

# Introduction to Machine Learning

Andreas Müller

# What and Why of Machine Learning

What is machine learning?



Search Facebook



Andreas

Home



Andreas Mueller

News Feed

...

Messenger

SHORTCUTS

Amazon Machine L...

4

Freunde +1

EXPLORE

Events

Pages

Groups

Friend Lists

Pokes

Photos

On This Day

Photos

Suggest Edits

Offers

Moments

See More...

4

20+

CREATE

Ad · Page · Group · Event ·

Fundraiser



untapt

Sponsored · 0

Like Page

Accomplish more in 2017. See this year's most coveted fintech engineering jobs.

add a   
to your resume.



### New job. New adventure.

We've trained computer models on thousands of successful job applications to understand how the best engineers land their dream jobs. Now it's time to land yours.

UNTAPT.COM

Learn More



4

3 Comments

Like

Comment

Share



Juliana Vanderlee is  watching Tom Price's confirmation hearing.  
6 mins · 

All of these people are literally some of the worst possible choices for their positions.

Like

Comment

Share



2



Write a comment...



Cédric Archambeau and 1 other

TRENDING

Gmail

190K people talking about this

Pierce Brosnan

9.8K people talking about this

Wii U

120K people talking about this

See More

CONNECT WITH FACEBOOK



Internet.org by Facebook

7,615,868 likes



Instagram

40,973,994 likes



Facebook Engineering

8,548,617 likes



Mark Zuckerberg

84,629,591 followers

PEOPLE YOU MAY KNOW

See All



Dmytro Pavlichenko

4 mutual friends

Add Friend



Lee Tanenbaum

2 mutual friends

Add Friend

English (US) · Español · Português (Brasil) ·  
Français (France) · DeutschPrivacy · Terms · Advertising · Ad Choices ·  
Cookies · More

Facebook © 2017

Search for people, places and things

Andreas | Home

Andreas Mueller  
Edit Profile

Update Status Add Photos/Videos

News Feed

Messages

Events

GROUPS

Amaz...  
Surf...  
Tob...  
VER...  
Hoc...  
Mei...  
Ber...  
Cre...

APPS

Gam...  
Gla...  
Phot...  
Peop...  
Mus...  
Citi...  
On...  
Gam...

FRIENDS

Amaz...  
Rud...  
Univ...  
Univ...  
Close...  
Ber...

INTERESTS

Page...

PAGES

June 6 at 5:07pm

What's on your mind?

9 CrossFit Icke Outdo... tomorrow

SPONSORED See All

Online Essen bestellen  
lieferheld.de

Da ist Abwechslung drin!  
franziskaner-weissbier.de

Andreas Mueller June 6

Add a description

Tag Photo Add Location Edit

Like · Comment · Stop Notifications · Share · Edit

Julia Hartmann likes this.

Write a comment...

Sponsored See All

Vielzahl-Pack gewinnen!  
franziskaner-weissbier.de

Wir verlosen Vielfalt-Packs von Franziskaner Weissbier. Hier klicken!

Franziskaner WEISSBIER

Online Essen bestellen  
lieferheld.de

Neu mit Lieferheld: PLZ eingeben, Lieferdienst finden und genießen!

car2go Deutschland

60 Freiminuten bei car2go

Finde die Diversity Day car2go in deiner Stadt! Foto machen und mit Glück 60 Freiminuten abstauben.

Like Page · 48,602 people like this page

1967年



Andreas Mueller  
Edit Profile

 Update Status  Add Photos/Videos

What's on your mind?



June 6 at 5:07pm · 

9 CrossFit Icke Outdoo... tomorrow

SPONSORED 

See All

Online Essen bestellen  
lieferheld.de

Da ist Abwechslung drin!

franziskaner-weissbier.de



Search for people, places and things



 Andreas | Home



 Andreas Mueller added 8 new photos.

June 6 at 5:07pm · 

Bye bye hong kong



Like · Comment · Share

 **1000** People, **Coco Antonietta, Wilson Wong and 6 others** like this.

 Louise Kastell Looks like such a vibrant city! I bet it was awesome 😊  
June 6 at 7:39pm · Like

 Write a comment...

