

ПРИМЕНЕНИЕ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ВОССТАНОВЛЕНИЯ ЗАШУМЛЕННОЙ ИНФОРМАЦИИ

В данной статье рассматривается возможность восстановления зашумленной информации с помощью рекуррентных нейронных сетей. Определены источники шума, его классификация, рассмотрены доступные методы восстановления информации. Представлены примеры задач восстановления зашумленной информации, решения которых возможно с помощью рекуррентных нейронных сетей. Выделены преимущества и недостатки их использования, описаны примеры успешной реализации. Выдвинуто предположение относительно дальнейших шагов в исследовании данного направления.

ВВЕДЕНИЕ

В нейронной сети каждый нейрон принимает входные данные, обрабатывает их с помощью весов и активационной функции, а затем передает результат на следующий уровень. Такой подход позволяет сети обучаться на основе примеров, адаптируя свои параметры для улучшения точности предсказаний [1]. Это делает нейронные сети мощным инструментом, позволяя моделировать им сложные зависимости и паттерны в данных. Они служат основой для разных приложений на основе технологий искусственного интеллекта, включая средства распознавания изображений, обработку естественного языка и генерацию контента.

Существуют различные типы нейронных сетей. Одним из них являются рекуррентные сети, где выход сети зависит от предыдущих вычислений [2].

Шум в данных – это случайные флуктуации значений признаков, которые не несут полезной информации. Причинами возникновения шума могут быть ошибки измерений, помехи в процессе передачи данных и переизбыток информации. Зашумленная информация может затруднять анализ, искажать результаты и приводить к неправильным решениям, принятым на основе неверных данных.

Среди основных подходов к восстановлению зашумленной информации можно выделить применение нейронных сетей, что широко используется для восстановления зашумленных сигналов и изображений.

Архитектура рекуррентных нейронных сетей основана на последовательном распространении информации, что делает их особенно подходящими для задач, связанных с временными рядами и последовательностями, такими как обработка естественного языка [1]. Это позволяет рекуррентным нейронным сетям учитывать контекст и предшествующие состояния, что критично для восстановления информации из зашумленных данных. Такая архитектурная особенность делает рекуррентные сети подходящим инструментом для восстановления зашумленной информации.

1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ

Концепция рекуррентных нейронных сетей была впервые предложена Дэвидом Румельхартом, в

журнале Nature в 1986 году, для описания новой процедуры обучения [3]. Цель ее разработчиков заключалась в имитации человеческой памяти.

Рекуррентные нейронные сети ориентированы на задачи, связанные с временными рядами и последовательностями. Обеспечение эффективной обработки последовательных данных является основной целью рекуррентных нейронных сетей.

Архитектура рекуррентных сетей, как и архитектура других нейронных сетей, включает в себя нейроны, слои и связи. Связи между нейронами формируют направленную последовательность, что позволяет им обрабатывать последовательные данные.

Нейроны являются основными строительными блоками любой нейронной сети, они обрабатывают входные данные и передают результаты. В рекуррентных сетях нейроны могут иметь связи не только с нейронами в предыдущем слое, но и с нейронами в том же слое [4]. Такая архитектура позволяет учитывать информацию из предыдущих временных шагов.

Для определения выхода на основе входных данных, нейроны могут использовать различные функции активации. Функции активации вводят нелинейность, ограничивают амплитуду выходного сигнала нейрона, и способны влиять на скорость обучения. В рекуррентных нейронных сетях используются различные активационные функции, включая сигмоиду, гиперболический тангенс и линейную функцию. Каждая из этих функций имеет свои особенности и может быть выбрана в зависимости от конкретной задачи и архитектуры сети.

