

Automatická kategorizace zákaznických požadavků

Autor: Filip Koukal

Vedoucí práce: doc. Ing. František Dařena, Ph.D.

Ústav informatiky, Provozně ekonomická fakulta, Mendelova univerzita v Brně



Motivace

Průlom v oblastech zpracování přirozeného jazyka a umělé inteligence spolu s trendem zapojování cloudových služeb do firemních procesů představují pro oblast zákaznické podpory mnoho příležitostí. Pomocí automatického zpracování zákaznických požadavků lze zefektivnit téměř všechny dílčí části v procesu řešení požadavku, což **usnadňuje práci operátorů a zlepšuje uživatelské zkušenosti** se zákaznickou podporou. Tato práce byla tvořena ve spolupráci s firmou ALVAO s.r.o. na projektu modernizace helpdesk systému.

Byly identifikovány tři hlavní úlohy, které se následně řešily pomocí strojového učení a umělé inteligence. Jedná se o úlohy **přiřazení kategorie k požadavku, přiřazení řešitele k požadavku a vyhledání a označení duplicitních požadavků**.

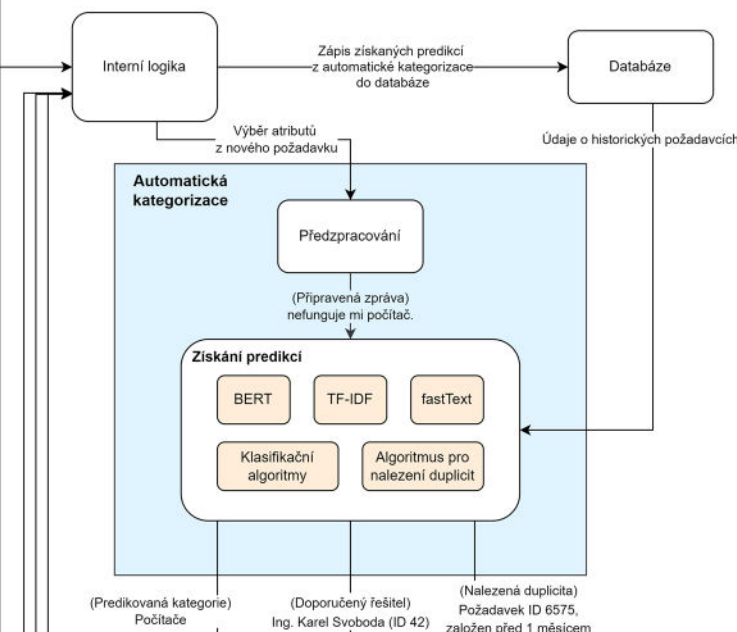
Navrhované řešení

14:01:11 08.03.2021
Ahoj, nefunguje mi počítač.
S pozdravem,
Jan Novák | jan.novak@email.cz

Vstup požadavku do systému

Systém zákaznické podpory

- Zákaznické požadavky jsou převážně v **textové podobě**. Pro doporučení řešitele jsou relevantní i některé netextové údaje.
- Data jsou velice **nevyvážená**, proto byl zkoumán dopad vyvažovacích technik pro zlepšení výsledků.
- Bylo srovnáno několik reprezentací textových dat, tradičních i moderních, mezi které patří zejména **TF-IDF, BERT a fastText**.
- U klasifikačních úloh bylo zkoumáno **11** různých klasifikátorů v kombinaci se všemi reprezentacemi.
- Po extensivních experimentech byly nejlepší přístupy pro všechny identifikované úlohy adaptovány do formy cloudové služby.



Výsledky klasifikačních úloh

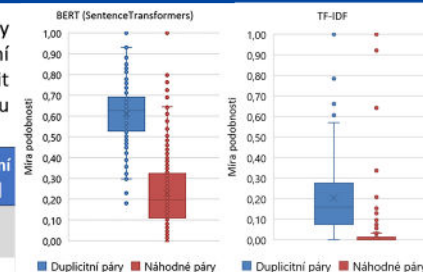
Bylo vykonáno více než **1000** dílčích experimentů. Nejlepší přístupy se liší podle daného úkolu. Pro predikce kategorie požadavku bylo dosaženo nejlepších výsledků s moderní textovou reprezentací BERT. Pro doporučení řešitele bylo výhodné tradiční řešení pomocí TF-IDF.

Úloha	Algoritmus	Makro F1 skóre [%]	Rychlost získání predikce [ms]
Přiřazení kategorie	BERT + XGBoost	86,94	312
Přiřazení řešitele	TF-IDF + XGBoost	74,11	224

Vyhledání duplicitních požadavků

Byly zkoumány 2 přístupy využívající předtrénované modely založené na modelu BERT a tradiční TF-IDF přístup. Tradiční přístup poskytuje dobré výsledky, nicméně vyhledání duplicit trvá více než **dvakrát déle** než moderní přístupy, kvůli vysokému počtu dimenzí.

Model	Podíl označených duplicit ve vyhledaných kandidátech [%]	Recall [%]	Rychlost získání predikce [ms]
BERT (SentenceTransformers)	25,25	25,49	517
TF-IDF	28,15	35,78	1366



Nasazení prototypu

- Prototyp služby poskytuje real-time predikce pro nově založené požadavky.
- Využito Platform-as-a-Service služby **Azure App Service** pro nasazení, což poskytuje kompromis mezi výkonem a náklady na provoz.
- Služba nasazena pomocí Docker Containeru pro snadné, opakovatelné a přenositelné nasazení.
- BERT modely byly optimalizovány pomocí **ONNX Runtime** frameworku pro dodatečné zrychlení poskytnutí predikce.
- Komunikace s existující aplikací pro zákaznickou podporu uskutečněna pomocí REST API.

