

<https://doi.org/10.31891/2307-5732-2026-363-76>

УДК 004.8

### ТЕРЕНТЬЄВ ОЛЕКСАНДР

Київський національний університет будівництва і архітектури

<https://orcid.org/0000-0001-9499-6635>

e-mail: [terentiev.oo@knuba.edu.ua](mailto:terentiev.oo@knuba.edu.ua)

### СОЛОВЕЙ БОГДАН

Київський національний університет будівництва і архітектури

<https://orcid.org/0009-0008-0328-1123>

e-mail: [controlerrf5@gmail.com](mailto:controlerrf5@gmail.com)

## БАГАТОЗАДАЧНІ АРХІТЕКТУРИ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ МОНІТОРИНГУ БЕЗПЕКИ МОНОРЕЙКОВОГО КРАНА

Забезпечення безпеки монорейкових кранів є критичним завданням у сучасному будівництві. Сучасним рішенням є застосування рекурентних нейронних мереж (GRU) для аналізу даних крана у реальному часі. Метою даної роботи є розробка та порівняльний аналіз архітектур багатозадачних рекурентних нейронних мереж для одночасного прогнозування стійкості монорейкового крана за критеріями на перекидання та на зсув. Для цього було розроблено та досліджено дві архітектури: багатозадачну GRU з механізмом самоуваги та багатозадачну GRU з байєсівським механізмом уваги.

Дослідження охоплює оцінювання регресійної точності (MAE, RMSE,  $R^2$ ), ефективності бінарної класифікації небезпечних станів (Precision, Recall, Specificity, F1) та характеристик невизначеності прогнозу (ширина 95% довірчого інтервалу, кореляція між помилкою та оцінкою невизначеності). Для статистичної перевірки відмінностей між моделями застосовано критерій Вілкоксона для пов'язаних вибірок.

Отримані результати показали, що обидві моделі забезпечують високу прогностичну здатність ( $R^2 > 0.85$ ). Для задачі прогнозування стійкості на перекидання модель з механізмом самоуваги продемонструвала статистично значуще вищу точність ( $p < 0.001$ ) та зменшення середньої абсолютної помилки на 25.8% порівняно з байєсівською моделлю. Для задачі зсуву статистично значущої різниці в регресійній точності не виявлено ( $p = 0.54$ ), однак байєсівська модель забезпечила вищу повноту виявлення небезпечних станів. Аналіз невизначеності засвідчив більш консервативну поведінку байєсівської моделі, яка формує ширші довірчі інтервали, тоді як модель із самоувагою продемонструвала кращу калібровку невизначеності для задачі зсуву ( $p = 0.93$ ). Статистичні тести підтвердили значущість відмінностей у поведінці моделей щодо оцінювання невизначеності ( $p < 0.001$ ).

Результати дослідження свідчать про доцільність комбінованого використання детермінованої та байєсівської архітектур у системах IT-моніторингу стабільності крана. Подальші дослідження будуть спрямовані на розроблення методу підтримки прийняття рішень, що інтегрує одночасний прогноз обох моделей для формування адаптивного індикатора ризику з урахуванням точності та невизначеності прогнозів.

**Ключові слова:** рекурентна нейронна мережа, GRU, механізм уваги, байєсівське наближення, прогнозування, невизначеність

TERENTYEV OLEKSANDR, SOLOVEI BOHDAN

Kyiv National University of Construction and Architecture

## MULTITASK RECURRENT NEURAL NETWORK ARCHITECTURES FOR INFORMATION TECHNOLOGY-BASED SAFETY MONITORING OF A MONORAIL CRANE

Ensuring the safety of monorail cranes is a critical task in modern construction. A contemporary solution involves the application of recurrent neural networks (GRU) for real-time data analysis. The aim of this study is to develop and comparatively analyze architectures of multitask recurrent neural networks for the simultaneous prediction of monorail crane stability with respect to overturning and sliding criteria. For this purpose, two architectures were developed and investigated: a multitask GRU with a self-attention mechanism and a multitask GRU with a Bayesian attention mechanism.

The study includes the evaluation of regression accuracy (MAE, RMSE,  $R^2$ ), the effectiveness of binary classification of hazardous states (Precision, Recall, Specificity, F1), and the characteristics of predictive uncertainty (width of the 95% confidence interval, correlation between prediction error and uncertainty estimate). The Wilcoxon signed-rank test was applied to statistically assess differences between the models.

The results demonstrate that both models provide high predictive performance ( $R^2 > 0.85$ ). For the task of overturning stability prediction, the self-attention model showed statistically significantly higher accuracy ( $p < 0.001$ ) and a 25.8% reduction in mean absolute error compared to the Bayesian model. For the sliding stability task, no statistically significant difference in regression accuracy was found ( $p = 0.54$ ); however, the Bayesian model achieved higher recall in detecting hazardous states. Uncertainty analysis revealed a more conservative behavior of the Bayesian model, which generates wider confidence intervals, whereas the self-attention model demonstrated better uncertainty calibration for the sliding task ( $p = 0.93$ ). Statistical tests confirmed significant differences in the models' behavior regarding uncertainty estimation ( $p < 0.001$ ).

The findings indicate the feasibility of combining deterministic and Bayesian architectures in IT-based crane stability monitoring systems. Future research will focus on developing a decision support method that integrates simultaneous predictions from both models to form an adaptive risk indicator that accounts for both prediction accuracy and uncertainty.

**Keywords:** recurrent neural network, GRU, attention mechanism, Bayesian approximation, prediction, uncertainty.

Стаття надійшла до редакції / Received 18.01.2026

Прийнята до друку / Accepted 11.02.2026

Опубліковано / Published 26.03.2026



This is an Open Access article distributed under the terms of the [Creative Commons CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)

© Терентьєв Олександр, Соловей Богдан

### Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

Інтелектуальні IoT-системи з використанням машинного навчання продемонстрували здатність аналізувати потоки даних із сенсорів і камер у режимі реального часу для виконання прогностичної аналітики та своєчасного виявлення потенційних загроз на будівельних майданчиках. Зокрема, у дослідженні [1] показано ефективність інтеграції IoT-датчиків і методів глибокого навчання для моніторингу ризиків і попередження небезпечних ситуацій на будівельних об'єктах, а система моніторингу безпеки будівельного майданчика з використанням комп'ютерного зору продемонструвала можливість застосування ML-методів у реальному часі для аналізу дій машин та персоналу [2].

У роботі [3] запропоновано архітектуру рекурентної нейронної мережі з механізмом уваги та байєсівським наближенням для прогнозування стійкості монорейкового крана. Під час навчання механізм уваги забезпечив виділення найбільш значущих ознак, що дозволило уникнути проблеми згасання градієнтів та засвідчило стабільність моделі, а байєсівське наближення дало змогу оцінювати невизначеність прогнозу. Експериментально досягнуто високу якість прогнозування без перенавчання ( $MSE = 0.009$ ,  $MAE = 0.066$ , точність  $\approx 0.96$ ). Запропоновану модель рекомендовано для використання у системах моніторингу стійкості монорейкового крана.

Проте задача оцінювання стійкості монорейкового крана є за своєю природою багатозадачною, оскільки втрата стійкості може відбуватися за різними, хоча й пов'язаними умовами: за умовою перекидання крана відносно ходового колеса (коли коефіцієнт стійкості  $\beta \leq 1,4$ ) та за умовою зрушення (зсуву) під дією горизонтальної сили  $H$  (коли коефіцієнт стійкості  $\gamma \leq 1,5 \dots 1,6$ ) [4]. Моделювання цих процесів окремо є неефективним, оскільки ігнорує спільні фізичні параметри та закономірності. Крім того, рекурентна нейронна мережа з механізмом уваги та байєсівським наближенням має обмеження щодо моделювання глобальних часових залежностей між ознаками, що ефективніше реалізується механізмом самоуваги [5].

Це створює передумови для даного дослідження, метою якого є визначення найбільш придатної архітектури багатозадачної рекурентної нейронної мережі GRU для реалізації в інформаційній технології моніторингу безпеки стрілоподібного пересувного монорейкового крана.

#### Аналіз досліджень та публікацій

Класичні моделі машинного навчання, навчаючись виключно на даних, можуть формувати фізично некоректні або неправдоподібні прогнози. Для розв'язання цієї проблеми активно розвивається напрям фізично-інформованого навчання, у межах якого нейронні мережі навчаються з урахуванням додаткової інформації, отриманої із застосуванням фізичних законів. Такий підхід інтегрує експериментальні дані та математичні моделі й реалізується засобами нейронних мереж [6-7].

