

Univerzita Hradec Králové
Fakulta informatiky a managementu
Katedra ekonomie

Informace a znalosti k podpoře firemní externí komunikace
na sociálních sítích

Disertační práce

Autor: Ing. Jan Hruška

Studijní program: Systémové inženýrství a informatika

Studijní obor: Informační a znalostní management

Školitel: doc. Ing. Mgr. Petra Marešová, Ph.D.

Katedra/pracoviště školitele: Katedra ekonomie

Prohlášení

Prohlašuji, že jsem disertační práci zpracoval/a samostatně a s použitím uvedené literatury a pramenů.

Hradec Králové, říjen 2020

Ing. Jan Hruška

Poděkování

Rád bych vyjádřil upřímnou vděčnost mé školitelce doc. Ing. Mgr. Petře Marešové, Ph.D., za její trpělivý, podporující, pozitivní a laskavý přístup a také za její cenné rady, které mi poskytovala během celého procesu výzkumu.

Také bych rád poděkoval své rodině za veškerou podporu.

Abstrakt

Během poslední dekády se sociální sítě vyvinuly v důležitou komponentu k zajištění šíření informací napříč různorodými oblastmi, jako jsou podnikání, zábava, krizový management nebo politika. Správným zpracováním a distribucí informací mohou podniky získat konkurenční výhodu, lépe uspokojovat zákazníky a snadněji řídit podnik. Cílem práce je navrhnout model využití informací a znalostí k podpoře externí komunikace firem na sociálních sítích. Pozornost byla mimo jiné zaměřena i na obsah a strukturu prezentovaných příspěvků. Kromě analýzy sentimentu jsou v práci použité metody standardizované PRISMA, systematická rešerše, statistické metody ANOVA s Bonferroni post hoc testy a Pearsonovou korelací. Neuronových sítí bylo využito pro testy více typů, ale hlavně byl využit typ multi-layer perceptron. Výzkum probíhal hlavně na sociálních sítích Facebook a Twitter s podporou softwarů WEKA, IBM SPSS Statistics a PyCharm k nastavení neuronových sítí a programování. Navržený model přispěje k ulehčení rozhodování podniků na sociálních sítích v ohledu, jak strukturovat svoje příspěvky, jak volit a připravovat jejich obsah a jaká kritéria dodržovat, aby se zvýšil dosah mezi uživatele a a docházelo ke zvyšování množství sdílení.

Klíčová slova: sociální média, podnik, informace, informační management

Title: Information and knowledge to support corporate external communication on social networks

Abstract

Social media has evolved over the last decade to become an important driver for acquiring and spreading information in different domains such as business, entertainment, crisis management and politics. By properly processing and distributing information, businesses can gain a competitive edge, better serve their customers, and manage the business more easily. The aim of the thesis is therefore to propose a model of using information and knowledge from social networks to support external communication in the company. The goal is to make external communication easier and more effective, especially from the point of view of the content and structure of the posts added. Besides sentiment analysis methods used are standardized PRISMA systematic review, statistic methods ANOVA with Bonferroni post hoc tests and Pearson correlation.

More types of neural networks will be used but most used network will be multi-layer perceptron. Most studied social media will be Facebook and Twitter with help of software WEKA, IBM SPSS Statistics and PyCharm. The proposed model will help to make social networking decisions easier for companies to structure their posts and what content will increase user reach and sharing.

Keywords: social media, business, information, information management

Obsah

1	Úvod	1
2	Cíl práce a metody	3
	2.1 Cíl práce	3
	2.2 Metodika práce.....	4
	2.2.1 Přístupy pro získávání a prezentaci poznatků	5
	2.2.2 Metody pro analýzu obsahu příspěvku.....	5
	2.2.3 Metody pro statistické zpracování dat.....	9
	2.2.4 Využitý software	9
3	Teoretická východiska a analýza současného stavu poznání	11
	3.1 Základní pojmy	11
	3.1.1 Data, informace a znalosti.....	11
	3.1.2 Učící se znalostní organizace nebo systém	12
	3.1.3 Sociální média a sociální sítě	14
	3.1.4 Podnik.....	20
	3.2 Komunikace	23
	3.2.1 Interní komunikace.....	24
	3.2.2 Externí komunikace.....	25
	3.2.3 Public relations.....	26
	3.2.4 Relevantní oblasti pro podnikovou externí komunikaci	27
	3.3 Vybrané sociální sítě pro externí komunikaci.....	31
	3.3.1 Obchodní model na sociálních sítích	32
	3.4 Využití sociálních sítí jako nástroje externí komunikace firem – současný stav poznání	34
4	Využití neuronových sítí v rámci sociálních médií – systematická rešerše	38
	4.1 Vyhledávací strategie a kritéria.....	38
	4.2 Hlavní zjištění – využívané metody v rámci neuronových sítí.....	40
	4.3 Metody neuronových sítí využívané v rámci analýzy sentimentu	46
5	Analýza struktury příspěvků.....	50
	5.1 Prvotní analýza proměnných struktury	50
	5.2 Obsahová analýza a analýza sentimentu	52
	5.3 Statistická významnost proměnných obsahové analýzy a analýzy sentimentu ..	57
	5.4 Analýza proměnných pomocí neuronové sítě MLP	58

5.5	Analýza proměnných pomocí neuronové sítě Keras.....	62
6	Dotazníkové šetření – analýza chování na sociálních sítích a motivace uživatelů pro sdílení.....	73
6.1	Rozdělení repondentů	73
6.2	Využívané sociální sítě	75
6.2.1	Motivace sdílení příspěvků	77
6.3	Souhrn výsledků z dotazníkového šetření.....	78
7	Návrh konceptuálního modelu externí komunikace firmy na sociálních sítích	79
7.1	Východiska modelu.....	79
7.2	Externí komunikace ve firmě a její vazby na vnější prostředí	81
7.3	Externí komunikace a toky informací a znalostí ve firmě	83
7.4	Klíčové prvky v rámci modelu externí komunikace	86
7.5	Vybraný proces metody externí komunikace.....	87
8	Doporučení, přínos a limity práce	95
9	Závěr	97
10	Seznam použité literatury	98
11	Uvedení vlastních publikací disertanta	112
	Seznam příloh	116
	Přílohy	117

Seznam zkratek

ANOVA: Analysis Of Variance

ARMA: Autoregressive Moving Average model

CNN: Convolutional Neural Networks

C-RBFNN: Cloud Radical Basis Function Neural Network

DAN2: Dynamic Architecture for Artificial Neural Networks

DBN: Deep Belief Network

DSS: Decision Support Systems

EACD: European Association of Communication Directors

EUPRERA: European Public Relations Education and Research Association

GNU: General Public License

GRU: Gated Recurrent Unit

LSD: Least Significant Difference

LSTM: Long Short-Term Memory

MLP: Multi-layer Perceptron

NB: Naive Bayes

PR: Public Relations

PRISMA: Preferred Reporting Item for Systematic Reviews and Meta-Analyses

RBF: Radical Basis Function

RNN: Recurrent Neural Network

SVM: Support Vector Machines

WEKA: Waikato Environment for Knowledge Analysis

Seznam obrázků

Obrázek 1: Průběh výzkumu	4
Obrázek 2: Návrh modelu využití neuronových sítí pro zpětnou vazbu.	7
Obrázek 3: Některé z nejlepších softwarů pro budování neuronových sítí podle stránky predictiveanalyticstoday.	8
Obrázek 4: Propojení vztahu dat, informací a znalostí.....	13
Obrázek 5: Interakce podniků se sociálními médii a veřejností	15
Obrázek 6: Klíčové oblasti informačního managementu.	21
Obrázek 7: Informační strategie podniku.	22
Obrázek 8: Obecný průběh komunikace	23
Obrázek 9. Klíčové aktivity externí komunikace na sociálních sítích.	33
Obrázek 10: Postup filtrování článků.	40
Obrázek 11: Zobrazení postavené neuronové sítě v programu NeurophStudio.....	61
Obrázek 12: Zobrazení vyhodnocené neuronové sítě a predikce výstupu.	61
Obrázek 13: Postupné učení pomocí algoritmu v Keras.	65
Obrázek 14: Učení se neuronové sítě na 40 tisíc opakování s optimálním nastavením pro tento problém.....	69
Obrázek 15: Učení se neuronové sítě na 40 tisíc opakování s druhým nejlepším nastavením pro tento problém.....	70
Obrázek 16: Zobrazení rozložení respondentů podle čistého měsíčního příjmu.....	74
Obrázek 17: Zobrazení rozložení respondentů podle počtu obyvatel místa bydliště.	74
Obrázek 18: Zobrazení rozložení používaných sociálních sítí u respondentů.	75
Obrázek 19: Zobrazení hlavních oblastí, na které respondenti reagují.	76
Obrázek 20: Zobrazení hlavních oblastí, které respondenti sdílí.	76
Obrázek 21: Vyobrazení, z jakého zařízení a odkud uživatelé sociálních sítí sdílí.....	77
Obrázek 22: Dopady zavedení nové kvalitní externí komunikace na související proměnné a činnosti firmy.....	82
Obrázek 23: Klíčové entity a propojení firmy, sociálních médií a široké veřejnosti.	84
Obrázek 24: Inicializace nového příspěvku.....	88
Obrázek 25: Proces „výběr tématu“.	90
Obrázek 26: Proces „analýza a výpočty“	91
Obrázek 27: Tvorba vhodné struktury pro vybrané téma.....	93

Seznam tabulek

Tabulka 1: Vazba využitých metod a SW na dílčí cíle.....	10
Tabulka 2: Rozložení zkoumaných článků (v období 2010 až květen 2019).....	39
Tabulka 3: Shrnutí finálního filtru zkoumaných článků.....	43
Tabulka 4: Hlavní využívané metody klasifikace a zkoumaná oblast sociálních sítí.	47
Tabulka 5: Nejúspěšnějších 20 článků podle počtu sdílení.....	53
Tabulka 6: Pořadí nejoblíbenějších kategorií pro všechny natěžené články a pro tisíc nejsdílenějších.	54
Tabulka 7: Porovnání průměru u proměnných pro všechny natěžené články a pro tisíc nejsdílenějších.	55
Tabulka 8: Porovnání proměnných pro tisíc nejméně sdílených článků a tisíc nejvíce sdílených článků.	56
Tabulka 9: Rozložení počtu sdílení během různých měsíců, kdy byl článek přidán.....	58
Tabulka 10: Testování neuronových sítí v programu WEKA.....	59
Tabulka 11: Porovnání přesnosti testů pro model s 27 proměnnými oproti modelu s 16 proměnnými.....	67
Tabulka 12: Optimalizační proces neuronové sítě pomocí Keras.	67
Tabulka 13: Shrnutí východisek pro tvorbu modelu a doporučení.	81
Tabulka 14: Regresní analýza pro všechny proměnné v datovém souboru.	120
Tabulka 15: Zobrazení regresní analýzy jen pro některé vybrané proměnné podle statistické významnosti.....	121

1 Úvod

Současná doba je specifická rychlým rozvojem moderních technologií, které zasahují do mnoha oblastí podnikových činností. Jednou z nich je i komunikace. Stále stoupající trend využívání sociálních médií umožňuje podnikům nové způsoby prezentace svých služeb, produktů, komunikace s cílovými skupinami a budování vztahů s veřejností. Sociální sítě fungují na internetu jako hlavní komunikační kanál, kde lidé hledají informace, hodnotí, sledují své oblíbené značky a sdílejí své názory i pocity. Sociální sítě proto v případě efektivního využití informací, které jsou na nich vyměňovány a sdíleny, představují obrovský potenciál pro firmy, umožňující vybudovat si lidské, přátelské a především loajální vztahy se svými zákazníky. Významnou roli v tomto kontextu představuje informační management, „jehož jedním z cílů je zajistit, poskytnout co nejrychleji a nejlépe pracovníkům organizace informaci potřebnou jako základ racionalizace“ (Eroshkin et al., 2017). Klíčové jsou také jeho funkce, zejména pak plánování a řízení informační činnosti, organizace a prezentace poznatků v podnikání a v neposlední řadě i management informačních a komunikačních technologií (Sigala & Chalkiti, 2015a).

Neustálý vývoj sociálních médií a interaktivního chování mezi uživateli často způsobuje, že se obsah vytvořený uživateli rychle šíří po internetu. Popularita sociálních médií a zde dostupný obsah nám umožňují porozumět těmto činnostem, které mají dopad na sociální, ekonomické a vládní činnosti. Modelování a predikce popularity online obsahu je důležitou oblastí průzkumu sociálních médií a může mít pozitivní dopad v mnoha ohledech na veřejnou správu, organizace, podniky a oblasti online bezpečnosti. Například může podporovat řešení krizí tím, že zná dopady přírodních katastrof, podvodů, falešných informací, terorismu nebo zločinu (G. Chen et al., 2019). V podnikání může tato predikce pomoci analyzovat současné trendy a obavy mezi uživateli a poskytnout cenné informace pro modelování rizik a zdůraznit potenciální zisky, tržby nebo návratnost akcií (Cerchiello et al., 2017a).

Využití různých metod pro zvýšení informovanosti podniku může zajistit klíčovou konkurenční výhodu. Mezi současné metody, které umožní z dat získat informace a znalosti, patří analýza současných dostupných materiálů jak od podniků, tak uživatelů pomocí analýzy sentimentu, neuronových sítí, a různé variace statistických metod jako analýza rozptylu či korelace. Takto získávané informace, znalosti a jejich využití a propojení v rámci vědních disciplín informačního a znalostního managementu umožňují firmám například zjistit, co se říká o jejich značce, co si myslí zákazníci a jaký to má dopad na prodeje.

Především umožní doporučit, jak změnit obchodní strategii či marketingovou kampaň tak, aby firma celkově měla efektivní externí komunikaci, která je i předmětem této práce.

Práce je rozdělena do pěti hlavních částí. Nejprve jsou vymezeny klíčové pojmy, využívané v celé práci, druhá kapitola specifikuje cíl, kroky k jeho dosažení a využívané metody. Třetí část popisuje současný stav poznání, zmapovaný na základě systematické rešerše. Čtvrtá kapitola seznamuje s prvotními výsledky zkoumání a na závěr je uvedeno další směřování práce.

2 Cíl práce a metody

Cílem práce je navrhnout konceptuální model využití informací a znalostí k podpoře firemního rozhodování o obsahu a formě externí komunikace na sociálních sítích. Pozornost je mimo jiné zaměřena i na obsah a strukturu prezentovaných příspěvků.

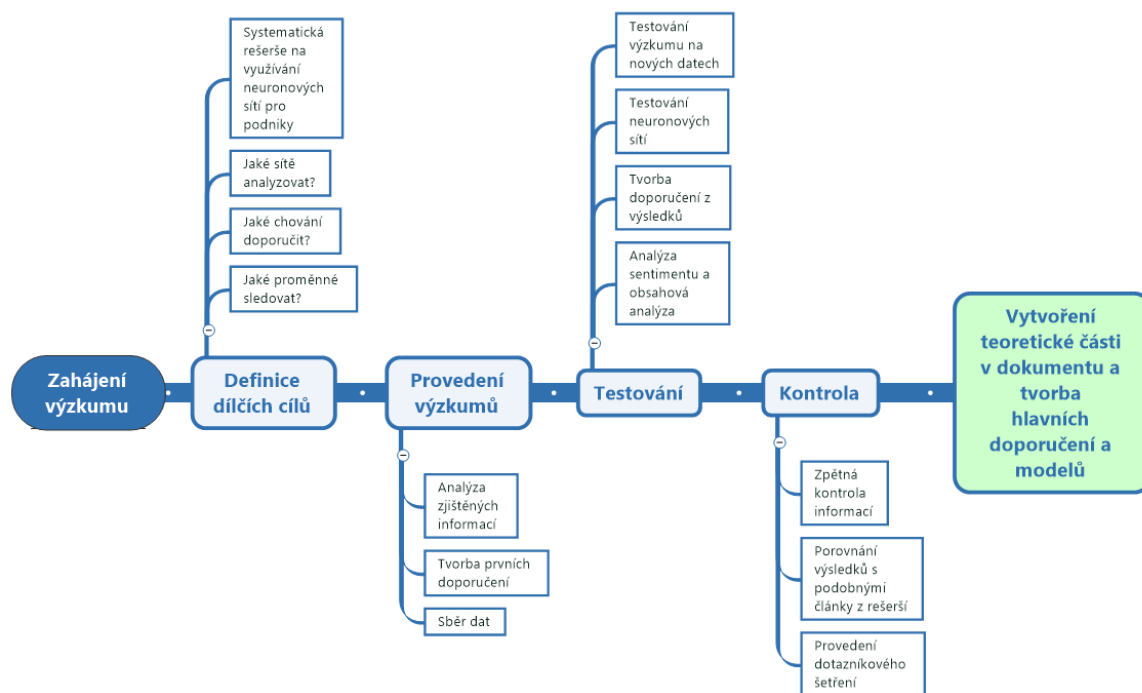
2.1 Cíl práce

Dosahování vytyčeného cíle „*návrhu konceptuálního modelu využití informací a znalostí k podpoře firemního rozhodování o obsahu a formě externí komunikace na sociálních sítích*“ bude probíhat v pěti navazujících dílčích krocích.

Dílčí kroky k dosažení cíle:

- analýza současného stavu poznání v oblasti firemní externí komunikace na sociálních sítích a metod používaných pro zvýšení efektivity komunikace na sociálních sítích,
- využití metod pro analýzu obsahu příspěvku a dotazování pro specifikaci klíčových informací a znalostí k podpoře firemního rozhodování o obsahu a formě externí komunikace,
- specifikace klíčových znalostí a informací ve vazbě na stanovený cíl,
- návrh prototypu konceptuálního modelu pro usnadnění podnikové externí komunikace,
- návrh finální verze konceptuálního modelu.

Syntézou získaných poznatků bude možné následně navrhnout konceptuální model zajišťující větší úspěch při externí komunikaci. Podrobněji lze průběh výzkumu vidět na obrázku níže (obr. 1).



Obrázek 1: Průběh výzkumu (vlastní dílo autora)

Před zahájením výzkumu byly definovány cíle, následně došlo na sběr a analýzu dat, následovalo testování pomocí neuronových sítí, analýza sentimentu, obsahová analýza, dotazníkové šetření a kontrola. Na základě tohoto postupu vznikla další dokumentace, byly vytvořeny hlavní modely a doporučení.

2.2 Metodika práce

Pro dosažení cíle byla zvolena systematická rešerše dle standardu PRISMA, statistická metoda analýza rozptylu (ANOVA), Pearsonova korelace, dále budou využity neuronové sítě multi-layer perceptron, analýza sentimentu, obsahová analýza a dotazníkové šetření s využitím softwarů WEKA, PyCharm, Stella Architect, Enterprise Architect a IBM SPSS Statistics.

2.2.1 Přístupy pro získávání a prezentaci poznatků

Rešerše (PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses)

Standardizovaný postup pro systematické rešerše PRISMA má důkladně zpracovanou dokumentaci, jak lze vytvářet kvalitní systematické rešerše a meta-analýzy. Dokumentace popisuje, jak postupovat při výběru článků pro rešerši v několika krocích, jak je následně filtrovat a na co se zaměřit při zkoumání finálního vzorku článků (Emily Jones, n.d.).

2.2.2 Metody pro analýzu obsahu příspěvku

Analýza sentimentu

Principem této analýzy je odhadování obecných „nálad“ v určité skupině lidí definované společným zájmem nebo činnostmi (Abid et al., 2019a). Analýza sentimentu se snaží o odhadování jejich nálad a z toho vyplývající predikci chování (nákup /doporučení /prodej). Výstupy analýzy sentimentu slouží aktérům na trhu a také samotným subjektům zkoumání – obchodníkům.

Analýza sentimentu se do online marketingu dostala jednoduchou cestou. Není zde možné dopředu přesně odhadovat výsledky vlastních akcí, například masivní reklamní kampaně, protože výsledky závisí na řadě velmi složitých interakcí mezi jednotlivci. Stejně jako na burze i zde existuje velké množství kvantitativních parametrů sloužících k měření stavu věci – ovšem ani zde tyto parametry nestačí, protože prostě nemohou stačit k popisu věci (Alharbi & Doncker, 2019).

Metodika práce obsahuje pozorování a následný sběr dat z různých sociálních médií (hlavně sociální sítě zmíněné v úvodu) pro zkoumanou oblast spolu s natěžením informací naprogramovaným robotem (Crawler) z různých článků online a poté testování pomocí neuronových sítí, jaké faktory/proměnné mají na počet sdílení klíčový vliv. Následovat bude tvorba neuronových sítí z nasbíraných dat a spuštění neuronových sítí v různém nastavení, zjištění a analýza současného stavu, ale i možné cesty vývoje do budoucna.

Testy umožňují vyjádření jevů a vztahů mezi nimi (například vztah mezi počtem klíčových slov na sociálních sítích jako Facebook nebo Twitter a mezi počtem sdílení nebo objednaných a zakoupených výrobků v určité oblasti).

Obsahová analýza

Obsahová analýza je metoda, kterou lze definovat jako rozbor obsahu záznamu určité komunikace. Někdy je obsahová analýza vymežována jako analýza textu či množiny textů. Hlavním účelem této metody je vyhledávání konkrétních slov a konceptů v analyzované komunikaci a stanovení četnosti jejich výskytu, významu, vzájemného vztahu atp. Nejčastěji je metoda využívána při analýze mediálních sdělení (Elo & Kyngäs, 2008).

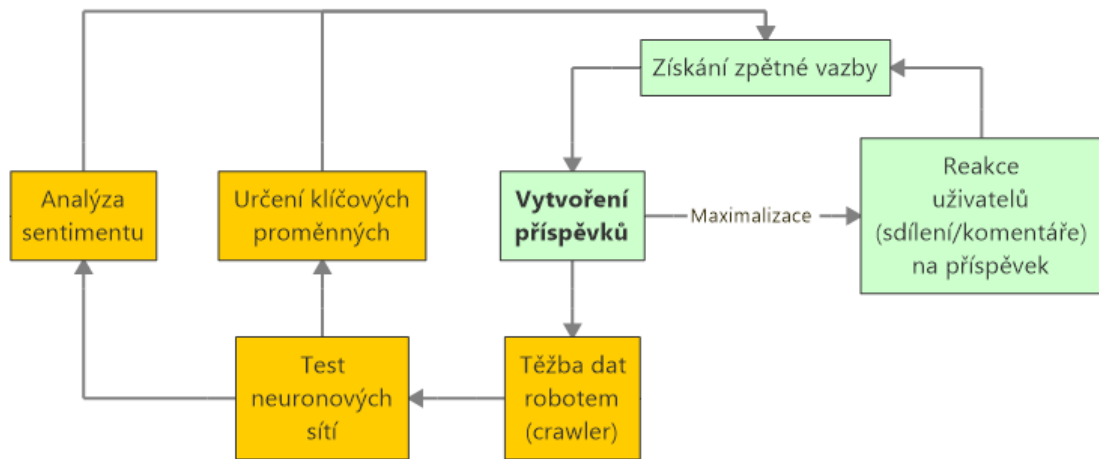
Zkoumané příspěvky (články) budou rozděleny do několika kategorií (podnikání, produkty, služby, zábava, technologie, věda atd.) a následně kvalitativně zkoumány pomocí vybraných metod analýzy textu.

Mezi proměnné, které budou zkoumány pomocí neuronových sítí, patří počet slov a počet znaků v nadpisu, počet slov v článku, počet znaků v článku, průměrná délka slova, počet vložených obrázků, počet vložených odkazů, počet emotikonů, počet pozitivních slov, počet negativních slov, počet charakterů v prvním odstavci, počet slov v prvním odstavci, počet pozitivních slov v prvním odstavci, počet negativních slov v prvním odstavci, poměr pozitivních slov ku celkovému počtu slov, poměr negativních slov ku celkovému počtu slov, počet pozitivních a negativních slov v nadpisu.

Neuronová síť – Multi-layer Perceptron (MLP)

Tato metoda byla vybrána na základě provedené systematické rešerše, kde ji autoři často využívají a uvádějí jako vhodnou pro tuto problematiku. Metoda MLP byla vybrána za účelem zkoumání struktury příspěvků ve smyslu zjistit, do jaké míry struktura článku ovlivňuje množství sdílení na sociálních sítích.

Metoda perceptron a vícevrstvý perceptron je nástroj pro tvorbu neuronových sítí. V případě, že se jedná o jednoduchý perceptron, má model vždy dvě vrstvy, a to vrstvu vstupů a výstupů. Model se následně na vložených datech trénuje a upravuje váhy vstupů. Podle těchto úprav poté zjišťuje, jak velký vliv mají úpravy vah na výstup.



Obrázek 2: Návrh modelu využití neuronových sítí pro zpětnou vazbu (vlastní dílo autora)

V případě, že se jedná o vícevrstvý perceptron, lze si zvolit neomezené množství vrstev a uzlů mezi vstupy a výstupy. Zpravidla se metoda MLP používá v případě, že jsou uzly z každé vrstvy propojené se všemi uzly z okolních vrstev, a čím více vrstev bylo v modelu, tím byla délka učení modelu delší (Kim & Adali, 2002a). Někteří autoři se domnívají, že i pro velmi složité problémy stačí maximálně pouze tři vrstvy mezi vstupy a výstupy (Hampshire & Pearlmutter, 1991a). V rámci této disertační práce bude otestována verze bez skrytých vrstev i varianta s jednou skrytou vrstvou či s více skrytými vrstvami.

Neuronová síť – Keras

Pro získání hlubších informací a podrobnější ladění nastavení neuronových sítí bude v této disertační práci využita volně dostupná knihovna pro Python – Keras. Ačkoliv hluboké neuronové sítě jsou stále velice populární, komplexnost hlavních frameworků zůstává bariérou pro jejich častější využívání vývojáři. Keras patří mezi nejoblíbenější vysokoúrovňové API pro složité neuronové sítě. Je napsán v Pythonu a podporuje back-endová výpočetní jádra neuronových sítí. Keras zdánlivě ostatní aplikační rozhraní hlubokých neuronových sítí porazil: stal se totiž vysokoúrovňovým standardem pro nadcházející TensorFlow 2.0 (Vasilev et al., 2019).

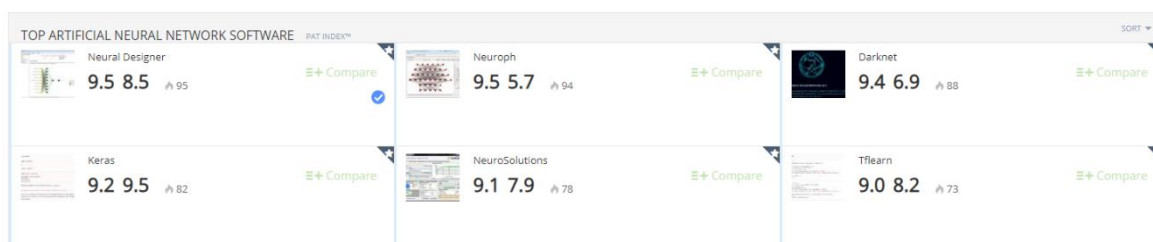
Základním principem vzniku Kerasu byla jednoduchost, modularita, snadná rozšiřitelnost a kompatibilita s Pythonem. Keras vznikl „pro člověka, ne pro stroj a řídí se nejlepšími metodami pro snížení kognitivní zátěže“ (Vasilev et al., 2019). Samozřejmostí je přidávání nových modulů definovaných v Pythonu. Jedním z hlavních důvodů, proč Keras používat, je jeho důraz na to, aby zůstal uživatelsky přívětivý. U této metody lze modifikovat celou řadu proměnných, jedná se o velikost sítě (počet vrstev a neuronů), nastavení zatížení

paměti, počet iterací učení, nastavení optimalizačního algoritmu, rychlost zjištění nepotřebných neuronů nebo nastavení aktivační funkce. Tyto proměnné budou dále v práci popsány a modifikovány, aby bylo dosaženo nejlepšího možného výsledku pro problematiku struktury článků.

Keras je také podporován řadou velkých IT korporací v čele s Googlem, Microsoftem, Amazonem, Applem, Nvidií nebo Uberem.

Dílčím cílem této práce je analyzovat data získaná naprogramovaným algoritmem. Data poskytují informace o článcích na internetu a s jejich pomocí se autor snaží predikovat úspěšnost článku z pohledu počtu sdílení (dostupná informace uvedena přímo u článku). K naprogramování algoritmu byl využit programovací jazyk Python a knihovna pro procházení webových stránek Selenium. Články jsou z internetové stránky mashable.com, která bohužel během tohoto výzkumu zrušila zobrazování počtu sdílení. Existují i další stránky, které uvádějí počet sdílení, ale často jsou zaměřené na jednu konkrétní oblast, jako je zdravotnictví (psychcentral.com) nebo marketing na sociálních sítích (<https://www.socialmediaexaminer.com/mark-zuckerberg-lays-out-future-vision/>). Další vhodnou stránkou pro tento výzkum by byla například thenextweb.com, ale pro zjednodušení byly získány proměnné o článcích jenom ze stránky Mashable.

Ze šesti nejlépe hodnocených softwarů byly vybrány dva pro testy. Jedná se o software NeuropStudio a Keras.



Software	Rating 1	Rating 2	Stars	Compare
Neural Designer	9.5	8.5	95	Compare
Neuroph	9.5	5.7	94	Compare
Darknet	9.4	6.9	88	Compare
Keras	9.2	9.5	82	Compare
NeuroSolutions	9.1	7.9	78	Compare
Tflern	9.0	8.2	73	Compare

Obrázek 3: Některé z nejlepších softwarů pro budování neuronových sítí podle stránky predictiveanalyticstoday

Dotazníkové šetření

Okruhy témat v dotazníku vycházejí z článku „Let’s Hate Together: How People Share News in Messaging, Social, and Public Networks“ z roku 2018 od autorů Danielle Lottridge and Frank R. Bentley. V rámci práce bylo provedeno dotazníkové šetření, které mělo za cíl zjistit motivaci uživatelů na sociálních sítích pro reagování na příspěvky a jejich sdílení. Sběr dat probíhal v období duben až srpen 2020. Šíření dotazníku se uskutečnilo primárně po sociálních sítích. Dotazník zahrnuje jak otázky otevřené, kde má respondent možnost popsat delším textem svůj názor, tak i uzavřené, kde lze vybrat jednu či více z uvedených možností.

2.2.3 Metody pro statistické zpracování dat

Po získání dat z různých příspěvků bude provedena statistická analýza ANOVA a Pearsonova korelace mezi proměnnými.

ANOVA

Analýza rozptylu (anglicky Analysis of variance – ANOVA) je metodou matematické statistiky, která umožňuje ověřit, zda na hodnotu náhodné veličiny pro určitého jedince má statisticky významný vliv hodnota některého znaku. Tento znak musí nabývat jen konečného počtu možných hodnot (nejméně dvou) a slouží k rozdělení jedinců do vzájemně porovnávaných skupin.

Pro zjištění odlišnosti skupin byl použit Bonferroniho, Tukey a LSD post hoc test. Použití Bonferroniho procedury je poměrně konzervativní, což znamená, že je při jejím použití relativně obtížné dosáhnout statistické významnosti.

2.2.4 Využitý software

Software WEKA 3.8

Weka je soubor algoritmů strojového učení pro úlohy dolování dat. Obsahuje nástroje pro přípravu dat, klasifikaci, regresi, clustering, asociační pravidla, dolování a vizualizaci. Weka je open source software vydávaný pod GNU General Public License.

PyCharm Community Edition 2018.3.4

K naprogramování robota (crawler) byl využit software PyCharm a využit programovací jazyk Python s několika rozšířeními pro automatické pohybování po internetu.

Výzkum byl realizován primárně na sociálních sítích Facebook a Twitter z toho důvodu, že Facebook je nejvyužívanější sociální síť na světě a na Twitteru je uzpůsobeno snadnější dolování dat.

IBM SPSS Statistics

IBM SPSS Statistics je komplexní soubor nástrojů určený analytikům a statistickým pracovníkům. Patří mezi celosvětově rozšířené statistické systémy pro aplikaci ve vědě, marketingu, personalistice a výzkumu i sumarizaci dat z velkých i menších databází různého druhu. IBM SPSS Statistics poskytuje softwarové prostředí pro vybrané metody v oblasti Data Miningu (DM), manažerských analýz a podporu business intelligence.

Mindjet MindManager

MindManager je software pro vizualizaci informací, a to zejména ve formě myšlenkových map, konceptuálních map a vývojových diagramů. Tento software patří mezi nejrozšířenější a nejvýkonnější programy pro podporu kreativního myšlení.

Enterprise Architect

Sparx Systems Enterprise Architect je vizuální modelovací a návrhový nástroj založený na OMG UML. Platforma podporuje návrh a konstrukci softwarových systémů; modelování obchodních procesů; modelování průmyslových domén. Podniky a organizace jej používají nejen k modelování architektury svých systémů, ale také ke zpracování implementace těchto modelů v průběhu celého životního cyklu vývoje produktu, služeb i organizačních metodik.

Celkově je vazba využitého SW a metod na dílčí cíle práce uvedena v tabulce číslo 1.

Tabulka 1: Vazba využitých metod a SW na dílčí cíle

Dílčí cíl	Využitá metoda/software
Analýza současného stavu poznání	Systematická rešerše dle standardu PRISMA
Specifikace klíčových znalostí a informací	Metody obsahové analýzy (neuronové sítě, analýza sentimentu), dotazování
Návrh konceptuálního modelu	Enterprise Architect

3 Teoretická východiska a analýza současného stavu poznání

3.1 Základní pojmy

Mezi základními pojmy je zde nutné definovat data, informace, znalosti, komunikace, podnik a vzájemné propojení.

3.1.1 Data, informace a znalosti

Základním stavebním kamenem jsou data, jež reprezentují objektivní fakta o událostech. Data mohou být jak čísla, tak písmena nebo symboly. Datům je nutné porozumět, je třeba je interpretovat a přiřadit jim význam. V praxi jde o něco, co lze získat experimentem, měřením, pozorováním či šetřením. Data jsou objektivní a zobrazují vlastnosti či stavy objektů. Mohou zobrazovat i probíhající procesy v reálném prostředí (Wu, Chiclana, Fujita, & Herrera-Viedma, 2017a). Data jsou jednoduché reprezentační nástroje faktů s jednosměrným a jedinečným významem. Data je možné rozdělit na:

- **Strukturovaná:** Zachycují fakta, atributy a objekty. Důležitým aspektem je existence elementů dat. Možností uložení těchto dat je relační databázový systém, využívající hierarchie elementů pole, záznam, relace a další databáze. Toto strukturování usnadňuje výběr relevantních dat.
- **Nestrukturovaná:** Data lze vyjádřit jako určitý tok bitů bez další specifikace, jako jsou například videozáznamy, zvukové nahrávky, obrázky nebo textové dokumenty.

Data jsou surovinou, z níž mohou vyvstávat informace. Pro ukázkou data 745486-P4X639 nebo Krrgn jistě reprezentují něco z reálného světa, ale bez dalšího popisu a vysvětlení kontextu nedávají smysl (Kratzwald et al., 2018a).

Informace

Informace jsou obtížněji uchopitelný termín. Je možné je najít za všemi oblastmi lidských činností a síla i problematičnost spojená s informacemi vychází právě z této všeobecnosti. Ve středověké filozofii tento termín znamenal „dát formu myšlence“. Už od samotného vzniku toto slova znamená jakési zhmotnění myšlenky do komunikovatelné podoby s cílem sdělovat, přenášet myšlenky a komunikovat (Xiang & Gretzel, 2010).

Jiný pohled na informace tvrdí, že jde vlastně o data, kterým příjemce přiřadí určitý význam na základě svých znalostí, zkušeností a vědomostí.

Znalost

Znalost neboli poznání či poznatek, vědění, dovednost vyplývá z porozumění zákonitostem. Informace s přidanou hodnotou je znalost. Pro lidskou mysl je připravena tak, aby ji bylo možné používat a na jejím základě se dál rozhodovat. Znalosti jsou založené na interpretaci, zkušenostech, poznávání a porozumění. Jsou také závislé na schopnostech dávat si věci do souvislostí (Sigala & Chalkiti, 2015b). Označit něco jako znalost znamená širší, hlubší a bohatší pochopení než data a informace. Znalosti se vyznačují také tím, že se těžko předávají. Životní cyklus znalosti je možné rozdělit do čtyř základních kategorií:

- Tvorba, získání a zjištění znalostí či poznatků;
- Uchování, uspořádání;
- Přenos, sdílení, šíření;
- Používání, aplikace.

Znalost zahrnuje mentální proces porozumění, chápání a učení.

Explicitní a tacitní znalost

Znalosti se dají dále rozdělit na explicitní a tacitní.

Explicitní znalost: Je kodifikovaná a schopná přenosu či sdílení. Tuto znalost je možné uchovávat a archivovat. Mezi takové znalosti patří fakta nebo teorie.

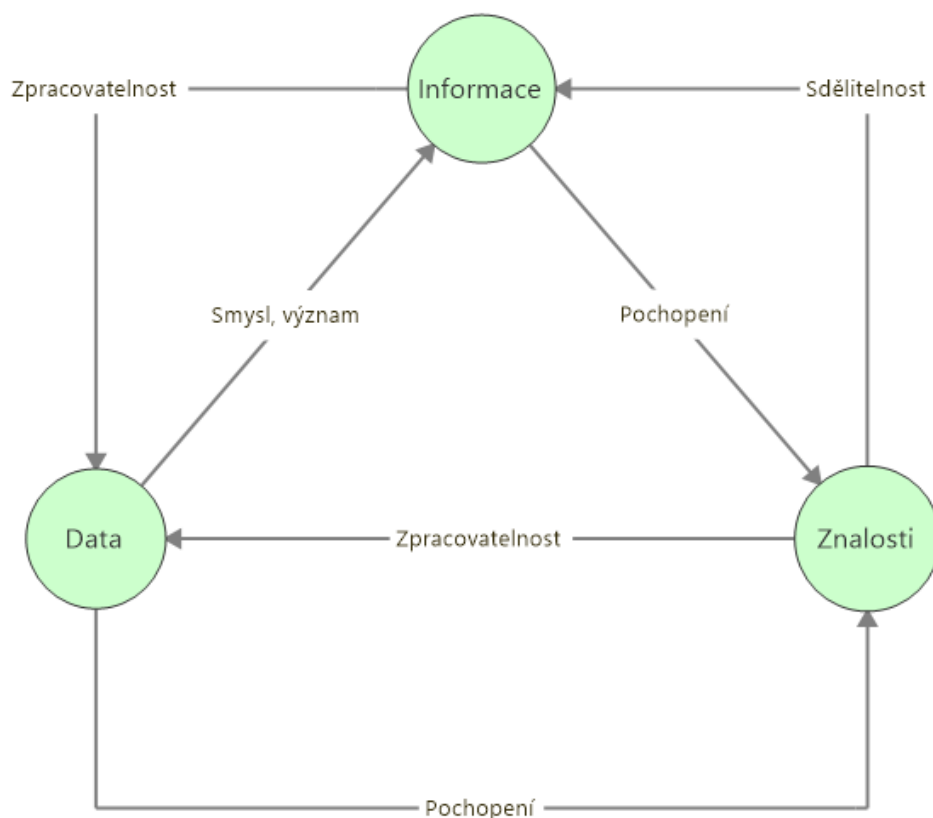
Tacitní znalost: Tato znalost je naopak těžko kodifikovatelná a těžko přenositelná nebo uzpůsobená pro sdílení. Je skrytá v mysli člověka a založená na osobních zkušenostech. Mezi tyto znalosti patří například hodnoty, zkušenosti nebo způsoby konání (Ngai, Tao, & Moon, 2015).

3.1.2 Učící se znalostní organizace nebo systém

Učící se organizace je pojem charakterizující zvyšování schopností dosahovat požadovaných výsledků prostřednictvím společného kolektivního učení. Jde o způsob, jak se podílet na vytváření a zlepšování reality. Díky trvalé podpoře učení u zaměstnanců roste vzdělání a dochází k samotné transformaci organizace.

Organizace lze podle tohoto vzdělávacího kritéria rozdělit do tří typů:

- **Tradiční organizace:** U tohoto typu prochází zaměstnanci určitým výcvikem.
- **Učící se organizace:** Zde se zaměstnanci učí sami navzájem, a to z toho důvodu, že chtějí.
- **Znalostní organizace:** Tento typ organizace vede své zaměstnance k plnému využití znalostního kapitálu v rámci naplňování požadavků zainteresovaných stran.



