

Педагогика и просвещение

Правильная ссылка на статью:

Кашпур В.В., Губанов А.В., Феценко А.В., Изофатова М.С., Кобенко А.В. — Взаимосвязь образовательных достижений старшеклассников и их цифрового следа в социальной сети. // Педагогика и просвещение. – 2020. – № 4. DOI: 10.7256/2454-0676.2020.4.33952 URL: https://nbpublish.com/library_read_article.php?id=33952

Взаимосвязь образовательных достижений старшеклассников и их цифрового следа в социальной сети.

Кашпур Виталий Викторович

кандидат социологических наук

доцент, кафедра социологии, Томский государственный университет

634050, Россия, Томская область, г. Томск, ул. Ленина, 36

✉ vitkashpur@mail.ru



Губанов Александр Валерьевич

Научный сотрудник, Sirius University of Science and Technology

354340, Россия, Краснодарский край, г. Сочи, Олимпийский проспект, 1

✉ derzhiarbuz@yandex.ru



Феценко Артем Викторович

старший преподаватель, кафедра Гуманитарной информатики, Томский государственный

университет 634050, Россия, Томская область, г. Томск, пр. Ленина, 36

✉ fav@ido.tsu.ru



Изофатова Мария Сергеевна

студент, Sirius University of Science and Technology

354340, Россия, Краснодарский край, г. Сочи, Олимпийский проспект, 1

✉ izofatova.mariya@mail.ru

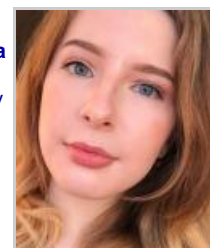


Кобенко Алина Владимировна

студент, Sirius University of Science and Technology

354340, Россия, Краснодарский край, г. Сочи, Олимпийский проспект, 1

✉ a.v.kobenko@outlook.com



[Статья из рубрики "Новые методики и технологии"](#)

DOI:

10.7256/2454-0676.2020.4.33952

Дата направления статьи в редакцию:

22-09-2020

Дата публикации:

09-11-2020

Аннотация.

В работе рассматриваются взаимосвязь между цифровым следом старшеклассников в социальной сети «ВКонтакте» с их высокими формальными образовательными достижениями (призовые места во всероссийских олимпиадах). Для этого с помощью API «ВКонтакте» выгружается цифровой след пользователей, строится классификатор, является ли школьник призёром олимпиад, и выясняется, какие группы переменных (личные характеристики, популярность и новизна постов на стене, подписки на сообщества) имеют наибольшее влияние на точность классификации. Для классификации используются пользовательские данные из социальной сети «ВКонтакте» 12 588 выпускников 11 класса в 2019 и 2020 гг., и алгоритмы машинного обучения. В результате исследования выявлена взаимосвязь между уровнем образовательного потенциала, как способности учащегося побеждать в олимпиадах, и тематикой контента интересного таким учащимся. Также определено 63 сообществ «ВКонтакте» наиболее значимые для классификации участников исследования. Среди сообществ, более популярных у победителей олимпиад, фигурирует тематика, имеющая отношение к науке и сдаче ЕГЭ, а также к интеллектуальным мемам. В подписках «не олимпиадников» выражено присутствие сообществ с юмором и развлекательным контентом. Это наблюдение опровергает распространённое в педагогическом сообществе мнение о негативном влиянии социальных сетей на успешность обучения школьников. Информационное пространство «ВКонтакте» содержит не только развлекательный контент, но информацию поддерживающую или стимулирующую познавательный интерес и учебную мотивацию.

Ключевые слова: олимпиады, образовательные достижения, социальная сеть, машинное обучение, цифровые следы, ВКонтакте, случайный лес, сообщество, юпитер, большие данные

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-313-51024 «Влияние структурных и содержательных характеристик интернет-активности старшеклассников и студентов на их образовательные достижения».

Введение

Тема образовательных достижений находится под пристальным вниманием субъектов образовательной политики всех уровней: **организационном** (школы и университеты); **институциональном** (муниципальные и региональные департаменты образования, министерства просвещения и высшего образования и науки); **общественном** (профессиональные и родительские сообщества и ассоциации). Политику государства в этом направлении субъекты образования реализуют путем выявления и поддержки талантов среди школьников и студентов, проведением профориентации и индивидуализации обучения. Одним из конкретных механизмов претворения в жизнь

политики образовательной эффективности является стимулирование образовательных учреждений в зависимости от успехов их обучающихся. Примером тому служит рейтингование субъектов образовательной деятельности в зависимости от среднего балла ЕГЭ выпускников школ или абитуриентов вузов.

