

Научная статья | Original paper

УДК 004.8.032.26:004.85:004.62

Моделирование и прогнозирование цифрового когнитивного выгорания у студентов на основе мультимодальных временных рядов и кросс-модальной архитектуры трансформера

Н.Е. Юрьева

Московский государственный психолого-педагогический университет
Москва, Российская Федерация

✉ yurieva.ne@gmail.com

Резюме

В статье рассматривается задача прогнозирования цифрового когнитивного выгорания (ЦКВ) у студентов — динамического психологического состояния, формируемого в цифровой образовательной среде. Задача формализуется как проблема многомерной регрессии на мультимодальных временных рядах, целью которой является предсказание траектории изменений показателей по шкале MBI-SS на заданном временном горизонте. Автором предложена математическая модель на основе иерархической кросс-модальной архитектуры трансформера¹, включающая специализированные модальные кодировщики, механизм кросс-модального внимания и рекуррентный прогнозный декодер. Для совместного решения задач регрессии (прогнозирования) и классификации (определения текущего уровня выгорания) введён составной функционал потерь, который учитывает не только точность прогноза, но и сглаженность его временной динамики, а также корректность классификации. Экспериментальная проверка модели на синтетических данных продемонстрировала снижение средней абсолютной ошибки прогноза по сравнению с базовыми подходами и подтвердила интерпретируемость межмодальных взаимодействий на основе анализа матриц внимания. Полученные результаты формируют формально-вычислительную основу для разработки систем раннего выявления риска цифрового когнитивного выгорания в образовательных платформах.

Ключевые слова: цифровое когнитивное выгорание, прогнозное моделирование, мультимодальные временные ряды, кросс-модальный трансформер, интеграция данных, математическая модель

¹ От англ. *transformer* — архитектура модели внимания



Для цитирования: Юрьева, Н.Е. (2026). Моделирование и прогнозирование цифрового когнитивного выгорания у студентов на основе мультимодальных временных рядов и кросс-модальной архитектуры трансформера. *Моделирование и анализ данных*, 16(3), 91—103. <https://doi.org/10.17759/mda.2026160304>

Modeling and forecasting digital cognitive burnout in students based on multimodal time series and a cross-modal transformer architecture

N.E. Yuryeva

Moscow State University of Psychology & Education, Moscow, Russian Federation

✉ yuryeva.ne@gmail.com

Abstract

The article discusses the problem of predicting digital cognitive burnout (CKD) in students, a dynamic psychological state formed in a digital educational environment. The problem is formalized as a multidimensional regression problem on multimodal time series, the purpose of which is to predict the trajectory of changes in indicators on the MBI-SS scale over a given time horizon. The author proposes a mathematical model based on a hierarchical cross-modal transformer architecture, including specialized modal encoders, a cross-modal attention mechanism, and a recurrent predictive decoder. To jointly solve the problems of regression (forecasting) and classification (determining the current level of burnout), a composite loss functional has been introduced that takes into account not only the accuracy of the forecast, but also the smoothness of its temporal dynamics, as well as the correctness of the classification. Experimental verification of the model on synthetic data demonstrated a decrease in the average absolute prediction error compared to the basic approaches and confirmed the interpretability of intermodal interactions based on the analysis of attention matrices. The results obtained form the formal computational basis for the development of early detection systems for the risk of digital cognitive burnout in educational platforms.

Keywords: digital cognitive burnout, predictive modeling, multimodal time series, cross-modal transformer, data fusion, mathematical model

For citation: Yuryeva, N.E. (2026). Modeling and forecasting digital cognitive burnout in students based on multimodal time series and a cross-modal transformer architecture. *Modelling and Data Analysis*, 16(3), 91—103. (In Russ.). <https://doi.org/10.17759/mda.2026160304>



Введение

Задачи прогнозирования динамики учебных и когнитивных показателей обучающихся в онлайн-средах в последние годы активно развиваются в рамках прикладных исследований образовательной аналитики. Однако большинство существующих работ ориентировано на предсказание академической успеваемости и агрегированных поведенческих метрик, а не на моделирование психологических состояний как непрерывных динамических процессов (Wang & Yu, 2025).

