

УДК 004.8

ПЕРСОНАЛІЗОВАНЕ НАДСИЛАННЯ СПОВІЩЕНЬ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

DOI 10 36994 / 2788- 5518- 2022-02-04-13

Матвієнко Н.Т. Національний технічний університет України «Київський політехнічний Інститут імені Ігоря Сікорського», Київ, Україна. matvienkont@yahoo.com

Могильний С.Б., к.т.н. Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна. isearch@ukr.net

Рожаловець К.С. Відкритий міжнародний університет розвитку людини «Україна», Київ, Україна. konstantin.rozahalavets@gmail.com

Анотація: Метою цієї статті є дослідити та спрогнозувати етапи та шляхи розвитку інфраструктури для баз даних фінансових установ, що пов'язані із індустріальними регуляціями. Розглянуто основні тенденції в ІТ в цілому та у вибраному домені, наведені порівняльні розрахунки для безсерверних хмарних сервісів та самоінстальованих баз даних.

Ключові слова: бази даних, AWS, публічна хмара, безсерверні технології, RDS, EC2, фінансовий сектор.

SENDING PERSONALIZED USER NOTIFICATIONS POWERED BY MACHINE LEARNING

Nazar Matvienko. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. matvienkont@yahoo.com

Sergij Mohylnyy, Ph.D. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. isearch@ukr.net

Konstantyn Rozhalovets. Open International University of Human Development "Ukraine", Kyiv, Ukraine. rozahalavets@gmail.com

Abstract: The article deals with the system of sending notifications to users. It is proposed to create a notification scheduling algorithm using deep learning, which includes: a data processing algorithm for training and the construction of a long-short-term memory (LSTM) network model that performs the scheduling most efficiently. The algorithm allows to improve the effectiveness of notifications through personalization, taking into account the gender, age and previous history of the target object.

One of the most important indicators of the notification system in the information and telecommunication system (ITS) is the timeliness of sending notifications to the target object, which is a dynamic variable in the system. This indicator of timeliness cannot be calculated absolutely precisely, since it has a large number of dependencies that cannot be tracked absolutely accurately.

The proposed model offers a technique to create personalized notifications to increase timeliness by considering gender, age, and previous activity history. The result of more accurate planning of notifications is an increase in the effectiveness of feedback, which increases the amount of time the user spends on a resource that provides services in one or another area, which is the main reason for the existence of a notification system. The improved scheduling system can be applied in many areas with high traffic: web services for educators and students, news/magazine portals, video hosting, online marketing, entertainment services and other areas.

For planning in the notification system, it is proposed to use LSTM recurrent neural network in combination with elements of feed forward neural network. An LSTM network is able to store its input data in internal memory, which plays a key and integral role in solving problems involving sequential data.

Key words: information and telecommunication system, notification system, machine learning, personalized notifications.

Вступ

Одним з найголовніших показників системи сповіщень в інформаційно-телекомунікаційній системі (ІКТС) є своєчасність відправки сповіщень цільовому об'єкту, що є динамічною змінною в системі. Даний показник своєчасності неможливо порахувати абсолютно точно, оскільки він має в собі велику кількість залежностей, які не можна відслідкувати достеменно.

Запропонована модель пропонує методику для створення персоніфікованих сповіщень, щоб збільшити показник своєчасності шляхом врахування статі, віку та попередньої історії активності. Оскільки дані параметри властиві всім користувачам ІКТС, вони дають змогу знайти схожі патерни активності, що дає цінну інформацію для планування надсилання сповіщень вчасно.

Результатом більш точнішого планування сповіщень є збільшення ефективності відгуків, що збільшує кількість часу користувач проводить на ресурсі, що надає послуги в тій чи іншій сфері, що є найголовнішою причиною існування системи сповіщення. Покращену систему планування можна застосовувати в багатьох сферах з великим трафіком: веб-сервіси для освітян, портали новин/журналів, відеохостинги, онлайн-маркетинг, сервіси розваг та інші сфери.

