AN ONGOING INDUSTRY/ RESEARCH INTERNSHIP REPORT ON

Smart Shoe for Predicting Ground Reaction Forces using Machine Learning

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

BATCHELOR OF TECHNOLOGY IN

COMPUTER ENGINEERING

BY

Dhruv Shambhu Thakkar (GR. No. 12111512)

UNDER THE MENTORSHIP OF

Prof. Pushkar Joglekar

DEPARTMENT OF COMPUTER ENGINEERING

BANSILAL RAMNATH AGARWAL CHARITABLE TRUST'S

VISHWAKARMA INSTITUTE OF TECHNOLOGY,

PUNE – 411037.

(An Autonomous Institute affiliated to Savitribai Phule Pune University)

A.Y. 2024 – 2025

DEPARTMENT OF COMPUTER ENGINEERING

BANSILAL RAMNATH AGARWAL CHARITABLE TRUST'S

VISHWAKARMA INSTITUTE OF TECHNOLOGY,

PUNE – 411037.

(An Autonomous Institute affiliated to Savitribai Phule Pune University)

CERTIFICATE

This is to certify that the industry/research internship report entitled **"Smart Shoe for Predicting Ground Reaction Force using Machine Learning"** submitted by **Dhruv Shambhu Thakkar** (**GR No: 12111512)** is approved for partial fulfillment of the requirements for the award of degree of Batchelor of Technology in Computer Engineering of Vishwakarma Institute of Technology, Savitribai Phule Pune University. This report is a record of bonafide work carried out as a part of his ongoing internship in HCR Lab, IIT Gandhinagar during the academic year 2024–2025, Semester–7.

Signature	Signature
Research Mentor	**Faculty Mentor**
Prof. Vineet Vashishta	Prof. Pushkar Joglekar
(HCR Lab, IIT Gandhinagar**)**	

Signature	Signature
Examiner	**Head of the Department**
	Prof. Sandeep Shinde

Date: November 2024

Place: Pune

Re: Inquiry Regarding Robotics Research Internship at HCR Lab IIT Gandhinagar

1 message

Vineet Vashista <vineet.vashista@iitgn.ac.in> Tue, 11 Jun 2024 at 10:03 am
To: DHRUV SHAMBHU <shambhu.dhruv211@vit.edu>

Dear Dhruv,

This is to confirm your internship at the HCR Lab, Mechanical Engineering at IIT Gandhinagar from August 5, 2024, to May 5, 2025. The internship does not offer any stipend. The internship will be conducted in person. Please note that hostel accommodation request on payment basis has to be made to IIT Gandhinagar. I will send your accommodation request to hostel office around June 15 - please remind me of the same.
You should book your travel once the hostel office confirm the accommodation.

Best,
Vineet

Vineet Vashista
Associate Professor
Indian Institute of Technology Gandhinagar

Website: Human-Centered Robotics Lab
Google Scholar Profile

ACKNOWLEDGEMENT

I wish to express my sincere gratitude to **Prof. Rajesh Jalnekar**, Director, VIT Pune, for providing the necessary facilities of the Institute and for his unwavering encouragement during the course of this research internship.
I am deeply indebted to **Prof. Sandeep Shinde**, Head of the Department of Computer Engineering, VIT Pune, for his invaluable guidance and continuous support throughout this journey.

My heartfelt thanks go to my research supervisor, **Prof. Vineet Vashishta**, for his enthusiastic mentorship, insightful discussions, and invaluable feedback, which greatly contributed to the progress and completion of this research work. His patience, guidance, and constructive suggestions have been instrumental in shaping this research internship.

Finally, I extend my sincere gratitude to all those who directly or indirectly supported me throughout this research internship.

Dhruv Thakkar

Date:

Table of Contents

1. Abstract

This project presents the design and implementation of a portable system to estimate **Ground Reaction Forces (GRFs)** using an innovative insole embedded with four pressure sensors configured as spiral silicon tube-type pressure pads. The aim is to provide an alternative to traditional force plate lab setups, which are cumbersome, expensive, and limited to controlled environments. This portable solution has significant potential for applications in sports, rehabilitation, and real-world biomechanical analysis.

Data was collected from multiple subjects performing a range of activities, including standing up, squats, and other motion-based tasks, to simulate diverse real-life scenarios. The pressure sensor data was processed to train various regression models such as Decision Trees, Random Forest, and Support Vector Regression (SVR). Deep learning models were also employed for comparison to evaluate performance across both traditional and advanced computational approaches.

The trained models were assessed on their ability to predict GRF components (Fx, Fy, and Fz) accurately. The evaluation highlighted trade-offs between computational efficiency and prediction accuracy, with certain regression models, such as Random Forest, achieving strong performance. To enhance practicality, the most effective model was deployed on a microcontroller, enabling real-time GRF estimation without reliance on external computing resources.

This study showcases the feasibility of using lightweight, sensor-embedded insoles for GRF estimation, providing a cost-effective and portable solution for biomechanical research and sports science. The system's versatility and ease of deployment pave the way for future advancements in wearable technology and human motion analysis.

Keywords: *Ground Reaction Forces (GRF), Pressure Sensors, Spiral Silicon Tube-Type Pressure Pads, Biomechanical Analysis, Rehabilitation, Regression Models, Decision Trees, Random Forest, Support Vector Regression (SVR), Deep Learning Models, Wearable Technology, Human Motion Analysis, Force plate*

2. Introduction

The global population is undergoing a significant and sustained aging process, leading to an increase in degenerative conditions of the musculoskeletal system (e.g., osteoporosis and arthritis) and nervous system (e.g., Alzheimer's disease, stroke, and Parkinson's disease). These conditions often result in walking difficulties, driving a growing demand for gait physical therapy.

Currently, gait rehabilitation involves therapists manually stimulating reflexes and guiding lower limb movements to help retrain the central nervous system for correct gait patterns. While effective, this approach is physically taxing for both patients and therapists, time-intensive, and costly. Moreover, clinical assessments of gait abnormalities often rely on visual observation, video analysis, physical tests, and patient self-reports, which can be subjective and lack precision. Objective metrics, such as ground reaction forces (GRF), joint rotations, and step length, are typically unavailable, limiting the accuracy of evaluations.

Additionally, patients with neurological impairments often struggle to understand their gait issues and rely solely on therapist instructions, with little insight into their progress. Providing intuitive visual feedback on walking behaviors and rehabilitation outcomes is crucial for improving patient awareness and engagement in their therapy.

2.1 Ground Reaction Force

Ground Reaction Forces (GRFs) are a cornerstone of biomechanical research, providing critical insights into human dynamics and kinematics. The measurement of GRFs has been pivotal in understanding activities such as walking, running, and lifting, with applications spanning fields like sports science, rehabilitation, and ergonomics. Accurate GRF measurement is vital for applications ranging from injury prevention and performance enhancement in sports to evaluating motor dysfunctions like Parkinson's disease or scoliosis. Additionally, GRFs play a central role in ergonomics, where they help assess physical stress during manual material handling tasks and in detecting risks of musculoskeletal disorders (MSDs). Despite their importance, traditional GRF measurement methods remain limited in scope and usability, especially in real-world scenarios.

The most widely used GRF measurement tool is the stationary **force plate**, developed by Herbert Elftman in the early 1930s. These plates have since evolved into highly accurate tools, such as the Kistler and AMTI platforms, widely used in laboratories. While these systems provide precise measurements, they suffer from significant drawbacks. They are expensive, bulky, and immobile, restricting their use to controlled environments. Consequently, real-time GRF measurement during diverse activities such as sports or outdoor walking is impractical. Moreover, force plates can only measure GRFs at specific locations, requiring multiple units to cover larger areas, further increasing costs and logistical challenges.

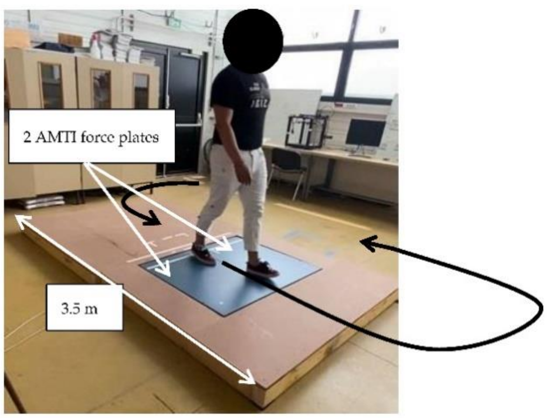


Fig 1: Force Plates in lab setup

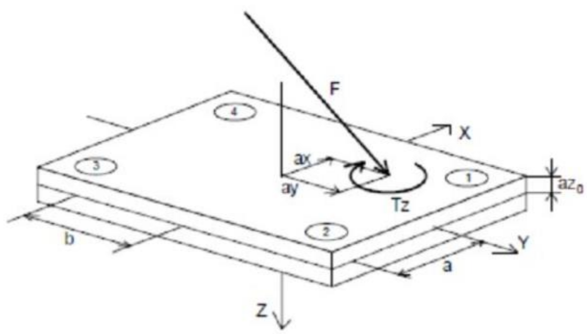


Fig 2: Force directions of Force Plate

2.2 Gait Cycle

The process of walking can be summarized as a sequential series of events starting with the registration and activation of the gait command within the central nervous system, followed by the transmission of signals to the peripheral nervous system. This leads to the activation of motor neurons, subsequent engagement of muscle fibers, and muscle contraction. The contraction generates forces across joints, which are regulated by skeletal segments, ultimately producing ground reaction forces (GRF). Normal gait is composed of two primary phases: the stance phase and the swing phase, which are further divided into a total of eight sub-phases. The gait cycle involves a combination of open- and closed-chain activities, reflecting the complex interplay of neurological, muscular, and skeletal systems.

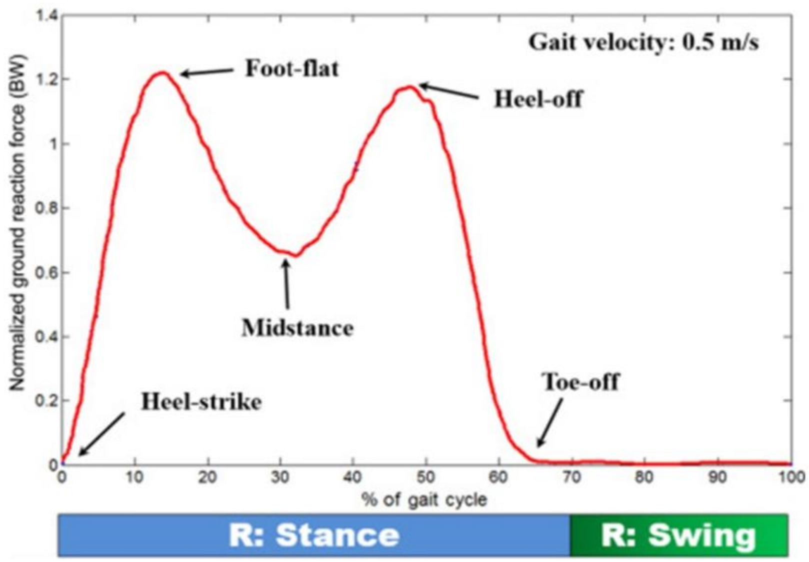


Fig 3: GRF of a normal Gait Cycle

2.2.1 Stance Phase

The stance phase occupies 60% of the total gait cycle, during which some part of the foot is in contact with the ground. It is further divided into five sub-phases:

1. Initial contact (heel strike)
2. Loading response (foot flat)
3. Mid-stance
4. Terminal stance (heel off)
5. Pre-swing (toe off)

2.2.2 Swing Phase

The swing phase occupies 40% of the total gait cycle , during which the foot is not in contact with the ground and the bodyweight is borne by the other leg and foot. It is further divided into three sub-phases:

1. Initial swing

2. Mid-swing
3. Late swing

In a complete two-step cycle both feet are in contact with the ground at the same time for 20% of the total gait cycle, 10% at the beginning of the stance phase and 10% at the end of the stance phase. These are termed 'double-support periods'. The rest of the time is spent in single support, when only one foot is in contact with the ground.