Obrázek 4: Propojení vztahu dat, informací a znalostí (vlastní dílo autora, inspirováno (Cheng et al., 2018; Cooper, 2017))

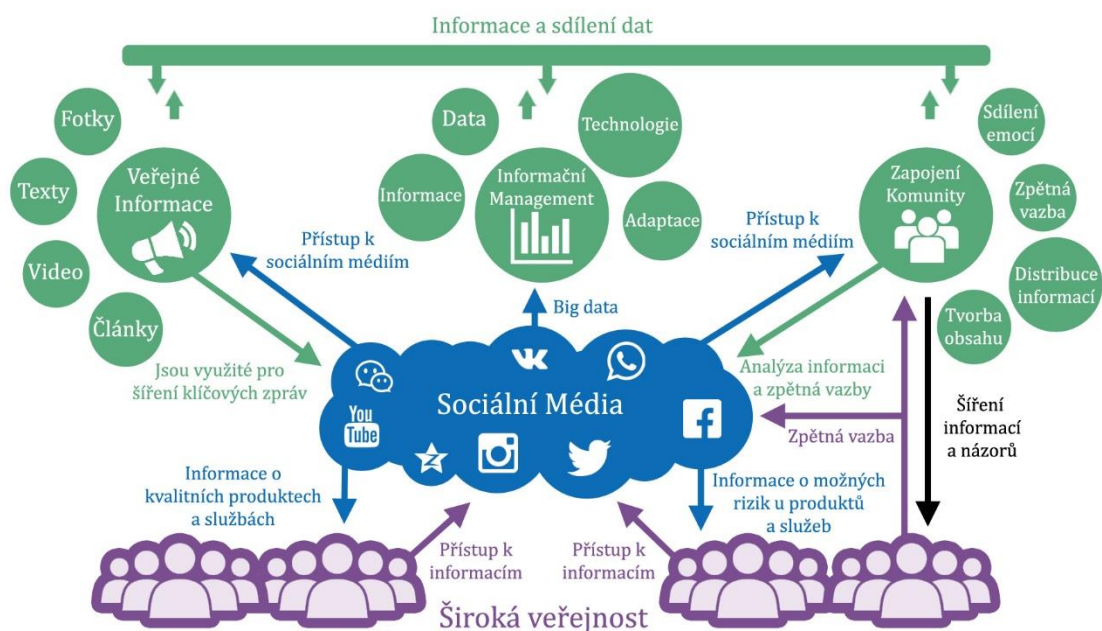
Dalším souvisejícím pojmem je moudrost spojená s procesem individuálního učení, kde je kontext příliš osobní. Moudrost je soubor znalostí vycházející z pochopení podstaty problematiky v daných souvislostech. Vychází z využití znalostní kompetence (rozumové, emoční inteligence) jednotlivce, vysokého stupně lidského poznání, z hodnoticích kritérií a individuálního vztahu k okolnímu prostředí, světu (Peláez et al., 2019a).

3.1.3 Sociální média a sociální sítě

Sociální média jsou využívána v mnoha odvětvích společnosti. Ve vzdělávání jsou označována za ideální prostředí pro kombinované vzdělávání, v sociální sféře jsou využívána pro zvýšení povědomí například o problematice závislých a zvýšení duševní pohody (sdílení osobních problémů a možností jejich řešení), ve zdravotnictví i ve veřejné sféře (Bachmann 2018; Best et al. 2016; Househ 2013). Tato práce se více zaměřuje na působení soukromých společností na sociálních médiích.

Sociální média lze definovat mnoha způsoby, všechna jsou založena na možnosti vytvářet multimediální obsah, jako jsou texty, videa, obrázky pro osobní použití, ale také na možnosti sdílení s přáteli či ostatními uživateli. V dnešní době chtějí být zákazníci dobře informovaní, již nestačí dostávat informace jednou za čas, zákazníci vyžadují aktuální informace (Ashley & Tuten, 2015). Noviny, televize a rozhlas jsou postupně nahrazovány internetem a sociálními médii. Pro splnění očekávání spotřebitelů se firmy musely naučit používat sociální média pro propagaci svých služeb či zboží a navázat komunikaci se zákazníky nebo potenciálními zákazníky. Propagace a marketing jsou dnes nedílnou součástí sociálních médií (Mynarzová a Kaňa 2014).

Mezi známá sociální média je možné zařadit Facebook, YouTube, Twitter, MySpace, Flickr, Instagram atd. Tyto sítě jsou vždy online a podporované na mobilních zařízeních. Sociální média mohou firmám pomoci ušetřit čas, finanční prostředky, díky nim lze získávat informace od zákazníků téměř zadarmo, navíc je zde výhoda toho, že si zákazníci mohou navzájem radit bez nutnosti zásahu firmy (Si 2015).



Obrázek 5: Interakce podniků se sociálními médii a veřejností (vlastní zpracování inspirováno (Eroshkin et al., 2017; Jue et al., 2010; Mayer, 2009; Phua et al., 2017; Prajogo et al., 2018))

Každým dnem se vliv sociálních médií rozšiřuje, protože se připojuje stále více lidí do tzv. online komunit. Facebook je sociální nástroj, který nabízí synchronní (psaní s přáteli v reálném čase) a asynchronní interakce (není zapotřebí, aby účastníci reagovali ihned, příkladem je zaslání zprávy a obdržení odpovědi druhý den), sdílení obsahu videí, obrázků, hudby, her, aplikací, skupin a mnoha dalších položek. V dnešní době už poskytuje kanál na sdílení fotografií, videí i textu. Navíc se snaží nahradit i prodejní a informativní (zpravodajské) kanály za pomoci svého Facebook Marketplace a Facebook Live (Tuten a Solomon 2018).

Stránky sociálních sítí jsou založeny na webové platformě, kde si mohou uživatelé vytvářet své profily, navazovat přátelství, komunikovat, nahrávat fotografie či videa, dávat „lajky“, komentovat obsah a mnoho dalšího (Phua et al. 2017).

Sociální sítě patří mezi aplikace tzv. Webu 2.0, které jsou založeny na tvorbě obsahu samotnými uživateli. Většinou zde neexistuje žádná redakce nebo autorita, která každý příspěvek musí před publikováním schválit. Obsah publikují výhradně uživatelé, kteří se starají také o jeho další distribuci (Bednář 2011).

Facebook s 2,13 mld. měsíčně aktivních uživatelů ukazuje, že sociální média jsou kriticky důležitá pro firmy a jejich značku (Popp a Wilson 2018). S růstem sociálních sítí za

posledních deset let se některé z nich staly nedílnou součástí každodenního života. Sociální sítě nahrazují způsoby zapojení do společenské interakce, získávání informací a jejich sdílení. Inspirované Facebookem začaly vznikat další formáty sociálních sítí, které měly a mají své jedinečné rysy. Od textové komunikace ke sdílení obrázků se sociální sítě staly místem, kde se mohou jednotlivci vyjádřit. Zvláště Instagram a Snapchat jsou dvě sociální sítě zaměřené na sdílení obrázků (Choi a Sung 2018). Nejvíce rozšířené sociální sítě jsou dostupné ve více jazycích a dovolují uživatelům spojit se s přáteli nebo lidmi napříč politickým, ekonomickým a geografickým hranicím.

Facebook je pro uživatele doména, na které mohou najít nové přátele, odesílat komentáře, sdílet fotografie, videa a odkazy, chatovat, sledovat obsah ostatních uživatelů a mnoho dalšího. Pro firmy je Facebook příležitostí pro zviditelnění a získání nových potenciálních zákazníků, s nimiž může poté pracovat. Facebook je díky svým informacím o uživateli jedním z nejlepších nástrojů pro cílenou reklamu (Sabaté et al.).

YouTube je sociálně mediální síť se zaměřením na videa. Firmy na YouTube mohou inzerovat třemi druhy videoreklam. Výhody YouTube pro firmy představují ukázky produktů v akci, budování komunity, ukázka odborných znalostí, předvedení značky, archivace prezentací a možnost jejich přehrání a řešení problémů uživatelů (Khan, 2017).

Instagram funguje obdobně jako Facebook. Když uživatel přidá fotografii, zobrazí se na daném profilu a v takzvaném feedu těch, kdo profil sledují. Jedná se o zjednodušenou verzi Facebooku s větším důrazem na mobilní zařízení, fotografie a videa. Na Instagramu je možné sledovat, být sledován, komentovat, posílat soukromé zprávy a ukládat fotografie. Uživatelé Instagramu jsou nejen spotřebitelé, ale také obchodníci, přesněji 86 % špičkových značek provozuje oficiální profil na Instagramu a zveřejňuje příspěvky o své značce nebo produktech ve snaze navázat komunikaci se současnými či potenciálními zákazníky. Instagram nabízí dvě umístění propagovaných příspěvků: ve feedu a stories (Rang Choi 2018; Phua et al. 2017).

Výhody marketingu na sociálních médiích

- **Velké publikum** – Twitter má více než 320 milionů měsíčních aktivních uživatelů, Instagram více než 500 milionů, což dělá ze sociálních médií velkou příležitost pro tvůrce po celém světě a napříč všemi průmysly.
- **Svobodné vytváření** – Začátek vytváření jakékoli marketingové strategie je na všech největších sociálních médiích zdarma. Placené marketingové nástroje samozřejmě garantují, že se počet potenciálních zákazníků bude drasticky zvyšovat, ale marketing je možné provádět zcela zdarma.
- **Podporuje sdílení** – Snad nejvýraznější a zcela jedinečnou výhodou sociálních médií je schopnost získat pomoc od svých následovníků. Lidé sdílejí věci se svými sítěmi od fotografií a receptů až po zajímavé články a aktuální nabídky. Na rozdíl od jiných forem internetového marketingu, jako jsou stránky a placené reklamy, je obsah na sociálních médiích často sdílen. Následovníci ho mohou sdílet se svými následovníky, kteří ho pak sdílejí se svými následovníky, což poskytuje širší dosah (s nižšími náklady) než tradiční marketingová kampaň (Ding et al., 2019).
- **Silnější loajalita značky** – Kromě zvýšení dosahu značky umožňují sociální média zvýšit loajalitu značky. Studie provedená společností The Social Habit ukazuje, že 53 % Američanů, kteří sledují určité značky na sociálních sítích, je vůči nim s větší pravděpodobností loajální. Sociální média jsou více než prodejní místo, usnadňují obousměrnou komunikaci, jež umožní budovat smysluplné vztahy se stávajícími a potenciálními zákazníky. Všechny tyto aspekty je činí jistějšími při jejich rozhodnutí důvěřovat vlastnímu podnikání a vybízí je, aby si v budoucnu vybrali stejnou značku.

Nevýhody marketingu na sociálních médiích:

- **Negativní zpětná vazba** – Největší nevýhodou marketingu na sociálních médiích je skutečnost, že negativní zpětná vazba může být ničující. Uživatelé sociálních médií mohou psát vše, co chtějí. Spokojený zákazník může zanechat perfektní recenze, ale nespokojení zákazníci mohou zanechat vzteklé, nenávislné komentáře a všichni na sociálních médiích si je mohou přečíst.
- **Rozpaky** – S různými marketingovými příspěvky a kampaněmi je snadné udělat chybu nebo prostě použít špatné dvojnásobné slovo, které může při rychlém šíření příspěvku způsobit firmě mnoho komplikací.
- **Časová náročnost** – Vytvoření a udržení interaktivního marketingového profilu na většině sociálních médií může vyžadovat mnoho času a úsilí. Pokud je tým sociálních médií malý s omezenými zdroji, může být obtížné udržet dobré výsledky.

Marketingovým oddělením se v posledních letech otevřel zcela nový propagační kanál. Firmy mají nyní jedinečnou možnost, jak velmi rychle a do jisté míry osobně reagovat na názory zákazníků, dotazy nebo reklamace závadného zboží. Velmi oblíbená je pro ně především reklama prostřednictvím word of mouth (WoM – ústní předávání). Jde o bezplatnou formu reklamy, kdy zákazníci ostatním sami sdělují, že mají rádi určitou firmu, produkt nebo službu. Pokud se informace pomocí osobního předávání šíří extrémně rychle, dá se dokonce mluvit o virálním marketingu. Přesněji může být definován jako „elektronické ústní předávání“, při němž se forma marketingové zprávy týkající se firmy, značky či produktu předává exponenciálně narůstající rychlostí, často pomocí sociálních médií.

Sociální média se dají využít i jako zákaznická podpora. Pro zákazníky není nic snazšího než napsat stížnost na profil firmy na dané platformě a čekat, až se ozve některý ze zaměstnanců. Pokud má firma založený profil na některé ze sociálních sítí a chce aktivně komunikovat se zákazníky, existuje několik základních možností, jak profil spravovat: vlastnoruční správa, delegování správy na některého ze zaměstnanců, založení speciální divize, správa prostřednictvím agentury a kombinace výše zmíněných možností.

Naplnění firemních cílů na sociálních sítích lze měřit pomocí několika různých metrik a na jejich základě pak dále upravovat aktuální firemní strategii.

Jednou z často používaných metrik je návratnost investic (return on investment, zkráceně ROI), která se počítá jako čistý zisk dělený kapitálem vynaloženým k jeho dosažení

a násobený 100. Výsledek se udává v procentech a dá se jím vyjádřit výkonnost jakéhokoli podnikání nebo úspěšnost jakékoli investice.

U marketingových kampaní na sociálních sítích ale bývá problém v tom, že lze jen velmi těžko zjistit, jestli vedly firmou zveřejněné příspěvky a placené reklamy ke zvýšení prodeje. I proto se používají jiné sociální metriky, díky kterým mají firmy přehled, jak se jim na sociálních sítích daří budovat základnu fanoušků a zaujmout uživatele.

Dosah udává procentní podíl fanoušků stránky, kteří viděli vybraný příspěvek. Důležité je rozlišovat dva typy dosahu: organický, tedy počet osob, kterým byl příspěvek zobrazen prostřednictvím neplacené distribuce, a placený, tedy počet jedinečných osob, kterým byl příspěvek zobrazen díky jeho finanční podpoře.

Firmy mohou zaplatit za to, aby se vybraný příspěvek zobrazoval více uživatelům, nebo dokonce konkrétním uživatelům na základě podrobného nastavení cílové skupiny. Například Facebook na speciální stránce pro podnikatele udává, že průměrný organický dosah příspěvků je přibližně 16 %. Údaj však pochází z roku 2012. Je všeobecně známo, že Facebook organický dosah příspěvků neustále snižuje, čímž nutí firmy k vyšším investicím.

Engagement rate (zkráceně ER, česky míra zapojení) je metrika sledující míru zapojení uživatelů. Jde o jeden z hlavních ukazatelů kvality obsahu a aktivity fanoušků.

Metrikami response rate a response time se zjišťuje, jak rychle a úspěšně zodpovídají jednotlivé firmy dotazy uživatelů. V rámci analýzy sociálních sítí byla využívána především specifická metrika response rate for questions, tedy procentuální hodnota vyjadřující, kolik dotazů položených uživateli přímo na zeď stránky na Facebooku stránka zodpověděla. Response time určuje, jak dlouho trvalo stránce na daný příspěvek či dotaz na zdi zareagovat.

Inzerenti a marketingoví odborníci také často využívají metriku zvanou conversion rate – ta se aplikuje při porovnání celkového počtu návštěvníků webové stránky s počtem návštěvníků, kteří se na základě návštěvy stránky později stali platícími zákazníky nebo odběrateli. Na Facebooku lze díky conversion rate sledovat například poměr uživatelů, kteří se po kliknutí na sponzorovaný příspěvek stali fanoušky stránky.

Ke sledování některých výše zmíněných metrik lze využít nástroje pro správu a monitorování působení firem na sociálních sítích. Tyto analytické nástroje nabízí podrobné statistiky a údaje o konkurenčních stránkách, mnoho z nich navíc umožňuje i plánování příspěvků, sledování uživatelských dotazů a přiřazování úkolů jednotlivým členům týmu, kteří správu sociálních profilů zajišťují.

On-line marketingem můžeme daleko lépe zacílit na naše potenciální zákazníky, je také lépe měřitelný než již zmíněná televizní reklama. Aby byl on-line marketing efektivní, je nutné správně zvolit a naplánovat marketingovou strategii. Na internetu je nepřehledné množství marketingových nástrojů, ale je nutné zvolit takové, které jsou vhodné pro cíl marketingové strategie i oblast jejího zaměření.

S využitím nástrojů jako je search engine optimization (SEO), pay per click (PPC) kampaně, kontextová reklama, bannerová reklama, newslettery, sociální sítě a další lze vytvořit mohutnou podporu pro každou organizaci. Velkou výhodou je široký dosah mezi uživateli a možnost sdílení příspěvků.

3.1.4 Podnik

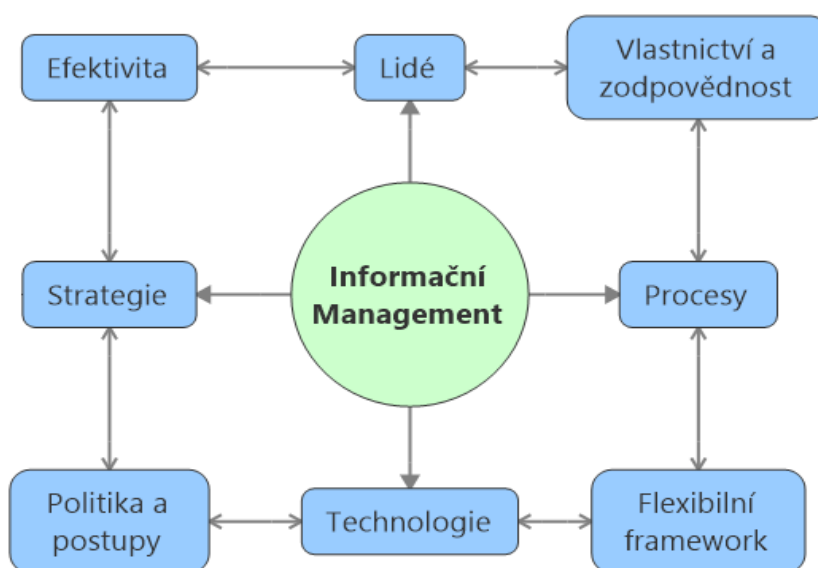
Podnik je pojem, který se používá ve dvou různých významech. V právním pojetí (obchodní firma) se jedná o název, pod kterým je podnikatel zapsán v obchodním rejstříku. V ekonomickém pojetí se používá jako synonymum k pojmu podnik společnost. Z hlediska velikosti (v podmínkách Evropské unie, vč. České republiky) se podnik dělí na malý, střední a velký. Tato disertační práce by svým výstupem měla být schopna pomoci všem třem typům, pokud bude podnik aktivní na sociálních sítích.

Za středního podnikatele se považuje podnikatel, který zaměstnává méně než 250 zaměstnanců a jeho roční obrat nepřesahuje 50 milionů EUR nebo jeho bilanční suma roční rozvahy nepřesahuje 43 milionů EUR. V rámci kategorie malých podniků jsou malé podniky vymezeny jako podniky, které zaměstnávají méně než 50 osob a jejichž roční obrat nebo bilanční suma roční rozvahy nepřesahuje 10 milionů EUR. V rámci kategorie malých a středních podniků jsou drobní podnikatelé vymezeni jako podnikatelé, kteří zaměstnávají méně než 10 osob a jejichž roční obrat nebo bilanční suma roční rozvahy nepřesahuje 2 miliony EUR (Czech Invest, 2019; Eurostat, 2019).

Informační management

Výchozí bod o „informačním managementu“ poukazuje na to, že efektivní dosažení cíle v informační činnosti předpokládá interní zpracovávání a zprostředkovávání informací stejně jako výměnu informací s okolním světem.

Nově jsou změněné podmínky organizace informací v přechodu od průmyslové k informační společnosti a dále ke společnosti znalostní. Změněné podmínky vedou k novým organizačním koncepcím (Jue et al., 2010). Na pojmu informační management a z něj odvozených koncepcích je změněné chápání organizace provozní informace a komunikace a také její nová hodnota.



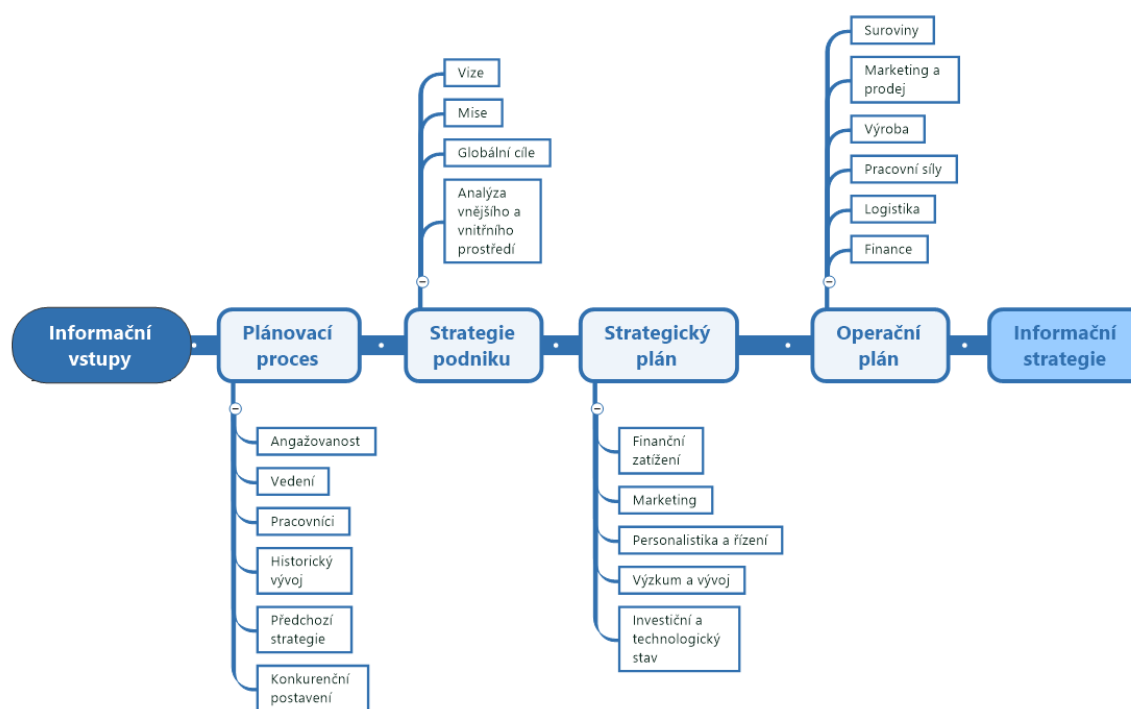
Obrázek 6: Klíčové oblasti informačního managementu (vlastní dílo autora inspirováno (Eroshkin et al., 2017; Prajogo et al., 2018))

V poslední době posilují úvahy, jak zlepšit metody managementu informačních procesů. „Informování“ nebo jinak řečeno „informační činnosti“ mají dle principu podléhat stejně jako další oblasti činnosti jednotlivým funkcím managementu – plánování, řízení, kontrole a organizaci. Zřetelná je nutnost integrace různých informačních a komunikačních úkolů do firemních procesů. Jako úkol managementu se chápe komplexní a mnohotvárné nasazení nových informačních a komunikačních technologií. Konečně vstupuje do praxe a vědy managementu vědomí, že člověk a podnikání se nyní nacházejí ve společnosti informací a vědění (ve společnosti znalostní), s níž přichází převrat veškerých pracovních

struktur a struktur vědomostí (Roberts et al., 2016a). S tím spojené zvažování kulturních, sociálních a ekonomických šancí a rizik představuje významný úkol managementu.

Hlavní funkce informačního managementu

- Funkce 1: Informační podpora organizačně propojeného jednání;
- Funkce 2: Vývoj racionálních informačních konceptů;
- Funkce 3: Plánování a řízení informační činnosti;
- Funkce 4: Organizace a prezentace poznatků v podnikání;
- Funkce 5: Management informačních a komunikačních technologií.



Obrázek 7: Informační strategie podniku (vlastní dílo autora inspirováno (Ghaisani et al., 2017; Hwang et al., 2018; Zhou et al., 2018))

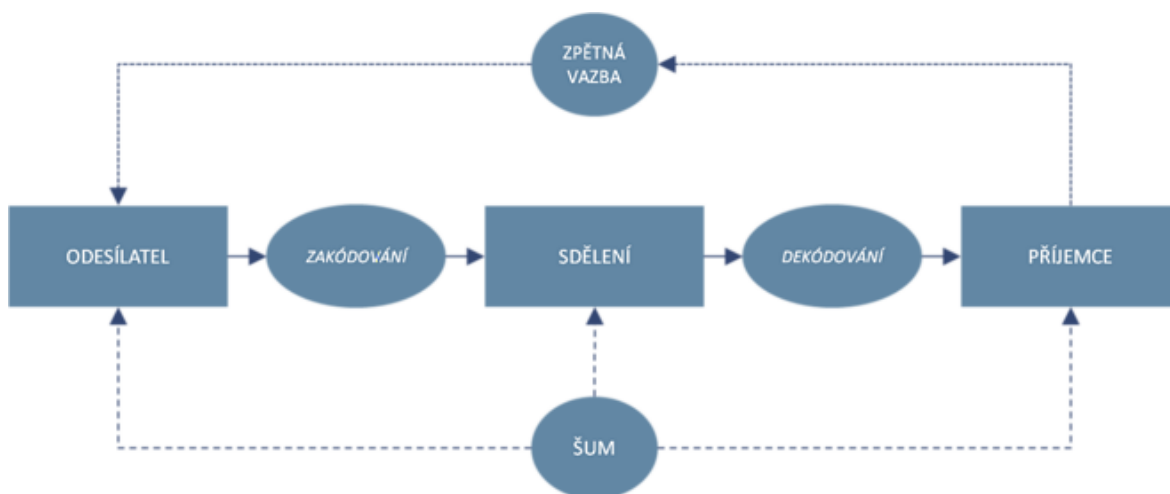
IM lze definovat jako transdisciplinárně pojatý soubor poznatků, metod a doporučení systémových přístupů a informatiky, které pomáhají vhodně realizovat informační procesy manažerského myšlení a jednání k dosažení cílů dané organizace (ManagementMania, n.d.).

Analyzováním dat od konkurence, vlastních úspěchů i neúspěchů z minulosti, příspěvků na sociálních sítích mohou podniky získat informační výhodu na trhu a modely v této práci by měly pro podniky fungovat ke zlepšení efektivity z hlediska většího sdílení jejich obsahu a zajištění více interakcí na sociálních sítích. Toto zlepšení u podniků pomůže rozvíjet jejich strategii a bude se jednodušeji dodržovat firemní vize.

3.2 Komunikace

Komunikace – původ tohoto slova lze nalézt v latině, v současné době je však jeho definice mnohem širší, než jak byla vnímána dříve. Původní význam slova komunikace – tedy *communicare* – byl *činit něco společným, společně sdílet, vespolečně se účastnit* (Vybíral, 2009, s. 25).

Podle G. A. Millera (1951) komunikace znamená, že *informace přechází z jednoho místa na druhé*. J. W. Wander Zanden (1987) používá následující definici – „*komunikace je proces, jímž lidé předávají informaci, ideje, postoje a emoce jiným lidem*“. Tento autor také předpokládá společné sdílení znaků (*common sense*), jimiž se lidé dorozumívají – například jazyk, gesta či způsob chování ovlivněný kulturou, ve které byli jedinci socializováni (Nakonečný, 2009, s. 288).



Obrázek 8: Obecný průběh komunikace (Management Mania, 2018)

Z pohledu směru a lokace se komunikace podniku dá rozdělit na interní a externí.

3.2.1 Interní komunikace

Schopnost jednat s lidmi je problematikou oblasti řízení lidských zdrojů. Úspěšnost efektivního vedení lidí závisí na schopnosti komunikovat danou problematiku firmy. Hlavní úlohou interní komunikace je propojení zaměstnanců, managementu, majitele podniku a jejich angažovanost. Pokud se podaří vytvořit efektivní a fungující komunikaci mezi těmito prvky, je mnohem jednodušší plnit operativní i strategické cíle firmy (Szarková 2008; Trnovcová 2016).

Interní komunikace, formu, toky a její nástroje by měla mít každá firma nastaveny individuálně podle stadia hospodářského cyklu, ve kterém se právě nachází (pokles, stabilita, růst) a podle charakteru podniku (struktura zaměstnanců, velikost firmy, firemní kultura a hodnota podniku). Kvalita interní komunikace firmy se pozná v průběhu zavádění změn, kdy musí být zaměstnanci informováni, aby se předešlo chaosu, dezorganizaci a šoku (Zerfass et al., 2014). Když změna proběhne plynule a bez problémů, je to ukazatel toho, že je interní komunikace nastavena dobře. Komunikace není jen o přidělování příkazů, ale je také o komunikaci s majitelem a o řešení aktuálních mezilidských vztahů na pracovišti. Kvalitní podniková komunikace nevyžaduje jen delegování od managementu k řadovým pracovníkům, ale také komunikaci mezi všemi odděleními (marketing, PR, IT, lidské zdroje aj.). Díky správně zavedené interní komunikaci se zvyšuje produktivita práce či motivace zaměstnanců a snižují se dny pracovní neschopnosti (Trnovcová 2016).

Nástroje interní komunikace

Komunikačními nástroji uvnitř podniku nemusí být jen osobní rozhovory, podnikové noviny nebo telefon. S nástupem nových technologií mohou firmy začít komunikovat rychleji, efektivněji a zábavněji. Nové komunikační kanály jsou například videokonference, e-mail, audio, web-casting, SMS, sociální sítě a další. V organizacích se stále využívají i tradiční metody, jako jsou schůze, porady, interní časopis, letáky, tabule a přímá komunikace manažera se zaměstnancem. Před implementací určitého nástroje je potřebné zvážit, jací jsou zaměstnanci, kteří budou nástroj používat, jaký mají přístup k technologiím, jakou formu nebo obsah zpráv jsou schopni přijmout či pochopit a jaká interakce, zpětná vazba může nastat (De Marco 2015; Trnovcová 2016).

3.2.2 Externí komunikace

Externí komunikace je veškerá komunikace, kterou firma vyvíjí navenek. Spadá do ní i komunikace komerční (myšleno placené komerční sdělení v médiích aj.) a komunikace přirozená (např. image firmy, inovace, firemní identita, kultura a styl) (Gómez et al., 2019; Hong et al., 2016). V užším významu označujeme externí komunikaci jako souhrn všech forem marketingové komunikace.

Internetová reklama je forma komunikace, kdy inzerent nebo dodavatel využívá internet k přenosu svých sdělení směrem k zákazníkovi. Uživatelé upouštějí od tradičních médií a migrují na internet, který jim nabízí mnohem větší interakci a individualizaci (Lei et al., 2019; Muninger et al., 2019). Uživatelé mají kontrolu nad tokem informací, což vede k reklamám a sdělením, jež jsou pro ně relevantní. S rozvojem sociálních sítí se individualizace a efektivita ještě více prohloubila. Sociální sítě shromažďují informace o návycích, chování a zájmech uživatelů. Reklama na internetu má svá pro a proti (Denić et al. 2018).

Mezi výhody internetové komunikace se řadí (Denić et al. 2018):

- globálnost – internet je na celém světě a prodej produktů či služeb je možné provádět po celém světě,
- zaměření na cílové skupiny – internet umožňuje zacílit na vybrané skupiny lidí už geograficky, nebo třeba i podle jejich zájmů,
- snížení nákladů – reklama v televizi je poměrně nákladná, oproti tomu na internetu lze za nižší částku oslovit poměrně velkou skupinu potenciálních zákazníků,
- vyhodnocení reklam – pokud je určen cíl internetové reklamy správně, dá se ze získaných dat dobře vyhodnotit,
- rychlé přijetí změn – reklamu je možné změnit pomocí několika kliků,
- interaktivita – zákazníci mohou na reklamy reagovat komentářem pod příspěvek, napsáním zprávy a dalšími způsoby, které jsou jim nabídnuty,
- vysoká úroveň měření – při správném nastavení webové stránky a propojení s analytickými nástroji je možné změřit téměř cokoliv (klik na určitý prvek dobu strávenou na stránce, dokončení objednávky a další).

Za nevýhody se považují (Denić et al. 2018):

- přeplněnost stránek reklamami – bohužel v dnešní době jsou webové portály přeplněné reklamami a zákazníci začínají bannery přehlížet – jedná se o takzvanou bannerovou slepotu (Patel 2015),
- krátká životnost reklamy – uživatelé nechtějí vidět jednu reklamu stále dokola, je zapotřebí ji obměňovat,
- při špatné volbě cílové skupiny je možný nulový dopad reklamy.

Externí komunikace může být vysvětlena jako strategie, která zahrnuje širokou škálu metod využívaných firmami, aby zaujaly veřejnost. Marketing je součástí externí komunikace, ale zaměřuje se spíše na předání své hodnoty zákazníkům, avšak externí komunikace je o propojení s kýmkoliv mimo podnik. Zahrnuje například marketing sociálních médií a video obsahu, ale také prezentace pro akcionáře nebo kampaně investorů (Harvey 2018).

Externí komunikační strategie jsou závislé na mnoha faktorech – od umístění značky přes marketingové cíle až po rozpočet. Všechny tyto aspekty mohou mít vliv na to, jak bude externí komunikace vypadat. V dnešní době je potřeba dávat si velký pozor na výskyt a využívání falešných zpráv, takzvaných fake news, případně fake reviews, tedy falešné recenze (Hájek et al., 2020b). Na strategii externí komunikace neexistuje univerzální řešení, ale většina strategií obsahuje tyto prvky: obsahový marketing (blog), konference a pořádání akcí (event marketing), e-mail marketing, sociální média a tiskové zprávy (Harvey 2018).

3.2.3 Public relations

Public relations jsou často uváděné pod zkratkou PR. Toto označení pochází z angličtiny a volně jej lze přeložit jako „vztahy s veřejností“. PR jsou techniky a nástroje, s jejichž pomocí instituce nebo firma buduje a udržuje vztahy se svým okolím a s veřejností, nahlíží jejich postoje a snaží se je ovlivňovat. Jedná se o dlouhodobou cílevědomou činnost, která by měla mimo jiné zajišťovat poskytování informací veřejnosti a zároveň získávat zpětnou vazbu a další informace od veřejnosti. Důležitým aspektem PR je zajištění obousměrné komunikace (například na rozdíl od reklamy). PR představuje důležitou složku sociální komunikace a kontroly.

PR je součástí komunikačního mixu, který dále obsahuje reklamu, osobní prodej, přímý marketing a podporu prodeje. Celý komunikační mix je součástí marketingového mixu.

PR se výrazně odlišuje od reklamy. Zatímco se reklama snaží hlavně o krátkodobý prodej určitého produktu nebo služby, PR se soustřeďuje na dlouhodobé cíle, jako jsou image, vztahy a komunikace s cílovými skupinami. Reklama i PR působí na širokou veřejnost, ale reklama pro svou komunikaci s veřejností využívá zásadně masová média (televize, tisk, rozhlas, internet, venkovní reklamu), zatímco PR komunikuje i jinými prostředky.

Mezi využití a cíle PR patří:

- **sdělování informací** potenciálním zákazníkům ve snaze podpořit prodej, udržet nebo zlepšit reputaci; získávání zpětné vazby od zákazníků následně umožňuje upravovat metody a techniky prodeje a komunikace,
- **řešení nepříznivých vlivů** – odvracení útoků na podnik (například očeňující články, nepravdivá prohlášení, zkreslené statistiky atd.),
- U zaměstnanců se snaží o **zlepšení informovanosti** a **zvýšení motivace** (firma funguje dobře – spokojení zaměstnanci pracují lépe).

3.2.4 Relevantní oblasti pro podnikovou externí komunikaci

Externí komunikace je druh komunikace, kterou firma vyvíjí navenek, a spadá do ní jak komunikace komerční (myšleno placené komerční sdělení v médiích aj.), tak i komunikace přirozená (např. image firmy, firemní identita, kultura a styl). V užším významu označujeme jako externí komunikaci souhrn všech forem marketingové komunikace (Corral de Zubielqui et al., 2019; Muninger et al., 2019). Mezi tyto formy patří:

- podpora prodeje zaměřená na spotřebitele,
- podpora distribuce,
- aktivity PR,
- přímý marketing.

Prodej je podporován formou komunikace, jejímž cílem je stimulovat prodej výrobků a služeb konečným spotřebitelům či distribučnímu článku prostřednictvím dodatečných podnětů, které nabízenému produktu dodávají přidanou hodnotu v časově omezené době (Din et al., 2015; El Ouiridi et al., 2016).

Podpora prodeje představuje nástroj akcelerující zvýšený prodej. Těto akcelerace je dosaženo využitím vhodných stimulů (například peníze, cena, výrobek poskytnutý zdarma, drobné dárkové předměty atd.), které vedou zákazníka k nákupu, návštěvě obchodu, pátrání po dalších informacích či k jiným žádoucím aktivitám. Tato část propagačního mixu je také jednou z nejdůležitějších částí tzv. podlinkové marketingové komunikace (Gómez et al., 2019).

Další způsob dělení nástrojů PR nabízí PENCILS, kde se jednotlivé aktivity zčásti překrývají, ale vždy jde o jiný pohled na danou oblast, a tudíž i o jiný přístup a navázání vztahu z jiné perspektivy:

- **publications:** výroční zprávy, podnikový časopis, časopis pro významné zákazníky, publikace k výročí společnosti nebo k jiné důležité události,
- **events:** akce, ať už veřejné, nebo vnitrofiremní – sponzorství kulturních, sportovních nebo charitativních aktivit, představení nového produktu, udělování odměn zaměstnancům apod.,
- **news:** materiály pro novináře a podklady pro tiskové konference, které zahrnují základní informace o podniku, o produktech, základní informace o vysokém managementu nebo o nově přichozících či právě odcházejících zaměstnancích,
- **community Involvement Activities:** angažovanost v lokální komunitě, ve snaze o začlenění do běhu komunity a porozumění jejím potřebám, uvědomění si způsobu, jakým postavení nové továrny nebo zvýšení výroby zasáhne jak do krajinného rázu, tak i vnitřního fungování společnosti,
- **identity média:** využití korporátní identity v celkové komunikaci od jednotného fontu v e-mailech přes hlavičku dopisního papíru a vzhled obálek až po podnikovou uniformu,
- **lobbying activity:** lobbování za cíle společnosti, krizové PR, regulační opatření apod.,
- **social Responsibility Activities:** společenská odpovědnost firmy, která sahá od ekologické výroby po budování dobrého jména aktivitami v sociální oblasti.

V posledních letech se vztah mezi výrobcí a distribučními články mění. Změna postavení distribuce, jakož i ekonomická situace a převis nabídky nad poptávkou vedou k posílení podlinkových aktivit ve vztahu k jednotlivým distribučním článkům, jejichž prostřednictvím se výrobky dostávají ke konečným spotřebitelům. Kvalita a loajalita jednotlivých distributorů velmi silně ovlivňuje dostupnost výrobku, jeho pozici v prodejnách

a jeho objednávané množství. Základními cíli podpory distribuce bývá zavedení nového výrobku, rozšíření sortimentu, získání nových velkoobchodů či maloobchodů, zvýšení numerické distribuce, zvýšení objemů prodeje, zvýšení tržního podílu a budování loajality distributorů (Hong et al., 2016; Lei et al., 2019).

Obsahem zprávy může být řada informací, podmínkou je, aby tyto informace byly pro veřejnost dostatečně zajímavé. Zpráva může zahrnovat informace o příchodu nového produktu na trh, výsledcích soutěže nebo významném úspěchu dosaženém pracovníkem podniku (Corral de Zubielqui et al., 2019; Muninger et al., 2019). Důležité je spojení zprávy s firmou nebo jejím produktem, stručnost, správné výtvarné řešení a úprava materiálu, dle možností jeho doplnění fotografií, vyloučení věcných nebo pravopisných chyb v textu atd.

Direct marketing bychom měli chápat jako určitou filozofii, která je založena na vybudování trvalé a pevné vazby mezi firmou a zákazníky – stávajícími i potenciálními. V řadě aspektů se odlišuje od ostatních forem marketingové komunikace. Jednak využívá dvousměrné komunikace, jež umožňuje velmi přesné vyhodnocení úrovně dosažení stanovených komunikačních cílů. Cílová skupina bývá obvykle menší než u ostatních forem, protože tato forma komunikace je úzce zaměřena na ty recipienty, kteří byli vyhodnoceni pomocí údajů příslušné databáze jako potenciálně perspektivní či podchycení zákazníci (El Ouiridi et al., 2016; van Esch & Mente, 2018). Přímý marketing je ve svých tradičních formách nákladnou formou komunikace, v případě e-shopů to ale neplatí úplně.

Externí komunikace na sociálních sítích je často zastoupena budováním dobrého jména firmy, důvěry, inovace, propagace/marketing, náborem zaměstnanců, informováním (slevy, plány, kupóny, soutěže) a inspirací (příspěvky, které nejsou nutně spojeny s firmou) (Gómez et al., 2019; Hong et al., 2016; Lei et al., 2019).

Vývoj technologií přinesl do jisté míry novou podobu interpersonální (mezilidské) a skupinové komunikace. *Přínosem internetu je umožnění informační dostupnosti a prostupnosti, interaktivnost při výměně zpráv a rychlost zprostředkování kontaktu.* Využívání internetu má dopady na kvalitu řeči a myšlení (ochuzování a zužování slovní zásoby, zpragmatictění projevu, roztěkanost, úsečnost atd.) a má dopad na životní styl. Odlišností digitální komunikace od komunikace tváří v tvář můžeme nalézt celou řadu. Jednotlivé dopady těchto způsobů komunikací také působí rozdílně na lidskou psychiku a vývoj osobnosti (Vybíral, 2009).

