



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

**ANALÝZA TÉMAT A ŠÍŘENÍ DEZINFORMACÍ Z PRO-
PAGANDISTICKÝCH WEBOVÝCH STRÁNEK**

ANALYSIS OF TOPICS AND SPREADING OF DISINFORMATION FROM PROPAGANDIST
WEB PAGES

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

DALIBOR KYJOVSKÝ

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. KAREL ONDŘEJ

BRNO 2023

Zadání bakalářské práce



148612

Ústav: Ústav počítačové grafiky a multimédií (UPGM)
Student: **Kyjovský Dalibor**
Program: Informační technologie
Specializace: Informační technologie
Název: **Analýza témat a šíření dezinformací z propagandistických webových stránek**
Kategorie: Umělá inteligence
Akademický rok: 2022/23

Zadání:

1. Seznamte se s metodami a nástroji pro obecné shlukování a extrakci témat z textu a pro odhalování dezinformačních zpráv a webových zdrojů.
2. Na základě získaných znalostí navrhnete a implementujete systém, který dokáže identifikovat nová témata dezinformačních kampaní se zaměřením na šíření propagandy
3. Vytvořte systém pro pravidlnou aktualizaci informací, aby bylo možné sledovat schémata konkrétních dezinformačních kampaní v čase.
4. Vyhodnoťte systém na shromážděných datech a demonstруйте jeho výhody na vhodně zvolených příkladech.
5. Vytvořte stručný plakát prezentující práci, její cíle a výsledky.

Literatura:

- dle doporučení vedoucího

Při obhajobě semestrální části projektu je požadováno:

- funkční prototyp řešení

Podrobné závazné pokyny pro vypracování práce viz <https://www.fit.vut.cz/study/theses/>

Vedoucí práce: **Ondřej Karel, Ing.**
Vedoucí ústavu: Černocký Jan, prof. Dr. Ing.
Datum zadání: 1.11.2022
Termín pro odevzdání: 10.5.2023
Datum schválení: 31.10.2022

Abstrakt

Tato práce se zabývá problematikou dezinformací na propagandistických webových stránkách v České republice a návrhem systému, který by měl sloužit k analýze a rozpoznání dezinformačního charakteru obsahu těchto stránek. Navrhovaný a implementovaný systém dokáže stahovat a archivovat data z propagandistických webů a následně analyzovat jejich obsah. Systém využívá především znalostní a síťové metody. Cílem práce je využití moderních technologií pro odhalování dezinformací v médiích. Výsledky práce mohou pomoci při dalším výzkumu a při prevenci šíření dezinformací.

Abstract

This thesis deals with the issue of disinformation on propaganda websites in the Czech Republic and the design of a system that should be used to analyze and recognize the disinformation character of the content of these websites. The proposed and implemented system is able to download and archive data from propaganda websites and then analyze their content. The system mainly uses knowledge-based and network analyses. The aim of the thesis is to use modern technologies for detecting disinformation in the media. The results may help in further research and in preventing the spread of disinformation.

Klíčová slova

Dezinformace, odhalování dezinformací, propagandistické stránky, strojové učení, BERT, datové sady

Keywords

Disinformation, disinformation detection, propaganda pages, machine learning, BERT, datasets

Citace

KYJOVSKÝ, Dalibor. *Analýza témat a šíření dezinformací z propagandistických webových stránek*. Brno, 2023. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Karel Ondřej

Analýza témat a šíření dezinformací z propagandistických webových stránek

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením Ing. Karla Ondřeje. Další informace mi poskytl doc. RNDr. Pavel Smrž, Ph.D. Uvedl jsem všechny literární prameny, publikace a další zdroje, ze kterých jsem čerpal.

.....
Dalibor Kyjovský
9. května 2023

Poděkování

Rád bych poděkoval doc. RNDr. Pavlu Smržovi, Ph.D. za vstřícný přístup a rady, které mi poskytl při tvorbě této bakalářské práce. Děkuji své rodině a Papčovi za podporu a trpělivost.

Obsah

1	Úvod	3
2	Dezinformace v kontextu moderních technologií	5
2.1	Definice klíčových pojmů	5
2.2	Společenský problém	8
2.3	Přístupy k odhalování dezinformací	9
3	Zpracování přirozeného jazyka	14
3.1	Předzpracování textu	15
3.2	Metody extrakce příznaků	16
3.3	Klasifikační techniky	20
3.4	Evaluační metriky	20
4	Datové sady	22
5	Návrh systému a jeho limity	24
5.1	Zvolený přístup k problému	25
5.2	Návrh řešení	26
5.3	Limity předkládané práce	31
6	Implementace	32
6.1	Struktura systému	32
6.2	Přístup k databázi	33
6.3	Stahování dat	34
6.4	Zpracování dat	35
6.5	Uživatelské prostředí	36
6.6	Vyhodnocení implementace systému	37
7	Závěr	38
	Literatura	40
	A Obsah příloženého paměťového média	44
	B Manuál pro spuštění systému	45
	C Seznam dezinformačních webů	46
	D Plakát	48

Seznam obrázků

2.1	Typy zkrácení informací	6
2.2	Trend vyhledávání dezinformací, převzato z Google Trends [5].	8
2.3	Různorodost přístupů k odhalování dezinformací, převzato z [12].	9
2.4	Typy dat v článku, převzato z [30]	10
3.1	Proces zpracování textu.	16
3.2	Bag-of-Words Model.	17
3.3	Word2Vec modely, převzato z [26].	19
3.4	BERT, převzato z [14].	20
5.1	Schéma propojení systému a uživatele.	25
5.2	Schéma systému a jeho jednotlivé závislosti.	26
5.3	Schéma databáze.	27
5.4	Schéma získávání html kódu.	28
6.1	Příklad staženého článku z webu.	37
6.2	Příklad nalezených url z textu článků.	37
D.1	Náhled vytvořeného plakátu.	48

Kapitola 1

Úvod

Fenomén dezinformací v posledních letech ve společnosti rezonuje čím dál intenzivněji. Bývá zmiňován zejména v souvislosti s možným ovlivňováním politického vývoje a nejen v českém prostředí v souvislosti s rozdělováním společnosti na více vyhraněných názorových skupin. Faktem zůstává, že dezinformace byly, jsou a zřejmě i nadále budou součástí lidské společnosti, otázkou zůstává, jakým způsobem nás budou v životech provázet a ovlivňovat a jak jim budeme schopni čelit.

Propagandistických webových stránek zveřejňujících dezinformační obsah je v České republice provozováno hned několik desítek. Tyto stránky jsou mnoha institucemi a neziskovými organizacemi sledovány a jejich seznam¹ je jimi pravidelně zveřejňován a aktualizován.

Teoretická část práce sestává ze tří kapitol. První kapitola vymezuje stěžejní terminologii a definuje různorodé přístupy k odhalování dezinformací. V této kapitole bylo čerpáno z výzkumu *A Survey on Automated Fact-Checking* [17] a *Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study* [12], které vznikly v posledních letech. Aktuálnost výzkumů je důležitá zejména z důvodu rychlého vývoje metod a technologií, které jsou k tomuto účelu využívány. Další kapitola rozebírá metody strojového učení, které jsou zmiňovány v nespočtu relevantních výzkumů a reflektují vývoj přístupu k této problematice v posledních letech. Datové sady potřebné pro využití analýzy dat k prevenci šíření dezinformací a pro trénování modelů jsou popsány ve třetí kapitole této části. Teoreticky se práce opírá o práce předních vědeckých odborníků z daného odvětví, jež zejména v posledních letech zaznamenalo značný boom.

Praktická část je věnována konkrétnímu návrhu systému, jehož úkolem bude stahovat a archivovat data z propagandistických webů a následně podrobně analyzovat jejich obsah za účelem rozpoznání jejich dezinformačního charakteru. Analýza bude prováděna různými metodami, které jsou teoreticky popsány v teoretické části práce. Lze předpokládat, že během analýzy vyvstanou možné komplikace, které zúží výběr využitelných metod pro analýzu získaných dat.

V poslední části práce autor přistupuje k detailnímu popisu implementace jím navrhovaného systému, který by měl sloužit k výše zmíněným účelům. Je zde popsána struktura systému a jeho základní moduly, které tvoří stěžejní součást systému. Práce se věnuje také konkrétnímu návrhu a realizaci podoby uživatelského prostředí webové aplikace.

¹Seznamy dezinformačních webů jsou pravidelně zveřejňovány mnoha subjekty, mezi nimi např. ve výroční zprávě *Bezpečnostního centra evropských hodnot* za rok 2022 (https://europeanvalues.cz/wp-content/uploads/2023/04/CS_Vyrocní_zprava_o_stavu_ceske_dezinformacni_sceny_pro_rok_2022.pdf), na webu *Nelež.cz* (<https://www.nelez.cz>) nebo na webu *Neovlivní.cz* (<https://neovlivni.cz/category/databaze-lzi/>)

Práce *Analýza témat a šíření dezinformací z propagandistických webových stránek* propojuje problematiku dezinformací ve společnosti a politice se studiem informačních technologií. Cílem práce je uvést čtenáře do problematiky odhalování dezinformací a využití moderních technologií k tomuto účelu.

Kapitola 2

Dezinformace v kontextu moderních technologií

V době pokročilých technologických řešení a vysokého výkonu výpočetních zdrojů by se mohlo zdát, že boj proti dezinformacím a manipulaci ve veřejném prostoru by měl být snadnější. Avšak v praxi je situace právě opačná. Dezinformace se v současné době šíří na internetu a sociálních sítích mnohem rychleji a efektivněji než kdy předtím. V České republice existuje množství dezinformačních webů, které se tváří jako ověřená média, avšak ve skutečnosti zkreslují fakta a manipulují s nimi, což ztěžuje boj proti nim.

2.1 Definice klíčových pojmů

V této kapitole budou vysvětleny zásadní pojmy, s nimiž práce operuje a jež je nutné znát pro správné uchopení problematiky dezinformací a jejich následnou analýzu.

Dezinformace

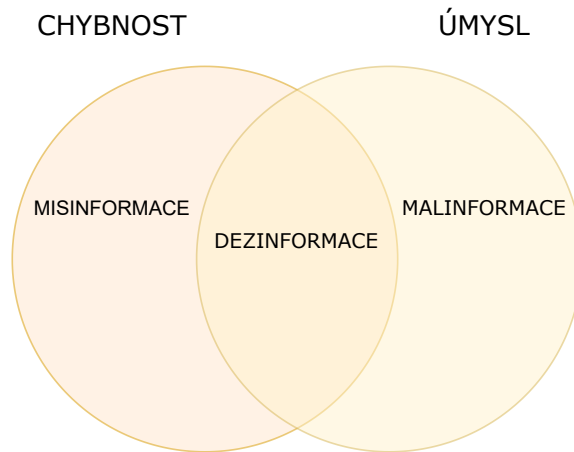
Dezinformace je kombinací **misinformace**¹ a **malinformace**². Je chybná (jako misinformace), ale je úmyslná (jako malinformace) a jejím cílem je někoho poškodit a/nebo obelhat. [10]

Různé zdroje dezinformaci jako takovou definují s mírnými odlišnostmi, proto zde bude uvedeno více příkladů těchto definic.

Atlas konspirací [3] dezinformaci chápe „jako ověřitelnou, falešnou nebo zavádějící informaci, která je vytvořena, prezentována a šířena za účelem ekonomického zisku nebo slouží jako nástroj pro záměrné zmatení veřejnosti s cílem ji poškodit.“ Oproti tomu pro Ministerstvo vnitra [7] „pojem dezinformace znamená šíření záměrně nepravdivých informací, obzvláště pak státními aktéry nebo jejich odnožemi vůči cizímu státu nebo vůči médiím, s cílem ovlivnit rozhodování nebo názory těch, kteří je přijímají.“ Podle České terminologické databáze knihovnictví a informační vědy [24] se jedná o „záměrně nepravdiv[ou] (falešná, lživá, nesprávná, zkreslená) informac[i] sdělovan[ou] s cílem uvést v omyl a ovlivnit příjemce tím, že ji bude považovat za pravdivou a důvěryhodnou.“ Don Fallis z Johns Hopkins University Press definici ještě rozšiřuje a tvrdí, že mezi prototypové případy dezinformací patří klamavá reklama, vládní propaganda, upravené fotografie, padělané dokumenty, fa-

¹Chybná informace, jejímž účelem není úmyslné poškození jedince.

²Jedná se o pravdivou informaci, která je vypuštěna za účelem někoho poškodit.



Obrázek 2.1: Typy zkreslení informací

lešné mapy, internetové podvody, falešné webové stránky a zmanipulované záznamy na Wikipedii. [15]

Pro potřeby této práce bude dezinformace definována několika následujícími zásadními body, jež vychází z výše zmíněných definic.

- Cílem dezinformací je negativně ovlivňovat jedince.
- Informace, které jsou jedinci sdělovány, jsou alespoň z části nepravdivé.
- Jedná se o úmyslně zavádějící informace.

Dezinformace lze rozdělit na dva základní typy:

- **Dezinformace aktivní**

Dochází k manipulaci již existující informace, ve smyslu jejího zpoždění, zatajení či likvidace. [33]

- **Dezinformace pasivní**

Dezinformace jsou vytvářeny, anebo záměrně pozměňovány. [33]

Propaganda

Termín propaganda ve veřejném prostoru v posledních letech rezonuje čím dál tím více. Tento termín je v dnešní době spojován především s šířením dezinformací a ovlivňováním veřejného mínění ve prospěch politických subjektů a režimů.

Dle Ministerstva vnitra České republiky je pojem propaganda v aktuálním kontextu mezinárodního vývoje nejčastěji spojován s politickým přesvědčováním a psychologickou válkou – ta je definována jako použití propagandy proti protivníkovi s cílem zlomit jeho vůli bojovat nebo se bránit, případně jej naklonit vlastní pozici. [7]

Propagandu lze rozdělit do tří základních skupin:

- **Bílá propaganda**

Objevuje se ve veřejně známých médiích, které využívají své pozice média všeobecně širokou veřejností považovaného za ověřené. Nejprve se snaží budovat svou kredibilitu, již následně využívají k rozšiřování propagandy.

- **Černá propaganda**

Snaží se vzbuzovat dojem, že informace pochází z ověřeného zdroje, nicméně skutečný zdroj je jiný, skryté jsou také záměry tohoto typu propagandy.

- **Šedá propaganda**

Kombinuje oba předešlé způsoby a snaží se těžit z jejich předností.

