



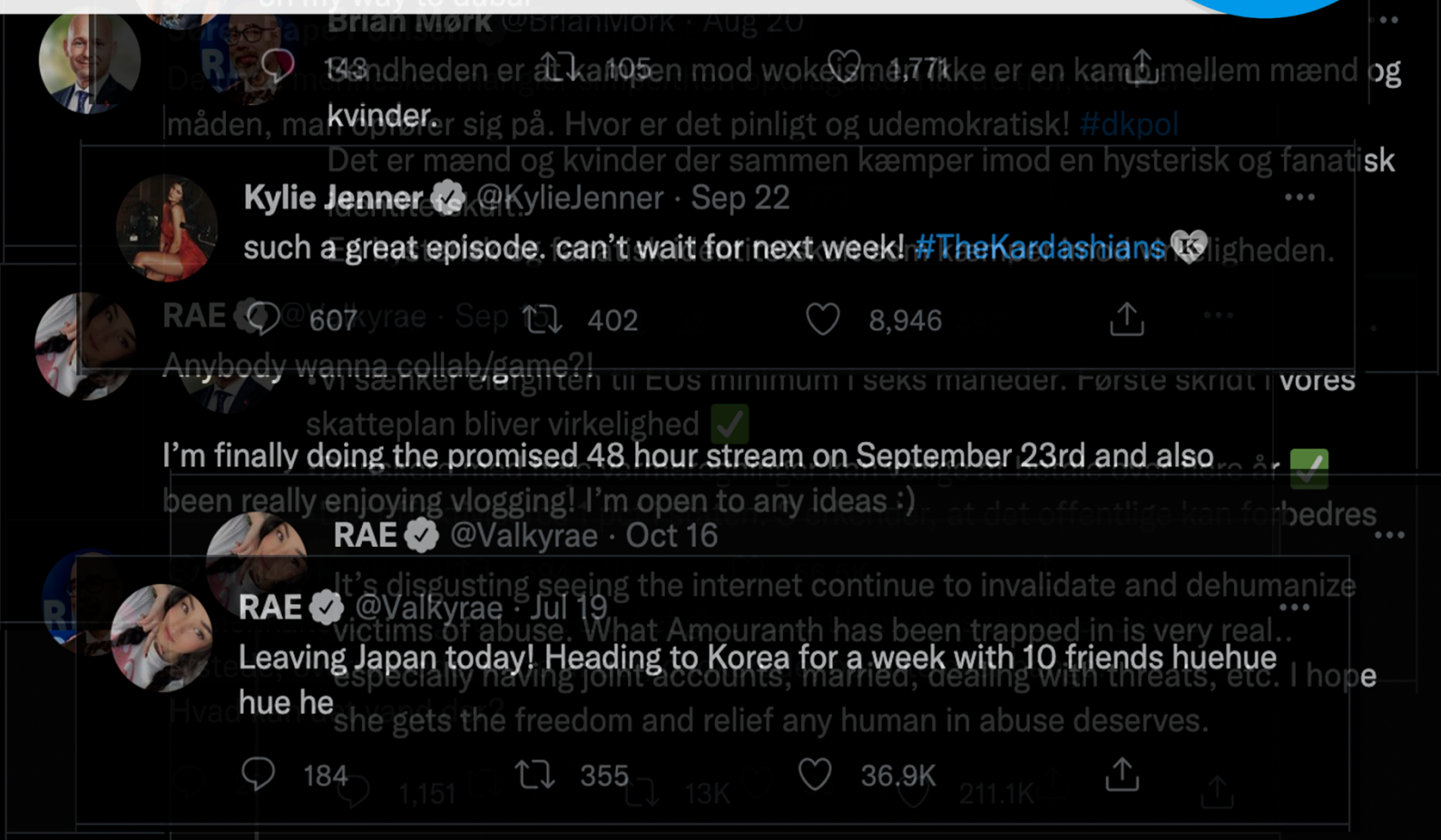
Forført af algoritmer

- en undersøgelse af algoritmernes favorisering af polariserende sprogbrug



Camilla Bargmann, [NGG](#)

Projekt Forskerspirer 2022, [HUM](#)



