

## Оригинальная статья / Original article

УДК 004.89

<https://doi.org/10.21869/2223-1560-2025-29-2-109-129>

## Формирование индивидуальной образовательной траектории в онлайн обучении на основе технологий скрытых марковских моделей

**И.П. Бурукина<sup>1</sup>✉, Л.Н. Горшенин<sup>1</sup>**

<sup>1</sup> Пензенский государственный университет  
ул. Красная, д. 40, г. Пенза 440026, Российская Федерация

✉ e-mail: burukinairina@gmail.com

### Резюме

**Цель исследования.** Целью настоящего исследования является разработка и обоснование методики формирования индивидуальной образовательной траектории в рамках онлайн курсов посредством анализа учебной активности учащихся и уровня их академической успеваемости.

**Методы.** В работе применяются скрытые марковские модели, которые хорошо сочетаются с современными подходами машинного обучения, что усиливает их потенциал в плане аналитики и точного подбора образовательных траекторий. Выделены ключевые характеристики учебной активности учащихся, которые можно использовать в качестве наблюдений, а также выбрано подходящее количество скрытых состояний, соответствующее разным уровням академической успеваемости учащихся.

**Результаты.** Для экспериментального построения модели использовалась библиотека scikit-learn, разработанная для языка программирования Python. Обучение модели осуществлялось на двух массивах данных: реальная выборка включала 48942 записей результатов студентов по онлайн курсу «Технологии разработки интернет ресурсов», а дополнительный набор данных содержал 18052 записей из открытого репозитория Kaggle. Проведенное тестирование подтвердило эффективность предлагаемой методики, продемонстрировав улучшение качества образования благодаря точной оценке текущего состояния учащегося (учебной активности, уровня академической успеваемости), гибкому подбору учебных материалов и иной формы взаимодействия.

**Заключение.** Полученные результаты доказали перспективность использования предлагаемого подхода, способствующего повышению вовлеченности учащихся за счет особенностей восприятия учебного материала, увеличению скорости освоения новых компетенций путем оптимизации последовательности подачи учебного материала и возможности автоматизации процессов мониторинга прогресса учащихся. Исследование представляет особый интерес для специалистов, работающих над повышением эффективности онлайн обучения, и разработчиков образовательных платформ, желающих интегрировать такие модели в свои сервисы для поддержки педагогов и организаторов образовательного процесса.

**Ключевые слова:** индивидуальная стратегия; онлайн обучение; скрытая марковская модель; академическая успеваемость; активность учащихся; машинное обучение.

**Конфликт интересов:** Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

**Для цитирования:** Бурукина И.П., Горшенин Л.Н. Формирование индивидуальной образовательной траектории в онлайн обучении на основе технологий скрытых марковских моделей // Известия Юго-Западного государственного университета. 2025. Т. 29, №2. С. 109-129. <https://doi.org/10.21869/2223-1560-2025-29-2-109-129>.

Поступила в редакцию 28.04.2025

Подписана в печать 05.06.2025

Опубликована 23.07.2025

## Individual educational trajectory in online learning based on hidden Markov model technologies

Irina P. Burukina<sup>1</sup> , Lev N. Goshenin<sup>1</sup>

<sup>1</sup> Penza State University

40 Krasnaya str., Penza 440026, Russian Federation

 e-mail: burukinairina@gmail.com

### Abstract

**Purpose of research.** The purpose of this study is to develop and substantiate a methodology for the formation of an individual educational trajectory in online courses by analyzing the educational activity of students and their level of academic achievement.

**Methods.** Hidden Markov models are used in the work, which are well combined with modern machine learning approaches, which enhances their potential in terms of analytics and accurate selection of educational trajectories. The key characteristics of students' learning activity that can be used as observations are highlighted, and a suitable number of hidden states corresponding to different levels of students' academic performance are selected.

**Results.** The scikit-learn library, developed for the Python programming language, was used for experimental model construction. The model was trained on two data arrays: the real sample included 48942 records of students' results in the online course «Internet Resource Development Technologies», and an additional data set contained 18052 records from the Kaggle open repository. The conducted testing confirmed the effectiveness of the proposed methodology, demonstrating an improvement in the quality of education due to an accurate assessment of the student's current state (academic activity, academic achievement), flexible selection of educational materials and other forms of interaction.

**Conclusion.** The obtained results proved the prospects of using the proposed approach, which helps to increase student engagement due to the peculiarities of the perception of educational material, increase the speed of mastering new competencies by optimizing the sequence of presentation of educational material and the possibility of automating the processes of monitoring student progress. The study is of particular interest to specialists working to improve the effectiveness of online learning, and developers of educational platforms who want to integrate such models into their services to support teachers and organizers of the educational process.

**Keywords:** individual strategy; online learning; hidden Markov model; academic performance; student activity; machine learning.

**Conflict of interest.** The Authors declare the absence of obvious and potential conflicts of interest related to the publication of this article.

**For citation:** Burukina I. P., Gorshenin L. N. Individual educational trajectory in online learning based on hidden Markov model technologies. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta = Proceedings of the Southwest State University*. 2025; 29(2): 109-129 (In Russ.). <https://doi.org/10.21869/2223-1560-2025-29-2-109-129>.

Received 28.04.2025

Accepted 05.06.2025

Published 23.07.2025

\*\*\*

## Введение

В условиях динамично развивающегося информационного общества образование приобретает особую значимость, становясь ключевым фактором успешного развития личности и государства. Доступность интернета и распространение мобильных устройств привели к стремительному развитию рынка онлайн курсов и дистанционных образовательных сервисов. По данным экспертов [1], ежегодно аудитория онлайн курсов увеличивается примерно на треть, предлагая на выбор широкий спектр интересующих направлений. Несмотря на доступность большого количества материалов, традиционные подходы к организации онлайн обучения часто оказываются неэффективными, поскольку не учитывают индивидуальные потребности учащихся [2]. Как следствие, многие учащиеся сталкиваются с трудностями освоения материала, теряют мотивацию и бросают обучение раньше завершения курса [3,4]. Другой проблемой является несбалансированность нагрузки на преподавателей, особенно в массовых открытых онлайн курсах (*MOOCs*), где число учащихся достигает тысячи человек одновременно. Преподаватели физически не способны оказать персональное внимание каждому слушателю, что снижает качество обратной связи и затрудняет своевременное выявление трудностей, возникающих в процессе обучения. Кроме этого, существующие курсы с индивидуальными образовательными траекториями, зачастую основаны на усред-

ненных показателях и общих тенденциях [5], что делает невозможным адекватную адаптацию учебной программы к потребностям учащегося и негативно оказывается на конечном результате обучения. Ученые активно исследуют способы построения индивидуальных образовательных траекторий посредством внедрения различных подходов к обработке больших объемов данных об учащихся [6,7]. Однако имеющиеся решения имеют ряд ограничений, связанных с необходимостью значительных вычислительных ресурсов и сложностью интерпретации полученных выводов [8], что снижает возможность массового использования данных методик при разработке образовательных ресурсов.

