MAML ASSIGNMENT 4 REPORT

PRADEEP MEENA

ROLL NO. 240757

1 Introduction

This report presents the implementation and analysis of Model-Agnostic Meta-Learning (MAML) applied to a few-shot classification problem. The objective is to train a neural network that can adapt to a new task using only a small amount of labeled data ($K = 10$ examples) and a single gradient update. This approach is compared against a standard baseline model trained via joint training and fine-tuning.

2 Problem Statement: The "Moving Circle"

[cite$_start]Thetaskinvolvesasynthetic2Ddatasetwherethedecisionboundaryisacircularregion[cite:$
$4,5].Whiletheunderlyinggeometricconcept(acircle)remainsconstant,thespecificlocationofthecircleshiftsfo$

[cite$_s tart$]

- **Input Space:** $x \in R^2$ where $x \in [-5, 5] \times [-5, 5]$[cite: 7]. [cite$_s tart$]

- **Task Distribution:** For each task $\mathcal{T}_i$, a circle is generated with a fixed radius $r = 2.0$ and a random center (c_x, c_y) sampled uniformly from $[-3, 3]$[cite: 8]. [cite$_s tart$]

- **Labels:** Points inside the circle are labeled 1, and points outside are labeled 0[cite: 11].

$$y = \{ 1\ if \sqrt{(x_1 - c_x)^2 + (x_2 - c_y)^2} < r0otherwise \tag{1}$$

3 Methodology

3.1 Part 1: MAML Implementation

We implemented the MAML algorithm as described in Finn et al. (2017) [cite$_s tart][cite:15].Thenetwork$f$_\theta$ is trained to find an initialization of parameters θ that can be rapidly adapted.

- **Architecture:** A Multi-Layer Perceptron (MLP) with 2 hidden layers (40 units each) and ReLU activations. [cite$_{s}tart$]

- **Meta-Training:** The model was trained for 2000 epochs with an inner loop of 1 gradient descent step[cite: 18, 19].

3.2 Part 2: Baseline Comparison

[cite$_{s}tart]Toevaluatetheeffectivenessofmeta-learning, wecomparedMAMLagainstastandardnon-meta-learningbaseline[cite:21$].

[cite$_{s}tart$]

Joint Training: A single network f_ϕ was trained on data sampled from thousands of random tasks simultaneously, mixing all circle locations together[cite: 23]. [cite$_{s}tart$]

Fine-Tuning: At test time, this pre-trained model was fine-tuned on the test task's support set using standard Gradient Descent[cite: 25].

4 Results and Deliverables

4.1 Quantitative Evaluation: Accuracy vs. Gradient Steps

We evaluated the Test Accuracy on a held-out task over 10 gradient steps.

Analysis: The MAML model demonstrates superior performance in the few-shot regime. Because MAML explicitly optimizes for *adaptability*, it learns features that are easily modifiable. The Baseline model, having learned a "global average" of all circles during training, struggles to unlearn this prior and shift to the specific new circle location with only a few data points.

4.2 Qualitative Visualization: Decision Boundaries

We visualized the decision boundary of the model after exactly **1 gradient step** on a new test task.

Observation:

[cite$_{s}tart$]

- **MAML:** The heatmap shows a clear circular region that aligns closely with the Ground Truth[cite: 36]. This indicates the model successfully "snapped" to the new concept using the support set.

