How to Get the Most Accurate Results for Your Data Using Ensembles for Numeric, Classification, Logistic and Time Series Data

Russ Conte, 2025 INFORMS Conference, Atlanta, Georgia

The vision: As easy to use as an LLM, but it’s for data.

Intro: The “Hello, world!” of ensembles

Live demo #1

NETFLIX
2009
DATE: 09 21 09
PAY TO THE ORDER OF BellKor's Pragmatic Chaos
$ 1,000,000 00
AMOUNT: ONE MILLION 00/100
FOR The Netflix Prize
Reed Hastings

How much time would it take to build from scratch:

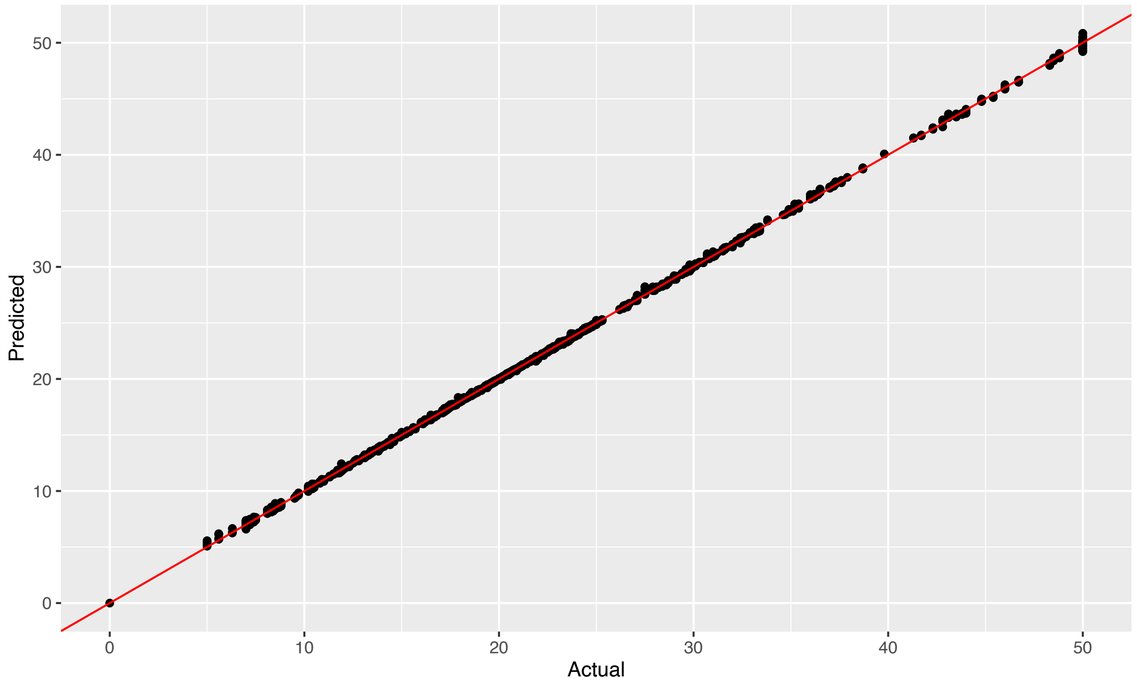
- A set of 18 individual models and 14 ensembles of models
- Automatically optimize all models which can be optimized
- Randomly resample the data as many times as the user requests
- Automatically produce all EDA, summary plots and reports
- Automatically get results that beat the best Kaggle results for 20 years
- Automatically run without any errors or warnings
- Be fully reproducible

Best results with NumericEnsembles package beats best results on Kaggle data science competitions for Boston Housing data