Celkový pohled na nejdůležitější a nejpoužívanější on-line metriky je rozdělený na staré neboli tradiční a nové neboli moderní metody měření. Staré metody jsou (Marketing Mojo 2013):

- počet zobrazení stránek – kolikrát byla stránka dohromady zobrazena,
- unikátní počet zobrazení stránek – kolikrát byla zobrazena stránka jednomu návštěvníkovi,
- počet návštěvníků – kolik návštěvníků iniciovalo přístup na webovou stránku, započteno s opakovanými návštěvami,
- unikátní počet návštěvníků (noví návštěvníci) – kolik návštěvníků iniciovalo přístup na webovou stránku, nezapočítává opakované návštěvy,
- čas strávený na webu – ukazatel toho, jak dlouho návštěvník zůstane na dané stránce nebo webu,
- bounce rate – procentuální ukazatel, kolik lidí ihned opustilo webovou stránku, aniž by provedli jakoukoliv interakci,
- počet stránek na jednu návštěvu – kolik stránek si zobrazí každý návštěvník.

Tyto metriky jsou často kritizovány z mnoha důvodů, ať už kvůli možnému zkreslení dat, nebo horší vypovídající hodnotě či kvůli datům, ze kterých jsou počítány. Přesto je nelze zcela zanedbat, pokud se správně interpretují, mohou být i přínosné. Mimo výše zmíněné tradiční metriky existují také moderní (Marketing Mojo 2013):

- Konverze lze napojit na dané podnikatelské cíle, jako jsou například sběr kontaktů, stažení souboru, objednávka.
- Konverzní poměr dává možnost zjistit slabé stránky webu a optimalizovat danou stránku např. pomocí nového designu, jiných textů a dalších věcí.
- Čas na pořízení ukazuje, za kolik dní či hodin nakoupí uživatel od první návštěvy. Tuto hodnotu lze využít v naskladňování zboží.
- Zdroj návštěvy je nejvíce využívaný ukazatel a vyjadřuje, z jakých zdrojů (vyhledávače, placená reklama, sociální sítě aj.) přichází návštěvníci. Proměnná zdroj návštěvy je velice zajímavá v souvislosti s konverzí.
- Asistovaná konverze je velice důležitý ukazatel pro pochopení chování zákazníka před jejím provedením. Pomocí asistovaných konverzí lze odhalit jednotlivé kroky a zdroje návštěvy uživatele před uskutečněním konverze.

3.3 Vybrané sociální sítě pro externí komunikaci

Celosvětově jsou nejvyužívanější sociální sítě Facebook, YouTube, Twitter, Instagram, WhatsApp a LinkedIn (Baird & Parasnis, 2011). Poznatky zjištěné v rámci této disertační práce lze v určité míře využít na jakékoliv sociální síti, kde lze přidávat text/popisky.

Facebook

Při tvorbě účtu na Facebooku si uživatel vytváří profil, ve kterém může uvést své jméno, místo a datum narození, náboženské vyznání, zaměstnání, dosažené vzdělání, rodinný stav a celou řadu dalších údajů. Po úspěšném založení účtu lze přidávat tzv. přátele. Facebook uživateli na základě e-mailových kontaktů automaticky nabídne seznam osob, které by mohl znát. Uživatel se může stát členem různých komunitních skupin a může zakládat nové. Skupiny slouží především ke sdružování uživatelů s podobnými zájmy. Uživatel si dále může „oblíbit“ některou z existujících stránek, případně založit novou. Vlastní stránky si zakládají například celebrity, firmy, televizní stanice či rádia. Facebook obsahuje také jednoduchý chat pro on-line komunikaci s přáteli, systém pro zasílání zpráv, události, fotky a videa. Po přihlášení uživatel vidí již zmiňovanou zeď, na níž se mu zobrazují nejnovější příspěvky přátel, stránek a skupin, jejichž je členem. Zeď je možné každých pár minut aktualizovat, a proto neustále udržuje uživatelskou pozornost.

Základním principem šíření příspěvků je sdílení, komentování nebo označení „To se mi líbí“ (like). Jakmile uživatel u příspěvku nějaké stránky či skupiny provede jednu ze tří výše zmíněných interakcí, je víc než pravděpodobné, že se někdo z jeho přátel dozví, že tak učinil. Stejně tak může daný uživatel sledovat, co si oblíbili jeho přátelé, případně přidávat na svůj profil aktualizace ve formě tzv. textových statusů. Facebook umožňuje i sdílení fotek a videí.

Twitter

Sociální médium, které je občas zařazováno do kategorie microbloggingu, je unikátní především originálním systémem aktualizací stavů – tzv. tweety, které omezují textové příspěvky registrovaných uživatelů na maximální délku 140 znaků.

Na Twitteru si uživatel nepřidává přátele, ale sledované. Pokud naopak někdo sleduje uživatele, je jeho sledujícím (follower). Každý 140znakový tweet lze „retweetnout“ (retweet), tedy přeposlat svým sledujícím, nebo označit jako oblíbený. Twitter podobně jako

Facebook nabízí na základě toho, co si oblíbili nebo „retweetnuli“ uživatelem sledovaní uživatelé, na titulní straně zajímavé novinky a profily, které se každých pár vteřin aktualizují.

Twitter rovněž značně zpopularizoval hashtagy, které obsahují znak „#“ a slouží pro označování příspěvků s podobnou tematikou. Díky hashtagům se také průběžně generuje seznam aktuálně nejpopulárnějších témat, o kterých se na Twitteru mluví.

YouTube

Služba YouTube byla založena v únoru 2005 a umožňuje sledovat, sdílet a nahrávat videa – i proto stojí na pomezí sociálních médií a sítí. Každý uživatel má možnost nahrávat videa na svůj kanál, který má určitý počet odběratelů, což jsou uživatelé, kteří se přihlásili k odběru kanálu, protože chtějí být informováni o nově nahraných videích. U každého videa je uveden jeho název, délka, popis a počet zhlédnutí. Uživatel má možnost video okomentovat, sdílet či ohodnotit pomocí palce nahoru, nebo dolů.

Instagram

Instagram je sociální síť a mobilní aplikace založená v roce 2010 a primárně určená pro sdílení obrázků. Každému uživateli se zobrazuje obsah podle toho, koho sleduje. V aplikaci je možné sdílet fotografie a videa do délky 60 vteřin. Uživatelé mohou veškerý svůj obsah z Instagramu sdílet na ostatních sociálních sítích, jako jsou Facebook či Twitter. Instagram měl dva měsíce po svém spuštění jeden milion uživatelů, za dva roky vzrostl počet uživatelů na 200 milionů s 1 miliardou sdílených fotografií.

LinkedIn

LinkedIn je největší profesní sociální síť na světě, na které se setkávají profesionálové a diskutují o svých pracovních zájmech. Mezi uživatele patří manažeři, konzultanti a odborníci z nejrůznějších oborů a mají zde své účty také firmy. V květnu roku 2020 má LinkedIn více než 690 milionů uživatelů ve více než 200 zemích, v ČR jich má 1,5 milionu.

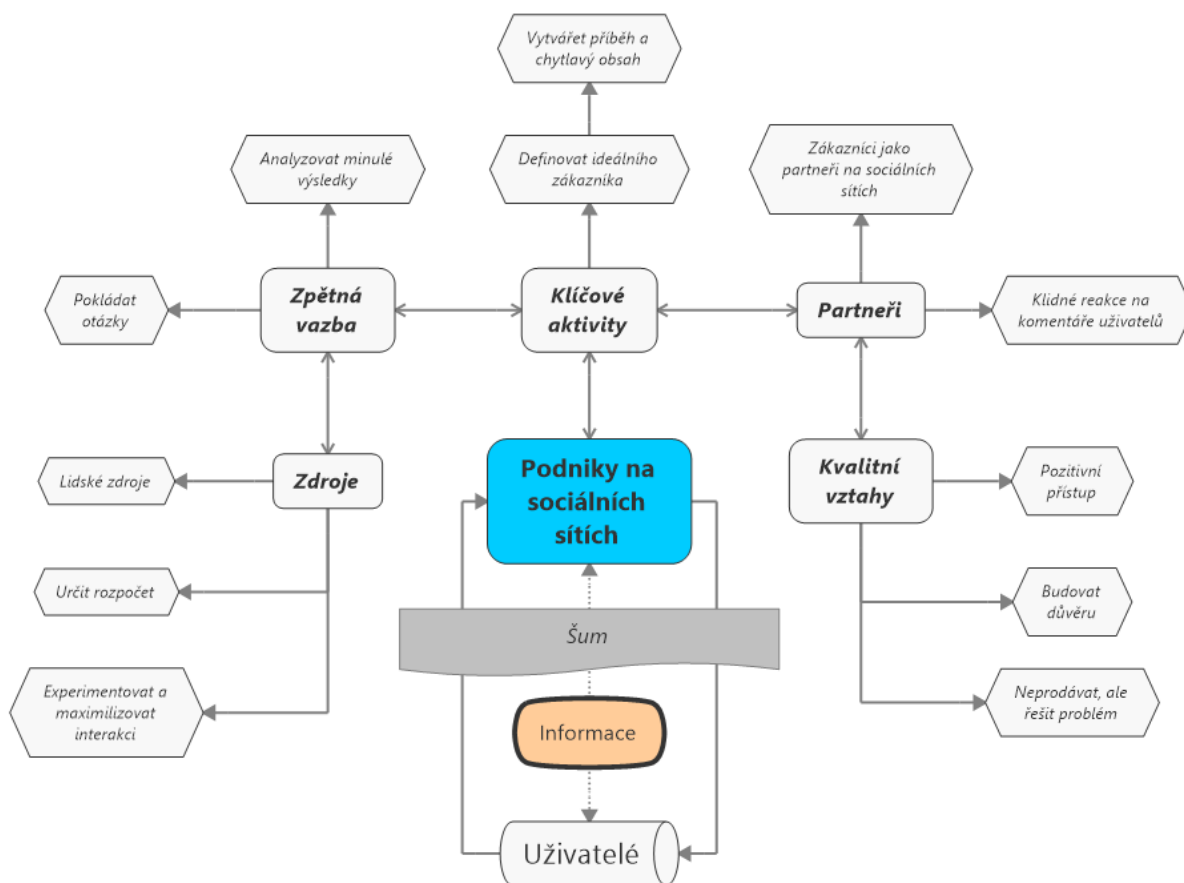
3.3.1 Obchodní model na sociálních sítích

Obchodní model je obecně způsob, jakým společnost vytváří hodnotu pro sebe a zároveň dodává produkty a služby zákazníkům. Produkty pro zákazníky zde mohou být považovány za řešení, která uspokojí jejich problém nebo potřebu.

Tyto dvě věci mohou být kombinovány na sociálních médiích, společnost může splnit individuální potřebu, která může být jen zábavná a může řešit problém s koherentní reklamou. Tento model bychom neměli chápat jako model pro organizování celé společnosti, ale speciálně pro sociální média.

Hodnota

První velkou skupinou je hodnota. Jedná se o cílené publikum. Facebook má především velmi propracovaný reklamní systém, a když společnost správně definuje své publikum, pak je snadné tak učinit na sociálních médiích.



Obrázek 9: Klíčové aktivity externí komunikace na sociálních sítích (vlastní zpracování inspirováno (Ashley & Tuten, 2015; Drury, 2008; Hofacker & Belanche, 2016))

Segmenty

Druhou velkou částí modelu jsou zákaznické segmenty, kde je postaven archetyp zákazníka. Kdo jsou nejčastější (typičtí) zákazníci a proč by nakupovali nebo využívali služby společnosti?

Kanály

Třetí část modelu ilustruje kanály. Facebook a YouTube dosáhly více než 1000 milionů uživatelů, což znamená, že více než třetina všech uživatelů internetu na světě a více než šestina světové populace jsou aktivní na těchto sítích. Před rokem 1990 byl jediným kanálem fyzický kanál, který obsahoval prodejce, obchody. Ale v této době máme virtuální kanály, web, mobilní a zejména sociální média s velmi složitým a specializovaným systémem přesně pro tento případ. Problém fyzického kanálu spočívá v tom, že je drahý a musí být na velmi dobrém místě, aby ho lidé mohli vidět, pokud se předpokládá, že fyzický kanál funguje. Se sociálními médii je možné se zaměřit přesně na potřebnou věkovou kategorii, pohlaví atd. Pokud chce společnost zvýšit své příjmy, může se zaměřit pouze na osoby, které si již něco koupily; může cílit na lidi ve specifických skupinách, geografických lokalitách a mnohem více. Důležitou otázkou v této části modelu je to, která sociální média měla společnost (firma, zájmová skupina) používat. Facebook je nutný, ale pokud by se měly přidat Instagram, Twitter a další, záleží na tom, co společnost dělá, co je jejím hlavním cílem a jak velká je.

Zákaznický vztah

Čtvrtou částí modelu je vztah se zákazníkem; v klasickém obchodním modelu je to způsob, jak získáváme zákazníky. Pokud kombinujeme obchodní model se sociálními médii záleží více na tom, jak udržujeme dobré vztahy se zákazníky, jak vytváříme zábavný obsah a reagujeme v komentářích a příspěvcích stejně, jak zákazníci chtějí, asertivní, klidný postoj vždy s trochou humoru v případě sociálních médií. Na sociálních médiích je důležitá otázka, jak vytvořit obsah, který by se šířil virálně, protože sociální média mohou velmi rychle získat společnosti tisíce potenciálních zákazníků s minimálními investicemi.

3.4 Využití sociálních sítí jako nástroje externí komunikace firem – současný stav poznání

Za pouhé desetiletí sociální média revolucionizovala život mnoha lidí, a tak přitahuje velkou pozornost nejen z oblasti průmyslu, ale také z akademické sféry. Příchod sociálních médií podstatně změnil způsob, jakým mnoho lidí, komunit nebo organizací komunikuje a spolupracuje (Ngai, 2015). Rok 2016 byl pro mnohé z největších platforem sociálních médií obdobím konsolidace. Článek „Budoucnost sociálních médií (a jak se na to připravit): Zpráva o stavu sociálních médií 2016“ popisuje význam sociálního marketingu jako

klíčového pro přežití v konkurenci. „Představte si, že dokážete předpovědět další velkou věc na sociálních médiích s vědomím, jak budou sociální média vypadat v roce 2017 a dále a jak budou značky používat sociální platformy pro spojení se svým publikem. Tyto informace by byly neocenitelné pro někoho, kdo denně pracuje se sociálními médii.“ Podle průzkumu společnosti State of Social Media 2016 83 % obchodníků uvedlo, že by v příštích 12 měsících chtěli vytvořit více videoobsahu. Stejná zpráva ukázala, že téměř 93 % obchodníků používá Facebook a 91 % z nich používá reklamy na Facebooku. Tři hlavní důvody, proč sociální média využívají, jsou povědomí o značce s 85 %, angažovanost komunity se 71 % a distribuce obsahu 61 %. Dalšími důvody pro využívání sociálních médií jsou prodej 54 % a zákaznická podpora 21 %. Prognózy poklesu marketingu na sociálních médiích jsou nejvýznamnější na Google+ s 27 %. Skutečný význam sociálních médií a v budoucnu ještě více ukazuje, že 93 % společností, které využívají sociální média pro marketing, v příštích 12 měsících zůstanou na svém současném rozpočtu, nebo ho zvýší. V roce 2014 obchodníci utratili 9 % svého celkového rozpočtu na sociální média, v roce 2015 to bylo 13 %. Podle některých výzkumů budou na ně obchodníci v roce 2020 utrácet více než 21 % svého celkového rozpočtu.

Konstantní vývoj sociálních médií a interaktivního chování často způsobuje, že se uživateli generovaný obsah rozšíří velmi rychle po internetu. Popularita sociálních médií a jejich obsahu nám může pomoci pochopit těchto aktivit, které vykazují vliv na sociální, ekonomické i vládní aktivity. Modelování a predikování popularity online obsahu představuje důležitou oblast zkoumání u sociálních médií a může mít příznivý dopad v mnoha ohledech u veřejného managementu, organizací, firem i domén zabývajících se online bezpečností. Například může podporovat krizový management tím, že bude znát dopad přírodních katastrof, terorismu nebo zločinu (G. Chen, Kong, Xu & Mao, 2019). V oblasti podnikání může predikce pomoci analyzovat aktuální trendy i obavy u uživatelů, poskytnout cenné informace pro modelování rizik i poukázat na možnosti zvýšení zisku (Cerchiello, Giudici, & Nicola, 2017).

Predikce popularity není triviální úloha. Množství generovaného obsahu na sociálních médiích je obrovské a každou sekundu se zvyšuje. Například jenom na Facebooku a Twitteru jsou stovky milionů aktivních uživatelů měsíčně a provázanost jejich vztahu je velmi komplikovaný problém. Data zde také nejsou nijak strukturována a je tady mnoho neformálních vyjádření. V neposlední řadě existuje ale mnoho faktorů, které ovlivňují úspěšnost příspěvků, jako jsou obsah textu, typ uživatele a čas přidání, délka příspěvku,

zvolený obrázek nebo video, pozitivní, nebo negativní naladění příspěvku, vhodnost a délka nadpisu, počet reakcí za prvních pár hodin a mnoho dalších proměnných.

Studie na predikci popularity primárně využívají metody založené na vlastnostech (Tsur & Rappoport, 2012; Weng, Menczer, & Ahn, n.d.). Tyto metody extrahují velké množství vhodných vlastností ohledně uživatele, typu sítě, složení textu, data přidání a poté predikují popularitu po testování na strojovém učení či jiných klasifikačních modelech.

Ty jsou značně závislé na volbě správných vlastností, což může být časově náročné, a také je zde nutný expertní pohled. V případě, že se zvolí nevhodné vlastnosti/proměnné, bude model nepřesný a křehký (Mishra, Rizoio, & Xie, 2016).

Prvotní studie se zaměřovaly na jednoduché regresní modely, ale některé z posledních let využívají modelování popularity pomocí stochastických procesů (Mishra, Rizoio, & Xie, 2016). Tyto modely ale poměrně často využívají jenom predikci pomocí časových řad a ignorují ostatní cenné informace pro predikci popularity.

Popularitou se u sociálních médií rozumí zvýšená pozornost věnovaná online obsahu vyjádřená počtem sdílení, přeposlání tweetů, počet zhlédnutí nebo označení reakce („to se mi líbí“, „super“, „paráda“). Predikce popularity míří na predikování těchto zmíněných proměnných s pomocí různých vlastností jak u textového příspěvku, události, tak i videa nebo obrázku. Existující přístupy v těchto metodách predikce spadají do čtyř kategorií: predikce čistě pomocí vlastností, analýza časových řad, kaskádová procesní analýza a hloubkové strojové učení (G. Chen et al., 2019).

Predikce pomocí vlastností extrahuje proměnné jako takzvané povrchové proměnné (například čas publikování, délka obsahu, délka názvu nebo počet obrázku a odkazů v příspěvku), akumulativní proměnné (počet článků publikovaných ve stejný čas), textové vlastnosti (například frekvence různých klíčových slov jak v názvu, tak v textu), sémantické vlastnosti (případně zvolené lokace, osoby nebo organizace a jejich případná aktuální oblíbenost) nebo fyzické vlastnosti (teplota, počasí) (Bandari, Asur, & Huberman, 2012).

Analýza pomocí časových řad má výhodu informace o vývoji v čase a zjištění popularity na začátku příspěvku. Základní studie popularity pomocí časových řad využívají jednoduchou lineární funkci (Szabo & Huberman, 2008). Některé další práce využívají multivariační lineární regresi (Pinto, Almeida, & Gonçalves, 2013), ARMA modely (Gürsun, Crovella, & Matta, 2011) nebo multiplikativní sezónní Holt-Winter model (Hu, Hu, Fu, Shi, & Ning, 2016). Tyto modely předpokládají, že budoucí popularita má lineární, nebo nelineární vztah s počáteční popularitou.

Kaskádová procesní analýza využívá pro predikci popularity podobnosti z minulých úspěšných a neúspěšných událostí. Zhang (Zhang et al., 2015) simuluje proces oblíbenosti u událostí používáním informací o podobných publikacích z minulosti, textový i další obsah u události k určení zájmu o ni.

Ohledně strojového učení („Machine learning“ nebo „Deep neural networks“) existuje několik možných neuronových sítí, které lze využít pro predikci popularity. Základní neuronovou sítí se ale stala opakující se neuronová síť (anglicky RNN – Recurrent Neural Network). Rozšiřující modely pro tuto základní neuronovou síť jsou pak LRCN (Donahue et al., 2014), DeepCas (Li, Ma, Guo, & Mei, 2016), LSTM (Serban, Sordoni, Bengio, Courville, & Pineau, 2015) a GRU (Chung, Gulcehre, Cho, & Bengio, 2014). Ovšem neuronové sítě předpokládají monotónní závislé změny u celé sekvence, přičemž tyto faktory mohou být komplikovanější na sociálních médiích, navíc zde platí proměnné jako čas sdílení (pracovní hodiny, večerní hodiny, víkend) a to může být obtížné zachytit s klasickou opakující se neuronovou sítí (RNN).

Nicméně menší počet studií věnuje pozornost výzkumu pro implementaci prediktivních systémů, které lze efektivně použít k předpovědi vývoje postu před jeho zveřejněním. Systém schopný předvídat dopad zveřejněných individuálních příspěvků může poskytnout cennou výhodu při rozhodování komunikovat prostřednictvím sociálních médií, přizpůsobit propagaci produktů a služeb. Reklamní manažeři by mohli rozhodovat na základě výsledků právě z těchto prediktivních systémů. Dolování dat poskytuje zajímavý přístup pro extrahování prediktivních znalostí ze surových dat. Většina studií je ale zaměřená na reaktivní hodnocení toho, co uživatelé říkají prostřednictvím různých odpovědí, postů či tweetů.

Cílem a jednou z metod disertační práce jsou i systematické rešerše a prozkoumání aktuálních metod spojených s klasifikací dat z různých sociálních sítí. Analýza a aplikace těchto metod mohou u firem zlepšit celou řadu aspektů, mezi něž patří:

- rozhodování v podobě rychlejšího šíření informací,
- reklama,
- zjišťování názorů na produkt,
- identifikace, v jaké nákupní fázi se uživatel nachází,
- využití pocitů, komentářů a dostupných informací o uživatelích pro nárůst sdílení i reakcí a tím potenciální rozšíření povědomí o firmě, produktu či pořádané akci.

4 Využití neuronových sítí v rámci sociálních médií – systematická rešerše

V rámci zkoumání byla provedena systematická rešerše dle standardů PRISMA (Emily Jones, n.d.) na neuronové sítě, klasifikační modely vzhledem k podnikům a marketingu, další statistické výzkumy ohledně využívání uživatelů dominantních sociálních sítí a vedlejší výzkum leteckých společností a jejich využívání a aktivita na sociálních médiích. Rozsah systematické rešerše je prováděn pro zjištění a shrnutí aktuálních výzkumů na různé druhy neuronových sítí a jiných klasifikačních modelů, jako jsou například rozhodovací stromy, regresní modely či metoda nejbližších sousedů a případně jejich kombinace, využívané pro data ze sociálních sítí. Systematická rešerše se skládala z několika kroků: nalezení a identifikování vhodných článků, vyfiltrování dle kritérií, prozkoumání konečného počtu vybraných článků pro extrakci hlavních informací. Těchto článků je v konečném filtru 22.

4.1 Vyhledávací strategie a kritéria

Pro nalezení vhodných publikací bylo zvoleno několik kritérií a strategií výběru. Postup filtrování publikací pro získání finálního výběru je následující:

- Nalezení článků dle klíčových slov „Neural Networks“, „Marketing“, „Machine Learning“, „Artificial Intelligence“ a „Social Media“ mezi roky 2010 a 2019 – počet publikací 1415 (bližší informace v tabulce 1).
- Odstranění duplikací – odstraněno 190 publikací z výběru.
- Odfiltrování nevhodných publikací dle klíčových slov a názvu – odstraněno 984 publikací z výběru.
- Vymazání publikací po přečtení abstraktu – odstraněno 194 publikací po přečtení abstraktu.
- Prozkoumání celého textu 47 publikací:
 - Odstraněno 8 publikací pro využití klasifikace jiným způsobem než neuronovými sítěmi;
 - Odstraněno 6 publikací z výběru pro nedostatečný popis řešení;
 - Odstraněno 6 publikací s hlavním zaměřením jiným než na podnikání (například predikce vývoje burzy nebo predikce věku a pohlaví uživatele na sociálních sítích);

- Odstraněno 5 publikací v jiném jazyce než v angličtině.
- Z celkového počtu 1415 publikací z kroku 1 zbývá 22 k analýze.

Tabulka 2: Rozložení zkoumaných článků (v období 2010 až květen 2019, vlastní dílo autora)

Klíčová slova s využitím „AND“ u všech slov	WOS	Scopus
Neural Networks Social Media	612	52
Neural Networks Marketing	577	76
Neural Networks Social Media Marketing	58	6
Deep Learning Social Media Marketing	22	4
Artificial Intelligence Social Media Marketing	8	0
Celkem	1277	138

Extrakce dat a způsob filtrování článků

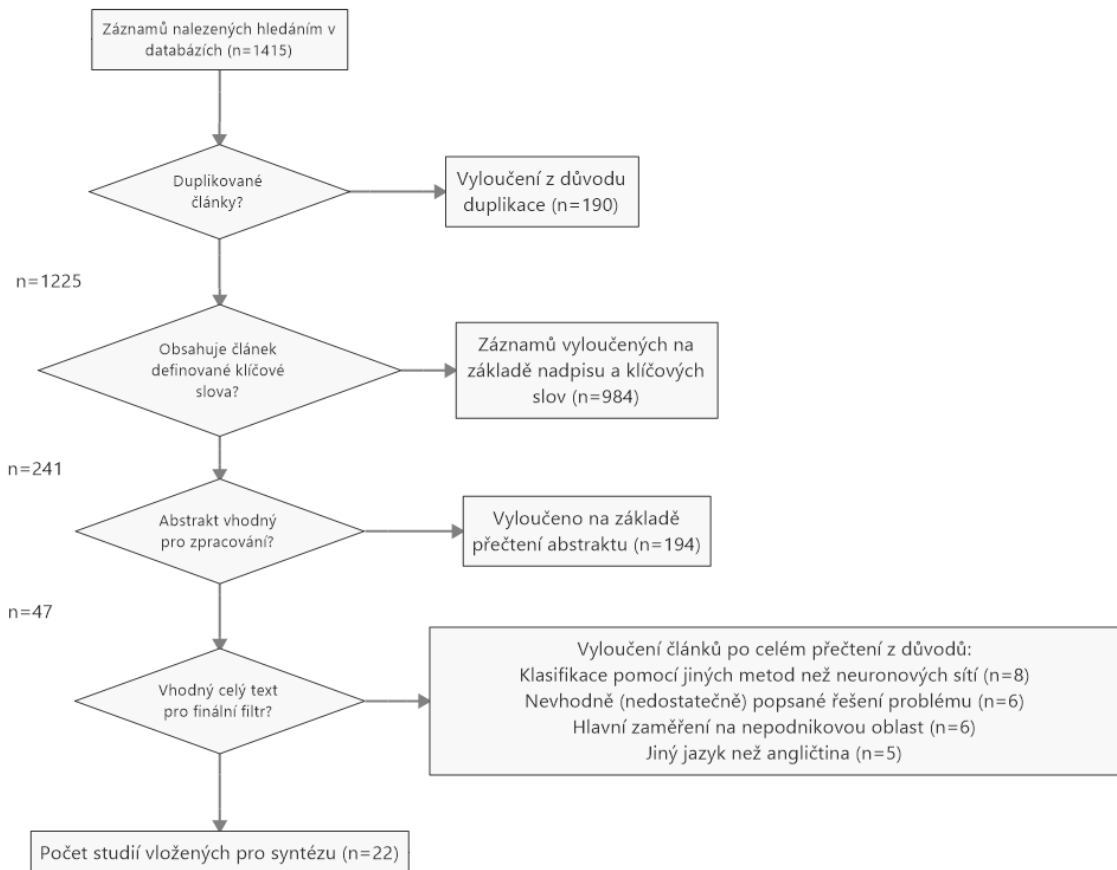
Články pro rešerši byly vybrány dle následujících kritérií:

- Publikováno v roce 2010 nebo novější;
- Zaměření na neuronové sítě a jejich modifikace;
- Zaměření na jiné klasifikační modely (rozhodovací stromy, regrese, metody nejbližších společných vlastností) spojené s metodikou neuronových sítí;
- Hlavní výsledky z článků by měly být využitelné na sociálních sítích například pro budování firmy, zlepšování image firmy, kontrolování psychického stavu uživatelů, k průzkumu spokojenosti, názorů zákazníka na výrobky, služby a podobně;
- Zaměření na jakékoliv sociální sítě, jako jsou Facebook, Twitter, Instagram;
- Autoři mohli analyzovat i data ze sítě YouTube, Amazon, blogů či online fóra;
- Text musel být napsán v anglickém jazyce.

Články byly vyřazeny na základě těchto kritérií:

- Článek je napsán v jiném jazyce než v angličtině;
- Publikace byla vydána před rokem 2010;
- Článek se zaměřuje na neuronové sítě bez spojitosti se sociálními sítěmi (například predikce nemocí, počasí, vývoje burzy apod.), v této fázi bylo vyřazeno nejvíce článků;
- Článek se zaměřuje jenom na neuronové sítě (bez spojitosti na sociální sítě);

- Zaměření publikace je pouze na sociální sítě (bez zaměření na neuronové sítě či jiné klasifikační modely);
- Zaměření článků na zdravotnictví, zeměpisné problémy či predikce vývoje na burze;
- Článek je systematická rešerše.



Obrázek 10: Postup filtrování článků (vlastní dílo autora)

4.2 Hlavní zjištění – využívané metody v rámci neuronových sítí

I když mnoho z vyfiltrovaných článků využívá metody a funkce, které se ve vybraných člancích objevily jen jednou, bylo nalezeno několik metod ke klasifikaci i přímo typy neuronových sítí, které se v člancích opakují. Mezi tyto metody patří metody Support Vector Machines (SVM), Multi-layer Perceptron (MLP), Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN), Naive Bayes (NB) a Long Short-Term Memory (LSTM). Dané metody a jejich využití budou popsány dále v textu.

Jednou z nejvyžívanějších metod je Support Vector Machines (SVM), což je metoda podpůrných vektorů. Jde o metodu strojového učení sloužící pro klasifikaci a také pro regresní analýzu. Základem metody SVM je lineární klasifikátor do dvou tříd. Cílem úlohy je nalézt nadrovinu, která prostor příznaků optimálně rozděluje tak, že trénovací data náležející odlišným třídám leží v opačných poloprostorech. Optimální nadrovina je taková, že hodnota minima vzdáleností bodů od roviny je co největší. Jinými slovy, okolo nadroviny je na obě strany co nejširší pruh bez bodů (maximální odstup, angl. maximal margin, česky je tento pruh někdy nazýván také pásmo necitlivosti nebo hraniční pásmo). Na popis nadroviny stačí pouze body ležící na okraji tohoto pásma a těch je obvykle menší počet – tyto body se nazývají podpůrné vektory (angl. support vectors) a odtud název metody.

Další často využívanou metodou je Multi-layer Perceptron (MLP). Metoda perceptron a vícevrstvý perceptron je základní typ neuronových sítí. V případě, že se jedná o jednoduchý perceptron, má model vždy dvě vrstvy, a to vrstvu vstupů a výstupů. Model se následně na vložených datech trénuje a upravuje váhy vstupů. Podle těchto úprav poté zjišťuje, jak velký vliv mají úpravy vah na výstup.

V případě, že se jedná o vícevrstvý perceptron, lze si zvolit neomezené množství vrstev a uzlů mezi vstupy a výstupy. Zpravidla se metoda MLP používá v případě, že jsou uzly z každé vrstvy propojené se všemi uzly z okolních vrstev, takže čím více vrstev v modelu, tím bude délka učení modelu delší (Kim & Adali, 2002a). Někteří autoři se domnívají, že i pro velmi složité problémy stačí maximálně pouze tři vrstvy mezi vstupy a výstupy (Hampshire & Pearlmutter, 1991a).

Třetí často využívaná metoda je Recurrent Neural Networks (RNN). Rekurentní neuronová síť (RNN) je třída umělé neuronové sítě, kde spojení mezi uzly tvoří směrovaný graf podél sekvence. To umožňuje vykazovat časové dynamické chování. Na rozdíl od dopředných neuronových sítí mohou RNN použít svůj vnitřní stav (paměť) ke zpracování sekvencí vstupů. To je činí použitelnými pro úkoly, jako jsou například nerozdělení, rozpoznávání rukopisu nebo rozpoznávání řeči.

Termín „rekurentní neuronová síť“ je používán na dvě široké třídy sítí s podobnou obecnou strukturou, kde jeden je konečný a druhý nekonečný impuls. Obě třídy sítí vykazují dynamické chování. Konečná impulsová rekurentní síť je řízený acyklický graf, který může být rozvinut a nahrazen striktně dopřednou neuronovou sítí, zatímco nekonečná impulsní opakující se síť je řízený cyklický graf. Tyto sítě jsou vhodné pro textová data, analýzu řeči, klasifikace a predikční problémy spolu s regresí, ale nevhodné pro rozklad a klasifikaci obrázků.

Autoři v rešerši také opakovaně používají metodu Convolutional Neural Networks (CNN). Konvoluční neuronová síť (ConvNet/CNN) je algoritmus Deep Learning, který může přijímat vstupní obraz, přiřadit důležitost (určitelné váhy a předsudky) různým aspektům/objektům v obraze a být schopen je od sebe odlišovat. Předběžné zpracování požadované v ConvNet je mnohem nižší ve srovnání s jinými klasifikačními algoritmy. Zatímco v primitivních metodách jsou filtry konstruovány ručně, s dostatečným tréninkem, ConvNets mají schopnost naučit se tyto filtry/charakteristiky.

Architektura ConvNet je analogická k architektuře struktury neuronů v lidském mozku a byla inspirována organizací Visual Cortex. Jednotlivé neurony reagují na podněty pouze v omezené oblasti vizuálního pole známého jako recepční pole. Sběrka takových polí se překrývá, aby pokryla celou vizuální oblast.

Předposlední často využívaná metoda je Naive Bayes (NB). Klasifikátor Naive Bayes je pravděpodobnostní model strojového učení, který se používá pro klasifikační úkoly. Jeho podstata je založena na Bayesově teorému. Pomocí Bayesovy věty můžeme najít pravděpodobnost jevu A vzhledem k tomu, že na jev B došlo. B je zde důkaz a A je hypotéza. Předpokládá se zde, že prediktory/rysy jsou nezávislé.

Využívaná je také metoda Long Short-Term Memory (LSTM). Long Short Term Memory je druh rekurentní neuronové sítě. V RNN je výstup z posledního kroku přiváděn jako vstup v aktuálním kroku. Řeší problém dlouhodobých závislostí RNN, ve kterých RNN nemůže předpovídat slovo uložené v dlouhodobé paměti, ale může poskytnout přesnější předpovědi z nedávných informací. Vzhledem k tomu, že se délka mezery zvětšuje, RNN neposkytuje dostatečný výkon. LSTM může standardně uchovávat informace po dlouhou dobu. Používá se pro zpracování, predikci a klasifikaci na základě údajů časové řady.

Tabulka 3: Shrnutí finálního filtru zkoumaných článků (vlastní dílo autora)
Specifikuje účel využití vybraných metod. Zobrazen je i cíl práce, problematika, výsledky a limitace vybraných prací.

Tabulka 3: Shrnutí finálního filtru zkoumaných článků (vlastní dílo autora)

Autoři (rok)	Název publikace	Cíl	Problém k vyřešení	Vzorek	Metody	Výsledky	Limitace
Silvia Vázquez et al. (2014)	A classification of user-generated content into consumer decision journey stages.	Předložit novou analýzu a klasifikaci obsahu generovaného uživatelem z pohledu zařazení do jedné ze čtyř fází rozhodovacího procesu spotřebitele.	Identifikovat sadu lingvistických vzorů pro každou fázi nákupu, která bude následně použita v klasifikátoru založeném na pravidlech.	36 701 textů napříč 72 značkami na sociálních sítích, v recenzích, fórech i na blogu či mikroblohu.	Techniky strojového učení, speciálně rozhodovací stromy	Klasifikace nákupních fází dosahuje průměrné přesnosti 74 %. Navrhovaná klasifikace textů v závislosti na vyjádřených prvcích marketingového mixu dosáhla průměrné přesnosti 75 % pro všechny analyzované prvky.	Jazykové vzorce nejsou schopny zachytit rozdíly mezi některými produkty a jejich fázemi.
Guandan Chen et al. (2018)	NPP: A Neural Popularity Prediction Model for Social Media Content	Předpovídat popularitu v vlastním modelem kombinující obsah textu, informace o uživateli a časové řadě ve srovnání s ostatními běžnými modely.	Predikovat popularitu tweetů na Twitteru	20 000 tweetů s méně než 5 retweets. Každý obsahuje text tweetu, počet retweets v době pozorování, autory retweetů a čas publikace.	Tsurova metoda, Aiello metoda, Hawkes, DeepHawkes, DeepCas pro porovnání s vlastní metodou.	Slova, která vyjadřují emoce, mají obvykle vyšší pozornost, např. „prosím“, „pán“, „hrozne“.	Zkoumání jen několika aspektů a funkcí u Tweetu. Nezohlednění jiných sociálních médií nebo různých platform.
Ahmed Sulaiman M Alharbi and Elise Doncker de (2018)	Twitter Sentiment Analysis with a Deep Neural Network: An Enhanced Approach using User Behavioral Information	Vytvořit konvoluční neuronové síť (CNN), které berou v úvahu nejen text (uživatelský tweet), ale také chování uživatele.	Porovnat několik metod s vlastní vyvinutou neuronovou sítí.	700 000 tweetů přidáno více než 3 500 uživateli.	Sentiment analýza na Twitteru pomocí Deep Learning Convolutional Neural Network (CNN) ve srovnání s Naive Bayes (NB) a Support Vector Machines (SVM).	Navrhovaný model překonává současné modely (včetně Naive Bayes a Support Vector Machines), což ukazuje, že „překračování“ obsahu dokumentu (tweet) je prospěšné při klasifikaci sentimentu.	Příspěvek by mohl prozkoumat další modely učení založené na neuronových sítích, jako jsou rekurentní neuronové síť (RNN).
Yanbing Liu et al. (2017)	C-RBFNN: A User Retweet Behavior Prediction Method for Hotspot Topics Based on Improved RBF Neural Network	Predikce chování uživatelů v retweetování na Twitteru.	Predikce popularity tweetu pomocí fuzzy teorie a algoritmu neuronových sítí.	9 262 zúčastněných uživatelů a 534 459 potenciálních uživatelů s 1 365 106 vztahy.	Radical Basis Function (RBF), Cloud Radical Basis Function Neural Network (C-RBFNN)	Vytvořený model používal behaviorální data a data o uživatelských vztazích v sociální síti pro efektivní predikci chování uživatelů u retweetování.	Pro model jsou testována pouze tři hlavní témata.
Pradip Kumar Sharma et al. (2017)	Multilevel learning based modeling for link prediction and users' consumption preference in Online Social Networks	Prozkoumat chování uživatelů v online platformě sociálních médií s modely. Predikce propojení a preference uživatelů.	Porovnat víceúrovňovou síť Deep Belief Network (DBN) s jinými metodami.	4 039 uzlů pro Facebook, miliony recenzí, uživatelů a produktů na Amazonu a 107 614 uzlů pro Google+	Multilevel Deep Belief Network (DBN) + ostatní metody pro porovnání jako Support Vector-based Classifier (SVC), Naive Bayes and C4.5	Byly vytvořeny modely pro predikci spotřebních preferencí u nových uživatelů.	Vzorky a jejich proměnné by mohly být vysvětleny lépe. Není jasné, proč si autoři vybírají vybrané metody ke srovnání.
Pietro Ducange et al. (2018)	An effective Decision Support System for social media listening based on cross-source sentiment analysis models	Odhad pocitů uživatelů z hlediska kladné, záporné nebo neutrální polarity vyjádřené v komentáři.	Charakterizovat konkrétní kontext v komentářích.	Více než 100 000 komentářů na Facebooku, Twitteru a Instagramu.	Sentiment Analysis Engine (SAE) spolu s Decision Support Systems (DSS)	Byl navržen DSS pro sociální média na základě analýzy sentimentu.	Výzkum je zaměřen především na dvě oblasti, jimiž jsou restaurace a internetové obchody se spotřební elektronikou.
Fazeel Abid et al. (2018)	Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter	Vytvořit nový, jednoduchý a efektivní rámec, který je společnou architekturou mezi variantami RNN, které zachytí dlouhodobé závislosti a sníží ztrátu informací pomocí CNN.	Kombinovat rekurentní neuronové síť s konvolučními neuronovými sítěmi.	1,6 milionů Twitter tweetů	Recurrent neural networks (RNN), Convolutional neural networks (CNN) a analýza sentimentu.	Výsledky autorů ukazují spolehlivou přesnost klasifikace na různých srovnávacích souborech dat.	Článek je zaměřen pouze na jednu sociální síť Twitter.

