

Рішення
разової спеціалізованої вченої ради
про присудження ступеня доктора філософії

Здобувачка ступеня доктора філософії Бочарова Майя Юріївна 1997 року народження, громадянка України освіта вища: закінчила у 2019 році Одеський національний економічний університет за спеціальністю «Міжнародні економічні відносини».

3 жовтня 2021 році є аспіранткою денної форми навчання за спеціальністю 122 «Комп'ютерні науки» кафедри математичного забезпечення комп'ютерних систем Факультету математики, фізики та інформаційних технологій Одеського національного університету імені І.І. Мечникова. Виконала акредитовану освітньо-наукову програму «Комп'ютерні науки».

Разова спеціалізована вчена рада, утворена наказом Одеського національного університету І.І. Мечникова від 23 квітня 2025 року № 74.122.2025

У складі:

Голови разової спеціалізованої вченої ради: Мещеряков Володимир Іванович – доктор технічних наук, професор, професор кафедри Інформаційних технологій Одеського національного університету імені І.І. Мечникова

Рецензента: Шпінарева Ірина Михайлівна – кандидат фізико-математичних наук, доцент, доцент кафедри математичного забезпечення комп'ютерних систем Одеського національного університету імені І.І. Мечникова

Офіційних опонентів:

— Бісікало Олег Володимирович – доктор технічних наук, професор, завідувач кафедри Автоматизації та інтелектуальних інформаційних технологій Вінницького національного технічного університету.

— Арсірій Олена Олександрівна – доктор технічних наук, професор, завідувачка кафедри інформаційних систем Національного університету «Одеська політехніка»

— Висоцька Вікторія Анатоліївна – доктор технічних наук, доцент, доцент кафедри інформаційних систем та мереж Інституту комп'ютерних наук та інформаційних технологій Національного університету «Львівська політехніка»

на засіданні «26» червня 2025 року прийняла рішення про присудження ступеня доктора філософії з галузі знань 12 «Інформаційні технології» Бочарової Майї Юріївни на підставі публічного захисту дисертації «Методи та інформаційна технологія предметно-орієнтованого аналізу природномовних текстів» за спеціальністю 122 «Комп'ютерні науки».

Дисертацію виконано в Одеському національному університеті імені І.І. Мечникова.

Науковий керівник: Малахов Євгеній Валерійович, доктор технічних наук, професор, завідувач кафедри математичного забезпечення комп'ютерних систем Одеського національного університету імені І.І. Мечникова.

Дисертацію подано у вигляді спеціально підготовленого рукопису українською мовою. У дисертації містяться нові науково обґрунтовані результати дослідження. Наукова новизна отриманих результатів полягає в у розробці нових та вдосконаленні наявних методів обробки природномовних текстів на прикладі резюме та вакансій, що забезпечують ефективніше моделювання текстів у сфері управління персоналом. Зокрема, вперше запропоновано підходи до контекстного подання фраз, структурування документів для некерованого навчання та векторизації довгих документів із використанням позиційного кодування й механізму самоуваги, а також удосконалено методи подання токенів у візуально насичених документах та скорочення тексту з урахуванням його структури. Наукові положення, висновки і рекомендації, сформульовані у дисертації, є достовірними, аргументованими та належним чином обґрунтованими. Наукові результати, представлені в роботі, є особистим внеском здобувача. Актуальність теми зумовлена необхідністю підвищення ефективності автоматизованого аналізу резюме та вакансій у сфері управління персоналом за допомогою штучного інтелекту, що особливо важливо в умовах динамічних змін на ринку праці. Запропоновані підходи спрямовані на покращення якості семантичного пошуку та відповідності між досвідом кандидатів і вимогами роботодавців, що сприяє розвитку конкурентоспроможних систем у сфері управління персоналом. Дисертаційна робота відповідає вимогам Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії, затвердженого постановою Кабінету Міністрів України від 12 січня 2022 року № 44 із змінами, внесеними Постановою КМУ № 341 від 21.03.2022 року та №502 від 19.05.2023 року, та оформлена згідно з вимогами наказу МОН України № 40 від 12 січня 2017 року «Про затвердження Вимог до оформлення дисертації» (зі змінами).

