

Continual Quality Estimation with Online Bayesian Meta-Learning

Abiola Obamuyide¹ Marina Fomicheva¹ Lucia Specia^{1,2}

¹Department of Computer Science, University of Sheffield

²Department of Computing, Imperial College London
United Kingdom

{a.obamuyide, m.fomicheva, l.specia}@sheffield.ac.uk

Abstract

Most current quality estimation (QE) models for machine translation are trained and evaluated in a static setting where training and test data are assumed to be from a fixed distribution. However, in real-life settings, the test data that a deployed QE model would be exposed to may differ from its training data. In particular, training samples are often labelled by one or a small set of annotators, whose perceptions of translation quality and needs may differ substantially from those of end-users, who will employ predictions in practice. To address this challenge, we propose an online Bayesian meta-learning framework for the continuous training of QE models that is able to adapt them to the needs of different users, while being robust to distributional shifts in training and test data. Experiments on data with varying number of users and language characteristics validate the effectiveness of the proposed approach.

1 Introduction

Quality Estimation (QE) models aim to evaluate the output of Machine Translation (MT) systems at run-time, when no reference translations are available (Blatz et al., 2004; Specia et al., 2009). QE models can be applied for instance to improve translation productivity by selecting high-quality translations amongst several candidates. A number of approaches have been proposed for this task (Specia et al., 2009, 2015; Kim et al., 2017; Kepler et al., 2019; Ranasinghe et al., 2020), and a shared task yearly benchmarks proposed approaches (Fonseca et al., 2019; Specia et al., 2020).

Different users of MT output have varying quality needs and standards, depending for instance on the downstream task at hand, or the level of their knowledge of the languages involved, and training for the task. Thus, the perception of the quality

of MT output can be subjective, and therefore the quality estimates obtained from a model trained on data from one set of users may not serve the needs of a different set users. However, most existing QE models are trained and evaluated in a static setting which assumes a fixed distribution of train and test data. This consequently leads to suboptimal performance when faced with test data from a different set of users in practice.

The few previous approaches to develop QE models that are able to learn from a continuous stream of data suffer from the following limitations: they do not have an explicit objective that encourages the model to exploit common structures shared among different users to continually adapt efficiently for new users (Turchi et al., 2014), or assume a fixed number of users, and that the identity of each user is known in advance (de Souza et al., 2015). In addition, these previous approaches do not explicitly account for the underlying uncertainties in the data in order to improve performance.

In contrast, we propose a continual meta-learning framework that makes none of the aforementioned assumptions, but instead considers each user as a task and explicitly meta-learns the common structure shared among different users. This approach further exploits the underlying uncertainties in the streaming data through Bayesian inference to improve performance. In addition, the proposed approach is applicable even in a setting where no user identities are available, for instance due to privacy concerns, but where we still want to learn and adapt as efficiently as possible from supervision data that arrives incrementally.

2 Background

2.1 Continual Learning

Continual learning (Ring, 1994; Thrun, 1996; Zhao and Schmidhuber, 1996) aims to develop mod-

els that are capable of learning from a continuous stream of sequential tasks, $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_T$, with each task \mathcal{T}_t having its associated train \mathcal{D}_t^{train} , validation \mathcal{D}_t^{val} and test \mathcal{D}_t^{test} splits. A major challenge associated with learning in this setting is the issue of *catastrophic forgetting*, where a model forgets knowledge of how to perform previous tasks as new tasks are encountered. Most recent work in lifelong learning has focused on ways of mitigating catastrophic forgetting, and approaches proposed include replay-based methods (Rebuffi et al., 2017; Lopez-Paz and Ranzato, 2017; Chaudhry et al., 2019), which replay either stored or generated samples to remind the model of how to perform previous tasks; regularization-based methods (Kirkpatrick et al., 2017; Zenke et al., 2017), which utilize an additional regularization term to enforce retaining knowledge learned from previous tasks; and parameter-isolation methods, which make use of dedicated parameters for each task to prevent interference among tasks (Rusu et al., 2016; Fernando et al., 2017). Lange et al. (2019) presents an overview of recent continual learning methods. Research in continual learning can generally be carried in one of two settings (Aljundi et al., 2019): in a *task-incremental* continual learning setting, where the learner is sequentially given access to all the data of each task and is allowed to make multiple passes over it, with task boundaries and identities known to the learner; or in an *online continual learning* setting, where the learner is only allowed a single pass over the data of each task, and with no task identities or boundaries known to the learner. In this work we conduct experiments in the online continual learning setting.

