РАЗДЕЛ IX. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Басов С.А, Юркин В.А, Сараджишвили С.Э.

Обучение нейронной сети с использованием ограниченной машины Больцмана для распознавания изображений

Санкт-Петербургский политехнический университет Петра Великого
(Россия, Санкт-Петербург)

doi:10.18411/spc-22-12-2017-16

idsp: 000001:spc-22-12-2017-16

Машины Больцмана (RBM), как подвид рекуррентных нейронных сетей, являются сетями с обратной связью. Процесс обучения данного вида сетей чуть более сложен, что привело к поиску новых алгоритмов. Первым этапом на пути к RBM стали сети Хопфилда, затем были придуманы обыкновенные машины Больцмана, отличающиеся стохастической структурой, и тем, что нейроны в такого рода сетях, разделяются на две группы, одна из которых отражает обозреваемые состояния, другая – скрытые (как в скрытых моделях Маркова).

На данном этапе, машина Больцмана представляет собой полносвязный неориентированный граф, такой что, любые две его вершины одной группы имеют зависимость друг от друга.

Изначально, подобного рода алгоритмы применялись для работы с бинарными данными, поскольку были разработаны на основе стохастических бинарных нейронов. Однако, в данном случае, было решено использовать сеть для работы с изображениями, которые в большинстве своем представлены в оттенках серого и обладают значением яркости каждого пикселя от 0 до 255. Одним из возможных вариантов решения такой проблемы – это изменение значения яркости таким образом, чтобы оно находилось в заданном промежутке от 0 до 1, для этого необходимо разделить исходные значения на 255, и допустить, что пиксели бинарны, а значения, получаемые на выходе, имеют вероятность установки каждого конкретного рассчитанного вновь пикселя, в единицу.

Главное проблемой является то, что пиксели в реальных изображениях имеют зависимость друг от друга, такую, что интенсивность каждого конкретного будет равняться средней интенсивности соседнего. Это говорит о том, что значение интенсивности должно иметь большую вероятность находиться в диапазоне близком к среднему значению, нежели быть отдаленным от него. Однако, логистическая функция имеет иное распределение и не позволяет достичь таких показателей, хотя и работает в некоторых задачах, где данный параметр является не столь важным.

Зададим значения интенсивности группы видимых нейронов такое что, нормальным будет считаться 0.61, менее вероятным 0.59 и 0.63, еще менее вероятным 0.5 и 0.7. В итоге получаем:

Рисунок 1 - функция плотности вероятности.

Сделав значения группы видимых нейронов случайными величинами, данное распределение будет использоваться для моделирования поведения нейронов. Такой подход можно использовать не только в задачах связанных с обработкой изображений, но и с другими данными, которые представлены действительными числами от $-\infty$ до $+\infty$. В данном случае, скрытые нейроны все также остаются бинарными, что позволяет получить Gaussian-Binary RBM. Значения распределения нейронов можно выразить формулами:

$$p(v_i | \mathbf{h}) = N(v_i | b_i + \sigma_i \sum_{j \in \text{hid}} h_j w_{i,j}, \sigma_i^2) \quad (1),$$

$$p(h_j = 1 | \mathbf{v}) = \frac{1}{1 + \exp(-\sum_{i \in \text{vis}} \frac{v_i}{\sigma_i} w_{i,j} - b_j)} \quad (2),$$

Энергия Больцмана, тогда:

$$E(\mathbf{v}, \mathbf{h}) = \sum_{i \in \text{vis}} \frac{(v_i - b_i)^2}{2\sigma_i^2} - \sum_{j \in \text{hid}} b_j h_j - \sum_{i,j} \frac{v_i}{\sigma_i} h_j w_{ij} \quad (3)$$

где hid – это множество индексов группы скрытых нейронов,

vis – множество индексов группы видимых,

b – коэффициент смещения,

σ_i – коэффициент стандартного отклонения видимого нейрона на каждом шаге,

$w_{i,j}$ – вес связи нейронов на i -ом и j -ом шаге,

$N(x | \mu, \sigma^2)$ – расчет значения вероятности для переменной, где μ – матожидание, σ^2 – дисперсия, при нормальном распределении.

На данном этапе, действительные данные представляются в виде видимых нейронов, внутреннее же их скрытое состояние бинарно. Далее необходимо «научить» нейроны отображать натуральные числа больше единицы. Для этого необходимо создать для скрытых нейронов множество копий с аналогичными весами видимых нейронов и обучаемым смещением, но при расчете вероятностей, необходимо будет от данного смещения вычесть заданные ранее фиксированные значения.

После данных преобразований, все наши скрытые нейроны стали из бинарных линейными с Гаусовским шумом, без изменения алгоритма. Теперь нейроны способны нести в себе информацию в виде любого неотрицательного действительного числа.

Реализация обучения. Значение видимого нейрона вычисляем по формуле:

$$v_i = b_i + \sum_{j \in \text{hid}} h_j w_{i,j} + N(0, 1) \quad (4),$$

Значение скрытого нейрона:

$$h_j = \max(0, b_j + \sum_{i \in \text{vis}} v_i w_{i,j} + N(0, \sigma(x))) \quad (5)$$

Тестирование. В качестве входных данных, были использованы готовые наборы изображений людей, выражающих какие-либо эмоции, в количестве 700 шт. Первым шагом на изображении выделяется лицо, полученный участок сжимается до заданного размера и переводится в оттенки серого. Затем переводим изображения в вектор размерностью 1 на 4480 и записываем в матрицу.

Для того чтобы избежать переобучения, зададим количество нейронов в слоях сети равное: 4480 – 200 – 300 – 500 – 8. Для наглядности, сначала проведем обучение сети стандартной сигмоидальной функцией, затем с использованием полученного алгоритма предобучения.

В первом случае, среднее значение ошибки составило ~36%, во втором случае при тех же условиях ~28%.

Рисунок 2 – средняя ошибка.

В результате работы, был написан алгоритм обучения нейронной сети на основе ограниченных машин Больцмана для работы с изображениями. Сравнительный анализ показал, что данный подход является более эффективным, нежели обучение с сигмоидальной функцией активации.

1. Alex Krizhevsky - Learning Multiple Layers of Features from Tiny Images, 2009г.
2. Geoffrey Hinton, Department of Computer Science, University of Toronto - A Practical Guide to Training Restricted Boltzmann Machines, Version 1, 2010г.
3. Басов С. А., Юркин В. А., Финогенов А. Э., Сараджишвили С. Э. Система биометрической идентификации личности по 3d модели лица, получаемой с помощью стереокамер, использующая математическую модель движения мимических мышц лица // Сборник докладов студенческой научной конференции Института компьютерных наук и технологий. Издательство: Федеральное государственное автономное образовательное учреждение высшего образования "Санкт-Петербургский политехнический университет Петра Великого" (Санкт-Петербург), 2016г.

Ворожцова Н.А.

Исследование методов и систем семантического анализа текстов

ФГБОУ ВО «ИжГТУ имени М.Т. Калашникова»

(Россия, Ижевск)

doi:10.18411/spc-22-12-2017-17

idsp: 000001:spc-22-12-2017-17

Аннотация

В связи с быстрым ростом количества текстовой информации все более актуальной становится проблема автоматической обработки текстов. Одной из основных задач в рамках данной проблемы является задача семантического анализа. В работе рассмотрены методы и существующие системы семантического анализа текстов, отмечены их основные достоинства и недостатки, сделан вывод по результатам исследования.

Семантический анализ – процесс выявления смыслового содержания слов и словосочетаний в предложении. Большинство методов семантического анализа работают со смыслом слов.