SPONSORED 

See All

Da Ist Abwechslung drin!

franziskaner-weissbier.de



3 Sorten Franziskaner Weissbier in einem Pack.  
Hier klicken und 1 von 100 Packs gewinnen!

Singles auf Facebook



Schau dir Dating-Profile von Singles in deiner Nähe an.

Online Essen bestellen

lieferheld.de



Neu mit Lieferheld: PLZ eingeben, Lieferdienst finden und genießen!

Globaler Chauffeurservice

blacklane.com



Fahren Sie eine Klasse besser zu günstigen Preisen – Blacklane ist weltweit verfügbar!

28,723 people like this

Sanssouci...

Geld vom Staat zurück!



Steuererklärung preiswert für Arbeitnehmer, Azubis, Arbeitsuchende, Rentner, Pensionäre etc.

 - 4 people like this page

Der neue WhatsApp Tarif!

eplus.de



WhatsApp SIM – der revolutionäre Prepaid Tarif

Der Prepaid Tarif 😊

Search

Andreas Mueller

Browse

Radio

YOUR MUSIC

Your Daily Mix

Songs

Albums

Stations

Artists

Local Files

PLAYLISTS

Discover Weekly

Your Top Songs 2016

loud clubby music

rammstein rope

Starred

rope\_hard

rope\_ebm

rope\_silly

DUBSTEP

Rammstein — Reise, ...

Rammstein — LIEBE IS...

Die Antwoord — Donk...

Combichrist — Everyb...

System Of A Down — ...

System Of A Down — ...

New Playlist

Electronic World Transmiss ✓  
Rotersand

Rammstein

PLAY FOLLOWING ...

OVERVIEW RELATED ARTISTS ABOUT CONCERTS

Latest Release

RAMMSTEIN XXI - Klavier 11 NOV 2016

Popular

1	+	Du hast	50,593,821
2	↻	Sonne	23,137,196
3	+	Feuer Frei	20,253,606
4	!	AMERIKA	18,832,022
5	+	Ich Will	18,126,559

SHOW 5 MORE

On Tour

JUN 25 View upcoming concerts in your country  
Rammstein Nikon at Jones Beach Theater (Wantagh)

Friend activity

Martin Kremers I See a Darkness Acid Pauli Bar 25 - Tage ausser... 7 h

More Friends

1112775392

1122190832

1122256379

Angelo Laub

Dominic Schrögendorfer

Florian Stemmer

jaymee

Julia Hartmann

larsborn

Player controls: back, forward, play/pause, volume, seek bar (0:05 to 4:06)

www.amazon.com/s/ref=nb\_sb\_noss\_1?url=search-alias%3Daps&field-keywords=machine learning&sprefix=machine+%

DuckDuckGo

Father's Day is June 15  
Sponsored by DeWalt >Shop now

amazon Try Prime

Andreas's Amazon.com Today's Deals Gift Cards Sell Help

Shop by Department Search All machine learning Go Hello, Andreas Your Account Try Prime Cart Wish List Choose a Department to sort

1-16 of 21,549 results for "machine learning"

Show results for Books > Machine Learning Computers & Technology Computer Science Reference Programming Algorithms + See more Kindle Store > Computers & Technology Computer Programming + See more + See All 30 Departments

Refine by Eligible for Free Shipping Free Shipping by Amazon Book Series Information Science and Statistics Adaptive Computation and Machine Learning series I Can Draw Use RI Book Language English Book Format Hardcover Paperback Kindle Edition HTML Avg. Customer Review ★★★★ & Up ★★★★★ & Up ★★★★★ & Up ★★★★★ & Up

**Machine Learning Resources**  
Find tips and tricks with these featured titles on machine learning, algorithms, sensors, and more. [Learn more](#)

**Related Searches:** [artificial intelligence](#), [data mining](#), [machine learning python](#).

**Practical Machine Learning: Innovations in Recommendation** by Ted Dunning and Ellen Friedman (Apr 17, 2014)  
**\$0.00 Kindle Edition**  
Auto-delivered wirelessly

**Machine Learning: The Art and Science of Algorithms that Make Sense of Data** by Peter Flach (Nov 12, 2012)  
**\$64.00 \$54.96 Paperback** Prime  
Only 18 left in stock - order soon.  
More Buying Choices - Paperback  
**\$46.98 new** (40 offers)  
**\$35.00 used** (18 offers)

