

ОБЗОР МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ МОДЕЛИРОВАНИЯ ПРОФИЛЯ СТУДЕНТОВ

Картанова А.Дж., Максатбеков Д.М.

КГТУ им. И. Раззакова, Бишкек, КР, a.kartanova@mail.ru, maksatbekov.dastan@gmail.com

***Аннотация.** В работе рассмотрены возможности и преимущества использования электронных образовательных ресурсов, проведен анализ информационной системы оценки психологического портрета студента, системы предоставляющей возможность анализировать академические данные, поведенческие и психологические характеристики студентов. Проведен обзор методов машинного обучения для выбора наиболее подходящих алгоритмов и рассмотрены особенности приложения для моделирования профиля студентов на основе методов машинного обучения.*

***Ключевые слова:** модель профиля студента; анализ; ЭОР; методы машинного обучения; наборы данных.*

ОКУУЧУЛАРДЫН ПРОФИЛДЕРИН МОДЕЛДӨӨ ҮЧҮН МАШИНАЛЫК ҮЙРӨТМӨ БОЮНЧА КЫСКАЧА КӨРГӨЗМӨ

Картанова А.Дж., Максатбеков Д.М.

И.Раззаков атындагы КМТУ, Бишкек, КР, a.kartanova@mail.ru, maksatbekov.dastan@gmail.com

***Аннотация.** Документте электрондук билим берүү ресурстарын колдонуунун мүмкүнчүлүктөрү жана артыкчылыктары каралат, студенттин психологиялык портретине баа берүүнүн маалымат системасы, академиялык маалыматтарды, окуучулардын жүрүм-турум жана психологиялык өзгөчөлүктөрүн талдоо мүмкүнчүлүгүн берген система талданат. Эң ылайыктуу алгоритмдерди тандоо үчүн машинаны үйрөнүү ыкмаларын карап чыгуу жүргүзүлдү жана машинаны үйрөнүү методдорунун негизинде окуучулардын профилин моделдөө үчүн колдонмонун өзгөчөлүктөрү каралды.*

***Ачык сөздөр:** студенттин профилинин модели; анализ; ЭОР; машинаны үйрөнүү ыкмалары; маалымат топтому.*

OVERVIEW OF MACHINE LEARNING TECHNIQUES FOR MODELING STUDENT PROFILE

Kartanova A.Dzh., Maksatbekov D.M.

KSTU n. a. I. Razzakov, Bishkek, Kyrgyz Republic a.kartanova@mail.ru, maksatbekov.dastan@gmail.com

***Annotation.** The possibilities and advantages of using electronic educational resources, an analysis of the information system for assessing the psychological portrait of a student, a system that provides an opportunity to analyze academic data, behavioral and psychological characteristics of students are considered in the work. A review of machine learning methods was carried out to select the most appropriate algorithms and the features of an application for modeling the profile of students based on machine learning methods were considered.*

***Keywords:** student profile model; analysis; EER; machine learning methods; data sets.*

Введение. В настоящее время использование электронных образовательных ресурсов (ЭОР) является крайне актуальным, поскольку электронное обучение

предоставляет множество возможностей как для обучаемых, так и для преподавателей. Во-первых, с помощью электронных образовательных ресурсов обучаемые и преподаватели могут получать широкий доступ к большому количеству информации, включая учебные материалы, научные статьи и другие источники к информации. Во-вторых, ЭОР позволяют обучаемым и преподавателям работать в своем собственном темпе и в удобное для них время, что особенно удобно для студентов, занятых работой или семейными обязанностями. В-третьих, ЭОР разрабатываются с использованием интерактивных элементов, таких как видеоуроки, тесты и игры, которые помогают учащимся лучше усваивать материал и принимать активное участие в учебном процессе. В-четвертых, использование ЭОР позволяет снизить затраты на учебные материалы и книги, что особенно важно для студентов, не имеющих возможности покупать дорогостоящие учебники. В-пятых, с помощью ЭОР преподаватели могут адаптировать, настраивать материал под потребности каждого конкретного обучаемого, что позволит им обучаться в соответствии со своими индивидуальными потребностями и темпом. Таким образом, ЭОР имеют огромный потенциал для обучения и развития учащихся, а также для повышения качества образования в целом [1].

