

AI and technology: Applications in educational assessment, recruitment and last-mile healthcare

Shashank Srikant
shash@mit.edu

Result of work done at



Varun Aggarwal

research.aspiringminds.com



Nishant Ojha

sevasetu.org

31 July 2017, MSR Bangalore



Information Asymmetry

31 July 2017, MSR Bangalore



Information Asymmetry



Technology to the rescue

Many million of graduates



I need a job.
Am I employable?

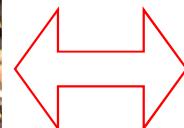
How do I signal my
employability to companies?

Millions of companies



How do we sift through
million of candidates?

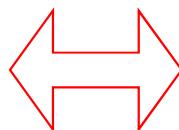
How do we hire those who
would succeed in the job?



Many million of graduates



Millions of companies



Can technology rescue this problem?

31 July 2017, MSR Bangalore

Many million of graduates



Millions of companies



Can technology rescue this problem?

New age assessments

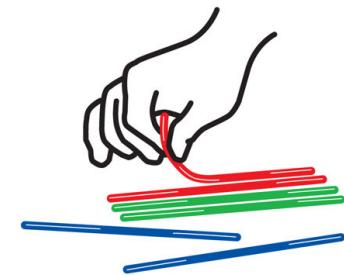
```
Python 3.4.1: File IO - Reading.py - I/1
File Edit Format Run Options Windows Help
def main():
    f = open('lines.txt')
    for line in f.readlines():
        print(line)

main()
```

Writing code



Spontaneous speech



Motor skills



Predicting hiring outcome

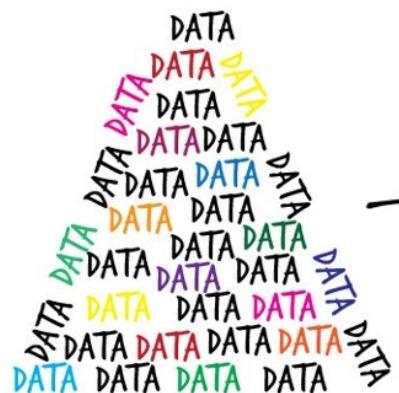


Labor market insights

Machine Learning

Key idea

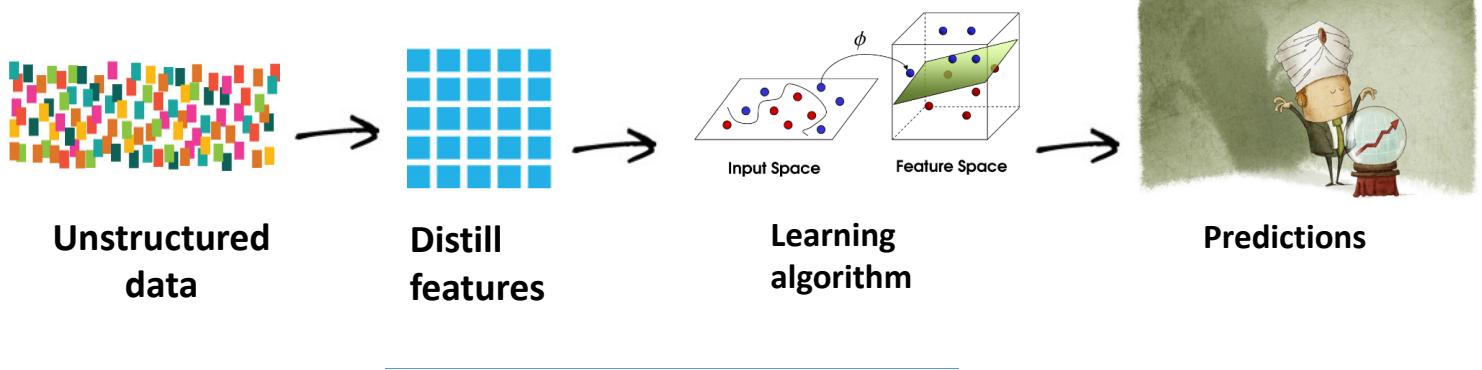
Predict an unknown from information you know



→ Learn
patterns →

Use these patterns on
new data to predict an
unknown

Machine Learning

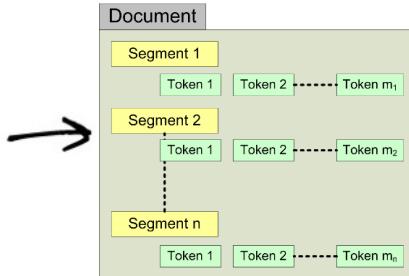


ENGLAND CAPTAIN David Beckham claims any different formation against Wales tomorrow will be down to Steve Grogan rather than a lack of player power - and called for "a little bit more respect" to be shown to the England manager. England have been experimenting with a 4-5-1 formation in training this week ahead of the World Cup quarter-final in Cardiff with Michael Owen set to lose his starting place even though Michael Owen has been suspended.

There has been speculation that the reason for a planned tactical shift is down to the England players. But Beckham said: "It's important to find different formations because you come up against different teams, especially when you have players out or injured. You have to compensate for that, so if you have to play different roles you do what the manager says."

"This whole player power thing has come up this week and it's been mentioned. "It happens all the time - if the manager decides he wants to play with players to focus on different formations, that's why if playing he calls a meeting. I think a little bit more respect has to be shown to the manager."

"He [Fransson] picks the team and formations. He discusses things with players and players agree with him. It's his decision. The decision is down to him and that's the way it's always been."



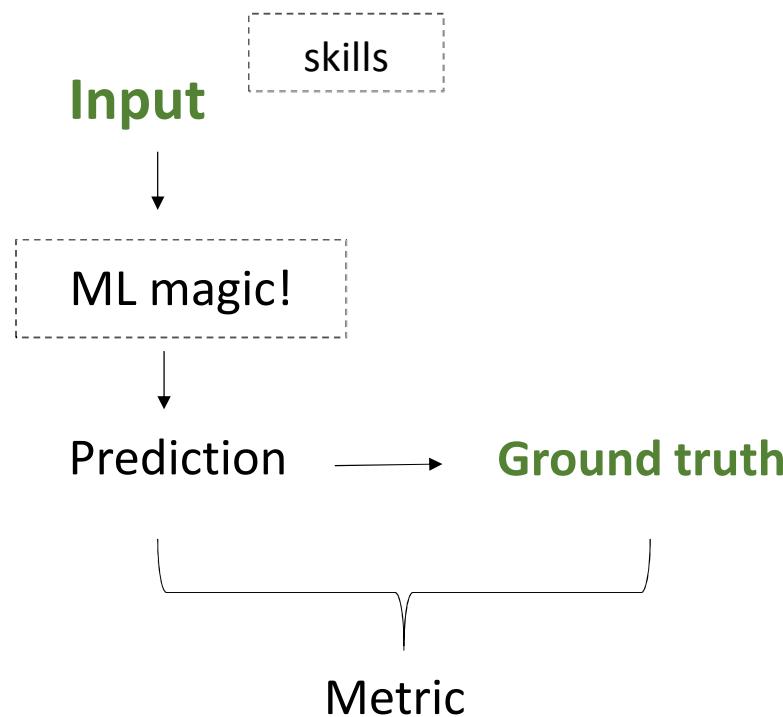
Automatically learn



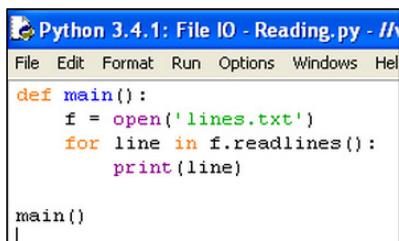
New sentence - *I am afraid of dogs* → Scared