У дослідженні [8] обґрунтовано підхід, згідно з яким прогноз нейронної мережі має узгоджуватися з фізичними законами, а відхилення від них повинні штрафуватися у процесі навчання. Це реалізується шляхом модифікації функції втрат:

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_{physics}, \quad (1)$$

де  $\mathcal{L}_1$  - відповідає за точність прогнозу відносно даних, а  $\mathcal{L}_{physics}$  - складова, що забезпечує узгодженість прогнозу з фізичними обмеженнями.

У дослідженні фізична модель інтегрується безпосередньо у прямий прохід рекурентної мережі через вихідні нейрони, що фактично еквівалентно введенню фізично обґрунтованого зсуву у процесі формування прогнозу. Такий підхід гарантує, що мережа формує результати, які не суперечать законам механіки, а байєсівське навчання додатково стабілізує оптимізацію та забезпечує ймовірнісну інтерпретацію результатів. Це дослідження методологічно обґрунтовує доцільність включення у функцію втрат для задач прогнозування стійкості монорейкового крана штрафного члена за порушення умов рівноваги та стійкості.

У статті [9] автори аналізують багатозадачне навчання на прикладі задач комп'ютерного зору (семантична сегментація, оцінювання глибини сцени, визначення напрямків поверхні). Показано, що просте додавання втрат окремих задач  $\sum_t \mathcal{L}_t$  або їх зважування за невизначеністю не гарантує узгодженості результатів між задачами, оскільки кожна з них може оптимізуватись у власному напрямку. Автори доводять необхідність введення додаткових обмежень або регуляризаційних членів, які враховують взаємозв'язок між виходами задач у функції втрат:

$$\mathcal{L} = \sum_t \mathcal{L}_t + \mathcal{L}_{corr}, \quad (2)$$

де  $\mathcal{L}_{corr}$  - контролює узгодженість між завданнями.

Для задачі прогнозування стійкості крана це має принципове значення, оскільки коефіцієнти стійкості на перекидання  $\beta$  та зсув  $\gamma$  фізично пов'язані однією геометрією та системою сил. Отже, введення у функцію втрат доданка виду  $\lambda \cdot \text{corr}(\hat{y}_\beta, \hat{y}_\gamma)$  має як фізичне, так і методологічне підґрунтя, де  $\hat{y}_\beta, \hat{y}_\gamma$  - прогнозовані моделлю значення коефіцієнтів стійкості  $\beta$  та  $\gamma$ .

У дослідженні [10] запропоновано гібридну модель на основі двонапрямлених архітектур довготривалої короткочасної пам'яті та керованих рекурентних блоків, доповнену механізмом самоуваги. Механізм самоуваги, що обчислює попарну взаємодію між усіма часовими кроками послідовності за допомогою матриць Запиту (Q), Ключа (K) та Значення (V), дозволяє моделі виявляти довгострокові та нелінійні залежності на основі матриці уваги A:

$$A = \text{softmax} \frac{QK}{\sqrt{d_k}}. \quad (3)$$

Показано суттєве покращення метрик коефіцієнта детермінації ( $R^2$ ), ефективності Неша–Саткліффа (NSE) та зниження середньоквадратичної помилки (RMSE). Робота підтверджує ефективність механізму самоуваги для часових рядів, однак не розглядає фізичні обмеження та багатозадачність.

В роботі [11] запропонована рекурентна мережа GRU з багатоголовим механізмом самоуваги та агрегуванням ознак для виявлення наявності серцево-судинного захворювання за табличними клінічними даними пацієнта. В архітектурі до матриці прихованих станів  $H \in \mathbb{R}^{T \times d}$  додається навчуваний вектор CLS, після чого формується розширена послідовність станів  $\hat{H}$ . Із цієї послідовності лінійними проєкціями отримуються матриці механізму самоуваги  $Q = \hat{H}W_Q, K = \hat{H}W_K, V = \hat{H}W_V$ . Матриця уваги (3) використовується для оновлення представлення CLS:

$$CLS^{(l)} = CLS^{(l-1)} + A^{(l)}V^{(l)}. \quad (4)$$

Крок (4) повторюється для  $l=1, \dots, L$  шарів, після чого вектор CLS стає узагальненим представленням усієї послідовності. Для нього застосовується нормалізація  $z = \text{LayerNorm}(CLS)$ , та двошаровий повнозв'язний перетворювач для формування прогнозу. Формула (4) є реалізацією залишкового з'єднання, яке дозволяє накопичувати інформацію з різних шарів механізму самоуваги без втрати попереднього стану, а нормалізація узгоджує масштаб агрегованих ознак перед подачею на багатощаровий перцептрон.

Запропонована в роботі [11] архітектура є концептуально придатною для задач прогнозування стійкості монорейкового крана, оскільки дозволяє поєднувати локальні залежності між фізичними параметрами (через GRU) та глобальну оцінку їх взаємного впливу (через механізм самоуваги). Водночас вона потребує модифікації з метою підтримки багатозадачного прогнозування та зміни функції втрат.

На основі проведеного аналізу сучасних досліджень з розробки багатозадачних рекурентних мереж, для прогнозування стійкості монорейкового крана необхідно розробити архітектуру, яка включатиме:

1. Паралельні процеси для спеціалізованої обробки ознак стійкості крана за умов на перекидання та на зсув.
2. Механізм самоуваги в кожному процесі для виявлення глобальних часових залежностей (на основі дослідження [10]).
3. Залишкові з'єднання та нормалізацію шару для стабільного та глибокого навчання (на основі дослідження [11]).
4. Композитну функцію втрат, що включає як фізично-обґрунтований штраф ([8]), так і регуляризатор кореляції між завданнями ([9]).

### Формулювання цілей статті

**Метою роботи є:** забезпечення оцінювання стійкості стрілоподібного пересувного монорейкового крана одночасно за критеріями на перекидання та на зсув шляхом розроблення архітектури рекурентної нейронної мережі GRU для реалізації в інформаційній технології моніторингу безпеки.

Для досягнення поставленої мети визначено такі завдання:

Розробити архітектуру багатозадачної рекурентної мережі GRU з механізмом самоуваги для прогнозування стійкості стрілоподібного пересувного монорейкового крана.

Удосконалити розроблену в роботі [3] архітектуру рекурентної нейронної мережі GRU з механізмом уваги та байєсівським наближенням з метою уможливлення багатозадачного прогнозування стійкості стрілоподібного пересувного монорейкового крана.

Виконати порівняльний аналіз двох архітектур багатозадачних рекурентних нейронних мереж GRU з метою визначення архітектури для реалізації в інформаційній технології моніторингу безпеки.

### Виклад основного матеріалу

**Розробка багатозадачних рекурентних нейронних мереж GRU, призначених для одночасного прогнозування стійкості монорейкового крана за двома критеріями — на перекидання та на зсув.**

Для архітектур багатозадачних рекурентних нейронних мереж GRU, призначених для одночасного прогнозування стійкості монорейкового крана за двома критеріями — на перекидання та на зсув, — вхідна послідовність  $X \in \mathbb{R}^{T \times F}$ , де  $T$  – довжина послідовності,  $F$  – кількість ознак, розділяється на дві під-послідовності, що відповідають окремим задачам:  $X^{(\beta)} \in \mathbb{R}^{T \times F_\beta}$ ,  $X^{(\gamma)} \in \mathbb{R}^{T \times F_\gamma}$ . Таким чином, архітектура складається з двох паралельних процесів  $t \in \{\beta, \gamma\}$ , кожен з яких спеціалізується на своєму завданні оцінювання стійкості монорейкового крана, з подальшим об'єднанням отриманої інформації у спільних шарах моделі.

Архітектура багатозадачної рекурентної мережі GRU з механізмом самоуваги представлена на рис. 1. Рівняння прямого поширення для кожного процесу з детермінованими вагами має вигляд:

$$h_t^{(\beta)} = \text{GRU}_\beta(x_t^{(\beta)}, h_{t-1}^{(\beta)}), h_t^{(\gamma)} = \text{GRU}_\beta(x_t^{(\gamma)}, h_{t-1}^{(\gamma)}), \forall t \in [1, T], \quad (5)$$

де  $h_t^{(\beta)}, h_t^{(\gamma)}$  – прихований стан GRU на кроці  $t$  для відповідного завдання  $t$ .

