# Detection and Estimation Theory Project Presentation

-Shannon Muthanna I B (IMT2022552)
- Margasahayam Venkatesh Chirag (IMT2022583)

# Introduction

- $H_0$- Null hypothesis, $H_1$- Alternative Hypothesis. NP rule: $\Lambda(\mathbf{z}) = \frac{f(\mathbf{z}|H_1)}{f(\mathbf{z}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta(P_{\mathrm{FA}})$
- For the NP detector to be implemented, both likelihood functions must be known and so, statistical models are assumed and parameters are estimated. Losses are calculated when this statistical model varies from the design.
- For the proposed approach, no statistical model is to assumed and a good approximation for the optimal NP detector is obtained from choosing a relevant error function while training the ML model.
- Drawback of the approach is the difficulty of obtaining training samples, and defining the most suitable learning machine architecture.

# Previous Works

- Basic NP Detector Approximation: Use a **single-output learning machine** compared to a **threshold** to decide between null ($H_0$) and alternative ($H_1$) hypotheses. The threshold sets the desired **Probability of False Alarm ($P_{FA}$)**.
- Alternative Implementations: Varies the **bias** of a single output neuron. Utilizes a **two-output NN** (outputs in (0,1)), comparing their subtraction to a threshold, which is equivalent to a single output with desired outputs {−1,1}.
- Advanced Techniques: **Radial Basis Function Neural Networks (RBFNN)**, **Support Vector Machines (SVMs)**, and **Committee Machines** have also been applied for detection.
- Supervised Training: Adaptive systems (learning machines) can approximate the **Neyman-Pearson (NP) detector** when trained in a supervised manner.
- Suitable Error Functions: The **Sum-of-Squares error** is suitable for approximating the NP detector (and optimal for **Gaussian signals**). The **Minkowski error** with R=1 (Mean Absolute Deviation) is suitable for approximating the **minimum probability of error classifier**.
- Future Work: Investigating **alternative error functions** suitable for non-Gaussian interference, ensuring they meet the sufficient condition for NP detector approximation.

# Problem Statement

- Consider a learning machine with one output, that is used to classify input vectors $z = [z_1, z_2, ..., z_L]^T$ into two hypotheses $H_0$ and $H_1$, that represent absence of target and its presence, respectively, in radar detection problems.
- Given a decision rule, let $Z_i$ be the set of all possible input vectors that will be assigned to hypothesis $H_i$,and $Z_0 \cup Z_1 = Z$ the ensemble of all possible input vectors. The output of the learning machine is represented by $F(z)$, and the desired output by $t_{Hi}$ . A training set, $Z = Z_0 \cup Z_1$,where $Z_1$ is composed of $N_1$ training patterns from hypothesis $H_1$,and $Z_0$ is composed of $N_0$ training patterns from hypothesis $H_0$ ($N = N_1 + N_0$), is available.
- The function the learning machine approximates to after training is obtained (as a function of the likelihood functions and the prior probabilities) and the implemented detector compares the learning machine output to a threshold $\eta_0$, which varies to fix the $P_{FA}$. The NP detector is usually implemented by comparing the likelihood ratio to a threshold $\eta_{lr}$, fixed according to the required $P_{FA}$.

# Proof of Concept

This section explains how the MLP's output can be treated as an approximation of the NP detector

## Law of Large Numbers

The law of large numbers states that given a sample of independent and identically distributed values, the sample mean converges to the true mean.

- Equation three here denotes the cross entropy loss function at an MLP's output.
- For N tends to infinity, the sample mean of the function becomes true expected value of the function as shown in equation 5.

$$E = -\frac{1}{N}\left[\sum_{\mathbf{z}\in H_1}\ln[F(\mathbf{z})] + \sum_{\mathbf{z}\in H_0}\ln[1-F(\mathbf{z})]\right] \quad (3)$$

If the number of patterns tends to infinity ($N \to \infty$), the error can be expressed as follows:

$$E_m = -\lim_{N\to\infty}\left[\frac{N_1}{N}\frac{1}{N_1}\sum_{\mathbf{z}\in H_1}\ln[F(\mathbf{z})] + \frac{N_0}{N}\frac{1}{N_0}\sum_{\mathbf{z}\in H_0}\ln[1-F(\mathbf{z})]\right] \quad (4)$$

Applying the strong law of large numbers, expression (5) is obtained:

$$E_m = -\int_{\mathcal{Z}} \big(P(H_1)f(\mathbf{z}|H_1)\ln(F(\mathbf{z})) + P(H_0)f(\mathbf{z}|H_0)\ln(1-F(\mathbf{z}))\big)\,d\mathbf{z} \quad (5)$$

# Proof of Concept

By Calculus of Variations, a function Fo(z) which minimizes the function/loss Em can be written using Euler-Lagrange:

$$\frac{\partial I}{\partial F} - \sum_{k=1}^{L} \frac{\partial}{\partial z_k}\left(\frac{\partial I}{\partial F'_k}\right) = 0$$

Here the function I is the function inside the integral of the loss Em, Solving further we get Fo(z) as follows:

$$\frac{\partial}{\partial F}\Big(-P(H_1)f(\mathbf{z}|H_1)\ln(F(\mathbf{z})) - P(H_0)f(\mathbf{z}|H_0)\ln(1 - F(\mathbf{z}))\Big) = 0 \tag{8}$$

$$F_0(\mathbf{z}) = \frac{P(H_1)f(\mathbf{z}|H_1)}{P(H_1)f(\mathbf{z}|H_1) + P(H_0)f(\mathbf{z}|H_0)}$$

$$\frac{P(H_1)f(\mathbf{z}|H_1)}{P(H_1)f(\mathbf{z}|H_1) + P(H_0)f(\mathbf{z}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_0$$

# Proof of Concept

Extracting the Likelihood ratio from the above Fo(z) equation we get our new NP detector threshold as given below:

$$\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)\eta_0}{P(H_1)(1-\eta_0)} = \eta_{lr}$$

$$\eta_0 = \frac{\eta_{lr}P(H_1)}{P(H_0) + \eta_{lr}P(H_1)}$$

# Experiments

**Objective:** Design **Multilayer Perceptrons (MLPs)** to approximate the performance of the **optimum NP detector**.

**Challenge:** Target parameters are **unknown**; therefore, the NP detector is approximated using an **Average-Likelihood Ratio** compared against a threshold.

