

# Utmaningar inom attitydanalys och automatisering av enkätanalyser

Oscar Norberg  
Merve Günes

**Systemvetenskap, kandidat**  
**2019**

Luleå tekniska universitet  
Institutionen för system- och rymdteknik

## **Förord**

Vi vill tacka alla som varit med och hjälpt oss genomföra denna kandidatuppsats. Ett stort tack till våra deltagare som hjälpte oss att genomföra undersökningen, utan er hade vi inte fått ihop ett resultat. Vi vill också tacka företaget Atea som faciliterade vårt examensarbete och väckte vårt intresse för forskningsområdet. Vi vill också tacka företaget Golfplisir AB och vår kontaktperson Robert Norberg som försedde oss med all data som har legat till grund för arbetet. Slutligen ett stort tack till vår handledare Svante Edzén vid Luleå Tekniska Universitet som väglett oss och svarat på alla våra frågor under arbetets gång.

Merve Günes & Oscar Norberg

*Luleå, Juni 2019*

## Sammanfattning

I studien utreds det mänskliga sättet att analysera text jämfört med ett attitydanalys-verktygs tolkning. Syftet med studien är att ta fram skillnaderna som uppmärksamhet vid jämförelsen av människans analys av text och attitydanalys-verktygets analys. Området blir intressant eftersom analyser av enkätundersökningar blir allt mer automatiserade och det finns ett stort utbud av olika verktyg som erbjuder tjänster för att analysera data. Attitydanalys-verktyget kan bland annat uppfatta om responsen från enkäter är positiv, neutral eller negativ. Det som skiljer attitydanalys-verktygen åt är oftast algoritmen som används för att utföra analysen. Problemområdet som finns är att algoritmen som undersökts har svårt att tolka kontexter, förstå innebörden av orden och dess styrka samt inflytande i meningar. En annan problematik är att det finns begränsningar inom språket för attitydanalys-verktyget. De teoriområden som användes för att utreda problemområdet var bland annat Natural Language Processing, Natural Language Understanding, attitydanalys, lingvistik samt tidigare studier inom forskningsområdet. Metoderna som användes för att ta fram resultatet var en undersökning som utfördes med hjälp av fyra oberoende deltagare och ett attitydanalys-verktyg. Samt ett experiment där data manipulerades. Resultatet visade att algoritmen har stor andel felaktiga svar och det är något som behöver förbättras för att användningen av verktyget ska kunna användas i en större utsträckning. Studien bidrar till forskningsområdet och är ett försök att förbättra algoritmen som undersökts genom att bidra med ett lösningsförslag.

*Nyckelord:* Natural Language Processing, Natural Language Understanding, Attitydanalys, Digitalisering, Artificiell intelligens, Lingvistik.

## Abstract

The study examines the human way of analysing text compared to sentiment analysis tool. The aim of the study is to bring out the differences that occur during the comparison of a human analysis and sentiment analysis. Since analysis of surveys are becoming more automated, there is a wide range of different tools that offers services for analysing data. The sentiment analysis tool can decide among whether an answer from the survey is positive, neutral or negative. The range of tools is countless and what distinguishes them is usually the algorithm used to perform the analysis. The problem area that exists is that the algorithm has difficulty interpreting contexts and understand the meaning of the words as well as its strength and influence in sentences. Another problem is that there are limitations in the language of the sentiment analysis tool. The theoretical foundations used to investigate the problem area were, among others, Natural Language Processing, Natural Language Understanding, sentiment analysis, linguistics and previous studies in the field of research. The methods used to bring out the results was an investigation performed with the help of four independent participants and a sentiment analysis tool and an experiment where the data was manipulated. The result showed that the algorithm has a large proportion of incorrect answers and it is something that needs to be improved, in order to use this tool to a greater extent. The study contributes to the research area and is an attempt to improve the algorithm that have been analyzed in this study.

*Keywords:* Natural Language Processing, Natural Language Understanding, Sentiment Analysis, Digitization, Artificial Intelligence, Linguistics.

# ***Innehållsförteckning***

<b>1. INLEDNING</b>	6
1.1 Bakgrund	6
1.2 Problembeskrivning	7
1.3 Syfte	7
1.4 Forskningsfråga	7
1.5 Avgränsning	8
1.5.1 Motivering till avgränsning	8
<b>2. TEORI</b>	9
2.1 Begreppslista	9
2.2 Artificiell intelligens	9
<b>2.2.1 Utmaningar inom artificiell intelligens</b>	9
2.3 Maskininlärning	10
2.4 Natural Language Processing	10
2.5 Natural Language Understanding	11
2.6 Attitydanalys	11
2.7 Meaningcloud Sentiment Analysis	11
2.8 Enkätundersökningar	12
2.9 Lingvistik	12
2.10 Översättningar inom språkteorier	12
2.11 Tidigare forskning	12
<b>3. METOD</b>	15
3.1 Bakgrund	15
<b>3.1.1 Motivering till kvalitativ metod</b>	15
<b>3.1.2 Motivering till experiment</b>	15
<b>3.1.3 Motivering till användningen av MeaningCloud</b>	15

3.2. Litteraturstudie	16
3.3 Datainsamling	16
3.4 Tillvägagångssätt	16
<b>3.3.1 Körning av verktyg</b>	17
<b>4. RESULTAT</b>	20
4.1 Bakgrund	20
4.2 Redovisning av matupplevelsen	21
4.3 Redovisning av Golfupplevelsen.	26
4.3 Jämförelse	31
4.4 Sammanställning av resultat	32
4.5 Sammanställning av Matupplevelsen.	32
4.6 Sammanställning av Golfupplevelsen	33
4.7 Experiment	33
<b>5. ANALYS</b>	34
5.1. ARBETETS BIDRAG	35
<b>6. DISKUSSION</b>	38
<b>7. SLUTSATS</b>	39
7.1 Övervägande vid användning	39
<b>8. VIDARE FORSKNING</b>	40
<b>KÄLLOR</b>	41

# 1. INLEDNING

*Det här kapitlet presenterar arbetets bakgrund, problembeskrivning, syfte, forskningsfråga samt arbetets avgränsningar.*

## 1.1 Bakgrund

Dagens samhälle blir mer och mer automatiserat för varje dag. Processer och tjänster som tidigare utförts av människor ersätts av automatiserade, digitala lösningar. Regeringskansliet (2017) menar att artificiell intelligens, förkortat AI, har blivit en stor del av vardagen då den blir mer och mer integrerad med näringslivet samt forskning. Ett av det större användningsområdena för artificiell intelligens är beslutsstöd då systemen kan ta in en stor mängd data och analysera den för att sedan använda det som underlag vid beslutsfattande processer. Det finns också en möjlighet att förutse framtida händelser om historisk data kan kopplas till data som analyseras i realtid. Den stora fördelen med AI är att det går väldigt snabbt att analysera komplexa och stora mängder data.

Ovchinnikova (2012) beskriver ett område inom AI som benämns "Natural Language Understanding" (NLU), som fokuserar på tolkning av textfragment mellan datorer och människor. Inom NLU finns ett digitalt verktyg som benämns attitydanalys (eng. Sentiment Analysis). Liu (2012) beskriver att verktyget analyserar textkommentarer av människors åsikter och känslor rörande exempelvis nya produkter, tjänster eller problem. Det här verktyget kan appliceras på ostrukturerad data vilket är data som inte är lagrats enligt en standard struktur.

Liu (2011) beskriver att webbplatser är fyllda av ostrukturerade data. Det är väldigt viktigt att analysera ostrukturerade data, kanske till och med viktigare än att analysera strukturerad data. Detta på grund av den stora mängden data som finns tillgänglig samt dess variation. Feldman (2013) skriver att en person som vill köpa en produkt online med stor sannolikhet kommer börja söka efter recensioner och åsikter (ostrukturerade data) skrivna av andra. En följd blir att attitydanalys är ett av de hetaste forskningsämnena inom NLU idag. SCB (u.å) beskriver att enkätundersökningar har traditionellt sett varit det vanligaste tillvägagångssättet att samla in information. Liu (2012) menar att enkätundersökningar inte längre kommer behövas göras av företag nu när attitydanalys finns.

Företag har sedan länge använt sig av enkätundersökningar för att tyda vad deras kunder tycker om deras produkter och tjänster. Enkätundersökningar har även använts för framtagandet av nya produkter och tjänster. Behandlingen av enkätundersökningar har länge utgjort ett stort manuellt arbete. Idag finns det AI verktyg som möjligtvis kan underlätta dessa processer. Det finns många aspekter varför det kan vara till väldigt stor nytta för företag att använda AI istället för människor till den här typen av uppgifter. Den ekonomiska aspekten är central, färre arbetare för uppgiften (sett till större företag), mindre tid, färre mänskliga fel och så vidare.

## 1.2 Problembeskrivning

Problemområdet som finns är att den manuella hantering av enkäter är väldigt tidskrävande, speciellt för större företag då den ostrukturerade datamängden kan vara enorm. Under arbetets gång har en kontakt med ett företag etablerats. Företaget berättar att de i dagsläget behandlar mer än 10000 enkäter årligen. Liu, (2011) beskriver att ostrukturerad data kan vara väldigt värdefull då den är mer varierad och subjektiv i jämförelse med strukturerad data. Det finns automatiserade lösningar som kan underlätta detta, dock medför dessa automatiseringar andra påföljder som exempelvis svårigheter vid tolkning av kontext, att förstå innebörden av orden samt dess styrka och inflytande i meningar. En annan aspekt är att automatiserade lösningarna som används i bredare utsträckning endast behandlar poäng från en skala t.ex. 1-10. Dessa poäng kan ge företaget en bild av hur produkten/tjänsten generellt uppfattas, fast ger ingen indikator vad som är bra och vad som kan förbättras.

Verktyget attitydanalys erbjuder en mer subjektiv lösning som kan analysera attityder från fritext, så ingen ostrukturerad data försvinner. Tillämpningen av verktyget är inte felfritt då det finns begränsningar inom språket. Verktyget är inte tillräckligt utvecklat på det svenska språket vilket gör att översättningen från svenska till engelska krävs, den översättningen kan dessutom lätt bli felaktig och översättningar som är inkorrekta går inte att korrigera för framtida bruk. En attitydanalys-verktygets analys kan skilja sig drastiskt från en människas analys eftersom människors analyser kan vara enligt egna preferenser medan attitydanalysen utgår från klassificeringar. Svårigheterna inom området är att få en utförlig och nyanserad bild av kommentarer utan att göra ett manuellt arbete som är väldigt tidskrävande. Samtidigt som ingen viktig data förloras då attitydanalysen ger en generell bild.

Tidigare forskning gjord av Karde och Sonawane (2016) menar att det finns ett flertal problemområden inom attitydanalys. Karde och Sonawane identifierade några riktlinjer för problematiken med attitydanalys på twitter-inlägg/kommentarer. Problemet med den forskningen är att de endast undersökt twitter-inlägg/kommentarer vilket innebär en sträng begränsning på ord. Kommentarer på enkätundersökningar erbjuder i regel större utrymme att skriva och därför borde även dessa riktlinjer undersökas för attitydanalys-analyser gjorda på enkätundersökningar.

## 1.3 Syfte

Syftet med studien är att undersöka ett attitydanalys-verktyg förmåga att tolka text i förhållande till människa. Efter att undersökningen är gjord utreds vad som behövs förbättras med attitydanalys-verktyget som en människa i dagsläget gör bättre. Undersökningen kan bidra till forskningsområdet samt hjälpa framtida utveckling av bättre attitydanalys-verktyg. Verktyget har potential att automatisera tjänster som enkätundersökare.

## 1.4 Forskningsfråga

Varför skiljer sig mänskliga tolkningar av text från automatiserade tolkningar av text?



## 1.5 Avgränsning

Vi har valt att endast undersöka ett attitydanalys-verktyg samt endast haft ett begränsat antal oberoende deltagare på 4 personer i vår undersökning.

### 1.5.1 Motivering till avgränsning

Motiveringen är att det finns för många attitydanalys-verktyg för att undersöka samtliga. Skillnaden mellan verktygen är oftast den algoritmen som de använder för sin analys. Syftet med denna studie är inte att hitta den bästa attitydanalys-verktyg utan att identifiera generella skillnader mellan mänskliga tolkningar och ett attitydanalys-verktygs tolkningar hos just den algoritmen vi valt att undersöka. Anledningen till att det endast är 4 personer med i undersökningen är delvis för att undersökningen var väldigt tidskrävande vilket gjorde det svårt att få fler att engagera sig. Vid ett tidigt skede upptäcktes även att resultaten som samlats in i stort sett var identiska.

## 2. TEORI

Kapitlet presenterar den litteratur som ligger till grund för arbetet. Litteraturen består av teorier och tidigare publiceringar som främst berör tidigare forskning, NLP, NLU, Attitydanalys och AI. De teoretiska områden som valts ut används för att kunna beskriva och motivera det resultatet som framställs i studien samt vara som en grund för problemformuleringen.

### 2.1 Begreppslista

NLP	Natural Language Processing.
NLU	Natural Language Understanding.
Attitydanalys	Ett Natural Language Understanding-verktyg som bl.a. kan identifiera positiva och negativa meningar.
Naturligt språk	Det språk som människor kommunicerar informellt till varandra.
AI	Artificiell intelligens.
Lingvistik	Vetenskapen om det naturliga språkets strukturer och funktioner.

**Tabell 2.1.1: Begreppslista.**

### 2.2 Artificiell intelligens

Artificiell intelligens, även känt som AI, undersöker hur datorer ska kunna göra saker som människan i dagsläget gör bättre. Rich & Knight (1991) beskriver att syftet med artificiell intelligens är att efterlikna hjärnans förmåga att planera, dra slutsatser, inhämta ny kunskap, lösa problem och förstå naturligt språk. Ett intelligent program ska inte bara kunna dra egna slutsatser utan även lära sig av sina misstag för att optimera sina lösningar. De största tidigare användningsområdena inom AI menar Nilsson (1993) var Natural Language Processing, automatisk programmering, robotik, maskinvision, automatiserad teoriprövning och intelligenta datainsamlingssystem.