Таким образом, находясь в ситуации конкуренции, образовательные учреждения и экспертные центры начинают собирать массивы информации, связанные с успеваемостью учащихся, и применять чаще сложные аналитические методы для определения наиболее важных факторов, способствующих успешности обучающихся и эффективности обучения. Однако аналитика в большей степени ограничивается оценкой зависимости учебной успеваемости от персональных (связанных с обучающимся) и институциональных (связанных с образовательным учреждением) факторов. Источником информации для такой аналитики выступают данные, генерируемые в рамках LMS образовательных учреждений, данные психологических тестирований, а также данные образовательной и социально-экономической статистики. При этом, игнорируется колоссальный объем данных, позволяющих оценить образовательные достижения учащихся, которые сгенерированы в сети и ключевом для современной молодежи интернет локусе - социальных сетях. Существующие в настоящее время попытки использовать эти данные для индивидуализации процесса обучения сводятся к применению данных LMS учебных заведений, информации о процессе и прогрессе обучения для выстраивания траектории в рамках изучаемого курса. Между тем, проникновение социальных медиа в повседневную и профессиональную жизнь современного общества актуализирует исследования методов и инструментов извлечения и анализа данных о пользователях социальных сетей, так называемых цифровых следах. Цифровой след (или цифровой отпечаток; англ. digital footprint) — совокупность информации о посещениях и вкладе пользователя во время пребывания в цифровом пространстве.

Изучение цифрового следа человека позволяет осуществлять анализ и моделирование его физиологических, психологических и когнитивных особенностей. Сформированная в результате исследования модель вполне применима для целей прогнозирования, программирования и эффективного управления образовательной деятельностью.

Данная работа направлена на изучение вышеназванной проблематики о сборе и анализе данных обучающихся в Интернет. Она посвящена исследованию взаимосвязи формальных образовательных достижений школьников-выпускников 11 классов с их цифровыми следами в социальной сети «ВКонтакте». Гипотеза исследования заключается в том, что характеристики профиля, связанные с поведенческими особенностями (наличие фото профиля, количество друзей, популярность постов, их уникальность и т.д.), а также подписки на сообщества выражены отличаются у школьников, занявших призовые места на олимпиадах, от таковых у школьников, не ставших призерами или победителями. Подтверждение данной гипотезы позволяет установить наличие связи цифрового следа с образовательными достижениями. Кроме этого, становится возможно делать выводы о том, какие компоненты цифрового следа наиболее значимы для этой связи.

Тема использования социальных сетей обучающимися в современных исследованиях представлена в основном в качестве эмпирического материала данных о студентах, особенности поведения которых измеряются преимущественно показателями времени проведенными ими в сети. Например, журнал записанных действий (Jacobsen и Forste 2011); минуты использования (Karpinski et al. 2013); частота и часы использования (Rouis 2012); количество времени, проведенного на Facebook и частота проверки Facebook (Lee 2014); среднее ежедневное использование Facebook (Michikyan et al.

2015); шкала интенсивности Facebook (Ainin et al. 2015); частота использования и средняя продолжительность использования Facebook (Lambić 2016); количество часов в неделю в социальной сети (Alwagait et al. 2014). Самая популярная техника измерения использования социальных медиа - самоотчеты участников исследований. Однако данная методика является неточной, показывает субъективные измерения. Дело в том, что в период проведения исследований 2011-2016 годах, уровень проникновения социальных медиа в повседневную жизнь человека не был настолько высоким как сейчас. Не случайно ранние исследовательские работы по настоящей теме рассматривали только вопросы использования и неиспользования социальных медиа (Kirschner and Karpinski 2010; Kolek and Saunders 2008). И это было актуально для времени становления пользовательской культуры Интернет и знакомства с ним. Спустя годы в литературе произошел отказ от самоотчетного подхода, поскольку социальные медиа укоренились в нашей повседневной жизни. Несколько исследований описывают использования методов измерения, отличных от самоотчетов (Rosen et al. 2013). По существу исследования ученые наблюдали за использованием участниками различных технологий социальных медиа во время обучения и дополнили их анкетой после обучения. Авторы других исследований анализировали журнал записей мероприятия по измерению времени, затраченного на социальные сети (Jacobsen and Forste 2011).

