is also safe for biological applications. Several methods for thermometry have been developed for the measurement of single cells. These include electron spin from nitrogen vacancies in diamond nanoparticles and fluorescent nanothermometers utilizing nanoparticles [5] [6]. However, these methods are complicated and costly. It has been shown that a simple method for measuring the thermal conductivity is utilizing a micropipette thermal sensor (MTS) and laser point heating thermometry (LPTH) [7]. While this method is effective it has a high computational cost because of the multiparameter fitting program used by LPTH to solve for the thermal conductivity [8]. Machine learning (ML) serves as an alternative to numerical analysis [9].

ML works on the concept of pattern recognition [10]. It utilizes mathematical minimization techniques to process large amounts of data in order to find the path which results in the least amount of error or cost. ML is used in many applications such as computer vision, speech recognition, and game play [11]. Different methods of ML have

been created such as linear regression, decision tree, and support vector machines. This study focuses on artificial neural networks (ANN). Based on a biological brain ANNs are a connection of nodes with weighted values placed on the synapses [12]. These weights are trained by a process of minimization known as gradient descent. Once trained ANNs can solve complex problems instantaneously.

ANNs have been shown to be used for a variety of thermal characterization related tasks, including the prediction of hybrid nanofluids and ethylene glycol thermal conductivities [13] [14]. These both utilize a multi-input and single output regression learning model, where the inputs are the concentration, density, and temperature of the fluids. The output is the thermal conductivity of the fluid.

This study will focus on a real time approach for the thermal conductivity prediction of a glycol solution measuring the time-series heat propagation profile. With this an ANN trained with the time-series temperature profiles of known thermal conductivities can be proposed to predict parameters (i.e., thermal conductivity) of a target chemical or a biological system including liquid or a biological cell. To that end, we first obtain massive heat propagation profiles using the partial differential equation (3.1) varying the thermal conductivity and train an ANN model for Sim-to-Real approach. Once trained, classification is instantaneous thereby solving the issues of computation time. Since the AI model is trained by the time-series heat propagation data, the proposed approach is unique and useful for real-time physical and biological property measurement for time-critical medical applications, in-situ biological screening, or for real-time physiological metabolism analysis.

1.1   Objectives

The primary objective for this research is to develop an ANN capable of classifying a 10% difference in concentration of propylene glycol utilizing time series temperature profiles. Along with this specific objectives are as follows:

- Train the ANN with simulated data for high classification accuracy

- Demonstrate the possibility of a simulation to real approach for ANN classification of thermal conductivity

- Produce a micro-pipette thermal sensor using nickel and bismuth wire

- Measure the experimental temperature profiles of propylene glycol mixtures and utilize trained ANN for classification

1.2   Organization of Thesis

Chapter 2 presents the literature research. The chapter starts by describing the thermal conductivity measurement technique, focusing on the LPHT that is utilized with the thermocouple. Along with these details over the thermocouple circuit is covered. Lastly ML techniques are discussed, covering ANNs and types of ML problems. This section also covers gradient descent and cost functions utilized in this research.

Chapter 3 covers the materials and method utilized in the experiment and ANN development. First the process for the fabrication and calibration of the MTS is given. Next a description of the ANN configuration is given as well as the training process used to optimize the network.

Chapter 4 discusses the results from calibration of the MTS, results from training, and the experimental results along with ANN classification results.

Chapter 5 concludes the thesis with a summary and provides the future work for this research.

CHAPTER 2

LITERATURE REVIEW

2.1   Heat Conduction

The transfer of heat energy from the interaction of molecules is known as conduction. Heat energy moves from more excited molecules to less excited molecules. Therefore, at higher temperatures the rotational and vibrational motions of the molecules are elevated. These excited molecules then collide with neighboring molecules and their energy is transferred [15].

This heat transfer process can be quantified with rate equations. Conduction follows Fouries law, where heat flux (q_x''), heat energy per unit area, is equal to the thermal conductivity (k) multiplied by the negative of the gradient of temperature (∇T). One dimensional heat transfer formula can be seen in equation (2.1) below.

$$q_x'' = -k\frac{dT}{dx} \tag{2.1}$$

The equation is very similar in cylindrical coordinates.

$$q_r'' = -k\frac{dT}{dr} \tag{2.2}$$

2.2   Thermocouple Method

Thermocouples are sensors that measure temperature. It is a junction of two different metals, and when this junction is heated, or cooled voltage is generated. This is known as the Seebeck effect, discovered in 1821 by Thomas Seebeck. [16] Change in voltage with respect to the change in temperature is the Seebeck coefficient, and represents the thermoelectric sensitivity of the metal. This is shown in equation (2.3) below [17].

$$\Delta T = S * \Delta V \tag{2.3}$$

where: ΔT = Change of Temperature.

S = Seebeck Coefficient

ΔV = Change of Voltage

2.3   Thermocouple Circuit

In order to measure the change in temperature from the thermocouple a circuit must be created. This is shown in figure 1. The two metals are joined at the ΔT_{sense} end and are connected to copper leads a the ΔT_{ref}. After this the copper leads are connected to a voltmeter in order to register the change of voltage. In this circuit the ΔT_{ref} must be a constant known temperature. This is because the Seebeck effects works on the change of temperature. Therefore, in order to solve for ΔT_{sense} the reference temperature must be known.
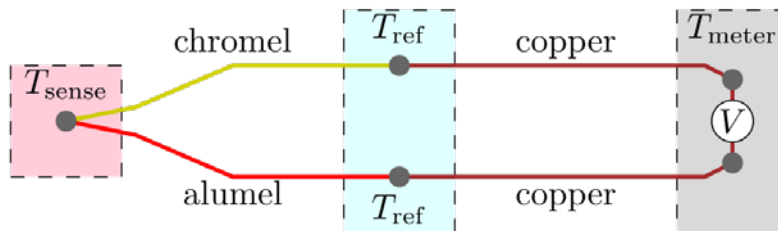
Figure 1: Image of thermocouple circuit example

The Seebeck for chromium, aluminum, and copper will be denoted by S_{Cr} , S_{Al} , and S_{Cu} respectively. [18] The voltage will then be written as:

$$V = \int_{Meter}^{Ref} S_{Cu}(T)\frac{dT}{dx}dx + \int_{Ref}^{Sense} S_{Cr}(T)\frac{dT}{dx}dx + \int_{Sense}^{Ref} S_{Al}(T)\frac{dT}{dx}dx +$$

$$\int_{Ref}^{Meter} S_{Cu}(T)\frac{dT}{dx}dx \tag{2.4}$$

It can be seen that this will simplify to the equation:

$$V = \int_{\text{Ref}}^{\text{Sense}} S_{\text{Cr}}\,(T)\frac{dT}{dx}\,dx + \int_{\text{Sense}}^{\text{Ref}} S_{\text{Al}}\,(T)\frac{dT}{dx}\,dx \qquad (2.5)$$

The Seebeck is mostly constant for small temperature ranges, and the combined Seebeck

$$V = (S_{cr} - S_{Al})\left(T_{sense} - T_{Ref}\right) \qquad (2.6)$$

$$T_{sense} = T_{ref} + \frac{V}{S_{Cr-Al}} \qquad (2.7)$$

2.4   Thermal Analysis Techniques

Thermal analysis is technique used to measure the temperature and time while a substance is heated or is cooled and to observe the physical changes that occur [19]. It is often used in heat transfer and thermodynamics applications to find material properties such as thermal conductivity, thermal diffusivity, and specific heat. In order to measure the thermal conductivity as proposed in this study two elements are needed a heat source and sensor. The most standard way to measure the thermal conductivity is the hot wire technique [20]. This is where a thin metal wire is submerged in a fluid and heated electrically while simultaneously monitoring the resistance variation. One application of this work is the ability to use this in a biological environment. Several thermal-analysis techniques have been used for the measurement of thermal properties of biological samples, such as differential scanning calorimetry (DSC), micro electro mechanical systems (MEMS), scanning thermal microscopy (SThM), laser point heating thermometry (LPHT) [21] [22] [23] [24] [7].

2.4.1   Differential Scanning Calorimetry

DSC is a technique used to determine the difference in heat energy required to increase the temperature of a sample and a reference. It is measured as a function of

specific heat vs. the temperature. It is used commonly in biochemical reactions to find transition temperature points. The method is done by monitoring the heat flow into the sample and reference and plotting with respect to temperature [25].
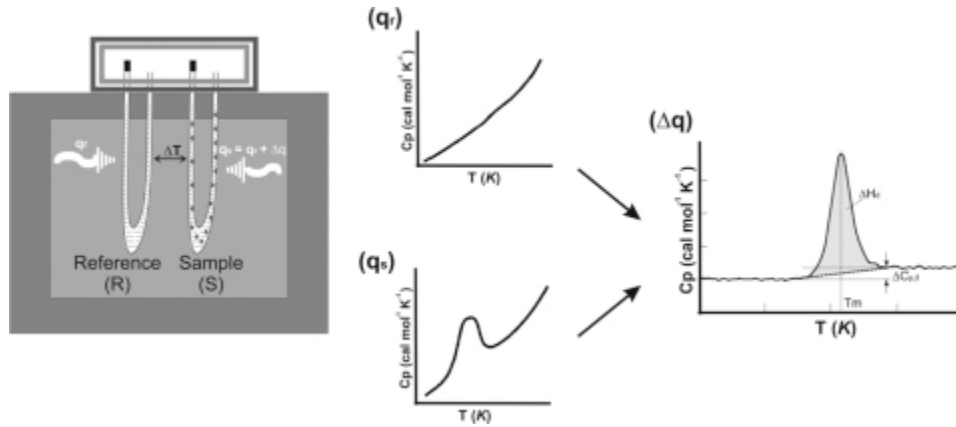
Figure 2: Experimental setup for a DSC experiment

2.4.2 Micro Electro Mechanical Systems

MEMS are miniaturized mechanical and electromechanical single devices and structures produced by microfabrication techniques. They are currently uses in a wide variety of microsensors, micro-actuators, and microsystems [26].
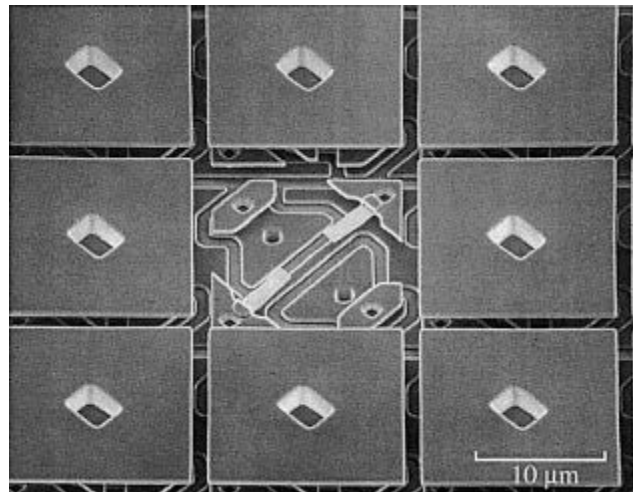
Figure 3: Detail of micromachined mirrors from an array used in a portable digital projector [27]

