

ИНТЕГРАЦИЯ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ В СОВРЕМЕННОЕ ОБРАЗОВАНИЕ ДЛЯ РАЗВИТИЯ «ГИБКИХ» НАВЫКОВ

Березина А. О.¹, студентка, ✉ bereanas@mail.ru

Медведев А. В.¹, студент, mdl-studio@yandex.ru

Шнайдер К. С.¹, студент, knhn0@yandex.ru

Митрохин М. А.¹, доктор техн. наук, доцент, mmax83@mail.ru

¹ Пензенский государственный университет, ул. Красная, д. 40, 440026, г. Пенза, Россия

Аннотация

В статье рассматривается подход к интеграции нейросетевых технологий для устранения дисбаланса между развитием жестких (hard skills) и гибких (soft skills) навыков в процессе образования. Современный рынок труда требует от специалистов высоких коммуникативных, креативных и логических навыков. Однако традиционные образовательные программы недостаточно акцентируют внимание на soft skills. Авторы статьи, опираясь на исследования РБК, Google и РСМ, подчеркивают важность гибких навыков для успешной карьеры и социальной адаптации, отмечая, что многие испытывают трудности именно с коммуникативным навыком как основополагающим soft skills.

Для решения этой проблемы предлагается инновационный подход с использованием нейросетевых технологий для диагностики и развития коммуникативных навыков. Авторы анализируют существующие модели глубокого обучения и предлагают свою структуру сверточно-рекуррентной нейронной сети (CRNN) для диагностики речевых недостатков в русском языке. Разработанная модель оценивает дефекты произношения и предоставляет персонализированные обучающие материалы.

Авторами предлагается интерактивная образовательная платформа, реализующая созданную модель в рамках технологии обучения hard skills совместно с программами развития soft skills. Нейросетевые алгоритмы платформы оптимизируют учебный процесс, адаптируя его к индивидуальным особенностям обучающегося, и могут использоваться как самостоятельно, так и в дополнение к занятиям с репетитором.

Ключевые слова: гибкие навыки, техника речи, нейросетевые технологии, дикционные недостатки, платформа.

Цитирование: Березина А. О., Медведев А. В., Шнайдер К. С., Митрохин М. А. Интеграция нейросетевых технологий в современное образование для развития «гибких» навыков // Компьютерные инструменты в образовании. 2025. № 1. С. 48–60. doi:10.32603/2071-2340-2025-1-48-60

1. ВВЕДЕНИЕ

Активная современная молодежь планирует свою карьеру еще в школе и вузе, создавая комплексный план, учитывающий цели, методы, риски и адаптацию. Однако традиционное образование фокусируется на жестких навыках (hard skills), оставляя без должного внимания мягкие навыки (soft skills) — коммуникацию, критическое мышление, креативность и др. Это создает дисбаланс, так как современный рынок труда требует универсальных специалистов, способных адаптироваться к изменениям.

Существует два противоположных взгляда на роль soft skills. Первая точка зрения, которую отстаивает, в частности, Сет Годин, утверждает, что soft skills — это фундаментальные, «настоящие» навыки (real skills), определяющие эффективность работы и взаимодействие в коллективе [1]. Подтверждением служит исследование начала XX века, показавшее, что 85 % успеха в карьере зависит от гибких навыков [2]. Противоположная точка зрения, высказанная К. Нордстремом и Й. Риддерстрале, заключается в том, что для работодателей первостепенное значение имеет профессиональная компетентность кандидата, а soft skills играют лишь дополнительную роль [3].

Несмотря на это разногласие, современное образование должно интегрировать развитие soft skills в учебный процесс, используя для этого инновационные методы, в том числе нейросетевые технологии для диагностики и развития индивидуальных качеств. Такой подход позволит повысить качество обучения.

2. ИННОВАЦИОННЫЙ ПОДХОД К РАЗВИТИЮ Soft Skills НА ОСНОВЕ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ

Аналитики из РСМ (Российского Союза Молодёжи) выделили 14 ключевых soft skills, поставив на первое место коммуникативные навыки. Исследования показали, что недостаточный уровень развития коммуникации является серьезным препятствием как для построения успешной карьеры, так и для социальной адаптации личности. Подтверждение этому можно найти и в результатах зарубежных исследований, проведенных компанией Google в 2017 году. Анализируя наиболее продуктивные рабочие команды, Google установила, что ключевым фактором успеха является наличие у сотрудников хорошо развитых soft skills, а в первую очередь — высокий уровень коммуникативных навыков [4]. Однако, несмотря на важность коммуникации, многие люди испытывают значительные трудности.

Перспективным направлением представляется использование нейросетевых технологий для обучения технике речи в качестве основополагающих soft skills. Данную технологию можно использовать для дополнения образовательного процесса школьников и студентов, основанного на классическом обучении, реализованную в качестве информационной системы по обучению детей и взрослых, которая сможет оценивать текущие навыки пользователей, рекомендовать материалы по развитию навыков, оценивать прогресс и выстраивать образовательную программу. Алгоритм работы образовательной платформы показан на рисунке 1.