2.2 Meta-Learning

The goal of meta-learning, also known as learning to learn (Schmidhuber, 1987; Thrun and Pratt, 1998), is to develop models that can learn more efficiently over time, by generalizing from knowledge of how to solve related tasks from a given distribution of tasks. Given a learner model f_w , for instance a neural network parametrized by w , and a distribution $p(\mathcal{T})$ over tasks \mathcal{T} , gradient-based meta-learning approaches such as MAML (Finn et al., 2017) seek to learn the parameters of the learner model which can be quickly adapted to new tasks sampled from the same distribution of tasks. In formal terms, these approaches seek parameters

that optimize the meta-objective:

$$\min_w \mathbb{E}_{\mathcal{T} \sim p(\mathcal{T})} [\mathcal{L}_{\mathcal{T}}(\mathcal{U}_k(w; \mathcal{D}_{\mathcal{T}}))] \quad (1)$$

where $\mathcal{L}_{\mathcal{T}}$ is the loss and $\mathcal{D}_{\mathcal{T}}$ is training data from task \mathcal{T} , and \mathcal{U}_k denotes k steps of a gradient descent learning rule such as SGD.

In order to account for uncertainty and improve robustness, Bayesian approaches to meta-learning have also been proposed (Kim et al., 2018; Finn et al., 2018; Ravi and Beatson, 2019; Wang et al., 2020; Nguyen et al., 2020).

2.3 Meta-Learning for Continual Learning

Meta-learning for continual learning methods generally make use of the meta-learning objective one task at a time to ensure that learning on the current task does not lead to catastrophic forgetting on previous tasks. For instance, both Riemer et al. (2019) and Obamuyide and Vlachos (2019) propose to combine REPTILE (Nichol and Schulman, 2018), a first order meta-learning algorithm, together with experience replay to improve performance during continual learning. Javed and White (2019) proposed an online-aware meta-learning (OML) objective for learning representations that are less prone to catastrophic forgetting during continual learning. Holla et al. (2020) proposed to combine the OML objective together with experience replay for improved continual learning performance. Recently, Gupta et al. (2020) proposed Look-Ahead MAML (LA-MAML), which meta-learns per-parameter learning rates to help adapt to changing data distributions during continual learning.

These approaches have demonstrated that meta-learning can yield performance improvements for continual learning. Our work builds on these approaches and additionally demonstrates that the performance of meta-learning for continual learning can be further improved with the combination of an adaptive learning rate and Bayesian inference.

2.4 Bayesian Inference with Stein Variational Gradient Descent

Stein Variational Gradient Descent (SVGD) (Liu and Wang, 2016) is a Bayesian inference method which works by initializing a set of samples, also known as particles, from a simple distribution and iteratively updating the particles to match samples from a target distribution. Because its particle update rule is deterministic and differentiable, it can

be used to perform Bayesian inference in the meta-learning inner loop, since the entire update process can still be differentiated through for gradient-based updates from the outer loop.

In order to obtain N samples from a posterior $P(\mathbf{w})$, SVGD maintains N samples of model parameters, and iteratively transports the samples to match samples from the target distribution. Let the samples be represented by $\mathbf{W} = \{\mathbf{w}^n\}_{n=1}^N$. At each successive iteration t , SVGD updates each sample with the following update rule:

$$\mathbf{w}_{t+1} \leftarrow \mathbf{w}_t + \alpha_t \phi(\mathbf{w}_t) \quad (2)$$

where $\phi(\mathbf{w}_t) =$

$$\frac{1}{N} \sum_{n=1}^N [k(\mathbf{w}_t^n, \mathbf{w}_t) \nabla_{\mathbf{w}_t^n} \log p(\mathbf{w}_t^n) + \nabla_{\mathbf{w}_t^n} k(\mathbf{w}_t^n, \mathbf{w}_t)] \quad (3)$$

α_t is a step-size parameter and $k(\cdot, \cdot)$ is a positive-definite kernel, such as the RBF kernel.