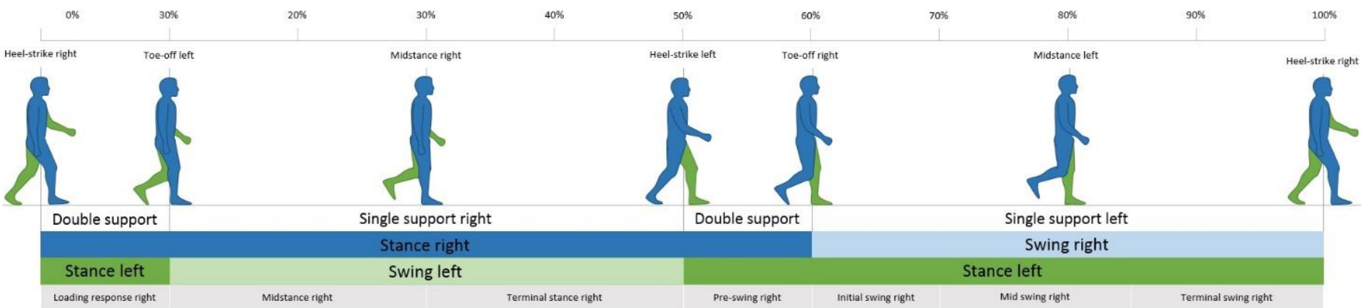


Fig 4: A Gait Cycle

2.3 Force Plates

Force plates are specialized mechanical sensing systems designed to measure ground reaction forces and moments generated during human movement. These devices use load cells, which may incorporate piezoelectric elements, strain gauges, or beam load cells, to detect forces. When force is applied to the plate, the sensors deform, producing voltage changes that correspond to the magnitude of the applied force. By positioning the sensors in various orientations, the system can capture both the direction and magnitude of forces in three dimensions. Additionally, force plates can provide valuable information, such as the center of pressure, center of force, and the moments acting around each axis.

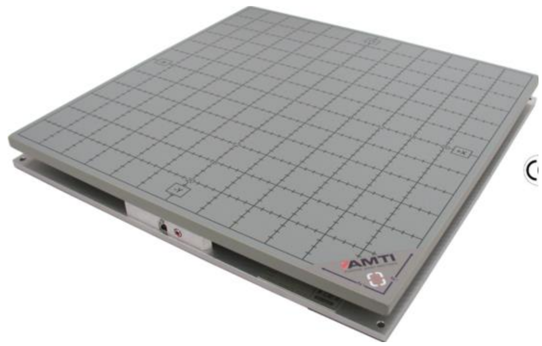


Fig 5: AMTI 6 Axis Force Plate

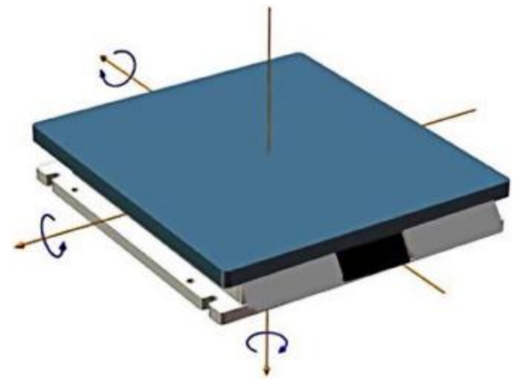


Fig 6: Six Axis of a Force Plate

Applications of force plates are vast, offering a wide range of possibilities.

Jump Tests: These include assessments such as countermovement jumps (CMJ), Abalakov jumps, single-leg jumps, squat jumps, and drop jumps.

Functional Tests: Examples include evaluations like squat assessments, push-ups, and sit-to-stand-to-sit movements.

Balance Tests: These cover activities such as quiet standing, single-leg standing, and single-leg range-of-stability tasks.

Isometric Tests: This category includes tests like the isometric mid-thigh pull (IMTP), isometric squats, single-leg isometric assessments, and the athletic shoulder (ASH) test.

This structured approach helps streamline the application of force plates across various performance and clinical evaluations.

2.4 Vicon Motion Capture System

Motion capture systems typically consist of a specialized array of cameras strategically arranged to cover a designated area where the tracked objects are located. These cameras work together to accurately determine the position of moving objects equipped with reflective markers. Using advanced tracking software, motion capture systems can identify and monitor various targets, including 3D glasses, handheld devices, human skeletons, UAVs, robots, or virtually any object of interest. Some systems achieve remarkable precision, with accuracy as fine as 1/5th of a human hair, enabling the collection of highly detailed movement data.

Often referred to as "mo-cap," motion capture is widely used to enhance animated characters, giving them a heightened sense of realism and authenticity.

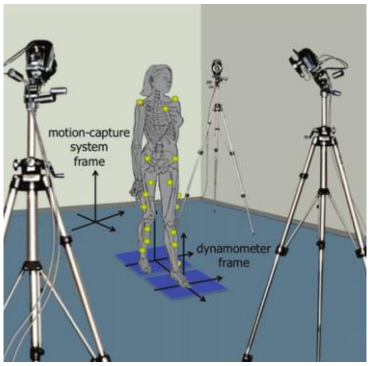


Fig 7: Vicon Motion Capture system

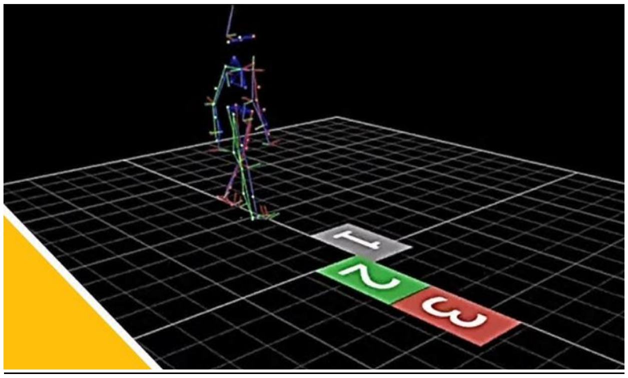
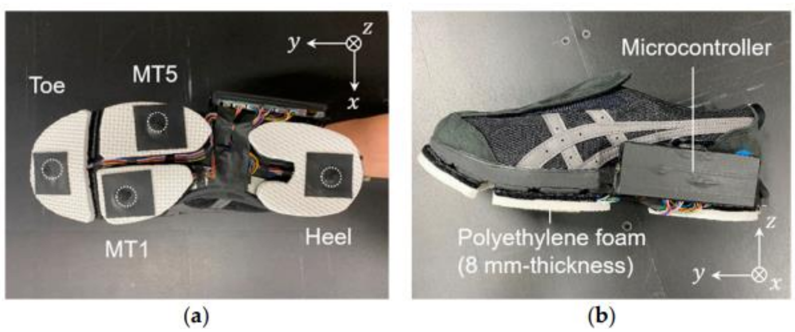


Fig 8: Markers and Force Plates in NEXUS Software

3. Literature Review

[1] This work develops a shoe sole sensor system with high-capacity, compact triaxial force sensors to estimate three-directional ground reaction forces (GRFs) during walking and turning movements. By utilizing Cr–N strain-sensitive thin films mounted at key points on the shoe sole, the system captures detailed GRF data, which is then processed using machine learning models, including multiple linear regression (MLR) and Gaussian process regression (GPR). The results demonstrate that the GPR model effectively predicts GRFs in the y and z directions with a prediction error of less than 15%, showing strong agreement with force plate measurements.

However, prediction accuracy in the x direction, especially during cross-step turning, was lower, highlighting the challenges of capturing complex movement patterns. Overall, this work provides a portable, accurate alternative to traditional force plates, with potential applications in gait analysis for sports, rehabilitation, and daily activities.



(a) Location of the four triaxial force sensors, and (b) side view of the sole sensor system.

[2] This paper presents a method for estimating three-axis ground reaction forces (GRFs) during walking using shoes equipped with three uniaxial EzForce-1D optical load cells and a sequence-to-sequence Long Short-Term Memory (seq2seq LSTM) model. The load cells, placed on the 1st metatarsal, 5th metatarsal, and

mid-heel regions, were designed to minimize distortion and adapt to different foot sizes. The collected data was used to predict vertical, anterior-posterior (AP), and medial-lateral (ML) GRFs. The experiment involved 81 healthy adults walking at their preferred speed, with GRF measured using force plates as a reference. The LSTM model was trained with resampled and filtered data from 109 sets and validated with separate test sets.

Fig 9: Load-cell based insole

The results showed strong accuracy, with correlation coefficients of 0.97, 0.96, and 0.90 for the vertical, AP, and ML GRFs, respectively. The root mean square errors were 65.12 N, 15.50 N, and 9.83 N, and the timing error was minimal at 0.06 s. A Bland-Altman analysis confirmed good agreement between measured and predicted values, especially for the maximum GRF. The study demonstrates that using an insole with load cells and a seq2seq LSTM model provides a viable, portable solution for accurate 3D GRF estimation, offering a practical alternative to traditional force plate systems for gait analysis in real-world environments.

[3] This paper presents a novel wearable six-axis sensor system designed for gait analysis, combining fiber-optic and FlexiForce sensors to measure ground reaction forces (GRF) and moments. The system is lightweight and compliant, which helps maintain natural gait. The sensor consists of two modules: a fiber-optic triaxial

force sensor and an array of FlexiForce pressure sensors that enhance the system's moment sensitivity. These sensors were mounted on shoes, with two units placed at the toe and heel, and tested in two configurations—floor-mounted and shoe-mounted. The shoe-mounted configuration enables the sensor to be used outside laboratory settings, providing practical solutions for real-world gait analysis and fall prevention. The sensor system was calibrated using both linear and nonlinear methods, with a neural network employed for nonlinear calibration, yielding improved accuracy.

