

Mendelova univerzita v Brně
Převádzkovo ekonomická fakulta

Digitální osobní asistent s emoční inteligencí pro aplikaci Moje MENDELU

Diplomová práce

Vedúci práce:
Ing. David Procházka, Ph.D.

Bc. Petra Javorková

Brno 2023

Podakovanie

Mám rada anonymný citát o tom, že strach je iba prezlečená nevedomosť. Do riešenia tejto diplomovej práce som vstupovala s minimálnou skúsenosťou z tvorby konverzačných agentov. Chcela by som preto vyjadriť vďaku všetkým výskumníkom, ktorí vkladajú svoju expertízu a vášeň do skúmaného problému, a vytvárajú tak vzácne výstupy, ktoré posúvajú vpred problematiku spracovania prirodzeného jazyka.

Čestné prehlásenie

Prehlasujem, že som prácu **Digitálny osobný asistent s emočnou inteligenciou pre aplikáciu Moje MENDELU** vypracovala samostatne a všetky použité zdroje a informácie uvádzam v zozname použitej literatúry. Súhlasím, aby moja práca bola zverejnená v súlade s § 47b zákona č. 111/1998 Zb., o vysokých školách v znení neskorších predpisov a v súlade s platnou *Směrnici o zveřejňování závěrečných prací*.

Som si vedomá, že sa na moju prácu vzťahuje zákon č. 121/2000 Zb., autorský zákon, a že Mendelova univerzita v Brne má právo na uzatvorenie licenčnej zmluvy a použitie tejto práce ako školského diela podľa § 60 odst. 1 autorského zákona.

Ďalej sa zaväzujem, že pred spísaním licenčnej zmluvy o použití diela inou osobou (subjektom) si vyžiadam písomné stanovisko univerzity, že predmetná licenčná zmluva nie je v rozpore s oprávnenými záujmami univerzity a zaväzujem sa uhradiť prípadný príspevok na úhradu nákladov spojených so vznikom diela, a to až do ich skutočnej výšky.

V Brně dne 30. decembra 2022

.....

podpis

Abstract

Javorková, P. My MENDELU Conversational Agent with Emotional Intelligence. Diploma thesis. Brno, 2023.

Day-to-day conversation with an emotional support on a family or friends level (ESC) has become a scientifically interesting area of affective programming for conversational agents. Currently, there is a gap of context-specific approaches for a closed domain chatbot and their experimentation with the empathy involved. The goal of this thesis is to propose and implement a chatbot with an ability to detect emotions and react to them appropriately. The existing COMET empathetic model was adopted for this work in order to meet the social and physical aspects of common human world experience (commonsense). Other state-of-the-art ESC models were used to encode strategies to alleviate stress or generate responses outside of the training data, with subsequent training in the relevant configuration. The chatbot has been implemented in the RASA tool and is capable of unimodally detecting emotions and reacting to them relevantly aiming to reduce emotional distress if user has experienced one.

Key words: COMET, chatbot, affective computing, RASA, empathy-driven dialogue

Abstrakt

Javorková, P. Digitálny osobný asistent s emočnou inteligenciou pre aplikáciu Moje MENDELU. Diplomová práca. Brno, 2023.

Každodenná konverzácia s laickou emocionálnou podporou na úrovni priateľov či rodiny (ESC) sa stala vedecky zaujímavou oblasťou afektívneho programovania konverzačných agentov. Raritou či súčasťou medzerou sú kontextovo špecifické prístupy pre uzatvorenú doménu a ich experimentovanie so začlenením empatie. Táto práca sa zaoberá návrhom a implementáciou chatbota so schopnosťou detekcie emócií a ich vhodnou reakciou. Existujúci empatický model COMET bol prevzatý pre túto prácu s cieľom splnenia sociálnych a fyzických aspektov bežnej ľudskej skúsenosti s reálnym svetom. K zakódovaniu stratégií k zmierneniu stresu či generovaniu odpovedí mimo tréningových dát boli použité ďalšie *state-of-the-art* modely ESC výskumu, s následným natrénovaním v relevantnej konfigurácii. Implementovaný chatbot v nástroji RASA je schopný emócie detegovať na singulárnej modalite a relevantne na nich reagovať s cieľom znížiť stres zo situácie v ktorej sa používateľ nachádza.

Kľúčové slová: COMET, chatbot, afektívne programovanie, RASA, dialóg riadený empatiou

Obsah

1	Úvod práce	12
2	Cieľ práce	14
3	Literárny prehľad	15
3.1	Emócie na vstupe	15
3.1.1	Empatia a sympatia	15
3.1.2	Rozpoznanie emócií na základe používateľských reportov . . .	16
3.1.3	Rozpoznanie emócií na základe aktivity a vlastnosti písania cez digitálne konverzačné prostredie	17
3.1.4	Rozpoznanie emócií senzorickými údajmi smart telefónu . . .	17
3.1.5	Rozpoznanie emócií na základe textu	18
3.2	Modely emócií	19
3.3	Dátové sady	20
3.3.1	Dialógový systém založený na úlohách	20
3.3.2	Dialógový systém chit-chat	21
3.3.3	Modalita datasetu	22
3.4	Commonsense v spracovaní prirodzeného jazyka	22
3.4.1	Kognitívne systémy	23
3.4.2	Commonsense reasoning	24
3.4.3	Knowledge base	26
3.5	Commonsense knowledge v predtrénovaných jazykových modeloch . .	29
3.5.1	BERT	30
3.5.2	Lexikálna nápoveda	31
3.5.3	Physical commonsense	31
3.5.4	Symbolic reasoning	32
3.6	Commonsense zdroje	33
3.7	Response design patterns	35
3.7.1	Rule-based	36
3.7.2	Retrieval-based	37
3.7.3	Generation-based	39
3.7.4	Trustworthy design	40
3.7.5	Non-progress design patterns	41
3.8	Dialog systems	42
3.8.1	Empathetic responding	43
3.8.2	Emotional support conversation	45
4	Metodika	48
4.1	Zber datasetov	48
4.1.1	Zabezpečený dataset na taks-oriented konverzáciu	48
4.1.2	Zabezpečený commonsense v porozumení prirodzeného jazyka	48
4.1.3	Zabezpečený dataset pre emocionálnu podporu v konverzácii .	48

4.2	Analýza a design	49
4.2.1	Požiadavky na schopnosť spracovania emócií	49
4.2.2	Identifikácia emočného stavu	49
4.2.3	Zvolený dialógový systém	49
4.3	Zostavenie konverzačného modulu	50
5	Architektúra konverzačného modelu	51
5.1	Popis implementačného nástroju	51
5.2	Konfigurácia modelu	51
5.3	Tréningové dáta	53
5.4	Problematika všeobecnej konverzácie	53
5.4.1	Closed-domain riešenie	54
5.4.2	Open-domain riešenie	56
5.5	Usporiadanie dát do príbehov	57
6	Testovanie konverzačného modelu	61
6.1	Evaluácia dialógového modelu	61
6.2	Evaluácia NLU modelu	63
6.2.1	Výsledky klasifikátoru zámerov	63
6.2.2	Výsledky retrieval zámerov	65
7	Workflow end-to-end konverzácií	67
8	Používateľské testovanie	74
9	Diskusia	76
9.1	Etická úvaha	76
9.2	Limity práce	77
9.3	Nadväzujúci výskum	77
10	Záver	79
11	Literatúra	81
	Prílohy	95
A	Textový súbor požiadaviek	96
B	Konfiguračný súbor config.yml	97
C	Evaluácia Response Selector časti NLU modelu	98
D	Evaluácia dialógového modelu	99
E	Evaluácia NLU modelu	100

Zoznam obrázkov

Obrázok 1: Popis vyťažených vlastností zo surových dát (Zhang et al., 2018)	18
Obrázok 2: OCC model (Christoph, Lyons a Saerbeck, 2008)	20
Obrázok 3: Russelov model (Russell, 1980)	21
Obrázok 4: Translokácia emocionálneho stavu po viacerých rovnakých stimuloch v AVS priestore (Jing et al., 2015)	22
Obrázok 5: Viacnásobné odpovede generované modelom postavenom na datasete z <i>Weibo</i> (Havasi et al., 2009)	23
Obrázok 6: Modality datasetov pre rozpoznávanie emócií využívané v literatúrach (Zhang et al., 2020)	24
Obrázok 7: Pre daný fyzický cieľ vyjadrený prirodzeným jazykom musí model zvoliť najrozumnejšie riešenie (Bisk et al., 2020)	25
Obrázok 8: Kognitívne systémy (Kahneman, 2003)	26
Obrázok 9: AGQA ako štandard na testovanie kompozičného časopriestorového uvažovania (Madeleine, Ranjay a Maneesh, 2021)	27
Obrázok 10: Triedenie prístupov k získaniu základu fenoménu logického uvažovania (Davis a Marcus, 2015)	28
Obrázok 11: Lexikálna jednotka a rámec pre anglické slovo "book" (Kundi a Chitchyan, 2017)	28
Obrázok 12: Diagram zobrazujúci vybrané uzly a prepojenia pre pojem "koláč" v ConceptNet (Davis a Marcus, 2015)	29
Obrázok 13: Dopytovanie sa znalostného grafu a jazykového modelu na faktické znalosti (Petroni et al., 2019)	30
Obrázok 14: Odvodzovanie o akciách aplikovateľných na objekt na základe jeho vlastností (Forbes, Holtzman a Choi, 2019)	33
Obrázok 15: Ukážka malej podmnožiny znalostí zdravého rozumu ATOMIC- 2020 – atlasu rozsiahlych textových popisov o každodenných inferenčných poznatkoch o entitách a udalostiach, ktoré kódujú sociálne a fyzické aspekty bežnej ľudskej skúsenosti (Hwang et al., 2021)	35
Obrázok 16: Konverzácie bez a so zapojením emócií (Zhou et al., 2018)	36
Obrázok 17: Výsledok výskumu možností dialógového systému (Khatua et al., 2017)	36
Obrázok 18: Textové správy na vstupe sú predmetom priamejšej interakcie s robotom a sú rozdelené na ich konceptný význam a detegované emócie (Justo et al., 2018)	37
Obrázok 19: 13 vyvinutých dialógov pre realizáciu chatbotovej konverzácie (Denecke, Vaaheesan a Arulnathan, 2021)	38
Obrázok 20: Systém architektúry pre retrieval-based model (Ji, Lu a Li, 2014)	39

Obrázok 21: Výber so sémantickou podobnosťou a parametrami emócií (Lubis et al., 2019)	39
Obrázok 22: Dvojitá architektúra modelu (Rashkin et al., 2019)	40
Obrázok 23: Dôležitosť relevantnosti témy v systéme emočného dialógu (Hou et al., 2020)	40
Obrázok 24: Framework dialógového systému rozšírený o mechanizmus lexikónu z dôvodu zapojenia požadovaných emocionálnych slov (Song et al., 2019) .	41
Obrázok 25: Adaptačné stratégie používateľa pri výskyte NP v úlohou orientovanom dialógu (Li, Chen a Chang, 2019)	42
Obrázok 26: Kompozícia state-of-the-art dialógového modelu (Liu et al., 2021)	43
Obrázok 27: Ukážka tréningovej sady pre daný emočný štítok (Rashkin et al., 2019)	44
Obrázok 28: Výmena emócií a zámerov medzi poslucháčom a rečníkom (Welivita a Pu, 2020)	45
Obrázok 29: Prehľad stratégií frameworku ESC (Liu et al., 2021)	46
Obrázok 30: Rozloženie stratégií počas rozhovoru (Liu et al., 2021)	46
Obrázok 31: Organizovanie reakcie na základe kombinácie blokov COMET a zmesi stratégií (Tu et al., 2022)	47
Obrázok 32: Diagram spúšťania ABC dialógu chatbota určeného na monitorovanie duševného zdravia používateľa (Denecke, Vaaheesan a Arulnathan, 2021)	50
Obrázok 33: Škálovateľná architektúra RASA (Rasa, 2022)	51
Obrázok 34: Graf príbehov v RASA (Zdroj: autorka práce)	60
Obrázok 35: 2D model s predikovanými a anotovanými pozíciami viacerých emócií (Xie a Park, 2021)	66
Obrázok 36: Interpretácia spracovaného emočného vstupu a prípravy parametrov na generovanie odpovede pre vstup "I have just checked UIS and I had not passed my exam" (Zdroj: autorka práce)	67
Obrázok 37: Interpretácia troch mechanizmov generovania odpovede (Zdroj: autorka práce)	68
Obrázok 38: Vizualizácia konverzácie s rozpoznaným emočným faq zámerom (Zdroj: autorka práce)	71
Obrázok 39: Scenár na spracovanie všeobecného vstupu u dvoch rôznych konverzáciách predikuje akciu na aplikovanie stratégie a vhodnej <i>empathetic policy</i> ako čiastočný nástroj na emočnú podporu základnej úrovne. Ručne vytvorené interpretácie <i>empathetic policies</i> sa povahou približuje k bežnej podpore rodiny alebo priateľov (Zdroj: autorka práce)	73
Obrázok 40: Ukážka dialógu z používateľského testovania (Zdroj: autorka práce)	75

Obrázok 41: Matica ukazuje, ako často boli konkrétne FAQ otázky či emócie správne predpovedané (Zdroj: autorka práce)	98
Obrázok 42: Matica ukazuje, ako často bola daná akcia v scenári správne predpovedaná a ako často bola namiesto toho predpovedaná nesprávna akcia (Zdroj: autorka práce)	99
Obrázok 43: Matica ukazuje, ktoré zámery sa mýlia s inými (Zdroj: autorka práce)	100
Obrázok 44: Histogram rozloženia istoty pri predikovaní zámeru (Zdroj: autorka práce)	101

Zoznam tabuliek

Tabuľka 1: Prehľad populárnych bázy znalostí	34
Tabuľka 2: Prehľad definovaných zámerov	53
Tabuľka 3: Prehľad akcií pre reakciu agenta	54
Tabuľka 4: Príklad tréningových dát pre otázku "What is UIS?"	55
Tabuľka 5: Príklad tréningových dát pre podanie výroku v emócií "surprised"	57
Tabuľka 6: Prehľad dostupných stratégií	69

1 Úvod práce

„Otázkou nie je, či inteligentné stroje môžu mať emócie, ale či môžu byť stroje inteligentné bez akýchkoľvek emócií“, pokladá intelektuálnu otázku Minsky (1988, s. 163). Vylučujúce významy pre výraz ”strojový” adresujú na jednej strane absenciu emócií, a na druhej strane určitú zaujatost' či oddanosť jednej stereotypnej činnosti. Na prichádzajúcom prelome desaťročia sa do popredia dostáva dôležitosť emócií výpočtových strojov pre nenarušené rozhodovanie s odkazom na neurologické štúdie (Picard, 1995). Zo záveru vyplýva podnet či návrh, že implementácia počítačov s inteligentnými rozhodnutiami môže byť podmienená výstavbou počítačov s emóciami. V nasledujúcej dekáde pokladá Reilly (1996) na protichodné osi mieru emocionálnej expresivity postáv a mieru ich celkovej kvality. Lineárna regresia znázornenej situácie ukazuje, že interakcia stelesnených postáv s ľuďmi má stavať základy na emóciách.

S pokrokom v technológiách zohráva interakcia človeka s počítačom a využívanie softvéru a hardvéru na určenie emocionálneho stavu človeka (affective computing) dôležitú úlohu pri určovaní komplexného ľudského správania (Jadhav a Sugandhi, 2018).

Často skloňovaná inteligencia nesie so sebou pre skúmanú problematiku schopnosť čítania emócií v texte rôzne konotácie (Park, Bae a Cheong, 2020). Vnímať, zistiť a pochopiť stavy emocionálneho charakteru či iné sociálne signály, je jadrom sociálnej a emocionálnej inteligencie. Agent zosobnený v digitálnej konverzácii pri pomoci správnych nástrojov, porozumení sociálneho kontextu a využitií vhodného prezentačného jazyka môže mať značnú hodnotu v oblastiach behaviorálnej vedy, antropológie, neurológie a psychológie (Pantic et al., 2005). Pri otázke zvyšovania produktivity sociálnej skupiny, vďaka včasnej detekcii problémov, je taký agent favoritom na odbúranie kolektívnych fenoménov – *groupthink* či *polarization* (Pentland, 2005).

Emócie, z náтуры ich vplyvu na interpretáciu výroku, majú dopad na generovanie prirodzeného jazyka (lexikálny, syntaktický, sémantický a spôsob reči) ale i porozumenie prirodzeného jazyka (upriamenie/neupriamenie pozornosti na rečnÍka, kompletná/čiastočná analýza zložitej výpovede) (Reilly, 1996).

Automatické rozpoznávanie emócií z textu je neoddeliteľnou súčasťou návrhu prirodzeného, používateľsky príjemnejšieho a inteligentnejšieho používateľského rozhrania (Namavar Jahromi a Homayounpour, 2012). S odkazom na súčasnosť je program so schopnosťou porozumenia používateľa a jeho vnútorného stavu vrcholom intuitívnosti komunikácie človek vz. počítač (HCI – Human Computer Interaction). Moschona (2020) ku svojej rozprave o systéme schopného emocionálnej inteligencie pridáva nový koncept detekcie emócií človeka vo fázach sebareflexie a vnútorného premýšľania, s ktorým sa pomyselne dostáva na úroveň stroja so schopnosťou čítať myšlienky. Pre daný koncept vymedzuje zber údajov o mozgových vlnách (konverzačná zložka) a EEG signáloch (zložka reči).

Správanie človeka podmienené emóciou prirodzene neprechádza z jedného neutrálného stavu do druhého, ale podlieha spontánne sa vyskytujúcim prejavom emócií. Realizácia interpretácie správania ovplyvneného emóciou si vyžaduje citlivosť na kontext v ktorom emocionálny model operuje [(Pantic et al., 2005), (Christoph, Lyons a Saerbeck, 2008)]. Ak pozornosť nie je upriamená na osobu, ale analýza smeruje do širšieho záberu, dokážeme hádať emocionálny stav bez viditeľného výrazu tváre. Pre potrebu takej analýzy (Calvo a D’Mello, 2010) bola vytvorená databáza obrázkov ľudí s veľkou škálou emocionálnych stavov vykonávajúcich rôzne činnosti na rôznych miestach.

2 Cieľ práce

Cieľom tejto práce je chatbot vyvinutý s emočnou inteligenciou ako základným determinantom pre všeobecnú konverzáciu v otvorenej doméne kde chatbot (*listener, supporter*) sprevádza používateľa (*speaker, help-seeker*) jeho situáciou v určitom emocionálnom stave. Pre podporu konceptu chatbota v univerzitnom prostredí je do implementácie zaradená *task-oriented* architektúra s doménou FAQ prevzatou od aplikácie *Moje MENDELU*. Chatbot bude vyvinutý frameworkom RASA a jeho implementácia bude abstrahovať produkčné prostredie pre prípadné potreby ďalšieho vývoja nadväzujúceho výskumu. Kým primárnym cieľom práce je predstaviť konverzačný modul zrovnateľný s *benchmarkom*, sekundárnym cieľom je poskytnutie základov pre budúci vývoj plne doménového chatbota pre aplikáciu *Moje MENDELU*. Autor práce na záver upozorňuje na limity a vznáša i etické otázky do diskusie.

3 Literárny prehľad

3.1 Emócie na vstupe

3.1.1 Empatia a sympatia

S pokrokom v spracovaní prirodzeného jazyka a strojového učenia sa otvárajú možnosti k vytvoreniu chatbota, ktorý emócie používateľa rozpoznáva a primeraným spôsobom na nich reaguje. Otázkou je, či porozumenie vyplývajúce z ľudskej komunikácie očakávajú používatelia aj od neživej entity agenta (Liu a Sundar, 2018). Pri tejto problematike sa ponúkajú dve teoretické a protichodné paradigmy:

- *Computers are social actors* (CASA): paradigma, ktorá naznačuje vnímanie partnera v konverzácii nezávisle od vnímanej identity – bez ohľadu či je partner považovaný za stroj alebo človeka (Ghandeharioun et al., 2019b).
- *Uncanny valley of mind* (UVM): paradigma, ktorá naznačuje, že robotický agent s vlastnou myslou môže pôsobiť pre používateľa rušivo. Zavedenie explicitného antropomorfizmu v *Human Computer Interaction* (HCI) je komplexný jav, ktorý môže vyvolať silné negatívne reakcie u ľudského respondenta (Larson a Csikszentmihalyi, 2014).

Výskumníci (Liu a Sundar, 2018) testovali vplyv troch typov empatického prejavu – sympatie, kognitívnej empatie a afektívnej empatie na vnímanie služby chatbota, ktorý sa vo svojej doméne venuje osobným problémom. Chatbot bol v štúdiu použitý v nasledujúcich skriptách:

- *Advice-only*: chatbot poskytuje výhradne len faktické informácie bez vyjadrenia súcitu – „Najlepšie bude ak pred použitím lieku navštívite lekára.“
- *Sympathy*: chatbot vyjadruje súcit a prenáša na seba pocity používateľa – „Vašej bolesti mi je ľúto.“
- *Cognitive-empathy*: chatbot vo svojej odpovedi rozpoznáva a uznáva pocity používateľa ale bez vyjadrenia ľútosti – „Svrbenie môže pôsobiť dosť rušivo.“
- *Affective-empathy*: chatbot v odpovedi pridáva pochopenie pre pocity používateľa – „Rozumiem tvojej obave z nakazenia.“

Pre potreby explicitného skúmania prirodzenosti či podobnosti stroja človeku boli respondenti vyzvaní k ohodnoteniu 7 stupňovou škálou i týchto premenných:

- Desivý či upokojujúci pocit z podobnosti ľudskej reakcie.
- Presvedčenie o pocitoch robota.
- Presvedčenie o inteligencii robota.

Experiment autorov dokazuje, že už existujúce presvedčenie používateľa o robotických schopnostiach, s ktorým vstupuje do konverzačného prostredia, následne

ovplyvňuje vnímanie o empatii či sympatii chatbota. V zmysle vnímania podpory zo strany chatbota nemá kognitívna empatia výraznú pridanú hodnotu v porovnaní so sympatiou a afektívnou empatiou. Autori predpokládajú, že absencia prejavu priateľstva alebo záujmu o ostatných zo strany chatbota neprináša pre používateľa uspokojenie, i keď dobre zodpovedá povahe chatbota ako objektívneho stroja. Dáta v štúdiu teda ukazujú podporu paradigmy CASA. V závere štúdie sa pozorný čitateľ dozvie, že vyjadrenie kognitívnej empatie sa všeobecne chatbotom neodporúča, a odporúčaný je naopak prejav empatie najmä pre tých respondentov, ktorí neveria emóciám a inteligencii robota (Liu a Sundar, 2018).

3.1.2 Rozpoznanie emócií na základe používateľských reportov

EMotion-Aware mHealth Agent (EMMA) je agent, ktorý poskytuje emocionálne vhodné mikroaktivity empatickým spôsobom s možnosťou zachytenia nálady pomocou údajov z polohy zo smart telefónu (Ghandeharioun et al., 2019a). Okrem geolokačných a identifikačných dát, či iných detailných aktivít zistených prostredníctvom aplikácie, obohacuje počiatočný zber dát i metóda *Positive and Negative Affect Scale* (PANAS). PA odráža mieru do akej sa človek cíti nadšený, aktívny, smutný či bdely. NA je dimenzia subjektívneho utrpenia a neprijateľného zapojenia, ktorá zahŕňa rôzne averzné stavy nálady, a nízka NA je stavom pokoja (Watson, Clark a Tellegen, 1988). *Depression, Anxiety and Stress Scale* (DASS) slúži pre zachytenie depresie, hnevu a stresu. Tieto metódy sú ale vo svojej podstate statickou časťou zbierania dát.

Pre pochopenie dynamiky duševného zdravia je nevyhnuté využiť inštrument na popisanie variácií formou tzv. *self-assessment*. Použitím metódy *Experience Sampling Method* (ESM) je možno získať empirické údaje o zmenách polohy, sociálnych interakcií či vzorcov psychologických stavov používateľa (Larson a Csikszentmihalyi, 2014).

Na vykonávanie ESM sa v poslednej dobe čoraz častejšie využívajú osobní asistenti, chatboty a virtuálni asistenti, a preto je nevyhnutné študovať vlastnosti takého agenta (Ghandeharioun et al., 2019a). EMS metóda bola u EMMA agenta implementovaná na mobilnom zariadení pomocou matice založenej na Russelovom dvojrozmernom modeli emócií. Ten slúžil pre overenie automatických predpovedí nálady na základe senzorov (údaje o polohe či iné mikroaktivity získané inteligentným mobilným zariadením). Aplikácia zaznamenávala vybraný emocionálny kvadrant, ktorý mal používateľ vyplniť 5-krát denne, aby EMMA reagovala emocionálne relevantnými frázami. Automatizácia chatbota k predvídaniu emočných štítkov používateľa (emotion labels) vychádza z ML modelu so vstupom senzorických dát mobilného zariadenia (prioritne údaje o polohe) a datasetom založenom na emočných štítkoch pre výpočet presnosti modelu.

Autori EMMA agenta dokázali, že existuje implementácia agenta, ktorý dokáže detegovať náladu používateľa zo senzorických dát mobilného telefónu. Rovnako poukazuje na možnosť redukovania opakovaného užívateľského vstupu metódou EMS,

nakoľko vnímanie agenta ktorý automaticky predpovedá emocionálny stav užívateľa sa zásadne nelíši od vnímania agenta, ktorý využíva pri interakcii pravidelné EMS. Existuje niekoľko obmedzení tejto metódy, ako napríklad náročnosť na čas a zdroje a pravdepodobnosť záťaže účastníka konverzácie. I z toho dôvodu musí taký návrh agenta zohľadňovať maximalizáciu spoľahlivosti zhromaždených údajov a minimalizáciu záťaže používateľa (Kolakowska, Szwoch a Szwoch, 2020).

3.1.3 Rozpoznanie emócií na základe aktivity a vlastnosti písania cez digitálne konverzačné prostredie

Rozšírenie chatbota, ktorý rozpoznáva emócie z vyplnených záznamov používateľmi (Ghandeharioun et al., 2019b), prichádza riešenie v zmysle zaznamenania aktivity písania používateľa, a nasadzuje model strojového učenia pre rozpoznanie emócie. Taký chatbot (Ghosh et al., 2019) zhromažďuje dáta o emóciách a koreluje ich s vlastnosťami textového vstupu (rýchlosť písania, množstvo klávesov, počet vymazaných textov, percento špeciálnych znakov zadanych v jednej relácii, dĺžka a trvanie relácie) a dokáže určiť 4 emočné stavy (šťastný, smutný, vystresovaný, uvoľnený). Na zostavenie modelu boli použité aj dáta o posledných emóciách, nakoľko stavy emócií pretrvávajú v priebehu času a navzájom sa ovplyvňujú.

Presnosť predikcie by sa dala zvýšiť pridaním ďalších kontextových vlastností písania (deň v týždni, kategória aplikácie) alebo predchádzaním šikmosti dát štatistického rozdelenia u klasifikácii emócií (u používateľov boli v priebehu experimentu zaznamenané dve dominantné emócie). Pre efektívne doručenie emócií v rámci mobilného četu je možná aj kombinácia emotikonov a použitého typu písma (Choi a Aizawa, 2019).

3.1.4 Rozpoznanie emócií senzorickými údajmi smart telefónu

Získavanie dát a následné vymedzenie tried na anotáciu dát môže byť v reálnom čase náročné. Dôležité senzory a vstupné zariadenia mobilného telefónu, ktoré k takému zberu dát slúžia, sú napríklad kamera, mikrofón, klávesnica, dotyková obrazovka, senzory hĺbky obrazu, akcelerometer a gyroskop, senzor geologickej polohy, barometer, kompas a iné (Kolakowska, Szwoch a Szwoch, 2020).

Aplikovaním takej myšlienky je výskum (Zhang et al., 2018), ktorý za účelom detekcie emócií implementuje aplikáciu na platformu Android, ktorá okrem intervalového vytvárania záznamu o emóciách užívateľom (EMS metóda), číta citlivé senzorické údaje a následne ich ukladá na vzdialený server pre analýzu. Pre zber dát využíva GPS polohu, kontroluje stav zapnutia/vypnutia obrazovky smart telefónu, skenuje WiFi signály v okolí a zaznamenáva ID skenovaných prístupových bodov (APs) ako aj silu prijímaného signálu (RSS), číta mikrofón, svetelný senzor, akcelerometer a elektronický kompas. S časovou značkou zaznamenáva využitie aplikácie, prijatie telefónneho hovoru, odoslanie a prijatie SMS. Zo surových dát sú následne vyťažené vlastnosti zahrňujúce nasledujúce typy momentu: prostredie, kontakt, využitie aplikácie a aktivity (Obr. 1).

Type	Data source category	Extracted features description
Environment	Microphone Audio (5)	Mean, Variance, NR, SR, NSR
	Light Sensor (5)	Mean, Variance, DR, BR, DBR
	GPS (3)	GPS Longitude, Altitude, Latitude
	WiFi (20)	Frequency of SSIDs of the top 20 APs for an user
Contact	Phone Call (*)	Call frequency and duration of each contact person
	SMS (*)	SMS frequency of each contact person
APP Usage	APP log (18)	Duration of 18 APP Categories
Activity	Accelerometer (7)	Mean and Variance of three axis (X, Y, Z), Step Count
	Compass (6)	Mean and Variance of three axis (X, Y, Z)
	Gyroscope (6)	Mean and Variance of three axis (X, Y, Z)
	Screen (4)	Screen on ratio, off ratio, Sleeping Duration, Usage Amount

Obr. 1: Popis vyťažných vlastností zo surových dát (Zhang et al., 2018)

Meranie tlaku vyvíjajúceho na dotykovú obrazovku pri písaní má značnú koreláciu so zaznamenanou emóciou stresu pri vyplňaní EMS dotazníku (Exposito, Hernandez a Picard, 2018). Hoci je pre túto premennú potrebné ďalšie štúdium, mohla by sa zaradiť medzi ďalšie smerodajné senzorické údaje.

3.1.5 Rozpoznanie emócií na základe textu

Rozpoznanie emócií je prvým krokom k tomu, aby bol počítač afektívne inteligentný. Systémy na vnímanie afektu môžu byť klasifikované podľa modality (Daily et al., 2017). Každá z modalít, teda nástrojov k prenosu informácií pri komunikácii, má svoje výhody a nevýhody, ktoré ovplyvňujú jej hodnotu (validita, spoľahlivosť, časové rozlíšenie, nákladovosť, rušivosť používateľa). Prístupy k výpočtu afektu sú kategorizované práve stupňom modality správy – hovorená, gestikulovaná či písaná forma (Calvo a D’Mello, 2010). Prechod od multimodálneho k unimodálnemu prístupu (Kratzwald et al., 2018) rovnako spadá pod analýzu *affective computing*, ktorá sa vzťahuje na schopnosti počítača rozpoznať emocionálne stavy používateľa vyjadriť vlastné emócie a reagovať na emócie používateľa (Picard, 1995).

Kým multimodálne formy detekcie emócií pomocou hlasu, reči tela, obrazu či výrazu tváre majú silné výskumné zastúpenie v publikáciách zastávaných poprednými inštitúciami (IEEE), detekcia emócií z textu vo svojich počiatkoch v štúdiách zaostáva. Tento nedostatok pozornosti štúdiám o rozpoznávaní emócií primárne z písaného textu je najmä kvôli nemožnosti odchytenia podnetov k emóciám, zložitosti gramaticky chybných textov, sarkazmu či absencii kontextu (Acheampong, Wenyu a Nunoo-Mensah, 2020). Pre taký náročný výskum je esenciálne obmedziť detekciu emócie na identifikáciu n-psychických stavov, a to v závislosti na zvolenom modeli (Nandwani a Verma, 2021).

Kao et al. (2009) zredukovali problém detekcie emócie z textu na hľadanie vzťahu medzi konkrétnym textovým vstupom a aktuálnymi emóciami autora textu. Takto zredukovaný problém môže byť matematicky sformulovaný ako:

$$r : A * T \rightarrow E \quad (1)$$

kde E predstavuje množinu všetkých emócií,
 A predstavuje množinu všetkých autorov textov,
 T predstavuje množinu všetkých reprezentácií textov vyjadrujúcich emócie,
 r vyjadruje emóciu e autora a z textu t .

Množina T podlieha modifikácii na dennej báze aj kvôli novým slangom či terminológiám (Nandwani a Verma, 2021). Nakoľko neexistuje štandardná klasifikácia všetkých ľudských emócií, je dôležité aby výskum na začiatku definoval klasifikačný model emócií (Kao et al., 2009).