Цифровизация образования привела к возникновению нового психологического феномена — цифрового когнитивного выгорания (ЦКВ). Оно определяется как устойчивое сочетание истощаемости, дистанцирования и снижения эффективности в цифровых учебных средах (Soldatova, 2018). В отличие от кратковременной усталости, ЦКВ является хроническим состоянием, требующим ранней диагностики и профилактики (Kuppens & Verduyn, 2017; Юрьева, 2025).

Работа (Юрьева, 2025) была посвящена задаче классификации дискретных когнитивных состояний. Однако такой подход не позволяет анализировать динамику изменения психометрических показателей и выявлять ранние субклинические тренды, что существенно ограничивает возможности профилактики цифрового когнитивного выгорания.

Данное исследование переходит к более сложной задаче прогнозирования непрерывной многомерной траектории психометрического показателя (MBI-SS) на заданном временном горизонте. Фактически, это задача регрессионного анализа мультимодальных, неоднородных и зашумленных временных рядов.

Пусть $y_t \in \mathbb{R}^3$ — вектор баллов по субшкалам MBI-SS (истощение, цинизм, эффективность) в момент времени t .

Дано мультимодальное временное окно истории $\mathcal{X}_{[t-L:t]} = \{X^{(A)}, X^{(V)}, X^{(T)}, X^{(K)}\}_{t-L}^t$, где индексы A, V, T, K обозначают аудио, видео, текст и кинестетику соответственно.

Требуется построить модель $\mathcal{F}_\Theta$, аппроксимирующую условное распределение:

$$P(y_{t+1}, \dots, y_{t+T} \mid \mathcal{X}_{[t-L:t]}; \Theta),$$

где T — горизонт прогнозирования, Θ — параметры модели.

Цель работы: разработка и математическое обоснование архитектуры $\mathcal{F}_\Theta$, обеспечивающей минимизацию ошибки прогноза при сохранении интерпретируемости и устойчивости к пропускам данных.

Формальная постановка задачи и функционал потерь

Математическая формализация

Рассмотрим множество студентов $\mathcal{U} = \{u_1, \dots, u_N\}$. Для каждого студента $u \in \mathcal{U}$ в момент времени t определены:

1. Мультимодальные наблюдаемые переменные: $\mathbf{X}_u^{(m)}(t) \in \mathbb{R}^{d_m \times s_m}$, где $m \in \{A, V, T, K\}$ — тип модальности, d_m — размерность признаков, s_m — длина временного ряда в окне (может различаться для разных модальностей).



2. Целевая переменная: $\mathbf{y}_u(t) \in \mathbb{R}^3$ — вектор баллов по трем субшкалам MBI-SS.

Задача 1 (прогнозирование). По заданной истории наблюдений $\mathcal{H}_u(t) = \{\mathbf{X}_u^{(m)}(\tau), \mathbf{y}_u(\tau)\}_{\tau=t-L}^t$ построить оценку:

$$\hat{\mathbf{y}}_u(t+1:t+T) = \mathcal{F}_\Theta(\mathcal{H}_u(t)),$$

минимизирующую функционал потерь $\mathcal{L}_{pred}$.

Задача 2 (классификация текущего состояния). На основе $\mathcal{H}_u(t)$ определить класс текущего уровня выгорания $c_u(t) \in \{\text{низкий, средний, высокий}\}$.

Следует отметить, что предполагаемая стационарность статистических свойств временных рядов на интервале наблюдения является приближением, допустимым лишь для относительно коротких временных окон, что накладывает ограничения на интерпретацию долгосрочных прогнозов.

