

EXAMENSARBETE Exploring the suitability of explainable AI for binary conflation of POI data**STUDENTER** Emil Aminy, Marcus Lissner**HANDLEDARE** Pierre Nugues (LTH), Frank Camara (AFRY)**EXAMINATOR** Jacek Malec (LTH)

Undersökning av explainable AI för sammanslagning av POI-data

POPULÄRVETENSKAPLIG SAMMANFATTNING **Emil Aminy, Marcus Lissner**

En undersökning av hur effektivt explainable AI är för binär sammanslagning av POI-data samt hur användbar dess förklarbarhet är för denna uppgift.

Det finns idag många företag och tjänster som tillgodoser diverse kartdata. Det kan till exempel handla om vägar, vegetation, eller platser av intresse (POI). Det finns dock ingen garanti att data från en enstaka leverantör är helt korrekt, fel kan alltid uppstå. En metod för att korrigera sådana fel är sammanslagning; kombinerande av data från flera källor för att mitigera eventuella fel.



Att utföra en sådan process på ett vis där slutprodukten är mer korrekt än alla dess delar är dock inte helt trivial. Någon form av beslutsprocess måste existera som kan avgöra vilken data från vilken källa kan anses vara av bättre kvalitet. Ett väl lämpat verktyg för denna process är maskininlärning. Men, om detta moment skulle genomföras med avancerade maskininlärningsalgoritmer, såsom neurala nätverk, kan algoritmens resonemang kring datasammanslagningen bli väldigt svårtolkad för en människa.

Att förstå sig på varför en algoritm gör som den gör kan vara användbart för vidareutvecklingen av algoritmen. Dessutom kan insikten ge människor förtroende i att algoritmen utför sitt jobb korrekt. Det är därför som vi i detta examensarbete har tittat på hur explainable AI (XAI) kan användas för att råda bot på detta insynsproblem. Vårt fokus ligger specifikt på XAIs användbarhet vid sammanslagningen av POI-data. Vi vill ta reda på hur pass jämförbar en XAI-lösning är för just denna uppgift, både i förhållande till en mindre förklarbar lösning men även jämfört med en människa. Vi vill även undersöka hur användbar och effektiv förklarbarheten är.

Vårt experiment gick ut på att utföra binär sammanslagning av POI-data med hjälp av ett beslutsträd och jämföra dess resultat med ett neuralt nätverk och mänsklig prestanda. POI-data för detta experiment insamlades från fem källor: Eniro, Foursquare, Hitta, OpenStreetMap, och Tripadvisor. Data begränsades endast till POI:ar inom Sverige som innehöll information om hotell- och restaurangverksamheter.

Resultaten pekar mot att beslutsträd är jämförbara med både neurala nätverk och människor för denna uppgift. Deras förklarbarhet är dessutom väldigt användbar vid utvecklingsstadiet då insynen tillåter snabb evaluering och förbättring av skapade modeller.