2.4.3 Scanning Thermal Microscopy

SthM is a combination of an atomic force microscope (AFM) and a temperature sensor such as a thermocouple or bolometer. AFM is where a nanosized tip is placed onto a cantilever beam attached to a piezoelectric device, while moving across a sample a laser is reflecting off the tip into a photoelectric sensor. This also the machine to detect deflections into the beam and correct with the piezoelectric device to avoid damaging the samples, as well as create a topological map of the sample. The temperature sensor is mounted on a cantilever beam which is attached to a piezoelectric device. Heat is supplied to the tip of the temperature sensor and its temperature is constantly monitored. When the sample moves the temperature of the tip will change based on the sample's thermal properties [28].
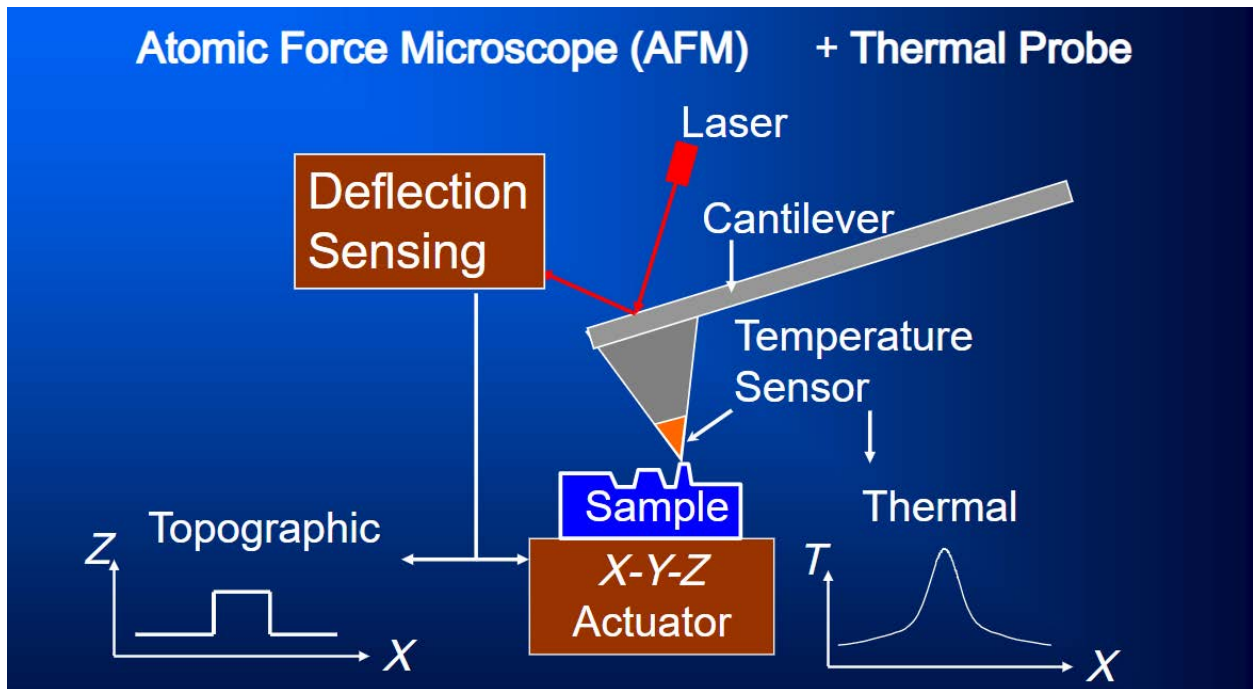
Figure 4: Scanning thermal microscope diagram [29].

2.4.4 Laser Point Heating

The technique used in this study is LPHT. [30] The main concept is using an optical point heating of a thermocouple tip. This in consequence causes a small amount of fluid around the thermocouple to be heated as well as simultaneously monitoring the fluids thermal response.
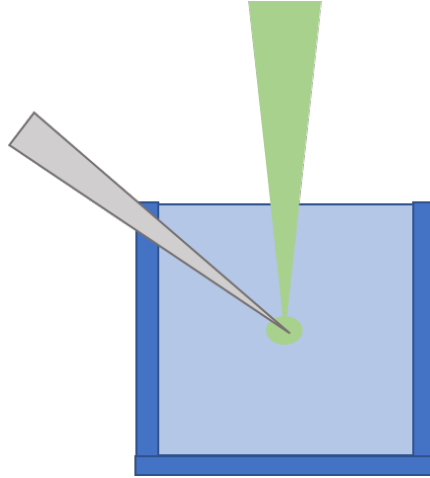
Figure 5: LPHT diagram, in this the grey shape is the thermocouple, green is the laser light, and blue is the fluid being measured

The equation below analytically expresses the temperature of thermocouple, assuming no heat loss through the thermocouple wires. [31]

$$T = \frac{q}{4\pi ak} * \left\{ \frac{1 + ac_1}{c_1} - \frac{2a^2 c_1^2 c_2^2}{\pi} \times \int_0^\infty \frac{e^{-\alpha u^2 t/a^2}}{[u^2(1 + ac_1) - ac_1 c_2]^2 + [u^3 - uac_1 c_2]^2} \right\}$$

$$c_1 = \frac{h}{k}, \qquad c_2 = \frac{4\pi a^3 \rho C}{M_s C_s} \tag{2.8}$$

where in this a, C_s, and M_s are the size, heat capacity, and mass of the TC tip, respectively, and C, k, and ρ are the heat capacity, thermal conductivity and density of the fluid sample, respectively. (q: heat transfer rate, h: contact thermal conductance at the TC tip – fluid.

The equation is a simple analytical approach to find the thermophysical

properties of the fluid. However, it does not account for the inevitable heat loss in the

TC wire. Therefore, a new approach is proposed using a FEM (finite element method) to

obtain simulated profiles and these profiles are used to train an Artificial Neural Network

for classification.

2.5   Machine Learning Method

Machine learning is a subset of artificial learning, where algorithms are improved

automatically through experience by using large data sets [32]. There are three different

approaches used in machine learning depending on the data that is used in the learning

system:

- Supervised learning: The model is presented with inputs and their outputs; its goal is to map the input to the output.

- Unsupervised learning: No output is given to the model; it then finds hidden patterns in the data.

- Reinforcement learning: The program interacts in each environment to perform a given goal or output. As the programs navigates its environment it is given feedback that it tries to maximize.

Once the approach is chosen these algorithms build models based on the

sample or training data. Similar research to ours was done over thermal properties

classification using a supervised artificial neural network approach [33].

2.5.1   Neural Networks

One model used in supervised learning is an artificial neural network (ANN). The

model is composed of connected artificial neurons (or perceptions) that are like

biological neurons in animal brains. Each perception has a weight and activation

function. This activation function is a type of normalization where the output is a number

between zero and one. ANNs have three different layers. The first is the input layer

where the inputs are fed into the network through input nodes. Next it goes into the

hidden layer where inputs are multiplied by an associated weight and summed together

at the node before the activation function is applied. The final layer is the output, where

it undergoes the same process as the hidden layer. Once it passes through the output

layer the result is obtained [12].
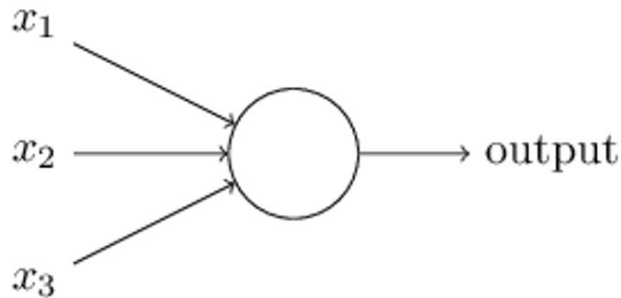
Figure 6: Image example of neuron for an ANN

The model is trained by a method known as gradient descent, where the error of

the model is used to adjust the weights inside the model. The error is calculated by a

cost function such as root mean square error (RMSE), and the partial derivative or

gradient $\left(\frac{\partial E}{\partial W_1}\right)$ is found with respect to each weight. This is multiplied by a

hyperparameter called the learning rate (α) and is subtracted from the weight

$(\theta_{Previous})$. This is done until the model has reached an acceptable margin of error [34].

This process in machine learning is called backpropagation.

$$\theta_{update}^{(1)} = \theta_{Previous}^{(1)} - \alpha \left(\frac{\partial E}{\partial \theta_1}\right) \qquad (2.9)$$

$$\theta_{update}^{(2)} = \theta_{Previous}^{(2)} - \alpha \left(\frac{\partial E}{\partial \theta_2}\right) \qquad (2.10)$$

2.5.2 Gradient Descent Methods

There are several variants of traditional gradient descent such as batch gradient

descent, mini-batch gradient descent, stochastic gradient descent, and conjugate gradient method [35]. Batch gradient descent finds the gradient of the cost function for the entire training set. This however can be ineffective as it takes longer to train because it requires a smaller learning rate [36]. Mini-batch takes the gradient of n training samples. Stochastic gradient descent takes a different approach as it finds the gradient of each training example. These three methods share a common method as they only take the first derivative of the cost function. Conjugate gradient is a second order technique, meaning that it takes the second derivative of the cost function [37]. Second order techniques will usually find a better path to the local minimum then first order methods but use higher computational power. These four methods all share the same challenge as they utilize the same learning rate no matter the size gradient. Therefore, a fifth method is presented scaled conjugate gradient method (SCGM). In SCGM the learning rate is varied based on the slope of the gradient [38]. Therefore, if the gradient is large, the learning rate increases and decreases if the gradient is small. This allows for faster and more accurate learning when compared to traditional gradient descent in which the learning rate is constant [39].

$$\theta_{k+1} = \theta_k - g_k H_k^{-1} \tag{2.11}$$

$$g_k = E'(\theta_k) \tag{2.12}$$

$$H_k = E''(\theta_k) \tag{2.13}$$

where, θ_k is the weight of the perception and $E(\theta_k)$ is the cost function.

In this H_k is the Hessian matrix which is a matrix of second order partial derivatives of the cost function. In equation (2.11) $g_k H_k^{-1}$ is defined as a newton step [40]. This newton step is impossible in large scale with multiple dimensions. Therefore,

Moller defined a temporal step $\theta_{t,k}$ which is between θ_{k+1} and θ_k. This is equation (2.14) below, where γ_k is the step size and $\boldsymbol{d}_k$ is the conjugate direction.

$$\theta_{t,k} = \theta_k + \gamma_k \boldsymbol{d}_k, \quad where\ 0 < \gamma_k \ll 1 \tag{2.14}$$

The conjugate direction is the minimization of all search directions. Thereby it means that one would find the local minimum for that iteration in the number of dimensions or search directions that are associated with the problem.

With this the newton step can be found using the approximation found in equation 2.15. s_k represents the second order information and α_k represents the learning rate.

$$s_k = E''(\theta_k)\boldsymbol{d}_k \approx \frac{E'(\theta_{t,k}) - E'(\theta_k)}{\gamma_k} \tag{2.15}$$

$$\alpha_k = \frac{-\boldsymbol{d}_k^T E'(\theta_1)}{-\boldsymbol{d}_k^T s_k} \tag{2.16}$$

Thus, the next weight value can be found using equation (2.17).

$$\theta_{k+1} = \theta_k - \alpha_k \boldsymbol{d}_k \tag{2.17}$$

2.5.2.1   Cost Functions

Just as with gradient descent there are few methods to find the cost or error of the network. One of the most common is RMSE [41]. This is the standard deviation of the errors. It can be seen below in equation (2.18) where x_i, $\hat{x}_i$, and N represent the observed value, expected value, and number of samples respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N}} \tag{2.18}$$

The cost function that was used to represent uncertainty in this study is the cross-entropy loss function (CELF) [42]. Before this function is discussed the concept of

entropy in machine learning must be covered. Entropy or $H(x)$ is the amount of

uncertainty given from a random variable across all the possible outcomes [43]. It is

based upon the probability density function and an equation is shown below. Where in

this $p(x)$ is the probability of a random variable.

$$H(x) = -\sum_x p(x) * \log(p(x)) \tag{2.19}$$

This loss function is based upon the concept of entropy or the uncertainty in

possible outcomes. When the probability of the ANN classifying the correct output is

high the loss of the function is minimized. In this t_i is the true label associated with that

sample and p_i is the output for the class.

This can be seen in equation (2.20) below.

$$L_{CE} = -\sum_{i=1}^{N} t_i \log(p_i) \tag{2.20}$$

2.5.2.2   Normalization

One important aspect of machine learning is normalization of data or feature

scaling. This is done before data is inputted into the network. Normalization is the

process of setting the data into a common scale. Rescaling or min-max normalization

sets the scale of the data in a range in [0,1]. Since learning is based upon calculating

the cost, the cost between two features that are not on the same scale will not be

accurate. Feature scaling data will converge faster than data without [44].

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{2.21}$$

The min-max normalization equation is shown above where x' is the normalized

value and x is the current feature.

2.5.3 Types of Network Problems

Mentioned above in supervised learning the network is given an input and an output, then it adjusts the weights connected between the input and output layers until an error threshold is met. With this learning method there are two main types of problems that can be solved: Classification and Regression.

2.5.3.1 Classification

Classification deals with mapping input variables to discreate output variables. It requires a problem whose outputs can be put in separate classes. When there are two classes for outputs this is known as a binary classification problem, and when there are more than two classes this is known as a multiclassification problem. A simple example of a binary classification would be a picture being classified as "dog" or "not dog". Classification problems commonly find the probability of the input belonging to a certain class.

There are a few metrics to measure the ability of the network for a classification problem. These metrics can be calculated by a confusion matrix.

2.5.3.2 Regression

Regression is when a network maps input variables to continuous output variables. This continuous is most commonly an integer value corresponding to a quantity. It is required that the regression predicts the correct quantity. An example of a regression problem would be predicting the number of people at a coffee shop depending on the time of day.

