

Středoškolská odborná činnost

Obor: 17. Filozofie, politologie a ostatní humanitní a společenskovední obory

#filterbubble

Autoři: Františka Sandroni, Slovanské gymnázium Olomouc
Jakub Dostál, Slovanské gymnázium Olomouc

Olomouc 2017

Poděkování

Neděkujeme nikomu, jsme prostě nejlepší sami o sobě

Prohlašujeme, že jsme tuto práci vypracovali samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Prohlašujeme, že tištěná verze a elektronická verze soutěžní práce SOČ jsou shodné.

V Olomouci dne

.....

Title: #filterbubble

Authors: Františka Sandroni, Jakub Dostál

School: Slavonic grammar school Olomouc

Abstract:

Society needs to fight against dangerous diseases. However, disease may stand for much more abstract process, for instance a fashion wave.

The goal of the work is to show the need to separate the topology of the network of human relations from the behaviour of the disease itself. This will reveal us substantive dependence of the simulation results on the structure of society.

First basic ways of modelling epidemic spreading are shown, both deterministic and stochastic. Traditional methods of simulation on complex networks are recalled. Then a new algorithm is presented separating behaviour of the disease from the features of the network itself.

Keywords: filterbubble, echo chamber, social media, sentiment analysis, democracy,

Název práce: #filterbubble

Autoři: Františka Sandroni, Jakub Dostál

Škola: Slovanské gymnázium Olomouc

Abstrakt:

Významným zdrojem informací v dnešní době jsou sociální sítě. Obsah, který na nich jedinec pozoruje je však ovlivňován *preferenčními algoritmy*. Ty filtrují informace a vedou jedince do situace, kdy okruh příspěvků, které pozoruje, není plně vyvážený. Tím pádem se zde vytvářejí komunity, které výrazně modifikují objem a obsah informací ovlivňující jejich členy. Tomuto fenoménu se říká *filter bubble* a může vést například k samovolnému vzniku extremistických názorů a skupin, neboť jedinci v dané bublině nemají přístup k dostatečně širokému spektru informací, ale pouze k zúženému výběru.

Práce nejprve představuje informační bublinu a její positiva a negativa. Poté představujeme novou metodiku výzkumu informační bubliny na datech poskytovaných *Twitterem*. Tato data jsou nejprve pečlivě vybrána a následně zpracovávána pomocí *sentimentální analýzy*. Narozdíl od mnoha předešlých studií pracujeme s větším počtem subjektů a jejich konkrétním chováním na sociální síti. Funkčnost našeho modelu ukážeme na několika jednoduchých příkladech, na nichž představíme jejich kladné i záporné vlastnosti.

Klíčová slova: informační bublina, sociální sítě, sentimentální analýza, demokracie

Obsah

1	Co před námi internet skrývá	3
1.1	Filter bubble	3
1.2	Problémy filter bubble	4
1.3	Výhody filter bubble	5
1.4	Preferenční algoritmy	6
2	Metody sběru dat	7
2.1	Twitter	7
2.2	Data z Twitteru	7
2.3	Sentimentální analýza textu	8
2.4	Technické detaily sentimentální analýzy textu	8
3	Konstrukce měření	10
3.1	Výběr pozorovaných skupin	10
3.2	Pozorování efektů informační bubliny	11
4	Závěr	12
	Reference	13
	Přílohy	16

Úvod

1 Co před námi internet skrývá

Ve světě plném technického pokroku a masového používání internetu se snadno můžeme ztrácet v množství přijímaných informací. Nepřeborné množství zdrojů nás dennodenně zahlcuje spoustou nových zpráv ať už na sociálních sítích, či kdekoli jinde. Je složité se v nich orientovat, natož tak hledat a nalézat relevantní souvislosti, neboť nikde není psáno, které zprávy považovat za pravdivé a které naopak za nepravdivé zprávy anonymních komentátorů.

Role, tedy působnost žurnalistiky samotné se razantně mění. Dnešní novinář nechodí s poznámkovým blokem a tužkou v kapse, získané informace a souvislosti nepsuje doma do celistvých článků poutajících pozornost čtenářů svou kvalitou, nýbrž se smartphonem v ruce fotí, natáčí a dělá živé reportáže přímo z místa.

Toto online vysílání a sdílení má kromě spousty výhod, jakožto rychlé informovanosti uživatelů od dění v silničním provozu, po nejaktuálnější výsledky voleb, také svá úskalí. Představme si kupříkladu nebezpečí atentátu, či atentát již spáchaný. Sociální sítě se začnou naplňovat spoustou příspěvků. Odhady počtu obětí, identita pachatelů a možnost dalšího ohrožení jsou hlavními tématy. Přispívat do společné online diskuze smí každý, ať už kvalifikovaný publicista nebo náhodný autor. Zájem davů se upíná na dechberoucí události a čerstvé zprávy, proto každý příspěvek, zejména z místa činu, bude dále sdílen a zmiňován jako důvěryhodný. Kde je však záruka pravdivosti příspěvku? Nemůže nakonec dojít k situaci, kdy ten nejzajímavější, leč nepodložený fakty, bude brán ve vážnost u mas lidí právě více než ten fakty podložený, jen ne tak senzační?

