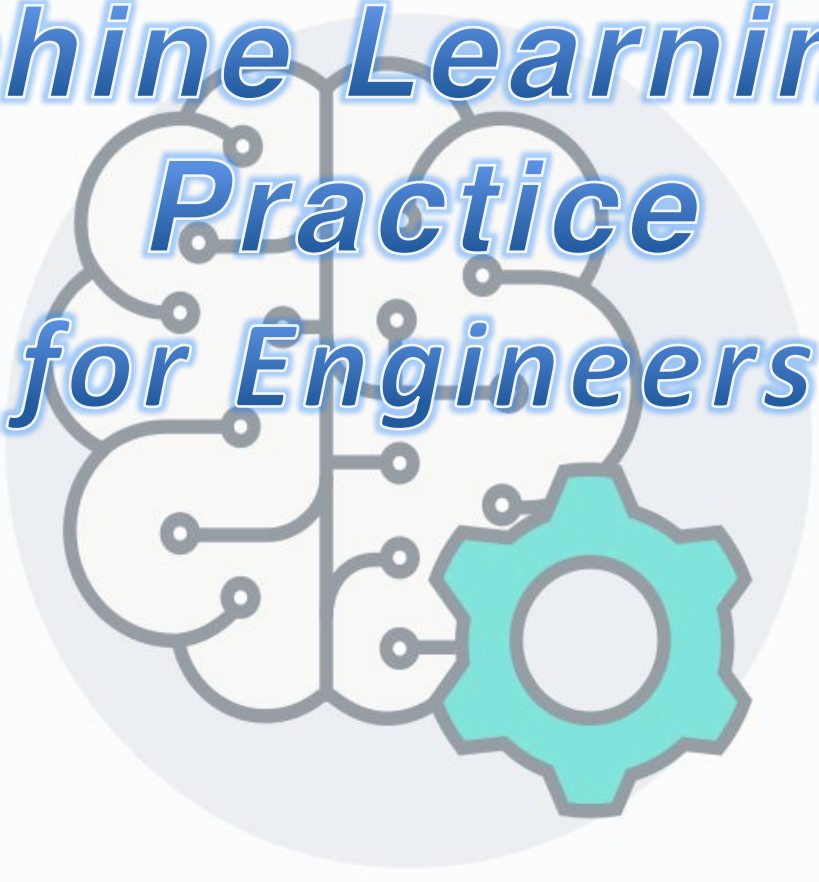
# Machine Learning in Practice for Engineers



*Chao Cai*
*@AICon*
*2018.1*


# 极客时间

重拾极客精神·提升技术认知

## 下载极客时间App

获取有声IT新闻、技术产品专栏，每日更新

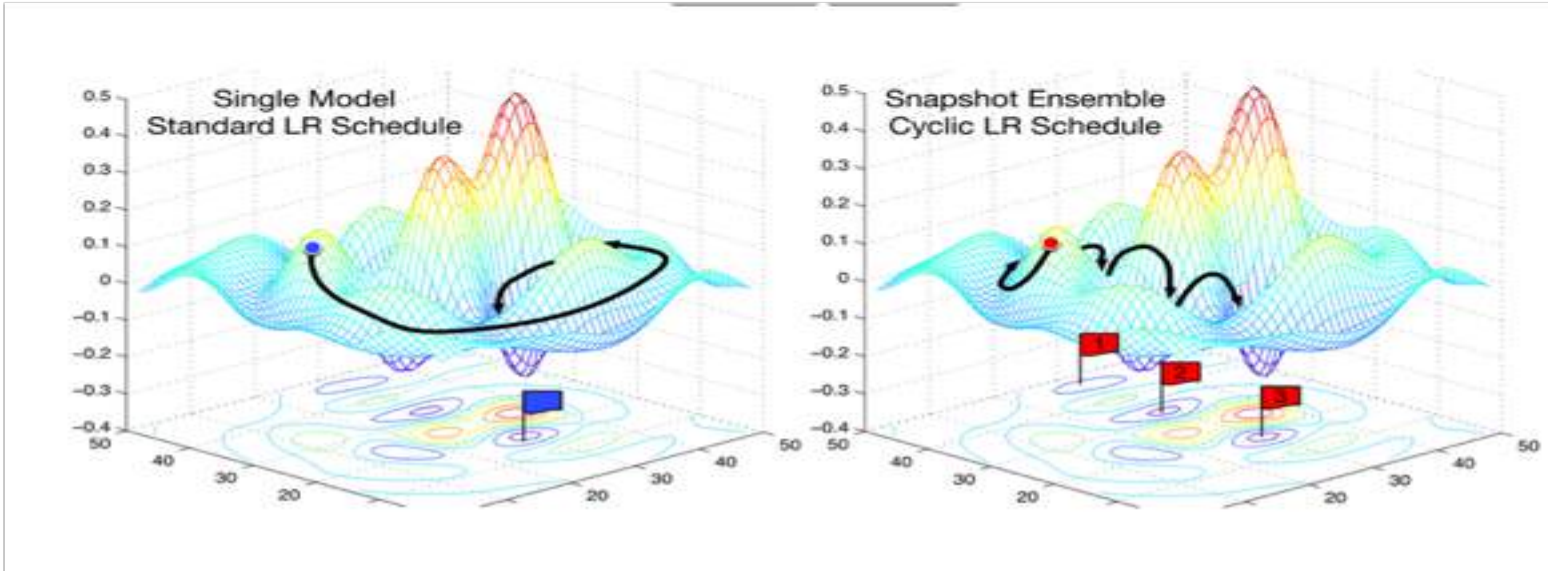
扫一扫下载极客时间App


人工智能基础课
“通俗易懂的人工智能入门课”
王天一
博士 副教授
扫一扫，免费试读


AI技术内参
你的360度人工智能信息助理
洪亮劼
Etsy 数据科学主管
扫一扫，免费试读

$$
\begin{aligned}
m_t &= \beta_1 m_{t-1} + (1-\beta_1) g_t \\
v_t &= \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \\
\hat{m}_t &= \frac{m_t}{1-\beta_1^t} \\
\hat{v}_t &= \frac{v_t}{1-\beta_2^t} \\
\theta_{t+1} &= \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t - \eta w_t \theta_t
\end{aligned}
$$

$$
\begin{aligned}
m_t &= \beta_1 m_{t-1} + (1-\beta_1) g_t \\
v_t &= \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \\
\hat{v}_t &= \max(\hat{v}_{t-1}, v_t) \\
\theta_{t+1} &= \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} m_t
\end{aligned}
$$