2.5.4 Confusion Matrix

Classification networks can be evaluated using a confusion matrix. This is a visual way to observe the performance of the network. It shows the number of true positives (*TP*), true negatives (*TN*), false positives (*FP*), and false negatives (*FN*) produced by the network from the training, validation, and testing data. Most confusion matrixes are shown from binary machine learning models, meaning only two outputs. This simple example can be seen in figure 7. In this study the confusion matrix generated is from a multiclass model.

Figure 7: Binary classification matrix example

There are several metrics that can be utilized to evaluate the performance of a classification mode: Precision, Accuracy, and F1 score [45]. Precision represents the number of positive classifications the network returned that were positive. Recall indicates the number of positive samples that were correctly classified. The F1 score is the harmonic mean or the reciprocal mean of the precision and recall of the model.

$$Recall = \frac{TP}{TP+FN} \tag{2.22}$$

$$Precision = \frac{TP}{TP+FP} \tag{2.23}$$

$$F1\ Score = 2 * \frac{Precision*Recall}{Precision+Recall} = \left(\frac{2*TP}{2*TP+FP+FN}\right) \qquad (2.24)$$

In the case of the multiclass model, each of the metrics can be found by the individual classification or by the total *TP*, *FP*, and *FN* of the model. When the totals are used to calculate the recall and precision, the following F1 score is known as the micro F1 score. This can be seen in equation (2.25) where $T.Precision$ and $T.Recall$ is the total precision and total recall, respectively. The total means these are metrics that are calculated using *TP* of the entire matrix over the amount of data sets used in the matrix, because the *FP* and *FN* are considered equivalent. Therefore, in multiclass models the accuracy, precision, and micro F1 score are all equal. The macro F1 score can also be used, as it calculates the average of the individual classes F1 score. Therefore, it depends on each class F1 more than the overall accuracy of the network. Lastly the weighted F1 score can be found by using the total number of samples for each class and multiplying by their F1 score and dividing by the total number of samples.

$$Micro\ F1\ Score = 2 * \frac{T.Precision*T.Recall}{T.Precision+T.Recall} \qquad (2.25)$$

$$Macro\ F1\ Score = \frac{1}{N}\sum_{i=1}^{N} F1\ Score_i \qquad (2.26)$$

$$Weighted\ F1\ Score = \frac{\sum_{i=1}^{N}(w_i*F1Score_i)}{\sum_{i=1}^{N} w_i} \qquad (2.27)$$

CHAPTER 3

METHODOLOGY AND MATERIALS

This chapter focuses on the development of the neural network and the training

methods employed for creating the model. Along with this, details over the fabrication

and calibration of the micropipette thermal sensor used in the experimental thermal

conductivity measurements. Lastly a description of the experimental setup and

procedure for the measurement of the propylene glycol mixtures.

3.1   Micropipette Thermal Sensor

The sensor used for the experimental measurement was a thermocouple

fabricated in the laboratory. The two metals used were nickel and a bismuth tin solder

alloy. The melting point of the alloy was 270°C, thereby making this the maximum

measurement temperature in the research.

The nickel was applied via sputtering deposition, which is a physical vapor

deposition process (PVD) that is primarily used to create thin films onto a substrate. It

occurs when an ionized gas molecule displaces atoms from a target or specific material.

These displaced atoms from the target then bond at the atomic level to the substrate.

When the thin film of nickel bonds to the inner core of bismuth, it creates a

thermocouple junction. Now with the applied Seebeck effect discussed in section 1

temperature rise can be solved from the increase in voltage of the MTS.

3.1.1 Fabrication

The scale of the thermocouple was on the micro size. Thick wall borosilicate

glass 2mm in outer diameter, commonly used in biomedical applications to inject

biological cells, was heated and pulled by a pipette puller (P-97, Sutter Instrument) to

create a sensor tip between 1-2 microns in diameter.

Figure 8: P-97 Pulling machine

Once pulled the pipette was filled with a bismuth (Bi) alloy by an injection molding process in conjunction with localized heating of the material. This was accomplished by a combination of a small heating chamber and mechanical pressurization (pushing). The heating chamber was maintained at 270°C; this melts the metal inside of the chamber. Next a steel rod pushed the molten metal to the tip by a servo motor.

After this the excess metal needs to be grinded off from the tip of the pipette. This is important to create a good connection between the inner metal and the outer metal. This was done by the Microforge-Grinding Center MFG-5APT designed for beveling micropipettes with tip diameters between 0.1 and 50 µm, which was used to sharpen and smoothen the tip. A 2-axis micromanipulator consisting of an angle plate to clamp the pipette was used to keep the bevel angle at 90°. The control of the grinding

19

advancement of the pipette to the grinding surface was performed with use of fine

control knobs mounted on the manipulator. This process was controlled by observation

through a 20x microscope mounted on the Microforge-Grinding Center.

Figure 9: Pipette filling set up
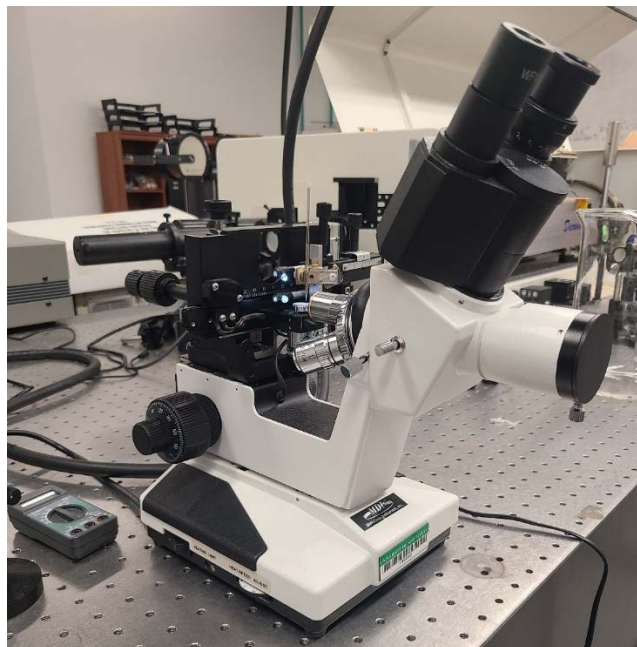
Figure 10: Microforge grinding center MFG-5APT

After the tip of the pipette was grinded to create a flat surface, it was sonicated to clean off the debris before sputtering. This was done with Sharptek ultrasonic cleaner Stamina XP. Once cleaned the pipette are ready for deposition of thin film. Pipettes are placed inside the sputtering vacuum chamber on a rotating plate, shown in figure 12. The chamber is then vacuumed down to 10^-6 Torr. Argon gas is pumped into the chamber in order to create the high energy plasma. 150 Watts of power is sent to the target, and the sputtering process is run for 35 min. This creates a coating thickness of 200-250 nm of nickel on the outside of the pipette.

Figure 11: Denton vacuum sputtering machine
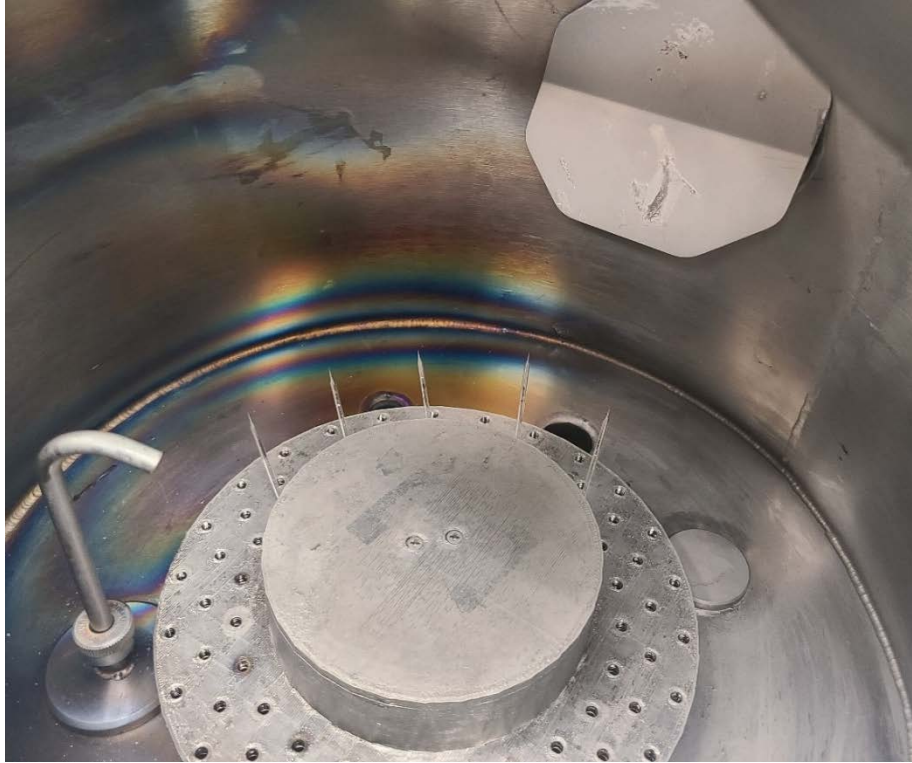
Figure 12: Pipettes placed inside of sputtering machine

Once the pipette is coated, copper wire is soldered to the inner core of bismuth. This is done by heating outside walls of the pipette with soldering iron and feeding wire into the open end of the borosilicate glass. Once cooled the MTS can be tested for resistance to ensure no short exists in the circuit. If there is a resistance calibration of the MTS can occur.
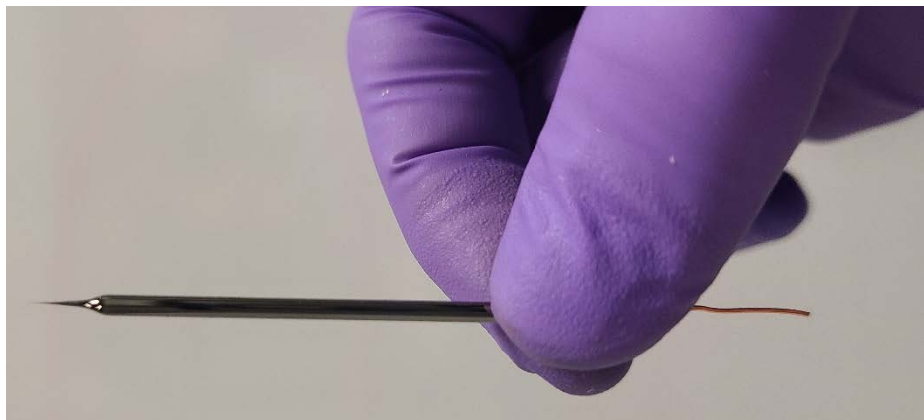
Figure 13: Completed MTS

3.1.2 Calibration

The sensor was calibrated using a thermally insulated heated chamber filled with

deionized water. The chamber was made of aluminum with an insulating lid of Teflon.

Schematics of both are shown in figure 14.

Temperature of chamber was controlled by a heater unit, with an accuracy of ± .1°C. The
chamber was heated from room temperature 21°C to a maximum of 40°C. A high-precision
digital thermometer and the MTS were submerged into the water and placed within a very close
vicinity to indicate the same temperature. The voltage of the MTS was measured by
oscilloscope (BK Precision 2190E) The voltage vs. the temperature were plotted and shown in
figures 17 and 18. The slope of this line is the Seebeck coefficient of the MTS.
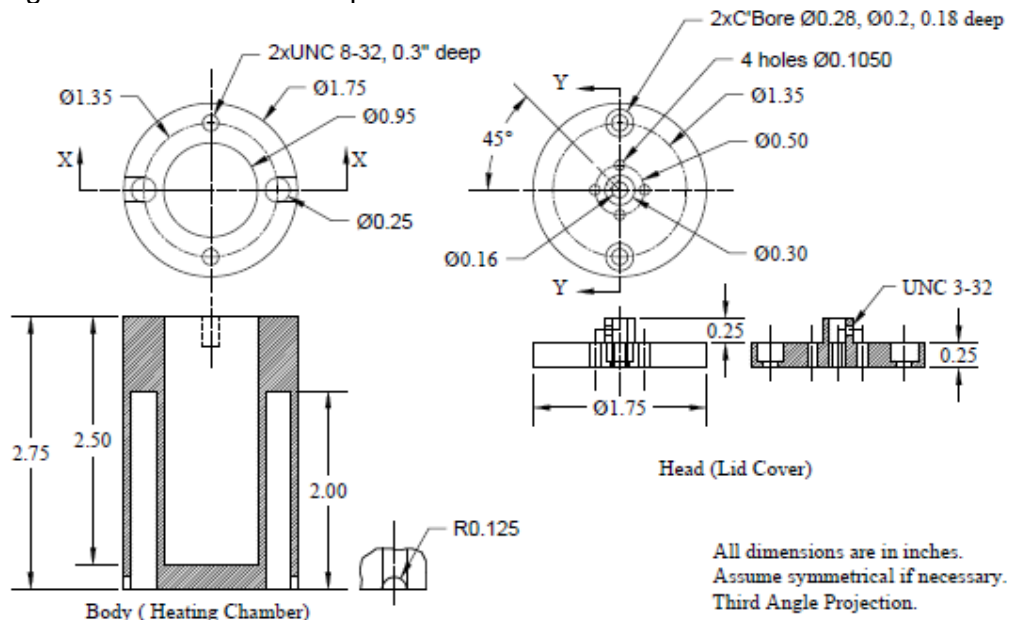
Figure 14: Schematic of calibration set up

3.2   Artificial Neural Network Model

The ANN model was created in MATLAB using the deep learning toolbox and

neural pattern recognition tool. This allows users to create a single hidden layer neural

network with specified inputs and outputs. The designed ANN was a classification

model. Input data for the model consisted of a temperature vector with respect to time.

