

Využití umělé inteligence při rozhodování o žádostech o azyl

Diplomová práce

Vedoucí práce:

Mgr. et Mgr. Robert Stojanov, Ph.D.

Bc. Kateřina Černá

Brno 2024

Poděkování

Chtěla bych poděkovat svému vedoucímu diplomové práce Mgr. et Mgr. Robert Stojanov, Ph.D. za odbornou pomoc, cenné rady a náměty. Rovněž bych ráda poděkovala pedagogům na univerzitě, kteří významně přispěli k mému vzdělání. Nicméně bych chtěla vyjádřit hlubokou vděčnost mé rodině, která mi vždy poskytovala nezbytnou podporu a umožnila mi věnovat se dlouhodobému studiu a úspěšnému dokončení studia. Bez jejich podpory bych tento cíl nedosáhla.

Čestné prohlášení

Prohlašuji, že jsem práci Využití umělé inteligence při rozhodování o žádostech o azyl vypracovala samostatně a veškeré použité prameny a informace uvádím v seznamu použité literatury. Souhlasím, aby moje práce byla zveřejněna v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách ve znění pozdějších předpisů a v souladu s platnou Směrnicí o zveřejňování závěrečných prací.

Jsem si vědoma, že se na moji práci vztahuje zákon č. 121/2000 Sb., autorský zákon, a že Mendelova univerzita v Brně má právo na uzavření licenční smlouvy a užití této práce jako školního díla podle § 60 odst. 1 autorského zákona.

Dále se zavazuji, že před sepsáním licenční smlouvy o využití díla jinou osobou (subjektem) si vyžádám písemné stanovisko univerzity, že předmětná licenční smlouva není v rozporu s oprávněnými zájmy univerzity, a zavazuji se uhradit případný příspěvek na úhradu nákladů spojených se vznikem díla, a to až do jejich skutečné výše.

V Brně dne 10. května 2024

.....

Abstract

Asylum decision-making is a critical process with a high level of responsibility that affects the lives of many individuals and the integrity of immigration systems. This research explores the potential of using artificial intelligence (AI) to enhance the efficiency and objectivity of the asylum decision-making process. The advantages of AI, particularly machine learning, enable the analysis of extensive and diverse data, contributing to predictive assessments of the likelihood of a successful asylum application. However, the deployment of AI presents technical and ethical challenges, including transparency, accountability, and the risk of bias. The aim of the thesis is to examine the use of artificial intelligence in asylum decision-making, analyze the current state, and provide recommendations for further development. Key questions include the utilization of AI, the advantages, and disadvantages of this approach. Literature review and analysis of the benefits and limitations of AI in asylum proceedings will lead to a comprehensive understanding of this issue.

Keywords

artificial intelligence, asylum proceedings, decision-making, migration, machine learning

Abstrakt

Rozhodování o azylu je kritickým procesem s vysokou mírou zodpovědnosti, který ovlivňuje životy mnoha jednotlivců a integritu imigračních systémů. Tento výzkum se zabývá potenciálem využití umělé inteligence (AI) k posílení efektivity a objektivity rozhodovacího procesu v oblasti azylu. Přednosti AI, zejména strojového učení, umožňují analýzu rozsáhlých a různorodých dat, což přispívá k prediktivnímu posuzování pravděpodobnosti úspěšné žádosti o azyl. Nicméně, nasazení AI přináší technické a etické výzvy, včetně transparentnosti, odpovědnosti a rizika předpojatosti.

Cílem diplomové práce je prozkoumat využití umělé inteligence v rozhodování o azylu, analyzovat současný stav a poskytnout doporučení pro další rozvoj. Klíčové otázky zahrnují způsob využití AI, výhody a nevýhody tohoto přístupu. Literární rešerše a analýza přínosů a omezení AI v azylových řízeních povedou k komplexnímu porozumění této problematice.

Klíčová slova

umělá inteligence, azylová řízení, rozhodování, migrace, strojové učení

Obsah

1 Úvod a cíl práce	14
1.1 Předmluva	14
1.2 Úvod	14
1.3 Cíl práce.....	15
2 Literární rešerše.....	16
2.1 Vymezení základních pojmů	16
2.2 Analýza procesu azylových řízení.....	18
2.3 Využití umělé inteligence	20
2.4 Využití umělé inteligence v procesu rozhodování o žádostech o azyl	26
2.4.1 Využití umělé inteligence v procesu azylových řízení.....	28
2.4.2 Kritická analýza stávajícího využití umělé inteligence	30
2.5 Analýza dat současného stavu rozhodování o žádostech o azyl.....	31
2.5.1 Analýza dat o migraci.....	31
2.5.2 Přehled počtu žádosti o azyl a rozhodnutí v EU.....	32
2.6 Shrnutí literární rešerše.....	33
2.6.1 Shrnutí klíčových poznatků z literatury	33
2.6.2 Identifikace mezery ve stávajících výzkumech a potenciální oblasti pro další analýzu	34
3 Metodologie.....	35
3.1 Popis metodologických postupů	35
3.2 Technologie	36
3.3 Metody	37
4 Analýza dat a implementace.....	40
4.1 Sběr a analýza dat o žadatelích o azyl	40
4.1.1 Sběr dat	41
4.1.2 Struktura dat.....	42
4.1.3 Důvody k zamítnutí žádostí	45
4.2 Návrh algoritmu pro rozhodování o žádostech o azyl.....	45

4.2.1	Struktura projektu.....	46
4.2.2	Implementace	46
4.3	Analýza a vizualizace výsledků.....	50
4.4	Zhodnocení výsledků	62
5	Diskuse.....	64
5.1	Etická stránka problému	64
5.2	Možnosti dalšího výzkumu v této oblasti.....	66
6	Závěr	68
7	Literatura	69
7.1	Seznam citací	69
8	Přílohy.....	74

Seznam obrázků

Obr. 1	Státy s největším počtem žadatelů o azyl v EU.....	33
Obr. 2	Naivní Bayesovi algoritmy	38
Obr. 3	K-Means algoritmy.....	38
Obr. 4	K-Nearest neighbor algoritmus.....	39
Obr. 5	Rozhodovací strom.....	39
Obr. 6	Úspěšnost jednotlivých žádostí v Sýrii	52
Obr. 7	Četnost slov žádostí v Sýrii	53
Obr. 8	Úspěšnost jednotlivých žádostí v Iráku.....	55
Obr. 9	Četnost slov žádostí v Iráku.....	56
Obr. 10	Úspěšnost jednotlivých žádostí v Tunisku.....	57
Obr. 11	Četnost slov žádostí v Tunisku	58
Obr. 12	Úspěšnost jednotlivých žádostí v Afghánistánu	60
Obr. 13	Četnost slov žádostí v Afghánistánu	61

Seznam tabulek

Tab. 1	Počet žádostí o azyl od roku 2013 do roku 2023.....	40
Tab. 2	Výsledky klasifikace – Sýrie	51
Tab. 3	Výsledky klasifikace – Irák.....	54
Tab. 4	Výsledky klasifikace – Tunisko	57
Tab. 5	Výsledek klasifikace – Afghánistán	59
Tab. 6	Souhrnný přehled výsledků úspěšnosti algoritmu	62

1 Úvod a cíl práce

1.1 Předmluva

Tato práce má za výsledek informační cvičení, ve kterém jsem aplikovala znalosti nabyté během studia na vysoké škole Mendelovy univerzity v Brně na magisterském oboru Otevřená informatika.

Je důležité zdůraznit, že tento projekt není reálnou aplikací pro vyhodnocování žádostí o azyl, nýbrž analytickým cvičením. Omezení tohoto cvičení spočívá v absenci všech proměnných, které by odpovídaly skutečnému procesu vyhodnocování žádostí. Můj přístup k této problematice je tedy především analytický a nemá za cíl simulovat komplexní realitu rozhodovacího procesu v této oblasti. Během práce jsem se snažila integrovat poznatky z oblasti umělé inteligence s problematikou azylových žádostí, a to s ohledem na současný vývoj a možnosti technologií.

1.2 Úvod

Řízení migrace a rozhodování o azylu jsou složité procesy, které vyžadují účinné a přesné posuzování, aby byla zajištěna ochrana práv jednotlivců a integrita imigračních systémů. V posledních letech roste zájem o využití technologií umělé inteligence (artificial intelligence – AI) ke zvýšení efektivity, objektivitu a prediktivních schopností rozhodovacích procesů o azylu.

Integrace algoritmů umělé inteligence, především techniky strojového učení, se ukázaly jako slibné pro automatizaci a zefektivnění různých aspektů rozhodovacího procesu v oblasti azylu. Tyto algoritmy mohou analyzovat velké objemy heterogenních dat s údaji týkajícími se azylu, včetně informací o zemi původu, socioekonomických faktorech a historických výsledcích azylového řízení, a předpovídat tak pravděpodobnost úspěšné žádosti o azyl. Využitím umělé inteligence mohou rozhodovací orgány potenciálně zvýšit objektivitu a efektivitu svých posouzení, což povede ke spolehlivějším a konzistentnějším výsledkům.

Nasazení umělé inteligence při rozhodování o azylu však vyvolává důležité technické a etické otázky. Složitost a nepřehlednost některých algoritmů strojového učení může ztěžovat interpretaci a vysvětlování důvodů, které stojí za automatizovanými rozhodnutími, což představuje potenciální problémy s transparentností a odpovědností. Kromě toho existuje riziko předpojatosti a diskriminace, pokud jsou tréninková data použita k vývoji modelů umělé inteligence nereprezentativní nebo odrážejí existující předsudky. Zajištění spravedlnosti a robustnosti systémů AI v této oblasti vyžaduje pečlivou pozornost věnovanou kvalitě dat, transparentnosti algoritmů a průběžnému monitorování a vyhodnocování.

1.3 Cíl práce

Hlavním cílem mé diplomové práce je pečlivě prozkoumat a analyzovat využití umělé inteligence v procesu žádostí o azyl. Mým záměrem je provést analýzu současného stavu a poskytnout stručný popis průběhu azylových řízení ve vybraných zemích. Zaměřím se zejména na analýzu stávajícího využití umělé inteligence v procesech migrace obyvatel, s důrazem na specifika azylových řízení.

Pro dosažení tohoto cíle byly formulovány následující výzkumné otázky: Jakým způsobem je umělá inteligence využívána v procesech rozhodování o udělení azylu? Jaké jsou výhody a nevýhody používání umělé inteligence v těchto procesech? Na základě těchto otázek bude provedena důkladná analýza a následně vypracována doporučení pro další využití umělé inteligence v oblasti azylových řízení.

V rámci práce vypracuji literární rešerši analyzující využití umělé inteligence v oblasti azylových řízení. Provedu analýzu výhod a nevýhod implementace umělé inteligence v azylových řízeních a zhodnotím výsledky. Zaměřím se na konkrétní aspekty procesu rozhodování o žádostech o azyl, které jsou relevantní pro využití umělé inteligence.

2 Literární rešerše

Tato část práce představuje klíčový teoretický rámec a kontext pro hlubší porozumění analýzy využití umělé inteligence (AI) v procesu rozhodování o žádostech o azyl. Začnu vymezením pojmů, které poskytnou lepší vhled do kontextu.

2.1 Vymezení základních pojmů

Migrace

„Migrace je jednou z forem prostorové mobility obyvatelstva. Tou rozumíme pohyb obyvatelstva v geografickém prostoru z důvodu uspokojování potřeb v nejširším významu tohoto pojmu. Prostorová mobilita není pouhým fyzickým přemístěním, ale je to též proces výběru místa spojený s emocionální vazbou k objektům či k prostorovým celkům.“ (Drbohlav, 1999)

„Lidská migrace je pohyb lidí z jedné oblasti do druhé s úmyslem se v nové oblasti dočasně či trvale usadit. Pohyb lidí obvykle probíhá na dlouhé vzdálenosti, přičemž překračuje hranice jednotlivých států, ale může zahrnovat i vnitřní migraci. Migrace může být dočasná nebo trvalá a mohou se jí účastnit jednotlivci, rodiny nebo větší skupiny. Součástí migrace mohou být různé typy migrace, jako je ekonomická migrace, uprchlická migrace, nebo migrační saldo, které vyjadřuje rozdíl mezi počtem přistěhovalých a vystěhovalých osob v dané oblasti. Migrace může být také vnitřní, kdy se lidé přesunují v rámci jednoho státu, nebo mezinárodní, kdy dochází k přesunu mezi různými státy.“ (Henig, 2007)

„Celkový počet migrantů je uváděn jako migrační objem, který je součtem počtu přistěhovalých a vystěhovalých v dané oblasti. Migrace může být krátkodobá, dlouhodobá, trvalá i opakovaná (cirkulární).“ (OAMP, 2024)

Podle Evropského parlamentu mohou být důvody pro lidskou migraci rozděleny do těchto sekcí (Evropský parlament, 2020):

- „*Politické důvody*: Lidé mohou opouštět svou zemi kvůli politickému pronásledování, válce nebo konfliktu.“
- „*Ekonomické důvody*: Hledání lepších pracovních příležitostí, vyšších mezd nebo lepšího ekonomického prostředí může být také důvodem pro migraci.“
- „*Sociální důvody*: Mezi sociální důvody patří hledání lepšího vzdělání, sjednocení rodin nebo útěk před chudobou.“
- „*Environmentální důvody*: Změna klimatu a životní prostředí může také vést k migraci, například kvůli přírodním katastrofám nebo nedostatku zdrojů.“

Uprchlík

Podle definice obsažené v Ženevské úmluvě z roku 1951 je uprchlíkem osoba, která "v důsledku oprávněných obav ze stíhání pro svoji rasu, náboženství, národnost, členství ve skupině sdílející určité politické názory, se nachází mimo území státu, jehož je občanem nebo kde se nacházelo bydliště, a nemůže či v důsledku těchto

obav nechce se užívat ochrany této země" (Ženevská úmluva, 1951, článek 1A (2)). Tato definice tvoří základ pro poskytování azylu v rámci mezinárodního práva.“

„Jejich rozhodnutí překročit státní hranice a hledat bezpečí v okolních zemích je motivováno nutností ochrany v životu ohrožujících nebezpečí ve své domovině. Tyto osoby jsou pak mezinárodně uznány za uprchlíky a přísluší jim přístup k humanitární pomoci od států, UNHCR a dalších relevantních organizací. Jejich status uprchlíka je potvrzen zejména proto, že návrat do jejich původní země by pro ně představovalo nebezpečí a ohrožení života. Odepření azylu by v jejich případech mohlo mít fatální následky, protože jsou postaveni před skutečností, že nemohou bezpečně a volně žít ve své mateřské zemi.“ (Dobrovolník, 2023)

Podle Jelínkové lze uprchlíky rozdělit do následujících kategorií (Jelínková, 2000):

- „*Statutární uprchlík* – je uprchlík, který splňuje kritéria Ženevské úmluvy z roku 1951 a Protokolu z roku 1967“
- „*Uprchlík de facto* – je uprchlík, který nemůže nebo nechce získat postavení statutárního uprchlíka“
- „*Uprchlík in orbit* – je uprchlík, kterému je stále odmítán azyl, nebo který nemůže nalézt zemi ochotnou posoudit jeho žádost, přestože nebyl vrácen přímo do země, kde mu hrozí pronásledování“
- „*Uprchlíci prima facie* – jsou uprchlíci, o jejichž uprchlickém statutu bylo rozhodnuto v rámci větší skupiny v situaci rozsáhlého přílivu, aniž by byly žádosti posouzeny jednotlivě“
- „*Mandátoví uprchlíci* – jsou uprchlíci, kteří spadají do kompetence UNHCR podle mezinárodního uprchlického práva, Statutu UNHCR a rezolucí Valného shromáždění OSN“
- „*Uprchlík „sur place“* - je osoba, která v okamžiku odchodu ze země nebyla uprchlíkem, stala se jím až později v důsledku okolností nastalých v době jeho nepřítomnosti“

Azyl

„Ochranný pobyt, který stát poskytuje státnímu příslušníku třetí země nebo osobě bez státní příslušnosti v souvislosti s jejím pronásledováním zpravidla z důvodů politických či humanitárních.“ (Dobrovolník, 2023) Azyl představuje právo nebo ochranu poskytovanou jednotlivci, kteří opustili svou původní zemi v důsledku oprávněných obav před pronásledováním. Toto právo je spojeno s mezinárodním humanitárním právem a slouží k zajištění bezpečného útočiště pro jedince, kteří čelí život ohrožujícím podmínkám ve svém domově. Získání azylu v jiné zemi umožňuje jednotlivcům uniknout nebezpečí a hledat bezpečí, kde mají právo na ochranu před dalším pronásledováním. (Dobrovolník, 2023)

Žadatel o azyl

Žadatel o azyl je osoba, která požádala o mezinárodní ochranu a žádá o udělení azylu v zemi, kde se nachází. Tento termín se často zaměňuje s termíny "uprchlík" nebo "azylant", ale sám o sobě neznamená automatické udělení statusu uprchlíka. Žadatel o azyl může být kdokoli, kdo tvrdí, že je uprchlíkem a žádá o mezinárodní ochranu před pronásledováním nebo nebezpečím vážné újmy v zemi. Každý žadatel o azyl je nejprve žadatelem, ale ne každý žadatel bude nakonec uznán jako uprchlík. O nároku na udělení statusu uprchlíka rozhoduje vnitrostátní právní úprava dané země. Efektivita azylového systému má klíčový význam, a pokud je systém rychlý, spravedlivý a účinný, každý žadatel má možnost získat ochranu, pokud mu náleží. (UNHRC, 2016)

2.2 Analýza procesu azylových řízení

Azylové řízení

Podle Jany Filipové je azylové řízení důležitým procesem, který zajišťuje ochranu a pomoc lidem hledajícím azyl. Mezinárodní právo chránící práva uprchlíků existuje ve formě Úmluvy o právním postavení uprchlíků z roku 1951 a jejího Protokolu z roku 1967. Tato úmluva stanoví, že uprchlíci jsou chráněni a mají přístup k pomoci států, UNHCR a příslušných organizací. Azylové řízení je tedy klíčovým prvkem ochrany a poskytování mezinárodní ochrany lidem, kteří jsou nuceni opustit svou rodnou zemi nebo stát kvůli různým důvodům, jako je represivní režim, válečná situace nebo ekonomické podmínky. (Filipová, 2008)

„Azylové řízení je proces, který provádí žadatelé o azyl, kteří požádali o mezinárodní ochranu v zemi, do které přišli, aby unikli pronásledování nebo jiným nebezpečným situacím ve své domovské zemi. Toto řízení má za cíl posoudit, zda žadatelé splňují kritéria pro udělení azylu a tím umožnit jim legálně zůstat v této zemi a získat ochranu a podporu.“ (EASO, 2016)

V rámci azylového řízení jsou žadatelé o azyl obvykle podrobeni procesu, který zahrnuje prověření jejich žádosti, výslech a posouzení, zda mají důvody pro udělení azylu. Výsledkem azylového řízení může být udělení azylu, což znamená, že žadatel může legálně zůstat v zemi a získat ochranu, nebo zamítnutí žádosti, což může vést k deportaci žadatele do jeho domovské země. (EASO, 2016)

„Azylové řízení má klíčový význam pro ochranu lidských práv a zajištění bezpečnosti těch, kteří utekli před pronásledováním a nebezpečím. Je také důležitou součástí mezinárodního práva a lidských práv, a jeho regulace a normy se vyvíjely v průběhu historie v souladu s mezinárodními událostmi a změnami ve světové migraci.“ (Evropský parlament, 2020)

Azylové řízení v České republice

Azylové řízení je proces, kterým prochází žadatelé o mezinárodní ochranu, kteří hledají azyl v cizí zemi. „V České republice je azylové řízení upraveno zákonem o azylu, který stanoví postup udělení mezinárodní ochrany.“ (Zákon pro lidi, 1999) Žádost

o udělení mezinárodní ochrany získává žadatel právo zůstat na území České republiky. Řízení může být ukončeno rozhodnutím ve věci, zamítnutím žádosti nebo zastavením řízení. V případě udělení azylu obdrží azylant průkaz povolení k pobytu, který mu umožňuje trvale pobývat v České republice. (OAMP, 2018).

Azylové řízení představuje klíčový proces, který poskytuje ochranu a podporu lidem, kteří hledají azyl. V případě, že žadatelé o mezinárodní ochranu nenalezli spravedlnost ve svých právech, mají možnost požádat o soudní přezkoumání své žádosti o udělení mezinárodní ochrany (Hrdlička, 2021).

Podle informací poskytnutých Ministerstvem vnitra České republiky, po podání žádosti o mezinárodní ochranu je žadatel obvykle přemístěn do pobytového střediska, kde čeká na rozhodnutí v prvním stupni. V případě, že je azyl udělen, azylant obdrží průkaz povolení k pobytu, který mu umožní trvale žít v České republice. Žadatelé o mezinárodní ochranu mají právo na bezplatné ubytování, stravu a základní zdravotní péči, za předpokladu, že pobývají v přijímacím či pobytovém středisku (OAMP, 2023a).

Zákon o azylu v České republice stanovuje, že ministerstvo může udělit azyl cizincům, pokud je v průběhu řízení o udělení mezinárodní ochrany zjištěn důvod pro udělení azylu. Dále ministerstvo může udělit humanitární azyl, pokud nejsou splněny podmínky pro udělení azylu, ale existují zvláštní důvody pro udělení azylu z humanitárního důvodu (OAMP, 2023a). Tato ustanovení jsou obsažena v zákoně č. 325/1999 Sb., o azylu a o změně zákona č. 283/1991 Sb.

Žadatelé o azyl v České republice mají podle zákona několik práv a povinností. K nim patří právo na jednání ve svém jazyce během řízení, nárok na bezplatnou základní zdravotní péči, ubytování, stravu a kapesné, právo na zastupování během řízení, právo na pomoc od fyzických nebo právnických osob specializujících se na poskytování pomoci uprchlíkům a další (OAMP, 2023a).

Popis procesu azylového řízení

Postup při azylovém řízení představuje komplexní a právně stanovený proces, jenž má za cíl posoudit, zda cizinec má nárok na mezinárodní ochranu v rámci České republiky. Azylové řízení vychází z národního zákona o azylu a relevantních mezinárodních právních předpisů, s ohledem na zásadu zachování lidské důstojnosti a základních práv jednotlivce, zejména v situacích, kdy je vystaven pronásledování či vážné újmě v zemi původu. (OAMP 2023b)

Průběh azylového řízení může být rozdělen do několika základních fází, a to následovně:

1. **Podání žádosti o udělení mezinárodní ochrany:** Cizinec, který hledá ochranu před pronásledováním nebo vážnou újmu, musí formálně podat žádost o mezinárodní ochranu. Tato žádost může být podána pouze na území České republiky na příslušných místech, včetně hraničních přechodů, přijímacích středisek, nebo u krajských ředitelství policie, pokud cizinec přijde dobrovolně. Žádost se zahajuje projevem vůle cizince hledat ochranu.