Zkreslení a mentální zkratky

- **Kotvení**

Toto kognitivní zkreslení se projevuje kladením důrazu na první dostupnou informaci o daném tématu. První informace funguje jako tzv. „kotva“. Vysvětluje to společenskou tendenci pamatovat si hoaxy, přestože jsou následně několikrát vysvětleny. [11]

- **Rámování**

Jedná se o vytváření vlastních konceptů a perspektiv reality. [11]

- **Stádový efekt**

Za tímto zkreslením stojí tendence přijímat myšlenky a názory, které zastávají i ostatní lidé. Pokud se jedinec ocitne v sociální bublině bez přítomnosti oponentů jeho názorů, bude mít větší sklony zastávat totožný názor jako tato skupina, již se cítí být členem. [11]

- **Konfirmační zkreslení**

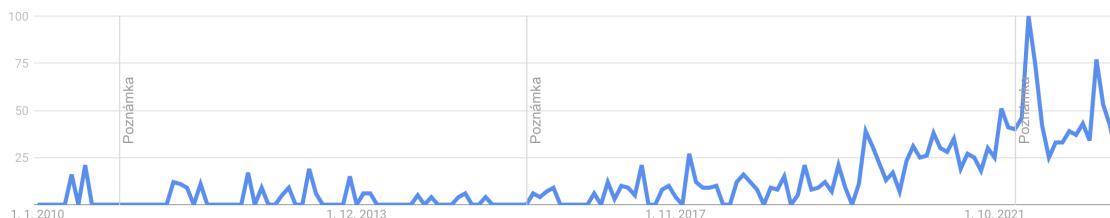
Tendence upřednostňovat informace, které podporují již existující názory a postoje jedince. [11]

Propagandistické stránky

Propagandistické stránky, které využívají buď bílou, černou, nebo šedou propagandu jsou označeny Ministerstvem vnitra za nebezpečné kvůli důvodnému podezření z šíření dezinformací. Seznam propagandistických stránek označených za dezinformační, lze nalézt v příloze C této práce.

2.2 Společenský problém

Dezinformace se ve veřejném prostoru vyskytují odjakživa. Již od 13. století evidujeme zmínky o dezinformacích. Hojně užívaným nástrojem se dezinformace staly také v průběhu první a druhé světové války. V posledních letech se rozšířily zejména s roustoucím připojením uživatelů k internetu. Na obrázku 2.2 lze vidět vzrůstající počty vyhledávání temínu *dezinformace* ve vyhledávači Google.



Obrázek 2.2: Trend vyhledávání dezinformací, převzato z Google Trends [5].

Boj proti dezinformacím v České Republice závisí především na neziskových organizacích jako je Demagog³, Hlídací pes⁴, Manipulátoři⁵ aj. Tyto organizace se snaží především objasňovat zavádějící informace, vyvracet nepravdy a uvádět je na pravou míru.

Způsoby, kterými se nejčasteji šíří dezinformace:

- Sociální sítě
- Řetězové e-maily
- Propagandistické webové stránky

Dle Českých elfů ovlivňuje fenomén řetězových e-mailů desetitisíce lidí. E-maily jsou založeny zejména na mechanickém přeposílání dalším uživatelům, přičemž obsahují většinou bulvární, dryáčnický či na emoce útočící obsah. [4] V roce 2021 Čeští elfové zpřístupnili široké veřejnosti databázi⁶, která obsahuje přes deset tisíc textů z řetězových e-mailů jak v českém, tak v evropském kontextu. Podle výzkumu organizace Elpida a Nadace O2 z roku 2018 se s řetězovými e-maily setkala přes 90 % seniorů, a 20 % z nich přiznává, že se podílí i na jejich dalším přeposílání. [2]

Dezinformace vyskytující se na sociálních sítích, v řetězových e-mailech i na propagandistických stránkách mohou vést k rozdělování společnosti a také ovlivňování demokratických voleb. Od roku 2014, kdy na jaře tohoto roku Ruská federace anektovala ukrajinský poloostrov Krym, nabíralo šíření dezinformací v českém prostoru na síle. [23]

Existují případy, kdy byly dezinformace, jež bývají šířené nejen skrz sociální sítě, citovány ověřenými médii, čímž vzniká jakási domněnka o pravdivosti informace, jež ovšem pravdivá není.

³<https://demagog.cz>

⁴<https://hlidacipes.org>

⁵<https://manipulatori.cz>

⁶<https://eldariele.cesti-elfove.cz>

Jak s dezinformacemi bojovat?

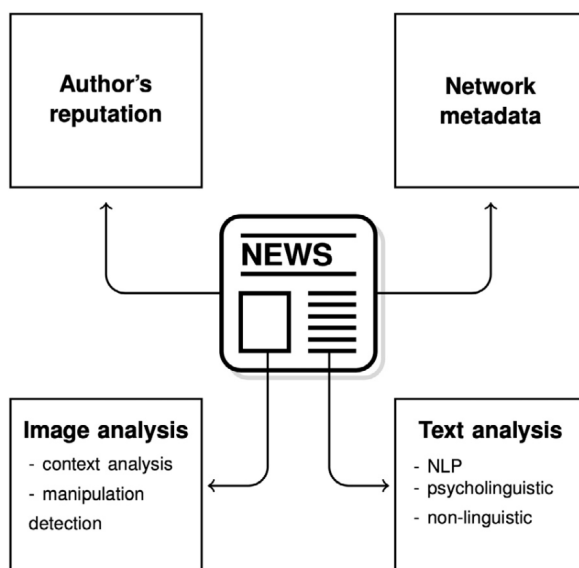
Nejčastěji se bojuje s dezinformacemi manuálně. Novináři ověřují informace a doplňují protipříklady, kterými dezinformace vyvracejí. Je to jeden ze zásadních úkolů žurnalistiky, současně je to ovšem úkol velmi časově náročný. [17] Manuálním ověřováním se v České republice zabývají i výše zmíněné webové stránky, které současně pracují na vyvracení dezinformací.

Od roku 2010 se začalo diskutovat o automatizaci procesu ověřování faktů v kontextu výpočetní žurnalistiky [16]. V následujících letech bylo provedeno několik výzkumů zaměřených na jednotlivé části procesu odhalování dezinformací.

Roste také počet výzkumů, které se zaměřují na různé přístupy a dílčí úkoly nezbytné pro automatizaci ověřování faktů. Jejich cílem je zlepšit kvalitu a rychlost procesu ověřování a pomoci společnosti získávat spolehlivé informace. [17]

2.3 Přístupy k odhalování dezinformací

Detekce dezinformací může zahrnovat ověřování různých typů dat, jako jsou texty, obrázky nebo videa, ale také může zahrnovat hodnocení reputace zdroje/autora, tj. zda je zdroj/autor spolehlivý, a také analýzu kontextu textu nebo obrázku. [12]

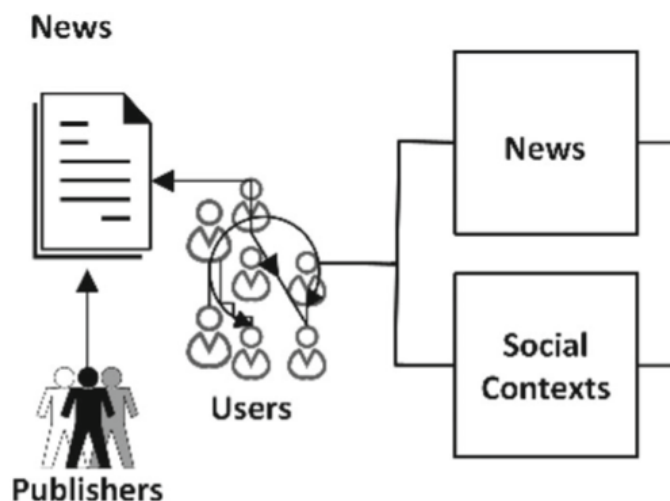


Obrázek 2.3: Různorodost přístupů k odhalování dezinformací, převzato z [12].

Typy dat v internetových člancích

V článku *Fake news detection based on news content and social contexts: a transformer-based approach* [30] jsou zmíněny dvě hlavní entity pro odhalování dezinformačních zpráv. První z nich se týká samotného obsahu článku, který je základem textu a zahrnuje informace jako zdroj, nadpis, autor, čas a datum vydání. Druhá entita se týká sociálního kontextu článku, kdy uživatel může sdílet článek na své sociální síti. Tento kontext zahrnuje metadata, jako je reputace uživatele, který článek sdílí, počet komentářů u článku,

hodnocení článku a další relevantní informace. Tento sociální kontext může být důležitý pro posouzení důvěryhodnosti a vlivu daného článku. Celkově je tedy důležité brát v úvahu jak samotný obsah, tak i sociální kontext, aby bylo možné efektivně identifikovat a odhalovat dezinformace.



Obrázek 2.4: Typy dat v článku, převzato z [30]

Podle článku *Media-Rich Fake News Detection: A Survey* [28] se v internetových článcích také často vyskytují multimédia, která dokážou zachytit pozornost uživatelů, zejména pak obrázky a zvukové záznamy. Dále jsou součástí článků často odkazy na jiný obsah, pomocí kterých autoři mohou podpořit svá tvrzení.

Analýza přístupu k detekci dezinformací

V této části práce budou zmíněny zejména průzkumy popisující tři různé přístupy k automatickému odhalování dezinformací. První z nich vyšel v roce 2021 pod názvem *A Survey on Automated Fact-Checking* [17], druhý vyšel v roce 2018 pod názvem *Media-Rich Fake News Detection: A Survey* [28] a třetí *Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study* [12] v roce 2020.

Autoři průzkumu *Automatizace odhalování dezinformací* [17] rozdělili odhalování dezinformací do tří hlavních oblastí, které sjednocují různé definice.

- **Detekce tvrzení**

Podle průzkumu *A Survey on Automated Fact-Checking* [17] je detekce tvrzení první fází ověřování dezinformací. Detekce často spočívá v tzv. konceptu ověření hodnoty tvrzení a posouzení přínosu tohoto ověření pro veřejnost. Další možností je detekce fámy, což je neověřený příběh nebo prohlášení. Obě kategorie mohou být subjektivní a záleží na kontextu a čase jejich výskytu (nelze například objektivně ověřit fakt, že někdo koupil deset rohlíků).

- **Vyhledávání důkazů (evidence retrieval)**

Základem je vyhledávání informací, které přesahuje pouhé tvrzení a zahrnuje samotný text, obrázky, tabulky, metadata a jiné. Důkazy jsou důležité pro generování vysvětlení ověření faktů, které mají přesvědčit uživatele. *Stance detection* může být považována za instanci *evidence retrieval*, která obvykle předpokládá omezené množství možných důkazů a předpovídá jejich postoj k tvrzení. Otázkou však zůstává, zda jsou tyto dostupné informace důvěryhodné. Nejlepších výsledků dosahují přístupy ověřování faktů, které předpokládají přístup k důvěryhodnému zdroji informací, jakými jsou například encyklopedie nebo výsledky poskytnuté ověřenými vyhledávači.

- **Ověření nároku**

Ověření nároku se dělí na dva základní přístupy, které se navzájem doplňují.

- **Předpověď verdiktu**

V rámci predikce pravdivosti tvrzení se využívá binární klasifikace tvrzení jako pravdivých či nepravdivých, případně klasifikace podpory/neobhajoby pomocí nalezených důkazů. Některé systémy využívají jemnější klasifikaci tvrzení, například pomocí klasifikace nedostatku informací nebo s použitím vícečetných kategorií zastupujících stupně pravdivosti.

- **Vytvoření odůvodnění**

Odůvodňování rozhodnutí je důležitou součástí ověřování faktů v žurnalistice, protože autoři tohoto ověření potřebují přesvědčit čtenáře o svých důkazech. Pouhé označení něčeho za nepravdivé často selhává jako argument v přesvědčování a může vyvolat „zpětnou reakci“. Tato potřeba je ještě větší pro automatizované ověřování faktů, které může využívat takzvané černé skříňky, do níž uživatel nevidí.

Vyhledávání důkazů a ověřování nároků jsou občas řešeny jako jeden úkol pod pojmem faktické ověřování.

Podle článku *Media-Rich Fake News Detection: A Survey* [28] lze nejmodernější přístupy k detekci dezinformací rozdělit do pěti základních skupin podle toho, jaké metody využívají:

1. **Metody založené na lingvistických rysech** jsou založeny na extrakci klíčových lingvistických rysů z textu. Například pomocí metody *N-grams*⁷, identifikace interpunkce v textu, psycholingvistických rysů, POS taggingu anebo analýzy syntaxe textu. Více jsou tyto metody vysvětleny v kapitole 2.3.
2. **Metody založené na modelování klamů** vychází z teorie rétorické struktury (RST) a vektorového prostorového modelování (VSM). Tato metoda spočívá v aplikaci RST, která převede text na hierarchickou sadu rétorických vztahů a následně je využito VSM k identifikaci vztahů této struktury.
3. **Metody založené na shlukování** umožňují porovnávat velké množství dat na základě zvoleného shlukového algoritmu. Například shlukování podobných zpráv na základě využití přístupu *k-nejbližších sousedů* viz kapitola 3.3. Nevýhodou tohoto přístupu může být obtížnost detekce dezinformací v nedávno zveřejněných případech, které mají odlišnou tematiku, jelikož v datové sadě nemusí být dostupné dostatečné množství podobných příkladů dezinformací.

⁷N-grams v kontextu zpracování přirozeného jazyka jsou sekvence sousedících slov o délce n .

4. **Metody založené na prediktivním modelování** umožňují předvídat budoucí hodnoty na základě historických dat. Tyto metody využívají matematické a statistické techniky k vytvoření modelu pro detekci dezinformací. Využívají regresní analýzu, neuronové sítě, rozhodovací stromy a jiné.
5. **Metody založené na obsahu podnětů** se zaměřují na identifikaci obsahu, který může nést dezinformační charakter. Využívají se například lexikální a sémantické analýzy, nebo syntaktické přístupy k analýze textu.

Podle systematického mapování studií provedeného v textu *Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study* [12] lze identifikovat tři hlavní přístupy. Prvním z nich je analýza textu, následovaná analýzou reputace a síťovou analýzou.

Analýza textu (jazykový přístup)

Asi nejjednoznačnějším přístupem pro automatickou detekci dezinformací je zpracování přirozeného jazyka, viz kapitola 3. Základním přístupem pro detekci dezinformací je extrakce vlastností přímo z obsahu článku. Teoreticky nejsnazší je analýza reprezentace dat bez lingvistického kontextu. Například pomocí modelu *bag-of-words* viz 3.2 nebo *N-grams*. Je důležité zmínit, že pouhý počet *N-grams* v každém dokumentu je vázán na jeho délku a pro potřeby rozpoznávání vzorů je nutné jej normalizovat *TF* 3.2, nebo pro sadu dokumentů aplikovat *TF-IDF* 3.2.

Navzdory tomu, že jde o jednoduché a osvědčené řešení, jsou tyto metody úspěšně využívány k řešení problému odhalování dezinformací. O důkazech této problematiky pojednává i text *Credibility Detection in Twitter Using Word N-gram Analysis and Supervised Machine Learning Techniques*. [19]

Reprezentace dat na bázi NLP

Jak již bylo zmíněno, jednou z nejzákladnějších metod v oblasti zpracování přirozeného jazyka je metoda *bag-of-words*, která spočívá v rozpoznávání počtu výskytů konkrétních termínů v textu. Další metodou je *N-grams*, která umožňuje analyzovat nejen jednotlivá slova, ale také celé sekvence, aniž by je rozdělovala na jednotlivé tokeny.

Dalším možným způsobem, jež lze využít při detekci dezinformací, je analýza neobvyklých znaků vyskytujících se v textu. Mezi takové znaky mohou patřit například opakované výskyty otazníků, vykřičníků, emotikonů a dalších symbolů. Tyto znaky jsou často odstraněny během předzpracování textu, jak je popsáno v části 3.1. To však může vést k odstranění prvků, které by mohly pomoci odhalit dezinformaci. Je tedy důležité zvážit, zda a jak tyto neobvyklé znaky zahrnout do analýzy textu.