Рассмотрим некоторые из наиболее основных видов ЭОР, это видеоуроки лекций, презентаций, демонстрационные материалы, позволяющие обучаемым просматривать уроки в записи и изучать материал в удобное время, интерактивные задания: тесты, викторины, игры и другие, которые помогают обучаемым усваивать материал и проверять свои знания, электронные учебники содержащие тексты, изображения, анимации и другие интерактивные элементы, которые помогают учащимся лучше понимать и запоминать учебный материал. ЭОР могут быть представлены как веб-сайты, с содержанием вышеперечисленных учебных материалов, а также могут быть представлены как социальные сети, которые используются в качестве платформы для общения и обсуждения заданий и проблем, для обмена знаниями между обучаемыми и преподавателями, как виртуальные классы позволяют обучаемым и преподавателям общаться в режиме реального времени в онлайн-среде, что включает видеоконференции, чаты, форумы и другие средства общения.

Применение электронных образовательных ресурсов в обучении преследует несколько целей с точки зрения педагогики и образования, это индивидуализация обучения, разнообразие учебных материалов, повышение эффективности обучения, доступность образования, стимулирование самостоятельности и самоорганизации.

Индивидуальный подход к обучению означает, что преподаватель учитывает индивидуальные потребности, интересы, способности и стиль обучения каждого студента. Например, рассмотрим некоторые методы индивидуального обучения,

которые включают оценку индивидуальных потребностей и целей, использование различных методов обучения, соответствующие разным стилям обучения студентов, адаптация учебных материалов в соответствии с индивидуальными потребностями каждого студента, предоставляя дополнительные материалы или объясняя материал более подробно, организация индивидуальных консультаций для обсуждения их индивидуальных потребностей и целей, использование технологий, такие как электронные образовательные ресурсы и онлайн-курсы, чтобы предоставить студентам возможность изучать материал в своем темпе и на своем уровне, формирование групп по уровню знаний и навыков, чтобы обеспечить оптимальные условия для обучения каждого студента.

Наряду с образовательными ЭОР существуют информационные системы поддержки принятия решений, которые опираясь на многолетний опыт функционирования при использовании ЭОР формируют огромную базу данных о достижениях и результатах обучения обучаемых. Такие информационные системы предоставляют возможности создавать электронные профили обучаемых, собирать данные по итогам обучения, а также на основе проведения опросов и анкетирования в ходе обучения на платформе, и в дальнейшем проводить анализ и оценку академических данных, поведенческих и психологических характеристик обучаемых.

Постановка задачи. Целью настоящей работы является разработка веб-приложения по оценке психологического портрета студента для выработки преподавателем эффективных методов обучения, индивидуального подхода к обучаемому для его поддержки и в целом для мотивации обучаемого. Студентам, в свою очередь, веб-приложение позволит определить его сильные и слабые стороны, а также, какие аспекты его психологического состояния нужно улучшить для достижения наилучших результатов в обучении и на практике. Как известно, от психологического состояния студента зависит его достижения в учебе, а также его работоспособность и эффективность в профессиональной деятельности.

Проектируемое веб-приложение представляет собой систему поддержки принятия решения для определения и оценки психологического портрета студента, позволяющая формировать профиль с личной информацией о студенте, его социальном положении, академической успеваемости, навыками, личностными качествами, предпочитаемыми стилями обучения, его поведением в Интернете, а также выявленные в ходе анкетирования потребностями и пожеланиями обучаемых.

Реализация веб-приложения позволит получать визуализацию данных анализа и на основе оценки выдавать рекомендации по каждому студенту, для дальнейших мероприятий по индивидуализации методов и подходов к обучению обучаемого. Таким образом, разрабатываемое веб-приложение будет предназначено для

моделирования профиля студентов и представляет собой мощный инструмент для анализа и поддержки образовательного процесса, способное адаптироваться к индивидуальным потребностям студентов.

Многие системные цели электронных образовательных ресурсов, такие как электронная рекомендация, электронная ориентация, электронный набор персонала и прогнозирование отсева, в основном основаны на профиле поддержки принятия решений. Машинное обучение играет важную роль в этом контексте, проведем обзор некоторых исследований с целью обзора методов машинного обучения, как классификации, прогнозирования или кластеризации.

Методы машинного обучения. Использование методов машинного обучения для образовательных предложений или интеллектуального анализа образовательных данных является новой областью, направленной на разработку методов изучения данных из вычислительных образовательных учреждений и обнаружение значимых закономерностей. Сохраненные данные, такие как виртуальные курсы, файл журнала электронного обучения, демографические и академические данные студентов, информация о приеме/регистрации могут быть полезны для алгоритмов машинного обучения.