Составной функционал потерь

Для совместного обучения модели регрессии и классификации вводится составной функционал:

$$\mathcal{L}(\Theta) = \lambda_1 \mathcal{L}_{MSE} + \lambda_2 \mathcal{L}_{Smooth} + \lambda_3 \mathcal{L}_{Clf} + \lambda_4 \mathcal{L}_{Reg}.$$

1. Функция потерь прогноза (MSE):

$$\mathcal{L}_{MSE} = \frac{1}{N \cdot T \cdot 3} \sum_{u=1}^N \sum_{\tau=1}^T \sum_{s=1}^3 (\hat{y}_{u,s}(t+\tau) - y_{u,s}(t+\tau))^2.$$

Функция сглаживания траектории: штрафует резкие, немотивированные изменения прогноза между соседними временными шагами, что соответствует естественной инерционности психологического состояния (Выготский, 1982).

$$\mathcal{L}_{Smooth} = \frac{1}{N \cdot (T-1) \cdot 3} \sum_{u=1}^N \sum_{\tau=1}^{T-1} \sum_{s=1}^3 (\hat{y}_{u,s}(t+\tau+1) - \hat{y}_{u,s}(t+\tau))^2.$$

Использование штрафа на резкие изменения согласуется с представлениями о динамике психических состояний как инерционных процессов (Выготский, 1982).

2. Функция потерь классификации:

$$\mathcal{L}_{Clf} = -\frac{1}{N} \sum_{u=1}^N \sum_{c=1}^C [\mathbb{I}[c_u(t) = c] \cdot \log p_\Theta(c | \mathcal{H}_u(t))],$$

где $C = 3$, p_Θ — вероятность класса, предсказанная моделью.

3. L2-регуляризация параметров:



$$\mathcal{L}_{Reg} = \|\Theta\|_2^2.$$

Коэффициенты $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ подбираются с помощью метода полного перебора на валидационной выборке.

Математическая модель архитектуры

Модальные кодировщики: отображение сырых данных в пространство признаков

Использование мультимодальных данных для анализа и прогнозирования психических и аффективных состояний опирается на результаты исследований, демонстрирующих, что сочетание поведенческих, речевых и физиологических сигналов позволяет более полно описывать динамику внутренних состояний по сравнению с одномодальными подходами (D'Mello & Kory, 2015).

Каждая модальность m обрабатывается специализированным кодировщиком $E_m: \mathbb{R}^{d_m \times s_m} \rightarrow \mathbb{R}^{d_{emb} \times s'_m}$, который преобразует исходные данные в последовательность скрытых представлений фиксированной размерности d_{emb} . В общем виде:

$$\mathbf{Z}^{(m)} = \text{Encoder}_m(\mathbf{X}^{(m)}; \Theta_m)$$

Например, для текстовой модальности может использоваться трансформер BERT: $\mathbf{Z}^{(T)} = \text{BERT}(\text{Tokenize}(X^{(T)}))$. Для аудио — сверточная сеть (1D-CNN) поверх мел-спектрограмм: $\mathbf{Z}^{(A)} = \text{CNN}_{1D}(\text{MelSpectrogram}(X^{(A)}))$.

Кросс-модальный модуль интеграции данных на основе механизма внимания

Выбор кросс-модального подхода к объединению разнородных источников данных соответствует современным представлениям мультимодального машинного обучения, в рамках которых особое внимание уделяется согласованию представлений различных модальностей и явному моделированию их взаимодействий (Baltrušaitis et al., 2019).

Полученные векторные представления $\mathbf{Z}^{(A)}, \mathbf{Z}^{(V)}, \mathbf{Z}^{(T)}, \mathbf{Z}^{(K)}$ конкатенируются по временной оси с добавлением модальных позиционных кодировок $\mathbf{P}_m$:

$$\mathbf{Z}^{concat} = [\mathbf{Z}^{(A)} + \mathbf{P}_A; \mathbf{Z}^{(V)} + \mathbf{P}_V; \mathbf{Z}^{(T)} + \mathbf{P}_T; \mathbf{Z}^{(K)} + \mathbf{P}_K] \in \mathbb{R}^{d_{emb} \times S},$$

где $[\cdot]$ — конкатенация по временной оси, $S = s'_A + s'_V + s'_T + s'_K$.

