ISSN 2079-3316 **ПРОГРАММНЫЕ СИСТЕМЫ: ТЕОРИЯ И ПРИЛОЖЕНИЯ** **т. 12**, № 4(51), с. 33–63

ББК 32.973.202-018.2
ГРНТИ 06.01.29
УДК 519.689.3:007.51

И. Е. Калабихина, Н. В. Лукашевич, Е. П. Банин,
К. В. Алибаева, С. М. Ребрей

Автоматическое извлечение мнений пользователей социальных сетей по вопросам репродуктивного поведения

Аннотация. В данной работе мы представляем специализированный датасет, с разметкой мнений пользователей о репродуктивном поведении. Мы анализируем особенности распределение оценок «за» и «против» по конкретным аспектам репродуктивного поведения. Созданный датасет используется для решения двух задач классификации: классификации сообщений по релевантности изучаемых тем и позиции автора по той или иной теме. Для классификации сообщений используются классические методы машинного обучения, а также нейросетевая модель BERT. Лучшие результаты классификации в обеих задачах достигаются на основе вариантов модели BERT с использованием в классификации пар предложений — варианты NLI (natural language inference — вывод по тексту) и QA (question-answering — вопросно-ответный подход). Кроме того, созданный датасет позволяет сделать содержательные выводы по вопросам отношения пользователей сети ВКонтакте к вопросам репродуктивного поведения. Выявлено, что феномен сознательной бездетности активно представлен в сети, а многодетность остается слабо распространенной моделью поведения. В рамках пронаталистской политики важно формировать позитивное общественное мнение о родительстве, смягчать дефицит времени у родителей.

Ключевые слова и фразы: анализ мнений, BERT, обучение с учителем, демографическая политика, ВКонтакте, репродуктивное поведение.

Работа выполнена в рамках НИР «Воспроизводство населения в социально-экономическом развитии» АААА-А17-117062610054-1[1].

Исследование выполнено при поддержке Междисциплинарной научно-образовательной школы Московского университета «Мозг, когнитивные системы, искусственный интеллект»[1, 2].

Частично поддержано грантом РНФ 21-71-30003 в части создания интерфейса разметки, собственно разметки и методологии применения методов машинного обучения[2].

10.25209/2079-3316-2021-12-4-33-63

Введение

Мнения пользователей социальных сетей по демографическим вопросам могут служить дополнительным источником информации в демографических исследованиях и в разработке научно обоснованной социально-демографической политики. Преимущество такого подхода в оперативности получаемой информации, в возможности получать мнение представителей различных социально-демографических групп практически в реальном режиме времени, в широком наборе вопросов, обсуждаемых пользователями, в выявлении новых феноменов в демографическом поведении. Это требует использования методов автоматической обработки текстов для извлечения мнений по важным для исследования вопросам. Такого рода задачи относятся к широкой области задач анализа тональности [1, 2] или точнее к задаче анализа мнения (позиции) по заданному вопросу (stance detection) [3–5].

В данном исследовании рассматриваются подходы к автоматическому извлечению и анализу мнений пользователей сети ВКонтакте по вопросам репродуктивного поведения, а именно отношение пользователей к рождению или не рождению детей, абортам, мерам государственной поддержки семей в области рождаемости.

Согласно Демографическому энциклопедическому словарю, «репродуктивное поведение — это система действий и отношений, опосредующих рождение или отказ от рождения ребенка в браке или вне брака» [6]. Информация о различных аспектах репродуктивного поведения собирается путем сбора мнений, оценочных суждений респондентов о личных событиях, связанных с планированием семьи, рождением или не рождением ребенка, высказываний о демографической ситуации в стране и демографической политике. Традиционно сбор мнений осуществляется путем социологических опросов. Метод автоматического извлечения текстов из социальных сетей является альтернативным способом сбора мнений (в данном случае мы собираем информацию из свободного обмена мнениями, из повествования о событиях, из оценочных суждений без заданного набора ответов, без довлеющей структуры и логики опроса).

Можно выделить несколько паттернов репродуктивного поведения: многодетность (отношение к рождению трех и более детей), малодетность (отношение к рождению одного или двух детей), бездетность (отношение к отказу от рождения детей). Мы также рассматриваем такие связанные паттерны, как отношение к абортам и направленность на индивидуализм (стремление «жить для себя»), которые существенно

влияют на отношение к рождению детей. Наименее распространенные паттерны (многодетность и бездетность) относительно легко вычленяются из текстов в силу своей исключительности. В отличие от социологических опросов, в которых задается прямой вопрос о том, сколько сам(а) респондент(ка) хочет иметь детей, мы изучаем мнение публики о таких паттернах репродуктивного поведения, как бездетность и многодетность, вне непосредственной зависимости от личной демографической судьбы человека.

По данным последней переписи 2010 года, у самого молодого поколения женщин, завершивших на момент переписи свою репродуктивную историю (45-49-летние женщины в 2010), доля бездетных составила 7%, доля многодетных — около 15%, остальные 78% имели 1 ребенка или 2 детей. Наиболее распространенный паттерн (малодетность) мы изучаем, приняв допущение, что мы можем анализировать отдельные аспекты репродуктивного поведения представителей всего массива родительских групп, принимая их с высокой вероятностью за представителей малодетного типа. В данном случае мы анализируем не столько выбор такой модели (как в маргинальных случаях), сколько детали массового типа репродуктивного поведения (отношение к аборту или рождению ребенка, отношение к мерам демографической политики).

Мнения о паттернах репродуктивного поведения мы собираем через высказывания пользователей социальной сети: 1) о многодетности и бездетности в личной репродуктивной судьбе, в судьбах знакомых, в оценочных суждениях по поводу этих относительно редких поведенческих паттернов, 2) о регулировании рождаемости на персональном уровне и в оценочном формате (о планах о рождении ребенка, об отказе от рождения ребенка, об абортах), 3) о двух блоках мер государственной демографической политики — пособиях по поводу рождения ребенка (включая материнский капитал) и отпусках, связанных с рождением ребенка.

В процессе исследования решается нетривиальная задача выбора демографических тем в области репродуктивного поведения, пригодных для обработки автоматическим методом. Создание автоматического классификатора мнений позволит проводить мониторинг изменений мнений по этим вопросам в реальном времени, отслеживать изменение мнений в зависимости от внешних событий (например, инициатив правительства в рамках современной демографической политики 2006-2025 гг.) за длительный период времени [**7**].

В работе используются мнения пользователей социальной сети ВКонтакте. Сеть занимает 4-место по охвату пользователей в России[1]. Выбор социальной сети ВКонтакте для анализа паттернов репродуктивного поведения представляется обоснованным, поскольку данную сеть устойчиво используют преимущественно лица основных репродуктивных возрастов (примерно половина пользователей сети в 2016 году принадлежало к поколениям, родившимся после 1989 года [8], большинство пользователей в 2020 году находилось в возрасте 25-34 года, 55% пользователей — женщины)[1]. По расчетам авторов на данных официальной статистики, женщины в возрастах 25-34 года вносят основной вклад в рождаемость (более 50% рожденных детей) в современной России.

В данной работе мы представляем специализированный датасет, с разметкой мнений пользователей о репродуктивном поведении (датасет находится в открытом доступе, DOI 10.5281/zenodo.5561126). Мы анализируем особенности распределение оценок «за» и «против» по конкретным аспектам репродуктивного поведения. Созданный датасет используется для решения двух задач классификации: классификации сообщений по релевантности изучаемых тем и позиции автора по той или иной теме. Для классификации сообщений используются классические методы машинного обучения, а также нейросетевая модель BERT. Лучшие результаты классификации в обеих задачах достигаются на основе вариантов модели BERT с использованием в классификации пар предложений — варианты NLI (natural language inference — вывод по тексту) и QA (question-answering — вопросно-ответный подход). Созданный датасет позволяет сделать содержательные выводы по вопросам отношения пользователей сети ВКонтакте к вопросам репродуктивного поведения.

1. Обзор литературы

Анализ тональности и извлечение мнений — это активно развивающаяся область исследований, которая анализирует мнения, настроения, оценки, отношения и эмоции людей на основе письменной или звучащей речи [1, 2].

[1] С. Гаитбаева. Аудитория шести крупнейших соцсетей в России в 2020 году: изучаем инсайты. 2020. URL https://ppc.world/articles/auditoriya-shesti-krupneyshih-socsetey-v-rossii-v-2020-godu-izuchaem-insayty/

Интенсивное изучение задачи определения точки зрения (stance detection) автора социальной сети по заданному вопросу началось в 2016 году, когда Mohammad et al. [3] создали набор данных SemEval-2016, содержащий пять независимых тем, например, «легализация абортов» или «Хиллари Клинтон». Каждая из пар «твит-тема», выбранных для аннотации, была размечена через краудсорсинговую систему CrowdFlower, не менее,чем восемью аннотаторами.