Инновационная платформа дополнительного обучения сочетает освоение общеобразовательных предметов с развитием гибких навыков (логика, коммуникация, управление проектами, творческое мышление). Обучение гибким навыкам строится на индивидуальных параметрах: нейросети подбирают контент, отслеживают прогресс и выстраивают долгосрочную программу. Это оптимизирует время, заменяя репетитора (рис. 1, верхняя ветвь алгоритма) либо дополняя его работу (рис. 1, нижняя

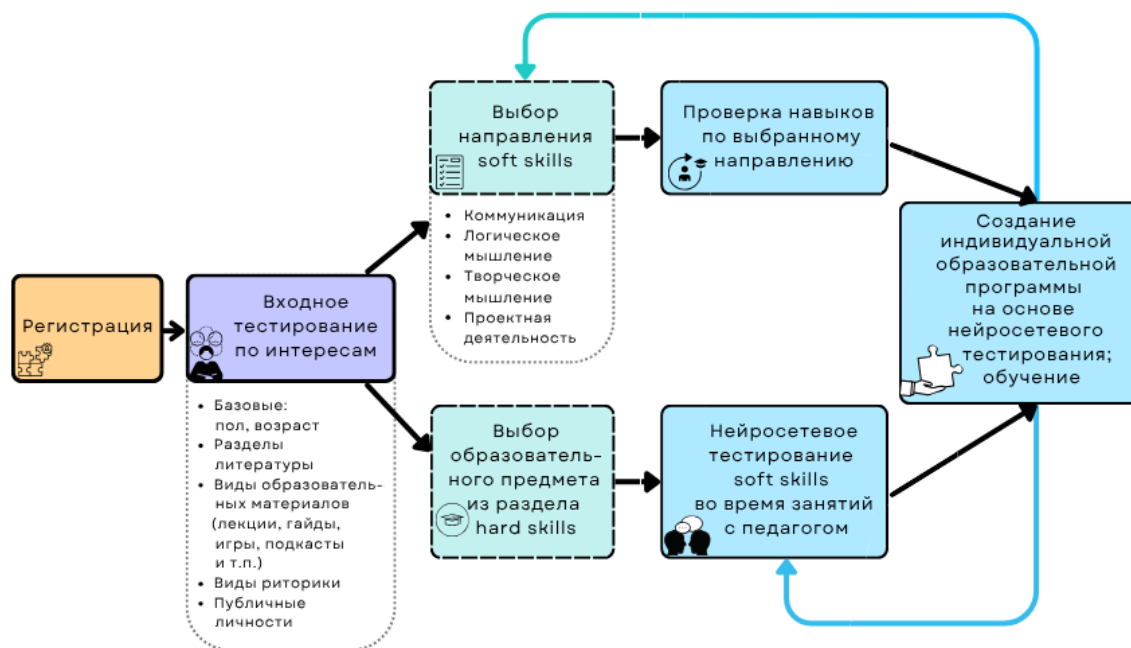


Рис. 1. Алгоритм работы образовательной платформы

ветвь алгоритма). В перспективе возможно расширение платформы за счёт новых нейросетевых модулей, обеспечивающих индивидуальную подготовку по лидерским, командным и другим социально-психологическим навыкам, повышая вовлечённость и эффективность обучения благодаря интеграции hard и soft skills.

Учитывая, что коммуникация является основополагающим навыком soft skills, авторы статьи начинают с изучения и создания необходимой технологии именно по развитию данного навыка. Согласно данным ИТМО [5], около 95 % населения планеты имеют проблемы с техникой речи, связанные с построением речи и дикционными недостатками. Отсюда возникают страхи к публичным выступлениям и сложности освоения навыка коммуникации. Традиционные методы обучения в данном случае оказываются малоэффективными, ограниченными формальными правилами и недостатком индивидуального подхода.

3. СУЩЕСТВУЮЩИЕ МОДЕЛИ

Нейросетевые технологии позволяют получать персонализированные обучающие программы, адаптированные к индивидуальным особенностям речи каждого ученика [6].

В качестве параметров для индивидуальной диагностики речи на первом этапе авторы статьи решили рассмотреть дикционные недостатки пользователей и разработать соответствующую нейронную сеть. Для начала важно выбрать модель и разработать структура под данную задачу. Для этого рассмотрим наиболее успешные модели глубокого обучения по анализу и корректировке речи на сегодняшний день.

Рибас и другие [7] предложили перспективный подход к использованию самонаблюдаемых представлений (SS) для классификации дефектов речи. SS-представления сначала обучаются на размеченных записях речи, а затем используются в качестве входных данных для модели классификации, которая обучается на размеченных речевых данных.

Для задачи классификации авторы применили архитектуру трансформер, которая основана на автоматическом выделении признаков в процессе обучения на размеченных данных.

В исследовании [8] авторы рассмотрели целый ряд методов извлечения признаков, включая MFCC, CQCC, I-векторы, признаки просодии, артикуляционные признаки и меры фонирования для представления различных аспектов речи и ее нарушений, в частности дизартрии. Авторами показано, что модели глубокого обучения, включая глубокие нейронные сети (DNN), сверточные (CNN), нейронные сети с долгой краткосрочной памятью (LSTM) и управляемые рекуррентные блоки (GRU) способны эффективно обрабатывать описанные признаки для решения задач определения дефектов речи.