Indholdsfortegnelse

<u>INDLEDNING, PROBLEMFORMULERING OG FORMÅL</u>	<u>3</u>
<u>AFGRÆNSNING</u>	<u>4</u>
<u>TEORI</u>	<u>5</u>
<u>METODE</u>	<u>6</u>
<u>FREMGANGSMÅDE</u>	<u>7</u>
<u>PILOTFORSØGET</u>	<u>7</u>
<u>TIDSRAMME</u>	<u>10</u>
<u>BUDGET</u>	<u>10</u>
<u>KONKLUSION</u>	<u>10</u>
<u>PERPEKTIVERING</u>	<u>10</u>
<u>KONTAKTER OG TAK</u>	<u>11</u>
<u>BILAG</u>	<u>12</u>
<u>LITTERATURLISTE</u>	<u>17</u>

INDLEDNING, PROBLEMFORMULERING OG FORMÅL

I debatten om hvordan sociale medier påvirker samfundet og menneskers velvære, glemmes ofte et essentielt led; *social media algorithms* (herefter: *SMA*). Formålet med disse var oprindeligt at sikre, at brugeren kunne forslås opslag på baggrund af de informationer, de ubevidst konstant efterlader online. Samtidig skulle algoritmerne bruges til at markedsføre produkter målrettet den enkelte forbruger, og udgør derfor en stor forretningsmæssig værdi for de sociale medier. Brugeren bliver altså præsenteret for personliggjort og biased¹ indhold.

Udfordringen opstår, når sociale medier påvirker samfundet negativt, eksempelvis når de opslag, brugeren eksponeres for, er polariserende og ensidige. Brugeren modtager unuanceret information, og der skabes en distancering og polarisering igennem sociale medier. Forskning i Facebook og Instagram fra 2020² viser derudover, at især unge bliver påvirket af det, de bliver eksponeret for på sociale medier, og at dette fører til en forværring af deres mentale helbred. Disse problemer har ført til en usikkerhed for, hvor fysisk-, emotionelt- og samfundsskadelige sociale medier er blevet.

Mere forskning indenfor dette er derfor nødvendigt, da resultaterne fra forskningen kan være med til at belyse problemets omfang, samt skabe et politisk fundament for f.eks. at kunne lovgive omkring *SMA*, for at modvirke de skadelige effekter af, at polariserende opslag eller opslag med voldsomt indhold, boostes af algoritmerne. En anden gevinst af forskningen kan være øget opmærksomhed på, hvordan *SMA* fungerer, så brugere forholder sig mere kritisk til det indhold, de eksponeres for. Dertil kan forskningsresultaterne også medvirke til indførelse af et undervisningsforløb i *critical digital literacy*³ for studerende på ungdomsuddannelser.

Sociale medier og polariserende opslag er et bredt emne at forske i, så projektet afgrænses til at undersøge polariserende sprogbrug på Twitter. Dette medie er oplagt at undersøge for polariserende sprogbrug, da de fleste tweets er ren tekst i modsætning til mange andre medier.

¹ Biased = farvet

² Wells, G. & Horwitz, J. & Seetharaman, D. (2021) Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show: *The Wall Street Journal*, 2021. Fundet d. 10/22/2022 på <https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739?mod=djemalertNEWS>.

³ (12/16/2016). Teaching critical digital literacy. *Folgerpedia*. Fundet [10/22/2022]. På [https://folgerpedia.folger.edu/Teaching_critical_digital_literacy].

Ud fra disse overvejelser er min problemformulering:

Favoriseres polariserende sprogbrug på Twitter via social media algorithms?

Problemformuleringen er opstået på baggrund af min hypotese: *Twitters SMA favoriserer og booster opslag med polariserende sprog, så de dermed rækker ud til en større mængde brugere, fordi polariserende tweets i sig selv skaber mere engagement end neutrale tweets, hvilket fanges af algoritmerne, som yderligere booster opslaget.*

Ud fra problemformulering undersøges tweets for polariserende sprogbrug ved analyse af en mængde tweets ud fra en definition af polariserende sprog. Dernæst analyseres tweetets *engagement rate*⁴ ved at sammenligne tweetets *engagement*⁵ i forhold til profilens gennemsnitlige *engagement*. Slutteligt foretages en komparativ analyse, hvor der sammenlignes *engagement* for et tweet med et neutralt indhold af samme bruger.

Projektet er blot et lille område indenfor studier af sociale medier, og dog kan det give et bedre indblik i, hvordan sociale medier påvirker brugere ubevidst igennem de opslag, de eksponeres for. Projektet kan endvidere give et indblik i, hvorfor flere brugere benytter polariserende sprog i deres tweets, med det formål at opnå højere *engagement*, fordi tweetet boostes af algoritmerne.

I synopsen gennemgås en detaljeret beskrivelse af projektet samt et pilotforsøg inklusive dettes resultater i bilag 3.