# Common Traps

Accuracy ≠ most important metric

$$\frac{2{,}000}{1{,}000{,}000}$$

No Information Rate = 99.8%

*Accuracy = (TP + TN) / (P + N)*

*Recall/Sensitivity/TPR = TP/ P*

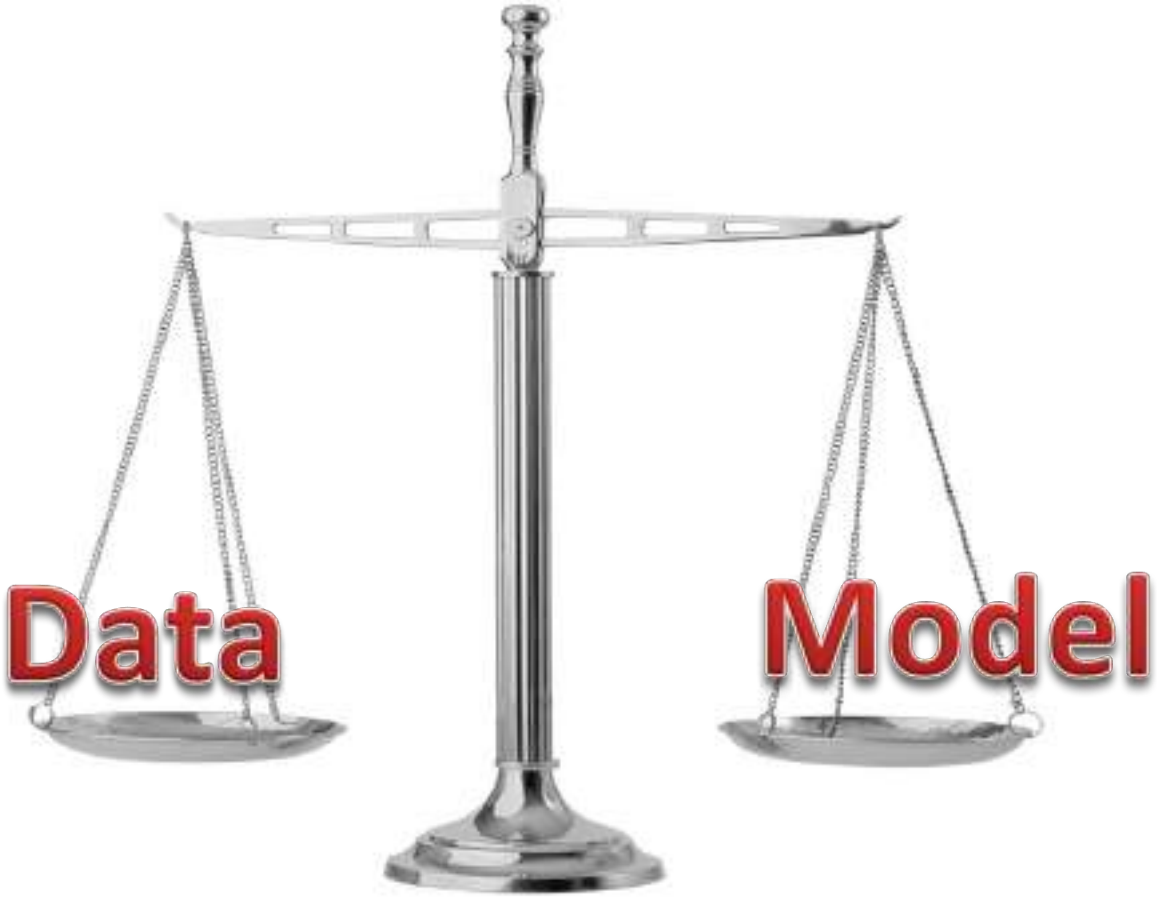
*Specificity = TN / N*

*FRP = FP/N = 1 - Specificity*

*Precision = TP / (TP + FP)*

*Informedness = Sensitivity + Specificity - 1*

*ROC Curve (x= FPR, y=TPR)*


Data
Model

修订您的亚马逊帐户 — 收件箱

**Amazon.cn**

收件箱 - iCloud　2017年1月17日 下午11:36

修订您的亚马逊帐户
收件人：chaocai2001@icloud.com,
回复－收件人：customer_service@cs.amazon.cn

感谢您访问 Amazon！ 按照您的要求，我们已经更改了与您的帐户关联的电子邮件地址。

与您的帐户关联的电子邮件地址已被更改。原地址为 chaocai2001@yahoo.com.cn。新地址为 chaocai2001@icloud.com。

请访问您在 Amazon 上的帐户查看您的订单，更改尚未进入发货流程的订单，更新订阅及其它许多内容。

如果您因任何原因需要联系我们，请注意，我们只能将订单信息提供给与您帐户关联的姓名和电子邮件地址。

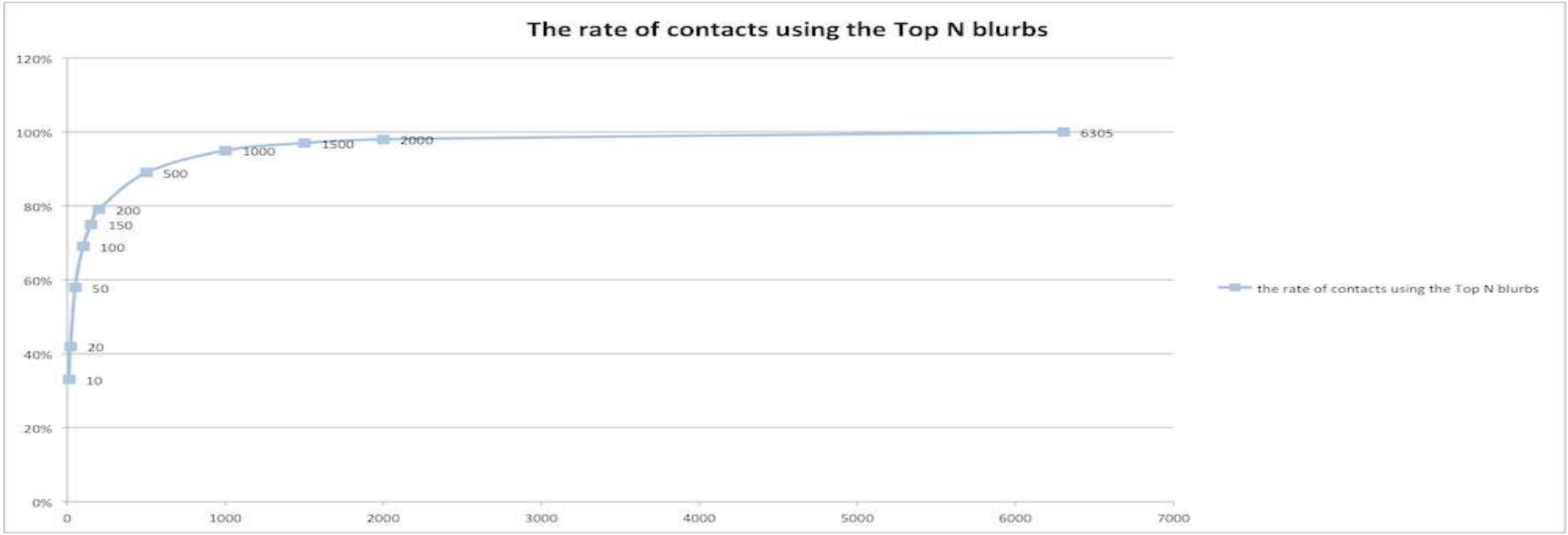
再次感谢您的惠顾。

# 100,000

NO

YES