EMOTION DETECTED

ML for skill identification – Basic setup

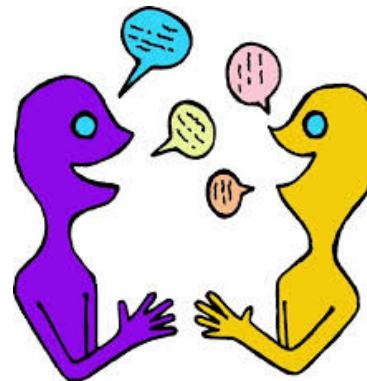


New age assessments

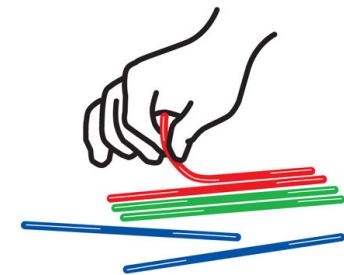


```
Python 3.4.1: File IO - Reading.py - I/Python 3.4.1: File IO - Reading.py - I/  
File Edit Format Run Options Windows Help  
def main():  
    f = open('lines.txt')  
    for line in f.readlines():  
        print(line)  
  
main()
```

Writing code



Spontaneous speech



Motor skills



OUTCOME

Predicting hiring outcome



Labor market insights

Grading programs

Every TA's nightmare



A 2-3 hour fling for test takers with no *objective* feedback to improve



You are stupid.
Goodbye.

Helps companies find talent at scale



Existing technology

- **Manual evaluation:** Doesn't scale; not standardized 
- **Test-case based evaluation**
 - High false-positives – hard code, not efficient 
 - High false-negatives – inadvertent errors

Rubric

OBJECTIVE

To print the pattern of integers

1

2 3

3 4 5

4 5 6 7

1. Are there loops? Are there print statements?
2. Is there a nested-loop structure?
3. Is the conditional in the inner loop dependent on
 - a variable modified in the outer loop?
 - a variable used in the conditional of the outer loop?

An implementation

```
void print_1(int N){
    for(i=1 ; i<=N; i++){
        print newline;
        count = i;
        for(j=0; j<i; j++)
            print count;
        count++;
    }
}
```

TARGET PROGRAM

```
void print(int N){
    for(i=1 ; i<=N; i++){
        print newline;
        count = i;
        for(j=0; j<i; j++)
            print count; count++;
    }
}
```

OBJECTIVE

To print the pattern of integers

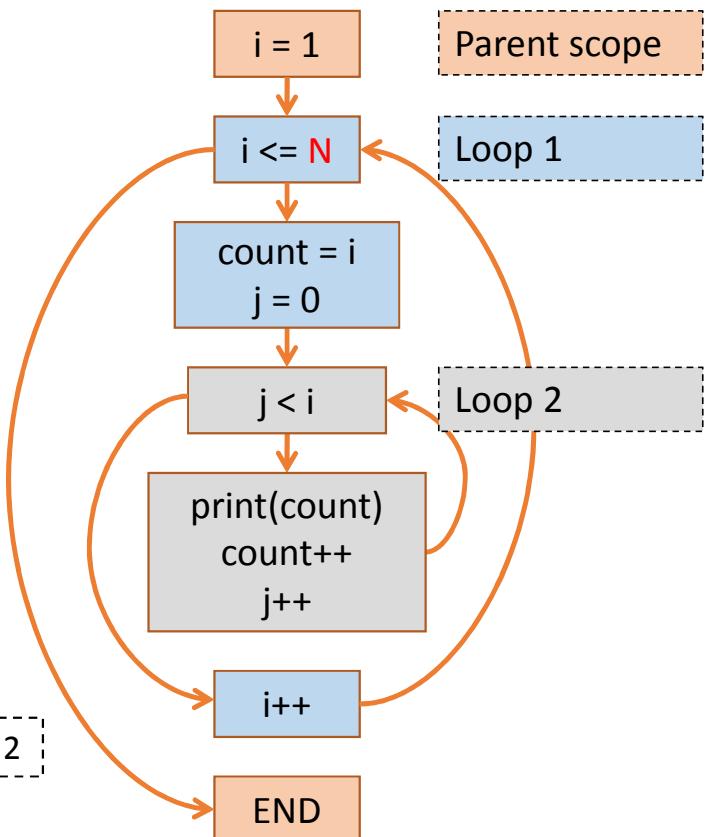
```

1
2 3
3 4 5
4 5 6 7

```

count(block1:loop(loop(++))) = 2

CONTROL FLOW GRAPH



Automata - Our technology

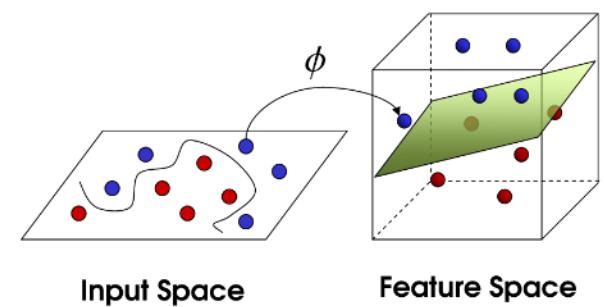


Get an expert to grade
these codes



$$A = \begin{bmatrix} 3 & 0 & 1 & 1 \\ 1 & 2 & 5 & 0 \\ 1 & 1 & 3 & 0 \\ 2 & 0 & 1 & 1 \end{bmatrix}$$

