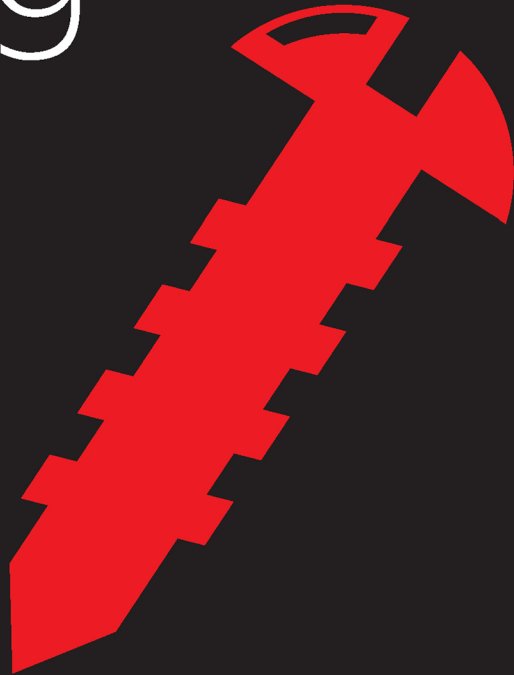


# Introducing Machine Learning



 Professional

Dino Esposito  
Francesco Esposito

# Introducing Machine Learning

Dino Esposito  
Francesco Esposito

# Introducing Machine Learning

## Table of Contents

Cover

Title Page

Copyright Page

Dedication

Contents at a Glance

Contents

Acknowledgments

About the Authors

Introduction

Part I: Laying the Groundwork of Machine Learning

Chapter 1 How Humans Learn

The Journey Toward Thinking Machines

The Dawn of Mechanical Reasoning

Godels Incompleteness Theorems

Formalization of Computing Machines

Toward the Formalization of Human Thought

The Birth of Artificial Intelligence as a Discipline

The Biology of Learning

What Is Intelligent Software, Anyway?

How Neurons Work

The Carrot-and-Stick Approach

Adaptability to Changes

Artificial Forms of Intelligence

# **Table of Contents**

Primordial Intelligence

Expert Systems

Autonomous Systems

Artificial Forms of Sentiment

Summary

## **Chapter 2 Intelligent Software**

Applied Artificial Intelligence

Evolution of Software Intelligence

Expert Systems

General Artificial Intelligence

Unsupervised Learning

Supervised Learning

Summary

## **Chapter 3 Mapping Problems and Algorithms**

Fundamental Problems

Classifying Objects

Predicting Results

Grouping Objects

More Complex Problems

Image Classification

Object Detection

Text Analytics

Automated Machine Learning

Aspects of an AutoML Platform

The AutoML Model Builder in Action

Summary

## **Chapter 4 General Steps for a Machine Learning Solution**

Data Collection

Data-Driven Culture in the Organization

Storage Options

Data Preparation

# **Table of Contents**

Improving Data Quality

Cleaning Data

Feature Engineering

Finalizing the Training Dataset

## **Model Selection and Training**

The Algorithm Cheat Sheet

The Case for Neural Networks

Evaluation of the Model Performance

## **Deployment of the Model**

Choosing the Appropriate Hosting Platform

Exposing an API

## **Summary**

## **Chapter 5 The Data Factor**

### **Data Quality**

Data Validity

Data Collection

### **Data Integrity**

Completeness

Uniqueness

Timeliness

Accuracy

Consistency

### **Whats a Data Scientist, Anyway?**

The Data Scientist at Work

The Data Scientist Tool Chest

Data Scientists and Software Developers

## **Summary**

## **Part II: Machine Learning in .Net**

### **Chapter 6 The .NET Way**

#### **Why (Not) Python?**

Why Is Python So Popular in Machine Learning?