2. **Identifikace a sběr biometrických dat:** Po podání žádosti probíhá identifikace a sběr biometrických dat, včetně otisků prstů a obrazového záznamu. Cizinec musí rovněž odevzdat cestovní doklad a podstoupit lékařské vyšetření.
3. **Zajištění v přijímacím středisku:** Po prvních krocích může být žadatel přemístěn do přijímacího střediska, kde čeká na rozhodnutí. Zde může také požádat o možnost bydlení na soukromé adrese na vlastní náklady.
4. **Podání detailních informací o žádosti:** Žadatel je povinen poskytnout detailní informace týkající se důvodů, proč opustil svou zemi a proč žádá o mezinárodní ochranu. Tyto informace jsou klíčové pro rozhodování.
5. **Pohovory a další důkazní prostředky:** Žadatel má možnost vést pohovory s úředníky, během nichž může blíže specifikovat důvody svého žádosti. Má také právo předložit další důkazní prostředky, které podporují jeho tvrzení.
6. **Právo na jazykovou podporu a právní pomoc:** Cizinec má právo jednat v mateřském jazyce nebo jazyce, ve kterém je schopen se dorozumět. Tlumočnick je k dispozici pro úkony, ke kterým byl předvolán úřadem. Žadatel může přizvat dalšího tlumočnicka na své náklady. Právní pomoc může poskytnout další osoba nebo advokát.
7. **Rozhodnutí v prvním stupni:** Ministerstvo vnitra musí rozhodnout o žádosti do 6 měsíců od poskytnutí detailních údajů. Tuto lhůtu lze v zákonem stanovených případech prodloužit. Rozhodnutí nabývá právní moci po doručení žadateli.
8. **Ukončení azylového řízení:** Azylové řízení může být ukončeno pozitivně, pokud je žadateli udělena mezinárodní ochrana, nebo negativně, pokud je jeho žádost zamítnuta. V případě zamítnutí má žadatel možnost podat odvolání.
9. **Odvolací řízení:** V případě zamítnutí může žadatel podat odvolání k Ministerstvu vnitra, které je nezávislým orgánem, nebo k soudu. Odvolání má odkladný účinek na vystěhování žadatele.
10. **Výkon rozhodnutí:** Pokud je rozhodnutí kladné, žadatel získá status uprchlíka nebo doplňkovou ochranu a má právo na pobyt v České republice. V případě negativního rozhodnutí může být přinucen opustit území České republiky.

Celý proces azylového řízení je pečlivě regulován právními předpisy a je založen na zásadách spravedlnosti a lidských práv. Žadatelé mají právo na přístup k právní pomoci a možnost odvolat se proti nepříznivým rozhodnutím, aby mohli efektivně prosazovat svá práva v souladu s mezinárodním právem. (OAMP, 2023b)

2.3 Využití umělé inteligence

Historie umělé inteligence

Historie umělé inteligence má kořeny až v antických dobách, kdy filozofové zkoumali povahu lidské mysli. Avšak skutečný základ pro rozvoj umělé inteligence byl

položen v první polovině 20. století s vývojem programovatelných strojů a prvními počítači (Kod'ousková, 2023). V roce 1943 nastal zlomový okamžik v historii umělé inteligence s vytvořením prvního matematického modelu neuronu, který položil základy pro další rozvoj této oblasti (Lukáš, 2024).

V padesátých a šedesátých letech byl vytvořen Turingův test, který se stal klíčovým pro teorii strojového učení. Tento test měl určit, zda lze stroje považovat za skutečně inteligentní, a dodnes zůstává jednou z tradičně používaných metodik v oblasti umělé inteligence. Během tohoto období byly představeny první produkty umělé inteligence, jako například počítač Arthura Samuela, schopný simulovat hru Dáma, a systém od IBM, který vynikal v překladu jazyků. (Kod'ousková, 2023) Současně se začaly objevovat prvotní aplikace umělé inteligence v oblasti medicíny a překladů. V této éře došlo také k průniku mezi matematikou a elektronikou a k využití termínu "umělá inteligence", a to včetně her jako piškvorky a stroje MENACE (Lukáš, 2024).

V sedmdesátých letech došlo k útlumu v oblasti umělé inteligence z důvodu nedostatku finančních a technických zdrojů. Nicméně v devadesátých letech se objevily chatboti, jako například A.L.I.C.E., a japonská společnost Sony představila domácího mazlíčka AIBO. V roce 1996 porazil šachový mistr Garry Kasparov počítač Deep Blue ve šestizápasovém duelu výsledkem 4:2. (Lukáš, 2024)

V roce 2005 se objevil první autonomní automobil Stanley, zahajující éru samostatně řídicích vozidel. V té době také společnosti Apple a Microsoft představily technologické inovace v oblasti rozpoznávání řeči a pohybového ovládání.

Druhá dekáda 21. století přinesla nástup virtuálních asistentů, jako jsou Siri, Cortana a Alexa, spolu s klíčovým rozvojem big data, analýzy uživatelského chování a strojového učení. V této době došlo k významným událostem v oblasti umělé inteligence, včetně trénování neuronových sítí k rozpoznávání obrazů (2012), protestů proti autonomním zbraním (2015), vytvoření humanoidního robota Sophii (2016), vzniku vlastního jazyka mezi AI chatboty na Facebooku (2017), překonání lidského intelektu v čtení a porozumění textům umělou inteligencí od Alibaba (2018), získání titulu velmistra AI AlphaStar ve videohře StarCraft 2 (2019), spuštění beta testování GPT-3 (2020) a vývoj AI DALL-E schopné produkovat popisky k obrázkům (2021) (Tableau, 2024).

Dnešní umělá inteligence se stále více integruje do běžného lidského života, s populárními asistentkami a rostoucím zájmem a investicemi v podnikovém sektoru i v rámci start-upů (Kod'ousková, 2023).

Umělá inteligence

„Umělá inteligence neboli Artificial Intelligence (AI) označuje vývoj počítačových systémů schopných vykonávat úkoly, které obvykle vyžadují lidskou inteligenci.“ (Russell, S., a Norvig, P. 2016) Mezi tyto úkoly patří rozpoznávání řeči, řešení problémů, učení, rozhodování a rozpoznávání vzorů. Umělá inteligence zahrnuje řadu technik a přístupů, včetně strojového učení, zpracování přirozeného jazyka, počítačového vidění a robotiky. (Russell, S., a Norvig, P. 2016)

Cílem umělé inteligence je napodobit a rozšířit lidskou kognitivní schopnost a umožnit strojům analyzovat obrovská množství dat, provádět predikce a automatizovat složité procesy. Systémy se učí zkušeností a upravují své chování na základě zpětné vazby. Využití umělé inteligence je široké a zahrnuje odvětví jako zdravotnictví, finance, dopravu, výrobu a zábavu. (Russell, S., a Norvig, P. 2016)

Koncept umělé inteligence vznikl v polovině 20. století a někteří vědci, jako Alan Turing a John McCarthy, přispěli k jejímu počátečnímu rozvoji (McCarthy, Minsky, Rochester, a Shannon, 2006). V posledních letech však došlo k významnému pokroku díky zlepšení výpočetního výkonu, dostupnosti velkého množství dat a průlomovým algoritmům.

Druhy umělé inteligence:

Typy umělé inteligence (AI) představují různé úrovně schopností a potenciálu, kterými tyto systémy disponují.

Podle společnosti SAP (2023) se dělí do těchto 3 částí:

- **Umělá úzká inteligence (ANI):** „Tato kategorie AI je momentálně přítomná a známá jako "slabá" inteligence. ANI je schopna řešit specifické úkoly, jako je rozpoznávání obličejů, internetové vyhledávání nebo autonomní řízení aut. I když je schopná využívat složité algoritmy a neuronové sítě, je omezená na specifické úkoly a chybí jí celková inteligence, jakou přisuzujeme lidskému myšlení.“ (SAP, 2023)
- **Umělá obecná inteligence (AGI):** „Tento druh umělé inteligence by měl být schopen provádět jakýkoliv intelektuální úkol, který zvládá člověk. AGI má schopnost generalizovat a extrapolovat informace přes různé úkoly a situace, což přesahuje schopnosti ANI.“ (SAP, 2023)
- **Umělá superinteligence (ASI):** „Tato teoretická kategorie AI by měla být plně sebevědomá a schopná chápat lidské chování na základní úrovni. ASI se zdá být dystopickým představením budoucnosti, kde by tyto systémy mohly předčit lidskou inteligenci ve všech směrech.“ (SAP, 2023)

V současné době se nejvíce využívá úzká umělá inteligence. Učení s učitelem a učení bez učitele spadají do kategorie umělé úzké inteligence (ANI). Učení se s učitelem a učení bez učitele jsou dvě základní metody v oblasti strojového učení, každá s odlišným přístupem k výuce algoritmů. (SAP, 2023)

Učení s učitelem

„Strojové učení s učitelem stanovuje jasně definované vstupní a výstupní parametry.“ Vědec či učitel poskytuje algoritmu informace, například identifikuje auto a ukazuje, jak s touto informací pracovat. Následně se výstup porovnává s očekáváním, využívající bohaté datové sady obsahující různé varianty objektu. Algoritmus se tak učí rozeznávat a pracovat s těmito variantami. (Kod'ousková, 2021)

Když je konfrontován s novými daty, algoritmus je řízen podle daných parametrů a neustále upravuje své znalosti. Strojové učení pomocí této metody ovládá

pouze to, co mu bylo explicitně předáno a nemůže provádět činnosti, které nebyly předem definovány. (Kod'ousková, 2021)

Učení bez učitele

Na rozdíl od učení s učitelem definuje metoda učení bez učitele pouze vstupní data. „Algoritmus se sám snaží nalézt postup, řešení a výsledek pomocí principu pokus – omyl. Tato metoda využívá shlukování dat, kde algoritmus hledá podobnosti a spojitosti mezi jednotlivými objekty a přiřazuje jim vlastní označení.“ (Kod'ousková, 2021)

Shlukování dat napomáhá vytvářet komplexní informace, což se v praxi využívá k tvorbě personalizovaných nabídek služeb nebo zboží. Příkladem může být Netflix, který analyzuje historii sledování uživatelů a na základě těchto dat doporučuje další pořady ke zhlédnutí. Obě metody strojového učení, bez učitele i s učitelem, mohou být také kombinovány pro dosažení komplexnějších výsledků. (Kod'ousková, 2021)

Strojové učení

„Strojové učení je podoblast umělé inteligence, která se zaměřuje na vývoj algoritmů a modelů umožňujících počítačovým systémům učit se a provádět předpovědi nebo rozhodnutí, aniž by byly explicitně naprogramovány.“ (Jordan a Mitchell, 2015) Zahrnuje použití statistických technik a matematických modelů k analýze a interpretaci vzorců v datech s cílem umožnit strojům zlepšit jejich výkon nebo chování v průběhu času.

Jedním ze základních konceptů strojového učení je proces trénování, kdy jsou modely vystaveny označeným nebo neoznačeným datům, aby se učily z příkladů a prováděly přesné předpovědi nebo klasifikace. Modely jsou poté vyhodnocovány a zdokonalovány na základě jejich výkonnosti pomocí metrik, jako je přesnost, odvolání nebo skóre F1. „Prostřednictvím tohoto opakujícího se procesu se algoritmy strojového učení mohou přizpůsobovat a zlepšovat svou schopnost zpracovávat nová a neznámá data.“ (Jordan a Mitchell, 2015)

Existují různé typy algoritmů strojového učení, včetně učení s učitelem, učení bez učitele a učení s posilováním. Učení s učitelem zahrnuje trénování modelů pomocí označených dat, kdy se algoritmus učí mapovat vstupní data na odpovídající výstupní označení. Učení bez učitele se naproti tomu zabývá neoznačenými daty a jeho cílem je objevit vzory, vztahy nebo struktury v datech. Učení s posilováním se zaměřuje na trénování agentů, aby dělali postupná rozhodnutí na základě zpětné vazby z prostředí.

„Strojové učení se uplatňuje v různých oblastech, včetně rozpoznávání obrazu a řeči, zpracování přirozeného jazyka, doporučovacích systémů a autonomních vozidel.“ Jeho potenciál pro transformaci průmyslových odvětví a společnosti je značný, protože umožňuje počítačům analyzovat a interpretovat složitá data, automatizovat úlohy a poskytovat inteligentní poznatky. (Jordan a Mitchell, 2015)

Hluboké učení

Umělá inteligence (AI) a hluboké učení jsou dva úzce propojené koncepty. Hluboké učení (Deep Learning) je disciplínou strojového učení, které využívá umělé neuronové sítě k učení a zpracování složitých úkolů. Tento přístup se zaměřuje na vytvoření modelů, které dokáží hierarchicky zpracovávat informace a automaticky extrahovat vysokoúrovňové reprezentace dat.

Článek Bernarda Marra (Marr 2020) poskytuje komplexní přehled o hlubokém učení (Deep Learning AI) a jeho praktických aplikacích. Marr se snaží odstranit nejasnosti kolem této technologie a zdůraznit její důležitost v dnešním technologickém prostředí.

Podstatou hlubokého učení je využití umělých neuronových sítí, které se učí z rozsáhlých souborů dat a plní úkoly, které obvykle vyžadují lidskou inteligenci. Jak uvádí Marr, "Hluboké učení je podmnožinou strojového učení, kdy se umělé neuronové sítě učí z rozsáhlých souborů dat plnit úkoly, které vyžadují lidskou inteligenci" (Marr, 2020). Citace připomíná, že kvalitní data jsou zásadní pro úspěch algoritmů hlubokého učení. Data slouží jako základní stavební kameny, které umožňují algoritmům učit se a zlepšovat své výsledky. Marr zdůrazňuje: "Data jsou zlato pro algoritmy hlubokého učení, protože čím větší a kvalitnější jsou soubory dat, tím lepší výsledky mohou tyto algoritmy dosáhnout" (Marr, 2020).

Dále v článku Marr přináší osm praktických příkladů využití hlubokého učení v různých oblastech. Tyto příklady zahrnují rozpoznávání obrazů, překlad textu, detekci podvodů, hlasové asistenty a další. Každý příklad ilustruje konkrétní aplikaci hlubokého učení a jeho potenciál v různých odvětvích.

Neuronové sítě

„Neuronové sítě (Neural Networks) jsou zahrnuty v kontextu umělé inteligence (AI) a jsou jedním z klíčových komponentů AI systému“ (Haykin 2005) AI systém musí být schopen ukládat znalosti, aplikovat tyto znalosti na řešení problémů a získávat nové znalosti zkušeností. „Neuronové sítě využívají symbolické struktury pro reprezentaci znalostí o dané problémové doméně a specifických znalostech o řešení problému. Symbolické reprezentace AI jsou srozumitelné pro lidského uživatele a umožňují efektivní komunikaci mezi lidmi a stroji“ (Haykin 2005). Neuronové sítě jsou také spojovány s učáním (learning) a schopností řešit problémy. Jak popsal Haykin „Učení v neuronových sítích probíhá na základě informací poskytnutých prostředím, které jsou následně využity ke zlepšování znalostní báze sítě.“ Tímto způsobem neuronová síť získává zpětnou vazbu (feedback) a přizpůsobuje své hypotézy.

IBM (2023) rozlišuje několik typů neuronových sítí:

- *Perceptron*: Nejstarší typ neuronové sítě, který vytvořil Frank Rosenblatt v roce 1958.
- *Feedforward neural networks (přepojovací neuronové sítě)*: Tyto sítě jsou složeny ze vstupní vrstvy, skryté vrstvy nebo vrstev a výstupní vrstvy. Jsou také známé jako multi-layer perceptrons (MLPs). Neurony v těchto

sítích jsou často sigmoidní, nikoli perceptrony, protože většina reálných problémů je nelineární. Tato síť je základem pro počítačové vidění, zpracování přirozeného jazyka a dalších aplikací neuronových sítí.

- *Konvoluční neuronové síť (CNNs)*: Tyto sítě se obvykle používají pro rozpoznávání obrazů, rozpoznávání vzorů anebo počítačové vidění. Využívají principy lineární algebry, zejména násobení matic, k identifikaci vzorů v obrazech.
- *Rekurentní neuronové síť (RNNs)*: Tyto sítě jsou identifikovány svými zpětnými smyčkami. Jsou využívány zejména při použití časových řadových dat pro předpovídání budoucích výsledků, například při předpovídání vývoje akciového trhu nebo předpovídání prodejů.

Jak zdůrazňuje v článku společnost IBM, je důležité si uvědomit, že termín "hluboké učení" (deep learning) často splývá s pojmem "neuronové sítě" (neural networks). "Hluboké" v hlubokém učení odkazuje na počet vrstev v neuronové síti. Pokud má síť více než tři vrstvy (včetně vstupu a výstupu), pak ji můžeme považovat za algoritmus hlubokého učení. Základní neuronová síť má obvykle jen dvě nebo tři vrstvy. (IBM, 2023a)

Přirozený jazyk

„Přirozený jazyk (Natural Language Processing) je oblastí umělé inteligence, která se zaměřuje na porozumění, generování a interakci s lidským jazykem. Využívá se k tomu široká škála technik a algoritmů, které umožňují počítačům analyzovat, interpretovat a generovat lidský jazyk.“ (Jurafsky a Martin 2020)

Přirozený jazyk je podle Goldberga (2017) klíčovým prvkem mnoha aplikací umělé inteligence, včetně:

- *Automatického překladu*: Přirozený jazyk umožňuje počítačům překládat texty z jednoho jazyka do druhého. Využívá se zde metody strojového učení a statistického modelování.
- *Chatbotů*: Přirozený jazyk umožňuje počítačům porozumět a odpovídat na lidské dotazy nebo příkazy. Chatboti využívají techniky zpracování přirozeného jazyka k analýze textových vstupů a generování přiměřených odpovědí. (Goldberg, Y., 2017)
- *Analýzy sentimentu*: Přirozený jazyk umožňuje počítačům rozpoznávat a analyzovat emoce, nálady a sentiment v textech. Tato technologie se využívá například v sociálních médiích nebo ve výzkumu trhu.
- *Extrahování informací*: Přirozený jazyk umožňuje počítačům analyzovat texty a extrahovat z nich klíčové informace, jako jsou jména, místa, datumy nebo události. Tato technologie je užitečná například při zpracování zpravodajství nebo automatizaci administrativních procesů. (Bird, Klein, a Loper, 2009)

Big data

Umělá inteligence a Big Data jsou úzce propojeny, přičemž každá z nich přispívá k té druhé a těží z ní. „Big Data označují obrovské množství strukturovaných i nestrukturovaných dat, která jsou generována z různých zdrojů, včetně sociálních médií, senzorů a online transakcí.“ (Mayer-Schönberger a Cukier 2013) Na druhé straně umělá inteligence zahrnuje řadu technik a algoritmů, které umožňují strojům provádět úkoly, jež obvykle vyžadují lidskou inteligenci, jako je porozumění přirozenému jazyku, rozpoznávání vzorců a vytváření předpovědí.

„Vztah mezi AI a Big Data je symbiotický, protože techniky AI hrají klíčovou roli při získávání smysluplných poznatků z Big Data, zatímco Big Data poskytují „palivo“ pro trénování a zlepšování modelů AI“ (Chen a kol. 2014). Dostupnost velkých a různorodých souborů dat umožňuje algoritmům AI identifikovat složité vzorce a korelace, které by pro lidské analytiky nemusely být zřejmé. Využitím dat velkého objemu se algoritmy AI mohou učit z obrovského množství informací a činit přesné předpovědi nebo rozhodnutí.

„Jednou z hlavních aplikací AI v kontextu Big Data je strojové učení, které zahrnuje trénování modelů na velkých souborech dat za účelem rozpoznávání vzorů a vytváření předpovědí nebo klasifikací.“ (Chen a kol., 2014) Algoritmy strojového učení, jako jsou neuronové sítě s hlubokým učením, prosperují díky dostupnosti rozsáhlých souborů dat pro trénování. Čím více dat mají k dispozici, tím lépe mohou zobecňovat a vytvářet přesné předpovědi. (Chen a kol., 2014)

2.4 Využití umělé inteligence v procesu rozhodování o žádostech o azyl

V kontextu migrace a azylových řízení by se dali využít různé druhy umělé inteligence (AI) k různým účelům. Můžeme je dělit například tímto způsobem:

1. Prediktivní analýza:

Prediktivní analýza představuje specializovaný obor vyhodnocování, který se zaměřuje na vytváření předpovědí ohledně budoucích událostí, chování a výsledků. Tento přístup využívá různé statistické metody, včetně algoritmů strojového učení a sofistikovaného modelování, s cílem analyzovat aktuální i historická data a odhadnout pravděpodobnost nastání specifické události. (SAP, 2024)

Moderní prediktivní analýza se rozvinula díky technologiím umělé inteligence (AI), jako jsou strojové učení, hluboké učení a neuronové sítě. „Tyto inovativní technologie umožňují rychlé zpracování velkých objemů dat, odhalování poznatků, které by byly obtížněji identifikovatelné lidským okem, a zvyšují přesnost předpovědí budoucích událostí. Automatizují také složité kroky v procesu prediktivní analýzy, včetně vytváření a testování prediktivních modelů.“ (SAP, 2024)

- **Použití:** Pomocí prediktivní analýzy mohou migrační úřady azylových řízení odhadovat pravděpodobnost úspěchu nebo zamítnutí žádosti na základě předešlých případů.

2. Rozhodovací systémy:

Tento typ AI se zaměřuje na automatizaci procesu rozhodování o azylových žádostech. „Může zahrnovat rozhodovací algoritmy a strojové učení pro efektivnější a rychlejší rozhodování. Rozhodovací systémy, které využívají umělou inteligenci, mají schopnost učit se z dat, identifikovat vzory a poskytovat personalizované a okamžité odpovědi.“ (Chatify, 2023)

- **Použití:** Rozhodovací systémy mohou pomoci identifikovat relevantní informace a kritéria pro posuzování žádostí a podporovat rozhodovací orgány v tomto procesu.

3. Biometrická identifikace:

Biometrická identifikace je technologie, která se používá k identifikaci, analýze a měření individuálních fyzických a behaviorálních charakteristik. Každá lidská osoba má unikátní biometrické charakteristiky, jako jsou otisky prstů, hlas, obličej, duhovka, nebo DNA. Tyto charakteristiky se mohou použít k jednoznačné identifikaci osoby. (Klímeck, 2012)

- **Použití:** V souvislosti s rozhodováním o žádostech o azyl se biometrická identifikace využívá k ověření totožnosti žadatelů o azyl a ke sledování pohybů osob v rámci azylového procesu. Tato technologie může pomoci s přesnějším a spolehlivějším zaznamenáváním údajů o žadatelích o azyl a s prevencí zneužití a podvodů v rámci azylového systému.

4. Chatboti:

„Chatboti jsou počítačové programy simulující lidskou konverzaci, často s využitím umělé inteligence (AI) a techniky zpracování přirozeného jazyka (NLP).“ Moderní chatboti umožňují okamžité odpovědi na otázky uživatelů, ať už pomocí textového nebo zvukového vstupu, bez lidské intervence. Tyto technologie jsou běžné v domácích chytrých reproduktorech, komunikačních aplikacích a pracovních prostředích. (IBM, 2024b)

„AI chatboti jsou pokročilejší, využívají strojové učení ke zlepšování odpovědí a technologie zpracování přirozeného jazyka (NLP) a porozumění přirozenému jazyku (NLU) k přesné interpretaci otázek uživatele.“ Díky schopnostem hlubokého učení jsou AI chatboti přesnější a umožňují uživatelům interagovat s nimi přirozeným způsobem bez nedorozumění. (IBM, 2024b)

„Virtuální agenti jsou další úroveň vývoje AI chatbotů, která kombinuje AI s robotickou automatizací (RPA) pro pokročilou interakci s uživateli.“ Zatímco AI chatboti mohou interpretovat a odpovídat na otázky, virtuální agenti mohou přímo reagovat na úmysly uživatele s minimální potřebou lidské intervence. (IBM, 2024b)

- **Použití:** Tento druh AI může zlepšit komunikaci s žadatelem o azyl, poskytovat jim nezbytné informace a usnadnit jim orientaci v procesu.

