

*А.Ю. Дорошенко, Р.В. Кушніренко*

## РЕКУРЕНТНІ НЕЙРОННІ МЕРЕЖІ ДЛЯ ЗАДАЧІ УТОЧНЕННЯ ЧИСЕЛЬНИХ МЕТЕОРОЛОГІЧНИХ ПРОГНОЗІВ

Зроблено короткий огляд застосування “глибокого навчання” до геонаукових задач. Порівняні два найпопулярніші види архітектур рекурентних нейронних мереж, а саме мережу довгої короткочасної пам’яті та вентильний рекурентний вузол. Показано, що усі моделі з вентильними рекурентними вузлами є більш ефективними за моделі довгої короткочасної пам’яті. На основі доступних даних спостережень здійснено чисельні експерименти з уточнення прогнозу за допомогою машинного навчання. Виявлено, що кращою архітектурою рекурентних нейронних мереж для розв’язання задачі уточнення чисельних метеорологічних прогнозів є вентильний рекурентний вузол.

Ключові слова: “глибоке навчання”, рекурентні нейронні мережі, метеорологічне прогнозування.

### Вступ

Протягом останнього десятиліття “глибоке навчання” стало важливою частиною дослідницьких і оперативних схем геонаукової обробки, що стосуються атмосфери, поверхні суші та океану. Цьому сприяли, поміж іншого, збільшення доступності даних спостережень, а також підвищення швидкості їхньої передачі, що вже перевищує сотні терабайт на день [1]. Ці дані надходять від безлічі датчиків, що вимірюють різні часово і просторово інтегровані величини. Зокрема, вони включають дані дистанційного зондування на висоті від кількох метрів до сотень кілометрів над Землею, а також спостереження на місці (на поверхні та під нею) за допомогою автономних датчиків.

Хоча “глибоке навчання” досягло помітних успіхів у моделюванні впорядкованих послідовностей і даних із просторовим контекстом у сферах комп’ютерного зору, систем розпізнавання мови та керування [2], а також у таких наукових галузях як фізика [3], хімія [4] та біологія [5], його застосування до проблем геонауки знаходиться в зародковому стані. Однак деякі спроби його застосування до таких ключових проблем як класифікація, регресія, виявлення аномалій та прогнозування залежного від простору або часу стану мають перспективні рішення. До прикладу, є декілька досліджень, що демонструють застосування “глибокого навчання” до проблеми прогнозування екстремальних погодних

умов [6,7]. Зауважимо, що ця задача є проблематичною для традиційного машинного навчання. Згадані дослідження свідчать про успіх у застосуванні архітектур “глибокого навчання” до виокремлення просторових і часових характеристик для визначення та класифікації екстремальних ситуацій (наприклад, штормів) у вихідних даних числової моделі прогнозування погоди. Такий підхід дозволяє швидко виявляти такі події та моделювати прогнози без використання суб’єктивних анотацій людини або методів, які покладаються на заздалегідь визначені порогові значення для швидкості вітру та інших метеорологічних величин.

Нагадаємо, що підходи “глибокого навчання” класично поділяються на просторові (наприклад, згорткові нейронні мережі [8] для класифікації об’єктів) і послідовні (наприклад, рекурентні нейронні мережі для розпізнавання мовлення [9]). Згорткові мережі являють собою стек фільтрів малого розміру з невеликою кількістю параметрів, які власне “навчаються”. Вони застосовуються до зображень або інших даних на прямокутній сітці для отримання узагальнених характерних особливостей досліджуваного об’єкта. У царині геонауки згорткові мережі можуть бути використані для виявлення просторових характеристик, наприклад, під час аналізу супутникових зображень [10]. На противагу рекурентні нейронні мережі були розроблені для вивчення залежних від часу особливостей

даних. Рекурентність — це лише загальна ідея, яка полягає у тому, що топологічно така архітектура нейронної мережі може бути представлена орієнтованим у часі графом. Завдяки цьому утворюється “пам’ять” мережі (вектор внутрішнього стану), що і дозволяє виявляти динамічні (у часовому вимірі) характеристики досліджуваних даних.