Рекуррентные нейронные сети, как и другие вариации нейронных сетей, состоят из входного, скрытых и выходного слоев. Они обрабатывают входные данные и передают результаты на следующий уровень. Организация скрытых слоев позволяет сети запоминать информацию о предыдущих входах, и изменять веса не только в зависимости от входных данных, но и от предыдущих выходов. Рекуррентная нейронная сеть разворачивается в полную сеть. Если последовательность представляет собой предложение из некоторого количества слов, то развертка будет состоять из количества слоев, равного количеству слов в предложении.

Связи в нейронных сетях представляют собой математические отношения между нейронами, ко-

которые позволяют передавать информацию от одного нейрона к другому [1]. Сигнал в рекуррентных нейронных сетях движется не только в прямом направлении, но и в обратном. Как следствие, результат выхода способен возвращаться на вход. Выход нейрона определяется весовыми характеристиками и входными сигналами, что также дополняется предыдущими выходами. За счет изменения веса не только в зависимости от входных данных, но и от предыдущих выходов, нейросетям присуща функция кратковременной памяти, что позволяет сигналам восстанавливаться и дополняться во время их обработки.

Такая архитектура позволяет рекуррентным сетям эффективно обрабатывать данные, где важен порядок. Например, в задачах обработки естественного языка, данные могут учитывать контекст слов в предложении, что способствует улучшению качества предсказаний.

В отличие от сверточных нейронных сетей, которые эффективно извлекают пространственные зависимости в изображениях и видео, рекуррентные нейронные сети способны учитывать порядок и контекст входных данных [5]. Это адаптирует их для задач обработки естественного языка и анализа временных рядов. Полносвязные нейронные сети не могут учитывать последовательность данных, что ограничивает их применение в ситуациях, где порядок имеет значение.

Обучение нейронных сетей – это процесс, в ходе которого модель искусственного интеллекта учится выполнять определенные задачи на основе предоставленных ей данных и меток, из которых можно сделать вывод о правильности обучения. Сеть учится на примерах и постепенно улучшает свои предсказания, минимизируя ошибки. Подход к обучению зависит от разных факторов, что включают цель обучения, тип решаемой задачи, характеристику и размер данных, выбранную архитектуру.

Обучение рекуррентных нейронных сетей является сложной задачей, так как они подвержены проблеме исчезающих градиентов, что приводит к медленному или нестабильному обучению, и трудностям в изучении долгосрочных зависимостей. Основополагающим методом для обучения рекуррентных нейронных сетей является обратное распространение ошибки, который позволяет оптимизировать веса и улучшать предсказательную способность модели. Этот метод основан на градиентном спуске, минимизирующем функцию потерь, которая измеряет, насколько хорошо модель предсказывает выходные данные.

Для решения проблемы затухания градиентов, на базе стандартной архитектуры, было разработано несколько расширенных вариантов, основанных на принципе сохранения памяти.

Долгая краткосрочная память (LSTM) является вариацией рекуррентной нейронной сети. Ее назначением является решение проблемы исчезающего градиента. В данной архитектуре увеличен объем

памяти для размещения данных за более длительный период, что позволяет эффективно запоминать долгосрочные зависимости [6].

Управляемые рекуррентные блоки (GRU) – это вариация, являющаяся упрощенной версией LSTM, также поддерживающая выборочное сохранение памяти. Данная модель объединяет некоторые функции и имеет меньшую вычислительную сложность. Обязательным элементом ее скрытого слоя являются плюзы, позволяющие хранить информацию в памяти и удалять ее [7]. Выбор между описанными архитектурами зависит от конкретной задачи и требований к модели. LSTM и GRU эффективнее работают с памятью и позволяют лучше справляться с долгосрочными зависимостями в данных. Стандартная модель больше подходит для простых задач, где не требуется запоминание долгосрочных зависимостей.

Рекуррентные нейронные сети представляют собой особый тип нейронных сетей, которые предназначены для обработки последовательных данных. В отличие от традиционных нейронных сетей, где информация передается только в одном направлении, рекуррентные сети имеют циклические связи, что позволяет им сохранять информацию о предыдущих входах и использовать ее для предсказания следующих значений в последовательности.

