



РЕЦЕНЗІЯ

Всеукраїнський конкурс
студентських робіт зі
штучного інтелекту 2023

Перший етап

на наукову роботу

назва Розробка додатку розпізнавання мовлення
з використанням штучного інтелекту
шифр Розпізнавання мовлення

представлену на Всеукраїнський конкурс студентських робіт зі штучного інтелекту

№ з/п	Характеристики та критерії оцінки рукопису наукової роботи	Рейтингова оцінка. Максимальна кількість балів (за 120-бальною шкалою)	Бали
1	Актуальність проблеми	10	9.0
2	Новизна та оригінальність ідей	15	6.0
3	Використані методи дослідження	15	11.3
4	Теоретичні наукові результати	10	4.3
5	Практична направленість результатів	20	14.7
6	Рівень використання наукової літератури та інших джерел інформації	5	2.3
7	Ступінь самостійності роботи	10	9.3
8	Якість оформлення	5	4.0
9	Наукові публікації	10	1.5
10	<p>Недоліки роботи (пояснення зниження максимальних балів у пунктах 1-9):</p> <ol style="list-style-type: none">Тема роботи актуальна.Використано алгоритми розпізнавання мови, реалізовані в бібліотеці SpeechRecognition python. В роботі немає / недостатньо власних оригінальних ідей. <p>Суть роботи: використано засоби бібліотеки PyAudio для зчитування даних з мікрофону; ці аудіо-дані надіслано в хмару google для розпізнавання з параметром «українська мова» (рядок коду <code>text=recognizer.recognize_google(audio, language="uk-UA").lower()</code> в програмі студента), розпізнавання мови робиться в хмарі Google алгоритмами, що реалізовані в бібліотеці SpeechRecognition.</p> <ol style="list-style-type: none">Використано сучасні методи, алгоритми, реалізовані іншими дослідниками в бібліотеці SpeechRecognition python. Бібліотека SpeechRecognition використовує Python 3.8+, PyAudio 0.2.11+ (використання мікрофону), PocketSphinx (Sphinx recognizer, працює offline), Google API Client Library for Python		

№ з/п	Характеристики та критерії оцінки рукопису наукової роботи	Рейтингова оцінка. Максимальна кількість балів (за 120-бальною шкалою)	Бали
	<p>(використовує Google Cloud Speech API), Vosk (Vosk API speech recognition, працює offline), Whisper API, OpenAI (в разі використання Whisper API speech recognition).</p> <p>В роботі аудіо-дані передано в хмару для розпізнавання алгоритмами, які реалізовано Google. Окремі відомі інструменти, такі як Sphinx та Vosk працюють offline.</p> <p>Бібліотеку gTTS (Google Text-to-Speech), згадану в теоретичному розділі роботи, не використано в розробленому додатку.</p> <p>4. Немає теоретичних наукових результатів.</p> <p>5. Продемонстровано вміння використовувати на практиці готові рішення з бібліотеки SpeechRecognition python для розпізнавання голосового повідомлення з мікрофону.</p> <p>6. Рівень використання наукової літератури та інших джерел інформації за темою розпізнавання мовлення – низький. Немає посилань на наукові статті в періодичних виданнях за останні 5 років за темою розпізнавання мовлення.</p> <p>9. Наукових публікацій немає.</p> <p>Рецензований текст не задовільняє критерії наукової публікації: поставлена задача не є новою, наводиться методика налаштування існуючого програмного забезпечення для розпізнавання голосових команд. Це, скоріше, методика виконання лабораторної роботи.</p> <p>Немає порівняння запропонованого рішення з існуючими аналогами, немає аналізу точності розпізнавання, не наведено алгоритму, відсутнє моделювання чи аналіз обчислювальної складності... Ані структура, ані стиль рукопису та його граматики не дозволяють мені рекомендувати дану роботу до публікації.</p> <p>Окремо варто звернути увагу на довільне розміщення посилань на публікації у тексті рукопису: жодне із перевірених посилань не відповідало контексту. Зокрема, виникли застереження стосовно наступних фрагментів тексту, його нумерації та ілюстрацій:</p> <p>1. Що означають фрази на кшталт:</p> <ul style="list-style-type: none"> - "Мало хто залишається невідомим"; - "...що розпізнавання мовлення стало надзвичайно актуальним явищем..."; - "Передавачем команд є конкретний модуль, 		