Однак спостерігається все більша зацікавленість у поєднанні цих двох підходів. Прототипним прикладом цього поєднання є прогнозування відео та руху [11], проблема, яка має разючу подібність до багатьох динамічних геонаукових проблем. Тут ми стикаємося з багатовимірними структурами, змінними в часі. Наприклад, параметри рослинного покриву, що впливають на вуглецевий цикл та випаровування. Вже існують дослідження, що починають застосовувати комбіновані згортково-рекурентні підходи до таких геонаукових проблем як прогнозування опадів [12]. Моделювання динаміки атмосфери та океану, моделювання поширення вогню чи руху ґрунту також є прикладами проблем, де важлива просторово-часова динаміка. Але наразі вони не отримали переваг від застосування комбінованих згортково-рекурентних підходів “глибокого навчання”.

Коротко кажучи, подібність між типами даних, притаманних класичним застосуванням “глибокого навчання”, і даних, з якими працює геонаука, є переконливим аргументом на користь інтеграції “глибокого навчання” в геонауки. Зображення є аналогом двовимірних полів даних, що містять певні змінні за аналогією з триплетами кольорів (значення RGB) на фотографіях, тоді як відео можна пов’язати з послідовністю зображень, тобто з двовимірними полями, які змінюються у часі. Подібним чином природна мова та мовлення мають такі ж характерні особливості динамічних часових рядів, що їх мають дані, притаманні геонауковій сфері. Крім того, класифікація, регресія, виявлення аномалій і динамічне моделювання є типовими проблемами як для класичних застосувань “глибокого навчання”, так і для геонаук.

Проте, як було показано у [13], ще рано говорити про повний перехід геонау-

ки, зокрема, сфери метеорологічного прогнозування, на методи, що базуються лише на “глибокому навчанні” і спостережуваних даних. Як було зазначено вище, попри те, що “глибоке навчання” останнім часом показує себе успішно у різних сферах, і попри те, що робляться спроби його застосування до метеорологічних задач, такі дослідження все ще знаходяться в зародковому стані. Це пов’язано передусім з тим, що характерні особливості метеорологічних даних вимагають розробки нових підходів поза межами класичних концепцій комп’ютерного зору, розпізнавання мовлення та інших типових задач, поставлених перед “глибоким навчанням”. На відміну від них, підвищення точності прогнозування є хоча і дуже важливим, однак недостатнім компонентом. Дуже важливим складником тут є також надання можливості інтерпретації та розуміння результатів, включаючи їх візуалізацію для аналізу людьми. А, як відомо, інтерпретованість була визначена як потенційна слабкість “глибоких” нейронних мереж, і досягнення цієї мети зараз є центральною проблемою для “глибокого навчання” [14]. Ця галузь все ще є далекою від створення зрозумілих моделей, а також від надання можливості визначення причин закономірностей на основі даних спостережень [15]. Однак ми маємо визнати, що на практиці, враховуючи складність сучасних моделей чисельного метеорологічного прогнозування, також нелегко відстежити зв’язок між результатами їхньої роботи та припущеннями, на яких їх побудовано, а це, очевидно, обмежує їхню інтерпретованість. Крім цього, моделі “глибокого навчання” можуть бути фізично непослідовними або неправдоподібними, хоча їхня статистична точність буде високою. Це може відбуватися через надмірну екстраполяцію та/або статистичну упередженість спостережень. Інтеграція знань предметної області та досягнення фізичної узгодженості моделей за допомогою навчання відповідно до законів фізики, може забезпечити дуже сильні теоретичні обмеження на додаток до спостережуваних даних. Треба зазначити, що робота над підходами, що розв’язували б ці та інші проблеми, триває, і навіть деякі проблеми

тією чи іншою мірою вже розв'язані, та попри це не існує єдиного методу, що розв'язував би усі проблеми одночасно, а саме він і потрібен аби успішно здійснити повний перехід метеорологічного прогнозування на методи “глибокого навчання”.

Однак відповідно до [16], є тенденція до розробки гібридних підходів моделювання, які поєднали б моделі фізичних процесів з універсальністю інструментів “глибокого навчання” для досягнення кращих результатів. Зокрема, існують дослідження можливості успішного використання “глибокого навчання” на різних стадіях чисельного прогнозу: обробка спостережень [17], асиміляція даних [18], прогностична модель [19] та постпроцесинг отриманих результатів [20].

Власне, дана стаття присвячена застосуванню “глибокого навчання” до постпроцесингу результатів прогнозу приземної температури, отриманого за допомогою чисельних гідродинамічних методів метеорологічного прогнозування, а саме порівнянню двох найпопулярніших видів рекурентних нейронних мереж у застосуванні до даної прикладної задачі. Це порівняння дасть змогу виявити оптимальну архітектуру нейронної мережі для подальшого досягнення належного ступеня точності метеорологічних прогнозів.

