НИТУ «МИСиС»

Институт ИТКН

Кафедра инженерной кибернетики

Направление подготовки: 01.03.04 прикладная математика

КУРСОВАЯ РАБОТА

Предмет: **«Искусственные нейронные сети»**

Тема: **«Однослойный перцептрон. Теорема о сходимости перцептрона»**

Работу выполнил:

Учащийся: **Евстигнеев С.А.**

Группа: **БПМ-17-2**

Проверила: **Кондыбаева А. Б.**

Москва 2020

Содержание

1. Введение

Нейронные сети были разработаны с целью имитации нервной системы человека для решения задач машинного обучения на основе вычислительных элементов, работа которых напоминала бы действие человеческих нейронов. Концептуальной задачей, стоящей перед разработчиками нейронных сетей, является создание искусственного интеллекта (ИИ) на основе машин, архитектура которых была бы способна имитировать процессы, происходящие в нервной системе человека. Решить подобную задачу не так просто, поскольку вычислительная мощность даже самых быстрых современных компьютеров составляет лишь ничтожную долю возможностей человеческого мозга. Первые упоминания о нейронных сетях появились после появления компьютеров. Алгоритм перцептрона Розенблатта вызвало сенсацию относительно перспектив искусственного интеллекта. Но алгоритм требовал завышенные запросы в отношении необходимых объемов данных и вычислительных ресурсов нейронных сетей, что считалось основной помехой на пути их дальнейшего развития. В конечном счете на рубеже тысячелетия повышение доступности данных и увеличение вычислительных мощностей обеспечили дальнейший прогресс нейронных сетей, и эта область возродилась под названием "глубокое обучение".

2. Актуальность выбранной темы

Потенциальными областями применения искусственных нейронных сетей являются те, где человеческий интеллект малоэффективен, а традиционные вычисления трудоёмки или не отражают реальные физические процессы и объекты. Актуальность применения нейронных сетей многократно возрастает, когда появляется необходимость решения плохо формализованных задач. Основные области применения нейронных сетей: автоматизация процесса классификации, автоматизация прогнозирования, автоматизация процесса распознавания и т.д. Тем самым разобраться в том, как устроен перцептрон, являющийся основной составляющей любой нейронной сети, одна из главных задач в изучении области «Нейронные сети и глубокое обучение».

3. Цель работы

Изучить принципы работы однослойного перцептрона, а также теорему о его сходимости.

4. Однослойный перцептрон

Перцептрон представляет собой простейшую форму нейронной сети, предназначенную для классификации линейно-разделимых сигналов, у которых образы можно разделить некоторой гиперплоскостью. Перцептрон состоит из одного нейрона с настраиваемыми синаптическими весами и порогами. Первый алгоритм настройки свободных параметров для такой нейронной сети был создан Розенблаттом. Перцептрон строится для модели Мак-Каллоки-Питца, которая представляет нелинейный нейтрон. Такая нейронная модель состоит из линейного сумматора и ограничителя, реализованного в виде пороговой функции вычисления знака. Суммирующий узел этой нейронной модели вычисляет линейную комбинацию входных сигналов, поступающих на синапсы с учетом внешнего возмущения(порога). Полученная сумма (индуцированное локальное поле) передается на узел ограничителя. Таким образом, выход нейрона принимает значение +1, если сигнал на выходе сумматора положителен, и -1, если отрицателен.

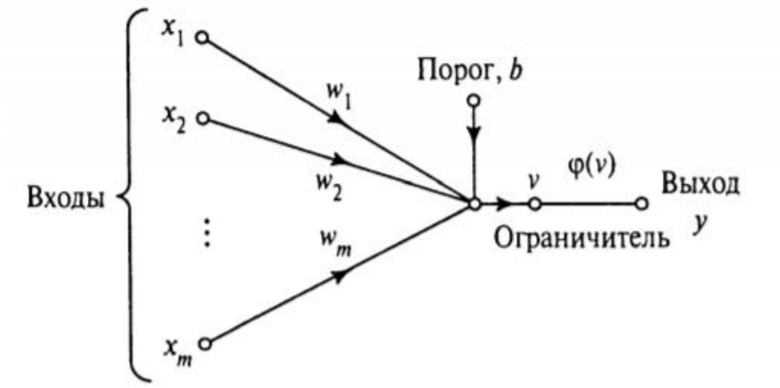

Рисунок 1 — Граф передачи сигнала для перцептрона

На Рисунке 1 синаптические веса перцептрона обозначены $w_1, w_2, \dots, w_m$, сигналы, поступающие на вход перцептрона, обозначены $x_1, x_2, \dots, x_m$, а пороговое значение – b. Согласно структуре модели, можно заключить, что входной сигнал ограничителя нейрона определяется выражением:

