

# **MACHINE-LEARNED PIEZORESISTIVE SENSORS FOR BILINGUAL SIGN LANGUAGE RECOGNITION**

**MECH0020 Final Report  
Student: Toby Bunn  
Supervisor: Professor Manish Tiwari**

October 11, 2023

## Abstract

Novel sensor development is a field which holds great future potential. Using different sensor types, various gestures and movements can be classified through datasets recorded from the sensors. These have proved to offer significant developments in many areas, for example aiding clinicians in healthcare applications. The main barrier to wide exploitation of these technologies is the price, which is usually unfeasible for common use, especially in lower-income countries. One growing area of interest in sensorised technology is aiding the deaf community. People with hearing impairments may struggle to communicate with the general population as most people do not know sign language, and it is difficult and inaccessible to learn. This study attempts to use cheap, novel, piezoresistive sensors to detect sign language gestures through classification of the time-dependent voltage signals, which has potential of broadening the knowledge of sign language worldwide if combined with an accessible user interface. This project is split into two distinct parts: experimental and computational. The experimental section aims to test new, novel approaches to detecting strain from the bending of finger joints. The requirements imposed for these are that they must be inexpensive and simple to make. Three different sensors are therefore tested and developed, one a commercial sensor that has been previously used in this type of application which should provide a good standard for comparison, and two novel fabricated sensors that have not been used in this setting before. In the computational part, machine learning is exploited to train and test various models for American and Arabic sign language gesture classification, and also a combined dataset to comprehensively evaluate the capabilities of the model. The variations used consist of two chosen sensor types from the experimental testing, used for three different datasets, and tested with three different data splits to mitigate overfitting and other model issues. This results in eighteen total models, each of which utilises five unique algorithms: Support Vector Machine, Random Forest, k-Nearest Neighbours, Artificial Neural Network, and Extreme Gradient Boosting. All models generally exhibit very high performance on all datasets, giving final classification accuracies over 85% in every case. Many even exhibited perfect performance on the testing data with learning curves suggesting well-fitted models. This shows that the piezoresistive sensors have high potential in gesture recognition applications with broader adoption of these technologies reliant on further developments of sensors and varied datasets, using low-cost, reproducible sensors.

All code and Machine Learning data used for the project is publicly available via the following link:

<https://github.com/Toby080402/SensorisedGloves.git>

## Acknowledgements

I would like to thank my supervisor Professor Manish Tiwari for his advice and supervision throughout the project, which helped keep the project on track and to a high quality.

I would also like to express my sincerest gratitude to Carmen Salvadores Fernandez, a PhD student in the mechanical engineering department, and Rohit Gupta, an academic researcher at UCL. Their guidance, advice, mentoring, and technical assistance in the different areas of my project were crucial to achieving the success I did, and I am very grateful for both of their contributions and eagerness to assist.

## Declaration

I, Toby Bunn, confirm that the work presented in this report is my own. Where information has been derived from other sources, I confirm that this has been indicated in the report.

## Word Count

Word Count: 7489 words (subtracting numbers within tables containing data)

# Contents

<b>List of Figures</b>	<b>5</b>
<b>List of Tables</b>	<b>6</b>
<b>List of Acronyms</b>	<b>7</b>
<b>Nomenclature</b>	<b>8</b>
<b>1 Introduction</b>	<b>9</b>
1.1 Background and Problem Statement . . . . .	9
1.2 Project Aim . . . . .	9
<b>2 Literature Review</b>	<b>10</b>
2.1 Tactile Sensors . . . . .	10
2.2 Theoretical Basis . . . . .	10
2.3 Wearable Sensor Review . . . . .	11
2.4 Machine-Learned Sensor Review . . . . .	12
2.5 Sign Language . . . . .	14
2.6 Project Novelties Conceived from Previous Literature . . . . .	15
<b>3 Methodology</b>	<b>16</b>
3.1 Project Plan . . . . .	16
3.2 Experimental Methods . . . . .	16
3.2.1 Materials . . . . .	16
3.2.2 Fabrication . . . . .	17
3.2.3 Instrumentation . . . . .	19
3.2.4 Test Strategy . . . . .	20
3.3 Computational Plan . . . . .	21
3.3.1 Coding Language and Libraries . . . . .	21
3.3.2 Data Collection . . . . .	21
3.3.3 Data Transformation . . . . .	21
3.3.4 Data Splitting . . . . .	21
3.3.5 Algorithms . . . . .	22
3.3.6 Hyper-Parameter Optimisation . . . . .	22
3.3.7 Validation . . . . .	22
3.3.8 Evaluation Performance Metrics . . . . .	22
<b>4 Signal Characterisation of Sensorised Gloves</b>	<b>23</b>
4.1 Commercial Sensors . . . . .	23
4.1.1 All Joint Tests . . . . .	23
4.1.2 Evaluation of Signal to Noise Ratio for Periodic Tests . . . . .	24
4.2 Foam Sensors . . . . .	26
4.2.1 Tension and Compression Tests . . . . .	26
4.2.2 Reproducibility Resistance Tests . . . . .	26
4.2.3 Evaluation of Signal to Noise Ratio for Periodic Tests . . . . .	27
4.3 Screen-Printed Sensors . . . . .	29
4.3.1 Tension and Compression Tests . . . . .	29
4.3.2 Reproducibility Resistance Tests . . . . .	29
4.3.3 Uniaxial Tensile Tests . . . . .	30
4.4 Decisions from Final Data Analysis . . . . .	31
<b>5 Machine Learning</b>	<b>33</b>
5.1 Statistical Exploration of Datasets . . . . .	33

5.1.1	Signals . . . . .	33
5.1.2	Algorithm Training . . . . .	34
5.2	Machine Learning Classification Results . . . . .	35
5.2.1	Foam Sensor with Singular Language Dataset . . . . .	35
5.2.2	Foam Sensor with Combined Datasets . . . . .	37
5.2.3	FlexiForce Sensor with Singular Language Dataset . . . . .	38
5.2.4	FlexiForce Sensor with Combined Datasets . . . . .	40
5.2.5	Best Performing Learning Curves . . . . .	40
5.3	Analysis and Evaluation of Results . . . . .	41
5.3.1	Observations from Results . . . . .	42
5.3.2	Evaluation of Data Limitations . . . . .	42
5.4	Live Recognition . . . . .	43
<b>6</b>	<b>Conclusion and Evaluation of Process and Outcomes</b>	<b>44</b>
6.1	Summary . . . . .	44
6.2	Strengths . . . . .	44
6.3	Weaknesses . . . . .	44
6.4	Limitations . . . . .	44
6.5	Future Work and Implications . . . . .	45
<b>References</b>		<b>46</b>
<b>Appendix A</b>		<b>49</b>
<b>Appendix B</b>		<b>49</b>

## List of Figures

1	Example of a sensorised glove known as WaveGlove [5]. It uses five inertial sensors to sense finger movement.	9
2	Screen-printed piezoresistive sensor design used for facial expression recognition [10].	11
3	AmSL alphabet [45].	14
4	Gestures for a) ILY sign [46], and b) Rock on sign [47].	15
5	a) Arabic symbol translations [48] with b) Arabic sign language alphabet [49].	15
6	Overall project plan.	16
7	FlexiForce A201 Sensor.	17
8	a) Foam CNF sensor, and b) foam CNF sensor after packaging.	17
9	a) Flowchart of the screen-printed sensor fabrication process, b) stencil CAD drawing, and c) screen-printed sensor.	18
10	Voltage divider circuit a) diagram and b) image.	19
11	Tensile testing setup a) image and b) schematic.	20
12	a) Finger showing joint numbering and b) hand showing numbering of each finger.	20
13	Initial FlexiForce A201 tests for consistent bending on a) joint 1 front, b) joint 1 back, c) joint 2 front, d) joint 2 back.	23
14	FlexiForce A201 best performing joint test for extended periods of strain.	24
15	FlexiForce A201 SNR test signals for a) little, b) ring, c) middle, d) index, e) thumb. Tests involve 5 equal periodic bends over a 37.5s time interval.	25
16	Initial foam sensor tests for consistent bending on joint 2 a) front and b) back.	26
17	Foam sensors SNR test signals for a) little, b) ring, c) middle, d) index, e) thumb. Tests involve five equal periodic bends over a 37.5s time interval.	28
18	Initial tension and compression screen-printed sensor tests on the joints a) front and b) back.	29
19	Screen-printed sensor double print with periodic bending test.	30
20	Uniaxial tensile tests using periodic incremental 1mm extension every 10-20 seconds for a) 1 print, b) 2 prints, and c) 3 prints.	31
21	Final glove fabrications with a) FlexiForce A201 sensors, and b) Foam sensors.	32
22	Bar charts for 'Dhaal' gesture for a) Foam sensors, and b) FlexiForce sensors.	33
23	Bar charts for a) Foam sensor baseline, b) Foam sensor 'Haa', c) FlexiForce sensor baseline, and d) FlexiForce sensor 'Haa'.	34
24	Increasing epoch effect on the accuracy of the ANN(2) algorithm for foam combined dataset.	35
25	Foam AmSL best-performing algorithm confusion matrix (kNN(2)).	36
26	Foam ArSL best-performing algorithm confusion matrix (ANN(2)).	37
27	Foam combined best-performing algorithm confusion matrix (kNN(1)).	38
28	FlexiForce AmSL best-performing algorithm confusion matrix (kNN(2)).	39
29	FlexiForce ArSL best-performing algorithm confusion matrix (ANN(2)).	39
30	FlexiForce combined best-performing algorithm confusion matrix (kNN(1)).	40
31	Learning curves for a) Foam AmSL (kNN(2)), b) FlexiForce AmSL (kNN(2)), c) Foam ArSL (ANN(2)), d) FlexiForce ArSL (ANN(2)), e) Foam Combined (kNN(1)), and f) FlexiForce combined (kNN(1)).	41
32	RF(1) learning curve for Foam AmSL dataset.	42

## List of Tables

1	Summary of related work in classification. . . . .	13
2	FlexiForce A201 peak and trough SNR signals. . . . .	24
3	Foam sensor base resistance values. . . . .	26
4	Foam sensors peak and trough SNR signals. . . . .	27
5	Screen-printed sensors base resistances. . . . .	29
6	Screen-printed sensors base resistances (multiple prints). . . . .	30
7	Best-performing classification report of foam sensors on singular datasets. . . . .	36
8	Best-performing classification report of foam sensors on combined dataset. . . . .	37
9	Best-performing classification report of FlexiForce A201 on singular datasets. . . . .	38
10	Best-performing classification report of FlexiForce A201 on combined dataset. . . . .	40
11	Full classification report of foam sensors on singular datasets. . . . .	49
12	Full classification report of foam sensors on combined dataset. . . . .	50
13	Full classification report of FlexiForce A201 on singular datasets. . . . .	50
14	Full classification report of FlexiForce A201 on combined dataset. . . . .	51

## List of Acronyms

**AmSL** American Sign Language.

**ANN** Artificial Neural Network.

**ArSL** Arabic Sign Language.

**CNF** carbon nanofiber.

**DT** Decision Tree.

**GF** gauge factor.

**HCI** human-computer interaction.

**ILY** I love you.

**kNN** k-Nearest Neighbours.

**ML** Machine Learning.

**NB** Naive Bayes.

**RF** Random Forest.

**SNR** Signal to Noise Ratio.

**SVM** Support Vector Machine.

**XGBoost** Extreme Gradient Boosting.

## Nomenclature

$\Delta$ or $d$	Change in
$\mu$	Sample Mean ( $V$ )
$\rho$	Resistivity ( $\Omega \cdot m$ )
$\sigma$	Sample Standard Deviation ( $V$ )
$\varepsilon$	Strain
$A$	Surface Area ( $m^2$ )
$GF$	Gauge Factor
$L$	Length ( $m$ )
$R$	Resistance ( $\Omega$ )
$R_0$	Initial Resistance ( $\Omega$ )
$SNR$	Signal to Noise Ratio
$V$	Voltage ( $V$ )

# 1 Introduction

## 1.1 Background and Problem Statement

Wearable sensor technology has gained popularity in recent years for various applications, including health monitoring and human-computer interaction (HCI). These sensors gather data and report it to a central computer. For example, commercially available smartwatches track and display the number of steps or calories burned to aid with fitness routines. As these sensors prove successful, they are being exploited in more areas.

Sensors are a transduction technology which can be fabricated using flexible materials for gesture recognition purposes, enabling automation in tasks related to HCI. In healthcare, hand gestures of surgeons can indicate critical procedures during surgery. This data can show contrasts between a trainee and an expert, and provide key areas for improvement. In addition, gesture recognition could also be used to aid doctors within procedures themselves, with a trained algorithm that can perform a task when it recognises a specific gesture being made. An applied example is using gestures to manipulate images during a procedure [1]. This could save valuable time in critical procedures. There are other uses for this technology within healthcare, like recognising shapes of different handheld objects, which has a clear significance towards developing assistive technologies for visually disabled personnel [2]. Furthermore, medical rehabilitation is an area of large potential for this technology. In physical therapy for hand functions, precise gesture classification can determine whether the patient is performing exercises correctly so that rehabilitation is effective. There is a lot of research into stroke rehabilitation in particular, generally producing excellent algorithm performance [3]. Another area of exploit is the computer gaming sector, where sensorised gloves are routinely used for HCI, like in virtual reality settings.

These gloves require a robust, reproducible, and accurate Machine Learning (ML) model, which converts time-dependent voltage signals to target quantities, which must then be output in an interpretable way. In this realm, implementation of interpretable deep learning models can offer significant accuracy improvement, while ensuring generalisability [4].

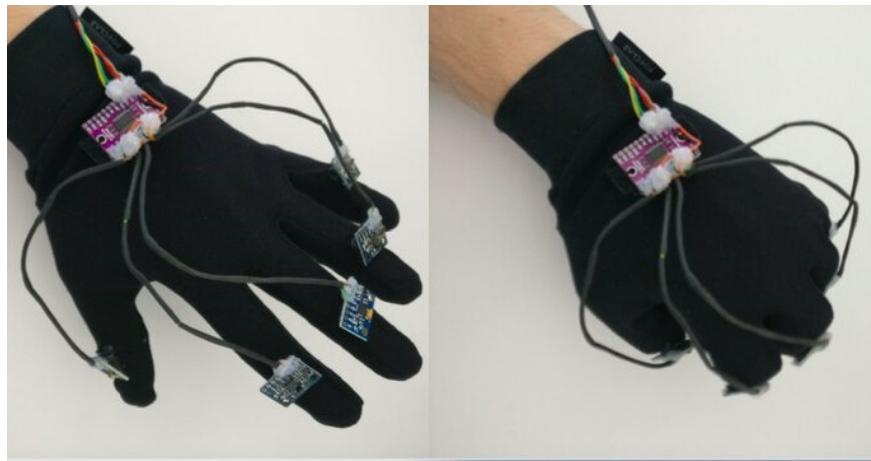


Figure 1: Example of a sensorised glove known as WaveGlove [5]. It uses five inertial sensors to sense finger movement.

## 1.2 Project Aim

This project aims to create sensorised surgical gloves, with affordable flexible strain sensors embedded onto finger joints, to reproducibly classify a variety of sign language gestures. The model will be trained using time-dependent voltage signals, obtained from commercial and fabricated sensors for comparison. These are fabricated and tested for sensor functionality and suitability.

## 2 Literature Review

The literature review examined key areas to be explored further. These areas include general tactile sensor research and their theoretical basis, research reviews of wearable sensors in different applications, and an ML review of related classification research.

### 2.1 Tactile Sensors

An appropriate tactile sensor must be chosen. The project aims to create wearable sensors that detect strain from bending when adhered to finger joints. Therefore, the main sensor types researched are triboelectric, piezoresistive, and capacitive.

**Triboelectric:** The triboelectric mechanism converts applied forces to electrical signals through the triboelectric effect. Contact electrification is a universal effect that occurs when two materials in contact are separated or rubbed against each other [6]. The triboelectric series ranks these materials based on how easily they lose or gain electrons [7].

**Capacitive:** For a capacitive sensor, a change in touch, pressure, or force would produce a change in distance between capacitor plates, therefore causing a change in capacitance. It is governed by a relationship dependent on plate size and sensor thickness, and the magnitude varies dependent on the permittivity of the material [8].

**Piezoresistive:** For a piezoresistive sensor, a change in touch, pressure, or force would produce a measurable change in resistance. Piezoresistive sensors are commonly made from semiconductor materials that conduct under certain conditions. These conditions cause a change in the resistance signal measured by altering the material's conductivity. For example, exerting a force on conductive nano-fibre material causes compression of the material and a decrease in resistance [9].

Piezoresistive sensors are stable, reproducible, and produce good results throughout their common use in literature. They are generally the easiest to fabricate at low costs, making them the focus of this project.

### 2.2 Theoretical Basis

To successfully fabricate piezoresistive sensors, the governing equations must be understood. This allows sensors to be designed to give desired signal ranges. The shape relationship is shown in Equation 1.

$$R = \frac{\rho l}{A} \quad (1)$$

Where  $R$  is resistance,  $\rho$  is electrical resistivity,  $l$  is length, and  $A$  is area [10]. The area is split into two components: the width and thickness. The thickness, length, and width depend on the sensor's design and shape. Finally, resistivity is a material property depending on the conductible ink that is used.

By taking natural logarithms and differentiating Equation 1, Equation 2 can be found.

$$\frac{1}{R} dR = \frac{1}{\rho} d\rho + \frac{1}{l} dL - \frac{1}{A} dA \quad (2)$$

Rearranging Equation 2 and knowing resistivity is a constant material parameter finds Equation 3. This shows resistance change is dependent on the sensors' base parameters along with the change in length and surface area when strained.

$$dR = R \left( \frac{dL}{L} - \frac{dA}{A} \right) \quad (3)$$

There are other relationships that the sensors follow. For example, Equation 4 is for gauge factor (GF), which represents the sensitivity of strain sensors to deformation.

$$GF = \frac{\Delta R}{\varepsilon R_0} \quad (4)$$

Where  $\Delta R$  represents resistance change,  $\varepsilon$  the strain, and  $R_0$  the initial resistance [11]. For straight-line printed sensors, the working strain range increases and gauge factor decreases as the line width increases [12]. Another important consideration for flexible sensors is microstructure. With large strains or poor printing, microcracks are generated and decrease the number of conductive pathways, increasing resistance [13].

Signal to noise ratio (SNR) tests must also be performed. This is defined in Equation 5, and shows how much the signal varies per standard deviation.

$$SNR = \frac{\mu}{\sigma} \quad (5)$$

Where  $\mu$  is the signal mean and  $\sigma$  is the standard deviation of the noise [14]. A ratio greater than 1 represents more signal than noise. As a result, a higher SNR suggests a less noisy signal and a higher sensor sensitivity where most of the measurement is from the signal produced, which is desired.