Предлагаемый авторами подход, основанный на применении скрытых марковских моделей (*CMM*), представляется эффективным способом построения индивидуальных образовательных траекторий при онлайн обучении. Скрытые марковские модели способны оперативно реагировать на изменение уровня знаний, мгновенно адаптируя дальнейший маршрут обучения с учетом показателя активности учащегося на образовательной платформе. Скрытые марковские модели хорошо сочетаются с современными подходами машинного обучения, что усиливает их потенциал в плане аналитики и точного подбора образовательных траекторий.

Целью настоящего исследования является разработка и обоснование методики автоматизированной коррекции

индивидуальной образовательной траектории в рамках онлайн курсов посредством анализа учебной активности учащихся и уровня их академической успеваемости с применением скрытых марковских моделей.

Для достижения обозначенной цели сформулированы следующие задачи исследования:

- проведение аналитического обзора современных методов формирования индивидуальных образовательных траекторий и выявление направлений их возможного совершенствования;
- построение математической модели поведения учащегося на основе скрытых марковских процессов, учитывающей факторы активности и успеваемости;
- реализация программного инструментария для обработки данных о поведении и успеваемости учащихся, обеспечивающего автоматическую коррекцию образовательной траектории;
- экспериментальная проверка эффективности разработанной методики.

### **Материалы и методы**

Интенсивные научные исследования в области психологии и педагогики, направленные на понимание индивидуальных особенностей восприятия, памяти, внимания и стиля мышления учащихся, убедительно продемонстрировали, что унифицированные подходы неэффективны для удовлетворения образовательных потребностей общества. В рамках современных научных направ-

лений активно разрабатываются модели дифференцированного и индивидуально-ориентированного обучения.

В работе [9] Лапенок М.В. и Макеевой В.В. предлагается оригинальная методика формирования индивидуальной траектории обучения, основанная на применении технологической и дидактической карт. Технологическая карта служит инструментом пошагового планирования и управления процессом формирования индивидуальной траектории обучения учащегося и включает три стадии: постановочную, практико-ориентированную и контрольно-оценочную. Дидактическая карта служит источником методологического сопровождения образовательного процесса, предоставляя рекомендации по формированию содержания и структуры учебного курса, предлагая конкретные варианты коррекции образовательной траектории, исходя из выявленных трудностей. Предложенная методика, несмотря на очевидные преимущества, среди педагогов-практиков вызывает дискуссии относительно того, насколько реально воплотить этот подход в массовых онлайн курсах, где ресурсы могут быть ограничены, а численность учащихся высока.

В работе [10] Троицкой Е.А. предлагается нестандартный подход к автоматизации формирования индивидуальной стратегии, на основе автоматизированных систем обучения. Цикл обучения подразделяется на два этапа: этап первичной диагностики и этап закрепления умений и навыков. Первичная

диагностика включает сбор и накопление данных о результатах выполнения стандартных задач. На данном этапе предлагается индивидуальный пакет типовых обучающих задач, автоматически формируемых из базы данных. Второй этап предполагает подготовку и реализацию индивидуальной последовательности упражнений, направленных на повторение и закрепление методов решения задач. Данный подход автоматизирует формирование индивидуальной стратегии обучения, однако вся коррекция сводится к генерации дополнительных заданий для закрепления ранее изученного материала, что создает предпосылки для последующего усовершенствования предложенной схемы.

Васильченко С.Х. в своей работе [11] предлагает алгоритм построения индивидуальной образовательной траектории путем формирования персональной образовательной среды с использованием информационных технологий. Алгоритм включает пять последовательных этапов (диагностика обучаемого, конкретизация целей и задач обучения, наполнение персональной образовательной среды компонентами, наполнение персональной образовательной среды содержанием и выбор способа управления средой) и предоставляет учащемуся вариант самостоятельного выбора образовательного пути. Среди возможных способов управления выделяются три основных типа:

– управление на основе обратной связи (консультации, интерактивное общение);

– отсроченное управление (коррекция курса по результатам финального тестирования);

– немедленное управление (коррекция обучения после завершения каждого этапа).

Каждый из указанных способов управления несет определенные ограничения, препятствующие эффективному использованию в массовых системах образования: зависимость от человеческого фактора, низкий уровень автоматизации процессов коррекции, риск возникновения ошибок при принятии управлений решений из-за ошибочных интерпретаций полученных данных. Для устранения вышеуказанных ограничений целесообразно разрабатывать комплексные информационные системы, интегрирующие элементы машинного обучения, механизмы больших данных и т.д.

Современные тенденции в образовательной сфере характеризуются активным внедрением методов искусственного интеллекта, направленных на повышение эффективности образовательного процесса и адаптацию учебных методик к индивидуальным потребностям учащихся [12].

Например, исследование Подколзина М.М. [13] посвящено разработке и внедрению в российские вузы интеллектуальной системы адаптивного обучения на основе сверточной (*Convolutional Neural Network, CNN*) и рекуррентной (*Recurrent Neural Network, RNN*) нейронных сетей. Основная цель заключается в создании инструментов

персонализации образовательного процесса, позволяющих учитывать индивидуальные особенности студентов и повышать качество и эффективность обучения. Хотя данная работа продемонстрировала высокую эффективность в конкретных условиях эксперимента, существуют вопросы, которые требуют дополнительного изучения и рассмотрения. Во-первых, предложенный подход обеспечивает высокий процент точности прогнозирования итоговых результатов обучения, однако не акцентирует внимание на построении моделей, способных предсказывать промежуточные этапы образовательного процесса и рекомендовать корректирующие меры для предотвращения возможных неудач. Во-вторых, представленная структура гибридной нейронной сети требует значительных вычислительных ресурсов для обработки больших объемов данных, особенно при увеличении числа студентов и дисциплин, что ставит под угрозу масштабируемость данного подхода и увеличивает затраты на поддержку инфраструктуры. Кроме этого, хотя предложенная система обеспечивает техническую поддержку процесса преподавания, присутствующие традиционные схемы оценивания отличаются высокой степенью субъективизма, создавая ложное впечатление об успехах или проблемах в обучении.

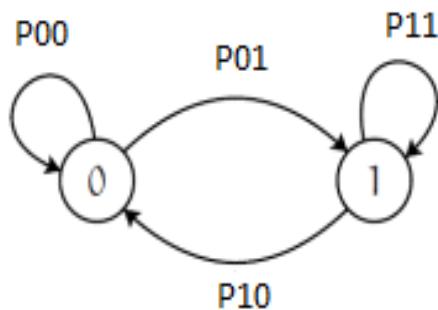
В условиях онлайн обучения, характеризующегося отсутствием личного взаимодействия преподавателя с учениками и значительным увеличением численности обучающихся, существенно повыша-

ется значение применения технологий искусственного интеллекта. Современные платформы, такие как *DreamBox Learning* (США), *Knewton* (США), *Smart Sparrow* (Австралия), *Geekie* (Бразилия), *Учи.ру* (Россия), *Яндекс.Учебник* (Россия) позволяют создавать адаптивные учебные среды, способные изменять содержание и структуру онлайн курсов в режиме реального времени, исходя из достижений учащихся. Подобные системы, как правило, характеризуются наличием большого количества подсказок и примеров с готовыми решениями, что допускает пропуск отдельных этапов обучения. Однако одним из основных требований любого учебного процесса является полное освоение всего объема необходимого материала. Несоблюдение этого принципа ведет к возникновению пробелов в знаниях, снижению общей компетентности и неэффективности дальнейшей профессиональной деятельности учащегося. Именно поэтому важным условием эффективной работы адаптивной системы должна стать обязательность изучения всех предусмотренных программой элементов.