AFGRÆNSNING

Projektet er en undersøgelse på Twitter, hvor *SMA* analyseres indirekte, da disse ikke er offentligt tilgængelige, men er firmaernes forretningshemmeligheder. I sin guideline skriver

⁴ Engagement rate: hvor stor en procentdel af likes, kommentarer og retweets det enkelte tweet har fået i forhold til hvor mange visninger tweetet har.

⁵ Engagement: se teoriafsnit

Twitter at tweets med høj grad af følgere⁶, popularitet og relevans⁷ vil udbredes (boostes) til flest brugere. Projektet vil derfor med komparativ analyse forsøge at skabe indblik i, hvordan disse algoritmer påvirker sociale mediers indhold, igennem en analyse af tweets' *engagement*. Derudover opdeles projektets anvendte tweets i to kategorier: polariserende og neutrale tweets. Der tages hermed ikke højde for, hvilken stemning, der ligger i et neutralt (ikke-polariserende) tweet. Om tweets er polariserende eller ej defineres⁸ i en selvudviklet software⁹. Projektet vil omfatte Twitter-brugere med over 10.000 følgere, da disse vil have et *reach*¹⁰ stort nok til at påvise, om polariserende tweets favoriseres eller ej.

TEORI

For at forstå, hvad der undersøges i dette projekt, er det nødvendigt at definere følgende centrale begreber: tweets, polariserende sprog og *engagement*.

Tweets

I dette projekt skal tweets forstås som opslag af ren skrift skrevet af en afsender. Dette vil dermed ekskludere profilens retweets, videoer, kommentarer til andre tweets, tags af andre personer og billeder. Jeg har vurderet det nødvendigt at ekskludere disse faktorer, da de automatisk booster tweetets *engagement* og dermed booster algoritmen uden nødvendigvis at have noget med deres sprogbrug at gøre.

Polariserende sprog

Polarisering betyder "at opdele eller opsplitte i skarpe modsætninger"¹¹, hvor polariserende sprog i dette projekt forstås som værdiladet og ensidigt sprog, der kan være nedsættende overfor andre med anvendelse af grove ord, hvilket resulterer i en dannelse af skarpe modsætninger. Sproget skaber "othering"¹² – en os-dem retorik. Softwarens¹³ definition af

⁶ Hvor mange følgere tweetets afsender har

⁷ (08/17/2021). About your Home timeline on Twitter. *Twitter*. Fundet [10/22/2022]. På <https://help.twitter.com/en/using-twitter/twitter-timeline>].

⁸ Se definition i teoriafsnit

⁹ Se metodeafsnit

¹⁰ Reach: tweetet vil række ud til andre brugere

¹¹ (05/07/2020). Polarisere. *Sproget.dk*. Fundet [10/22/2022]. På <https://sproget.dk/lookup?SearchableText=polarisere>].

¹² Cherry, Kendra. (12/13/2020). What is othering?. *Very well mind*. Fundet [10/22/2022]. På <https://www.verywellmind.com/what-is-othering-5084425>].

¹³ Se metodeafsnit

polariserende sprog vil ske på grundlag af en *Polarization Dictionary*¹⁴ inspireret af en tilsvarende ordbog, der blev opstillet i en forskningsartikel fra 2022¹⁵ for Reddit-opslag.

Engagement

Den totale mængde af et tweets interaktioner. Disse findes ved at aflæse tweetets statistik i form af likes, kommentarer og retweets.¹⁶

METODE

I forskningsprojektet anvendes den kvantitative og komparative metode, hvor der arbejdes indenfor feltet digital humaniora. Her vil jeg forholde mig til en mængde tweets fra twitter-brugere med mere end 10.000 følgere, gennem en relevant tidsperiode. Via selvudviklet software analyseres store mængder data (tweets), så mængden af tweets med polariserende sprogbrug og tweets med neutralt sprogbrug let kan optælles. Softwaren kan foretage sætningsanalyse og yderligere være programmeret til at bedømme, hvorvidt ordene i tweetet indeholder polariserende eller neutralt sprog. Projektets brug af kvantitativt metodearbejde vil gøre processen langt lettere og mere overskuelig. Der kan dog opstå fejkilder, når en computer skal vurdere, hvorvidt et ord eller en sætning opfattes neutralt eller polariserende, hvilket i højere grad vil tale for kvalitativt metodearbejde. Dog har jeg vurderet at jeg igennem *beta-tests* af softwaren, samtidigt med den programmeres, vil kunne finpudse programmet til ikke at begå den slags fejl. Derudover vil der løbende blive foretaget stikprøver af softwarens sorteringer af tweets, for at kvalitetssikre kriterierne.