#### 2.2.1 Utmaningar inom artificiell intelligens

Mönsterigenkänning är en stor utmaning för AI, då människor till exempel lyckas känna igen ansikten oavsett ljus, tidpunkt, avstånd eller vinkel, detta är väldigt svårt att få en dator att göra (Balkenius, Skeppstedt, Gärdenfors, u.å). Davidov (2010) beskriver att ett av AIs och NLUs största utmaningar är tolkning av ironi och sarkasm. Dessa områden är så pass komplext att det till och med kan vara svårt för oss människor att förstå. Om det skulle finnas ett verktyg som direkt skulle kunna avgöra hurvida en person är ironisk/sarkastisk eller seriös så skulle det underlätta väldigt mycket för AI-verktyg som textsammanfattning, rankingssystem och dialogsystem.

### 2.2.2 Sverige och artificiell intelligens

Regeringskansliet (2018) redogör att “Sverige ska vara bäst i världen på att använda framtidens digitalisering och i internationell jämförelse ligger Sverige i absolut framkant”. Redan idag finns exempel på hur AI kan bidra till att bättre identifiera sjukdomar, effektivisera industriell produktion, minska energianvändning osv. AI-lösningar kan därmed bidra till ekonomisk tillväxt såväl som miljöutmaningar och sociala samhällsproblem. AI är ett brett område som kräver många tekniker, inte minst maskininlärning. Balkenius, Skeppstedt & Gärdenfors (u.å) menar att artificiell intelligens har sedan länge fått kritik för att dess forskare har gjort optimistiska uttalande om att AI-lösningar ska kunna förstärka och ersätta människor. Skeptiker har funnits med sedan tidigt stadie.

## 2.3 Maskininlärning

Mitchell (1997) berättar att alltsedan datorer blev uppfunna har människan undrat om datorer varit gjorda för att själva lära sig. Om vi människor skulle förstå hur vi ska programmera datorer för att datorerna själva ska kunna lära sig automatiskt av erfarenhet, så kan påverkan bli dramatisk. Enligt Skeppstedt (u.å) är maskininlärning studien av algoritmer för datorinlärning baserat på stora mängder data i syfte att göra goda förutsägelser och prognoser. Exempel på framgångsrik användning är självkörande fordon, autonoma robotar och framtagning av medicinska diagnoser.

## 2.4 Natural Language Processing

Natural Language Processing (NLP) handlar om design och implementering av datorprogram som kommunicerar med människor genom ett naturligt språk. Människor kommunicerar lättast och mest effektivt med varandra genom att använda sitt naturliga språk. Forskning inom NLP går främst ut på att designa input/output-komponenter av artificiell intelligens som gör att systemen kan använda språk lika flytande och välfungerande som människor. (Dale, Somers, Moisl, 2000). Liddy (2001) beskriver att målet med NLP handlar om att uppnå mänsklig språkbehandling med fokus på bearbetning. NLP konstaterades som en viktig del vid de första utvecklingsstadierna av AI. Det har satts upp mål för NLP, dock är dessa inte uppfyllda än men det är något som man strävar mot. Mål 1-3 går att uppfylla med endast NLP, men för att kunna se och hitta kopplingarna som finns i mål fyra krävs det NLU.

Målen är:

- 1: parafrasera en inmatad text
- 2: översätta texten till ett annat språk
- 3: besvara frågor om innehållet i texten
- 4: hitta kopplingar från texten

## 2.5 Natural Language Understanding

Ovchinnikova (2012) beskriver att Natural Language Understanding (NLU) är en del av Natural Language Processing (NLP) som hanterar maskinläsningskompetens. Syftet med NLU är att tolka ett inmatat textfragment för att översätta en text från ett naturligt språk till en representant i ett entydigt formellt språk. Representationen av det formella språket ska uttrycka textens innehåll samt hjälpa till vid utförandet av konkreta uppgifter som har blivit outtalade vid användarförfrågan. Vid användning av datorprogram går det att se över processen på en mer detaljerad nivå och ta fram representationen av meddelandet som programmet skapade av sig själv. Utvärderingen av det automatiskt skapade meddelandet inom processen görs genom en jämförelse mot mänsklig intuition om innehållet i det behandlade textfragmentet.

## 2.6 Attitydanalys

Attitydanalys kommer från engelskans "Sentiment Analysis" och har växt till ett av de största forskningsområdena inom Natural Language Processing. Liu (2012) berättar att attitydanalys går ut på att analysera människors åsikter, attityder och känslor gentemot produkter, tjänster, problem och händelser. Attitydanalys omfattar ett stort problemområde. Tystarau och Palpanas (2011) beskriver att attitydanalys skapades för att vara ett verktyg till NLP i syfte med att identifiera känslor i uttryckta texter. Attitydanalysen gör mer specifika identifieringar av subjektiva uttalanden, inneboende åsikter och känslor. Författarna beskriver att attitydanalys har fått ett omfattande forskningsintresse de senaste åren. Vilket har gjort det möjligt att klassificera metoder inom området som delas upp i fyra kategorier: Maskininläring, ordboksbaseade, statistiska och semantiska tillvägagångssätt.

Liu (2012) beskriver att vid appliceringar av attitydanalys är åsikter centrala i nästan all mänsklig aktivitet eftersom åsikter influerar vårt beteende. Vi människor rådfrågar nästan alltid vid ett beslut och vill gärna veta vad andra tycker innan ett beslut tas. Från att tidigare endast rådfrågat familj och vänner så finns nu recensioner och kommentarer från en större mängd människor. Genom att applicera attitydanalys på sociala medier så behöver företag inte längre göra enkätundersökningar eller åsiktsundersökningar eftersom det redan finns tillgänglig data att analysera.

## 2.7 Meaningcloud Sentiment Analysis

Enligt MeaningCloud (2019) erbjuder Sentiment Analysis flerspråkig attitydanalys som går att applicera flera olika källor. Verktyget identifierar positiva, negativa och neutrala meningar både på sociala medier och enkäter. MeaningClouds Algoritm extraherar aspektbaserade attityder och kan enligt MeaningCloud själva upptäcka ironi och skilja fakta från åsikter. Skillnaderna mellan de olika meningarna utvärderas separat av algoritmen, sedan skapas en helhetsbild för texten.

## 2.8 Enkätundersökningar

Enligt SCB (u.å) är enkätundersökningar det vanligaste sättet att fråga efter information. Utskick av enkäter via post eller e-post kan vara tidskrävande, det är dessutom viktigt att motivet med undersökningen framgår och beskrivning av hur enkäten ska fyllas i. Enkäter som sker på webbplatser är allt vanligare eftersom det blir billigare och enklare än att göra postutskick.

## 2.9 Lingvistik

Lingvistik är vetenskapen om det naturliga språkets strukturer och funktioner som utgör det teoretiska grundvalen för forskning inom enskilda språk. Termen lingvistik är en synonym till ordet språkvetenskap som handlar om forskning och den organiserade kunskapen om språk samt användningen av språk. I Sverige berör lingvistiken förutom den allmänna språkvetenskapen, också fonetiken vilket är vetenskapen om talet eller talljudet. Lingvistik kan också beskrivas som strukturell lingvistik. Det är språkvetenskapen som ser språken som ett system av olika enheter eller strukturer på olika nivåer (Sigrud, u.å). *Linguistics Journal* (2019) beskriver att lingvistik består av studier av strukturer och utvecklingen av ett språk och dess förhållande till andra språk, sambandet mellan sinnen och sambandet mellan språk och samhälle. Manning & Schütze (1999) nämner att det enklaste exemplet på användning av lingvistik är att ord kan delas in i olika ordklasser som anger om ett ord är substantiv, adjektiv eller verb.

## 2.10 Översättningar inom språkteorier

Wollin (u.å) definierar översättning inom språkteorier som en överföring av ett budskap från ett språk till ett annat som vanligtvis sker i skriftlig form. Vid en översättning finns det ett källspråk som är utgångsspråket och sedan ett målspråk som är det språket som översättningen sker till. De två olika typerna av språk har olika grammatiska system samt olika lexikala strukturer där en översättning innebär att en språklig omfattning av vidare räckvidd är det enklare utskriften av enstaka ord och fraser. Det innebär också att en textnära överföring är möjlig, dock kan det kulturella skillnaderna mellan språken göra att det blir svårbegripligt samt kräva att översättningen inkluderar förklarande begrepp för att göra det mer begripligt. Teoretisk är det möjligt att göra en översättning läsbar och kan avläsas i grader av ekvivalens som betyder likvärdighet mellan källtexten och måltextern.

## 2.1 Tidigare forskning

Glant (2018) menar att data som finns på svenska kan översättas till engelska vid användning av befintliga engelska attitydanalys-verktyg. För att få verktyget att uppnå en optimal klassificeringsförmåga krävs det 33% mer träningsdata för att få ett omdöme. Träningsdata är den data som algoritmen använder för att lära upp sig och bli bättre. Trots den något större mängd data som krävs för att uppnå optimal klassificeringsförmåga, är det fullt möjligt att använda ett engelskt verktyg för svenska produktomdömen. Det blir sannolikt eftersom det finns mycket tillgänglig data på engelska som gör att användningen av ett engelskt attitydanalys-verktyg tillsammans med maskinöversättning blir ett rimligt alternativ för bearbetning av produktomdömen på svenska.

Taboada (2016) menar att stora delar av tal förmedlas med hjälp av substantiv som exempelvis mästerverk och katastrof, genom verb som till exempel älskar och hatar, eller genom adverb som dålig och skicklig. Detta blir intressant då verben används mer exklusivt och i somliga sammanhang uttrycks verben som positiva eller negativa adjektiv. Ordböcker innehåller listor på positiva och negativa ord som inkluderar en beskrivning av styrkan på orden. Detta underlättar för attitydanalys som baseras på ordböcker och lexikon. Exempelvis ordet ”bra” är svagt positivt medan ordet ”utmärkt” är starkt positivt. Ordet ”dåligt” är starkt negativt, ”förfärligt” är också ett starkt negativt ord.

De klassificeringar som finns på orden är:

- Stark Positiv
- Svag Positiv
- Neutral
- Svag Negativ
- Stark Negativ

Intensifiering och nedtoning som författaren nämner i rapporten går att modifiera alla delar av tal. Författaren menar att modifieringen har konsekvent visat att det förbättrar prestandan för attitydanalys. Genom att använda en metod som innefattar multiplikation går det att mäta och väga om orden är positiva eller negativa. Intensifiering genomförs genom att presentera de orden som ofta återkommer i feedbacksammanhang med procentsatser som används vid övervägningen. Procentsatsen av orden är:

- Mest + 100%
- Verkligen + 25%
- Våldigt + 15%
- Något - 30%
- Argumenterbart - 20%

Karde och Sonawane (2016) nämner att Attitydanalys underlättar i många sammanhang dock några utmaningar inom attitydanalys. Dessa utmaningar är listade nedan:

- Identifieringen: Subjektiva delar från en text eller ett ord som använd inom ett objekt, går att användas för att beskriva en känsla
- Domänberoende: Ett ord kan vara positivt inom vissa domäner/sammanhang men negativt i ett annat sammanhang.
- Sarkasmdetektion: Attitydanalys-verktyg har svårt att förstå när en människa uttrycker något positivt, men menar det på ett negativt sätt.
- Motsatta uttryck: Attitydanalys-verktyg har svårt att förstå ett negativt omdöme om det är skrivet positivt. Exempelvis när en person beskriver “Den här filmen borde vara fantastisk” kan verktyget uppfatta det som något positivt trots det är negativt.
- Följdordningen: Kan spela roll för hur verktyget tolkar en mening, A är bättre än B eller B är sämre än A innebär att verktyget kan tolka positivt i första meningen medan det tolkar negativt i andra meningen
- Entitetigenkänning: Är en utmaning då det oftast är utifrån specifika entiteter det går få ut kvalitativ data ifrån, därför måste analysen göras utifrån varje entitet.
- Jämförelser: Kan vara en utmaning då verktyget inte känner till bakgrunden. Att säga A var bättre än B behöver varken betyda att det är positivt eller negativt. Exempelvis “Det var bättre än på stenåldern” kan vara svårt att tyda då det varken är positivt eller negativt.

## 3. METOD

*Det här kapitlet presenterar bakgrunden till arbetet, vald forskningsmetod, hur informationsinsamlingen gått till samt en motivering till varför forskningsmetoden valdes.*

### 3.1 Bakgrund

För att utföra vår forskning på ett lämpligt sätt valde vi att göra en kvalitativ fallstudie samt en undersökning och ett experiment. Enligt Scott, J (under utgivning) är en fallstudie en forskningsdesign som tar ut ett enskilt fall hos ett fåtal individer inom en viss social entitet, sedan appliceras metoder för att undersöka dem. Vi har valt att göra en fallstudie för eftersom vi ville undersöka ett visst fenomen inom automatisering. Experimentet gjordes genom att manipulera felaktig analys gjorda av attitydanalys-verktyget. Jämförelsen kategoriseras i positivt, negativt eller neutralt. Eftersom varje människa har olika preferenser för vad som är positivt, negativt och neutralt har vi även gjort en undersökning på människors analyser av kommentarer mot en körning gjord av verktyget. Denscombe (2018) beskriver att fokus vid ett experiment är att besvara frågor som varför saker händer och hur det hänger ihop. Detta är något som vi följt i vårt experiment för att säkerställa att det utförs på ett korrekt sätt. En bild från experimentet har även redovisats för att ge en tydligare bild.

#### 3.1.1 Motivering till kvalitativ metod

Vid kvalitativa metoder försöker forskare att fånga in användarens handlingar i syfte med att tolka samt förstå dem. För att lyckas fånga in användarens handlingar på detta sätt krävs det byggstenar som ömsesidig tillit och en förståelse mellan forskare och användare (Holme & Krohn Solvang, 1997) Vi valde att använda en kvalitativ metod för att få en större förståelse för våra deltagare och lättare kunna identifiera skillnader uppkommer vid jämförelsen av människans analys av text och attitydanalys-verktygets analys.