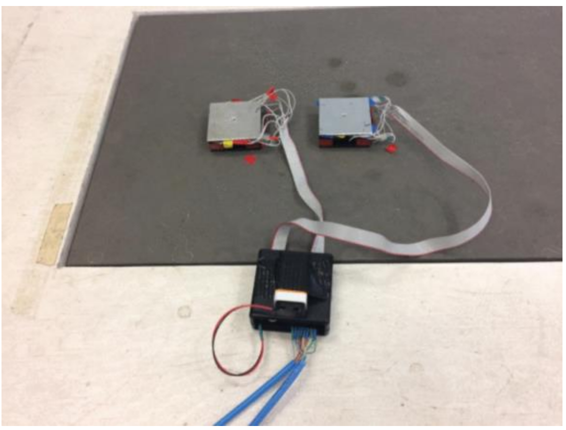


Fig 10:

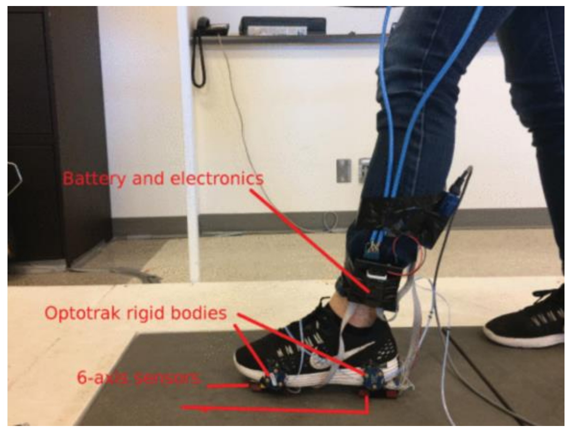


Fig 11:

The experimental results demonstrated that the sensor system performs well in both configurations, with linear calibration achieving an RMSE of 4.49% for floor-mounted tests and 9.39% for shoe-mounted tests. The neural network calibration further improved accuracy, reaching an RMSE of 2.68% for floor-mounted tests and 5.21% for shoe-mounted tests. Additionally, including sensor orientation in the calibration model resulted in a further reduction in RMSE to 3.25%. The study concludes that the proposed sensor system is capable of accurately measuring the full GRF and moments during gait, making it suitable for applications in gait analysis, center of pressure measurement, and fall prevention.

[4] This paper investigates the use of wearable in-shoe pressure sensors combined with machine learning models (Multiple Linear Regression, MLR, and Artificial Neural Networks, ANN) to predict the three components of Ground Reaction Force (GRF) in middle-distance running. The objective was to explore a portable

solution for GRF measurement, replacing traditional force plate systems. The study used hydrocell in-shoe pressure sensors placed on key foot locations to capture pressure data, which was then used as input for both MLR and ANN models. These models were trained and validated with data from elite middle-distance runners to predict vertical, anterior-posterior, and medio-lateral GRF components.

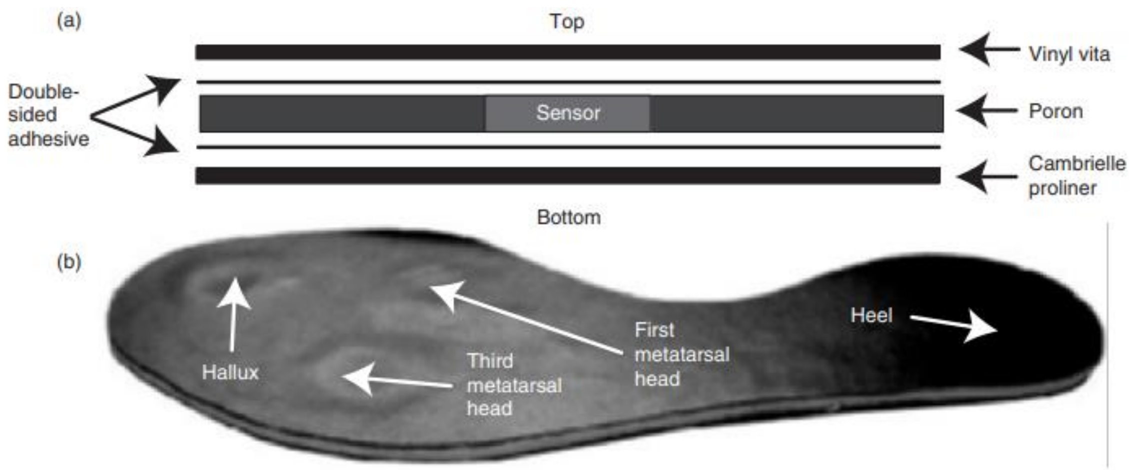


Fig 12:

The results showed that both models performed well, with MLR being more accurate for vertical and medio-lateral GRF components, and ANN excelling at predicting the anterior-posterior GRF component. The ANN model demonstrated flexibility in capturing non-linear relationships, while the MLR model was more straightforward and efficient. In general, the models were able to predict vertical and anterior-posterior GRF with high accuracy, though the medio-lateral component presented more challenges. These findings suggest that wearable sensors coupled with machine learning offer a viable method for real-time GRF measurement, with potential applications in sports performance monitoring, though further validation across different conditions and subject populations is necessary.

[5] This study presents a novel, low-cost wearable sensor system designed to estimate ground reaction forces (GRF) and ankle joint torque (AJT) during human motion tasks, such as walking and calf raises. The sensor system, composed of a pressure insole and tendon sensor, was tested on six healthy subjects and trained

using data from multiple tasks across different days. The results demonstrated that the multi-task, intra-day model produced normalized root mean square error (NRMSE) values under 10% for all components of the GRF, including vertical force, anterior-posterior force, and center of pressure (COP) location. The vertical torque predictions slightly exceeded this threshold, but the overall accuracy was still comparable to that of commercial devices. Notably, the insole system's high accuracy in predicting vertical force suggests it could potentially be used for accurate estimation of contact time and stance duration in dynamic activities.

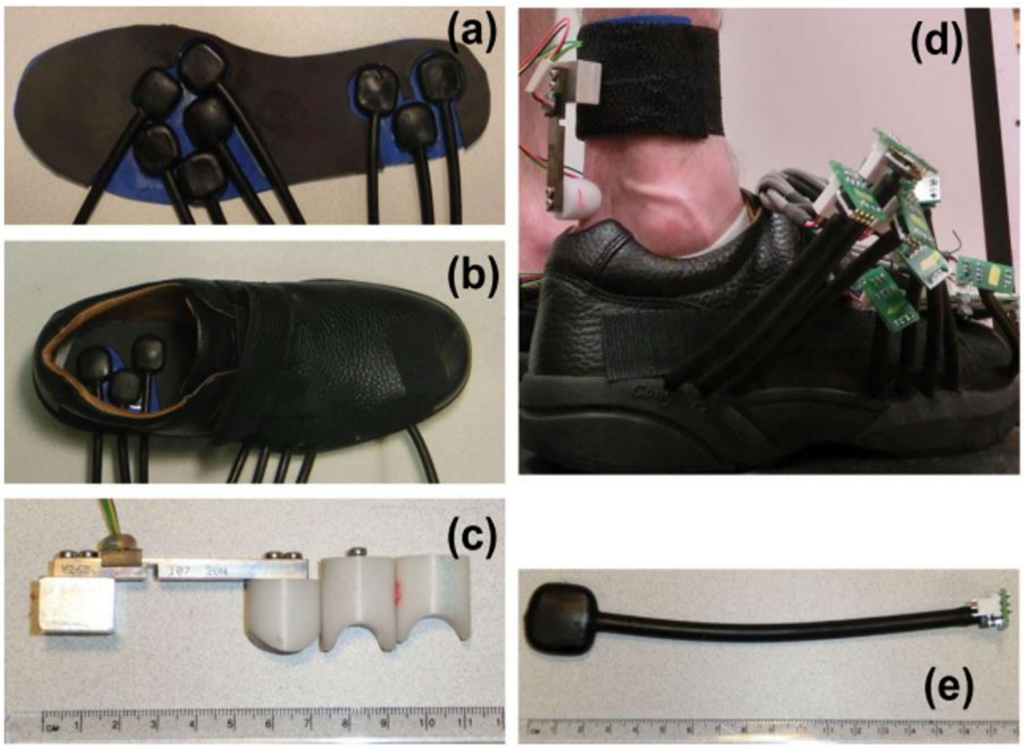


Fig 13:

The performance of the prototype sensors was favorably compared with previous studies using commercial systems like the Pedar insoles, which reported higher RMSE values for similar tasks. The custom sensors in this study offer a key advantage over commercial systems, as they respond to three-dimensional axial and shear stresses, rather than just uniaxial forces, allowing for a more accurate representation of plantar pressure. Despite the increased variability in the data due to multi-day and multi-task testing, the results indicated that the sensor's accuracy was limited more by the sensor's inherent precision than by task variability. This suggests that further refinement of the sensor system could improve its performance, expanding its potential applications in clinical settings and sports performance analysis. The study also highlighted that the system could be used to accurately predict spatial variables, such as peak force and COP, which are crucial for understanding gait dynamics.

[6] This paper provides a new analysis method of the GCF signals is discussed for detection of the gait phases. Human gaits are complicated, and the gait phases cannot be exactly distinguished by comparing sensor outputs to a threshold. This paper proposes a method by fuzzy logic for detecting the gait phases continuously and smoothly. The smooth and continuous detection of the gait phases enables a full use of information obtained from GCF sensors. For advanced rehabilitation systems, this paper also introduces a higher level algorithm that quantitatively monitors the amount of abnormalities in a human gait. The abnormalities detected by the proposed method include an improper GCF pattern as well as an incorrect sequence of the gait phases. To realize the monitoring algorithm, the gait phases are analyzed as a vector and the abnormalities are detected by simple kinematic equations. The proposed methods are implemented by using signals from sensor-embedded shoes called smart shoes. Each smart shoe has four GCF sensors installed between the cushion pad and the sole. The GCF sensor applies an air pressure sensor connected to an air bladder. A gait monitoring system that integrates the proposed methods is shown in this paper and verified for both normal and abnormal gaits.

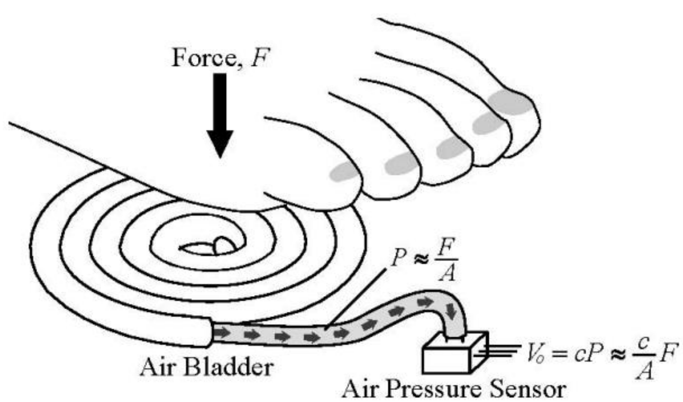


Fig 14:

4. Problem Statement

The problem addressed in this study is the development of a portable, cost effective shoe insole system equipped with pressure sensors to predict ground reaction forces (GRF) using machine learning. Current methods for measuring GRF, such as force plates, are expensive and require specialized laboratory setups, limiting their accessibility for real-time or field-based applications, particularly in sports or rehabilitation settings.