### 2.3 Wearable Sensor Review

Sensorised gloves collect data about hand movements and gestures. Careful consideration is required on sensor type, placement, and materials to ensure accuracy, reliability, and reproducibility. Despite their usefulness in various applications, they usually have crucial limitations, like cost. Universalisation of these technologies, especially in lower-income countries, requires costs to be greatly reduced. As a result, the concepts studied must be cheap and easy to fabricate. These requirements are most important for healthcare applications [15]. This is a key area, as there are currently no sensorised surgical gloves specifically as a tool to aid clinicians throughout procedures.

In a study by Ramli et al. [10], published in 2020, screen-printed piezoresistive sensors were detailed, which can be fine-tuned within a lab for different applications. It is simple, inexpensive, and allows the use of stretchable conductive materials. The focus is facial expressions recognition, a transferable area to hand gesture recognition. Signals could improve due to the higher strain exhibited by finger joints than facial muscles. The sensor's shape, shown in Figure 2, involves two rectangles made of stretchable conductive material, connected by a thin line of the material. Equation 1 confirms that resistance changes will be much higher within the thin line of material than the two squares that it is connecting. These changes can be considered negligible, with the squares acting as terminals in which resistance changes can be measured across. The stencil design and ink used can tune the parameters as desired. For example, Equation 4 suggests a smaller line width leads to higher sensitivity.

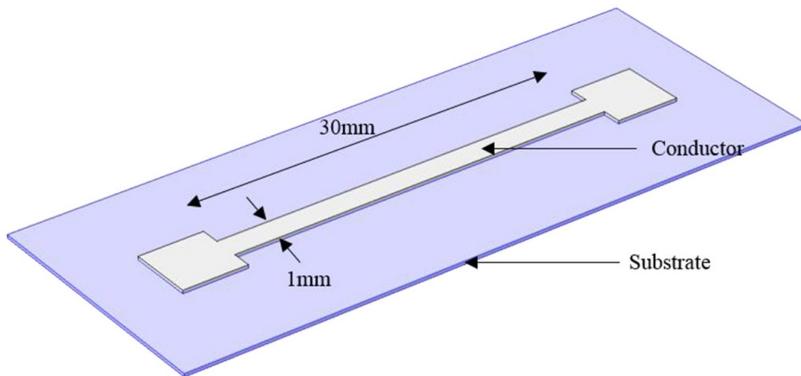


Figure 2: Screen-printed piezoresistive sensor design used for facial expression recognition [10].

Conductive foam piezoresistive sensors are also commonly used, due to their low cost, large surface area, low weight, and high flexibility [16]. By soaking melamine foam in conductive ink, carbon nanofiber (CNF) flakes are deposited throughout the melamine structure [17]. Compression increases the number of conductive pathways and therefore

decreases resistance. These have success in gesture recognition, for example, a 98.9% recognition accuracy for Arabic sign language detection [18]. These sensors can be biocompatible [19], which combined with soft foam offers potential in clinical applications.

Other studies have been conducted using commercially available sensors. These are simpler due to only needing to adhere them to gloves. They are known to work well for the application, offering reliability. The FlexiForce A201 sensors were found to have high precision when tested against other commercial sensors, which would be a good fit for a repeatable ML classifier [20]. They are preferred over others in the market as it is thin, flexible, convenient, and easy to use [21]. These have also been used successfully in gesture recognition applications, like detecting the movement of tendons in a smartwatch [22], VR applications [23], and sensing micro-gestures whilst cycling [24]. However, they are significantly more expensive than other options discussed.

## 2.4 Machine-Learned Sensor Review

The project will use labelled data to train a classifier, which entails supervised machine learning. The categories researched are gesture recognition using different methods, and general classification of datasets. The algorithms explored consist of Artificial Neural Network (ANN), k-Nearest Neighbours (kNN), Support Vector Machine (SVM), Naive Bayes (NB), Random Forest (RF), Decision Tree (DT), and Extreme Gradient Boosting (XGBoost).

Table 1: Summary of related work in classification.

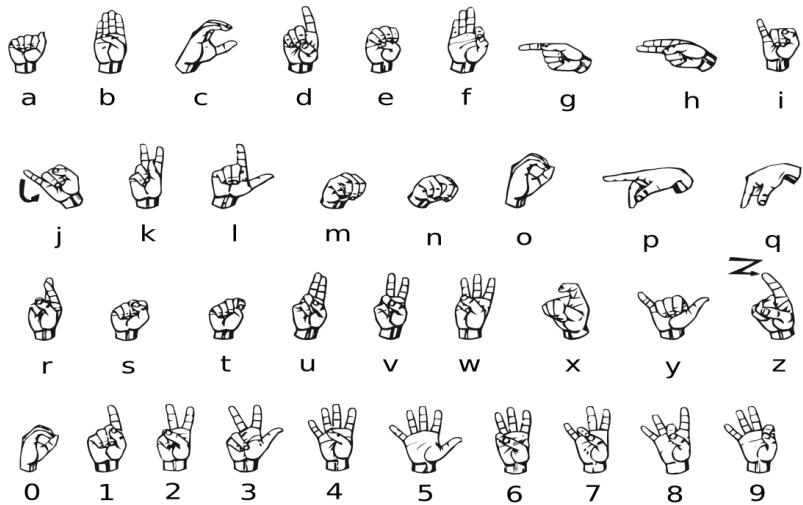
Reference	Algorithms used							Category	Results
	ANN	kNN	SVM	NB	RF	DT	XGBoost		
[3]		✓	✓					Gesture (sensor)	kNN and SVM (linear) perform excellently
[18]						✓		Gesture (sensor)	Achieves 98.9% accuracy
[25]		✓	✓	✓	✓			Gesture (sensor)	SVM best, NB worst, others perform well
[26]		✓			✓	✓		Gesture (sensor)	kNN best, RF similar, DT clear worst
[27]	✓	✓	✓		✓			Gesture (sensor)	All perform similarly
[28]							✓	Gesture (sensor)	High accuracy of 94%
[29]		✓	✓			✓		Gesture (sensor)	kNN best, DT and SVM similar
[30]			✓					Gesture (sensor)	Achieves >97% accuracy
[31]			✓	✓	✓			Gesture (sensor)	RF best, NB worst, SVM computationally expensive
[32]	✓							Gesture (sensor)	Achieves >90% accuracy
[33]	✓	✓			✓	✓		Gesture (sensor)	ANN and RF best, DT worst
[34]			✓		✓			Gesture (sensor)	RF best , SVM performs well
[35]	✓	✓	✓	✓				Gesture (image)	ANN, kNN, SVM all great performance, NB worst
[36]		✓	✓		✓			Gesture (image)	RF, Linear SVM best performing algorithms
[37]		✓	✓	✓	✓		✓	Gesture (image)	RF, XGBoost, kNN best, NB worst
[38]			✓				✓	Gesture (waves)	XGBoost outperforms SVM
[39]					✓		✓	Gesture (waves)	RF and XGBoost both perform excellently
[40]		✓	✓					Data	RBF SVM best performing kernel and algorithm
[41]		✓	✓	✓	✓	✓		Data	RF best, NB worst, others perform similarly
[42]			✓	✓	✓			Data	SVM and RF best performing overall
[43]			✓	✓	✓			Data	RF and SVM most often performed best
[44]	✓	✓	✓		✓	✓		Data	RF and ANN best, all good convergence speed

## 2.5 Sign Language

The approach uses strain sensors which bend at the respective joints they are mounted to. Sign languages that involve motion therefore cannot be detected without an alternative sensor, like an accelerometer. For this project, only a small range of gestures will be used, restricted to ones solely requiring the bending of fingers for simplicity.

American Sign Language (AmSL) and Arabic Sign Language (ArSL) use a singular hand to make letters with fingers. In AmSL, only two gestures involve motion (J, Z), and only one letter requires motion in ArSL. As a result, a mixture of letters, numbers, and phrases will be used to try and obtain a set of gestures which utilise the different joints in many separable ways, so that many distinguishable combinations are tested. In AmSL, initial signs best for classifying are 1, 2, 3, 4, 7, 8, A, F, I, L, S, W, and Y, shown in Figure 3. These are easily distinguishable as they involve bending different fingers.

### American Sign Language Alphabet



wplipart.com - public domain images

Figure 3: AmSL alphabet [45].

The phrases used are in Figure 4, which show (a) 'I love you (ILY)' and (b) 'rock on' signs to gain a full range of gestures for the available joints.

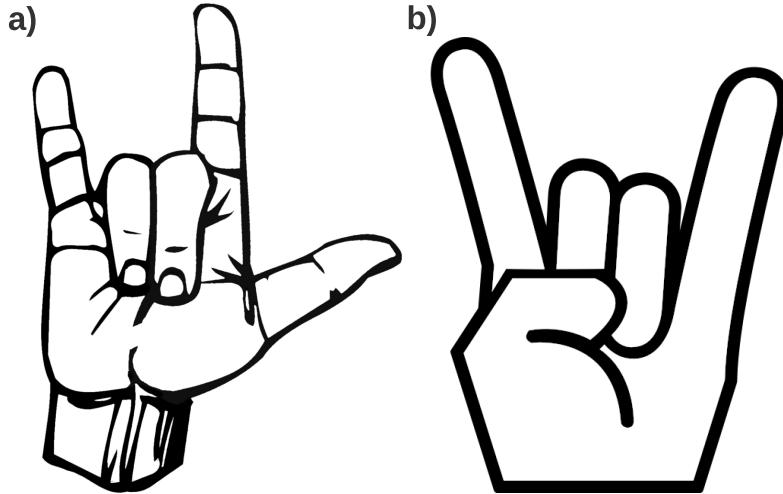


Figure 4: Gestures for a) ILY sign [46], and b) Rock on sign [47].

In ArSL, the letters chosen for classification are Jiim, Thaa, Taa, Baa, Alif, Dhaal, Daal, Qaaf, Ghayn, Zaa, Yaa, Haa, Miim, Laam, and Kaaf. These are shown in Figure 5 with translations.

Khaa	Haa	Jiim	Thaa	Taa	Baa	Alif
خ	ح	ج	ت	ب		
Saad	Shiin	Siin	Zay	Raa	Dhaal	Daal
ص	ش	س	ز	ر	ذ	د
Qaaf	Faa	Ghayn	Ayn	Zaa	Taa	Daad
ض	ف	غ	ع	ظ	ط	ل
Yaa	Waaw	Haa	Nun	Miim	Laam	Kaaf
ي	و	ه	ن	م	ل	ك

**b)**

Figure 5b displays the Arabic sign language alphabet. It shows a grid of 28 hand gestures, each corresponding to one of the 28 letters listed in the table above. The letters are arranged in four rows: the first row contains ح, ش, ك, ث, ت, ب; the second row contains د, ذ, ر, ز, س, ص; the third row contains ض, ف, غ, ع, ط, ظ; and the fourth row contains ي, و, ه, ن, م, ل, ك.

Figure 5: a) Arabic symbol translations [48] with b) Arabic sign language alphabet [49].

## 2.6 Project Novelties Conceived from Previous Literature

The primary novelty of the experimental phase is to develop new, low-cost, reproducible sensors for use in strain applications. The main limitations in this area are cost, fabrication time, size, and reproducibility. Strain measurement sensors like the ones studied have not been utilised on surgical gloves in this manner with implications for clinician aid before, which could lead to significant future advancements after biocompatibility tests. Computationally, the novelty is performing bilingual sign language detection. This is ideal for testing the concept of classifying distinguishable signals with affordable, fabricated sensors, which can be utilised in many different settings. There is a gap in this area in literature, highlighting the need for strain sensor development for gesture recognition purposes.

### 3 Methodology

The planning of work can be split into two main areas: practical work encompassing experiments undertaken, and computational work encompassing the ML process. A general plan is described for each, with more in depth analysis of processes to be undertaken.

#### 3.1 Project Plan

Figure 6 shows the flowchart for the project plan.

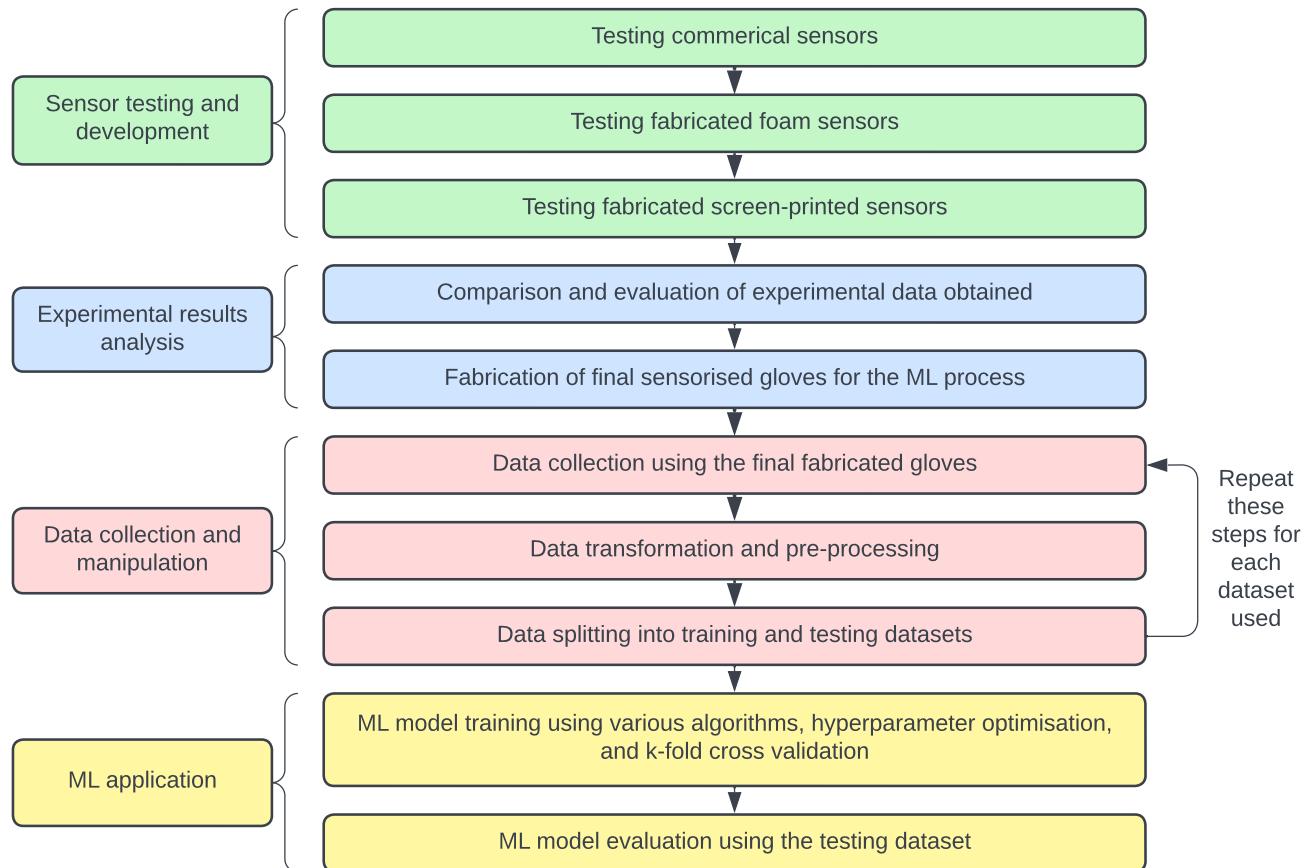


Figure 6: Overall project plan.

#### 3.2 Experimental Methods

Experimentally, the project aims to find ideal sensor types and configuration on gloves.

##### 3.2.1 Materials

- General

All tests require connecting sensors to a measuring device using wires. The method for attaching wires, if required, will be using conductive copper tape, and pins are clamped onto the wire ends.

- Foam Sensors

Doktor Power Magic Eraser is used as a cheap, readily available, melamine foam. A liquid comprised of CNF and Tannic acid is used to soak the foam in, embedding conductive particles throughout the structure. This cures at  $80^{\circ}\text{C}$  in an oven.

- Screen-Printed Sensors

These are printed using a conductive silver nanoparticle ink, DM-SIP-2002 by Dycotec Materials. This cures at  $80^{\circ}\text{C}$ , a temperature below the melting point of surgical gloves. It has an electrical resistivity value of around  $1.15 \times 10^{-6} \Omega \cdot \text{m}$ . A silver conductive adhesive, DM-SAS-10010 by Dycotec Materials, is used to adhere the copper tape, which cures at  $80^{\circ}\text{C}$ .

### 3.2.2 Fabrication

- Commercial Sensors

These sensors need adhering to the joint using an instant adhesive, for example, Loctite 406. Figure 7 shows the FlexiForce A201 sensor.



Figure 7: FlexiForce A201 Sensor.

- Foam Sensors

The methodology followed is the same as the group who developed them as pressure sensors [19], aiming to extend the capabilities into strain measurements. It can be seen in Figure 8a. As a result, they are covered using electrical tape, as shown in Figure 8b, and adhered to gloves using Loctite 406.

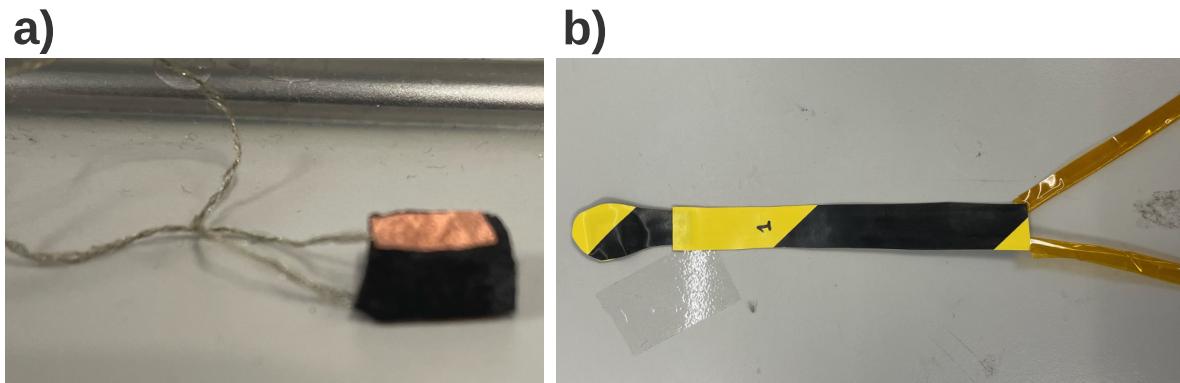


Figure 8: a) Foam CNF sensor, and b) foam CNF sensor after packaging.

- Screen-Printed Sensors

The fabrication process is shown in Figure 9a. It is difficult to screen print onto medical gloves due to their shape and surface porosity, so the roller and repetitions are used to maintain accuracy. The shape was designed on Fusion 360, with initial variables of 16mm length, 1mm width, and  $6\mu\text{m}$  thickness, providing a base resistance of  $\approx 3\Omega$ . The stencil is shown in Figure 9b, alongside the sensor in Figure 9c.