The output was a class label representing a range of potential thermal conductivities.

3.2.1 Simulating Training Data

Since one of the objectives is to demonstrate the capability of a simulation to real approach for thermal conductivity classification, simulated data must be used to train the ANN. Therefore, a finite element model was created in order to produce the necessary data.
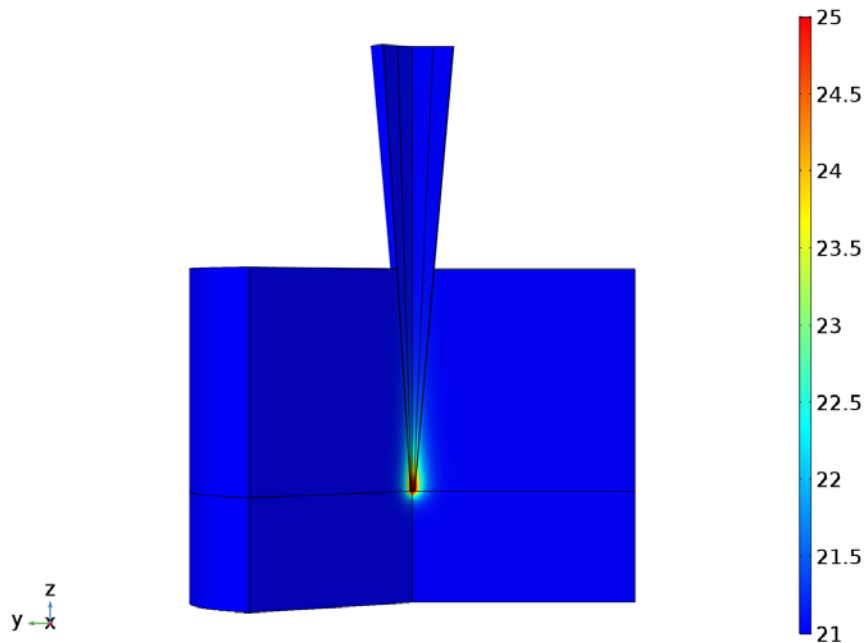
Figure 15: 3D cut away COMSOL multiphysics simulation of MTS with temperature scale

This data were generated using a Partial Differential Equation (PDE) Solver. COMSOL Multiphysics was chosen to calculate transient temperature profiles given a parameter - thermal conductivity for the model. A 2D axis model was made to show the MTS and propylene glycol mixture. Next a point source heating element was added to emulate a laser heating of the MTS. The parameters of this heating source were 100 mW for a 500 μs pulse. The temperature of the liquid adjacent to the MTS was sampled every 4 μs to get a vector of 126 temperature inputs. Evolution of this temperature was dependent on the thermal conductivity of the surrounding liquid, water and propylene

glycol in this thesis. The point at which temperature was sampled was important as with LPHT the temperature measured from the MTS isn't directly at the point of contact with liquid. Therefore, an analysis must be performed to find the optimal distance.

$$\rho c_p \frac{\partial T}{\partial x} + \nabla \cdot (-k\nabla T) = Q \qquad (3.1)$$

Concentration of propylene glycol in water directly impacts the mixture thermal properties. Thermal conductivity decreases with the increase in of glycol. Table 1 shows the distinction of 10% concentration and the thermal conductivity range.

3.2.2 Designing and Training the Network

3.2.2.1 Design

In this study the ANN utilized was a classification network that is trained with SCGD. With this the learning rate is scaled based off the gradient of the cost. In order to create the optimal network configuration, the cost function and hyperparameters must be chosen. In this study the cost function used was CELF. The number of inputs to the network is the 126 transient temperature profile. There were nine classes in which the temperature profile can be assigned corresponding to the thermal conductivity range of the glycol concentrations. The class output or classification label was a binary vector, this can be seen in table 1.

Table 1: Classification based on thermal conductivity range and glycol concentration

Glycol %	Thermal Conductivity Range	Classification Label
0% - 10%	.608 - .542	[1 0 0 0 0 0 0 0 0]
10% - 20%	.541 - .484	[0 1 0 0 0 0 0 0 0]
20% - 30%	.483 - 432	[0 0 1 0 0 0 0 0 0]
30% - 40%	.431 - .385	[0 0 0 1 0 0 0 0 0]
40% - 50%	.384 - .342	[0 0 0 0 1 0 0 0 0]

Glycol %	Thermal Conductivity Range	Classification Label
50% - 60%	.341 - .303	[0 0 0 0 0 1 0 0 0]
60% - 70%	.302 - .268	[0 0 0 0 0 0 1 0 0]
70% - 80%	.267 - .238	[0 0 0 0 0 0 0 1 0]
80% - 90%	.237 - .214	[0 0 0 0 0 0 0 0 1]

The last hyperparameter that the user can optimize in this network is the hidden layer size. Six different networks were created with hidden layer sizes ranging from 25 - 150 and compared to find the number of nodes that would yield the lowest error in classification. The micro and macro F1 score were calculated using each networks confusion matrix. Since each class had the same amount of data the Weighted F1 score would be equal to macro F1. A diagram of the network can be seen in figure 16 below.
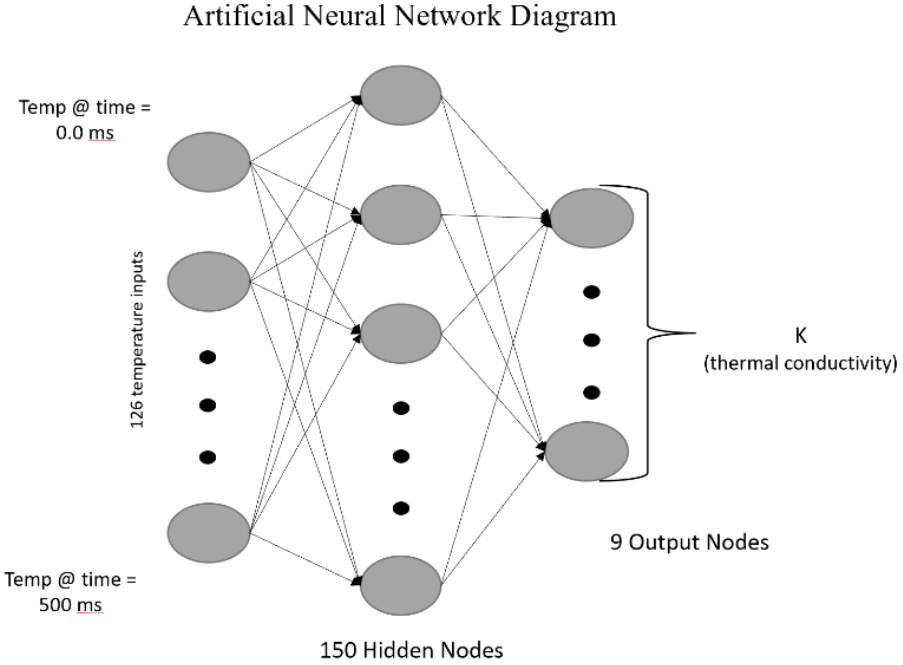
Figure 16: Neural network diagram, left side is the 126 temperature imputes connected by a system of weights to hidden layer. Lastly the hidden layer is connected to the output nodes that represent the thermal conductivity ranges.

3.2.2.2   Training

First before training the data was normalized by using the min-max method described in the literature review. The normalized data were randomized and separated into 70% training, 15% validation, and 15% testing data sets. Training data sets are used with optimization methods such as gradient descent. Validation sets provides a way to evaluate the model during training. This prevents overfitting of the data by early stopping. The final testing data is used to evaluate the trained model.
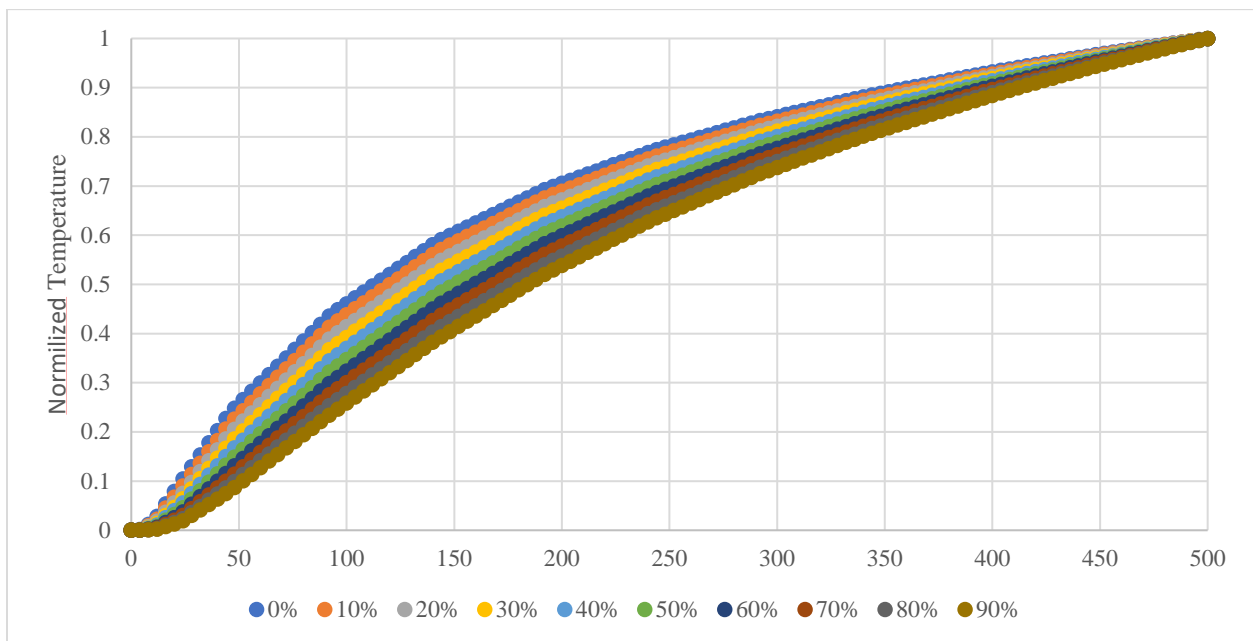
Figure 17: Normalized profiles of propylene glycol

During training the network utilizes the training set and performs SCGD until either the network reaches the performance goal, maximum epoch number, or validation failures. The performance goal is the gradient amount the developer wants to achieve in this study .005 was chosen. When training the network will continue to iterate until this goal is reached. Therefore, a maximum epoch or iteration number is chosen to stop training once this value is met. This was chosen to be 300000. Lastly while training

27

the network is evaluated with the validation set. In order to prevent divergence a number is chosen for the number of times the network can fail the evaluation; this was 300 in the current study.

Using the results from the training a classification matrix can be created. This as described in the literature review is used to evaluate the performance of the network. Once the network is trained newly simulated or experimental data can be imputed into the matrix, which results in instantaneous classification of the thermal conductivity.

3.3   Experimental Process

This next section will describe the experimental process followed to prepare the testing samples as well as the connection of the MTS to laser and oscilloscope.

3.3.1 Preparation of Propylene Glycol

In order to verify the ANN the transient temperature profiles of real propylene glycol water mixtures were collected, normalized, and inserted into the network. This experiment consisted of 6 different concentrations of propylene glycol, shown in table 2 below. Samples were prepared in 100 ml graduated cylinder and transferred to beaker for mixing. The water glycol samples were then mixed for 60 seconds via magnetic stirrer (Corning stirrer/Hot plate).

