

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Одеський національний технологічний університет
Університет Інформатики і прикладних знань, м.Лодзь, Польща
Національний технічний університет України «Київський
політехнічний інститут»
Навчально-науковий інститут комп'ютерних систем і технологій
«Індустрія 4.0» ім. П.М. Платонова

XXIII Всеукраїнська науково-технічна конференція
молодих вчених, аспірантів та студентів

«СТАН, ДОСЯГНЕННЯ ТА ПЕРСПЕКТИВИ
ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ»

Матеріали конференції



Одеса

20-21 квітня 2023 р.

Стан, досягнення та перспективи інформаційних систем і технологій / Матеріали XXIII Всеукраїнської науково-технічної конференції молодих вчених, аспірантів та студентів. Одеса, 20-21 квітня 2023 р. - Одеса, Видавництво ОНТУ, 2023 р. – 449 с.

Збірник включає матеріали доповідей учасників конференції, які об'єднані за тематичними напрямками конференції.

Збірник буде корисним як для фахівців і працівників фірм, зайнятих в області ІТ, так і для викладачів, магістрів і студентів вищих навчальних закладів, які навчаються за напрямками і спеціальностями програмного забезпечення, обчислювальної техніки і автоматизованих систем, прикладної математики та обробки інформації, буде корисним професіоналам з комп'ютерного моделювання та розробки комп'ютерних ігор.

Результати досліджень у збірнику представляють собою своєрідний зріз сучасного стану справ в перерахованих галузях знань, який може допомогти як фахівцям, так і студентам університетів скласти загальну картину розвитку інформаційних технологій та пов'язаних з ними питань.

Наукові праці згруповані за напрямками роботи конференції та наведені в алфавітному порядку прізвищ авторів.

Матеріали (тези доповідей) друкуються в авторській редакції. Відповідальність за якість та зміст публікацій несе автор.

Матеріали подано українською та англійською мовами.

Редактор збірника Котлик С.В.

7. Порівняльний аналіз сучасних шляхів діагностики складних технічних виробничих систем. Лактіонов О. (Національний університет «Полтавська політехніка») 93	93	
8. Optimization of paths, taking into account the significance of intermediate points. Мазурок І.Є., Веремйов К.В. (Одеський національний університет ім. Мечникова) 95	95	
9. Методика навчання фахівців із інформаційної безпеки соціальної інженерії з використанням OSINT і мови SIEVE. Міронов І. В., Болтач С. В. (Одеський національний технологічний університет) 97	97	
10. Дослідження факторів впливу на безпеку мобільних застосунків на прикладі клієнтської частини кіберфізичної системи розумної парковки. Павлова О.О., Авсієвич В.Р., Кузьмін А.А. (Хмельницький національний університет) 98	98	
11. Парсинг тексту: використання потужностей NLP задля підвищення точності отримуваних даних. Пелович Д. В., Смиш О. Р. (Національний університет «Києво-Могилянська академія») 100	100	
12. Захист підприємств від кібератак на корпоративні мережі. Петрук Д. С. (Волинський національний університет імені Лесі Українки) 102	102	
13. Використання мобільних застосунків у роботі з документацією ТОВ "Агрона Фрут Україна". Погоріла Ю. В. (Донецький національний університет імені Василя Стуса) 103	103	
14. Технологія HDR у моніторах. Романюк О. Н., Захарчук М. Д., Романюк О.В., Коробейнікова Т. І. (Вінницький національний технічний університет, Національний університет «Львівська політехніка») 105	105	
15. Проектування інформаційної системи управління сегрегаційним комплексом збору відходів оперативної поліграфії. Сторожук Д.І. (Українська академія друкарства) 107	107	
16. Дослідження методів перетворення повідомлень у бортових автомобільних системах. Чайковський О.Р., Селіванова А.В. (Одеський національний технологічний університет) 109	109	
17. Процес безпечної передачі інформації у мобільному додатку “Студент ЧДТУ” з Використанням Spring Security на основі JWT. Куницька С.Ю., Архіпов М.О., Чоповенко В.М. (Черкаський державний технологічний університет) 110	110	
18. Захист даних та вихідних файлів від несанкціонованого доступу та копіювання комп’ютерних відеоігор. Шаповал В.В. (Київський національний університет імені Тараса Шевченка) 112	112	
19. Програмне забезпечення для забезпечення безпеки резервного архівування даних у хмарних системах. Шевчук Р.П., Заріцький О.І. (Західноукраїнський національний університет) 114	114	
20. Вплив війни в Україні на кібербезпеку. Шередега Р.О., Бутенко Т.А. (Харківський державний біотехнологічний університет) 116	116	
21. Дослідження застосування стандартів PAPERLESS у закладах вищої освіти. Чіклікчі О.С., Лукашенко Д.О., Ольшевська О.В. (Одеський національний технологічний університет) 117	117	
22. 3-D візуалізація авторадіограмм радіоактивних частинок. Новіков А.М. (Інститут проблем безпеки атомних електростанцій Національної академії наук України) 119	119	
Розділ 3: Нові інформаційні технології в освіті		121
1. Development of a methodology for evaluating the efficiency of ship operator model. Nosov P.S., Masonkova M.M., Diahyleva P.S., Solovey O.S. (Херсонська державна морська академія) 121	121	
2. Optimization of management processes for maritime transport personnel qualification. Nosov P.S., Ponomaryova V.P., Diahyleva O.S., Ben A.P. (Херсонська державна морська академія) 123	123	
3. Using SolidWorks in modern education and science. Rudyk O.Yu., Baranov I.I., Gereta M.M., Dytynyuk V.O., Fedoryshyn S.I. (Хмельницький національний університет) 125	125	