В литературе по изучаемой теме встречаются работы, в которых освещается анализ академической успеваемости и методов ее измерения. Под академической успеваемостью чаще всего принимается общая средняя оценка за семестр. Академическая успеваемость в исследованиях измерялась с помощью: самооценки (Bellur et al. 2015), данных, полученных от преподавателя или из LMS вуза, и предполагаемая субъективная оценка (e.g., Ainin et al. 2015; Janković et al. 2015). В подавляющем большинстве случаев в исследовательских работах использовалась самооценка успеваемости и это, вероятно, говорит о сложности получения точных данных об академической успеваемости. В большинстве исследований сообщалось о демографических переменных участников, таких как пол и возраст. Женщины были лучше представлены среди респондентов, чем мужчины, разброс возраста среди участников исследований от 12 до 67. Размер выборки исследований варьировался от 93 до 1839 человек. В большинстве работ представлен описательный анализ данных. В ранних исследованиях по этой теме преимущественно показывался анализ через корреляции (Pasek et al. 2009). В чуть более поздних - применялся регрессионный анализ (Junco and Cotten 2012; Rosen et al. 2013). Выводы исследователей относительно взаимосвязи использования социальных медиа и академической успеваемостью противоречивы. Их позиции условно делятся на три группы. К первой группе относится только одно исследование, которое обнаружило позитивные отношения указанной взаимосвязи (Ainin et al. 2015). Вторая группа составляет большую часть исследований, которые выявили негативную связь (Glass et al. 2014; Jacobsen and Forste 2011; Janković et al. 2015; Junco and Cotten 2012; Junco, 2012; Junco 2015; Karpinski et al. 2013; Kirschner and Karpinski 2010; Lee 2014; Michikyan et al. 2015; Paul et al. 2012; Rosen et al. 2013; Rouis et al. 2011; Wohn and LaRose 2014). Третья группа - это материалы ученых, не нашедших значимой связи (Alwagait et al. 2014; Bellur et al. 2015; Kabre and Brown 2011; Kolek and Saunders 2008; Lambić 2016; Pasek et al. 2009; Rahman and Stephen 2016; Rouis 2012).

Противоречивость выводов, к которым пришли ученые, можно объяснить субъективной природой данных, полученных ими для анализа от участников как об академической успеваемости, так и характеристиках использования социальных сетей. В то же время в исследованиях, посвященных теме влияния социально-демографических характеристик и

психологических черт личности на поведение в социальных сетях, подтверждается гипотеза о наличии взаимосвязи между особенностями личности и поведением, отраженным в цифровых следах (Gosling et al. 2011; Miller et al. 2012; Markovikj et al. 2013; Schwartz et al. 2013; Kosinski et al. 2014; Arroju et al. 2015). При этом исследователи используют объективные данные о действиях пользователей выгруженных из социальной сети (тексты сообщений, подписки на сообщества, лайки, количество друзей и фотографий) и алгоритмы машинного обучения для их анализа.

Некоторые исследования отечественных авторов подтверждают наличие взаимосвязи поведения в социальной сети не только с психологическими чертами личности, но и когнитивными: академическая успеваемость (Поливанова 2017), диагностика уровня интеллекта и креативности (Фещенко, Кабанова 2018) и образовательных интересов (Резанова 2018). При этом основу данных о поведении пользователей в социальных сетях составили сведения о подписках на тематические сообщества, от нескольких десятков до нескольких сотен на одного пользователя. Эти сведения выгружались с помощью API российской популярной социальной сети «ВКонтакте».

Обзор представленных исследований показывает, что в работах по теме влияния социальных сетей на образовательные достижения не предлагается анализ положительной и отрицательной взаимосвязи использования социальных сетей на оценки учащихся, так как для этого не достаточно данных об особенностях использования контента студентами в сети. Исходя из анализа названных выше научных работ, представляется недостаточным искать взаимосвязь между временем нахождения в сети и образовательными достижениями учащихся. Важно понимать какому именно содержанию обучающийся отдает предпочтение в социальной сети. Поскольку знание о существе интереса обучающегося пользователя в Интернет позволит объяснить как позитивное, так и негативное его влияние на достижения школьника или студента в формальном образовании. Для проверки гипотезы о существовании взаимосвязи цифрового следа и образовательных достижений предлагается использовать анализ не только подписок на тематические сообщества, но и личные данные (пол, возраст, число друзей и публикаций) из профиля пользователя «ВКонтакте».

