Аппроксимация многомерной функции перцептроном с двумя скрытыми слоями.

Баталов Семен

02.03.2021

1. Перцептрон

Для аппроксимации любой непрерывной вектор-функции достаточно использовать перцептрон с двумя скрытыми слоями. Это следует из теоремы Колмогорова.

При решении поставленной задачи использовался перцептрон с 2000 нейронами в первом и 40 нейронами во втором внутренних слоях (количесво нейронов во входном и выходном слоях варьировалось в зависимости от вида аппроксимируемой функции). В качестве активационной функции использовалась функция «**RELU**». В качестве функции ошибки использовалась «**MAE**» (Mean Absolute Error). Для обучения был выбран алгоритм «**ADAM**» (Усовершентсвованный стохастический градиентный спуск). Количество эпох при обучении составило 60.

2. Обучение и оценка

Сначала производилась подготовка данных для обучения. Она заключалась в создании в области определения функции сетки, размерность которой совпадала с размерностью области определения, далее в узлах сетки расчитывались значения этой функции. Таким образом, получался датасет для обучения, то есть координаты узлов – входные данные, значения функции в них – выходные.

Далее из датасета случайно извлекалась обучающая и тестировочная выборки. Оценкой качества работы перцептрона был результат, возвращаемый функцией «**MAE**», примененной к тестировочной выборке. Для наглядности были также построены графики (и их приближения) некоторых одномерных функций.

3. Инструменты

Языком разработки был «**Python**». Для работы с нейросетью использовалась подбиблиотека «**Keras**» библиотеки «**TensorFlow**». Также для разделения датасета на обучающую и тестировочную выборки использовалась библиотека «**Sklearn**». Подробнее о программе можно узнать в папке «**source**» проекта.

4. Эксперименты и результаты

Здесь приведено описание некоторых функций, на которых проверялась корректность работы программы. Также изображены графики некоторых из них и приближенно вычисленные значения.

4.1. $f(x) = -x^2 + 5$

Непрерывная одномерная функция. Область аппроксимации: $[-1, 1]$; шаг сетки: 0.01; размер обучающей выборки: 0.6 $-$ 0.8.

Величина ошибки (MAE): 0.009 - 0.085.

4.2. $f(x) = sin(x) + 1$

Непрерывная одномерная функция. Область аппроксимации: $[0, 6.5]$; шаг сетки: 0.01; размер обучающей выборки: 0.6 $-$ 0.8.

Величина ошибки (MAE): 0.021 - 0.100.

4.3. $f(x) = \sqrt{x}$

Непрерывная одномерная функция. Область аппроксимации: $[0, 6.5]$; шаг сетки: 0.01; размер обучающей выборки: 0.6 $-$ 0.8.

Величина ошибки (MAE): 0.006 - 0.015.

4.4. $f(x, y) = x + y$

Непрерывная функция двух переменных. Область аппроксимации: $[0, 2] \times [0, 2]$; шаг сетки: 0.1; размер обучающей выборки: 0.6 $-$ 0.8.

Величина ошибки (MAE): 0.007 - 0.05.

4.5. $f(x, y) = (x + y, x^2 \cdot y)$

Непрерывная функция двух переменных. Область аппроксимации: $[0, 2] \times [0, 2]$; шаг сетки: 0.1; размер обучающей выборки: 0.6 $-$ 0.8.

Величина ошибки (MAE): 0.029 - 0.036.

5. Выводы

На рисунках ниже представлены результаты работы программы для различных функций. Можно заметить, что перцептрон приближает функцию кусочно-линейно, это отчетливо видно на (Рис. 5.1). При разных запусках нейросеть обучается по-разному (это связано с попаданием в разные локальные минимумы при градиентном спуске), поэтому график ошибки каждый раз меняется (функция ошибки возвращает модуль разности значений оригинальной функции и полученной при обучении перцептрона).

В целом удалось решить задачу аппроксимации функции с приемлемой точностью, однако стоит отметить, что точность в данной задаче зависит от количества нейронов в скрытых слоях перцептрона и способа его обучения.