Авторы данной научной работы предлагают использовать для формирования индивидуальной образовательной траектории учащегося с обеспечением полного охвата содержания учебной программы возможности скрытых марковских моделей [14].

Цепь Маркова представляет собой стохастическую последовательность случайных событий с ограниченным числом

возможных исходов [15]. При этом вероятность наступления каждого следующего события зависит только от настоящего состояния, т.е. при фиксированном настоящем будущее состояние независимо от прошлого. Математически дискретную марковскую цепь удобно интерпретировать как ориентированный граф, вершины которого соответствуют возможным состояниям системы, а дуги отображают возможные переходы между этими состояниями. Вероятности соответствующих переходов задаются весами указанных дуг (рис.1). Таким образом, полная характеристика такого графа содержится в матрице вероятностей переходов, строки которой соответствуют начальному состоянию, столбцы – конечному, а элементы матрицы показывают вероятность перемещения из одного состояния в другое. Важное условие для такой матрицы заключается в том, что сумма элементов каждой строки равняется единице, отражая полную определенность возможного направления дальнейшего движения из любого начального состояния.



**Рис. 1.** Пример марковской модели с двумя состояниями

**Fig. 1.** An example of a Markov model with two states

Скрытая модель Маркова [16] отличается от обычной тем, что истинные состояния и их количество скрыты, известны лишь наблюдаемые значения. Формально модель можно описать с помощью тройки  $(A, B, \pi)$ , где  $A$  – матрица вероятностей переходов между состояниями;  $B$  – матрица вероятностей появления наблюдаемых значений в различных состояниях;  $\pi$  – начальное распределение состояний.

На основании анализа научных трудов и публикаций авторы систематизировали основные типы задач в рамках теории скрытых марковских моделей, каждая из которых обладает уникальной спецификой и находит применение в различных областях исследований:

1. Оценка вероятности наблюдений: определение вероятности конкретной наблюдаемой последовательности в заданной скрытой марковской модели с параметрами  $(A, B, \pi)$ . Эта проблема возникает, например, при анализе акустических сигналов и находит применение в распознавании речи и обработке звукового сигнала [17].

2. Определение оптимальной траектории состояний: установление наиболее вероятной последовательности внутренних состояний модели, соответствующей заданной последовательности наблюдений. Данная задача встречается, например, в биомедицинском моделировании, включая исследование побочных эффектов лекарственных препаратов [18].

3. Подбор оптимальных параметров модели: выбор лучших значений мат-

риц переходов  $A$ , эмиссионных вероятностей  $B$  и начальных распределений  $\pi$ , обеспечивающих максимальное правдоподобие наблюдаемой последовательности. Такая постановка характерная для многих прикладных областей, включая метеорологию и климатический мониторинг [19].

Очевидно, что широкий спектр возможностей, представленных скрытыми марковскими моделями, делает их перспективным инструментом для реализации инновационных подходов и в области дистанционного образования. Они отлично подходят для моделирования процессов, развивающихся во времени, так как позволяют учитывать последовательность действий, предпринимаемых учащимся, отслеживать прогресс и выявлять этапы, на которых учащийся испытывает наибольшие трудности. Кроме того, скрытые марковские модели относительно просты в реализации и обучении, по сравнению с более сложными моделями, например, рекуррентными нейронными сетями [16]. Это делает СММ доступными для использования в широком спектре образовательных сред, даже с ограниченными вычислительными ресурсами.

Несмотря на указанные преимущества, важно отметить проблему выбора оптимального числа скрытых состояний в СММ, поскольку недостаточное их количество ведет к излишнему упрощению структуры модели, тогда как избыточное – способствует переобучению и затрудняет последующий анализ ре-

зультатов. Скрытые марковские модели также имеют ограничения в описании нелинейной траектории обучения, поэтому категорию учащихся, нуждающихся в повторении пройденного материала, целесообразно исключить из рассмотрения возможных состояний. Во избежание риска недообучения или переобучения предложен метод многократного обучения различных экземпляров модели с различными начальными параметрами, среди которых выбирается оптимальная версия согласно критерию максимального логарифмического правдоподобия [20].

## Результаты и их обсуждение

Начальным этапом работы с СММ является формирование четкого представления о структуре наблюдаемых данных и внутренней динамики изучаемой системы. Учитывая результаты научных исследований [21], демонстрирующие наличие статистически значимых корреляций между отдельными показателями учебной активности учащихся и уровнем их академической успешности, авторы определили ключевые характеристики поведения учащихся, пригодные для наблюдения и последующего анализа.

Наблюдаемые значения выступают в роли объективных индикаторов, отображающих особенности поведения учащегося в ходе изучения выбранного онлайн курса, и к ним относятся:

- Результаты промежуточного тестирования. Для упрощения обработки результатов предлагается применить шкалу из четырех уровней («неудовлетворитель-

но», «удовлетворительно», «хорошо», «отлично»), каждый из которых отражает степень готовности учащегося к прохождению следующих этапов онлайн курса.

- Частота обращений к учебно-методическим материалам. По данному параметру целесообразно выделить три градации активности, характеризующие интенсивность использования цифрового ресурса каждым участником образовательного процесса.

- Продолжительность сеансов на платформе. Регистрация временных интервалов пребывания учащегося на учеб-

ной платформе служит важным показателем вовлеченности в учебный процесс.

Обоснование оптимального числа скрытых состояний выполнено посредством классического метода Байесовского информационного критерия (*BIC*) [22]. Модели оценивались при различных количествах скрытых состояний – двух, трех, четырех и пяти – применительно к обучающему набору данных, включавшему 526 наблюдений. Анализ представленных в табл.1 результатов расчета свидетельствует о минимальном значении показателя *BIC* при выборе трех скрытых состояний.

**Таблица 1.** Значения Байесовского информационного критерия

**Table 1.** Bayesian Information Criterion Values

Количество скрытых состояний / Number of hidden states	Значение <i>BIC</i> / <i>BIC</i> meaning
2	1262.776
3	1258.89
4	1263.124
5	1304.362

Следовательно, оптимально выбрать именно три скрытых состояния, соответствующих следующим категориям учащихся:

- Учащиеся, которым не требуются дополнительные меры поддержки. Они демонстрируют высокий уровень самостоятельности и легко усваивают предлагаемый учебный контент без посторонней помощи.

- Учащиеся, нуждающиеся в углубленном изучении отдельных тем. Эта группа характеризуется потребностью в дополнительных материалах (видеолек-

ции, учебники, учебные пособия и т.д.) для улучшения понимания предмета, однако не требуют вмешательства преподавателя для консультации.

- Учащиеся, испытывающие большие трудности в освоении учебного материала и остро нуждающиеся в консультативной помощи преподавателя. Данная категория требует комплексных мер сопровождения, включая и предоставление дополнительных дидактических материалов, и непосредственное взаимодействие с преподавателем, направленное

на устранение пробелов в понимании учебной информации.

