

Информационная система помощи выбора специальности для абитуриентов на основе нейронной сети

*М.А. Корнаухова, С.С. Основин, Е.В. Баскаков, Е.С. Савин,
Н.С. Полусмакова*

Волгоградский государственный университет, Волгоград

Аннотация: Важнейшей задачей приемной комиссии любого вуза является консультирование абитуриентов по вопросам поступления на имеющиеся в вузе специальности. В данной статье была разработана автоматизированная информационная система, которая на основе анкетирования абитуриента позволяет сформировать рекомендации при выборе специальности обучения, с учетом личных предпочтений абитуриента. В качестве метода для выбора предлагается использовать искусственные нейронные сети, а моделью нейронной сети являлся многослойный персептрон (MLP).

Ключевые слова: нейронная сеть, профориентационная работа, многослойный персептрон, обучение нейросетей, методы, автоматизированная информационная система, направления подготовки.

Одним из направлений деятельности вуза является профориентационная работа. В процессе приемной кампании вуз обязан обеспечить заполняемость всех бюджетных мест по образовательным программам. Для этого вузу необходимо найти оптимальный подход к целевой аудитории, с применением современных инструментов взаимодействия [1]. Одним из таких инструментов с абитуриентом может являться система помощи осуществления выбора направления подготовки обучающегося, в основе которой лежит концепция нейронных сетей [2].

Описание и принцип работы системы

Разработанная система предназначена для помощи абитуриентам, которые не могут определиться с выбором направления подготовки, в большей степени подходящего под их интересы.

Данная система была разработана и задействована в профориентационной работе университета, в котором реализуются десять

направлений подготовки в области IT-технологий по программам высшего образования, объединенных в пять укрупнённых групп специальностей.

Принцип работы системы представлен на рисунке 1. Обучение нейросети осуществляется на основе данных, полученных путем анкетирования студентов, обучающихся на рассматриваемых направлениях подготовки. Входные данные для обучения сети были составлены на основе опроса, состоящего из 20 пунктов, направленного на выявление образовательного потенциала абитуриентов и его профессиональных интересов, соответствующих определенным направлениям подготовки. Кроме того некоторые вопросы напрямую относятся к областям деятельности, связанной с определенными направлениями подготовки. Ответы на вопросы такого рода имеют больший вес входного параметра при обучении нейронной сети, что позволяет провести лучшую кластеризацию абитуриентов по специальностям [2]. Входные параметры оцениваются по десятибалльной шкале.