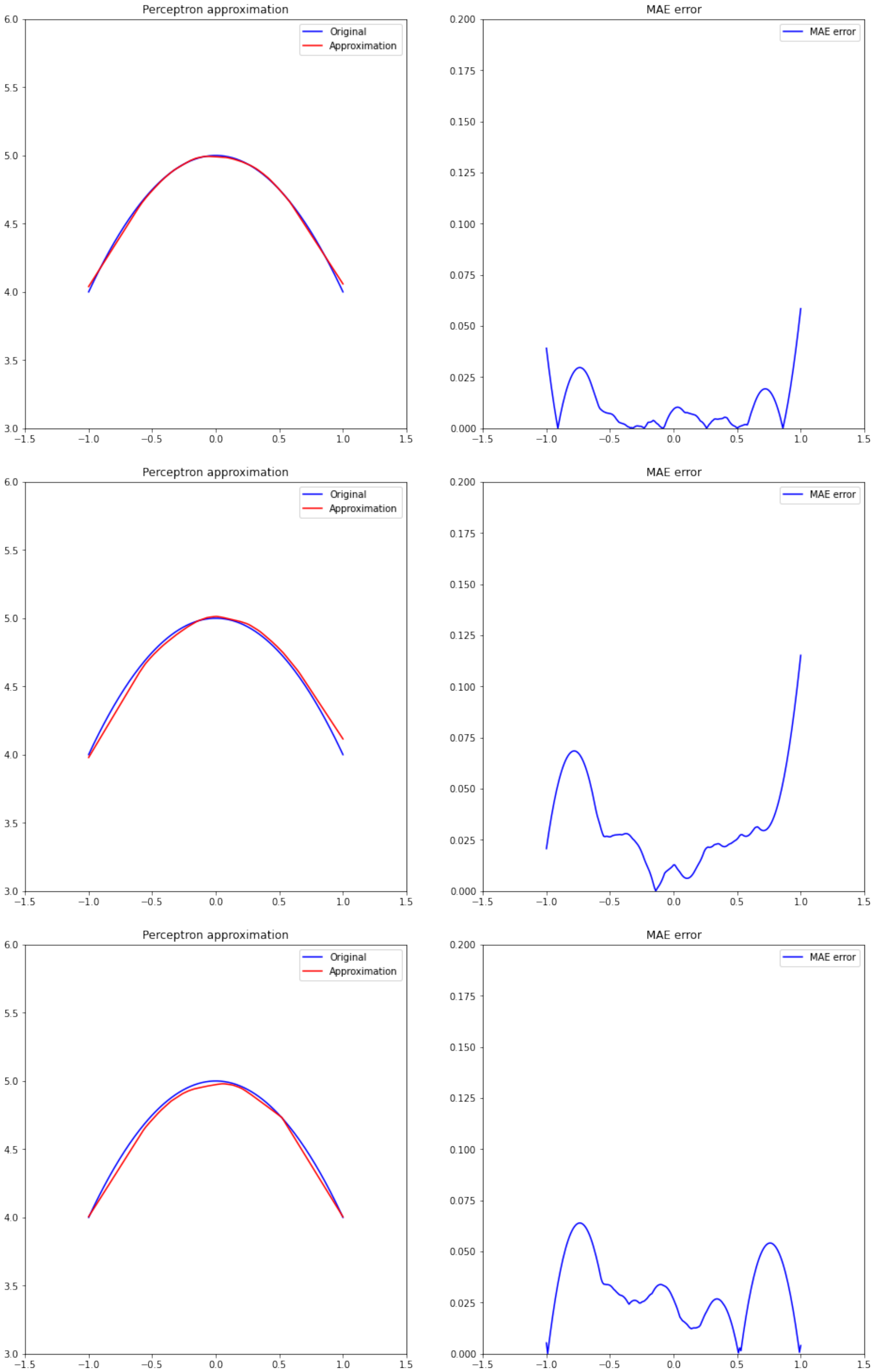

Рис. 1. Аппроксимация функции $f(x) = -x^2 + 5$.