Затем $\mathbf{Z}^{concat}$ подается на вход стека из L_{tr} кросс-модальных блоков трансформера. В каждом блоке выполняется операция многоголового кросс-модального внимания (Cross-Modal Multi-Head Attention, СММНА). Механизм внимания, лежащий в основе трансформерных архитектур, позволяет моделировать зависимости между элементами последовательностей без явного задания локальной структуры, что делает его удобным инструментом для анализа сложных временных и кросс-модальных зависимостей (Vaswani et al., 2017).



Для головы h :

$$\text{Attention}^{(h)}(\mathbf{Q}^{(m_i)}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}^{(m_i)} \mathbf{K}^\top}{\sqrt{d_k}} + \mathbf{M}\right) \mathbf{V}.$$

Здесь:

- $\mathbf{Q}^{(m_i)} = \mathbf{Z}^{(m_i)} \mathbf{W}_Q^{(h)}$ — запрос, получаемый из токенов модальности m_i .
- $\mathbf{K}, \mathbf{V} = \mathbf{Z}^{concat} \mathbf{W}_K^{(h)}, \mathbf{Z}^{concat} \mathbf{W}_V^{(h)}$ — ключи и значения из всех модальностей.
- $\mathbf{M}$ — маска, позволяющая, например, текстовым токенам «видеть» только аудио и видео за тот же временной интервал, что эмпирически снижает количество некоррелированных межмодальных взаимодействий и упрощает интерпретацию матриц внимания.

Результаты всех голов конкатенируются и проходят через линейное преобразование:

$$\text{CMMHA}(\mathbf{Z}^{(m_i)}, \mathbf{Z}^{concat}) = \text{Concat}(\text{head}_1, \dots, \text{head}_H) \mathbf{W}_O.$$

Это позволяет каждой модальности «запрашивать» релевантную информацию из других модальностей.

Альтернативой кросс-модальному вниманию могли бы выступать методы тензорной интеграции или графовые модели взаимодействия модальностей, однако они либо существенно увеличивают размерность параметрического пространства, либо затрудняют интерпретацию межмодальных связей.

Прогнозный декодер и классификатор

Скрытое состояние после трансформер-блоков $\mathbf{H} \in \mathbb{R}^{d_{emb} \times S}$ агрегируется (например, глобальным усреднением) в вектор контекста $\mathbf{c} \in \mathbb{R}^{d_{emb}}$.

Прогнозный декодер представляет собой рекуррентную сеть (GRU/LSTM), которая по вектору контекста $\mathbf{c}$ и предыдущему прогнозу генерирует последовательность на горизонте T :

$$\begin{aligned} \mathbf{h}_0 &= \mathbf{c}, \mathbf{h}_\tau = \text{DecoderRNN}(\mathbf{h}_{\tau-1}, \hat{\mathbf{y}}_{\tau-1}; \Theta_{dec}), \\ \hat{\mathbf{y}}_\tau &= \mathbf{W}_{out} \mathbf{h}_\tau + \mathbf{b}_{out}, \tau = 1, \dots, T. \end{aligned}$$

Классификатор — линейный слой с функцией активации *Softmax* над вектором $\mathbf{c}$:

$$p(c|\mathcal{H}) = \text{Softmax}(\mathbf{W}_{clf} \mathbf{c} + \mathbf{b}_{clf}).$$



Алгоритм обучения и оптимизация

Алгоритм 1. Обучение кросс-модальной прогнозной модели.

