Singapore Institute of Technology - University of Glasgow
Joint Degree in Computing Science Degree Programme

CSC3001 Capstone Project

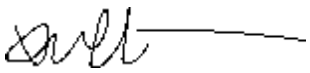
Please complete the following form and attach it to the Capstone Report submitted.

Capstone Period: 28 August 2023 to 11 August 2024
Assessment Trimester: Final Trimester
Project Type: Academic

Academic Supervisor Details:
Name: Seow Chee Kiat, Lawrence
Designation: Associate Professor
Institution: University of Glasgow
Email Address: cheekiat.seow@glasgow.ac.uk
Contact Number: +65 6908 6043

Student Particulars & Declaration:
Name of Student: Xu Xueli
Student ID: 2101812

I hereby acknowledge that I have engaged and discussed with my **Academic Supervisor** on the contents of this Final Capstone Report and have sought approval to release the report to the Singapore Institute of Technology and the University of Glasgow.

Signature 

Date: 15/7/2024

END OF FORM

Singapore Institute of Technology - University of Glasgow
Joint Degree in Computing Science Degree Programme

Final Capstone Report

Encoder-Based LOS/NLOS and Single-Bounce Detection using Channel Impulse Response

For the **Final Trimester** from 12 Apr 2024 to 1 Aug 2024

Xu Xueli
Student ID: 2101812

Academic Supervisor: Seow Chee Kiat, Lawrence

Submitted as part of the requirement for CSC3001 Capstone Project

Table of Contents

List of Figures

List of Tables

Acknowledgements

This project would not have been possible without the support of many people. I would like to thank my academic supervisor, Professor Lawrence, for his guidance, patience, encouragement, and committed assistance during my capstone project. I would also like to thank Professor Lou Xin for his constructive recommendations during the preliminary stages of this project.

1. Introduction

Indoor localization technology is revolutionizing navigation within buildings, akin to the impact of GPS on outdoor navigation. Despite GPS and GNSS's effectiveness outdoors, their indoor performance is hindered by signal weakening and multipath effects, driving the need for alternative indoor solutions. As indoor localization advances, it necessitates either the integration with existing infrastructure or the installation of new systems, along with thorough calibration for peak performance. This technology proves indispensable in environments where GPS is ineffective, such as malls and hospitals, and is particularly crucial for emergency services. For instance, in Singapore, it enables firefighters to accurately track their location indoors, enhancing both their safety and accountability. This underscores its importance to governments and organizations globally, highlighting its potential to revolutionize emergency protocols. The continuous development of indoor localization promises to enhance navigation, safety, and efficiency across various industries, indicating its broad impact and future growth potential.

However, technical obstacles such as Non-Line of Sight (NLOS) conditions, and diverse environments significantly challenge the advancement of ranging and localization technologies. Within these challenges, the distinction between Dominant Direct Path Propagation (DDPP) and Non-Dominant Path Propagation (NDPP) becomes crucial. DDPP refers to the scenario where the signal travels directly from the transmitter to the receiver without any obstruction, ensuring the strongest and most reliable signal path. Conversely, NDPP occurs in NLOS conditions where signals take indirect paths, bouncing off surfaces before reaching the receiver, often resulting in signal loss, or weakening and, consequently, inaccuracies in distance measurements and localization efforts.

This research focuses on enhancing the precision of indoor localization by developing advanced techniques to accurately identify Non-Line of Sight (NLOS) conditions, where Non-Dominant Path Propagation (NDPP) is prevalent, and to pinpoint the Time of Arrival (TOA) of single-bounce signals in the Channel Impulse Response (CIR). By effectively distinguishing between Direct Dominant Path Propagation (DDPP) and NDPP, the research aims to mitigate the NLOS effects on localization accuracy. Moreover, identifying single-bounce signals is crucial, as existing research highlights the potential of localization strategies based on single-bounce signals, underlining their significance in improving indoor localization technologies. This initiative is vital for unlocking the full potential of indoor localization technologies, thereby

paving the way for safer and more efficient indoor navigation.

1.1 Problem Formulation

The challenge of achieving precise indoor localization lies in the confluence of complex environmental conditions and the limitations of current technologies. At the heart of this problem are two primary issues: the impact of non-line-of-sight (NLOS) conditions on signal accuracy and the dynamic nature of indoor environments that continuously alter the efficacy of localization systems.

1.1.1 NLOS Conditions

Ultra-wideband (UWB) indoor positioning is hindered by non-line-of-sight (NLOS) conditions, leading to inaccuracies due to signal disruption. Research has improved localization accuracy by using multiple devices for redundancy, excluding NLOS-affected signals [1]. Efforts to identify NLOS conditions have employed various methods, ranging from statistical approaches and traditional machine learning to advanced deep learning strategies, including Support Vector Machines (SVMs) [2] and Convolutional Neural Networks (CNNs). Despite improvements, deploying many devices at various angles in critical service scenarios remains impractical.

1.1.2 Dynamic Environments

Dynamic indoor environments, with their changing layouts, moving objects, and fluctuating conditions, challenge the accuracy and adaptability of indoor localisation systems. Current methods, often reliant on fixed beacon placement, struggle due to the assumption of static environments, highlighting a gap in real-time adaptive solutions for such dynamic settings [3]. Existing evaluation methodologies for indoor localization systems primarily target interference-free, line-of-sight environments, which limits their applicability and generalizability to real-world, dynamic spaces [3]. This unresolved subproblem necessitates innovative solutions and adaptive algorithms that can effectively tackle the intricacies of dynamic indoor environments, enabling reliable and accurate indoor localization despite continuous changes in layouts, moving objects, and environmental variations.

The persistent challenges in achieving precise indoor localization primarily revolve around non-line-of-sight (NLOS) conditions and dynamic environments. NLOS conditions pose a significant hurdle, profoundly impacting accuracy and requiring rigorous methodologies for identification and mitigation. Meanwhile, dynamic indoor settings, characterised by shifting layouts and environmental conditions, demand innovative solutions. Traditional methods

based on static assumptions struggle to cope with these challenges, highlighting the urgent need for adaptive approaches that ensure localization systems are effective and dependable in complex, evolving indoor spaces.

1.2 Project Objectives/Project Specifications

The primary objective of this project is to enhance the identification of Non-Line-of-Sight (NLOS) conditions and single-bounce signals from Channel Impulse Response (CIR) data through the development of models. This project aims to achieve a milestone by reaching a classification accuracy of at least 90% for both Line-of-Sight (LOS) and NLOS conditions, surpassing existing models within the one-year timeframe. Additionally, the project targets identifying single bounce conditions from the CIR data with an accuracy of at least 90%, based on our experimental setup.

A key component of this endeavour involves creating and annotating a comprehensive dataset that includes Ultra-Wideband (UWB) readings and measurements of a single bounce path across a minimum of ten different scenarios. These scenarios will incorporate various types of obstacle materials. This dataset, crucial for the effective training of the model, is targeted to be completed within the first six months of the project.

In terms of project scope, our efforts will concentrate on the development and testing of machine learning models specifically designed for classifying LOS and NLOS conditions, as well as identifying single bounce paths in indoor environments. However, the project will not cover localization technology after signal identification and will not integrate with the device. Additionally, while performance testing of the model will be conducted across varied indoor settings, extensive field testing beyond the selected environments is not included.

2. Literature Review

2.1 Indoor Localization Technology

The literature review on indoor localization technology over the last decade emphasizes the advancements and applications of wireless communication technologies, with a significant focus on Ultra-Wideband (UWB) [4], Wi-Fi [5], and Bluetooth Low Energy (BLE) [6]. UWB is distinguished for its exceptional features, including high accuracy, fast communication rates, low energy consumption, and effective multipath interference mitigation due to its broad bandwidth, making it highly suitable for indoor navigation and adept at addressing non-line-of-sight (NLOS) condition challenges [7]. Furthermore, a comparative study revealed Decawave's UWB system to surpass Ubisense and BeSpoon systems in an indoor industrial environment by offering superior precision and reliability in both LOS and NLOS conditions. With its advanced antenna technology, Decawave's system achieved ±10 cm accuracy using two-way ranging (TWR) time-of-flight (TOF) measurements. It could measure up to three hundred meters in optimal conditions, showing a slight edge over BeSpoon and a significant advantage in accuracy and outlier management compared to Ubisense [8].

A key advantage of UWB in indoor localization is its use of the channel impulse response (CIR), which provides intricate details about the propagation channel's characteristics. The CIR analysis, benefiting from UWB's wide bandwidth, offers insights into signal propagation changes, including delays, amplitude variations, and phase shifts, with a high temporal resolution. This resolution, with sampling intervals as short as 1ns between CIR values [9], allows for a detailed representation of environmental features along the propagation path. This detailed mapping significantly enhances localization performance by distinguishing between line-of-sight (LOS) and non-line-of-sight (NLOS) conditions, making UWB an invaluable tool in the realm of indoor localization technologies.

2.2 Technical Approaches to NLOS Identification

In the realm of LOS and NLOS detection using CIR data, an array of statistical and machine-learning techniques has been deployed to augment both accuracy and reliability.

2.2.1 Statistical Approaches

This section explores statistical methods for distinguishing between line-of-sight (LOS) and non-line-of-sight (NLOS) conditions in multipath channels, focusing on two distinct studies that leverage channel statistics for effective differentiation.

The first study presents an innovative approach using amplitude, kurtosis, mean excess delay and RMS delay spread statistics to identify key features for distinguishing LOS from NLOS conditions. It proposes utilizing the joint probability density function (PDF) of kurtosis (κ), mean excess delay (τm), and RMS delay spread (τrms) to enhance accuracy, acknowledging the practical challenges of this method. As a feasible alternative, it suggests considering κ, τm, and τrms as independent variables and creating a composite metric (J) by multiplying their likelihood ratios. This method marks a departure from prior approaches that relied on a history of range measurements, requiring only amplitude and delay statistics from the multipath channel for NLOS detection. [10]

The second study investigates five channel statistics: kurtosis, mean excess delay, and RMS delay spread, applying a binary hypothesis test to distinguish between LOS and NLOS scenarios. It calculates a likelihood ratio (L) based on the statistics' probability density functions under LOS (H0) and NLOS (H1) conditions. With a specific decision threshold (λ), values of $L > 1$ indicate LOS, whereas $L < 1$ suggest NLOS. The study highlights kurtosis and skewness as particularly accurate for classification, albeit computationally demanding, while noting that other metrics may lead to less precise outcomes. [11]

Both studies underscore the potential of statistical methods to refine decision-making in identifying LOS and NLOS conditions. However, they also point to inherent challenges, such as suboptimal detection rates and the need for extensive manual calibration, which may limit their practical application.

2.2.2 Traditional Machine Learning

Exploring traditional machine learning, specifically Support Vector Machines (SVM), has shown promising results in overcoming limitations of statistical approaches for LOS/NLOS detection. SVM's appeal lies in its lack of dependence on prior data, making it well-suited for integrating three innovative features to improve accuracy under challenging conditions. These features are FCN (False Crest Number), which addresses high-energy false crests in NLOS scenarios by focusing on weakened Time of First Path (TFP) signals; FPE (First Path Energy), aimed at mitigating the impact of ambient noise on true first path signals; and FDE (First Path Energy Difference), which identifies energy variations indicative of NLOS paths due to their longer distances compared to LOS paths. This approach has led to a notable increase in performance, achieving a mean accuracy of 93.27%. [12]

2.2.3 Deep Learning

Deep learning approaches, particularly Convolutional Neural Networks (CNNs), have shown exceptional promise in enhancing LOS/NLOS classification in challenging environments. This exploration encompasses several studies that demonstrate the robustness and versatility of CNNs in indoor localization and the classification of wireless channel conditions.

One innovative method input Channel Impulse Response (CIR) graph images into a CNN, leveraging its resistance to distortion and strong generalization abilities. A short-time Fourier transform on the signal's impulse response creates a spectrum, using colour to represent energy intensities, thus framing channel identification as an image recognition task. This technique successfully distinguishes four channel environments—LOS, NLOS within 0-4 meters, NLOS within 4-10 meters, and extreme NLOS—achieving a remarkable 98.24% accuracy rate. This approach demonstrates the power of applying CNNs to wireless channel classification by transforming a complex signal analysis problem into a more tractable image classification challenge. [13]

Another study enhances indoor positioning accuracy by contrasting CNN's analysis of CIR graph images and raw CIR data. A novel approach is proposed, using raw CIR data to create a feature that represents CIR values based on a probability estimate of Line-Of-Sight (LOS) conditions. This feature, combined with manually selected features, is processed through a Multi-Layer Perceptron (MLP) to improve location accuracy.
The experimental setup spanned three environments: an office with iron and glass obstructions, a teaching building affected by multipath effects from concrete, wood, and pedestrian movement, and an underground mine with substantial multipath interference from

metal structures. In each case, data were collected with fixed anchor and tag placements over distances up to eleven meters, amassing over 15,000 data points per scenario. The findings reveal that using CNN to analyse raw CIR data is not only more time-efficient but also significantly more accurate than image-based CNN approaches. This method notably improves LOS/NLOS identification by 26.9% and achieves a 44.16% accuracy improvement over traditional image-based CNN techniques, highlighting its effectiveness in diverse and challenging indoor environments. [14]

Further expanding the deep learning horizon, a subsequent study evaluates localization system performance using MLP and CNN models. The methodology incorporates eight specific input features derived from data symbol transmission: firstPathAmp1, firstPathAmp2, firstPathAmp3, stdNoise, maxGrowthCIR, firstPathIndex, rxPreamCount, and Channel Impulse Response Power (C). These features were collected at the tag position through a setup involving three anchors and one tag, with data for each anchor recorded separately on a laptop. The system's performance was assessed through static and dynamic experiments to ensure accuracy and reliability. The initial performance metrics revealed that the CNN model outperformed the MLP model across all evaluated aspects: achieving an overall accuracy of 0.91, precision of 0.82, recall of 0.72, and an F1-Score of 0.76, compared to the MLP model's overall accuracy of 0.88, the precision of 0.80, recall of 0.56, and F1-Score of 0.66. This highlights the CNN model's superior classification accuracy and balance between precision and recall. [15]

Lastly, an exploration of three CNN models—A fully Convolutional Network (FCN), a Residual Network (ResNet), and an Encoder model that incorporates convolutional and dense layers with an attention mechanism before the final classification layer, for classifying line-of-sight (LOS) and non-line-of-sight (NLOS) conditions, the CNNs demonstrated superior accuracy in complex environments like corridors with reflective walls and industrial areas with varied propagation obstacles. The study emphasized the importance of optimizing the number of filters in convolutional layers to capture signal non-linearities effectively, with larger models yielding better results. The Encoder model achieved the highest classification accuracy at 94%, slightly outperforming the FCN and ResNet models at 93%. In contrast, traditional Support Vector Machine (SVM) classifiers, both linear and radial basis function (RbfSVM), significantly underperformed, with accuracies of 84% and 83%, respectively. This illustrates CNNs' superior capability in complex signal classification over classical machine learning methods, particularly in distinguishing between LOS and NLOS conditions with high precision

[16].

