ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«Поволжский государственный университет

телекоммуникаций и информатики»

 Кафедра ПОУТС

Практическая работа

по дисциплине «Математическое программирование»

Выполнил: студент 3 курса,

гр. ИВТ-93 Балабин Валерий

Проверил: Вержаковская М.А.

Самара, 2022

**Теория**

Персептрон – Модель искусственной нейронной сети, была предложена Фрэнком Розенблаттом в 1957 году

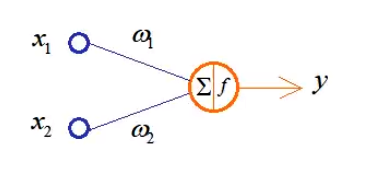
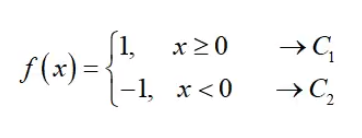


Рисунок 1 – Простейший персептрон

Такой персептрон способен решать задачу разделения сущностей с характеристиками x1, x2 на два класса С1 и С2

В основе работы нейрона нашей сети лежит активационная функция вида

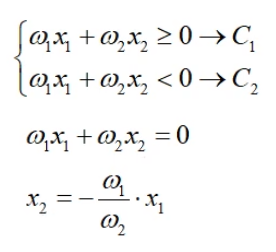


На вход в нее подаются 2 аргумента, которые умножаются на коэффициенты ω, называющиеся веса. Если функция будет давать

значение 1 то мы сделаем вывод, что этот входной сигнал принадлежит

классу С1, ну а если активационная функция будет давать -1, то значит этот

сигнал принадлежит классу С2.



Выразив x2 через x1, мы получим уравнение вида y = kx, т.е. уравнение прямой проходящей через центр координат. Она называется разделяющей прямой

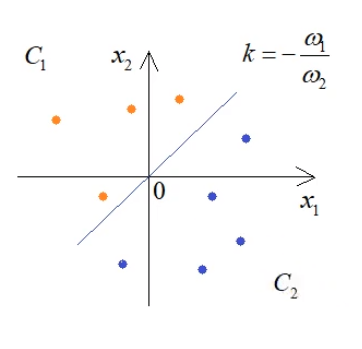


Рисунок 2 – координатная плоскость, случай 1

Однако, что если наши элементы невозможно классифицировать прямой, проходящей только через центр координат? Например, как в случае, показанном на рисунке 3.

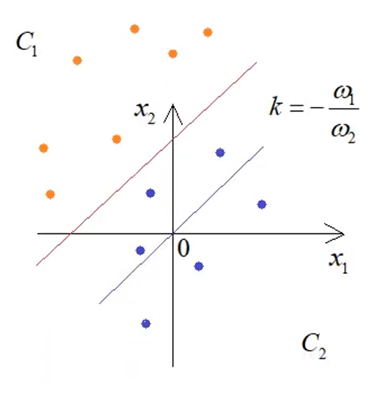


Рисунок 3 – координатная плоскость, случай 2

Как бы мы не проводили прямую, у нас не получится корректно классифицировать элементы. В этом случае нам нужен еще один аргумент, который называется bias или вход смещения.

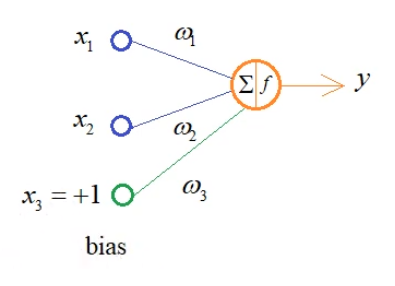
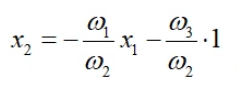


Рисунок 4 - bias



Выразив x2 через x1, мы получим уравнение вида y = kx + b, а значит наша разделяющая прямая больше не привязана к центру координат.

Но что если наш случай еще сложнее и элементы распределены, например, как на рисунке 5?

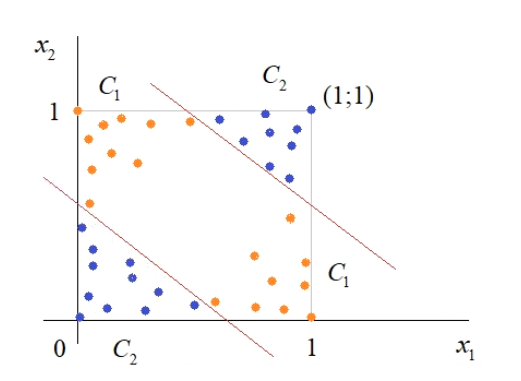


Рисунок 5 – координатная плоскость, случай 3

Решение тоже есть, нам нужно использовать 2 нейрона

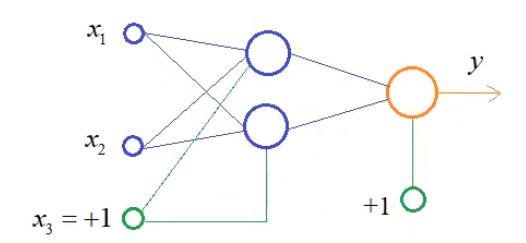


Рисунок 6 – персептрон из двух нейронов

Добавляя новые нейроны, мы сможем получать все более сложные формы

разделяющих областей, которые будут получены на основе комбинации

разделяющих прямых или гиперплоскостей, если речь идет о многослойных нейронных сетях.