Intuitively, the first term in Equation 3 implies that a particle determines its update direction through a weighted aggregate of the gradients from the other particles, with the kernel distance between the particles serving as the weight. Thus, closer particles have more weight in the aggregate. The second term of the equation can be understood as a repulsive force that prevents the particles from collapsing to a single point. For the case when the number of particles is one, the SVGD update procedure reduces to standard gradient ascent on the objective $p(\mathbf{w})$ for any kernel with the property $\nabla_{\mathbf{w}} k(\mathbf{w}, \mathbf{w}) = 0$, such as the RBF kernel. SVGD has been applied in a wide range of settings, including reinforcement learning (Liu et al., 2017; Haarnoja et al., 2017), uncertainty quantification (Zhu and Zabarar, 2018) and to improve performance in an offline meta-learning setup (Kim et al., 2018) which requires all tasks ahead of training. In this work we adapt SVGD to an online continual meta-learning setting for a natural language task.

3 Meta-Learning for Continual Learning with Adaptive SVGD

Learning continually from a stream of observations with varying underlying distributions involves dealing with various sources of uncertainty, which a model should properly account for in order to enhance its continual learning performance. One source of uncertainty is in the learning rate, that is, how fast learning should proceed on new data

in order to both reduce catastrophic forgetting and enhance performance on the current task. Another source is the inherent uncertainty in the values of the model’s parameters themselves. Learning an adaptive learning rate, for instance as proposed in Gupta et al. (2020), can help account for the first source of uncertainty, and Bayesian inference can be used to help a model account for the other source of uncertainty. In order to properly model both sources of uncertainty during continual learning, we propose to both perform inference of model parameters with SVGD, and meta-learn an adaptive per-parameter learning rate for SVGD updates. Thus, the SVGD update in Equation 2 now becomes:

$$\mathbf{w}_{t+1} \leftarrow \mathbf{w}_t + \alpha_t \cdot \phi(\mathbf{w}_t) \quad (4)$$

where α_t is a learnable parameter containing per-parameter learning rates, and \cdot is the dot product.

The aim is then to meta-learn both the parameters of the model and the per-parameter learning rates that enhance continual learning performance. The advantage of this approach is that it allows for greater flexibility to adapt to non-stationary data distributions during continual learning. In the experiments, we demonstrate that this change leads to improved performance for the task of continual quality estimation. The proposed approach is illustrated in Algorithm 1.

We first initialize the parameters of the QE model, and the learning rate (line 1). Then for each mini-batch in a task t that arrives, we store its training instances in the buffer with a probability p (lines 2-6). In the inner loop, we perform K SVGD updates (using Equation 4) starting from the initial model parameters W_0 (lines 7-9). In the outer loop, instances in the current mini-batch are augmented with instances sampled from the buffer (line 10). Finally, the augmented mini-batch is used to perform a meta-update on the learning rate (line 11), and on the parameters of the QE model (line 12). Because this approach can also be considered the online counterpart to the Bayesian Model Agnostic Meta-Learning approach of Kim et al. (2018), we refer to it as Continual Quality Estimation with Online Bayesian Meta-Learning (*CQE-OBML*).

4 Experiments and Results

The QT21 Dataset We evaluate our approach with the publicly available QT21 (Specia et al., 2017), a large-scale dataset containing translations

Algorithm 1 Continual Quality Estimation with Online Bayesian Meta-Learning (*CQE-OBML*)

Require: QE model f_{W_0} , learning rates α_0, β **Require:** Buffer B , update probability p

```
1: Initialize  $W_0, \alpha_0$ 
2: for  $t = 1, 2, 3, \dots$  do
3:   for each  $(X_t, Y_t)$  in  $D_t^{train}$  do
4:     if  $random() < p$  then
5:       Update  $B \leftarrow B \cup (X_t, Y_t)$ 
6:     end if
7:     for  $k = 1, \dots, K$  do
8:        $W_k = SVGD(W_{k-1}, \alpha_0, X_t, Y_t)$ 
9:     end for
10:     $(X_v, Y_v) \leftarrow (X_t, Y_t) \cup sample(B)$ 
11:     $\alpha_0 \leftarrow \alpha_0 - \beta \nabla_{\alpha_0} \mathcal{L}_t(f_{W_k}(X_v), Y_v)$ 
12:     $W_0 \leftarrow W_0 - \alpha_0 \cdot \nabla_{W_0} \mathcal{L}_t(f_{W_k}(X_v), Y_v)$ 
13:  end for
14: end for
```