Dernæst foretages den komparative analyse, hvor der opstilles to grafer for, hvor stor en procentdel af hhv. tweets med polariserende sprogbrug og tweets med neutralt sprogbrug bliver favoriseret af *SMA* på Twitter. Her udregnes hver af de to kategoriers procenter for, hvor mange polariserende og neutrale tweets med en *engagement rate*¹⁷ på min. 1% der er. Hvorvidt et tweet bliver favoriseret eller ej, bedømmes ud fra tweetets *engagement rate*. Hvis

¹⁴ Inspiration til min egen Polarization Dictionary: se bilag 1

¹⁵ Simchon, Almog. (2022) Troll and divide: the language of online polarization : *Oxford Academic*, 2022. Fundet d. 10/22/2022 på <https://academic.oup.com/pnasnexus/article/1/1/pgac019/6546199>.

¹⁶ (10/15/2013). About your activity dashboard. *Twitter*. Fundet [10/22/2022]. På [<https://help.twitter.com/en/managing-your-account/using-the-tweet-activity-dashboard>].

¹⁷ Engagement rate: hvor stor en procentdel af likes, kommentarer og retweets det enkelte tweet har fået i forhold til hvor mange visninger tweetet har

tweetet har fået en høj *engagement rate* (over 1%¹⁸), betyder det altså, at tweetet er blevet favoriseret af Twitters algoritmer. Resultaterne fra dette giver dermed mulighed for at foretage en komparativ analyse, for at kunne indikere, om Twitters *SMA* særligt fremhæver tweets med polariserende sprogbrug. Projektets brug af komparativt metodearbejde er uundværligt, da dette visualiserer databehandlingen.

FREMGANGSMÅDE

Projektet vil forløbe således, at jeg først vil indsamle data i form af tweets via det selvudviklede software. Dernæst vil softwaren analysere tweets ud fra den foreliggende definition af polariserende ord, neutrale ord, polariserende sætninger og neutrale sætninger. Ud fra dette vil mit software kunne identificere polariserende tweets via. sprog- og sætningsanalyse.

Dernæst vil software opgøre, hvor mange procent polariserende tweets har en *engagement rate* på over 1% i forhold til antal polariserende tweets brugeren har i alt samt hvor mange af de neutrale tweets har fået en *engagement rate* på over 1% i forhold til antal neutrale tweets brugeren har i alt. Ved at opstille hvor mange procent af alle polariserende opslag, der har en *engagement rate* på over 1% samt hvor mange procent af alle neutrale tweets, der har en *engagement rate* på over 1%, gøres projektets endelige resultater nemme og overskuelige at forstå og vurdere et resultat ud fra.

Slutteligt vil jeg sammenligne de to resultater jeg er nået frem til og dermed forholde mig til og analysere eventuelle forskelle og/eller ligheder mellem de to kategorier og konkludere, hvorvidt polariserende tweets bliver favoriseret af Twitters *SMA*. Jeg vil beskrive mine fund i en rapport eller videnskabelig artikel, så den kan offentliggøres.

PILOTFORSØGET

Formålet med pilotforsøget er at få en indikation af, om min hypotese kan be- eller afkræftes. Dette vil danne grundlag for at beslutte, om det på baggrund af pilotforsøgets resultater giver

¹⁸ (5/8/2022). Your guide to social media engagement rates.. *Adobe Express*. Fundet [10/22/2022]. På [<https://www.adobe.com/express/learn/blog/what-is-a-good-social-media-engagement-rate#what-does-social-media-engagement-rate-mean>].

mening at udføre det fulde projekt. Derudover er formålet med pilotforsøget at afprøve min definition af polariserende sprog¹⁹.

AFGRÆNSNING

Pilotforsøget afgrænses yderligere end det endelige projekt, da det anvendte softwareprogram stadig er under udvikling. Her er pilotforsøget en undersøgelse på Twitter, der ønsker at finde svar på det samme som projektet. Der undersøges blot fem Twitter brugere fra fem forskellige områder, gennem det sidste halve år. Brugerne har alle følgerantal på over 20000 følgere, da de, ligesom i projektet, skal have et stort nok *reach* til at påvise, om polariserende tweets favoriseres eller ej.

