ДИПЛОМНА РОБОТА
БАКАЛАВРА

ДРБ. ІП - 82.00.00.000 ПЗ

Група ІП-19-1К

Слободян Ігор

2022

Міністерство освіти і науки України
Івано-Франківський національний технічний університет нафти і газу

Інститут інформаційних технологій

Кафедра інженерії програмного забезпечення

УДК 004.4

ДИПЛОМНА РОБОТА БАКАЛАВРА

Тема: " Інформаційні моделі візуалізації програмних сервісів та рішень"

(назва згідно з наказом ректора)

спеціальність – 121 "Інженерія програмного забезпечення"

ПОЯСНЮВАЛЬНА ЗАПИСКА

ДРБ.ІП - 82.00.00.000 ПЗ

(позначення)

Консультанти:

Студент

ІП-19-1К_____ /І.В.Слободян/
(шифр групи) (підпис) (дата) (розшифровка підпису)

Керівник роботи

к.т.н.,доц. _____ /М.М.Яцишин/
(посада) (підпис) (дата) (розшифровка підпису)

Допускається до захисту
Завідувач кафедри

з питань нормоконтролю

доц., к.т.н_____ /В.Я. Піх/
(посада) (підпис) (дата)(розшифровка підпису)

д.т.н.,проф. _____ /В.І.Шекета/
(посада) (підпис) (дата) (розшифровка підпису)

2022
(рік)

Івано-Франківський національний технічний університет нафти і газу

Інститут, факультет інформаційних технологій

Кафедра інженерія програмного забезпечення

Освітньо-кваліфікаційний рівень бакалавр

Спеціальність 121 – Інженерія програмного забезпечення

ЗАТВЕРДЖУЮ:
Зав. кафедрою **ІПЗ**
проф. В.І. Шекета
“ ____ ” _____2022 р.

ЗАВДАННЯ

НА ДИПЛОМНУ РОБОТУ БАКАЛАВРА СТУДЕНТОВІ

Слободяну Ігорю Васильовичу

(прізвище, ім’я, по-батькові)

1. Тема проекту (роботи) “ Інформаційні моделі візуалізації програмних сервісів та рішень”

керівник проекту (роботи) Яцишин Микола Миколайович, к.т.н.

затверджені наказом закладу вищої освіти від **“** ___ ” __травня__ 2022 р. № _____

2. Строк подання студентом проекту (роботи) 20 червня 2022 р.

3. Вихідні дані до проекту (роботи) Результати і матеріали отримані під час проходження переддипломної практики

4. Зміст розрахунково - пояснювальної записки (перелік питань, які потрібно розробити)

1. Огляд існуючих програмних рішень та сервісів в предметній області візуалізації
2. Алгоритмізація процесу вилучення та аналізу RST CUE
3. Програмна реалізація та результати виконання імітаційного моделювання
4. Представлення інтерфейсу
5. Виконання процесу тестування розробленого програмного забезпечення

5. Перелік графічного матеріалу (з точним зазначенням обов’язкових креслень)

1. Приклад публікації CMV разом із коментарями, наданими для зміни вигляду (рис. 1.1)
2. Щільність розподілу оцінок релевантності для переконливих коментарів (рис. 1.2)
3. Щільність розподілу часових оцінок для переконливих коментарів (рис. 1.3)
4. Середня оцінка релевантності в кожному квартальному періоді (рис. 1.4)
5. Результат аналізу ознак (рис. 1.5)

6. Консультанти розділів проекту (роботи)

Розділ	Консультант	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання **25 травня 2022 р.**

Керівник _____
(підпис)

Завдання прийняв до виконання_____
(підпис)

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту	Примітка
1	Огляд існуючих програмних рішень та сервісів в предметній області візуалізації	26.05.2022	виконано
2	Алгоритмізація процесу вилучення та аналізу RST CUE	29.05.2022	виконано
3	Програмна реалізація та результати виконання імітаційного моделювання	30.05.2022	виконано
4	Програмна реалізація інтерфейсу користувача	04.06.2022	виконано
5	Тестування ефективності розробленого програмного забезпечення	10.06.2022	виконано
6	Оформлення пояснювальної записки дипломної роботи завідувачем кафедри	19.06.2022	виконано

Студент – дипломник _____
(підпис)

Керівник роботи _____
(підпис)

АНОТАЦІЯ

В дипломній роботі на першому етапі була вивчена попередня текстова інтерактивна система виявлення вимог і на основі базових концепцій розробки цієї системи був розроблений та реалізований найбільш підходящий графічний інтерфейс як з точки зору семантики, так і з точки зору зручності використання. Крім того, було проведено дослідження зручності використання на групі студентів з різним академічним досвідом, щоб підтвердити, що запропонований дизайн може позитивно покращити зручність використання інтерфейсу користувача.

КЛЮЧОВІ СЛОВА: ПРОЦЕС ВІЗУАЛІЗАЦІЇ, ІНТЕРФЕЙС КОРИСТУВАЧА, ДЕТЕРМІНОВАНІ ВИЗНАЧЕННЯ, АРХІТЕКТУРА СИСТЕМИ, КОНЦЕПЦІЯ.

ANNOTATION

In the diploma thesis, at the first stage, the previous text-based interactive system for identifying requirements was studied, and on the basis of the basic concepts of the development of this system, the most suitable graphical interface was developed and implemented, both from the point of view of semantics and from the point of view of ease of use. In addition, a usability study was conducted on a group of students with different academic backgrounds to confirm that the proposed design can positively improve the usability of the user interface.

KEY WORDS: VISUALIZATION PROCESS, USER INTERFACE, DETERMINISTIC DEFINITIONS, SYSTEM ARCHITECTURE, CONCEPT.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ 7
ВСТУП 8
1. ОГЛЯД ІСНУЮЧИХ ПРОГРАМНИХ РІШЕНЬ ТА СЕРВІСІВ В ПРЕДМЕТНІЙ ОБЛАСТІ ВІЗУАЛІЗАЦІЇ 9
1.1. Детермінанти онлайн-переконання 9
1.2. Аналіз пов'язаних робіт за темою дипломного дослідження 11
1.4 Опис процесу застосування коментарів 18
1.5 Лінгвістичні властивості мови коментарів 22
1.6. Дослідження процесу зрозумілості написання коментарів 29
1.5. Опис процедури “переконаних користувачів” 37
2. АЛГОРИТМІЗАЦІЯ ПРОЦЕСУ ВИЛУЧЕННЯ ТА АНАЛІЗУ RST CUE 45
2.1. Основні теоретичні положення 45
2.2 Аналіз робіт по темі дипломної роботи 46
2.3. Сутність запропонованої методології 48
2.4. Результати попереднього імітаційного експерименту 50
2.5. Опис результатів моделювання 62
3. ПРОГРАМНЕ ТА АЛГОРИТМІЧНЕ РІШЕННЯ ЗАДАЧІ ВІЗУАЛІЗАЦІЇ ПРОГРАМНИХ СЕРВІСІВ 64
3.1. Семантичне обґрунтування розробки 64
3.2. Опис пропонованого підходу 69
3.4. ProjectTales: Доказ концепції 76
3.5. Дизайн інтерфейсу 79
ВИСНОВКИ 83
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ 85
БІБЛІОГРАФІЧНА ДОВІДКА

ПЕРЕЛІК СКОРОЧЕНЬ

ADO – ActiveX Database Objects

COM – Component Object Model

DCOM – Distributed Component Object Model

IEC - International Electrotechnical Commission

ISO – International Organization for Standardization

UML – Unified Modelling Language

VCL – Visual Component Library

ВСТУП

В процесі розробки програмного забезпечення необхідно скоротити вартість, зусилля на розробку програмних продуктів. Розмір і складність програмних систем збільшуються з кожним днем, і важливо, щоб розроблене програмне забезпечення було достатньо гнучким, щоб адаптуватися до змін ринку та нових технологій . Таким чином, покращення можливості повторного використання, а отже, і якості підвищить продуктивність команд програмного забезпечення та зменшить час і витрати на вихід на ринок нових програмних продуктів.

Збірка продукту з компонентів, зменшення трудомістких завдань за допомогою автоматизації настроювання ліній програмних продуктів і ланцюгів поставок, а також стандартизації інтерфейсів, архітектур і процесів можуть підвищити індустрію програмного забезпечення до набагато вищої продуктивності. Це можна зробити за допомогою програмної продуктової лінії, яка виробляє набір різних, але схожих продуктів, і сервіс-орієнтованої архітектури, які разом вирішують проблеми інтеграції та взаємодії. Обидва вони обіцяють розробку гнучких, економічно ефективних і багаторазових програмних систем.

З іншого боку, щоб зробити процес розробки програмного забезпечення менш складним і більш зрозумілим для розробників програмного забезпечення, важливо використовувати деякі інструменти для підтримки завдань, які включені в процес розробки програмного забезпечення, спеціально для найбільш обмеженої фази, яка є фазою виявлення вимог. Одним з можливих аспектів такої підтримки є візуалізація програмного забезпечення . Останнім часом доступно дуже багато методів та інструментів візуалізації програмного забезпечення, але важливо вибрати найбільш підходящий для відповідної діяльності в процесі розробки програмного забезпечення, щоб зробити найбільш ефективну візуалізацію для конкретної програмної системи.

1. ОГЛЯД ІСНУЮЧИХ ПРОГРАМНИХ РІШЕНЬ ТА СЕРВІСІВ В ПРЕДМЕТНІЙ ОБЛАСТІ ВІЗУАЛІЗАЦІЇ

1.1. Детермінанти онлайн-переконання

Наші повсякденні соціальні взаємодії часто включають складний процес переконання та впливу. Такі взаємодії варіюються від неформальних обстановок, таких як розмови в коридорі, до офіційних ситуацій, як-от ділові переговори та президентські виборчі кампанії. Тому вміння переконувати є ключовою навичкою, якою потрібно володіти, і вона може суттєво сприяти нашому соціальному житті. Починаючи з часів Аристотеля, у таких традиційних комунікаціях було проведено багато досліджень щодо переконання та переконливого дискурсу [23, 7]. Однак із появою, постійним розвитком і зростанням веб-сайтів соціальних мереж, значна частина нашої соціальної взаємодії та взаємодії відбувається через нові та передові форми спілкування. Згенеровані комунікаційні дані, переважно у формі текстових повідомлень та коментарів, забезпечують чудовий ресурс для розуміння механізмів онлайн-переконання та його подібності та відмінності від традиційних процесів переконання. Проте такі повідомлення не були достатньо використані для вивчення переконливих дій в середовищі онлайнового спілкування, в основному через відсутність позначених даних.

Нішева спільнота субредітів пропонує значний ресурс для вивчення онлайн-переконання, сприяння зміні думок за допомогою онлайн-коментарів і вимоги до учасників коментувати успішні коментарі. Цей субреддіт під назвою Change My View (CMV) призначений для користувачів, які мають свою думку на тему, але бажають прислухатися до різних голосів і змінити свою думку. Користувачі, які публікують свої погляди в CMV, відомі як оригінальні плакати (OPs) у спільноті.

Версія цієї глави була подана для публікації в ACM Transactions on Social Computing. 109 користувачів надають коментарі з метою змінити думку ОП. Правила CMV вимагають, щоб ОП позначав(і) коментар(и), які успішно змінили його/її погляд, і коротко пояснював, чому подання змінено. У попередніх роботах, присвячених переконуванню та тексту, переконання широко визначають як інтерактивний процес, за допомогою якого дане повідомлення змінює точку зору людини, змінюючи знання, переконання чи інтереси, які лежать в основі цієї точки зору [32]. Це визначення спирається на відповідну літературу в галузі грамотності, психології та філософії і застосовне до того, чого просять досягти користувачі CMV. Тому субреддіт CMV надає набір коментарів, які написані з метою переконання. Крім того, правило CMV щодо позначення коментарів дозволяє нам швидко розробити корпус переконання, оскільки коментарі вже опосередковано позначені ОП як переконливі або непереконливі.

Зосереджуючись на цьому корпусі, ми досліджуємо різні виміри переконливих коментарів у порівнянні з непереконливими. Спочатку ми вивчаємо текст коментарів щодо їх відповідності змісту оригінального допису. Через важливість часу та порядку в асинхронних комунікаціях між кількома учасниками, ми також враховуємо тимчасові атрибути коментарів та порядок їх введення. Згідно з нашими результатами, ці дві характеристики відіграють важливу роль в успіху коментаря щодо зміни переконань. Тому спочатку ми контролюємо ці дві змінні, а потім аналізуємо лінгвістичні показники, які не обов'язково доповнюють пропозиційний зміст коментаря, але мають на меті допомогти читачеві організувати, інтерпретувати та оцінити надану інформацію. Ці лінгвістичні особливості часто називають метадискурсом або не тематичним матеріалом і, як відомо, впливають на процес переконання [12, 27].

У цій роботі багато з таких лінгвістичних показників фіксуються у словнику LIWC, який надає психологічно значущі категорії слів [45]. Крім того, відомо, що якість написання та організація тексту полегшують розуміння та

пригадування аргументованого тексту, що згодом сприяє переконанню [14]. Тому ми проаналізували набір атрибутів, що стосуються узгодженості та витонченості написаного, таких як використання перехідних фраз, використання розділових знаків, а також лексичне різноманіття та перекриття в наших контрольованих наборах, і виявили відмінності між переконливими та не-переконливими. переконливі коментарі. Крім того, враховуючи, що текст, наданий із надійного джерела, швидше за все, спричинить переконання, ми беремо до уваги набір атрибутів користувача. Прогностична сила, яку має кожна з цих ознак, далі аналізується.

Нарешті, як згадувалося раніше, CMV вимагає від користувачів включати пояснення, чому і як коментар може змінити їх погляд. З метою усвідомлення передбачуваних причин зміни переконань ми провели попередній аналіз зібраних пояснень. Спочатку ми виділили окремі речення з пояснень, які ми вважали

1.2. Аналіз пов'язаних робіт за темою дипломного дослідження

Кілька напрямків дослідження пов’язані з цим дослідженням: роботи з виявлення та аналізу впливових користувачів у соціальних мережах, дослідження, зосереджені на виконанні запитів в Інтернеті, дослідження переконливих есе та оцінки есе, і, нарешті, спроби проаналізувати та знайти переконливий текст у онлайн-пости та коментарі.

Вплив у соцмережах. Очікується, що впливові користувачі будуть переконувати частіше, ніж маловпливові. Таким чином, унікальні атрибути в їхніх текстах можуть дати розуміння лінгвістичних показників переконливих текстів. У [8] вивчили лінгвістичні характеристики твітів користувачів і виявили, що впливовими користувачами, як правило, є особи, які частково висловлюють негативні настрої у своїх твітах.

В [5] висунули гіпотезу, що люди, що впливають, частіше братимуть участь у певній поведінці в розмові, як-от переконання, домовленості та розбіжності. Щоб виявити впливових осіб у розмовах у LiveJournal та Вікіпедії, вони спершу створили набір функцій, заснований на анотаціях людей щодо цих трьох розмовних моделей поведінки та діалогової моделі обговорень. Потім вони розробили класифікатори для виявлення впливових осіб у цих двох середовищах. Їхній найкращий класифікатор привів до оцінки F 59,3% для обговорень у Вікіпедії та F-оцінки 74,3% для веб-журналів LiveJournal.

Виконання запиту на базі Інтернету. Інший напрямок досліджень зосереджений на аналізі онлайн-запитів. Виконання запитів є центральним для багатьох соціальних платформ, таких як веб-сайти запитань і відповідей (наприклад, StackOverflow.com) і спільноти краудфандингу та філантропії (наприклад, Kickstarter.com). Щоб такі запити виконували інші учасники, запитувачу зазвичай потрібно використовувати стратегії переконання, переконуючи інших допомогти.

У дослідженні на краудфандинговій платформі [31] досліджуються мовні характеристики, які спонукали натовп фінансувати конкретні проекти. Проаналізувавши 9 мільйонів фраз, витягнутих з Kickstarter.com, вони виявили, що мова, яка використовується для опису проектів, має прогнозну силу, що становить приблизно 58% розбіжностей щодо успішного фінансування. Вони також проаналізували 59 інших функцій, зазвичай присутніх на краудфандингових веб-сайтах, таких як наявність відео, кількість коментарів і кількість оновлень, і оцінили їх передбачувану силу.

Такі веб-сайти, як Kickstarter.com, надають стимули для підвищення рівня внесків. Інший набір платформ дозволяє виконувати альтруїстичні запити. В [3] досліджували, що спонукає людей давати, коли вони не отримують відчутної винагороди натомість. Вивчивши текст запиту, вони виявили, що надання розповіді, яка чітко передає потреби, є важливим фактором успіху запитів. Крім того, подяка, а також взаємність підвищували

шанси на успіх. Нарешті, високий статус запитувача був пов’язаний із процесом виконання запиту. Цей набір функцій розширено в [22] шляхом вивчення центральної ролі та ролі запитувача, а також актуальних аспектів запиту. Центральність характеризує ступінь взаємодії користувачів у спільноті, а атрибут role відображає роль взаємодії користувачів у мережевих спільнотах. Додавши ці функції до функції, запропонованої в [3], вони покращили ефективність класифікації для прогнозування успіху невидимого запиту.

Оцінка та оцінка есе. Автоматизоване оцінювання есе є одним із основних застосувань дослідження обробки природної мови в освіті і часто включає методи оцінки та оцінки якості тексту. Дослідження оцінки есе проаналізували письмовий текст з різних вимірів, таких як швидке дотримання [5], узгодженість і витонченість мови [6, 13], технічна якість і релевантність [14], а також сила і якість аргументів [6, 3].

Наприклад, в [6] запропонували багатий набір функцій для вивчення та оцінювання аргументованих есе. Цей набір функцій включає функцію POS n-gram як синтаксичне узагальнення n-грам слів, а також семантичні кадри тексту як семантичне узагальнення. Крім того, вони дослідили використання перехідних фраз (наприклад, крім того і аналогічно) та кореференцій як ознаки зв’язності. Також досліджуються такі ознаки, пов’язані з аргументацією, як швидка згода та помилки аргументації.

Переконання в соціальних мережах. Була обмежена робота, зосереджена на автоматичному виявленні дій переконання в соціальних мережах. В [4] представив корпус дописів у блозі, які позначені спробами переконати в онлайн-взаємодії. Вони розробили класифікацію цих спроб на основі попереднього вивчення даних та літератури. Ця класифікація включала шість принципів впливу [11], дванадцять типів стратегій [29] для забезпечення поведінкової відповідності, а також деякі аргументаційні схеми [7], зокрема, аргументи.

1.3. Дослідження набору даних Reddit

Найближчим до нашої роботи є недавнє дослідження, проведене в [4], яка зосереджена на аналізі та виявленні переконливих коментарів від CMV. У CMV користувачі можуть коментувати вихідну публікацію або створити ланцюжок обговорення, відповідаючи на коментарі інших. Вивчаючи ці дискусії, автори виявили динаміку взаємодії, яка може вплинути на процес переконання, наприклад, порядок входу користувачів і ступінь обміну та обміну. Крім того, вони дослідили мовну взаємодію між ОП та коментарями, досліджуючи лексичне перекриття та подібність двох текстів. Їхня спроба передбачити, який із двох коментарів вдасться змінити переконання, призвела до точності приблизно 65% для основної відповіді та 70% при розгляді повного шляху обговорення. Крім того, вони дослідили набір мовних особливостей з оригінальних публікацій і знайшли невеликий набір атрибутів, які можуть сигналізувати про те, чи є ОР податливим чи стійким до переконання.

Наша робота ґрунтується на цих напрямках дослідження та ґрунтується на літературі соціальної психології про зміну переконань та переконання. На відміну від [4], ми зосередилися виключно на кореневих коментарях, але провели більш всебічне та поглиблене вивчення таких коментарів. Наш експериментальний контроль є більш надійним, оскільки ми включаємо часовий аспект коментарів, окрім відповідності вихідній публікації. За винятком кількох мовних особливостей, які є спільними між двома роботами, ми вводимо новий набір змінних, які досліджують психологічні аспекти мови, а також читабельність, згуртованість та витонченість тексту. Крім того, наші методи вимірювання певних атрибутів (наприклад, релевантності та схожості вмісту) відрізняються від їхніх. Наша модель прогнозування покращує ефективність їхнього класифікатора у визначенні переконливих коментарів на 10%. Нарешті, ми вивчили пояснення користувачів, щоб отримати уявлення про переконливі фактори, які сприймають самі користувачі. Потім ми порівняли

передбачувані причини з ознаками, знайденими під час прямого аналізу та порівняння переконливих і непереконливих коментарів.

Change My View (CMV) — це активний субреддит із понад 23 тисячами передплатників. Користувачі CMV висловлюють свої переконання на тему, відкриті до різних поглядів і змінять свої переконання, якщо їх переконають коментарі інших. Згідно з правилами CMV, оригінальний пост має пояснювати аргументи, що стоять за поданням, а не лише те, що таке подання. Тому пости мають бути не менше 500 символів. Крім того, правила CMV чітко вказують, що ОП «повинен особисто дотримуватися точки зору та бути готовим до її зміни».