PE ID	Train	Dev	Test
PE1	1440	360	200
PE2	2160	540	300
PE3	1444	361	195
PE4	1834	459	244
PE5	4866	1217	617
PE6	1677	420	203
PE7	1567	392	241
Total	14988	3749	2000

(a) QT21 en-lv (nmt)

PE ID	Train	Dev	Test
PE1	9952	2488	559
PE2	3445	862	193
PE3	8770	2193	537
PE4	4579	1145	276
PE5	7651	1913	435
Total	34397	8601	2000

(b) QT21 en-cs (smt)

Table 1: Number of instances per post-editor (PE) for the QT21 dataset.

from both statistical (smt) and neural (nmt) machine translation systems in multiple language directions.¹ This is the largest dataset with annotator information available. We use data from the English-Latvian (en-lv) and English-Czech (en-cs) language pairs. These languages were chosen as they contain the largest number of annotators. Each instance in the dataset is a tuple of source sentence, its machine translation, the corresponding post-edited translation by a professional translator (post-editor), a reference translation and other information such as (anonymized) post-editor identifier. We construct a QE dataset from this corpus by

¹<http://www.qt21.eu/resources/data/>

computing the HTER (Snover et al., 2006) values between each source sentence and its post-edited translation. We thereafter split the data into train, dev and test splits for each post-editor. A breakdown of the number of train, dev and test instances per post-editor is available in Table 1.

Benchmark Approaches SEQUENTIAL is a baseline trained sequentially over the streaming data of each task. In each round, the model parameters are initialized from that of the previous round; A-GEM (Chaudhry et al., 2019) is a continual learning method which utilizes the gradients of samples of previous tasks saved in a buffer as an optimization constraint to prevent catastrophic forgetting; OML-ER (Holla et al., 2020) augments the Online-Aware Meta-Learning approach of Javed and White (2019) with experience replay from a buffer; LA-MAML (Gupta et al., 2020) learns per-parameter learning rates using meta-learning; MTL-IID is trained on the concatenated and shuffled data from all users for multiple epochs in multi-task fashion. It assumes i.i.d access to the data from all users, and thus serves as an upper-bound for the performance.

QE Model The quality estimation model used by all continual learning methods is based on multi-lingual DistilBERT (Sanh et al., 2019), a smaller version of multi-lingual BERT (Devlin et al., 2019) trained with knowledge distillation (Buciluă et al., 2006; Hinton et al., 2015). It accepts as input the source and machine translation outputs concatenated as a single text, separated by a ‘[SEP]’ token and prepended with a ‘[CLS]’ token. The representation of the ‘[CLS]’ token is then passed to a linear layer to predict HTER (Snover et al., 2006) values as regression targets.

Evaluation We report Pearson’s r correlation scores and Mean Absolute Error (MAE) between model output and gold labels, both standard evaluation metrics in QE.

Each experiment is repeated across five (5) different orders of the tasks and five (5) different random seeds, and we report the average.

4.1 Comparison with Benchmark Approaches

The results of our approach in comparison with other benchmark approaches are presented in Table 2. We can observe that naively training sequentially on the data of each task as it arrives (SEQUENTIAL) leads to poor results.

Method	en-lv		en-cs	
	Pearson \uparrow	MAE \downarrow	Pearson \uparrow	MAE \downarrow
MTL-IID	59.17	0.1450	54.79	0.1547
SEQUENTIAL	47.07	0.1773	50.08	0.1689
A-GEM	46.29	0.1794	46.49	0.1736
OML-ER	52.58	0.1621	50.40	0.1635
LA-MAML	52.86	0.1621	50.56	0.1631
CQE-OBML	53.67	0.1596	51.19	0.1619

Table 2: Comparison with benchmark approaches.

OML-ER outperforms both SEQUENTIAL and A-GEM, likely because of its combination of meta-learning and experience replay, which makes it better able to combat forgetting. LA-MAML slightly improves over the results of OML-ER, as a result of its meta-learned learning rate. We find that our approach, CQE-OBML, which combines a meta-learned adaptive learning rate together with Bayesian inference, outperforms previous approaches. This demonstrates the effectiveness of adequately modelling the various sources of uncertainty in continual meta-learning.