Соббани и др. [9] представили задачу извлечения позиций автора по нескольким темам (Multi-target) и создали датасет, который состоит из трех наборов твитов, соответствующих целевым парам (кандидатам в президенты США), отношение автора к которым оценивается в одном и том же твите: Дональд Трамп и Хиллари Клинтон, Дональд Трамп и Тед Круз, Хиллари Клинтон и Берни Сандерс. Для формирования набора данных были извлечены твиты с хэштегами, относящимися к двум политикам. Задача состоит в определении позиции автора (за, против или прочее) к каждому из политиков, упомянутых в твите. Аннотаторы должны были отвечать на два вопроса о позиции по отношению к каждому из упомянутых кандидатов в президенты в целевой паре интересов. Наилучшие результаты в исходной статье [9], полученные для этого набора данных, составили около 54,81%, что значительно ниже, чем для датасетов о высказанных позициях, имеющих только одну тему. В более поздней работе [10] были получены результаты 65,87% F-меры для данного датасета.

В статье [11] авторы описывают набор данных Will-They-Won't-They (WT–WT), который содержит 51 284 твита на английском языке. Датасет основан на твитах, в которых обсуждаются операции по слиянию и поглощению (M&A) компаний. Для датасета были извлечены твиты, которые помечены по крайней мере двумя хештегами, связанными с разными организациями. Данные размечались финансовыми экспертами. Авторы классифицируют полученные твиты на четыре класса: «поддержка», «опровержение», «комментарий» и «несвязанные», поскольку твиты, упоминающие две организации могли не относиться к теме слияний и поглощений.

В работе [12] авторы представили политематический набор данных на основе высказываний, написанных кандидатами на выборах в Швейцарии. Набор данных состоит из текстов на немецком, французском и итальянском языках, что позволяет проводить кроссязыковую оценку определения позиций авторов. Набор данных состоит из ответов кандидатов на 150 политических вопросов (целей) и содержит 67 000

высказываний. В отличие от обычных подходов к автоматическому определению авторской позиции, которые обучают отдельные модели для каждой темы авторы используют созданный набор данных для обучения единой модели по всем темам. Последние исследования, ориентированные на позицию, посвящены извлечению точек зрения по различным социально-экономическим и демографическим аспектам эпидемии COVID [2, 13–16].

Среди извлечения авторских мнений по демографическим вопросам чаще всего встречается обсуждение проблемы абортов [17–22], разных аспектов родительства [23], проблем здравоохранения [24], влияния различных факторов на демографические процессы, например, природных катастроф [25] и пандемии COVID-19. Помимо авторской позиции в исследовании [17] размечаются доводы в пользу той или иной позиции. В частности, для мнений по поводу абортов 8 наиболее часто используемых доводов за аборты и 5 доводов против абортов. В экономических науках анализ тональностей используется также ранней идентификации трендов финансовых рынков на основе не только социальных сетей и микроблогов, но и отчетов компаний и финансовых учреждений, новостей СМИ и пр. [26].

Наилучшие результаты для задачи извлечения авторской позиции в экспериментах SemEval-2016 были получены с помощью классификатора SVM-ngrams, который в качестве признаков использовал пословные и символьные n-граммы [3]. В последних работах было обнаружено, что наилучшие результаты в обнаружении авторских позицией достигаются подходами на основе нейросетевой модели BERT [27]. Гнош и др. [28] сравнили предыдущие подходы извлечению авторских позиций и выяснили, что модель BERT является лучшей моделью для определения точки зрения авторы для задачи в рамках SemEval2016. В работе [14] сравниваются три группы методов определения авторской позиции в отношении аспектов COVID: на основе сетей LSTM и CNN, а также на основе модели BERT. Наилучшие результаты с большим запасом дают модели на основе BERT.

Для русского языка в 2015-2016 годах было организовано тестирование подходов к извлечению позиций авторов по отношению к мобильным операторам и банкам из постов Твиттера [29]. В 2019 году результаты на этих датасетах были значительно улучшены на основе классификаторов на основе модели BERT [27] и дополнительной автоматической разметки данных [30–32]. Лучшие результаты были получены с использованием русскоязычного варианта модели

BERT RuBERT [33] и подхода к классификации, основанного на парах предложений (подход на основе вывода по тексту — natural language inference (NLI)), при котором специальные дополнительные предложения добавляются к исходным предложениям [34].

В работе [4], авторы изучают автоматическое извлечение авторских позиций о вакцинации детей. Датасет состоит из постов социальной сети ВКонтакте, разделенных на два класса: «за» и «против». Лучшие результаты (84.3 F-меры) были получены классификатором SVM с rbf ядром. В последующей работе [35] были рассмотрены две дополнительных темы: «единый государственный экзамен» и «клонирование человека». Лучшие результаты были получены классификатором на основе голосования нескольких базовых классификаторов (kNN, SVM, Naive Bayes, и др.).

В работе [36] была рассмотрена задача выявления проявлений ненависти на национальной почве, которая была сформулирована как классификация на три класса. Был создан датасет RuEthnoHate dataset, содержащий 5.5 тысяч высказываний в социальных сетях. Лучшие результаты были получены на основе модели BERT с дополнительным предобучением и настройкой, а также лингвистическими признаками и признаками на основе оценочной лексики.

2. Данные

2.1. Сбор данных

Сбор данных осуществлялся в два этапа. На первом этапе для анализа мнений в области репродуктивного поведения отобрано девять групп социальной сети ВКонтакте, в названиях или описаниях которых явно присутствовали слова «чайлдфри» и «childfree» и их вариации, и 341 группа, в названиях или описаниях которых присутствовали ключевые слова «мама», «мамочки», «дети», и др. и количество подписчиков которых было больше 10 000 человек [7]. Глубина поиска составляла 5000 постов. Использование данных из разных групп социальной сети ВКонтакте позволяет избежать гомогенности данных — одного из слабых мест анализа тональностей [37]. Количество групп было выбрано исходя из достаточного объема текста в них. Сторонники бездетного образа (ключевые слова «чайлдфри» и «childfree» и их вариации) жизни публиковали значительно больше постов и комментариев, чем представители остальных групп. Ключевые слова на первом этапе подбирались, исходя из частоты упоминаний. Наиболее подходящие для

дальнейшего анализа слова и направления исследований (редкие типы поведения, регулирование рождаемости, отношение к основным мерам демографической политики) определялись экспертами. Таблица 1 содержит список тем и соответствующие ключевые слова к ним. По выборке данных на первом этапе производился поиск релевантных текстов по ключевым словам из таблицы 1 и последующее их разделение на отдельные группы предложений (по разделителю «.»).

Таблица 1. Списки ключевых слов для извлечения предложений по темам

Тема	Характерные слова
Бездетность	бездетный, нет детей, без детей, childfree, чайлдфри
Индивидуализм	в свое, эго, эгоист, ответственность, для себя, личность, развиваться
Многодетность	многодетный, многодетность, много детей
Аборт	аборт, прерывание
Выплаты	маткапитал, материнский капитал, выплаты, пособие
Отпуска	декрет, отпуск

Наборы ключевых слов по каждой теме анализировались посредством оценки частоты использования каждого ключевого слова или словосочетания по темам в группах антинаталистов и пронаталистов. Некоторые темы сохранялись и при невысокой частоте употребления по причине их содержательной значимости для исследования. В этом случае темы объединялись в одну тему (например, «материнский капитал» и «семейные/детские пособия»). В настоящей работе исследуемые темы были систематизированы по направлениям, как показано в таблице 2.

Таблица 2. Соотношение направлений исследования и выбранных тем

Направления исследования	Темы
Редкие паттерны репродуктивного поведения	«бездетность» «многодетность» «индивидуализм»
Регулирование рождаемости на индивидуальном уровне	«аборт»
Отношение к текущей демографической политике	«материнский капитал + семейные/детские пособия», «родительский отпуск» (связанный с рождением детей)

В дополнение к теме «бездетность» отдельно разрабатывалась тема «индивидуализм» (в контексте «пожить для себя»). Пристальное внимание к паттерну «бездетность» связано с прогнозами распространения такой модели поведения в России. Такой прогноз основан на теоретической концепции второго демографического перехода [**38**] и на исторических аналогиях модернизированных в демографическом отношении стран, в которых уровень бездетности может превышать 20% [**39**]. Выбор темы «индивидуализм» основан на гипотезе мотивации сознательной бездетности — изменении системы ценностей, увеличения набора жизненных траекторий, конкуренция ценностей самореализации и семейных ценностей, рост индивидуализации и приоритетов саморазвития. Данные темы отражают как отношение пользователей к (не)рождению детей, так и их оценку мер, предпринимаемых государством для повышения рождаемости.