Karren Howells and Ahmet Ertugan (2017)	Applying fuzzy logic for sentiment analysis of social media network data in marketing	Spojit fuzzy logiku spolu s analýzou sentimentu a vytvořit nástroj pro zlepšení marketingu u firem.	Analyzovat data ze sociálních médií a analyzovat je pro analýzu sentimentu.	V článku není popsán žádný soubor dat.	Fuzzy logic model	Model je speciálně zaměřen na aplikace v oblasti řízení vztahů se zákazníky, udržení zákazníků a dalších aspektů marketingu, tomto ohledu poskytuje doporučení.	Informace a modely popsané v dokumentu nejsou testovány na konkrétní datové sadě. Model jako plně funkční softwarová aplikace musí být ještě postaven.
Michel Ballings et al. (2015)	Social media optimization: Identifying an optimal strategy for increasing network size on Facebook.	Cílem tohoto příspěvku je vytvořit expertní systém, který poskytne optimální strategii pro zvýšení velikosti sítě na Facebooku.	Odhad prediktivního modelu velikosti sítě, který je následně použit v modelu.	Data byla získána od 5488 uživatelů Facebooku s 426 proměnnými.	Random Forest a Genetic Algorithm.	Z analýzy vyplývá, že velikost sítě může být v případě přijetí optimální strategie zvýšena o 61 %. Výsledky ukazují, že systém je účinný a životaschopný strategický nástroj pro zvýšení velikosti sítě.	Některé proměnné mohou být ovlivněny závislou proměnnou, již je počet přátel. Také některé proměnné mají omezený počet hodnot.
José I. Peláez et al. (2019)	Decision making in social media with consistent data	V této práci je navržen model rozhodování, jehož cílem je získat alternativní hodnocení, které by vycházelo z pocitu / názorů uživatelů zastoupených v intervalech pocitu konzistentnosti s využitím údajů získaných ze sociálních médií.	Aplikace na reálný případ a ukázat odpovídající výsledky podle trhu.	60 000 párových komentářů pro společnosti Movistar, Orange, Jazztel a Vodafone.	Consistency index CI+, ISMA-OWA, a C-OWG.	Model pro rozhodování tak, aby bylo možné získat hodnocení, které bude určovat pocity a názory na uživatele.	Automatizace rozhodovacího procesu. Autoři neznají rozhodovací kritéria, která rozhodovatelé používají ve skutečnosti.
Jose Luis Jimenez-Marquez et al. (2018)	Towards a big data framework for analyzing social media content	Navrhnout rámec pro řešení tohoto problému: první fáze je zaměřena na přípravu dat a nalezení optimálního modelu strojového učení pro data; druhá etapa se opírá o zavedení vrstvy velkých datových architektur, které jsou zaměřeny na získávání výsledků dat tím, že berou většinu modelu strojového učení první etapy.	Analyzovat velké a malé datové sady v prostředí bez velkých dat, zatímco druhá etapa analyzuje velké datové sady použitím modelu strojového učení v první fázi.	66 410 recenzí s 13 282 recenzemi pro každou hvězdu (kvalita hotelu)	Machine Learning classifiers. Multi layer perceptron (MLP), Logistic Regression, Naive Bayes a Stochastic Gradient Descent.	Multi Layer Perceptron (MLP) zobrazuje nejlepší přesnost pro tuto oblast.	Příspěvek je omezen na studium služeb cestovního ruchu, konkrétně hotelů. Nízká doba tréninku pro modely.
Jaiteng Singh and Gaurav Goyal (2018)	Anticipating movie success through crowdsourced social media videos	Vytvořit rámec pro identifikaci a kvantifikaci emocionální hodnoty každého filmového přívěsu.	Správně nastavit proměnné a funkce pro identifikaci emotivních aspektů filmového přívěsu.	Případová studie zahrnující 141 filmových trailerů vydaných od 1. ledna 2017 do 31. dubna 2018.	Dlib-ml (Machine Learning Toolkit), Genetic Algorithm (GA) a Support Vector Machine algoritmus (SVM)	Výsledky ukázaly přímou souvislost mezi emotivním skóre filmového traileru a finančními výnosy. Dále se dospělo k závěru, že emocionálně intenzivní filmový trailer by mohl mít za následek vysoké finanční výnosy ve srovnání s ne-emocionálně intenzivními přívěsy.	Malý rozsah vzorku.
Sérgio Moro et al. (2016)	Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach	Studie prezentuje výzkumný přístup využívající datamining pro predikci výkonostních ukazatelů příspěvků publikovaných na facebookových stránkách značek.	Jaké proměnné použít v konečné klasifikaci.	790 publikovaných příspěvků na Facebooku.	Linear Transformation a Support Vector Machine Model (SVM).	Autoři vypracovali rozhodovací proces z modelu „Lifetime Post Consumers“, který může doplněním informací o analýze citlivosti sloužit k podpoře rozhodnutí manažerů o zveřejnění příspěvku.	Datový soubor s 790 posty obsahuje pouze jednu společnost – celosvětově uznávanou kosmetickou značku.
Ana Carolina E.S. Lima and Leandro Nunes de Castro (2014)	A multi-label, semi-supervised classification approach applied to personality prediction in social media.	Předpověď osobnosti z Tweetu.	Vytvořit kvalitní vhodný model pro mnoho různých proměnných a témat příspěvků.	18 435 Tweetů o různých tématech.	Naive Bayes (NB) classifier, a Support Vector Machine (SVM) a Multilayer Perceptron (MLP) network.	Nejlepší model vyústil v přibližně 83% přesnou předpověď, s některými rysy osobnosti představoval lepší individuální klasifikační míry než jiní.	Omezené typy informací v souborech dat (politika, elektronika)

Byeongki Jeong et al. (2017)	Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis.	Identifikace produktových příležitostí na základě modelování témat a analýzy sentimentu sociálních dat.	Jaká klíčová slova používat v jakých dokumentech, jak určit, zda uživatelé mluví o tématu a jak jsou s ním spokojeni.	Ke studiu bylo získáno 32 637 klíčových slov a 11 123 dokumentů.	Sentiment Analysis, Opportunity algorithm.	Bylo identifikováno několik témat, kterým by se marketéři měli věnovat více než ostatním.	Pro budoucí výzkum je třeba definovat standardní stupně důležitosti a spokojenosti.
Souljanya Poria et al. (2016)	Ensemble Application of Convolutional Neural Networks and Multiple Kernel Learning for Multimodal Sentiment Analysis.	Multimodální analýza afektivních datových analýz pro extrakci názorů uživatelů a emocí z video obsahu.	Které metody jsou pro tuto studii optimální.	80 videí, 12 hodin dat rozdělených po 5minutových částech.	Kernel learning	Nejzajímavější částí této práce je, že společný multimodální rámec pro analýzu dat je schopen vytěžit emoce a sentiment z různých souborů dat.	V datasetu převažují ženy, a tak výsledky mohou být lehce zkršené.
Bernhard Kratzwald et al. (2018)	Deep learning for affective computing: text-based emotion recognition in decision support.	Cíl přizpůsobení podpory rozhodování emocionálními stavům jednotlivců.	Specifická povaha tohoto úkolu vyžaduje přizpůsobení rekurentních neuronových sítí s ohledem na obousměrné zpracování.	100 000 vzorků bylo označeno podle pozitivního, nebo negativního sentimentu.	Naive Bayes (NB), Long Short-Term Memory (LSTM) a variace. Vlastní přístup a metoda „sent2affect“.	Zlepšení výkonnosti může být dokonce až 23,2 % u F1 skóre pro klasifikaci a 11,6 % u MSE pro regresi (ve srovnání s jinými studii).	Výzkum by se měl zaměřit na vytváření větších datových souborů s cílem umožnit efektivní využití hlubokého učení (plán budoucího výzkumu autorů).
Yao Liu et al. (2017)	Using contextual features and multi-view ensemble learning in product defect identification from online discussion forums.	Autoři navrhuji novou metodu pro identifikaci vad výrobku z online fór.	Vyvinout novou metodu identifikace defektů a porovnání s existujícími modely.	10 000 diskusních příspěvků.	Naive Bayes (NB), Random Forest, Logistic Regression, Support Vector Machine (SVM), K-Nearest-Neighbors.	Výsledky z případové studie v automobilovém průmyslu ukazují, že autorova metoda zlepšila výkonnost identifikace defektů ve srovnání s existujícími metodami.	Studie je zaměřena na webové stránky pro automobilové výrobky v Číně.
Tommy Tandra et al. (2017)	Personality Prediction System from Facebook Users.	Studie se pokouší vybudovat systém, který dokáže předpovědět osobnost osoby na základě informací o uživateli Facebooku.	Správně klasifikovat osobnost člověka do osobnosti z „Big Five“.	400 Facebook profilů s více než 10 000 statusy.	Multi-Layer Perceptron (MLP), Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU) a Convolutional Neural Network (CNN) plus jejich kombinace.	Výsledky se podařilo překonat přesnost předchozích podobných výzkumů s průměrnou přesností 74,17 %.	Výzkum se skládá pouze ze 400 uživatelů Facebooku.
Shiyang Liao et al. (2016)	CNN for situations understanding based on sentiment analysis of twitter data.	Autoři navrhuji přístup k pochopení situací v reálném světě s analýzou sentimentu.	Rozumět náladě uživatelů a předvídat spokojenost uživatelů s produktem, službou nebo jinými aspekty.	12 696 vět.	Convolutional neural network (CNN) v porovnání s Naive Bayes (NB) a Support Vector Machine	Model má průměrnou přesnost 75,39 %.	Mohlo by být provedeno mnoho dalších výzkumů, například zvážení nástroje word2vec, vícevrstvé konvoluční neuronové sítě, většího souboru učicích dat a další analýzy situace nebo stavu.
M. Ghiassi et al. (2013)	Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network.	Provedení analýzy sentimentu z tweetů na Twitteru.	Správně vytvořit model pro analýzu sentimentu.	10 345 184 Twitter tweetů.	Support Vector Machine (SVM), Dynamic Architecture for Artificial Neural Networks (DAN2)	Analýza je schopna charakterizovat 97,3 % všech zpráv v 10 345 184 příspěvcích.	Všechny tweety jsou o značce Justin Bieber.
Jinpeng Chen et al. (2017)	User emotion for modeling retweeting behaviors.	Autoři navrhuji retweeting predikční rámec.	Vytvořit model detekce emocí pomocí zvažování dvou druhů emočních signálů.	22 262 tweetů a 8 467 uživatelů.	Support Vector Machine (SVM), Subdomain Generation Method (SGM), Level Sets (LS), Naive Bayes (NB)	Výsledky poukazují na to, že se autorům podařilo vytvořit model pro predikci retweetování kvalitnější než aktuálně běžně používané predikční modely.	Článek je omezen na Twitter a tweety.

Z literární rešerše vyplynulo, že jednou z nejčastěji řešených metod napříč články byla analýza sentimentu (Sentiment Analysis). Obvykle se řešila pomocí metody Support Vector Machine, Naive Bayes, Convolutional Neural Networks a Multi-Layer Perceptron.

4.3 Metody neuronových sítí využívané v rámci analýzy sentimentu

Analýza se používá pro zjišťování názorů z velkého množství dat, jimiž mohou být jak text, tak video nebo zvuk. Nejčastěji se jedná o analýzu textu, která se často používá právě pro různá odvětví marketingu, zjišťování potřeb a služeb zákazníka.

Jedním z problémů analýzy je detekce subjektivity. Uživatel může někde na sociálních sítích napsat – „Švédsko je země v Evropě“ a někdo jiný může napsat – „Švédsko je nejlepší země v Evropě“. Další problém, se kterým se museli autoři článků vypořádat, je detekce polarity, zda uživatel svůj příspěvek myslí pozitivně, či negativně. Pro vyřešení těchto problémů autoři často vyberou skupiny slov a roztrídí je do potřebných skupin. Například skupina pozitivních slov „nejlepší, super, skvělé“ a skupiny negativních slov jako „nejhorší, nemožně, nechutně“ a tyto skupiny pak přes vybrané metody strojového učení porovnávají se svým vzorkem dat, případně využijí už připravené tyto skupiny dat.

Při analýze sentimentu se autoři z rešerše uvedené výše museli rozhodnout, jestli budou k příspěvkům od uživatelů přistupovat na úrovni slov, vět nebo celých dokumentů a následně podle toho své metody přizpůsobit. Každý přístup má svoje výhody a svoje nevýhody. Přístup na úrovni slov bude mít například problém u příspěvku hodnocení – „Je to pěkně blbě.“, protože označí slovo pěkně za pozitivní a následně může nesprávně vyhodnotit názor či náladu uživatele. Problémové jsou také věty typu – „Za málo peněz hodně muziky“ či časté komentáře ve stylu „*SUPEER*!!!:-D“ a celou řadou dalších emotikonů.

Nálada nebo vztah zákazníků či uživatelů k produktu/službě se ve zkoumaných článcích provádí analýzou především pomocí textů, jako jsou příspěvky, hodnocení a komentáře (Alharbi & de Doncker, 2019a; Ducange, Fazzolari, Petrocchi, & Vecchio, 2019a; Jimenez-Marquez, Gonzalez-Carrasco, Lopez-Cuadrado, & Ruiz-Mezcua, 2019a; Lima & de Castro, 2014a; Vázquez et al., 2014a), přičemž se zde provádí i analýza sentimentu ze zvukového zdroje či videa (Poria, Peng, Hussain, Howard, & Cambria, 2017a; Singh & Goyal, 2018a). Autoři uvádí, že nejdůležitější faktor pro zjištění polarity je právě text, ale shodují se na tom, že pro odhadnutí emocí jsou nejlepší vizuální data, tedy foto nebo

video. Autoři testují přesnost svých neuronových sítí a jejich následná úspěšnost predikce někdy přesahuje 80 i 90 % (Alharbi & de Doncker, 2019a; Kratzwald, Ilic, Kraus, Feuerriegel, & Prendinger, 2018b; Sharma, Rathore, & Park, 2019a). Mezi preferované sociální sítě využívané pro výzkum jsou Twitter a Facebook. Zobrazení nejčastěji podobně zvolených metod lze nalézt v tabulce níže.

Tabulka 4: Hlavní využívané metody klasifikace a zkoumaná oblast sociálních sítí (vlastní dílo autora)

Autoři (rok)	Support Vector Machines (SVM)	Multi-layer Perceptron (MLP)	Recurrent Neural Networks (RNN)	Convolutional Neural Networks (CNN)	Naive Bayes	Long Short-Term Memory (LSTM)	Ostatní	Vybraná sociální síť
Silvia Vázquez et al. (2014)							Decision Tree (DT)	Twitter, Facebook, Forums.
Guandan Chen et al. (2018)							Tsur's, Aiello's, Hawkes, DeepCas Methods.	Twitter.
Ahmed Sulaiman et al. (2018)	*			*	*			Twitter.
Yanbing Liu et al. (2017)							Radical Basis Function (RBF)	Weibo.
Pradip Kumar Sharma et al. (2017)	*		*		*		Multilevel Deep Belief Network (DBN)	Facebook, Amazon, Google+.
Pietro Ducange et al. (2018)							Sentiment Analysis Engine (SAE)	Facebook, Twitter, Instagram.
Fazeel Abid et al. (2018)			*	*				Twitter
Karren Howells and Ahmet Ertugan (2017)							Fuzzy logic model	Twitter
Michel Ballings et al. (2015)							Random Forest and Genetic Algorithm	Facebook
José I. Peláez et al. (2019)							Consistency index CI+ and the aggregation operators ISMA-OWA, and C-OWG.	Neuvedeno
Jose Luis Jimenez-Marquez et al. (2018)		*			*	*		Facebook, Twitter, Instagram.
Jaiteg Singh and Gaurav Goyal (2018)	*							YouTube
Sérgio Moro et al. (2016)	*							Facebook
Ana Carolina E.S. Lima et al. (2014)	*	*			*			Twitter
Byeongki Jeong et al. (2017)			*				Opportunity Algorithm	Facebook, Twitter, Reddit
Souljanya Poria et al. (2016)				*			Kernel Learning	YouTube

Bernhard Kratzwald et al. (2018)					*	*		Facebook, Twitter
Yao Liu et al. (2017)	*				*		Random Forest, K-Nearest Neighbors.	Online Forums
Tommy Tandra et al. (2017)		*		*		*		Facebook
Shiyang Liao et al. (2016)	*			*	*			Twitter
M. Ghiassi et al. (2013)	*						Dynamic Architecture for Artificial Neural Networks (DAN2)	Twitter
Jinpeng Chen et al. (2017)	*				*			Twitter

Mezi sociální sítě, na nichž autoři nejčastěji prováděli výzkumy, patří Twitter. Důvodem je pravděpodobně snadnější dolování dat z této domény. Twitteru se věnovalo 13 z 22 vybraných článků. Sociální síť Facebook analyzovalo mimo jiné 8 z 22 vybraných článků a v menšině jsou potom platformy jako Instagram, YouTube, blogy, fóra či Reddit. Z dalších výzkumů bylo zjištěno, že množství využívání sociálních sítí roste s vyšším dosaženým vzděláním a stejně tak s větším příjmem domácnosti. Součástí zkoumání byla i aktivita a využívání sociálních sítí u dominantních leteckých společností. Z tohoto výzkumu bylo zjištěno, že je výhodné být aktivní na sociálních sítích, jako jsou Facebook, Instagram a Twitter, i několikrát denně (několik foto nebo video příspěvků za den, ideálně 1–3 příspěvky).

Z rešerše zároveň vyplynulo, že se autoři zaměřují na analyzování dat přímo ze sociálních sítí k určení sentimentu příspěvku, názoru na výrobek nebo službu, zjišťování defektu u výrobků a obecně názorů uživatelů ze sociálních sítí. Často je tendence sdílení zkoumána pouze na jedné sociální síti, a to primárně Twitter pro takzvané retweetování (sdílení na Twitteru). Prostor pro výzkum je v tomto ohledu využívat data z příspěvků/článků mimo sociální média a analyzovat, jak jejich struktura i sentiment ovlivňují počet sdílení a jak tento počet sdílení maximalizovat pro vybrané oblasti. Autoři se v člancích zaměřují na témata: automobilový průmysl, bankovníctví, nápoje, sport, telefonní služby, potraviny, maloobchod, filmový průmysl, finance, vzdělávání.

Z této rešerše vyplynulo, že zhruba ve třetině zkoumaných článků se autoři zabývají analýzou sentimentu a shodují se na tom, že znalost správné nálady a pocitů uživatelů je aktuálně i do budoucna klíčem ke kvalitnímu rozhodování na sociálních sítích pro expanzi na vybraném trhu podniku i pro udržení a zvyšování image firmy. Spolu s tím je spojena i loajálnost uživatelů/zákazníků. Na základě analyzovaných článků se pro účely komunikace firmy navenek nejčastěji provádějí různé variace analýzy sentimentu.

Testování je autory prováděno několika metodami, kde mezi ty nejčastěji využívané patří „Support Vector Machines (SVM)“, „Multi-layer Perceptron (MLP)“, „Recurrent Neural Networks (RNN)“, „Convolutional Neural Networks (CNN)“, „Naive Bayes“ a „Long Short-Term Memory (LSTM)“ (Jinpeng Chen, Liu, & Zou, 2017; Ghiassi, Skinner, & Zimbra, 2013a; Yao Liu, Jiang, & Zhao, 2018a; Tandra, Hendro, Suhartono, Wongso, & Prasetio, 2017a).

Ve vztahu k cíli byla provedena obsahová analýza článků/příspěvků sdílených na sociálních sítích. Analýza má za cíl zkoumat obsah příspěvku i jeho strukturu.

5 Analýza struktury příspěvků

Za účelem zjištění, jaké proměnné mohou ovlivňovat počet sdílení, bude provedena obsahová analýza a analýza sentimentu. V jazyku Python bude naprogramován algoritmus, který bude z volně dostupných internetových článků získávat informace o struktuře článků, ty bude ukládat do textového souboru, jenž bude následně upraven pro další analýzu. Po získání těchto informací bude proveden prvotní průzkum o rozložení dat a budou provedeny statistické výpočty v programu SPSS Statistics. V návaznosti budou provedeny výpočty neuronových sítí pomocí metody MLP a pro zjištění, jaké je optimální nastavení neuronové sítě pro tuto problematiku, bude také postavena neuronová síť pomocí knihovny Keras.

5.1 Prvotní analýza proměnných struktury příspěvků

Cílem testů neuronových sítí v této práci je přispět k hlubšímu pochopení chování spotřebitele (čtenáře) a jeho motivace pro sdílení na sociálních sítích. Pomocí proměnných získaných naprogramovaným algoritmem v programovacím jazyku Python spolu s vytvořenými neuronovými sítěmi budou hledány rovnice a váha různých proměnných, u kterých je pravděpodobné, že počet sdílení navyšují, či naopak snižují. Validita neuronových sítí byla testována pomocí křížové validace a testování na datech, na nichž ještě dané neuronové sítě neměly možnost trénovat. Naprogramovaným algoritmem bylo natěženo 27 proměnných popisujících článků a získáno téměř 17 tisíc článků. Otestovaných neuronových sítí bylo přes 150 v několika různých programech a mnoha různých nastaveních. Mezi hlavní typy testovaných neuronových sítí patří Multi-layer Perceptron (MLP). Zkoumána byla také analýza sentimentu textu, a to tím způsobem, že algoritmus pro získávání informací z článku zjišťoval, kolik pozitivních a negativních slov se v článku vyskytuje a jestli má počet pozitivních či negativních slov v názvu článku nebo v prvním odstavci vliv na počet sdílení. Bylo zjištěno, že v některých dnech a měsících lidé sdílí články častěji, ale nebylo možno naučit neuronovou síť přesnou kombinaci proměnných pro zvýšení počtu sdílení, i když některé proměnné mají z hlediska sdílení větší váhu než ostatní. Mezi takové proměnné patří celkový počet slov v článku, počet negativních slov v článku, počet vložených odkazů do článku nebo počet negativních slov v prvním odstavci.

Algoritmem natěžené proměnné u článků

- TitleWords = počet slov v názvu článku
- TitleCharacters = počet znaků v názvu článku
- TotalWordsInArticle = počet slov v celém článku
- TotalCharactersInArticle = počet znaků v názvu článku
- TotalParagraphs = celkový počet odstavců
- AverageWordLengthInTitle = průměrná délka slova v názvu článku
- AverageWordLengthInArticle = průměrná délka slova v celém článku
- NumberOfImages = počet obrázků v článku
- NumberOfHrefs = počet odkazů v článku
- PositiveWordsInArticle = počet pozitivních slov v článku
- NegativeWordsInArticle = počet negativních slov v článku
- PositiveDividedByNegativePolarityInArticle = počet pozitivních slov vydělený počtem negativních slov v článku (polarita článku)
- PositiveDividedByAllWordsInArticle = počet pozitivních slov vydělený počtem všech slov v článku
- NegativeDividedByAllWordsInArticle = počet negativních slov vydělený počtem všech slov v článku
- PositiveWordsInTitle = počet pozitivních slov v názvu článku
- NegativeWordsInTitle = počet negativních slov v názvu článku
- PositiveWordsInTitleDividedWordsInTitle = počet pozitivních slov v názvu děleno počtem slov v názvu článku
- NegativeWordsInTitleDividedWordsInTitle = počet negativních slov v názvu děleno počtem slov v názvu článku
- WordsInFirstParagraph = počet slov v prvním odstavci
- CharactersInFirstParagraph = počet znaků v prvním odstavci
- PositiveInFirstParagraph = počet pozitivních slov v prvním odstavci
- NegativeInFirstParagraph = počet negativních slov v prvním odstavci
- PositiveInFirstDividedByWordsInFirst = počet pozitivních slov v prvním odstavci děleno počtem slov v prvním odstavci
- NegativeInFirstDividedByWordsInFirst = počet negativních slov v prvním odstavci děleno počtem slov v prvním odstavci

- DayPosted = den v měsíci, kdy je článek přidán (1–31)
- MonthPosted = měsíc, kdy je článek přidán
- YearPosted = rok, kdy je článek přidán

Poslední zkoumaná proměnná

- NumberOfShares = celkový počet sdílení na všech sociálních sítích.

Pro získání počtu pozitivních a negativních slov byl celý analyzovaný text porovnáván slovo za slovem s externími dokumenty `positiveWords.txt` a `negativeWords.txt`. V těchto dokumentech jsou uvedena anglicky pozitivní a negativní slova. Tyto seznamy jsou volně dostupné na <https://positivewordsresearch.com/sentiment-analysis-resources/>. Po stažení souborů bylo nutné ještě dokumenty projít ručně a odstranit slova, u kterých je polarita sporná. Například slovo „joke“ bylo uvedeno jako negativní a podobně.

5.2 Obsahová analýza a analýza sentimentu

Za účelem analyzování struktury článků a vlivu struktury na počet sdílení byl vytvořen datový soubor obsahující proměnné o článcích. Dataset obsahuje 17 tisíc článků z let 2013 a 2014. Vybrány byly tyto roky pro ujištění, že se počet sdílení již nebude měnit, je tak finální. Stejně tak počet lidí na sociálních sítích stále roste a tím pádem roste i počet sdílení a kombinovat roky 2013 a například 2019 by mohlo ve výpočtech vytvářet nepřesnosti. Podle výsledků z práce bude rozhodnuto, zda natěžit i tisíce článků z dalších let a i z dalších webových stránek poskytujících různé články.

Pokud byl počet slov v prvním odstavci nula, článek se přeskočil (jednalo se například pouze o vložené video nebo obrázky). Počet znaků u proměnných je vždy včetně mezer.

Minimum slov v názvu článku představují dvě slova. Maximum slov v názvu je 16 slov. Průměr u všech 16783 článků je 8,528 slov v názvu článku.

Podle označení kategorie u článku patří mezi nejčteněji zastoupené skupiny uvedené v následující tabulce. Jedná se o slova, u kterých je v souboru více jak tisíc případů. Jeden článek může mít několik kategorií, kam spadá. V celém souboru článků je více než 7000 kategorií. Přehled nejúspěšnějších článků a oblastí lze vidět v tabulkách níže.

Tabulka 5: Nejúspěšnějších 20 článků podle počtu sdílení (vlastní dílo autora)

Název	Počet sdílení
Leaked: More Low-Cost iPhone Photos	843300
Dove Experiment Aims to Change the Way You See Yourself	690400
Kanye West Lectures at Harvard About Creativity	652900
Viral Video Shows the Extent of U.S. Wealth Inequality	617900
Roomba 880 Has More Sucking Power Trouble With Obstacles	441000
It's Hot as Hell in Australia Right Now	310800
BlackBerry Sold 1 Million BlackBerry 10 Smartphones in Q4	306100
IBM Brings Watson to the Masses and Other News You Need to Know	298400
Your Childhood Will Never Be the Same After These 12 Mashups	233400
Sprint's New Plans Guarantee Unlimited Data for Life	210300
Paul Rudd is a One Direction Fangirl at Heart	205600
What to Do With Your New Xbox One	197600
No Movie Trailer Is Complete Without This One Line	196700
McDonalds Kills Site That Advised Employees to Eat Healthy Meals	193400
Wearably Brings RSS Feeds to Google Glass	180600
All The Christmas Movies You Need in One Mashup	158900
How to Photoshop People Into Pictures	143100
Apple to Return \$32.5 Million for Accidental App Purchases	139500
How a \$6000 Video Got 6 Million Views and Launched a Business	138700
Unbelievable Ping Pong Trick Shots Paddles Amateurs	128900

Tabulka 6: Pořadí nejoblíbenějších kategorií pro všechny natěžené články a pro tisíc nejsdílenějších (vlastní dílo autora)

Pořadí	Všechny články	Tisíc nejsdílenějších
1	tech	video
2	video	tech
3	apps	culture
4	business	apps
5	world	world
6	culture	business
7	mobile	mobile
8	entertainment	entertainment
9	software	software
10	media	videos
11	videos	gadgets
12	gadgets	media
13	social	social
14	google	design
15	design	travel
16	photography	photography
17	travel	dev
18	dev	google
19	lifestyle	viral
20	viral	work

Porovnání proměnných u všech článků oproti tisícům nejlepších dle sdílení:

Tabulka 7: Porovnání průměru u proměnných pro všechny natěžené články a pro tisíc nejsdílenějších (vlastní dílo autora)

Seznam proměnných	Všechny články	Tisíc nejsdílenějších
Počet slov v názvu článku	8,53	8,55
Počet znaků v názvu článku	51,51	52,16
Počet slov v celém článku	391,72	384,02
Počet znaků v názvu článku	2344,77	2295,75
Celkový počet odstavců	11,06	11,26
Průměrná délka slova v názvu článku	5,24	5,31
Průměrná délka slova v celém článku	5	5
Počet obrázků v článku	4,78	5,51
Počet odkazů v článku	16,34	18,95
Počet pozitivních slov v článku	16,8	16,49
Počet negativních slov v článku	5,29	5,37
Počet pozitivních slov děleno počtem negativních slov v článku (polarita článku)	5,48	5,26
Počet pozitivních slov děleno všemi slovy v článku	0,043	0,044
Počet negativních slov děleno všemi slovy v článku	0,013	0,014
Počet pozitivních slov v názvu článku	0,45	0,47
Počet negativních slov v názvu článku	0,24	0,24
Počet pozitivních slov v názvu děleno počtem slov v názvu článku	0,053	0,055
Počet negativních slov v názvu děleno počtem slov v názvu článku	0,028	0,028
Počet slov v prvním odstavci	31,98	30,91
Počet znaků v prvním odstavci	189,61	183,12
Počet pozitivních slov v prvním odstavci	1,4	1,32
Počet negativních slov v prvním odstavci	0,5	0,56
Počet pozitivních slov v prvním odstavci děleno počtem slov v prvním odstavci	0,044	0,042
Počet negativních slov v prvním odstavci děleno počtem slov v prvním odstavci	0,016	0,019
Celkový počet sdílení na všech sociálních sítích	3700,15	29618,4

Bylo provedeno ještě jedno porovnání, a to tisíc nejméně sdílených oproti tisíci nejsdílenějších.

Tabulka 8: Porovnání proměnných pro tisíc nejméně sdílených článků a tisíc nejvíce sdílených článků (vlastní dílo autora)

Seznam proměnných	Tisíc nejhorších	Tisíc nejsdílenějších
Počet slov v názvu článku	8,51	8,55
Počet znaků v názvu článku	51,58	52,16
Počet slov v celém článku	403,81	384,02
Počet znaků v názvu článku	2425,63	2295,75
Celkový počet odstavců	11,307	11,26
Průměrná délka slova v názvu článku	5,27	5,31
Průměrná délka slova v celém článku	5,02	5
Počet obrázků v článku	3,88	5,51
Počet odkazů v článku	15,31	18,95
Počet pozitivních slov v článku	17,14	16,49
Počet negativních slov v článku	5,58	5,37
Počet pozitivních slov děleno počtem negativních slov v článku (polarita článku)	5,55	5,26
Počet pozitivních slov děleno všemi slovy v článku	0,042	0,044
Počet negativních slov děleno všemi slovy v článku	0,013	0,014
Počet pozitivních slov v názvu článku	0,45	0,47
Počet negativních slov v názvu článku	0,23	0,24
Počet pozitivních slov v názvu děleno počtem slov v názvu článku	0,053	0,055
Počet negativních slov v názvu děleno počtem slov v názvu článku	0,028	0,028
Počet slov v prvním odstavci	32,16	30,91
Počet znaků v prvním odstavci	192,14	183,12
Počet pozitivních slov v prvním odstavci	1,42	1,32
Počet negativních slov v prvním odstavci	0,48	0,56
Počet pozitivních slov v prvním odstavci děleno počtem slov v prvním odstavci	0,044	0,042
Počet negativních slov v prvním odstavci děleno počtem slov v prvním odstavci	0,015	0,019
Celkový počet sdílení na všech sociálních sítích	443,5	29618,4

Největší rozdíly u proměnných mezi tisíci nejsdílenějšími články a tisícem těch nejméně sdílených je hlavně v počtu obrázků. Ty nejvíce sdílené příspěvky obsahují podstatně více obrázků než ty téměř nesdílené. Další velké rozdíly jsou u počtu odkazů v článku, ty úspěšnější jich obsahují opět podstatně více.

Obráceně to je u počtu pozitivních slov v celém článku, ty úspěšnější články obsahují méně pozitivních slov. Další větší rozdíl je u slov v prvním odstavci. Úspěšnější články jich obsahují méně, to může souviset s tím, že uživatelé sociálních sítí jsou zahlceni informacemi a neradi čtou delší texty.

5.3 Statistická významnost proměnných obsahové analýzy a analýzy sentimentu

Jako první byla v rámci statistických výpočtů spuštěna korelace, při níž byla zjištěna ve spojitosti s celkovým počtem sdílení statisticky významná velmi slabá korelace u pěti proměnných. Mezi tyto proměnné patří celkový počet obrázků v článku (0,031), počet pozitivních slov v názvu (0,02), počet negativních slov v prvním odstavci (0,019), počet pozitivních slov děleno celkovým počtem slov v článku (0,016) a měsíc sdílení (0,026). I když jsou tyto korelační koeficienty statisticky významné, jedná se o velmi slabou hodnotu.

ANOVA

Po spuštění ANOVA testu pro všechny dny na celkovém vzorku téměř 17 tisíc článků nebylo prokázáno statisticky významně větší nebo menší sdílení v nějaký den v měsíci a bylo vyvráceno, že se například první, nebo poslední den v měsíci sdílí více než nějaký jiný den. I když se například v daném vzorku 19. den v měsíci sdílí téměř dvojnásobně více než poslední den v měsíci, není prokázán statisticky významný rozdíl u sdílení během různých dní.

Při spuštění dalšího ANOVA testu pro zjištění, zda existuje rozdíl v počtu sdílení v různých měsících, se ukázalo, že se v listopadu sdílí téměř dvojnásobně více než v únoru (v průměru 2 215 více sdílení na příspěvek). Stejně tak v květnu, červnu, červenci i srpnu se sdílí statisticky méně než v listopadu.

V analyzovaném souboru příspěvků bylo zjištěno, že se nejvíce sdílí v listopadu a prosinci, naopak nejméně v únoru a červnu. V případě, že bychom analyzovali čtvrtletí, nejvíce sdílení bylo zjištěno ve čtvrtém čtvrtletí (4599,62), poté v prvním čtvrtletí (3604,52) a nejméně ve třetím (3186,45) a v druhém čtvrtletí (3157,18). Přitom rozdíl mezi čtvrtým a druhým čtvrtletím činí 31,36 %.

Tabulka 9: Rozložení počtu sdílení během různých měsíců, kdy byl článek přidán (vlastní dílo autora)

Počet sdílení podle měsíců			
Měsíc přidání příspěvků	Průměr	N	Odchylka
11	5161,6293	2107	27309,63961
12	4780,7917	1181	15044,73075
3	4071,0854	2178	16066,53652
4	3765,4643	1120	21310,09341
Celkem průměr všech měsíců	3699,9644	16783	14718,14370
10	3573,8606	1363	10210,92595
1	3522,7190	1612	10844,08489
9	3403,1948	1073	6544,23255
8	3176,4452	1242	8561,07504
5	3084,0189	1218	6148,26806
7	3011,3158	1257	9566,30699
2	2946,0573	1343	6133,00116
6	2613,4151	1089	5905,75555

Regresní analýza

Při spuštění regrese v prvním kroku pro všechny proměnné je označena nejméně důležitá proměnná „průměrná délka slova v článku“ a „den sdílení“. Z modelu bylo odstraněno několik nejméně důležitých proměnných a model byl spuštěn znovu. Obě spuštění regresního modelu lze najít na konci práce v příloze. Regresní analýzou vybrané proměnné budou zmíněny dále v práci u výpočtů neuronových sítí.

Při tomto odebrání bylo také zjištěno, že podle regresní analýzy má na počet sdílení větší vliv počet negativních slov v článku než počet slov pozitivních. To samé se potvrdilo pro pozitivní a negativní slova v prvním odstavci článku.

5.4 Analýza proměnných pomocí neuronové sítě MLP

Pro zjištění, jaké proměnné nebo jejich kombinace ovlivňují nejvíce počet sdílení, bylo postaveno několik desítek neuronových sítí pomocí metody MLP. Prvotní testy neuronových sítí probíhaly v softwaru Weka, což je populární balík programů strojového učení napsaný v Javě a vyvinutý na University of Waikato, Nový Zéland. Výsledky testování jsou zobrazeny v tabulce níže. V označení sítě znamená MLP = Multi-Layer Perceptron, číslo za L označuje learning rate, L03 = Learning rate 0,3. Číslo za M označuje momentum, M02 = Momentum 0,2. Číslo za N označuje počet epoch, respektive na kolika opakováních

se daná síť učila, N10K = 10 tisíc opakování. Následně číslo v závorce poukazuje na to, jak síť sama označila svoji přesnost. Poslední údaj u označení sítě je délka trénování v minutách. Ve sloupci WEKA je potom přesnost sítě po otestování na datech, na nichž síť nemohla trénovat, ale předtím je neviděla.