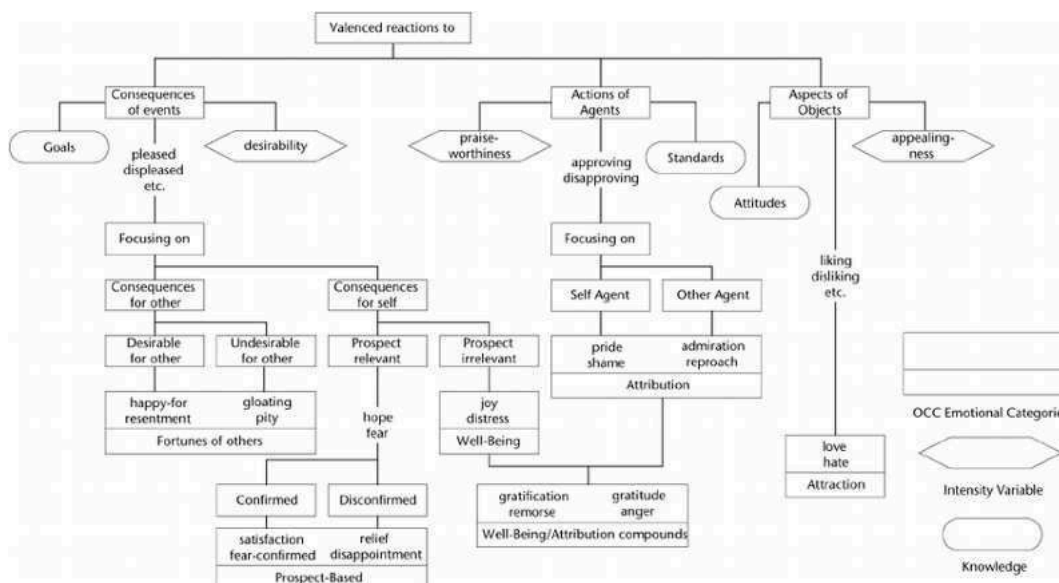
3.2 Modely emócií

Zložitosť javu *affective computing* so sebou prináša evolúciu odlišných teórií pre modelovanie výpočtu afektu. Taký model je základom pre každý systém zaoberajúci sa rozpoznávaním emócie (emotion detection). Väčšina algoritmov zaoberajúcich sa detekciou emócií používajú diskretný alebo dimenzionálny model pre prezentáciu stavu afektu – *discret/categorical emotion model* (DEM) a *dimensional emotion model* (DiEM) [(Acheampong, Wenyu a Nunoo-Mensah, 2020),(Hudlicka, 2017),(Xu et al., 2020)].

Emócie zvažované v teórii diskretného modelu sú zvyčajne základné a intuitívne, zahŕňajúce radosť, smútok, strach, hnev a znechutenie (Hudlicka, 2017). Teória zaradovania emócií do diskretných tried a kategórií má ale obmedzenú vyjadrovaciu schopnosť o intenzite emócie (Yang a Sun, 2017). Spomedzi skorých teoretikov vystupoval Ekman (1993) so šiestimi jednoduchými a intuitívnymi kategóriami: hnev, šťastie, smútok, prekvapenie, nechúť a strach. Tento zoznam slúži pre prípady, v ktorých má emócia charakteristický výraz a umožňuje pozorovateľovi odvodiť informácie o tom, o akú emóciu ide, bez poznania kontextu v ktorom sa výraz nachádza. K tak explicitnému určeniu dopomáha daná skladba 6 odlišných emócií bez vzájomnej závislosti. Ak charakter agenta používa ku spätnej väzbe iba výrazy tváre, odporúča sa zamerať na základných 6 emócií (Christoph, Lyons a Saerbeck, 2008). Ekmanove emócie sú preto v štúdiách využívané najmä pre skúmanie výrazu tvare (facial expression) [(Daily et al., 2017),(Feidakis, 2016)].

Colby et al. (1989) vyvinuli výpočtový model 22 emócií s označením OCC, často chápaný ako model pre syntézu emócií (Obr. 2). V tomto prípade platí fakt, že existuje iba jeden pozitívny výraz tváre – úsmev (Christoph, Lyons a Saerbeck, 2008), ku ktorému je možné priradiť všetkých 11 pozitívnych kategórií OCC modelu. Každý výraz sa objavuje v určitom kontexte, ktorý divákovi poskytuje ďalšie informácie. Model je vhodný na použitie v konverzačných agentoch (Ranade, Patel a Magare, 2018), nakoľko sa emócia vytvára ako výsledok situácie udalosti – objektu – agenta a dokáže odvodiť kognitívne emócie a spustiť tak subjektívne prežívanie iných emócií (Jing et al., 2015).

Podstata druhej metódy na charakteristiku stavu afektu spočíva v malom súbore základných dimenzií, ktoré definujú priestor pre umiestnenie konečného zoznamu emócií. Dimenzionálny model predstavuje konkrétny emočný stav ako bod v 2D, 3D

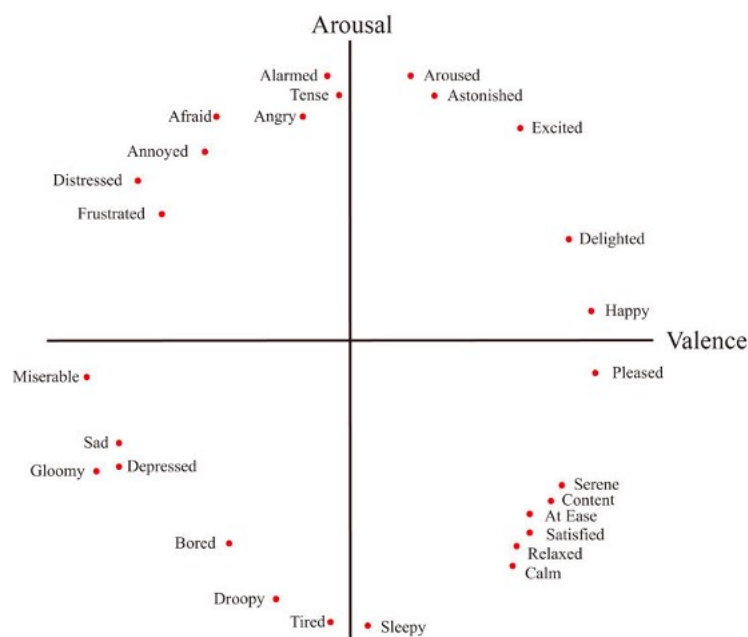


Obr. 2: OCC model (Christoph, Lyons a Saerbeck, 2008)

či 4D mierkach, ktoré sa líšia vo svojej interpretácii (Kolakowska, Szwoch a Szwoch, 2020). Táto metóda predpokladá, že emócie sú závislé a že existuje medzi nimi vzťah, ktorý je možno zobrazit v priestore (Acheampong, Wenyu a Nunoo-Mensah, 2020). Priestorový model dokáže lepšie zohľadniť multidimenzionálne charakteristiky ľudskej emócie. Štúdie uvádzajú nasledujúce dimenzionálne charakteristiky:

- *Arousal – Valence* (AV): dvojdimenzionálna charakteristika, ktorá popisuje ako je emócia príjemná/nepříjemná a aktívna/neaktívna. Túto charakteristiku popísal Russell pre svoj model *Russell's circumplex model of affect* (Obr. 3). Model patrí medzi najčastejšie využívané [(Seo a Huh, 2019), (Preotiuc-Pietro et al., 2016)].
- *Pleasure/Valence – Arousal – Dominance* (PAD): tretia dimenzia popisuje mieru, do akej má zažívajúci kontrolu nad svojimi emóciami. Medzi najznámejší model patrí Mehrabianov a Russellov (Mehrabian a Russell, 1974).
- *Arousal – Valence – Stance* (AVS): emočný priestor (Obr. 4.) vďaka ktorému je možné označiť určité vonkajšie emocionálne podnety pomocou troch druhov emocionálnych črt (vzrušenie, valencia a postoj), čím sa aktivuje zodpovedajúci citový stav (Jing et al., 2015).

Za účelom rozpoznávania emócií v kontexte (Kosti et al., 2017), bola vytvorená databáza *Emotions in Context Database* (EMOTIC), dátová sada obrázkov anotovaných kombináciou 2 vyššie vymenovaných prístupov: 26 emočných kategórií a 3 kontinuálne dimenziami *valence*, *arousal*, *dominance*. Dimenzionálne modely sú odporúčané pre také projekty, ktoré skúmajú podobnosť emócií (Acheampong, Wenyu a Nunoo-Mensah, 2020).



Obr. 3: Russelov model (Russell, 1980)

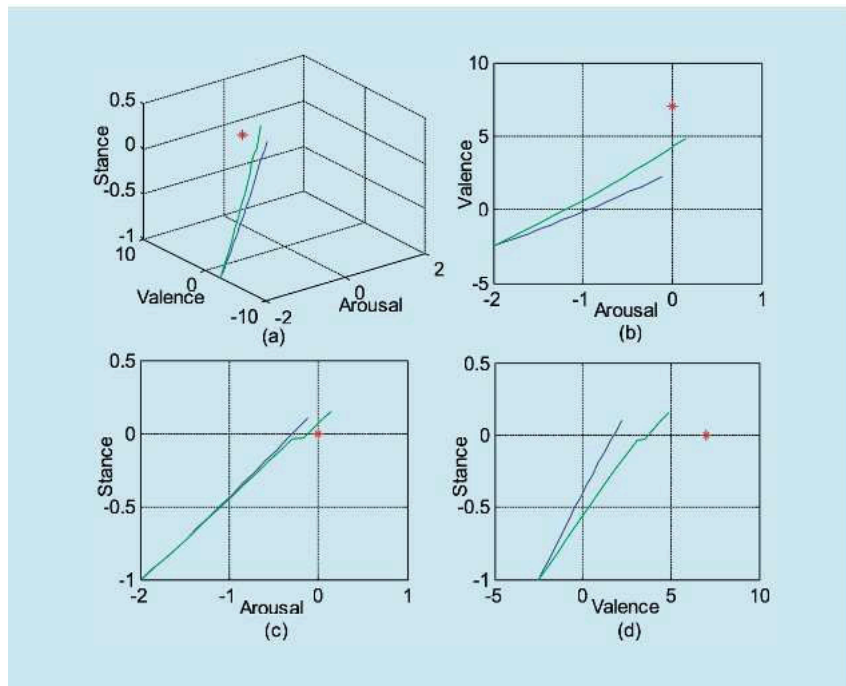
3.3 Dátové sady

Pri konverzačných agentoch založených na detekcii emócie je podstatné operovať s datasetom, ktorý obsahuje emočné konverzačné dáta (Yoo a Jeong, 2021). Avšak získať kvalitné emočne označené dáta vo veľkom rozsahu je subjektívna úloha a klasifikácia emócií je svojou podstatou náročná (Zhou et al., 2018). Autor, ktorý sa zaoberá problémom *text-based emotion detection* v konverzačnom prostredí najprv určuje, či ide o systém dialógu založenom na plnení úloh alebo na všeobecnom dialógu.

3.3.1 Dialógový systém založený na úlohách

Problém je zvyčajne formalizovaný ako úloha vyplnenia slotu, kde sa agenti pokúšajú predpovedať cieľ používateľa počas konverzácie. V tejto kategórii môže byť chatbot s neštruktúrovaným dialógom (bez logickej reprezentácie informácií počas rozhovoru) alebo so štruktúrovaným dialógom (Lowe et al., 2015).

Príkladom je *ATIS Database*, ktorá obsahuje výroky používateľov dopytujúcich sa otázky súvisiace s letom (Witbrock, 2010). Konverzačný spoločník *Xiaolce* integruje zručnosti IQ a EQ aby mohol používateľovi pomôcť s dokončením úlohy – a teda integrovať sa do prostredia používateľa a pristupovať k znalostiam reálneho sveta (napríklad prostredníctvom volania API). Pre tento prípad je používaná modulárna architektúra kde pre každú úlohu prináleží jeden modul (Zhou et al., 2020).



Obr. 4: Translokácia emocionálneho stavu po viacerých rovnakých stimuloch v AVS priestore (Jing et al., 2015)

Ubuntu Dialogue Corpus pozostáva z takmer jedného milióna dvojčlenných konverzácií extrahovaných z chatovacích protokolov *Ubuntu1*, ktoré sa používajú na získanie technickej podpory pre rôzne problémy súvisiace s Ubuntu (Chen, Agbodike a Wang, 2020).

3.3.2 Dialógový systém chit-chat

Všeobecne sú sociálni chatboty založení na krátkej konverzácii s mimikou ľudskej konverzácie bez nutnosti operácie s prostredím používateľa. Takéto scenáre vedú k veľmi jednoduchej architektúre a ich modely pre generovanie odpovedí sú často trénované na dátach zhromaždených zo sociálnych sietí. To umožňuje chatbotom komunikovať na akúkoľvek tému (Zhou et al., 2020).

Táto populárna téma výskumu v oblasti spracovania prirodzeného jazyka je nazývaná ako *Short-Text Conversation* (STC). Princíp je založený na 1 odpovedi k 1 príspevku používateľa – *Post-Comments* forma (Li et al., 2020). Obr. 5 ukazuje viacnásobné odpovede generované modelom postavenom na datasete z *Weibo*, populárna blogová služba v Číne, ktorá má limit dĺžky príspevku a odpovede na 140 znakov. Vďaka tomu je tento dataset populárny pre konverzáciu v krátkom kontexte (Havasi et al., 2009).

K datasetom získaných zo sociálnych sietí patrí aj *AffectiveTweets4*, ktorý poskytuje viac ako 7 000 emóciami označených výrokov (Chowanda et al., 2021). Hoci existujú verejné súbory nazbieraných emočne anotovaných dát zo sociálnych médií,

nezachytávajú ale výrazy emócií špecifické pre danú doménu. To potvrdzuje obmedzenia pri zovšeobecnení do rôznych domén (rozdielne výrazy na politickom fóre, fóre o depresii atď.) (Adikari et al., 2021).

Post	吃素第一天，坚持住，崔朵拉。 First day of being a vegetarian. Hold on, CuiDuoLa
R1	坚持住，坚持素食！ Hold on, Keep up your vegetarian diet.
R2	想要健康吗？想要长寿吗？那就关注我的微博吧。 Would you like to be healthy? Would you like to live long? If so, follow me on Weibo.
R3	这就是传说中的吃素吗？ This is the so-called vegetarian diet ?
R4	吃饱了才有力气减肥 Eat up so that you will have enough energy to go for a diet!
R5	身体是革命的本钱，不要着急！ Health is the essence for winning your future battles. No rush!

Obr. 5: Viacnásobné odpovede generované modelom postavenom na datasete z *Weibo* (Havasi et al., 2009)

3.3.3 Modalita datasetu

Trénovací a validačný dataset by mal disponovať takou modalitou, aká je pre skúmaný problém vhodná. Obr. 6 poskytuje prehľad datasetov vhodných pre určité stupne modality.

3.4 Commonsense v spracovaní prirodzeného jazyka

Keď ľudia medzi sebou komunikujú, spoliehajú sa na podobné základné znalosti, ktoré takmer nikdy výslovne neuvádzajú. Vyplýva to z pragmatickej zásady, že sa ľudia vyhýbajú uvádzaniu informácií, ktoré sú poslucháčovi zrejmé. A tak približujú Havasi et al. (2009) význam všeobecnej (často nevyslovenej) základnej znalosti, ktorú ľudský mozog vníma implicitne v nových situáciách. Aby systém umelej inteligencie porozumel svetu ľudí, potrebuje zdroj informácií o základných vzťahoch medzi vecami v explicitnej forme pre výpočtové použitie.

Znalosť „zdravého sedliackeho rozumu“ je súbor bežne zdieľaných vedomostí o tom, ako svet funguje. V súvislosti s daným kontextom je možno použiť logické uvažovanie na vyvodenie záverov o tom, čo sa mohlo stať aby sa entita dostala do určitého stavu a aké rozumné ďalšie kroky by boli založené na tom, čo sa doteraz stalo (Ammanabrolu et al., 2021).

Data modality	Ref.
CP + C + PP	Koelstra et al. (2012) [11]
A + PP	Kim (2007) [56]
V + PP	Bailenson et al. (2008) [57]
CP + PP	Khalali and Moradi (2009) [58]
A + PP	Kim and Lingenfelter (2010) [59]
CP + PP	Chanel et al. (2011) [60]
A + PP	Walter et al. (2011) [61]
V + PP	Hussain et al. (2012) [62]
V + PP	Monkaresi et al. (2012) [63]
CP + Gaze	Soleymani et al. (2012) [64]
CP + C	Wang et al. (2014) [65]

Legenda: V = Video, A = Audio, C = Content/Context, CP = Central Physiology, PP = Peripheral Physiology; Gaze = eye Gaze.

Obr. 6: Modality datasetov pre rozpoznávanie emócií využívané v literatúrach (Zhang et al., 2020)

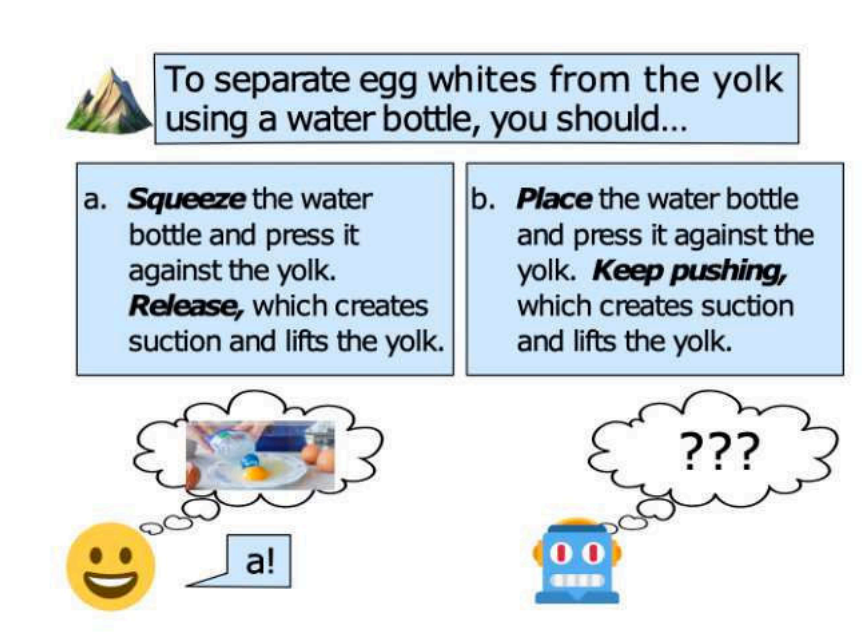
Hlavnou výzvou AI na ceste ku skutočnej úplnosti, ktorá interaguje so svetom a rozumie prirodzenému jazyku, je modelovanie znalostí o fyzikálnych vlastnostiach každodenných objektoch (Obr. 7). Táto subsekcia prináša prehľad o formách zberu základu pre „commonsense knowledge“, jej existujúcich zdrojoch, kognitívnych systémoch človeka, a v neposlednom rade i oblasti zautomatizovaného uvažovania zdravým rozumom.

3.4.1 Kognitívne systémy

S rýchlym pokrokom hlbokých neurónových sietí (DNN) a učeníu s učiteľom sa za posledných niekoľko rokov dosiahol pokrok v detekcii a lokalizácii objektov. Od skutočnej strojovej inteligencie nás ale delí absencia všeobecného pochopenia bežných faktov. Mostom medzi moderným počítačovým videním a ľudským videním je už spomínaná chýbajúca dimenzia – zdravý rozum podobný ľudskému. K hodnoteniu inteligencie systému môžeme využiť schopnosť dobre vykonávať rôzne úlohy, vnímanie prostredia, plánovanie akcií, predpovede správania iných agentov a schopnosť rýchlo prispôbiť naučené poznatky novému prostrediu pre nové úlohy (Zellers et al., 2019a).

Obr. 8 približuje dva typy kognitívnych procesov a ich charakteristiky, ktoré nesú označenie Systém 1 (intuícia) a Systém 2 (uvažovanie). Uvažovanie sa aktivuje napríklad pri výpočte matematického súčinu, a teda zámerne a s námahou, zatiaľ čo intuitívne myšlienky prichádzajú na myseľ spontánne a bez vedomého hľadania alebo výpočtov - bez námahy. Operácie systémov sú nasledovné:

- Systém 1: operuje rýchlo, automaticky, nenáročne, asociatívne a často emocionálne, riadiaci sa zvykom.
- Systém 2: operuje pomalšie, sériovo, namáhavo, zámerne kontrolovane, a potenciálne sa riadi pravidlami.



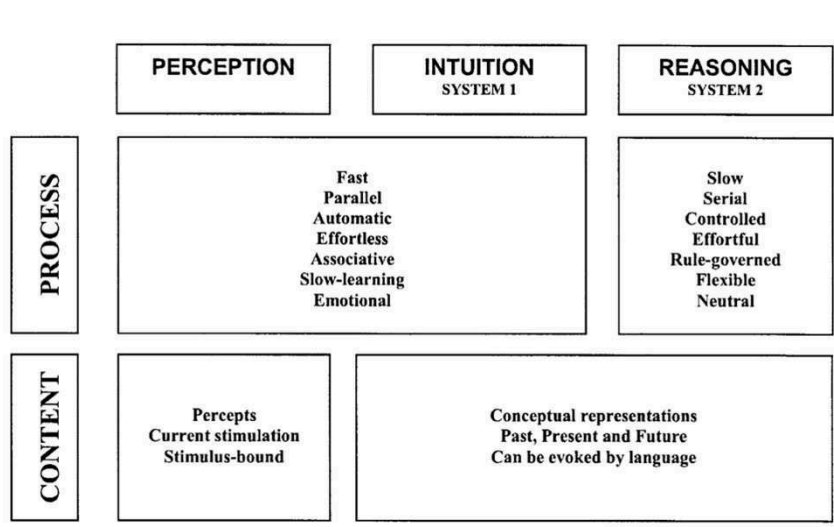
Obr. 7: Pre daný fyzický cieľ vyjadrený prirodzeným jazykom musí model zvoliť najrozumnejšie riešenie (Bisk et al., 2020)

Kahneman (2003) pripomína, že rozhodovanie o tom, či k danému mentálnemu procesu priradíme označenie jedného či druhého systému, rozhoduje rozdiel v mentálnom úsilí, ktoré je obmedzené. Namáhavé procesy majú tendenciu sa navzájom rušiť, zatiaľ čo procesy bez námahy netrpia veľkým rušením v kombinácii s inými úlohami.

3.4.2 Commonsense reasoning

„Kto je vyšší, princ William alebo jeho malý syn princ George?“. „Dokážete pripraviť šalát z polyesterovej košeľe?“. Mnohé inteligentné úlohy (porozumenie textu, počítačové videnie, plánovanie a vedecké uvažovanie) vyžadujú podobné druhy vedomostí a rozumových schopností v reálnom svete. Automatizované logické uvažovanie je pre AI úlohy dôležité a podstatný pokrok dosiahlo v štyroch nasledujúcich oblastiach:

- Taxonomické uvažovanie: porozumenie sémantickým sieťam. Taxonómia je množina kategórií, individuálnych inštancií a vzťahov medzi nimi. Taký systém AI rozumie dedeniu vlastností, označeniu vlastností či vytvorením podmnožín pre určenú doménu, kde napríklad **Pes** je podmnožinou **Cicavcov**.
- Časové uvažovanie: mnohé dôležité časové vzťahy nie sú v textoch explicitne uvedené a proces dedukcie tak môže byť náročný. Priradenie časových pečiatok udalostiam sa nedá v súčasnosti vykonať s vysokou mierou presnosti (napríklad priradená časová pečiatka udalosti v novinách). V roku 2021 do tejto oblasti pribudol nový štandard AGQA (Madeleine, Ranjay a Maneesh, 2021), ktorý



Obr. 8: Kognitívne systémy (Kahneman, 2003)

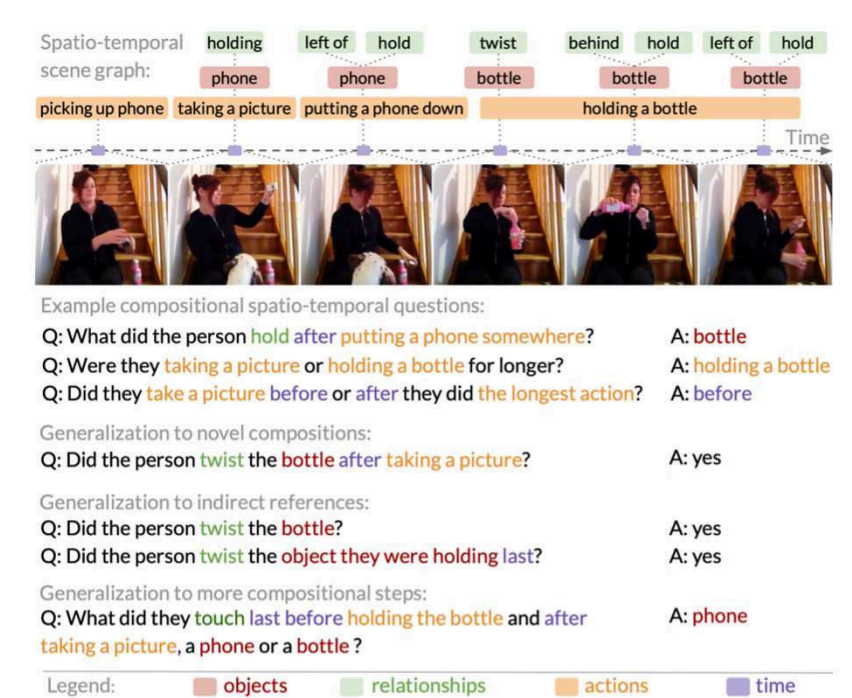
používa jazyk na hodnotenie vizuálnych kompozičných požiadaviek (Obr. 9).

- Akcia a zmena: problém reprezentácie a dôležitých foriem uvažovania akými sú predikcia a plánovanie (simultánne udalosti, teória rozhodovania, domény nedokonalých znalostí a domény viacerých agentov).
- Kvalitatívne uvažovanie: teória, ktorá bola aplikovaná v mnohých oblastiach fyziky, inžinierstva, biológie, ekológie či porozumenia textu. Týka sa smeru zmeny vzájomne súvisiacich veličín, napríklad ak ekosystém obsahuje líšky a králikov, a počet líšok sa zníži, potom sa úhyn králikov tiež zníži.
- Vizuálne uvažovanie: ľudský rozum dokáže jediným pohľadom na obrázok odvodit činy, ciele a duševné stavy ľudí, napríklad prečo človek 1 ukazuje prstom na človeka 2? Na rozdiel od systémov AI je táto úloha pre ľudí jednoduchá. Vďaka *R2C Engine* od (Zellers et al., 2019a) sa dokázali posunúť k porozumeniu na úrovni poznania, kedy sa na náročnú otázku o obrázku engine snaží odpovedať správne a následne poskytnúť odôvodnenie svojej odpovede. Presnosť výstupu dokázali oproti stávajúcim riešeniam zvýšiť o približne 20 %.

(Davis a Marcus, 2015)

Z Obr. 10 je zrejmé, že štúdium problematiky *commonsense reasoning* je do značnej miery rozdelené na prístupy založené na:

- Znalostiach (knowledge-based).
- Strojovom učení s veľkými dátovými korpunami (machine learning).
- Vybudovaní vedomostnej základne spojením kolektívneho poznania a participácie mnohých neodborníkov (crowd-sourcing).



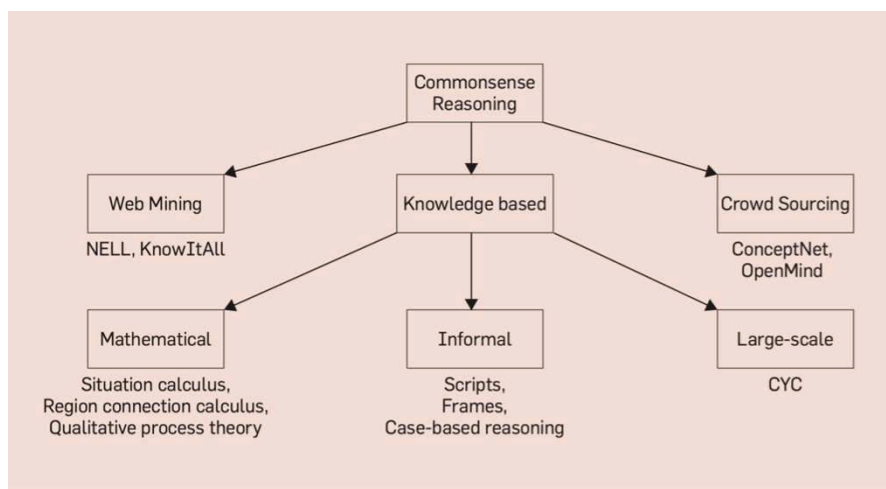
Obr. 9: AGQA ako štandard na testovanie kompozičného časopriestorového uvažovania (Madeleine, Ranjay a Maneesh, 2021)

3.4.3 Knowledge base

Pri absencii ľudského uvažovania v systéme umelej inteligencie musí mať taký systém nevysslovené znalosti o svete ako zdroj informácií v explicitnej forme, aby ich mohol použiť (Havasi et al., 2009). Nedávno sa objavil záujem o inžinierstvo požiadaviek na základe ontológie. Ontológia sa vo všeobecnosti používa na prezentáciu konceptov v rámci domény, takže by mohla pomôcť pri špecifikácii úplnejších a menej jednoznačných požiadaviek. Lingvistické ontológie zahŕňajú glosáre, slovníky, riadené slovníky, taxonómie alebo lexikálne databázy. Lingvistický ontologický prístup sa navrhuje ako dobrý nástroj na získanie požiadaviek, pretože je určený na zachytenie popisov konkrétnych domén z pohľadu spracovania prirodzeného jazyka (Kundi a Chitchyan, 2017).

Cyc projekt

Jedným z prvých pokusov ručne zostaviť veľmi rozsiahle vedomostné bázy zdravého rozumu je dodnes udržiavaný projekt Cyc. Najnovšia verejná verzia OpenCyc 4.0, vydaná v júni 2012, obsahuje 239 000 konceptov a 2 039 000 faktov (Davis a Marcus, 2015). Projekt Cyc sa avšak zameriava na vývoj zovšeobecnených polarizovaných lexikónov a teda neposkytuje veľa informácií v rôznych iných oblastiach podľa špecifickej domény. Použitie takého lexikónu na klasifikáciu textu môže ovplyvniť presnosť predikcie (Yildirim et al., 2019). Táto vedomostná база primárne slúži na



Obr. 10: Triedenie prístupov k získaniu základu fenoménu logického uvažovania (Davis a Marcus, 2015)

podporu budúcich reprezentácií znalostí a úloh k riešeniu *commonsense reasoning* (Matuszek et al., 2006).

FrameNet

Projekt Berkeley FrameNet (Kundi a Chitchyan, 2017) je lexikálna databáza, ktorá zohráva úlohu vedomostnej bázy pre spracovanie prirodzeného jazyka. Cieľom projektu je porozumieť syntaktickým a sémantickým hodnotám slova pretože inak nemožno úplne pochopiť význam jedného slova. V prirodzenom jazyku môže mať 1 slovo mnoho lingvistických asociácií, ako sú homonymá, synonymá a polysémia, a preto je význam slova vo FrameNet opísaný z hľadiska jeho vzťahu k sémantike daného rámca. Kľúčové komponenty sú *frame*, *lexical unit* a *frame element* (Obr. 11).

LU as Polysemy and homonyms)	Frame
book.n	Text
book.n	Part ordered segments
book.n	Aggregate
book.n	Records
book.v	Arrest
book.v	Reserving
booking.n	Reserving
booklet.n	Text

LU as Synonyms
book
booking
Hold
Reserve

Obr. 11: Lexikálna jednotka a rámec pre anglické slovo "book" (Kundi a Chitchyan, 2017)

PropBank

Hlavnou súčasťou pochopenia prirodzeného jazyka je pochopenie udalostí a ich aktérov. Generovanie takej sémantickej reprezentácie vety (*semantic role labelling*) rozpoznáva sémantické argumenty (kľúčové odpovede na otázky kto, čo, prečo, kde atď.) spojené so slovesom vo vete a klasifikuje ich do špecifických rolí. Narouei, Takabi a Nielsen (2020) uvádzajú modelový príklad schémy:

$$[Arg0 \textit{John}] [ArgM - MOD \textit{can}] [V \textit{assign}] [Arg1 \textit{clerk}] \quad (2)$$

$$[Arg2 \textit{to users from department A}] \quad (3)$$

kde V predstavuje verb,
 $ArgM - MOD$ predstavuje modal,
 $Arg0$ predstavuje assigner,
 $Arg1$ predstavuje thing assigned,
 $Arg2$ predstavuje assigned to.