После формирования выборки на первом этапе производился анализ количества текстов по каждой теме. В результате был собран датасет, характеризующий общий интерес к заявленным в таблице 1 темам в целевых группах, но, как оказалось, не сбалансированный по классам: основной дисбаланс вносило преобладание тональностей в высказываниях по теме из таблицы 1, чаще связанной с антинаталистскими взглядами. Для выравнивания выборки был проведен второй этап сбора данных из пронаталистских групп.

На втором этапе использовался уточненный список целевых групп для сбора релевантных текстов пронаталистской направленности. В уточненный список целевых групп вошли 15 групп, в названиях и описаниях которых присутствовали слова «чайлдфри», «childfree», «не хочу детей», «не нужны дети», и 42 группы, в названиях или описаниях которых присутствовали ключевые слова «мама», «мамочки», «дети», и др. На втором этапе глубина поиска составляла 20000 тысяч постов с отступом в 5000 постов для групп, по которым уже был проведен поиск на первом этапе. Из общего набора текстов отбирались тексты, соответствующие пронаталистской направленности. Для некоторых групп глубина поиска доходила до 100000 постов (например, https://vk.com/club34677924 — «Беременность»), что охватывало период до 2018 года, таким образом были выявлены группы с длинной «историей» обсуждения исследуемых тем. Отметим, что в группе пронаталистов в значительной степени присутствуют и представители малодетной модели репродуктивного поведения. Выборка первого

этапа сбора данных сформирована двумя авторами данной статьи и находятся в открытом доступе [41, 42], выборка второго этапа сбора данных размещена по DOI 10.5281/zenodo.5561126.

2.2. Разметка данных

Предложения из собранной выборки размечались тремя аннотаторами. Поскольку в каждом предложении могли обсуждаться несколько вопросов, то аннотатор каждое предложение размечал по всем шести темам. Предложения размечались преимущественно профессиональными демографами и лингвистами.

В отличие от других наборов данных, для разметки которых использовались три оценки для аннотации авторской позиции, в данном наборе использовались шесть типов меток, а именно:

- Оценка «нерелевантно», используется для разметки нерелевантных к теме предложений, поскольку вхождение в предложение ключевого слова не гарантирует релевантность предложения теме. Кроме того, предложения, извлеченные по конкретному ключевому слову конкретной темы, опрашиваются на соответствие другим темам;
- Оценки «за» или «против»;
- Оценка «нейтрально», используется для разметки фактографических предложений без каких-либо видимых оценок;
- Оценка «положительно и отрицательно», используется для предложений, в которых упомянуты и позитивные, и негативные мнения по теме;
- Оценка «неясно», используется для явно оценочных, эмоциональных предложений, в которых контекст предложения не дает возможности определить направленность мнения (за или против).

После разметки три оценки «нейтрально, «положительно и отрицательно» и «неясно» были объединены в единый класс «Прочее». Таким образом, на текущем этапе рассматривается классификация на три класса («За», «Против» или «Прочее») или на четыре класса («За», «Против», «Прочее», «Нерелевантно»). Вместе с тем более конкретные оценки в классе «Прочее» предполагается использовать в дальнейшем, для извлечения аргументации.

Оценка предложения выводится на основании оценок нескольких аннотаторов посредством голосования, по большинству голосов.

Встречались предложения, при разметке которых, все три голоса были отнесены к разным классам («За, «Против», «Прочее»), это обычно означает, что обрабатываемое предложение имеет сложную структуру, например пересказ чужого мнения с собственной оценкой. На текущем этапе такие предложения исключались из выборки. Таблица 3 содержит результаты разметки по темам и классам оценок. Текущий датасет содержит 5413 предложений, предложения могут содержать оценки по нескольким темам.

Таблица 3. Распределение авторских оценок по темам, относящимся к (не)рождению детей

Тема	Релевантно	За	Против	Прочее
Бездетность	1636	655	413	568
Индивидуализм	615	373	142	100
Многодетность	464	205	158	101
Аборты	1399	677	187	535
Выплаты	789	106	302	381
Отпуска	970	119	271	580
Всего	5873	2135	1473	2265

Демографическая интерпретация полученных данных была выполнена как на основе числовых соотношений различных авторских позиций, так и содержательном анализе собранных высказываний.

Из таблицы 3 видно, что оценки «за» и «против» достаточно неравномерно распределены по темам, несмотря на два этапа сбора данных. Превалирование оценок «за» в теме «аборт» связано с позицией пользователей соцсети в отношении того, что женщины должны иметь право на аборт; с отношением населения к аборту как к приемлемому средству регулирования рождаемости. Уровень абортов значительно снизился в настоящее время в России, но отношение к этому способу регулирования рождаемости показывает не столько готовность к действиям, сколько признание репродуктивных прав. Были сделаны дополнительные усилия на втором этапе, чтобы найти высказывания против абортов для балансировки датасета, однако все равно таких высказываний оказалось меньше, чем в других классах.

Относительно небольшое число высказываний по теме «многодетность» связано с современным распределением женщин по числу рожденных детей. Рост доли многодетных семей среди семей с несовершеннолетними детьми в период демографической политики (с 7 до 9 процентов, по данным переписи 2010 г. и микропереписи 2015 г.) не меняет того факта, что такой тип репродуктивного поведения остается редким. Присутствие этого феномена в социальной сети также относительно редкое, даже в специализированных группах.

Преобладание негативной оценки по теме «родительские отпуска» связано с двумя сюжетами: 1) декретный отпуск, отпуск по уходу за детьми негативно воспринимается работодателями и коллегами; 2) сами женщины часто сетуют на возросшую нагрузку в этот период, на изменение образа жизни в течение таких отпусков, на дефицит свободного времени. Мнения о пособиях также скорее негативные, поскольку их размер невелик для получателей. А сторонники чайлдфри не хотят, чтобы их налоги шли на семейные пособия.

По теме «бездетность» видно превалирование позитивных оценок, поскольку сторонники этого паттерна достаточно эмоциональны, тема, видимо, активна в период роста представителей такой модели. Кроме того, мы должны помнить о селекции — мы работаем с текстами групп антинаталистов и пронаталистов. В среде антинаталистов тема выражена преимущественно позитивно, в последней группе тема о бездетности менее популярна. Тем же можно объяснить и превалирование позитивных оценок в теме «индивидуализм», которая трактуется как саморазвитие, направление ресурсов на собственные удовольствия. Эта тема часто появляется в контексте обоснования положительной позиции по бездетности, наша гипотеза о таком обосновании подтвердилась.

Помимо селекции, вызванной спецификой данной работы (опора на данные специфических социально-демографических групп пронаталистов и антинаталистов) и спецификой пользователей ВКонтакте (преимущественно лица в возрастной группе 25-34 года), для адекватного использования результатов исследования важно помнить и о характере высказываний в соцсетях. Не следует отождествлять срез мнений пользователей соцсетей с персональным планированием пользователей сети своей демографической судьбы или с результатами соцопросов по причине преобладания оценочных критических

высказываний в сетях по многим вопросам. Скорее, это инструмент отслеживания сигналов о критической позиции общественного мнения для своевременного снятия проблем в области социально-демографической политики или инструмент выявления новых социальных процессов [39]. Например, само группообразование в стиле чайлдфри и большое число высказываний о бездетности (даже в пронаталистских группах) может свидетельствовать о возникновении и росте популярности такого паттерна репродуктивного поведения (до 2010-х гг. в России бездетность не была массовым сознательным выбором). Количество высказываний о бездетности свидетельствует о популярности феномена, что соответствует данным последних опросов населения, которые отмечают значительный рост людей, заявляющих о нежелании иметь детей [40].

В контексте специфики высказываний пользователей соцсетей можно сделать вывод о некоторых рекомендациях в отношении социально-демографической политики, скорее, в контексте «что не делать», что не является популярным. Например, о непопулярности темы запрета абортов, о слабой популярности модели многодетной семьи, о необходимости формирования позитивного общественного мнения к поддержке родительства, о развитии государственной поддержки родительства (в первую очередь мер по развитию баланса «работа-семья» или «жизнь-семья» для сохранения времени на саморазвитие у молодых родителей).

