Искусственные нейронные сети

Здравствуйте, уважаемые слушатели! Тема нашей лекции – Искусственные нейронные сети.

План лекции:

- Искусственные нейронные сети
- Математическое описание искусственной нейронной сети

1. Вступительное слово

Как вы помните из прошлых уроков машинное обучение всего лишь часть искусственного интеллекта. Но мы не описали происхождение самого искусственного интеллекта как науку. Есть предположение что все началось с теста Тьюринга, который был опубликован 1950 году в статье « Computing Machinery and Intelligence». При этом сама идея обсуждалась примерно 10 годами ранее самим Тьюрингом и его коллегами. На сегодняшний день суть теста заключается в следующем: компьютер должен выдать себя за человека в диалоге между судьями. Другими словами, компьютер должен владеть естественным языком, который не отличается от человеческого. Хотя изначальная формулировка задачи отличается от нынешней, где основной идеей была умение судьей различить мужчин от женщин: задача мужчин выдать себя за женщин, а задача женщин помочь разобраться судьям кто есть кто.

Благодаря данному тесту появилась ненаучная деятельность по созданию чатботов. А в 2014 году данный тест успешно прошел чатбот под названием «Женя Густман».

Так же как машинное обучение искусственные нейронные сети (Artificial Neural Networks – ANN – ИНС) являются частью большой науки под названием искусственный интеллект.

2. Искусственные нейронные сети

Искусственные нейронные сети – аппарат, который активно исследуется начиная с 40-х годов прошлого столетия. Немаловажный вклад в развититие этой сферы включил Ф. Розенблат, который разработал в 1950-х годах однослойный персептрон [1]. Недостатки однослойного персептрона отражены в книге М. Минского и С. Пейперта [2, 3]. В этом уроке мы рассмотрим ограничения однослойной нейронной сети и покажем, что она не способна решать некоторые классические логические задачи, в частности, обозначена знаменитая проблема неразрешимости функции XOR для однослойной нейронной сети. Преодолеть этот недостаток можно было путем использования многослойных нейронных сетей. Однако в конце 60-х годов было еще неясно, как обучать многослойные нейронные сети.

В 1974 году был предложен алгоритм, который впоследствии получил название «алгоритм обратного распространения» (backpropagation) [4, 5], или «алгоритм обратного распространения ошибки», пригодный для автоматического подбора весов (обучения) многослойного персептрона или многослойной нейронной сети прямого распространения. Этот алгоритм стал базой для бурного развития нейросетевых методов вычислений.

Наиболее популярная архитектура ANN – сеть прямого распространения, в которой нелинейные элементы (нейроны) представлены последовательными слоями, а информация распространяется в одном направлении (Feed Forward Neural Networks) [6]. В 1989 году в

---

1 Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain // Psychological Review. – 1958. – Vol. 65 (6). – P. 386–408.

2 Minsky M. L., Papert S. A. Perceptrons: An Introduction to Computational Geometry. – MIT, 1969. – 252 p.

3 Marvin Minsky, Seymour Papert. Perceptrons, expanded edition. – The MIT Press, 1987. – 308 p.

4 Werbos P. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. – Harvard University, 1974. – 38 p.

5 Werbos P. J. Backpropagation: past and future // IEEE International Conference on Neural Networks. – San Diego, 1988. – Vol. 1. – P. 343–353.

6 David Saad. Introduction. On-Line Learning in Neural Networks. – Cambridge University Press, 1998. – P. 3–8.

работах G. Gybenco [7], K. Hornik [8] и др. показано, что такая сеть способна аппроксимировать функции практически любого вида. Однако в тот период теоретическая возможность была существенно ограничена вычислительными мощностями. Преодолеть этот разрыв удалось в 90-х годах, когда были предложены сети новой архитектуры, получившие впоследствие название глубоких нейронных сетей.

3. Математическое описание искусственной нейронной сети

Рассмотрим ANN с прямым распространением сигнала. В такой сети отдельный нейрон представляет собой логистический элемент, состоящий из входных элементов, сумматора, активационного элемента и единственного выхода.