1. Вход: мультимодальный набор данных $\mathcal{D} = \{\mathcal{H}_u(t), \mathbf{y}_u(t+1:t+T), c_u(t)\}$, гипер-параметры $\lambda_{1,4}, L, T, d_{emb}, L_{tr}$.
2. Инициализация: параметры $\Theta = \{\Theta_m, \Theta_{tr}, \Theta_{dec}, \Theta_{clf}\}$ инициализируются согласно He et al. (2015).
3. Для каждой эпохи:
 1. Разбить $\mathcal{D}$ на мини-пакеты $\mathcal{B}$.
 2. Для каждого мини-пакета $\mathcal{B}_j$:
 - Прямой проход: вычислить $\hat{\mathbf{y}}, \hat{c} = \mathcal{F}_\Theta(\mathcal{H})$.
 - Вычислить составную функцию потерь $\mathcal{L}(\Theta)$.
 - Обратный проход: вычислить градиенты $\nabla_\Theta \mathcal{L}$.
 - Шаг оптимизации: обновить параметры с помощью AdamW: $\Theta \leftarrow \Theta - \eta \text{AdamW}(\nabla_\Theta \mathcal{L}, \Theta)$, где η — скорость обучения.
4. Ранняя остановка: если значение функции потерь $\mathcal{L}_{val}$ не улучшается P эпох.
5. Выход: обученные параметры Θ^* .

Использование оптимизатора AdamW с отдельным затуханием весов обусловлено его более стабильными обобщающими свойствами при обучении глубоких нейронных моделей по сравнению со стандартной версией Adam (Loshchilov & Hutter, 2019).

Экспериментальное моделирование и результаты

Синтетические данные и базовые модели

Для валидации математической корректности модели на первом этапе использован синтетический набор данных, где целевая переменная $\mathbf{y}_i$ генерируется линейно-динамической системой, возмущаемой скрытыми факторами, зависящими от смоделированных мультимодальных признаков. Использование синтетических данных на данном этапе обусловлено отсутствием открытых лонгитюдных мультимодальных наборов данных, включающих валидированные психометрические шкалы выгорания.

Для оценки эффективности предложенной архитектуры ее сравнивали с тремя классическими подходами (базовыми моделями):

- BL1: линейная регрессионная модель с квадратичной регуляризацией (Ridge Regression), обученная на объединенном наборе признаков. Используется как простая интерпретируемая модель сравнения.
- BL2: рекуррентная нейронная сеть с долговременной памятью (типа LSTM), обученная на объединенных временных рядах признаков, с учетом их последовательной структуры.
- BL3: ансамблевая модель с отдельной обработкой модальностей и последующим объединением результатов (Late Fusion). Комплексная модель, где для каждого



типа данных (аудио, видео и т.д.) используется своя LSTM, а результаты их работы затем объединяются общим классификатором.

Метрики оценки

1. Для задач прогнозирования использовались следующие показатели: *средняя абсолютная ошибка* (MAE), *среднеквадратичная ошибка* (RMSE) и *коэффициент детерминации* (R^2).
2. Качество классификации оценивалось с использованием:
 - взвешенной F1-меры (weighted F1-score), учитывающей распределение классов;
 - макроусреднённой площади под ROC-кривой (macro-averaged ROC-AUC).

Для оценки интерпретируемости моделей применялась *норма Фробениуса матриц внимания* $\|A_{m_i, m_j}\|_F$ между различными модальностями, рассматриваемая как количественная мера их взаимного влияния.

Предыдущие исследования показывают, что анализ синхронизированных мультимодальных потоков данных позволяет выявлять скрытые закономерности, связанные с процессами обучения и вовлечённости, недоступные при анализе отдельных каналов информации (Grafsgaard et al., 2014).

Результаты и анализ

Таблица 1/ Table 1

Сравнение моделей по метрике MAE (прогноз) и F1 (классификация)

Comparison of models using the MAE (prediction) and F1 (classification) metrics

Модель / Model	MAE (Истощение/ Exhaustion)	MAE (Цинизм/ Cynicism)	MAE (Эффект./ Effect.)	MAE (Средн./ Average)	F1- score
BL1: Ridge	3.21	2.98	2.75	2.98	0.65
BL2: LSTM	2.45	2.67	2.30	2.47	0.72
BL3: Late Fusion	2.20	2.31	2.05	2.19	0.78
Предложенная модель / The proposed model	1.85	1.92	1.73	1.83	0.84

Анализ: предложенная модель показывает снижение MAE на ~16% относительно лучшей базовой модели (BL3). Функция $\mathcal{L}_{Smooth}$ снизила дисперсию ошибки прогноза на последовательных шагах на 30%.