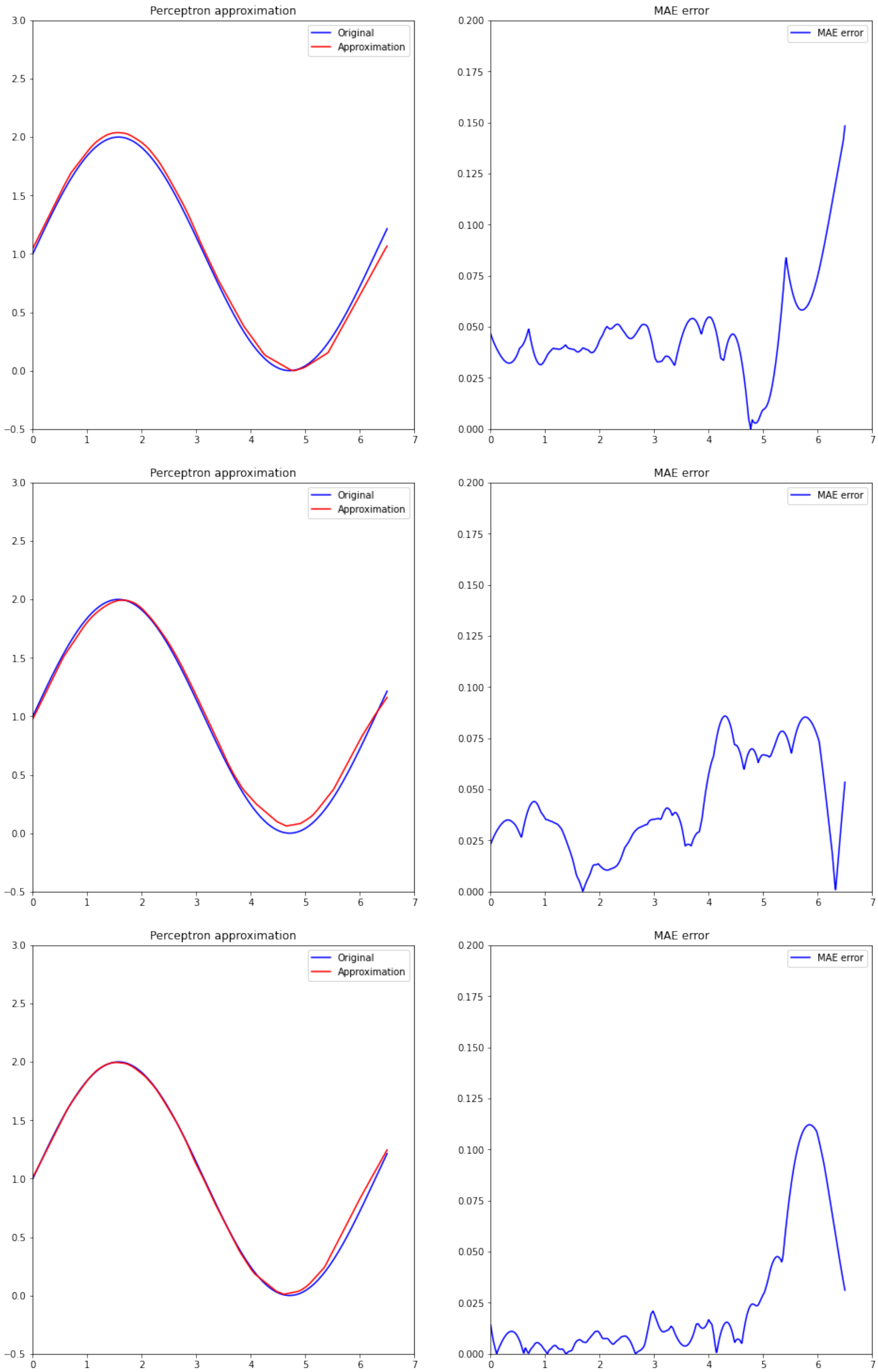

Рис. 2. Аппроксимация функции $f(x) = sin(x) + 1$.

