Key Findings

▷ OT-based ensemble models outperform single methods

▷ OT-based methods reach better average performance than the Factorized Attention model

▷ the Factorized Attention model reaches the best score on the Hard set

▷ **feature-free** OT-based methods are very close to the feature-informed approaches on the Hard set

▷ poor generalization on a dataset from a specialized domain (StackExchange) (best **ensemble** model: 2×Unbalanced OT + Multi-OT (5 epochs) + Multi-OT (7 epochs) achieved only **0.4211**)

▷ the methods outperform naïve baselines (≈ 0.45) but do not reach state-of-the-art results (0.863 macro F1 score) from the similar 2024 PAN shared task

▷ **future research directions:** computational diachronic semantics, extension to paragraph-based authorship detection, feature-free ensemble models