У результаті формуються дві послідовності прихованих станів  $H_\beta$  та  $H_\gamma$ , які відображають локальну рекурентну пам'ять процесу:

$$H^{(\beta)} = [h_1^{(\beta)}, \dots, h_T^{(\beta)}], H^{(\gamma)} = [h_1^{(\gamma)}, \dots, h_T^{(\gamma)}]. \quad (6)$$

Для кожної послідовності  $H^{(\beta)}, H^{(\gamma)}$  прихованих станів обчислюються матриці запитів, ключів та значень:

$$Q^{(\beta)} = H^{(\beta)}W_Q^{(\beta)}, K^{(\beta)} = H^{(\beta)}W_K^{(\beta)}, V^{(\beta)} = H^{(\beta)}W_V^{(\beta)}, \quad (7)$$

$$Q^{(\gamma)} = H^{(\gamma)}W_Q^{(\gamma)}, K^{(\gamma)} = H^{(\gamma)}W_K^{(\gamma)}, V^{(\gamma)} = H^{(\gamma)}W_V^{(\gamma)}, \quad (8)$$

де  $W_Q^{(\beta)}, W_K^{(\beta)}, W_V^{(\beta)}, W_Q^{(\gamma)}, W_K^{(\gamma)}, W_V^{(\gamma)}$  - вагові матриці механізму самоуваги для відповідного завдання, що навчаються.

Глобальні часові залежності по часовій осі визначається на основі механізму самоуваги як:

$$H_A^{(\beta)} = softmax\left(\frac{Q^{(\beta)}K^{(\beta)T}}{\sqrt{d_K}}\right)V^{(\beta)}, H_A^{(\gamma)} = softmax\left(\frac{Q^{(\gamma)}K^{(\gamma)T}}{\sqrt{d_K}}\right)V^{(\gamma)}. \quad (9)$$

Матриця  $H_A^{(\beta)} \in R^{T \times H^{(\beta)}}$ ,  $H_A^{(\gamma)} \in R^{T \times H^{(\gamma)}}$  - містить оновлені стани, у яких кожен момент часу враховує інформацію з усієї послідовності. Для отримання послідовності, яка поєднує рекурентно накопиченої фізичної динаміки  $H^{(\tau)}$  з глобальною оцінкою важливості подій у часі  $H_A^{(\tau)}$ , що формується механізмом самоуваги застосовується залишкове з'єднання з наступною нормалізацією:

$$\tilde{h}_t^{(\beta)} = \alpha \odot \frac{(H_t^{(\beta)} + H_{t,A}^{(\beta)}) - \mu_t}{\sqrt{\sigma_t^2 + \epsilon}} + \varphi, \tilde{h}_t^{(\gamma)} = \alpha \odot \frac{(H_t^{(\gamma)} + H_{t,A}^{(\gamma)}) - \mu_t}{\sqrt{\sigma_t^2 + \epsilon}} + \varphi, \forall t \in [1, T], \quad (10)$$

де  $\mu_t = \frac{1}{H} \sum_{i=1}^H (h_{t,i}^{(\tau)} + h_{t,A,i}^{(\tau)})$  - середнє значення;  $\sigma^2 = \frac{1}{H} \sum_{i=1}^H ((h_{t,i}^{(\tau)} + h_{t,A,i}^{(\tau)}) - \mu)^2$  - дисперсія.

Функція (9) виконує нормалізацію кожного вектора ознак по його власних компонентах, забезпечуючи нульове середнє та одиничну дисперсію з подальшим масштабуванням і зсувом за допомогою параметрів  $\alpha \in R^H$ ;  $\varphi \in R^H$  значення яких оптимізуються в процесі навчання моделі шляхом мінімізації функції втрат  $\mathcal{L}$  за допомогою градієнтного спуску:

$$\alpha^{(i+1)} = \alpha^{(i)} - \frac{\partial \mathcal{L}}{\partial \alpha}, \varphi^{(i+1)} = \varphi^{(i)} - \frac{\partial \mathcal{L}}{\partial \varphi}. \quad (11)$$

Інтегральне представлення часової послідовності формується як:

$$z^{(\beta)} = \frac{1}{T} \sum_{t=1}^T \tilde{h}_t^{(\beta)} + \max_{t \in [1, T]} \tilde{h}_t^{(\beta)}, z^{(\gamma)} = \frac{1}{T} \sum_{t=1}^T \tilde{h}_t^{(\gamma)} + \max_{t \in [1, T]} \tilde{h}_t^{(\gamma)}. \quad (12)$$

Отримані вектори об'єднуються:

$$z = [z^{(\beta)} || z^{(\gamma)}]. \quad (13)$$

Спільний блок перетворення ознак реалізовано у вигляді резидуального багатошарового перцептрона:

$$z_{k+1} = z_k + f_k(z_k), \quad (14)$$

Де  $f_k$  - нелінійні шари багатошарового перцептрона (MLP) з залишковим з'єднанням.

Модель прогнозує параметри нормального розподілу для завдань визначення стійкості на перекидання та на зсув:

$$\hat{\beta} = \mu_\beta, s_\beta = \log \sigma_\beta^2. \quad (15)$$

$$\hat{\gamma} = \mu_\gamma, s_\gamma = \log \sigma_\gamma^2. \quad (16)$$

Функція втрат включає доданки, які відповідають регресії з неоднорідною дисперсією залишків, фізичному штрафу за порушення умов стійкості та регуляризації узгодженості між завданнями.

$$\mathcal{L} = e^{-s_\beta} (\beta - \mu_\beta)^2 + s_\beta + e^{-s_\gamma} (\gamma - \mu_\gamma)^2 + s_\gamma + \lambda_\beta [\max(0, 1.4 - \hat{\beta})]^2 + \lambda_\gamma [\max(0, 1.5 - \hat{\gamma})]^2 + \lambda_{corr} (1 - \text{corr}(\hat{\beta}, \hat{\gamma})), \quad (17)$$

Де  $\beta, \gamma$  - істинні значення коефіцієнтів стійкості монорейкового крана на перекидання та зсув;  $\mu_\beta, \mu_\gamma$  - передбачені середні коефіцієнтів стійкості монорейкового крана на перекидання та зсув;  $s_\beta, s_\gamma$  - передбачені логарифми дисперсії ( $s = \log \sigma^2$ );  $\hat{\beta}, \hat{\gamma}$  - прогнозовані значення коефіцієнтів стійкості монорейкового крана на перекидання та зсув;  $\lambda_\beta, \lambda_\gamma, \lambda_{corr}$  - вагові коефіцієнти доданків;  $\text{corr}(\cdot)$  - коефіцієнт кореляції між прогнозами.

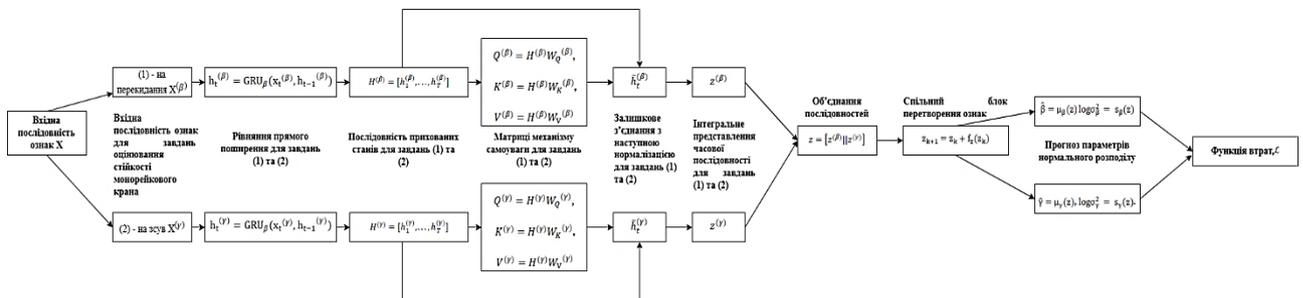


Рис. 1. Архітектура багатозадачної рекурентної мережі GRU з механізмом самоуваги

Удосконалена архітектура рекурентної нейронної мережі GRU з механізмом уваги та байєсівським наближенням представлена на рисунку 2. В архітектурі, запропонованій у попередній роботі [3], основні байєсівські принципи параметризації ваг та механізму уваги залишаються концептуально незмінними, проте в даній роботі вони поширюються на два паралельні процеси GRU та два незалежні механізми уваги.

