

ještě promyslete cíl a výsledek, od zvoleného cíle je ke zvolenému výsledku poměrně daleko

Předkládaná bakalářská práce se zabývá hláskovými změnami v historické češtině. Cílem je vytvořit seznam regulárních výrazů, který je založený na změnách od 15. století po současnost. V práci se pracuje s ručně vybranou stovkou nejčastěji vyskytovaných podstatných jmen a sloves z korpusu Staročeská textová banka. Regulární výrazy jsou vytvořeny podle výukových materiálů z předmětu PLIN007. Webová aplikace je vytvořena v javascriptovém frameworku Angular. Výsledkem práce je webová aplikace, která prostřednictvím slideru vizualizuje a demonstruje hláskové změny. Výsledná aplikace je určena především pro studenty českého jazyka.

lepší uvést až do práce, ne do anotace

ANOTACE: pozor na sazbu

Cílem předkládané bakalářské práce je provést analýzu hodnocení u produktů na webovém e-shopu

Heureka.cz. Hlavním zaměřením je identifikace recenzí s podobným významem, a to metodou jazykového modelu word2vec. K dosažení vyšší přesnosti a eliminaci potenciálních problémů jako jsou např. překlepy či nerelevantní informace je do analýzy zahrnut i velký jazykový model BERT.

Výsledkem této bakalářské práce je seznam obsahující extrakci klíčových slov, který poskytne uživatelům snadnější přehled o tom, v čem daný produkt vyniká, a to jak v pozitivních, tak záporných aspektech.

Anotace je příliš krátká a po přečtení vyvstávají spíše otázky, než abychom odpovědi dostali. Zkuste doplnit část o metodě

Anotace

Nedá! V odborné texty pouze lze nebo je možné :-)

Cílem bakalářské práce bylo dotrénovat mezijazykový model pro rozpoznávání řeči a následnou transkripci na fonémy. Model byl trénován na zvukových nahrávkách nerodilých mluvčích češtiny. Při fine-tuningu se proces několikrát opakoval, aby došlo k co nejpřesnější shodě se zvukovou nahrávkou. K dotrénování byla využita technika CTC (konekcionistická temporální klasifikace), která se uplatňuje při rozpoznávání řeči nebo překladu mezi jazyky. Tato technika byla vybrána, protože umožňuje učení jazykových modelů v případě, kdy má výsledný text odlišnou délku od vstupu, což je významné pro rozpoznávání řeči. Výsledkem práce je natrénovaný model, který dokáže přepsat mluvenou řeč nerodilých mluvčích češtiny na fonémy. Tento model se dá využít v různých aplikacích určených k výuce českého jazyka, nebo kde je potřebná automatická transkripce.

užití různých mluvinických časů lze odůvodnit

Bakalářská práce se zabývá generováním českých textů splňujících určité formální i obsahové požadavky pomocí generativního předtrénovaného transformátoru. Jejím cílem je vytvořit nástroj pro generování pravopisných diktátů v češtině. Podle zvoleného tématu a procvičovaných pravopisných jevů se algoritmem na bázi word embeddings modelu vyberou slova na procvičení významově a kontextově nejbližší danému tématu. Tato slova spolu s tématem se pak pošlou jako vstup ChatuGPT, který díky vhodně zvolenému dotazu vygeneruje několik možných diktátů. Ty se následně překontrolují, jestli splňují vstupní požadavky a jsou smysluplným textem, a vybere se z nich jeden nejvhodnější diktát. Výsledkem práce je webová aplikace, kde si uživatel zvolí vlastní parametry diktátu (obtížnost a délka textu, téma, procvičované jevy) a dostane diktát splňující tyto parametry. Nástroj může urychlit a usnadnit přípravu na hodinu vyučujícím češtiny.

Do zpětné vazby jsem napsala, že jsou souvětí dlouhá, když jsem je vybarvila, zjistila jsem, že nejsou, ale přesto se to špatně čte.

Tato bakalářská diplomová práce si klade za cíl navrhnout způsob, jakým je možné zakomponovat funkci podobnostního vyhledávání do korpusu dat. Specificky se jedná o řešení pro korpus chemických dat, a to z důvodu potenciálně nejvíce relevantního užití v tomto oboru, nicméně daná funkce je významná také pro čistě lingvistické účely.

Konkrétně se práce zaměřuje na Corpus Query Language (CQL) a možnosti jeho rozšíření pro potřeby cíle práce tak, aby bylo místo exaktního dotazu možné využít také přibližného dotazování. Výsledkem tedy bude návrh implementace funkce podobnostního vyhledávání v korpusu chemických dat. **Používejte jeden mluvnický čas, nepoužívejte budoucí čas, když někdo bude anotaci číst, všechno už bylo.**

není to projekt

Bakalářská práce se zaměřuje na analýzu sentimentu českých titulků různých žánrů filmů. V průběhu projektu je implementováno několik fází, přičemž první zahrnuje pečlivý sběr dat. Tato fáze spočívá ve stažení českých titulků filmů různých žánrů. Následně jsou tyto titulky podrobeny úpravám za využití vlastního skriptu, jehož funkcionalita zahrnuje odstranění nežádoucích znaků, korekci chybných mezer a přiřazení každé věty vlastnímu řádku. Současně jsou data obohacena o relevantní informace, jako jsou například žánr filmu, jméno režiséra nebo rok uvedení filmu. Analýza sentimentu, která chronologicky následuje, se zaměřuje na zkoumání emocionálního kontextu v jednotlivých žánrech. Pro tento účel jsou využity specializované nástroje pro sentimentální analýzu textu. **Hlavním cílem bakalářské práce je identifikovat potenciální variace v projevu sentimentu mezi různými filmovými žánry prostřednictvím českých titulků. Výsledky této analýzy poskytují cenné poznatky o způsobu, jakým se sentiment projevuje v českých titulcích s ohledem na konkrétní žánry filmů. Bakalářská práce přináší komplexní pohled na problematiku sentimentu v českých titulcích a může sloužit jako základ pro další výzkum v oblasti jazykové analýzy filmů.** **Pokud zmiňujete hlavní cíl, předpokládá se, že máte i vedlejší.**

Já bych možná řekla, že spíše než pro jazykovou analýzu filmů to bude dobré pro analýzu sentimentu.

Anotace:

Asi cíl?

This bachelor thesis examines interpretations of plurality in the Czech language. In particular, it focuses on sentences containing conjunctions coordinating noun phrases (so-called non-Boolean conjunction). There are two possible interpretations: the cumulative, wherein both conjunct nouns are understood to have performed an action together, and the distributive, wherein each conjunct noun is understood as performing the same action as the other, but separately. First, all relevant sentences are listed in their most narrow forms together with the author's intuition as to their likelier interpretation. Afterwards, corpus queries are formulated which find occurrences of these sentences with or without other syntactic elements. In the end, the data acquired using these queries is generalized into categories by the types of surrounding elements and filtered through the lens of language intuition and formal semantics. **This way the thesis aims to reveal how syntactic position and other elements affect the resulting interpretation** of sentences that contain non-Boolean conjunction.

Taky cíl?

Co je potom výsledkem práce?