

## I. IDENTIFIKAČNÍ ÚDAJE

<b>Název práce:</b>	Contextual Embeddings for Predictions Based on Log Files
<b>Jméno autora:</b>	Petra Vaňková
<b>Typ práce:</b>	diplomová
<b>Fakulta/ústav:</b>	Fakulta elektrotechnická (FEL)
<b>Katedra/ústav:</b>	Katedra počítačů
<b>Vedoucí práce:</b>	Ing. Jan Drchal, PhD.
<b>Pracoviště vedoucího práce:</b>	Katedra počítačů

## II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

<b>Zadání</b>	<b>průměrně náročné</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání vyžaduje zvládnutí moderních metod používaných v NLP a jejich aplikaci pro analýzu souborů logů. Použité datové sady jsou rozsáhlé, což vede na vysoké výpočetní nároky a nutnost dobře plánovat experimenty.	

<b>Splnění zadání</b>	<b>splněno</b>
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Zadání bylo splněno bez výhrad.	

<b>Aktivita a samostatnost při zpracování práce</b>	<b>A - výborně</b>
<i>Posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posuďte schopnost studenta samostatně tvůrčí práce.</i>	
Studentka byla na konzultace vždy perfektně připravena. Oceňuji rychlost, s jakou byla schopna implementovat nové experimenty, na kterých jsme se průběžně domlouvali.	

<b>Odborná úroveň</b>	<b>A - výborně</b>
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Z hlediska odborné úrovně je práce na vysoké úrovni. V predikcích sice nebylo dosaženo příliš dobrých výsledků, ale tento výsledek byl dán spíše problematickou datovou sadou (viz Další komentáře a hodnocení). Studentka postupovala systematicky a v textu dobře motivovala svůj postup. Ve výsledku považuji za nejpřínosnější podrobnou analýzu dat (Exploratory Data Analysis) – ta v budoucnu pomůže s lepší volbou zdrojů dat. Ve finální fázi studentka odhalila implementační chybu předzpracování dat pro modely postavené na architektuře Temporal Convolutional Network – v omezeném zbývajícím čase se jí podařilo část zásadních experimentů zopakovat. Problém je v textu dobře zdokumentován, což oceňuji.	

<b>Formální a jazyková úroveň, rozsah práce</b>	<b>A - výborně</b>
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je psána kvalitní angličtinou s minimem překlepů a je dobře strukturována. Drobnou výhradu mám k tabulkám, u kterých by bylo vhodné zvýraznit nejdůležitější dosažené hodnoty.	

<b>Výběr zdrojů, korektnost citací</b>	<b>A - výborně</b>
<i>Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.</i>	
Zdroje jsou citovány správně.	

#### **Další komentáře a hodnocení**

*Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.*

Cílem práce bylo otestovat schopnosti neuronových modelů typu BERT pro zpracování souborů logů. Nejčastěji je nad podobnými daty řešena úloha typu anomaly detection, v našem případě se jednalo o úlohu predikce. V datech poskytnutých Škoda Auto, a.s. jsme se pokoušeli o predikci náročnosti výrobního procesu, jež se měl hypoteticky promítat do délky zpracování databázových operací. Ve výsledku se tato hypotéza nepotvrdila a dosažené výsledky nejsou příliš použitelné. Musím však zdůraznit, že studentka postupovala metodicky správně. V rámci následného rozvíjení tohoto tématu navrhneme rozšíření datové sady o další relevantnější prediktory.

### **III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE**

*Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.*

Studentka odvedla velký objem zejména experimentální práce a vše velmi kvalitně zdokumentovala. Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Datum: 2.9.2021

Podpis: