



International Science Group

ISG-KONF.COM

IV

**INTERNATIONAL SCIENTIFIC
AND PRACTICAL CONFERENCE
«TECHNOLOGIES, THEORIES AND DEVELOPMENTS:
MODERN SCIENTIFIC TEACHING»**

Valencia, Spain

September 23-26, 2025

ISBN 979-8-89814-218-6

DOI 10.46299/ISG.2025.2.4

TECHNOLOGIES, THEORIES AND DEVELOPMENTS: MODERN SCIENTIFIC TEACHING

Proceedings of the IV International Scientific and Practical Conference

Valencia, Spain
September 23-26, 2025

UDC 01.1

The 4th International scientific and practical conference “Technologies, theories and developments: modern scientific teaching” (September 23-26, 2025) Valencia, Spain. International Science Group. 2025. 183 p.

ISBN – 979-8-89814-218-6

DOI – 10.46299/ISG.2025.2.4

EDITORIAL BOARD

<u>Pluzhnik Elena</u>	Professor of the Department of Criminal Law and Criminology Odessa State University of Internal Affairs Candidate of Law, Associate Professor
<u>Liudmyla Polyvana</u>	Department of accounting, Audit and Taxation, State Biotechnological University, Kharkiv, Ukraine
<u>Mushenyk Iryna</u>	Candidate of Economic Sciences, Associate Professor of Mathematical Disciplines, Informatics and Modeling. Podolsk State Agrarian Technical University
<u>Prudka Liudmyla</u>	Odessa State University of Internal Affairs, Associate Professor of Criminology and Psychology Department
<u>Marchenko Dmytro</u>	PhD, Associate Professor, Lecturer, Deputy Dean on Academic Affairs Faculty of Engineering and Energy
<u>Harchenko Roman</u>	Candidate of Technical Sciences, specialty 05.22.20 - operation and repair of vehicles.
<u>Belei Svitlana</u>	Ph.D., Associate Professor, Department of Economics and Security of Enterprise
<u>Lidiya Parashchuk</u>	PhD in specialty 05.17.11 "Technology of refractory non-metallic materials"
<u>Levon Mariia</u>	Candidate of Medical Sciences, Associate Professor, Scientific direction - morphology of the human digestive system
<u>Hubal Halyna Mykolaiivna</u>	Ph.D. in Physical and Mathematical Sciences, Associate Professor

10.	Водотика Я., Яковлева О.В. СТРУКТУРНИЙ ПІДХІД ДО РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ НА ОСНОВІ КЛАСТЕРИЗАЦІЇ КЛЮЧОВИХ ТОЧОК	35
11.	Голубов В.Р., Молчанов А.О. СИСТЕМА АВТОМАТИЗАЦІЇ ЖИТЛОВОГО ПРИМІЩЕННЯ ЗА КОНЦЕПЦІЄЮ "РОЗУМНИЙ БУДИНОК" З ІНТЕГРАЦІЄЮ МОДУЛІВ ІОТ ТА АІ-ПОМІЧНИКОМ	39
12.	Колева А.М., Апенько Н.В. УТИЛІТА ДЛЯ ЗАБЕЗПЕЧЕННЯ КОНФІДЕНЦІЙНОСТІ ФАЙЛІВ СИСТЕМИ WINDOWS	41
13.	Колісниченко М. РОЛЬ ПРОЦЕДУРНИХ ГЕНЕРАТИВНИХ МЕТОДІВ У СТВОРЕННІ ОПТИМІЗОВАНИХ 3D-МОДЕЛЕЙ ДЛЯ ІГОР	44
14.	Нечаєва Я. АНАЛІЗ ОСОБЛИВОСТЕЙ МЕТОДІВ РОЗПІЗНАВАННЯ ІМЕНОВАНИХ СУТНОСТЕЙ ДЛЯ ТЕКСТІВ СТРАХОВОГО СЕКТОРУ	48
15.	Фещенко О.О., Марченко Н.Б. ЗАСТОСУВАННЯ РІВНІВ ДЕТАЛІЗАЦІЇ ДЛЯ ПІДВИЩЕННЯ ПРОДУКТИВНОСТІ 3D-СЦЕН	51
16.	Харченко В.В., Любченко В.А. АНАЛІЗ СУЧАСНИХ ПІДХОДІВ ДЕТЕКЦІЇ ТА РОЗПІЗНАВАННЯ ТЕКСТУ НА ЗОБРАЖЕННЯХ	54
COMPUTER-INTEGRATED TECHNOLOGIES		
17.	Михайленко В.В., Марченко Н.Б. ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ УПРАВЛІННЯ ЯК ОСНОВА КОМП'ЮТЕРНО-ІНТЕГРОВАНИХ ТЕХНОЛОГІЙ	60
CONSTRUCTION AND CIVIL ENGINEERING		
18.	Burda Y., Vanian A. QUALITY CONTROL IN HEAT, GAS SUPPLY, AND VENTILATION SYSTEMS: ENSURING ENERGY EFFICIENCY, RELIABILITY, AND SUSTAINABILITY IN CONSTRUCTION	62

