

## Abstrakt

Tato práce se zabývá možnostmi využití strojového učení (ML) při analýze zpravodajských informací z otevřených zdrojů (OSINT). Vzhledem k tomu, že data jsou společným vstupem obou oborů, jsou data hlavní optikou, kterou je téma zkoumáno. Pro pochopení celého procesu nasazení ML modelu od sběru dat až po jejich analýzu byl vycvičen a otestován obrazový klasifikátor ztrát ruských vozidel při invazi na Ukrajinu. Tento klasifikátor, vycvičený na více než 50 000 obrázcích z databáze WarSpotting, dosáhl slušné přesnosti 79 % na tréninkových datech pěti nejpočetnějších kategorií snímků. Na testovacích datech z pozdějšího období klesl výkon na 62 %. Jedno z nabízených vysvětlení je, že statické frontové linie a rozšíření dronů vedly k tomu, že většina nedávných snímků byla pořízena ze vzduchu, zatímco tréninková data byla pořízena převážně ze země. Tento výsledek ukázal, jak nevyhnutelné změny i ve zdánlivě dobře spravovaných datech mohou vést k nízkému výkonu ML modelů při nasazení. Kromě změn na bojišti vyšly najevo i hlubší problémy s daty, včetně kaskádových účinků raných rozhodnutí o procesování dat a jejich nevyváženosti. Celkově lze říci, že současné metody klasifikace obrazu nefungují dobře na dostupných nedokonalých datech.