



POSUDEK VEDOUcí BAKALÁŘSKÉ PRÁCE

Jméno studenta: Zdeněk Archleb

Název práce: Predikce zpoždění letů metodou strojového učení

Autor posudku: doc. RNDr. Kamila Štekerová, Ph.D., MSc.

Cíl práce: Cílem práce je popsat vybranou klasickou metodu strojového učení (rozhodovací stromy) a realizovat vlastní aplikaci (předpověď zpoždění letů).

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 type="checkbox"/>	<input checked="" 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 type="checkbox"/>	<input checked="" type="checkbox"/>	<input 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 type="checkbox"/>	<input type="checkbox"/>	<input checked="" 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

Práce byla vyhodnocena službou Odevzdej.cz a byla zjištěna celková podobnost 0%. Student práci intenzivně konzultoval, postupně zapracovával nápady a připomínky, k nimž jsme společně došli. Jedná se o jeho samostatné vlastní dílo.

Díličí připomínky:

- V tištěné verzi práce chybí číslování stránek.
- Text obsahuje jisté množství překlepů, nejasných či nevhodně přeložených formulací (např. „Algoritmus je znázorněn přerušovanou čarou a snaží se najít přímku...“).
- Ve vizualizaci výsledků jsou kombinovány české a anglické popisky, vhodnější by bylo držet se jednoho jazyka.

Celkové posouzení práce a zdůvodnění výsledné známky:

Student si coby téma závěrečné práce zvolil strojové učení, přestože není v bakalářském studiu na FIM vyučováno. Samostatně si nastudoval základy statistického strojového učení a při průzkumu aplikací objevil zajímavou úlohu: predikci zpoždění v letecké dopravě. Vyhledal volně dostupné datasey (databáze vnitrostátních letů v USA), prozkoumal vhodné metody řešení a z nich zvolil rozhodovací stromy. Aby se mohl pustit do implementace, seznámil se se souvisejícími knihovnamy jazyka Python (Scikit, Pandas, Matplotlib) a naučil se používat Jupyter Notebook. Tomuto sledu prací odpovídá uspořádání textu, který má dvě části, teoretickou a praktickou.

V teoretické části práce jsou nejprve vysvětleny příčiny a důsledky zpoždění letů, což je motivací pro pokusy zpoždění predikovat. Následuje úvod do strojového učení a vysvětlení principu rozhodovacích stromů. Tento výklad není příliš dobře uspořádaný a formulace jsou neobratné, spíše na úrovni popularizačního než odborného textu (např. vysvětlení pojmu Gini impurity na konci s. 11).

V praktické části (kap. 3) je představen dataset a jeho předzpracování, implementace algoritmu CART rozšířeného o gradient boosting (za využití knihovny CatBoost) a testování modelu. Autor zkoumal i korelace mezi proměnnými, snažil se odůvodnit výsledky procesu učení a pokusil se také o ověření pomocí jiného datasetu. Přílohou práce je soubor Jupyter Notebook, obsahující návod na stažení datasetu a práci s modelem.

Bakalářská práce může posloužit jako jednoduchý úvod do problematiky implementace rozhodovacích stromů a lze na ni různými způsoby dále navázat, jednou z možností je vytvoření mobilní aplikace pro cestující, což autor navrhuje v samém závěru práce.

Otázky k obhajobě:

Vysvětlete principy algoritmu CART. Proč jste použil právě tento algoritmus (a ne některý z těch, které uvádíte v kapitole 2.3.8)?

Práci doporučuji k obhajobě.

Navržená výsledná známka: B

V Hradci Králové , dne 5. září 2023



podpis