Так как параллельно наблюдаются разные наборы значений, каждый из которых может говорить о внутреннем состоянии модели – решено было обучить три модели, а в качестве наиболее вероятного скрытого состояния брать наиболее часто встречающийся ответ моделей, а в конфликтных ситуациях – ответ модели, обученной на результатах промежуточного контроля. При этом после обучения первой модели, матрица вероятностей переходов должна фиксироваться, так как все наблюдаемые значения зависят от одних и тех же скрытых состояний, а вероятности переходов между состояниями не зависят от наблюдаемых значений. Начальное распределение состояний должно инициализироваться случайным образом перед началом обучения, как и матрицы вероятностей появления наблюдаемых значений, так как отсутствуют данные, позволяющие определить для них подходящие начальные значения.

Математически описанные скрытые марковские модели можно представить следующими формулами:

$$X=\{K1,K2,K3\}, \quad (1)$$

$$Y_1=\{R1,R2,R3,R4\}, \quad (2)$$

$$Y_2=\{F1,F2,F3\}, \quad (3)$$

$$Y_3=\{T1,T2,T3\}, \quad (4)$$

$$\lambda_i=\{A,B_i,\pi\}, \quad (5)$$

где  $X$  – скрытые состояния модели ( $K1$  – не требуются дополнительные меры поддержки,  $K2$  – требуются дополнительные материалы для изучения отдельных тем,

$K3$  – требуются комплексные меры, включая и предоставление дополнительных материалов, и непосредственное взаимодействие с преподавателем);

$Y_1$  – наблюдаемые результаты промежуточного тестирования модели ( $R1$  – оценка «неудовлетворительно»,  $R2$  – оценка «удовлетворительно»,  $R3$  – оценка «хорошо»,  $R4$  – оценка «отлично»);

$Y_2$  – наблюдаемая частота обращения к материалам ( $F1$  – низкая,  $F2$  – средняя,  $F3$  - высокая);

$Y_3$  – наблюдаемое время пребывания на платформе ( $T1$  – короткое,  $T2$  – среднее,  $T3$  – длительное);

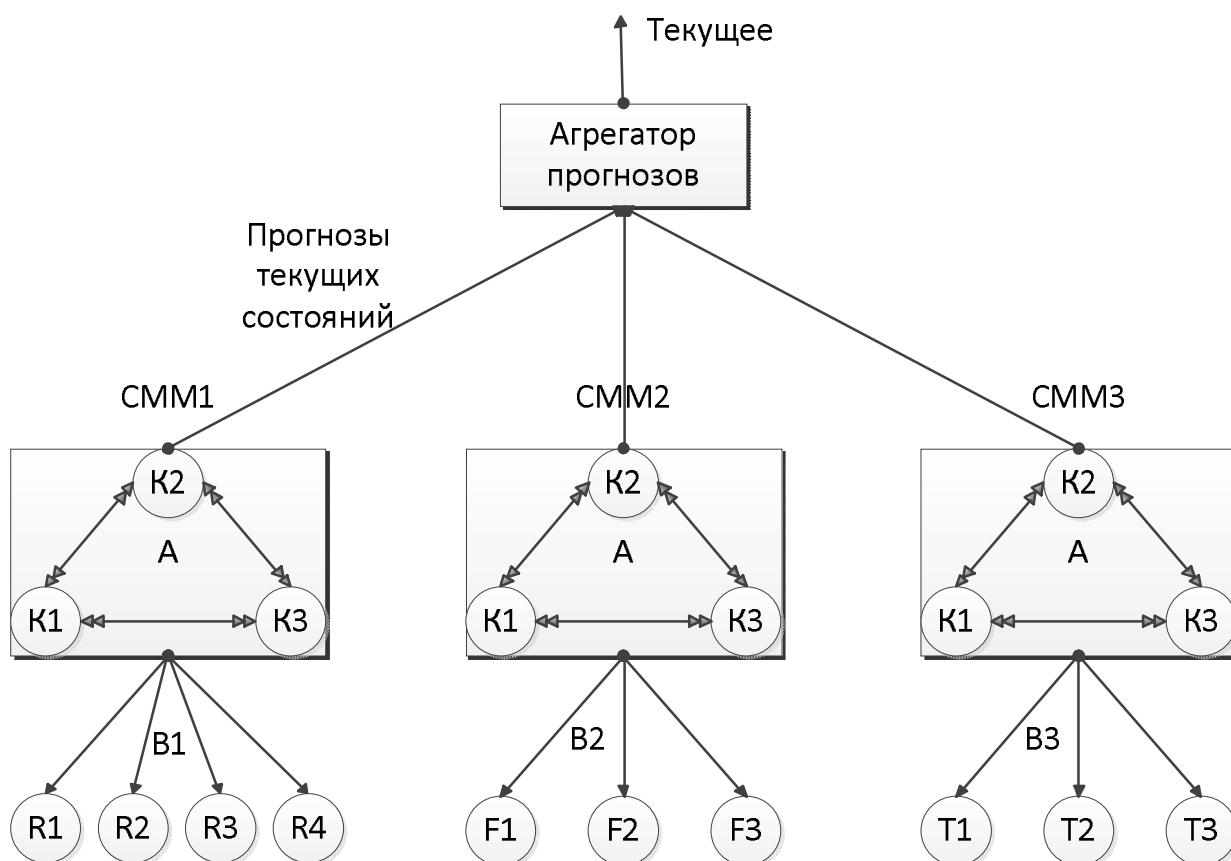
$\lambda_i$  – параметры модели ( $A$  – неизвестная матрица вероятностей перехода,  $B_i$  – матрица вероятностей появления наблюдаемых значений в скрытых состояниях,  $\pi$  - вектор начальных вероятностей).

Структура ансамбля моделей для результатов промежуточного контроля представлена на рис.2, где через  $A$  обозначены дуги с вероятностями перехода из одного состояния в другое, а через  $B_i$  – вероятности наблюдения значений из множества  $Y_i$  при определённых скрытых состояниях.

Реализация предлагаемой методики включает два этапа. Первый этап – обучение модели, заключающееся в подборе оптимальных параметров применительно к особенностям учебной дисциплины. Второй этап предусматривает использование обученной модели для анализа результатов промежуточного тестирования и поведения учащегося на

образовательной платформе. Цель такого анализа – выявление наиболее вероятного скрытого состояния обучаемого в текущий момент времени, что позволит оперативно скорректировать индивидуальную

образовательную стратегию путем подбора соответствующих дополнительных материалов и/или организации консультативной поддержки преподавателя.