Венугопалан и др. [9] использовали большой набор данных, состоящий из 756 147 устных высказываний 677 дикторов, каждый из которых имел различные нарушения речи. Они классифицировали речь на типичную (здоровую) и нетипичную (нарушенную). Они разработали модели классификации на основе различных архитектур глубокого обучения. Авторы использовали уникальный подход к обработке входных данных. В этой модели вместо пакетов фильтров используется метод, известный как LEAF. LEAF динамически настраивает фильтры, процессы пулинга и нормализации, основанные на специфических характеристиках входных данных. Эти низкоуровневые представления данных затем обрабатываются двумерной CNN, которая работает в двух измерениях — временном и частотном. Автор утверждает, что эта адаптивная система может эффективно обрабатывать широкий спектр паттернов и характеристик данных, что делает ее мощным классификатором.

Существуют мультимодальные модели, например, GPT-4 и Claude 3.5, демонстрирующие высокую эффективность в различных областях обработки естественного языка. Однако их применение для анализа дикционных недостатков речи ограничено отсутствием специализированного обучения и высокой стоимостью создания необходимых обучающих данных, включающих тонкие нюансы произношения.

Наиболее близкую по структуре разработку, предлагаемую авторами статьи, описала команда разработчиков «NovatorSpace», которые специализируются на разработках в сфере логопедии, связанные с диагностикой и помощью в коррекции речевых недостатков. Однако сравнить архитектуру нейронной сети не предоставляется возможным, так как компания не предоставила данные в свободных источниках. Недостатками технологии «NovatorSpace» оказалось отсутствие индивидуальной программы и трудности в определении нарушенного звука при произнесении его верно в одной из позиций, а в другой — искаженно [10].

Методика, представленная авторами статьи, позволяет оценить звук в разных позициях говорящего благодаря обучению нейронной сети на фразах, насыщенных проблемным звуком в различных постановках. При этом обучение происходило на звукозаписях людей в возрасте от 12 до 50 лет, что позволило нейронной сети диагностировать звук как у детей, так и взрослых, что не подразумевалось у «NovatorSpace» ввиду отсутствия такой необходимости.

Еще одной современной конкурирующей моделью выступает «Алгоритм постановки произношения на основе информационной теории восприятия речи и сверточных нейронных сетей» [11]. Представлен итеративный метод обучения распознаванию речи, основанный на сверточных нейронных сетях. На первом этапе формируется база эталонных слов, включающая только те, для которых вероятность, выдаваемая CNN, максимальна. Далее CNN дообучается на основе отобранных эталонов. Качество произношения

оценивается посредством распознавания вновь произнесенных слов дообученной CNN. Экспериментальная проверка на десяти английских словах-командах (“down”, “go”, “left”, “no”, “off”, “on”, “right”, “stop”, “up”, “yes”) показала самую большую точность распознавания — 72 % для модели Conv. В эксперименте участвовали также модели: low_latency_conv, low_latency_conv (использует нейронную сеть cnn-one-fstride4), low_latency_svdf (использует топологию rank-constrained, сжатие нейронных сетей), tiny_conv (состоит из одно-сверточного нейронного слоя).

Таким образом, DNN, в частности RNN и CNN, широко применяются в автоматическом распознавании речи (APP). RNNs отличаются циклической структурой связей между нейронами, что позволяет им «запоминать» предыдущую информацию и учитывать контекст. Однако высокая вычислительная сложность RNNs, требующая мощных вычислительных ресурсов, ограничивает их применение. RNNs, в отличие от многослойных перцептронов, обладают циклическими связями, позволяющими учитывать временную динамику речи [12].

CNNs, в свою очередь, демонстрируют более высокую скорость обработки и точность распознавания, превосходя при этом RNNs в плане вычислительной эффективности [13]. CNN анализирует речевой сигнал по частям, обнаруживая характерные особенности на разных уровнях детализации, что позволяет более точно распознавать речь.

Несмотря на преимущества CNN, эффективность всех типов DNN существенно зависит от качества и репрезентативности обучающих данных. Недостаточная репрезентативность приводит к снижению точности распознавания нестандартной речи (с шумами, дефектами произношения). Так как анализ дефектного произношения речи — ключевая задача, авторы статьи определенным образом подходят к сбору и разметки обучающих данных и выбирают архитектуру, способную минимизировать недостатки.

4. РАЗРАБОТКА МОДЕЛИ

Авторы статьи предлагают использовать ряд готовых технологий в создании новой модели по диагностированию речевого дефекта. Под речевым дефектом или речевым отклонением понимается неправильная постановка конкретного согласного звука в русскоговорящей речи — дикционные недостатки. После диагностирования пользователь получит образовательный материал на постановку конкретного согласного звука, что позволит работать над речью в любое удобное время через функциональные устройства: мобильный телефон, планшет, ноутбук, компьютер. Индивидуальные образовательные материалы по технике речи, основанные на нейросетевой проверке речевых дефектов, позволят обучаться основному soft skills — коммуникации через образовательную платформу.

