STATISTICAL ANALYSIS AND DATA MINING
Original Article

CASOS: a subspace method for anomaly detection in high dimensional astronomical databases

Marc Henrion

Corresponding Author

E-mail address:marc.henrion03@imperial.ac.uk

Department of Mathematics, Imperial College London, London SW7 2AZ, UK

Department of Mathematics, Imperial College London, London SW7 2AZ, UK
Search for more papers by this author
David J. Hand

Department of Mathematics, Imperial College London, London SW7 2AZ, UK

Search for more papers by this author
Axel Gandy

Department of Mathematics, Imperial College London, London SW7 2AZ, UK

Search for more papers by this author
Daniel J. Mortlock

Department of Mathematics, Imperial College London, London SW7 2AZ, UK

Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK

Search for more papers by this author
Marc Henrion

Corresponding Author

E-mail address:marc.henrion03@imperial.ac.uk

Department of Mathematics, Imperial College London, London SW7 2AZ, UK

Department of Mathematics, Imperial College London, London SW7 2AZ, UK
Search for more papers by this author
David J. Hand

Department of Mathematics, Imperial College London, London SW7 2AZ, UK

Search for more papers by this author
Axel Gandy

Department of Mathematics, Imperial College London, London SW7 2AZ, UK

Search for more papers by this author
Daniel J. Mortlock

Department of Mathematics, Imperial College London, London SW7 2AZ, UK

Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK

Search for more papers by this author
First published: 06 November 2012
https://doi.org/10.1002/sam.11167
Cited by: 2

Password

Enter your password

Forgot password?

Log In

NEW USER > Purchase instant access

- Recommend to a librarian

PDF

PDF Tools

- Request permission
- Export citation
- Add to favorites
- Track citation

Share

Give access

Share full text access

Share full text access

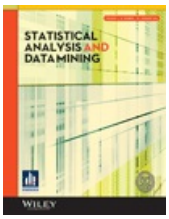
Please review our Terms and Conditions of Use and check box below to share full-text version of article.

I have

Abstract

We develop a novel algorithm for detecting anomalies. Our method has been developed to suit the challenging task of detecting anomalous sources in cross-matched astronomical survey data. Our algorithm computes anomaly scores in lower-dimensional subspaces of the data. By subspaces we mean, in this work, subsets of the original data variables. Our technique presents several advantages over existing methods: it can work directly on data with missing values; it addresses some of the problems posed by high-dimensional data spaces; it is less susceptible to a masking effect from irrelevant features; it can be easily adapted to suit specific needs and it allows an easier interpretation of why a given object has a high combined anomaly score. One drawback of our method is that it cannot detect outliers that are only apparent in high-dimensional spaces. Anomaly scores are computed using a nearest neighbor (NN) technique, but the algorithm works with any other method computing numerical anomaly scores. We demonstrate the properties of our algorithm and evaluate its performance on both simulated and real datasets. We show that it is capable of outperforming state-of-the-art, full-dimensional approaches in some situations. © 2013 Wiley Periodicals, Inc. Statistical Analysis and Data Mining 6: 53–72, 2013

Citing Literature



Volume6, Issue1

Special Issue: Special Issue on Astrostatistics

January 2013

Pages 53-72

- Access
- Related
- Information

Caption

Additional links

About Wiley Online Library

- Privacy Policy
- Terms of Use
- Cookies
- Accessibility

Help & Support

- Contact Us

Opportunities

- Subscription Agents
- Advertisers & Corporate Partners

Connect with Wiley

- The Wiley Network
- Wiley Press Room

Password

Enter your password

Forgot password?

Log In

NEW USER > INSTITUTIONAL LOGIN >

Wiley Online Library

Change Password

Old Password

New Password

Too Short Weak Medium Strong Very Strong Too Long

Submit

Congrats!

Your password has been changed

Create a new account

Email or Customer ID

Register

Returning user

Wiley Online Library

Forgot your password?

Enter your email address below. If your address has been previously registered, you will receive an email with instructions on how to reset your password. If you don't receive an email, you should register as a new user

Email or Customer ID

Enter your email

RESET PASSWORD

Please check your email for your password reset instructions.

Request Username

Can't sign in? Forgot your username?

Enter your email address below and we will send you your username

Email or Customer ID

Submit

Close

If the address matches an existing account you will receive an email with instructions to retrieve your username