3. Модели

В исследовании рассматривались две задачи классификации: классификация высказываний на релевантные/нерелевантные и классификация релевантных высказываний на три класса тональности позиций. Классификация сообщений по релевантности важна, поскольку сообщения извлекались не по хэштегам, как во многих других работах, а по ключевым словам, которые не всегда точно характеризуют тему сообщений.

В качестве базовых моделей используются классические методы машинного обучения: наивный байесовский классификатор в двух вариантах мультиномиальный (MNB) и Бернулли (BNB), метод опорных векторов (SVC), Gradient Boosting (GB), случайный лес (Random

Forest). В качестве основного метода использовалась нейросетевая модель BERT [28], в версии Conversational RuBERT, для создания которой использовалась русскоязычная модель RuBERT [34], которая была дообучена на русскоязычных диалогах и текстах социальных сетей.

Использовалось три варианта обучения модели BERT: классификация целевого высказывания, а также так называемые NLI (Natural Language Inference — вывод по тексту) и QA (question-answering — вопросно-ответный) подходы. В NLI и QA подходах модель получала пары (текст, предположение). Для классификации релевантности NLI и классификации позции QA этим предположением был сам аспект («Аборты», «Выплаты» и т.д.), для классификации позиции NLI предположение включало в себя еще и саму позицию («Негативно к абортам», «Нейтрально к выплатам» и т.д.)

Примеры входа для классификации высказываний по модели BERT описаны в таблице 4. Первый столбец содержит задачу классификации, названия моделей и количество классов, на которые идет классификация. RuBERT обозначает базовую модель, на вход которой подается одно предложение. Классификация по тональности позиции NLI производится на два класса: следует ли из анализируемого предложения добавленное предположение. Так производится классификация по всем трем классам и выбирается наиболее вероятное второе предложение.

Третий столбец таблицы 4 показывает примеры данных для классификации по каждой моделью и токенизированные входные данные, которые подаются на вход классификатора, включая служебные токены CLS и SEP, которые являются стандартные компонентами подачи данных при обучении модели BERT. В BERT используется так называемая WordPiece токенизация. WordPiece-токенизатор создает последовательность токенов из входных предложений следующим образом: если словарь токенизатора содержит текущее слово, то его представление не изменится. Словарь обычно содержит слова, частотные в коллекции, на которых обучалась модель. Если же слово отсутствует, то токенизатор делит его на подслова таким образом, чтобы полученные токены были как можно более распространены в корпусе. Первое подслово — это слово или префикс, часто встречающееся в корпусе, а другие подслова будут иметь префиксные символы «##», указывающие на то, что они следуют за некоторыми другими подсловами.

Таблица 4. Классификаторы на основе модели BERT и представление входных данных

Задача Модель	Предложение Токенизированый вход
Релевантность RuBERT (2)	(солидарен! всегда считал, что все эти маткапиталы только вредят) ['[CLS]','солидарен', '!', 'всегда', 'считал', ',', 'что', 'все', 'эти', 'матка', '##пита', '##лы', 'только', 'вредят', '[SEP]']
Релевантность NLI (QA) (2)	('солидарен! всегда считал, что все эти маткапиталы только вредят', 'Выплаты') ['[CLS]', 'солидарен', '!', 'всегда', 'считал', ',', 'что', 'все', 'эти', 'матка', '##пита', '##лы', 'только', 'вредят', '[SEP]', 'Выплаты', '[SEP]']
Позиция RuBERT (3)	солидарен! всегда считал, что все эти маткапиталы только вредят ['[CLS]','солидарен', '!', 'всегда', 'считал', ',', 'что', 'все', 'эти', 'матка', '##пита', '##лы', 'только', 'вредят','[SEP]']
Позиция NLI (2)	('солидарен! всегда считал, что все эти маткапиталы только вредят', 'Положительно к бездетности') ['[CLS]','солидарен', '!', 'всегда', 'считал', ',', 'что', 'все', 'эти', 'матка', '##пита', '##лы', 'только', 'вредят', '[SEP]', 'Положи', '##тельно', 'к', 'бездет', '##ности', '[SEP]']
Позиция QA (3)	('солидарен! всегда считал, что все эти маткапиталы только вредят', 'Бездетность') ['[CLS]','солидарен', '!', 'всегда', 'считал', ',', 'что', 'все', 'эти', 'матка', '##пита', '##лы', 'только', 'вредят', '[SEP]' , 'бездет', '##ность', '[SEP]']

4. Эксперименты

В работе использовались реализации классических методов машинного обучения из пакета scikit-learn[2]. Обучение алгоритмов производилось на основе векторных представлений предложений в виде вектора слов с минимальной подокументной частотой в данных 5, с весами tf-idf. Для настройки параметров алгоритмов использовалась процедура grid-search на валидационных данных. Лучшие результаты по разным темам для классических подходов были получены на основе методов: Байесовский классификатор Бернулли (BNB), метод опорных векторов (SVC) и Gradient Boosting (GB), поэтому в дальнейших таблицах результатов приводятся только результаты этих методов из всех опробованных классических подходов.

Для оценки качества классификации используются меры: правильность классификации (Accuracy) и F-мера.

Для экспериментов датасет был разделен на обучающее, валидационное и тестовое множества. Валидационное и обучающие множества имеют размер по 10% от объема все коллекции и имеют распределение классов позиций, сходное распределению классов позиций по темам во всем датасете.

При применении модели BERT использовалась всегда одна полносвязная сеть: dropout с вероятностью 0.5, линейный слой размера 768, размер скрытого состояния — 256, функция активации ReLU, dropout, линейный слой размера 256, на выходе количество меток (2 или 3), learning rate = 0.0005, batch = 64. Количество эпох при применении нейронных сетей выбиралось по лучшему показателю F-меры на валидационном множестве.

Результаты классификации текстов по релевантности представлены в таблице 5 (для классических методов машинного обучения) и таблице 6 (для моделей BERT). Величина меры Accuracy намного выше, чем F-мера для всех тем. Это связано с тем, что в величине Accuracy учитывается доля правильных классификаций как релевантных, так и нерелевантных высказываний, при этом нерелевантных высказываний для каждой темы намного больше, чем релевантных. F-мера рассчитывается как среднее гармоническое точности и полноты только для релевантных сообщений. Во втором столбце Таблицы 5 показано качество извлечения релевантных высказываний по ключевым словам.

Таблица 5. Результаты определения релевантности по темам классическими методами

	Ключевые слова		GB		SVC		BNB	
	ACC	F1	ACC	F1	ACC	F1	ACC	F1
Бездетность	**96,68**	**91,35**	90,41	84,15	91,14	84,62	89,85	82,65
Индивидуализм	90,04	54,24	92,25	58,82	92,07	58,25	87,82	38,89
Многодетность	95,02	58,46	93,73	43,33	93,54	33,96	89,48	19,72
Аборты	**99,63**	**99,3**	97,97	96,0	97,97	96,0	97,79	95,62
Выплаты	97,23	90,32	**98,52**	**95,06**	97,97	93,17	92,44	71,72
Отпуска	96,68	91,35	**97,23**	**92,06**	97,05	91,49	94,28	81,66
Вся выборка	94,96	85,20	**95,02**	**78,2**4	94,96	76,25	91,94	65,04

Таблица 6. Результаты определения релевантности по темам моделями BERT в сравнении с ключевыми словами

	Ключевые слова		RuBERT		NLI	
	ACC	F1	ACC	F1	ACC	F1
Бездетность	**96,68**	91,35	92,99	87,25	95,27	**93,62**
Индивидуализм	90,04	54,24	93,54	71,07	**94,83**	**75,44**
Многодетность	95,02	58,46	97,05	80,95	**97,23**	**82,35**
Аборты	**99,63**	**99,3**	98,52	97,12	98,71	97,49
Выплаты	**97,23**	90,32	96,86	89,94	**97,23**	**90,68**
Отпуска	96,68	91,35	95,39	86,91	**97,79**	**93,62**
Вся выборка	94,96	85,20	95,73	85,54	**96,83**	**90,91**

Наилучшие результаты по извлечению релевантных сообщений по темам получены методами машинного обучения на основе модели BERT. Наиболее простой для определения релевантности является тема «Аборты», поскольку в подавляющем большинстве случаев тема определяется по самому слову «аборт» или образованных от него слов. В этом случае тема с высоким качеством определяется также и ключевыми словами. Наиболее сложной для определения релевантности методами машинного обучения является тема «Индивидуализм». Данная тема имеет разнообразные формы выражения в высказываниях, что подтверждается и относительно низкими значениями F-меры по ключевым словам.

