Exploring the implementation blind spots

Selective Decoupling of Freedom of Information

Kuk, George; Chim, Jimmy; Giamporcaro, Stephanie; Janssen, Marijn

Sentiment Analysis and Google Trends Data for Predicting Car Sales

Completed Research Paper

Fons Wijnhoven
University of Twente
P.O. Box 217, 7500 AE Enschede,
Netherlands
a.b.j.m.wijnhoven@utwente.nl

Olivia Plant
University of Twente
P.O. Box 217, 7500 AE Enschede,
Netherlands
o.h.plant@student.utwente.nl

Abstract

This article explores the usefulness of sentiment analysis and Google trends data for car sales forecasting. Previous research has demonstrated the use of both techniques for sales forecasting, but current literature is more ambiguous in its results for forecasting the sales of high involvement goods like cars. In this study, about 500,000 social media posts for eleven car models on the Dutch market are analyzed using linear regression models. Furthermore, this study compares these outcomes to the predictive power of Google Trends. The results suggest that social media sentiments have little predictive power towards car sales while Google Trends data and social mention volume show significant results and can be incorporated into an effective prediction model. A prediction model with time lags using decision tree regression is built that can be used by the car industry as an addition to traditional forecasting methods.

Keywords: Consumer decision making, predictive modelling, search engines, sentiment analysis, social media, 15. Data science, decision analytics and visualization

Introduction

The automotive industry is an important industry in the Netherlands and employs about 50.000 people. Sales forecasting in the automotive industry is particularly important since cars in the current system are either built-to-delivery or built-to-forecast. However, the latter one often leads to a bullwhip effect due to uncertainty in demand and inaccurate forecasting (Suthikarnnarunai 2008). Even if cars are built-to-delivery, accurate forecasting can still help managers to plan and allocate their resources better.

Social media act as word of mouth and allow companies to collect large-scale and up to date data that represents honest consumer opinions (Ceron et al. 2013; Tuarob et al. 2014). Many companies have understood this development and pay increasing attention to social media content for making better decisions (Karlgren et al. 2012; Liu 2012; Wijnhoven and Bloemen 2014). A popular method used by companies to analyze these data is sentiment analysis, which analyzes people's opinions, sentiments, evaluations, attitudes, and emotions from natural language (Pang and Lee 2008). Understanding these sentiments is important for product and service evaluation and these sentiments also may have an impact on people's future purchases. Among studies that use sentiment analysis are predictions of movie sales (Asur and Huberman 2010), stock price movements (Bing et al. 2014; Nguyen et al. 2015), book sales (Dijkman et al. 2015), and iPhone sales (Lassen et al. 2014). Nevertheless, literature still lacks information about the validity and reliability of sentiment analysis in the context of more expensive products like cars.

Consumers tend to spend a considerable amount of time searching for information about a potential vehicle. A study by Kandaswami and Tiwar (2014) showed that a majority of customers spend more than 10 hours to identify the best vehicle for their requirements. In China, this number even reached 70%, while in western countries like Germany and the USA, a percentage of 40-50% stated that they spend at least this amount of time. Consequently, search volumes —as registered by search engine firms like Google.com— may represent a large part of people's searches for decisional information for their car

purchases. Therefore, previous research has studied the predictive power of Google Trends data for car sales, although with different success (Barreira et al. 2013; Geva et al. 2017). Geva et al (2017) found that the best car sales prediction models combine Google Trends data, Forum sentiment and Forum mention volumes as predictors. In their study, Google Trends has a similar predictive power as Forum sentiments and volume together. Our aimed at contribution here is a further identification of the meaning of these predictors. For this we use the AIDA model of the customer journey (Gensler et al. 2017). AIDA stands for four stages in a customer's buying process. The first A stand for attention. Attention is raised especially by sales promotion and publishing activities, which may for example be found back in mention volumes of social media. The I stands for interest, which may be found back in search behavior of customers, as can be registered via Google Trends. The D stands for desire, which can be found back in the sentiment of people which can be registered by the ratio of positive and negative social media expressions, so-called P/N ratios. The last A stands for action, which is the actual purchasing by the customer.

A research question therefore is *'What is the predictive power of sentiments for car models expressed on social media towards car sales in the Netherlands?'*

This research question, therefore, calls for an exploration of the applicability of sentiment analysis results as an indicator of a product desire, and thus may be useful for building a prediction formula by these findings. In this prediction, time lags have a central role. If the time lags between sentiment expressions and sales moments is positive (i.e. sales happen after the expressions), this could indicate a desire. If the time lags are negative, these expressions could indicate an evaluation. We will also compare the predictive power of sentiments to the predictive power of Google Trends because search could be an indicator of interests with possible predictive power.

Therefore, our second research question is *'What is the best predictor of car sales in The Netherlands: social media sentiments or Google Trends data?'*

We expect that if the AIDA activities run via a sequential process, that Google Trends volume peaks will have a larger time lag and weaker correlations with actual sales peaks then sentiment peaks.

This article will start by providing an overview of relevant literature and theories. From this, a research model is constructed that intends to explain possible relations between Google Trends, social media data and sales. Subsequently, data from social media and Google Trends is analyzed with regards to sales prediction abilities, and the outcomes are compared. Finally, a prediction model is constructed including the variables that were assessed as most useful. The results will then be discussed in the context of academic and practical implications and limitations.

Theory and Hypotheses

Sentiment analysis refers to extracting sentiments and opinions from written text (Liu 2012)(Pang and Lee 2008). This process requires natural language processing (NLP) which is a research area that explores how computers can be used to understand and manipulate natural language text or speech (Chowdhury, 2003). Sentiment analysis has been researched at a document level, a sentence level and an aspect/entity level of which the latter one is the most fine-grained level (Pang and Lee 2008).

Serrano-Guerrero et al. (2015) identify five main tasks of sentiment mining tools: The first is sentiment classification, often termed sentiment polarity identification. Common problems with identifying and classifying sentiments as positive, negative or neutral arise when the author expresses multiple opinions or when there is more than one source of opinion mentioned in the text. These opinions can contradict each other or refer to different attributes of the target (Liu, 2012). The second challenge is subjectivity classification. This means that the tool needs to define whether a text contains factual data or expresses the subjective belief of the author. Thirdly, the tool needs to summarize the given opinion of the author. The fourth challenge is to extract the opinion from the text. Finally, a task of sentiment analysis is sarcasm or irony identification for avoiding textual sentiment misclassifications. This is a difficult problem for automatic sentiment analyzers (Reyes et al. 2012) but a comparison of sentiment analysis tools by Serrano-Guerrero et al (2015) has shown that some tools have made great progress in this area.

Many of these challenges have already been dealt with to some extent, and the accuracy of sentiment analysis tools is steadily improving.

