

**МОДЕЛІ ВІДСТЕЖЕННЯ ЗНАНЬ
В АДАПТИВНИХ НАВЧАЛЬНИХ СИСТЕМАХ:
ПОРІВНЯННЯ ВКТ І САКТ**

Бондаренко А.А.

e-mail: anna.bondarenko1@nure.ua

Науковий керівник – к.т.н., доц. Лещинський В.О.

Харківський національний університет радіоелектроніки, каф. ПІ
м. Харків, Україна

This study compares Self-Attentive Knowledge Tracing (SAKT) and Bayesian Knowledge Tracing (BKT) within adaptive learning systems. Using the ASSISTments 2017 dataset, a manually parameterised BKT baseline reached a Test AUC of 0.5715, while SAKT achieved a mean Test AUC of 0.669 ± 0.001 over three runs. The resulting 0.097 AUC gap suggests attention-based models better capture sequential learner behaviour. Results support integrating transformer-based components in educational platforms. Future work should include comprehensive hyperparameter optimisation, significance testing, and cross-dataset validation.

Сучасні інтелектуальні навчальні системи потребують надійних моделей для відстеження динаміки знань учня та прогнозування його результатів з метою побудови адаптивних траєкторій навчання. Формалізовані підходи до моделювання знань у рекомендаційних системах забезпечують прозорість формування рекомендацій з урахуванням характеристик об'єктів і вимог користувача [1].

Проблема відстеження знань (knowledge tracing, KT) полягає у прогнозуванні прихованого стану знань на основі послідовності взаємодій учня з навчальними завданнями. До класичних моделей KT належать імовірнісні підходи, зокрема Bayesian Knowledge Tracing (BKT). Натомість в останнє десятиліття з'явилися нейромережеві підходи (Deep Knowledge Tracing), а згодом – архітектури на основі механізму уваги (Self-Attentive Knowledge Tracing, SAKT), що ефективніше враховують складні послідовні залежності між взаємодіями здобувача освіти. Мета цієї роботи полягає в емпіричному порівнянні ВКТ та САКТ на репрезентативному датасеті, аналізі відмінностей у якості прогнозування та наданні рекомендацій для практичного застосування.

Використано публічний набір даних ASSISTments (Data Mining Competition 2017) [2] після стандартного препроцесингу: видалено артефакти, сесії впорядковано хронологічно. Для САКТ послідовності нарізано у вікна фіксованої довжини ($L=50$), де кожна послідовність містить вхідні пари (ідентифікатор завдання, правильність відповіді), а метою є прогноз правильності наступної відповіді. Для ВКТ використано формат: (user, skill, correct), що відповідає традиційному застосуванню моделі. Усі перетворення збережено для забезпечення відтворюваності.

ВКТ – класична ймовірнісна модель на основі прихованих ланцюгів Маркова, що відстежує бінарний стан знання навички $L_t \in \{0,1\}$ (не засвоєно, засвоєно). Модель оперує чотирма параметрами: початкова ймовірність знання $P(L_0)$, ймовірність засвоєння після завдання $P(T)$, ймовірність помилки при наявності знання $P(S)$ та ймовірність вгадування $P(G)$ [3]. Ймовірність засвоєння навички на даному кроці (t), оновлюється за формулою Байєса, а перехід до наступного кроку ($t + 1$) – за рівнянням:

$$P(L_{t+1}) = P((L_t | obs_t) + 1 - P(L_t | obs_t)) \cdot P(T).$$

У дослідженні було використано ручні початкові параметри:

$$P(L_0) = 0.2, P(T) = 0.3, P(S) = 0.1, P(G) = 0.2.$$

SAKT базується на механізмі self-attention трансформерів і паралельно обробляє всю послідовність взаємодій, обчислюючи ваги уваги між усіма токенами (питання, відповідь):

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) \cdot V,$$

де Q проєктує поточне питання, а K та V є проєкціями минулих взаємодій [4]. Модель реалізовано в середовищі для глибинного навчання PyTorch на основі відкритого коду з репозиторію [5] із такими налаштуваннями: розмірність векторного представлення – 128, кількість голів уваги – 4, коефіцієнт регуляції випадкового вимнення (dropout) – 0.2, розмір навчального пакету – 64, швидкість навчання – 0.001, кількість епох – 10. Стабільність результатів оцінено за трьома незалежними запусками.

