

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Student: MICHALOV Ján, Bc.

Oponent: Hashim Habiballa

Studijní program: Informační technologie

Specializace: Softwarové inženýrství

Akademický rok: 2024/2025

Téma diplomové práce: **Comparison of Deep Neural Networks and Regression Models for Software Effort Estimation**

Hodnocení práce:

Student ve své práci provedl srovnání klasických regresních metod a neuronových sítí (NN) při odhadu náročnosti vývoje softwarových projektů na základě konkrétního datasetu obsahujícího 70 projektů. Výsledky ukazují, že jednoduchá neuronová síť s jednou skrytou vrstvou o čtyřech neuronech, implementovaná v TensorFlow/Keras, dosahuje srovnatelné nebo mírně lepší efektivity oproti klasickým regresním modelům, zejména Ridge regresi, která byla mezi klasickými přístupy nejúspěšnější. Práce je po stránce náročnosti úkolu průměrná, zadání bylo splněno, i když by bylo vhodné prohloubit experimentální část o testování dalších architektur neuronových sítí a jejich optimalizace, případně porovnat výsledky s dalšími pokročilými ML modely, jako jsou Random Foresty či Support Vector Machines, které jsou v této oblasti běžně využívány a v práci jsou zmíněny v teoretické části.

Silné stránky práce:

Student prokázal schopnost analyzovat a předzpracovat reálný dataset, včetně identifikace a ošetření chybějících hodnot, normalizace a výběru relevantních rysů (feature selection). Zvolil systematický přístup, kdy nejprve ověřil, jaký vliv má rozpad složených proměnných (např. rozpad environmentálních faktorů na jednotlivé složky) na přesnost modelů, a správně identifikoval, že detailnější popis environmentálních faktorů zlepšuje prediktivní schopnosti modelů.

Experimentální část je přehledně strukturovaná, využívá standardní validační postupy (3-fold cross-validation), což je vzhledem k malému rozsahu dat vhodné. Výsledky jsou prezentovány pomocí přehledných tabulek a grafů, které umožňují srovnání jednotlivých modelů na základě metrik R^2 a RMSE.

Práce je psána v angličtině, má standardní strukturu a v úvodních kapitolách dobře shrnuje základní pojmy, popisuje dataset i použité metodologie.

Slabší stránky a doporučení:

Experimentální část by šla rozšířit o testování dalších architektur neuronových sítí (např. více vrstev, různé počty neuronů, různé regularizační techniky), případně i o další typy ML modelů, které jsou v literatuře běžné (Random Forest, SVM apod.). Tím by bylo možné lépe posoudit, zda jednoduchá NN skutečně představuje optimální řešení, nebo zda by složitější modely přinesly další zlepšení.

Práce se opírá pouze o jeden dataset se 70 projekty, což omezuje zobecnitelnost výsledků. Student sice správně uvádí důvody, proč nebylo možné použít některé proměnné (např. DataDonator kvůli riziku



"leakage"), nicméně nebyla provedena žádná evaluace na jiných datech nebo simulace přenositelnosti modelu.

Výběr literatury je spíše obecný, chybí odkazy na konkrétní studie, které by analyzovaly využití ML metod v odhadu SW náročnosti na obdobných datech, případně srovnání výsledků s publikovanými benchmarky.

Dotazy a připomínky:

V tabulce 2.4 je u typu aplikace "Real-time application" záznam dvakrát – jednou s velkým a jednou s malým písmenem, což může být důsledek nekonzistence v zadání dat. Je vhodné sjednotit názvy kategorií, aby nedocházelo k duplicitám a zkreslení analýzy.

Grafy 3.1 a 3.2 obsahují málo informací – pro lepší interpretaci by bylo vhodné použít například bodové grafy (scatter plot) s jednotlivými projekty, případně boxploty rozptylu predikovaných a skutečných hodnot, což by lépe ukázalo, kde modely selhávají nebo naopak excelují.

Použití konkrétní konfigurace NN na jiných datasetech nebylo v práci testováno, student tedy nemůže garantovat zobecnitelnost výsledků. Bylo by vhodné v budoucnu ověřit model na jiných datech, případně použít transfer learning nebo jiné techniky pro zvýšení robustnosti modelu.

Práci považuji za průměrnou, dobře splňující zadání v rámci možností dostupných dat. Student prokázal schopnost aplikovat základní i pokročilejší metody strojového učení, systematicky analyzovat data a interpretovat výsledky. Rezervy vidím zejména v širší experimentální části (další modely, hlubší analýza architektur NN) a v otázce zobecnitelnosti. Přesto je práce pro daný dataset použitelná a doporučuji ji k obhajobě s hodnocením C.

Celkové hodnocení práce:

Známku uvede oponent dle svého uvážení dle klasifikační stupnice ECTS:

A – výborně, B – velmi dobře, C – dobře, D – uspokojivě, E – dostatečně, F – nedostatečně.

Stupeň F znamená též „nedoporučuji práci k obhajobě“.

Předloženou diplomovou práci doporučuji k obhajobě a navrhuji hodnocení

C - dobře.

V případě hodnocení stupněm „F – nedostatečně“ uveďte do připomínek a slovního vyjádření hlavní nedostatky práce a důvody tohoto hodnocení.

Datum 03.06.2025

Podpis oponenta diplomové práce