Данные и методы

Проверка данной гипотезы осуществлялась с использованием методов машинного обучения. В качестве источников данных выступала следующая информация: данные Приемной комиссии ТГУ об абитуриентах, занявших призовые места на всероссийских олимпиадах; пользовательские данные этих школьников, выгружаемые из социальной сети «ВКонтакте». В качестве первоначальных данных были взяты id социальной сети «ВКонтакте» школьников - выпускников Томской области и информация о том, являются ли они призерами олимпиад. Из соображений конфиденциальности id были пронумерованы и отсоединены от информации, полученной в дальнейшем из личных профилей участников исследования, таким образом, данные были обезличены. В выборке присутствуют школьники - выпускники 2019 (450 человек) и 2020 (624 человека) годов. Для решения задачи сравнительного анализа цифровых следов, данные «олимпиадников» дополнены данными о школьниках, не являющимися призерами олимпиад, но из тех же регионов и того же возраста, что и олимпиадники (4879 человек для 2019 года и 6635 человек для 2020 года).

О всех участниках исследования были выгружены данные из профиля «ВКонтакте»: тип профиля (открытый или закрытый); наличие фотографии; наличие информации о дате рождения; пол; число друзей; среднее число просмотров постов на странице; число

репостов; число собственных постов; коэффициент уникальности постов – отношение числа собственных постов к их общему числу. Также из пользовательских профилей были выгружены данные о подписках на тематические сообщества и выделение наиболее популярных сообществ у всех школьников. Сообщества были разделены на два класса: сообщества, на которые подписаны в большей степени призеры олимпиад, и сообщества, на которые подписаны в большей степени школьники, не являющиеся призерами олимпиад. Насколько значимо в сообществе преобладает одна или другая категория школьников определяется точным тестом Фишера.

На основе двух конечных классов сообществ (в скольких суммарно сообществах одного и другого класса состоит школьник) было создано два новых признака количества подписок на сообщества. Сформированы новые признаки из сообществ, которые были выделены ранее в отдельные классы (бинарная матрица подписок школьников на сообщества). Число признаков равно числу сообществ. Школьник либо состоит в сообществе (значение признака равно единице), либо не состоит в сообществе (значение признака равно нулю).

Подготовленные наборы данных за 2019 и 2020 год оказались не сбалансированными (менее 10% учащихся – «олимпиадники»), поэтому производилось семплирование призеров олимпиад, чтобы сбалансировать выборку. Семплирование проводилось после разбиения выборок на тестовую и обучающую, чтобы избежать попадания одинаковых копий в обе выборки. После был обучен классификатор (случайный лес с ограничением высоты деревьев равным 6) для разных наборов параметров и вычислены для них метрики качества. Затем классификаторы сравнивались между собой (кросс-валидация случайными разбиениями с дальнейшим сравнением метрик качества по критерию Стюдента) и формулировался вывод о вкладе тех или иных составляющих цифрового следа в точность классификации.

Выгрузка данных о школьниках 2020 года происходила с помощью API «ВКонтакте» в среде разработки Jupyter с использованием id их личных страниц. Выгрузка данных о школьниках 2019 происходила при помощи специальной платформы ТГУ и базы данных, сохранившей информацию о профилях «ВКонтакте» на период 2019 года. Однако в базе данных не содержится информации о постах на стене пользователей, поэтому данная группа признаков в классификаторах для 2019 года не используется. Обработка данных осуществляется в среде разработки Jupyter.

Результаты

В математической модели объектом (элементом выборки) является выпускник 11 класса. Выборка состоит из всех школьников выпускников – участников исследования. В качестве признаков объектов (независимых переменных), отвечающих за личные характеристики, были выбраны: тип профиля (открытый или закрытый); наличие фотографии; наличие информации о дате рождения; пол; число друзей. Также в наборе данных выпускников 2020 года дополнительно использовались следующие признаки: среднее число просмотров постов; число репостов; число собственных постов; коэффициент уникальности постов.