Use our grammar to
extract features from
each student's code



Input Space Feature Space

Build a model for each
question in our
question bank

Results

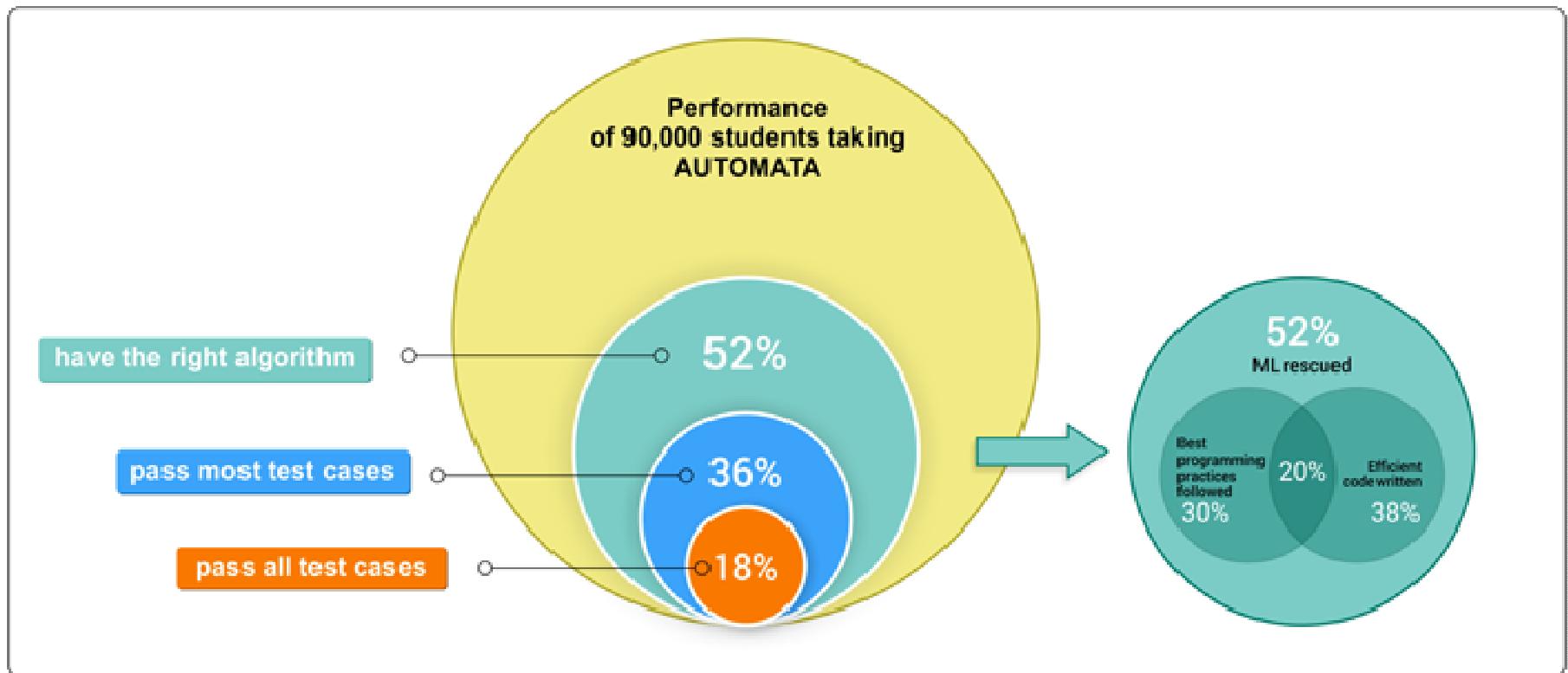
Work published at KDD 2014 and KDD 2016

PROBLEM	# of features	Cross-val correl	Train correl	Validation correl
1	80	0.61	0.85	0.79
2	68	0.77	0.93	0.91
3	193	0.91	0.98	0.90
4	66	0.90	0.94	0.90
5	87	0.81	0.92	0.84

Validation correlation > 0.79

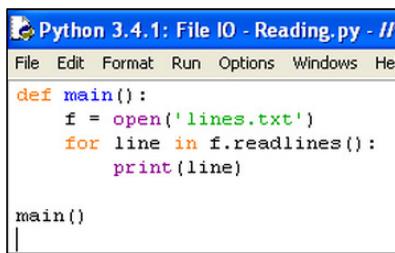
Matches inter-rater correlation between two human raters

Impact



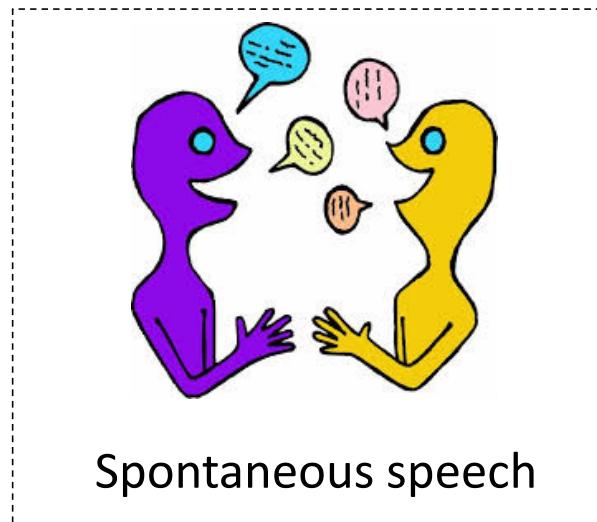
90K programs written by undergraduates in the US

New age assessments

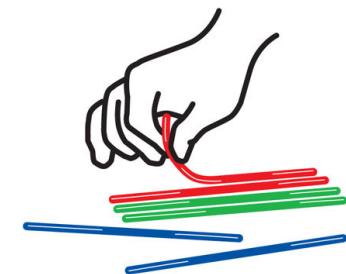


```
Python 3.4.1: File IO - Reading.py - I/Python 3.4.1: File IO - Reading.py - I/  
File Edit Format Run Options Windows Help  
  
def main():  
    f = open('lines.txt')  
    for line in f.readlines():  
        print(line)  
  
main()
```

