



2016 杭州·云栖大会  
THE COMPUTING CONFERENCE

云栖社区  
yq.aliyun.com

# Data-Driven Innovation

Tokopedia Use Cases

2016  
The Computing Conference

主办单位:



战略合作伙伴:



扫码观看大会视频

**Leontinus Alpha Edison**  
Co-Founder & COO tokopedia.com



I am the technical co-founder & COO of Tokopedia, an open marketplace that provides people in Indonesia an opportunity to build and grow their online business easily, while also providing better online shopping experience.

Prior to Tokopedia, I has various experience in software development on different kind of industries, including ISP, Insurance, Airlines and several government projects.

Now, in Tokopedia, I am focusing on mixture of technical, operation and people management. I also care about security on all aspect.





Big Data is like teenage sex:  
everyone talks about it, nobody  
really knows how to do it, everyone  
thinks everyone else is doing it, so  
everyone claims they are doing it.

— Dan Ariely —

AZ QUOTES



扫码观看大会视频



# TVC Optimization





## What did we do?

We found that people tends to play with their phone while watching TV.

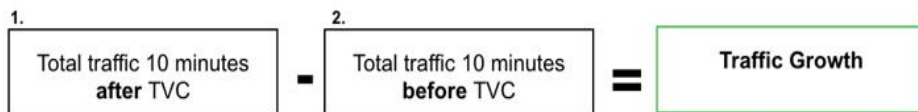
IF we can captured number of people who visits our site, after watching our TVC. Means, our TVC works on that niche of people in that particular program.

## How?

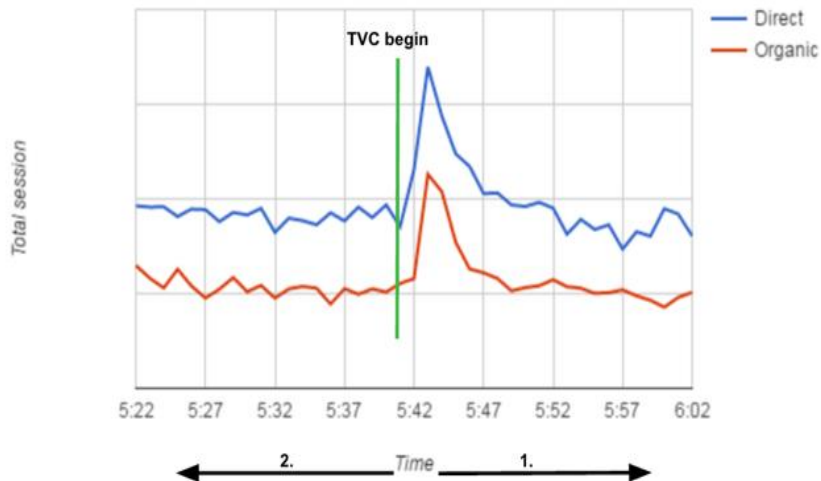
1. Filter traffic only from Direct and Organic.
2. Filter traffic only from Mobile, WAP, Android and iOS.
3. Use Google Trend (to analyse search terms).



## Calculation Method:



### Direct and Organic Combine



## Calculation Method:

		1-Jun.	2-Jun.	3-Jun.	5-Jun.	6-Jun.	7-Jun.	8-Jun.	9-Jun.	10-Jun.	11-Jun.	12-Jun.	13-Jun.	14-Jun.	15-Jun.	16-Jun.	17-Jun.
NET   BREAK OUT	15-18	269	-26	191													
NET   BUKAN SEKEDAR WAYANG	12-15																
NET   CELEBRITY LIPSYNC(R)	9-12	210	119	-460		138	-326	318	381	414			273	332	196	-140	560
NET   ENTERTAINMENT E	9-12	-184	534	-459		-274	-1,824	263	-343	-994			373	686	-167	-210	138
NET   INDONESIA MORNING SHOW	6-9	810	563	121		750	87	797	427	553			365	717	654	242	-172
NET   INI SAHUR	24.4 4.6					798	470		1,043				380	62	236	-217	-127
NET   KELAS INTERNASIONAL SEASON 2	18-22	102	-18,800	117		-1,086	-817	492	595	633			-202	-91	1,605	-299	864
NET   NET.10	9-12												-202	259	464	62	347
NET   NET.12	12-15												1,197	-518	931	298	132
NET   PAGI PAGI	6-9	397	403	732		889	1,009	559	1,045	921			558	893	832	921	905
NET   SARAH SECHAN	12-15	-163															
NET   TENTANG MASA GITU?	18-22																
NET   TENTANG MASA GITU? SEASON 3	18-22	-12	121	-186		920	2,464	3,094	2,239	1,976			1,651	1,562	1,898	1,308	1,397
NET   TENTANG MASA GITU? SEASON 3(R)	9-12					1,166	345	244	780	-16			646	622	639	-139	15





With the same  
budget

We've achieved:

↑ 50% number  
of spot

↑ 30% number  
of TVR

### Other metrics that we try to optimise:

- Duration selection : 15 sec vs 30 sec. 15s cost **40% less** than 30s and drive the same amount of traffic. However, it will affect our brand. We tend to use 30 sec duration for 3-5 days after launching.
- **15%** from total spot on **Primetime** (18:00 - 20:00). It is better to use less frequency on primetime. Even though PT program drive lots of traffic compared to non-PT, the cost on PT tends to be a lot higher. Making it not efficient enough.

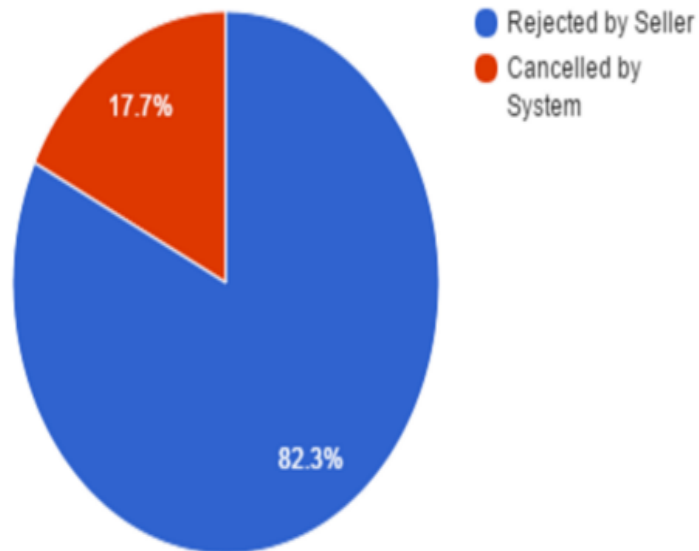




