WOCMAN:

Harnessing the Wisdom of the

Crowd for High-Quality Estimates

DANIEL W. BAROWY EMERY D. BERGER

UMassAmherst

DANIEL G. GOLDSTEIN SIDDHARTH SURI

Microsoft



"Is it healthy to eat this donut?"

"Is it within my calorie budget?"







calorimeter

• \$30,000.00

- advanced degree
- several hours (calibration, running, analysis)
- expert labor
 \$50-100/hr
- · Lavoisier, c1780

- cellphone app
- free-ish
- Homer-approved
 button-pushing
- · seconds to minutes



• ???





Facts There exists a true estimate of

the donut's calories.

None of my friends are donut

experts.

Unlikely that any given friend can

tell me precise calorie count.

Unlikely that I can discern which

friend knows the right answer.









"How many calories in this donut?"



"How many calories in this donut?"



Is this better? Maybe.



median=275.5

"How many calories in this donut?"



There's something different about this

answer...

median=275.5







a proxy for the real thing







Insight:

inference about pop.'s opinion



a proxy for the real thing

small n: median $(X_1) = 300.5 \text{ kcal } \pm 76$

largen: median $(X_2) = 275.5 \text{ kcal } \pm 55$



Insight:

inference about pop.'s opinion



a proxy for the real thing

small n: median $(X_1) = 300.5 \text{ kcal } \pm 76$

large n: median $(X_2) = 275.5 \text{ kcal } \pm 55$



Insight: inference about pop.'s opinion a proxy for the real thing $median(X_1) = 300.5 kcal \pm 76$ small n: large n: median $(X_2) = 275.5 \text{ kcal } \pm 55$ 300 200 401 To get a better estimate, ask more Homers! Where can we find more Homers ...?

- lots of problems, e.g., payment, scheduling, quality control (esp. bad actors!)
- small fees for simple tasks available via API call
- human task market; lots of Homers!







get calories(

UTOMAN



Pr[correct] = nil Pr[agree] = nil

Different kind of consensus needed.

get calories(

UTOMAN

Pr[correct] = nil Pr[agree] = nil

Different kind of consensus needed.

def numCalories(url: String) = Estimate (
 budget = 5.00,
 confidence_interval = SymmetricCI(50),
 confidence = 0.95,
 text = "How many calories are in the food
 pictured?",
 image_url = url,
 estimator = Mean
}

WoCMan increases sample size until either

1) interval meets user-defined "tightness"

2) budget is exhausted





nose finder

calorie counter



where's waldo?



NPR planet money's ox

MAE with fixed n=200 was 83.6 kcal (±41) run 30x w/varying CI (±250 to ±50 kcal)



vs Google's IM2Calories

calorie counter

· IM2Calories needs access to your GPS

huh?

· essentially a constraint solver: searches

nearby restaurant menus using labels

generated by vision algorithm

IM2Calories MAE: 152.95, SE: 15.61

WOCMAN MAE: 103.08, SE: 6.00

nose finder (aka nose picker) 15 images from Google image search





· 100% correct; Cls are in PIXELS

· Cls ranged from #3.8 to #41.3 (asked: #15)

cost ranged from \$0,72 to \$5.76

where's waldo? 5 very large where's waldo images



· 100% correct

Cls ranged from ±22 to ±943 (asked: ±15)

• all cost max of \$5.76



NDB No. *	Description	Food Group	•
18248	Doughnuts, cake-type, plain (includes unsugared, old-fashioned)	Baked Products	
18249	Doughnuts, cake-type, plain, chocolate-coated or frosted	Baked Products	
18250	Doughnuts, cake-type, plain, sugared or glazed	Baked Products	
18251	Doughnuts, cake-type, chocolate, sugared or glazed	Baked Products	
18253	Doughnuts, french crullers, glazed	Baked Products	
18254	Doughnuts, yeast-leavened, with creme filling	Baked Products	
18255	Doughnuts, yeast-leavened, glazed, enriched (includes honey buns)	Baked Products	
18256	Doughnuts, yeast-leavened, with jelly filling	Baked Products	
18436	Doughnuts, yeast-leavened, glazed, unenriched (includes honey buns)	Baked Products	

UDSA: 450 kcal; Dunkin' Donuts: 280 kcal

WOCMAN

- · easy to use
- · enforces quality on input data
- · allows cost/accuracy tradeoff

http://automan-lang.org

DANIEL W. BAROWY EMERY D. BERGER

UMassAmherst

DANIEL G. GOLDSTEIN SIDDHARTH SURI

Microsoft