Psycholingvistické rysy

Podle studií *What Online Reviewer Behaviors Really Matter? Effects of Verbal and Nonverbal Behaviors on Detection of Fake Online Reviews* [34] a *Following Linguistic Footprints: Automatic Deception Detection in Online Communication* [35] se v dezinformačních sděleních často vyskytují dlouhá souvětí, která obsahují časté opakování klíčových slov anebo používají méně formální jazyk. Často se jedná o měřitelné faktory, které lze využít při návrhu systému pro rozpoznávání vzorů.

Dalším užitečným nástrojem může být *POS tagging*, který přiřazuje jednotlivá slova slovním druhům a vrací jejich procentuální podíl v dokumentu. Ty jsou pak doplněny o další informace, jako je označení slova za slangové, hovorový výraz a jiné. Zajímavým trendem v této oblasti může být použití behaviorálních informací, které nejsou přímo obsaženy v textu.

Přístup na základě syntaxe textu

Důležitým aspektem při detekci dezinformací je i syntaxe textu, jelikož se zabývá vztahy mezi slovy v textu a jejich gramatickou strukturou. Pokud je tedy text napsán nesprávným nebo matoucím způsobem, tak se může jednat o známku dezinformace.

Analýza reputace (znalostní přístup)

Z pohledu sociologie je reputace jedince důležitým mechanismem sociální kontroly. Tento koncept má rovněž klíčový význam v politice, vzdělávání a dalších oblastech. Pro měření reputace se používá koncept reputačního skóre, které je číselně vyjádřeno. Skóre reputace lze vypočítat na základě lokálních nebo globálních metrik důvěryhodnosti. V oblasti zpravodajství existují dva typy zdrojů reputace: ty, které vycházejí z recenzí, a ty, které jsou založeny na IP adrese nebo doméně, na které se článek vyskytuje.

Existuje několik klíčových vlastností článků, které mohou odlišit důvěryhodný článek od falešného. Častou charakteristikou falešných článků je anonymita autora článku nebo popularita domény, která se odráží v návštěvnosti webové stránky.

Analýza reputace je systémový proces, který automaticky ověřuje sadu důvěryhodných databází blacklistů a provádí křížovou kontrolu.

Síťová analýza (síťový přístup)

Tato analýza se odkazuje na teorii sítí, která studuje grafy, jež představují buď symetrické nebo asymetrické vztahy mezi jednotlivými objekty.

Zkoumá tedy dva druhy sítí: homogenní a heterogenní sítě. Homogenní sítě obsahují jeden druh uzlů a hran, oproti tomu heterogenní sítě mají více uzlů a hran a jejich výhodou je, že dokážou reprezentovat vztahy z různých pozic.

Detekce dezinformací pomocí síťové analýzy se provádí pro vyhodnocení pravdivostní hodnoty zprávy.

Příkladem síťové analýzy detekce falešných zpráv může být například článek *Exploiting Tri-Relationship for Fake News Detection* [32], kde je navržen rámec pro detekci dezinformací využívající model tří vztahů – propojení mezi články, vydavateli a šířiteli. Tento rámec se zaměřuje na vkládání a reprezentaci entit, modelování vztahů mezi nimi a polosupervizované učení. Podle publikace je tento model vhodný i pro detekci nově šířených dezinformačních zpráv.

Kapitola 3

Zpracování přirozeného jazyka

Využitím zpracování přirozeného jazyka můžeme díky souboru technik a metod umožnit počítačům zpracovávat přirozený jazyk.

Strojové učení

V této části jsou rozepsány metody strojového učení.

Strojové učení je podoblastí umělé inteligence, zabývající se algoritmy a technikami, které umožňují počítačovému systému „učit se“. [25] Strojové učení dokáže na základě předešlých zkušeností klasifikovat data a agregovat je. Značně se prolíná s oblastmi statistiky a dobývání znalostí. [25]

Tři základní algoritmy strojového učení:

- **Učení s učitelem**

Jedná se o nejčastěji používanou formu strojového učení. Pro tento algoritmus je důležité mít předem vytvořenou velkou sadu trénovacích dat, která obsahuje základní rozdělení prvků.

Systém během tréninku prochází jednotlivé případy, které jsou uvedeny v trénovací sadě a vytvoří výstup ve formě vektoru hodnocení pro každou kategorii. Cílem je mít co nejvyšší hodnotící skóre pro každou kategorii, a proto je nutné trénování daného systému.

Nevýhodou tohoto algoritmu je již právě zmíněná trénovací sada, která se ve velké většině případů připravuje ručně. [25]

- **Učení bez učitele**

Na rozdíl od učení s učitelem nemá tento algoritmus provázána trénovací data s cílovými proměnnými, nemá tedy k dispozici žádné kritérium správnosti hledané transformace vstupních dat. Na základě principu shlukování hledá sobě podobné elementy a s ohledem na vzorky vstupních dat provádí jejich třídění do skupin, nicméně jedná bez možnosti posouzení správnosti zatřídění. Do adaptace nevstupuje žádný arbitr, celé učení je založeno pouze na informacích obsažených ve vstupních datech. Tento algoritmus si lze tedy představit jako kompresi vstupních dat či jejich vyhlazení. [1]

- **Učení posilováním - zpětnovazební učení**

Jedná se o typ strojového učení, kdy se agent nacházející se v neznámém prostředí rozhoduje, jaké akce zvolí, aby z nich maximalizoval svůj užitek. Pro usnadnění rozhodování je mu v daný okamžik poskytnuto jen určité penzum informací, ostatní nebere v úvahu. Od učení s učitelem se odlišuje tím, že je agentovi při jeho rozhodování poskytnuta jen částečná zpětná vazba. V tomto algoritmu existují jen dva typy ohodnocení, jimiž jsou odměna a trest. [13]

3.1 Předzpracování textu

Předzpracování textu je součástí oblasti zpracování přirozeného jazyka. Zakládá se na souboru technik a metod, které optimalizují následné využití hodnotících funkcí.

- **Tokenizace**

Tokenizace je proces, při němž se řetězec rozdělí na skupinu podřetězců, které reprezentují slova, číslice, speciální znaky nebo jiné významové elementy, ty pak nazýváme tokeny. Věta „*Dnes je krásný slunečný den.*“ bude tedy rozdělena na tuto posloupnost tokenů:

["Dnes", "je", "krásný", "slunečný", "den", "."]

Hlavním účelem tokenizace je identifikace významově důležitých slov. Tokenizace je v jednotlivých jazycích specifická. Například při rozčlenění textu v češtině by se daly použít bílé znaky, avšak v čínštině se bílé znaky k rozdělení slov nepoužívají. [18]

Dalším problémem při tokenizaci je rozdělení slov, která k sobě neoddělitelně patří a jejich rozdělení může vést ke změně významu.

- **Odstranění stop-slov**

Stop-slova jsou slova, která se v textu vyskytují nejčastěji, avšak nenesou žádnou důležitou informaci. Nejčastěji používaná slova lze dohledat ve frekvenčních seznamech daného jazyka a jedná se zejména o funkční slova – předložky, spojky a zájmena. V českém jazyce jsou *stop-slova* typicky například „a“, „být“, „v“ atd.

- **Převod na malá písmena**

Jedná se o jednoduchou metodu, kterou lze lehce implementovat. Zaručuje jednotnost textu i v případě, že je využito hůlkové písmo.

- **Stematizace**

Stematizace je proces redukování variací slov do základního tvaru s cílem zredukování počtu pojmů. Tento proces vede ke zkrácení tvaru slova na kořen pomocí odstranění sufixu a prefixu, nebo pomocí nalezení kořene ve slovníku. Například slova „*prezentační*“, „*prezentující*“ a „*prezentovaný*“ lze převést na základní tvar „*prezentace*“.

Při stematizaci dochází ke dvěma základním chybám.

- **Over stemming**

Proces, při němž jsou dvě slova jiného významu stematizací konvergovány ke stejnému základnímu tvaru.

– **Under stemming**

Proces, při němž nejsou dvě slova stejného významu převedena na stejný základní tvar. Takový proces je také označován jako falešně negativní.

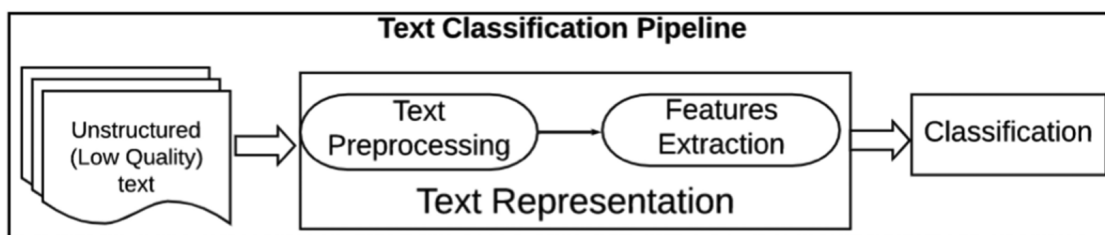
• **Lemmatizace**

Lemmatizace je podobný proces jako stematizace, ale liší se v logice převedení slova. Základem lemmatizace je extrahování slov a číslic do jednotného tvaru. Výsledná množina slov se skládá z lemmat.

Existují různé nástroje pro lemmatizaci slov a následné získání lemmat. Český jazyk je ovšem mnohoznačný a lemmatizace je v takovém případě náročná, protože kontext slova může odkazovat k několika základním tvarům.

3.2 Metody extrakce příznaků

Dalším důležitým krokem zpracování přirozeného jazyka je převedení slov na počítačově srozumitelnou posloupnost. Tyto metody řeší různým způsobem semantické a syntaktické vazby mezi jednotlivými slovy. V následujících podkapitolách budu vycházet z článku *A Comprehensive Survey on Word Representation Models: From Classical to State-of-the-Art Word Representation Language Models*. [27]



Obrázek 3.1: Proces zpracování textu.

Klasické modely

Klasické modely transformují text do vektoru, který obsahuje počet slov vyskytujících se v dokumentu. Můžeme je rozdělit na dva základní přístupy:

Kategorická reprezentace slov

Jedná se o nejsnazší způsob jak reprezentovat text. Tato metoda využívá symbolické vyjádření buď "1", nebo "0". K této reprezentaci lze využít dvou známých modelů:

• **One-hot-encoding**

Každé slovo ve slovníku je reprezentováno binární proměnnou, což znamená, že každé slovo je složeno z nul a jedniček. Index odpovídajícího slova je označen "1", zatímco ostatní jsou označeny "0".

• **Bag-of-Words**

Bag-of-Words (BoW) je model, který je rozšířením modelu *One-hot-encoding*. Jedná se o zpracování textu, jež reprezentuje výskyt slov za pomoci dvou faktorů:



Obrázek 3.2: Bag-of-Words Model.

- slovníku známých slov,
- míry přítomnosti známých slov.

Veškeré informace o pořadí a struktuře textu v dokumentu jsou zahozeny, proto se model nazývá „*Bag of Words*“. Jednoduchost tohoto modelu tedy závisí na definování slovníku známých slov a míře přítomnosti známých slov.

Vážená reprezentace slov

Metody vážené reprezentace slov jsou podobné kategorické reprezentaci slov, avšak odlišují se numerickými reprezentacemi založenými na frekvenci slov.

- **TF**

Pro tuto metodu je zásadní četnost slova v dokumentu. Celkový počet výskytů slova je roven podílu četnosti výskytu slova v dokumentu a celkového počtu slov. Záleží tedy na délce dokumentu.

- **TF-IDF**

V této metodě se využívá *TF metoda* současně s *IDF*. Nezáleží tedy pouze na četnosti slov v dokumentu. *TF* představuje počet výskytů daného slova v dokumentu a *IDF* představuje měřítko obecné důležitosti slova. Potlačuje tedy efekt častých slov v dokumentu, který je tímto způsobem potlačen.

Matematická definice metod:

$$TF(t) = \frac{\text{number of term } t \text{ occurrence in a document}}{\text{total number of terms}}$$

$$IDF(t) = \ln\left(\frac{\text{total number of documents}}{\text{total number of documents with term } t}\right)$$

Učení reprezentací (Representation Learning)

Na rozdíl od klasicických modelů se tyto modely učí reprezentaci dat přímo ze vstupních dat bez potřeby ručně navrhovat definované funkce. Zároveň umožňují vysokou míru generalizace, díky níž jsou schopny provádět úlohy na nových datech.

Mapování slov

Mapování slov, v angličtině též nazýváno *Word embedding*, je proces zpracování slov z textové reprezentace do číselné podoby – nízkorozměrového vektoru. Ten popisuje jak syntaktický, tak sémantický význam slova. Jedná se o naučenou reprezentaci textu, v níž mají slova stejného významu podobnou reprezentaci. [9] Vektorová reprezentace umožňuje celou řadu matematických operací, které se dají využít pro porovnávání, tvorbu modelu či pro klasifikaci textu.

Možnou nevýhodou mapování slov je existence pouze jednoho reprezentujícího vektoru ke slovu, které může nést více významů.

Existují dvě základní kategorie metod pro ohodnocování slov: vnitřní a vnější hodnocení.

Mapování slov lze řešit pomocí těchto dostupných algoritmů:

- **Word2Vec**

Word2Vec je statistická metoda pro efektivní učení samostatných vektorových reprezentací slov z textového korpusu. Jedná se o bezkontextový model, který byl publikován v roce 2013 výzkumníkem Tomášem Mikolovem. [26]

Byly představeny dvě modelové architektury, které mohou být součástí *Word2Vec* metody pro učení mapování slov, jež se snaží minimalizovat výpočetní složitost. [26]

- **Continuous Bag-of-Words Model (CBOW Model)**
- **Continuous Skip-gram Model**

V *modelu CBOW* se učí predikovat aktuální slovo v závislosti na jeho kontextu, tedy na základě okolních slov. Naopak v *modelu Continuous Skip-gram* se učí predikovat okolní slova na základě aktuálního slova.

- **GloVe**

Algoritmus GloVe, celým názvem *Globální vektory pro reprezentaci slov*, je rozšířením pro metodu *Word2Vec* pro efektivní učení vektorů slov.