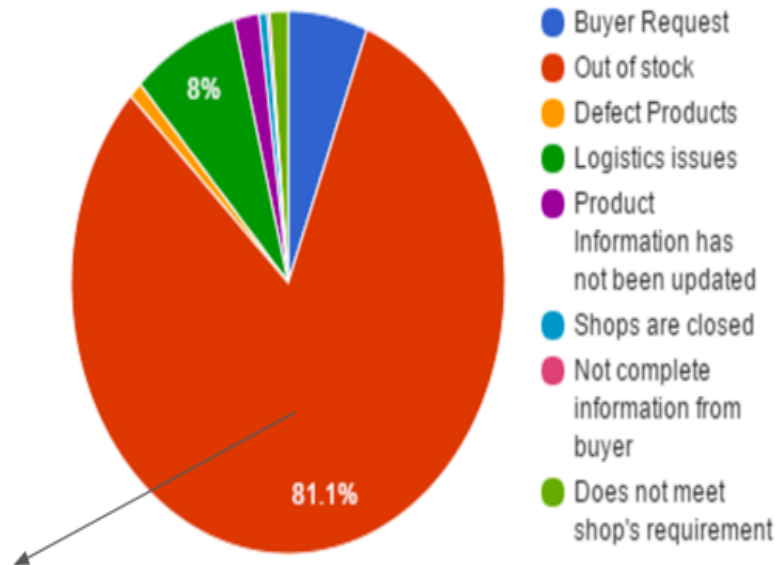
# Cancellation Rate



## Cancellation Overview



## Reason of Cancel Order



80% of the cancellation was due to out of stock

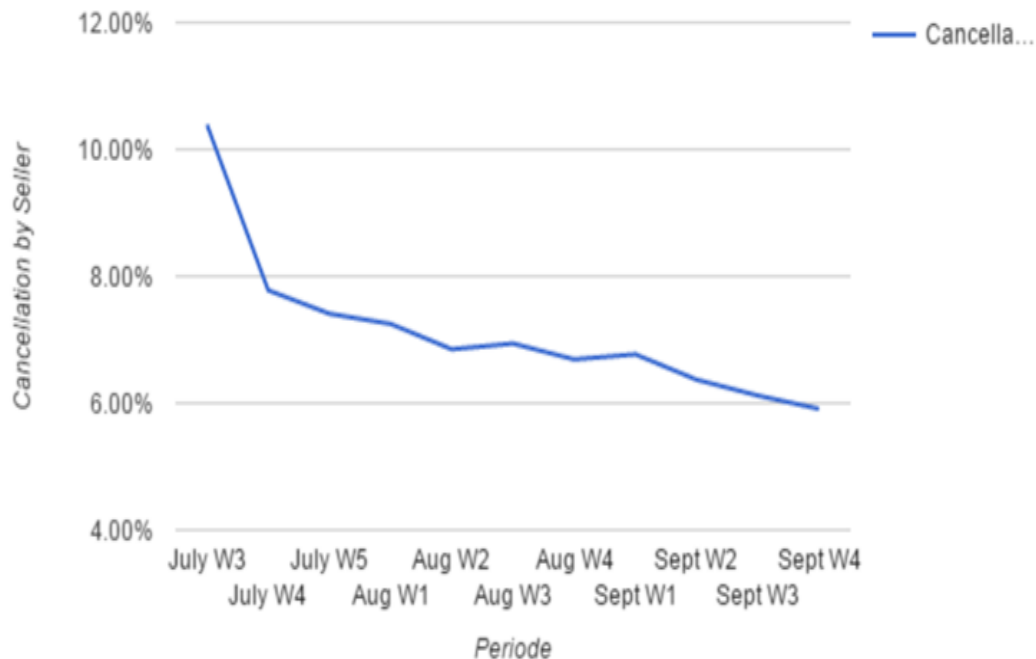


扫码观看大会视频

1. Automated closing the shops which are not log in in 7 days
2. Making close shop schedule so sellers can set the time period when they want to close their shops due to holiday or other personal reasons
3. Making Seller of the Week Program. We gave some appreciation to seller who gave the best performance on certain week. The criteria is based on the shop's success rate, its speed, and also shop reviews by buyers.



