

Detecting novel associations in large data sets.

Title	Detecting novel associations in large data sets.
Publication Type	Journal Article
Year of Publication	2011
Authors	Reshef, DN [1], Reshef, YA [2], Finucane, HK [3], Grossman, SR [4], McVean, G [5], Turnbaugh, PJ [6], Lander, ES [7], Mitzenmacher, M [8], Sabeti, PC [9]
Journal	Science
Volume	334
Issue	6062
Pagination	1518-24
Date Published	2011 Dec 16
ISSN	1095-9203
Keywords	Algorithms [10], Animals [11], Baseball [12], Data Interpretation, Statistical [13], Female [14], Gene Expression [15], Genes, Fungal [16], Genomics [17], Humans [18], Intestines [19], Male [20], Metagenome [21], Mice [22], Obesity [23], Saccharomyces cerevisiae [24]
Abstract	<p>Identifying interesting relationships between pairs of variables in large data sets is increasingly important. Here, we present a measure of dependence for two-variable relationships: the maximal information coefficient (MIC). MIC captures a wide range of associations both functional and not, and for functional relationships provides a score that roughly equals the coefficient of determination (R^2) of the data relative to the regression function. MIC belongs to a larger class of maximal information-based nonparametric exploration (MINE) statistics for identifying and classifying relationships. We apply MIC and MINE to data sets in global health, gene expression, major-league baseball, and the human gut microbiota and identify known and novel relationships.</p>
DOI	10.1126/science.1205438 [25]
Alternate Journal	Science
PubMed ID	22174245 [26]

PubMed
Central ID PMC3325791

Grant List 090532 // Wellcome Trust / United Kingdom
P50 GM068763 / GM / NIGMS NIH HHS / United States
P50 GM068763 / GM / NIGMS NIH HHS / United States
P50 GM068763-09 / GM / NIGMS NIH HHS / United States
T32 GM007753 / GM / NIGMS NIH HHS / United States
U54 GM088558 / GM / NIGMS NIH HHS / United States
U54 GM088558-03 / GM / NIGMS NIH HHS / United States
U54GM088558 / GM / NIGMS NIH HHS / United States

- [Contact Us](#)
- [Twitter](#)
- [UCSF Main Site](#)

Search

© 2014 The Regents of the University of California

Source URL: <http://turnbaughlab.ucsf.edu/content/detecting-novel-associations-large-data-sets>

Links:

- [1] <http://turnbaughlab.ucsf.edu/Papers?f%5Bauthor%5D=816>
- [2] <http://turnbaughlab.ucsf.edu/Papers?f%5Bauthor%5D=821>
- [3] <http://turnbaughlab.ucsf.edu/Papers?f%5Bauthor%5D=826>
- [4] <http://turnbaughlab.ucsf.edu/Papers?f%5Bauthor%5D=831>
- [5] <http://turnbaughlab.ucsf.edu/Papers?f%5Bauthor%5D=836>
- [6] <http://turnbaughlab.ucsf.edu/Papers?f%5Bauthor%5D=6>
- [7] <http://turnbaughlab.ucsf.edu/Papers?f%5Bauthor%5D=841>
- [8] <http://turnbaughlab.ucsf.edu/Papers?f%5Bauthor%5D=846>
- [9] <http://turnbaughlab.ucsf.edu/Papers?f%5Bauthor%5D=851>
- [10] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=626>
- [11] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=1>
- [12] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=821>
- [13] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=826>
- [14] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=181>
- [15] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=831>
- [16] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=836>
- [17] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=301>
- [18] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=86>
- [19] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=256>
- [20] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=36>
- [21] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=96>
- [22] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=46>
- [23] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=421>
- [24] <http://turnbaughlab.ucsf.edu/Papers?f%5Bkeyword%5D=841>
- [25] <http://dx.doi.org/10.1126/science.1205438>
- [26] <http://www.ncbi.nlm.nih.gov/pubmed/22174245?dopt=Abstract>