

УДК 004.42

РАЗРАБОТКА ИТ-СРЕДСТВ ДЛЯ ЛЮДЕЙ С ОГРАНИЧЕННЫМИ
ПО СОСТОЯНИЮ ЗДОРОВЬЯ ВОЗМОЖНОСТЯМИ

Н. В. ВЫГОВСКАЯ, Е. И. ГАЙЧУКОВ, Н. Ю. ШЕКУНОВ

Белорусско-Российский университет

Могилев, Беларусь

Есть категория людей, которые перенесли тяжелые заболевания и травмы и для которых обычное общение не доступно, которым по некоторым причинам нелегко или невозможно коммуницировать. Для такой категории людей требуется специальное программное обеспечение, которое может помочь им в этом. Синтезаторы речи полезны тем, у кого имеются проблемы с речью, программы с распознаванием образов – для слабовидящих людей.

Развитие машинного обучения позволяет создавать программные средства для распознавания рукописных цифр и текстов, даже небрежно написанных, а затем озвучивать их. Написанные на экране планшета или телефона рукописные цифры могут помочь донести важную информацию. Предлагается программа – скринридер цифр с синтезатором речи.

Первый этап в разработке таких программ – распознавание образа знака, в нашем случае – цифры. Для этого была использована бесплатная библиотека для машинного обучения ML.NET, которая предназначена для языков программирования C# и F#.

ML.NET предоставляет .NET-разработчикам возможности анализа и прогнозирования машинного обучения на основе моделей. Фреймворк построен на .NET Core и .NET Standard, унаследовав возможность кроссплатформенной работы в Linux, Windows и macOS.

ML.NET была разработана как расширяемая платформа, с возможностью использования популярных фреймворков ML (TensorFlow, ANNEX, Infer.NET), а также для еще большего количества сценариев машинного обучения, таких как классификация изображений, обнаружение объектов.

Структура кода приложения и итеративный процесс разработки модели:

- сбор и загрузка обучающих данных в объект IDataView;
- указание конвейера операций для извлечения функций и применение алгоритма машинного обучения;
- обучение модели путем вызова функции Fit() для конвейера;
- оценка модели и итерации для ее улучшения;
- сохранение модели в двоичном формате для использования в приложении;
- загрузка модели обратно в объект Itransformer;

– прогнозирование с помощью функции `CreatePredictionEngine.Predict()`.

Входными данными для нейронной сети являются изображения цифр. Для каждой цифры предусмотрено свыше двух тысяч изображений для обучения нейронной сети.

Для корректного функционирования нейронная сеть должна пройти этап обучения. В нашем случае используется база данных MNIST, которая содержит 60000 изображений для обучения и 10000 изображений для тестирования.

Данные состоят из заранее подготовленных примеров изображений, на основе которых проводятся обучение и тестирование систем.

База данных была создана после переработки оригинального набора чёрно-белых образцов размером 20×20 пикселей NIST. Создатели базы данных NIST, в свою очередь, использовали набор образцов из Бюро переписи населения США, к которому были добавлены ещё тестовые образцы, написанные студентами американских университетов.

Образцы из набора NIST были нормализованы, прошли сглаживание и приведены к серому полутоновому изображению размером 28×28 пикселей.

При разработке программного обеспечения (ПО) была использована платформа WPF, где главным элементом интерфейса приложения является элемент управления InkCanvas. Главное предназначение InkCanvas заключается в обеспечении перьевого ввода. Перо (stylus) – это подобное карандашу устройство ввода, используемое в планшетных ПК. С помощью него пользователь может рисовать линии или выбирать и манипулировать элементами. В нашем случае пользователь пишет цифры, как на бумаге. Разработанное ПО распознает изображение цифры.

Второй этап разработки ПО – синтез речи для распознанных цифр. Речь – это также часть WPF, а синтез речи – это средство, генерирующее аудиосигнал разговора на основе предоставленного текста. С помощью объекта PromptBuilder можно не только воспроизвести (озвучить) цифры, но и изменить интонацию и скорость при их озвучивании, увеличить или уменьшить громкость.

ПО аналогичного назначения было разработано на платформе Everland компанией Билайн для программы «Особый взгляд» благотворительного фонда. Компания Билайн разработала специальные программы, которые используются при разных видах инвалидности, будь то проблемы со слухом, зрением или речью. Эти программы помогают наладить общение, но они требовательны к среде и ресурсам, т. к. им нужен скоростной интернет, тихое место, они могут дорого стоить.

В отличие от названного выше ПО, для нашей программы требования к техническим средствам минимальны. Приложение является десктопным, имеет простой интерфейс и внешние звуки не являются помехой для ее использования.