Synthesis Lectures on
Engineering, Science, and Technology

SYNTHESIS
COLLECTION OF TECHNOLOGY

Pramod Gupta · Naresh Kumar Sehgal ·
John M. Acken

# Introduction to Machine Learning with Security

## Theory and Practice Using Python in the Cloud

*Second Edition*

Springer

## Synthesis Lectures on Engineering, Science, and Technology

The focus of this series is general topics, and applications about, and for, engineers and scientists on a wide array of applications, methods and advances. Most titles cover subjects such as professional development, education, and study skills, as well as basic introductory undergraduate material and other topics appropriate for a broader and less technical audience.

Pramod Gupta · Naresh Kumar Sehgal ·
John M. Acken

# Introduction to Machine Learning with Security

## Theory and Practice Using Python in the Cloud

Second Edition

Springer

Pramod Gupta
Brentwood, CA, USA

John M. Acken
Portland State University
Vancouver, WA, USA

Naresh Kumar Sehgal
NovaSignal Corp
Santa Clara, CA, USA

ISSN 2690-0300 ISSN 2690-0327 (electronic)
Synthesis Lectures on Engineering, Science, and Technology
ISBN 978-3-031-59169-3 ISBN 978-3-031-59170-9 (eBook)
https://doi.org/10.1007/978-3-031-59170-9



This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*Pramod dedicates this book to his late parents who always encouraged him to pursue his goals.*

*Naresh dedicates this book to his father, late Shri Pawan Kumar Sehgal, whose inspiration encouraged Naresh to continue his lifelong education.*

*John dedicates this book to his wife, Dr. Charlotte Acken, and his daughter, McKinsey Schenker.*

# Foreword to the Second Edition

In the recent past, artificial intelligence and machine learning has deeply impacted every industry: it is now used in virtually every industry, namely, web search, smart phone, speech recognition, medical and life sciences, self-driving cars to name a few.

Pramod et al. have written an excellent and comprehensive textbook enabling readers to understand and comprehend backgrounds to each of these domains. The book takes the students to comprehend and learn the concepts of data science and statistics, thereby setting up their own machine learning platform with open-source tools. The intention behind the book is to concentrate more on the usage and application of machine learning. The book covers a wide base of techniques, from the simplest and most commonly used algorithms to complex machine learning algorithms.

The authors of this book have leveraged their hands-on experience with solving real-world problems using Python and the Machine Learning ecosystem to help the readers gain solid knowledge needed to apply essential concepts, methodologies, tools, and techniques for solving their own real-world problems.

The book aims to cater to readers with varying skill levels ranging from beginners to experts and enable them in structuring and building practical Machine Learning and AI solutions.

This book is appropriate for both advanced undergraduate or master's students who want to work in this domain, or for individuals working in the area of machine learning. It is an excellent resource for those who wish to start learning data science and machine learning so as to understand and use these powerful techniques in their work area.

By the end of the book, readers will have the knowledge on the tools needed to begin their journey in the domain of machine learning and artificial intelligence.

The book will also support students to treat information securely as both AI and ML are intertwined due to data confidentiality, model integrity, and system availability.

Solutions such as encryption, hashing, and secure computations have been used for both AI and ML.

Readers will find their knowledge significantly enhanced after going through the book.

Mumbai, India

Dr. Praveer Sinha
CEO and MD
The Tata Power Company Limited

# Foreword to the First Edition

The invention of the cloud brought massive advancement in computer architecture. Seemingly overnight the delivery and consumption of services were transformed. Cloud computing has been an accelerant to the pace of technology adoption. In fact, it is cloud computing that has enabled artificial intelligence to move from the world of sci-fi to mainstream consumption. The cloud delivers the massive data storage and computes capacity needed for cost-effective analytics and insights.

As the promises of reduced time to deployment, lower cost of operation, and increased accessibility have been realized, adoption of the cloud has moved from novel to normal. And yet the theoretical understanding of the underlying technology remains locked in the brains of a relatively small group. The availability of cloud and AI curriculum in all educational settings needs to match the availability of cloud and AI services if the current pace of innovation is to continue.

For those looking to gain insight into the complex algorithms of artificial intelligence and the cloud architecture that fuels them, this book fills the void with highly valuable instruction.

Los Angeles, California, USA
January 2021

Diane Bryant
CEO and Chairman
https://www.NovaSignal.com

# Preface to the Second Edition

Since the publication of our book's first edition in 2021, the AI world has changed in many ways. It is akin to the Gold Rush of California in 1849, where prospectors came from all over the world to become rich. Some were lucky and found gold, but most of the riches went to the companies that were providing supplies to the prospectors. One of them, Levi Strauss, started by an eponymous immigrant from Bavaria still survives. He recognized a need among the hardworking miners for clothing to endure anything. This need was met with blue jeans. Similarly, in the modern-day gold rush of AI, a few companies are doing very well supplying GPUs, Cloud, and other tools to the folks currently doing datamining. It remains to be seen how many of these will last as long as Levi Strauss has endured.

