

О ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ И МАШИННОГО ОБУЧЕНИЯ ДЛЯ ВЫЯВЛЕНИЯ ОБСТРУКЦИОННЫХ ПРИЕМОВ В КОММУНИКАЦИЯХ ПРОЕКТОВ ПО СОЗДАНИЮ СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ ДЛЯ ИНДУСТРИИ 4.0

USING NEURAL NETWORKS AND MACHINE LEARNING TO IDENTIFY OBSTRUCTIVE TECHNIQUES IN COMMUNICATIONS IN LARGE SCALE SYSTEMS IMPLEMENTATION PROJECTS FOR INDUSTRY 4

S. Smirnov

Summary. Less than a third of the global volume of completed projects are completely successful. As one of the key problems that actually kill projects, the problem of poor-quality project communications is described in the literature. The paper presents the results of the analysis of the possibility of using neural networks and machine learning methods to automate the recognition of obstructive techniques in project communications, which was not previously described in the literature.

Keywords: project management, project communications, project risks, obstructive techniques, neural networks, machine learning.

Смирнов Сергей Павлович

Аспирант, Российский технологический университет МИРЭА (г. Москва); Руководитель департамента управления проектами, ПАО «Вымпелком» (г. Москва)
smirnov.s.p1@edu.mirea.ru

Аннотация. Менее трети мирового объема выполняемых проектов являются полностью успешными. В качестве одной из ключевых проблем, фактически убивающих проекты, в литературе описана проблема некачественных коммуникаций проекта. В работе представлены результаты анализа возможности применения нейронных сетей и методов машинного обучения для автоматизации распознавания obstructive приемов в проектных коммуникациях, что ранее не описывалось в литературе.

Ключевые слова: управление проектами, проектные коммуникации, риски проектов, obstructive приемы, нейронные сети, машинное обучение.

Введение

Ежегодно в мире инициируется и выполняется более миллиона проектов общей стоимостью более триллиона долларов США и лишь 30% из них завершаются успешно, т.е. созданием продукта или услуги с ожидаемыми свойствами, в установленное время, в рамках установленного бюджета. Более половины проектов считаются условно успешными: полученный результат частично удовлетворяет потребностям и требуемым характеристикам; и-или срок выполнения проекта превысил ожидания, но оказался приемлем; и-или бюджет проекта превышен, но итоговая стоимость принята заказчиком проекта. Около 20% проектов, в среднем, признаются полностью неуспешными [1]. Отмечается существенная обратная корреляция объема проекта и вероятности его успеха: если для небольших, или относительно небольших, проектов процент полной неуспешности составляет 4–5%, то для больших проектов (стоимость трудозатрат проекта сто миллионов рублей и более) эта величина достигает уже 38% [1].

Для увеличения вероятности успешного получения необходимого результата для сложных систем успешно применяются различные техники и технологии, например, в качестве наиболее распространенных можно назвать методологии управления проектами (PMBOK, PRINCE2, Agile/Scrum, Agile/Kanban, Agile/Lean) и Системную инженерию. Метод Системной инженерии охватывает более широкий спектр вопросов, чем методологии управления проектами, поскольку Системная инженерия, в первую очередь, вводит понятие и рассматривает полный жизненный цикл результата (изделия, продукта, услуги), в то время как проектная деятельность (в силу заданного определением проекта границ времени существования), ограничивается процессом получения результата и не отвечает на вопросы как полученный продукт будет эксплуатироваться, развиваться, улучшаться и заменяться следующими поколениями продукта или другими продуктами. Тем не менее, и метод Системной инженерии и методологии управления проектами имеют дело с одним и тем же негативным воздействием факторов окружения, отличающимся в деталях. По результа-

там анализа литературных источников можно привести следующие статистические данные о причинах неудач проектов по данным ведущих организаций, занимающихся консалтингом и методологией в области проектного управления:

1. 46% компаний признают, что недооценили значимость и необходимость управления проектами в части управления ресурсами [2];
2. 27% проектов обычно выходят за рамки бюджета [3];
3. 41% малоэффективных компаний и 17% чрезвычайно эффективных компаний отмечают, что ключевой причиной неуспешности проектов является недостаточная поддержка проекта ключевыми заинтересованными сторонами, в том числе в контексте авторизации и выделения ресурсов для выполнения проекта [4];
4. 47% проектов имеют проблемы с достижением заявленных целей из-за существенных недостатков в управлении требованиями [5];
5. 38% организаций полагают, что наиболее значимым барьером на пути к успеху проекта является непонимание задач и разделения ответственности команды проекта [6];
6. 80% организаций отмечают, что по крайней мере половина всех трудозатрат приходится на переделки [6];
7. 37% организации отмечают, что параллельное выполнение слишком большого количества проектов ведет к существенному увеличению вероятности провала, в том числе из-за конфликта с доступностью ресурсов [7];
8. 50% проектов перерасходуют запланированный бюджет как минимум на 1 миллион долларов [8].

Из сказанного выше, можно заключить, что управление ресурсами (включающее в себя управление взаимодействием с заинтересованными сторонами, идентификацию трудоемкости и потребности в ресурсах, планирование использования ресурсов и планирование решения конфликтов, управление рисками и оказание влияния на владельцев ресурсов) является ключевым фактором для успешного выполнения проекта. При этом, необходимо отметить, что большинство применяемых методов и стандартов управления проектами [9][10] в части управления ресурсами основываются на нечеткой логике, субъективных суждениях, а также навыках ведения переговоров у менеджера проекта, что создает отдельное множество рисков, связанных с возможной ошибкой.

В области исследований социальных взаимодействий описано не очень большое количество моделей [11]: модель ДеГрута [12], модель Фридкина-Джонсона [13] и модель Краузе [14]. Модель ДеГрута [11], считаю-

щаяся на данный момент уже классической, рассматривает переход системы мнений из состояния А в состояние В. В модели рассматривается пошаговый алгоритм при котором каждый участник одновременно с другими изменяет свое мнение, используя взвешенное среднее значение мнений других участников. Ключевым недостатком модели является предположение об однородности лиц принимающих решение (агентов). Также недостатком модели является отсутствие возможности понять (учесть в модели) близость мнения индивидуума к истине — модель не позволяет учитывать влияние внешних факторов или взаимодействие мнений. В работе [15] показано, что модель ДеГрута использует чересчур упрощенный (наивный) подход к определению мнения и не может быть применена для решения практических задач. В то же время, в работе [16] отмечается, что существуют определенные виды сетей, с помощью которых, в ряде случаев, удается достигнуть приемлемого решения. В работе [17] представлено расширение модели ДеГрута путем анализа влияния гетерогенности агентов для процесс принятия решения. Для этого рассматривается упрощенная модель, учитывающая два вида агентов: конформистов, стремящихся к достижению результата, и противников — стремящихся к противоположному общему мнению. Другие расширения и модификации модели ДеГрута представлены в работах [18–32], в том числе в работах [29][30] рассматривается влияние внутренних факторов, таких как психоэмоциональные особенности человека, на достижение общего консенсуса при принятии решения в группе.

Модель Фридкина-Джонсона [13] фактически является расширением модели ДеГрута, в которую вводится некоторая устойчивость мнения агента к изменению (упрямство). Модель Краузе [14] определяет множество соседей по-другому. Каждый агент принимает в качестве своих соседей тех индивидуумов, чьи мнения отличаются от него не более чем на определенный уровень доверия. В результате, оказывается, что модель Краузе нелинейна, в то время как первые модели ДеГрута и Фридкина относятся к линейным.