This project aims to create an alternative solution by using a set of pressure sensors embedded in a custom shoe insole to measure plantar pressure changes during various dynamic activities. The data collected from these sensors is then used to train machine learning models, specifically regression-based and deep learning approaches, to predict GRF components such as vertical force, anterior-posterior force, and center of pressure with high accuracy. By leveraging machine learning techniques, the shoe insole system seeks to provide an affordable, portable, and real-time solution for accurate ground reaction force estimation, enabling broader accessibility and applications in biomechanics, sports science, and clinical rehabilitation.

5. Scope and Objectives

This project aims to design and implement a portable, wireless insole system that leverages pressure sensors to measure localized plantar pressure and predicts the ground reaction forces (GRF) using machine learning techniques. The system will offer an affordable alternative to traditional force plate setups commonly used in biomechanics research and clinical assessments. The system is designed to be lightweight, non-intrusive, and suitable for continuous monitoring in real-world conditions, making it applicable in various contexts such as sports, rehabilitation, and assistive technologies. The scope includes data collection from subjects performing dynamic activities like walking and squats, training machine learning models on this data, and evaluating the system's performance in predicting GRF with high accuracy.

5.1 Objectives:

5.1.1. Development of Pressure Sensing Insole:

To design and prototype a wireless insole system equipped with pressure sensors to measure plantar pressure during dynamic activities such as walking and squats.

5.1.2. Data Collection and Preprocessing:

To collect and preprocess data from pressure sensors and use it to predict ground reaction forces (GRF) during activities. The data will be processed to ensure synchronization with motion capture and force plate data.

5.1.3. Application of Machine Learning Models:

To apply machine learning techniques (e.g., neural networks, regression models) to predict the GRF based on pressure sensor data and evaluate the performance of these models through metrics such as root mean square error (RMSE) and normalized RMSE.

5.1.4. Evaluation and Comparison:

To evaluate the model's accuracy in comparison to conventional GRF measurement methods (such as force plates) and analyze the effectiveness of the wireless insole system in predicting GRF across various activities and subjects.

5.1.5. Exploration for Rehabilitative and Assistive Use:

To explore the potential applications of this system in rehabilitative and assistive settings, where real-time feedback on GRF could be used for improving performance, guiding rehabilitation, or preventing injury.

5.1.6. Cost-Effective and Portable Solution:

To design a cost-effective system that is portable and provides real-time data, making it accessible for widespread use outside of laboratory environments, particularly in sports and rehabilitation.

6. Methodology

6.1 Hardware Development

6.1.1 Hardware Components Used:

- Teensy 4.1
- I2C Multiplexer
- Sparkfun MicroPressure sensor
- Sparkfun ICM 20948 9 – dof IMU
- Li-ion cells
- 5V step down module
- Silicon tubes
- SD Card

6.1.2 Circuit Diagram

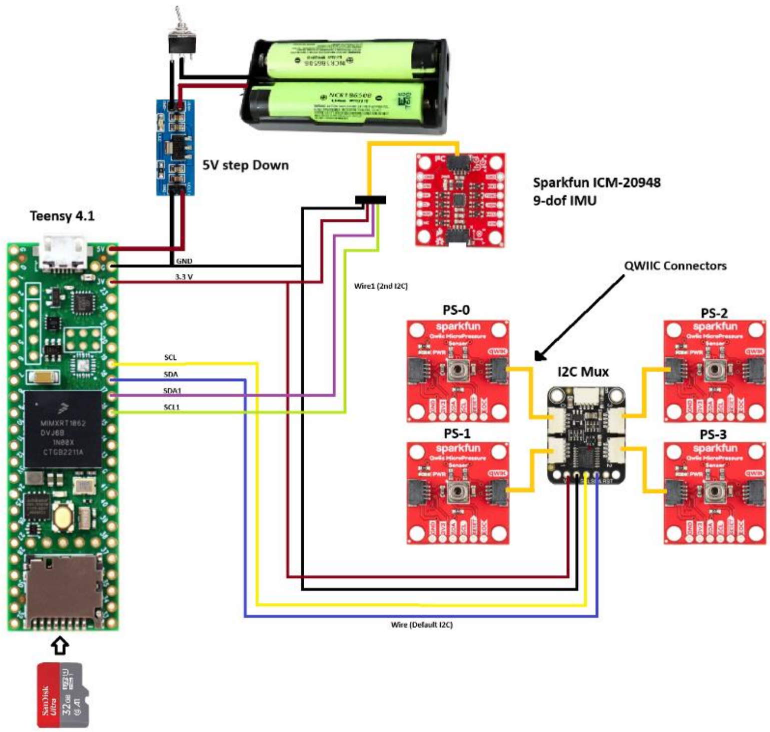


Fig 15: Final Circuit

6.1.3 CAD Models

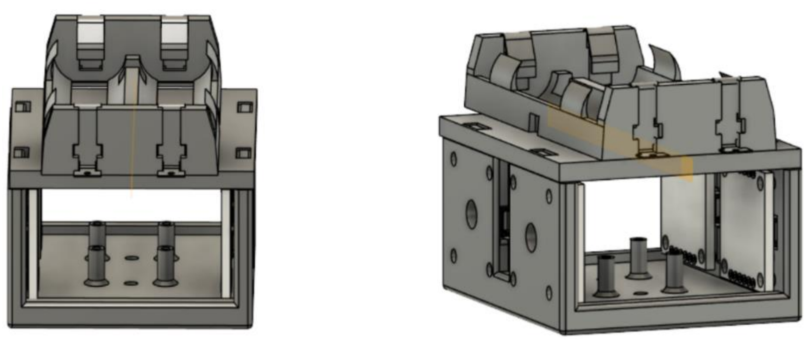


Fig 16: CAD

The design was made on Fusion 360. Dimensions were manually checked using the hardware present.

The battery holder was attached with a plate on the model which was removable and was connected to the main body using locking slots.4 slots were provided for the pressure sensor and a mount for I2C Mux. A slot for SD card was also provided such that it is easy to insert and remove after each experiment without disturbing the assembly.

It was made sure that the complete circuit fits properly inside the proposed CAD model without any extra wires hanging outside. It was then 3D printed using PLA Material with an infill percentage of 60%.

6.1.4 The Insole Fabrication

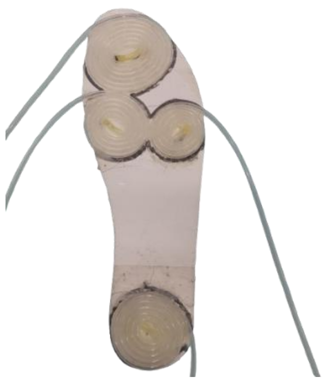


Fig 17: Silicon pressure pads

Fig 18: Feet contact representation

Shoe Insole with Silicone tube coiled pressure pads placed at: [i] Toe [ii] Metatarsophalengeal 1-2 [iii] Metatarsophalengeal 3-4 [iv] Heel

The insole design employs an innovative air pressure-based sensing unit to measure Ground Reaction Forces (GRF). The sensing unit comprises an air bladder, constructed by winding soft silicone tubes, and an air pressure sensor. When a foot applies force to the air bladder, it deforms, causing a pressure change proportional to the applied force, modeled by $P=F/A$, where P is the pressure, F is the applied force, and A is the effective area of contact. Silicone tubes were chosen for their minimal viscoelastic effects and faster response time compared to other materials like nylon. To optimize durability and sensitivity, tubes with an outer diameter of 4 mm and an inner diameter of 2 mm were selected, ensuring that pressure changes under normal body loads remain within the sensor’s measurable range (0–50 mbar). This design minimizes discomfort, provides linear pressure-to-force correlation, and ensures durability for practical use.

Four spiral patterns were created and adhered to the insole of a standard shoe using adhesive glue that did not deform the tubes. The insole was reinserted into the shoe such that the silicone tubes faced downwards while the flat surface remained in contact with the foot. This ensured that there was no difference in the feel of the shoe's insole and no discomfort to the wearer.

The other ends of the silicone tubes were routed out through a hole made in the shoe, positioned such that they did not extend below the sole and had no contact with the ground. All four tubes were secured with threads stitched to the shoe and connected to pressure sensors attached to the back of the shoe near the heel.

6.1.5 Final Assembly

Fig 19: Side View

Fig 20: Top View

6.2 Experimental Setup

To ensure a comprehensive and reliable dataset for ground reaction force (GRF) prediction, the experimental setup was meticulously designed and implemented. The primary equipment included AMTI force plates integrated with the Vicon NEXUS software for synchronized data recording. The Vicon system provided high-accuracy motion capture capabilities, while the AMTI force plates delivered precise GRF measurements. Synchronization between the shoe-mounted pressure sensors and the force plate data was achieved using a trigger signal, ensuring temporal alignment of all recorded data points.

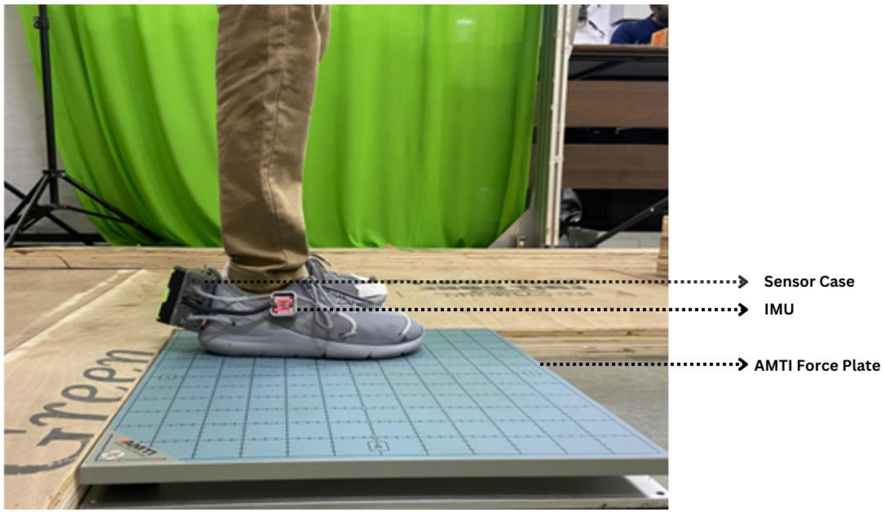


Fig 21: Shoe on Force Plate for data collection

6.2.1 Subject Recruitment and Consent

A total of four subjects participated in the study. Before the experiment, each participant was briefed on the motivation behind the research, its objectives, and the methodology. Written informed consent was obtained from all participants, adhering to ethical standards for human subject research.

6.2.2 Experimental Protocol

The experiment was divided into three main activities: step-ups, walking, and squatting. For the final squatting phase, the data collection protocol was as follows:

Squatting Task: Each subject performed 10 squats while wearing the instrumented shoe system. Subjects were instructed to place each foot on a separate force plate, ensuring that GRF data could be recorded independently for each leg. This setup facilitated the accurate measurement of forces and their distribution during the activity.