$$v = \sum_{i=1}^{m} w_i x_i + b$$

Целью перцептрона является корректное отнесение множества внешних стимулов $x_1, x_2, \dots, x_m$ к одному из двух классов: A или B. Решающее правило такой классификации заключается в том, что входной сигнал относится к классу A, если выход y равен $+1$, и к классу B, если равен -1.

В том случае, когда в качестве классификатора выступает перцептрон, имеется всего две области решения, разделенные гиперплоскостью, определяемые формулой:

$$\sum_{i=1}^{m} w_i x_i + b = 0$$

На Рисунке 2 проиллюстрирован пример, когда для переменных x_1 и x_2 разделяющая гиперплоскость вырождается в прямую. Пороговое значение определяет смещение разделяющей поверхности по отношению к началу координат.

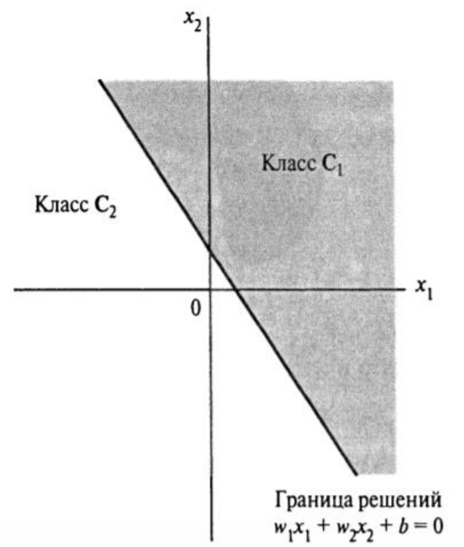

Рисунок 2 — Разделяющая поверхность в виде гиперплоскости для двумерной задачи классификации образов на две задачи

Синаптические веса перцептрона $w_1, w_2, \dots, w_m$ можно адаптировать итеративным методом. В частности, для настройки весовых коэффициентов можно использовать алгоритм сходимости перцептрона.

5. Теорема о сходимости перцептрона

5.1.Необходимые обозначения

Введем необходимые обозначения и описания к элементам перцептрона, которые потребуются для алгоритма обучения перцептрона, основанного на коррекции ошибок. На Рисунке 3 порог $b(n)$ рассматривается как синаптический вес связи с фиксированным входным сигналом +1. Для входного вектора размерности (m+1) запишем как:

$$x(n) = [+1, x_1(n), x_2(n), \dots, x_m(n)]^T,$$

Где n – номер итерации алгоритма. Аналогично можно определить (m+1)-мерный вектор весовых коэффициентов:

$$w(n) = \ [b(n), w_1(n), w_2(n), \dots, w_m(n)]^T$$

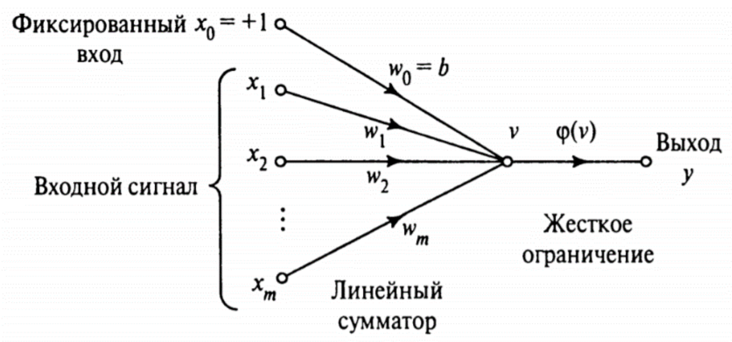

Рисунок 3 — Граф передачи сигнала для перцептрона

Таким образом, выход линейного сумматора можно записать в более компактной форме

$$v(n) = \sum_{i=0}^{m} w_i(n)x_i(n) = \ w^T(n)x(n),$$

Где $w_0(n)$ – пороговое значение b(n). При фиксированном значении n уравнение $w^T x \ = \ 0$ в m-мерном пространстве с координатами $x_1, x_2, \dots, x_m$ определяет гиперплоскость, которая является поверхностью решений для двух различных классов входных сигналов.