# Recommend by TopN



# More is Better?

```
> head(Advertising)
     TV Radio Newspaper Sales
1 230.1  37.8      69.2  22.1
2  44.5  39.3      45.1  10.4
3  17.2  45.9      69.3   9.3
4 151.5  41.3      58.5  18.5
5 180.8  10.8      58.4  12.9
6   8.7  48.9      75.0   7.2
```

...

model<-train(**Sales ~.**, data=Advertising, method="glm",trControl=ctrl)

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.938889   0.311908   9.422   <2e-16 ***
TV           0.045765   0.001395  32.809   <2e-16 ***
Radio        0.188530   0.008611  21.893   <2e-16 ***
Newspaper   -0.001037   0.005871  -0.177     0.86
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

| RMSE | Rsquared |
|---|---|
| 1.723615 | 0.8914224 |

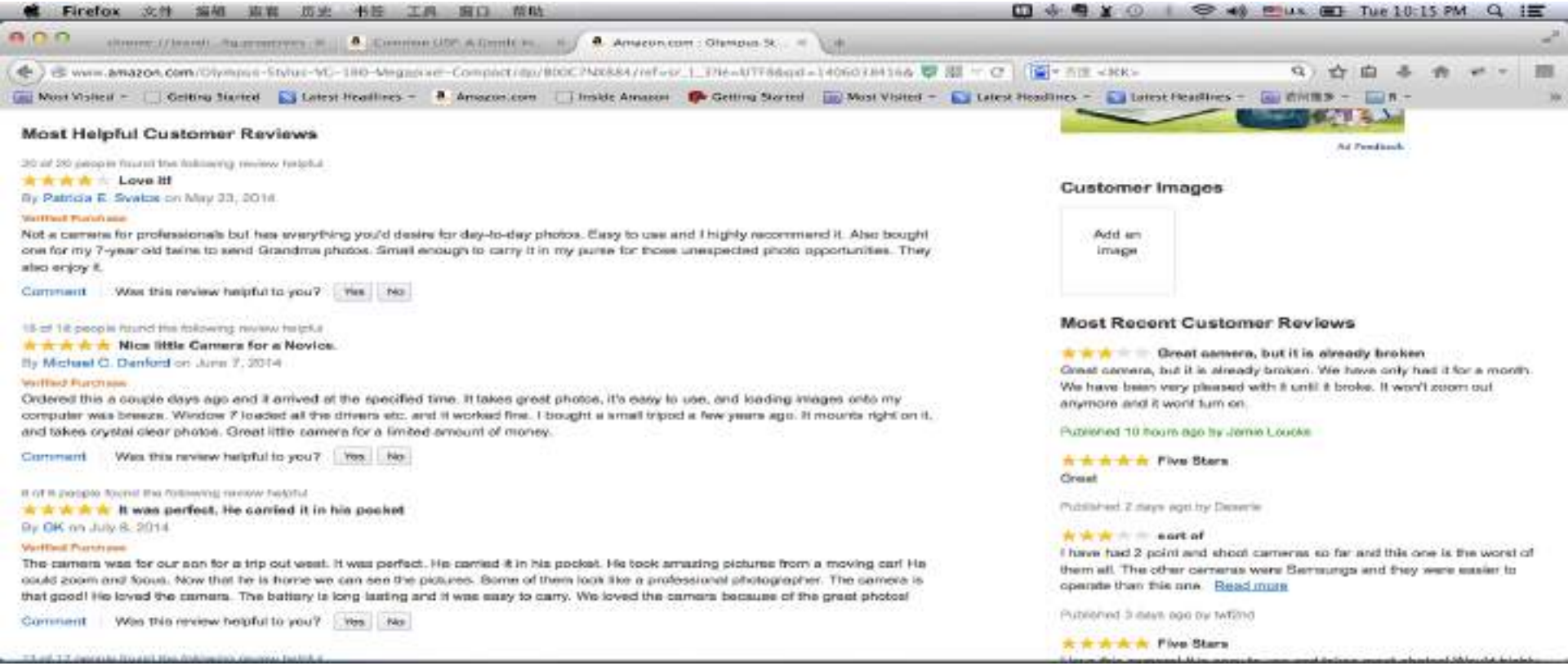
model<-train(**Sales ~TV+Radio** , data=Advertising, method="glm",trControl=ctrl)