#### 3.1.2 Motivering till experiment

Denscombe (2018) beskriver att syftet med experiment är att upptäcka nya förhållanden och egenskaper mellan specifika faktorer. Då studien undersökt egenskaperna som attitydanalys erbjuder samt varför människan i dagsläget utför bättre analyser, finner vi att ett experiment varit en lämplig undersökningsmetod. Förhoppningen med experimentet var att ta fram de klassificeringar som blir fel under en körning av attitydanalys-verktyget.

#### 3.1.3 Motivering till användningen av MeaningCloud

En av de främsta anledningarna att vi valde MeaningCloud var att verktyget hade ett inbyggt översättningssystem. Andra anledningar var att MeaningCloud erbjöd en gratis provperiod, gick att applicera direkt på Microsoft Excel och hade flera bra specifikationer jämfört med många andra verktyg vi undersökte. Dock fungerade inte den inbyggda svenska översättningen, vilket upptäcktes i ett senare skede. Det gjorde att all data fick översättas manuellt i programmet Google translate, för att sedan göra en körning av algoritmen.



## 3.2. Litteraturstudie

Den tidigare litteratur som använts är baserad på Grants (2018) examensarbete 2018, där han motiverat att det går att använda ett engelskt attitydanalys-verktyg på svenska produktomdömen. Taboada (2016) beskriver styrka och klassificering av ord vid attitydanalys. Taboda nämner också intensifiering och nedtoning. Karde och Sonawane (2016) Beskriver utmaningarna som finns inom attitydanalys. Dessa studier användes var och en för att forma en generell tolkning till vad som leder till felaktiga analyser. I Grants examensarbete fann vi forskning gjord på möjligheterna och problemen med maskinöversatt text från svenska till engelska. I Taboadas arbete fann vi en modell för hur intensifiering av ord kan användas för att förbättra attitydanalys-verktyg. I Karde och Sonawanes forskning fann vi riktlinjer för vad som oftast blir fel vid attitydanalys-tolkningar.

## 3.3 Datainsamling

Datainsamlingen skedde genom att de mänskliga deltagarna fick ta del av Microsoft-excel filen där svaren från enkätundersökningarna fanns sparade. De blev sedan försedda med penna och papper och fick göra bedömningar på kommentarerna som företagets kunder hade skrivit. Deltagarna klassade kommentarerna som positiva, negativa eller neutrala. Attitydanalys-verktyget som testades erbjöd en så kallad "plug-in" till Microsoft Excel som är programkomponent som lägger till funktioner till ett befintligt datorprogram. Detta gjorde att enkätundersökningarna kunde integrera med verktyget på ett simpelt sätt och ingen data gick förlorad. Algoritmen i verktyget gjorde sedan sin bedömning av kommentarerna och klassade dem på samma sätt som det mänskliga deltagarna. Svaren som det mänskliga deltagaren och verktyget gav, sammanställdes och jämfördes sedan med varandra för att kunna framställa ett resultat.

## 3.4 Tillvägagångssätt

Första steget i processen var att sammanställa data från datasetet och avgränsa den. Studien avgränsas genom att vi valde att fokusera på två olika områden. Dessa områden är Matupplevelsen och Golfupplevelsen hos fyra olika hotell som företaget vi varit i kontakt med anordnar resor till. Vi använde oss av de kommentarer som företaget samlat in om hotellen/golfen genom enkätundersökningar. Då enkäterna är omfattande fokuserade vi endast på "ge oss gärna en kommentar gällande måltiderna och restaurangen" vid matupplevelsen och "ge oss gärna en kommentar gällande golfen" vid golfupplevelsen.

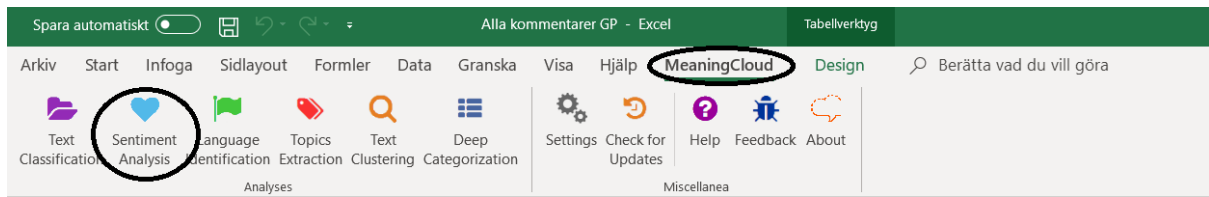
För att öka kvalitetssäkringen med undersökningsdeltagarna, samt få dem att känna sig trygga med formatet så förklarades ändamålet med undersökningen och visade hur svaren från undersökningen skulle behandlas. Deltagarna försäkrades även om att deras svar var anonyma och att inte finns någon möjlighet att koppla deras svar tillbaka till dem.

Deltagare blev försedda med penna och papper och sedan blev de placerade framför ett antal recensioner och fick i uppdrag att analysera dessa svar och kategorisera dem

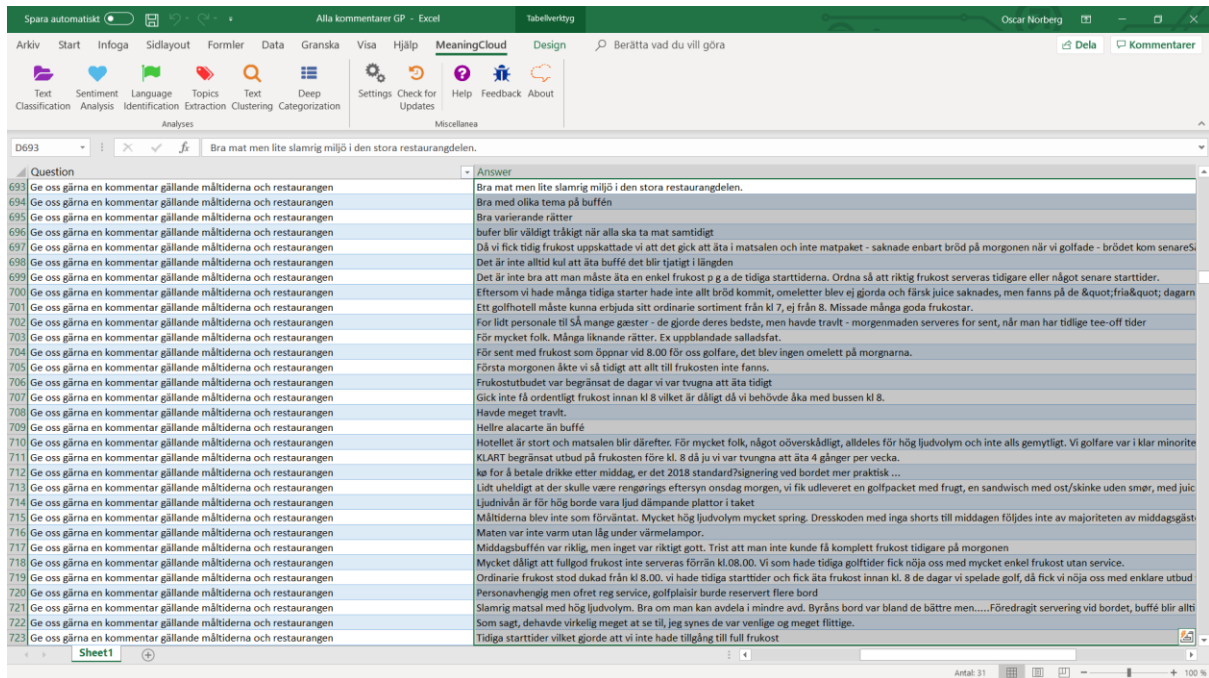
Mätningen av undersökningen genomfördes genom att sortera data efter tre kategorier. Kategorierna var positivt, negativt och neutralt. Motiveringen till kategorivalet var att attitydanalys-verktyget vi använde oss av använde den kategoriseringen. Därför behövde även deltagarna använda sig av samma kategoriseringssystem för att göra undersökningen legitim. Datat som var i form av kommentarer granskades av fyra oberoende deltagare som gjorde sin personliga analys av data baserat på egna erfarenheter och åsikter. Deltagarna fick inte samtala under undersökningen, för att de inte skulle ha någon möjlighet att påverka varandra. Det gjorde att resultatet blev opartiskt och lättare att analysera. Attitydanalys verktyget fick sedan göra sin genomgång på datat och utföra sin egen analys.

Mätningen av undersökningen skedde genom att resultatet sammanställdes och en jämförelse genomfördes i Microsoft Excel. Jämförelsen gick ut på att gå igenom alla kommentarer för att se om människa och attitydanalys-verktyg tolkat kommentarerna på liknande sätt. Då människans svar representerades av fyra olika deltagare tog vi det genomsnittliga svaret. De skillnader och likheter som fanns beräknades och presenterades sedan i form av tabeller och diagram. Presentationen av resultatet delades upp i två delar där den första delen berörde Matupplevelsen och den andra delen handlade om Golfupplevelsen. Vid sammanfattningen av resultatet ställde vi en fråga till våra deltagare som lydde "Hur tycker ni det var att göra den här analysen?" Därefter sammanställdes svaren på papper som vi sedan sammanfattade under resultatet. Analysen av resultatet gjordes genom en jämförelse där vårt resultat tolkades utifrån de felmönster som kunde identifieras, efter felmönstret identifierats så gjordes en motivering till felet utifrån den tidigare forskning samt teori som använts.

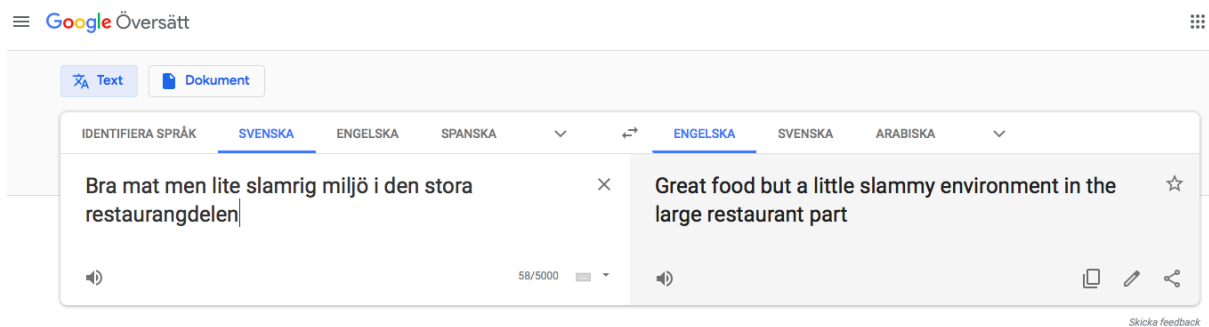
### 3.3.1 Körning av verktyg



Figur 3.3.2: Figuren visar hur verktyget integreras med Microsoft Excel.



Figur 3.3.3: Figuren visar hur selekteringen av kommentarer som skall analyseras sker.



Figur 3.3.4: Visar ett exempel på den manuella översättningen av kommentarer i Google translate.

	A	B	C	D	E	F	G
	Text	Polarity	Agreement	Subjectivity	Confidence	Irony	
2	Too boring food for a fourteen day trip! More variety is required, especially week two.	N	AGREEMENT	SUBJECTIVE	100	NONIRONIC	
3	The unweiling was bad for several days	N	AGREEMENT	SUBJECTIVE	100	NONIRONIC	
4	Limited breakfast times due to early start times.	P	AGREEMENT	OBJECTIVE	100	NONIRONIC	
5	Great food but a little slummy environment in the large restaurant part.	P+	AGREEMENT	OBJECTIVE	100	NONIRONIC	
6	Good with different themes on the buffet	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC	
7	Good variety of dishes	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC	
8	buffer becomes very boring when everyone is going to eat food at the same time	N	AGREEMENT	SUBJECTIVE	100	NONIRONIC	
	When we got early breakfast we appreciated that it was possible to eat in the dining room and not food parcels - lacking only bread in the morning when we golfed - the bread came later	P	DISAGREEMENT	SUBJECTIVE	94	NONIRONIC	
	The way with the notes and the notification for payment was complicated and people were generally very annoyed at this in the dining room - no big deal for us						
9	Good varied food at dinner - nothing to complain about						
10	It is not always fun to eat the buffet it gets boring in the long run	N	AGREEMENT	SUBJECTIVE	89	NONIRONIC	
	It is not good that you have to eat a simple breakfast because of the early start times. Arrange so that proper breakfast is served earlier or some later start times.	NEU	DISAGREEMENT	SUBJECTIVE	86	NONIRONIC	
11	Since we had many early starts, not all bread had come, omelets were not made and fresh juices were missing, but were on the "free" days. Friendly staff!	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC	
12	A golf hotel must be able to offer its regular range from 7 am, not from 8. Missed many good breakfasts	N	AGREEMENT	SUBJECTIVE	100	NONIRONIC	
13	Too little staff for so many guests - they did their best but were busy	P	DISAGREEMENT	SUBJECTIVE	94	NONIRONIC	
14	- Breakfast is served late when you have early tee-off times						
15	Too many people. Many similar dishes. Ex mixed salad dishes.	N	AGREEMENT	OBJECTIVE	100	NONIRONIC	
16	Too late with breakfast opening at 8.00 for us golfers, there was no omelette in the mornings.	N	AGREEMENT	SUBJECTIVE	100	NONIRONIC	
17	The first morning we went so early that everything for breakfast was not there.	NONE	AGREEMENT	OBJECTIVE	100	NONIRONIC	
18	The breakfast was limited the days we had to eat early	NONE	AGREEMENT	OBJECTIVE	100	NONIRONIC	

**Figur 3.3.5:** Figuren visar alla kommentarer som är översatta och efter ett klick på "Sentiment Analysis"-knappen började verktyget bearbeta frågorna och efter några sekunder gav attitydanalys-verktyget ett resultat på de olika kommentarerna. Den gröna markeringen med bokstaven P är meningar som är positiva, röda markeringen med bokstaven N innebär negativa meningar och gråa markeringen med bokstäverna NEU innebär neutrala meningar

Efter att sammanställningen och jämförelsen av datat var gjord så började vi experimentera med verktyget. Vi manipulerade parametrar i verktyget för att kunna urskilja vad de felaktiga analyserna berodde på. Manipuleringen gjordes genom att lägga till och byta ut deskriptiva ord som kommentarerna bestod av för att sedan göra nya körningar med verktyget. Syftet med experimentet var att få en förståelse för när klassificeringarna blir rätt och fel. Vi analyserade vår data genom att undersöka vår sammanställning och vårt experiment utifrån den teori och tidigare forskning som finns inom attitydanalys.

