Automatic decision development 2016 – 2017

MICRO MONEY

Four key areas for development

Development of data model

connecting new sources of customer data

Development of automated strategies

Anti-fraud, scoring, blacklists, minimum requirements

Introduction and piloting of automated strategies

Implementation of scoring checks, anti-fraud rules and other automatic customer checks, pilots

Monitoring

Development of quality monitoring systems for scoring models and auto-tests, as well as the correctness of its work

MICRO MONEY

Four key areas for development

Development of data model

connecting new sources of customer data

Development of automated strategies

Anti-fraud, scoring, blacklists, minimum requirements

Introduction and piloting of automated strategies

Implementation of scoring checks, anti-fraud rules and other automatic customer checks, pilots

Monitoring

Development of quality monitoring systems for scoring models and auto-tests, as well as the correctness of its work

MICRO MONEY

Main analytical tasks 2016

- VN Scoring
- ID Scoring + New TS Process
- Calculation of profit
- AF Rules VN
- AF Rules ID & PH
- New process PH Site
- PH Scoring

VN Application Scoring

- VN Scoring
- ID Scoring + New TS Process
- Calculation of profit
- AF Rules VN
- AF Rules ID & PH
- New process PH Site
- PH Scoring

MICRO MONEY

VN Application Scoring

Problems

- First trying to create scoring model
- Big problems with data (excel, master file and other)
- Little historical period
- Non stable risk strategy in training period
- Low quality of application data
- Only basic application fields
- Low level of understanding, how to implement model in Terrasoft
- Model was developed for DSA channel

Results

- We implemented this model 5.5 months
- Model didn’t work properly because we have tried to use it on other channel
- We couldn't change model in TS quickly
- We didn't have normal point in TS for managing all our features
- We understand how what we need to do

MICRO MONEY

ID Application Scoring

- VN Scoring
- ID Scoring + New TS Process
- Calculation of profit
- AF Rules VN
- AF Rules ID & PH
- New process PH Site
- PH Scoring

Result of implementation

- We spent 3 month for implementation
- We realized not only scoring model it TS, but new scoring process, which gave us:
  - Possibility to change model and model parameters very fast
  - Possibility to manage all our features as Trusting Social, Scoring, BL process from one point
  - We created new strategy “skip pv” without verification procedure

Result of model working

- We can say now that model work properly on production and quality is stable
- We see that our strategies which are connected with scoring model work properly too
  - we have now lower BR than it could be
  - We reduce the costs
  - We increase conversation in skip pv segment
- Next week we launch strategy pilot which can help us to reduce more costs and to reject more bad clients without verification

MICRO MONEY

ID Application Scoring

	SCORE INTERVAL	APP	AGR	APP - > AGR	AGR EXP	BR EXP Real	BR EXP Predicted
SKIP PV	[0.99; 2.00)	2039	1618	0.79	281	0.18	0.25
PROCEED / SKIP PV	[0.65; 0.99)	5015	2026	0.40	320	0.24	0.32
PROCEED	[0.27; 0.65)	8018	3009	0.38	491	0.27	0.39
PROCEED	[-0.14; 0.27)	5626	1790	0.32	282	0.33	0.47
REJECT	[-2.00; -0.14)	2695	0	0.00	-	-	-

MICRO MONEY

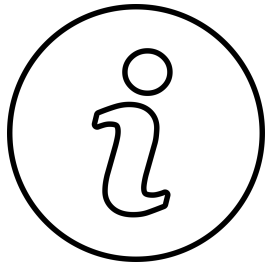
Calculation of external data profit

or Trusting Social and Iamreal, two projects – one fate

- VN Scoring
- ID Scoring + New TS Process
- Calculation of profit
- AF Rules VN
- AF Rules ID & PH
- New process PH Site
- PH Scoring

Facts:

- We had two same projects
  - Iamreal – integration with FB for ID
  - Trusting social – integration with telecom operator in VN
- Request cost was around 2-3
- Integration through website
- We do request firstly for all long applications and than for all accepted applications



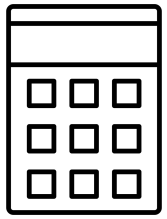
MICRO MONEY

Calculation of external data profit

Results:

- Both two projects give us the same quality – around GINI = 10
- For better understanding is it good or bad result was created model in excel which can help us to create analysis “what if” for such projects

Calculation result:



- This model give us that with average amount = 100$ and with our Bad Rate ~ 30% each GINI = 10 give us around 0.6 $ per each application with request, for situation when the decision is depends only on scoring model
- We also have sales funnel, and, because of funnel, we need to compare this 0.6$ with request cost price * 5
- So such external data sources is to expensive for us
- We understand that we need to focus on free and very cheep data sources

MICRO MONEY

Four key areas for development

Development of data model

connecting new sources of customer data

Development of automated strategies

Anti-fraud, scoring, blacklists, minimum requirements

Introduction and piloting of automated strategies

Implementation of scoring checks, anti-fraud rules and other automatic customer checks, pilots

Monitoring

Development of quality monitoring systems for scoring models and auto-tests, as well as the correctness of its work

MICRO MONEY

Source of data about client

Data type	Description	VN	ID	PH	MY
Application data	Clean data from client	prod	prod	prod	prod
Iovation	Device id, device information, antifraud cheks	prod	prod	prod	prod
Facebook 1	Autorization fact + email, name, link, photo	development (IT side)	development (IT side)	prod	development (IT side)
Facebook 2	Friends info, additional personal info, timeline				
Historical web data	Application data, Black Lists	prod (AF, BL)	prod (AF, BL)	prod (AF, BL)	
Social Vector 1	30 universal cheks	prod	prod	prod	prod
Social Vector 2	Additional local sites checks	development (RD side)			
How client fills application	Total time, time for each page, count of correction				
Historical Terrasoft Data	Data from Terrasoft about clients, defaults and other information				
Geolocation	IP Geolocation , Google API, GPS coordinates				
IP	Client IP	prod	prod	prod	prod
User Agent	Device type, operation system type, browser type	prod	prod	prod	prod
UTM	Marketing data	prod	prod	prod	prod