Виконання досліджень

У статті буде використано анонімна база даних з інтернет-ресурсу, який надає послуги з навчання освітян на різні тематики. Особливістю алгоритму те, що таку базу даних можна зібрати для кожної ІКТС окремо, оскільки необхідні дані зазвичай присутні в системі логування та інформації користувача, при цьому дотримуючись приватності цих користувачів через узагальненість таких даних. В результаті обробки та очистки, вигляд даних на вхід навчальної моделі буде мати наступний вигляд:

$$\{ s, a, t1, t2, t3, t4, t5, t6, t7, t8, t9 \}, \quad (1)$$

де s — стать користувача,

a — вік користувача,

$t1, \dots, t9$ — часові мітки останньої активності користувача в послідовному порядку.

Експериментальним шляхом було прийнято рішення вибрати 9 часових міток останньої активності, а на виході надавати 3 обраховані часові мітки наступного вигляду:

$$\{ p1, p2, p3 \}, \quad (2)$$

де $p1, p2, p3$ — обраховані часові мітки моделлю.

Змінні $t1, \dots, t9$ обробляються таким чином, щоб кожне з цих значень представляло рівно 1 годину на проміжку 1 рік (представлений датасет в статті дозволяє взяти такий проміжок) за формулою 2 для кожного значення:

$$(\text{TIMESTAMP} - \text{START_OF_YEAR}) // \text{ONE_HOUR},$$

де TIMESTAMP — одинарна мітка часу (Unix стандарт) активності користувача,

START_OF_YEAR - початок року в рік одинарної мітки TIMESTAMP ,

ONE_HOUR — 3600 секунд в одній годині,

// — ділення з остачею.

Наприклад, $\text{TIMESTAMP} = 1480453179$ (Tuesday, November 29, 2016 20:59:39 UTC), $\text{START_OF_YEAR} = 1451606400$ (Friday, January 1, 2016 0:00:00 UTC). Звідси, перетворене число для навчання буде рівне 3579. Дане форматування має такі переваги:

- Модель навчання працює з невеликими числами, що зменшує навантаження на навчання, прискорюючи роботу;
- Дає змогу робити планування маючи активності користувачів наступних років, а не тільки обмежується роком коли навчали модель.

Недоліком є ускладнене планування на перехідному періоді між роками, проте її можна уникнути якщо взяти базу даних з даними більше 1 року.

Для планування в системі сповіщень пропонується використовувати LSTM рекурентну нейронну мережу в комбінації з елементом нейронної мережі прямого поширення. Мережа LSTM має змогу пам'ятати свої вхідні дані у внутрішній пам'яті, що відіграє ключову та невід'ємну роль у вирішенні проблем, що пов'язані з послідовними даними. Тому рекурентні мережі віддається перевага як алгоритму для послідовних даних: часових рядів, розпізнавання мови, тексту фінансові дані, аудіо- та відеоряди, погода та інші типи даних. LSTM мережа виділяється серед конкурентів точнішим вивченням важливих залежностей, які мають більшу відстань один від одного. Обрана мережа здатна охоплювати складніші задачі, використовуючи її ключові елементи для контролю попередніх даних для наступних значень, що називаються затвори.

Ключовим компонентом LSTM є стан елемента, що є горизонтальною шиною, що проходить через весь нейрон. Тому ця мережа здатна віднімати або додавати інформацію зі стану елемента шляхом відрегульованих структур – затворами:

- Затвор забуття – вирішує яку інформацію потрібно видалити з стану елемента;
- Затвор оновлення – сукупність двох функцій, використана для оновлення загального стану елемента;
- Вихідний затвор – сукупність функцій, що вирішують які дані будуть передані на вихід.

Експериментальним шляхом було знайдено оптимальну структуру моделі, яка забезпечує високу швидкість навчання та найкращу точність знаходження наступних часових міток на оціночних даних.

Model: "sequential_8"

Layer (type)	Output Shape	Param #
bidirectional_8 (Bidirectional)	(None, 40)	3840
dense_16 (Dense)	(None, 50)	2050
dense_17 (Dense)	(None, 3)	153

Total params: 6,043
Trainable params: 6,043
Non-trainable params: 0

Рис. 1. Структура побудованої нейронної мережі

З структури бачимо, що двонаправленому режимі вхідні дані проходять у двох напрямках, що робить BiLSTM відмінним від звичайної LSTM. При глибокому навчанні модель повинна підібрати 6043 параметри, щоб робити ефективні обрахунки для різних потоків даних, які не пов'язані один з одним.