II. ПОНЯТИЕ ЗАШУМЛЕННОЙ ИНФОРМАЦИИ

Развитие информационных технологий ведет к увеличению объемов данных и источников информации. Огромное количество информации образуется в таких источниках, как социальные сети, интернет вещей и медицинская диагностика. Накопление информации требует решения задач фильтрации шума и восстановления полезного сигнала.

Шум в данных – это случайные флуктуации значений признаков, которые не несут полезной информации и могут иметь различные источники. Причинами возникновения шума являются ошибки измерений, помехи в процессе передачи данных и переизбыток информации, что ведет к трудностям в анализе и интерпретации данных. Классификация шума способствует пониманию того, как он влияет на восприятие информации. В контексте обработки сигналов, шум представляет собой составляющую сигнала, имеющую случайный характер.

Основные источники шума могут включать в себя внешние и внутренние факторы. Внешние источники могут быть связаны с условиями сбора данных, такими как плохое освещение для изображений или фоновый шум в аудиозаписях. Внутренние источники шума могут возникать из-за ошибок в алгоритмах обработки данных или неправильной настройки параметров модели.

Важной характеристикой является соотношение сигнала и шума, которое определяет силу полезного сигнала относительно фонового шума. Это соотношение критично для оценки качества передачи и обработки данных.

Шум в данных может быть случайным и систематическим. Случайный шумом называется тот, который имеет случайный характер и не несет полезной информации. Систематический шум, в отличие от случайного, представляет собой предсказуемые и повторяющиеся ошибки, и может возникать из-за неправильной калибровки оборудования или других систематических факторов, что способны привести к систематическим ошибкам в выводах. Оба типа шума могут существенно влиять на качество данных и результаты их анализа, поэтому важно применять соответствующие методы для их идентификации и устранения.

Зашумленной информацией называется информация, содержащая элементы, искажающие ее смысл или усложняющие понимание. Данный термин часто используется для обозначения информационного шума, который описывает наличие ненужных данных, мешающих восприятию смысла.

В более широком смысле, зашумленность может относиться к любым факторам, которые ухудшают качество передачи или восприятия информации. Это может включать технические помехи, ошибки в данных или даже культурные и языковые барьеры [6]. Одним из основных типов является информационный шум, который представляет собой элементы, усложняющие понимание текста и искажающие его смысл. Этот шум может возникать из-за избыточной информации, неуместных данных или нечетких формулировок, что затрудняет восприятие и анализ информации.

К распространенным типам искаженной информации относятся: искажения или помехи в аудиозаписях, потеря четкости видео, шум на спутниковых изображениях и результатах МРТ, искажение данных из-за ошибок измерений и случайных колебаний, а также аномалии, возникающие в процессе передачи информации по каналам связи, вызванные электромагнитными помехами и внутренними ошибками в кодировании. Каждый из этих типов зашумленной информации требует специфических методов для восстановления и очистки.

Зашумленная информация может затруднять анализ, искажать результаты и приводить к неправильным решениям, принятым на основе неверных данных. Это обосновывает актуальность процесса фильтрации и очистки информации для достижения более точных результатов.

III. ВОССТАНОВЛЕНИЕ ЗАШУМЛЕННОЙ ИНФОРМАЦИИ

Восстановление зашумленной информации – это процесс, направленный на улучшение качества данных, которые были искажены или затенены шумом.

Существует несколько методов и алгоритмов, которые могут быть использованы для этой цели, среди них можно выделить:

- алгоритмы фильтрации;
- статистические методы;

- алгоритмы машинного обучения.

Одним из основных подходов является использование фильтров, предназначенных для обработки данных с целью удаления шума или выделения полезной информации. Они могут быть как простыми, так и сложными.

Алгоритмы фильтрации могут использоваться в различных направлениях, включая обработку сигналов, компьютерное зрение и анализ временных рядов. Примерами являются фильтры Калмана и медианные фильтры, способные уменьшить шум и сохранить важные детали информации.

