



УДК 004.42

ПРОЕКТИРОВАНИЕ СИСТЕМЫ ДЛЯ ИСПРАВЛЕНИЯ ДЕФЕКТОВ РЕЧИ У ДЕТЕЙ

Н.Д. Курин, *kurin.kolyan@yandex.ru*

Южно-Российский государственный политехнический университет (НПИ) имени М.И. Платова

Данная работа рассматривает проблему распространенности дефектов речи и описывает принцип и реализацию приложения, разрабатываемого для компании TikTok с целью облегчения работы детских логопедов.

В статье описываются основные проблемы, с которыми пришлось столкнуться при разработке системы, и использованные решения. Некоторые возможности системы реализованы при помощи свободно распространяемой библиотеки для распознавания речи CMUSphinx.

Ключевые слова: распознавание речи, CMUSphinx.

DESIGNING THE SYSTEM FOR CORRECTION OF CHILDREN'S SPEECH DEFECTS

N.D. Kurin, *kurin.kolyan@yandex.ru*

Platov South-Russian State Polytechnic University (NPI), Novocherkassk

This work examines the problem of the prevalence of speech defects and describes the principle and implementation of an application being developed for TikTok company in order to facilitate the work of children's speech therapists.

The article describes the main problems encountered during the development of the system, and the solutions used. Some features of the system are implemented using a freely distributed library for speech recognition CMUSphinx.

Keywords: speech recognition, CMUSphinx.

В современном мире нейронные сети проникают в самые различные сферы жизни человека. Применение методов машинного обучения еще в конце XX века позволяло решать множество задач, а стремительный рост вычислительной мощности компьютеров позволил сделать их использование доступнее и распространеннее, появилось большое количество свободно распространяемых библиотек на их основе. Можно привести множество примеров их использования как для развлечения, так и для решения проблем, так или иначе затрагивающих множество людей по всему земному шару. Одно из перспективных и активно развивающихся направлений использования нейронных сетей – медицина.

В сфере медицины на сегодняшний день существует огромное количество вопросов, для решения нейронные сети уже активно используются [1], но, тем не менее, остаются проблемы, которых подобные инновации еще не успели коснуться. Одним из таких направлений является логопедия. Согласно исследованиям российских логопедов, наибольший процент дефектов речи наблюдается у детей, и количество таких случаев в современном мире стремительно растет. Занимающиеся проблемой эксперты утверждают, что без своевременного вмешательства речевые проблемы могут отразиться как на развитии, так и на социальном самоощущении ребенка, поэтому очень важно вовремя заметить отклонения и принять своевременные меры [2].

Рост количества детей с проблемами произношения приводит к тому, что количество пациентов на одного специалиста (логопеда) становится слишком большим, что неизбежно влечет за собой снижение качества работы с пациентом. Кроме того,



многие семьи, имеющие невысокий доход или живущие в слаборазвитых районах зачастую просто не имеют возможности обратиться к логопеду.

Решение, которое было предложено компанией TikTalk заключается в том, что часть работы с пациентом можно переложить на некоторую специально для этого разработанную систему. Принимая во внимание растущую популярность мобильных приложений, в качестве платформы для разработки системы были избраны мобильные устройства.

Прежде, чем приступить к разбору проблем, возникших при разработке такой системы, и их решений, рассмотрим, как в общем случае происходит работа логопеда с пациентом. Таблица 1 иллюстрирует основные шаги, которые можно выделить в работе логопеда с пациентом.

