**MAML-Based Models**

Algorithm 1: Model-Agnostic Meta-Learning

1. **Input:** Dataset $D$, hyperparameters $h$.
2. **Output:** Trained model $M$.
3. **Procedure:**
   - Randomly split $D$ into training set $D_{train}$ and test set $D_{test}$.
   - For each $i$, do:
     - On $D_{train}$, train $M_i$.
     - On $D_{test}$, evaluate $M_i$.
   - Return $M$.

**Parallel MAML**

- Ensemble baseline with a MAML model at eval time.
- Pick the result with the highest confidence.

<table>
<thead>
<tr>
<th>Version</th>
<th>F1</th>
<th>EM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>43.887</td>
<td>24.337</td>
</tr>
<tr>
<td>Vanda-MAML</td>
<td>47.021</td>
<td>26.692</td>
</tr>
<tr>
<td>Transformer</td>
<td>47.858</td>
<td>31.475</td>
</tr>
</tbody>
</table>

---

**Data Augmentation**

- Correct Masking

- Incorrect Masking

---

**Transformer Level Attention (Clustering)**

**Architecture**

**Model Interpretation**

**Results**

<table>
<thead>
<tr>
<th>Model Type</th>
<th>EM</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>43.887</td>
<td>24.337</td>
</tr>
<tr>
<td>Baseline + OOD Detection</td>
<td>42.74</td>
<td>26.86</td>
</tr>
<tr>
<td>Baseline + OOD Detection + OOD Detection</td>
<td>42.74</td>
<td>26.86</td>
</tr>
<tr>
<td>Clustering + OOD Detection</td>
<td>50.00</td>
<td>26.84</td>
</tr>
<tr>
<td>Clustering + OOD Detection + OOD Detection</td>
<td>50.00</td>
<td>26.84</td>
</tr>
<tr>
<td>Clustering + OOD Detection + OOD Detection</td>
<td>50.00</td>
<td>26.84</td>
</tr>
<tr>
<td>Clustering + OOD Detection + OOD Detection</td>
<td>50.00</td>
<td>26.84</td>
</tr>
</tbody>
</table>

**Next Steps:**

- Tune $a$ and $Y$ and experiment with $Y$ being based on ground truth answer length in sigmoid loss method.
- Try to make task distribution for MAML.
- Try synonym replacement for data augmentation.
- Combine these approaches (e.g., MAML w/ data augmentation, sigmoid loss w/ anything other than baseline).

---

**Exploration of Techniques for Robust Question Answering**

**Introduction**

The use of MAML in non-exam tasks, such as question answering, can be challenging. MAML is a popular approach for learning to learn, which allows models to quickly adapt to new tasks with a small number of training examples. However, MAML's performance can be limited by the difficulty of the tasks it is applied to. This paper explores techniques for improving MAML's performance in question answering, as well as other tasks that require fast adaptation.

One challenge with MAML is that it is often difficult to predict the performance of the final model before training. To address this, we propose a new method that uses a combination of MAML and data augmentation to improve performance. Our experiments show that this method performs well on a variety of tasks, including both natural language processing and computer vision.

**Conclusions:**

- MAML seems to underperform due to the distribution of tasks not being as clear when compared to something like a DRNN.
- Sigmoid loss made F1 scores higher on the imdb task.
- Having organized code is very important for running tests efficiently.