Following the Theory of Planned Behavior (1991), behavior is influenced by an intention which again is influenced by three different factors. These are the person's attitude towards the behavior, a subjective norm and the perceived behavioral control (describing the perceived easiness of fulfilling the behavior). Applying this theory to social media and consumer buying decisions, social media can help to shape the subjective norm that consumers' experience. If a large number of users' posts contain negative comments about a car, the consumer might decide not to buy the respective car.

The degree of influence, however, depends very much on the circumstances as well as the content and source of the information received. Based on the Theory of Reasoned Action (Ajzen and Fishbein 1980) and the Technology Acceptance Model (Davis 1989), Erkan and Evans (2016) established the Information Acceptance Model (IACM) which states that purchase intentions are also influenced by the type of information a consumer receives and whether this information is adopted. Information adoption depends mainly on information usefulness which is influenced by three factors: information quality, information credibility, and the need for information. These three factors therefore also play a crucial role when it comes to the influence of social media on purchase decisions. During this information collection and information evaluation process, the consumer goes through stages of the decision-making process in which a consumer can be influenced.

Bing et al. (2014) state that social media sentiments affect sales only with a delay, which we name the time lag. The time lag between an increase in positive or negative comments and an increase/decrease in sales is variable since social media may influence the consumer at any stage (early or late) of the buying process. The nature of this time lag becomes clearer when observing the consumer buying decision process. According to Kotler (1994), consumers go through five stages when buying a product. The initial buying decision is started by the consumer through recognizing a need or problem instead of being persuaded by a product. Next the customer goes to the following successive stages: information search, evaluation of alternatives, actual purchase decision; and post purchase behavior. A time lag will take place between the information search phase and the purchase decision while the consumer is evaluating alternatives.

A purchasing model that describes the customer's decision process in terms of information search, processing and decisions is the AIDA model. The AIDA model describes the funnel from attention, to interest, desire and action that consumers go through when they are drawn to a product and ultimately decide to buy it. According to Lassen et al. (2014), social media can play a part in all steps of this model. While mentions on social media are the rate of attention that a product receives, Google Trends may represent the interest that potential customers have for the product. Research has shown that search activities can represent buying intention and even predict consumer behavior and sales of both lower and higher involvements purchases (Choi and Varian 2012; Goel et al. 2010; Yang et al. 2015). However, Google Trends do not represent positive or negative sentiments and therefore are not suitable to indicate a desire for a product. Desire of sentiment can be measured by the positive/negative ratio of subjective expressions on social media. Following the AIDA model, an interest gives a stronger indication of an intention to buy than attention, and desire is a stronger intention to buy than interest. Therefore, we would expect stronger correlations of desire with sales volume than for interest and attention. We also would expect shorter time lags between the moment of desire peaks and the moment of sales peaks than for interest and attention. However, Google can still strongly influence the decisions made by the presentation order of search results. This is especially strong in situations where the searcher is more indecisive at the search stage (Epstein and Robertson 2015). Since high involvement purchases require more research and more difficulty in making a choice than low involvement and routine purchases, it is possible that Google Trends is even a better predictor of sales volumes of high involvement goods than sentiment (Yang et al. 2015).

There has been previous research that explored the relevance of social media to predict sales, such as the work of Asur and Huberman (2010) who predicted box office sales remarkably accurate by including many variables such as sentiments and the frequency of tweets into their prediction model. Various other prediction research has followed Asur and Huberman's (2010) approach. A study based on this method by Lassen et al. (2014) predicted quarterly iPhone sales by analyzing the sentiments of tweets and using a seasonal weighting of tweets to calculate the given quarter's proportion of the last calendar year. However, predicting car sales from sentiment is not a trivial problem. Much of sentiment expressions are not necessarily a desire expression but also could be an evaluation. If an evaluation, this will indirectly

influence the desire, but evaluations may also be of use for knowing the service need of a product (Pang and Lee 2008). One could state therefore that sentiment expressions before the purchase (i.e. with a positive time lag) indicate desire, whereas sentiment expressions after the sales are evaluations or expressions of service needs.

Both types of research used the following definition: p= Tweets with positive sentiment; n= Tweets with negative sentiment; and o= Tweets with a neutral sentiment, with Subjectivity being: $Subjectivity = \frac{p+n}{o}$ and the Positivity to Negativity Ratio (PNR) being $PNR = \frac{p}{n}$.

A similar approach will be taken in this research where the PNR will be the independent variable. However, p will not be defined as the number of active posts but as the percentage of positive posts from all posts about a particular car model over the time span of one month. This also yields for n being the percentage of negative posts. This is done to prevent the increase of social media usage and the subsequent increase of posts over the past years to influence the outcome of the research. The PNR will be the same, regardless if it is calculated with percentages or absolute number since it only describes the ratio between those.

The first hypothesis to be researched therefore is:

H1: The PNR of social media mentions about a car model has a positive influence on sales for this model.

Bataineh (2015) concluded that three factors of electronic word of mouth (eWOM, i.e. sentiment and mentions) have a significant and positive impact on consumer purchase intention, i.e. eWOM credibility, quality, and quantity. This finding is also supported by Cheung and Thadani (2012). Consequently, we hypothesize that not only the sentiment of eWOM but also the attention as present in the volume of mentions has a positive influence on purchase volumes. Therefore, the following hypothesis is established:

H2: The number of total mentions about a car model on social media correlates positively with the number of car sales.

Equally to positive reviews having a positive impact on sales, it is assumed that negative reviews have a negative impact. Lee, Park, and Han (2008) found that the consumer attitude towards a product becomes more unfavorable as the proportion of negative online consumer reviews increase. Since they also stated that customers tend to believe negative comments more than positive ones it is expected that this relationship is even stronger than the one between positive comments and sales resulting in:

H3: The percentage of negative mentions about a car model has a negative influence on the number of sales of this model.

The previous theories have started from the assumption that people who have already bought a car, place social media posts with reviews of it online which then influence other people to buy the same item. However, this does not necessarily hold for high end priced cars. Especially luxury cars tend to receive a lot of attention with only a small and exquisite clientele purchasing them. It is, therefore, possible that an increase of positive comments from 'fans' of a high-end car does not lead to an increase in purchases. This leads to the following hypothesis:

H4: The higher the price of a car, the weaker the correlation between the social media data and the sales.

When comparing studies about Google Trends and Twitter as predictors for sales, Twitter usually provided higher R Square values (Asur & Huberman, 2010; Choi & Varian, 2012). Previous studies on car sales predictions using Google Trends by Barreira et al. (2013) and Fantazzini and Toktamysiva (2015) did not result in effective prediction models. A more recent study of Geva et al (2017) which found successful predictions models for car sales, did not distinguish between models for different car models. This is problematic, because the car market has a highly monopolist competitive differentiation nature, and thus customer behavior and customer decision processes may be much different per car model (Brander and Spencer 2015; Cattani et al. 2017; Hunt and Morgan 1995). We, therefore, want to test if:

H5: The correlation of sentiments with car sales for models with different prices is higher than the correlation of relative search volume with car sales for models with different prices.

