Thesis Topics: Legal Data Science

Daniel Saggau

LMU

2022

Discussion Topics

1. Topics Lit. Review
2. Eyecite Tool
3. Supervisors and next steps
4. Phd Applications
5. Tokyo

Overview Models for Legal Text

- Multi-Task Pre- Training **(transfer learning)**
- Efficient Transformers **(knowledge distillation)**
- Interpretable Knowledge distillation

Initial Thoughts

- Most added value from multi task training with approx. 100+ tasks
- Feasibility concern for thesis !
- Interesting feasible problems: how do we ensure that we keep the good (transferable) elements and reduce the bad learned elements ?
- What do we actually attend to in the mechanisms: in depth analysis of what we do in the attention mechanism for different models
- how does multi-task learning fare if we only train for specific tasks that are related to current challenges in legal tech vs broader multi-task learning
- another idea is the combination of an in depth analysis of the attention mechanism and looking deeper into the student teacher relationship

Supervisors

- Contacted Mina Rezaie as 1. Supervisor (meeting end of January) (Deep Learning)
- Contacted Elliot Ash at ETH as interesting external supervisor(Applications Legal Tech and Econ) (pending)
- Potentially can contact Christoph Kern as 2. Supervisor if needed within our department (Statistics for Social Science)
- Considering applying to Christof Monz at UVA in Amsterdam (Machine Translation)

Literature List

- Jiao, X., Yin, Y., Shang, L., Jiang, X., Chen, X., Li, L., ... & Liu, Q. (2019). Tinybert: Distilling bert for natural language understanding. arXiv preprint arXiv:1909.10351.
- Gou, J., Yu, B., Maybank, S. J., Tao, D. (2021). Knowledge distillation: A survey. International Journal of Computer Vision, 129(6), 1789-1819.
- Tay, Y., Dehghani, M., Bahri, D., & Metzler, D. (2020). Efficient transformers: A survey. arXiv preprint arXiv:2009.06732.
- Sanh, V., Webson, A., Raffel, C., Bach, S. H., Sutawika, L., Alyafeai, Z., ... Rush, A. M. (2021). Multitask prompted training enables zero-shot task generalization. arXiv preprint arXiv:2110.08207.
- Ghader, H., & Monz, C. (2017). What does attention in neural machine translation pay attention to?. arXiv preprint arXiv:1710.03348.
- Ash, E., Chen, D. L., & Galletta, S. (2021). Measuring Judicial Sentiment: Methods and Application to US Circuit Courts. Economica.
- Gu, N., Ash, E., & Hahnloser, R. H. (2021). MemSum: Extractive Summarization of Long Documents using Multi-step Episodic Markov Decision Processes. arXiv preprint arXiv:2107.08929. Chicago
- Chalkidis, I., Fergadiotis, M., & Androutsopoulos, I. (2021). MultiEURLEX–A multi-lingual and multi-label legal document classification dataset for zero-shot cross-lingual transfer. arXiv preprint arXiv:2109.00904.
- Crawshaw, M. (2020). Multi-task learning with deep neural networks: A survey. arXiv preprint arXiv:2009.09796.

Tokyo

- Application deadline?
- Funding
- Timing