

O'REILLY®

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

Concepts, Tools, and Techniques  
to Build Intelligent Systems

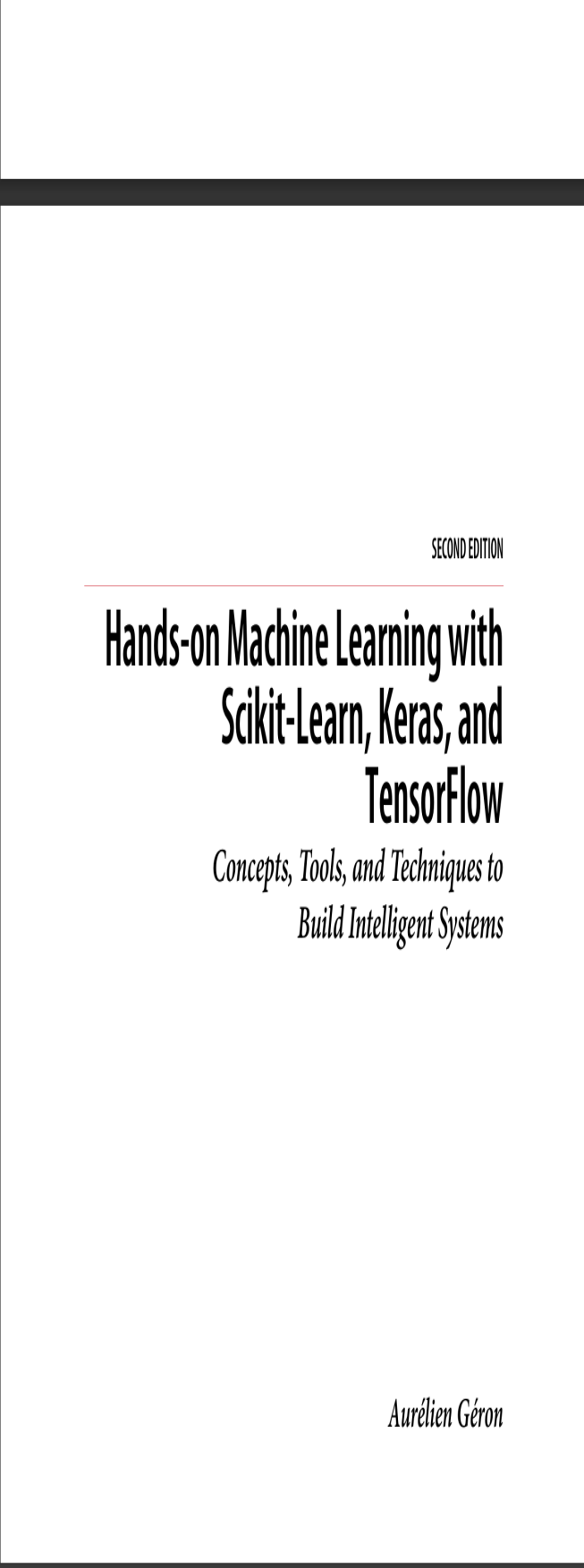
powered by



Early  
Release  
RAW &  
UNEDITED



2nd Edition  
Updated for  
TensorFlow 2



Start Screenshot

Screenshot 2/5 taken...

978-1-492-03264-9  
[LSI]

# Table of Contents

Preface..... xi

**Part I. The Fundamentals of Machine Learning**

1. The Machine Learning Landscape..... 3

- 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

Enter number of pages:

5

Start Screenshot

Screenshot 3/5 taken...

3. Classification..... 87

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

---

iv | Table of Contents

Multilabel Classification 108

Multioutput Classification 109

Exercises 110

4. Training Models..... 113

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

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