

ВИКОРИСТАННЯ МЕТОДІВ ГЛИБОКОГО НАВЧАННЯ ДЛЯ ОЦІНКИ ЯКОСТІ МЕДИЧНИХ ЗАПИСОК ПАЦІЄНТІВ

¹Вінницький національний технічний університет

Анотація

У статті розглянуто роль машинного та глибокого навчання в сфері освіти. Показано приклад використання сучасних моделей глибокого навчання для задачі обробки природної мови, а саме оцінки екзаменаційних медичних записок пацієнтів.

Ключові слова: машинне навчання; глибоке навчання; обробка природної мови; медичні записки пацієнтів.

Abstract

The article examines the role of machine and deep learning in the education field. The work shows an example of using modern models of deep learning for the natural language processing task, namely the evaluation of examination of clinical patient notes.

Keywords: machine learning; deep learning; natural language processing; clinical patient notes.

Машинне і глибоке навчання у нинішній час використовуються у дедалі більшій кількості сфер нашого життя. Не виключенням є і сфера освіти [1]. При наявності достатньо великої кількості гарних даних, що мають підходящу розмітку та ця розмітка є узгодженою, багато задач, які вирішуються зазвичай експертом, можна автоматизувати.

Автоматизація процесів оцінки деякого виду робіт, може дозволити виключити людину із процесу, зробивши цей процес більш справедливим. Наприклад, оцінка твору учня на тестуванні, може бути досить суб'єктивною за рахунок відносин між студентом та викладачем, або просто поганим/гарним настроєм викладача.

Одна із задач, що може бути автоматизована – задача оцінки якості написання медичних записок пацієнтів, при навчанні експертів-лікарів. Лікар при первинній розмові з пацієнтом повинен чітко записати всі основні елементи розповіді пацієнту: вік, стать, сімейний стан, симптоми пацієнта, перелік хронічних захворювань, основні хвороби батьків, чи присутній активний образ життя, та багато інших важливих деталей.

Ще одна важлива проблема – проблема часу та вартість оцінки великої кількості робіт. Щоб оцінити якісно записку лікаря, що навчається, потрібен як мінімум один експерт який може її проаналізувати. Тому автоматизація цього процесу є дуже важливою задачею, для навчання або перевірки лікарів.

Базові підходи обробки природної мови, можуть бути застосовані для великої кількості шаблонних фраз при опису характеристики, наприклад віку. Якщо відомо, що пацієнту 17 років, то потрібно шукати в тексті «17 years old», «17 yo», «17 y.o.», «17», та інші варіації. Але іноді, якась характеристика може бути в декількох реченнях відразу, або дуже сильно перефразована, і щоб її зрозуміти потрібно дивитись на текст цілком а не окремі шматочки. Наприклад, ознака «втрата інтересу до діяльності» може бути виражена як «більше не грає в баскетбол».

Для рішення цієї задачі вибрано набір даних з платформи Kaggle [2], що містить 10 історій пацієнтів та для кожної історії написано 100 записок від різних лікарів, що писали записку під час розповіді пацієнта [3]. Приклад записки пацієнта та ключових фраз зображено на рисунку 1. Задача – знайти всі ключові фрази що є в тексті, та є обов'язковими на думку експерту-лікаря. Після знаходження, можна порівняти кількість описаних характеристик, з кількістю яка повинна була бути із історії пацієнта.

HPI: 17yo 11 M 12 presents with palpitations 9 . Patient reports 3-4 months of 10 intermittent episodes 3 of " heart beating/pounding 9 out of my chest." 2 days ago during a soccer game had an episode 3 , but this time had chest pressure 2 and felt as if he were going to pass out 4 (did not lose consciousness). Of note patient endorses abusing adderall 6 , primarily to study (1-3 times per week). Before recent soccer game, took adderall 6 night before and morning of game. Denies shortness of breath, diaphoresis, fevers, chills, headache, fatigue, changes in sleep, changes in vision/hearing, abdominal pain, changes in bowel or urinary habits.

PMHx: none

Rx: uses friends adderall 6

FHx: mom with "thyroid disease 1 , " dad with recent heart attack 0

All: none

Immunizations: up to date

SHx: Freshmen in college. Endorses 3-4 drinks 3 nights / week (on weekends), denies tobacco, endorses trying marijuana. Sexually active with girlfriend x 1 year, uses condoms

Рисунок 1 – приклад записки пацієнту

Задача вирішувалась з використанням мови програмування Python [4], моделі глибоко навчання Bidirectional Encoder Representations from Transformers (BERT) [5] та фреймворку transformers [6]. Вибрана модель може вирішувати проблему відповіді на запитання. В випадку використаних в роботі даних запитанням буде ключова фраза, що повинна бути в записці, наприклад «вік пацієнта», або «хронічна хвороба батька», відповіддю – фрагмент тексту, де вказана необхідна інформація «17yo» або «dad with recent heart attack».