Výhoda této metody spočívá v tom, že nepoužívá okno pro definování lokálního kontextu, *GloVe* namísto toho vytvoří explicitní matici slova, či matici společného významu pomocí statistik v celém textovém korpusu. Výsledkem je model, který může vést k obecně lepšímu mapování slov. [29]

- **FastText**

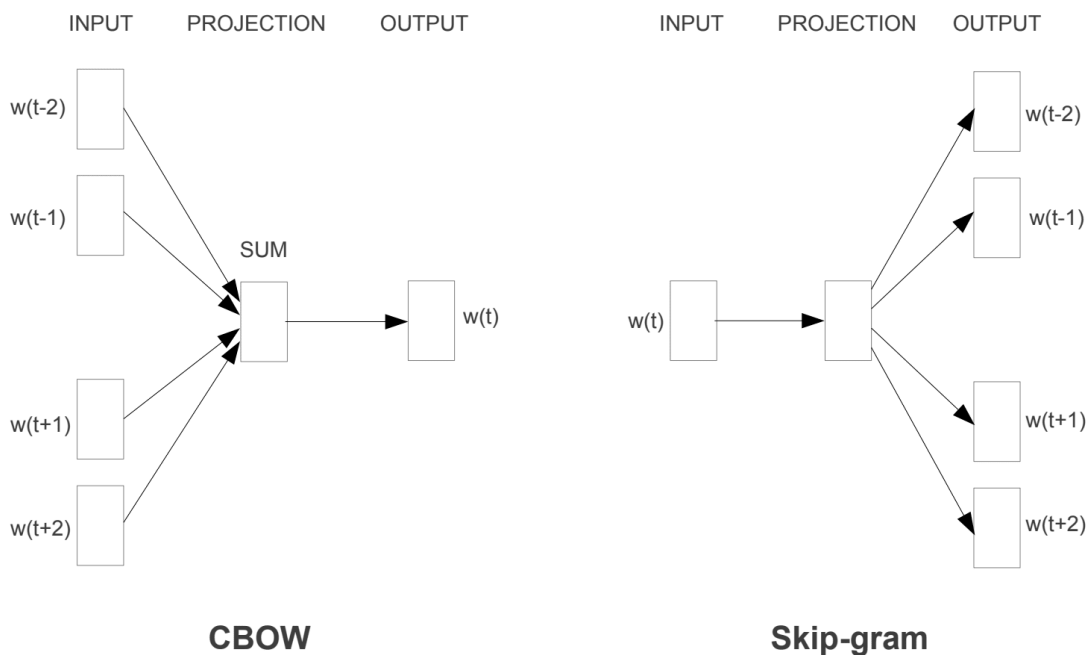
FastText vychází z modelu *Word2Vec* 3.2, ale navíc zahrnuje i informace o podvýrazech slov.

Kontextuální reprezentace slov

Kontextuální reprezentace slov se snaží zachytit význam slov v kontextu, tedy zohledňovat slovo v jeho okolí, aby byla získána co nejpřesnější reprezentace.

- **Context2Vec**

Tento model je založen na *Word2Vec CBOW modelu*, ale nahrazuje průměrný výskyt slova fixním oknem s lepší a výkonnější rekurentní neuronovou sítí.



Obrázek 3.3: Word2Vec modely, převzato z [26].

- **CoVE**

Princip *CoVE* spočívá v získání vektorové reprezentace pro každé slovo věty a následné kombinaci těchto reprezentací s reprezentací celé věty pomocí neuronové sítě.

- **ELMo**

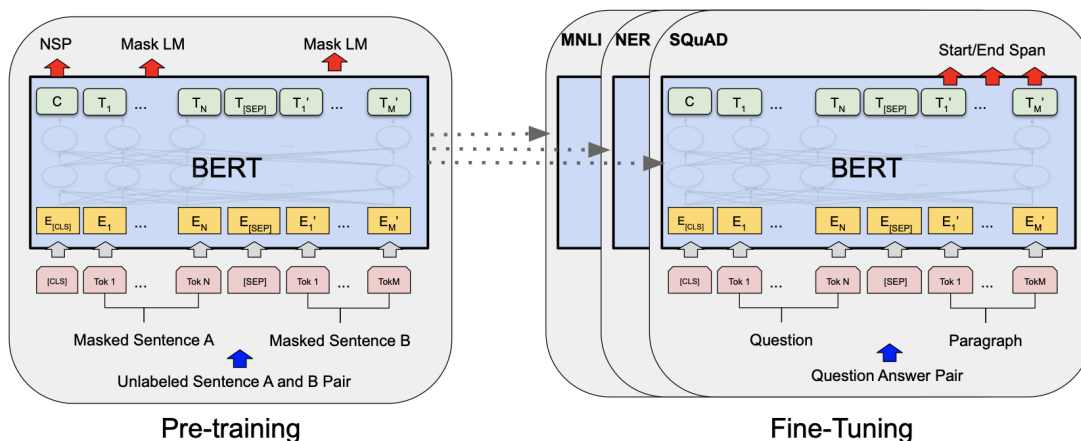
Jedná se o hlubokou kontextualizovanou reprezentaci slov. Na rozdíl od tradičních reprezentací, které zobrazují každé slovo jako fixní vektor bez ohledu na jeho kontext, *ELMo* zahrnuje okolní slova a vytváří dynamickou, kontextově specifickou reprezentaci. [31]

- **BERT**

BERT (*Bidirectional Encoder Representations from Transformers*) je jednou z nejnovějších inovací v oblasti strojového učení pro zpracování přirozeného jazyka. Patří do kategorie předtrénovaných jazykových modelů založených na architektuře *Transformer*.

Obsahuje dvě základní fáze: předtrénink a doladění. Během předtréninku je model trénován na neoznačených datech s různými předtrénovacími úlohami. Při doladění je model *BERT* nejprve inicializován s předtrénovanými parametry a poté jsou všechny parametry doladěny pomocí označených dat z cílových úloh.

Jedním z charakteristických rysů modelu *BERT* je jeho jednotná architektura napříč různými úlohami. Existuje minimální rozdíl mezi předtrénovanou architekturou a finální architekturou pro konkrétní úlohy. [14]



Obrázek 3.4: BERT, převzato z [14].

3.3 Klasifikační techniky

Jedná se o metody, které slouží k přiřazení vstupních dat do určitých tříd nebo kategorií. Níže jsou zmíněné některé z nich.

- **Metoda rozhodovacích stromů**

Cílem je vytvořit model, který předpovídá hodnotu cílové proměnné na základě několika vstupních proměnných. Každý vnitřní uzel odpovídá jedné ze vstupních proměnných; existují hrany vedoucí k potomkům pro každou možnou hodnotu této vstupní proměnné. [22]

- **K-nejbližších sousedů**

Klasifikační technika *k-nejbližších sousedů* (*k-nn*) hledá řešení podobných problémů, které již předtím vyřešila. Spočítá počet případů pro každou třídu a nový případ přiřadí stejné třídě, k níž patří většina jeho sousedů. [22]

- **Neuronové sítě**

Neuronové sítě jsou nelineární statistické nástroje pro modelování dat. Mohou být použity k modelování složitých vztahů mezi vstupy a výstupy nebo k hledání vzorců v datech. Dochází k manipulaci a křížení dat, což uživatelům umožňuje rozhodovat se na základě většího množství informací. [22]

3.4 Evaluační metriky

Evaluační metriky jsou určeny k optimalizaci a hodnocení modelů strojového učení. Při hodnocení textových klasifikačních modelů se často používají metriky jako *accuracy*, *precision*, *recall* a *F1 score*. Tyto metriky slouží k měření úspěšnosti modelu a k určení jeho schopnosti klasifikovat vstupní data do správných tříd.

Confusion matrix je unikátní tabulka, která prezentuje efektivnost daného klasifikačního algoritmu. Obsahuje informace o počtu správně a nesprávně klasifikovaných instancí v jednotlivých třídách. Je popsána v tabulce 3.1.

		Aktuální třída	
		Pozitivní	Negativní
Predikovaná třída	Pozitivní	Pravdivě Pozitivní (TP)	Falešně Negativní (FN)
	Negativní	Falešně Pozitivní (FP)	Pravdivě Negativní (TN)

Tabulka 3.1: Confusion matrix.

- **Pravdivě Pozitivní (TP):** pravdivě označeny za pozitivní výsledky,
- **Pravdivě Negativní (TN):** pravdivě označeny za negativní výsledky,
- **Falešně Pozitivní (FP):** falešně označeny za pozitivní výsledky,
- **Falešně Negativní (FN):** falešně označeny za negativní výsledky.

Jednotlivé evaluační metriky:

- **Accuracy**
Poměr počtu správných predikcí ke všem predikcím.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

- **Precision**
Poměr pravdivě pozitivních predikcí ke všem pozitivně predikovaným predikcím.

$$Precision = \frac{TP}{TP + FP}$$

- **Recall**
Poměr pravdivě pozitivních predikcí ke všem predikcím.

$$Recall = \frac{TP}{TP + FN}$$

- **F1 score**
Vážený průměr jak odvolání, tak preciznosti.

$$F1score = 2 * \frac{Recall * Precision}{Recall + Precision}$$

Kapitola 4

Datové sady

V této kapitole jsou detailně vysvětleny přístupy k vytváření datových sad pro odhalování dezinformací a jsou zde také zmíněny některé existující datové sady, které jsou v tomto kontextu používány. Dále jsou představeny metody, jak lze datové sady pro tyto úlohy vytvořit.

Datové sady obecně

Datová sada je (velká) kolekce dat, od níž se čeká konzistentnost a vyváženost. Pro nejlepší výsledky systému je důležité, aby byla data reprezentativní a velikost datasetu dostatečná pro dané použití.

Přístup k vytváření datové sady

Výzkum *Survey on Automated Fact-Checking* [17] rozděluje datové sady do tří skupin v závislosti na přístupu k odhalování dezinformací. Dále jsou detailně rozepsány jednotlivé sady.

Vstupy

Do procesu detekce tvrzení a ověřování faktů vstupují obvykle textová tvrzení získaná z různých zdrojů, jimiž jsou sociální sítě, debaty, webové stránky a další. Některé datové sady jsou vytvořeny ručním označováním tvrzení, zatímco jiné jsou získány pomocí strojového učení. Tyto vstupy jsou poté analyzovány pomocí různých metod a algoritmů, aby bylo možné určit, zda jsou tvrzení pravdivá, nepravdivá, nebo je nemožné je ověřit. Mezi různými typy vstupů jsou také věty obsahující citace z *Wikipedie*, otázky z diskuzních fór a další.

Důkazy

V rámci ověřování informací se často používají metadata jako zdroje, datum publikování a další informace. Tato metadata mohou být užitečná při odhalování dezinformací, ale sama o sobě nestačí jako důkaz pro potvrzení informace. Články na *Wikipedii* nebo akademické články jsou častými typy důkazů, u kterých se předpokládá důvěryhodnost pro celou škálu znalostí, to ale nemusí platit vždy. Strukturovaná data jsou dalším typem důkazů, které umožňují získání znalostí stručným a flexibilním způsobem.

Ohodnocení s vysvětlením

Dříve se běžně používalo binární označení - pravda/nepravda, aby se určilo, zda je tvrzení pravdivé či nikoliv. Avšak v současné době se stále více upřednostňuje používání nebinárního označení pro přesnější určení stupně pravdivosti informace. Tento přístup může zahrnovat různé stupně nejistoty, například „pravděpodobně pravda“, „pravděpodobně nepravda“ nebo „nevíme“. Dále se často přidávají vysvětlení nebo protitvrzení k danému tvrzení, což může pomoci k většímu porozumění situace a zvýšení objektivitu ověřování informací.

Většina existujících datasetů nezahrnuje vysvětlení ohodnocení přímo od novinářů.

Používání článků ověřujících fakta není vždy možné, protože v době vyvozování závěrů nemusí být k dispozici. V takových případech systém není schopen poskytnout důkladné vysvětlení na základě získaných důkazů.

Příprava datové sady pro trénování modelu

Pro trénování modelu *BERT*, jak je popsáno v kapitole 3.2, je podle *Automatic Fake News Detection with Pre-trained Transformer Models* [31] nutné nejprve připravit datovou sadu pro dosažení lepších výsledků.

Prvním krokem je odstranění všech titulků, které mají méně než 20 znaků nebo více než 120 znaků. Titulky s méně než 20 znaky jsou většinou nepodstatné informace pro trénování modelu – například název webové stránky. Naopak titulky delší než 120 znaků bývají chybové hlášky, které vznikly při stahování dat z internetu.

Podobně je nutné postupovat i s texty článků. Texty kratší než 1000 znaků nebo naopak delší než 10 000 znaků jsou vymazány.

Dalším krokem může být odstranění některých HTML řetězců, které se vyskytují v některých textech z důvodu, že byly automaticky stahovány pomocí crawleru.

Poslední částí přípravy textu pro trénování modelu je odstranění všech znaků, které nejsou ASCII znaky, a vymazání značek nových řádků.

Existující datové sady pro odhalování dezinformací

V průběhu let vzniklo již mnoho různých datových sad, které lze využít pro odhalování dezinformací. Problém je však v tom, že většina z nich je v anglickém jazyce, což není vhodně k trénování modelů pro odhalování českojazyčných dezinformací. Dalším problémem, jež tyto datové sady provází, je zastaralost dezinformačních témat. Jak bylo zmíněno výše, pro mnoho přístupů k dohalování dezinformací je důležitá aktualizovaná datová sada, kde se vyskytují i aktuální témata.

Dataset pro trénování detekce dezinformací

V českém prostředí existuje databáze řetězových e-mailů, která skýtá přes 10000 jednotlivých textů. Obdobná databáze pro české dezinformační texty vyskytující se převážně na propagandistických stránkách však není dostupná a nelze tedy žádnou takovou databázi využít. Z toho důvodu bylo nutné ručně zpracovat datovou sadu pro následnou možnost trénování modelu.

Kapitola 5

Návrh systému a jeho limity

Systém je navržen s ohledem na stávající možnosti ověřování dezinformací a na možnosti využití již existujících prostředků pro tuto úlohu. V podkapitole 5.1 jsou představeny čtyři hlavní body, které vyplývají ze zvoleného přístupu k řešení. Dále v kapitole 5.2 je popsáno schéma systému, na němž jsou podrobněji znázorněny záměry jednotlivých částí systému, jejich úlohy, vstupy a výstupy. Nakonec v kapitole 5.3 jsou popsány limity předkládané práce.

Systém je primárně zaměřen na stahování článků z propagandistických webových stránek, následné ukládání těchto dat a provedení jejich analýzy. Toto řešení bylo zvoleno z několika důvodů, které vyplynuly v průběhu výzkumu.

Prvním z důvodů je nedostatečná dostupnost zpracovaných dat z propagandistických webů v České republice pro další využití, například k trénování modelů, jak je popsáno v kapitole 3. Dalším důvodem je fakt, že propagandistické weby mohou mít významný dopad na společenské a politické dění, a tudíž je nutné jim věnovat pozornost a analyzovat jejich obsah. Často se u těchto typů webů stává, že mažou svůj obsah anebo jej pozměňují. Jedná se zejména o problematický obsah, který nese dezinformace. Tento systém tudíž slouží i pro archivaci již smazaných článků a uchovává i další metadata, která se dají dále zpracovávat.

