

# **Practical AI for Healthcare Professionals**

**Machine Learning with Numpy,  
Scikit-learn, and TensorFlow**

**Abhinav Suri**

**Apress®**

# ***Practical AI for Healthcare Professionals: Machine Learning with Numpy, Scikit-learn, and TensorFlow***

Abhinav Suri  
San Antonio, TX, USA

ISBN-13 (pbk): 978-1-4842-7779-9  
<https://doi.org/10.1007/978-1-4842-7780-5>

ISBN-13 (electronic): 978-1-4842-7780-5

Copyright © 2022 by Abhinav Suri

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

Trademarked names, logos, and images may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, logo, or image we use the names, logos, and images only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Managing Director, Apress Media LLC: Welmoed Spahr  
Acquisitions Editor: Aaron Black  
Development Editor: James Markham  
Coordinating Editor: Jessica Vakili

Distributed to the book trade worldwide by Springer Science+Business Media New York, 233 Spring Street, 6th Floor, New York, NY 10013. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail [orders-ny@springer-sbm.com](mailto:orders-ny@springer-sbm.com), or visit [www.springeronline.com](http://www.springeronline.com). Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please e-mail [booktranslations@springernature.com](mailto:booktranslations@springernature.com); for reprint, paperback, or audio rights, please e-mail [bookpermissions@springernature.com](mailto:bookpermissions@springernature.com).

Apress titles may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Print and eBook Bulk Sales web page at <http://www.apress.com/bulk-sales>.

Any source code or other supplementary material referenced by the author in this book is available to readers on GitHub via the book's product page, located at [www.apress.com/978-1-4842-7779-9](http://www.apress.com/978-1-4842-7779-9). For more detailed information, please visit <http://www.apress.com/source-code>.

Printed on acid-free paper

*This book is dedicated to all the mentors and friends who have given me support throughout my education and beyond. A special thank you goes to my parents, whom I cannot thank enough for encouraging me and helping me down this path. Sic Itur Ad Astra.*

# Table of Contents

<b>About the Author .....</b>	<b>ix</b>
<b>About the Technical Reviewer .....</b>	<b>xi</b>
<b>Foreword to <i>Practical AI for Healthcare Professionals</i> .....</b>	<b>xiii</b>
<b>Chapter 1: Introduction to AI and Its Use Cases .....</b>	<b>1</b>
The Healthcare Information Paradox .....	2
AI, ML, Deep Learning, Big Data: What Do the Buzzwords Mean? .....	3
AI Considerations.....	11
Summary.....	14
The Rest of the Book.....	14
<b>Chapter 2: Computational Thinking .....</b>	<b>17</b>
How Computers “Think” .....	18
What “Can” and “Cannot” Be Solved .....	20
Algorithmic Alternatives.....	25
Stable Matching .....	26
Activity Selection.....	29
Analysis of Algorithms and Other Algorithms.....	34
Conclusion .....	40

TABLE OF CONTENTS

- Chapter 3: Overview of Programming .....41**
  - But First, What Are Programs? ..... 41
  - Getting Started with Python ..... 43
    - What Just Happened? ..... 45
    - Stepping It Up a Bit..... 45
  - Variables, Methods/Functions, String Operations, Print String Interpolation Applied ..... 47
  - Minor Improvements: If Statements..... 54
  - More Improvements: File Input and For Loops/Iteration ..... 59
  - File Output, Dictionaries, List Operations ..... 66
  - Cutting This Down with Pandas ..... 74
  - Summary..... 77
- Chapter 4: A Brief Tour of Machine Learning Algorithms .....79**
  - ML Algorithm Fundamentals ..... 79
  - Regression ..... 82
    - Linear Regression (for Classification Tasks)..... 83
    - Logistic Regression ..... 85
    - LASSO, Ridge, and Elastic Net for Regression, the Bias-Variance Trade-Off..... 89
  - Instance Learning ..... 93
    - k-Nearest Neighbors (and Scaling in ML)..... 93
    - Support Vector Machines..... 95
  - Decision Trees and Tree-Based Ensemble Algorithms ..... 98
    - Classification and Regression Trees ..... 100
    - Tree-Based Ensemble Methods: Bagging, Random Forest, and XGBoost.... 101
  - Clustering/Dimensionality Reduction ..... 103
    - k-Means Clustering ..... 104
    - Principal Component Analysis ..... 108

Artificial Neural Networks and Deep Learning .....	110
Fundamentals (Perceptron, Multilayer Perceptron) .....	110
Convolutional Neural Networks .....	116
Other Networks (RCNNs, LSTMs/RNNs, GANs) and Tasks (Image Segmentation, Key Point Detection, Image Generation) .....	120
Other Topics .....	121
Evaluation Metrics .....	121
k-Fold Cross Validation .....	124
Next Steps .....	125
<b>Chapter 5: Project #1 Machine Learning for Predicting Hospital Admission .....</b>	<b>127</b>
Data Processing and Cleaning .....	128
Installing + Importing Libraries .....	129
Reading in Data and Isolating Columns .....	130
Data Visualization .....	133
Cleaning Data .....	136
Dealing with Categorical Data/One-Hot Encoding .....	138
Starting the ML Pipeline .....	140
Training a Decision Tree Classifier .....	144
Grid Searching .....	146
Evaluation .....	147
Visualizing the Tree .....	151
This Seems Like a Lot to Do .....	152
Moving to PyCaret .....	152
Extra: Exporting/Loading a Model .....	161
Summary and What's Next .....	161