The above mentioned hypothesizes are summarized in Figure 1.

Figure 1. Research model
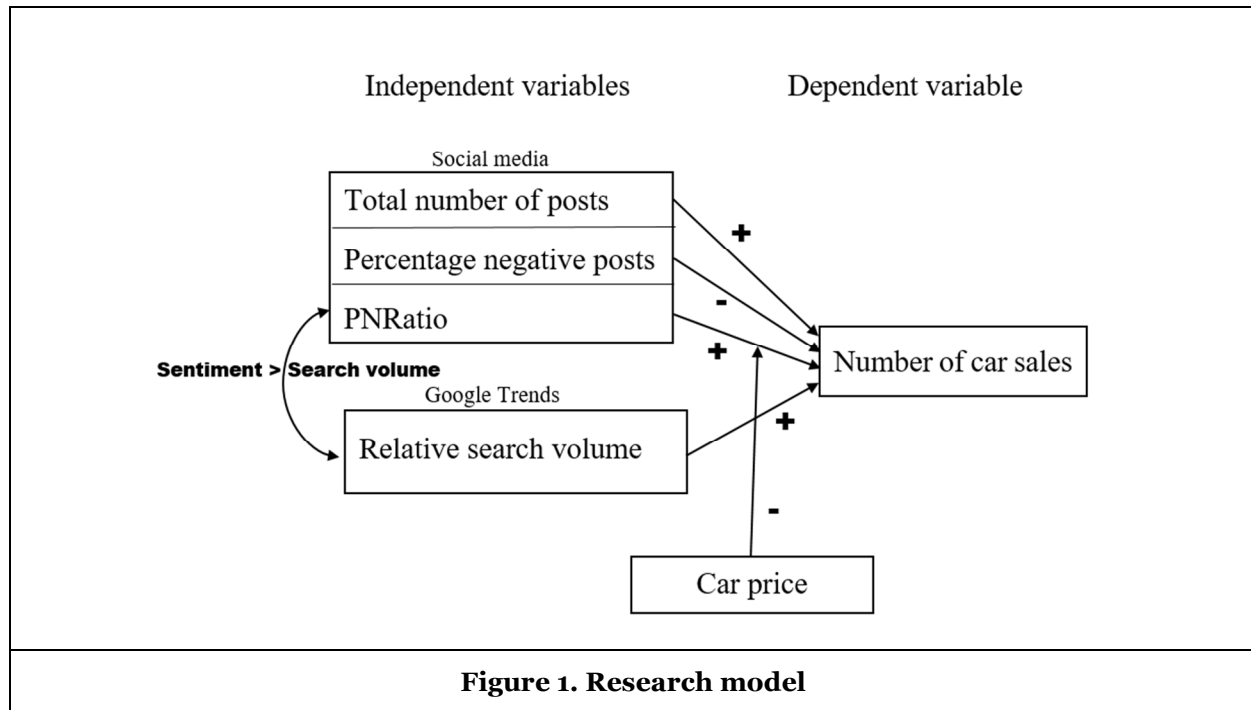
Methodology

For testing the research hypotheses and answering the research questions, it is important to know how the group of opinion expressing people corresponds with the target group of the research (Wijnhoven and Bloemen 2014). Obviously not all people who intend to buy a car have a social media account; however, in the Netherlands already 80% of the population with internet access also uses social media (CBS 2015). The vast majority of car buyers is therefore assumed to also have access to social media.

The required data was gathered for a total of eleven models over a period of 52 months, ranging from January 2012 until April 2016. The cars were chosen according to the European system of car classification (Office for Official Publications of the European Communities 1999) which divides cars based on their size and specifications. So-called mini-cars are labeled as A-Class while bigger cars are ranked as B-Class and above. The list ends with the F-Class which describes luxury cars. Extra classes include among others S (sports cars) and J (off-roaders). This research analyzes two models each from class B to E and one model each from class A, F, and S which covers the most common cars as well as two luxury cars. The full list of car models in this study is given in Table 1.

Car model	Class	Starting price in €	Search term Coosto	Search term Google Trends
Table 1: car models included in this study				
Fiat Panda	A	13,675	Fiat Panda	Fiat Panda
Ford Fiesta	B	13,995	Ford Fiesta	Ford Fiesta
Opel Corsa	B	20,950	Opel Corsa	Opel Corsa
Honda Civic	C	21,050	Honda Civic	Honda Civic
VW Golf	C	21,050	(Volkswagen Golf) OR (VW Golf)	"Volkswagen Golf" + "VW Golf" -Trends
Ford Mondeo	D	29,575	Ford Mondeo	Ford Mondeo
Volkswagen Passat	D	31,450	(Volkswagen Passat) OR (VW Passat)	"Volkswagen Passat" + "VW Passat" -Trends
BMW 5 Series	E	47,990	"BMW 5-serie" –GT	"BMW 5-serie" –GT
Mercedes E Class	E	46,800	(MB OR Mercedes-Benz OR Mercedes) "E-Klasse"	MB + Mercedes Benz + Mercedes "E-Klasse."
Porsche Panamera	F	106,400	Porsche Panamera	Porsche Panamera
Porsche 911	S	115,000	Porsche 911	Porsche 911

For each car model the following variables were collected: the total number of posts about this model per month, the number of social media positive posts per month, the number of social media negative posts per month, the Google Trends score per month, and the number of cars sold per month.

The social media feeds were analyzed through the use of sentiment analysis tool Coosto. This tool allows to analyze social media posts placed in the Netherlands and classifies each post as either positive, neutral or negative. As sources, Coosto gives eight social media sites, as well as various news sites, blog sites, and forums. This means that data collected via Coosto are less biased by the social dynamics and membership population of one platform, like Twitter (Mislove et al. 2011; Wijnhoven and Bloemen 2014; Wilson et al. 2012). Coosto has a high reputation and is widely used. Its sentiment classifier has an accuracy of 80% according to Team Nijhuis (2013) —an independent social media consultancy firm— which is very high compared to other tools considering the results of the research from Serrano-Guerrero et al. (2015). Since the research is determined to analyze social media sentiments, news sites, blogs and forums have been excluded from the search. The remaining social media sites that were searched are Twitter, Facebook, LinkedIn, YouTube, Google+, Hyves, Instagram and Pinterest. In total, 502,681 social media posts were analyzed.

The search was conducted by entering the Dutch name of each car model into Coosto and taking down the monthly number of all comments related to this car model as well as the number of positive and negative comments as determined by Coosto's sentiment analyzer. Included in the count of posts are both the original posts as well as retweets (on Twitter). This increases the validity of the measurement method since a post that is often retweeted is read by more people and could also influence more consumers in their buying decisions.