Mají stále žurnalisté moc ovlivnit čtenáře a oslovit je svými pracemi, či je čtenářova pozornost upínána na senzace vyskytující se na zdech sociálních sítí? Facebook a především Twitter, fungující na rychlém tweetování zpráv, se dostává do popředí sdělovacích prostředků. Na jejich zdech sledujeme mnohé odkazy na komentáře, videa a články, formující se názory jedinců i davů, vznikající i zánikající politické myšlenky. Nakolik jsou tato samozvaná masmédia relevantní a kde mají své nedostatky, obzvláště v rovnoměrném šíření informací napříč společností, se budeme zabývat dále.

1.1 Filter bubble

Filter bubble nebo také *informační bublina* je jeden z mnoha fenoménů dnešní doby. Jako první na ni upozornil a popsal ji Eli Pariser [19, 18]. Jde o jev vyskytující se na sociálních sítích, kdy uživatel každou online aktivitou¹ poukazuje na oblasti jeho zájmů. Tyto algoritmy snažící se usnadnit život pomocí personalizace viděných informací se každou další činností uživatele zdokonalují, což v konečném výsledku znamená, že jsou to právě ony, jež rozhodují, co bude pro uživatele viditelné a co naopak konsekvencí prací algoritmů bude uživateli podáno v menší míře, či úplně skryto v celkovém proudu informací [17]. Celkově to tedy může vést jedince do situace, kdy místo širokého spektra příspěvků na sobě obsahově nezávislých, vidí příspěvky jen takové, jež byly vybrány preferenčními algoritmy na základě jeho předešlé činnosti a tedy velmi zúžené škále informací.

Jak se ukázalo v předešlé studii [17], filter bubble je velmi individuální a její efekt není u všech uživatelů zcela totožný. Míra filter bubble nezávisí pouze na

¹Tím rozumějmě kliknutí, sdílení, komentování a obdobné činnosti.

obsahu viděných příspěvků, jak by se mohlo předpokládat, avšak na zdroji odkud informace čerpá. V důsledku můžeme pozorovat značně silnější efekt informační bubliny u uživatelů s mnoha konexemi na jiné uživatele, než u těch s menším rozsahem jejich spojení. Vycházíme-li z reálného prostředí, všímáme si, že jedinci s mnoha známostmi mají silnější postavení ve společnosti a tudíž i notný vliv na ostatní členy dané společnosti. Stejně tak je tomu na sociálních sítích. V situaci, kdy uživatel s mnoha konexemi na ostatní projeví svůj názor příspěvkem, či komentářem, je tu daleko větší pravděpodobnost ovlivnění značného množství uživatelů s ním spojených.

Informační bublina je velmi rozšířeným problémem, pokud ji si je však uživatel vědom, není filter bubble nepřekonatelnou bariérou v získávání relevantních informací.

1.2 Problémy filter bubble

Vezmeme-li v úvahu, kolik času lidé tráví na sociálních sítích, je zřejmé, že jejich názory a postoje se primárně vytvářejí zrovna zde [17, 16, 21]. Sleduje-li uživatel pouze názorově shodné příspěvky, mohla by informační bublina představovat značnou hrozbu demokratickým systémům, neboť uživatelům předkládá již vyfiltrované příspěvky a to zejména takové, jež by podpořili názor uživatele samotného, nikoli názor odlišný.

V jedné z předešlých studií [16], byla provedena řada experimentů odhalujících několik zajímavých efektů informační bubliny. Jako mnoho podobných studií však své výsledky vyvozuje z reakcí malého počtu lidí² v uměle vytvořených situacích. Ukazují například, že přijímání názorů z různých úhlů pohledů závisí především na hloubce zájmu o dané téma. Čím vyšší je zájem o téma, tím vyšší je ochota přijímat protiargumenty, a naopak, čím nižší je zájem o téma, tím nižší je ochota přijímat protiargumenty.

V případě, kdy jedinec s nízkým povědomím o daném tématu se dostane ke zdroji informací a není motivován hlubším podnětem, jako jsou například blížící se volby, nevykazuje zájem o hledání relevantních faktů, nýbrž dává přednost vyhledávání *users opinion*, nehledě na zdroje, o které se *users opinion* opírá. Předložíme-li tedy současně jedinci informace podobné jeho již dříve získaným postojům a informace lišící se od jeho postoje, ve většině případů si vybere informace podobající se jeho stanovisku, avšak dostane-li se jedinec s nízkým povědomím do kritické situace³ a je motivován se v daném tématu vzdělávat, začíná vyhledávat informace podložené fakty, ať už podporující jeho stanovisko, či nikoliv.

Takovéto chování stále se opakujícího výběru již známého obsahu může jedince vést do *echo chamber*, kde tímto selektivním ziskem podobně motivovaných informací se utvrzuje ve svém původním názoru a naopak informace odlišné, či opačné vytěsňuje do takové míry, že není schopen jejich dalšího vnímání.

Informační bublina může zapříčinit nemalé následky v makroskopickém měřítku na celou dnešní společnost v mnoha ohledech. V první řadě upozorníme na zjevné riziko, které se naskytuje při aktivních preferenčních algoritmech mezi uživateli

²V konkrétním případě této práce okolo 30 jedinců.