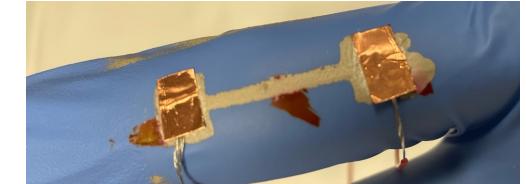
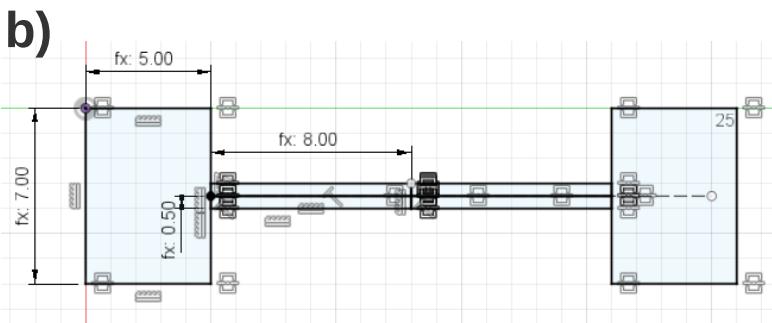
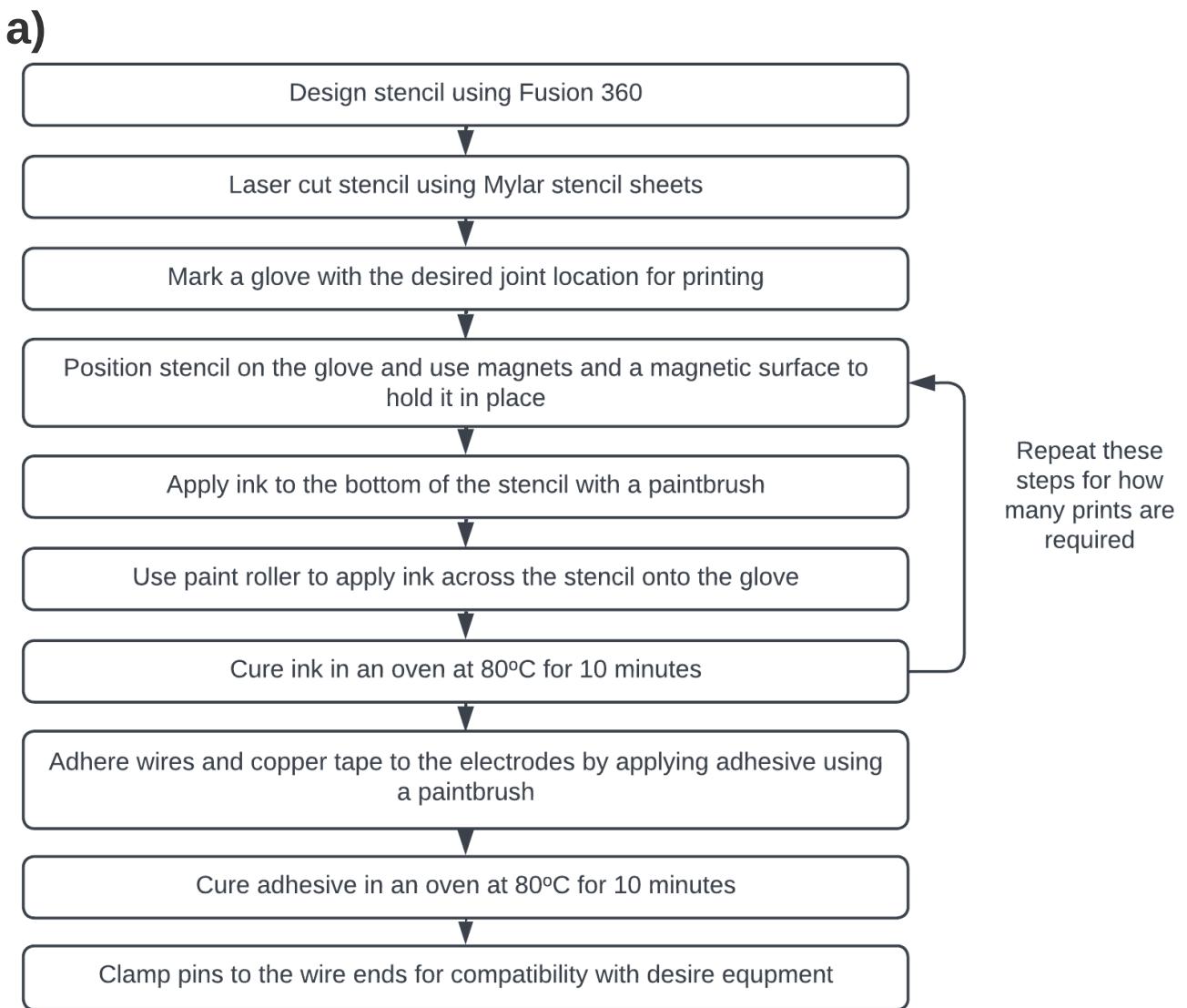


Figure 9: a) Flowchart of the screen-printed sensor fabrication process, b) stencil CAD drawing, and c) screen-printed sensor.

### 3.2.3 Instrumentation

- Lab Measurements (Hardware)

Sensor testing is recorded using various equipment. One is the Multimeter RSDM 3055. Due to the software used, it samples at  $\approx 3.3$  S/s. This is used to record resistance signals. The next is the NI USB 6212 DAQ Card, a data acquisition card from National Instruments with a sampling rate  $\approx 400$  S/s, which uses a variable controller gain to obtain time-dependent voltage signals. Last is an Arduino Uno with a voltage divider circuit, which has a sampling rate  $\approx 8$  S/s. This obtains accurate time-dependent voltage signals in accordance with Equation 6.

- Lab Measurements (Software)

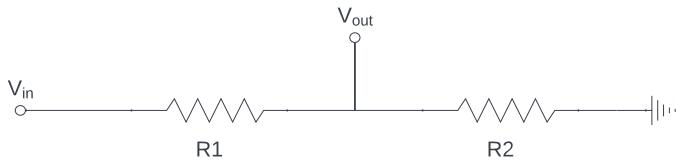
The multimeter and DAQ card record signals using NI LabView 2020 software. For the Arduino, a script on Python is used to record signals. Both save the data to .csv files.

- Voltage Divider Circuit

The circuit follows Equation 6, and a circuit diagram can be seen in Figure 10a, with R1 representing the sensor and R2 a fixed resistor. Using sensors, the varying resistance through bending at a joint will cause a measurable voltage change which can be classified. The base resistance through the sensor must be similar to the fixed resistor for sensitivity in measurements. As a result, base resistances are found for fabricated sensors using the Multimeter, and  $10\text{ k}\Omega$  resistors are used for FlexiForce sensors. This is recommended for gripping force which is most applicable [23]. Forming 5 of these circuits using a breadboard, each separate voltage output can be measured simultaneously. The constructed circuit is shown below in Figure 10b.

$$V_{out} = V_{in} * \frac{R_1}{R_1 + R_2} \quad (6)$$

a)



b)

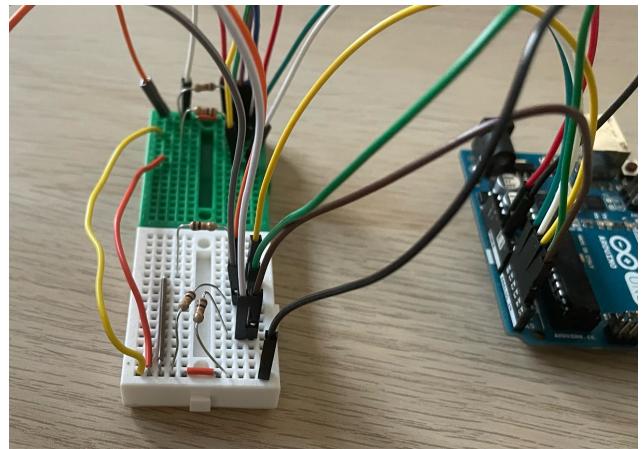


Figure 10: Voltage divider circuit a) diagram and b) image.

- Tensile Testing Setup

This uses a precision 3-axis (XYZ) motorised stage from Aerotech, driven using G-code. By fixing the sensor within the setup, as shown in the schematic and image in Figure 11a and 11b, increments can be set to stretch the sensor, so strain effects on electrical properties can be tested. The increment is set to 1mm displacement, with time steps of every 10-20 seconds, depending on how much the signal has settled.

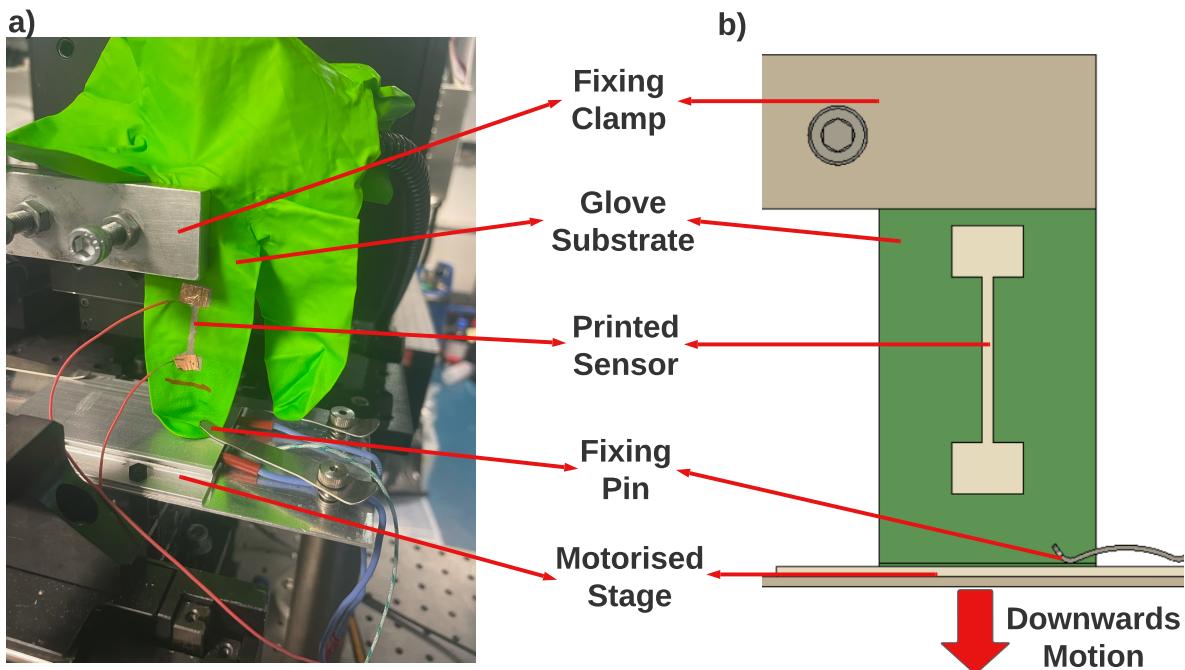


Figure 11: Tensile testing setup a) image and b) schematic.

### 3.2.4 Test Strategy

The project focuses on sensing the bending strain of finger joints. Medical gloves are used because they are cheap, easy to adhere to, and provide future clinical potential. Figure 12a shows the joints on a finger labelled where bending can be sensed. These joints can include a sensor on either the front or back of the finger. Figure 12b shows the numbering of each finger on the hand. Both labels will be used throughout the report.

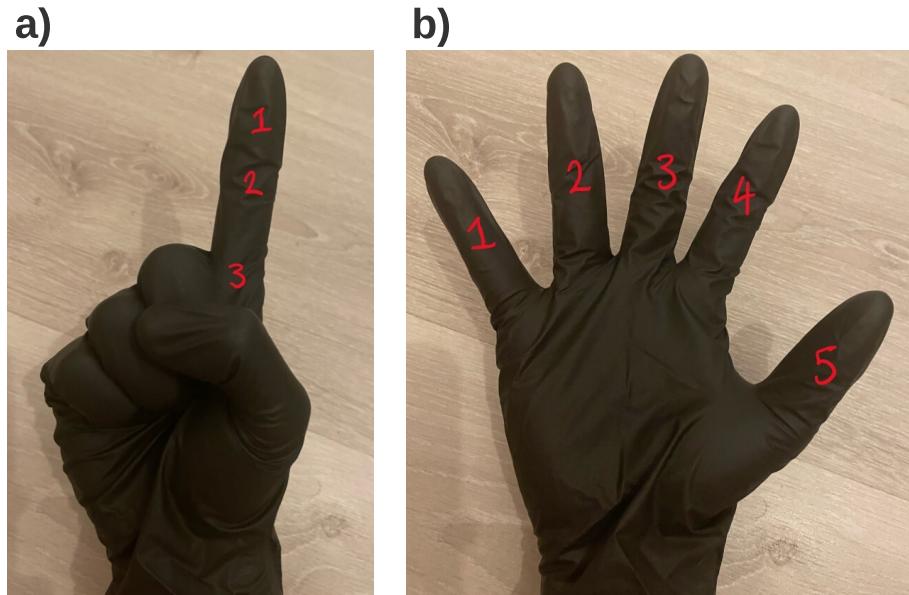


Figure 12: a) Finger showing joint numbering and b) hand showing numbering of each finger.

Testing will need to be undertaken to find the best type of sensor and their positions. Based on the sensors used, joint 3 is disregarded due to requiring a larger sensor and being less easy to adhere and keep in place, providing less consistency in testing.

Concepts tested on one finger only are transferable to others as they have the same joint configuration, where the thumb joint is assumed to provide similar signals to joint 2. The index finger (4) is chosen for all initial tests for consistency. All tests will be performed on the best-performing joint found from an initial proof of concept test for consistency and performance. The protocol for testing will be a full bend of the finger at consistent time intervals.

General tests for all sensors consist of proof of concept in compression (front) and tension (back), testing extended bending periods, base resistance tests (not required for FlexiForce A201 as their base is  $0\Omega$ ), and evaluating SNR for periodic bending. The screen-printed sensors also had different parameters tested, like the number of prints and strain variation of electrical properties in a uniaxial tensile test.

### 3.3 Computational Plan

The methods used for the ML section of the project are detailed as follows.

#### 3.3.1 Coding Language and Libraries

ML models are implemented in Jupyter Notebooks. Libraries used consist of Keras, XGBoost, TensorFlow, and scikit-learn. Keras provides an interface for building and training neural networks, TensorFlow offers efficient computation capabilities, XGBoost offers an optimised implementation of the XGBoost algorithm, and scikit-learn offers data preprocessing, algorithm implementation, and evaluation.

#### 3.3.2 Data Collection

Data is collected on the fifteen AmSL and fifteen ArSL signals from Section 2.5 for fifty frames while making the gesture and a mean is recorded for each finger. Each gesture is then repeated ten times. The result will be four distinct datasets, and combining the languages for each sensor gives a total of six datasets to be classified.

Using the data collected, a bar graph is plotted made of five bars for each gesture using the mean of the ten data points per finger. Error bars are then plotted showing the standard deviation in each measurement. This allows visual comparison of gestures and signals obtained for each finger, so conclusions can be made knowing discrepancies between different joints and sensors.

#### 3.3.3 Data Transformation

Collection of the data set performs most of the cleaning and pre-processing required, by finding means of the signals, and the data is labelled within the code, with the finger being sensed in columns and the gesture made in rows.

Gesture labels are digitised by assigning each an integer from 0-15 through one-hot encoding. Data is then stratified to ensure the same number of symbols in each set are used for the data split into training and testing data.

Feature scaling then transforms the features of the dataset so that they are all on the same scale, ensuring none are dominating when comparing similarity and finding the distance between data points. This is particularly important for sensors with varying signal ranges. The method used is normalisation, which scales the range to be from 0-1 by subtracting the features' minimum value and dividing by the total range for every data point.

#### 3.3.4 Data Splitting

The data is split into training and testing sets, with the testing set being unseen by the model. Three random splits will be used to test the effect on algorithm performance:

- (1) 80% training, 20% testing.
- (2) 70% training, 30% testing.
- (3) 60% training, 40% testing.

### 3.3.5 Algorithms

Based on the algorithm review in Section 2.4, the algorithms chosen for the ML comparison study are listed below.

- **SVM:** SVM finds a separating boundary between the types of data. This problem involves non-linearly separable data with multiple classes, so kernelised multi-class SVM is used, where a kernel is a hyperparameter that measures the similarity between data points.
- **kNN:** kNN arranges known data into a space defined by selected features, and new data is classified depending on the classes of the 'k' closest data points. The 'k' chosen represents the number of nearest neighbours for the algorithm to consider when a prediction is made.
- **RF:** RF trains multiple decision trees on different subsets of the data provided, an approach known as bagging, and then takes an average of the predictions that the trees find to make a final prediction.
- **ANN:** ANN is the simplest class of neural networks. They simulate the behaviour of human brain neurons to learn from input data.
- **XGBoost:** XGBoost uses decision trees built additively to improve model performance, an approach known as boosting. It avoids overfitting by constraining tree size.

These are chosen mainly for their success in different applications or their frequency of use. Most notably, NB and DT are disregarded, because, despite their frequent use, they frequently give worse results.

### 3.3.6 Hyper-Parameter Optimisation

All ML models used have hyperparameters, which are not learnt through data but are set in the code to determine the learning behaviour of the model. Hyperparameters need different values depending on the data set and the problem being solved. The grid search method used consists of specifying a range of values for each hyperparameter and training the model for every single combination. As a result, every point in the hyperparameter space needs to be trained and evaluated, therefore this is the most systematic and thorough method with the highest computational cost. Considering the relatively small data set used and the computational power accessible, this is the best method available for accuracy.

### 3.3.7 Validation

K-fold cross-validation provides a robust evaluation of the model. The data is divided into  $k$  'folds', where each fold is used as a validation set in the training once, and the rest is used as training data. This means that there would be  $k$  validations of the algorithm. The model is therefore trained and validated on different subsets of the data.  $k = 4$  is chosen, which is a compromise between computational time and robustness of the training and evaluation.

### 3.3.8 Evaluation Performance Metrics

The algorithms will be evaluated using six main performance metrics.

- **Accuracy:** Percentage of correct predictions.
- **Precision:** Percentage of correct predictions relative to total predictions.
- **Recall:** Percentage of correct predictions relative to the actual number of results.
- **F1 Score:** Weighted harmonic mean of precision and recall.
- **Confusion Matrix:** Shows the accuracy for each gesture by plotting the frequency of actual against predicted gesture classifications.
- **Learning Curves:** Graphs showcasing model performance on training and validation over the training data used. Useful for diagnosing overfitting, underfitting, and evaluating learning and generalization. High scores and a small gap between curves indicate a balanced model.

The best models exhibit excellent performance, with learning curves suggesting a balanced, well-fitted model.

## 4 Signal Characterisation of Sensorised Gloves

This details results and analysis of experiments using sensorised gloves.

### 4.1 Commercial Sensors

This details the testing and analysis of FlexiForce A201 sensors.

