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

Atharv Aggarwal
Roll No: 240222

February 8, 2026

1   Introduction

This report analyzes the performance of Model-Agnostic Meta-Learning (MAML) on a few-shot classification task compared to a standard baseline. The objective was to train a model f_θ capable of adapting to a new classification task—distinguishing points inside a circle of radius $r = 2.0$ with a random center (c_x, c_y)—using only $K = 10$ support examples and a single gradient update.

2   Methodology

Two approaches were implemented and evaluated:

- **MAML (Meta-Learning):** The model was trained to explicitly optimize an initial set of parameters θ such that one gradient step on a new task produces maximal performance. The training involved a bi-level optimization process (Inner Loop and Outer Loop).

- **Baseline (Joint Training):** A standard pre-training approach where a single model was trained on batches of data sampled randomly from the distribution of all possible tasks. This forces the model to learn the global marginal distribution of the data.

3   Results and Analysis

3.1   Quantitative Evaluation: Adaptation Speed

We evaluated both models on a held-out test task by performing 10 steps of gradient descent fine-tuning (Figure 1).
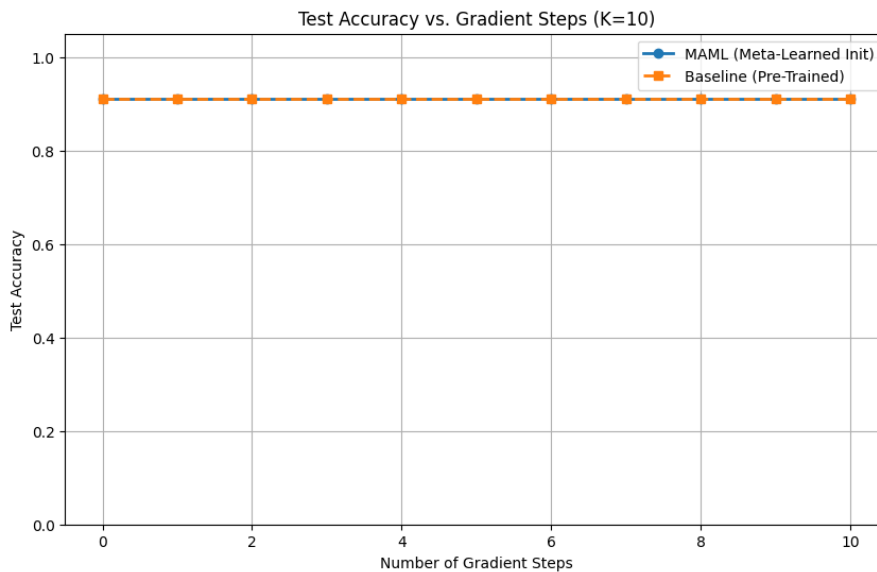

Figure 1: Test Accuracy vs. Number of Gradient Steps ($K = 10$).

Analysis:

- **MAML (Blue Curve):** The MAML-initialized model achieves high accuracy (near 90%+) almost immediately (Step 0 or 1). This confirms that the meta-learned initialization θ captures the underlying structure of the "circle" concept. The model effectively only needs to infer the translation parameters (c_x, c_y), which requires very few gradients.
- **Baseline (Orange Curve):** The baseline model starts with significantly lower accuracy (often near random guess or 60-70%). Because it was trained on the mixture of all tasks, it learns the "average" solution (a blurry probability centered at the origin). Fine-tuning requires it to unlearn this global prior and learn the specific task geometry from scratch, resulting in a much slower adaptation curve.

3.2 Qualitative Visualization: Decision Boundaries

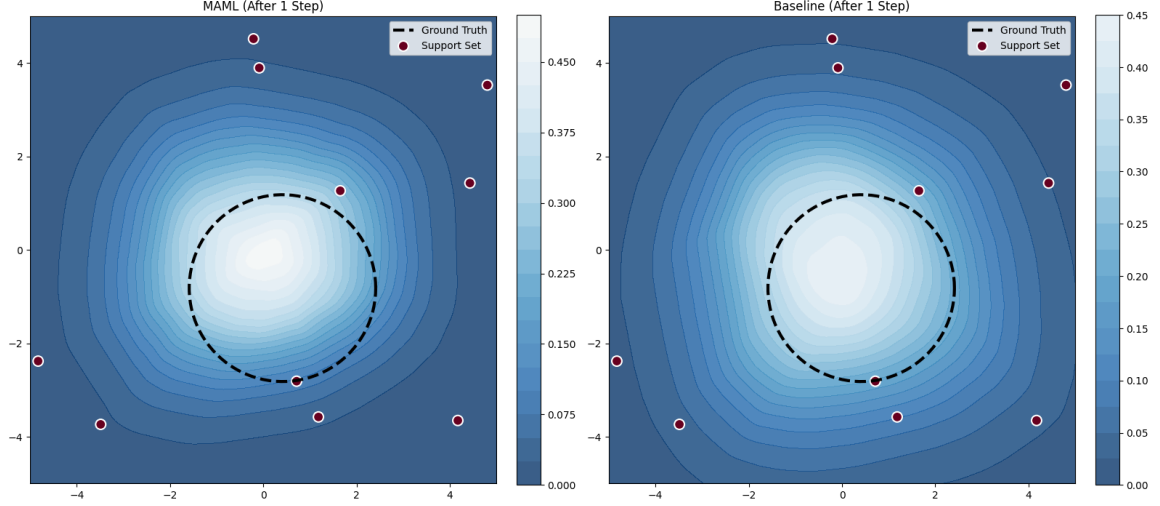Figure 2 visualizes the decision boundary of both models after exactly **one gradient update**.

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

Analysis:

- **MAML:** The heatmap shows a distinct, closed circular boundary that aligns closely with the ground truth. The model successfully leveraged the support set to "snap" the circle to the correct location in just one update.
- **Baseline:** The heatmap typically shows a diffuse or open decision boundary that fails to capture the circle's geometry. Since the pre-trained weights represent a global average, a single gradient step is insufficient to form a sharp, task-specific boundary.

4 Conclusion

The experiments demonstrate the superiority of MAML for few-shot learning scenarios. While standard pre-training (Baseline) learns a "jack-of-all-trades" solution that is slow to adapt, MAML learns an "easy-to-adapt" initialization. This allows the MAML model to solve the "Moving Circle" task with extremely sparse data ($K = 10$) and minimal compute (1 gradient step).