У ході дослідження було виявлено, що один фактор може спричиняти декілька наслідків, так само як один наслідок може бути спричинений декількома факторами. Тому напрямками подальших досліджень є аналіз кореляції факторів, які впливають на безпеку мобільних застосунків та наслідків, до яких вони призводять для розробки ефективних методів підвищення безпеки мобільних застосунків.

Список використаних джерел:

1.CVEdetails.com the ultimate security vulnerability data source URL: <https://www.cvedetails.com/> (Доступ 11.03.2023)

2.A decade in, how safe are your iOS and Android apps? URL: <https://www.nowsecure.com/blog/2018/07/11/a-decade-in-how-safe-are-your-ios-and-android-apps/> (Доступ 11.03.2023)

3.Understanding OWASP Mobile Top 10 Risks with Real-world Cases URL: <https://appinventiv.com/blog/owasp-mobile-top-10-real-world-cases/> (Доступ 11.03.2023)

4. Авсієвич В., Кузьмін А. Дослідження вразливостей системи розумної парковки та способи їх усунення. Актуальні Проблеми Комп'ютерних Наук (АПКН-2022), Хмельницький, Україна, 18-19 листопада 2022. Хмельницький: ХНУ, 2022. С. 11-14.

ПАРСИНГ ТЕКСТУ: ВИКОРИСТАННЯ ПОТУЖНОСТЕЙ NLP ЗАДЛЯ ПІДВИЩЕННЯ ТОЧНОСТІ ОТРИМУВАНИХ РЕЗУЛЬТАТІВ

ПЕЛОВИЧ Д. В, СМІШ О. Р.

Національний університет «Києво-Могилянська академія

Аналіз текстів судових рішень (далі — СР) є невіддільною частиною процесів оцінювання якості, ефективності та прозорості судочинної системи України — за різними параметрами. Утім, тексти СР, як і будь-яких інших правових документів, є доволі специфічними для сприйняття. У таких документах використовуються спеціалізована лексика й терміни, складні синтаксичні структури, а також існує певний стандарт в оформленні змісту, який є відмінним від звичних. Ці фактори ускладнюють мануальний процес аналізу подібних текстів. Наразі найбільш поширеним способом отримання сирих текстів СР з бази ЄДРСР [1] є відкритий офіційний ресурс ЄДРСР ДСАУ [2] (релевантні пошукові запити, що стосуються СР, надають попередньо згаданий ресурс першим у вибірці, що свідчить про його популярність).

Парсинг [3], як інструмент аналізу, може запропонувати можливість вицнення інформації з СР та подальше формування з неї певної уніфікованої структури (напр., у форматі JSON чи XML) задля отримання потрібної інформації з СР. Для *парсингу* використовуються різні підходи. Одним із таких підходів є *rule-based approach* [4], який орієнтується на певні закономірності в розташуванні слів у тексті, граматику тощо. Зазвичай, такі правила визначаються людьми емпірично після попереднього огляду певної вибірки текстів. У цьому випадку, проблема полягає в тому, що практично неможливо передбачити всі можливі комбінації розташування слів у реченні, словоформи, граматичні складові тощо. Цей чинник безпосередньо впливає на отримувані результати *парсингу* — менша кількість і досконалість визначених правил дає зменшення точності результатів.