METODE

I pilotforsøget anvendes den kvalitative og komparative metode, hvor jeg forholder mig til en mængde tweets gennem en relevant tidsperiode. Jeg har her valgt at forholde mig til ti udvalgte tweets, som er blevet lagt op i løbet af de sidste 6 måneder (i perioden 17/4-2022 – 17/10-2022). Jeg har valgt at foretage undersøgelsen af fem brugere indenfor hvert sit område: En politiker, en komiker, en gamer, en influencer og en musiker. For hver bruger er to tweets udvalgt. De ti tweets er udvalgt på baggrund af en kombination af en stemningsanalyse (*sentiment analysis*) og en tematisk analyse af alle profilens tweets i den givne periode.

Dernæst er de analyseret ideografisk ved at foretage en grundig sproglig analyse af hvert tweet, hvor de er kodet i forhold til polariserende sprogbrug, og hvorvidt der fra tweetet udledes en positiv eller negativ tendens og/eller tone.

Slutteligt har jeg induktivt forsøgt at påvise en tendens, der enten af- eller bekræfter min hypotese.

Havde jeg på forhånd udviklet det software, jeg ønsker at udvikle i selve forskningsprojektet, ville pilotforsøget og projektet bestå af samme metoder, og jeg ville dermed via pilotforsøget opnå guidelines til at forbedre mit software. Fokus for dette pilotprojekt er at afprøve og sikre definitionen af polariserende sprog, og sikre, at der ikke opstår fejl i den kommende software

¹⁹ Se teoriafsnit

på baggrund af upræcise definitioner. Dette sikres igennem den tematiske analyse, kombineret med den sproglig detaljerede analyse. Jeg vil derudover anvende pilotforsøget til at afgøre, om jeg kan be- eller afkræfte min hypotese om, at polariserende tweets bliver favoriseret af Twitters *SMA*. Samtidigt kan pilotforsøgets resultat, til trods for at have en lav repræsentativitet, grundet den lille mængde data der analyseres, give et kvalificeret bud på, hvad der kan forventes at blive forskningsprojektets resultat. Jeg vil i min konklusion via induktion forsøge at forudsige, hvilke resultater der kan forventes af projektet, på baggrund af mine fund i pilotforsøget²⁰.

FREMGANGSMÅDE

Pilotforsøget gennemførtes ved, at jeg først valgte fem brugere fra fem forskellige områder, jeg ønskede at undersøge; politik, komik, gaming, influencing og musik. Disse brugere blev valgt ud fra forskellige kategorier, for at undersøge, om alle kategorier havde samme tendens og dermed gøre forsøget mere repræsentativt. Udover kategorierne blev brugerne valgt ud fra et kriterium om, at de min. skulle have 20.000 følgere, så de havde et stort nok *reach* til at påvise en tendens.

Dernæst foretog jeg en semantisk analyse af den enkelte brugers tweets gennem det sidste halve år (i perioden 17/4-17/10-2022). Herfra vurderede jeg, hvilke tweets var neutrale, og hvilke tweets var polariserende, hvor vurderingen tog udgangspunkt i min definition af polariserende sprog. Ud fra dette blev to repræsentative tweets valgt; en polariserende, en neutralt. Disse to blev grundigt sprogligt analyseret via. autopsy, for at tydeliggøre præcis, hvad karakteriserede de to tweets som polariserende og neutralt.

Jeg opgjorde, hvor mange retweets, likes og kommentarer (*engagement*) de to tweets havde fået, og sammenlignede resultaterne. Derefter afgjorde jeg, om det polariserende sprog havde haft en indvirkning på tweetets *engagement*, som dermed kunne indikere, om polariserende tweets favoriseres af Twitters *SMA* for den konkrete bruger. Jeg gentog denne proces med de fire andre brugere, for at se om der var en tendens i polariserende tweets, og hvor stort *engagement* disse tweets opnåede.

Slutteligt konkluderedes, om polariserende tweets favoriseres af Twitters *SMA*.

²⁰ Pilotforsøgets resultater: se bilag 4

PILOTFORSØGETS RESULTATER

Set bilag 4.

TIDSRAMME

Se bilag 2.

BUDGET

Se bilag 3.

KONKLUSION

På baggrund af tidligere forskning²¹ og mit pilotforsøgs resultater²² forventes det, at svaret til problemformuleringen: *Favoriseres polariserende sprogbrug på Twitter af social media algorithms?* vil være ja. Både pilotforsøget og tidligere forskning i sociale medier viser, at brugere af sociale medier bliver påvirket af polariserende og provokerende opslag, som medfører en reaktion fra mange brugere og derudover har negative konsekvenser for velvære, trivsel og mental velbefindende, især for den unge generation. Her forventes Twitter ikke at være en undtagelse. Hvis flere og flere brugere bliver provokerede af et polariserende tweet vil dette kunne ses på tweetets *engagement rate*. En forhøjet *engagement rate* betyder, at Twitters *SMA* har favoriseret tweetet, og at det når ud til flere brugere.