6.2.3 Data Synchronization:

The pressure sensor data from the shoe and the force plate readings were recorded simultaneously, ensuring that each timestamp contained paired data points. This alignment provided a robust dataset for training and evaluating machine learning models, with the force plate readings serving as the ground truth.

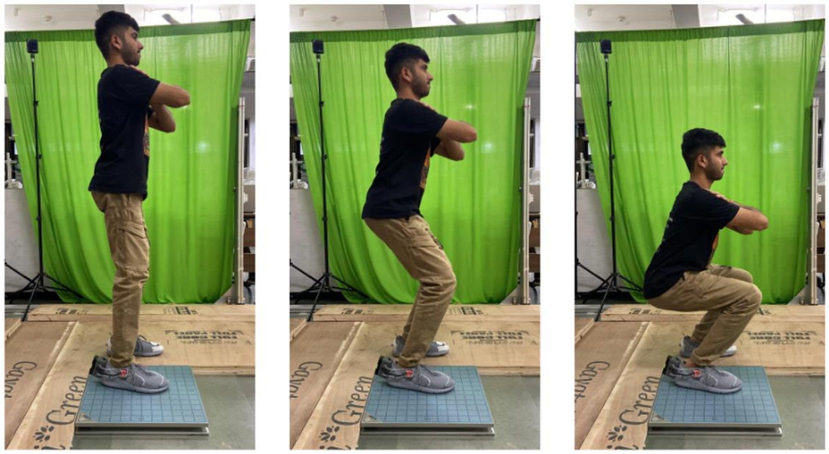


Fig 22: Squatting experiment on subject

6.2.4 Data Recording

During the experiments, the shoe insole sensors captured dynamic pressure variations, while the force plates measured corresponding GRFs. These synchronized recordings allowed for precise validation of the wearable system's performance in real-world scenarios. Data was collected at various frequencies to ensure compatibility with the machine learning algorithms used in the analysis.

This setup not only ensured the integrity of the collected data but also laid a strong foundation for evaluating the performance of regression and deep learning models in predicting GRF values. By involving different activities and multiple subjects, the experiment simulated diverse conditions, making the system adaptable to practical applications like sports and rehabilitation.

6.3 Data Collection

6.3.1 Data Synchronization

Data synchronization is a critical step in biomechanical studies, ensuring that data collected from multiple sensors or systems are temporally aligned for accurate analysis and interpretation. In wearable systems, precise synchronization is essential to correlate sensor readings with the reference signals or ground truth values. This alignment not only ensures the reliability of the model but also enables meaningful comparisons between datasets, especially in applications like gait analysis or sports performance monitoring.

In this study, data synchronization was achieved between the reference force plate system and the in-shoe pressure sensors. The force plate, a gold-standard tool for measuring ground reaction forces (GRFs), provides high-fidelity data that serve as the *ground truth*. However, discrepancies in data sampling rates or recording timestamps between the force plate and the shoe-mounted sensors can introduce errors in the analysis.

The force plate recorded data at 1000Hz whereas teensy with four pressures integrated recorded at a rate of around 30-40Hz.

To address this, both systems were synchronized using a common trigger mechanism or time alignment strategy, ensuring that the force plate data and shoe sensor data corresponded to the same instances of foot contact and movement. This step is vital as it allows the machine learning model to accurately map the in-shoe sensor data to the GRF measurements from the force plate, ultimately improving the prediction accuracy of the system. Proper synchronization also enhances the credibility and applicability of the developed system in real-world scenarios, where timing discrepancies can significantly impact the results.

At the start of the experiment, the trigger output of the Force Plate software was connected to the shoe using a DC jack. Whenever the Force Plate recording started, it sent a high pulse to the shoe, which served as a flag to initiate recording. This ensured that the recordings of both systems began simultaneously.

As soon as recording starts, we can remove the DC jack so that the subject can perform the experiment wirelessly.

6.3.2 Data Collection

For data collection, a Teensy 4.1 microcontroller was used, equipped with an SD card module to facilitate onboard data storage. The pressure sensor values from the insole's silicone tubes, along with the IMU (Inertial Measurement Unit) readings, were recorded in real time. These measurements were saved in a CSV file format, ensuring structured and accessible data storage. Each data entry was accompanied by a corresponding timestamp, enabling precise synchronization and temporal analysis of sensor outputs. This setup allowed efficient, portable data acquisition without reliance on external data transmission, making the system robust and independent.

currTime	s1	s2	s3	s4	s	AccX	AccY	AccZ	GyroX	GyroY	GyroZ
22	754.57	756.86	751.5	753.78	754.18	-60.55	-887.21	518.55	-0.02	-1.27	-2.05
53	754.56	756.85	751.51	753.8	754.18	-55.18	-881.84	499.51	1.69	-1.17	-2.65
84	754.55	756.82	751.48	753.81	754.16	-60.06	-897.46	506.84	0.11	-0.38	0.11
114	754.47	756.82	751.56	753.89	754.18	-54.69	-890.14	497.07	-1.05	2.91	0.19
145	754.44	756.87	751.51	753.88	754.17	-52.73	-892.58	500.98	-1.24	-0.01	0.17
176	754.43	756.88	751.49	753.8	754.15	-50.29	-902.34	512.7	-0.35	-2.4	-0.11
206	754.2	756.87	751.46	753.83	754.09	-45.41	-896.48	497.56	-1.16	0.05	-1.98
240	754.16	756.86	751.56	753.83	754.1	-46.88	-879.39	495.61	1.16	-1.76	-2.85
271	754.07	756.77	751.49	753.84	754.04	-64.94	-894.04	489.26	1.34	-0.06	-0.85
302	754.08	756.75	751.59	753.95	754.09	-56.15	-895.51	506.35	2.91	-5.22	0.57
333	754.09	756.77	751.56	754.07	754.12	-55.66	-884.77	516.11	4.02	-1.57	-0.31

Fig 23: Sample data from Shoe

Teensy created a new file for each experiment so that there is no overlapping and loss of data between multiple experiments.

For data collection, a Vicon motion capture system integrated with force plates was employed to accurately record ground reaction forces during various activities. The system operated at variable frequencies ranging from 30 Hz to 1000 Hz, depending on the specific requirements of the activity and the precision needed for the measurements. The high-frequency capabilities ensured that even rapid changes in force could be captured with fine temporal resolution. The recorded data from the force plates, corresponding to the three orthogonal components of force (Fz, Fx, and Fy), was exported using the ASCII pipeline feature of the Vicon system. This export process converted the data into a structured CSV file format, providing clear and accessible records of the force data synchronized with corresponding timestamps. The CSV files contained detailed information about the vertical force (Fz), anterior-posterior force (Fx), and medial-lateral force (Fy) at every recorded instant, ensuring a comprehensive dataset for subsequent analysis and model training. This meticulous data collection process played a critical role in creating a reliable ground truth for the validation of the insole-based GRF prediction system.

	Sub Frame	Fx	Fy	Fz	Mx	My	Mz	Cx	Cy	Cz
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0

Fig 24: Sample data frame of Force Plate

6.4. Data Preprocessing

The Data Preprocessing phase is a critical step in ensuring that raw data from the shoe sensors and the force plate is cleaned, synchronized, and transformed for analysis. This phase involves multiple steps including time alignment, visualization, removal of redundant or invalid data, interpolation, and augmentation. The detailed methodology for preprocessing is described below:

6.4.1 Loading and Structuring Data

The first step involves loading raw data from both the force plate and shoe sensors into MATLAB. Data from the force plate includes components like vertical force (Fz), while the shoe data contains pressure sensor readings (s1, s2, s3, s4) and their corresponding timestamps. After loading, both datasets are structured into tables, making it convenient for further processing and visualization.

6.4.2 Adding Time Information

To analyze data over time, a time column is added to the force plate dataset. Assuming a fixed sampling frequency (fs=200 Hz), the time column is calculated using:

$$t = (0, 1, 2, \ldots\ldots\ldots, n-1) \times 1/fs \times 1000$$

where nnn is the total number of samples, and the time values are converted to milliseconds. This ensures both datasets have a comparable temporal reference.

6.4.3 Visualization

To understand the trends and correlations, a dual-plot visualization is employed:

- The Force Plate (Fz) data is plotted against its time column to observe vertical force variations.
- The Shoe Sensor Data (s1, s2, s3, s4) is plotted together, showcasing how individual sensors behave during activity.

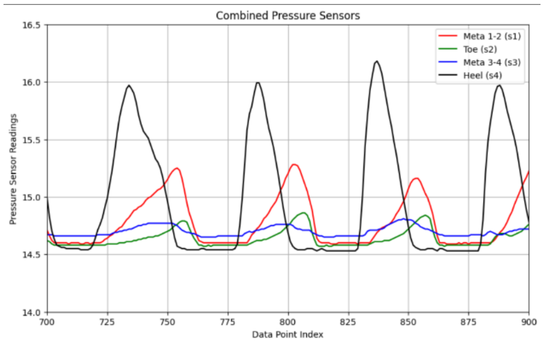


Fig 25: Pressure sensor data while walking

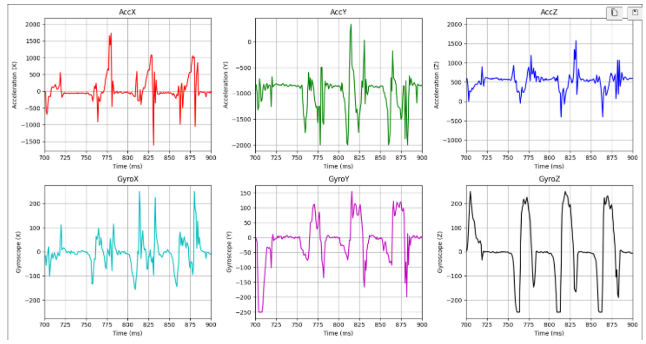


Fig 26: IMU data while walking

6.4.4 Synchronization and Cleaning

Synchronization is performed to align the timestamps of the shoe and force plate data:

1. Identifying Zero-Force Intervals: The force plate's Fz values are examined to find timestamps where no significant force was detected (indicating no activity). Corresponding timestamps in the shoe dataset are also identified.
2. Data Cleansing: At these timestamps, sensor values (s1, s2, s3, s4) are set to zero to remove noise and irrelevant data points. This ensures only meaningful activity data remains.