TABLE OF CONTENTS

- Chapter 6: Project #2 CNNs and Pneumonia Detection from Chest X-Rays .....163**
  - Project Setup ..... 163
    - Colab Setup ..... 164
    - Downloading Data ..... 165
    - Splitting Data ..... 168
    - Creating Data Generators and Augmenting Images..... 172
    - Your First Convolutional Neural Network: SmallNet ..... 182
    - Callbacks: TensorBoard, Early Stopping, Model Checkpointing, and Reduce Learning Rates..... 193
    - Defining the Fit Method and Fitting Smallnet..... 197
    - Your Second Convolutional Neural Network: Transfer Learning with VGG16 ..... 204
    - Visualizing Outputs with Grad-CAM..... 213
    - Evaluating Performance of SmallNet vs. VGG16..... 218
    - Evaluating on “External” Images ..... 223
  - Things to Improve ..... 225
  - Recap ..... 227
- Chapter 7: The Future of Healthcare and AI.....229**
  - Starting Your Own Projects ..... 230
  - Debugging..... 231
  - Considerations ..... 238
    - Patient Privacy..... 238
    - Algorithmic Bias ..... 239
    - Snake Oil + Creating Trust in the Real World..... 241
    - How to Talk About AI ..... 243
  - Wrap Up ..... 246
- Index.....247**

# About the Author

**Abhinav Suri** is a current medical student at the UCLA David Geffen School of Medicine. He completed his undergraduate degree at the University of Pennsylvania with majors in Computer Science and Biology. He also completed a master's degree in Public Health (MPH in Epidemiology) at Columbia University Mailman School of Public Health. Abhi has been dedicated to exploring the intersection between computer science and medicine. As an undergraduate, he carried out and directed research on deep learning algorithms for the automated detection of vertebral deformities and the detection of genetic factors that increase risk of COPD. His public health research focused on opioid usage trends in NY State and the development/utilization of geospatial dashboards for monitoring demographic disease trends in the COVID-19 pandemic.

Outside of classes and research, Abhi is an avid programmer and has made applications that address healthcare worker access in Tanzania, aid the discovery process for anti-wage theft cases, and facilitate access to arts classes in underfunded school districts. He also developed (and currently maintains) a popular open source repository, Flask Base, which has over 2,000 stars on GitHub. He also enjoys teaching (lectured a course on JavaScript) and writing. So far, his authored articles and videos have reached over 200,000 people across a variety of platforms.



# About the Technical Reviewer

**Vishwesh Ravi Shrimali** graduated in 2018 from BITS Pilani, where he studied mechanical engineering. Since then, he has worked with BigVision LLC on deep learning and computer vision and was involved in creating official OpenCV AI courses. Currently, he is working at Mercedes Benz Research and Development India Pvt. Ltd. He has a keen interest in programming and AI and has applied that interest in mechanical engineering projects. He has also written multiple blogs on OpenCV and deep learning on LearnOpenCV, a leading blog on computer vision. He has also coauthored *Machine Learning for OpenCV 4* (Second Edition) by Packt. When he is not writing blogs or working on projects, he likes to go on long walks or play his acoustic guitar.

# Foreword to *Practical AI for Healthcare Professionals*

Over the years ahead, artificial intelligence (AI) will play an ever-increasing and ultimately a transformative role for medicine's future. Nearly every week, we are seeing peer-reviewed studies that demonstrate the potential of deep neural networks for improving the accuracy of interpretation of medical images, from scans to slides to skin abnormalities to real-time machine vision pickup of colon polyps during endoscopy. Beyond medical images, algorithms are getting validated for patients, capturing their own data, coupled with algorithmic assistance, to facilitate the diagnosis of heart rhythm abnormalities, urinary tract infections, ear infections in children, and many other common reasons that would require a visit to a doctor. This early phase of medical AI will inevitably progress with validation via prospective and randomized clinical trials that are sorely lacking at this juncture. As Antonio de Leva wrote in *The Lancet*, "Machines will not replace physicians, but physicians using AI will soon replace those not using it."

But how will physicians get up to speed and learn about this field, which has undergone so much rapid change in the past decade owing to the subtype of AI known as deep learning (DL)? In this new book, Abhinav Suri, a medical student at UCLA, has provided an outstanding primer for uninitiated clinicians. Abhinav has the perfect background for this: a double degree in computer science and biology from Penn, an MPH degree from Columbia, and additional experience leading medical scan AI research at the Perelman School of Medicine. In just seven chapters, he succinctly lays out the basics and delineates the limits and potential flaws

## FOREWORD TO PRACTICAL AI FOR HEALTHCARE PROFESSIONALS

of AI, the different types of machine learning (ML) algorithms and deep neural networks, and “snake oil” AI. We’ve needed such a book for the medical community to get grounded, not so that physicians can code, but rather to understand the power, nuances, and limitations as AI makes its way deeper into the practice of medicine.

Undoubtedly, we will see more educational tools to help promote understanding and optimal use of AI in healthcare over the years ahead. The main textbook of the overall field is *Deep Learning* by Ian Goodfellow and colleagues, but it is quite comprehensive and well suited for people who intend to code and get deep into neural networks. Suri’s new book sets a very good standard for the goal of getting a quick and pragmatic introduction into AI, catering to the specific needs of clinicians. It will get you to “think like a computer” which is a requisite step to get grounded. Awareness of the basics and nuances of AI will eventually become a standard part of every medical school curriculum, and this primer may be considered a very good start in that direction.

—Eric J. Topol, MD  
Professor and EVP of Scripps Research  
Author, *Deep Medicine*  
La Jolla, California