Projektet vil på sigt kunne bidrage til, at man på sociale medier kan bruge min selvudviklede software som udgangspunkt for at lave et filter, som udelukker polariserende tweets (og på sigt også tweets med andre typer af voldsomt og stødende indhold), således at de ikke længere boostes af *SMA*. Dette vil også bidrage til større trivsel blandt unge.

PERSPEKTIVERING

Forskningen kan danne grund for en ny lovgivning, der sikrer, at tweets med polariserende sprogbrug ikke bliver favoriseret af algoritmer, så konsumerne ikke ender med at få et

²¹ Simchon, Almog. (2022) Troll and divide: the language of online polarization : *Oxford Academic*, 2022. Fundet d. 10/22/2022 på <https://academic.oup.com/pnasnexus/article/1/1/pgac019/6546199>.

²² Se bilag 3

entydigt billede af, hvordan verden ser på forskellige problemer. Derudover kan forskningsresultaterne medvirke til at indføre et undervisningsforløb i *critical digital literacy*²³ for studerende på ungdomsuddannelser.

KONTAKTER OG TAK

Forskerkontakt

Kim Ebensgaard Jensen, lektor i engelsk sprog på KU.

Først og fremmest ønsker jeg at sige en kæmpe tak til min forskerkontakt, Kim Ebensgaard Jensen, for alt hans hjælp gennem hele forløbet.

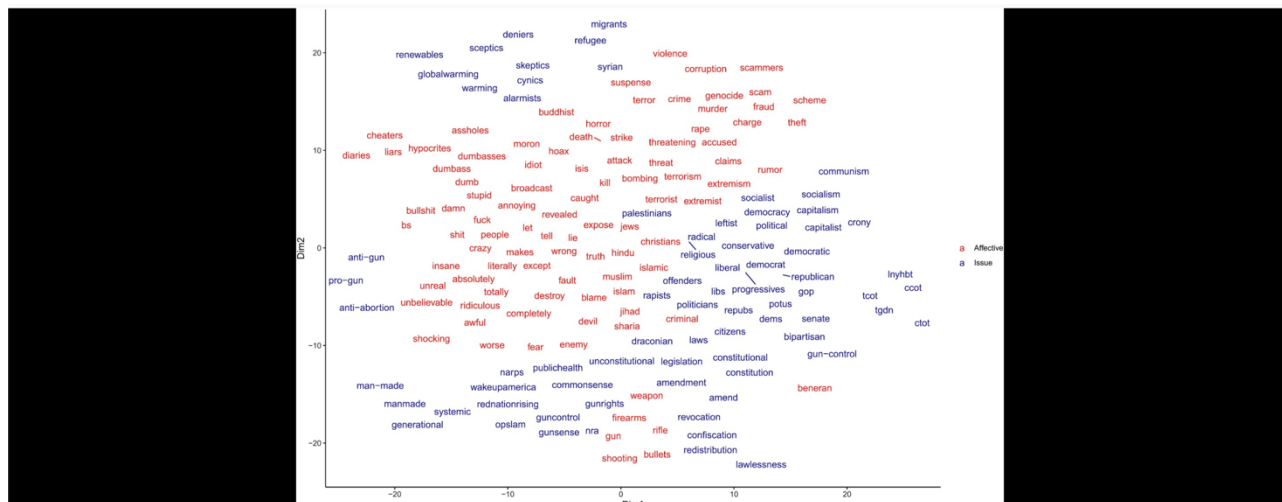
Derudover en kæmpe tak til min koordinator, Marianne Johansson, for støtten, de mange input og for at give mig muligheden for at deltage.

Slutteligt en kæmpe tak til mine forældre og min medspirer, Zacharias Beck, for altid at være motiverende, støttende og utroligt hjælpsomme gennem hele forløbet.

²³ (12/16/2016). Teaching critical digital literacy. *Folgerpedia*. Fundet [10/22/2022]. På [https://folgerpedia.folger.edu/Teaching_critical_digital_literacy].

BILAG

Bilag 1 - inspiration til egen *Polarization Dictionary*



Semantic space representation of the Polarization Dictionary. X- and Y-axes represent t-SNE dimensionality reduction of GloVe embeddings. Words in red mark the first cluster ("Affective") and words in blue mark the second cluster ("Issue").