АНАЛІЗ ОСОБЛИВОСТЕЙ МЕТОДІВ РОЗПІЗНАВАННЯ ІМЕНОВАНИХ СУТНОСТЕЙ ДЛЯ ТЕКСТІВ СТРАХОВОГО СЕКТОРУ

Нечаєва Ярослава

студент групи ІНФм-24-2

Харківський національний університет радіоелектроніки

Науковий керівник:

Творошенко Ірина Сергіївна

к.т.н., доц., доцент кафедри інформатики

Харківський національний університет радіоелектроніки

У сучасному світі обробка текстових даних у страховому секторі набуває критичного значення через експоненційне зростання обсягів документів, таких як страхові поліси, заяви на відшкодування та звіти про ризики.

Розпізнавання іменованих сутностей (Named Entity Recognition, NER) дозволяє автоматично витягувати ключові елементи, зменшуючи помилки ручної обробки та підвищуючи операційну ефективність. Однак, методи NER стикаються з викликами, пов'язаними з домен-специфічною термінологією (наприклад, регуляторні терміни, типи ризиків), неоднозначністю контексту та потребою в високій точності для юридично чутливих текстів.

Сучасні підходи варіюються від правилкових систем до трансформерів та великих мовних моделей (LLM), їх емпіричний аналіз на фінансових та страхових корпусах допомагає обрати оптимальні рішення.

Розвиток NER для фінансових і страхових текстів зумовлений необхідністю швидкої обробки неструктурованих даних, де традиційні методи поступаються глибокому навчанню. Дослідження показують, що BERT-based моделі досягають F1-score понад 85% у фінансових задачах на наборі даних BraFiNER [1], тоді як LayoutLM на страховому корпусі PAYSIPS демонструє F1 ~ 64% з перевагою попереднього тренування на подібних документах [2].

Гібридні підходи, такі як BiLSTM-CRF, покращують результати на 5-10% у порівнянні з базовими моделями на фінансових наборах даних [3].

Ключові сутності у страховому домені включають номери полісів, імена клієнтів, дати, інформацію про претензії, організації та регуляторні терміни. Для оцінки використовуються анотовані корпуси, такі як PAYSIPS (страхові виписки, 611 сторінок) [2] та BraFiNER (фінансові транскрипти, 384 документи) [1], а також FiNER-ORD (фінансові звіти з XBRL-тегами) [4].

Правилкові (rule-based) методи демонструють високу точність для стандартних шаблонів (формати дат чи номерів полісів) з низькими обчислювальними вимогами, досягаючи точність більше 95% на фіксованих паттернах фінансових текстів. Однак, вони проявляють низьку гнучкість для

неоднозначних сутностей (F1 <70% на варіативних описах ризиків) та вимагають трудомісткого створення правил.

Conditional Random Fields (CRF) моделюють послідовності міток з урахуванням залежностей. Формула ймовірності [5]:

$$P(y|x) = \frac{1}{Z(x)} \exp(\sum_{i=1}^n \sum_{k=1}^K \lambda_k f_k(y_{i-1}, y_i, x, i)), \quad (1)$$

де y – послідовність міток;

x – вхідний текст;

$Z(x)$ – нормалізуючий фактор;

f_k – функції ознак;

λ_k – ваги.