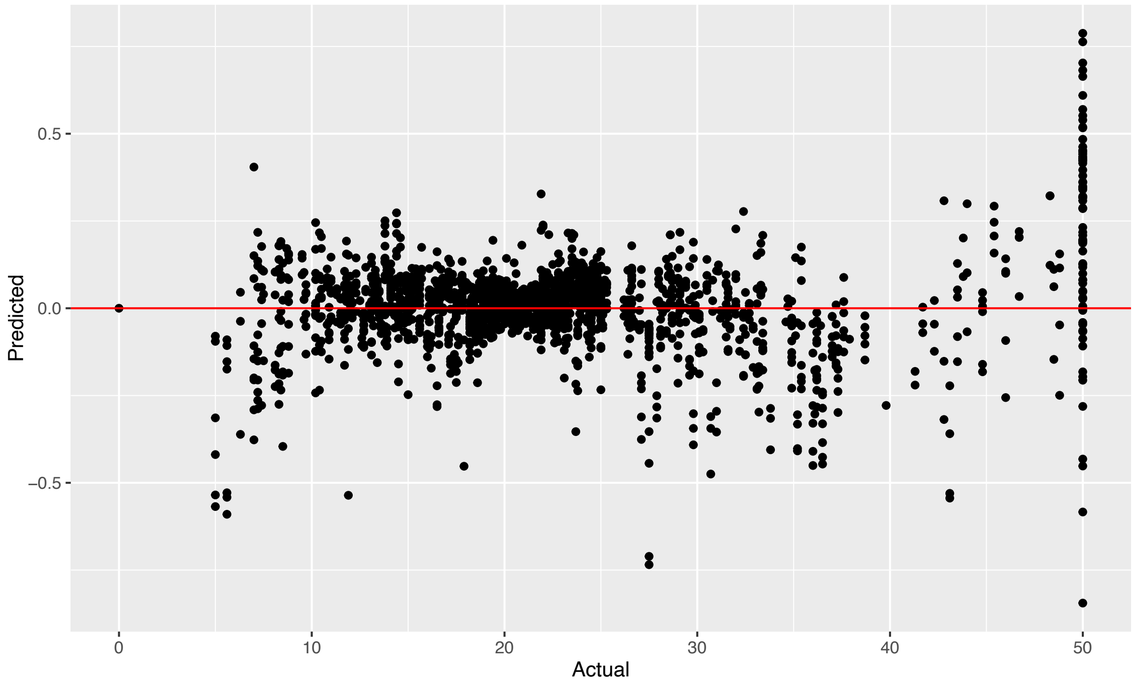
Kaggle Contest Name	Best Score (lowest RMSE)	# of teams	# of entries
Reitaku University	0.80946	12	32
Veronica	1.78911	41	328
UOU G03784	2.41242	49	699
SC201 June 2004	2.56893	17	402
SC201 Oct 2024	3.05221	10	220
Dupanya	3.09643	18	63
Total		147	1744
NumericEnsembles	Live demo right now	1	1

Live demo #2

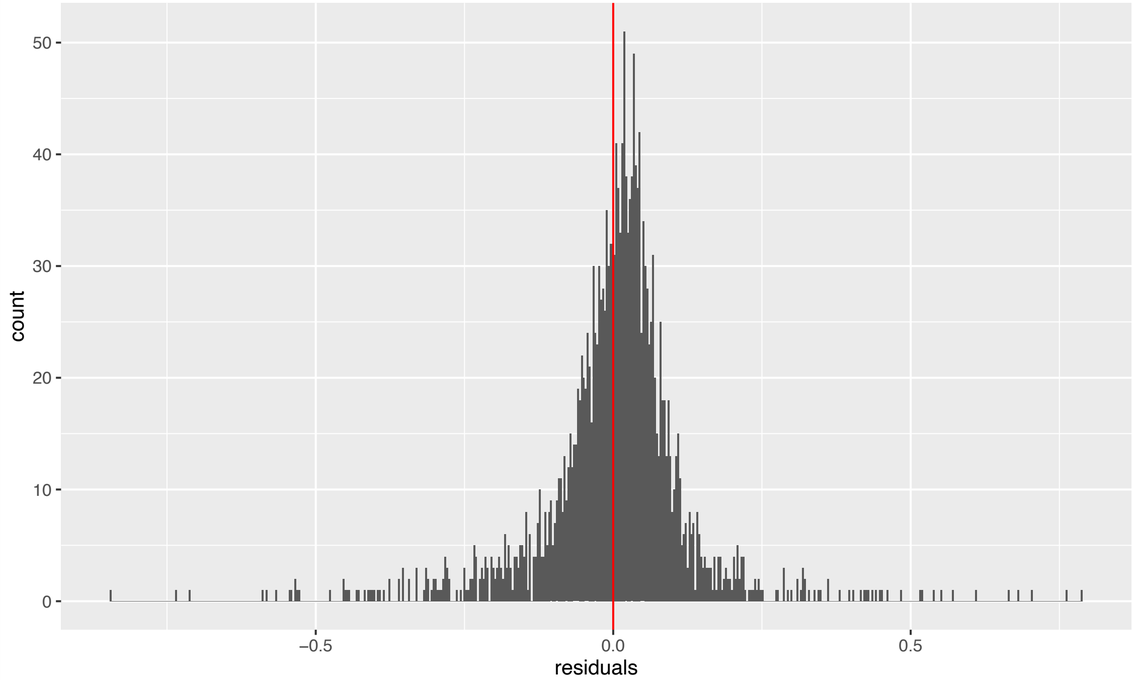
Ensemble Earth model: Predicted vs actual