**MLP Structure:** Real arithmetic MLP with **L/M/1** architecture.

- **Input Layer (L=16):** 8 complex radar echoes → 16 real inputs (in-phase and quadrature).
- **Hidden Layer (M):** Number of neurons (M) is a studied parameter.
- **Output Layer (1):** Compared to a hard threshold for detection.
- **Activation Function:** Sigmoid.

# Experiments

**Three Case Studies of Detection Problems:** The studies assess the MLP's ability to approximate the NP optimum detector, focusing on c

| Case Study | Target Type & Interference | Sub-Cases (Target Unknowns) |
|---|---|---|
| **Case 1** | Gaussian Fluctuating Target in **AWGN** (Clear Conditions) | Unknown **correlation coefficient** OR Unknown **Doppler shift** (Swerling I). |
| **Case 2** | Gaussian Fluctuating Target in **Correlated Gaussian Clutter + AWGN** (Sea/Land Clutter, Low Resolution) | Unknown **correlation coefficient** OR Unknown **Doppler shift** (Swerling I). |
| **Case 3** | **Non-Fluctuating** (Swerling V) Target in **K-distributed Clutter** ($\nu = 0.5$) (Spiky Clutter, High Resolution, Low Grazing Angle) | Unknown **Doppler shift**. |

# Experiments

### Performance Comparison Strategies

The proposed MLP detectors are benchmarked using two primary strategies:

1. **MLP vs. NP Approximation:**
    - MLP detectors (trained via supervised learning) are compared against the optimum **NP Detector approximation**.
    - The NP approximation uses a **Constrained Generalized Likelihood Ratio (CGLR)**.
    - **Metric:** Receiver Operating Characteristic (**ROC** curves) are used to show validity (Probability of Detection ($P_D$) vs. Probability of False Alarm ($P_{FA}$)).
2. **MLP Error Function Comparison:**
    - Compare detectors trained with **Cross-Entropy Error** against those trained with:
        - **Sum-of-Squares Error**
        - **Minkowski Error** (R=1)
    - **Note:** This comparison is performed only for **Case Study 1** due to space limitations, but similar trends were observed for the other cases.
    - **Metric:** PD versus Signal-to-Noise Ratio (SNR) for the best detectors.

# Experiments

### II. Training and Testing Parameters

| Parameter | Value / Method | Details |
|---|---|---|
| **Training Set** | 50,000 patterns | Equal patterns for $H_0$ (No Target) and $H_1$ (Target). |
| **Validation** | 10,000 patterns | Used for **cross-validation** ($k = 5$) to prevent overfitting. |
| **Testing Set** | $2 \times 10^7$ ($H_0$) and $5 \times 10^4$ ($H_1$) patterns | Allows accurate estimation of very low $P_{FA}$ values (down to $10^{-6}$) and $P_D$. |
| **Training Algorithm** | **Cross-Entropy:** Algorithm from [43] | **Sum-of-Squares/Minkowski:** Conjugate Gradient method [44]. |

# Experiments

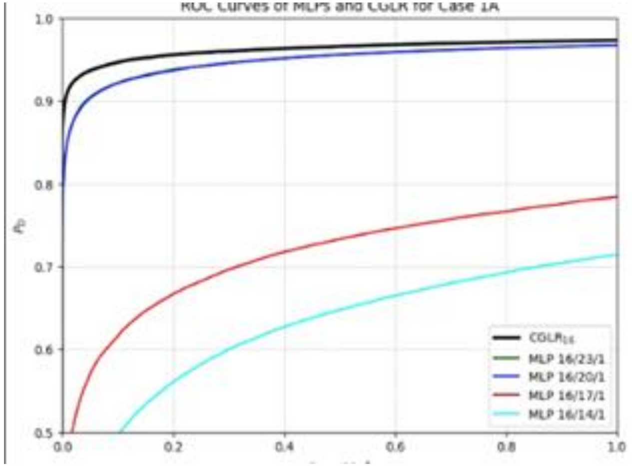
- The final approximation error depends on the **error function**, **training/validation sets**, system **structure**, and **training algorithm**.
- For a good approximation, the training set must be a subset of the input, the predicted function (implemented by the learning machine) must be **sufficiently general** to minimize the error function and the **learning algorithm** must effectively find the **minimum** of the error function.