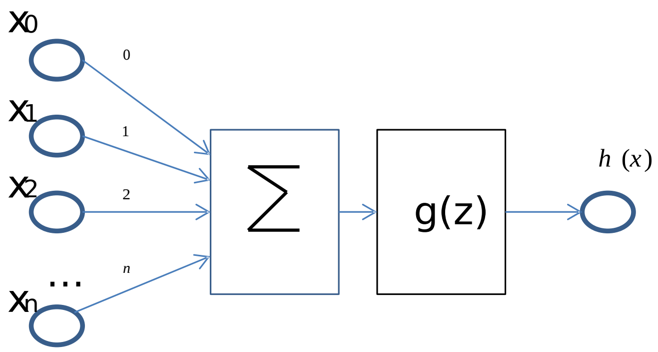

Схема классического нейрона

Выход нейрона определяется формулами:

$$z = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \ldots + \theta_n x_n,$$
$$h_\theta(x) = g(z),$$

где *g(z)* – сигмоидальная функция.

Выражение функции гипотезы классического нейрона идентично выражению функции гипотезы логистической регрессии.

Часто в качестве активационной функции применяется сигмоидальная функция, описанная ранее «Логистическая регрессия».

В последнее время в литературе веса θ нейронной сети чаще обозначают символом w, подчеркивая тем самым преемственность естественных нейронных сетей и искусственных нейронных сетей, где широко используется понятие синаптического коэффициента или веса (weight). Кроме того, такое обозначение показывает разницу между множеством параметров или весов (W) и гиперпараметрами модели. Гиперпараметры определяют общие свойства модели, и к ним относят коэффициент обучения, алгоритм оптимизации, число эпох обучения, количество скрытых слоев сети, количество нейронов в слоях и т.п.

Для упрощения схемы сумматор и активационный элемент объединяют, тогда многослойная сеть может выглядеть так, как показано на рисунке ниже. Сеть содержит четыре входных нейрона, четыре нейрона в скрытом слое и один выходной нейрон.
На рисунке входные нейроны обозначены символом x, нейроны скрытого слоя – символами $a_1^{[1]}$, $a_2^{[1]}$, $a_3^{[1]}$, $a_0^{[1]}$ и выходного слоя – символом $a_1^{[2]}$. Если нейронная сеть имеет несколько слоев, то первый слой называют входным, а последний – выходным. Все слои между ними

---

7 Cybenco G. Approximation by superpositions of a sigmoidal function // Mathematics of Control, Signals, and Systems. – 1989. – Vol. 4. – P. 304–314.

8 Hornik K. et al. Multilayer feedforward networks are universal approximators // Neural Networks. – 1989. – Vol. 2. – P. 359–366.

называются скрытыми. Для нейронной сети с L-слоями выход входного или нулевого слоя нейронов определяется выражением $a^{[0]} = x$.

На входе следующего или первого скрытого слоя имеем

$$z^{[1]} = w^{[1]} a^{(0)}.$$

Выход первого слоя:

$$a^{[1]} = g\ (z^{[1]}).$$

Для любого нейрона j, находящегося в скрытом слое i:

$$a_j^{[i]} = g\ (\quad w_j^{[i]} a^{[i-1]} + w_{j0}^{[i]} a_0^{[i]})\ . \quad \text{(Eq. 2.13)}$$

В этом выражении значение bias и его вес упомянуты отдельно как произведение $w_{j0}^{[i]} a_0^{[i]}$, где $w_j^{[i]}$ – вектор весов нейрона j.