Було здійснено процес навчання, який зайняв 64 хвилини, навчившись на 825 000 різних мультиміжних послідовностях. В результаті навчання на тренувальних даних було зроблено оцінку точності планування провівши 21000 прогнозувань, вхідними даними яких при цьому було використано тестувальні дані приведені в аналогічний формат, результати яких зображено на рис. 2.

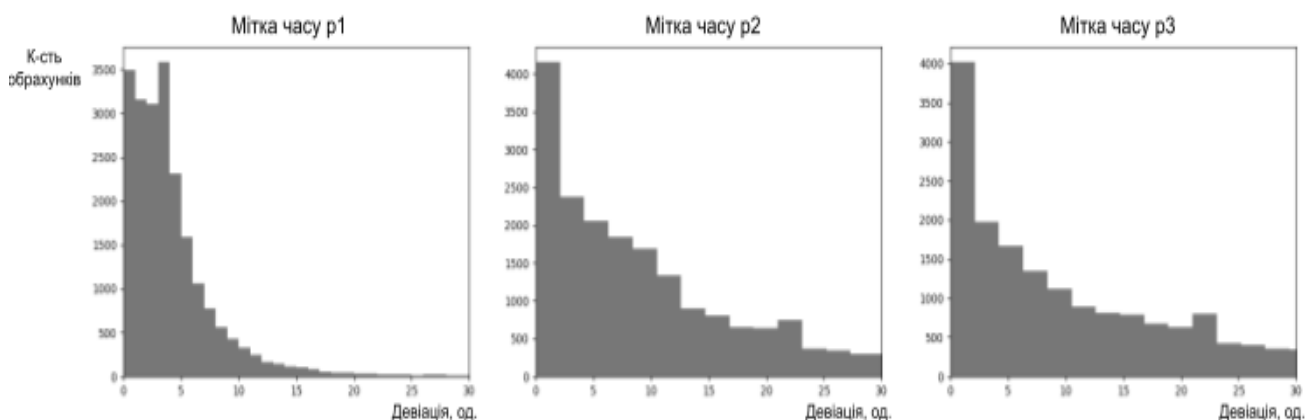


Рис. 2. Девіація мультикрокового прогнозування для 500 випадкових користувачів

З графіків видно, що більша частина обрахунків припадає на незначну девіацію в часі. В першому кроці обрахунку припадає 80% прогнозованих результатів з 23600 на девіацію 0-5 годин відносно зарані відомої правильної відповіді.

В табл.1 наведені статистичні значення прогнозувань для першого кроку.

Таблиця 1.

Статистичні значення прогнозувань для першого кроку

Девіація, год	0	1	2	3	4	5	6	7	8	9	10
К-ість прогнозувань	4018	7533	11027	15090	17541	19096	20155	20919	21459	21854	22157
Всього прогнозувань	23600	23600	23600	23600	23600	23600	23600	23600	23600	23600	23600
Частка з включно попередніми девіаціями, %	17,03	31,92	46,72	63,94	74,33	80,92	85,40	88,64	90,93	92,60	93,89

Значення з табл. 1 вказують на те, що побудована система сповіщень у 80% випадках правильно прогнозує наступний вхід користувача в систему.

Таблиця 2.

Статистичні значення прогнозувань для другого кроку

Девіація, год	0	1	2	3	4	5	6	7	8	9	10
К-ість прогнозувань	1915	3592	4999	6409	7791	9020	10123	11103	12066	12943	13858
Всього прогнозувань	23580	23580	23580	23580	23580	23580	23580	23580	23580	23580	23580
Частка з включно попередніми девіаціями, %	8.12	15.23	21.20	27.18	33.04	38.25	42.93	47.09	51.17	54.89	58.77

В табл. 2 зображена девіація для другого кроку прогнозувань, звідки видно, що відсоток спадає до 38% точності.