### Опис даних

“Глибоке навчання” як техніка вивчення характерних особливостей даних суттєво залежить від якості, репрезентативності та цілісності використовуваних даних. Тому правильний відбір і підготовка даних є важливими факторами для отримання хороших і узагальнюючих результатів.

Зокрема, відбір даних має бути спрямований на охоплення якнайбільше повної варіативності значень змінних, на яких базуватиметься власне навчання нейромережевої моделі. Хороші дані мають дозволяти моделі охоплювати зв'язки між змінними, на основі яких робиться прогноз. Водночас важливим є уникнення надлишковості у даних.

Нижче поданий опис даних, використаних для дослідження, описаного у

даній статті. Ці дані склалися з чотириелементних кортежів і містили наступну інформацію:

- дата,
- час за Гринвічем,
- прогнозоване значення температури (Fcst), завчасністю в одну добу від моменту ініціалізації чисельної регіональної моделі,
- спостережуване значення температури (Obs).

Чисельною моделлю прогнозу погоди, результат роботи якої ми хочемо покращити, є модель однойменного європейського консорціуму COSMO (Consortium for Small-scale Modelling). Ця модель використовується в Українському гідрометеорологічному інституті ДСНС України та НАН України для наукових та прикладних задач, починаючи із липня 2011 р. [21]. Нагадаємо, що COSMO є негідростатичною моделлю, яка здатна ефективно відтворювати широкий спектр атмосферних процесів у масштабі мезо- $\beta$  та мезо- $\gamma$ . В основу динамічного ядра моделі покладено рівняння термо- та гідродинаміки, що описують потік у вологій атмосфері. Різноманітні фізичні процеси враховуються схемами параметризації [22].

Рис. 1 зображує розрахункову область чисельної регіональної моделі: кількість вузлів із заходу на схід – 209; кількість вузлів із півдня на північ – 101; кількість рівнів по вертикалі – 50; крок  $\sim$  14 км.



Рис.1. Розрахункова область моделі прогнозу погоди COSMO

Наявні дані охоплюють проміжок часу від 01.07.2012 до 31.03.2014, або 639 днів. Спостереження проводилися кожні три години, а саме о 00:00, 03:00, 06:00, 09:00, 12:00, 15:00, 18:00 і 21:00 за Гринвічем. Для цих же моментів часу обчислювався і прогноз регіональної моделі.

Таким чином, для кожної дати маємо по вісім кортежів. Відповідно 639 днів дають 5112 кортежів.

Що до просторової приналежності, то дані охоплюють спостережувані значення і прогнози для станцій “Біла Церква”, “Бориспіль”, “Київ”, “Миронівка”, “Тетерів”, “Фастів”, “Чорнобиль” та “Яготин”.

### Рекурентні нейронні мережі

Як зазначалося вище, поточне дослідження має за мету порівняння двох найпопулярніших видів архітектур рекурентних нейронних мереж у застосуванні до задачі постпроцесингу результатів прогнозу приземної температури, отриманого за допомогою чисельних гідродинамічних методів метеорологічного прогнозування.

Зауважимо, що використання концепції рекурентних нейронних мереж зумовлене тим, що метеорологічні дані мають природу часового ряду, тобто містять залежні від часу особливості. А рекурентні нейронні мережі були розроблені саме для роботи з даними такого роду.

Двома найпопулярнішими видами рекурентних нейронних мереж, про які йшлося вище, є мережа довгої короткочасної пам’яті (long short-term memory, LSTM) [23] та вентильний рекурентний вузол (gated recurrent unit, GRU) [24]. Ці методи завдячують своєю популярністю тому, що дозволяють уникнути основних проблем, які виникають під час аналізу довгострокових залежностей. Цими проблемами є ефект зникнення градієнтів (vanishing gradient effect) і безпосередньо пов’язаний з ним ефект вибуху градієнтів (exploding gradients effect).

Наведемо теоретичні відомості, що стосуються згаданих вище архітектур нейронних мереж.

Вузол GRU працює наступним чином. У кожний момент часу йому на вхід

подаються вхідний вектор  $x_t$  і вихідний вектор з попереднього моменту часу  $h_{t-1}$ . Вихідний вектор  $h_t$  обчислюється як лінійна інтерполяція між  $h_{t-1}$  і поточним кандидатом  $\tilde{h}_t$ :

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t,$$

де  $z_t$  – вектор вузла уточнення. Він визначає, які долі першого і другого векторів впливатимуть на поточне значення.  $\odot$  позначає добуток Адамара.