While searching for Google Trends data, various spellings or expressions that a user could use while posting about a car model were checked. For example, the query for the car model Volkswagen Passat was: *"VW Passat" OR "Volkswagen Passat*." Furthermore, cars with a similar name but different specifications were explicitly excluded from the search, such as the BMW 5-serie GT. Posts about this car model would also show up when searching for the BMW 5-series, however, the GT differs from the standard BMW 5-Series which means that posts about this version are unlikely to influence consumers considering to buy the standard 5-Series. The query was therefore defined as *"BMW 5-serie" –GT*.

As a comparison factor, the relative search volume per month from Google Trends was collected for the same car models using the same search terms. These sometimes had to be adjusted slightly in order to match the search language of Google Trends, for example by replacing the Boolean search term OR with the sign +. Google Trends only gives the relative search volume per month which is the query share of the searched term. The query share is calculated by dividing the query volume of the searched term by the total number of searches in the specified region and the given time frame. The month with the highest relative search volume is then normalized to 100. With regards to this research this means that the search volume of a particular car model was divided by all searches in the Netherlands between January 2012 and April 2016. After this, the month with the highest average then received the score while the scores of the other months were adjusted according to this maximum (Choi and Varian 2012). This normalization implied the score of 100 represents a different query share (and the absolute number of searches) for every model, depending on what the monthly maximum of searches was.

For the dependent variable, the monthly number of car sales was deducted from the website of BOVAG, the Dutch Federation of Automotive Dealers and Garage Holders which were also taken down into SPSS. The source website is https://www.bovag.nl/pers/personenauto/verkoopcijfers-personenauto-s-naar-merk-model-per (accessed may 10, 2016).

A study revealed that 60% of the buyers ('normal buyers') needed between one and six months from first thinking about buying a new car and the actual purchase while 16% needed less than a month for this decision. Only 9% needed more than a year to buy a new vehicle (Putsis and Srinivasan 1994, 1995) The decision duration (time lag in our model) will, therefore, be tested until a maximum of twelve months since the moment of sales.

Analysis and Results

Descriptives and linear regression analysis

Before analyzing each car model separately, the averages of each model were calculated for the variables *PNR, the total number of mentions, percentage negative mentions* and *Google Trends*. Each variable was mapped onto a scatterplot to examine their relationship with sales. Since most findings in earlier research were based on linear models (Asur and Huberman 2010; Barreira et al. 2013; Fantazzini and Toktamysova 2015; Goel et al. 2010), a linear regression analysis for the PNR and the sales was conducted in order to analyze the first three hypothesizes. While conducting the regression analyses, the residuals histogram and the PP plot were examined to ensure that the criteria of a linear regression analysis were fulfilled. The diagrams indicated that the errors were independent from each other, were approximately normally distributed and had a constant variance. This means that the relationships between the variables were indeed approximately linear and the conditions for a linear regression fulfilled. The use of linear models would therefore not bias the outcome. The outcomes of these linear regression analyses are summarized in Table 2.

Table 2. Linear regression analyses for averages of variables with sales			
Variable	R	R²	Significance
PNratio	.111	.012	.746
Total mentions	.804	.646	.003*
Percentage negative comments	.253	.064	.454
Google Trends	.320	.102	.338

The regression analysis for the PNR with sales indicates a weak, negative correlation that is not significant and obtains a very low R² value. The negative direction of the correlation does not fit the research model which stated that an increase in positive sentiments, as well as a decrease in negative sentiments, will lead to an increase in sales. As indicated by the graph, the sales, therefore, might not be causally related to the preceding social media sentiments. However, consumer behavior may vary per car model a person wishes to purchase. Therefore, although a correlation of the PNratio and sales cannot be found at this general level it is possible that the ratio serves as a predictor for sales if it is analyzed per car model and by including a time lag between social media data and sales into the model.

To test the second hypothesis, a regression analysis between the total number of mentions (regardless of the type of sentiment) about a car model and the number of sales was conducted. This showed a strong

correlation of 0.804, significant at p < 0.01 level, and a quite high R² value of 0.606. This means that about 60.6% of all variance can be calculated through the use of this variable. The third hypothesis is tested by conducting a regression analysis for the percentage of negative comments and the number of car sales. This analysis showed a weak and nonsignificant but negative correlation which fits the expectation of the research model. Finally, an estimate of the predictive power of Google Trends was obtained by conducting a regression analysis for Google Trends and sales which also showed a not significant but positive correlation.

Inclusion of time lags into the model

The dataset is split per car model into smaller sets to analyze each car model separately. Furthermore, a time lag was incorporated. This was done by using the cross-correlation function of SPSS and checking, for which time lag (smaller than or equal to 12 months positive or negative) between the independent variable and the car sales the correlation was the strongest. Table 3 shows per car model and variable, for which time lag the strongest correlation was found. In case no number is given, a correlation fitting the research model could not be found. An asterisk marks whether this correlation was significant at p ≤ 0.05 or not.

Table 3. Optimal time lags and Pearson's correlations per car model								
	PNratio x Sales		Number total mentions x Sales		Google Trends Score x Sales		Negative mentions x Sales	
Car model	*Lag*	*Corr*	*Lag*	*Corr*	*Lag*	*Corr*	*Lag*	*Corr*
Fiat Panda	11	0.027	3	0.424*	3	0.694*	-	-
Ford Fiesta	8	0.223	12	0.247	12	0.257	8	-0.164
Opel Corsa	3	0.176	7	0.155	2	0.271	10	-0.152
Honda Civic	9	0.136	12	0.256	1	0.519*	9	-0.206
VW Golf	0	0.284*	7	0.388*	4	0.473*	8	-0.212
VW Passat	-5	0.146	4	0.508*	5	0.486*	2	-0.207
Ford Mondeo	9	0.132	9	0.499*	5	0.344*	2	-0.306*
BMW 5 Serie	9	0.152	-	-	12	0.015	11	-0.064
Mercedes-Benz E-Class	9	0.205	5	0.063	12	0.107	5	-0.118
Porsche Panamera	10	0.055	9	0.422*	9	0.348*	11	-0.092
Porsche 911	9	0.244	4	0.245	9	0.277	2	-0.227
Averages	6.5	0.162	6.5	0.192	6.7	0.345	6.1	-0.158

The correlations varied widely per car model. This supports the assumption that consumer behavior does indeed vary per model. Contrary to our hypotheses, the PNR is a weak predictor for car sales in our dataset, even with the inclusion of a time lag. Only one out of eleven found correlations were significant and this correlation had a time lag of 0 which means that the social media sentiments occurred in the same month as the corresponding car sales. It can therefore not be used to predict sales. In comparison to the Google Trends values, which showed many moderate and significant correlations, the PNratio of the analyzed models does not seem to have any predictive power towards car sales. H_1 can therefore only be accepted for the model VW Golf.

