

Een voorspelling van de Tweede Kamer- en Provinciale Statenverkiezingen aan de hand van tweets met demografische gegevens

In hoeverre kunnen demografische gegevens, ‘geslacht’ en ‘leeftijd’, van twitteraars met behulp van TweetGenie de schattingen van verkiezingsuitslagen, middels tweets, van de Tweede Kamer (2012) en de Provinciale Staten (2015) nauwkeuriger maken?

Masterthesis

Nieuwe Media, Taal en Communicatie

Naam: Michelle de Gier

Studentnummer: s4494571

E-mail: mc.degier@student.ru.nl

Begeleiders: Drs. E. Sanders & Prof. Dr. A.P.J. van den Bosch

15 juli 2016

Samenvatting

Voorgaande onderzoeken laten zien dat er min of meer gedegen voorspellingen van politieke verkiezingsuitslagen gemaakt kunnen worden aan de hand van Twitterdata. Echter, veel kritische onderzoekers betogen dat in deze onderzoeken geen rekening wordt gehouden met een belangrijk criterium: een representatieve steekproef, vaak gecontroleerd aan de hand van demografische gegevens. Bij Twitterdata zijn de demografische gegevens daarentegen vaak onbekend. Hierdoor zouden de onderzoeken niet representatief zijn voor de stemgerechtigde populatie. De huidige studie onderzocht in hoeverre demografische gegevens, ‘geslacht’ en ‘leeftijd’, van twitteraars de schattingen van de verkiezingsuitslagen aan de hand van tweets over de Tweede Kamer (2012) en de Provinciale Staten (2015) nauwkeuriger konden maken. TweetGenie is gebruikt om politieke twitteraars te classificeren op leeftijd en geslacht. Dit is een automatische methode die ingezet kan worden bij het labelen van Twitteraccounts op demografische gegevens. Er is een voorspelling gedaan van de politieke verkiezingsuitslagen voor 2012 en 2015 door enkel frequenties te tellen en daarnaast is een voorspelling gedaan door de tweets recht te trekken middels de demografische gegevens die voortkwamen uit TweetGenie. Deze voorspellingen zijn afgezet tegen de daadwerkelijke verkiezingsuitslagen en de opiniepeilingen. Het blijkt dat opiniepeilingen nog altijd een nauwkeurigere voorspelling weergeven, maar dat de voorspellingen middels de Twitterdata met en zonder aanpassing op demografie ook nauwkeurig zijn. Daarnaast is te zien dat mannelijke en vrouwelijke twitteraars verschillend politiek Twittergedrag vertonen, alsook twitteraars in verschillende leeftijdscategorieën. Twitter kan dienen als een analysetool, waarmee het mogelijk is te monitoren welke mensen waar over praten op Twitter. Het correct voorspellen van verkiezingsuitslagen blijft vooralsnog complex, mede doordat het uiteindelijk een voorspelling met terugwerkende kracht betreft. Desalniettemin is met deze studie wel een belangrijke bias betreffende social media data gereduceerd, waar dit zowel in het politieke alsook in onder andere het gezondheids - en marketingdomein waardevol is.

Inleiding

In 2006 werd Twitter gelanceerd en inmiddels zijn er in totaal 2,6 miljoen Nederlandse gebruikers (RTL Nieuws, 2016). Per dag worden er wereldwijd zo'n 500 miljoen tweets verstuurd. Vanaf 2010 zijn Nederlandse gebruikers samen goed voor de productie van zo'n 2,7 miljard tweets (Tjong Kim Sang & Van den Bosch, 2013). Er is tot legio data tegelijk toegang; Twitter is hiermee een enorme bron van informatie voor vele onderzoekers. Mede doordat 91% van de Twittergebruikers een openbaar profiel heeft. Dit houdt in dat iedereen deze berichten kan lezen (Mislove, Lehmann, Ahn, Onnela & Rosenquist, 2011). Het is een veelgebruikt medium om berichten, meestal persoonsgebonden, te versturen. Tweets gaan veelal over het dagelijks leven van de schrijvers van die tweets. Dit is gerelateerd aan Twitters vraag "What's happening?", maar tweets kunnen ook de opinie van de schrijver bevatten. Het medium is hierdoor aantrekkelijk voor bijvoorbeeld *opinion mining* en sentimentanalyses (Pak & Paroubek, 2010). Ook kan Twitterdata heel waardevol zijn voor onderzoek naar real world events. Zo wordt geprobeerd real world events als politieke verkiezingen te monitoren aan de hand van Twitter (Sanders & Van den Bosch, 2013). De laatste keer dat een Nederlandse verkiezing te monitoren was, was toen de stembus mochten voor de Provinciale Statenverkiezingen op 18 maart 2015 naar de stembus mochten (Kiesraad, z.d.a). De meest recente Tweede Kamerverkiezingen vonden plaats op 12 september 2012 (Kiesraad, z.d.b). In Nederland is iedereen vanaf achttien jaar stemgerechtigd, waarvan 75% stemde tijdens de Tweede Kamerverkiezingen in 2012 (Kiesraad, 2012) en zo'n 47% een stem uitbracht tijdens de Provinciale Statenverkiezingen in 2015 (NOS, 2015). In die perioden worden opiniepeilingen veelal gebruikt als een indicatie van de uitslagen.

De opiniepeilingen van bijvoorbeeld Maurice de Hond geven inzicht in de politieke voorkeuren van de kiezers. Hierbij wordt een representatieve groep gevraagd op welke partij zij van plan is te stemmen. Rond verkiezingstijd wordt de berichtgeving van de media overladen met dergelijke opiniepeilingen. Met die opiniepeilingen, ook wel polls genoemd, wordt geprobeerd een voorspelling van de uiteindelijke verkiezingsuitslag te maken. Toch zitten deze polls er nogal eens naast, het blijven immers steekproeven (NRC, 2012). Met de komst van nieuwe media, zoals Twitter, hebben zich ook nieuwe methoden ontwikkeld om voorspellingen te doen over de "echte" wereld, mede doordat het medium veel data bevat (Nguyen, Gravel, Trieschnigg & Meder, 2013a; Culotta, Ravi & Cutler, 2015). Zo kunnen naast de opiniepeilingen ook de opinies van twitteraars gemonitord worden om een beeld te vormen van een waarschijnlijke verkiezingsuitslag (Tumasjan, Sprenger, Sandner & Welp, 2010; Tjong Kim Sang & Bos, 2012; Sanders & Van den Bosch, 2013).

Ook in andere domeinen is Twitter gebruikt om real world events te analyseren. Zo beoogden Signorini, Segre en Polgreen (2011) met hun onderzoek een relatie tussen tweets en de epidemische curve van H1N1 (influenza, ofwel de Mexicaanse griep) te vinden. Zij monitorde de epidemie achteraf, waaruit bleek dat er een relatie bestond tussen het aantal tweets dat werd verstuurd, met trefwoorden gerelateerd aan H1N1, en de epidemische curve van H1N1. Naast epidemieën is er ook onderzoek gedaan naar de voorspelbaarheid van Twitter tijdens talentenjachten zoals American Idol. Met succes is een schatting gemaakt van de winnaar van de talentenjacht van 2012 aan de hand van de frequentie van tweets, middels gebruikersnamen en hashtags over American Idol, die in de Verenigde Staten werden verstuurd. Zoals de onderzoekers zelf stelden, hebben zij een ogenschijnlijk eenvoudig onderzoek uitgevoerd. Zij namen namelijk aan dat de twitteraars representatief zijn voor de stemmers. Onder andere doordat de finalisten door de stemmers zelfgekozen zijn en de uitkomst van American Idol niet direct invloed heeft op henzelf, waar dit bij bijvoorbeeld politieke verkiezingen wel het geval is. Mensen stemmen vaak op de politieke partij die het meest hun belangen behartigt: dit is bij de winnaar van American Idol niet van toepassing (Ciulla et al., 2012).

In diverse disciplines is Twitter dus ingezet om zogeheten voorspellingen te doen van real world events. In dit onderzoek ligt de focus op het politieke domein. Daarbij is de vraag in hoeverre Twitter een voorspellende kracht kan hebben bij verkiezingsuitslagen, expliciet voor de Tweede Kamerverkiezingen in 2012 en de Provinciale Statenverkiezingen in 2015. Tevens is in deze discipline nog onbekend of de twitteraars representatief zijn voor de stemgerechtigde populatie, zoals dit bij de voorspelling van de winnaar van American Idol geen beperking was. Oftewel in hoeverre twitteraars, betreffende demografie, een afspiegeling van de daadwerkelijke stemmers in Nederland vormen. TweetGenie kan helpen in het verkrijgen van inzicht in deze demografische gegevens van twitteraars. Door middel van automatische classificatie kunnen twitteraars ingedeeld worden op onder andere geslacht en leeftijd. Hoe nauwkeurig kunnen de verkiezingsuitslagen voorspeld worden met deze nieuwe inzichten? En zouden daarmee de voorspellingen van opiniepeilingen overtroffen kunnen worden?

Voorspelbaarheid politieke verkiezingen aan de hand van Twitter

Twitter is in het politieke domein al vaker ingezet om voorspellingen te doen. Zo is Twitter gebruikt om de Duitse verkiezingen in 2009 te voorspellen. Er kon een weerspiegeling gecreëerd worden van de voorkeuren van stemmers. Gesuggereerd werd, dat middels de

frequenties van tweets per partij, een voorspelling gedaan kon worden van de verkiezingsuitslag. De tellingen die gedaan werden, kwamen dicht bij de voorspellingen van de polls (Tumasjan et al., 2010). Jüngherr, Jürgens en Schoen (2012) reageren met een artikel op dat onderzoek. Zij hebben commentaar op de studie van Tumasjan et al., omdat één partij (de Piraten Partij) niet was opgenomen in de resultaten, terwijl die het hoogst zou eindigen. Daarnaast viel de tijdsperiode waarin de tweets zijn verzameld, volgens Jüngherr et al., te overdenken: acht dagen voor de verkiezingen werd gestopt met het verzamelen van tweets.

Daarnaast deden Sanders en Van den Bosch (2013) onderzoek naar de voorspelling van de Tweede Kamerverkiezingen in 2012 op basis van tweets en vergeleken dit met de voorspelling door polls. Zij telden hoe vaak iedere politieke partij in de tien dagen voor de verkiezing genoemd werd in tweets. De frequentie van de tweets over een politieke partij correleerde met de voorspellingen die in de polls werden gedaan. De studie liet zien dat polls vooralsnog een betere voorspelling geven over de verkiezingsuitslag, waarbij de correlatie 0.98 was. Toch was de correlatie (0.96) van de verkiezingsuitslag met tweets ook hoog. Wel zijn er een aantal uitschieters in de frequentie van tweets te zien, die vaak verklaard konden worden door opvallendheden in het nieuws over verschillende partijen. Een voorbeeld hiervan was dat op de dagen dat de PVV een zeer hoge frequentie aan tweets had, zij in het nieuws waren betreffende het onderwerp “vals geld”. Ook lag de voorspelling voor GroenLinks standaard hoger, waarschijnlijk door het actieve social mediabeleid dat zij voeren.

Hoewel er studies zijn waarin het tellen van tweets succesvol blijkt te zijn om politieke verkiezingsuitslagen te voorspellen, stellen verschillende onderzoekers (Tumasjan et al., 2010; Tjong Kim Sang & Bos, 2012; Schoen et al., 2013) de kanttekening dat de demografische gegevens van de twitteraars niet bekend zijn. Hierdoor kan niet met stelligheid gezegd worden dat de onderzoeken representatief zijn voor de populatie; de stemgerechtigden. Het is een uitdaging van *big data*, in dit geval de data van Twitter, om uitspraken te kunnen doen over real world events.

Demografische gegevens stemgerechtigden versus twitteraars

Uit verschillende onderzoeken blijkt dat er significante verschillen te vinden zijn in het stemgedrag tussen verschillende groepen mensen. Naar voren komt dat het stemgedrag uiteenloopt, gekeken naar demografische gegevens, tussen bijvoorbeeld mannen en vrouwen, en jongeren en ouderen. Zo blijkt dat Vlaamse mannen en vrouwen anders kiezen, wat tevens geldt voor stemmers in verschillende leeftijdscategorieën (Deschouwer, Delwit, Hooghe & Walgrave, 2010). Inglehart en Norris (2000) deden een wereldwijd onderzoek waarbij ze het

verschil tussen mannen en vrouwen in negentien landen vergeleken gedurende de jaren '90. Zij onderzochten of er een verschil is tussen de politieke oriëntatie in geslacht, dus meer links of meer rechts georiënteerd. Het blijkt dat vrouwen en mannen wel degelijk verschillend georiënteerd zijn van elkaar. Daarnaast blijken mannen meer politieke interesse te vertonen dan vrouwen (Mestre & Marín, 2012). Zo variëren de politieke interesse en het stemgedrag in geslacht, maar het stemgedrag blijkt ook uiteen te lopen tussen verschillende leeftijden. In Groot-Brittannië blijkt bijvoorbeeld dat men steeds conservatiever gaat stemmen naarmate men ouder wordt (Tilley & Evans, 2014). Tevens is men meer geneigd om te stemmen naarmate men ouder wordt. Die verklaring is te vinden in een sociaal perspectief waarin de sociale norm rondom stemmen een belangrijke rol speelt. Ook de levensfase waarin ouderen zich bevinden draagt hieraan bij, waarbij gedacht moet worden aan bijvoorbeeld gezondheidskwesties (Goerres, 2007).