Таблица 1

Анализ шагов по исправлению дефекта речи

Название	Краткое описание	Назначение	Роль врача	Повторяемость
Первичная консультация	Пациент обращается к врачу для определения наличия дефектов речи и получения методики лечения	Получение рекомендаций по лечению (набор упражнений)	Диагностирование дефектов, формирование рекомендаций	Нет
Выполнение упражнений	Пациент выполняет назначенные ему упражнения	Исправление дефекта	-	Да
Повторная консультация	Пациент обращается к врачу для проверки динамики лечения	Контроль состояния исправления дефекта	Диагностирование состояния исправления дефекта	Да

При первом обращении человека, имеющего дефект речи, к логопеду, первым делом врач стремится понять, с каким дефектом приходится иметь дело и его степень, а затем назначается ряд упражнений для его устранения, с периодической проверкой результатов. Если упростить описанные выше шаги, получается, что специалист сначала сравнивает речь пациента с речью среднестатистического здорового человека, после чего посредством занятий приближает дефектное произношение к считающемуся нормальным уровнем, делая промежуточные проверки, чтобы убедиться в эффективности терапии.

В описанной методике можно четко выделить повторяющиеся действия, такие как выполнение упражнений и проверка произношения. Если рассмотреть роль логопеда в каждом из этих процессов, можно сделать вывод, что при выполнении упражнений она полностью отсутствует, а при проверке, в общем случае, сводится к сверке речи пациента с нормальным, или эталонным произношением.

Исходя из рассмотренного выше, можно сделать вывод, что наибольшую значимость играет первичное общение врача с пациентом, в ходе которого будут назначены требуемые упражнения, тогда как остальные шаги многократно повторяются и могут быть заменены разрабатываемой для этого системой.



Таким образом, первичный осмотр по-прежнему остается обязательным, однако же по его итогу пациенту вместо самостоятельных занятий с последующим повторным визитом назначается подбираемый индивидуально лечебный план в разрабатываемой системе. План составляется таким образом, что в ходе работы по нему пациент должен будет произносить определенные слова или фразы, которые были подобраны логопедом. В некоторых случаях, например, при наличии нескольких дефектов речи или особо сложных случаев, план может быть разбит на фазы, каждая из которых будет направлена на исправление одного из дефектов, или постепенное ослабление одного.

Так как целевую аудиторию представляют дети, процесс выполнения упражнений было решено выполнить в игровой форме. Идея состоит в том, что ребенку в ходе игры, например, классического раннера (от англ. runner – бегун, игра, в которой игрок, управляя бегущей фигуркой, собирает предметы и преодолевает препятствия), будут встречаться преграды, представляющие собой слова или фразы из лечебного плана, которые необходимо будет произносить, чтобы продвинуться дальше в игровом процессе. Чем лучше произношение – тем больше будет начислено игровых очков, и тем выше будет общий счет по окончании игры. Такой подход имеет сразу несколько преимуществ, которые должны повысить эффективность терапии:

- Повышенный интерес – в первую очередь, приложение воспринимается ребенком именно как игра
- Соревновательный подход – каждый из детей будет стремиться набрать больше очков, чем остальные

Разобравшись с планированием занятий и их представлением в системе, встает главный вопрос: каким образом система может контролировать правильность произношения? Обратившись повторно к описанию того, как проходит первый прием у логопеда, получаем, что нужно сравнивать произносимые ребенком слова с некоторым «нормальным» произношением. Таким образом, для реализации контроля правильности произношения в системе нужно реализовать распознавание речи пациента и его сравнение с неким эталоном.

Решение для задачи по распознаванию речи было найдено в свободно распространяемой библиотеке для распознавания речи CMUSphinx [3]. Данная библиотека свободно распространяется по BSD лицензии и представляет большие возможности по работе с распознаванием голоса. Для распознавания речи используются акустические модели, которые размещены в свободном доступе для некоторых популярных языков. Библиотека также имеет несколько дополнительных пакетов, расширяющих ее функционал. Основной возможностью, нашедшей применение в системе, является возможность адаптирования уже существующей модели под особенности произношения человека, для распознавания речи которого она будет использоваться. Также важной особенностью является то, что, помимо простого распознавания речи, библиотека может возвращать некоторое целочисленное значение, которое содержит информацию о точности результата распознавания.