Анализ матриц внимания: вычислены средние матрицы внимания $A = \mathbb{E}[A_{m_i, m_j}]$ по валидационной выборке. Обнаружено, что:

Аудио-токены проявляют высокое внимание к видео-токенам ($A_{A \rightarrow V} = 0.31$), что соответствует интеграции голосового стресса и мимики.

Текстовые токены «циничных» высказываний сильно взаимодействуют с кинестетическими токенами замедления печати ($A_{T \rightarrow K} = 0.28$), формируя паттерн поведенческой демотивации.



Следует отметить, что интерпретация матриц внимания и прямое трактование весов как свидетельства причинно-следственных связей не получила однозначного подтверждения и продолжает обсуждаться в научной литературе (Jain & Wallace, 2019).

Обсуждение, ограничения и заключение

Результаты исследований аффективных состояний в образовательных контекстах показывают, что эмоциональные и когнитивные компоненты обучения тесно взаимосвязаны и проявляются в динамике поведения обучающихся, что делает задачу их формального моделирования особенно сложной (Reis et al., 2018). Результаты прогнозирования цифрового когнитивного выгорания целесообразно рассматривать в более широком контексте регуляции психических состояний в учебной деятельности, где когнитивные, эмоциональные и поведенческие компоненты образуют взаимосвязанную динамическую систему. Современные исследования показывают, что устойчивые изменения в структуре учебной активности и саморегуляции обучающихся формируются постепенно и опосредуются как индивидуальными, так и средовыми факторами, что накладывает дополнительные ограничения на интерпретацию краткосрочных прогнозных моделей (Карташева, 2022).

Разработанная математическая модель формализует задачу прогнозирования ЦКВ как задачу условной регрессии на гетерогенных временных рядах. Введение составного функционала потерь $\mathcal{L}$ и механизма кросс-модального внимания позволило учесть временную динамику, согласованность модальностей и дискретную классификацию одновременно.

Полученные результаты не следует интерпретировать как готовую диагностическую методику, а лишь как формальный и вычислительный каркас для последующей клинической и психометрической валидации. Подобное ограничение интерпретации соответствует современной методологической позиции в психологии, согласно которой прогнозные модели временной динамики психических состояний не могут рассматриваться как прямые инструменты каузального вывода или индивидуальной диагностики без дополнительной теоретической и эмпирической валидации (Namaker et al., 2015; Карпов, 2021).

Основные ограничения

Вычислительная сложность модели $\mathcal{F}_\Theta$ составляет $O(L_{tr} \cdot S^2 \cdot d_{emb})$, что требует оптимизации для реального времени.

Ограниченная доступность больших размеченных образовательных наборов данных является общей проблемой исследований в области образовательной аналитики и интеллектуальных обучающих систем, что существенно влияет на устойчивость и воспроизводимость результатов моделей машинного обучения (Baker, 2019).

Дополнительным ограничением практического применения прогнозных моделей в образовательных средах являются этические и нормативные аспекты использования



методов искусственного интеллекта, включая вопросы интерпретации результатов, ответственности за принимаемые решения и защиты персональных данных обучающихся, что подчеркивается в современных обзорных работах по данной проблематике (Савгильдина, 2024).

Направления будущих исследований

Применение байесовских методов для количественной оценки неопределённости прогноза, что позволит получать не только точечные оценки траектории выгорания, но и доверительные интервалы $\sigma(\hat{y}_t)$.

Отдельным направлением является интеграция вариационных автоэнкодеров (VAE) в модальные кодировщики, что может повысить устойчивость моделей при работе с зашумлёнными и неполными данными.

Заключение

Разработанная в данной работе математическая модель служит теоретическим фундаментом и практическим инструментом для построения систем раннего обнаружения риска цифрового когнитивного выгорания (ЦКВ).

