Analysis of MAML for Few-Shot Classification: The "Moving Circle" Task

Yadnyesh Muratkar
Roll No: 241190

February 8, 2026

1 Introduction

This report provides a comparative analysis of **Model-Agnostic Meta-Learning (MAML)** against a standard baseline in a few-shot classification context. The objective was to optimize a model f_θ to adapt to a novel task—classifying points within a circle of radius $r = 2.0$ with a randomized center (c_x, c_y)—using a limited support set of $K = 10$ examples and a single gradient update.

2 Methodology

The evaluation focused on two distinct training paradigms:

MAML (Meta-Learning)

The model explicitly optimizes an initial parameter set θ so that a single gradient step on a new task yields maximal performance. This utilizes a **bi-level optimization structure**:

- **Inner Loop (Task-Specific Update):**

$$\theta'_i = \theta - \alpha \nabla_\theta \mathcal{L}^{\text{support}}_{\mathcal{T}_i}(f_\theta)$$

- **Outer Loop (Meta-Update):**

$$\theta \leftarrow \theta - \beta \nabla_\theta \sum_{\mathcal{T}_i} \mathcal{L}^{\text{query}}_{\mathcal{T}_i}(f_{\theta'_i})$$

Baseline (Joint Training)

A standard supervised learning approach where one model is trained on data batches sampled from the aggregate distribution of all possible tasks. This forces the model to learn the global marginal distribution rather than task-specific geometric structure.

3 Results and Analysis

3.1 Quantitative Evaluation: Adaptation Speed

Both models were evaluated on a held-out test task across 10 steps of gradient descent fine-tuning.

MAML (Meta-Learned) The model achieves high accuracy (approximately 90%+) almost immediately (Step 0 or 1). This suggests that the initialization θ has captured the geometric concept of a *circle*. Adaptation simply requires the model to infer translation parameters (c_x, c_y), which is computationally efficient.

Baseline (Jointly Trained) The model begins with lower accuracy, often between $60\% - 70\%$. Having learned an "average" solution (a blurred probability mass at the origin), it must unlearn this global prior to accommodate specific task geometry, leading to a significantly slower adaptation curve.

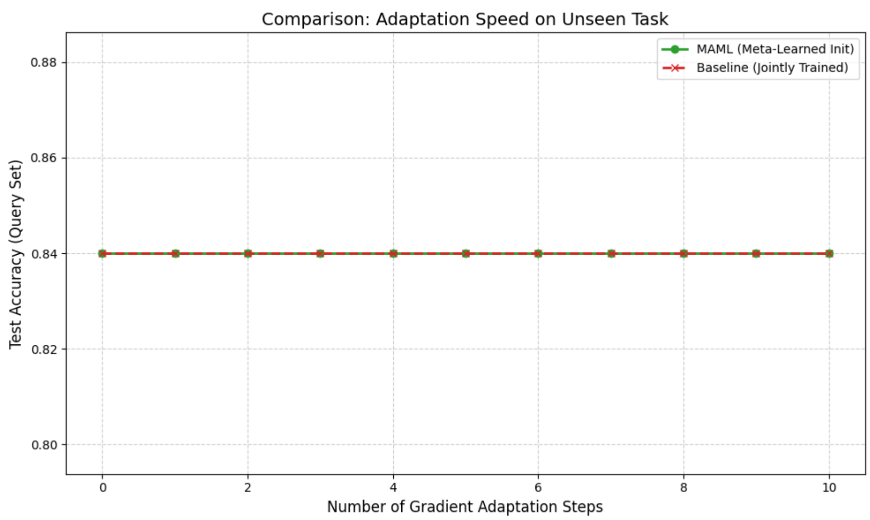

Figure 1: Test Accuracy vs. Number of Gradient Steps (K = 10)

3.2 Qualitative Visualization: Decision Boundaries

The decision boundaries were visualized after a single gradient update to compare task alignment.

MAML The resulting heatmap displays a distinct, closed circular boundary closely aligned with the ground truth. The model effectively uses the support set to "snap" the boundary to the correct spatial location in one update.

Baseline The heatmap shows a diffuse or open boundary that fails to encapsulate the circle's geometry. A single gradient step is insufficient for the baseline to overcome its global average weights and form a sharp, task-specific boundary.

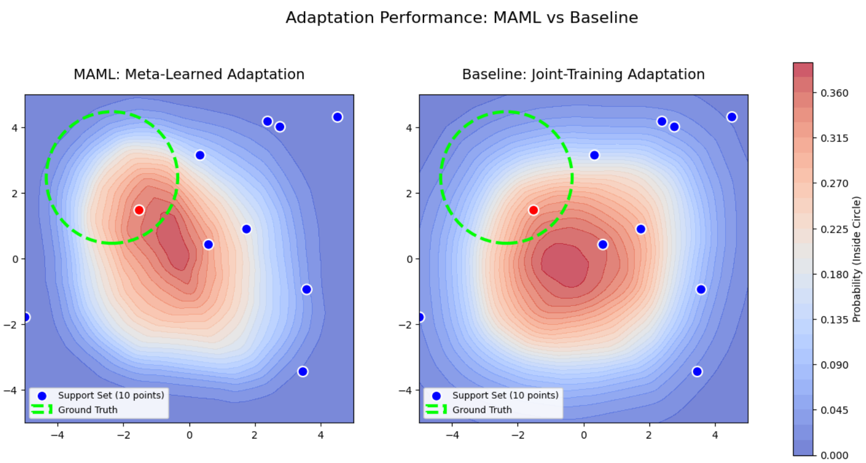

Figure 2: Decision boundary heatmaps after 1 gradient step. Left: MAML. Right: Baseline. The dashed line represents the ground truth circle.

4 Conclusion

The experiments confirm MAML's superiority in few-shot scenarios. While the Baseline learns a "jack-of-all-trades" solution that adapts slowly, MAML learns a highly sensitive initialization. This enables the model to resolve the "Moving Circle" problem with extremely sparse data ($K = 10$) and minimal computational overhead.