Кроме того, как для данных 2019 так и для данных 2020 годов добавлены признаки подписки на сообщество наиболее популярное у «олимпиадников» и «не олимпиадников» (всего 63 сообщества выделенные при помощи теста Фишера. Число новых признаков равно числу сообществ (63), и каждый признак является бинарным, принимая значение равное "1" если школьник состоит в сообществе и значение равное "0", если не состоит.

Исследование производилось над выпускниками 2020 года, количество участников исследования составляло 7259 школьников, из которых 624 являются призерами олимпиад. В таблицах ниже приведены средние значения метрик качества для кросс-валидации методом случайных разбиений (разбиение 80/20, $k=10$).

Таблица 1. Признак для каждого сообщества и признаки личных характеристик после семплирования (2020 год, 7259 школьников).

accuracy	precision	recall	ROC-AUC
0.771	0.822	0.704	0.865

Из таблицы видно, что получено положительное значение чувствительности (recall). Это означает, что классификатор хорошо выделяет настоящих призеров олимпиад и тех, кто на них сильно похож, но при этом не является призером олимпиады. Высокое значение ROC-AUC говорит о том, что классы хорошо разделимы.

Если убрать признак подписок школьников на сообщества и оставить только признаки с их личной информацией на странице, то точность уменьшается, но классы по-прежнему разделимы.

Таблица 2. Только признаки личных характеристик (2020 год, 7259 школьников).

accuracy	precision	recall	ROC-AUC
0.684	0.647	0.830	0.758

Если убрать из признаков личную информацию страницы и оставить только признак о подписках на сообщества, точность уменьшится, однако уменьшается чувствительность классификатора.

Таблица 3. Признак для каждого сообщества без личных характеристик (2020 год, 7259 школьников).

accuracy	precision	recall	ROC-AUC
0.692	0.780	0.549	0.758

Следовательно, как личностные характеристики школьника, так и его подписки имеют влияние на точность классификатора.

Для наглядности ниже приведены сообщества, отнесенные к классу наиболее характерных для призеров олимпиад 2020 года: Мемуары ценителей научных мемов; «FEMALE MEMES»; «Рифмы и Панчи»; «VACUUM»; «MARVEL/DC»; «Палата №6»; «голуби папка»; «че»; «Строки пыльных книг|Литература»; «Netflix, Reddit; КБ»; «Словарный запас (EN)»; «околоинтеллектуальные мемы»; «Begin English. Английский язык для всех»; «ЕГЭ 2020»; «my astrology»; «МХК»; «Литература»; «Поступашки - Олимпиады, ЕГЭ и ДВИ»; «Лентач»; «это моя вечеринка»; «CLIQUE»; «Hollywood по-русски»; «Dr»; «Добрые мемы»; «potrachenó»; «ALMERIA»; «Словарный запас»; «какие-то биохимические мемы»; «ЕГЭ | Онлайн школа ЗНАНИЕ»; «NR.Music»; «Репетиторская империя: битва за сотку | ЕГЭ 2021»; «как в книге написано»; «твоей прекрасной юности момент».

Ниже представлены сообщества, определенные к классу наиболее характерных для

школьников, которые не были призерами олимпиад 2020 года: «Леонардо Дай Винчик»; «Овсянка, сэр!»; «Бот Максим»; «ПРИКОЛЫ | Смеяка»; «Новинки Музыки 2020 | Лучшая Музыка»; «Смейся до слёз :D»; «Команда ВКонтакте»; «Четкие Приколы»; «Бог фотошопа»; «СТЫД»; «Музыка ВКонтакте»; «Арт Бот; ОРУ, СЭР!»; «Комментатор от Бога»; «Мембот»; «Сигнал»; «Действительно смешно»; «HD Кино - Фильмы онлайн 2020»; «MORGENSHTERN»; «Я ♥ КИНО»; «ШБТ»; «Вся жизнь в скринах»; «абсурдные комедии» «Знакомо?»; «ОБОИ»; «Правда жизни».

При сравнении характерных для «олимпиадников» и «не олимпиадников» сообществ, можно сделать вывод, что среди сообществ, более популярных у победителей олимпиад, фигурируют сообщества, имеющие отношение к науке и сдаче ЕГЭ, а также к интеллектуальным мемам. Во втором классе «не олимпиадников» выражено присутствие сообществ с юмором и развлекательным контентом.