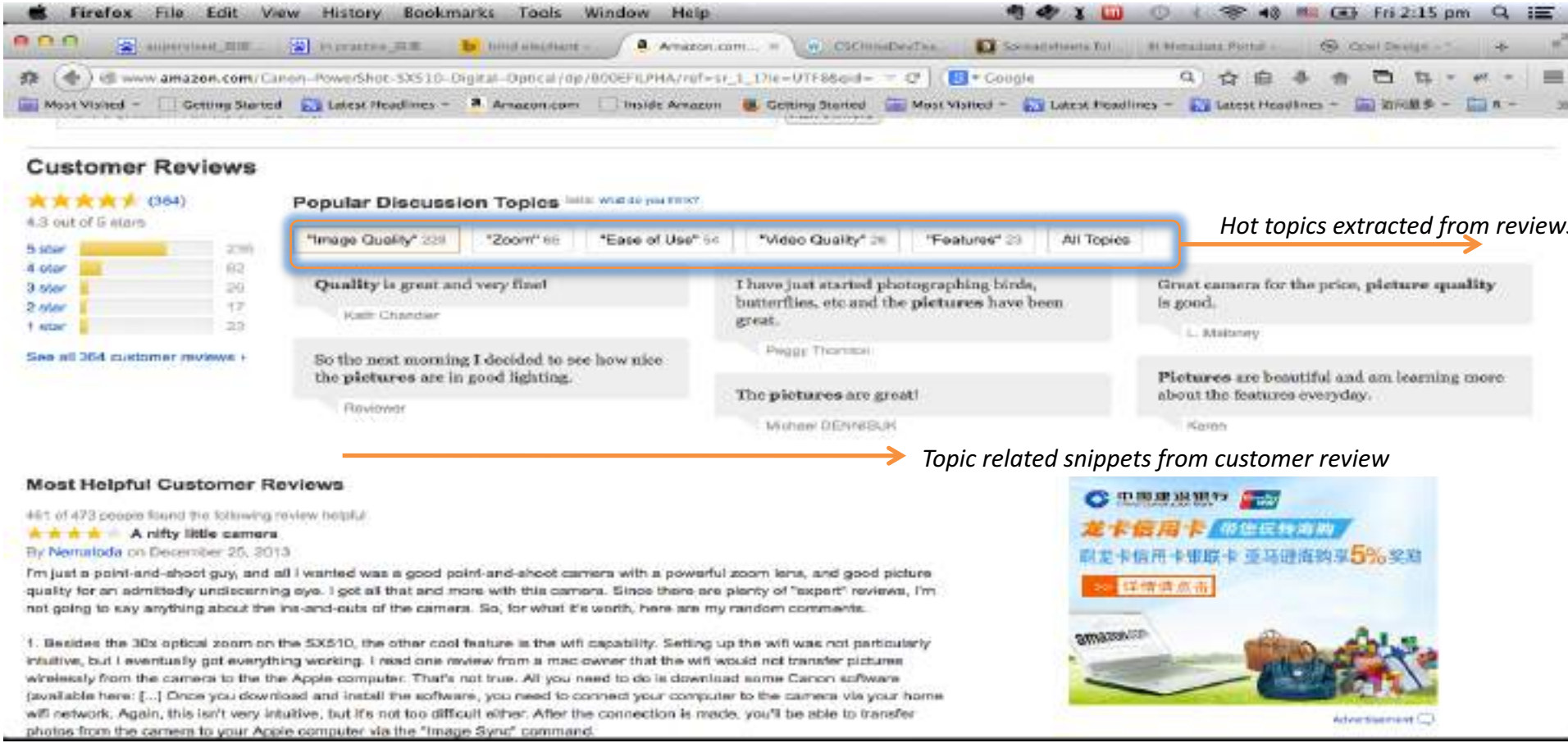
```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.92110    0.29449   9.919   <2e-16 ***
TV           0.04575    0.00139  32.909   <2e-16 ***
Radio        0.18799    0.00804  23.382   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**RMSE Rsquared**
**1.676055 0.8975927**

# Only Supervised?

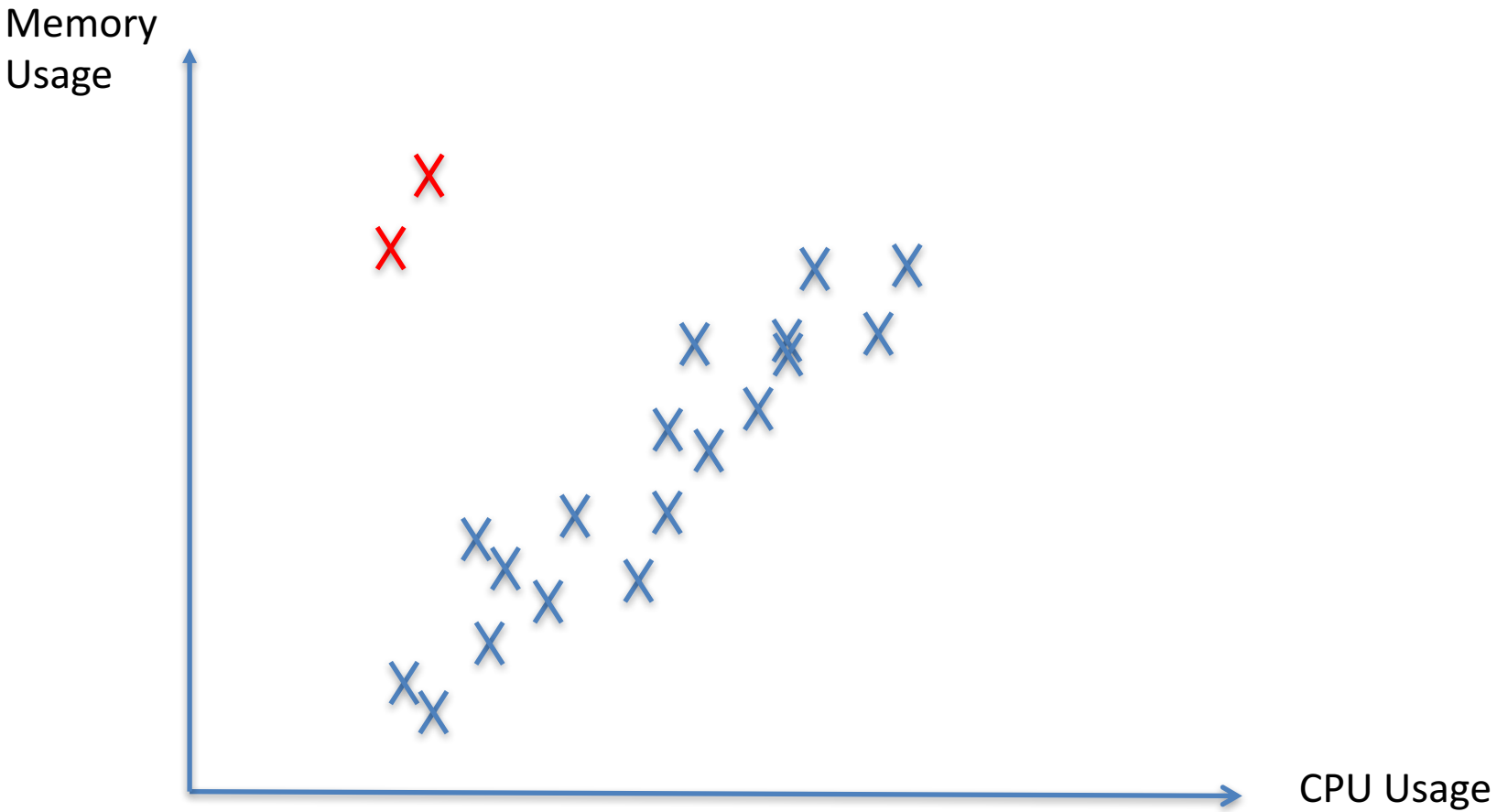

# Clustering

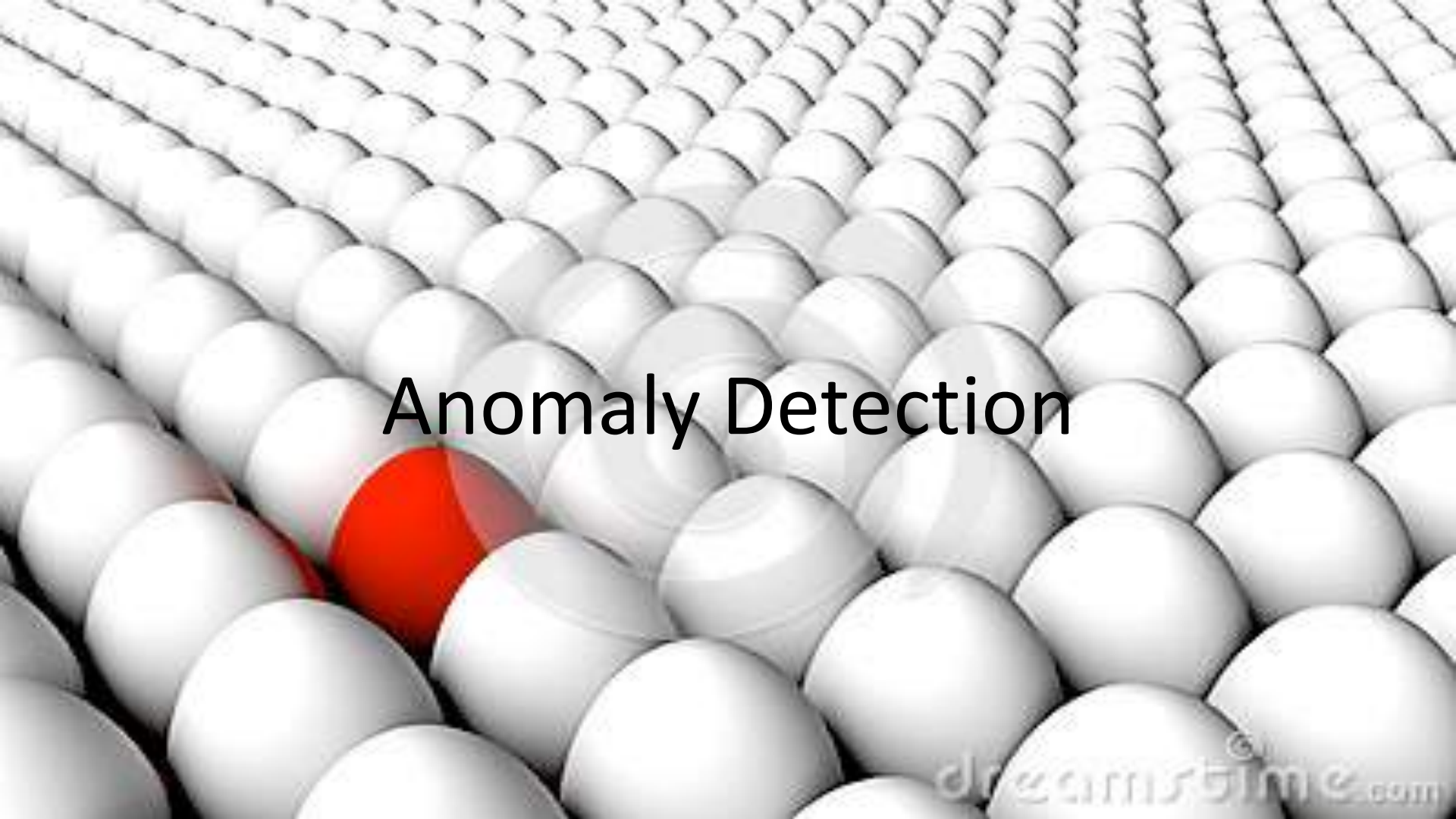

Firefox 文件 编辑 查看 历史 书签 工具 窗口 帮助
Tue 10:15 PM
Amazon.com : Olympus St
www.amazon.com/Olympus-Stylus-VG-180-Megapixel-Compact/dp/B00C7N0684/ref=sr_1_3?ie=UTF8&qid=1406018416&
Most Visited
Getting Started
Latest Headlines
Amazon.com
Inside Amazon
Most Helpful Customer Reviews
20 of 20 people found the following review helpful
Love it!
By Patricia E. Svatos on May 23, 2014
Verified Purchase
Not a camera for professionals but has everything you'd desire for day-to-day photos. Easy to use and I highly recommend it. Also bought one for my 7-year old twins to send Grandma photos. Small enough to carry it in my purse for those unexpected photo opportunities. They also enjoy it.
Comment
Was this review helpful to you?
Yes
No
18 of 18 people found the following review helpful
Nice little Camera for a Novice.
By Michael C. Danford on June 7, 2014
