

# Experience in using AI and Big Data at Trusting Social to change the Finance Industry

Mai Hai Thanh

Lead Data Scientist, Trusting Social

2019-11

Do you have a bank account?

Have you ever sent or received money?

Have you ever borrowed money from a financial institution?

If yes, finance is already an important part of your daily life.

Bank account, send/receive money transactions,  
borrowing money process...

Are they safe, fast, easy, and convenient?

Not sure, it depends. Fintech will play an important role.

# Trusting Social - An AI Fintech company

## MISSION

Advancing data science and technology to deliver **financial access** for all

**200+** members

Working hard everyday with **1.5PB** of data  
in billions of billions of records

**60+**

Data Scientists,  
Research engineers,  
Data analytic engineers

to create better products for the above mission

# Trusting Social - An AI Fintech company

## MISSION

Advancing data science and technology to deliver **financial access** for all

Do you know that billions of people still do not have financial access because financial institutions cannot assess their credit risks?

Cannot borrow even 20M VND for improving their lives (health, education, business, travel...)

### First success that we are proud of:

**500M+**

Scored Consumers

**\$600M+**

Disbursed Loans

**1B+ by 2020**

# Trusting Social solutions

## Credit Insight



## E-KYC



## Lead Generation



## Lending Platforms



## Credit score

- Predict the probability that a borrower will pay back his/her debt

## Lead generation

- Find people who want to borrow money or open credit cards

## eKYC (Know Your Customer)

- Make sure that you are [who you say you are]
  - Face recognition, face retrieval

TRUST INSIGHT

# Credit score

---

## Credit Score

- Predict the probability that a borrower will pay back his/her debt
  - Very important for banks and financial institutions
  - Credit score
    - > custom evaluation process
    - > financial access decision
    - > interest rate, loan amount, tenure
    - > loss, profit

# Our products help lenders reduce ~50% of loan losses

Across 100+ consumer financing portfolios in Vietnam, Indonesia and India



**63%**

Reduction in Credit Loss



**48%**

Reduction in Credit Loss



**50%**

Reduction in Credit Loss



**63%**

Reduction in Credit Loss

GINI\* of .50

GINI\* of .42

GINI\* of .64

GINI\* of .76

GINI is the standard measure of accuracy in credit scoring  
GINI =  $2 \times \text{ROCAUC} - 1$

## Telco data is one of the main data sources for our scoring models

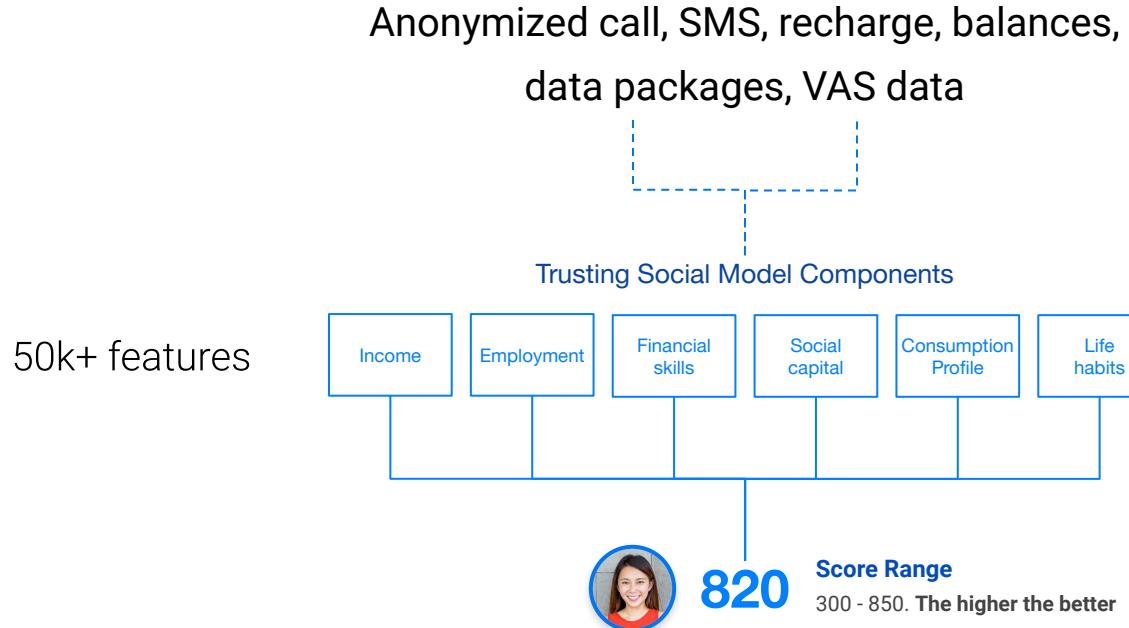
- Call, SMS, recharge, balances, data packages, VAS histories
  - All are anonymized
- Advantages
  - Great coverage (a major part of the whole population)
  - Good density
- Challenges
  - Very large scales (PBs of data, 10+ billions of rows per day)
  - Not really human readable, hard to analyze and extract insights
  - Noise