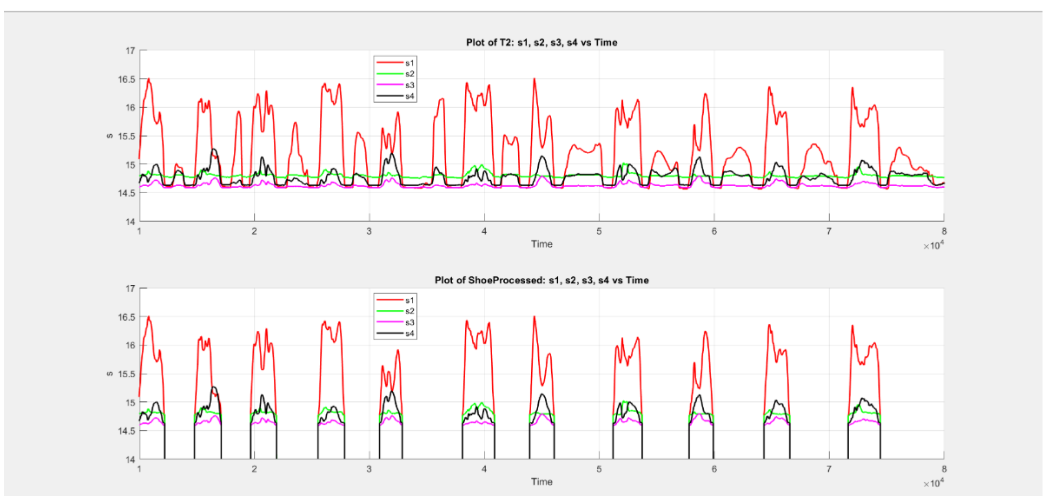


Fig 27: Data after removing instances when shoe was not on Force plate

6.4.5 Interpolation

Since the force plate operates at a higher sampling frequency than the shoe sensors, interpolation is used to match timestamps:

1. For each timestamp in the shoe dataset, the nearest corresponding timestamp in the force plate data is identified.
2. Data from the force plate is interpolated and aligned with the shoe timestamps to create a unified dataset.

6.4.6 Data Augmentation

To account for subject-specific variations, additional features such as the subject's weight are added to the combined dataset. For instance, a constant weight of 70kg is appended as a new column, which could be used as a factor during machine learning model training.

6.4.7 Final Cleanup and Export

Invalid or redundant rows, such as those where all sensor readings are zero, are filtered out. The cleaned, synchronized dataset is then exported as a CSV file for further analysis. This final dataset includes:

- Time-aligned force plate data (Fx, Fy, Fz)
- Shoe sensor data (s1, s2, s3, s4)
- Derived features like subject weight

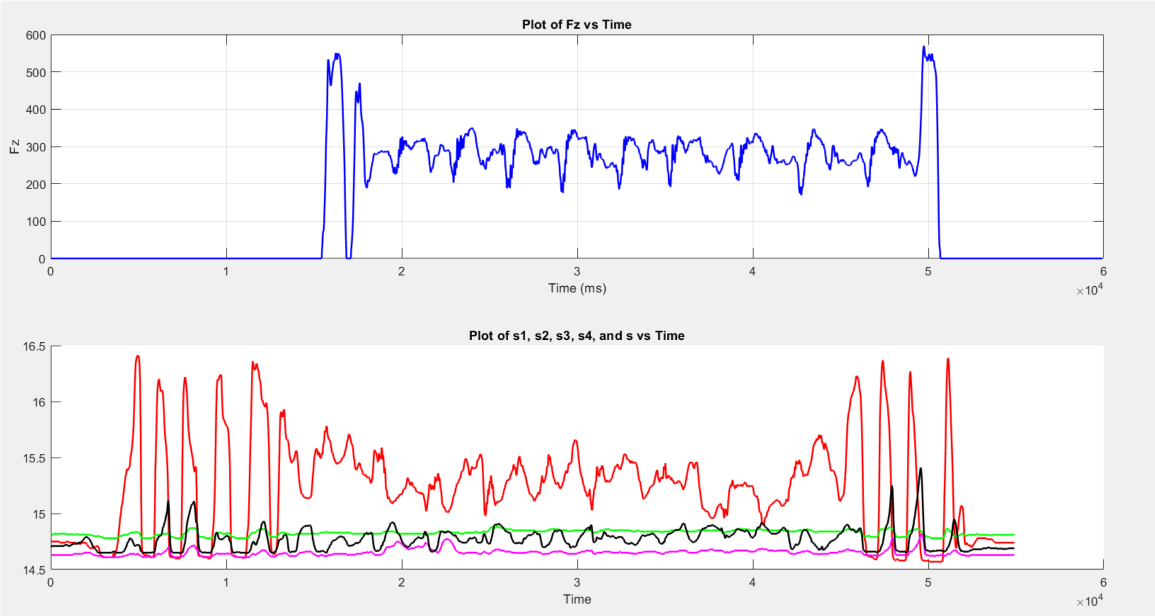


Fig 28: Squatting data before pre-processing

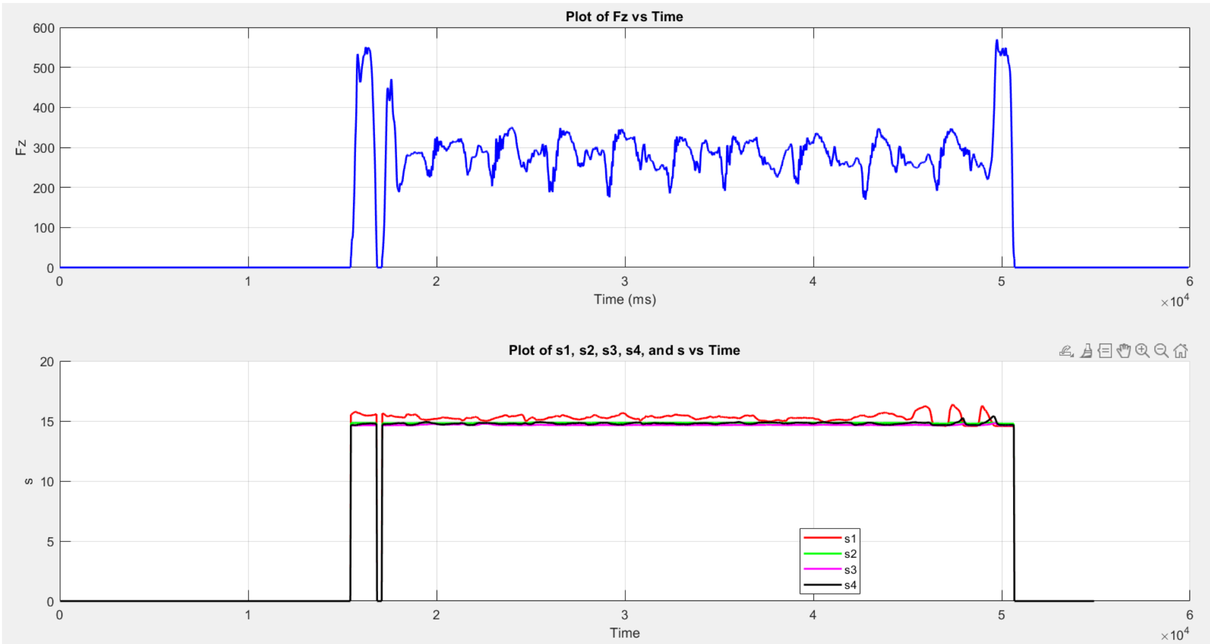


Fig 29: Squatting data after pre-processing

6.5. Machine Learning Model Development

The Machine Learning Model Development section focuses on leveraging both traditional regression techniques and deep learning models to predict Ground Reaction Forces (GRFs) from the insole sensor data. The process includes data visualization, model implementation, evaluation, and comparison to identify the most effective approach.

6.5.1 Visualization

To begin, the relationship between the input features (pressure sensor data) and the target variables (Fx, Fy, Fz from the force plate) was explored through data visualization. Scatter plots, correlation heatmaps, and time-series overlays were created to understand trends, correlations, and data distributions. These visualizations helped identify patterns, non-linear relationships, and potential outliers, guiding the selection of regression techniques.

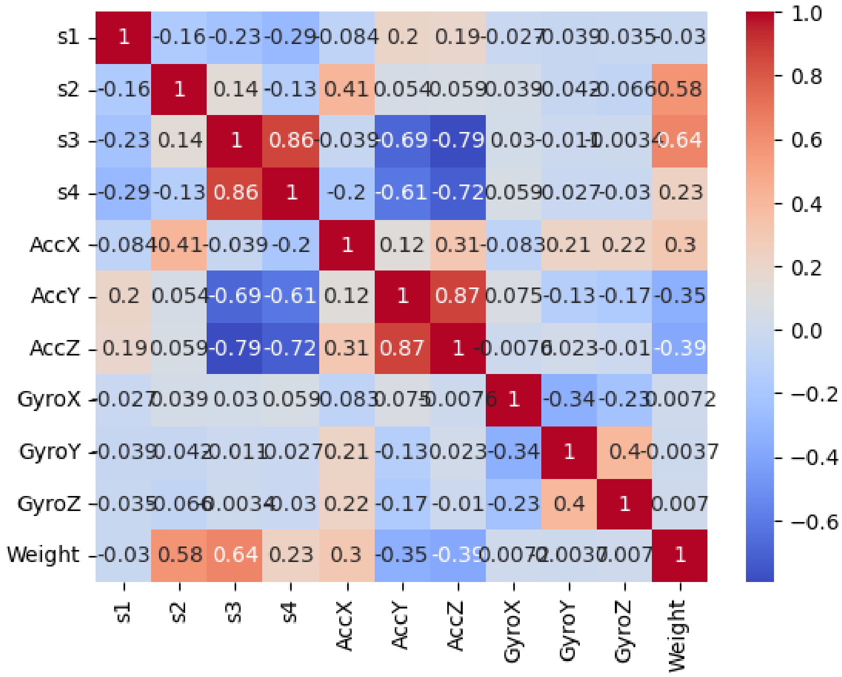


Fig 30: Heat map on correlation of features

6.5.2 Pipeline for Scaling

To standardize the dataset, we used a scikit-learn pipeline with a StandardScaler to ensure all features were on a similar scale, which is critical for the performance of regression models, especially those sensitive to feature scaling (e.g., SVR). The pipeline was simple and modular, providing easy integration into subsequent steps.

6.5.3 Dimensionality Reduction

Unnecessary columns that were not required for prediction were dropped to reduce dimensionality and improve model performance. This included outputs and redundant variables like timestamps and intermediary features.

Dimensionality Reduction for better performance

```
data = data.drop(["s", "currTime","SubFrame", "Frame","Fx","Fy","Fz","Mx","My","Mz","Cx","Cy","Cz"], axis=1)
data.head()
```

[8]

	s1	s2	s3	s4	AccX	AccY	AccZ	GyroX	GyroY	GyroZ	Weight
0	754.18	758.27	754.73	757.57	-215.82	-1025.88	333.01	0.34	-34.50	-54.49	70
1	755.85	758.69	754.79	758.29	-35.64	-990.72	399.90	33.84	-35.31	-52.63	70
2	758.30	759.12	754.98	758.66	109.86	-835.45	577.64	12.05	-6.95	-5.08	70
3	761.97	759.41	755.08	759.25	-91.80	-945.80	397.46	18.98	-16.66	-5.27	70
4	765.97	759.53	755.19	759.81	-111.33	-989.75	360.84	15.03	-3.83	-0.82	70

Fig 31: Dimensionality Reduction

6.5 4 Feature Correlation Analysis

A correlation heatmap was generated to evaluate the relationship between input features and GRF outputs. This helped in identifying highly correlated features, which could improve model efficiency by focusing on the most relevant inputs.