Cancellation by Seller vs. Periode



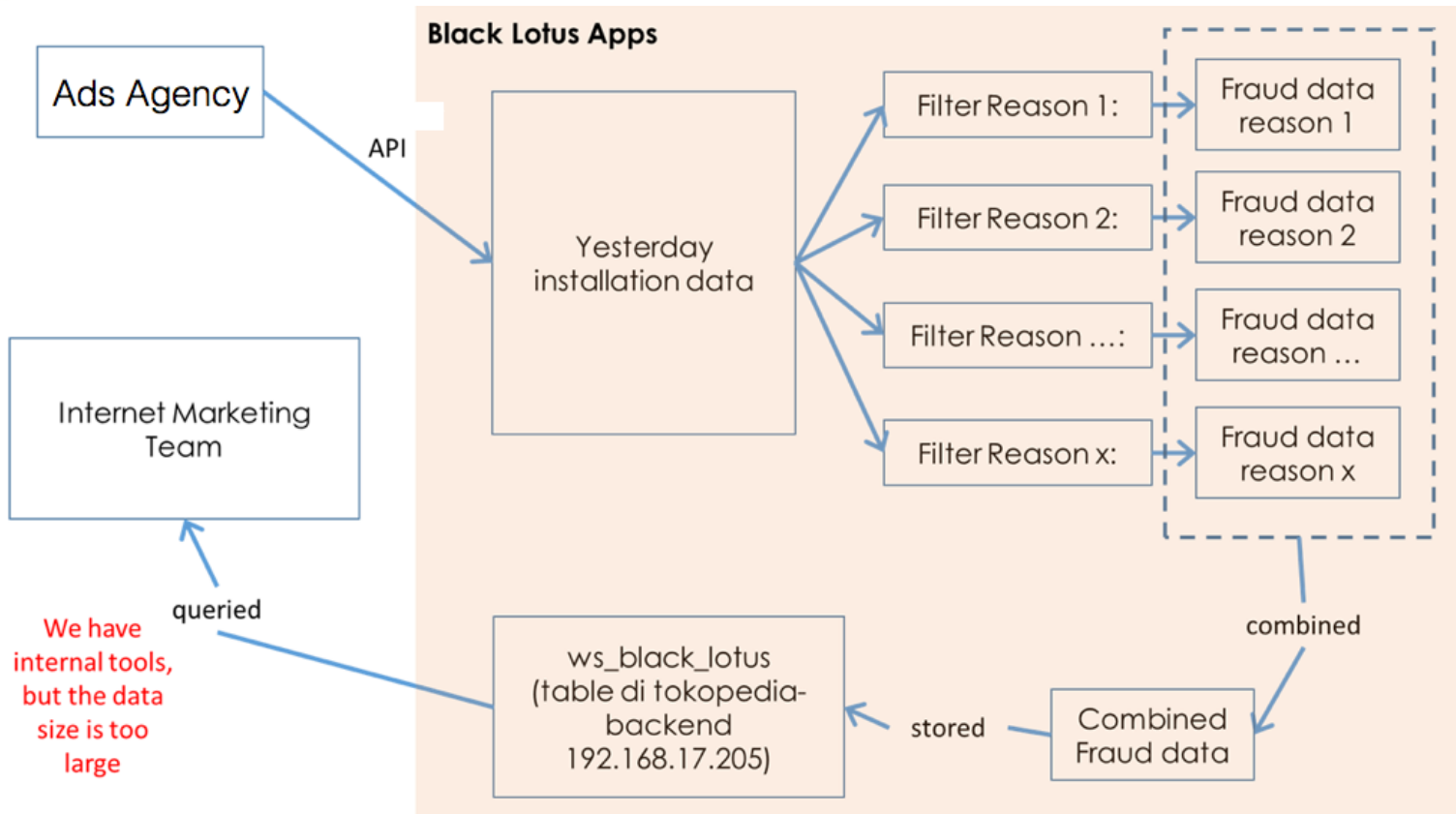
Cancellation rate is showing decreasing trends on each week. Currently stable at 6%, decreased by almost 50% compared to before we done some improvements on it.





# Fraud Install Detector





# The Variables

In order to detect fraud installs, we used several variables which are:

1. Reason 1 => 'Invalid IMEI', Reason 2 => 'Duplicate IMEI', IMEI duplicated in the last 30 days
2. Reason 3 => 'Invalid Android ID',
3. Reason 4 => 'Duplicate Android ID'
4. Reason 5 => 'Emulator',
5. Reason 6 => 'Invalid App. Version',
6. Reason 7 => 'Non Indonesian Operator'
7. Reason 8 => 'Outside Indonesia'
8. Reason 9 => 'Duplicate Advertising ID'





## Result

After socializing this method to all our ad network partners, we have saved this amount of money (we are not paying for fraud installs):

Partner	Saved Amount
A	\$ 25,000
B	\$ 9,000
C	\$ 4,000





# Machine Learning @Tokopedia



## Fighting transactional fraud at Tokopedia

Fraudulent transactions are handled **MANUALLY** using a backend system

Only using **Rule Engine** and later on we counting on **MANUAL INSPECTION**

However, number of cases overwhelms fraud team, it leaves a lot of unhandled cases.