Входными данными для математической модели являются звуковые сигналы с обозначениями типа дефекта речи, установленного логопедом. Для описания дефектов в речи была предложена следующая модель. Речевые нарушения представлены в виде набора:

$$M = \{A, B, C, F\},$$

где M — множество дефектов устной речи; A — дикционное нарушение звуков [с], [с'] и [з]; B — дикционное нарушение звуков [р] и [р']; C — дикционное нарушение звуков [ш] и [ж]; F — дикционное нарушение звуков [л] и [л']. Каждое из обозначенных множеств представляет собой целевое нарушение, диагностируемое сетью по аудиозаписи речи тестируемого человека.

4.1. Моделирование нейронной сети для распознавания дефектов речи

Исходя из особенностей задачи и анализа литературы, в качестве нейронной сети выбрана сверточно-рекуррентная нейронная сеть (Convolutional Recurrent Neural Networks, CRNN). Она обладает рядом преимуществ для решения задачи распознавания речи [14]. Сверточные слои функционируют как экстракторы признаков, стремясь изучить отличительные особенности посредством последовательных сверток и нелинейных преобразований, применяемых к частотно-временному представлению на входе сети [15].

В рамках выбранной архитектуры разработана структура сверточно-рекуррентной нейронной сети, которая используется для обучения модели для решения задачи определения типа дефектов речи из набора M . Нулевой слой нейронной сети — это слой с входными данными (слой “Ввод”) показан на рисунке 2. Предложенная структура CRNN состоит из трех сверточных блоков, за которыми следуют два рекуррентных и два линейных слоя. Первый слой сверточной нейронной сети — “Свертка № 1” (2D CNN). Его основными входными параметрами являются количество фильтров, размер из окна свертки и массива со значением размера входных данных. Значение для количества фильтров было установлено равным 128, размер окна свертки — 3 на 3, максимальное объединение — 1×5 .

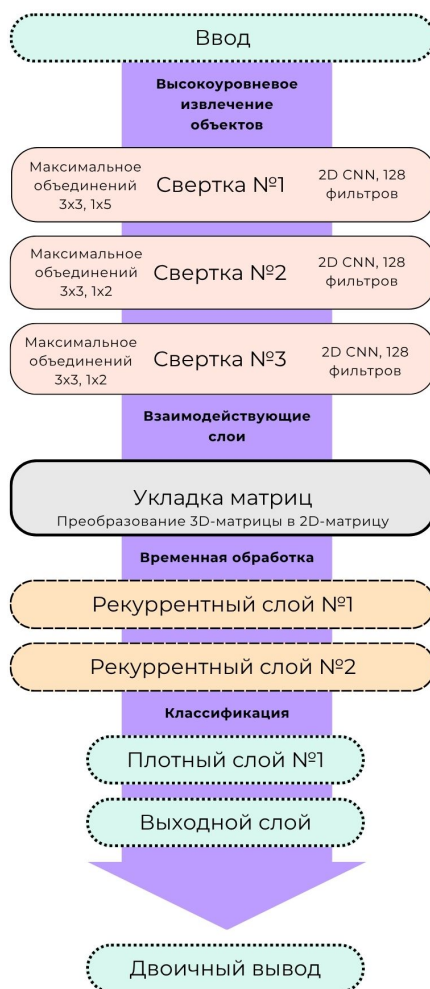


Рис. 2. Алгоритм выявления речевых дефектов

На вход сети поступает частотно-временное представление данных. Для уменьшения размерности данных используются объединение и субдискретизация. Поскольку обнаружение звуковых событий требует оценки временного положения, необходимо сохранить временную ось, и поэтому операция объединения выполняется только по частотной оси. Повторяющиеся блоки свертки (Свертка № 1, Свертка № 2, Свертка № 3) включают операции свертки, нелинейной активации и объединения. Они позволяют выполнять последовательные преобразования входных данных, соответствующие сверточным блокам. Второй слой предназначен для уменьшения размерности данных и получения важной информации. Максимальное значение второго и третьего слоев — 1×2 . Третий слой является активным. Функция активации ReLU была выбрана в качестве основной. Выходные данные последнего сверточного блока складываются по оси частот, образуя двумерное представление объекта, которое затем передается в качестве входных данных для первого рекуррентного слоя.

4.2. Инженерия признаков и обучение модели

4.2.1. Стратегия выбора признаков

Процесс выявления речевых дефектов включал в себя несколько этапов. Для разработки функционирующей модели были рассмотрены все признаки речевого нарушения. Предварительно была осуществлена подготовка конкретных предложений и фраз из сборника Л. Д. Алферовой [16], содержащие повторение проблемного звука в разных частях слов. Данный подход позволил исключить трудности в определении нарушенного звука при произнесении его верно в одной из позиций, а в другой искаженно. Согласно исследованиям Л. В. Лопатиной [17], чаще всего встречаются нарушения в группах: свистящих — 22 %, шипящих — 24 %, [р] — 26 %, [л] — 10 %; реже встречаются нарушения заднеязычных звуков, а также нарушения смягчения и озвончения — от 1 % до 4,5 %. Искажения таких звуков, как [м], [н], [п], [т], и гласных звуков наблюдаются редко, из-за чего на этапе апробации ими можно было пренебречь. Данный подход позволил изначально сегментировать речевой дефект, чтобы увеличить точность обучения конкретного звука. Платформа, предложенная авторами статьи, рассчитана как на детей, так и на подростков, поэтому важно привести результаты исследования Л. В. Лопатиной не только взрослого человека, но и детей: нарушение свистящих звуков — у 95,7 %, шипящих — у 82,4 %; звуков [р], [р'] — у 50 %, [л], [л'] — у 46,7 % [17]. Таким образом, авторами статьи были приняты к рассмотрению дефектное произношение (не соответствующее фонетическим нормам русского языка) звуков: [р-р'] — ротацизм; [л-л'] — ламбдацизм; [с-с'-з], [ш-ж] — сигматизм.