2.2.4 Transformers

The promising results achieved by the Encoder model, which incorporates attention mechanisms before its classification layer, have sparked interest in the potential application of transformer models for indoor localization.

A study explored the use of Transformer-based models for Non-Line-of-Sight (NLOS) detection in ultra-wideband (UWB) localization systems [18]. The implementation of the Transformer, specifically its encoder architecture, has proven effective in identifying NLOS conditions by processing the raw channel impulse response (CIR) data. Unlike conventional methods that rely on manually extracted features, the Transformer model autonomously learns to discern the intricate patterns indicative of NLOS conditions, improving the accuracy and reliability of NLOS detection. This approach not only enhances the model's performance over traditional convolutional neural network models but also adapts the advanced feature extraction capabilities of Transformers to a critical application in wireless communications, addressing one of the key challenges in indoor localization systems.

Given the versatility and success of transformers in various domains, their use in indoor localization represents an exciting frontier. Transformers, known for their ability to manage sequential data and capture long-range dependencies through self-attention mechanisms, could offer significant improvements in modelling complex indoor environments. This exploration into transformers aims to leverage their advanced feature extraction capabilities and robust handling of spatial and temporal relationships, which could enhance the accuracy and reliability of indoor localization systems. By adapting transformer architectures, which have shown remarkable performance in areas such as natural language processing and computer vision, there's potential for groundbreaking advancements in accurately mapping and navigating indoor spaces.

In summary, the evolution of LOS/NLOS (Line of Sight/Non-Line of Sight) detection technologies has developed significantly from statistical methods and today involves sophisticated machine learning and deep learning models, and now they are driven by innovative transformer models. This evolution reflects a relentless pursuit of improved accuracy, reliability, and adaptability to complex environments. In comparison to models like CNN, which would necessitate converting CIR graphs to images as an intermediate step,

numerical data offers more direct processing. These qualities are crucial for practical applications such as indoor positioning systems, where accurately distinguishing between LOS and NLOS conditions and system portability are paramount for optimal performance.

2.3 Single-Bounce Signal-Based Localization for NLOS Environments

In recent advancements within non-line-of-sight (NLOS) localization methodologies, studies have emerged, focusing on the utilization of single-bounce signal reflections to enhance accuracy and feasibility in environments where direct line-of-sight (LOS) signals are obstructed, rather than on the removal of NLOS signals from the localization techniques.

The first study introduces a hybrid time-of-arrival/angle-of-arrival (TOA/AOA) method that identifies single-bounce scattering signals from a singular scatterer, thereby simplifying the localization process. By recognizing these signals and pinpointing their origin, the scatterers are treated as virtual stations for mobile station localization. This algorithm is distinct in its reliance solely on TOA/AOA measurements without necessitating prior knowledge about the scatterer, base station, or mobile station. Simulation outcomes indicate an impressive 95% success rate in identifying single-bounce signals from the same scatterer, laying a robust groundwork for virtually eliminating localization errors that arise from scattering signal misidentification [19].

Conversely, the second study proposes a beacon-based approach tailored for outdoor NLOS environments, particularly enhancing RF source localization in search and rescue operations. This novel method integrates readings from RSS-AOA sensors with the dynamic positioning of beacons, which function as mobile helpers to delineate the search space and paths necessary for localizing both the reflector and the RF source. The deployment of two beacons, moving along predefined trajectories, facilitates the accurate localization of the reflector, thereby determining the RF source's bearing and position [20].

The third paper introduces an algorithm for 3D non-line-of-sight (NLOS) localization that significantly emphasizes the use of single-bounce NLOS paths alongside line-of-sight (LOS) paths in structured multipath environments. The algorithm constructs Lines of Possible Mobile Device Positions (LPMDs) for each single bounce reflection detected by reference devices, enhancing the accuracy of mobile device localization. By resolving the error variance through

these single bounce paths, the system mitigates location errors, demonstrating a significant improvement in positional accuracy [21].

In this fourth study, the focus is on utilizing the concept of image points derived from single-bounce NLOS reflections to improve localization accuracy in dense multipath environments. The paper details a method that uses image theory to reconstruct the probable position of NLOS single bounce reflections, which is pivotal for overcoming the traditional challenges posed by multipath effects. The approach shows promise through simulation results, which exhibit distinct image point shapes under various noise scenarios, suggesting a new avenue for NLOS localization refinement [22].

This research paper presents a novel peer-to-peer NLOS localization technique that leverages single bounce reflections to accurately determine the position of mobile devices in multipath environments. The method effectively utilizes the Time of Arrival (TOA) and Angle of Arrival (AOA) of single bounce NLOS signals to construct a defined area for mobile device positioning. This new approach markedly outperforms traditional methods, especially in scenarios with high levels of TOA and AOA measurement noises, by simplifying the localization process to require fewer signal paths while maintaining high accuracy. [23]

Lastly, this paper develops a two-step weighting process for Non-Line-of-Sight (NLOS) localization in multipath environments, focusing on single bounce signals to improve localization accuracy. The method employs a Line of Possible Mobile Device (LPMD) concept that only utilizes Line-of-Sight (LOS) and single-bounce-reflection paths to calculate the mobile device's position. Multiple-bounce paths are excluded due to their tendency to introduce errors into the system. By applying weights to these LPMD paths based on the signal strength (proportional to the inverse square of the distance travelled by the signal), and subsequently weighting the proximate points derived from these paths, the system more accurately predicts the mobile device's location by focusing on the most likely single bounce paths, significantly enhancing the localization process in complex environments.[24]

The studies highlighted exemplify significant advancements in handling non-line-of-sight (NLOS) challenges within localization technologies, particularly emphasizing the importance of single-bounce signal identification. The first two studies utilize hybrid TOA/AOA and beacon-based strategies, focusing on the precise detection and utilization of single bounce signals to overcome obstructions in direct line-of-sight (LOS) environments. These methods

display not only algorithmic ingenuity but also practical applications in scenarios like search and rescue operations, where accuracy is critical. Similarly, the papers on three-dimensional NLOS localization in indoor multipath environments, image-based NLOS localization, and the two-step weighting process for NLOS environments all stress the utility of single bounce paths. These approaches significantly enhance localization accuracy by carefully reconstructing the probable positions from these single reflections, effectively minimizing errors caused by multipath distortions. Collectively, these studies underline the transformative impact of accurately identifying and leveraging single bounce signals to robustly improve the precision and reliability of NLOS localization systems across various environmental settings.

3. Methodology

3.1 Data Collection

This study employed two DWM1001 modules, set up as a transmitter and receiver to gather data using Two-Way Ranging technology. The transmitter dispatches a signal which the receiver then acknowledges by sending a return signal. The primary dataset acquired was the Channel Impulse Response (CIR) at the transmitter's end, stored in the accumulator memory. These values are recorded at 1 ns interval, correlating to half the period of the 499.2 MHz fundamental frequency, with the memory covering one symbol time of 1016 samples at a 64 MHz mean PRF. The register file intricately logs each 8-bit segment of these complex values, providing a comprehensive analysis of both real and imaginary parts of every sample [9].

In the transmitter code, a poll message is configured and transmitted. Following this, the function continuously monitors the system status register to detect incoming frames or identify errors and timeouts. If a frame is correctly received, it processes this frame to extract and compute timestamps, which are then used to determine the time of flight and calculate the distance between the transmitter and receiver devices. After calculating the distance, the CIR data is read from the accumulator, computing amplitude values for each successful transmission. This process outputs the distance and a sequence of CIR values when connected via a serial connection, as shown in Figure 11. After the data collection process, each dataset was initially saved in a text file format. Before analysis, a data cleaning procedure was implemented to remove any records containing fewer than 1016 Channel Impulse Response (CIR) values and to eliminate duplicate entries. After cleaning, the datasets were consolidated and converted into a CSV format to facilitate easier manipulation and analysis.

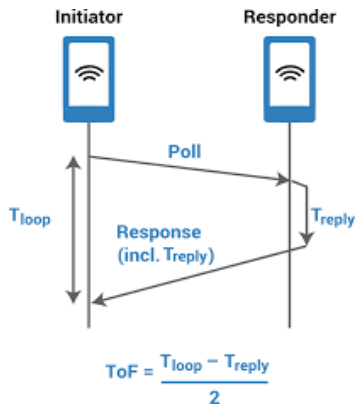


Figure 1. Time-of-flight Calculation [25]

Measurements were conducted across fourteen different scenarios, with the layout of the environments shown in the appendix, detailed in Tables 1 and 2: ten involved an unobstructed line of sight (LOS) between the modules, and four featured non-line of sight (NLOS) conditions with obstacles to simulate signal attenuation. These scenarios were carefully selected to represent a variety of indoor environments, each with unique multipath propagation characteristics influenced by dynamic furniture layouts. This diverse data collection aimed to prevent the resulting model from being overly fitted to specific environmental conditions. For accuracy in modelling single-bounce signal paths, distances in all scenarios were measured using a laser device, CONDTROL Traccer 30, with a precision of ±3mm.

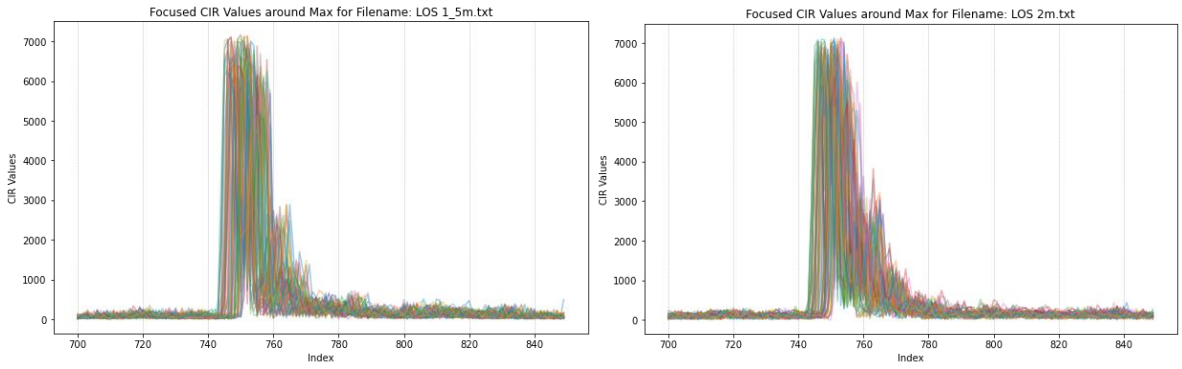
Table 1: Summary of Non-Line-of-Sight (NLOS) Scenarios with Corresponding Counts

NLOS Scenarios	Count
NLOS 5_7m.txt	176
NLOS corner door 13May.txt	79
NLOS 3_8m.txt	47
SIT DR NLOS 1_1m from ipad.txt	12
Total:	314

Table 2: Summary of Line-of-Sight (NLOS) Scenarios with Corresponding Counts

LOS Scenario	Count
LOS 2_5m.txt	95
LOS plaster 1_72M.txt	93
LOS plaster.txt	87
LOS 5m.txt	75
LOS 2m.txt	72
LOS 1_5m.txt	63
LOS plaster 1_62M.txt	63
LOS plaster 2_14M.txt	53
LOS 2_5M_LR.txt	20
SIT DR LOS.txt	17
Total:	638

The exploratory data analysis of the Channel Impulse Response (CIR) data revealed that the initial signal path typically emerged around the 700th index and then declined to baseline levels within the following 150 steps, as seen in Figures 1 and 2, indicating this segment as a focal point for detailed study.