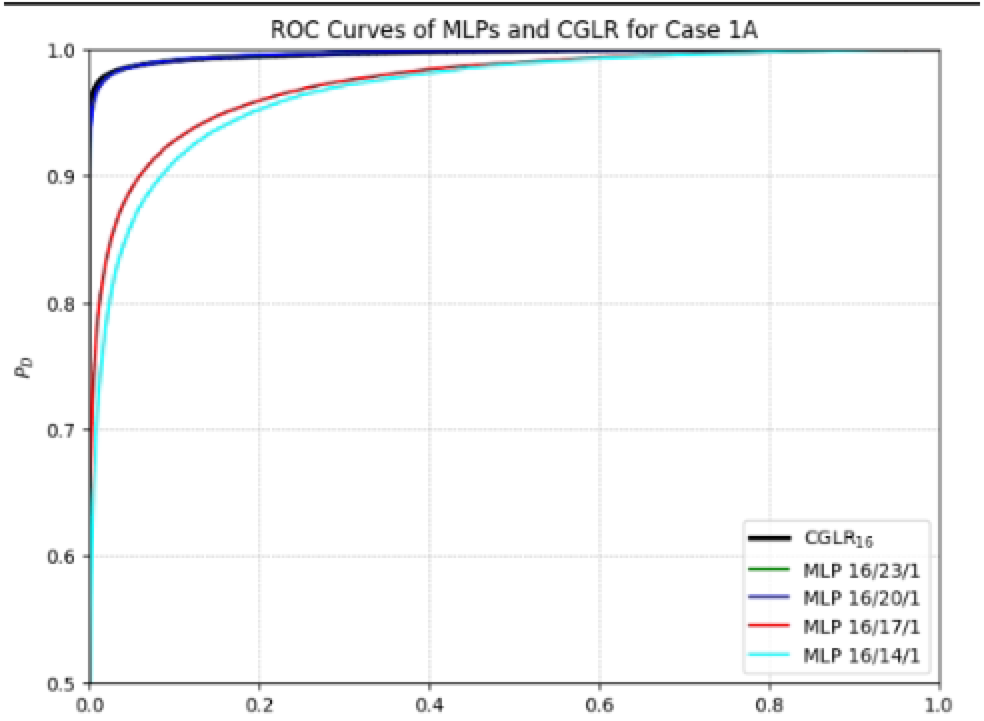
# Results: Case 1A

**ROC curve for Gaussian Targets with**

**Cross-Entropy error**

Zoomed version given below

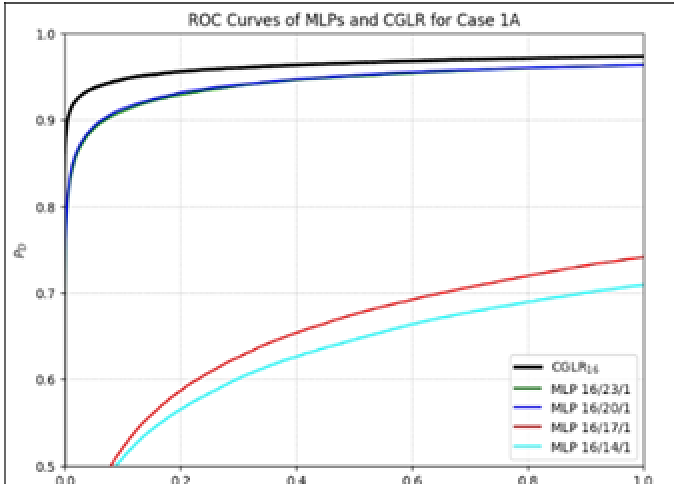


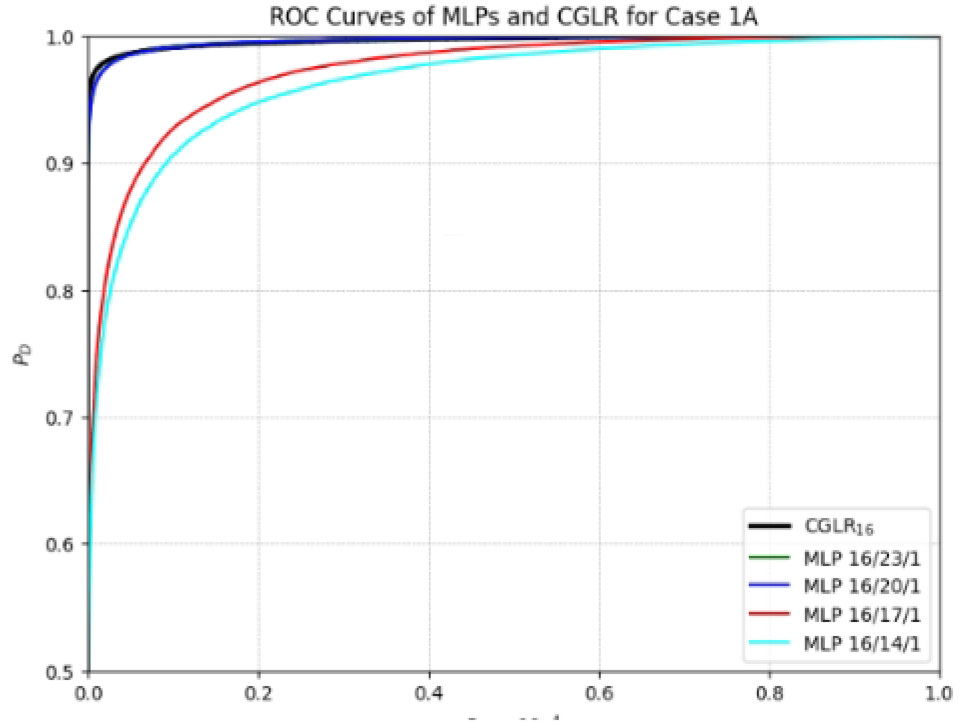



# Results: Case 1A

**ROC curve for Gaussian Targets with**

**Sum of