https://nrat.ukrintei.ua/longeval-dlya-oczinky-uzagalnen-vid-shi/

LONGEVAL ДЛЯ ОЦІНКИ УЗАГАЛЬНЕНЬ ВІД ШІ

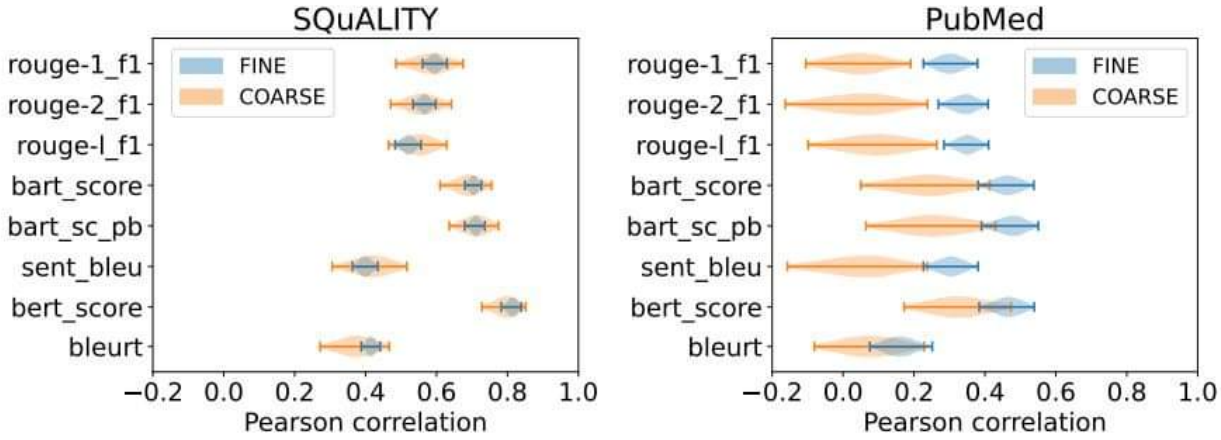
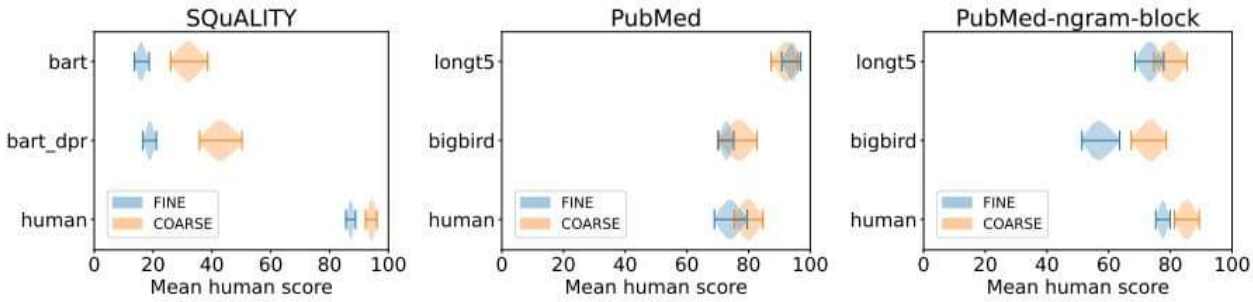


Figure 2: 95% confidence intervals of Pearson correlations between various automatic evaluation metrics and using human evaluation data collected with FINE (blue) and COARSE (orange) annotation methods. In both datasets, FINE annotations lead to much narrower CIs than COARSE annotations. See Appendix G for plot with Kendall's Tau.



На порталі Arxiv опублікована стаття групи дослідників «LongEval: рекомендації щодо людської оцінки узагальнень».
У матеріалі йдеться про те, як оцінка, що здійснюється людиною (найкраща практика для точної оцінки достовірності автоматично згенерованих зведень/резюме) та інші рішення дозволяють впоратися зі складністю та навантаженням, які супроводжують процес оцінювання. Автори вивчили 162 статті, присвячені підготовці переказу у вигляді короткого резюме, узагальнили поточну практику людського оцінювання зведеної інформації, згенерованої ШІ. Вони використовували LongEval – набір посібників з оцінювання людиною точності докладних резюме, які вирішують такі проблеми: забезпечення достовірності; мінімізація навантаження на людину за збереження високої точності достовірності; забезпечення цінності автоматичного резюме. Застосування LongEval у дослідженнях анотацій двох наборів даних результатів узагальнення у різних галузях знань (SQuALITY та PubMed) дозволило виявити, що перехід до більш тонкої деталізації суджень знижує дисперсію результатів.
Детальніше: https://arxiv.org/pdf/2301.13298.pdf, https://arxiv.org/abs/2301.13298, https://is.gd/7TCIQN, https://doi.org/10.48550/arXiv.2301.13298

Фото: скріншот
#НРАТ_Усі_новини #НРАТ_Науковцям_новини #НРАТ_Освітянам_новини
2023-08-09

Інформація з офіційного вебпорталу Національного репозитарію академічних текстів