

idealized user. Therefore, we recommend that further research into sequential decision-making for inline text autocomplete should *not* pursue the goal of solely increasing text entry speed, but rather aim to make real users enjoy the text entry experience.

There are several important directions for future work, to address some of the limitations of this paper. (1) It would be interesting to perform our theoretical analysis on the outputs of a real language model, to connect better with practice. (2) On the computational side, we hope to experiment with feeding the LM probability values as part of the input into the PPO and DQN agents. (3) Our user study only considered prompted writing tasks, where our software told users what sentences to write. The results may be different if users were asked to do freeform writing. (4) Most importantly, we believe that introducing more realistic conditions into our simulations would provide greater scope for RL-based agents to improve over threshold-based methods. Examples include stochasticity in the user suggestion acceptance behavior, typos in their text input, or deciding suggestion acceptance based on semantic matching with the target sentence rather than hard matching. This research would align well with other work that has found users to prefer autocomplete suggestions due to lower cognitive and physical burden, even when text entry speed is impaired (Quinn & Zhai, 2016).

References

- Tiago A Almeida, José María G Hidalgo, and Akebo Yamakami. Contributions to the study of sms spam filtering: new collection and results. In *Proceedings of the 11th ACM symposium on Document engineering*, pp. 259–262, 2011.
- Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. A brief survey of deep reinforcement learning. *arXiv preprint arXiv:1708.05866*, 2017.
- Leif Azzopardi and Guido Zuccon. An analysis of the cost and benefit of search interactions. In *Proceedings of the 2016 acm international conference on the theory of information retrieval*, pp. 59–68, 2016.
- Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016.
- Fei Cai, Maarten De Rijke, et al. A survey of query auto completion in information retrieval. *Foundations and Trends® in Information Retrieval*, 10(4):273–363, 2016.
- Lili Chen, Kevin Lu, Aravind Rajeswaran, Kimin Lee, Aditya Grover, Misha Laskin, Pieter Abbeel, Aravind Srinivas, and Igor Mordatch. Decision transformer: Reinforcement learning via sequence modeling. *Advances in neural information processing systems*, 34:15084–15097, 2021.
- CTRL-labs at Reality Labs, David Sussillo, Patrick Kaifosh, and Thomas Reardon. A generic noninvasive neuromotor interface for human-computer interaction. *bioRxiv*, pp. 2024–02, 2024.
- Nicolas Fiorini and Zhiyong Lu. Personalized neural language models for real-world query auto completion. *arXiv preprint arXiv:1804.06439*, 2018.
- Andrew Fowler, Kurt Partridge, Ciprian Chelba, Xiaojun Bi, Tom Ouyang, and Shumin Zhai. Effects of language modeling and its personalization on touchscreen typing performance. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pp. 649–658, 2015.
- Yury Gorishniy, Ivan Rubachev, and Artem Babenko. On embeddings for numerical features in tabular deep learning. *Advances in Neural Information Processing Systems*, 35:24991–25004, 2022.
- Shengyi Huang, Rousslan Fernand Dossa, Chang Ye, Jeff Braga, Dipam Chakraborty, Kinal Mehta, and João G.M. Araújo. Cleanrl: High-quality single-file implementations of deep reinforcement learning algorithms. *Journal of Machine Learning Research*, 23(274):1–18, 2022. URL <http://jmlr.org/papers/v23/21-1342.html>.

