

O'REILLY®

2nd Edition  
Updated for  
TensorFlow 2

# Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow

Concepts, Tools, and Techniques  
to Build Intelligent Systems

powered by



**Early  
Release**

**RAW &  
UNEDITED**

Aurélien Géron



SECOND EDITION

---

# Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow

*Concepts, Tools, and Techniques to  
Build Intelligent Systems*

*Aurélien Géron*

Beijing • Boston • Farnham • Sebastopol • Tokyo

**O'REILLY®**

# Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow

by Aurélien Geron

Copyright © 2019 Aurélien Geron. All rights reserved.

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (<http://oreilly.com>). For more information, contact our corporate/institutional sales department: 800-998-9938 or [corporate@oreilly.com](mailto:corporate@oreilly.com).

**Editor:** Nicole Tache

**Interior Designer:** David Futato

**Cover Designer:** Karen Montgomery

**Illustrator:** Rebecca Demarest

June 2019:

Second Edition

## Revision History for the Early Release

2018-11-05: First Release

2019-01-24: Second Release

2019-03-07: Third Release

2019-03-29: Fourth Release

2019-04-22: Fifth Release

See <http://oreilly.com/catalog/errata.csp?isbn=9781492032649> for release details.

The O'Reilly logo is a registered trademark of O'Reilly Media, Inc. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*, the cover image, and related trade dress are trademarks of O'Reilly Media, Inc.

While the publisher and the author have used good faith efforts to ensure that the information and instructions contained in this work are accurate, the publisher and the author disclaim all responsibility for errors or omissions, including without limitation responsibility for damages resulting from the use of or reliance on this work. Use of the information and instructions contained in this work is at your own risk. If any code samples or other technology this work contains or describes is subject to open source licenses or the intellectual property rights of others, it is your responsibility to ensure that your use thereof complies with such licenses and/or rights.

978-1-492-03264-9

[LSI]

---

# Table of Contents

Preface.....	xi
--------------	----

---

## Part I. The Fundamentals of Machine Learning

<b>1. The Machine Learning Landscape.....</b>	<b>3</b>
What Is Machine Learning?	4
Why Use Machine Learning?	4
Types of Machine Learning Systems	8
Supervised/Unsupervised Learning	8
Batch and Online Learning	15
Instance-Based Versus Model-Based Learning	18
Main Challenges of Machine Learning	24
Insufficient Quantity of Training Data	24
Nonrepresentative Training Data	26
Poor-Quality Data	27
Irrelevant Features	27
Overfitting the Training Data	28
Underfitting the Training Data	30
Stepping Back	30
Testing and Validating	31
Hyperparameter Tuning and Model Selection	32
Data Mismatch	33
Exercises	34
<b>2. End-to-End Machine Learning Project.....</b>	<b>37</b>
Working with Real Data	38
Look at the Big Picture	39

Frame the Problem	39
Select a Performance Measure	42
Check the Assumptions	45
Get the Data	45
Create the Workspace	45
Download the Data	49
Take a Quick Look at the Data Structure	50
Create a Test Set	54
Discover and Visualize the Data to Gain Insights	58
Visualizing Geographical Data	59
Looking for Correlations	62
Experimenting with Attribute Combinations	65
Prepare the Data for Machine Learning Algorithms	66
Data Cleaning	67
Handling Text and Categorical Attributes	69
Custom Transformers	71
Feature Scaling	72
Transformation Pipelines	73
Select and Train a Model	75
Training and Evaluating on the Training Set	75
Better Evaluation Using Cross-Validation	76
Fine-Tune Your Model	79
Grid Search	79
Randomized Search	81
Ensemble Methods	82
Analyze the Best Models and Their Errors	82
Evaluate Your System on the Test Set	83
Launch, Monitor, and Maintain Your System	84
Try It Out!	85
Exercises	85
<b>3. Classification.....</b>	<b>87</b>
MNIST	87
Training a Binary Classifier	90
Performance Measures	90
Measuring Accuracy Using Cross-Validation	91
Confusion Matrix	92
Precision and Recall	94
Precision/Recall Tradeoff	95
The ROC Curve	99
Multiclass Classification	102
Error Analysis	104

Multilabel Classification	108
Multioutput Classification	109
Exercises	110
<b>4. Training Models.....</b>	<b>113</b>
Linear Regression	114
The Normal Equation	116
Computational Complexity	119
Gradient Descent	119
Batch Gradient Descent	123
Stochastic Gradient Descent	126
Mini-batch Gradient Descent	129
Polynomial Regression	130
Learning Curves	132
Regularized Linear Models	136
Ridge Regression	137
Lasso Regression	139
Elastic Net	142
Early Stopping	142
Logistic Regression	144
Estimating Probabilities	144
Training and Cost Function	145
Decision Boundaries	146
Softmax Regression	149
Exercises	153
<b>5. Support Vector Machines.....</b>	<b>155</b>
Linear SVM Classification	155
Soft Margin Classification	156
Nonlinear SVM Classification	159
Polynomial Kernel	160
Adding Similarity Features	161
Gaussian RBF Kernel	162
Computational Complexity	163
SVM Regression	164
Under the Hood	166
Decision Function and Predictions	166
Training Objective	167
Quadratic Programming	169
The Dual Problem	170
Kernelized SVM	171
Online SVMs	174