- For China data sources are differ because of external factors
- For new countries we are trying to realize these data sources in first time after launch
- Also we are planning to create with IT minimal data kit

MICRO MONEY

Source of data about client

2016–2017

Data type	Description	VN	ID	PH	MY
Application data	Clean data from client	prod	prod	prod	prod
Iovation	Device id, device information, antifraud cheks	prod	prod	prod	prod
Facebook 1	Autorization fact + email, name, link, photo	development (IT side)	development (IT side)	prod	development (IT side)
Facebook 2	Friends info, additional personal info, timeline				
Historical web data	Application data, Black Lists	prod (AF, BL)	prod (AF, BL)	prod (AF, BL)	
Social Vector 1	30 universal cheks	prod	prod	prod	prod
Social Vector 2	Additional local sites checks	development (RD side)			
How client fills application	Total time, time for each page, count of correction				
Historical Terrasoft Data	Data from Terrasoft about clients, defaults and other information				
Geolocation	IP Geolocation , Google API, GPS coordinates				
IP	Client IP	prod	prod	prod	prod
User Agent	Device type, operation system type, browser type	prod	prod	prod	prod
UTM	Marketing data	prod	prod	prod	prod

- For China data sources are differ because of external factors
- For new countries we are trying to realize these data sources in first time after launch
- Also we are planning to create with IT minimal data kit

MICRO MONEY

Source of data about client

Data type	Description
Application data	Clean data from client
Iovation	Device id, device information, antifraud cheks
Facebook 1	Autorization fact + email, name, link, photo
Facebook 2	Friends info, additional personal info, timeline
Historical web data	Application data, Black Lists
Social Vector 1	30 universal cheks
Social Vector 2	Additional local sites checks
How client fills application	Total time, time for each page, count of correction
Historical Terrasoft Data	Data from Terrasoft about clients, defaults and other information
Geolocation	IP Geolocation , Google API, GPS coordinates
IP	Client IP
User Agent	Device type, operation system type, browser type
UTM	Marketing data

- For China data sources are differ because of external factors
- For new countries we are trying to realize these data sources in first time after launch
- Also we are planning to create with IT minimal data kit

iovation

What is it? How does it work?

- The iovation module is installed on the site or in the app
- The module collects information about the device used by the client, the device is assigned a unique identifier if it is not in the external database iovation; If the device is contained in the database, the frequency of the institution of applications from this device is analyzed.
- Iovation provides device id, device information, calculated own anti-fraud rules

What and where is realized

- Implemented on all prod sites VN, ID, PH, MY
- implemented on dev CH

What is planned to be realized

- Implement prod CH
- Implement in new countries send default data to iovation

How is used

- In anti-fraud rules of the form: more than one application for 21 days from one device, and with different client data for PH, VN, ID - work at the pilot stage
- In the scoring card by PH

How is planned to be used

In anti-fraud inspections and in scoring models of all countries

Source of data about client

Data type	Description
Application data	Clean data from client
Iovation	Device id, device information, antifraud cheks
Facebook 1	Autorization fact + email, name, link, photo
Facebook 2	Friends info, additional personal info, timeline
Historical web data	Application data, Black Lists
Social Vector 1	30 universal cheks
Social Vector 2	Additional local sites checks
How client fills application	Total time, time for each page, count of correction
Historical Terrasoft Data	Data from Terrasoft about clients, defaults and other information
Geolocation	IP Geolocation , Google API, GPS coordinates
IP	Client IP
User Agent	Device type, operation system type, browser type
UTM	Marketing data

- For China data sources are differ because of external factors
- For new countries we are trying to realize these data sources in first time after launch
- Also we are planning to create with IT minimal data kit

Facebook

How does it work?

- Receive data:
  - Email
  - Name
  - Page link
  - Gender
  - Facebook_id

What and where is realized

- Implemented on all prod sites PH
- In the process of implementation on MY, ID, VN

What is planned to be realized

- To expand the volume of FB data

How is used

- Accumulation of statistics

How is planned to be used

- In anti-fraud inspections and in scoring models of all countries
- mandatory authorization via FB on one of the sites
- collection

Source of data about client

Data type	Description
Application data	Clean data from client
Iovation	Device id, device information, antifraud cheks
Facebook 1	Autorization fact + email, name, link, photo
Facebook 2	Friends info, additional personal info, timeline
Historical web data	Application data, Black Lists
Social Vector 1	30 universal cheks
Social Vector 2	Additional local sites checks
How client fills application	Total time, time for each page, count of correction
Historical Terrasoft Data	Data from Terrasoft about clients, defaults and other information
Geolocation	IP Geolocation , Google API, GPS coordinates
IP	Client IP
User Agent	Device type, operation system type, browser type
UTM	Marketing data

- For China data sources are differ because of external factors
- For new countries we are trying to realize these data sources in first time after launch
- Also we are planning to create with IT minimal data kit

Historical data

How does it work?

- The information available on the site is used to find applications in the past that are associated with the application being processed by one of the parameters

What and where is realized

- implemented in the form of AF rules iovation on prod sites PH, ID, VN

What is planned to be realized

- New types of anti-fraud rules and other rules related to social ties

How is planned to be used

- As scoring variables
- For rejection rules

Source of data about client

Data type	Description
Application data	Clean data from client
Iovation	Device id, device information, antifraud cheks
Facebook 1	Autorization fact + email, name, link, photo
Facebook 2	Friends info, additional personal info, timeline
Historical web data	Application data, Black Lists
Social Vector 1	30 universal cheks
Social Vector 2	Additional local sites checks
How client fills application	Total time, time for each page, count of correction
Historical Terrasoft Data	Data from Terrasoft about clients, defaults and other information
Geolocation	IP Geolocation , Google API, GPS coordinates
IP	Client IP
User Agent	Device type, operation system type, browser type
UTM	Marketing data

- For China data sources are differ because of external factors
- For new countries we are trying to realize these data sources in first time after launch
- Also we are planning to create with IT minimal data kit

Social Vector

How does it work?

