

ТЕХНОЛОГИЧЕСКИЕ ПРИЕМЫ ГЕНЕРИРОВАНИЯ ОБУЧАЮЩИХ ВЫБОРОК ДЛЯ МАШИННОГО ОБУЧЕНИЯ

АНДРЕЕВ А. И.

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Дубовик Д.А. – доктор технических наук, доцент

Целью работы являлось проектирование технологических приемов для автоматизации генерирования обучающих выборок с целью классификации объектов в машинном обучении для улучшения скорости генерирования выборок при сопоставимом качестве. Для достижения поставленной цели были решены следующие задачи:

- изучены основные тенденции генерирования обучающих выборок на основе литературы и исследованы программные средства-аналоги для возможного использования при генерации выборки;
- спроектированы технологические приемы генерирования обучающих выборок для машинного обучения и разработан прототип программного средства на их основе;
- с использованием прототипа улучшена скорость генерирования выборки.

При анализе существующих аналогов и применяемых методов, пользовательских потребностях и особенностях целевой аудитории был сделан следующий вывод: при генерировании обучающих выборок не обойтись без ручного труда (своеобразной биржи), но можно его минимизировать и сделать из работника — подтверждающего решения компьютера. Применяв обучение с подкреплением, машина, с реализованной нейронной сетью, будет предлагать уже распознанные параметры объектов, а работник только проверять соответствуют ли они действительности [1]. Таким образом возможно повысить скорость выполнения работы, тем самым увеличим доход работников за единицу времени и скорость получения заказчиками готового результата.

Итоговые технологические приемы должны поддерживать широкий ряд функций для уменьшения вовлеченности человека в процесс машинного обучения, а также быть доступными для широкой массы рядовых пользователей (работников). В работе предлагается применить следующие технологические приемы для эффективного генерирования обучающих выборок являются:

1 Привлечение более широкой аудитории для генерации объектов и определении их параметров. Данный прием может: ускорить генерацию объектов по сравнению с работой нескольких специально обученных людей, сократить расходы на генерацию выборки, дать доступ к различным объектам разных слоев населения. Для привлечения широкой аудитории необходимо понизить порог входа – разработать платформу или приложение для генерации обучающей выборки. Так же необходимо поощрять работников денежным или иным вознаграждением за выполнения заданий по генерации объектов. При этом может пострадать качество итоговой обучающей выборки.

2 Использование обучения с подкреплением – машина может вычислять предполагаемые параметры объекта, а работнику остается только доопределить их или исправить [1]. Данный прием представляет собой автоматизацию части процесса генерации выборки, он позволит сократить время на генерацию одного объекта, а также контролировать итоговое качество выборки.

3 Упрощение генерации объектов без участия конечного пользователя. Для изображений, например, это автоматическая обрезка заднего фона изображения, это позволит получать изображение только с конечным объектом, что ускорит и упростит обучение конечной нейронной сети – на вход будут подаваться объекты без заднего фона, что можно предусмотреть при обучении и улучшить его качество. Так же обрезка фона необходима при обучении с подкреплением.

Был разработан прототип программного средства, реализующий данные технологические приемы. По результатам испытаний прототипа выявлено, что за счет привлечения широкой аудитории генерирование ускоряется на 77%, а обучение с подкреплением дополнительно ускоряет генерирование выборок на 38%, при этом результаты генерирования дают расхождение около 5% в распознанных параметрах.

Список использованных источников:

1. Саттон Р.С. Обучение с подкреплением / Р.С. Саттон, Э.Г. Барто. – М. : Бином. Лаборатория знаний, 2017. – 400 с.