Здобувачка має 8 наукових публікацій за темою дисертації, з них 4 – статті у наукових виданнях, включених до переліку наукових фахових видань України, зокрема 1 стаття в українському періодичному науковому виданні, проіндексованому у базі даних Scopus; 4 – публікації, що засвідчують апробацію матеріалів дисертації:

1. Bocharova M. Y. VacancySBERT: the approach for representation of titles and skills for semantic similarity search in the recruitment domain / M. Y.

Bocharova, E. V. Malakhov, V. I. Mezhuyev // Applied Aspects of Information Technology. – 2023. – Vol. 6, no. 1. – P. 52–59

2. Bocharova M. General-purpose text embeddings learning for Ukrainian language / M. Bocharova, E. Malakhov // Advanced Information Technology. – 2024. – No 1(3). – P. 6–12.

3. Bocharova M. Improving information theory of context-aware phrase embeddings in HR domain / M. Bocharova, E. Malakhov // Eastern-European Journal of Enterprise Technologies. – 2024. – Vol. 5, iss. 2(131). – P. 53–60.

4. Bocharova M. Y. ResJobFit – end-to-end artificial neural networks based technology for job-resume matching / M. Y. Bocharova, E. V. Malakhov // Applied Aspects of Information Technology. – 2024. – Vol. 7, no. 4. – P. 378–391.

У дискусії взяли участь (голова, рецензенти, офіційні опоненти, інші присутні) та висловили зауваження:

Шпінарева Ірина Михайлівна – кандидат фізико-математичних наук, доцент, доцент кафедри математичного забезпечення комп'ютерних систем Одеського національного університету імені І.І. Мечникова. Висловлені рецензентом зауваження:

1. У роботі досліджуються моделі базових розмірів (як от BERT-base), а також вказано, що проводиться дистиляція моделей. Було б доцільно дослідити також більші розміри моделей (BERT-large) і навести порівняння якості та швидкості в залежності від розміру моделей.

2. На рисунку 4.7 “Метрики втрат на валідаційному наборі даних при тренуванні” зазначено не відсотки, а частину цілого.

3. У розділі 4 зазначено, що дані походять з різних форматів документів (зокрема, pdf, docx, rft, html), але не описано процес їх обробки та конвертації у формат, який використовується в подальших етапах і містить текст та стиліві ознаки.

4. Для деяких з розглянутих методів не приведено повний опис експериментів, зокрема, експерименти з виділення ключової інформації з кожної з секцій резюме та вакансій.

5. У роботі розглянуто питання прискорення процесу сприйняття рекрутером резюме, але такі аспекти, як врахування вподобань споживачів (наприклад, форматування та структура анотації) не розкрито повною мірою. Було б доцільно розглянути вплив різних факторів на результат сприйняття резюме рекрутером.

Бісікало Олег Володимирович – доктор технічних наук, професор, завідувач кафедри Автоматизації та інтелектуальних інформаційних технологій Вінницького національного технічного університету. Висловлені опонентом зауваження:

1. У роботі не наведено ефективний час застосування розробленої технології на етапі векторного пошуку, хоча вказано, що складність алгоритму лінійна.

2. Недостатньо розкрито етап зберігання та реалізації пошуку подібності у базі даних на основі векторів.

3. У роботі майже не висвітлена оцінка експертами якості запропонованих анотацій при використанні технології з етапом сумаризації.

4. Виходячи з тексту дисертаційної роботи не зовсім зрозуміло, які саме програмні засоби та технології використано для реалізації прототипу, зокрема в тексті описано вибір параметрів, але не ясно, які бібліотеки та мова програмування застосовано для програмної імплементації методів.

