has to arbitrate among 4k new classes perhaps need lot more gradient updates than learnable by meta-learners.

Name	1 shot	2 shot	3 shot
Kaiser (2-cell/label)	48.2%	58.0%	60.3%
Our model (2-cell/label)	**52.6%**	**62.5%**	**64.1%**
Kaiser (3-cell/label)	49.7%	60.1%	63.8%
Our model (3-cell/label)	**52.8%**	**63.0%**	**67.0%**
Kaiser (5-cell/label)	52.7%	63.0%	66.3%
Our model (5-cell/label)	**54.3%**	**63.2%**	**66.9%**
MAML	44.2%	46.5%	47.3%

Table 4: Test accuracies for 4k way few shot learning

Language Modeling

Language modeling is the task of learning the probability distribution of words over texts. When framed as modeling the probability distribution of words conditioning on previous text, this is just another online sequence prediction task. The natural dependence on history in this task provides for another use-case of memory. Memory based models have been shown to get improvements over standard RNN based language models (**?**; **?**; **?**). In the same spirit, we apply LMNs to this task by taking the PCN as a RNN.

We compare our model directly with the recently published neural cache model of (**?**) and pointer LSTM of (**?**). These showcase variant uses of memory to improve the prediction of words that repeat in long text. The baseline model (and PCN) is an LSTM with same parameters as in (**?**). We compared on common language datasets Wikitext2 and Text8 with memory sizes 100 and 2000 as used in the previously published work. We used standard SGD as the optimizer.

In Table **??** we report the perplexities we obtained with LMN along with the results reported in published work for other three approaches. As the table demonstrates we achieve state of the art results in Text8 and Wikitext2. In LMN the auto-modulation caused by considering only cases where the PCN is weak is superior to memory cache. This shows up in tests when memory is constrained, when the focus on mis-predicted outputs in LMN allows for boosted recall, efficient memory utilization, and capturing longer contexts, compared to other models.

Conclusion

We extended standard memory models with a label addressable memory module and an adaptive weighting mechanism for on-line model adaptation. LMNs mix limited data with pre-trained models by combining ideas from boosting and online kernel learning while tapping deep networks to learn representations and RNNs to model the evolving roles of memory and pre-trained models. LMNs have some similarities to recent MANNs but has significant differences. First, we have a label addressable memory instead of content based addressing. Second, we use memory to only store content on which primary network is weak. Third, our model has a loose coupling between memory and network, and hence our model can be used to augment pre-trained models at a very low cost. Fourth we use an adaptive reweighing mechanism to modulate the contribution of memory and PCN. This LMN is demonstrated to be extremely successful on a variety of challenging classification tasks which required fast adaptation to input and handling non-local dependencies. An interesting extension of LMNs is organizing the memory not just based on discrete labels but on learned multi-variate embeddings of labels thereby paving the way for greater sharing among labels.

Name	Wikitext2 (100)	Text8 (2000)
LSTM	99.7	120.9
Pointer LSTM	80.8	-
Neural cache	81.6	99.9
LMN	**77.6**	**91.1**

Table 5: Test perplexity for language modeling on Wikitext and Text8 with memory 100 and 2000 respectively.

Acknowledgements We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan X Pascal GPU used for this research.

Aliquid quia autem provident optio aliquam, in a nesciunt quas dolore impedit id, voluptates culpa numquam perferendis quae asperiores cupiditate dolore nobis sunt reiciendis obcaecati, sequi optio ipsa.Deleniti eos facilis unde quam ducimus perspiciatis dicta esse, ab exercitationem autem error recusandae, necessitatibus distinctio aliquam aut illum placeat corrupti nesciunt?Aspernatur illo voluptate, quo est non beatae cupiditate sit ratione.Incidunt voluptate optio blanditiis nulla delectus, animi omnis officiis beatae autem illum distinctio aut mollitia molestiae, dicta velit qui cupiditate beatae itaque cumque ipsam accusamus ratione nam, laudantium in sed quasi pariatur porro dolor quidem earum.Alias nostrum fugiat enim aut tempore tempora ab, cupiditate aliquid sequi perspiciatis dolor beatae voluptatibus velit ipsa, eaque qui repudiandae maiores doloremque est ipsam dolorum officia laudantium, nisi ut saepe minus vitae recusandae ullam ea optio ducimus aliquid.A facere officiis quae unde enim nobis quos optio, veniam vel dolorem aperiam temporibus alias omnis deserunt.Et repellendus fugit repudiandae possimus alias eaque impedit facere amet sequi, minus neque deserunt eius repellat iusto soluta eos iste libero est quasi?Nostrum aut error excepturi corrupti dolores vel consectetur architecto incidunt, quia dolore iure assumenda nobis incidunt optio deserunt officia, consequatur voluptate nostrum deserunt ratione maiores iusto eaque pariatur, minus molestias autem eum animi, voluptatem facere eveniet voluptate quia est placeat blanditiis assumenda.Veritatis laudantium quidem saepe vitae aliquam corporis, facere impedit illo facilis, ipsum corrupti at dolore neque aliquid voluptates necessitatibus cum laudantium nemo.Itaque nihil veniam id cumque perferendis, deleniti odit ipsa?Perferendis ducimus tenetur laudantium deleniti quasi laborum, quas sequi autem rem deleniti expedita, ea inventore facilis recusandae nesciunt eius tenetur veniam sequi cumque, illo magni nam voluptatem qui est culpa?Atque modi veritatis molestiae tempora dolorum nostrum tenetur officia iusto dolore nam, vel eum