De grootste groep Nederlandse twitteraars zou zich bevinden in de leeftijdscategorie van 15 tot en met 19 jaar (50% daarvan gebruikte Twitter in 2014) en daaropvolgend neemt het Twittergebruik per volgende leeftijdscategorie af. In de leeftijdscategorie van 65 tot en met 79 jaar blijkt bijvoorbeeld maar 8% gebruik te maken van Twitter (Boekee, Engels & Van der Veer, 2014). Daarnaast verstuurden Nederlandse mannen in 2015 iets meer berichten ten opzichte van vrouwen, maar dit verschil was slechts 2% (Dietz, 2015). Hoewel Twitter uit *naturally occurring* data bestaat, is het, buiten steekproefsgewijze mediaonderzoeken, onder andere een beperking dat de demografische gegevens van de twitteraars normaliter niet bekend zijn (Sloan, Morgan, Burnap & Williams, 2015). Naast dat sommige onderzoekers hierom sceptisch zijn, is een andere beperking dat niet iedereen Twitter gebruikt. Zo'n 15% van de Nederlandse bevolking maakt gebruik van Twitter (RTL Nieuws, 2016; CBS, 2016). "Het is dus een deelverzameling en waarschijnlijk een beperkte", stelt Gayo-Avello (2011, p. 3). Daarom is de vraag in hoeverre Twitter te gebruiken is als representatie van de populatie (Mislove et al., 2011).

Deze vraag stelde Gayo-Avello (2013) ook. Hij nam negen onderzoeken met betrekking tot het voorspellen van verkiezingsuitslagen onder de loep die tussen 2010 en 2012 zijn uitgevoerd. Hij stelde dat Twitter als medium voor voorspellingen wordt overschat en dat de onderzoeken die hij analyseerde, dit bevestigden. Eén van de kritiepunten die hij had bij deze geanalyseerde onderzoeken, net zoals in zijn artikel een jaar daarvoor (Gayo-Avello, 2012), is dat er geen rekening is gehouden met de demografische bias. Daarom suggereert hij dat het een uitdaging en tevens van belang is dat de demografische gegevens van twitteraars worden meegenomen wanneer er verder onderzoek gedaan wordt naar de voorspelbaarheid

van Twitterdata bij verkiezingsuitslagen. Gayo-Avello geeft aan dat deze kwestie aangepakt kan worden. Dit kan ten eerste logischerwijs door het vergaren van de demografische gegevens van twitteraars. Ten tweede kan dit probleem worden tegengegaan door voor iedere categorie de tweets te herwegen gebaseerd op de verhoudingen in het stemgedrag van de populatie bij eerdere verkiezingen. Dit laatste is in overeenstemming met wat Schoen et al. (2013) ook beschrijven in hun artikel betreffende de biases, maar tevens de voorspellende kracht van social media. Ook zij stellen dat door de toepassing van een weging in tweets de Twitterdata meer valide is. Deze weging is toe te passen door het verschil tussen de daadwerkelijke stemverhouding, in het geval van dit onderzoek, en de stemverhouding volgens Twitter te reduceren. Dit kan middels *weighting* als er sprake is van één onafhankelijke variabele, zoals alleen geslacht (Culotta, 2014). Het wordt iets ingewikkelder wanneer er meerdere onafhankelijke variabelen, zoals zowel geslacht als leeftijd, in het spel zijn. Dan wordt *post-stratification* toegepast (Gelman, 2007). Daarnaast moet toekomstig soortgelijk onderzoek controleerbaarheid en reproduceerbaarheid handhaven (Huberty, 2015).

Tjong Kim Sang en Bos (2012) poogden hierin al stappen te maken met hun studie die gebaseerd is op het werk van Tumasjan et al. (2010). Tijdens de Eerste Kamerverkiezingen in 2011 telden Tjong Kim Sang en Bos tweets, zoals Tumasjan et al. ook al deden. Tjong Kim Sang en Bos plaatsten hier echter een aantal kanttekeningen bij. Zo heeft iedere stemgerechtigde maar één stem tijdens de verkiezingen, terwijl meerdere tweets per persoon verstuurd kunnen worden. Daarnaast is het mogelijk dat een tweet negatief is. Daarom behielden ze iedere keer maar één tweet per persoon in de dataset en verwijderden ze negatieve tweets. Echter, de sentimentanalyse zorgde er niet voor dat de voorspelling middels tweets meer overeenkwam met de voorspelling door de polls. Ze gaven aan dat dit te wijten zou kunnen zijn aan de ontbrekende demografische gegevens van de twitteraars. Hierdoor was onbekend of de twitteraars een afspiegeling van de stemgerechtigden vormden. Aangezien zij niet de beschikking hadden over data om de demografische gegevens te filteren, verkleinden ze de bias door correcties in het gewicht van hun voorspelling aan te brengen. Deze gewichten waren gebaseerd op de polls, waardoor het verschil tussen de pollvoorspellingen en hun voorspelling van het aantal zetels per partij kleiner werd. Zij plaatsten hier zelf als kritische noot bij dat zij hiermee in plaats van de verkiezingsuitslag meer neigden de polluitslag te voorspellen. Het bleek wel dat het aantal zetels dat uiteindelijk voorspeld werd middels tweets, dicht bij de daadwerkelijke verkiezingsresultaten lagen.

Tjong Kim Sang en Bos deden dus al een poging om de bias van de demografische gegevens van twitteraars te corrigeren. Met inzicht in de demografische gegevens van de

twitteraars kan de bias van de ongecontroleerde data verminderd worden (Culotta et al., 2015). De vraag is alleen: wat verstaan we onder de demografische gegevens van de Twittergebruikers? Want wie tweeten er en zijn diegenen representatief voor de stemgerechtigden tijdens de Tweede Kamer- en Provinciale Statenverkiezingen? In dit onderzoek wordt dit beperkt tot twee demografische gegevens: geslacht en leeftijd.

Detectie demografische gegevens twitteraars

In verschillende studies is getracht deze demografische gegevens, ‘geslacht’ en ‘leeftijd’, vast te stellen middels automatische profilering, wat in meerdere en in mindere mate lukte. Aan de hand van automatische profilering kunnen tweets met trefwoorden gefilterd en vervolgens ook geclassificeerd worden. Mislove et al. (2011) en Beretta, Maccagnola, Cribbin en Messani (2015) toonden aan dat het demografisch gegeven geslacht, in tweets bijvoorbeeld te herleiden is, gebaseerd op namen. Middels namenlijsten trachtten ze het geslacht van twitteraars te detecteren. Naast geslacht in tweets, hebben verscheidene onderzoekers ook gepoogd de leeftijd van twitteraars te detecteren. Onderzoek van Sloan et al. (2015), naar de leeftijd van Britse twitteraars, laat zien dat de gemiddelde leeftijd van twitteraars lager ligt dan die van de bevolking. Zij hebben leeftijd uit tweets geprobeerd te halen door woorden op te stellen die geassocieerd worden met leeftijd, zoals “verjaardag” en “jaar”. Het blijkt dat twee derde van de Twittergebruikers onder de 35 jaar is. Een belangrijke mogelijke verklaring die zij daarbij stellen, is dat ouderen minder vaak hun leeftijd bekendmaken op Twitter. Beretta et al. (2015) poogden naast geslacht, ook leeftijd te detecteren. Aan de hand van het trefwoord “birthday” haalden zij tweets op, die zij vervolgens handmatig indeelden op jongere mannen, jongere vrouwen, oudere mannen en oudere vrouwen, waarbij de grens tussen jong en oud was gesteld op 30 jaar. De detectie van geslacht gaat sneller dan de detectie van leeftijd en ook de nauwkeurigheid van de voorspelling van het geslacht bleek beter dan leeftijd (Beretta et al., 2015). Dit beaamt Mislove et al. (2011); voornamelijk voor de detectie van geslacht blijken er effectieve manieren te zijn.

Eerder onderzoek toont aan dat machine learning ook gebruikt kan worden om demografische gegevens van twitteraars automatisch vast te stellen (Nguyen et al., 2013a; Rao & Yarowsky, 2010). Hoe gaat deze methode in zijn werk? Hiervoor zijn eerst geannoteerde data, ofwel voorbeelden, nodig, om de computer deze gegevens te leren detecteren. Met machine learning kan middels gelabelde, ofwel geannoteerde, voorbeelden de computer geleerd worden teksten te classificeren (Nigam, McCallum, Thrun, & Mitchell, 1998). Dus bijvoorbeeld ook tweets. Allereerst wordt gelabelde data ingevoerd. Deze data

bestaat uit een trainingset en een testset. Met de trainingset wordt de computer labels geleerd en leert het dus de classificatie. Op basis van features leert de computer die labels. Dit gaat met behulp van tekstuele eigenschappen; het gebruik van individuele woorden of combinaties van woorden. Aan de hand van de testset wordt de prestatie van de computer geanalyseerd in vergelijking met de data waarvan bekend is wat de classificatie is. Als deze classificatie goed bevonden wordt, kan de classificatie toegepast worden op de gehele dataset (Frankwatching, 2014). Het idee is dat de computer leert van “ervaring” (Kirranne, 1990). Oftewel, hoe meer voorbeelden ingegeven worden, des te meer ervaring de computer heeft en hoe beter de computer de tweets, in het geval van dit onderzoek, bijvoorbeeld kan classificeren.

In een onderzoek naar in hoeverre met behulp van machine learning gebruikers geclassificeerd konden worden op onder andere geslacht, is tekstanalyse en fotoanalyse gecombineerd. Het bleek dat de classificatie met tekstanalyse de hoogste nauwkeurigheid behaalde (Sakaki, Miura, Ma, Hattori & Ohkuma, 2014). Eerder pasten Rao en Yarowsky (2010) al binaire classificatie toe op onder andere geslacht en leeftijd (waarbij ze een grens aanhielden van 30 jaar), waarin onder meer naar voren kwam dat vrouwen vaker emoticons gebruiken dan mannen. Nguyen et al. (2013a) gebruikten machine learning om leeftijd te classificeren middels taalgebruik. Zij ontwikkelden TweetGenie in 2013, een systeem dat geslacht en leeftijd kan schatten van Nederlandse twitteraars. Allereerst verzamelden ze Nederlandse Twitteraccounts. Voor hun corpus hebben ze twee annotatoren ingezet die in totaal 3185 accounts annoteerden op leeftijd en geslacht (Nguyen, Gravel, Trieschnigg & Meder, 2013b). Nguyen et al. gebruikten drie manieren om leeftijd te annoteren: een indeling op basis van leeftijdscategorie (20-, 20-40 en 40+), exacte leeftijd en levensfase. Automatische classificatie hebben zij vervolgens toegepast om geslacht (man en vrouw) en leeftijd (de drie aspecten) te labelen. Dit deden ze met logistische en lineaire regressie waarbij een cut-off van 200 tweets per gebruiker werd ingesteld. De automatische classificatie werd vergeleken met de handmatige annotaties. Nguyen et al. (2013b) stippen aan dat de classificatie van leeftijd door TweetGenie preciezer is dan de classificatie door annotatoren.

Daarnaast beschrijven Nguyen et al. (2013a) onder andere dat jongeren meer informele taal gebruiken zoals “haha” en “xd”. Oudere personen gebruikten daarentegen formeler taalgebruik. Kenmerkende woorden voor jongere personen waren onder andere: “school”, “ik”, “:)” en “stages”. Voor oudere personen waren dit bijvoorbeeld: “verdomd”, “dochter”, “wens”, “zoon” en “geniet”. Ze stellen dat vanaf 35 jaar en ouder het voorspellen van de leeftijd van twitteraars steeds lastiger wordt. Dit is wat Beretta et al. (2015) ook ondervonden. De verschillen tussen leeftijdsgroepen worden dan steeds kleiner. De verklaring

die Nguyen et al. (2013b) hiervoor geven, is dat de ontwikkelingen op taalkundig gebied voornamelijk optreden bij jongeren; naarmate mensen ouder worden, wordt dit steeds minder en lijken de verschillen in taal kleiner te worden. Daarnaast trachtten Nguyen et al. (2013b) geslacht te detecteren. Verschillen die zij vonden tussen mannen en vrouwen waren bijvoorbeeld dat mannen meer hyperlinks en intensiverders zoals “echt” gebruiken in hun tweets dan vrouwen en dat vrouwen meer de eerste persoon enkelvoud, zoals “ik” en “mijn” gebruiken. Daas, Burger, Le, Ten Bosch en Puts (2016) gebruikten het systeem van Nguyen et al. (2013a) al voor de classificatie van Nederlandse twitteraars waarmee ze leeftijd trachtten vast te stellen middels schrijfstijl in tweets.