- get information on the list

What and where is realized

- Implemented on all prod sites PH
- In the process of implementation on MY, ID, VN

What is planned to be realized

- To expand the volume of FB data

How is used

- Accumulation of statistics

How is planned to be used

- In anti-fraud inspections and in scoring models of all countries
- mandatory authorization via FB on one of the sites
- collection

Source of data about client

Data type	Description
Application data	Clean data from client
Iovation	Device id, device information, antifraud cheks
Facebook 1	Autorization fact + email, name, link, photo
Facebook 2	Friends info, additional personal info, timeline
Historical web data	Application data, Black Lists
Social Vector 1	30 universal cheks
Social Vector 2	Additional local sites checks
How client fills application	Total time, time for each page, count of correction
Historical Terrasoft Data	Data from Terrasoft about clients, defaults and other information
Geolocation	IP Geolocation , Google API, GPS coordinates
IP	Client IP
User Agent	Device type, operation system type, browser type
UTM	Marketing data

- For China data sources are differ because of external factors
- For new countries we are trying to realize these data sources in first time after launch
- Also we are planning to create with IT minimal data kit

UTM

How does it work?

- Information about the marketing source

What and where is realized

- Implemented on all prod sites PH, MY, ID, VN

What is planned to be realized

- Together with the marketing department to fix the rules for filling UTM tags
- to collect detailed information about the launched companies

How is planned to be used

- As scoring variables
- Analyze the quality of marketing segments by recurrence / default

Source of data about client

Data type	Description
Application data	Clean data from client
Iovation	Device id, device information, antifraud cheks
Facebook 1	Autorization fact + email, name, link, photo
Facebook 2	Friends info, additional personal info, timeline
Historical web data	Application data, Black Lists
Social Vector 1	30 universal cheks
Social Vector 2	Additional local sites checks
How client fills application	Total time, time for each page, count of correction
Historical Terrasoft Data	Data from Terrasoft about clients, defaults and other information
Geolocation	IP Geolocation , Google API, GPS coordinates
IP	Client IP
User Agent	Device type, operation system type, browser type
UTM	Marketing data

- For China data sources are differ because of external factors
- For new countries we are trying to realize these data sources in first time after launch
- Also we are planning to create with IT minimal data kit

New data source

How client fills application

- Parameterize the features of filling the application by the client
  - Time to fill each field
  - Number of fixes for each field
  - Time between fields filling
  - Other features
- Use in scoring models and anti-fraud rules

Historical Terrasoft Data

- Integrate the site with Terrasof in terms of receiving additional data on the client
  - Receive data about delays of this client
  - receive data about delays of related persons
- Use in scoring models, anti-fraud rules and behavioral scoring
-

Geolocation

- Integrate with Google service to retrieve geolocation data using Google API Geolocation
- Use in anti-fraud rules

Source of data about client

2016–2017

Data type	Description	VN	ID	PH	MY
Application data	Clean data from client	prod	prod	prod	prod
Iovation	Device id, device information, antifraud cheks	prod	prod	prod	prod
Facebook 1	Autorization fact + email, name, link, photo	development (IT side)	development (IT side)	prod	development (IT side)
Facebook 2	Friends info, additional personal info, timeline				
Historical web data	Application data, Black Lists	prod (AF, BL)	prod (AF, BL)	prod (AF, BL)	
Social Vector 1	30 universal cheks	prod	prod	prod	prod
Social Vector 2	Additional local sites checks	development (RD side)			
How client fills application	Total time, time for each page, count of correction				
Historical Terrasoft Data	Data from Terrasoft about clients, defaults and other information				
Geolocation	IP Geolocation , Google API, GPS coordinates				
IP	Client IP	prod	prod	prod	prod
User Agent	Device type, operation system type, browser type	prod	prod	prod	prod
UTM	Marketing data	prod	prod	prod	prod

- For China data sources are differ because of external factors
- For new countries we are trying to realize these data sources in first time after launch
- Also we are planning to create with IT minimal data kit

MICRO MONEY

Four key areas for development

Development of data model

connecting new sources of customer data

Development of automated strategies

Anti-fraud, scoring, blacklists, minimum requirements

Introduction and piloting of automated strategies

Implementation of scoring checks, anti-fraud rules and other automatic customer checks, pilots

Monitoring

Development of quality monitoring systems for scoring models and auto-tests, as well as the correctness of its work

MICRO MONEY

New Anti-Fraud rules for VN

- VN Scoring
- ID Scoring + New TS Process
- Calculation of profit
- **AF Rules VN**
- AF Rules ID & PH
- New process PH Site
- PH Scoring

Have been realized

Rule Type: applications for which we find applications for last 2 weeks with same field:	Bad rate if AFrule = 0	Bad rate if AFrule = 1	Hit Rate
iovation_device_alias	0.22	0.33	9.8%
IP	0.23	0.31	7.1%
IP without last block	0.21	0.26	47.9%
mobile_phone	0.23	0.26	7.8%
document_number	0.23	0.26	8.5%

In plan

Rule Type: applications for wich we find applications with same device id for two last 2 weeks but with differ field:	Bad rate if AFrule = 0	Bad rate if AFrule = 1	Hit Rate	Information value
date_of_birth	0.22	0.50	3.7%	0.07
document_number	0.23	0.47	3.6%	0.05
mobile_phone	0.23	0.44	4.2%	0.05
company_phone	0.22	0.41	6.5%	0.05
guarantor_phone	0.22	0.39	7.1%	0.05
IP without last block	0.23	0.36	5.5%	0.03
email	0.23	0.33	2.4%	0.01

New Anti-Fraud rules for PH and ID

- VN Scoring
- ID Scoring + New TS Process
- Calculation of profit
- AF Rules VN
- AF Rules ID & PH
- New process PH Site
- PH Scoring

Have been realized

Rule Type: applications for wich we find applications with same device id for two last 3 weeks but with differ field:	Timelag	Bad rate if AFrule = 0	Bad rate if AFrule = 1	Hit Rate
date_of_birth	21	0.36	0.46	4.1%
mobile_phone	21	0.36	0.49	4.2%
company_phone	21	0.36	0.49	5.5%
guarantor_phone	21	0.36	0.48	5.5%
account_number	21	0.36	0.50	4.3%
email	21	0.36	0.45	4.3%
document_number	21	0.36	0.46	4.9%
IP4	21	0.36	0.44	3.8%
IP3	21	0.36	0.49	3.2%
IP2	21	0.36	0.47	1.2%