5. Jazyková analýza:

Zpracování přirozeného jazyka (NLP) je odvětvím umělé inteligence, které umožňuje počítačům analyzovat a generovat lidský jazyk. Chatboti s umělou inteligencí,

jako Mondly, Duolingo a Replika, hrají klíčovou roli v interaktivním učení jazyků, simulující reálné rozhovory a zlepšující mluvní dovednosti. Technologie rozpoznávání hlasu v nástrojích pro výuku jazyků umožňuje analýzu a zpětnou vazbu k výslovnosti, což podporuje efektivní rozvoj mluvních schopností. (Talkpal, 2024)

„Umělá inteligence se využívá také v oblasti překladu jazyků. Technologie strojového překladu se stává stále výkonnější a přesnější, a to díky pokroku v oblasti umělé inteligence a strojového učení.“ (Meriva, 2023) Překladové systémy využívají algoritmy strojového učení, které se učí na základě velkého množství trénovacích dat a jsou schopny překládat texty do různých jazyků. Tyto systémy jsou schopny rozpoznat gramatiku a slovní spojení v textu a přeložit je do cílového jazyka. Přestože jsou překladové systémy stále daleko od dokonalosti, umělá inteligence v této oblasti poskytuje uživatelům rychlý a pohodlný způsob překladu textů do různých jiných jazyků. (Meriva, 2023)

- **Použití:** Tato technologie může urychlit a zjednodušit vyhodnocování obsahu žádostí a zajistit efektivní komunikaci mezi úřady a žadateli o azyl s různými jazykovými rozdíly či bariérami.

Ve studii Nalbandiana (2022) je vysvětleno, jak technologie umělé inteligence zpracovávají a analyzují obrovské množství dat a vytvářejí předpovědi. Podrobně vysvětluje klíčové pojmy, jako jsou big data, strojové učení, prediktivní analýza, hluboké učení a neuronové sítě. Diskutuje se také o výzvách, které představují systémy černých skříněk, kde je rozhodovací proces algoritmů AI neprůhledný a obtížně interpretovatelný.

2.4.1 Využití umělé inteligence v procesu azylových řízení

V souvislosti s řízením migrace v rámci Evropské unie (EU) se stále častěji využívá umělá inteligence. Podle studie provedené Agatou Szwed (2022) se AI využívá především k automatizaci hraničních kontrol, identifikaci osob, ověřování cestovních dokladů a algoritmickému rozpoznávání objektů, například při celních kontrolách. „EU se snaží regulovat využití AI prostřednictvím různých sdělení a nařízení, včetně nedávno přijatého Aktu o umělé inteligenci z roku 2021“. AI může zvýšit bezpečnost hranic, automatizovat rozhodovací procesy a zlepšit účinnost migračních politik a komunikaci. Technologizace migrace a usnadnění mobility osob jsou významnými faktory, které AI přináší. (Agata Szwed, 2022) V kontextu těchto praktických aplikací AI, práce Schöpflina (2020) nabízí kritický pohled na možnosti a rizika spojená s rozvojem AI v migraci. Schöpflin analyzuje pokročilé technologie, jako jsou strojové učení a neuronové sítě, a prostřednictvím fiktivního příběhu ilustruje potenciální budoucí aplikace AI v této oblasti. Klíčovým bodem jeho analýzy je výzva k etickému a kritickému přístupu při využívání AI, zdůrazňující nutnost pečlivě zvážit rizika, včetně možnosti zneužití a porušení práv jednotlivců.

Projekt AFAR („*Algorithmic Fairness for Asylum Seekers and Refugees*“) zkoumá využití nových technologií v řízení migrace a azylu, zejména automatizaci nebo částečnou automatizaci rozhodnutí obvykle přijímaných lidmi, stejně jako složitější

systemy strojového učení a umělé inteligence a související použití mechanismů digitální identity. Jeho jádrem je zkoumání konceptu „spravedlnosti“ jako právního, normativního a politického konceptu. (Ruhs a Dražanová, 2023) Tímto tématem se také zabýval Daniel Ghezelbash (2020) ve svém článku, kde zdůrazňuje důležitost zapojení migrantů a dalších zúčastněných stran do procesu vývoje a implementace systémů umělé inteligence, které se využívají v řízení migrace. Podle Ghezelbasha je takové zapojení klíčové pro zajištění, že výstupy systémů AI budou kvalitnější a skutečně reflektovat potřeby a specifika různých migračních situací. To pomáhá zvýšit transparentnost, odpovědnost a spravedlnost v rozhodovacích procesech, a tím chránit lidská práva a důstojnost migrantů. (Ghezelbash, 2020)

Umělá inteligence se ukazuje jako klíčový nástroj v oblasti migrace a přesídlení uprchlíků. Studie Thielemanna (2019) ve svém článku "*Artificial intelligence and migration governance*" zkoumá, jak AI může změnit rozhodovací procesy a politiku v oblasti migrace, zlepšovat pochopení migračních trendů a předpovědi, ale zároveň upozorňuje na rizika jako diskriminace a porušování práv migrantů. Naopak, podle Ineliho má využití umělé inteligence (AI) při přesídlování uprchlíků potenciál výrazně usnadnit a zlepšit celý proces. Tato technologie umožňuje rychlé rozhodování o velkém množství případů přesídlení a současně snižuje náklady spojené s tímto procesem. „Díky AI je také možné identifikovat nejzranitelnější uprchlíky, kteří potřebují trvalá řešení, a spojit je s vhodnými hostitelskými státy a komunitami. Tímto způsobem se přispívá k efektivní integraci uprchlíků do nového prostředí.“ (Ineli, 2023) Oba přístupy poukazují na potřebu etického a transparentního využívání AI, aby se zajistilo spravedlivé a efektivní řízení migrace.

Digitální nástroje, včetně chatbotů založených na umělé inteligenci, mají zásadní význam pro šíření informací během migrace. Tyto chatboti mohou poskytovat důležité informace migrantům týkající se průběhu celého procesu a jejich práv. Tím se zvyšuje transparentnost a srozumitelnost celého procesu pro uprchlíky, což napomáhá k jejich lepší integraci do nového prostředí. (Ineli, 2023)

Použití umělé inteligence k párování uprchlíků s různými lokalitami ve státě přesídlení je dalším způsobem, jak tato technologie přispívá k úspěšné integraci. Tento proces nepředstavuje zásadní problém lidských práv a neporušuje principy mezinárodního práva o uprchlících. Je však důležité, aby byla jasně definována role a odpovědnosti konečného uživatele. Navíc je nutné chránit data zranitelných uprchlíků, jako jsou děti, osoby potřebující lékařskou péči, ženy v ohrožení nebo přeživší mučení a násilí, a omezit přístup třetích stran k těmto informacím. (Ineli, 2023)

Další studie, provedená Ho (2017) s názvem „*Using Machine Learning to Predict Asylum Outcomes*“, významně přispívá do oblasti zkoumání potenciálu technik strojového učení při předpovídání výsledků azylových případů, konkrétně v kontextu využití v USA. S využitím rozsáhlých souborů dat a pokročilých algoritmů se autor pokouší vyvinout prediktivní modely, které mohou pomoci rozhodovacím orgánům při odhadu pravděpodobnosti úspěšné žádosti o azyl. Hoův výzkum se zabývá potřebou přesnějších a efektivnějších rozhodovacích procesů v oblasti azylu prostřednictvím využití možností strojového učení. Článek zdůrazňuje potenciál těchto technik při zvyšování efektivity a spravedlnosti rozhodování o azylu prostřednictvím

zkoumání vzorců a korelací v datech. Zjištění naznačují, že techniky strojového učení mohou poskytnout přesnější a efektivnější predikce výsledků azylových případů. Tyto poznatky ukazují na potenciál těchto metod zlepšit rozhodovací procesy v oblasti azylu a zajištění spravedlivějšího zacházení s žadateli o azyl. (Ho, 2017)

V rámci studie Nalbandiana (2022) se zkoumá využití umělé inteligence Úřadem Vysokého komisaře OSN pro uprchlíky (UNHCR), který byl založen Valným shromážděním OSN dne 14. prosince 1950 s cílem poskytnout ochranu a pomoc uprchlíkům. UNHCR implementoval umělou inteligenci s cílem posílit důstojnost uprchlíků, zefektivnit poskytované služby, omezit podvody, zvýšit efektivitu a centralizovat informace (Nalbandian 2022). Biometrické technologie, včetně skenerů oční duhovky a otisků prstů, jsou využívány prostřednictvím ekosystému pro registraci obyvatelstva a správu identit (PRIMES). Cílem tohoto systému je poskytovat uprchlíkům a žadatelům o azyl snadno dostupné identifikační doklady a usnadnit jejich interakci se státy a organizacemi. (Nalbandian 2022)

V souvislosti s tímto tématem uvádí Ana Beduschi (2021), že některé země, jako Kanada a Německo, již integrují algoritmická rozhodnutí do procesů imigrace a azylu. Tyto státy využívají algoritmy umělé inteligence k předpovídání migračních trendů, ověřování totožnosti migrantů a zlepšení kontroly hranic. Konkrétní výsledky takového využití zahrnují přesnější predikce migračních trendů, což umožňuje státům lépe plánovat a připravovat se na příchod migrantů. Dále tyto algoritmické systémy pomáhají efektivně ověřovat totožnost migrantů prostřednictvím srovnání biometrických dat a dalších identifikačních informací. Integrace algoritmů umělé inteligence do procesů kontroly hranic přináší efektivnější monitorování a správu hranic, což pomáhá státům lépe řídit tok migrace a zlepšit bezpečnost.

2.4.2 Kritická analýza stávajícího využití umělé inteligence

V rámci kritické analýzy stávajícího využití umělé inteligence v oblasti rozhodování o azylu vystupují několik klíčových problémů, které je třeba pečlivě zvážit. Autoři Gavaghan a kolektiv (2019) výrazně kladou důraz na přesnost, lidskou kontrolu, transparentnost, předpojatost a ochranu soukromí v prediktivních algoritmech. Jejich analýza zdůrazňuje potřebu nezávislého dohledu nad těmito technologiemi, zahrnutí lidského faktoru do rozhodovacího procesu, transparentních postupů a eliminace předpojatosti. Tato problematika je důležitým prvkem i ve výzkumu Maxe Borkerta (2020), který se zaměřuje na etické aspekty používání umělé inteligence v rozhodování o migraci. Borkertův výzkum přináší podrobný pohled na různé etické výzvy spojené s implementací AI v rozhodovacích procesech týkajících se migrace. Jeho práce podtrhuje nutnost transparentnosti a odpovědnosti v AI systémech, stejně jako důležitost zapojení relevantních zúčastněných stran při navrhování a provádění těchto technologií. (Borkert, 2020)

Dále se zaměříme na výzkumy od Samuela a Caselliho (2021) a týmu Kammerera, Bruniho a Ruedina (2022), kteří se věnují lidské odpovědnosti a rozhodování v kontextu umělé inteligence při procesu rozhodování o azylu. Oba týmy zdůrazňují klíčový význam transparentnosti, odpovědnosti a lidského dohledu při im-

plementaci těchto technologií. Samuels a Caselli (2021) navrhují transparentní postupy a mechanismy pro kontrolu a vysvětlení rozhodnutí, zatímco Kammerer, Bruni a Ruedin (2022) poukazují na potřebu lidského dohledu pro zajištění spravedlivého a důstojného rozhodování.

Další důkladná analýza od Kinchina a Mougoueiho (2022) se zaměřuje na integraci umělé inteligence do procesů určování statusu uprchlíka, zejména v oblasti posuzování důvěryhodnosti. Autoři upozorňují na subjektivní povahu současného posuzování důvěryhodnosti a klade otázku, zda by umělá inteligence mohla zlepšit efektivitu a přesnost tohoto procesu. Zároveň zdůrazňují možná rizika spojená s používáním umělé inteligence a potřebu stanovení standardních kritérií pro posuzování důvěryhodnosti.

Celkově tyto analýzy poskytují hlubší vhled do klíčových otázek spojených s využíváním umělé inteligence v rozhodování o azylu. Klade se důraz na potřebu transparentnosti, odpovědnosti a lidského dohledu jako základních prvků při implementaci těchto technologií v citlivém rozhodovacím procesu. Tato zdokonalená perspektiva také naznačuje potenciál umělé inteligence přinést vylepšení v posuzování důvěryhodnosti, ačkoliv současně zdůrazňuje nutnost pečlivého přístupu k rizikům, která mohou s používáním této technologie v této oblasti souviset.

2.5 Analýza dat současného stavu rozhodování o žádostech o azyl

2.5.1 Analýza dat o migraci

Podle Evropského parlamentu je migrace komplexní jev, který je ovlivňován různými faktory, známými jako push a pull faktory. Push faktory jsou důvody, proč lidé opouštějí svou zemi, zatímco pull faktory jsou důvody, proč se stěhují do konkrétní země. Mezi hlavní push patří sociální a politické faktory, jako je válka, konflikt, pronásledování či obava o bezpečí. Těmto lidem hrozí porušování lidských práv nebo politické perzekuce, a proto hledají útočiště v jiných zemích. Naopak mezi pull faktory spadají důvody, proč se lidé stěhují do konkrétní země. Sem patří ekonomické příležitosti, lepší životní podmínky, vyšší mzdy a možnosti vzdělání. Lidé hledají lepší budoucnost pro sebe a své rodiny a migrují do zemí, kde mají vyhlídky na zlepšení svého života. (European Parliament, 2020)

Rozdělení faktorů do tří kategorií podle Evropského parlamentu:

Sociální a politické faktory

Mezi push faktory patří sociální a politické faktory, jako je pronásledování kvůli etnické příslušnosti, náboženství, rase, politickým názorům nebo kultuře. Vládní konflikty, válečné situace a obecné riziko pronásledování mohou také přimět lidi k opuštění své země. Tito jednotlivci jsou často považováni za humanitární uprchlíky a hledají útočiště v zemích, které mají liberálnější politiku vůči uprchlíkům. (European Parliament, 2020)

Demografické a ekonomické příčiny

Demografické změny, jako je růst nebo úbytek populace a stárnutí společnosti, mohou mít vliv na ekonomiku a pracovní příležitosti v dané zemi. Ekonomické faktory, jako je vysoká nezaměstnanost a nízké příjmy, mohou vést k hledání lepších životních podmínek a pracovních příležitostí v jiných zemích. (European Parliament, 2020)

Environmentální a klimatické faktory

Změna klimatu a přírodní katastrofy, jako jsou povodně a hurikány, mohou také přimět lidi k migraci, když se snaží uniknout před nepříznivými životními podmínkami. Tito migranti jsou často označováni jako environmentální migranti. (European Parliament, 2020)

2.5.2 Přehled počtu žádosti o azyl a rozhodnutí v EU

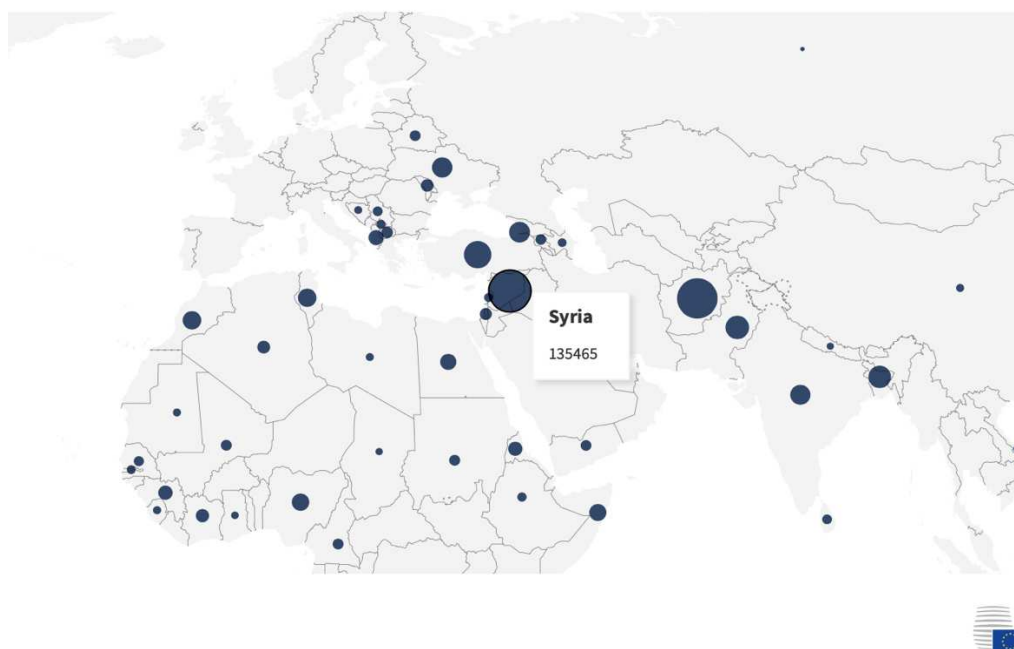
Podle fakt a čísel zjištěných na stránkách Evropského parlamentu bylo v roce 2022 v Evropské unii podáno téměř milion žádostí o azyl, což představuje 52,1% nárůst oproti předchozímu roku a nejvyšší úroveň od roku 2016. V době vrcholící migrační krize v letech 2015–2016 dosáhl počet žadatelů až 1 221 690. (European Parliament, 2017)

„Počet prvožadatelů o azyl v EU v roce 2022 dosáhl 881 220, což je o 64 % více než v předchozím roce.“ (European Parliament, 2017) Prvožadatelé jsou lidé, kteří poprvé podávají žádost o azyl v určité zemi EU.

Země s největším relativním nárůstem prvožadatelů v roce 2022 byly Irsko (+421,8 %), Chorvatsko (+367,9 %) a Rakousko (+181,4 %). Německo mělo nejvíce prvožadatelů, tvořících 24,7 % z celkového počtu v EU, následovaly Francie, Španělsko a Rakousko. (European Parliament, 2017)

Podle nedávných zpráv Agentury Evropské unie pro azyl (EUAA) vykazují Syřané i Afghánci v roce 2022 výrazný nárůst v podání žádostí o azyl v zemích EU+. Syřané představují největší skupinu s 135 465 žádostmi (obr. 1), což představuje 24% nárůst oproti roku 2021, nejvíce od roku 2016. Na druhém místě jsou Afghánci se 129 000 žádostmi, což je nárůst o 29 % a nejvyšší úroveň od roku 2016. Obě země jsou také mezi třemi nejčastějšími národnostmi, které se dopouštějí nelegálního překračování hranic v EU+. Míra uznání žadatelů je vysoká, zejména u Syřanů, kde dosahuje 94 %, zatímco u Afghánců je 54 %. Tento trend zahrnuje i větší podíl humanitární pomoci.

Zároveň je zaznamenáno, že nejčastěji žádali o azyl Syřané, Afghánci, Venezuela a Turci, kteří dohromady tvořili téměř 40 % všech prvožadatelů. Tuto informaci poskytuje European Union agency for asylum na svých webových stránkách. (EUAA, 2022) „Celkem bylo v roce 2022 v EU schváleno 384 245 žádostí, přičemž asi 44 % z nich obdrželo status uprchlíka, 31 % doplňkovou ochranu a 25 % humanitární ochranu.“ Tuto statistiku zpracoval Evropský parlament ve své analýze o migraci a azylu. (European Parliament, 2017)



Obr. 1 Státy s největším počtem žadatelů o azyl v EU (European Council, 2023)

2.6 Shrnutí literární rešerše

2.6.1 Shrnutí klíčových poznatků z literatury

Azylové řízení je komplexní proces, který hraje klíčovou roli v posuzování oprávněnosti mezinárodní ochrany pro žadatele o azyl. Tento proces je regulován jak národními zákony, tak mezinárodními normami, které se vyvíjely v reakci na mezinárodní události a změny ve světové migraci. Klíčové poznatky z literatury naznačují, že azylové řízení není pouze právní procedurou, ale zahrnuje i sociální a humanitární aspekty. Umělá inteligence v tomto kontextu nabízí mnoho možností, jako je prediktivní analýza, rozhodovací systémy, biometrická identifikace, chatboti a jazyková analýza, které mohou zlepšit efektivitu, transparentnost a spravedlnost tohoto procesu.

Během rešerše byly identifikovány klíčové faktory ovlivňující rozhodování o žádostech o azyl. Migrace je komplexní fenomén, který je ovlivněn různými push a pull faktory. Mezi push faktory patří sociální a politické faktory, jako je pronásledování kvůli etnické příslušnosti, náboženství, rase, politickým názorům nebo kultuře, zatímco pull faktory zahrnují ekonomické příležitosti, lepší životní podmínky a vyšší mzdy. Zároveň je zaznamenán významný nárůst počtu žádostí o azyl v Evropské unii v posledních letech, převážně Syřané a Afghánci, což naznačuje rostoucí potřebu efektivního azylového systému.

2.6.2 Identifikace mezery ve stávajících výzkumech a potenciální oblasti pro další analýzu

I přesto, že azylové řízení je důležitou součástí mezinárodního práva a lidských práv, existují oblasti, které vyžadují další zkoumání. Mezi tyto oblasti patří efektivita azylových systémů, zajištění spravedlivého a nestranného posuzování žádostí o azyl, ochrana zranitelných skupin a dětí v azylových procesech, a zajištění práva na odvolání a soudní přezkum.

Zároveň je z literatury patrná mezera v hlubší analýze dopadů využití umělé inteligence na lidská práva a důstojnost žadatelů o azyl. Potenciální oblasti pro další výzkum zahrnují zkoumání možností vylepšení procesu posuzování důvěryhodnosti, stanovení standardních kritérií pro používání umělé inteligence v azylových řízeních a posílení ochrany osobních údajů a soukromí žadatelů o azyl.

Zjištěné informace ukazují, že i přes rozsáhlé množství literatury v oblasti migrace a azylových procesů existuje potřeba dalšího výzkumu. Hluběji by měly být zkoumány faktory ovlivňující míru uznání žádostí o azyl, dlouhodobé dopady migrační politiky na společnost a ekonomiku, a možné souvislosti mezi migrací a dalšími sociálními jevy. Tyto oblasti představují perspektivní směry pro budoucí výzkum, který může přispět k lepšímu porozumění a efektivnějšímu řešení problematiky migrace a žádostí o azyl.

3 Metodologie

Tato část mé diplomové práce popisuje metodiku, kterou jsem použila pro provedení výzkumu, včetně výběru vhodných metod a technik pro sběr a analýzu dat, vedoucí k vytvoření algoritmu pro rozhodování o azylových žádostech.

3.1 Popis metodologických postupů

V prvním kroku jsem prováděla důkladnou rešerši, kde jsem zkoumala existující literaturu, dokumenty a další informace související s umělou inteligencí a procesem žádání o azyl. Tato fáze zahrnovala rozsáhlé procházení různých zdrojů, včetně webových stránek, databází, odborných článků a veřejně dostupných zpráv. Pro účely sběru dat jsem analyzovala obsah veřejně dostupných rozhodnutí o udělení či neudělení azylu, která jsem našla online. Tato data jsem následně využila pro vytvoření analýz a modelů, které jsem konzultovala s přední českou nevládní organizací zaměřenou na problematiku migrace a uprchlíků, jejíž feedback mi napomohl ve zdokonalení výzkumu.

Ve druhém kroku jsem vytvořila datový set obsahující podrobné informace o žadatelích o azyl o velikosti několika stovek záznamů. Nejdříve jsem data rozdělila podle zemí, a poté podle výsledku validace, tedy zda byla žádost schválena nebo zamítnuta. Data jsem následně rozdělila na trénovací a testovací sady s poměrem 80:20, na kterých jsem použila metodu učení s učitelem. Tento přístup mi umožnil učit algoritmus na trénovacích datech a ověřovat ho na testovacích datech. Vybrala jsem klasifikační metody, které mi pomohly kategorizovat žádosti o azyl na základě různých kritérií, jako jsou například země původu žadatelů.

Pro návrh algoritmu umělé inteligence, který rozhoduje o žádostech na základě relevantních kritérií, jsem použila metody zpracování přirozeného jazyka (NLP). Tyto metody mi umožnily analyzovat textové informace z žádostí, včetně rozpoznávání jazyka a identifikace klíčových slov. Dále jsem využila Bayesovské modely pro odhad pravděpodobnosti udělení azylu na základě různých faktorů. Algoritmus byl trénován na vzorových datech s využitím knihoven pro strojové učení jako Numpy, Scikit-learn a matplotlib, které jsem zvolila pro jejich rozsáhlé možnosti a podporu komplexních modelů učení. Do rozhodovacího procesu jsem zahrнула metodu shody slov za použití Levenshteinovi vzdálenosti, pro kvalitnější výsledek rozhodování.

Využití Cloud Computingu jsem zvažovala především pro jeho škálovatelnost a výkon, které poskytují platformy jako AWS, Google Cloud nebo Microsoft Azure. I když tyto nástroje jsou vhodné pro zpracování velkých datových sad, v mé práci jsem se na ně nezaměřila.

V dalším kroku jsem provedla detailní vyhodnocení úspěšnosti algoritmu na testovacích datech. Tato analýza zahrnovala hodnocení přesnosti, úplnosti a F1 skóre, což poskytlo komplexní pohled na úspěšnost algoritmu a jeho dopady. Kromě toho jsem prováděla analýzu shody slov v žádostech, což odhalilo důležitá témata a pomohlo identifikovat případné nepřesnosti ve výsledcích rozhodování.

V poslední fázi práce jsem se zaměřila na interpretaci výsledků algoritmu. Tento proces zahrnoval analýzu rozhodnutí algoritmu a jeho dopadu na žádosti o azyl. Zkoumala jsem faktory, které ovlivňovaly rozhodování algoritmu, a zhodnotila jsem, jaké praktické dopady mohly mít výsledky algoritmu na konkrétní žadatele o azyl. Tato interpretace byla klíčová pro posouzení relevance a efektivity algoritmu v kontextu rozhodování o žádostech o azyl.

V diplomové práci jsem použila několik pokročilých metod a technologií v oblasti strojového učení a textové analýzy. Doplněním jsou informace o konkrétních technologiích a metodách použitých v práci a o dalších metodách a technologiích, které by mohly být vhodné pro další vylepšení algoritmu:

3.2 Technologie

1. **Python a knihovny pro strojové učení:** Python je široce využívaný programovací jazyk pro implementaci algoritmů strojového učení. Jeho aktuální verze je 3.12.2. Existuje několik populárních knihoven, které jsou klíčové pro práci v oblasti strojového učení:
 - **TensorFlow:** TensorFlow poskytuje výkonnou platformu pro vývoj modelů umělé inteligence vhodných pro zpracování přirozeného jazyka a rozpoznávání obrazu. Jeho klíčové výhody zahrnují schopnost provádět nízkourovňové operace přes mnoho akceleračních platform a automatický výpočet gradientů, což usnadňuje tvorbu a optimalizaci modelů. (NVIDIA, 2024b)
 - **Scikit-learn:** Další důležitá knihovna pro strojové učení v Pythonu, která nabízí mnoho algoritmů a nástrojů pro analýzu dat. Aktuální verze knihovny Scikit-learn, kterou používám je 0.24.0. S její pomocí můžeme snadno implementovat širokou škálu modelů pro regresi, klasifikaci a shlukování, stejně jako provádět statistickou analýzu těchto modelů. Je to užitečný nástroj pro tvorbu a testování algoritmů (Ashis, 2022)
 - **Keras:** Specializovaná knihovna pro hluboké učení, která usnadňuje tvorbu neuronových sítí. „Keras je rozhraní API pro hluboké učení na vysoké úrovni vyvinuté společností Google pro implementaci neuronových sítí. Je napsán v Pythonu a používá se k usnadnění implementace neuronových sítí.“ (Simplilearn, 2024)
 - **PyTorch:** Další populární knihovna, která je plně vybavená pro vytváření modelů hlubokého učení, což je typ strojového učení, který se běžně používá v aplikacích, jako je rozpoznávání obrázků a zpracování jazyků. (NVIDIA, 2024a)
2. **Cloud Computing:** Diskutujte o výhodách využití cloudových platform pro ukládání a zpracování dat, zejména pokud jde o velké objemy dat. Cloud Computing je poskytování výpočetních služeb, jako jsou servery, úložiště, data-báze a software, přes internet („cloud“), což umožňuje rychlejší inovace a flexibilní zdroje. (Azure, 2024)

3. **OpenAI GPT:** OpenAI GPT (Generative Pre-trained Transformer) je pokročilý generativní model založený na transformerech, který je široce využíván pro textovou generaci a analýzu. Tento model byl předtrénován na velkém množství textových dat a je schopen vytvářet texty v odpověď na zadaný vstup nebo provádět různé úkoly, jako je sumarizace, překlad nebo odpovídání na otázky. (Platform OpenAI, 2024)

3.3 Metody

1. **Metoda F1 score:** Datová sada je rozdělena na dvě části (trénovací a testovací). Testovací sada se používá k výpočtu výkonu a vyhodnocení modelu. Vyhodnocení modelu se aktivuje automaticky po úspěšně dokončení trénování. Podle Microsoftu (2024) pro vyhodnocení používá vlastní klasifikace textu následující metriky:

- **Přesnost:** Měří, jak přesný model je. Jedná se o poměr mezi správně identifikovanými pozitivními výsledky (pravdivě pozitivními výsledky) a všemi identifikovanými pozitivními výsledky. Metrika přesnosti ukazuje, kolik predikovaných tříd je správně označeno.
- **Úplnost:** Měří schopnost modelu předpovědět skutečné kladné třídy. Jedná se o poměr mezi predikovanými pravdivě pozitivními výsledky a skutečným příznakem. Metrika úplnosti ukazuje, kolik predikovaných tříd je správných.
- **F1 skóre:** Skóre F1 je funkce přesnosti a úplnosti. Je potřeba, když hledáte rovnováhu mezi přesností a úplností.

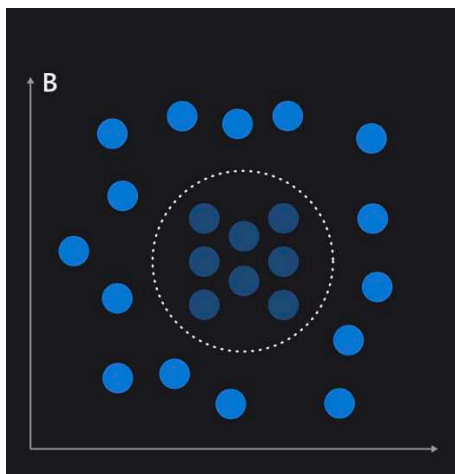
$$F1 = 2 \times \frac{\text{přesnost} \times \text{úplnost}}{\text{přesnost} + \text{úplnost}}$$

- „Přesnost, úplnost a skóre F1 se počítají pro každou třídu samostatně (vyhodnocení na úrovni třídy) a pro model souhrnně (vyhodnocení na úrovni modelu).“ (Microsoft, 2024)
2. **Levenshteinova vzdálenost:** je metoda, která udává nejmenší počet znakových operací, které jsou potřebné k převodu jednoho řetězce na druhý. Znakové operace zahrnují vložení, vymazání a substituci tokenů. Použití této metriky poskytuje dobrou představu o míře podobnosti mezi dvěma řetězci. Jednoduše řečeno, tato metoda spočítá počet kroků, které bylo zapotřebí k přizpůsobení jednoho textu druhému. (Vodsloň, 2007)

Tento přehled prezentuje algoritmy strojového učení, ukazující metody používané pro klasifikaci a shlukování dat:

3. **Naivní Bayesovy klasifikátory:** „Klasifikátor Naïve Bayes je řízený algoritmus strojového učení, který se používá pro klasifikační úlohy, jako je klasifikace textu. Je také součástí rodiny generativních algoritmů učení, což znamená, že se snaží modelovat distribuci vstupů dané třídy nebo kategorie. Na rozdíl od diskriminačních klasifikátorů, jako je logistická regrese, se neučí,

keré vlastnosti jsou nejdůležitější pro rozlišení mezi třídami.“ (IBM, 2024a) Algoritmy odhadují pravděpodobnost výskytu události na základě souvisejících událostí.



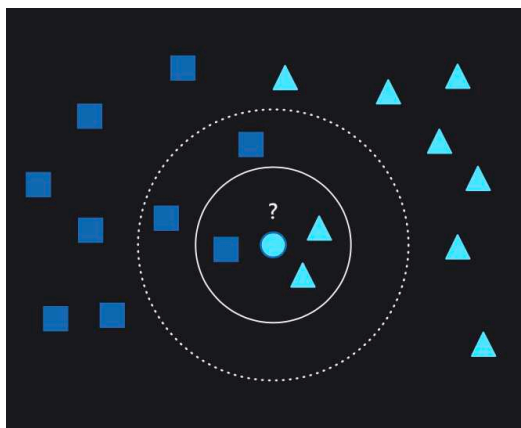
Obr. 2 Naivní Bayesovi algoritmy (Azure, 2024)

- 4. Metoda k-means:** Algoritmus K-means je nehierarchická metoda shlukování, která data rozděljuje do předem určeného počtu shluků (k). Každý datový bod je přiřazen k nejbližšímu středu shluku, a středy jsou následně aktualizovány jako průměr všech bodů ve shluku. Algoritmus si klade za cíl minimalizovat rozdíly uvnitř jednotlivých shluků. (Kučera, 2024) Algoritmy K-Means rozdělují data do skupin, přičemž K označuje počet těchto skupin. Data v rámci jednotlivých skupin jsou si podobná (homogenní), zatímco data mezi různými skupinami se liší (heterogenní). (Azure, 2024)



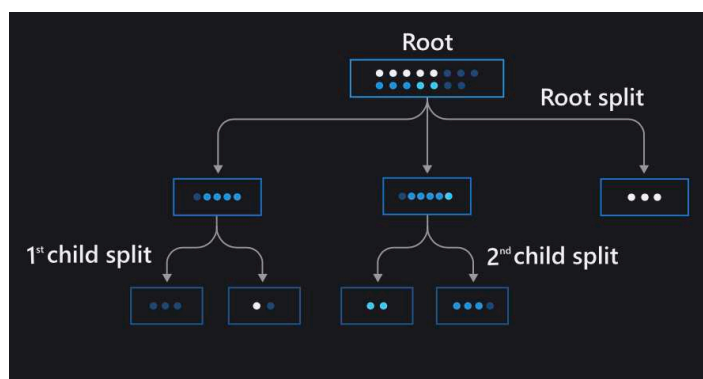
Obr. 3 K-Means algoritmy (Azure, 2024)

5. **Support vector machine (SVM):** „Support vector machine je dohlížený algoritmus strojového učení, který klasifikuje data nalezením optimální linie nebo nadroviny, která maximalizuje vzdálenost mezi každou třídou v N-rozměrném prostoru.“ (IBM, 2023b)
6. **k-Nejblížších sousedů (kNN):** Algoritmus K-nearest neighbors (KNN) je metoda řízeného učení, založená na principu, že objekty blízko sebe mají tendenci patřit do stejné kategorie. Základním předpokladem je, že věci, které jsou blízko sebe, jsou si navzájem podobné. (Airon, 2022)



Obr. 4 K-Nearest neighbor algoritmus (Azure, 2024)

7. **Rozhodovací strom:** Algoritmy rozhodovacího stromu rozdělují data do dvou nebo více homogenních skupin pomocí sérií rozhodovacích pravidel typu když-tak. Tyto pravidla identifikují nejvýznamnější rozdíly mezi datovými body a používají je k efektivnímu oddělení dat do jednotlivých skupin nebo tříd. (Azure, 2024)



Obr. 5 Rozhodovací strom (Azure, 2024)

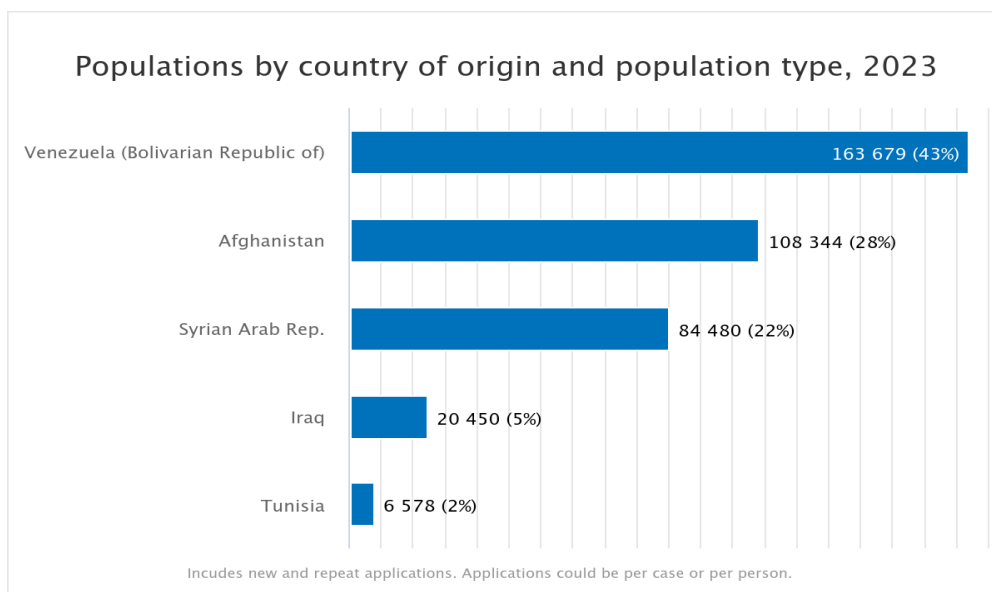
4 Analýza dat a implementace

Ve čtvrté kapitole této práce se zaměřím na analýzu dat a implementaci algoritmu pro vyhodnocování žádostí o azyl. Nejprve popíši složení datové sady, kterou využívám pro hodnocení těchto žádostí, a zdůvodním volbu konkrétního formátu dat. Poté se podrobněji zaměřím na implementaci algoritmu, který slouží k automatickému vyhodnocování žádostí o azyl, a vysvětlím jeho programování. V rámci popisu implementace budu pojednávat o použitých technologiích a metodách a jejich vlivu na výsledky algoritmu. Nakonec provedu zhodnocení získaných výsledků a představím zjištěné poznatky z vyhodnocování žádostí o azyl.

4.1 Sběr a analýza dat o žadatelích o azyl

Jak bylo zmíněno v podkapitole 2.5 „Analýza dat současného stavu rozhodování o žádostech o azyl“, žadatelé o azyl odcházejí ze zemí z mnoha důvodů – politických, demografických, ekonomických a dalších. Každý žadatel, tedy jedinec, má individuální a specifický problém, a je nezbytné k němu i tak přistupovat.

Důvody, proč jsem si vybrala právě tyto čtyři země, jsou založeny na statistických datech, která ukazují historicky významný nárůst žádostí o azyl. Cílem bylo zahrnout data z různých oblastí s rozmanitými problémy. Mezi tyto země patří Sýrie, Irák, Tunisko a Afghánistán.



Tab. 1 Počet žádostí o azyl od roku 2013 do roku 2023 (UHRC, 2024)

Lidé opouštějí své rodné země z různorodých důvodů. Většina z nich se stěhuje kvůli pronásledování, etnické příslušnosti, náboženství, rasové diskriminaci, politickým

názorům, kulturním rozdílům, porušování lidských práv, sexuální orientaci, politické perzekuci, obavám o život a bezpečí a mnoha dalším důvodům, které by bylo vhodné zde zmínit.

Když mluvím o individuálních problémech žadatelů o azyl, je důležité si uvědomit, že každý případ je unikátní. Například někteří lidé opouštějí své domovy kvůli politické perzekuci, kdy jsou pronásledováni za své politické názory nebo členství v určitých politických skupinách. Jiní se pak stěhují kvůli náboženskému pronásledování, kdy jsou ohroženi kvůli své víře nebo náboženské příslušnosti.

Kromě toho můžeme pozorovat i situace, kdy lidé opouštějí své země kvůli ekonomickým problémům. Nadměrná chudoba, nedostatek pracovních příležitostí a ekonomická nejistota mohou rovněž přivést k migraci.

Dalším aspektem jsou obavy o život a bezpečí. Někteří lidé se stěhují z důvodu násilí, sexuálního obtěžování nebo obecného nebezpečí, které ohrožuje jejich životy a životy jejich rodin.

Zkrátka, existuje mnoho faktorů, které ovlivňují rozhodnutí jednotlivců opustit své domovy a hledat azyl v jiných zemích. Rozličnost těchto důvodů vyžaduje individuální a citlivý přístup při posuzování jejich žádostí o azyl.

4.1.1 Sběr dat

Sběr dat pro tuto práci byl neodmyslitelnou součástí celého výzkumu a současně jednou z nejnáročnějších částí práce. Vzhledem k platným právním předpisům v Evropské unii, zejména ochraně osobních údajů dle GDPR, bylo získání relevantních dat obtížné. Avšak i přes tyto překážky jsem se pokusila získat co nejkvalitnější a reprezentativní soubor informací. Data, která jsem použila pro analýzu, jsem čerpala převážně z veřejně dostupných zdrojů na internetu. Patří sem například soudní rozhodnutí týkající se udělení či neudělení azylu, příspěvky na blogových stránkách, diskusní fóra a další.

Je třeba poznamenat, že ne všechna data pochází ze zdrojů, které jsou plně ověřené. Některé informace jsem často musela kombinovat z různých zdrojů a sestavovat vlastní dataset. Datový soubor, který jsem vytvořila, zahrnuje údaje i z jiných zemí, neboť pro Českou republiku nebylo možné nalézt dostatečné množství relevantních informací. Cílem bylo vytvořit dataset obsahující dostatečné množství vzorových příkladů, aby byl výstup analýzy co nejkvalitnější a co nejpřesnější.

Jedním z důležitých aspektů bylo shromažďování informací o důvodech, proč lidé opouštějí své rodné země, jak bylo zmíněno výše. Definovala jsem základní faktory, které představují hlavní důvody, proč lidé opouštějí konkrétní zemi.

Nejčastějším důvodem, proč lidé opouštěli Sýrii, bylo nebezpečí ze strany islámských extremistů. Tento faktor představoval pro mnoho jedinců zásadní hrozbu a často vedl k jejich rozhodnutí hledat azyl v jiných zemích. Muži i ženy byli vystaveni riziku teroristických útoků a pronásledování ze strany extrémistů, což vedlo k utíkání ze země. Jedním z hlavních problémů, kterým muži čelili, bylo nejen nebezpečí ze strany islámských extremistů, ale také dlouhodobý ozbrojený konflikt, politické aktivity, sexuální orientaci a násilí ze strany vojenských skupin. Ženy se zase často potýkaly s rodinným násilím, perzekucí ze strany vlády, aktivismu za lidská

práva a ohrožením ze strany ozbrojených skupin, což způsobilo, že hledaly útočiště mimo Sýrii. Tyto skutečnosti jim nedávají jinou možnost než opustit svou zemi a hledat bezpečí jinde.

Muži i ženy v Afghánistánu čelili různým formám perzekuce a ohrožení, což je vedlo k rozhodnutí hledat bezpečí mimo zemi. Muži byli často pronásledováni kvůli svému náboženství, politické aktivitě nebo sexuální orientaci. Ženy čelily perzekuci z důvodu jejich aktivismu, novinářské činnosti nebo kvůli sexuálnímu zneužívání. Teroristické útoky a domácí násilí dále zhoršovaly jejich situaci. Tato neustálá hrozba a nedostatek bezpečí vedly mnoho lidí k opuštění Afghánistánu.

V Tunisku se muži i ženy ocitají v obtížné situaci, kdy čelí různým formám perzekuce a ohrožení, což je vedou k hledání bezpečí jinde. Muži jsou konfrontováni s náboženskou perzekucí, politickým pronásledováním a hrozbou extremistických skupin kvůli jejich politickým postojům a aktivitám. Novináři jsou vystaveni riziku útoků a perzekuce kvůli své práci. Ženy, na druhou stranu, čelí domácímu násilí, kritice ze strany společnosti kvůli svému rodinnému stavu a sexuální orientaci a mohou být terčem perzekuce kvůli svému aktivismu.

Důvodem pro muže i ženy v Iráku je násilí ze strany vojenských a ozbrojených skupin, politickou perzekucí a útoky teroristických skupin. Muži čelí také falešným obviněním z terorismu a ohrožením kvůli své novinářské práci a politickým aktivitám. Ženy jsou terčem útoků ze strany ozbrojených skupin a čelí násilí v humanitárních organizacích. Problematika spojená s jejich sexuální orientací a náboženskou příslušností dále zhoršuje jejich bezpečnostní situaci.

Po sběru dat jsem vytvořila CSV tabulky, kde jsem systematicky rozdělila a uspořádala veškerá data podle jednotlivých zemí. Tyto tabulky byly upraveny do formátu, který je vhodný pro použití v algoritmu využívající umělou inteligenci k vyhodnocování žádostí o azyl. Takto připravené tabulky umožňují efektivní a strukturovanou analýzu dat, což je klíčové pro následné zpracování a fungování algoritmu.

4.1.2 Struktura dat

Informace, které jsem shromáždila, pocházely z různých zdrojů, převážně z veřejně dostupných zdrojů na internetu. Bylo nutné je kombinovat a strukturovat do vlastního datového souboru. Rozdělila jsem je podle jednotlivých zemí, konkrétně Sýrie, Irák, Tunis a Afghánistán. Každá z těchto zemí má vlastní CSV soubor podle země, který jsem dále rozdělila na další soubory, například "zadostSyrieAno" a "zadostSyrieNe".

V souboru "zadostSyrieAno" jsou obsažena data, která slouží k prezentaci žádostí o azyl, které jsou schváleny. Naopak v souboru "zadostSyrieNe" jsou obsaženy žádosti, které budou zamítnuty. Tato struktura souborů byla zvolena s ohledem na potřeby umělé inteligence, která se na základě těchto dat bude učit.

Ve struktuře každého CSV souboru, který jsem připravila pro trénování modelu umělé inteligence, jsem kladla velký důraz na detailní a přehlednou organizaci dat. Hlavička souboru obsahuje klíčová slova, která jednoznačně definují obsah každého sloupce, což usnadňuje navigaci v datovém souboru a zpracování dat algoritmem.

První sloupec s klíčovým slovem „zkratka“ je navržen tak, aby poskytoval rychlou a efektivní identifikaci jednotlivých žádostí. Používám systém zkratky, kde první písmeno označuje zemi původu (například „S“ pro Sýrii), následuje písmeno označující pohlaví žadatele („M“ pro muže, „W“ pro ženy), a zakončuje to unikátním identifikačním číslem, čímž vznikne kód jako „SM001“ pro syrského muže. Tento přístup zajišťuje snadnou orientaci v datech a jejich anonymizaci.

Dalším klíčovým slovem ve sloupci je „datum_narozeni“, který udává datum a rok narození žadatele. Sloupec „misto_narozeni“ dále specifikuje místo narození žadatele, zatímco „statni_prislusnost“ identifikuje jejich státní příslušnost, což je nezbytné pro zjištění původu žádostí. Poté by měl uživatel uvádět kontaktní adresu, kde se v současné situaci nachází a kam migroval. Pro tyto informace slouží „kontaktni_adresa“. V rámci anonymizace jsem se rozhodla, že všechny kontaktní adresy budou uvedeny jako „fiktivní adresa, fiktivní země“, aby byla zachována bezpečnost a soukromí fiktivních datových profilů.