**Рис. 2. Структура ансамбля моделей для коррекции индивидуальной стратегии обучения**

**Fig. 2. The structure of an ensemble of models for correcting an individual learning strategy**

Для нахождения неизвестных параметров модели часто применяется алгоритм Баума-Велша [23,24]. Основная часть алгоритма заключается в последовательных процедурах подсчёта ожидания данных при условии модели и изменения модели при условии данных. По сути, этот алгоритм является частным случаем алгоритма ожидания-максимизации (*Expectation-Maximization*). Начальные значения  $\lambda$  задаются случайным образом. В алгоритме используют-

ся значения вероятности получения заданной наблюдаемой последовательности  $Y$  из заданного начального состояния  $i$  (всего  $N$  состояний) в момент времени  $t$  (6) и вероятности получения заданной наблюдаемой последовательности  $Y$  к моменту времени  $T$  при условии начального состояния  $i$  в момент времени  $t$  (7). Также используются оценки вероятности состояния  $i$  в момент времени  $t$  (8) и вероятности перехода из состояния  $i$  в состояние  $j$  в момент времени  $t$  (9).

$$\begin{aligned} \alpha_i(1) &= \pi_i * b_i(Y_1); \alpha_i(t+1) = \\ &= b_j(Y_{t+1}) \sum_{i=1}^N (\alpha_i(t) * a_{i,j}). \end{aligned} \quad (6)$$

$$\begin{aligned} \beta_i(1) &= 1; \beta_i(t) = \\ &= \sum_{j=1}^N (\beta_j(t+1) * a_{i,j} * b_j(Y_{t+1})). \end{aligned} \quad (7)$$

$$\gamma_i(t) = \frac{\alpha_i(t) * \beta_i(t)}{\sum_{j=1}^N (\alpha_j(t) * \beta_j(t))}. \quad (8)$$

$$\varphi_{i,j}(t) = \frac{\alpha_i(t) * a_{i,j} * \beta_j(t+1) * b_j(Y_{t+1})}{\sum_{i=1}^N \sum_{j=1}^N \alpha_i(t) * a_{i,j} * \beta_j(t+1) * b_j(Y_{t+1})}. \quad (9)$$

Для пересчёта параметров модели на очередной итерации используются следующие формулы:

$$\pi_i = \gamma_i(1), \quad (10)$$

$$a_{i,j} = \frac{\sum_{t=1}^{T-1} \varphi_{i,j}(t)}{\sum_{t=1}^{T-1} \gamma_i(t)}, \quad (11)$$

$$b_i(k) = \frac{\sum_{t=1}^T \delta(Y_t, Y_k) * \gamma_i(t)}{\sum_{t=1}^T \gamma_i(t)}, \quad (12)$$

где  $\delta$  – символ Кронекера, результат которого равен единице, если аргументы равны, и ноль – в противном случае;  $Y$  – вектор возможных наблюдаемых значений;  $k$  – индекс наблюдаемого значения.

Обучение скрытой марковской модели может осуществляться с использованием различных инструментальных средств. В рамках настоящего исследования для экспериментального построения и обучения модели была выбрана библиотека *hmmlearn* версии 0.3.3, разработанная для языка программирования *Python* версии 3.10 – высокуюровневого интерпретируемого языка, широко применяемого в области машинного обучения [25]. Библиотека *hmmlearn* предоставляет широкий спектр модулей, предназначенных для реализации алгоритмов машинного обучения моделей СММ [26].

Для тестового обучения модели использовались два набора данных. Массив реальных данных, собранный в процессе преподавания учебного онлайн курса «Технологии разработки интернет ресурсов» для студентов направления подготовки 09.03.01 «Информатика и вычислительная техника» в Пензенском государственном университете автором И.П. Бурукиной. Массив включает 48942 последовательности, зафиксированные в течение двух лет преподавания среди 526 учащихся на платформе Moodle. Дополнительно был использован массив «*student-semester-result*» из репозитория *Kaggle*, состоящий из 18052 последовательностей, и отражающих, так же как в первом случае, оценочные показатели промежуточного контроля, активность учащихся при обращении к учебному контенту и продолжительность сессий на цифровой образовательной платформе. Наборы данных решено было разделить на обучающую и валидационную выборки в пропорции 9 к 1.

Разработанная на языке *Python* программа реализует процедуру обучения 10000 моделей для каждого набора входных данных, начиная с произвольно инициализированных начальных распределений и матриц вероятностей. Пример произвольных начальных матриц представлен формулами (13), (14), (15). Каждая модель проходит процесс обучения, состоящий из 15 итераций, при этом алгоритм останавливается досрочно, если прирост логарифма правдоподобия становится меньше заданного порога в

0.01. Для демонстрации процесса сходимости приведены примеры матриц перехода на различных этапах итерации (формулы (16), (17), (18)). По окончании процедуры обучения осуществляется отбор оптимальной модели путем оценки ее производительности на предварительно выделенной валидационной

выборке. Оптимальность определяется максимальным значением логарифмического правдоподобия среди всех обученных моделей, результаты которого зафиксированы в табл.2. Согласно представленным данным, наилучшие характеристики демонстрирует первая модель.

**Таблица 2.** Таблица значений логарифмического правдоподобия

**Table 2.** Table of log-likelihood values

Модель / Model	Информационный критерий Акаике / Akaike Information Criterion	Логарифмическое правдоподобие / Log-likelihood
1. Результаты промежуточного тестирования	34.83	0.36
2. Частота обращений к учебно-методическим материалам	30.34	-1.1
3. Продолжительность сеансов на платформе	29.4	-0.74

Матрица вероятностей переходов между состояниями обученной модели имеет следующий вид (19). Матрицы вероятностей выдачи значений приведены в формулах (20), (21), (22).

$$(0.359 \quad 0.476 \quad 0.165) \quad (13)$$

$$\begin{pmatrix} 0.1 & 0.67 & 0.22 \\ 0.004 & 0.006 & 0.99 \\ 0.97 & 0.02 & 0.01 \end{pmatrix} \quad (14)$$

$$\begin{pmatrix} 0.228 & 0.297 & 0.25 & 0.226 \\ 0.177 & 0.269 & 0.182 & 0.372 \\ 0.361 & 0.144 & 0.297 & 0.198 \end{pmatrix} \quad (15)$$

$$\begin{pmatrix} 0.114 & 0.658 & 0.227 \\ 0.004 & 0.006 & 0.99 \\ 0.978 & 0.021 & 0.001 \end{pmatrix} \quad (16)$$

$$\begin{pmatrix} 0.113 & 0.660 & 0.225 \\ 0.004 & 0.006 & 0.99 \\ 0.979 & 0.020 & 0.001 \end{pmatrix} \quad (17)$$

$$\begin{pmatrix} 0.116 & 0.662 & 0.221 \\ 0.004 & 0.006 & 0.99 \\ 0.979 & 0.020 & 0.001 \end{pmatrix} \quad (18)$$

$$A = \begin{pmatrix} 0.8 & 0.199 & 0.001 \\ 0.095 & 0.722 & 0.183 \\ 0.114 & 0.048 & 0.838 \end{pmatrix} \quad (19)$$

$$B1 = \begin{pmatrix} 0.001 & 0.094 & 0.804 & 0.101 \\ 0.002 & 0.773 & 0.224 & 0.001 \\ 0.362 & 0.597 & 0.040 & 0.001 \end{pmatrix} \quad (20)$$

$$B2 = \begin{pmatrix} 0.737 & 0.26 & 0.003 \\ 0.014 & 0.942 & 0.044 \\ 0.006 & 0.546 & 0.448 \end{pmatrix} \quad (21)$$

$$B3 = \begin{pmatrix} 0.003 & 0.663 & 0.334 \\ 0.033 & 0.941 & 0.026 \\ 0.773 & 0.219 & 0.008 \end{pmatrix} \quad (22)$$

Из значений матрицы  $A$  можно определить следующее: крайне мала вероятность перехода из скрытого состоя-

ния  $K1$  в  $K3$  (0.001); вероятность перехода из  $K2$  в  $K1$  значительно меньше, чем в  $K3$ ; высока вероятность остаться в текущем скрытом состоянии. Из значений матрицы  $B1$  видно, что если обучающийся находится в скрытом состоянии  $K1$  (первая строка), то он имеет очень высокую вероятность получить оценки «отлично» и «хорошо» (4-ый и 3-ий столбец соответственно). Если обучающийся находится в состоянии  $K2$  (вторая строка), он с большей вероятностью получит оценки «удовлетворительно» (2-ой столбец) и «хорошо» (3-ий столбец). В состоянии  $K3$  (третья строка) – «неудовлетворительно» (1-ый столбец) и «удовлетворительно» (2-ой столбец). Значения матрицы  $B2$  показывают, что при скрытом состоянии  $K1$  обучающиеся не слишком часто обращаются к материалам, возможно, из-за

более быстрого понимания материала. Из значений матрицы  $B3$  можно понять, что чем больше времени обучающийся провёл на цифровой образовательной платформе – тем выше будет его категория (от  $K3$  до  $K1$ ). В целом можно заключить, что полученные значения матриц адекватно отражают действительность.