We felt a need to update our book to reflect this new development in the AI world, including a not-so-broadly recognized gap with security of data and code. Hence, we included a new co-author, Dr. John M. Acken. He is the professor of hardware security at Portland State University in Oregon, before finishing his Ph.D. at Stanford and then working for over half-a-century at various organizations such as US Army, Sandia National Labs, and Intel. The new Security chapters and sections that John has added to our book fill an important gap in the current practices of AI in the real world. Pramod and Naresh have expanded their existing chapters to include new material related to AI algorithms, practices to remove noisy data, optimizations in the Cloud, and an introduction to the new products related to Large Language Models (LLMs). Special thanks to Shiva Kintali for introducing us to jailbreaking LLMs. A tutorial for implementing Transformers Architecture has been added in the Appendix.

We hope that our readers will find this updated edition useful in their continuing learning and professional practices. As always, please reach out to us for your suggestions for our future editions.

Brentwood, CA, USA — Pramod Gupta
Santa Clara, CA, USA — Naresh Kumar Sehgal
Portland, OR, USA — John M. Acken

# Preface to the First Edition

The idea for this book came from our lead co-author, Pramod Gupta, who has been teaching Data Sciences and related classes at University of California, Santa Cruz Extension, and most recently at University of California, Berkeley for several years. Prior to that his hands-on experience in the industry uniquely qualifies him to write the AI and ML parts of this book. Pramod had met with Naresh K. Sehgal decades ago during their undergraduate studies at Punjab Engineering College, Chandigarh in India. Since then, Naresh's career path took him through Chip design and later on a journey through Cloud Computing. The amalgamation of their respective work experiences has resulted in this book in your hands. It would not have been possible without the guidance and inspiration of Prof. PCP Bhatt, to whom the authors would like to dedicate this book.

It starts with an introduction to Machine Learning in Chap. 1 and laying down a deeper foundation of ML algorithms in the Chap. 2. Then Chap. 3 serves as a bridge to the Cloud, making a case for using the abundant compute and storage availability for training and inference purposes in the context of Deep Learning. Chapter 4 further explains some basic concepts of Cloud Computing with emphasis on the key characteristics that differentiate it from the enterprise computing, or even running AI algorithms on a laptop. Chapter 5 expands the usage of Cloud for Machine Learning by enumerating its data pipeline stages. Chapter 6 touches on a very important aspect of security in the Cloud for AI and ML algorithms as well as datasets. Chapter 7 delves into some practical aspects of running ML in Amazon's Cloud setup. Chapter 8 gives an example of using Cloud for health care-based AI and ML solutions. Lastly, in the Chap. 9, we look at efforts underway to speed up AI and ML using various hardware-based solutions. No engineering book can be complete without some practical problems and solutions. To meet that expectation, we present three real-life projects that Pramod's students had implemented using Python in Appendices A through C. For each of these, migration of these projects' code to a commercial Public Cloud is illustrated for the reader to practice. Appendix D has solutions to various Points to Ponder, which were posed at the end of each chapter. The motivation here is for the reader to think and then compare one's own answers with our proposed solutions. It can also be the basis of discussion in a classroom setting. The book wraps up with additional questions in Appendix E, the answers to which we leave for the readers to complete.

As with any major project, writing this book took months of planning and over a year to execute it. Even though only two co-authors are listed, there was a major contribution by Prof. PCP Bhatt, who met and reviewed our progress every week for over a year. In addition, we are thankful to our colleague, Aditya Srinivasan, who wrote two sections in Chap. 8 on Multi-Cloud Solutions and UCSD Antibiogram case study. We also used coding examples of several students from Pramod's classes in the Appendices. Naresh picked some ideas and sections from his earlier books with Dr. John M. Acken and Prof. Bhatt on Cloud Computing and Security. NovaGuide View application developed by Shiv Shankar has been used to illustrate NovaSignal's growing presence in the Cloud. Needless to say, several other resources and sites were used to learn and leverage educational material that has been duly acknowledged in our reference sections.

We sincerely hope that readers will have as much fun reading it as we had in writing the varied material in this book. We accept ownership for all the mistakes in this book, but please don't miss sending these to us for corrections, in addition to any suggestions for a future edition. If you include at least part of the sentence the error appears in, that makes it easy for us to search. Page and section numbers are fine, too. Thanks!

San Jose, CA, USA — Pramod Gupta
Santa Clara, CA, USA — Naresh Kumar Sehgal

# About This Book

## Objective

The purpose of this book is to introduce Machine Learning and Cloud Computing, both from a conceptual level and its Usages with underlying infrastructure. The focus areas of this book include Best Practices for using AI and ML in a Dynamic Infrastructure with Cloud Computing and high Security.

Target audiences are as follows:

1. Senior UG students who have studied programming languages, operating systems.
2. Senior UG and PG Students in software engineering or Information Technology Disciplines.
3. SW developers engaged in migrating in-house ML applications to Public Cloud with Security.
4. Information Technology managers for improving AI/ML performance in Cloud with Security.
5. Professionals who want to learn about the ML, Cloud, Security and technologies behind them.

Level of the book: Mostly at the senior UG or first semester of PG in software engineering, data science, Machine Learning or IT systems.

# Contents