Table 2: Experimental propylene glycol mixtures. Row 1 shows the images of the petri dishes, row 2 shows the corresponding percent concentration, and row 3 shows the thermal conductivity.

	Dish 1	Dish 2	Dish 3	Dish 4	Dish 5	Dish 6
Glycol Concentration	90%	75%	50%	25%	10%	0%
Thermal Conductivity	.214	.2525	.341	.4575	.541	.608
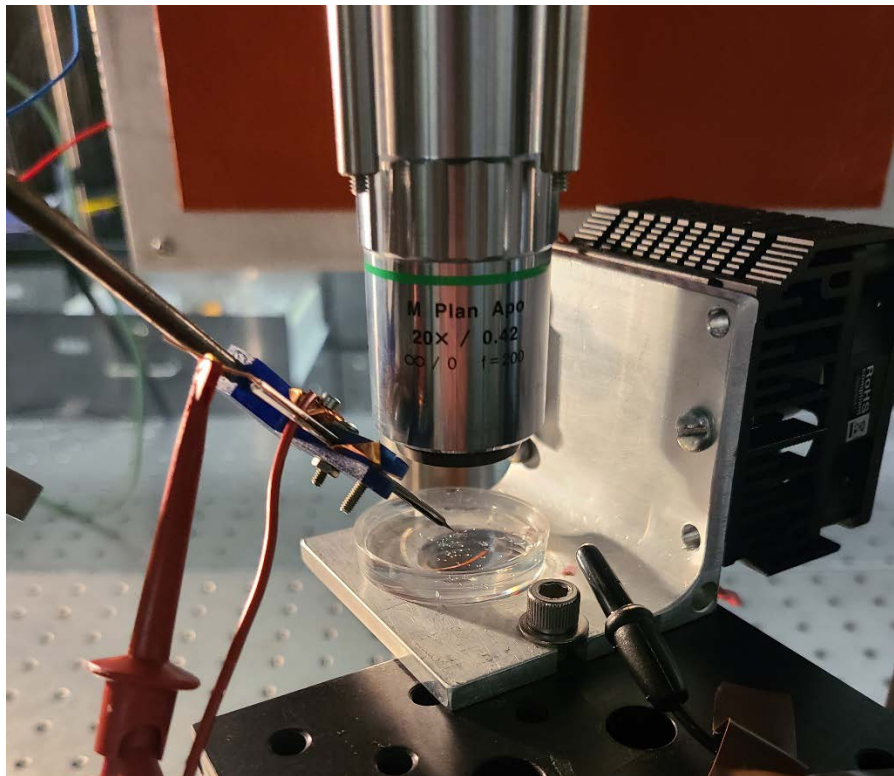
Figure 18: Propylene glycol water mixture placed on MTS stage

Once samples were mixed the sample was poured into a 35 mm cell culture dish,

to fit onto servo-controlled stage in the MTS thermal center. This is shown in figure 18.

3.3.2 Experimental Setup

The experimental setup consists of the following equipment:

- MTS actuator

- Stage actuator

- Computer

- Function Generator

- Oscilloscope

- Pre-amplifier

- Laser

- Laser shutter

- Migtex camera

- 20x lens

A schematic of this setup can be seen in figure (15). The laser is a 532 nm wavelength laser, with controllable power. This was set at a constant power of 100mW for the experiment. Laser pulse was controlled by the function generator and Benchvue application on PC. This pulse was set at 500 µs to match that of the simulation.
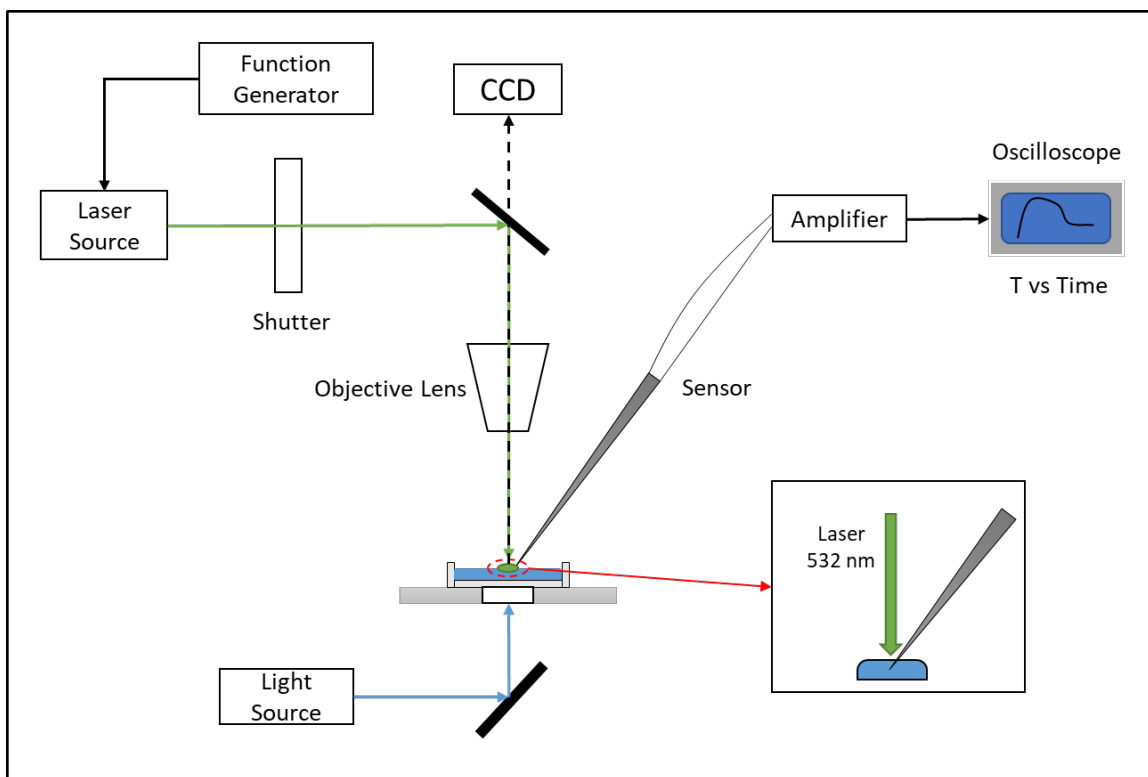
Figure 19: Schematic of the MTS experimental setup

The laser travels the path shown in the schematic and through the 20x lens down to the stage. The Migtex camera was mounted directly above the lens to view the MTS tip and laser spot.

The MTS was connected to a system which had 3 axes of manual movement: left/right, up/down, and forward/backward. This system had a linear actuator servo that

could translate the MTS up and down at a 45° angle. This is important as the tip is on the micron scale and needs to have small and accurately controlled movement. Next the stage has 2 axes of manual movement, and one linear actuator servo for up and down.

The MTS was then connected to the pre-amplifier for a 500x gain on the input voltage. The pre-amplifier was then connected to the oscilloscope to collect the voltage data. Along with this the function generator was connected to the oscilloscope to create the trigger for data collection.

Once everything was connected the experiment consisted of focusing continuous laser on tip of the pipette. This would cause the voltage on the oscilloscope to increase. The sample that is to be tested is placed on the stage and lifted upward until the tip is submerged in the liquid. This will cause an instantaneous decrease in the voltage due to the drastic difference in thermal conductivity of air vs. water. Once in the liquid the function generator is switched on and the pulse can be sent to the tip of the laser and the transient temperature profile can be collected and saved from the oscilloscope.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter has three sections. The first covers the calibration of the sensors. The next goes over the results from training the ANN, discussing the chosen hidden layer size as well as the results from inputting new simulated data. The last section presents the results of the experiment and the classification of the ANN with real data.

4.1   Calibration of Sensor

The first step in the process was to create the MTS that can be used for testing the water/glycol mixtures. In this study we wanted to show the improvement of utilizing bismuth alloy in replacement of the tin used in previous research.

Table 3 shows a clear improvement in the Seebeck coefficient with the addition of bismuth. This is logical as the difference in the Seebeck between bismuth and nickel is 57 µV/K, while tin and nickel is around 14 µV/K. The reason that we are not receiving a Seebeck on the scale of the literature value is due to the thin film deposition [46]. Conditions for sputtering must be improved to increase the sensitivity of the MTS. Annealing of the deposited thin film may enhance the sensitivity.

Table 3: Calibration results table of Sn vs. Bi inner core metal with Seebeck coefficient and tip size

Inner core material	Pipette Number	Seebeck	Tip Size
Sn Pipettes	1	2.7959	3.15
	2	1.8712	3.15
	3	3.0979	3.15
Bi Pipettes	1	10.386	3.15
	2	11.458	3.6
	3	11.809	3.6
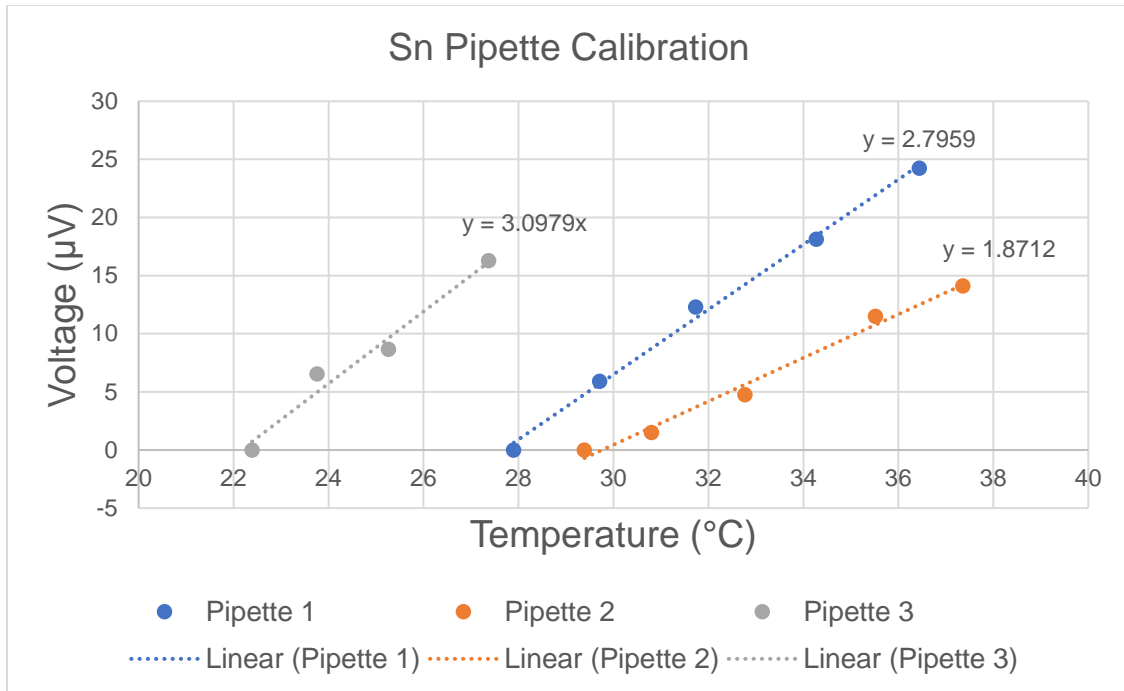
Figure 20: Calibration of Sn MTS with similar tip sizes the slope of each of the lines shows the corresponding Seebeck coefficient of the MTS
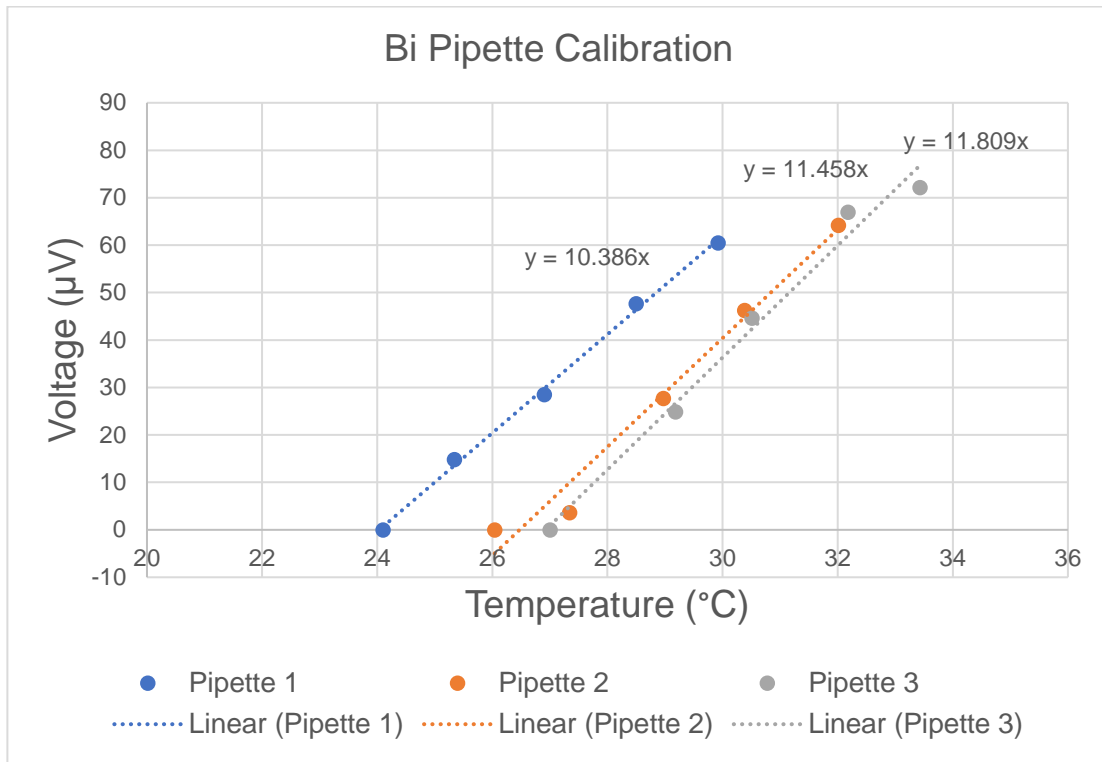
Figure 21: Calibration of Bi MTS with similar tip sizes, the slope of each of the lines shows the corresponding Seebeck coefficient of the MTS