[2]https://scikit-learn.ru/

Таблицы 7 и 8 представляют результаты, полученные для классификации по тональности мнений по всем темам и в среднем по коллекции. Используются меры качества: Accuracy и macro F-мера, которая вычисляется усреднением F-меры по трем классам тональности авторской позиции. В таблице 8 представлено 4 варианта методов на основе модели BERT: RuBERT — это классификация на основе отдельного предложения, NLIsingle — это модель NLI, обученная на обучающих данных только соответствующего аспекта, модели NLI и QA обучаются на всем объеме обучающих данных — это возможно, поскольку в процессе обучения указывается аспект, который должна оценить модель.

Таблица 7. Результаты определения тональности позиции по темам классическими методами машинного обучения

	GB		SVC		BNB	
	ACC	F1	ACC	F1	ACC	F1
Бездетность	56,17	55,77	**57,41**	**57,96**	55,56	55,0
Индивидуализм	**54,84**	**42,22**	51,61	22,94	50,0	32,0
Многодетность	**61,36**	**56,75**	50,0	43,0	59,09	51,51
Аборты	48,94	39,37	53,19	23,15	**53,9**	**43,41**
Выплаты	**61,25**	**53,74**	61,25	52,56	56,25	46,69
Отпуска	56,25	38,21	59,38	39,59	**65,62**	**53,3**
Вся выборка	56,47	**47,68**	55,47	39,87	**56,74**	46,97

Таблица 8. Результаты определения тональности позиции по темам моделями типа BERT

	RuBERT		NLIsingle		NLI		QA	
	ACC	F1	ACC	F1	ACC	F1	ACC	F1
Бездетность	**69,14**	**71,37**	69,14	71,09	68,52	70,27	61,11	64,27
Индивидуализм	**66,13**	48,71	59,68	50,77	64,52	**57,9**	61,29	55,86
Многодетность	59,09	45,04	63,64	61,1	**65,91**	**65,31**	43,18	44,65
Аборты	61,7	58,46	63,12	60,99	**68,09**	**65,12**	63,12	61,62
Выплаты	52,5	34,37	**65,0**	**45,21**	61,25	44,77	58,75	51,65
Отпуска	56,25	34,97	**66,67**	51,76	64,58	**55,43**	60,42	49,97
Вся выборка	60,80	48,82	64,54	56,80	**66,15**	**65,88**	59,83	48,82

Лучшие результаты классификации получены моделью BERT NLI, обученной на парах предложений, обучение выполнено на полной

выборке для каждой темы. Следующие по качеству результаты получены моделью NLIsingle, которая для каждой темы обучалась только на обучающих данных своей темы. Достигнутые результаты сопоставимы с результатами на некоторых других наборах данных [10-11]. Отметим, что мы применили такую же модель BERT NLI к данным работы [15], в которых оценивается отношение пользователей к аспектам, связанным с коронавирусной инфекцией, и получили результаты, сравнимые с лучшими результатами моделей, представленными в этой статей — около 80% меры Accuracy.

Результаты, представленные в таблицах 7 и 8, получены на выделенной тестовой выборке, в которой сохранены пропорции всех классов исходной выборки. Дополнительно было произведено тестирование лучшего метода NLI с помощью метода кросс-валидации на 10 частях. Были получены похожие результаты Accuracy 65.44, F-мера — 65.12.

Текущее качество работы модели связано со сложностью данных, извлекаемых по темам в области репродуктивного поведения в конкретной социальной сети. Во-первых, в датасете работы [15], как и многих других датасетах в этой задаче, анализируемые высказывания извлекаются на основе хештегов, т.е. анализируемый аспект находится в основном фокусе высказывания. Мы извлекали высказывания по ключевым словам (у комментариев сети ВКонтакте нет хештегов), что влечет за собой менее явное упоминание темы, хотя позиция по теме все равно может быть извлечена. Во-вторых, объемы датасетов по каждой теме относительно небольшие. Это связано с тем, что некоторые позиции достаточно редко высказываются, поэтому требуется достаточно много усилий, чтобы их собрать. Поэтому почти во всех темах присутствует дисбаланс по классам тональности позиции, когда один из классов представлен значительно меньше в обучающей выборке, чем остальные.

5. Анализ ошибок

Для анализа ошибок классификации позиции автора была выбрана модель BERT NLI. В таблице 9 представлена матрица ошибок по этой

Таблица 9. Матрица ошибок классификации по тональности авторских позиций

	Негативно	Прочее	Положительно
Негативно	98	25	27
Прочее	27	177	53
Положительно	14	52	112

модели. Видно, что модель относительно редко путает позитивный и негативный классы. Основные ошибки связаны с классом Прочее.

Рассмотрим некоторые примеры высказываний, получивших оценки классификатора, противоположные оценкам аннотаторов.

Следующие высказывание получает позитивную оценку классификатора по отношению к абортам, оценка аннотаторов «против». Ошибка видимо связана с тем, что слово аборт отделено знаками препинания от упоминания проблем, связанных с ним:

*а что ей надо было сделать???**аборт**???чтоб она потом не смогла иметь детей?и жалела всю оставшуюся жизнь!!!!!!!*

В следующем высказывании классификатор «не замечает» негативное отношение к отсутствию детей, выставляет позитивную оценку «бездетности». Ошибка может быть объяснена далеким расположением негативного слова «непродуктивный» и более близким расположением слова «смело», которое обычно носит явно позитивную оценку:

*по моему опыту, если у кандидата на вакансию в моей компании в таком возрасте **нет детей**, не закончено образование, можно смело предположить: это непродуктивный работник, не способный завершать начатое, придерживаться плана и т*

В следующем высказывании аннотаторы отмечают негативное отношение к бездетным (чайлдфри). Модель ставит оценку «за». Видимая причина состоит в наличии ярких позитивных слов (добрые и всепомогающие). Однако в тексте фактически такие добрые и всепомогающие люди противопоставляются чайлдфри, что не удается уловить классификатору.

слава богу они ошибаются, и добрые и всепомогающие люди действительно есть, и их намного больше, чем этих «бабкаешек» — ***чайлдфри***.

В сложных с содержательной точки зрения предметных полях, каким, безусловно, является репродуктивное поведение, много высказываний трудно классифицируемых, содержащих иронию, иносказательные выражения.

6. Заключение

В настоящее время существенным направлением в анализе общественного мнения является извлечение мнений пользователей из текстов социальных сетей. В данной работе описано создание специализированного датасета на русском языке с классификацией позиций пользователей по репродуктивному поведению. После анализа имеющейся литературы можно сказать, что это первый набор данных, посвященный анализу мнений по нескольким аспектам в конкретной подобласти демографии на русском языке.

На собранных данных были исследованы две задачи, существенные для анализа мнений в социальных сетях в реальном времени: классификация высказываний по релевантности и классификаций релевантных мнений по тональности позиции.

В обеих задачах лучшие результаты получены на основе варианта модели NLI BERT, на вход которой данные подаются в виде двух предложений, и обучение для классификации по конкретным темам производится на всем объеме обучающих данных.

Основные выводы по изучаемым аспектам репродуктивного поведения следующие. Паттерн «многодетность» редко обсуждается пользователями и остается менее популярной моделью поведения. Паттерн «бездетность» активно обсуждается в соцсети ВКонтакте, преимущественно в соответствующих образованных группах типа чайлдфри, в которых имеет позитивную окраску. Данный паттерн обсуждается и в пронаталистских группах. Это свидетельствует о присутствии феномена сознательной бездетности в современной России.

Превалирование позитивных оценок в теме «индивидуализм», которая трактуется как саморазвитие, направление ресурсов на собственные удовольствия, часто появляется в контексте обоснования положительной позиции по бездетности. Выявлено позитивное отношение к теме «аборт» в контексте репродуктивных прав женщин и допустимого средства регулирования рождаемости в разных группах.

Преобладание негативных оценок в отношении пособий и родительских отпусков связано либо с неготовностью публики поддерживать родительство собственными ресурсами (налоги, рабочее время, хлопоты с сотрудниками-родителями), либо с персональными трудностями воспитания маленьких детей, дефицита времени и запросом на большую помощь со стороны государства.

С учетом описанной селекции в используемых базах и специфики высказываний пользователей соцсетей можно сделать вывод о некоторых рекомендациях в отношении социально-демографической политики в контексте определения потенциально непопулярных мер или оценки недостатков текущей политики. Например, о непопулярности темы запрета абортов, об отсутствии выраженной поддержки многодетности, о необходимости формирования позитивного общественного мнения к поддержке родительства, о развитии государственной поддержки родительства (в первую очередь мер по развитию баланса «работа-семья» или «жизнь-семья» для сохранения времени на саморазвитие у молодых родителей).