В наборе данных о выпускниках школ 2019 года, количество участников исследования составило 5329 школьников, из которых 450 являются призерами олимпиад. В таблицах ниже приведены средние значения метрик качества для кросс-валидации методом случайных разбиений (разбиение 80/20, $k=10$).

Так как информация о профиле должна соответствовать тому моменту времени, когда школьник ещё являлся выпускником общеобразовательного учреждения, данные были выгружены из базы данных ТГУ с информацией по профилям 2019 года.

Таблица 4. Признак для каждого сообщества и признаки личных характеристик после семплирования (2019 год, 5329 школьников).

accuracy	precision	recall	ROC-AUC
0.649	0.640	0.701	0.708

Несмотря на высокий recall, accuracy и precision получаются несколько меньше, чем в аналогичном случае для 2020 года. Предположительные причины этого может подсказать следующая таблица 6, из которой видно, что если убрать признак о подписках на сообщества и оставить только личную информацию на странице точность классификации уменьшается, чем в аналогичном классификаторе для 2020 года. В 2019 году среди признаков о личных характеристиках школьников не учитывалось число просмотров их страниц, число постов, число собственных постов и коэффициент уникальности постов. Можно сделать вывод, что эти признаки оказывают значительное влияние на разделимость классов.

Таблица 5. Только личные характеристики (2019 год, 5239 школьников).

accuracy	precision	recall	ROC-AUC
0.563	0.565	0.572	0.581

Если исключить из признаков личную информацию страницы и оставить только признак о подписках на сообщества, точность снижается.

Таблица 6. Признак для каждого сообщества (2019 год, 5239 школьников).

accuracy	precision	recall	ROC-AUC
0.645	0.627	0.728	0.701

Ниже представлены сообщества, определенные как наиболее характерные для призеров олимпиад 2019 года: «Палата №6»; «VLCUUM»; «че»; «ЕГЭ 2021»; «голубая папка»; «Мемуары ценителей научных мемов»; «Dr»; «ЕГЭ Итоговое сочинение | «ЗНАНИЕ»»; «Лучшие стихи великих поэтов | Литература»; «Наука и Техника»; «Словарный запас (EN)»; «Стихи давно забытого поэта»; «Литература»; «Русский язык ЕГЭ 100БАЛЛОВ»; «Словарный запас». Далее представлены сообщества, определенные как наиболее характерные для школьников, которые не были призерами олимпиад 2019 года: «4ch»; «MDK»; «Леонардо Дай Винчик»; «НОВИНКИ МУЗЫКИ 2020»; «Я тебя хочу»; «Бс Максим»; «Смейся до слёз»; «Лайфхак»; «ПРИКОЛЫ | Смеяка»; «КиноКайф - Лучшие фильмы»; «Четкие Приколы»; «Киномания ► Новинки 2020»; «Бог фотошопа»; «Чёткие приколы»; «GIRL MEMES»; «Арт Бот»; «Киномания | Фильмы | Сериалы»; «А ты знал?»; «Команда ВКонтакте»; «ЁП»; «СТЫД»; «Я тебя люблю»; «СМС приколы»; «KFC Россия»; «Vine Video»; «Телеканал ТНТ»; «Я ♥ КИНО»; «НЕНОРМАЛЬНО»; «Школьные истории»; «HD Кино - Фильмы онлайн 2020»; «Мои аудиозаписи»; «Музыка ВКонтакте»; «Действительно смешно»; «Tattoo»; «Книга Рекордов»; «Бесплатные стикеры | Кролик Кир»; «Эскизы»; «Академия Порядочных Парней»; «Улётный Юмор»; «станция №12»; «ты сохранишь»; «ЗАСТЫДИ И ХВАСТАЙ»; «Восемнадцать плюс»; «Филиал сообщений о родственниках»; «Vulgar»; «КАЕФ»; «Записки сумасшедших»; «Videosos»; «Грубо? Простите»; «Комментатор от Бога»; «Улыбнуло».

