

our models  
le one<sup>4</sup>—in  
al tricks for  
I show how  
e Dropout,  
ng stability



## CHAPTER 4

# Extensions

In the last chapter, after having spent two chapters reasoning from first principles about what deep learning models are and how they should work, we finally built our first deep learning model and trained it to solve the relatively simple problem of predicting house prices given numeric features about houses. On most real-world problems, however, successfully training deep learning models isn't so easy: while these models can conceivably find an optimal solution to any problem that can be framed as a supervised learning problem, in practice they often fail, and indeed there are few theoretical guarantees that a given model architecture will in fact find a good solution to a given problem. Still, there are some well-understood techniques that make neural network training more likely to succeed; these will be the focus of this chapter.

We'll start out by reviewing what neural networks are "trying to do" mathematically: find the minimum of a function. Then I'll show a series of techniques that can help the networks achieve this, demonstrating their effectiveness on the classic MNIST dataset of handwritten digits. We'll start with a loss function that is used throughout classification problems in deep learning, showing that it significantly accelerates learning (we've only covered regression problems thus far in this book because we hadn't yet introduced this loss function and thus haven't been able to do classification problems justice). On a similar note, we'll cover activation functions other than sigmoid and show why *they* might also accelerate learning, while discussing the trade-offs involved with activation functions in general. Next, we'll cover momentum, the most important (and straightforward) extension of the stochastic gradient descent optimization technique we've been using thus far, as well as briefly discussing what more advanced optimizers can do. We'll end by covering three techniques that are unrelated to each other but that are all essential: learning rate decay, weight initialization, and dropout. As we'll see, each of these techniques will help our neural network find successively more optimal solutions.

odel to fail



## **8. Time Restriction**

This product has no time limits in terms of its usage.

## **9. Limited Warranty**

Contractor makes no warranties or representations, express or implied, in fact or law, including the implied warranties of merchantability and fitness for a particular purpose. Contractor does not warrant that the operation of Software will be uninterrupted or error free. Customer agrees that all warranties are immediately null and void in the event of modification or alteration of Software and Source code.

## **10. Limitation of Liability**

Notwithstanding the form (e.g. contract, negligence, or otherwise) in which any legal or equitable action may be brought against Contractor, Contractor will not be liable for damages which exceed the amount of the license fees paid by the Customer for the Software. In no event will Contractor be liable for any damages, whether arising from tort or contract, including loss of data, lost profits, cost of cover, or other special, incidental, consequential, or indirect damages arising out of the use or inability to use the Software or accompanying documentation, however caused and on any theory of liability. This limitation will apply even if Contractor has been advised of the possibility of such damage. The Customer acknowledges that the license fee reflects this allocation of risk.

## **11. Miscellaneous**

This Agreement represents the complete agreement concerning this license between the parties and supersedes all prior agreements and representations between them. It may be amended only by a writing executed by both parties. This Agreement shall be governed by and construed under Russian Federation laws.

## **12. Contractor**

\_\_\_\_\_  
Technical support: \_\_\_\_\_

Web site: \_\_\_\_\_

## **13. Data details of the parties**

(this part of the Agreement is to be filled out by Customer and Contractor at the time of making contract)





MAIL TO:  
DMV CHANGE OF ADDRESS  
P. O. BOX 942859  
SACRAMENTO, CA 94238-0001

**A SEPARATE FORM IS NEEDED  
FOR EACH DRIVER OR VEHICLE OWNER**  
**NOTICE OF CHANGE OF ADDRESS**

DMV USE ONLY  
DL address  
updated by FO

24002

Please Print Characters in Capital Letters Using Black or Dark Blue Ink only.

1 2 3 4 5 6 7 8 9 0

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

**INSTRUCTIONS:**

- Enter the information as shown on the document, i.e. California driver license, ID card, or vehicle registration card, for which a change is being requested.
- Names not matching DMV records and/or unreadable information cannot be updated.
- Type or write your new address on a small piece of paper with your signature and date. Keep it with your driver license or ID card.
- A commercial licensed driver must maintain a California residence address or the driver license will be downgraded to non-commercial status.

**Personal Information**



LAST NAME  
**MALLICK**  
FIRST  
**SATYA**

INITIAL

BIRTH DATE  
**01 01 1990**

**D1234567**

**Voter Change of Address**

Use only with  
DL Change of Address

**New or Correct  
Residence  
Address**



Do Not Use P. O. Box  
in this space

**New or Correct  
Mailing Address**



If Different From  
Residence Address

**Vehicles,  
Vessels, or  
Placards  
Owned By You**

Use Additional  
Forms if Necessary

**Leased Vehicles**

**Location of  
Trailer Coach  
or Vessel**

If Different From  
Residence Address

**Old Address**

Your mailing address may be given to requesters providing a valid reason for requesting the information. If you receive mail at your residence, then giving DMV a separate mailing address is optional. Your residence address is restricted to authorized requesters per Vehicle Code Section 1808.21. I am the person whose name appears on the record(s) above and the mailing address shown is valid, existing and accurate. I consent to receive service of process at this mailing address pursuant to 415.20(b), 415.30, and 416.90 of the Civil Procedure Code. I certify (or declare) under penalty of perjury under the laws of the State of California that the foregoing is true and correct.

SIGNATURE  
**X**