

Posudek diplomové práce

Název práce: Jazyková arbitrárnost pohledem kvantitativních metod

Jméno diplomantky: Bc. Klára Hájková

Cílem předložené diplomové práce je zabývat se otázkou existence a průkaznosti vztahů kvantifikovatelných vlastností jmen – především vlastností týkajících se jejich fonologie či délky – a důležitých vlastností objektů, na která se tato jména odkazují. V práci postupně nalezneme analýzu vztahů jmen měst a obcí s velikostí jejich populace, analýzu vzorů využití grafémů ve vztahu s tím, zda se jedná o jméno hlavního města, nebo tzv. periferního města v dané zemi, analýzu vztahů délky názvů filmů a jejich oblíbenosti a další, širší analýzy týkající se především toponym.

Provedení této práce je bráno primárně technicky, tedy za použití statistického aparátu tam, kde je možné vybrané kvantifikované vztahy zkoumat a neporušovat tím prerekvizity daných testů a, co je zajímavější, tam, kde by tyto prerekvizity byly potenciálně porušeny nebo se jedná o otázky vyžadující hledání skrytých vzorů, je využito strojové učení, především pak formou klasických SVM, včetně jejich plné evaluace pomocí trénovacího a testovacího datasetu a tam, kde je to možné i interpretace koeficientů. Veškeré výsledky jsou statisticky podepřené a veškeré jevy – z technického hlediska – prokazatelné. Tím více jsou výsledky, které v práci nalezneme, překvapující. Výsledky souhrnně říkají, že pro řadu jmen, ať už se jedná o jména měst nebo řek, nalezneme prokazatelný, nenáhodný, a tedy systematický vztah s kritickou vlastností pojmenovaného objektu. Jaké jsou důvody existence takových vztahů jsou pak v práci diskutovány z široké škály hledisek od jazykové ekonomie, vlivů původních indoevropských jazyků, až k vlivům kolonialismu. Práci tak nelze upřít jasný cíl, technicky skvělé provedení, evaluaci, nabídnutá vysvětlení, diskuze, a především pak závěrem, který shrnuje, že problematiku vztahů jmen a jejich objektů nelze řešit pouze technicky a kvantitativně, ale vyžaduje i nutný kvalitativní vhled. Otázka, zda jsou *označující* a *označované* vzájemně *arbitrární* ve zkoumaných případech, pak zůstává otevřená – zároveň je tato otázka zvýrazněna a podtržena zjištěními, že už jen pro některé vybrané vztahy (natož pak všechny) vysvětlit a vyloučit jejich arbitrárni či systematickou povahu nepůjde nijak snadno. Důležité je také zmínit to, že práce obsahuje – krom samotného textu – i kompletní zdrojové kódy pro trénování a evaluaci strojového učení v jazycích Python/R.

Práci lze z technického i formálního hlediska vytknout relativně málo. Práce je dobře uspořádána, psána srozumitelným jazykem a vhodně graficky upravena. Obsahově však práci chybí dva důležité body. První bod je chybějící větší vhled do literatury a do zdrojů, které by zkoumanou problematiku lépe podepřely, především z hlediska sémiotiky a otázek definic arbitrárnosti. Druhý bod je, že některé interpretace a závěry jsou velmi stručné a postrádají hlubší analýzy. Je však nutné dodat, že rozsah práce je již teď značný a její rozšíření o další diskuze by tak nemuselo být vhodné.

Celkově se však jedná o fascinující práci, a i přes některé potíže ji lze doporučit k obhajobě se stupněm A.

Otázky

- Na základě všech zjištění Vaší práce a na základě Vašeho názoru, lze podle Vás říci, že – z fonologického hlediska – existuje v některém případě nějaký vztah jména vůči označovanému?

Mgr. Vladimír Matlach, Ph.D.
V Olomouci dne 23. 8. 2021