Verified Purchase
Ordered this a couple days ago and it arrived at the specified time. It takes great photos, it's easy to use, and loading images onto my computer was breeze. Window 7 loaded all the drivers etc. and it worked fine. I bought a small tripod a few years ago. It mounts right on it, and takes crystal clear photos. Great little camera for a limited amount of money.
Comment
Was this review helpful to you?
Yes
No
8 of 8 people found the following review helpful
It was perfect. He carried it in his pocket
By OK on July 8, 2014
Verified Purchase
The camera was for our son for a trip out west. It was perfect. He carried it in his pocket. He took amazing pictures from a moving car! He could zoom and focus. Now that he is home we can see the pictures. Some of them look like a professional photographer. The camera is that good! He loved the camera. The battery is long lasting and it was easy to carry. We loved the camera because of the great photos!
Comment
Was this review helpful to you?
Yes
No
Ad Feedback
Customer Images
Add an image
Most Recent Customer Reviews
Great camera, but it is already broken
Great camera, but it is already broken. We have only had it for a month. We have been very pleased with it until it broke. It won't zoom out anymore and it wont turn on.
Published 10 hours ago by Jamie Loucks
Five Stars
Great
Published 2 days ago by Deserie
sort of
I have had 2 point and shoot cameras so far and this one is the worst of them all. The other cameras were Samsungs and they were easier to operate than this one. Read more
Published 3 days ago by tw5hd
Five Stars

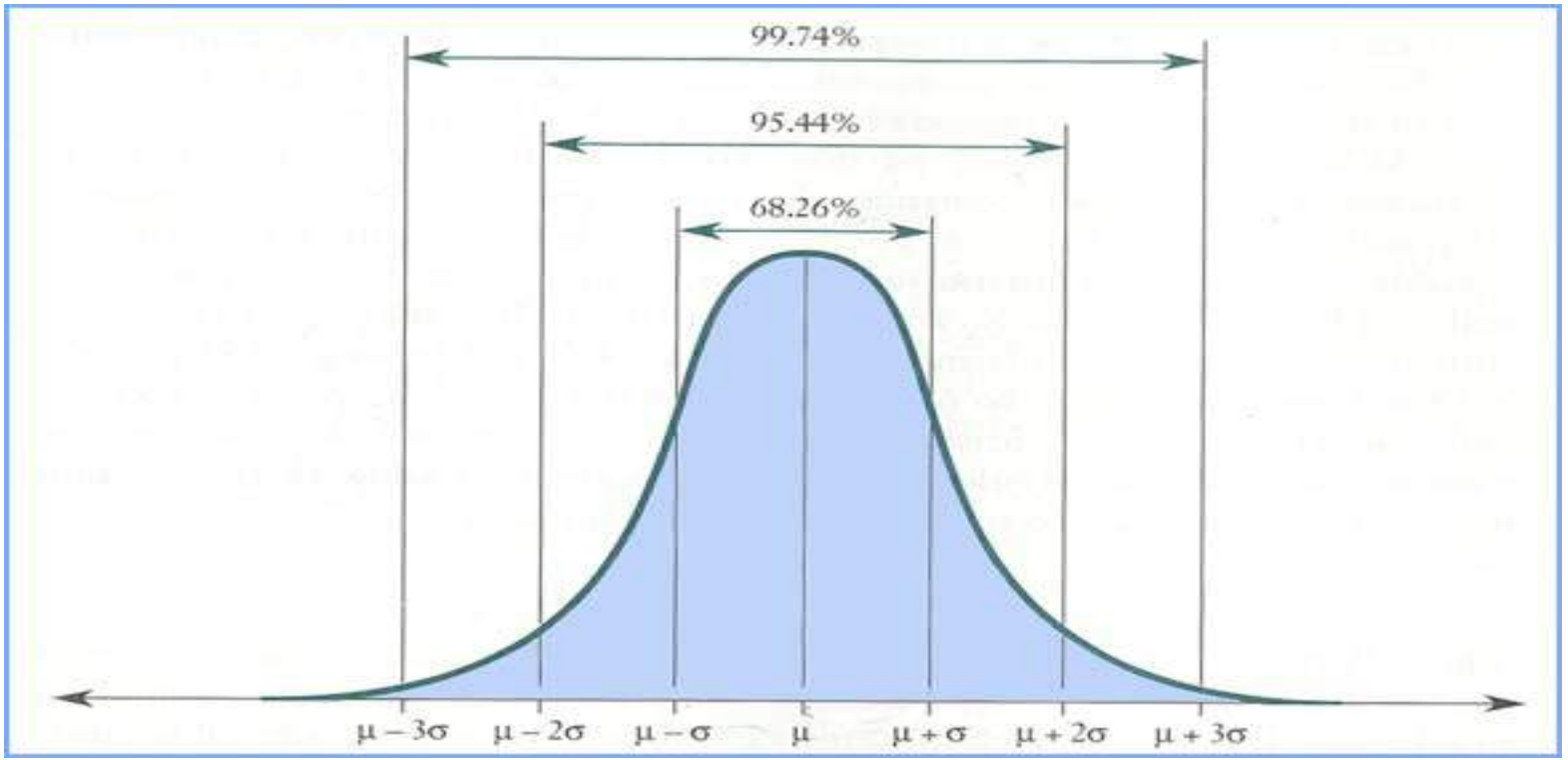

Customer Reviews
(364)
4.3 out of 5 stars
5 star
4 star
3 star
2 star
1 star
See all 364 customer reviews
Popular Discussion Topics
"Image Quality"
"Zoom"
"Ease of Use"
"Video Quality"
"Features"
All Topics
Quality is great and very fine!
Keith Chandler
So the next morning I decided to see how nice the pictures are in good lighting.
Reviewer
I have just started photographing birds, butterflies, etc and the pictures have been great.
Peggy Thornton
The pictures are great!
Great camera for the price, picture quality is good.
L. Maloney
Pictures are beautiful and am learning more about the features everyday.
Karen
Hot topics extracted from review
Topic related snippets from customer review
Most Helpful Customer Reviews
461 of 473 people found the following review helpful
A nifty little camera
By Nemaloda on December 25, 2013