4.2.2. Сбор и анализ данных для обучения нейронной сети

Обучающая выборка получена с помощью ручного сбора 4380 тысяч секунд аудиозаписей, отобранных по виду дефектного произношения звуков русскоговорящих людей. Эти аудиоматериалы обрабатывались и разбивались на фразы и слова. Разбиение речевого аудиофайла производилось с помощью инструмента для маркировки данных Label.Studio [<https://labelstud.io/>] и библиотек Python. Затем звуковые сигналы преобразовывались в спектрограмму, которая позволила проводить дальнейший анализ. Звуковая спектрограмма используется для фонетической идентификации произносимых слов и для их анализа.

На рисунке 3 представлена модель предварительной обработки входного сигнала.

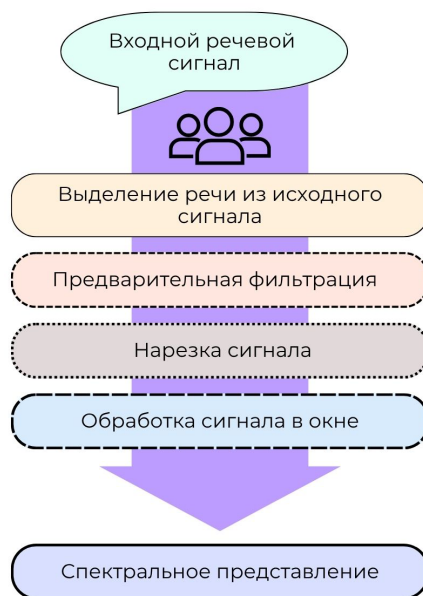


Рис. 3. Схема предварительной обработки речевого сигнала

Диаграмма прохождения сигнала аудиозаписи с речевым дефектом для детектирования речевого дефекта (звуки [с] и [з]) через нейронную сеть представлена на рисунке 4. Вход — спектрограмма сигнала, выход — разметка ошибок по времени аудиозаписи.

4.2.3. Обучение и тестирование нейронной сети

После обучения сверточно-рекуррентной нейронной сети она была протестирована на данных, полученных с дефектами речи и без них.

Для обучения использовался алгоритм прямого распространения с предварительным разделением на сверточные слои для выделения признаков, функция ошибки — бинарная кроссэнтропия. Обучение проводилось в течение 30–50 эпох.

5. РЕЗУЛЬТАТЫ

Средняя точность оценки звукопроизношения нейросетью составила 81 %, что является успешным результатом для апробации. Качественными показателями стали: удобство использования и быстрое получение результата с пояснительной запиской (удобство и понятность результатов) о проблематике звукопроизношения конкретного пользователя. Удобство использования проверялось по реакциям людей после тестирования данного функционала платформы.

Сравнительный анализ точности и времени распознавания для акустических моделей (в том числе отдельных преимуществ) наиболее близких к разработке авторов статьи, показан в таблице 1.

Оба конкурирующих метода со своими моделями ограничены анализом изолированных слов и не учитывают контекст речи или появление новых слов с аналогичным звучанием. Предложенный авторами статьи алгоритм, в отличие от разработок «NovatorSpace»