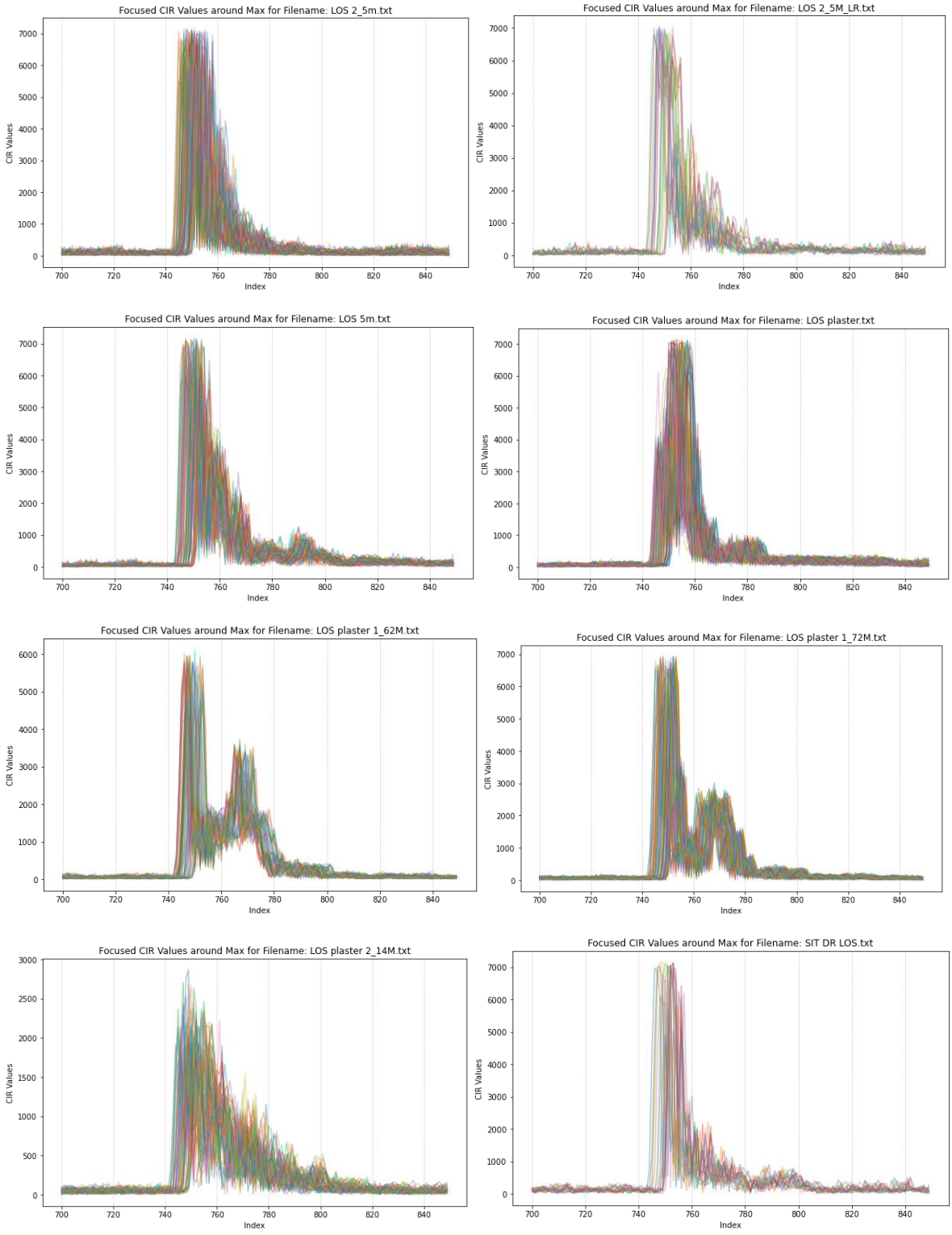


Figure 2: Combine CIR plots for LOS scenarios

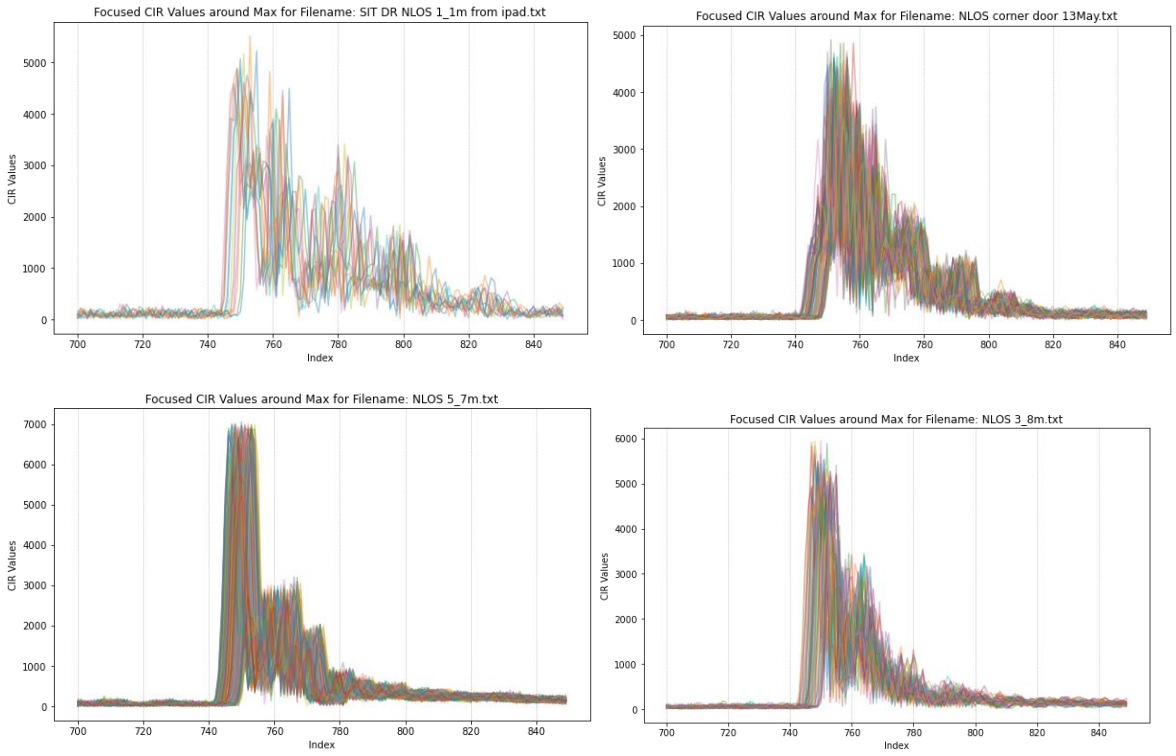


Figure 3: Combine CIR plots for NLOS scenarios

3.1.1 Enhanced Signal Analysis

In addition to extracting direct Channel Impulse Response (CIR) data from the DWM1001 modules for the model, several derived metrics were computed to capture the characteristics of the dataset more comprehensively. These metrics include the mean and standard deviation of the initial 700 CIR values, as well as those metrics calculated for the data points beyond the 800th index. The maximum CIR value was determined, along with the index at which this maximum occurs, and the mean and standard deviation of the overall signal. To further analyse signal dynamics, the index difference between the initial detected path and the point of maximum value was identified, as well as the index and value of this first detected path. Additionally, a threshold was established to identify significant peaks in the data; this threshold was set at the mean of the first 700 CIR values plus fifteen times their standard deviation. These metrics are essential for a detailed understanding of the signal characteristics and

support robust analysis.

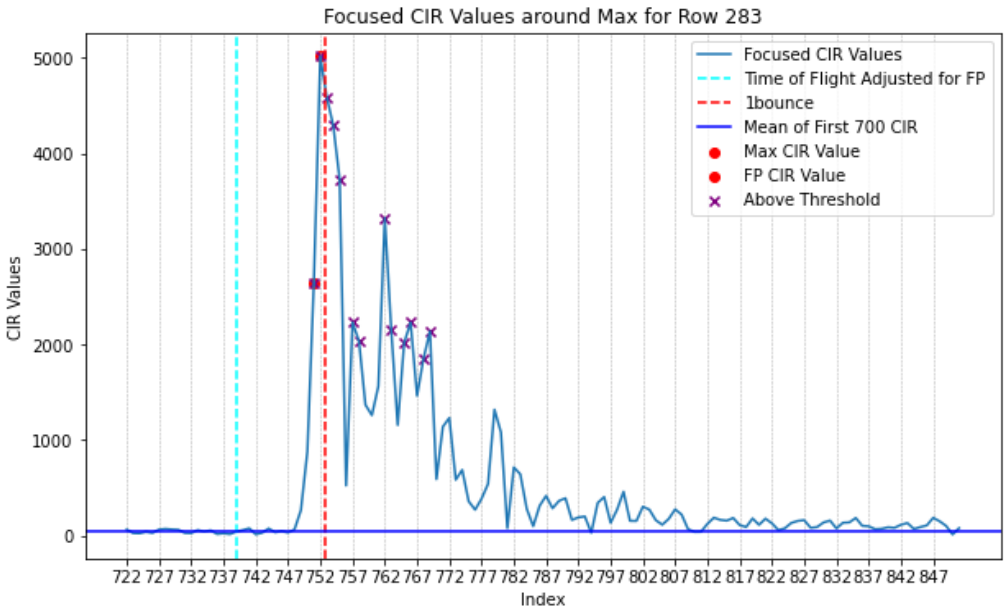


Figure 4: Focused CIR Plot

3.1.2 Time-of-Flight Distance Calculation

In our study, distance measurement between two devices utilizes the time-of-flight (ToF) method, based on the leading edge of the radio signal—the first detectable part that minimizes errors from signal reflections. According to the DWM1000 user manual, this approach ensures the ToF calculation is highly accurate by focusing on the initial signal received. The ToF is then calculated by multiplying the speed of light (approximately 299,702,547 meters per second) by the time it takes the signal to travel from the transmitter to the receiver, with the distance displayed through serial communication.

For practical analysis, the measured distance is converted back into Time of Flight (ToF) to align with the Channel Impulse Response (CIR) data sampling rate. This conversion involves dividing the distance by the speed of light to revert to time in seconds, which is then multiplied by 10^9To convert it to nanoseconds, matching the CIR sampling interval.

By precisely mapping the estimated ToF to the corresponding index within the CIR data, the research can accurately identify the point at which the signal is sent. This exact indexing is made possible because the starting point of the ToF within the CIR data is known. Consequently, this allows for converting the manually measured distance of the single bounce signal into a specific index in the CIR data.

3.2 Model Architecture

The study introduces a multi-model system designed based on the encoder components of the Transformer model, to address critical challenges in indoor localization. The modified encoder outputs embeddings that capture the relationship in the input data, via the self-attention mechanism. The first model takes the embeddings to assess whether the data belongs to Non-Line-of-Sight (NLOS) or Line-of-Sight (LOS) conditions, which can directly impact the accuracy and reliability of indoor positioning systems. Upon detection of NLOS conditions, which are prevalent in complex indoor environments and significantly hinder accurate localization, the data is advanced to a secondary model. This model is specifically designed to identify single-bounce signals, which is crucial for enhancing localization accuracy in non-line-of-sight (NLOS) environments. By distinguishing these single reflections from multiple signal paths, the model significantly reduces errors caused by multipath distortions, thereby improving the reliability and precision of localization technologies in scenarios where a clear line of sight is obstructed. This capability is particularly important in applications such as search and rescue operations, where precise location information can be lifesaving. Finally, the third model calculates the distance of the detected single-bounce signal. The overall structure of the multi-model system is shown in Figure 5.

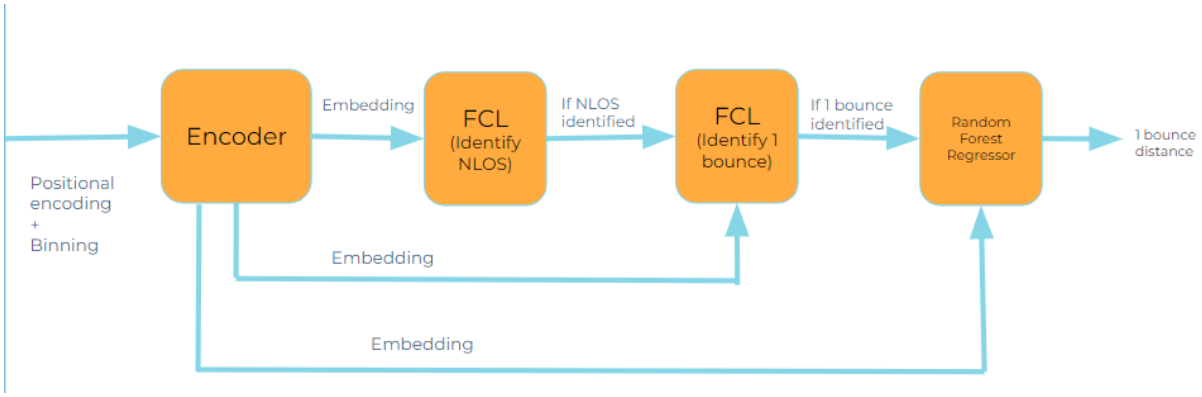


Figure 5: Multi-model System

For the training of these models, the UWB CIR dataset was segmented into three parts: 70% allocated for training, and 15% each for testing and validation. This distribution was maintained to preserve the original distribution of LOS and NLOS labels, with 638 cases of LOS and 314 cases of NLOS. The training process commenced with the models being trained on the training set, followed by evaluations on the validation set to fine-tune and adjust model parameters. Finally, the models were tested on the unseen test set to assess their performance and robustness across new data, ensuring a comprehensive evaluation of their capabilities.

3.2.1 Modified Encoder - Novelty

The modified encoder in this study incorporates an additional step before the application of positional embeddings, by binning CIR values into defined bins. Bins are set to include the left boundary and exclude the right, with ranges labelled 0, 200, 500, 800, 3000, 5000, and 6000. Binning categorizes the CIR data into discrete intervals, simplifying complex patterns into more manageable formats that are easier for models to interpret and learn from. This minimizes noise impact by grouping similar data points, reducing variability, and decreasing the number of distinct variable values, which simplifies the problem's dimensionality. It also stabilizes model performance by grouping data within specific ranges, lessening the effects of minor input fluctuations. This results in the reduction of the unique values from 9394 to 4831.

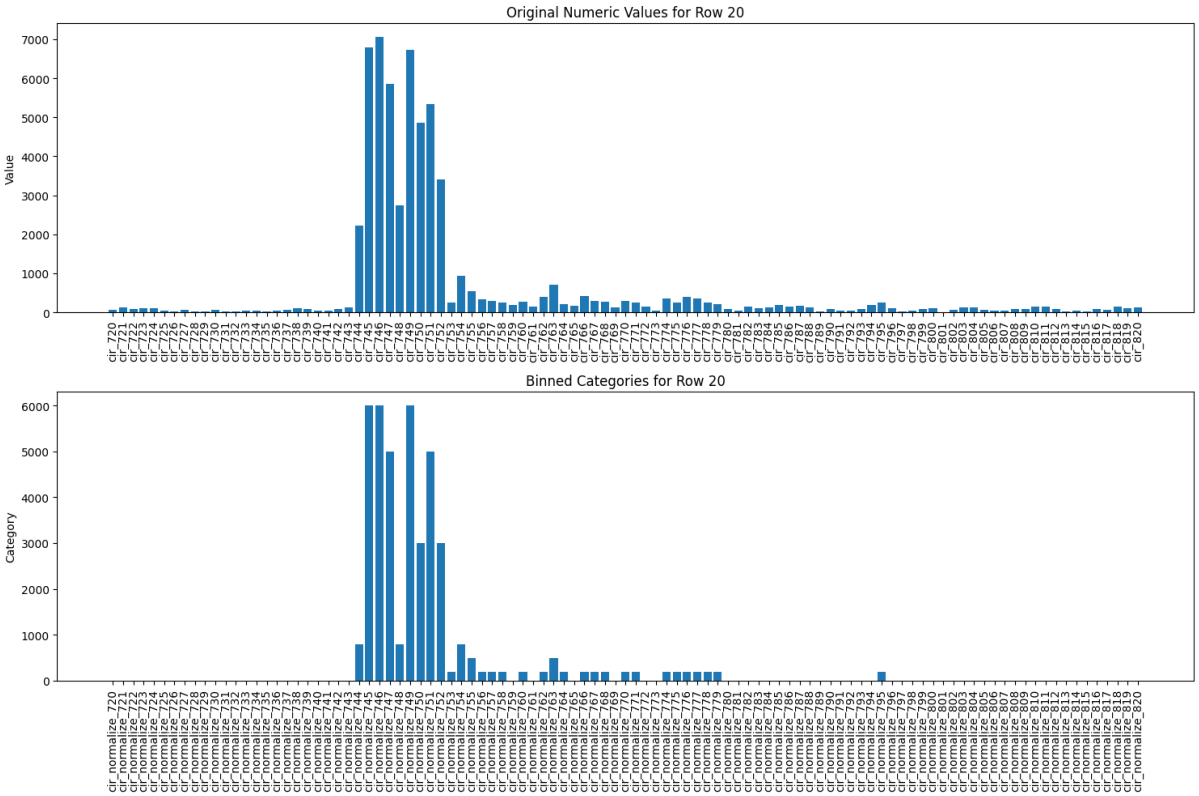


Figure 6: Effects on Binning CIR Data

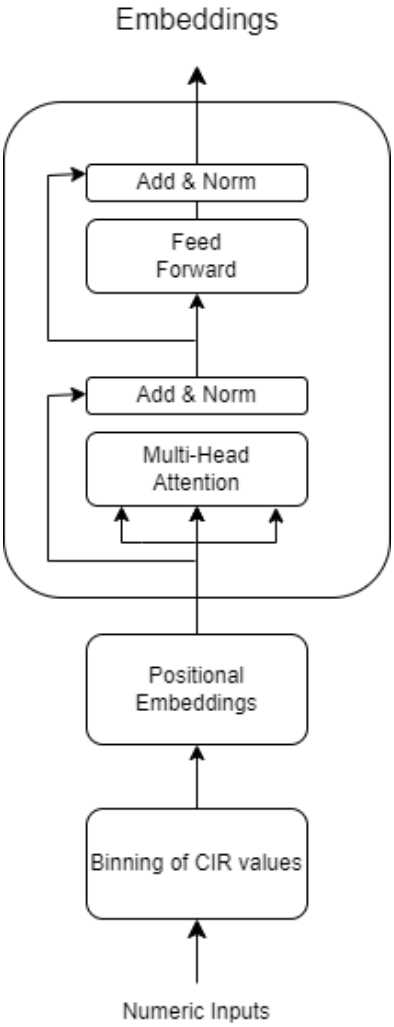


Figure 7: Architecture of Modified Encoder

The binned CIR values and derived metrics, such as the mean and standard deviation of the signal, utilize positional encodings with sinusoidal functions. These encodings embed the sequence order, which is essential for accurate data interpretation.

