

Глубокое машинное обучение

ПРИ-120
Баймурзин Д.А.

Оглавление

1. Введение
2. Алгоритм обратного распространения ошибки
3. Компоненты нейронной сети
4. Процесс обучения глубоких сетей
5. Применение
6. Польза
7. Преимущества
8. Будущее сетей
9. Заключение

Введение

Глубокое машинное обучение - это подраздел машинного обучения, который уделяет особое внимание использованию искусственных нейронных сетей для решения сложных задач. Оно основывается на принципах и алгоритмах обратного распространения ошибки, которые позволяют нейронным сетям самостоятельно извлекать иерархические признаки из больших объемов данных.

Алгоритм обратного распространения ошибки

Этот алгоритм начинается с передачи входных данных через нейронную сеть, после чего происходит сравнение полученных выходных значений с ожидаемыми целевыми значениями. Разница между предсказанными и целевыми значениями называется ошибкой.

Затем происходит процесс обратного распространения ошибки, где эта ошибка передается назад через сеть от последнего слоя к первому. При этом каждый нейрон в сети вносит свой вклад в ошибку в соответствии с его весами и активационной функцией. Веса каждого нейрона корректируются таким образом, чтобы уменьшить ошибку.

Алгоритм выполняется в цикле, называемом эпохой, на протяжении нескольких повторений. За каждую эпоху веса нейронной сети корректируются с учетом ошибок и обновленных целевых значений.

Компоненты нейронной сети

Входной слой

Искусственная нейронная сеть содержит несколько узлов, через которые в нее поступают данные. Эти узлы составляют входной слой системы.

Скрытый слой

Входной слой обрабатывает и передает данные на слои, расположенные дальше в нейронной сети. Эти скрытые слои обрабатывают информацию на разных уровнях, адаптируя свое поведение по мере получения новых данных. Сети глубокого обучения имеют сотни скрытых слоев, которые они могут использовать для анализа проблемы с разных точек зрения.

Выходной слой

Выходной слой состоит из узлов, которые выводят данные. Модели глубокого обучения, которые выводят ответы «да» или «нет», имеют на выходном слое только два узла. С другой стороны, модели, которые выводят более широкий диапазон ответов, содержат больше узлов.

Процесс обучения глубоких сетей

1. Подготовка данных
2. Создание архитектуры сети
3. Инициализация весов
4. Прямой проход
5. Подсчет функции потерь
6. Обратное распространение ошибки
7. Повторение шагов 4-6
8. Оценка модели
9. Подстройка гиперпараметров
10. Использование модели

1 шаг. Подготовка данных

Для обучения глубоких нейронных сетей необходимо подготовить данные, которые будут использоваться в процессе обучения. Это может включать в себя загрузку данных, предварительную обработку, нормализацию и разделение на тренировочную и тестовую выборки.

2 шаг.Создание архитектуры сети

Для каждой задачи выбирается соответствующая архитектура сети, которая определяет количество слоев, количество нейронов в каждом слое и тип функций активации. Архитектура сети должна быть оптимальной для конкретной задачи.

3 шаг. Инициализация весов

Перед началом обучения сети веса всех нейронов инициализируются случайными значениями, обычно из нормального распределения. Это позволяет сети иметь некоторую начальную "способность обучения".

4 шаг. Прямой проход

Во время прямого прохода данные подаются на вход сети, проходят через слои нейронов и вычисляются выходные значения. В каждом слое происходит линейная комбинация входных значений с весами нейронов и применение функции активации, чтобы получить выходной сигнал.

5 шаг. Подсчет функции потерь

После прямого прохода сравниваются выходные значения сети с ожидаемыми выходами и вычисляется ошибка. Функция потерь, такая, как среднеквадратичная ошибка (MSE) или перекрестная энтропия, используется для оценки разницы между предсказанными и истинными значениями.

6 шаг. Обратное распространение ошибки

Основным шагом в обучении нейронных сетей является обратное распространение ошибки. Ошибка вычисляется на каждом слое и распространяется обратно, чтобы обновить веса нейронов. В каждом слое происходит вычисление градиента функции потерь по весам и обновление весов с помощью алгоритма градиентного спуска или его вариантов.

7 шаг. Повторение шагов 4-6

Процессы прямого прохода, вычисления ошибки и обратного распространения повторяются на каждой эпохе обучения, пока не достигнута заранее заданная точность или количество эпох.

8 шаг. Оценка модели

После завершения обучения проводится оценка модели на тестовой выборке или на новых данных, чтобы получить информацию о ее обобщающей способности и точности.

9 шаг. Подстройка гиперпараметров

При необходимости можно провести оптимизацию гиперпараметров модели, таких как скорость обучения, размер пакета и количество слоев. Это поможет улучшить производительность и точность сети.

10 шаг.Использование модели

После успешного обучения сеть готова к использованию для предсказания новых данных или для решения задачи, для которой она обучалась.

Обработка естественного языка и распознавания речи

Глубокие нейронные сети были успешно применены в области распознавания речи. Они могут обучаться на больших объемах аудиоданных и извлекать высокоуровневые признаки, что облегчает автоматическое распознавание произношения и перевод речи в текст.

Также они обладают способностью обрабатывать последовательности и последовательно связывать знания в тексте, что позволяет им эффективно работать с естественным языком.

Применение в области компьютерного зрения

Распространение глубокого машинного обучения в области компьютерного зрения и распознавания образов привело к значительному прогрессу в различных задачах, таких как классификация изображений, семантическая сегментация, детектирование объектов и распознавание лиц.

Польза от глубокого обучения

- В беспилотных автомобилях используются модели глубокого обучения для автоматического обнаружения дорожных знаков и пешеходов.
- Системы защиты используют глубокое обучение для автоматического обозначения областей интереса на спутниковых снимках.
- В анализе медицинских изображений глубокое обучение используется для автоматического обнаружения раковых клеток в ходе медицинской диагностики.
- Заводы используют приложения глубокого обучения, чтобы автоматически выявлять, когда люди или объекты находятся на небезопасном расстоянии от машин.

Преимущества и ограничения глубокого машинного обучения:

- Высокая точность и способность к обучению на больших данных
- Сложность интерпретации и объяснения решений модели
- Необходимость больших вычислительных ресурсов

Будущее глубокого машинного обучения:

- Развитие и использование более сложных архитектур нейронных сетей
- Применение в автономных транспортных средствах и робототехнике
- Влияние на область труда и потенциал автоматизации

Заключение

Глубокое машинное обучение продолжает развиваться и исследоваться, привлекая внимание и инвестиции со стороны многих компаний и исследовательских центров. Его потенциал в области решения сложных задач и создания новых технологий делает его одной из самых захватывающих и перспективных областей в сфере искусственного интеллекта.

Список литературы

1. "Deep Learning" by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.
2. "Recurrent Neural Networks" by Juergen Schmidhuber.
3. deep learning // techtarget URL:
<https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network> (дата обращения: 01.10.20223).
4. J. J. Hopfield. [[PNAS Reprint \(Abstract\)](#) [PNAS Reprint \(PDF\)](#) Neural networks and physical systems with emergent collective computational abilities] // Proceedings of National Academy of Sciences. — April 1982. — С. vol. 79 no. 8 pp. 2554—2558.
5. David Barber, Bayesian Reasoning and Machine Learning [greenteapress.com] - David Verber - Режим доступа: <https://greenteapress.com/wp/think-bayes> (дата обращения: 01.10.20223).