Рівняння прямого поширення для кожного процесу з детермінованими вагами має вигляд:

$$h_t^{(\beta)} = GRU_\tau(x_t^{(\beta)}, h_{t-1}^{(\beta)}), h_t^{(\gamma)} = GRU_\tau(x_t^{(\gamma)}, h_{t-1}^{(\gamma)}). \quad (18)$$

Оцінки значущості  $e_t^{(\beta)}$ ,  $e_t^{(\gamma)}$  для кожного прихованого стану  $h_t^{(\beta)}$ ,  $h_t^{(\gamma)}$ :

$$e_t^{(\beta)} = v_a^{(\beta)T} \tanh(W_{a(\beta)} h_t^{(\beta)}), \quad e_t^{(\gamma)} = v_a^{(\gamma)T} \tanh(W_{a(\gamma)} h_t^{(\gamma)}). \quad (19)$$

Вагові коефіцієнти уваги  $\alpha_t^{(\beta)}$ ,  $\alpha_t^{(\gamma)}$  обчислюються за функцією softmax:

$$\alpha_t^{(\beta)} = \frac{\exp(e_t^{(\beta)})}{\sum_{j=1}^T \exp(e_j^{(\beta)})}, \quad \alpha_t^{(\gamma)} = \frac{\exp(e_t^{(\gamma)})}{\sum_{j=1}^T \exp(e_j^{(\gamma)})}. \quad (20)$$

Котекстні вектори для кожного процесу  $c^{(\beta)}$ ,  $c^{(\gamma)}$  є зваженою сумою відповідних прихованих станів  $h_t^{(\beta)}$ ,  $h_t^{(\gamma)}$ :

$$c^{(\beta)} = \sum_{t=1}^T \alpha_t^{(\beta)} h_t^{(\beta)}, \quad c^{(\gamma)} = \sum_{t=1}^T \alpha_t^{(\gamma)} h_t^{(\gamma)}. \quad (21)$$

Багатошарові перцептрони відповідних задач включають стохастичну активацію Dropout, що реалізує Monte Carlo апроксимацію варіаційного байєсівського висновку:

$$(\mu_\beta, \log \sigma_\beta^2) = \text{MLP}_\beta(c^{(\beta)}; W_\beta). \quad (22)$$

$$(\mu_\gamma, \log \sigma_\gamma^2) = \text{MLP}_\gamma(c^{(\gamma)}; W_\gamma). \quad (23)$$

Модель прогнозує параметри нормального розподілу для завдань визначення стійкості на перекидання та на зсув:

$$\hat{\beta} = \mu_\beta, \quad s_\beta = \log \sigma_\beta^2. \quad (24)$$

$$\hat{\gamma} = \mu_\gamma, \quad s_\gamma = \log \sigma_\gamma^2. \quad (25)$$

Функція втрат включає від'ємне ELBO, яка узагальнюється на дві правдоподібності  $p(\beta_t | X_t, \tilde{W})$  та  $p(\gamma_t | X_t, \tilde{W})$ , тоді як регуляризатор Кульбака–Лейблера між апостеріорним та апіорним розподілами застосовується до сукупності ваг обох гілок. Додатково до негативного ELBO вводяться фізично-обґрунтовані штрафи за порушення порогів стійкості та регуляризація узгодженості між задачами через кореляційний доданок.

$$\mathcal{L}_{\text{loss}}^{(i)} = -\mathbb{E}_{q(\tilde{W})} \left[ \sum_{t=t_i}^{t_i+k-1} (\log p(\beta_t | X_t, \tilde{W}) + \log p(\gamma_t | X_t, \tilde{W})) \right] + \text{KL}(q(\tilde{W}) || p(\tilde{W})) + \lambda_\beta [\max(0, 1.4 - \hat{\beta})]^2 + \lambda_\gamma [\max(0, 1.5 - \hat{\gamma})]^2 + \lambda_{\text{corr}} (1 - \text{corr}(\hat{\beta}, \hat{\gamma})), \quad (26)$$

Де  $\tilde{W}$  - набір усіх вагових коефіцієнтів моделі.

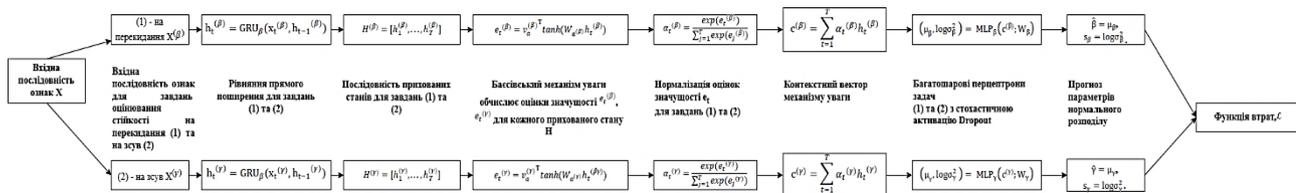


Рис. 2. Архітектура багатозадачної рекурентної мережі GRU з механізмом уваги та байєсівським наближенням

### Методи порівняння ефективності розроблених архітектур багатозадачної рекурентної мережі GRU.

Для порівняння ефективності розроблених архітектур багатозадачної рекурентної мережі GRU з механізмом самоуваги (далі - модель 1) та багатозадачної рекурентної мережі GRU з механізмом уваги та байєсівським наближенням (далі - модель 2) визначимо три групи метрик: 1) метрики регресійної точності (27-29) [12]; 2) метрики бінарної класифікації безпеки (30-33) [13]; 3) метрики оцінювання невизначеності прогнозу (34-39) [14]. Сформулюємо дослідницькі гіпотези:  $H_1$  - одна з моделей демонструє статистично значуще вищу точність у завданні регресійного прогнозування коефіцієнтів стійкості.  $H_2$  - одна з моделей є більш надійною у завданні класифікації безпеки, демонструючи кращий баланс між здатністю виявляти небезпечні стани та мінімізацією хибних тривог.  $H_3$  - моделі демонструють статистично значущі відмінності у своїй поведінці щодо оцінки невизначеності, зокрема у консервативності прогнозів.

Середню величину відхилення прогнозів моделей 1–2 від істинного значення визначимо на основі середньої абсолютної помилки (MAE) та середньоквадратичної помилки (RMSE):

$$MAE_\beta = \frac{1}{N} \sum_{i=1}^N |\beta_i - \hat{\beta}_i|, \quad MAE_\gamma = \frac{1}{N} \sum_{i=1}^N |\gamma_i - \hat{\gamma}_i|, \quad (27)$$

$$RMSE_\beta = \sqrt{\frac{1}{N} \sum_{i=1}^N (\beta_i - \hat{\beta}_i)^2}, \quad RMSE_\gamma = \sqrt{\frac{1}{N} \sum_{i=1}^N (\gamma_i - \hat{\gamma}_i)^2}. \quad (28)$$

Частку дисперсії істинних даних, пояснених моделями 1-2, обчислимо на основі коефіцієнта детермінації  $R^2$ :

$$R_\beta^2 = 1 - \frac{\sum_{i=1}^N (\beta_i - \hat{\beta}_i)^2}{\sum_{i=1}^N (\beta_i - \bar{\beta})^2}, \quad R_\gamma^2 = 1 - \frac{\sum_{i=1}^N (\gamma_i - \hat{\gamma}_i)^2}{\sum_{i=1}^N (\gamma_i - \bar{\gamma})^2}, \quad (29)$$

де  $\bar{\beta}$ ,  $\bar{\gamma}$  - середні значення істинних даних.

На основі граничних значень коефіцієнтів стійкості ( $\beta=1.4$  та  $\gamma=1.5$ ) регресійна задача перетворюється на задачу бінарної класифікації. Перевірка гіпотези  $H_2$  здійснюється шляхом порівняльного аналізу метрик, отриманих із матриць помилок, побудованих окремо для прогнозів моделей 1 та 2.

Нехай матриця помилок моделі 1 визначає:  $TP_\beta$  – небезпечні стани, коректно визначені моделлю 1,  $TN_\beta$  – безпечні стани, коректно визначені моделлю 1,  $FP_\beta$  – хибні тривоги (безпечний стан визначено як

небезпечний) моделлю 1;  $FN_{\beta}$  – пропущені моделлю 1 небезпечні стани. Аналогічно, матриця помилок для моделі 2 включає:  $TP_{\gamma}$ ,  $TN_{\gamma}$ ,  $FP_{\gamma}$ ,  $FN_{\gamma}$ . Тоді, наслідком сценаріїв з  $FN_{\beta}$  або  $FN_{\gamma}$  є втрата стійкості монорейкового крана, яку відповідна модель пропустила, тоді як  $FP_{\beta}$  або  $FP_{\gamma}$  призводять до необґрунтованої зупинки крана – модель помилково визначає втрату стійкості, яка фактично відсутня.