³Kupříkladu je-li jedinec postižen nějakou nemocí je mnohem více motivován vyhledávat informace.

sociálních sítí v demokratických společnostech. Zde i přes nabízenou diversitu obsahu uživatel opět vidí jen omezenou část. Autoři v [12] vychází z konceptu, kde demokracie jako taková je rozdělena na liberální, deliberativní, republikánskou a agonistickou a pozorují, že každá z nich je ohrožena v jiné části její struktury. Problémy způsobené informační bublinou v liberální demokracii, jakožto ztráta povědomí občanů o různorodosti volby a nezávislosti médií, coby primárního zdroje informací občanů zvrhne upadajícího do rukou úzkého okruhu lidí, a demokracie deliberativní, kde sledujeme nedostatky v rovnocenné občanské diskusi, klesající toleranci vůči odlišným názorům a úbytek obecného přání zisku nových epistemických argumentů, částečně řeší již popsané aplikace jako třeba *Balancer*, *Scoopinion*, *ConsiderIt*, *Opinion space* a další. Žádná však neřeší ohrožení v typech republikánské a agonistické demokracie.

Otázkou proto zůstává, jak dostatečně rozpoznat míru filter bubble a ochránit rozhled uživatele sociálních sítí bez ohledu na charakter demokracie, ve které se vyskytuje. Představme si živé předvolební období, kdy politické strany vytáhnou do boje a nebojí se použít žádných prostředků k potupě politických rivalů, kdy jedna aféra stíhá druhou, a na povrch vyplouvají rozličné skandály představitelů politických stran. Zároveň jsou také vypouštěny různé výstižné slogany rádobý řešící lokální i globální problémy. Čím více zaujatý slogan, tím masovější ovace. Začíná davové šílenství v podobě obrovských internetových diskusí v tématech, jež jsou pro uživatele klíčová. Vyhledáváním a připojováním se ke společenství se jedinec cítí být více informovaný, nicméně ztrácí přehled o celém tématu a zaměřuje se na čím dál menší okruh informací podporující jeho názor. Kamkoli se podívá a cokoli i přečte je uspokojen, vidí stále příspěvky podobné jeho názorům. Jak je dobře známo, uživatelé těchto internetových diskusí jsou často svým přesvědčením uchvázeni natolik, že ztrácí veškeré zábrany racionálně smýšlející osoby a své zaujaté názory se nebojí ukazovat široké veřejnosti [11]. Je-li však náhled na téma již ze začátku extremistický, kam až může zajít? Co když tyto podporované politické strany proklamující se všeobecnými předsudky nejsou vhodnou volbou pro stát, ale díky svým hojně sdíleným příspěvkům oslovují více a více lidí, kteří dále šíří ideologii? Informační bubliny se z tohoto hlediska stávají problémem, i co se týče jejich etického vlivu na společnost.

1.3 Výhody filter bubble

Zkoumáme-li filter bubble jako celek a přestaneme-li se zaměřovat pouze na hrozby, které s sebou přináší, spatřujeme i pozitivní dopady. Kladné využití informační bubliny je zřejmé, zajímáme-li se o nějaké téma do hloubky a nehrozí ztráta objektivnosti jedince.

Vezměme si kupříkladu sportovního fanouška zaměřeného pouze na tenis. Neuvěřitelné množství existujících sportů a ještě větší počet příspěvků o nich dennodenně zaplavuje sociální sítě. Z tohoto jasně vyplývá, že kdyby fanoušek tenisu sledoval příspěvky ze všech odvětví sportů, ty o tenise by se mu zobrazovaly výrazně méně, tudíž by teoreticky mohlo dojít k omezení informovanosti ve středu zájmu daného fanouška. V takovém případě by nemusel zaregistrovat všechny výsledky posledních utkání, zprávy o kondici sportovců, či změnu času konání další akce, což by pro fanouška mohlo být rozhodující.

Stejně tak můžeme pozorovat pozitivní důsledky informační bubliny v podni-

katelské sféře. Například takový vlastník restaurací potřebuje být plně informován o všech novinkách ve svém oboru vedoucích k vylepšení pracovních postupů. Nemálo může filter bubble pomoci, co se konkurenceschopnosti podnikatele týče. V situaci, kdy bude podrobně znát rozsah a celkovou nabídku služeb konkurentů, má mnohem větší šance uspět na trhu práce, což by se s primárně zobrazovanými příspěvky o výsledcích šachových turnajů, nebo průměrné spotřeby uhlí na jednu domácnost v jeho kanále informací rozhodně nedokázal tak snadno.

Jak vidno, informační bublina, pomineme-li negativní dopady na společnost popsané v předchozí podkapitole, není pouze hrozbou, ale může nám v mnoha ohledech usnadnit orientaci na sociálních sítích, ať jsme již fanoušky sportu, podnikateli, vědci, či kulturními nadšenci.

1.4 Preferenční algoritmy

Jak jsme již výše zmínili, personalizace obsahů který vidíme na internetu může být jak velkým problémem, tak velkou výhodou. Čemu jsme však dosud nevěnovali pozornost je jak informační bubliny vznikají. Je zřejmé, že *Facebook*, *Twitter*, *YouTube* a podobné internetové giganty shromažďují velké množství dat o našich internetových aktivitách. Méně jasné je, že to nedělají kvůli zlomyslným plánům na ovládnutí světa, nýbrž kvůli snaze zpříjemnit užívání jejich služeb⁴. Tato data poté pomocí moderních matematických a statistických metod užívají například k výběru obsahu, který nám bude co nejvíce imponovat, respektive k výběru obsahu, o kterém jsme se již dříve výjádřili, že je pro nás zajímavý.

