



POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Jméno studenta: Martin Skalický

Název práce: Statistický model kvality vody ve vodní nádrži

Autor posudku: Hana Skalská

Cíl práce: Data mining dat měřených ve vodní nádrži. Sestavení modelů pro predikování kyslíku, pH a teploty vody rozšířené o termoklinu ve vodní nádrži Vrchlice.

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 checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování teoretických aspektů tématu	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování praktických aspektů tématu	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" 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 type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková a terminologická úroveň	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Formální úprava a náležitosti práce	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vlastní přínos studenta	<input checked="" 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 type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vyjádření k výsledku anti-plagiátorské kontroly

Nepodstatná shoda 1 % s texty na Odevzdej.cz.

Dílčí připomínky a náměty:

Formální a jazyková stránka: Opakují se problémy interpunkce (například str. 7, 9, 19, 20, 36, 37, 116 a mnohé další), nevhodná stavba věty (například str. 8 poslední věta v prvním odstavci, str. 9 druhý odstavec, str. 25 první odstavec – „z hlediska teorie představ“, „data je potřeba poznat“, „...četnosti výskytu reziduí na str. 64, text na str. 116 a dalších), formální úpravy v tabulkách 1 a 2 jsou neprakticky formátované a tím jsou přehledné.

Položky 4 až 8 nevhodně zařazené v citované literatuře (osobní kontakty), naproti tomu odkaz na použitý softwar IBM SPSS Modeler chybí v seznamu použité literatury.

Termín termoklina nepatří do uvozovek, je uvedený v ČNK a užíváný v odborné literatuře. Některé informace v teoretické části jsou nepřesné (str. 25, EDA).

Kapitola 4 je obecná (problémy dat) a text se podobá některým necitovaným zdrojům.

Definice neuronové sítě (str. 32) nevystihuje moderní pojetí vícevrstevných neuronových sítí a mohla rovnou směřovat na konkrétní metodu MLP, která byla v práci použita.

Důvody pro uvádění mnoha tabulek (heatmaps) v textu výsledků nejsou jasné. Pro účel odborné kvalifikační práce měly být vybrány reprezentativní ukázky s přesnějším vysvětlením jejich významu. Popis výsledků by byl přehlednější a srozumitelnější, neopakovaly by se podobné, nepřesné až nesprávné formulace (například vysvětlení relativní chyby modelu – relative error).

Chybí ukázka nastavení parametrů modelu řešených úloh, náhled struktury dat v Modeleru, výpočtový model (datový stream) a podrobnější ukázka výstupu z Modeleru.

Celkové posouzení práce a zdůvodnění výsledné známky:

Obsah práce neodpovídá názvu práce („Statistický model...“) ani cíli: „...sestavit modely pro predikování...“. Popsaná a užitá metoda MLP poskytne řešení, ne statistický model. V práci chybí vysvětlení zvoleného přístupu, výhody i nevýhody jiných variant řešení.

Popis dat mohl být pestřejší o charakteristiky, které byly důležité pro rozhodnutí o způsobu a realizaci řešení (korelace atributů, autokorelace, popis vývoje v čase).

Implementace zvolené metody v IBM SPSS Modeler není popsána.

Výsledky v podobě heatmap jsou málo přehledné. Samotná relativní chyba vícerozměrného modelu nemůže být jediným indikátorem jeho kvality, navíc je ukazatel nepřesně interpretovaný.

V textu teoretické části práce je obtížné odlišit nepodstatné informace od zásadních, které byly důležité a využité pro návrhy modelů a jejich řešení.

Přínosem práce je seznámení autora s věcnou problematikou, datovými zdroji a jejich vazbami, provedení popis dat a příprava dat pro analýzu.

Kladně hodnotím solidně zpracovaný obecný popis metody MLP, využití odvozených veličin v časové řadě a promyšlený přístup k poměrně složité problematice přípravy dat.

Popis výsledků je nepřehledný, jejich finální prezentování je rozsáhlé a málo přínosné.

Otázky k obhajobě:

1. Které typy alternativních řešení mohly být uvažované pro danou úlohu?
2. Jak je definovaná hodnota Relative Error ve výstupu MLP v IBM SPSS Modeler?
3. Jaký potenciál má metoda MLP pro řešení daného problému? V čem vidíte přínos zvoleného postupu?

Práci doporučuji k obhajobě.

Navržená výsledná známka: C

V Hradci Králové, dne 25. srpna 2020

podpis