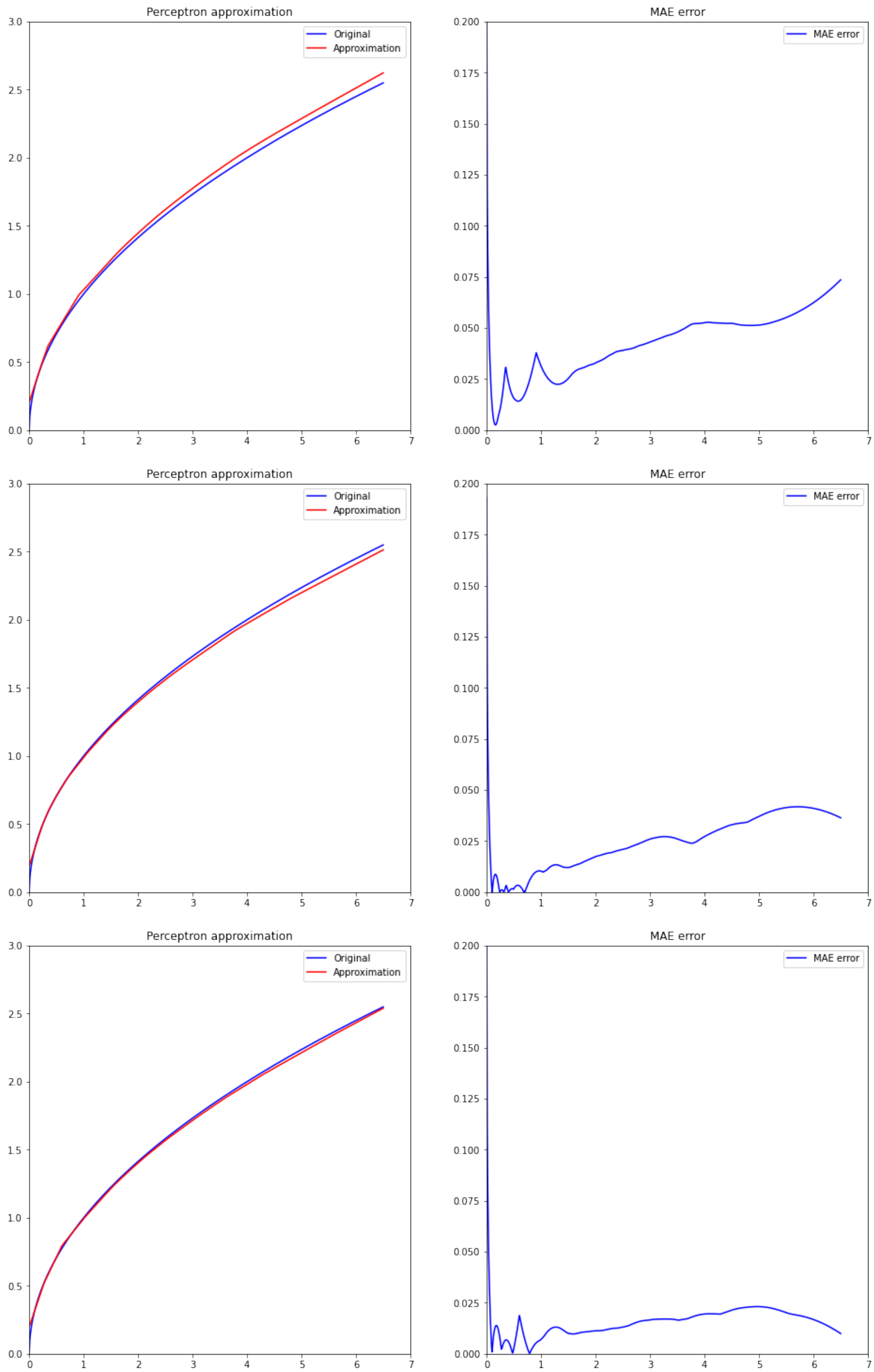

Рис. 3. Аппроксимация функции $f(x) = \sqrt{x}$.