Основною метрикою якості обрано площу під ROC кривою (AUC), оскільки вона інваріантна до порогу класифікації. В освітньому дата-майнінгу AUC 0.6-0.7 вважається базовим рівнем, тоді як ≥ 0.8 вказує на високоточну модель.

Таблиця 1 – Результати порівняння моделей

Модель	Налаштування	Test AUC
ВКТ (baseline)	$P(L_0) = 0.2,$ $P(T) = 0.3$	0.5716
SAKT (best run)	emb=128, heads=4	0.668–0.669
SAKT (mean, 3 runs)	emb=128, heads=4	0.669±0.001

Різниця за показником AUC (≈ 0.097) зумовлена архітектурними відмінностями, адже SAKT використовує повну історію взаємодій та добирає релевантний попередній контекст для кожного прогнозу, тоді як ВКТ моделює еволюцію знання навички локально й обмежується фіксованим набором параметрів, що спрощує модель і знижує її виражальну спроможність у присутності перехресних залежностей між завданнями. Як показано на рисунку 1, крива SAKT розташована вище

кривої ВКТ вздовж усього відрізка порогів, що наочно підтверджує перевагу attention-based архітектури.

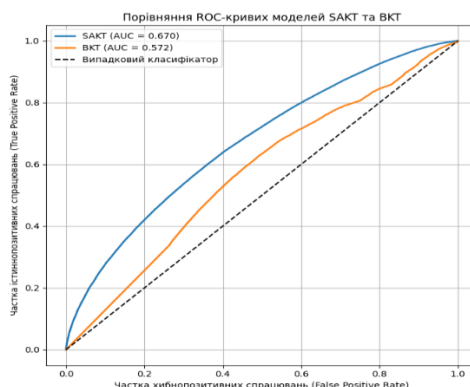


Рисунок 1 – Порівняння ROC-кривих моделей SAKT та ВКТ

Серед обмежень дослідження слід зазначити, що для ВКТ не було виконано повної статистичної підгонки параметрів, а SAKT не пройшла систематичного пошуку гіперпараметрів. У подальших дослідженнях слід провести крос-датасетну валідацію, повну оптимізацію гіперпараметрів та оцінку статистичної значущості результатів.

Проведене емпіричне порівняння показує, що SAKT забезпечує суттєве покращення у задачах КТ порівняно з ВКТ на датасеті ASSISTments. Це підтверджує доцільність використання трансформерних архітектури при побудові інтелектуальних компонентів адаптивних освітніх платформ.

Список використаних джерел:

1. Chalyi S., Leshchynskiy V. Knowledge representation in the recommendation system based on the white box principle. Advanced information systems. 2019. T. 3, № 3. С. 82–89. DOI: <https://doi.org/10.20998/2522-9052.2019.3.12>
2. The 2017 assistments datamining competition – dataset. URL: <https://sites.google.com/view/assistmentsdatamining/dataset> (дата звернення: 05.03.2026).
3. Hawkins W. J., Heffernan N. T., Baker R. S. J. D. Learning bayesian knowledge tracing parameters with a knowledge heuristic and empirical probabilities. Intelligent tutoring systems. Cham, 2014. С. 150–155. DOI: https://doi.org/10.1007/978-3-319-07221-0_18
4. Ghosh A., Heffernan N., Lan A. S. Context-Aware attentive knowledge tracing. KDD '20: the 26th ACM SIGKDD conference on knowledge discovery and data mining, м. Virtual Event CA USA. New York, NY, USA, 2020. DOI: <https://doi.org/10.1145/3394486.3403282>
5. GitHub - arshadshk/SAKT-pytorch: Implementation of paper "A Self-Attentive model for Knowledge Tracing". GitHub. URL: <https://github.com/arshadshk/SAKT-pytorch> (дата звернення: 05.03.2026).