6.5.5 Regression Techniques

Four regression models were tested to evaluate their effectiveness:

- Linear Regression:
  - Served as the baseline model.
  - Captured basic trends but lacked the capability to model the dataset's complexity.
  - R^2: 28.21%, MSE: 841.47.
- Decision Tree:
  - Displayed overfitting tendencies and struggled with generalization.
  - R^2: 15.75%, MSE: 1731.94.
- Random Forest:
  - Improved generalization compared to Decision Tree but still fell short in accuracy.
  - R^2: 21.22%, MSE: 1157.88.
- Support Vector Regressor (SVR):

- Outperformed other methods by capturing non-linear relationships effectively.
- R^2: 58.40%, MSE: 487.66.

The SVR model emerged as the most effective traditional approach, demonstrating its ability to model the dataset's complexities. However, the limitations of traditional regression methods highlighted the need for more advanced techniques, such as deep learning, to achieve higher accuracy and generalization.

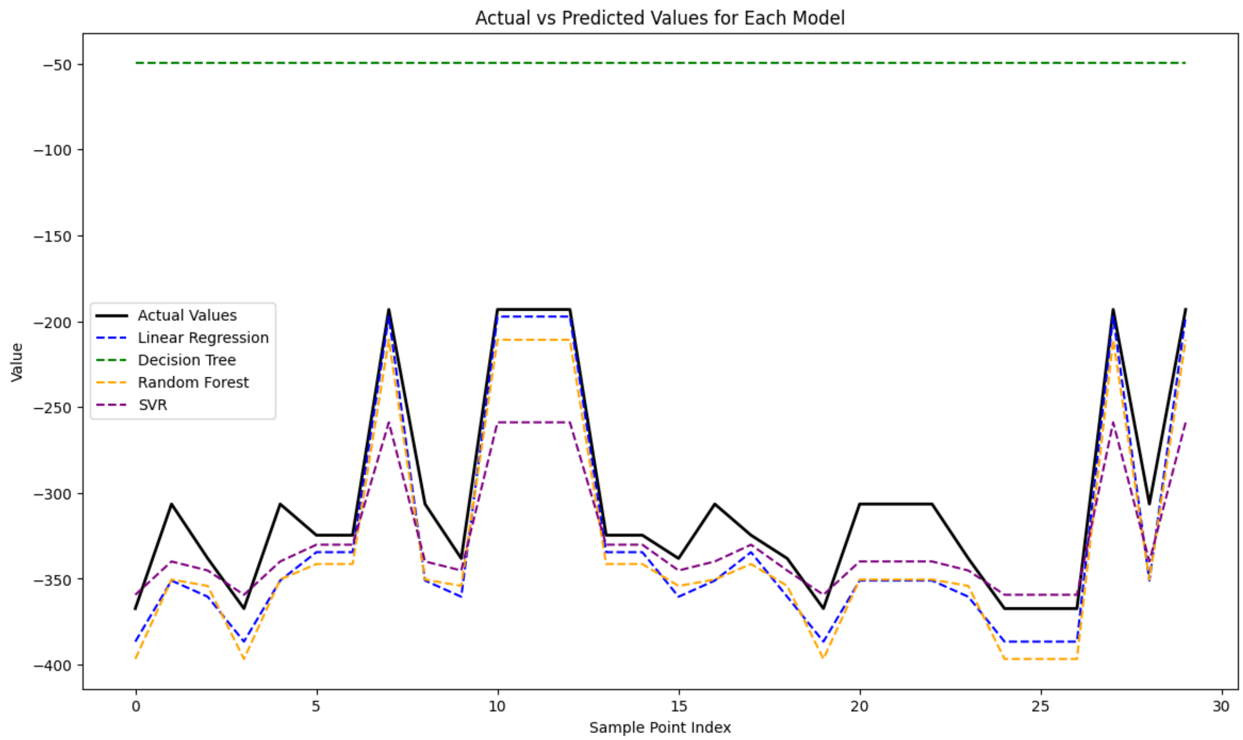


Fig 32: MSE of all regression models

6.5.6 Deep Learning Model Development

Deep learning is a powerful approach for modelling complex, non-linear relationships in data, making it well-suited for predicting Ground Reaction Forces (GRFs) based on sensor inputs. Unlike traditional regression models, deep learning can learn intricate patterns and interactions within the dataset, improving accuracy and generalization.

Advantages of Deep Learning

- Ability to handle non-linear relationships.
- Robust performance with large datasets.
- Automated feature extraction through its layered architecture.

1. *Feature Scaling*:
   All input features were scaled using the StandardScaler to standardize their range. This step ensured consistent gradients during training and faster convergence.
2. *Model Architecture*:
   - Input Layer: The model accepted input features corresponding to sensor data.
   - Hidden Layers: Two fully connected layers with 64 and 32 neurons, respectively, each using ReLU activation. These layers captured non-linear dependencies within the data.
   - Output Layer: A single neuron with linear activation for regression output.
   - The architecture was optimized for efficiency and accuracy, balancing complexity with the risk of overfitting.
3. *Model Compilation*:
   - Optimizer: Adam optimizer with a learning rate of 0.001 was used for adaptive learning.
   - Loss Function: Mean Squared Error (MSE), appropriate for regression tasks, minimized prediction errors.
   - Metrics: Mean Absolute Error (MAE) provided additional insights into model performance.
4. *Training*:
   The model was trained over 100 epochs using a batch size of 32. A validation split of 20% was employed to monitor performance on unseen data during training.

5. *Evaluation*:
   The model's performance was evaluated on test data, yielding the following metrics:
   - Test Loss: Low MSE indicated minimal deviation from actual values.
   - Test MAE: Provided insight into the average error magnitude.
6. *Performance Metrics*:
   - R^2: 92.22%, indicating a high degree of variance explained by the model.

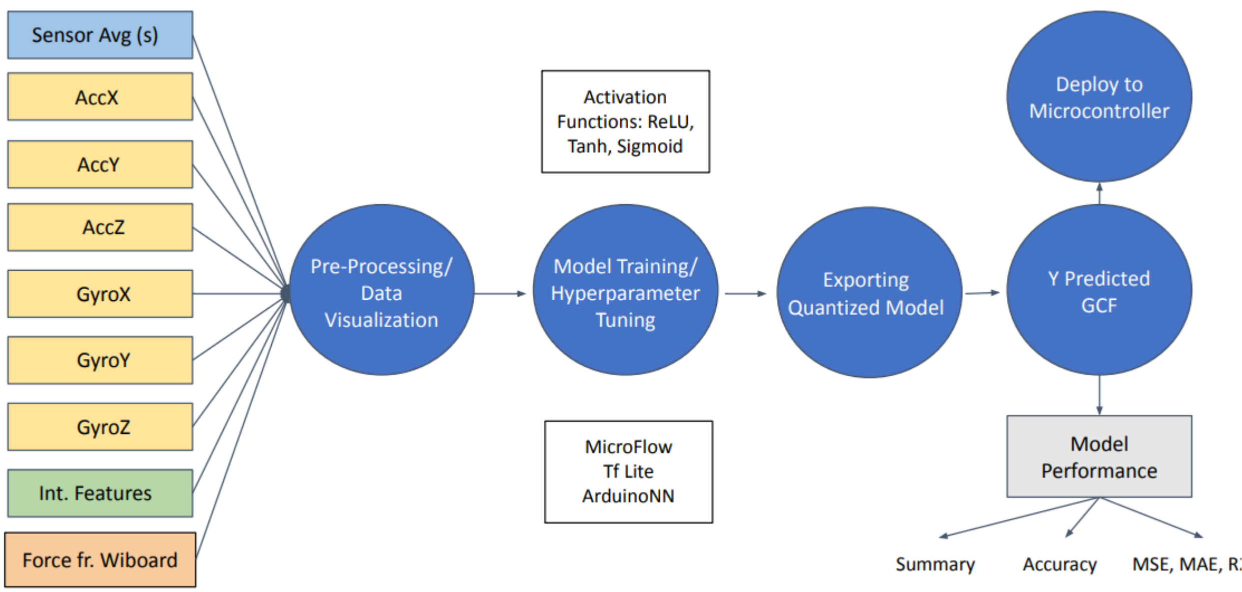


Fig 33: Flow chart of model training

7. System Testing and Validation

The system was deployed on participants during various physical activities, such as squats, walking, and step-ups, to assess its real-time data collection and prediction capabilities.

Deployment of the System on Participants for Real-Time Data Collection and Prediction:

The system was successfully deployed on multiple participants during the data collection phase. Each participant wore the instrumented shoes, which were connected to pressure sensors and a microcontroller for data acquisition. The real-time data collection process involved recording pressure sensor readings and IMU data from the shoes while the participants performed different activities. This setup enabled the real-time prediction of ground reaction forces (GRF) based on the collected sensor data. Throughout the experiment, the system provided continuous and uninterrupted data flow, demonstrating its capability to monitor the subjects' movements and generate predictions in real time.

```
Predictions vs Actual Values:

Data Point 1: Predicted = -331.55, Actual = -328.25
Data Point 2: Predicted = -389.84, Actual = -386.93
Data Point 3: Predicted = -355.04, Actual = -358.24
Data Point 4: Predicted = -201.61, Actual = -173.82
Data Point 5: Predicted = -296.75, Actual = -299.87

Predicted values: [-331.5521  -389.84354 -355.03867 -201.61392 -296.75223]
Actual values: [-328.249 -386.925 -358.241 -173.821 -299.872]
```

Fig 34: Results of Deep Learning model

Validation of Predictions Against Laboratory-Grade Systems (e.g., Force Plates):

To ensure the reliability and accuracy of the predictions made by the integrated system, the data collected from the shoe sensors was validated against data from laboratory-grade equipment, such as AMTI force plates. The force plate data served as the ground truth for comparing the predicted GRF values from the shoe-based system. In various activities like squats and step-ups, the predicted forces from the shoe system showed promising results when compared to the force plate measurements. The models demonstrated a good level of agreement with the force plate data, with R-squared values indicating moderate to strong correlation for some activities. This validation confirmed the potential of the portable system as a reliable alternative to traditional force plates for GRF estimation.