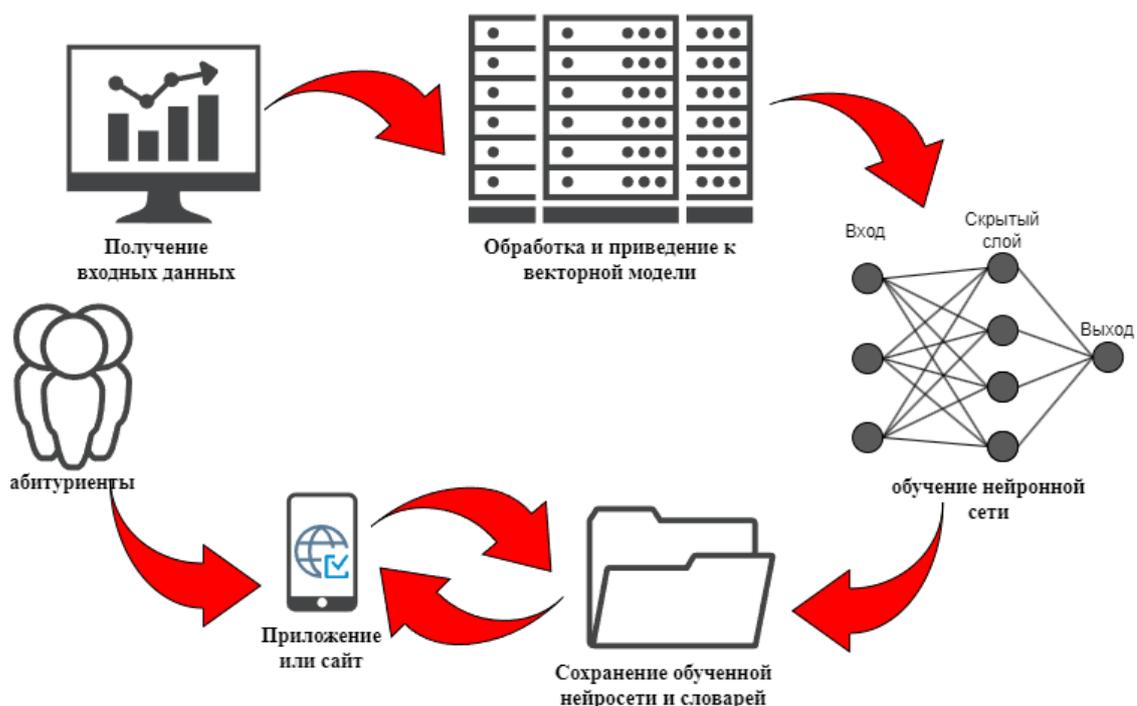


Рис. 1. – Принцип работы автоматизированной информационной системы

По результатам опроса и обработки данных была сформирована выборка по данным 309 абитуриентов, но для исключения стратегической близорукости и достижения более точного результата нейронной сети планируется провести этот опрос несколько раз, при новых наборах в вуз.

Выбор модели нейронной сети

В качестве модели нейронной сети был выбран многослойный персептрон (MLP), который представляют собой классические нейронные сети прямого распространения, которые применяются как в задачах регрессии, так и в задачах кластеризации и классификации [3]. MLP широко используются в различных областях, например, в медицине, в науках о растениях и окружающей среде [4 – 6].

На рисунке 2 представлена схема разработанной нейросети. Входной сигнал, который представляет собой вектор значений, сформированный на основе результатов анкетирования, подаётся на первый слой, после чего переходит по стадиям (скрытым слоям), итогом каждого из которых получается некоторый промежуточный результат.

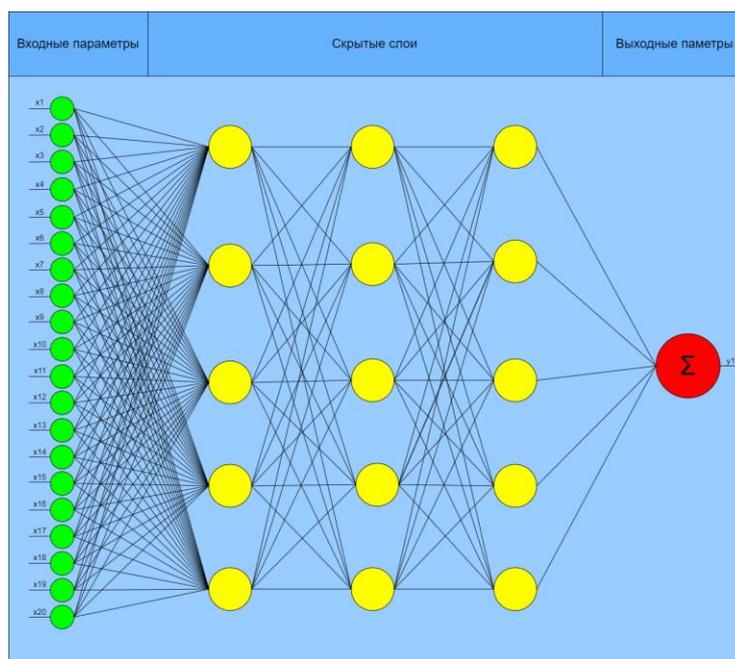


Рис. 2. – Схема многослойного персептрона

Персептрон относится к нейронным сетям, обучаемым с учителем. Процесс ее обучение заключается в корректировке синоптических весов таким способом, чтобы обучаемая система на входной набор параметров выдавала определенное выходное значение [7].

Обучение нейросети

Данная система была разработана на языке Python, выбор которого обуславливается тем, что данный язык программирования удобен для обработки и анализа больших массивов данных из-за наличия исчерпывающего количества средств для работы со структурами данных различной сложности, кроме того большую роль сыграла возросшая популярность этого языка для решения задач, связанных с искусственным интеллектом.

Перед началом обучения нейросети, необходимо разделить выборку на две части: первая используется для обучения, а другая предназначена для тестирования и оценки точности. Соотношение между объемами выборок равняется 70:30 соответственно.

Для данной системы был использован MLPClassifier – классификатор, предназначенный для многослойного персептрона [8]. Его особенность заключается в том, что он реализует многоуровневый алгоритм персептрона, который использует метод обратного распространения. Обучение происходит благодаря двум массивам: первый представляет собой входные векторы признаков, в данном случае это вектор, содержащий ответы на проводимый опрос. Второй содержит в себе метки классов – целевых значений нейронной сети. Для модификации весов, используются различные методы их оптимизации, которые могут быть определены одним из параметров при создании классификатора.

Обучение нейронной сети означает поиск комбинации параметров модели, которые будут минимизировать функцию потерь для набора входных и соответствующих им выходных данных. Функция потерь — это величина, которую требуется свести к минимуму в ходе обучения, поэтому она должна представлять собой меру успеха [9].

Оценка результатов

Рассмотрим оценку эффективности обучения с помощью среднеквадратичной погрешности (RMSE).