I'm just a point-and-shoot guy, and all I wanted was a good point-and-shoot camera with a powerful zoom lens, and good picture quality for an admittedly undiscerning eye. I got all that and more with this camera. Since there are plenty of "expert" reviews, I'm not going to say anything about the ins-and-outs of the camera. So, for what it's worth, here are my random comments.
1. Besides the 30x optical zoom on the SX510, the other cool feature is the wifi capability. Setting up the wifi was not particularly intuitive, but I eventually got everything working. I read one review from a mac owner that the wifi would not transfer pictures wirelessly from the camera to the the Apple computer. That's not true. All you need to do is download some Canon software (available here: [...] Once you download and install the software, you need to connect your computer to the camera via your home wifi network. Again, this isn't very intuitive, but it's not too difficult either. After the connection is made, you'll be able to transfer photos from the camera to your Apple computer via the "Image Sync" command.


Challenges and Solutions

# Imbalanced Data


Memory
Usage
CPU Usage


Anomaly Detection




$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Feature Vector X $\{X_1, X_2, X_3 \ldots X_n\}$

Given new example $x$, compute $p(x)$:

$$p(x) = \prod_{j=1}^{n} p(x_j; \mu_j, \sigma_j^2) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(x_j - \mu_j)^2}{2\sigma_j^2}\right)$$