#### 4.1.1 All Joint Tests

Figure 13 shows the signals produced from the FlexiForce A201 sensors on each joint position specified in Section 3.2.4 with periodic bending over a 20-second interval. The DAQ card was used for recordings, which tunes ranges due to the gain. The front joints show very small ranges, with a maximum value between both tests  $\approx 0.57\text{V}$ , showing that these sensors are not ideal in compression. Back joints show much larger, distinctive ranges, suggesting tension is ideal for these sensors. Despite the same base signal, joint 1 has a peak value  $\approx 0.69\text{V}$ , which is lower than joint 2 which has a maximum  $\approx 0.73\text{V}$ .

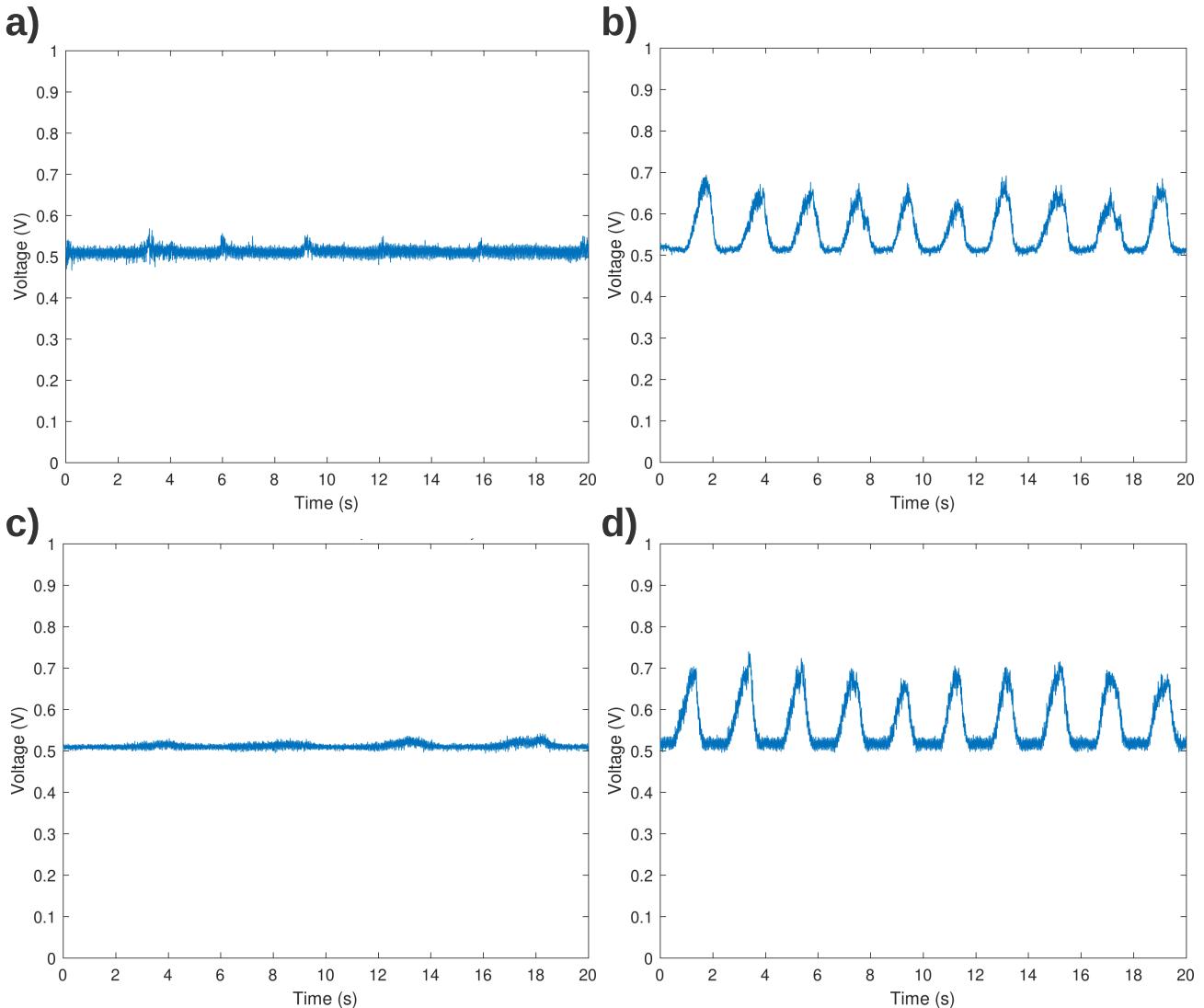


Figure 13: Initial FlexiForce A201 tests for consistent bending on a) joint 1 front, b) joint 1 back, c) joint 2 front, d) joint 2 back.

Anatomically, joint 1 cannot be bent uniquely, as it requires the bending of joint 2. Joint 2 can also naturally bend further. Along with the results, this confirms the use of joint 2 exclusively.

The test in Figure 14 shows three occasions of holding the joint bent for an extended period with the chosen joint. Results show some noise and some time to settle, with an average peak standard deviation  $\approx 0.015V$ . Low deviation when a gesture is made concludes the feasibility of sign language classification.

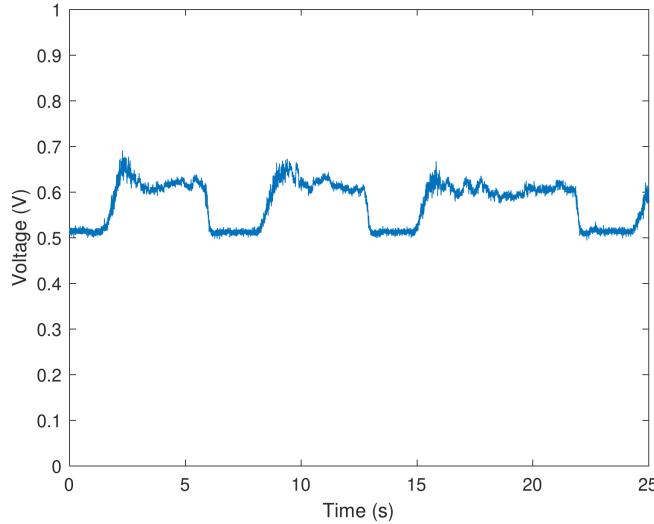


Figure 14: FlexiForce A201 best performing joint test for extended periods of strain.

#### 4.1.2 Evaluation of Signal to Noise Ratio for Periodic Tests

Figure 15 and Table 2 show the signals for periodic bending from which SNR was calculated. From observing Table 2, all trough SNRs are equal to 0. This is because the sensors do not produce any signal in their base state, and there is no noise, which is beneficial. The peak SNR values are mostly similar, with the index and ring fingers slightly outperforming others. This is related to the higher signals they produce with lower noise. There can be many reasons for deviations, like small errors in the positioning of the sensors and natural geometry of joints producing different signal ranges. No sensor surpasses the 3V signal range despite 5V input, suggesting low sensor sensitivity, supported by relatively low SNR values. The sensitivity however is adequate for the requirement.

Table 2: FlexiForce A201 peak and trough SNR signals.

Finger	Peak Mean (V)	Peak SD (V)	Trough Mean (V)	Trough SD (V)	Peak SNR	Trough SNR
Pinky	2.067	0.189	0	0	10.921	0
Ring	2.041	0.103	0	0	19.779	0
Middle	1.819	0.186	0	0	9.782	0
Index	2.127	0.135	0	0	15.750	0
Thumb	1.689	0.171	0	0	9.875	0

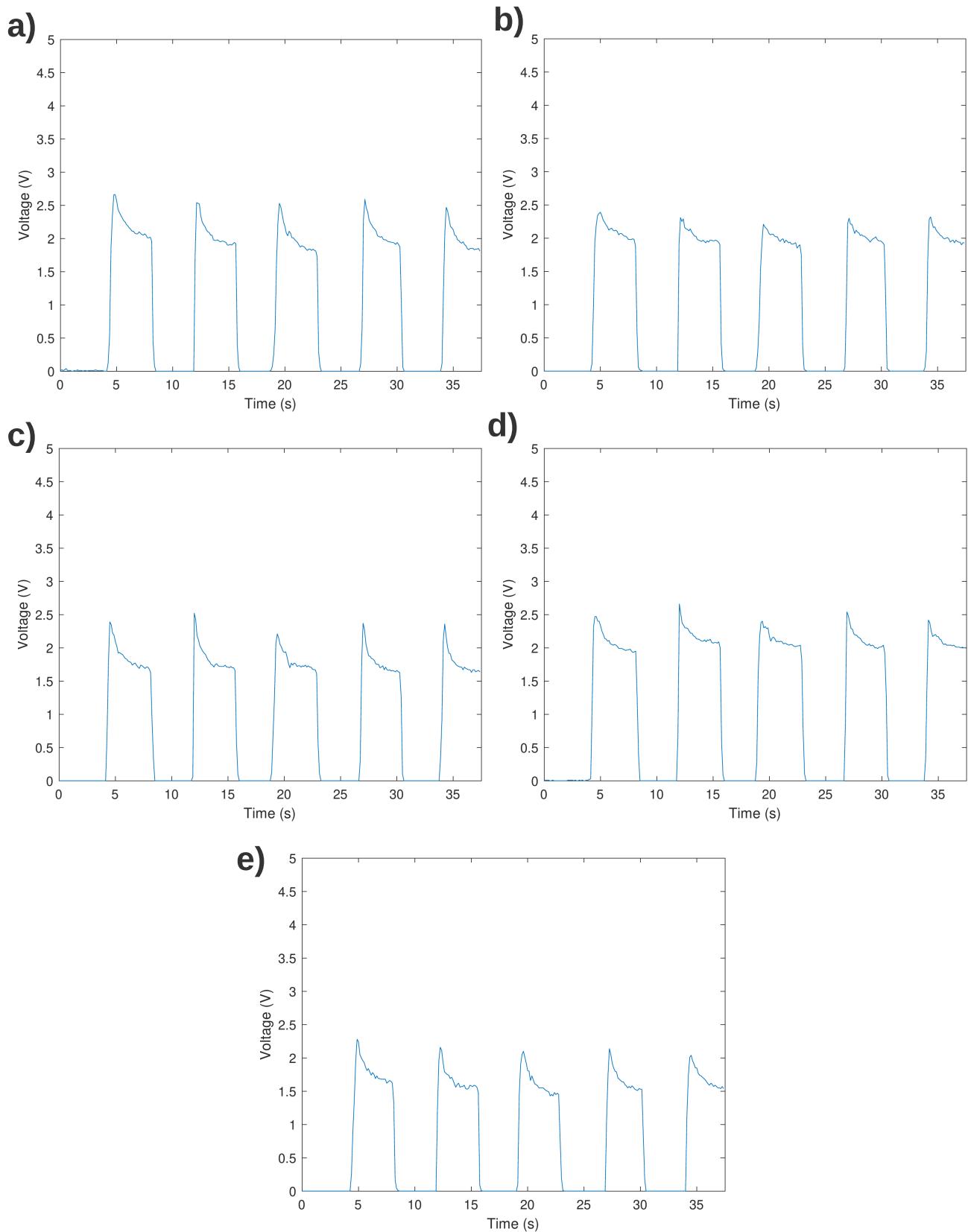


Figure 15: FlexiForce A201 SNR test signals for a) little, b) ring, c) middle, d) index, e) thumb. Tests involve 5 equal periodic bends over a 37.5s time interval.

## 4.2 Foam Sensors

This details testing and analysis of foam sensors.

### 4.2.1 Tension and Compression Tests

Figure 16 shows initial proof of concept tests on foam sensors in compression (a) and tension (b) with consistent bending over a 22-second interval. While both have similar trough values, the front has a maximum peak value  $\approx 3.1\text{V}$  and a minimum peak value  $\approx 2.3\text{V}$ . The back has a maximum peak value  $\approx 3.6\text{V}$  and a minimum peak value  $\approx 3.3\text{V}$ . Foam sensors, therefore, perform better in tension, with higher ranges and more consistent results.

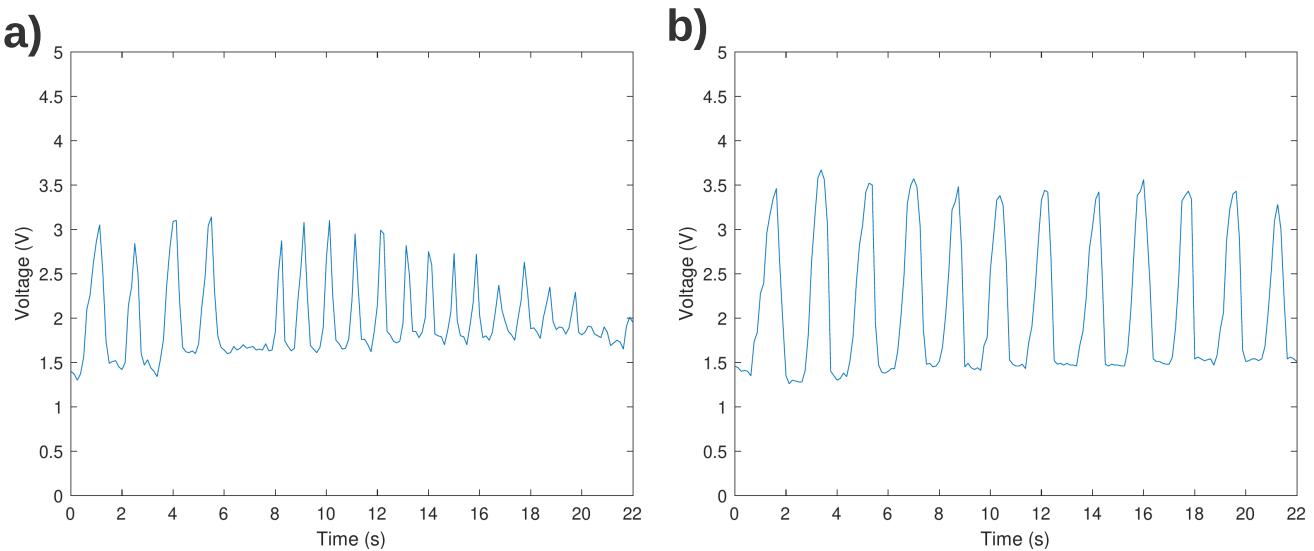


Figure 16: Initial foam sensor tests for consistent bending on joint 2 a) front and b) back.

### 4.2.2 Reproducibility Resistance Tests

Five sensors were fabricated, with base resistances shown in Table 3 along with resistor values for the circuit. Sensors all have resistances in the same order of magnitude so some repeatability can be observed. However, there is still a significant enough difference to conclude that there is difficulty in reproducing the sensors to have the same properties, despite sensors 2, 3, and 5 being similar. For classification, the changes depend on the voltage divider circuit, so this should not be problematic, as all resistors used have very similar resistance, as can be observed.

Table 3: Foam sensor base resistance values.

Sensor Number	Resistance ( $\Omega$ )	Resistor used in circuit ( $\Omega$ )
1	750	750
2	330	330
3	200	220
4	1000	1000
5	350	330

#### 4.2.3 Evaluation of Signal to Noise Ratio for Periodic Tests

The sensors above were adhered to the gloves in the order above, with 1 adhered to the pinky and 5 adhered to the thumb. Figure 17 and Table 4 show the signals for periodic bending from which SNR was calculated. From observing Table 4, the peak SNR values are mostly similar, with the index and ring fingers outperforming others again, for similar reasons as discussed in Section 4.1.2. They can also be observed to have the largest ranges, with peak maximums  $>4.5V$ , with others being  $<3V$ . The thumb however shows to have high noise signals compared to other sensors, although still an adequate SNR. Figure 17e shows the deviations at the peaks being higher than other sensors. Trough SNR is lower for all sensors, with the middle finger giving the highest, showing more noise dominating the signals than for the peaks. Overall, the SNR values suggest good sensor sensitivity with low settling time, and the sensors are adequate for classification despite possible fabrication inconsistencies. Normalisation should fix result range differences.

Table 4: Foam sensors peak and trough SNR signals.

Finger	Peak Mean (V)	Peak SD (V)	Trough Mean (V)	Trough SD (V)	Peak SNR	Trough SNR
Pinky	2.543	0.024	1.372	0.044	107.728	31.462
Ring	4.436	0.037	0.862	0.027	120.084	31.692
Middle	2.506	0.025	1.825	0.032	101.749	57.594
Index	4.420	0.027	1.197	0.031	161.555	38.466
Thumb	2.468	0.064	0.551	0.019	38.770	28.307

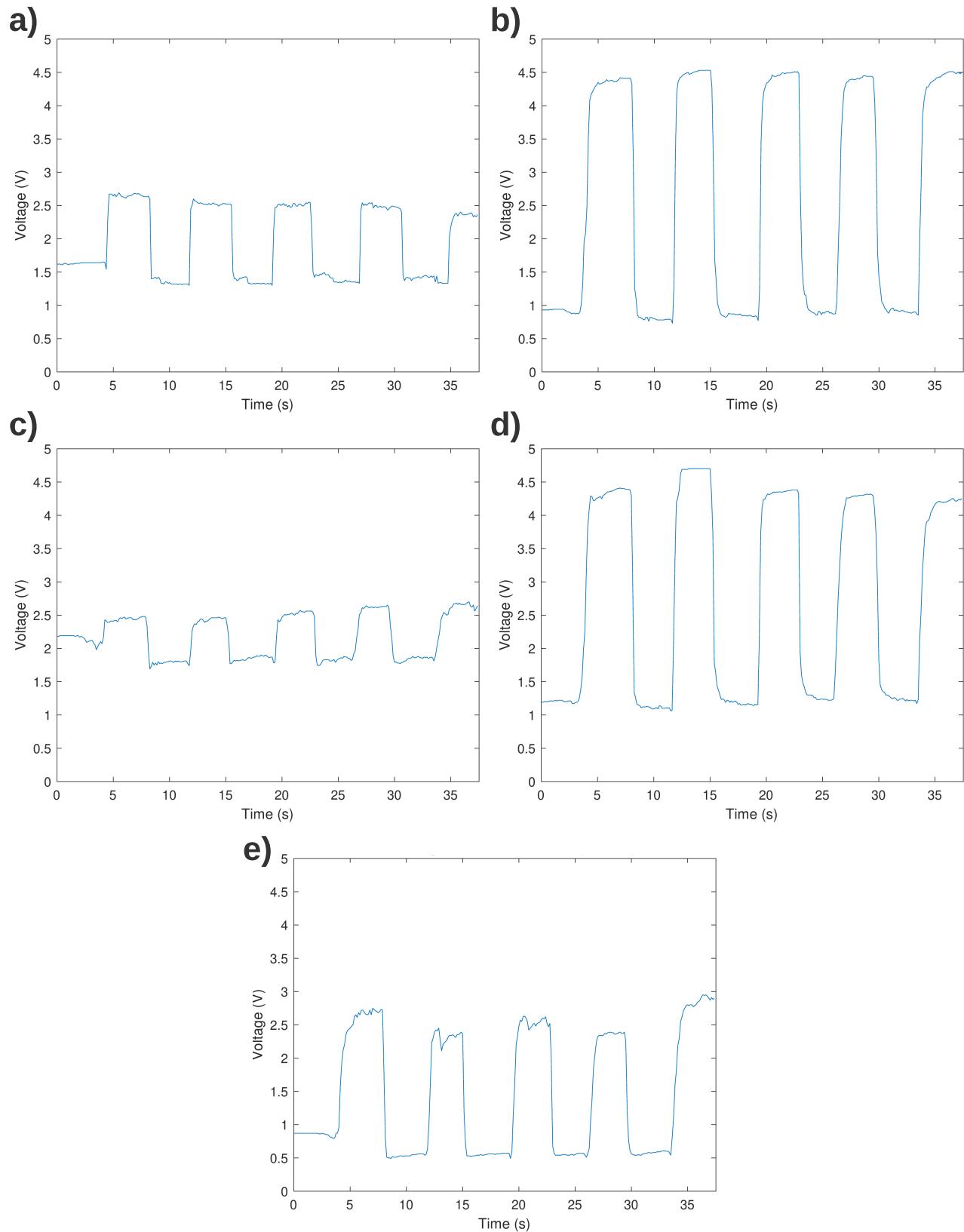


Figure 17: Foam sensors SNR test signals for a) little, b) ring, c) middle, d) index, e) thumb. Tests involve five equal periodic bends over a 37.5s time interval.