Теперь, когда мы разобрали основы, поговорим о персептроне более предметно, он состоит из:

* входного вектора (входные данные),
* выходного вектора (выходные данные),
* вектора(ов) промежуточного представления (скрытый(ые) слои).

Вычисления в такой модели распространяются от входа к выходу. Связям между нейронами на разных слоях соответствуют веса.

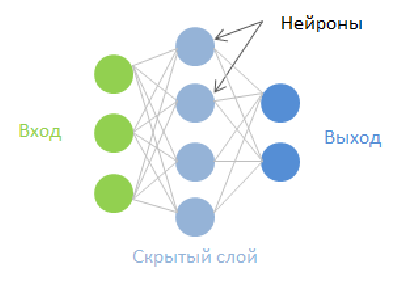
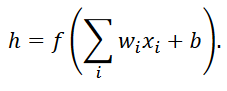


Рисунок 7 – многослойный персептрон

В нейрон поступают входные значения, 𝑥1, 𝑥2, 𝑥3 (элементы входного вектора) с соответствующими весами ω1, ω2, ω3. Далее, внутри нейрона происходит вычисление двух операций, а именно, композиции линейной и нелинейной функции:

* сначала мы вычисляем взвешенную сумму входных значений и добавляем некоторый параметр смещения 𝑏,
* далее, от полученных на предыдущем шаге значений, берём нелинейную функцию 𝑓, функцию активации.

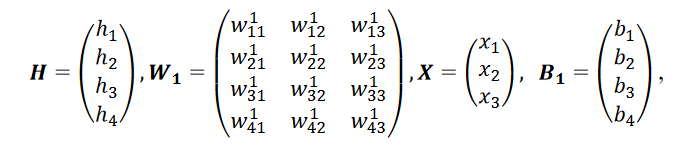
Таким образом, выходное значение ℎ из одного нейрона вычисляется по формуле



Обратим внимание на то, что одному слою в нейронной сети соответствует уже целая матрица весов 𝑾1 и некоторый вектор смещений 𝑩. Поэтому вычисления в одном слое персептрона можно представить в виде композиции матричного умножения, прибавления вектора смещений и поэлементного взятия нелинейной функции



Где



В многослойный персептроне, где выход каждого предыдущего слоя является входными данными для последующего слоя, вычисления можно представить в виде следующего рекуррентного соотношения



Где 𝑮𝒘 – функция, зависящая от параметров модели (от весов, связывающих нейроны и смещений). Чем больше слоев в нашей нейронной сети, тем более сложное промежуточное представление она имеет и тем более сложные зависимости между входными и выходными данными она способна описать.

Основное свойство функции активации – ее нелинейность. Если бы мы использовали линейную функцию, то, во-первых, мы могли бы решать только узкий класс задач, где зависимость между входными и выходными данными описывается линейной функцией, а, во-вторых, увеличение числа скрытых слоев не повышало бы эффективность нашей модели, поскольку композиция линейных функций – это все еще линейная функция. Задача функции активации – помочь принять локальное решение в каждом из нейронов. В практической части мы используем функцию типа sigmoid, она отображает значения на выходе из нейрона во что-то большее или меньшее нуля.

Нейронные сети обучаются с помощью, так называемого обучения с учителем. Для этого необходима обучающая выборка – размеченные данные, которые состоят из пар «входной объект – выходной объект». Мы подаём эту обучающую выборку в процесс обучения, который состоит в том, чтобы найти такие параметры (веса) модели 𝑾, чтобы наша нейронная сеть предсказывала правильно те ответы, которые мы уже знаем. Таким

образом, процесс обучения сводится к решению задачи минимизации или оптимизации функции ошибки на тех примерах, что есть в нашей обучающей выборке. Функцию ошибки можно записывать по-разному. Например,



где (𝑫 − 𝑮𝒘(𝒁)) есть разница между правильным ответом 𝑫 и предсказанием сети 𝑮𝒘(𝒁). От этой разницы мы берем норму, которую хотим минимизировать по всей обучающей выборке. Затем мы ищем такие параметры модели 𝑾∗, которые минимизируют эту ошибку. Изначально веса модели инициализируются случайными значениями

**Практика**

**Список используемых источников**

1. Галушкин, А.И. Нейронные сети: история развития теории: Учебное пособие для вузов. / А.И. Галушкин, Я.З. Цыпкин. - М.: Альянс, 2015. - 840 c.

2. Галушкин, А.И. Нейронные сети: основы теории. / А.И. Галушкин. - М.: РиС, 2015. - 496 c.

3. Каллан, Р. Нейронные сети: Краткий справочник / Р. Каллан. - М.: Вильямс И.Д., 2017. - 288 c.

4. Редько, В.Г. Эволюция, нейронные сети, интеллект: Модели и концепции эволюционной кибернетики / В.Г. Редько. - М.: Ленанд, 2015. - 224 c.

5. Хайкин, С. Нейронные сети: полный курс / С. Хайкин. - М.: Диалектика, 2019. - 1104 c.