## 4. RESULTAT

### 4.1 Bakgrund

Arbetet har utförts tillsammans med företaget Golf Plaisir AB som har delat med sig av information och data. Kontakten med företaget etablerades eftersom de hade ett problemområde som passar väl mot det syfte och den forskningsfråga som denna uppsats behandlar.

Under möten med företaget framgick det att de vill hantera och utvärdera den stora mängd enkäter som de årligen samlar in från sina resenärer på ett bättre samt smidigare sätt. De förklarade att i dagsläget är det ca 18 000 resenärer årligen och de behandlar ca 10 000 enkäter från dessa resenärer. Enkäten täcker ett flertal områden så som t.ex. resenärens uppfattning om boendet, maten och golfen som är delar av den produkt som de tillhandahåller sina kunder. Enkäten distribueras ut via e-post och en länk till ett frågeformulär bifogas detta mail. Enkäten består av ett antal frågor där respondenten skall betygsätta olika aspekter av respektive område med en poängsättning mellan 1-10. Baserat på de poäng respondenten ger inom de olika områdena räknas en snittpoäng fram. Ju högre poäng desto högre nöjdhet. Denna behandling görs automatiskt av enkätprogrammet som används. En allmän bedömning av resenärens nöjdhet framkommer genom detta - men ger inga svar på vad eller varför kunden ger en viss bedömning.

Fallet har varit att behandla ett stort dataset som Golf Plaisir har försett. Datasetet bestod av cirka 3000 kommentarer från 6 olika destinationer. Datasetet bestod av korta kommentarer som svarade på frågor rörande vistelsen. Begränsningen för varje kommentar var 100 tecken och formatet var endast webbenkäter. Den nuvarande behandlingen av datasetet har varit att personalen läser igenom samtliga kommentarer som består av stora mängder fritextsvar vilket gör att analys och utvärdering blir mycket tidskrävande och komplex

Företaget Golf Plaisir är intresserade av att veta om man genom analysera all den fritext man samlar in kan nyttja ett NLU-verktyg för att på ett enklare och mer kostnadseffektivt sätt identifiera vad och varför deras resenärer tycker något är bra eller dåligt. Man vill även kunna få bättre överblick och sammanställning av vad resenären uppfattar som positivt, negativt eller är neutralt.

## 4.2 Redovisning av matupplevelsen

Hotell	Antal kommentarer
Hotell 1	12
Hotell 2	29
Hotell 3	21
Hotell 4	39

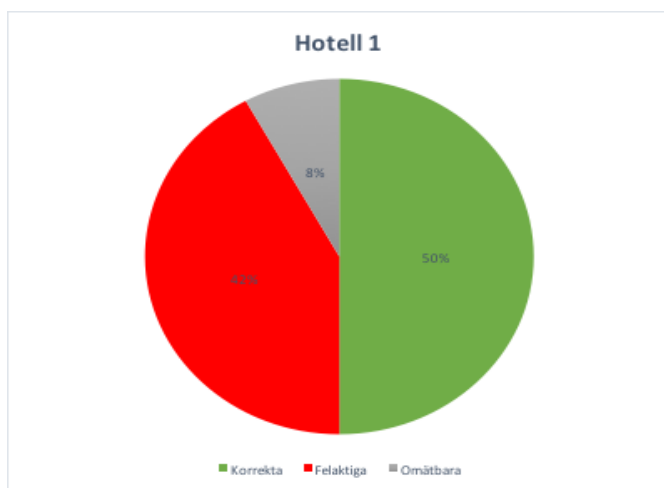
**Tabell 4.2.1:** Antalet kommentarer angående matupplevelsen.



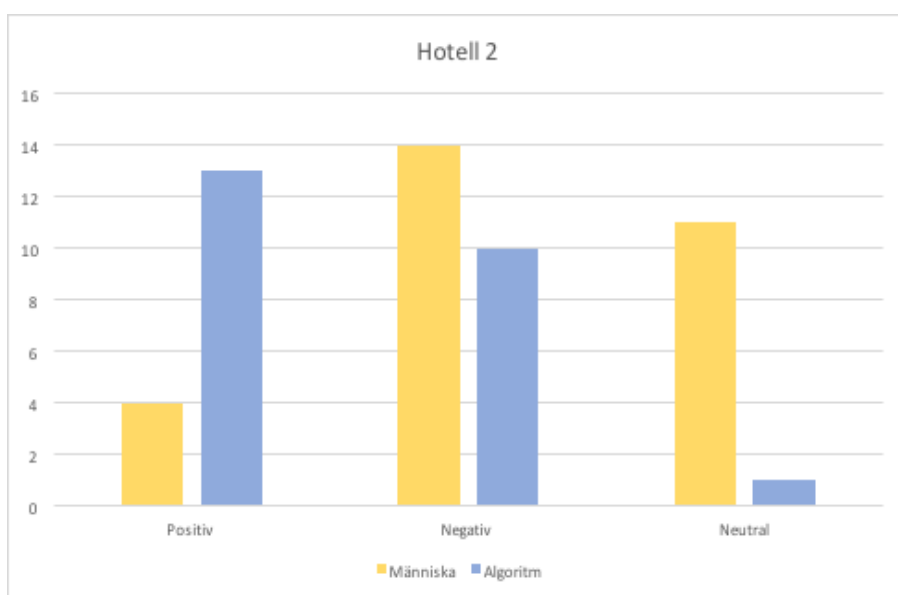
**Figur 4.2.2:** Fördelningen av Positiva, Negativa och Neutrala kommentarer mellan människa och algoritm. De kommentarer som algoritmen inte kunde mäta och klassificerades som NONE är borträknade.

Hotell 1	Människa	Algoritm	Avvikelser
Positiv	7	8	1
Negativ	2	2	0
Neutral	3	1	2
Omätbara	0	1	1

**Tabell 4.2.3:** Sammanställning av positiva, negativa, neutrala och omätbara svar av människa och algoritm. Kolumnen som heter avvikelser visar antalet svar som skiljer sig mellan människan och algoritmen.



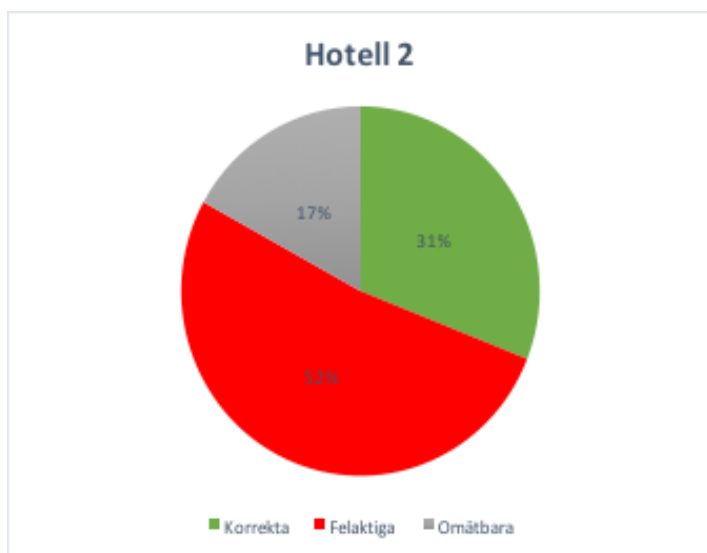
**Figur 4.2.4:** Resultatet av antalet korrekta, felaktiga och omätbara mätningar som algoritmen gjort i jämförelse med människan.



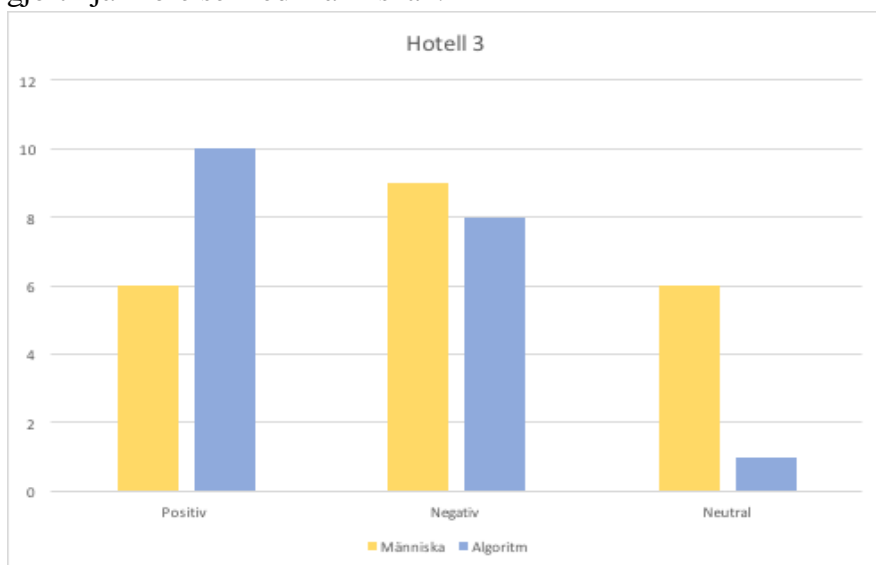
**Figur 4.2.5:** Fördelningen av Positiva, Negativa och Neutrala kommentarer mellan människa och algoritm. De kommentarer som algoritmen inte kunde mäta och klassificerades som NONE är borträknade.

Hotell 2	Människa	Algoritm	Avvikelser
Positiv	4	13	9
Negativ	14	10	4
Neutral	11	1	10
Omätbara	0	5	5

**Tabell 4.2.6:** Sammanställning av positiva, negativa, neutrala och omätbara svar av människa och algoritm. Kolumnen som heter avvikelser visar antalet svar som skiljer sig mellan människan och algoritmen.



**Figur 4.2.7:** Resultatet av antalet korrekta, felaktiga och omätbara mätningar som algoritmen gjort i jämförelse med människan.

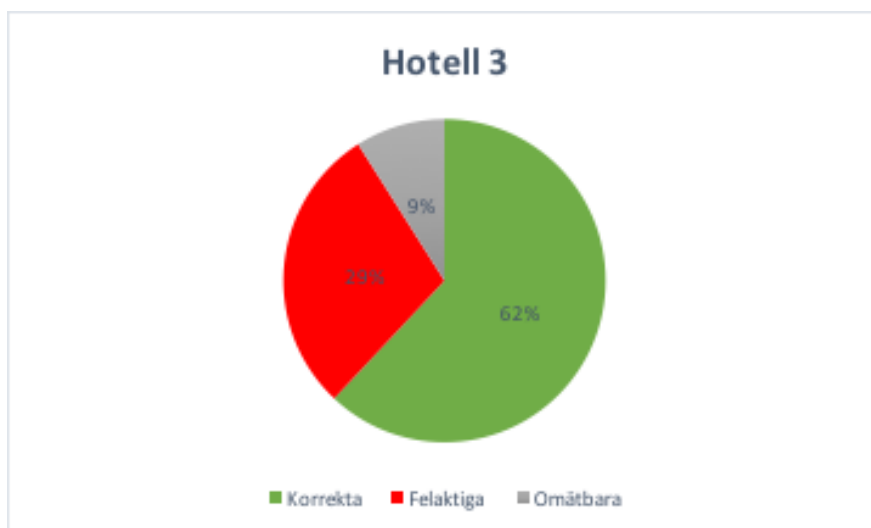


**Figur 4.2.8:** Fördelningen av Positiva, Negativa och Neutrala kommentarer mellan människa och algoritm. De kommentarer som algoritmen inte kunde mäta och klassificerades som NONE är borträknade. .

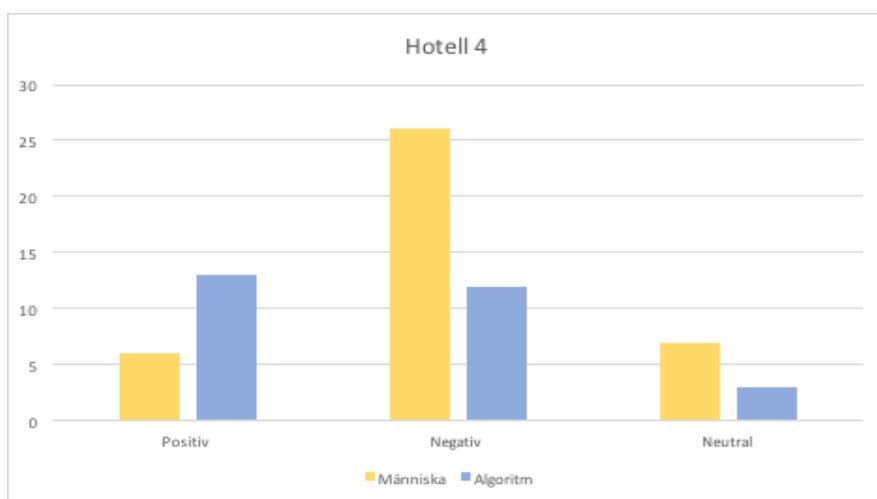
Hotell 3	Människa	Algoritm	Avvikelser
Positiv	6	10	4
Negativ	9	8	1
Neutral	6	1	5
Omätbara	0	2	2

**Tabell 4.2.9:** Sammanställning av positiva, negativa, neutrala och omätbara svar av människa och algoritm. Kolumnen som heter avvikelser visar antalet svar som skiljer sig mellan människan och algoritmen.