Так как система ориентирована на работу с детьми, простое использование готовой акустической модели было бы не совсем корректно из-за некоторых



особенностей детского произношения. Для улучшения работы уже существующей модели с детской речью, была проведена ее адаптация. Вся суть процесса адаптации сводится к тому, что для уже существующей модели создается дополнительная матрица трансформаций. Для проведения адаптации необходимы: начальная модель, достаточное количество аудиозаписей и файл со списком соответствий имени аудиозаписи с текстом, в ней произнесенным. Помимо этого, также используются некоторые другие файлы, но их роль в данном процессе значительно меньше. Когда все необходимые файлы подготовлены – достаточно запустить необходимую команду и CMUSphinx на выходе выдаст обновленную акустическую модель. Для проведения адаптации было собрано достаточно много записей произношения многими детьми различных слов, общая продолжительность которых составила более 10 часов, после чего был запущен сам процесс адаптации. Путем подбора параметров конфигурации CMUSphinx [4], была подобрана такая их комбинация, что получаемая модель выдавала необходимую для работы системы точность распознавания. Таким образом, проблема с отсутствием эталонной модели была решена.

Так как речь пациентов может содержать несколько дефектов, каждый из которых, в свою очередь, может иметь различную степень выраженности, для того, чтобы система смогла корректно распознавать речь, также должна создаваться адаптированная модель. Каждый план направлен на исправление одного дефекта, поэтому такая модель должна перезаписываться в начале каждого нового плана. Такая модель создается по той же схеме, что и эталонная, с той лишь разницей, что слова произносятся непосредственно пациентом по некоторому списку, определенному для конкретного плана. После того, как записи были получены, проходит процесс, аналогичный описанному ранее для эталонной модели.

Для процесса распознавания речи во время занятий (игры) было принято решение использовать обе описанные модели. Первая – эталонная модель, пытается распознать речь ребенка относительно произношения, принятого за правильное. Вторая – модель, создаваемая для ребенка в начале плана и адаптируемая к его индивидуальным особенностям произношения, используется для подсчета очков и в некоторых случаях для помощи в определении правильности произношения. В случаях, когда произношение очень далеко от правильного, эталонная модель не всегда способна понять речь ребенка, тогда сравнение результатов точности распознавания позволяет определить наличие или отсутствие прогресса при продвижении по плану.

Для облегчения задачи отслеживания прогресса была введена формула, учитывающая результаты точности распознавания обеих моделей и вводящая некоторые поправки, получаемое в итоге значение стало возможным использовать не только для проверки результатов в ходе выполнения лечебного плана, но и как количество очков, набранное пациентом в ходе занятий. Благодаря такому подходу, одно и то же значение становится значимым как для пациента (в роли игрока), так и для врача.

Таким образом, в ходе работы были выделены основные задачи для создания системы для исправления дефектов речи у детей. Каждая задача была тщательно проанализирована, и, в конечном итоге, решена. Результатом работы



стала система, позволяющая врачам-логопедам в игровой форме проводить работу с пациентами, даже в удаленной форме.

Использование CMUSphinx и сотрудничество с реальными специалистами в сфере логопедии открывает несколько возможностей для дальнейшего развития системы. Например, для реализации мультязычности системы был реализован алгоритм, позволяющий собирать акустические модели для языка, который их не имеет, а внедрение машинного обучения для анализа прогресса планов позволило бы предсказывать сроки исправления того или иного дефекта.

Несмотря на то, что момент написания статьи система все еще улучшается, она уже запущена с поддержкой двух различных языков и активно используется.

Список цитируемой литературы

1. Нейронные сети в медицине [электронный ресурс] – URL: <https://www.osp.ru/os/1997/04/179201/> (дата последнего обращения – 09.01.2019).
2. Докшина А.А., Мехрякова Т.И. – Актуализация вопроса этиологии речевых нарушений у детей – 2013 – с.99-101.
3. Официальная документация CMUSphinx [электронный ресурс] – URL: <https://cmusphinx.github.io/> (дата последнего обращения – 10.01.2019)
4. Шумопонижение в CMUSphinx [электронный ресурс] – URL: <https://habr.com/ru/post/227099/> (дата последнего обращения – 11.01.2019)

© Н.Д. Курин, 2019