Tabulka 10: Testování neuronových sítí v programu WEKA (vlastní dílo autora)

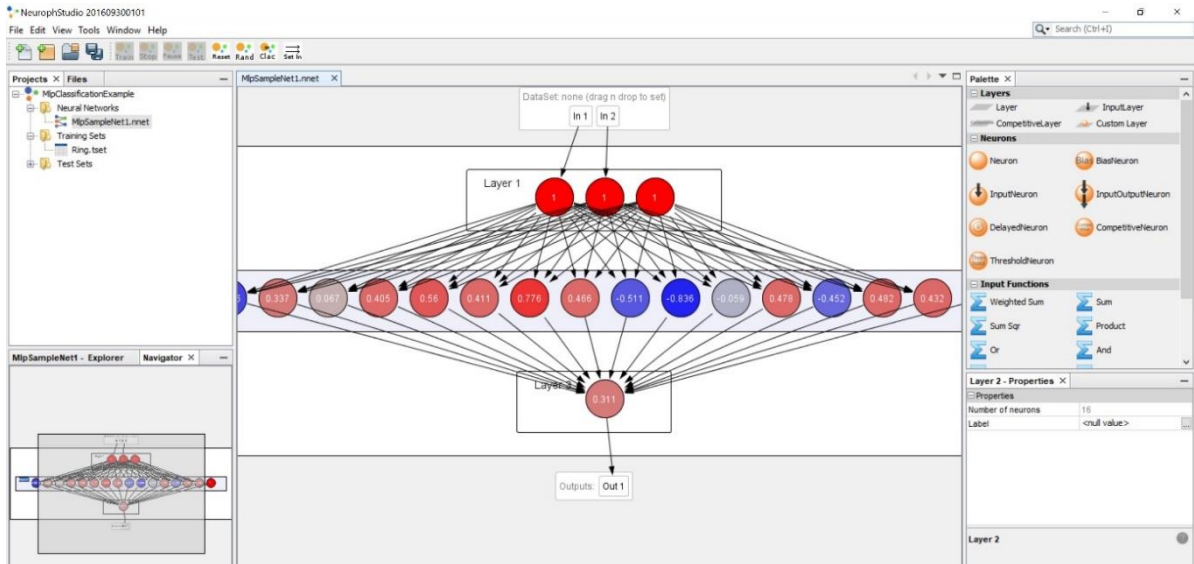
Zobrazení přesnosti různých sítí	
Nastavení sítě	WEKA Přesnost
MLP 28-0-1 L03 M02 N10K (79,33 %) 3 m	79,10 %
MLP 28-1-1 L03 M02 N10K (79,28 %) 3 m	79,10 %
MLP 28-2-1 L03 M02 N10K (79,52 %) 4 m	78,82 %
MLP 28-3-1 L03 M02 N10K (79,52 %) 5 m	78,82 %
MLP 28-4-1 L03 M02 N10K (79,42 %) 6 m	78,81 %
MLP 28-5-1 L03 M02 N10K (79,42 %) 6 m	78,81 %
MLP 28-6-1 L03 M02 N10K (79,62 %) 7 m	78,81 %
MLP 28-7-1 L03 M02 N10K (79,72 %) 7 m	78,81 %
MLP 28-8-1 L03 M02 N10K (79,62 %) 8 m	78,47 %
MLP 28-9-1 L03 M02 N10K (79,52 %) 8 m	78,47 %
MLP 28-10-1 L03 M02 N10K (79,52 %) 11 m	78,47 %
MLP 28-10-1-1 L02 M03 N10K (79,52 %) 11 m	78,47 %
MLP 28-10-2-1 L02 M03 N10K (79,72 %) 12 m	78,81 %
MLP 28-10-3-1 L002 M03 N10K (79,72 %) 13 m	78,81 %
MLP 28-10-4-1 L02 M03 N10K (79,72 %) 13 m	78,82 %
MLP 28-10-5-1 L02 M03 N10K (79,72 %) 15 m	77,99 %
MLP 28-10-6-1 L03 M02 N10K (80,17 %) 18 m	77,99 %
MLP 28-10-7-1 L002 M003 N10K (80,17 %) 18 m	77,99 %
MLP 28-10-8-1 L002 M003 N10K (80,17 %) 20 m	78,13 %
MLP 28-10-9-1 L002 M003 N10K (80,17 %) 21 m	79,10 %
MLP 28-10-10-1 L02 M03 N10K (80,47 %) 23 m	78,13 %
MLP 28-0-1 L0001 M0001 N10K (79,33 %) 3 m	79,10 %
MLP 28-1-1 L0001 M0001 N10K (79,33 %) 3 m	79,10 %
MLP 28-2-1 L0001 M0001 N10K (79,52 %) 4 m	78,82 %
MLP 28-3-1 L0001 M0001 N10K (79,52 %) 4 m	78,82 %
MLP 28-4-1 L0001 M0001 N10K (79,42 %) 6 m	78,47 %
MLP 28-5-1 L0001 M0001 N10K (79,42 %) 6 m	78,47 %
MLP 28-6-1 L0001 M0001 N10K (79,42 %) 6 m	78,47 %
MLP 28-7-1 L0001 M0001 N10K (79,72 %) 7 m	78,47 %
MLP 28-8-1 L0001 M0001 N10K (79,72 %) 8 m	77,64 %
MLP 28-9-1 L0001 M0001 N10K (79,72 %) 10 m	79,17 %
MLP 28-10-1 L0001 M0001 N10K (79,33 %) 12 m	79,17 %
MLP 28-11-1 L03 M02 N20K (80,4 %) 18 m	77,99 %

MLP 28-12-1 L03 M02 N20K (80,6 %) 19 m	77,99 %
MLP 28-13-1 L03 M02 N20K (80,6 %) 20 m	77,99 %
MLP 28-14-1 L03 M02 N20K (80,6 %) 21 m	77,64 %
MLP 28-15-1 L03 M02 N20K (80,4 %) 22 m	77,64 %
MLP 28-16-1 L03 M02 N20K (80,4 %) 24 m	78,47 %
MLP 28-17-1 L03 M02 N20K (80,4 %) 26 m	78,47 %
MLP 28-18-1 L03 M02 N20K (80,2 %) 28 m	78,47 %
MLP 28-19-1 L03 M02 N20K (80,2 %) 30 m	78,47 %
MLP 28-20-1 L03 M02 N20K (80,2 %) 32 m	78,47 %
MLP 28-20-10-1 L03 M02 N20K (81,38 %) 42 m	77,92 %
MLP 28-20-11-1 L03 M02 N20K (81,38 %) 43 m	77,92 %
MLP 28-20-12-1 L03 M02 N20K (79,72 %) 43 m	77,92 %
MLP 28-20-13-1 L03 M02 N20K (79,72 %) 43 m	78,47 %
MLP 28-20-14-1 L03 M02 N20K (79,72 %) 44 m	78,47 %
MLP 28-20-15-1 L03 M02 N20K (79,42 %) 45 m	78,47 %
MLP 28-20-16-1 L03 M02 N20K (79,42 %) 46 m	77,64 %
MLP 28-20-17-1 L03 M02 N20K (79,42 %) 47 m	77,64 %
MLP 28-20-18-1 L03 M02 N20K (79,42 %) 49 m	78,47 %
MLP 28-20-19-1 L03 M02 N20K (79,42 %) 52 m	78,47 %
MLP 28-20-20-1 L03 M02 N20K (79,42 %) 52 m	77,64 %

MLP = Multi-Layer Perceptron, číslo za L označuje learning rate, L03 = Learning rate 0,3. Číslo za M označuje momentum, M02 = Momentum 0,2. Číslo za N označuje počet epoch, respektive na kolika opakováních se daná síť učila, N10K = 10 tisíc opakování. Následně číslo v závorce poukazuje na to, jak síť sama označila svoji přesnost. Poslední údaj u označení sítě je délka trénování v minutách. Ve sloupci WEKA je potom přesnost sítě po otestování na datech, na kterých síť nemohla trénovat a předtím je neviděla.

NeurophStudio

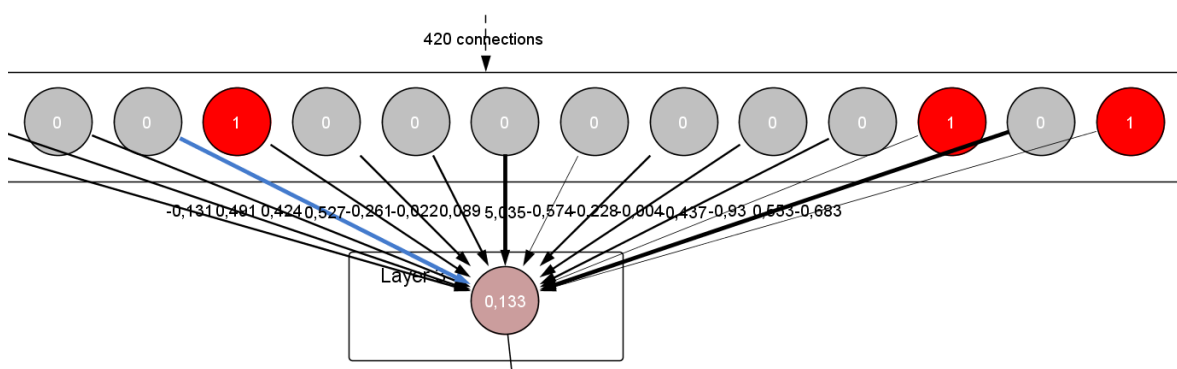
Tento software je podle listu PredictiveAnalyticsToday druhý nejlepší na analyzování neuronových sítí. V tomto softwaru se dá také postavit libovolně velká neuronová síť s jakýmkoliv počtem skrytých vrstev i uzlů. Neuronovou síť vytvořenou v tomto programu lze vidět na následujícím obrázku.



Obrázek 11: Zobrazení postavené neuronové sítě v programu NeurophStudio

Při testování se přepočítáváním vah u proměnných snaží neuronová síť minimalizovat chybovost. Síť následně označí klíčové proměnné s největší vahou a zobrazí pravděpodobnost, s jakou bude mít výstup (množství sdílení) nadprůměrnou hodnotu (čím více blíže k číslu 1, tím větší pravděpodobnost na nadprůměrné sdílení).

Rozdělení počtu sdílení podle průměru je počáteční nejjednodušší varianta. Stejně tak by se dal počet sdílení rozdělit na tři, čtyři nebo více skupina i podle toho by se dala testovat úspěšnost neuronových sítí. S každým dalším rozdělením by bylo obtížnější predikovat počet sdílení, protože by musela síť správně zařadit výsledek do určité skupiny podle počtu sdílení.



Obrázek 12: Zobrazení vyhodnocené neuronové sítě a predikce výstupu

Po natrénování sítě se dají zobrazit uzly, spojení a k nim přidělená důležitost. Při predikování více instancí v programu NeurophStudio je zobrazena hned pravděpodobnost, s jakou algoritmus určí, že se daná hodnota výstupu bude blížit k hodnotě 1 (větší sdílení příspěvku), nebo 0 (spíše menší sdílení). Tento výstup je následně potřeba upravit například v Excelu a porovnat se skutečným výsledkem a tak určit úspěšnost dané neuronové sítě.

Při zadání nových dat, pro která chceme hodnotu sdílení predikovat, je vždy následně zobrazena hodnota „Output“ a je uvedena pravděpodobnost, na kolik se daná síť s aktuálním nastavením domnívá, že se článek bude, nebo nebude sdílet nad průměrnou hodnotu.

V návrhu této práce bylo uvedeno testování v některých dalších programech, mezi něž patří například Simbrain, JustNN, Sharky Neural Network, MemBrain nebo Spice-Neuro. Neuronové sítě by se zde pravděpodobně také stejně jako dva testované programy (WEKA a NeurophStudio) předtím naučily, že nejlepší je predikovat všude menší než průměrné sdílení. Proto bylo testování změněno na podrobné testy v specializovaném programu na neuronové sítě Keras, kde se dá podrobně zkoumat celkové nastavení neuronové sítě a lépe optimalizovat než v programech, kde jsou nějaká nastavení skrytá před uživateli. V softwaru Keras si uživatel musí své nastavení „naprogramovat“ celé sám.

5.5 Analýza proměnných pomocí neuronové sítě Keras

Pro podrobnější prozkoumání, jaké nastavení neuronové sítě je pro tuto problematiku optimální, je využita volně dostupná knihovna Keras. Po naprogramování neuronové sítě v jazyku Python je možné optimalizovat nastavení pomocí úprav různých vlastností. Mezi hlavní nastavení patří:

- **Momentum** (hybnost) je koeficient, který pomáhá při učení sítě v udržení zpětného šíření algoritmu. V praxi to zhruba znamená, že udržuje učící algoritmus v dynamice učení.
- **Learning rate** je rychlost učení konfigurovatelným hyperparametrem používaným při tréninku neuronových sítí, který má malou kladnou hodnotu často v rozsahu mezi 0,0 a 1,0. Rychlost učení určuje, jak rychle je model přizpůsoben problému.

Vysvětlení dalších možných nastavení neuronových sítí:

Optimalizační algoritmy pro neuronové sítě popisují metody, jak zapracovat vypočítaný gradient do modelu. Zde je uveden přehled některých jednoduchých optimalizačních algoritmů, které lze použít.

- **Stochastický gradientní sestup** (Stochastic gradient descent, SGD) – SGD je asi nejpoužívanější optimalizační algoritmus, protože dovede efektivně urychlit učení. Ideou metody je, že chceme upravovat váhy podle nejobvyklejšího gradientu, který se určí jako průměr všech gradientů v dávce. Navíc pokud jsou trénovací příklady nezávisle a stejně rozloženy, odpovídá tento průměrný gradient skutečnému gradientu na cost funkci.
- **RMSprop** – Efektivní algoritmus pro hluboké neuronové sítě vynalezený T. Tielemanem není citlivý na hyperparametry, i když zavádí nový hyperparametr ρ . Kombinuje použití znaménka gradientu a úpravu kroku pro každý parametr zvlášť. Idea je, že pokud se znaménka posledních dvou gradientů shodují, zvětší krok úpravy vah. Pokud ne, sníží krok úpravy vah. Dále je vhodné určit minimum a maximum pro velikost kroku. Aby ale metoda fungovala dobře i pro dávkové učení, je nutné aplikovat kvadratický průměr na každý gradient.
- **Adaptivní gradient (AdaGrad)** – Další z algoritmů, který upravuje každý parametr zvlášť. V podstatě má tato metoda takový efekt, že zvětšuje malé gradienty a zmenšuje velké gradienty. Pro hluboké neuronové sítě ale není tento algoritmus moc efektivní, protože snižuje velikost kroku příliš brzy.
- **AdaDelta** - Metoda, která se pokouší řešit problémy metody adaptivního gradientu s příliš agresivním snižováním learning rate. Při použití této metody se nepoužívá learning rate, její funkcionalitu si řeší algoritmus sám, nicméně potřebuje zavést konstantu k , pro lepší formování jmenovatele. Počet hyperparametrů tak fakticky zůstává stejný.
- **Adam** – Adam byl navržen jako kombinace adaptivního gradientu a RMSprop. Metoda má v podstatě v sobě implementovanou funkcionalitu momenta. Adam se ve svém výpočtu odstiňuje bias z gradientů, dále používá dvě cache pro gradienty ($cg1$, $cg2$), hyperparametr korigující velikost kroku (v podstatě learning rate) α , dále je hyperparametr ρ pro každý gradient zvlášť a konstantu k . Adam má výhodu v tom, že jeho hyperparametry obecně nepotřebují složité ladění.

Aktivační funkce (systém přepočítávání vah u uzlů):

- **Usměrněná lineární funkce** (Rectified linear unit, ReLU) – Velmi často používaná aktivační funkce. Funkce není shora ohraničená a díky tomu umí síť odhalit míru chyby.
- **Sigmoida** – Běžně používaná aktivační funkce. Nejčastěji se používá sigmoida se střední hodnotou v bodě [0,0]. Takové sigmoidy jsou liché funkce a díky svému tvaru výborně modelují výstup (vstup do další vrstvy), protože jeho hodnota bude okolo nuly. Jejich problémem ale je, že vytvářejí téměř plochou oblast u středu chybové funkce, což nejenom může zpomalovat učení, ale také může způsobit mizející gradient v hlubších sítích. K sigmoidě lze přidat i parametr α , který ovlivňuje zakřivení sigmoidy. Tento parametr lze učit a adaptovat.
- **Hyperbolický tangens** – Další z běžně používaných funkcí. Má podobné vlastnosti jako sigmoida. U většiny klasifikačních úloh konverguje rychleji jak sigmoida.
- **Softsign** – Funkce podobná hyperbolickému tangentu, ale má polynomiální asymptoty. Je robustnější, co se týká inicializace vah.
- **Softplus** – Jedná se v podstatě o vyhlazenou ReLU funkci a díky tomu je bez problémů diferencovatelná.

Mezi další možná nastavení patří optimální výběr skrytých vrstev a počet neuronů v nich, velikost využívané paměti počítače, počet iterací učení či rychlost identifikace nepotřebných neuronů.

První test s nastavením sítě 27-10-10-10-1 se neuronová síť znovu naučila, že bezpečná cesta, jak mít kvalitní úspěšnost, je všude predikovat menší počet sdílení než průměr a to není vhodný výstup, takže je potřeba síť dále upravit.

Kvalita v modelu se popisuje pomocí chyby neboli ztrátovosti (loss a val_loss) a pomocí přesnosti (acc a val_acc). V případě „val“ hodnot se jedná o testování na vzorku stranou, na kterém se síť neučí.

```

9100/12366 [=====>.....] - ETA: 0s - loss: 3.3706 - acc: 0.7909
9810/12366 [=====>.....] - ETA: 0s - loss: 3.3698 - acc: 0.7909
10530/12366 [=====>.....] - ETA: 0s - loss: 3.3430 - acc: 0.7926
11240/12366 [=====>.....] - ETA: 0s - loss: 3.3283 - acc: 0.7935
11930/12366 [=====>.....] - ETA: 0s - loss: 3.3520 - acc: 0.7920
12366/12366 [=====] - 1s 73us/step - loss: 3.3394 - acc: 0.7928
Epoch 83/150

 10/12366 [.....] - ETA: 2s - loss: 3.2236 - acc: 0.8000
 730/12366 [>.....] - ETA: 0s - loss: 3.2899 - acc: 0.7959
1440/12366 [==>.....] - ETA: 0s - loss: 3.4587 - acc: 0.7854
2130/12366 [====>.....] - ETA: 0s - loss: 3.3598 - acc: 0.7915
2800/12366 [====>.....] - ETA: 0s - loss: 3.2812 - acc: 0.7964
3490/12366 [=====>.....] - ETA: 0s - loss: 3.2698 - acc: 0.7971
4210/12366 [=====>.....] - ETA: 0s - loss: 3.2428 - acc: 0.7988
4920/12366 [=====>.....] - ETA: 0s - loss: 3.3186 - acc: 0.7941
5630/12366 [=====>.....] - ETA: 0s - loss: 3.3324 - acc: 0.7933
6330/12366 [=====>.....] - ETA: 0s - loss: 3.2975 - acc: 0.7954
7040/12366 [=====>.....] - ETA: 0s - loss: 3.3152 - acc: 0.7943
7710/12366 [=====>.....] - ETA: 0s - loss: 3.2947 - acc: 0.7956
8400/12366 [=====>.....] - ETA: 0s - loss: 3.3023 - acc: 0.7951
9100/12366 [=====>.....] - ETA: 0s - loss: 3.3263 - acc: 0.7936
9810/12366 [=====>.....] - ETA: 0s - loss: 3.3058 - acc: 0.7949
10530/12366 [=====>.....] - ETA: 0s - loss: 3.3262 - acc: 0.7936
11240/12366 [=====>.....] - ETA: 0s - loss: 3.3369 - acc: 0.7930
11930/12366 [=====>.....] - ETA: 0s - loss: 3.3195 - acc: 0.7940
12366/12366 [=====] - 1s 73us/step - loss: 3.3394 - acc: 0.7928
Epoch 84/150

 10/12366 [.....] - ETA: 2s - loss: 1.6118 - acc: 0.9000
 720/12366 [>.....] - ETA: 0s - loss: 3.5370 - acc: 0.7806
1430/12366 [==>.....] - ETA: 0s - loss: 3.3927 - acc: 0.7895
2150/12366 [====>.....] - ETA: 0s - loss: 3.3811 - acc: 0.7902
2850/12366 [====>.....] - ETA: 0s - loss: 3.3480 - acc: 0.7923
3540/12366 [=====>.....] - ETA: 0s - loss: 3.3830 - acc: 0.7901
4240/12366 [=====>.....] - ETA: 0s - loss: 3.3947 - acc: 0.7894
4960/12366 [=====>.....] - ETA: 0s - loss: 3.3796 - acc: 0.7903

```

Obrázek 13: Postupné učení pomocí algoritmu v Keras

Po zmenšení sítě na 27-14-1 a zvýšení „batch-size“ na 32 má síť úspěšnost 76 % na neznámých datech.

Bylo potřeba rozdělit dataset místo 80 na 20 (podle průměru), celý vzorek na 50 na 50, aby síť automaticky nesklouzávala k tomu predikovat všude jednu hodnotu (menší než průměr).

Následně byla provedena série testů pro variantu se všemi proměnnými a další série pro variantu s výběrem několika proměnných podle IMB SPSS Statistics regresní analýzy. V druhém případě se jedná o variantu s 16 proměnnými. Kategorické proměnné a proměnné, které jsou podle regresní analýzy nejméně důležité, byly odebrány.

Protože hodnota acc ve většině výsledků předbíhá hodnotu val_acc, znamená to, že se síť nedokáže z aktuálních dat naučit souvislosti. Pro další experimenty a testování bylo podle regresní analýzy z programu SPSS odebráno několik proměnných.

Mezi proměnné, které zůstaly pro další testování patří:

- Počet slov v celém článku;

- Počet znaků v názvu článku;
- Počet odkazů v článku;
- Počet pozitivních slov v článku;
- Počet negativních slov v článku;
- Počet pozitivních slov děleno počtem negativních slov v článku (polarita článku);
- Počet pozitivních slov děleno všemi slovy v článku;
- Počet negativních slov děleno všemi slovy v článku;
- Počet negativních slov v názvu článku;
- Počet pozitivních slov v názvu děleno počtem slov v názvu článku;
- Počet negativních slov v názvu děleno počtem slov v názvu článku;
- Počet znaků v prvním odstavci;
- Počet pozitivních slov v prvním odstavci;
- Počet negativních slov v prvním odstavci;
- Počet pozitivních slov v prvním odstavci děleno počtem slov v prvním odstavci;
- Počet negativních slov v prvním odstavci děleno počtem slov v prvním odstavci;
- Celkový počet sdílení na všech sociálních sítích.

Znovu byla provedena stejná série testů a podle úspěšnosti bylo vybráno, zda se zaměřit a blíže prozkoumat variantu se 27 proměnnými, nebo variantu s 16 proměnnými.

Ve všech devíti testech se ukázala jako přesnější neuronová síť s větším počtem proměnných, a tak byly následně spuštěny testy, jak optimalizovat celkově síť a její nastavení přesně zaměřené na tento problém.

Tabulka 11: Porovnání přesnosti testů pro model s 27 proměnnými oproti modelu s 16 proměnnými (vlastní dílo autora)

27 Proměnných		16 Proměnných
loss: 0.6573 - acc: 0.5935 - val_loss: 0.7618 - val_acc: 0.5344	VS	loss: 0.6820 - acc: 0.5492 - val_loss: 0.6987 - val_acc: 0.5101
loss: 0.6231 - acc: 0.6364 - val_loss: 0.7360 - val_acc: 0.5461	VS	loss: 0.6738 - acc: 0.5752 - val_loss: 0.7139 - val_acc: 0.5040
loss: 0.6402 - acc: 0.6181 - val_loss: 0.7345 - val_acc: 0.5303	VS	loss: 0.6502 - acc: 0.5913 - val_loss: 0.7727 - val_acc: 0.5061
loss: 0.6538 - acc: 0.5997 - val_loss: 0.8209 - val_acc: 0.5348	VS	loss: 0.6852 - acc: 0.5508 - val_loss: 0.7245 - val_acc: 0.5121
loss: 0.6203 - acc: 0.6076 - val_loss: 1.0889 - val_acc: 0.5445	VS	loss: 0.6817 - acc: 0.5448 - val_loss: 0.7444 - val_acc: 0.5000
loss: 0.6653 - acc: 0.5899 - val_loss: 0.7181 - val_acc: 0.5497	VS	loss: 0.6839 - acc: 0.5523 - val_loss: 0.6935 - val_acc: 0.5085
loss: 0.6181 - acc: 0.6149 - val_loss: 1.1300 - val_acc: 0.5303	VS	loss: 0.6759 - acc: 0.5616 - val_loss: 0.7171 - val_acc: 0.5117
loss: 0.6653 - acc: 0.5866 - val_loss: 0.7069 - val_acc: 0.5449	VS	loss: 0.6837 - acc: 0.5490 - val_loss: 0.6947 - val_acc: 0.5125
loss: 0.6139 - acc: 0.6362 - val_loss: 0.7599 - val_acc: 0.5392	VS	loss: 0.6932 - acc: 0.5036 - val_loss: 0.6930 - val_acc: 0.5125

Nástroj Keras v programovacím jazyku Python je schopen pomocí knihoven scikit-learn a KerasClassifier testovat najednou několik různých nastavení neuronových sítí a následně zobrazit porovnání i přesnost pro nejlepší nastavení. Jedná se o sedmikrokový proces a po následném optimalizování každé části byla postavena optimální neuronová síť pro tento daný problém.

Tabulka 12: Optimalizační proces neuronové sítě pomocí Keras (vlastní dílo autora)

Krok	Cíl	Výstup	Výběr
1	Zjistit optimální velikost sítě (počet neuronů ve skryté vrstvě) pro daný problém - Tune the Number of Neurons in the Hidden Layer.	Best: 0.521753 using {'neurons': 25}	Jako nejlepší se ukázala síť s 25 neurony ve skryté vrstvě.
2	Zjistit „Batch-size“ (velikost paměti) a počet epoch (iterací učení) – Tune Batch Size and Number of Epochs.	Best: 0.545609 using {'batch_size': 40, 'epochs': 50}	Nejlépe vyšla batch_size na hodnotu 40 a počet epoch na 50. Velikost batch_size byla následně upravena, ale počet epoch pro větší vzorek zůstal větší než 50.
3	Zjistit vhodný optimalizační algoritmus – Tune the Training Optimization Algorithm.	Best: 0.506874 using {'optimizer': 'RMSprop'}	Z několika možných optimalizačních algoritmů byl vybrán RMSprop.
4	Zjistit nejlepší rychlost učení a momentum – Tune Learning Rate and Momentum.	Best: 0.557011 using {'learn_rate': 0.001, 'momentum': 0.9}	Pro tuto problematiku vyšla rychlost učení nejlépe pro hodnotu 0,001. Jedná se o nízkou hodnotu a velmi pomalé učení. Nejlepší momentum připadlo na hodnotu 0,9.
5	Zjistit ideální množství neuronů, které jsou časem nepotřebné – Tune Dropout Regularization.	Best: 0.512939 using {'dropout_rate': 0.3, 'weight_constraint': 5}	Hodnota dropout, tedy jak rychle má algoritmus přestat počítat s některými neurony vyšla na 0,3.
6	Zjištění optimální aktivační funkce – Tune the Neuron Activation Function.	Best: 0.455119 using {'activation': 'tanh'}	Jako nejlepší aktivační funkce se s neuspokojivou hodnotou úspěšnosti 0,455 ukázala funkce „tanh“.
7	Zjistit vhodnou inicializaci vah – Tune Network Weight Initialization.	Best: 0.544234 using {'init_mode': 'lecun_uniform'}	Nejlepší byl označen mód „lecun_uniform“.

Podle těchto nastavení byla následně vytvořena neuronová síť (podle zjištěných optimálních parametrů pro tento problém).

```
# neural network with keras – Disertační práce Hruška
```

```
from numpy import loadtxt
import numpy
import matplotlib.pyplot as plt
from keras.models import Sequential
from keras.layers import Dense
from keras.optimizers import SGD
from keras.layers import Dropout
from keras.constraints import maxnorm
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
```

```
# load the dataset
```

```
= loadtxt('KerasRozdeleno50-50.csv', delimiter=',')
```

```
# split into input (X) and output (y) variables
```

```
X = dataset[:,0:27]
```

```
y = dataset[:,27]
```

```
optimizer = SGD(lr=0.001, momentum=0.9)
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)
```

```
scaler = StandardScaler()
```

```
X_train = scaler.fit_transform(X_train)
```

```
X_test = scaler.fit_transform(X_test)
```

```
seed = 7
```

```
numpy.random.seed(seed)
```

```
# define the keras model
```

```
= Sequential()
```

```
model.add(Dense(25, input_dim=27, kernel_initializer = 'lecun_uniform', activation='tanh',  
kernel_constraint=maxnorm(5)), Dropout(0.3))
```

```
model.add(Dense(1, activation='sigmoid', kernel_initializer = 'lecun_uniform'))
```

```
# compile the keras model
```

```
model.compile(loss='binary_crossentropy', optimizer=optimizer, metrics=['accuracy'])
```

```
history = model.fit(X_train, y_train, validation_data=(X_test, y_test), batch_size=40,  
epochs=40000, verbose=0)
```

```
model.save('modelKerasSuper-Last.h5')
```

```
# evaluate the keras model
```

```
_, accuracy = model.evaluate(X, y)
```

```
print('Accuracy: %.2f' % (accuracy*100))
```

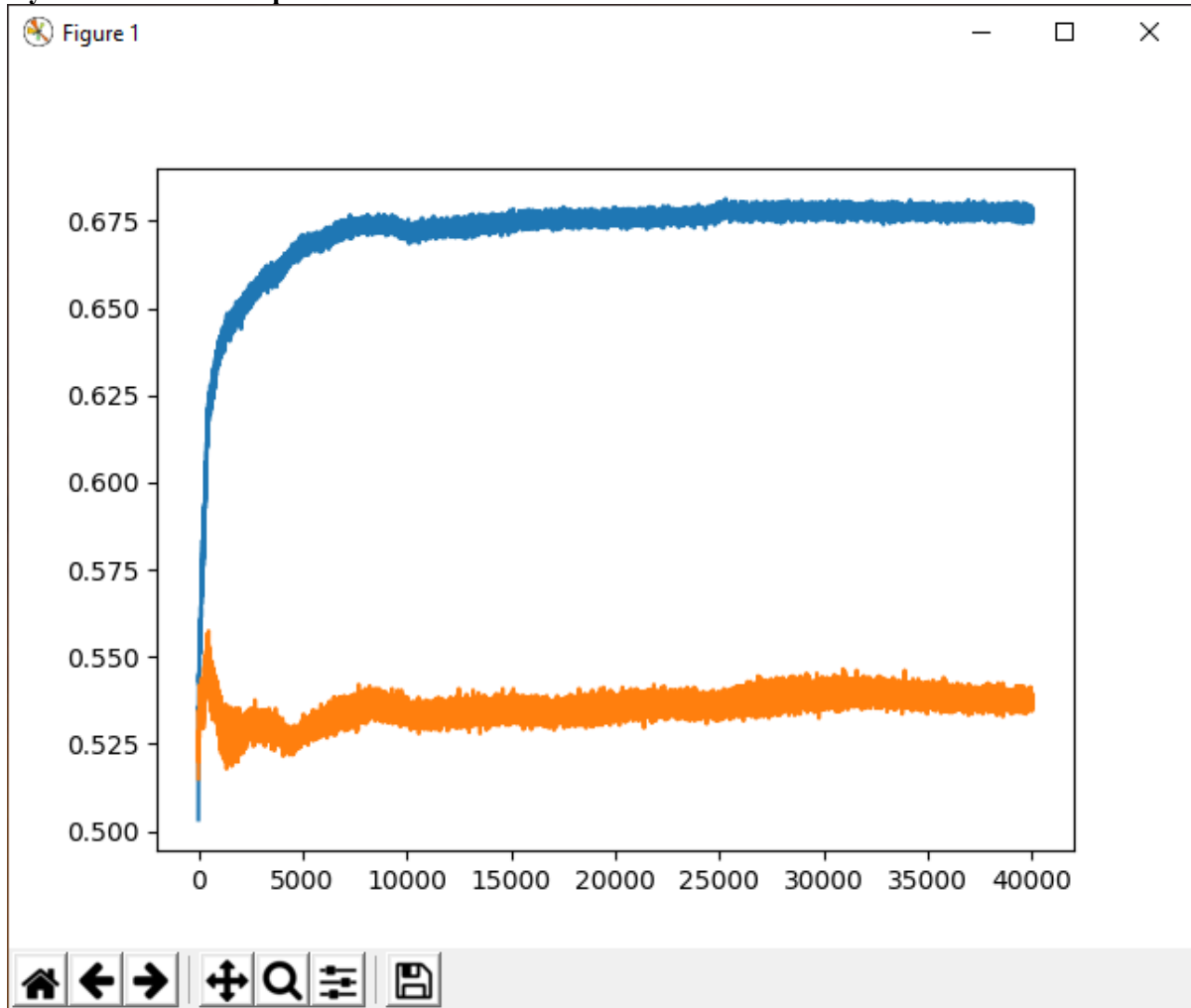
```
# Zobrazení grafu
```

```
plt.plot(history.history['acc'])
```

```
plt.plot(history.history['val_acc'])
```

```
plt.show()
```

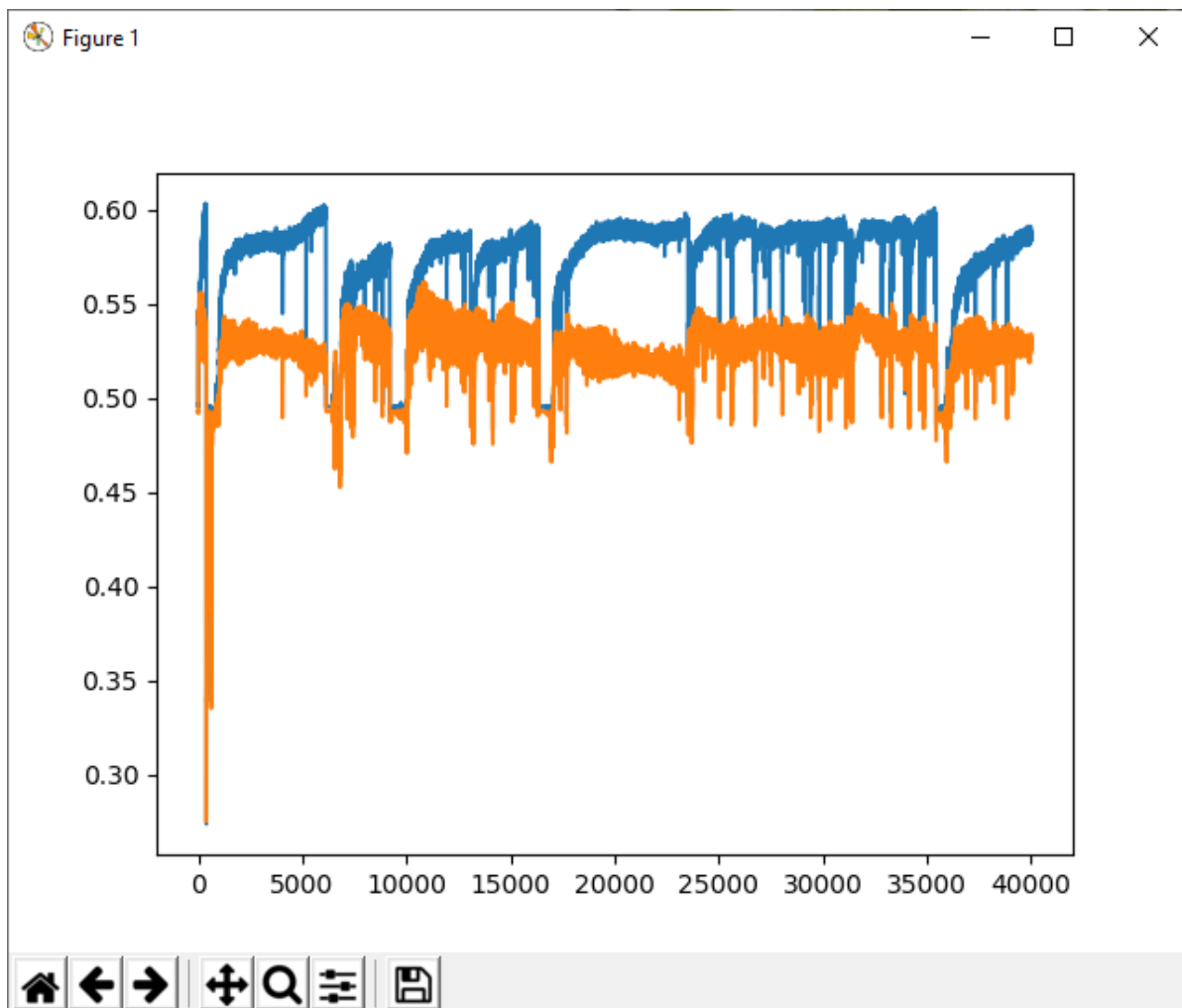
Výsledek na 40 tisíc epoch:



Obrázek 14: Učení se neuronové sítě na 40 tisíc opakování s optimálním nastavením pro tento problém (vlastní dílo autora)

Test 2 podle nastavení jako druhé v pořadí.

```
model = Sequential()
model.add(Dense(25, input_dim=27, kernel_initializer = 'uniform', activation='softsign',
kernel_constraint=maxnorm(3))), Dropout(0.6)
model.add(Dense(1, activation='softsign', kernel_initializer = 'uniform'))
# compile the keras model
model.compile(loss='binary_crossentropy', optimizer=optimizer, metrics=['accuracy'])
history = model.fit(X_train, y_train, validation_data=(X_test, y_test), batch_size=64,
epochs=40000, verbose=0)
```



Obrázek 15: Učení se neuronové sítě na 40 tisíc opakování s druhým nejlepším nastavením pro tento problém (vlastní dílo autora)

Po celkových testech se všemi proměnnými a po testech s vybranými proměnnými podle regreseční analýzy byl vybrán pro další detailnější testování model se všemi proměnnými, protože byl napříč všemi devíti neuronovými sítěmi přesnější. Následně bylo optimalizováno sedm různých nastavitelných parametrů v nastavení neuronové sítě. Po těchto testech byla postavena optimální síť pro tento problém. I po následném testování dvou nejlepších nastavení podle optimalizačních testů nebyla síť schopna se naučit a efektivně predikovat počet sdílení. Z těchto výsledků se dá usoudit, že vybrané proměnné mají jenom velmi malý vliv na celkový počet sdílení. Hlavním faktorem pro větší sdílení může být následně celá řada v tomto výzkumu neznámých proměnných, které se velmi těžko získávají. Mezi ně může patřit například aktuální naladění čtenáře, jeho frekvence sdílení, věk, pohlaví, jestli je doma, či v práci, jak na čtenáře působí uvedené obrázky a videa v článku a další proměnné tohoto typu.

V případě složitějších neuronových sítí a určitých nastavení sítí predikuje pravděpodobnost, s jakou nastane u počtu sdílení hodnota 1 = pravděpodobně bude více sdíleno. Při otestování pravděpodobnosti nad 60 %, nad 70 %, nad 80 % nebo nad 90 % se žádné sítí nepodařilo dostat nad úspěšnost 80 % při rozložení sdílení podle průměru (zastoupení ve vzorku 80 %/20 %) a stejně tak se nepodařilo získat lepší úspěšnost než 55 % při rozložení sdílení na dvě stejně velké poloviny.

Ze zjištěných proměnných a postavených neuronových sítí nebyla zjištěna ideální kombinace nastavení proměnných pro pravděpodobné zvýšení počtu sdílení. Bylo otestováno, že neuronové sítě se s aktuální velikostí datasetu a vybranými proměnnými nedokáží naučit predikovat počet sdílení. Některé sítě měly během testování úspěšnost predikce okolo 79 %, ale to je z toho důvodu, že přidělily většině článků variantu „nebude se sdílet více jak průměr“ a těch je v daném výběru zhruba 80 %. Statistické testy poukázaly na fakt, že tyto proměnné mají šanci zvýšit počet sdílení:

- Celkový počet slov v článku;
- Počet odkazů v článku;
- Počet negativních slov v článku;
- Počet negativních slov v prvním odstavci;
- Měsíc sdílení.

V případě těchto proměnných se jedná o velmi malou váhu, a tudíž se na ně nelze spolehnout, z toho důvodu nelze v tomto případě efektivně predikovat úspěšnost sdílení. Bylo zde potvrzeno takzvané Parretovo pravidlo, že jenom 20 % článků se dostane nad průměr sdílení a má šanci na virální šíření. Ostatních 80 % článků není potom tolik sdíleno.

Předchozí výzkum a s ním spojená systematická rešerše odhalila, že tyto sítě by mohly být pro tento problém vhodné řešení, to se při testování v této práci nepotvrdilo.

Mezi limity této analýzy patří zaměření na menší počet internetových stránek poskytujících články a počty sdílení jejich článku. Mezi limity práce může také patřit zaměření prozatím jenom na roky 2013 a 2014 a neanalyzování celého textu složitějšími algoritmy, mezi něž patří například TF/IDF (Tf-idf je metodika hodnocení relevance při vyhledávání textu). Název je spojením zkratkou dvou termínů: Term Frequency – četnost slova v dokumentu Inverse document frequency – převrácená četnost slova ve všech dokumentech).

S limity práce jsou spojena možná rozšíření do budoucna. Bylo by vhodné získat další články z různých webových stránek a jejich počet sdílení z dalších let a provést podobné testy a rozšířit je o nějaké ještě komplexnější.

Shrnutí

V rámci výše uvedených testů proběhly aktivity v návaznosti na stanovené kroky vedoucí k dosažení cíle:

- Specifikace klíčových znalostí a informací ve vazbě na stanovený cíl, tedy efektivní externí komunikaci firmy. Specifikace byla prováděna v kontextu vybraných témat (budováním dobrého jména firmy, budování důvěry, inovace, propagace/marketing, nábor zaměstnanců, informování (slevy, plány, kupóny, soutěže) a inspirace (příspěvky, které nejsou nutně spojeny s firmou).
- Provedení analýzy sentimentu ve vazbě na sdílené příspěvky.
- Provedení obsahových analýz.
- Určení faktorů, které v určitých případech ovlivňovaly sdílení příspěvků. Ty je nutné více rozpracovat, aby bylo možné testovat reakci uživatelů spolu s tím, jak se mají proměnné v příspěvku a struktura příspěvku měnit pro různé firmy a vybrané oblasti.
- Určení nedůležitých aspektů ve struktuře příspěvků, které mají nulový nebo minimální vliv na množství reakcí od uživatelů.
- Naprogramování algoritmu pro získání potřebných proměnných.

6 Dotazníkové šetření – analýza chování na sociálních sítích a motivace uživatelů pro sdílení

Vzhledem k slabším závislostem zjištěným v předchozích kapitolách bylo provedeno dotazníkové šetření pro ověření získaných informací a byl vytvořen přehled, co uživatelé sdílí a co je ke sdílení motivuje k popsání jejich chování na sociálních sítích.