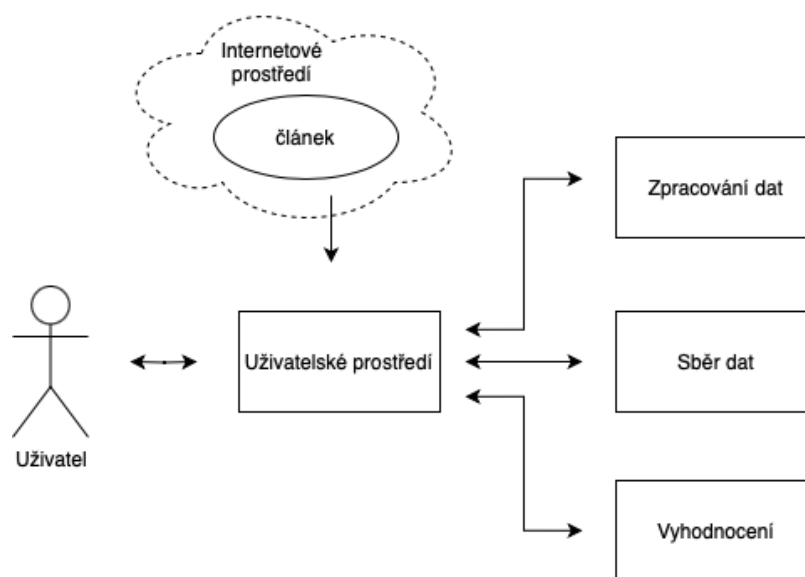
Systém je především navržen pro uživatele, kteří se věnují odhalování dezinformací na propagandistických webových stránkách a využívají při tom konvenční metody a nástroje, jako je například ruční ověřování faktů a další. Systém umožňuje uživatelům propojovat jednotlivé informace a zdroje v člancích a poskytuje také analýzu, která pomáhá odhalování dezinformací v textu.

Velkou výhodou systému je, že dokáže pracovat s velkým množstvím dat a zdrojů a dokáže tyto zdroje propojovat a porovnávat mezi sebou. To umožňuje uživatelům získat komplexní přehled o tématu a rychleji identifikovat dezinformace a manipulace v textu.

Cílem navrženého systému je poskytnout uživatelům komplexní a účinný nástroj pro odhalování dezinformací a zlepšení ověřování informací a zejména pro vytváření velké datové sady pro trénování modelů, viz 3.4, které se dají v budoucnu využít pro rozšíření systému a zlepšení jeho schopností.

S ohledem na etický rámec této práce jsou používány pouze články z veřejně přístupných zdrojů. Pro potřeby této práce sbírá autor rovněž data o uživatelích z důvodu častého odstranění článků na propagandistických webových stránkách jeho správci. Tyto údaje však

nejsou dále zpracovávány a je postupováno s ohledem na zákon č. 101/2000 Sb., o ochraně osobních údajů neboli GDPR¹.



Obrázek 5.1: Schéma propojení systému a uživatele.

5.1 Zvolený přístup k problému

Jak již bylo zmíněno výše v 2.3, přístupů k odhalování dezinformací může být více a je dobré je kombinovat pro lepší výsledky. V této práci bylo zvoleno základní paradigma pro odhalování dezinformací skládající se ze 4 hlavních bodů:

- **Sběr dat**

Sběr dat je klíčovým prvkem v odhalování dezinformací. Systém umožňuje stahovat články z propagandistických webových stránek viz C, v určených periodách. To znamená, že může být použit pro průběžné monitorování webových stránek a jejich obsahu.

- **Analýza dat**

Analýza dat je důležitá pro získání relevantních informací ze stahovaných článků. Systém analyzuje a extrahuje informace z dat pro odhalení dezinformací v textu. To zahrnuje identifikaci klíčových slov, frází a vztahů mezi nimi.

- **Detekce dezinformací**

Detekce dezinformací je primárním cílem této práce a je jí dosaženo pomocí příznaků, které se ve člancích vyskytují, a využívá také strojového učení. Detekce dezinformací může být obtížná, protože dezinformace mohou být velmi sofistikované a obtížné na odhalení.

¹GDPR (General Data Protection Regulation) je anglická zkratka Obecného nařízení, která představuje nový právní rámec ochrany osobních údajů v evropském prostoru.

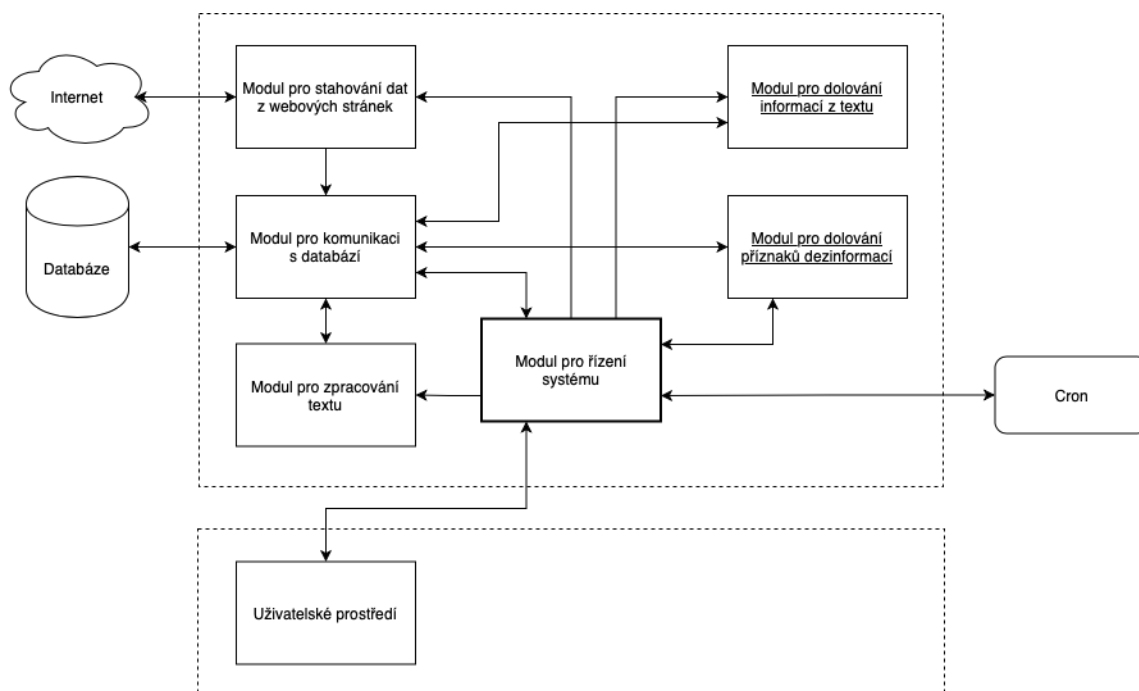
- **Vyhodnocení**

Vyhodnocení je klíčové pro zajištění toho, že uživatelé mohou snadno interpretovat výsledky analýzy a detekce dezinformací. Systém uživateli srozumitelně poskytne výsledky své analýzy a detekce dezinformací pomocí webové aplikace. To umožňuje uživatelům okamžitě reagovat na nalezené dezinformace a přispět tak k prevenci šíření nepravdivých informací.

5.2 Návrh řešení

V této podkapitole jsou detailněji rozepsány jednotlivé části systému, které byly představeny v předchozí kapitole pomocí schématu 5.2. V první části, konkrétně v sekci 5.2, je popsán návrh a struktura databáze, která slouží jako úložiště pro data, s nimiž systém pracuje.

Dále se podkapitola věnuje návrhům jednotlivých modulů, které zpracovávají text a tvoří jádro celého systému. Nakonec je popsán základní princip systému a jeho automatické spuštění, které zajišťuje pravidelný sběr dat a jejich následnou analýzu a detekci dezinformací. Celkově se tato podkapitola zaměřuje na popis jednotlivých částí systému a jejich funkčnost.



Obrázek 5.2: Schéma systému a jeho jednotlivé závislosti.

Databáze

Databáze slouží k ukládání seznamu dezinformačních webových stránek, jejich obsahu a dalších informací, které jsou získány z dat.

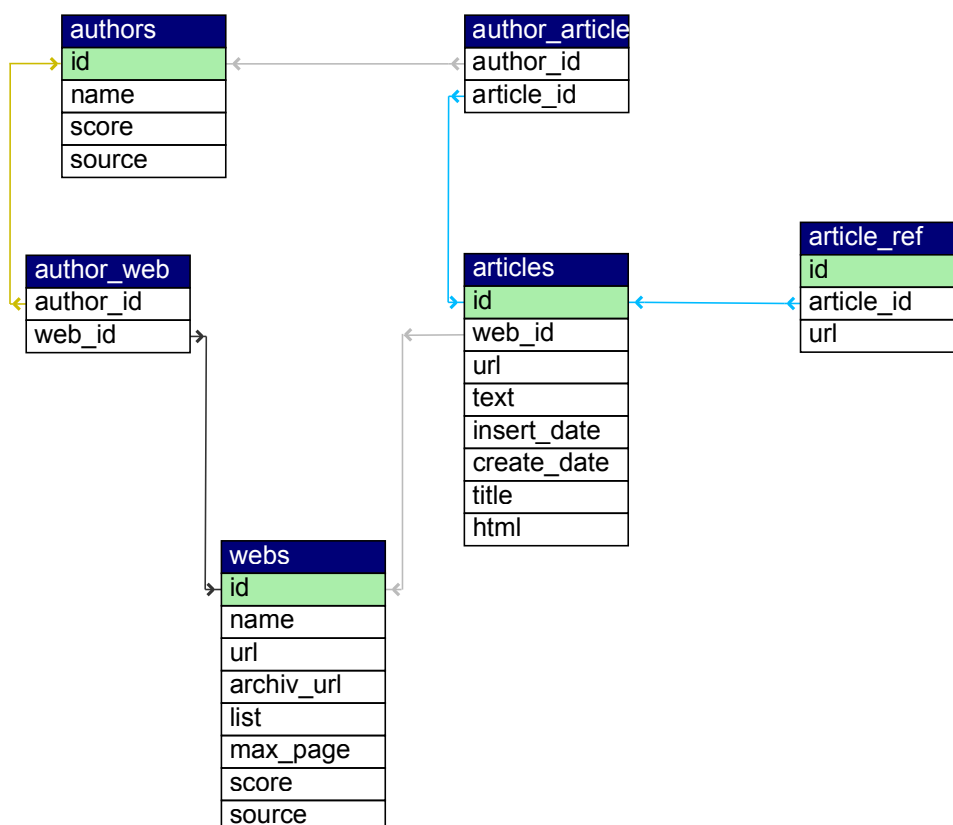
Tabulka **webs** obsahuje seznam dezinformačních webových stránek a potřebné informace pro získávání článků z těchto stránek. Každý záznam v tabulce **webs** je propojen pomocí cizího klíče se záznamem v tabulce **articles**, který je stažen z odpovídající webové stránky.

Tabulka `articles` obsahuje informace o stažených článcích, jako jsou nadpis, obsah, datum publikace a další atributy. Tato tabulka je také propojena pomocí relační tabulky `author_article` s tabulkou `authors`, která nese záznamy o jednotlivých autorech, kteří články publikují.

Dále se v databázi nachází tabulka `article_ref`, do které jsou ukládány odkazy a citace, které byly nalezeny v daném článku. Tato tabulka je propojena s tabulkou `articles` pomocí cizího klíče, aby bylo možné identifikovat, z kterého článku odkazy a citace pocházejí.

Tabulka `authors` je propojena s tabulkou `webs` pomocí relační tabulky `author_web`. V této tabulce je vždy uloženo id autora a id webové stránky.

Tento databázový návrh umožňuje ukládání a organizaci dat v systému pro odhalování dezinformací, což je klíčové pro jeho správné fungování a využití.



Obrázek 5.3: Schéma databáze.

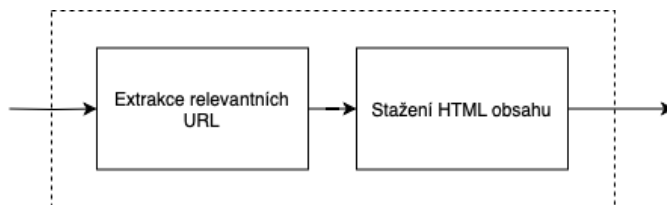
Modul pro stahování dat z webových stránek

Pro pochopení fungování modulu pro stahování dat z webových stránek je nutné vědět, že relevantní odkazy jsou odkazy na články, které se nacházejí na dezinformačních webových stránkách. Tyto odkazy jsou získávány z archivu webové stránky, což je seznam odkazů na starší články, které se již nenacházejí na úvodní stránce dané webové stránky.

Po získání relevantních odkazů je druhá část modulu určena pro stažení veškerého obsahu ze stránky, která se nachází pod danou url adresou. Tento obsah zahrnuje nejen sa-

motný textový obsah článku, ale také metadata, jako jsou název článku, datum publikace, autor a další informace, které jsou důležité pro následné zpracování a analýzu.

Po stažení obsahu je tento obsah vkládán do databáze, kde se ukládá společně s informacemi o zdrojové webové stránce a dalšími relevantními informacemi. Tento proces umožňuje uživatelům systému rychle a efektivně získat relevantní informace o dezinformačních webových stránkách a zdrojích.



Obrázek 5.4: Schéma získávání html kódu.

Modul pro zpracování html

Tento modul slouží pro extrakci důležitých informací z webových stránek k následné analýze. Konkrétně se jedná o extrakci textu, autora článku, data publikace a titulku článku.

Jelikož každá webová stránka je postavena jiným způsobem, vzniká problém s přesnou extrakcí informací. Modul je proto navržen tak, aby dokázal co nejpřesněji extrahovat požadované informace ze všech podobných webových stránek.

Obsah webových stránek

Webová stránka je v kontextu internetu označována jako soubor dat zobrazovaný uživateli webovým prohlížečem. Tyto soubory dat jsou obvykle vytvářeny pomocí HTML kódu, který popisuje strukturu a obsah stránky. [20] Pro získání dat z webových stránek stačí právě tento HTML dokument.

Stahování dat z webových stránek provází mnoho problémů, jako jsou nejednotné kódování stránek, dynamický obsah generovaný na straně klienta, nejednoznačné identifikátory elementů a další.

Struktura webových stránek

Webové stránky jsou tvořeny pomocí HTML kódu, který slouží jako základní stavební kámen celé stránky. Pomocí *CSS (Cascading Style Sheets)* lze dále upravovat jednotlivé aspekty vzhledu stránky, například barvy, fonty, velikosti textů a umístění prvků. *JavaScript* je poté využíván pro tvorbu interaktivního obsahu a umožňuje například animaci obsahu nebo interakci uživatele s webovou stránkou.

Obecně se webová stránka skládá z následujících částí:

- **Hlavička stránky**

Hlavička se nachází v horní části stránky a často obsahuje navigační menu, logo a další prvky, které slouží k lepší orientaci uživatele na webových stránkách.

- **Obsah stránky**

Obsah představuje hlavní informace, které se na webu nacházejí a může zahrnovat text, obrázky, videa a další prvky.

- **Postranní panel**

Postranní panel se často využívá pro umístění reklam a dalších navigačních komponent.

- **Patička stránky**

Ve spodní části stránky se nachází patička stránky, která obsahuje základní informace o webové stránce, jako jsou kontaktní údaje, copyright a další.

HTML

HTML neboli hypertextový značkovací jazyk slouží pro vytváření HTML dokumentů, které vytváří základ webových stránek. Aktuálně se využívá verze *HTML5*. Díky možnosti znorovávání značek dokument tvoří stromovou strukturu a je tak přehlednější.

```
1      <!DOCTYPE html>
2      <html>
3          <head>
4              <title>Nazev stranky</title>
5          </head>
6          <body>
7              <h1>Nadpis</h1>
8              <p>Tady je odstavec.</p>
9          </body>
10     </html>
```

Každý HTML dokument se skládá z několika základních komponent, které pak obsahují detailnější části webové stránky. Pro získávání článků tj. čistého textu je důležitá část `<body>`, v níž se nachází tzv. kontejnery označované značkou `<div>` a v těch jsou zanořeny dané články.