# Telco data

sub	contact	call_time	call_duration	sub_imei	call_direction
4804150255503569679	1339765896486592202	20170518232109		11	352049075322550
2084368117628033730	1854693525159154669	20170518232047		38	357572064134710
2602643891195640943	6410044028384707014	20170518231847		159	355253053995240
-8757467228756585742	-4879630390135123088	20170518232105		26	863660031735520
-7230572333357253088	4854405124613406877	20170518231921		153	863195033995960
7667159449759858436	6132939848853171710	20170518232150		25	862646039630450
-3742368084637748235	584598293138804427	20170518232105		72	869769020654100
-2325794995460161467	-7215786075539507436	20170518232159		20	862651038535310
540747165654862090	-254987573649354387	20170518231541		407	354390067802910
-7000900723990919856	2459754691345755725	20170518232117		95	355423075840930
-7805869792159353889	-7027659564921962173	20170518232256		13	356850070132810
-3970250341884060254	-2416992529900082725	20170518232131		104	352212080158030
-6532793627135447994	-6059056852056697698	20170518231745		342	354448060825020
1965791373525624058	7951789323882974040	20170518230732		958	013334002138130
-7310834866323406411	1511184906935057953	20170518232325		7	355388071564000
-5615792583501681453	8711891512480282720	20170518232317		17	355385071203010
-1187957499667698058	1754622475937024017	20170518231811		325	351853075647830
-2640423561644015226	-4237199024801984056	20170518232342		20	359300067430850
4804150255503569679	-3830314811254042505	20170518232345		23	352049075322550
-7868560825803839047	5840534556390925697	20170518232241		87	357224070766740

only showing top 20 rows

# TS Credit Score



# TS Credit Score

## Validation results

Model performs extremely well with GINI reaching 0.8

Score range 300 - 850

Input → Output

📞 Phone no. ⚙️ Credit score

👤 National ID ⚠️ Fraud warning signals

📋 Past inquiries

# Credit score backtest - Example of Lender A

Lender A:  
30K applications;  
mix of CD loans  
and Cross-sale  
cash Loan;  
average bad rate:  
8.3%

TS Score Range	% of Population	Actual Default Rate	Cumulative Default Rate	
[300-499)	10%	37.17%	44.99%	Reduce 63% credit loss
[499-547)	10%	15.17%	62.90%	
[547-585)	10%	9.04%	73.97%	
[585-615)	10%	6.86%	82.01%	
[615-642)	10%	4.72%	87.61%	
[642-668)	10%	4.04%	92.50%	
[668-693)	10%	2.39%	95.34%	
[693-720)	10%	1.97%	97.75%	
[720-753)	10%	1.22%	99.15%	
[753-850)	10%	0.82%	100%	

Define Leadgen customer segments

Enable risk-based pricing

Fast track, good product to attract good customers & improve customer experience

## The best practice of using TS scores

- Combining TS credit score and TS fraud score with financial institution's internal score and other data sources gives even more powerful solutions
  - Application information, national credit bureau data, internal data
  - Up to 20% increase in GINI -> significantly reduce loss, improve profit, better customer experience (less document, faster process)