Точність з якою моделі 1-2 ідентифікують небезпечний стан крана для кожного завдання:

$$Recall_{\beta} = \frac{TP_{\beta}}{TP_{\beta} + FN_{\beta}}, Recall_{\gamma} = \frac{TP_{\gamma}}{TP_{\gamma} + FN_{\gamma}}. \quad (30)$$

Частка істинно небезпечних станів серед усіх, що були визначені моделями 1-2 як небезпечні:

$$Precision_{\beta} = \frac{TP_{\beta}}{TP_{\beta} + FP_{\beta}}, Precision_{\gamma} = \frac{TP_{\gamma}}{TP_{\gamma} + FP_{\gamma}}. \quad (31)$$

Здатність моделей 1-2 коректно ідентифікувати безпечні стани:

$$Specificity_{\beta} = \frac{TN_{\beta}}{TN_{\beta} + FP_{\beta}}, Specificity_{\gamma} = \frac{TN_{\gamma}}{TN_{\gamma} + FP_{\gamma}}. \quad (32)$$

Загальна точність моделей 1-2 буде визначена як гармонійне середнє між точністю та повнотою, що дає збалансовану оцінку класифікатора:

$$F_{1,\beta} = \frac{2 \cdot Precision_{\beta} \cdot Recall_{\beta}}{Precision_{\beta} + Recall_{\beta}}, F_{1,\gamma} = \frac{2 \cdot Precision_{\gamma} \cdot Recall_{\gamma}}{Precision_{\gamma} + Recall_{\gamma}}. \quad (33)$$

Консервативність моделей 1–2 щодо ризику невизначеності, пов'язаної з прогнозом, визначимо на основі ширини 95% довірчого інтервалу прогнозу:

$$CI_{95\%,\beta} = [\hat{\beta}_i - z\sigma_{i,\beta}, \hat{\beta}_i + z\sigma_{i,\beta}], CI_{95\%,\gamma} = [\hat{\gamma}_i - z\sigma_{i,\gamma}, \hat{\gamma}_i + z\sigma_{i,\gamma}], \quad (34)$$

Де  $z=1.96$  – квантиль стандартного нормального розподілу для довірчої ймовірності 95%.  $2z\sigma_{i,\beta}$  – ширина  $i$ -го інтервалу для моделі 1;  $2z\sigma_{i,\gamma}$  – ширина  $i$ -го інтервалу для моделі 2.

Для оцінки калібрування невизначеності моделей 1–2 використовується коефіцієнт кореляції Пірсона між абсолютною помилкою прогнозу та оцінкою невизначеності:

$$\rho_{|e|,\sigma}^{(\beta)} = \text{corr}(|\beta_i - \hat{\beta}_i|, \sigma_{i,\beta}), \rho_{|e|,\sigma}^{(\gamma)} = \text{corr}(|\gamma_i - \hat{\gamma}_i|, \sigma_{i,\gamma}) \quad (35)$$

Високе значення коефіцієнтів  $\rho_{|e|,\sigma}^{(\beta)}$ ,  $\rho_{|e|,\sigma}^{(\gamma)}$  свідчить про те, що модель у складних фізичних режимах генерує ширші довірчі інтервали, тобто коректно «усвідомлює» власну помилку.

Оскільки розподіл похибок прогнозування моделей 1-2 може бути ненормальним, як основний статистичний критерій для парного порівняння моделей обрано критерій Вілкоксона для пов'язаних вибірок (Wilcoxon signed-rank test) [15].

Нехай для кожного спостереження абсолютні похибки моделей 1–2 для завдання прогнозування стійкості монорейкового крана на перекидання обчислюються як:  $e_{i,\beta}^{(1)} = |\beta_i - \hat{\beta}_i^{(1)}|$ ,  $e_{i,\beta}^{(2)} = |\beta_i - \hat{\beta}_i^{(2)}|$ . Тоді парні різниці абсолютних похибок моделей 1-2:  $d_{i,\beta} = e_{i,\beta}^{(1)} - e_{i,\beta}^{(2)}$ . Нульова гіпотеза  $H_{0d}^{(\beta)}$  полягає в тому, що медіана різниць  $d_{i,\beta}$  дорівнює нулю. Відхилення  $H_{0d}^{(\beta)}$  свідчатиме про наявність статистично значущої різниці в точності прогнозів моделей 1-2 для завдання прогнозування стійкості крана на перекидання:

$$H_{0d}^{(\beta)} = \text{median}(d_{i,\beta}) = 0. \quad (36)$$

Аналогічно, для завдання визначення стійкості крана на зсув парні різниці абсолютних похибок моделей 1-2:  $d_{i,\gamma} = e_{i,\gamma}^{(1)} - e_{i,\gamma}^{(2)}$ . Нульова гіпотеза  $H_{0d}^{(\gamma)}$  для порівняння моделей 1-2 для завдання прогнозування стійкості крана на зсув полягає в тому, що медіана різниць  $d_{i,\gamma}$  дорівнює нулю. Відхилення  $H_{0d}^{(\gamma)}$  свідчатиме про наявність статистично значущої різниці в точності прогнозів моделей 1-2 для завдання прогнозування стійкості крана на перекидання:

$$H_{0d}^{(\gamma)} = \text{median}(d_{i,\gamma}) = 0. \quad (37)$$

Нульова гіпотеза  $H_{0\Delta}^{(\beta)}$  для перевірки значущості різниці поведінки моделей 1-2 щодо невизначеності прогнозів для завдання прогнозування стійкості крана на перекидання на основі ширини 95% довірчих інтервалів  $\Delta_{i,\beta} = w_{i,\beta}^{(1)} - w_{i,\beta}^{(2)}$  полягає в тому, що медіана різниць  $\Delta_{i,\beta}$  дорівнює нулю. Відхилення  $H_{0\Delta}^{(\beta)}$  доведе, що моделі мають статистично різну поведінку щодо оцінювання невизначеності:

$$H_{0\Delta}^{(\beta)} = \text{median}(\Delta_{i,\beta}) = 0. \quad (38)$$

Аналогічно, для перевірки значущості різниці поведінки моделей 1-2 щодо невизначеності прогнозів для завдання прогнозування стійкості крана на зсув, відхилення нульова гіпотеза  $H_{0\Delta}^{(\gamma)}$  доведе, що моделі мають статистично різну поведінку щодо оцінювання невизначеності:

$$H_{0\Delta}^{(\gamma)} = \text{median}(\Delta_{i,\gamma}) = 0. \quad (39)$$

Для проведення експериментальних досліджень із моделями 1–2 були розроблені синтетичні набори даних за методом, описаним у роботі [3].

## Результати

Обидві моделі продемонстрували прогностичну здатність, що підтверджується коефіцієнтами детермінації  $R^2 > 0.85$  для всіх завдань. Однак кількісний аналіз метрик (таблиці 1–2) та візуальний аналіз графіків (рис. 3–4) виявляють різну спеціалізацію моделей.

Для завдання прогнозування стійкості монорейкового крана на перекидання модель 1 показала статистично значуще вищу точність (таблиця 3,  $p < 0.001$ ). Її середня абсолютна помилка ( $MAE_{\beta} = 0.049$ ) є на 25.8% нижче, ніж у Моделі 2 ( $MAE_{\beta} = 0.066$ ). Це свідчить про те, що архітектура GRU з механізмом самоуваги краще апроксимує фізичну FEM-модель для задачі прогнозування стійкості на перекидання.

Для завдання прогнозування стійкості на зсув статистично значущої різниці в точності між моделями не виявлено (таблиця 3,  $p=0.54$ ). Водночас Модель 2 демонструє кращу середню точність ( $MAE_{\gamma} = 0.080$  проти  $MAE_{\gamma} = 0.087$  у Моделі 1), що вказує на її кращу регресійну поведінку для цього критерію.

Візуально це підтверджується графіками точності прогнозу: для Моделі 1 (рис. 3) прогнози стійкості на перекидання практично збігаються з істинними значеннями, тоді як для Моделі 2 (рис. 4) спостерігається більший розкид. Криві навчання демонструють стабільну динаміку функції втрат та швидку збіжність протягом перших 15 епох.