В работе [2] авторы применили семь различных алгоритмов классификации контролируемого обучения, используя личностные характеристики обучаемого в качестве целевых значений, построили прогноз для определения психологического портрета обучаемого, а также построили модель обучаемого, используя особенности данных обучения, в результате разработали автоматический классификатор, который ненавязчиво предсказывает портрет обучаемого на основе его достижений в системе онлайн-обучения.

В статье [3] авторы представляют сравнительное исследование различных алгоритмов бустинга, которые успешно используются во многих областях и для многих целей. Кроме того, авторы применили методы выбора признаков Fisher Score, Information Gain в сочетании с рекурсивным устранением признаков для повышения производительности задач предварительной обработки и моделей. Используя многометочный набор данных, предсказывающий класс успеваемости учащихся по математике, результаты этой статьи показывают, что алгоритм Light Gradient Boosting Machine (LightGBM) достиг наилучшей производительности при использовании усиления информации с помощью метода рекурсивного исключения признаков по сравнению с другими алгоритмами бустинга.

Идея прогнозирования предпочтений пользователей и соответствующей адаптации платформы электронного обучения является центральным моментом статьи [4]. В частности, данная работа начинается с основных требований к продвинутой

системе электронного обучения, объясняет, как пользователь ориентируется в такой системе, представляет архитектуру соответствующей системы электронного обучения и описывает ее основные компоненты. Исследование сосредоточено на компоненте User Model, его роли в системе электронного обучения и составляющих его параметрах. В этом контексте байесовские сети используются в качестве инструмента для кодирования, изучения и обоснования вероятностных отношений с целью эффективного прогнозирования предпочтений пользователей. В поддержку этого видения представлены четыре различных сценария, чтобы проверить, как байесовские сети применяются в области электронного обучения.

В статье [5] представляем пример прогнозирования оценок учащихся, где оценки в небольшом количестве письменных заданий вместе с ключевыми демографическими характеристиками обучаемых составили обучающий набор для регрессионного метода с целью прогнозирования успеваемости учащихся. В результате был построен прототип программного средства поддержки для репетиторов.

В исследовании [6] для изучения поведения учащихся в процессе обучения и прогнозирования их результатов обучения, применяются технологии машинного обучения, при анализе данных интерактивных сред обучения использовались различные методы классификации машинного обучения, викторины и журналы систем программирования, в результате обнаружено, что характеристики обучения учащихся коррелировали с их успеваемостью, когда они сталкивались с аналогичной практикой программирования. В этом исследовании использовались алгоритмы случайного леса (RF), машины опорных векторов (SVM), логистической регрессии (LR) и нейронной сети (NN), чтобы предсказать, вовремя ли студенты сдадут экзамен на курс. Среди них алгоритм нейронной сети (NN) показал лучшие результаты прогнозирования.

В работе [7] были определены мотивационные профили студентов с помощью лично-ориентированного подхода и проанализировать связи с удовлетворением их основных психологических потребностей, стратегиями обучения, успеваемостью, самооценкой и жизненной силой. В общей сложности 924 студента из Университета Сан-Себастьяна (Чили) приняли участие в этом кросс-секционном корреляционном исследовании весной 2016 года. Данные были собраны с помощью опроса по 5 видам анкет в дополнение к успеваемости студентов. Были рассчитаны альфа Кронбаха, показатели корреляции, описательная статистика набора данных показана в таблице 1.

Таблица 1.

Описательная статистика набора данных

Характеристика (Features)	Среднее значение (Mean)	Стандартное отклонение (SD)	Минимальное значение (Min)	Максимальное значение (Max)
Внутренняя мотивация	4,98	0,61	2,17	6,58
Внешняя мотивация	5,25	0,75	2,42	7,00
Автономность	5,01	0,89	2,00	6,25
Коммуникабельность	4,46	0,90	1,5	6,25
Компетенция	4,77	0,84	2,25	6,25
Чувство собственного достоинства	4,17	0,17	1,75	7,00
Стратегия углубленного обучения	4,11	0,72	2,00	6,25
Стратегия поверхностного обучения	3,32	0,79	1,50	6,25
Учебный год	3,24	1,48	1	6
Возраст	22,83	3,36	18,00	44,00
Академическая успеваемость	4,72	0,54	2,92	6,40

Для выявления различных мотивационных профилей был проведен кластерный анализ k-средних с внутренней и контролируемой мотивацией. Мотивационные профили студентов характеризовались разной степенью внутренней и управляемой мотивации. Группы с высокой внутренней мотивацией продемонстрировали более высокое восприятие своих базовых психологических характеристик, большую склонность к углубленному, а не поверхностному обучению, лучшую академическую успеваемость и более высокие баллы по самооценке и жизненной силе, чем группы с низкой внутренней мотивацией, независимо от степени вовлеченности. Получили результаты, что студенты с высоким профилем внутренней мотивации, независимо от их контролируемых показателей мотивации, сообщили о лучших характеристиках обучения. Поэтому особое внимание следует уделять мотивационным профилям студентов, поскольку качество мотивации может служить основой для вмешательств, направленных на поддержку их академических успехов и благополучия.