Для проверки работоспособности модели был составлен ряд запросов в виде последовательностей результатов промежуточного контроля, количества обращений к учебному контенту и времени сессий на платформе. Запросы и результаты представлены на рис. 3. Наблюдаемые значения результатов промежуточного контроля представлены в форме:  $R1 - 0, R2 - 1, R3 - 2, R4 - 3$ ; значения из остальных наборов наблюдаемых значений – от 0 до 2; предсказанные ансамблем модели состояния:  $K1, K2, K3$ .

```

ensemble.get_result([[2,1,2,2,1,2,2,1,0]], [[1,2,1,1,1,0,1,2,2]], \
[[1,2,1,1,0,0,1,2]])
'K3'
ensemble.get_result([[3,2,3,2,2,3,3,1,2]], [[1,2,1,1,0,0,1,1,2]], \
[[1,2,1,0,1,0,1,2,1]])
'K1'
ensemble.get_result([[3,2,3,2,2,3,2,1,1]], [[1,2,1,1,0,0,1,1,2]], \
[[1,2,1,0,1,0,1,2,1]])
'K2'
ensemble.get_result([[2,0,2,0,2,0,2,0,2]], [[1,2,1,1,2,1,1,2,1]], \
[[1,2,1,1,1,0,0,1,2]])
'K1'
ensemble.get_result([[0,2,0,2,0,2,0,2,0]], [[2,1,2,1,2,1,2,1,2]], \
[[1,2,1,1,1,0,0,1,2]])
'K3'

```

**Рис. 3.** Предсказания состояний модели

**Fig. 3.** Predictions of model states

Анализ представленных данных (рис. 3) демонстрирует способность модели эффективно реагировать на измене-

ния результатов промежуточных оценочных мероприятий учащихся и уровня их активности на цифровой образо-

вательной платформе. При регулярном появлении наблюдаемых значений  $R3$  («хорошо») и  $R2$  («удовлетворительно»), особенно если данные показатели совмещены с показателями активности – учащийся обращается к учебно методическим материалам, выполняет практические задания, участвует в форумах и т.д., наблюдается переход в состояние  $K2$  (требуются дополнительные учебные материалы). Если же, несмотря на значительную активность учащегося, отмечается последовательное ухудшение академической успеваемости от уровня  $R2$  до  $R1$  («неудовлетворительно»), диагностируется более критическое состояние  $K3$ . Для устранения такой ситуации необходимо проведение радикальных мер коррекции стратегии обучения, включающие предоставление дополнительных учебных материалов и обязательную консультацию с преподавателем. При преобладании высоких показателей  $R4$  («отлично») и  $R3$  («хорошо») система остается в стабильном состоянии  $K1$ , которое свидетельствует об отсутствии трудностей у учащегося при обучении, материал ему ясен и понятен, все задания студент выполняет, потребности во вмешательстве и дополнительном контроле нет.

Отдельного внимания заслуживают случаи циклических колебаний между независимыми показателями  $R3$  («хорошо») и  $R2$  («удовлетворительно»), сопровождающиеся низким уровнем участия обучающегося в изучении базовых учебно-методических ресурсов и редким по-

сещением электронной среды обучения. Такая динамика расценивается моделью как сигнал перехода в состояние  $K3$ , подчеркивая важность своевременного выявления низкой вовлеченности учащегося и инициирования комплексного подхода к исправлению ситуации посредством сочетания обязательной консультации с преподавателем, а после представления дополнительных материалов. Подчеркнем, что предлагать дополнительную литературу без проведения консультаций неэффективно, поскольку учащийся проявляет пассивность в освоении основного учебного материала.

В двух последних запросах проводилось поведение модели на неординарных входных последовательностях: чередование оценок «неудовлетворительно» и «хорошо». На эти запросы модель ответила в соответствии с последней оценкой: если «неудовлетворительно» – то  $K3$  (требуются комплексные меры), если «хорошо» – то  $K1$  (мер не требуется). Однако, в любом случае, такие последовательности могут свидетельствовать о нечестном прохождении некоторых тестирований и, следовательно, о непонимании материала и требуют комплексных мер. Скорее всего, модель ответила не вполне корректно на один из запросов из-за очень малого количества подобных неординарных последовательностей в обучающем наборе данных.

Эффективность разработанной модели оценивалась путем сравнения академической успеваемости эксперимен-

тальной группы, проходившей дистанционное обучение с применением *CMM*, и контрольной группы, обучавшейся онлайн по той же дисциплине традиционным способом. Исследование проводилось среди студентов Пензенского государственного университета направления подготовки бакалавров 09.03.01 «Информатика и вычислительная техника». Чтобы обеспечить высокий уровень надежности и валидности результатов эксперимента, предварительно было выполнено статистическое обоснование объема выборки [27], определившее минимальную численность каждой группы – 22

участника. Для равномерного распределения студентов между двумя группами, перед началом обучения они прошли процедуру предварительного электронного тестирования, предназначенного для объективной оценки исходного уровня знаний каждого учащегося. Полученные данные подверглись однофакторному дисперсионному анализу, подтвердившему (высокое значение показателя F-статистики (9.2656) при низком значении вероятности p-value (0.0001)) положительное влияние использования *CMM* на повышение уровня академической успеваемости учащихся, что видно из табл. 3.

**Таблица 3.** Результаты однофакторного дисперсионного анализа

**Table 3.** The results of a one-way analysis of variance

Источник изменчивости / Source of Variability	Сумма квадратов (SS) / Sum of Squares (SS)	Степень свободы (DF) / Degrees of Freedom (DF)	Средние квадраты (MS) / Mean Squares (MS)	F	p-value
Между группами	626.087	1	626.087	9.2656	0.0001
Внутри групп	1486.5652	43	33.7856		
Всего	2112.6522	44			

Таким образом, разработанная скрытая марковская модель обеспечивает надежную идентификацию текущего образовательного статуса учащегося на основании динамического анализа результатов промежуточных контрольных испытаний и показателей активности, позволяя своевременно определить наиболее целесообразные меры требуемой коррекции в индивидуальной стратегии обучения.

## Выводы

В настоящей работе предложен подход к формированию индивидуальных образовательных траекторий на основе скрытых марковских моделей, применяемых в системах онлайн обучения. Полученные результаты подтверждают возможность существенного улучшения качества образования благодаря точной оценке текущего состояния учащегося и

гибкому подбору материалов для изучения или иной формы взаимодействия.