### 4.3 Screen-Printed Sensors

This details testing and analysis of screen-printed sensors.

#### 4.3.1 Tension and Compression Tests

Figure 18 shows initial proof of concept tests on screen-printed sensors in compression (a) and tension (b) for periodic bending. The DAQ cards gain alters the peak values. The front shows no signals at all, as any bending in this direction causes electrodes to short circuit, therefore this direction is not feasible. The back shows signals to instantly saturate to the maximum range when bent, meaning the sensors are highly sensitive, which can be investigated further.

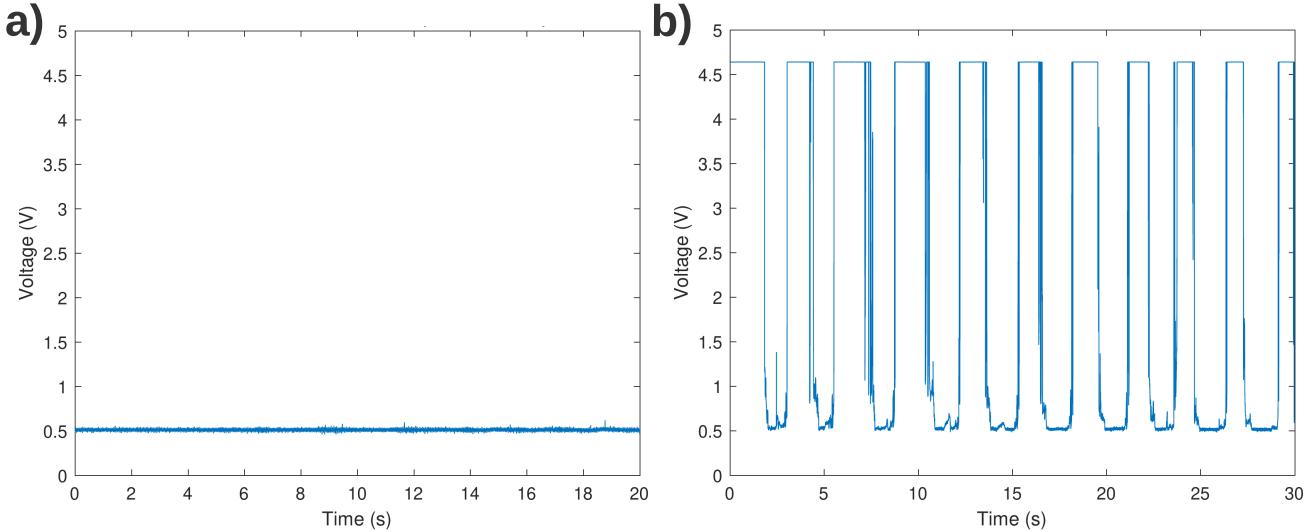


Figure 18: Initial tension and compression screen-printed sensor tests on the joints a) front and b) back.

#### 4.3.2 Reproducibility Resistance Tests

Table 5 shows the base resistances of four screen-printed sensors. The  $392\Omega$  range shows that there are inconsistencies in the printing technique when using a single print. This is likely due to the porosity of the paint roller and glove used, which may not provide a good surface with a single print. The actual base resistance has been designed for  $3\Omega$ . Values fluctuate a lot when testing from different points on the electrodes, further implying the points discussed previously.

Table 5: Screen-printed sensors base resistances.

Sensor Position	Base Resistance ( $\Omega$ )
1	400
2	8
3	14
4	41

Table 6 shows base resistances for screen-printing multiple layers, using 1, 2, and 3 prints. The 1 print sensor is over  $5\Omega$  higher than the others and also showed more fluctuation while testing. This suggests that the optimal is using two prints as it preserves time while optimising sensor performance, whereas the 3 print performance benefits are not worth the fabrication time.

Table 6: Screen-printed sensors base resistances (multiple prints).

Sensor Position	Base Resistance ( $\Omega$ )
1 Print	8.8
2 Prints	2.6
3 Prints	2.5

Figure 19 shows a proof of concept test for the 2 print sensor within the finger bending application, holding the gesture three times for an extended period. The resistances are recorded using a multimeter and show very high ranges. Once the strain is applied, the sensor's resistance has a range of  $\approx 200\Omega$ . The peaks also range from  $\approx 550\Omega$  initially to  $\approx 730\Omega$  finally. This shows very high sensitivity, but potential microstructural changes of conductive pathways when strained which leads to repeatability and extended use issues.

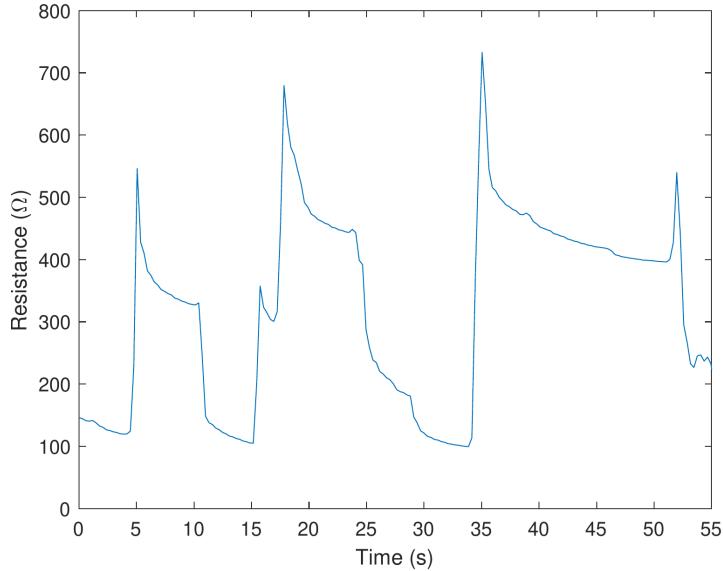


Figure 19: Screen-printed sensor double print with periodic bending test.

#### 4.3.3 Uniaxial Tensile Tests

Figure 20 shows uniaxial tensile tests completed for the different number of prints tested, the method described in Section 3.2.3. 2 print signals stay within a 0.5V range with low fluctuation. 3 print signals have more fluctuation within a 1V range. 1 print shows high variations and fluctuation within a 6V range. Overall, the results suggest 2 prints as optimal, and that the ink is suitable for strain applications. Differences between 2 and 3 prints may be due to issues with clamping the glove within the equipment, or even due to sensor design, where strain of electrodes affects measurements due to not being negligible. 3 prints may also lead to an undesired thickness increase.

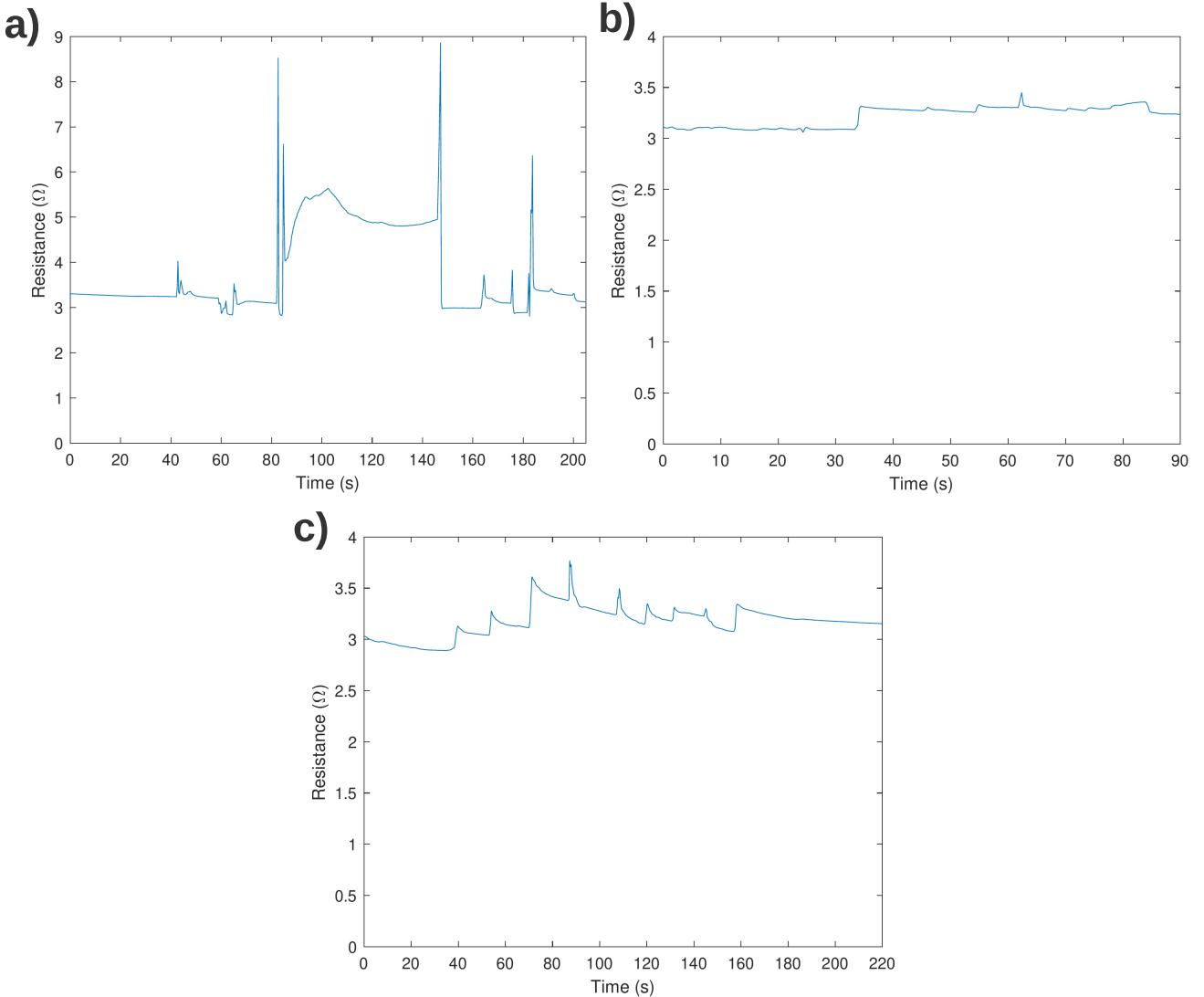


Figure 20: Uniaxial tensile tests using periodic incremental 1mm extension every 10-20 seconds for a) 1 print, b) 2 prints, and c) 3 prints.

#### 4.4 Decisions from Final Data Analysis

The commercial and foam sensors are consistent, reliable, and reproducible, so will be used for the classifier. The final gloves can be observed in Figure 21. This results in four distinct datasets to collect, due to two languages and two sensor types. Each data point will consist of five signals, from a sensor on joint 2 of each finger.

The screen-printed sensor requires more development before use, which time does not permit, but the design has the most potential. Due to their thinness, they would not interfere with the tactile perception of the user, which is an important requirement for wider use of the concept [50]. Furthermore, it does so while being inexpensive and relatively simple to fabricate with reproducible base resistances.

**a)****b)**

Figure 21: Final glove fabrications with a) FlexiForce A201 sensors, and b) Foam sensors.

## 5 Machine Learning

This section details the entire ML process.

### 5.1 Statistical Exploration of Datasets

Datasets must first be scrutinised.

#### 5.1.1 Signals

Plotting bar charts for each signal allowed for many noticeable observations, and these are all found in the GitHub repository. Firstly, both sensor types can have large deviations for fully bent signals, which is likely attributed to human error in performing the gesture or positioning the sensors each time. A good example is the Dhaal symbol in Figure 22. Large deviations can be observed on the index and ring fingers in the error bars. These sensors also have the largest range and peaks, so deviations in their signals may have a larger magnitude but are of a similar proportion, as they are known to have similar SNR.

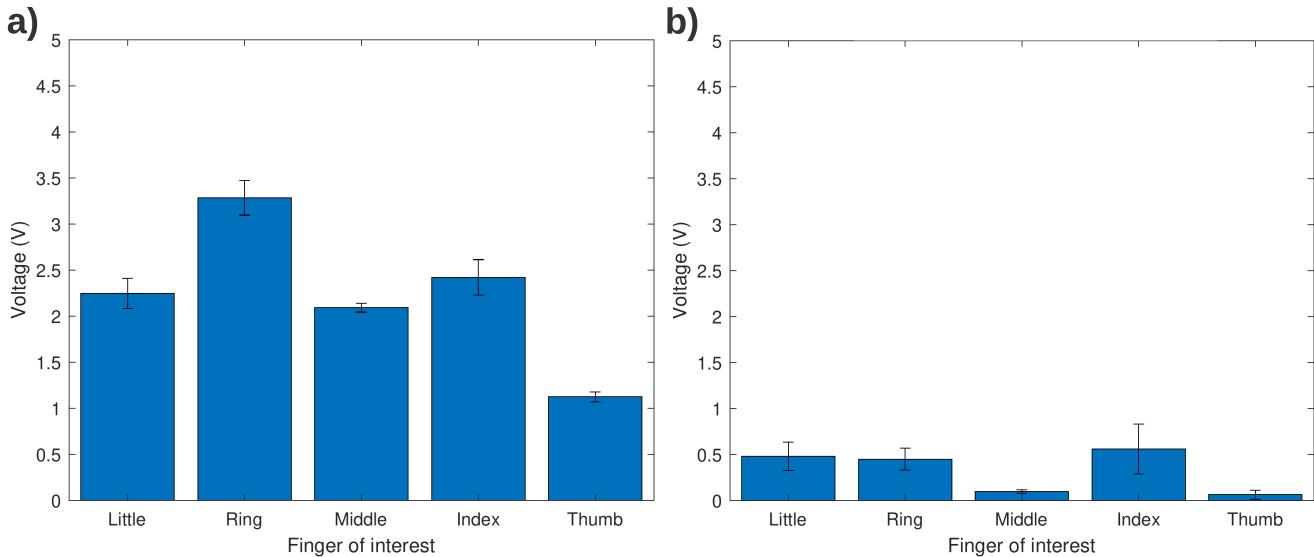


Figure 22: Bar charts for 'Dhaal' gesture for a) Foam sensors, and b) FlexiForce sensors.

Another noticeable observation was the lack of deviation of the 'Haa' signal from the baseline signal for both sensors, shown in Figure 23. This is due to a lack of bending in the required joint for the gesture, which requires joint 3 (bottom) bending only, which is not sensed. As a result, FlexiForce sensors pick up no signal at all for this gesture. The 'Haa' gesture was removed from the dataset in order to ensure gestures are distinguishable for classifier accuracy. This is a limitation of the final fabricated glove.

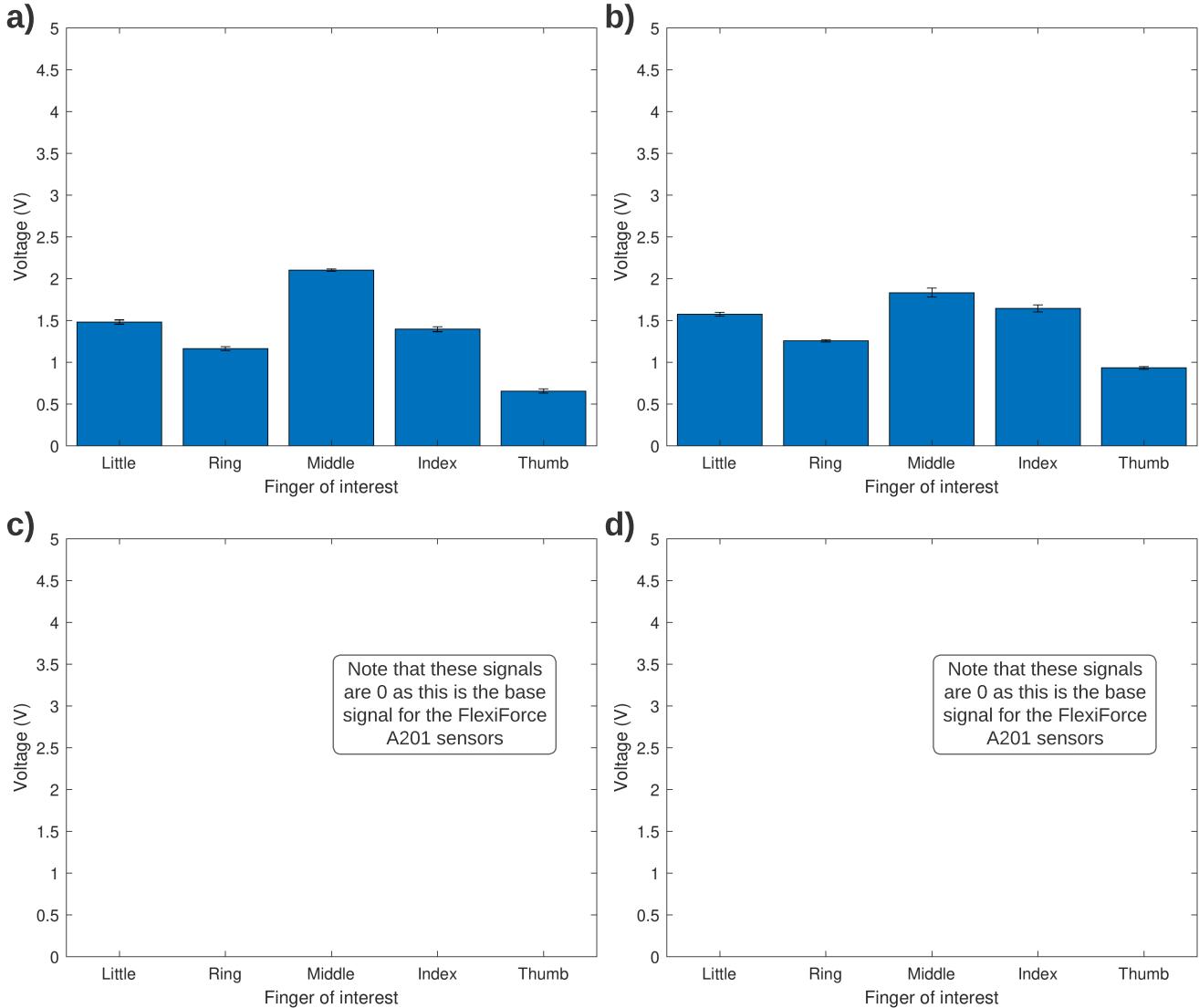


Figure 23: Bar charts for a) Foam sensor baseline, b) Foam sensor 'Haa', c) FlexiForce sensor baseline, and d) FlexiForce sensor 'Haa'.