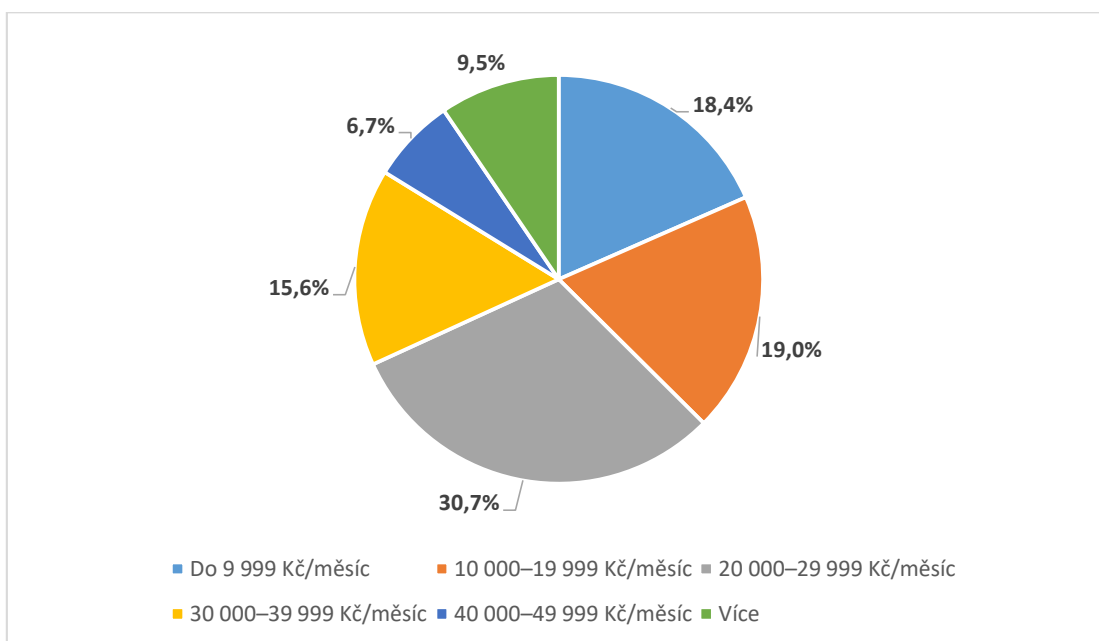
Dotazník obsahoval celkem 18 otázek a k jeho vyplnění bylo získáno 197 respondentů. Otázky byly rozděleny na dva okruhy. První okruh je charakteristika respondentů a druhý je zaměřen především na téma, co uživatele motivuje ke sdílení, tedy jestli je nějak ovlivňují pozitivní, nebo negativní slova v nadpisech a textu nebo co nejčastěji sdílejí. Dotazník byl vytvořen jako podpora pro řešení i metody strojového učení. Dotazník měl také za úkol poukázat na to, s jakou frekvencí uživatelé sdílí.

Jednou podmínkou pro respondenty bylo to, aby měli vlastní účet alespoň na jedné sociální síti. Dotazník byl šířen primárně po sociální síti Facebook a sběr dat probíhal v období duben až srpen 2020. Struktura dotazníku je uvedena v přílohách práce.

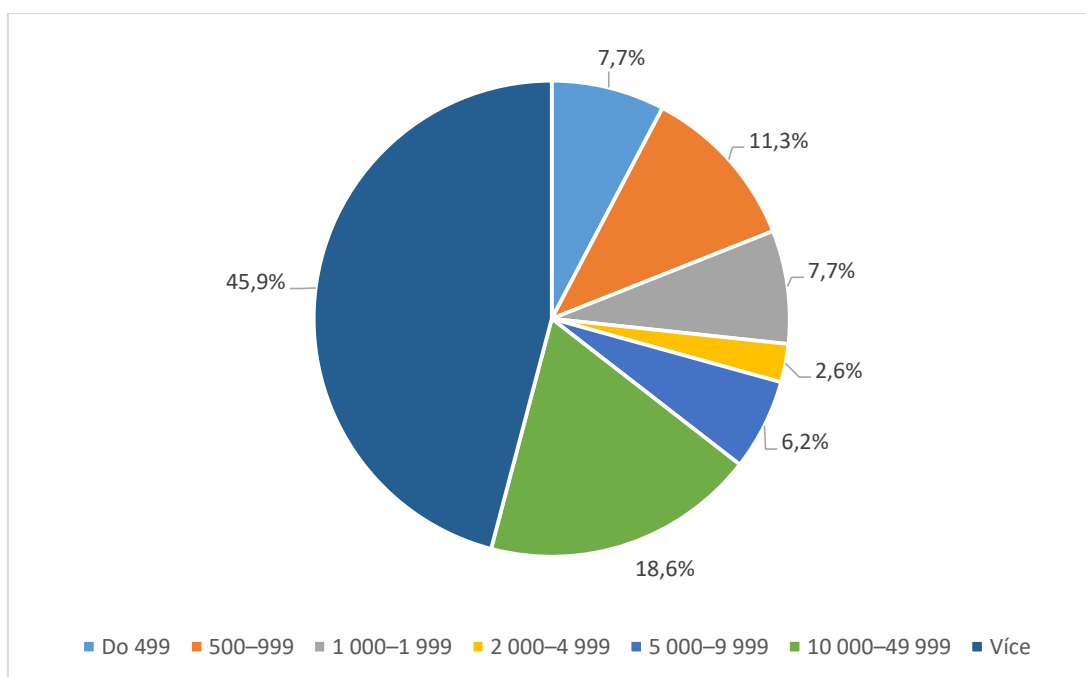
6.1 Rozdělení respondentů

Mezi respondenty je 35,5 % mužů a 64,5 % žen s průměrným věkem 29,43 let. Napříč všemi respondenty bylo 71,1 % svobodných, 25,4 % vdaných/ženatých a 3,6 % rozvedených. Z celkového počtu respondentů bylo 48,2 % s dokončeným vysokoškolským vzděláním, 36 % s dokončeným středoškolským vzděláním s maturitou, 10,7 % se základním a 5,1 % se středoškolským bez maturity. Celkem 30,4 % respondentů stále studuje.

Rozdělení respondentů podle výše příjmů a místa bydliště je následující:



Obrázek 16: Zobrazení rozložení respondentů podle čistého měsíčního příjmu (vlastní dílo autora)



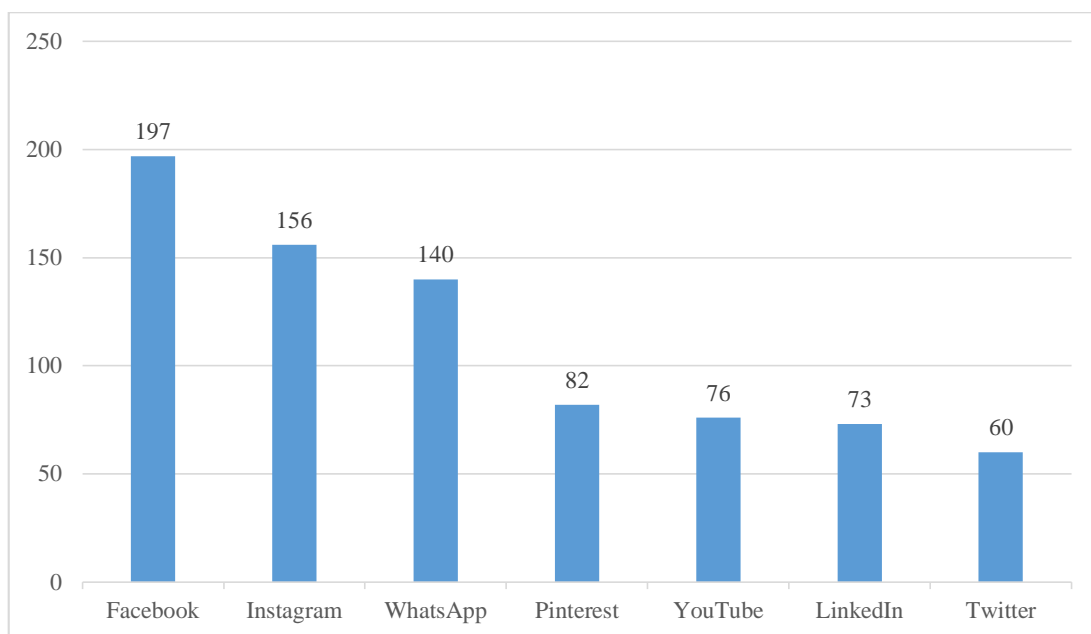
Obrázek 17: Zobrazení rozložení respondentů podle počtu obyvatel místa bydliště (vlastní dílo autora)

Největší zastoupení u kategorií příjmu připadlo na rozpětí 20–30 tisíc korun měsíčně (30,7 %), nejméně zastoupená byla skupina 40–50 tisíc měsíčně (6,8 %). U počtu obyvatel

místa bydliště byla respondenty nejvíce označena kategorie s více jak 50 tisíci obyvatel a nejméně zastoupená skupina je 2–5 tisíc obyvatel.

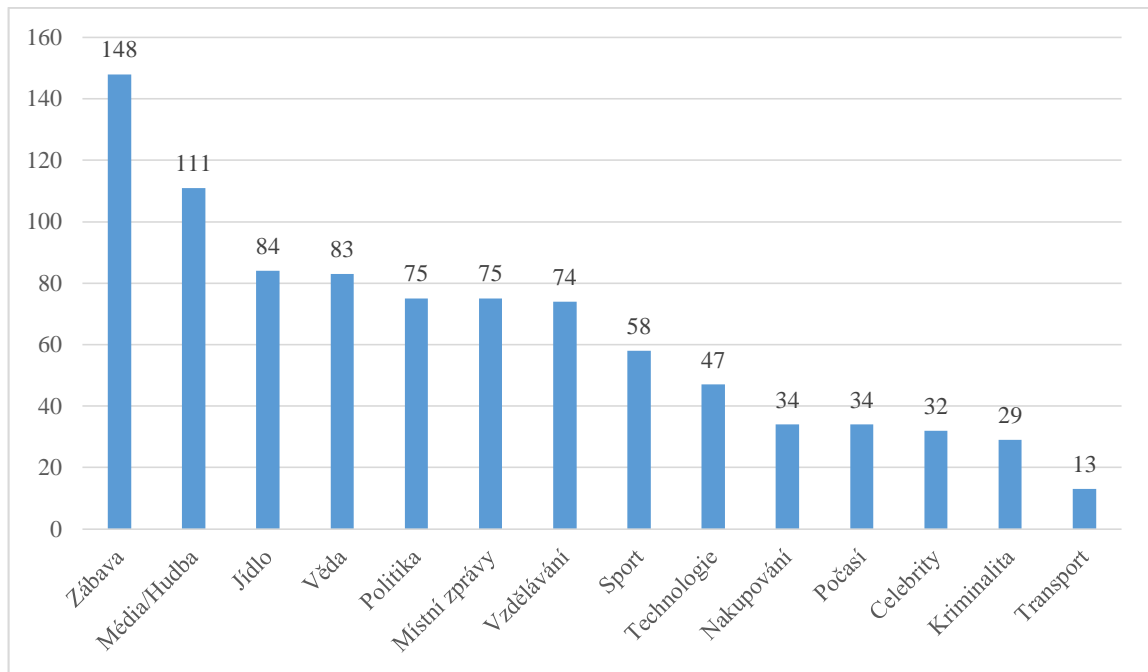
6.2 Využívané sociální sítě

Rozdělení používání sociálních sítí lze vidět na následujícím Obrázek 18.



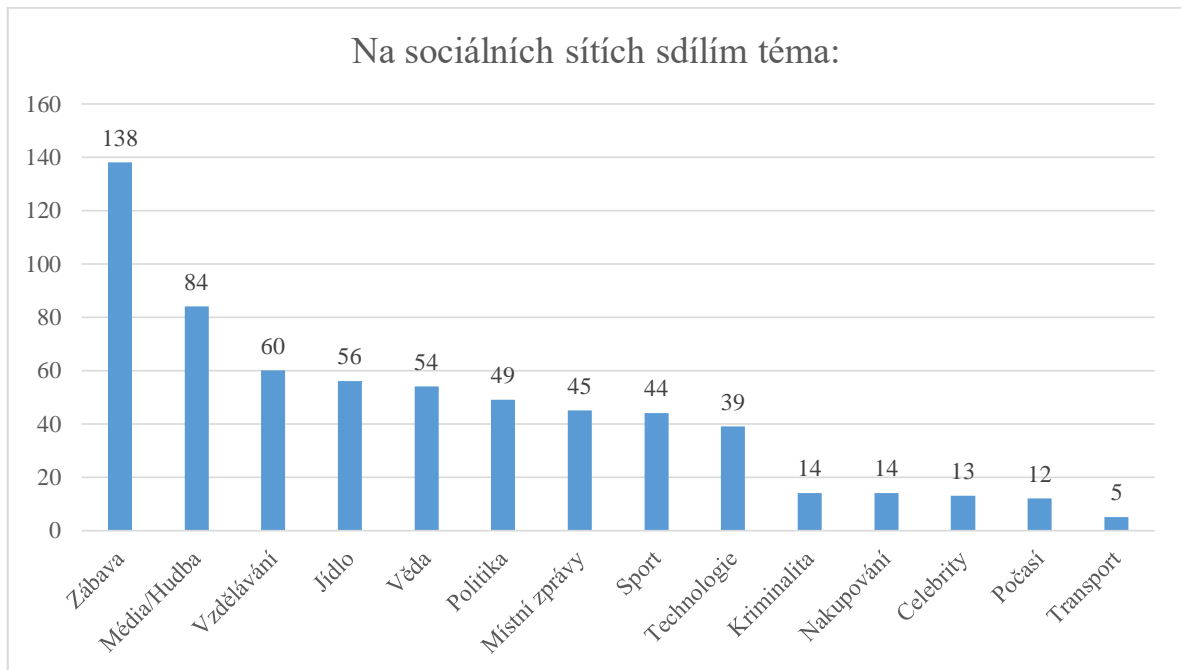
Obrázek 18: Zobrazení rozložení používaných sociálních sítí u respondentů (vlastní dílo autora)

Při položení otázky, jaké téma na sociálních sítích řeší, respondenti uvedli nejčastěji, že zábavu se zastoupením 77,5 % a na druhém místě média, hudbu nebo TV se zastoupením 58,1 %. Další rozložení lze vidět na Obrázek 19.



Obrázek 19: Zobrazení hlavních oblastí, na které respondenti reagují (vlastní dílo autora)

U této otázky byla možnost vlastní odpovědi, kde respondenti zmiňují v minoritním zastoupení například víru, zvířata, děti, knihy, zahrady nebo auto/moto. Pro porovnání s tím, co respondenti řeší na sociálních sítích a co sdílí, byla později v dotazníku položena podobná otázka. U zábavy se poměr moc nezměnil, ale u některých skupin, jako je například politika, se poměr výrazně snížil. Tuto otázku lze vidět na obrázku níže.



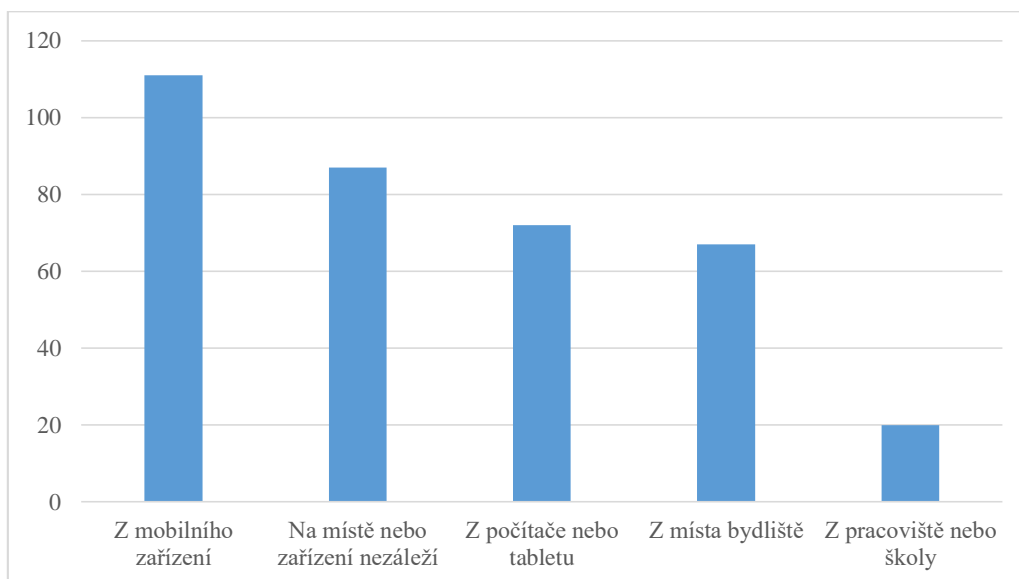
Obrázek 20: Zobrazení hlavních oblastí, které respondenti sdílí (vlastní dílo autora)

6.2.1 Motivace sdílení příspěvků

Nejčastěji respondenti jako autora příspěvku, který sdíleli, uvádějí přátele, které osobně znají, se zastoupením 59,5 %, následně sdílení příspěvků přímo od nich, jako jsou například selfie, foto rodiny a podobně se zastoupením 56,4 %, a poté zcela náhodné autory, které vidí někde na jakékoli sociální síti, a to 55,9 %.

Při položení otázky, co uživatele nejvíce motivuje ke sdílení, byla reakce, že se chtějí podělit o radost a pobavit ostatní se zastoupením 75,9 %, následováno sdílením pro kontakt s přáteli, chtít se podělit s přáteli o právě zjištěné informace – s výskytem 55,4 %, poté sdílením příspěvků pro pochlubení se (např. dovolenou, úspěchem, romantickou večeří...) s 45,6 %.

Pro zjištění, odkud a z jakého zařízení respondenti nejčastěji sdílejí, byla uvedena následující otázka – viz Obrázek 21 na následující stránce.



Obrázek 21: Vyobrazení, z jakého zařízení a odkud uživatelé sociálních sítí sdílí (vlastní dílo autora)

U otázky ohledně ověřování informací 53,6 % respondentů uvedlo, že si občas ověřují informace z příspěvků na sociálních sítích, 30,6 % je pravidelně kontroluje, 13,3 % si informace většinou neověřuje a 2,6 % pak nikdy neověřuje informace z příspěvků na sociálních sítích. Respondenti v průměru sdílí 155,44 příspěvků za rok.

Při prozkoumání průměrů počtu sdílení a stavem respondenta lze vidět, že nejvíce sdílejí rozvedení, následují svobodní a poté ženatí/vdaní sdílí nejméně.

6.3 Souhrn výsledků z dotazníkového šetření

V návaznosti na zkoumané téma lze uvést, že v dotazníku pouze 61 respondentů jasně uvedlo, že pokud příspěvek obsahuje negativitu nebo vulgarismy, získá více jejich pozornost. U optimistických a radostných příspěvků uvedlo 123 respondentů, že příspěvek upoutá jejich pozornost. Je složité jednoznačně definovat, jak takové příspěvky mohou vypadat, protože možností pro jejich strukturu je nespočet, ale potvrzují to výpočty z neuronových sítí, že motivace ke sdílení závisí na mnoha různých faktorech, z nichž jedním je motivace ke sdílení. Co vede uživatele ke sdílení, zajímá organizace po celém světě už od počátků sociálních sítí. Z dotazníků i dalších zdrojů (Ballings et al., 2016a; Ducange et al., 2019a) vyplývá, že sociální sítě jsou primárně o zábavě, a i proto u otázky, co je hlavní motivací pro sdílení, nejvíce respondentů odpovědělo, že chce pobavit přátele a podělit se o radost a v určitých dalších případech poznat názor kamarádů.

Při prozkoumání posledních sdílených příspěvků u respondentů převažují zábavné příspěvky, ale odpovědi jsou velmi variabilní. Protože analýza byla prováděna v roce 2020, je velmi často sdílen virus COVID-19 jako aktuální situace, jež se musí řešit a je nebezpečná. Také je často uveden jako téma sdílení sport, výlet, vlastní myšlenka, domácí mazlíček (45,6 % respondentů uvedlo, že sdílejí, protože se chtějí pochlubit, a tak i tuto metodu využívají často organizace pro rozdávání dáreků a sdílení foto s výhercem a podobně).

7 Návrh konceptuálního modelu externí komunikace firmy na sociálních sítích

V následující kapitole budou popsána východiska pro konceptuální model spolu s klíčovými prvky v rámci modelu externí komunikace. Popsány budou také klíčové procesy v souvislosti s externí komunikací.

7.1 Východiska modelu

Současný návrh modelu vychází z poznatků teorie informačního managementu, externí komunikace na sociálních médiích. Mezi východiska pro návrh modelu patří prvotní systematická rešerše na aktuálně využívané metody pro zlepšení externí komunikace, statistické výpočty pro sdílené příspěvky na sociálních sítích, postavené neuronové sítě a klasifikační modely pro sdílené příspěvky, obsahová analýza a analýza sentimentu spolu s podpůrným dotazníkovým výzkumem a zjištění z vlastních odborných publikací.

Východiska modelu na základě systematické rešerše dle standardu PRISMA

Na základě analyzování 1415 odborných publikací byly zjištěny nejčastěji využívané metody pro podporu rozhodování na sociálních sítích. Mezi tyto metody patří Support Vector Machines (SVM), Multi-layer Perceptron (MLP), Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN), Naive Bayes a Long Short Term Memory (LSTM). Autoři napříč články využívali mnoho dalších metod a jejich modifikace, ale zde výše zmíněné jsou ty nejvíce zastoupené. Z těchto zjištění bylo rozhodnuto využít pro analýzu metodu MLP. Mezi vstupy u této metody patří získané proměnné naprogramovaným algoritmem a výstupy z této analýzy jsou dále v práci využity pro model externí komunikace.

Východiska modelu na základě výpočtů neuronových sítí a klasifikačních metod

Na základě postavených neuronových sítí a následně spuštěných milionů výpočtů nebyla z vybraných proměnných (viz předcházející podkapitola 4.3) zjištěna velká spojitost s celkovým počtem sdílení příspěvků. To platí i pro kombinace proměnných. Bylo zjištěno několik proměnných, které mohou mít z vybraných ukazatelů největší potenciál ovlivnit počet sdílení.

Mezi tyto proměnné patří:

- Celkový počet slov v článku;
- Počet odkazů v článku;
- Počet negativních slov v článku;
- Počet negativních slov v prvním odstavci;
- Měsíc sdílení.

Postavených a otestovaných neuronových sítí bylo přes 150 a výsledky z těchto analýz jsou použity dále v práci u modelu optimalizace externí komunikace.

Východiska modelu na základě analýzy sentimentu a obsahové analýzy

V rámci disertační práce bylo analyzováno, jak velký vliv má zastoupení pozitivních a negativních slov v nadpisu a textu příspěvku. Bylo zjištěno, že zastoupení negativních slov má potenciál lehce ovlivnit množství sdílení, ale nebylo to statisticky potvrzeno a ani výsledky z výpočtů neuronových sítí toto zjištění nepotvrdily. Je možné, že výsledky na jiném, a ještě větším vzorku by odhalily něco jiného, ale pro následný model externí komunikace nebude doporučováno příliš se této problematice věnovat, ale naopak se má zaměřit na jiné aspekty, které mají větší potenciál oslovit uživatele sociálních sítí pro sdílení příspěvku.

V tomto ohledu nebylo potvrzeno, že by množství pozitivních nebo negativních slov v článku mělo velký vliv na celkový počet sdílení.

Východiska modelu na základě podpůrného dotazníkového výzkumu

Z dotazníkového šetření vyplynulo potvrzení předešlých výpočtů z neuronových sítí. Respondenti ve většině případů uvádějí, že je jim lhostejné, jestli jsou uvedena pozitivní nebo negativní slova, pokud to není přehnané. Stejně tak uvedli, že nejčastěji sdílí příspěvky, které je pobaví, tedy kategorii zábava. Respondenti také v průměru sdílí 155 příspěvků za rok. Mezi další oblasti, kterým se respondenti věnují a jež sdílí, patří média/hudba, věda, politika a gastronomie. Respondenti také uvedli, že nejčastěji sdílí příspěvky na sociálních sítích z mobilního zařízení a že mezi hlavní využívané sociální sítě patří Facebook, Instagram a WhatsApp.

Z tohoto pohledu budou zjištění využita v modelu externí komunikace v procesu „definice formy sdělení“.

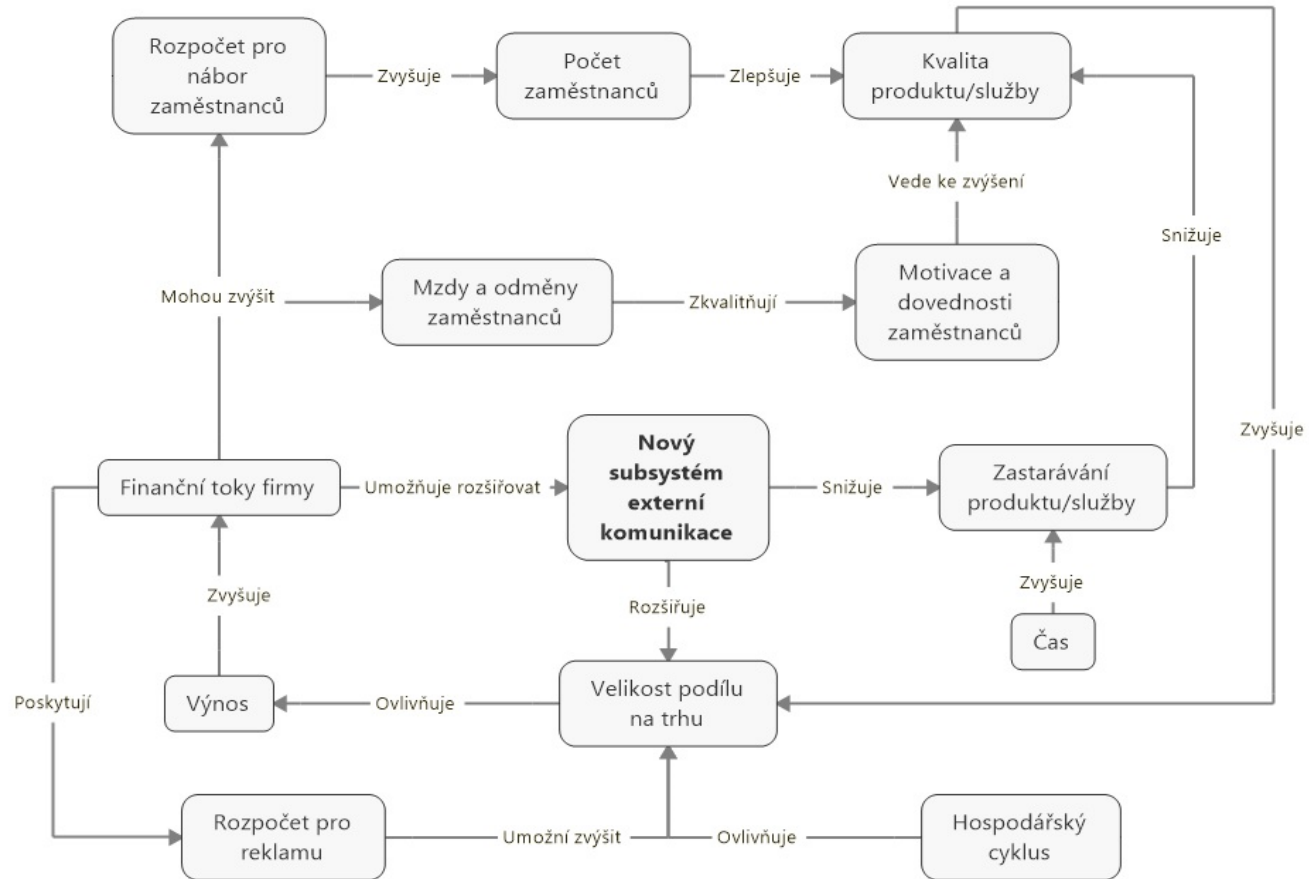
Shrnutí východisek pro uvedené metody je uvedeno v následující tabulce:

Tabulka 13: Shrnutí východisek pro tvorbu modelu a doporučení (vlastní dílo autora)

Oblast výzkumu	Otázka	Zjištění
Systematická rešerše dle standardu PRISMA	Jaké metody jsou nejčastěji využívány pro podporu rozhodování na sociálních sítích?	Mezi tyto metody patří Support Vector Machines (SVM), Multi-layer Perceptron (MLP), Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN), Naive Bayes a Long Short Term Memory (LSTM)
Neuronové sítě a klasifikační metody	Ovlivňuje struktura příspěvku množství sdílení?	Výpočty u vybraných proměnných nepotvrdily ovlivňování počtu sdílení. Nejvíce ovlivňující počet sdílení byly označeny následující proměnné: celkový počet slov v článku, počet odkazů v článku, počet negativních slov v článku, počet negativních slov v prvním odstavci a měsíc sdílení. Nejúspěšnější články také využívaly v průměru více obrázků v textu.
Analýza sentimentu a obsahová analýza	Ovlivňují pozitivní a negativní slova počet sdílení příspěvku?	Počet sdílení u příspěvku není primárně ovlivňován zastoupením pozitivních a negativních slov. Bylo poukázáno na to, že zastoupení negativních slov má větší váhu pro počet sdílení než pozitivní slova.
Podpůrný dotazníkový výzkum	Co uživatele motivuje ke sdílení a co se nejčastěji sdílí?	Jako nejčastější oblast, kdy uživatelé sdílí, byla uvedena „zábava“ a motivace ke sdílení „chci se podělit o radost, pobavit ostatní“. Také bylo zjištěno, že uživatelé nejčastěji sdílí příspěvky z mobilního zařízení.

7.2 Externí komunikace ve firmě a její vazby na vnější prostředí

Obrázek 22: Dopady zavedení nové kvalitní externí komunikace na související proměnné a činnosti firmy (Obr. 23) popisující souvislosti tvorby a zavedení externí komunikace do firmy. Obrázek 22: Dopady zavedení nové kvalitní externí komunikace na související proměnné a činnosti firmy ukazuje základní interakce mezi zavedením nové metodiky externí komunikace s možným cash flow a výkonem na trhu. Udržení vysokého podílu na trhu je klíčové pro dlouhodobý úspěch organizace. Zásadní roli ve firmě hraje lidský kapitál, a to na všech úrovních. Motivace má velký vliv na kvalitu produktů a služeb, což platí všemi úrovněmi pozic ve firmě, včetně těch oddělení, která se zaměřují na externí komunikaci.



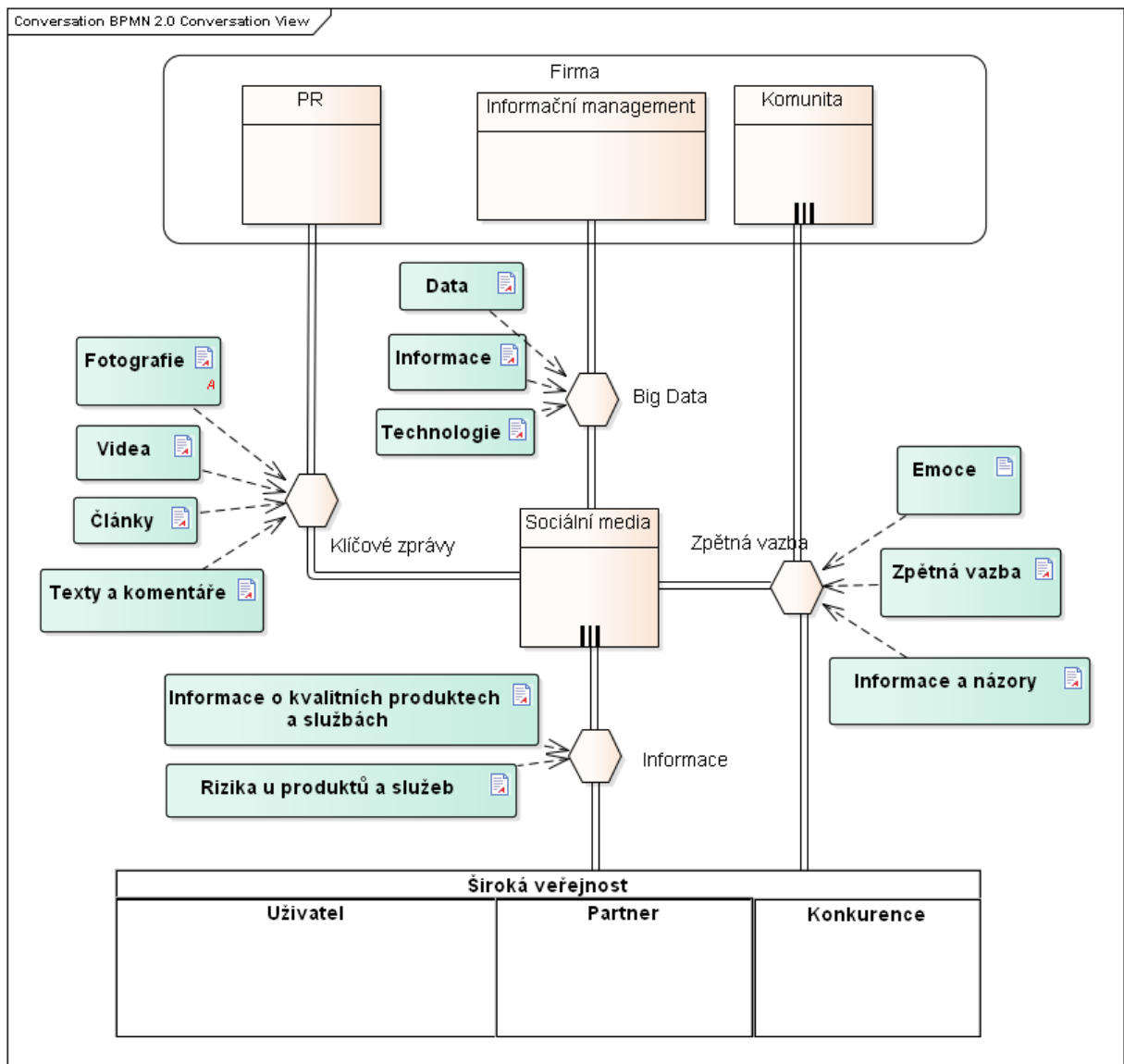
Obrázek 22: Dopady zavedení nové kvalitní externí komunikace na související proměnné a činnosti firmy (vlastní zpracování inspirováno (AbdelShafy et al., 2015))

Autor se ve své práci věnuje samotnému subsystému externí komunikace, který je analyzován a popsán v následujících kapitolách.

7.3 Externí komunikace a toky informací a znalostí ve firmě

Analyzováním dat od konkurence, vlastních úspěchů i neúspěchů z minulosti, příspěvků na sociálních sítích mohou podniky získat informační výhodu na trhu. Modely a přístupy zpracované v této práci by mohly být pro podnik funkčním východiskem ke zlepšení efektivity z hlediska většího sdílení obsahu jeho příspěvků a zajištění více interakcí na sociálních sítích, což by mu pomohlo rozvíjet jeho strategii a jednodušeji dodržovat firemní vize. Externí komunikace je veškerá komunikace, kterou firma vyvíjí navenek, s její pomocí může komunikovat se zákazníky, dodavateli a obchodními partnery, akcionáři, médii, státními a kontrolními orgány, investory a bankami, potenciálními zaměstnanci, místním společenstvím či širokou veřejností. Tyto skupiny jsou nazývány jako zájmové skupiny.

Při zaměření ještě o jednu úroveň konkrétněji do interakce firmy, sociálních médií a široké veřejnosti lze poukázat na to, jaké klíčové entity v tomto propojení jsou (Obrázek 23: Klíčové entity a propojení firmy, sociálních médií a široké veřejnosti (vlastní dílo autora)).



Obrázek 23: Klíčové entity a propojení firmy, sociálních médií a široké veřejnosti (vlastní dílo autora)

Při vložení podniku a jeho chování na sociálních sítích do širšího měřítka lze vidět cyklus vkládání informací, reakci na sociálních sítích, zpětnou vazbu a poučení podniků s jejich následnou aktivitou.

Na internetu je k dispozici velké množství volně dostupných veřejných informací ve formě fotografií, videí, článků, různých blogů a textů všeho druhu. Podniky z těchto informací mohou analyzovat chování trhu, konkurence i spotřebitelů a podle své podnikové strategie a informačního managementu si vytvořit propagační materiál například ve formě prezenčního webu, e-shopu, příspěvku na sociálních sítích nebo videa. Z tohoto podnikem vytvořeného propagačního materiálu se následně stává také veřejná informace, kterou může buď sám podnik, nebo jakýkoliv uživatel sdílet na libovolnou sociální síť a tím se informace dostane k dispozici

dále pro širokou veřejnost. S postupem času uživatelé na tento příspěvek reagují a sdělují si názory s ostatními uživateli. Touto interakcí a zapojením komunity vzniká pozitivní, negativní nebo případně neutrální zpětná vazba, a to nejenom na sociálních sítích, ale k šíření dochází i prostřednictvím mluveného slova ve školách, při práci, na návštěvách a společných akcích. Reagováním tvoří uživatelé další obsah například na svoje blogy, na svůj prostor na sociálních sítích a tím sdělují svůj názor, nálady a emoce svému okolí. Z těchto názorů a emocí si udělají podniky představu, jak se jejich sdělení, produkt nebo služba veřejnosti zamlouvá a adaptují se podle reakcí uživatelů i podle konkurence. Vytvoří další příspěvek ze zjištěných informací a celý cyklus se znovu opakuje, firma je tak v neustálém vývoji tlačena názory uživatelů, zákazníků i konkurencí. Cílem pro podniky v tomto ohledu je vytvořit kvalitní vztahy s uživateli, maximalizovat dosah svých sdělení a tím i prodeje nebo využívání svých služeb při vhodném rozložení lidských i finančních zdrojů. Cíle musí korespondovat se strategií, vizí a posláním podniku.

Při zapojení neuronových sítí do tohoto cyklu lze analyzovat již vytvořené příspěvky, články, novinky i videa a podle toho, jak uživatelé reagují, komentují a sdílejí, analyzovat s jistou mírou pravděpodobnosti, na co reagují více a na co méně. Touto analýzou je možné zjistit, zda uživatelé sociálních sítí více reagují a sdílejí příspěvky plné optimismu a dobrého naladění, nebo jestli naopak sdílejí články, které v nich vyvolávají nenávisť či strach. S tím je spojena i struktura článků. Analyzován by měl být počet slov v nadpisu příspěvku, počet slov v příspěvku, počet vložených obrázků nebo videí, zvolená klíčová slova, počet odstavců, počet znaků, emotikonů a podobně. Z této analýzy se dá následně vytvořit doporučení, jakým příspěvkům by měl daný podnik věnovat největší pozornost a jaký obsah vytvářet. Na nastavené neuronové síti se provedou výpočty na trénovací množině z vložených dat, která činí zpravidla 80 % z celkového objemu dat a na zbylých 20 % se provede test kvality modelu. Pokud chybovost modelu MSE (Mean Square Error) nepřesahuje hodnotu 0,5, lze model využít pro upravení vytvářeného příspěvku. Na získaných datech byla provedena korelace, jak silně ovlivňuje každá proměnná počet sdílení, a následně také ANOVA test pro zjištění, jak se liší množství sdílení pro určité téma nebo jak se množství sdílení liší pro den v týdnu nebo měsíc v roce. Na zkoumaných příspěvcích byla provedena obsahová analýza s ohledem na počet sdílení článku. Obsahová analýza byla použita pro redukci textu a pro věcné pořádkání informací: systematické a předmětové třídění.

7.4 Klíčové prvky v rámci modelu externí komunikace

Na sociálních sítích lze v kontextu řešeného cíle identifikovat celou řadu relevantních entit významných při externí komunikaci firmy, jako jsou například skupiny uživatelů, podniků, jednotlivé zainteresované skupiny subjektů (stakeholdeři). V práci bude následovat definice klíčových znalostí, entit a informací v rámci externí komunikace.

Za klíčové entity budou tedy v procesu externí komunikace na sociálních sítích považováni:

- Uživatelé;
- Konkurence;
- Partneři;
- Sociální síť;
- Veřejné informace.

Mezi klíčové informace v procesu externí komunikace na sociálních sítích z hlediska strategie podniku patří:

- Strategie podniku, vize, mise, globální cíle, analýza vnějšího prostředí;
- Cíl sdělení;
- Cílová skupina;
- Typ sociální sítě;
- Forma sdělení;
- Komunikace se zákazníkem.

Z hlediska **plánování** v podniku:

- Předchozí strategie;
- Konkurenční postavení;
- Vedení;
- Angažovanost;
- Historický vývoj.

Mezi klíčové aktivity na sociálních sítích patří:

- Sběr informací;
- Sdílení informací;
- Analýza informací (aktuálních i z minulosti);
- Budování důvěry;
- Rozložení lidských i finančních zdrojů;
- Analýza zpětné vazby;
- Plánovací proces pro externí komunikaci, strategie podniku, tvorba strategického a operačního plánu.