Squared Error**

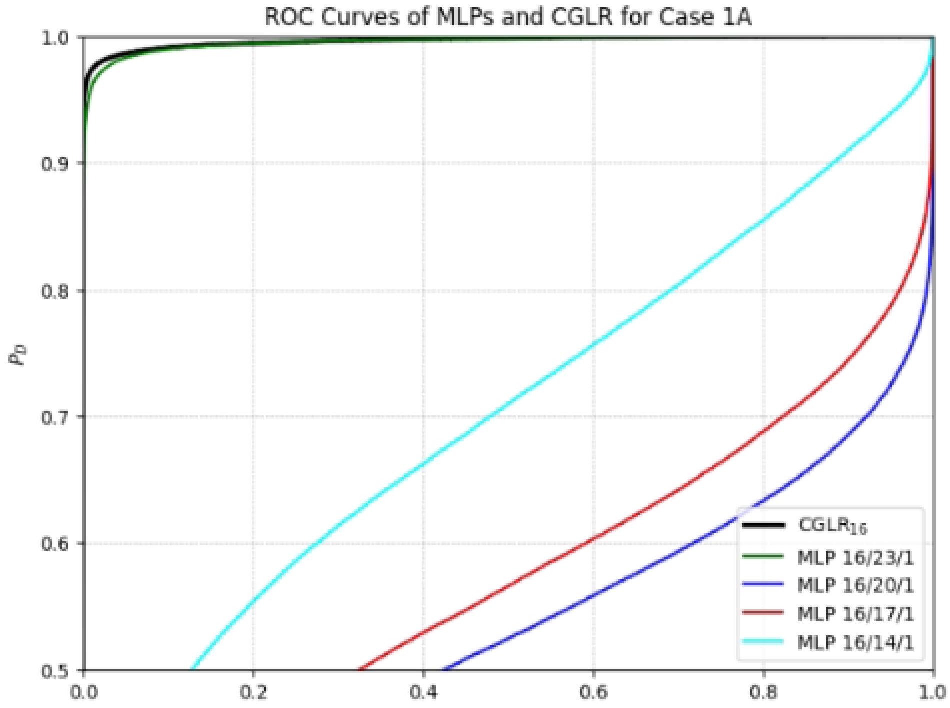


# Results: Case 1A

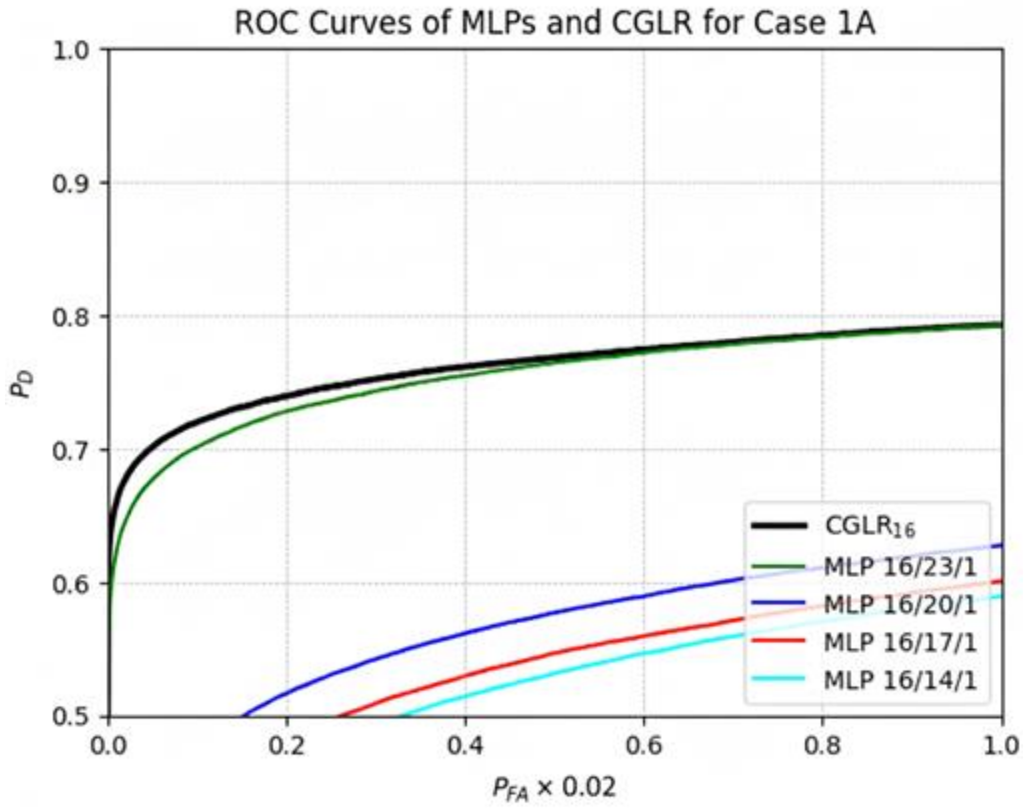
**ROC curve for Gaussian Targets with**

**Minkowski error R = 1**



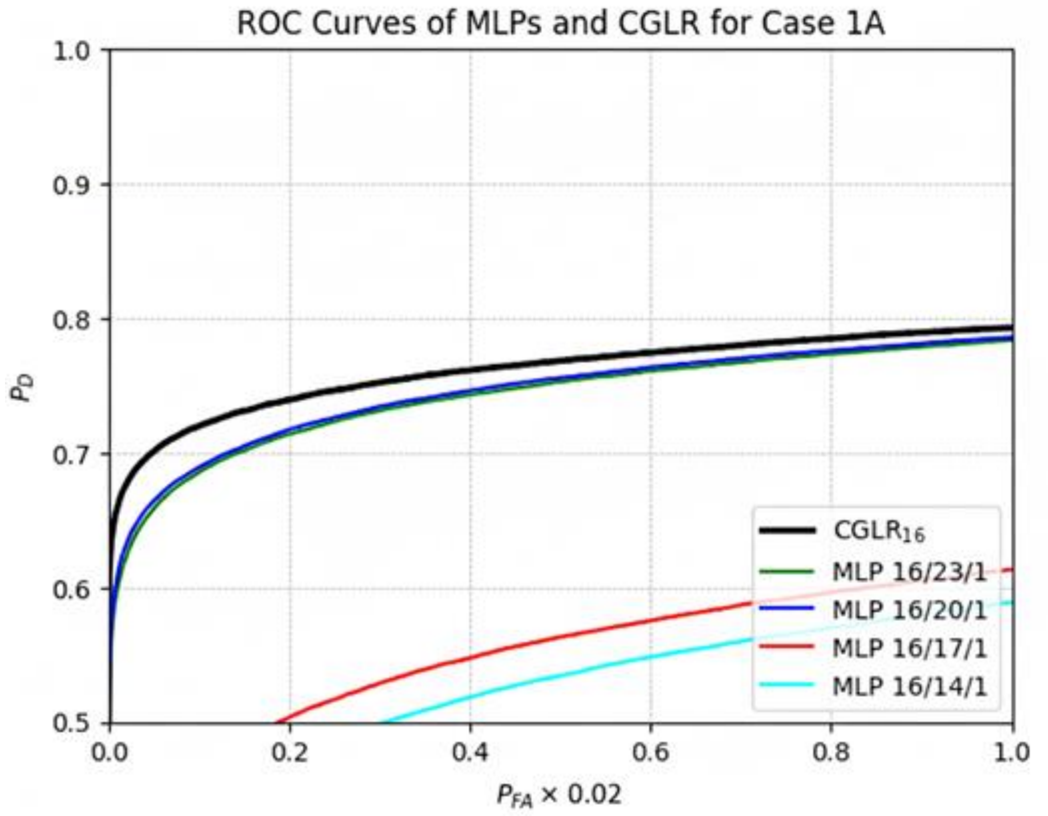
# Case 1B: ROC curve for SW1 with Cross-Entropy Error

Local Trend (Zoomed Version)



