

Průběh obhajoby – A20B0373P KAŠÁKOVÁ Tereza

Automatické rozpoznávání typů otázek v různých jazycích.

Intonace otázek.

Ukázka dat.

Pravidla, tázací zájmena, konfigurační soubor.

Automatická pravidla.

SVC, T5.

Výsledky.

Nejčastější chyby. Eliptické otázky.

Ve shrnutí výsledků se objevuje pojem "Tf-idf transformace". Můžete objasnit, co zkratka "Tf-idf" znamená a o co se jedná?

Uvádíte, že pro ruská data neby ručně vytvořen seznam tázacích zájmen, a proto ruský jazyk nejsou k dispozici výsledky s ručně odvozenými pravidly. Zkoušela jste tento seznam vytvořit automatickým překladem z jiného jazyka?

Neuvažovala jste použít embedding?

K vyhodnocení úspěšnosti testovaných metod pro rozlišení typů otázek používáte míru accuracy. Co byste použila v případě, že byste řešila klasifikaci do dvou tříd, kdy jedna z nich je mnohem častější než druhá?

Binární klasifikace. Co se stane s vylučovacími otázkami?

Kolik jste měla dat pro přetrénování modelu?

K čemu je ve vaší úloze dekodér?

Signál v práci, jaké jsou jednotky na x-ové ose a y-ové ose obrázku?

Klasifikace – výborně

Datum obhajoby – 30.8.2023