Writing code



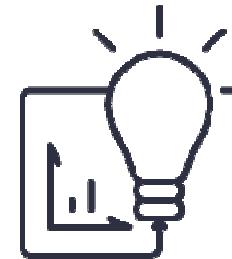
Spontaneous speech



Motor skills



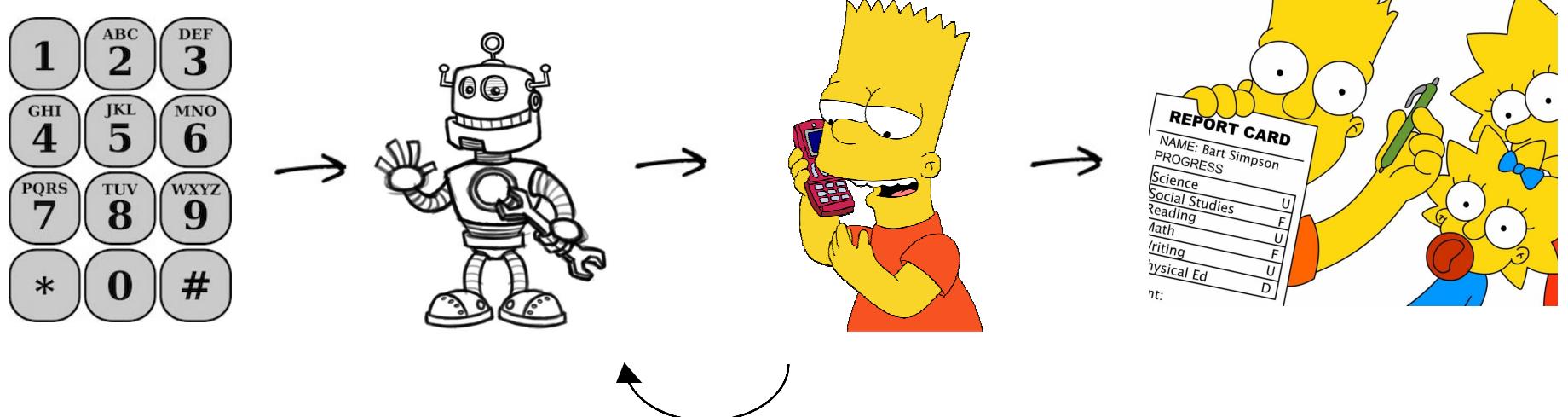
Predicting hiring outcome



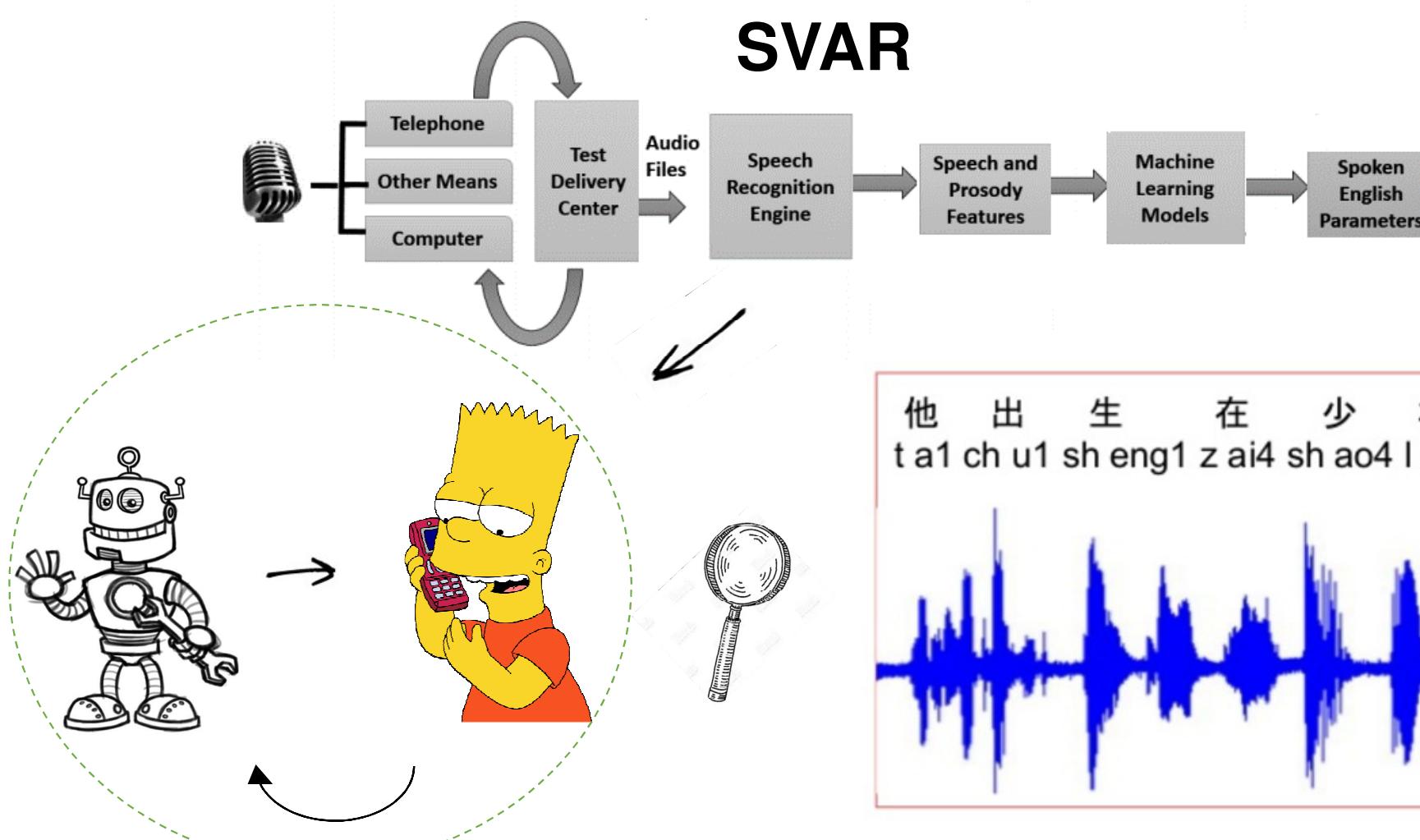
Labor market insights

SVAR

- A first-of-its-kind IVR based Spoken English Assessment tool
- Built in-house
- Live since 2010



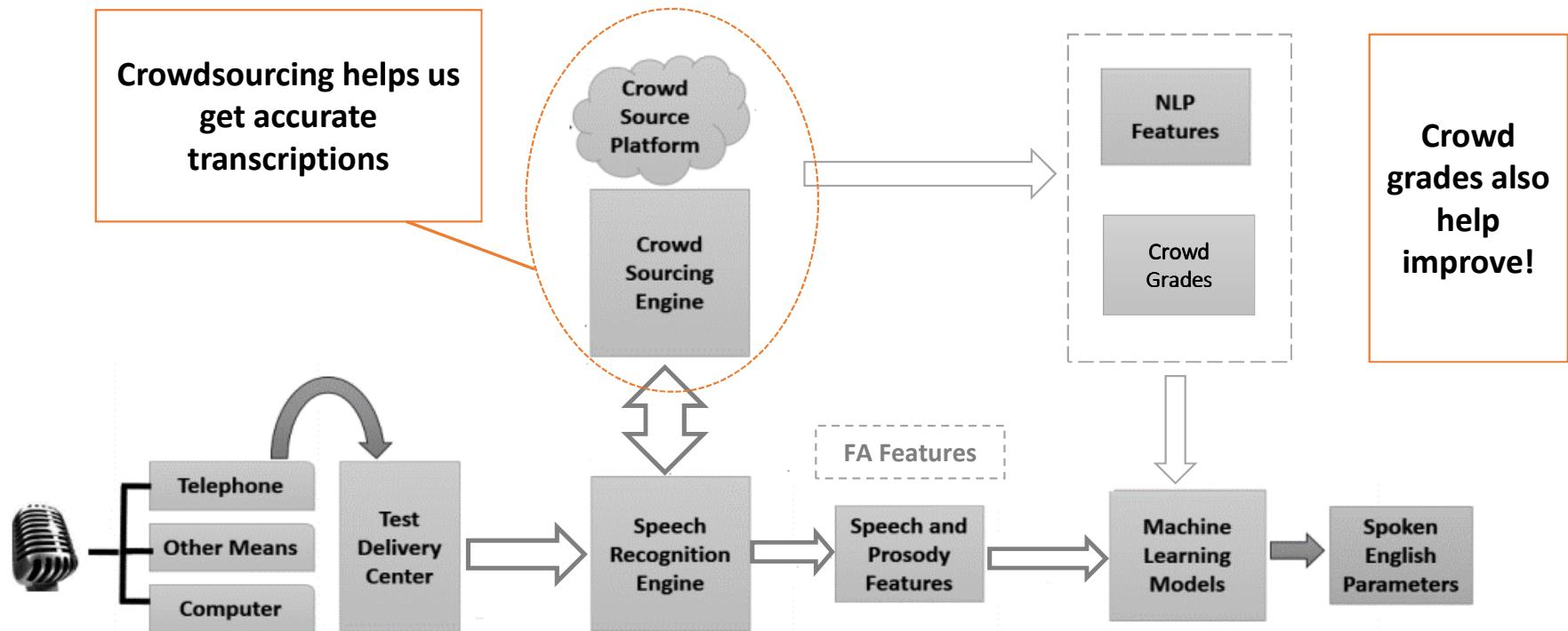
SVAR



他 出 生 在 少 林 寺
t a1 ch u1 sh eng1 z ai4 sh ao4 l in2 s iy4



SVAR



Results

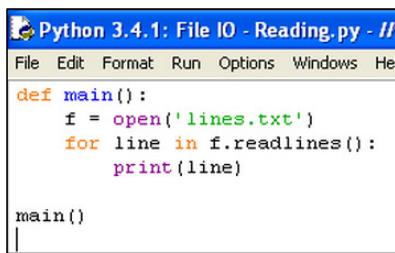
Work published at KDD 2015 and ACL 2015

Technique	Model Code	Feature Type	IN Set		PH Set	
			Train r	Validation r	Train r	Validation r
Ridge Regression	RR-1	RS/LR	0.42	0.51	0.47	0.44
	RR-2	Pure ML	0.46	0.48	0.60	0.54
	RR-3	Crowd Grades	0.61	0.67	0.61	0.71
	RR-4	ML-CS	0.64	0.70	0.77	0.60
	RR-5	All	0.74	0.74	0.76	0.79

Crowdsourced transcriptions + Crowd grades outperforms all other methods

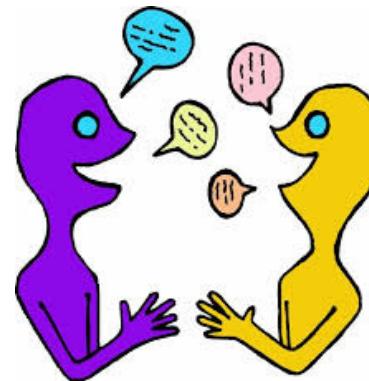
Accuracy nears inter expert agreement (~0.8).