Формализация задачи как прогноза многомерных траекторий на основе гетерогенных временных рядов, а также предложенная кросс-модальная архитектура позволяют не только выявлять характерные изменения в динамике когнитивных и поведенческих показателей, но и рассматривать их во взаимосвязи, что принципиально важно для своевременной диагностики рисков снижения учебной и профессиональной эффективности.

Практическая значимость модели заключается в возможности её использования в качестве аналитического ядра интеллектуальных образовательных сред, ориентированных на поддержку обучающихся и профилактику перегрузки в условиях цифрового обучения. При этом модель рассматривается не как инструмент формального контроля, а как средство повышения адаптивности образовательных систем и более бережного отношения к когнитивным ресурсам человека.

Полученные на синтетических данных результаты подтверждают перспективность подхода, однако их нельзя напрямую экстраполировать на реальные образовательные контексты, поэтому дальнейшие исследования будут направлены на апробацию предложенной модели на масштабных эмпирических данных, собранных в условиях реальной цифровой образовательной среды. Особое внимание планируется уделить вопросам интерпретируемости результатов, корректной работы с неопределённостью прогнозов и строгому соблюдению этических принципов, включая прозрачность алгоритмов, защиту персональных данных и недопущение стигматизации обучающихся. Окончательной целью является создание на основе



данной модели этических, прозрачных и полезных инструментов, которые могут быть интегрированы в образовательные платформы.

В перспективе развитие данного направления может способствовать формированию нового класса интеллектуальных образовательных систем, ориентированных не только на оценку результатов обучения, но и на поддержку устойчивого когнитивного и психологического состояния участников образовательного процесса.

Список источников / References

1. Выготский, Л.С. (1982). *Проблемы общей психологии* (Собр. соч., Т. 2). Москва: Педагогика. Vygotsky, L.S. (1982). *Problemy obshchei psikhologii* [Problems of general psychology] (Vol. 2). Pedagogika.
2. Карпов, А.В. (2021). *Методологические основы психологического анализа информационной деятельности*. Филлигрань. Karpov, A.V. (2021). *Metodologicheskie osnovy psikhologicheskogo analiza informatsionnoi deiatel'nosti* [Methodological foundations of the psychological analysis of informational activity]. Filigran'.
3. Карташева, М.И. (2022) Роль системы Я в структуре ментальной регуляции психических состояний в процессе учебной деятельности. Ярославский педагогический вестник. 2022. № 1 (124). С. 162–174. <http://dx.doi.org/10.20323/1813-145X-2022-1-124-162-174> Kartasheva, M.I. (2022). *The role of the self-system in the structure of mental states regulation during the educational activities*. *Yaroslavl Pedagogical Bulletin*, (1), 162–174. <https://doi.org/10.20323/1813-145X-2022-1-124-162-174>
4. Коровникова, Н.А. Искусственный интеллект в образовательном пространстве: проблемы и перспективы. Социальные новации и социальные науки. — Москва: ИНИОН РАН, 2021. — № 2. — С. 98–113. Korovnikova, N.A. (2021). Artificial intelligence in the educational space: Problems and prospects. *Social Innovations and Social Sciences*, 2, 98–113. (In Russ.). Moscow: INION RAS.
5. Солдатова, Г.У. (2018). Цифровая социализация в культурно-исторической парадигме: изменяющийся ребенок в изменяющемся мире. *Социальная психология и общество*, 9(3), 71–80. <https://doi.org/10.17759/sps.2018090308> Soldatova, G.U. (2018). Digital socialization in the cultural-historical paradigm: A changing child in a changing world. *Social Psychology and Society*, 9(3), 71–80. (In Russ.). <https://doi.org/10.17759/sps.2018090308>
6. Юрьева, Н.Е. (2025). Искусственный интеллект в психодиагностике: когнитивные состояния в цифровой образовательной среде. *Моделирование и анализ данных*, 15(3), 47–55. <https://doi.org/10.17759/mda.2025150303> Yuryeva, N.E. (2025). Artificial intelligence in psychodiagnostics: cognitive states in a digital educational environment. *Modelling and Data Analysis*, 15(3), 47–55. (In Russ.). <https://doi.org/10.17759/mda.2025150303>
7. Baker, R.S. (2019). Challenges for the future of educational data mining: The Baker Learning Analytics Prizes. *Journal of Educational Data Mining*, 11(1), 1–17. <https://doi.org/10.5281/zenodo.3554745>



