

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Tomáš Novotný
Název práce Predikce výrobních časů v průmyslu pomocí metod strojového učení
Rok odevzdání 2023
Studijní program Informatika **Studijní obor** Umělá inteligence

Autor posudku Roman Neruda **Role** Oponent
Pracoviště Ústav informatiky AVČR

Text posudku:

Ve své práci Predikce výrobních časů v průmyslu pomocí metod strojového učení se student Tomáš Novotný zabývá aplikacemi několika metod strojového učení v doméně průmyslové výroby. Z hlediska aplikace jde o důležitou oblast uplatnění strojového učení, protože dobrý odhad času výroby je klíčový pro plánování efektivního výrobního procesu.

Předložená práce má devět kapitol včetně úvodu a závěru a obsahuje Apendix s popisy dat, vyvinutého software a výsledků experimentů. V úvodních kapitolách autor seznamuje s kontextem problému a přehledně zpracovává relevantní existující publikace. Kapitola 3 se zabývá konkrétními daty a popisuje jejich zpracování včetně detekce odlehlých pozorování a volby hodnotící funkce. V dalších třech kapitolách jsou popsány tři přístupy strojového učení k řešení daného problému. Autor postupně popisuje využití variant lineární regrese, shlukovou analýzu a použití hlubokých neuronových sítí. V sedmé kapitole jsou popsány výsledky experimentů a v závěru autor shrnuje dosažené výsledky a naznačuje další možné směry pokračování.

Za hlavní klady práce považuji následující:

1. Uzpůsobení existujících přístupů strojového učení pro konkrétní problém. Autor prozkoumal několik základních metod učení s učitelem i učení bez učitele a na základě existující literatury je přizpůsobil pro řešení specifického problému. V rámci své práce musel řešit několik netriviálních problémů. Data, která jsou k dispozici pro danou úlohu jsou heterogenní, včetně textových popisů, takže bylo třeba zvolit vhodné metody jejich preprocessingu pro dané modely strojového učení. Také velký rozptyl odhadů času výroby představoval problém, který vedl k horším výsledkům daných modelů.
2. Navržený model hluboké neuronové sítě, který pracuje s konkrétními embeddingy s ohledem na hybridní vstupní data.
3. Dobře zpracovanou experimentální část, která podrobně vyhodnocuje relativní úspěšnost modelů a přístupů.

K obhajobě bych rád, aby student okomentoval následující otázky a poznámky:

1. Zvažoval jste v části věnované lineárním regresním modelům použití dalších možností předzpracování některých datových typů, jako například TF-IDF pro textové popisy?

2. V závěru zmiňujete, že popisy operací obsahují určité klíčové prvky, které jsou obtížné pro indentifikaci, ale ovlivňují výrazně časy operací, můžete uvést nějaký příklad?

3. Jsou v literatuře popsány podobné experimenty, se kterými byste mohl své výsledky srovnat? Jak úspěšné by vaše nejlepší modely případně byly vzhledem k existujícím přístupům?

Závěrem shrnuji, že považuji předloženou práci za kvalitní a dobře zpracovanou a rád doporučuji její uznání jako diplomové práce.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

Pokud práci navrhuje na zvláštní ocenění (cena děkana apod.), prosím uveďte zde stručné zdůvodnění (vzniklé publikace, významnost tématu, inovativnost práce apod.).

Datum 25.1.2023

Podpis