A difference in the sensor signals can be observed throughout all signals, in which the peak, trough, and signal ranges alter slightly after extended use, suggesting sensitivity changes. As a result, the signals for later gestures appear to have consistently lower peaks for similar amounts of bending, and this is present for both FlexiForce and foam sensors. This should not affect the accuracy of the model as the gestures are still very distinct, made at the same time, and normalised, but may cause issues if used for live recognition due to different magnitudes of signals being produced for the same gestures. The FlexiForce trough is always at zero voltage, so this sensor is affected less. An example of this is the small deviation between Figure 23a and 23b despite that they should have equal signals.

### 5.1.2 Algorithm Training

SVM, RF, kNN, and XGBoost are non-parametric, meaning hyperparameters are altered using grid search. However, ANN uses iterative training and is parametric, therefore hyperparameters control the number of iterations used for training, which needs to be high enough so it converges to a number. An example of this for ANN is shown in Figure 24. It shows by increasing the number of epochs, or iterations, the accuracy of the model increases until a point where it begins to converge to the maximum accuracy, which usually occurs within 1000 epochs.

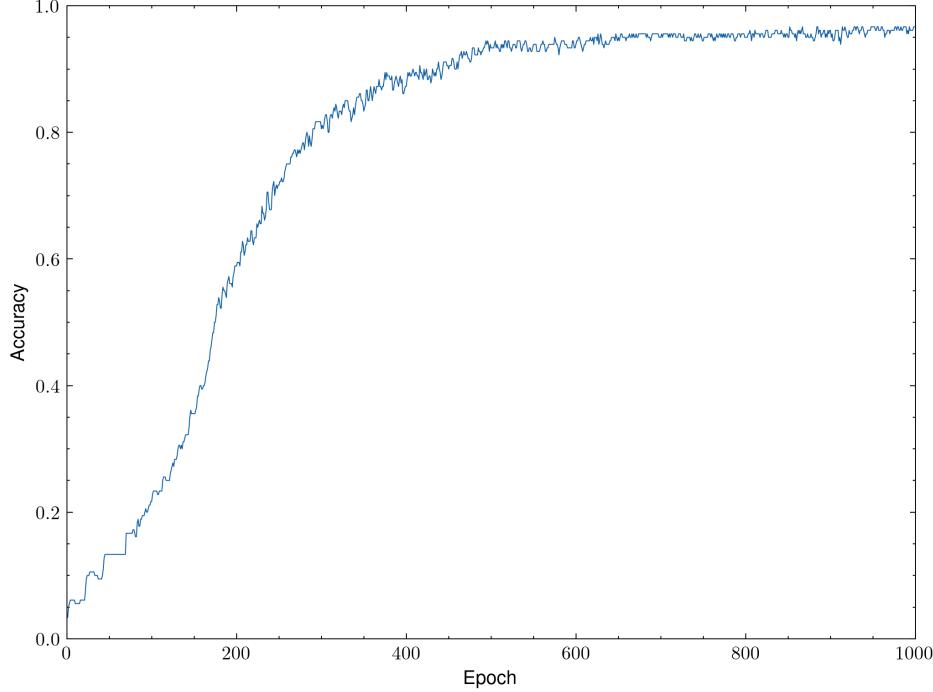


Figure 24: Increasing epoch effect on the accuracy of the ANN(2) algorithm for foam combined dataset.

## 5.2 Machine Learning Classification Results

The results obtained are from algorithm performance on testing data. They are presented for foam sensors then FlexiForce sensors, detailing AmSL, ArSL, and combined in that respective order. Tests were conducted for all cases as explained in Section 3.3.4, and it showed some disparity between the splits. For brevity, classification reports only include the best performing data splits, coloured from purple (1.0) to white (0.80). Confusion matrices are displayed for only the best performing algorithms for each dataset.

Full classification reports can be seen in Appendix B.

### 5.2.1 Foam Sensor with Singular Language Dataset

Table 7 shows the classification report for foam sensor singular datasets. kNN(2) and ANN(2) perform best for AmSL and ArSL respectively, showing optimal performance alongside the excellent fitting of the model within learning curves, shown later in Figure 31a and 31c. Figures 25 and 26 show the confusion matrices, where only the 'Baa' sign was misclassified, which occurred for many algorithms on the ArSL dataset. Most algorithms exhibit excellent performance, XGBoost being an exception which performs consistently worse on all metrics. ANN(1) and kNN(1) perform much worse for ArSL, suggesting incompatibility with this dataset. Many high-performing algorithms are overfitting, so the performance would be limited on new, varied data.

Table 7: Best-performing classification report of foam sensors on singular datasets.

Algorithm	Foam Sensor AmSL				Foam Sensor ArSL			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
SVM (2)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
kNN (2)	1.00	1.00	1.00	1.00	0.98	0.98	0.98	0.98
RF (3)	1.00	1.00	1.00	1.00	0.98	0.99	0.98	0.98
ANN (2)	1.00	1.00	1.00	1.00	0.98	0.98	0.98	0.98
XGBoost (1)	1.00	1.00	1.00	1.00	0.93	0.96	0.93	0.93

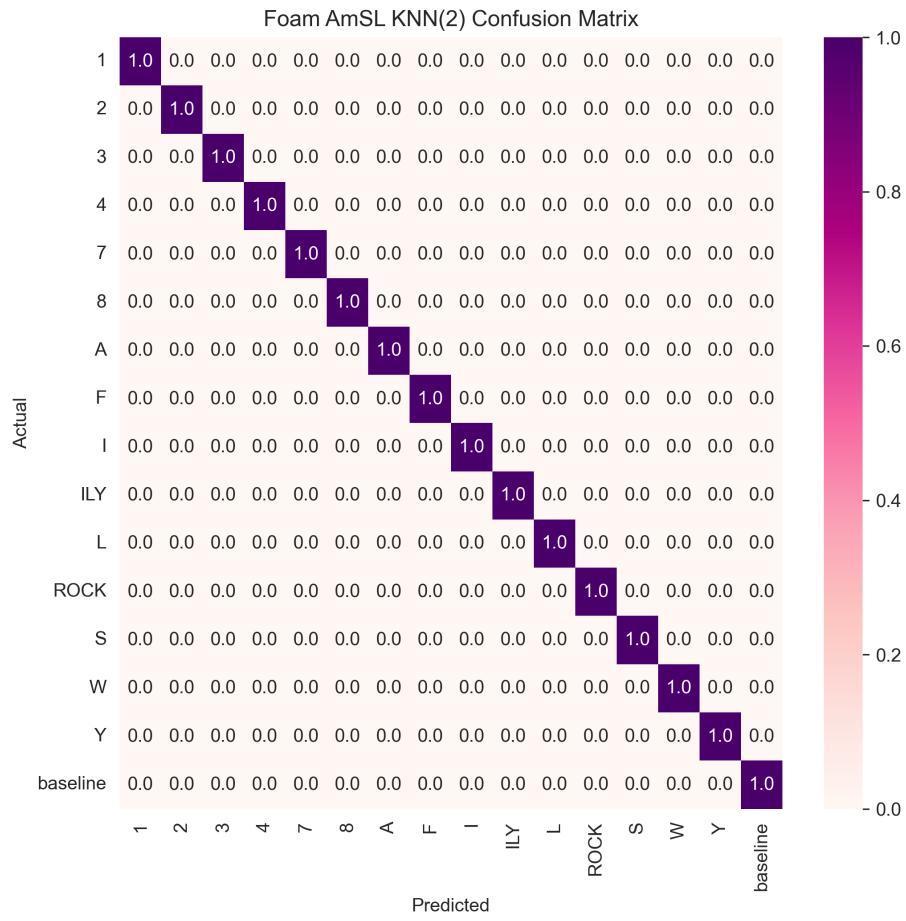


Figure 25: Foam AmSL best-performing algorithm confusion matrix (kNN(2)).

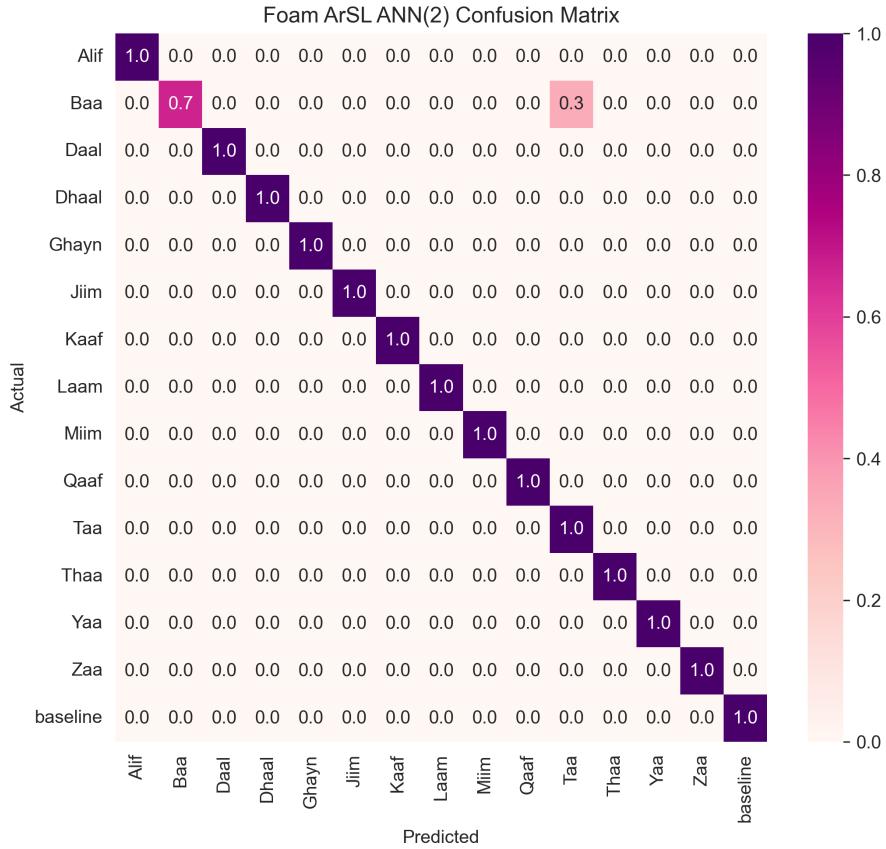


Figure 26: Foam ArSL best-performing algorithm confusion matrix (ANN(2)).

### 5.2.2 Foam Sensor with Combined Datasets

Table 8 shows the classification report for FlexiForce A201 combined dataset. kNN(1) performs best in this case, showing optimal performance alongside the excellent fitting of the model within learning curves, shown later in Figure 31e. Figure 27 shows the confusion matrix for the algorithm, with only one misclassification occurring of the '3' gesture. Due to good model fitting, this suggests a bias and limited dataset as some gestures are repeated between languages, although not the one misclassified, so different misclassifications should occur. Most algorithms exhibit excellent performance, XGBoost (2) and (3) being exceptions which perform worse on all metrics, and ANN, kNN, and XGBoost appear to be affected most by less training data.

Table 8: Best-performing classification report of foam sensors on combined dataset.

Foam Sensor Combined				
Algorithm	Accuracy	Precision	Recall	F1-Score
SVM (1)	1.00	1.00	1.00	1.00
kNN (1)	0.98	0.99	0.98	0.98
RF (1)	1.00	1.00	1.00	1.00
ANN (2)	0.96	0.96	0.96	0.95
XGBoost (1)	0.95	0.97	0.95	0.95

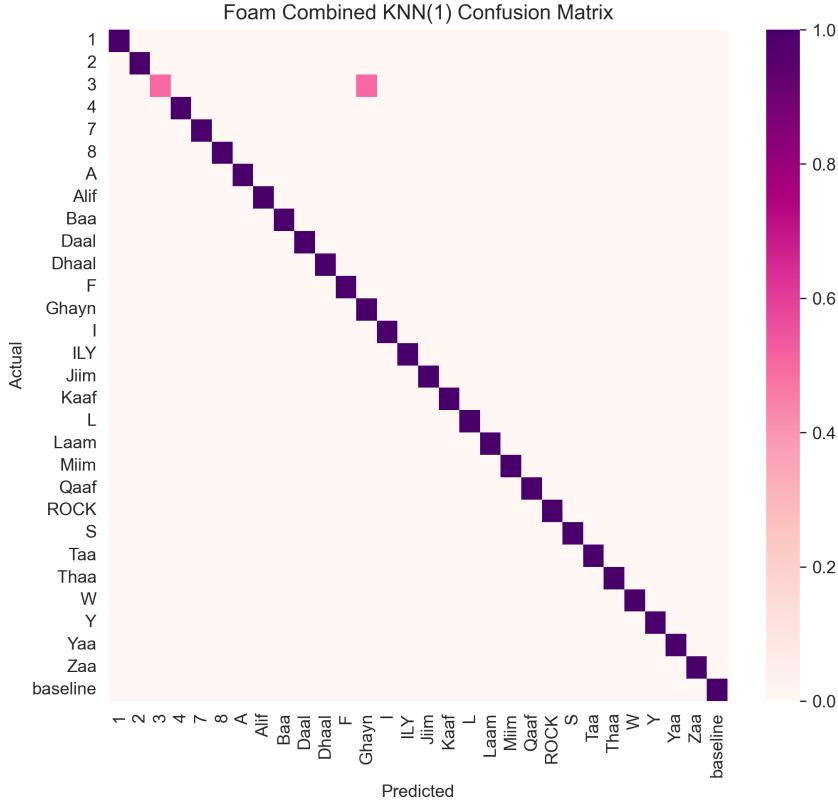


Figure 27: Foam combined best-performing algorithm confusion matrix (kNN(1)).

### 5.2.3 FlexiForce Sensor with Singular Language Dataset

Table 9 shows the classification report for FlexiForce A201 singular datasets. kNN(2) and ANN(2) perform best for AmSL and ArSL respectively, showing optimal performance alongside the excellent fitting of the model within learning curves, shown later in Figure 31b and 31d. Figures 28 and 29 show the confusion matrix for each algorithm, where only the 'Zaa' sign was misclassified, which occurred for many algorithms on the ArSL dataset. Most algorithms exhibit excellent performance, XGBoost being an exception which performs worse on all metrics. However, most algorithms are overfitting in this case, so the performance would be limited on new, varied data.

Table 9: Best-performing classification report of FlexiForce A201 on singular datasets.

Algorithm	FlexiForce Sensor AmSL				FlexiForce Sensor ArSL			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
SVM (3)	1.00	1.00	1.00	1.00	0.98	0.99	0.98	0.98
kNN (2)	1.00	1.00	1.00	1.00	0.93	0.96	0.93	0.93
RF (3)	1.00	1.00	1.00	1.00	0.98	0.99	0.98	0.98
ANN (2)	1.00	1.00	1.00	1.00	0.98	0.98	0.98	0.98
XGBoost (3)	0.92	0.94	0.92	0.92	0.92	0.93	0.92	0.92

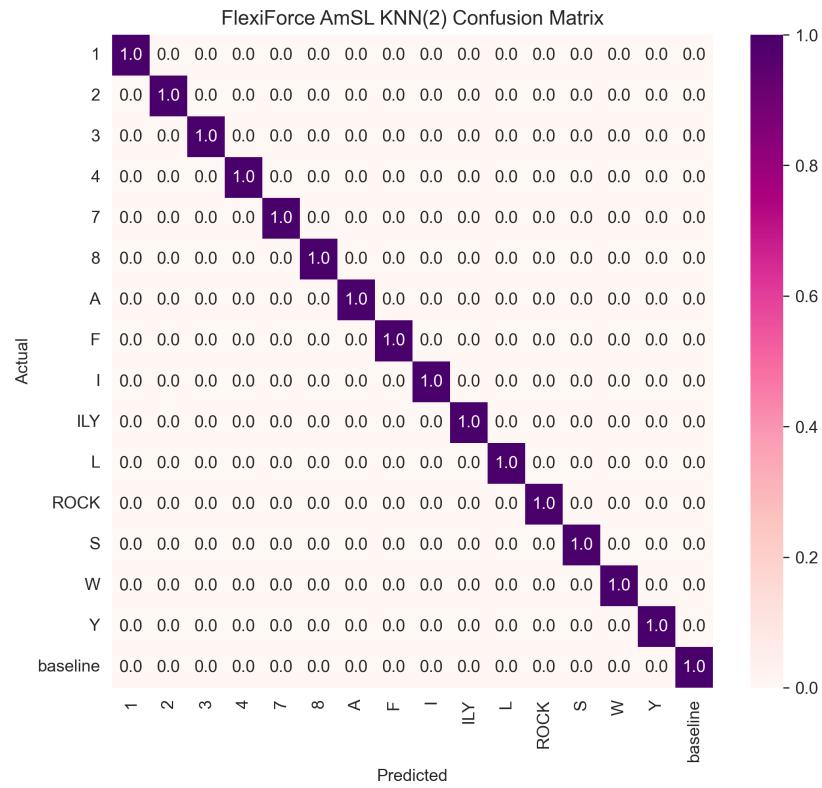


Figure 28: FlexiForce AmSL best-performing algorithm confusion matrix (kNN(2)).