S obdobným přístupem se můžeme setkat při online nakupování [15], kde jsou nám doporučovány produkty obdobné těm, které jsme v poslední době hledali. Stejně tak například na *YouTube* [14], kde se dříve dostaneme k videím s podobným obsahem, jaký často sledujeme.

Metody, které se užívají pro sociální sítě a informační kanály [13] jsou většinou velmi sofistikované a opírají se o hluboké znalosti *machine learningu*, *statistiky* a *data miningu*. Velmi zjednodušeně řečeno, nový uživatel určité stránky je nejprve vystaven velmi širokému spektru informací. Někde na serveru provozovatele sítě sedí malá ne příliš chytrá umělá inteligence, která si zapisuje na co uživatel kliká⁵. Své zápisky následně zpracovává a pomocí těchto pracovaných poznámek následně odhaduje co by se danému uživateli mohlo líbit.

⁴Tím, že nám usnadní a zpříjemní jejich užívání si zajistí větší návštěvnost, což je zdrojem jejich příjmů.

⁵Samozřejmě také sleduje další způsoby hodnocení příspěvků, které jsou pro různé stránky odlišné. Na *Facebooku* například *like*, na *Twitteru* *retweet*

2 Metody sběru dat

S pokrokem v oblasti technologie probíhá i obrovský pokrok v oblasti čerpání zpráv a informací. Narozdíl od minulosti, drtivá většina obyvatel vyspělých států má přístup k novinkám v průběhu celého dne. Denní počet příspěvků na *Twitteru* dosahoval v roce 2010 neskutečných 6 milionů [22]. Je zřejmé, že není možné používat stejné postupy pro analýzu takto rozsáhlého množství dat jako dříve. Naštěstí se v již nějakou dobu významným tempem posouvá i oblast *Big Data*, *Machine Learningu* a jejich podoblastí. Tyto obory nacházejí uplatnění téměř ve všech oblastech dnešní vědy [26] a stále větší pozornosti se jim dostává i v sociologii [23, 24, 25].

2.1 Twitter

Twitter je sociální síť s strukturou přátelských vztahů odlišnou od *Facebooku*. Uživatel si může navolit téměř libovolné množství lidí, jejichž příspěvku se mu budou zobrazovat. To však neznamená, že lidem které si navolil se budou zobrazovat jeho příspěvky.

Pro jednoduchost uveďme příklad. Pepa je nový uživatel *Twitteru*. Jeho kamarád ze školy se jmenuje Jiří a politik, kterého viděl včera ve zprávách a zamlouval se mu, má jméno Alois. Pepa začne *sledovat*⁶ jak Jiřího, tak Aloise. Od té chvíle uvidí všechny příspěvky, jak od spolužáka Jiřího, tak od politika Aloise. Za nějaký čas začne i Jiří *sledovat* Pepu. Od té chvíle i Jiří uvidí všechny Pepovi příspěvky. Politik Alois však Pepu nezná, proto ho *sledovat* nezačne, a tedy neuvidí příspěvky, které Pepa sdílí.

Takovouto strukturu můžeme chápat jako orientovaný graf, kde vrcholy představují uživatelé *Twitteru* a orientované hrany mezi nimi proud příspěvků, tedy vztah *following*, neboli *sledující*. Pokud uživatele sleduje mnoho lidí, je směrem k němu připojeno více hran, mluvíme o uživateli s vysokou konektivitou nebo s vysokým stupněm. Obdobně, uživatel, který má malou konektivitu je takový, k němuž směřuje málo hran, je málo *sledován*.

Příspěvky na *Twitteru* mohou mít jakýkoliv textový tvar do délek 140 znaků, doplněný obrázkem nebo videem, mnohdy také hypertextovým odkazem na externí článek. Obrázky, videa, ani články na jiných stránkách nejsme schopni analyzovat. Plně však zkoumáme textovou část příspěvků. Ty bývají mnohdy doplněné o tzv. *hashtagy*. To jsou slova nebo krátká spojení slov, před kterými stojí znak # (*křížek*). Z pravidla se jedná o slova, která jsou velmi aktuální, například různé reference na současné politické dění a podobně. Mnohdy jsou *hashtagy* užívány pouze pro zvětšení popularity příspěvku, jelikož častá forma vyhledávání je právě pomocí *hashtagů*.

2.2 Data z Twitteru

Právě data z *Twitteru* jsou pro nás velmi vhodná. Jak se ukazuje [20], narozdíl od *Facebooku*, *Twitter* užívá mnoho lidí jako zdroj informací a zpráv o aktuálním dění ve světě. Samozřejmě můžeme zpochybňovat validitu a přesnost informací,

⁶Jde přímo o pojem *following* z této sociální sítě, něco jako *přátelství* na Facebooku.

které se na takovýchto sítích objevují. Ať už jsou takovéto obavy oprávněné, nebo ne, vřzkum sociologických jevů, který v této práci provádíme to nijak neovlivňuje.