Исследование [8] посвящено разработке моделей машинного обучения для прогнозирования успеваемости студентов и стратегий обучения. При построении моделей использовались ключевые атрибуты обучения (внутренние, внешние, автономия, связанность, компетентность и самооценка), необходимые для процесса обучения студентов. Для анализа влияния ключевых атрибутов на академическую успеваемость студентов и стратегию обучения использовали пакет Scikit-learn в Python для создания пяти моделей машинного обучения, таких как дерево решений, k-

ближайший сосед, случайный лес, линейная/логистическая регрессия и машина опорных векторов, как для задач регрессии, так и для задач классификации. Модели были обучены, оценены, и проверены на точность с использованием данных 924 студентов-стоматологов университетов, собранных чилийскими авторами [6] в рамках количественного исследования, распределение студентов по возрасту и по годам обучения показаны на рисунках 1-2.

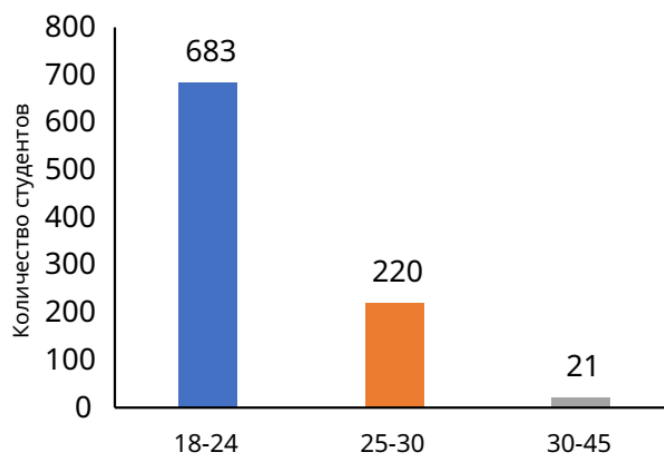


Рис. 1. Распределение студентов по возрасту

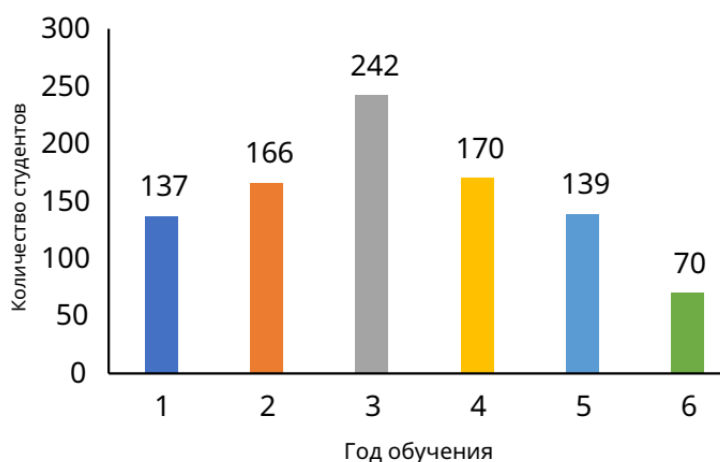


Рис. 2. Распределение студентов по годам обучения

Сравнительный анализ моделей показал, что древовидные модели, такие как случайный лес (с точностью предсказания 94,9%) и дерево решений, показывают наилучшие результаты по сравнению с линейным, опорным вектором и k-ближайшими соседями. Модели, построенные в этом исследовании, можно использовать для прогнозирования успеваемости учащихся и стратегии обучения, чтобы можно было принять соответствующие меры для улучшения успеваемости обучаемых. Это также предоставляет информацию для разработчиков онлайн-образовательных систем о

динамической роли включения стратегий, которые могут учитывать различные аспекты мотивации в содействии обучению.