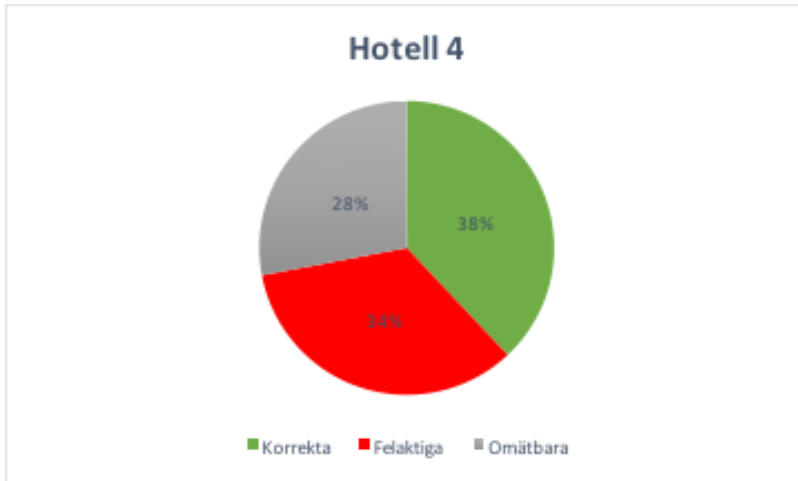
**Figur 4.2.10:** Resultatet av antalet korrekta, felaktiga och omätbara mätningar som algoritmen gjort i jämförelse med människan.



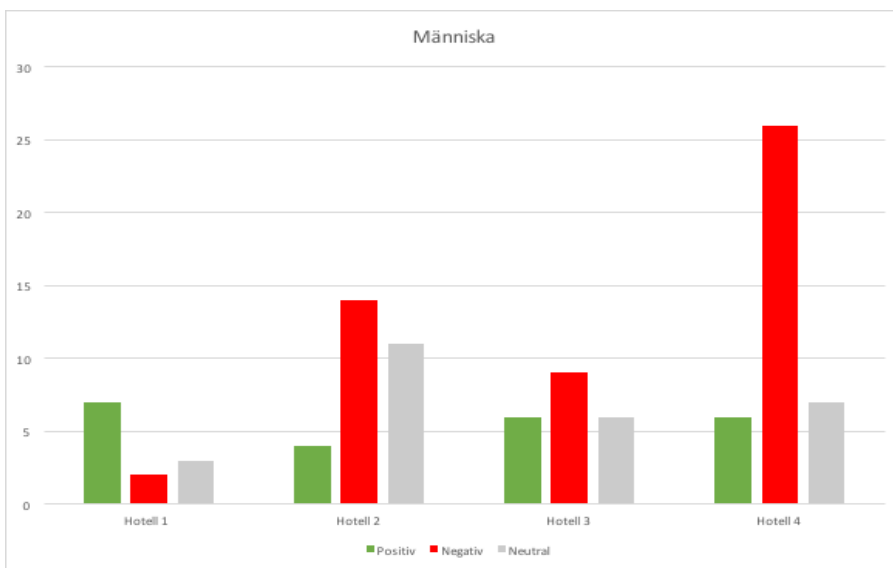
**Figur 4.2.11:** Fördelningen av Positiva, Negativa och Neutrala kommentarer mellan människa och algoritm. De kommentarer som algoritmen inte kunde mäta och klassificerades som NONE är borträknade.

Hotell 4	Människa	Algoritm	Avvikelser
Positiv	6	13	7
Negativ	26	12	14
Neutral	7	3	4
Omätbara	0	11	11

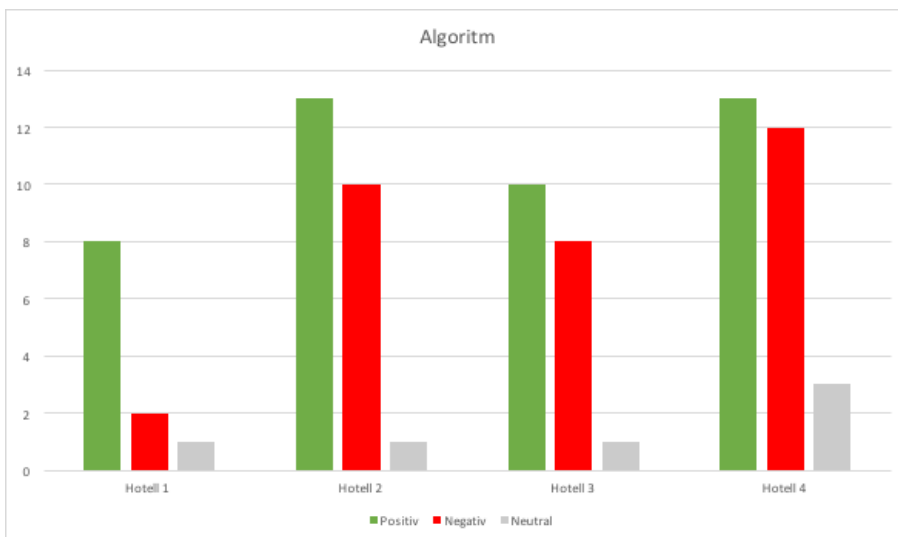
**Tabell 4.2.12:** Sammanställning av positiva, negativa, neutrala och omätbara svar av människa och algoritm. Kolumnen som heter avvikelser visar antalet svar som skiljer sig mellan människan och algoritmen.



**Figur 4.2.13:** Resultatet av antalet korrekta, felaktiga och omätbara mätningar som algoritmen gjort i jämförelse med människan.



**Figur 4.2.14:** Sammanställning av positiva, negativa och neutrala svar av människan.

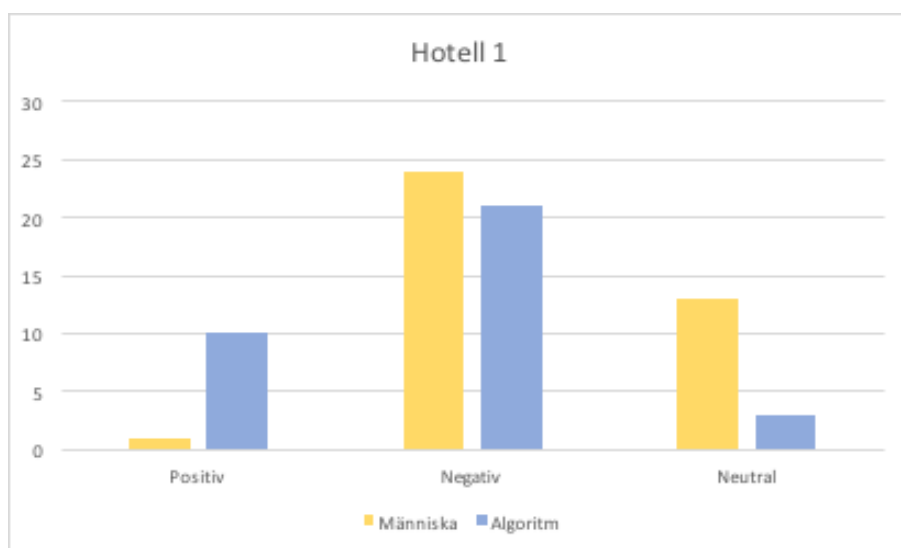


**Figur 4.2.15:** Sammanställning av positiva, negativa och neutrala svar av algoritmen.

### 4.3 Redovisning av Golfupplevelsen.

Hotell	Antal kommentarer
Hotell 1	38
Hotell 2	39
Hotell 3	23
Hotell 4	40

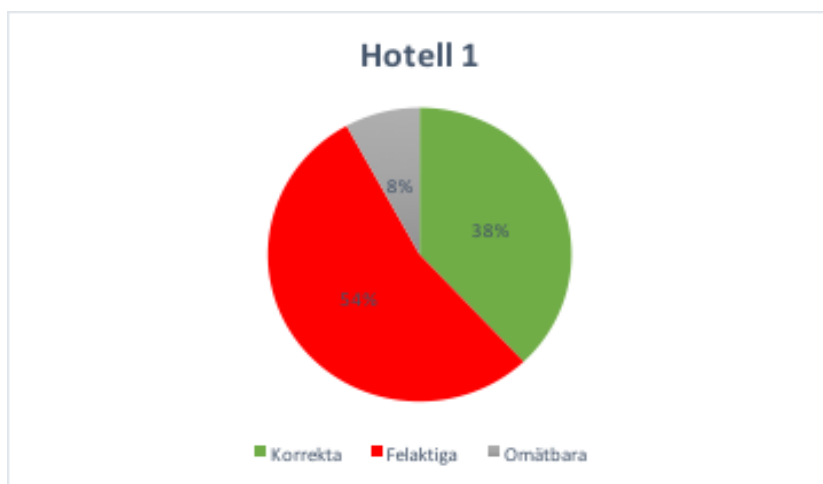
**Tabell 4.2.16:** Redovisning av antalet kommentarer angående golfupplevelser.



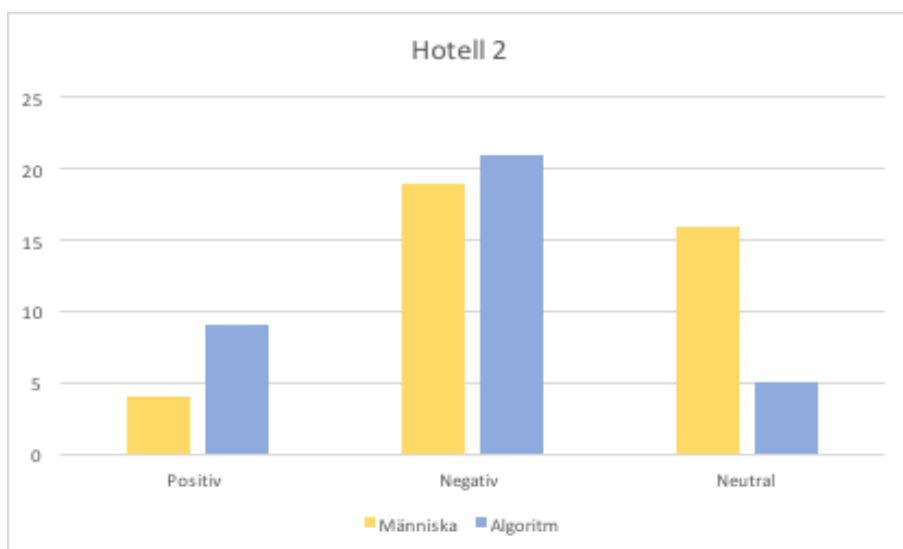
**Figur 4.2.17:**Fördelningen av Positiva, Negativa och Neutrala kommentarer mellan människa och algoritm. De kommentarer som algoritmen inte kunde mäta och klassificerades som NONE är borträknade.

Hotell 1	Människa	Algoritm	Avvikelser
Positiv	1	10	9
Negativ	24	21	3
Neutral	13	3	10
Omätbara	0	4	4

**Tabell 4.2.18:** Sammanställning av positiva, negativa, neutrala och omätbara svar av människa och algoritm. Kolumnen som heter avvikelser visar antalet svar som skiljer sig mellan människan och algoritmen.



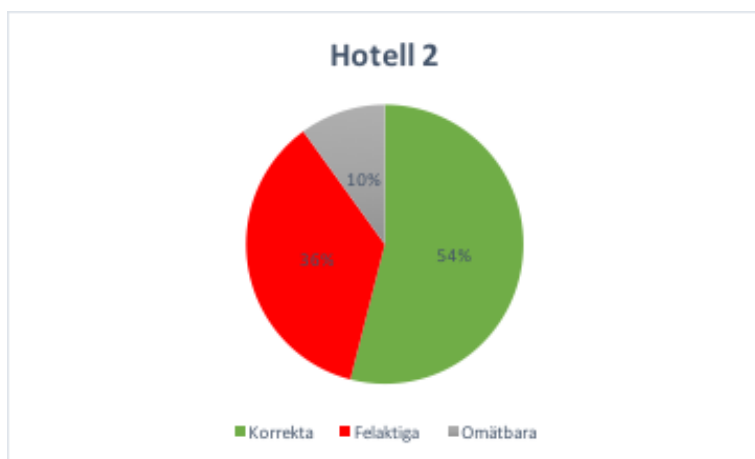
**Figur 4.2.19:** Resultatet av antalet korrekta, felaktiga och omätbara mätningar som algoritmen gjort i jämförelse med människan.



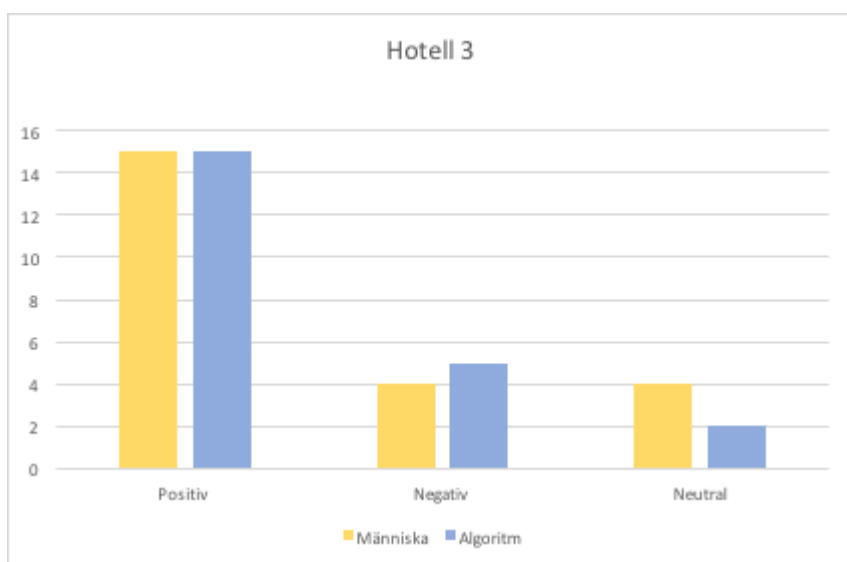
**Figur 4.2.20:** Fördelningen av Positiva, Negativa och Neutrala kommentarer mellan människa och algoritm. De kommentarer som algoritmen inte kunde mäta och klassificerades som NONE är borträknade.

