

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce: Václav Stibor

Název práce: Financial News Sentiment Analysis

Rok odevzdání: 2024

Studijní program: Informatika

Specializace: Databáze a web

Autor posudku: Ing. Pavel Koupil, Ph.D.

Role: Oponent

Pracoviště: Katedra softwarového inženýrství, MFF UK

Text posudku:

| K celé práci | lepší | obvyklé | horší | nevyhovuje |
|------------------|-------------------------------------|-------------------------------------|--------------------------|--------------------------|
| Obtížnost zadání | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Splnění zadání | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Rozsah práce | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Zadání práce považuji za splněné. I přes drobné nedostatky (viz dále) se jedná o velmi kvalitní práci, a to zvláště po technické stránce. Oceňuji zejména velmi podrobnou diskusi problému a implementační detaily v příloze B.

| Textová část práce | lepší | obvyklé | horší | nevyhovuje |
|----------------------------|-------------------------------------|-------------------------------------|-------------------------------------|--------------------------|
| Formální úprava | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Struktura textu | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Analýza | <input type="checkbox"/> | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| Programátorská dokumentace | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Uživatelská dokumentace | <input type="checkbox"/> | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

Text práce je psán velmi čtivou angličtinou s minimem překlepů. Jednotlivé kapitoly na sebe logicky navazují. Velká část práce je věnována velmi detailnímu popisu problému a diskusi jednotlivých částí řešení, např. různým způsobům získání dat.

Sazba je na dobré úrovni, v práci jsem našel pouze drobné nedostatky (slug na konci nadpisu na str. 117 a obrázky nejsou vektorové).

Práce s (odbornou) literaturou je adekvátní. Pouze bych doporučil odkázat se např. pomocí poznámky pod čarou na dokumentaci u použitých technologií a upřesnil bych, jaké verze technologií jsou použity (např. z textu nelze rozlišit, zda je řešení postaveno na programovacím jazyku Python verzi 2 nebo 3 – lze pouze intuitivně předpokládat). Dále není jasný původ grafů (např. obrázky 1, 3.1 a další). Jedná se o vlastní, převzaté obrázky nebo screenshoty aplikace?

Z analýzy jsou uvedeny pouze funkční a nefunkční požadavky. Chybí např. analýza uživatelských rolí nebo diagramy případů užití. Vzhledem k tomu, že grafické rozhraní je stěžejní součástí této aplikace, očekával bych také alespoň jeden mockup nebo storyboard.

Architektura řešení je ilustrována pomocí diagramu komponent a detailně popsána. Backend aplikace je vhodně rozdělen do jednotlivých modulů. Autor práce využil Apache Airflow pro automatizaci sběru dat z různých (webových) zdrojů a rozhraní. Nesouhlasím pouze s použitím XCom pro přenos souborů o velikosti 1-2 MB, což je v rozporu s best practices (str. 70). Součástí kapitoly je ještě popis services a REST rozhraní.

Data jsou ukládána v systému PostgreSQL a datový model je vizualizován pomocí (proprietárního) diagramu (obrázek 7.3), který se ale nezobrazuje správně. (Jako náhrada diagramu tříd?)

Programátorská dokumentace neobsahuje sekvenční diagramy, nicméně ty jsou částečně nahrazeny diagramy workflow z Apache Airflow. V textu jsem také nenašel seznam použitých technologií (na jednom místě).

Uživatelská dokumentace předpokládá, že je aplikace již nasazena. Nejsou v ní tak uvedeny minimální HW a SW prostředky pro spuštění aplikace a její nasazení. Dále se jedná spíše o tutoriál ilustrující jeden ze scénářů použití aplikace. O klasickou uživatelskou dokumentaci se nejedná.

| Implementační část práce | lepší | obvyklé | horší | nevyhovuje |
|--------------------------|-------------------------------------|-------------------------------------|--------------------------|--------------------------|
| Kvalita návrhu | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Kvalita zpracování | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Stabilita implementace | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Návrh aplikace včetně použitých technologií je vhodný. Zdrojový kód je kvalitní, vhodně členěný a obsahuje dokumentaci. Součástí odevzdané přílohy je také vygenerovaná programátorská dokumentace.

Nenalezl jsem ovšem jednotkové ani integrační testy, a tedy nelze ověřit stabilitu implementace.

Otázky:

1. Je tato aplikace multiplatformní? Pokud ano, jak je toho dosaženo?
2. Jak je aplikace konfigurovatelná? Pokud ano, jakým způsobem a co lze konfigurovat a co je naopak pevně nastaveno?
3. Jaké jsou uživatelské role?
4. Má aplikace administrátorské rozhraní pro její správu?
5. Jakým způsobem aplikace reaguje na změny API nebo struktury dat z webových služeb?

Celkové hodnocení Výborně

Práci navrhuji na zvláštní ocenění Ne

21. srpna 2024

Podpis: