

## АВТОМАТИЧЕСКАЯ ФИЛЬТРАЦИЯ НЕЦЕЛЕВЫХ ЗАДАЧ В РАБОТЕ АНАЛИТИКА БИЗНЕС-ПРОЦЕССОВ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ С УЧИТЕЛЕМ

Демьянов А.С. (Аналитик данных, Т-Банк)

Научный руководитель – кандидат физико-математических наук, доцент Азимов Р.Ш. (ИТМО)

**Введение.** Для достижения эффективного использования ресурсов в компании необходимо сосредоточить работу аналитиков бизнес-процессов на задачах, которые соответствуют их компетенциям и приносят наибольшую пользу бизнесу. В нашем случае значительная часть заявок, поступающих от коллег, оказывается нецелевой и требует обработки, сопоставимой по времени с целевыми запросами, что затрудняет достижение максимальной производительности. В связи с этим, наше исследование направлено на разработку метода, который использует алгоритмы машинного обучения для распознавания и фильтрации нецелевых запросов на ранних этапах. Это позволит аналитикам сосредоточиться на более критически важных задачах, уменьшить издержки на обработку и повысить эффективность [1].

**Основная часть.** В качестве первого алгоритма был выбран Наивный байесовский классификатор, так как он часто используется для спам-фильтров и тематической классификации, а также прост в реализации и быстр в обучении. Была проведена разметка запросов на целевые и нецелевые, предобработка текста. Для обучения использовалось 17 тысяч запросов, в обучающем наборе класс 0 составил 33%, класс 1 – 66%. Для данной задачи допустимый FPR было решено взять не более 5-10%. Полученные результаты не отвечали изначальным ожиданиям, так как точность попадания была не слишком высокой [2].

Для повышения точности было принято решение использовать CatBoost. На ранее подготовленных данных была обучена модель и опробованы различные параметры для улучшения качества предсказания. Данная модель показала на обучающем наборе ROC AUC 0.73, Ассигасу 0.763. Полученные результаты оказались лучше предыдущего алгоритма, что позволило собрать весь материал исследования и презентовать руководству для дальнейшего проведения теста на ограниченной выборке с целью последующего внедрения в рабочий процесс [3].

**Выводы.** В результате проведенного исследования был разработан эффективный метод, направленный на снижение доли нецелевых заявок, поступающих к аналитикам бизнес-процессов. Данный подход учитывает специфику работы аналитиков, основывается на анализе часто обрабатываемых запросов и позволяет фильтровать нецелевые обращения на ранних этапах. Это способствует оптимизации рабочего процесса и освобождает время для более важных задач. В дальнейшем планируется проведение пилотного тестирования на ограниченной выборке, результаты которого будут использованы для уточнения и доработки метода. При положительных результатах метод будет интегрирован в операционную деятельность, что значительно повысит производительность аналитического подразделения.

### Список использованных источников:

1. Т. Н. Davenport, R. Ronanki. Artificial Intelligence for the Real World - Harvard Business Review, 2018
2. D. Jurafsky, J. H. Martin. Speech and Language Processing – 3<sup>rd</sup> edition, 2020 – 84 p.
3. Dorogush, Ershov, et al. CatBoost: unbiased boosting with categorical features - <https://arxiv.org/abs/1810.11363>, 2018