4.2 Results from Training

The simulation was run 900 times with 100 data sets for each class of glycol concentration. The transient temperature profiles were collected and normalized using the min-max method outlined in chapter 2. The data was then randomized and split into the 70% training, 15% validation, and 15% testing. These data sets were then used to train the ANN.

4.2.1 Hidden Layer Size

The first part of optimizing the ANN is to find the hidden layer size that yields the highest F1-Score. This would mean that it had low FNs and FPs in the confusion matrix. Six different network configurations were created with hidden layer size staring at 25 nodes and increasing by 25 until 150 nodes in the hidden layer. All the networks received the same amount of training data for each class representing the 10% change in glycol concentration. Next the micro and macro F1 score was calculated using each networks confusion matrix. Since each class had the same amount of data the Weighted F1 score would be equal to macro F1.

Table 4: Results of hidden layer size testing from 10% change in glycol concentration where column one represents the hidden layer size, two is the micro F1 score, and three is the macro F1 score

Hidden Layer Size	Micro F1 Score	Macro F1 Score
25	.9988	.9988
50	.9933	.9962
75	.9988	.9988
100	1.00	1.00
125	.9888	.9888
150	.9977	.9977

It can be seen in table 4 that all networks performed similarly with a 98% accuracy or

above and no clear correlation between accuracy and hidden layer size. However, the

100 hidden layer size with a perfect performance of 0 FP and FN.
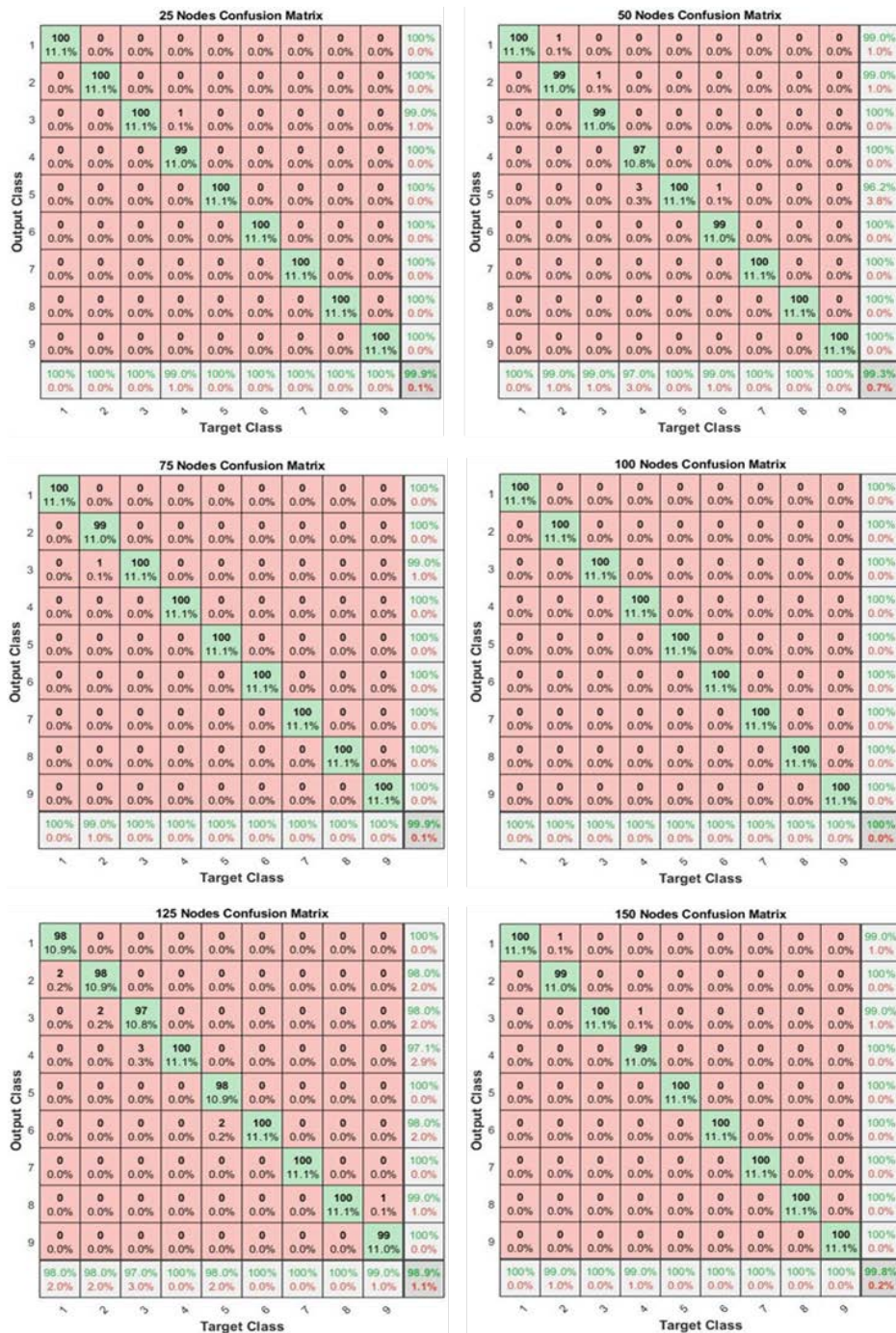
Figure 22:Total data confusion matrix for varying hidden layer size. The green diagonal represents the data sets the network got correct. The far-right column shows the precision for each class, while the bottom row shows the recall. (A) 25 Hidden Nodes (B) 50 Hidden Nodes (C) 75 Hidden Nodes (D) 100 Hidden Nodes (E) 125 Hidden Nodes (F) 150 Hidden Nodes

This configuration reached the minimization at epoch number 109. The validation performance of this epoch was .00852 using the cost function CELF discussed in chapter 2. Confusion matrices for each hidden layer size can be seen in figure 22. These matrices are for the entire set of data meaning the training, validation and testing are included in the final accuracy. The column on the far right of each matrix shows the precision of the class, or the percentage of the classes that were correctly identified. While the bottom row shows the recall of the examples or the percentage of examples that were correctly identified as positive. The bottom right corner shows the overall accuracy of the ANN, which is equal to the Micro F1 score as discussed in the method.

4.2.2 Sensitivity Analysis

The next part of the ANN results was to find the sensitivity of the network or the smallest percent change in glycol that the network can detect. In order to reduce the output size a 0-10% glycol concentration data set was simulated in COMSOL.
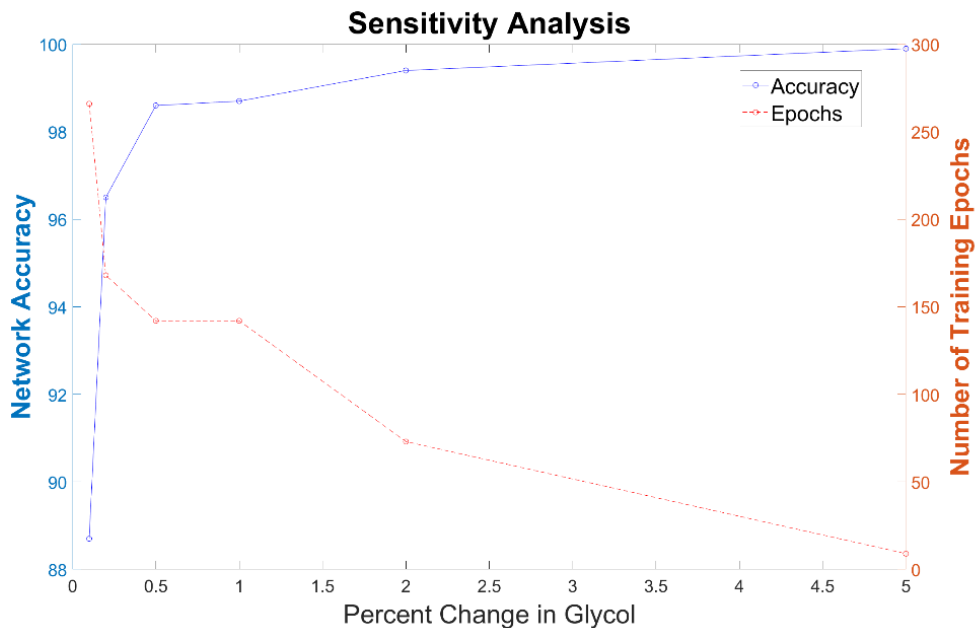
Figure 23: Sensitivity analysis from ANN where the network accuracy and number of training epochs were measured against the percent change in glycol

This data was used to train five different networks: .1%, .2%, .5%, 1%, 2%, and 5% change in the concentration of glycol. It was found that the sensitivity had an inverse effect on the accuracy. It means that as the ANN was trained to classify smaller percent changes in glycol the overall accuracy declined. Along with this the number of epochs increased as the sensitivity increased. This study shows promise that the ANN can classify small changes in glycol concentrations up to .2%, with a 96.5% accuracy. This can be seen in figure 23.

4.3 Experimental Results

This next section will cover the experimental results from measuring the propylene glycol water mixtures. First the optimal distance was found in order match the simulation to the experiment. Next the ANN classification results will be given for the 6 glycol concentrations.

4.3.1 Optimal Distance Testing

Before the ANN can be utilized for testing glycol samples the optimal distance must be found. This is the distance from the tip of the MTS that is being measured. Water with a known thermal conductivity of .608 (W/mK) was tested and compared to simulated data with varied point of measurement. The mean absolute percent error (MAPE) was used to evaluate the difference between the simulated and experimental. It was found that the optimal distance lies between 4.2 and 4.7 µm. These results can be seen in table 5.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - \hat{x}_i}{\hat{x}_i}\right|$$

Table 5: MAPE results from optimal distance testing. Distance of 4.2 μm yielded lowest percent error, leading this to be optimal distance of measurement.