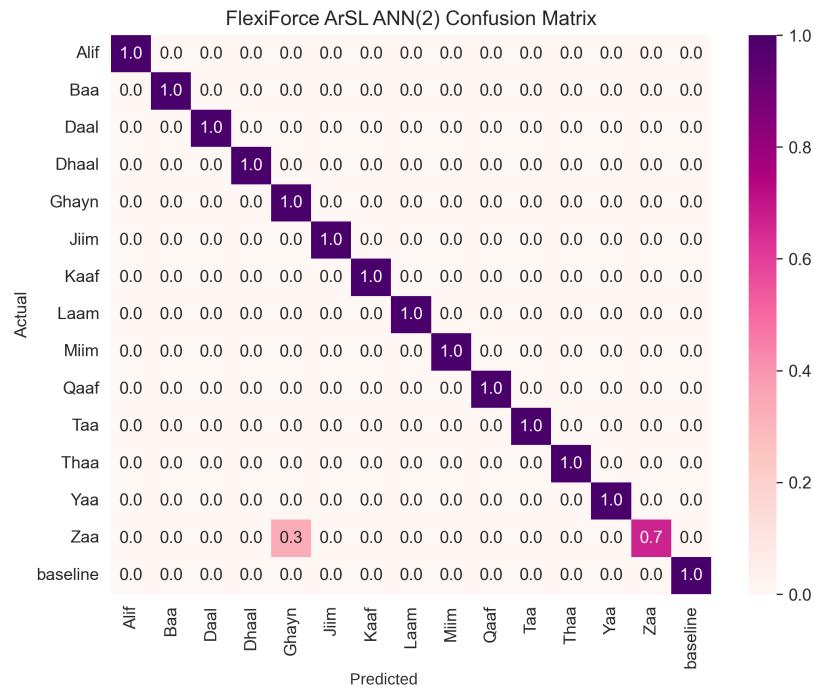


Figure 29: FlexiForce ArSL best-performing algorithm confusion matrix (ANN(2)).

### 5.2.4 FlexiForce Sensor with Combined Datasets

Table 10 shows the classification report for FlexiForce A201 combined dataset. kNN(1) performs best in this case, showing optimal performance alongside the excellent fitting of the model within learning curves, shown later in Figure 31f. Figure 30 shows the confusion matrix for the algorithm, with no misclassifications occurring. Due to good model fitting, this suggests a bias and limited dataset as some gestures are repeated between languages, so misclassifications should occur. Most algorithms perform well, XGBoost being an exception which performs worse on all metrics, and ANN appears to be affected most by less training data.

Table 10: Best-performing classification report of FlexiForce A201 on combined dataset.

FlexiForce Sensor Combined				
Algorithm	Accuracy	Precision	Recall	F1-Score
SVM (1)	1.00	1.00	1.00	1.00
kNN (1)	1.00	1.00	1.00	1.00
RF (1)	0.98	0.99	0.98	0.98
ANN (1)	0.97	0.98	0.97	0.96
XGBoost (1)	0.88	0.88	0.88	0.87

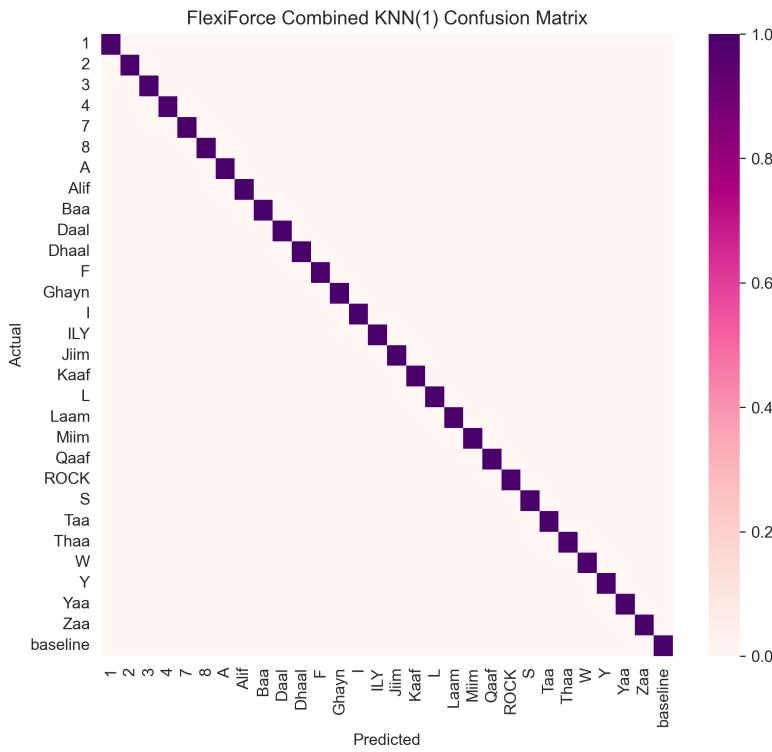


Figure 30: FlexiForce combined best-performing algorithm confusion matrix (kNN(1)).

### 5.2.5 Best Performing Learning Curves

Figure 31 shows the learning curves for each of the best-performing algorithms. For every case, they suggest a well-fitted model, with training and validation curves following a similar pattern. Validation accuracy takes longer to converge as expected, and there is little variance in training and validation curves after convergence. This suggests that the model has generalised to the underlying patterns in the data well, and did not fit the noise for all datasets.

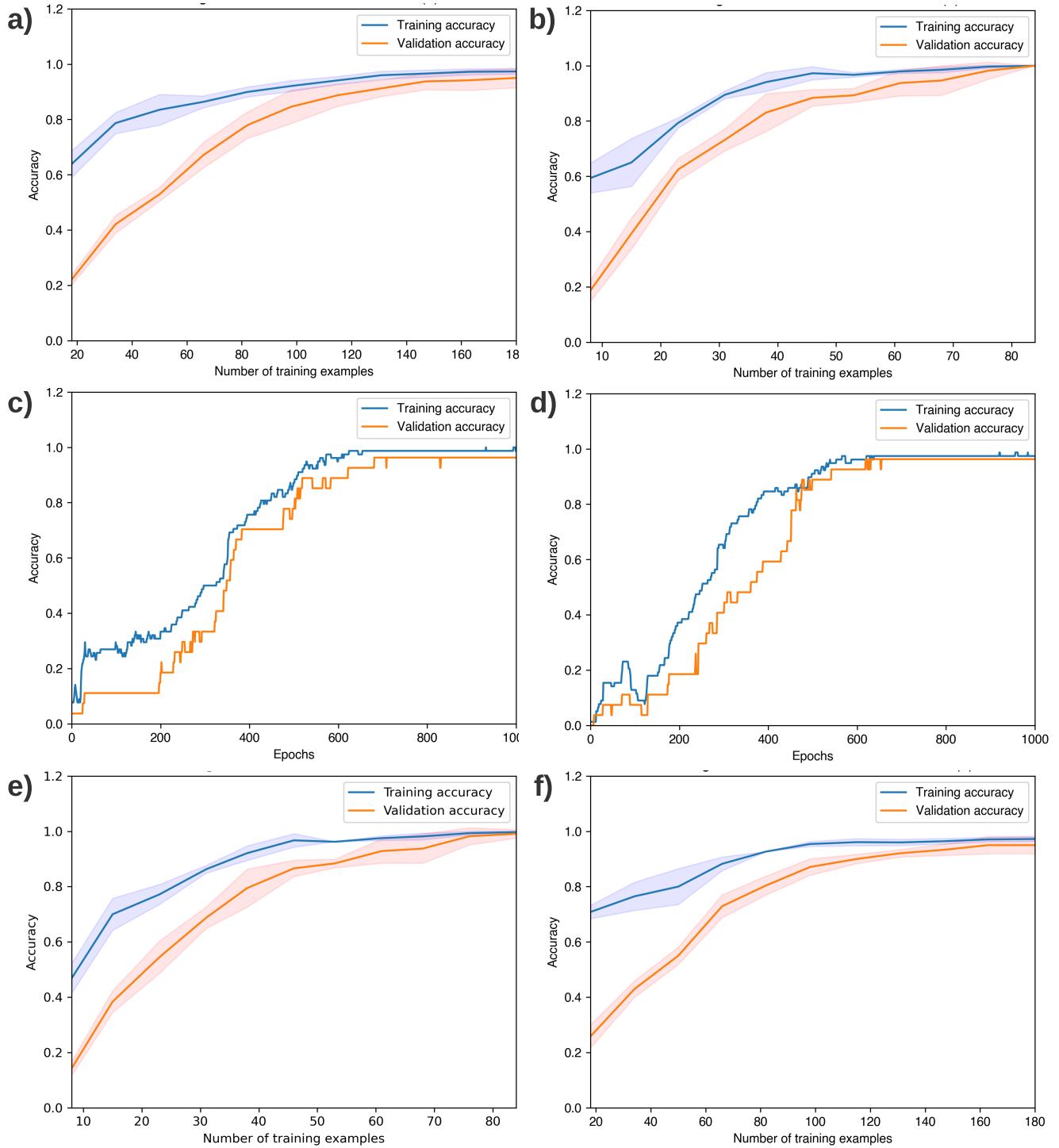


Figure 31: Learning curves for a) Foam AmSL (kNN(2)), b) FlexiForce AmSL (kNN(2)), c) Foam ArSL (ANN(2)), d) FlexiForce ArSL (ANN(2)), e) Foam Combined (kNN(1)), and f) FlexiForce combined (kNN(1)).

### 5.3 Analysis and Evaluation of Results

The results obtained can be further analysed and evaluated.

### 5.3.1 Observations from Results

The best-performing algorithms exhibit excellent performance, and the learning curves suggest they are well-fitted. These consist of only the kNN and ANN algorithms, which means they are well-suited for problems involving the type of dataset used.

The best data split is the 70/30 split (case 2) overall, with enough training points to allow good generalisation, and enough testing points to thoroughly test the trained model. A 60/40 split generally gave too little training data, and an 80/20 split too little testing data.

High performance was expected for single language classification as easily distinguishable signs formed simple datasets. However, the excellent performance of the combined set was not anticipated. When combining languages, many gestures chosen are the same, for example, 'Laam' and 'L'. As a result, misclassifications were expected and desired, and the lack of them in similar gestures suggests issues with the dataset.

XGBoost often exhibits worse performance, suggesting it is not an optimal algorithm in this scenario. Furthermore, RF and SVM often have excellent performance metrics, but their learning curves suggest potential overfitting. Figure 32 below shows overfitting of the training data initially with no variance, but validation accuracy still converges. In this case, it is likely that the models are too complex and capture noise in the training data, and although it generalises, it may not be optimal. For example, SVMs form a separating boundary between classes to classify the data, therefore very distinct data points are likely to overfit and perform highly on unvaried data, as is used here. This suggests dataset limitations, which are discussed later.

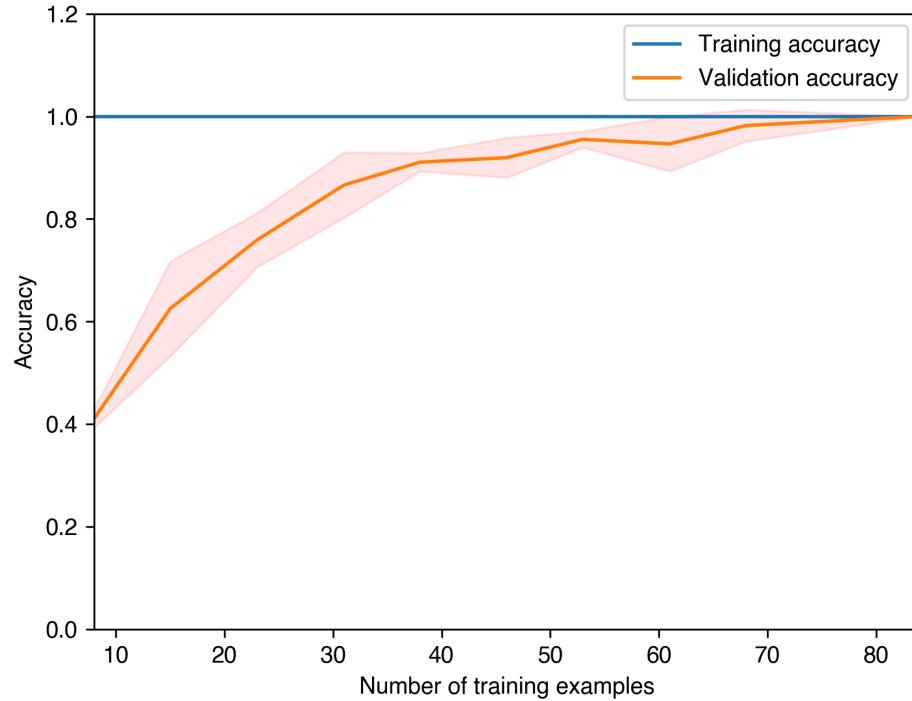


Figure 32: RF(1) learning curve for Foam AmSL dataset.

Despite some limitations, the best models appear to generalise well, suggesting they learn underlying patterns in the data effectively. Most misclassifications are of gestures that differ by a single finger signal, with very few that differ greatly, showing the models to be effective.

### 5.3.2 Evaluation of Data Limitations

There are many limitations with the datasets. Signals deviate due to extended sensor use, positioning of sensors, and how the user performs the gesture. All gestures were performed by a single person consistently, producing a

lack of variation and real-world implications in the datasets with different people performing gestures differently. Despite attempts to keep the glove taut, sensor positioning also deviates naturally. Sensor ranges changing with extended use largely impacts results due to lack of breaks and randomness in the data collection, so all gestures would follow similar signals at the time they were performed, but may not transfer to other scenarios. This is likely due to hysteresis of the sensors, where resistance temporarily changes due to mechanical stress effects on the microstructure. It would also be beneficial to sample with different fabricated gloves.

The ten repetitions used is low for a detailed analysis. The use of just joint 2 also limits the number of distinct gestures possible to classify. A full analysis requires more variation and randomness in the data collection process, along with more attempts to obtain a well-rounded algorithm for wider use. One method of improving the generalisation of the algorithm is to manually add identical and independently distributed Gaussian or Laplacian noise to the dataset, which may produce stronger models more capable of dealing with real-world inputs and variation. Furthermore, more gestures should be added to the datasets using more sensors.

## 5.4 Live Recognition

Live recognition was tested for the models. An example can be seen in 'LiveRecognitionRecording.mp4' in the GitHub repository, using the FlexiForce A201 sensors. It confirmed that the best models chosen were able to generalise the data to an extent, and had acceptable accuracy with new, unseen data. For the 5 gestures chosen: Baseline, I, Y, 'Rock on', and 'ILY', it was able to correctly classify them with only small errors as the signals settled. In general, the FlexiForce sensors performed better due to having a more consistent baseline and range of signals. As a result, more development of the foam sensor models, in terms of gaining more random data for models to cope with varied signals, would be beneficial.

## 6 Conclusion and Evaluation of Process and Outcomes

This details further critical discussion of the project and evaluates the process.

### 6.1 Summary

This study made significant contributions to novel sensor development, by fabricating and implementing novel sensors in bi-lingual sensor-based sign language classification. The experiments have demonstrated the feasibility of using novel piezoresistive sensors for gesture recognition purposes, achieving high classification accuracy with generalised, well-fitted sign language models. kNN and ANN algorithms performed best in general, with RF and SVM often appearing to overfit the data, and XGBoost appearing to be unsuitable. Due to a lack of variation in the data, the performance is average in real-time, despite excellent performance on the testing dataset. The lowest accuracy observed across all tests was 86%, and many achieved perfect performance. Despite the model and sensor limitations discussed, the study still gave promising results. The developed foam sensors exhibit great potential for strain applications, with high sensitivities and excellent SNRs achievable. The printed sensors' potential is high, but the performance and reproducibility required still must be questioned. Overall, the study shows capability of excellent performance and highlights the necessity for continued research in both experimental and computational aspects.

### 6.2 Strengths

- **Work process:** Work was completed in a logical order.
- **Thorough literature review:** All knowledge required for the project was obtained.
- **Depth of experimental work:** A detailed study with a high number of successful tests was completed.
- **Good engineering practice:** Detailed experimental logs were crucial to refer to.
- **Analysis of data:** Limitations with models and sensors were easily detected.
- **Depth of ML study:** The ML study tuned many optimisable parameters, which led to excellent performance.
- **Mitigation of data issues:** Data processing steps like normalisation mitigated issues with different sensor ranges and base values.
- **Novelty:** Most concepts were untested in this application, therefore can aid advancement in sensorised glove technology for HCI.

### 6.3 Weaknesses

- **Inefficiency:** Practical work is unpredictable, and issues are common in novel sensor development. Where the lab can be busy and experiments face issues, the project schedule can be delayed. Implementing a booking system would solve this, or allowing longer times for experiments on the day. There were many times processes like screen-printing or other fabrications would fail, delaying progression.
- **Glove type:** Nitrile medical gloves are cheap and excellent for testing, but not designed to be reused. This resulted in gloves tearing, despite mitigation attempts like using strengthened gloves. For the ML application pursued, a reusable alternative, like fabric, should have been considered.
- **Datasets:** Biased sampling methods without randomness were used, which led to limited data sets. Different people should collect data, with differently made gloves to ensure variation in fabrication and placement, and signs should be made in a random order to avoid extended sensor usage issues. It is important to include these variances for all gestures. More repetitions should also be used.

### 6.4 Limitations

- **Sensors used:** Different sensor types and positions would allow a wider range of gestures and signals to be detected.

- **Resources Available:** The project is mostly novel, so there is limited research for comparison. It can therefore be difficult to judge the results.
- **Computational Power:** The computational power available limited the size of the grid search, meaning not all hyperparameters could be tuned together, limiting performance.
- **Extended Use:** The large foam sensors, combined with an adhesive, add stiffness which makes it unfeasible for long-term use due to comfort. Miniaturisation and alternative adhering methods are required.

## 6.5 Future Work and Implications

The sensors each have many practical implications, but further research and development is first required, as detailed below.

- Wireless gloves using a Bluetooth low-energy transmitter and microcontroller attached to the wrist.
- More sensors to detect different gestures.
- Testing different inks, fabrication methods, and shapes for the screen-printed sensors. High-resolution 3D printing using stretchable conductive inks could lead to miniaturisation and reproducibility in an automated process.
- Testing different materials for foam sensors.
- More random, varied data collection.
- Biocompatibility tests for clinical applications.

The different sensors can be tuned for different potential applications by altering the sensitivity. This allows detection of smaller or larger movements at appropriate resolutions.

- Surgical scenarios, like interacting with surgical images by performing gestures.
- Rehabilitation or disease monitoring, like monitoring tremors produced by recovering patients.

A key area of focus regardless of the application is developing the HCI aspect to increase usability.

- Test methods of translating classified data into a user-friendly interface to obtain information from sensors.
- Perform feasibility studies to see how different methods are perceived by others.

Sign language recognition was used as a proof of concept in the report, but other applications mentioned are suitable for the sensors. The sign language route has large future potential, as it could provide a cheap, reproducible technology that can revolutionise methods for learning sign language, a necessity for people affected by hearing loss. Duolingo is a global tool for learning languages, so combining reproducible sensors, a reliable ML model of different sign languages, and a popular interface has excellent prospects.

