

EXAMENSARBETE Sentiment Analysis and Aspect Extraction for Business Intelligence**STUDENTER** Fredrik Olsson, Gustav Handmark**HANDLEDARE** Pierre Nugues (LTH)**EXAMINATOR** Jacek Malek (LTH)

Sentimentanalys av kundåsikter för strategiskt beslutsfattande

POPULÄRVETENSKAPLIG SAMMANFATTNING **Fredrik Olsson, Gustav Handmark**

Mängden tillgänglig data i form av användarrecensioner, inlägg och kommentarer på sociala medier har ökat kraftigt på senaste tiden. För många företag är det en överväldigande uppgift att manuellt granska denna data för att få en uppfattning om vad kunder tycker. Hur kan vi automatisera denna process?

Att undersöka kunddata och kunna förutse trender är viktiga instrument vid affärsanalys och kan leda till identifiering av strategiska affärsmöjligheter. I detta arbete har vi utforskat modern språkteknologi för bearbetning av företagsriktade kundrecensioner, och på så sätt lyckats automatisera analysen och möjliggöra ett datadrivet beslutsfattande.

I vårt arbete har vi behandlat kundrecensioner på två nivåer. Först tittar vi på recensionerna som helhet och klassificerar dessa som positiva eller negativa. Detta innebär att vi kan få ett omdöme som är frikopplat från ett eventuellt stjärnbetyg, och som istället är helt baserat på innehållet i texten.



“ **Maten var god, men vår servitör var rätt otrevlig** ”

Vi arbetar därefter med aspekt-baserad sentimentanalys för att extrahera mer detaljerad information från varje recension. Detta innebär att vi på detaljerad nivå kan få ut aspekter och till-

hörande beskrivande ord. Istället för att enbart se hela recensionen som positiv eller negativ, kan vi till exempel få ut att kunden tyckte maten var god, men var missnöjd med sin servitör som var rätt otrevlig.

Sedan aggregerar vi resultaten från tusentals användarrecensioner för ett visst företag och presenterar dessa i diverse grafer och interaktiva tabeller. Detta möjliggör enkla tolkningar av resultaten och observerade trender över tid.

Vi kan se hur antalet positiva och negativa recensioner har varierat över tid. Vi kan också se fördelningen mellan positiva och negativa omnämningen för de tio vanligaste aspekterna och få ut ett stjärnbetyg för dessa. Slutligen kan vi också undersöka de ord som använts för att beskriva de vanligaste aspektorden.

För att öka användbarheten av vårt system har vi skapat modeller som klarar av att hantera såväl engelska som svenska texter. I framtiden hade vi gärna titta närmare på möjligheterna att använda våra nuvarande modeller för att skapa syntetiska svenska dataset, vilka kan användas för att träna upp mer kraftfulla modeller för det svenska språket. Dessutom hade det varit intressant att se om det finns något samband mellan kundernas åsikter om ett företag och dess aktiekurs.