Distance from tip(μm)	0.7	1.2	1.7	2.2	2.7	3.2	3.7	**4.2**	4.7	5.2
MAPE (%)	79.867	58.539	42.514	30.285	20.180	12.976	7.1765	3.9760	5.9831	10.841

Although this distance had the lowest error, it the experiment was still 4% off from the simulated profile. A comparison of both profiles can be seen in figure 24.
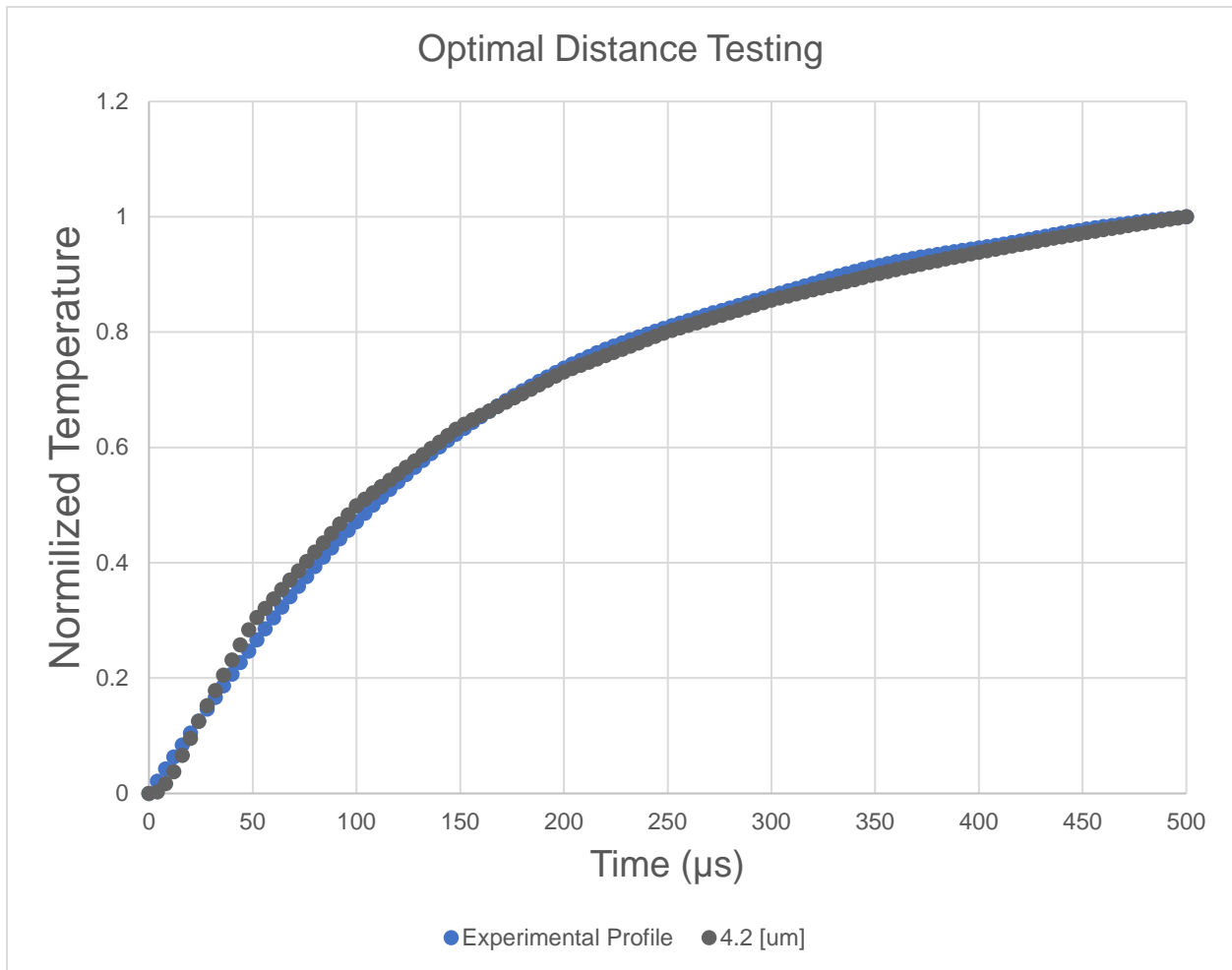
Figure 24:Comparison of a simulated normalized temperature profile of water with a optimal distance of 4.2μm vs. an experimental normalized temperature profile of water

Along with this different optimal distances were tested for different thermal conductivities and plotted. It was found that the distance stayed constant, however there

was a drop after the thermal conductivity fell below 2.5 W/mK. These results can be seen in figure 25.
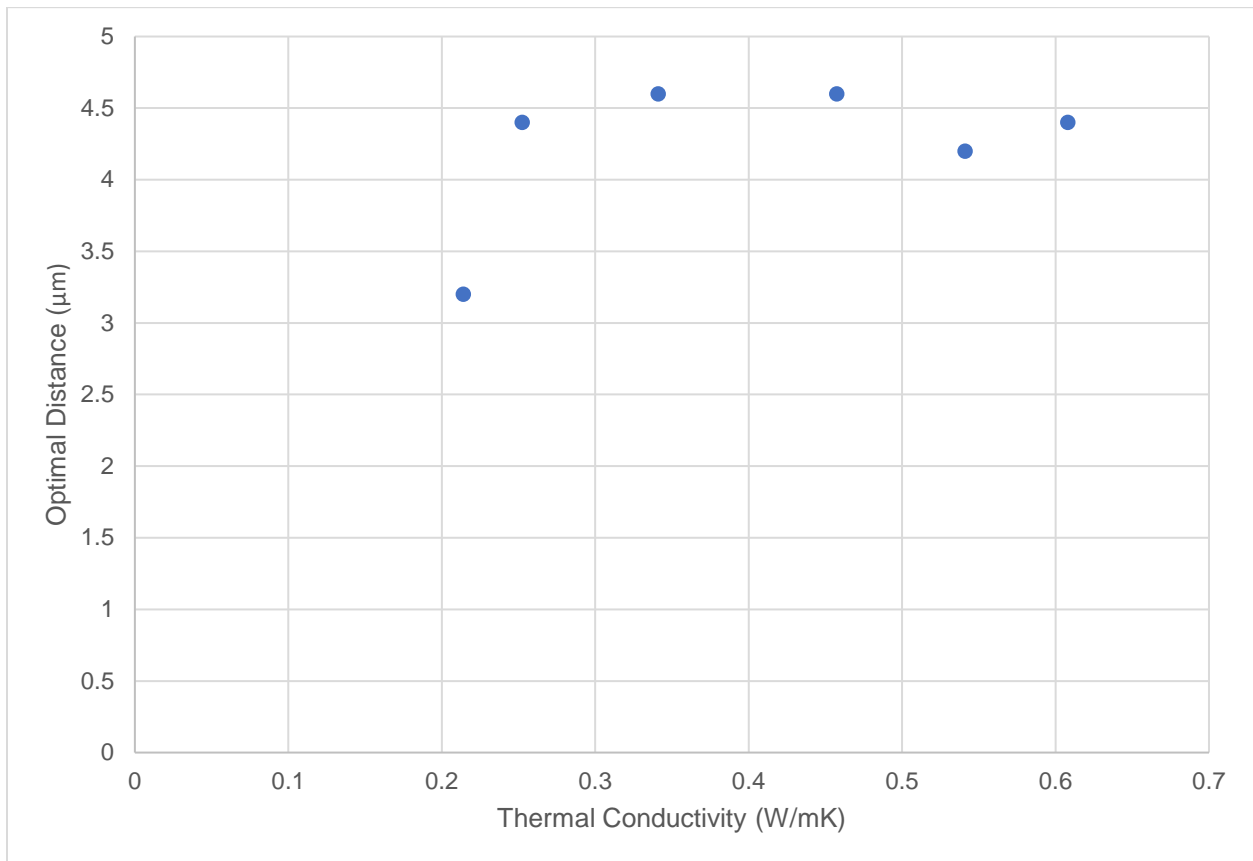
Figure 25: Optimal distance of MTS vs. the thermal conductivity of measured fluid

4.3.2 Experimental Glycol Profiles

In this study 6 different concentrations of glycol were tested and their transient temperature profiles for a 500 µs laser pulse were collected and plotted. It can be seen that 0% and 10% had very clear distinction in the profiles. However, as the concentration increased the difference in thermal conductivity decreases therefore distinction between profiles become blurred. This can especially be seen in 25% and 50% concentrations. It should also be noted that 90% concentration had a very high slope almost following the 25% profile. This may be caused by the sensitivity of the

sensor. As the thermal conductivity of the liquid decreases the accuracy of the sensor seems to decrease as well. These plots can be seen in figure 25.
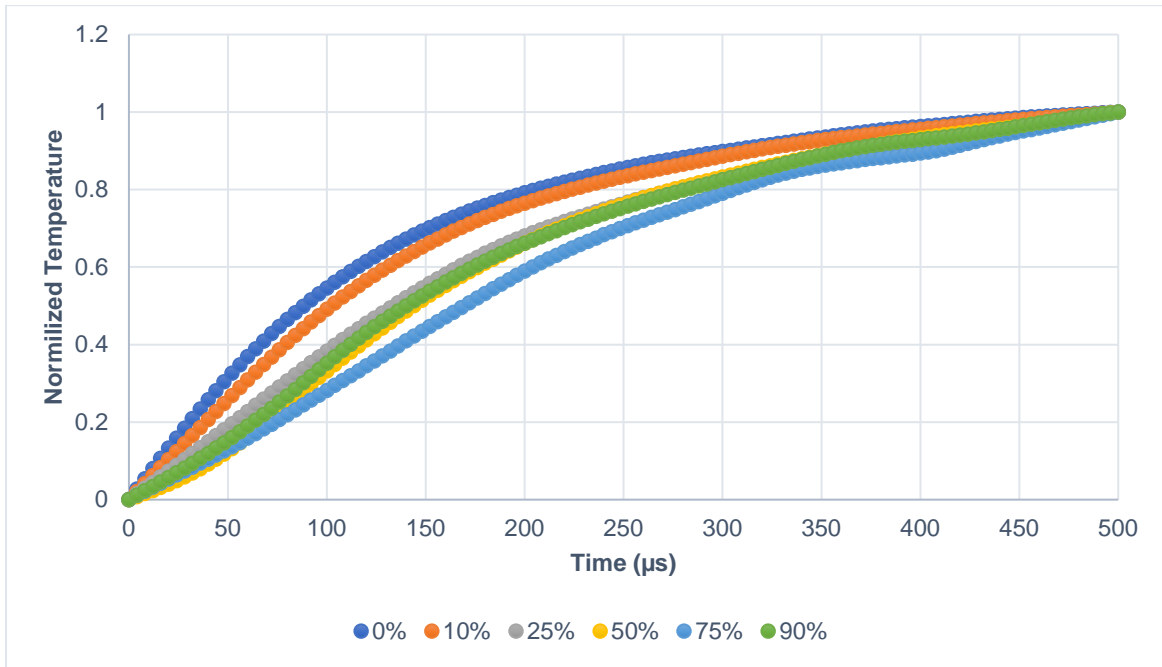
Figure 26: Experimental normalized temperature profiles of water propylene glycol mixtures with concentrations of: 0%, 10%, 25%, 50%, 75%, and 90%

4.3.3 ANN Classification Results

Once the optimal distance was found the networks can be trained with the simulated profiles for this distance. Two different networks were created with optimal distances of 4.2 μm and 4.4 μm. The configuration for both networks had a 100-node hidden layer size. This training showed similar results to the previous section, in which the overall network had a 98.7% and 97.2% accuracy with simulated results. Experimental profiles from the propylene glycol mixtures were inputted into the networks and the classification vectors were collected. Three trials for each mixture concentration of: 0%, 10%, 25%, 50%, 75%, and 90% were tested, and the results for ANN for optimal distance 4.2 μm and 4.4 μm are shown in tables 6 and 7, respectively.

Table 6: Confusion matrix for the ANN with optimal distance of 4.2 µm, This network shows an overall accuracy of 22.2% based upon experimental results.

