

**I. IDENTIFIKAČNÍ ÚDAJE**

<b>Název práce:</b>	<b>Sequence translations and their applications</b>
<b>Jméno autora:</b>	<b>Katka Morovicová</b>
<b>Typ práce:</b>	bakalářská práce
<b>Fakulta:</b>	Fakulta jaderná a fyzikálně inženýrská (FJFI)
<b>Katedra:</b>	Katedra matematiky
<b>Oponent práce:</b>	Ing. Petr Pošík, Ph.D.
<b>Pracoviště oponenta práce:</b>	ČVUT FEL, katedra kybernetiky

**II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ**

<b>Zadání</b>	<b>náročnější</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Odhaduji, že téma práce není pokryto standardním obsahem bakalářského studia na FJFI a studentka tedy musela vynaložit značné úsilí k proniknutí do problematiky. Z toho důvodu hodnotím téma jako náročnější.	

<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>	
Všechny požadavky uvedené v zadání považuji za splněné.	

<b>Zvolený postup řešení</b>	<b>vhodný</b>
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Studentka provedla rešerši existujících přístupů pro překlad sekvencí se zaměřením na neuronové sítě. Vybrané modely aplikovala a porovnávala na 3 úlohách: překlad z angličtiny do francouzštiny, detekce konců slov v textu a prepis hieroglyfického textu, z nichž poslední dvě jsou velmi nestandardní a značně obtížné kvůli nedostatku dat. Porovnávané architektury modelů byly vždy testovány s několikerym nastavením hyperparametrů.	

<b>Odborná úroveň</b>	<b>výborná</b>
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Splnění zadání vyžadovalo nastudovat nejen literaturu o neuronových sítích, ale také proniknout do základů zpracování egyptských hieroglyfů. Obojí se studentce povedlo a přehledová kapitola, kde obě oblasti představuje, je stručná, srozumitelná a obsahuje potřebné informace. Studentka také prokázala schopnost implementovat modely známé z literatury a díky spolupráci s Egyptologickým ústavem AV ČR získala i zpětnou vazbu na dosažené výsledky.	

<b>Formální a jazyková úroveň</b>	<b>výborná</b>
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce má logickou strukturu a je psána velmi čtivou angličtinou. Po formální stránce nemám k práci připomínek. (Jedinou malou výtku mám k uspořádání částí textu: studentka zvolila členění dle použitých typů NN a pro mě jako čtenáře bylo trochu obtížnější uvědomit si, na jaké úloze je daný typ NN testován. Osobně bych preferoval členění spíše podle řešených úloh.)	

<b>Výběr zdrojů, korektnost citací</b>	<b>výborné</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</i>	

*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.*

Seznam literatury obsahuje 34 položek, které jsou v práci řádně citovány. Literatura týkající se překladu sekvencí je nesmírně bohatá a neumím posoudit, zda studentka využila *\*všechny\** relevantní zdroje; nicméně vybrané publikace, z nichž práce čerpá, považuji za dostatečně reprezentativní. Některé zdroje jsou články ze serveru arxiv, k nimž možná existuje publikace v časopise nebo na konferenci, nicméně je jich jen zlomek celkového počtu. Nezaznamenal jsem žádné porušení citační etiky ani plagiátorství: podle služby iThenticate je skóre podobnosti tohoto dokumentu menší než 20 %, přičemž shody byly nalezeny jen v krátkých fragmentech vět o několika slovech.

#### **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.*

**Klepněte sem a zadejte text.**

### **III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE**

*Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.*

I přes doplňující dotazy níže považuji tuto práci za výbornou, a to z následujících důvodů:

- \* Náročnost zadání přesahuje obvyklou úroveň bakalářských prací.
- \* Úvodní přehledové kapitoly o zpracování egyptských textů a o architekturách neuronových sítí pro překlad sekvencí jsou stručné, jasné a obsahují relevantní informace.
- \* Velmi oceňuji, že každou porovnanou architekturu neuronových sítí studentka otestovala s několika sadami hodnoty hyperparametrů.
- \* Dosažené výsledky jsou smysluplně porovnány a diskutovány.

Dotazy k obhajobě:

- 1) V sekci 3.1.1 je poprvé použit termín „Encoder Only“, aniž by byl předtím definován. Můžete přesněji vysvětlit, co přesně tento termín znamená? Pokud se jedná jen o část Encoder-Decoder modelu, přičemž dekodér chybí, dává stále smysl mluvit o enkodéru? V jakém smyslu je tento model enkodérem?
- 2) V sekci 3.2 zmiňujete, že jste implementovala RNN, LSTM a Transformer sama. V popisu jednotlivých vrstev se ale odkazujete na dokumentaci Kerasu. Rozumím dobře, že tím „implementovala sama“ máte na mysli, že jste využila existující vrstvy/bloky v Kerasu, které jste sama sestavila do konkrétních NN modelů?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Datum: 12.8.2024

Podpis: Ing. Petr Pošík, Ph.D.