Hotell 2	Människa	Algoritm	Avvikelser
Positiv	4	9	5
Negativ	19	21	2
Neutral	16	5	9
Omätbara	0	4	4

**Tabell 4.2.21:** Sammanställning av positiva, negativa, neutrala och omätbara svar av människa och algoritm. Kolumnen som heter avvikelser visar antalet svar som skiljer sig mellan människan och algoritmen.



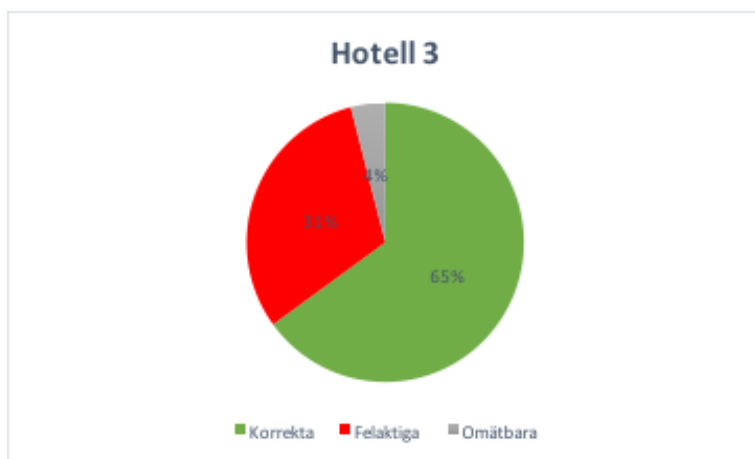
**Figur 4.2.22:** Resultatet av antalet korrekta, felaktiga och omätbara mätningar som algoritmen gjort i jämförelse med människan.



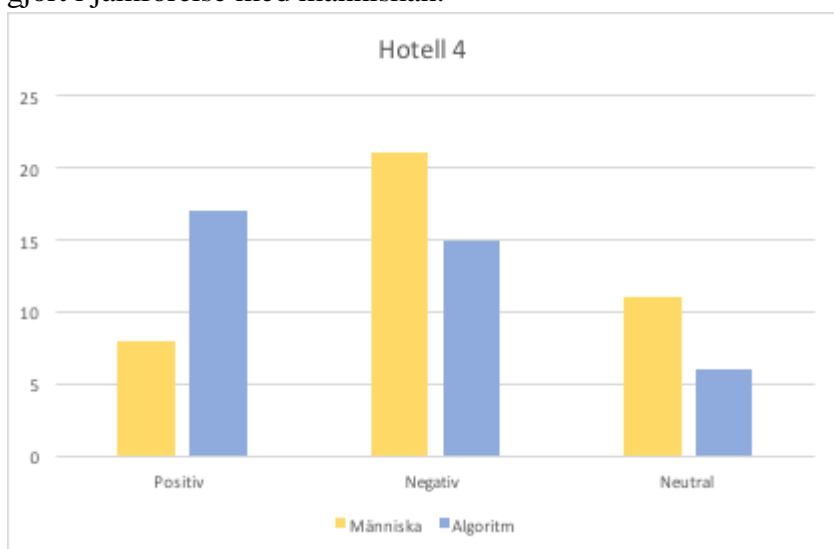
**Figur 4.2.23:** Fördelningen av Positiva, Negativa och Neutrala kommentarer mellan människa och algoritm. De kommentarer som algoritmen inte kunde mäta och klassificerades som NONE är borträknade.

Hotell 3	Människa	Algoritm	Avvikelser
Positiv	15	15	0
Negativ	4	5	1
Neutral	4	2	2
Omätbara	0	1	1

**Tabell 4.2.24:** Sammanställning av positiva, negativa, neutrala och omätbara analyser av människa och algoritm. Kolumnen som heter avvikelser visar antalet svar som skiljer sig mellan människan och algoritmen.



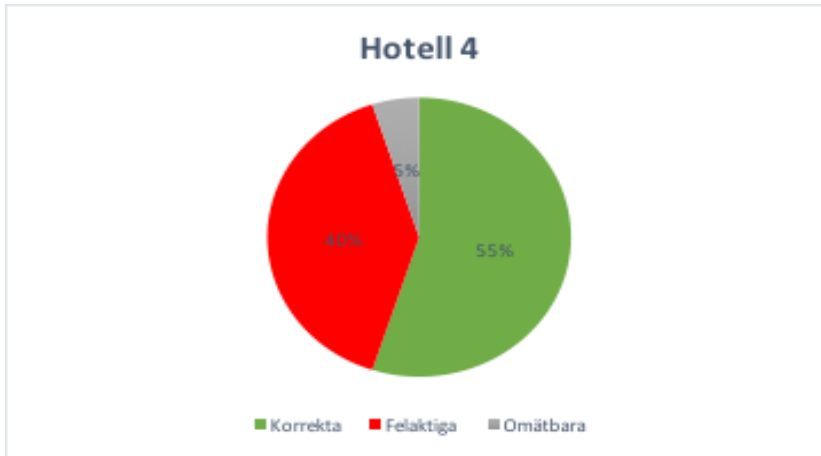
**Figur 4.2.25:** Resultatet av antalet korrekta, felaktiga och omätbara analyser som algoritmen gjort i jämförelse med människan.



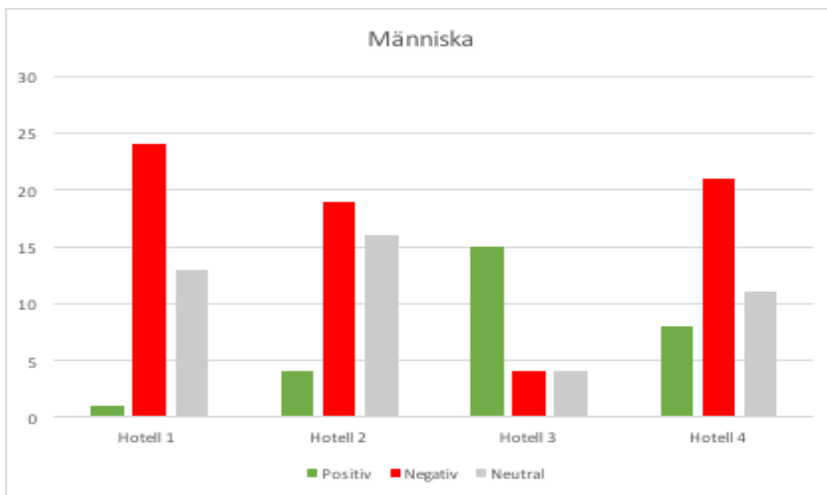
**Figur 4.2.26:** Fördelningen av Positiva, Negativa och Neutrala kommentarer mellan människa och algoritmen. De kommentarer som algoritmen inte kunde mäta och klassificerades som NONE är borträknade.

Hotell 4	Människa	Algoritm	Avvikelser
Positiv	8	17	9
Negativ	21	15	6
Neutral	11	6	5
Omätbara	0	2	2

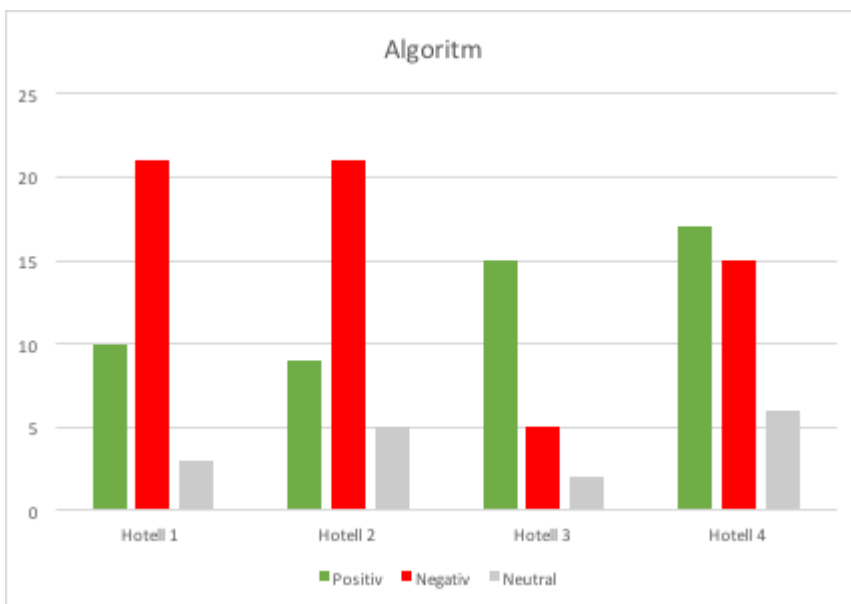
**Tabell 4.2.27:** Sammanställning av positiva, negativa, neutrala och omätbara analyser av människa och algoritmen. Kolumnen som heter avvikelser visar antalet svar som skiljer sig mellan människan och algoritmen.



**Figur 4.2.28:** Resultatet av antalet korrekta, felaktiga och omätbara analyser som algoritmen gjort i jämförelse med människan.



**Figur 4.2.29:** Sammanställning av positiva, negativa och neutrala svar av människan.



**Figur 4.2.29:** Sammanställning av positiva, negativa och neutrala svar av algoritm.

### 4.3 Jämförelse

Det som fungerar till högre grad hos attitydanalys-verktyget var klassificeringen av deskriptiva ord. Attitydanalys-verktyget har en bättre förståelse för vad som är positivt/negativt vid korta meningar. Figuren nedanför visar ett exempel på korta kommentarer där verifieringen fungerar. Den röda markeringen med bokstaven N illustrerar att algoritmen tolkat kommentaren som negativ och den gröna markeringen med bokstaven P illustrerar att algoritmen tolkat kommentaren som positiv. Detta stämmer överens med deltagarnas uppfattningen av meningen.

Too boring food for a fourteen day trip! More variety is required, especially week two.	N	AGREEMENT	SUBJECTIVE	100	NONIRONIC
The unveiling was bad for several days	N	AGREEMENT	SUBJECTIVE	100	NONIRONIC
Good with different themes on the buffet	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC
Good variety of dishes	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC

**Figur 4.3.1:** En lyckad verifiering

Det som fungerade mindre bra var kontextuppfattningen av längre meningar. Bilden nedan visar ett exempel på när en mening blir för lång och komplex för algoritmen. Kontexten från denna kommentar tolkas som negativ av våra deltagare medan algoritmen tolkar detta som något positivt. En förklaring till detta är att algoritmen har lagt stor tyngd på ordet “best” utan att förstå omständigheterna. Förutom att kommentaren är negativ och dåligt översatt, är även meningen med ordet “best” feltolkad. “They did their best but were busy” innebär inte i direkt mening någonting positivt.

Too little staff for so many guests - they did their best but were busy - Breakfast is served late when you have early tee-off times	P	DISAGREEMENT	SUBJECTIVE	94	NONIRONIC
-----------------------------------------------------------------------------------------------------------------------------------------	---	--------------	------------	----	-----------

**Figur 4.3.2:** Misslyckad verifiering av positiv/negativ kommentar

Algoritmen hade även svårt att tyda neutrala kommentarer. Bilden nedan visar ett exempel där algoritmen försökt balansera ut ett negativt ord med ett positivt ord. Det algoritmen misstolkar är att “Good if you can divide into smaller departments” inte är en positiv mening utan endast en rekommendation och att hela meningen faktiskt är negativ. Huruvida kategorisering bör gå till på det sättet, att en positiv och negativ kommentar blir en neutral kommentar är tveksamt. Det borde snarare vara att en kommentar som inte direkt innehåller någon typ av feedback som berör ämnet, bör tolkas som neutral. Finns det blandad återkoppling bör den övervägande känslan som kommentaren förmedlar bli den dominanta.

Slammy dining room with high volume. Good if you can divide into smaller departments. The agency's table was among the better ones but .... Preferred dining at the table, buffet always becomes enahanda despite the abundance. The proximity to the airport was disturbing with constant over-flight	NEU	DISAGREEMENT	SUBJECTIVE	94	NONIRONIC
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----	--------------	------------	----	-----------

**Figur 4.3.3** Misslyckad verifiering av neutral kommentar

Kommentarerna är textbaserade och därmed blir frågan “hur fungerar algoritmen på text?” något viktigt att utreda. Enligt det teoretiska avsnittet som finns under rubrik ” 2.11



**MEANINGCLOUD SENTIMENT ANALYSIS**” ska verktyget utan problematik kunna avgöra huruvida en mening är positiv/negativ/neutral, extrahera aspektbaserade attityder, upptäcka ironi, skilja fakta från åsikter och slutligen kunna skapa en helhetsbild för texten. Det resultatet visar oss är att verktyget inte helt lever upp till dessa utmärkelser. Ironi/sarkasm var något som inte förekom så ofta i dessa enkäter vilket gjorde att det inte gick att utreda fullständigt. Tidigare forskning som nämns i stycke 2.11 som gjort på till exempel Twitter-inlägg har dock visat att det är ett aktuellt problem. Verktyget kan förvisso utreda om en kommentar är positiv/negativ/neutral på ett korrekt sätt i 38-65% av fallen, i snitt 53% av gångerna i de fall vi testat. Däremot att verktyget kan skapa kontextuppfattningar på ett fungerande sätt är enligt vår undersökning någonting helt inkorrekt.

#### 4.4 Sammanställning av resultat

##### *Analys av enkäter med människor*

Del ett av undersökningen utfördes av människor. Deltagarnas ålder varierade mellan 21-51 och majoriteten av deltagarna hade en akademisk bakgrund inom digitalisering och turism. Svaren från deltagarna var näst intill identiska. Deltagarna ansåg att de positiva svaren som fanns i enkäterna var enkla att identifiera. Svårigheterna som fanns var att tyda om en kommentar var negativ eller neutral, deltagarnas svar skiljer sig mest åt vid denna aspekt. Deltagarnas uppfattning av analyser på enkäter är att analyserna är tidskrävande. En annan punkt de lyfte var att det kan vara svårt att tyda svaren från enkäterna eftersom det kunde vara dåligt formulerade eller innehålla data som inte var relevant för ämnet. Flertal av svaren var endast abstrakta åsikter snarare än feedback.