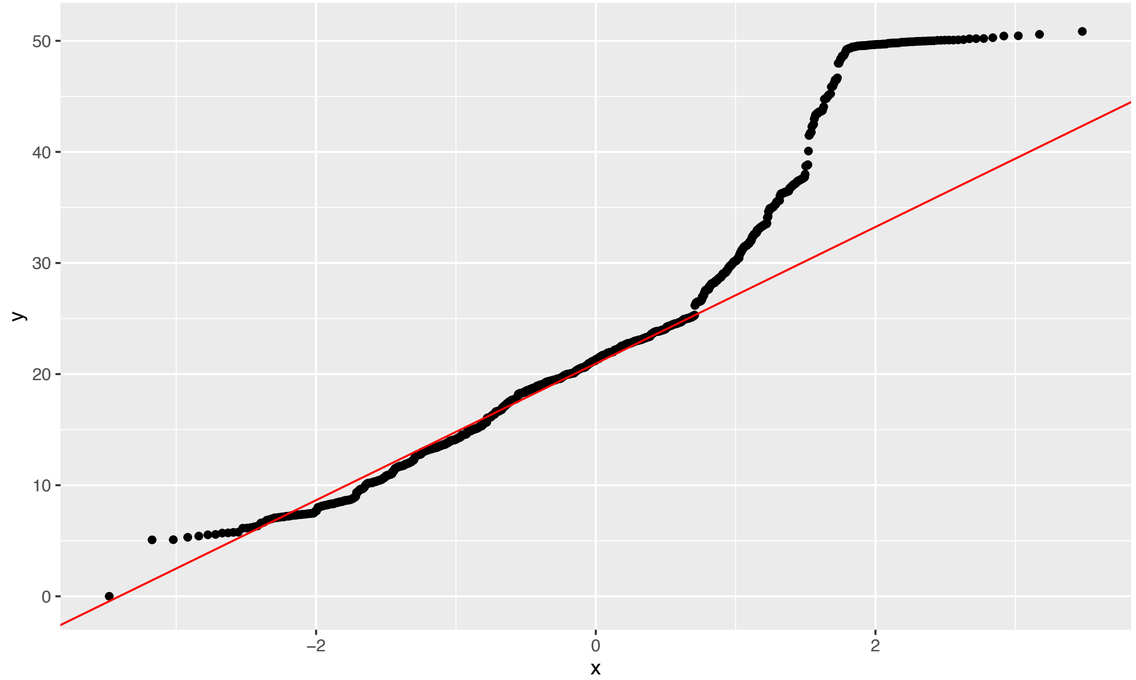
Ensemble Earth model: Residuals



Ensemble Earth model: Histogram of residuals



Ensemble Earth model: Q–Q plot



First (of many) pro features: How the package helps you write a professional paper using the NumericEnsembles package and helps you tell the story in your data.

Live Demo #3

Addressing the most difficult (and important) problem in data science today using the NumericEnsembles package.

Replication crisis

18 languages

Article Talk | Read Edit View history Tools

From Wikipedia, the free encyclopedia

The **replication crisis**, also known as the **reproducibility** or **replicability crisis**, is the growing number of published scientific results that other researchers have been unable to reproduce. Because the reproducibility of empirical results is a cornerstone of the scientific method,[2] such failures undermine the credibility of theories that build on them and can call into question substantial parts of scientific knowledge.

