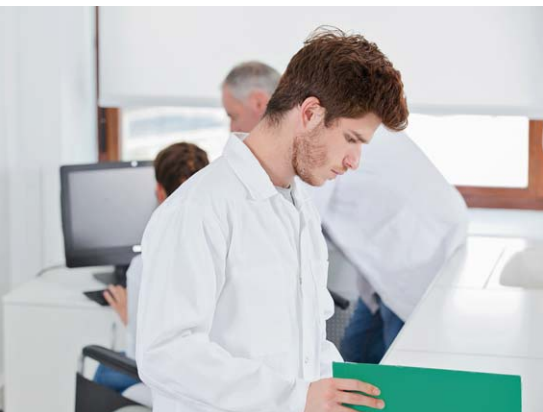


Využití strojového čtení textu pro lokální přehled rezistencí v Thomayerově nemocnici

Dominik Herman



Rezistence patogenů na antibiotika je problematika, která pro smysluplné vyhodnocování vyžaduje práci s velkým množstvím dat. Pro analýzu tohoto typu se tak nabízí využití inteligentních softwarových řešení, protože eliminuje subjektivní vstupy a lidské chyby.

Skupina Zdravel, věnující se širokému portfoliu produktů v oblasti elektronizace zdravotnictví, provedla takovou analýzu počátkem roku 2020 v Thomayerově nemocnici v Praze, a jejím výstupem byla dokumentace s názvem: „*Lokální přehled rezistencí: Místní přehled nejčastějších patogenů a jejich citlivostí na antibiotika za roky 2017 a 2018 na lůžkových odděleních Thomayerovy nemocnice*“. Za pomoci strojového čtení textu byl vytvořen obsáhlý přehled, na jehož základě již dnes dochází k úpravě antimikrobiálního zabezpečení chirurgických výkonů a racionalizaci podávání antibiotik. Taková cílená opatření chrání zdraví pacientů a šetří čas a finanční prostředky zdravotnického zařízení.

V rámci analýzy bylo zpracováno přes 70 tisíc XML souborů z laboratorního informačního systému obsahujících výsledky mikrobiologických vyšetření. Výsledky uložené ve formě volného textu byly strukturovány pomocí nástroje HarveyX, který Zdravel vyvinul pro účely strojového čtení nestruturovaných částí

zdravotnické dokumentace. Řešení postavené na platformě IBM Watson dokáže v tomto případě spolehlivě určit typ odebraného vzorku, pozitivní výskyt jednotlivých druhů patogenů, antimikrobiální rezistenci, citlivost na antibiotika a jejich další vlastnosti.

Poptávka po strojově zpracované analýze vzešla z absence dostatečně široké palety statistických přehledů v rámci nemocničního a laboratorního informačního systému, které by umožnily sledovat trendy vybraných ukazatelů a na jejich základě činit opatření pro zlepšení lokální epidemiologické situace jak na úrovni nemocnice jako celku, tak po jednotlivých odděleních.

Základní prerekvizitou pro sestavení požadovaných tabulkových a grafických přehledů bylo strukturování informací. Primární XML soubory ve formátu DASTA obsahovaly v rámci bloku <v> (dedikovanému formalizovaným výsledkům laboratorních vyšetření) veškeré podstatné údaje pouze ve formě volného textu uloženého v elementu <ptext>. Nástroj HarveyX však dokáže univerzálně číst zdravotnickou dokumentaci v českém jazyce jak s ohledem na skloňování, časování, synonyma, zkratky či negace, tak na odborné zdravotnické slovníky anatomie, diagnóz, výkonů, léků a účinných látek aj.

Díky revizi algoritmů doc. MUDr. Pavlem Čermákem, CSc., primářem Oddělení klinické mikrobiologie Thomayerovy nemocnice, tak vznikl mechanismus pro sestavení přehledu osvětlujícího lokální epidemiologickou situaci, který:

- je rigorózně definován za přímé kooperace s expertem v daném odvětví,
- poskytuje spolehlivé a přesné výsledky,
- je zpětně auditovatelný a v budoucnu reprodukovatelný,
- šetří čas a finanční prostředky v porovnání s manuálním vyhodnocením.

Snaha z uplynulých let o vytvoření přehledu manuálně si žádala zapojení lékařů v takovém rozsahu, že ekvivalent dnešního strojového výstupu sestavovaly dvě osoby bezmála půl roku. Navíc se s ručním zpracováním

takového množství dat pojila neodvratitelná chybovost. V porovnání s tím byla společná revize algoritmů pro třídění informací hotova v horizontu týdnů, přičemž přímo porovnatelné výsledky lze v budoucnu nad novými daty dodat během několika dní včetně rigorózních validací výstupu.

Tento projekt spočíval v dodání jednorázové analýzy, kterou je třeba každoročně opakovat. Původně však Zdravel společně s Hartmann-Rico oslovil zástupce Thomayerovy nemocnice s nabídkou implementace nástroje pro aktivní surveillance nozokomiálních nákaz Epidis, který v reálném čase zpracovává významnou většinu celé zdravotnické dokumentace z nemocničního a laboratorního informačního systému. Výše popsaná analýza je pouze jednou z integrálních součástí daného nástroje, který umožňuje sledovat trendy průběžně. Důvodem jednoduššího zadání byl problém s dostupností dat ze zastaralého nemocničního informačního systému, který zadavatele i řešitele odkázal pouze na práci s daty z laboratorního informačního systému. Podaří-li se v budoucnu efektivně napojit všechna potřebná data, mohli by zdravotničtí profesionálové pracovat s mnohem obsáhlejší analýzou, která jim bude k dispozici stále aktuální a online. ■

Dominik Herman



Autor článku je Chief Analytics Officer projektu Zdravel.