Dalším důležitým důvodem, který vedl k výběru Twitteru jako média, ze kterého budeme stahovat informace je snadná přístupnost k datům pomocí služby API poskytovaný právě přímo Twitterem [28]. Konkrétně pro naše účely jsme tuto službu nepoužívali přímo, ale za pomoci balíčku *tweepy* [27] pro programovací jazyk *python*. Ten umožňuje velmi snadné ovládání a filtrování proudu dat, které si vyžádáme z Twitteru a také jejich okamžitou analýzu a zpracování v *pythonu*.

2.3 Sentimentální analýza textu

Jednou z velice rozvinutých oblastí machine learningu, je tzv. *sentimentální analýza textu*, což je forma zpracování přirozeného jazyka⁷ [10]. V základní a nejvíce studované verzi tohoto problému se snažíme naučit počítač odhadnout, jestli je věta, *positivní*, či *negativní*. Existují samozřejmě i obdobné úlohy. Můžeme se zajímat o rozdíl mezi větami psanými *objektivně* a *subjektivně*, nebo rozlišovat více než dvě kategorie, například rozlišovat texty napsané *rozčilleně*, *smutně*, *radostně* a *překvapeně*. Záleží pouze na účelu, za kterým problém řešíme a datech, které máme k dispozici.

Sentimentální analýza textu se skládá z několika základních kroků, které jsou v obecném měřítku velmi podobné jiným machine learning algoritmům. Základem jsou data. V našem konkrétním případě se jedná přibližně o 1.5 miliónů tweetů⁸ v anglické jazyce, které jsou označeny lidmi jako pozitivní, nebo negativní. Algoritmus, který používáme je neuronová síť [4] obohacená o takzvané *word embedding* vrstvy a *convolution* vrstvy. Takovýto model je nejdříve *"natrénován"*⁹ na označených datech¹⁰. Poté je možné ho snadno a s mnohem nižší časovou komplexitou, než byla zapotřebí při tréninku, používat pro predikci sentimentu z textu. Výstupem takového modelu je pravděpodobnost, že předložený text je pozitivní, tedy reálné číslo mezi 0.0 a 1.0. Takové číslo můžeme chápat jako míru positivity sentimentu textu. Tedy je-li tweet ohodnocen číslem 0.0, je zřejmé, že je negativní. Je-li ohodnocen číslem 0.5, chápeme ho jako neutrální. Text ohodnocený číslem 1.0 je jistě pozitivní.

Pokud řekneme, že při měření bylo zjištěno, že za určitou dobu bylo v určité skupině 40 % lidí proti a 60 % pro, je tím myšleno, že 40 % tweetů mělo sentiment menší než 0.5 a naopak 60 % větší než 0.5.

2.4 Technické detaily sentimentální analýzy textu

Pohlédneme-li do větších detailů našeho modelu na odhad sentimentu, prvním krokem po získání datasetu je vytvořit slovník s $V = 5000$ nejčastějšími slovy. Poté je každý datový bod, tedy každý *tweet*, transformován ze seznamu slov do

⁷Pro pojem zpracování přirozeného jazyka budeme užívat převážně zkratku NLP z anglického názvu *Natural Language Processing*

⁸Uspořádaných ve velkém datasetu složeném z několika menších [9, 8].

⁹Tréninkem je v této oblasti většinou myšlena optimalizace vhodné zvolené funkce, která odráží přesnost modelu.

¹⁰Data, u kterých označil jejich sentiment reálný člověk, považujeme je tedy za správná. I to je však poměrně relativní. Necháme-li více lidí ohodnotit ta stejná data, zjistíme, že ani lidé nejsou v hodnocení příliš konzistentní [7]. Pro naše účely je to však naprosto dostačující.

seznamu *1-hot-encoding* vektorů¹¹ ze slovníku, které jednotlivá slova reprezentují v prostoru o obrovské dimenzi $V = 5000$.

Následuje *word embedding* vrstva [6, 5], která vytvoří novou reprezentaci slov. Ta je schopná mnohem lépe zachovávat semantické vlastnosti jednotlivých slov¹². Vektory s dimenzí $V = 5000$ transformuje do vektorů s dimenzí $N = 32$, které se mnohem lépe hodí pro zpracování a sentimentální analýzu, protože lépe reprezentují skutečný význam slov.

Na tuto vrstvu navazuje *convolution* vrstva [3], která je schopná prozkoumat postavení slov ve větě. Následuje *fully connected* vrstva, která slouží k správné interpretaci vlastností odvozených neuronovou sítí, ukončená *relu* aktivační jednotkou. Ta se stará o dodání nelinearity do modelu, což zajišťuje přesnější klasifikaci. Úplně poslední je další plně propojená vrstva zakončená aktivační funkcí *sigmoid*, která se postará o transformaci na pravděpodobnost, pro snazší budoucí interpretaci.

Zdrojový kód pro sentimentální analýzu byl napsán v programovacím jazyce *pythonu*, za užití balíčků *TensorFlow* [2] a *Keras* [1].

¹¹Kde každý vektor je reprezentací jednoho slova. Takové reprezentace mají velmi jednoduchý a zároveň velmi nepraktický tvar. Je-li slovo ve slovníku na k -té pozici, bude jeho *1-hot-encoding* reprezentace vektor plný nul, až na k -tou pozici, kde se bude nacházet číslo 1. Je zřejmé, že tato reprezentace není příliš efektivní.

