

I've tried using KL Divergence for learning to rank (see here: <https://theiconic.tech/learning-to-rank-is-good-for-your-ml-career-part-2-lets-implement-listnet-11af69d1704>). Ended up getting slightly worse results than just regular MSE so I didn't explore it too much. I might come back to it eventually, seemed cool at the time