Після того, як представлення опубліковано в CMV, інші користувачі можуть надати аргументи та аргументацію проти початковий погляд ОП, відображений у цій публікації. Потім постер може прийняти твердження коментарів і змінити свої переконання, ігнорувати їх або брати участь у дискусії у формі ланцюжків розмов і відкласти дії, доки він/вона не переконається в цьому. Коли коментар успішно змінює погляд постера, він/вона зобов'язаний відповісти на цей коментар. Ця відповідь повинна включати символ дельта (Δ), який вказує на зміну переконань і повинен давати пояснення, чому і як коментар змінив його/її погляд. CMV може функціонувати як контрольована експериментальна установка для досліджень переконання, оскільки вона контролює набір змінних, пов'язаних із переконанням, таких як переконання, яке має змінитися, і атрибути користувача, який дотримується переконань. Зокрема, оскільки ця платформа контролює особисту релевантність переконання, яке виявилося дуже впливовим у процесі переконання [8]. Тим часом CMV допускає варіації іншого набору змінних, включаючи атрибути користувачів, які публікують коментарі, а також різні типи мови та стилів письма, які використовуються в коментарях. CMV ретельно модерується, контролює будь-які дописи чи коментарі, які не відповідають правилам.

Дельта-механізми CMV разом із його строгими правилами надають унікальний і цінний набір даних для вивчення онлайн-переконання.

На рисунку 1.1 показано приклад публікації, надісланої до CMV, разом із трьома вибраними коментарями, які надаються для зміни початкового вигляду. У верхній частині першого блоку показано назву публікації разом із часом подання та ім’ям користувача ОП. Нижня частина - це погляд і міркування ОП. Наступні три блоки із суцільною рамкою є коментарями, наданими для зміни вигляду. Як видно, кожен коментар пов’язаний з ім’ям користувача коментатора та загальним дельта-рахунком (тобто, скільки разів цей користувач отримував дельту на даний момент), а також оцінкою коментаря (тобто кількістю голосів «за», вирахуваною на кількість голоси проти коментаря, отриманого на даний момент), і позначку часу подання коментаря. Нарешті, блок із пунктирною межею є відповіддю на перший коментар із символом дельта (Δ) та обґрунтуванням надання дельти (Δ) коментарю. Слід зазначити, що метадані, представлені в прикладі, та їх розміщення збігаються з фактичним інтерфейсом CMV.

Як уже згадувалося, коментарі CMV можуть відбуватися у формі ланцюжкових обговорень, де користувачі CMV можуть відповідати на інші коментарі та будувати дерево обговорень. Дельта може бути присуджена коментарям на різних рівнях p дерева. Однак у нашому дослідженні ми зосередилися лише на коментарях верхнього рівня, які є прямими відповідями на публікацію і тому служать основним коментарем у кожній темі. Зосереджуючись лише на кореневих коментарях, ми можемо переконатися, що переконливі аспекти лежать у самому коментарі, але не в контексті розмови, який відбувається через потік коментарів. Враховуючи, що ми зосереджені на кореневих коментарях, відтепер, коли ми говоримо коментар, ми посилаємося на кореневий коментар. Крім того, під переконливими або успішними коментарями ми посилаємось на коментарі, які виграли дельту.

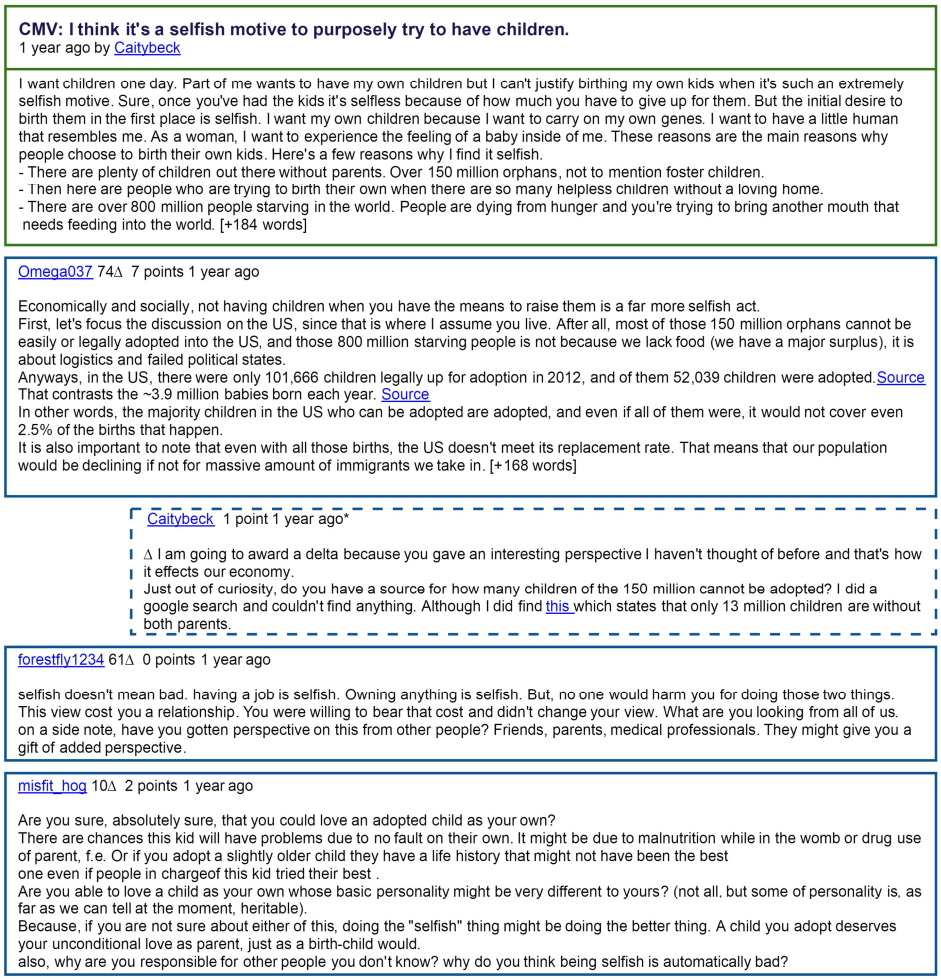


Рисунок 1.1 - Приклад публікації CMV разом із коментарями, наданими для зміни вигляду

За певний проміжок ми збирали повідомлення та коментарі до них від CMV. У Reddit, коли користувач або текстовий вміст коментаря чи публікації більше не доступні, відповідному полю даних призначається значення [видалено]. Ми повідомляємо статистику даних після очищення таких точок даних. Крім того, лише невелика частина (менше 5%) спроб переконати за допомогою кореневих коментарів сприймаються як переконливі, і тому оцінюються. Щоб достовірно порівняти переконливі та непереконливі коментарі, нам потрібно контролювати набір змінних. Одним із найважливіших факторів, про які слід подбати, є публікація, до якої надаються коментарі, оскільки тема публікації та характеристики ОП можуть вплинути на процес

переконання. Тому ми зосередилися лише на публікаціях, які пов’язані хоча б з одним успішним коментарем. Крім того, для порівняння ми відфільтрували цей набір, щоб включати лише повідомлення, які мають принаймні 5 коментарів. Цей процес фільтрації привів до 1249 дописів і 18534 коментарів, пов’язаних з цими публікаціями, серед яких 1690 коментарів отримали дельту.

1.4 Опис процесу застосування коментарів

Відповідність оригінальній публікації

Одним з основних показників, які ми вважали доречними, є ступінь тематики коментарів і пов’язаність із публікацією, наданою ОР. Щоб обчислити релевантність коментарів до основної публікації, ми використали стандартну техніку пошуку інформації tf-idf, яка широко використовується для отримання та ранжирування документів за запитом. tf-idf визначає важливість даного слова запиту для документа на основі частоти слова в документі (tf), а також розподілу слова в базовому корпусі (idf). На основі tfidf словам, які частіше зустрічаються в документі, але рідко зустрічаються в корпусі, призначаються більші значення.

Щоб визначити релевантність коментаря до публікації, ми спочатку провели все необхідне етапи очищення та попередньої обробки як для публікацій, так і для коментарів, наприклад перетворення в нижній регістр, заміна смайликів і контроль вилучення. Далі ми розглядали колекцію дописів як основний корпус, а кожен коментар — як запит. Релевантність коментаря до публікації (тобто документа) потім обчислюється як сума балів tf-idf слів у коментарі.

Оскільки необроблена оцінка коментарів tf-idf залежить від особливостей основної публікації, пряме порівняння показників релевантності може бути ненадійним підходом. Замість цього ми окремо проаналізували кожен набір коментарів, наданих для публікації, та вивчили, як переконливі

коментарі оцінюються порівняно з непереконливими на основі їхньої оцінки tf-idf. Зосередження на первинному рейтингу коментарів призведе до упередження для наборів коментарів різного розміру. Замість цього ми обчислюємо відсоток непереконливих коментарів, які мають вищий бал за переконливі, використовуючи наступне:

$$Relevance\ Grade = \frac{Rank_{\Delta} - 1}{N} * 100 \quad (1.1)$$

RankΔ — це необроблений ранг переконливого коментаря, коли коментарі відсортовані від найвищого до найнижчого балу tf-idf. N – загальна кількість коментарів у наборі коментарів. Наприклад, розглянемо публікацію з 20 кореневими коментарями. Якщо переконливий коментар має другий найвищий бал tf-idf у наборі, йому буде присвоєно оцінку релевантності 5%. Цей показник означає, що 5% коментарів розміщуються над переконливим коментарем щодо їхньої оцінки tf-idf. На рисунку 1.2 показано розподіл оцінки релевантності для різних переконливих коментарів. Сюжет чітко зображує низхідну схему, що вказує на те, що найбільш релевантні коментарі в наборі мають вищі шанси отримати дельту. Вертикальна червона лінія показує середнє значення і означає, що в середньому переконливі коментарі знаходяться в межах 25% найвищих коментарів відповідно до їхньої релевантності. Можна стверджувати, що результати tfidf є упередженими і віддають перевагу довшим коментарям. Хоча це твердження загалом вірне, для нашого набору даних немає кореляції між оцінкою релевантності коментаря та кількістю слів (Пірсон r=0,002). Однак варто зазначити, що необроблений бал tf-idf корелює з кількістю слів (Пірсон r=0,65). Це показує ще одну перевагу використання оцінки релевантності замість вихідного рангу або необробленої оцінки tf-idf.

Щоб проаналізувати взаємодію коментарів і дописів, використовували показники подібності, засновані на накладанні слів. У результаті вони виявили,

що коментарі, які відрізняються від оригінальної публікації з точки зору використання слів у змісті, мають вищі шанси на успіх. Однак цей показник не враховує той факт, що деякі слова можуть бути поширеними в цій конкретній спільноті і не повинні надавати однакове значення у разі збігу у зв'язку з рідкісними словами.

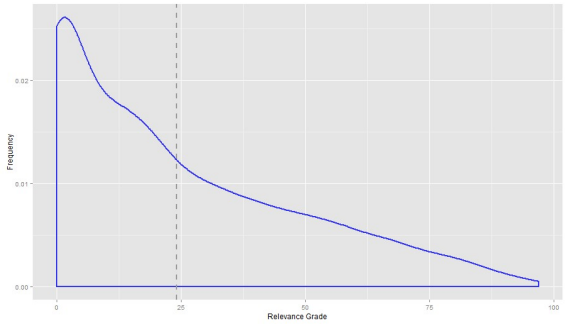


Рисунок 1.2 - Щільність розподілу оцінок релевантності для переконливих коментарів

Наприклад, виходячи з нашого розрахунку IDF, слова люди та вірять часто з'являються в публікаціях CMV загалом, тому їхнє накладання не обов'язково вказує на сильну тематичну схожість. Незважаючи на те, що загальна несхожість може позитивно вплинути на переконання, наші результати показують, що дуже важливо використовувати слова, специфічні для точки зору. Цей висновок узгоджується з попередніми дослідженнями переконання, які стверджують, що відповідність змісту переконливого повідомлення функціональній основі ставлення людей підвищує шанси переконання.

Терміни та порядок

Як і більшість соціальних платформ для запитань і відповідей та обміну новинами, Reddit насамперед сприяє асинхронному спілкуванню. Таким чином, час і порядок вступу можуть вплинути на процес переконання. Спочатку ми дослідили часовий лаг між поданням основної публікації та поданням

коментарів. Результати показують, що більшість коментарів надається досить швидко після розміщення основного перегляду, причому понад 90% коментарів надходять протягом доби після подання публікації. Іншими словами, час для всіх коментарів у ланцюжку однаковий.

Тому замість часового лагу ми проаналізували порядок введення коментарів. Враховуючи публікацію, ми обчислювали тимчасову оцінку для переконливих коментарів так само, як і оцінку релевантності. Кожному переконливому коментарю надається значення, яке є відсотком непереконливих коментарів, які були надіслані до публікації до коментаря у фокусі. На рисунку 1.3 показано розподіл часової оцінки між переконливими коментарями. Незважаючи на те, що частота падає із збільшенням тимчасової оцінки, зміна не є ні настільки стійкою, ні настільки різкою, як оцінка релевантності. Зокрема, спостерігається відносно значне зростання, оскільки ми переходимо від перших 10% коментарів до тих, які становлять 10-20%. Однак загалом частота оцінок падає зі збільшенням тимчасової оцінки, що вказує на те, що коментар, надісланий після більшого набору коментарів, має менший шанс для отримання дельти.

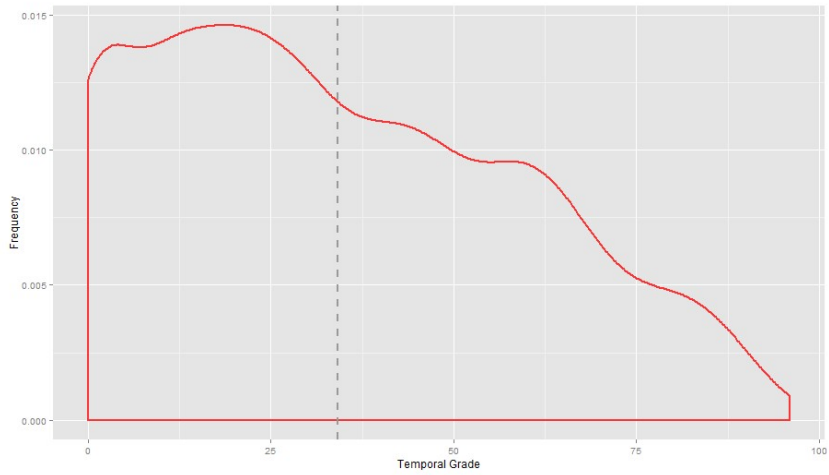


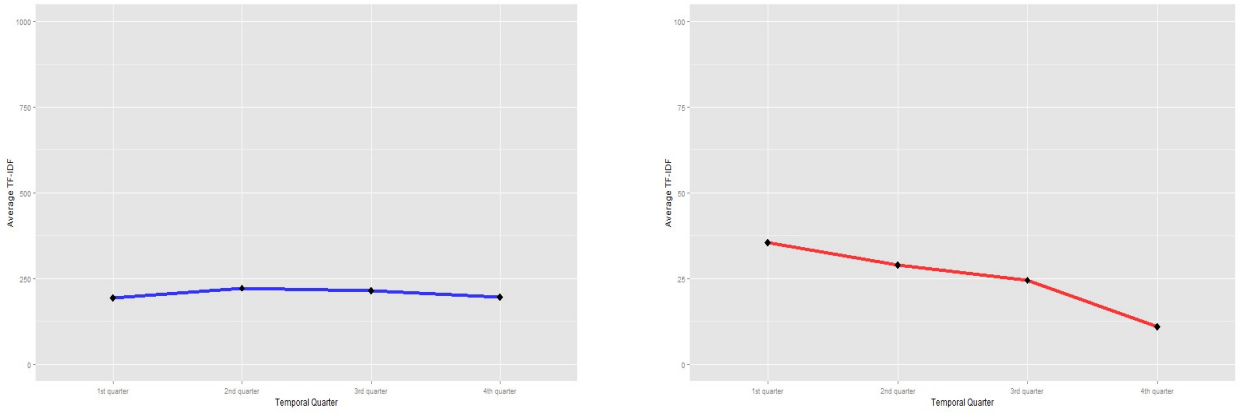
Рисунок 1.3 - Щільність розподілу часових оцінок для переконливих коментарів

Середнє значення для тимчасового рівня становить 34%, як показано вертикальною червоною лінією. Можна стверджувати, що ця тенденція пов’язана з можливою зміною актуальності в міру подання нових коментарів. Оскільки більше користувачів коментують публікацію, може відбутися зміщення теми через потенційний вплив попередніх коментарів. Щоб перевірити цю гіпотезу, ми поквартували всі наші набори коментарів на основі їх тимчасового порядку та обчислили середній бал tf-idf у кожному кварталі. Як показано на рисунку 1.4 (a), зміни релевантності є незначними, що переконує аргумент, що релевантність може бути прихованою змінною, що впливає на наші тимчасові результати. Однак, як випливає з рисунка 1.4 (b), відсоток коментарів, які сприймаються як переконливі, постійно зменшується, коли ми переходимо від одного кварталу до іншого. Кілька можливих причин можуть викликати цей ефект. Наприклад, ця зміна може бути пов’язана з ефектом зворотного ефекту [33], який свідчить про те, що переконання можуть стати сильнішими перед обличчям суперечливих доказів. Збільшення кількості суперечливих коментарів і аргументів може збільшити ймовірність зворотного ефекту. Потрібні подальші розслідування, щоб підтвердити цю заяву в контексті ЦМВ. Крім того, користувачі можуть втратити віру в спроможність спільноти, побачивши кілька невдалих коментарів, і, таким чином, можуть не так уважно стежити за майбутніми коментарями. Нарешті, ті, хто вже змінив свої погляди на попередній коментар, можуть бути менш мотивованими або навіть, можливо, неохоче стежити за майбутніми коментарями.

1.5 Лінгвістичні властивості мови коментарів

Згідно з попереднім аналізом, як релевантність, так і часові аспекти коментарів можуть вплинути на те, чи сприйматимуться вони як переконливі. Тому, щоб проаналізувати потенційні відмінності у використанні не тематичних лінгвістичних індикаторів та якості написання коментарів,

необхідно спочатку контролювати ці дві змінні. Для кожного переконливого коментаря ми вибираємо непереконливий коментар, який 1) найбільш схожий на переконливий за оцінкою релевантності та 2) передує переконливому. Додавши тем ми гарантуємо, що ОП побачив непереконливих, але не змінив його/її переконання. Нехтуючи цим аспектом, коментар може не отримати дельту просто тому, що він міг бути пропущений ОП. Коли переконливе повідомлення є найпершим коментарем, ми вибираємо найбільш схожий коментар з точки зору релевантності з наступних трьох надісланих коментарів. Дотримання цього методу призвело до вибору набору непереконливих коментарів того ж розміру, що й переконливі, що дозволило подальше порівняння та аналіз кореневих коментарів.



(a) (b)

Рисунок 1.4 - Середня оцінка релевантності в кожному квартальному періоді

На рисунку 1.4 a) є подібною, b) відсоток переконливих коментарів постійно зменшується, коли ми переходимо від однієї чверті до наступної.

Ми використовували LIWC як один із наших інструментів лінгвістичного аналізу, оскільки він здатний надати широкий спектр соціальних і психологічних ідей з мови. Програма LIWC має два центральні компоненти, а саме компонент обробки та словники. Компонент обробки обробляє текстовий

файл слово за словом. Потім кожне слово порівнюється зі словниками, щоб визначити, до якої категорії LIWC належить слово. Після обробки всіх слів у тексті LIWC обчислює відсоток кожної категорії LIWC. Категорії LIWC належать до чотирьох основних процесів: мовні процеси (наприклад, прийменники, займенники та допоміжні дієслова), психологічні процеси (наприклад, сім'я, тривога та здоров'я), особисті проблеми (наприклад, смерть, робота та досягнення) та розмовні категорії (наприклад, погодження, невільні знання та фільтри). Визначення та приклади кожної з цих категорій можна знайти на веб-сайті LIWC.

Крім того, набір змінних LIWC вимірюється незалежно від словників. Використання різноманітних розділових знаків досліджується в тексті, обчислюючи їх відсоток використання. Крім них, враховуються чотири непрозорі мовні змінні (аналітичне мислення, вплив, автентичність та емоційний тон) та загальні описові змінні (слів у реченні та відсоток слів, які довші за шість літер). Хоча більшість змінних LIWC проаналізовано тут, ті, які можуть значно покращити якість написання (наприклад, розділові знаки), включені в наступний розділ разом із нещодавно введеним набором функцій.

Щоб використовувати інструмент LIWC, ми розглядали кожен коментар у колекції як окремий вхід. Потім ми проаналізували кожну категорію LIWC за допомогою двостороннього t-критерію з двома зразками. Атрибути та ознаки, для яких відмінності є статистично значущими, представлені в таблиці 7.2. Різниця також є практично суттєвою для змінних, які позначені зірочкою (d Коена > 0,2).