## References

- [1] Oudah M, Al-Naji A, Chahl J. Hand Gesture Recognition Based on Computer Vision: A Review of Techniques. *Journal of Imaging*. 2020 Jul 23;6(8):73. Available from: doi: 10.3390/jimaging6080073
- [2] Sundaram S, Kellnhofer P, Li Y, Zhu JY, Torralba A, Matusik W. Learning the signatures of the human grasp using a scalable tactile glove. *Nature*. 2019 May;569(7758):698–702. Available from: doi: 10.1038/s41586-019-1234-z
- [3] Li WJ, Hsieh CY, Lin LF, Chu WC. Hand gesture recognition for post-stroke rehabilitation using leap motion. 2017 International Conference on Applied System Innovation (ICASI). 2017 May; Available from: doi: 10.1109/ICASI.2017.7988433
- [4] Sciforce. Introduction to the White-Box AI: the Concept of Interpretability. Medium. Sciforce; 2020. Available from: <https://medium.com/sciforce/introduction-to-the-white-box-ai-the-concept-of-interpretability-5a31e1058611>
- [5] Fadelli I. WaveGlove: A glove with five inertial sensors for hand gesture recognition. Techxplore.com: Tech Xplore; 2021. Available from: <https://techxplore.com/news/2021-06-waveglove-glove-inertial-sensors-gesture.html>
- [6] Wang L, Lou Z, Jiang K, Shen G. Bio-Multifunctional Smart Wearable Sensors for Medical Devices. *Advanced intelligent systems*. 2019 Sep 1;1(5):1900040–0. Available from: doi: 10.1002/aisy.201900040
- [7] Zou H, Zhang Y, Guo L, Wang P, He X, Dai G, et al. Quantifying the triboelectric series. *Nature Communications*. 2019 Mar 29;10(1). Available from: doi: 10.1038/s41467-019-09461-x
- [8] Puers R. Capacitive sensors: When and how to use them. *Sensors and Actuators A: Physical*. 1993 Jun;37-38:93–105. Available from: doi: 10.1016/0924-4247(93)80019-D
- [9] Fiorillo AS, Critello CD, Pullano SA. Theory, technology and applications of piezoresistive sensors: A review. *Sensors and Actuators A: Physical*. 2018 Oct;281:156–75. Available from: doi: 10.1016/j.sna.2018.07.006
- [10] Ramli NA, Nordin AN, Zainul Azlan N. Development of low cost screen-printed piezoresistive strain sensor for facial expressions recognition systems. *Microelectronic Engineering*. 2020 Oct;234:111440. Available from: doi: 10.1016/j.mee.2020.111440
- [11] Liu H, Zhang H, Han W, Lin H, Li R, Zhu J, et al. 3D Printed Flexible Strain Sensors: From Printing to Devices and Signals. *Advanced Materials*. 2021 Jan 14;33(8):2004782. Available from: doi: 10.1002/adma.202004782
- [12] Tian B, Yao W, Zeng P, Li X, Wang H, Liu L, et al. All-printed, low-cost, tunable sensing range strain sensors based on Ag nanodendrite conductive inks for wearable electronics. *Journal of Materials Chemistry C*. 2019;7(4):809–18. Available from: doi: 10.1039/C8TC04753G
- [13] Li Y, Liu Y, Bhuiyan SRA, Zhu Y, Yao S. Printed Strain Sensors for On-Skin Electronics. *Small Structures*. 2021 Nov 29;3(2):2100131. Available from: doi: 10.1002/sstr.202100131
- [14] Sharma C, Ojha CSP. Statistical Parameters of Hydrometeorological Variables: Standard Deviation, SNR, Skewness and Kurtosis. *Lecture Notes in Civil Engineering*. 2019 Jun 26; 29:59–70. Available from: doi: 10.1007/978-981-13-8181-2\_5
- [15] Wang L, Lou Z, Jiang K, Shen G. Bio-Multifunctional Smart Wearable Sensors for Medical Devices. *Advanced intelligent systems*. 2019 Sep 1;1(5):1900040–0. Available from: doi: 10.1002/aisy.201900040
- [16] Li J, Fang L, Sun B, Li X, Kang SH. Review—Recent Progress in Flexible and Stretchable Piezoresistive Sensors and Their Applications. *Journal of The Electrochemical Society*. 2020 Jan 2;167(3):037561. Available from: doi: 10.1149/1945-7111/ab6828
- [17] Bhatt BB, Kumar L, Kushwaha A, Gupta D. An ultra-compressible piezoresistive strain and pressure sensor based on RGO-CNT-Melamine foam composite for biomedical sensing. *Sensors and Actuators A: Physical*. 2021 Nov;331:112875. Available from: doi: 10.1016/j.sna.2021.112875
- [18] Su Y, Ma K, Mao X, Liu M, Zhang X. Highly Compressible and Sensitive Flexible Piezoresistive Pressure Sensor Based on MWCNTs/Ti3C2Tx MXene @ Melamine Foam for Human Gesture Monitoring and Recognition. *Nanomaterials*. 2022 Jun 29;12(13):2225. Available from: doi: 10.3390/nano12132225

- [19] Hugo Layard Horsfall, Salvadores Fernandez C, Bagchi B, Datta P, Gupta P, Chan Hee Koh, et al. A Sensorised Surgical Glove to Analyze Forces During Neurosurgery. *Neurosurgery*. 2022 Dec 12;92(3):639–46. Available from: doi: 10.1227/neu.0000000000002239
- [20] Hollinger A, Wanderley M. Evaluation of Commercial Force-Sensing Resistors. *Input Devices and Music Interaction Laboratory (IDMIL) Report*; January 2006. Available from: <https://www.semanticscholar.org/paper/Evaluation-of-Commercial-Force-Sensing-Resistors-Hollinger/d6a38535098c2a08479e67d068798474994c426c>
- [21] Surbhi Singh, BS. Force measurement, simulation and implementation using Flexiforce A201 sensor on a mobile robot. Masters Degree Thesis. The University of Texas at Dallas; 2015.
- [22] Morganti E, Angelini L, Adami A, Lalanne D, Lorenzelli L, Mugellini E. A Smart Watch with Embedded Sensors to Recognize Objects, Grasps and Forearm Gestures. *Procedia Engineering*. 2012;41:1169–75. Available from: doi: 10.1016/j.proeng.2012.07.297
- [23] Lim WN, Lee Y, Yap KM, Yen CC. Weight Perception Simulation in Virtual Reality with Passive Force using Force Sensing Resistors. 2022 IEEE International Conference on Computing (ICOCO). 2022 Nov 14; Available from: doi: 10.1109/ICOCO56118.2022.10031797
- [24] Tan Y, Yoon SH, Ramani K. BikeGesture. Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. ACM Conferences. 2017. Available from: doi: 10.1145/3027063.3053075
- [25] Wahid MF, Tafreshi R, Al-Sowaidi M, Langari R. An efficient approach to recognize hand gestures using machine-learning algorithms. 2018 IEEE 4th Middle East Conference on Biomedical Engineering (MECBME). 2018 Mar; Available from: doi: 10.1109/MECBME.2018.8402428
- [26] Siddiqui UA, Ullah F, Iqbal A, Khan A, Ullah R, Paracha S, et al. Wearable-Sensors-Based Platform for Gesture Recognition of Autism Spectrum Disorder Children Using Machine Learning Algorithms. *Sensors*. 2021 May 11;21(10):3319. Available from: doi: 10.3390/s21103319
- [27] Carnevale A. Machine learning algorithms for facial gesture recognition: a first analysis based on event-driven sEMG acquisition. Masters Degree Thesis. Politecnico di Torino; 2021.
- [28] Liu C, Al Alif M, He G. Shoulder Motion Detection Algorithm Based on MPU6050 Sensor and XGBoost Model. 2022 International Conference on Computing, Communication, Perception and Quantum Technology (CCPQT). 2022 Aug; Available from: doi: 10.1109/CCPQT56151.2022.00068
- [29] Al-Handarish Y, Omisore OM, Chen J, Cao X, Akinyemi TO, Yan Y, et al. A Hybrid Microstructure Piezoresistive Sensor with Machine Learning Approach for Gesture Recognition. *Applied Sciences*. 2021 Aug 6;11(16):7264. Available from: doi: 10.3390/app11167264
- [30] Lee BG, Lee SM. Smart Wearable Hand Device for Sign Language Interpretation System With Sensors Fusion. *IEEE Sensors Journal*. 2018 Feb 1;18(3):1224–32. Available from: doi: 10.1109/JSEN.2017.2779466
- [31] Mummadri C, Leo F, Verma K, Kasireddy S, Scholl P, Kempfle J, et al. Real-Time and Embedded Detection of Hand Gestures with an IMU-Based Glove. *Informatics*. 2018 Jun 11;5(2):28. Available from: doi: 10.3390/informatics5020028
- [32] Oz C, Leu MC. Linguistic properties based on American Sign Language isolated word recognition with artificial neural networks using a sensory glove and motion tracker. *Neurocomputing*. 2007 Oct;70(16-18):2891–901. Available from: doi: 10.1016/j.neucom.2006.04.016
- [33] Ullah F, AbuAli NA, Ullah A, Ullah R, Siddiqui UA, Siddiqui AA. Fusion-Based Body-Worn IoT Sensor Platform for Gesture Recognition of Autism Spectrum Disorder Children. *Sensors*. 2023 Feb 3;23(3):1672. Available from: doi: 10.3390/s23031672
- [34] Alosail D, Aldolah H, Alabdulwahab L, Bashar A, Khan M. Smart Glove for Bi-lingual Sign Language Recognition using Machine Learning. 2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT). 2023 Jan 5; Available from: doi: 10.1109/IDCIoT56793.2023.10053470
- [35] P. Trigueiros, F. Ribeiro, L. P. Reis. A comparison of machine learning algorithms applied to hand gesture recognition. 7th Iberian Conference on Information Systems and Technologies (CISTI 2012); 2012 June 20-23; Madrid, Spain. Madrid: IEEE; 2012.

- [36] Almasre MA, Al-Nuaim H. A comparison of Arabic sign language dynamic gesture recognition models. *Heliyon*. 2020 Mar;6(3):e03554. Available from: doi: 10.1016/j.heliyon.2020.e03554
- [37] Bhushan S, Alshehri M, Keshta I, Chakraverti AK, Rajpurohit J, Abugabah A. An Experimental Analysis of Various Machine Learning Algorithms for Hand Gesture Recognition. *Electronics*. 2022 Mar 21;11(6):968. Available from: doi: 10.3390/electronics11060968
- [38] Ding X, Jiang T, Xue W, Li Z, Zhong Y. A New Method of Human Gesture Recognition Using Wi-Fi Signals Based on XGBoost. 2020 IEEE/CIC International Conference on Communications in China (ICCC Workshops). 2020 Aug 9; Available from: doi: 10.1109/ICCCWorkshops49972.2020.9209953
- [39] Zhao Y, Sark V, Krstic M, Grass E. Novel Approach for Gesture Recognition Using mmWave FMCW RADAR. 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring). 2022 Jun; Available from: doi: 10.1109/VTC2022-Spring54318.2022.9860976
- [40] Srivastava D, Bhambhu L. Data classification using support vector machine. *JATIT*. 2010; 12(1). Available from: <http://www.jatit.org/volumes/Vol12No1/1Vol12No1.pdf>
- [41] Ul Hassan CA, Khan MS, Shah MA. Comparison of Machine Learning Algorithms in Data classification. 2018 24th International Conference on Automation and Computing (ICAC). 2018 Sep; Available from: doi: 10.23919/IConAC.2018.8748995
- [42] Osisanwo F.Y., Akinsola J.E.T., Awodele O., Hinmikaiye J. O., Olakanmi O., Akinjobi J. Supervised Machine Learning Algorithms: Classification and Comparison. *IJCTT journal*. 2017. Available from: doi: 10.14445/22312803/IJCTT-V48P126
- [43] Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making*. 2019 Dec;19(1). Available from: doi: 10.1186/s12911-019-1004-8
- [44] Vakili M, Ghamsari M, Rezaei M. Performance Analysis and Comparison of Machine and Deep Learning Algorithms for IoT Data Classification. *arXivorg*. 2020; Available from: doi: 10.48550/arXiv.2001.09636
- [45] Unknown. Deaf Sign Language's. *Blogspot.com*. 2014. Available from: <http://sadlierstore.blogspot.com/2014/05/deaf-sign-languages.html>
- [46] Cliparts.co. Sign Language Clip Art. 2023. Available from: <http://cliparts.co/sign-language-clip-art>
- [47] Freesvg.org. Vector image of rock on hand sign in black and white. 2021. Available from: <https://freesvg.org/vector-image-of-rock-on-hand-sign-in-black-and-white>
- [48] Bejarano M. Moroccan Arabic: Phrases and Vocabulary. *Feel Morocco*. 2015. Available from: <https://www.feelmorocco.travel/morocco/moroccan-arabic-phrases-vocabulary/>
- [49] Momosign Language Learning. Arabic sample. 2023. Available from: <https://www.momosignlanguagelearning.com/arabic-sample.html>
- [50] Lee S, Franklin S, Faezeh Arab Hassani, Takao Someya, Osman Goni Nayeem, Wang Y, et al. Nanomesh pressure sensor for monitoring finger manipulation without sensory interference. *Science*. 2020 Nov 20;370(6519):966–70. Available from: doi: 10.1126/science.abc9735

## Appendix A

Supporting material for the project can be found in the following GitHub repository.

<https://github.com/Toby080402/SensorisedGloves.git>

First, all scripts can be found, which consists of Python scripts for collecting data using Arduino Uno, Matplotlib plots based off of the ML algorithms, and a script using the trained models for live gesture recognition. Many other plots were made and recorded throughout the script to visualise the data that was recorded, and these are included in the 'archive' section of the script folders.

Jupyter Notebooks were used for ML to allow easy visualisation and outputs after each process. All 18 notebooks are found with the scripts. This includes three for each dataset (FoamAmSL, FoamArSL, FoamCombined, FlexiForceAmSL, FlexiForceArSL, FlexiForceCombined), as each is tested using three different cases for training and testing data splits.

Next, the Matlab code used to plot graphs from data collected within .xlsx and .csv files is included. This allowed all plots to be observed in a single location.

The datasets collected can then be observed. This includes the csv files used for ML model training, consisting of AmSL, ArSL, and combined language datasets for both foam and FlexiForce sensors.

## Appendix B

This details the full classification reports, referenced in Section 5.

Table 11: Full classification report of foam sensors on singular datasets.

Algorithm	Foam Sensor AmSL				Foam Sensor ArSL			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
SVM (1)	1.00	1.00	1.00	1.00	0.97	0.98	0.97	0.96
SVM (2)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
SVM (3)	1.00	1.00	1.00	1.00	0.98	0.99	0.98	0.98
kNN (1)	1.00	1.00	1.00	1.00	0.97	0.98	0.97	0.96
kNN (2)	1.00	1.00	1.00	1.00	0.98	0.98	0.98	0.98
kNN (3)	1.00	1.00	1.00	1.00	0.98	0.99	0.98	0.98
RF (1)	1.00	1.00	1.00	1.00	0.97	0.98	0.97	0.96
RF (2)	1.00	1.00	1.00	1.00	0.98	0.98	0.98	0.98
RF (3)	1.00	1.00	1.00	1.00	0.98	0.99	0.98	0.98
ANN (1)	1.00	1.00	1.00	1.00	0.97	0.98	0.97	0.96
ANN (2)	1.00	1.00	1.00	1.00	0.98	0.98	0.98	0.98
ANN (3)	1.00	1.00	1.00	1.00	0.97	0.97	0.97	0.97
XGBoost (1)	1.00	1.00	1.00	1.00	0.93	0.96	0.93	0.93
XGBoost (2)	1.00	1.00	1.00	1.00	0.84	0.84	0.84	0.82
XGBoost (3)	0.95	0.97	0.95	0.95	0.88	0.92	0.88	0.87

Table 12: Full classification report of foam sensors on combined dataset.

Foam Sensor Combined				
Algorithm	Accuracy	Precision	Recall	F1-Score
SVM (1)	1.00	1.00	1.00	1.00
SVM (2)	0.99	0.99	0.99	0.99
SVM (3)	0.98	0.99	0.98	0.98
kNN (1)	0.98	0.99	0.98	0.98
kNN (2)	0.94	0.93	0.94	0.93
kNN (3)	0.92	0.95	0.92	0.91
RF (1)	1.00	1.00	1.00	1.00
RF (2)	0.98	0.98	0.98	0.98
RF (3)	0.98	0.98	0.97	0.98
ANN (1)	0.93	0.94	0.93	0.93
ANN (2)	0.96	0.96	0.96	0.95
ANN (3)	0.90	0.92	0.90	0.90
XGBoost (1)	0.95	0.97	0.95	0.95
XGBoost (2)	0.86	0.88	0.86	0.86
XGBoost (3)	0.87	0.91	0.87	0.87

Table 13: Full classification report of FlexiForce A201 on singular datasets.

Algorithm	FlexiForce Sensor AmSL				FlexiForce Sensor ArSL			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
SVM (1)	1.00	1.00	1.00	1.00	0.97	0.98	0.97	0.96
SVM (2)	1.00	1.00	1.00	1.00	0.98	0.98	0.98	0.98
SVM (3)	1.00	1.00	1.00	1.00	0.98	0.99	0.98	0.98
kNN (1)	1.00	1.00	1.00	1.00	0.90	0.88	0.90	0.88
kNN (2)	1.00	1.00	1.00	1.00	0.93	0.96	0.93	0.93
kNN (3)	1.00	1.00	1.00	1.00	0.95	0.96	0.95	0.95
RF (1)	1.00	1.00	1.00	1.00	0.97	0.98	0.97	0.96
RF (2)	1.00	1.00	1.00	1.00	0.98	0.98	0.98	0.98
RF (3)	1.00	1.00	1.00	1.00	0.98	0.99	0.98	0.98
ANN (1)	1.00	1.00	1.00	1.00	0.90	0.88	0.90	0.88
ANN (2)	1.00	1.00	1.00	1.00	0.98	0.98	0.98	0.98
ANN (3)	0.98	0.99	0.98	0.98	0.97	0.97	0.97	0.97
XGBoost (1)	0.91	0.93	0.91	0.90	0.90	0.94	0.90	0.90
XGBoost (2)	0.90	0.91	0.90	0.89	0.93	0.96	0.93	0.93
XGBoost (3)	0.92	0.94	0.92	0.92	0.92	0.93	0.92	0.92

Table 14: Full classification report of FlexiForce A201 on combined dataset.

FlexiForce Sensor Combined				
Algorithm	Accuracy	Precision	Recall	F1-Score
SVM (1)	1.00	1.00	1.00	1.00
SVM (2)	0.97	0.97	0.97	0.97
SVM (3)	0.97	0.98	0.97	0.97
kNN (1)	1.00	1.00	1.00	1.00
kNN (2)	0.96	0.97	0.96	0.96
kNN (3)	0.98	0.98	0.97	0.97
RF (1)	0.98	0.99	0.98	0.98
RF (2)	0.96	0.96	0.96	0.95
RF (3)	0.94	0.95	0.94	0.94
ANN (1)	0.97	0.98	0.97	0.96
ANN (2)	0.96	0.97	0.96	0.95
ANN (3)	0.93	0.93	0.93	0.93
XGBoost (1)	0.88	0.88	0.88	0.87
XGBoost (2)	0.86	0.89	0.86	0.85
XGBoost (3)	0.86	0.89	0.86	0.85