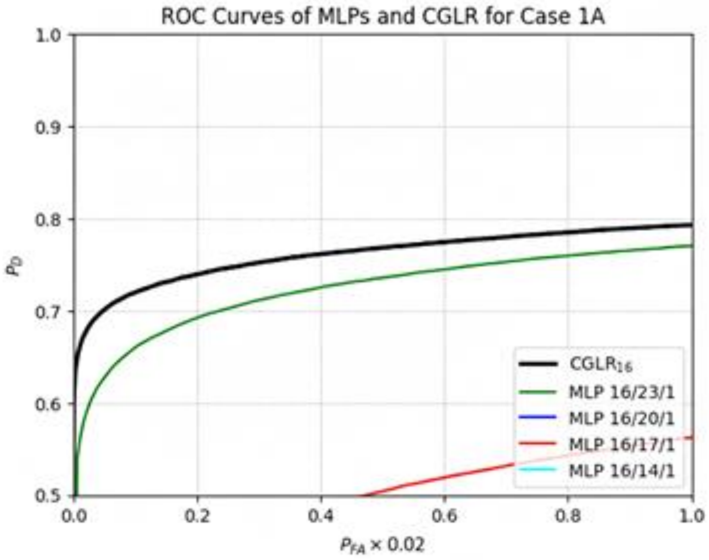
# Case 1B: ROC curve for SW1 with Sum-of-Squares Error

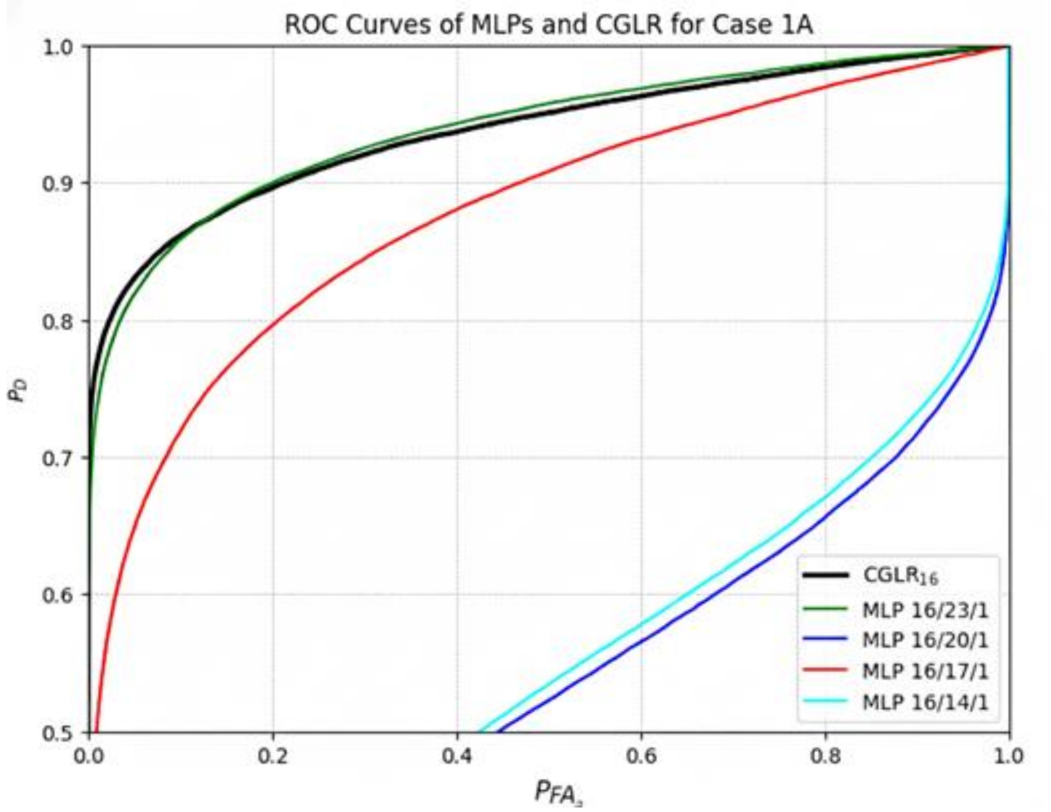
Local Trend (zoomed version)



# Case 1B: ROC curve for SW1 with Minkowski Error (for R=1)

Local Trend (zoomed version)



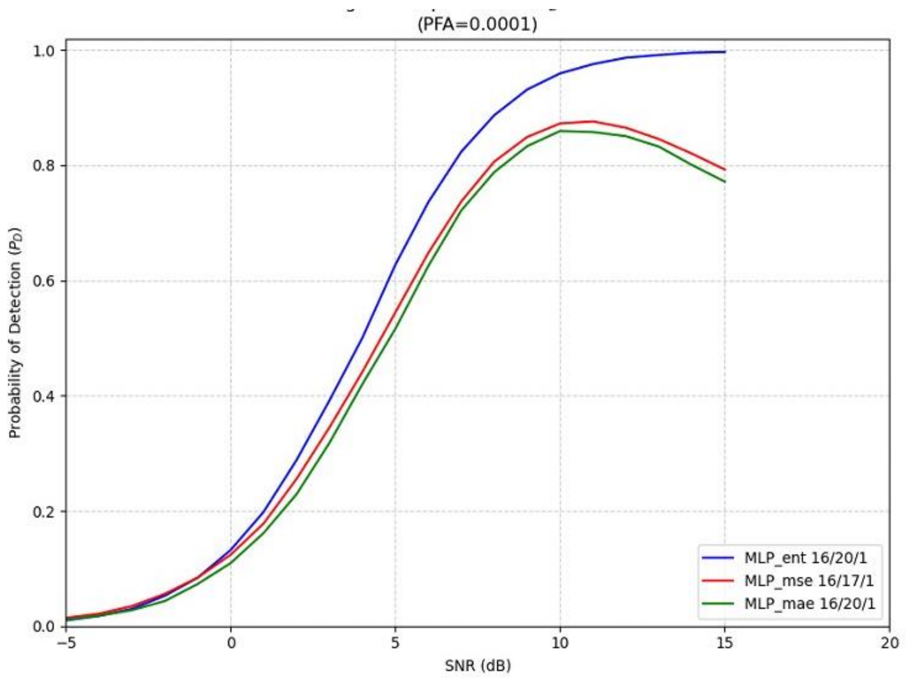
# Future Works

- Generate the ROC plots for Case-2 and Case-3 for A and B subcases.
- Plot $P_D$ VS SNR for all the 3 types of error
- Finding an innovation for this proposal.