- Danyang Jiang, Wanyu Chen, Fei Cai, and Honghui Chen. Neural attentive personalization model for query auto-completion. In *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pp. 725–730. IEEE, 2018.
- Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4:237–285, 1996.
- Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1-2):99–134, 1998.
- Michael N Katehakis and Arthur F Veinott. The Multi-Armed bandit problem: Decomposition and computation. *Mathematics of OR*, 12(2):262–268, May 1987.
- Ilya Kostrikov, Ashvin Nair, and Sergey Levine. Offline reinforcement learning with implicit q-learning. *arXiv preprint arXiv:2110.06169*, 2021.
- Mina Lee, Tatsunori B Hashimoto, and Percy Liang. Learning autocomplete systems as a communication game. *arXiv preprint arXiv:1911.06964*, 2019.
- Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu. Offline reinforcement learning: Tutorial, review, and perspectives on open problems. *arXiv preprint arXiv:2005.01643*, 2020.
- Bhaskar Mitra and Nick Craswell. Query auto-completion for rare prefixes. In *Proceedings of the 24th ACM international on conference on information and knowledge management*, pp. 1755–1758, 2015.
- Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.
- Kenton Murray and David Chiang. Correcting length bias in neural machine translation. *arXiv preprint arXiv:1808.10006*, 2018.
- Fred Paas, Juhani E. Tuovinen, Huib Tabbers, and Pascal W. M. Van Gerven. Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1):63–71, January 2003. ISSN 1532-6985. doi: 10.1207/s15326985ep3801_8. URL http://dx.doi.org/10.1207/s15326985ep3801_8.
- Dae Hoon Park and Rikio Chiba. A neural language model for query auto-completion. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1189–1192, 2017.
- Martin L Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.
- Philip Quinn and Shumin Zhai. A cost-benefit study of text entry suggestion interaction. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pp. 83–88, 2016.
- Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9, 2019.
- Keith Rayner and Charles Clifton, Jr. Language processing in reading and speech perception is fast and incremental: implications for event-related potential research. *Biol. Psychol.*, 80(1):4–9, January 2009.
- Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*, 2019.
- John Schulman, Sergey Levine, Pieter Abbeel, Michael Jordan, and Philipp Moritz. Trust region policy optimization. In *International conference on machine learning*, pp. 1889–1897. PMLR, 2015.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

Richard S Sutton, David McAllester, Satinder Singh, and Yishay Mansour. Policy gradient methods for reinforcement learning with function approximation. *Advances in neural information processing systems*, 12, 1999.

Denis Tarasov, Alexander Nikulin, Dmitry Akimov, Vladislav Kurenkov, and Sergey Kolesnikov. CORL: Research-oriented deep offline reinforcement learning library. In *3rd Offline RL Workshop: Offline RL as a "Launchpad"*, 2022. URL <https://openreview.net/forum?id=SyAS49bBcv>.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Michael Volske, Martin Potthast, Shahbaz Syed, and Benno Stein. TL;DR: Mining Reddit to learn automatic summarization. In *Proceedings of the Workshop on New Frontiers in Summarization*, pp. 59–63, Copenhagen, Denmark, September 2017. Association for Computational Linguistics. doi: 10.18653/v1/W17-4508. URL <https://www.aclweb.org/anthology/W17-4508>.

Po-Wei Wang, J Zico Kolter, Vijai Mohan, and Inderjit S Dhillon. Realtime query completion via deep language models. *Proceedings of SIGIR Workshop On eCommerce (SIGIR eCom'18)*, 2018.

Sida Wang, Weiwei Guo, Huiji Gao, and Bo Long. Efficient neural query auto completion. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pp. 2797–2804, 2020.

Yingfei Wang, Hua Ouyang, Hongbo Deng, and Yi Chang. Learning online trends for interactive query auto-completion. *IEEE Transactions on Knowledge and Data Engineering*, 29(11):2442–2454, 2017.

Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8:279–292, 1992.

A Reward Derivation

Our objective is to minimize the total time it takes for the user to enter the input. Hence, the instantaneous reward will be proportional to the *saved time* due to a correct suggestion that the user accepted minus the *lost time* due to cognitive load imposed on the user. For simplicity, we assume all characters take the same amount of time to write ($\Delta t_{char-write}$) and to be read ($\Delta t_{char-read}$). We also associate an additional fixed time loss for every suggestion, where the user has to pause and move their gaze ($\Delta t_{distraction}$). Hence we will have:

$$r \propto \text{Saved Time} - \text{Lost Time} \quad (1)$$

$$\propto \text{len}(a) \times \text{accepted} \times \Delta t_{char-write} - \text{len}(a) \times \Delta t_{char-read} - \Delta t_{distraction} \times (\text{len}(a) \neq 0) \quad (2)$$

$$\propto \text{len}(a) \left(\text{accepted} - \frac{\Delta t_{char-read}}{\Delta t_{char-write}} \right) - \frac{\Delta t_{distraction}}{\Delta t_{char-write}} \times (\text{len}(a) \neq 0) \quad (3)$$

$$\propto \text{len}(a)(\text{accepted} - \alpha) - \beta \times (\text{len}(a) \neq 0), \quad (4)$$