Згідно з посібником LIWC, високе значення змінної аналітичного мислення відображає формальне, логічне та ієрархічне мислення, тоді як менші цифри відображають більш неформальне, особисте та розповідне мислення. Автентична категорія фіксує ступінь, до якого мова є більш чесною, особистою та відкритішою. Риторика Аристотеля описує три способи переконання: логос, пафос і етос. Логос — це логічна привабливість акту переконання, яку можна

охарактеризувати за допомогою оцінки аналітичного мислення LIWC. Як і очікувалося, переконливі коментарі мають вищу оцінку аналітичного мислення (як статистично, так і практично значуще), що підтверджує, що користувачі часто цінують аргументи, засновані на фактах, логіці та аналітиці. Пафос відноситься до вчинків переконання, які мають на меті сколихнути почуття та емоції аудиторії. Згідно з аналізом, переконливі коментарі оцінюються нижче за автентичність, що вказує на те, що логотипи цінуються більше в онлайн-комунікації, ніж пафос.

Емоційний тон фіксує рівень настроїв мови. Велике число асоціюється з більш позитивним стилем, тоді як низьке число показує більшу тривогу, смуток або ворожість. Число близько 50 говорить про відсутність емоційності або різні рівні амбівалентності. Незважаючи на те, що емоційний тон є нижчим для переконливих коментарів, середні значення для обох наборів коментарів близькі до 50, що ускладнює інтерпретацію результатів. Крім того, не було помічено різниці щодо використання позитивних і негативних слів, записаних у словнику LIWC.

Тому, щоб зробити надійні висновки щодо функцій, пов’язаних із настроями та переконання в онлайн-тексту, можуть знадобитися більш складні інструменти аналізу настроїв. Більш ранні роботи щодо переконання стверджують, що повідомлення, які відповідають емоційному стану одержувача, мають вищі шанси переконання [15]. Таким чином, замість аналізу настроїв коментарів окремо, вивчення взаємодії настроїв основної публікації та коментарів може дати нам уявлення про потенційні зв’язки переконання та настроїв в онлайн-спільнотах.

Хоча ми аналізували вміст коментарів на основі їх відповідності до основної публікації, лінгвістичні процеси LIWC дозволяють перевіряти функціональні слова, також відомі як стиль словами, в тексті.

Таблиця 1.2 - Т-тест Уелча для двох груп коментарів

Підсумкові змінні Welch T-Test P-value	Підсумкові змінні Welch T-Test P-value
аналітичне мислення*	+++
Автентичний	- -
Емоційний тон	- -
Мовні процеси	
Функціональні слова	- - -
Усього займенників	- - -
Особисті займенники	- -
1 особа однини (I)	- -
2 особа (ви)	- -
Безособові займенники	- -
Прийменники	- -
Допоміжні дієслова*	- -
Заперечення	- -
Питальні	- -
Числа	+++
Психологічні процеси	
Когнітивні процеси	- - -
Причина	- -
Розбіжності	- - -
Необережність	- -
Диференціація	- - -
Постійні процеси	- -
Орієнтація в часі	
Минулий фокус	++
Теперішній фокус	- - -

P < 0,005, коли значення є більшими для переконливих коментарів, тоді як - - і - - - використовуються, коли значення менші для переконливих коментарів. Змінні з невеликим розміром ефекту (d Коена > 0,2) позначені зірочкою.

З психологічної та соціальної точки зору слова стилю відображають те, як люди спілкуються, тоді як змістові слова передають те, що вони говорять. Тому функціональні слова набагато тісніше пов’язані з мірками соціального та

психологічного світу письменників і читачів. Ми виявили статистично значущі відмінності між двома групами коментарів у використанні функціональних слів. До переконливих коментарів, зокрема, відноситься менше займенників (як особових, так і безособових), прийменників і допоміжних дієслів.

У коментарях, які не змогли переконати, частіше з'являються заперечення. Згідно з попередніми психолінгвістичними дослідженнями, негативні граматичні трансформації складніші, ніж позитивні, і вони часто потребують більшого часу та більше когнітивних ресурсів для обробки [19]. Питальні, які часто використовуються при постановці питань, також частіше спостерігаються в непереконливих повідомленнях. З іншого боку, цифри частіше з'являються в переконливих коментарях, які можуть бути пов'язані з посиланнями на кількісні докази та статистичні дані. Вважається, що використання кількісної оцінки замість описових модифікаторів забезпечує цілісність комунікації через достовірність, пов'язану з числами [18]. Незважаючи на те, що деякі попередні дослідження продемонстрували негативний вплив кількісної оцінки на переконання, коли джерело має низьку довіру (переважно реклама) [21], кількісна оцінка позитивно пов'язана з переконанням у CMV, де довіра джерела в основному невідома.

Слова, що представляють когнітивну обробку, здається, менше з'являються в переконливих коментарях. Зокрема, рідше вживаються в успішних коментарях слова, пов'язані з розбіжністю (наприклад, слід і хотів би) та орієнтовними (наприклад, можливо і можливо). Висновки, пов'язані з цими двома категоріями, можна пояснити попередніми дослідженнями переконання з двох різних точок зору. По-перше, ці дві категорії демонструють вагання в мові, що пов'язано з безсилим стилем спілкування. Безсила мова може негативно вплинути на судження читачів про довіру до автора і згодом вплинути на процес переконання або може мати прямий негативний вплив на переконання [21, 22]. По-друге, на переконання може вплинути те, наскільки ясно або нечітко подано повідомлення. Деякі роботи виявили, що розпливчаста

мова негативно пов’язана з переконанням інших, тоді як деякі виявили стратегічну двозначність, яка підвищила шанси переконання [21]. Причина останнього висновку полягає в тому, що одержувач повідомлення не міг так легко відхилити нечіткі повідомлення. Ця причина, однак, непридатна в контексті CMV, коли ОП особисто дотримується точки зору, і мета полягає в тому, щоб змінити цю точку зору. Отже, очікується, що неоднозначність повідомлення матиме негативний вплив на CMV. Слова, пов’язані з диференціацією (наприклад, без і незважаючи), також частіше зустрічаються в невдалих коментарях, що може бути пов’язано з додатковою синтаксичною складністю використання негативної граматики. На додаток до потенційних наслідків синтаксичної складності, риторичні стратегії, пов’язані з цими словами, можуть значно вплинути на процес переконання.

Аналогічно, процеси сприйняття (тобто процеси, пов’язані з баченням, слухом і відчуттями) використовуються рідше в переконливих коментарях. Цей висновок також може бути пов’язаний з домінуючою силою логотипів у порівнянні з пафосом на онлайн-платформах, оскільки процеси сприйняття часто пов’язані з описом особистого досвіду та наративів у порівнянні з аналітичним і логічним текстом. Нарешті, минулий час частіше використовується в переконливих коментарях, тоді як теперішній час частіше використовується в непереконливих. Наше спостереження за коментарями вказує на те, що минулий час вживається в основному, коли наводяться докази та приклади з минулого.

Незважаючи на те, що наявність доказів і фактичної інформації може безпосередньо призвести до зміни думки, надання таких історичних випадків також може сприяти створенню словесних образів. Показано, що здатність слів викликати у читачів образи позитивно впливає на переконання порівняно з абстрактною мовою [9, 8]. Однак теперішній час також може запропонувати словесні образи. Потрібні подальші експерименти, щоб підтвердити, що саме

словесний образний аспект цих історичних випадків призводить до переконання.

1.6. Дослідження процесу зрозумілості написання коментарів

Відомо, що довіра до джерела є одним із головних впливових факторів у процесі переконання [2]. Варіації в мові можуть вплинути на судження читачів про достовірність і привабливість автора [21]. Тому ми аналізуємо коментарі з точки зору витонченості їх мови. Витонченість мови можна виміряти за довжиною тексту [13], лексичною складністю та різноманітністю [13, 20], а також за наявністю розмовних та неформальних мовних елементів. Крім того, добре написаний текст покращує запам’ятовування та розуміння, дозволяючи читачам глибоко обробляти текст і збільшуючи шанси на розуміння та переконання [17]. Правильне використання розділових знаків і використання зв’язної мови може сприяти зрозумілості. Згуртованість може характеризуватися використанням сполучників і перехідних фраз [13], а також лексичним перекриттям між послідовними реченнями [13, 20].

Згідно з результатами, спонукальні повідомлення, як правило, довші та містять довші речення. Крім того, вони частіше включають складні слова (тобто слова, що містять більше шести літер). Крім того, переконливі коментарі різноманітніші за лексичним вибором. Ми обчислили лексичне різноманіття, розділивши кількість унікальних слів (типів) на загальну кількість слів (лексем) у коментарі, метрику, відому як співвідношення тип-токен. Нижчий показник різноманітності вказує на те, що мова є відносно надлишковою, тоді як більша цифра вказує на різноманітний і багатий словниковий запас. Відомо, що лексична різноманітність і складність впливають на судження читачів про автора через принцип переваги складності [21, 6]. Пізніше цей висновок підтвердився, оскільки лексична різноманітність і витонченість безпосередньо пов’язані з судженнями читачів про компетенцію та соціально-економічний

статус автора [21, 6]. Крім того, неформальні категорії слів, включаючи netspeak (наприклад, plz і ppl), згоду (наприклад, так і добре) і невільні слова (наприклад, mmm і er), часто асоціюються з невдалими коментарями. Ці особливості не тільки пов'язані з недостатньою витонченістю мови, але й використання невільних мов також є ознакою вагання та безсилого стилю мови, що призводить до негативного впливу на переконання.

Далі ми розрахували оцінку читаності за Флеша-Кінкейда та рівень оцінки [30]. Оцінка читаності Флеша-Кінкейда вказує на те, наскільки важко зрозуміти уривок для читання на основі кількості речень, слів і складів. Вищі бали вказують на те, що текст легше читати, тоді як нижчі числа позначають уривки, які важче читати. Рівень читаності Flesch-Kincaid – це перетворення оцінки читабельності, яка безпосередньо представляє оцінку як рівень оцінки в США, необхідний для розуміння тексту. Можна очікувати, що коментарі, які легше читати, будуть переконувати частіше, оскільки їх легко зрозуміти. З іншого боку, складні коментарі можуть бути пов'язані з довірою та рівнем освіти автора і можуть частіше призводити до переконання. Наші результати, пов'язані з читабельністю, відповідають попереднім висновкам про співвідношення складної мови та ефективності повідомлення. Загалом, результати вказують на те, що складність і витонченість мови впливають на судження користувачів про довіру, статус і навіть рівень освіти автора. Таке сприйняття може бути ще більш важливим для онлайн-переконання, коли інформації про автора мало або зовсім немає, а мова часто є основним відображенням довіри до автора.

Одним із основних лінгвістичних засобів, що сприяє розумінню та зрозумілості тексту, є використання сполучників і перехідних фраз. Тому ми проаналізували частоту перехідних фраз у переконливих і непереконливих коментарях. Набір використовуваних перехідних фраз складений Навчальними посібниками та стратегіями і включає 14 різних типів.

Таблиця 1.2 - Складність Welch T-Test P-value

Підрахунок слів**	+++
Слово в реченні	+++
Слова>6 букв	+++
Лексичне різноманіття	++
Netspeak	- -
Згода	- -
Неплинності	- -
Читабельність	
Показник читабельності*	- - -
Рівень читаності	+++
Згуртованість	
Напрямні фрази	+++
Фрази з наголосом	++
Наведіть приклади фраз	+++
Узагальнення фраз	+++
Лексичне перекриття	- -
Розділові знаки	
Усі розділові знаки	+++
Коми	++
Двокишкові	+++
Знаки питання	- - -
Тире	+++
дужки (пари)	++
Інші розділові знаки*	+++

Т-тест Уелча для двох груп коментарів. ++ для P < .05 і +++ для P < .005, коли значення більші для переконливих коментарів, тоді як - - і - - - використовуються, коли значення менші для переконливих коментарів. * і **, розміщені поруч із іменами змінних, вказують на малий розмір ефекту (d>Cohen’s>0,2) і помірний розмір ефекту (Cohens’d>0,5), відповідно.

Наш аналіз показує, що чотири типи перехідних фраз зустрічалися в переконливих коментарях частіше, ніж непереконливі: фрази напряму (наприклад, за межами і майже), фрази з наголосом (наприклад, перш за все та зокрема), фрази-екземпляри (наприклад, у т.ч. і такі як), і узагальнюючі фрази

(наприклад, все-таки і на завершення). На додаток до внеску цих фраз у згуртованість мови, ці слова можуть бути пов’язані з набором риторичних стратегій, які впливають на процес переконання. Наприклад, узагальнення фраз може допомогти у розумінні тексту, а приклади фраз можуть привести до словесних образів.

Іншим показником для оцінки згуртованості мови є лексичне перекриття між сусідніми реченнями. Більший ступінь накладання між сусідніми реченнями вказує на більш зв’язну мову. Щоб обчислити лексичне перекриття між двома реченнями, ми спочатку лексемизували іменники зі основою, знайдені в двох реченнях. Потім, щоб врахувати довжину речень, ми використали метрику подібності Жаккара між двома наборами іменників і виміряли середню подібність між реченнями коментаря:

$$Lexical\ Overlap = \frac{\sum_{N\,i=1}^{N-1} |S_i \cap S_{i+1}|}{|S_i \cup S_{i+1}|} \qquad (1.2)$$

де N — загальна кількість речень у коментарях, а S і — набір іменників із основою і-го речень коментаря. Цікаво, що непереконливі коментарі демонструють більший ступінь збігу між реченнями. Цей несподіваний результат можна пояснити різницею в довжині переконливих і непереконливих коментарів. Враховуючи, що переконливі загалом довші, вони можуть представляти більше аргументів з різних точок зору, що призводить до меншого загального лексичного перекриття. Тому ми відфільтрували набори коментарів, щоб включити лише коментарі, які мають більше 10 речень, і повторно запустили експеримент. Як і очікувалося, істотної різниці в двох наборах не спостерігалося.

Наші результати також показують, що переконливі коментарі, як правило, мають більше розділових знаків. Враховуючи, що дужки, двокрапки та тире зазвичай використовуються як лінгвістичні стратегії для надання

додаткової інформації, пов’язаної з ядром тексту, можливо, що переконливі коментарі передбачають більше розробок, пояснень та роз’яснень, що робить їх більш переконливими. На відміну від усіх розділових знаків, переконливі коментарі, як правило, містять менше знаків питання. Крім того, наше спостереження за підмножиною коментарів зі знаками питання показує, що те, що слідує за запитаннями, схоже, також відрізняється в двох групах. У групі переконливих коментарів особа, яка поставила запитання, часто надавала пояснення та/або відповіді одразу після запитання, тоді як у групі непереконливих коментарів користувач часто не надавав додаткової інформації. Наприклад, перший пункт нижче є частиною коментаря, який отримав дельту на рисунку 1.1 (ця частина прихована на малюнку через обмеженість простору), а другий — із невдалого коментаря (останнє речення останнього коментар на рисунку 1.1):

• «Чому це важливо? Тому що наші економічні і особливо наші соціальні системи залежать від нових продуктивних працівників, які входять до робочої сили, коли старі виходять на пенсію. Без них економіка впаде, соціальні програми збанкрутують, і ми, по суті, переживатимемо тривалу депресію, поки рівень населення не стабілізується або не підвищиться».

• «Також чому ви несете відповідальність за інших людей, яких ви не знаєте? Чому ти думаєш, що бути егоїстом автоматично погано?»

Окрім мови, яка використовується в коментарях, дельта-показник користувача може служити показником довіри користувачів у спільноті. Дельта-показник користувача відображає загальну кількість разів, коли він/вона успішно змінював переконання ОР, і відображається в інтерфейсі CMV як частина метаданих коментарів (див. Рисунок 7.1), що потенційно впливає на те, як інші оцінюють коментар. Ми дослідили дельта-показники, пов’язані з користувачами в двох групах коментарів, і виявили статистично значущі відмінності ($P < 0,05$).

На додаток до дельта-показника, характерного для CMV, користувачі Reddit пов’язані з двома іншими показниками, пов’язаними з їхньою історією в спільноті Reddit, а саме з кармою посилань і кармою коментарів. Карма посилань — це загальна кількість балів, яку наразі отримали публікації, надіслані основним користувачем у різних субредитах. Бали розраховуються на основі загальної кількості голосів «за», відрахованої на загальну кількість голосів «проти». Аналогічно, карма коментарів відноситься до загальної кількості балів, які коментарі, опубліковані основним користувачем, отримали в різних субредитах. Незважаючи на те, що загальна кількість балів, отриманих коментарем, відображається в інтерфейсі CMV, інформація про посилання та карму коментарів автора не доступна безпосередньо (див. Рисунок 7.1). Тому ми очікуємо, що ці дві функції мало впливатимуть на успіх коментаря. Коефіцієнти кореляції оцінки дельта та карми зв’язку (Pearson r=0,002) та дельта-оцінки та карми коментарів (Pearson r=0,11) користувачів підтвердили наші очікування.

Оцінка прогнозу

Аналіз переконливих і непереконливих коментарів свідчить про наявність певних ознак, характерних для кожної групи коментарів. Щоб вивчити прогнозну силу цих ознак, ми провели кілька експериментів. Для кожного переконливого коментаря ми вибирали випадковим чином непереконливий з того самого ряду аргументів, наданих для зміни точки зору, що призводить до збалансованого набору коментарів. Потім ми використали алгоритми навчання під наглядом (використовуючи Weka), щоб передбачити, чи отримає невидимий коментар дельту. Ми оцінили їх за допомогою 10-кратної перехресної перевірки. Ці моделі вибираються з різних категорій алгоритмів (наприклад, байєсівських, на основі дерева та на основі регресії) і побудовані на 49 функціях, які пов’язані з релевантністю коментарів, порядком їх введення, а також наявністю лінгвістичних індикатори, які показали свою

ефективність у зміні думки. У таблиці 1.3 представлені результати класифікаторів.

Таблиця 1.3 - Результати оцінки набору контрольованих алгоритмів машинного навчання для виявлення переконливих коментарів.

Algorithm	Precision	Recall	F-score
Наївний Байєс	0.609	0.779	0.684
LibSVM	0.523	0.996	0.686
Класифікація за допомогою регресії	0.714	0.740	0.727
Ada Boost	0.744	0.679	0.710
Випадковий ліс	0.744	0.760	0.752

Результати оцінки контрольованих алгоритмів є багатообіцяючими, оскільки ми отримали F-бал 0,752 за допомогою алгоритму Random Forest (як наш найкращий класифікатор) і середній F-показник 0,713 для всіх класифікаторів. Ці результати свідчать про те, що поверхневі атрибути можуть внести значний внесок у прогнозування переконливості повідомлення, незалежно від основних претензій та аргументів. Точність Random Forest становить 0,749, що приблизно на 10% більше, ніж набір функцій, використаний в [4] для передбачення успішних відповідей. Однак, враховуючи те, що наш базовий набір даних відрізняється від їхнього, результати можуть бути непорівнянними. Використання нашого набору функцій у їх наборі даних може пролити світло на внесок наших нещодавно введених функцій і включено в наші майбутні плани досліджень.

Щоб отримати уявлення про передбачувану силу окремих ознак, ми використали алгоритм вибору та ранжування ознак. Враховуючи, що найкраща продуктивність досягається Random Forest, для такого аналізу використовується в [26]. В алгоритмі Random Forest екземпляри класифікуються за голосами кількох дерев рішень, побудованих незалежно на різних вибірках даних для навчання. Мірим важливості атрибута є втрата точності класифікації, викликана випадковою перестановкою значень

атрибутів. Алгоритм Борути працює в ітераційному режимі, коли важливість оригінальних атрибутів порівнюється з важливістю їх рандомізованих копій [26].

