trustingsocial

FINANCIAL INCLUSION FOR ONE BILLION UNBANKED PEOPLE

Short Biography

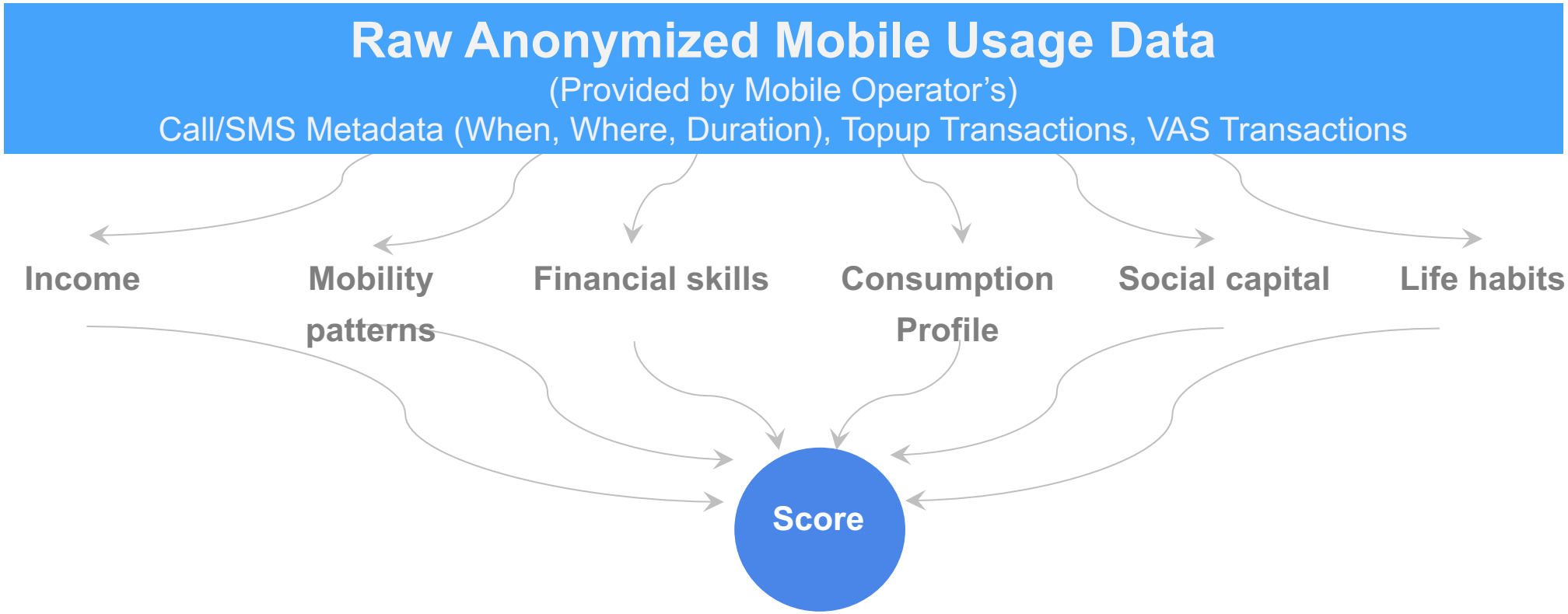
- Education:
  - Bachelor of Engineering, Bach Khoa University, 2004
  - Master in Computer Science, University of Texas at Austin, 2007
  - PhD in Computer Science, University of Texas at Austin, 2011
- Work:
  - 2011-2014: Computer Scientist, SRI International
  - 2014-2015: Research Scientist, JVN Institute
  - 2015-Present: Chief Scientist and CTO, Trusting Social

FINANCIAL EXCLUSION PROBLEM IN DELOPING COUNTRIES

Banks avoid lending to the unbanked consumers

because it is hard to assess their credit risk

TRUSTING SOCIAL: TELCO-BASED CREDIT SCORE



Data Volume & Data Cleaning

- Data volume:
  - 400M records/day → 12B records/month
  - 1TB data/month
  - 2-year of data in our Hadoop production cluster
- Data cleaning:
  - Remove ill-formatted records
  - Remove duplicated records

Supervised Learning Problems

- Credit scoring:
  - Goal: Predict the probability of default of a customer
  - Binary classification problem:
    - Input: Good/Bad loans from banks and mobile usage data of customers who made loan before they made loans
    - Output: A model for predicting the probability of default of a user based on his/her mobile usage data

Supervised Learning Problems (cont.)

- Target marketing:
  - Goal: Predict the probability of interest of a customer to a product
  - Binary classification problem:
    - Input: List of interested/uninterested customers from a random campaign along with their mobile usage data before receiving the promotional messages
    - Output: A model for predicting the probability of default of a user based on his/her mobile usage data

Supervised Learning Problems (cont.)

- Other supervised learning problems:
  - Income prediction
  - Employment prediction
  - Age-range prediction
  - Gender prediction
  - Customer identity

Unsupervised Learning Problems

- Mobility Analysis
  - Identify different mobility pattern based on customers' location data (location of the cell tower when customers make/receive call)
- Graph Analysis on Telco's Social Graph
  - Community Detection:
    - Group SIMs of a call center into one-group
  - Node Embedding:
    - Embed each node into a common subspace so that nodes that have similar connectivity will be close in that subspace

Thank you

We are hiring

https://trustingsocial.com/jobs