Таблиця 1

**Оцінка точності регресійних прогнозів та ефективності бінарної класифікації для прогнозу стійкості монорейкового крана на перекидання**

Метод оцінювання	Модель 1	Модель 2
$MAE_{\beta}$	0.049	0.066
$RMSE_{\beta}$	0.0609	0.0794
$R^2_{\beta}$	0.916	0.856
$Precision_{\beta}$	0.955	0.89
$Recall_{\beta}$	0.955	0.982
$Specificity_{\beta}$	0.944	0.853
$F1_{\beta}$	0.955	0.933

Таблиця 2

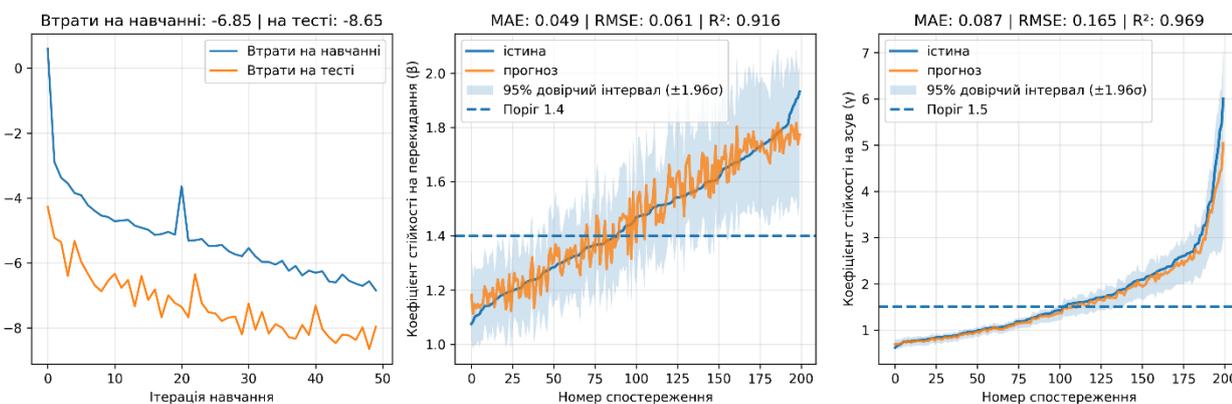
**Оцінка точності регресійних прогнозів та ефективності бінарної класифікації для прогнозу стійкості монорейкового крана на зсув**

Метод оцінювання	Модель 1	Модель 2
$MAE_{\gamma}$	0.087	0.08
$RMSE_{\gamma}$	0.165	0.191
$R^2_{\gamma}$	0.969	0.958
$Precision_{\gamma}$	1	0.993
$Recall_{\gamma}$	0.918	0.965
$Specificity_{\gamma}$	1	0.994
$F1_{\gamma}$	0.957	0.979

Таблиця 3

**Критерій Вілкоксона (Wilcoxon) для порівняння похибок прогнозу**

Нульова гіпотеза	T-статистика	p-value
$H_{0d}^{(\beta)}$	6512	$2.4e^{-10}$
$H_{0d}^{(\gamma)}$	6517	0.54



**Рис. 3. Багатоцільна рекурентна мережа GRU з механізмом самоуваги – криві навчання, та прогнозування стійкості монорейкового крана на перекидання та на зсув.**

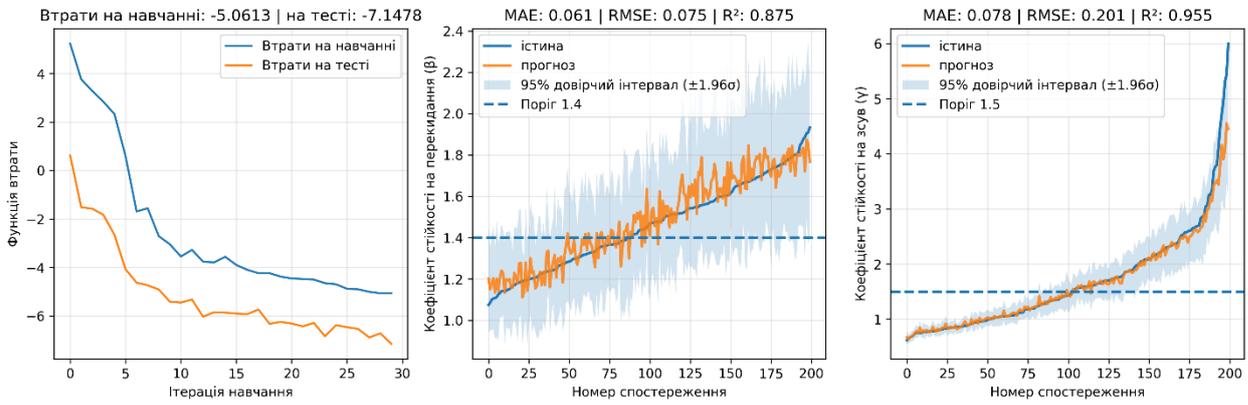


Рис 4. Багатоцільова рекурентна мережа GRU з механізмом уваги і байєсівським наближенням – криві навчання, та прогнозування стійкості монорейкового крана на перекидання та на зсув.

Аналіз матриць помилок (рис. 5 - 6) та метрик класифікації (таблиці 1 - 2) вказує на відмінності у стратегіях прийняття рішень моделей. Для завдання прогнозування стійкості крана на перекидання Модель 1 демонструє збалансовану продуктивність з високим  $F1_{\beta} = 0.955$ . Натомість Модель 2 реалізує більш ризик-орієнтовану стратегію, що підтверджується високою повнотою ( $Recall_{\beta}=98.2\%$ ). Модель 2 пропустила 4 небезпечних випадки ( $FN_{\beta}=4$ ) проти  $FN_{\beta}=5$  у Моделі 1 (рис. 5-6), що досягається ціною збільшення кількості хибних тривог ( $FP_{\beta} = 14$ ), які знижують  $Precision_{\beta}$  до 0.89. Для стійкості на зсув Модель 2 демонструє вищий показник повноти ( $Recall_{\gamma} = 96.5\%$  проти 91.8% для Моделі 1), пропустивши лише 3 небезпечних випадків у порівнянні з 8 у Моделі 1 (рис. 5-6). При цьому Модель 1 виявилася ідеальною у специфічності, не згенерувавши жодної хибної тривоги ( $FP_{\gamma}=0$ ).

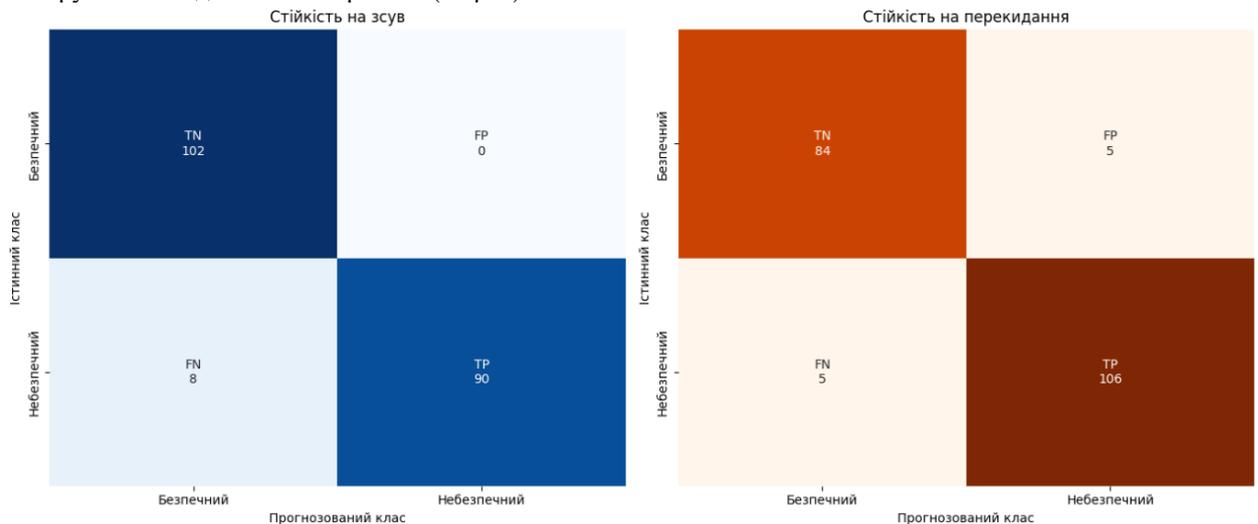


Рис 5. Багатоцільова рекурентна мережа GRU з механізмом самоуваги – матриця помилок