**Understanding Machine Learning: From Theory to Algorithms** by Shai Shalev-Shwartz and Shai Ben-David (May 19, 2014)  
**\$60.00 \$50.92 Hardcover** Prime  
Order in the next **11 hours** and get it by Tuesday, Jun 10.  
**\$34.35 Kindle Edition**  
Auto-delivered wirelessly  
More Buying Choices - Hardcover  
**\$50.92 new** (8 offers)  
**\$52.02 used** (4 offers)

**Learning From Data** by Yaser S. Abu-Mostafa, Malik Magdon-Ismail and Hsuan-Tien Lin (Mar 27, 2012)  
**\$28.00 new** (9 offers)  
**\$40.00 used** (13 offers)

**★ ★ ★ ★ & Up** **#1 Best Seller** in Machine Learning  
Trade-in eligible for an Amazon gift card  
Books: See all 18,805 items

[www.amazon.com/Fathers-Day-Gifts-Sale/b/ref=n](http://www.amazon.com/Fathers-Day-Gifts-Sale/b/ref=n)

www.amazon.com/s/ref=nb\_sb\_noss\_1?url=search-alias%3Daps&field-keywords=machine learning&sprefix=machine+%2Caps&rh=

DuckDuckGo Father's Day Is June 15 Sponsored by DeWalt

Amazon Try Prime Andreas's Amazon.com Today's Deals Gift Cards Sell Help

Shop by Department Search All machine learning Go Hello, Andreas Your Account Try Prime Cart Wish List Choose a Department to sort

1-16 of 21,549 results for "machine learning"

Show results for Books > Machine Learning Computers & Technology Computer Science Reference Programming Algorithms + See more Kindle Store > Computers & Technology Computer Programming + See more + See All 30 Departments Refine by Eligible for Free Shipping Free Shipping by Amazon Book Series Information Science and Statistics Adaptive Computation and Machine Learning series I Can Draw Use R! Book Language English Book Format Hardcover Paperback Kindle Edition HTML Avg. Customer Review ★★★★☆ & Up ★★★★★ & Up ★★★☆☆ & Up ★☆☆☆☆ & Up

**Introduction to Machine Learning with Python: A Guide for Data Scientists** 1st Edition

by Andreas C. Müller (Author), Sarah Guido (Author)  
★★★★★ 9 customer reviews  
[Look inside](#)

**Kindle** \$24.99    **Paperback** \$41.23    **Other Sellers** from \$26.53

**Buy new**    **In Stock.**    **Prime** \$41.23    List Price: \$49.99 Save: \$8.76 (18%)  
 Ships from and sold by Amazon.com. Gift-wrap available.  
Prime

Want it tomorrow, Jan. 19? Order within **5 hrs 20 mins** and choose **One-Day Shipping** at checkout. [Details](#)

**Add to Cart**    **Turn on 1-Click ordering**    **Ship to:** Andreas C Mueller- New York - 10009

**More Buying Choices**    51 used & new from \$26.53    [See All Buying Options](#)

ISBN-13: 978-1449369415    ISBN-10: 1449369413  
[Why is ISBN important?](#)

Have one to sell? [Sell on Amazon](#)

[Add to List](#)

Share [Email](#) [Facebook](#) [Twitter](#) [Pinterest](#)

Machine learning has become an integral part of many commercial applications and research projects, but this field is not exclusive to large companies with extensive research teams. If you use Python, even as a beginner, this book will teach you practical ways to build your own machine learning solutions. With all the data available today, machine learning applications are limited only by your imagination.