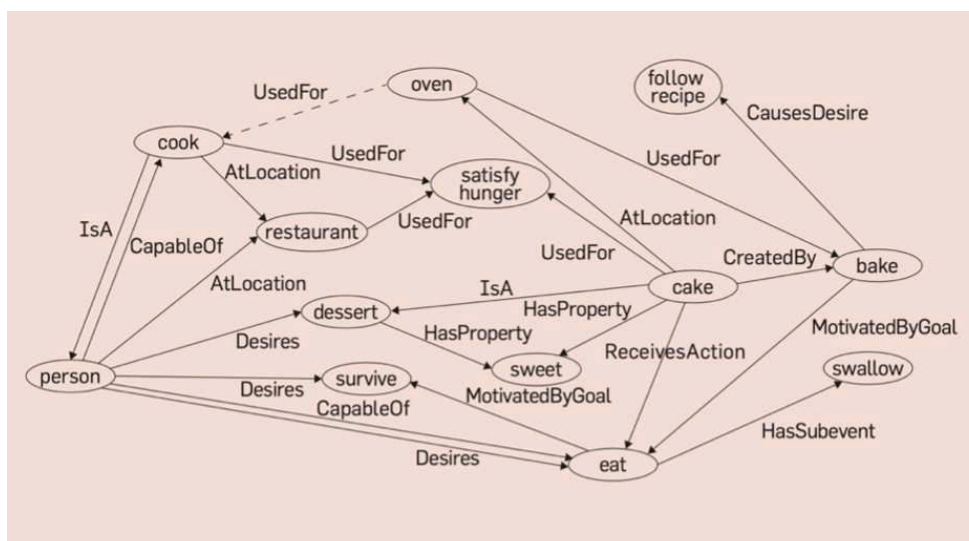
Pri tvorbe jedného z najväčších človekom označených video datasetov *Moments in Time* (Monfort et al., 2020), boli použité generovacie funkcie PropBank a FrameNet na zoskupenie sloviess a získanie informácií o význame a použití každého slovesa.

ConceptNet

ConceptNet je voľne dostupná viacjazyčná vedomostná základňa, ktorá je navrhnutá tak, aby pomohla strojom porozumieť znalostiam zdravého rozumu, ktoré ľudia používajú. Také znalosti popisujú vzťahy vychádzajúce zo znalostí zdravého rozumu medzi slovami a frázami v prirodzenom jazyku. Ukážky súvislosti porozumenia reálneho sveta sú vizualizované na Obr. 12. Jej dáta sú zozbierané dobrovoľníkmi (crowdsourcing) pomocou Wikislovníku či účelne vytvorených herných užívateľských rozhraní. API od ConceptNet neposkytuje slovník a pre tieto účely sa v aplikáciách používa WordNet (Sigdel et al., 2020).

WordNet

WordNet je najpoužívanejšia lexikálna databáza pre prirodzený jazyk. Často sa používa v lexikálnych požiadavkách, napríklad na pomoc pri zisťovaní synonym alebo definovaní sémantickej vzdialenosti medzi párom slov (Kundi a Chitchyan, 2017). Inými slovami, WordNet je databáza anglických slov obsahujúca podstatné mená, prídavné mená, definície a príklady použitia slov, a taktiež poskytuje niekoľko lexikálnych vzťahov (Sigdel et al., 2020). Jej obsah je vytvorený lexikografmi – ľudskými expertami. Konverzia textu na emotikony a vice versa (Jha, 2018), či integrácia sémantickej podobnosti slov odvodennej od taxonomických vzťahov (Aouicha, Taieb a Hamadou, 2016), i to sú príklady výskumov efektívne využívajúcich túto databázu. K ďalším zaujímavým „korpusom“ (zdrojom) znalostí patria Google-RE, T-REx, SQuAD a iné (Petroni et al., 2019).



Obr. 12: Diagram zobrazujúci vybrané uzly a prepojenia pre pojem "koláč" v ConceptNet (Davis a Marcus, 2015)

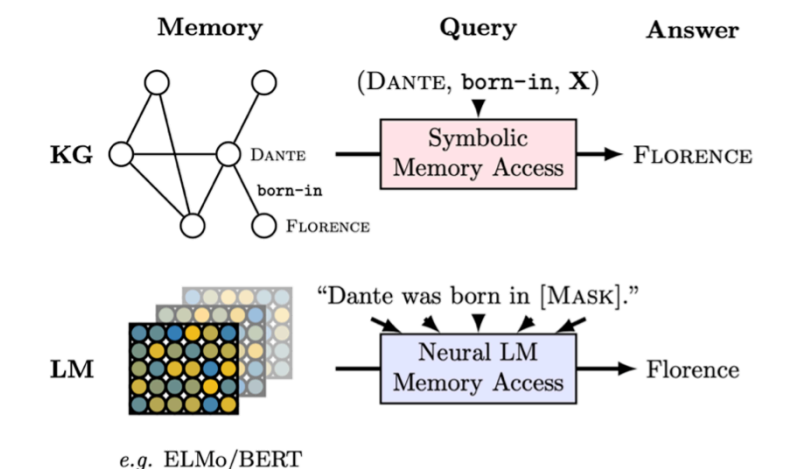
3.5 Commonsense knowledge v predtrénovaných jazykových modeloch

Trénovanie jazykových modelov (LMs) na veľkých textových datasetoch vedie k prudkému nárastu zlepšenia úloh z oblasti spracovania prirodzeného jazyka (NLP). Počas tréovania môžu tieto modely ukladať aj relačné znalosti prítomné práve v tréovacích dátach, čo im umožňuje odpovedať na štruktúrované otázky. Oproti samostatným štruktúrovaným znalostným bázam (KBs) majú LMs výhodu v tom, že sú ľahko rozšíriteľné o viac údajov, podporujú otvorenú množinu *queries* a pri tréovaní nepotrebujú anotované dáta dobrovoľníkmi. (Petroni et al., 2019)

KBs sú vhodným nástrojom pre prístup k anotačným relačným údajom vďaka tomu, že umožňujú dopyty formou symbolického odkazovania do pamäti (Obr. 13). Na druhej strane vysokokapacitné LMs ako ELMo a BERT sú optimalizované tak, aby predpovedali ďalšie slovo v sekvencii pomocou masky (masked language model).

3.5.1 BERT

Tento model jazykovej reprezentácie s *Bidirectional Encoder Representations from Transformers* (Devlin et al., 2019) je navrhnutý tak, aby predtrénoval hlboké obojsmerné reprezentácie bez učiteľa spoločným podmienením ľavého aj pravého kontextu vo všetkých vrstvách. Výhodou pri aplikovaní je, že model sa doladí len jednou ďalšou výstupnou vrstvou aby dosiahol výkon *state-of-the-art* modelu vhodného na širokú škálu NLP úloh bez podstatnej modifikácie architektúry pre danú úlohu. Obecne existujú 2 stratégie prístupov k aplikácii BERTu:



Obr. 13: Dopytovanie sa znalostného grafu a jazykového modelu na faktické znalosti (Petroni et al., 2019)

- *Feature-based*: napríklad ELMo, ktorý používa svoju architektúru pre také úlohy, ktoré obsahujú prednatrénované jazykové reprezentácie ako doplnkové funkcie.
- *Fine-tuning*: napríklad OpenAI GPT, ktorý zavádza minimálne parametre pre každú úlohu a svoje tréningovanie spúšťa jednoduchým doladením všetkých pôvodných parametrov.

BERT je prvý model reprezentácie založený na *fine-tuning* procese, ktorý dosahuje špičkový výkon na veľkom súbore dát, čím prekonáva mnohé doterajšie architektúry (jeho GLUE skóre je rekordných 80.5 %). Jeho implementácia spočíva z 2 častí:

- *Pre-training*: proces tréningovania na neoznačených dátach sa skladá z dvoch úloh bez učiteľa. Podstatou prvej úlohy je náhodne maskovanie 15 % vstupných tokenov a ich následné predpovedanie. Tento postup sa označuje ako *masked language model*. Podstatou druhej úlohy je predpovedanie ďalšej vety, tzv. *next sentence prediction*, ktoré je založené na pochopení vzťahu medzi dvoma vetami.
- *Fine-tuning*: najprv je model inicializovaný s predtrénovanými parametrami. Tento proces je zvyčajne veľmi rýchly a preto autori odporúčajú hľadať vhodnú kombináciu parametrov na danom datasete.

3.5.2 Lexikálna nápoveda

Pri obhajobe modelov, ktoré riešia *commonsense natural language inference* problém, vyzdvihujú výskumníci BERT kvôli jeho výkonu pri odvodzovaní zdravého rozumu takmer na úrovni človeka (86 %) (Zellers et al., 2019b). K takému výkonu sa populárne pripisoval titulok: „Finally, a machine that can finish your sentence.” Aplikáciou BERT na datasete SWAG (Zellers et al., 2019b) dospeli autori k nasledujúcim zisteniam:

- Model sám o sebe nepreukazuje robustnú schopnosť logického uvažovania.
- Model funguje ako rýchly *surface learner* na konkrétnom datasete.
- Výkon modelu závisí od procesu doladovania (fine-tuning).
- Model dosahuje dobrú *validation accuracy* (cca 50 %) už pri 16 000 vzorkách.
- Model by dosiahol výkon na úrovni človeka na danom datasete za 109 GPU hodín, resp. 100 000 GPU rokov, ak nedôjde k zlepšeniu algoritmov.
- Model sa pri voľbe možnosti rozhoduje podľa lexikálnej nápovedi (volí možnosť kde sa výskyt posledných slov zhoduje s kontextom otázky).
- Toho času neexistuje model, ktorý by vyriešil *commonsense reasoning*.

3.5.3 Physical commonsense

„Dokážu systémy AI spoľahlivo odpovedať na otázky týkajúce sa sveta a jeho fyzikálnych vlastností bez toho, aby tento fyzický svet skutočne zažívali?“. Fyzické chápanie je široká doména, ktorá sa dotýka všetkého od vedeckých poznatkov až po interaktívne získavanie vedomostí multimodálnym prístupom (Bisk et al., 2020). *Symbol grounding problem* predstavuje mapovanie predikátov (podstatné a prídavné mená, napríklad modrý pohár) k objektom z reálneho sveta. Zahŕňa prepojenie informácií zo zmyslového vnímania (vizuálne, haptické, sluchové) s údajmi zo skutočného sveta (Thomason et al., 2016).

Pri predmetoch reálneho sveta môžeme uvažovať o ich vlastnostiach (properties: „uses electricity“), akciách ktoré sú na nich aplikovateľné vyjadrené slovesami (affordances: „plug-in“), a ich vzájomnej implikácii (inferences: „plug-in means it uses electricity“). Absolútne implikácie (výrazy 4 až 6), hoci ich treba brať s rezervou, zachytávajú našu intuíciu o tom, čo očakávame za reálne a pravdivé. Avšak, ich modelácia na základe samotného textu je ťažká, pretože akákoľvek triviálne pochopiteľná implikácia je druh informácie, ktoré zostávajú explicitne nevyvodené (Forbes, Holtzman a Choi, 2019).

$$\text{robot.plug} - \text{in}(x) \rightarrow \text{uses} - \text{electricity}(x) \quad (4)$$

$$\text{keyhole.look} - \text{through}(x) \rightarrow \text{transparent}(x) \quad (5)$$

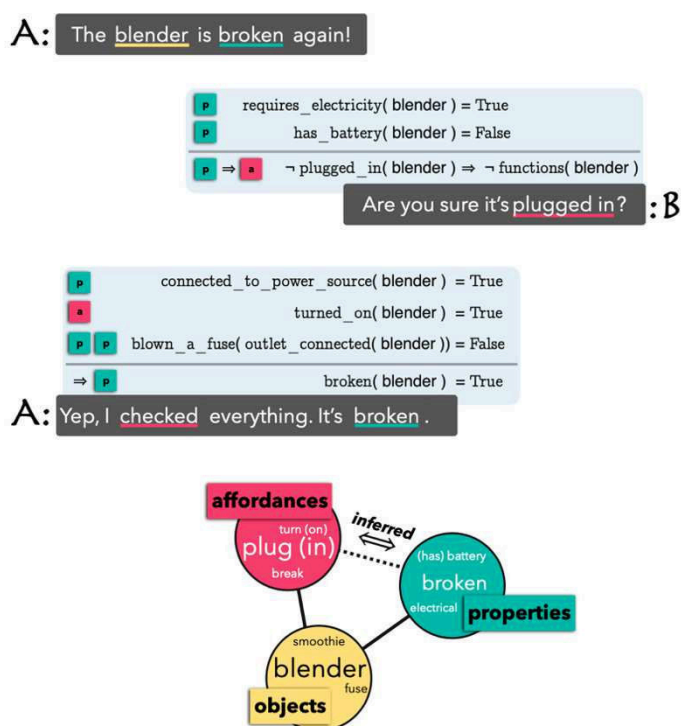
$$\text{pout} - \text{into}(x) \rightarrow \text{holds} - \text{liquid}(x) \quad (6)$$

Súčasný modely dosahujú primeraný výkon pri predpovedaní kompatibility medzi vlastnosťami a akciami objektov. Na základe výsledkov výkonu (Forbes, Holtzman a Choi, 2019) rozdelili vlastností objektov do 4 tried (kde prvé 3 triedy nie sú percepčné), zoradených podľa výkonu zostupne:

- Funkčná („is used for cooking“): priamo viazaná na činnosti objektu.
- Encyklopedická („is an animal“): zahŕňa základné vlastnosti objektu.

- Opierajúca sa o zdravý rozum („comes in pairs“): korelujúca s viacerými slovesami.
- Percepčná („is smooth“): o ktorej črtách možno písať jednoduchými zmyslovými slovesami, čo jej dáva menej implicitných dôkazov.

Obr. 14 demonštruje príklad odvodzovania vlastností zo slova „blender“. Súčasnú neurónové modely sú zásadne obmedzené vo svojej schopnosti fyzikálne uvažovať o svete. Nie veľkosť datasetu, ale komunikácia pomocou implicitného modelu sveta (priamym dopytovaním ľudí) by umožnila neurónovým modelom prístup k takej zručnosti (Forbes, Holtzman a Choi, 2019).



Obr. 14: Odvodzovanie o akciách aplikovateľných na objekt na základe jeho vlastností (Forbes, Holtzman a Choi, 2019)

3.5.4 Symbolic reasoning

Úspechy predtrénovaných jazykových modelov podnietili záujem o zistenie ďalších jazykových schopností, ktorými disponujú. To, či sú súčasné LMs užitočné pre úlohy symbolického uvažovania sa rozhodli výskumníci (Talmor et al., 2020) demonštrovať na 8 úlohách, ktoré si vyžadujú operácie ako porovnávanie, spájanie a skladanie. Schopnosťou symbolického uvažovania sú schopnosti jazykových modelov porovnávať čísla, porozumieť vyvodeným vlastnostiam objektu, či vykonávať viacsokové zloženie faktov. Autori úlohám podrobili jazykové modely ako sú BERT (trénovaný na veľkom datasete pomocou MLM jazyka) a RoBERTa (podobná architektúra,

avšak trénovaný na 10-krát väčšom a optimalizovanejšom datasete). Pri rôznych architektúrach a stratových funkciách boli modely podnecované k rôznym úlohám:

- Dokáže LM vykonávať robustné numerické porovnanie?
- Dokáže LM porovnávať veľkosti objektov?
- Dokáže LM odlíšiť „vždy“ od „často“?
- Dokáže LM porovnávať vekové rozdiely?
- Dokáže pri predikcii LM vziať do úvahy vyskytnutú negáciu?
- Dokáže LM zachytiť sémantiku antoným?
- Dokáže LM porozumieť výslednej konjunkcii faktov?
- Dokáže LM nájsť spoločné vlastnosti dvoch objektov?

Hoci RoBERTa preukázala symbolické uvažovanie (rozumové schopnosti) ktoré v iných modeloch chýbali, výkon LM v mnohých úlohách uvažovania je zatiaľ stále slabý (Talmor et al., 2020).

3.6 Commonsense zdroje

K formalizovaniu a kodifikovaniu znalostí *commonsense knowledge acquisition* (CSKA) podobným ľuďom (*commonsense knowledge*) sa v počiatočných štádiách využívala manuálna anotácia systémových expertov. S pokrokom strojového učenia sa spôsob presunul na odvodzovanie znalostí z textových zdrojov pomocou porovnávania vzorov. Tieto zdroje sú robustné a ich znalosti sú vyjadrené explicitným spôsobom. Pri priamej extrakcii z textu sú počítačové systémy schopné učiť sa nepretržite a postupne, a svoje znalosti získavajú z webu každý deň do nikdy nekončiacej znalostnej bázy (KB) (Batista Dos Santos et al., 2019).

Iným prístupom je automatické odvodzovanie znalostí na základe už existujúcich. Tento spôsob smeruje k implicitným tvrdeniam a reprezentáciou znalostnej bázy ako grafu (knowledge graph) sa stáva pre bázu veľmi obohacujúcim zdrojom encyklopedickej znalosti. Aby sa zvýšila efektívnosť manuálneho zhromažďovania vedomostí, využívalo sa kolektívne úsilie prostredníctvom verejných platforiem, ako sú webové stránky a hry zamerané na *crowdsourcing* (Alhussien, Cambria a Nengsheng, 2018). Prechodom od explicitného k implicitnému sa zlepšila i sémantická analýza textu (Cambria et al., 2018).

Vývoj nových grafov znalostí zdravého rozumu (CSKG) bol ústredným pokrokom pri reprezentácii takej znalosti v oblasti porozumenia prirodzeného jazyka. Ich rôznorodé fakty môžu byť modelmi strojového učenia použité a odkazované na riešenie náročných úloh. Avšak tie manuálne skonštruované nemajú také pokrytie aby boli použité agentami v úlohe NLP pretože sa z nich nedá efektívne naučiť implicitná

Tab. 1: Prehľad populárnych bázy znalostí

Metóda extrakcie znalostí	Báza znalostí
Z textu – automatická	NELL (Batista Dos Santos et al., 2019) SenticNet 5 (Cambria et al., 2018) Dbpedia (Alhussien, Cambria a Nengsheng, 2018)
Od ľudí – experti	Cyc (Davis a Marcus, 2015)
Od ľudí – bez expertízy	ConceptNet (Alhussien, Cambria a Nengsheng, 2018) ATOMIC (Alhussien, Cambria a Nengsheng, 2018)

znalosť (Hwang et al., 2021). Predtrénované jazykové modely síce reprezentujú znalosti zdravého rozumu (Petroni et al., 2019) – avšak tieto výsledky pozorovania sú vymedzené najmä na podskupiny znalostí, ktoré zahŕňajú taxonomické znalosti explicitne uvedené v texte (Davis a Marcus, 2015). K znalostným grafom taxonómnych a lexikálnych znalostí patria napríklad Nell, OpenCyc 4.0 či ConceptNet. Označujú sa ako „knowledge of what“ (Hwang et al., 2021).

COMET

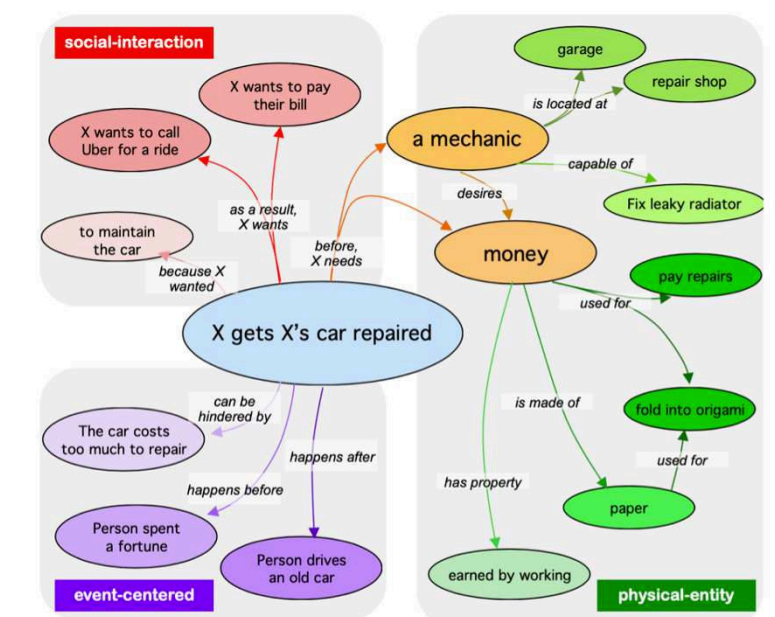
Transformátor COMET prispôsobuje predtrénované jazykové modely trénovaním na vzorkách grafov znalostí zdravého rozumu – konkrétne ConceptNet a ATOMIC, zobrazené na Obr. 15. Tieto modely boli aplikované na nové úlohy, ako je generovanie sarkastických komentárov, terapeutické chatboty a automatizované generovanie príbehovej zápletky. Takto vzniknuté znalostné modely sa označujú ako „knowledge of why and how“ – schopnosť vyvodenia príčin a následkov (Hwang et al., 2021). Vstupom transformátoru je hlava, skúmaný vzťah a výstupom je cieľová fráza:

$$\textit{Take a nap (head) + Causes (relation) = have energy (tail)} \quad (7)$$

Konverzační agenti sa stávajú prominentnými v našom každodennom živote vďaka pokroku v rozpoznávaní reči, spracovaní prirodzeného jazyka a strojového učenia. K vedeniu bohatého rozhovoru s ľuďmi patria i nevy povedané domnienky (*common-sense presumption*) o *if-then-because* podmienkach. Pre túto úlohu uvažovania je COMET veľmi vhodný, pretože je možný priamy dopyt na predpoklad, následok, akciu a cieľ, napríklad nalievanie kávy je základným predpokladom pre pitie kávy (Denecke, Vaaheesan a Arulnathan, 2021).

3.7 Response design patterns

Väčšina neurónových modelov trénovaných na rozsiahlych dátach výrazne pokročila v generovaní konverzácie za cieľom zlepšiť kvalitu jej obsahu (relevantnosť a gramatika). Avšak, faktor emócií pri generovaní vhodnej odpovede nebol v súčasných neurónových modeloch výrazne riešený. Pri zohľadnení emócií je model emocionálne zapojený a empatickejší (Obr. 16). Využitie emočnej inteligencie pri generovaní odozvy je zásadné pre napodobenie ľudskej konverzácie (Khatua et al., 2017).



Obr. 15: Ukážka malej podmnožiny znalostí zdravého rozumu ATOMIC-2020 – atlasu rozsiahlych textových popisov o každodenných inferenčných poznatkoch o entitách a udalostiach, ktoré kódujú sociálne a fyzické aspekty bežnej ľudskej skúsenosti (Hwang et al., 2021)

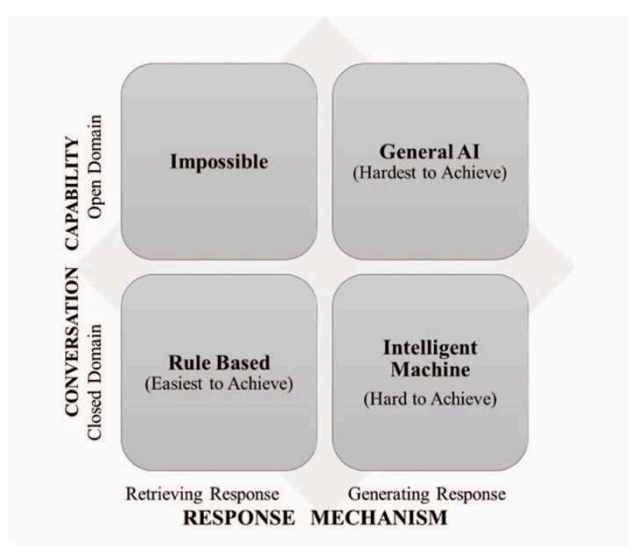
<p>User: Worst day ever. I arrived late because of the traffic. Basic Seq2Seq: You were late. ECM (<i>Like</i>): I am always here to support you. ECM (<i>Happy</i>): Keep smiling! Things will get better. ECM (<i>Sad</i>): It's depressing. ECM (<i>Disgust</i>): Sometimes life just sucks. ECM (<i>Angry</i>): The traffic is too bad!</p>

Obr. 16: Konverzácie bez a so zapojením emócií (Zhou et al., 2018)

Schopnosti konverzácie s ohľadom na rozsah diskusie a mechanizmus odozvy sú 2 dôležité aspekty dialógového systému. Taký systém sa skladá z viacerých komponentov, ako sú rozpoznávanie reči, interpretátor prirodzeného jazyka, sledovateľ emočného stavu, generátor odozvy a generátor prirodzeného jazyka (Obr. 17).

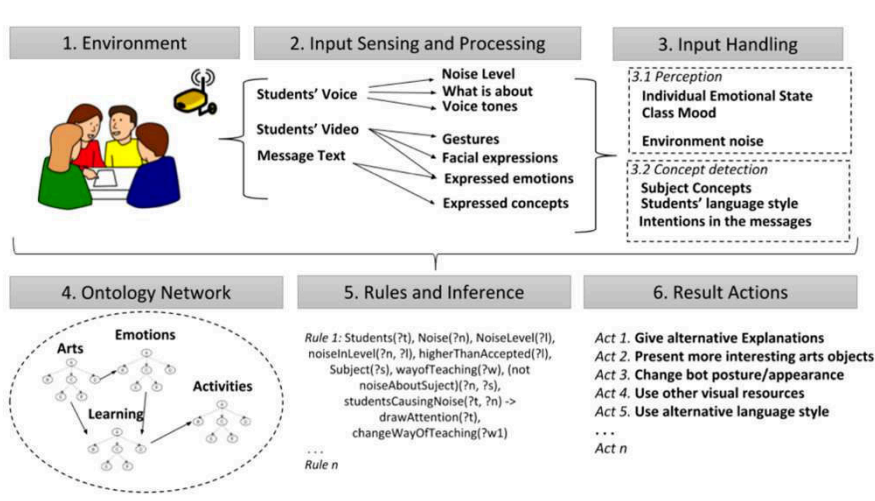
3.7.1 Rule-based

Boti, ako virtuálni agenti s ktorými ľudia môžu komunikovať pomocou textových správ, sú väčšinou vyrobení s cieľom napodobniť osobu v rozhovoroch. Kľúčovou vlastnosťou takých agentov je schopnosť odpovedať na otázky a plniť úlohy prostredníctvom konverzačných dialógov. Jeden z konceptov vyjadrenia afektívneho správania s cieľom zlepšiť empatiu agentov sú formálne pravidlá (*rule-based, hand-crafted, templates*).



Obr. 17: Výsledok výskumu možností dialógového systému (Khatua et al., 2017)

Sociálny kontext zohráva v konverzáciách kľúčovú rolu. Je to prostredie, v ktorom agent jedná podľa situácie na základe jeho definovaných rolí – vlastnej kultúrnej normy. Justo et al. (2018) definovali pravidlá pre vyjadrenie afektívnych vzťahov a určenie emocionálneho stavu osoby. Cieľom je umožniť agentom vybrať si vhodné správanie, aby si zachovali kontextové, pohodlné a príjemné zapojenie s používateľom. Model integrovaním ontologickej siete obohatili o zlepšenie porozumenia synonymým, vzťahov medzi slovami a analýze jazyka bez spoliehania na databázu otázok a odpovedí (Obr. 18).



Obr. 18: Textové správy na vstupe sú predmetom priamejšej interakcie s robotom a sú rozdelené na ich konceptný význam a detegované emócie (Justo et al., 2018)

Menej náročná implementácia *rule-based* konceptu je založená na jednoduchých

pravidlách porovnávania vzorov (pattern matching rules). Umožňuje eliminovať problémy so spracovaním prirodzeného jazyka a rýchlo nastaviť jednoduché scenáre dialógu. Augello et al. (2016) navrhli model sociálneho chatbota pre hru na učenie komunikačných znalostí založenú na sociálnej praxi, ktorá dokáže vybrať najvhodnejší dialóg. Kritika autorov o zahrnutí sociálneho kontextu upozorňuje na nadmernú zložitosť implementácie chatbotov a komplikovanosti písania znalostnej bázy. Dialógy určujú, ktoré odpovede musia byť vrátené na užívateľský vstup chatbota (Denecke, Vaaheesan a Arulnathan, 2021). Vo vzťahu k určenej emócií prebieha dialóg tak, ako je určené vo vytvorených dialógoch pre danú emóciu (Obr. 19). V závislosti na organizácii toku konverzácie, ktorá poskytuje vhodné reakcie na intenty a entity, má každý uzol logické kritéria, ktoré musia byť splnené k úspešnému spusteniu príslušnej reakcie (Ralston et al., 2019). Ak je emócia používateľa detegovaná napríklad ako strach alebo smútok, spustí sa príslušný uzol ktorý ponúka odpoveď, ktorá je vhodná pre náladu užívateľa.

Welcome Dialogue	In this dialogue, the welcoming of the user is administered. If started for the first time, the chatbot asks for the user's name and consent to the privacy policy. Then, the user is asked about his mood. If the mood was already entered up to three hours ago, the welcome dialogue is based on the last mood.
Joy Dialogue	As soon as the user states that he is in a good mood, the Joy Dialogue is called. SERMO asks for the reason of the good mood and finally asks if he wants to do a task. If so, the HashTag dialogue is started.
Normal Dialogue	This dialogue is started when the user is in a balanced mood. SERMO asks for the reason of the mood and starts the Emotion Dialogue to determine the emotion.
Sadness Dialogue	This dialogue is started when the user states he is sad. The Emotion Dialogue is started to determine the emotion.
Emotion Dialogue	After the mood has been selected, the emotion dialogue is executed. This dialogue implements the ABC theory. The user is asked about the situation or event, his thoughts and feelings. Based on the replies, the emotion is recognized and passed forward to the appropriate emotion dialogue (i.e. Fear Dialogue, Anger Dialogue, Grief Dialogue, Sadness Dialogue, Joy Dialogue).
Anger Dialogue	The dialogue handles the emotion anger. The user is informed on the different types of anger (appropriate anger and inadequate anger). Further, a pleasant activity is suggested.
Fear Dialogue	This dialogue concerns the emotion fear. The user is provided with information on reasonable and inadequate fear. Finally, he is asked to transform the fear-provoking thoughts into positive thoughts.
Grief Dialogue	The dialogue handles the emotion grief. The user is informed on the different phases of grief. Further, activities for distraction are suggested.
Other Dialogue	Further measures are proposed to the user in this dialogue. One dialogue is about improving the user's mood and the other allows the user to plan the day. In addition, the user can select mindfulness exercises or the option Nothing.
Improved Mood Dialogue	In this dialogue, various activities are suggested to the user which could improve his current mood. After having carried out an activity, he has the possibility to carry out another activity.
HashTag Dialogue	This dialogue manages specific interactions that are triggered by the user using a hashtag. In its current implementation two interactions are available: #todo show a list of tasks for the current day, #strengths shows a list of strengths of the user. Both lists can be adapted by the user.
Activity Dialogue	In this dialogue, mindfulness exercises are suggested and, if selected, the user is redirected directly to the exercise on the Activities page and the exercise is started.
Goodbye Dialogue	This dialogue manages the ending of the conversation.

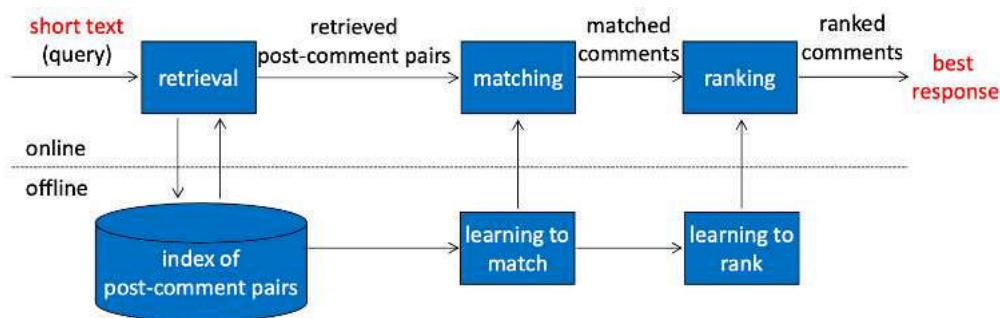
Obr. 19: 13 vyvinutých dialógov pre realizáciu chatbotovej konverzácie (Denecke, Vaaheesan a Arulnathan, 2021)

Hoci prepracované pravidlá môžu fungovať dobre, problémom konceptu založenom na pravidlách je vysoká cena a neschopnosť zvládnuť nečakané výroky. I napriek tomu je tento spôsob štandardom pre konvenčné prístupy vytvárania konverzačného systému v otvorenej doméne (Higashinaka et al., 2014). Generovanie textovej odpovede sa v implementáciách môže vyskytovať i s generátormi gest, pohľadov a výrazov tváre (Cassell et al., 1944).

3.7.2 Retrieval-based

Non-task oriented a open-domain konverzačné systémy priťahujú pozornosť vedy pre ich funkčné, sociálne a zábavné role v aplikáciách reálneho sveta. Je to zaujímavý, ale notoricky náročný problém vzhľadom k tomu, že používateľ môže systému potvrdzovať akýkoľvek vstup. Taký spôsob nešpecifikuje doménové znalosti a počet možných kombinácií stavov konverzácie je nekonečný. Je to predpoklad, žeby konvenčne vytvorené pravidlá a šablóny pri generovaní odpovedí zlyhali (Yan, Song a Wu, 2016).

