Calculator for predicting the probability of faculty promotion in an academic medical centre

May May Yeo [1], Shih-Hui Lim,[1] Anshul Kumar,[2] Anne W Thompson[2]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/leader-2023-000861).

[1]Office of Academic Medicine, Duke-NUS Medical School, Singapore
[2]Health Professions Education, MGH Institute of Health Professions, Boston, Massachusetts, USA

Correspondence to
Dr May May Yeo; gmsymm@nus.edu.sg

Received 23 June 2023
Accepted 9 September 2023
Published Online First 6 October 2023

ABSTRACT

Objective The academic medical centre (AMC), with over 2200 faculty members, annually manages approximately 300 appointments and promotions. Considering these large numbers, we explored whether machine learning could predict the probability of obtaining promotional approvals.

Methods We examined variables related to academic promotion using predictive analytical methods. The data included candidates' publications, the H-index, educational contributions and leadership or service within and outside the AMC.

Results Of the five methods employed, the random forest algorithm was identified as the 'best' model through our leave-one-out cross-validation model evaluation process.

Conclusions To the best of our knowledge, this is the first study on the AMC. The developed model can be deployed as a 'calculator' to evaluate faculty performance and assist applicants in understanding their chances of promotion based on historical data. Furthermore, it can act as a guide for tenure and promotion committees in candidate review processes. This increases the transparency of the promotion process and aligns faculty aspirations with the AMC's mission and vision. It is possible for other researchers to adopt the algorithms from our analysis and apply them to their data.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Advances in machine learning, a subset of artificial intelligence, have been explored to improve patient care quality, reduce clinician workload and revolutionise the practice of medicine. Concurrently, interest in the application of algorithms when making hiring decisions is growing, and studies have documented the association of academic productivity factors with the promotion or attainment of senior academic ranks.

WHAT THIS STUDY ADDS

⇒ Our best algorithms yielded the predicted probabilities of faculty promotion to a high degree of accuracy.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ We created a tool to predict the probability of faculty promotions using data on faculty research, education, services and leadership contributions. This tool provides evidence-based guidance for candidates seeking academic promotions. It aids the leadership, tenure and promotion committee members within academic medical centres in reviewing and advising aspiring candidates for promotion.

INTRODUCTION

The application of algorithms in making hiring decisions has attracted considerable interest.[1] Chalfin *et al*[2] used machine learning to predict worker productivity in hiring decisions, including teacher tenure decisions. Punnoose and Ajit predicted employee turnover by using machine learning algorithms. They cited several studies that predicted the same regarding employee turnover.[3] Despite these developments, there is a paucity of literature on the application of algorithms to tenure and promotion decisions in an academic medical centre (AMC). We aimed to address this gap by answering the following research question: Can the probability of senior faculty appointment or promotion be predicted with reasonable and significant accuracy using faculty research, education, service and leadership contributions within an AMC? We examined the variables related to these contributions using predictive analytical methods.

Gennatas and Chen termed machine learning as 'procedures that allow learning from data' (Gennatas and Chen, p.7).[4] Chalfin *et al* argued that prediction via machine learning was better than the traditional causal inference approach as it could '… trade off bias and variance to maximise out-of-sample prediction accuracy' (Chalfin *et al*, p.124).[2]

As an extension of Chalfin *et al*'s premise, we sought to understand whether machine learning could help predict the probability of obtaining academic promotion approval in applications received by the faculty affairs department of the AMC. This could form the basis of a 'tenure and promotion calculator' that faculty candidates could use to evaluate their probability of being promoted. This tool can provide evidence-informed decision-making support for tenure and promotion committees. The output from the calculator could be combined with an additional advisory programmed into the calculator or mentoring sessions with the faculty preparing for promotion.

Fox and Bunton advocated clear articulation and consistent application of tenure and promotion policies in strategic talent management to meet imminent national healthcare needs.[5] Our tool facilitates transparency and consistency in the application of tenure and promotion policies in an AMC.