Kódování stránky

Různé kódování webových stránek může být problémem při stahování jejich obsahu. Nejčastěji používaným zápisem pro české webové stránky je UTF-8, Windows-1250 a ISO 8859-2.

UTF-8 je nejčastějším zápisem znakové sady Unicode, která se čím dál tím víc rozšiřuje. Unicode je určený pro všechny světové jazyky najednou na rozdíl od ostatních zmíněných metod kódování. [21]

Metodu kódování webové stránky lze vyčíst z hlavičky HTML dokumentu, kde se nachází tag `meta` s parametrem `charset`.

Nepotřebné prvky webové stránky

Na webových stránkách se kromě textu vyskytuje i mnoho obsahu, který nesouvisí s jeho obsahem nebo funkcí a může ovlivnit uživatelskou zkušenost. Jedná se například o grafiku, logo či vyhledávací mechanismy, které bývají situované v hlavičce webové stránky;

dále také reklamní sdělení a patičku se základními informacemi o webové stránce, jež bývají umístěny zpravidla v postranním panelu.

Existující nástroje

Existuje mnoho nástrojů pro web scraping, které mohou být použity pro extrakci dat z různých typů webových stránek. Nejpopulárnějšími nástroji pro *Python*² jsou kupříkladu *Python knihovna – BeautifulSoup*³ a *Python framework Scrapy*⁴.

Modul pro dolování informací z textu

Tento modul se specializuje na získávání informací z textu, které přispívají k ověření dezinformačních článků. Součástí tohoto procesu je také sběr citací a odkazů na jiné články a webové stránky, které mohou být také zdrojem dezinformací. Tyto informace jsou dále analyzovány pomocí modulu pro dolování příznaků dezinformací, jak je popsáno v 5.2.

Modul pro dolování příznaků dezinformací v textu

Modul se skládá z více entit, které jsou na sobě nezávislé, avšak dohromady tvoří zásadní složku systému.

Ověření velikosti písmen

Velikost písma často slouží jako prostředek zdůraznění důležitých informací a názorů autora v textu, a proto je často využívána i v dezinformačních člancích. Je důležité být obezřetný při interpretaci velikosti písma, jelikož sama o sobě není dostatečným indikátorem důvěryhodnosti textu. Při analýze textu je nutné brát v úvahu i další faktory a kontext, aby bylo možné adekvátně posoudit její význam a pravdivost.

Ověření reputace webové stránky

Reputace webových stránek lze ověřit pomocí různých metod a technik, které byly již získány a ověřeny. Existují seznamy webových stránek, které jsou považovány za dezinformační a nevyhovují standardům kvalitních informačních zdrojů (např. na webu *Neovlivní.cz* [6], *Nelež.cz* [8] či ve výroční zprávě *Bezpečnostního centra evropských hodnot* [36]). Seznam z webu *Nelež.cz*, z něhož bylo čerpáno, je k nahlédnutí v příloze C, je vytvářen ověřenou neziskovou organizací, která sleduje a analyzuje různé faktory, jako je kvalita obsahu, nezávislost, zdroje informací a další. V situaci, kdy nelze stránky považovat za čistě dezinformační, nýbrž mohou se na ně často například odkazovat ty objektivně pravdivé, jsou weby začleňovány na „šedý seznam“. Nicméně je třeba mít na paměti, že reputaci webových stránek nelze posuzovat pouze na základě jednoho faktoru, ale je nutné zohlednit celkový kontext a další faktory.

Ověření reputace autora příspěvku

Tato úloha je velmi podobná předešlé úloze, která se zabývá ověřením reputace webové stránky. Na základě předem získaných znalostí lze reputaci autora soudit.

²<https://www.python.org>

³<https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

⁴<https://scrapy.org>

Uživatelské prostředí

Cílem je vytvoření uživatelsky přívětivého rozhraní, které usnadní uživatelům zadávání úkolů do systému a sníží nároky na jejich kognitivní zátěž.

Tato aplikace bude koncipována jako webová aplikace, kde uživatel bude moci prostřednictvím ní komunikovat se systémem.

5.3 Limity předkládané práce

Předložená práce má řadu limitací, vzhledem k rozsahu předmětu výzkumu a množství dat, která byla v průběhu práce získána.

Jedním z limitů systému zpracování článků na propagandistických webových stránkách je vysoká náročnost na výpočetní výkon vzhledem k celkovému množství dat. Překážkou v získání dostatečného množství dat k získání vypovídajících výsledků jsou i faktory jako časová či paměťová složitost, které jsou determinovány velikostí celého systému a počtem zpracovávaných článků, jejichž počet může přesahovat stovky tisíc textů.

Limitem může také být chybovost systému, kdy je potřeba brát v úvahu systémové problémy a aplikovat korekční opatření, aby se chybovost minimalizovala. Je také důležité, aby byly výstupy systému pravidelně kontrolovány a ověřovány, aby nemohlo dojít k chybnému nemotivovanému označení článku a tím ke vzniku misinformace.

Vzhledem ke zmíněným limitům vyvstávají další otázky, které by mohly být předmětem budoucího výzkumu. Jednou z možností by mohlo být využití velkého množství zpracovaných dat pro trénování modelu *BERT* a následné využití tohoto modelu pro úlohy odhalování dezinformací. Takový přístup by mohl přinést významné zlepšení v přesnosti rozpoznávání a klasifikaci dezinformací a zároveň by bylo možné zkoumat, jaký vliv má velikost trénovací množiny na výkon modelu.

Kapitola 6

Implementace

V této kapitole je popsána implementace systému a nástroje, které byly využity při jeho realizaci.

Pro implementaci byl zvolen programovací jazyk *Python 3*¹, který byl vybrán z několika důvodů. Prvním z nich je, že *Python* je vysokoúrovňový programovací jazyk a poskytuje nástroje pro jednodušší a přehlednější implementaci.

Další velkou výhodou je velké množství dostupných knihoven právě pro *Python*, které jsou taktéž užitečné. Existuje mnoho knihoven, které umožňují jednoduchou manipulaci s daty, tvorbu grafů, statistické analýzy, strojové učení a mnoho dalšího.

Python také poskytuje možnost využití objektového programování, což umožňuje snadnou správu a organizaci kódu.

Vzhledem k výše uvedeným výhodám byl *Python* vhodnou volbou pro implementaci systému, který pracuje s daty a provádí jejich analýzu a zpracování.

V následujících podkapitolách je popsána struktura systému v 6.1 s názorným schématem, dále je v 6.2 představena komunikace s databází a operace s daty. V podkapitole 6.3 jsou popsány moduly, pomocí kterých se stahují data z webových stránek. V podkapitole 6.4 jsou představeny moduly pro zpracování dat. Dále v kapitole 6.5 je prezentována implementace uživatelského prostředí. Každá podkapitola obsahuje detailní popis použitých technologií a postupů.

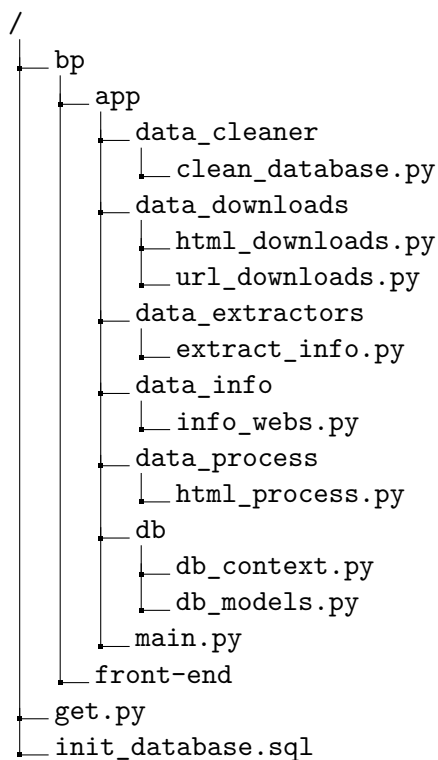
6.1 Struktura systému

Systém je implementován ve složce s názvem `bp`. Skript `get.py`, který se nachází v kořenové složce, pravidelně spouští tento systém. Všechny moduly systému se nacházejí v adresáři `app` a vzájemně komunikují. Hlavní řídicí skript systému je nazvaný `main.py`. Uživatelské rozhraní je implementováno v adresáři `frontend`, který obsahuje soubor `app.py`, a další konfigurační a pomocné soubory.

Systém se skládá z celkem šesti základních modulů. Prvním je adresář `data_cleaner`, který obsahuje skript `clean_database.py`. Tento skript je zodpovědný za přetřídění jednotlivých záznamů, mazání duplikátů, odstranění nevhodných znaků a další úpravy. Druhým modulem je `data_downloads`, který má za úkol stáhnout požadovaná data z webových stránek. Tento modul se skládá ze dvou skriptů: `url_downloads.py` a `html_downloads.py`. První ze zmíněných skriptů má za cíl extrahovat všechny url adresy z webové stránky, druhý poté samotnou extrakci html kódu z webové stránky. Třetí z modulů, `data_extractors`,

¹<https://www.python.org>

obsahuje skript `extract_info.py`, který extrahuje všechny důležité informace z html kódu, jako je například text článku. Dalším z modulů je `data_info` se skriptem `info_webs.py`, který získává potřebná informace, pro práci a funkčnost ostatních modulů. Modul pro zpracování dat, `data_process` obsahuje skript `html_process`. Modul `db`, tedy poslední modul, se zabývá komunikací systému s databází. Obsahuje dva základní skripty `db_context.py` a `db_models.py`.



6.2 Přístup k databázi

V této podkapitole jsou popsány způsoby, jakými se přistupuje k *MySQL* databázi. Pro tento účel jsou využity dva moduly.

Modul `db_context`

Jedná se o první z modulů pro přístup k databázi. Má na starosti navazování spojení s databází a jeho následné ukončení. Navázání spojení s databází probíhá pomocí funkce `pymysql.connect`, kde jsou jako jednotlivé parametry uvedeny konkrétní hodnoty pro přístup k databázi.

```
connection = pymysql.connect (host="127.0.0.1", user="root",
database="data_disinfo")
```

Modul `db_modules`

Druhý modul zajišťuje práci s daty v databázi. To provádí pomocí vytvořených tříd, které mapují databázové tabulky na objekty *Pythonu*. Tyto třídy zahrnují *CRUD* (*create*, *read*, *update*, *delete*) operace pro manipulaci s daty v databázi.

Díky využití objektově orientovaného přístupu je práce s databází přehlednější a strukturovanější. Toto řešení zajišťuje také bezpečnost a snižuje pravděpodobnost vzniku chyb, které by mohly být způsobeny neopatrným zacházením s databázovými dotazy.

6.3 Stahování dat

Implementace stahování dat z webových stránek byla nejprve řešena pomocí *RSS*², ale nakonec bylo rozhodnuto o implementaci stahování webových stránek pomocí sekvenčního procházení a následného stahování. Hlavním důvodem nevyužití *RSS* byla nedostupnost této technologie u části dezinformačních webových stránek.

Stahování dat je implementováno modulem `data_downloads`, který obsahuje skripty `url_downloads.py` a `html_downloads.py`

Dotazování webových serverů

Pro získávání dat z webových stránek byla implementována funkce `get_html`, která využívá knihovnu `requests`³. Tato knihovna poskytuje nástroje pro odesílání HTTP požadavků na webové servery.

Požadavek na webový server se zasílá pomocí funkce `get`, kde je prvním parametrem předána URL adresa stránky, ze které se stahuje obsah. Dalšími nepovinnými parametry jsou například vytvoření falešného uživatelského agenta viz 6.3, nastavení proxy adresy nebo nastavení časového intervalu pro případ, že server neodpovídá.

Funkce `get_html` vrací stav požadavku pomocí `response.status_code` a obsah získané stránky pomocí `response.content`.

```
response = requests.get(url, headers=self.headers, timeout=10)
```

Prevence blokování

Využití náhodného výběru hlaviček přispívá k prevenci blokování ze strany webových serverů, protože se tím snižuje pravděpodobnost detekce automatizovaného skriptu, který by mohl být považován za útok na server.

Jelikož je cílem stáhnout všechny články z webových stránek, které byly publikovány od vzniku webové stránky a některé stránky obsahují i více než tisíc článků, je potřeba měnit identitu klienta, aby nedošlo k detekci automatizovaného skriptu a blokování ze strany webových serverů, kterému se může dotazování jevit jako útok na něj. Jednoduchou metodou, jak tomuto předejít, je zamezit identifikaci skriptu a simulovat prohlížeč, který žádá o zobrazení webové stránky. Každý prohlížeč má svou vlastní hlavičku a pro dotazování je využít náhodný výběr ze seznamu hlaviček. Každá hlavička obsahuje agenta, který posílá dotaz na webový server, předešlý vychozí bod, požadovaný typ dat a cookies.

Agent je identifikátor, který říká webovému serveru, jaký typ prohlížeče a operačního systému se používá. Předešlý vychozí bod říká serveru, odkud se stránka načetla. Je tedy vhodné měnit url adresy vychozího bodu. Cookies jsou malé soubory, které webové servery ukládají do prohlížeče, aby si pamatovaly informace o uživateli a jeho předchozích interakcích se stránkou.

²RSS je určeno pro čtení novinek na internetu, viz <https://www.jaknainternet.cz/page/1640/rss-kanaly/>

³<https://pypi.org/project/requests/>

V případě blokování ze strany webových serverů je také možné využít proxy serverů, které umožňují dotazování na web z jiné IP adresy, než je adresa počítače uživatele.

Stažení URL adres

Stažení URL adres probíhá pomocí skriptu `url_downloads.py`, který sbírá relevantní odkazy z archivů webových stránek. Funkce `get_web_info()` analyzuje strukturu webu a určuje, zda se archiv webové stránky listuje po jednotlivých stránkách s iterací jedna, nebo je archiv uspořádán dle data publikace článků. V případě druhého se musí odkazy listovat po měsících a letech existence webové stránky.

Po analýze struktury webové stránky následuje samotné shromažďování a ukládání odkazů na publikované články. Protože při vyhledávání těchto odkazů vzniká mnoho redundancí, jsou ověřována dvě fakta. První z nich se týká toho, zda je URL adresa adekvátní a představuje odkaz na článek na webové stránce. To je řešeno pomocí regulárních výrazů a následně při zpracování a vyhodnocování dat, viz 6.4. Druhý fakt se týká již uložených URL adres v databázi. Systém tomu předchází pomocí dotazování se databáze na již uložené URL adresy.

Stažení HTML

Po stažení všech URL adres z archivu webové stránky se spustí skript `html_downloads.py`, který obsahuje třídu `HTMLs`. Pomocí funkce `get_html_text` postupně stahuje všechna data z dané URL adresy, které jsou poté uložena do databáze. Před uložením HTML kódu do databáze k dané URL adrese jsou všechny textové řetězce upraveny pomocí funkce `html.escape` z knihovny `html`. To zamezuje problémům při vkládání dat do databáze.