At its core, a multi-head attention mechanism simultaneously addresses various input segments, enhancing the model's ability to uncover complex dependencies. Each data segment is transformed through a dual-layer feed-forward network with ReLU activation, applied individually to each sequence position, facilitating detailed transformations.

The encoder architecture integrates these elements with residual connections and binning layers, crucial for training stability and convergence. It comprises three stacked encoder layers, each tailored with specific parameters for optimized performance.

The configuration includes handling up to 10,000 unique tokens and embedding them into a 504-dimensional space to aid in pattern recognition. The attention mechanism is divided into twelve heads for efficient parallel processing. Each encoder layer features a feed-forward network with a 3072-dimensional hidden layer, larger than the embedding dimension, to capture intricate features. Input sequences are limited to a maximum of 160 tokens to maintain consistency.

This encoder produces a rich embedding that encapsulates well-represented data attributes essential for further classification. These embeddings are adept at capturing complex patterns within the data, paving the way for enhanced performance in analytical stages.

3.2.1.1 Multihead Attention Mechanism

The multi-head attention mechanism, a pivotal feature of models like Transformers, allows the system to attend simultaneously to different parts of the sequence, ensuring a comprehensive capture of nuances and dependencies.

Principle of Operation:

Multihead attention is built on the idea of scaled dot-product attention, which calculates the attention scores based on the query (Q), key (K), and value (V) matrices. The formula for scaled dot-product attention is:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Figure 8: Scaled Dot-product Attention Formula

Where dk is the dimensionality of the keys and queries, and the SoftMax function is applied over each row, ensuring that they sum to one and are proportionally scaled, which allows the model to weigh the importance of each data point effectively.

1. **Linear Projections**: Each component of the queries, keys, and values transforms each attention head, defined by:

$$Q_i = QW_i^Q, \quad K_i = KW_i^K, \quad V_i = VW_i^V$$

Figure 9: Linear Projection of Q, K, V

Here,Wi^Q ,Wi^K,Wi^V Are parameter matrices for the *i*-th attention head. Each head processes the data independently, allowing the model to explore different perspectives of the information.

2. **Attention Computation in Each Head**: Within each head, attention scores are computed, focusing on different segments of the input data. This stage emphasizes the tailored treatment of data, segment by segment, thereby enhancing the detection of interdependencies among the input features.

3. **Concatenation of Heads**: The outputs of all heads are concatenated and linearly transformed again.

3.2.2 Classification Layers - Neural Networks

The embeddings produced by the encoder are then passed to a neural network with two linear layers. The first linear layer maps the input embeddings to a hidden space of size 150, followed by a ReLU activation function for introducing non-linearity, and a dropout layer with a dropout rate of 0.6 to prevent overfitting. The second linear layer maps the hidden representations to an output layer of size two. CrossEntropyLoss is used, which is suitable for multi-class classification tasks. The AdamW optimizer, known for its effectiveness in handling sparse gradients and adaptive learning rate capabilities, is chosen for optimization. A StepLR scheduler is used to decay the learning rate by a factor of 0.1 every 10 epochs, assisting in the fine-tuning of the model as training progresses. The model is set to train for up to 15 epochs with early stopping implemented to prevent overtraining.

This neural network architecture is used to build two classifiers one for Non-Line-of-Sight (NLOS) and another for single-bounce signal presence, where the input feature is the embedding from the encoder. The primary difference between model one and model two lies in the target values (y-values) used during training. For the training of model two, the dataset is specifically filtered to include only those instances labelled as NLOS. This filtering is critical

because, in line-of-sight (LOS) scenarios, the LOS distance is more accurate, and calculating its single-bounce signal becomes unnecessary, thus optimizing computational efficiency. However, in NLOS scenarios, not all signals exhibit a single bounce, many displays multiple bounces. In these cases, it is deemed that no single bounce signal is present.

3.2.2.1 Neural Network

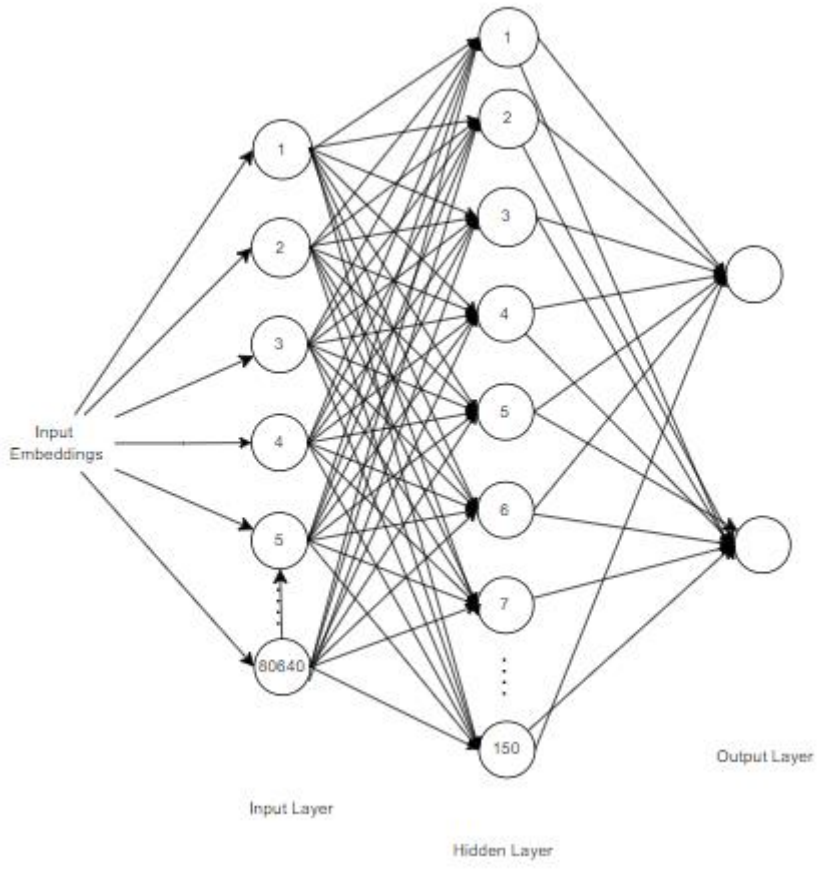


Figure 10: Neural Network Architecture

Neural networks are mathematical models designed to recognize patterns and make predictions based on input data by simulating the behaviour of interconnected neurons in a biological brain. The fundamental operation of a neural network involves a series of linear transformations and non-linear activations. Each neuron in a network computes a weighted sum of its inputs, This weighted sum is then passed through an activation function ϕ(z) (such as the ReLU, sigmoid, or tanh function) to introduce non-linearity, represented mathematically as shown in the equation, $y = \phi(w^T x + b)$, enabling the network to capture complex relationships in the data. Layers of neurons are stacked, allowing for the composition of multiple transformations, with the output of one layer serving as the input to the next. The network's parameters (weights and biases) are optimized during training using algorithms like

gradient descent, where the gradient of a loss function concerning the parameters is computed using backpropagation. This process iteratively adjusts the parameters to minimize the loss, thereby enhancing the model's predictive accuracy.

3.2.2 Regression layer - Predicting distance.

After confirming the presence of a single bounce signal with model two, the embeddings are processed by a Random Forest Regressor to predict the distance.

Configured with two hundred estimators, this model builds a robust collection of decision trees that delve deep into the data without a maximum depth constraint, capturing complex patterns effectively. It requires a minimum of ten samples to split an internal node and four at a leaf node, reducing the risk of overfitting. This method systematically estimates distances in scenarios involving NLOS and single-bounce signals. The Random Forest handles datasets with numerous features efficiently, which is well-suited for inputs that utilize embeddings, as each tree assesses a random subset of features, maintaining performance even with extensive input spaces. Additionally, it can identify and rank the most predictive features, enhancing model interpretability and focus.

3.3 Scalability of System

Ensuring scalability in the system is crucial for managing the complexities and variability of real-world environments where the system will be deployed. Scalable systems can efficiently handle increases in workload and adapt to expanding data volume without losing performance. This capability is especially important in indoor localization, where environments can vary widely from small rooms to large complexes, and the number of users or devices can fluctuate significantly.

To ensure the scalability of the system, the DWM1000 devices have been programmed such that the transmitter collects the CIR data from various receivers in a round-robin style, displaying it via a serial connection as shown in Figure 11. These CIR values are then processed and passed through our multi-model system to determine or predict whether there is a Line-of-Sight (LOS) path or a single-bounce path, and to calculate the distance if present.

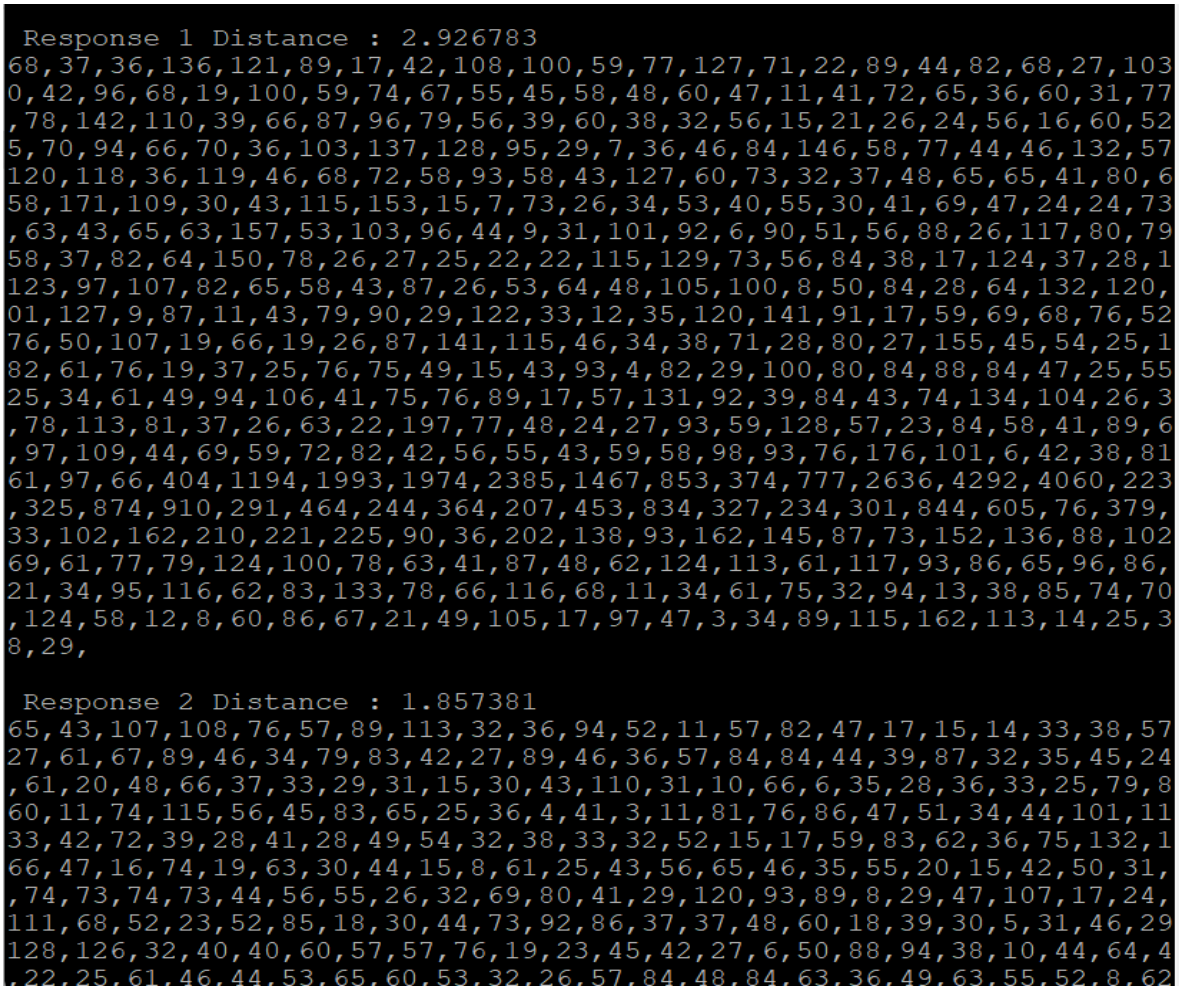


Figure 11: Data Collection of Multiple Devices via Serial Connection

The CIR values collected are subsequently processed through our multi-model system, designed to swiftly determine whether a Line-of-Sight (LOS) path or a single-bounce path is present and to calculate the distance if applicable.

4. Results and Analysis

4.1 Results

This section evaluates the results of the modified encoder by comparing its performance with several established classifiers. These classifiers include a traditional image-based Convolutional Neural Network (CNN), a basic Transformer model with an encoder-decoder structure, a BERT model [18] tailored for sequence classification (BertForSequenceClassification), and the unmodified encoder. Performance was assessed using accuracy and macro averages for various metrics like precision, recall, and F1 score,

ensuring that all classes are treated equally and evaluated fairly, regardless of their frequency in the dataset. The generalization abilities of the models were analysed based on loss graphs during training. Additionally, the computational complexity was measured by the time taken to train each model.

4.1.1 Model Performance

Table 3: Macro Average for Model 1

Model	ACC	Precision	Recall	F1-Score
CNN	0.94	0.94	0.93	0.94
Encoder-Decoder	0.70	0.35	0.50	0.41
BERT	0.97	0.96	0.97	0.97
Encoder w/o binning	0.97	0.97	0.97	0.97
Encoder w/ binning	0.99	1	0.99	0.99

From the results in Table 3, the CNN model demonstrates an accuracy (ACC), precision, and F1-score all at 0.94, with a recall of 0.93. The Encoder-Decoder model shows metrics with an accuracy of 0.70, precision of 0.35, recall at 0.50, and F1-score of 0.41. BERT and the Encoder models (both with and without binning) achieve scores above 0.96 in all metrics, with the binned encoder achieving near perfect scores in the precision, recall, and F1 score.