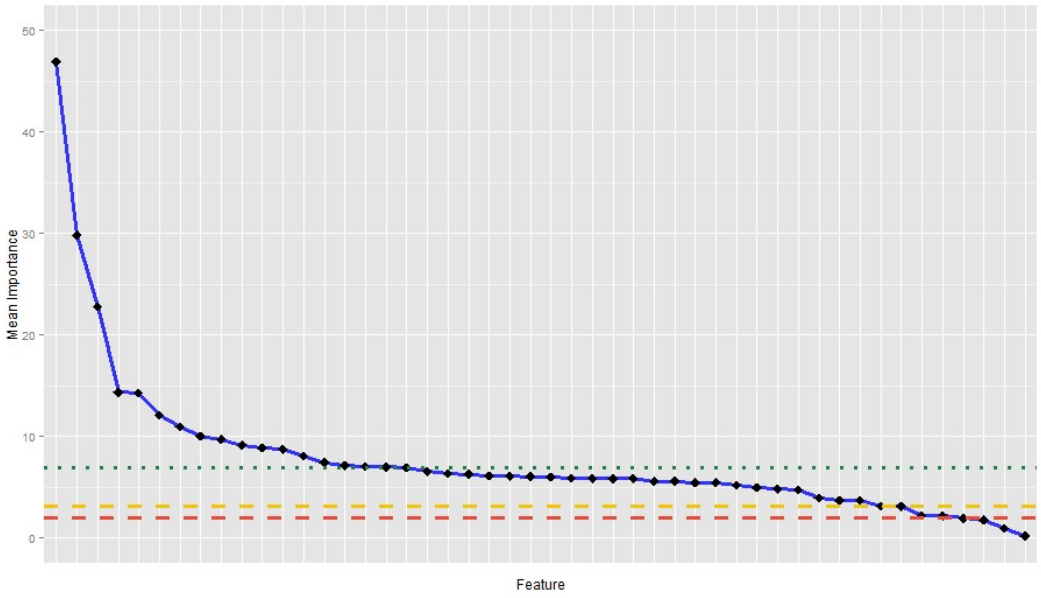


Рисунок 1.5 - Результат аналізу ознак

Таким чином, очікується, що потенційний вплив випадкових коливань і кореляцій можна мінімізувати. Результати, отримані за допомогою алгоритму Борути, представлені на рисунку 1.5, на якому вісь x представляє ознаки, а вісь ординат — середню важливість ознак, призначених Борутою. Залежно від середньої важливості функції порівняно з важливістю всіх функцій, Boruta також вирішує або відхилити, або підтвердити функцію. Крім того, алгоритм може бути не в змозі прийняти чітке рішення щодо функції і позначає її як Орієнтовну. Ознаки під нижньою горизонтальною лінією на рисунку 1.5 відхилені, що вказує на те, що вони не мають передбачуваної сили. Такими особливостями є емоційний тон мови, відсоток мови, специфічної для Інтернету (відомий як netspeak у LIWC), а також частота узагальнення та наголосу фраз. Крім того, три об’єкти, розміщені між середньою лінією та нижньою лінією, позначені як попередні. Такими ознаками є відсоток погодження, I займенник і оцінка читабельності тексту.

Незважаючи на те, що решту ознак підтверджує Борута, багато з них дуже схожі за середньою важливістю. Однак середня важливість починає помітно зростати після того, як мине сім (як позначено верхньою зеленою лінією). Як правило, коли кількість класифікаційних ознак занадто велика щодо розміру навчальної вибірки, ефективність класифікації може зменшитися, оскільки набір даних може бути заниженим у порівнянні з набором ознак [24]. Отже, ми оцінили класифікатори, використовуючи 18 ознак, розміщених над верхнім рядком. Цікаво, що середня продуктивність (F-score=0,718) і продуктивність Random Forest (F-score=0,763) трохи підвищилися.

Найпопулярніші вибрані функції представляють усі обговорювані категорії функцій. Наприклад, оцінка дельта користувача розміщується вгорі, що служить явним індикатором довіри до автора. Усі функції, пов’язані з релевантністю вмісту (тобто необроблений tf-idf, необроблений ранг релевантності та рейтинг релевантності), входять до п’ятірки найкращих функцій. Часова оцінка також займає високе місце в списку. Кількість слів, лексична різноманітність, кількість напрямних фраз і відсоток

1.5. Опис процедури “переконаних користувачів”

Фактичні причини, що призводять до зміни переконань і переконання, можуть відрізнятися від причин, які люди сприймають як переконливі [16]. Хоча аналіз груп коментарів допоміг нам зрозуміти фактичні причини переконання, аналіз пояснень користувачів може виявити передбачувані причини. Як обговорювалося раніше, у CMV ті, чия віра змінилася, зобов’язані надати коментар, пояснюючи, як і чому інший коментар змінив свою точку зору.

Ці пояснення часто є короткими, а іноді включають подальші обговорення, пов’язані з основною темою. Тому ми спочатку визначили відповідні речення, які стануть предметом подальшого аналізу. Ці речення

прямо посилаються на коментар, який змінив їхнє переконання, і пояснюють, як коментар був успішним. Тому ми виділили речення з явними посиланнями на попередні коментарі (наприклад, використання вас, коментар і аргумент), де ці лексеми є номінальним суб'єктом речення. Стенфордський аналізатор залежностей [25] потім використовується для виділення граматичної структури речень.

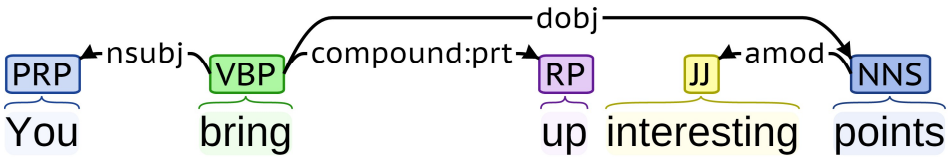


Рисунок 1.6 - Приклад пояснення, аналізованого аналізатором залежностей

Крім того, відсоток заперечень, розбіжностей і запитань є одними з ознак, які можуть мати психологічний вплив на процес зміни переконань.

Результатом аналізу залежностей є набір бінарних відношень, які є граматичними відносинами між керуючим і залежним. Губернатор іменного підмета часто є дієсловом. Однак, коли дієслово є дієсловом copula, корінь речення є доповненням до дієслова copula, яке може бути прикметником або іменником. Тому, аналізуючи керуючих вибраними токенами, ми можемо отримати уявлення про сприйняті причини та атрибути переконання в коментарях. Приклад виділеного речення та його дерева, що аналізується залежностями, показано на рисунку 1.6.

Після виділення з пояснень 2554 відповідних речень ми проаналізували губернатори предметів і підрахували їх частоту. Перед підрахунком ми переконалися, що губернатори не працюють. Крім того, коли губернатор є дієсловом, ми перевірили, чи є це складене дієслово, і витягли обидва слова, якщо це так (наприклад, підняти в реченні екземпляра). На рисунку 1.7 (а) представлено хмару тегів 50 найбільш частих керуючих (а) Представлення

хмари тегів для управління; (b) Представлення хмари тегів об'єктів, витягнутих з пояснень. витягнуто з пояснень.

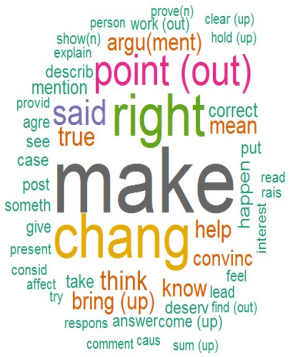


Рисунок 1.7 - Граматичний аналіз пояснень користувача може бути корисним для розуміння причин переконання

Крім того, ми виділили об’єкти, коли такими керуючими є перехідні дієслова (наприклад, точка в реченні екземпляра). Хмара тегів об'єктів показано на рисунку 1.7 (б).

Найчастіші слова пов’язані з реченнями, які користувачі надають, щоб описати, що відбулася зміна переконань. Приклади цих слів включають make і change з набору регуляторів і delta і view з набору об’єктів. Наприклад, такі речення, як «Ви змінюєте мій погляд» або «Ви заслуговуєте на дельту» є серед таких описів. Однак деякі з вилучених слів виявляють передбачувані причини переконання. Деякі з цих причин є спільними з реальними причинами, знайденими в аналізі коментарів, в той час як деякі можуть дати нові перспективи щодо того, чому люди вирішують змінити свої переконання. Наприклад, дієслова present і provide разом з прикладом об’єктів і джерелом означають, що користувачі цінують приклади та статистичні дані в процесі переконання. Цей висновок відповідає поширеній присутності чисел, ілюстративних фраз, а також минулого часу в мові вдалих коментарів. Крім того, слово sum up часто зустрічається в наборі губернаторів, що підтверджує

наші попередні висновки щодо частого використання узагальнюючих фраз у переконливих коментарях. Крім того, деякі фактично знайдені причини можуть бути факторами нижчого рівня, що сприяють тому, що сприймається як причина. Наприклад, правильне використання розділових знаків або відсутність безсилих лінгвістичних індикаторів може сприяти чіткості повідомлення, яке може бути пов’язано з дієсловом clear up, часто використовуваним у поясненнях.

З іншого боку, усвідомлені причини можуть виявити певні особливості переконливого тексту, які важко або навіть неможливо визначити за допомогою обчислювальних лінгвістичних засобів. Наприклад, наявність правильних, істинних, правильних і доказових серед губернаторів свідчить про важливість достовірності та обґрунтованості аргументів і міркувань у суспільстві. Крім того, дієслова, такі як підняти, вказати та підняти, разом із такими об’єктами, як перспектива, інсайт та вказують, показують цінність надання аргументів з нової точки зору.

Велика кількість досліджень була присвячена розумінню переконання та процесів переконання в традиційних умовах. Незважаючи на відмінності, очевидні в літературі, існують певні принципи переконання, які, як правило, спільні для різних дослідницьких спільнот. В [12] виділяють чотири широко спостережувані особливості тексту, які, як відомо, призводять до переконання. До них належать: структура аргументу, зміст, зрозумілість і достовірність. Відомо, що структура аргументу є однією з ключових ознак переконливого тексту [10]. Незважаючи на обґрунтовані теорії, запропоновані в [18], автоматизований аналіз і виявлення цих елементів все ще є складним завданням, і не існує надійних реалізацій таких структур. Тому тут ми обговорюємо наші висновки щодо інших трьох факторів і виділяємо обмеження дослідження.

Зміст аргументів виявився глибоко ефективним у процесі переконання [9]. Ми проаналізували вміст коментарів з точки зору його відповідності вмісту

публікації та виявили, що найрелевантніші мають вищі шанси виграти дельту. Крім того, за словами Борути, наші функції, пов'язані з вмістом, мають сильну прогнозну силу для переконливих коментарів. Крім того, перевірка пояснень показала важливість обґрунтованості аргументів. Попередня робота показала, що зміст, який містить обидві сторони аргументу та доводить хибність одного, має тенденцію бути більш переконливим [2]. Проте більшість цих досліджень зосереджені на науковому читанні та тексті. Незважаючи на те, що для прямого аналізу такого тексту спростування потрібні складні інструменти обробки мови, певні ознаки, знайдені в цьому дослідженні, можуть свідчити про те, що це може бути не так для обговорень в Інтернеті. Наприклад, слова-заперечення та слова для диференціації (наприклад, хоч і тоді, коли) є одними з лінгвістичних індикаторів, які можна використовувати, щоб відхилити аргумент. Однак такі показники частіше спостерігаються в непереконливих коментарях.

Іншим фактором, що сприяє процесу переконання, є зрозумілість тексту [1, 2]. Зрозумілість означає те, наскільки текст в цілому легко читається, і це дозволяє читачеві без зусиль зрозуміти сенс [2]. Деякі з мовних особливостей, виявлених у цій роботі, як-от використання розділових знаків та використання перехідних фраз, можуть призвести до зрозумілості. Таким чином, ці висновки узгоджуються з попередніми дослідженнями переконання в тексті та вказують на важливість зрозумілості зміни переконань в онлайн-середовищі. Фактично, ефект зрозумілості може посилюватися в контексті онлайн-спільнот, де користувачі можуть отримати велику кількість коментарів за короткий проміжок часу. Тому вони можуть вибірково приділяти більше уваги коментарям, для обробки яких потрібно менше зусиль і когнітивного навантаження. Подальше дослідження щодо прямого порівняння зрозумілості в традиційних і онлайн-формах тексту та його впливу на переконання є виправданим, щоб підтвердити це твердження.

Нарешті, усвідомлена достовірність комунікації відіграє істотну роль у процесі переконання. Дельта-оцінка користувача є єдиною інформацією, яка

може служити явним індикатором довіри до автора. Як і очікувалося, були виявлені значні відмінності між двома наборами коментарів щодо дельта-балла автора, і цей показник виявився найбільш передбачуваним для переконання. У зв'язку з відсутністю прямої інформації про автора, мовні ознаки, які можуть сигналізувати про достовірність, виявляються дуже цінними та важливими. Наприклад, переконливі коментарі оцінюються вище за оцінкою їх читабельності, що може бути потенційним показником рівня освіти автора. Безсилий стиль мови (наприклад, використання орієнтовних індикаторів) рідше зустрічається в успішних коментарях. Нарешті, всі особливості, які сприяють витонченості письма (наприклад, лексична різноманітність і складність), є більш домінуючими в переконливих коментарях.

API Reddit обмежує кількість публікацій, які можна отримати з веб-сайту, до останніх 1000 публікацій. Таким чином, після збору дописів і пов'язаних з ними коментарів протягом року, ми зібрали 1690 переконливих коментарів, які відповідають вимогам, щоб стати предметом дослідження. Однак для отримання надійних результатів можуть знадобитися більші набори даних. Наразі збір даних триває, і перевірка обґрунтованості наших висновків на більших наборах даних є частиною наших майбутніх планів дослідження. Окрім обмеження розміру даних, ми зосереджені на єдиній платформі з певним дизайном та аудиторією. Незважаючи на те, що CMV забезпечує привабливе середовище для вивчення онлайн-змін віри та переконання, те, як розроблена соціальна платформа, може формувати та впливати на поведінку користувачів [28]. Отже, ще одним обмеженням дослідження є той факт, що результати можуть бути специфічними для ЦМВ і не можуть бути узагальнені для інших онлайн-платформ. Тому ми плануємо перевірити та випробувати наші моделі в різних наборах даних, а також дослідити застосування моделей в інших онлайн-середах обговорення.

Незважаючи на те, що в традиційних умовах вже давно створено різноманітні основи переконання та теорії, онлайн-переконання та зміна

переконань приділяють відносно мало уваги з боку різних дослідницьких спільнот. У цьому дослідженні ми в основному зосередилися на варіантах онлайн-коментарів, які можуть призвести до переконання. Проаналізувавши набір коментарів із людськими анотаціями, витягнутих з Reddit.com, ми визначили набір атрибутів, характерних для переконливих коментарів, і дослідили їх передбачувану силу для переконання.

За винятком однієї тимчасової функції та одного атрибута, пов'язаного з історією користувачів у спільноті, інші функції пов'язані з переконливим впливом різних компонентів мови. Ці особливості привели до класифікації переконливих і непереконливих коментарів з розумною продуктивністю 75%. Нарешті, попередній аналіз пояснень користувачів дозволив проаналізувати передбачувані причини зміни їхнього переконання.

Майже всі варіації мови є важливими в процесі переконання [21]. Тому, хоча ми враховуємо 47 лінгвістичних особливостей, існують інші виміри мови, які потребують подальшого вивчення. Наприклад, синтаксичні варіанти, теги частини мови та рівні інтенсивності мови є одними з особливостей, які ми плануємо вивчати в майбутньому. У літературі переконання, крім особливостей тексту та письменника, значущими виявилися характеристики читача [17]. Однією з примітних властивостей є базові знання читача щодо теми, яка є предметом уваги. Зокрема, люди з більшими попередніми знаннями про цю тему менш схильні змінювати свої переконання. В [4] стверджували, що сила ставлення людини залежить від кількості вхідних повідомлень, пов'язаних з цим ставленням, яке людина обробляла. Оскільки автор оригінального допису та читач, який присуджує дельту, у CMV однакові, можна розглядати довжину оригінального допису як показник базових знань читача та аналізувати його потенційний вплив на процес переконання. Аналіз тематичних особливостей та їх потенційного впливу на переконливість повідомлень є виправданим.

Нарешті, дослідження переконання показали, що існують різні контекстуальні фактори, які впливають на переконливість вчинку [11]. Такі

фактори включають соціальний вплив, завдання зосередженості, коментарі, зроблені іншими, і дизайн основного середовища. Як відправну точку для всеосяжної обчислювальної моделі виявлення переконливих повідомлень у онлайн-взаємодіях, ми зосередилися на кореневих коментарях у контексті зміни переконань у Reddit. Ми плануємо покращити нашу модель, беручи до уваги ці контекстуальні фактори.

2. АЛГОРИТМІЗАЦІЯ ПРОЦЕСУ ВИЛУЧЕННЯ ТА АНАЛІЗУ RST CUE

2.1. Основні теоретичні положення

Риторичні відносини, також відомі як відносини дискурсу та відносини когерентності, є паратактичними або гіпотаксичними відношеннями, які зберігаються між проміжками тексту, пояснюючи побудову когерентності в дискурсі [17]. Протягом багатьох років дослідники були зосереджені на створенні надійних теоретичних основ і таксономій для риторичних відносин [8, 9, 10, 12]. Проте автоматична ідентифікація таких зв'язків між текстовими проміжками залишалася складним і непростим завданням.

Полегшення автоматичної ідентифікації риторичних зв'язків може сприяти вирішенню різних завдань обробки природної мови (NLP), зокрема узагальнення тексту, генерації тексту та оцінки есе [19]. Тому необхідно провести ретельні дослідження на основі корпусу, щоб отримати знання про набори ознак, які мають значення для виявлення як явних, так і неявних відносин.

Явні відношення – це ті, про які сигналізують сигнали, тоді як в неявних зв'язках жодної репліки немає. Загалом, різні види сигналів можуть сигналізувати про існування відношення, включаючи лексичні підказки, настрій, модальність та інтонацію [17]. У цьому дослідженні ми зосереджені на явних відносинах у письмовому тексті, які сигналізуються наявністю лексичних ознак. Ми прагнемо зрозуміти здатність таких сигналів автоматичного виявлення риторичних відносин. Ми також сподіваємося зрозуміти, чи залежить ефективність підходів на основі репліки залежно від природи відношення та основного жанру тексту.

Це дослідження є частиною більш масштабної роботи з визначення обґрунтування в тексті дискурсу. Враховуючи, що три риторичні відносини:

обставини, оцінка та розробка зазвичай присутні в обґрунтуваннях [22], ми зосередимося на цих трьох відносинах у цьому дослідженні. Ми надаємо початкові результати для наступних завдань:

1) розуміння особливостей корпусних підказок

2) дослідження значення таких лексичних підказок у виявленні різних риторичних зв'язків

3) аналіз потенційних відмінностей і подібностей між репліки, витягнуті з двох корпусів, що належать до різних жанрів.

Експерименти проводяться з двома анотованими корпусами RST: корпусом RST [6] та набором оглядових даних (SFU) [18]. Змінена версія tf-idf, запропонована в [3], використовується для вилучення набору ключових нграм з кожного корпусу як потенційних лексичних ознак, що сигналізують про наявність риторичних зв'язків. Аналізуючи такі базовані на корпусі сигнали, це дослідження виходить за рамки фіксованих списків сигналів, які зазвичай пропонуються та використовуються в попередній літературі. Такі набори сигналів зазвичай пов'язані з кількома чітко визначеними синтаксичними категоріями, такими як сполучники (наприклад, але і тому, що) і прийменники (наприклад, з і через) [15], і називаються маркерами дискурсу або дискурсивними сполучниками. Щоб отримати емпіричні докази щодо потенційної цінності використання лексичних підказок для ідентифікації риторичних зв'язків, для виявлення випадків такого ж відношення в іншому корпусі було використано список основних ознак для кожного відношення та витягнутий з кожного корпусу.

2.2 Аналіз робіт по темі дипломної роботи

Робота з автоматичного виявлення риторичних зв'язків поділяється на дві категорії досліджень. Низка досліджень спиралася на анотовані людиною корпуси та запропонувала різні моделі з різними наборами ознак для виявлення

риторичних відносин. Інші дослідники дотримувалися підходу, заснованого на шаблонах, коли набір шаблонів використовується для створення мічених штучних корпусів. Побудовані корпуси потім використовуються для побудови моделей вилучення відносин.

Використовуючи банк діаграм дискурсу [21], [20] розробив класифікатор на основі багатий набір синтаксичних і лексико-семантичних ознак, деякі з яких побудовані на основі зовнішніх джерел знань. Зосереджуючись на розпізнаванні неявних риторичних відносин, Лін та ін. [11] запропонував класифікатор, заснований на Penn Discourse Treebank, беручи до уваги контекст двох діапазонів, інформацію про пари слів, а також аналіз внутрішньої конституції проміжків і залежностей. В [14] для побудови моделі автоматичного виявлення зв'язків використовуються декілька лінгвістично інформованих ознак, таких як полярність слів, класи дієслова та пари слів.

Дотримуючись підходу, заснованого на шаблонах, для виявлення неявних відношень, в [13] використовували невеликий набір шаблонів, побудованих на наборі заздалегідь визначених маркерів дискурсу, щоб виділити відносини та автоматично сформувати анотований корпус речень. Корпус невідношень також утворюється шляхом випадкового вибору несуміжних речень на відстані не менше трьох речень. Ці навчальні набори потім використовуються для вивчення пар слів, які, ймовірно, зустрічаються разом із кожним відношенням. Нарешті, для виявлення зв'язків між текстовими проміжками використовується інформація про пару слів разом із набором маркерів дискурсу. Blaire та ін. [4] розширив цей підхід, уточнивши процес навчання та класифікації за допомогою оптимізації параметрів, сегментації тем і синтаксичного аналізу. В [16] також розширив підхід [13], показавши, що фразові шаблони, виділені з текстових проміжків, можуть бути цінними для завдання ідентифікації відношень.