З метою вдосконалення програми задля збільшення відсотка випадків отримання коректних результатів, пропонується інтеграція NLP [5]. Natural Language Processing (далі — NLP) передбачає розробку алгоритмів і обчислювальних моделей, які можуть обробляти, розуміти та відтворювати природну людську мову. NLP має широкий спектр застосування. Розглянемо саме такі субкатегорії NLP, як Named Entity Recognition (далі — NER) і Grammatical Tokenization and Tagging (далі — GTT).

Якими є варіанти застосування наданого інструментарію? До прикладу, поставлена задача проаналізувати, якої статі найчастіше є позивач, а якої — відповідач. Розв'язати задачу лише за допомогою *rule-based approach* доволі складно, адже специфіка текстів СР послуговується такими словами, як *особа, позивач, відповідач, свідок* тощо, що об'єктивно не дозволяє безперешкодно визначити стать фігурантів. Натомість можна скомбінувати кілька моделей в один функціонал. Розглянемо українську NER-модель [6], надану бібліотекою Hugging Face Transformers [7], і GTT-модель UDPipe 2.10 [8]. За допомогою NER-моделі можна виокремлювати в текстах певних осіб. Своєю чергою, GTT-модель може визначити зв'язки слів у реченнях і зв'язки між реченнями таким чином, що можна буде прирівняти ПІБ фігуранта до його ролі в СР, і, зважаючи на ПІБ фігуранта, стверджувати, що останній є чоловіком або жінкою. Важливо зазначити, що для реалізації поставленої задачі важливо узгодити з відповідними інституціями надання вибірки СР, які містять неконфіденційні дані фігурантів, тобто такі, що охоплюють ПІБ фігурантів (згідно ст. 31-32 КУ [9], ст. 182 ККУ [10] та Закону України «Про захист персональних даних» [11]).

Інший приклад: необхідно статистично визначити, у які дні інституції продукують найбільшу кількість проваджень. Для такої задачі використаємо згадану раніше NER-модель, яка також має функціонал розпізнавання дат. Це фактично весь функціонал моделі, який потрібен для вирішення задачі. Усе, що залишається, це ввести обмеження на опрацювання лише верхніх титульних частин документів (зазвичай, дата оприлюднення СР розташована саме там). Це обмеження є важливою частиною, адже у СР є й інші дати, які не стосуються факту оприлюднення. Далі потрібно за вичленими датами визначити дні тижня для кожного СР, згрупувати ці дані й сформувати відповідний графік співвідношення дня тижня з вихідною кількістю СР.

Підсумовуючи, варто зазначити, що використання наявних потужностей NLP, зокрема NER і GTT, може стати вирішальним фактором у підвищенні точності отримуваних результатів, порівняно з такими підходами, як *rule-based approach*.

Список використаних джерел

1. Закон України від 22 грудня 2005 року № 3262-IV «Про доступ до судових рішень».
2. Єдиний державний реєстр судових рішень. ЄДРСР ДСАУ. URL: <https://reyestr.court.gov.ua/>.
3. Nilesh Parashar. (2018). What is Parsing in NLP? URL: <https://medium.com/@niitwork0921/what-is-parsing-in-nlp-69f41b3e5620>.
4. Suresh Krishnan. (2021). Data with AI, Rule-Based & ML. URL: <https://www.linkedin.com/pulse/data-ai-rule-based-ml-suresh-krishnan>.
5. What is NLP? | IBM - Deutschland. URL: <https://www.ibm.com/topics/natural-language-processing>.
6. Auto Classes. Transformers. Hugging Face – The AI community building the future. URL: https://huggingface.co/docs/transformers/model_doc/auto.
7. Transformers. Hugging Face – The AI community building the future. URL: <https://huggingface.co/docs/transformers>.
8. UDPipe 2 Models | ÚFAL. URL: https://ufal.mff.cuni.cz/udpipe/2/models#universal_dependencies_210_models.
9. Статті 31-32, Конституція України.
10. Кримінальний кодекс України, стаття 182.
11. Закон України від 27 жовтня 2010 року № 2297-VI «Про захист персональних даних».