Ключевым недостатком всех перечисленных моделей является необходимость нескольких итераций для достижения целевого состояния или близкого к нему. Это приводит к существенной неопределенности, поскольку переходные состояния системы, по факту, не являются статическими и, де-факто, могут быть изменены между выходом N-ной итерации и входом N+1-й итерации в совершенно произвольном порядке. Например, в работе [33], в которой с использованием методов математической биологии, путем переноса модели совместной миграции живых организмов на миграцию мнений в информационно-коммуникационном пространстве, рассматривается аспект, фактически определяющий

ограниченную применимость многоитерационных методов достижения консенсуса: перегрузка механизмов восприятия человека приводит к тому, что чем большим количеством информации обладает каждый член группы, тем большие противоречия будут формироваться и мешать достижению общего единого мнения. В этой связи достижение соглашения за один шаг выглядит более перспективным, чем решение этой же задачи за несколько итераций.

В ранее проведенной работе [34] представлены результаты анализа нескольких сотен коммуникаций 48 проектов одного направления деятельности с целью выявления приемов недружественных коммуникаций, оказывающих негативное влияние на проект, описана методика проведения анализа коммуникаций и анализ большой выборки коммуникаций, описаны исходные данные для анализа и критерии их отбора, выявлены наиболее часто встречающиеся методы недружественной коммуникации в проектах. При этом предполагается, что полученные результаты могут быть использованы менеджерами проектов для идентификации приемов недружественной коммуникации в ручном режиме (идентификация с использованием собственного опыта). Ключевым недостатком такого метода является высокая чувствительность к субъективности суждений менеджера в силу его опыта, настроения или иных факторов.

В развитие алгоритма и метода, предложенных в работе [34], в настоящей работе представлен метод, который можно использовать для снижения степени субъективности суждения менеджера при идентификации приемов недружественной коммуникации.

Подходы и материалы

При проектировании метода сделаны следующие допущения и предположения:

1. Метод предполагает использование менеджером проекта или иным заинтересованным лицом при подготовке встреч.
2. Метод предполагает использование менеджером проекта или иным заинтересованным лицом для анализа конструктивности встреч.
3. Метод предполагает использование менеджером проекта для разработки плана коррекционных мероприятий по итогам встреч.
4. Разрабатываемый метод не рассчитан на онлайн применение в процессе проведения встреч.
5. Метод не предполагает непосредственную работу с сырыми материалами встреч — презентациями, письмами, записями.
6. Метод предполагает работу с предварительно обработанными источниками данных.

7. На данном этапе развития, метод не предполагает возможности идентификации всех выявленных в [34] приемов недружественной коммуникации.

Поиск возможных источников данных для обучения [35–55] показал, что в открытом доступе готовых наборов данных (datasets), содержащих расшифровки совещаний в технической области, не найдено.

Выдвинута гипотеза о возможности выявления определенного подмножества методов недружественных коммуникаций путем анализа эмоционального статуса коммуникации. Таким образом, можно предположить возможность идентификации таких обструкционных приемов, описанных в [34], как «демонизация явления или предмета», «демонизация оппонента», «оскорбительный стиль», «переход на личности». По экспертному мнению автора, указанные приемы недружественной коммуникации встречаются достаточно часто (особенно в Enterprise бизнес-культуре или государственных предприятиях), а влияние на неэффективность встреч и на невозможность достижения консенсуса существенное, поэтому настройка модели на выявление указанных приемов представляется целесообразной.

В качестве исходных данных для обучения модели с целью определения эмоционального статуса коммуникации использована выборка маркированных токсичных (имеющих негативную эмоциональную окраску) комментариев на русском языке [56] в объеме 14412 единиц.

В качестве инструмента для проведения эксперимента выбрана площадка Kaggle (www.kaggle.com) — облачная платформа, поддерживающая Jupiter Notebook, интегрированная с платформой Google, широко используемая для простых экспериментов. Язык программирования — Python.

Выборка тестирования сделана вручную из материалов, проанализированных ранее в работе [34] и включает в себя 3 протокола, 3 электронных письма и 3 расшифровки встречи.

Результаты

В результате работы получена и размещена в публичном доступе в сети интернет обучающая выборка [57].