##### *Attitydanalys med en algoritm*

Algoritmen uppfattade fler svar som positiva än negativa. Den hade även svårigheter att läsa av viss text vilket gjorde att den inte kunde utföra en analys. Algoritmen klassade kommentarer som den inte kunde svara på med “NONE”, Detta gav en felaktig analys och ett ogiltigt resultat, det gjorde att värdefull data inte kunde utnyttjas. Det som nämns i rubrik “**4.3 Jämförelse**” var också några av de svårigheter som upptäcktes vid attitydanalys med en algoritm.

#### 4.5 Sammanställning av Matupplevelsen.

Vid undersökningen av Hotell 1 visade sig 50% av analyserna vara likvärdiga mellan människa och algoritm. 8% av svaren blev omätbart vilket endast motsvarade 1 fråga. 42% var felaktiga. Hotell 2 visade en större andel felaktiga analyser motsvarande 52% samt ett större antal omätbara resultat på 17%. Hotell 2 fick ett sämre resultat på undersökning då antalet likvärdiga analyser endast var 31%. En orsak till att resultatet skiljer sig åt är att antalet neutrala kommentarer, människan har tolkat 10 svar fler som neutrala i jämförelse med algoritmen som endast tolkade 1 svar som neutralt. Hotell 3 gav det mest likvärdiga resultat på 62% korrekt, 29% felaktigt och 9% omätbart. Det som framförallt skiljer sig i denna undersökning är antalet neutrala och positiva kommentarer. Människan har uppfattat fler frågor som neutrala, medans verktyget tolkar densamma som positiva. Hotell 4 hade 38% korrekta, 34% felaktiga och 28% omätbara svar. Det som utmärker undersökningen för detta hotell i jämförelse med de 3 andra hotellen är det höga antalet omätbara svar.

## 4.6 Sammanställning av Golfupplevelsen

Hotell 1 gav 38% korrekta, 54% felaktiga och 8% omätbara analyser. Det höga antalet felaktiga analyser har främst berott på antalet kommentarer som algoritmen tolkat som positiva och där människor har tolkat kommentarer som neutrala. Hotell 2 hade ett resultat på 54% korrekta, 36% felaktiga och 10% omätbara svar. Även i denna undersökning har antalet neutrala analyser varit betydligt fler hos människor än hos algoritmen. Hotell 3 hade 65% korrekta, 31% felaktiga och 4% omätbara. Människan och algoritmen hade tolkat samma andel av kommentarerna som positiva och det var ytterst få svar som skiljde dem åt. Hotell 4 visar 55% korrekt, 40% felaktigt och 5% omätbart. Det som främst har blivit felaktigt är antalet positiva kommentarer, dessa kommentarer har människan tolkat antingen som negativa eller neutrala.

## 4.7 Experiment

There was not so often clean meat much chicken and fish	P	AGREEMENT	OBJECTIVE	100	NONIRONIC
Much chicken and fish, there was not so often clean meat	P	AGREEMENT	OBJECTIVE	100	NONIRONIC
Too much chicken and fish, there was not enough clean meat	N	AGREEMENT	OBJECTIVE	92	NONIRONIC
We hated that it was too much chicken and fish, we missed clean meat so much.	N	AGREEMENT	SUBJECTIVE	100	NONIRONIC

**Figur 4.7.1** Manipulering av data

Experimentet visade att manipuleringen av datat ger olika resultat även fast innebörden av meningen egentligen är densamma. Figuren ovan visar att de deskriptiva orden har för påverkan på resultatet av en körning. Den första meningen är inte manipulerad, utan är i sitt original medan de tre resterande meningarna har blivit manipulerade. Andra meningen är innehåller exakt samma ord som den första, fast i en annan följd. Detta visar att meningsföljden inte har haft en stor påverkan i detta fallet. De två sista meningarna visar att desto mer deskriptivitet meningen har desto troligare är det att meningen påverkas.

Text	Polarity	Agreement	Subjectivity	Confidence	Irony
The restaurant have good staff, unlike the food	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC
not to complain too much but i have had better dinners	P	AGREEMENT	SUBJECTIVE	92	NONIRONIC
so small amount of food, i could have been happier with it	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC
multiple times i have been treated better within this price-category	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC
The "good" standard of yours is negotiable	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC

**Figur 4.7.2** Vidare manipulering av data

Vidare experimentering visade att många fall där meningar med utfyllnadsord förekommer blir meningarna feltolkade. Figuren ovan visar att algoritmen har klassat kommentarerna som positiva då den endast har fokuserat på det positiva orden istället för att ta med utfyllnadsord vid bedömningen.

## 5. ANALYS

*Detta avsnitt analyserar resultatet av studien med hjälp av den teori som har presenterats tidigare i rapporten. Analysen inkluderar konstateranden som gjorts av resultatet.*

Klassificeringen av ord var en utmaning för algoritmen då den endast söker efter positiva eller negativa ord för att göra sin bedömning. För en människa är det lättare att urskilja detta eftersom individen ser helheten vid bedömningen. Leverantören av verktyget MeaningClouds Sentiment Analysis motiverar att attitydanalys-verktyget kan urskilja helhetsperspektiv utifrån kontexter. Detta stämmer inte då resultatet i studien visar motsatsen. Däremot stämmer det MeaningCloud (2019) påstår om att verktyget är smart nog att kunna skilja på objekt och åsikter. Detta sker genom vanligtvis genom lingvistisk klassificering av ord vid attitydanalys. Dessvärre finns det förbättringsområden då algoritmen inte använder sig tillräckligt av lingvistik. Sigrud (u.å) beskriver lingvistik som vetenskapen om det naturliga språkets strukturer. Språkvetenskapen är bristande i algoritmen och en konsekvens av det är att algoritmen inte har möjligheten att ge korrekta tolkningar.

Skeppstedt (u.å) beskriver att maskininlärning har målet att ge bra prognoser vid ett lyckat resultat. Algoritmen som användes i studien behandlar maskininlärning, dessvärre kan den inte uppfylla målet då resultatet visar att den har över 50% felaktiga bedömningar i vissa undersökningar. Rich & Knight (1991) nämner att maskininlärning ingår i artificiell intelligens som har en avsikt att dra slutsatser samt förstå det naturliga språk. Maskininlärningen ger algoritmen förutsättningar att genomföra en analys av kommentarerna, dock finns det fortfarande brister. I rubriken **”2.2.1 Utmaningar inom artificiell intelligens”** nämns mönsterigenkänningen som en utmaning (Balkenius, Skeppstedt, Gärdenfors, u.å). I studien framkommer felaktiga bedömningar samt omätbara kommentarer vilket visar att mönsterigenkänningen inte fungerar perfekt. Ovchinnikova (2012) beskriver att NLU hanterar maskininlärningskompetensen som påverkar förståelsen av textfragmentet. Baserat på resultatet verkar det finnas en del problem inom det området.

Taboada (2016) nämner att vid användning av intensifiering och nedtoning går det att med hjälp av multiplikation att mäta och väga om orden är positiva eller negativa. Med det i åtanke, kan detta vara ett utvecklingsområde inom attitydanalys och det kan också användas för att förbättra analysen. Om intensifiering varit tillräckligt utvecklad vid körningen av algoritmen hade resultatet med stor sannolikhet blivit annorlunda då antalet felaktiga svar haft en lägre procentsats. Andra faktorer som också påverkade resultatet var de utmaningar som Karde & Sonawane (2016) skriver i deras rapport om domänberoende, motsatta uttryck och jämförelser. Dessa utmaningar för algoritmen gick att tyda i resultatet. Deltagarna kunde också ha svårt att för dessa utmaningar under undersökningen, dock inte lika mycket som algoritmen.

Algoritmen kunde endast köras på engelska vilket kräver att det sker en översättning från ursprungsspråket till engelska. Dale, Somers & Moisl (2000) säger att NLP ska göra det möjligt

att kommunicera med människor genom ett naturligt språk. Det går inte att uppfyllas fullständigt i studien då översättningen tar bort en del av det naturliga språket. Resultatet i studien visar att vid översättningen uppstår några kommunikationsbrister och meningsuppbyggnaden kan variera. Detta kan påverka resultatet då algoritmen får svårt att uppfatta kontexten eftersom den inte är grammatiskt korrekt, och detta resulterar i ökad andel felaktiga svar.

Enligt Regeringskansliet (2018) vill Sverige vara i framkant av digitalisering och AI. Baserat på resultatet av studien kan ett steg till detta vara att förbättra de verktyg som finns idag och göra dem bättre. I rubrik ”4.4 Sammanställning av resultat” nämns det att de mänskliga deltagarna ansåg att analys av kommentarer var väldigt tidskrävande. Det borde vara i Sveriges intresse att utveckla attitydanalys-verktyg som fungerar bra på det svenska språket för att vara i fas med omvärldens utveckling inom AI.

## 5.1. ARBETETS BIDRAG

Sedan tidigare har både fördelar och nackdelar med attitydanalys diskuterats av forskare. I vår fördjupning har vi försökt utreda vad som blir fel vid klassificeringarna mellan positivt, negativt och neutralt. Det vi har kommit fram till i vår undersökning är att det borde finnas ett bättre system för uppfattning av kontext. Det attitydanalys-verktyg vi testat ger efter sig ett visst felmönster. När en längre skrift både behandlar positiv och negativ information får verktyget det svårt att göra en korrekt analys. Vårt bidrag blir därför att kontextuppfattningen för den här typen av algoritmer måste förbättras om de ska gå att användas på ett förtroendefullt sätt av företag. Vårt lösningsförslag är att förbättra klassificeringen av ord för verktyget då utfyllnadsorden måste identifieras och kontext uppfattas bättre. Tidigare forskning menar att deskriptiva ord ska ställas mot varandra för att få kontexten, det vårt resultat pekar mot är att utfyllnadsorden har en viktigare påverkan på kontexten än enbart de deskriptiva orden.

	P	AGRE	OBJE	100	NON
		EMEN	CTIV		IRO
There was not so often clean meat much chicken and fish		T	E		NIC

**Figur 7.0.1** Figuren visar en misslyckad verifiering av neutral kommentar

Exemplet i **Figur 7.0.1** tolkar algoritmen kommentaren som positiv. Dock orden “there”, “was”, “not” samt “so” är ord som både kan ha positiv och negativ innebörd beroende på kontext. Algoritmen tolkar detta som positivt eftersom den fokuserar på “clean meat much chicken and fish”.

Text	Polarity	Agreement	Subjectivity	Confidence	Irony
The restaurant have good staff, unlike the food	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC
not to complain too much but i have had better dinners	P	AGREEMENT	SUBJECTIVE	92	NONIRONIC
so small amount of food, i could have been happier with it	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC
multiple times i have been treated better within this price-category	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC
The "good" standard of yours is negotiable	P	AGREEMENT	SUBJECTIVE	100	NONIRONIC

**Figur 7.0.2** Figuren visar 5 meningar där utfyllnadsorden inte identifieras och kontextuppfattningen blir felaktigt. (Bild från experiment).

Motsatt uttryck går att se på första raden på **Figur 7.0.2** och är ett problem som går att verifiera från tidigare forskning av Karde och Sonawane (2016) En negativ mening som är uttryckt på ett positivt sätt blir ofta feltolkad. Den här undersökningen tar ett steg längre eftersom Karde och Sonawanes forskning endast visar på ett exempel med endast en mening “This movie should be fantastic”, meningen i **Figur 7.0.2** är inte bara ett motsatt uttryck utan ger även 2 typer av feedback. “Personalen är bra, till skillnad från maten” i ett försök av algoritmen att kontextuppfatta blir det tyvärr motsatt effekt.

There was not so often clean meat much chicken and fish	P	AGREEMENT	OBJECTIVE	100	NONIRONIC
Much chicken and fish, there was not so often clean meat	P	AGREEMENT	OBJECTIVE	100	NONIRONIC
Too much chicken and fish, there was not enough clean meat	N	AGREEMENT	OBJECTIVE	92	NONIRONIC
We hated that it was too much chicken and fish, we missed clean meat so much.	N	AGREEMENT	SUBJECTIVE	100	NONIRONIC

**Figur 7.0.3** Figuren visar att utfyllnadsorden inte har någon betydelse utan att tolkningen enbart baseras på de deskriptiva orden.(Bild från experiment).

Exemplet som visas i **Figur 7.0.3** testas meningsföljden. I detta test spelade inte ordningen någon roll eftersom de två övre meningarna innehåller samma ord. Däremot de två nedre meningarna testas att lägga till mer deskriptivitet vilket direkt gör att algoritmen kommer in på rätt spår. Tidigare forskning har nämnt att meningsföljden har en påverkan. Det var inte fallet i vårt test utan snarare att algoritmen i dagsläget behöver tränas upp mot mer naturligt språk. Antalet deskriptiva ord borde också övervägas bättre av algoritmen. Vi har sett meningar där det till exempel sägs flera negativa ord följt av ett enda positivt, algoritmen har sedan tolkat detta som en neutral/positiv kommentar och vice versa, vilket kan bli fel.

The hotel is big and the dining room also. Too many people, P DISAGRE OBJEC  
 something rude, far too loud, and not at all cozy. We golfers were in EMENT TIVE  
 a clear minority and in addition the families with small children were  
 many. Apparently autumn school holidays in England.  
 Would rather stay at a smaller hotel inside Caleta Fuste.

**Figur 7.0.4:** Figuren visar en positiv mening följt av sex negativa meningar. Helheten borde därmed inte bli positiv.

En insikt vi fått efter att ha testat en attitydanalys-algoritm är att de flesta av dessa algoritmerna är upptränade genom att applicera dem på större dataset, sedan sker det inte mycket mer i utvecklingen. En idé som slår oss är att man borde nyttja användarna mer. Om det skulle gå att identifiera och rapportera feltolkningar av kontexter på ett smidigare sätt, skulle det gå betydligt snabbare att träna upp algoritmerna till att bli välfungerande och användbara för företag, privatpersoner etc. Ett återkommande problem i en kontextuppfattning ska inte behöva identifieras enbart av utvecklarna av verktyget, då kommer uppträningen ta väldigt lång tid och bli väldigt omfattande för utvecklarna. Det är bättre att ta till användare för att snabbt kunna identifiera fel. För att kunna nyttja användarna mer så krävs det mer än att de lägger till ord och kategoriseringar i sina egna “ordböcker” - ordböcker är individuella anpassningar man gör till sitt program.