Besides the relative search volume, the total number of mentions also correlates significantly with car sales in five cases. Although the relationships are not as strong as initially suggested by the average data, this makes it the second strongest predictor of sales from the analyzed variables. Furthermore, as expected, the percentage of negative mentions negatively correlates with car sales; however, this is only significant in one case.

We expected weaker correlation Google Trends (interest) than for the PNR (desire), but the reverse is the case and except one all PNR correlations with sales volume are not significant. We also would expect a larger time lag for mentions than for PNR, but again this is not true, their summed averages are equal. As expected, negative sentiments would result in a negative correlation with sales, but this was only significant for one car model.

Influence of price on correlations

To analyze H_4, the eleven car models were split into two price classes. The lower price class contained all cars with a starting price of less than 30,000€ (five models) while the higher price class contained the cars who started at 30,000€ or more (six models). The prices were taken from the Top Gear website (http://www.topgear.nl/koopgids/nieuw/ accessed May 10, 2017). The correlations found in Table 5 are then compared in an independent t-test between the two groups. Since a t-test requires the variables to be normally distributed, the Shapiro-Wilk test for normality was conducted for every group (each correlation divided into high- and low-priced cars) whereby all variables were found to be normally distributed. The results for the four independent t-tests are shown in Table 4.

Table 4. t-tests for higher and lower priced car models' correlations with sales		
Correlations with sales	Mean difference	Significance
PNratio	-0.0006	0.990
Total mentions	-0.02333	0.985
Negative mentions	0.69	0.320
Google Trends	-0.1704	0.950

Since the PNR already showed no significant connection to car sales, it also was not surprising that the car price had no significant influence on this correlation either. The mean correlation for lower priced cars with sales is 0.160, and the mean correlation for higher priced cars is 0.161. This leads to a mean difference of 0.0006 which is almost negligible. Furthermore, since Levene's test showed insignificant results, equal variances between the high- and low-priced group exist.

Equally, when looking at the differences between high and low priced cars concerning the correlations of total mentions, the percentage of negative mentions, and Google Trends with sales, no significant difference is visible. Equal to the first test, Levene's test indicated equal variances in all cases which leads to the conclusion that there is no difference between the high and low priced car samples. H_5 is therefore rejected for all analyzed cars.

Establishing a prediction formula

After testing the variables for each car model separately and finding some moderate correlations, the question remains whether a combination of independent variables can lead to a more reliable prediction of sales. Previously, the total number of mentions about a car model and the Google Trends score were identified as the strongest predictors.

A decision tree regression is performed by using the M5P classifier of data mining tool WEKA, which also includes the different car models as a variable. If the classifier is evaluated as a successful improvement, a decision time lag is implemented to establish the prediction formula's, resulting in a decision tree regression estimation with three independent variables (car model, number of mentions, and Google Trends score) and the dependent variable car sales per car model.

Before performing any parametric tests with combinations of variables, the independent variables were tested for multicollinearity. In case a multicollinearity was found this could reduce the impact of a linear regression analysis since the variables are dependent on each other. This was tested by determining the Variance Inflation Factor (VIF) for a regression analysis of the two variables included in the model. The VIF of the Google Trends score and the total number of mentions was 1,110. Since multicollinearity is only considered to be a problem in case the VIF is greater than 10, both variables can be used with confidence in combinations with each other.

As stated above, when trying to find a useful prediction formula it is important to include an appropriate time lag that gives car sellers and manufacturers enough time to react to the forecast. WEKA does not have the option to include an optimum time lag for each model into one decision tree. Therefore, a time lag suitable for all models needs to be defined: Production planning in the automotive industry usually starts three months before the manufacturing and can be changed at the latest one month beforehand (Suthikarnnarunai 2008). The lead time varies from a few weeks to a few months per car producer. For this formula, the time lag will be set to four months which gives manufacturers time to plan one month ahead of production and three months for production and distribution of the cars. An optimal time lag of

four or five months was also found in Table 3 and is therefore expected to bring the best results. Nevertheless, the same procedure can also be used to establish a decision tree with another time lag. The time lag was added by shifting the number of sales in the dataset upwards. The m5p analysis is given in Figure 2.

Scheme: weka.classifiers.trees.M5P -M 4.0
Relation: Data incl flexible time lags-weka.filters.unsupervised.attribute.Remove-R2,4-5,8-10
Instances: 466 (this is the sum of the number of months (total period covered was 52 months), times the number of car models (i.e. 1 for A, 1 for F 1 for S and two for B, C, D, and E categories). These are the instances used in the model training.
Attributes: 4
           Car model; Number total mentions; Google Trends Score; Sales
Test mode: split 66.0% train, remainder test
=== Classifier model (full training set) ===
M5 pruned model tree (using smoothed linear models)
Car model=Volkswagen Passat, Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf <= 0.5: LM1 (260/12.826%)
Car model=Volkswagen Passat, Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf > 0.5:
| Number total mentions <= 1471.5 : LM2 (133/51.484%)
| Number total mentions > 1471.5 : LM3 (73/123.165%)
LM num: 1 Sales =
   - 23.8761 * Car model=Honda Civic, Mercedes Benz E-Klasse, Ford Mondeo, BMW 5 series, Volkswagen Passat, Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 22.6675 * Car model=Mercedes Benz E-Klasse, Ford Mondeo, BMW 5 series, Volkswagen Passat, Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 134.4791 * Car model=Ford Mondeo, BMW 5 series, Volkswagen Passat, Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 6.1353 * Car model=Volkswagen Passat, Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 13.5125 * Car model=Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - - 7.7129 * Car model=Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 12.4077 * Car model=Ford Fiesta, Volkswagen Golf
   - + 0.0053 * Number total mentions
   - + 2.6319 * Google Trends Score
   - - 183.4712
LM num: 2 Sales =
   - 16.736 * Car model=Ford Mondeo, BMW 5 series, Volkswagen Passat, Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 7.6344 * Car model=Volkswagen Passat, Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 413.2376 * Car model=Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - - 167.065 * Car model=Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 30.5117 * Car model=Ford Fiesta, Volkswagen Golf
   - + 0.0143 * Number total mentions
   - + 15.7286 * Google Trends Score
   - - 931.9067
LM num: 3 Sales =
   - 16.736 * Car model=Ford Mondeo, BMW 5 series, Volkswagen Passat, Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 7.6344 * Car model=Volkswagen Passat, Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 79.8739 * Car model=Fiat Panda, Opel Corsa, Ford Fiesta, Volkswagen Golf
   - - 62.1503 * Car model=Opel Corsa, Ford Fiesta, Volkswagen Golf
   - + 259.1746 * Car model=Ford Fiesta, Volkswagen Golf
   - + 0.0196 * Number total mentions
   - + 29.0436 * Google Trends Score
   - - 1709.9792
Number of Rules : 3