Table 4: Macro Average for Model 2

Model	ACC	Precision	Recall	F1-Score
CNN	0.97	0.98	0.93	0.95
Encoder-Decoder	0.75	0.37	0.50	0.43
BERT	1	1	1	1

Encoder w/o binning	0.98	0.99	0.96	0.98
Encoder w/ binning	1	1	1	1

In Table 4 for Model 2, the CNN model performed strongly with an accuracy of 0.97, precision of 0.98, recall of 0.93, and an F1-score of 0.95. The Encoder-Decoder model showed lower performance, with an accuracy of 0.75, precision of 0.37, recall of 0.50, and an F1-score of 0.43. The BERT model excelled, achieving perfect scores across all metrics, with 1 in accuracy, precision, recall, and F1-score. The Encoder without binning also performed well, with an accuracy of 0.98, precision of 0.99, recall of 0.96, and an F1-score of 0.98. Similarly, the Encoder with binning achieved perfect scores, with 1 in accuracy, precision, recall, and F1-score.

4.1.2 Generalization Capability

The analysis of training and validation loss, detailed in Appendix Figure 21, indicates that the binned CIR Encoder model demonstrated the lowest losses and exhibited the smallest disparity between training and validation losses. This was observable in the context of both identifying the NLOS condition and detecting single bounce conditions.

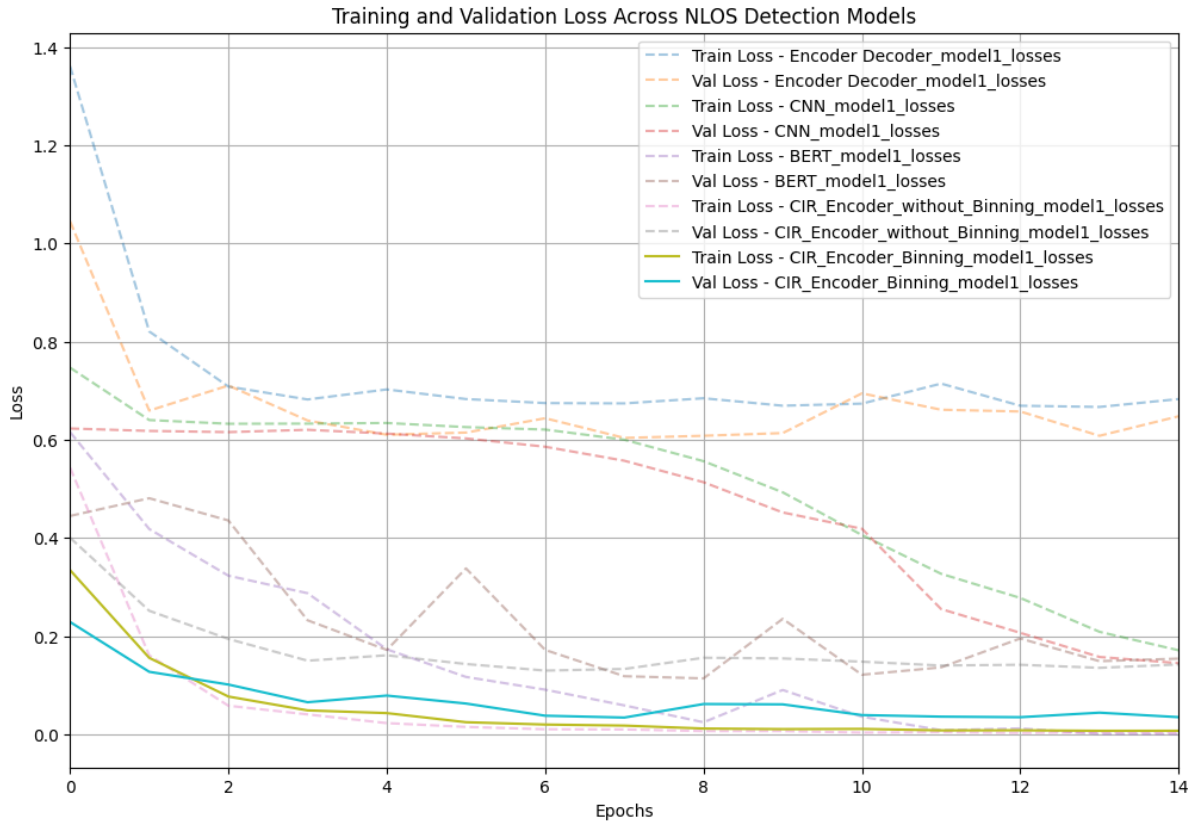


Figure 12: Combined Graph for NLOS Identification

As depicted in the combined graph, the binned CIR Encoder model showed a consistent and rapid decline in loss, stabilizing at a lower value early in the training phase. In contrast, the other models sustained higher loss values or exhibited spikes, particularly evident during the validation phases.

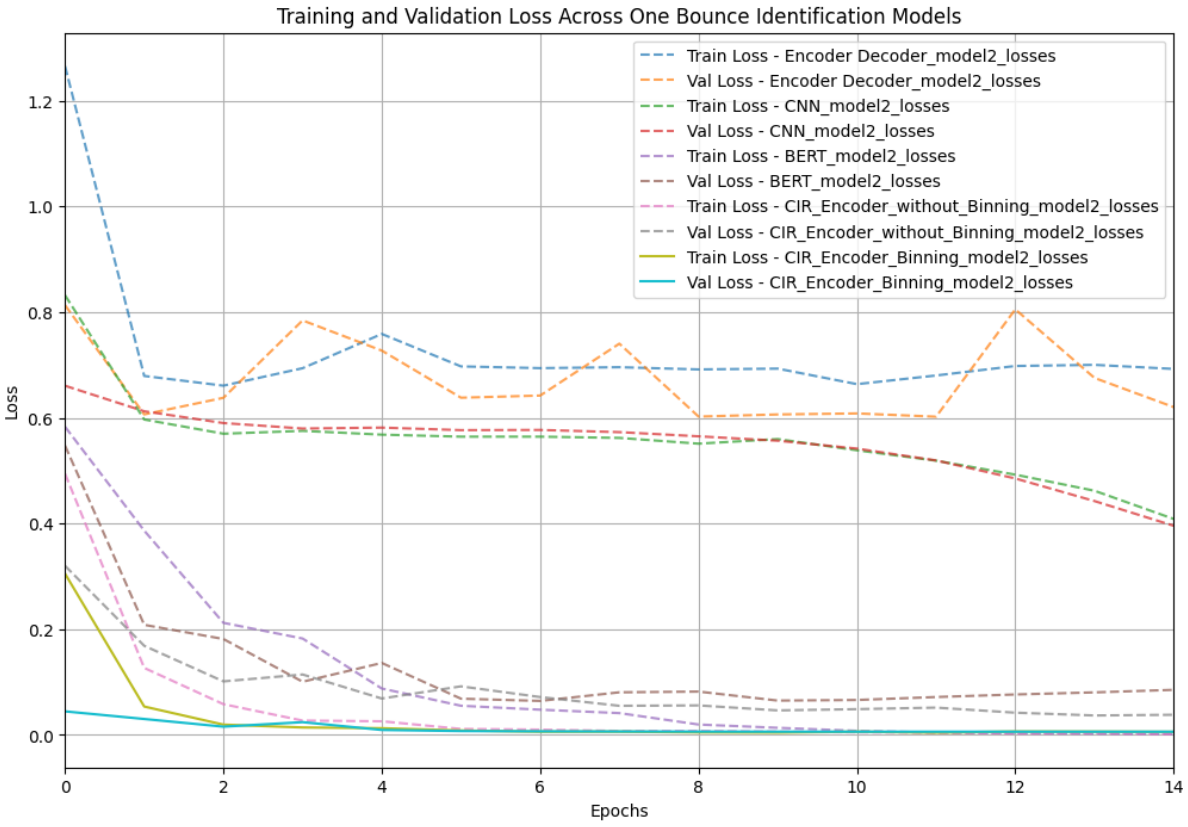


Figure 13: Combined Graph For 1 Bounce Identification

The training and validation loss trends for the single-bounce signal identification models over epochs highlight the effectiveness of the binning process. The binned CIR Encoder model achieved an immediate reduction in losses at the start and maintained low loss values throughout the entire training process. In contrast, the non-binned model exhibited higher loss values, especially in validation. The graph also shows the relative performance of other models like the Encoder-Decoder, BERT, and CNN, all of which showed improvements over time. The standout was the binned CIR Encoder model.

4.1.3 Computational Complexity

In assessing the computational efficiency of several machine learning models trained on Google Colab, distinct variations in training times were observed across different model architectures and tasks. All models utilized Google Colab's CPU except for the BERT model, which was trained on a Colab T4 GPU.

For identifying non-line-of-sight (NLOS) signals, the training durations were as follows:

- Encoder with binning: 257.59 seconds
- Encoder without binning: 266.47 seconds
- Encoder-Decoder: 2576.23 seconds
- BERT (on GPU): 736.78 seconds
- CNN: 51.59 seconds

For identifying single-bounce signals:

- Encoder with binning: 70.53 seconds
- Encoder without binning: 83.18 seconds
- Encoder-Decoder: 268.47 seconds
- BERT: 244.42 seconds
- CNN: 33.17 seconds

The CNN model required 539.28 seconds to process and save 952 plots as images.

4.1.4 Random Forest regressor

The Random Forest regressor's performance was detailed with several key metrics rounded to five decimal places. The Mean Squared Error (MSE) was 0.36010, the Mean Absolute Error (MAE) was 0.34594, and the Root Mean Squared Error (RMSE) reached 0.60008. Additionally, the R-squared (R^2) value was 0.61868, indicating the proportion of variance in the dependent variable that is predictable from the independent variables.

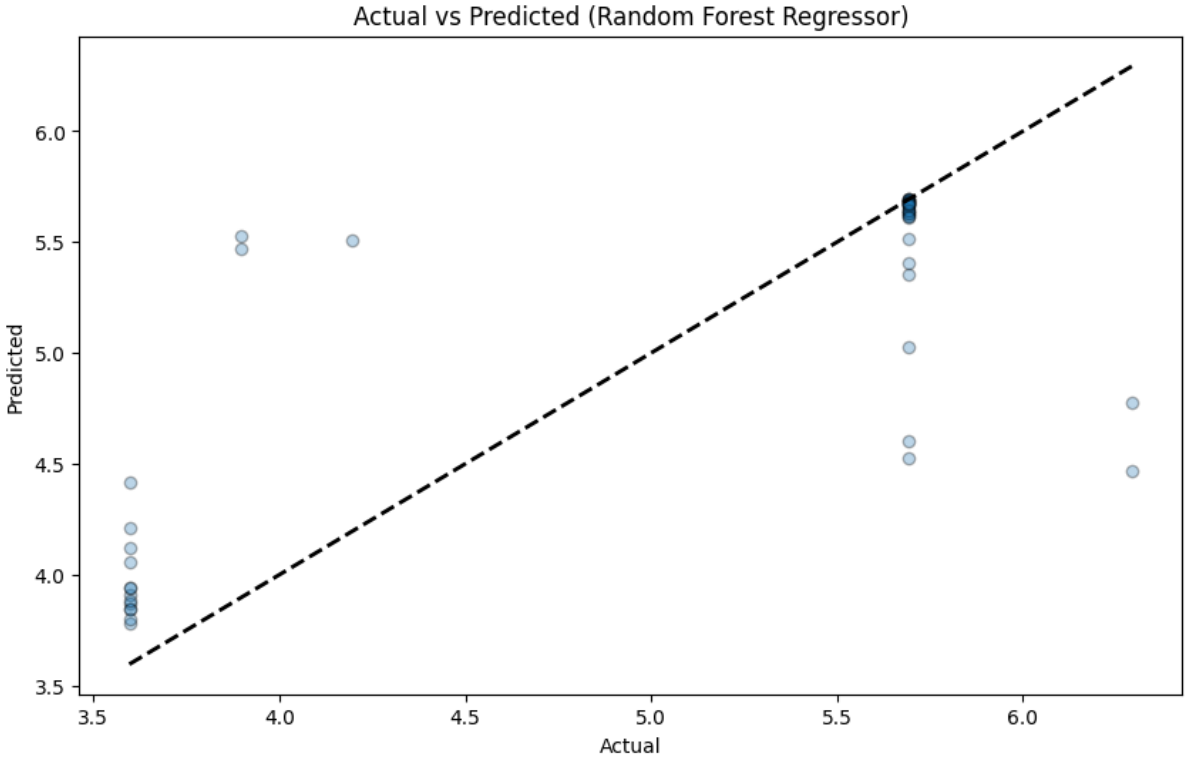


Figure 14: Actual vs Predicted Graph for 1 Bounce Distance Prediction

4.2 Analysis

4.2.1 Comparative Analysis of Model Performance

In the task of identifying Non-Line-of-Sight (NLOS) signals, the performance varies significantly among different models. The Convolutional Neural Network (CNN) model demonstrates robustness, yet it is surpassed by both versions of the BERT and the Encoder models, especially the version with binning, which approaches perfect metrics across all categories. In contrast, the Encoder-Decoder model underperforms markedly, suggesting its limitations in handling this specific task.

When focusing on the identification of single-bounce signals, the trend remains consistent. Both the BERT model and the binned Encoder model excel, showcasing their strong predictive capabilities. The CNN model performs better in this context than in the NLOS task, indicating a potentially better fit or more effective learning from the training data. However, the Encoder-Decoder model's continued suboptimal performance underscores its unsuitability for this use case, likely due to the complexity and increased parameter count inherent in its dual phases of encoding and decoding, which may also necessitate a larger or more varied dataset to effectively learn the task.

4.2.2 Insights from Training and Validation Loss Trends

The analysis of graphs from Appendix Figure 21 offers insightful revelations into the generalization capabilities of various models, specifically under two conditions: NLOS signal identification and single-bounce signal detection. For instance, after the fifteenth epoch, the CNN model's training and validation losses stabilize at 0.2, indicating effective learning from the training data and good generalization to validation data. However, the BERT models dedicated to NLOS detection initially show a decrease in validation loss but subsequently exhibit significant fluctuations from around the sixth to the fifteenth epoch. This erratic pattern, coupled with a large disparity between training and validation losses, signals potential issues in generalizing to new, unseen data. Similarly, the encoder model without binning demonstrates a larger gap between training and validation losses, suggesting that while it learns the training data well, it struggles to generalize this knowledge to new situations.

Among the evaluated models, the encoder with binning shows superior performance. Binning simplifies complex patterns into more manageable chunks, facilitating easier interpretation and learning by the models. This method effectively minimizes noise impact by grouping similar data points, which reduces variability and the number of distinct variable values, thereby simplifying the problem's dimensionality. It also stabilizes model performance by categorizing data within specific ranges, reducing the influence of minor input fluctuations.

