

[Home](#) > [SN Computer Science](#) > Article

Original Research | [Published: 01 April 2023](#)

Machine Learning Based Approach for Future Prediction of Authors in Research Academics

[Sovan Bhattacharya](#) , [Ayan Banerjee](#), [Abhik Goswami](#), [Subrata Nandi](#) & [Dinesh Kumar Pradhan](#)

[SN Computer Science](#) **4**, Article number: 306 (2023)

79 Accesses | [Metrics](#)

Abstract

In the present time, it is important to predict superior authors from the huge research community, which determines their performance with their future prospects and opportunities, in the field of scientific research.

There is a need to solve the authors' future prediction problem for equally assessing the performance of the researchers that helps to improve research quality highly, influence other types of research, that will also help to identify the research carrier and other research parameters, that will further influence budding researchers. In our proposed model, to solve the authors' future prediction problems, five machines learning models i.e., SVM, Logistic Regression, Naive Bayes, Decision Tree, and Random Forest are implemented on a data set of 2750 authors where 1000 authors are the ranked authors. So, we have collected 2750 authors' data

from their bibliographic data sets. Then used feature engineering and feature scaling to build the desired model and prepare the final data set used to build the model. There after varied data sizes and training ratios against the five ML models and measured all the evaluation metrics in different cases. We have seen that the Decision Tree and Random Forest model outperform the other models with an accuracy score of 1.0 and the other scores also performed well.

This is a preview of subscription content, [access via your institution](#).

Access options

Buy article PDF

39,95 €

Price includes VAT (India)

Instant access to the full article PDF.

[Rent this article via DeepDyve.](#)

[Learn more about Institutional subscriptions](#)

Notes

1. <https://www.adscientificindex.com/top-100-scientist/>.
2. <https://lfs.aminer.cn/lab-datasets/citation/acm.v9.zip>.

References

1. Ding Y, Yan E, Frazho A, Caverlee J. Pagerank for ranking authors in co-citation networks. *J Am Soc Inf Sci Technol*. 2009;60(11):2229–43.
 2. Zhao P, Han J, Sun Y. P-rank: a comprehensive structural similarity measure over information networks. In: *Proceedings of the 18th ACM conference on information and knowledge management*. 2009; p. 553–562.
 3. Gollapalli SD, Mitra P, Giles CL. Ranking authors in digital libraries. In: *Proceedings of the 11th annual international ACM/IEEE joint conference on digital libraries*. 2011; p. 251–254.
 4. Senanayake U, Piraveenan M, Zomaya AY. The p-index: ranking scientists using network dynamics. *Procedia Comput Sci*. 2014;29:465–77.
 5. Liu Z, Huang H, Wei X, Mao X. Tri-rank: an authority ranking framework in heterogeneous academic networks by mutual reinforce. In: *2014 IEEE 26th international conference on tools with artificial intelligence*. IEEE; 2014. p. 493–500.
-

6. Pradhan D, Paul PS, Maheswari U, Nandi S, Chakraborty T. C3-index: revisiting author's performance measure. In: Proceedings of the 8th ACM conference on web science. 2016; p. 318–319.

7. Franceschet M, Colavizza G. Timerank: a dynamic approach to rate scholars using citations. *J Informetr.* 2017;11(4):1128–41.

8. Zhang J, Ning Z, Bai X, Kong X, Zhou J, Xia F. Exploring time factors in measuring the scientific impact of scholars. *Scientometrics.* 2017;112(3):1301–21.

9. Zhang J, Hu Y, Ning Z, Tolba A, Elashkar E, Xia F. AIRank: author impact ranking through positions in collaboration networks. *Complexity.* 2018;2018:4697485.

10. Zhang C, Yu L, Zhang X, Chawla NV. Task-guided and semantic-aware ranking for academic author-paper correlation inference. *International Joint Conferences on Artificial Intelligence.* 2018.

11. Bornmann L, Leydesdorff L, Wang J. How to improve the prediction based on citation impact percentiles for years shortly after the publication date? *J Informetr.* 2014;8(1):175–80.

12. Varlamis PGTG. I detecting rising stars in dynamic collaborative networks. *J Informetr.* 2017;11(1):198.

13. Daud A, Aljohani NR, Abbasi RA, Rafique Z, Amjad T, Dawood H, Alyoubi KH. Finding rising stars in co-author networks via weighted mutual influence. In: *Proceedings of the 26th international conference on world wide web companion.* 2017; p. 33–41.

14. Amjad T, Daud A, Aljohani NR. Ranking authors in academic social networks: a survey. *Library Hi Tech.* 2018.

15. Zhu L, Zhu D, Wang X, Cunningham SW, Wang Z. An integrated solution for detecting rising technology stars in co-inventor networks. *Scientometrics.* 2019;121(1):137–72.

16. Zhang F, Bai X, Lee I. Author impact: evaluations, predictions, and challenges. *IEEE Access.* 2019;7:38657–69.

17. Wang M, Zhang J, Jiao S, Zhang T. Evaluating the impact of citations of articles based on knowledge flow patterns hidden in the citations. *PLoS One.* 2019;14(11):0225276.