You'll learn the steps necessary to create a successful machine-learning application with Python and the [Read more](#)

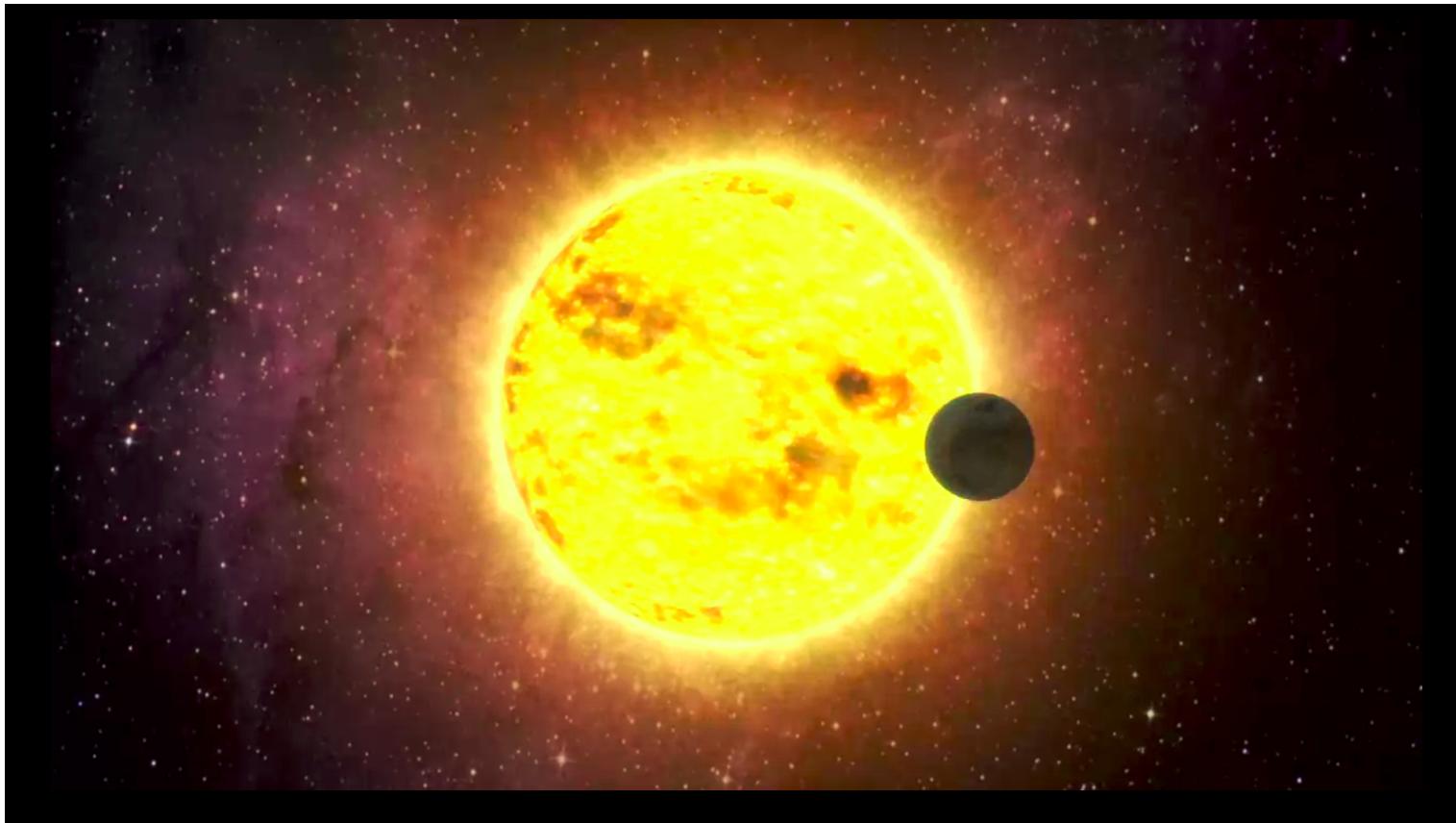
[Report incorrect product information.](#)

**Save up to 90% on textbooks** [Shop now](#)

**Frequently Bought Together**

Total price: \$108.29  
[Add all three to Cart](#)    [Add all three to List](#)

# Science!



Types of machine learning:

- supervised
- unsupervised
- reinforcement

# Supervised Learning

$$(x_i, y_i) \propto p(x, y) \quad \text{i.i.d.}$$

$$x_i \in \mathbb{R}^n$$

$$y_i \in \mathbb{R}$$

$$f(x_i) \approx y_i$$

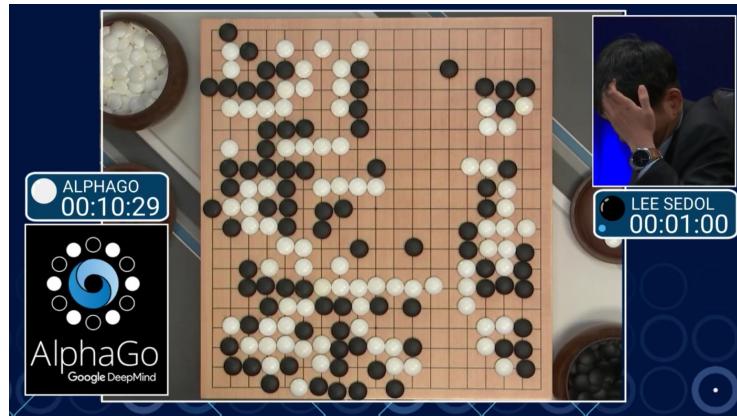
# Examples of Supervised Learning

# Unsupervised Learning

$$x_i \propto p(x) \quad \text{i.i.d.}$$

Learn about p

# Reinforcement Learning



Explore & Learn

# Other kinds of learning

- Semi-Supervised
- Active Learning
- Forecasting
- ...

# Classification and Regression

Classification:

- $y$  discrete

Will you pass?

Regression:

- $y$  continuous

How many points will  
you get in the exam?

# Generalization

Not only

$$f(x_i) \approx y_i$$

Also for new data:

$$f(x) \approx y$$

# Relationship to Statistics

## Statistics

- Model first
- Inference

## Machine Learning

- Data first
- Prediction
- Generalization

# Relationship to Statistics

## Statistics

- Model first
- Inference

## Machine Learning

- Data first
- Prediction
- Generalization

# Guiding principles in machine learning

# Goal considerations

## The cost of complex systems

Data driven first: yes! (or maybe)  
Machine Learning first: No!

Thinking in Context!  
What is the baseline?  
What is the benefit?

Thinking in Context!  
What is the baseline?  
What is the benefit?

# Good and bad substitutes

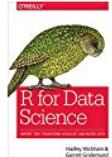
# Communicating Results

# Explainable Results

These recommendations are based on items you own and more.

view: All | New Releases | Coming Soon

1.



**R for Data Science: Import, Tidy, Transform, Visualize, and Model Data**

by Hadley Wickham (January 5, 2017)

Average Customer Review: ★★★★☆ (4)

In Stock

List Price: \$39.99

Price: \$32.91

[28 used & new from \\$28.91](#)

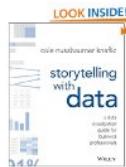
Add to Cart Add to Wish List

I own it  Not interested

Rate this item

Recommended because you purchased **Data Science from Scratch: First Principles with Python** and more ([Fix this](#))

2.



**Storytelling with Data: A Data Visualization Guide for Business Professionals**

by Cole Nussbaumer Knaflic (November 2, 2015)

Average Customer Review: ★★★★☆ (137)

In Stock

List Price: \$39.95

Price: \$29.71

[94 used & new from \\$20.00](#)

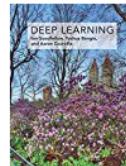
Add to Cart Add to Wish List

I own it  Not interested

Rate this item

Recommended because you purchased **Mindset: The New Psychology of Success** and more ([Fix this](#))

3.



**Deep Learning (Adaptive Computation and Machine Learning series)**

by Ian Goodfellow (November 18, 2016)

Average Customer Review: ★★★★☆ (20)

In Stock

List Price: \$80.00

Price: \$68.34

[18 used & new from \\$68.34](#)

Add to Cart Add to Wish List

I own it  Not interested

Rate this item

Recommended because you purchased **Data Science from Scratch: First Principles with Python** and more ([Fix this](#))

# Sidebar: Ethical Considerations



<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

Ethics: It's in the application!