Check for updates

To cite: Yeo MM, Lim S-H, Kumar A, *et al*. *BMJ Leader* 2024;**8**:203–207.

Table 1 Criteria for academic appointment and promotion of faculty (adapted from the AMC's website)

Criteria	Academic level			
	Instructor	Assistant professor	Associate professor	Professor
Hospital grade	Associate consultant	Associate consultant*/consultant/ senior consultant	Consultant/senior consultant	Senior consultant
Reputation	Not specified	Not specified	National/regional/international	Regional or international (clinical); international (academic/scientific)
Publications in peer-reviewed journals			30 or more; < 30 if they have other significant credentials.	75 or more; <75 if they have other significant credentials.
Grant funding†			Sustained funding, including major national grants as a principal investigator (PI)/co-principal investigator (Co-PI) for research pathways.	Sustained funding, including major national grants as PI/Co-PI for research pathways.
Education and/or service/ leadership‡			Significant contributions in education and/or service/leadership. Well-developed education portfolio for education pathways.	Significant contributions in education and/or service/leadership. Well-developed education portfolio for education pathways.

*Associate consultants with 15 or more publications may be promoted to the clinical assistant professor level.
†Those in the research pathway or pursuing research.
‡Service/leadership includes membership in the AMC, national committees, medical associations, national specialist accreditation boards and national and international journal editorial boards. However, this list is not exhaustive.
AMC, academic medical centre.

MATERIALS AND METHODS

Dataset

The faculty structure of the AMC includes regular and clinical ranks, adjuncts, and visiting appointments. Regular rank tracks comprise research and educational pathways. Clinical rank appointments are accorded to clinicians who focus on clinical/ medical care within the AMC, though they may have quantifiable academic or leadership involvement. Both regular and clinical rank appointments have four levels: instructor, assistant professor, associate professor and professor. This study focused on (1) regular non-tenure-track appointments and (2) clinical associate professors and professor appointments. Accordingly, the dataset comprises candidates who were nominated for promotion to associate professors, professors, clinical associate professors and clinical professors between January 2018 and June 2021.

Promotion from instructor to assistant professor was not studied as associate consultants and consultants could obtain appointments as instructors and assistant professors, respectively. The promotion criteria and pathways are summarised in table 1. The AMC promotion process is illustrated in figure 1.

In total, 255 candidates were presented for promotion (units of observation). The 21 independent variables are listed in table 2, including a relevant explanation. The independent variables included the candidates' information, such as department and title to be promoted, research information, including number of publications in peer-reviewed journals and grant funding, teaching contributions, and leadership or services roles. The dependent variable was promotion approval.

Tools

The R software was used for all analyses.[6] We used cross-validation to rigorously test the predictions of five algorithms: standard random forest, standard logistic regression, standard K nearest neighbours (KNN), standard KNN regression and 'adaptive minimum match' KNN.[7,8]

We ascribed Browne's philosophy that it is more desirable to compare and choose from a set of competing models than to decide whether a single model is good (Browne, p.110).[9] Therefore, we presented five algorithms or models for which we selected the 'best' as explained in the Model evaluation criteria section.

Model evaluation

The leave-one-out cross-validation method yields a predicted probability of promotion for each nomination that the computer calculates without knowing the true outcome of the nomination