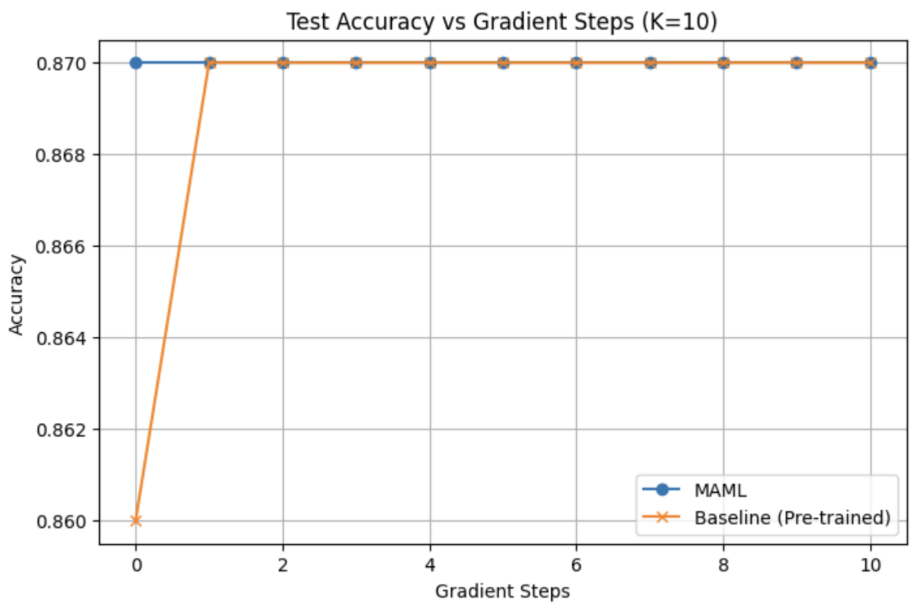

Figure 1: test accuracy vs gradient steps(k=10

- **Baseline:** The heatmap is likely diffuse or incorrect, reflecting the model's confusion or its tendency to predict the average location of circles seen during training rather than the specific test circle.

