

Методы повышения качества краудсорсинга

Катунина Евгения Артёмовна

E-mail: evgeniya-katunina@yandex.ru

Университет ИТМО, г. Санкт-Петербург

Научный руководитель – Фильченков Андрей Александрович, кандидат физико-математических наук, доцент

Университет ИТМО, г. Санкт-Петербург

Введение

Краудсорсинг является широко используемым способом решения многих задач сбора и агрегации данных. Примерами таких задач является обработка естественного языка, инженерия знаний, анализ данных, и др. Несмотря на это, до сих пор чрезвычайно актуальна проблема максимально эффективного использования труда заинтересованных участников. Их доступность и свободное время крайне ограничены, поэтому необходимо извлечь максимально возможную пользу от вклада каждого вовлеченного участника.

В рамках данного исследования планируется изучить существующие методы повышения эффективности краудсорсинга и предложить новые методы.

Цель работы

Целью работы является разработка методов повышения качества краудсорсинга. Для достижения указанной цели требуется решить следующие задачи:

1. Исследовать существующие методы повышения качества краудсорсинга;
2. Предложить новые методы повышения качества краудсорсинга.

Обзор существующих методов повышения качества краудсорсинга

Исследования вопросов управления процессом краудсорсинга посвящены построению типовых схем краудсорсинга для решения распространенных практических задач. Ранние работы в этой области относились преимущественно к задачам машинного зрения и разметки изображений при помощи тегов. Одной из первых и наиболее цитируемых работ является статья Л. вон Ана и Л. Дэббис из Университета Карнеги-Меллон (США) об игрофицированном способе разметки изображений путем согласования одновременных ответов игроков. В 2009 году, Дж. Уайтхилл и группа исследователей из Калифорнийского университета в Сан-Диего (США) представили модель GLAD, позволяющую объединить в единую систему такие параметры, как сложность назначенных заданий, квалификацию участников, и получаемые ответы. В 2010 г., М. Бернстейн и группа исследователей из нескольких университетов США представили текстовый процессор Soylent, являющийся тесно интегрированной с платформой Amazon Mechanical Turk надстройкой над редактором Microsoft Word. Soylent позволяет прозрачным образом привлекать участников платформы с целью форматирования, перефразирования и доработки текста. Важным результатом данной работы является подход Find-Fix-Verify, разделяющий задания с открытым вопросом на независимые части. В 2012 г. Г. Демартини, Дж. Дифаллах и Ф. Кудре-Мару из Университета Фрибура (Швейцария) разработали модель ZenCrowd, изначально предназначенную для крупномасштабной разметки ссылок новостей при помощи краудсорсинга на MTurk. Модель предполагает автоматическую предварительную обработку данных и применяет вероятностную сеть для агрегации ответов участников.

Вопрос автоматической оценки производительности участника хорошо представлен в литературе; существуют методы, гарантирующие оптимальный результат в определённых частных случаях. Наиболее популярным на практике подходом является метод Давида-Скина,

предложенный в 1979 г. для задач медицинской диагностики задолго до возникновения краудсорсинга. Данный метод основан на EM-алгоритме и строит матрицы ошибок участников, на основании которой можно вычислить аккуратность участников. Другим популярным методом является Wawa (англ. *worker agreement with aggregate*), определяющий производительность участника как отношение количества ответов этого участника, совпавших с ответами других участников, к количеству всех ответов данного участника.

Алгоритм WorkerRank, созданный М. Дальтаянни, Л. де Альфаро и П. Пападимитроу из Калифорнийского университета в Санта-Крузе и компании Elance-oDesk (США), позволяет одновременно оценить репутацию участников и относительную важность заданий с учётом заранее заявленных компетенций. Представленный алгоритм предназначен для рекомендации исполнителей заданий в области информационных технологий и показывает лучший результат, чем методы коллаборативной фильтрации.

Новый метод повышения качества краудсорсинга

Многие платформы выдают исключительно ответы пользователей. При этом теряется информация о поведении пользователей: кто-то может отказаться от задания, кто-то может выполнять его долго, и т.д. Использование этой информации может повысить качество агрегации ответов при помощи взвешенного голоса большинства.

Увеличить эффективность краудсорсинга можно также добавлением опции, при которой процесс разметки не останавливается до тех пор, пока не получено заданное значение меры согласованности.

Результаты

Были изучены существующие методы повышения качества краудсорсинга, проведён их анализ, были предложены новые методы повышения качества краудсорсинга.