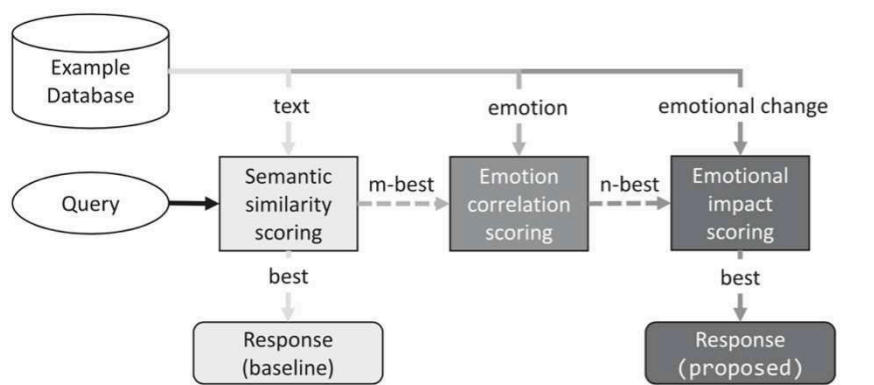
Pri koncepte modelu založenom na vyhľadávaní je odpoveď definovaná dopredu pomocou vytvorených pravidiel, pravidelného párovania alebo trénovaného klasifikátora hlbokého učenia pre výber najlepšej odpovede z databázy (Sun et al., 2018). Pre vytvorenie databázy sú extrahované dva po sebe idúce výroky ako vstupné-výstupné páry a indexované pomocou textového vyhľadávača. Pre daný výrok sa prvých desať párov výrokov získa na základe podobnosti medzi požiadavkou a vstupnou časťou indexovaných párov. Následne sa náhodne vyberie jeden zo získaných párov, aby vytvoril ďalší výrok systému (Higashinaka et al., 2014). Iným spôsobom spracovania získaných kandidátskych výrokov je pridanie ich kontextov – jedného alebo viacerých predchádzajúcich výrokov v aktuálnej relácii konverzácie. Pomocou hlbokéj neurónovej siete je teda možné získať množinu preformulovaných otázok s pôvodnými otázkami a týmto spôsobom je do konverzačného systému začlenených viacero dimenzií hodnotiacich dôkazov vrátane otázok, kontextov, príspevkov kandidátov a odpovedí (Yan, Song a Wu, 2016). Zadaním správy systém získa súvisiace odpovede z veľkého úložiska údajov o konverzácii, zoradí komentáre kandidátov na základe ich skóre a vyberie komentár s najvyšším skóre pre očakávanú odpoveď (Obr. 20).



Obr. 20: Systém architektúry pre retrieval-based model (Ji, Lu a Li, 2014)

Lubis et al. (2019) rozširujú kritériá výberu odpovede o emocionálne parametre, aby vybrali odpoveď, ktorá má najväčší potenciál vyvolať pozitívne emócie. Do svojej architektúry integrujú korpus na vyvolanie pozitívnych emócií získaním nových odpovedí, ktoré vyvolajú pozitívnejšie emócie pomocou navrhovanej metódy vyhľadávania. Takouto optimalizáciou dosahujú neurónový dialógový systém orientovaný

na chat, ktorý zachytáva emocionálny stav používateľa a zohľadňuje ho pri vytváraní dialógovej odpovede (Obr. 21).



Obr. 21: Výber so sémantickou podobnosťou a parametrami emócií (Lubis et al., 2019)

3.7.3 Generation-based

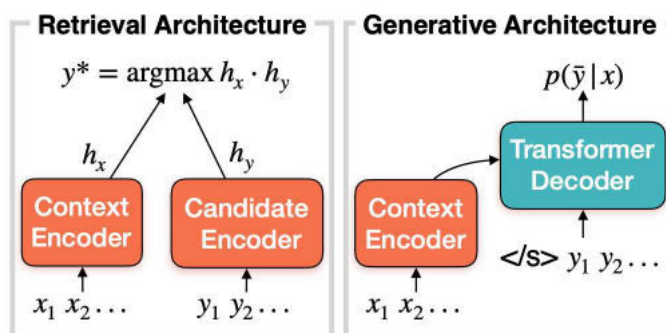
Generovanie reakcie chatbota v tomto prípade nezávisí na vopred definovaných odpovediach ale na procese tréovania, vďaka ktorému pre danú skupinu algoritmov systém vygeneruje odpoveď z daného vstupu (Sun et al., 2018). Systém (Hasegawa et al., 2013) aplikuje model štatistického strojového prekladu – *statistical machine translation* (STM) pre generovanie odpovede, ktorá zodpovedá cieľovej emócií užívateľa. V tomto prípade je odpoveď vnímaná ako preklad vstupného výroku, a pomocou bežných nástrojov na strojový preklad sa model učí z párov výroku a odpovede. Autori skúmajú oblasť predpovedania a vyvolávania emócií pomocou veľkého množstva dialógových údajov získaných z mikrobloginých príspevkov.

Experimenty ukazujú (Ritter, Cherry a B. Dolan, 2011), že techniky STM sú vhodnejšie pre úlohy generovania reakcie a tréovaním na rozsiahlych neobmedzených korpusoch, a je možné vytvoriť systém, ktorý je flexibilný a rozširiteľný pri prevádzke v otvorenej doméne. Obe štúdie [(Sun et al., 2018) a (Ritter, Cherry a B. Dolan, 2011)] používajú na nájdenie najlepšej odpovede pre daný vstup Mosesov dekodér.

Dvojitá architektúra *retrieval a generation* sa v experimentoch používa pre variabilitu modifikovateľnosti modelu do evaluačných procesov. Experimentovaním s BERT modelom či iným transformátorom v rámci *retrieval-based* architektúry, model vyberá kandidáta s výrokom s najväčším *softmax* koeficientom. U *generation* architektúry, tzv. *full transformer architecture* existuje dekodér, ktorý predikuje sekvenciu slov y (Obr. 22).

3.7.4 Trustworthy design

Hoci súčasné modely emocionálnej generatívnej konverzácie dosiahli sľubné výsledky, autori (Hou et al., 2020) upozorňujú na problém tendencie generovať triviálne alebo



Obr. 22: Dvojitá architektúra modelu (Rashkin et al., 2019)

všeobecne relevantné odpovede s malým významom (napr. „Haha“, alebo „I love you“) práve kvôli pridaniu emocionálneho faktora. Vysvetľujú, že ďalším problémom môže byť absencia relevantnosti témy v odpovediach chatbota (Obr. 23).

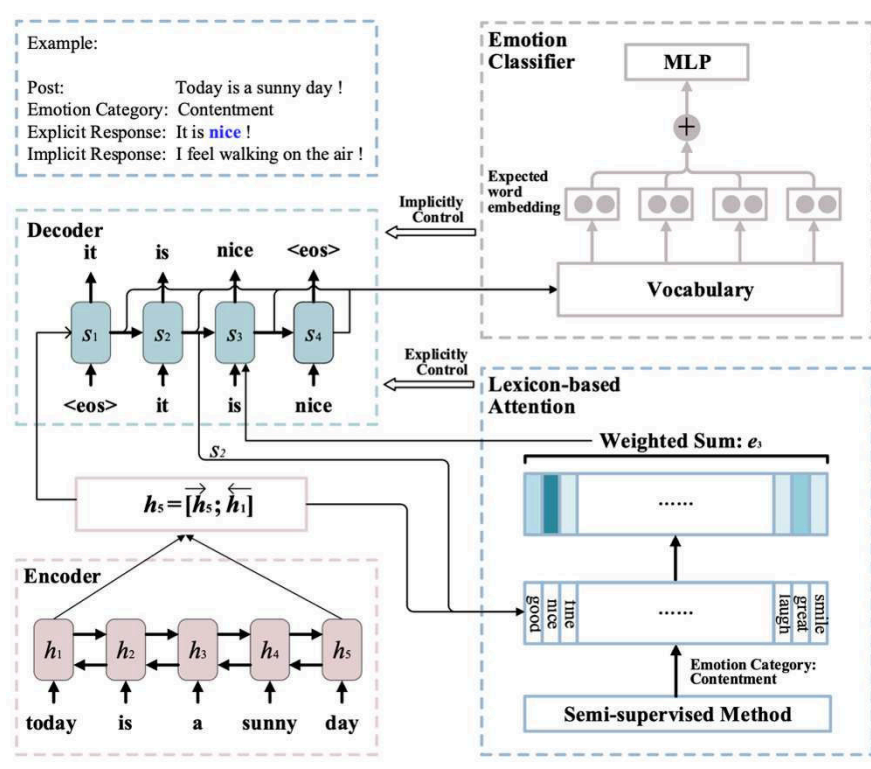
Emotion Label	Message: My favorite sport is playing basketball.	
	Generated responses	Real-life responses
Happy	What a happy day!	I will be very excited after playing basketball.
Like	I like it.	I also like playing basketball.
Disgust	I hate playing piano.	The basketball game is too difficult for me.
Sad	I am so sad.	The team lost yesterday, and the players were frustrated.
Angry	I am so angry.	I am very annoyed because I cannot throw the ball into the basket.
Null	We will not play basketball tomorrow.	No one doubts Kobe's talent on basketball.

Obr. 23: Dôležitosť relevantnosti témy v systéme emočného dialógu (Hou et al., 2020)

Na usmernenie procesu generovania odpovedí ako prvá štúdia (Benke, 2020) použila klasifikátor emócií, ktorý zaisťuje, že špecifická emócia je vo vygenerovaných odpovediach primerane vyjadrená buď explicitným alebo implicitným spôsobom. Architektúra systému emočného dialógu pozostáva z dekodéra, ktorého výstupom je zmysluplná odpoveď so špecifickou emóciou za pomoci klasifikátoru emócií a zapojením lexikónu (Obr. 24).

3.7.5 Non-progress design patterns

Súčasťou *task-oriented* konverzačných agentov sú i prípady, keď sa plnenie potrieb používateľa nestretne s úspechom. Agent buď nesplní žiadosť používateľa alebo nesprávne intent klasifikuje, čo môže spôsobiť opustenie konverzácie alebo aplikovanie vyrovnávajúcich stratégií používateľa. Identifikovanie konverzačných nápodiev



Obr. 24: Framework dialógového systému rozšírený o mechanizmus lexikónu z dôvodu zapojenia požadovaných emocionálnych slov (Song et al., 2019)

o opustení dialógu s chatbotom, a skúmanie mentálnych modelov interakcií smerom k spolupráci s chatbotom na dosiahnutí výsledku, sú súčasťou viacerých štúdií [(Weisz et al., 2019), (Li, Chen a Chang, 2019), (Shevat, 2017)].

Nájdienie optimálnej sady konverzačných interakcií na dokončenie užívateľskej požiadavky je kľúčom k navrhovaniu úlohou orientovaného dialógu (Shevat, 2017). Každý vstup používateľa umožní chatbotovi posunúť konverzáciu ďalej. Pokrokom sa rozumie poskytnutie informácií používateľovi od chatbota a vice versa s cieľom posunúť sa na ďalšiu žiadosť (Li, Chen a Chang, 2019). Ak pokrok nenastáva, zdá sa, že vina smeruje na chatbota a nie na formulovaný ľudský výrok. Ľudskí účastníci konverzácie (Weisz et al., 2019) uvádzajú stratégie pre používateľa, akými minimalizovať nedostatočný pokrok a zvýšiť pravdepodobnosť pochopenia žiadosti: používať jednoduché slová, neuvádzať zbytočné podrobnosti, preformulovanie otázky či explicitné vyjadrenie zámeru. V prípade výskytu absencie pokroku – *non-progress* (NP), sa správanie užívateľov zaraďuje do viacerých stratégií podľa zastúpenia výskytu (Obr. 25).

Chatbot v reálnom svete reaguje na nejednoznačné otázky odhadom (výber najpravdepodobnejšieho intentu, ktorý sa zhoduje s výrokom) alebo výzvou (k úprave nejednoznačného kontextu či k zjednodušeniu otázky). Užívateľská analýza (Weisz

Message reformulation		
C1	add words	6.68%
C2	remove words	4.76%
C3	rephrase	8.82%
C4	repeat	5.75%
C5	ask new topic	5.48%
C6	others	3.56%
Quitting		
C7	quit subject temporarily	27.16%
C8	quit conversation temporarily	6.74%
C9	switch subject	13.47%
C10	abandon chatbot service	17.58%

Obr. 25: Adaptačné stratégie používateľa pri výskyte NP v úlohou orientovanom dialógu (Li, Chen a Chang, 2019)

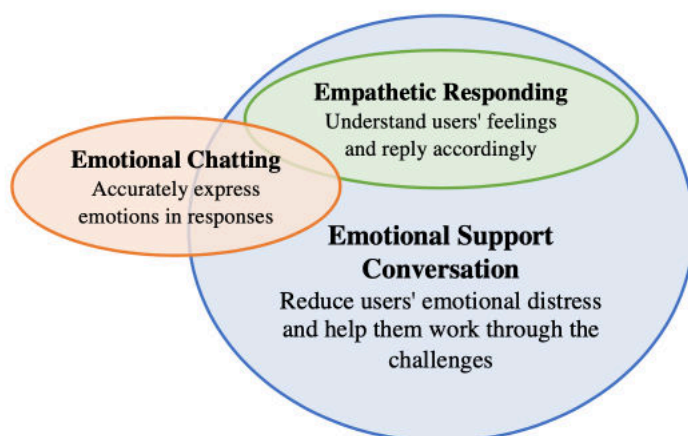
et al., 2019) ukazuje, že používateľ pravdepodobne opustí konverzáciu pri viacnásobnom využití odhadu či úprave kontextu. Poznanie kontextu ale i tak nie je determinantom k úspešnému splneniu žiadosti užívateľa. Jedným zo spôsobov ako predísť NP je uchopenie princípu transparentnosti v užívateľskom rozhraní o vnútornom fungovaní systému (Li, Chen a Chang, 2019), a to deklarováním účelu, ku ktorému bol chatbot vytvorený (formou onboardingu v aplikácii) a edukáciou používateľa ako chatbota používať (Shevat, 2017).

3.8 Dialog systems

Profilovaním schopností konverzačného modulu na *chit-chat* konverzácie vzniká relevantný predpoklad, aby systém empaticky reagoval na bežné spoločenské konverzácie. Tie sa v interakcii človek - stroj vyskytujú často, a vedie to k tomu, že ľudia interagujú so strojmi prirodzeným a sociálnym spôsobom (Rashkin et al., 2019). V nasledujúcich podsekcích sa zástupca dialógu vyskytne v sociálnej roli *listener*, ďalej aj ako *supporter*, čo odráža jeho empiricky odvodené vlastnosti: vhodne vyjadriť emócie v odpovediach, porozumieť pocitom a relevantne odpovedať, až po schopnosť redukcie emočného stresu (Obr. 26).

3.8.1 Empathetic responding

Referenčným bodom pre novú generáciu empatických dialógov bol dataset *EmpatheticDialogues* (ED) (Rashkin et al., 2019), ako nový štandard pre empatickú reakciu na konverzačného partnera, ktorý do konverzácie vstupuje s určitou osobnou skúsenosťou. Tento dataset s 25 000 konverzáciami sa zameriava na rozhovory o emocionálne podložených osobných situáciách, kde súbor emócií je rozšírený na 32 typov



Obr. 26: Kompozícia state-of-the-art dialógového modelu (Liu et al., 2021)

(Obr. 27). Cieľom bolo predstaviť dataset s veľkým potenciálom na meranie schopnosti chat modelov reagovať empatickým spôsobom.

Varianty modelov vytvorených týmto datasetom boli autormi ohodnotené podľa ich schopnosti reakcie poslucháča na príbeh hovoriaceho. Klasifikátor bol trénovaný s učiteľom (predikcia emócie a témy/kontextu z popisu situácie napísanej pred dialógom ako vstupná *supervised* informácia), čo je možné považovať za predikát k zvýšeniu výkonnosti vysokokapacitných modelov založených na BERT. I vďaka tomu sa hodnotenie empatie blížilo k ľudskému výkonu (Rashkin et al., 2019). Ľudské hodnotenie ako súčasťou metrik zahŕňovalo nasledovné:

- Ukázala odpoveď porozumenie pocitov osoby hovoriacej o svojej skúsenosti?
- Boli odpovede relevantné k téme konverzácie?
- Bol jazyk odpovede presný a zrozumiteľný?

Label: Proud

Situation: Speaker felt this when...

“My little dog learned to sit!”

Conversation:

Speaker: I finally taught my new little puppy his first trick!

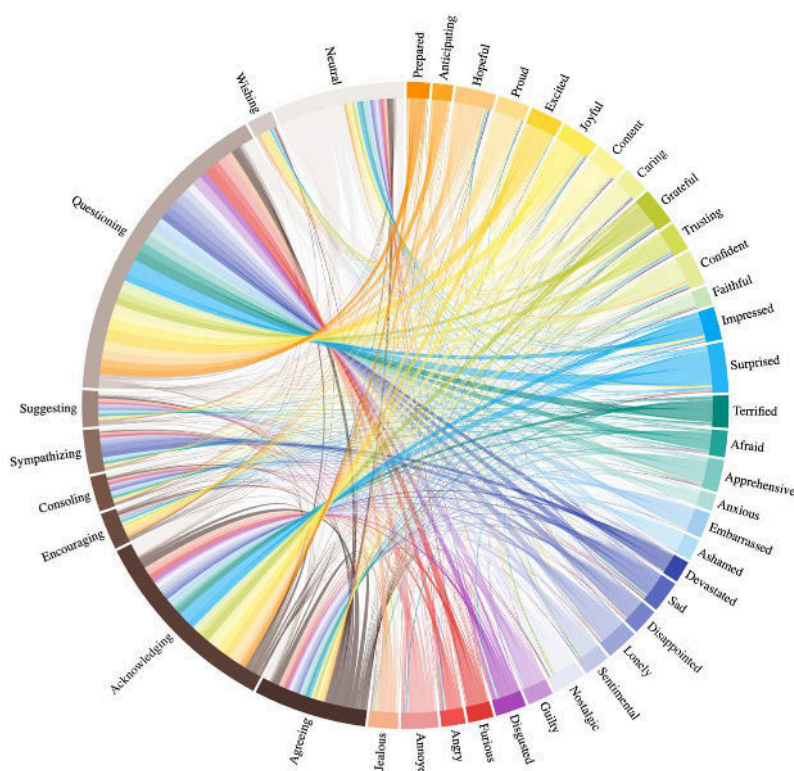
Listener: What trick did you teach him?

Speaker: I taught him to sit for a treat, its so cute.

Listener: That is good, do you plan to teach him more tricks?

Obr. 27: Ukážka tréningovej sady pre daný emočný štítok (Rashkin et al., 2019)

EmpatheticDialogues (ED) podnietili radu ďalších výskumov na ceste k empatickejším dialógovým systémom. Doplňkovou prácou, ktorá ale neprehlbuje využitie datasetu ED, je anotačný mechanizmus najčastejších vzorcov výmeny emócií a ich zmeny v priebehu interakcie. Tieto výsledky objasňujú časté vzorce empatickej konverzácie ľudí (Obr. 28), ktorí sa zapájajú do spoločenských debát (Welivita a Pu, 2020). Opísané metódy možno použiť ako anotačnú schému aplikovateľnú na viacero datasetov s cieľom jednotlivého posudzovania ich kvality empatie.



Obr. 28: Výmena emócií a zámerov medzi poslucháčom a rečníkom (Welivita a Pu, 2020)

Roller et al. (2021) svojim výskumom o dva roky neskôr ukazujú, že veľké modely sa môžu naučiť zručnosti otvorenej/všeobecnej domény s udržaním kontextu ak majú vhodné tréningové dáta a *generation-based* architektúru. K posudzovaniu kvality konverzácie agenta pripisujú priamu úmeru k dĺžke výrokov počas konverzácie. Pre účely *fine-tuning* využili okrem vyššie opisovaného *Empathetic Dialogues* – vhodné pre modely prejavujúce väčšiu empatiu (Rashkin et al., 2019), i tieto účelné, a pre tému spoločenského rozhovoru významné dátové sady:

- ConvAI2: vhodný pre modely podmieňujúce osobnosť a históriu dialógu.
- *Wizard of Wikipedia* (WoW): hlboká znalosť vhodná pre modely keď sa diskutuje o témach do hĺbky.

- *Blended Skill Talk* (BST): vhodný pre modely, ktoré kombinujú vlastnosti predošlých dvoch datasetov (osobnosť z ConvAI2, empatiu od ED, znalosti z WoW).

Výsledný vysokokapacitný *open-domain* BlenderBot model sa výkonom približuje ľudskosti, avšak autori vnášajú do diskusie očakávanie, že riešenie problému všeobecnej konverzácie je ešte len na dosah. Zároveň upozorňujú na limity existujúcich riešení spočívajúc v často opakovaných frázach v reakciách chatbota, používaním príliš jednoduchého jazyka a halucinovaním vedomostí (Shuster et al., 2021), kde modely generujú hodnoverne vyzerajúce tvrdenia, ktoré sú fakticky nesprávne a často si zamieňajú fakty medzi dvoma podobnými entitami. Ak nie je daná len otázka, ale celý kontext dialógu ako vstup, úloha získania odpovede je ťažšia jednak kvôli dlhšiemu kontextu a jednak kvôli potrebe nájsť podporné znalosti na vedenie rozhovoru, a nie jeden fakt na zodpovedanie otázky. Takéto konverzačné modely musia preukázať znalosť a vecnosť odpovedí.

3.8.2 Emotional support conversation

Emocionálna podpora má za cieľ znížiť emocionálny stres jednotlivca a pomôcť mu pochopiť a zvládnuť výzvy, ktorým čelí. Liu et al. (2021) navrhujú framework *Emotional Support Conversation* (ESC), ktorý pri oboznamovaní sa so situáciou žiadateľa o pomoc vyjadruje porozumenie a empatiu pomocou rôznych vyrovnávacích stratégií (Obr. 29). Tento mílnik výskumu konverzačných agentov kategorizuje strany dialógu ako *seeker* a *supporter*. Ako kostry porovnávaných variantných modelov využili dva najmodernejšie predtrénované modely: BlenderBot (Roller et al., 2021) a DialogPT (Zhang et al., 2020). Autori preukázali, že chatbot vo všeobecnosti dodržiava poradie fáz a ich stratégií a tak sa distribúcia stratégií v rôznych fázach konverzácie dá paušalizovať (Obr. 30).

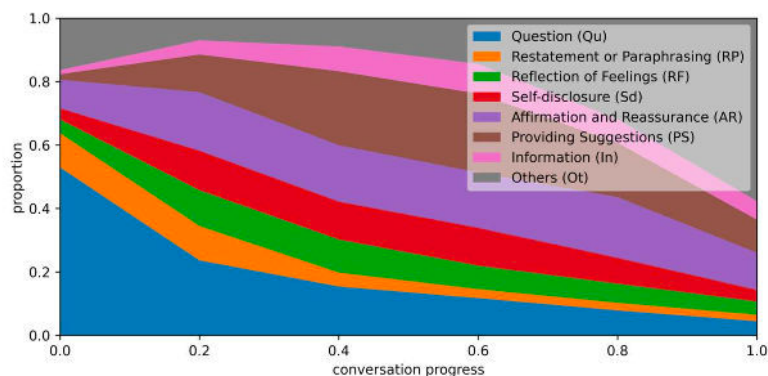
Výber a realizácia stratégie v konverzácii ako aj modelovanie stavu používateľa sú dôležitými smermi pre ďalší výskum emocionálnej podpory chatbota (Liu et al., 2021). Tu et al. (2022) o rok neskôr nadväzujú na nový smer empatickej konverzácie svojím riešením navrhovaného modelu *Mixed Strategy-Aware Model* (MISC) integrujúci framework COMET (Hwang et al., 2021), ktorý pri každej odpovedi zvažuje veľmi detailný (jemnozrnný) dušený stav používateľa (*seeker*) a reaguje pomocou kombinácie stratégií. Autori namiesto modelovania odozvy singulárnou stratégiou ju formulujú ako rozdelenie pravdepodobnosti v číselníku stratégií a vedú generovanie odozvy pomocou ich kombinácií (Obr. 31). O modely MISC je možno tvrdiť nasledovne:

- Je založený na BlenderBot-Small modely (Roller et al., 2021), tzv. model s jednou stratégiou, kde na začiatku generovania odozvy najprv dekoduje špeciálny token stratégie.
- K ohodnoteniu používa porovnanie s ESCConv (Liu et al., 2021), ktorý považujú za *benchmark* medzi najmodernejšími empatickými chatbotmi.

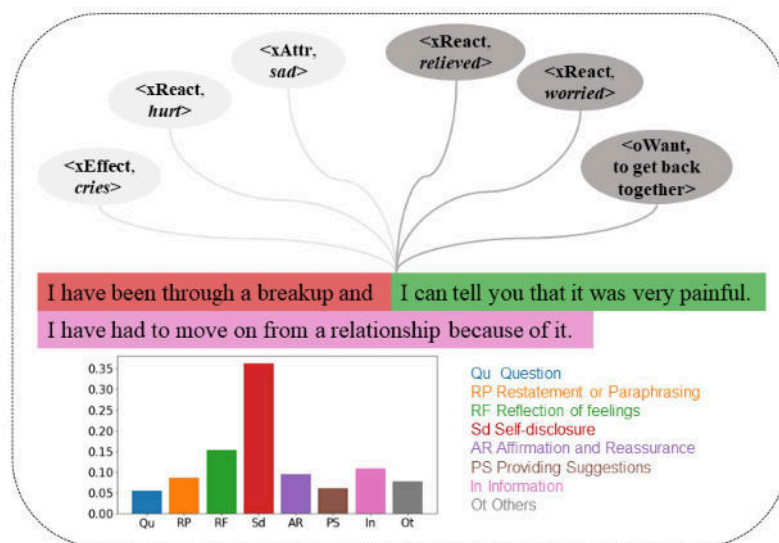
Strategies	Stages	Examples	Lexical Features
Question		<i>Can you talk more about your feelings at that time?</i>	do you (15.0), are you (13.8), how (13.7), what (12.3), do (11.5)
Restatement or Paraphrasing		<i>It sounds that you feel like everyone is ignoring you. Is it correct?</i>	is that (8.2), so you (8.2), it sounds (7.1), correct (7.1), so (6.6)
Reflection of Feelings		<i>I understand how anxious you are.</i>	can tell (7.4), understand how (5.8), are feeling (5.1), tell (5.1), understand (4.9)
Self-disclosure		<i>I feel the same way! I also don't know what to say to strangers.</i>	my (15.3), was (10.5), me (10.2), had (9.7), myself (7.8)
Affirmation and Reassurance		<i>You've done your best and I believe you will get it!</i>	its (5.7), thats (5.6), will (5.4), through this (5.1), you will (4.7)
Providing Suggestions		<i>Deep breaths can help people calm down. Could you try to take a few deep breaths?</i>	maybe (7.3), if (6.5), have you (6.4), talk to (5.8), suggest (5.8)
Information		<i>Apparently, lots of research has found that getting enough sleep before an exam can help students perform better.</i>	there are (4.4), will (3.8), available (3.7), seen (3.3), possible (3.3)
Others		<i>I am glad to help you!</i>	welcome (9.6), hope (9.6), glad (7.3), thank (7.0), hope you (6.9)

Obr. 29: Prehľad stratégií frameworku ESC (Liu et al., 2021)

- K doplneniu konverzačných prejavov o fyzické znalosti (physical commonsense) (Bisk et al., 2020) využívajú COMET (Hwang et al., 2021).
- Takmer vo všetkých fázach konverzácie je predpovedaná stratégia Others menej pravdepodobná ako v BlenderBot.
- Na základe automatických metrick sú generované odpovede relevantnejšie a empatičejšie ako v BlenderBot.
- Prostredníctvom rozsiahlych experimentov dokazuje, že zmiešaná stratégia je efektívnejšia v empatičej open-domain konverzácii ako jedna stratégia.



Obr. 30: Rozloženie stratégií počas rozhovoru (Liu et al., 2021)



Obr. 31: Organizovanie reakcie na základe kombinácie blokov COMET a zmesi stratégií (Tu et al., 2022)

Za účelom strategického plánovania konverzačného postupu v *multi-turn* systémoch emocionálnej podpory, navrhli o rok neskôr Cheng et al. (2022) MultiESC systém, ktorý vykonáva strategické plánovanie s poprednou heuristikou na odhadnutie dlhodobého účinku aplikovania danej stratégie na používateľa. Taký systém má tendenciu proaktívne skúmať situáciu používateľa na začiatku konverzácie namiesto priamej interakcie s jeho aktuálnym stavom. Oproti generovaniu empatickej odpovede (Rashkin et al., 2019) je táto architektúra vhodná pre zmiernenie negatívnej emócie namiesto singulárneho generovania odpovedí konzistentných s emočným stavom používateľa. Autori v roku 2022 vyslovujú limity, že výroky chatbotov sú zvyčajne všeobecné a opakujúce sa a nedokážu ukázať personalizované a hlboko pochopené skúsenosti používateľa a teda poskytnúť konštruktívne návrhy, ako situáciu zmeniť.

4 Metodika

Stanovený problém návrhu a implementácie empatického konverzačného agenta s emočnou inteligenciou pre aplikáciu *Moje MENDELU*¹ prešiel niekoľkými fázami, ktorými sú zber korpusov pre *chit-chat* a *task-oriented* doménu agenta, analýza účelu agenta v zmysle riešenia emocionálneho stavu používateľa a jeho používateľskú privetivosť, zostavenie konverzačného modulu s návrhom architektúry a testovaním na vzorke respondentov.

4.1 Zber datasetov

Pre tento skúmaný problém prináleží najlepšie ten dataset, ktorý disponuje zoznamom týchto vlastností:

- Dvojsmerná/dvojčlenná komunikácia speaker-listener.
- Veľký počet rozhovorov.
- Rozhovory s niekoľkonásobnými obratmi (multi-turn).
- Doména špecifická pre konkrétnu úlohu (FAQ).

4.1.1 Zabezpečený dataset na task-oriented konverzáciu

Neodmysliteľným datasetom pre zabezpečenie *task-oriented* agenta je korpus určený na dopytovanie sa používateľa v rámci domény FAQ zvolenej aplikácie. Tento dataset bude dodaný vo formáte JSON a spracovaný k identifikácii intentu.

4.1.2 Zabezpečený commonsense v porozumení prirodzeného jazyka

Štúdiami často skloňovaná a skúšaná *commonsense knowledge* je v tejto práci zabezpečená frameworkom COMET² (Hwang et al., 2021), ktorý reprezentuje implicitnú znalosť (sociálnu, fyzikálnu a riadenú udalosťami) vo svojich parametroch neurónovej siete. Túto znalosť vyjadruje pomocou generovaných textových novelových popisov, a tak je možné požadovať od agenta popis podmieneného stavu a možného následku pre udalosť, stav alebo cieľ. Znalostná báza na ktorej je framework trénovaný ostáva pre túto prácu bez modifikácie, nakoľko transformovaný jazykový model na model znalostný je už vo vysokej kvalite a konfigurácii.

4.1.3 Zabezpečený dataset pre emocionálnu podporu v konverzácii

Pre simuláciu ľudského poskytovateľa podpory v danej emocionálnej situácii s bohatými podpornými stratégiami bol zvolený dostupný korpus *Emotion Support Conversation* (ESConv) od Liu et al. (Liu et al., 2021). Existuje mnoho typov a úrovní

¹<https://my.mendelu.cz>

²<https://github.com/kearnsw/comet-atomic-2020>

podpory, ktoré môže model poskytnúť, avšak táto práca sa nesnaží suplovať profesionálnu podporu s možnými rizikami, ale aplikuje iba základnú úroveň podporných vlastností agenta pre *day-to-day* aktivitu univerzitého prostredia.

4.2 Analýza a design

4.2.1 Požiadavky na schopnosť spracovania emócií

Rozpoznanie emočného stavu sa pre túto prácu bude definovať na modalite 1 – a teda výhradne z textového vstupu. Získavanie iným spôsobom je pre tento výskum nevhodné, nakoľko ďalší výskumník má po prečítaní tejto štúdie dostať fundamentálny základ pre produkčné prostredie chatbota pre aplikáciu *Moje MENDELU* so všetkými jej doménami a ekosystémom. Nakoľko je mobilná aplikácia toho času singulárna platforma pre takého používateľa, iný spôsob získavania nekoreluje s používateľským očakávaním mobilnej aplikácie.