New age assessments

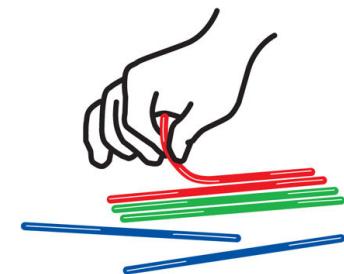


```
Python 3.4.1: File IO - Reading.py - I/Python 3.4.1: File IO - Reading.py - I/  
File Edit Format Run Options Windows Help  
  
def main():  
    f = open('lines.txt')  
    for line in f.readlines():  
        print(line)  
  
main()
```

Writing code



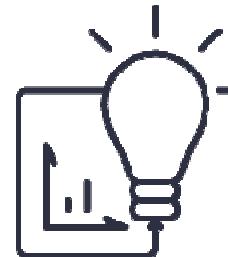
Spontaneous speech



Motor skills

OUTCOME

Predicting hiring outcome

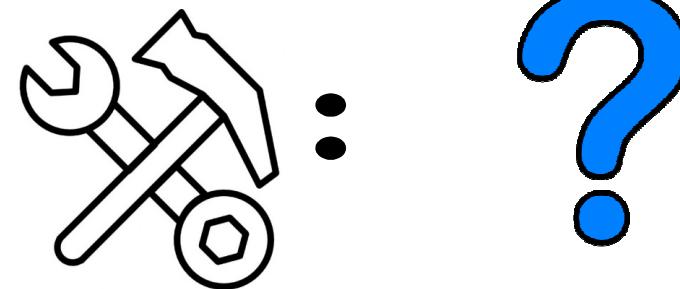


Labor market insights

Many standardized tests
predict performance of
knowledge workers.

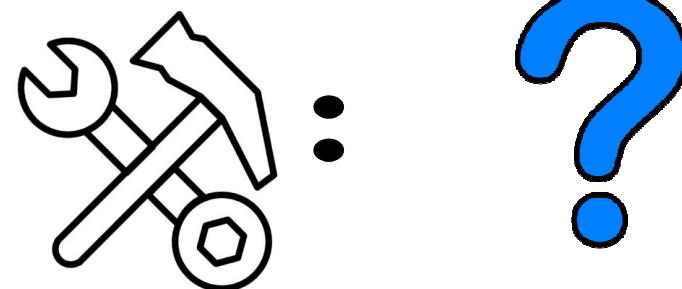


No automated tests for
vocational job workers



350 / 1650 jobs

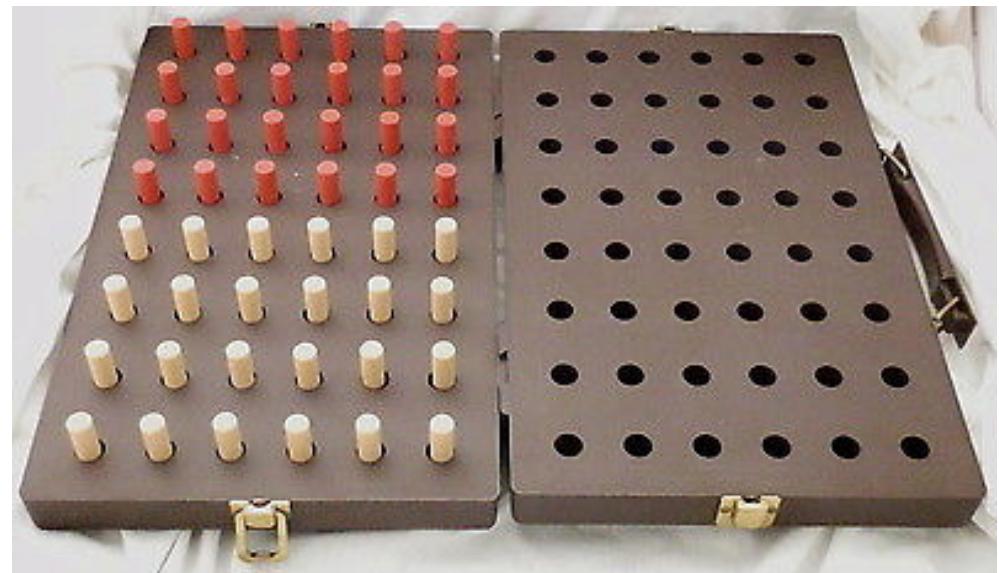
require motor-skills



O'Connor Tweezer Dexterity Test

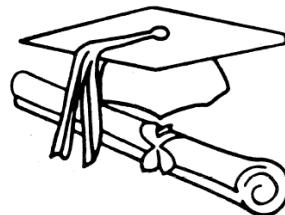


GATB Manual Dexterity Scale



350 / 1650 jobs

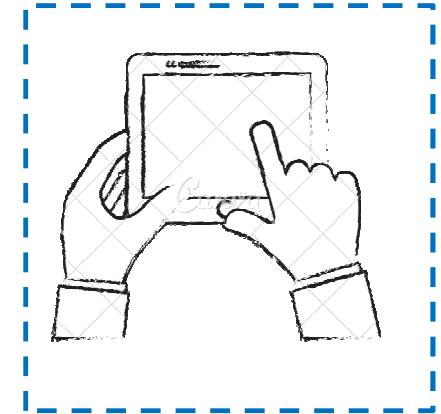
require motor-skills



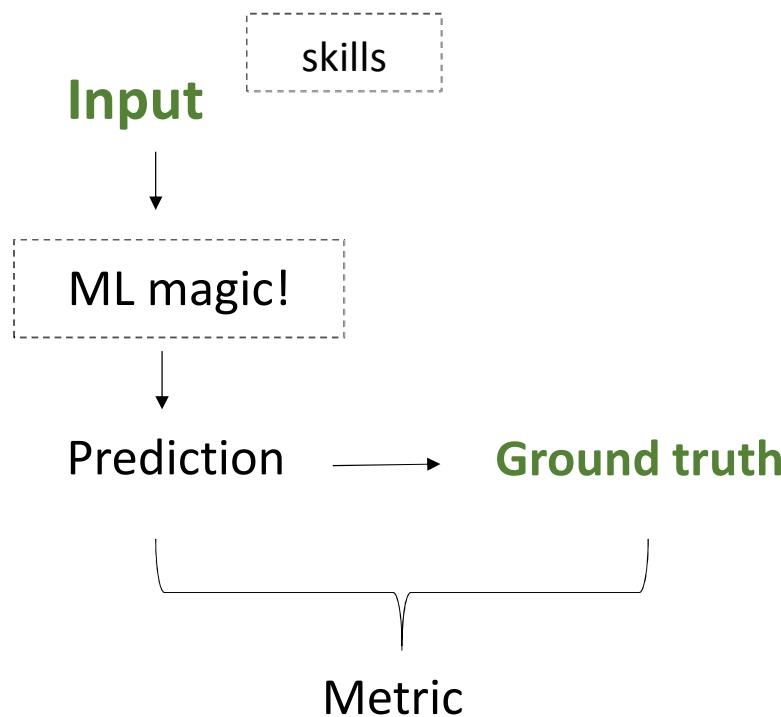
:



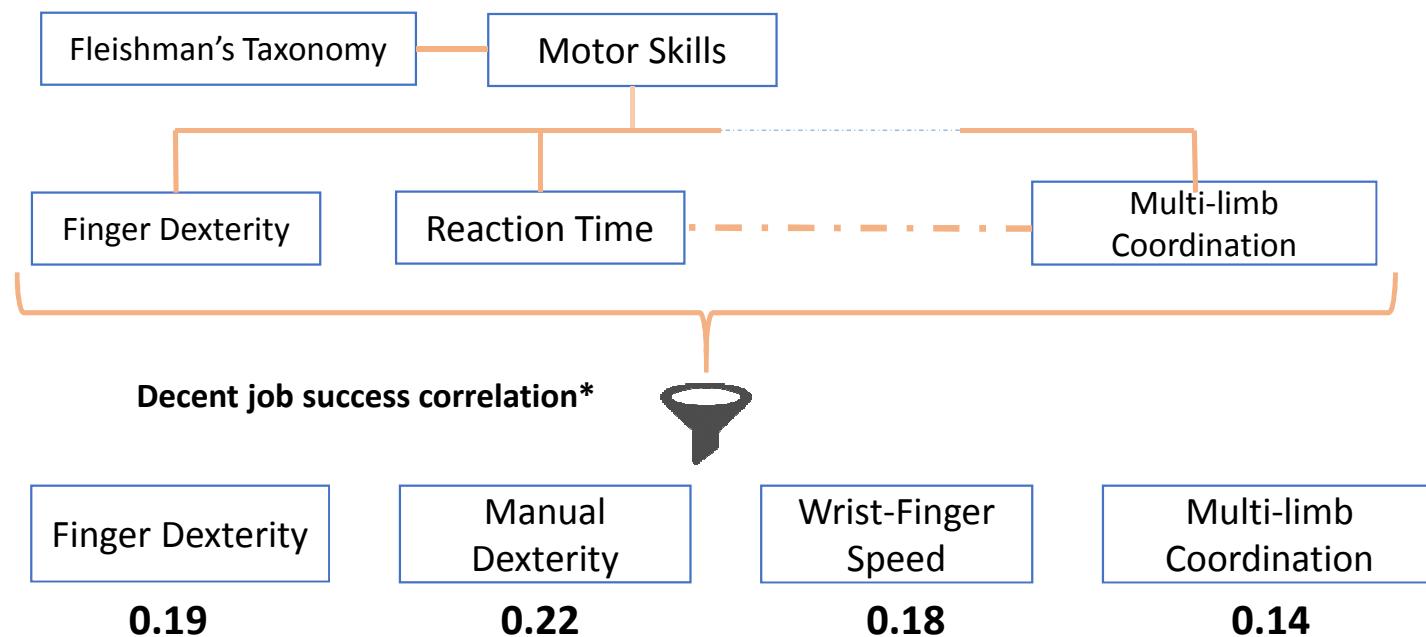
:



ML for skill identification – Basic setup

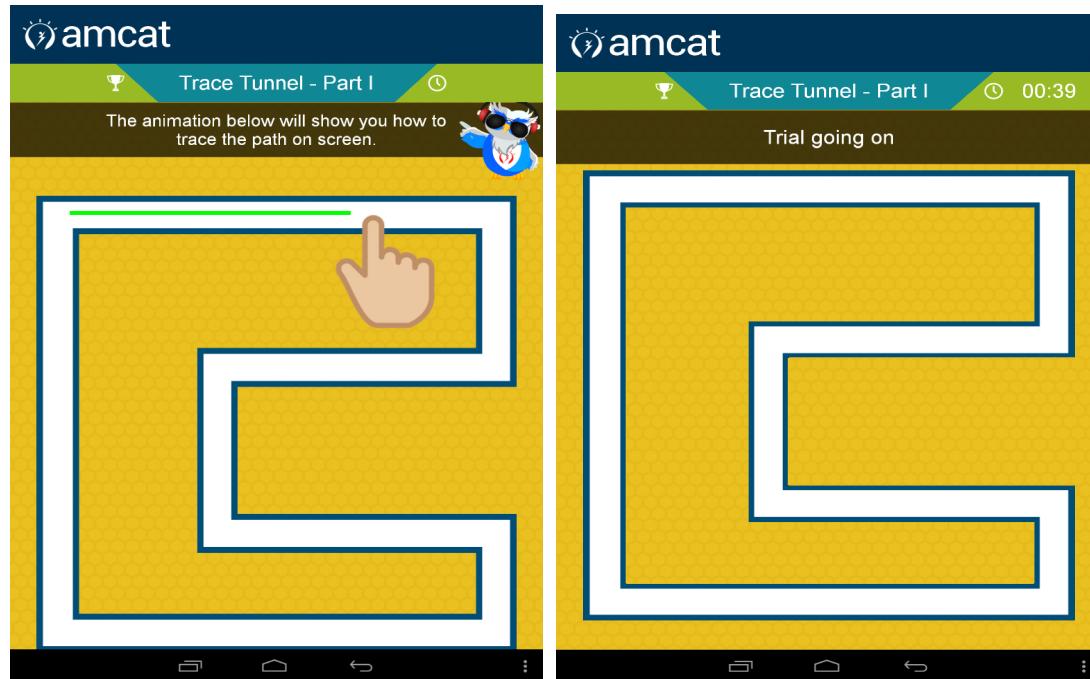


Which motor skills to measure?

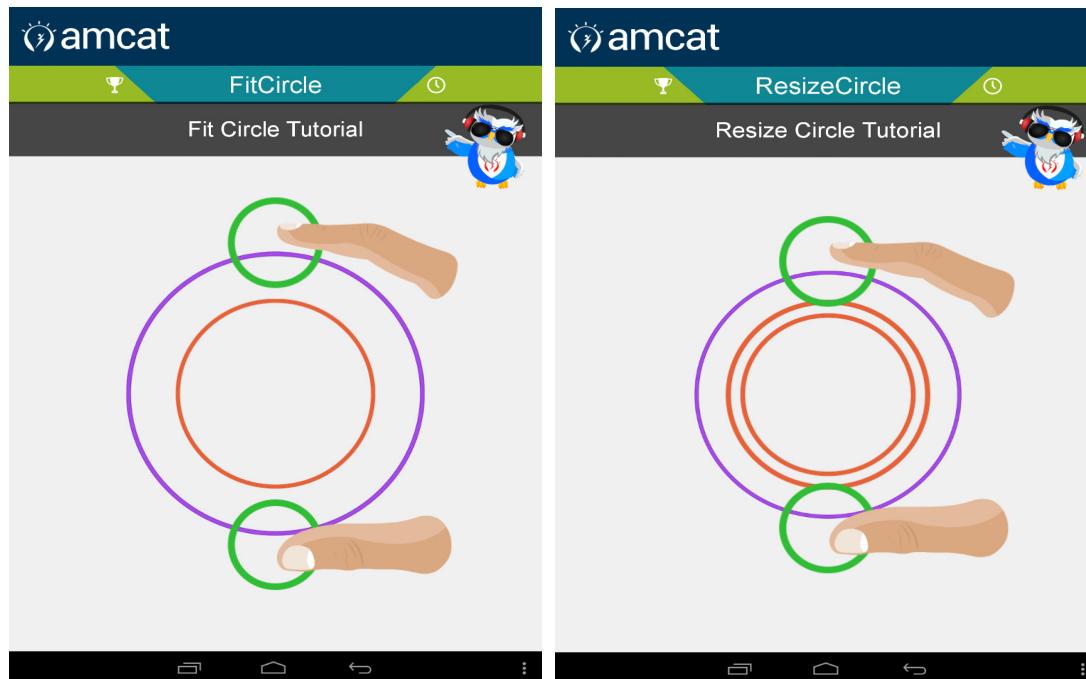


*J. J. McHenry and S. R. Rose. Literature review: Validity and potential usefulness of psychomotor ability tests for personnel selection and classification. Technical report, DTIC Document, 1988