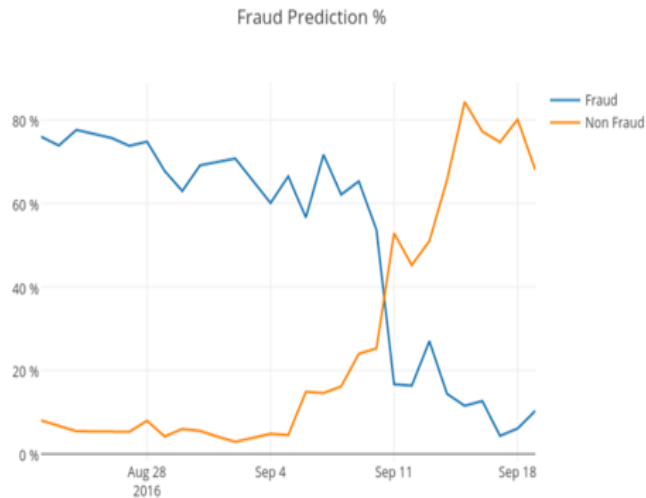
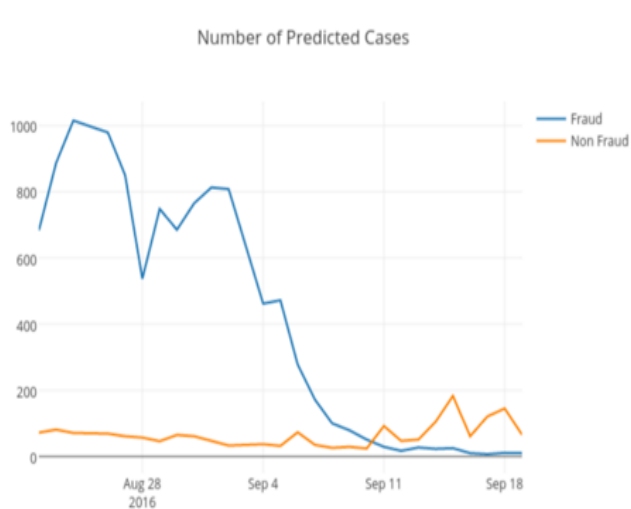
## Solutions?

Using **machine learning** to discover fraudulent pattern rather than manually checking similar cases over and over again.

We use machine learning model alongside the existing rule engine.



## .. result



These graphs display data from 22 Aug' 2016 to 19 Sep' 2016. The fraud cases captured are **high** late august until early september because Tokopedia is holding a **free shipment promo**. During this period, **16000 cases in total** were captured and our machine learning model got **78% accuracy** and **6% error rate**.





**Transaction are not the only place for fraud**

**What about MESSAGING?**



## challenges..

- Most of the time fraudster will post messages that is claiming their victims as swindler, this psychologically will make the victim curious or angry about the claim, and open the link that is given to them, for example:

"sorry gan,kok toko agan masuk berita  
penipuan ?  
saya liat disini  
www\*tokped\*pe\*hu  
ubah bintang jadi titik,soalnya gk bisa  
share link di sini  
thanks"

"sorry sir, why does your shop got listed  
in this fraudster list announcement?  
I saw it here  
www\*tokped\*pe\*hu  
substitute stars to dots,because we  
can't share link here  
thanks"



## more examples..

"BE CAREFUL

TOKOPEDIA IS TAKING ITS SELLER'S SALDO  
WITHOUT YOU KNOWING IT.

I AM THE VICTIM,

I SENT COMPLAIN AND TOKOPEDIA CAN  
NOT DENY IT.

HERE IS THE CHRONOLOGY ON HOW I MAKE  
CHAOS IN TOKOPEDIA'S OFFICE

CHECK IT OUT HERE

www\*tokped\*pe\*hu

substitute stars with dots, tokped blocked

my news ,they might be embarrassed !

wkakwka<laughing intensively>"

"An info for tokopedia's member

Just for today's transaction in tokopedia, you can get  
bonus pulsa in tokopedia

Here is the steps, click the web link given bellow,  
and put your login email that you use to login to  
tokopedia, then follows the next steps in the web

<http://pulsaalloperator.890m.com/index.php>"





Phishing, spamming, &  
masking

Promo abuse



Credit card advance



20 The  
16 Computing  
Conference  
**THANKS**