	0% - 10%	10% - 20%	20% - 30%	30% - 40%	40% - 50%	50% - 60%	60% - 70%	70% - 80%	80% - 90%	
0% - 10%	0	0	0	0	0	0	0	0	0	100%
10% - 20%	0	0	0	0	0	0	0	0	0	100%
20% - 30%	6	0	2	0	0	0	0	0	3	22.2%
30% - 40%	0	0	0	0	0	0	0	0	0	100%
40% - 50%	0	0	0	0	0	1	0	0	0	0.0%
50% - 60%	0	0	0	0	0	2	0	0	0	100%
60% - 70%	0	0	1	0	0	0	0	3	0	0.0%
70% - 80%	0	0	0	0	0	0	0	0	0	100%
80% - 90%	0	0	0	0	0	0	0	0	0	100%
	0.0%	100%	66.6%	100%	100%	66.6%	100%	0.0%	0.0%	**22.2%**

Table 7: Confusion matrix for the ANN with optimal distance of 4.4 µm, this network shows an overall accuracy of 33.3% based upon experimental results

	0% - 10%	10% - 20%	20% - 30%	30% - 40%	40% - 50%	50% - 60%	60% - 70%	70% - 80%	80% - 90%	
0% - 10%	6	0	0	0	0	3	0	3	0	50.0%
10% - 20%	0	0	0	0	0	0	0	0	0	100%
20% - 30%	0	0	0	0	0	0	0	0	0	100%
30% - 40%	0	0	1	0	0	0	0	0	3	0.0%
40% - 50%	0	0	2	0	0	0	0	0	0	0.0%
50% - 60%	0	0	0	0	0	0	0	0	0	100%
60% - 70%	0	0	0	0	0	0	0	0	0	100%
70% - 80%	0	0	0	0	0	0	0	0	0	100%
80% - 90%	0	0	0	0	0	0	0	0	0	100%
	100%	100%	0.0%	100%	100%	0.0%	100%	0.0%	0.0%	**33.3%**

The ANN for 4.2 μm showed good classification for 25% and 50% concentrations with 66.6% precision for each class. Although 75% concentration has a 0% precision, it was put its classification between 60%-70%, which is close to the correct class. All the trials for 0% and 10% were not classified correct and ended with 0% precision for this network. However, when these profiles were tested for the ANN with optimal distance of 4.4 μm, this network correctly classified these concentrations. With 100% precision for this class. This ANN had 0% precision with all other temperature profiles. Both networks had 0% precision with 90%, which could be attributed to this profile overlapping 25% and 50% profile. We can see from the bottom right corner of the confusion matrixes the accuracy of 4.2 μm and 4.4 μm optimal distance were 22.2% and 33.3%, respectively.

If we take the classification of the 0% and 10% concentration from the 4.4 μm optimal distance ANN and the classification of 25% and 50% concentration from the 4.2 μm optimal distance ANN, the overall accuracy of the network would be 55.5%. The reason for the low accuracy of the network might be due to a shifting of the optimal distance at the thermal conductivity of the liquid changes. This method has validity if the change in optimal distance is considered when training the network.

CHAPTER 5

CONCLUSION

In this study we discussed the capability of an ANN to classify up to 10%

differences in propylene glycol concentration. This was done by using a PDE solver to

simulate a laser pulse unto a bismuth and nickel MTS and collecting the transient

temperature profile every 4 μs to train the ANN. The network was designed to have 126

inputs representing the temperature profile, a single hidden layer, and 9 outputs for the

range's thermal conductivity of corresponding to the concentration of glycol. Training

consisted of utilizing SCGD meth with a CELF for the cost. The optimized configuration

of the ANN was found by comparing different hidden layer sizes varying from 25-150. It

was found that 100 nodes in the hidden layer produced the highest accuracy based

upon the confusion matrix. The sensitivity of the network was also found to be .2%,

meaning that with simulated data the network could identify a 0.0066 W/m*K change in

thermal conductivity.

The next part of the study was to identify if the network can replicate these

results with real experimental data in order to demonstrate a sim to real ML approach.

In order to complete this the optimal distance of measurement from the tip of the MTS

had to be found. This was done by measuring a liquid with known parameters such as

water. Utilizing this method, 4.2 μm resulted in a 3.98% MAPE. Two networks were then

trained with an optimal distance of 4.2 μm and 4.4 μm. Next six different concentrations

of propylene glycol water mixtures were tested with 3 profiles collected from each. The

confusion matrix was generated for both networks resulting in an accuracy of 22.2%

and 33.3%, for 4.2 μm and 4.4 μm optimal distance ANNs respectively. The accuracy of

the networks can be improved by considering a varying optimal distance depending on

the thermal conductivity of the fluid being measured.

5.1   Recommendations for Future work

The following topics are recommendations to improved and continue the direction

of the research:

- Utilize different materials for the MTS coating such as gold which would increase the overall Seebeck coefficient. Along with this perform coating inside a clean room in which contaminates from the environment can be removed, thereby improving the coating quality.

- Improve grinding quality of the tip with a laser etching technique, as this directly effects the performance of the MTS.

- Compare accuracy of the ANN with a regression model instead of classification.

- Gather a large amount of experimental data and train the ANN with these profiles and compare to ANN trained with simulated data.

- Improve the simulated training data sets by include a secondary parameter of the optimal distance.

REFERENCES

[1] M. Abdolahad, M. Janmaleki, M. Taghinejad, H. Taghnejad, F. Salehia and S. Mohajerzadeh, "Single-cell resolution diagnosis of cancer cells by carbon nanotube electrical spectroscopy," *Nanoscale,* no. 8, 2018.

[2] D.-K. Kang, M. M. Ali, K. Zhang, E. J. Pone and W. Zhao, "Droplet microfluidics for single-molecule and single-cell analysis in cancer research, diagnosis and therapy," *TrAC Trends in Analytical Chemistry,* vol. 58, pp. 145 - 153, 2014.

[3] B. K. Park, N. Yi, J. Park and D. Kim, "Thermal conductivity of single biological cells and relation with cell viability," *Applied Physics Letters,* vol. 102, 2013.

[4] a. Y. W. a. D. J. Byoung Kyoo Park1, J. Park, T.-Y. Choi, D. P. Simmons, J. Ha and D. Kim, "Thermal conductivity of biological cells at cellular level and correlation with disease state," *Journal of Applied Physics,* vol. 119, 2016.

[5] G. M. P. Y. N. e. a. Kucsko, "Nanometre-scale thermometry in a living cell," *Nature,* vol. 500, p. 54–58, 2013.

[6] F. Vetrone, R. Naccache, A. Zamarrón, A. J. d. l. Fuente, F. Sanz-Rodríguez, L. M. Maestro, E. M. Rodriguez, D. Jaque, J. G. Solé and J. A. Capobianco, "Temperature Sensing Using Fluorescent Nanothermometers," *ASC Nano,* vol. 4, no. 6, p. 3254–3258, 2010.

[7] R. Shrestha, R. Atluri, D. P. Simmons, D. S. Kim and T. Y. Choi, "Thermal conductivity of a Jurkat cell measured by a transient laser point heating method," *International Journal of Heat and Mass Transfer,* vol. 160, p. 120161, 2020.

[8] R. Shrestha, R. Atluri, D. P. Simmons, D. S. Kim and T. Y. Choi, "A micro-pipette thermal sensing technique for measuring the thermal conductivity of non-volatile fluids," *Review of Scientific Instruments,* vol. 89, 2018.

[9] M. Frank, D. Drikakis and V. Charissis, "Machine-Learning Methods for Computational Science and Engineering," *Computation,* vol. 8, 2020.

[10] J.-H. Lee, J. Kang, W. Shim, H.-S. Chung and T.-E. Sung, "Pattern Detection Model Using a Deep Learning Algorithm for Power Data Analysis in Abnormal Conditions," *Electronics,* vol. 9, no. 7, 2020.

[11] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. Ab, E. Mohamed and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heilyon,* vol. 4, no. 11, 2018.

[12] S.-C. Wang, "Artificial Neural Network," in *Interdisciplinary Computing in Java Programming*, Boston, Springer, 2003, pp. 81-100.

[13] P. Sharma, K. Ramesh, R. Parameshwaran and S. S. Deshmukh, "Thermal conductivity prediction of titania-water nanofluid: A case study using different machine learning algorithms," *Case Studies in Thermal Engineering,* vol. 30, 2022.

[14] H. Kurt and M. Kayfeci, "Prediction of thermal conductivity of ethylene glycol–water solutions by using artificial neural networks," *Applied Energy,* vol. 86, no. 10, pp. 2244-2248, 2009.

[15] F. P. Incopera, D. P. Dewitt, T. L. Bergman and A. S. Lavine, Fundamentals of Heat and Mass Transfer, Danvers: John Wiley & Sons, 2007.

[16] T. Seebeck, "Magnetische Polarisation der Metalle und Erze durch Temperatur-Differenz", Berlin: Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin, 1825.

[17] S. Blundell, S. J. Blundell and K. M. Blundell, Concepts in Thermal Physics, Oxford University Press, 2010.

[18] S. Blundell and K. M. Blundell, Concepts in Thermal Physics, Oxford: Oxford Univeristy Press, 2006.

[19] T. Analysis, Bernhard Wunderlich, San Diego: Academic Press, 1990.

[20] H. M. Roder, "A Transient Hot Wire Thermal Conductivity Apparatus for Fluids," *J Res Natl Bur Stand,* vol. 86, no. 5, p. 457–493, 1981.

[21] C. Rossmann and D. Haemmerich, "Review of Temperature Dependence of Thermal Properties, Dielectric Properties, and Perfusion of Biological Tissues at Hyperthermic and Ablation Temperatures," *Critical Review in Biomedical Engineering,* vol. 42, no. 6, pp. 467-492, 2014.

[22] M. W. Ashraf, S. Tayyaba and N. Afzulpurkar, "Micro Electromechanical Systems (MEMS) Based Microfluidic Devices for Biomedical Applications," *International Journal of Moleculer Science,* vol. 12, no. 6, pp. 3648-3704, 2011.

[23] H. O'Hanley, J. Buongiorno, T. McKrell and L.-w. Hu, "Measurement and Model Validation of Nanofluid Specific Heat Capacity with Differential Scanning Calorimetry," *Advances in Mechanical Engineering,* vol. 2012, 2011.

[24] L. Shi and A. Majumdar, "Thermal Transport Mechanisms at Nanoscale Point Contacts," *Journal of Heat Transfer,* vol. 124, no. 2, pp. 329-337, 2002.

[25] P. Gill, T. T. Moghadam and B. Ranjbar, "Differential Scanning Calorimetry Techniques: Applications in Biology and Nanoscience," *Jouranal of Biomolecular Techniques,* vol. 21, no. 4, p. 167–193, 2010.

[26] J. W. Judy, "Microelectromechanical systems (MEMS): fabrication, design and applications," *Smart Materials and Structures,* vol. 10, no. 6, 2001.

[27] S. Spearing, "Materials issues in microelectromechanical systems (MEMS)," *Acta Materialia,* vol. 48, no. 1, pp. 179-196, 2000.

[28] R. Kohli, "Developments in Imaging and Analysis Techniques for Micro- and Nanosize Particles and Surface Features," in *Developments in Surface Contamination and Cleaning*, William Andrew, 2012, pp. 215-306.

[29] L. Shi, *Micro-Nano Thermal Fluid: Physics, Sensors, Measurements,* Austin: Department of Mechanical Engineering University of Texas at Austin, 2016.

[30] S. R. Choi and D. Kim, "Measurement of thermal properties of microfluidic samples using laser point heating thermometry," *Thermochimica Acta ,* vol. 455, pp. 11-15, 2007.

[31] H. Carslaw and J. Jaeger, Conduction of Heat in Solids, Oxford: Clarendon Press, 1959.

[32] T. Mitchell, Machine Learning, New York: McGraw Hill, 1997.

[33] H. Kurt and M. Kayfeci, "Prediction of thermal conductivity of ethylene glycol–water solutions by using artificial neural networks," *Applied Energy,* vol. 86, no. 10, pp. 2244-2248, 2009.

[34] I. Goodfellow, Y. Bengio and A. Courville, "6.5 Back-Propagation and Other Differentiation Algorithms," in *Deep Learning*, MIT Press, 2016, pp. 200-220.

[35] S. Ruder, "An overview of gradient descent optimization algorithms," 2016.

[36] D. R. Wilson and T. R. Martinez, "The general inefficiency of batch training for gradient descent learning," *Neural Networks,* vol. 16, no. 10, pp. 1429-1451, 2003.

[37] P. D. Wasserman, Advanced methods in neural computing, Van Nostrand Reinhold, 1993.

[38] L. Bottou, "Large-Scale Machine Learning with Stochastic Gradient Descent," in *COMPSTAT*, Paris, 2010.

[39] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks,* vol. 6, no. 4, pp. 525-533, 1993.

[40] J. Jang, C. Sun and E. Mizutani, Neuro-Fuzzy and Soft Computing—A Computational Approach to Learning and Machine Intelligence—, Englewood Cliffs: Prentice-Hall, 1997.

[41] W. Wang and Y. Lu, "Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model," in *IOP Conference Series: Materials Science and Engineering*, 2018.

[42] K. P. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.

[43] P. D. Beale, Statistical Mechanics, Academic Press, 2011.

[44] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," 2015.

[45] I. Markoulidakis, G. Kopsiaftis, I. Rallis and I. Georgoulas, "Multi-Class Confusion Matrix Reduction method and its application on Net Promoter Score classification problem," in *PETRA 2021*, 2021.

[46] F. G. Cougnon and D. Depla, "The Seebeck Coefficient of Sputter Deposited Metallic Thin Films: The Role of Process Conditions," *Coatings,* vol. 9, no. 5, 2019.