Je nutno poznamenat, že ke spuštění systému dochází periodicky v určitých časových intervalech dle nastavení, a proto je stahován obsah všech stránek, které ještě nebyly navštíveny a uloženy. Nemá tedy vliv na kvalitu shromáždění webových článků a i v případě, že je spuštěno stahování obsahu webových stránek na nově vložené webové stránce, je z ní stažen veškerý obsah.

6.4 Zpracování dat

Modul `data_process` se zabývá především extrakcí důležitých informací z HTML kódu. Jedná se zejména o 4 hlavní úkoly:

- Extrakce textu článku – stěžejní část stahovaných dat.
- Extrakce titulku článku – většinou velmi výstižně shrnuje obsah článků a často se skládá z klíčových slov, vyskytujících se v něm. Občas ale mohou být titulky zavádějící, jelikož jejich primárním cílem je zapůsobit na čtenáře.
- Extrakce autora článku – autor článku či editor, který článek publikoval.
- Extrakce data publikování článků – datum publikování článků.

Pomocí skriptu `html_process.py`, ve kterém se nachází třída `ProcessText`, jež obsahuje funkce, se extrahují z html kódu již zmíněné atributy. Pro extrakci se využívá především knihovna *BeautifulSoup*⁴, která je určena pro extrakci dat z HTML a XML. Funkci je předán html kód v odescapované podobě.

⁴<https://beautiful-soup-4.readthedocs.io/en/latest/>

```
SOUP = BeautifulSoup(unescape_html_string, 'html.parser')
```

Modul `data_cleaner`

V modulu `data_cleaner` jsou řešeny požadavky na integritu dat, která jsou uložena v databázi. Tento modul zajišťuje odstranění nevhodných dat, která vznikla při automatickém stahování obsahu webových stránek. Kromě toho jsou z databáze odstraňovány duplicity, které mohou vzniknout v důsledku automatického stahování stejného obsahu vícekrát.

Základem je skript `clean_database.py`, který obsahuje třídu `DataClean`. V tomto modulu jsou implementovány funkce pro mazání duplicit v databázi – `clean_duplicates`, mazání všech znaků, které nejsou *ASCII*⁵ – `clean_non_ASCII`, a mazání všech nových řádků v textu – `clean_new_lines`.

Modul `data_extractors`

Jedná se o modul, který slouží ostatním modulům a poskytuje pomocné funkce pro extrahování informací z textu. Tento modul obsahuje několik funkcí, které jsou určeny k extrakci určitých informací z textových dat. A skládá se ze dvou tříd `DateExtractor` a `NameExtractor`.

Funkce v obou třídách jsou implementovány pomocí pravidel pro rozpoznávání vzorů v textu a následného extrahování odpovídajících informací. Pro rozpoznávání vzorů je využito zejména regulárních výrazů.

Modul `data_info`

V modulu `data_info` se nachází třídy `GetDataWebs` a `GetDataArticles`, které slouží k získávání informací o webových stránkách a článcích z databáze. Tyto třídy komunikují s modulem pro komunikaci s databází a získávají potřebná data, která následně poskytují ostatním modulům projektu.

Modul primárně poskytuje data a informace o jednotlivých článcích a webových stránkách ostatním modulům projektu. Pro získání dat využívá modulu `db` a poté je zpracuje do podoby, která je vhodnější pro ostatní moduly.

6.5 Uživatelské prostředí

Za účelem snazší uživatelské komunikace se systémem je vytvořena webová aplikace postavená na frameworku *Streamlit*⁶, jež umožňuje rychlý a jednoduchý vývoj webových aplikací s přehledným uživatelským rozhraním. Implementace uživatelského prostředí se nachází v adresáři `front-end`, kde jsou umístěny soubory pro správný běh aplikace. Hlavní část uživatelského prostředí je implementována ve skriptu `app.py`.

⁵Jedná se o kódovou tabulku, která definuje jednotlivé znaky.

⁶<https://streamlit.io>

6.6 Vyhodnocení implementace systému

System disponuje funkcí sběru a ukládání velkého množství dat z webových stránek. To zahrnuje stahování dat z těchto stránek a jejich následné ukládání do databáze pro další zpracování.

Příkladem úspěšně staženého a zpracovaného článku z webové stránky *protiproud.info* může být záznam v tabulce `articles`. Každý záznam v této tabulce je identifikován unikátním `id`, který slouží jako primární klíč. Obrázek 6.1 ukazuje příklad takového záznamu v tabulce `articles`.

id	web_id	url	text	insert_date	create_date	title	html
63279	2	https://protiproud.info/politika/6228-byl-to-pekny...	Byl to fajny rok, bo zme se mohli v mnohem poucit...	2023-04-19 22:00:00	7. 1. 20227. 1. 2022	1. Byl to pěkný rok: Ten další bude horší. Méli bycho...	<!DOCTYPE html><html xmlns="http:...

Obrázek 6.1: Příklad staženého článku z webu.

Na dalším obrázku 6.2 lze vidět nalezené url adresy z textu článků, které jsou ověřeny a považovány za funkční url adresy. Každá url adresa má následně přiřazené `id` článků, ve kterém byla nalezena a také své `id`, které adresu jednoznačně identifikuje.

id	article_id	url
1	24310	http://www.idnes.cz
3	272156	http://Brighteon.tv
4	272156	http://bunkr.Pro
5	272156	http://CurrencyReset.news
6	272156	http://Brighteon.tv

Obrázek 6.2: Příklad nalezených url z textu článků.

Kapitola 7

Závěr

Kromě propojení problematiky dezinformací ve společnosti a politice a studia informačních technologií bylo cílem této práce především představit čtenáři problematiku odhalování dezinformací a využití moderních technologií za tímto účelem.

Motivací autora této bakalářské práce byla nejen snaha přispět k výzkumu v této oblasti, ale také touha přijít s nástrojem, který by mohl mít praktický společenský přesah a jehož by bylo možné využívat při rozpoznávání dezinformací a tím i k prevenci jejich šíření.

Teoretická část v první řadě definuje terminologii, s níž práce operuje. Jsou to především pojmy *dezinformace* a *propaganda* a jejich druhy, také je zde vymezen pojem *propagandistická stránka* a jsou ustanoveny zdroje seznamů webových stránek označených za dezinformační, s nimiž je v práci dále nakládáno. Práce dále uvádí zmiňované pojmy do společenského kontextu a pojednává o možných negativních dopadech šíření dezinformací na společnost. K šíření dezinformací dochází zejména skrze sociální sítě, řetězové e-maily a propagandistické webové stránky, na poslední zmiňovaný způsob šíření se tato práce orientuje. Boj s dezinformacemi je zdoluhavý a komplikovaný proces a nejúčinnějším nástrojem v tomto boji je systematické ověřování informací a následné odhalování dezinformací.

K odhalování dezinformací se používá nepřehledné množství metod, jež vychází z různých výzkumů, které vznikly v posledních letech. V podkapitole *Přístupy k odhalování dezinformací* jsou vybrané metody popsány a uchopeny teoreticky, v následující kapitole je potom podrobněji zpracováno jejich praktické fungování. V průběhu vývoje systému se ukázalo, že metody založené na analýze jazyka či syntaxe jsou sice vhodné, ale jsou náročné na potřebné množství dat. Pro účely této práce byl nakonec zvolen znalostní a síťový přístup, které analyzují stránky na základě informací o kontextu, v němž se objevují.

Původně bylo pro vývoj systému zamýšleno využití modelu *BERT*, který je jedním z nejnovějších přístupů v oblasti strojového učení. Tento model potřebuje pro své učení rozsáhlý vzorek dat, což je jedním z limitů této práce zmíněných podrobněji níže.

Model *BERT* ke svému učení využívá rozsáhlé datové sady, které je pro tyto účely potřeba vytvořit a připravit pro trénování modelu. Konkrétní postup vytváření datových sad a jednotlivé kroky přípravy pro trénování jsou popsány v kapitole *Datové sady*. V rámci vývoje systému byla vytvořena datová sada, která obsahuje jednotky stovek tisíc článků stažených z propagandistických webových stránek.

Kapitola *Návrh systému a jeho limity* přináší ucelený návrh systému sloužícího ke stahování a archivaci dat z propagandistických webových stránek. Systém se pravidelně spouští a je implementován tak, že postupně prochází jednotlivé weby ze stanoveného seznamu webových stránek označených za dezinformační. Při procházení těchto zdrojů vyhodnocuje, zda články už jsou v systému evidovány. Pokud narazí na nový článek, který ještě není

stažen, dojde k jeho stáhnutí včetně velkého množství metadat. Po uložení do databáze proběhne extrakce samotného textu článku a informací potřebných pro jeho identifikaci (datum zveřejnění, autor a titul). Extrahovaná data poté podléhají analýze, která hledá spojitosti s jinými články v databázi a případné odkazy na jiné problematické texty či weby. Tyto vazby jsou následně ukládány do databáze a vytvářejí síť propojující jednotlivé problematické texty.

Jak již bylo zmíněno výše, práce naráží na limity, které jsou způsobeny nedostatečným výpočetním výkonem či vysokou časovou a paměťovou náročností při sběru dat. Limitem může být také možná chybovost systému, která vyžaduje pravidelnou kontrolu a napravování takových chyb. Práce však současně otevírá možnost dalšího výzkumu, nabízí se například využití modelu pro zpracování přirozeného jazyka např. modelu *BERT*.

Praktické provedení implementace systému je do detailu popsáno v kapitole *Implementace*. V průběhu implementace došlo ke zjištění výše zmíněných limitů, což následně značně ztížilo vývoj přesnějšího a funkčnějšího systému. Přesto byla vytvořena datová sada čítající několik stovek tisíc textů z dezinformačních webových stránek.

Finální verze navrhovaného systému dokáže při spuštění stahovat články ze stanovených webů, archivovat je, analyzovat a extrahovat jejich obsah a ukládat tyto informace do databáze viz výše. Implementovaný vzorec vykazuje jistou míru chybovosti při stahování dat z jednotlivých webových stránek, proto je zde ještě prostor ke zlepšení vzorce tak, aby dokázal lépe extrahovat informace ze stránek, které mají různorodou stavbu, což by vedlo i ke zpřesnění extrakce informací z textu.

Součástí práce je také návrh uživatelského rozhraní, které je primárně určené pro uživatele zabývající se problematikou dezinformací na odbornější rovině. Vyžaduje totiž základní znalost programování, ale i systému samotného. Uživatelské prostředí by bylo možné dále vyvíjet a zlepšovat (např. zlepšit podobu uživatelského prostředí tak, aby uživatelé nebyli nuceni pracovat se zdrojovým kódem).

Závěrem lze tedy říci, že původní cíl práce vytvořit systém aktivně vyhledávající a rozpoznávající dezinformace byl naplněn dílčím způsobem, protože se autor během vývoje tohoto systému setkal s těžko překonatelnými limity (nedostatek výpočetního výkonu atd.). Byl však navržen funkční systém, který dokáže stahovat, archivovat a analyzovat informace stahované z dezinformačních webů. Tento systém by mohl pomoci při dalším vývoji nástroje určeného pro odhalování dezinformací. Vytvořit a zcela dokončit takový nástroj však nebylo v silách autora práce. Systém je nicméně k dispozici dalším výzkumníkům, kteří mohou na práci navázat. Veškeré podklady k celému systému a projektu (schémata, zdrojové kódy, komentáře kódu) jsou volně k dispozici na *GitHubu*¹. Boj s dezinformacemi je velmi náročná disciplína, která vyžaduje nejen množství lidských zdrojů, ale je také náročná na výpočetní výkon, kapacity úložišť atd. Systém, který by odhaloval dezinformace na základě jazykové, nebo znalostní analýzy, by byl jistě velkým přínosem pro společnost, šíření dezinformací má totiž prokazatelně negativní společenský vliv. Vznik takového systému, který by zcela samostatně a bezchybně rozpoznával fakta od dezinformací zůstává otázkou budoucího vývoje. Autor je přesvědčen, že při dostatku potřebných zdrojů a po odstranění jmenovaných limitů, je něco takového jistě možné.

¹https://github.com/kiwibojovnik/B_THESIS.git

Literatura

- [1] *Učení bez učitele* [online]. Institut biostatistiky a analýz Lékařské fakulty Masarykovy univerzity [cit. 2023-03-10]. Dostupné z: <https://portal.matematickabiologie.cz/index.php?pg=analiza-a-hodnoceni-biologickych-dat--umela-inteligence--neuronove-site-jednotlivy-neuron--adaptacni-dynamika-neuronu--uceni-bez-ucitele>.
- [2] *Řetězové e-maily dostává devět z deseti seniorů* [online]. Elpida, o. p. s., září 2018 [cit. 2023-03-29]. Dostupné z: <https://www.elpida.cz/fake-news-po-cesku>.
- [3] *Kategorie:Dezinformace* [online]. Atlas Kospirací, 22. září 2019 [cit. 2023-01-16]. Dostupné z: <https://atlaskonspiraci.cz/Kategorie:Dezinformace>.
- [4] *Řetězové e-maily* [online]. Čeští elfové, 2022 [cit. 2023-02-28]. Dostupné z: <https://cesti-elfove.cz/retezove-e-maily/>.
- [5] 2023 [cit. 2023-05-06]. Dostupné z: <https://trends.google.com/home?geo=CZ&hl=cs>.
- [6] *Databáze dezinformačních webů* [online]. 2023 [cit. 2023-05-08]. Dostupné z: <https://neovlivni.cz/vnitro-ma-databazi-dezinformacnich-webu-koho-na-ni-zaradilo/>.
- [7] *Definice dezinformací a propagandy* [online]. Ministerstvo vnitra České republiky, 2023 [cit. 2023-01-16]. Dostupné z: <https://www.mvcr.cz/chh/clanek/definice-dezinformaci-a-propagandy.aspx>.
- [8] *Seznam dezinformačních webů* [online]. 2023 [cit. 2023-05-08]. Dostupné z: <https://www.nelez.cz>.
- [9] BROWNLEE, J. *What Are Word Embeddings for Text?* [online]. Machine learning mastery, červenec 2019 [cit. 2023-03-06]. Dostupné z: <https://machinelearningmastery.com/what-are-word-embeddings/>.
- [10] BURÝŠEK, J. *CO JE TO DEZINFORMACE?* [online]. Bezfaulu.net, červenec 2019 [cit. 2023-03-22]. Dostupné z: <https://bezfaulu.net/clanky/o-manipulaci/co-je-to-dezinformace/>.
- [11] BURÝŠEK, J. *RÉTORIKA PROPAGANDY* [online]. Bezfaulu.net, červenec 2019 [cit. 2023-03-22]. Dostupné z: <https://bezfaulu.net/clanky/o-manipulaci/retorika-propagandy/>.
- [12] CHORAŚ, M., DEMESTICHAS, K., GIELCZYK, A., HERRERO Álvaro, KSIENIEWICZ, P. et al. Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study. *Applied Soft Computing*. 2020, sv. 101, s. 15, [cit. 2023-04-14]. DOI: 10.1016/j.asoc.2020.107050.