4.2 Analysis of Model Components

We investigate the effect of the various components of our approach through an ablation study. As shown in Table 3, our approach (CQE-OBML) without the adaptive learning rate (-LR (α)) has a drop in performance, especially for en-cs. Without inference with SVGD (-SVGD), we observe a larger reduction in performance on both datasets, demonstrating the usefulness of incorporating Bayesian inference into the continual meta-learning of quality estimation models.

Method	en-lv		en-cs	
	Pearson \uparrow	MAE \downarrow	Pearson \uparrow	MAE \downarrow
CQE-OBML	53.67	0.1596	51.19	0.1619
- LR (α)	53.48	0.1598	50.94	0.1623
- SVGD	52.86	0.1621	50.56	0.1631

Table 3: Ablation of model components.

5 Conclusions

We proposed a framework for the continual meta-learning of machine translation quality estimation models, which is able to learn continually from the streaming data of multiple quality estimation users. We further incorporate an adaptive learning rate together with online Bayesian inference for improved

performance. In experiments on quality estimation data from two language directions, we demonstrate improved performance over recent state-of-the-art continual learning methods.

Acknowledgements

This work was supported by funding from the Bergamot project (EU H2020 grant no. 825303).

References

- Rahaf Aljundi, Klaas Kelchtermans, and T. Tuytelaars. 2019. Task-free continual learning. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11246–11255.
- John Blatz, Erin Fitzgerald, George Foster, Simona Gandrabur, Cyril Goutte, Alex Kulesza, Alberto San-chis, and Nicola Ueffing. 2004. [Confidence estimation for machine translation](#). In *COLING 2004: Proceedings of the 20th International Conference on Computational Linguistics*, pages 315–321, Geneva, Switzerland. COLING.
- Cristian Buciluă, Rich Caruana, and Alexandru Niculescu-Mizil. 2006. Model compression. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 535–541.
- Arslan Chaudhry, Marc’Aurelio Ranzato, Marcus Rohrbach, and Mohamed Elhoseiny. 2019. Efficient lifelong learning with A-GEM. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186. Association for Computational Linguistics.
- Chrisantha Fernando, Dylan Banarse, Charles Blundell, Yori Zwols, David Ha, Andrei A. Rusu, Alexander Pritzel, and Daan Wierstra. 2017. Pathnet: Evolution channels gradient descent in super neural networks. *CoRR*, abs/1701.08734.
- Chelsea Finn, Pieter Abbeel, and Sergey Levine. 2017. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017*, volume 70 of *Proceedings of Machine Learning Research*, pages 1126–1135. PMLR.