В первой группе повторяются некоторые сообщества из исследования для 2020 года, а именно: «Палата №6»; «VLCUUM»; «Мемуары ценителей научных мемов»; «Dr»; «Словарный запас (EN)»; «Словарный запас». Во второй же группе повторяются: «Леонардо Дай Винчик»; «Бот Максим»; «ПРИКОЛЫ | Смеяка»; «Смейся до слёз»; «Команда ВКонтакте»; «Четкие Приколы»; «Бог фотошопа»; «СТЫД»; «Музыка ВКонтакте»; «Арт Бот»; «Комментатор от Бога»; «Действительно смешно»; «HD Кино - Фильмы онлайн 2020»; «Я ♥ КИНО».

Таким образом, характерная зависимость сохраняется: на сообщества с юмором и развлекательным контентом чаще подписаны школьники, не являющиеся призерами олимпиад. Призеров олимпиад больше в сообществах, которые связаны с обучением и наукой.

Заключение и выводы

Сравнительный анализ цифровых следов школьников-призеров олимпиад со старшеклассниками, без подтверждения высоких образовательных достижений, показал, что информация о личных характеристиках учащихся и их интересах позволяет создать точный классификатор, определяющий по цифровому следу факт их победы в олимпиадах. Классификаторы, не учитывающие личные характеристики, показали меньшую точность. Кроме того, наилучшую точность показал классификатор, использующий отдельные признаки принадлежности к определенным сообществам. Это означает, что существует связь между уровнем образовательного потенциала, как способности учащегося побеждать в олимпиадах, и тематикой контента интересного таким учащимся. На наш взгляд это наблюдение опровергает распространенное в педагогическом сообществе мнение о негативном влиянии социальных сетей на успешность обучения школьников. Информационное пространство «ВКонтакте» содержит не только развлекательный контент, но и информацию, поддерживающую или стимулирующую познавательный интерес и учебную мотивацию. Пока мы можем констатировать только взаимосвязь подписок на сообщества определенной тематики и успешного участия в олимпиадах. Лонгитюдное наблюдение за цифровым следом учащегося до и после участия в олимпиаде, позволит изучить степень влияния

потребляемого контента в социальной сети на вероятность победы в олимпиаде.

Таким образом, можно предположить возможность в дальнейшем составить цифровую модель школьника-призера олимпиады по цифровому следу «Вконтакте». Сравнение этой модели со следами новых пользователей «Вконтакте» позволит делать прогноз о схожести образовательного потенциала такого учащегося с победителем предметных олимпиад. Прикладное применение такой прогнозной модели можно найти в цифровых рекомендательных сервисах для школьников, где идет речь о их возможности принять участие в олимпиаде с высокой вероятностью победы в ней, а также для университетов, приглашающих старшеклассников к участию в олимпиадах.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-313-51024 «Влияние структурных и содержательных характеристик интернет-активности старшеклассников и студентов на их образовательные достижения».

Библиография

1. Jacobsen W. , Forste R. The wired generation: Academic and social outcomes of electronic media use among university students // *Cyberpsychology, Behavior, And Social Networking*, no. 14(5), 2011.-pp. 275–280.
2. Karpinski A., Kirschner P., Ozer I., Mellott J., Ochwo P. An exploration of social networking site use, multitasking, and academic performance among United States and European university students// *Computers in Human Behavior*, no. 29(3), 2013.-pp. 1182–1192.
3. Rouis S., Limayem M., Salehi-Sangari E. Impact of Facebook usage on students' academic achievement: Role of self-regulation and trust // *Electronic Journal of Research in Educational Psychology*, no. 9(3), 2011.-pp. 961–994.
4. Lee E. Facebook use and texting among African American and Hispanic teenagers: An implication for academic performance // *Journal of Black Studies*, no. 45(2), 2014.-pp. 83–101.
5. Michikyan M., Subrahmanyam K., Dennis J. Facebook use and academic performance among college students: A mixed-methods study with a multi-ethnic sample // *Computers in Human Behavior*, no. 45, 2015.-pp. 265–272.
6. Ainin S., Naqshbandi M., Moghavvemi S., Jaafar N. Facebook usage, socialization and academic performance// *Computers & Education*, no. 83, 2015.-pp. 64–73.
7. Lambić D. Correlation between Facebook use for educational purposes and academic performance of students // *Computers in Human Behavior*, no. 61, 2016.-pp. 313–320.
8. Alwagait E., Shahzad B., Alim S. Impact of social media usage on students academic performance in Saudi Arabia // *Computers in Human Behavior*, no. 51, 2014.-pp. 1092–1097.
9. Kirschner P., Karpinski A. Facebook and academic performance // *Computers in Human Behavior*, no. 26(6), 2010.-pp. 1237–1245,
10. Kolek E., Saunders D. Online disclosure: An empirical examination of undergraduate Facebook profiles// *NASPA Journal*, no. 45(1), 2008.-pp. 1–25,
11. Rosen L., Carrier L., Cheever N. Facebook and texting made me do it: Media-induced taskswitching while studying// *Computers in Human Behavior*, no. 29(3), 2013.-pp. 948–958.
12. Jacobsen W., Forste R. The wired generation: Academic and social outcomes of electronic media use among university students// *Cyberpsychology, Behavior, And Social Networking*, no. 14(5), 2011.-pp. 275–280.