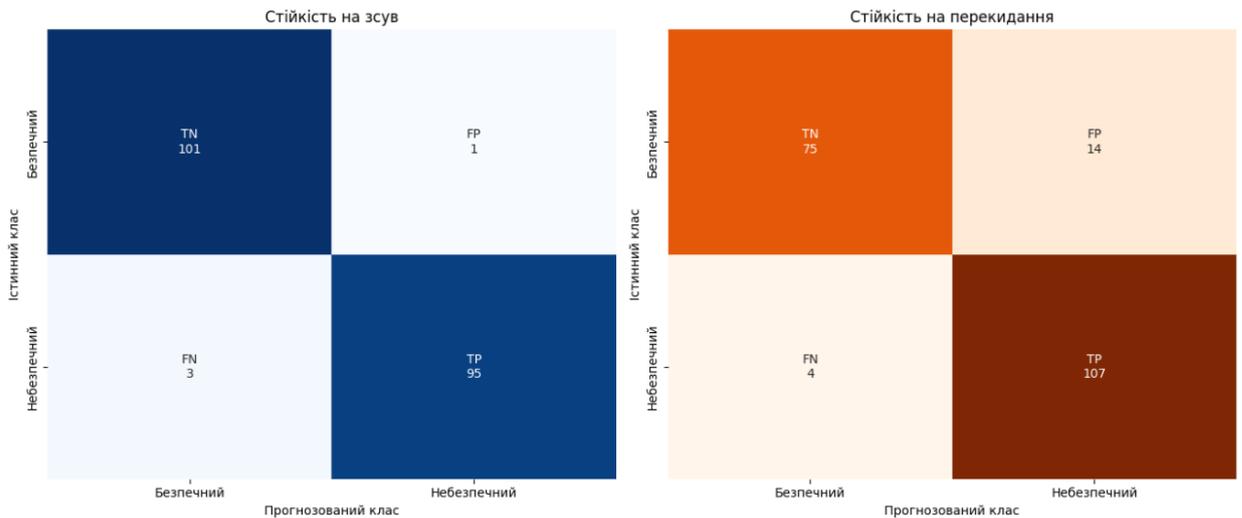


Рис 6. Багатоцільова рекурентна мережа GRU з механізмом уваги і байєсівським наближенням – матриця помилок

Аналіз оцінок невизначеності (таблиці 4–5) показує принципово різну поведінку моделей. Модель 2 є більш консервативною: середня ширина її 95% довірчого інтервалу  $CI_{95\%,\beta}$  є на 59.8% більшою для перекидання та на 21.4% більшою для зсуву. Це свідчить про те, що Модель 2 "визнає" більшу потенційну невизначеність у прогнозах, що є перевагою для систем безпеки. Модель 1 демонструє кращу калібровку невизначеності для задачі зсуву: коефіцієнт кореляції між помилкою та невизначеністю становить  $\rho_{|e|,\sigma}^{(\gamma)} = 0.93$ , що означає майже лінійну відповідність між величиною помилки та шириною інтервалу. Для Моделі 2 цей показник також є високим ( $\rho_{|e|,\sigma}^{(\gamma)} = 0.72$ , проте на 22.6% нижчим).

Таблиця 4

**Оцінки невизначеності прогнозу стійкості монорейкового крана на перекидання**

Метод оцінювання	Модель 1	Модель 2
$CI_{95\%,\beta}$	0.381	0.609
$\rho_{ e ,\sigma}^{(\beta)}$	0.227	0.26

Таблиця 5

**Оцінки невизначеності прогнозу стійкості монорейкового крана на зсув**

Метод оцінювання	Модель 1	Модель 2
$CI_{95\%,\gamma}$	0.529	0.642
$\rho_{ e ,\sigma}^{(\gamma)}$	0.93	0.72

Статистичні тести (таблиця 6) для перевірки значущості різниці поведінки моделей 1-2 щодо невизначеності прогнозів підтверджують, що різниця є статистично значущою ( $p\text{-value} < 0.001$ ) для завдання оцінювання стійкості крана на перекидання.

Таблиця 6

**Критерій Вілкоксона (Wilcoxon) для перевірки значущості різниці поведінки моделей щодо оцінювання невизначеності прогнозу**

Нульова гіпотеза	T-статистика	p-value
$H_{0\Delta}^{(\beta)}$	0	$1.4e^{-34}$
$H_{0\Delta}^{(\gamma)}$	2125	$4e^{-22}$

Таким чином, результати критерію Вілкоксона (таблиця 6) підтверджують, що моделі використовують різні підходи до кількісної оцінки невизначеності, причому Модель 2 є системно більш консервативною.

**Висновки з даного дослідження**

**і перспективи подальших розвідок у даному напрямі**

У дослідженні було розроблено, реалізовано та порівняно дві архітектури багатозадачних рекурентних нейронних мереж GRU для прогнозування стійкості монорейкового крана за критеріями перекидання та зсуву з урахуванням регресійної точності, якості бінарної класифікації та характеристик невизначеності прогнозу.

Обидві моделі продемонстрували високу прогностичну здатність ( $R^2 > 0.85$ ), що підтверджує адекватність апроксимації фізичної FEM-моделі засобами глибинного навчання. Водночас встановлено їхню функціональну спеціалізацію.

Для задачі прогнозування стійкості на перекидання багатозадачна рекурентна нейронна мережа GRU з механізмом самоуваги показала статистично значуще вищу регресійну точність ( $p < 0.001$ ), зменшивши середню абсолютну помилку на 25.8% порівняно з моделлю нейронної мережі GRU з механізмом уваги та байєсівським наближенням. Вона також забезпечує збалансовану якість класифікації ( $F1\beta = 0.955$ ), що свідчить про стабільність її рішень у задачах оцінювання ризику втрати стійкості.

Для задачі прогнозування стійкості на зсув статистично значущої різниці в регресійній точності не виявлено ( $p = 0.54$ ), однак модель нейронної мережі GRU з механізмом уваги та байєсівським наближенням демонструє кращу повноту ( $\text{Recall}\gamma = 96.5\%$ ) та вищий показник  $F1\gamma = 0.979$ , що свідчить про її чутливість до небезпечних станів.

Аналіз невизначеності прогнозів показав різні стратегії моделей. Модель нейронної мережі GRU з механізмом уваги та байєсівським наближенням є більш консервативною, формуючи ширші 95% довірчі інтервали (на 59.8% для перекидання та на 21.4% для зсуву), що підвищує рівень безпеки в умовах ризику. Натомість модель нейронної мережі GRU з механізмом самоуваги демонструє кращу калібровку невизначеності для задачі зсуву ( $\rho = 0.93$ ), що означає високу узгодженість між величиною помилки та шириною інтервалу. Статистичні тести підтвердили значущість відмінностей у поведінці моделей щодо оцінювання невизначеності ( $p < 0.001$ ).

Таким чином, результати дослідження доводять, що жодна з розроблених моделей не є абсолютно кращою за всіма критеріями. Це обґрунтовує головний висновок роботи: для створення надійної та ефективної інформаційної технології моніторингу безпеки крана необхідно використовувати гібридний підхід, що поєднує обидві розроблені архітектури багатозадачних рекурентних нейронних мереж GRU.

Перспективи подальших досліджень полягають у практичній реалізації цього гібридного підходу. Наступним кроком є розробка методу підтримки прийняття рішень, який би інтегрував результати одночасного прогнозу обох моделей в рамках єдиної інформаційної технології. Такий метод має використовувати точні та надійні прогнози моделі багатозадачної рекурентної нейронної мережі GRU з механізмом самоуваги для безперервного оперативного моніторингу, а у випадках, коли прогнози наближаються до критичних порогів, активувати модель багатозадачної рекурентної нейронної мережі GRU з механізмом уваги та байєсівським наближенням для отримання більш консервативної оцінки ризиків та глибшого аналізу невизначеності. Це дозволить створити дворівневу систему, що поєднує швидкість і точність з високою чутливістю до потенційних загроз.