MICRO MONEY

New Anti-Fraud rules for PH and ID

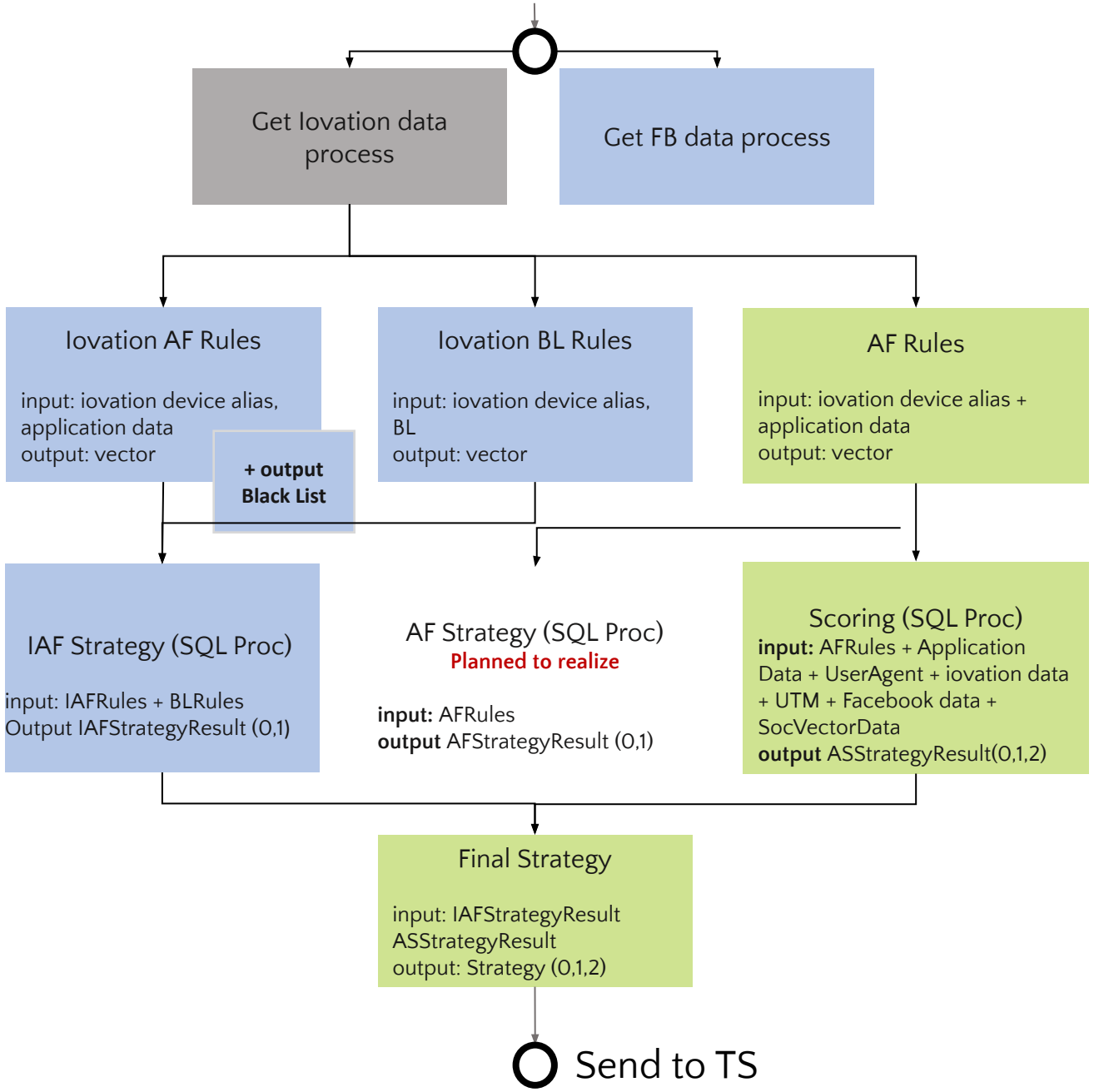
In plan

Rule Type: applications for wich we find applications for last 3 weeks with same field:	Bad rate if AFrule = 0	Bad rate if AFrule = 1	Hit Rate
mobile_phone	0.38	0.46	0.14
email	0.38	0.44	0.13
guarantor_phone	0.38	0.49	0.09
document_number	0.38	0.47	0.10
account_number	0.38	0.47	0.12
ip	0.37	0.42	0.34
full_name	0.38	0.49	0.10
company_phone	0.38	0.42	0.18
living_home_phone	0.38	0.44	0.13
company_name	0.37	0.43	0.34
iovation_device_alias	0.39	0.44	0.03

MICRO MONEY

PH Scoring

- VN Scoring
- ID Scoring + New TS Process
- Calculation of profit
- AF Rules VN
- AF Rules ID & PH
- New process PH Site
- PH Scoring



PH Scoring

- VN Scoring
- ID Scoring + New TS Process
- Calculation of profit
- AF Rules VN
- AF Rules ID & PH
- New process PH Site
- PH Scoring

Result of implementation

- We spent 2-3 week for implementation
- We realized not only scoring model on WEB, but new scoring process, which gave us:
  - Possibility to change model and model parameters very fast
  - Possibility to manage all our features as from one point

Result of modeling

- We can say that model work properly on production and quality is stable
- We can get such results:
  - Reduce Bad Rate by 10% (43 -> 33)
  - Reduce by 40% our vinificators' load
  - Save AR on current level

MICRO MONEY

PH Scoring

Period 2016w35 - 2016w50

score interval		count apps	count agreement	conversion
[min	max]			
0.00	0.30	**8522**	**1523**	**18%**
0.30	0.35	**8522**	**1595**	**19%**
0.35	0.40	**8523**	**1336**	**16%**
0.4	0.46	**8522**	**1345**	**16%**
0.46	1.00	**8523**	**1292**	**15%**