Overview

Het is dus een beperking dat demografische gegevens van twitteraars niet bekend zijn. Om een gedegen voorspelling van een verkiezingsuitslag te vormen, is het nodig om de demografische gegevens inzichtelijk te maken; dan kan de voorspellingen mogelijk ook nauwkeuriger maken (Sanders & Van den Bosch, 2013; Mislove et al., 2011). TweetGenie zal gebruikt worden om deze gegevens te genereren. Met die gegevens kan ook het Twittergedrag tussen verschillende demografische dimensies geanalyseerd worden, want hoe tweeten gebruikers in verkiezingstijd? Een onderliggend doel is om voor toekomstig onderzoek handvatten te bieden om in het vervolg niet alleen achteraf een voorspelling te kunnen doen, maar om ook een daadwerkelijke voorspelling te kunnen doen (Schoen et al. 2013). Dit mondt uit in de volgende onderzoeksvraag: “In hoeverre kunnen demografische gegevens, ‘geslacht’ en ‘leeftijd’, van twitteraars met behulp van TweetGenie de schattingen van de verkiezingsuitslagen, aan de hand van tweets, van de Provinciale Staten en de Tweede Kamer nauwkeuriger maken?” Daarbij zijn drie deelvragen geformuleerd, ten eerste: “In hoeverre verhouden de dimensies in stemgedrag op Twitter, volgens TweetGenie, gekeken naar de demografische gegevens, ‘geslacht’ en ‘leeftijd’, zich tot elkaar ten opzichte van de daadwerkelijke verhoudingen in stemgedrag?”, ten tweede: “In welke mate kan de verkiezingsuitslag nauwkeurig voorspeld worden wanneer de demografische gegevens van de twitteraars bekend zijn in vergelijking met de opiniepeilingen en wanneer de demografische gegevens van de twitteraars niet bekend zijn?” en ten derde: “In hoeverre zijn er verschillen binnen de demografische dimensies in politiek Twittergedrag?”

Methode

Deze methodesectie is tweeledig. Allereerst werd aan de hand van een corpusanalyse handmatig data geannoteerd en gebruikt voor een vergelijking tussen menselijke annotatoren en TweetGenie. Een set politieke Twitteraccounts, een tweet met daarin de benoeming van een politieke partij, is handmatig geannoteerd op de demografische gegevens geslacht en leeftijd. Hierna classificeerde TweetGenie dezelfde set op geslacht en de leeftijd. Aan de hand van deze gegevens kon bepaald worden hoe nauwkeurig TweetGenie ten opzichte van menselijke annotatoren Twitteraccounts classificeert en op welke punten de annotaties van elkaar verschillen. Dit eerste deel van de methodesectie is aangegeven met: *Deel I*.

Ten tweede is TweetGenie ingezet voor een grotere dataset om de onderzoeksvragen te beantwoorden. TweetGenie heeft die dataset met politieke twitteraars geclassificeerd op geslacht en leeftijd. Zo konden de verhoudingen van de demografische gegevens op Twitter vergeleken worden met de daadwerkelijke verhoudingen. Met dezelfde dataset werden louter de frequenties van benoemingen van politieke partijen in tweets per jaar geteld voor een voorspelling van de verkiezingsuitslagen. Daarna werden de demografische gegevens van de politieke twitteraars, vergaard middels TweetGenie, aangepast aan de daadwerkelijke demografische verhoudingen van de stemmers. De nauwkeurigheid van de verkiezingsuitslagen volgens Twitter met en zonder aanpassing werd getoetst door deze af te zetten tegen de nauwkeurigheid van de opiniepeilingen en de daadwerkelijke verkiezingsuitslagen per jaar. Ten slotte zijn de verschillen in Twittergedrag binnen de demografische dimensies per jaar geanalyseerd. Met *Deel II* wordt dit deel van deze methodesectie aangeduid.

Deel I - Materiaal vergelijking handmatige classificatie en TweetGenie classificatie

Om te beginnen werd een corpus aangelegd van Nederlandse Twitterprofielen waarin een politieke tweet voorkwam. Het corpus werd aangelegd met behulp van Twiqs.nl (Tjong Kim Sang & Van den Bosch, 2013). De tweets zijn verzameld aan de hand van de Twitter API. Daaruit zijn Nederlandse tweets gefilterd middels veelgebruikte Nederlandse woorden (229) die niet veelvoorkomend zijn in andere talen, zoals “het”, “nooit” en “zijn”. Deze data was reeds verzameld en bevatte alle tweets met politieke partijen. De tweets zijn afkomstig uit de perioden: tien dagen voor de Tweede Kamerverkiezingen 2012 en tien dagen voor de Provinciale Statenverkiezingen 2015. De accounts van de politieke twitteraars zijn verzameld door te zoeken op naam van de politieke partijen die met één of meerdere zetels in de Tweede Kamer vertegenwoordigd zijn. Hierbij werd een cut-off van honderd tot tweehonderd tweets

ingesteld. Op die manier werden niet alleen Twitteraccounts met slechts enkele tweets buiten beschouwing gelaten, maar werd ook getracht, door een maximum van tweehonderd in te stellen, te voorkomen dat automatische Twitteraccounts in de data terechtkwamen. De tweets werden middels de reguliere expressies van partijnamen opgehaald (zie tabel 1).

Tabel 1. De reguliere expressies van de partijnamen zoals opgesteld door Sanders en Van den Bosch (2013)

Partijnaam	Reguliere expressies partijnamen
VVD	“vvd”
PVDA	"pvda", "partij\s+v(oor\s+ an\s+ .)?d(e .)?\s+arbeid"
SP	“sp”
PVV	"pvv", "partij\s+v(oor\s+ an\s+ .)?d(e .)?\s+vrijheid"
CDA	“cda”
D66	"d\'?66"
GroenLinks	"gl", "groen.?links"
CU	"cu", "christen.?unie"
SGP	“sgp”
PVDD	"pvdd", "partij\s+v(oor\s+ an\s+ .)?d(e .)?\s+dieren"
50PLUS	"50[^\d]?(\+ plus)"

De annotaties voor dit onderzoek zijn door twee annotatoren uitgevoerd in een *tailor made* webtool: een webtool die speciaal voor dit onderzoek werd ingericht. Een randomisatie in de webtool zorgde ervoor dat allereerst honderd dezelfde willekeurige Twitteraccounts aan de twee annotatoren afzonderlijk, werden getoond. Deze zijn gebruikt voor een vergelijking in nauwkeurigheid. Hierna zijn 2800 willekeurige Twitteraccounts aan de twee annotatoren getoond, die ieder een gedeelte annoteerde. De grotere dataset werd gebruikt voor een meer diepgaande vergelijking in de classificaties. Dong Nguyen heeft haar systeem TweetGenie beschikbaar gesteld om diezelfde, vooraf geselecteerde Twitterdata te classificeren op geslacht en leeftijd.

Deel I - Procedure handmatige classificatie en TweetGenie classificatie

De data werd door twee annotatoren geannoteerd op het demografische gegeven geslacht in twee categorieën, te weten: man en vrouw. Het geslacht is middels namen van twitteraars herleid. In het codeerveld in de annotatietool was ook een link beschikbaar naar het profiel.

Dit was vooral van belang bij ambigue namen, zoals Robin of Jamie, die voor mannen en vrouwen gebruikt kunnen worden. Vaak kon aan de hand van de profielfoto van de twitteraar herleid worden of het een man of vrouw betrof. Wanneer het geslacht niet te herleiden was, bleven de codeervelden blanco.

Daarnaast zijn dezelfde Twitteraccounts handmatig op leeftijd gelabeld. De exacte leeftijd van de Twitteraccounts is door de annotatoren ingevoerd. Soms was de leeftijd zichtbaar, maar het kwam ook voor dat twitteraars hun geboortjaar of de gehele geboortedatum vermeldden, zoals: “Ik ben X jaar oud” en “Geboren in X”. Hierbij kon X een nummer zijn of een specifieke geboortedatum (Sloan et al. 2015). Een voorbeeld van een inforegel met daarin de leeftijd, is: “Lars | 22 | Napoli | Ajax”. Dit account is geannoteerd als een man van 22 jaar. Ook voor het inschatten van de leeftijd, indien deze ontbrak in de inforegel, was de link naar het profiel in het codeerveld nuttig. Met behulp van de profielfoto of tweets kon vaak een inschatting van de leeftijd gemaakt worden. De exacte leeftijd werd in het codeerveld ingevoerd in cijfers. Wanneer de leeftijd niet te herleiden was, bleef het veld oningevuld. De leeftijden werden bijna altijd van de profielfoto's afgeleid, omdat de leeftijden vaak niet werden vermeld in de inforegel.

TweetGenie classificeerde dezelfde Twitteraccounts van politieke twitteraars uit 2012 en 2015, op geslacht en leeftijd. Aan de hand van kenmerkende woorden (beschreven in de introductie), konden Twitteraccounts ingedeeld worden op geslacht (man of vrouw) en exacte leeftijd. De verdere werking van TweetGenie werd beschreven in de introductieparagraaf: *Detectie demografische gegevens twitteraars.*

TweetGenie kan alleen bestaande accounts classificeren en doet dit op basis van het taalgebruik van de politieke twitteraars. Om de analyse betreffende de overlap uit te kunnen voeren, werd de voorwaarde gesteld dat de Twitteraccounts door zowel de annotatoren als TweetGenie geannoteerd waren. Hierdoor bleven van de 100 Twitteraccounts 72 accounts over voor geslacht en 61 voor leeftijd om vast te kunnen stellen hoe nauwkeurig TweetGenie classificeert. Deze leeftijden zijn ingedeeld op leeftijdscategorieën: < 18, 18-34, 35-54 en 55 +, conform gegevens van TNS NIPO. Vervolgens hebben de twee annotatoren ieder een gedeelte van de 2800 Twitteraccounts handmatig geannoteerd om de demografische classificatie tussen beide annotatoren en TweetGenie te kunnen analyseren. Hierin werden ook de 100 Twitteraccounts voor de overlap meegenomen, waardoor het totale aantal Twitteraccounts dus op 2900 kwam. Annotator 1¹ annoteerde 1463 accounts en annotator 2²

¹ Annotator 1 betroffen de annotaties van Eline Pilaet

² Annotator 2 betroffen de annotaties van Michelle de Gier

1537. Ook zijn deze Twitteraccounts geclassificeerd door TweetGenie, zoals eerder beschreven. 1815 Twitteraccounts konden zowel door de annotatoren samen als TweetGenie geclassificeerd worden.

Deel I - Vergelijking handmatige classificatie met TweetGenie classificatie

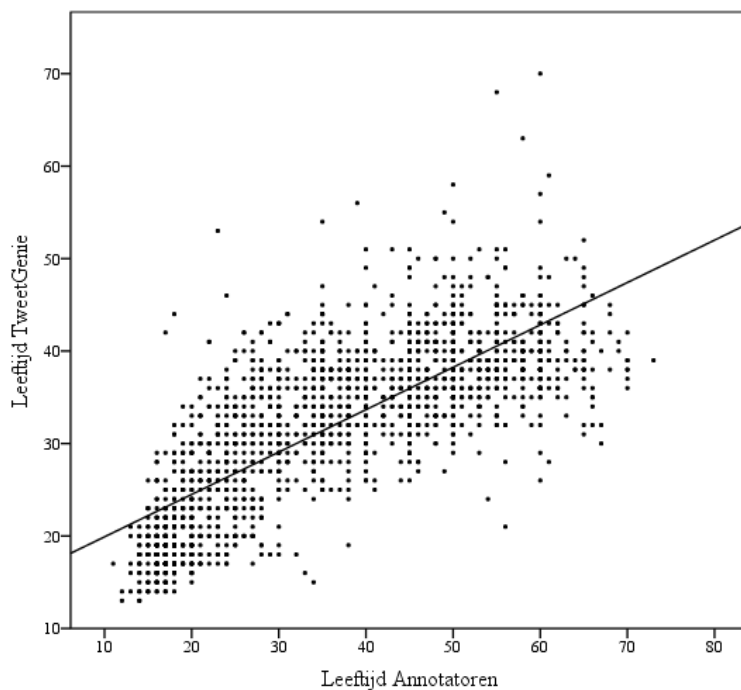
Honderd Twitteraccounts zijn dubbel geannoteerd om de interbeoordelaarsbetrouwbaarheid tussen de twee annotatoren en TweetGenie te bepalen. Hierbij zijn de annotaties van twee humane annotatoren vergeleken met TweetGenie. De interbeoordelaarsbetrouwbaarheid bij de annotatie van geslacht tussen annotator 1 en annotator 2 was bijna perfect ($\kappa = .97, p < .001$). Tussen annotator 1 en TweetGenie was deze redelijk ($\kappa = .52, p < .001$) en tussen annotator 2 en TweetGenie ook redelijk ($\kappa = .56, p < .001$). Ook is de interbeoordelaarsbetrouwbaarheid van de leeftijdscategorieën bepaald, waaruit bleek dat de annotatie van leeftijd tussen annotator 1 en annotator 2 voldoende tot goed was, ($\kappa = .64, p < .001$). De annotaties tussen annotator 1 en TweetGenie ($\kappa = .09, p < .001$) en tussen annotator 2 en TweetGenie ($\kappa = .14, p < .001$) waren beide gering. Uit deze vergelijkingen bleek dat de annotaties door mensen preciezer was dan de annotaties door TweetGenie.