Проведен анализ выборки с точки зрения ее качества для обучения. Обучающая выборка содержит 35% нетоксичных и 65% токсичных комментариев. Разработанный на Python инструмент для анализа размещен в публичном доступе в сети интернет [58].

Реализован стенд на основе Convolution Network для обучения и дальнейшего анализа тестовой выборки. Тестовые данные размещены внутри инструмента [59].

Получена приемлемая корреляция между экспертным ожиданием и результатами работы конволюционной нейросети на тестовых данных.

Выводы

В рамках работы выдвинута гипотеза о возможности выявления определенного подмножества методов недружественных коммуникаций путем анализа эмоционального статуса коммуникации.

Осуществлено экспериментальное подтверждение гипотезы, для чего разработано модельное программное обеспечение на Python и опубликовано в сети Интернет, выбрана подходящая для анализа эмоционального статуса выборка достаточного для обучения модели объема, подтверждено совпадение экспертной оценки тестовой выборки и полученных результатов.

В результате длительного поиска источников данных для обучения показано, что готовых наборов размеченных данных по расшифровке совещаний, посвященных разработке ПО или выполнению ИТ-проектов, не выявлено в общедоступных источниках. Предположительно, это может быть связано с двумя факторами: 1. отсутствием научно-исследовательской составляющей в проектной деятельности компаний реального сектора экономики; 2. раскрытие информации может нарушать как коммерческую тайну, так и требования по защите персональных данных.

Наборов данных на русском языке в общем количестве выборки данных менее 1%.

Примененная обучающая выборка может быть рассмотрена как best of breed, но является специфичной для общения в Интернет, хотя и показала себя применимой для обучения с целью дальнейшего анализа корпоративных коммуникаций.

Таким образом, подтверждена возможность использования нейросетей и машинного обучения для выявления обструкционных приемов в коммуникациях в ИТ-проектах

Направление дальнейших работ по данной тематике автор формулирует следующим образом:

1. Переход на более аккуратные выборки для обучения [60];
2. Маркировка существующих материалов, полученных в работе [34] в разрезе множества идентифицируемых приемов обструкции;
3. Разработка базы данных маркированных выборок данных для обучения моделей в разрезе различных предметных областей;
4. Разработка и внедрение портала общего доступа, в который можно будет загружать маркированные материалы других авторов;
5. Разработка сетей для идентификации других типов недружественных коммуникаций;
6. Сравнительный анализ различных форм нейронных сетей и их параметров для выбора наиболее подходящего набора сетей для классификации расшифровок для идентификации обструкционных приемов;

ЛИТЕРАТУРА

1. The Standish Group report. Chaos. 2011. [Электронный ресурс] Режим доступа <https://www.projectsmart.co.uk/white-papers/chaos-report.pdf>. Дата визита 01.01.2022.
2. PMI's pulse of the profession: The high cost of low performance. 2014. [Электронный ресурс] Режим доступа <https://www.pmi.org/-/media/pmi/documents/public/pdf/learning/thought-leadership/pulse/pulse-of-the-profession-2017.pdf>. Дата визита 01.01.2022.
3. PMI's pulse of the profession. 9th Global Project management survey. 2017. [Электронный ресурс] Режим доступа <https://www.pmi.org/-/media/pmi/documents/public/pdf/learning/thought-leadership/pulse/pulse-of-the-profession-2017.pdf>. Дата визита 01.01.2022.
4. PMI's pulse of the profession: Success in disruptive times. 2018. [Электронный ресурс] Режим доступа <https://www.pmi.org/-/media/pmi/documents/public/pdf/learning/thought-leadership/pulse/pulse-of-the-profession-2018.pdf>. Дата визита 01.01.2022.
5. PMI's pulse of the profession: Requirements management — a core competency for Project and Program success. 2014 [Электронный ресурс] Режим доступа <https://www.pmi.org/-/media/pmi/documents/public/pdf/learning/thought-leadership/pulse/requirements-management.pdf>. Дата визита 01.06.2020.
6. Geneca (2011). Doomed From the Start? Why a Majority of Business and IT Teams Anticipate Their Software Development Projects Will Fail. // Geneca. Oakbrook terrace, IL.
7. The state of project management. Annual survey. 2018. [Электронный ресурс] Режим доступа <https://www.wellington.co.uk/wp-content/uploads/2018/05/The-State-of-Project-Management-Survey-2018-FINAL.pdf>. Дата визита 01.06.2020.
8. Gartner survey shows why project fails. 2012. [Электронный ресурс] Режим доступа <https://thisiswhatgoodlookslike.com/2012/06/10/gartner-survey-shows-why-projects-fail/>. Дата визита 01.06.2020.
9. A guide to the project management body of knowledge (PMBOK guide), 6th edition. 2017/ Newton-square, PA: Project management institute, 2017.