- Chelsea Finn, Kelvin Xu, and S. Levine. 2018. Probabilistic model-agnostic meta-learning. In *Advances In Neural Information Processing Systems*.
- Erick Fonseca, Lisa Yankovskaya, André F. T. Martins, Mark Fishel, and Christian Federmann. 2019. [Findings of the WMT 2019 shared tasks on quality estimation](#). In *Proceedings of the Fourth Conference on Machine Translation (Volume 3: Shared Task Papers, Day 2)*, pages 1–10, Florence, Italy. Association for Computational Linguistics.
- G. Gupta, Karmesh Yadav, and Liam Paull. 2020. Look-ahead meta learning for continual learning. In *Advances In Neural Information Processing Systems (NeurIPS)*.
- Tuomas Haarnoja, Haoran Tang, Pieter Abbeel, and Sergey Levine. 2017. Reinforcement learning with deep energy-based policies. In *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017*, volume 70 of *Proceedings of Machine Learning Research*, pages 1352–1361. PMLR.
- Geoffrey E. Hinton, Oriol Vinyals, and Jeffrey Dean. 2015. Distilling the knowledge in a neural network. *CoRR*, abs/1503.02531.
- Nithin Holla, Pushkar Mishra, H. Yannakoudakis, and Ekaterina Shutova. 2020. Meta-learning with sparse experience replay for lifelong language learning. *ArXiv*, abs/2009.04891.
- Khurram Javed and Martha White. 2019. [Meta-learning representations for continual learning](#). In H. Wallach, H. Larochelle, A. Beygelzimer, F. de Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 1820–1830. Curran Associates, Inc.
- Fabio Kepler, Jonay Trénous, Marcos Treviso, Miguel Vera, and André F. T. Martins. 2019. [OpenKiwi: An open source framework for quality estimation](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 117–122, Florence, Italy. Association for Computational Linguistics.
- Hyun Kim, Jong-Hyeok Lee, and Seung-Hoon Na. 2017. Predictor-estimator using multilevel task learning with stack propagation for neural quality estimation. In *Proceedings of the Second Conference on Machine Translation, WMT 2017, Copenhagen, Denmark, September 7-8, 2017*, pages 562–568. Association for Computational Linguistics.
- Taesup Kim, Jaesik Yoon, O. Dia, S. Kim, Yoshua Bengio, and Sungjin Ahn. 2018. Bayesian model-agnostic meta-learning. In *Advances In Neural Information Processing Systems*.
- James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, and Others. 2017. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences*, 114(13):3521–3526.
- Maarit Koponen, W. Aziz, L. Ramos, Lucia Specia, J. Rautio, M. González, L. Carlson, and C. España-Bonet. 2012. Post-editing time as a measure of cognitive effort. In *AMTA 2012 Workshop on Post-editing Technology and Practice (WPTP)*.
- Matthias De Lange, Rahaf Aljundi, Marc Masana, Sarah Parisot, Xu Jia, Ales Leonardis, Gregory G. Slabaugh, and Tinne Tuytelaars. 2019. [Continual learning: A comparative study on how to defy forgetting in classification tasks](#). *CoRR*, abs/1909.08383.
- Qiang Liu and Dilin Wang. 2016. [Stein variational gradient descent: A general purpose bayesian inference algorithm](#). In *Advances in Neural Information Processing Systems 29*, pages 2378–2386. Curran Associates, Inc.
- Yang Liu, Prajit Ramachandran, Qiang Liu, and Jian Peng. 2017. Stein variational policy gradient. In *Proceedings of the Thirty-Third Conference on Uncertainty in Artificial Intelligence, UAI 2017, Sydney, Australia, August 11-15, 2017*. AUAI Press.
- David Lopez-Paz and Marc’Aurelio Ranzato. 2017. Gradient episodic memory for continual learning. In *Advances in Neural Information Processing Systems*.
- Cuong Nguyen, Thanh-Toan Do, and Gustavo Carneiro. 2020. Uncertainty in model-agnostic meta-learning using variational inference. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 3090–3100.
- Alex Nichol and John Schulman. 2018. Reptile: a scalable metalearning algorithm. *arXiv preprint arXiv:1803.02999*, 2(2):1.
- Abiola Obamuyide and Andreas Vlachos. 2019. [Meta-learning improves lifelong relation extraction](#). In *Proceedings of the 4th Workshop on Representation Learning for NLP (RepL4NLP-2019)*, pages 224–229, Florence, Italy. Association for Computational Linguistics.
- Tharindu Ranasinghe, Constantin Orasan, and Ruslan Mitkov. 2020. [Transquest at wmt2020: Sentence-level direct assessment](#). In *Proceedings of the Fifth Conference on Machine Translation*, pages 1049–1055, Online. Association for Computational Linguistics.
- Sachin Ravi and Alex Beatson. 2019. [Amortized bayesian meta-learning](#). In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net.
- Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H Lampert. 2017. icarl: Incremental classifier and representation learning. In *Proceedings of the IEEE Conference*