Dalšími klíčovými hesly jsou „telefon“ a „email“, který žadatel musí vyplnit, jako kontaktní prostředek, což umožňuje modelovat reálnější scénáře komunikace s žadatelem o azyl, i když jsou tyto informace fiktivní.

Klíčovým prvkem v datovém souboru je sloupec „duvod_o_azyl“, který obsahuje detailní popis důvodů, proč se žadatel rozhodl opustit svou domovinu a usiluje o azyl v jiné zemi. Je nesmírně důležité, aby každý z těchto důvodů odpovídal kritériím pro udělení azylu podle zákona přijímající země. Tato část dat je nezbytná pro pochopení individuální situace žadatele a posouzení jeho nároku na ochranu.

Další důležitý sloupec, „rodinni_prislusnici“, nám dává informace o rodinném zázemí žadatele a uvádí, zda jsou jeho blízcí v podobném nebezpečí. S tím souvisí sloupec „historie_cestovani_transitu“, kde je popis cesty z rodné země do země, kde podává žádost o azyl.

Velmi významným je také sloupec „doklady_dukazy_dokumenty“, který obsahuje seznam všech dokumentů a materiálů přiložených k žádosti. Mezi tyto důkazy patří zprávy, právní dokumenty, videozáznamy, fotografie, smlouvy, dopisy a další, které slouží k podpoře tvrzení žadatele. Tyto materiály jsou fundamentální pro ověření autenticity a vážnosti důvodů pro azyl.

Závěrečným, ale neméně důležitým sloupcem, je „podepsane_prohlaseni“. Každá žádost musí být opatřena podpisem žadatele, čímž získává oficiální status a stává se závaznou pro další právní úkony. Podpis zajišťuje autentičnost podaných informací a je klíčovým prvkem v procesu vyhodnocení žádostí o azyl.

Zde je ukázka žádosti, která byla v trénovacích datech zařazena jako pozitivně hodnocená:

- **Zkratka:** SM426
- **Datum narození:** 3.9.1990
- **Místo narození:** Aleppo
- **Státní příslušnost:** Sýrie
- **Kontaktní adresa:** fiktivní adresa, fiktivní země
- **Telefon:** 123456789

- **Email:** ahmed.alsayed@email.com
- **Důvod:** Ahmed uprchl ze Sýrie kvůli pronásledování, kterému byl vystaven ze strany vlády kvůli svému odvážnému postoji proti porušování lidských práv. Boj za spravedlnost a lidskou důstojnost ho vystavil riziku, a proto se obává, že by mu hrozilo nebezpečí, pokud by se do Sýrie vrátil.
- **Rodinný příslušníci:** Ahmedova sestra se v současné době skrývá v Sýrii, protože jí rovněž hrozí pronásledování.
- **Historie cestování/transitu:** Ahmed uprchl přes Turecko a lodí se dostal do Řecka a poté do Německa a dále.
- **Doklady/Dokumenty:** Ahmed má policejní záznamy a lékařské zprávy, v nichž je podrobně popsáno pronásledování, kterému čelil v Sýrii.
- **Podepsané prohlášení:** Ano

Celkově každý ze zmíněných sloupců představuje nezbytnou součást komplexního vyhodnocování žádostí o azyl, které umožňuje umělé inteligenci zpracovávat a analyzovat data s přesností a citlivostí na kontext každé individuální situace.

V rámci trénování umělé inteligence je klíčové, aby trénovací data byla rovnoměrně distribuována mezi pozitivní (schválené) a negativní (zamítnuté) příklady. V tomto případě je každý soubor s žádostmi o azyl strukturován tak, aby obsahoval 200 příkladů, z nichž 100 by mělo být schváleno a 100 zamítnuto.

Kromě toho bylo důležité rozdělit data na trénovací a testovací množinu, kde je poměr mezi trénovacími a testovacími daty nastaven na 80:20. Tento poměr znamená, že větší část, tedy 200 trénovacích dat pro každý stát, se používá pro trénink modelu, zatímco 40 dat je vyhrazeno pro testování modelu pro každý stát. Proces je přizpůsoben tak, aby zahrnoval všechny čtyři státy (Sýrie, Irák, Tunisko a Afghánistán), což znamená, že pro každý stát se připravuje samostatná sada testovacích dat. Tento poměr byl vypočítán takto:

- 80 % pro trénink: 200 dat pro každý stát
- 20 % pro testování: $0.2 \times 200 = 40$ dat pro každý stát

Celkově pro všechny čtyři státy je potřeba 800 dat pro trénink a 160 dat pro testování. Trénovací množina slouží k vlastnímu učení modelu umělé inteligence na základě poskytnutých příkladů pro daný stát. Po trénování je testovací množina použita k ověření, jak dobře se model naučil a jak dobře dokáže klasifikovat nová, dosud neviděná data. Každá testovací množina obsahuje stejný poměr schválených a zamítnutých žádostí jako odpovídající trénovací data pro daný stát, což zajišťuje, že model je testován na relevantním a reprezentativním vzorku.

Kromě strukturovaných dat v CSV souborech jsem se rozhodla přidat ještě jednu vrstvu analýzy, která se zaměřuje na porozumění kontextu a specifikací jednotlivých zemí z pohledu bezpečnostní situace a běžných problémů. Tuto vrstvu tvoří textové soubory, každý specifický pro danou zemi, které obsahují klíčová slova

a fráze, jež vystihují typické problémy, s nimiž se obyvatelé dané země mohou potýkat. Tyto soubory slouží jako důležitý nástroj pro automatické porovnání a analýzu důvodů uvedených v žádostech o azyl s obecným kontextem problémů dané země.

Například, pro Sýrii jsem vytvořila textový soubor pojmenovaný **syrie.txt**, který obsahuje seznam slov a frází souvisejících s občanskou válkou, politickým pronásledováním a dalšími problémy, které mohou být důvody pro udělení azylu. Podobné soubory jsem připravila i pro ostatní země, s nimiž se naše analýza zabývá, jako jsou Irák, Tunisko a Afghánistán, každý s vlastním souborem reflektujícím specifické problémy, které jsou pro danou zemi charakteristické.

Tento přístup umožňuje umělé inteligenci nejen analyzovat strukturovaná data z žádostí, ale také porovnávat obsah žádostí s typickými problémy dané země, jak jsou definovány v textových souborech. Díky tomu může AI lépe rozumět kontextu žádostí a zvýšit přesnost svého hodnocení, zda jsou důvody pro azyl oprávněné a zda by žádosti měly být schváleny či zamítnuty. Vytvořením těchto souborů a integrací do procesu hodnocení žádostí o azyl tak přispívám k zefektivnění a zpřesnění rozhodovacího procesu.

4.1.3 Důvody k zamítnutí žádostí

V rámci procesu udělování azylu je zásadní, aby žadatelé předkládali validní a přesvědčivé důvody, proč hledají ochranu v hostitelské zemi. Zde jsem podrobně popsal typické příklady neuznávaných důvodů, které jsem v žádostech použila:

- **Ekonomické důvody:** Hledání lepších ekonomických příležitostí a důvody založené čistě na touze po ekonomickém zlepšení.
- **Kriminální činnost:** Útěk před stíháním za trestné činy.
- **Zdravotní problémy:** Léčba nebo péče o zdravotní stav bez spojení s pronásledováním.
- **Přírodní katastrofy:** Hledání ochrany před dopady přírodních katastrof.
- **Odmítnutí vojenské služby:** Neochota sloužit v armádě bez politického pronásledování.
- **Nedoložené podklady:** Žádost nepodložená konkrétními důkazy, pouze obecné informace.
- **Rozpory ve výpovědích a dokumentech:** Nesrovnalosti v informacích poskytnutých žadatelem.
- **Nesrovnalosti ve výpovědích:** Rozdíly v datování událostí, nekonzistentní vyprávění.
- **Nepravdivé události:** Případy, kdy žadatelé uvádějí události, které se ve skutečnosti nestaly nebo jsou značně zkreslené.

Tyto neuznávané důvody jsou klíčovými kritérii pro posouzení žádostí o azyl a pomáhají k lepší identifikaci žádostí, které jsou v procesu vyhodnocování zamítnuty.

4.2 Návrh algoritmu pro rozhodování o žádostech o azyl

V této části diplomové práce se zaměřuji na návrh algoritmu pro vyhodnocování žádostí o azyl. Pro vývoj tohoto algoritmu jsem zvolila programovací jazyk Python ve

verzi 3.12.2, který používám v prostředí Visual Studio Code. Tato volba jazyka Python se ukázala být velmi vhodná z několika důvodů. Python je vysoce oblíbený pro svou snadnou čitelnost a jednoduchost, což zjednodušuje vývoj a údržbu kódu. Díky široké paletě knihoven, jako jsou NumPy pro numerické operace, Pandas pro manipulaci s daty, a scikit-learn pro strojové učení, mohu efektivně implementovat složité analýzy a prediktivní modely. Tato robustní podpora umožňuje rychlou integraci funkcí pro zpracování textu, klasifikaci a vizualizaci dat, což je zásadní pro práci s různorodými datovými sadami z různých zemí. Algoritmus, který jsem vyvinula, je komplexním skriptem určeným k analýze žádostí o azyl z různých zemí. Na základě provedeného průzkumu jsem pro analýzu vybrala čtyři země: Sýrii, Irák, Tunisko a Afghánistán. Kód nejen zpracovává data, ale také predikuje výsledky žádostí na základě předem stanovených kritérií a poskytnutých dat.

4.2.1 Struktura projektu

Projekt, který jsem vyvinula, je strukturován tak, aby byl co nejpřehlednější a nejintuitivnější pro další vývoj a kontrolu. Veškeré zdrojové soubory jsou systematicky organizovány podle své funkčnosti. Data, která jsou používána pro trénink a testování modelu, jsou uložena v souborech s příponou `.csv` a jsou rozdělena podle zemí, přičemž každý soubor obsahuje příslušná trénovací nebo testovací data. Například `zadostSyrieAno.csv` pro pozitivní případy a `zadostSyrieNe.csv` pro negativní případy ze Sýrie.

Další klíčové soubory obsahují seznamy důvodů pro azyl a seznam stop slov, které jsou ve formátu `.txt` nebo `.json`, jako je `syrie.txt` pro důvody a `stopwords-cs.json` pro stop slova. Tyto soubory jsou nezbytné pro předzpracování textových dat a zajišťují, že analýza je co nejpřesnější. Před spuštěním algoritmu je třeba nainstalovat potřebné závislosti uvedené v souboru `requirements.txt`, který je součástí projektu.

Celý algoritmus je spuštěn přes skript `main.py`, který slouží jako vstupní bod celého projektu. Tento skript koordinuje načítání dat, jejich zpracování, spouštění analýz, vyhodnocování výsledků a vizualizaci. Struktura a pojmenování souborů jsou navrženy tak, aby zajišťovaly snadnou orientaci v projektu a umožňovaly efektivní správu a aktualizace kódu.

4.2.2 Implementace

Implementace algoritmu zahrnuje několik klíčových kroků, jako je zpracování textu, vektorizace, klasifikace, vizualizace dat a výpočty pro predikci úspěšnosti žádostí. V následujících odstavcích detailně popíši jednotlivé části kódu a metody, které jsem při vývoji použila.

Knihovny

K základním nástrojům, které jsem použila, patří knihovna `numpy`, která mi umožňuje efektivně manipulovat s datovými strukturami. Dále využívám `CountVectorizer`

zer z balíčku **scikit-learn**, který slouží k transformaci textových dat na číselnou vektorovou formu, což je klíčové pro další zpracování. Pro klasifikaci dat a predikci výsledků žádostí o azyl používám **MultinomialNB**, naivní Bayesův klasifikátor. Výsledky a analýzy pak vizualizují pomocí knihovny **matplotlib**.

Načítání dat

Klíčovou součástí zpracování dat je načítání a předzpracování souborů. Pro načítání dat jsem definovala několik funkcí:

- **load_csv_file(file_name)**: Funkce načte data z CSV souboru. Otevře soubor v režimu čtení s kódováním UTF-8, přeskočí hlavičku a data z každého řádku uloží do seznamu pro snadné zpracování.
- **load_reasons(reason_file)** a **load_stopwords(stopwords_file)**: Tyto funkce slouží k načítání specifických dat z textových nebo JSON souborů. **load_reasons** extrahuje důvody pro udělení azylu, které jsou uloženy ve formě textu, zatímco **load_stopwords** načítá seznam běžných, ale pro analýzu irelevantních slov. Oba seznamy jsou převedeny do množin pro efektivnější manipulaci a vyloučení z analýzy.

Předzpracování textu

Předzpracování textu zahrnuje několik klíčových technik, které umožňují efektivní analýzu a zpracování textových dat:

- **Lowercase**: Neboli minuskule je funkce pro převod textu na malá písmena nazývá. Tato funkce je užitečná pro efektivní práci s daty.
- **Stematizace**: Tento proces převede slova do jejich základního tvaru, což v praxi znamená redukci na kořen slova. Pro stematizaci využívám funkci **cz_stem**, která je specificky navržena pro český jazyk. Tímto způsobem je možné snížit množství variací jednotlivých slov, což značně zvyšuje přesnost další textové analýzy.
- **Vektorizace**: Pro transformaci textů na numerické reprezentace využívám knihovnu **CountVectorizer** z **scikit-learn**. Tato metoda převádí texty na vektory, kde každý element vektoru reprezentuje četnost výskytu slova ve slovníku. Vektorizace je nezbytná, protože modely strojového učení pracují s číselnými hodnotami, nikoli s přímým textem.
- **Levenshteinova vzdálenost**: Tato technika se používá k měření rozdílů mezi dvěma textovými řetězci. Implementovala jsem algoritmus, který vytváří matici vzdáleností, kde každý prvek na pozici $[i][j]$ reprezentuje vzdálenost mezi prvními i znaky řetězce s_1 a prvními j znaky řetězce s_2 . Výsledná hodnota v pravém dolním rohu matice ukazuje celkovou Levenshteinovu vzdálenost mezi oběma řetězci. Tato metoda je užitečná při porovnávání slov v textech žádostí s klíčovými slovy, která mohou naznačovat validní důvody pro udělení azylu.

Klasifikace

Klasifikace textů je realizována pomocí multinomického Naivního Bayesova klasifikátoru, což je metoda vhodná pro zpracování dat s diskrétními charakteristikami, jako je text. Model jsem natrénovala na souboru trénovacích dat, který obsahuje příklady žádostí s předem známými výsledky. Po natrénování modelu jej používám pro predikci na neznámých testovacích datech. Naivní Bayesův klasifikátor je založen na principu pravděpodobnosti a umožňuje klasifikaci textů na základě četností slov, což je ideální pro naše účely, jelikož pracujeme s textovými daty popisujícími různé situace a důvody pro udělení azylu. V kontextu tohoto kódu Multinomický Bayesovský klasifikátor klasifikuje žádosti o azyl do dvou tříd – pozitivních (žadatel splňuje podmínky pro udělení azylu) a negativních (žadatel nesplňuje podmínky pro udělení azylu). Klasifikace je založena na analýze obsahu žádostí o azyl a srovnání s trénovacími daty.

Tato metodologie mi umožňuje efektivně zpracovat a analyzovat žádosti o azyl a poskytuje pevný základ pro následné vyhodnocení a predikci výsledků těchto žádostí. V dalších částech práce se budu podrobněji věnovat výsledkům klasifikace a diskusi.

Evaluace

Součástí mého kódu je funkce **evaluate_application**, která je klíčovou součástí mého systému pro vyhodnocování žádostí o azyl. Tato funkce integruje několik pokročilých technik analýzy dat, a tak umožňuje komplexní hodnocení každé jednotlivé žádosti.

Proces začíná inicializací a přípravou nezbytných datových struktur. Následně načítám trénovací data, která mi umožňují analyzovat, jaké dokumenty a důvody jsou obvykle spojeny s úspěšnými nebo neúspěšnými žádostmi. Jednou z klíčových částí vyhodnocení je analýza přítomnosti důležitých dokumentů u každé žádosti. Pokud chybí klíčové dokumenty, jako jsou doklady, fotografie nebo podepsané prohlášení, žádost je automaticky vyhodnocena jako méně pravděpodobná k úspěchu.

Další zásadní částí mého přístupu je analýza podobnosti textu žádosti s trénovacími daty, která zahrnuje stematizaci slov a výpočet Levenshteinovy vzdálenosti. Tento proces mi umožňuje zjistit, do jaké míry text žádosti odpovídá důvodům uvedeným v databázi pozitivních případů. Přítomnost klíčových slov, která se shodují s důvody v pozitivních datech, výrazně zvyšuje šanci žádosti na úspěch. Naopak, pokud žádost obsahuje slova typická pro negativně hodnocené případy, její šance na úspěch se snižuje. Toto hodnocení je umocněno specifickými textovými soubory vytvořenými pro každý stát zvlášť, které reflektují regionální specifika a často se vyskytující důvody pro udělení, nebo zamítnutí azylu.

Integrace všech těchto faktorů do finálního hodnocení je završena výpočtem celkové pravděpodobnosti úspěchu žádosti. Algoritmus je navržen tak, aby výsledek sloužil spíše jako doporučení, nikoli jako konečné rozhodnutí, s maximální hranicí úspěšnosti nastavenou na 99 %. Toto omezení reflektuje skutečnost, že algoritmus má své limity a konečné rozhodnutí by mělo být vždy ověřeno lidským zdrojem. Vy-

sledky této funkce jsou konsolidovány do jednotného výstupu, který poskytuje ucelený pohled na úspěšnost každé jednotlivé žádosti, čímž zajišťuji, že systém je používán odpovědně a s potřebnou opatrností.

Tento úryvek kódu demonstruje, jak funkce vyhodnocuje jednotlivé žádosti a přiřazuje jim míru úspěšnosti.

```
# Proces kontroly
if data_yes[i]['doklady_dokumenty'] == "" or data_yes[i]['podepsane_prohlaseni'].lower() == "ne":
    success_rate = 0
else:
    matching_words = set()
    for word in re.findall(r'\b\w+\b', data_yes[i]['duvod_o_azytl'].lower()):
        stemmed_word = cz_stem(word) # Stemmatizace slova
        min_distance = min(levenshtein_distance(stemmed_word, reason) for reason in reasons)
        if min_distance <= 3:
            matching_words.add(word)

    # Zjištění, zda je text podobný trénovacím datům
    similar_to_training = bool(matching_counts.get(abbreviation, 0) > 0)
    # Zjištění skutečného počtu shodných slov
    matching_words_count = len(matching_words)

    # Zjištění, zda je text podobný negativním trénovacím datům
    similar_to_negative_data = any(levenshtein_distance(data_yes[i]['duvod_o_azytl'].lower(),
row['duvod_o_azytl'].lower()) <= 3 for row in data_no)

    # Vyhodnocení úspěšnosti žádosti na základě nových kritérií
    if y_predicted[i] == 0 or similar_to_negative_data:
        success_rate = 0 # Klasifikátor předpověděl negativní výsledek nebo data jsou podobná
negativním trénovacím datům -> 0%
    elif similar_to_training:
        success_rate = 70 # Data jsou shodná nebo podobná pozitivním trénovacím datům -> 70%
    else:
        success_rate = 0 # Defaultní hodnota úspěšnosti

    # Přičtení dalších procent za shodná slova
    success_rate += matching_words_count * 5
    # Omezení maximální hodnoty na 99%
    success_rate = min(success_rate, 99)
```

Vizualizace

Abych získala lepší přehled o výsledcích vyhodnocení dat, navrhla jsem vizualizaci pomocí dvou grafů, které ilustrují různé aspekty analýzy. Každý graf poskytuje unikátní pohled na specifické informace vyplývající z našich dat.

- **plot_similar_words(similar_words)**: Tato funkce vizualizuje frekvenci specifických slov z žádostí v podobě sloupcového grafu. Tímto způsobem lze snadno identifikovat, která slova jsou v kontextu žádostí o azyl nejčastější a jak se mohou lišit podle země nebo typu žádosti. Grafy jsou vytvořeny s využitím knihovny **matplotlib**, která umožňuje výstup dat v přehledné a vizuálně atraktivní formě.
- **plot_success_failure_by_country(success_rates)**: Další vizualizační funkce je schopna ilustrovat úspěšnost a neúspěšnost žádostí, rozdělených podle jednotlivých zemí. Tato vizualizace je klíčová pro pochopení geografických trendů a efektivity procesu vyhodnocování v různých regionech.

Hlavní funkce a spuštění

Klíčovou součástí mého projektu je funkce **process_and_evaluate_applications(...)**, která efektivně integruje všechny předchozí kroky vyhodnocovacího procesu. Tato funkce je zodpovědná za načítání dat, jejich zpracování, provádění predikcí a následné vyhodnocení výsledků. Je navržena tak, aby byla spouštěna pro jednotlivé země s příslušnými datovými soubory, což umožňuje přizpůsobení a relevantnost analýzy vzhledem ke specifikům v určité oblasti. Dále zahrnuje důkladnou segmentaci dat dle státní příslušnosti a pokročilé zpracování textu, které obnáší odstranění stop slov a vektorizaci textu s využitím knihovny CountVectorizer.

Kód je komplexní příklad aplikace strojového učení a textové analýzy v kontextu reálného problému, kterým je klasifikace žádostí o azyl na základě textových dat, reflektující specifické důvody a okolnosti uvedené v každé žádosti.

4.3 Analýza a vizualizace výsledků

V této podkapitole jsem provedla analýzu a vizualizaci výsledků klasifikace žádostí o azyl ze zemí Sýrie, Iráku, Tuniska a Afghánistánu. Získané výsledky jsou zaznamenány ve formě tabulek nebo obrázků, které poskytují detailní přehled úspěšnosti klasifikace každé žádosti. Tyto zobrazení jsou užitečným nástrojem pro porovnání a zhodnocení výkonu klasifikátoru v různých kontextech a pro různé země původu žadatelů o azyl.

Dále jsem provedla podrobnou analýzu testovacích dat a srovnání s výsledky klasifikace a skutečnými výsledky. Tato analýza umožňuje hlouběji porozumět chybám a úspěchům klasifikátoru a poskytuje důležité poznatky pro budoucí optimalizaci procesu klasifikace žádostí o azyl.

