Университет ИТМО, факультет ПИиКТ

Лабораторная работа №3

Дисциплина: Системы Искусственного Интеллекта

Выполнил: Чайка Алексей

Группа: P33113

Преподаватель: Болдырева Е.А.

г. Санкт-Петербург

2020 г.

Цель лабораторной работы

Learning neural networks: binary classification and CIFAR100 dataset classification with certain hyperparameters.

Задание

There are 2 parts of work in files named Lab1-Part1 and Lab1-Part2 respectively. Both parts represent work with collections of training and test data.

Part 1 represent recognition of basic math functions with illustrations of neural net vision of the functions for training.

Part 2 represent recognition of simple images for making following work.

Var	Part1 func	Part2 data	Hyperparameters
1	Absolute(Sin(x)) X: 6,3..6.3 Y: 0..1.2	CIFAR10	Layers count, neurons count per layer
2	Cos(x) X: -9..9 Y: -1..1	CIFAR100	Learn rate, regularization L1
3	Absolute(Sin(x)) X: 6,3..6.3 Y: 0..1.2	Handwritten digits	Regularization L2, output layer activation type
4	Cos(x) X: -9..9 Y: -1..1	Fashion articles	Layer activation type, loss function type

There are represented such hyperparameters as

- Layer count

- Neurons count per layer (actually it's not hyperparameter but structure parameter)

- Learn rate

- Regularization L1 and L2

- Output layer activation type

- Layer activation type

- Loss function type

- Epoch count

1) By changing these hyperparameters try to reach max accuracy value(at least 0.95) for Part2 model with fixed epoch count 20

2) Change 1st hyperparameter's value from min to max with minimal step depends on your variant

3) Show impact on result using graphs

4) Describe impact of each hyperparameter on accuracy.

5) Set hyperparameter value back to one which produced max accuracy

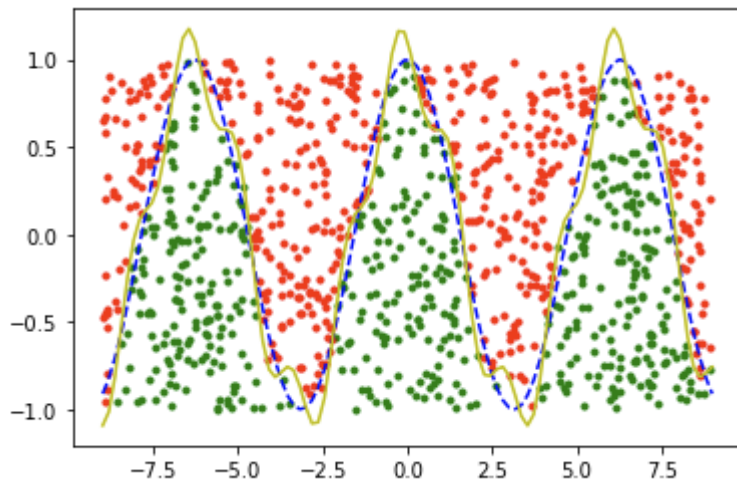6) Repeat 2-5 steps for second hyperparameter

Make a report including:

- Each hyperparameter description and its impact on accuracy.