10. The standart for project management. 2017. Newton-square, PA: Project management institute, 2017.
11. Xu, Zhi & Liu, Ji & Başar, Tamer. (2015). On a Modified DeGroot-Friedkin Model of Opinion Dynamics. Proceedings of the American Control Conference. 2015. DOI: 10.1109/ACC.2015.7170871.
12. Morris H. DeGroot (1974). Reaching a consensus. // Journal of American statistical association, Volume 69, Issue 345, pp. 118–121.
13. N. E. Friedkin and E.C. Johnsen. Social influence networks and opinion change. Advances in Group Processes, 16(1):1–29, 1999.
14. R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis, and simulation. Journal of Artificial Societies and Social Simulation, 5(3), 2002.
15. Sobel J, Economists' models of learning, Journal of Economic Theory, 2000, 94(2): 241–261.
16. Golub B and Jackson M O, Naive learning in social networks and the wisdom of crowds, American Economic Journal: Microeconomics, 2010, 2(1): 112–149.
17. Cao, Zhigang & Jiao, Fengliang & Qu, Xinglong & Wang, Wen-Xu & Yang, Mingmin & Yang, Xiaoguang & Zhang, Boyu. (2018). Rebels Lead to the Doctrine of the Mean: A Heterogeneous DeGroot Model. Journal of Systems Science and Complexity. 31. 1498–1509. DOI: 10.1007/s11424–018–7136–6.
18. Mandyam, Sridhar & Sridhar, Usha. (2011). Endogenous Control of DeGroot Learning. Social Network Analysis and Mining. 3. 22–35. DOI: 10.1007/s13278–013–0094-z.
19. Meza, David & Reyniers, Diane. (2013). Debiasing the Becker-degroot-marschak valuation mechanism. Economics Bulletin. 2. 1446–1456.
20. Han, Huawei & Qiang, Chengcang & Wang, Caiyun & Han, Jing. (2015). Intervention of DeGroot model by soft control. 1291–1296. DOI: 10.1109/ChiCC.2015.7259820.
21. Xia, Weiguo & Liu, Ji & Johansson, Karl & Basar, Tamer. (2016). Convergence rate of the modified DeGroot-Friedkin model with doubly stochastic relative interaction matrices. 1054–1059. 10.1109/ACC.2016.7525054.
22. Ye, Mengbin & Liu, Ji & Anderson, Brian & Yu, Changbin & Başar, Tamer. (2017). On the Analysis of the DeGroot-Friedkin Model with Dynamic Relative Interaction Matrices. IFAC-PapersOnLine. 50. DOI: 10.1016/j.ifacol.2017.08.1426.
23. Romeijn, Jan-Willem & Roy, Olivier. (2017). All agreed: Aumann meets DeGroot. Theory and Decision. DOI: 10.1007/s11238–017–9633–9.
24. Burfurd, Ingrid & Wilkening, Tom. (2018). Experimental guidance for eliciting beliefs with the Stochastic Becker–DeGroot–Marschak mechanism. Journal of the Economic Science Association. 4. 10.1007/s40881–018–0046–5.
25. Gyftopoulos, Sotirios & Efraimidis, Pavlos & Katsaros, Panagiotis. (2018). Formal Analysis of DeGroot Influence Problems using Probabilistic Model Checking. Simulation Modelling Practice and Theory. 89. 10.1016/j.simpat.2018.09.009.
26. Halder, Abhishek. (2019). DeGroot–Friedkin Map in Opinion Dynamics Is Mirror Descent. IEEE Control Systems Letters. PP. 1–1. 10.1109/LCSYS.2019.2900452.
27. Xue, Min & Fu, Chao & Yang, Shan-Lin. (2019). Group consensus reaching based on a combination of expert weight and expert reliability. Applied Mathematics and Computation. 369. 124902. DOI: 10.1016/j.amc.2019.124902.