Прежде всего моделирование профиля студента основано на наборе данных характерных для студента, в качестве которых могут выступать физиологические данные, или данные о взаимодействии с системой, которые показывают многообещающие результаты, чем при использовании их по отдельности, поэтому многие исследователи применяют мультимодальный подход (комбинацию двух источников данных) при создании моделей профилей студента для повышенной точности.

При мультимодальном подходе в исследованиях объединялись журналы отслеживания взгляда и взаимодействия с системой, в частности, такие характеристики, как отслеживание взгляда, мимика, выражение лица, ЭЭГ, браслет движения мыши и нажатия клавиш. Для создания мультимодальных моделей для прогнозирования аффективных состояний, эмоций, мотивации и вовлеченности студентов объединяли журналы позы и взаимодействия с системой. Некоторые из этих исследований показали, что мультимодальный метод работает лучше, чем модель, основанная на одном подходе. Многие исследования в основном сосредоточены на автоматическом моделировании шести основных характеристик студентов. Характеристики включают аффективное состояние обучаемого, эмоции, вовлеченность, стиль обучения, мотивацию и личность. как показано на рисунке 3.

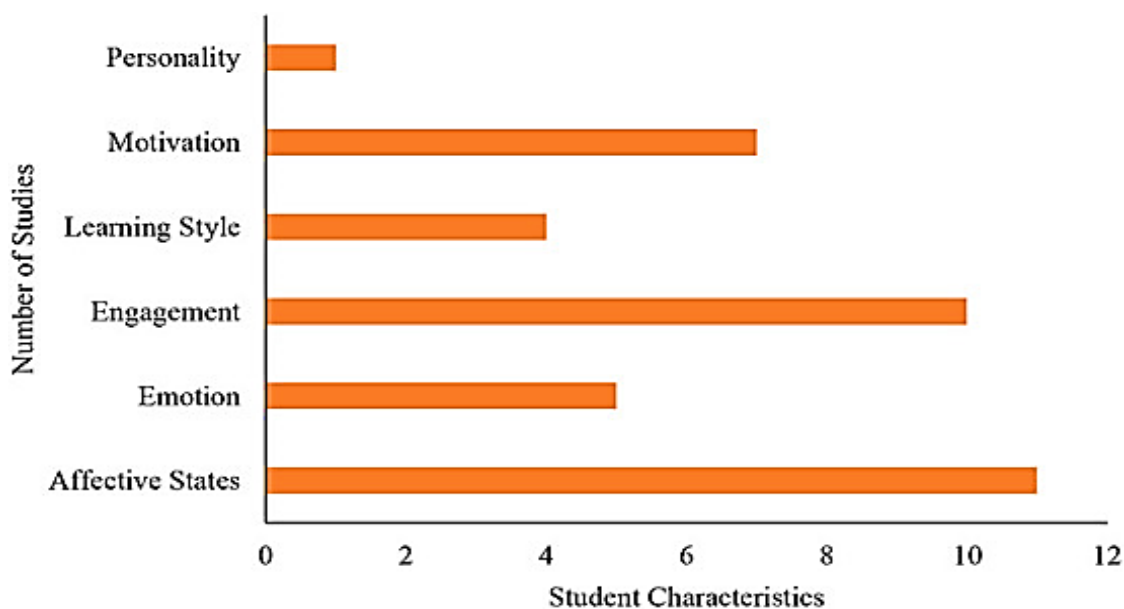


Рис.3. Характеристики студентов

Среди 38 исследований, в которых моделировались различные характеристики студентов с использованием методов машинного обучения и данных, основанных на взаимодействии и/или физиологических данных, аффективные состояния возглавляли

список с 29% из всех исследований, за которыми следовала вовлеченность с 26%, мотивация занимает третье место с 18%, эмоции с 13% занимают четвертое место, 10% - стиль обучения и 2,5% личные качества.

В целом охарактеризованные выше алгоритмы машинного обучения наряду с традиционными методами выступают надежными прикладными инструментами исследования инновационных процессов, так как анализ актуальной информации составляет фундамент для принятия обоснованных решений.

На основе моделирования профиля студента, разработанное приложение будет предоставлять предсказания и рекомендации по каждому студенту. Это позволит вести мониторинг и отслеживание образовательного процесса по каждому студенту во времени, тем самым позволяет измерять эффективность предоставляемой поддержки и проводить соответствующую коррекцию в образовательном процессе и адаптировать методику обучения под индивидуальные потребности и т.п.