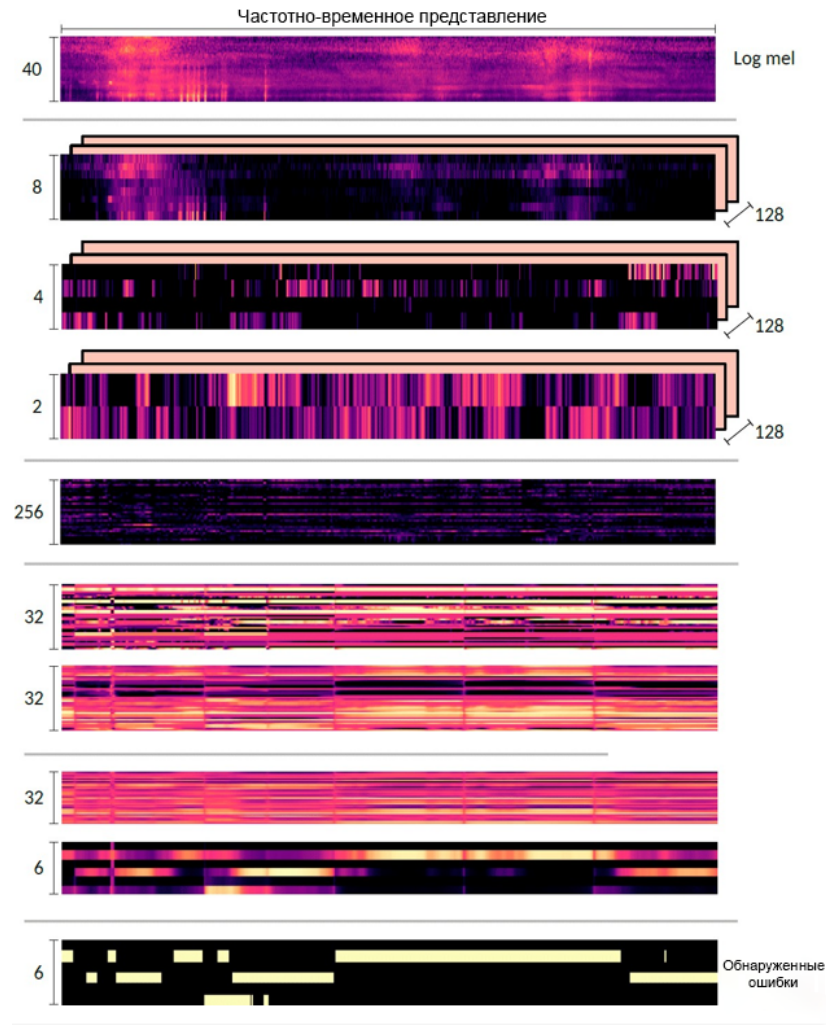


Рис. 4. Диаграмма прохождения сигнала аудиозаписи через нейронную сеть

Таблица 1. Сравнительный анализ эффективности различных акустических моделей

Модель	Целевая аудитория продукта	Точность распознавания, % (без дообучения)	Время распознавания, мс	Возможность анализировать звук в контексте предложений в различных позициях	Язык
Conv	дети / взрослые	72	7,5	нет	англ.
low_latency_conv	дети / взрослые	27	7	нет	англ.
low_latency_svd	дети / взрослые	46	11	нет	англ.
tiny_conv	дети / взрослые	20	5	нет	англ.
NovatorSpace	только дети	—	—	нет	русский
Предлагаемая модель	дети / взрослые	81	5	да	русский

и модели, основанной на информационной теории восприятия речи, демонстрирует более высокую адаптивность. Ключевое преимущество заключается в использовании фразового анализа, а не анализа отдельных звуков. Оба конкурирующих метода, как было отмечено, сталкиваются с трудностями при идентификации искаженного звука, если он произносится в различных позициях. Это обусловлено анализом изолированных фонем, что не учитывает контекста речи. В естественной речи звуки влияют друг на друга, изменяя свою артикуляцию и акустические характеристики. Разработка авторов статьи учитывает эти взаимодействия. Тем самым повышается точность распознавания даже при незначительных искажениях звука в определенном фонетическом окружении.

Разработка авторов статьи демонстрирует более перспективный подход к диагностике и мониторингу речевых нарушений по сравнению с известными аналогами. Использование фразового анализа, обучение на данных большого объема и разнообразия, а также широкий возрастной диапазон обеспечивают более высокую точность, адаптивность и универсальность алгоритма. Дальнейшие исследования могут быть направлены на расширение базы данных, включение новых типов речевых нарушений.

6. ЗАКЛЮЧЕНИЕ

Таким образом, полученные результаты показали, что диагностика с применением нейросетей эффективна и открывает новые перспективы в «Образовании 4.0», представляя собой инновационный подход. Предложенный подход к обучению технике речи с использованием нейросетей — это дополнение к классическому образованию, которое помогает не только улучшить технику речи, но и развивает уверенность в себе, способность к публичным выступлениям — важные компоненты успешной самореализации. Интеграция нейросетевых технологий в образовательный процесс позволит подготовить новое поколение специалистов, обладающих как *hard skills*, так и развитыми *soft skills*, необходимыми для успеха в динамично меняющемся мире. Дополнительным преимуществом для пользователя будет инновационность диагностики, поскольку современные технологии вызывают интерес и желание попробовать новые разработки, особенно учитывая спрос общества на нейросети.

Список литературы

1. Godin S. Let's stop calling it soft skills. Talent Management // Medium. 2017. [Электронный ресурс]. URL: <https://itsyourturnblog.com/lets-stop-calling-them-soft-skills-9cc27ec09ecb> (дата обращения: 03.03.2025).
2. Mann C. R. A study of engineering education, Joint committee on engineering education of the national engineering societies. New York: New York city, 1918.
3. Нордстрем К., Риддерстрале Й. Бизнес в стиле фанк. Капитал пляшет под дудку таланта. М.: Манн, Иванов и Фербер, 2013. 288 с.: ил.
4. Гибкие навыки (soft skills) — компетенции будущего: метод. пособие / Дальневост. гос. науч. б-ка, отд. науч.-исслед. и науч.-метод. работы; сост. М. П. Лопушенко; редкол.: Т. Ю. Якуба и др. Хабаровск: ДВГНБ, 2023. 36 с.
5. Никулина А. 7 советов: как побороть страх публичных выступлений // Новости Университета ИТМО. 2020 [Электронный ресурс]. URL: <https://news.itmo.ru/ru/news/9396/> (дата обращения: 03.03.2025).
6. Alam M., Samad M., Vidyaratne L., Glandon A., Iftekharuddin K. Survey on Deep Neural Networks in Speech and Vision Systems. Neurocomputing, 2020. 417. P. 302–321. [Электронный ресурс]. URL: <https://doi.org/10.1016/j.neucom.2020.07.053> (дата обращения: 10.03.2025).