Om de annotaties tussen beide annotatoren en TweetGenie te vergelijken op exacte leeftijd, is een grotere dataset gebruikt. Deze bestond uit 1815 Twitteraccounts die zowel door de annotatoren, ieder een deel, als TweetGenie op exacte leeftijd gelabeld waren. Wanneer gekeken werd naar de correlatie voor de Leeftijd volgens TweetGenie en de Leeftijd volgens de annotatoren bleek er een significant, positief verband te bestaan ($r(1815) = .77, p < .001$). Dit houdt in dat er een sterk effect is tussen de leeftijd die TweetGenie schat en de leeftijd die de annotatoren schatten voor de Twitteraccounts. Vervolgens laat een lineaire regressie zien dat de Leeftijd voor de Twitteraccounts voor 59% te verklaren was door de ingebrachte variabele ($F(1, 1813) = 2579,42, p < .001$). De Leeftijd geannoteerd door de annotatoren bleek een significante voorspeller voor de Leeftijd volgens TweetGenie ($\beta = .77, p < .001$) (zie tabel 2). Als de annotatoren de leeftijd annoteerden was deze hoger dan de leeftijd die TweetGenie aan een Twitteraccount toekende. Gekeken naar de descriptieve statistieken, laten deze tevens zien dat de annotatoren ($M = 36.14, SD = 15.03$) de leeftijden gemiddeld hoger annoteerden dan TweetGenie ($M = 31.90, SD = 8.99$), waarbij de annotatoren maximaal 73 jaar annoteerden en TweetGenie 70 jaar. Figuur 1 laat in een scatterplot de leeftijd volgens TweetGenie en de leeftijd volgens de annotatoren zien.

Tabel 2. Regressie-analyse voor de leeftijd volgens TweetGenie en de leeftijd volgens de annotatoren ($N = 1815$)

Variabele	B	$SE B$	β
Intercept	15.33	.35	
Age annotator	.77	.01	.77*
R^2	.59		
F	2579.42*		

* $p < .001$



Figuur 1. Scatterplot voor de leeftijd volgens TweetGenie en de leeftijd volgens de annotatoren

In tabel 3 zijn de percentuele verhoudingen in demografische gegevens van de handmatige annotaties en TweetGenie weergegeven. Dit zijn de classificaties die gedaan konden worden over de dataset van 2900 politieke twitteraars. De Twitteraccounts werden ingedeeld op jaar van het versturen van een politieke tweet: 2012 of 2015. Wanneer een twitteraar meerdere politieke tweets had verstuurd, werden deze allen apart in dataset opgenomen. Dit zorgde voor dubbele accounts die daarom per jaar zijn verwijderd. Er bleven 2286 accounts over voor 2012 en 1699 accounts voor 2015. Dit zijn opgeteld 3985 Twitteraccounts, omdat er 1085

twitteraars in beide jaren een politieke tweet verstuurd. De annotatoren wisten voor 2012 van 86% accounts het geslacht en van 73% de leeftijd vast te stellen en in 2015 van 83% het geslacht en van 68% de leeftijd. TweetGenie kon in 2012 van 87% het geslacht en de leeftijd vaststellen en in 2015 van 93% zowel geslacht als leeftijd. Te zien is dat de percentuele verhoudingen in de classificatie van geslacht enigszins hetzelfde lijkt te zijn tussen de handmatige annotaties en TweetGenie. De classificatie van leeftijden lijkt iets verder uiteen te lopen tussen de annotatoren en TweetGenie. Voornamelijk voor 2012 bij de 55-plus categorie en voor 2015 bij de 55-plus categorie, de categorie van 35 tot en met 54 jaar en de jongste categorie, zijn er verschillen te zien in de percentuele verhoudingen.

Tabel 3. Verhoudingen twitteraars naar demografische gegevens, ‘geslacht’ en ‘leeftijd’, per jaar in percentages volgens de resultaten van de handmatige annotaties en TweetGenie

Demografische gegevens	2012 handmatige annotaties	2012 TweetGenie	2015 handmatige annotaties	2015 TweetGenie
<i>Geslacht</i>				
Man	69.3	72.9	72.5	76.5
Vrouw	30.7	27.1	27.5	23.5
<i>Leeftijdscategorieën</i>				
< 18	15.6	18.2	23.3	0.3
18 - 34	38.1	40.0	21.6	32.6
35 - 54	33.7	41.0	38.9	65.1
55 +	12.6	0.8	16.3	2.0

Ten slotte is gekeken hoe de annotatoren en TweetGenie omgaan met niet-menselijke Twitteraccounts. De annotatoren hebben over een dataset van 2878 Twitteraccounts bij 13% aangegeven dat het ging om een niet-menselijk Twitteraccount. TweetGenie heeft 90% van deze niet-menselijke accounts geclassificeerd op geslacht en leeftijd. Daarvan werd 78% als man geclassificeerd en 72% werd ingedeeld in de leeftijdscategorie van 35 tot en met 54 jaar. Zie bijlage 1 voor een totaaloverzicht.

Deel II - Materiaal TweetGenie classificatie politieke tweets

Dong Nguyen heeft TweetGenie nogmaals beschikbaar gesteld om een grotere dataset met 71.410 politieke twitteraars te classificeren op geslacht en leeftijd. Het materiaal is verder op dezelfde manier vergaard als de data van TweetGenie, besproken in de paragraaf: *Materiaal handmatige classificatie en TweetGenie classificatie*, van deze methodesectie.

De daadwerkelijke verhoudingen van de stemmers tussen mannen en vrouwen en de leeftijdscategorieën zijn via TNS NIPO verkregen. Dezelfde leeftijdscategorieën zoals eerder besproken zijn aangehouden: < 18, 18 - 34, 35 - 54 en 55 +. De daadwerkelijke verhoudingen zijn vergeleken met de verhoudingen in de Twitterdata middels TweetGenie.

Voor een vergelijking tussen de nauwkeurigheid van opiniepeilingen en van de voorspelling van de verkiezingsuitslagen aan de hand van Twitteraccounts, waren de volgende gegevens benodigd: de opiniepeilingen, de verkiezingsuitslagen van zowel de Tweede Kamer- alsook de Provinciale Statenverkiezingen en het stemgedrag per partij per demografische dimensie. Conform het onderzoek van Sanders en Van den Bosch (2013) zijn de opiniepeilingen voor de Tweede Kamerverkiezingen in 2012 van de website allepolitiekepeilingen.nl verkregen. Ook de opiniepeilingen van de Provinciale Statenverkiezingen in 2015 werden op die manier vergaard. Voor de opiniepeiling van de Tweede Kamerverkiezingen in 2012 zijn de peilingen van TNS NIPO en De Stemming op 11 september 2012 gebruikt. Gegevens voor de Provinciale Statenverkiezingen in 2015 betroffen de peilingen van de Politieke Barometer, I&O Research en De Stemming op 17 maart 2015. Enkel de opiniepeilingen van één dag voor de verkiezingsdag zijn meegenomen in het onderzoek, waarvan de gemiddelden per jaar zijn berekend (zie bijlage 2). Tevens zijn de verkiezingsuitslagen in percentages van de Provinciale Statenverkiezingen van Kiesraad (z.d.c.) gebruikt, evenals de verkiezingsuitslagen van de Tweede Kamerverkiezingen (Kiesraad, 2012). Het stemgedrag per partij, onderverdeeld in geslacht en leeftijd op het moment van de Tweede Kamerverkiezingen in 2012 en de Provinciale Statenverkiezingen in 2015, werd via TNS NIPO vergaard.

Vervolgens was voor de vergelijking van het politieke Twittergedrag tussen de demografische dimensies per jaar het aantal genoemde politieke partijen per demografische dimensie benodigd. Deze gegevens werden vergaard uit de reeds verzamelde Twitterdata.

Deel II- Procedure TweetGenie classificatie politieke tweets

In totaal zijn 442.698 politieke tweets van 71.411 Twitteraccounts opgehaald. De Twitteraccounts werden weer ingedeeld op jaar van het versturen van een politieke tweet:

2012 of 2015. Wanneer een twitteraar meerdere politieke tweets had verstuurd, stonden deze allen in de dataset. Daarom zijn de dubbele accounts per jaar verwijderd, waardoor ieder account per jaar maar één keer werd meegenomen. Voor 2012 bleven er 48.614 accounts over en voor 2015 31.458. Doordat 8.661 twitteraars in beide jaren politiek getwitterd hebben, werden dit samen 80.072 Twitteraccounts. De classificaties door TweetGenie werden gebruikt om de demografische gegevens uit te draaien. De accounts werden ingedeeld op geslacht (man of vrouw). TweetGenie baseert de leeftijden op de actuele tweets en het huidige onderzoek is uitgevoerd in 2016. Voordat de accounts ook werden ingedeeld op leeftijdscategorie, is daarom vier jaar van de leeftijden van 2012 afgehaald en één jaar van 2015. De verschillen in de jaren zijn daarmee opgeheven. Deze classificaties werden gebruikt om te zien hoe de demografische gegevens van de daadwerkelijke stemmers en die van politieke twitteraars zich tot elkaar verhouden, ter beantwoording van deelvraag 1.

Om vervolgens antwoord te kunnen geven op de vraag of met de kennis van de demografische gegevens van de twitteraars de voorspellingen van de verkiezingsuitslagen nauwkeuriger gemaakt kunnen worden (deelvraag 2), zijn politieke tweets gebruikt om de frequenties van de politieke partijen per jaar te berekenen. Wanneer een twitteraar meerdere politieke tweets had verstuurd, werden al deze tweets in dit geval wel meegenomen in de data. Dit resulteerde in 204.914 politieke tweets in 2012 en 237.784 in 2015. Deze frequenties zijn afgezet tegen het totale aantal tweets waarin partijen voorkomen, die resulteerden in percentages in verhouding tot het totaal. De percentages per partij zijn als volgt berekend: het aantal tweets van een specifieke partij werd gedeeld door de som van alle tweets met politieke partijen, vermenigvuldigd met honderd. Dit werd voor iedere partij herhaald.

Hierna is de data van enkel de frequenties aangepast aan de daadwerkelijke demografische verhoudingen van de kiezerspopulatie. Dit is gedaan door een gewicht in de berekening toe te wijzen aan het geslacht en de leeftijdscategorieën. De verdeling van het stemgedrag per partij per demografische dimensie is vergeleken met de demografische verhoudingen van de politieke twitteraars. Hiervoor is *post-stratification* toegepast (Gelman, 2007). In de berekening heeft een correctie plaatsgevonden tussen de daadwerkelijke percentuele verdeling van mannen en vrouwen opgedeeld in de drie leeftijdscategorieën in 2012 en 2015 en de percentuele verdeling van mannen en vrouwen in diezelfde leeftijdscategorieën volgens de Twitterdata geïdentificeerd door TweetGenie op datzelfde moment. De niet-stemgerechtigden, iedereen jonger dan 18 jaar, werden verwijderd.

In tabel 4 zijn de wegingen van de demografische dimensies weergegeven. Dit wordt toegelicht met een voorbeeld: het daadwerkelijke percentage van mannen tussen de 18 en 34

jaar in 2012 is 6.08. Uit de Twitterdata geïnclassificeerd door TweetGenie blijkt dat voor dezelfde groep het percentage 51.25 is. De weging van de Twitterdata is aangepast aan dat verschil. Dit houdt in dat de tweets van mannen tussen de 18 en 34 jaar in 2012 vermenigvuldigd werden met 0.12 (6.08/51.25). Zo werd dit voor de andere categorieën ook uitgevoerd. Vervolgens zijn per demografische dimensie en per politieke partij de wegingen doorgevoerd, waarna per politieke partij de som van de demografische dimensies is genomen om tot het totale aantal tweets per partij te komen. Ten slotte zijn de percentages per partij na aanpassing op dezelfde manier berekend zoals hierboven beschreven voor louter de tellingen van de politieke partijen in tweets.

Tabel 4. De absolute aantallen en de procentuele verhoudingen per demografische dimensie, ingedeeld op geslacht en leeftijd met de gewichten

Geslacht en leeftijdscategorie	Populatie <i>N</i>	Percentage van de populatie	Sample <i>n</i>	Percentage van de sample	Gewicht
<i>Tweede Kamerverkiezingen 2012</i>					
Man 18 - 34	259860	6.08	21453	51.25	0.12
Man 35 - 54	669085	15.66	8934	21.34	0.73
Man 55 +	1060317	24.82	397	0.95	26.17
Vrouw 18 - 34	239083	5.60	7373	17.61	0.32
Vrouw 35 - 54	747601	17.50	3545	8.47	2.07
Vrouw 55 +	1295616	30.33	157	0.38	80.87
Totaal	4271562		41859		
<i>Provinciale Statenverkiezingen 2015</i>					
Man 18 - 34	179902	6.23	8969	28.71	0.22
Man 35 - 54	392716	13.61	14219	45.51	0.30
Man 55 +	774790	26.84	439	1.41	19.10
Vrouw 18 - 34	186752	6.47	2542	8.14	0.80
Vrouw 35 - 54	440971	15.28	4930	15.78	0.97
Vrouw 55 +	911176	31.57	144	0.46	68.49
Totaal	2886307		31243		

Getracht is een standaard te hanteren die een correcte weergave vormt van de voorspelling middels Twitterdata geïnclassificeerd door TweetGenie. De voorspelling van de verkiezingsuitslag met de aanpassing is vergeleken met de voorspelling van de

verkiezingsuitslag zonder deze aanpassing, met de daadwerkelijke verkiezingsuitslag en met de opiniepeilingen.