5. На жаль, текст роботи не вільний від орфографічних, синтаксичних та стилістичних похибок, зокрема введений авторкою термін «контекстно-освічений (метод)» був би більш зрозумілим читачу як «контекстно-залежний» за сутністю запропонованого методу.

Арсирій Олена Олександрівна – доктор технічних наук, професор, завідувачка кафедри інформаційних систем Національного університету «Одеська політехніка». Висловлені опонентом зауваження:

1. В роботі на сторінці 23 представлений занадто великий перелік поставлених задач, що створює враження їх фрагментарності та розпорошеності. Крім того це ускладнює аналіз виконання цих задач для опонента. Наприклад, вираз «інформаційну теорію багатомовної дистиляції текстових подань» із задачі 9 взагалі не знайдено по тексту дисертації.

2. Розділ 2 дисертаційної роботи присвячений створенню 3-х методів обробки природномовних текстів (згідно із науковою новизною). Доцільно було б описати методи у вигляді послідовності кроків виконання. Крім того при викладені методів недостатньо повно описані обмеження їх використання щодо обчислювальної складності та використання пам'яті при збереженні векторних представлень документів.

3. Схема інформаційної технології предметно-орієнтованого аналізу природномовних текстів AI ResJobFit, яку представлена на сторінці 91 потребує пояснень щодо наявності двох однакових паралельних послідовностей кроків та відсутності блока прийняття рішень.

4. Не заважаючи на той факт, що набір пов'язаних між собою метрик для оцінки якості сумаризації ROUGE (сторінка 103) є найбільш популярним доцільно було б навести також і інші автоматизовані метрики оцінки, а також вказати на можливість експертних оцінок (Human Evaluation) та спеціалізованих завдань (QA-based Evaluation).

5. Виходячи з тексту дисертаційної роботи не зовсім зрозуміло, які саме програмні засоби та технології використано для реалізації прототипу. В тексті описано вибір параметрів, але не ясно, які саме бібліотеки та мова програмування використовуються для реалізації методів та підходів.

Висоцька Вікторія Анатоліївна – доктор технічних наук, доцент, Доцент кафедри інформаційних систем та мереж Інституту комп'ютерних

наук та інформаційних технологій Національного університету «Львівська політехніка». Висловлені опонентом зауваження:

1. Некоректна назва дисертації, так як інформаційна технологія це сукупність методів і засобів.

2. досліджені моделі авторкою на основі архітектури типу «Трансформер» з класичним механізмом самоуваги (Multihead self attention), проте було б доцільно проаналізувати можливість використання альтернативних методів самоуваги, зокрема, Flash Attention, Paged Attention.

3. У роботі розглянуто питання прискорення процесу сприйняття рекрутером резюме, але такі аспекти, як врахування вподобань споживачів (наприклад, форматування та структура анотації) не розкрито повною мірою. Було б доцільно розглянути вплив різних факторів на результат сприйняття резюме рекрутером.

4. Розділ 4 роботи є перевантаженим за обсягом для кращого сприйняття було б доцільно розділити його на два, наприклад: основні етапи технології та опціональні.

5. Згідно назви, мети та деяких задач дослідження полягає ще в розробленні інформаційної технології, яка, в свою чергу, передбачає розроблення засобів (програм, системи, модулів тощо), але в роботі мало приділено увагу цим розробленим засобам, опису їх технічних характеристик, інтерфейсів, архітектури тощо.

6. У роботі виявлено ряд технічних помилок граматичного та синтаксичного характеру.

Результати відкритого голосування:

«За» – п'ять членів ради,

«Проти» – немає.

На підставі результатів відкритого голосування разова спеціалізована вчена рада присуджує Бочаровій Майї Юріївні ступінь доктора філософії з галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки».

Відеозапис трансляції захисту дисертації додається.

Голова разової спеціалізованої
вченої ради

Мещеряков
Володимир Іванович

