

Распознавание и перевод языка жестов методами машинного обучения

Тараненко Алексей Дмитриевич

Удмуртский государственный университет, Ижевск

E-mail: ikodane@yandex.ru

По состоянию на 2020 год в России количество глухонемых и слабослышащих людей, которые знают русский жестовый язык, составляет около 240 тысяч человек [2]. Это можно сравнить с населением не очень крупного города, например Сургута или Йошкар-Олы. В обычной жизни люди с нарушениями слуха часто сталкиваются с барьером в общении и коммуникации. Проблемы с адаптацией людей с ограниченным слухом начинаются с самого детства. Как правило, подавляющее большинство таких детей получают образование в специализированных школах, что лишает их возможности получить социально-коммуникативные навыки, научиться взаимодействовать со слышащими людьми. Это приводит к ограничениям в выборе, как образовательных траекторий, так и профессиональных сфер без учета личных предпочтений. Ситуация осложняется еще и тем, что почти никто, кроме самих глухонемых людей, не знает языка жестов. Появление своеобразных переводчиков с языка жестов на голосовой русский язык и наоборот позволило бы решить проблемы, связанные с коммуникацией, благодаря чему можно было сделать доступным образование и другие сферы деятельности для людей с дефектами слуха. В качестве решения проблемы коммуникации и адаптации глухонемых людей в обществе предлагается разработка доступного программного продукта, который мог бы распознавать жестовый язык с помощью web-камеры или других средств регистрации изображения, и автоматически переводить его на естественные языки, прежде всего русский. Для реализации идеи используется модель полносвязной нейронной сети, интерпретацией которой является вычислительный граф, показанный на рисунке 1. В качестве модели нейронной сети предполагается использовать классический метод обратного распространения ошибки, основанный на методе стохастического градиентного спуска, так как он уже хорошо зарекомендовал себя в сфере обработки изображений. Например, с помощью этого метода были реализованы программы для распознавания написанных от руки цифр, букв и знаков, а также детектирования человека или руки на видео или изображении. Кроме метода градиентного градиентного спуска также стоит рассмотреть возможность использования метода случайного леса. Алгоритм применяется для решения задач классификации, регрессии и кластеризации. Данный метод имеет более высокую скорость обработки поданной на вход информации, однако является менее "гибкий", по сравнению с методом классического градиентного спуска.

С алгоритмом оптимизации все обстоит не так однозначно. В случае использования метода градиентного спуска, здесь можно перебрать несколько алгоритмов и посмотреть на скорость и качество выполнения работы. Для начала следует протестировать нейросеть с такими алгоритмами оптимизации, как ADAM, и разделения датасета на мини батчи (от англ. mini batch). Далее, в зависимости от скорости работы модели можно будет попробовать некоторые другие оптимизаторы, например, такие как: метод инерции или ускоренный градиентный спуск (Моменты Нестерова). В качестве функции активации скрытых слоев (то есть первого слоя нейронной сети и всех последующих, не включая выходного) предполагается взять функцию Relu. ReLu не линейна по своей природе, а комбинация ReLu также не линейна. Такая функция является хорошим аппроксиматором [4].

Благодаря такому междисциплинарному подходу, сочетающему теоретические основы лингвистики и практические знания в области информационных технологий, возможны:

- снижение дискриминации по отношению к этой группе населения,
- устранение барьера в общении между глухонемыми и слышащими людьми,
- построение образовательных траекторий (доступ к общеобразовательным школам, вузам и колледжам) и карьеры в профессиональной деятельности для людей с нарушениями слуха.

Источники и литература

- 1) Брюс, П. Практическая статистика для специалистов Data Science: Пер. с англ. / П. Брюс, Э. Брюс, П. Гедек. — 2-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2021. — 352 с.: ил.
- 2) Итоги Всероссийской переписи населения 2020 года // Всероссийское общество глухих: официальный сайт. - 2012. - URL: <https://voginfo.ru/society/2023/02/09/itogi-vserossijskoj-perepisi-naselenija-2020-goda/> (дата обращения: 29.10.2023)
- 3) Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. — СПб.: Питер, 2018. — 480 с.: ил. — (Серия «Библиотека программиста»).
- 4) Милятин И. Функции активации нейросети // Neurohive - Нейронные сети: официальный сайт. URL: <https://neurohive.io/ru/osnovy-data-science/activation-functions/> (дата обращения: 01.11.2023).

Иллюстрации

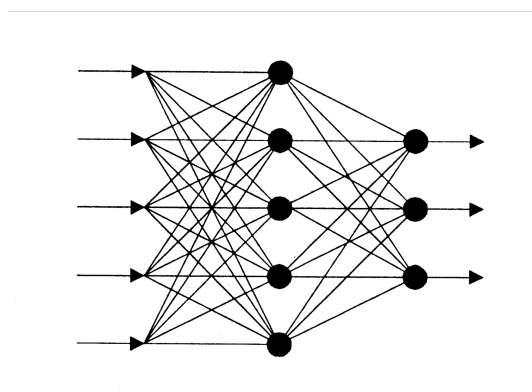


Рис. : 1. Вычислительный граф