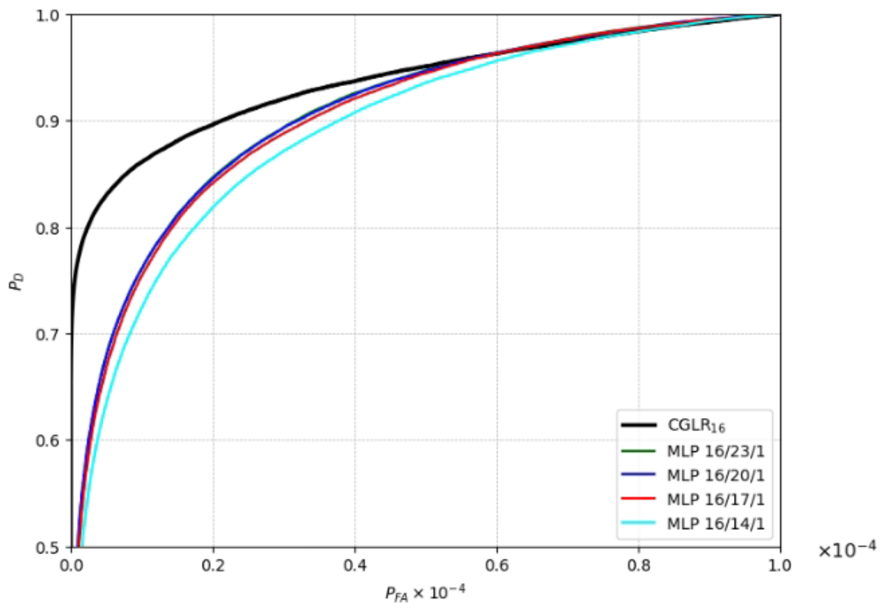
# TILL MIDTERM

# POST-MIDTERM PROGRESS

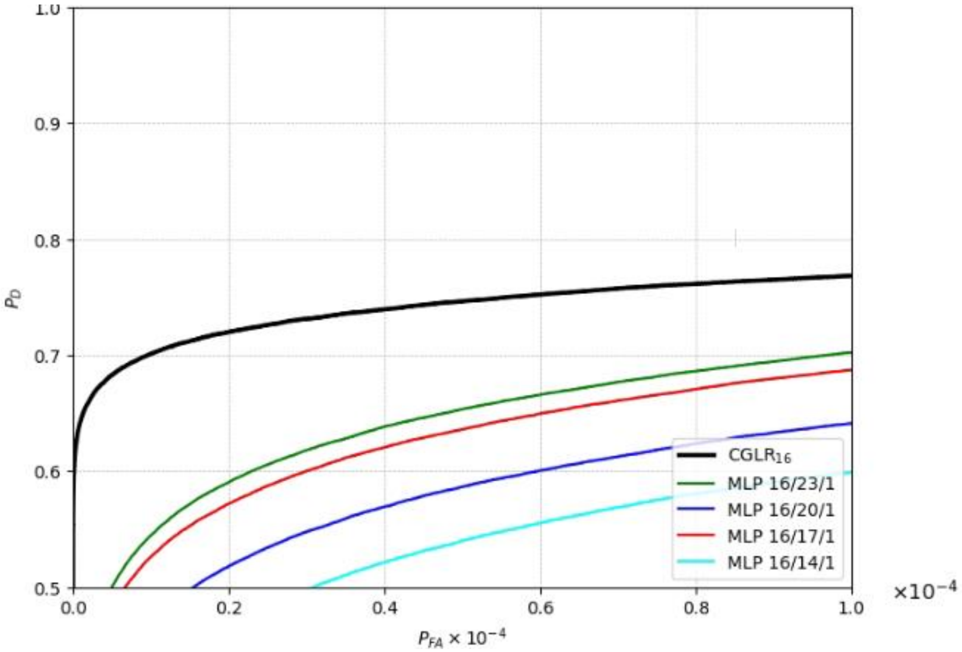
# Variation of Pd vs SNR



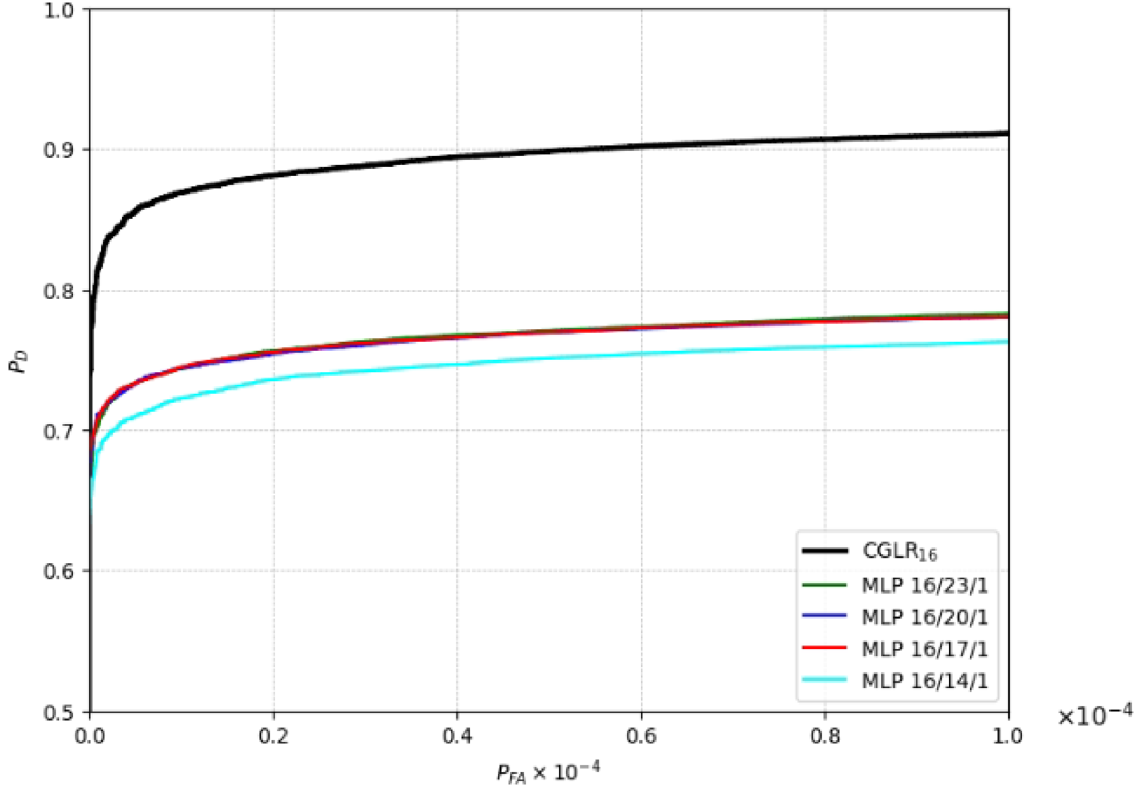
Detection of Gaussian fluctuating targets in presence of correlated Gaussian clutter and AWGN (CNR = 20 dB, ρ c = 0.7, SIR = 0 dB)