# **Table of Contents**

Taxonomy of Python Machine Learning Libraries

End-to-End Solutions on Top of Python Models

## **Introducing ML.NET**

Creating and Consuming Models in ML.NET

Elements of the Learning Context

## **Summary**

## **Chapter 7 Implementing the ML.NET Pipeline**

### **The Data to Start From**

Exploring the Dataset

Applying Common Data Transformations

Considerations on the Dataset

### **The Training Step**

Picking an Algorithm

Measuring the Actual Value of an Algorithm

Planning the Testing Phase

A Look at the Metrics

### **Price Prediction from Within a Client Application**

Getting the Model File

Setting Up the ASP.NET Application

Making a Taxi Fare Prediction

Devising an Adequate User Interface

Questioning Data and Approach to the Problem

## **Summary**

## **Chapter 8 ML.NET Tasks and Algorithms**

### **The Overall ML.NET Architecture**

Involved Types and Interfaces

Data Representation

Supported Catalogs

### **Classification Tasks**

Binary Classification

Multiclass Classification

# Table of Contents

## Clustering Tasks

- Preparing Data for Work

- Training the Model

- Evaluating the Model

## Transfer Learning

- Steps for Building an Image Classifier

- Applying Necessary Data Transformations

- Composing and Training the Model

- Margin Notes on Transfer Learning

## Summary

## Part III: Fundamentals of Shallow Learning

### Chapter 9 Math Foundations of Machine Learning

#### Under the Umbrella of Statistics

- The Mean in Statistics

- The Mode in Statistics

- The Median in Statistics

#### Bias and Variance

- The Variance in Statistics

- The Bias in Statistics

#### Data Representation

- Five-number Summary

- Histograms

- Scatter Plots

- Scatter Plot Matrices

- Plotting at the Appropriate Scale

## Summary

### Chapter 10 Metrics of Machine Learning

#### Statistics vs. Machine Learning

- The Ultimate Goal of Machine Learning

- From Statistical Models to Machine Learning Models

#### Evaluation of a Machine Learning Model

# **Table of Contents**

From Dataset to Predictions

Measuring the Precision of a Model

## **Preparing Data for Processing**

Scaling

Standardization

Normalization

## **Summary**

## **Chapter 11 How to Make Simple Predictions: Linear Regression**

### **The Problem**

Guessing Results Guided by Data

Making Hypotheses About the Relationship

### **The Linear Algorithm**

The General Idea

Identifying the Cost Function

The Ordinary Least Square Algorithm

The Gradient Descent Algorithm

How Good Is the Algorithm?

### **Improving the Solution**

The Polynomial Route

Regularization

## **Summary**

## **Chapter 12 How to Make Complex Predictions and Decisions: Trees**

### **The Problem**

Whats a Tree, Anyway?

Trees in Machine Learning

A Sample Tree-Based Algorithm

### **Design Principles for Tree-Based Algorithms**

Decision Trees versus Expert Systems

Flavors of Tree Algorithms

### **Classification Trees**

How the CART Algorithm Works



# Table of Contents

How the ID3 Algorithm Works

## Regression Trees

How the Algorithm Works

Tree Pruning

## Summary

## Chapter 13 How to Make Better Decisions: Ensemble Methods

### The Problem

### The Bagging Technique

Random Forest Algorithms

Steps of the Algorithms

Pros and Cons

### The Boosting Technique

The Power of Boosting

Gradient Boosting

Pros and Cons

## Summary

## Chapter 14 Probabilistic Methods: Naïve Bayes

### Quick Introduction to Bayesian Statistics

Introducing Bayesian Probability

Some Preliminary Notation

Bayes Theorem

A Practical Code Review Example

### Applying Bayesian Statistics to Classification

Initial Formulation of the Problem

A Simplified (Yet Effective) Formulation

Practical Aspects of Bayesian Classifiers

### Naïve Bayes Classifiers

The General Algorithm

Multinomial Naïve Bayes

Bernoulli Naïve Bayes

Gaussian Naïve Bayes

# **Table of Contents**

## Naïve Bayes Regression

- Foundation of Bayesian Linear Regression

- Applications of Bayesian Linear Regression

## Summary

## Chapter 15 How to Group Data: Classification and Clustering

### A Basic Approach to Supervised Classification

- The K-Nearest Neighbors Algorithm

- Steps of the Algorithm

- Business Scenarios

### Support Vector Machine

- Overview of the Algorithm

- A Quick Mathematical Refresher

- Steps of the Algorithm

### Unsupervised Clustering

- A Business Case: Reducing the Dataset

- The K-Means Algorithm

- The K-Modes Algorithm

- The DBSCAN Algorithm

## Summary

## Part IV: Fundamentals of Deep Learning

### Chapter 16 Feed-Forward Neural Networks

#### A Brief History of Neural Networks

- The McCulloch-Pitt Neuron

- Feed-Forward Networks

- More Sophisticated Networks

#### Types of Artificial Neurons

- The Perceptron Neuron

- The Logistic Neuron

#### Training a Neural Network

- The Overall Learning Strategy

- The Backpropagation Algorithm

# Table of Contents

Summary

## Chapter 17 Design of a Neural Network

Aspects of a Neural Network

Activation Functions

Hidden Layers

The Output Layer

Building a Neural Network

Available Frameworks

Your First Neural Network in Keras

Neural Networks versus Other Algorithms

Summary

## Chapter 18 Other Types of Neural Networks

Common Issues of Feed-Forward Neural Networks

Recurrent Neural Networks

Anatomy of a Stateful Neural Network

LSTM Neural Networks

Convolutional Neural Networks

Image Classification and Recognition

The Convolutional Layer

The Pooling Layer

The Fully Connected Layer

Further Neural Network Developments

Generative Adversarial Neural Networks

Auto-Encoders

Summary

## Chapter 19 Sentiment Analysis: An End-to-End Solution

Preparing Data for Training

Formalizing the Problem

Getting the Data

Manipulating the Data

Considerations on the Intermediate Format

# **Table of Contents**

## Training the Model

- Choosing the Ecosystem

- Building a Dictionary of Words

- Choosing the Trainer

- Other Aspects of the Network

## The Client Application

- Getting Input for the Model

- Getting the Prediction from the Model

- Turning the Response into Usable Information

## Summary

## Part V: Final Thoughts

### Chapter 20 AI Cloud Services for the Real World

#### Azure Cognitive Services

#### Azure Machine Learning Studio

- Azure Machine Learning Service

- Data Science Virtual Machines

#### On-Premises Services

- SQL Server Machine Learning Services

- Machine Learning Server

#### Microsoft Data Processing Services

- Azure Data Lake

- Azure Databricks

- Azure HDInsight

- .NET for Apache Spark

- Azure Data Share

- Azure Data Factory

## Summary

### Chapter 21 The Business Perception of AI

#### Perception of AI in the Industry

- Realizing the Potential

- What Artificial Intelligence Can Do for You

# **Table of Contents**

Challenges Around the Corner

End-to-End Solutions

Lets Just Call It Consulting

The Borderline Between Software and Data Science

Agile AI

Summary

Index