The replication crisis is frequently discussed in relation to psychology and medicine, wherein considerable efforts have been undertaken to reinvestigate the results of classic studies to determine whether they are reliable, and if they turn out not to be, the reasons for the failure.[3][4] Data strongly indicate that other natural and social sciences are also affected.[5]

The phrase "replication crisis" was coined in the early 2010s as part of a growing awareness of the problem.[6] Considerations of causes and remedies have given rise to a new scientific discipline known as metascience,[7] which uses methods of empirical research to examine empirical research practice.[8]

Considerations about reproducibility can be placed into two categories. *Reproducibility* in a narrow sense refers to reexamining and validating the analysis of a given set of data. The second category, *replication*, involves repeating an existing experiment or study with new, independent data to verify the original conclusions.

Open access, freely available online

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. In this essay, I discuss the implications of these problems for the conduct and interpretation of research.

Published research findings are sometimes refuted by subsequent evidence, with ensuing confusion and disappointment. Refutation and controversy is seen across the range of research designs, from clinical trials and traditional epidemiological studies [1–3] to the most modern molecular research [4,5]. There is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims [6–8]. However, this should not be surprising. It can be proven that most claimed research findings are false. Here I will examine the key factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a *p*-value less than 0.05. Research is not most appropriately represented and summarized by *p*-values, but, unfortunately, there is a widespread notion that medical research articles

It can be proven that most claimed research findings are false.

should be interpreted based only on *p*-values. Research findings are defined here as any relationship reaching formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. "Negative" research is also very useful. "Negative" is actually a misnomer, and the misinterpretation is widespread. However, here we will target relationships that investigators claim exist, rather than null findings.

As has been shown previously, the probability that a research finding is indeed true depends on the prior probability of it being true (before doing the study), the statistical power of the study, and the level of statistical significance [10,11]. Consider a 2 × 2 table in which research findings are compared against the gold standard of true relationships in a scientific field. In a research field both true and false hypotheses can be made about the presence of relationships. Let *R* be the ratio of the number of "true relationships" to "no relationships" among those tested in the field. *R*

is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R + 1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, α. Assuming that relationships are being probed in the field, the expected values of the 2 × 2 table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Wacholder et al. have called the false positive report probability [10]. According to the 2 × 2 table, one gets PPV = $(1 - \beta)R/(R - \beta R + \alpha)$. A research finding is thus

Citation: Ioannidis JPA (2005) Why most published research findings are false. PLoS Med 2(8): e124.

Copyright: © 2005 John P. A. Ioannidis. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abbreviations: PPV, positive predictive value

John P. A. Ioannidis is in the Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, and Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts-New England Medical Center, Tufts University School of Medicine, Boston, Massachusetts, United States of America. E-mail: jioannid@cc.uoi.gr

Competing Interests: The author has declared that no competing interests exist.

DOI: 10.1371/journal.pmed.0020124

PLoS Medicine | www.plosmedicine.org 0696 August 2005 | Volume 2 | Issue 8 | e124

Ioannidis (2005): "Why Most Published Research Findings Are False".[1]

Live demo #4

The vision: As easy to use as an LLM, but it’s for data.

The packages are designed to be:

Accurate
Fast
Easy to use

The goal is to make ensembles available to everyone.

No trackers, no cache, no saved data, no shared data, no LLMs, no agents, no coding assistants, no calls to any other systems.

Everything is done by the package on your local machine.

‘All truths are easy to understand, once they are discovered; the point is, to discover them.’

Galileo Galilei

Dialogue Concerning the two Chief World Systems (1632) ‘The Second Day’ tr. Stillman Drake

Solve real problems and/or make great opportunities with Ensembles for Everyone:

NumericEnsembles

ClassificationEnsembles

LogisticEnsembles

ForecastingEnsembles

All packages published on CRAN

Message me/feedback/collaborate, etc:

russconte@mac.com

Github (including slides) : InfiniteCuriosity

Russ Conte, October 28, 2025 for INFORMS Annual Meeting in Atlanta, Georgia, USA