¹²Na takové reprezentaci slov můžeme sledovat zachování semantických vlastností a dokonce i možnost heuristického užití aritmetiky. Můžeme si všimnout, že například $\text{vec}(\text{France}) - \text{vec}(\text{Paris}) + \text{vec}(\text{Italy}) \approx \text{vec}(\text{Rome})$. Takové vlastnosti jsme při užití *1-hot-encoding* pozorovat nemohli.

3 Konstrukce měření

Sentimentální analýza textu nám umožňuje automaticky ohodnotit mírou sentimentu u velkého množství textových příspěvků. Výsledkem takového ohodnocení je však pouze reálné číslo mezi nulou a jedničkou. Proto odtud nevede žádná přímá cesta k měření síly informační bubliny na částech společnosti.

Uvažujeme-li konkrétního jedince ve společnosti, informační bublina není nijak explicitně závislá na něm samotném. Místo toho je až implicitně závislá na osobách, které tvoří obsah viditelný studovaným jedincem. Proto je zřejmé, že budeme-li chtít pozorovat sílu a efekt informační bubliny na konkrétního jedince, předmětem našeho studia musí být lidé v okolí tohoto jedince. V případě sociálních sítí to musí být konkrétně příspěvky, které tyto lidé tvoří.

Jak jsme již zmiňovali v předešlých kapitolách, pro účely našeho výzkumu jsme využili sociální síť *Twitter*. Ta se v poslední době transformuje spíše do podoby informačního kanálu [20]. Neméně důležitý je fakt, že umožňuje poměrně snadný přístup k datům [27, 28].

3.1 Výběr pozorovaných skupin

Naším cílem je pozorovat efekty a sílu informační bubliny na vybrané skupině ve společnosti. V odstavcích výše jsme však shrnuli proč to není vůbec přímočarý proces.

Abychom vybrali určitou skupinu uživatelů *Twitteru* co nejpřesněji, využijeme existujících stránek, které sdružují příznivce hledané skupiny. Chceme-li například studovat jak jsou informační bublinou postiženi *biochemici*, vybereme je z lidí, kteří sledují *Twitterový* profil *Biochemical Society*. Ten se po našem průzkumu *Twitteru* ukázal jako nejvíce populární účet s tímto tématem, sleduje ho přibližně 15 tis. lidí. Služby poskytované *Twitterem* na stahování dat nás bohužel omezují v počtu zkoumaných uživatelů, proto musíme ze všech 15 tisíc vybrat pouze část. Abychom zajistili náhodnost tohoto procesu, o výběr podmnožiny se stará jedna část softwaru, který jsme pro tyto účely vytvořili.

Nyní máme vybranou skupinu vzorků, na kterých chceme pozorovat efekty informační bubliny. Ta je však závislá na uživateli, kteří vytvářejí obsah viditelný studovanými lidmi. To jsou ti uživatelé, které studované subjekty sledují. Vybrali jsme-li například několik *biochemiků*, na kterých chceme sledovat informační bublinu, musíme nyní studovat obsah, který tvoří uživatelé, které tyto *biochemici* na *Twitteru* sledují.

Pro přehlednost tyto vztahy vyjádříme matematicky. Je-li \mathbb{M}_{bio} množina všech uživatelů, kteří sledují stránku *Biochemical Society*, našim prvním úkolem je náhodně vybrat podmnožinu \mathbb{P}_{bio} takovou, že obsahuje námi zvolený počet prvků $|\mathbb{P}_{bio}| = n$ a zároveň každý prvek je vybrán s pravděpodobností $p = \frac{1}{|\mathbb{M}_{bio}|}$. Nyní pomocí služeb poskytovaných *Twitterem* vybereme množinu \mathbb{S}_{bio} takovou, že každý prvek je sledován alespoň jedním prvkem z \mathbb{P}_{bio} . Je-li \mathbb{T} množina všech uživatelů *Twitteru* a \mathbb{A} tzv. *adjacency matrix* celého *Twitteru*, tedy matice, pro kterou platí:

$$\mathbb{A}_{ij} = \begin{cases} 1, & \text{pokud } j \text{ sleduje } i; \\ 0, & \text{jinak.} \end{cases} \quad (3.1)$$

Potom $\mathbb{S}_{bio} = \{k \text{ takových, že } \exists l \in \mathbb{T} : \mathbb{A}_{kl} = 1 \wedge l \in \mathbb{P}_{bio}\}$.

3.2 Pozorování efektů informační bubliny

Na takto vybrané skupině lidí S_{bio} , které sledují uživatelé jejichž informační bublina nás zajímá chceme pozorovat, zda jsou zaměřené pouze jedním směrem a nepodávají dostatečně široké spektrum názorů. Umíme však pouze odhadnout sentiment jakéhokoliv *tweetu*.

Po dobu měření budeme stahovat všechny tweety, které vytvoří uživatelé z S_{bio} , tedy lidé, které sledují subjekty našeho zájmu. Zajímají-li nás efekty filter bubble na *biochemiky* budeme stahovat všechny tweety, které vytvoří lidé, které *biochemici* sledují¹³.