Tot slot zijn, in het kader van deelvraag 3, de verschillen in het politieke Twitergedrag binnen de verschillende demografische dimensies met elkaar vergeleken. Hiervoor is dezelfde data gebruikt als voor enkel het tellen van de politieke partijen. Twitteraars konden namelijk meerdere politieke partijen benoemen in hun tweets; daar is geen onderscheid in gemaakt. Deze data is daarom ingedeeld naar de categorieën in het geslacht en de leeftijd per partij.

Deel II - Statistische toetsing

Om antwoord te kunnen geven op deelvraag 1, zijn de percentuele daadwerkelijke stemverhoudingen en de verhoudingen vanuit Twitter na classificatie door TweetGenie vergeleken met elkaar in de vorm van descriptieve statistieken. Daarnaast zijn, voor deelvraag 2, de uitslagen van beide verkiezingen met elkaar vergeleken. Hiervoor zijn de daadwerkelijke uitslagen en de uitslagen volgens Twitter met en zonder aanpassing via TweetGenie naast elkaar gelegd. Om te bepalen hoe nauwkeurig de verkiezingsuitslag voorspeld werd, werd de Mean Absolute Error (MAE) berekend tussen het voorspelde aantal zetels per partij en het daadwerkelijk behaalde aantal zetels. Door het verschil tussen de MAE te berekenen van de voorspelling van de verkiezingsuitslagen van de Tweede Kamerverkiezingen in 2012 en de Provinciale Statenverkiezingen in 2015 met en zonder aanpassing op demografie, kon de effectiviteit van deze methode bepaald worden. Om te analyseren of de verkiezingsuitslag volgens tweets waarvan de demografische gegevens van de schrijver bekend zijn, nauwkeuriger is dan de verkiezingsuitslag volgens tweets waarbij deze niet bekend zijn en deze nauwkeuriger is dan de opiniepeilingen, zijn de correlaties van deze variabelen ten opzichte van de daadwerkelijke verkiezingsuitslag berekend. Hiervoor werd de Pearson correlatie gehanteerd die met behulp van SPSS 22 is berekend, om zo ook de significantie te kunnen bepalen. De samenhang tussen de variabelen is met de correlatie aangeduid in effectsterkte, ofwel de correlatiecoëfficiënt. Ten slotte is een Chi-kwadraat gebruikt om te toetsen of er significante verschillen zijn tussen de nominale variabelen, in het kader van deelvraag 3, te weten: de politieke partijen in tweets en de verschillende demografische dimensies. Met de gestandaardiseerde residuen is aangetoond waar de significante verschillen precies voorkomen.

Resultaten

Verhoudingen demografische gegevens politieke verkiezingen

In tabel 5 zijn de verhoudingen in demografische dimensies te zien van de populatie die gestemd heeft tijdens de Tweede Kamerverkiezingen in 2012 en de Provinciale Statenverkiezingen in 2015. De verhoudingen vanuit gegevens van TNS NIPO die de daadwerkelijke stemverhoudingen aantonen, maar ook volgens de resultaten die verkregen zijn vanuit de Twitterdata, zijn weergegeven. De Twitterdata bevatten de verhoudingen vanuit TweetGenie. Hierin is zichtbaar dat de verschillen in verhoudingen tussen mannen en vrouwen zowel in 2012 en 2015 groot zijn tussen de daadwerkelijke verhouding en de verhouding die TweetGenie heeft geclassificeerd. Gekeken naar de daadwerkelijke verhoudingen in geslacht, was de verdeling tussen mannen (46.7%) en vrouwen (53.3%) in 2015 redelijk gelijk. Dit terwijl de data vanuit TweetGenie laat zien dat mannen aanzienlijk meer (75.6%) hebben getweet over politieke partijen dan vrouwen (24.4%) in 2015. Deze verschillen zijn ook te zien bij 2012. Daarnaast kwam naar voren dat de leeftijdscategorie van 18 tot en met 34 jaar ondervertegenwoordigd was bij de daadwerkelijke stemmingen, maar dat in 2012 de grootste groep politieke twitteraars juist te vinden was in die leeftijdscategorie. In zowel 2012 als 2015 waren de politieke twitteraars in de 55-plus categorie ondervertegenwoordigd. Ten slotte is te zien dat, in tegenstelling tot 2012, de meeste politieke twitteraars zich in 2015 in de leeftijdscategorie van 35 tot en met 54 jaar bevonden.

Tabel 5. Verhoudingen stemmers Tweede Kamer- en Provinciale Statenverkiezingen naar demografische gegevens, ‘geslacht’ en ‘leeftijd’, in percentages volgens de resultaten van TNS NIPO en TweetGenie

Demografische gegevens	2012		2015	
	(TNS NIPO)	2012 TweetGenie	(TNS NIPO)	2015 TweetGenie
<i>Geslacht</i>				
Man	46.6	73.5	46.7	75.6
Vrouw	53.4	26.5	53.3	24.4
<i>Leeftijdscategorieën</i>				
18 - 34	11.7	68.9	12.7	36.8
35 - 54	33.2	29.8	28.9	61.3
55 +	55.2	1.3	58.4	1.9

Een voorspelling van de verkiezingsuitslagen

In tabel 6 en tabel 7 zijn de daadwerkelijke uitslag en de uitslag volgens de opiniepeilingen te zien van zowel de Tweede Kamer- alsook de Provinciale Statenverkiezingen. Vervolgens is de uitslag van beide verkiezingen volgens tweets weergegeven. Allereerst bevat dit een uitslag aan de hand van de frequentie van de politieke partijen in tweets. Hierna is dit nogmaals weergegeven, maar dan met een aanpassing in demografie op geslacht en leeftijd die middels TweetGenie verkregen zijn.

Een correlatie voor de Verkiezingsuitslag in 2012 en de Opiniepeilingen op één dag voor de Tweede Kamerverkiezingen in 2012 liet een significant, positief verband zien ($r(11) = .98, p < .001$). De daadwerkelijke verkiezingsuitslag van de Tweede Kamer correleert sterk met de opiniepeilingen die gedaan zijn op de dag voor de verkiezing. Vervolgens bleek er uit een correlatie voor de Verkiezingsuitslag in 2012 en Twitter zonder aanpassing ook een significant, positief verband te bestaan ($r(11) = .97, p < .001$). Dit houdt in dat er een sterk effect is tussen de daadwerkelijke uitslag van de Tweede Kamerverkiezingen in 2012 en de uitslag volgens Twitter zonder aanpassing. Tevens geeft een correlatie voor de Verkiezingsuitslag 2012 en Twitter met aanpassing middels TweetGenie, een significant, positief verband weer ($r(11) = .97, p < .001$). Hieruit bleek dat de daadwerkelijke uitslag in 2012 en de uitslag volgens Twitter, nadat er aanpassingen in demografie op leeftijd en geslacht hebben plaatsgevonden, sterk met elkaar correleren. De betrouwbaarheidsintervallen (95%) zijn in de tabel tussen haakjes weergegeven. De MAE is lager bij de opiniepeilingen dan bij de verkiezingsuitslagen volgens Twitter met en zonder aanpassing. Dit houdt in dat de voorspelde uitslag volgens de opiniepeilingen minder afwijkt van de daadwerkelijke verkiezingsuitslag dan die van Twitter.

Tabel 6. De daadwerkelijke uitslag, de uitslag volgens de opiniepeilingen en de uitslag volgens Twitter met en zonder aanpassing op basis van demografie van de Tweede Kamerverkiezingen in 2012 in percentages

Partij	Verkiezing 12 sep. 2012	Opiniepeilingen 11 sep. 2012	Twitter 2 - 11 sep. 2012	TweetGenie demografie 2 - 11 sep. 2012
VVD	26.8	23.4	22.7	23.1
PVDA	25.1	22.7	20.1	21.9
SP	9.8	14.3	11.2	11.8
PVV	10.2	11.5	10.2	7.7
CDA	8.6	8.1	8.1	8.7
D66	8.1	7.9	9.0	7.5
GroenLinks	2.4	2.7	8.1	8.1
CU	3.2	4.1	3.6	4.6
SGP	2.1	1.8	2.6	1.9
PVDD	2.0	1.6	3.0	3.2
50PLUS	1.9	2.1	1.4	1.6
Correlatie Verkiezingen	-	0.98*	0.97*	0.97*
MAE Verkiezingen	-	1.3	1.8	1.9
Correlatie Opiniepeilingen	0.98*	-	0.96*	0.96*
MAE Opiniepeilingen	1.3	-	1.6	1.3

* $p < .001$

Gekeken naar de Provinciale Statenverkiezingen, bleek er uit een correlatie voor de Verkiezingsuitslag in 2015 en de Opiniepeilingen op één dag voor de Provinciale Statenverkiezingen een significant, positief verband te bestaan ($r(11) = .96, p < .001$). De daadwerkelijke verkiezingsuitslag van de Provinciale Staten correleert sterk met de opiniepeilingen die gedaan zijn op de dag voor de verkiezing. Een correlatie voor de Verkiezingsuitslag in 2015 en Twitter zonder aanpassing gaf ook een significant, positief verband weer ($r(11) = .84, p = .001$). Dit houdt in dat er een sterk effect is tussen de

daadwerkelijke uitslag van de Provinciale Statenverkiezingen en de uitslag volgens Twitter zonder aanpassing. Ook liet een correlatie voor de Verkiezingsuitslag 2015 en Twitter met aanpassing middels TweetGenie een significant, positief verband zien ($r(11) = .86, p = .001$). Hieruit blijkt dat er nogmaals een sterk effect is, nu tussen de daadwerkelijke uitslag in 2015 en de uitslag die aan de hand van Twitter voorspeld kon worden, na aanpassing in demografie middels TweetGenie. In de tabel zijn tussen haakjes de betrouwbaarheidsintervallen (95%) te zien. De voorspellingen van de opiniepeilingen zijn nauwkeuriger dan de andere twee voorspellingen, gekeken naar de MAE. De uitslag volgens Twitter met aanpassing op demografie is wel nauwkeuriger dan die van Twitter zonder de aanpassing.

Tabel 7. De daadwerkelijke uitslag, de uitslag volgens de opiniepeilingen en de uitslag volgens Twitter met en zonder aanpassing op basis van demografie van de Provinciale Statenverkiezingen in 2015 in percentages

Partij	Verkiezing 18 mrt. 2015	Opiniepeilingen 17 mrt. 2015	Twitter 8 - 17 mrt. 2015	TweetGenie demografie 8 - 17 mrt. 2015
VVD	16.7	16.7	26.1	25.3
PVDA	10.6	10.1	14.1	13.9
SP	12.2	13.4	7.0	8.8
PVV	12.3	15.5	11.2	8.8
CDA	15.4	12.4	9.9	11.6
D66	13.0	15.5	15.6	13.3
GroenLinks	5.6	4.1	5.7	6.1
CU	4.2	4.0	3.4	4.3
SGP	2.9	2.6	2.2	2.1
PVDD	3.6	2.9	3.3	3.6
50PLUS	3.5	2.9	1.2	2.3
Correlatie Verkiezingen	-	0.96*	0.84**	0.86**
MAE Verkiezingen	-	1.3	2.9	2.3
Correlatie Opiniepeilingen	0.96*	-	0.83***	0.81****
MAE Opiniepeilingen	1.3	-	2.9	2.7

* $p < .001$, ** $p < .002$, *** $p < .003$, **** $p < .004$

Verschillen in Twittergedrag over politieke partij per demografische dimensie

Tevens is voor beide verkiezingen bestudeerd of er verschillen zijn tussen de demografische dimensies in het politiek tweeten van de twitteraars.

Uit een χ^2 -toets tussen Geslacht en Politieke partij in 2012 bleek er een significant verband te bestaan ($\chi^2(10) = 2147.26$, $p < .001$). Vrouwen blijken een andere politieke voorkeur te hebben dan mannen. In tabel 8 zijn de gestandaardiseerde residuen weergegeven.

Wanneer het gestandaardiseerde residu groter is dan 1.96 of -1.96, houdt dit in dat het verschil tussen de verwachte en de geobserveerde waarde significant is. In tabel 8 is te zien dat mannen significant meer tweetten over de VVD, PVV en CDA dan vrouwen. Vrouwen tweetten significant meer over GroenLinks, CU, Partij voor de Dieren en 50Plus dan mannen. Bij de PVDA, SP, D66 en SGP zijn er geen significante verschillen in de tweets te zien.

Vervolgens is gekeken naar de verschillen tussen leeftijdscategorieën en daarbij bleek volgens een χ^2 -toets een significant verband te bestaan tussen de Leeftijdscategorieën en Politieke partij in 2012 ($\chi^2 (30) = 4665.11, p < .001$). Er zijn verschillen tussen de leeftijdscategorieën en de politieke partijen waarover zij tweetten. Wanneer ingezoomd wordt op de gestandaardiseerde residuen, is te zien dat 55-plussers significant meer tweetten over de PVDA dan mensen onder de 18 jaar en mensen tussen de 18 en 34 jaar en tussen de 35 en 54 jaar. De jongste categorie tweette significant meer over de SP, PVV, Partij voor de Dieren en 50Plus dan de andere drie categorieën, terwijl de leeftijdscategorie van 35 tot en met 54 jaar meer tweette over de CDA, D66, GroenLinks en CU en de twitteraars tussen de 18 tot en met 34 jaar meer over de VVD en SGP, ten opzichte van de overige categorieën.