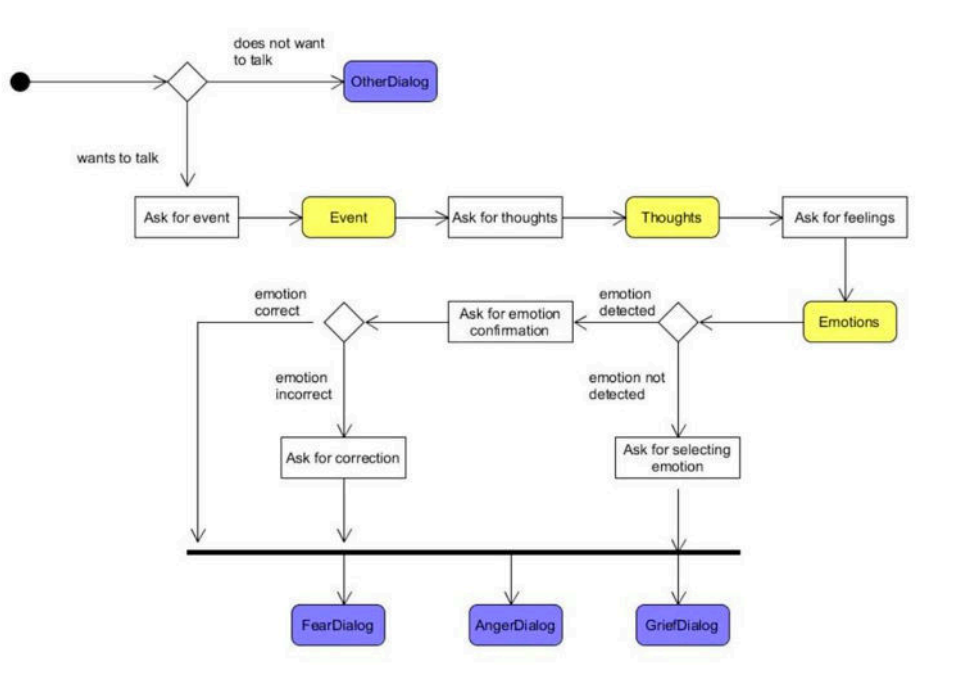
4.2.2 Identifikácia emočného stavu

K zberu dát o duševnom stave nebude v práci využitá metóda *self-assessment* (Larson a Csikszentmihalyi, 2014) dotazníkov ani metóda *ABC dialógu* (Denecke, Vaheesan a Arulnathan, 2021), v ktorých sa agent dopytuje na aktuálnu náladu, na základe ktorej spustí dialóg, aby získal informácie o aktuálnej udalosti, ktorá ovplyvnila používateľa, ako aj o emóciách spojených s udalosťou a na základe toho navrhuje vhodné aktivity a cvičenia (Obr. 32). Cieľom tejto práce teda nie je kontrola a monitorovanie emócií ale práve zmiernenie či zlepšenie afektívneho stavu používateľa takým spôsobom, aby bol zohľadnený kontext sociálno-interakčných vzťahov zahrňujúcich sociálne spúšťané stavy a správanie. Informácie o emóciách používateľa sú spracovávané z každého vstupu, ktorý tvorí kontext dialógu (história) a textový vstup používateľa. Takýmto spôsobom disponuje program chronologickým vývojom duševného stavu a môže nad tým overovať rôzne hypotézy o vypovedanom vplyve stratégií.

4.2.3 Zvolený dialógový systém

Každá reakcia agenta v *chit-chat* konverzácii spadá pod 1 zo 7 definovaných stratégií (Liu et al., 2021) vďaka korpusu ESConv³, ktoré determinujú postup na zmiernenie stresu používateľa v situácii, v ktorej sa nachádza. Keďže tento výskum sa zaoberá konceptom všeobecnej konverzácie, existuje podmnožina stratégií, ktorá udáva experimentálny smer pre odpovedanie agenta. Agent sa teda pohybuje medzi 3 architektúrami generovania odpovede: *task-oriented* definované odpovede, *retrieval-based* odpovede zo zvoleného datasetu, a definované odpovede s maskou, ktorá pokrýva cenné vzťahy získane COMET frameworkom.

³<https://github.com/thu-coai/Emotional-Support-Conversation>



Obr. 32: Diagram spúšťania ABC dialógu chatbota určeného na monitorovanie duševného zdravia používateľa (Denecke, Vaaheesan a Arulnathan, 2021)

4.3 Zostavenie konverzačného modulu

Zámer používateľa sa podľa predloženého návrhu vetví do rôznych úrovní. Každá úroveň má definovanú vlastnú množinu príbehov k úspešnému zvládnutiu pôvodného zámeru. Systémy založené na zámeroch sú zároveň systémy založené na prirodzenom jazyku človeka. Vďaka nim dokážeme identifikovať a predpovedať čo chce používateľ skutočne povedať v konkrétnom kontexte (Shabbir, 2021).

Navrhovaný agent (chatbot) je vyvinutý pomocou frameworku RASA⁴ ako nástroj na riadenie konverzácie, pochopení prirodzeného jazyka, kategorizáciu zámerov a extrakciu entít. RASA je aktuálne často používaný nástroj na vývoj konverzačného systému (Shabbir, 2021). V porovnaní s nástrojom DialogFlow⁵ je RASA *open-source* (Turganaliyev, 2021) modulárna, a teda umožňuje integráciu voľne prepojených modulov pre zabezpečenie všeobecnej konverzácie s empatickým porozumením a riešením emočnej situácie. Hoci RASA, na rozdiel od DialogFlow, neposkytuje komplexné používateľské rozhranie, poskytuje používateľom možnosť prispôbovať *Python* skripty.

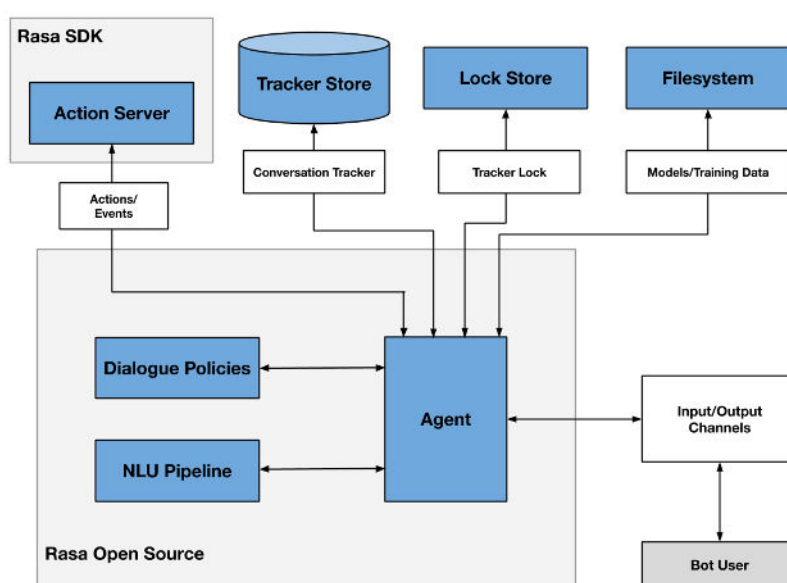
⁴<https://rasa.com>

⁵<https://cloud.google.com/dialogflow>

5 Architektúra konverzačného modelu

5.1 Popis implementačného nástroju

RASA poskytuje silný základ pre vytváranie NLU modelov pre klasifikáciu zámerov, extrakciu entít a pridávanie vlastných komponentov, aby sa vývoj AI agentov posunul na úplne novú úroveň (Rasa, 2022). Architektúra navrhovaného modelu vychádza zo samotnej architektúry RASA. Na Obr. 33 sú znázornené 2 kľúčové komponenty: *Natural Language Understanding* (NLU) a *Dialogue Management*, vďaka ktorým je možné škálovať modulárny prístup vývoja.



Obr. 33: Škálovateľná architektúra RASA (Rasa, 2022)

Zvolené integrované vývojové prostredie, ktoré poskytlo dostatočné komplexné vybavenie, je *PyCharm*⁶ od vývojárskej spoločnosti *JetBrains*. Využitá konfigurácia pre túto prácu je Rasa 3.3.0 a Python 3.8.12. Celý projekt bol spúšťaný a jeho knižnice a balíčky inštalované vo virtuálnom prostredí *Pipenv* a zoznam požiadaviek pre prostredie je uvedený v prílohe A.

5.2 Konfigurácia modelu

Vypis z konfiguračného súboru **config.yml** v prílohe B popisuje nastavenie *pipeline*, vďaka ktorej agent dokáže porozumieť a predpokladať používateľský *intent* (zámer). Poradie komponentov je dôležité, nakoľko každý komponent má požadované vstupy z výstupu predchodcu. Konkrétne *policy* kľúče pomáhajú modelu predikovať ďalšiu akciu agenta.

⁶<https://www.jetbrains.com/pycharm/>

Trénovanie NLU je v tejto práci ovplyvnené dostupnými⁷ a vlastnými komponentami:

- *Tokenizer*: z dostupných komponentov bol použitý *WhitespaceTokenizer*, ktorý generuje tokeny pre každú sekvenciu znakov oddelených medzerami.
- *Featurizer*: v tejto práci využívajú *RegexFeaturizer* a *LexicalSyntacticFeaturizer* a *CountVectorsFeaturizer*.
- *Intent classifier*: využíva *DIETClassifier*, ktorý klasifikuje zámer používateľa z predpovedaných možností a extrahuje entity.
- *components.EmotionalStrategyClassifier*: vlastný komponent postavený na ESC frameworku (Liu et al., 2021) trénuje a následne načíta model založený na Blenderbot-small modely, ktorý každý používateľský vstup (s chit-chat zámerom) klasifikuje z pohľadu 7 rôznych stratégií k zmierneniu emocionálneho stavu.
- *components.RelationsClassifier*: vlastný komponent postavený na COMET frameworku trénuje a následne načíta model založený na znalostnom grafe ATOMIC (Hwang et al., 2021) s jazykovým modelom BART. Z používateľskej správy generuje 3 *commonsense* vzťahy z kategórie sociálno-interakčných vzťahov: vzťah duševného stavu **xReact** ako emocionálnu reakciu zo strany X, behaviorálny vzťah **xWant** so zameraním na sociálno-relevantné reakcie na udalosť, a posledný vzťah **xAttr** opisuje osobnosť alebo atribút X tak, ako ho vnímajú iní pri danej udalosti. Výstupom je pole týchto vzťahov, ktoré sa spracovávajú pri generovaní odpovede.
- *ResponseSelector*: definovaný pre celú doménu FAQ a špecifických *chit-chat* tém predpovedá odpoveď agenta zo súboru odpovedí kandidátov.
- *FallbackClassifier*: klasifikuje *nlu fallback intent* používateľa v prípade, že predchádzajúci klasifikátor intentov nedokázal klasifikovať zámer s istotou väčšou alebo rovnou zvolenému prahu.

Rozhodovanie o tom, akú akciu agent pre svoju odpoveď zvolí, ovplyvňujú konkrétne *policies*. Tie definujú, ktorú akciu je potrebné zvoliť tak, aby konverzácia pokračovala. Zvolené *policies* sú kombináciou *machine-learning based* (napr. *TEDPolicy* a *MemoizationPolicy*) a *rule-based* (napr. *RulePolicy*). *TEDPolicy* rozpoznáva entity a predpovedá ďalšiu najlepšiu akciu (python funkcie definované v súbore **actions.py**). *MemoizationPolicy* sa snaží nájsť zhodu príbehu opísanú v tréningových dátach a ak sa nájde, predpovedá ďalšiu akciu z definovaného príbehu (story). Ten je esenciálny pre vetvenie konverzácie na základe detegovaného emočného stavu a zvolenej stratégie. Všetky 3 spomínané *policies* sa navzájom dopĺňajú a každá z nich umožňuje prispôbiť sa danej situácii.

⁷<https://rasa.com/docs/rasa/components/>

5.3 Tréningové dáta

Doménové tréningové dáta pre modul FAQ tvoria väčšinou bázu zámerov, nakoľko ku každej inštancii otázok je ručne vytvorených 25 príkladov na polozenie danej otázky. Vo všeobecnosti sú zábery rozdelené na pozdrav, súhlas, nesúhlas, nečitateľné, mimo kompetencie chatbota, a empatické pre zabezpečenie *open-domain* konverzácie. Pre túto prácu bol použitý *ResponseSelector*, slovník s kľúčmi zámerov, aby sa predikcia modelu zúžila na priame predpovedanie zo súboru kandidátskych odpovedí. Výpis tréningových dát je v súboroch **nlu.yml** a ich kľúče v **domain.yml**. V Tab. 2 je prehľad možných kategórií používateľských vstupov a počet tréningových dát v projekte.

Tab. 2: Prehľad definovaných zámerov

Intent	Popis	Počet dát
faq	Kategorizuje jednu z 86 rôznych faq príkladov	2 150
empathetic	Kategorizuje 32 možných chit-chat vstupov s predikciou nosnej emócie	64 000
greet	Kategorizuje začiatok konverzácie	20
goodbye	Kategorizuje ukončenie konverzácie	10
affirm	Kategorizuje súhlas s ponukou agenta (primárne pre feedback formulár)	20
deny	Kategorizuje nesúhlas s ponukou agenta (primárne pre feedback formulár)	20
bot challenge	Kategorizuje sponchybňovanie používateľa o jeho neľudskosti	10
out of scope	Kategorizuje pokusy používateľa nespádajúce do kompetencií agenta	20

Každý zámer (intent) používateľa je nasledovaný reakciou agenta (utterance) z určitej množiny definovaných odpovedí v súbore **domain.yml**, alebo je odpoveď generovaná funkciou zo súboru **actions.py**. Zoznam akcií, ktorými *policies* disponujú pri predikcii odpovede a počet variantných textových odpovedí je uvedený v Tab. 3.

5.4 Problematika všeobecnej konverzácie

Open-domain system, ako spomínaná navrhovaná vlastnosť agenta, referuje na schopnosť pochopiť akúkoľvek tému a vrátiť relevantnú odpoveď. Taká konverzácia môže pokračovať viacerými smermi. Od "I have not passed my exam..." až po "My brother set me up! I am so furious!". Navrhovaný systém agenta je heterogénny, nakoľko má preddefinované odpovede v doméne FAQ, preddefinované šablóny odpovedí s maskami pre COMET vstupy v chit-chat priestore, a nakoniec rýdzo generované odpovede pre čistý *chit-chat* dialóg (platí pre konkrétnu **strategy**). Toto je priestor, kedy existujú pre agenta dva hlavné smery ako generovať odpoveď, a to spomínaný *retrieval-based/closed-domain* a *generative-based/open-domain* systém.

Tab. 3: Prehľad akcií pre reakciu agenta

Action	Popis	Počet
utter did faq help	Úvodná fráza feedback formuláru po zodpovedaní otázky	4
utter happy faq	Reakcia agenta na potvrdenie spokojnosti	2
utter bot challenge	Objasnenie humánnosti agenta	2
utter sad faq	Reakcia agenta na nespokojnosť s odpovedanou otázkou	2
utter happy faq	Reakcia agenta na spokojnosť s odpovedanou otázkou	2
utter sad faq rephrase	Výzva agenta k preformulácii otázky	2
utter goodbye	Rozlúčenie s používateľom	4
utter out of scope	Reakcia na neschopnosť zodpovedania zámeru používateľa z dôvodu nekompetencie agenta	2
utter ask rephrase	Výzva na zopakovanie zámeru	2
utter welcome user	Pozdravenie používateľa	3
utter eased stress	Rozlúčenie s používateľom po aplikovanej policy	2
utter {empathetic} policy	Aplikovanie policy podľa vstupnej emócie	2
action read faq	Vracia <i>questionAnswer</i> na základe <i>questionId</i>	3
action offer other faq	Ponuka na zobrazenie ďalších možných otázok z pôvodnej kategórie otázky	1
action show other faqs	Výber vo forme tlačidiel na ďalšie možné otázky z pôvodnej kategórie otázky	1
action default fallback	Vyjadrenie ľútosti nad nečitateľným zámerom používateľa (potenciálny human hand-off)	2
action {strategy}	Generovaná odpoveď na základe stratégie s integráciou viacerých entít	3

5.4.1 Closed-domain riešenie

Keď je rad na agentovi aby odpovedal, siahne model do množiny dostupných definovaných odpovedí a náhodne vyberie jedného kandidáta. *Retrieval-based* systémy môžu pre odpoveď použiť iba text, ktorý majú k dispozícii, a nemôžu generovať novú odpoveď. Keďže v tomto prípade iterujeme nad dopredu známou množinou zámerov a odpovedí, nie je bežné vo FAQ module používať vyhodnotenú stratégiu k zmierneniu emocionálneho stavu. Výnimky, pri ktorých je relevantné vziať do úvahy sentiment FAQ sú uvedené v Diskusii tejto práce. V sekcii 5.2 je zdôraznená závislosť na poradí komponentov v konfiguračnom súbore. Ak *DIETClassifier* klasifikuje vstup ako *faq*, *components.EmotionalStrategyClassifier* ani *components.RelationsClassifier* nie sú potrebné.

Nepopierateľnou výhodou tejto architektúry je pozornosť iba na konkrétne témy (o ktorých má model dopredné vedomosti). Vyhodnotený vstup s *confidence* pod

hodnotu *threshold* 0.3, je podľa *FallbackPolicy* označený ako *nlu fallback intent* a odpoveď agenta generuje vlastná *action* s textom "I am sorry, there is nothing I can do". V opačnom prípade je k zámeru *faq/question101* nájdená odpoveď z tabuľky. Pre všetky výskyty 86 otázok boli tréningové dáta zostavené ručne s odporúčaným počtom príkladov cca 25 pre každé potenciálne podanie zámeru. V prípade *retrieval* zámeru **faq/question101** je šablóna variant k jeho podaniu v Tab. 4.

Tab. 4: Príklad tréningových dát pre otázku "What is UIS?"

No.	Príklad
1	What is UIS?
2	UIS?
3	Tell me what is the University information system
4	Information system of university
5	I am looking for the university system
6	What does uis stand for?
7	I want UIS information
8	Do you know what is university information system?
9	Where is UIS?
10	Where is university information system?
11	I have heard little about UIS
12	I have heard none of UIS
13	My friends and I want to know about university information system
14	I have no clue about UIS
15	University information system is a secret to me
16	Do you know something about school system?
17	What is my school information system?
18	Where is my school's system located?
19	Is my university system located on web?
20	What is the website of UIS?
21	I wish I knew what is UIS
22	I have never heard of UIS
23	University information system
24	Do you know where is UIS?
25	I want to know what is UIS

5.4.2 Open-domain riešenie

Táto časť architektúry agenta generuje nové odpovede v závislosti od danej histórie konverzácie – kontextu. Zvládne bežné a nepredvídateľné otázky, čo zvyšuje zložitosť implementácie vrátane jej časovej náročnosti. Aby RASA prostredníctvom triedy *CollectingDispatcher* a jeho metódy *utter message* vygenerovala odpoveď naspäť

k používateľovi, musí rozpoznať konkrétnu *story* v ktorej je daný *intent* zahrnutý. Príprava tréningových dát pre všeobecnú chat konverzáciu je z toho dôvodu náročná, takmer nemožná, úloha. Bez integrovania *retrieval* zámerov *empathetic intents* podľa Tab. 2, používateľský vstup "I have not passed my exam..." by bol klasifikovaný ako *nlu fallback*. Spôsob zachovania dialógu a predchádzania *fallback* stavu je nasledovný:

- Manuálnou analýzou datasetu *EmpatheticDialogues* (ED) obsahujúcich 25 000 empatických konverzácií, spracovali (Rashkin et al., 2019)⁸ dataset do rôznych csv súborov na základe 32 emócií, ktorými sú dialógy podmienené.
- Všetky súbory boli mnou vyfiltrované na výskyty "Speaker" a prevedené do NLU formátu RASA.
- Generovanie odpovedí, ktoré poslucháč (chatbot) používa na vyjadrenie svojej empatie voči rečníkovi, táto práca abstrahuje nakoľko tento *generation-based* prístup by spôsobil utterance nad ktorou neexistuje v dialógovom manažmente kontrola.
- Vzniknuté NLU tréningové dáta v projekte boli označené *retrieval* zámerom **empathetic**, kde jeden zámer má približne 2 000 tréningových dát. Pre zjednotenie tréningovania boli tieto dáta v určitých fázach experimentu skrátené na 200 vzoriek bez ovplyvnenia hodnoty *precision*.

Výsledkom tohto postupu je zozbieranie vzácných dát, ktoré umožnili konverzáčnemu modelu klasifikovať zámer takmer na akúkoľvek emocionálnu situáciu. Pri vstupe "Someone shot the door when we were taking an exam!" sa tak klasifikovanie výroku pohybuje vďaka *DIETClassifier* v slovníku **empathetic** medzi týmito kategóriami emócie rečníkovho výroku: *surprised, excited, angry, proud, sad, annoyed, grateful, lonely, afraid, terrified, guilty, impressed, disgusted, hopeful, confident, furious, anxious, anticipating, joyful, nostalgic, disappointed, prepared, jealous, content, devastated, embarrassed, caring, sentimental, trusting, ashamed, apprehensive* a *faithful*. Pravdepodobnosť predošlého vysokého klasifikovania *nlu fallback* sa radikálne znížila na minimum. Ukážku tréningových dát pre zámer emócie *empathetic/surprised* uvádza Tab. 6.

5.5 Usporiadanie dát do príbehov

S klasifikáciou zámerov a generovanými *entities* od *RelationsClassifier* a *Emotional-StrategyClassifier*, potrebuje agent vedieť tok konverzácie aby vedel predpovedať správnu akciu za behu kontextu. *Stories*, ktoré takýto tok definujú, predstavujú východiskový bod a konečný cieľ jednej ucelenej výmeny zámeru a odpovede. Sú najväčším súborom pravidiel a skladajú sa z menších príbehov. Pri každom tréningovaní príkazom `rasa train`, používajú vytvorené modely príbehy na zovšeobecnenie

⁸<https://github.com/anuradha1992/EmpatheticIntents>

Tab. 5: Príklad tréningových dát pre podanie výroku v emócií "surprised"

No.	Príklad
1	i really like the feeling of the unknown on christmas morning
2	Just came home to so much cake
3	I was shocked when I wsa told my boss was let go
4	I can't believe Matt Ryan got 30 million.
5	went to the bathroom and saw something you don't see everyday.
6	Whoa so that was crazy. You'll never guess what just happened.
7	I recently saved money by switching to a different insurance company.
8	My dog found a cat in our backyard and has been nursing it.
9	Tv has gotten so good in the past few years
10	I haven't slept in 3 days, yet don't feel tired.
11	I knew I was smart but I had no clue I was that smart.
12	i didnt think i could jump as high as I did when I stepped on that snake
13	I got into Harvard. Can you believe it?
14	The other day I found out that my sister is having twins
15	I got lots of cool stuff and I had no idea they were planning it
16	I found 6 puppies on my doorstep.
17	I got the wrong part for my mower.
18	My daughter drove into town and paid us a visit
19	My Mother in Law offered to buy us almost all the furniture.
20	first time making a cake. it turned out pretty well.
21	I was sitting here playing video games and heard something loud
22	I was given a huge discount on my groceries today.
23	A monkey jumped on my back and stole my banana
24	No, didn't really get a chance. It happened too fast.
25	I woke up and my car had a flat tire

a predpovedanie cesty doposiaľ neviditeľnej konverzácie na základe príbehov, na ktorých boli vyškolené.

Implementácia formulárov so sekvenčným vyžadovaním rôznych vstupov od používateľa (forms) bola pre architektúru agenta zamietnutá, nakoľko pre povahu cieľa práce nie je kritické preukazovať schopnosť agenta pracovať s komplexnou objednávkou, či iným formulárovým prístupom. Tento krok umožňuje voľný dialóg, ktorý môže s akýmkoľvek používateľským vstupom začínať u akejkoľvek stratégie. Pre ukážku všetkých 7 stratégií, ktoré majú rovnaký scenár, je na Obr. 34 vizualizácia rôznych konverzácií so všetkými definovanými zámermi (intents).

Nasledujúca ukážka príbehu v doméne FAQ definuje situáciu, ak odpoveď na otázku bola uspokojujúca, a zároveň používateľ súhlasí so zobrazením pridružených otázok na základe pôvodnej kategórie otázky. V tomto príbehu neexistuje korelácia medzi neutrálne podanou otázkou a tou so silným emocionálnym zámerom. Scenár nižšie teda rovnako predikuje správanie agenta pri vstupe "What is uis" a "I

have spent three hours looking for what the UIS is. Help!!!”.

```

1 # Multi-turn FAQ interactions
2 stories:
3 - story: faq positive show story
4   steps:
5     - intent: faq
6     - action: action_set_faq_slot
7     - action: action_read_faq_answer
8     - action: utter_did_faq_help
9     - intent: affirm
10    - action: utter_happy_faq
11    - action: action_offer_other_faq
12    - intent: affirm
13    - action: action_show_other_faqs

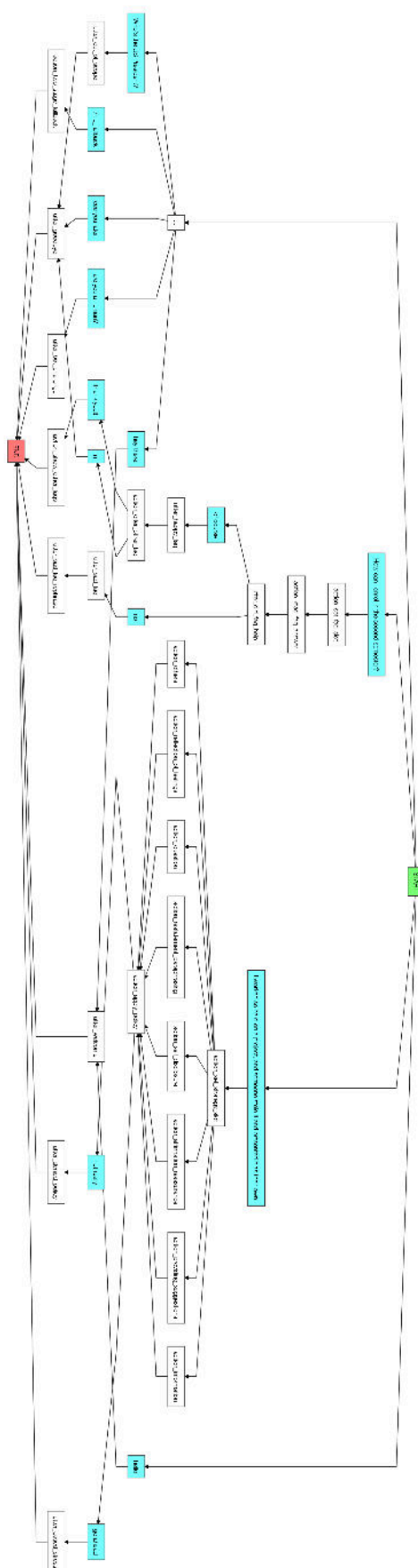
```

V prípade, keď prevládajúca *confidence* predikovaného zámeru nie je v prospech FAQ, je potreba u každého *chit-chat* vstupu aplikovať jednu zo siedmich definovaných stratégií vzhľadom na vyhodnotenú entitu **strategy**. V prípade stratégie *Affirmation and Reassurance* je reakcia agenta podmienená aplikovaním špecifickej textovej odpovede – empathetic policy. Agent následne použije konkrétnu *policy* ako spôsob na zmiernenie stresu používateľa (každá *policy* sa líši v závislosti od empatického zámeru). Príkladom mnou vytvorenej laickej *empathetic policy* môže byť URL na hudobný playlist alebo navrhnutie fyzického cvičenia.

```

1 # Single-turn empathetic interaction
2 - story: empathetic policy
3   steps:
4     - intent: empathetic # do not disclose the exact emotion suffix
5       entities:
6         - strategy: Affirmation and Reassurance
7         - action: action_set_strategy_slot
8         - slot_was_set:
9           - strategy: affirmation_reassurance
10        - action: action_affirmation_reassurance
11        - checkpoint: checkpoint_apply_policy
12 - story: checkpoint apply policy affirm
13   steps:
14     - checkpoint: checkpoint_apply_policy
15     - action: action_apply_policy
16     - intent: affirm # response for policy
17     - action: utter_eased_stress # goodbye
18 - story: checkpoint apply policy deny
19   steps:
20     - checkpoint: checkpoint_apply_policy
21     - action: action_apply_policy
22     - intent: deny # response for policy
23     - action: utter_denied_policy # goodbye

```



Obr. 34: Graf príbehov v RASA (Zdroj: autorka práce)

6 Testovanie konverzačného modelu

Povaha tejto práce vyžaduje dôkladnú validáciu, testovanie a evaluáciu NLU dát (všetky tréningové dáta ktorými pripravujeme agenta na vstupy v reálnom prostredí) a *end-to-end* konverzácie (príbehy, pravidlá a doménový súbor). Vďaka tomu je možné predísť nekonzistenciám či odchyteniu pretrénovaného neurónového modelu. Samotné tréningovanie oboch častí súbežne – *NLU* a *Dialogue management* trvalo v priemere 2 až 3 hodiny na CPU, na grafickej karte o takmer polovicu času menej. Aj z toho dôvodu je nevyhnutné celkové testovanie modelu, aby sa pravdepodobnosť nesprávne natrénovaných modelov zminimalizovala.

Predtým ako sa evaluovali obe spomínané časti architektúry, bolo prostredníctvom príkazu `rasa data validate` overené, že sa v doménovom a príbehovom súbore nevyskytujú žiadne chyby alebo kritické nekonzistencie.

6.1 Evaluácia dialógového modelu

RASA umožňuje overiť a otestovať dialógy prejdením testovacích príbehov. Nasledujúci testovací príbeh otestovaný na natrénovanom modeli je najlepším spôsobom ako zvýšiť *confidence* v to, že bude agent reagovať na *chit-chat* konverzáciu so vstupnou emóciou *afraid* a predikovanou stratégiou na *Reflection of feelings* v reálnom prostredí správne. Pomocou `rasa test core --stories test_stories.yml --out results` bolo otestované, že scenár je reprezentantom skutočného rozloženia skutočných konverzácií. Nasledovná ukážka demonštruje základný priebeh *single-turn* dialógu vo vstupnou emóciou *afraid*:

```
1 version: "3.1"
2 # Single-turn empathetic interactions
3 stories:
4   - story: empathetic policy
5     steps:
6       - user: |
7         I am so worried about my job
8         intent: empathetic/afraid
9         entities:
10        - strategy: Reflection of feelings
11        - action: action_set_strategy_slot
12        - slot_was_set:
13          - strategy: reflection_of_feelings
14        - action: action_reflection_of_feelings
15        - action: action_apply_policy
16        - user: |
17          indeed
18          intent: affirm
19        - action: utter_eased_stress
```

Spôľahivosť dialógového manažmentu pri výbere správnej akcie v reálnom prostredí som podnietila rôznymi prerušovanými stavmi na strane používateľa, tzv.

User Interjections. Ukážkou takého prerušenia môže byť plynulý multi-turn dialóg vo všeobecnej konverzácii:

```

1 version: "3.1"
2 # Multi-turn empathetic interactions
3 - story: empathetic policy interjected emotion
4   steps:
5     - user: |
6       I dropped my laptop at canteen and broke it.
7       intent: empathetic/angry
8       entities:
9         - strategy: Question
10      - action: action_set_strategy_slot
11      - slot_was_set:
12        - strategy: question
13      - action: action_question
14      - action: action_apply_policy
15      - user: |
16        What is more, our school has been closed due pandemic without
17        no prior warning!
18        intent: empathetic/angry
19        entities:
20          - strategy: Self-disclosure
21          - action: action_set_strategy_slot
22          - slot_was_set:
23            - strategy: self_disclosure
24          - action: action_self_disclosure
25          - action: action_apply_policy
26          - user: |
27            I will definitely do that
28            intent: affirm
29            - action: utter_eased_stress

```

Obdobne boli pripravené testy na prípad ak klasifikovanie faq zámeru je natoľko vysoké, že sa stane súčasťou *multi-turn flow* vo všeobecne empatickej konverzácii:

```

1 version: "3.1"
2 # Multi-turn empathetic interactions
3 - story: empathetic policy interjected by FAQ
4   steps:
5     - user: |
6       Someone took my phone when I was at toilet. It was during the
7       class and I have no clue who did it.
8       intent: empathetic/surprised
9       entities:
10        - strategy: Restatement or Paraphrasing
11        - action: action_set_strategy_slot
12        - slot_was_set:
13          - strategy: restatement_paraphrasing
14        - action: action_restatement_paraphrasing
15        - action: action_apply_policy
16        - user: |

```



```
16         And what is more, I have spent three hours looking for my UIS
17         credentials.
18         intent: faq/question103
19         - action: action_set_faq_slot
20         - slot_was_set:
21           - faq: [ "Get in touch with the systems integrator of your
22             faculty. You will find the contact details here: https://is.mendelu
23             .cz/dok/integratori.pl?lang=en", "UIS" ]
24           - action: action_read_faq_answer
25           - action: utter_did_faq_help
26           - user: |
27             not really
28             intent: deny
29           - action: utter_sad_faq
30           - action: utter_sad_faq_rephrase
```

Výsledné evaluačné výsledky na úrovni predpokladaných akcií v konverzácii boli nasledovné:

- Počet správne predpovedaných akcií: 148/148
- F1-Score: 1.000
- Precision: 1.000
- Accuracy: 1.000

Vyššie dosiahnuté výsledky tak potvrdzujú správne fungujúci dialógový model, ktorý vizualizuje *confusion matrix* na Obr. 42 v prílohe D.

6.2 Evaluácia NLU modelu

Keďže implementácia všeobecnej konverzácie je v tejto práci podmienená súborom tréningových dát zámerov o 32 rôznych emóciách, je nutné evaluovať ako by agent v reálnom svete spracoval správy, ktoré nevidel v tréningových dátach. K takej simulácii bola pomocou `rasa data split nlu` odložená časť údajov na testovanie, ktoré sa vykonalo pomocou `rasa test nlu -u /train_test_split/test_data`.