28. Klickstein, Isaac & Pecora, Louis & Sorrentino, Francesco. (2019). Symmetry induced group consensus. Chaos: An Interdisciplinary Journal of Nonlinear Science. 29. 073101. DOI: 10.1063/1.5098335.
29. Gvirts, Hila. (2019). Alexithymia and Reaching Group Consensus. Cognition and Emotion. DOI: 10.1080/02699931.2019.1675600.
30. Shamay-Tsoory, Simone & Saporta, Nira & Marton, Inbar & Gvirts, Hila. (2019). Herding Brains: A Core Neural Mechanism for Social Alignment. Trends in Cognitive Sciences. 23. DOI: 10.1016/j.tics.2019.01.002.
31. Xue, Min & Fu, Chao & Yang, Shan-Lin. (2020). Dynamic Expert Reliability Based Feedback Mechanism in Consensus Reaching Process with Distributed Preference Relations. Group Decision and Negotiation. DOI: 10.1007/s10726–020–09660–8.
32. Fu, Chao & Chang, Wenjun & Yang, Shanlin. (2020). Multiple criteria group decision making based on group satisfaction. Information Sciences. 518. DOI: 10.1016/j.ins.2020.01.021.
33. Hiroki Sayama (2020). Enhanced ability of information gathering may intensify disagreement among groups. // Physical Review. E102. DOI: 10.1103/PhysRevE.102.012303
34. Смирнов С.П. (2021). Алгоритм и метод выявления обструкционных приемов в коммуникациях ИТ-проектов / Современная наука: актуальные проблемы теории и практики. Серия «Естественные и Технические науки», 2020 № 5–2, 88–92. DOI 10.37882/2223–2966.2021.05–2.25
35. The 50 Best Free Datasets for Machine Learning, Lionbridge AI, [Электронный ресурс] Режим доступа <https://lionbridge.ai/datasets/the-50-best-free-datasets-for-machine-learning/> Дата визита 01.01.2022.
36. Google Cloud Public Datasets, Google, [Электронный ресурс] Режим доступа <https://cloud.google.com/public-datasets/> Дата визита 01.01.2022
37. Machine Learning and AI Datasets, Carnegie Mellon University, [Электронный ресурс] Режим доступа <https://guides.library.cmu.edu/c.php?g=844845&p=6191907> Дата визита 01.01.2022.
38. Big Data and AI: 30 Amazing and Free Public Data Sources, Forbes, [Электронный ресурс] Режим доступа <https://www.forbes.com/sites/bernardmarr/2018/02/26/big-data-and-ai-30-amazing-and-free-public-data-sources-for-2018/#f3bdeb5f8aес> Дата визита 01.01.2022.
39. Awesome Autonomous Vehicles Datasets, Github, [Электронный ресурс] Режим доступа <https://github.com/takeitallsource/awesome-autonomous-vehicles#datasets> Дата визита 01.01.2022
40. Fueling the Gold Rush, The Greatest Public Datasets for AI, StartupGrind, [Электронный ресурс] Режим доступа <https://medium.com/startup-grind/fueling-the-ai-gold-rush-7ae438505bc2> Дата визита 01.01.2022
41. Places to Find Free Datasets for Data Science Projects, Dataquest, [Электронный ресурс] Режим доступа <https://www.dataquest.io/blog/free-datasets-for-projects/> Дата визита 01.01.2022
42. The Best Datasets for Natural Language Processing, Gengo AI, [Электронный ресурс] Режим доступа <https://gengo.ai/datasets/the-best-25-datasets-for-natural-language-processing/> Дата визита 01.01.2022