Anomaly if $p(x) < \varepsilon$

**Multivariate Gaussian (Normal) distribution**

Parameters $\mu, \Sigma$

$$p(x; \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right)$$

# Anomaly detection vs. Supervised ML

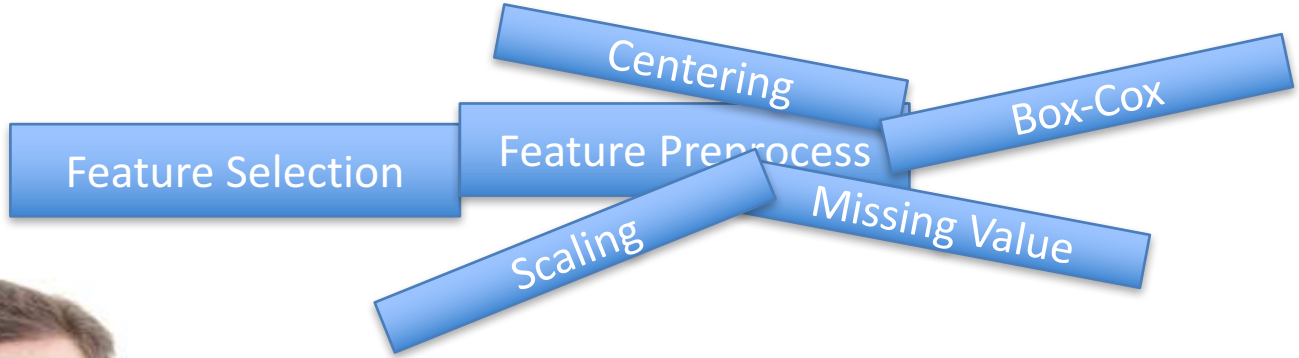
**Anomaly detection**

- Very small number of the anomaly (positive) samples
- Large number of the normal (negative) samples

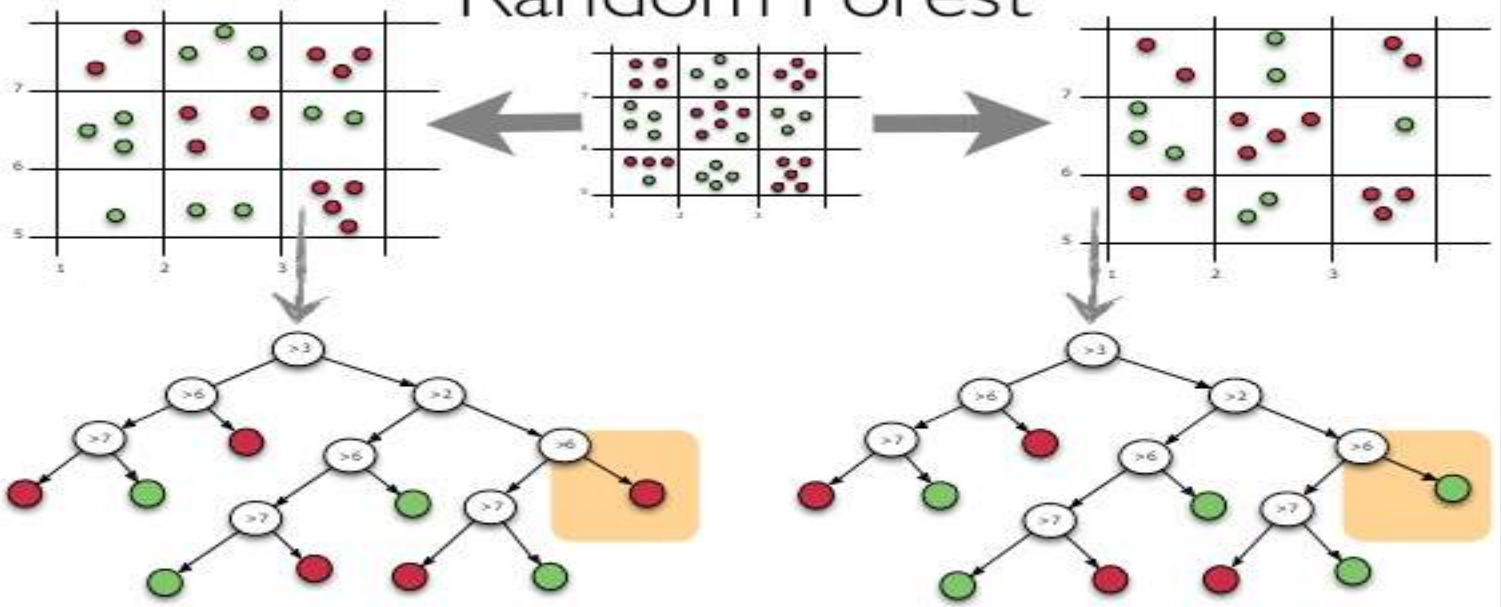
**Supervised ML**

- Large number of the positive and negative samples.

# Feature preprocessing is hard




# Random Forest



- Each tree sees part of the training sets and captures part of the information it contains

# Amazon Content Compliance

amazon

Content Compliance Service

COMPLIANCE

# No enough Labeled Data

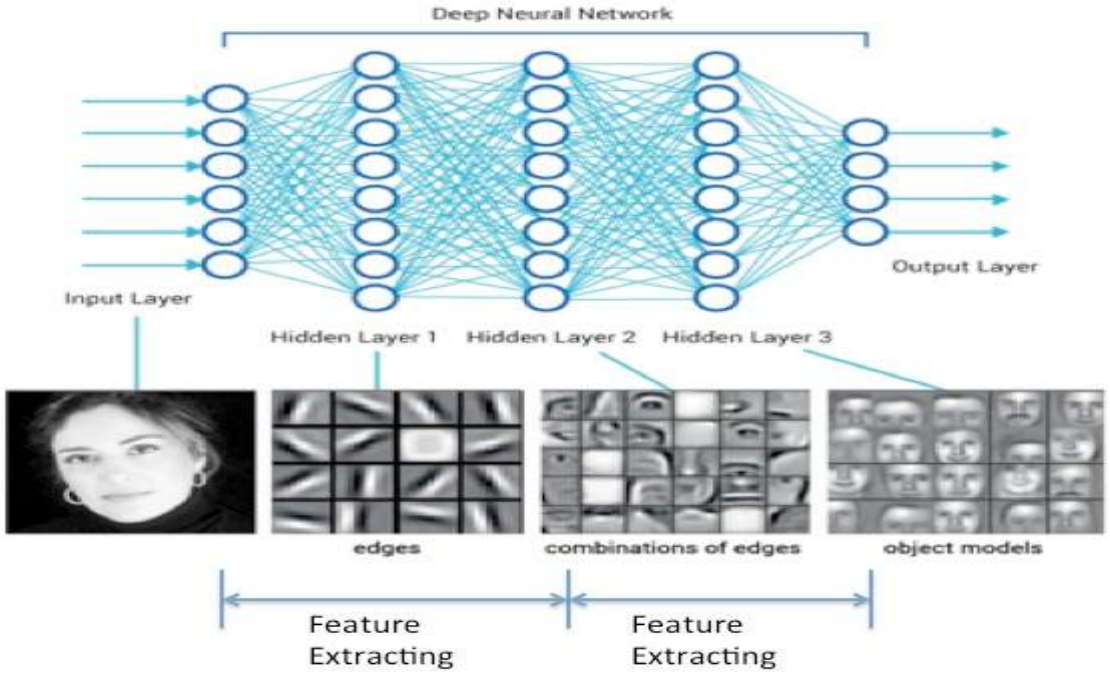
# Image Augmentation

# Image Compliance D

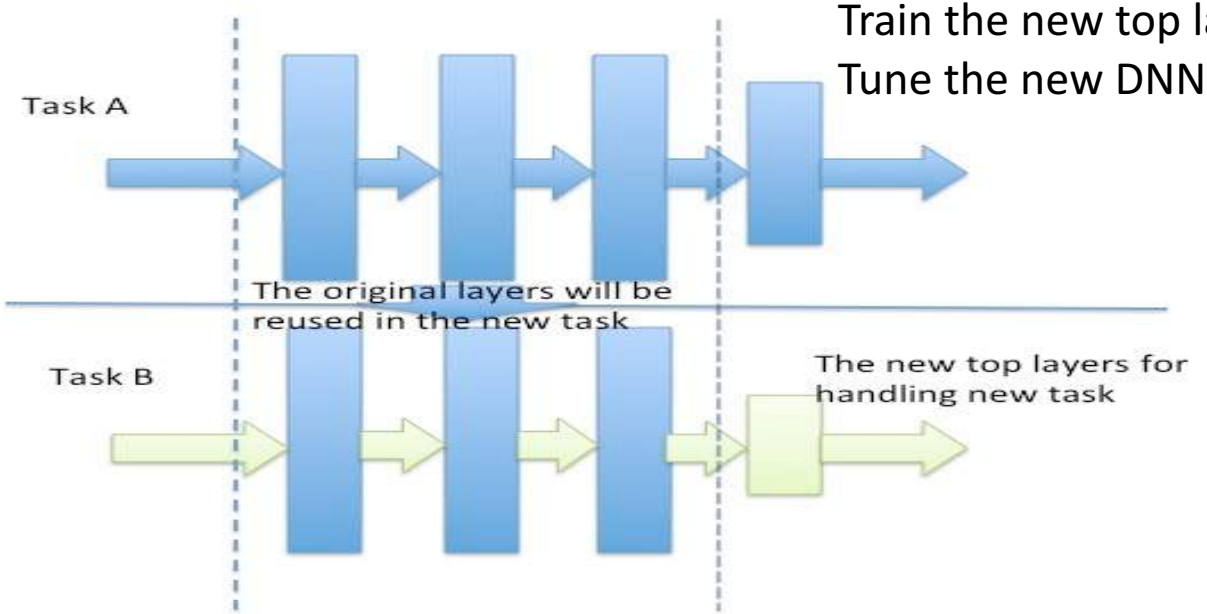
I see you.

# No Bible for constructing DNN

# Transfer Learning



# Transfer Learning




Remove the top layer of the original DNN
Build the new top layers on the original DNN
Train the new top layers
Tune the new DNN finely

# Where do you start from…

- Start from your case
- Start from your data
- Start from the simple

# For More

My Public Account