## The best practice of using TS scores

- Example 1: Consumer durable and two-wheeler products
  - Combining TS credit score with internal score helps a lender to offer approval within **10 minutes** for 100% Consumer Durable & Two-wheeler applications and control credit risk
- Example 2: Cash loan & Credit card
  - Combining TS credit score with internal data helps another lender to cross-sell Cash loans & Credit cards to existing customers **without a gap time** to observe performance as usual and have better credit quality
  - Combining TS credit score with application data and browsing data helps another lender to offer pre-approval of cash loans for online customers within **15 minutes**



TRUST INSIGHT

# Lead generation

---

## Lead generation - Finding customers

- Lending is one of the most important businesses of most financial institutions (FIs)
- But never easy for any FI
  - Data is sparse & low-coverage
  - Customer acquisition costs of both offline & online channels are high
- We help FIs to find & approach the right people at the right time at a low cost

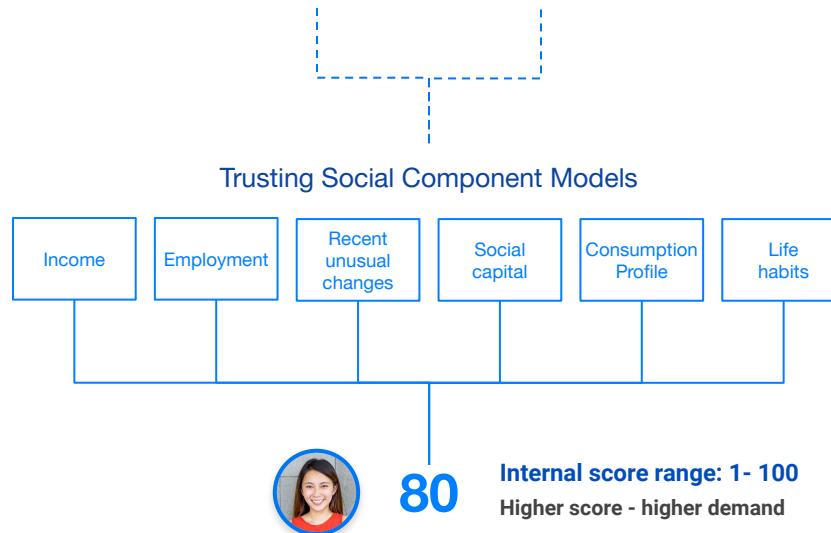
## Lead generation

- The data science problem: find the probability that a person wants to borrow money or open credit card at a given time point

# Lead generation - Propensity To Borrow score

Similar inputs we have for credit scores - looking at different viewpoints

Anonymised call, SMS, recharge, balances,  
data packages, VAS data



# Lead generation - Propensity To Borrow score

## Validation results

Model performs well with **GINI 0.6**

Top 10% have upto **10x demand higher** than average

-> Lower cost, higher profit, better customer experience

TRUST VISION

## E-KYC Selfie ID Matching

---

Get ready for a frictionless onboarding process

A smaller, abstract geometric graphic located at the bottom right of the slide, mirroring the style of the one at the top left. It consists of a dense cluster of light-grey lines and triangles.

# Is the current user the same as the face in the ID card?

Smart lenders and telcos use our E-KYC Selfie ID Matching to:

- ✓ Quickly verify a loan applicant and speed up the on-boarding process (paperless)
- ✓ Optimize lending operation by reducing human errors and operating cost
- ✓ Fast & scalable solution, easy integration



# Selfie ID Matching

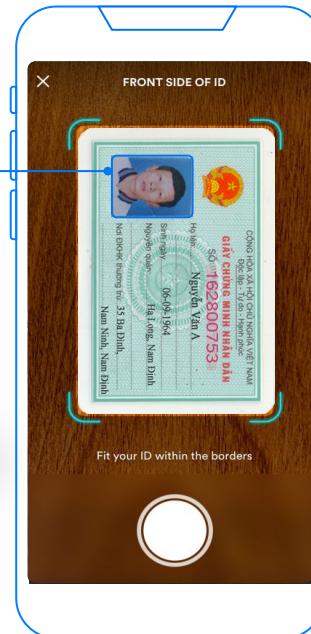


Very  
similar

## Similarity score

## Matched

Very  
Dissimilar



40x

Faster

10%

More accurate

than an average  
human in a face  
matching contest!