Статистические методы основываются на статистических моделях и теории вероятностей, и используются для анализа данных, выявления закономерностей и принятия решений на основе данных. Они могут включать в себя методы, такие как регрессионный анализ, байесовские методы, методы максимального правдоподобия. Данные методы применены для оценки вероятности различных состояний данных и выбора наиболее вероятного.

Нейронные сети широко используются для восстановления зашумленных сигналов и изображений. Распознавание паттернов и восстановление оригинальных данных возможно после обучения сетей на больших наборах данных. Методы, основанные на моделях, используют предположения о структуре данных для восстановления. Для примера, сверточные нейронные сети часто используются для обработки изображений, позволяя эффективно удалять шум и улучшать качество [5].

Указанные методы и алгоритмы могут быть использованы в различных областях, включая обработку сигналов, анализ данных и восстановление зашумленной информации. Важным шагом в процессе восстановления информации является выбор подходящего метода, который зависит от типа данных и характера шума. Корректный выбор позволит достичь наилучших результатов.

IV. ПРИМЕНЕНИЕ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ

Машинное обучение играет ключевую роль в восстановлении зашумленной информации. Методы машинного обучения являются высокоэффективными инструментами для обработки и анализа данных. Эти методы активно развиваются и применяются в различных областях, включая обработку данных и телекоммуникации. Восстановление зашумленных данных с использованием рекуррентных нейронных сетей предполагает несколько ключевых шагов:

- сбор и подготовку данных;
- создание модели;
- обучение модели;
- восстановление данных;
- оценку производительности.

Данные играют ключевую роль в обучении. Они должны быть предварительно подготовлены для обучения, и содержать как чистые, так и зашумленные примеры. Это могут быть временные ряды, аудиосигналы или другие последовательные

данные. Их обработка подразумевает нормализацию и разделение данных на обучающую и тестовую выборки.

Рекуррентные нейронные сети могут иметь различную архитектуру. Выбор архитектуры зависит от сложности задачи и характера данных. Важным шагом является определение как количества слоев, так и количества нейронов в каждом слое. Более сложные архитектуры способны лучше справляться с шумом, но требуют больше вычислительных ресурсов.

Целью обучения является минимизация разницы между восстановленными и чистыми данными. Подобный эффект достижим с помощью функции потерь. В процессе обучения модель учится выявлять паттерны в данных и игнорировать шум, используя информацию из предыдущих временных шагов.

Результатом завершения процесса обучения будет являться готовность модели к использованию для восстановления данных, что предполагает поступление на вход сети зашумленных данных, и последующую генерацию предсказаний, представляющих собой восстановленные данные. Для достижения наилучших результатов может потребоваться дополнительная обработка выходных данных, то есть фильтрация или сглаживание.

По завершению процесса восстановления данных необходимо оценить качество результатов с помощью различных метрик, таких как соотношение уровней сигнала и шума, или структурное сходство. Сравнение восстановленных данных с оригинальными позволит определить эффективность обученной модели.

На основе результатов оценки можно внести изменения в архитектуру модели, параметры обучения или методы обработки данных для улучшения качества восстановления. Возможно комбинирование рекуррентных нейронных сетей с другими типами нейронных сетей для повышения точности.

Восстановление зашумленных данных с помощью рекуррентных нейронных сетей – это итеративный процесс, который требует тщательной настройки и оценки на каждом этапе.

ЗАКЛЮЧЕНИЕ

Рекуррентные нейронные сети представляют собой вариацию архитектуры нейронных сетей. Циклические связи отличают их от других нейронных сетей, позволяя учитывать предшествующую информацию, что делает рекуррентные нейронные сети актуальными для задач, связанных с временными рядами и последовательностями. Рекуррентные сети имеют множество преимуществ для восстановления зашумленной информации, включая обработку последовательностей переменной длины, способность к запоминанию контекста и адаптивность.