Sýrie

Výstup z klasifikace ukazuje, že z celkového počtu 40 žádostí o azyl, klasifikátor předpověděl 35 jako pozitivní (1) a 5 jako negativní (0). Přesnost klasifikace pozitivních žádostí je 100 %, což znamená, že všechny žádosti označené jako pozitivní

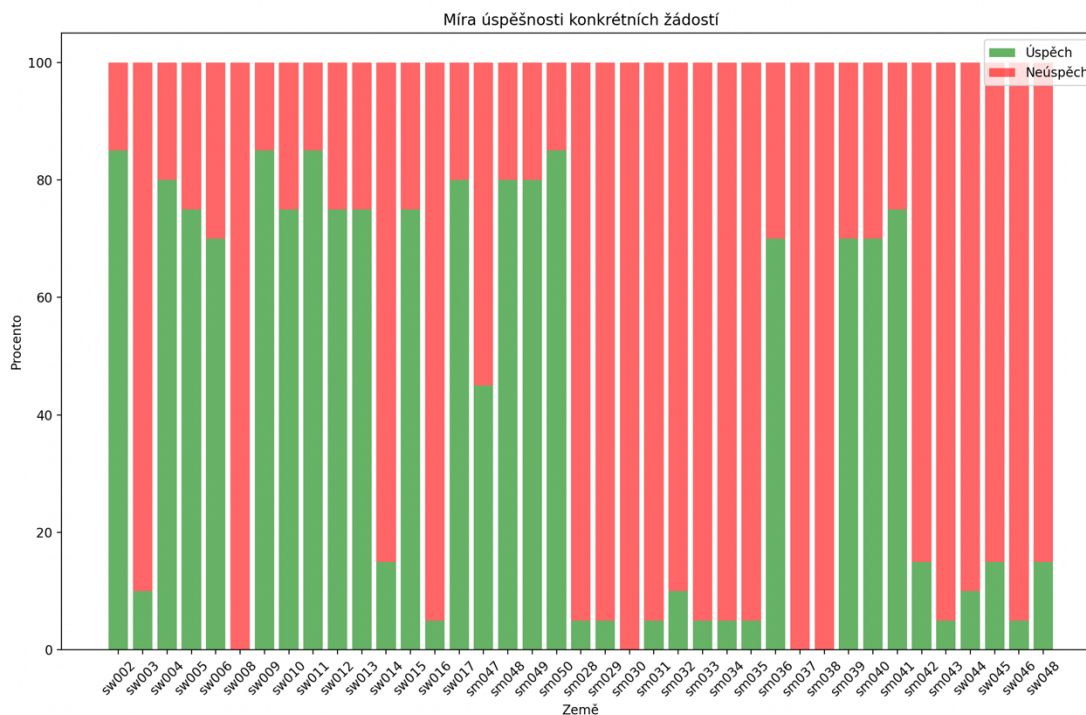
skutečně splňovaly podmínky pro udělení azylu. Nicméně, přesnost klasifikace negativních žádostí je pouze 29 %, což znamená, že pouze 29 % žádostí označených jako negativní skutečně nesplňovalo podmínky pro azyl. Celková přesnost klasifikace, která kombinuje obě třídy, je 70 %. Tento výsledek poukazuje na důležitost zvážení opatrnosti při použití tohoto algoritmu, kde přesnost není příliš vysoká.

Sýrie	Precision	Recall	F1-Score
0	0.29	1.00	0.45
1	1.00	0.66	0.79
Accuracy			0.70

Tab. 2 Výsledky klasifikace – Sýrie

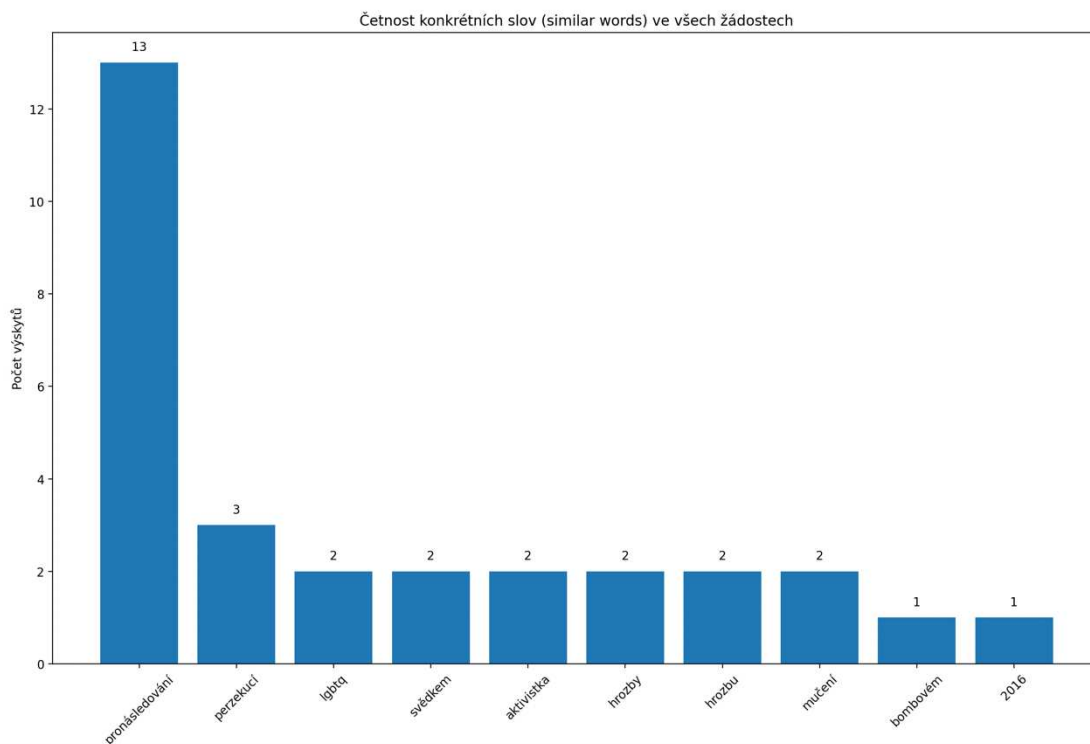
Žádost SW014 poskytuje zajímavý příklad, kdy byla vyhodnocena žádost negativně, i když obsahuje důležitá data, která by mohla být považována za důvod pro její schválení. Jak ukazuje obrázek (Obr. 6), žádost byla schválena pouze na 15 %. Algoritmus identifikoval pozitivní slova spojená s perzekucí a pronásledováním, jako jsou "isis", "perzekucí" a "pronásledování". To naznačuje, že žadatelka mohla čelit značnému nebezpečí a byla ohrožena v důsledku svého členství v menšinové náboženské skupině. Přestože předložila důkazy a informace podporující její žádost, algoritmus ji vyhodnotil negativně. Tento výsledek může být způsoben nedostatkem dostatečných dat použitých při tréninku algoritmu. Nedostatečný počet informací mohl omezit schopnost algoritmu správně identifikovat situaci a rizika, kterým žadatelka čelila.

Žádost SW003 představuje další zajímavý případ, kdy by měla být schválena, avšak algoritmus ji vyhodnotil jako negativní s pouhými 10 %. Tato žádost pochází od osoby, která opustila Sýrii kvůli útokům z důvodu své LGBTQ+ identity. Tato osoba se cítí ohrožena a obává se o svou bezpečnost, zejména v případě návratu do Sýrie. Navíc i její bratr je ohrožen z podobného důvodu. V žádosti jsou obsaženy lékařské zprávy, které detailně popisují fyzické a emoční útrapy, jimž tato osoba čelila v Sýrii kvůli své identitě. Navzdory těmto důkazům ji algoritmus vyhodnotil negativně, což může být chybné rozhodnutí, které může mít vážné následky pro žadatele. Možné vysvětlení tohoto nesprávného rozhodnutí algoritmu může být, že žádost SW003 mohla být nesprávně klasifikována jako negativní kvůli podobnosti s jinými žádostmi, které byly zařazeny mezi negativní trénovací data, avšak s pouhými drobnými odlišnostmi. Bohužel, algoritmus není dostatečně přesný na to, aby tyto nuance rozpoznal a zohlednil je při rozhodování.



Obr. 6 Úspěšnost jednotlivých žádostí v Sýrii

Na základě analýzy žádostí o azyl se ukazuje, že některá klíčová slova naznačují specifické problémy a hrozby, kterým lidé čelí ve svých domovských zemích. Mezi tato slova patří "pronásledování", "hrozby", "lgbtq", "perzekuce", "mučení" a "aktivista". Dále se v analýze vyskytují termíny spojené s konkrétními situacemi, jako jsou "isis" a "bombardování", což naznačuje specifické bezpečnostní riziko.



Obr. 7 Četnost slov žádostí v Sýrii

Slova, která jsou vyobrazena na obrázku (Obr. 7) ukazují na různé formy násilí, diskriminace a pronásledování, kterým lidé čelí ve svých domovských zemích. To je motivuje k hledání bezpečí jinde. Problémová slova naznačují, že mnoho lidí uteklo z politického pronásledování, náboženské diskriminace, sexuálního násilí nebo terorismu organizací jako ISIS. Tato slova jsou klíčová pro pochopení motivací a potřeb žadatelů o azyl. Ukazují na rozsáhlé bezpečnostní problémy a otázky lidských práv, které mohou vést k opuštění domovské země a hledání ochrany v jiných zemích.

Irák

V kontextu souboru dat z Iráku, klasifikátor poskytuje podrobné vyhodnocení úspěšnosti klasifikace jednotlivých žádostí. Například, žádosti s identifikátory označenými jako "0" byly klasifikovány s úspěšností 91 %. To znamená, že klasifikátor správně identifikoval téměř všechny žádosti jako negativní (nevyhovující podmínkám pro udělení azylu) s přesností 91 %. Dále se uvádí, že všechny tyto žádosti měly plnou podporu v testovacím souboru, což znamená, že všechny žádosti byly správně klasifikovány.

Naopak žádosti s identifikátory označenými jako "1" byly klasifikovány s úspěšností 100 %. To znamená, že klasifikátor správně identifikoval všechny žádosti jako pozitivní (splňující podmínky pro udělení azylu) s přesností 100 %. Stejně jako u

žádostí označených jako "0", i zde byla plná podpora v testovacím souboru, což znamená, že všechny žádosti byly správně klasifikovány.

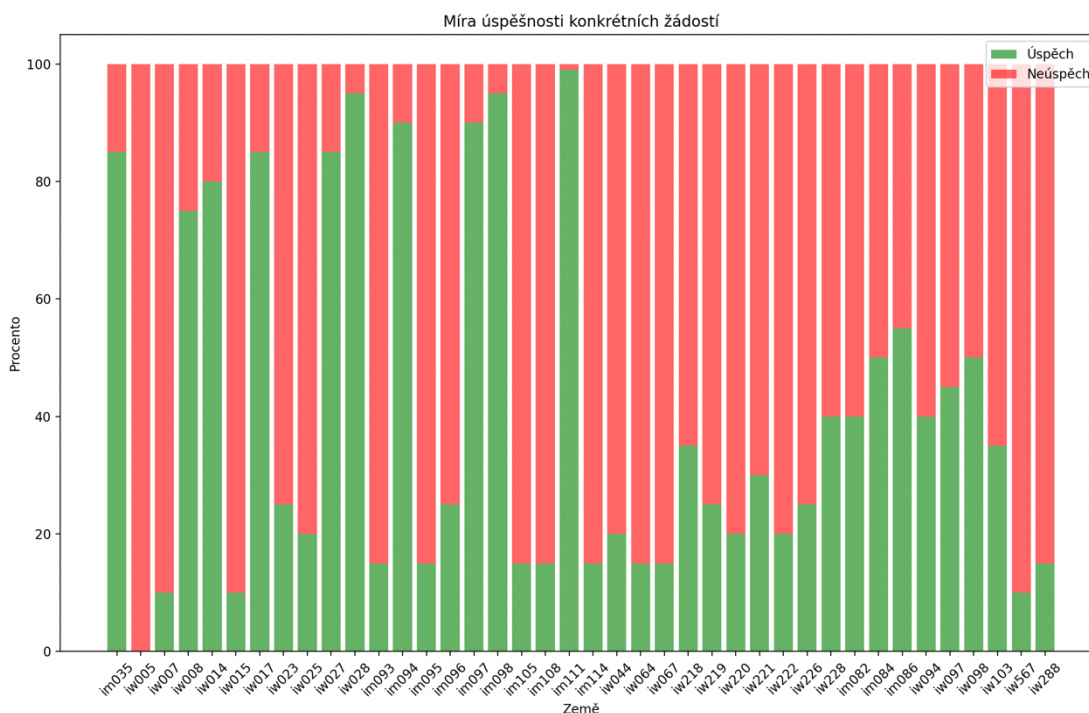
Celková přesnost klasifikace v testovacím souboru dosáhla hodnoty 95 %.

Irák	Precision	Recall	F1-Score
0	0.91	1.00	0.95
1	1.00	0.90	0.95
Accuracy			0.95

Tab. 3 Výsledky klasifikace – Irák

Žádosti IW023 a IW025 poskytují zajímavé příklady, kdy byly vyhodnoceny negativně, přestože obsahují podobná klíčová slova týkající se pronásledování a násilí ze strany ISIS. Oba žadatelé čelili vážnému ohrožení svého života jako členové jezídské komunity v Iráku. Jejich rozhodnutí hledat mezinárodní ochranu je zřejmě založeno na oprávněných obavách o svou bezpečnost a životní situaci. Přestože předložili dokumentaci o pronásledování a násilí, algoritmus je doporučil ke kontrole lidským faktorem.

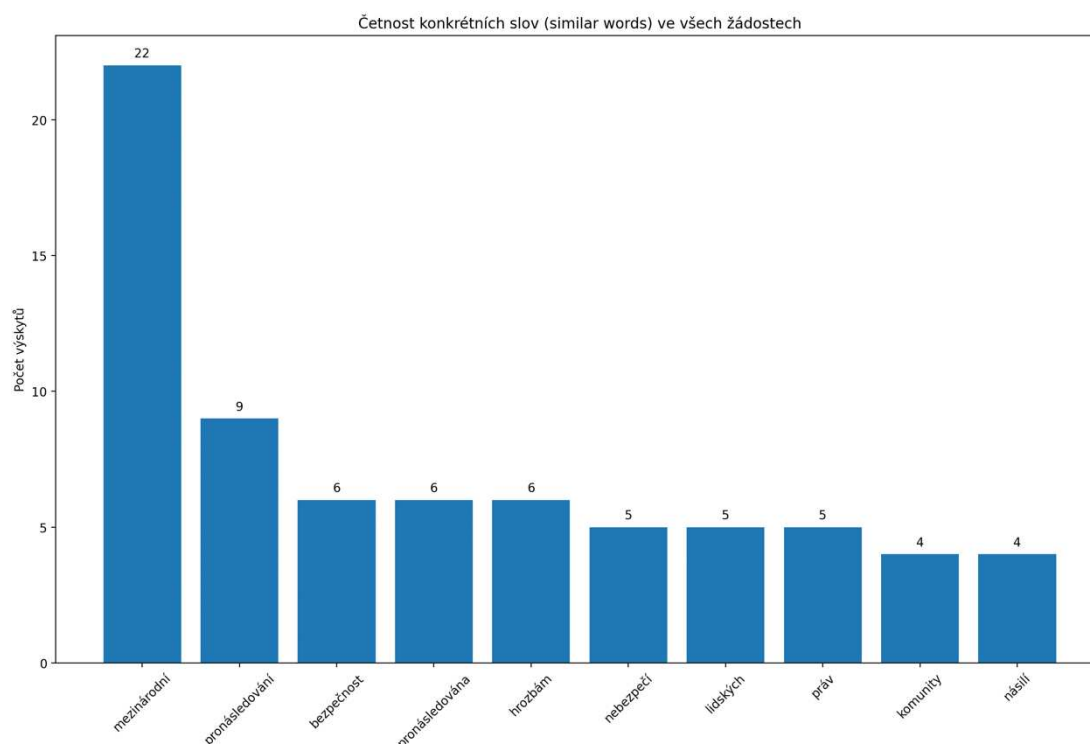
Žádost IM096 představuje další příklad, kdy byla žádost vyhodnocena negativně, i když měla být schválena. Důvodem žádosti bylo násilí a perzekuce, kterým žadatelka čelila jako žena a příslušník menšinové skupiny v Iráku. Hrozby a útoky na její život ji vedly k hledání azylu v jiné zemi, kde by mohla žít bezpečně a bez obav. Navíc i ostatní členové její rodiny, včetně sestry a mladé neteře, jsou v ohrožení. Algoritmus označil tuto žádost za potřebnou kontroly lidským faktorem, což poukazuje na nutnost detailního posouzení případu v případech, kdy dochází k shodě klíčových slov a důležitých informací.



Obr. 8 Úspěšnost jednotlivých žádostí v Iráku

Tyto výsledky ukazují na důležitost lidského posouzení v procesu hodnocení žádostí o azyl. I přesto, že algoritmus může být efektivním nástrojem pro rychlé zpracování velkého množství dat, rozhodování o osudu jednotlivých žadatelů vyžaduje lidskou empatii a schopnost porozumět jejich individuálním situacím a potřebám.

Ve 14 případech žádostí o azyl v Iráku bylo doporučeno provést důkladné posouzení lidským faktorem. Tato doporučení byla založena na výskytu významného množství shodných slov v žádostech, což může poukazovat na důvěryhodnost těchto žádostí. Nicméně, shoda slov sama o sobě nemusí vždy znamenat správnost žádosti. Je klíčové, aby každá žádost byla pečlivě posouzena lidským faktorem, a to bez ohledu na výsledky, které mohou vyplývat z hodnocení provedeného pomocí umělé inteligence. Tímto způsobem lze zajistit, že rozhodnutí o udělení azylu bude spravedlivé a pravdivé pro každého žadatele.



Obr. 9 Četnost slov žádostí v Iráku

Z analýzy poskytnutých informací na obrázku (Obr. 9) vyplývá, že žadatelé o azyl čelí různým formám pronásledování, násilí a ohrožení v jejich domovských zemích. Klíčová slova obsažená v žádostech o azyl naznačují porušování lidských práv, politické a náboženské perzekuce, útoky na novináře a aktivisty a další formy bezpečnostního rizika. Tato slova ukazují na komplexní bezpečnostní situaci a nedostatek ochrany v těchto zemích, což žadatele nutí hledat azyl v jiných zemích, kde mohou žít bezpečněji a bez obav.

Tunis

V kontextu souboru dat z Tuniska poskytuje klasifikátor podrobné vyhodnocení úspěšnosti klasifikace jednotlivých žádostí. Z přiloženého výstupu klasifikace vidíme, že z celkového počtu 40 žádostí o azyl bylo 14 klasifikováno jako negativní (nevyhovující podmínkám pro udělení azylu) a 26 jako pozitivní (splňující podmínky pro udělení azylu).

Pro negativní žádosti (označené jako "0") klasifikátor dosáhl přesnosti 70 % a úplnosti (recall) 100 %. To znamená, že z 14 negativně klasifikovaných žádostí bylo 100 % z nich správně identifikováno jako negativní. F1-skóre pro tuto třídu je 0.82, což je harmonický průměr precision a recall.

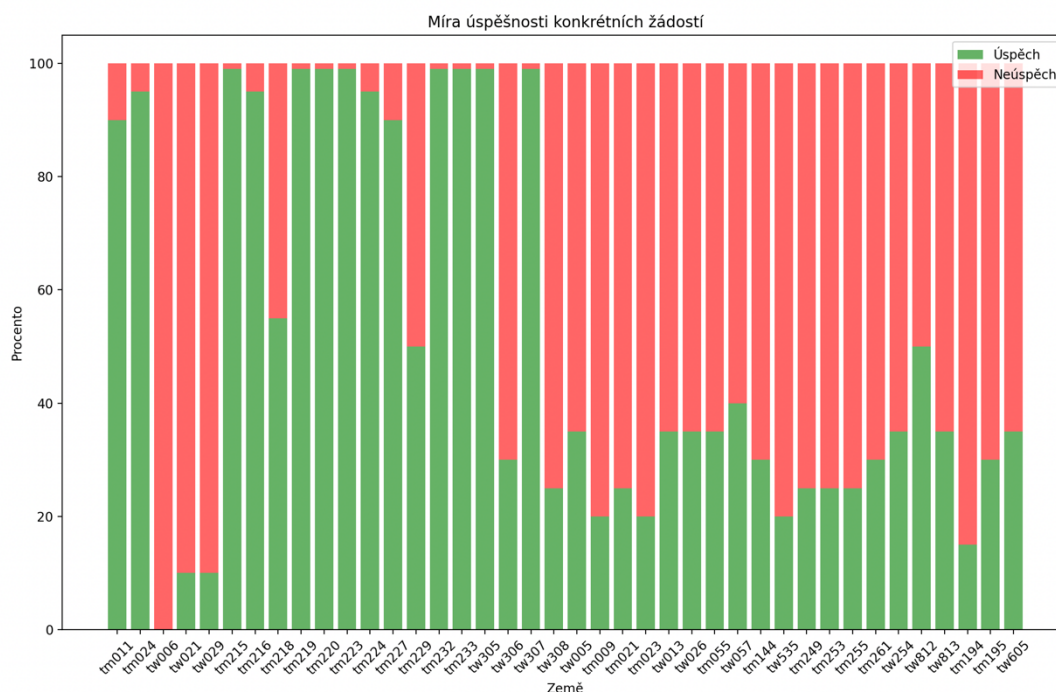
Pro pozitivní žádosti (označené jako "1") dosáhl klasifikátor přesnosti 100 % a úplnosti 77 %. To znamená, že ze 26 pozitivně klasifikovaných žádostí bylo 100 % z nich správně identifikováno jako pozitivní. F1-skóre pro tuto třídu je 0.87.

Celková přesnost klasifikace v testovacím souboru dosáhla hodnoty 85 %.

Tunisko	Precision	Recall	F1-Score
0	0.70	1.00	0.82
1	1.00	0.77	0.87
Accuracy			0.85

Tab. 4 Výsledky klasifikace – Tunisko

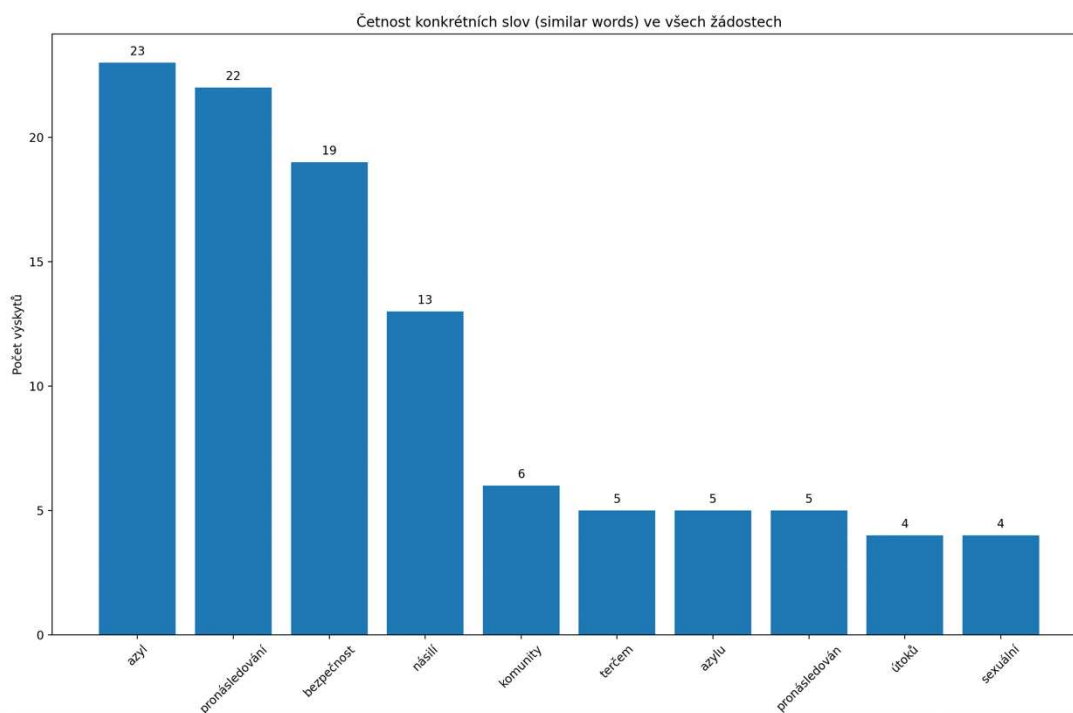
To naznačuje, že klasifikační algoritmus je schopen správně identifikovat většinu žádostí a poskytnout jim adekvátní hodnocení. Je nutné brát v úvahu, že i přesnost algoritmu není stoprocentní a stále existují situace, kdy může docházet k chybám ve vyhodnocování. V případě Tuniska algoritmus pravděpodobně přistupoval klasifikaci striktněji, což vedlo k vyššímu počtu negativně vyhodnocených žádostí než pozitivně hodnocených, což je patrné zobrazením v obrázku (Obr. 10).