Так як для кожного тексту є декілька запитань, то потрібно забезпечити тренування таким чином, щоб один і той самий текст був тільки у одному наборі даних (тренувальному, валідаційному, або тестовому), щоб отримати правдиву точність моделі на даних, які вона не бачила взагалі. В якості метрики вибору кращої моделі використовується F1 міра [7]. Ця метрика застосовується у випадку, коли дані дуже не збалансовані, як в випадку задачі, що вирішується, відповідь на запитання є дуже маленькою, по кількості символів, в порівнянні зі всім текстом. Тренувальні і валідаційні метрики, під час навчання, показані на рисунку 2, модель навчалась на протязі 5 епох. Тестове значення F1 вийшло 0.82, що нижче ніж валідаційний показник, що може свідчити про різницю даних що використовувались при тренуванні і при тесті.

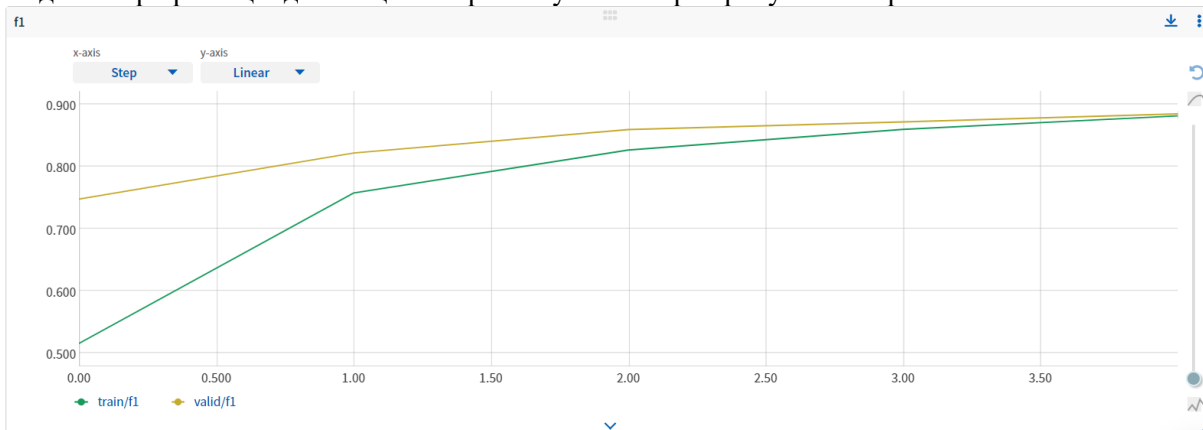


Рисунок 2 – метрики навчання моделі

Побудована модель має досить велике значення F1 метрики. Подальший план покращення моделі може бути пов'язаний з більш детальним аналізом даних та покращенням їх узгодженості. Також є необхідність проведення більш детального аналізу помилок, для розуміння які проблеми виникли у моделі, щоб вирішити їх у наступній ітерації.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Webb, M.E., Fluck, A., Magenheimer, J. et al. Machine learning for human learners: opportunities, issues, tensions and threats. *Education Tech Research Dev* 69, 2109–2130 (2021). <https://doi.org/10.1007/s11423-020-09858-2> .
2. Kaggle. [Online]. Available: <https://www.kaggle.com/> . Accessed on: May 16, 2022.
3. NBME - Score Clinical Patient Notes. [Online]. Available: <https://www.kaggle.com/competitions/nbme-score-clinical-patient-notes> . Accessed on: May 16, 2022.
4. Python. [Online]. Available: <https://www.python.org/> . Accessed on: May 16, 2022.

5. J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, “ BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in arXiv e-prints, 2018. [Online]. Available: <https://arxiv.org/pdf/1810.04805.pdf> . Accessed on: May 16, 2022.

6. Transformers. [Online]. Available: <https://huggingface.co/transformers/v3.0.2/index.html> . Accessed on: May 16, 2022.

7. Precision, recall and F-measures. [Online]. Available: https://scikit-learn.org/stable/modules/model_evaluation.html#precision-recall-f-measure-metrics . Accessed on: May 16, 2022.

Ісаснков Ярослав Александрович, аспірант, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: oiuygl@gmail.com

Isaienkov Yaroslav, graduate student, Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: oiuygl@gmail.com