Exercises	175
<b>6. Decision Trees</b>	<b>177</b>
Training and Visualizing a Decision Tree	177
Making Predictions	179
Estimating Class Probabilities	181
The CART Training Algorithm	182
Computational Complexity	183
Gini Impurity or Entropy?	183
Regularization Hyperparameters	184
Regression	185
Instability	188
Exercises	189
<b>7. Ensemble Learning and Random Forests</b>	<b>191</b>
Voting Classifiers	192
Bagging and Pasting	195
Bagging and Pasting in Scikit-Learn	196
Out-of-Bag Evaluation	197
Random Patches and Random Subspaces	198
Random Forests	199
Extra-Trees	200
Feature Importance	200
Boosting	201
AdaBoost	202
Gradient Boosting	205
Stacking	210
Exercises	213
<b>8. Dimensionality Reduction</b>	<b>215</b>
The Curse of Dimensionality	216
Main Approaches for Dimensionality Reduction	218
Projection	218
Manifold Learning	220
PCA	222
Preserving the Variance	222
Principal Components	223
Projecting Down to $d$ Dimensions	224
Using Scikit-Learn	224
Explained Variance Ratio	225
Choosing the Right Number of Dimensions	225
PCA for Compression	226



Randomized PCA	227
Incremental PCA	227
Kernel PCA	228
Selecting a Kernel and Tuning Hyperparameters	229
LLE	232
Other Dimensionality Reduction Techniques	234
Exercises	235
<b>9. Unsupervised Learning Techniques.....</b>	<b>237</b>
Clustering	238
K-Means	240
Limits of K-Means	250
Using clustering for image segmentation	251
Using Clustering for Preprocessing	252
Using Clustering for Semi-Supervised Learning	254
DBSCAN	256
Other Clustering Algorithms	259
Gaussian Mixtures	260
Anomaly Detection using Gaussian Mixtures	266
Selecting the Number of Clusters	267
Bayesian Gaussian Mixture Models	270
Other Anomaly Detection and Novelty Detection Algorithms	274
<hr/>	
<b>Part II. Neural Networks and Deep Learning</b>	
<b>10. Introduction to Artificial Neural Networks with Keras.....</b>	<b>277</b>
From Biological to Artificial Neurons	278
Biological Neurons	279
Logical Computations with Neurons	281
The Perceptron	281
Multi-Layer Perceptron and Backpropagation	286
Regression MLPs	289
Classification MLPs	290
Implementing MLPs with Keras	292
Installing TensorFlow 2	293
Building an Image Classifier Using the Sequential API	294
Building a Regression MLP Using the Sequential API	303
Building Complex Models Using the Functional API	304
Building Dynamic Models Using the Subclassing API	309
Saving and Restoring a Model	311
Using Callbacks	311

Visualization Using TensorBoard	313
Fine-Tuning Neural Network Hyperparameters	315
Number of Hidden Layers	319
Number of Neurons per Hidden Layer	320
Learning Rate, Batch Size and Other Hyperparameters	320
Exercises	322
<b>11. Training Deep Neural Networks. ....</b>	<b>325</b>
Vanishing/Exploding Gradients Problems	326
Glorot and He Initialization	327
Nonsaturating Activation Functions	329
Batch Normalization	333
Gradient Clipping	338
Reusing Pretrained Layers	339
Transfer Learning With Keras	341
Unsupervised Pretraining	343
Pretraining on an Auxiliary Task	344
Faster Optimizers	344
Momentum Optimization	345
Nesterov Accelerated Gradient	346
AdaGrad	347
RMSProp	349
Adam and Nadam Optimization	349
Learning Rate Scheduling	352
Avoiding Overfitting Through Regularization	356
$\ell_1$ and $\ell_2$ Regularization	356
Dropout	357
Monte-Carlo (MC) Dropout	360
Max-Norm Regularization	362
Summary and Practical Guidelines	363
Exercises	364
<b>12. Custom Models and Training with TensorFlow. ....</b>	<b>367</b>
A Quick Tour of TensorFlow	368
Using TensorFlow like NumPy	371
Tensors and Operations	371
Tensors and NumPy	373
Type Conversions	374
Variables	374
Other Data Structures	375
Customizing Models and Training Algorithms	376
Custom Loss Functions	376