Obr. 10 Úspěšnost jednotlivých žádostí v Tunisku

Jako příklad udáváním žádost TW006 o azyl, která byla označena jako negativní, i přesto, že obsahovala důvěryhodné informace o pronásledování ze strany bývalého manžela a jeho rodiny. Přestože obsahuje důležité informace a důkazy, jako jsou lékařské záznamy a policejní zprávy, klasifikační algoritmus ji vyhodnotil jako negativní. Toto rozhodnutí může být způsobeno nedostatkem informací nebo počtu vzorových trénovacích dat, což vedlo k chybné klasifikaci žádosti.

Další zajímavý případ představuje žádost o azyl TM261 od Karima Benjemiai. Původně tvrdil, že utekl z důvodu pronásledování kvůli své sexuální orientaci. Nicméně podrobnější analýza odhalila, že skutečným důvodem jeho útěku byla nespokojenost se svou prací a pracovními vztahy. I když žádost obsahovala klíčové prvky, jako je zmínka o "pronásledování" a "azyly", nedostatek potřebné dokumentace a důkazů vedl k její negativní klasifikaci v rámci klasifikačního algoritmu. Tento příklad ukazuje, že i při zdánlivé shodě s pozitivními žádostmi může nedostatek informací nebo dokumentace vést k negativní klasifikaci.



Obr. 11 Četnost slov žádostí v Tunisku

Analýza dat četnosti slov v důvodech ukazuje na obrázku (Obr. 11), že většina žádostí o azyl z Tuniska se opírá o několik klíčových slov a témat. Nejčastějšími tématy jsou "azyl", "pronásledování", "bezpečnost" a "násilí". Tyto termíny naznačují hlavní důvody, proč lidé hledají azyl, a poukazují na oblasti, ve kterých se cítí ohroženi ve své domovské zemi. Další častá slova zahrnují "komunity", "útoky", "aktivista", "politický", "sexuální" a "domácí". Tato slova naznačují různé okolnosti a motivace, které vedou k žádosti o azyl, včetně politické perzekuce, genderového násilí

a násilí ve společnosti obecně. Je zajímavé, že slova jako "lidských práv", "demokracie" a "menšinového" se vyskytují jen v několika žádostech. To může naznačovat, že tato témata nejsou tak častá nebo nejsou považována za primární důvody žádosti o azyl z Tuniska. Celkově lze konstatovat, že analýza četnosti slov poskytuje užitečný náhled na hlavní motivace a témata, která vedou k žádostem o azyl z Tuniska, a může pomoci při pochopení specifických potřeb a obav žadatelů o azyl z této země.

Afghánistán

V kontextu souboru dat z Afghánistánu poskytuje klasifikátor podrobné vyhodnocení úspěšnosti klasifikace jednotlivých žádostí. Z přiloženého výstupu klasifikace vidíme, že z celkového počtu 40 žádostí o azyl bylo 17 klasifikováno jako negativní (nevyhovující podmínkám pro udělení azylu) a 23 jako pozitivní (splňující podmínky pro udělení azylu).

Pro negativní žádosti (označené jako "0") klasifikátor dosáhl přesnosti 85 % a úplnosti (recall) 100 %. To znamená, že ze 17 negativně klasifikovaných žádostí bylo 100 % z nich správně identifikováno jako negativní. F1-skóre pro tuto třídu je 0.92, což je harmonický průměr precision a recall.

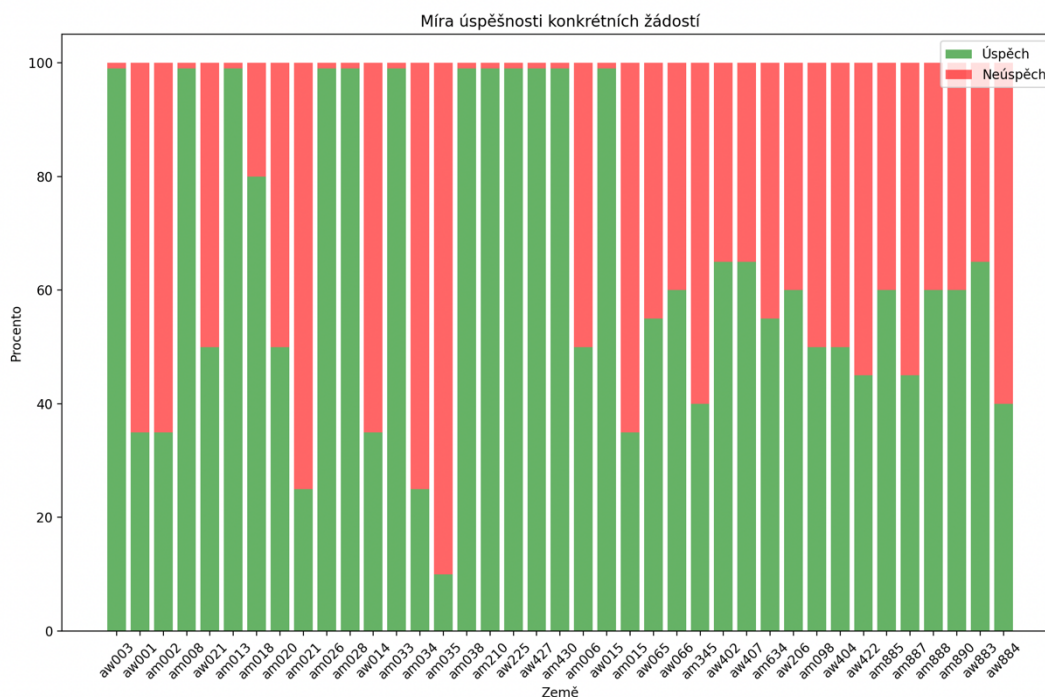
Pro pozitivní žádosti (označené jako "1") dosáhl klasifikátor přesnosti 100 % a úplnosti 87 %. To znamená, že z 23 pozitivně klasifikovaných žádostí bylo 100 % z nich správně identifikováno jako pozitivní. F1-skóre pro tuto třídu je 0.93.

Celková přesnost klasifikace v testovacím souboru dosáhla hodnoty 93 %.

Afghánistán	Precision	Recall	F1-Score
0	0.85	1.00	0.92
1	1.00	0.87	0.93
Accuracy			0.93

Tab. 5 Výsledek klasifikace – Afghánistán

Z analýzy výsledků klasifikace žádostí o azyl z Afghánistánu vyplývá, že klasifikační algoritmus dosáhl vysoké úspěšnosti u většiny žádostí. Nicméně, existuje značný podíl žádostí, které byly nekvalitně klasifikovány jako negativní nebo pozitivní. To naznačuje, že klasifikátor příliš striktně vyhodnotil tyto žádosti a přiřadil jim negativní hodnocení, i když mohly být potenciálně pozitivní.



Obr. 12 Úspěšnost jednotlivých žádostí v Afghánistánu

Z obrázku (Obr. 12) vyplývá, že 25 žádostí z celkového počtu 40 bylo klasifikováno jako negativní, ačkoliv jejich úspěšnost byla menší než 70 %. Tato skutečnost naznačuje potřebu dalšího zkoumání a optimalizace klasifikačního algoritmu, aby bylo dosaženo vyváženějšího a přesnějšího hodnocení.

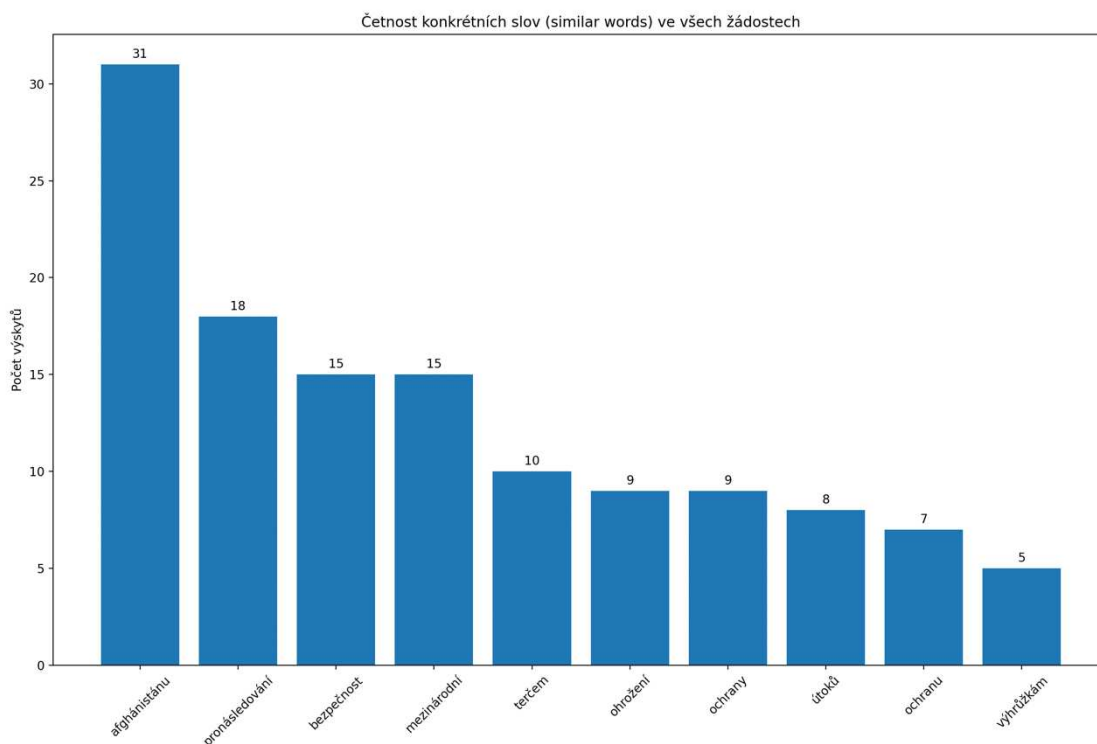
Žádost s identifikátorem AM035 byla vyhodnocena s úspěšností 10,00 % a byla klasifikována jako negativní. Tato žádost obsahuje detailní popis situace žadatelky o azyl, která uvádí neustálé hrozby, obtěžování a fyzické násilí, které utrpěla v Afghánistánu kvůli svému boji za práva žen a genderovou rovnost. Žadatelka žádá o azyl v Evropě s cílem získat mezinárodní ochranu a uniknout stálému nebezpečí. Přestože žádost obsahuje silné argumenty a zdůvodnění pro žádost o azyl, byla vyhodnocena jako negativní kvůli nízké úspěšnosti 10,00 %. Příčinou tohoto negativního vyhodnocení může být podobnost s jinými negativně vyhodnocenými žádostmi v trénovacím modelu. Tyto žádosti, které byly označeny jako negativní, mohou mít podobné charakteristiky jako žádost AM035, ačkoliv skutečně mohou být oprávněné.

Dalším příkladem je žádost s identifikátorem AW001, která má úspěšnost klasifikace 35 % a obsahuje sedm shodných slov s daty. Tato žádost byla přesto klasifikována jako negativní. Žadatelka se uchází o azyl z důvodu pronásledování ze strany své vlastní rodiny. Jako členka LGBT komunity čelila zneužívání, násilí a výhrůžkám kvůli své sexuální orientaci. Má obavy o svou bezpečnost a život v případě návratu

do Afghánistánu. Přestože poskytla důkladné dokumenty jako lékařské zprávy, psychologické posudky a výpovědi svědků, kteří mohou dosvědčit zneužívání a nebezpečí, jimž v Afghánistánu čelila kvůli své sexuální orientaci, byla její žádost označena jako negativní. Toto rozhodnutí může být důsledkem nízké úspěšnosti klasifikace a podobnosti s jinými negativně klasifikovanými žádostmi v trénovacím modelu.

Opačným příkladem je žádost s identifikátorem AW015, která dosáhla úspěšnosti klasifikace 99 % a obsahuje sedm shodných slov. Tato žádost byla klasifikována jako pozitivní, ale měla být hodnocena jako negativní. Přestože data obsažená v žádosti mohou být podobná některým pozitivně klasifikovaným žádostem a mohou obsahovat velký počet shodných slov, klasifikační algoritmus vyhodnotil tuto žádost jako pozitivní.

Analýza žádostí o azyl z Afghánistánu odhalila četné klíčové termíny a fráze, které se v těchto žádostech často objevují. Tyto termíny poskytují důležitý vhled do motivací a důvodů, proč lidé žádají o azyl, a odkazují na specifické výzvy a hrozby, kterým čelí v jejich domovské zemi. Z těchto dat byl vytvořen graf zobrazující podobné slova v každé žádosti.



Obr. 13 Četnost slov žádostí v Afghánistánu

Při analýze těchto dat jsem zjistila, že mnoho žadatelů o azyl uvádí jako důvod své žádosti diskriminaci, pronásledování a útoky v Afghánistánu. Tato data z obrázku (Obr. 13) také zobrazují, že žadatelé o azyl často zdůrazňují své úsilí o ochranu lidských práv, zejména práv žen a genderovou rovnost. Mezi další často

zmiňované motivy patří obavy o osobní bezpečnost, napadání náboženskými extremisty a obavy o budoucnost v rámci konfliktního prostředí. Například některé žádosti zmiňují násilí a pronásledování kvůli sexuální orientaci (AW001, AW015), útoky na aktivisty za lidská práva (AW021, AW407) nebo obavy z teroristických útoků a politické nestability (AM026, AM210). Jako ilustraci pozitivního příkladu jsem vybrala žádost o azyl s identifikátorem AM002. Tato žádost obsahuje klíčová shodná slova, jako jsou "tlumočnick", "zahraničními", "Afghánistánu", "pronásledování", "Talibanu", "bezpečnost", "útoků" a "terčem". Tyto informace poskytují cenný kontext pro pochopení individuálních příběhů žadatelů o azyl. Přestože žádost obsahuje tyto důležité prvky, byla nečekaně vyhodnocena jako negativní v rámci klasifikačního modelu.

4.4 Zhodnocení výsledků

V této kapitole jsem úspěšně vyvinula algoritmus pro vyhodnocení žádostí o azyl, který byl zaměřen na žadatele z vybraných zemí. Kvalita a účinnost algoritmu byly podpořeny důkladným průzkumem problematiky a vytvořením simulovaného datasetu, který reflektoval kritéria nutná pro udělení azylu. Analýza pomocí tohoto algoritmu ukázala jeho schopnost s celkovou úspěšností přibližně 85 %, což znamená, že dokáže přesně klasifikovat žádosti o azyl, zejména z Iráku a Afghánistánu s úspěšností 95 % a 93 % podle informací zobrazených v tabulce (Tab. 6). Výsledky klasifikace byly prezentovány pomocí grafů a obrázků, což ukázalo důležitost zahrnutí lidského úsudku do procesu rozhodování v situacích, kdy mohou být interpretace výsledků nejednoznačné. Tím se zdůraznilo, že lidská intuice je klíčová při vyhodnocování komplexních dat a rozhodování v podmínkách, kdy automatizované systémy mohou selhat.

Zemé	Precision (0)	Recali (0)	FI-Score (0)	Precision (1)	Recali (1)	FI-Score (1)	Accuracy
Sýrie	0.29	1.00	0.45	1.00	0.66	0.79	0.70
Irák	0.91	1.00	0.95	1.00	0.90	0.95	0.95
Tunisko	0.70	1.00	0.82	1.00	0.77	0.83	0.85
Afghánistán	0.85	1.00	0.92	1.00	0.87	0.93	0.93

Tab. 6 Souhrnný přehled výsledků úspěšnosti algoritmu

Přestože algoritmus prokázal vysokou míru přesnosti, je důležité mít na paměti, že není dokonalý a byl trénován pouze na základě 800 dat. Pro dosažení robustnějšího a spolehlivějšího vyhodnocovacího systému by bylo nezbytné rozšířit dataset o desítky tisíc dalších datových bodů. V současné době dochází k chybám

způsobeným podobnostmi mezi žádostmi o azyl v trénovacích datech, které jsou určeny ke schválení nebo zamítnutí. Tyto chyby jsou často spojeny s nepatrnými změnami v textu nebo přidáním vyvracejících dokumentů. V takových případech je doporučeno vyhodnocení zkontrolovat lidským faktorem pro zajištění správnosti rozhodnutí.

Tato práce zdůrazňuje zásadní význam integrace strojového učení a lidského hodnocení v procesu vyřizování žádostí o azyl. Zjistila jsem, že spojení těchto dvou prvků je klíčové pro zajištění správnosti a spravedlnosti rozhodování. Moje zjištění také poukazují na nutnost neustálých inovací a vylepšení těchto procesů, což je nezbytné pro zajištění efektivního a spravedlivého vyřizování azylu v budoucnosti.

5 Diskuse

V této kapitole se zabývám porovnáním výsledků svého výzkumu s existující literaturou a diskutuji o etických aspektech a potenciálu dalšího výzkumu v oblasti využití umělé inteligence pro vyhodnocování žádostí o azyl. Cílem je nejen posoudit shodu mých zjištění s již publikovanými studiemi, ale také identifikovat klíčové oblasti, ve kterých by mohl být výzkum rozšířen.

5.1 Etická stránka problému

V této části diplomové práce se budu zabývat možným zneužitím algoritmu z etického hlediska. Algoritmy mohou být manipulovány nebo nesprávně využívány, což vyvolává závažné otázky o možných rizicích a etických důsledcích jejich použití. Tato problematika je obzvláště kritická v oblasti azylu, kde rozhodnutí algoritmu může mít výrazný dopad na životy jednotlivců. Z tohoto důvodu je zásadní prozkoumat, jak lze zajistit, aby bylo technologické řešení vždy používáno s nejvyšší odpovědností a respektem k lidským právům.

S ohledem na vývoj mého algoritmu `evaluate_application`, jsem se v této podkapitole zaměřila na etické aspekty využití umělé inteligence pro vyhodnocování žádostí o azyl. Inspirací mi byly studie autorů jako Gavaghan a kolektiv (2019) a Samuels a Caselli (2021). Tyto zdroje zdůrazňují důležitost přesnosti, transparentnosti, lidské kontroly, ochrany soukromí a eliminace předpojatosti. Tyto klíčové prvky jsem se snažila reflektovat ve svém nástroji, což je zásadní pro minimalizaci rizika zneužití a zajištění, že rozhodování zůstává spravedlivé a eticky odpovědné.

Spravedlnost a rovnost přístupu

Důraz na spravedlnost a rovnost přístupu je zásadní. Při vývoji mého algoritmu jsem byla ovlivněna kritickými poznámkami Gavaghana a kolektivu (2019), kteří upozornili na rizika spojená s neobjektivními datovými sadami. Práce na tomto projektu mi umožnila hlubší pochopení toho, jak důležité je, aby data, na kterých je algoritmus trénován, neodrážela žádné předsudky nebo diskriminaci. Jsem si vědoma toho, že návrh algoritmů musí být prováděn s maximální opatrností, aby nedocházelo k nerovnosti nebo diskriminaci žadatelů o azyl.

Transparentnost a odpovědnost

Transparentnost a odpovědnost jsou pro mě klíčové. Samuels a Caselli (2021) navrhují transparentní postupy a mechanismy pro kontrolu a vysvětlení rozhodnutí, což jsem začlenila do svého nástroje. Můj algoritmus poskytuje detailní informace o tom, jak jsou data zpracována a jak dochází k rozhodnutím. Je nezbytné pravidelně provádět audit kritérií a vah použitých v modelu s cílem eliminovat jakoukoli formu zkreslení.

Ochrana osobních údajů a soukromí

Osobní údaje žadatelů jsou velmi citlivé informace. Kinchin a Mougouei (2022) upozorňují na potřebu pečlivého zacházení s těmito daty. Můj algoritmus je navržen tak, aby minimalizoval jakékoliv riziko úniku osobních informací. Použila jsem anonymizaci žadatelů, které chrání citlivé informace.

Empatie a lidský faktor

Přestože algoritmy mohou zefektivnit proces rozhodování, nesmí se zapomínat na lidský rozměr a empatii, která je v procesu nezbytná. Jak Kammerer, Bruni a Ruedin (2022) zdůrazňují, konečná rozhodnutí musí být vždy podpořena lidskou intervencí. Můj nástroj umožňuje, aby konečné rozhodnutí bylo vždy přezkoumáno člověkem, což zajišťuje, že každá žádost je posuzována s pochopením a soucitem.

Důvěryhodnost a nezávislost

Zajištění důvěryhodnosti a nezávislosti algoritmu je klíčové pro jeho etické použití, zvláště v citlivé oblasti, jakou je rozhodování o azylu. Ve svém výzkumu jsem identifikovala potřebu nezávislého ověřování algoritmu, jak upozorňují Kinchin a Mougouei (2022). Tito autoři zdůrazňují, že algoritmy by měly být pravidelně prověřovány externími odborníky, aby byla zajištěna jejich objektivita a spravedlnost.

Respektování lidského práva na azyl

Vývoj mého nástroje **evaluate_application** byl řízen principy a doporučeními, které jsou v souladu s mezinárodním právem a základními lidskými právy. Inspirací mi byly studie a doporučení odborníků, jako jsou Kinchin a Mougouei (2022) a Kammerer, Bruni a Ruedin (2022), kteří se zabývají integrací etických principů do AI systémů v kontextu azylu.

Kladu důraz na to, aby rozhodnutí generovaná mým algoritmem byla nejen spravedlivá a objektivní, ale také aby respektovala právo jednotlivce na azyl. Zahrnula jsem do algoritmu mechanismy, které zajišťují, že každé rozhodnutí respektuje Úmluvu o právním postavení uprchlíků. To zahrnuje zásady nediskriminace a objektivního posouzení každé žádosti, což je základním pilířem pro zajištění spravedlnosti a ochrany lidských práv v procesu rozhodování o azylu.

S ohledem na důležitost těchto etických prvků je jasné, že etické řízení využití umělé inteligence v procesu rozhodování o azylu má zásadní význam. Snažila jsem se, aby můj algoritmus nejen zvyšoval efektivitu a přesnost rozhodování, ale aby také respektoval základní lidské hodnoty. Můj záměr je, aby tento nástroj byl využíván s nejvyšší možnou mírou spravedlnosti, transparentnosti a citlivosti k lidským právům, což je podle mého názoru nezbytné pro zachování důstojnosti a práv všech žadatelů o azyl.

5.2 Možnosti dalšího výzkumu v této oblasti

V oblasti vyhodnocování žádostí o azyl pomocí umělé inteligence se nacházíme na nevyzkoumané půdě. Zatímco umělá inteligence nabízí slibné možnosti, ještě není žádná země kompletně vybavena systémem, který by tento proces úplně zautomatizoval. Tato oblast je nesmírně složitá a zahrnuje mnoho faktorů, které jsou s tímto tématem spojeny. Tento komplexní proces zahrnuje mnoho faktorů, které lze zkoumat a zdokonalovat. Například, technologie umělé inteligence může být využita k autentizaci žadatelů o azyl pomocí různých biometrických identifikátorů, jako jsou otisky prstů, dentální ověření a další, nebo jejich kombinace, včetně hlasového a obličejového rozpoznávání či ověřování úředně ověřených dokumentů.

Další možností pro výzkum je využití umělé inteligence k řízení celého procesu vyhodnocování žádostí o azyl. Tím lze zlepšit efektivitu a rychlost posuzování případů a současně poskytnout žadatelům veškeré potřebné informace a podporu během procesu. Umělá inteligence může například pomáhat žadatelům s vyplněním formulářů, poskytovat důležité informace o jejich právech a povinnostech a generovat odpovědi na jejich dotazy.