Основным классом задач, которые решаются рекуррентными нейронными сетями, является обработка текстов [8]. Обучаясь на больших объемах

данных, модели способны создавать логичные тексты, восстанавливать пропущенные слова или фразы в тексте с учетом контекста и структуры предложений. Они эффективно обрабатывают последовательности слов, учитывая контекст и порядок, что позволяет создавать более точные модели для понимания и генерации языка.

Однако, рекуррентные нейронные сети также сталкиваются с определенными ограничениями, среди которых были выделены: проблемы с долгосрочной зависимостью, сложность обучения, затухающий градиент. Данные ограничения способны замедлить процесс восстановления данных.

Рекуррентные нейронные сети отличаются от других типов нейронных сетей по нескольким ключевым аспектам:

- ориентированы на обработку последовательных данных;
- позволяют учитывать временные зависимости;
- используют обратные связи для адаптации к изменениям во входных данных;
- имеют внутреннюю память.

Существует множество успешных проектов и исследований, где задействованы рекуррентные нейронные сети, включая проекты компаний Google и Apple. Рассматриваемая технология способствует повышению качества перевода, позволяет учитывать контекст и структуру данных. Также она применяется для прогнозирования цен на акции, анализа климатических данных или диагностики здоровья пациентов.

В дальнейшем предполагается практическая реализация шагов, описанных в данной статье, то есть обучение модели и проведения эксперимента по восстановлению данных. Ожидается, что результаты эксперимента позволят оценить эффективность различных архитектур рекуррентных нейронных сетей и выявить оптимальные подходы для восстановления зашумленной информации.

Таким образом, благодаря своей способности обрабатывать последовательные данные и учитывать контекст, рекуррентные нейронные сети представляют собой перспективный подход для восстановления зашумленной информации. Необходимость обработки растущих объемов данных и популяризация технологий, включая IoT и большие данные, позволяет сделать вывод об актуальности рассмотренного направления как сейчас, так и в будущем.

Список литературы

1. Haykin, S. Neural Networks: A Comprehensive Foundation / S. Haykin // Prentice Hall, 1998. – P. 842.
2. Krauss, P. Artificial Intelligence and Brain Research / P. Krauss // Springer Berlin, 2024. – P. 260.
3. Rumelhart, D. Learning representations by back-propagating errors / D. Rumelhart, G. Hinton, R. Williams // Nature. – 1986. – Vol. 323. – P. 533–536.
4. Tyagi, A. Recurrent Neural Networks: Concepts and Applications / A. Tyagi, A. Abraham // CRC, 2022. – P. 396.

5. Patil, V. Restoration of Images using only noisy data A Deep Learning Approach using U-NET / V. Patil, P. Katte // IJRAR. – 2019. – Vol. 6. – P. 988–993.
6. Maher, S. Recurrent Neural Networks for Chatbot Excellence: Examining the Power of LSTM Architecture / S. Maher, R. Gaikwad, S. Nimbhore // ICT for Intelligent Systems. – 2024. – P. 513–522.
7. Das, S. Recurrent Neural Networks (RNNs): Architectures, Training Tricks, and Introduction to Influential Research / S. Das, A. Tariq, T. Santos // Machine Learning for Brain Disorders. Neuromethods. – 2023. – Vol. 197. – P. 117–138.
8. Lai, S. Recurrent Convolutional Neural Networks for Text Classification / S. Lai, K. Lui, J. Zhao // AAAI Conference on Artificial Intelligence. – 2015. – P. 2267–2273.

Славинский Григорий Александрович, магистрант кафедры информатики Белорусского государственного университета информатики и радиоэлектроники, gr.slavin.sci@mail.ru.

Научный руководитель: Боброва Наталья Леонидовна, доцент кафедры информатики Белорусского государственного университета информатики и радиоэлектроники, bobrova@bsuir.by.