8. Baltrušaitis, T., Ahuja, C., & Morency, L.-P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2), 423–443. DOI: 10.1109/TPAMI.2018.2798607
9. D'Mello, S., & Kory, J. (2015). A review and meta-analysis of multimodal affect detection systems. *ACM Computing Surveys*, 47(3), Article 43. <https://doi.org/10.1145/2682899>. Открытый PDF (author version): <https://people.ict.usc.edu/~gratch/CSCI534/Readings/dmello-MMML-Survey.pdf>
10. Grafsgaard, J.F., Wiggins, J.B., Boyer, K.E., Wiebe, E.N., & Lester, J.C. (2014). Predicting learning and affect from multimodal data streams in task-oriented tutorial dialogue. In J. Stamper, Z. Pardoc, M. Mavrikis, & B.M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014)* (pp. 122–129). International Educational Data Mining Society. <https://learndialogue.org/pdf/LearnDialogue-Grafsgaard-EDM-2014.pdf>
11. Hamaker, E.L., Kuiper, R.M., & Grasman, R.P.P.P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, 20(1), 102–116. <https://doi.org/10.1037/a0038889>
12. He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 1026–1034). <https://doi.org/10.1109/ICCV.2015.123>
13. Jain, S. & Wallace, B.C. (2019). Attention is not explanation. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)* (pp. 3543–3556). <https://arxiv.org/abs/1902.10186>
14. Kuppens, P. & Verduyn, P. (2017). Emotion dynamics. *Current Opinion in Psychology*, 17, 22–26. <https://doi.org/10.1016/j.copsyc.2017.06.004>
15. Loshchilov, I., & Hutter, F. (2019). Decoupled weight decay regularization. In 7th International Conference on Learning Representations, ICLR 2019. <https://doi.org/10.48550/arXiv.1711.05101>
16. Reis, R.C.D., Isotani, S., Rodriguez, C.L., Jaques, P. A. & Bittencourt, I.I. (2018). Affective states in computer-supported collaborative learning: Studying the past to drive the future. *Computers & Education*, 120, 29–50. <https://doi.org/10.1016/j.compedu.2018.01.015>. Открытый PDF: <https://isiarticles.com/bundles/Article/pre/pdf/126884.pdf>
17. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems 30* (pp. 5998–6008). Curran Associates, Inc. URL: <https://arxiv.org/abs/1706.03762>
18. Wang, J., Yu, Y. (2025). Machine learning approach to student performance prediction of online learning. *PLOS ONE*, 20(1), e0299018. <https://doi.org/10.1371/journal.pone.0299018>

Информация об авторах

Юрьева Наталья Евгеньевна, кандидат технических наук, научный сотрудник, Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0003-1419-876X>, e-mail: yurieva.ne@gmail.com

Information about the authors

Nataliya E. Yuryeva, PhD (Engineering), Research Fellow, Moscow State University of Psychology and Education (MSUPE), Moscow, Russia, ORCID: <https://orcid.org/0000-0003-1419-876X>, e-mail: yurieva.ne@gmail.com



Юрьева Н.Е. (2026)

Моделирование и прогнозирование цифрового
когнитивного выгорания у студентов на основе...

Моделирование и анализ данных, 2026, 16(3), 91 – 103.

Yuryeva N.E. (2026)

Modeling and forecasting digital cognitive burnout
in students based on multimodal time series ...

Modelling and Data Analysis, 2026, 16(3), 91 – 103.

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.

Conflict of interest

The author declares no conflict of interest.

Поступила в редакцию 12.02.2026

Поступила после рецензирования 23.07.2026

Принята к публикации 10.08.2026

Опубликована 30.09.2026

Received 2026.02.12

Revised 2026.07.23

Accepted 2026.08.10

Published 2026.09.30