Tabel 8. Twittergedrag opgedeeld in geslacht en leeftijdscategorieën per politieke partij in 2012 met de percentages van tweets uitgedrukt en de gestandaardiseerde residuen tussen haakjes

	VVD	PVDA	SP	PVV	CDA	D66	GL	CU	SGP	PVDD	50+
<i>Geslacht</i>											
Man	23.1 (4.0)	20.0 (-0.8)	11.2 (0.6)	10.7 (5.9)	8.5 (6.2)	9.1 (1.4)	7.6 (-7.5)	3.6 (-2.1)	2.6 (0.8)	2.3 (-16.9)	1.3 (-3.9)
Vrouw	20.9 (-7.6)	20.4 (1.5)	11.0 (-1.1)	8.5 (-11.4)	6.4 (-12)	8.6 (-2.7)	10.1 (14.5)	4.0 (4.0)	2.5 (-1.6)	5.8 (32.6)	1.8 (7.6)
<i>Leeftijdscategorieën</i>											
< 18	19.6 (-7.1)	17.6 (-6.0)	16.9 (18.6)	20.5 (35.5)	3.2 (-18.7)	5.7 (-11.9)	4.6 (-13.3)	3.8 (1.1)	1.9 (-4.8)	3.9 (5.7)	2.1 (6.8)
18-34	23.2 (4.2)	20.3 (1.4)	10.9 (-3.2)	10.9 (7.0)	6.6 (-17.1)	9.1 (1.3)	8.1 (0.6)	3.3 (-6.8)	2.9 (6.4)	3.2 (3.6)	1.5 (1.3)
35-54	22.2 (-2.7)	20.1 (0.0)	10.7 (-3.9)	7.5 (-22.7)	11.2 (29.8)	9.4 (4.0)	8.6 (5.2)	4.2 (7.3)	2.3 (-5.5)	2.6 (-6.9)	1.2 (-5.0)
55 +	23.7 (1.1)	22.7 (2.9)	12.6 (2.1)	7.4 (-4.5)	8.4 (0.6)	6.6 (-4.1)	6.7 (-2.5)	0 (6.0)	1.4 (-3.9)	2.7 (-1.0)	1.9 (2.3)

Dezelfde toetsing is gedaan voor de tweets tijdens de Provinciale Statenverkiezingen in 2015. Er bleek een significant verband te bestaan tussen Geslacht en Politieke partij in 2015 gebruikmakend van een χ^2 -toets ($\chi^2(10) = 5113.32, p < .001$). Zoals ook in tabel 9 te zien is, waren er verschillen tussen het Twittergedrag van mannen en vrouwen over politieke partijen. Zo is te zien dat mannen significant meer tweets verstuurd over de VVD, PVV en CDA. Vrouwen verstuurd significant vaker een tweet met daarin de politieke partijen: PVDA, D66, GroenLinks, CU, Partij voor de Dieren en 50Plus. Bij de SP en SGP waren geen significante verschillen te zien.

In nogmaals een χ^2 -toets kwam naar voren dat er een significant verband was tussen de Leeftijdscategorieën en Politieke partij in 2015 ($\chi^2(30) = 2882.23, p < .001$). De vier leeftijdscategorieën lijken verschillend Twittergedrag betreffende politieke partijen te vertonen. Gekeken naar de gestandaardiseerde residuen, komt naar voren dat twitteraars in de jongste leeftijdscategorie significant vaker tweetten over GroenLinks en de CU; de twitteraars tussen de 18 tot en met 34 jaar tweetten vaker over de VVD, PVV, SGP en Partij voor de

Dieren; twitteraars tussen 35 tot en met 54 jaar tweekten significant meer over de PVDA en CDA en de 55-plussers over de SP en 50Plus, telkens in vergelijking met de andere drie leeftijdscategorieën.

Tabel 9. Twittergedrag opgedeeld in geslacht en leeftijdscategorieën per politieke partij in 2015 met de percentages van tweets uitgedrukt en de gestandaardiseerde residuen tussen haakjes

	VVD	PVDA	SP	PVV	CDA	D66	GL	CU	SGP	PVDD	50+
<i>Geslacht</i>											
Man	26.9 (7.3)	14.1 (-3.3)	7.0 (-0.8)	12.3 (14.9)	10.1 (2.7)	15.5 (-1.4)	5.1 (-11.7)	3.2 (-2.6)	2.2 (-0.7)	2.5 (-21)	1.1 (-2.7)
Vrouw	22.1 (-15.9)	15.8 (7.2)	7.2 (1.8)	5.8 (-32.5)	9.0 (-5.8)	16.2 (3.0)	8.7 (25.5)	3.9 (5.7)	2.3 (1.5)	7.4 (45.8)	1.5 (6.0)
<i>Leeftijdscategorieën</i>											
< 18	13.9 (-4.8)	8.6 (-3.1)	8.6 (1.2)	17.1 (3.5)	8.1 (-1.2)	12.1 (-1.8)	9.6 (3.2)	11.6 (9.0)	1.3 (-1.3)	3.0 (-0.3)	6.3 (9.4)
18-34	26.5 (2.0)	13.8 (-4.3)	6.9 (-0.9)	14.6 (26.8)	6.8 (-26.4)	15.5 (-0.9)	5.5 (-2.8)	2.9 (-6.1)	2.4 (4.3)	4.0 (9.8)	1.1 (-0.9)
35-54	25.9 (-1.2)	14.7 (3.2)	7.0 (-0.6)	9.8 (-16.8)	11.2 (16.6)	15.8 (1.7)	5.8 (1.9)	3.5 (2.7)	2.1 (-2.8)	3.1 (-5.9)	1.1 (-1.8)
55 +	26.6 (0.6)	13.1 (-2.0)	10.7 (8.0)	7.4 (-6.5)	12.7 (5.1)	10.8 (-7.0)	5.1 (-1.4)	5.3 (6.0)	2.4 (0.9)	2.2 (-3.6)	3.7 (13.4)

Conclusie

Deze studie beoogde antwoord te geven op de hoofdvraag: “In hoeverre kunnen demografische gegevens, ‘geslacht’ en ‘leeftijd’, van twitteraars met behulp van TweetGenie de schattingen van de verkiezingsuitslagen, aan de hand van tweets, van de Provinciale Staten en de Tweede Kamer nauwkeuriger maken?”

Allereerst werd onderzocht in welke mate demografische gegevens, ‘geslacht’ en ‘leeftijd’, op Twitter verschillen van de daadwerkelijke verhoudingen in stemgedrag. Dit werd onderzocht voor twee politieke verkiezingen in Nederland, te weten: de Tweede Kamerverkiezingen in 2012 en de Provinciale Statenverkiezingen in 2015. Om hier antwoord op te geven, zijn de verhoudingen in percentages geanalyseerd. Daarin komt naar voren dat er grote verschillen op te merken zijn tussen de daadwerkelijke verhoudingen mannen en vrouwen die een stem hebben uitgebracht in 2012 en 2015 en de verhouding mannen en vrouwen die TweetGenie classificeerde. Mannen zijn in beide jaren oververtegenwoordigd als het gaat om de frequentie van politieke tweets ten opzichte van vrouwen. De verschillen in de leeftijdscategorieën liggen iets anders. Waar de verhoudingen in de daadwerkelijke stemgedragingen enigszins constant lijken te zijn over 2012 en 2015, is dit bij de classificatie van TweetGenie anders. De 55-plus categorie is bij beide jaren ondervertegenwoordigd, maar de jongste categorie is in 2012 het grootst en in 2015 is dit de leeftijdscategorie van 35 tot en met 54 jaar. De verhoudingen in demografische gegevens op Twitter ten opzichte van de daadwerkelijke verhoudingen lijken dus verschillen te vertonen.

Daarnaast was de vraag in hoeverre de verkiezingsuitslag nauwkeuriger voorspeld kon worden wanneer de demografische gegevens van de twitteraars bekend zijn in vergelijking met de opiniepeilingen en wanneer de demografische gegevens van de twitteraars niet bekend zijn. Wanneer gekeken wordt naar de Tweede Kamerverkiezingen in 2012 blijkt uit de resultaten dat de opiniepeilingen nog altijd sterker correleren met de daadwerkelijke verkiezingsuitslag van dat jaar ten opzichte van de voorspelling middels Twitter met en zonder aanpassing naar demografie. De voorspelling zonder aanpassing is net iets nauwkeuriger dan de voorspelling met aanpassing gebaseerd op de classificatie door TweetGenie. Ook de MAE toonde aan dat de opiniepeilingen de meest nauwkeurige voorspelling geven van de daadwerkelijke verkiezingsuitslag. Uit de resultaten blijkt dat de Provinciale Statenverkiezingen in 2015 ook het beste worden voorspeld door de opiniepeilingen, gekeken naar de correlaties. De voorspelling met aanpassing op demografie is nauwkeuriger dan de voorspelling zonder demografie. Volgens de MAE is de voorspelling volgens de opiniepeilingen het meest nauwkeurig, maar is de uitslag van Twitter met

aanpassing nauwkeuriger dan enkel de frequenties van politieke tweets. Uit deze analyse blijkt dus dat de verkiezingsuitslag in 2012 en 2015 door de opiniepeilingen het meest nauwkeurig wordt voorspeld, maar in 2015 de voorspelling met aanpassing op demografie een betere voorspelling geeft dan de voorspelling zonder aanpassing. Hiermee is antwoord gegeven op deelvraag 2: “In welke mate kan de verkiezingsuitslag nauwkeurig voorspeld worden wanneer de demografische gegevens van de twitteraars bekend zijn in vergelijking met de opiniepeilingen en wanneer de demografische gegevens van de twitteraars niet bekend zijn?”

Ten slotte zijn de verschillen in Twittergedrag tussen mannen en vrouwen en de leeftijdscategorieën geanalyseerd. Uit de resultaten blijkt dat er significante verschillen zijn tussen de politieke tweets van mannen en vrouwen en tussen de leeftijdscategorieën. Dit geldt zowel voor 2012 als voor 2015. Bij beide jaren komt naar voren dat vrouwen meer politieke tweets verstuurden waarin GroenLinks, CU, Partij voor de Dieren en 50Plus voorkwamen dan mannen en dat mannen bij beide jaren meer tweets over de VVD, PVV en CDA verstuurden ten opzichte van vrouwen. De leeftijdscategorie van 18 tot en met 34 jaar tweette in beide jaren het meest over de VVD en SGP, vergeleken met de andere leeftijdscategorieën. Dit terwijl de leeftijdscategorie van 35 tot en met 54 jaar beide jaren meer tweets over de CDA verstuurde. Bij de jongste en de 55-plus leeftijdscategorie zijn er geen patronen over beide jaren te zien. Gerelateerd aan deelvraag 3: “In hoeverre zijn er verschillen binnen de demografische dimensies in politiek Twittergedrag?”, kan geantwoord worden dat er in dat opzicht significante verschillen te zien zijn tussen de diverse demografische dimensies.

Discussie

Verschillende onderzoekers (Mislove et al., 2011; Tumasjan et al., 2010; Schoen et al., 2013; Tjong Kim Sang & Bos, 2012) plaatsten de kanttekening, bij het gebruik van Twitterdata in onderzoek, dat de demografische gegevens niet bekend zijn. In de huidige studie zijn de juiste gegevens hiervoor wel gegenereerd. Dat mannen zowel in 2012 als in 2015 bijna driekwart van de politieke twitteraars vormen is opmerkelijk ten opzichte van de cijfers met betrekking tot de activiteit in verdeling van mannen en vrouwen op Twitter. Daaruit bleek namelijk dat deze verschillen niet zo groot waren (Dietz, 2015). Mogelijk komt dit doordat mannen meer geïnteresseerd zijn in politiek dan vrouwen (Mestre & Marín, 2012). Ook is de zwijgspiraal onder vrouwen wellicht groter dan onder mannen; het uitkomen voor de politieke preferentie kan soms lastig zijn, omdat politieke kwesties vaak wat gevoelig liggen. Dat de oudste groep politieke twitteraars (de 55-plus categorie) ondervertegenwoordigd was in beide jaren, komt overeen met het nationaal social mediaonderzoek van Boeke et al. (2014) die ook inzichtelijk maakten dat de grootte van de groep twitteraars per oudere leeftijdscategorie afneemt. In 2012 bevinden de meeste politieke twitteraars zich inderdaad in de leeftijdscategorie van 18 tot en met 34 jaar, maar in 2015 is de grootste groep te vinden in de leeftijdscategorie van 35 tot en met 54 jaar. Dit zou kunnen komen doordat de leeftijden op Twitter opschuiven tussen 2012 en 2015. Iemand die in 2012 32 is, komt in 2015 in de volgende leeftijdscategorie terecht. Ook kan het liggen aan het type verkiezing in 2015: de Provinciale Statenverkiezingen, die over het algemeen ook een lagere opkomst hadden dan de Tweede Kamerverkiezingen in 2012 (Kiesraad, 2012; NOS, 2015).