Выявленные преимущества включают:

- повышение вовлеченности учащихся за счет особенностей восприятия учебного материала;
- увеличение скорости освоения новых компетенций путем оптимизации последовательности подачи учебного материала;
- возможность автоматизации процессов мониторинга прогресса учащих-

ся и рекомендаций по дальнейшим этапам обучения.

Настоящее исследование представляет интерес для специалистов, занятых вопросами совершенствования технологий дистанционного обучения, а также разработчиков электронных образовательных платформ, стремящихся интегрировать соответствующие аналитические модели в свои системы с целью оптимизации деятельности преподавателей и повышения эффективности учебного процесса.

### **Список литературы**

1. Li D., Xing W. A comparative study on sustainable development of online education platforms at home and abroad since the twenty-first century based on big data analysis // Education and Information Technologies. 2025. P. 1-22. <https://doi.org/10.1007/s10639-025-13400-3>.
2. Бурукина И.П. LX Design в разработке онлайн курсов: принципы, методы и практика // Педагогическая информатика. 2025. № 1. С. 117-123.
3. Rajabalee Y.B., Santally M.I. Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy analysis // Education and Information Technologies. 2021. №3. P. 2623-2656. <https://doi.org/10.1007/s10639-020-10375-1>.
4. Kedia P., Mishra L. Exploring the factors influencing the effectiveness of online learning: a study on college students // Social Sciences & Humanities Open. 2023. №1. P. 100559. <https://doi.org/10.1016/J.SSAHO.2023.100559>.
5. Lalitha T.B., Sreeja P.S. Personalised self-directed learning recommendation system // Procedia Computer Science. 2020. Vol.171. P. 583-592. <https://doi.org/10.1016/j.procs.2020.04.063>.
6. Alper A., Okyay S., Nihat A. Hybrid course recommendation system design for a real-time student automation application // Avrupa Bilim ve Teknoloji Dergisi. 2021. T.26. P. 85-90. <https://doi.org/10.31590/ejosat.944596>.
7. Когнитивные технологии в принятии управленческих решений / С.П. Серегин, А.Р. Федорова, Ю.А. Халин, А.И. Катыхин // Известия Юго-Западного государственного университета. 2024; 28(4): 57-66. <https://doi.org/10.21869/2223-1560-2024-28-4-57-66>.
8. Thaipisutikul T., Shih T.K., Enkhbat A., Aditya W. Exploiting long- and short-term preferences for deep context-aware recommendations // IEEE Transactions on Computational Social Systems. 2021. Vol. 9, №4. P. 1237-1248.

9. Лапенок М.В., Макеева В.В. Формирование индивидуальной траектории обучения в информационно-образовательной среде школы // Педагогическое образование в России. 2016. №7. С. 37-43.
10. Троицкая Е.А. Психологическая устойчивость и субъективное благополучие личности как ресурсы для проявления эмпатии // Вестник Московского государственного лингвистического университета. 2014. №. 7. С. 46-59.
11. Васильченко С. Х. Функциональные особенности формирования персональной образовательной среды как средства индивидуализации обучения на основе информационных технологий // Информатика и образование. 2010. №. 12. С. 104-108.
12. Wang S., Wang F., Zhu Z. Artificial intelligence in education: A systematic literature review // Expert Systems with Applications. 2024. Р. 124-167. <https://doi.org/10.1016/j.eswa.2024.124167>.
13. Подколзин М.М. Интеллектуальная система адаптивного обучения на основе нейронных сетей для персонализации образовательных траекторий студентов российских вузов // Информатика и образование. 2024. 39(6). С. 65–81. <https://doi.org/10.32517/0234-0453-2024-39-6-65-81>.
14. Варнухов А.Ю. Скрытая марковская модель: метод построения модели бизнес-процесса // Бизнес-информатика. 2024. Т. 18, №. 3. С. 41-55. DOI: 10.17323/2587-814X.2024.3.41.55.
15. Земсков А.В. Аналитический подход к исследованию неоднородных марковских цепей с кусочно-постоянными изменениями переходных вероятностей // Известия высших учебных заведений. Приборостроение. 2024. Т. 67, №. 8. С. 657-669. <https://doi.org/10.17586/0021-3454-2024-67-8-657-669>
16. Ravari A., Ghoreishi S.F., Imani M. Optimal inference of hidden Markov models through expert-acquired data // IEEE Transactions on Artificial Intelligence. 2024. Vol. 5, № 8. P. 3985-4000. <https://doi.org/10.1109/TAI.2024.3358261>.
17. Pan W. Research on pig sound recognition based on deep neural network and hidden Markov models // Sensors. 2024. Vol. 24, № 4. P. 1269. <https://doi.org/10.3390/s24041269>
18. Jithendar T.R., Devi M.T., Saritha G. Determination of viterbi path for 3 hidden and 5 observable states using hidden Markov model // Reliability: Theory & Applications. 2024. Vol. 19, №. 2 (78). P. 509-515.
19. Zhu K. Physics-informed hidden markov model for tool wear monitoring // Journal of Manufacturing Systems. 2024. Vol. 72. P. 308-322. <https://doi.org/10.1016/j.jmsy.2023.11.003>.
20. Харин Ю.С. Критерий отношения правдоподобия проверки сложных гипотез суммарной равномерности двоичных последовательностей // Вероятностные методы в дискретной математике. 2024. Т. 1. С. 122.
21. Wong Z.Y. et al. Student engagement and its association with academic achievement and subjective well-being: A systematic review and meta-analysis // Journal of Educational Psychology. 2024. №. 1. P. 48–75. <https://doi.org/10.1037/edu0000833>.

22. Zhao J. et al. Choosing the number of factors in factor analysis with incomplete data via a novel hierarchical Bayesian information criterion // Advances in Data Analysis and Classification. 2024. P. 1-27. <https://doi.org/10.1007/s11634-024-00582-w>.
23. Yang F., Balakrishnan S., Wainwright M.J. Statistical and computational guarantees for the Baum-Welch algorithm // Journal of Machine Learning Research. 2017. Vol. 18, № 125. P. 1-53.
24. Wang C., Li K., He X. Network risk assessment based on Baum Welch algorithm and HMM // Mobile Networks and Applications. 2021. Vol. 26, №. 4. P. 1630-1637. <https://doi.org/10.1007/s11036-019-01500-7>.
25. Anderson K.S. Python: 2023 project update // Journal of Open Source Software. 2023. Vol. 8, №. 92. P. 5994.
26. Библиотеки python для начинающих / В.В. Сааков, Л.Х. Кучмезова, А.А. Дзамихова, З.Х. Шаущева // Молодой учёный: сборник статей III Международной научно-практической конференции. Пенза, 2023. С. 19-21.
27. Сальников А.В. Верификация и валидация компьютерных моделей // Известия высших учебных заведений. Машиностроение. 2022. №. 9 (750). С. 100-115. <https://doi:10.18698/0536-1044-2022-9-100-115>.

