**### Task 1 ###**

**Source Documents:**

Link to OpenReview: <https://openreview.net/forum?id=Arn2E4IRjEB>

**Meta-review:**

After an extensive back-and-forth discussion, all reviewers felt positively about the paper and its contribution, and lean towards acceptance.\n\nHowever, there were issues regarding clarity in the submitted draft — this has been largely addressed through clarifications from the authors, but it is essential that the discussion here is incorporated into the final version of the manuscript (and in particular, the responses to reviewer 1AuD).

**### Task 2 ###**

**Source Documents:**

Link to OpenReview: <https://openreview.net/forum?id=SJgmR0NKPr>

**Meta-review:**

The paper proposes an alternative to BPTT for training recurrent neural networks based on an explicit state variable, which is trained to improve both the prediction accuracy and the prediction of the next state. One of the benefits of the methods is that it can be used for online training, where BPTT cannot be used in its exact form. Theoretical analysis is developed to show that the algorithm converges to a fixed point. Overall, the reviewers appreciate the clarity of the paper, and find the theory and the experimental evaluation to be reasonably well balanced. After a round of discussion, the authors improved the paper according to the reviews. The final assessments are overall positive, and I’m therefore recommending accepting this paper.

**### Task 3 ###**

**Source Documents:**

Link to OpenReview: <https://openreview.net/forum?id=cu7IUiOhujH>

**Meta-review:**

This paper introduces supervised contrastive learning loss on top of typical cross-entropy loss for fine-tuning language model for downstream tasks. While the idea is simple and has been used in vision literature (as pointed out by R1 & R4), its application LM is first introduced in this paper. The experimental gain is small in the regular setting but clearer gains in a few-shot learning setting and noisy training dataset (through back translation) setting. Overall the paper is clearly written and experiments are carefully studied. During the discussion phase, the authors provided results on the full GLUE dataset as well as other ablation studies (e.g., CE+CE recommended by R2), improving the paper.

**### Task 4 ###**

**Source Documents:**

Link to OpenReview: <https://openreview.net/forum?id=ryenvpEKDr>

**Meta-review:**

The paper presents a very interesting idea for estimating the held-out error of deep models as a function of model and data set size. The authors intuit what the shape of the error should be, then they fit the parameters of a function of the desired shape and show that this has predictive power. I find this idea quite refreshing and the paper is well written with good experiments. Please make sure that the final version contains the cross-validation results provided during the rebuttal.

**### Task 5 ###**

**Source Documents:**

Link to OpenReview: <https://openreview.net/forum?id=f9AIc3mEprf>

**Meta-review:**

This paper introduces an ImageNet-scale benchmark UIMNET for uncertainty estimation of deep image classifiers and evaluates prior works under the proposed benchmark. Two reviewers suggest reject, and one reviewer does acceptance. In the discussion period, the authors did not provide any response for many concerns of reviewers, e.g., weak baselines, weak novelty, and lack of justification for the current design. Hence, given the current status, AC recommends reject.

**### Task 6 ###**

**Source Documents:**

Link to OpenReview: <https://openreview.net/forum?id=wZYWwJvkneF>

**Meta-review:**

This paper generated an involved discussion between the reviewers and the authors, as well as between the reviewers themselves.\nThe paper essentially combines two well-known baselines in MARL domain, and therefore was judged as a report of experiments for which reproducibility is of particular importance. The reviewers had an intense discussion about reproducibility, two of them tried the code provided by the authors and one of them commented on it based on their experience. The analysis raised significant doubts about reproducibility (it is unfortunate that the authors did not provide configs on their experiments, to enable to rerun them exactly).