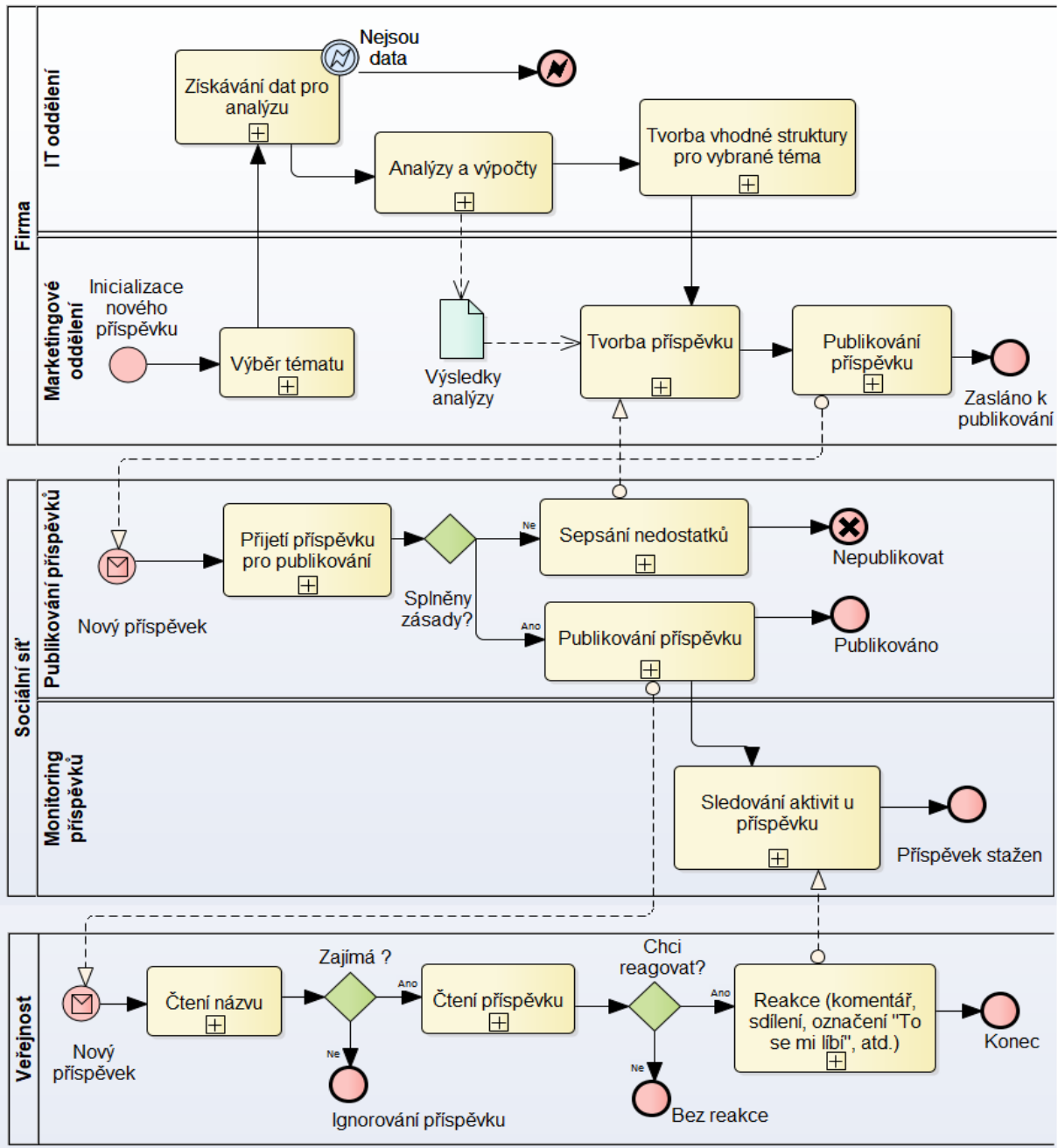
Zmíněné entity a aktivity vychází z rešerše odborné literatury, která se již promítla do obrázků: Obrázek 5: Interakce podniků se sociálními médii a veřejností (vlastní zpracování inspirováno (Eroshkin et al., 2017; Jue et al., 2010; Mayer, 2009; Phua et al., 2017; Prajogo et al., 2018)) Obrázek 6: Klíčové oblasti informačního managementu (vlastní dílo autora inspirováno (Eroshkin et al., 2017; Prajogo et al., 2018)) Obrázek 9: Klíčové aktivity externí komunikace na sociálních sítích (vlastní zpracování inspirováno (Ashley & Tuten, 2015; Drury, 2008; Hofacker & Belanche, 2016)) Obrázek 23: Klíčové entity a propojení firmy, sociálních médií a široké veřejnosti (vlastní dílo autora)

7.5 Vybraný proces metody externí komunikace

Provádět správně aktivity na sociálních médiích je pro podniky komplexní a velmi důležitý úkol. Přidané příspěvky mohou mít pro podnik obrovskou přidanou hodnotu jako málokterý marketingový nástroj v minulosti. Stejně jako může mít příspěvek kvalitní odezvu a přínos, může být pro firmu při nesprávném provedení destruktivní a z důvodu rychlého šíření informací po sociálních sítích obtížné takovou chybu napravit. Podniky musí na sociálních sítích správně rozvrhnout jak lidské zdroje, které se této oblasti budou věnovat, tak finanční zdroje optimální pro daný podnik a jeho vizi i strategii. Podniky musí velmi pečlivě zkoumat a analyzovat reakce uživatelů na jejich aktivitu a podle toho upravovat svoje budoucí chování. Uživatelé a zákazníci na sociálních sítích je nezbytné vnímat jako plnohodnotné partnery a tak se k nim chovat, přijímat zpětnou vazbu, reagovat s rozmyslem na reakce uživatelů, ať už jsou jakkoliv negativní, a v případě, že negativní komentáře převažují, je potřeba rychle změnit aktuální chování na sítích. Vybudovaná důvěra s uživateli a kvalitní vztahy v mnoha případech zaručí loajálnost zákazníků a jejich dobrovolné sdílení a šíření pozitivních informací o podniku. Aby uživatel dobrovolně sdílel a šířil takové informace (i mimo sociální sítě), je jeden

z hlavních cílů podniků na sociálních sítích a má největší přidanou hodnotu. Mezi podnikem a uživateli neustále probíhá výměna informací, ať už se jedná o přímou výměnu (komentář uživatele na příspěvek podniku) nebo nepřímou výměnu (podniky vidí, jak se chovají uživatelé na sociálních sítích a podle toho upravují strategii). Tuto výměnu neustále narušuje šum, jako je například fakenews, cílené poškozování, drby, chybné úsudky a podobně.

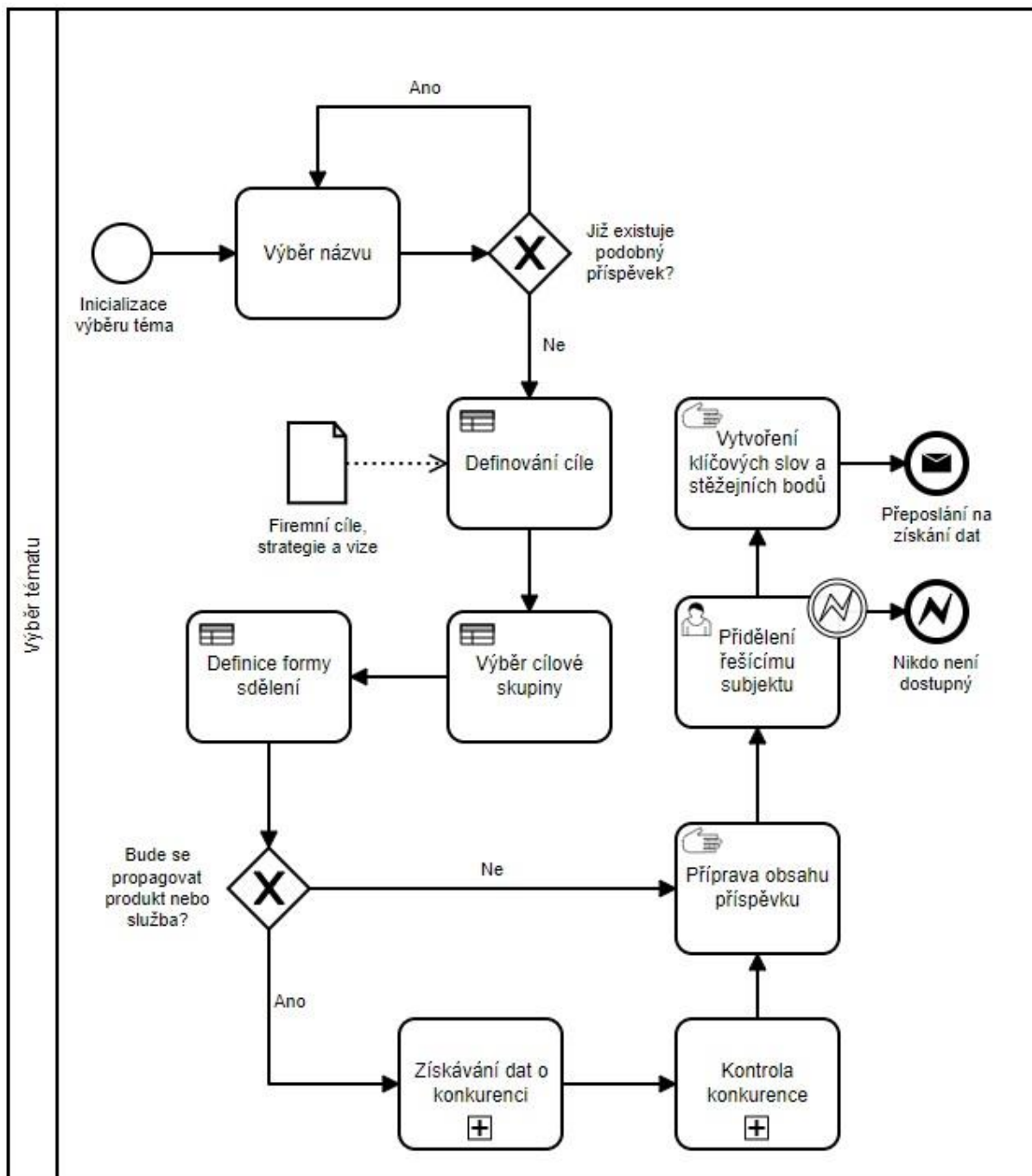
Obrázek 24 zobrazuje, jak probíhá ve firmě tvorba nového příspěvku a výše popsání analýza s postupným sdílením a reagováním na sociálních sítích.



Obrázek 24: Inicializace nového příspěvku (vlastní dílo autora)

Na procesním modelu „inicializace nového příspěvku“ je zobrazen postup tvorby příspěvku od výběru tématu až po reakci uživatelů na sociálních sítích. Pro definici doporučení následuje bližší zobrazení tří procesů. Tyto tři procesy jsou „výběr tématu“, „analýzy a výpočty“ a „tvorba vhodné struktury pro vybrané téma“. U procesu s názvem výběr tématu nejdříve dochází k volbě názvu celého příspěvku. Výběr tématu by se měl řídit podle toho, na jakou cílovou skupinu míříme a co očekáváme, že získáme vytvořením nového příspěvku.

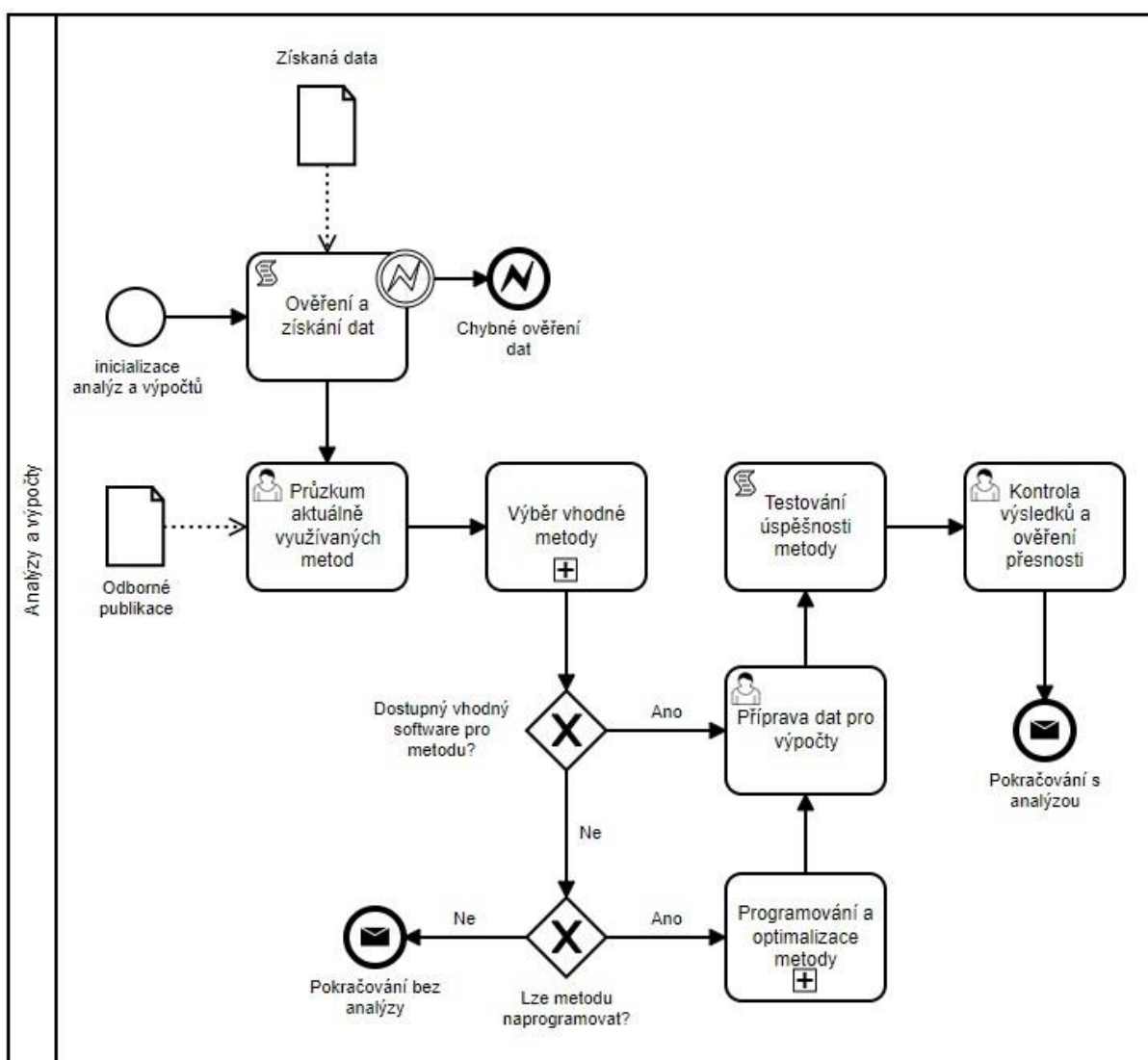
V procesu je uvedena kontrola, jestli už organizace nebo konkurence stejný či podobný článek nevytvořila, a pokud ano, vybírá se jiný název. Následně je definován cíl příspěvku, tedy čeho se tímto příspěvkem snažíme dosáhnout. Může se jednat o zvýšení image firmy, prodej produktu, poskytnutí služby, budování důvěry a podobně. Tento cíl by měl ideálně korespondovat s firemní strategií, vizí a posláním. S tím je spojen výběr cílové skupiny a definice, pro koho je článek určen. Následuje definice formy sdělení, tím je myšleno, zda bude článek seriózní, nebo má spíše pobavit, případně kombinace obojího. Pokud se jedná o propagaci produktu nebo služby, je vhodné analyzovat konkurenci, zjistit, jak propaguje svůj produkt nebo službu, a také se dozvědět, jak reagují uživatelé, aby se případně předešlo některým chybám. Stejně tak může organizace u konkurence zjistit, že nějaký styl propagace funguje kvalitně, a použít podobný ve vylepšené podobě. Následuje příprava obsahu příspěvku, předání zodpovědnému řešiteli a vytvoření klíčových slov, podle kterých lze provést získání dat z již existujících příspěvků pro analýzu sentimentu, obsahovou analýzu i další testování.



Obrázek 25: Proces „výběr tématu“ (vlastní dílo autora)

Po získání potřebných dat je potřeba provést analýzu a výpočty pomocí vybrané metody machine learning nebo jiné vhodné klasifikační metody. V případě, že nedojde k chybě u ověření dat v první fázi analýzy, je následně vybrána vhodná metoda pro analýzu. Pokud organizace tyto testy ještě neprováděla, je vhodné udělat průzkum aktuálně využívaných metod a ověřit, zda neexistuje nějaká nová metoda, jež je lepší než testované. Toto zjištění by mohlo poskytnout klíčovou konkurenční výhodu. Z tohoto průzkumu je vybrána nejvhodnější metoda, co lze doporučit k aktuálnímu datu (vhodné metody v době tvorby této práce jsou uvedeny

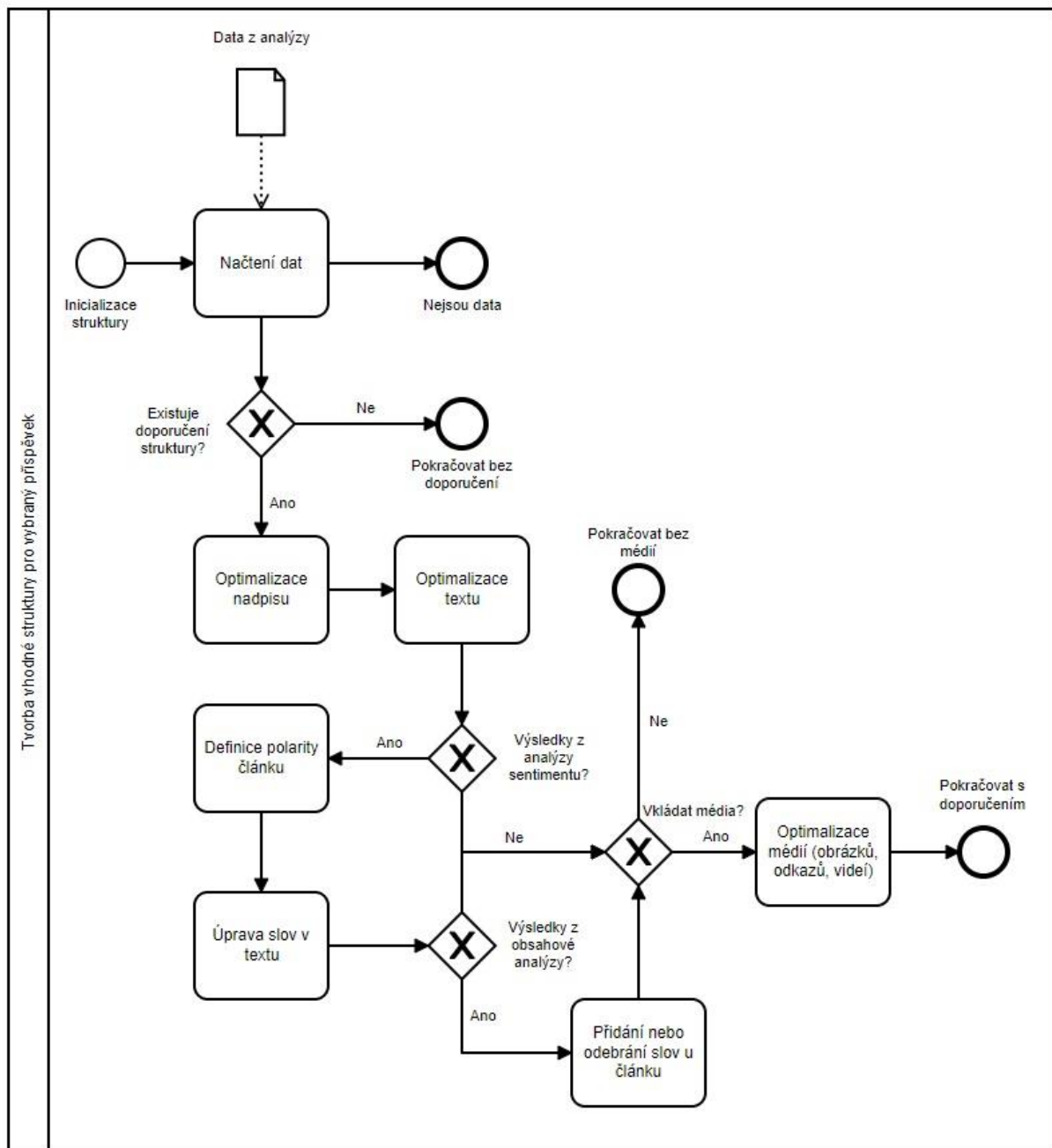
v kapitole 7.1 – východiska modelu). Pokud má organizace k dispozici software pro analýzu této metody, provede analýzu v softwaru. Pokud software není dostupný, lze provést analýzu ve vlastním naprogramovaném modifikovatelném algoritmu. Pokud organizace nemá ani tuto možnost, pokračuje se na další proces s aktuálně zjištěným doporučením, případně na zvolení jiné metody. Po provedení přípravy dat následuje testování úspěšnosti vybrané metody pomocí křížové validace. Následná kontrola výsledků a ověření přesnosti poskytne doporučení, čemu se u příspěvku věnovat a jaké proměnné nemají pro model a případné sdílení na sociálních sítích žádnou váhu.



Obrázek 26: Proces „analýza a výpočty“ (vlastní dílo autora)

Při přiblížení o úroveň konkrétněji do procesu „Tvorba vhodné struktury pro vybrané téma“ lze vidět další množinu procesů, které se týkají doporučení vhodné struktury příspěvku pro zvýšení pravděpodobnosti sdílení uživateli na sociálních sítích. V první fázi probíhá načtení dat z analýzy doporučení pro optimalizaci různých proměnných (délka článku, počet slov

v nadpisu, zastoupení pozitivních a negativních slov, množství obrázků a podobně). Pokud není dostupné žádné doporučení (tedy prozatím firma neanalyzovala dosavadní příspěvky na sociálních sítích), celý tento proces není v daný moment relevantní, firma se k němu vrátí později. Pokud již výstupy analýz existují, dochází k optimalizaci (ve vazbě na výsledky) nadpisu a celého textu příspěvku, nastavení optimálního počtu slov v textu a nadpisu, počet odstavců a podobně. Pokud jsou dostupné výsledky z analýzy sentimentu, upravuje se článek podle zjištěné ideální polaridy článku a vkládají se, nebo se odebírají pozitivní a negativní slova. Pokud existují výsledky z obsahové analýzy, upravuje se text znovu – doplňují se slova a slovní spojení, která mají potenciál zvýšit počet sdílení a oblíbenost na sociálních sítích. Nakonec se optimalizují média v příspěvku. Vloží se optimální počet obrázků, odkazů a případně videí.



Obrázek 27: Tvorba vhodné struktury pro vybrané téma (vlastní dílo autora)

Na výše uvedených částech procesu lze vidět možnou hloubku přípravy některých z nich. Externí komunikace zasahuje i do dalších částí, jak lze vidět na obrázku č. 25, tedy v částech popsaných procesů pro „širokou veřejnost“ a „sociální sítě“.

Bylo by možné dále zkoumat hloubku procesu a správného nastavení v rámci těchto dílčích oblastí. To představuje možnost dalšího zkoumání, nicméně to není součástí hlavního cíle disertační práce, představuje to však potenciál pro další zkoumání.

Celkově lze říci, že výzkumem opírajícím se o neuronové sítě a další klasifikační metody bylo zjištěno, že samotná struktura článku umístěného na internetu neovlivňuje

primárně počet sdílení. Ani statistické výpočty tyto předpoklady nepotvrdily. Na druhou stranu zjištění z dotazníkového šetření poukazují na určité skutečnosti. Čím vyšší vzdělání uživatel má, tím více sociálních sítí používá. Stejně pravidlo platí i pro zvyšující se příjem uživatelů sociálních sítí. I když nebyl potvrzen přímý vztah mezi počtem sdílení, velmi úspěšné články (z pohledu počtu sdílení) využívaly v textu zhruba 5–6 obrázků na článek. Ty méně úspěšné články měly většinou menší počet obrázků (například 4 a méně), byly tak pravděpodobně pocitově obtížnější ke čtení. Uživatelé sociálních sítí nejčastěji sdílí příspěvky, které je pobaví, a chtějí pobavit i svoje okolí. Druhou nejčastější oblastí sdílení jsou videa jako například hudební klipy. V dotazníku bylo také uvedeno časté sdílení různých kulturních zážitků, kde se uživatel chce pochlubit svojí zkušeností z cestování. Mezi další zjištění patří, že uživatelé rádi sdílí příspěvky slavných osobností, ale ve většině případů se jedná o naprosto náhodné autory příspěvků, a jde tedy hlavně o obsah a sdělení. Některé metody použité v rámci disertační práce poukázaly na cesty, jež nevedou k jednoznačnému cíli, a není tedy možné se pouze o ně samotné opřít jako o klíčové podklady k rozhodování.

Pokud se organizace snaží prosadit na sociálních sítích, je jasné, že se nejedná o triviální úkol a je zapotřebí být kreativní a experimentovat.

8 Doporučení, přínos a limity práce

Kvalitně zavedená a rozvíjející se firemní externí komunikace může pozitivně ovlivnit celou řadu aspektů, jako je expanzivní politika firmy pro velikost na trhu, kladně zvednout motivaci a dovednosti zaměstnanců, zlepšit kvalitu produktů a služeb, ušetřit čas a zvýšit tržby.

V prostředí firmy, sociálních sítí a informačního managementu je mnoho navzájem působících entit a stakeholderů. Mezi tyto entity patří široká veřejnost (uživatelé, partneři, konkurence), využívaná technologie, veřejné informace (fotografie, videa, články, texty a komentáře) nebo zpětná vazba. Oddělení, které má na starost rozhodování o externí komunikaci, je seznámeno s jejich vzájemným působením a hledá optimální způsoby, jak zajišťovat pozitivní dopad této komunikace.

Jako klíčové procesy v externí komunikaci na sociálních médiích lze označit výběr tématu pro příspěvek, získání dat pro analýzu a následné provedení analýz a výpočtů, z nichž lze vytvořit vhodnou strukturu příspěvku pro vybrané téma. V rámci jeho výběru je nutné jednoznačně definovat cíle, jasné určení cílové skupiny a specifikace formy sdělení. V návaznosti na to je pak řešena struktura příspěvků. Průzkum jejich struktury ve spojitosti s počtem sdílení na sociálních sítích metodou MLP neposkytl jednoznačné výsledky a je možné, že toto je slepá ulička, která nikam nevede. Přesto lze na základě výsledků této práce doporučit tvorbu spíše kratších příspěvků, kratších prvních odstavců, zábavných příspěvků nebo prokládání textu více obrázky. Firmy mohou tento postup z uvedeného konceptuálního modelu i z procesních modelů využít pro své vlastní analýzy například na datech, která nejsou veřejnosti dostupná (například doba setrvání uživatele na firemním webu nebo reklamy umístěné na sociálních sítích), a upravovat svoji externí komunikaci podle oboru, v němž podnikají, a podle cílové skupiny. Každý obor a každá cílová skupina vyžaduje individuální přístup.

Při hlubším zkoumání výsledků disertační práce lze popsat doporučení ve vztahu k možným metodám využití pro hlubší zkoumání tématu příspěvku a jeho struktury. Z provedené systematické rešerše vyplynulo několik vhodných metod pro možné usnadnění rozhodování firem na sociálních sítích, mezi něž patří metody strojového učení SVM, MLP, RNN, CNN, LSTM a Naive Bayes. Lze konstatovat, že zmíněné metody neuronových sítí ve vazbě na zkoumané proměnné v kapitole 4.3 nelze generalizovaně doporučit jako řešení k zjištění klíčových proměnných ovlivňujících sdílení příspěvků, protože vykázaly menší význam zkoumaných proměnných při využití metody MLP. Z návazné analýzy sentimentu bylo zjištěno, že polarita článku má minoritní vliv na jeho sdílení.

Vhodným vstupem pro rozhodování o nastavení komunikace a sdílení příspěvků jsou informace od samotných potenciálních uživatelů. Na základě vyhodnocení dotazování totiž lze doporučit příspěvky, které určitým způsobem uživatele pobaví. Mají totiž větší šanci na sdílení, protože uživatelé nejčastěji sdílí právě zábavné příspěvky.

Navržený konceptuální model celkově poskytuje popis metod a návaznosti procesů pro efektivní externí komunikaci, pro tvorbu příspěvku na sociální síť, který bude cílit na konkrétní skupinu uživatelů, kterou následně dokáže vhodně zaujmout. Popisuje i vhodný způsob zajišťování zpětné vazby tak, aby bylo možné podle následných zjištění dále upravit výsledky a ještě lépe připravit nová sdělení.

Limity práce

Autor si je v práci vědom určitých limitů výzkumu, za něž lze považovat to, že byly vybrány pouze dva určité typy neuronových sítí ke zkoumání vybraných proměnných struktury příspěvku. Je možné, že by další metody poukázaly na význam některých proměnných, které nyní nebyly označeny jako signifikantní pro návrh struktury příspěvku. Stejně tak by bylo možné, že při využití dalších sad dat v jiných tematických oblastech by bylo dosaženo nových výsledků.

Druhým limitem příspěvku je tvorba a rozsah dotazování, kdy sada otázek byla za účelem dosažení uspokojivé míry návratnosti limitována svým počtem. Autor vycházel z literatury a do výběru zahrnul zejména ty otázky, které v minulých průzkumech a jejich výsledcích poukázaly na určité významné faktory. Nicméně je možné, že nemusely být postiženy všechny skutečnosti.

Přínos práce v praktické a teoretické rovině

Organizace a firmy soupeřící o pozornost na sociálních sítích se musí vypořádávat s mnoha aspekty, které mohou ovlivnit jejich úspěšnost. Disertační práce může v praktické rovině pomoci v tom, že přináší jednak souhrn doporučení, jak publikovat příspěvek, který má být na sociálních médiích úspěšný, ale také poskytuje představu o velké šíři metod, které lze v tomto směru aplikovat. Ačkoli některé z nich i při rozsáhlé těžbě dat nevykazovaly statisticky významné výsledky, jedná se o zjištění, které pomůže firmě i správně zacílit a vybrat vhodnou metodu pro své užití.

Přínos v teoretické rovině spočívá v rozšíření metod informačního managementu o aspekty externí komunikace organizací, efektivního řízení a podpory rozhodování.

9 Závěr

Předložená disertační práce si klade za cíl navrhnout konceptuální model využití informací a znalostí k podpoře firemního rozhodování o obsahu a formě externí komunikace na sociálních sítích.

V tomto kontextu byla nejprve charakterizována klíčová teoretická východiska týkající se komunikace, informací a znalostí, byla provedena analýza současného stavu poznání v oblasti firemní externí komunikace na sociálních sítích a metod používaných pro zvýšení efektivity komunikace na sociálních sítích. Následně byl analyzován aktuální stav využívání metod pro analýzu obsahu příspěvku a dotazování pro specifikaci klíčových informací a znalostí k podpoře firemního rozhodování o obsahu a formě externí komunikace. V neposlední řadě byla provedena analýza dat a výpočty pomocí neuronových sítí a klasifikačních modelů, obsahová analýza a analýza sentimentu. V práci je také uvedeno dotazníkové šetření, ve kterém byla zjištěna motivace uživatelů pro sdílení příspěvků a také jaké je téma nejčastěji sdílených příspěvků.

Na základě realizace výše uvedených metod a z nich vyplývajících zjištění je v práci vytvořen model klíčových informací a znalostí pro externí komunikaci firmy na sociálních sítích spolu se specifikací procesů, které doporučují a poukazují na to, jak lze zefektivnit firemní externí komunikaci.

Analyzováním dat od konkurence, vlastních úspěchů i neúspěchů z minulosti, příspěvků na sociálních sítích mohou podniky získat zásadní informační výhodu na trhu. Modely a přístupy zpracované v této práci by mohly být pro podnik funkčním východiskem ke zlepšení efektivity z hlediska většího sdílení obsahu jejich příspěvků a zajištění více interakcí na sociálních sítích. Toto zlepšení u podniků pomůže rozvíjet jejich strategii a bude se jednodušeji dodržovat firemní vize.

Disertační práce se zabývá velmi komplexním problémem. Další rozvoj výzkumného tématu a možné pokračování této práce je zcela jistě možné. Může spočívat například jen v opakovaném použití postupu na širším datovém vstupu a následné porovnání výsledků. Efektivní metodou by mohlo být použití metod machine learning na celý text příspěvků pro zjištění, jestli kombinace nějakých slov nebo slovních spojení je závislá na počtu sdílení. Dala by se také provést analýza uvedená v této práci pro různé tematické bloky v rámci typu konverzací, či typu subjektů, které dané téma sdílejí a diskutují. I pro predikci počtu sdílení by se také dalo otestovat mnoho dalších metod jako Random Forrest, Naive Bayes nebo K-Nearest Neighbors (KNN) a zkoumat jejich úspěšnost.

10 Seznam použité literatury

- AbdelShafy, A., Romme, A. G. L., & Walrave, B. (2015). Why some small and medium enterprises grow and others do not? *Proceedings of the 33rd International Conference of the System Dynamics Society, 19–23 July 2015, Cambridge, Massachusetts.* conference; International Conference of the System Dynamics Society; 2015-07-19; 2015-07-23. <https://research.tue.nl/en/publications/why-some-small-and-medium-enterprises-grow-and-others-do-not>
- Abid, F., Alam, M., Yasir, M., & Li, C. (2019a). Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter. *Future Generation Computer Systems, 95*, 292–308. <https://doi.org/10.1016/j.future.2018.12.018>
- Agichtein, E., Castillo, C., Donato, D., Gionis, A., & Mishne, G. (2008). Finding High-quality Content in Social Media. *Proceedings of the 2008 International Conference on Web Search and Data Mining*, 183–194. <https://doi.org/10.1145/1341531.1341557>
- Alagöz, S. B., & Ekici, N. (2014). Experiential Marketing and Vacation Experience: The Sample of Turkish Airlines*. *Procedia – Social and Behavioral Sciences, 150*, 500–510. <https://doi.org/10.1016/j.sbspro.2014.09.065>
- Alharbi, A. S. M., & de Doncker, E. (2019). Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information. *Cognitive Systems Research, 54*, 50–61. <https://doi.org/10.1016/j.cogsys.2018.10.001>
- Aral, S., Dellarocas, C., & Godes, D. (2013). Introduction to the Special Issue– Social Media and Business Transformation: A Framework for Research. *Information Systems Research, 24*, 3–13. <https://doi.org/10.1287/isre.1120.0470>
- Ashley, C., & Tuten, T. (2015). Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement. *Psychology & Marketing, 32*. <https://doi.org/10.1002/mar.20761>
- Bachman, P. (n.d.). *Social Media in Public Marketing: Facebook Pages of Regional Authorities.* 11–19. https://uni.uhk.cz/hed/site/assets/files/1072/proceedings_2018_1.pdf
- Baird, C. H., & Parasnis, G. (2011). From social media to Social CRM: Reinventing the customer relationship. *Strategy & Leadership.* <https://doi.org/10.1108/10878571111176600>

- Ballings, M., Van den Poel, D., & Bogaert, M. (2016a). Social media optimization: Identifying an optimal strategy for increasing network size on Facebook. *Omega*, *59*, 15–25. <https://doi.org/10.1016/j.omega.2015.04.017>
- Bandari, R., Asur, S., & Huberman, B. A. (2012). The Pulse of News in Social Media: Forecasting Popularity. *ArXiv:1202.0332 [Physics]*. <http://arxiv.org/abs/1202.0332>
- Bednář, V. (2011). *Marketing na sociálních sítích: Prosaďte se na Facebooku a Twitteru*.
- Berthon, P., Pitt, L., & Campbell, C. (2008). Ad Lib: When customers create the ad. *California Management Review*, *50*(4), 6–30.
- Best, P., Manktelow, R., & Taylor, B. J. (2016). Social Work and Social Media: Online Help-Seeking and the Mental Well-Being of Adolescent Males. *The British Journal of Social Work*, *46*(1), 257–276. <https://doi.org/10.1093/bjsw/bcu130>
- Braun, P., Cuzzocrea, A., Doan, L. M. V., Kim, S., Leung, C. K., Matundan, J. F. A., & Robby Singh, R. (2017). Enhanced Prediction of User-Preferred YouTube Videos Based on Cleaned Viewing Pattern History. *Procedia Computer Science*, *112*, 2230–2239. <https://doi.org/10.1016/j.procs.2017.08.129>
- Camarero, C., & San José, R. (2011). Social and Attitudinal Determinants of Viral Marketing Dynamics. *Comput. Hum. Behav.*, *27*(6), 2292–2300. <https://doi.org/10.1016/j.chb.2011.07.008>
- Cerchiello, P., Giudici, P., & Nicola, G. (2017). Twitter data models for bank risk contagion. *Neurocomputing*, *264*, 50–56. <https://doi.org/10.1016/j.neucom.2016.10.101>
- Chen, J., Liu, Y., & Zou, M. (2017). User emotion for modeling retweeting behaviors. *Neural Networks : The Official Journal of the International Neural Network Society*, *96*, 11–21. <https://doi.org/10.1016/j.neunet.2017.08.006>
- Cheng, Y., Chen, K., Sun, H., Zhang, Y., & Tao, F. (2018). Data and knowledge mining with big data towards smart production. *Journal of Industrial Information Integration*, *9*, 1–13. <https://doi.org/10.1016/j.jii.2017.08.001>
- Chu, S.-C., & Kim, Y. (2011). Determinants of consumer engagement in electronic word-of-mouth (eWOM) in social networking sites. *International Journal of Advertising*, *30*(1), 47–75. <https://doi.org/10.2501/IJA-30-1-047-075>
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *ArXiv:1412.3555 [Cs]*. <http://arxiv.org/abs/1412.3555>
- Cooper, P. (2017). Data, information, knowledge and wisdom. *Anaesthesia & Intensive Care Medicine*, *18*(1), 55–56. <https://doi.org/10.1016/j.mpaic.2016.10.006>

- Corral de Zubielqui, G., Fryges, H., & Jones, J. (2019). Social media, open innovation & HRM: Implications for performance. *Technological Forecasting and Social Change*, *144*, 334–347. <https://doi.org/10.1016/j.techfore.2017.07.014>
- Czech Invest. (n.d.). *Definice malého a středního podnikatele*. Retrieved 12 July 2019, from <https://www.czechinvest.org/cz/Sluzby-pro-male-a-stredni-podnikatele/Chcete-dotace/OPPI/Radce/Definice-maleho-a-stredniho-podnikatele>
- De Marco, A. (2015, January 27). 2015: The Year of Smarter Internal Communications. *Tech.Co*. <https://tech.co/news/2015-will-year-smarter-internal-communications-2015-01>
- Denić, N., Petković, D., Vujović, V., Spasić, B., & Vujičić, I. (2018). A survey of internet marketing by small and medium-sized enterprises for placing wine on the market. *Physica A: Statistical Mechanics and Its Applications*, *506(C)*, 718–727.
- Din, S. Z. M., Anuar, R. H. M., Omar, N., Omar, H., & Dahlan, J. M. (2015). Discovering the Use of Online Recruitment via Social Media of Student Internship. *Procedia Economics and Finance*, *31*, 856–860. [https://doi.org/10.1016/S2212-5671\(15\)01181-8](https://doi.org/10.1016/S2212-5671(15)01181-8)
- Ding, K., Wang, R., & Wang, S. (2019). Social Media Popularity Prediction: A Multiple Feature Fusion Approach with Deep Neural Networks. *Proceedings of the 27th ACM International Conference on Multimedia*, 2682–2686. <https://doi.org/10.1145/3343031.3356062>
- Donahue, J., Hendricks, L. A., Rohrbach, M., Venugopalan, S., Guadarrama, S., Saenko, K., & Darrell, T. (2014). Long-term Recurrent Convolutional Networks for Visual Recognition and Description. *ArXiv:1411.4389 [Cs]*. <http://arxiv.org/abs/1411.4389>
- Drury, G. (2008). Opinion piece: Social media: Should marketers engage and how can it be done effectively? *Journal of Direct, Data and Digital Marketing Practice*, *9(3)*, 274–277. <https://doi.org/10.1057/palgrave.ddmp.4350096>
- Ducange, P., Fazzolari, M., Petrocchi, M., & Vecchio, M. (2019a). An effective Decision Support System for social media listening based on cross-source sentiment analysis models. *Engineering Applications of Artificial Intelligence*, *78*, 71–85. <https://doi.org/10.1016/j.engappai.2018.10.014>
- El Ouiridi, M., El Ouiridi, A., Segers, J., & Pais, I. (2016). Technology adoption in employee recruitment: The case of social media in Central and Eastern Europe. *Computers in Human Behavior*, *57*, 240–249. <https://doi.org/10.1016/j.chb.2015.12.043>
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62(1)*, 107–115. <https://doi.org/10.1111/j.1365-2648.2007.04569.x>