Dále lze zkoumat možnosti začlenění žadatelů do nového prostředí a společnosti pomocí technologií jako jsou chatboti, mobilní aplikace a další. Tím lze podpořit integraci žadatelů do nového prostředí a poskytnout jim přístup k potřebným službám a informacím. Při dalším výzkumu je důležité brát v úvahu aspekty jako ochrana osobních údajů, transparentnost, bezpečnost a kvalifikovanost. Je nezbytné, aby nové technologie respektovaly práva a potřeby žadatelů o azyl a zároveň byly schopny poskytnout spravedlivé a objektivní rozhodnutí.

Kvalitní a pravdivá trénovací data jsou základním stavebním kamenem pro úspěšné využití umělé inteligence při vyhodnocování žádostí o azyl. Algoritmus se na základě těchto dat učí správně rozhodovat. Aby bylo zajištěno, že se algoritmus neustále zdokonaluje a stává se přesnějším, je klíčové zařadit do trénovacích dat nově vyhodnocené žádosti, které prošly ověřením lidským faktorem. Tyto žádosti by měly být klasifikovány jako pozitivní nebo negativní a začleněny do databáze. Dále, s přechodem žádostí o azyl do digitální formy, je důležité vytvořit uživatelsky přívětivé a intuitivní grafické rozhraní. Toto rozhraní by mělo umožňovat rychlý a snadný přístup ke klíčovým informacím a nástrojům potřebným pro rozhodovací proces příslušných orgánů.

Dále je nezbytné zdůraznit význam databáze, která by sloužila jako úložiště veškerých potřebných dat. Tato databáze by měla být navržena tak, aby byla škálovatelná a zabezpečená, a zahrnovala by různé typy informací potřebných pro vyhodnocování žádostí o azyl. To může zahrnovat osobní údaje žadatelů, informace o jejich životní situaci, předchozí zkušenosti a další relevantní údaje.

Důležité je také zajistit, aby proces sběru dat byl transparentní a zabezpečený, a aby žadatelé měli jasné a jednoduché možnosti, jak poskytnout potřebné informace. Zavedení online žádostí může být efektivním způsobem, jak usnadnit tento proces a zajistit, aby data byla kompletní a přesná. Takový systém by mohl umožnit žadatelům vyplnit žádost online a automaticky odeslat data do centrální databáze pro další vyhodnocování.

Dalším krokem může být implementace mechanismů pro zajištění kvality dat a ochranu osobních údajů žadatelů. To může zahrnovat validaci a ověřování poskytnutých informací a zavedení přísných bezpečnostních opatření pro ochranu dat před neoprávněným přístupem a zneužitím. Takové opatření by pomohlo zajistit důvěryhodnost a integritu dat v celém procesu vyhodnocování žádostí o azyl.

Celkově lze konstatovat, že oblast vyhodnocování žádostí o azyl pomocí umělé inteligence je plná možností a výzev, které vyžadují další zkoumání a inovace. Je nezbytné pokračovat v průzkumu této problematiky s cílem dosáhnout spravedlivých, transparentních a lidsky respektujících procesů.

6 Závěr

V závěru této diplomové práce jsem zkoumala, jak umělá inteligence ovlivňuje procesy rozhodování o udělení azylu a jaké jsou její výhody a nevýhody. Z literární rešerše vyplývá, že umělá inteligence nabízí různé přístupy, včetně prediktivní analýzy, rozhodovacích systémů, biometrické identifikace, chatbotů a jazykové analýzy, které mohou posílit azylové procesy.

Mé vlastní zkušenosti s vytvořením algoritmu s umělou inteligencí pro vyhodnocování žádostí o azyl potvrzují, že tato technologie může výrazně zrychlit hodnocení žádostí a zlepšit organizaci dat. Na základě literární rešerše a mé vlastní praxe lze konstatovat, že technologie umělé inteligence má potenciál výrazně zlepšit efektivitu a spravedlnost rozhodování v procesech udělování azylu. Nicméně je nezbytné provést další zkoumání a zajistit transparentnost, etiku a zapojení relevantních stran, aby bylo dosaženo optimálních výsledků. Jak uvedl Marr je zásadní také zvýšit množství dostupných dat, ze kterých se umělá inteligence učí, jelikož to je klíčové pro rozhodování a přesnost algoritmu.

V průběhu výzkumu jsem identifikovala i několik zásadních nevýhod a výzev spojených s používáním AI. Mezi hlavní patří nedostatek transparentnosti algoritmů, což může vést k zneužití a komplikacím v interpretaci jejich rozhodnutí. Tento nedostatek transparentnosti ztěžuje ověřitelnost výsledků, což může ohrozit spravedlnost rozhodovacích procesů. Další výzvou je neosobní povaha rozhodování AI, která může postrádat empatii a hluboké porozumění jedinečným situacím žadatelů o azyl.

Zejména znepokojující je riziko, že se AI systémy stávají takzvanými „černými skříňkami“, kde procesy rozhodování, které algoritmy používají, nejsou plně transparentní nebo srozumitelné. Tato nejasnost může vyvolat nedůvěru veřejnosti ohledně etičnosti a spravedlnosti těchto systémů. Kromě toho vyžadují AI systémy pravidelné aktualizace, aby adekvátně reflektovaly dynamické změny v sociálním a legislativním prostředí.

V souhrnu, ačkoliv AI přináší značné výhody pro efektivitu a rychlost rozhodovacích procesů v oblasti azylu, její použití přináší závažné výzvy. Tyto výzvy zahrnují otázky transparentnosti, etičnosti a schopnosti algoritmů adekvátně reagovat na komplexní lidské situace. Tyto problémy vyžadují pečlivé zvažování a řízení, aby se zajistilo, že technologické inovace slouží ke zlepšení, a ne k zhoršení spravedlnosti.

V závěru mé diplomové práce musím zdůraznit, že neustálé zdokonalování AI systémů a jejich doplnění o lidský úsudek je klíčové pro dosažení rovnováhy mezi technologickými inovacemi a zachováním základních lidských hodnot. Jako studentka vysoké školy IT jsem pochopila, že transparentnost, empatie a pravidelné aktualizace jsou nezbytné k tomu, aby AI systémy sloužily jako nástroje pro spravedlivé rozhodování, a ne jako překážky. Tento přístup nám umožňuje maximalizovat přínosy umělé inteligence, zatímco minimalizujeme rizika spojená s jejím využitím, a zajišťujeme, že její aplikace je etická a spravedlivá.

7 Literatura

7.1 Seznam citací

- AGATA SZWED, *The use of artificial intelligence in migration-related procedures in the European Union - opportunities and threats*, Procedia Computer Science, Volume 207, 2022, Pages 3645-3651, ISSN 1877-0509 [cit. 2023-11-13].
- AIRON, Palak. *Rychlý průvodce porozuměním algoritmu KNN*. Online. Unite. 2022. Dostupné z: <https://www.unite.ai/cs/rychlý-průvodce-algoritmem-knn/>. [cit. 2024-03-02].
- ANA BEDUSCHI, *International migration management in the age of artificial intelligence*, Migration Studies, Volume 9, Issue 3, September 2021, Pages 576–596
- ASHISH. 15 Most Important Features of Scikit-Learn!. *Analytics Vidhya* [online]. 2022 [cit. 2024-02-17]. Dostupné z: <https://www.analyticsvidhya.com/blog/2021/07/15-most-important-features-of-scikit-learn/>
- AZURE. *Algoritmy strojového učení*. Online. Microsoft Azure. 2024. Dostupné z: <https://azure.microsoft.com/cs-cz/resources/cloud-computing-dictionary/what-are-machine-learning-algorithms>. [cit. 2024-04-22].
- AZURE. *What is cloud computing?* Online. Azure. 2024. Dostupné z: <https://azure.microsoft.com/en-in/resources/cloud-computing-dictionary/what-is-cloud-computing>. [cit. 2024-03-02].
- BIRD, S., KLEIN, E., & LOPER, E. (2009). *Natural Language Processing with Python*. O'Reilly Media. [cit. 2024-03-02].
- BORKERT, M a SCHOLTEN, P. In: *The ethics of using AI in migration management: A literature review*. *Migration and Development*. 2020. 9(3) S. 437-454. [cit. 2024-03-02].
- DOBROVOLNIK. *KDO JE UPRCHLÍK*. Online. 2023. Dostupné z: <https://www.idobrovolnik.cz/cs/kdo-je-uprchlik>. [cit. 2023-12-21].
- DRBOHLAV, D.: Geografické aspekty v rámci interdisciplinárního výzkumu migrace obyvatelstva. In: *Geografie – Sborník české geografické společnosti, rok 1999, č. 2, roč. 104, s. 73 – 88*
- EASO. *Přístup k azylovému řízení*. Online. 2016. Dostupné z: https://euaa.europa.eu/sites/default/files/publications/Practical_Tools-Access_To_Procedures-Practical-Guide-CS.pdf.pdf. [cit. 2024-01-21].
- EUAA. *Latest Asylum Trends - Annual Overview 2022*. Online. EUAA. European Union Agency for Asylum. 2022. Dostupné z: <https://euaa.europa.eu/latest-asylum-trends-annual-overview-2022>. [cit. 2024-02-17].
- EUROPEAN COUNCIL. *Infographic - Asylum applications in the EU*. Online. European Council. 2023. Dostupné z: <https://www.consilium.europa.eu/en/infographics/asylum-applications-eu/>. [cit. 2024-03-02].

- EUROPEAN PARLIAMENT. *Asylum and migration in the EU: facts and figures*. Online. European Parliament. 2017, 05-08-2022. Dostupné z: <https://www.europarl.europa.eu/topics/en/article/20170629STO78630/asylum-and-migration-in-the-eu-facts-and-figures>. [cit. 2024-02-17].
- EUROPEAN PARLIAMENT. *Exploring migration causes: why people migrate*. Online. European Parliament. 2020, 02-05-2023. Dostupné z: <https://www.europarl.europa.eu/topics/en/article/20200624STO81906/exploring-migration-causes-why-people-migrate>. [cit. 2024-02-17].
- FILIPOVÁ, Jana. *Azylové právo v ČR*. Online, Diplomová práce. Brno: Masarykova univerzita, 2008. Dostupné z: https://is.muni.cz/th/dvvee/azylove_pravo.pdf. [cit. 2024-01-21].
- GAVAGHAN, C., KNOTT, A., MACLAURIN, J., ZERILLI, J., & LIDDICOAT, J. (2019). *Government Use of Artificial Intelligence in New Zealand: Final report on phase 1 of the New Zealand Law Foundation's Artificial Intelligence and Law in New Zealand Project*. New Zealand Law Foundation. <https://tinyurl.com/3w9ybbud>. Retrieved February 4, 2021. [cit. 2023-11-13].
- GHEZELBASH, D. In: *The promise and peril of using artificial intelligence in migration governance*. *Journal of Refugee Studies*. 2020. 33(3) S. 385-406.
- GOLDBERG, Y. (2017). *Neural Network Methods for Natural Language Processing*. Morgan & Claypool Publishers. [cit. 2023-11-13].
- HENIG, David. In: *Migrace* [online]. 2-3. AntropoWeb: AntropoWeb, 2007, s. 130 [cit. 2023-10-09]. ISBN 1801-8807. Dostupné z: http://www.antropoweb.cz/webzin/achive_old/webzin_2-3_2007/AntropoWEBZIN%202-3_2007.pdf, 35-41.
- HO, D. E. (2017). *Using Machine Learning to Predict Asylum Outcomes*. *The Journal of Politics*, 79(2), 691-705. [cit. 2023-11-13].
- HRDLIČKA, VLADAN. *Soudní přezkum žádosti o udělení mezinárodní ochrany (azyly)*. Online, Diplomová práce. Brno: Masarykova univerzita, 2021. Dostupné z: https://is.muni.cz/th/ngwtg/Diplomova_prace.pdf. [cit. 2024-01-21].
- CHATIFY. *Jak umělá inteligence ovlivňuje rozhodovací procesy ve velkých korporacích?* Online. Chatify. 2023. Dostupné z: <https://chatify.cz/blog/jak-umela-inteligence-ovlivnuje-rozhodovaci-procesy-ve-velkych-korporacich>. [cit. 2024-01-25].
- CHEN, C., ZHANG, C., CHEN, T., WANG, L., & YU, Y. (2014). *Big Data Challenges and Opportunities in Building a Learning Healthcare System*. *Proceedings of the IEEE*, 102(12), 1947-1962. [cit. 2024-03-02].
- IBM. *Naïve Bayes classifiers*. Online. IBM. 2024. Dostupné z: <https://www.ibm.com/topics/naive-bayes>. [cit. 2024-03-02].
- IBM. *What is a chatbot?* Online. IBM. 2024. Dostupné z: <https://www.ibm.com/topics/chatbots>. [cit. 2024-01-25].

- IBM. *What is a neural network?* IBM [online]. IBM, 2023 [cit. 2023-06-07]. Dostupné z: <https://www.ibm.com/topics/neural-networks>
- IBM. *What is support vector machine.* Online. IBM. 2023. Dostupné z: <https://www.ibm.com/topics/support-vector-machine>. [cit. 2024-03-02].
- INELI CIGER, Meltem, *Artificial intelligence and resettlement of refugees : implications for the fundamental rights*, EUI, RSC, Working Paper, 2023/44, Migration Policy Centre - <https://hdl.handle.net/1814/75689> [cit. 2024-01-25].
- JELÍNKOVÁ, Radka. *Právní postavení uprchlíků v České republice.* Online. Tábor, 2000. Dostupné z: https://aa.ecn.cz/img_upload/213998dd557a6ecf241d80d7748bd811/postaveni_uprchliku.pdf. [cit. 2023-11-13].
- JORDAN, M. I., & MITCHELL, T. M. (2015). *Machine learning: trends, perspectives, and prospects.* Science, 349(6245), 255-260.
- JURAFSKY, D., & MARTIN, J. H. (2020). *Speech and Language Processing.* Cambridge University Press. [cit. 2023-11-13].
- KLÍMEK, Martin. *Biometrická identifikace a verifikace.* Online, Bakalářská práce. Zlín: Univerzita Tomáše Bati, 2012. Dostupné z: http://digilib.k.utb.cz/bitstream/handle/10563/21926/kl%EDmek_2012_bp.pdf?sequence=1. [cit. 2024-01-25].
- KOŘOUSKOVÁ, Barbora. *CO JE STROJOVÉ UČENÍ A JAK SOUVISÍ S UMĚLOU INTELI-GENCÍ?* Online. Rascasone. 2021. Dostupné z: <https://www.rascasone.com/cs/blog/strojove-uceni-ml-metody-klasifikace>. [cit. 2024-01-25].
- KOŘOUSKOVÁ, Barbora. *UMĚLÁ INTELIGENCE (AI): HISTORIE A TRENDY PRO ROK 2023.* Online. Rascasone. 2023. Dostupné z: <https://www.rascasone.com/cs/blog/umela-intelligence-ai-trendy>. [cit. 2024-01-25].
- KUČERA, Jiří. *Algoritmus k-means.* Online. MUNI. 2024. Dostupné z: https://is.muni.cz/th/172767/fi_b/5739129/web/web/kmeans.html. [cit. 2024-03-02].
- LUKÁŠ, Ondřej. *What is the history of artificial intelligence (AI)?* Online. AI Dětem. 2024. Dostupné z: <https://aidetem.cz/obecny-uvod-do-umele-intelligence/strucna-historie-umele-intelligence/>. [cit. 2024-03-02].
- MARR, B. (2018). *What is deep learning AI? A simple guide with 8 practical examples.* <https://tinyurl.com/3hb8uzup>. Retrieved August 9, 2020. [cit. 2023-11-13].
- MAYER-SCHÖNBERGER, V., & CUKIER, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think.* Houghton Mifflin Harcourt. [cit. 2023-11-13].
- MCCARTHY, J., MINSKY, M. L., ROCHESTER, N., & SHANNON, C. E. (2006). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.* AI Magazine, 27(4), 12-14. [cit. 2023-11-13].

- MERIVA. *Poskytují AI překladače skutečně kvalitní překlady?* Online. https://www.meriva-preklady.cz/clanek/jsou-preklady-umele-intelligence-ai-kvalitni-jako-ziveho-prekladatele_178. 2023. Dostupné z: https://www.meriva-preklady.cz/clanek/jsou-preklady-umele-intelligence-ai-kvalitni-jako-ziveho-prekladatele_178. [cit. 2024-01-25].
- MICROSOFT. *Metriky vyhodnocení*. Online. MICROSOFT. Microsoft Learn. 2024. Dostupné z: <https://learn.microsoft.com/cs-cz/azure/ai-services/language-service/custom-text-classification/concepts/evaluation-metrics>. [cit. 2024-03-02].
- NALBANDIAN, L. *An eye for an 'I': a critical assessment of artificial intelligence tools in migration and asylum management*. CMS 10, 32 (2022). [cit. 2023-11-13].
- NVIDIA. *PyTorch*. Online. NVIDIA. 2024. Dostupné z: <https://www.nvidia.com/en-us/glossary/pytorch/>. [cit. 2024-03-02].
- NVIDIA. *TensorFlow*. Online. NVIDIA. 2024. Dostupné z: <https://www.nvidia.com/en-us/glossary/tensorflow/>. [cit. 2024-03-02].
- OAMP. *Mezinárodní ochrana*. Online. Odbor azylové a migrační politiky, Ministerstvo vnitra ČR, Praha. 2018. Dostupné z: <https://www.mvcr.cz/migrace/clanek/nase-hlavni-temata-mezinarodni-ochrana-mezinarodni-ochrana.aspx?q=Y2hudW09NQ%3D%3D>. [cit. 2024-01-21].
- OAMP. *Průběh řízení o udělení mezinárodní ochrany*. Online. Odbor azylové a migrační politiky, Ministerstvo vnitra ČR, Praha. 2023. Dostupné z: <https://www.mvcr.cz/clanek/prubeh-spravniho-rizeni-o-udeleni-mezinarodni-ochrany.aspx>. [cit. 2024-01-21].
- OAMP. *Průběh správního řízení o udělení mezinárodní ochrany*. Online. Odbor azylové a migrační politiky, Ministerstvo vnitra ČR, Praha. 2023. Dostupné z: <https://www.mvcr.cz/migrace/clanek/podrobnejsi-informace-o-prubehu-spravniho-rizeni-o-udeleni-mezinarodni-ochrany.aspx>. [cit. 2024-01-21].
- OAMP. *Slovníček pojmů*. Online. Odbor azylové a migrační politiky, Ministerstvo vnitra ČR, Praha. 2024. Dostupné z: <https://www.mvcr.cz/migrace/clanek/slovnicek-pojmu.aspx>. [cit. 2024-01-21].
- PLATFORM OPENAI. *Text generation models*. Online. Platform OpenAI. 2024. Dostupné z: <https://platform.openai.com/docs/guides/text-generation>. [cit. 2024-03-02].
- RUHS, MARTIN A DRAŽANOVÁ, Lenka. *Algorithmic Fairness for Asylum-Seekers and Refugees*. Online. Migration policy center. 2023. Dostupné z: <https://migrationpolicycentre.eu/projects/algorithmic-fairness-for-asylum-seekers-and-refugees-afar/#overview>. [cit. 2024-01-25].
- RUSSELL, S., & NORVIG, P. (2016). *Artificial Intelligence: A Modern Approach*. Pearson. [cit. 2023-11-13].
- SAMUELS, M., & CASELLI, M. (2021). *The challenges and potential of using AI in asylum adjudication*. Migration Studies, 9(3), 576-593. [cit. 2023-11-13].

- SAP. *Co je to umělá inteligence?* Online. SAP. 2023. Dostupné z: <https://www.sap.com/cz/products/artificial-intelligence/what-is-artificial-intelligence.html>. [cit. 2023-12-28].
- SAP. *Co jsou prediktivní analytické nástroje?* Online. SAP. 2024. Dostupné z: <https://www.sap.com/cz/products/technology-platform/cloud-analytics/what-is-predictive-analytics.html>. [cit. 2024-01-25].
- SCHÖPFLIN, A. In: *Can artificial intelligence enhance migration governance in the EU? A critical assessment. Journal of Ethnic and Migration Studies*. 2020. 46(10) S. 385-406.
- SIMON HAYKIN (2005). *Neural Networks: A Comprehensive Guide*. Přístupné z: https://cdn.preterhuman.net/texts/science_and_technology/artificial_intelligence/Neural%20Networks%20-%20A%20Comprehensive%20Foundation%20-%20Simon%20Haykin.pdf. [cit. 2023-11-13].
- SIMPLILEARN. *What Is Keras: The Best Introductory Guide To Keras*. Online. Simplilearn. 2024. Dostupné z: <https://www.simplilearn.com/tutorials/deep-learning-tutorial/what-is-keras>. [cit. 2024-03-02].
- TABLEAU. *What is the history of artificial intelligence (AI)?* Online. Tableau. 2024. Dostupné z: <https://www.tableau.com/data-insights/ai/history>. [cit. 2024-03-02].
- TALKPAL. *Jak umělá inteligence mění učení jazyků*. Online. Talkpal. 2024. Dostupné z: <https://talkpal.ai/cs/uceni-jazyku-nebylo-s-umelou-inteligenci-nikdy-tak-snadne/>. [cit. 2024-01-25].
- THIELEMANN, E. R. In: *Artificial intelligence and migration governance. Migration Studies*. 2019. 7(1) S. 1-15. [cit. 2024-03-02].
- UHRC. *Refugee Data Finder*. Online. 2024. Dostupné z: <https://www.unhcr.org/refugee-statistics/download/?url=OyZ4me>. [cit. 2024-04-23].
- UNHCR. *Žadatel o azyl*. Online. 2016. Dostupné z: <https://www.unhcr.org/cz/98-czkomu-pomahamezadatele-o-azyl-html.html>. [cit. 2024-01-21].
- UNHCR. *Ženevská úmluva*. Online. 1951. Dostupné z: https://www.unhcr.org/cz/wp-content/uploads/sites/20/2016/12/Umluva_1951_a_Protokol_1967.pdf. [cit. 2024-01-21].
- VODSLOŇ, František. *Vyhodnocování podobnosti zdrojových textů*. Online, Bakalářská práce. Praha: Univerzita Karlova v Praze Matematicko-fyzikální fakulta, 2007. Dostupné z: <https://www1.cuni.cz/~obo/vyuka/projekty/vodslon-ukazkova-bc.pdf>. [cit. 2024-04-23].
- ZÁKONY PRO LIDI. *Zákon č. 325/1999 Sb.* Online. 1999. Dostupné z: <https://www.zakonyprolidi.cz/cs/1999-325>. [cit. 2024-01-21].
- ZPRAVODAJSTVÍ EVROPSKÝ PARLAMENT. *Migrace: Proč lidé opouštějí své domovy?* [online]. 2020, 21.12.2023 [cit. 2024-01-21]. Dostupné z: <https://www.europarl.europa.eu/news/cs/headlines/world/20200624STO81906/migrace-proc-lide-opousteji-sve-domovy>

8 Přílohy

V této části je odkaz na dostupné úložiště: <https://github.com/kaciicerna/AIApplyForAsylum.git>, které obsahuje další materiály související s touto prací. Uložiště nese název AIApplyForAsylum a obsahuje seznam souborů, které jsou klíčové pro správnou funkci algoritmu, na kterém tato práce závisí. Mezi tyto soubory patří datové soubory, textové dokumenty a samotný soubor s implementovaným algoritmem. Na tomto úložišti naleznete také podrobný návod, který pomůže s procesem stahování a instalace potřebných knihoven, stejně jako s následným spuštěním algoritmu.