На рисунке 3 представлен график зависимости среднеквадратичной погрешности от количества нейронов в скрытом слое.

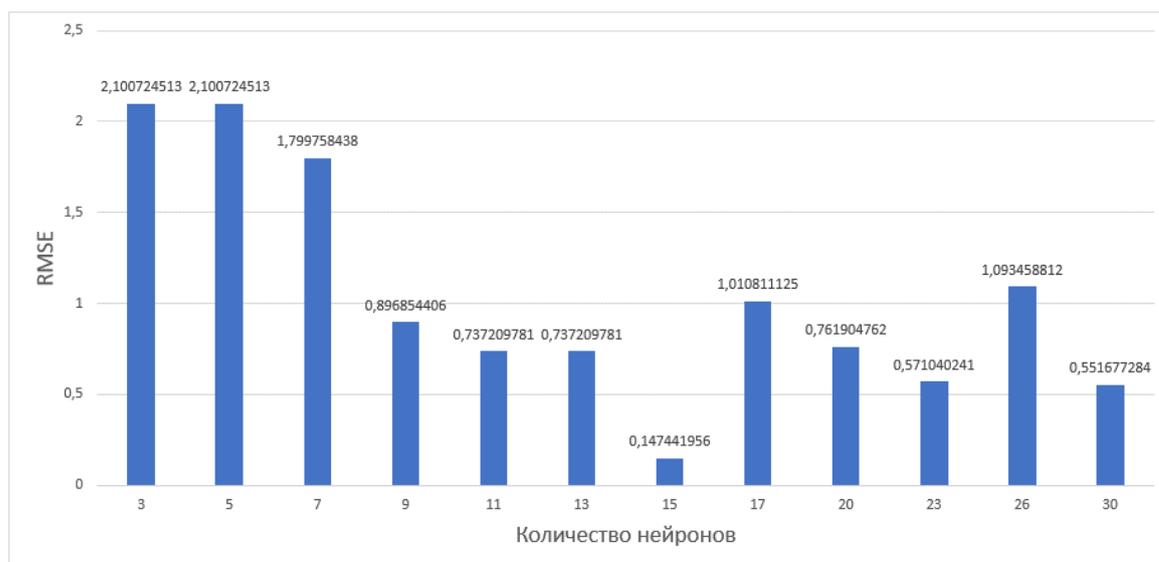


Рис. 3. – Зависимость среднеквадратичной погрешности от количества нейронов в скрытом слое

Из рисунка 3 видно, что при количестве скрытых нейронов равных 15 система имеет наименьшую ошибку.

Для того чтобы визуализировать результат работы нейронной сети, использовалась библиотека TSNE [10]. Одной из важных особенностей библиотеки является то, что при кластеризации похожие друг на друга объекты группируются близко друг к другу.

На рисунке 4 представлена визуализация разделения данных опроса. В легенде классами обозначены укрупненные группы направлений в рассматриваемом университете.

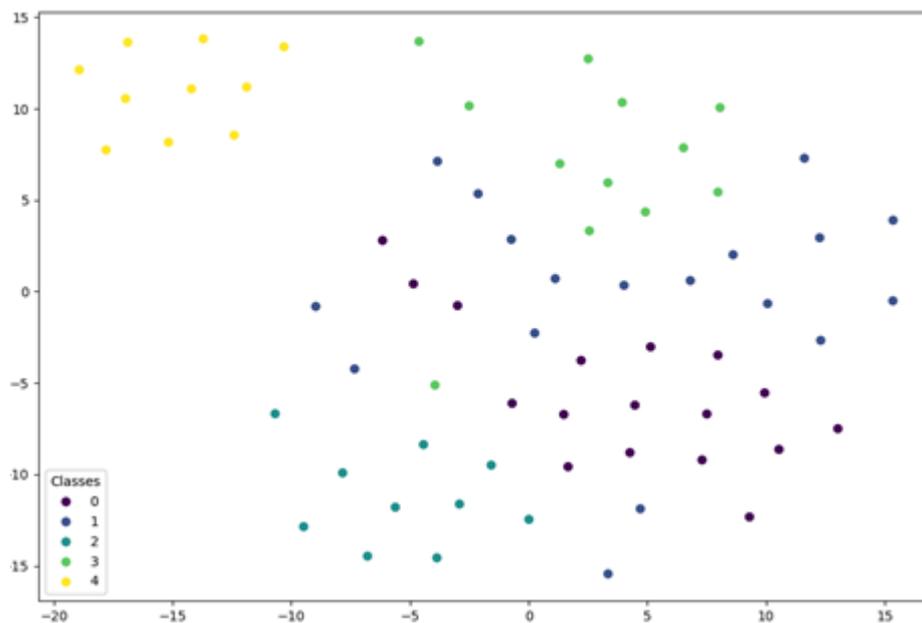


Рис. 4. – Кластеризация данных опроса

В процессе работы с полученной нейронной сетью удалось добиться точности приблизительно 0,92. Полученная в результате модель была сохранена для дальнейшего использования в системе помощи выбора направления подготовки абитуриентам.