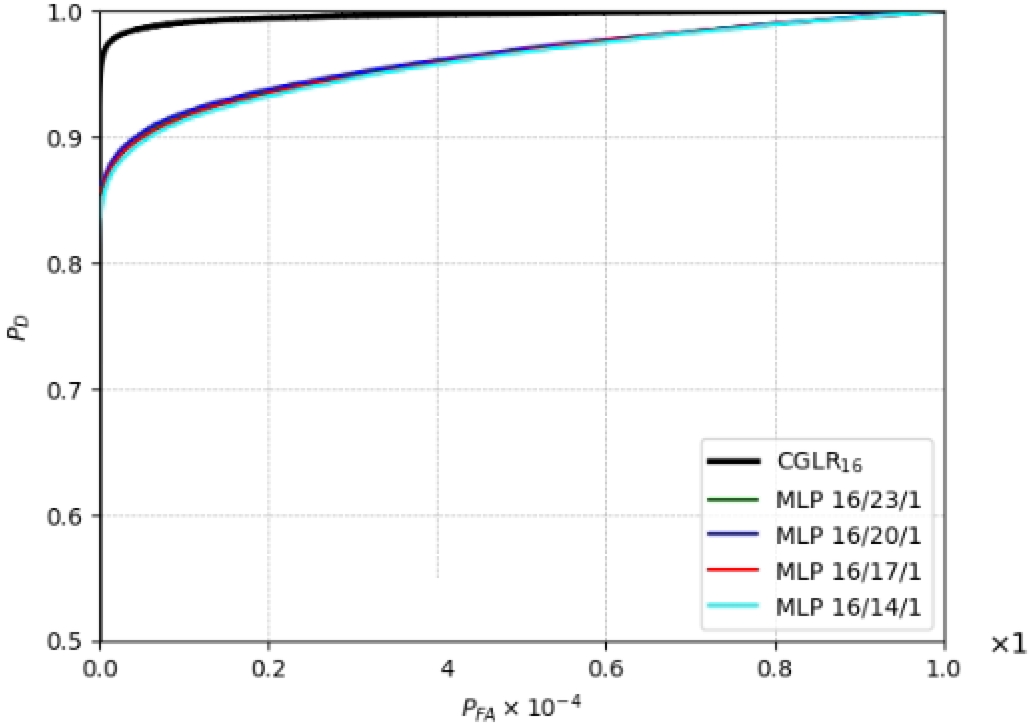
Detection of Gaussian fluctuating targets in presence of correlated Gaussian clutter and AWGN (CNR = 20 dB, ρ c = 0.995, SIR = −10 dB)



Detection of Gaussian fluctuating SW targets in presence of correlated Gaussian clutter and AWGN (CNR = 20 dB, ρ c = 0.7, SIR = 13 dB)



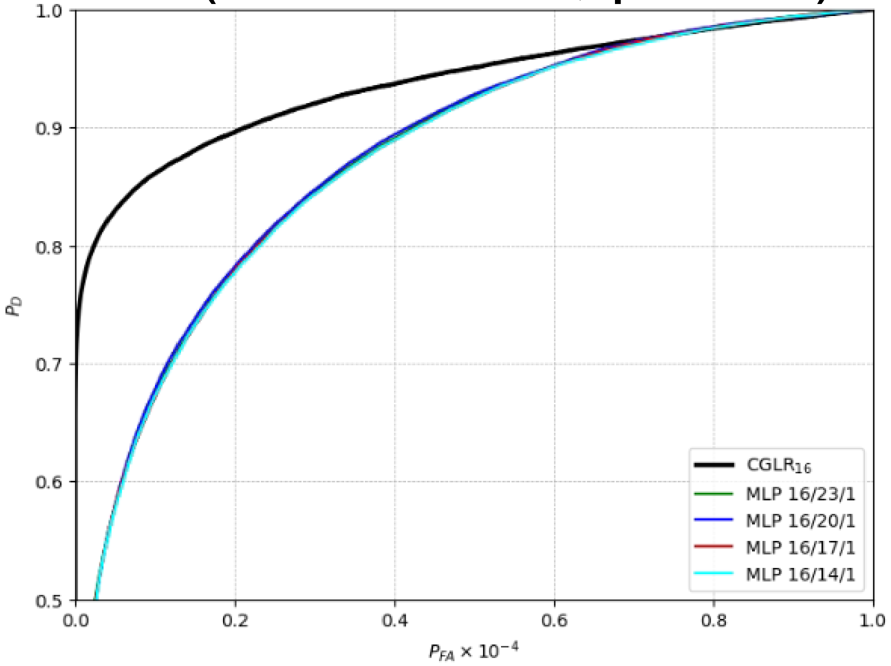
Detection of Gaussian fluctuating SW targets in presence of correlated Gaussian clutter and AWGN (CNR = 20 dB, ρ c = 0.995,
SIR = 1 dB)