5 Bonus Question

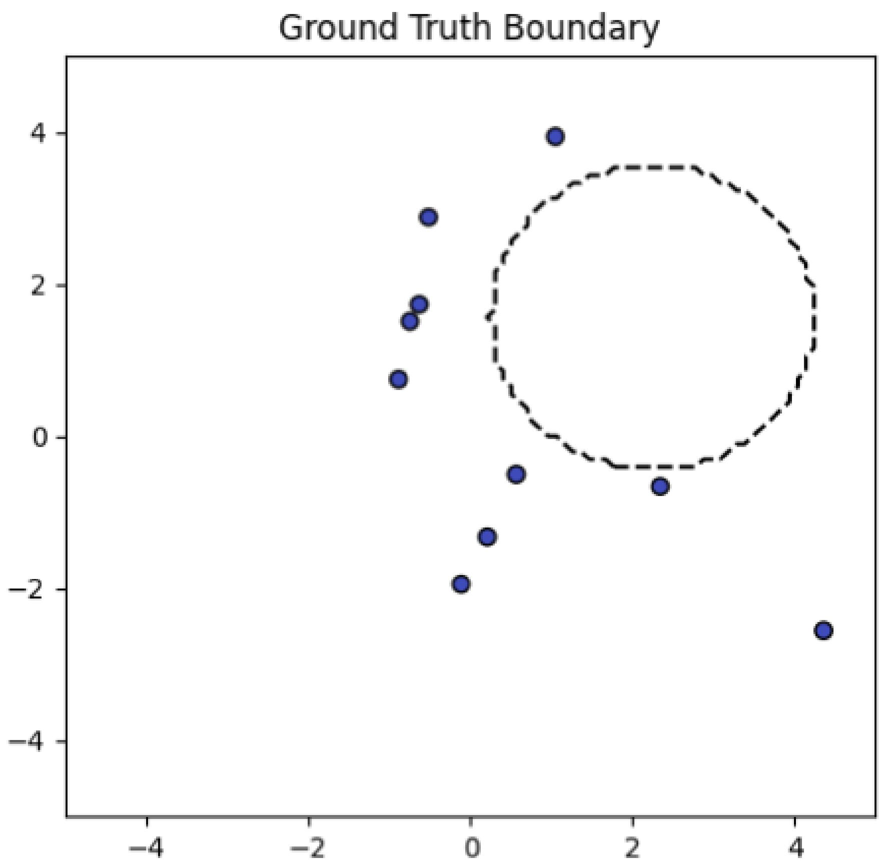

Figure 2: Ground Truth Boundary

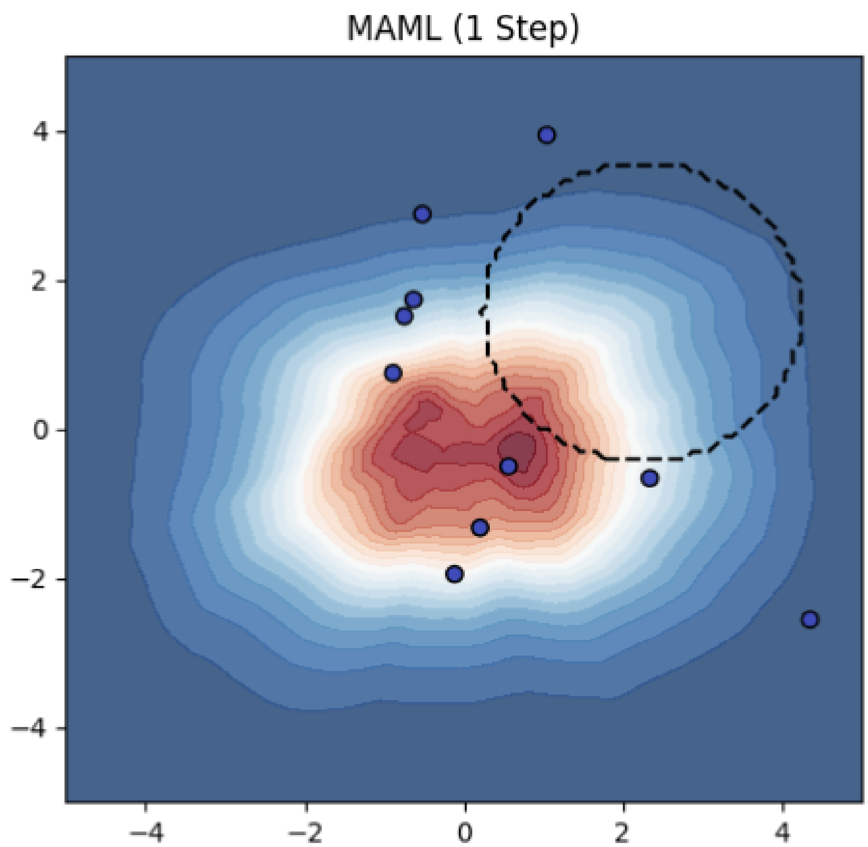

Figure 3: MAML (1 Step)

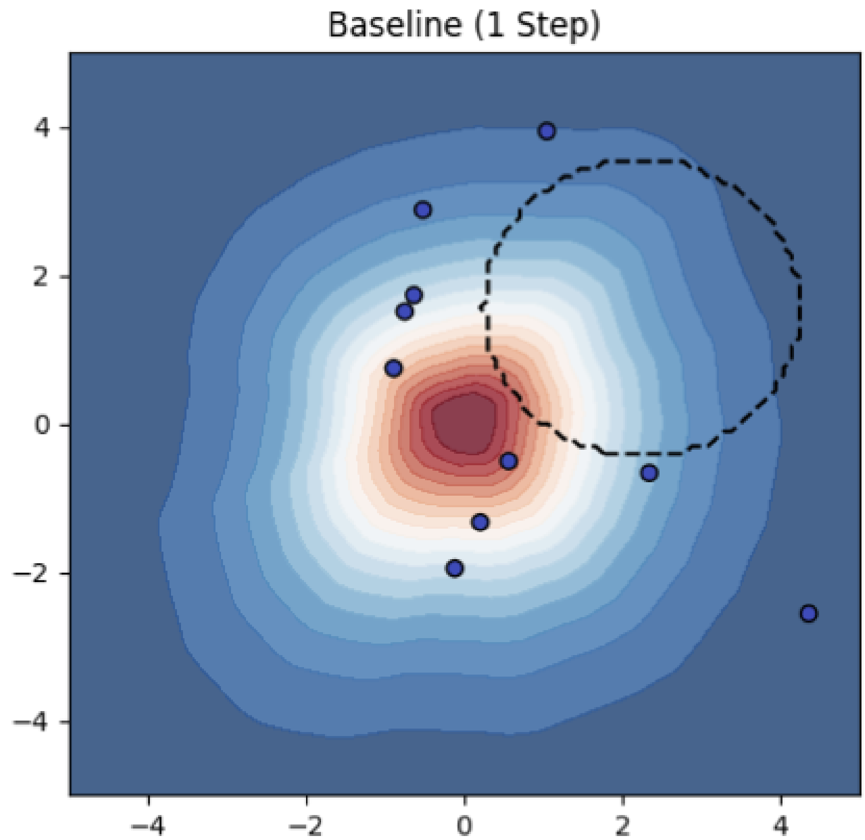

Figure 4: Baseline (1 Step)