While models such as the Encoder-Decoder, BERT, and CNN demonstrate progressive improvements in their loss values, their higher and more variable rates compared to the binned encoder suggest they are less adept at generalizing from training data to novel scenarios.

4.2.3 Assessment of Computational Complexity

The analysis of computational complexity among various machine learning models, as evidenced by their training durations, uncovers significant disparities that illuminate the influence of model architecture and task specificity on processing times.

In the task of identifying non-line-of-sight (NLOS) signals, the training durations exhibited wide variation across models. The Convolutional Neural Network (CNN) model completed its training in just over 51 seconds, showcasing its high efficiency for tasks involving image or signal processing using standard computational resources. On the other end of the spectrum, the Encoder-Decoder model demanded the most time, taking over 2576 seconds to train. This considerable duration can be attributed to its intricate architecture that incorporates both encoding and decoding processes. Such dual-component systems inherently add complexity, which tends to slow down the training process.

For the task of identifying single-bounce signals, the CNN model continued to demonstrate its efficiency with the shortest training time, clocking in at 33.17 seconds. The binned and non-binned Encoder models also performed efficiently, albeit with slightly longer training durations compared to CNN. Meanwhile, the BERT model, despite being trained on a more powerful T4 GPU, required more time than the CNN for NLOS signal identification. This duration reflects the intensive computation demands of its architectures, which are more resource intensive.

Moreover, the additional time the CNN model spent on processing and saving a large number of plots introduces another dimension of computational demand that extends beyond mere training. This consideration becomes especially important for potential AIoT implementations, where the device's memory and processing capabilities are as crucial as the efficiency of model training itself.

Overall, the encoder-decoder exhibited the lowest accuracy among the evaluated models, followed by the CNN. Although the CNN is noted for its speed, it requires an intermediate step of converting values into images, which adds complexity. In contrast, encoder-based models, such as the pre-trained BERT and those with and without binning, demonstrated superior performance according to the classification reports from the test dataset. When considering factors such as performance on the test dataset, generalizability, and operational speed, the binned encoder model stands out as the most effective, outperforming the other four models.

4.2.3 Reasons for BERT's Suboptimal Performance

The BERT model [26], designed for robust contextual understanding using its multi-head self-attention mechanism, excels at processing complex language patterns and is pre-trained on vast text datasets. However, it's primary training on text and reliance on WordPiece tokenization make it less effective for tasks involving numerical data, such as indoor localization, where data characteristics differ significantly from natural language.

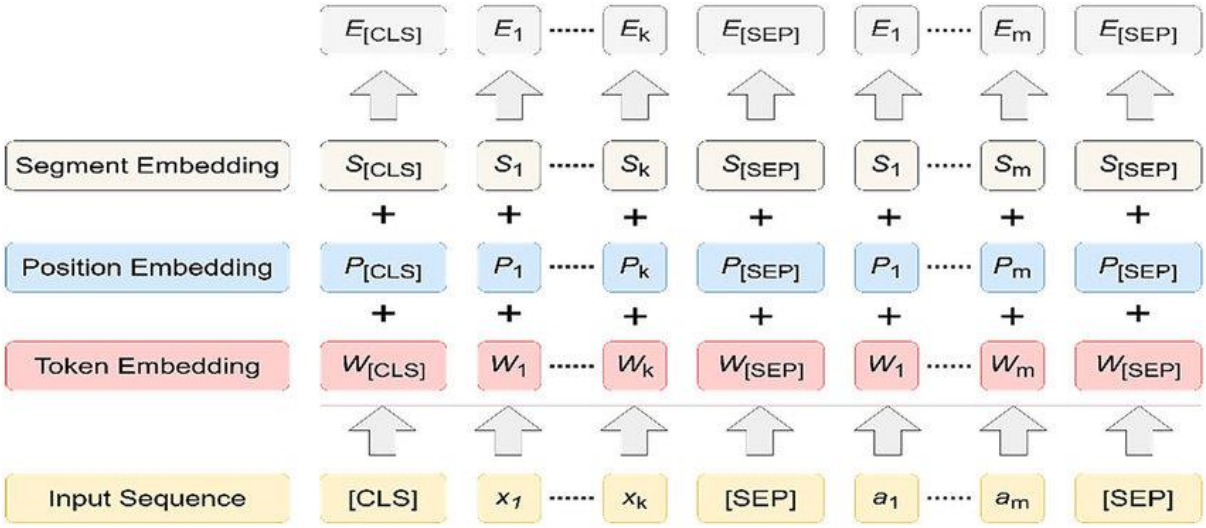


Figure 15: Input Embeddings in BERT [27]

An experiment investigating how BERT processes long numerical strings showcased that the model tends to split numbers into smaller tokens. For example, the number "2345" is tokenized into "234" and "##5". This method of tokenization treats sequences that are numerically connected as distinct entities, which could lead to a loss of contextual continuity and complicate the interpretation of data critical for Channel Impulse Response (CIR) analysis.

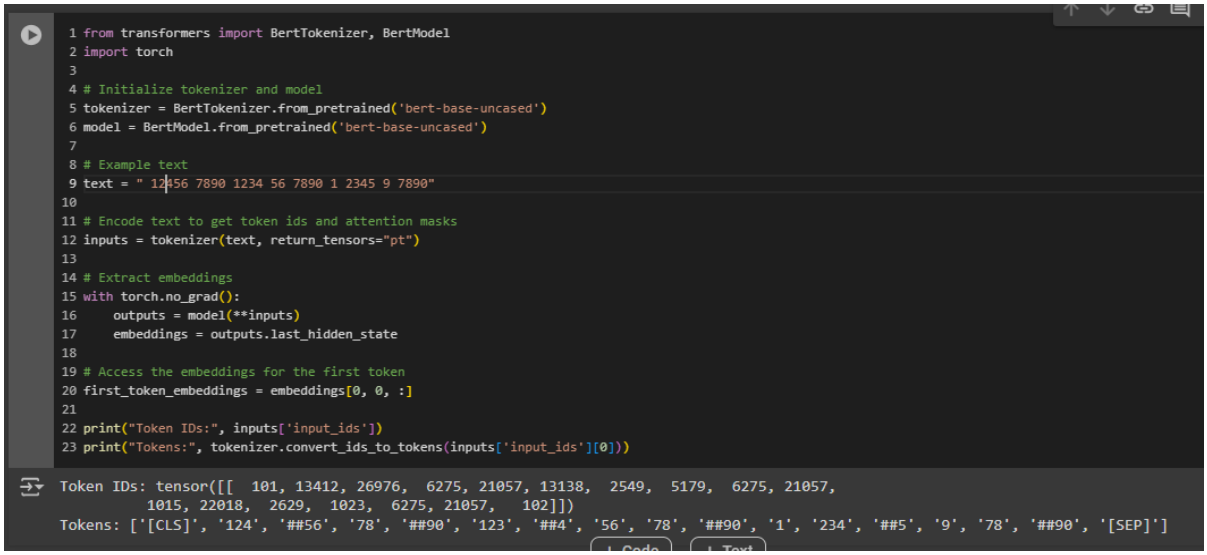


Figure 16: Tokenization of Numerical Values Using BERT

This tokenization issue underscores BERT's limitations when applied to numerical data, where the continuity and exact sequencing of numbers are crucial for accurate interpretation. The transition from BERT's training on natural language data to handling numerical sequences highlights a significant mismatch between the model's capabilities and the requirements of tasks like indoor localization.

A notable difference between BERT and the encoder model with binning is their approach to embedding. Both models use positional encoding and do not incorporate segment embedding, but BERT's reliance on WordPiece contrasts with the encoder's binning technique, which effectively handles numerical data by preserving integral sequences.

Further examination of architectural distinctions between BERT, the original Transformer (Encoder-Decoder), and GPT [28] models reveal their tailored design objectives and their suitability for specific tasks. Unlike the Transformer, designed for translation tasks by transforming input directly to output without a focus on bidirectional context, BERT includes a bidirectional training strategy. This strategy enables BERT to interpret context from both sides of a token, significantly enhancing its language processing capabilities. In contrast, GPT is designed for generative tasks, predicting the next token in a sequence from a one-way perspective, which may not capture the surrounding context as effectively as BERT's bidirectional approach.

These architectural nuances underscore why BERT, despite its advanced capabilities in language understanding, faces challenges when tasked with numerical data, particularly in specialized applications like indoor localization where precise data interpretation is crucial.

4.2.4 Criticism of the Dataset

The criticism of the dataset focuses on its imbalance, with 638 cases of LOS and 314 cases of NLOS. Despite this disparity, the classification metrics—precision, recall, and F1 scores—are high for both classes. This suggests that the minority class has enough data to learn distinguishing features, leading to strong performance overall. This is confirmed by robust macro and weighted averages, which affirm the models' effectiveness and indicate that the evaluation metrics do not disproportionately favour one class over the other.

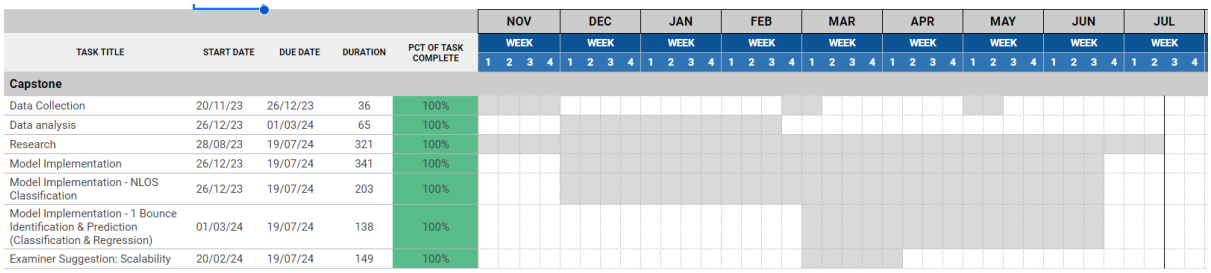
However, despite these positive outcomes, there are still some challenges associated with imbalanced data. These challenges include metrics that might not fully capture real-world performance, generalization capabilities that may be restricted when adapting to new, balanced datasets, and less reliable predictions for minority classes. These points highlight the importance of focusing on dataset diversity and balance to improve model robustness and reliability across various real-world scenarios.

5. Project Management

In managing the project, a systematic approach was adopted, incorporating a specific allocation of resources and tools. The project utilized a laser measuring device to ensure precise data collection, with a budget set at $79.90 for necessary materials and expenses. For computational tasks, including data processing and model training, Google Colab software was employed, providing a robust and accessible platform for executing machine learning algorithms and handling large datasets. DWM1001 modules provided by the school facilitated essential data gathering.

Here is an overview of the key stages involved in the project lifecycle, from initial data collection to implementing enhancements for scalability:

1. **Data Collection** - This initial stage focused on gathering all necessary data to support the entire project, ensuring that the subsequent analysis and model implementations were grounded in comprehensive and relevant datasets.
2. **Data Analysis** - Following data collection, this phase involved a detailed examination of the data to identify patterns and prepare it for effective application in modelling processes.
3. **Research** - Extensive research was conducted to explore existing models and methodologies, aiming to pinpoint areas where improvements could be made, or new approaches developed.
4. **Model Implementation - NLOS Classification** - This task was dedicated to developing a model for Non-Line-Of-Sight (NLOS) classification, tackling one of the project's significant challenges.
5. **Model Implementation - 1 Bounce Identification & Prediction** - This stage focused on creating a model capable of identifying and predicting single-bounce signals, which are critical for accurate localization and mapping applications.
6. **Examiner Suggestion: Scalability** - The final phase centred on implementing the examiner's suggestion by enhancing the system's scalability to ensure its efficient application in more extensive and complex real-world scenarios.



TASK TITLE	START DATE	DUE DATE	DURATION	PCT OF TASK COMPLETE
Capstone				
Data Collection	20/11/23	26/12/23	36	100%
Data analysis	26/12/23	01/03/24	65	100%
Research	28/08/23	19/07/24	321	100%
Model Implementation	26/12/23	19/07/24	341	100%
Model Implementation - NLOS Classification	26/12/23	19/07/24	203	100%
Model Implementation - 1 Bounce Identification & Prediction (Classification & Regression)	01/03/24	19/07/24	138	100%
Examiner Suggestion: Scalability	20/02/24	19/07/24	149	100%

Figure 17: Gantt Chart of the Capstone Project

Strategies and Skills:

- **Machine Learning Technical Expertise:** The project demanded the application and understanding of advanced principles in machine learning with emphasis on encoder-based models. Such models did play a very important role in the detection and classification of the signal data and, therefore, had their tuning done carefully to adapt and be robust to complex cases of LOS/NLOS and single-bounce signal scenarios.
- **Expertise in Transformer Models:** The expertise of the project lies in the extensive use of transformer models to leverage their advanced capabilities in the handling of sequence-to-sequence tasks. This needed adaptation of the native transformer architecture, so it could be better used to handle the signal specifics in indoor localization. This approach has highlighted the novelty applied in considering how model adaptation and optimization are approached.
- **Analytical and Problem-Solving Skills:** These have been key at the stage in which data analysis was being conducted. The team has been good at mining values out of complex datasets to be able to identify subtle patterns that may be a critical factor in the accuracy of the localisation process. This rigour of analysis ensured that the resulting models were robust and capable of working in many different tough conditions.

Project Deliverables:

1. Comprehensive UWB Dataset:
    - Minimum of ten different scenarios
    - Diverse types of obstacle materials
2. Improved Model for NLOS and LOS Identification:
    - Model achieving at least 90% in classification accuracy.
    - Documentation detailing model architecture, training process, and evaluation metrics

3. Identification of single-bounce Signals:
    - Model achieving at least 90% accuracy in identifying single bounce conditions from CIR data.
    - Documentation detailing model development and performance evaluation
4. Technical Report:
    - Comprehensive report summarising project objectives, methodology, findings, and recommendations.
    - Includes insights on dataset creation, model development, and experimental results.
5. Presentation:
    - Slide deck for presenting project overview, methodology, and key findings.
    - Presentation materials suitable for stakeholders and technical audiences

6. Conclusion

This study significantly advances non-line-of-sight (NLOS) indoor localization by integrating a multi-model system featuring a binned encoder architecture, achieving high accuracy and efficiency. It effectively addresses critical NLOS challenges, enhancing localization precision in obstructed environments essential for applications like emergency response and navigation. The utilization of binned Channel Impulse Response (CIR) data has improved model performance and generalization, proving superior to traditional methods such as CNNs, which demonstrated limitations in handling detailed numerical data. Our approach aligns with findings from Si, Wang, Siljak, Seow, and Yang (2023). By processing raw CIR data directly, our model retains detailed signal characteristics, achieving near-perfect accuracy, and is more efficient and less computationally demanding. Additionally, the model's capability in identifying and predicting single-bounce signals significantly enhances its utility for accurate localization and mapping applications. Future work will focus on expanding dataset diversity and exploring the integration of the model with IoT devices for broader applications in complex environments, potentially benefiting sectors like healthcare and emergency services. This progression promises to further refine indoor localization technologies, enhancing both accuracy and practical applicability.