Period 2016w35 - 2016w50

score interval		count of agrements	% of agreements	count of defaults	BR	New strategy
[min	max]					
0.00	0.30	213	0.19	50	23%	low level of defaults, can be without "pv verification" which can increase the conversion twice
0.31	0.35	210	0.19	77	37%	clients for normal procced strategy
0.35	0.40	238	0.22	100	42%	
0.41	0.46	188	0.17	95	51%	
0.46	1.00	251	0.23	159	63%	very high level of defaults, should be rejected

MICRO MONEY

Four key areas for development

Development of data model

connecting new sources of customer data

Development of automated strategies

Anti-fraud, scoring, blacklists, minimum requirements

Introduction and piloting of automated strategies

Implementation of scoring checks, anti-fraud rules and other automatic customer checks, pilots

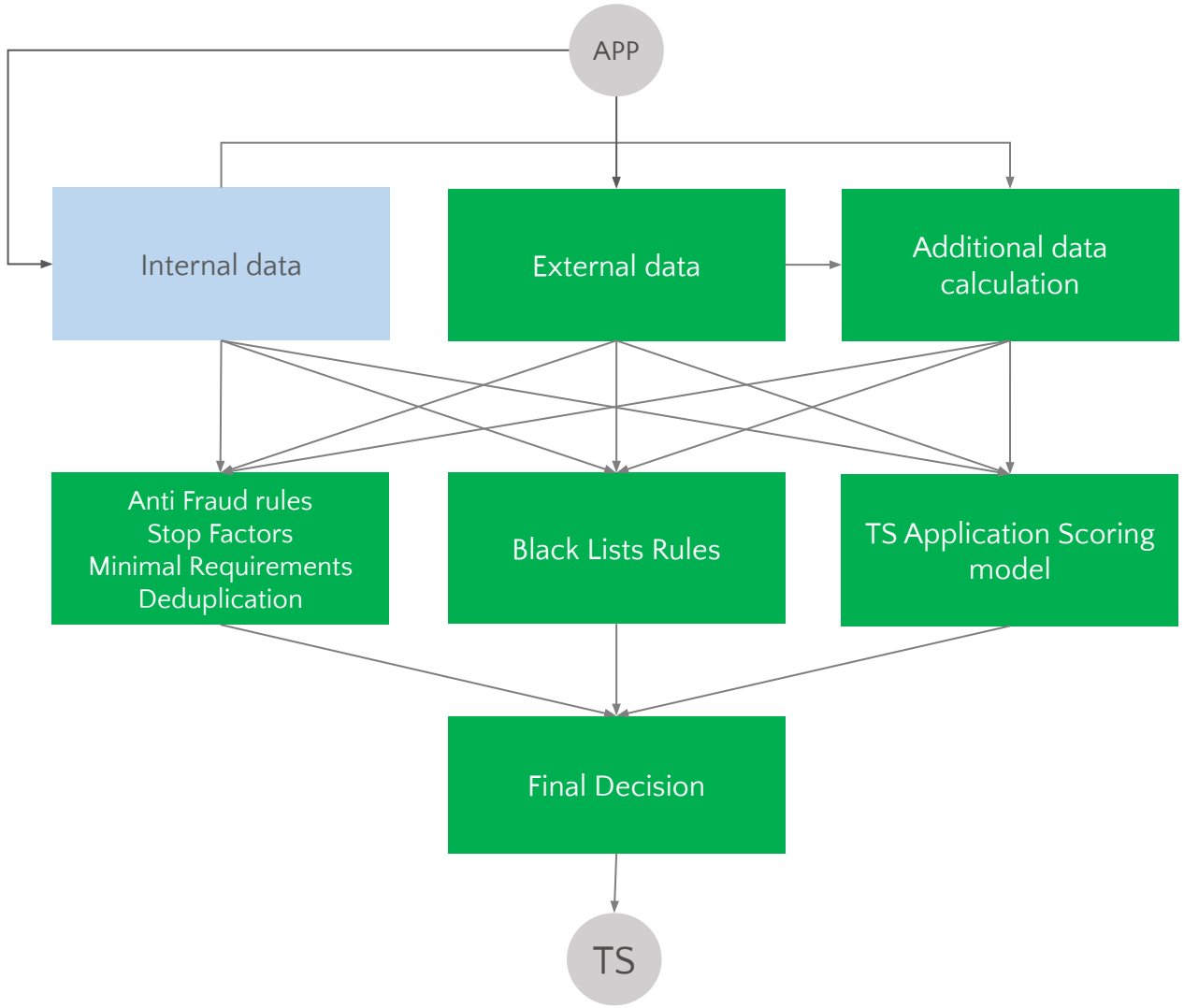
Monitoring

Development of quality monitoring systems for scoring models and auto-tests, as well as the correctness of its work

