

ПРЕДСКАЗАНИЕ ЦЕЛЕВЫХ ХАРАКТЕРИСТИК ПОВЕДЕНИЯ ОБУЧАЮЩЕГОСЯ В УНИВЕРСИТЕТЕ НА ОСНОВЕ ЦИФРОВОГО ПРОФИЛЯ СТУДЕНТА

Панфилов А.Н. (Университет ИТМО, Санкт-Петербург)

Научный руководитель – к.т.н., доцент Хлопотов М.В. (Университет ИТМО,
Санкт-Петербург)

Аннотация. Предсказание поведения обучающегося, его успеваемости, академических задолженностей, факта окончания Университета – это амбициозная задача, решение которой ведет к качественному улучшению образовательного процесса и является шагом на пути цифровой трансформации сферы образования. В докладе описаны возможные методы применения алгоритмов искусственного интеллекта и промежуточные результаты их применения для решения данных задач, на основе цифрового профиля учитывающего социо-демографические и психометрические характеристики студента.

Введение

С учетом характеристик имеющегося цифрового профиля были рассмотрены и вынесены на обсуждение наиболее релевантные для Университета целевые характеристики студента для предсказания которых могут быть использованы алгоритмы искусственного интеллекта и машинного обучения. Далее описаны целевые характеристики и результаты предсказательных моделей.

Основная часть

Цель работы – рассмотреть возможность предсказания целевых характеристик поведения студента и построить прототипы предиктивных моделей. Основной целевой характеристикой поведения студента в Университете является факт отчисления или успешного окончания Университета. Предсказание данной характеристики можно осуществлять как динамически, с учетом поведения и результатов обучающегося за уже известные семестры, так и интегрально, только с входной информацией о студенте. Имеющийся цифровой предоставляет данные для решения обеих этих задач, однако вторая задача является более полезной, так как позволит сразу же при поступлении оценить шансы студента на успешное окончание университета, кластеризовать потенциально неуспевающих студентов и с учетом вероятностей

принадлежности студента к успевающим или неуспевающим, сформировать группы обучающихся или порекомендовать соответствующую образовательную траекторию. Точность предсказаний в данной задаче классификации предиктивной модели основанной на градиентном бустинге над решающими деревьями является 89%. Результаты данного алгоритма машинного обучения превосходили результаты логистической регрессии и случайного решающего леса, поэтому было решено остановить выбор на этой модели. Примерами целевых нединамических характеристик может являться факт научной или спортивной деятельности студента в течение курса обучения, студентов, отмеченных моделью как потенциально успешных в этих направлениях так же можно выделять во отдельные группы.

Другими релевантными целевыми характеристиками являются динамические характеристики, такие как успеваемость или количество долгов в следующем семестре, факт отчисления в следующем семестре. Для предсказания этих характеристик было рассмотрено два подхода: построение предиктивных моделей на каждый отдельный семестр с параметрами, соответствующими известным результатам и построение общей модели, которая в качестве предыдущих результатов студента использует усредненные динамические характеристики цифрового портрета. Первый подход показал более высокую точность и используя модели градиентного бустинга над деревьями в задаче предсказания количества долгов, рассмотренной как задача регрессии, равную 1.06 по метрике средней абсолютной ошибки. Примером другой целевой динамической характеристики может являться прогнозирование посещаемости студента или группы обучающихся.

Заключение

В результате работы описаны и продемонстрированы примеры выбора целевых характеристик с учетом особенностей цифровой профиля студента, результаты применения методов машинного обучения над цифровым профилем студента для предсказания различных целевых характеристик и возможные прикладные применения данных предсказаний.

Панфилов А.Н.

Хлопотов М.В.