Sanders en Van den Bosch (2013) en Tumasjan et al. (2010) trachtten de politieke verkiezingsuitslagen te voorspellen aan de hand van de frequenties van politieke tweets. De nauwkeurigheden van de voorspellingen in die studies waren goed, maar waren nog niet beter dan de opiniepeilingen. Dit geldt ook voor de voorspellingen in de huidige studie. De voorspellingen zonder aanpassing lijken dicht bij de nauwkeurigheid van de eerdere onderzoeken te komen, waarin de baseline al hoog was. De huidige studie laat zien dat de aanpassing op demografie de voorspellingen vervolgens niet veel onnauwkeuriger maakt en in 2015 zelfs een preciezere voorspelling weergeeft dan de uitslag zonder aanpassing. Tjong Kim Sang en Bos (2012) poogden ditzelfde al te doen, maar hadden de geschikte data hier niet voor, waar dat in dit onderzoek wel gerealiseerd is.

Overeenkomend met de bevindingen van Deschouwer et al. (2009) en Inglehart en Norris (2000), is ook in deze studie te zien dat het stemgedrag (gemeten aan politieke tweets) verschilt tussen mannen en vrouwen en tussen verschillende leeftijdscategorieën. Tevens geeft

dit weer dat Twitter kan dienen als analysetool waarmee inzicht verkregen kan worden in welke groepen politiek tweeten tijdens de verkiezingen in Nederland. Hiermee is het vervolgens mogelijk de publieke opinie te monitoren.

Implicaties

Voor zover bekend is dit het eerste wetenschappelijke onderzoek dat Nederlandse verkiezingen vergelijkt en daarin ook demografische gegevens meeneemt in combinatie met TweetGenie. Deze studie kan als fundering dienen voor verder onderzoek in dit thema, aan de hand van een soortgelijke methode die onder andere de mogelijkheid biedt om veel data in één keer te analyseren.

Zoals in de introductie van dit onderzoek beschreven, overladen media het publiek met opiniepeilingen in verkiezingstijd. Het is mogelijk om de voorspellingen aan de hand van social media mee te nemen. Media, als klassieke media (kranten en televisie) en online media (nieuwssites en blogs), kunnen deze voorspellingen aan de hand van Twitter gebruiken. Het zou als een tweede exitpoll kunnen dienen, wanneer geselecteerd wordt op een tijdrange van een aantal uur op de dag van verkiezingen. Hierbij moet men wel bedacht zijn op het feit dat dit gaat om real-time voorspellingen; dit was in het huidige onderzoek nog niet mogelijk.

De resultaten voortkomend uit dit onderzoek zijn in het politieke domein van belang. Voor politieke partijen zijn deze gegevens nuttig, omdat hiermee ingespeeld kan worden op de actualiteiten onder twitteraars. Vooral als dit real-time ingezet kan worden, is het mogelijk om doelgroeptargeting in te zetten. Hiermee wordt bedoeld dat er gericht (geadverteerde) berichten geplaatst worden die aansluiten bij de doelgroep op Twitter bij een desbetreffende politieke partij. Gekeken naar alleen de frequenties van politieke partijen, kan er gemonitord worden wat er speelt onder de politieke twitteraars. Dit is onder andere van belang voor de leden van de politieke partijen en degenen die de campagnes opzetten. Ook voor andere politici is het nuttig om te monitoren wat er onder de Twitterpopulatie speelt. Het creëert enigszins een afspiegeling van de populatie, waardoor er waardevolle informatie uit gefilterd kan worden waarbij bijvoorbeeld vroegtijdig crises op de radar kunnen komen.

De inzichten in de demografische gegevens van de twitteraars die via TweetGenie verkregen zijn, zijn niet alleen nuttig in het politieke domein. Dit is breder inzetbaar en wordt in meerdere en mindere mate al toegepast. In het gezondheidsdomein zijn er praktische implicaties mogelijk. Met een methode als in het huidige onderzoek kan eventueel ook geanalyseerd worden of vooral ouderen of jongeren over een epidemie praten, of vooral mannen of vrouwen. Daarmee zouden op een snelle manier risicogroepen geprofileerd kunnen

worden. Het is ook voor andere doeleinden zeer bruikbaar, bijvoorbeeld in het marketingdomein, zoals Rao en Yarowsky (2010) ook al stelden. Door kennis van de onderwerpen waar mensen over praten en welke mensen daarover praten kan er gericht marketing ingezet worden. Hierbij valt te denken aan reputatiemanagement en tevens doelgroepgerichte (geadverteerde) berichten. Er moet dan wel rekening gehouden worden met het ethische aspect rondom het vergaren van persoonsgegevens van mensen. Deze worden niet vrijwillig gegeven, maar zijn wel openbaar. Te allen tijde moet gewaarborgd worden dat deze anoniem behandeld worden.

Dit onderzoek is over het geheel gezien een stap verder in het monitoren van de publieke opinie middels social media. Het geeft weer dat het onder andere ingezet kan worden als analysetool in verschillende vakgebieden. Hiermee kunnen de discrepanties tussen groepen naar demografie weergegeven worden, zoals het met de data van dit onderzoek mogelijk was weer te geven dat de leeftijdscategorie van 18 tot en met 34 jaar significant meer politiek tweet over de VVD en SGP en dat mannen significant meer tweeten over de VVD, PVV en CDA. Deze informatie is zeer bruikbaar voor onder andere sociologen.

Kritische reflectie

Betreffende de verhoudingen van demografische gegevens is het van belang aan te duiden dat de afwezigheid van vrouwen en oudere politieke twitteraars een groot effect heeft gehad op de weging van de tweets. Vrouwen en de oudste leeftijdscategorie waren ondervertegenwoordigd, waardoor de weging van de tweets zeer zwaar werd. De categorie “vrouw 55-plus” van 2015 moest daarom bijvoorbeeld herwogen worden met een gewicht van 80.87. Zoals Gayo-Avello (2011) al stelde, blijft het gebruik van Twitterdata een deelverzameling, omdat niet iedereen het gebruikt.

Gekeken naar enkel de frequenties van tweets is hierop aan te merken dat de lijsttrekker groter kan zijn dan de partij. Wanneer een lijsttrekker rumoer veroorzaakt, zoals Geert Wilders, kan dit van invloed zijn op de stemmen voor de PVV, maar ook op het aantal Twitterberichten dat gestuurd wordt over de partij. Tevens kan een kleine achterban met veel activiteit op social media zorgen voor effect op de resultaten vanuit Twitter. Een voorbeeld is GroenLinks, die online zeer actief is en daardoor ook meer aandacht online genereert, zo stelden ook Sanders en Van den Bosch (2013) al. Die trend is in de huidige studie ook terug te zien; de voorspelde percentages voor GroenLinks zijn aanzienlijk hoger dan de daadwerkelijke verkiezingsuitslagen en de opiniepeilingen. Anderzijds zal de 50Plus partij minder online actievelingen hebben. Gekeken naar de resultaten van Boekee et al. (2014),

bleek dat oudere leeftijdsgroepen minder actief zijn op Twitter. Aansluitend daarop was in de data van TNS NIPO te zien dat de groep van 55-plus veruit het meeste stemde op de 50Plus partij, die ook in de data van het huidige onderzoek ondervertegenwoordigd was. Deze punten zijn van belang bij de interpretatie van de resultaten.

De leeftijd van mensen boven de 30 jaar is lastig in te schatten, bleek uit eerder onderzoek van Nguyen et al. (2013a). Mislove et al. (2011) en Beretta et al. (2015) beaamden eerder al dat voornamelijk geslacht goed te detecteren is en dat leeftijd lastiger is en meer tijd kost. Dit is terug te zien in de scatterplot in de methode. Zo classificeerde TweetGenie een maximale leeftijd van 70 jaar en de annotatoren 73 bij de overlap. TweetGenie classificeert anders dan een humane annotator. Een verschil hierin was onder andere dat de annotator profielfoto's kon bekijken voor een inschatting en TweetGenie dit enkel op basis van tekst kon doen. Deze verschillen leverden ruis op, wat ook te zien was in de interbeoordelaarsbetrouwbaarheid van leeftijden. Dit liet zien dat die vergelijking lastig is te maken, omdat de classificaties door beide op geheel andere wijzen zijn gemaakt. Uiteindelijk bleken de annotaties van de annotatoren en TweetGenie toch sterk te correleren, wat liet zien dat TweetGenie een degelijke classificatie kon maken. Echter, deze discrepanties kunnen van invloed zijn geweest op de classificatie. Een verklaring voor de "jongere" classificatie van TweetGenie is dat snelle trends en veranderingen in taal een rol spelen. Neem het gebruik van het woord "YOLO". Het gebruik van woorden verandert snel en het komt voor dat een woord korte tijd 'hip' is, zoals "YOLO", maar daarna ook weer uit de mode raakt. Taal is heel dynamisch en subtiel en voor TweetGenie zijn cynisme en sarcasme daarom vaak ook lastig te herkennen. Deze kleine subtiliteiten en de complexiteit die het meebrengt voor TweetGenie, zullen ook voorkomen bij het classificeren van geslacht en leeftijd aan de hand van taalgebruik.

Wat tevens van invloed kan zijn geweest op de classificatie van leeftijden is dat bij de leeftijden van 2012 vier jaar is geminderd en bij 2015 één jaar. Deze keuze is gebaseerd op het feit dat TweetGenie de classificatie op basis van de actuele tweets vaststelt; in het huidige onderzoek (in 2016 uitgevoerd) dus één en vier jaar nadat de twitteraars hun politieke tweets hebben verstuurd. Het is mogelijk dat deze vereffening ervoor gezorgd heeft dat twitteraars in andere leeftijdscategorieën zijn vastgesteld dan indien deze aanpassing niet gedaan was.

In dit onderzoek wordt een tweet waarin een partij voorkomt als stem geteld. Echter, de tweet hoeft niet per definitie te gaan over de partij waar de twitteraar daadwerkelijk op stemt. Dit in tegenstelling tot opiniepeilingen, waarin mensen gevraagd wordt op welke partij men stemt. Deze kritiek met betrekking tot de biases in social media data werd ook al gegeven

door Schoen et al. (2013). Hierbij was de assumptie in het huidige onderzoek dat deze discrepantie zichzelf op een natuurlijke manier uitbalanceerde, doordat dit voor alle partijen het geval is geweest.

Er zijn verschillen tussen de demografische gegevens van de twitteraars tussen 2012 en 2015 en doorpakkend op dat punt: per verkiezing fluctueert de correlatie inclusief de nauwkeurigheid. Dit heeft te maken met de demografische gegevens (oftewel bevolkingssamenstelling) per jaar, de soort verkiezing, het stemgedrag, de behandeling van de data en de keuze van de methode. Zo is met terugwerkende kracht een voorspelling gemaakt van de Tweede Kamerverkiezingen en de Provinciale Statenverkiezingen, maar deze twee verkiezingen verschillen alleen in opkomstpercentages al veel. Dit maakt het lastig om één standaard door te voeren. Echter, het heeft wel de mogelijkheid gegeven om twee verkiezingen en twee jaren samen in één onderzoek te bestuderen.

Suggesties voor toekomstig onderzoek

In de huidige studie zijn de demografische gegevens tot het classificeren van geslacht en leeftijd beperkt. Uiteraard zijn dit niet de enige demografische gegevens die van twitteraars te herleiden zijn. Dit laat ruimte open voor onderzoek waarin de demografische bias mogelijk nog verder gereduceerd kan worden. Een studie waarin andere demografische gegevens worden meegenomen, zoals: geo-locatie, etniciteit, beroep, inkomen en gezinssituatie. Hiermee kan getracht worden een studie met een hogere representativiteit op te zetten en tevens kan gekeken worden welke dimensies de meest gedegen predictoren zijn.

Zoals Gayo-Avello (2012) het in zijn artikel noemt; studies als deze zijn post-hoc analyses. Onderzoek dat zich richt op het real-time voorspellen van verkiezingen en niet enkel achteraf zou een toevoeging zijn op het huidige werk. Dit blijft complex, omdat het analyseren en het verwerken van demografische gegevens vaak veel tijd kost. Juist daarom is aanvullend onderzoek nodig naar de mogelijkheden van real-time voorspellingen middels Twitter. Dit zodat er een daadwerkelijke voorspelling gecreëerd kan worden, waar Schoen et al. (2013) al op wezen. Het is een groot voordeel van *big data* en TweetGenie dat een zeer grote hoeveelheid tweets geanalyseerd kan worden. Vele malen meer en sneller dan humane annotatoren en goedkoper dan opiniepeilingen. Dit maakt het rendabel om dit te toetsen tijdens bijvoorbeeld de Britse parlementaire verkiezingen in 2016 of de Nederlandse Tweede Kamerverkiezingen in 2017.

Tjong Kim Sang en Bos (2012) merkten het belang ervan al op en pasten het toe in hun studie: sentimentanalyse. Een tweet over een partij is niet per definitie een positieve

tweet, deze kan ook negatief of neutraal zijn. In het huidige onderzoek is geen sentimentanalyse toegepast, maar dit zou in de toekomst wel kunnen zorgen voor een reducering van de biases in social media data wanneer dit gecombineerd wordt met de werkwijze van deze studie. Daarnaast zijn nu alle politieke tweets van een persoon meegenomen. In het vervolg zou het een waardevolle toevoeging kunnen zijn slechts een enkele tweet per persoon mee te nemen. Tijdens de daadwerkelijke verkiezingen kan er namelijk ook slechts één stem per persoon uitgebracht worden.