- on *Computer Vision and Pattern Recognition*, pages 2001–2010.
- Matthew Riemer, Ignacio Cases, Robert Ajemian, Miao Liu, Irina Rish, Yuhai Tu, and Gerald Tesauro. 2019. [Learning to learn without forgetting by maximizing transfer and minimizing interference](#). In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net.
- Mark Bishop Ring. 1994. *Continual learning in reinforcement environments*. Ph.D. thesis, University of Texas at Austin Austin, Texas 78712.
- Andrei A Rusu, Neil C Rabinowitz, Guillaume Desjardins, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, and Raia Hadsell. 2016. Progressive neural networks. *arXiv preprint arXiv:1606.04671*.
- Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *ArXiv*, abs/1910.01108.
- Jurgen Schmidhuber. 1987. Evolutionary principles in self-referential learning. *On learning how to learn: The meta-meta-... hook.) Diploma thesis, Institut f. Informatik, Tech. Univ. Munich*.
- Matthew Snover, Bonnie Dorr, Richard Schwartz, Linnea Micciulla, and John Makhoul. 2006. A study of translation edit rate with targeted human annotation. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers*. Cambridge, MA.
- José Guilherme Camargo de Souza, Matteo Negri, Elisa Ricci, and Marco Turchi. 2015. [Online multitask learning for machine translation quality estimation](#). In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 1: Long Papers*, pages 219–228. The Association for Computer Linguistics.
- Lucia Specia, Frédéric Blain, Marina Fomicheva, Erick Fonseca, Vishrav Chaudhary, Francisco Guzmán, and André F. T. Martins. 2020. [Findings of the WMT 2020 shared task on quality estimation](#). In *Proceedings of the Fifth Conference on Machine Translation*, pages 743–764, Online. Association for Computational Linguistics.
- Lucia Specia, Kim Harris, Aljoscha Burchardt, Marco Turchi, Matteo Negri, and Inguna Skadina. 2017. Translation quality and productivity: A study on rich morphology languages. In *Machine Translation Summit XVI*, pages 55–71.
- Lucia Specia, Gustavo Paetzold, and Carolina Scarton. 2015. Multi-level translation quality prediction with quest++. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, System Demonstrations*, pages 115–120. The Association for Computer Linguistics.
- Lucia Specia, Marco Turchi, Nicola Cancedda, Marc Dymetman, and Nello Cristianini. 2009. Estimating the sentence-level quality of machine translation systems. In *13th Conference of the European Association for Machine Translation*, pages 28–37.
- Sebastian Thrun. 1996. Is learning the n-th thing any easier than learning the first? In *Advances in neural information processing systems*, pages 640–646.
- Sebastian Thrun and Lorien Pratt. 1998. [Learning to Learn: Introduction and Overview](#). In *Learning to Learn*, pages 3–17. Springer US, Boston, MA.
- Marco Turchi, Antonios Anastasopoulos, José G. C. de Souza, and Matteo Negri. 2014. Adaptive quality estimation for machine translation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014, June 22-27, 2014, Baltimore, MD, USA, Volume 1: Long Papers*, pages 710–720. The Association for Computer Linguistics.
- Zhenyi Wang, Yang Zhao, Ping Yu, Ruiyi Zhang, and Changyou Chen. 2020. Bayesian meta sampling for fast uncertainty adaptation. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*.
- Friedemann Zenke, Ben Poole, and Surya Ganguli. 2017. [Continual Learning Through Synaptic Intelligence](#). In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 3987–3995, International Convention Centre, Sydney, Australia.
- Jieyu Zhao and Jurgen Schmidhuber. 1996. Incremental self-improvement for life-time multi-agent reinforcement learning. In *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, Cambridge, MA, pages 516–525.
- Yinhao Zhu and Nicholas Zabarar. 2018. Bayesian deep convolutional encoder-decoder networks for surrogate modeling and uncertainty quantification. *J. Comput. Phys.*, 366:415–447.

A Additional Results

We present additional results on the **WPTP12** dataset (Koponen et al., 2012), which is a small English-Spanish (en-es) translation dataset consisting of documents from the news domain. It features translations from eight different machine translation systems. Each instance in the dataset includes the corresponding post-edited translation along with post-editing time and HTER scores computed between the translation and the corresponding post-edit. Statistics about the number of instances per post-editor are in Table 4.

Table 5 contains the results obtained on this dataset. As a result of its size, all methods generally find it challenging, with reduced performance across-the-board. Despite reduced performance in terms of mean absolute error, our approach obtains better Pearson correlation than all previous methods.

PE ID	Train	Dev	Test
A1	121	40	42
A2	121	40	42
A3	121	40	42
A4	121	40	42
A5	121	40	42
A6	121	40	42
A7	121	40	42
A8	121	40	42
Total	968	320	336

Table 4: Number of instances per Post Editor (PE) for the WPTP12 dataset.

Hyper-parameter	Value
Learning rate	3e-5
Mini-batch size	16
Max. sequence length	100

Table 6: Hyper-parameter values for all compared approaches

Method	WPTP12	
	Pearson \uparrow	MAE \downarrow
SEQUENTIAL	33.05	0.2061
A-GEM	38.95	0.2066
OML-ER	39.17	0.1786
LA-MAML	38.89	0.1772
CQUEST-OBML	40.11	0.1780

Table 5: Averaged performance for all methods.

B Additional Experimental Details

All models make use of the same values for hyper-parameters such as learning rate and batch size, selected by manual search in initial experiments. These are provided in Table 6.