6.2.1 Výsledky klasifikátoru zámerov

Testovací skript vo vygenerovanom reporte včas odhalil, že zámer **empathetic** bol v jednom prípade zamenený so zámerom **greet**. Táto chyba nastala kvôli príkladom v *Empathetic datasets* (Rashkin et al., 2019), ktoré pre túto prácu boli spracované a normalizované. Niektoré robustné súbory emócií mali v sebe príklady reakcií používateľa v kontexte celej konverzácie (napr.: "Yes, he is 3 or 4 years old" alebo "Thank you!!!"). Ukážka nižšie teda odhalila, že v súbore `nlu/empathetic/joyful.yml` sa vyskytovala kontextová reakcia "Hello there. How are you feeling today?" ako odpoveď používateľa na pozdravenie asistenta. Kvôli tomu boli následne tréningové

dáta opätovne normalizované tak, aby neobsahovali zameniteľné reakcie so zámerom na pozdravenie, súhlas alebo nesúhlas.

```

1 intent_errors: [
2   {
3     "text": "Hello there. How are you feeling today?",
4     "intent": "empathetic",
5     "intent_prediction": {
6       "name": "greet",
7       "confidence": 0.7086317539215088
8     }
9   }
10 ]

```

```

1 "empathetic": {
2   "precision": 0.0,
3   "recall": 0.0,
4   "f1-score": 0.0,
5   "support": 1,
6   "confused_with": {
7     "greet": 1
8   }
9 },

```

Po normalizácii tréningových dát bola pri evaluácii odhalená ďalšia chyba, ktorá reflektuje relevantnú zameniteľnosť domény FAQ so všeobecnou *open-domain* konverzáciou. Nasledujúci výrok z tréningových dát emócie **ashamed** bol zamenený so zámerom **faq/question123** – "Where can I borrow books or study materials?". To bol podnet k zvýšeniu tréningových dát v doméne FAQ a opätovnej normalizácii empatických tréningových dát. Keďže výrok "About why I needed to borrow money." je kontextovou prerušenou výpoveďou v roli *speaker*, odstránením neúplnosti celého výroku bola dosiahnutá nasledujúca podoba tréningového výroku: "I can not believe I lied to my mother About why I needed to borrow money." To výrazne pomohlo odstráneniu popisovanej chyby v evaluácii.

```

1 [
2   {
3     "text": "About why I needed to borrow money.",
4     "intent": "empathetic",
5     "intent_prediction": {
6       "name": "faq",
7       "confidence": 0.753279447555542
8     }
9   }
10 ]

```

```

1 "empathetic": {
2   "precision": 1.0,
3   "recall": 0.9991379310344828,
4   "f1-score": 0.9995687796463993,
5   "support": 1160,

```

```
6   "confused_with": {
7     "faq": 1
8   }
9 },
```

Na Obr. 43 v prílohe E je vyhodnotená *confusion matrix*, ktorá potvrdzuje, že predikcia klasifikácie zámeru funguje dobre a pri klasifikácii zámeru došlo len k jednej vyššie opísanej chybe.

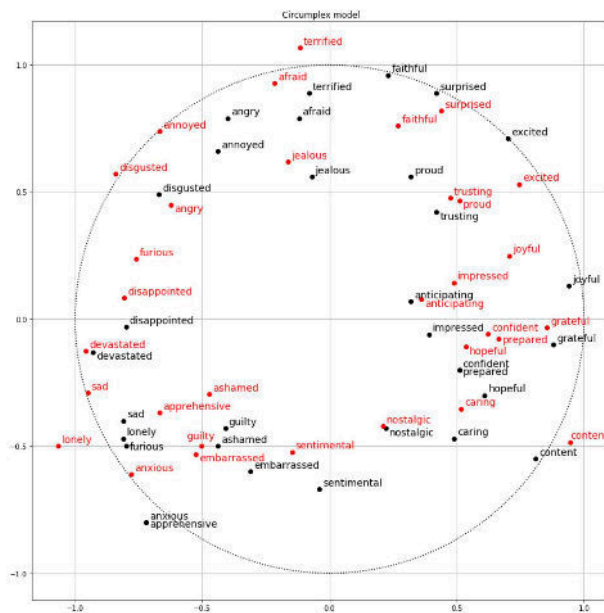
6.2.2 Výsledky retrieval zámerov

Kým klasifikátor obecných zámerov dosiahol hodnotou *accuracy* 0.9992 (percento testovacích viet, ktoré sa správne zhodujú so základným zámerom), klasifikovanie konkrétnych zámerov zo skupiny **faq** alebo **empathetic**, dosiahol hodnotu o desatinu nižšiu, a to 0.9888. Ukážky nesprávne klasifikovaných zámerov nižšie demonštrujú podobnosť rôznych emócií, ktoré sú pre natrénovaný model v niektorých prípadoch ťažko rozlíšiteľné, s najnižšie dosahovanou confidence hodnotou 0.3982 (surprised-dissatisfied) pri testovacom výroku: "The other day I won 100 dollar on a scratch off lottery ticket and I couldn't believe my luck".

```
1   {
2     "text": "I walked up on a snake in my yard.",
3     "intent_response_key_target": "empathetic/terrified",
4     "intent_response_key_prediction": {
5       "name": "empathetic/surprised",
6       "confidence": 0.5535579323768616
7     }
8   },
9
10  {
11   "text": "Someone stole my car recently. I am so upset about it!",
12   "intent_response_key_target": "empathetic/furious",
13   "intent_response_key_prediction": {
14     "name": "empathetic/angry",
15     "confidence": 0.5476170778274536
16   }
17 },
18
19 {
20   "text": "The other day I won $100 on a scratch off lottery ticket
21   and I couldn't believe my luck",
22   "intent_response_key_target": "empathetic/surprised",
23   "intent_response_key_prediction": {
24     "name": "empathetic/disappointed",
25     "confidence": 0.3981646001338959
26   }
27 }
```

Vizualizácia kompletnej confusion matrix pre je na Obr. 42 v prílohe C. Vlastná voľná interpretácia chýb neurčitosti rozpoznania surprised-terrified alebo furious-angry zámeru je podložená 2D emočným modelom 2 dimenzií (Russell, 1980), kedy

sa oba páry emócií nachádzajú ako pár v rovnakej dimenzii hodnôt *arousal* i *valence*, ktoré popisujú ako je emócia príjemná/neprijemná a aktívna/neaktívna. Obdobnú anomáliu zistil aj výskum (Xie a Park, 2021), kde pozície spomínaných emócií furious-angry zostali príliš ďaleko od anotovaných pozícií z tréningového datasetu. Vysvetľujú, že emócia *angry* má podobné charakteristiky ako furious, disgusted a to spôsobuje nesprávnu klasifikáciu (Obr. 35). Z toho dôvodu nebola potreba pretrénovania pre môj model tak alarmujúca.

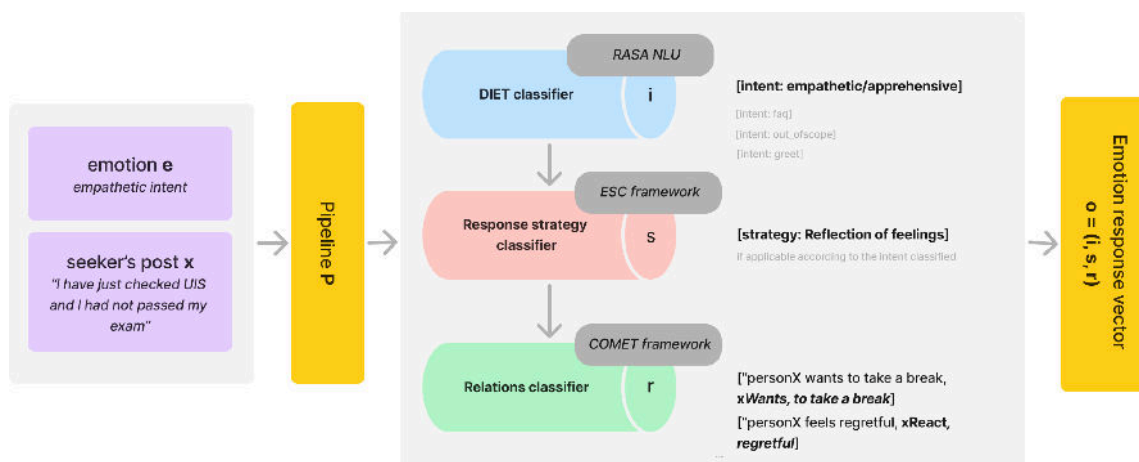


Obr. 35: 2D model s predikovanými a anotovanými pozíciami viacerých emócií (Xie a Park, 2021)

Z množiny všetkých vygenerovaných testovacích výsledkov nasleduje Obr. 44 v prílohe E, kde je vidieť počet uchovávaných vzoriek a ich spoľahlivosť pre všetky odpovede, pričom správne a nesprávne predpovede sú zobrazené zelenými a červenými pruhmi. Dôvod presahu nespoľahlivosti pre odpovede (červené pruhy) spočíva vo vyššie uvedenej anomálii podobnosti emócií.

7 Workflow end-to-end konverzácií

Vyhodnotené *entities* (generovaná hodnota z každého vstupu), *intents ranking* (určitosť zámerov) a *responses* (určitosť zámerov z danej kategórie) nepochybne ovplyvňujú scenár a dialógový vývoj. Na Obr. 36 je súbor komponent a ich výstupy, ktoré sú povahovo dôležité pre vysvetlenie návrhu agenta. Na vizualizácii abstrahujem od typu komponent *Tokenizers* alebo *Featurizers*, pretože pre túto prácu nie sú vo vysvetlení architektúry kritické. Výsledky entít alebo zámerov ďalej spracúvajú zvolené *polícies* (predpovedajú ďalšiu akciu s určitou úrovňou spoľahlivosti).

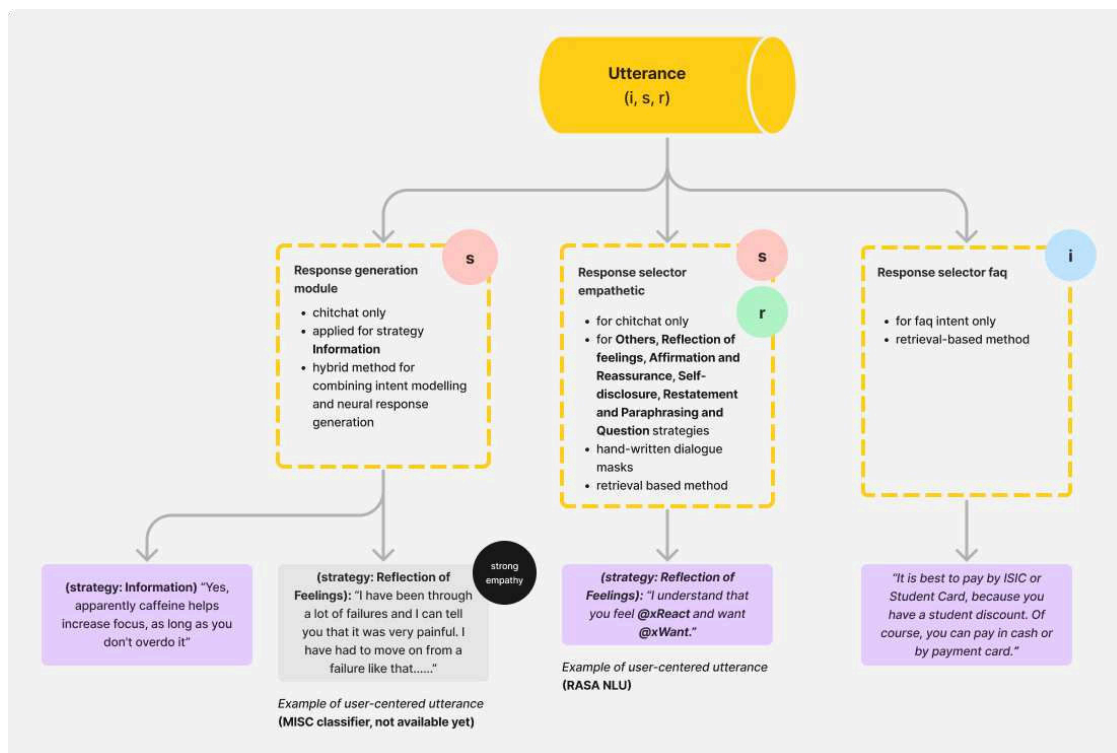


Obr. 36: Interpretácia spracovaného emočného vstupu a prípravy parametrov na generovanie odpovede pre vstup "I have just checked UIS and I had not passed my exam" (Zdroj: autorka práce)

Získaná hodnota entity **strategy** ovplyvňuje konverzáciu len v prípade zámeru **empathetic**, nakoľko zámer **faq** primárne reaguje iba v rámci svojej domény. Táto práca plne nevyužíva *generation-based* prístup, pretože RASA stráca kontrolu nad správou používateľa na takú reakciu, nakoľko *engine* na generovanie textu, ktorý nie je v tréningových dátach, môže vyzývať používateľa ku krátkej odpovedi ktorej kontext je pre RASA neznámy. Výnimkou je stratégia *Information*, u ktorej generovaný text má iba informatívny charakter. *Emotional response vector* sa spracováva podľa Obr. 37 troma rôznymi mechanizmami reakcií:

- *Response selector faq*: platí pre najvyššie klasifikovaný *i* (zámer) **faq**. Odpoveď používateľovi sa prehľadáva v tabuľke najčastejších odpovedí. Ak sa medzi najvyššími 5 vyhodnotenými zámermi nachádza aj **empathetic**, použije agent v šablóne odpovedi hodnoty *r*.
- *Response generation module*: platí pre najvyššie klasifikovaný *i* (zámer) **empathetic** s extrahovanou *s*="Information" (stratégia). Tento experimentálny prístup je v projekte použitý ako *proof-of-concept*.

- *Response selector empathetic*: platí pre najvyššie klasifikovaný *i* (zámer) **empathetic**. Odpoveď závisí od extrahovanej *s* (stratégia) a jej dedikovanej šablóny (Tab. 6 sumarizuje prehľad akcií, ktoré definujú text odpovedi pre ich jedinečnú stratégiu s pomocou získaných *entities* z frameworku COMET).



Obr. 37: Interpretácia troch mechanizmov generovania odpovede (Zdroj: autorka práce)

Aplikácia empatie v prípade *task-oriented* používateľského zámeru využíva dostupné prostriedky vlastných komponent (xReact a xWant). Na demonštráciu emočnej inteligencie v triviálnej úlohe FAQ bol použitý nasledujúci vstup: **"I have spent three days looking for uis credentials. It takes ages! please help"**.

S takmer úplnou presnosťou klasifikoval DIETClassifier zámer podľa predpokladaného správania, a príslušnú konkrétnu otázku taktiež správne ako: "How shall I get the UIS credentials?":

```

1 "response_selector": {
2   "faq": {
3     "response": {
4       "responses": [
5         {
6           "text": "faq/question102"
7         }
8       ],
9       "confidence": 0.9992315769195557,
10      "intent_response_key": "faq/question102",

```

Tab. 6: Prehľad dostupných stratégií

Názov stratégie a maska odpovedi
(Reflections of feelings) "I understand that you feel {xReact[0]}"
(Reflections of feelings) "It looks like you are {xReact[0]} and want {xWant[0]}."
(Question) "May I ask you why are you feeling {xAttr[0]}?"
(Question) "Tell me more about it please "
(Restatement or Paraphrasing) "I can see how that would make you {xReact[0]} and why you would like to {xWant[0]}"
(Self-disclosure) "I was in the same situation as you and felt {xReact[1]}"
(Self-disclosure) "I feel the same way. I would also be {xAttr[0]} and wanted {xWant[0]}"
(Affirmation or Reassurance) "I understand that you feel {xAttr[0]} and {xAttr[1]}"
*Providing suggestions - "I wish I could help you..."
*Information - uplatňuje sa generovaná odpoveď extraktorom Blenderbot-small (ESConv framework)
*Others - spracovaná stratégia ako fallback

```

11     "utter_action": "utter_faq/question102"
12   },
13   }
14 },
15 "intent_ranking": [
16   {
17     "name": "faq",
18     "confidence": 0.9997625946998596
19   },
20   {
21     "name": "empathetic",
22     "confidence": 0.00021055272372905165
23   },
24   {
25     "name": "out_of_scope",
26     "confidence": 1.7808446500566788e-05
27   },
28   {
29     "name": "affirm",
30     "confidence": 8.0698482634034e-06
31   },
32   {
33     "name": "deny",
34     "confidence": 3.203894323178247e-07
35   }
36 ],

```

Klasifikátor ako druhý (výrazne nepravdepodobný) zámer označil **empathetic**, čo podnecuje príslušnú akciu k použitiu empathickej šablóny v ktorej sú konkrétne hodnoty RelationsClassifier nasledovné:

```

1 "entities": [
2   {
3     "entity": "xReact",
4     "value": [
5       [
6         "frustrated",
7         "relieved",
8         "tired",
9         "sad",
10        "happy"
11      ]
12    ],
13    "extractor": "COMET-ATOMIC 2020 bart"
14  }
15 ],
```

Spracovaním vyššie získaných RASA NLU dát sa Dialogue Management postará o priebeh konverzácie u ktorej bol daný dôraz na vhodnú *user experience*. Na Obr. 38 je ukážka *empathy-driven* dialógu, ktorý je opakom situácie ak používateľ podá svoju otázku neutrálne, konkrétne: "How shall I get the UIS credentials?".

Na Obr. 39 môžeme sledovať niekoľko postupov. Ak v používateľskom vstupe **"I have just checked UIS and I had not passed my exam"** je najvyššie hodnotený empathický zámer, generuje komponenta EmotionalStrategyClassifier svoj predpoklad o vhodnej stratégii:

```

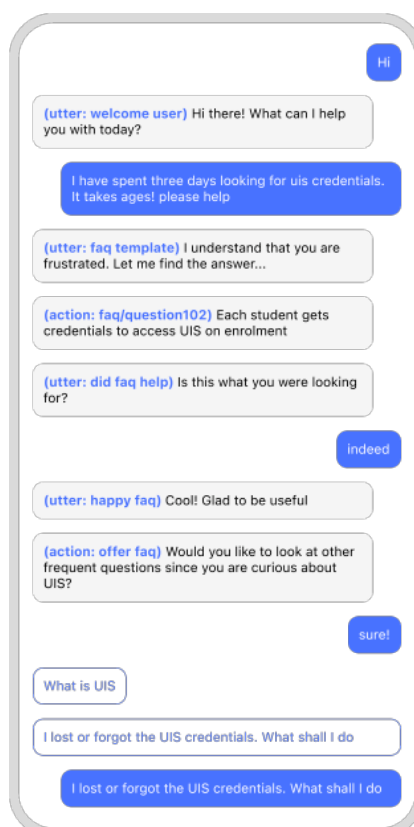
1 "entities": [
2   {
3     "entity": "strategy",
4     "value": "Restatement or Paraphrasing",
5     "extractor": "Blenderbot-small"
6   },
7 ],
```

Ako podporný nástroj emočnej podpory v situácii, ktorú používateľ prežíva, bola vytvorená množina *empathetic policies*, ktorá názvom reflektuje zistený *retrieval* zámer a dopĺňa tak stratégiu v závislosti na klasifikovanej emócii:

```

1 "response_selector": {
2   "faq": {
3     "response": {
4       "responses": [
5         {
6           "text": "empathetic/apprehensive"
7         }
8       ],
9       "confidence": 0.930866539478302,
10      "intent_response_key": "empathetic/apprehensive",
11      "utter_action": "utter_empathetic/apprehensive"

```

Obr. 38: Vizualizácia konverzácie s rozpoznaným emočným faq zámerom (Zdroj: autorka práce)

```

12     },
13     }
14     },

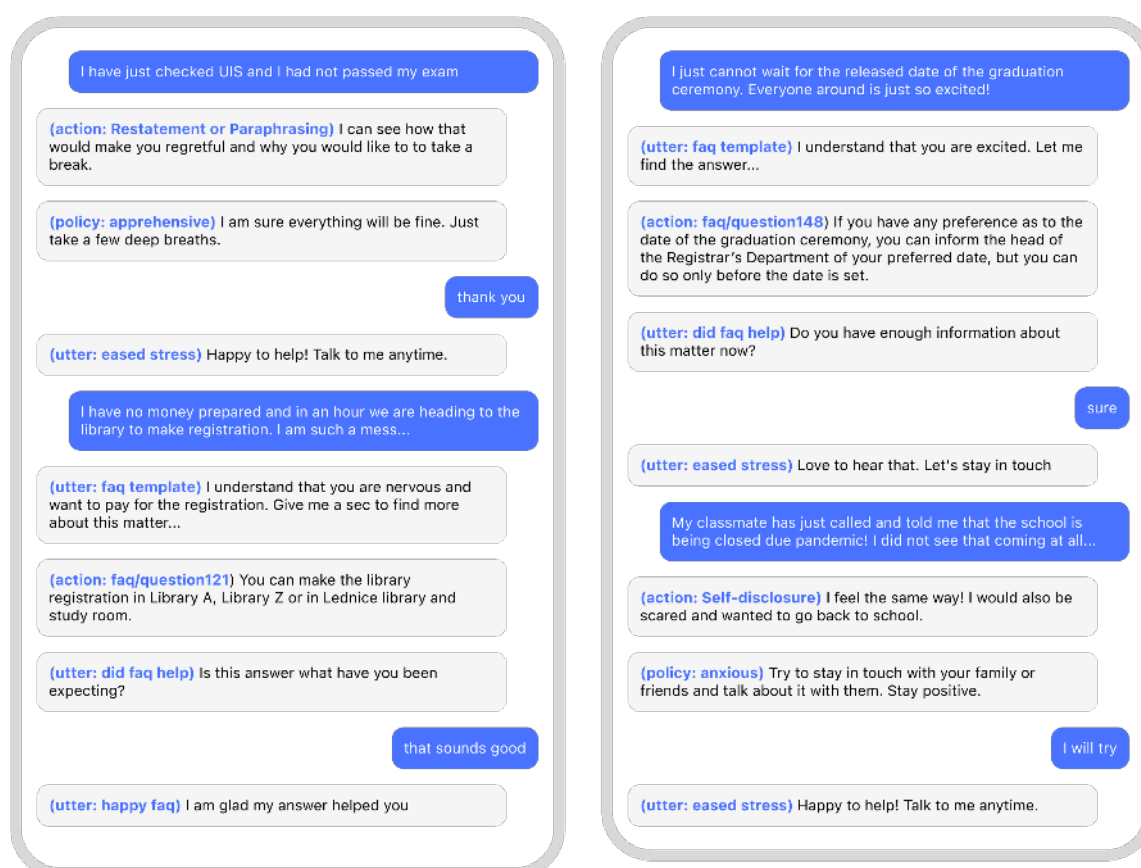
```

Na Obr. 39 je kombinácia niekoľkých mechanizmov odpovedí, kde môžeme v jednoduchosť uvažovať o pravidle:

- Ak prevyšuje **empathetic** zámer, aplikuj odpoveď podľa doporučenej stratégie ako primárnu interpretáciu empatie a vyber zo šablóny inštanciu *empathetic policies* podľa detegovanej emócie (policy: apprehensive) z pôvodného zámeru.
- Ak prevyšuje **faq** zámer, vráť odpoveď prehliadaním tabuľky z dostupných odpovedí na často kladené otázky.
- Ak prevyšuje **faq** zámer a zároveň bol klasifikovaný aj empatický zámer, použi empatickú šablónu pre faq a vráť odpoveď prehliadaním tabuľky z dostupných odpovedí na často kladené otázky.

Jeden z čisto empatických vstupov na Obr. 39: "My classmate has just called and told me that the school is being closed due pandemic! I did not see that coming at all..." spustil stratégiu, ktorá využila hodnotu *xWant*:

```
1  "entities": [  
2    {  
3      "entity": "xReact",  
4      "value": [  
5        [  
6          "scared",  
7          "sad",  
8          "shocked",  
9          "scared",  
10         "upset"  
11        ]  
12      ],  
13      "extractor": "COMET-ATOMIC 2020 bart"  
14    },  
15    {  
16      "entity": "xWant",  
17      "value": [  
18        [  
19          "to go back to school",  
20          "to go to school",  
21          "to find out what happened",  
22          "to go to the doctor",  
23          "to call their parents"  
24        ]  
25      ],  
26      "extractor": "COMET-ATOMIC 2020 bart"  
27    },
```



Obr. 39: Scenár na spracovanie všeobecného vstupu u dvoch rôznych konverzáciách predikuje akciu na aplikovanie stratégie a vhodnej *empathetic policy* ako čiastočný nástroj na emočnú podporu základnej úrovne. Ručne vytvorené interpretácie *empathetic policies* sa povahou približuje k bežnej podpore rodiny alebo priateľov (Zdroj: autorka práce)

8 Používateľské testovanie

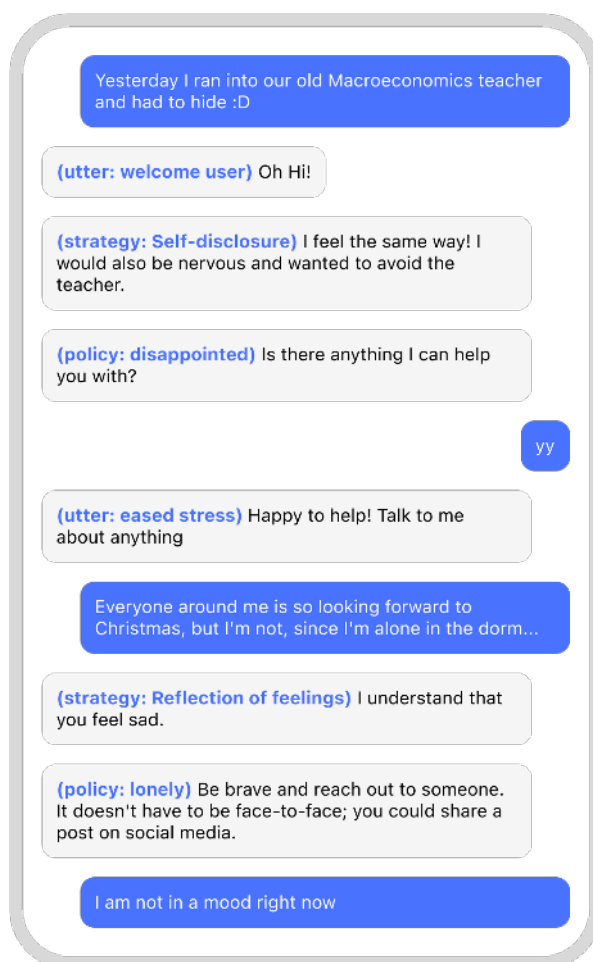
Cieľom používateľského testovania bolo potvrdiť alebo vyvrátiť schopnosť chatbota identifikovať emočný stav a empaticky naňho reagovať. K tomu, aby sa u chatbota minimalizovala jeho neschopnosť správne odhadnutého zámeru či predikovania nesprávnej akcie, bol v kapitole 6 vykonaný hĺbkový test jeho NLU častí a dialógového manažmentu. Používateľské testovanie absolvovalo 5 respondentov v roli aktívne študujúcich vysokoškolákov. Testovanie prebehlo online pomocou vzdialeného zdieľania obrazovky, nakoľko nástroj RASA vo voľne dostupnej verzii *open-source* neposkytuje rozhranie na zdieľanie agenta. Primárnym merateľným cieľom bolo reportovanie miery empatie agenta po bloku vymenených správ a reportovanie emočnej intenzity používateľa na začiatku konverzácie a na konci. Na začiatku konverzácie boli respondenti vyzvaní k vyplneniu nasledujúcich otázok:

- Popíšte problém s akým sa na chatbota obraciate.
- Popíšte jedným slovom emóciu, s ktorou do konverzácie vstupujete.
- Ohodnoťte intenzitu emócie, ktorú prežívate na stupnici 1 (málo aktívna) až 5 (veľmi aktívna).

Na konci konverzácie ma u respondentov zaujímalo, ako by znovu ohodnotili mieru intenzity prežívania emócie a či chatbot poskytol podporu k emočnému prežívaniu situácie na stupnici od 1 (určite nepodporil) až 5 (určite podporil). Testovanie prinieslo nasledujúce kľúčové zistenia:

- Počas testovania nebola zistená obava z etiky či ochrany súkromia.
- Respondenti boli príjemne prekvapení empaticky vedeným dialógom v prípade často kladených otázok.
- Respondenti uviedli, že chatbot relevantne identifikoval emočný stav a empaticky reagoval.
- U prežívaných emócií *joyful*, *proud* a *excited* síce neprišlo k zníženiu intenzity prežívania emócie, ale chatbot dokázal správnu stratégiu emóciu zdieľať a vytvoriť empaticky riadený dialóg.
- Miera intenzity sa priemerne znížila o 3 body.
- Empatia chatbota sa dostalo na hodnotu 4/5.

Obr. 40 ukazuje dialóg z testovania, u ktorého nebola zaznamenaná znížená intenzita prežívania, ale práve najlepšie ohodnotená podpora chatbota k emočnému prežívaniu situácie respondenta.



Obr. 40: Ukážka dialógu z používateľského testovania (Zdroj: autorka práce)

9 Diskusia

V posledných rokoch rastúci počet štúdií potvrdil výhody empatie a začal ju do konverzačných systémov začleňovať (Saha a Ananiadou, 2022). Aktuálny rozsiahly prieskum o empatii v chatbotoch (Raamkumar a Yang, 2022) naprieč výskytmi v digitálnych knižniciach (Scopus, Google Scholar a IEEE Xplore) svojimi dátami potvrdzuje, že konverzačná umelá inteligencia (CAI) má populárny presah v komerčných aplikáciách vo forme personalizovaných digitálnych agentov, virtuálnych asistentov a chatbotov. Aby výsledky tejto práce konceptuálne splňali celú množinu tém implementácie CAI – ktoré ako uchopený celok vedú k schopnosti simulácie efektívneho zážitku z reálneho sveta, musela by moja práca ešte implementovať konverzačné vyhľadávacie či odporúčacie systémy. Kľúčový cieľ CAI avšak výsledky práce splňujú – chatbot identifikuje prevládajúci pocit používateľa a empaticky naňho reaguje. Vnímanie takejto afektívnej empatie sa podľa výskumu preukázalo ako pozitívna skúsenosť používateľov v oblasti zdravia a marketingu, čím sa ale téma môjho univerzitného chatbota dostáva do zanedbateľnej podmnožiny relevantných tém.

Realizácia detekcie emócie na modalite 1 – z textu, nie je pre túto prácu postihnutím, nakoľko je tento prístup v súčasných ESC štúdiách (Emotional Support Conversation) stále raritou (Raamkumar a Yang, 2022). Dominantným prínosom môjho konverzačného agenta je generovanie empatickej reakcie v *closed-domain* prostredí. Nástroj RASA sa potvrdil ako vhodne zvolený nástroj na tvorbu chatbota, nakoľko pri definovaní empatických odpovedí zohráva kľúčovú úlohu detekcia zámeru používateľa, ktorú tento nástroj v uzatvorenej doméne podporuje, a ja som ho pre tento účel docielila s presnosťou 0.9888. Začlenenie *commonsense knowledge* do modulu chatbota (v mojom prípade docielené pomocou COMET komponenty) na extrakciu *if-then* vzťahov z textového vstupu na generovanie empatických odpovedí, je rovnako teoretickou medzerou v implementáciách CAI, ktorú som naplnila. Z používateľského testovania vyplýva, že kombinácia vyššie uvedených prístupov a splnených výziev je v mojom agentovi interpretovaná s ohľadom na prívetivú *user experience*.

9.1 Etická úvaha

Mimo technických obmedzení, ktoré sú v nasledujúcej kapitole, je v uvažovaní o prínose tejto práce aj etický aspekt. Súbežným cieľom tejto práce bolo sledovanie emocionálneho vývoju v roli *speaker* a vyhodnocovanie zmiernenia stresu na konci konverzácie. Termínom emocionálna podpora, o ktorú sa cieľ tejto práce opiera, je v rovine uvažovania bežnej empatickej podpory nezainteresovaných strán. Implementovaný agent nie je z oblasti profesionálneho psychologického poradenstva, ani s používateľom následne nenadväzuje žiadny kontakt s vlastným spúšťačom. V prípade nadväzujúceho výskumu a prípadného produkčného nasadenia, je dôležitá implementácia rôznych krízových mechanizmov na minimalizovanie duševne alarmujúcich stavov

používateľa. V takom prípade je rozhodujúca hlbšia literárna rešerš z oblasti chatbotov pre psychologické účely.

Navrhované *empathetic policies*, ako pridružené odpovede agenta s účelom explicitného zmiernenia odchytenej negatívnej emócie, sú vytvorené bez psychologického vzdelania a poskytujú tak laickú množinu podporných procedúr. Generované odpovede a prvky konverzačnej podpory sú len základnou úrovňou podpory, ktoré môžeme považovať za opatrné, skromné a konzervatívne. Hoci použitý verejne dostupný *benchmark* ESConv dataset (Liu et al., 2021) pri zverejnení odfiltroval citlivé informácie, stále môže obsahovať výroky spúšťajúce určité emócie, ktoré som pri spracovaní datasetov mohla prehliadnúť.