18. Daud A, Song M, Hayat MK, Amjad T, Abbasi RA, Dawood H, Ghani A. Finding rising stars in bibliometric networks. *Scientometrics*. 2020;124(1):633–61.
-
19. Amjad T, Munir J. Investigating the impact of collaboration with authority authors: a case study of bibliographic data in field of philosophy. *Scientometrics*. 2021;126(5):4333–53.
-
20. Xiao S, Yan J, Li C, Jin B, Wang X, Yang X, Chu SM, Zha H. On modeling and predicting individual paper citation count over time. In: *Ijcai*. 2016; p. 2676–2682.
-
21. Ashraf S, Iqbal HR, Nawab RMA. Cross-genre author profile prediction using stylometry-based approach. In: *CLEF (Working Notes)*. Citeseer; 2016. p. 992–999.
-
22. Reddy TR, Vardhan BV, Reddy PV. Author profile prediction using pivoted unique term normalization. *Indian J Sci Technol*. 2016; 9(46).
<https://doi.org/10.17485/ijst/2016/v9i46/99404>.
-
23. Nie Y, Zhu Y, Lin Q, Zhang S, Shi P, Niu Z. Academic rising star prediction via scholar's evaluation model and machine learning techniques. *Scientometrics*. 2019;120(2):461–76.
-

24. Bin-Obaidellah O, Al-Fagih AE. Scientometric indicators and machine learning-based models for predicting rising stars in academia. In: 2019 7th international conference on smart computing & communications (ICSCC). IEEE; 2019. p. 1–7.

25. Bütün E, Kaya M. Predicting citation count of scientists as a link prediction problem. IEEE Trans Cybern. 2019;50(10):4518–29.

26. Shoaib M, Daud A, Amjad T. Author name disambiguation in bibliographic databases: a survey. arXiv preprint [arXiv:2004.06391](https://arxiv.org/abs/2004.06391). 2020.

27. Su Z. Prediction of future citation count with machine learning and neural network. In: 2020 Asia-Pacific conference on image processing, electronics and computers (IPEC). IEEE; 2020. p. 101–104.

28. Fujita M, Inoue H, Terano T. Analyzing promising researchers using network centralities of co-authorship networks from academic literature. New Gener Comput. 2021;39(1):181–97.

29. Khan ZY, Niu Z, Sandiwarno S, Prince R. Deep learning techniques for rating prediction: a survey of the state-of-the-art. Artif Intell Rev. 2021;54(1):95–135.

30. Daud A, ul Islam N, Li X, Razzak I, Hayat MK. Identifying rising stars via supervised machine learning. *IEEE Trans Comput Soc Syst.* 2022. <https://doi.org/10.1109/TCSS.2022.3178070>.

31. Ávila-Toscano JH, Romero-Pérez I, Saavedra-Guajardo E, Marenco-Escuderos A. Determinants of colombian scientific production in social sciences articles indexed in wos, scopus and other databases: tree of classification and regression. *Rev Interam Bibliotecol.* 2022;45(1). <https://doi.org/10.17533/udea.rib.v45n1e339712>.

32. Bhattacharya S. Discoveries of research genealogy from large-scale academic dataset: issues, challenges and application. *Int J Comput Sci Eng.* 2019;07:262–7.

33. Wang X, van Harmelen F, Huang Z. Recommending scientific datasets using author networks in ensemble methods. 2022.

34. Nawaz A, Malik M. Rising stars prediction in reviewer network. *Electron Commer Res.* 2022;22(1):53–75.

35. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci.* 2005;102(46):16569–72.

Funding

We do not have any funding for this article.

Author information

Sovan Bhattacharya, Ayan Banerjee, Abhik Goswami, Subrata Nandi and Dinesh Kumar Pradhan contributed equally to this work.

Authors and Affiliations

Department of CSE, Dr. B. C. Roy Engineering College, Jemua Road, Durgapur, West Bengal, 713206, India

Sovan Bhattacharya, Ayan Banerjee, Abhik Goswami & Dinesh Kumar Pradhan

Department of CSE, National Institute of Technology, Mahatma Gandhi Road, Durgapur, West Bengal, 713209, India

Sovan Bhattacharya & Subrata Nandi

Contributions

All authors have equal contributions to the article.

Corresponding author

Correspondence to [Sovan Bhattacharya](#).

Ethics declarations

Conflict of interest

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Ethics approval

The Author confirms that the work described has not been published before, and it is not under consideration for publication elsewhere.

Consent to participate

The Authors agree to participate in the conference.

Additional information

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This article is part of the topical collection “Enabling Innovative Computational Intelligence Technologies for IOT” guest edited by Omer Rana, Rajiv Misra, Alexander Pfeiffer, Luigi Troiano and Nishtha Kesswani.

Rights and permissions

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

[Reprints and Permissions](#)

About this article

Cite this article

Bhattacharya, S., Banerjee, A., Goswami, A. *et al.* Machine Learning Based Approach for Future Prediction of Authors in Research Academics. *SN COMPUT. SCI.* **4**, 306 (2023).

<https://doi.org/10.1007/s42979-023-01692-6>

Received

Accepted

Published

30 September 2022

09 January 2023

01 April 2023

DOI

Keywords

Machine learning

Citation networks

Bibliographic information

Author impact evaluation

Author future prediction

Rising star prediction