43. Awesome Public Datasets, Github, [Электронный ресурс] Режим доступа <https://github.com/awesomedata/awesome-public-datasets#machinelearning> Дата визита 01.01.2022
44. StatLib Datasets Archive, Carnegie Mellon, [Электронный ресурс] Режим доступа <http://lib.stat.cmu.edu/datasets/> Дата визита 01.01.2022
45. Institutional Research and Analysis, Common Datasets, [Электронный ресурс] Режим доступа <https://www.cmu.edu/ira/CDS/index.html> Дата визита 01.01.2022
46. Datasets and Project Suggestions, Andrew W. Moore, [Электронный ресурс] Режим доступа <http://www.cs.cmu.edu/~awm/15781/project/data.html> Дата визита 01.01.2022
47. Datasets, Machine Learning Repository, MIT | [Электронный ресурс] Режим доступа <https://ocw.mit.edu/courses/sloan-school-of-management/15-097-prediction-machine-learning-and-statistics-spring-2012/datasets/> Дата визита 01.01.2022
48. Datasets, MIT Lincoln Laboratory, [Электронный ресурс] Режим доступа <https://www.ll.mit.edu/r-d/datasets> Дата визита 01.01.2022
49. Stanford Large Network Dataset Collection, Stanford University [Электронный ресурс] Режим доступа <https://snap.stanford.edu/data/> Дата визита 01.01.2022
50. Stanford Common Dataset, Stanford University [Электронный ресурс] Режим доступа <https://snap.stanford.edu/data/> Дата визита 02.01.2022
51. Datalab, UC Berkeley [Электронный ресурс] Режим доступа <http://www.lib.berkeley.edu/libraries/data-lab> Дата визита 02.01.2022
52. Exploring Datasets, Data Science at Berkeley [Электронный ресурс] Режим доступа <https://datascience.berkeley.edu/open-data-sets/> Дата визита 02.01.2022
53. DeepDrive, UC Berkeley [Электронный ресурс] Режим доступа <https://bdd-data.berkeley.edu/> Дата визита 03.01.2022
54. Machine Learning Datasets and Project Ideas — Work on real-time Data Science Projects, Data Flair [Электронный ресурс] Режим доступа <https://data-flair.training/blogs/machine-learning-datasets/> Дата визита 03.01.2022
55. Google Dataset search engine [Электронный ресурс] Режим доступа <https://datasetsearch.research.google.com/> Дата визита 09.01.2022
56. Russian Language Toxic Comments [Электронный ресурс] Режим доступа <https://www.kaggle.com/blackmoon/russian-language-toxic-comments/version/1> Дата визита 10.01.2022.
57. Обучающая выборка для анализа эмоциональной составляющей текстов [Электронный ресурс] Режим доступа <http://www.e-calc.ru/Obstruct/labeled.csv> Дата визита 10.01.2022
58. Инструмент для анализа обучающей выборки [Электронный ресурс] Режим доступа <http://www.e-calc.ru/Obstruct/toxic-feeling-comments.ipynb> Дата визита 10.01.2022
59. CCN для анализа токсичности текста [Электронный ресурс] Режим доступа <http://www.e-calc.ru/Obstruct/toxic-feeling-CCN.ipynb> Дата визита 10.01.2022
60. Корпус коротких текстов на русском языке. [Электронный ресурс] Режим доступа <https://study.mokoron.com/> Дата визита 12.01.2022.

© Смирнов Сергей Павлович (smirnov.s.p1@edu.mirea.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»