№ з/п	Характеристики та критерії оцінки рукопису наукової роботи	Рейтингова оцінка. Максимальна кількість балів (за 120-бальною шкалою)	Бали
	<p>який приймає голосові команди"?</p> <p>2. Де Рис.1.1.?</p> <p>3. Схема роботи процесу розпізнавання мовлення на Рис.1.2 надто абстрактна.</p> <p>4. "1.1.2. Методи розпізнавання мовлення." - нумерація розділів безладна: немає п. 1.1.1</p> <p>5. Стор.8.: "...знаходиться новий сигналу відліку..." - автору слід було вчитати власний текст та відредагувати його!</p> <p>6. Стор.8.: "...система "Одрі" [8]" - джерело не стосується згаданої системи безпосередньо: «8. Kabisatpathy P. Fault detection and diagnosis in analog integrated circuits using artificial neural network in a pseudorandom testing scheme / Kabisatpathy P., Barua A., Sinha S. // 3rd International Conference on Electrical & Computer Engineering ICECE 2004, 28-30 December 2004, Dhaka, Bangladesh, — pp. 52-55.»</p> <p>7. Стор.8.: "У 1997 році розпізнавач безперервної мови ragon Dictate починає підкорювати світ своєю інновацією [13]." - посилання не стосується теми, адже у джерелі йдеться про контроль зношування інструменту в процесі різання - 13. Krzysztof Jemielniak. Tool Wear Monitoring by Means of Artificial Neural Networks. [Електронний ресурс] — режим доступу: http://www.eng.utoledo.edu/pmmc/issue7-4.pdf (URL - адреса недійсна).</p> <p>8. Стор.9.: посилання на російські підручники не вітається: [2]. Николенко С., Кадурич А., Архангельская Е. Глубокое обучение. – СПб.: Питер, 2018. – 480 с.</p> <p>9. Стор.9.: посилання [7] - також явно помилкове ; Рис. 1.5. - логотип Siri недоречний, адже носить суто ілюстративний характер.</p> <p>10. Стор.13.: після Рис.1.5 йде одразу Рис. 1.15. Підпис "Схема реалізації машинного та глибокого навчання" не відповідає наведеному зображенню. Вона є суто символічною. Загальний аналіз вказує на системний характер недоліків та недоцільність наведення повного переліку зауважень.</p> <p>1. Тема актуальна, але дуже загальна 2. В роботі використані типові і готові рішення, оригінальні ідеї відсутні 3. Наукового пошуку немає, аналітичного</p>		

№ з/п	Характеристики та критерії оцінки рукопису наукової роботи	Рейтингова оцінка. Максимальна кількість балів (за 120-бальною шкалою)	Бали
	<p>порівняння готових засобів теж немає, ні моделі, ні результатів.</p> <p>4. Наукові результати не сформульовані і реально відсутні.</p> <p>5. Робота практично спрямована, але без врахування специфіки жодної конкретної сфери</p> <p>6. Джерела загальні, посилання не відповідають джерелам, жодного по власне розпізнаванню мови.</p> <p>8. Посилання не відповідають списку літератури</p> <p>9. Публікації відсутні</p> <p>2 – Ідеї не є оригінальними та новими. 6 – є застарілі джерела, є рос. мовою. 9 – Наукові публікації (автора) не наведено, тому вони не можуть бути оцінені.</p>		
10.1	Наявність розробленого програмного забезпечення, у т.ч. веб-сервісу, мобільного застосунку	10	6.3
10.2	Рівень отриманих практичних результатів наукової роботи	10	5.0
Сума балів			73.8

Загальний висновок

не рекомендується для захисту на науково-практичній конференції

(рекомендується, не рекомендується для захисту на науково-практичній конференції)

Рецензенти члени Конкурсної комісії напряму «Розпізнавання мови»

(підпис)

(Ім'я ПРІЗВИЩЕ)

(місце роботи, посада, науковий ступінь)

«08» листопада 2023 року