А также приложение будет предоставлять возможность отправки уведомлений и сообщений между преподавателями, кураторами и студентами. Это способствует более эффективной коммуникации и обмену информацией о состоянии и прогрессе каждого студента. Приложение может быть интегрировано с существующими системами управления образованием, такими как электронные журналы, системы оценок и расписания. Это обеспечивает более полное представление о студентах и упрощает доступ к информации.

В разработке инновационного приложения для моделирования профиля студентов будут использованы следующие инструментальные средства: язык программирования Java, Python, фреймворки машинного обучения TensorFlow, PyTorch, scikit-learn и другие. Эти инструменты предоставляют готовые алгоритмы и функции для обработки данных и построения моделей машинного обучения. Для хранения данных и управления данными применим систему управления данными MySQL.

Для создания пользовательского интерфейса и веб-серверной части приложения будет использован веб-фреймворк, такой как Django, которые облегчают разработку и управление веб-приложением, предоставляя инструменты для маршрутизации, обработки запросов и создания пользовательского интерфейса. Для создания наглядных графиков и диаграмм, представляющих результаты анализа данных, могут использоваться библиотеки визуализации данных, такие как Matplotlib, Plotly или D3.js. Эти инструменты позволяют создавать графические представления данных для более наглядного и понятного отображения результатов. Для развертывания приложения и хранения данных можно использовать облачные сервисы, такие как Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP) и другие.

Эти сервисы предоставляют инфраструктуру для развертывания приложений в облаке, обеспечивают масштабируемость и высокую доступность.

Выводы. Таким образом, проектируемое приложение предлагает простой и удобный интерфейс, позволяющий студентам заполнять анкеты и проходить тесты онлайн, а также оно автоматически собирает разнообразные данные из различных источников, таких как академические показатели, социальные аспекты и другие. С помощью мощных алгоритмов машинного обучения, приложение анализирует собранные данные и моделирует психологический портрет студента. Путем анализа статистических данных, предоставленных системой, можно получить ценную информацию о студентах, их успехах, потребностях и возможностях для дальнейшего развития.

Предложенное решение проблемы моделирования профиля студентов, основанное на системе определения психологического портрета студента, обладает потенциалом для повышения качества образовательного процесса путем индивидуальной поддержки студентов. Это позволит создать полную и объективную картину о каждом студенте и выявлять общие характеристики, определять типы личности, мотивационные особенности и стили обучения для каждого обучаемого.

СПИСОК ЛИТЕРАТУРЫ

1. **Гюльбякова Х.Н., Масловская Е.А.** Электронная форма обучения: особенности и перспективы // *Современные проблемы науки и образования*. – 2018. – № 4.; URL: <https://science-education.ru/ru/article/view?id=27812> (дата обращения: 21.05.2023).
2. **Abyaa, A., Idrissi, M. K., & Bennani, S.** (2018, October). Predicting the learner's personality from educational data using supervised learning. In *Proceedings of the 12th international conference on intelligent systems: theories and applications* (pp. 1-7).
3. **Hamim, T., Benabbou, F., & Sael, N.** (2022). Student Profile Modeling Using Boosting Algorithms. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, 17(5), 1-13. <http://doi.org/10.4018/IJWLTT.20220901.0a4>
4. **Kritikou, Y., Demestichas, P., Adamopoulou, E., Demestichas, K., Theologou, M., Paradia, M.** (2008). User profile modeling in the context of web-based learning management systems. *J. Netw. Comput. Appl.* 31(4), pp. 603–627.
5. **Kotsiantis, S. B.** (2012), 'Use of machine learning techniques for educational proposes: a decision support system for forecasting students' grades', *Artificial Intelligence Review*, vol. 37, no. 4, pp. 331–344, <https://doi.org/10.1007/s10462-011-9234-x>.
6. **Su Y-S, Lin Y-D and Liu T-Q** (2022) Applying machine learning technologies to explore students' learning features and performance prediction. *Front. Neurosci.* 16:1018005. doi: 10.3389/fnins.2022.1018005
7. **Orsini, C. A., Binnie, V. I., and Tricio, J. A.** (2018) Motivational Profiles and their Relationships with Basic Psychological Needs, Academic Performance, Study Strategies, Self-esteem, and Vitality in Dental Students in Chile," *J. Educ. Eval. Health Prof.*, vol. 15, p. 11.
8. **Orji, Fidelia & Vassileva, Julita.** (2022). Machine Learning Approach for Predicting Students Academic Performance and Study Strategies based on their Motivation. 10.48550/arXiv.2210.08186.