



POSUDEK VEDOUCÍHO DIPLOMOVÉ PRÁCE

Jméno studenta: Bc. Václav Tomíček
Název práce: Koncept využití Umělé inteligence k rozpoznání patternů a logických celků textu
Autor posudku: Ing. Pavel Blažek, Ph.D.
Cíl práce: Cílem diplomové práce je seznámit se s koncepty optického rozpoznávání znaků založenými na strojovém učení a navrhnout postup, jak tyto koncepty využít k efektivní a cenově nejméně náročné migraci dokumentů z grafické do znakové podoby.

| Povinná kritéria hodnocení práce | Stupeň hodnocení (známka) | | | | | |
|--|-------------------------------------|-------------------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | A | B | C | D | E | F |
| Práce svým zaměřením odpovídá studovanému oboru | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Vymezení cíle a jeho naplnění | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Zpracování teoretických aspektů tématu | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Zpracování praktických aspektů tématu | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Adekvátnost použitých metod, způsob jejich použití | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Hloubka a správnost provedené analýzy | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Práce s literaturou | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Logická stavba a členění práce | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Jazyková a terminologická úroveň | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Formální úprava a náležitosti práce | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Vlastní přínos studenta | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Využitelnost výsledků práce v teorii (v praxi) | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Vyjádření k výsledku anti-plagiátorské kontroly

Při kontrole nebyla nalezena shoda s jinými dokumenty.

Dílní připomínky a náměty:

Na základě dobře postavené osnovy je sepsána práce, jejíž sloh dovoluje plynulé čtení a snadné porozumění textu. Občas lze najít věty, kde se opakují výrazy v míře vyšší než by bylo nutné. Zkratky jsou vysvětlené, je API ne při prvním výskytu v textu. V úvodních teoretických kapitolách by bylo vhodné mít více citovaných zdrojů. Obrázky oken v praktické části jsou místy hůře čitelné, bylo by vhodné dávat spíš jejich výřezy, které by neměly problematickou velikost písma.

Celkové posouzení práce a zdůvodnění výsledné známky:

Práce vychází z praxe, byla vypracována podle zadání společnosti Quadient. Zabývá se pokročilou digitalizací velkých objemů dokumentů, které jsou uloženy v grafickém formátu, s důrazem na optimalizaci procesu. V úvodní části jsou zmíněny různé metody rozpoznávání textu a znaků, strojového učení včetně deep learningu a využití neuronových sítí. Navazuje popis produktů výše

uvedené společnosti a nástrojů a služeb třetích stran, které efektivně zpracovávají text, a které jsou využité dále v praktické části. V té jsou v úvodu nástroje porovnány a otestovány na testovacích dokumentech. Výsledkem pak je jejich integrace do Quadiant řešení a ověření funkčnosti. Student v implementační části popisuje proces od návrhu po hodnocení úspěšnosti načtení bloků a rychlosti provedení operace. Koncept sice používá nové přístupy, přesto není výstup 100%, což bude potřeba řešit v aktualizacích a nových verzích. Ze závěru vyplývá, že cíl práce byl splněn, a že vzniklo cenově přijatelné a tím pádem konkurence schopné řešení.

Otázky k obhajobě:

Jakou úspěšnost převodu na text měly dokumenty obsahující tabulku? Mělo by smysl testované externí nástroje ve finálním produktu kombinovat?

Práci doporučuji k obhajobě.

Navržená výsledná známka: B

V Hradci Králové, dne 6. září 2022

podpis