Trace Path A & B

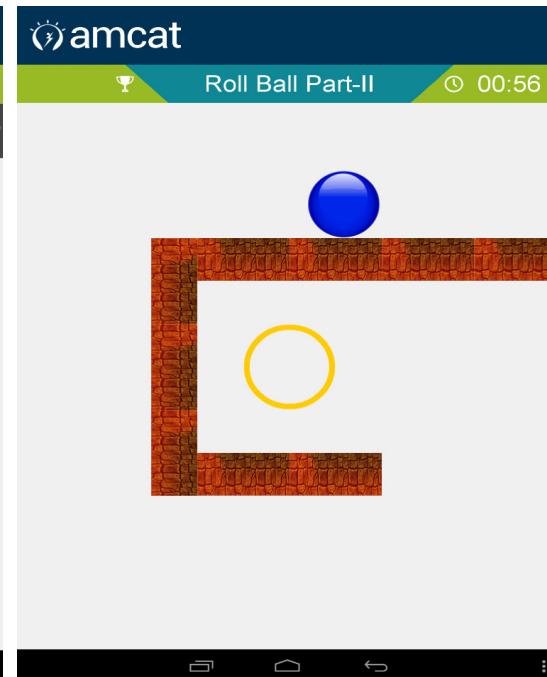
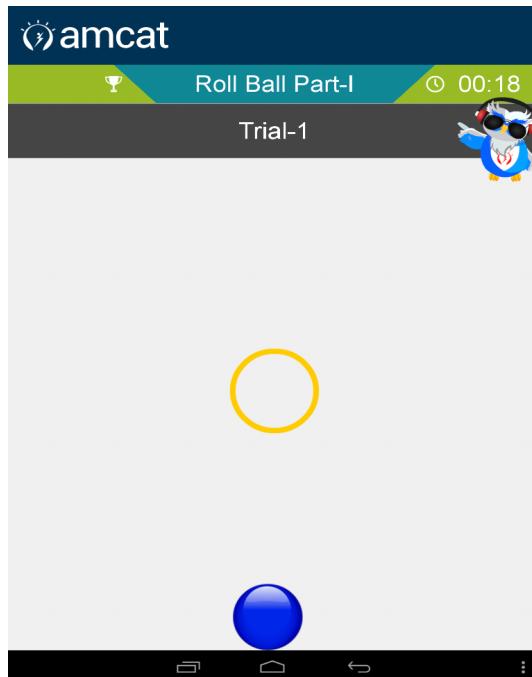


Fit and Resize Circle



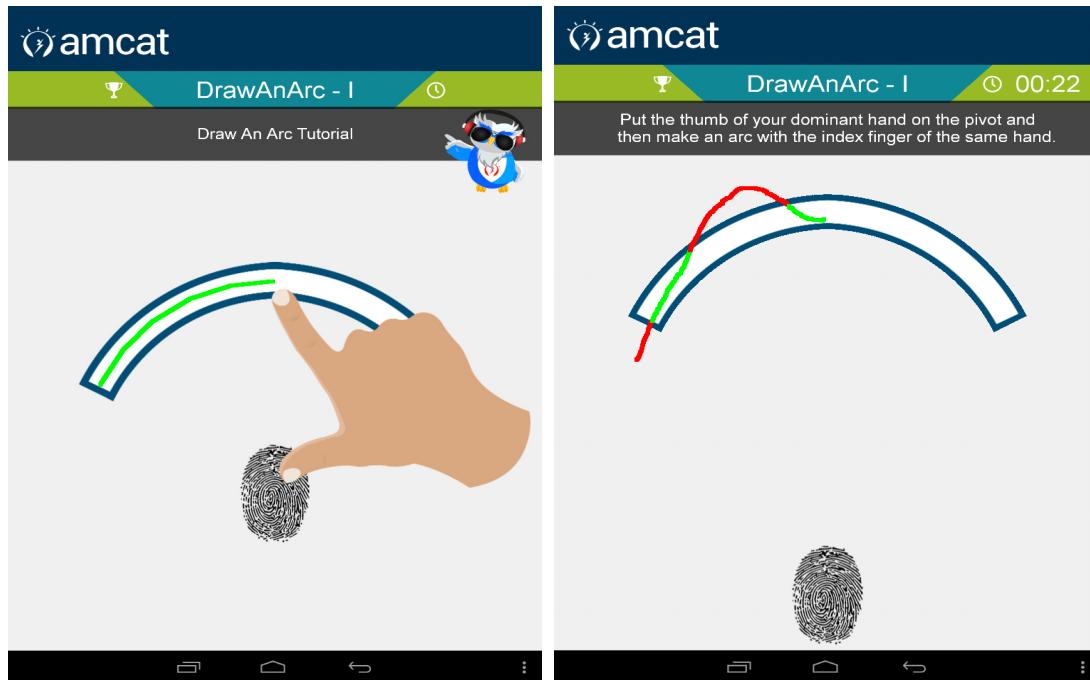
Requires **finger and wrist** movements.

Roll Ball A & B



Requires **coordinated arm and shoulder movements.**

Draw an Arc A & B



Results

Work published at UbiComp 2016

Tablet
vs
Job performance

0.19 to 0.38

Manual
vs
Job performance

0.19 to 0.33

- For each performance metric, **atleast one tablet** score shows a significant correlation.
- There is maximum support for **Trace app** which shows the highest correlation with job performance in **five out of the seven** metrics.

Outcome predictions using test scores

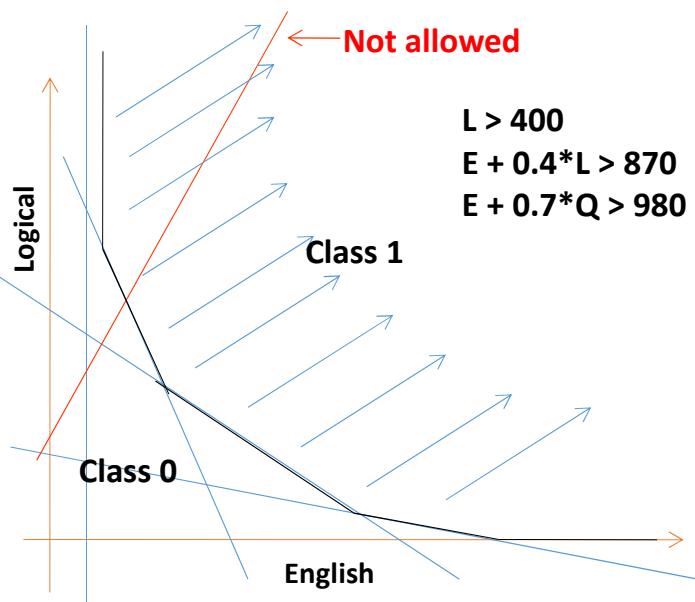
Client: What's the right cut-off to have on our test scores?

Conditions:

- I get to choose roughly how many candidates I want to talk to
- Your recommendations give the least type-1 and type-2 errors
- Your recommendations make sense



Outcome predictions using test scores



Outcome predictions using test scores



- $0.42 * \text{English} + \text{Logical} > 722$

- $\text{Logical} > 464$

OR



- $\text{Computer} > 360$

- $0.47 * \text{English} + 0.98 * \text{Computer} + \text{Logical} > 1196$

- Logical ability needs to be above the mean.
- Above a minimum benchmark, lack of logical Ability can be offset by high English comprehension skills

- The first filter eliminates bottom 28% candidates in Programming.
- Above a minimum benchmark, lack of Programming knowledge can be offset by high Logical Ability and English skills.

Insights into the labor market

Who gets a job?

Variable	coefficient	p-value	Unit of change*	Odds ($e^{(\text{coefficient} * \text{unit})}$)
English score	0.0026	0.00	100	1.29
Quantitative Ability score	0.0003	0.38	100	1.03
Logical Ability score	0.0014	0.01	100	1.15
Domain Percentile	0.0037	0.04	10	1.04
10th class percentage	0.0083	0.16	10	1.09
12th class percentage	-0.0086	0.08	10	0.92
College Percentage	0.0151	0.01	10	1.16
Gender	-0.0442	0.60	1	0.96
Tier of college	-0.1270	0.03	1	0.88
Branch of study	0.1515	0.05	1	1.16
Tier of city	-0.0026	0.96	1	1.00
Openness to Experience score	-0.0253	0.58	1	0.98
Extraversion + Agreeableness Score	0.0001	1.00	1	1.00
Polychronicity score	0.0175	0.66	1	1.02
Constant	-4.1389	0.00		

Merit

A candidate with an AMCAT English & Logical score higher by 100 points each and domain percentile up by 10 points has 54% higher odds to get a job.