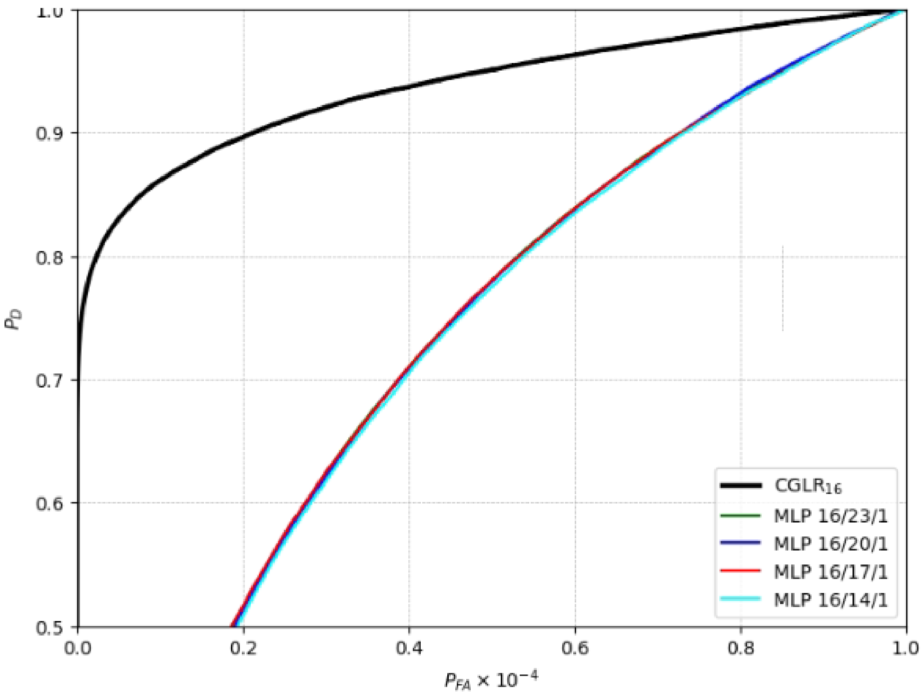
# Spiky K-distributed clutter

Detection of non fluctuating targets in presence of spiky K -distributed clutter (v = 0.5, where v is the shape parameter of the K -distribution): This model is suitable for target detection in sea/land clutter with high resolution radar systems and low grazing angles. In this case, the problem of detecting Swerling V (SWV) targets with unknown Doppler shift is considered.

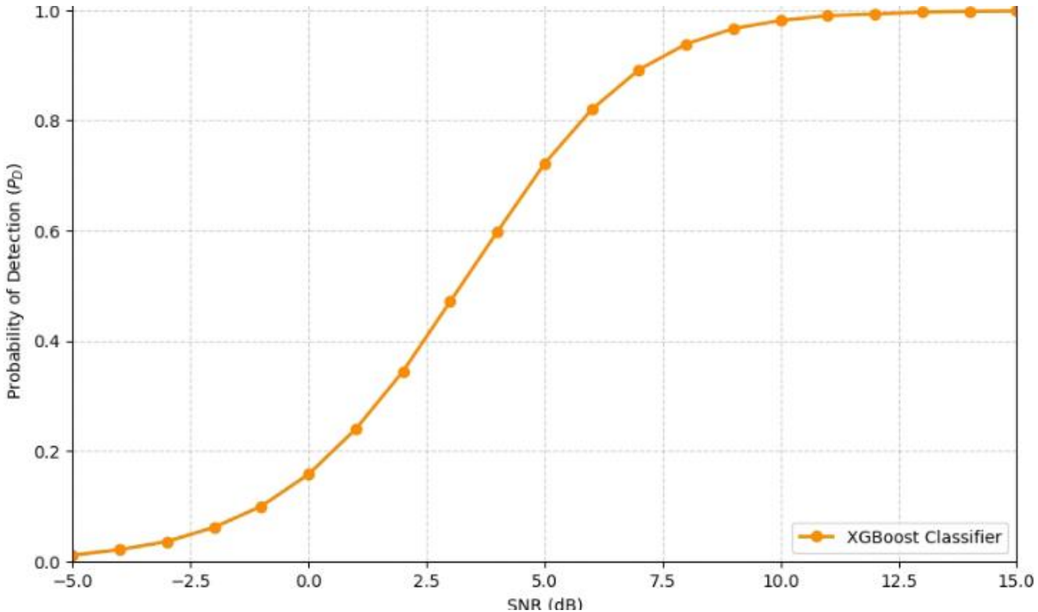
# Detection of non fluctuating targets in presence of spiky K-distributed clutter (SCR = 9 dB, ρ c = 0)



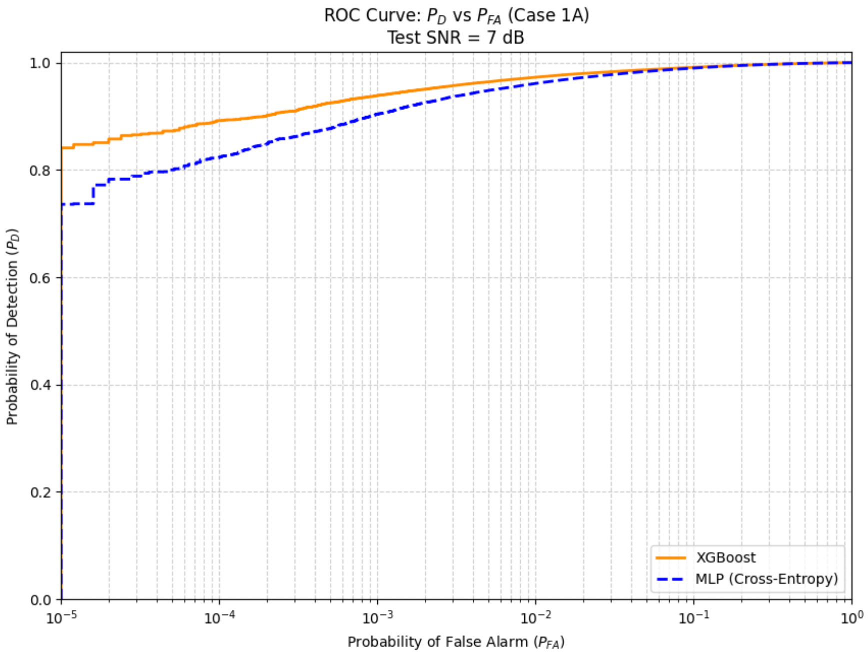
# Detection of non fluctuating targets in presence of spiky K-distributed clutter (SCR = -3 dB, ρ c = 0.9)



# XGBOOST (Novelty)



# XGBOOST (Pd vs Pfa)



THANK YOU