Вектор вузла уточнення обчислюється наступним чином:

$$z_t = \sigma_g(W_z x_t + U_z h_{t-1} + b_z).$$

Поточний кандидат  $\tilde{h}_t$  обчислюється подібно до стандартної рекурентної нейронної мережі:

$$\tilde{h}_t = \sigma_h(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h),$$

де  $r_t$  – вектор вузла скидання, що обчислюється подібно до вектора вузла уточнення:

$$r_t = \sigma_g(W_r x_t + U_r h_{t-1} + b_r).$$

$W_z, U_z, W_h, U_h, W_r, U_r$  позначають матриці параметрів,  $b_z, b_h, b_r$  позначають вектори параметрів. Вектори вузлів уточнення і скидання активуються сигмоїдною функцією, а вектор поточного кандидата – гіперболічним тангенсом.

Що до LSTM, то ця архітектура являє собою дещо ускладнену версію GRU. Замість двох вузлів (уточнення та скидання) вона має три — вузли входу, забування та виходу. Звичайно, це робить мережі довгої короткочасної пам’яті повільнішими для навчання та використання, але разом з тим, вони можуть бути ефективнішими, коли йдеться про зберігання та доступ до довгострокових залежностей.

Зауважимо, що і мережі довгої короткочасної пам’яті, і вентильні рекурентні вузли можуть розв’язувати широкий спектр задач, зокрема, розпізнавання мовлення, машинний переклад та прогнозування часових рядів. Як правило, LSTM-



мережі більш ефективні в задачах, які вимагають зберігання та доступ до довгострокових залежностей. З іншого боку, GRU-мережі ефективніші в задачах, які вимагають швидкого навчання та адаптації до нових вхідних даних. Проте варто пам'ятати, що не існує єдиного найбільш ефективного типу рекурентних нейронних мереж, який підходив би для усіх завдань. Тому вибір між LSTM і GRU залежатиме від конкретних вимог розв'язуваної задачі.

Як правило, доцільно спробувати обидві архітектури та порівняти їхню ефективність у застосуванні до тієї чи іншої конкретної задачі. Власне, це і є мета проведення даного дослідження, а саме порівняння цих двох архітектур у застосуванні до задачі постпроцесингу результатів чисельного метеорологічного прогнозування.

### Опис експерименту

Як було зазначено вище, порівняння архітектур рекурентних нейронних мереж буде здійснено у застосуванні до прикладної задачі уточнення прогнозу приземної температури, отриманого за допомогою чисельної моделі прогнозу погоди COSMO [25].

Для кожної метеорологічної станції було натреновано дві нейромережеві моделі (одна з GRU-шаром, інша з LSTM-шаром), які мали б якнайкраще виокремити фізичні особливості конкретного пункту спостереження. Тож, ми маємо 16 натренованих моделей.

Нейромережева топологія та інші конфігураційні параметри були однаковими для усіх моделей. Першим шаром нейронної мережі (після вхідного) було взято рекурентний шар (GRU або LSTM), що складався з 64 вузлів для GRU та 32 вузлів для LSTM, оскільки менша і більша кількість давали гірші результати. Після нього було додано один прихований повнозв'язний шар із 64 вузлами. Останній шар запропонованої архітектури був також повнозв'язним і мав 8 вузлів, оскільки вихідний вектор повинен мати розмірність добового вектора прогнозу (8 значень).

Рис. 2 зображує описану топологію нейромережевих моделей.

Що стосується даних, то для кожної метеорологічної станції уся їх сукупність була розбита на три класи: тренувальні (період з 01.07.2012 до 30.06.2013; 365 днів), валідаційні (період з 01.07.2013 до 31.10.2013; 123 дні) і тестувальні (з 01.11.2013 до 01.04.2014; 151 день).