### References

1. Li D., Xing W. A comparative study on sustainable development of online education platforms at home and abroad since the twenty-first century based on big data analysis. *Education and Information Technologies*. 2025. P. 1-22. <https://doi.org/10.1007/s10639-025-13400-3>.
2. Burukina I.P. LX Design in the development of online courses: principles, methods and practice. *Pedagogicheskaya informatika = Pedagogical informatics*. 2025; (1): 117-123. (In Russ.).
3. Rajabalee Y.B., Santally M.I. Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy analysis. *Education and Information Technologies*. 2021; (3): 2623-2656. <https://doi.org/10.1007/s10639-020-10375-1>.
4. Kedia P., Mishra L. Exploring the factors influencing the effectiveness of online learning: a study on college students. *Social Sciences & Humanities Open*. 2023; (1): 100559. <https://doi.org/10.1016/J.SSAHO.2023.100559>.
5. Lalitha T.B., Sreeja P.S. Personalised self-directed learning recommendation system. *Procedia Computer Science*. 2020; 171: 583-592. <https://doi.org/10.1016/j.procs.2020.04.063>.
6. Alper A., Okyay S., Nihat A. Hybrid course recommendation system design for a real-time student automation application. *Avrupa Bilim ve Teknoloji Dergisi*. 2021; 26: 85-90. <https://doi.org/10.31590/ejosat.944596>.
7. Seregin S. P., Fedorov A. R., Khalin Y. A., Katykhin A. I. Cognitive technologies in management decision-making. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta =*

8. Thaipisutikul T., Shih T.K., Enkhabat A., Aditya W. Exploiting long- and short-term preferences for deep context-aware recommendations. *IEEE Transactions on Computational Social Systems.* 2021; 9(4): 1237-1248.
9. Lapenok M.V., Makeeva V.V. Formation of an individual learning trajectory in the information and educational environment of the school. *Pedagogicheskoe obrazovanie v Rossii = Pedagogical education in Russia.* 2016; (7): 37-43. (In Russ.)
10. Troitskaya E.A. Psychological stability and subjective well-being of an individual as resources for the manifestation of empathy. *Vestnik Moskovskogo gosudarstvennogo lingvisticheskogo universiteta = Bulletin of the Moscow State Linguistic University.* 2014; (7): 46-59. (In Russ.)
11. Vasilchenko S.Kh. Functional features of the formation of a personal educational environment as a means of individualization of learning based on information technology. *Informatika i obrazovanie = Computer Science and Education.* 2010: (12): 104-108. (In Russ.)
12. Wang S., Wang F., Zhu Z. Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications.* 2024. P. 124-167. <https://doi.org/10.1016/j.eswa.2024.124167>.
13. Podkolzin M.M. Intelligent adaptive learning system based on neural networks for personalization of educational trajectories of students of Russian universities. *Informatika i obrazovanie = Computer Science and Education.* 2024; (39): 65–81. (In Russ.). <https://doi.org/10.32517/0234-0453-2024-39-6-65-81>.
14. Varnukhov A.Yu. Hidden Markov Model: Method for Constructing a Business Process Model. *Biznes-informatika = Business Informatics.* 2024; 18(3): 41–55. (In Russ.). <https://doi.org/10.17323/2587-814X.2024.3.41.55>.
15. Zemskov A.V. Analytical approach to the study of non-homogeneous Markov chains with piecewise constant changes in transition probabilities. *Izvestiya vysshikh uchebnykh zavedenii. Priborostroenie = Bulletin of higher educational institutions. Instrument engineering.* 2024; 67(8): 657-669. (In Russ.). <https://doi.org/10.17586/0021-3454-2024-67-8-657-669>.
16. Ravari A., Ghoreishi S.F., Imani M. Optimal inference of hidden Markov models through expert-acquired data. *IEEE Transactions on Artificial Intelligence.* 2024; 5 (8): 3985-4000. <https://doi.org/10.1109/TAI.2024.3358261>.
17. Pan W. Research on pig sound recognition based on deep neural network and hidden Markov models. *Sensors.* 2024; 24(4): 1269. <https://doi.org/10.3390/s24041269>
18. Jithendar T.R., Devi M.T., Saritha G. Determination of viterbi path for 3 hidden and 5 observable states using hidden Markov model. *Reliability: Theory & Applications.* 2024; 19(2): 509-515.

19. Zhu K. Physics-informed hidden markov model for tool wear monitoring. *Journal of Manufacturing Systems*. 2024; 72: 308-322. <https://doi.org/10.1016/j.jmsy.2023.11.003>.
20. Kharin Yu.S. Likelihood ratio criterion for testing complex hypotheses of s-dimensional uniformity of binary sequences. *Veroyatnostnye metody v diskretnoi matematike = Probabilistic methods in discrete mathematics*. 2024; 1: 122. (In Russ.).
21. Wong Z.Y., et al. Student engagement and its association with academic achievement and subjective well-being: A systematic review and meta-analysis. *Journal of Educational Psychology*. 2024; (1): 48–75. <https://doi.org/10.1037/edu0000833>.
22. Zhao J., et al. Choosing the number of factors in factor analysis with incomplete data via a novel hierarchical Bayesian information criterion. *Advances in Data Analysis and Classification*. 2024: 1-27. <https://doi.org/10.1007/s11634-024-00582-w>.
23. Yang F., Balakrishnan S., Wainwright M.J. Statistical and computational guarantees for the Baum-Welch algorithm. *Journal of Machine Learning Research*. 2017; 18 (125): 1-53.
24. Wang C., Li K., He X. Network risk assessment based on Baum Welch algorithm and HMM. *Mobile Networks and Applications*. 2021; 26 (4): 1630-1637. <https://doi.org/10.1007/s11036-019-01500-7>.
25. Anderson K.S. Python: 2023 project update. *Journal of Open Source Software*. 2023; 8 (92): 5994 p.
26. Saakov V.V., Kuchmezova L.Kh., Dzamikhova A.A., Shaushcheva Z.Kh. Python libraries for beginners. In: *Molodoi uchenyi: sbornik statei III Mezhdunarodnoi nauchno-prakticheskoi konferentsii = Young scientist: collection of articles of the III International Scientific and Practical Conference*. Penza, 2023. P. 19-21. (In Russ.)
27. Salnikov A.V. Verification and validation of computer models. *Izvestiya vysshikh uchebnykh zavedenii. Mashinostroenie = News of higher educational institutions. Mechanical engineering*. 2022; (9): 100-115. (In Russ.). <https://doi.org/10.18698/0536-1044-2022-9-100-115>.

---

## Информация об авторах / Information about the Authors

**Бурукина Ирина Петровна**, кандидат технических наук, доцент, заведующий кафедрой "Системы автоматизированного проектирования", Пензенский государственный университет, г. Пенза, Российская Федерация, e-mail: burukinairina@gmail.com, ORCID: <http://orcid.org/0009-0006-1953-2914>

**Горшенин Лев Николаевич**, аспирант, Пензенский государственный университет, г. Пенза, Российская Федерация, e-mail: burukinairina@gmail.com

**Irina P. Burukina**, Cand. of Sci. (Engineering), Associate Professor, Head of the Computer-Aided Design Systems Department, Penza State University, Penza, Russian Federation, e-mail: burukinairina@gmail.com, ORCID: <http://orcid.org/0009-0006-1953-2914>

**Lev N. Gorshenin**, Post-Graduate Student, Penza State University, Penza, Russian Federation, e-mail: burukinairina@gmail.com,