Figure 2. Formatted output of M5P classifier in WEKA without time lag

---

For the interpretation of this result, we give the decision tree in Figure 3 and an example formula that can be derived for Sales volume of the VW Passat (which applies model 1). The formula is
Sales VW Passat = 23.8761
  + 22.6675
  + 6.1353
  + 0.0053 * Number total mentions
  + 2.6319 * Trends_score
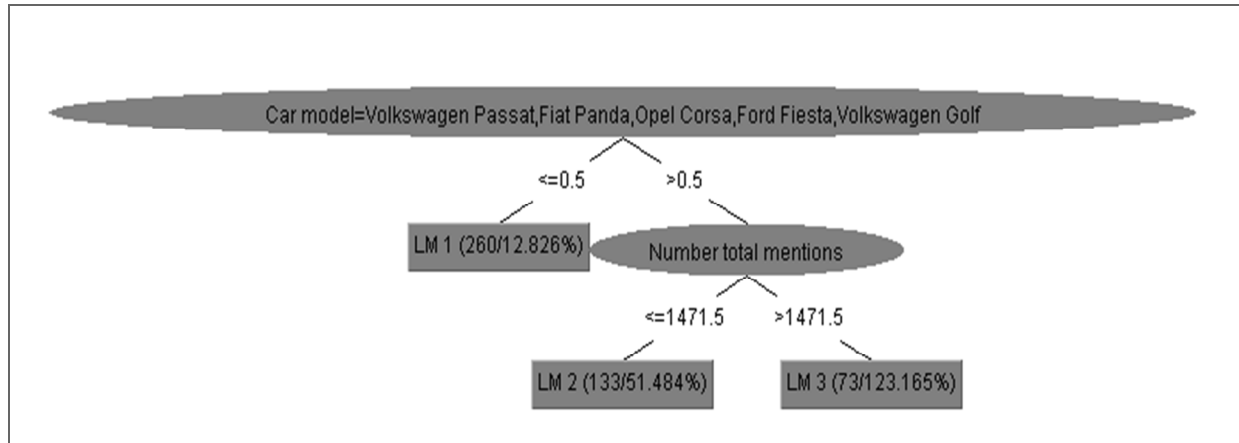 - 260.1935

Figure 3. Decision tree output M5P classifier including a four-month time lag

The decision tree of Figure 3 was built on 66% of the data and tested on the remaining 34% to ensure that the model is not only applicable to this specific dataset but also to data of the holdout set. The outcome of the decision tree is one of three linear models that have to be used depending on the car model.

As displayed in Table 5, the results were good with a correlation of 0.7313 and an applicability of 100%. Since the simple linear regression in Table 5 already brought significant results and the decision tree is also based on linear models it is assumed that this improved model is also significant. The root mean squared error of 328.9 is still quite high but is much lower than in the previous models. Practically, the model could perhaps be used as an addition to sales planning if it is further tested and improved.

Table 5. Summary of M5P classifier decision tree regressions output	
Correlation coefficient	0.7313
Mean absolute error	141.4526
Root mean squared error	328.8579
Relative absolute error	41.4809 %
Root relative squared error	68.5924 %
Total Number of Instances	158 (Note: this is the number of instances of the hold out set)

Conclusions, Discussion, and Limitations

Conclusions

Our first research questions was *'What is the predictive power of sentiments for car models expressed on social media towards car sales in the Netherlands?'* and our section question was *'What is the best predictor of car sales in The Netherlands: social media sentiments or Google Trends data?'.* The first question is answered by searching for significant and meaningful correlations between the PNratio (quotient of positive and negative social media posts) and sales of particular car models. The results showed that sentiments have little predictive power towards car sales in the Netherlands for the period we studied. Although the directions of the relationships seemed to fit the research model (with the PNR correlating positively and negative comments correlating negatively with sales), the found relationships are weak. Per variable, only for one car model, a significant correlation was found although these were

equally weak. Hypothesizes 1 and 4 were therefore rejected for ten car models and both only accepted for one.

While sentiments showed weak results, the general attention about a car model (represented by the total number of mentions) and the search volume (Google Trends score) correlated better with sales and in many cases significantly. Hypothesis 2 is therefore corroborated and specifically accepted for four car models.

For answering the second research question, we stated in H5 that PNratio would be a better sales predictor than Google trends. However, hypothesis 5 is rejected since Google Trends showed to be a much stronger predictor of sales than the PNRatio and the percentage of negative mentions in all cases. This result matches other research such as Wu and Brynjolfsson (2015), Choi and Varian (2012) and Yang et al. (2015) on Google Trends predictions. A combination of the two strongest predictors of a decision tree regression led to a prediction model that approximated the sales quite well. It could be used on 100% of the data in the dataset and showed a correlation of about 0.85. Although the root mean squared error was quite high, the model could be used as an addition to traditional sales forecasting methods if tested further.

Furthermore, a comparison between higher priced and lower priced cars showed that there was no significant difference concerning the strength of the correlations between the analyzed variables and sales. The assumption that data of higher priced cars would show weaker correlations with sales (H_4) is therefore rejected. All conclusions are summarized in Table 6.

Table 6. Conclusion of hypothesizes per car model		
H_1	The PNR of social media mentions about a car model has a positive influence on sales of this model	Accepted for VW Golf, but with a time lag of zero
H_2	The number of total mentions about a car model on social media correlates positively with the number of car sales	Accepted for VW Golf, VW Passat, Ford Mondeo, Porsche Panamera, with positive time lags
H_3	The percentage of negative mentions about a car model has a negative influence on the number of sales of this model.	Accepted for the Ford Mondeo only
H_4	The higher the price of a car, the weaker the correlation between the social media data and the sales.	Rejected for all analyzed correlations
H_5	The correlation of sentiments with car sales is higher than the correlation of relative search volume with car sales.	Rejected

While other scholars showed that sales of lower priced items such as movie tickets and iPhones could be predicted very accurately, no evidence is found that this also yields for cars. Speculations can be made on why this is the case although they will have to be verified in other research. The results suggest that consumers do not let the opinion of other people ('subjective norm' according to Ajzen (1991)) in an online environment influence their buying choices when it comes to more expensive items. This assumption also questions the usefulness of some part of the Theory of Planned Behavior in the context of high involvement purchases and virtual communities, although we only did observations for car models, this leaves food for further research on other products. Another assumption is that other decision factors outweigh the subjective norm when deciding for a car model. Car sales seem to be strong policy driven which means that consumers buying a car might consider factors such as the tax class of a model as more important than the opinion of other people (Barreira et al. 2013). Especially in the Netherlands, taxes between car models vary significantly due to the green policy of the Dutch government (Crisp 2014). The data also showed that car sales seem to vary per season although this was not the same in every year. While mostly car sales dropped towards the end of the year, they increased rapidly at the end of the year 2015 (see Figure 2), presumably because a new policy starting in the next year would require buyers to pay much higher taxes (Barreira et al. 2013)(Automotiveimport 2015).