Для выходного слоя:

$$a^{[L]} = h_w(x) = g\ (z^{[L]}). \qquad \text{(Eq. 2.14)}$$

Например, для сети на рисунке 3.8 выход каждого нейрона скрытого слоя можно рассчитать так же, как и для одиночного нейрона:

$$a_1^{[1]} = g\ (w_{10}^{[1]} x_0 + w_{11}^{[1]} x_1 + w_{12}^{[1]} x_2 + w_{13}^{[1]} x_3). \qquad \text{(Eq. 2.15)}$$
$$a_2^{[1]} = g\ (w_{20}^{[1]} x_0 + w_{21}^{[1]} x_1 + w_{22}^{[1]} x_2 + w_{23}^{[1]} x_3).$$
$$a_3^{[1]} = g\ (w_{30}^{[1]} x_0 + w_{31}^{[1]} x_1 + w_{32}^{[1]} x_2 + w_{33}^{[1]} x_3).$$

Выход нейронной сети определяется выражением:

$$a^{[2]} = h_\theta(x) = g\ (w_{10}^{[2]} a_0^{[1]} + w_{11}^{[2]} a_1^{[1]} + w_{12}^{[2]} a_2^{[1]} + w_{13}^{[2]} a_3^{[1]}). \qquad \text{(Eq. 2.16)}$$

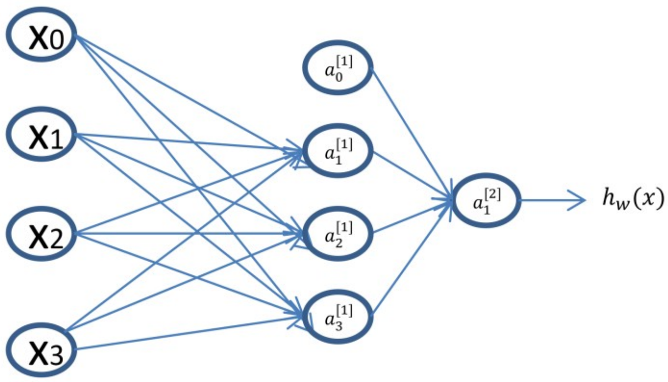

Схема многослойной сети с одним скрытым слоем

Для настройки весов w нейронной сети (обучения сети) используют функцию стоимости, напоминающую функцию стоимости для логистической регрессии (Eq. 2.12).

$$J(W) = \frac{-1}{m}, \qquad \text{(Eq. 2.17)}$$

где L – количество слоев нейронной сети; s_l – количество нейронов в слое l; K – количество классов (равно количеству нейронов в выходном слое); W – матрица весов.

Достоинством нейронной сети является возможность классификации с несколькими классами. В случае классификации объектов одного класса, то есть тогда, когда мы должны отделить условно «положительные» объекты от всех остальных, количество нейронов в выходном слое может быть равным и 1 (рисунок ниже). В этом случае принадлежность объекта к классу «положительных» определяется значением функции гипотезы, то есть если $h_\Theta(x^{(i)}) > 0.5$, то объект принадлежит к искомому классу. Однако чаще, в том числе с целью унификации, используется метод голосования («победитель забирает все»), когда сеть имеет в выходном слое 2 нейрона для двух классов объектов, три для трех и т. д.

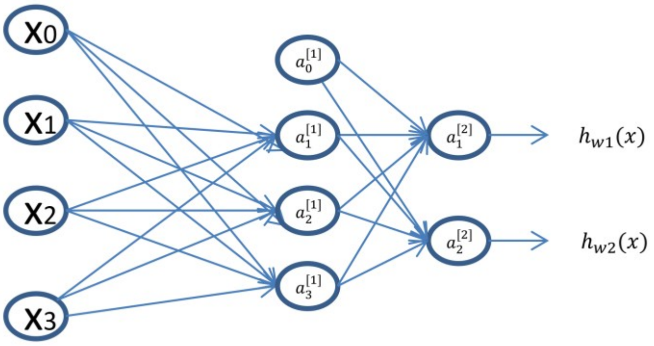

Схема многослойной сети с двумя выходами

Для обучения, то есть минимизации функции ошибки многослойной ИНС, используют алгоритм обратного распространения ошибки (Backpropagation of errors – BPE) [9] и его модификации, направленные на ускорение процесса обучения.

9 Werbos P. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. – Harvard University, 1974. – 38 p.