9.2 Limity práce

Dialógy z využitého datasetu *ESConv* založenom na pomocných stratégiách (podobné interakciám s priateľmi alebo rodinou) často vyžadujú *multi-turn* konverzáciu, aby poskytli efektívnu emocionálnu podporu. Z povahy univerzitného chatbota nemusí byť primárny *use-case* dlhá všeobecná konverzácia, s ktorou sa spája zložitosť uchovávanía kontextu pri prerušení konverzácie. U spomínaných stratégií typu *Others* zavádzajú do procesu učenia nejasnosť a znehodnocujú pravdepodobnosť inej stratégie lepšie reflektujúcej vecnú emóciu.

Táto práca sa technicky zrealizovala adoptovaním najpoužívanejšieho datasetu pre koncept *Emotional Support Conversation* (ESC) – *EmpatheticDialogues* s 32 kategóriami emócií. V súčasnosti avšak neexistuje taká adaptácia tohto datasetu, ktorá by sa venovala hypotéze prebytočnosti a efektívnosti takého počtu emócií (Raamkumar a Yang, 2022).

Výrazným limitom tejto práce je abstrakcia napojenia na server *Moje MENDELU* s cieľom dopytovania sa odpovedí FAQ a jeho integrita v čase. Pripojenie ostatných modulov (mapa, jedálniček, študijné oddelenie, rozvrh či dôležitá správa) vyžadujú komplexnú analýzu z pohľadu relevantnosti a nasadenia konverzačného modelu do architektúry mobilného a webového vývoja. Integrita a perzistencia novej doménovej expertízy je ale veľmi nákladová. Rozširovanie doménovej expertízy u chatbota je v priamej úmere s narastajúcou robustnosťou scenárov, pravidiel a trénovacích dát. Ďalším limitom je klasifikátor 32 emócií *EmpatheticDialogues* (Rashkin et al., 2019), ktorý nedokáže presne rozlíšiť rovnaké výroky, ktorých emócie sa líšia od kontextu. Adoptovaním iného klasifikátoru by bolo rozlišovanie viet povýšené na úroveň úplného kontextu.

9.3 Nadväzujúci výskum

Smer, ktorým by sa navrhovaný chatbot mohol uberať ďalej v zmysle univerzitného produktu, je rozšírenie domény FAQ. Zo sémantiky domény je možno hovoriť o často pokladaných otázkach, ktoré sú ako otázky pevne formulované. Interaktívne

učenie v implementácii by ale dopomohlo rozšíriť množinu o nie tak často pokladané otázky, u ktorých by sa mohol monitorovať ich sentiment. Záujem o "Jak podat přihlášku k promoci?" by *support-seeker/speaker* (používateľ) pravdepodobne vo väčšine výskytov niesol emóciu radosti, zatiaľ čo záujem o "Jaké jsou následky disciplinárního řízení?" by mohol nieť spektrum odlišných emócií. Nový korpus, ktorý by výskumom mohol vzniknúť, by tak priniesol Študijnému oddeleniu nové poznatky o emocionálnom prežívaní v rôznych fázach študijného života, či úplne novú množinu kladených otázok. I v tomto prípade ale platí, že otázky neutrálne a tzv. "Others" výskum nikam zásadne neposúvajú.

Problematika navrhovaného agenta by sa v budúcnosti mohla prehĺbiť o zastrešenie konverzácie pod používateľský profil, ktorý by umožnil retrospektívne naviazanie na konverzáciu a zhodnotenie vývoja emócií v čase. Jedna z možných *empathetic policies*, ktorá presahuje kontinuálny vývoj dialógu, je naviazanie spojenia s používateľom na základe vytvorenej pripomienky k odpočinku, cvičeniu či inej stratégii v rámci emocionálnej podpory. V prípade iterácie na doménovej expertíze je možno ďalej uvažovať pri opakujúcej sa nízkej *confidence* (zámer pod danou *threshold* hodnotou) o odbavení takého *fallback* stavu práve ľudským účastníkom (human hand-off).

Zaujímavou nadstavbou chatbota s emocionálnou podporou je využitie detegovaných emócií k prispôbeniu používateľského rozhrania konverzácie: využiť *button selection* namiesto výzvy k textovej odpovedi, zapínanie/vypínanie emotikonov v odpovediach agenta, či úplne prispôbenie persóny agenta k persóne používateľa. Pri potrebe takého používateľského testovania sa môže potvrdiť hypotéza, že určitá persóna v emočnom stave "angry" nevyžaduje explicitnú empatickú podporu v odpovedi, ale skôr vecnú rýchlu odpoveď agenta na konkrétny zámer. V tomto prípade je možné hovoriť o implicitnom empatickom prejave agenta.

10 Záver

V tejto práci som skúmala problematiku emočnej inteligencie konverzačných agentov so zameraním na emocionálnu podporu a strategického plánovania reakcií agenta. Cieľ práce som naplnila hĺbkovou analýzou niekoľkých relevantných dostupných korpusov, hľadaním najlepšej konfigurácie jazykových modelov, prípravou implementácie na *open-source* platforme RASA, až po výskum a adaptáciu vhodných stratégií pri generovaní odpovedí agenta. Navrhnutý a implementovaný agent bol riadne otestovaný v zmysle jeho 2 častí: NLU a dialógový manažment. Presnosť, s ktorou agent klasifikuje používateľský zámer dosiahol hodnotu 0.9888. Rovnako kvalitne dopadla aj presnosť predpovedania scenáru konverzácie s každým používateľským vstupom, ktorá má hodnotu 1.00.

Kľúčovým prínosom agenta je generovanie empatickej reakcie v *closed-domain* prostredí, čomu predchádzalo netriviálne riešenie v nástroji RASA. Evaluácia vhodnej konfigurácie zvolenej architektúry bola časovo náročná, nakoľko tréning konverzačného modelu v úlohe spracovania prirodzeného jazyka (NLP) zahrňovalo analýzu datasetov, ich následné normalizovanie, tréning vlastných komponent nad objemovo náročnými korpusmi a samotnú integráciu komponent do projektu. Prezentované *state-of-art* empatické modely a frameworky (COMET a ESConv) boli výskumníkmi dostupné vo viacerých konfiguráciách nad jazykovými modelmi, čo spôsobovalo ďalšie experimentálne tréningy a integrovanie do konverzačného prostredia. Pri definovaní empatických odpovedí zohrávala kľúčovú úlohu detekcia zámeru používateľa v *closed-domain* prostredí, ktorú som s výslednou architektúrou pre účel tejto diplomovej práce docielila s presnosťou 0.9888, čím sa táto práca stáva potvrdením raritného konceptu.

Stanovená vlastnosť digitálneho asistenta – emočná inteligencia, bola nepochybneým determinantom toho, aby riešený systém umelej inteligencie rozpoznával emócie, porozumel svetu ľudí a mal zdroj informácií o základných vzťahoch medzi vecami v explicitnej forme pre výpočtové použitie (Ammanabrolu et al., 2021). Pre účel tejto práce vznikla rozsiahla literárna rešerš, ktorá chronologicky po časovej osy dáva do vzájomnej syntézy výsledky výskumných prác od prvotného využitia sémantiky v texte, cez tvorbu prvých emočných datasetov, identifikácie emočných stavov používateľa, až po *state-of-the-art* riešenie s využitím efektívnych mechanizmov na dynamické modelovanie emočného stavu používateľa v *multi-turn* dialógu. Súčasťou literárnej rešerše bolo vytýčenie dôležitých bodov, ktoré znižujú dôveryhodnosť chatbota ako aj spôsoby k docieleniu opačného požadovaného správania. Dôležité oporné body rešerše mi pomohli s orientáciou medzi benchmarkami afektívneho programovania, ktoré determinovali samotnú emočnú inteligenciu a implicitné vlastnosti takého systému.

Ďalej bola navrhnutá architektúra digitálneho asistenta (chatbota) pre aplikáciu *Moje MENDELU*. Návrh jeho *closed* aj *open* domén zahrňoval schopnosť zrealizovať identifikáciu výrazne vychýleného emočného stavu používateľa, ktorá by umožňovala ovplyvnenie scenáru, ktorým sa konverzácia v priebehu uberá.

Na zvolenej platforme RASA bol implementovaný navrhnutý chatbot, naplnený FAQ dátami a následne otestovaný na 5 používateľoch k vyhodnoteniu presnosti identifikácie emočného stavu a vyhodnoteniu jeho schopnosti zmierňovať emocionálny stav vďaka jeho persóny a dialógového manažmentu. Keďže sa v riešení architektúry navrhovaného agenta zaoberám implementáciou všeobecnej konverzácie, jej hlavnou výzvou je jej nasadenie v RASA. Výzva spočíva v riešení *chit-chat* dialógu, kedy natrénovaný model musí byť pripravený na akýkoľvek vstup a zároveň vo svojej konfigurácii musí mať zoznam tréningových výrokov na takú konverzáciu. Komunita RASA poskytuje len malú podporu na vytváranie vlastných komponent k spracovaniu správy používateľa, a tak mnou vytvorené komponenty založené na vyhodnocovaní *commonsense* vedomosti (spoločenské, priestorové a fyzické vzťahy k reálnemu svetu) a *generation-based* odpovedí (odpovede, ktoré sa nenachádzajú v tréningových dátach) sú netriviálnou výzvou v príprave navrhovaného agenta.

Výsledky práce sa opierajú o detekciu emócií iba z textu (singulárna modalita), čo povahu problematiky z výskumného hľadiska sťažuje, a vďaka využitiu niekoľkých benchmarkoch z oblasti korpusov a frameworkov, sú výsledky tejto práce vedecky aktuálne. Jedinečným prínosom práce je integrácia detegovaného afektu do doménového *task-oriented* chatbota, kde konverzačný modul aplikuje emočnú inteligenciu do *closed-domain* architektúry – FAQ. Modul je tak pripravený na otázku v akomkoľvek emocionálnom stave a je schopný na takto podanú otázku empaticky reagovať.

11 Literatúra

- ADIKARI, ACHINI ET AL. 2021. *A self structuring artificial intelligence framework for deep emotions modeling and analysis on the social web. Future Generation Computer Systems* [online]. Victoria, Australia: Research Centre for Data Analytics and Cognition, **116**(1), 302-315. [cit. 2022-10-15]. Dostupné z: <https://doi.org/10.1016/j.future.2020.10.028>.
- ADOMA, ACHEAMPONG FRANCISCA; HENRY, NUNOO-MENSAH; CHEN, WENYU. 2020. *Comparative Analyses of Bert, Roberta, Distilbert, and Xlnet for Text-Based Emotion Recognition*. In: *2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)* [online]. Chengdu, China: IEEE, 2020-12-18, s. 117-121. [cit. 2022-10-16]. ISBN 978-1-6654-0503-4. ISSN 2576-8964. Dostupné z: <https://doi.org/10.1109/ICCWAMTIP51612.2020.9317379>.
- ACHEAMPONG, FRANCISCA ADOMA; WENYU, CHEN; NUNOO-MENSAH, HENRY. 2020. *Advances, challenges, and opportunities. Engineering Reports*. John Wiley & Sons Ltd, **2**(7), Review. 2, 7. ISSN 2577-8196. Dostupné z: <https://doi.org/doi.org/10.1002/eng2.12189>.
- ALHUSSIEN, IKHLAS; CAMBRIA, ERIK; NENGSHENG, ZHANG. 2018. *Semantically Enhanced Models for Commonsense Knowledge Acquisition*. In: *2018 IEEE International Conference on Data Mining Workshops (ICDMW)* [online]. Singapore: IEEE, 2018, s. 1014-1021. [cit. 2022-10-16]. ISBN 978-1-5386-9288-2. ISSN 2375-9259. Dostupné z: <https://doi.org/10.1109/ICDMW.2018.00146>.
- AMMANABROLU, PRITHVIRAJ ET AL. 2021. *Automated Storytelling via Causal, Commonsense Plot Ordering. Proceedings of the AAI Conference on Artificial Intelligence* [online]. AAI, **35**, 5859—586. [cit. 2022-10-16]. ISSN 2159-5399. Dostupné z: <https://doi.org/10.1609/aaai.v35i7.16733>.
- AOUICHA, MOHAMED BEN; TAIEB, MOHAMED ALI HADJ; HAMADOU, ABDELMAJID BEN. 2016. *Taxonomy-based information content and wordnet-wiktionary-wikipedia glosses for semantic relatedness. Applied Intelligence* [online]. **45**, 475—511. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.1007/s10489-015-0755-x>.
- AUGELLO, AGNESE ET AL. 2016. *A Model of a Social Chatbot. Smart Innovation, Systems and Technologies* [online]. Springer, Cham, **55**, 637—647. [cit. 2022-10-16]. ISSN 2190-3026. Dostupné z: https://doi.org/10.1007/978-3-319-39345-2_57.
- BATISTA DOS SANTOS, EDIMILSON ET AL. 2019. *Bayesian Networks for Inference and Discovery of Semantic Relations in a Never-Ending Learning System*. In: *2019 8th Brazilian Conference on Intelligent Systems (BRACIS)*

- [online]. Salvador, Brazil: IEEE, 2019, s. 687-692. [cit. 2022-10-16]. ISBN 978-1-7281-4253-1. ISSN 2643-6264. Dostupné z: <https://doi.org/10.1109/BRACIS.2019.00125>.
- BENKE, IVO. 2020. *Towards Design Principles for Trustworthy Affective Chatbots in Virtual Teams. Proceedings of the Twenty-Eighth European Conference on Information Systems (ECIS2020): Designing Chatbots & Conversational Agents* [online]. Marrakesh, Morocco. [cit. 2022-10-16].
- BISK, YONATAN ET AL. 2020. *PIQA: Reasoning about Physical Commonsense in Natural Language. Proceedings of the AAAI Conference on Artificial Intelligence* [online]. AAAI, **34**(5), 7432—7439. [cit. 2022-10-16]. ISSN 2159-5399. Dostupné z: <https://doi.org/10.1609/aaai.v34i05.6239>.
- CALVO, RAFAEL A; D'MELLO, SIDNEY. 2010. *Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. IEEE Transactions on Affective Computing* [online]. **1**(1), 18-37. [cit. 2022-10-15]. ISSN 1949-3045. Dostupné z: <https://doi.org/10.1109/T-AFFC.2010.1>.
- CAMBRIA, ERIK. 2016. *Affective Computing and Sentiment Analysis. IEEE Intelligent Systems* [online]. **31**(2), 102-107. [cit. 2022-10-15]. ISSN 1541-1672. Dostupné z: <https://doi.org/10.1109/MIS.2016.31>.
- CAMBRIA, ERIK ET AL. 2018. *SenticNet 5: Discovering Conceptual Primitives for Sentiment Analysis by Means of Context Embeddings. Proceedings of the AAAI Conference on Artificial Intelligence* [online]. **32**(1). [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.1609/aaai.v32i1.11559>.
- CASSELL, JUSTINE ET AL. 1994. *Animated conversation*. In: *Proceedings of the 21st annual conference on Computer graphics and interactive techniques - SIGGRAPH '94* [online]. New York, New York, USA: ACM Press, 1994, s. 413-420. [cit. 2022-10-16]. ISBN 0897916670. Dostupné z: <https://doi.org/10.1145/192161.192272>.
- CLARK DAVIDSON, ELLIOTT. 1992. *The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System* [online]. Dizertačná práca. United States: Northwestern University. [cit. 2022-10-15]. Dostupné z: <https://dl.acm.org/doi/book/10.5555/142741>.
- COLBY, B. N. ET AL. 1989. *The Cognitive Structure of Emotions. Contemporary Sociology* [online]. **18**(6), 957-958. [cit. 2022-10-15]. ISSN 00943061. Dostupné z: <https://doi.org/10.2307/2074241>.
- DAILY, SHAUNDR A. B. ET AL. 2017. *Affective Computing: Historical Foundations, Current Applications, and Future Trends*. In: *Emotions and Affect in Human Factors and Human-Computer Interaction* [online]. United States: Elsevier, 2017, s. 213-231. [cit. 2022-10-15]. ISBN 978-0-12-801851-4.

- Dostupné z: <https://doi.org/10.1016/B978-0-12-801851-4.00009-4>.
- DAVIS, ERNEST; MARCUS, GARY. 2015. *SHARE ON Commonsense reasoning and commonsense knowledge in artificial intelligence. Communications of the ACM* [online]. New York, United States: Association for Computing Machinery, **58**(9), 92-103. [cit. 2022-10-16]. ISSN 0001-0782. Dostupné z: <https://doi.org/10.1145/2701413>.
- DAY, MIN-YUH; HUNG, CHI-SHENG. 2019. *AI Affective Conversational Robot with Hybrid Generative-Based and Retrieval-Based Dialogue Models* [online]. Los Angeles, CA, USA: IEEE, 403-409. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.1109/IRI.2019.00068>.
- DENECKE, KERSTIN; VAAHEESAN, SAYAN; ARULNATHAN, AAGANYA. 2021. *A Mental Health Chatbot for Regulating Emotions (SERMO) - Concept and Usability Test. IEEE Transactions on Emerging Topics in Computing* [online]. **9**(3), 1170-1182. [cit. 2022-10-16]. ISSN 2168-6750. Dostupné z: <https://doi.org/10.1109/TETC.2020.2974478>.
- DEVLIN, JACOB ET AL. 2019. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding: Google AI Language. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Volume 1* [online]. Minneapolis, Minnesota: Association for Computational Linguistics, 4171–4186. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.18653/v1/N19-142>.
- EKMAN, PAUL. 1993. *Facial Expression and Emotion. American Psychologist*. **48**(4), 384-392.
- EXPOSITO, MARC; HERNANDEZ, JAVIER; W. PICARD, ROSALIND. 2018. *Affective keys: towards unobtrusive stress sensing of smartphone users. Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* [online]. MobileHCI, **18**(1), 139–145. [cit. 2021-03-15]. Dostupné z: <https://doi.org/10.1145/3236112.3236132>.
- FEIDAKIS, M. 2016. *A Review of Emotion-Aware Systems for e-Learning in Virtual Environments*. In: *Formative Assessment, Learning Data Analytics and Gamification* [online]. United States: Elsevier, 2016, s. 217-242. [cit. 2022-10-15]. ISBN 978-0-12-801851-4. Dostupné z: <https://doi.org/10.1016/B978-0-12-803637-2.00011-7>.
- FORBES, MAXWELL; HOLTZMAN, ARI; CHOI, YEJIN. 2019. *Do Neural Language Representations Learn Physical Commonsense?. The Proceedings of the 41st Annual Conference of the Cognitive Science Society* [online]. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.48550/arXiv.1908.02899>.

- GHANDEHARIOUN, ASMA ET AL. 2019a. *EMMA: An Emotion-Aware Wellbeing Chatbot*. 8th International Conference on Affective Computing and Intelligent Interaction (ACII) [online]. IEEE, **8**, 1-7. [cit. 2021-03-09]. Dostupné z: <https://doi.org/10.1109/ACII.2019.8925455>.
- GHANDEHARIOUN, ASMA ET AL. 2019b. *Towards Understanding Emotional Intelligence for Behavior Change Chatbots*. International Conference on Affective Computing and Intelligent Interaction (ACII) [online]. IEEE, **8**(8), 8-14. [cit. 2021-03-11]. Dostupné z: <https://doi.org/10.1109/ACII.2019.8925433>.
- GHOSH, SURJYA ET AL. 2019. *EmoKey: An Emotion-aware Smartphone Keyboard for Mental Health Monitoring*. International Conference on Communication Systems & Networks (COMSNETS) [online]. IEEE, **11**(11), 496-499. [cit. 2021-03-14]. Dostupné z: <https://doi.org/10.1109/COMSNETS.2019.8711078>.
- HASEGAWA, TAKAYUKI ET AL. 2013. *Predicting and Eliciting Addressee-s Emotion in Online Dialogue*. Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics [online]. Sofia, Bulgaria: Association for Computational Linguistics, **1**, 964—972. [cit. 2022-10-16]. Dostupné z: <https://aclanthology.org/P13-1095/>.
- HAVASI, CATHERINE ET AL. 2009. *Digital Intuition: Applying Common Sense Using Dimensionality Reduction*. IEEE Intelligent Systems [online]. IEEE, **24**(4), 24 - 35. [cit. 2022-10-16]. ISSN 1941-1294. Dostupné z: <https://doi.org/10.1109/MIS.2009.72>.
- HEBBAR, AKSHAY. 2017. *Augmented intelligence: Enhancing human capabilities*. In: *2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)* [online]. 3. Kolkata, India: IEEE, 2017, s. 251-254. [cit. 2022-10-15]. ISBN 978-1-5386-1931-5. Dostupné z: <https://doi.org/10.1109/ICRCICN.2017.8234515>.
- HIGASHINAKA, RYUICHIRO ET AL. 2014. *Towards an open-domain conversational system fully based on natural language processing*. Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers [online]. Dublin, Ireland: Dublin City University and Association for Computational Linguistics, 928—939. [cit. 2022-10-16]. Dostupné z: <https://aclanthology.org/C14-1088/>.
- HUDLICKA, EVA. 2017. *Computational Modeling of Cognition—Emotion Interactions: Theoretical and Practical Relevance for Behavioral Healthcare*. In: *Emotions and Affect in Human Factors and Human—Computer Interaction*. [online]. United States: Academic Press, s. 383-436. [cit. 2022-11-26]. ISBN 978-0-12-801851-4. Dostupné z:

- <https://doi.org/10.1016/B978-0-12-801851-4.00016-1>.
- HUO, PEI ET AL. 2020. *TERG: Topic-Aware Emotional Response Generation for Chatbot*. In: *2020 International Joint Conference on Neural Networks (IJCNN)* [online]. IEEE, 2020, s. 1-8. [cit. 2022-10-16]. ISBN 978-1-7281-6926-2. Dostupné z: <https://doi.org/10.1109/IJCNN48605.2020.9206719>.
- HWANG, JENA D. ET AL. 2021. *COMET-ATOMIC 2020: On Symbolic and Neural Commonsense Knowledge Graphs. AAI Conference on Artificial Intelligence* [online]. AAI, **35**(7), 6384-6392. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.48550/arXiv.2010.05953>.
- CHAKRABARTY, TUHIN ET AL. 2020. *R3 : Reverse, Retrieve, and Rank for Sarcasm Generation with Commonsense Knowledge. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* [online]. Association for Computational Linguistics, 7976—7986. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.18653/v1/2020.acl-main.711>.
- CHENG, YI ET AL. 2022. *Improving Multi-turn Emotional Support Dialogue Generation with Lookahead Strategy Planning* [online]. [cit. 2022-11-23]. Dostupné z: <https://doi.org/10.48550/ARXIV.2210.04242>.
- CHEN, JENHUI; AGBODIKE, OBINNA; WANG, LEI. 2020. *Memory-Based Deep Neural Attention (mDNA) for Cognitive Multi-Turn Response Retrieval in Task-Oriented Chatbots. Applied Sciences* [online]. **10**(17). [cit. 2022-10-16]. ISSN 2076-3417. Dostupné z: <https://www.mdpi.com/2076-3417/10/17/5819>.
- CHOI, SAEMI; AIZAWA, KIYOHARU. 2019. *Emotype: Expressing emotions by changing typeface in mobile messenger texting. Multimedia Tools and Applications* [online]. **78**(1), 14155—14172. [cit. 2021-03-14]. Dostupné z: <https://doi.org/10.1007/s11042-018-6753-3>.
- CHOWANDA, ANDRY ET AL. 2021. *Exploring Text-based Emotions Recognition Machine Learning Techniques on Social Media Conversation. Procedia Computer Science* [online]. **179**(1), 821-828. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.1016/j.procs.2021.01.099>.
- CHRISTOPH, BARTNECK; LYONS, MICHAEL; SAERBECK, MARTIN. 2008. *The Relationship Between Emotion Models and Artificial Intelligence* [online]. Workshop on The Role of Emotion in Adaptive Behavior and Cognitive Robotics. Osaka, Japonsko: Department of Industrial Design Eindhoven University of Technology, College of Image Arts and Sciences Ritsumeikan University. [cit. 2022-11-26]. Dostupné z: <https://ir.canterbury.ac.nz/handle/10092/16708>.
- JADHAV, NAGESH; SUGANDHI, REKHA. 2018. *Survey on Human Behavior*

- Recognition Using Affective Computing*. In: *2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN)* [online]. IEEE, 2018, s. 98-103. [cit. 2022-10-15]. ISBN 978-1-5386-5201-5. Dostupné z: <https://doi.org/10.1109/GWCWN.2018.8668632>.
- JHA, NITESH KUMAR. 2018. *An Approach towards Text to Emoticon Conversion and Vice-Versa Using NLTK and WordNet*. In: *2018 2nd International Conference on Data Science and Business Analytics (ICDSBA)* [online]. Changsha, China: IEEE, 2018, s. 161-166. [cit. 2022-10-16]. ISBN 978-1-5386-8431-3. Dostupné z: <https://doi.org/10.1109/ICDSBA.2018.00036>.
- JING, HAN ET AL. 2015. *Cognitive emotion model for eldercare robot in smart home*. *China Communications* [online]. **12**(4), 32-41. [cit. 2022-10-15]. ISSN 1673-5447. Dostupné z: <https://doi.org/10.1109/CC.2015.7114067>.
- JI, ZONGCHENG; LU, ZHENGDONG; LI, HANG. 2014. *An Information Retrieval Approach to Short Text Conversation* [online]. [cit. 2022-10-16]. Dostupné z: <https://arxiv.org/abs/1408.6988>.
- JUSTO, ANDREY VICTOR ET AL. 2018. *Exploring Ontologies to Improve the Empathy of Interactive Bots*. In: *2018 IEEE 27th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)* [online]. Paris, France: IEEE, 2018, s. 261-266. [cit. 2022-10-16]. ISBN 978-1-5386-6916-7. Dostupné z: <https://doi.org/10.1109/WETICE.2018.00057>.
- KAHNEMAN, DANIEL. 2003. *Maps of Bounded Rationality: Psychology for Behavioral Economics*. *The American Economic Review* [online]. American Economic Association, **93**(5), 1449-1475. [cit. 2022-10-16]. Dostupné z: <https://www.jstor.org/stable/3132137>.
- KAO, EDWARD CHAO-CHUN ET AL. 2009. *Towards Text-based Emotion Detection A Survey and Possible Improvements*. *2009 International Conference on Information Management and Engineering* [online]. Kuala Lumpur, Malaysia: IEEE, 2009, **2009**(1), 70-74. [cit. 2022-10-15]. ISBN 978-0-7695-3595-1. Dostupné z: <https://doi.org/10.1109/ICIME.2009.113>.
- KEARNS, WILL. 2021. *Beyond Sentiment Analysis: Creating Engaging Conversational Experiences through Empathy*. *Rasa.com* [online]. [cit. 2022-10-16]. Dostupné z: https://www.slideshare.net/rasa_technologies/beyond-sentiment-analysis-creating-engaging-conversational-experiences-through-empathy-rasa-summit.
- KHATUA, APARUP ET AL. 2017. *Let-s Chat about Brexit! A Politically-Sensitive Dialog System Based on Twitter Data*. In: *2017 IEEE International Conference on Data Mining Workshops (ICDMW)* [online]. New Orleans, LA, USA: IEEE, 2017, s. 393-398. [cit. 2022-10-16]. ISBN 978-1-5386-3800-2. ISSN

- 2375-9259. Dostupné z: <https://doi.org/10.1109/ICDMW.2017.57>.
- KOŁAKOWSKA, AGATA; SZWOCH, WIOLETA; SZWOCH, MARIUSZ. 2020. *A Review of Emotion Recognition Methods Based on Data Acquired via Smartphone Sensors. Emotion Monitoring System Based on Sensors and Data Analysis* [online]. *Sensors*, **2020**(6367). [cit. 2021-03-09]. Dostupné z: <https://doi.org/10.3390/s20216367>.
- KOSTI, RONAK ET AL. 2017. *Emotion Recognition in Context*. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. Honolulu, HI, USA: IEEE, 2017, s. 1960-1968. [cit. 2022-10-15]. ISBN 978-1-5386-0457-1. ISSN 1063-6919. Dostupné z: <https://doi.org/10.1109/CVPR.2017.212>.
- KRATZWALD, BERNHARD ET AL. 2018. *Deep learning for affective computing: Text-based emotion recognition in decision support. Decision Support Systems* [online]. **115**(1), 24-35. [cit. 2022-10-15]. ISSN 01679236. Dostupné z: <https://doi.org/10.1016/j.dss.2018.09.002>.
- KRUGER, HENNIE ET AL. 2020. *Acquiring sentiment towards information security policies through affective computing. 2020 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC)* [online]. IEEE, 2020-11-25, **2020**(2), 1-6. [cit. 2022-10-15]. ISBN 978-1-7281-9520-9. Dostupné z: <https://doi.org/10.1109/IMITEC50163.2020.9334134>.
- KUNDI, MAHWISH; CHITCHYAN, RUZANNA. 2017. *Use Case Elicitation with FrameNet Frames*. In: *2017 IEEE 25th International Requirements Engineering Conference Workshops (REW)* [online]. Lisbon, Portugal: IEEE, 2017, s. 224-231. [cit. 2022-10-16]. ISBN 978-1-5386-3488-2. Dostupné z: <https://doi.org/10.1109/REW.2017.53>.
- LARSON, REED; CSIKSZENTMIHALYI, MIHALY. 2014. *Validity and Reliability of the Experience-Sampling Method. Flow and the Foundations of Positive Psychology* [online]. Springer, Dordrecht, **First online**, 35-54. [cit. 2021-03-11]. Dostupné z: https://doi.org/10.1007/978-94-017-9088-8_3.
- LI, CHI-HSUN; CHEN, KEN; CHANG, YUNG-JU. 2019. *When There is No Progress with a Task-Oriented Chatbot*. In: *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* [online]. New York, NY, USA: ACM, 2019, s. 1-6. [cit. 2022-10-16]. ISBN 9781450368254. Dostupné z: <https://doi.org/10.1145/3338286.3344407>.
- LIU, BINGJIE; SUNDAR, S. SHYAM. 2018. *Should Machines Express Sympathy and Empathy? Experiments with a Health Advice Chatbot. Cyberpsychology, Behavior, and Social Networking* [online]. **21**(10), 625-636. [cit. 2021-03-07].