Implications

In contrast to what was expected, social media sentiments have little predictive power towards Dutch car sales. The research has primarily helped to further explore the boundaries of sentiment mining. This research contrasts other research within the field of sentiment analysis such as Asur and Huberman (2010) and Lassen et al. (2014) to name just a few researchers who received very good results when using

this technique. Furthermore, research using search engine volumes to forecast high priced items such as house sales (Wu & Brynjolfsson, 2013) or tourist volume (Yang et al., 2015) equally showed significant results with high R² values which makes it even more surprising that the same does not yield for sentiment analysis in this context. The research, therefore, has found new limitations that were previously unknown and should certainly be further investigated.

The decision tree regression leading to specific prediction formulas for car sales is one of the practical contributions of this research. Although it is advised to test and improve the model on more data and with other analytic techniques (like neural networks, gradient boosted trees, and random forest algorithms) before using it in sales forecasting, the results found in this paper look useful. It is important for dealers to know that social media sentiments do not necessarily directly represent or influence the performance of their company. A lot of positive attention will therefore not necessarily lead to high sales within a few months without the car dealer making a good effort. Equally, negative comments are no reason to expect a significant decrease in sales. Nevertheless, general, extensive and longtime taking negative publicity will have a negative impact at the end.

Further research

This study provided unexpected results that open up further questions. Firstly, the question whether sentiment mining has an equally weak predictive power when applied to other car models or other high priced items should be further investigated. The car industry was chosen as representative for high involvement and expensive purchases, however, the findings cannot be generalized for other high priced items. As mentioned previously, this study is based on the Dutch car market, which has its peculiarities and thus generalization over other car markets needs other research.

Second, there is no proof resulting from this study that the research model is also the actual causal model. Based on the previous speculations it is, therefore, advisable for further research to focus on the nature of the causal relationships between the applied variables and on their directions. For this we suggested and tried to operationalize the AIDA purchasing model, in which we suggested mention volumes to correspond with attention, Google Trends data with interest, and PNR with desire. However, the weak correlation of PNR with sales would be against expectations (high desire likely results in high sales). This may two implications. One is that PNR is not necessarily desire but a possible indication of service needs (Pang et al. 2002) and two that in our model we did not include macro-economic like oil prices (Geva et al. 2017)(Elshendy et al. 2017) and political indicators

Third, the relationships among the analyzed variables were found using linear models since these were determined as most appropriate and the requirements for the use of linear regression were fulfilled. Initial tests with the data set showed that using quadratic or cubic equations would not improve the results significantly and in some cases, distort the direction of the data (for example by creating an upside-down parabola which would mean that sales drop again if the PNratio increases beyond a certain turning point). However, since the errors with the linear regression were quite big (although normally distributed) and many correlations not very strong further research could also use nonparametric models to test the strength of the relationships.

Lastly, the conclusions made are based on data for eleven car models. The findings are therefore also restricted to these car models. Since the car models among each other showed great differences in terms of which factor has the strongest predictive power, further research is needed that includes more car models.

Acknowledgements

References

Ajzen, I. 1991. "The theory of planned behavior," *Organizational Behavior and Human Decision Processes* (50:2), pp. 179–211 (doi: 10.1016/0749-5978(91)90020-T).

Ajzen, I., and Fishbein, M. 1980. *Understanding attitudes and predicting social behavior*, Englewood

Cliffs, N.J.: Prentice-Hall (available at https://books.google.nl/books?id=AnNqAAAAMAAJ).

Asur, S., and Huberman, B. A. 2010. "Predicting the future with social media," in *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on* (Vol. 1), IEEE, pp. 492–499.

Automotiveimport. 2015. "BPM 2016: De BPM gaat weer fors omhoog, dit moet je weten," (available at http://www.automotiveimport.nl/bpm/bpm-2016-fors-omhoog; retrieved May 10, 2016).

Barreira, N., Godinho, P., and Melo, P. 2013. "Nowcasting unemployment rate and new car sales in south-western Europe with Google Trends," *NETNOMICS: Economic Research and Electronic Networking* (14:3), Springer, pp. 129–165.

Bing, L., Chan, K. C. C., and Ou, C. 2014. "Public sentiment analysis in Twitter data for prediction of a company's stock price movements," *2014 Ieee 11th International Conference on E-Business Engineering (Icebe)*, pp. 232–239 (doi: 10.1109/icebe.2014.47).

Brander, J. A., and Spencer, B. J. 2015. "Intra-industry trade with Bertrand and Cournot oligopoly: The role of endogenous horizontal product differentiation," *Research in Economics* (69:2), Elsevier, pp. 157–165.

Cattani, G., Porac, J. F., and Thomas, H. 2017. "Categories and competition," *Strategic Management Journal* (38:1), Wiley Online Library, pp. 64–92.

CBS. 2015. "Gebruik sociale netwerken sterk toegenomen," (available at https://www.cbs.nl/nl-nl/nieuws/2015/27/gebruik-sociale-netwerken-sterk-toegenomen; retrieved May 10, 2016).

Ceron, A., Curini, L., Iacus, S. M., and Porro, G. 2013. "Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France," *New Media & Society* (16:2), pp. 340–358 (doi: 10.1177/1461444813480466).

Choi, H., and Varian, H. 2012. "Predicting the present with Google Trends," *Economic Record* (88:special issue SI), pp. 2–9.

Crisp, J. 2014. "Dutch car tax regime leaves Germany far behind in curbing CO2 emissions," *EurActiv* (available at http://www.euractiv.com/section/transport/news/dutch-car-tax-regime-leaves-germany-far-behind-in-curbing-co2-emissions/; retrieved May 10, 2016).

Davis, F. D. 1989. "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly* (13:3), Management Information Systems Research Center, University of Minnesota, pp. 319–340 (doi: 10.2307/249008).

Dijkman, R., Ipeirotis, P., Aertsen, F., and van Helden, R. 2015. "Using twitter to predict sales: a case study," Beta Research School, Eindhoven (available at https://arxiv.org/ftp/arxiv/papers/1503/1503.04599.pdf).

Elshendy, M., Colladon, A. F., Battistoni, E., and Gloor, P. A. 2017. "Using four different online media sources to forecast the crude oil price," *Journal of Information Science*, SAGE Publications Sage UK: London, England, p. 165551517698298 (doi: 10.1177/0165551517698298).

Epstein, R., and Robertson, R. E. 2015. "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections.," *Proceedings of the National Academy of Sciences of the United States of America* (112:33), pp. E4512-21 (doi: 10.1073/pnas.1419828112).

