

Datadrivet lärande av vägbeskrivningar

Dmitrij Lioubartsev och Kristoffer Hallqvist har vid CSC, KTH genomfört ett kandidatexamensarbete inom datalogi som behandlar ett sätt att lära en maskin betydelsen av diverse objekt och rörelsemönster baserat på insamlad data. Målet för projektet var att specifikt utveckla en programvara som klarar inlärning av vägbeskrivningar. Författarna har använt sig av mättalet tf-idf (term frequency – inverse document frequency) som huvudredskap för att vikta betydelsen av olika ord för mappning mot koncept och objekt. För varje koncept har ord rankats efter tf-idf och den utvecklade programvaran har verifierats genom manuell kontroll om "rätt" ord (valt och presenterat av författarna) har valts av datorsystemet. Resultatet av arbetet är att det går att lära en dator ord som representerar specifika, fördefinierade objekt och koncept givet tillräcklig datamängd.

Dispositionen är tyvärr undermålig, rapporten inleds med en sedvanlig introduktion som leder in i en bakgrund. I bakgrunden presenteras förutom liknande arbete och den studie som data inhämtats ifrån, även specifika detaljer kring data och datastruktur. Denna redogörelse bör finnas senare i dokumentet under någon sektion som presenterar testmiljön som undersökningen utförts i. En pedagogisk redogörelse över verktyget td-idf följer. Sektion 3 är betitlad "Procedure", men inleds inte med testmetodik eller testförfarande, utan definierar begreppet "concept" som används framåt i rapporten. Sektion 3.1 avslutas med en problemformulering som inte känns helt genomtänkt då den syftar till att utröna hur väl ett än icke-definierat datorprogram kan lära sig de korrekta orden för de 16 fördefinierade koncept från det projekt som data tagits från. Frågeställningen borde finnas långt tidigare än i mitten av dokumentet och dessutom försöka utröna något större. Ett lämpligt förslag vore något som "Kan tf-idf användas för att filtrera ord i ett system som ska vara kapabelt att lära sig ord relaterade till vägbeskrivningar?" Arbetet som helhet skulle tjäna väldigt mycket på en explicit avgränsning, tydlig problemformulering och omarbetad disposition. Även språket håller en anmärkningsvärd låg kvalitet, med vissa felstavningar, men även felaktiga ordval, stundvis märkliga tempusval samt användning av felaktiga prepositioner. Den främsta kritiken mot språket är dock att författarna skriver med ett språk som starkt påminner om talspråk.

Givet den problemformulering som presenteras kan det dock anses att författarna på ett kort och koncist, sammanfattat adekvat sätt, motiverar sin metod till att lösa problemet. Dock utelämnas all form av motivering till vetenskaplig metod för att verifiera sin egen lösning.

Det finns en anmärkningsvärd brist på referenser som stödjer påståenden och fakta som presenteras. Rapporten refererar endast 5 källor, som hastigt presenteras under rubriken "Similair work", i övrigt saknas källhänvisningar nästan helt genom dokumentet. Även motiveringar saknas, i sektion 2.2 finns ett stycke som i sin helhet ser ut såhär:

"The sound files containing the actual recordings are not available, but they are not relevant to this project either."

Utsagan må visserligen vara sann och framgår om man läser hela dokumentet men läsaren ges en känsla av att dokumentet är ofärdigt.

Ett resultat som författarna når är att en kvadrering av idf ger ett bättre resultat för de specifika data som vederbörande opererat på, men att en kubisk viktning på samma delmått ger en poängfördelning som är sämre för samma data. Det är en sund slutsats som stöds av presenterad data. Dock gör författarna en logiskt kullerbytta när de påstår att deras formel 3.2 är optimal för att beräkna tf-idf värde. Dels saknar de data till att göra en allmän utsaga. Om det ändå är så att de uttalar sig om sin specifika data har de ändå fel, då de inte har undersökt alla möjliga koefficienter och exponenter de kan använda i sin funktion, de har inte utfört en korrekt optimering och kan därmed inte uttala sig om funktionens optimalitet.

Rapporten beskriver en (egenutvecklad) algoritm för att detektera rörelser i en generell riktning och sätta ut punkter som sammanfattar en rörelse genom att dra raka linjer mellan punkterna. Dock visas i figur 5.1 att denna algoritm inte fungerar helt tillfredsställande. Läsaren undrar om det inte finns några standardmetoder och standardalgoritmer för att lösa detta problem, då rapporternas författare förmodligen inte är de första som behövt lösa detta problem. Vid en mer noggrann utläggning kring varför denna algoritm ser ut som den gör och varför de valt den kanske läsaren skulle känna sig tryggare i författarnas metod.

Författarna har visat att det i en kontrollerad miljö med en uppsättning kontrollerade koncept går att skapa ett system som kan lära sig ord som motsvarar koncepten med en tillräcklig mängd data kommande från ett verkligt experiment. De har nått en nykter slutsats där de reserverar sig för tillämpning utanför kontrollerad laborationsmiljö. Dock kan man fråga sig om inte "Datadriven inlärning av vägbeskrivningsatomer" hade varit en mer passande titel, då arbetet utelämnar utvärdering av datorns möjlighet att följa eller skapa vägbeskrivningar efter utförd inlärning.