- Hyperparameters' values which were used to reach accuracy value 0.95
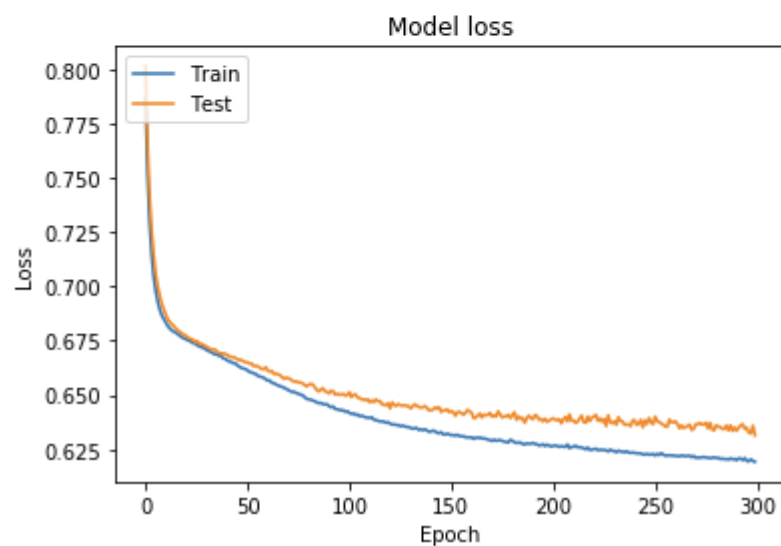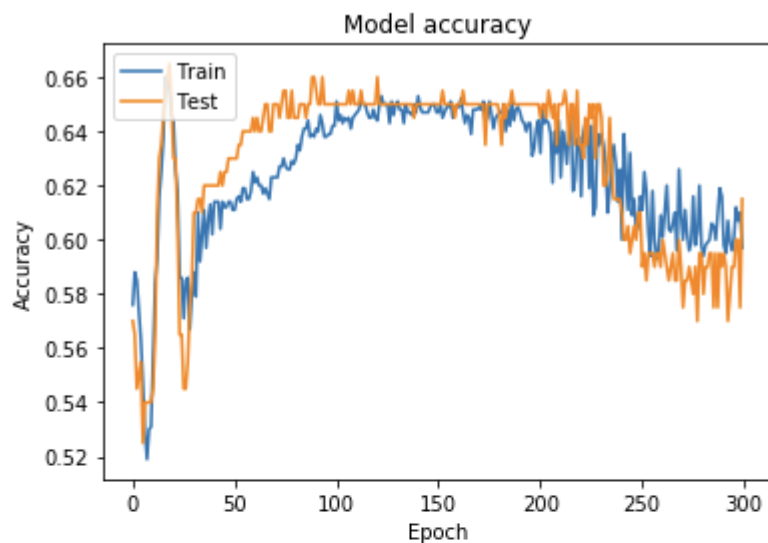- Graphs for these hyperparameters' values
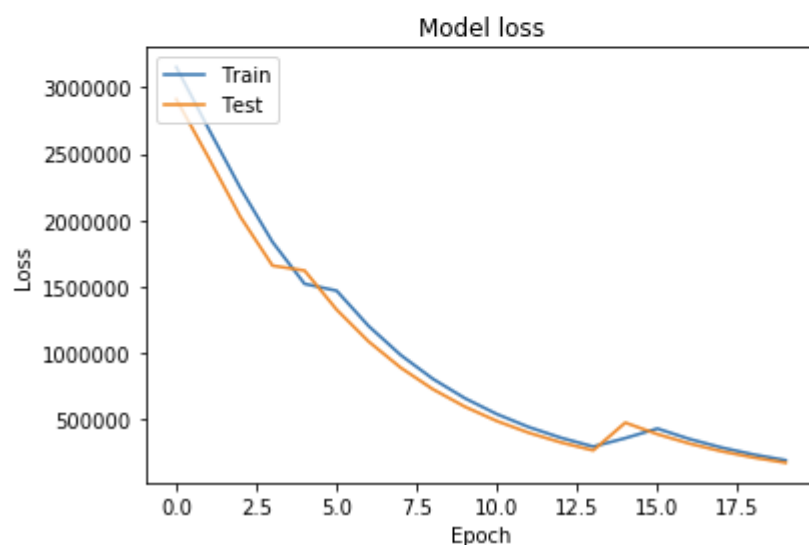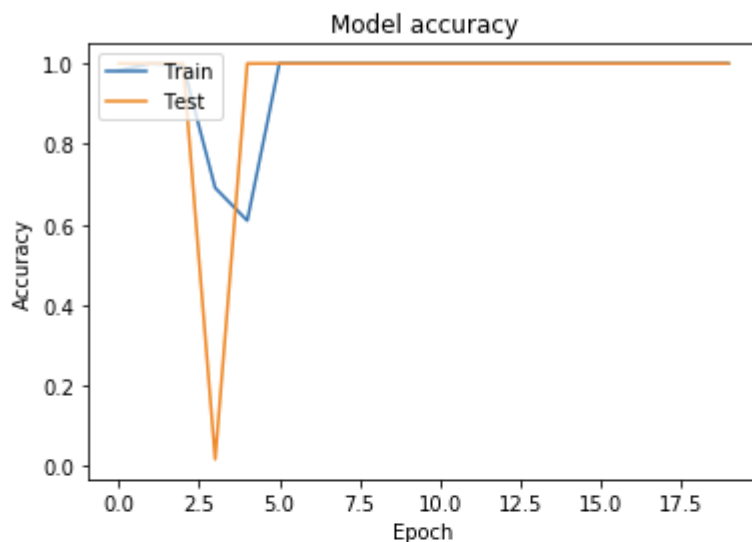
Порядок выполнения

Часть 1

Blue dash-line – real function, yellow line – noised function. You should be able to determine higher or lower new point is presented on the graph.

Accuracy: 0.6150000095367432

Часть 2

Here we have CIFAR100 small images classification dataset. The most valuable hyperparameter's values were:

- Learn rate = 1;
- Regularization L1 = 0.01;

Speaking of hyperparameters,

- The **learning rate** is a hyperparameter that controls how much to change the model in response to the estimated error each time the model weights are updated. Choosing the learning rate is challenging as a value too small may result in a long training process that could get stuck, whereas a value too large may result in learning a sub-optimal set of weights too fast or an unstable training process.

- **L1** (*also knows as Lasso regression*) adds "*absolute value of magnitude*" of coefficient as penalty term to the loss function. This variant pushes the parameters towards smaller values, even completely canceling the influence of some input features on the network output, which

implies an automatic **feature selection**. The result is a better generalization, but only to a certain extent

Вывод

During this laboratory, I have recalled several features working with Neural Network, as well as have trained two simple classificators.