7. Ribas D., Pastor M. A., Miguel A., Martinez D., Ortega A., Lleida E. Automatic voice disorder detection using self-supervised representations // IEEE Access. 2023. Vol. 11. P. 14915–14927.
8. Joshy A. A., Rajan R. Automated dysarthria severity classification: A study on acoustic features and deep learning techniques, IEEE Trans. Neural Syst. Rehabil. Eng. 2022. Vol. 30. P. 1147–1157.
9. Venugopalan S. et al. Speech intelligibility classifiers from 550 disordered speech samples // arXiv [eess.AS], 2023.
10. Хоменко А. А., Зинченко И. В., Брызгалова Ю. В. Диагностика речевых нарушений у дошкольников с помощью искусственного интеллекта // Педагогическая перспектива. 2024. № 1(13). С. 58–65. [Электронный ресурс]. URL: [https://doi.org/10.55523/27822559_2024_1\(13\)_58](https://doi.org/10.55523/27822559_2024_1(13)_58) (дата обращения: 5.04.2025).
11. Савченко Л. В. Система постановки произношения на основе сверточных нейронных сетей и информационной теории восприятия речи = Computer-assisted language based on convolutional neural networks and information theory of speech perception // Информационные технологии. 2019. Т. 25, № 5. С. 313–319.
12. Graves A., Mohamed A., Hinton G. Speech recognition with deep recurrent neural networks // Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 2013. P. 6645–6649.
13. Sainath T. N., Mohamed A., Kingsbury B., Ramabhadran B. Deep Convolutional Neural Networks for LVCSR // Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 2013. P. 8614–8618.
14. Surbhi G., Sham K., Adam T. Kalai, Cyril Z. Recurrent Convolutional Neural Networks Learn Succinct Learning Algorithms // arXiv, preprint arXiv:2209.00735v2, 2022.
15. Pronina O., Piatykov O. The recognition of speech defects using convolutional neural network // CTE Workshop Proceedings. 2023. Vol. 10. P. 153–166. doi:10.55056/cte.554.
16. Алферова Л. Д. Пособие для самостоятельной работы «Речевой тренинг: дикция и произношение», 2003.
17. Лопатина Л. В. Проявления и диагностика фонетических нарушений при стертой дизартрии // Известия РГПУ им. А. И. Герцена. 2006. № 14. С. 219–230.

Поступила в редакцию 25.03.2025, окончательный вариант — 15.04.2025.

Березина Анастасия Олеговна, студентка программы специалитета 5 курс, кафедра радиотехника и радиоэлектронные системы ПГУ, ✉ bereanas@mail.ru

Медведев Алексей Валерьевич, студент магистратуры, кафедра «Вычислительная техника» ПГУ, mdl-studio@yandex.ru

Шнайдер Кирилл Сергеевич, студент 3-го курса бакалавриата, кафедра «Вычислительная техника», ПГУ, knhn0@yandex.ru

Митрохин Максим Александрович, доктор техн. наук, доцент, заведующий кафедрой «Вычислительная техника», ПГУ, mmax83@mail.ru

Computer tools in education, 2025

№ 1: 48–60

<http://cte.eltech.ru>

[doi:10.32603/2071-2340-2025-1-48-60](https://doi.org/10.32603/2071-2340-2025-1-48-60)

Integration of Neural Network Technologies into Modern Education

Berezina A. O.¹, Student, ✉ bereanas@mail.ru

Medvedev A. V.¹, Student, mdl-studio@yandex.ru

Schneider K. S.¹, Student, knhn0@yandex.ru

Mitrokhin M. A.¹, Doctor Sc., Docent, mmax83@mail.ru

¹Penza State University, 40 Krasnaya st., 440026, Penza, Russia

Abstract

The article considers an approach to the integration of neural network technologies to eliminate the imbalance between the development of hard (hard skills) and flexible (soft skills) skills in the educational process. The modern labor market requires high communication, creative and logical skills from specialists, but traditional educational programs do not sufficiently focus on soft skills. The authors of the article, based on research by RBC, Google and RSM, emphasize the importance of flexible skills for successful careers and social adaptation, noting that many people have difficulties with communication skills as the fundamental soft skills.

To solve this problem, an innovative approach is proposed using neural network technologies to diagnose and develop communication skills. The authors analyze existing deep learning models and propose their own convolutional recurrent neural network (CRNN) structure for diagnosing speech deficiencies in the Russian language. The developed model evaluates pronunciation defects and provides personalized learning materials.

The authors propose an interactive educational platform that implements the created model within the framework of hard skills training technology in conjunction with soft skills development programs. The neural network algorithms of the platform optimize the learning process, adapting it to the individual characteristics of the student, and can be used both independently and in addition to classes with a tutor.