CRF ефективні на невеликих наборах даних (F1 ~ 69% на фінансовому корпусі FiNER [4]) та інтегруються з BiLSTM для підвищення точності. Недоліки включають потребу в ручному інжинірингу ознак та чутливість до помилок у анотаціях.

BERT-based трансформери використовують bidirectional контекст для вбудовувань. Вони досягають високої точності (F1 98,99% для BERTimbau на BraFiNER [1]; ~ 64% для LayoutLM на PAYSIPS [2]) та адаптуються через fine-tuning на FinBERT для фінансових текстів [6]. Недоліки: високі обчислювальні ресурси та обмежена інтерпретованість для регуляторних вимог.

LLM-based методи (GPT, T5) дозволяють zero/few-shot навчання з F1 до 97% у кейсах [5], інтегруючи промпти для домен-специфічних завдань. Перевагами є гнучкість для нових сутностей (weighted F1 ~ 67% для GPT-4o на FiNER-ORD [4]). Недоліки: ризик галюцинацій (помилки в сумах виплат) та залежність від API з високою вартістю.

Для порівняння метрик схожості множин сутностей застосовується Jaccard similarity [7]:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad (2)$$

де A, B – множини витягнутих сутностей.

Гібридні підходи (CRF + BERT) оптимізують продуктивність (F1 +5-10%), рекомендуються для страхового сектору з fine-tuning на спеціалізованих корпусах. Очікується, що гібридні моделі забезпечать F1 >85% на страхових текстах з динамічним розширенням бази знань. Перспективи: інтеграція з LayoutLM для документів та RL для вкладених сутностей. Подібні системи матимуть застосування у автоматизації страхових процесів та InsurTech, перспективою є впровадження засобів комп'ютерного зору [8-11].

Список літератури:

1. Oliveira D. et al. Evaluating Named Entity Recognition: A Comparative Analysis of Mono- and Multilingual Models in the Financial Domain // arXiv preprint arXiv:2403.12212. 2024.
2. Uthayasooriyar B. et al. Training LayoutLM from Scratch for Efficient Named-Entity Recognition in the Insurance Domain // arXiv preprint arXiv:2412.09341. 2024.
3. Huang Z. et al. Bidirectional LSTM-CRF Models for Sequence Tagging // arXiv preprint arXiv:1508.01991. 2015.
4. Malmasi S. et al. FiNER-ORD: Financial Named Entity Recognition Open Research Dataset // arXiv preprint arXiv:2302.11157. 2023.
5. Sutton C., McCallum A. An Introduction to Conditional Random Fields // Foundations and Trends in Machine Learning. 2012. Vol. 4, No. 4. P. 267-373.
6. Wang S. et al. GPT-NER: Named Entity Recognition via Large Language Models // arXiv preprint arXiv:2304.10428. 2023.
6. Devlin J. et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // arXiv preprint arXiv:1810.04805. 2018.
7. Niwattanakul S. et al. The Application of Similarity Measure of Jaccard Coefficient in Medical Field // Proc. of IEEE Conf. on ICT and Knowledge Engineering. 2013. P. 1-4.
8. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., Hudáková M., and Gorokhovatskyi O. (2024) Application a committee of Kohonen neural networks to training of image classifier based on description of descriptors set, *IEEE Access*, vol. 12, pp. 73376-73385.
9. Gorokhovatskyi V., Chmutov Y., Tvoroshenko I., and Kobylin O. (2025) Reducing computational costs by compressing the structural description in image classification methods, *Advanced Information Systems*, vol. 9, no. 1, pp. 5-12.
10. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., and Hudáková M. (2025) Image description compression in classification structural methods, *IEEE Access*, vol. 13, pp. 43631-43641.
11. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O. (2024) Transforming image descriptions as a set of descriptors to construct classification features, *Indonesian Journal of Electrical Engineering and Computer Science*, 33 (1), 113-125.