В [3] дотримувалися гібридного підходу. У їхньому дослідженні замість використання попередньо визначеного списку маркерів дискурсу, список

лексичних ознак (так звані індикатори відношень) витягується з анотованого людиною корпусу (тобто корпусу RST). Цей список базованих на корпусі сигналів використовується для побудови шаблонів і для формування штучного корпусу. Потім функції пар слів використовуються для виділення підмножини риторичних відносин.

Подібно до [3], у нашій роботі ми вирішили використовувати анотовані корпуси для виділення набору ознак відношень, виходячи за межі заздалегідь визначених списків маркерів дискурсу, які використовувалися в більшості попередніх робіт. Однак замість того, щоб використовувати цей список для створення штучного корпусу, ми зосередилися на здатності базованих на корпусі лексичних підказок ідентифікувати різні відносини та потенційний вплив жанрових факторів на якість та продуктивність цих реплік.

2.3. Сутність запропонованої методології

Ми використали два анотовані людьми корпуси як наші базові набори даних для експериментів: дані RST [6] та набір даних огляду ДФУ [18]. Обидва корпуси анотовані в структурі RST і створені за допомогою RSTTool [2]. RST є однією з найбільш поширених платформ для аналізу та розуміння дискурсу. У структурі RST елементарні дискурсивні одиниці (тобто атомні проміжки) можуть бути пов’язані через певні дискурсивні відносини, щоб утворити вбудовані текстові проміжки. Потім вбудовані проміжки можуть брати участь у новому відношенні, що веде до ієрархії відносин, яку можна представити у вигляді дерева дискурсу..

Корпус RST, який протягом багатьох років був доступним Консорціумом лінгвістичних даних, включає 385 статей Wall Street Journal і охоплює понад 178 000 слів. Серед відношень, що використовуються для анотації, існують три різні варіанти оціночного відношення. evaluation-s використовується, коли оцінка відбувається в супутнику (тобто, менш істотний

діапазон у відношенні), а evaluation-n використовується, коли оцінка відбувається в ядрі (тобто, більш істотний діапазон у відношенні). Якщо оцінка має однакову вагу в обох діапазонах, екземпляр позначається міткою оцінки. У цьому дослідженні текстові проміжки, позначені кожною з трьох міток, вважаються прикладом оцінки. Крім того, для анотування корпусу використовуються шість підвідношень розробки (наприклад, розробка-додаткова, розробка-набір та розробка-частина-ціле). Екземпляри зв'язку, позначені кожним із цих варіантів, розглядаються як екземпляр розробки.

Рецензійний корпус SFU — це колекція з 400 оглядових документів з фільмів, книг і товарів народного споживання. Цей набір даних містить понад 303 000 слів і був зібраний у 2004 році з веб-сайту Epinions [1]. Щоб спростити процес порівняння та аналізу для певних жанрів в решті рукопису, ми називаємо корпус RST корпусом новин, а набір даних SFU будемо називати корпусом огляду. Таблиця 2.1 показує відсоток відносин фокусування між двома корпусами.

Таблиця 2.1 - Відсоток відносин фокусування в базових корпусах

Відношення	Огляди новин	Огляди новин
обставина	3%	8%
оцінка	1%	2%
опрацювання	36%	7%

Вилучення лексичного сигналу

Щоб виділити лексичні ознаки, пов'язані з даним відношенням, спочатку створюється документ відношення, позначений відповідно до його відповідного риторичного відношення (наприклад, обставина, оцінка та розробка), щоб включити всі проміжки тексту, які пов'язані з цим відношенням у базовому корпусі. Подібним чином ми створили документ без зв'язку, що містить усі текстові проміжки, які беруть участь у будь-якому відношенні, крім

відношення фокуса. Кожна пара проміжків, пов'язаних зв'язком, об'єднана в правильному порядку і представлена в одному рядку.

2.4. Результати попереднього імітаційного експерименту

Формування цих двох документів для кожного відношення дозволяє нам використовувати кожен корпус або як навчальний набір для вилучення сигналів, або як тестовий набір для аналізу та оцінки використання лексичних підказок у процесі виділення відношень.

Таблиця 2.2 - Зразки лексичних підказок, витягнутих із двох корпусів

Відношення	Підказки для перегляду новин	Підказки для перегляду новин
обставина	коли, зараз, відколи, поки, раз	коли, зараз, відколи, поки, раз
оцінка	добре, високо, добре, подобається, вражений	добре, високо, добре, подобається, вражений
опрацювання	хто, який, як який, де, ніби	хто, який, як який, де, ніби

Щоб виділити потенційні лексичні ознаки, було дотримано підходу, запропонованого в [3]. Усі n-грами (аж до триграм) спочатку були витягнуті з документа відношення. Потім для кожного ngram була розрахована змінена версія метрики tf-idf. Показник idf все ще обчислювався на основі кількості документів, які містять ngram, та загальної кількості документів у корпусі. Однак, оскільки кожен рядок відповідає одному екземпляру відношення, метрика tf розраховується на основі кількості рядків, які містять принаймні один екземпляр ngram. Ця змінена метрика дозволяє нам компенсувати потенційне упередження, яке може бути викликано метрикою tf для слів, які з'являються більше одного разу в екземплярі відношення. Список витягнутих нграм (тобто лексичних сигналів) був відсортований на основі зміненої міри tf-idf у порядку спадання. Потім цей список було відфільтровано, щоб не включати займенники, модальні дієслова чи допоміжні дієслова.

З метою виділення відношень відсортований список лексичних ознак, витягнутих з кожного корпусу, був застосований до іншого корпусу. Таким чином, лексичні підказки, витягнуті з корпусу новин, були застосовані до корпусу огляду, а лексичні підказки, витягнуті з набору оглядів, були використані в корпусі новин. Потім вимірювання точності, точності, запам’ятовування та F були розраховані для кожного з сигналів незалежно. Малюнок 8.1 ілюструє загальний процес виділення сигналу та класифікації відношень для відношення, що цікавить. Виходячи з результатів експерименту, можна було побачити, що для всіх відносин і для обох наборів даних ці сигнали, ранжовані після 120, мали нульову ефективність. Таким чином, лише перші 120 сигналів, замовлених tf-idf, були залучені до процесу аналізу. Таблиця 8.2 містить зразки цих сигналів, виділені з двох корпусів. Пізніше ми обговоримо, як ці два набори сигналів працювали при виділенні відносин з іншого корпусу і як вони перекриваються.

Результати експерименту представлені у вигляді лінійних діаграм, де вертикальна вісь представляє міру фокусування (тобто tf-idf, точність, запам'ятовування та оцінку F), а горизонтальна вісь вісь представляє 120 вибраних сигналів, відсортованих на основі міри фокусування в порядку спадання. Тому подібні числові позначки на діаграмах можуть вказувати на різні лексичні ознаки. Таблиця 2.3 показує перші два ознаки в діаграмах обставин. Мітка набору даних на діаграмах позначає навчальний набір, тобто звідки витягуються сигнали. На всіх діаграмах темний колір представляє набір даних новин, що вказує на те, як лексичні підказки, витягнуті з набору даних новин, виконуються на корпусі огляду. Світліший колір представляє набір даних огляду, що ілюструє заходи для основних сигналів, витягнутих із колекції оглядів. Таке уявлення полегшує аналіз тенденцій і можливостей сигналів між цими відносинами та між жанрами.

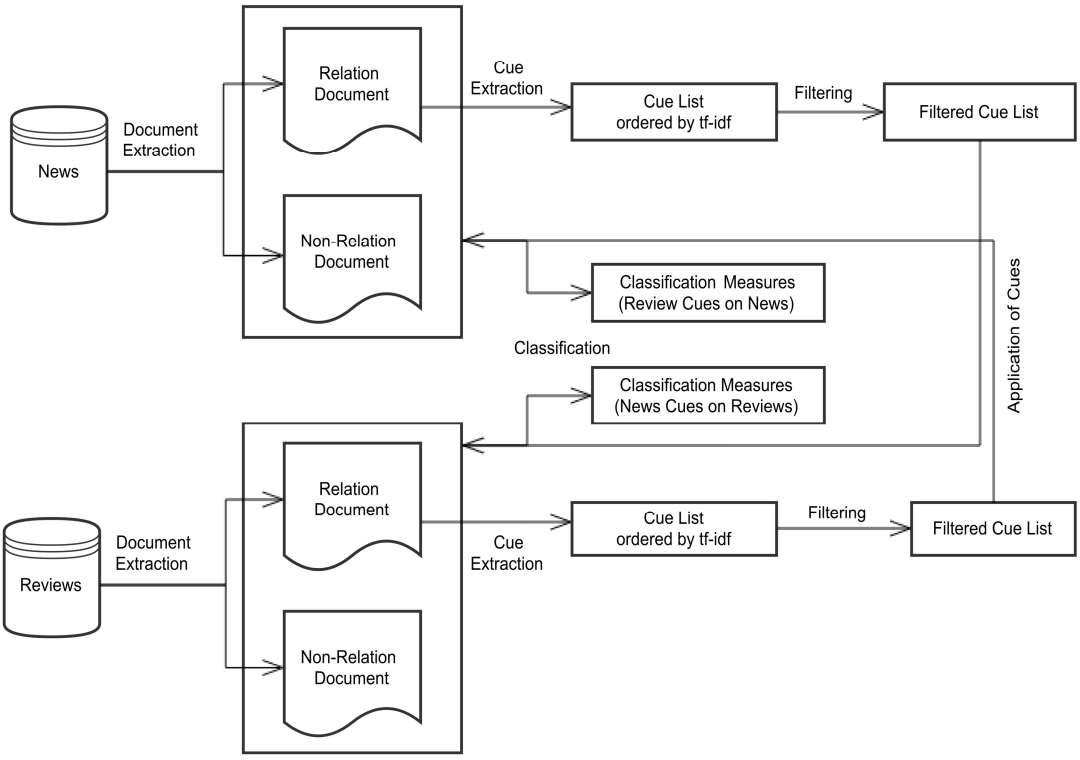


Рисунок 2.1 - The extracted cues from each corpus are applied to the other corpus to classify риторичні відносини

Як згадувалося раніше, ми створили документи без зв'язку, зібравши всі екземпляри, які пов'язані будь-яким відношенням, крім відношення фокуса. Таким чином, документи без зв'язку містили набагато більшу кількість екземплярів порівняно з документами про відношення, що призвело до великої кількості справжніх негативів для кожного сигналу в наших результатах класифікації. Тому результати точності були недостовірними і виключені з аналізу.

На рисунку 2.2 показано міру tf-idf, розраховану для основних сигналів, витягнутих для всіх трьох відносин і з обох корпусів. Як видно, показник tf-idf стабільно нижчий для лексичних ознак, витягнутих з оглядів. Цю знахідку можна віднести до той факт, що текст новин – це добре структурований офіційний текст, тоді як онлайн-огляди є відносно менш структурованими та неформальними, написаними користувачами з широким спектром умінь писати. Таким чином, лексичні підказки можуть використовуватися більш послідовним

чином у збірці новин порівняно з набором оглядів, що свідчить про те, що підходи, засновані на репліках, можуть краще працювати на офіційному та структурованому тексті.

Таблиця 2.3 - Мітки сигналів та відповідні їм лексичні ознаки для відношення обставин

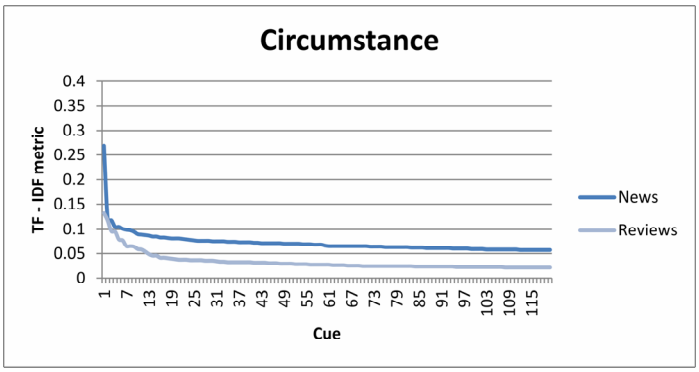
	№1	№2	
Виміряти	Огляди новин	Огляди	новин
TF-IDF	коли коли	як коли я	як коли я
Точність	коли справа доходить до	коли він прийде	коли він прийде
Пригадайте	коли коли	на після	на після

Порівнюючи сюжети для трьох відношень, можна побачити, що сигнали розробки мають відносно нижчий tf-idf для обох наборів даних, що вказує на те, що незалежно від жанру розробка може погано сигналізуватися лексичними ознаками. Цей показник є вищим як для обставин, так і для оцінки, особливо для сигналів з найвищим рейтингом; однак різниця між ознаками, витягнутими з двох корпусів, є більш значною для оціночного відношення. Ці результати можуть вказувати на те, що лексичні ознаки можуть бути більш специфічними для жанру для відношення оцінки.

Після виділення лексичних підказок з кожного набору ми провели подальші експерименти, використовуючи ці підказки для вилучення риторичних зв’язків з іншого набору. Показники точності, відкликання та F були розраховані для кожної репліки незалежно. Рисунки 3.3 і 3.4 ілюструють показники точності та відкликання для верхніх сигналів відповідно. Результати для обставин узгоджуються з висновком tf-idf, оскільки лексичні підказки, виділені з набору новин, мають відносно високу точність. Результати оцінки також можна очікувати, оскільки вони значно падають після перших кількох сигналів, що підтверджує, що ознаки оцінки можуть бути дуже специфічними для основного жанру. Однак результати розробки не узгоджуються з метрикою

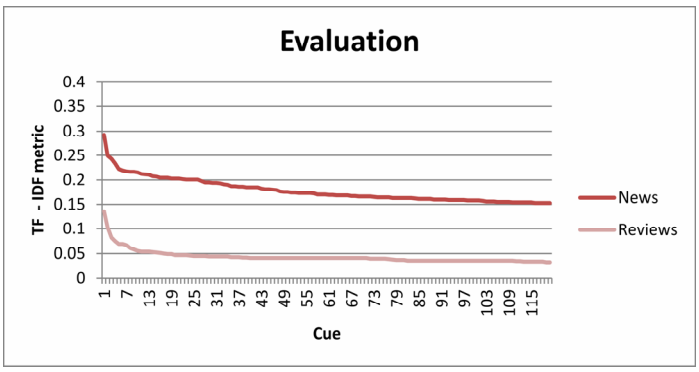
tf-idf. Незважаючи на те, що показники tf-idf були низькими для ознак розробки, точність відносно висока і значно вища для лексичних підказок, витягнутих із колекції оглядів. Цей несподіваний результат можна пояснити високим відсотком розробок у наборі даних новин. Як видно з таблиці 3.1, більше третини екземплярів відношень набору новин анотовано як розробка. Ця велика частка екземплярів розробки призводить до відносно високої точності навіть випадкового сигналу.

На відміну від tf-idf і точності, метрика відкликання має дуже подібну тенденцію для всіх відносин і наборів даних (див. рисунок 2.4). Для кожної репліки екземпляри класифікуються як помилкові негативні складаються з тих явних відношень, які сигналізуються будь-яким лексичним сигналом, крім сигналу фокусування, а також неявних екземплярів, які взагалі не сигналізуються. Відносно велика частка зв'язків, як правило, є неявними [17], що може призвести до того, що кількість істинних позитивних і хибних негативів, про які сигналізують інші ознаки, мало впливатиме на результати відкликання. Тому міра відкликання в основному характеризується неявними відношеннями.

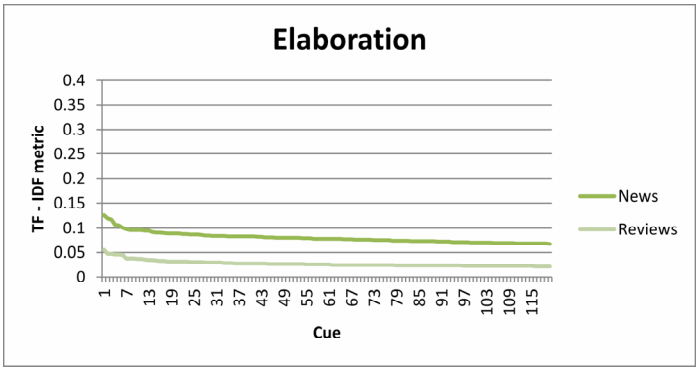


(а)

Рисунок 2.2 - Показник tf-idf нанесено на графік (початок)



(b)



(c)

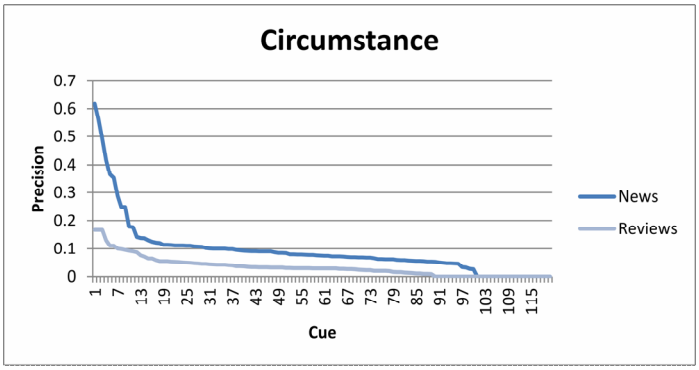
Рисунок 2.2 - Показник tf-idf нанесено на графік для основних сигналів, витягнутих із наборів даних новин та оглядів, для співвідношення обставин (a), оцінки (b) та розробки (c).

Нарешті, результати розрахунку бала F, як найнадійнішого показника продуктивності, наведені на рисунку 2.5. За винятком розробки, більшість лексичних підказок, витягнутих із набору даних новин, ефективніші у виділенні риторичних зв'язків. Ця краща продуктивність для цих двох відносин узгоджується з результатами tf-idf. Крім того, кращу продуктивність вилучених сигналів із набору оглядів у корпусі новин можна пояснити тим, що значна частина набору даних новин позначена як розробка. Порівняння результатів для різних відношень показує, що головні ознаки обставин мають кращу продуктивність у порівнянні з основними сигналами з двох інших відносин, тоді як різниця між оцінкою та розробкою менш суттєва.

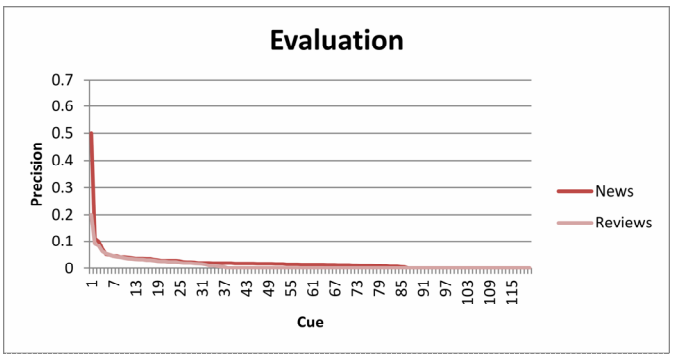
Результати експериментів демонструють, що основні сигнали для відношення обставин працювали краще, ніж основні сигнали з двох інших

відносин, і результати узгоджуються для обох наборів даних. Цей висновок можна пояснити тим, що відношення, які зазвичай виражаються через підрядність, сильніше сигналізуються у порівнянні з відношеннями, які існують між двома чи більше реченнями (наприклад, розробка та оцінка) [17]. Що стосується відношення обставин, основні підказки беруться з маркерів загального дискурсу, запропонованих і використаних у попередній літературі. Наприклад, коли, з, після і до є одними з основних ознак для обох корпусів. Ця велика кількість ознак, які є маркерами дискурсу, робить відношення обставин відносно незалежним від жанру і призвело до однакової продуктивності для обох наборів даних. Підрахунок кількості перекриваються сигналів, виділених з обох корпусів, підтверджує цей висновок, оскільки в двох списках сигналів є 29 подібних сигналів, причому всі вони є маркерами дискурсу.

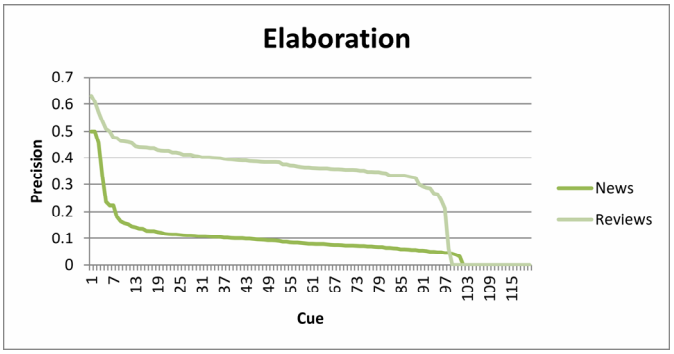
Виходячи з результатів експерименту для оцінки, можна зробити висновок, що відношення оцінки зазвичай сигналізується залежно від жанру тексту, а лексичні підказки можуть бути більш специфічними для жанру порівняно з двома іншими відношеннями. Більшість ознак оцінки, витягнутих з обох наборів даних, не є традиційними маркерами дискурсу (див. таблицю 2.2 для зразків), а походять із недостатньо вивчених синтаксичних груп, таких як прикметники, що робить розумним мати різні ознаки з двох різних текстових жанрів. Наприклад, любов – це лексичний знак, який зазвичай зустрічається в прикладах оцінки в рецензії.