Erkan, I., and Evans, C. 2016. "The influence of eWOM in social media on consumers' purchase intentions: An extended approach to information adoption," *Computers in Human Behavior* (61), pp. 47–55 (doi: http://dx.doi.org/10.1016/j.chb.2016.03.003).

Fantazzini, D., and Toktamysova, Z. 2015. "Forecasting German car sales using Google data and multivariate models," *International Journal of Production Economics* (170), pp. 97–135 (doi: 10.1016/j.ijpe.2015.09.010).

Gensler, S., Neslin, S. A., and Verhoef, P. C. 2017. "The Showrooming Phenomenon: It's More than Just About Price," *JOURNAL OF INTERACTIVE MARKETING* (38), 360 PARK AVE SOUTH, NEW YORK, NY 10010-1710 USA: ELSEVIER SCIENCE INC, pp. 29–43 (doi:

10.1016/j.intmar.2017.01.003).

Geva, T., Oestreicher-Singer, G., Efron, N., and Shimshoni, Y. 2017. "Using Forum and Search Data for Sales Prediction of High-Involvement Products," *Management Information Systems Quarterly* (41:1), pp. 65–82.

Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M., and Watts, D. J. 2010. "Predicting consumer behavior with web search," *Proceedings of the National Academy of Sciences of the United States of America* (107:41), National Academy of Sciences, pp. 17486–17490 (available at http://www.jstor.org/stable/20780485).

Hunt, S. D., and Morgan, R. M. 1995. "The comparative advantage theory of competition," *The Journal of Marketing*, JSTOR, pp. 1–15.

Kandaswami, K., and Tiwar, A. 2014. "Deloitte. Driving through the consumer's mind: Steps in the buying process," Deloitte Touch India (available at http://www2.deloitte.com/content/dam/Deloitte/in/Documents/manufacturing/in-mfg-dtcm-steps-in-the-buying-process-noexp.pdf).

Karlgren, J., Sahlgren, M., Olsson, F., Espinoza, F., and Hamfors, O. 2012. "Usefulness of sentiment analysis," *Lecture Notes in Computer Science* (7224), pp. 426–435.

Kotler, P. J. 1994. *Marketing management : analysis, planning, implementation, and control* (8th ed.), Englewood Cliffs, N.J.: Prentice Hall.

Lassen, N. B., Madsen, R., and Vatrapu, R. 2014. "Predicting iphone sales from iphone tweets," in *2014 IEEE 18th International Enterprise Distributed Object Computing Conference*, IEEE (doi: 10.1109/edoc.2014.20).

Lee, J., Park, D.-H., and Han, I. 2008. "The effect of negative online consumer reviews on product attitude: An information processing view," *Electronic Commerce Research and Applications* (7:3), pp. 341–352 (doi: 10.1016/j.elerap.2007.05.004).

Liu, B. 2012. "Sentiment analysis and opinion mining," *Synthesis Lectures on Human Language Technologies* (5:1), pp. 1–167 (doi: 10.2200/S00416ED1V01Y201204HLT016).

Mislove, A., Lehmann, S., Ahn, Y., Onnela, J., and Rosenquist, J. 2011. "Understanding the Demographics of Twitter Users," *Fifth International AAAI Conference on Weblogs and Social Media* (N. Nicolov and J. Shanaha, eds.), Barcelona: AAAI digital library, pp. 17–21.

Nguyen, T. H., Shirai, K., and Velcin, J. 2015. "Sentiment analysis on social media for stock movement prediction," *Expert Systems with Applications* (42:24), pp. 9603–9611 (doi: 10.1016/j.eswa.2015.07.052).

Office for Official Publications of the European Communities. 1999. "Regulation (EEC) No 4064/89 merger procedure," (available at http://ec.europa.eu/competition/mergers/cases/decisions/m1406_en.pdf).

Pang, B., and Lee, L. 2008. "Opinion Mining and Sentiment Analysis," *Foundations and Trends in Information Retrieval* (2:2), pp. 91–231 (doi: 10.1561/1500000001).

Pang, B., Lee, L., and Vaithyanathan, S. 2002. "Thumbs up? Sentiment classification using machine learning techniques," *Proceedings of the Conference on Empirical Methods in Natural*, pp. 79–86.

Putsis, W. P., and Srinivasan, N. 1994. "Buying or just browsing? The duration of purchase deliberation," *Journal of Marketing Research* (31:3), US: American Marketing Association, pp. 393–402 (doi: 10.2307/3152226).

Putsis, W. P., and Srinivasan, N. 1995. "So, how long have you been in the market? The effect of the timing of observation on purchase," *Managerial and Decision Economics* (16:2), John Wiley & Sons, Ltd., pp. 95–110 (doi: 10.1002/mde.4090160202).

Reyes, A., Rosso, P., and Buscaldi, D. 2012. "From humor recognition to irony detection: The figurative language of social media," *Data & Knowledge Engineering* (74), pp. 1–12.

Serrano-Guerrero, J., Olivas, J. A., Romero, F. P., and Herrera-Viedma, E. 2015. "Sentiment analysis: A review and comparative analysis of web services," *Information Sciences* (311:August 2015), pp. 18–38 (doi: 10.1016/j.ins.2015.03.040).

Suthikarnnarunai, N. 2008. "Automotive supply chain and logistics management," *Imecs 2008: International Multiconference of Engineers and Computer Scientists, Vols I and Ii*, pp. 1800–1806.

Tuarob, S., Tucker, C. S., and Asme. 2014. "Fad or here to stay: predicting product market adoption and longevity using large scale, social media data," *Proceedings of the Asme International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2013, Vol 2b* (doi: V02bt02a012).

Wijnhoven, F., and Bloemen, O. 2014. "External validity of sentiment mining reports: Can current methods identify demographic biases, event biases, and manipulation of reviews?," *Decision Support Systems* (59:1), pp. 262–273 (doi: 10.1016/j.dss.2013.12.005).

Wilson, R. E., Gosling, S. D., and Graham, L. T. 2012. "A Review of Facebook Research in the Social Sciences," *Perspectives on Psychological Science* (7:3), pp. 203–220 (doi: 10.1177/1745691612442904).

Wu, L., and Brynjolfsson, E. 2015. "The future of prediction: How Google searches foreshadow housing prices and sales," in *Economic Analysis of the Digital Economy*A. Goldfarb, S. Greenstein, and C. Tucker (eds.), Chicago, Illinois: University of Chicago Press, pp. 89–118.

Yang, X., Pan, B., Evans, J. A., and Lv, B. F. 2015. "Forecasting Chinese tourist volume with search engine data," *Tourism Management* (46), pp. 386–397 (doi: 10.1016/j.tourman.2014.07.019).