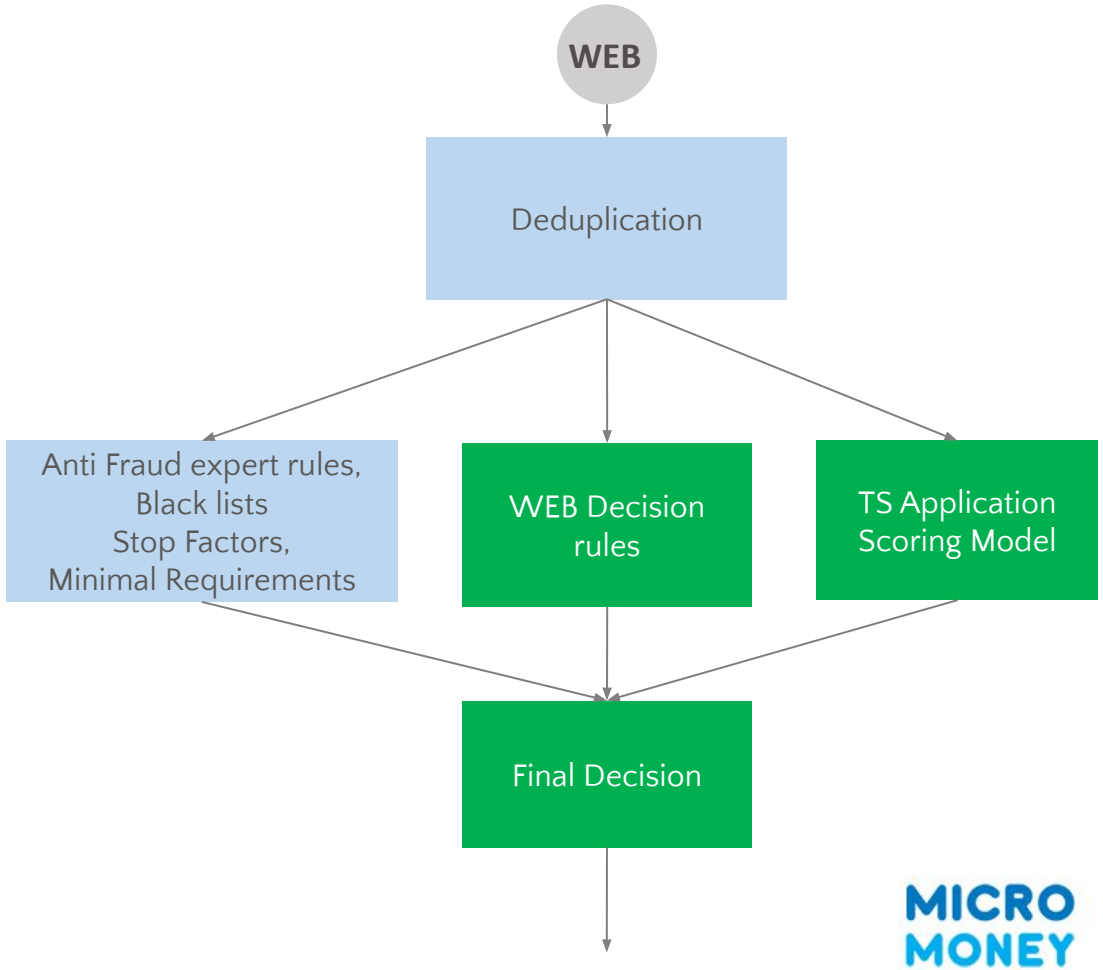
MICRO MONEY

Analytical module

WEB Analytical Module



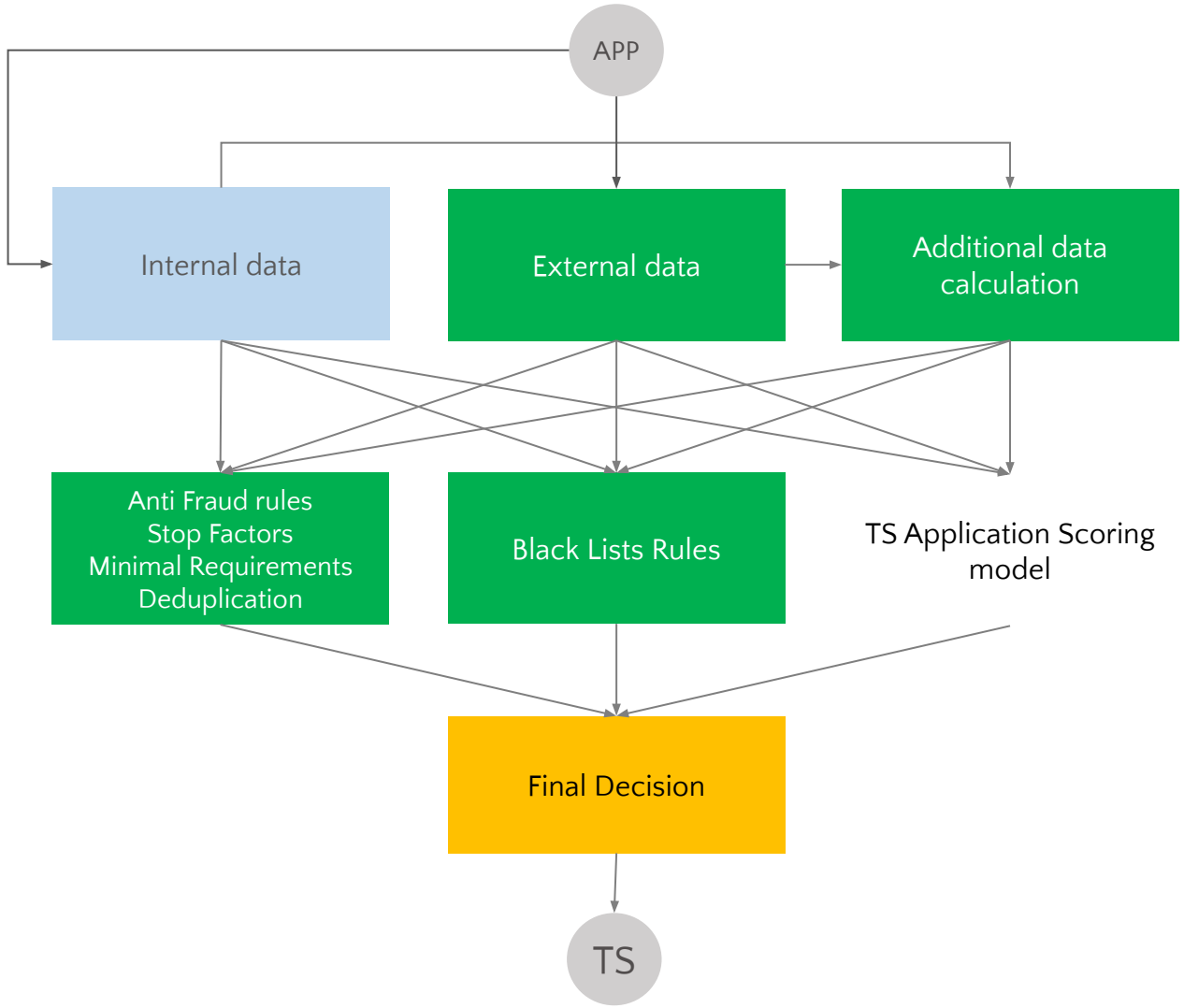
TS Analytical Module



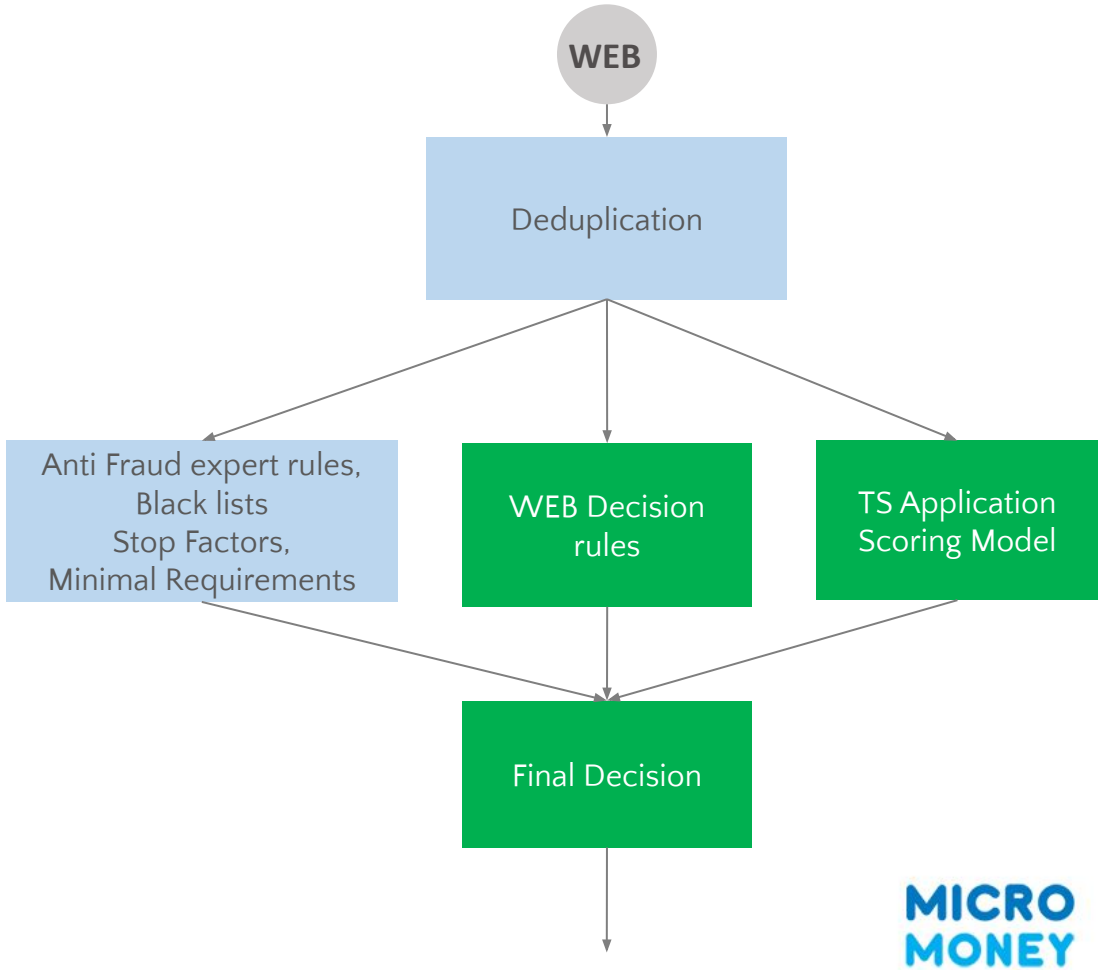
MICRO MONEY

Analytical module ID

WEB Analytical Module



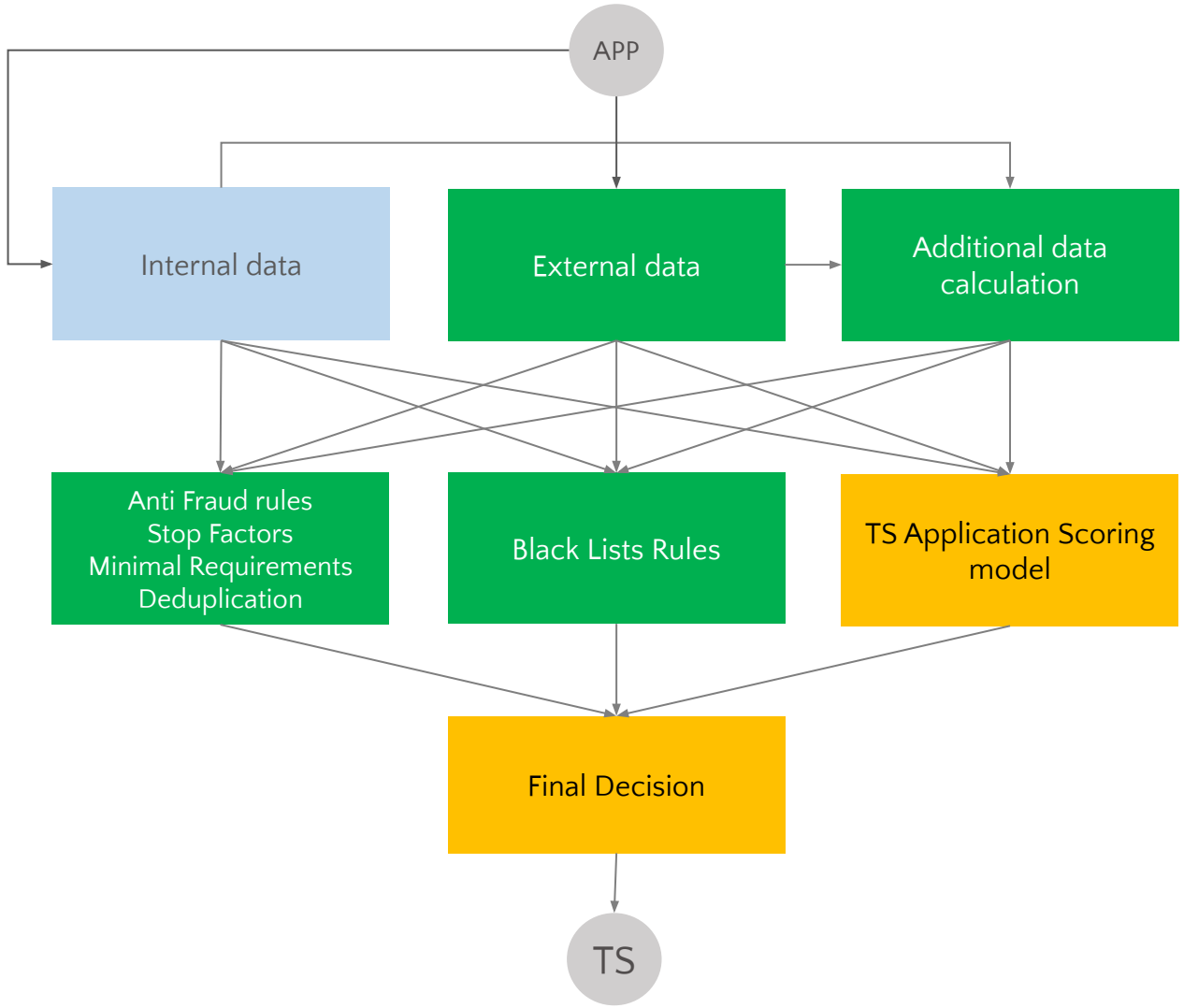
TS Analytical Module



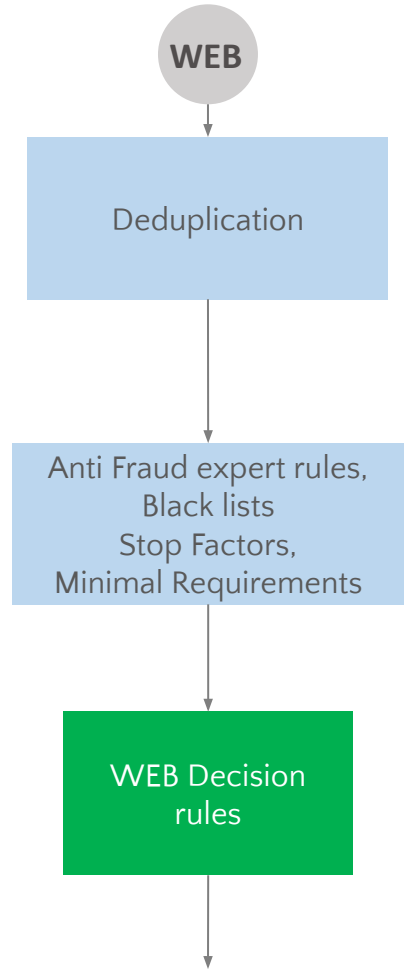
MICRO MONEY

Analytical module PH

WEB Analytical Module



TS Analytical Module



MICRO MONEY

Four key areas for development

Development of data model

connecting new sources of customer data

Development of automated strategies

Anti-fraud, scoring, blacklists, minimum requirements

Introduction and piloting of automated strategies

Implementation of scoring checks, anti-fraud rules and other automatic customer checks, pilots

Monitoring

Development of quality monitoring systems for scoring models and auto-tests, as well as the correctness of its work

MICRO MONEY

Conclusions

Aims:

- We do not want just to implement some analytics, we want to create **analytical system** for each country
  - which is consist from simple independent blocks with different functions
  - which can give us possibilities do any changes as fast as possible
  - with all free and cheep data sources, which we find