(а)

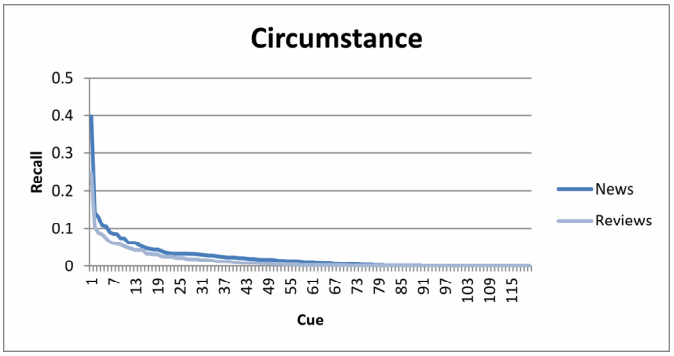


(b)

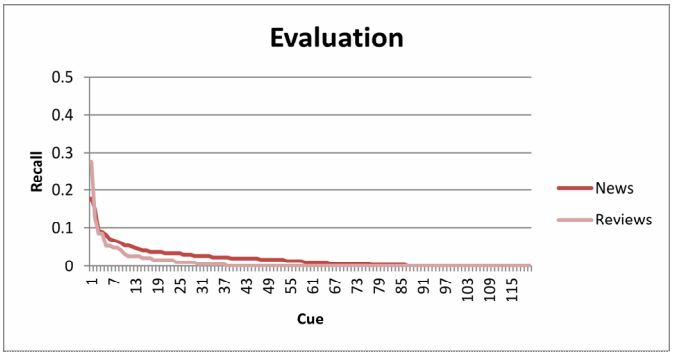


(c)

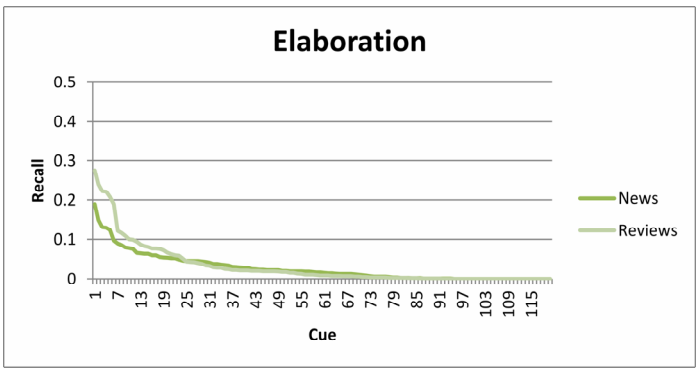
Рисунок 2.3 - Показник точності нанесено на графік для верхніх сигналів, витягнутих із навчального набору, який представлений темнішим кольором, для відношень обставин (a), оцінки (b) і розробки (c).



(a)

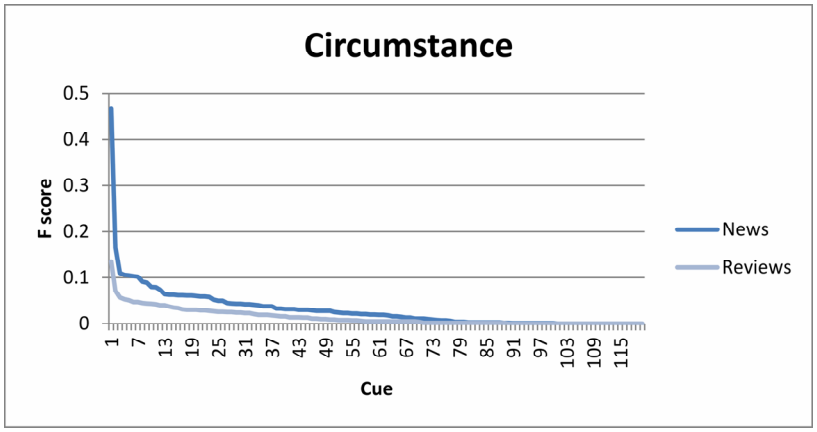


(b)

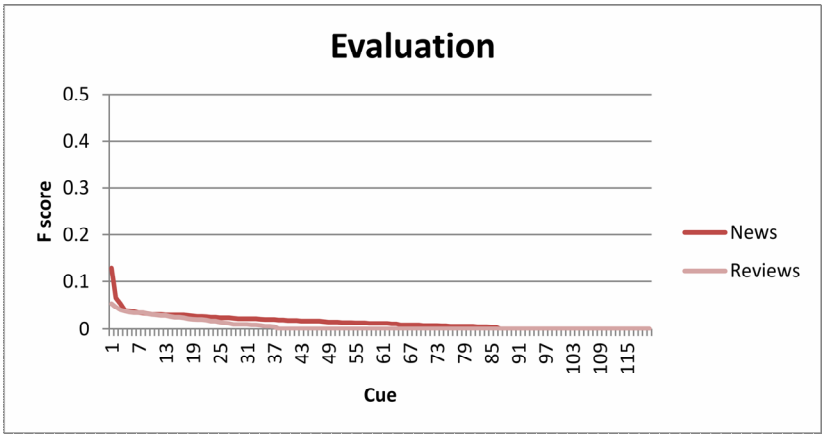


(c)

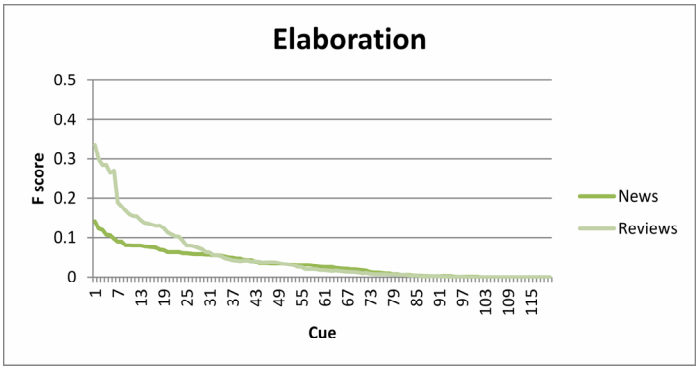
Рисунок 2.4 - Показник відкликання зображений для основних сигналів, витягнутих із навчального набору, який представлений темнішим кольором, для відношень обставин (a), оцінки (b) та розробки (c)



(a)



(b)



(c)

Рисунок 2.5 - Показник F нанесено на графік для основних сигналів, витягнутих із навчального набору, який представлений темнішим кольором, для відношень обставин (a), оцінки (b) та розробки (c).

Результати для відношення розробки можуть бути не такими надійними через незбалансовану кількість екземплярів розробки між двома корпусами (див. таблицю 2.1). Велика частка випадків опрацювання в корпусі новин є результатом структури та характеру новини, де кожна частина тексту надається для розробки попередньої частини. Однак з аналізу метрики tf-idf можна зробити висновок, що відношення розробки легко сигналізується, в результаті чого методи, засновані на сигналах, не дуже підходять для вилучення екземплярів розробки. Хоча серед двох наборів сигналів є 42 перекриваються підказки, вони не мають великої ваги, щоб на них покладатися. Цей висновок узгоджується з попередньою лінгвістичною літературою, де стверджується, що розробка є занадто погано визначеною, щоб її навіть розглядати як відношення [9].

Загалом, результати свідчать про те, що природа відношення відіграє значну роль в ефективності використання лексичних підказок для задачі виявлення відношення. Крім того, синтаксична група, до якої належать ці сигнали, може визначити, наскільки відношення є специфічним для основного жанру. Коли загальні маркери дискурсу є домінуючими ознаками, відношення можна вважати жанрово незалежним, тоді як у випадку прикметників чи дієслів

як ознак, відношення є більш жанрово-специфічним. Наше дослідження може стати основою для подальших досліджень щодо автоматичного виявлення риторичних відносин. Для розробки алгоритмів, заснованих на репліках, слід думати про різні відношення (або, принаймні, категорії відношень) незалежно і слід враховувати специфічні для жанру фактори для відносин, які сигналізуються іншими типами сигналів за межами маркерів дискурсу.

Під час розгляду підходів, заснованих на сигналах, вибір оптимальної кількості сигналів може бути складним завданням. Ми провели деякі експерименти з різною кількістю сигналів і розрахували сукупну точність, запам'ятовування та оцінку F для різних наборів сигналів. Таблиця 2.4 містить результати для наборів сигналів для обставин різних розмірів, витягнутих із колекції новин і застосованих до корпусу огляду. Як можна помітити, точність постійно зменшується в міру додавання нових сигналів, що свідчить про те, що кількість хибних спрацьовувань, доданих новими сигналами, більша в порівнянні з доданою кількістю справжніх сигналів. Проте міра відкликання збільшується в міру збільшення набору сигналів. Результати для відкликання були очікуваними, оскільки ми можемо витягти більше справжніх позитивів, додаючи нові лексичні підказки. Оцінка F зменшується, що вказує на те, що вплив точності більший, ніж відкликання.

Для цього експерименту з частинами, як не дивно, оптимальний показник F був досягнутий лише за допомогою першого сигналу (коли). Коли була додана друга ознака (як), велика кількість помилкових спрацьовувань призвела до зміни точності і, врешті, оцінки F, незважаючи на покращення людей справжніх позитивних результатів і метрики відкликання. Ці результати вимагають подальших обчислювальних кроків у підходах на основі сигналів, щоб використовувати неоднозначні сигнали та ідентифікувати в таблиці 2.4: Точність, запам'ятовування та оцінка F розраховуються для набору сигналів різних розмірів. Результати повідомляються для ознаки, витягнутих з корпусу новинок і застосованих до огляду корпусу для відношення стану.

Таблиця 2.4 – Результати ознак

Виміряти	15	45	60	75	90	105	120
Точність	0.24	0.11	0.10	0.9	0.09	0.08	0.08
Пригадайте	0.59	0.85	0.91	0.94	0.94	0.95	0.95
F Оцінка	0.34	0.20	0.18	0.16	0.16	0.16	0.15

Після додавання таких додаткових кроків необхідні подальші експерименти для аналізу та вивчення оптимальної кількості сигналів для різних відносин і текстових жанрів.

Таблиця 2.5 - Розподіл атомарних і вбудованих відношень для трьох відносин і між двома корпусами.

	Огляди новин		Огляди новин	
Відношення	Atomic	Embedded	Atomic	Embedded
обставина	46%	54%	78%	22%
оцінка	21%	79%	69%	31%
опрацювання	48%	52%	70%	30%

У RST атомарні текстові одиниці пов’язані один з одним за допомогою риторичних відносин. Ці пов’язані одиниці потім можуть брати участь у нових відносинах з іншими проміжками, що призводить до ієрархії відносин. Таблиця 2.5 показує розподіл атомарних і вбудованих відношень для всіх трьох відносин і між обома корпусами. Як видно, кількість вбудованих екземплярів відношень становить значну частку екземплярів у всіх випадках. Цей відсоток стабільно вищий для корпусу новин. Цей висновок був очікуваним, оскільки, на відміну від набору даних новин, корпус огляду анотується на рівні речення. Одним із потенційних обмежень цього підходу є відсутність акценту на атомарних і вбудованих відносинах як двох різних типах відносин. Однак наявність множинних відносин у відношенні дійсно може вплинути на підходи до вилучення сигналів. Наприклад, розглянемо два наступні випадки відносини обставин із корпусу новин:

Приклад №1:

[одягнена в джинси та світшот] [поки вона пробирається крізь постійний післяобідній дощ]

Приклад №2:

[нещодавно підрядник врятував її від провалу трьох поверхів][[оскільки вона досліджувала, що залишилося від старого вікторіанського будинку][спаленого підпалом]розробки]обставини

Перший екземпляр складається з двох атомних проміжків. Другий екземпляр, однак, складається з одного атомного діапазону як його першого діапазону та неатомного діапазону як його другого.

2.5. Опис результатів моделювання

Використовуючи наш підхід, as і by буде присвоєно однакову вагу tf для другого екземпляра, тоді як by явно не залежить від обставини і використовується для розробки попередньої частини. Отже, необхідно провести глибокий аналіз, щоб зрозуміти, як така ієрархічна структура може вплинути на підходи, засновані на репліках, і як виявити та переоцінити менш релевантні частини екземпляра під час пошуку лексичних підказок.

Ми описали експерименти, проведені для покращення нашого розуміння потенціалу та обмежень використання лексичних підказок для автоматичного визначення риторичних відносин у тексті. Витягуючи лексичні підказки з анотованих корпусів, ми вийшли за межі звичайного використання заздалегідь визначених маркерів дискурсу і дослідили, чи можуть лексичні підказки з інших синтаксичних категорій сигналізувати про наявність відносин. Крім того, ми мали на меті зрозуміти, як сигналізуються різні відносини. Крім того, ми використали два різних корпуси з двох різних жанрів, щоб дослідити, чи можуть специфічні для жанру фактори впливати на те, як сигналізуються

відносини. Загалом, результати свідчать, що те, як сигналізуються відносини, як загалом, так і між жанрами, значною мірою залежить від природи відносин.

Однією з головних відмінностей RST від інших теорій є те, що риторичні відносини акцентують увагу на намірах письменників і вплив цього відношення на читача [17]. Таким чином, анотації RST за своєю суттю є суб'єктивними і засновані на розумінні тексту читачами [17]. Тому деякі відмінності, виявлені в двох корпусах, можуть бути результатом потенційно різних знань кожного набору анотаторів щодо культури, ситуації та мови, які представляє текст.

RST постулює ієрархічну структуру тексту; таким чином, проміжки тексту можуть бути як атомарними, так і складатися з кількох атомних проміжків. Один із наших майбутніх дослідницьких активів буде зосереджений на незалежному аналізі вбудованих та атомарних речень. Крім того, зосереджуючись на добре повідомлених зв'язках, таких як обставини, ми плануємо покращити підходи, засновані на репліках, шукаючи нові методи для роз'яснення лексичних сигналів. Зараз ми створюємо моделі на основі графіків, щоб усунути неоднозначність сигналів, беручи до уваги їх синтаксичний контекст. Крім того, ми проведемо подальші експерименти, щоб визначити оптимальну кількість лексичних підказок для різних відносин.

Як зазначалося у вступному розділі, обставина, оцінка та розробка є трьома поширеними відношеннями RST в обґрунтуванні [22]. Як перший підхід до визначення обґрунтування в онлайн-корпусах, ми спробували виявити ці три риторичні зв'язки. Обгрунтування — це пояснення причин, що лежать в основі рішень, висновків та тлумачень. Попередні дослідження артикуляції та обміну обґрунтуваннями свідчать про те, що це сприяє контролю якості, управлінню знаннями та повторному використанню, витягуючи обґрунтування з тексту, створеного в результаті цих широкомасштабних ідей та обговорень [24, 23]. Було кілька спроб застосування обчислювальних методів для визначення обґрунтування погано структурованого тексту, такого як онлайн-дискурс та огляди [3, 7, 5].

3. ПРОГРАМНЕ ТА АЛГОРИТМІЧНЕ РІШЕННЯ ЗАДАЧІ ВІЗУАЛІЗАЦІЇ ПРОГРАМНИХ СЕРВІСІВ

3.1. Семантичне обґрунтування розробки

Семантично обґрунтований текст складається з дискурсивних одиниць, пов'язаних між собою дискурсивними відносинами, які також називають риторичними відносинами. Незважаючи на зусилля зі створення надійних теоретичних основ і таксономій для таких відносин [8, 13, 14, 16], сучасні методи їх автоматичного аналізу та виявлення в письмовому дискурсі ще мають удосконалити. Однак надання надійних моделей для аналізу та ідентифікації риторичних зв'язків може принести користь різним напрямкам дослідження в комп'ютерній лінгвістиці, таким як генерація тексту [9] та узагальнення [17], а також машинний переклад [18].

Однією з широко визнаних структур для аналізу та розуміння дискурсу є теорія риторичної структури (RST) [16]. У RST структура дискурсу має форму дерева, де листи відповідають елементарним одиницям дискурсу, а внутрішні вузли відповідають суміжним текстовим проміжкам. Кожен внутрішній вузол позначений риторичним відношенням, яке зберігається між його дочірніми вузлами. На малюнку 9.1 наведено приклад дерева RST, взятого з корпусу RST [3]. Однією з помітних відмінностей RST від інших подібних теорій є те, що вона побудована на намірах авторів використовувати ці відносини [24]. Ця відмінна риса може ще більше ускладнити побудову моделей для автоматичної ідентифікації та класифікації риторичних відносин у контексті RST.

Риторичні відносини можуть бути як явними, так і неявними. Експліцитні відношення – це ті, які сигналізуються за допомогою реплік, таких як лексичні підказки, настрій, модальність та інтонація [24], тоді як в неявних зв'язках жодної репліки немає. У цьому дослідженні ми зосередимося на явних відносинах у письмовому тексті, про які сигналізує наявність лексичних ознак.

Лексичні ознаки визначаються як мовні вирази, які функціонують як явні індикатори дискурсивного відношення [7]. Наприклад, у реченні, представленому на рисунку 3.1, але і тому можна вважати лексичні ознаки, що сигналізують про існування відношення поступки та відношення пояснення-аргументація відповідно.

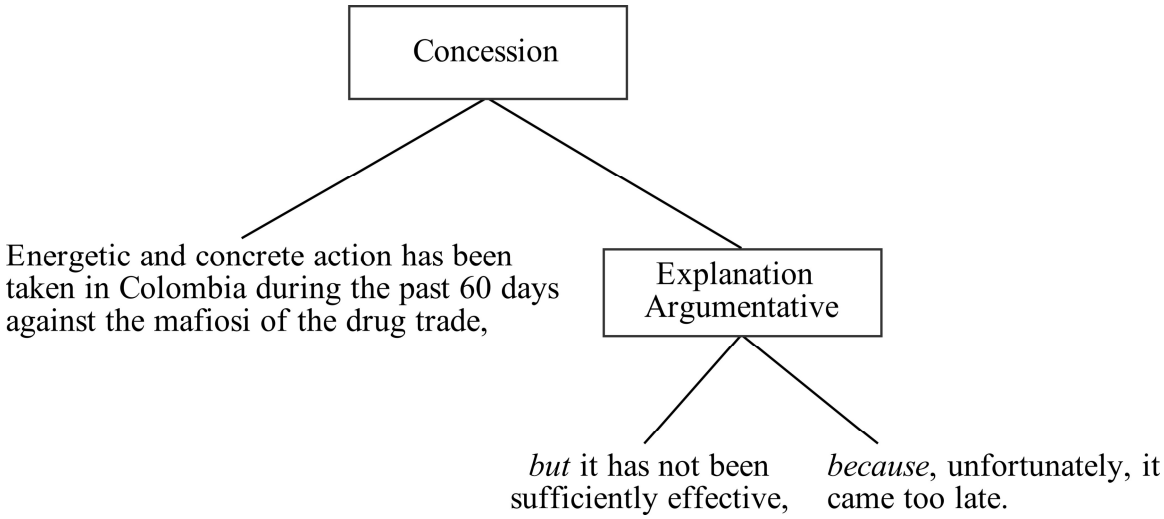


Рисунок 3.1 - Приклад речення, розібраного у формі RST

Оскільки це дослідження є частиною більшого проекту з визначення обґрунтування в письмовому дискурсі, ми зосереджуємось на трьох відносинах обставин, оцінки та розробки, які зазвичай присутні в обґрунтуваннях [28]. З метою запропонувати підхід, заснований на репліках, до виділення риторичних зв’язків, ми провели деякі експерименти на основі корпусів над анотованими корпусами RST. В результаті цих експериментів ми створили список ключових n-грам як потенційних лексичних ознак для кожного відношення. Такий метод на основі корпусу може призвести до виявлення недостатньо вивчених лексичних ознак.

Незважаючи на те, що лексичні ознаки можна використовувати для позначення риторичних відносин, вони не завжди є однозначними [20]. Деякі мовні вирази можуть функціонувати або не функціонувати як лексичний

підказка, або вони можуть сигналізувати про різні типи відносин у різних реченнях. Тому тут ми пропонуємо імовірнісну модель на основі графів, яка враховує синтаксичні особливості речень. Ці моделі мають на меті визначити, в якому синтаксичному контексті лексичний сигнал дійсно сигналізує про наявність певного відношення.

Наведено та обговорено результати оцінки нашого підходу для відношення обставин. обставина вибирається як відношення фокусу, оскільки [11] виявило, що підходи, засновані на репліках, можуть бути добре підходять для виявлення обставин у різних жанрах, тоді як зв'язок розробки зазвичай не сигналізується [11]. Крім того, особливості основного текстового жанру можуть суттєво впливати на те, як оцінюється налед [11].