### Література

1. Lung L. W. Leveraging Deep Learning and Internet of Things for Dynamic Construction Site Risk Management / L. W. Lung, Y. R. Wang, Y. S. Chen // *Buildings*. – 2025. – Vol. 15, № 8. – Art. 1325. DOI: <https://doi.org/10.3390/buildings15081325>.
2. Zhang S. Construction Site Safety Monitoring and Excavator Activity Analysis System / S. Zhang, L. Zhang // *Construction Robotics*. – 2022. – Vol. 6, № 2. – P. 151–161. DOI: <https://doi.org/10.1007/s41693-022-00077-0>.
3. Терентьев О. О. Рекуррентна нейронна мережа з механізмом уваги та байєсівським наближенням для прогнозування стійкості монорейкового крана на перекидання / О. О. Терентьев, Б. А. Соловей // *Управління розвитком складних систем*. – 2025. – № 64, С. 217–224. DOI: <https://doi.org/10.32347/2412-9933.2025.64.217-224>
4. Волянюк В. О. Розрахунок механізмів вантажопідіймальних машин : навч. посіб. / В. О. Волянюк, Є. В. Горбатюк. – Київ : КНУБА, 2021. – 136 с.
5. Han X. Multi-task Time Series Forecasting Based on Graph Neural Networks / X. Han, Y. Huang, Z. Pan, W. Li, Y. Hu, G. Lin // *Entropy*. – 2023. – Vol. 25, № 8. – Art. 1136. DOI: <https://doi.org/10.3390/e25081136>.
6. Karniadakis G. E. Physics-Informed Machine Learning / G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang // *Nature Reviews Physics*. – 2021. – Vol. 3, № 6. – P. 422–440. DOI: <https://doi.org/10.1038/s42254-021-00314-5>.
7. Zhou Z. Physics-Informed Neural Network for Load Sway Prediction in Travelling Autonomous Mobile Cranes / Z. Zhou, B. Johns, Y. Fang, Y. Bai, E. Abdi // *Advanced Engineering Informatics*. – 2025. – Vol. 65. – Art. 103269. – DOI: <https://doi.org/10.1016/j.aei.2025.103269>
8. Fernández J. Physics-Guided Recurrent Neural Network Trained with Approximate Bayesian Computation: A Case Study on Structural Response Prognostics / J. Fernández, J. Chiachío, J. Barros, M. Chiachío, C. S. Kulkarni // *Reliability Engineering & System Safety*. – 2024. – Vol. 243. – Art. 109822. DOI: <https://doi.org/10.1016/j.ress.2023.109822>.
9. Kirchdorfer L. Investigating Uncertainty Weighting for Multi-Task Learning: Insights and Analytical Alternative / L. Kirchdorfer, T. Sesterhenn, C. Bartelt, H. Stuckenschmidt, L. Schott, J. M. Köhler // *International Journal of Computer Vision*. – 2026. – Vol. 134, № 1. – Art. 8. DOI: <https://doi.org/10.1007/s11263-025-02625-x>.
10. Kilinc H. C. Multimodal Fusion of Optimized GRU–LSTM with Self-Attention Layer for Hydrological Time Series Forecasting / H. C. Kilinc, S. Apak, F. Ozkan, M. E. Ergin, A. Yurtsever // *Water Resources Management*. – 2024. – Vol. 38, № 15. – P. 6045–6062. DOI: <https://doi.org/10.1007/s11269-024-03943-4>.
11. Dash T. Residual GRU + MHSA: A Lightweight Hybrid Recurrent Attention Model for Cardiovascular Disease Detection / T. Dash, G. Datla, A. Vurity, T. Ahmad, M. Adnan, S. Rafi [та ін.] // *arXiv preprint*. – 2025. – arXiv:2512.14563.
12. Khoshvaght H. A Critical Review on Selecting Performance Evaluation Metrics for Supervised Machine Learning Models in Wastewater Quality Prediction / H. Khoshvaght, R. R. Permala, A. Razmjou, M. Khiadani // *Journal of Environmental Chemical Engineering*. – 2025. – Vol. 13, № 6. – Art. 119675. – DOI: <https://doi.org/10.1016/j.jece.2025.119675>.
13. Solovei O. Maximum Expected Differences in the Scores of MCC, Cohen's Kappa and Balanced Accuracy / O. Solovei, T. Honcharenko // *International Conference on Computer Science, Engineering and Education Applications*. – Cham : Springer Nature Switzerland, 2024. – P. 168–177. – DOI: [https://doi.org/10.1007/978-3-031-84228-3\\_14](https://doi.org/10.1007/978-3-031-84228-3_14).
14. Gawlikowski J. A Survey of Uncertainty in Deep Neural Networks / J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng [та ін.] // *Artificial Intelligence Review*. – 2023. – Vol. 56 (Suppl. 1). – P. 1513–1589. – DOI: <https://doi.org/10.1007/s10462-023-10562-9>.
15. Rainio O. Evaluation Metrics and Statistical Tests for Machine Learning / O. Rainio, J. Teuho, R. Klén // *Scientific Reports*. – 2024. – Vol. 14, № 1. – Art. 6086. – DOI: <https://doi.org/10.1038/s41598-024-56706-x>.

### References

1. Lung L. W. Leveraging Deep Learning and Internet of Things for Dynamic Construction Site Risk Management / L. W. Lung, Y. R. Wang, Y. S. Chen // *Buildings*. – 2025. – Vol. 15, № 8. – Art. 1325. DOI: <https://doi.org/10.3390/buildings15081325>.
2. Zhang S. Construction Site Safety Monitoring and Excavator Activity Analysis System / S. Zhang, L. Zhang // *Construction Robotics*. – 2022. – Vol. 6, № 2. – P. 151–161. DOI: <https://doi.org/10.1007/s41693-022-00077-0>.

3. Terentiev O. O. Rekurentna neironna merezha z mekhanizmom uvahy ta baiiesivskym nablyzhenniam dlia prohnozuvannia stiikosti monoreikovoho krana na perekydannia / O. O. Terentiev, B. A. Solovei // Upravlinnia rozvytkom skladnykh system. – 2025. – № 64, C. 217–224. DOI: <https://doi.org/10.32347/2412-9933.2025.64.217-224>.
4. Volianiuk V. O. Rozrakhunok mekhanizmv vantazhopidiimalnykh mashyn : navch. posib. / V. O. Volianiuk, Ye. V. Horbatiuk. – Kyiv : KNUBA, 2021. – 136 s.
5. Han X. Multi-task Time Series Forecasting Based on Graph Neural Networks / X. Han, Y. Huang, Z. Pan, W. Li, Y. Hu, G. Lin // Entropy. – 2023. – Vol. 25, № 8. – Art. 1136. DOI: <https://doi.org/10.3390/e25081136>.
6. Karniadakis G. E. Physics-Informed Machine Learning / G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang // Nature Reviews Physics. – 2021. – Vol. 3, № 6. – P. 422–440. DOI: <https://doi.org/10.1038/s42254-021-00314-5>.
7. Zhou Z. Physics-Informed Neural Network for Load Sway Prediction in Travelling Autonomous Mobile Cranes / Z. Zhou, B. Johns, Y. Fang, Y. Bai, E. Abdi // Advanced Engineering Informatics. – 2025. – Vol. 65. – Art. 103269. – DOI: <https://doi.org/10.1016/j.aei.2025.103269>
8. Fernández J. Physics-Guided Recurrent Neural Network Trained with Approximate Bayesian Computation: A Case Study on Structural Response Prognostics / J. Fernández, J. Chiachío, J. Barros, M. Chiachío, C. S. Kulkarni // Reliability Engineering & System Safety. – 2024. – Vol. 243. – Art. 109822. DOI: <https://doi.org/10.1016/j.ress.2023.109822>.
9. Kirchdorfer L. Investigating Uncertainty Weighting for Multi-Task Learning: Insights and Analytical Alternative / L. Kirchdorfer, T. Sesterhenn, C. Bartelt, H. Stuckenschmidt, L. Schott, J. M. Köhler // International Journal of Computer Vision. – 2026. – Vol. 134, № 1. – Art. 8. DOI: <https://doi.org/10.1007/s11263-025-02625-x>.
10. Kilinc H. C. Multimodal Fusion of Optimized GRU–LSTM with Self-Attention Layer for Hydrological Time Series Forecasting / H. C. Kilinc, S. Apak, F. Ozkan, M. E. Ergin, A. Yurtsever // Water Resources Management. – 2024. – Vol. 38, № 15. – P. 6045–6062. DOI: <https://doi.org/10.1007/s11269-024-03943-4>.
11. Dash T. Residual GRU + MHSA: A Lightweight Hybrid Recurrent Attention Model for Cardiovascular Disease Detection / T. Dash, G. Datla, A. Vurity, T. Ahmad, M. Adnan, S. Rafi [та ін.] // arXiv preprint. – 2025. – arXiv:2512.14563.
12. Khoshvaght H. A Critical Review on Selecting Performance Evaluation Metrics for Supervised Machine Learning Models in Wastewater Quality Prediction / H. Khoshvaght, R. R. Permala, A. Razmjou, M. Khiadani // Journal of Environmental Chemical Engineering. – 2025. – Vol. 13, № 6. – Art. 119675. – DOI: <https://doi.org/10.1016/j.jece.2025.119675>.
13. Solovei O. Maximum Expected Differences in the Scores of MCC, Cohen’s Kappa and Balanced Accuracy / O. Solovei, T. Honcharenko // International Conference on Computer Science, Engineering and Education Applications. – Cham : Springer Nature Switzerland, 2024. – P. 168–177. – DOI: [https://doi.org/10.1007/978-3-031-84228-3\\_14](https://doi.org/10.1007/978-3-031-84228-3_14).
14. Gawlikowski J. A Survey of Uncertainty in Deep Neural Networks / J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng [та ін.] // Artificial Intelligence Review. – 2023. – Vol. 56 (Suppl. 1). – P. 1513–1589. – DOI: <https://doi.org/10.1007/s10462-023-10562-9>.
15. Rainio O. Evaluation Metrics and Statistical Tests for Machine Learning / O. Rainio, J. Teuvo, R. Klén // Scientific Reports. – 2024. – Vol. 14, № 1. – Art. 6086. – DOI: <https://doi.org/10.1038/s41598-024-56706-x>.