- We want to do the same system for CH, MY and for new countries
- Next year we also want to focus on repeat sales to create the same process for them
- And we are planning to create good monitoring system for it

MICRO MONEY

Conclusions

Aims:

- We do not want just to implement some analytics, we want to create **analytical system** for each country
  - which is consist from simple independent blocks with different functions
  - which can give us possibilities do any changes as fast as possible
  - with all free and cheep data sources, which we find
- We want to do the same system for CH, MY and for new countries
- Next year we also want to focus on repeat sales to create the same process for them
- And we are planning to create good monitoring system for it

MICRO MONEY

Source of data about client. Annex

Applicant Info

Result	R
Trackingnumber	404097317130247683
Endblackbox	bLqMq9zYqKdX1uPVlwsfS8PUJtpEH5G3C5CZ5K+OvREd
Reason	Accounts Per Device_All Time
Device_alias	811575546835164156
Device_browser_lang	EN-US
Device_browser_type	MOBILE
Device_browser_ver...	14A456
Device_cookie_enab...	1
Device_firstseen	2016-11-16T16:38:38+00:00
Device_flash_enabled	0
Device_flash_installed	0
Device_flash_version	
Device_js_enabled	1
Device_new	0
Device_os	CPU IPHONE OS 10_0_2 LIKE MAC OS X