Більшість досліджень, зосереджених на аналізі дискурсу та класифікації дискурсу, повідомляють про результати як явних, так і неявних відносин [23, 27, 26]. Серед робіт, які особливо зосереджені на розшифровці лексичних ключів, велика частка проведена на Penn Discourse TreeBank (PDTB) [21], тоді як для вивчення інших теорій і рамок дискурсу було проведено менше досліджень.

Анотація PDTB є лексично обґрунтованою та теоретично нейтральною щодо структури дискурсу вищого рівня [22]. У ході анотації анотаторам було запропоновано шукати лексичні елементи, які можуть сигналізувати дискурсні відносини, а потім анотувати відповідні аргументи та відносини [22]. Навіть для неявних відношень анотаторів просили шукати сусідні речення, у яких відсутній один із цих сигналів. Коли можна було зробити висновок про дискурсне відношення, їх попросили спочатку позначити відношення лексичним елементом, який міг би служити сигналом, а потім анотувати значення відношення. Такий лексично орієнтований підхід до анотування дискурсивних відносин спонукав до роботи над розшифровкою лексичних ознак у PDTB.

Наприклад, в [19] використали набір синтаксичних особливостей разом із моделлю максимальної ентропії, щоб усунути три ознаки дискурсу про те, як, відколи й коли. Їхній набір ознак включає форму допоміжного слова have, форму допоміжного слова be, форму голови та наявність модального. Вони отримали точність класифікації в 75,5% з урахуванням трьох можливих відчуттів, 71,8% для врахування чотирьох можливих відчуттів і 61,6% для трьох можливих відчуттів.

В [20] також використовували набір синтаксичних ознак, щоб усунути неоднозначність сигналів дискурсу щодо їхнього дискурсного та недискурсового використання, а також розмежування сенсу. Їх особливості фокусування складаються з синтаксичної категорії самого маркера, його безпосереднього батька, а також його лівих і правих братів і сестер. Крім того, дві бінарні ознаки враховуються, щоб вказати, чи правий брат/сестра містить VP і чи правий брат/сестра містить слід. Їх найкращий набір функцій також включав функції парної взаємодії між сигналами та синтаксичними функціями, а також попарні взаємодії між синтаксичними функціями.

Незважаючи на те, що RST є однією з найбільш широко прийнятих платформ для аналізу дискурсу, відносно мало уваги приділено анотованим корпусам RST щодо аналізу лексичних сигналів і тлумачення. На відміну від PDTB, анотації після RST не є лексично обґрунтованими, і кожне відношення визначається з точки зору намірів, які спонукають авторів використовувати ці специфічні відносини [24]. Таким чином, діаграма RST представляє деякі цілі або наміри авторів щодо включення кожної частини тексту [24]. Такі атрибути анотацій RST роблять складним завданням вивчення ролі лексичних одиниць у класифікації відношень і потенційне їхнє визначення неоднозначності.

В [17] спробував створити алгоритм риторичного аналізу, який бере необмежений вільний текст і генерує дійсне дерево риторичної структури. Корпусне дослідження було проведено, щоб зрозуміти, як можна використовувати сигнали для визначення елементарних дискурсивних одиниць

і висунення гіпотези щодо їх відповідного співвідношення. Використовуючи попередні дослідження з аналізу дискурсу, він створив список із 450 ознак дискурсу для початку. Потім із корпусу Брауна було зібрано в середньому 17 текстових проміжків, пов'язаних з кожною реплікою.

Потім усі речення були анотовані двома наборами метаданих: інформація, пов'язана з дискурсом (наприклад, маркер, використання та позиція) та алгоритмічні функції. Використовуючи ці анотації, які здебільшого фіксують орфографічне середовище сигналів, набір регулярних виразів був створений вручну для розпізнавання потенційних сигналів дискурсу. Якщо репліка дискурсу мала різні функції дискурсу в різних орфографічних середовищах, для кожного випадку створювався окремий регулярний вираз. Алгоритм привів до 84,9% F-показу від відкликання 80,8% і точності 89% для підзавдання ідентифікації сигналу. Для підзавдання класифікації відношень вони досягли відкликання 47,0% і точності 78,4%, що призвело до оцінки F 58,76%.

У [4] набір неоднозначних сигналів дискурсу вперше витягується з бази даних іспанських сигналів дискурсу. Контекст кожного сигналу потім витягується з RST Spanish Treebank і передається синтаксичному аналізатору.

Синтаксичні особливості контексту кожної репліки потім вручну аналізуються, щоб визначити потенційні мовні закономірності та закономірності. Використовуючи результати аналізу, розробляються лінгвістичні правила, щоб усунути неоднозначність лексичних ознак. Оцінюючи свої правила на тестовому корпусі RST Spanish TreeBank, вони досягли точності 60,65%.

Обидві роботи над анотаціями RST є напівавтоматичними та включають кроки вручну. Наша робота має на меті забезпечити повністю автоматизований підхід до виявлення потенційних лексичних сигналів, які можуть вказувати на риторичні відносини, і проаналізувати, чи може їхній синтаксичний контекст бути корисним для розшифровки сигналів і класифікації сенсу.

3.2. Опис пропонованого підходу

Ми використовували два анотовані людиною корпуси як наші базові набори даних для експериментів: корпус RST [3] та набір даних огляду SFU [25]. Обидва корпуси анотовані в структурі RST і створені за допомогою RSTTool.

Корпус RST, який протягом багатьох років був доступним Консорціумом лінгвістичних даних, включає 385 статей Wall Street Journal і охоплює понад 178 000 слів. Серед екземплярів відносин у корпусі RST існує близько 700 випадків обставин, що становить майже 3% від загальної кількості екземплярів відносин.

Рецензійний корпус SFU — це колекція з 400 оглядових документів з фільмів, книг та товарів народного споживання. Цей набір даних містить понад 303 000 слів і був зібраний у 2004 році з веб-сайту Epinions. Існує близько 1300 випадків обставин, що становлять майже 7% анотованих випадків у корпусі.

Вибір лексичного сигналу

Текст новин має добре структурований офіційний стиль написання, тоді як онлайн-огляди відносно менш структуровані та неформальні, написані користувачами з широким спектром письмових здібностей. Тому для виділення лексичних ознак, пов’язаних із даним відношенням, ми використовували корпус RST.

Спочатку всі екземпляри відношень витягуються з корпусу RST і збираються в документ відношення, названий на честь відповідного відношення. Потім, дотримуючись підходу, запропонованого в [1], усі n-грами (аж до триграм) витягуються із складеного документа відношення. Для кожного n-грама потім обчислюється змінена версія метрики tf-idf. Показник idf все ще обчислюється на основі кількості документів, які містять n-грам, та загальної кількості документів у корпусі. Однак, оскільки кожен рядок відповідає одному екземпляру відношення, метрика tf розраховується на основі кількості рядків,

які містять принаймні один екземпляр n-грама. Ця змінена метрика дозволяє нам компенсувати потенційне упередження, яке може бути викликано метрикою tf для слів, які з'являються більше одного разу в екземплярі відношення.

Потім список витягнутих n-грам (тобто лексичних ознак) фільтрується, щоб включати лише n-грами з їх tf-idf вище 0,5. Щоб відфільтрувати будь-які специфічні для корпусу n-грами, які можуть з'явитися в списку, n-грами, витягнуті з корпусу RST, застосовуються до набору даних огляду SFU, щоб ідентифікувати відповідне відношення. Потім F-показник кожного n-грама обчислюється незалежно. Нарешті, n-грами з F-показом вище 0,1 вибираються як потенційні лексичні ознаки. Нижче наведено покрокове пояснення процесу для відношення обставин, що призвело до виділення семи лексичних ознак: Коли, після, на, перед, з, поза, як. Ці результати детально проаналізовані в іншому документі.

• Витягніть усі екземпляри обставин із корпусу RST та сформуйте документ про обставини

• Обчисліть змінену версію tf-idf для всіх n-грам (до триграмів), які з'являються в документі обставин.

• Опрацюйте список n-грам, щоб видалити ті, у яких tf-idf нижче певного порогу.

• Використовуйте оновлений список n-грам, щоб витягти обставини з корпусу огляду SFU та самостійно обчислити їх F-показ.

• Виберіть остаточні сигнали з F-оцінкою вище певного порогу.

Наш підхід до визначення неоднозначності в основному натхненний роботою [6] щодо виявлення речень із установками. Під час їхнього дослідження фрагмент тексту, який містить другий займенник, спочатку виділяється як найбільш релевантна частина речення. Потім ці фрагменти представляють за допомогою різних шаблонів, фіксуючи їх синтаксичні особливості та семантичну орієнтацію. Для кожного типу шаблонів моделі

графіків будуються на основі пропозицій з ставленням і без нього. Нарешті, ймовірність створення нового речення на основі цих моделей використовується для прогнозування існування ставлення. Ми скористалися їхнім підходом для розмежування лексичних ознак. Наші графічні моделі побудовані на корпусі RST та оцінені на корпусі огляду SFU і навпаки. Тому процедура побудови графіка, описана нижче, проводиться для обох базових корпусів.Data Collection

Для кожного витягнутого сигналу ми спочатку створюємо два відповідних документи з анотованих корпусів. Один документ складається з усіх екземплярів відношень, які містять сигнал і анотовані відношенням фокусу (наприклад, усі екземпляри обставин, про які сигналізує коли). Відтепер такі випадки називатимуться позитивними. Інший документ складається з усіх екземплярів відношень, які містять сигнал і анотовані як будь-яке відношення, за винятком відношення фокусу (наприклад, усі екземпляри не-обставин, які містять коли). У решті цього рукопису ми називатимемо ці випадки негативними.

RST постулює ієрархічну структуру в тексті, де екземпляр відношення може бути вбудований в інші екземпляри. Тому під час вилучення екземплярів ми переконалися, що не збираємо негативні екземпляри, які включають будь-який позитивний чи негативний підекземпляр. Ми також переконалися, що не збираємо жодних позитивних екземплярів, які включають негативні підекземпляри. Включення таких вбудованих екземплярів призвело б до надлишкових і неправильних точок даних.

Під час збору негативних екземплярів було виявлено, що цей екземпляр був вбудований у десять негативних екземплярів. Однак, оскільки фактично функціонує як ознака обставин у всіх з них, ці десять випадків не могли кваліфікуватися як негативні випадки, тому були виключені.

Після створення документів кожен екземпляр обробляється і перетворюється на два різних представлення, фіксуючи синтаксичні особливості екземпляра. Щоб створити перше синтаксичне представлення, слова в екземплярах

замінюються відповідними тегами Part-OfSpeech (POS), а сам сигнал залишається таким, як є. Друге представлення включає найкоротший шлях від кореневого елемента до сигналу в дереві розбору залежностей.

Ми використовували інструментарій OpenNLP для токенізації та POS-тегування екземплярів, а аналізатор залежностей Стенфордського університету для створення дерев розбору [12].

Графове моделювання

Ми закодували синтаксичну інформацію екземплярів у графічних моделях. Будуємо орієнтований зважений графік G = (V, E),w, де:

• V — набір усіх можливих лексем, які можуть з'являтися в представленнях. Наприклад, для представлень POS, V є об'єднанням набору всіх POS-тегів і набору сигналів. • E = V × V — множина всіх можливих упорядкованих переходів між будь-якими двома лексемами.

• w → [0 − 1] — це зважувальна функція, яка призначає значення ймовірності ребру (i, j), що представляє ймовірність переходу від лексеми i до лексеми j.

Враховуючи набір синтаксичних уявлень, ймовірність переходу від лексеми i до лексеми j розраховується за оцінкою максимальної правдоподібності. Таким чином, ймовірність обчислюється шляхом ділення кількості разів, коли токен і одразу слідує за маркером j, на кількість разів, коли сам маркер і з'являється в наборі.

Цей метод побудови графіків подібний до мовного моделювання, але проводиться за набором синтаксичних уявлень [6]. Для кожного виду представлення ми будуємо один графік на основі набору позитивних екземплярів і один на основі набору негативних екземплярів. В результаті, враховуючи сигнал (наприклад, коли) і відповідне йому відношення (наприклад, обставина), ми будуємо чотири моделі графіка на основі наступних наборів: POS-представлення позитивних екземплярів, POS-подання негативних

екземплярів, представлення позитивних екземплярів, аналізованих залежностями. екземпляри та залежність аналізували уявлення заперечних речень.

Нарешті, для нашої остаточної моделі розшифровки сигналів ми використовуємо значення ймовірності, отримані з наших моделей графіків, як набір функцій для стандартної моделі машинного навчання. За даними екземпляра та графіка ми обчислюємо ймовірність того, що його синтаксичні уявлення будуть згенеровані з відповідних синтаксичних графіків. Ймовірність того, що синтаксичне представлення R, яке складається з послідовності лексем T1,T2,...,Tn, буде сформовано з графіка G, оцінюється за допомогою наступної формули. Зауважте, що W – це зважувальна функція або функція переходу ймовірності.

3.3. Оцінка підходу візуалізації

Враховуючи, що у нас є чотири моделі графіків, ми можемо створити чотири значення ймовірності як наш набір ознак. Ці функції далі використовуються в стандартному алгоритмі навчання з наглядом, щоб усунути неоднозначність сигналу та класифікувати відношення даного екземпляра. На рисунку 3.2 наведено високорівневий опис всієї процедури виділення та розшифровки лексичних ключів.

З огляду на те, що наша кінцева мета полягає в тому, щоб виявити обґрунтування з письмового дискурсу, наш підхід оцінюється для відношення обставин як єдиного відношення на основі репліки, яке, як відомо, часто присутнє в обґрунтуваннях [11, 28]. Ми проводили експерименти з використанням різних форм уявлень POS на основі кількості POS-тегів, що оточують сигнал, і деталізації тегів. Ми проводили експерименти, використовуючи весь екземпляр POS-тегів, використовуючи два POS-теги до і два теги після сигналу, а також використовуючи один до і один після сигналу.

Ми також використовували три рівні деталізації POS-тегів, включаючи найтонший набір тегів Penn English Treebank, який використовується OpenNLP.

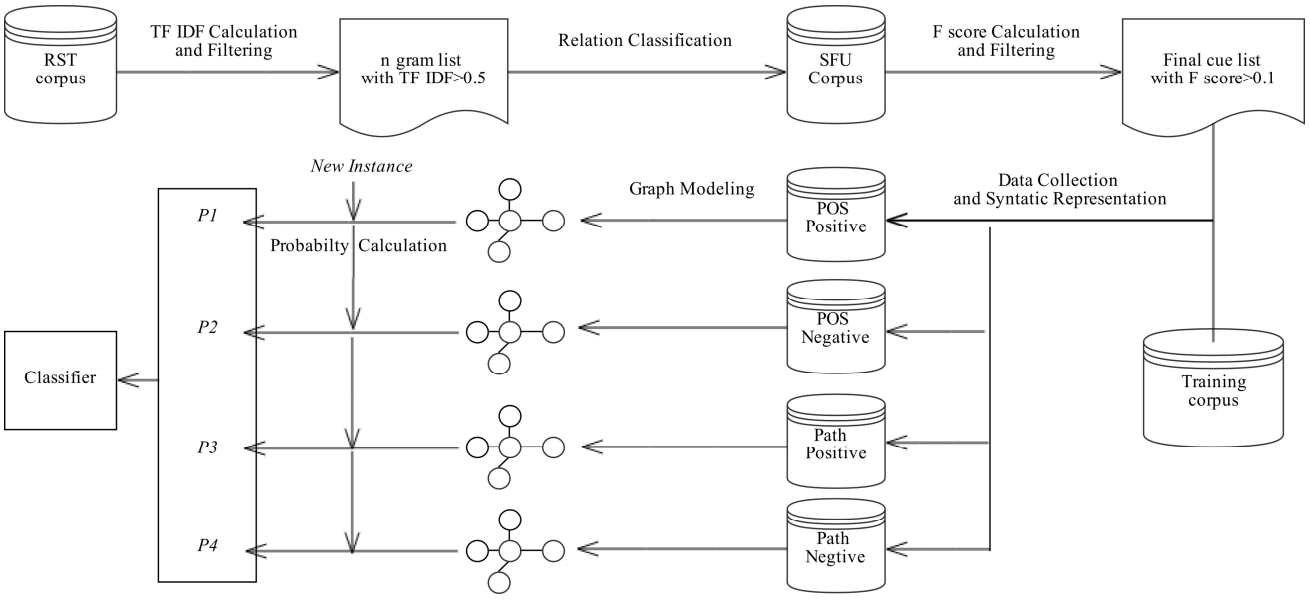


Рисунок 3.2 - Огляд високого рівня підходу до виділення сигналу та розмежування

Таблиця 3.1 – Додаток до POS-тегів у наборі тегів Penn English Treebank

Label	Medium Granularity	Coarse Granularity
JJ	JJ, JJR, JJS	JJ, JJR, JJS, DT, WDT
NN	NN, NNS, NNP, NNPS	NN, NNS, NNP, NNPS
PRP	PRP, PRP$	PRP, PRP$, WP, WP$
RB	RB, RBR, RBS	RB, RBR, RBS, WRB
WP	WP, WP$	-
VB	VB, VBD, VBN, VBP,	VB, VBD, VBN, VBP,
	VBZ	VBZ, MD, VBG

Ми також використовували середню та грубу деталізацію, які створюються шляхом категоризації подібних тегів в один високорівневий тег. Таблиця 3.1 показує теги, які класифікуються на кожному з цих двох рівнів деталізації. Зауважте, що теги, не згадані в рівнях деталізації, використовуються як є.

Використання цих трьох варіантів двох атрибутів POS-тегів призвело до дев’яти різних налаштувань експерименту. Ми досягли найкращих результатів на обох корпусах, використовуючи один тег до і один тег після сигналу та середній рівень деталізації.

Остаточний алгоритм, побудований на значеннях ймовірності, виконується та оцінюється за допомогою Weka Workbench та класифікує екземпляри за допомогою регресії. Для оцінки моделі використовується стратифікований підхід десятикратної перехресної перевірки. Щоб отримати уявлення про ефективність нашої моделі у розрізненні різних сигналів, результати звітуються для кожного з семи ознак незалежно. Фільтр SMOTE використовувався, коли ми зіткнулися зі значним дисбалансом класів.

Два основних корпуси, використані в цьому дослідженні, належать до дуже різних текстових жанрів. Крім того, оскільки RST робить акцент на намірах письменника та ефекті відношення на читача [24], анотації RST за своєю суттю є суб’єктивними і засновані на розумінні тексту читачами [24]. Тому між двома корпусами можуть бути відмінності через потенційно різні знання, якими володіє кожен набір анотаторів щодо культури, ситуації та мови, які представляє текст [24]. Незважаючи на жанрову диспропорцію та різницю в анотаціях, за допомогою запропонованої моделі ми отримали обнадійливі результати. Однак очікується, що результати покращаться, коли моделі будуть побудовані на корпусах подібних жанрів і анотовані з використанням основних правил істини.

Вивчення та аналіз риторичних відносин, як будівельних блоків узгодженості в дискурсі, може сприяти розробці складних лінгвістичних додатків і алгоритмів. З метою полегшення автоматичного виявлення явних риторичних зв’язків у тексті, ми розробили алгоритм, щоб спочатку виявити потенційні лексичні ознаки, а потім усунути їх, передбачивши відношення.

Змінена версія tf-idf була використана для виділення потенційних сигналів, а модель на основі графіка, побудована на синтаксичних ознаках, була

використана для вирішення завдання розшифровки сигналу. Загалом, результати оцінювання вказують на ефективність синтаксичних ознак у розкритті лексичних ознак та передбаченні явних риторичних зв'язків у різних жанрах. Наші експерименти показали перевагу нашого підходу до кодування таких синтаксичних ознак у імовірнісному графі порівняно з їх прямим використанням.

Як згадувалося раніше, це дослідження є нашою першою спробою визначити обґрунтування в тексті. Обгрунтування — це пояснення причин, що лежать в основі рішень, висновків та тлумачень. Попередні дослідження обґрунтування формулювання та обміну свідченнями свідчать про те, що це сприяє контролю якості, управлінню знаннями та повторному використанню знань [30, 29]. Проте було лише кілька спроб застосування обчислювальних методів для визначення обґрунтування погано структурованого тексту [1, 5, 2]. Наші майбутні дослідницькі зусилля спрямовані на розробку алгоритмів для виявлення легкосигнальних і неявних відносин, а також на подальше дослідження потенціалу та обмежень використання таких риторичних зв'язків у визначенні обґрунтування в тексті дискурсу. Крім того, ми плануємо оцінити розширюваність нашого підходу до PDTB та безпосередньо порівняти його з належним базовим рівнем серед попередніх досліджень щодо PDTB.

3.4. ProjectTales: Доказ концепції

За останні кілька десятиліть було розроблено і розроблено багато програмних програм для вирішення різних аспектів управління проектами, таких як оцінка зусиль проекту [4], управління спільною діяльністю [13] та моніторинг змін і ризиків проектів [7]. У цьому дослідженні ми зацікавлені в наданні допомоги керівникам проектів у прийнятті ефективних рішень щодо змін, пов'язаних з проектом. Проекти не проводяться у вакуумі і зазвичай на них впливають різні динамічні фактори, такі як наявність бюджетних ресурсів,

рівень пріоритету в організації та члени проекту. Отже, очікується, що менеджерам проектів зазвичай потрібно вносити зміни в проекти під час процесу управління.