Een laatste aanvulling op het huidige onderzoek zou een studie in een ander domein zijn. Zoals bij de implicaties besproken, kunnen deze nieuwe inzichten in een breder kader geplaatst worden en is het mogelijk zeer waardevol in het gezondheids- of marketingdomein. Een additioneel onderzoek hierin zou specifieke inzichten voor deze domeinen kunnen geven, dat het daarmee direct applicabel maakt.

Tot slot

Twitter kan in zekere mate dienen als een indicator voor real world events. Dit onderzoek heeft nieuw inzicht gegeven in de monitorende en voorspellende kracht van Twitter. Opiniepeilingen voorspellen nog altijd beter, hoewel de voorspellingen aan de hand van Twitter met en zonder aanpassing op demografie tevens behoorlijk nauwkeurig waren. Belangrijke eerste stappen naar een representatief Twitteronderzoek zijn gezet, hoewel dit onderzoeksveld nog niet is gecompleteerd. Het blijft complex en er blijven kritische kanttekeningen bij te stellen, wat nader onderzoek nodig maakt. Dit ook om weer te kunnen geven of *big data* vanuit Twitter inderdaad kan dienen als een blauwdruk voor een “glazenbol”.

Dankwoord

Graag wil ik in dit dankwoord de gelegenheid nemen om een aantal mensen te bedanken die een bijzondere bijdrage aan mijn onderzoek hebben geleverd. Allereerst wil ik Drs. Eric Sanders bedanken voor de mogelijkheid om met dit interessante onderwerp aan de slag te gaan, maar ook voor de gerichte begeleiding die ik gekregen heb om deze scriptie uit te kunnen voeren die bestond uit adviezen, feedback en supervisie. Ook wil ik Prof. Dr. Antal van den Bosch bedanken voor zijn vakkundige feedback en adviezen en tevens ook de inbreng van nieuwe inzichten. Daarnaast had ik dit onderzoek niet uit kunnen voeren zonder de hulp van Dong Nguyen die TweetGenie beschikbaar stelde voor de data van dit onderzoek, daar wil ik haar heel erg voor bedanken. Tevens ben ik TNS NIPO ook zeer dankbaar voor het beschikbaar stellen van de daadwerkelijke demografische verhoudingen van stemmers. Dit was erg belangrijk in het proces om de data recht te trekken. Tot slot wil ik masterstudent Eline Pilaet bedanken voor haar bereidheid om samen met mij zo'n 2900 Twitteraccounts te annoteren; haar werk heeft tevens een belangrijk deel van mijn onderzoek mogelijk gemaakt.

Literatuur

- Beretta, V., Maccagnola, D., Cribbin, T., & Messina, E. (2015). An interactive method for inferring demographic attributes in Twitter. *In Proceedings of the 26th ACM Conference on Hypertext & Social Media*, 113-122.
- Boekee, S., Engels, C., & Van der Veer, N. (2014). *Nationale Social Media Onderzoek 2014*. Amsterdam: Newcom Research & Consultancy BV.
- CBS (2008, 10 november). Betrokkenheid nieuws en politiek, 2006. CBS. Geraadpleegd van <http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=71722NED&D1=4&D2=1-9&HDR=T&STB=G1&VW=T>
- CBS (2016, 11 maart). Bevolkingsteller. CBS. Geraadpleegd van <http://www.cbs.nl/nl-NL/menu/themas/bevolking/cijfers/extra/bevolkingsteller.htm>
- Ciulla, F., Mocanu, D., Baronchelli, A., Gonçalves, B., Perra, N., & Vespignani, A. (2012). Beating the news using social media: the case study of American Idol. *EPJ Data Science*, 1, 1-11.
- Culotta, A. (2014). Reducing sampling bias in social media data for county health inference. *In Joint Statistical Meetings Proceedings*, 1-12.
- Culotta, A., Ravi, N. K., & Cutler, J. (2015). Predicting the demographics of Twitter users from website traffic data. *In Proceedings of the International Conference on Web and Social Media (ICWSM)*, in press, 72-78.
- Daas, P. J., Burger, J., Le, Q., Ten Bosch, O., & Puts, M. J (2016). Profiling of Twitter users: a big data selectivity study. *CBS*, 1-25.
- Deschouwer, K., Delwit, P., Hooghe, M., & Walgrave, S. (2010). *De stemmen van het volk. Een Analyse van het Kiesgedrag in Vlaanderen en Wallonië op 7 juni 2009*. Brussel: VUBPRESS
- Dietz, R. (2015). Hoe gedragen we ons op Twitter? *Marketingfacts*. Geraadpleegd van <http://www.marketingfacts.nl/berichten/hoe-gedragen-we-ons-op-twitter>
- Frankwatching (2014, 19 juni). Onmenselijke voorspellingen doen met machine learning. *Frankwatching*. Geraadpleegd van <http://www.frankwatching.com/archive/2014/06/19/onmenselijke-voorspellingen-doen-met-machine-learning/>
- Gayo-Avello, D. (2011). Don't turn social media into another 'Literary Digest' poll. *Communications of the ACM*, 54, 121-128.
- Gayo-Avello, D. (2012). "I wanted to predict elections with Twitter and all I got was this lousy paper" A balanced survey on election prediction using Twitter data. arXiv preprint arXiv:1204.6441.

- Gayo-Avello, D. (2013). A meta-analysis of state-of-the-art electoral prediction from Twitter data. *Social Science Computer Review*, 31, 649-679.
- Gelman, A. (2007). Struggles with survey weighting and regression modeling. *Statistical Science*, 22, 153-164.
- Goerres, A. (2007). Why are older people more likely to vote? The impact of ageing on electoral turnout in Europe. *The British Journal of Politics & International Relations*, 9, 90-121.
- Huberty, M. (2015). Can we vote with our tweet? On the perennial difficulty of election forecasting with social media. *International Journal of Forecasting*, 31, 992-1007.
- Inglehart, R., & Norris, P. (2000). The developmental theory of the gender gap: Women's and men's voting behavior in global perspective. *International Political Science Review*, 21, 441-463.
- Jüngherr, A., Jürgens, P., & Schoen, H. (2012). Why the pirate party won the german election of 2009 or the trouble with predictions: A response to Tumasjan, A., Sprenger, T.O., Sander, P.G., & Welpe, I.M. "Predicting elections with twitter: What 140 characters reveal about political sentiment". *Social science computer review*, 30, 229-234.
- Kiesraad (2012). Officiële uitslag Tweede Kamerverkiezing 12 september 2012. *Kiesraad*. Geraadpleegd van <http://www.verkiezingsuitslagen.nl/Na1918/Verkiezingsuitslagen.aspx?VerkiezingsTypeId=1>
- Kiesraad (z.d.a). Provinciale Staten. *Kiesraad*. Geraadpleegd van <https://www.kiesraad.nl/verkiezingen/inhoud/provinciale-staten>
- Kiesraad (z.d.b). Tweede Kamer. *Kiesraad*. Geraadpleegd van <https://www.kiesraad.nl/verkiezingen/inhoud/tweede-kamer>
- Kiesraad (z.d.c). Verkiezingsuitslagen Provinciale Staten 2015 - Nederland. *Kiesraad*. Geraadpleegd van <http://www.verkiezingsuitslagen.nl/Na1918/Verkiezingsuitslagen.aspx?verkiezingsTypeId=4>
- Kirrane, D. E. (1990). Machine Learning. *Training and Development Journal*, 44, 24-29
- Mestre, T. V., & Marín, R. T. (2012). The persistence of gender difference in political interest. *Revista Española de Investigaciones Sociológicas (REIS)*, 138, 89-108.
- Mislove, A., Lehmann, S., Ahn, Y. Y., Onnela, J. P., & Rosenquist, J. N. (2011). Understanding the demographics of Twitter users. *ICWSM*, 11, 5th.
- Nigam, K., McCallum, A., Thrun, S., & Mitchell, T. (1998). Learning to classify text from labeled and unlabeled documents. *AAAI/IAAI*, 792.

- Nguyen, D., Gravel, R., Trieschnigg, D., & Meder, T. (2013a). TweetGenie: Automatic age prediction from tweets. *ACM SIGWEB Newsletter, Autumn*, 1-6.
- Nguyen, D., Gravel, R., Trieschnigg, D., & Meder, T. (2013b). "How old do you think I am?"; A study of language and age in Twitter. In *Proceedings of the Seventh International AAI Conference on Weblogs and Social Media*. AAAI Press.
- NOS (2015, 18 maart). Opkomst dicht bij laagste niveau ooit. NOS. Geraadpleegd van <http://nos.nl/artikel/2025437-opkomst-dicht-bij-laagste-niveau-ooit.html>
- NRC (2012, 14 september). De peilingencarrousel: De Wereld Draait Door was het ergst. NRC. Geraadpleegd van <http://www.nrc.nl/nieuws/2012/09/14/de-peilingencarrousel-de-wereld-draait-door-was-het-ergst>
- Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. *LREc*, 10, 1320-1326.
- RTL Nieuws (2016, 25 januari). 'Jongeren kiezen voor snapchat, halen neus op voor Twitter'. *RTL Nieuws*. Geraadpleegd van <http://www.rtlnieuws.nl/economie/home/jongeren-kiezen-voor-snapchat-halen-neus-op-voor-twitter>
- Rao, D., & Yarowsky, D. (2010). Detecting latent user properties in social media. In *Proceedings of the NIPS MLSN Workshop*, 1-7.
- Sakaki, S., Miura, Y., Ma, X., Hattori, K., & Ohkuma, T. (2014). Twitter user gender inference using combined analysis of text and image processing. *Proceedings of the 25th International Conference on Computational Linguistics*, (p. 54-61).
- Sanders, E.P., & Bosch, A.P.J. Van den (2013). Relating political party mentions on Twitter with polls and election results. *CEUR Workshop Proceedings*, 986, 68-71.
- Schoen, H., Gayo-Avello, D., Takis Metaxas, P., Mustafaraj, E., Strohmaier, M., & Gloor, P. (2013). The power of prediction with social media. *Internet Research*, 23, 528-543.
- Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PloS one*, 6, e19467.
- Sloan, L., Morgan, J., Burnap, P., & Williams, M. (2015). Who tweets? Deriving the demographic characteristics of age, occupation and social class from twitter user meta-data. *PloS one*, 10, 1-20.
- Tilley, J., & Evans, G. (2014). Ageing and generational effects on vote choice: Combining cross-sectional and panel data to estimate APC effects. *Electoral Studies*, 33, 19-27.

- Tjong Kim Sang, E., & Bos, J. (2012). Predicting the 2011 dutch senate election results with twitter. *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, 53-60.
- Tjong Kim Sang, E., & Bosch, A.P.J. Van den (2013). Dealing with big data: the case of Twitter. *Computational Linguistics in the Netherlands Journal*, 3, 121-134.
- Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welppe, I. M. (2010). Predicting elections with Twitter: What 140 characters reveal about political sentiment. *ICWSM*, 10, 178-185.

Bijlagen

Bijlage 1: Niet-menselijke Twitteraccounts

Tabel 10. Overzicht van de percentuele verhoudingen van de classificatie door TweetGenie van de niet-menselijke Twitteraccounts ($N = 346$)

Demografische gegevens	Classificatie TweetGenie
<i>Geslacht</i>	
Man	77.5
Vrouw	22.5
<i>Leeftijdscategorieën</i>	
< 18	1.2
18 - 34	17.9
35 - 54	72.3
55 +	8.7

Bijlage 2: Opiniepeilingen

Tabel 11. Overzicht opiniepeilingen (TNS NIPO en De Stemming) op 11 september 2012 waarbij het gemiddelde van de opiniepeilingen en de percentages per partij weergegeven zijn

Partij	TNS NIPO	De Stemming	Gemiddelde zetels	Percentage per partij
VVD	35	35	35.17	23.43
PVDA	34	34	34.10	22.72
SP	21	22	21.43	14.28
PVV	17	17	17.27	11.50
CDA	12	12	12.10	8.06
D66	13	11	11.83	7.88
GroenLinks	4	4	4.00	2.66
CU	6	7	6.10	4.06
SGP	2	3	2.67	1.78
PVDD	2	2	2.33	1.55
50PLUS	4	3	3.10	2.07

Tabel 12. Overzicht opiniepeilingen (Politieke Barometer. I&O Research en De Stemming) op 17 maart 2015 waarbij het gemiddelde van de opiniepeilingen en de percentages per partij weergegeven zijn

Partij	Politieke			Gemiddelde zetels	Percentage per partij
	Barometer	I&O Research	De Stemming		
VVD	26.4	24.5	24.2	24.93	16.67
PVDA	15.4	17.8	12	15.05	10.07
SP	18.6	19.2	23.2	20.10	13.44
PVV	25.3	20.5	24.2	23.15	15.48
CDA	18.4	16.9	20.8	18.53	12.39
D66	24	24.4	21.8	23.20	15.52
GroenLinks	5	7.7	4.4	6.05	4.05
CU	5.6	5.9	6.4	5.98	4.00
SGP	4	3.9	3.4	3.83	2.56
PVDD	4.6	4.1	4.6	4.35	2.91
50PLUS	2.6	5.1	4.8	4.38	2.93