Kilde: Simchon, Almog. (2022) Troll and divide: the language of online polarization : *Oxford Academic*, 2022.

Fundet d. 10/22/2022 på <https://academic.oup.com/pnasnexus/article/1/1/1/pgac019/6546199>.

Bilag 2 – projektets tidsramme

Projektet forventes at vare 12 måneder, hvor de første 6 måneder bruges på at programmere og række ud til andre programmører, de efterfølgende 2 måneder bruges på at beta-teste min software, og de sidste 4 måneder bruges på at udføre projektet. Rapportskrivningen er ikke medtaget i projektets tidsramme.

Varighed	
6 måneder	Programmering og udvikling af software (inkl. kursus i programmering)
2 måneder	Beta-testing af software
4 måneder	Udførelse af projektet

Bilag 3 – budget

	Pris
Bog om <i>Machine Learning</i>	120 kr.
Bog om <i>Python Machine Learning</i>	435 kr.
Bog om <i>Python</i>	310 kr.
Online kurser i <i>Python</i> og <i>Machine Learning</i> (medlemskab til websitet i 6 måneder)	1350 kr. (6 måneder)
Uforudsete udgifter	3000 kr.
<i>Cloud space</i> i Heruko (til opbevaring og publicering af færdigudviklet software) ²⁴	14000 kr. (14 måneder)

Bilag 4 – pilotforsøgets resultater

a) Tweets fra kategorien politik



De to ovenstående tweets repræsenterer området politik. Her ses et polariserende (øverste) og et neutralt (nederste) tweet.

²⁴ Mit software vil indgå i min rapport, hvor den opbevares på platformen *Heruko*, så den vil være tilgængelig for alle uanset tidspunkt på dagen

I det øverste tweet ses det polariserende sprogbrug tydeligt, idet politikeren retter sit tweet mod den unge del af befolkningen. Her beskriver han gruppens handlinger med adjektiverne ”pinligt og udemokratisk”, hvilket er værdiladede ord med negative konnotationer. Dette tweet danner hermed skarpe modsætninger: befolkningens ældre og unge, hvor han prøver at overbevise modtagerne om, hans holdning er den rigtige.

I det nederste tweet reklamerer politikeren udelukkende for sit parti, hvor han undgår værdiladet sprog og at skabe nogen skarpe modsætninger.

Ved at sammenligne de to tweets’ engagement, ses det tydeligt, at det polariserende tweet har fået flere likes og retweets. Andelen af kommentarer er næsten ens for de to tweets, men det generelle billede er klart, at det polariserende tweet får langt højere engagement end det neutrale tweet.

b) Tweets fra kategorien komik



De to ovenstående tweets repræsenterer området komik. Her ses et polariserende (øverste) og et neutralt (nederste) tweet.

I det øverste tweet ses det polariserende sprogbrug tydeligt, idet komikeren retter sit tweet mod tilhængere af wokeismen, hvor han gør brug af stærkt værdiladede, nedladende og provokerende ord som ”hysterisk”, ”fanatisk” og ”identitetskult”. Dette tweet danner dermed

to skarpe modsætninger: tilhængere af wokeismen og resten af befolkningen. Han beskriver dem både implicit og eksplicit som værende de forkerte og vanvittige.

I det nederste tweet joker komikeren om amerikanske film, hvilket ikke vil virke stødende for hans målgruppe, som primært vil bestå af danskere.

Ved at sammenligne de to tweets' *engagement*, ses det tydeligt, at det polariserende tweet har fået flere likes og retweets. Andelen af kommentarer er næsten ens for de to tweets, men det generelle billede er klart, at det polariserende tweet får langt højere *engagement* end det neutrale tweet.

c) Tweets fra kategorien gaming



De to ovenstående tweets repræsenterer området gaming. Her ses et polariserende (øverste) og et neutralt (nederste) tweet.

I det øverste tweet ses det polariserende sprogbrug tydeligt, idet gameren retter sit tweet mod alle brugere på internettet, hvor hun bruger stærkt værdiladede ord som "invalidate" og "dehumanize". Her bliver der stillet skarp kritik af internettet, og der bliver hermed dannet to skarpe modsætninger: brugere på internettet og alle andre.

I det nederste tweet reklamerer gameren for sin stream og sit brand som gamer. Der bliver på intet tidspunkt brugt stærkt værdiladede ord som i det ovenstående tweet, og der er ingen skarpe holdninger, der kommer til udtryk.