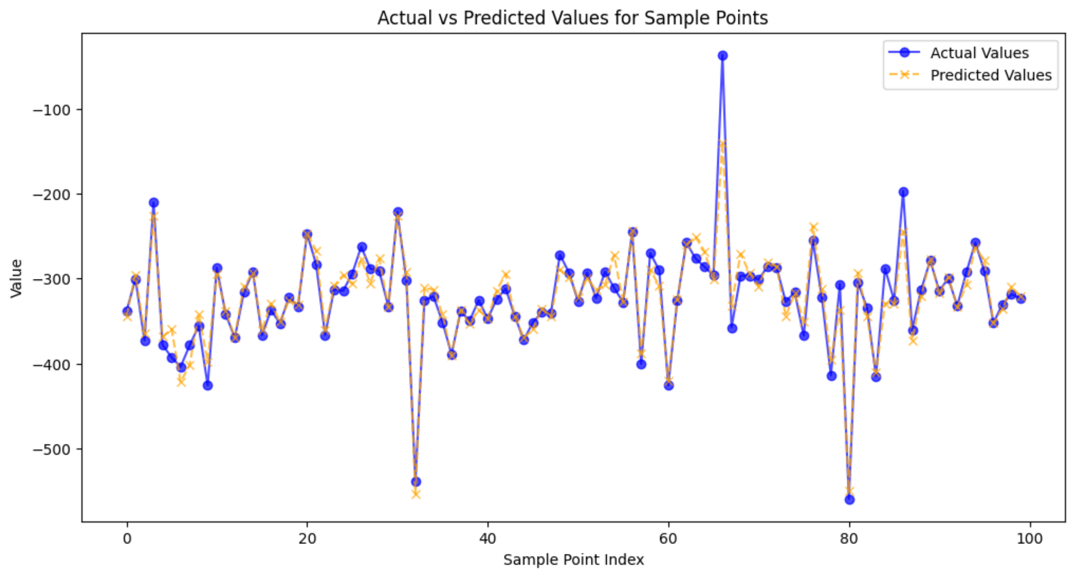


Fig 35: MSE of random points predicted using DL

Assessment of Portability, Ease of Use, and Data Transmission Reliability:

The system was designed with portability in mind, enabling easy deployment in real-world environments. The lightweight design of the shoe and the compact microcontroller system allowed participants to move freely without experiencing discomfort or hindrance during the experiments. The system was also designed to be user-friendly, with simple setup and operation instructions for both researchers and participants. During the testing, participants were able to wear the system for extended periods, performing various activities without issue.

Data transmission was carried out wirelessly, and the system was assessed for its reliability in transmitting sensor data to the connected processing unit. The wireless connection demonstrated robustness during all trials, with minimal data loss or delays, ensuring consistent and accurate data collection. This was particularly important for real-time predictions, where the timely transfer of data is crucial for accurate GRF estimation.

In conclusion, the deployment of the integrated system for real-time data collection and prediction showed that it is a viable, portable alternative to laboratory-grade systems like force plates. With validation against ground truth measurements, the system proved capable of reliably estimating GRF. The system's portability, ease of use, and dependable data transmission further highlighted its potential for practical applications in sports, rehabilitation, and other domains requiring movement analysis.

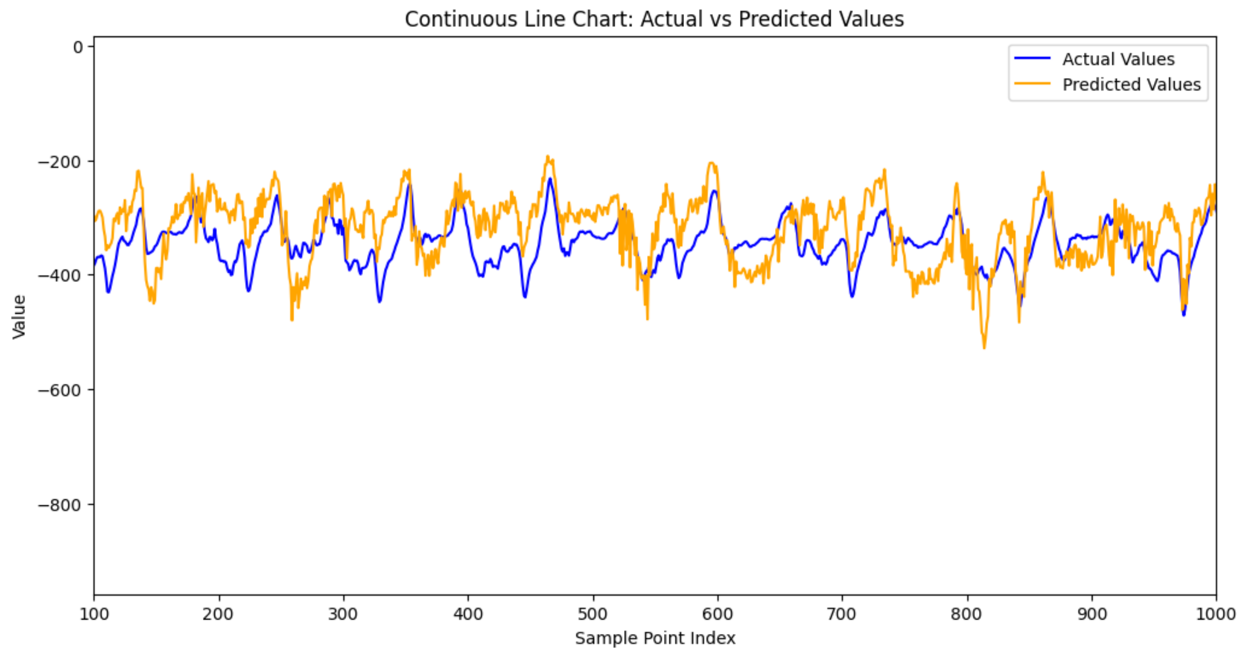


Fig 36: Complete Squat cycle predicted

8. Applications and Potential Use Cases

The developed wearable system for predicting Ground Reaction Forces (GRF) has significant potential across various fields due to its portability, cost-effectiveness, and real-time capabilities. By eliminating the need for traditional, stationary force plate setups, the system opens new avenues for practical applications, including:

8.1. Sports Performance and Training

- Biomechanical Analysis: The system can be used to monitor athletes' GRF during dynamic activities such as running, jumping, or weightlifting, providing insights into their biomechanics and areas for improvement.
- Injury Prevention: Real-time GRF monitoring allows coaches and trainers to detect abnormal force patterns that may lead to injuries, enabling corrective measures to be implemented promptly.
- Customized Training Programs: The data collected can help design personalized training regimens based on an athlete's unique force distribution and movement patterns.

8.2. Rehabilitation and Physiotherapy

- Post-Injury Recovery Monitoring: The system can track a patient's recovery progress by evaluating their GRF during various rehabilitation exercises.
- Gait Analysis: For individuals recovering from surgeries or injuries, the system provides a non-invasive method to assess walking patterns and ensure proper weight distribution.
- Assistive Feedback: Real-time feedback on movement can guide patients in performing exercises correctly, accelerating the rehabilitation process.

8.3. Clinical Research

- Understanding Movement Disorders: The system can assist researchers in studying conditions such as Parkinson's disease or cerebral palsy by analyzing how force distribution changes during movements.

- Orthopedic and Prosthetic Design: Data from the system can help in designing better orthopedic solutions or optimizing prosthetic devices by evaluating their performance during dynamic use.

8.4. Ergonomics and Workplace Safety

- Workforce Monitoring: In industrial settings, the system can assess workers' movements and detect improper lifting techniques, reducing the risk of musculoskeletal injuries.
- Design Optimization: Data-driven insights can inform the design of ergonomic footwear and workplace environments to enhance safety and efficiency.

8.5. Everyday Fitness and Health

- Wearable Fitness Devices: The system's technology can be integrated into consumer-grade wearable devices, providing users with advanced insights into their force distribution during daily activities.
- Weight Management and Posture Correction: Individuals can monitor their GRF to identify weight imbalances or poor posture and make necessary adjustments.

8.6. Military and Tactical Applications

- Performance Optimization: Monitoring the force exerted during marches or exercises can help improve soldiers' endurance and efficiency.
- Injury Risk Assessment: Identifying overexertion patterns can reduce the risk of injuries during training or missions.

9. Future Potential

With advancements in machine learning and sensor technologies, this system can be further miniaturized and enhanced for broader applications. The potential integration with wireless networks and smartphones could make it a widely accessible tool for both professional and personal use.

By addressing the limitations of traditional force plates, this system holds the promise of transforming the fields of sports, healthcare, and ergonomics, making advanced GRF monitoring more accessible and practical than ever before

10. References

[1] Prediction of Three-Directional Ground Reaction Forces during Walking Using a Shoe Sole Sensor System and Machine Learning

https://www.mdpi.com/1424-8220/23/21/8985#:~:text=The%20three%2Ddirectional%20GRFs%20during,sensor%20system%20for%20gait%20analysis.

[2] A Deep Learning Model for 3D Ground Reaction Force Estimation Using Shoes with Three Uniaxial Load Cells

https://www.mdpi.com/1424-8220/23/7/3428

[3] A 6 DoF, Wearable, Compliant Shoe Sensor for Total Ground Reaction Measurement

https://ieeexplore.ieee.org/abstract/document/8365154

[4] Predicting ground reaction forces in running using micro-sensors and neural networks

https://link.springer.com/article/10.1007/BF02844259

[5] Estimation of ground reaction forces and ankle moment with multiple, low-cost sensors

https://link.springer.com/article/10.1186/s12984-015-0081-x

[6] A Gait Monitoring System Based on Air Pressure Sensors Embedded in a Shoe

https://www.sciencedirect.com/science/article/pii/S1877705812021121

Image Citations:

Fig 1: https://www.mdpi.com/1424-8220/24/16/5318

Fig 2:
https://www.researchgate.net/publication/237050331_Ground_Reaction_Force_Analysis_of_Golf_Swings_using_Force_Plate_Data/figures?lo=1&utm_source=google&utm_medium=organic

Fig 3:
https://www.researchgate.net/publication/359163505_Investigation_of_Vertical_Ground_Reaction_Force_during_Walking_with_the_Exoskeleton_for_Patient_with_

Unilateral_Lower_Limb_Weakness/figures?lo=1&utm_source=google&utm_medium=organic

Fig 4: https://www.physio-pedia.com/index.php?title=Joint_Range_of_Motion_During_Gait&veaction=edit&section=8

Fig 5: http://www.graphicdevices.co.in/amti.html

Fig 6: https://www.sciencedirect.com/topics/engineering/force-plate

Fig 7: https://link.springer.com/chapter/10.1007/978-3-031-32781-0_6

Fig 8: https://www.youtube.com/watch?v=hM7xEoyP-4o

Fig 9: https://www.researchgate.net/publication/369517433_A_Deep_Learning_Model_for_3D_Ground_Reaction_Force_Estimation_Using_Shoes_with_Three_Uniaxial_Load_Cells/figures?lo=1&utm_source=google&utm_medium=organic

Fig 10, Fig 11: https://www.semanticscholar.org/paper/A-6-DoF%2C-Wearable%2C-Compliant-Shoe-Sensor-for-Total-Eng-Al-Mai/237dfdf240283556c7961359c1e736f42e9ab558

Fig 12: https://link.springer.com/article/10.1007/BF02844259

Fig 13: https://link.springer.com/article/10.1186/s12984-015-0081-x

Fig 14, Fig 18: https://www.semanticscholar.org/paper/Development-of-a-Smart-Insole-System-for-Real-Time-Stalin/83d572538b6a1fefe24f4ec09054c4988d068d81/figure/11