- Emily Jones, M. (n.d.). *LibGuides: Creating a PRISMA flow diagram: Creating a PRISMA flow diagram*. Retrieved 5 October 2020, from <https://guides.lib.unc.edu/c.php?g=216835&p=1432026>
- Eroshkin, S. Yu., Kameneva, N. A., Kovkov, D. V., & Sukhorukov, A. I. (2017). Conceptual System in the Modern Information Management. *Procedia Computer Science*, *103*, 609–612. <https://doi.org/10.1016/j.procs.2017.01.079>
- Ghaisani, A. P., Handayani, P. W., & Munajat, Q. (2017). Users' Motivation in Sharing Information on Social Media. *Procedia Computer Science*, *124*, 530–535. <https://doi.org/10.1016/j.procs.2017.12.186>
- Ghiassi, M., Skinner, J., & Zimbra, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with Applications*, *40*(16), 6266–6282. <https://doi.org/10.1016/j.eswa.2013.05.057>
- Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science (New York, N.Y.)*, *333*(6051), 1878–1881. <https://doi.org/10.1126/science.1202775>
- Gómez, M., Lopez, C., & Molina, A. (2019). An integrated model of social media brand engagement. *Computers in Human Behavior*, *96*, 196–206. <https://doi.org/10.1016/j.chb.2019.01.026>
- Guandan, C., Kong, Q., Xu, N., & Mao, W. (2018). NPP: A Neural Popularity Prediction Model for Social Media Content. *Neurocomputing*. <https://doi.org/10.1016/j.neucom.2018.12.039>
- Gürsun, G., Crovella, M., & Matta, I. (2011). Describing and forecasting video access patterns. *2011 Proceedings IEEE INFOCOM*, 16–20. <https://doi.org/10.1109/INFOCOM.2011.5934965>
- Hajek, P., Barushka, A., & Munk, M. (2020). Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining. *Neural Computing and Applications*. <https://doi.org/10.1007/s00521-020-04757-2>
- Hampshire, J. B., & Pearlmutter, B. (1991b). Equivalence Proofs for Multi-Layer Perceptron Classifiers and the Bayesian Discriminant Function. In D. S. Touretzky, J. L. Elman, T. J. Sejnowski, & G. E. Hinton (Eds.), *Connectionist Models* (pp. 159–172). Morgan Kaufmann. <https://doi.org/10.1016/B978-1-4832-1448-1.50023-8>
- Hanna, R., Rohm, A., & Crittenden, V. L. (2011). We're all connected: The power of the social media ecosystem. *Business Horizons*, *54*(3), 265–273. <https://doi.org/10.1016/j.bushor.2011.01.007>

- Harvey, S. (2018, May 16). *External Communication Strategies: Finding Your Marketing Megaphone*. <https://fabrikbrands.com/external-communication-strategies/>
- Hautz, J., Füller, J., Hutter, K., & Thürridl, C. (2014). *Let Users Generate Your Video Ads? The Impact of Video Source and Quality on Consumers' Perceptions and Intended Behaviors*. <https://doi.org/10.1016/j.intmar.2013.06.003>
- Hay, B. (2010). *Twitter Twitter - But who is listening? A review of the current and potential use of Twittering as a tourism marketing tool*. <https://eresearch.qmu.ac.uk/handle/20.500.12289/1500>
- Hennig-Thurau, T., Hofacker, C. F., & Bloching, B. (2013). Marketing the Pinball Way: Understanding How Social Media Change the Generation of Value for Consumers and Companies. *Journal of Interactive Marketing*, 27(4), 237–241. <https://doi.org/10.1016/j.intmar.2013.09.005>
- Hofacker, C. F., & Belanche, D. (2016). Eight social media challenges for marketing managers. *Spanish Journal of Marketing – ESIC*, 20(2), 73–80. <https://doi.org/10.1016/j.sjme.2016.07.003>
- Hong, Y. J., Shin, D., & Kim, J. H. (2016). High/low reputation companies' dialogic communication activities and semantic networks on Facebook: A comparative study. *Technological Forecasting and Social Change*, 110, 78–92. <https://doi.org/10.1016/j.techfore.2016.05.003>
- Hou, R., Wu, J., & Du, H. S. (2017). Customer social network affects marketing strategy: A simulation analysis based on competitive diffusion model. *Physica A: Statistical Mechanics and Its Applications*, 469(C), 644–653.
- Househ, M. (2013). The use of social media in healthcare: Organizational, clinical, and patient perspectives. *Studies in Health Technology and Informatics*, 183, 244–248.
- Howells, K., & Ertugan, A. (2017). Applying fuzzy logic for sentiment analysis of social media network data in marketing. *Procedia Computer Science*, 120, 664–670. <https://doi.org/10.1016/j.procs.2017.11.293>
- Hu, Y., Hu, C., Fu, S., Shi, P., & Ning, B. (2016). Predicting the popularity of viral topics based on time series forecasting. *Neurocomputing*, 210, 55–65. <https://doi.org/10.1016/j.neucom.2015.10.143>
- Hwang, Y., Lin, H., & Shin, D. (2018). Knowledge system commitment and knowledge sharing intention: The role of personal information management motivation. *International Journal of Information Management*, 39, 220–227. <https://doi.org/10.1016/j.ijinfomgt.2017.12.009>

- Infographic: Online Marketing Metrics That Matter. (n.d.). *Marketing Mojo*. Retrieved 12 July 2019, from <https://www.marketing-mojo.com/infographic/marketing-metrics-that-matter/>
- Jeong, B., Yoon, J., & Lee, J.-M. (2017). Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2017.09.009>
- Jimenez-Marquez, J. L., Gonzalez-Carrasco, I., Lopez-Cuadrado, J. L., & Ruiz-Mezcua, B. (2019). Towards a big data framework for analyzing social media content. *International Journal of Information Management*, *44*, 1–12. <https://doi.org/10.1016/j.ijinfomgt.2018.09.003>
- Jue, A. L., Marr, J. A., & Kassotakis, M. E. (2010). *Social media at work: How networking tools propel organizational performance*. Jossey-Bass.
- Kang, J., & Lee, H. (2017). Modeling User Interest in Social Media Using News Media and Wikipedia. *Inf. Syst.*, *65*(C), 52–64. <https://doi.org/10.1016/j.is.2016.11.003>
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59–68. <https://doi.org/10.1016/j.bushor.2009.09.003>
- Khan, G. F., & Vong, S. (2014). Virality over YouTube: An empirical analysis. *Internet Research*. <https://doi.org/10.1108/IntR-05-2013-0085>
- Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in Human Behavior*, *66*, 236–247. <https://doi.org/10.1016/j.chb.2016.09.024>
- Kietzmann, J. H., Hermkens, K., McCarthy, I., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, *54*(3), 241–251.
- Kim, T., & Adali, T. (2002b). Fully Complex Multi-Layer Perceptron Network for Nonlinear Signal Processing. *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, *32*(1), 29–43. <https://doi.org/10.1023/A:1016359216961>
- Kiráľová, A., & Pavlíček, A. (2015). Development of Social Media Strategies in Tourism Destination. *Procedia - Social and Behavioral Sciences*, *175*, 358–366. <https://doi.org/10.1016/j.sbspro.2015.01.1211>

- Kratzwald, B., Ilic, S., Kraus, M., Feuerriegel, S., & Prendinger, H. (2018a). Decision support with text-based emotion recognition: Deep learning for affective computing. *Decision Support Systems, 115*. <https://doi.org/10.1016/j.dss.2018.09.002>
- Kratzwald, B., Ilic, S., Kraus, M., Feuerriegel, S., & Prendinger, H. (2018b). Deep learning for affective computing: Text-based emotion recognition in decision support. *Decision Support Systems, 115*, 24–35. <https://doi.org/10.1016/j.dss.2018.09.002>
- Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R., & Kannan, P. K. (2016). From Social to Sale: The Effects of Firm-Generated Content in Social Media on Customer Behavior. *Journal of Marketing, 80*(1), 7–25. <https://doi.org/10.1509/jm.14.0249>
- Lee, I. (2018). Social media analytics for enterprises: Typology, methods, and processes. *Business Horizons, 61*(2), 199–210. <https://doi.org/10.1016/j.bushor.2017.11.002>
- Lei, L. (Gillian), Li, Y., & Luo, Y. (2019). Production and dissemination of corporate information in social media: A review. *Journal of Accounting Literature, 42*, 29–43. <https://doi.org/10.1016/j.acclit.2019.02.002>
- Li, C., Ma, J., Guo, X., & Mei, Q. (2016). DeepCas: An End-to-end Predictor of Information Cascades. *ArXiv:1611.05373 [Cs]*. <http://arxiv.org/abs/1611.05373>
- Liao, S., Wang, J., Yu, R., Sato, K., & Cheng, Z. (2017). CNN for situations understanding based on sentiment analysis of twitter data. *Procedia Computer Science, 111*, 376–381. <https://doi.org/10.1016/j.procs.2017.06.037>
- Lima, A. C. E. S., & de Castro, L. N. (2014). A multi-label, semi-supervised classification approach applied to personality prediction in social media. *Neural Networks, 58*, 122–130. <https://doi.org/10.1016/j.neunet.2014.05.020>
- Liu, Yanbing, Zhao, J., & Xiao, Y. (2018). C-RBFNN: A user retweet behavior prediction method for hotspot topics based on improved RBF neural network. *Neurocomputing, 275*, 733–746. <https://doi.org/10.1016/j.neucom.2017.09.015>
- Liu, Yao, Jiang, C., & Zhao, H. (2018). Using contextual features and multi-view ensemble learning in product defect identification from online discussion forums. *Decision Support Systems, 105*, 1–12. <https://doi.org/10.1016/j.dss.2017.10.009>
- Lüders, M. (2008). Conceptualizing personal media. *New Media & Society, 10*(5), 683–702. <https://doi.org/10.1177/1461444808094352>
- Luo, L., Wang, Y., & Han, L. (2013). Marketing via social media: A case study. *Library Hi Tech*. <https://doi.org/10.1108/LHT-12-2012-0141>

- ManagementMania. (n.d.). *Základní model sociální komunikace (Basic Model of Social Communication)*. ManagementMania.com. Retrieved 12 July 2019, from <https://managementmania.com/cs/zakladni-model-socialni-komunikace>
- Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons*, 52(4), 357–365.
- Mayer, A. (2009). Online social networks in economics. *Decision Support Systems*, 47(3), 169–184. <https://doi.org/10.1016/j.dss.2009.02.009>
- Miroslava Szarková, 1951-. (2008). *Komunikačné nástroje v systéme riadenia podniku*.
- Mishra, S., Rizoïu, M.-A., & Xie, L. (2016). Feature Driven and Point Process Approaches for Popularity Prediction. *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management – CIKM '16*, 1069–1078. <https://doi.org/10.1145/2983323.2983812>
- Moro, S., Rita, P., & Vala, B. (2016). Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach. *Journal of Business Research*, 69(9), 3341–3351.
- Munar, A. M., & Jacobsen, J. K. S. (2014). Motivations for Sharing Tourism Experiences through Social Media. *Tourism Management*, 43, 46–54. <https://doi.org/10.1016/j.tourman.2014.01.012>
- Muninger, M.-I., Hammedi, W., & Mahr, D. (2019). The value of social media for innovation: A capability perspective. *Journal of Business Research*, 95, 116–127. <https://doi.org/10.1016/j.jbusres.2018.10.012>
- Murdough, C. (2009). Social Media Measurement. *Journal of Interactive Advertising*, 10(1), 94–99. <https://doi.org/10.1080/15252019.2009.10722165>
- Musa, H. (2016). *Analyzing the Effectiveness of Social Media Marketing Original Research Article*. 4–10. <https://doi.org/10.15405/epsbs.2016.08.2>
- Mynarzová, M., & Kaňa, R. (2014). Theory and Practice of Industrial Policy of the EU in the Context of Globalization Challenges. *International Conference on European Integration*, 499–507.
- Nakonečný, M. (n.d.). *Psychologie osobnosti*. Retrieved 12 July 2019, from <https://www.kosmas.cz/knihy/146252/psychologie-osobnosti/>
- Nezakati, H., Amidi, A., Jusoh, Y. Y., Moghadas, S., Aziz, Y. A., & Sohrabinezhadtalemi, R. (2015). Review of Social Media Potential on Knowledge Sharing and Collaboration in Tourism Industry. *Procedia - Social and Behavioral Sciences*, 172, 120–125. <https://doi.org/10.1016/j.sbspro.2015.01.344>

- Ngai, E. W. T., Moon, K. K., Lam, S. S., Chin, E. S. K., & Tao, S. S. C. (2015). Social media models, technologies, and applications. *Industrial Management & Data Systems*. <https://doi.org/10.1108/IMDS-03-2015-0075>
- Ngai, E. W. T., Tao, S. S. C., & Moon, K. K. L. (2015). Social media research: Theories, constructs, and conceptual frameworks. *International Journal of Information Management*, 35(1), 33–44. <https://doi.org/10.1016/j.ijinfomgt.2014.09.004>
- Pan, B., & Crotts, J. (2012). Theoretical Models of Social Media, Marketing Implications, and Future Research Directions. *Social Media in Travel, Tourism and Hospitality: Theory, Practice and Cases*, 73–85.
- Patel, N. (2015, December 6). *Your Ads are Getting Ignored: 5 Smart Strategies to Overcome Banner Blindness*. Neil Patel. <https://neilpatel.com/blog/your-ads-are-getting-ignored-5-smart-strategies-to-overcome-banner-blindness/>
- Peláez, J. I., Martínez, E. A., & Vargas, L. G. (2019a). Decision making in social media with consistent data. *Knowledge-Based Systems*, 172, 33–41. <https://doi.org/10.1016/j.knosys.2019.02.009>
- Peláez, J. I., Martínez, E. A., & Vargas, L. G. (2019b). Decision making in social media with consistent data. *Knowledge-Based Systems*, 172, 33–41. <https://doi.org/10.1016/j.knosys.2019.02.009>
- Phua, J., Jin, S. V., & Kim, J. (Jay). (2017). Uses and gratifications of social networking sites for bridging and bonding social capital: A comparison of Facebook, Twitter, Instagram, and Snapchat. *Computers in Human Behavior*, 72, 115–122. <https://doi.org/10.1016/j.chb.2017.02.041>
- Pinto, H., Almeida, J., & Gonçalves, M. (2013). Using early view patterns to predict the popularity of YouTube videos. *WSDM 2013 - Proceedings of the 6th ACM International Conference on Web Search and Data Mining*. <https://doi.org/10.1145/2433396.2433443>
- Popp, B., & Wilson, B. (2018). Investigating the role of identification for social networking Facebook brand pages. *Computers in Human Behavior*, 84, 141–152. <https://doi.org/10.1016/j.chb.2018.01.033>
- Poria, S., Peng, H., Hussain, A., Howard, N., & Cambria, E. (2017). Ensemble application of convolutional neural networks and multiple kernel learning for multimodal sentiment analysis. *Neurocomputing*, 261, 217–230. <https://doi.org/10.1016/j.neucom.2016.09.117>

- Prajogo, D., Toy, J., Bhattacharya, A., Oke, A., & Cheng, T. C. E. (2018). The relationships between information management, process management and operational performance: Internal and external contexts. *International Journal of Production Economics*, *199*, 95–103. <https://doi.org/10.1016/j.ijpe.2018.02.019>
- Proceedings_2018_1.pdf*. (n.d.). Retrieved 11 July 2019, from https://uni.uhk.cz/hed/site/assets/files/1072/proceedings_2018_1.pdf
- Rang Choi, T., & Sung, Y. (2018). Instagram versus Snapchat: Self-expression and privacy concern on social media. *Telematics and Informatics*, *35*. <https://doi.org/10.1016/j.tele.2018.09.009>
- Roberts, D. L., Piller, F. T., & Lüttgens, D. (2016a). Mapping the Impact of Social Media for Innovation: The Role of Social Media in Explaining Innovation Performance in the PDMA Comparative Performance Assessment Study: MAPPING THE IMPACT OF SOCIAL MEDIA FOR INNOVATION. *Journal of Product Innovation Management*, *33*, 117–135. <https://doi.org/10.1111/jpim.12341>
- Roberts, D. L., Piller, F. T., & Lüttgens, D. (2016b). Mapping the Impact of Social Media for Innovation: The Role of Social Media in Explaining Innovation Performance in the PDMA Comparative Performance Assessment Study: MAPPING THE IMPACT OF SOCIAL MEDIA FOR INNOVATION. *Journal of Product Innovation Management*, *33*, 117–135. <https://doi.org/10.1111/jpim.12341>
- Rucker, D. D., Dubois, D., & Galinsky, A. D. (2011). Generous Paupers and Stingy Princes: Power Drives Consumer Spending on Self versus Others. *Journal of Consumer Research*, *37*(6), 1015–1029.
- Sabaté, F. M., Berbegal-Mirabent, J., Cañabate, A., & Lebherz, P. R. (2014). *Factors influencing popularity of branded content in Facebook fan pages*. <https://doi.org/10.1016/j.emj.2014.05.001>
- Saboo, A. R., Kumar, V., & Ramani, G. (2016). Evaluating the impact of social media activities on human brand sales. *International Journal of Research in Marketing*, *33*(3), 524–541.
- Sajid, S. I. (2016). *Social Media and Its Role in Marketing*. <http://repository.embuni.ac.ke/handle/123456789/810>
- Saravanakumar, M., & SuganthaLakshmi, T. (2012). Social Media Marketing. *Life Science Journal*, *9*, 4444–4451.
- Serban, I. V., Sordoni, A., Bengio, Y., Courville, A., & Pineau, J. (2016, March 5). Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models. *Thirtieth AAAI Conference on Artificial Intelligence*. Thirtieth AAAI Conference on

- Artificial Intelligence. <https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/11957>
- Sharma, P. K., Rathore, S., & Park, J. H. (2019). Multilevel learning based modeling for link prediction and users' consumption preference in Online Social Networks. *Future Generation Computer Systems*, *93*, 952–961. <https://doi.org/10.1016/j.future.2017.08.031>
- Sigala, M., & Chalkiti, K. (2015a). Knowledge management, social media and employee creativity. *International Journal of Hospitality Management*, *45*, 44–58. <https://doi.org/10.1016/j.ijhm.2014.11.003>
- Sigala, M., & Chalkiti, K. (2015b). Knowledge management, social media and employee creativity. *International Journal of Hospitality Management*, *45*, 44–58. <https://doi.org/10.1016/j.ijhm.2014.11.003>
- Singh, J., & Goyal, G. (2018). Anticipating movie success through crowdsourced social media videos. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2018.08.050>
- Sponcil, M., & Gitimu, P. (n.d.). *Use of social media by college students: Relationship to communication and self-concept*. 13.
- Statista. (n.d.). *Topic: Social Media Statistics*. www.Statista.Com. Retrieved 11 July 2019, from <https://www.statista.com/topics/1164/social-networks/>
- Steinfeld, C., Ellison, N. B., & Lampe, C. (2008). Social capital, self-esteem, and use of online social network sites: A longitudinal analysis. *Journal of Applied Developmental Psychology*, *29*(6), 434–445. <https://doi.org/10.1016/j.appdev.2008.07.002>
- Stephen, A. T. (2016). The role of digital and social media marketing in consumer behavior. *Current Opinion in Psychology*, *10*, 17–21. <https://doi.org/10.1016/j.copsyc.2015.10.016>
- Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics – Challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management*, *39*, 156–168. <https://doi.org/10.1016/j.ijinfomgt.2017.12.002>
- Sutcliffe, A. G., Binder, J. F., & Dunbar, R. I. M. (2018). Activity in social media and intimacy in social relationships. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2018.03.050>
- Szabo, G., & Huberman, B. A. (2008). Predicting the popularity of online content. *ArXiv:0811.0405 [Physics]*. <http://arxiv.org/abs/0811.0405>

- Tandera, T., Hendro, Suhartono, D., Wongso, R., & Prasetyo, Y. L. (2017). Personality Prediction System from Facebook Users. *Procedia Computer Science*, *116*, 604–611. <https://doi.org/10.1016/j.procs.2017.10.016>
- Thackeray, R., L Neiger, B., Hanson, C., & McKenzie, J. (2008). Enhancing Promotional Strategies Within Social Marketing Programs: Use of Web 2.0 Social Media. *Health Promotion Practice*, *9*, 338–343. <https://doi.org/10.1177/1524839908325335>
- Tiago, M. T. P. M. B., & Verissimo, J. M. C. (2014). Digital marketing and social media: Why bother? *Business Horizons*, *57*(6), 703–708.
- Trnovcová, D. (2016). INTERNÁ KOMUNIKÁCIA A JEJ STRATEGICKÁ FUNKCIA PRE PODNIK. *Hradecké Ekonomické Dny 2016*, 211–217.
- Tsikerdekis, M., & Zeadally, S. (2014). Online Deception in Social Media. *Commun. ACM*, *57*(9), 72–80. <https://doi.org/10.1145/2629612>
- Tsur, O., & Rappoport, A. (2012). *What's in a Hashtag? Content based Prediction of the Spread of Ideas in Microblogging Communities*. 643–652. <https://doi.org/10.1145/2124295.2124320>
- Tuten, T. L., & Solomon, M. R. (2018). *Social media marketing*.
- van Esch, P., & Mente, M. (2018). Marketing video-enabled social media as part of your e-recruitment strategy: Stop trying to be trendy. *Journal of Retailing and Consumer Services*, *44*, 266–273. <https://doi.org/10.1016/j.jretconser.2018.06.016>
- Vasilev, I., Slater, D., Spacagna, G., Roelants, P., & Zocca, V. (2019). *Python Deep Learning: Exploring deep learning techniques and neural network architectures with PyTorch, Keras, and TensorFlow, 2nd Edition*. Packt Publishing Ltd.
- Vázquez, S., Muñoz-García, Ó., Campanella, I., Poch, M., Fisas, B., Bel, N., & Andreu, G. (2014). A classification of user-generated content into consumer decision journey stages. *Neural Networks*, *58*, 68–81. <https://doi.org/10.1016/j.neunet.2014.05.026>
- Vybíral, Z. (2009). *Psychologie komunikace*. <https://www.martinus.cz/?uItem=63541>
- Wagner, T. F. (2017). Promoting Technological Innovations: Towards an Integration of Traditional and Social Media Communication Channels. In G. Meiselwitz (Ed.), *Social Computing and Social Media. Human Behavior* (pp. 256–273). Springer International Publishing.
- Warner-Søderholm, G., Bertsch, A., Sawe, E., Lee, D., Wolfe, T., Meyer, J., Engel, J., & Fatilua, U. N. (2018). Who trusts social media? *Computers in Human Behavior*, *81*, 303–315. <https://doi.org/10.1016/j.chb.2017.12.026>

- Weng, L., Menczer, F., & Ahn, Y.-Y. (2014, May 16). Predicting Successful Memes Using Network and Community Structure. *Eighth International AAAI Conference on Weblogs and Social Media*. Eighth International AAAI Conference on Weblogs and Social Media. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8081>
- Wright, K. B., Bell, S. B., Wright, K. B., & Bell, S. B. (2003). Health-related Support Groups on the Internet: Linking Empirical Findings to Social Support and Computer-mediated Communication Theory. *Journal of Health Psychology*, 8(1), 39–54. <https://doi.org/10.1177/1359105303008001429>
- Wu, J., Chiclana, F., Fujita, H., & Herrera-Viedma, E. (2017). A Visual Interaction Consensus Model for Social Network Group Decision Making with Trust Propagation. *Know.-Based Syst.*, 122(C), 39–50. <https://doi.org/10.1016/j.knosys.2017.01.031>
- Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management*, 31(2), 179–188. <https://doi.org/10.1016/j.tourman.2009.02.016>
- Zeng, B., & Gerritsen, R. (2014). What do We Know About Social Media in Tourism? A Review. *Tourism Management Perspectives*, 10, 27–36. <https://doi.org/10.1016/j.tmp.2014.01.001>
- Zeng, D. D., Chen, H., Lusch, R., & Li, S.-H. (2011). Social Media Analytics and Intelligence. *Intelligent Systems, IEEE*, 25, 13–16. <https://doi.org/10.1109/MIS.2010.151>
- Zerfass, A., Tench, R., Verčič, D., Verhoeven, P., & Moreno, Á. (2014). *European Communication Monitor 2014. Excellence in Strategic Communication—Key Issues, Leadership, Gender and Mobile Media. Results of a Survey in 42 Countries*.
- Zhang, X., Chen, X., Chen, Y., Wang, S., Li, Z., & Xia, J. (2015). Event detection and popularity prediction in microblogging. *Neurocomputing*, 149, 1469–1480. <https://doi.org/10.1016/j.neucom.2014.08.045>
- Zheng, W., Yuan, C.-H., Chang, W.-H., & Wu, Y.-C. J. (2016). Profile pictures on social media: Gender and regional differences. *Computers in Human Behavior*, 63, 891–898. <https://doi.org/10.1016/j.chb.2016.06.041>
- Zheng, X., QianZhou, D., YuFeng, M., & WeiGuo, F. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*, 58, 51–65.

Zhou, L., Zhang, D., Yang, C. C., & Wang, Y. (2018). Harnessing social media for health information management. *Electronic Commerce Research and Applications*, 27, 139–151. <https://doi.org/10.1016/j.elerap.2017.12.003>

11 Uvedení vlastních publikací disertanta

Publikace disertanta souvisejících s tématem:

Časopisecké články indexované v JCR WoS

Hruška, J., & Pásková, M. (2018). How do the National Tourism Organizations use the social media? *E+M Ekonomie a Management*, 21(4), 226–240. <https://doi.org/10.15240/tul/001/2018-4-015>

Časopisecké články indexované v SJR SCOPUS

Pásková, M., Hruška, J., & Zelenka, J. (2018). YouTube as an Airlines Marketing Tool. *Czech Journal of Tourism*, 7(1), 45–69. <https://doi.org/10.1515/cjot-2018-0003>

Maresova, P., Hruska, J., & Kuca, K. (2020). Social Media University Branding. *Education Sciences*, 10(3), 74. <https://doi.org/10.3390/educsci10030074>

Hruska, J., & Maresova, P. (2020). Use of Social Media Platforms among Adults in the United States—Behavior on Social Media. *Societies*, 10(1), 27. <https://doi.org/10.3390/soc10010027>

Zelenka, J., & Hruška. (2018). Ways and effectiveness of social media utilization by airlines. *Tourism: An International Interdisciplinary Journal*, 66(2), 227–238.

Konferenční publikace

Hruška, J. (2020). Sentiment Analysis of National Tourism Organizations on Social Media (P. Maresova, P. Jedlicka, K. Firlej, & I. Soukal, Eds.; pp. 250–256). <https://doi.org/10.36689/uhk/hed/2020-01-028>

Hruska, J., & Maresova, P. (2019). Design of Business Canvas Model for Social Media. In A. Abraham, P. Dutta, J. K. Mandal, A. Bhattacharya, & S. Dutta (Eds.), *Emerging Technologies in Data Mining and Information Security* (pp. 63–69). Springer. https://doi.org/10.1007/978-981-13-1498-8_6

Hruska, J. (2017, August 20). CONSUMER POWER AND EMPOWERMENT IN THE CONTEXT OF MARKETING ON SOCIAL NETWORKS. *4th SGEM International Multidisciplinary Scientific Conferences on SOCIAL SCIENCES and ARTS Proceedings*. <https://doi.org/10.5593/sgemsocial2017/41/S16.016>

Hruška, J., & Marešová, P. (2017). ANALYZE AND PREDICTIONS FOR MARKETING ON SOCIAL MEDIA. *Hradec Economic Days (HED)*, (pp. 303–309). Czech Republic.

Seznam všech publikovaných prací během dosavadního doktorského studia:

Impaktované časopisy

- Maresova, P., Hruska, J., Klimova, B., Barakovic, S., & Krejcar, O. (2020). Activities of Daily Living and Associated Costs in the Most Widespread Neurodegenerative Diseases: A Systematic Review. *Clinical Interventions in Aging*, Volume 15, 1841–1862. <https://doi.org/10.2147/CIA.S264688>
- Hruška, J., & Pásková, M. (2018). How do the National Tourism Organizations use the social media? *E+M Ekonomie a Management*, 21(4), 226–240. <https://doi.org/10.15240/tul/001/2018-4-015>

Časopisy SJR

- Maresova, P., Hruska, J., & Kuca, K. (2020). Social Media University Branding. *Education Sciences*, 10(3), 74. <https://doi.org/10.3390/educsci10030074>
- Hruška J., & Zelenka J. (2020). Success factors of airlines YouTube videos. (odesláno 9. 1. 2020 do „e-Review of Tourism Research“). - „under review“
- Hruška, J., & Marešová, P. (2020). Use of neural networks and classification models for more effective decision making on social media - systematic review. - „under review“
- Zelenka, J., & Hruška. (2018). Ways and effectiveness of social media utilization by airlines. *Tourism: An International Interdisciplinary Journal*, 66(2), 227–238

Konference

- Marešová, P., & Hruška, J. (2020). Management of Overweight and Obesity: Technology-Based Interventions Among Generation Y. *Smart Homecare Technology and TeleHealth*, Volume 7, 1–8. <https://doi.org/10.2147/SHTT.S214606>
- Hruska, J., & Maresova, P. (2020). Use of Social Media Platforms among Adults in the United States—Behavior on Social Media. *Societies*, 10(1), 27. <https://doi.org/10.3390/soc10010027>
- Hruška, J. (2020). Sentiment Analysis of National Tourism Organizations on Social Media (P. Maresova, P. Jedlicka, K. Firlej, & I. Soukal, Eds.; pp. 250–256). <https://doi.org/10.36689/uhk/hed/2020-01-028>
- Hruska, J., & Maresova, P. (2019). Design of Business Canvas Model for Social Media. In A. Abraham, P. Dutta, J. K. Mandal, A. Bhattacharya, & S. Dutta (Eds.), *Emerging Technologies in Data Mining and Information Security* (pp. 63–69). Springer. https://doi.org/10.1007/978-981-13-1498-8_6

- Matulová, P., & Hruška, J. (2019). Effective Support of Transfer Technology on Regional Level (P. Maresova, P. Jedlicka, & I. Soukal, Eds.; pp. 79–86). <https://doi.org/10.36689/uhk/hed/2019-02-007>
- Zelenka, J. & Hruška, J. (2018) Youtube as airlines information and marketing tool. *Hradec Economic Days (HED)*, In: HED 2018. Czech Republic, pp 514-526, ISBN 978-80-7435-701-5.
- Pásková, M., Hruška, J., & Zelenka, J. (2018). YouTube as an Airlines Marketing Tool. *Czech Journal of Tourism*, 7(1), 45–69. <https://doi.org/10.1515/cjot-2018-0003>
- Hruska, J. (2017, August 20). CONSUMER POWER AND EMPOWERMENT IN THE CONTEXT OF MARKETING ON SOCIAL NETWORKS. *4th SGEM International Multidisciplinary Scientific Conferences on SOCIAL SCIENCES and ARTS Proceedings*. <https://doi.org/10.5593/sgemsocial2017/41/S16.016>
- Hruška, J., & Marešová, P. (2017). ANALYZE AND PREDICTIONS FOR MARKETING ON SOCIAL MEDIA. *Hradec Economic Days (HED)*, (pp. 303–309). Czech Republic.

Přehled odborných vědecko-výzkumných aktivit vykonaných během mého studia v doktorském studijním programu:

Externí grantové projekty

- Projekt OP VVV 4823/02 IT4Neuro(degeneration). Junior Researcher – analytik na projektu, VZ2, Prevence a léčba osob s neurodegenerativním onemocněním.
- TAČR ÉTA TL01000300 Léčba a péče o osoby s Alzheimerovou chorobou – ekonomická zátěž v kontextu perspektiv vývoje nových léků: Junior Researcher – analytik na projektu.

Interní grantová schémata

- Excelence FIM 2017/18: Pokročilé metody zhodnocení ekonomické efektivity v oblasti zdravotnictví (řešitelka doc. Ing. Mgr. Petra Marešová, Ph.D.) – Zpracovávání a vyhodnocování dat.
- SPEV (2018–2019): Information and knowledge management and cognitive science in tourism – výzkum chování dominantních leteckých společností na sociální síti YouTube a tvoření doporučení pro úspěch na této sociální síti v této oblasti a v turismu celkově (řešitel prof. RNDr. Josef Zelenka, CSc.).
- SPEV (2017–2019): Investice v podmínkách konceptu Industry 4.0 – analýza změn v oblasti inovací se speciálním zaměřením na chování na sociálních sítích (řešitelka doc. Ing. Mgr. Petra Marešová, Ph.D.).

Smluvní výzkum

Inovační voucher (2018) „Sociální média pro zvýšení efektivity uvnitř firmy" pro firmu Savoir-faire s. r. o. – Sběr a analýza dat, tvorba tabulek, grafů a doporučení (vedoucí doc. Ing. Mgr. Petra Marešová, Ph.D.).

Inovační voucher (2019) „Model pro efektivní interní komunikaci uvnitř firmy" pro firmu Savoir-faire s.r.o. – Sběr a analýza dat, tvorba tabulek, grafů a doporučení (vedoucí doc. Ing. Mgr. Petra Marešová, Ph.D.).

Spolupráce Univerzity Hradec Králové se Škoda Auto: Statistické řešení nákupu součástek na after-sales – Člen týmu pro statistické zpracování (vedoucí prof. RNDr. Hana Skalská, CSc.).

Spolupráce na projektu „Výzkum a vývoj nového SW řešení – Inteligentní robotické řízení dokumentů“ jako výzkumník pro firmu Onlio (řešitel prof. Ing. Vladimír Bureš, Ph.D., MBA).

Seznam příloh

Příloha č. 1: Dotazník	121
Příloha č. 2: Regresní analýza	124

Přílohy

Příloha č. 1

Dotazník

Uvedené otázky v dotazníku:

- 1) Pohlaví?
 - a. Muž
 - b. Žena
- 2) Věk?
- 3) Stav (Ženatý, svobodný/á, rozvedený/á)?
 - a. Vdaná (ženatý)
 - b. Svobodná (svobodný)
 - c. Rozvedená (rozvedený)
- 4) Stále studuji?
 - a. Ano
 - b. Ne
- 5) Dokončené vzdělání?
 - a. Základní
 - b. Středoškolské bez maturity
 - c. Středoškolské s maturitou
 - d. Vysokoškolské
- 6) Čistý měsíční příjem?
 - a. Do 9 999 Kč/měsíc
 - b. 10 000–19 999 Kč/měsíc
 - c. 20 000–29 999 Kč/měsíc
 - d. 30 000–39 999 Kč/měsíc
 - e. 40 000–49 999 Kč/měsíc
 - f. Více
- 7) Počet obyvatel místa bydliště?
 - a. Do 499
 - b. 500 až 999
 - c. 1 000 až 1 999
 - d. 2 000 až 4 999
 - e. 5 000 až 9 999
 - f. 10 000 až 49 999
 - g. Více
- 8) Na jakých sociálních sítích máte profil?
 - a. Facebook
 - b. YouTube (vlastní kanál)
 - c. WhatsApp
 - d. Instagram

- e. Pinterest
 - f. Twitter
 - g. WeChat
 - h. QZone
 - i. Tumblr
 - j. QQ
 - k. Snapchat
 - l. LinkedIn
 - m. Žádné z uvedených
- 9) Na sociálních sítích řeším témata (komentuji, lajkuji, reaguji)?
- a. Politika
 - b. Kriminalita
 - c. Věda
 - d. Média/Hudba/TV
 - e. Transport
 - f. Technologie
 - g. Zábava
 - h. Sport
 - i. Nakupování
 - j. Celebrity
 - k. Místní zprávy
 - l. Počasí
 - m. Jídlo
 - n. Vzdělávání
 - o. Jiná...
- 10) Kdo je většinou autorem vámi sdílených příspěvků?
- a. Já sám jako jednotlivec (např. selfie, názory)
 - b. Přátelé, které osobně znám
 - c. Přátelé, které znám pouze virtuálně (ze sítí)
 - d. Zcela náhodní autoři (např. humorných příspěvků nebo videí)
 - e. Známé osobnosti
 - f. Firmy a podnikatelé nabízející své služby nebo výrobky
 - g. Státní instituce
 - h. Jiná...
- 11) Z jakého důvodu sdílíte příspěvky, co Vás motivuje?
- a. Chci se pochlubit (např. dovolenou, úspěchem, romantickou večeří...)
 - b. Chci se podělit o strasti, aby mě druzí pochopili a možná politovali
 - c. Chci poznat názory ostatních, např. sdílením kontroverzních témat
 - d. Chci se podělit o radost/pobavit ostatní
 - e. Pro kontakt s přáteli, chci se podělit s přáteli o právě zjištěné informace
 - f. Chci podpořit své přátele (např. v nouzi, v ohrožení, při důležitém životním kroku...)

- g. Nabízím své vlastní výrobky, služby, umění...
 - h. Jiná...
- 12) Na sociálních sítích sdílím většinou příspěvky týkající se témat?
- a. Politika
 - b. Kriminalita
 - c. Věda
 - d. Média/Hudba/TV
 - e. Transport
 - f. Technologie
 - g. Zábava
 - h. Sport
 - i. Nakupování
 - j. Celebrity
 - k. Místní zprávy
 - l. Počasí
 - m. Jídlo, zdravý životní styl
 - n. Vzdělávání
 - o. Ekonomika, finance, bankovníctví
 - p. Kultura
 - q. Podnikání
 - r. Jiná...
- 13) Odkud a z jakého zařízení sdílíte příspěvky?
- a. Z pracoviště/školy
 - b. Z místa bydliště
 - c. Z mobilního zařízení
 - d. Z počítače nebo tabletu
 - e. Jak chci (na místě nebo zařízení nezáleží)
- 14) Prohlížíte si sdílené příspěvky podrobněji v případě, že obsahují vulgarismy /negativitu? (Otevřená odpověď)
- 15) Prohlížíte si sdílené příspěvky podrobněji v případě, že obsahují pozitivní slova/optimismus? (Otevřená odpověď)
- 16) Kolik myslíte, že zhruba sdílíte příspěvků za rok? Je jedno, na jaké sociální síti. (Otevřená odpověď)
- 17) Ověřujete si pravidelně informace z příspěvků na sociálních sítích?
- a. Ano
 - b. Občas ano
 - c. Převážně ne
 - d. Nikdy
- 18) Jaké příspěvky jste naposledy sdílely a na jaké sociální síti to bylo? (Otevřená odpověď)

Dotazník lze nalézt na této adrese: <https://forms.gle/hX5nwVtQ1pze5aDW6>.

Regresní analýza

Tabulka 14: Regresní analýza pro všechny proměnné v datovém souboru (vlastní dílo autora)

Model	Nestandardizované koeficienty		Standardizované koeficienty	t	Sig.
	B	Chyba	Beta		
(Konstanta)	-	739784,698		-2,361	,018
	1746370,392				
TitleWords	164,501	385,019	,021	,427	,669
TitleCharacters	-23,950	62,729	-,017	-,382	,703
TotalWordsInArticle	10,970	7,819	,264	1,403	,161
TotalCharactersInArticle	-2,187	1,303	-,313	-1,679	,093
TotalParagraphs	8,586	26,847	,005	,320	,749
AverageWordLengthInTitle	259,401	493,359	,016	,526	,599
AverageWordLengthInArticle	8,003	668,977	,000	,012	,990
NumberOfImages	-11,866	22,104	-,007	-,537	,591
NumberOfHrefs	16,941	12,860	,020	1,317	,188
PositiveWordsInArticle	19,153	26,057	,022	,735	,462
NegativeWordsInArticle	51,568	38,046	,024	1,355	,175
PositiveDividedByNegativePolarityInArticle	-40,301	28,885	-,015	-1,395	,163
PositiveDividedByAllWordsInArticle	7949,223	10706,388	,010	,742	,458
NegativeDividedByAllWordsInArticle	-39858,449	16328,018	-,032	-2,441	,015
PositiveWordsInTitle	-203,535	801,997	-,009	-,254	,800
NegativeWordsInTitle	810,812	1057,471	,027	,767	,443
PositiveWordsInTitleDividedWordsInTitle	4823,171	6656,435	,025	,725	,469
NegativeWordsInTitleDividedWordsInTitle	-6048,068	8648,946	-,025	-,699	,484
WordsInFirstParagraph	24,883	46,829	,030	,531	,595
CharactersInFirstParagraph	-7,660	7,711	-,055	-,993	,321
PositiveInFirstParagraph	160,149	179,224	,016	,894	,372
NegativeInFirstParagraph	722,468	299,761	,041	2,410	,016

PositiveInFirstDividedByWordsInFirst	-4224,376	5142,602	-,012	-,821	,411
NegativeInFirstDividedByWordsInFirst	-7645,830	8554,988	-,014	-,894	,371
DayPosted	1,213	12,710	,001	,095	,924
MonthPosted	141,420	35,865	,035	3,943	,000
YearPosted	868,352	367,427	,021	2,363	,018

Dependent Variable: NumberOfShares

Tabulka 15: Zobrazení regresní analýzy jen pro některé vybrané proměnné podle statistické významnosti (vlastní dílo autora)

Model	Nestandardizované koeficienty		Standardizované koeficienty	t	Sig.
	B	Chyba	Beta		
(Constant)	3209,441	329,083		9,753	,000
TotalWordsInArticle	13,363	5,083	,321	2,629	,009
TotalCharactersInArticle	-2,463	,856	-,352	-2,878	,004
NumberOfHrefs	16,239	6,888	,019	2,358	,018
NegativeWordsInArticle	62,614	36,135	,029	1,733	,083
NegativeDividedByAllWordsInArticle	32915,968	15083,964	-,027	-2,182	,029
NegativeInFirstParagraph	422,571	156,072	,024	2,708	,007
MonthPosted	103,775	31,597	,025	3,284	,001

Dependent Variable: NumberOfShares