



Posudek oponenta závěrečné práce

Oponent práce: Mgr. Petr Novák, Ph.D.
Student: Marek Jirman
Název práce: Vliv počasí na spotřebu tankového piva
Obor / specializace: Umělá inteligence 2021
Vytvořeno dne: 9. června 2024

Hodnotící kritéria

1. Splnění zadání

- [1] zadání splněno
- ▶ [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Zadání specifikovalo zaměřit se na týdenní a měsíční předpověď počasí, v práci se zkoumá čtrnáctidenní.

Spolehlivost finálních modelů je ověřena jen zběžně a chybí podrobnější interpretace výsledků.

2. Písemná část práce

65 / 100 (D)

Text je psán srozumitelnou češtinou, problematika je dobře představena. Členění práce není vždy přehledné, název kapitoly "Teoretická část" není příliš informativní. Naopak kapitola "Použité technologie", obsahující jen výčet použitých knihoven, by mohla být součástí buďto metodologie nebo experimentů. Chybí informace o použité verzi jazyka Python.

Problematika předpovědi počasí je představena srozumitelně. Rešerše dostupných zdrojů by mohla být obsažnější včetně zhodnocení a srovnání spolehlivosti. Je zvláštní, že pro experimenty a modelování je použit zdroj C3S CDS, který v rešerši není zmíněn.

V metodické části chybí diskuse vhodnosti a použitelnosti jednotlivých metod na zkoumaný problém. Není vysvětlené, proč se kromě metod pro spojitá data uvažují i klasifikátory a metody pro detekci anomálií.

Metodika experimentů není vždy úplně šťastně volena. Například detekce odlehklých pozorování v částech 5.4.2 a 5.4.3 je prováděna strojově pomocí Isolation Forest, přičemž není zmíněno, kolik hodnot bylo odstraněno - z grafů se může zdát, že odstraněných

hodnot bylo více a mohlo dojít k vychýlení dat. Data o počasí jsou k dispozici hodinově, ale zkoumaná proměnná je výtoč za minutu.

Není jasné, v jakých jednotkách je měřen vliv příznaků v grafech v sekci 6. U regresních metod by bylo dobré vliv příznaků přímo kvantifikovat uvedením odhadnutých regresních koeficientů a diskutovat jejich významnost. Experimentální analýza by mohla být obsažnější - daly by se uvažovat interakce mezi příznaky, zajímavé by také bylo prozkoumat, jaká část výtoče je ovlivněna počasím a jaká ostatními faktory. Klíčový faktor velikosti jednotlivých hospod není diskutován. Kategorizace vysvětlované proměnné na nižší, střední a vysokou výtoč a následné využití klasifikačních metod působí arbitrárně. Odvolává se na heterogenitu dat v histogramu 5.5, který ale vypadá relativně souvisle. Očekával bych spíše využití pokročilejších regresních metod nebo modelů časových řad.

Poslední část, kdy se predikce testuje proti reálným datům, je velice zajímavá, ale poněkud krátká. Ocenil bych hlubší diskusi výsledků.

3. Nepísemná část, přílohy

80/100 (B)

Pomocí knihoven dostupných v Pythonu bylo natrénováno a porovnáno několik metod strojového učení pro predikci výtoče piva na základě informací o počasí. Podařilo se propojit data z meteorologických serverů s údaji o výtoči poskytnutými z praxe. Jupyterovské notebooky pro předzpracování dat, experimenty a predikce jsou přehledně strukturované a dobře komentované.

4. Hodnocení výsledků, jejich využitelnost

75/100 (C)

Po hlubší analýze výsledků modelů s důrazem na interpretaci vlivu počasí by mohly být metody nasazeny v praxi.

Celkové hodnocení

70/100 (C)

Student nastudoval metody předpovědi počasí a na základě dat poskytnutých z praxe úspěšně natrénoval a porovnal modely strojového učení pro odhad výtoče piva. Práce by mohla být přehledněji členěna, metodika není vždy volena zcela adekvátně a výsledky by zasloužily hlubší diskusi a interpretaci. Přes uvedené nedostatky se domnívám, že dílo splňuje podmínky kladené na bakalářskou práci a doporučuji ji za ní uznat.

Otázky k obhajobě

Kolik hodnot (absolutně i relativně) bylo odstraněno v rámci předzpracování v sekcích 5.4.2 a 5.4.3?

V jakých jednotkách je měřen vliv příznaků v grafech v sekci 6?

Instrukce

Splnění zadání

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.