Список литературы

[1] F. A. Pozzi, E. Fersini, E. Messina, B. Liu. “Challenges of sentiment analysis in social networks: an overview”, *Sentiment Analysis in Social Networks*, Elsevier, 2017, ISBN 978-0-12-804412-4, pp. 1–11. ↑34,36

[2] B. Liu. “Sentiment analysis and opinion mining”, *Synthesis Lectures on Human Language Technologies*, **5**:1 (2012), pp. 1–167. ↑34,36,38

[3] S. Mohammad, S. Kiritchenko, P. Sobhani, X. Zhu, C. Cherry. “Semeval-2016 task 6: detecting stance in tweets”, *Proceedings of the 10th International Workshop on Semantic Evaluation*, SemEval-2016 (June 2016, San Diego, California), ACL, 2016, pp. 31–41. ↑34,37,38

[4] S. V. Vychegzhanin, E. V. Kotelnikov. “Stance detection based on ensembles of classifiers”, *Programming and Computer Software*, **45**:5 (2019), pp. 228–240. ↑34,39

[5] D. Küçük, F. Can. “Stance detection: A survey”, *ACM Computing Surveys*, **53**:1 (2021), 12, 37 pp. ↑34

[6] *Демографический энциклопедический словарь*, ред. Валентей Д.И., Советская энциклопедия, М., 1985, 608 с. ↑34

[7] I. E. Kalabikhina, E. P. Banin, I. A. Abduselimova, G. A. Klimenko, A. V. Kolotusha. “The measurement of demographic temperature using the sentiment analysis of data from the social network VKontakte”, *Mathematics*, **9**:9 (2021), 987, 25 pp. ↑35,39

[8] В. Григорьев, Д. Разумова. «Даты рождений и православное мировоззрение у пользователей сети ВКонтакте», *Демографическое обозрение*, **4** (2017), с. 110–120. ↑36

[9] P. Sobhani, D. Inkpen, X. Zhu. “Exploring deep neural networks for multitarget stance detection”, *Computational Intelligence*, **35**:1 (2019), pp. 82–97. ↑37

[10] M. Hardalov, A. Arora, P. Nakov, I. Augenstein. *Cross-domain label-adaptive stance detection*, 2021, 18 pp. arXiv 2104.07467 ↑37

[11] C. Conforti, J. Berndt, M. T. Pilehvar, C. Giannitsarou, F. Toxvaerd, N. Collier. *Will-They-Won't-They: A very large dataset for stance detection on twitter*, 2020, 10 pp. arXiv 2005.00388 ↑37

[12] J. Vamvas, R. Sennrich. *X-stance: A multilingual multi-target dataset for stance detection*, 2020, 12 pp. arXiv 2003.08385 ↑37

[13] L. Miao, M. Last, M. Litvak. “Twitter data augmentation for monitoring public opinion on COVID-19 intervention measures”, *Proceedings of the 1st Workshop on NLP for COVID-19*. V. 2, EMNLP 2020, 2020, 7 pp. ↑38

[14] C. Zong, F. Xia, W. Li, R. Navigli (eds.). *Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing*. V. 1: *Long Papers*, ACL, 2021 ↑38

[15] D. T. Huerta, J. Hawkins, J. Brownstein, Y. Hswen. “Exploring discussions of health and risk and public sentiment in MA during COVID-19 pandemic mandate implementation: A twitter analysis”, *SSM-Population Health*, **15**:1 (2021), 100851, 9 pp. ↑38

[16] S. Abosedra, N. T. Laopodis, A. Fakih. “Dynamics and asymmetries between consumer sentiment and consumption in pre-and during-COVID-19 time: evidence from the US”, *The Journal of Economic Asymmetries*, **24** (2021), e00227. ↑38

[17] K. S. Hasan, V. Ng. "Stance classification of ideological debates: data, models, features, and constraints", *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, Asian Federation of Natural Language Processing, 2013, pp. 1348–1356. ↑38

[18] E. Sharma, K. Saha, S. K. Ernala, S. Ghoshal, M. De Choudhury. "Analyzing ideological discourse on social media: A case study of the abortion debate", *Proceedings of the 2017 International Conference of The Computational Social Science Society of the Americas*, CSS 2017 (October 19–22, 2017, Santa Fe, NM, USA), ACM, 2017, ISBN 978-1-4503-5269-7, 8 pp. ↑38

[19] K. J. LaRoche, K. N. Jozkowski, B. L. Crawford, K. R. Haus. "Attitudes of US adults toward using telemedicine to prescribe medication abortion during COVID-19: A mixed methods study", *Contraception*, **104**:1 (2021), pp. 104–110. ↑38

[20] P. R. Roldán-Robles, A. C. Umaquinga-Criollo, J. A. García-Santillán, I. D. Herrera-Granda, á D. García-Santillán. "A conceptual architecture for content analysis about abortion using the twitter platform", *Revista Ibérica de Sistemas e Tecnologias de Informação*, 2019, no. E22, pp. 363–374. ↑38

[21] N. Hopkins, S. Zeedyk, F. Raitt. "Visualising abortion: emotion discourse and fetal imagery in a contemporary abortion debate", *Social Science & Medicine*, **61**:2 (2005), pp. 393–403. ↑38

[22] E. Ntontis, N. Hopkins. "Framing a 'social problem': emotion in anti-abortion activists' depiction of the abortion debate", *British Journal of Social Psychology*, **57**:3 (2018), pp. 666–683. ↑38

[23] D.I.H. Farías, M. Lai, L. Mencarini, M. Mozzachiodi, V. Patti, E. Sulis, D. Vignoli. "Happy parents' tweet? An exploration of 3 million Italian Twitter data", 2017 International Population Conference (29 October–04 November 2017, Cape Town, South Africa), 2017, 5722, 4 pp. ↑38

[24] Z. Shah, P. Martin, E. Coiera, K. D. Mandl, A. G. Dunn. "Modeling spatiotemporal factors associated with sentiment on Twitter: synthesis and suggestions for improving the identification of localized deviations", *Journal of Medical Internet Research*, **21**:5 (2019), e12881. ↑38

[25] B. Mandel, A. Culotta, J. Boulahanis, D. Stark, B. Lewis, J. Rodrigue. "A demographic analysis of online sentiment during hurricane Irene", *Proceedings of the Second Workshop on Language in Social Media*, LSM 2012, 2012, pp. 27–36. ↑38

[26] T. Daudert. "Exploiting textual and relationship information for fine-grained financial sentiment analysis", *Knowledge-Based Systems*, **230** (2021), 107389, 12 pp. ↑38

[27] J. Devlin, M. Chang, K. Lee, K. Toutanova. *Bert: pre-training of deep bidirectional transformers for language understanding*, 2018, 14 pp. arXiv 1810.04805 ↑38

[28] S. Ghosh, P. Singhania, S. Singh, K. Rudra, S. Ghosh. "Stance detection in web and social media: a comparative study", International Conference of The Cross-Language Evaluation Forum for European Languages, Lecture Notes in Computer Science, vol. **11696**, Springer, Cham, 2019, ISBN 978-3-030-28577-7, pp. 75–87. ↑38

[29] N. Loukachevitch, Y. Rubtsova. "Entity-oriented sentiment analysis of tweets: results and problems", International Conference on Text, Speech, and Dialogue, Lecture Notes in Computer Science, vol. **9302**, Springer, Cham, 2015, ISBN 978-3-319-24033-6, pp. 551–559. ↑38

[30] A. Golubev, N. Loukachevitch. "Improving results on Russian sentiment datasets", Conference on Artificial Intelligence and Natural Language (October 7–9, 2020, Helsinki, Finland), Communications in Computer and Information Science, vol. **1292**, Springer, Cham, 2020, ISBN 978-3-030-59081-9, pp. 109–121. ↑38

[31] A. Golubev, N. Loukachevitch. "Multi-Step transfer learning for sentiment analysis", International Conference on Applications of Natural Language to Information Systems, Lecture Notes in Computer Science, vol. **12801**, Springer, Cham, 2021, ISBN 978-3-030-80599-9, pp. 209–217. ↑38

[32] S. Smetanin, M. Komarov. "Deep transfer learning baselines for sentiment analysis in Russian", *Information Processing & Management*, **58**:3 (2021), 102484, 19 pp. ↑38

[33] Y. Kuratov, M. Arkhipov. *Adaptation of deep bidirectional multilingual transformers for Russian language*, 2019, 8 pp. arXiv 1905.07213 ↑39

[34] C. Sun, L. Huang, X. Qiu. *Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence*, 2019, 6 pp. arXiv 1903.09588 ↑39

