

# The Data Farm

Neil D. Lawrence

Department of Computer Science  
Sheffield Institute for Translational Neuroscience  
University of Sheffield  
Festival of Science & Engineering

13th March 2015







HIERONIMUS BOSCHUS PICTOR

COLLEGIUM S. IUDYTHAE PRIMVS ORDEM





facebook

1 Hacker Way



















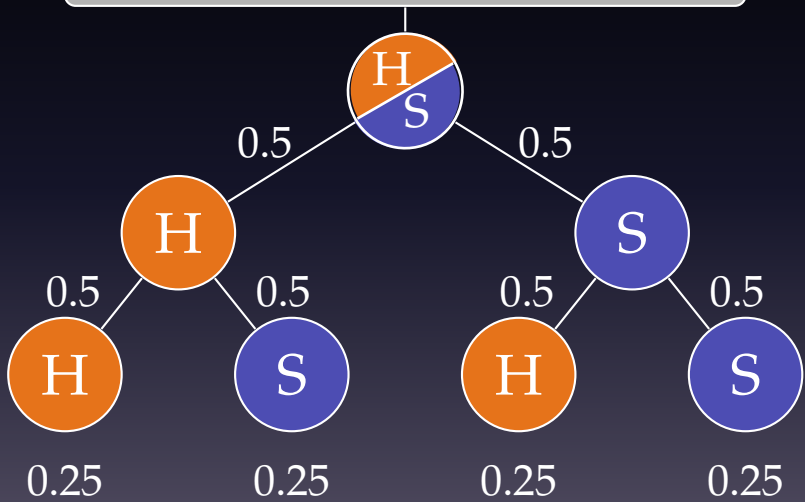








coin flipping









# Computer-based personality judgments are more accurate than those made by humans

Wu Youyou<sup>a,1,2</sup>, Michal Kosinski<sup>b,1</sup>, and David Stillwell<sup>a</sup>

<sup>a</sup>Department of Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom; and <sup>b</sup>Department of Computer Science, Stanford University, Stanford, CA 94305

Edited by David Funder, University of California, Riverside, CA, and accepted by the Editorial Board December 2, 2014 (received for review September 28, 2014)

**Judging others' personalities is an essential skill in successful social living, as personality is a key driver behind people's interactions, behaviors, and emotions. Although accurate personality judgments stem from social-cognitive skills, developments in machine learning show that computer models can also make valid judgments**

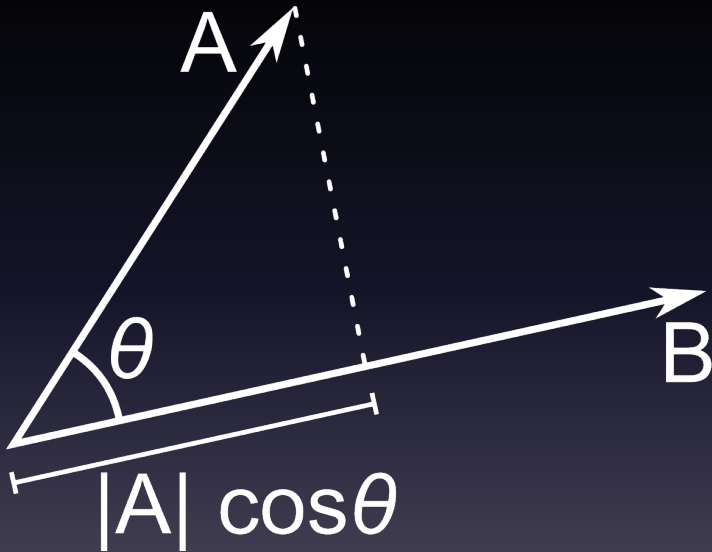
psychological traits (11). We used LASSO (Least Absolute Shrinkage and Selection Operator) linear regressions (16) with 10-fold cross-validations, so that judgments for each participant were made using models developed on a different subsample of participants and their Likes. Likes are used by Facebook













## The Data Farm: Science Week Presentation ¶

### Learning from Data

[Neil D. Lawrence](#) and the [Sheffield Machine Learning Research Group](#)

5th March 2014

This notebook has been made available as part of our **Open Data Science** agenda. If you want to read more about this agenda there is a [position paper/blog post available on it here](#).

This session is about 'learning from data'. How do we take the information on the internet and make sense of it. The answer, as you might expect, is using computers and mathematics. Luckily we also have a suite of tools to help. The first tool is a way of programming in python that really facilitates interaction with data. It is known as the "[IPython Notebook](#)", or more recently as the "[Jupyter Project](#)".

#### Welcome to the IPython Notebook

The notebook is a great way of interacting with computers. In particular it allows me to integrate text descriptions, maths and code all together in the same place. For me, that's what my research is all about. I try to take concepts that people can describe, then I try to capture the essence of the concept in a *mathematical* model. Then I try and implement the model on a computer, often combining it with data, to try and do something fun, useful or, ideally, both.

For the Science Week lecture on "The Data Farm" we are looking at *recommender systems*.

#### Recommender Systems

Do you watch Netflix? Have you ever rated a movie there? Do you buy books or electronics on Amazon? How about grocery shopping? All these companies want you to buy more, watch more or listen more. The best way of getting you to do that is by showing you more of what you like. But what do you like? What sort of person are you? Can the computer tell? It can certainly try! And it does so with a 'Recommender System'. Recommender systems are so important to Netflix that they offered a [\\$1 million dollar prize](#) for improving theirs.





‘Leave the THINKING to US.’

THE OLIGARCHY



‘Indeed it has been said that democracy is the worst form of Government except for all those other forms that have been tried from time to time.’

Winston Churchill 11th November 1947

ODSI  
open data science



sign in



subscribe



search

jobs dating more UK edition

# theguardian

Winner of the Pulitzer prize

home UK world sport football opinion culture economy lifestyle fashion environment tech money travel all

home

## media network

### ☞ Beware the rise of the digital oligarchy

**Neil Lawrence**

Powerful algorithms and the concentration of data in the hands means we need better models of data-ownership

0 comments



### Eight lessons political parties need to learn to woo young voters

**Matthew Hook**

2 comments



### Mobile World Congress 2015: what it means for marketing pros

**James Hilton**



### How we made MailMen for Royal Mail



### PocketHighStreet: linking bricks and clicks at a local level



### The return of the full-service agency approach

**Olly Markeson**

0 comments