Bias

A candidate from a tier 2 campus has 12% (25%) lower odds and tier 3 campus has 24% (33%) lower odds to get a job even if he/she has the equal merit.

Insights into the labor market

What salary does one get?

Variable	Coefficient	p-value	Unit of change*	Change in salary (in INR thousands)
English score	0.0013	0.01	100	13325
Quant Ability score	-0.0002	0.64	100	-1883
Logical Ability score	0.0014	0.04	100	14164
Domain Percentile	0.0013	0.57	10	1322
10th class percentage	0.0067	0.39	10	6709
12th class percentage	0.0007	0.91	10	732
College Percentage	0.0220	0.00	10	21979
Gender	0.1563	0.12	1	15631
Tier of college	-0.3326	0.00	1	-33257
Branch of study	-0.0376	0.69	1	-3762
Tier of city	0.0341	0.63	1	3407
Openness to Experience score	0.0157	0.78	1	1570
Extraversion + Agreeableness Score	-0.0393	0.28	1	-3934
Polychronicity score	0.0105	0.83	1	1049
Constant	-0.3968	0.52	1	

Merit

Bias

Higher English & Logical scores by 100 points each and higher college percentage by 10%, translate into a higher annual salary by Rs. 49500 (15% of standard package)

A candidate from a tier 2 campus has **12% (25%) lower odds** and tier 3 campus has **24% (33%) lower odds** to get a job even if he/she has the **equal merit**.

31 July 2017, MSR Bangalore



Information Asymmetry

Seva Setu

Mother care



Each one, reach one



Auto Aid



Monitor & improve Village Health
and Nutrition Days (VHNDs)

Child care



NRC rehabilitation of SAM
children

Citizen care



Pension schemes



Aam Sabha



Audits of existing state machinery
RTIs, Publications

Skills training



Training and employing



The problem

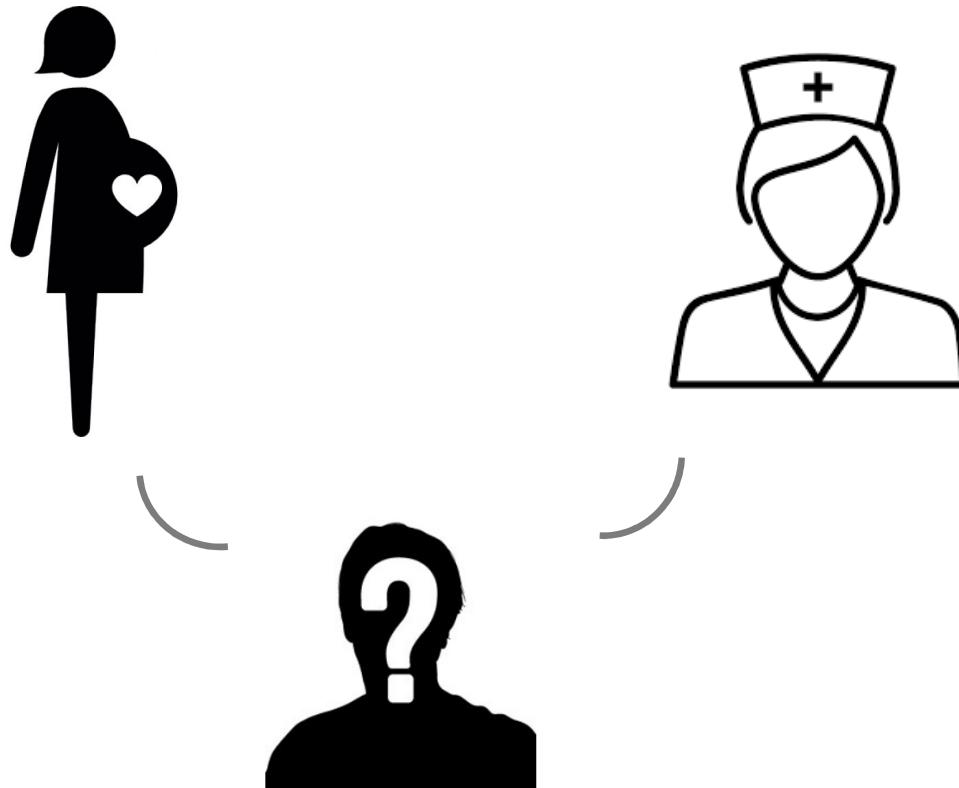
Rapid survey on Children tells us
10 million kids in Bihar are
malnourished

Once a month government program
where a nurse visits each village. Not
personalized enough

MMR in Bihar is 219 per 100,000
births

IMR in Bihar is 43 per 1000 births

The problem



Each One, Reach One



Seva Setu's Each One Reach One
program

Login

FAQ



1145
Mothers

MOTHERS FROM RURAL INDIA
REGISTERED ON OUR PLATFORM



134
Call Champions

YOUNG WOMEN FROM
CORPORATE INDIA WHO TALK TO
MOTHERS FROM RURAL INDIA



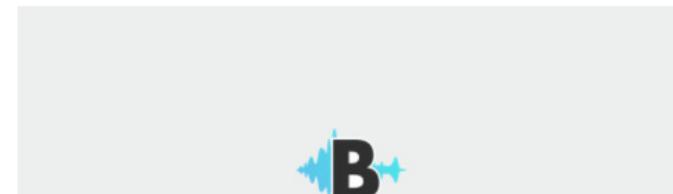
30100
minutes of phone
conversation

A FULLY AUTOMATED SYSTEM TO
TRACK AND MANAGE PHONE
CALLS



You?
next call champion

DO YOU WANT TO BE A CALL
CHAMPION TOO? LOG IN RIGHT
AWAY!



sevasetu.org/mother_care

Seva Setu's Mother care tool

↳ Mothers 2 assigned

📞 My calls 25 scheduled

☰ Action items

☰ Check list

⚙️ Settings

🔊 Feedback and Bugs

You are a Call Champion to the following mothers

Munni Kumari from Panchai Mubarak

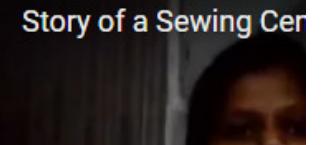
Anju Kumari from Panchai Mubarak

[Be a call champion to another mother](#)

What is the Mother Care program?



Other stories from Seva



sevasetu.org/mother_care

Seva Setu's Mother care tool Logout

Call My calls 25 scheduled

List Action items

List Check list

Gear Settings

Microphone Feedback and Bugs

Expected date of delivery: 08 Jul 16
Husband's name: Ramesh Rai

This call is scheduled for 01 Jan 16

Record your notes View previous notes Calls completed and scheduled

Please record the status of the call

Not called

Please record the status of expected date of delivery

Correct

Action items to be discussed

Register & ANC

- Get Registered with AWC
- Get MCP card
- Get blood, urine, weight, BP [ANC] checkup done

Village: Panchai Mubarak
Seva Setu's field worker associated: Mahendra Paswan

5 previous calls were not made. Please discuss action items from those calls too.

Action points for Seva Setu's field worker based on this call Important

[[Mention only those notes on which Seva Setu needs to act on.
Place all other general notes related to the call in the textbox below]]

General notes about Munni Kumari

Submit note

Takeaway

Bunch of hard problems out there for us to deal with

Key is to identify the right problem and articulate it right

Hope that the articulation can leverage technology to make it sustainable, easily implementable