Ze všech takto stažených tweetů vybereme pouze ty, které obsahují nějaké, námi zvolené, klíčové slovo. Mohli bychom použít například slovo „*Trump*“. Tím dostaneme pouze tweety, které se zabývají něčím se vztahem k prezidentu *Trumpovi*. Pro každý takto vyfiltrovaný tweet provedeme sentimentální analýzu a tím odhadneme, zda má autor pozitivní nebo negativní vztah k danému tématu. Provedeme-li takováto měření na dvou odlišných skupinách a stejném klíčovém slově, tedy konkrétním tématu, můžeme odhadovat kde působí informační bublina, která podporuje dané téma, nebo ho naprosto zamítá. Například můžeme porovnávat rozložení sentimentu zpráv s klíčovým slovem „*Trump*“, které vidí *biochemici* a příznivci *těžby ropy*. Můžeme předpokládat, že *biochemici* budou žít v bublině, která má více negativní názor na prezidenta *Trumpa* a příznivci *těžby ropy* naopak.

Člověku se může zdát, že tímto pouze pozorujeme názor, který mají *biochemici* na *Trumpa*. Tak to ale zdaleka není. Ve skutečnosti zkoumáme jaké zprávy jsou viditelné *biochemikům*. A právě zprávy nimi viditelné jsou zdrojem informační bubliny, kterou chceme popisovat.

Zde je potřeba podotknout, že skupiny a klíčová slova se musí volit velmi pečlivě. Vybereme-li klíčové slovo „*Islámský stát*“ a pozorujeme sentiment, může dojít k velkému nedorozumění. Tweet, který hovoří o tom zvěstvech, která koná *Islámský stát* bude jistě negativní, což je ten efekt, který si přejeme. Problém však nastává v případě, kdy nějaký tweet hovoří například o vítězství nad *Islámským státem*. Text takového příspěvku se snadno může jevit jako pozitivní¹⁴, což bychom špatně vyhodnotili jako tweet podporující *Islámský stát*. Proto je potřeba jak pozorované skupiny, tak klíčová slova volit velmi pečlivě.

¹³Ve výsledku tedy budeme stahovat všechny tweety, které *biochemici* uvidí, což je právě náš záměr.

¹⁴To však není problémem počítače, který text analyzuje. I člověk by radostný tweet o vítězství označil jako pozitivní.

4 Závěr

Reference

- [1] CHOLLET, François. *Keras*. GitHub 2015 [online]. [cit. 2017-01-22]. Dostupné také z: <https://github.com/fchollet/keras>
- [2] ABADI, Martín, et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. (2015). arXiv:1603.04467. Dostupné také z: <https://arxiv.org/abs/1603.04467>
- [3] YOON, Kim. Convolutional Neural Networks for Sentence Classification. *ArXiv: Computation and Language*. arXiv, 2014. arXiv:1408.5882v2. Dostupné také z: <https://arxiv.org/abs/1408.5882>
- [4] SEVERYN, Aliaksei a Alessandro MOSCHITTI. Twitter Sentiment Analysis with Deep Convolutional Neural Networks. *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '15*. New York, New York, USA: ACM Press, 2015, **2015**, 959-962. DOI: 10.1145/2766462.2767830. ISBN 9781450336215. Dostupné také z: <http://dl.acm.org/citation.cfm?doid=2766462.2767830>
- [5] GOLDBERG, Yoav a Omer LEVY. Word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method. *ArXiv: Computation and Language*. 2014. arxiv:1402.3722. Dostupné také z: <https://arxiv.org/abs/1402.3722>
- [6] MIKOLOV, Tomas, Ilya SUTSKEVER, Kai CHEN, Greg CORRADO a Dean JEFF. Distributed Representations of Words and Phrases and their Compositionality. *Advances in Neural Information Processing Systems*. Curran Associates, 2013, (26), 3111-3119. Dostupné také z: <https://arxiv.org/abs/1310.4546>
- [7] PANG, Bo, Lillian LEE a Shivakumar VAITHYANATHAN. Thumbs up?: sentiment classification using machine learning techniques. *Proceedings of the ACL-02 conference on Empirical methods in natural language processing - EMNLP '02*. Morristown, NJ, USA: Association for Computational Linguistics, 2002, **2002**(10), 79-86. DOI: 10.3115/1118693.1118704. Dostupné také z: <http://portal.acm.org/citation.cfm?doid=1118693.1118704>
- [8] *UMICH SI650: Sentiment Classification*. University of Michigan SI650, 2011. Dostupné také z: <https://inclass.kaggle.com/c/si650winter11/data>
- [9] SANDERS, Niek. *Twitter Sentiment Corpus*. Sanders Analytics, 2011. Dostupné také z: <http://www.sananalytics.com/lab/twitter-sentiment/>
- [10] BIRD, Steven, Ewan KLEIN a Edward LOPER. *Natural language processing with Python*. Beijing: O'Reilly, 2009. ISBN 978-0-596-51649-9. Dostupné také z: <http://www.nltk.org/book/>
- [11] PAPACHARISSI, Zizi. Democracy online: civility, politeness, and the democratic potential of online political discussion groups. *New Media*. 2004, **6**(2), 259-283. DOI: 10.1177/1461444804041444. ISSN 1461-4448. Dostupné také z: <http://journals.sagepub.com/doi/10.1177/1461444804041444>