TRUST VISION

# E-KYC Identity Fraud Detection

---

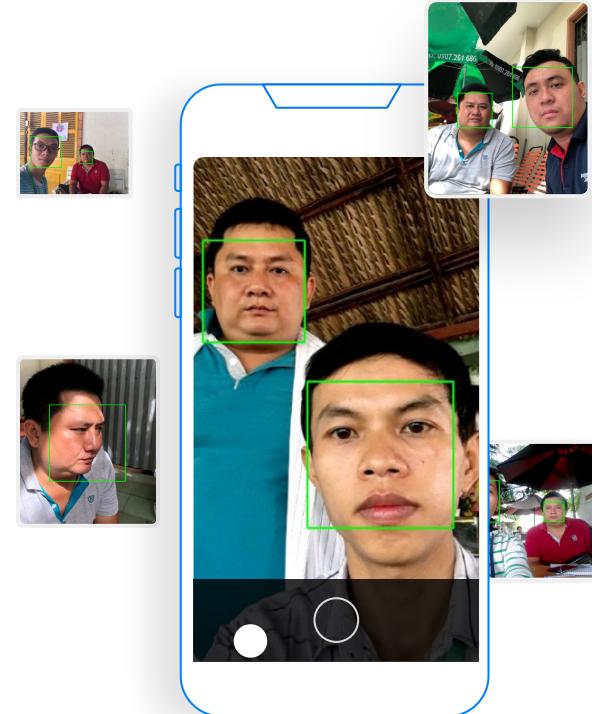
Identity solution for every enterprise

# Effectively prevent fraudulent acts of using multiple fake “identities”

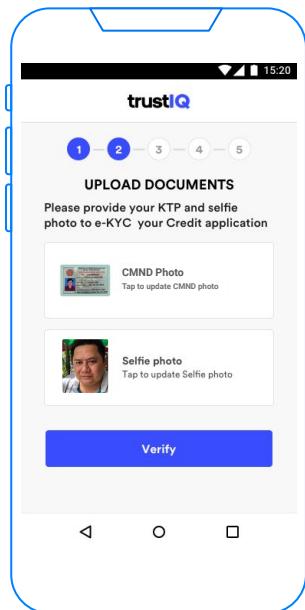
What if the same person takes different (fake) national IDs to apply for multiple loans? Fraud Detection with advanced facial recognition technology is an answer for this complex issue.

## Lenders and telcos use our E-KYC Fraud Detection because:

- ✓ Our face recognition AI accurately identifies similar faces across databases
- ✓ Actively prevent fraudulent from the beginning
- ✓ Provide early warning signs and during loan application processes



# How Fraud Detection works



- Criminal databases
- Existing customers databases
- Government tax, voter, business registration databases

**Fraud Detection report**

**95% Fraudulent act**

Profile Picture	Name	Status	Document Preview
	TRẦN THANH PHÚ	Disbursed	
	NGUYỄN NGỌC TUẤN	Disbursed	
	LÊ MINH MÃN	Disbursed	
	NGUYỄN VĂN ANH	Disbursed	

TRUST VISION

## ID Reader - OCR

---

Saving millions of hours of manual work



# Smartly detect, read and extract national ID in various formats

Our machine learning model has exceeded human in recognizing some of the poorest conditioned documents

## Smart lenders and telcos use our E-KYC ID Reader OCR

- ✓ Verify “liveness” of the ID to prevent fraud
- ✓ Auto extract data and populate application form - enable a paperless consumer lending process
- ✓ Optimize lending operation by reducing human key-in errors and operating cost



**98%**

accuracy on reading ID no. from Vietnamese National ID Card (1999 format)

# Thank you

---

**Trusting Social**  
8th Floor, Havana Tower,  
District 1  
Hochiminh City, Vietnam

[www.trustingsocial.com](http://www.trustingsocial.com)