# Data and Data Collection

# Free vs Expensive Data

## Free:

Predict observable  
events

- Stock market
- Clicks
- House numbers

# Free vs Expensive Data

## Free:

Predict observable events

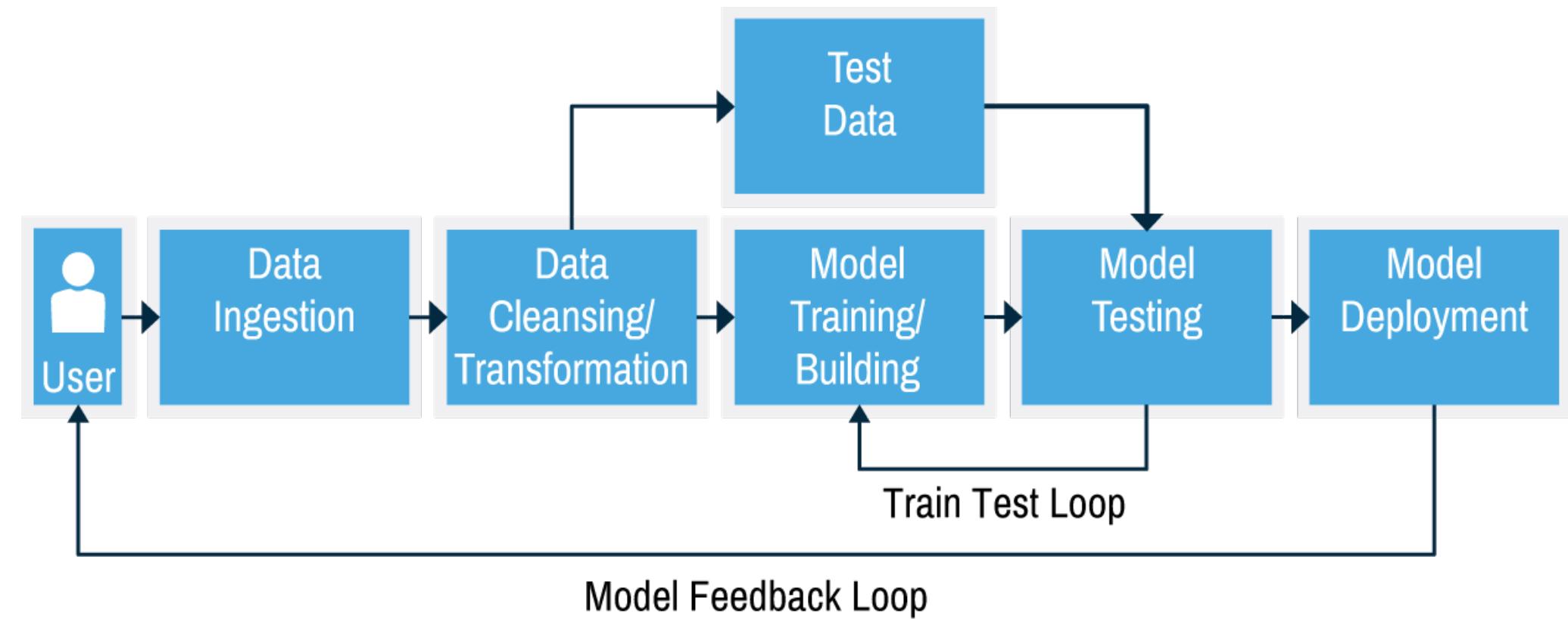
- Stock market
- Clicks
- House numbers

## Expensive:

Automate complex process

- Diagnosis
- Drug Trial
- Chip Design

# The Machine Learning Work-Flow



Taken from MAPR <https://www.mapr.com/ebooks/spark/08-recommendation-engine-spark.html>

# Representing Data

one sample

$X =$

1.1	2.2	3.4	5.6	1.0
6.7	0.5	0.4	2.6	1.6
2.4	9.3	7.3	6.4	2.8
1.5	0.0	4.3	8.3	3.4
0.5	3.5	8.1	3.6	4.6
5.1	9.7	3.5	7.9	5.1
3.7	7.8	2.6	3.2	6.3