13. Bellur S., Nowak K., Hull K. Make it our time: In class multitaskers have lower academic performance//Computers in Human Behavior, no. 53, 2015.-pp. 63-70.
14. Junco R., Cotten S. The relationship between multitasking and academic performance// Computers & Education, no. 59(2), 2012.-pp. 505-514.
15. Janković B., Nikolić M., Vukonjanski J., Terek E. The impact of Facebook and smart phone usage on the leisure activities and college adjustment of students in Serbia// Computers in Human Behavior, no. 55, 2015.-pp. 354-363.
16. Pasek J., More E., Hargittai E. Facebook and academic performance: Reconciling a media sensation with data// First Monday, no. 14(5), 2009.
17. Michikyan M., Subrahmanyam K., Dennis J.//Facebook use and academic performance among college students: A mixed-methods study with a multi-ethnic sample//Computers in Human Behavior, no. 45, 2015.-pp. 265-272.
18. Wohn D., LaRose R. Effects of loneliness and differential usage of Facebook on college adjustment of first-year students// Computers & Education, no. 76, 2014.-pp. 158-167.
19. Glass R., Li S., Pan R. Personality, problematic social network use and academic performance in China// Journal of Computer Information Systems, no. 54(4), 2014.-pp. 88-96.
20. Gosling S., Augustine A., Vazire S., Holtzman N., Gaddis S. Manifestations of personality in online social networks: Self-reported Facebook-related behaviors and observable profile information// Cyberpsychology, Behavior, and Social Networking, vol. 14, no. 9, 2011.-pp. 483-488.
21. Miller Z., Dickinson B., Hu W. Gender prediction on twitter using stream algorithms with N-gram character features// International Journal of Intelligence Science, vol. 2, no. 4A, 2012.-pp. 143-148.
22. Markovikj D., Gievska S., Kosinski M., Stillwell D. Mining Facebook Data for Predictive Personality Modeling// Proceedings of the 7th international AAAI conference on Weblogs and Social Media, 2013.-pp. 23-26.
23. Schwartz A., Eichstaedt J., Kern M., Dziurzynski L., Ramones S. et al. Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach// PLOS ONE, vol. 8, no.9, 2013.-p. 73791.
24. Kosinski M., Bachrach Y., Kohli P., Stillwell D., Graepel T. Manifestations of user personality in website choice and behaviour on online social networks// Machine learning, vol. 95, no. 3, 2014.-pp. 357-380.
25. Arroju M., Hassan A., Farnadi G. Age, gender and personality recognition using tweets in a multilingual setting// 6th Conference and Labs of the Evaluation: Experimental IR meets multilinguality, multimodality, and interaction, 2015.-pp. 23-31.
26. Поливанова К. Н., Смирнов И. Б. Что в профиле тебе моем данные «ВКонтакте» как инструмент изучения интересов современных подростков//Вопросы образования, № 2, 2017,-с. 134-152.
27. Feshchenko A., Goiko V., Matsuta V., Stepanenko A., Kiselev P. Modelling of an educational profile of a student by analyzing public user data from social networks// INTED2018 Proceedings, 2018.-pp.640-646.
28. Кабанова Т.В., Корепанов К.В., Мацута В.В., Можаяева Г.В., Фещенко А.В. Моделирование признаков одаренности учащихся по цифровым следам в социальной сети «ВКонтакте»// eLearning Stakeholders and Researchers Summit 2018, 2018, с.111-118.
29. Степаненко А.А., Резанова З.И., Гойко В.Л. Автоматическая классификация контента персональных страниц пользователей социальной сети «ВКонтакте» как