Keywords: *Soft skills, speech technique, neural network technologies, diction deficiencies, platform.*

Citation: A. O. Berezina, A. V. Medvedev, K. S. Schneider, M. A. Mitrokhin, "Integration of Neural Network Technologies into Modern Education," *Computer tools in education*, no. 1, pp. 48–60, 2025 (in Russian); [doi:10.32603/2071-2340-2025-1-48-60](https://doi.org/10.32603/2071-2340-2025-1-48-60)

References

1. S. Godin, "Let's stop calling it soft skills. Talent Management," in *Medium*, 2017. [Online]. Available: <https://itsyourturnblog.com/lets-stop-calling-them-soft-skills-9cc27ec09ecb>
2. C. R. Mann, *A study of engineering education, Joint committee on engineering education of the national engineering societies*, New York, USA: New York city, 1918.
3. K. Nordstrom and J. Ridderstrale, *Business in the style of funk. Capital dances to the tune of talent*, Moscow: Mann, Ivanov and Ferber, 2013 (in Russian).

4. T. Yu. Yakuba et al. eds., *Soft skills — competencies of the future: a methodological manual* / Far Eastern State Scientific Library, department of scientific research and scientific-methodological work, Khabarovsk, Russia: DVGBN, 2023 (in Russian).
5. A. Nikulina, “7 tips: how to overcome the fear of public speaking,” in *ITMO News*, 2020. [Online]. Available: <https://news.itmo.ru/en/education/trend/news/9396/>
6. M. Alam, M. D. Samad, L. Vidyaratne, A. Glandon, and K. M. Iftekharruddin, “Survey on Deep Neural Networks in Speech and Vision Systems” *Neurocomputing*, vol. 417, pp. 302–321, 2020; doi: 10.1016/j.neucom.2020.07.053
7. D. Ribas, M. A. Pastor, A. Miguel, D. Martinez, A. Ortega, and E. Lleida, “Automatic voice disorder detection using self-supervised representations,” *IEEE Access*, vol. 11, pp. 14915–14927, 2023.
8. A. A. Joshy and R. Rajan, “Automated dysarthria severity classification: A study on acoustic features and deep learning techniques,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1147–1157, 2022.
9. S. Venugopalan et al., “Speech intelligibility classifiers from 550 disordered speech samples,” 2023, arXiv:2303.07533v2 [eess.AS].
10. A. A. Khomenko, I. V. Zinchenko, and Yu. V. Bryzgalova, “Diagnostics of speech disorders using artificial intelligence as a diagnostic technology for detecting speech disorders,” *Pedagogical Perspective*, vol. 13, no. 1, pp. 58–65, 2024 (in Russian); doi:10.55523/27822559_2024_1(13)_58
11. L. V. Savchenko, “Computer-Assisted Language Learning Based on Convolutional Neural Networks and Information Theory of Speech Perception,” *Information Technologies*, vol. 25, no. 5, pp. 313–319, 2019 (in Russian); doi:10.17587/it.25.313-318
12. A. Graves, A. Mohamed, and G. Hinton, “Speech recognition with deep recurrent neural networks,” in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 2013*, pp. 6645–6649, 2013; doi:10.1109/ICASSP.2013.6638947
13. T. N. Sainath, A. Mohamed, B. Kingsbury, and B. Ramabhadran, “Deep Convolutional Neural Networks for LVCSR,” in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 2013*, pp. 8614–8618, 2013; doi:10.1109/ICASSP.2013.6639347
14. G. Surbhi, K. Sham, A. T. Kalai and Z. Cyril, “Recurrent Convolutional Neural Networks Learn Succinct Learning Algorithms,” 2022, arXiv:2209.00735v2 [cs.LG].
15. O. Pronina and O. Piatyko, “The recognition of speech defects using convolutional neural network,” *CTE Workshop Proceedings*, vol. 10, pp. 153–166, 2023; doi:10.55056/cte.554
16. L. D. Alferova, *Rechevoj trening: dikcija i proiznoshenie. Posobie* [Diction and pronunciation. Manual], St. Petersburg, Russia: RGSI, 2003 (in Russian).
17. L. V. Lopatina, “Proyavleniya i diagnostika foneticheskikh narusheniya pri stertoj dizartrii” [Manifestations diagnosis of phonetic disorders in erased dysarthria], *Izvestiya RGPU im. A. I. Gertsena*, vol. 6, no. 14, pp. 219–230, 2006 (in Russian).

Received 25-03-2025, the final version — 15-04-2025.

Anastasia Berezina, 5th year Student of the Department of Radio Engineering and Radioelectronic Systems, Penza State University, ✉ bereanas@mail.ru

Aleksey Medvedev, Master's Degree student of the Department of Computer Engineering, Penza State University, mdl-studio@yandex.ru

Kirill Schneider, 3rd year Student of the bachelor's degree program of the Department of Computer Engineering, Penza State University, knhn0@yandex.ru

Maksim Mitrokhin, Doctor of Sciences (Tech.), Docent, Head of the Department of Computer Engineering, Penza State University, mmax83@mail.ru