Для того, чтобы перцептрон работал корректно, два класса A и B должны быть линейно-разделимы. Другими словами, для правильной классификации образы должны быть значительно отдалены друг от друга, чтобы поверхность решений могла представлять собой гиперплоскость.

5.2.Формулировка теоремы

Теорема сходимости для алгоритма обучения перцептрона с фиксированным приращением для перцептрона сформулирована так: «Пусть подмножество векторов обучения X_1 и X_2 линейно-разделимы и входные сигналы поступают перцептрону только из этих подмножеств. Тогда алгоритм обучения перцептрона сходится после некоторого числа n_0 итераций в том смысле, что

$$w(n_0) = w(n_0 + 1) = w(n_0 + 2) = \cdots$$

является вектором решения для$n_0 \leq n_{max}$»

5.3. Абсолютная процедура адаптации однослойного перцептрона на основе коррекции ошибок

Рассмотрим абсолютную процедуру адаптации однослойного перцептрона на основе коррекции ошибок, в которой $\eta(n) -$ переменная величина. Пусть $\eta(n)$ – наименьшее целое число, для которого выполняется соотношение:

$$\eta(n)x^T(n)x(n) > |w^T(n)x(n)|$$

Согласно этой процедуре, если скалярное произведение $w^T(n)x(n)$ на шаге n имеет неверный знак, то $w^T(n+1)x(n)$ на итерации n+1 будет иметь правильный знак. Предполагается, что если знак произведения $w^T(n)x(n)$ некорректен, то можно изменить последовательность обучения для итерации n+1, приняв $x(n+1) = x(n)$. Другими словами, каждый из образов представляется перцептрону до тех пор, пока он не будет классифицирован корректно.

В Таблице 1 представлен алгоритм сходимости перцептрона. Символ $sgn(\cdot)$ на третьем шаге алгоритма означает функцию вычисления знака (сигнум).

Таблица 1. Алгоритм сходимости перцептрона

<u>Переменные и параметры</u> $x(n) = [+1, x_1(n), x_2(n), \dots, x_m(n)]^T$ – вектор-строка размерности m+1; $w(n) = [b(n), w_1(n), w_2(n), \dots, w_m(n)]^T$ – вектор-строка размерности m+1; b(n) – порог; y(n) – фактический отклик; d(n) – желаемый отклик; $0 < \boldsymbol{\eta} \leq \mathbf{1} -$ параметр скорости обучения.
1. <u>Инициализация</u> Пусть w(0) = 0. Последующие вычисления выполняются для шагов n = 1,2, …
2. <u>Активация</u> На шаге n активируем перцептрон, используя вектор x(n) с вещественными компонентами и желаемый отклик d(n)
3. <u>Вычисление фактического ответа</u> Вычисляем фактический отклик перцептрона: $y(n) = sgn(w^T(n)x(n))$
4. <u>Адаптация вектора весов</u> Изменяем вектор весов перцептрона: $w(n+1) = w(n) + \eta[d(n) - y(n)]x(n),$ где $d(n) = \begin{Bmatrix} +1, \text{если } x(n) \in A \\ -1, \text{если } x(n) \in B \end{Bmatrix}$
5. <u>Продолжение</u> Увеличиваем номер итерации n на единицу и возвращаемся к п. 2 алгоритма.

Алгоритм адаптации вектора весовых коэффициентов w(n) соответствует правилу обучения на основе коррекции ошибок:

$$w(n+1) = w(n) + \eta[d(n) - y(n)]x(n),$$

Где $\boldsymbol{\eta}$ – параметр скорости обучения, а разность $d(n) - y(n)$ сигнал ошибки. Параметр скорости является положительной константой, принадлежащей интервалу $0 < \boldsymbol{\eta} \leq 1$. Выбирая значение параметра скорости обучения из этого диапазона, следует учитывать два взаимоисключающий требования:

1) Усреднение предыдущих входных сигналов, обеспечивающее устойчивость оценки вектора весов, требует малых значений η.

2) Быстрая адаптация к реальным изменениям распределения процесса, отвечающего за формирование векторов сходного сигнала x, требует больших значений η.

Список литературы

1. Ф. Розенблатт, «Перцептрон: вероятностная модель для хранения и организации информации в мозге» *Психологический обзор, 1958*
2. Николенко С. И. – «Глубокое обучение», Библиотека программиста, 2018