- [13] DAŇHELOVÁ, J. *Zpětnovazební učení pro řešení herních algoritmů* [online]. Brno, 2018. [cit. 2023-03-10]. 48 s. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce KOLAŘÍK, I. M. Dostupné z: https://www.vut.cz/www_base/zav_prace_soubor_verejne.php?file_id=175619.
- [14] DEVLIN, J., CHANG, M.-W., LEE, K. a TOUTANOVA, K. {BERT}: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics, červen 2019, s. 4171–4186 [cit. 2023-04-15]. DOI: 10.18653/v1/N19-1423. Dostupné z: <https://aclanthology.org/N19-1423>.
- [15] FALLIS, D. What Is Disinformation? *Library Trends*. Johns Hopkins University Press. Duben 2015, sv. 63, č. 3, s. 401–426, [cit. 2023-03-29]. DOI: 10.1353/lib.2015.0014.
- [16] FLEW, T., SPURGEON, C., DANIEL, A. a SWIFT, A. The promise of computational journalism. *Journalism Practice*. Duben 2012, sv. 6, s. 157–171, [cit. 2023-04-15]. DOI: 10.1080/17512786.2011.616655.
- [17] GUO, Z., SCHLICHTKRULL, M. S. a VLACHOS, A. A Survey on Automated Fact-Checking. *CoRR*. 2021, abs/2108.11896, [cit. 2023-04-15]. Dostupné z: <https://arxiv.org/abs/2108.11896>.
- [18] GURUSAMY, V. a KANNAN, S. Preprocessing Techniques for Text Mining. In: říjen 2014.
- [19] HASSAN, N., GOMAA, W., KHORIBA, G. a HAGGAG, M. Credibility Detection in Twitter Using Word N-gram Analysis and Supervised Machine Learning Techniques. *International Journal of Intelligent Engineering and Systems*. Únor 2020, sv. 13, s. 291–300, [cit. 2023-04-14]. DOI: 10.22266/ijies2020.0229.27.
- [20] HOLEŠINSKÁ, A. Návrh tvorby webových stránek destinace. *Brno, ESF MU*. 2005, [cit. 2023-05-06].
- [21] JANOVSKÝ, D. *Diakritika na HTML stránkách*. [cit. 2023-05-06]. Dostupné z: <https://www.jakpsatweb.cz/cestina.html>.
- [22] KESAVARAJ, G. a SUKUMARAN, S. A study on classification techniques in data mining. In: *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*. 2013, s. 1–7 [cit. 2023-04-15]. DOI: 10.1109/ICCCNT.2013.6726842.
- [23] KIURCHEVA, O. *Fenomén hybridní informační války a obraz anexe Krymu v ruských médiích působících na území Německa* [online]. 2019 [cit. 2023-05-06]. [cit. 2023-05-06]. Diplomová práce. Masarykova univerzita, Fakulta sociálních studií, Brno. SUPERVISOR : Petra Kuchyňková. Dostupné z: <https://is.muni.cz/th/psmb0/>.
- [24] KUČEROVÁ, H. *Dezinformace* [online]. KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV), 2003 [cit. 2023-02-28]. Dostupné z: https://aleph.nkp.cz/F/?func=direct&doc_number=000000095&local_base=KTD.

- [25] MATOUŠEK, V. *Strojové učení* [online]. Plzeň: Katedra informatiky a výpočetní techniky, Fakulta aplikovaných věd, Západočeská univerzita v Plzni, 2022 [cit. 2023-03-10]. Dostupné z: https://www.kiv.zcu.cz/studies/predmety/uzi/Folie_ZS/Stroj_uceni.pdf.
- [26] MIKOLOV, T., CHEN, K., CORRADO, G. a DEAN, J. *Efficient Estimation of Word Representations in Vector Space*. 2013 [cit. 2023-04-15].
- [27] NASEEM, U., RAZZAK, I., KHAN, S. K. a PRASAD, M. A Comprehensive Survey on Word Representation Models: From Classical to State-of-the-Art Word Representation Language Models. In: [online]. Woodstock, NY. ACM, New York, NY, USA: [b.n.], červen 2018, s. 46 [cit. 2023-03-29]. DOI: 10.1145/1122445.1122456. Dostupné z: <https://arxiv.org/abs/2010.15036>.
- [28] PARIKH, S. a ATREY, P. Media-Rich Fake News Detection: A Survey. In:.. Duben 2018 [cit. 2023-04-14]. DOI: 10.1109/MIPR.2018.00093.
- [29] PENNINGTON, J., SOCHER, R. a MANNING, C. GloVe: Global Vectors for Word Representation. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, říjen 2014, s. 1532–1543 [cit. 2023-04-15]. DOI: 10.3115/v1/D14-1162. Dostupné z: <https://aclanthology.org/D14-1162>.
- [30] RAZA, S. a DING, C. Fake news detection based on news content and social contexts: a transformer-based approach. *International Journal of Data Science and Analytics*. Květen 2022, sv. 13, s. 1–28, [cit. 2023-04-17]. DOI: 10.1007/s41060-021-00302-z.
- [31] SCHÜTZ, M., SCHINDLER, A., SIEGEL, M. a NAZEMI, K. Automatic Fake News Detection with Pre-trained Transformer Models. In: DEL BIMBO, A., CUCCHIARA, R., SCLAROFF, S., FARINELLA, G. M., MEI, T. et al., ed. *Pattern Recognition. ICPRI International Workshops and Challenges*. Cham: Springer International Publishing, 2021, s. 627–641 [cit. 2023-04-15]. ISBN 978-3-030-68787-8.
- [32] SHU, K., WANG, S. a LIU, H. Exploiting Tri-Relationship for Fake News Detection. Prosinec 2017, [cit. 2023-04-15].
- [33] ZACHAR KUCHTOVÁ, J. J., ed. *Bezpečnosť elektronickej komunikácie* [online]. Bratislava: Akadémie Políciejního sboru v Bratislavě, Květen 2022, s. 247 [cit. 2023-03-22]. Dostupné z: https://www.bezpek.eu/wp-content/uploads/2023/03/Bezpecnost-elektronickej-komunikacie_2022.pdf#page=41.
- [34] ZHANG, D., ZHOU, L., KEHOE, J. a KILIC, I. D. What Online Reviewer Behaviors Really Matter? Effects of Verbal and Nonverbal Behaviors on Detection of Fake Online Reviews. *Journal of Management Information Systems*. Duben 2016, sv. 33, s. 456–481, [cit. 2023-04-14]. DOI: 10.1080/07421222.2016.1205907.
- [35] ZHOU, L. a ZHANG, D. Following Linguistic Footprints: Automatic Deception Detection in Online Communication. *Commun. ACM*. New York, NY, USA: Association for Computing Machinery. Září 2008, sv. 51, č. 9, s. 119–122, [cit. 2023-04-14]. DOI: 10.1145/1378727.1389972. ISSN 0001-0782. Dostupné z: <https://doi.org/10.1145/1378727.1389972>.

- [36] ŠPALKOVÁ, V. K., PASTÍR, M., PŘÍVRATSKÝ, J., ČINČEROVÁ, Z. a ŠERÁK, V. Výroční zpráva o stavu české dezinformační scény pro rok 2022. [online]. Bezpečnostní centrum Evropské hodnoty. 2023, [cit. 2023-05-08]. Dostupné z: <https://www.nelez.cz>.

Příloha A

Obsah příloženého paměťového média

Na příloženém paměťovém médiu se nachází:

- Soubor `bp.pdf`, který obsahuje technickou zprávu určenou pro čtení práce na počítači.
- Soubor `bp_print.pdf`, který obsahuje technickou zprávu určenou pro tisk.
- Složka `bp`, která obsahuje zdrojové kódy a ostatní data potřebné pro vytvoření `bp.pdf` a `bp_print.pdf`.
- Soubor `plakat.pdf`, který obsahuje plakát prezentující tuto práci.
- Složka `source_codes`, která obsahuje zdrojové kódy implementovaného systému a další podsložky. Taktéž obsahuje své vlastní `README`, ve kterém se popisuje, jakým způsobem systém spustit.

Příloha B

Manuál pro spuštění systému

Pro správnou funkčnost systému je nutné zajistit dostupnost konkrétních prostředků a nástrojů, jak je uvedeno v tomto manuálu. Další podrobnosti se lze dočíst v samotném README implementovaného systému. V následujícím seznamu jsou popsány nezbytné nástroje a knihovny, které je nutné nainstalovat.

- Python 3.11.1,
- databázový systém MySQL,
- následující Python knihovny:
 - Requests,
 - Beautiful Soup,
 - html.parser,
 - Random,
 - Streamlit,
 - Logging,
 - Regular Expressions (re),
 - Datetime,
 - Os,
 - Sys.

Postup pro spuštění systému:

1. Nainstalujte Python 3.11.1 a potřebné knihovny pro něj pomocí příkazu `pip install`.
2. Nainstalujte databázi MySQL a ověřte její správnou instalaci.
3. Nainportujte inicializační skripty do databáze, které vytvoří databázi *dezinformace*.
4. Upravte soubor `config.py` s přístupovými údaji pro připojení k databázi MySQL.
5. Spusťte aplikaci pomocí příkazu `streamlit run app.py` v příkazovém řádku.
6. Otevřete prohlížeč a přejděte na adresu `http://localhost:8501`.
7. Po úspěšném spuštění byste měli vidět hlavní stránku aplikace v prohlížeči.

Příloha C

Seznam dezinformačních webů

Seznam dezinformačních webů byl převzat z iniciativy *nelez.cz*. Seznam webových stránek byl zevrubně zanalyzován a nefunkční webové stránky byly vyřazeny. Všechny zmíněné webové stránky jsou aktivní k 9. květnu 2023.

Id	Název webu	URL	Aktivita
1	Ac24	https://ac24.cz	aktivní
2	Aeronet	https://aeronet.news	aktivní
3	Bezpolitické korektnosti	https://bezpolitickékorektnosti.cz	aktivní
4	Ceskezpravy	https://ceskezpravy.eu	aktivní
5	Ceskavec	https://ceskavec.com	aktivní
6	Cz24	https://cz24.news	aktivní
7	Dt24	https://dt24.cz	aktivní
8	Důležité 24	https://zpravy.dt24.cz	aktivní
9	E-republika	http://e-republika.cz	aktivní
10	Islamizace	https://islamizace.cz	aktivní
11	Jihoceskenovinky	http://jihoceskenovinky.cz	aktivní
12	Jihomoravskénovinky	http://jihomoravskénovinky.cz	aktivní
13	Karlovarskénovinky	http://karlovarskénovinky.cz	aktivní
14	Kralovehradeckénovinky	http://kralovehradeckénovinky.cz	aktivní
15	Moravskoslezskénovinky	http://moravskoslezskénovinky.cz	aktivní
16	Narodni-noviny	https://narodni-noviny.cz	aktivní
17	Necenzurovaná pravda	https://necenzurovanápravda.cz	aktivní
18	Necenzurujeme	https://necenzurujeme.cz	aktivní
19	Nejvic-info	https://nejvic-info.cz	aktivní
20	Novarepublika	https://novarepublika.online	aktivní
21	Novinky vysocina	http://novinkyvysocina.cz	aktivní
22	Otevrisvoumysl	https://otevrisvoumysl.cz	aktivní
23	Outsidermedia	https://outsidermedia.cz	aktivní
24	Pardubickénovinky	http://pardubickénovinky.cz	aktivní

Tabulka C.1: Tabulka dezinformačních webových stránek.

Id	Název webu	URL	Aktivita
25	Parlamentnilisty	https://parlamentnilisty.cz	aktivní
26	Plzenskenovinky	http://plzenskenovinky.cz	aktivní
27	Pokec24	https://pokec24.cz	aktivní
28	Pravyprostor	https://pravyprostor.cz	aktivní
29	Prvnizpravy	https://prvnizpravy.cz	aktivní
30	Pravdive	https://pravdive.eu	aktivní
31	Protiproud	https://protiproud.info	aktivní
32	Prazskenovinky	http://prazskenovinky.cz	aktivní
33	Raptor-tv	https://raptor-tv.cz	aktivní
34	Regionalninovinky	http://regionalninovinky.cz	aktivní
35	Stredoceskenovinky	http://stredoceskenovinky.cz	aktivní
36	Tadesco	https://tadesco.org	aktivní
37	Tydenikobcanskepravo	https://tydenikobcanskepravo.cz	aktivní
38	Usteckenovinky	http://usteckenovinky.cz	aktivní
39	Veksvetla	https://veksvetla.cz	aktivní
40	Vipnoviny	https://vipnoviny.cz	aktivní
41	Zlinskenovinky	http://zlinskenovinky.cz	aktivní
42	Zvedavec	https://zvedavec.news	aktivní

Tabulka C.2: Tabulka dezinformačních webových stránek.

Příloha D

Plakát

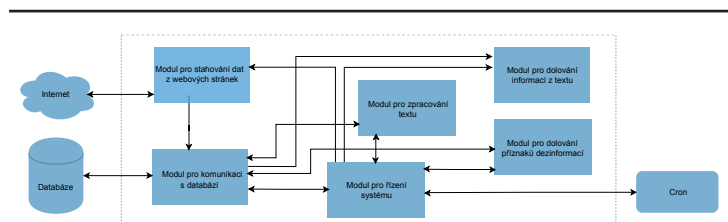


ANALÝZA TÉMAT A ŠÍŘENÍ DEZINFORMACÍ Z PROPAGANDISTICKÝCH WEBOVÝCH STRÁNEK

ANALYSIS OF TOPICS AND SPREADING OF DISINFORMATION
FROM PROPAGANDIST WEB PAGES

Tato práce se zabývá problematikou dezinformací na propagandistických webových stránkách v České republice a návrhem systému, který by měl sloužit k analýze a rozpoznání dezinformačního charakteru obsahu těchto stránek. Navrhovaný a implementovaný systém dokáže stahovat a archivovat data z propagandistických webů a následně analyzovat jejich obsah. Systém využívá především znalostní a síťové metody. Cílem práce je využít moderních technologií pro odhalování dezinformací v médiích. Výsledky práce mohou pomoci při dalším výzkumu a při prevenci šíření dezinformací.

This thesis deals with the issue of disinformation on propaganda websites in the Czech Republic and the design of a system that should be used to analyze and recognize the disinformation character of the content of these websites. The proposed and implemented system is able to download and archive data from propaganda websites and then analyze their content. The system mainly uses knowledge-based and network analyses. The aim of the thesis is to use modern technologies for detecting disinformation in the media. The results may help in further research and in preventing the spread of disinformation.



Fenómén dezinformací v posledních letech ve společnosti rezonuje čím dál intenzivněji. Bývá zmiňován zejména v souvislosti s možným ovlivňováním politického vývoje a nejen v českém prostředí v souvislosti s rozdělováním společnosti na více vyhraněných názorových skupin. Faktem zůstává, že dezinformace byly, jsou a zřejmě i nadále budou součástí lidské společnosti, otázkou zůstává, jakým způsobem nás budou v životech provázet a ovlivňovat a jak jim budeme schopni čelit.

AUTOR
VEDOUcí PRÁCE
BRNO 2023

DALIBOR KYJOVSKÝ
Ing. KAREL ONDŘEJ

Obrázek D.1: Náhled vytvořeného plakátu.