Device_screen	667X375
Device_type	IPHONE
Device_tz	-480
IPaddress	203.87.156.123
Realipaddress	203.87.156.123
ipaddress_isp	SMART BROADBAND
ipaddress_loc_city	
ipaddress_loc_country	PHILIPPINES
ipaddress_loc_count...	PH
ipaddress_loc_lat	14.595505
ipaddress_loc_lng	120.97211
ipaddress_loc_region	
ipaddress_org	SMART BROADBAND
realipaddress_isp	SMART BROADBAND
realipaddress_loc_city	
realipaddress_loc_c...	PHILIPPINES
realipaddress_loc_c...	PH
realipaddress_loc_lat	14.595505
realipaddress_loc_lng	120.97211
realipaddress_loc_re...	
realipaddress_org	SMART BROADBAND
iovation_realipaddre...	subscriber
device_browser_cha...	
device_browser_con...	EN-US
device_flash_storag...	
device_device_trust...	
device_ipaddress_pr...	
device_realipaddres...	
Rule reasons	Accounts Per Device_All Time,Accounts Per Device_All Time,Accounts Created Per Device_Monthly
Rule scores	-1,-1
Rule types	Accounts Per Device,Accounts Created Per Device
Ruleset_rulesmatched	2
Rule set score	-2

	Logged in	Not logged in
Youtube	✔	
Gmail	✔	
Facebook		✔
Skype		✔
Airbnb		✔
Dropbox		✔
Google Plus		✔
Twitter		✔
Paypal		✔
Amazon.com		✔
Indeed		✔
Blogger		✔
Expedia		✔
Khan Academy		✔
Academia.edu		✔
EdX		✔
Github		✔
Flickr		✔
Spotify		✔
Tumblr		✔
Foursquare		✔
Reddit		✔
Pinterest		✔
Steam		✔
Bitbucket		✔

MICRO MONEY