Табл. 3 показує, що девіація зростає, а її відсоткове відношення при девіації 0-5 годин сягає 30%. З цього можна зробити висновок, що найбільша точність отримується при першому кроці, а точність другого та третього – зменшується.

Таблиця 3

Статистичні значення прогнозувань для третього кроку

Девіація, год	1	2	3	4	5	6	7	8	9	10	11
К-ість прогнозувань	1828	3472	4860	6084	7150	8150	9019	9742	10399	11001	11606
Всього прогнозувань	23318	23318	23318	23318	23318	23318	23318	23318	23318	23318	23318
Частка з включно попередніми девіаціями, %	7.84	14.89	20.84	26.09	30.66	34.95	38.68	41.78	44.60	47.18	49.77

Для оцінки впливу даних на роботу моделі було погруповано дані по даті народження – 1970-1975, 1985-1990, 1990-1995. При цьому графік розподілу кількості навчальних даних по віку зображено на рис. 3, що вказує на те що основна маса даних припадає на користувачів 1990-1995 року народження

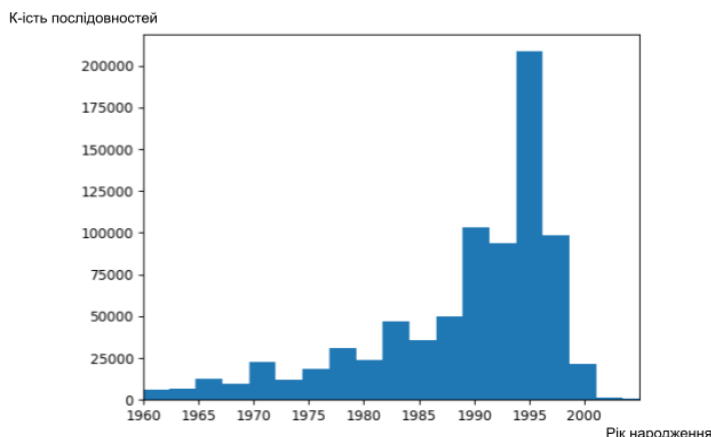


Рис. 3. Розподіл даних за роком народження

На рис. 4 зображено прогнозування для випадкових користувачів 1970-1975 року з даних для тестування.

Згідно рис. 4, для першого кроку прогнозування значення сума охоплених результатів з девіацією 0-5 рівна 77%, для другого – 38%, для третього – 30%.

Графік рис. 5 показує, що для першого кроку прогнозування значення сума охоплених результатів з девіацією 0-5 рівна 84%, для другого – 41%, для третього – 38%.

На рис.6 наведені дані прогнозування для користувачів 1990-1995 року.