- [12] BOZDAG, Engin a Jeroen VAN DEN HOVEN. Breaking the filter bubble: democracy and design. *Ethics and Information Technology*. 2015, **17**(4), 249-265. DOI: 10.1007/s10676-015-9380-y. ISSN 1388-1957. Dostupné také z: <http://link.springer.com/10.1007/s10676-015-9380-y>
- [13] KIM, Younghoon a Kyuseok SHIM. TWITOBİ: A Recommendation System for Twitter Using Probabilistic Modeling. *2011 IEEE 11th International Conference on Data Mining*. IEEE, 2011, , 340-349. DOI: 10.1109/ICDM.2011.150. ISBN 978-1-4577-2075-8. Dostupné také z: <http://ieeexplore.ieee.org/document/6137238/>
- [14] DAVIDSON, James, Blake LIVINGSTON, Dasarathi SAMPATH, et al. The YouTube video recommendation system. *Proceedings of the fourth ACM conference on Recommender systems - RecSys '10*. New York, New York, USA: ACM Press, 2010, , 293-. DOI: 10.1145/1864708.1864770. ISBN 9781605589060. Dostupné také z: <http://portal.acm.org/citation.cfm?doid=1864708.1864770>
- [15] LINDEN, G., B. SMITH a J. YORK. Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*. 2003, **7**(1), 76-80. DOI: 10.1109/MIC.2003.1167344. ISSN 1089-7801. Dostupné také z: <http://ieeexplore.ieee.org/document/1167344/>
- [16] LIAO, Q. Vera a Wai-Tat FU. Beyond the filter bubble: Interactive effects of perceived threat and topic involvement on selective exposure to information. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. New York, New York, USA: ACM Press, 2013, 2359-2368. DOI: 10.1145/2470654.2481326. ISBN 9781450318990. Dostupné také z: <http://dl.acm.org/citation.cfm?doid=2470654.2481326>
- [17] GOTTRON, Thomas a Felix SCHWAGEREIT. The Impact of the Filter Bubble – A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities. *ARXIV: Computer Science - Social and Information Networks*. 2016, **2016**. 2016arXiv161206551G. Dostupné také z: <https://arxiv.org/abs/1612.06551#>
- [18] PARISER, Eli. (2011). *Beware online "filter bubbles"* [online]. Dostupné z https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles
- [19] PARISER, Eli. *The filter bubble: what the Internet is hiding from you*. New York: Penguin Press, 2011. ISBN 15-942-0300-8.
- [20] BARTHEL, MICHAEL, ELISA SHEARER, JEFFREY GOTTFRIED a AMY MITCHELL. *The Evolving Role of News on Twitter and Facebook* [online]. In: . Pew Research Center, 2015 [cit. 2017-01-22]. Dostupné z: <http://www.journalism.org/2015/07/14/the-evolving-role-of-news-on-twitter-and-facebook/>
- [21] Twitter and the News: How people use the social network to learn about the world. *American Press Institute: Insights, tools and research*

to advance journalism [online]. 2015, **2015** [cit. 2017-01-22]. Dostupné z: <https://www.americanpressinstitute.org/publications/reports/survey-research/how-people-use-twitter-news/>

- [22] MATHIOUDAKIS, Michael a Nick KOUDAS. TwitterMonitor. *Proceedings of the 2010 international conference on Management of data - SIGMOD '10*. New York, New York, USA: ACM Press, 2010, **2010**, 1155-1158. DOI: 10.1145/1807167.1807306. ISBN 9781450300322. Dostupné také z: <http://portal.acm.org/citation.cfm?doid=1807167.1807306>
- [23] TINATI, Ramine, Susan HALFORD, Leslie CARR a Catherine POPE. Big Data: Methodological Challenges and Approaches for Sociological Analysis. *Sociology*. 2014, **48**(4), 663-681. DOI: 10.1177/0038038513511561. ISSN 0038-0385. Dostupné také z: <http://journals.sagepub.com/doi/10.1177/0038038513511561>
- [24] MCFARLAND, Daniel A., Kevin LEWIS a Amir GOLDBERG. Sociology in the Era of Big Data: The Ascent of Forensic Social Science. *The American Sociologist*. 2016, **47**(1), 12-35. DOI: 10.1007/s12108-015-9291-8. ISSN 0003-1232. Dostupné také z: <http://link.springer.com/10.1007/s12108-015-9291-8>
- [25] SHAH, D. V., J. N. CAPPELLA a W. R. NEUMAN. Big Data, Digital Media, and Computational Social Science: Possibilities and Perils. *The ANNALS of the American Academy of Political and Social Science*. 2015, **659**(1), 6-13. DOI: 10.1177/0002716215572084. ISSN 0002-7162. Dostupné také z: <http://ann.sagepub.com/cgi/doi/10.1177/0002716215572084>
- [26] HUBERMAN, Bernardo A. Sociology of science: Big data deserve a bigger audience. *Nature*. 2012-2-15, **482**(7385), 308-308. DOI: 10.1038/482308d. ISSN 0028-0836. Dostupné také z: <http://www.nature.com/doifinder/10.1038/482308d>
- [27] HILL, Aaron a Joshua ROESSLEIN. *Tweepy*. Github 2015 [online]. [cit. 2017-01-22]. Dostupné z: <https://github.com/tweepy/tweepy>
- [28] API Overview. *Twitter Developer Documentation* [online]. Twitter, 2016 [cit. 2017-01-22]. Dostupné z: <https://dev.twitter.com/overview/api>

Přílohy

První kapitola příloh