- Dostupné z: <https://doi.org/10.1089/cyber.2018.0110>.
- LIU, SIYANG ET AL. 2021. *Towards Emotional Support Dialog Systems. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, [online]. Association for Computational Linguistics, **59**, 3469—3483. [cit. 2022-11-19]. Dostupné z: <https://arxiv.org/abs/2106.01144>.
- LI, XIN ET AL. 2020. *Relevance-Promoting Language Model for Short-Text Conversation. AAAI Technical Track: Natural Language Processing* [online]. **34**(5), 8253-8260. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.1609/aaai.v34i05.6340>.
- LOWE, RYAN ET AL. 2015. *The Ubuntu Dialogue Corpus: A Large Dataset for Research in Unstructured Multi-Turn Dialogue Systems. Proc. SIGDIAL 16* [online]. 285-294. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.48550/arXiv.1506.08909>.
- LUBIS, NURUL ET AL. 2019. *Positive Emotion Elicitation in Chat-Based Dialogue Systems. IEEE/ACM Transactions on Audio, Speech, and Language Processing* [online]. **27**(4), 866-877. [cit. 2022-10-16]. ISSN 2329-9290. Dostupné z: <https://doi.org/10.1109/TASLP.2019.2900910>.
- MADELEINE, GRUNDE-MCLAUGHLIN; RANJAY, KRISHNA; MANEESH, GRAWALA. 2021. *AGQA: A Benchmark for Compositional Spatio-Temporal Reasoning. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. Nashville, TN, USA: IEEE, 11282-11292. [cit. 2022-10-16]. ISSN 2575-7075. Dostupné z: <https://doi.org/10.1109/cvpr46437.2021.01113>.
- MATUSZEK, CYNTHIA ET AL. 2006. *An Introduction to the Syntax and Content of Cyc. Formalizing and Compiling Background Knowledge and Its Applications to Knowledge Representation and Question Answering*. [online]. Stanford, California, USA: AAAI, 44-49. [cit. 2022-10-16].
- MEHRABIAN, ALBERT; RUSSELL, JAMES A. 1974. *The basic emotional impact of environments. Perceptual and Motor Skills*. **38**(1), 283—301. Dostupné z: <https://doi.org/10.2466/pms.1974.38.1.283>.
- MENDELOVA UNIVERZITA V BRNĚ 2019. *My MENDELU*. [software]. Verzia 2.3.0. [cit. 2022-10-26]. Dostupné z: <https://apps.apple.com/us/app/moje-mendelu/id1467807496?l=cs>.
- MINSKY, MARVIN. 1988. *Society of mind* In: *Society of mind*. New York: Simon & Schuster Paperbacks, s. 163. ISBN 978-0671657130.
- MONFORT, MATHEW ET AL. 2020. *Moments in Time Dataset: One Million Videos for Event Understanding. IEEE Transactions on Pattern Analysis and*

- Machine Intelligence* [online]. IEEE, **42**(2), 502-508. [cit. 2022-10-16]. ISSN 0162-8828. Dostupné z: <https://doi.org/10.1109/TPAMI.2019.2901464>.
- MOSCHONA, DANAI STYLIANI. 2020. *An Affective Service based on Multi-Modal Emotion Recognition, using EEG enabled Emotion Tracking and Speech Emotion Recognition*. In: *2020 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia)* [online]. Seoul, Korea (South): IEEE, 2020-11-1, s. 1-3. [cit. 2022-10-15]. ISBN 978-1-7281-6164-8. ISSN 20257698. Dostupné z: <https://doi.org/10.1109/ICCE-Asia49877.2020.9277291>.
- NAMAVAR JAHROMI, AMIR; HOMAYOUNPOUR, MOHAMMAD MEHDI. 2012. *Emotions from Farsi texts with mutual-word-counting and word-spotting*. In: *The 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012)* [online]. 16. Shiraz, Iran: IEEE, 2012, s. 339-342. [cit. 2022-10-15]. ISBN 978-1-4673-1479-4. Dostupné z: <https://doi.org/10.1109/AISP.2012.6313769>.
- NANDWANI, PANSY; VERMA, RUPALI. 2021. *A review on sentiment analysis and emotion detection from text*. *Social Network Analysis and Mining* [online]. **11**(1), Article number: 81. [cit. 2022-10-15]. ISSN 1869-5450. Dostupné z: <https://doi.org/10.1007/s13278-021-00776-6>.
- NAROUEI, MASOUD; TAKABI, HASSAN; NIELSEN, RODNEY. 2020. *Automatic Extraction of Access Control Policies from Natural Language Documents*. *IEEE Transactions on Dependable and Secure Computing* [online]. **17**(3), 506-517. [cit. 2022-10-16]. ISSN 1545-5971. Dostupné z: <https://doi.org/10.1109/TDSC.2018.2818708>.
- PANTIC, MAJA ET AL. 2005. *Affective multimodal human-computer interaction*. *Proceedings of the 13th ACM International Conference on Multimedia* [online]. **13**(1), 669-676. [cit. 2022-10-15]. Dostupné z: <https://doi.org/10.1145/1101149.1101299>.
- PARK, SEO-HUI; BAE, BYUNG-CHULL; CHEONG, YUN-GYUNG. 2020. *Emotion Recognition from Text Stories Using an Emotion Embedding Model*. In: *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)* [online]. Busan, Korea (South): IEEE, 2020, s. 579-583. [cit. 2022-10-15]. ISBN 978-1-7281-6034-4. ISSN 2375-9356. Dostupné z: <https://doi.org/10.1109/BigComp48618.2020.00014>.
- PENTLAND, ALEX. 2005. *Socially aware, computation and communication*. *Computer* [online]. IEEE, **38**(3), 33-40. [cit. 2022-10-15]. Dostupné z: <https://doi.org/10.1109/MC.2005.104>.
- PETRONI, FABIO ET AL. 2019. *Language Models as Knowledge Bases?*. Dostupné také z: <https://arxiv.org/abs/1909.01066>.

- PICARD, ROSALIND W. 1995. *Affective computing*. MIT Media Laboratory Perceptual Computing. 321. Cambridge.
- PREOȚIU-C-PIETRO, DANIEL ET AL. 2016. *Modelling Valence and Arousal in Facebook posts. Proceedings of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* [online]. San Diego, California: Association for Computational Linguistics, 7(1), 9—15. [cit. 2022-10-15]. Dostupné z: <https://doi.org/10.18653/v1/W16-0404>.
- RAAMKUMAR, ARAVIND SESAGIRI; YANG, YINPING. 2022. *Empathetic Conversational Systems: A Review of Current Advances, Gaps, and Opportunities. IEEE Transactions on Affective Computing* [online]. IEEE, 1-20. [cit. 2022-12-26]. ISSN 1949-3045. Dostupné z: <https://doi.org/10.1109/TAFFC.2022.3226693>.
- RALSTON, KENNEDY ET AL. 2019. *A Voice Interactive Multilingual Student Support System using IBM Watson*. In: *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)* [online]. IEEE, 2019, s. 1924-1929. [cit. 2022-10-16]. ISBN 978-1-7281-4550-1. Dostupné z: <https://doi.org/10.1109/ICMLA.2019.00309>.
- RANADE, ADVAIT GOPAL; PATEL, MAITRI; MAGARE, ARCHANA. 2018. *Emotion Model for Artificial Intelligence and their Applications*. In: *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)* [online]. Solan, India: IEEE, s. 335-339. [cit. 2022-10-15]. ISBN 978-1-7281-0646-5. ISSN 2573-3079. Dostupné z: <https://doi.org/10.1109/PDGC.2018.8745840>.
- RASA TECHNOLOGIES GMBH. 2022. [online]. [cit. 2022-11-30]. Dostupné z: <https://rasa.com/>.
- RASHKIN, HANNAH ET AL. 2019. *Towards Empathetic Open-domain Conversation Models: a New Benchmark and Dataset. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* [online]. 57. [cit. 2022-11-19]. Dostupné z: <https://doi.org/10.48550/arXiv.1811.00207>.
- REILLY, W. SCOTT NEAL. 1996. *Believable Social and Emotional Agents* [online]. Dizertačná práca. USA: School of Computer Science, Carnegie Mellon University. [cit. 2022-10-15]. Dostupné z: <https://api.semanticscholar.org/CorpusID:12898304>.
- RITTER, ALAN; CHERRY, COLIN; B. DOLAN, WILLIAM. 2011. *Data-Driven Response Generation in Social Media. Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* [online]. Edinburgh, Scotland, UK: Association for Computational Linguistics, 583—593. [cit. 2022-10-16]. Dostupné z: <https://aclanthology.org/D11-1054/>.

- ROLLER, STEPHEN ET AL. 2021. *Recipes for Building an Open-Domain Chatbot. Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume* [online]. Association for Computational Linguistics, 2021, **16**, 300—325. [cit. 2022-11-19]. Dostupné z: <https://doi.org/10.18653/v1/2021.eacl-main.24>.
- RUSSELL, JAMES A. 1980. *A circumplex model of affect. Journal of Personality and Social Psychology* [online]. **39**(6), 1161-1178. [cit. 2022-10-15]. ISSN 1939-1315. Dostupné z: <https://doi.org/10.1037/h0077714>.
- SEO, YEONG-SEOK; HUH, JUN-HO. 2019. *Automatic Emotion-Based Music Classification for Supporting Intelligent IoT Applications. Electronics 2019* [online]. **8**(2), 164. [cit. 2022-10-15]. Dostupné z: <https://www.mdpi.com/2079-9292/8/2/164>.
- SEYEDITABARI, ARMIN; TABARI, NARGES; ZADROZNY, WLODEK. 2018. *Emotion Detection in Text: a Review*. UNC Charlotte. doi.org/10.48550/arXiv.1806.00674. Dostupné také z: <https://arxiv.org/abs/1806.00674>.
- SHABBIR, JOHAR; ARSHA, MUHAMMAD UMAIR; SHAHZAD, WASEEM. 2021. *NUBOT: Embedded Knowledge Graph With RASA Framework for Generating Semantic Intents Responses in Roman Urdu*. [online]. [cit. 2022-11-27]. Dostupné z: <https://arxiv.org/abs/2102.10410>.
- SHEVAT, AMIR. 2017. *Designing Bots: Creating Conversational Experiences*. O'Reilly Media. ISBN 978-1491974827.
- SHUSTER, KURT ET AL. 2021. *Retrieval Augmentation Reduces Hallucination in Conversation. Conference on Empirical Methods in Natural Language Processing* [online]. [cit. 2022-11-22].
- SIGDEL, BIJAY ET AL. 2020. *Testing QA Systems- ability in Processing Synonym Commonsense Knowledge*. In: *2020 24th International Conference Information Visualisation (IV)* [online]. Melbourne, Australia: IEEE, 2020, s. 317-321. [cit. 2022-10-16]. ISBN 978-1-7281-9134-8. ISSN 2375-0138. Dostupné z: <https://doi.org/10.1109/IV51561.2020.00059>.
- SONG, ZHENQIAO ET AL. 2019. *Generating Responses with a Specific Emotion in Dialog. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* [online]. Florence, Italy: Association for Computational Linguistics, 3685—3695. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.18653/v1/P19-1359>.
- SUN, XIAO ET AL. 2018. *Emotional Human Machine Conversation Generation Based on SeqGAN*. In: *2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia)* [online]. Beijing, China: IEEE, 2018,

- s. 1-6. [cit. 2022-10-16]. ISBN 978-1-5386-5311-1. Dostupné z: <https://doi.org/10.1109/ACIIAsia.2018.8470388>.
- TALMOR, ALON ET AL. 2020. *OLMpics-On What Language Model Pre-training Captures*. *Transactions of the Association for Computational Linguistics* [online]. **8**(23), 743-758. [cit. 2022-10-16]. Dostupné z: https://doi.org/10.1162/tacl_a_00342.
- THOMASON, JESSE ET AL. 2016. *Learning multi-modal grounded linguistic semantics by playing "I Spy."* *IJCAI'16: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence* [online]. AAAI Press, 3477—3483. [cit. 2022-10-16]. Dostupné z: <https://dl.acm.org/doi/abs/10.5555/3061053.3061107>.
- TU, QUAN ET AL. 2022. *MISC: A Mixed Strategy-Aware Model Integrating COMET for Emotional Support Conversation*. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics* [online]. Association for Computational Linguistics, **60**(1), 308 - 319. [cit. 2022-11-19]. Dostupné z: <https://doi.org/10.48550/arXiv.2203.13560>.
- SAHA, TULIKA; ANANIADOU, SOPHIA. 2022. *Emotion-aware and Intent-controlled Empathetic Response Generation using Hierarchical Transformer Network*. *2022 International Joint Conference on Neural Networks (IJCNN)* [online]. IEEE, 2022-7-18, 1-8. [cit. 2022-12-26]. ISBN 978-1-7281-8671-9. Dostupné z: <https://doi.org/10.1109/IJCNN55064.2022.9892592>.
- TURGANALIYEV, ALISHER. 2021. *Osobní asistent pro cizince pracující v České republice*. Diplomová práce. Brno: Mendelova univerzita v Brně.
- WATSON, DAVID; CLARK, LEE ANNA; TELLEGEN, AUKE. 1988. *Development and validation of brief measures of positive and negative affect: The PANAS scales*. *Journal of Personality and Social Psychology* [online]. **54**(6), 1063-1070. [cit. 2021-03-09]. Dostupné z: <https://doi.org/10.1037/0022-3514.54.6.1063>.
- WEISZ, JUSTIN D. ET AL. 2019. *BigBlueBot*. In: *Proceedings of the 24th International Conference on Intelligent User Interfaces* [online]. New York, NY, USA: ACM, 2019-03-17, s. 448-459. [cit. 2022-10-16]. ISBN 9781450362726. Dostupné z: <https://doi.org/10.1145/3301275.3302290>.
- WELIVITA, ANURADHA; PU, PEARL. 2020. *A Taxonomy of Empathetic Response Intents in Human Social Conversations*. *Proceedings of the 28th International Conference on Computational Linguistics* [online]. Barcelona, Spain: International Committee on Computational Linguistics, **28**, 4886–4899. [cit. 2022-11-19]. Dostupné z: <https://doi.org/10.18653/v1/2020.coling-main.429>.

- WITBROCK, MICHAEL. 2010. *Acquiring and using large scale knowledge*. *Proceedings of the ITI 2010, 32nd International Conference on Information Technology Interfaces* [online]. Cavtat, Croatia: IEEE, 37-42. [cit. 2022-10-16]. ISSN 1330-1012. Dostupné z: <https://ieeexplore.ieee.org/document/5546360>.
- XIE, BAIJUN; PARK, CHUNG HYUK. 2021. *Empathetic Robot With Transformer-Based Dialogue Agent*. 18th International Conference on Ubiquitous Robots (UR) [online]. Gangwon-do, Korea, 18. [cit. 2022-12-19].
- XU, SHIHAO ET AL. 2020. *Emotion Recognition From Gait Analyses: Current Research and Future Directions*. Human-Computer Interaction (cs.HC). Dostupné také z: <https://arxiv.org/abs/2003.11461>.
- YANG, YONG; SUN, YUE. 2017. *Facial Expression Recognition Based on Arousal-Valence Emotion Model and Deep Learning Method*. 2017 International Conference on Computer Technology, Electronics and Communication (ICCTEC) [online]. IEEE, 2017, **2017**(1), 59-62. [cit. 2022-10-15]. ISBN 978-1-5386-5784-3. Dostupné z: <https://doi.org/10.1109/ICCTEC.2017.00022>.
- YAN, RUI; SONG, YIPING; WU, HUA. 2016. *Learning to Respond with Deep Neural Networks for Retrieval-Based Human-Computer Conversation System*. *SIGIR '16: Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* [online]. New York, NY, USA: Association for Computing Machinery, 55—64. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.1145/2911451.2911542>.
- YILDIRIM, SAVAS ET AL. 2019. *Building Domain-Specific Lexicons: An Application to Financial News*. In: *2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)* [online]. Istanbul, Turkey: IEEE, 2019, s. 23-26. [cit. 2022-10-16]. ISBN 978-1-7281-2914-3. Dostupné z: <https://doi.org/10.1109/Deep-ML.2019.00013>.
- YOO, SOYEOP; JEONG, OKRAN. 2021. *EP-Bot: Empathetic Chatbot Using Auto-Growing Knowledge Graph*. *Computers, Materials & Continua* 2021 [online]. **67**(3), 2807-2817. [cit. 2022-10-15]. Dostupné z: <https://doi.org/10.32604/cmc.2021.015634>.
- ZELLERS, ROWAN ET AL. 2019a. *From Recognition to Cognition: Visual Commonsense Reasoning*. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [online]. Long Beach, CA, USA: IEE, 6713-6724. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.1109/CVPR.2019.00688>.
- ZELLERS, ROWAN ET AL. 2019b. *HellaSwag: Can a Machine Really Finish Your Sentence?*. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* [online]. Florence, Italy: HellaSwag: Can a Machine

- Really Finish Your Sentence?, 4791—4800. [cit. 2022-10-16]. Dostupné z: <https://doi.org/10.18653/v1/P19-1472>.
- ZHANG, JIANHUA ET AL. 2020. *Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review*. *Information Fusion* [online]. **59**, 103-126. [cit. 2022-11-17]. ISSN 15662535. Dostupné z: <https://doi.org/10.1016/j.inffus.2020.01.011>.
- ZHANG, XIAO ET AL. 2018. *MoodExplorer: Towards Compound Emotion Detection via Smartphone Sensing*. *Proceedings of the ACM on Interactive Mobile Wearable and Ubiquitous Technologies* [online]. **1**(4), 1-30. [cit. 2022-10-13]. Dostupné z: <https://doi.org/10.1145/3161414>.
- ZHANG, YIZHE ET AL. 2020. *DIALOGPT : Large-Scale Generative Pre-training for Conversational Response Generation*. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* [online]. Association for Computational Linguistics, 270-278. [cit. 2022-11-19]. Dostupné z: <https://doi.org/10.18653/v1/2020.acl-demos.30>.
- ZHOU, HAO ET AL. 2018. *Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory*. *AAAI'18: AAAI Conference on Artificial Intelligence* [online]. New Orleans, Louisiana, USA, 730-738. [cit. 2022-10-16]. Dostupné z: <https://arxiv.org/abs/1704.01074>.
- ZHOU, LI ET AL. 2020. *The Design and Implementation of XiaoIce, an Empathetic Social Chatbot*. *Computational Linguistics* [online]. **46**(1), 1-62. [cit. 2022-10-15]. Dostupné z: https://doi.org/10.1162/COLI_a_00368.
- ZHOU, XIANDA; WANG, WILLIAM YANG. 2018. *MojiTalk: Generating Emotional Responses at Scale*. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* [online]. Melbourne, Australia: Association for Computational Linguistics, **1**, 1128-1137. [cit. 2022-11-19]. Dostupné z: <https://doi.org/10.18653/v1/P18-1104>.
- ZHU, YIXIN ET AL. 2020. *Dark, Beyond Deep: A Paradigm Shift to Cognitive AI with Humanlike Common Sense*. *Engineering* [online]. **6**(3), 310-345. [cit. 2022-10-16]. ISSN 2095-8099. Dostupné z: <https://doi.org/10.1016/j.eng.2020.01.011>.

Prílohy

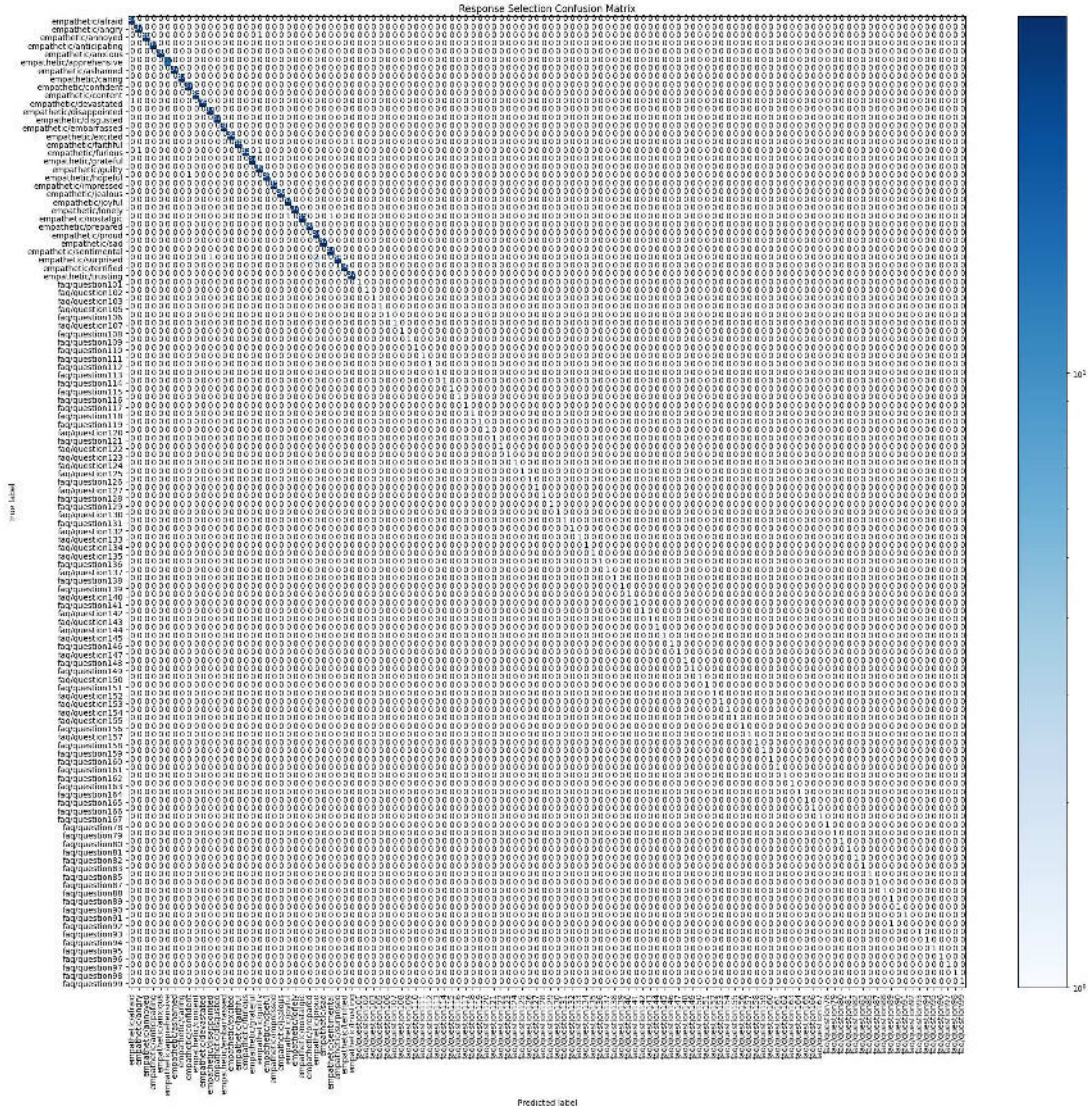
A Textový súbor požiadaviek

```
1  pytorch-lightning-0.8.5
2  torch-1.12.1
3  torchvision-0.13.1
4  filelock-3.8.0
5  huggingface-hub-0.0.12
6  sacremoses-0.0.53
7  tokenizers-0.10.3
8  transformers-4.9.2
9  nltk-3.7
10 rouge_score-0.1.2
11 colorama-0.4.6
12 lxml-4.9.1
13 portalocker-2.6.0
14 sacrebleu-2.3.1
```

B Konfiguračný súbor config.yml

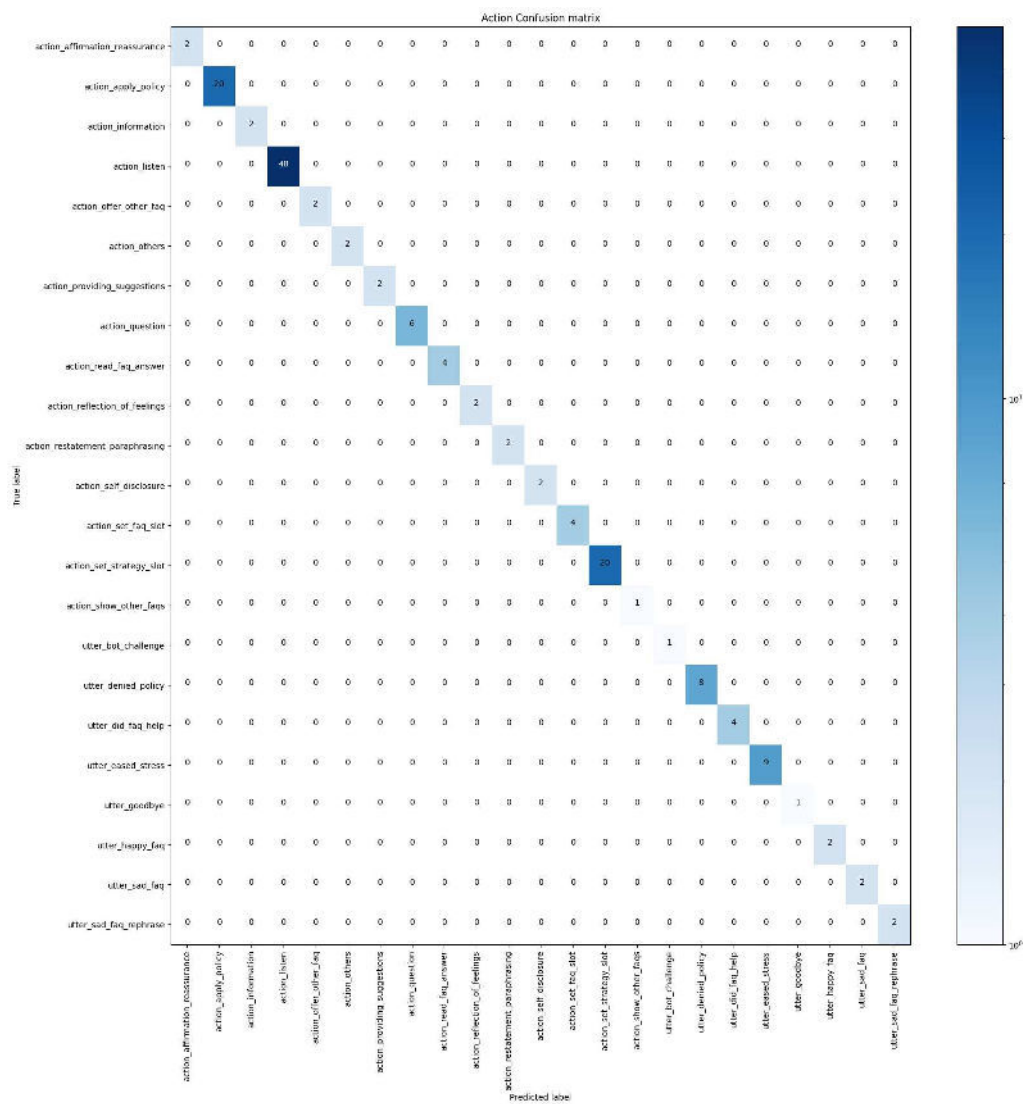
```
1 # The config recipe.
2 # https://rasa.com/docs/rasa/model-configuration/
3 recipe: default.v1
4
5 # Configuration for Rasa NLU.
6 # https://rasa.com/docs/rasa/nlu/components/
7 language: en
8 pipeline:
9   - name: WhitespaceTokenizer
10  - name: RegexFeaturizer
11  - name: LexicalSyntacticFeaturizer
12  - name: CountVectorsFeaturizer
13  - name: CountVectorsFeaturizer
14    analyzer: "char_wb"
15    min_ngram: 1
16    max_ngram: 4
17  - name: DIETClassifier
18    epochs: 200
19    ranking_length: 5
20  - name: components.RelationsClassifier
21  - name: components.EmotionalStrategyClassifier
22  - name: EntitySynonymMapper
23  - name: ResponseSelector
24    epochs: 100
25    retrieval_intent: faq
26    constrain_similarities: true
27  - name: ResponseSelector
28    epochs: 100
29    retrieval_intent: empathetic
30    constrain_similarities: true
31  - name: FallbackClassifier
32    threshold: 0.3
33    ambiguity_threshold: 0.1
34
35 # Configuration for Rasa Core.
36 # https://rasa.com/docs/rasa/core/policies/
37 policies:
38   - name: MemoizationPolicy
39     max_history: 5
40   - name: AugmentedMemoizationPolicy
41     max_history: 5
42   - name: RulePolicy
43     core_fallback_threshold: 0.3
44     max_history: 3
45   - name: TEDPolicy
46     max_history: 5
47     epochs: 100
```

C Evaluácia Response Selector časti NLU modelu



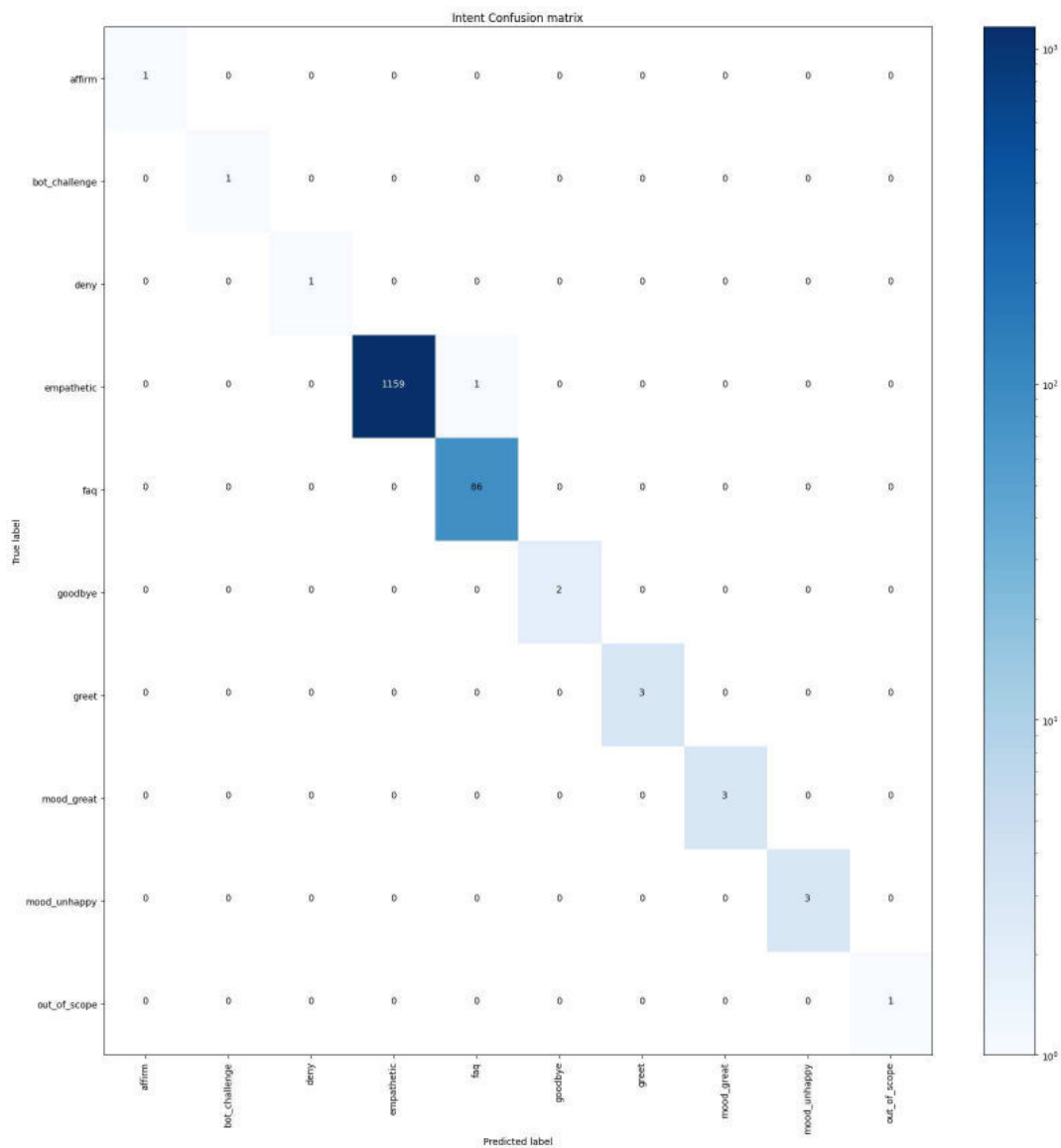
Obr. 41: Matica ukazuje, ako často boli konkrétne FAQ otázky či emócie správne predpovedané (Zdroj: autorka práce)

D Evaluácia dialógového modelu

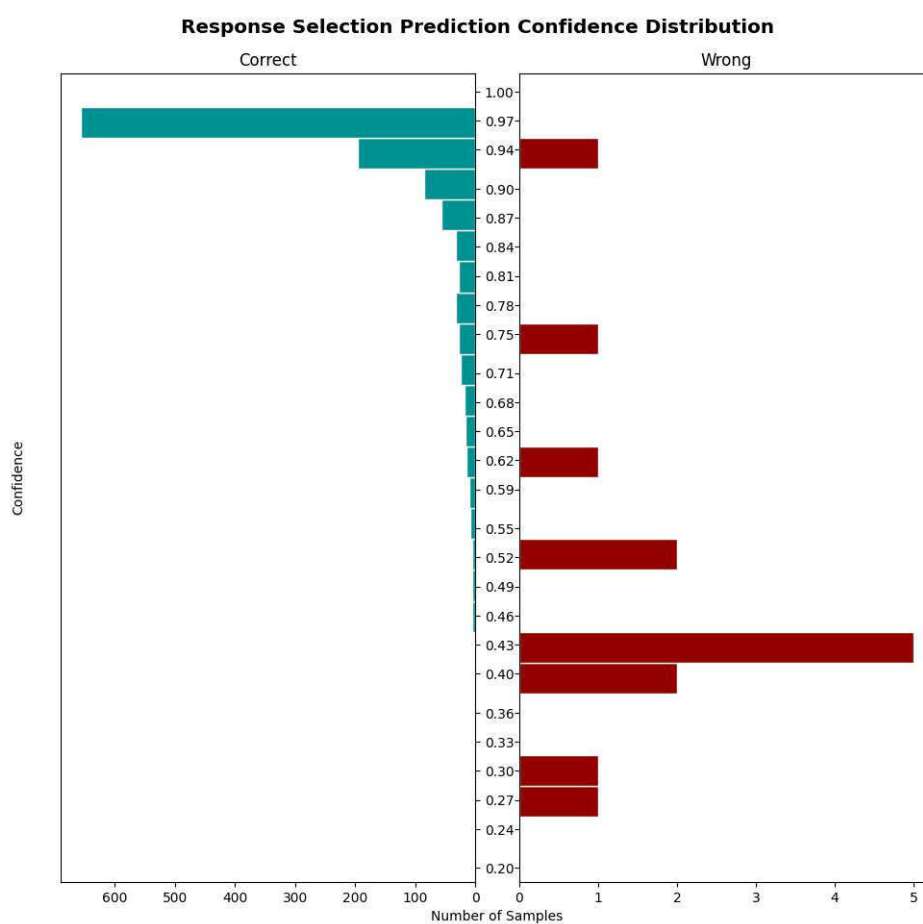


Obr. 42: Matica ukazuje, ako často bola daná akcia v scenári správne predpovedaná a ako často bola namiesto toho predpovedaná nesprávna akcia (Zdroj: autorka práce)

E Evaluácia NLU modelu



Obr. 43: Matica ukazuje, ktoré zámery sa mýlia s inými (Zdroj: autorka práce)



Obr. 44: Histogram rozloženia istoty pri predikovaní zámeru (Zdroj: autorka práce)