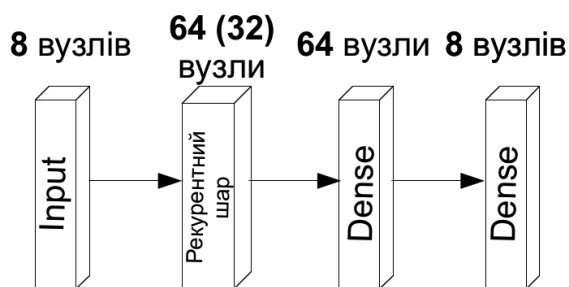


Рис. 2. Запропонована архітектура нейронної мережі

Зауважимо, що поділ на три класи (набори) — це загальна практика для “глибокого навчання” [26]. Тренувальний набір є найбільшим і використовується для оновлення вагових коефіцієнтів моделі шляхом зворотного розповсюдження помилки або інших алгоритмів навчання. Другий набір, валідаційний, використовується виключно для налаштування гіперпараметрів: кількості шарів, типів шарів, функцій активації, цільових функцій, швидкості навчання тощо. Ключовою метою цього налаштування є підвищення здатності мережі до узагальнення для гарантії, що мережа добре функціонуватиме на невідомих для неї даних. Третій набір даних — це тестовий набір, раніше невідомі дані, які використовуються для оцінювання мережі після налаштування.

Значимо, що усі експерименти проводилися з використанням відкритого нейромережевого інтерфейсу Keras [27] і відкритої програмної бібліотеки для “глибокого навчання” TensorFlow [28]. Програмний код був написаний мовою Python [29].

### Отримані результати

Порівняння ефективностей двох архітектур рекурентних нейронних мереж здійснювалось за допомогою стандартного способу оцінювання похибок, а саме за

допомогою середнього квадратичного відхилення (RMSE). Як бачимо з таблиці, усі моделі з вентильними рекурентними вузлами (GRU) є оптимальнішими за моделі довгої короткочасної пам'яті (LSTM).

Таблиця. Порівняння архітектур рекурентних нейронних мереж

Станція	RMSE (GRU), °C	RMSE (LSTM), °C
Біла Церква	1.8865	2.1451
Бориспіль	1.9958	2.1456
Київ	1.9824	2.1334
Миронівка	2.0099	2.4484
Тетерів	2.0215	2.3027
Фастів	1.8565	2.1394
Чорнобиль	2.0094	2.0355
Яготин	2.1157	2.3793

### Висновки

На прикладі прогнозів моделі COSMO приземної температури повітря для восьми метеорологічних станцій Київської області та відповідних їм даних фактичних спостережень було порівняно два найпопулярніші види архітектур рекурентних нейронних мереж, а саме мережу довгої короткочасної пам'яті та вентильний рекурентний вузол.

Показано, що усі моделі з вентильними рекурентними вузлами є ефективнішими за моделі довгої короткочасної пам'яті.

Таким чином, виявлено найкращу архітектуру рекурентних нейронних мереж для розв'язання задачі постпроцесингу результатів чисельного метеорологічного прогнозування.

### References

1. Agapiou, A., 2017. Remote sensing heritage in a petabyte-scale: satellite data and heritage Earth Engine© applications. *International Journal of Digital Earth*, 10(1), pp.85-102.
2. LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. *nature*, 521(7553), pp.436-444.
3. Bhimji, W., Farrell, S.A., Kurth, T., Paganini, M., Prabhat and Racah, E., 2018, September. Deep neural networks for physics analysis on low-level whole-detector data at the LHC. In *Journal of Physics: Conference Series* (Vol. 1085, p. 042034). IOP Publishing.
4. Schütt, K.T., Arbabzadah, F., Chmiela, S., Müller, K.R. and Tkatchenko, A., 2017. Quantum-chemical insights from deep tensor neural networks. *Nature communications*, 8(1), p.13890.
5. Alipanahi, B., Delong, A., Weirauch, M.T. and Frey, B.J., 2015. Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning. *Nature biotechnology*, 33(8), pp.831-838.
6. Liu, Y., Racah, E., Correa, J., Khosrowshahi, A., Lavers, D., Kunkel, K., Wehner, M. and Collins, W., 2016. Application of deep convolutional neural networks for detecting extreme weather in climate datasets. *arXiv preprint arXiv:1605.01156*.
7. Racah, E., Beckham, C., Maharaj, T., Ebrahimi Kahou, S., Prabhat, M. and Pal, C., 2017. Extremeweather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. *Advances in neural information processing systems*, 30.
8. LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), pp.2278-2324.
9. Sak, H., Senior, A. and Beaufays, F., 2014. Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. *arXiv preprint arXiv:1402.1128*.
10. Zhu, X.X., Tuia, D., Mou, L., Xia, G.S., Zhang, L., Xu, F. and Fraundorfer, F., 2017. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE geoscience and remote sensing magazine*, 5(4), pp.8-36.
11. Oh, J., Guo, X., Lee, H., Lewis, R.L. and Singh, S., 2015. Action-conditional video prediction using deep networks in atari