Дослідники досліджували комп'ютерні засоби, які допомагають керівникам проектів приймати обґрунтовані рішення, коли проекти потрібно адаптувати до нової ситуації.

Наш підхід допомагає менеджерам проектів приймати рішення щодо змін, представляючи інтерактивне візуальне уявлення про історію змін попередніх проектів. Наше припущення полягає в тому, що, зіткнувшись із ситуацією змін, менеджери проектів можуть стати більш ефективними

Тому наш дизайн інтерактивного інструменту візуалізації був зосереджений на представленні та створенні асоціацій 1) між історією змін попередніх проектів, їх причинами та обґрунтуваннями їхніх рішень; 2) між рішеннями та статусом проекту як індикатором впливу рішень на проект. У цьому плакаті ми описуємо наш поточний прототип системи ProjectTales. У решті розділів ми спочатку обговорюємо пов'язані роботи, а потім представляємо обґрунтування дизайну прототипу. Ми закінчуємо нашими планами оцінки користувачів і підсумком дослідження.

Після огляду інструментів управління проектами, таких як TeamSpace і TeamSCOPE запропонував рекомендації щодо проектування інструменту управління проектами для розподілених команд дизайнерів. Вони стверджували, що для полегшення повторного використання знань з управління проектами з попередніх проектів важливо заархівувати як знання про продукти, так і знання про процес. Щоб проілюструвати свій підхід до конфігурації інформаційних систем управління проектами, представили, як знання попередніх запитів на зміни можна поєднати із запитом у поточному проекті. У їхньому підході запит має дві конфігурації в системі: атрибути запиту (наприклад, пріоритет, опис, залишок, термін виконання) і робочі процеси стану (наприклад, закрито, відкрито, погоджено). Коли зроблено запит у

поточному проекті, система надасть запропоновані описи атрибутів і робочих процесів стану на основі попередніх проектів, тобто сховища знань з управління проектом.

У порівнянні з системою, представленою в [5], яка полегшує конфігурацію поточних змін на основі історії попередніх проектів, наша система зосереджена на полегшенні процесів прийняття рішень керівниками проектів, використовуючи знання про зміни з попередніх проектів. . Крім того, ми використовуємо методи інтерактивної візуалізації, щоб дозволити керівникам проектів переглядати та витягувати інформацію про зміни, що стосується поточної ситуації проекту, а також міркувати про зв'язок між змінами та їх впливом на проекти.

Наш прототип, так званий ProjectTales, розроблений і побудований на базі даних історії програмного забезпечення з відкритим вихідним кодом EGroupware. Програмне забезпечення призначене для компаній, щоб керувати контактами, зустрічами, проектами та списками справ. Це активно використовувана програма з надійною репутацією користувача.

База даних історії надає різноманітні атрибути для проектів (наприклад, назва, опис, пріоритет, використаний бюджет) і кілька атрибутів для змін (наприклад, проект, змінений атрибут проекту, мітка часу). Ми змінили цю базу даних історії, щоб включити причини змін. Завдяки причинам, доступним керівникам проектів, вони можуть легко отримати рішення щодо змін, які відповідають тим же причинам, що й поточна ситуація. Під час модифікації проекту менеджерів проекту просять призначити одну або кілька причин зміни. Ці причини можна вибрати із попередньо визначеного списку потенційних причин, витягнутих з літератури [2, 8, 12], включаючи зовнішні причини (наприклад, політичні, економічні, потреби клієнтів, ринкова сила), а також внутрішні причини (наприклад, дефіцит персоналу або бюджету, стратегічні рішення, контроль якості). Щоб дозволити керівникам проектів адаптувати

причини до своєї конкретної організаційної ситуації, їм дозволено додавати нові до списку.

Стверджуючи про важливість надання обґрунтування змін керівнику проекту, ми також включили цю інформацію до бази даних. У нашому дослідженні ми визначаємо обґрунтування зміни як інформацію, яка обґрунтовує рішення про зміну. Соціальні психологи досліджували види інформації в спілкуванні, які впливають на ставлення людей [1, 6]. Ці дослідження показали, що поширена інформація, яка є правдоподібною та логічною та додає щось нове до проблеми, швидше за все, спричинить зміну ставлення [5, 3]. В [9] виявили, що коли люди мали більший доступ до різних поглядів, якість аргументації повідомлення мала більший вплив на переконання. Ми віримо, що представлення не тільки рішень щодо змін, але й їх обґрунтування може бути корисним для менеджерів проектів при вирішенні поточної ситуації. Дослідження [10] вказує на те, що на навички міркування людей впливає чітке обґрунтування завдання під час групової діяльності. Очікується, що завдяки архівації та поширенню обґрунтування змін у проектах організації керівники проектів покращать свої навички міркування щодо прийняття рішень у довгостроковій перспективі. Покращені навички міркування можуть позитивно вплинути на якість рішень щодо змін, оскільки рішення приймаються після кращих процесів міркування.

3.5. Дизайн інтерфейсу

Інтерфейс ProjectTales розділений на два окремі компоненти, розташовані вертикально на екрані: компонент огляду історії та компонент детального перегляду проекту. Кожен компонент системи надає певні скоординовані уявлення про різні виміри базового набору даних, підтримуючи менеджерів проектів у глибшому уявленні про історичні дані [10]. На рисунку 3.3 показано скріншот системи

Компонент огляду історії складається з трьох різних уявлень, а саме: сітка історії, панель причин і панель обґрунтування. Сітка історії призначена для наочного огляду змін, внесених у різні атрибути різних проектів. У цьому поданні кожен рядок представляє проект у базі даних, а кожен стовпець — атрибут.

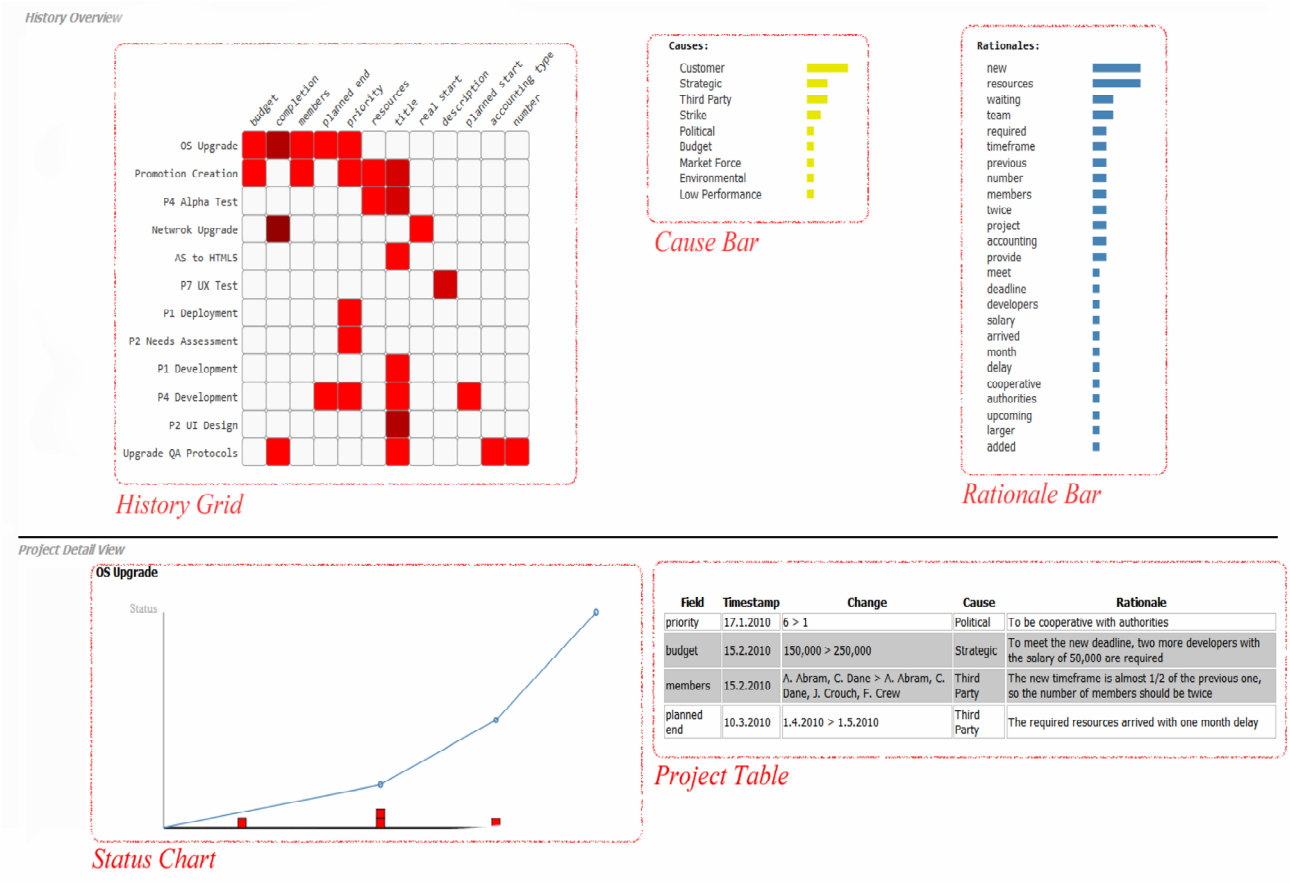


Рисунок 3.3 - Знімок екрана ProjectTales

ProjectTales: інтерактивний інструмент візуалізації проект. Оскільки колір є візуальною ознакою попередньої уваги і може бути окремо декодований від просторового положення зоровою системою людини [11], кольорове кодування використовується для представлення кількості змін в різних атрибутах проектів. Якщо зміни не відбулося, клітинка забарвлюється в світло-сірий колір. У разі зміни кількість змін кодується за допомогою каналу яскравості червоного кольору; таким чином, клітинки з високою частотою зміни виглядають темнішими, а клітини з низькою частотою змін — яскравішими.

Сітка історії додатково збагачується механізмом взаємодії, що дозволяє менеджерам проектів сортувати проекти за кількістю змін у певному атрибуті, клацнувши мітку атрибута. Аналогічно, клацання на мітці проекту призводить до сортування списку атрибутів відповідно до кількості змін для цього проекту. Таке візуальне уявлення дозволяє менеджерам проектів швидко отримати уявлення про частоту змін у різних проектах і різних атрибутах, а також дослідити цю інформацію для проектів і атрибутів, що цікавлять, за допомогою функції сортування.

На додаток до сітки історії, компонент огляду історії пропонує дві вертикальні гістограми, щоб наочно відобразити огляд причин та обґрунтування змін. Панель причин представляє відсортований вигляд причин від найчастішої до найменш поширеної, тоді як панель обґрунтування показує найчастіші слова, які використовуються в обґрунтуваннях змін. Оскільки стоп-слова не мають значення при дослідженні визначних слів у обґрунтуванні, вони вилучаються з панелі обґрунтування. Історична сітка тісно пов’язана з причиною та планкою обґрунтування за допомогою техніки зв’язування та розчісування. Наведення курсору на кожну клітинку в сітці історії виділяє відповідні причини й обґрунтування в обох гістограмах. Крім того, якщо менеджери проекту наводять курсор на причину або обґрунтування, відповідні клітинки в сітці історії будуть виділені.

Вивчаючи огляд історії змін, менеджери проекту можуть бути зацікавлені детально зосередитися на конкретному проекті. Подвійне клацання на мітці проекту завантажує компонент детального перегляду проекту з двома пов’язаними видами діаграми рядка стану та таблиці проекту. На діаграмі рядка стану вісь x представляє тривалість проекту як часову шкалу. Ця вісь доповнена червоними гліфами, кожен з яких представляє зміни, що відбулися в проекті в певний час. Представлення змін на часовій шкалі може дозволити менеджерам проекту дослідити послідовність змін і побачити, як конкретна зміна викликала наступні зміни. Вісь Y діаграми відображає статус проекту, що дозволяє

керівникам проекту оцінити вплив певної зміни або серії змін на статус проекту.

Нарешті, таблиця проекту надає менеджерам проекту детальну інформацію про зміни, включаючи змінений атрибут, час зміни, старі та нові значення, причини зміни та повне обґрунтування, у текстовому форматі. Ці два уявлення також пов’язані з технікою зв’язування та використання пензля, оскільки наведення курсора миші на гліф у діаграмі чи рядок у таблиці викликає виділення відповідного представлення змін в іншому поданні. Ці два представлення дозволяють менеджеру проекту визначити потенційно цікаву зміну в одному поданні, а потім виявити ту саму зміну в іншому поданні для подальшого дослідження.

ВИСНОВКИ

В дипломній рооті було проведено оціночне дослідження прототипу в контрольованих лабораторних умовах щодо візуалізації програмних сервісів. Це дослідження розроблено як дослідження в межах предмета, в якому учасникам пропонується виконати різні завдання щодо прийняття рішень з кожним інтерфейсом (ProjectTales і базова система на основі таблиць). Обидва завдання мають однаковий рівень складності і складаються з трьох підзадач щодо отримання та інтерпретації інформації про зміни та обґрунтування змін. Ми набирали менеджерів проектів для використання ProjectTales та порівняння їх із базовою системою.

У цьому дослідженні було проведено аналіз, щоб зробити висновок, коли обрана техніка графічної візуалізації та подальші вдосконалення застосовуються до попередньо розробленої текстової системи виявлення вимог, це дає краще розуміння користувачам системи та зменшує час і зусилля, які їм потрібні. витратити на отримання бажаних вимог.

Це дослідження було виконано в кілька етапів. Спочатку була вивчена попередня текстова інтерактивна система виявлення вимог і на основі базових концепцій розробки цієї системи був розроблений та реалізований найбільш підходящий графічний інтерфейс як з точки зору семантики, так і з точки зору зручності використання. Крім того, було проведено дослідження зручності використання на групі студентів з різним академічним досвідом, щоб підтвердити, що запропонований дизайн може позитивно покращити зручність використання інтерфейсу користувача.

Результати дослідження показують, що в цілому користувачі мали позитивну думку щодо використання як графічного, так і текстового інтерфейсу. Однак на основі думок користувачів у середньому всі параметри юзабіліті, як-от навчальність, ефективність, частота помилок і задоволеність користувачів, були покращені в порівнянні з текстовою системою інтерфейсу. Крім того, в середньому користувачі графічного інтерфейсу виконували ті ж

завдання швидше і точніше, ніж користувачі текстової системи. Загалом більшість користувачів обох систем віддають перевагу концепції програмної автоматизації, яка є основою обох систем. над класичним процесом розробки програмного забезпечення.

Оскільки ми виявили певні протиріччя в оцінках бізнес-користувачів та їх оцінці легкості та їхньому розумінні системи, можна зробити висновок, що потрібен подальший аналіз. Також з самого початку система була розроблена для людей з певним досвідом розробки програмного забезпечення, які зазвичай бізнесмени не входять до цієї групи. Тому спочатку очікувалося, що результати студентів-комп'ютерників будуть більш послідовними, ніж студентів бізнес-класу, що можна вважати позитивним результатом, виходячи з нашої основної мети.

Незважаючи на те, що графічна візуалізація покращила зручність та якість користувальницького інтерфейсу, результати дослідження юзабіліті показують, що є ще багато місця для вдосконалення. Перш за все, інтерфейс може бути розроблений таким чином, щоб користувачі могли мати пряму взаємодію з графічним інтерфейсом замість непрямого спілкування на основі діалогу. Це дасть користувачам краще відчуття контролю над системою. Крім того, різні графічні інтерфейси можуть бути розроблені для людей з різними очікуваннями та досвідом у сфері розробки програмного забезпечення. Це може бути настільки ж просто, як перетягування для користувачів із базовими знаннями комп'ютера чи зовсім без них, або настільки ж складним, як багатошаровий тривимірний графічний інтерфейс, який навіть може проілюструвати деталі про взаємозв'язки між службами та архітектурою системи.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. R. Bastide, J. Billington, E. Kindler, F. Kordon, and K. H. Mortensen, editors. Meeting on XML/SGML based Interchange Formats for Petri Nets, Arhus, Denmark, June 2000. 21st ICATPN.

2. F. Bause, P. Kemper, and P. Kritzinger. Abstract Petri net notation. Petri Net Newsletter, 49:9–27, October 1995.

3. G. Berthelot, J. Vautherin, and G. Vidal-Naquet. A syntax for the description of Petri nets. Petri Net Newsletter, 29:4–15, April 1988.

4. B. Bos, H. W. Lie, C. Lilley, and I. Jacobs (eds.). Cascading Style Sheets, level 2 – CSS2 Specification. URL http://www.w3.org/TR/CSS2, 1998.

5. J. Clark. TREX – tree regular expressions for XML. URL http://www.thaiopensource.com/trex/. 2001/01/20.

6. J. Clark (ed.). XSL Transformations (XSLT) Version 1.0. URL http://www.w3.org/TR/XSLT/xslt.html, 1999.

7. J. Clark and M. Murata (eds.). RELAX NG specification. URL http://www.oasis-open.org/committees/relax-ng/. 2001/12/03.

8. CPN Tools. URL http://www.daimi.au.dk/CPNtools. 2001/09/11.

9. J. Desel, G. Juh´as, R. Lorenz, and C. Neumair. Modelling and validation with VipTool. In Conference on Business Process Management, Tool Presentation, 2003.

10. Design/CPN. URL http://www.daimi.au.dk/designCPN/. 2001/09/21.

11. J. Ferraiolo, F. Jun, and D. Jackson (eds.). Scalable Vector Graphics (SVG) 1.1 Specification. URL http://www.w3.org/TR/SVG11/, 2003.

12. ISO/IEC/JTC1/SC7. Subdivision of project 7.19 for a Petri net standard. ISO/IEC/JTC1/SC7 N1441, October 1995.

13. ISO/IEC/JTC1/SC7. Software Engineering - High-Level Petri Nets - Concepts, Definitions and Graphical Notation. ISO/IEC 15909-1, Final Committee Draft, May 2002.

14. ISO/IEC/JTC1/SC7 WG19. New proposal for a standard on Petri net techniques. ISO/IEC/JTC1/SC7 N2658, June 2002.

15. M. Jungel, ¨ E. Kindler, and M. Weber. The Petri Net Markup Language. Petri Net Newsletter, 59:24–29, 2000.

16. S. Katz and O. Grumberg. VeriTech: Translating among specifications and verification tools. Technical report, The Technion, Haifa, Israel, March 1999.

17. E. Kindler and M. Weber. A universal module concept for Petri nets. An implementation-oriented approach. Informatik-Berichte 150, Humboldt-Universit¨at zu Berlin, June 2001.

18. A. M. Koelmans. PNIF language definition. Technical report, Computing Science Department, University of Newcastle upon Tyne, UK, July 1995. version 2.2.

19. R. B. Lyngsø and T. Mailund. Textual interchange format for high-level Petri nets. In Proc. Workshop on Practical use of Coloured Petri Nets and Design/CPN, pages 47–63, Department of Computer Science, University of˚Arhus, Denmark, 1998.

20. T. Mailund and K. H. Mortensen. Separation of style and content with XML in an interchange format for high-level Petri nets. In Bastide et al. [1], pages 7–11.

21. The Model-Checking Kit. URL http://wwwbrauer.in.tum.de/gruppen/ theorie/ KIT/. 2003/02/18.

22. The PEP Tool. URL http://parsys.informatik.uni-oldenburg.de/~pep/.

23. The Petri Net Kernel. URL http://www.informatik.hu-berlin.de/top/pnk/.

24. Petri Net Markup Language. URL http://www.informatik.hu-berlin.de/top/ pnml/.

25. Renew: The Reference Net Workshop. URL http://www.renew.de. 2002/03/04. 26. M. Sperberg-McQueen and H. Thompson (eds.). XML Schema. URL http://www.w3.org/XML/Schema, April 2000.

27. C. Stehno. Petri Net Markup Language: Implementation and Application. In J. Desel and M. Weske, editors, Promise 2002, Lecture Notes in Informatics P-21, pages 18–30. Gesellschaft fur ¨ Informatik, 2002.

28. M. Weber and E. Kindler. The Petri Net Kernel. In H. Ehrig, W. Reisig, G. Rozenberg, and H. Weber, editors, Petri Net Technology for Communication Based Systems, Lecture Notes in Computer Science 2472. Springer, Berlin Heidelberg, 2002.

29. G. Wheeler. A textual syntax for describing Petri nets. Foresee design document, Telecom Australia Research Laboratories, 1993.

БІБЛІОГРАФІЧНА ДОВІДКА

Тема дипломної роботи бакалавра: "Інформаційні моделі візуалізації програмних сервісів та рішень"

Обсяг пояснювальної записки: 87 аркушів

Дата закінчення дипломної роботи 13 червня 2022р.

Підпис студента _____