[35] S. V. Vychegzhanin, E. V. Kotelnikov. "Stance detection in Russian: a feature selection and machine learning based approach", *Supplementary Proceedings of AIST 2017*, CEUR Workshop Proceedings, vol. **1975**, 2017, pp. 166–177. ↑39

[36] E. Pronoza, P. Panicheva, O. Koltsova, P. Rosso. "Detecting ethnicity-targeted hate speech in Russian social media texts", *Information Processing & Management*, **58**:6 (2021), 102674. ↑39

[37] M. B. Nasreen Taj, G. Girisha. "Insights of strength and weakness of evolving methodologies of sentiment analysis", *Global Transitions Proceedings*, **2**:2 (2021), pp. 157–162. ↑39

[38] D. J. Van de Kaa. "Europe's second demographic transition", *Population Bulletin*, **42**:1 (1987), pp. 1–59. ↑41

[39] А. Н. Расходчиков. Как управлять неуправляемым? *Сети 4.0. Управление сложностью*, ВЦИОМ, 2020, ISBN 978-5-906345-24-0, с. 12–17. ↑41,45

[40] А. О. Макаренцева, Н. И. Галиева, Д. М. Рогозин. «(Не) желание иметь детей в зеркале опросов населения», *Мониторинг общественного мнения: экономические и социальные перемены*, 2021, №4. ↑45

[41] I. E. Kalabikhina, E. P. Banin. "Database "Childfree (antinatalist) communities in the social network VKontakte"", *Population and Economics*, **5**:2 (2021), pp. 92–96. ↑42

[42] I. E. Kalabikhina, E. P. Banin. "Database "Pro-family (pronatalist) communities in the social network VKontakte"", *Population and Economics*, **4**:3 (2020), pp. 98–103. ↑42

Поступила в редакцию 10.11.2021
Переработана 14.12.2021
Опубликована 23.12.2021

Рекомендовал к публикации *к.т.н. Е. П. Куршев*

Пример ссылки на эту публикацию:

И. Е. Калабихина, Н. В. Лукашевич, Е. П. Банин, К. В. Алибаева, С. М. Ребрей. «Автоматическое извлечение мнений пользователей социальных сетей по вопросам репродуктивного поведения». *Программные системы: теория и приложения*, 2021, **12**:4(51), с. 33–63.
10.25209/2079-3316-2021-12-4-33-63
http://psta.psiras.ru/read/psta2021_4_33-63.pdf

Об авторах:

Ирина Евгеньевна Калабихина
заведующая кафедрой народонаселения экономического факультета МГУ имени М.В. Ломоносова, доктор экономических наук, главный редактор журнала Population and Economics; исследовательские интересы: измерение демографического поведения и результативности социальной и демографической политики.

0000-0002-3958-6630
e-mail: ikalabikhina@yandex.ru

Наталья Валентиновна Лукашевич

ведущий научный сотрудник НИВЦ МГУ имени М.В. Ломоносова, профессор филологического факультета МГУ имени М.В. Ломоносова, доктор технических наук; научные интересы: автоматическая обработка текстов, представление знаний, онтологии.

0000-0002-1883-4121
e-mail: louk_nat@mail.ru

Евгений Петрович Банин

инженер-исследователь НИЦ «Курчатовский институт»; исследовательские интересы: алгоритмы машинного обучения в области NLP, сбор и обработка данных, исследование социальных сетей.

0000-0002-7006-2990
e-mail: bonziinc@mail.ru

Камила Винеровна Алибаева

студентка факультета вычислительной математики и кибернетики МГУ имени М.В. Ломоосова; научные интересы: классификация текстов, извлечение мнений, вопросно-ответные системы.

0000-0002-0047-907X
e-mail: camalibi@yandex.ru

Софья Михайловна Ребрей

доцент кафедры мировой экономики МГИМО МИД России, кандидат экономических наук, заместитель главного редакторв научного журнала "Мировое и национальное хозяйство"; научные интересы: социально-экономические факторы и последствия гендерного неравенства.

0000-0002-6244-4497
e-mail: sofiarebrey@gmail.com

CSCSTI 06.01.29
UDC 519.689.3:007.51

Irina E. Kalabikhina, Natalia V. Loukachevitch, Eugene P. Banin, Kamila V. Alibaeva, Sofia M. Rebrey. *Automatic extraction of social network users' attitudes on reproductive behavior issues.*

Abstract. This paper presents a specialized dataset with annotation of user attitudes on reproductive behavior. We analyze the features of the "for" and "against" stance distribution for specific aspects of reproductive behavior. The created dataset solves two classification problems: classifying messages by the relevance to a topic being studied and the author's stance on a particular issue. We use classical machine learning methods and the BERT-based neural network classified messages models. The best classification results in both tasks are achieved based on variants of the BERT model using pairs of sentences in the classification — variants of NLI (natural language inference) and QA (question-answering). In addition, the created dataset makes it possible to draw meaningful conclusions on the attitudes of VKontakte users to reproductive behavior issues. It was revealed that the phenomenon of deliberate childlessness is actively represented in VKontakte groups while having many children remains a poorly widespread model of behavior. Within the framework of the pro-natalist policy, it is crucial to form a favorable public opinion about parenting, to alleviate the deficiency of time for parents.

Key words and phrases: opinion analysis, BERT, supervised learning, demographic policy, VKontakte, reproductive behavior.

2020 *Mathematics Subject Classification*: 97P30; 97P20, 97R40

References

[1] F. A. Pozzi, E. Fersini, E. Messina, B. Liu. "Challenges of sentiment analysis in social networks: an overview", *Sentiment Analysis in Social Networks*, Elsevier, 2017, ISBN 978-0-12-804412-4, pp. 1–11. ↑34,36

[2] B. Liu. "Sentiment analysis and opinion mining", *Synthesis Lectures on Human Language Technologies*, **5**:1 (2012), pp. 1–167. ↑34,36,38

The work was a part of the research work "Reproduction of the population in the context of socio-economic development" AAAA-A17-117062610054-1)[1]

This research has been supported by the Interdisciplinary Scientific and Educational School of Moscow University «Brain, Cognitive Systems, Artificial Intelligence»[1,2]

The work was supported by the Russian Science Foundation grant 21-71-30003 in terms of creating the annotation interface, annotation itself, and methodology for applying machine learning methods[2]

© I. E. Kalabikhina[1], N. V. Loukachevitch[2], E. P. Banin[3], K. V. Alibaeva[4], S. M. Rebrey[5], 2021
© Lomonosov Moscow State University[1,2,4], 2021
© National Research Center "Kurchatov Institute"[3], 2021
© Moscow State Institute of International Relations (MGIMO)[5], 2021
© Program Systems: Theory and Applications (design), 2021

10.25209/2079-3316-2021-12-4-33-63

[3] S. Mohammad, S. Kiritchenko, P. Sobhani, X. Zhu, C. Cherry. "Semeval-2016 task 6: detecting stance in tweets", *Proceedings of the 10th International Workshop on Semantic Evaluation*, SemEval-2016 (June 2016, San Diego, California), ACL, 2016, pp. 31–41. ↑34,37,38

[4] S. V. Vychegzhanin, E. V. Kotelnikov. "Stance detection based on ensembles of classifiers", *Programming and Computer Software*, **45**:5 (2019), pp. 228–240. ↑34,39

[5] D. Küçük, F. Can. "Stance detection: A survey", *ACM Computing Surveys*, **53**:1 (2021), 12, 37 pp. ↑34

[6] *Demografic Dictionary*, ed. Valentay D.I., Sovetskaya encyclopediya, M., 1985 (in Russian), 608 pp. ↑34

[7] I. E. Kalabikhina, E. P. Banin, I. A. Abduselimova, G. A. Klimenko, A. V. Kolotusha. "The measurement of demographic temperature using the sentiment analysis of data from the social network VKontakte", *Mathematics*, **9**:9 (2021), 987, 25 pp. ↑35,39

[8] V. Grigor'yev, D. Razumova. "Orthodox self-identification and the distribution of birthdays of VK users", *Demographic Review*, **4** (2017), pp. 110–120 (in Russian). ↑36

[9] P. Sobhani, D. Inkpen, X. Zhu. "Exploring deep neural networks for multitarget stance detection", *Computational Intelligence*, **35**:1 (2019), pp. 82–97. ↑37

[10] M. Hardalov, A. Arora, P. Nakov, I. Augenstein. *Cross-domain label-adaptive stance detection*, 2021, 18 pp. arXiv 2104.07467 ↑37

