

Extrakcia právnych odkazov pomocou modelov transformérov

Vedúci práce: RNDr. Peter Gurský, PhD.

Konzultant: RNDr. Dávid Varga

Študent: Bc. Nicol Fedurcová



Motivácia

- Aktuálna tématika modelov transformérov
- Práca na datase Slovenských súdnych rozhodnutí
- Využitelnosť výsledku práce v súdnej vede
- Aplikácia metódy v systéme na vyhľadávanie súdnych rozhodnutí

Ciele práce

1. Vypracovať prehľad metód extrakcie z textov s využitím transformérov a prehľad transformer-modelov vhodných pre slovenské právne texty.

- Špecifická štruktúra právnych textov – písané prirodzeným jazykom
- Jazykové modely Mistral, Slovak-Bert

Ciele práce

2. Identifikovať a analyzovať dôvody neúspešných extrakcií v existujúcom pravidlovom systéme extrahujúcom odkazy na súdne rozhodnutia a právne predpisy.

- Vyplýva z bakalárskej práce Martina Gojdiča
„Automatizovaná extrakcia odkazov na právne predpisy zo súdnych rozhodnutí“

Ciele práce

3. Navrhnuť a implementovať rozšírenie existujúceho riešenia s využitím transformer modelu.

Ciele práce

4. Porovnať úspešnosť novej metódy s existujúcim pravidlovým systémom na rozšírenom anotovanom datasete právnych textov.

- Súčasná anotovaná sada má cca 40 rozhodnutí
- Pravidlový systém bol aplikovaný na cca 4 000 000 rozhodnutí

Prvé kroky

- Zorientovanie sa v problematike právnych textov
 - => Štuktúra rozhodnutí, štruktúra právnych predpisov
 - => Systémy zápisov odkazov - odkazy na právne predpisy nemajú jednotnú formu
 - => Oboznámenie sa s existujúcim pravidlovým systémom
- Vytvorenie prehľadu modelov tranformerov
- Zoznámenie sa s anotovanými sadami, prispôsobenie anotácie na vstup do modelu transformera

Literatúra

1. Varga, Dávid, et al. Extraction of Legal References from Court Decisions. Vol. 3498, 2023, pp. 89–95,
2. Darji, Harshil, Jelena Mitrović, and Michael Granitzer. "A Dataset of German Legal Reference Annotations." Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law. 2023.
3. Jurafsky, Daniel, and James H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models. 3rd ed., 2024.
4. Agnoloni, Tommaso, et al. "Linking european case law: BO-ECLI parser, an open framework for the automatic extraction of legal links." Legal Knowledge and Information Systems. IOS Press, 2017. 113-118.