Заключение

В данной работе для создания системы помощи выбора специальности были использованы искусственные нейронные сети, а для осуществления классификации выбрана модель многослойного персептрона. Использование предложенной АИС позволит оказать поддержку при принятии решения абитуриентом о выборе наиболее подходящей специальности на основе оценки интеллектуальных и личностных качеств.

Литература

1. Чернов А.В., Бутакова М.А., Шевчук П.С. Кластеризация данных методом растущего нейронного газа // Инженерный вестник Дона, 2020. № 7. URL: ivdon.ru/ru/magazine/archive/N7y2020/6537
2. Корнаухова М.А., Полусмакова Н.С. Применение цифрового маркетинга в системе высшего образования // Современный менеджмент: проблемы, анализ тенденций, перспективы развития. Волгоград, 2022. С. 227-230.
3. Замечник Т.В., Лосев А.Г., Петренко А.Ю. Управляемый классификатор в диагностике рака молочной железы по данным микроволновой радиотермометрии // Математическая физика и компьютерное моделирование, 2019. № 3. С. 53-67.
4. Бабушкина Н.Е., Ляпин А.А. Сравнительный анализ использования нейронной сети в задаче идентификационных свойств материалов // Инженерный вестник Дона, 2021. № 11. URL: ivdon.ru/ru/magazine/archive/n11y2021/7305
5. Федутинов К.А. Машинное обучение в задачах поддержки принятия решений при управлении охраной природы // Инженерный вестник Дона, 2021. № 9. URL: ivdon.ru/ru/magazine/archive/n9y2021/7186
6. Wang Y., Gao W. Prediction of the water content of biodiesel using ANN-MLP: An environmental application // Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 2018. V. 40. № 8. pp. 987-993.
7. Balasamy P.K., Athiyappagounder K. An Optimized Feature Selection Method for E-Learning Recommender System Using Deep Neural Network based on Multilayer Perceptron // International Journal of Intelligent Engineering and System, 2022. V.15. № 5. pp. 461-472.

8. Hernandez R, Atienza R. Career track prediction using deep learning model based on discrete series of quantitative classification // Applied Computer Science, 2021. V. 17. № 4. pp. 55-74.

9. Bendangnuksung, Prabu D. Students' Performance Prediction Using Deep Neural Network // International Journal of Applied Engineering Research, 2018. Volume 13. № 2. pp. 1171-1176.

10. Zahour O., Benlahmar E., Eddaouim A., Hourrane O. A comparative study of machine learning methods for automatic classification of academic and vocational guidance questions // International Journal of Interactive Mobile Technologies, 2020. V. 14. № 8. pp. 43-60.

References

1. Chernov A.V., Butakova M.A., Shevchuk P.S. Inzhenernyj vestnik Dona, 2020. № 7. URL: ivdon.ru/ru/magazine/archive/N7y2020/6537

2. Kornauhova M.A., Polusmakova N.S. Sovremennyy menedzhment: problemy, analiz tendencij, perspektivy razvitiya [Modern management: problems, trend analysis, development prospects]. Volgograd, 2022. pp. 227-230.

3. Zamechnik T.V., Losev A.G., Petrenko A.YU. Matematicheskaya fizika i komp'yuternoe modelirovanie [Mathematical Physics and Computer Simulation], 2019. T. 22. № 3. pp. 53-67.

4. Babushkina N. YE., Lyapin A.A. Inzhenernyj vestnik Dona, 2021. № 11. URL: ivdon.ru/ru/magazine/archive/n11y2021/7305.

5. Fedutinov K.A. Inzhenernyj vestnik Dona, 2021. №9. URL: ivdon.ru/ru/magazine/archive/n9y2021/7186

6. Wang Y., Gao W. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 2018. V. 40. № 8. pp. 987-993.

7. Balasamy P.K., Athiyappagounder K. International Journal of Intelligent Engineering and System, 2022. V.15. № 5. pp. 461-472.



8. Hernandez R, Atienza R. Applied Computer Science, 2021. V. 17. № 4. pp. 55-74.
9. Bendangnuksung, Prabu D. International Journal of Applied Engineering Research, 2018. Volume 13. № 2. pp. 1171-1176.
10. Zahour O., Benlahmar E., Eddaouim A., Hourrane O. International Journal of Interactive Mobile Technologies, 2020. V. 14. № 8. pp. 43-60.