[11] C. Conforti, J. Berndt, M. T. Pilehvar, C. Giannitsarou, F. Toxvaerd, N. Collier. *Will-They-Won't-They: A very large dataset for stance detection on twitter*, 2020, 10 pp. arXiv 2005.00388 ↑37

[12] J. Vamvas, R. Sennrich. *X-stance: A multilingual multi-target dataset for stance detection*, 2020, 12 pp. arXiv 2003.08385 ↑37

[13] L. Miao, M. Last, M. Litvak. "Twitter data augmentation for monitoring public opinion on COVID-19 intervention measures", *Proceedings of the 1st Workshop on NLP for COVID-19*. V. 2, EMNLP 2020, 2020, 7 pp. ↑38

[14] C. Zong, F. Xia, W. Li, R. Navigli (eds.). *Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing*. V. 1: *Long Papers*, ACL, 2021 ↑38

[15] D. T. Huerta, J. Hawkins, J. Brownstein, Y. Hswen. "Exploring discussions of health and risk and public sentiment in MA during COVID-19 pandemic mandate implementation: A twitter analysis", *SSM-Population Health*, **15**:1 (2021), 100851, 9 pp. ↑38

[16] S. Abosedra, N. T. Laopodis, A. Fakih. "Dynamics and asymmetries between consumer sentiment and consumption in pre-and during-COVID-19 time: evidence from the US", *The Journal of Economic Asymmetries*, **24** (2021), e00227. ↑38

[17] K. S. Hasan, V. Ng. "Stance classification of ideological debates: data, models, features, and constraints", *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, Asian Federation of Natural Language Processing, 2013, pp. 1348–1356. ↑38

[18] E. Sharma, K. Saha, S. K. Ernala, S. Ghoshal, M. De Choudhury. "Analyzing ideological discourse on social media: A case study of the abortion debate",

Proceedings of the 2017 International Conference of The Computational Social Science Society of the Americas, CSS 2017 (October 19–22, 2017, Santa Fe, NM, USA), ACM, 2017, ISBN 978-1-4503-5269-7, 8 pp. ↑38

[19] K. J. LaRoche, K. N. Jozkowski, B. L. Crawford, K. R. Haus. "Attitudes of US adults toward using telemedicine to prescribe medication abortion during COVID-19: A mixed methods study", *Contraception*, **104**:1 (2021), pp. 104–110. ↑38

[20] P. R. Roldán-Robles, A. C. Umaquinga-Criollo, J. A. García-Santillán, I. D. Herrera-Granda, á D. García-Santillán. "A conceptual architecture for content analysis about abortion using the twitter platform", *Revista Ibérica de Sistemas e Tecnologias de Informação*, 2019, no. E22, pp. 363–374. ↑38

[21] N. Hopkins, S. Zeedyk, F. Raitt. "Visualising abortion: emotion discourse and fetal imagery in a contemporary abortion debate", *Social Science & Medicine*, **61**:2 (2005), pp. 393–403. ↑38

[22] E. Ntontis, N. Hopkins. "Framing a 'social problem': emotion in anti-abortion activists' depiction of the abortion debate", *British Journal of Social Psychology*, **57**:3 (2018), pp. 666–683. ↑38

[23] D.I.H. Farías, M. Lai, L. Mencarini, M. Mozzachiodi, V. Patti, E. Sulis, D. Vignoli. "Happy parents' tweet? An exploration of 3 million Italian Twitter data", 2017 International Population Conference (29 October–04 November 2017, Cape Town, South Africa), 2017, 5722, 4 pp. ↑38

[24] Z. Shah, P. Martin, E. Coiera, K. D. Mandl, A. G. Dunn. "Modeling spatiotemporal factors associated with sentiment on Twitter: synthesis and suggestions for improving the identification of localized deviations", *Journal of Medical Internet Research*, **21**:5 (2019), e12881. ↑38

[25] B. Mandel, A. Culotta, J. Boulahanis, D. Stark, B. Lewis, J. Rodrigue. "A demographic analysis of online sentiment during hurricane Irene", *Proceedings of the Second Workshop on Language in Social Media*, LSM 2012, 2012, pp. 27–36. ↑38

[26] T. Daudert. "Exploiting textual and relationship information for fine-grained financial sentiment analysis", *Knowledge-Based Systems*, **230** (2021), 107389, 12 pp. ↑38

[27] J. Devlin, M. Chang, K. Lee, K. Toutanova. *Bert: pre-training of deep bidirectional transformers for language understanding*, 2018, 14 pp. arXiv:1810.04805 ↑38

[28] S. Ghosh, P. Singhania, S. Singh, K. Rudra, S. Ghosh. "Stance detection in web and social media: a comparative study", International Conference of The Cross-Language Evaluation Forum for European Languages, Lecture Notes in Computer Science, vol. **11696**, Springer, Cham, 2019, ISBN 978-3-030-28577-7, pp. 75–87. ↑38

[29] N. Loukachevitch, Y. Rubtsova. "Entity-oriented sentiment analysis of tweets: results and problems", International Conference on Text, Speech, and Dialogue, Lecture Notes in Computer Science, vol. **9302**, Springer, Cham, 2015, ISBN 978-3-319-24033-6, pp. 551–559. ↑38

[30] A. Golubev, N. Loukachevitch. "Improving results on Russian sentiment datasets", Conference on Artificial Intelligence and Natural Language (October 7–9, 2020, Helsinki, Finland), Communications in Computer and Information Science, vol. **1292**, Springer, Cham, 2020, ISBN 978-3-030-59081-9, pp. 109–121. ↑38

[31] A. Golubev, N. Loukachevitch. "Multi-Step transfer learning for sentiment analysis", International Conference on Applications of Natural Language to Information Systems, Lecture Notes in Computer Science, vol. **12801**, Springer, Cham, 2021, ISBN 978-3-030-80599-9, pp. 209–217. ↑38

[32] S. Smetanin, M. Komarov. "Deep transfer learning baselines for sentiment analysis in Russian", *Information Processing & Management*, **58**:3 (2021), 102484, 19 pp. ↑38

[33] Y. Kuratov, M. Arkhipov. *Adaptation of deep bidirectional multilingual transformers for Russian language*, 2019, 8 pp. arXiv 1905.07213↑39

[34] C. Sun, L. Huang, X. Qiu. *Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence*, 2019, 6 pp. arXiv 1903.09588↑39

[35] S. V. Vychegzhanin, E. V. Kotelnikov. "Stance detection in Russian: a feature selection and machine learning based approach", *Supplementary Proceedings of AIST 2017*, CEUR Workshop Proceedings, vol. **1975**, 2017, pp. 166–177. ↑39

[36] E. Pronoza, P. Panicheva, O. Koltsova, P. Rosso. "Detecting ethnicity-targeted hate speech in Russian social media texts", *Information Processing & Management*, **58**:6 (2021), 102674. ↑39

[37] M. B. Nasreen Taj, G. Girisha. "Insights of strength and weakness of evolving methodologies of sentiment analysis", *Global Transitions Proceedings*, **2**:2 (2021), pp. 157–162. ↑39

[38] D. J. Van de Kaa. "Europe's second demographic transition", *Population Bulletin*, **42**:1 (1987), pp. 1–59. ↑41

[39] A. N. Raskhodchikov. "How to manage unmanaged?", *Seti 4.0. Upravleniye slozhnost'yu*, VTsIOM, 2020, ISBN 978-5-906345-24-0, pp. 12–17 (in Russian). ↑41,45

[40] A. O. Makarentseva, N. I. Galiyeva, D. M. Rogozin. "Desire (Not) To Have Children in the Population Surveys", *The Monitoring of Public Opinion: Economic and Social Changes Journal*, 2021, no. 4 (in Russian). ↑45

[41] I. E. Kalabikhina, E. P. Banin. "Database "Childfree (antinatalist) communities in the social network VKontakte"", *Population and Economics*, **5**:2 (2021), pp. 92–96. ↑42

[42] I. E. Kalabikhina, E. P. Banin. "Database "Pro-family (pronatalist) communities in the social network VKontakte"", *Population and Economics*, **4**:3 (2020), pp. 98–103. ↑42

Sample citation of this publication:

Irina E. Kalabikhina, Natalia V. Loukachevitch, Eugene P. Banin, Kamila V. Alibaeva, Sofia M. Rebrey. "Automatic extraction of social network users' attitudes on reproductive behavior issues". *Program Systems: Theory and Applications*, 2021, **12**:4(51), pp. 33–63. (*In Russian*). 10.25209/2079-3316-2021-12-4-33-63

http://psta.psiras.ru/read/psta2021_4_33-63.pdf