Text	Polarity	Agreement	Subjectivity	Confidence	Irony
There was not so often clean meat much chicken	P	AGREEMENT	OBJECTIVE	100	NONIRONIC
Much chicken and fish, there was not so often clean	P	AGREEMENT	OBJECTIVE	100	NONIRONIC
Too much chicken and fish, there was not enough	N	AGREEMENT	OBJECTIVE	92	NONIRONIC
We hated that it was too much chicken and fish, w	N	AGREEMENT	SUBJECTIVE	100	NONIRONIC

**Figur 7.0.5:** Figuren visar 5 meningar där utfyllnadsorden inte identifieras och kontextuppfattningen blir felaktigt

Ett typiskt exempel på hur lingvistikens ordklassificering ska användas är genom uppdelning av substantiv, adjektiv och verb (Manning, Schütze 1999). För att förbättra MeaningCloud Sentiment Analysis bör mer användning av lingvistik användas. Detta på grund av att substantiven (som ofta nämns utfyllnadsorden i denna rapport) inte identifieras och kategoriseras på ett korrekt sätt.

## 6. DISKUSSION

Resultatet av attitydanalys med människor har fler negativa kommentarer än attitydanalysen med algoritmen. I problembeskrivningen nämns det att algoritmen har svårt att tolka kontexten i meningarna och det blir evident då den inte uppfattar sammanhanget på rätt sätt och därmed klassar kommentaren felaktigt. Ett återkommande misstag som algoritmen gör är att klassa kommentarer som positiva i sammanhang där majoriteten av texten är negativ. Taboada (2016) nämner klassificeringar av ord vid attitydanalys och i denna studie väger de positiva orden tyngre i jämförelse med de negativa orden. Därmed blir kommentaren klassad som positiv. Detta blir också felaktigt eftersom kontexten blir missuppfattad. Orsaken till problemet är att algoritmen inte tar med de deskriptiva orden vid bedömningen, utan fokuserar på att hitta de negativa och positiva orden. Detta är något som utreds i rubrik ” **5.1. Arbetets bidrag**”

Motivationen att genomföra arbetet kom från upptäckten av bristerna som fanns inom forskningsområdet. En av de största bristerna var att NLP inte kunde utnyttjas till dess fulla potential då det krävdes översättningar av text. Wollin (u.å) nämner att översättningar kan innefatta kulturskillnader som gör det svårbegripligt, vilket vi också kunde konstatera efter vårt resultat. Det var även svårt att hitta NLU-verktyg som var ordentligt utvecklade på svenska. Lingvistikerna var bristande då verktygen saknade vetenskapen om språket som Sigrud (u.å) har beskrivit tidigare i studien och specifikt för detta fall var det svenska språket.

Artificiell intelligens är ett växande område och det sker ständigt utvecklingar som tar forskningsområdet vidare, detta beskriver Rich & Knight (1991). Vi ville vara en del av denna process och valde därför att utforska attitydanalys, NLU, och lingvistik på en mer detaljerad nivå. Vår studie visade att AI har kommit långt vid sin utveckling men än går det inte att ersätta människan med nya innovativa lösningar inom analyser av enkäter. Människans kunskaper inom kommunikation och lingvistik är för utvecklade och har flera generationers kunskaper samt erfarenheter med sig som inte kan ersättas av digitala lösningar än så länge. För att visa detta valde vi att genomföra ett experiment där vi testade att manipulera vår data för att undersöka vad som blir fel vid en analys gjord av attitydanalys-verktyget. Dock testades endast ett attitydanalys-verktyg i denna undersökning och det går inte att med säkerhet att avgöra om andra verktyg skulle klarat uppgiften bättre. Eftersom syftet med undersökningen inte var att jämföra olika attitydanalys-verktyg, fokuserade vi endast på att undersöka ett verktyg grundligt. I efterhand märkte vi att det är svårt att bekräfta misslyckade fall för attitydanalys-verktygen, eftersom fallen inte behöver vara generella. Därför kan vi med denna forskning endast bekräfta vad som behöver förbättras med MeaningCloud Sentiment Analysis-algoritmen för framtida bruk.

## 7. SLUTSATS

*”Varför skiljer sig mänskliga tolkningar av text från automatiserade tolkningar av text?”*

Frågan kan enkelt besvaras genom att det går att konstatera att människan har lättare att förstå kontexter och användningen av lingvistik som leder till korrekta analyser. Detta blir tydligt då de fyra deltagarna i studien tolkade alla kommentarer på ett snarlikt sätt medan attitydanalys-verktygets tolkning skiljer sig drastiskt i jämförelse med deltagarna. Detta beror på att det finns brister inom algoritmens användning av lingvistik, NLU, dubbla budskap, domänberoende och jämförelser. Då det finns många utmaningar inom attitydanalys blir det svårt för större verksamheter att förlita sig på det attitydanalys-verktyg som testades i studien. Om människan ska ersättas av automatiserade, digitala lösningar krävs det att denna typ av algoritm förbättras för att få ut rätt analyser. Forskningsområdet kräver vidare forskning för att kunna förbättra algoritmerna till att förstå kontexter av kommentarer. Bidraget i denna forskning har varit att nyttja användarna mer i utvecklingsprocessen samt att detta attitydanalys-verktyg bör bli bättre på att utvärdera utfyllnadsord snarare än endast deskriptiva ord. För att utvärdera algoritmernas egenskaper mer på detaljnivå hade data behövts användas på fler verktyg/algoritmer. Denna studie har endast berört en specifik algoritm vilket inte med säkerhet kan intyga att detta är fallet för samtliga verktyg och algoritmer på marknaden, endast MeaningCloud Sentiment Analysis.

### 7.1 Övervägande vid användning

Ett övervägande som krävs, är att se om det är möjligt att använda ett attitydanalys-verktyget för att ersätta den mänskliga analys av enkäter. Frågor som ” Går det att lita på verktyget fullt ut?” behöver besvaras och följande ställningstaganden bör tas: Ett verktyg som kan hantera det svenska språket skall användas för en svensk verksamhet, vice versa för det språk som verksamheten opererar inom. Attitydanalys-verktyget har fortfarande en del förbättringar att göra inom kontextuppfattning och användning av lingvistik vilket gör det svårare att analysera längre texter, vissa verksamheter har inte något behov av att analysera längre texter (t.ex. analysering av Twitter flöden) och borde kunna ha en större nytta av verktygets aktuella potential. Ett annat övervägande är om formuleringen på frågorna kan utvecklas för att responsen av frågorna blir lättare att analyseras av attitydanalys-verktyget och ge respondenter tydliga riktlinjer. Om det finns en möjlighet att formulera frågorna så att det blir mer användarvänligt för attitydanalys-verktyget borde även det övervägas för att få ett bättre resultat.



## 8. VIDARE FORSKNING

*Vidare forskning på områden kan vara:*

- Att genomföra en liknande studie som denna men jämföra ett flertal attitydanalys-verktyg för att få en större förståelse för hur mycket det skiljer sig mellan de olika verktygen. Vissa slutsatser uteblev i denna studie eftersom det saknas belegg att verifiera att fallen är generella hos attitydanalys-verktyg. Till exempel går det inte att verifiera att samtliga verktyg har svårt att identifiera utfyllnadsord, utan endast det verktyg som testats.
- Hur ska verktyget bli bättre på att uppfatta ironi och sarkasm? Problemet med ironi och sarkasm finns det mycket tidigare forskning inom, problemet kvarstår dock och är en stor utmaning för maskininläringen. Denna studie har inte berört ironi och sarkasm specifikt men problematiken är relaterad till den problematik vi stött på med svårigheter att träna attitydanalys-verktyg till att förstå vardagligt språk med mycket utfyllnadsord. Därför tror vi att lösningen för problemet med ironi och sarkasm även kan vara relaterad till lösningen för att kunna kategorisera utfyllnadsord.
- Hur ska verktyget enkelt kunna läras upp i olika språk för att undvika översättningar? Tidigare forskning har försökt utreda hur detta ska kunna gå till, bland annat genom att lingvistikmetoder för att kunna göra träningen av data mer effektiv. Eftersom översättningar ofta blir fel så är det bättre med språkspecifika verktyg som är upplärda i var enskilt språk. Implementeringen verkar inte vara helt testad vilket innebär att ett experiment alternativt en artefakt skulle behöva testas.

# KÄLLOR

Balkenius, C., Skeppstedt, J., & Gärdenfors, P. (u.å.). Artificiell intelligens. Hämtad 2019-05-07. Från <https://www.ne.se/uppslagsverk/encyklopedi/l%C3%A5ng/artificiell-intelligens>

Dale.R., Somers, H.L., Moisl, H.(2000) HandBook of Natural Language Processing (1.uppl) New York. Marcel Dekker, Inc.

Denscombe,M.D.(2018).Forskningshandboken för småskaliga forskningsprojekt inom samhällsvetenskaperna.(4.1 uppl).Lund: Studentlitteratur.

Dmitry Davidov, Oren Tsur, and Ari Rappoport.( 2010). Semi-supervised recognition of sarcastic sentences in twitter and amazon. In Proceedings of the Fourteenth Conference on Computational Natural Language Learning, CoNLL '10, pages 107–116, Stroudsburg, PA, USA. Association for Computational Linguistics.

Fledman.R (2013) Techniques and Applications for Sentiment Analysis. Communication of the ACM (Volym .56) sid.82. doi:10.1145/2436256.2436274

Glant, O (2018). *Attitydanalys av svenska produktomdömen – behövs språkspecifika verktyg?*.Examensarbete.Stockholms Universitet, Stockholm. Hämtad 2019-04-15. Från <http://www.diva-portal.org/smash/get/diva2:1215641/FULLTEXT01.pdf>

Holme, I.M. & Solvang, B.K. (1997). *Forskningsmetodik: om kvalitativa och kvantitativa metoder*. (2., [rev. och utök.] uppl.) Lund: Studentlitteratur

Kharde,V,A., & Sonawane, S.S.(2016) Sentiment Analysis of Twitter Data: A Survey of Techniques. (Volume139.NR 11) Pune. Pune Institute of Computer Technology. Pune University of Pune.

Liddy, E, D.(2001). Natural Language Processing. Syracuse University. Hämtad 6 mars. Från <https://surface.syr.edu/istpub/63/>

Linguistics Journal (2019) Linguistics Journal. Hämtad: 2019-04-15 Från: <https://www.linguistics-journal.com/>

Liu, B., (2011) Web Data Mining - Exploring Hyperlinks, Contents and Usage Data(Andra utgåvand) : Springer-Verlag Berlin Heidelberg

Liu, B., (2012) Web Data Mining - Exploring Hyperlinks, Contents and Usage Data(Andra utgåvan) : Springer-Verlag Berlin Heidelberg.

- Manning, C.D, Schütze. H, (1999) Foundations of Statistical Natural Language Processing(1.uppl). Cambridge, Massachusetts: The MIT Press
- Meaningcloud. (2019). Sentiment Analysis API. Hämtad: 2019-04-15 Från: <https://www.meaningcloud.com/developer/sentiment-analysis>
- Mitchell, M. Tom. (1997). Machine Learning (1.uppl) New York: McGraw-Hill Science/Engineering/Math.
- Nilsson, N, J (1993) Principles of Artificial Intelligence.( [omarb. och utvidgade] uppl).Palo Alto: Tioga publishing company.
- Ovchinnikova, E.(2012). Integration of the world knowledge for natural language understanding.(3uppl). Paris, France: Atlantis Press
- Regeringskansliet.(2017). Artificiell intelligens- en nyckel för att stärka svensk konkurrens. Hämtat 2019-02-10. Från <https://www.regeringen.se/artiklar/2017/12/artificiell-intelligens--en-nyckel-for-att-starka-svensk-konkurrens/>
- Regeringskansliet.(2018). Sverige ska bli ledande på ansvarsfull utveckling och användning av framtidens teknologier. Hämtat 2019-02-11. Från <https://www.regeringen.se/artiklar/2018/08/sverige-ska-bli-ledande-pa-ansvarsfull-utveckling-och-anvandning-av-framtidens-teknologier/>
- Rich.E., Knight.K.(1991) Artificial Intelligence (Second Edition) New York. McGraw-Hill, Inc.
- SCB (u.å) Att välja med metod och intervjupersoner. Stockholm - SCB. Hämtad 6 mars. Från <https://www.scb.se/dokumentation/statistikguiden/undersokning-och-urval/att-valja-metod-och-intervjupersoner/>
- Scott, J (2014) A Dictionary of Sociology. Oxford: Oxford University Press
- Sigrud.B.(u.å). Strukturell lingvistik. Hämtad 2019-04-15. Från <https://www.ne.se/uppslagsverk/encyklopedi/lång/strukturell-lingvistik>
- Skeppstedt.J.(u.å)Maskininlärning. Hämtad 2019-03-13. Från <http://www.ne.se/uppslagsverk/encyklopedi/lång/maskininlärning>
- Taboada, M. (2016) Sentiment analysis: An overview from linguistics. *Annual Review of Linguistics* . Simon Fraser University, Burnaby. Hämtad 2019-05-22 Från: [https://www.researchgate.net/publication/283954600\\_Sentiment\\_Analysis\\_An\\_Overview\\_from\\_Linguistics](https://www.researchgate.net/publication/283954600_Sentiment_Analysis_An_Overview_from_Linguistics)
- Tsytsarau, M., & Palpanas, T. (2011). Survey on mining subjective data on the web. Springer,

Wollin.L.(u.å). Översättningar .Hämtad 2019-04-25.Från  
<http://www.ne.se/uppslagsverk/encyklopedi/lång/översättning>

Yin, R. K. (2003). Case study research: design and methods. California: Sage Publications  
Inc.Holme, I.M. & Solvang, B.K. (1997). Forskningsmetodik: om kvalitativa och kvantitativa  
metoder. (2., [rev. och utök.] uppl.) Lund: Studentlitteratur.