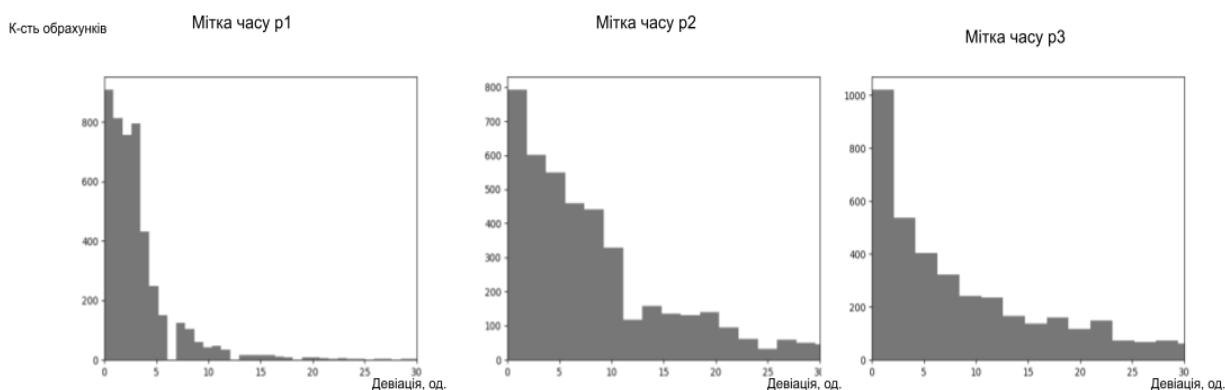


Рис. 4. Прогнозування для користувачів 1985-1990 року

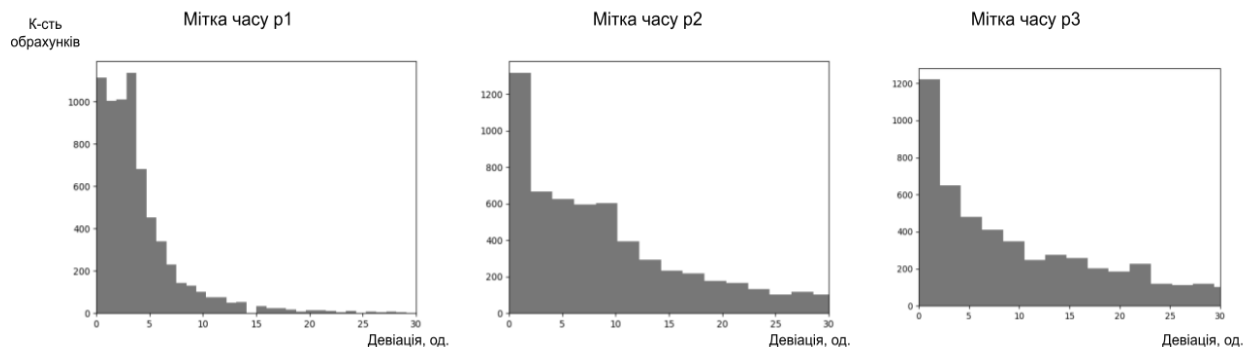


Рис.5. Прогнозування для користувачів 1990-1995 року

З рис. 6, для першого кроку прогнозування значення сума охоплених результатів з девіацією 0-5 рівна 79%, для другого – 33%, для третього – 31%. Що показує, що для даних 1990-1995 року результати прогнозування приблизно такі ж самі як і для даних 1970-1975. Але варто зауважити,

згідно графіку розподілу даних за роком народження, що дані 1990-1995 року мають в 20 разів більший об'єм, ніж дані 1970-1975 року. Схожі результати вказують на те, що модель з використанням LSTM нейронної мережі здатна адаптуватися до складних даних з великою кількістю варіативності, даючи гнучкість системі для мультизмінних послідовностей, що відрізняються один від одного.

Висновки

1. Завдяки глибокому навчанню було створено методику планування системи сповіщень з метою підвищення своєчасності їх надсилання.
2. Робота пропонує метод врахування персоналізації на базі існуючих знань про стать, вік та історію активності користувачів.
3. Вихідний алгоритм крім гнучкості в персоніфікації пропонує також мультикрокове планування сповіщень.
4. Експериментальним шляхом було прийдено до висновку, що найбільша точність отримується при першому кроці, а точність другого та третього – зменшується.
5. Експериментально продемонстровано використання LSTM нейронної мережі, що здатна адаптуватися до складних даних з великою кількістю варіативності.

References

1. Bzdok, D., Altman, N., Krzywinski, M. Statistics versus machine learning. 2018. pp. 233–234. [in English]. URL: <https://doi.org/10.1038/nmeth.4642>
2. Brownlee J. *Multi-Step LSTM Time Series Forecasting Models for Power Usage*. [in English]. URL: <https://machinelearningmastery.com/how-to-develop-lstm-models-for-multi-step-time-series-forecasting-of-household-power-consumption/>
3. Verma Y. Complete Guide To Bidirectional LSTM (With Python Codes). [in English]. URL: <https://analyticsindiamag.com/complete-guide-to-bidirectional-lstm-with-python-codes/>
4. Feng W., Tang J., Xiao Liu T., Zhang S., Guan J. Proceedings of the 33rd AAAI Conference on Artificial Intelligence. 2019.Dataset. [in English]. URL: <http://moocdata.cn/data/user-activity#User%20Activity>
5. Dolphin R. LSTM Networks | A Detailed Explanation. [in English]. URL: <https://towardsdatascience.com/lstm-networks-a-detailed-explanation-8fae6aefc7f9>