one feature

$y =$

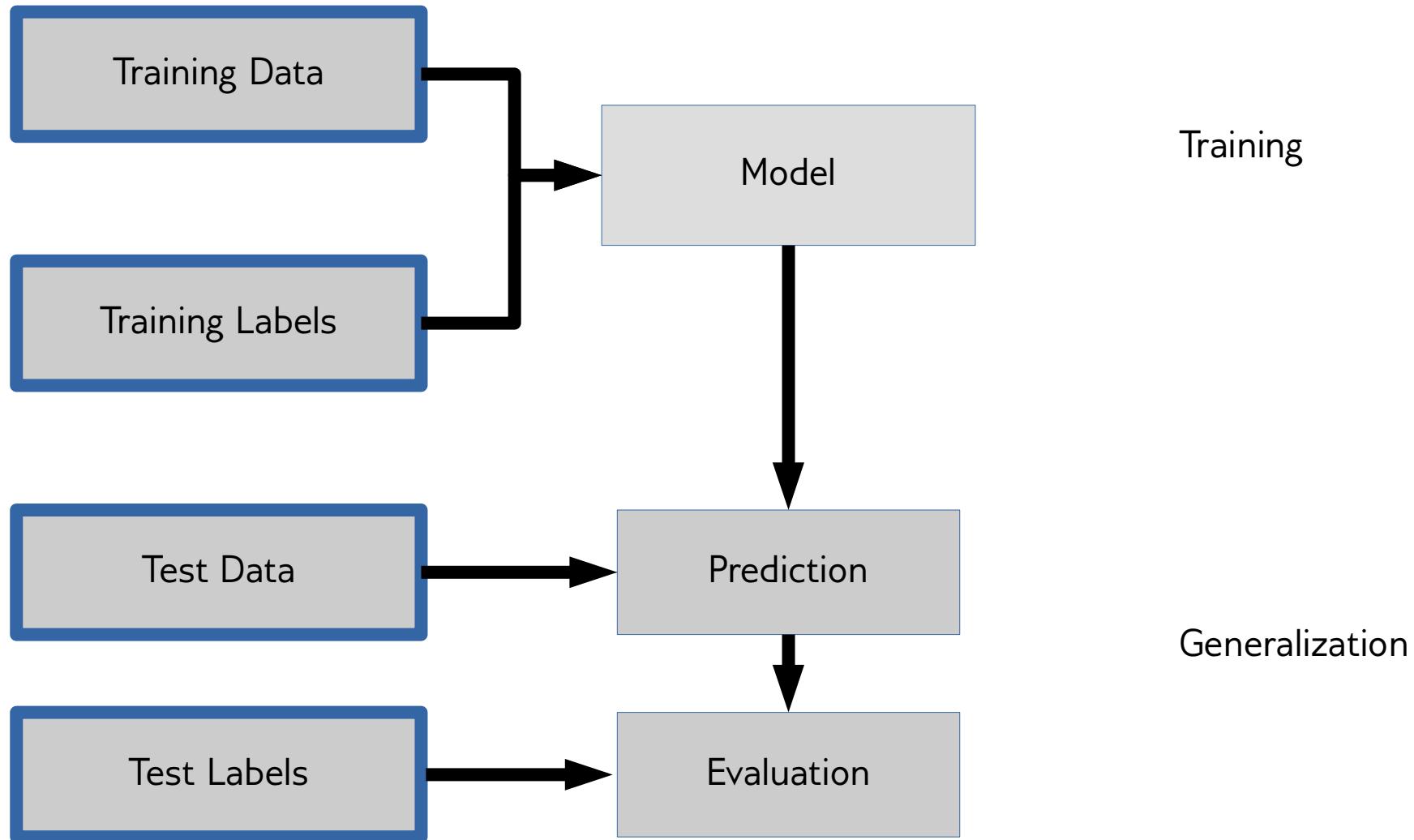
1.6
2.7
4.4
0.5
0.2
5.6
6.7

outputs / labels

# Training and Testing Data

training set					
X =	1.1	2.2	3.4	5.6	1.0
	6.7	0.5	0.4	2.6	1.6
	2.4	9.3	7.3	6.4	2.8
	1.5	0.0	4.3	8.3	3.4
	0.5	3.5	8.1	3.6	4.6
test set					
y =	1.6	2.7	4.4	0.5	0.2
	5.6	6.7			

# Supervised Machine Learning



```
clf = RandomForestClassifier()
```

```
clf.fit(X_train, y_train)
```

```
y_pred = clf.predict(X_test)
```

```
clf.score(X_test, y_test)
```