7. References

[1] A. Poulose and D. S. Han, "Feature-based DEEP LSTM network for indoor localization using UWB measurements," Presented at ICAIIC, 2021. [Online]. Available: https://doi.org/10.1109/ICAIIC51459.2021.9415277 [Accessed: Jul. 18, 2024].

[2] J. B. Kristensen, M. Massanet Ginard, O. K. Jensen, and M. Shen, "Non-Line-of-sight identification for UWB indoor positioning systems using support vector machines," presented at IEEE MTT-S International Wireless Symposium (IWS), 2019. [Online]. Available: https://doi.org/10.1109/IEEE-IWS.2019.8804072. [Accessed: Jul. 18, 2024].

[3] A. L. Crețu-Sîrcu et al., "Evaluation and comparison of Ultrasonic and UWB technology for indoor localization in an industrial environment," Sensors, vol. 22, no. 8, p. 2927, 2022. [Online]. Available: https://doi.org/10.3390/s22082927 [Accessed: Jul. 18, 2024].

[4] Y. Luo and C. L. Law, “Indoor positioning using uwb-ir signals in the presence of dense multipath with path overlapping,” IEEE Transactions on Wireless Communications, vol. 11, no. 10, pp. 3734–3743, 2012. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6287526 [Accessed: Jul. 18, 2024].

[5] H. Zou, M. Jin, H. Jiang, L. Xie, and C. J. Spanos, “Winips: Wifi-based non-intrusive indoor positioning system with online radio map construction and adaptation,” IEEE Transactions on Wireless Communications, vol. 16, no. 12, pp. 8118–8130, 2017. [Online]. Available: https://ieeexplore.ieee.org/document/8057286/ [Accessed: Jul. 18, 2024].

[6] X.Yang, Z. Wu and QQ.Zhan g, “Bluetooth Indoor Localization with Gaussian–Bernoulli Restricted Boltzmann Machine Plus Liquid State Machine,” IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1–8, 2022. [Online]. Available: https://www.semanticscholar.org/paper/Bluetooth-Indoor-Localization-With-Restricted-Plus-Yang-Wu/bb88a28c758b73a825e68cfda713fde39020fb7a [Accessed: Jul. 18, 2024].

[7] R. Bharadwaj, C. Parini, and A. Alomainy, “Experimental investigation of 3-d human body localization using wearable ultra-wideband antennas,” IEEE Transactions on Antennas and Propagation, vol. 63, no. 11, pp. 5035–5044, 2015. [Online]. Available: https://ieeexplore.ieee.org/document/7265001 [Accessed: Jul. 18, 2024].

[8] A. R. Jimenez Ruiz and F. Seco Granja, “Comparing Ubisense, BeSpoon, and DecaWave UWB location systems: Indoor Performance Analysis,” IEEE Transactions on Instrumentation and Measurement, vol. 66, no. 8, pp. 2106–2117, 2017. [Online]. Available: https://doi.org/10.1109/tim.2017.2681398 [Accessed: Jul. 18, 2024].

[9] DW1000 User Manual, https://www.qorvo.com/products/d/da007967 [Accessed: Jul. 18, 2024].

[10] I. Guvenc, C. -C. Chong and F. Watanabe, "NLOS Identification and Mitigation for UWB Localization Systems," 2007 IEEE Wireless Communications and Networking Conference, Hong Kong, China, 2007, pp. 1571-1576, [Online]. Available: https://doi.org/10.1109/WCNC.2007.296 [Accessed: Jul. 18, 2024].

[11] B. Silva and G. P. Hancke, "IR-UWB-Based Non-Line-of-Sight Identification in Harsh Environments: Principles and Challenges," in IEEE Transactions on Industrial Informatics, vol. 12, no. 3, pp. 1188-1195, June 2016, [online]. Available: https://doi.org/10.1109/TII.2016.2554522 [Accessed: Jul. 18, 2024].

[12] H. Yang, Y. Wang, C. K. Seow, M. Sun, M. Si and L. Huang, "UWB Sensor-Based Indoor LOS/NLOS Localization with Support Vector Machine Learning," in IEEE Sensors Journal, vol. 23, no. 3, pp. 2988-3004, 1 Feb.1, 2023, [Online]. Available: https://doi.org/10.1109/JSEN.2022.3232479 [Accessed: Jul. 18, 2024].

[13] F. Wang, Z. Xu, R. Zhi, J. Chen, and P. Zhang, "LOS/NLOS Channel Identification Technology Based on CNN," 2019 6th NAFOSTED Conference on Information and Computer Science (NICS), Hanoi, Vietnam, 2019, pp. 200-203, [Online]. Available: https://doi.org/10.1109/NICS48868.2019.9023805 [Accessed: Jul. 18, 2024].

[14] M. Si, Y. Wang, H. Siljak, C. Seow and H. Yang, "A Lightweight CIR-Based CNN With MLP for NLOS/LOS Identification in a UWB Positioning System," in IEEE Communications Letters, vol. 27, no. 5, pp. 1332-1336, May 2023, [online]. Available: https://doi.org/10.1109/LCOMM.2023.3260953 [Accessed: Jul. 18, 2024].

[15] Zhao, Y.; Wang, M. The LOS/NLOS Classification Method Based on Deep Learning for the UWB Localization System in Coal Mines. Appl. Sci. 2022, 12, 6484. [Online]. Available: https://doi.org/10.3390/app12136484 [Accessed: Jul. 18, 2024].

[16] M. Stahlke, S. Kram, C. Mutschler and T. Mahr, "NLOS Detection using UWB Channel Impulse Responses and Convolutional Neural Networks," 2020 International Conference on Localization and GNSS (ICL-GNSS), Tampere, Finland, 2020, pp. 1-6, [Online]. Available: https://doi.org/10.1109/ICL-GNSS49876.2020.9115498 [Accessed: Jul. 18, 2024].

[17] A Vaswani, N Shazeer, N Parmar, “Attention Is All You Need” [Online]. Available: https://doi.org/10.48550/arXiv.1706.03762 [Accessed: Jul. 18, 2024].

[18] S. Tomović, K. Bregar, T. Javornik, and I. Radusinović, "Transformer-based NLOS detection in UWB localization systems," in *2022 30th Telecommunications Forum (TELFOR)*, Belgrade, Serbia, Nov. 2022, pp. 1-4. [Online]. Available: https://ieeexplore.ieee.org/document/9983765. [Accessed: Jul. 18, 2024].

[19] Luo, R.; Yan, L.; Deng, P.; Kuang, Y. Hybrid TOA/AOA Virtual Station Localization Based on Scattering Signal Identification for GNSS-Denied Urban or Indoor NLOS Environments. *Appl. Sci.* **2022**, *12*, 12157. [Online]. Available: https://doi.org/10.3390/app122312157 [Accessed: Jul. 18, 2024].

[20] Fares, M.H. , Moradi, H. , Shahabadi, M. and Mohanna, Y. 2021. A Beacon-based Approach for RF Source localization in Outdoor NLOS Environment for Search and Rescue Missions . Trends in Sciences. 18, 23 (Nov. 2021), 685. [Online]. Available: https://doi.org/10.48048/tis.2021.685 [Accessed: Jul. 18, 2024].

[21] J. C. S. Tai, S. Y. Tan, and C. K. Seow, "Three-dimensional non-line-of-sight localisation in an indoor multipath environment," in *2009 7th International Conference on Information, Communications and Signal Processing (ICICS)*, Macau, China, 2009, pp. 1-5. [Online]. Available: https://ieeexplore.ieee.org/document/5397607. [Accessed: Jul. 18, 2024].

[22] S.W. Chen, C. K. Seow, Wen, Kai "Concept of image based Non-Line-of-Sight (NLOS) localization in multipath environments," NTU, 2009. [Online]. Available: https://dr.ntu.edu.sg/bitstream/10356/79493/1/Concept%20of%20image%20based%20Non-Line-of-Sight%20(NLOS)%20localization%20in%20multipath%20environments.pdf. [Accessed: Jul. 18, 2024].

[23] Chen, S. W., Seow, C. K. and Wen, K "Peer-to-peer non-line-of-sight localization in multipath environment," NTU, 2009. [Online]. Available: https://dr.ntu.edu.sg/bitstream/10356/104955/1/Peer-to-peer%20Non-line-of-sight%20localization%20in%20multipath%20environment.pdf. [Accessed: Jul. 18, 2024].

[24] H. Raharjo, S. Y. Tan, and C. K. Seow, "Non-Line-Of-Sight localization scheme using two-steps weighting process," in *2009 7th International Conference on Information, Communications and Signal Processing (ICICS)*, Macau, China, 2009, pp. 1-5. [Online]. Available: https://ieeexplore.ieee.org/document/5397607. [Accessed: Jul. 18, 2024].

[25]FiRa Consortium, "How UWB Works," FiRa Consortium, [online]. Available: https://www.firaconsortium.org/discover/how-uwb-works. [Accessed: Jul. 18, 2024].

[26] J Devlin, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”. [Online]. Available: https://doi.org/10.48550/arXiv.1810.04805 [Accessed: Jul. 18, 2024].

[27] H. Candemir, "Advancing natural language processing (NLP) applications of morphologically rich languages with bidirectional encoder representations from transformers (BERT): an empirical case study for Turkish," ResearchGate, 2021. [Online]. Available: https://www.researchgate.net/publication/351386823_Advancing_natural_language_processing_NLP_applications_of_morphologically_rich_languages_with_bidirectional_encoder_representations_from_transformers_BERT_an_empirical_case_study_for_Turkish. [Accessed: Jul. 18, 2024].

[28] OpenAI, “GPT-4 Technical Report” [Online]. Available: https://doi.org/10.48550/arXiv.2303.08774 [Accessed: Jul. 18, 2024].

8. Knowledge and Training Requirements

The section list all the knowledge, skillsets, and certifications both from the degree programme and beyond that was necessary for the successful completion of the capstone projects.

8.1. Applicable Knowledge from the Degree Programme

The prerequisite knowledge and skillsets from the degree programme that was necessary for the capstone projects are as follows:

No.	Module(s)	Knowledge(s) Applied
1	CSC2003: Embedded System Programming	Programming the DWM1001 device was a new challenge for me, but my basic understanding of embedded systems helped me navigate this endeavour. Working with this module deepened my knowledge of IoT devices, and I developed grit and determination, which enabled me to complete the data collection aspect of the project.
2	CSC3005: Data Analytics	This module provided us with fundamental data processing skills, which I applied effectively. It also exposed me to traditional machine-learning techniques.
3	CSC3009: Machine learning	This module equipped me with fundamental machine-learning skills, which I applied effectively. Additionally, it introduced me to deep learning and advanced machine learning techniques.

8.2. Additional Knowledge, Skillsets, or Certifications Required

The following are the additional requirements and knowledge that were required for the capstone projects:

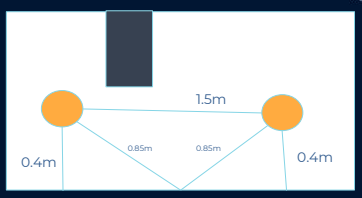
No.	Additional Requirement(s)	Knowledge(s) Applied
1	Deep Learning and Transformers	I leveraged my self-taught knowledge of, Transformers, and BERT architecture characteristics to develop a novel algorithm.
2	Signal Processing	I acquired foundational knowledge in signal processing, enabling me to identify non-line-of-sight (NLOS) and line-of-sight (LOS) conditions, as well as single-bounce characteristics from the collected signal data. Additionally, I learned to recognize outliers and effectively remove them from the dataset.

9. Appendix

```
#define ACCUM_DATA_LEN 1016


for (int i=0;i<ACCUM_DATA_LEN;i++){
  uint8 *cir;
  cir = (uint8 *)malloc(5*sizeof(uint8));
  dwt_readaccdata(cir,5, 4*i);
  int16 real = 0;
  int16 imag = 0;
   int16 amp = 0;

  real = (int16) (cir[2]<<8) | (int16)(cir[1]);
  imag = (int16) (cir[4]<<8) | (int16)(cir[3]);
  amp = MAX(abs(real),abs(imag)) + 1/4*MIN(abs(real),abs(imag));
  printf("%d,", amp);

  free(cir);

}
```

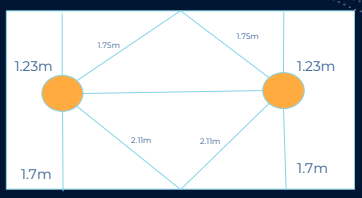
Figure 18: Code for Reading the CIR Values from The Accumulator

