*УДК* [*004.89*](http://teacode.com/online/udc/00/004.891.html) Научная статья

*DOI: 10.35330****/****1991-6639-2023-2-112- 18-29*

EDN: GXEHAC

**Интеллектуальный анализ образовательных данных**

**для прогноза успеваемости студентов вуза**

**Н. А. Попова1, Е. С. Егорова2**

1 Пензенский государственный университет

440026, Россия, г. Пенза, ул. Красная, 40

2 Пензенский государственный технологический университет

440039, Россия, г. Пенза, проезд Байдукова/ул. Гагарина, 1а/11

***Аннотация.*** Прогресс в области интеллектуального анализа данных делает возможным использование образовательных данных для повышения качества образовательных процессов. В статье рассмотрены различные методы анализа данных об успеваемости студентов. Основное внимание уделено двум аспектам: во-первых, прогнозирование академических достижений студентов в конце четырехлетней учебной программы по программам бакалавриата; во-вторых, изучение типичных прогрессий учащихся и объединение их с результатами прогнозирования. При прогнозировании было использовано порядка 10 алгоритмов классификации. Предложен подход к улучшению производительности методов классификации, когда атрибуты классификаторов выбираются в процессе их обучения. Определены две важные группы учащихся – с низкими и высокими достижениями. Результаты показывают, что, сосредоточив внимание на небольшом количестве курсов, которые являются показателями особенно хорошей или плохой успеваемости, можно своевременно предупреждать и поддерживать студентов с низкой успеваемостью, а также давать советы и возможности студентам с высокой успеваемостью.

***Ключевые слова:*** анализ образовательных данных, дерево решений, кластеризация, прогнозирование, успеваемость, дистилляция

**СПИСОК ЛИТЕРАТУРЫ**

1. *Белоножко П. П., Карпенко А. П., Храмов Д. А.* Анализ образовательных данных: направления и перспективы применения // Интернет-журнал «Науковедение» Том 9. № 4 (2017). URL: http://naukovedenie.ru/PDF/15TVN417.pdf

2. *Мамонтова М. Ю.* Качество учебных достижений: оценка и прогноз на основе результатов критериально-ориентированного тестирования // Образование и наука. Известия   
УрО РАО. 2009. № 3(60). С. 18–26.

3. *Русаков С. В.,* *Накарякова Н. Н.* Прогнозирование успеваемости студентов первого курса с помощью дерева решений на основе их результатов сдачи ЕГЭ // Наука. Информатизация. Технологии. Образование: Материалы XI международной научно-практической конференции. Екатеринбург: Российский государственный профессионально-педагогический университет, 2018. С. 589–594.

4. *Фирстов В. Е.* Социометрические и информационные аспекты кластеризации обучаемого контингента при организации и оптимизации группового сотрудничества в учебном процессе в школе и вузе // Известия Саратовского университета. Новая серия. Серия: Философия. Психология. Педагогика. 2014. Т. 14. № 1. С. 110–118.

5. *Medvedev D., D’yakonov A*. New Properties of the Data Distillation Method When Working with Tabular Data // Conference proceedings “Analysis of Images, Social Networks and Texts”. *Lecture Notes in Computer Science*. Vol. 12602. Springer, Cham. 2021. <https://doi.org/> 10.1007/978-3-030-72610-2\_29

6. *Sucholutsky I., Schonlau M.* Soft-Label Dataset Distillation and Text Dataset Distillation // International Joint Conference on Neural Networks, Shenzhen, China, 2021. Pp. 1–8.   
DOI: 10.1109/IJCNN52387.2021.9533769.

7. *Никонорова М. Л.* Компьютерная модель решения задач классификации в программной среде Rapid Miner // Медицинское образование и профессиональное развитие. 2017. № 2–3(28–29). С. 24–33.

8. *Филяк П. Ю.,* *Виноградов М. А.* Применение Rapid miner и открытых сред как инструментов интеллектуального анализа данных для обеспечения безопасности // Информация и безопасность. 2017. Т. 20. № 4. С. 552–555.

9. *Maimon O., Rokach L.* Data Mining and Knowledge Discovery Handbook. Springer Science, Business Media, 2010. 1285 p. ISBN: 978-0-387-09822-7.

**Информация об авторах**

**Попова Наталия Александровна,** канд. техн. наук, доцент кафедры «Математическое обеспечение и применение ЭВМ», Пензенский государственный университет;

440026, Россия, г. Пенза, ул. Красная, 40;

popov.tasha@yandex.ru, ORCID: https://orcid.org/0000-0001-9713-4897

**Егорова Екатерина Сергеевна,** канд. экон. наук, доцент кафедры «Прикладная информатика», Пензенский государственный технологический университет;

440039, Россия, г. Пенза, проезд Байдукова/ул. Гагарина, 1а/11;

katepost@yandex.ru, ORCID: https://orcid.org/0000-0002-0816-0944