Mixture of Readers (MoRe): Enabling Scanpath Generation for Different Reader Groups

Anna Bondar[1*], David R. Reich[1,2], Lena A. Jäger[1]

[1]Department of Computational Linguistics, University of Zurich
[2]Department of Computer Science, University of Potsdam
*contact: anna.bondar@uzh.ch

Eye movements in reading provide a rich behavioural signal that reflects moment-to-moment cognitive processing. Because reading behaviour depends not only on the text, but also on reader characteristics such as language background, reading proficiency, and cognitive profile, scanpaths vary systematically across reader groups. At the same time, current generative models of eye movements are often trained on heterogeneous populations with a single shared set of parameters. While such training can improve overall performance, it may also encourage the model to prioritise what is common across groups and to underrepresent patterns that are specific to individual populations, effectively producing an averaged scanpath. Modelling group-specific scanpaths is important both for basic research and for downstream applications: preserving population-specific variation can improve tasks that aim to infer reader characteristics from gaze data and can make synthetic scanpaths more useful as cognitively informative signals for NLP. In this work, we present Mixture of Readers (MoRe), an extension of the state-of-the-art scanpath generation model Eyettention (Deng *et al.*, 2023) for group-specific scanpath generation. For this we leverage a Mixture-of-Experts architecture (Shazeer *et al.*, 2017), originally proposed as a way of introducing sparse conditional computation into neural networks by replacing fully shared dense processing with a set of specialised experts, only some of which are activated for each input. We evaluate the model on CopCo (Hollenstein *et al.*, 2022), a Danish eye-tracking corpus comprising 58 readers and 2,032 sentence-level instances from three populations: L1 non-dyslexic readers, L1 readers with dyslexia, and L2 non-dyslexic readers. MoRe introduces reader-group-conditioned routing into different scanpath-processing components of Eyettention. In this way, rather than processing scanpaths from all reader groups with a single shared set of parameters, MoRe allows the model to select different weights depending on the reader group or on characteristics of the scanpath. We experiment with introducing mixture-of-experts routing into the computation of the cross-attention query and into the decoder, that is, the prediction head responsible for generating the next fixation. We also investigate full routing, in which both components are routed. Furthermore, we compare several routing strategies: hard routing, where the expert is selected deterministically based on the reader-group label, and soft routing, where a learned gate assigns weights over experts, the input is processed by all experts, and the final output is computed as a weighted combination of their outputs. In addition, we examine the use of a shared expert, which is activated independently of the routed group and is intended to capture patterns common across groups, while allowing the group-specific experts to focus on differences in reading behaviour. We conduct experiments under both new-sentence and new-reader data splits, considering training settings that include all reader populations, L1 non-dyslexic and dyslexic readers, and L1 and L2 non-dyslexic readers. MoRe consistently improves over the Eyettention baseline across all data splits. The strongest performance is achieved by configurations using either soft decoder routing or soft full routing with shared experts. These findings suggest that group-conditioned routing is particularly beneficial when the model is given the flexibility to select experts based on scanpath characteristics. Overall, MoRe provides a way to exploit shared regularities in reading behaviour across groups while still generating more population-sensitive synthetic scanpaths and achieving stronger overall performance.

References

Deng S., Reich D. R., Prasse P., Haller P., Scheffer T., Jäger L. A. (2023). Eyettention: An attention-based dual-sequence model for predicting human scanpaths during reading. *Proc. ACM Hum.-Comput. Interact. 7*.

Hollenstein N., Barrett M., Björnsdóttir M. (2022). The Copenhagen corpus of eye tracking recordings from natural reading of Danish texts. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, Marseille, France, pp. 1712–1720.

Shazeer N., Mirhoseini A., Maziarz K., Davis A., Le Q., Hinton G., Dean J. (2017). Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. In *International Conference on Learning Representations*.