Experimental Setup LOS 1.5m



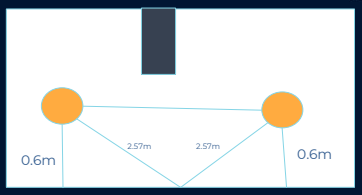
1 bounce distance = 1.7 m

Experimental Setup LOS 2.5m



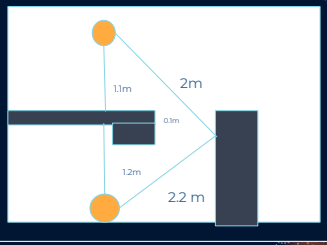
1 bounce distance = 4.22 m and 3.5m

Experimental Setup LOS 5m



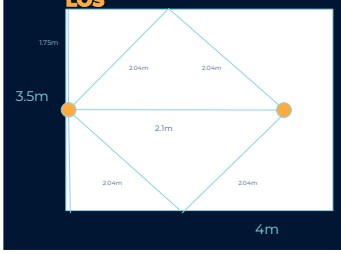
1 bounce distance = 5.14 m

Experimental Setup Plaster 2.4 m



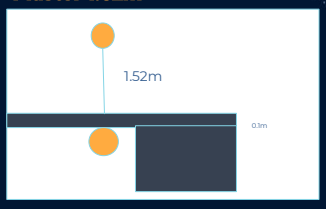
1 bounce distance = 4.2 m

Experimental Setup SIT DR LOS

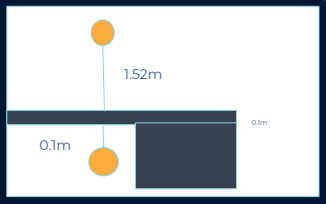


1 bounce distance = 4.08 m

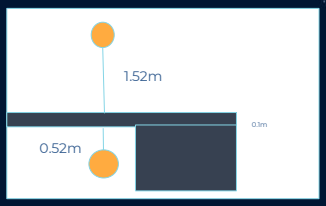
Experimental Setup Plaster 1.62m



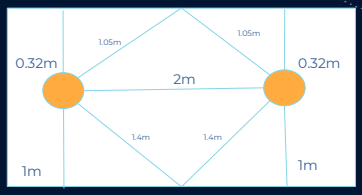
Experimental Setup Plaster 1.72m



Experimental Setup Plaster 2.14 m

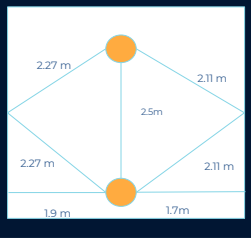


Experimental Setup LOS 2m



1 bounce distance = 2.1 m and 2.8 m

Experimental Setup LOS 2.5m Living Room



1 bounce distance = 4.22 m And 4.54m

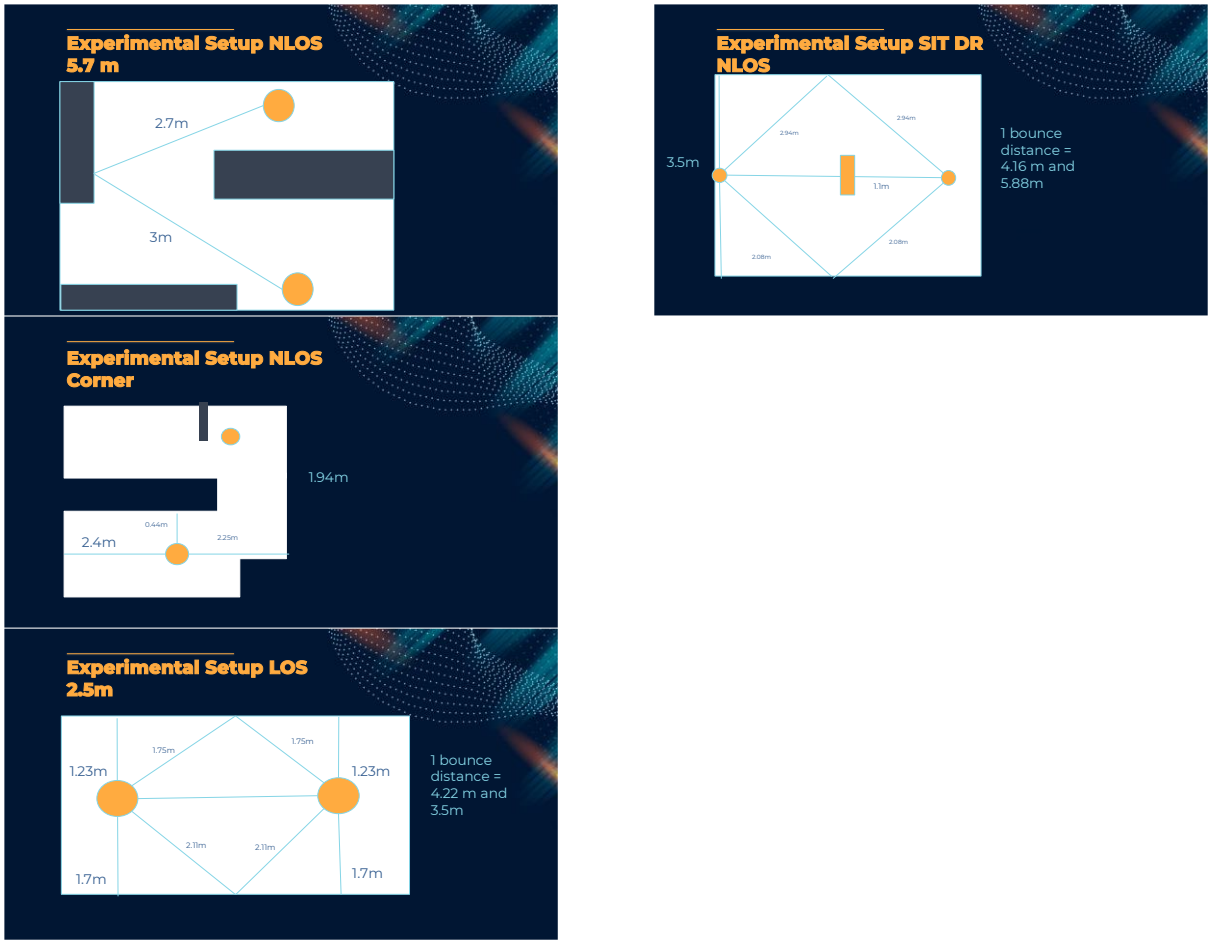


Figure 19: Environmental Setups for Data Collection

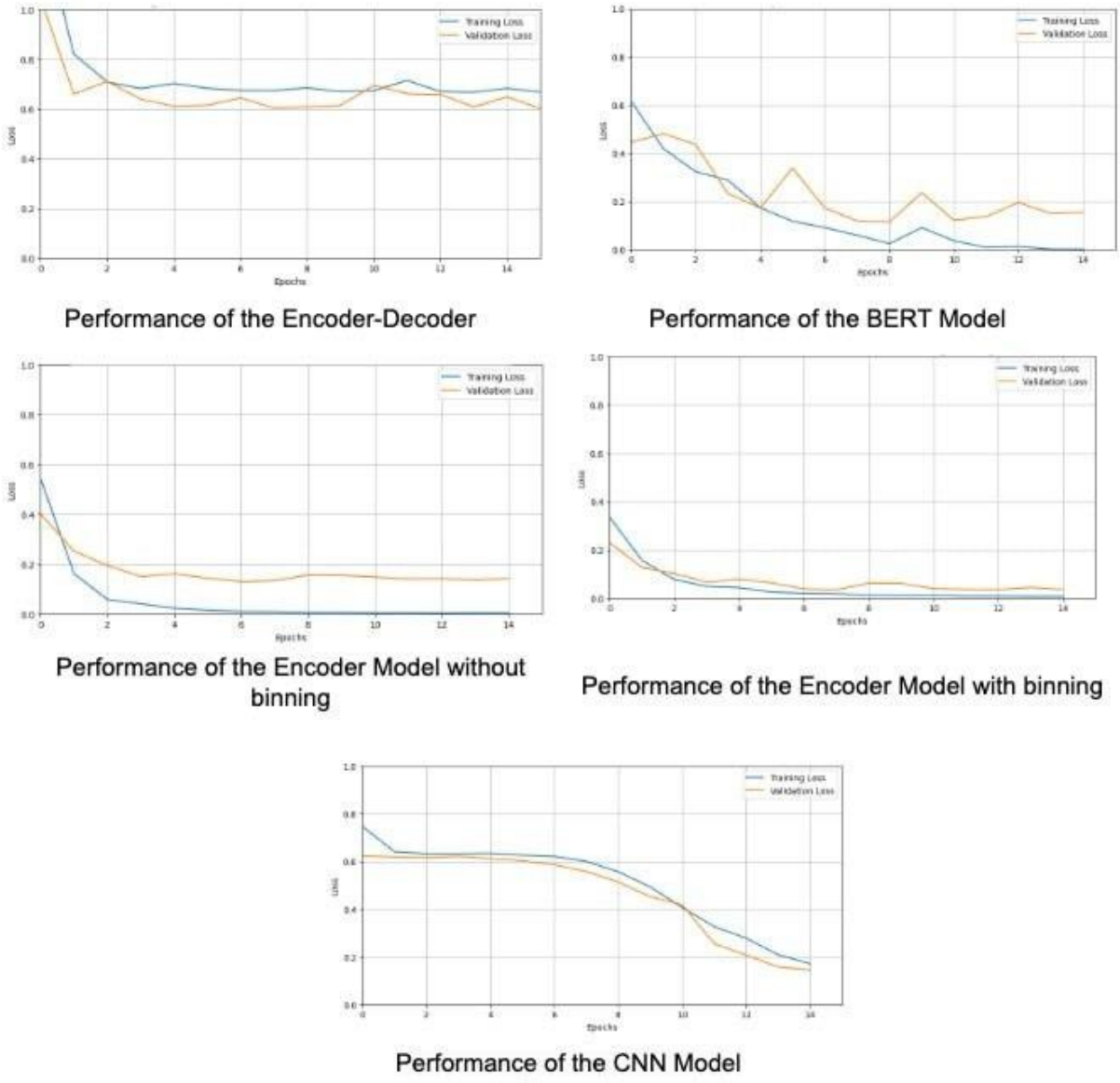


Figure 20: Combined Loss Graphs - Identifying NLOS and LOS Conditions

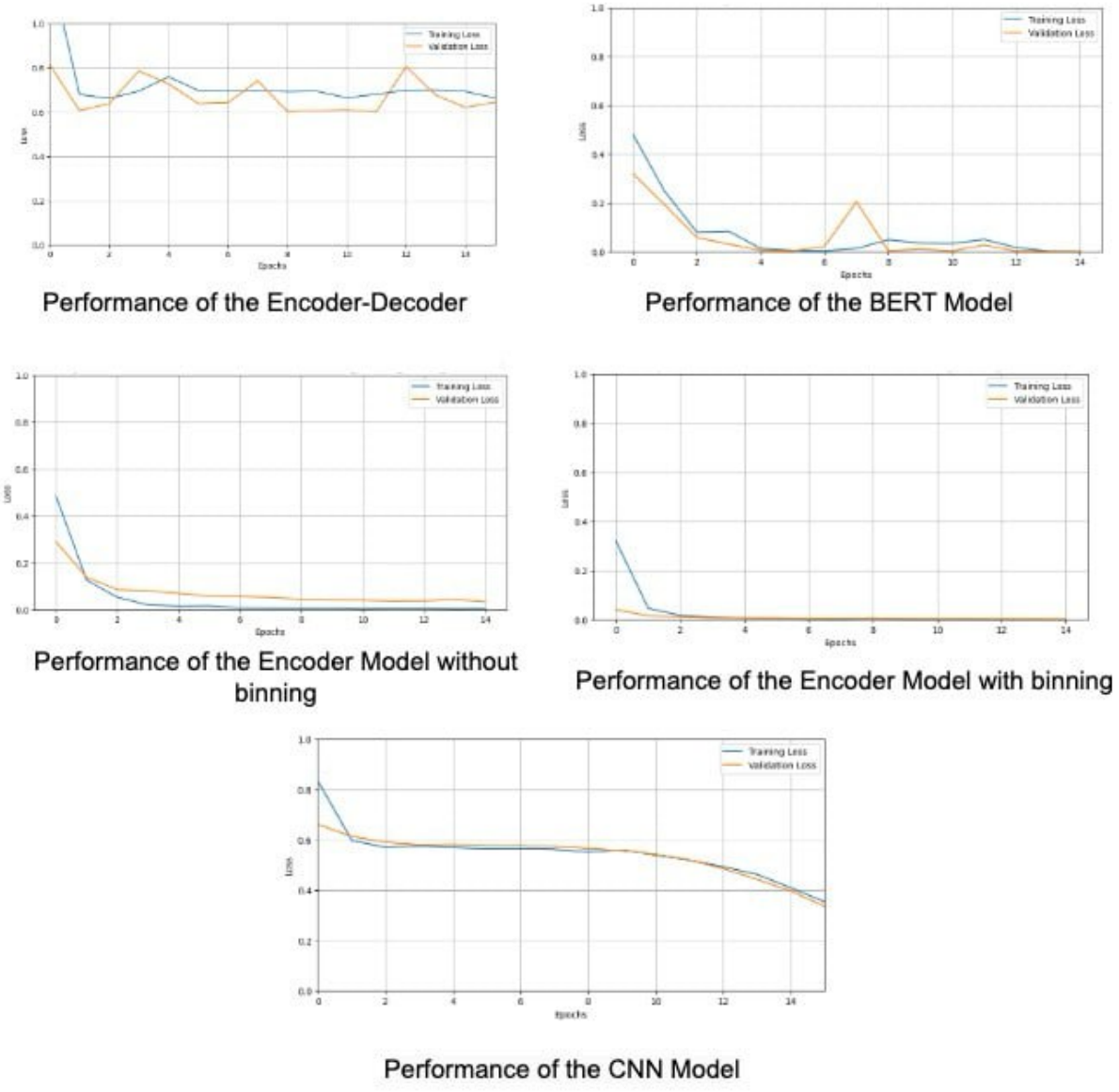


Figure 21: Combined Loss Graphs - Identifying Single-Bounce Signal Conditions

	precision	recall	f1-score	support
0	0.99	1.00	1.00	125
1	1.00	0.98	0.99	66
accuracy			0.99	191
macro avg	1.00	0.99	0.99	191
weighted avg	0.99	0.99	0.99	191

Encoder Model with binning

	precision	recall	f1-score	support
0	0.98	0.98	0.98	125
1	0.97	0.95	0.96	66
accuracy			0.97	191
macro avg	0.97	0.97	0.97	191
weighted avg	0.97	0.97	0.97	191

Encoder Model without binning

	precision	recall	f1-score	support
0	0.98	0.97	0.98	125
1	0.94	0.97	0.96	66
accuracy			0.97	191
macro avg	0.96	0.97	0.97	191
weighted avg	0.97	0.97	0.97	191

BERT

	precision	recall	f1-score	support
0	0.94	0.98	0.96	124
1	0.95	0.88	0.91	67
accuracy			0.94	191
macro avg	0.94	0.93	0.94	191
weighted avg	0.94	0.94	0.94	191

CNN

	precision	recall	f1-score	support
0	0.70	1.00	0.82	133
1	0.00	0.00	0.00	58
accuracy			0.70	191
macro avg	0.35	0.50	0.41	191
weighted avg	0.48	0.70	0.57	191

Encoder-Decoder

Figure 22: Classification Reports for Identifying NLOS Conditions

	precision	recall	f1-score	support
0	1.00	1.00	1.00	14
1	1.00	1.00	1.00	49
accuracy			1.00	63
macro avg	1.00	1.00	1.00	63
weighted avg	1.00	1.00	1.00	63

Encoder Model with binning

	precision	recall	f1-score	support
0	1.00	0.93	0.96	14
1	0.98	1.00	0.99	49
accuracy			0.98	63
macro avg	0.99	0.96	0.98	63
weighted avg	0.98	0.98	0.98	63

Encoder Model without binning

	precision	recall	f1-score	support
Class 0	1.00	1.00	1.00	14
Class 1	1.00	1.00	1.00	49
accuracy			1.00	63
macro avg	1.00	1.00	1.00	63
weighted avg	1.00	1.00	1.00	63

BERT

	precision	recall	f1-score	support
0	1.00	0.86	0.92	14
1	0.96	1.00	0.98	49
accuracy			0.97	63
macro avg	0.98	0.93	0.95	63
weighted avg	0.97	0.97	0.97	63

CNN

	precision	recall	f1-score	support
0	0.00	0.00	0.00	16
1	0.75	1.00	0.85	47
accuracy			0.75	63
macro avg	0.37	0.50	0.43	63
weighted avg	0.56	0.75	0.64	63

Encoder-Decoder

Figure 23: Classification Reports for Identifying Single Bounce Conditions

END OF REPORT