- games. Advances in neural information processing systems, 28.
12. Shi, X., Chen, Z., Wang, H., Yeung, D.Y., Wong, W.K. and Woo, W.C., 2015. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Advances in neural information processing systems, 28.
  13. Schultz, M.G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L.H., Mozaffari, A. and Stadler, S., 2021. Can deep learning beat numerical weather prediction?. Philosophical Transactions of the Royal Society A, 379(2194), p.20200097.
  14. Montavon, G., Samek, W. and Müller, K.R., 2018. Methods for interpreting and understanding deep neural networks. Digital signal processing, 73, pp.1-15.
  15. Runge, J., Petoukhov, V., Donges, J.F., Hlinka, J., Jajcay, N., Vejmelka, M., Hartman, D., Marwan, N., Paluš, M. and Kurths, J., 2015. Identifying causal gateways and mediators in complex spatio-temporal systems. Nature communications, 6(1), p.8502.
  16. Bauer, P., Dueben, P.D., Hoefler, T., Quintino, T., Schulthess, T.C. and Wedi, N.P., 2021. The digital revolution of Earth-system science. Nature Computational Science, 1(2), pp.104-113.
  17. Prudden, R., Adams, S., Kangin, D., Robinson, N., Ravuri, S., Mohamed, S. and Arribas, A., 2020. A review of radar-based nowcasting of precipitation and applicable machine learning techniques. arXiv preprint arXiv:2005.04988.
  18. Bonavita, M. and Laloyaux, P., 2020. Machine learning for model error inference and correction. Journal of Advances in Modeling Earth Systems, 12(12), p.e2020MS002232.
  19. Krasnopolsky, V.M., Fox-Rabinovitz, M.S. and Chalikov, D.V., 2005. New approach to calculation of atmospheric model physics: Accurate and fast neural network emulation of longwave radiation in a climate model. Monthly Weather Review, 133(5), pp.1370-1383.
  20. Rasp, S. and Lerch, S., 2018. Neural networks for postprocessing ensemble weather forecasts. Monthly Weather Review, 146(11), pp.3885-3900.
  21. Shpyg, V., Budak, I., Pishniak, D. and Poperechnyi, P., 2013, November. The application of regional NWP models to operational weather forecasting in Ukraine. In CAS Technical Conference (TECO) on "Responding to the Environmental Stressors of the 21st Century" Available from: <http://www.wmo.int/pages/prog/arep/cas/documents/Ukraine-NWPMODELS.pdf> [Accessed 27/02/2020].
  22. Doms, G. and Baldauf, M., 2011. A description of the nonhydrostatic regional COSMO-Model Part I: dynamics and numerics. Deutscher Wetterdienst, Offenbach.
  23. Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. Neural computation, 9(8), pp.1735-1780.
  24. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y., 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.
  25. Doroshenko, A.Y., Shpyg, V.M. and Kushnirenko, R.V., 2023. Deep learning-based approach to improving numerical weather forecasts. PROBLEMS IN PROGRAMMING, (3), pp.91-98.
  26. Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep learning. MIT press.
  27. <https://keras.io/>
  28. <https://www.tensorflow.org/>
  29. <https://www.python.org/>

Одержано: 01.12.2023

### *Про авторів:*

Дорошенко Анатолій Юхимович, доктор фізико-математичних наук, професор, завідувач відділу ІПС НАНУ, професор кафедри інформаційних систем та технологій КПІ імені Ігоря Сікорського. Кількість наукових публікацій в українських виданнях – понад 200.

Кількість наукових публікацій  
в зарубіжних виданнях – понад 90.  
Індекс Гірша – 7.  
<http://orcid.org/0000-0002-8435-1451>,

Кушніренко Роман Владиславович,  
аспірант.  
Кількість наукових публікацій  
в українських виданнях – 3.  
<https://orcid.org/0000-0002-1990-8727>.

***Місце роботи авторів:***

Інститут програмних систем  
НАН України,  
03187, м. Київ-187,  
проспект Академіка Глушкова, 40.  
Тел.: (38)(044) 526-60-33.  
E-mail:  
[doroshenkoanatoliy2@gmail.com](mailto:doroshenkoanatoliy2@gmail.com),  
[roman.kushnirenk@gmail.com](mailto:roman.kushnirenk@gmail.com).