Using Data Science for Social Good

Rayid Ghani

Center for Data Science & Public Policy
University of Chicago
Impaired Attention

Lack of Motor Skills

Hearing Loss

Learning Disability

Lower IQ

Memory Problems
WHO IS DROPPING OUT?

30% of high school students in the U.S. drop out.

Every year, 1.3 million students drop out of school.

7,000 every day of school or 1 every 26 seconds.
Density plot for predicted probabilities

ROC curve

Actual BLL
- red: over 5
- green: under 5

Density

Fitted value

True positive rate

False positive rate

AUC = 0.754

(Joe Brew, Alex Loewi, Subhabrata Majumdar, Andrew Reece, Eric Rozier)
Area Under ROC Curve

Time of Predicted Dropout (Months)

(Sarah Abraham, Jeff Lockhart, Sarah Tan, Rafael Turner, Young-Jin Kim)
(Everaldo, Aguiar, Nasir Bhanpuri, Hima Lakarajju, David Miller, Ben Yuhas)
Education/Training

Collaborative Projects:
Real Projects, Real Data, Real Partners

Reusable Software
and Data
The Eric & Wendy Schmidt Data Science for Social Good Summer Fellowship

http://dssg.uchicago.edu

@datascifellows
Goal of the Fellowship

• Train the next generation of data scientists who care about and understand how to solve social problems

• Expose and train governments and non profits to better use data to make better decisions

• Seed a community of people and organizations working together to make an impact
Fellowship by the Numbers

2013
- 36 Fellows
- 6 Mentors
- 12 Projects
- 12 Weeks

2014
- 48 Fellows
- 8 Mentors
- 14 Projects
- 12 Weeks

2015
- 42 Fellows
- 10 Mentors
- 12 Projects
- 14 Weeks
Fellows

~1800 Applicants 45 countries ~300 Universities

126 fellows

Computer Scientists
Statisticians
Social Scientists
Public Policy
(and other computational and quantitative fields)

CMU  U. of Chicago  Northwestern
Harvard  MIT  Stanford  Princeton  ITAM  Cornell  Yale  Villanova  Ohio State

USC  U Penn  Notre Dame
U of Minnesota  U of Michigan  Cambridge  McGill  UC Berkeley  U of Colorado  Swarthmore  Oberlin

UIUC  Emory  Duke
Fordham  Johns Hopkins  IIT  SAIC  NYU

Penn State  Simon Fraser  UC Santa Barbara
Project Partners

Agentis
Cincinnati
Perspective Charter School
Police
Vancouver Public Schools
KIPP Chicago
Chicago Public Schools
The University of Chicago
Chapin Hall
City of Memphis
World Bank Group
NOBLE
City of Chicago
Chicago Alliance to End Homelessness
Chapin Hall at the University of Chicago
Get Covered Illinois
Chicago Harris
Public Policy | The University of Chicago
Australian Conservation Foundation
Sunlight Foundation
The Case Foundation
EPIC
Energy Policy Institute at the University of Chicago
Pecan Street Inc.
Wiki Energy
Enroll America
Healthy Chicago
Health Leads
Ushahidi
careerbuilder
Skills for Chicagoland’s Future
MESA Public Schools
CDOT
NorthShore University HealthSystem
IDHS
Wake County Public School System
Institute for Housing Studies
cclba
Arlington Public Schools
Conservation International
Rayid Ghani
@rayidghani
Critical Research Gaps

- Problem Formulation
- Communication
- Experimental Design
- Databases
- Computer Science & Programming

Making an Impact with Data

Machine Learning

Social Science

Rayid Ghani
The Data Science for Social Good Summer Fellowship

The Eric & Wendy Schmidt
Data Science for Social Good
Summer Fellowship 2013

Rayid Ghani
@rayidghani
How to get involved

Students => Data Science Fellows
Recent PhDs => Apply to be Post-docs
Practitioners in Industry => Apply to be Visiting Data Scientists and/or Summer Mentors
Software Developers => Apply to be Data Engineers
Organizations with a problem and data => Apply to be a Project Partner

http://dssg.uchicago.edu
Contact Information

Rayid Ghani
rayid@uchicago.edu

For more information on the Summer Fellowship
http://dssg.uchicago.edu

For more information on the Center for Data Science & Public Policy
http://dspplab.com