Ved at sammenligne de to tweets' *engagement*, ses det tydeligt, at det polariserende tweet har fået flere likes og retweets. Andelen af kommentarer er næsten ens for de to tweets, men det generelle billede er klart, at det polariserende tweet får langt højere *engagement* end det neutrale tweet.

De sidste to kategorier musik og influencing (repræsenteret af sangerinde Taylor Swift og influencer Kylie Jenner) kunne der ikke findes tweets med polariserende sprogbrug. Dermed kan disse to kategorier ikke være med til at drage en konklusion på svaret. Årsagen til dette kan være, at brugerne ønsker at beskytte deres brand og omdømme.

Ud fra pilotforsøgets resultater, konkluderes det, at polariserende tweets favoriseres af Twitters *SMA*. Dog er pilotprojektets størrelse for småt til at kunne konkludere noget endeligt om projektet, men resultaterne kan stadig vise en tendens blandt polariserende vs. neutrale tweets.

LITTERATURLISTE

Forskningsartikler

- Adoui, A. & Makdad, M. E. (2021) Social Media Algorithms and Decision Making in the Digital Age: *Research Gate*, 2021. Fundet d. 10/22/2022 på https://www.researchgate.net/publication/364213487_Social_Media_Algorithms_and_Decision_Making_in_the_Digital_Age.
- Imana, B. & Korolova, A. & Heidemann, J. (2022) Having your Privacy Cake and Eating it Too: Platform-supported Auditing of Social Media Algorithms for Public Interest: *Research Gate*, 2022. Fundet d. 10/22/2022 på https://www.researchgate.net/publication/362090077_Having_your_Privacy_Cake_and_Eating_it_Too_Platform-supported_Auditing_of_Social_Media_Algorithms_for_Public_Interest.
- Johnson, N. F. & Manrique, P. & Zheng, M. & Cao, Z. (2017) Population polarization dynamics and next-generation social media algorithms: *Research Gate*, 2017. Fundet d. 10/22/2022 på https://www.researchgate.net/publication/321903129_Population_polarization_dynamics_and_next-generation_social_media_algorithms.
- Sîrbu, A. & Pedreschi, D. & Kertész, J. & Giannotti, F. (2019) Algorithmic bias amplifies opinion fragmentation and polarization: A bounded confidence model: *Plos One*, 2019. Fundet d. 10/22/2022 på <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0213246>.
- Simchon, Almog. (2022) Troll and divide: the language of online polarization : *Oxford Academic*, 2022. Fundet d. 10/22/2022 på <https://academic.oup.com/pnasnexus/article/1/1/pgac019/6546199>.
- Harriger, J. A. & Evans, J. A. & Thompson, J. K. & Tylka, T. L. (2022) The dangers of the rabbit hole: Reflections on social media as a portal into a distorted world of edited bodies and eating disorder risk and the role of algorithms: *Elsevier*, 2022. Fundet d. 10/22/2022 på .

Internetsider

- (10/15/2013). About your activity dashboard. *Twitter*. Fundet [10/22/2022]. På [\[https://help.twitter.com/en/managing-your-account/using-the-tweet-activity-dashboard\]](https://help.twitter.com/en/managing-your-account/using-the-tweet-activity-dashboard).
- (12/16/2016). Teaching critical digital literacy. *Folgerpedia*. Fundet [10/22/2022]. På [\[https://folgerpedia.folger.edu/Teaching_critical_digital_literacy\]](https://folgerpedia.folger.edu/Teaching_critical_digital_literacy).

- (5/8/2022). Your guide to social media engagement rates.. *Adobe Express*. Fundet [10/22/2022]. På [<https://www.adobe.com/express/learn/blog/what-is-a-good-social-media-engagement-rate#what-does-social-media-engagement-rate-mean>].
- (05/07/2020). Polarisere. *Sproget.dk*. Fundet [10/22/2022]. På [<https://sproget.dk/lookup?SearchableText=polarisere>].
- Cherry, Kendra. (12/13/2020). What is othering?. *Very well mind*. Fundet [10/22/2022]. På [<https://www.verywellmind.com/what-is-othering-5084425>].
- (08/17/2021). About your Home timeline on Twitter. *Twitter*. Fundet [10/22/2022]. På [<https://help.twitter.com/en/using-twitter/twitter-timeline>].
- Wells, G. & Horwitz, J. & Seetharaman, D. (2021) Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show: *The Wall Street Journal*, 2021. Fundet d. 10/22/2022 på <https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739?mod=djemalertNEWS>.