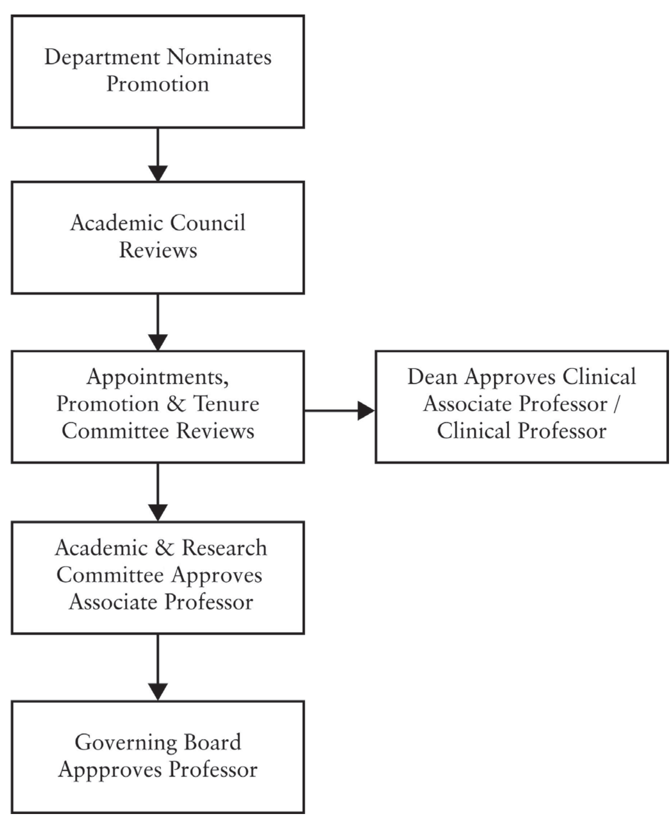

Figure 1 General promotion process of associate professors and professors in the AMC (adapted from the AMC'S website). AMC, academic medical centre.

Table 2 List of variables and description

Variable	Variable	Definition	Type
Dependent variable	Approval	Promotion outcome	Binary (promoted/ not promoted)
Independent variables			
Candidate related	Clinical programme	Candidate department or specialty.	Categorical
	Title	Academic title candidate is promoted to, that is, associate professor, professor clinical associate professor, clinical professor.	Categorical
	Level	Additional variable assigned by authors for grouping of academic titles into four levels by title.	Categorical
	UNI title	Candidate concurrently holds an academic title with a local competitor medical school.	Binary (yes/no)
	Other UNI title	Candidate concurrently holds an academic title with other universities or AMCs.	Binary (yes/no)
	Additional qualifications	Candidate has an advanced degree in addition to a medical degree (eg, Master's of Science in Medical Education, Master's of Business Administration and Ph.D.).	Binary (yes/no)
Research	Publications	Candidate's number of publications in peer-reviewed journals.	Numeric
	Regional international	Candidate's number of publications in regional or international peer-reviewed journals.	Numeric
	First last author	Candidate's number of publications in which he/she is first or senior author.	Numeric
	H-index	H-index of candidate from Google Scholar Citations at the time of promotion review.	Numeric
	FWCI	Field-weighted Citation Impact from SciVAL Copyright 2022 Elsevier at time of promotion review.	Numeric
	Book chapters	Candidate's number of book chapters authored.	Numeric
	Total grants	Sum of grants held by candidate.	Numeric
	Total grants as PI	Sum of grants where candidate is PI.	Numeric
Education	Teaching excellence	Candidate had provided a well-developed education portfolio to demonstrate educational excellence. The portfolio is included for promotion review only when cleared by the academic education institute of the AMC for completeness and effectiveness.	Binary (yes/no)
	International and/or regional teaching	If candidate has conducted international and/or regional teaching.	Binary (yes/no)
	National teaching	If candidate has conducted national level teaching.	Binary (yes/no)
	Undergraduate teaching	If candidate has conducted undergraduate level teaching.	Binary (yes/no)
	Postgraduate teaching	If candidate has conducted postgraduate level teaching.	Binary (yes/no)
Service/leadership	Leadership (AMC)	Candidate holds leadership role within the AMC. These roles may be clinical that is, clerkship coordinator or academic, that is, academic chair or vice chair.	Binary (yes/no)
	Other leadership	Candidate holds leadership or significant roles outside of AMC. These roles include residency programme director, member of national specialist accreditation committee and associate editor of local/international peer-reviewed journal.	Binary (yes/no)

AMC, academic medical centre; PI, principal investigator.

candidate. We classified the predicted probability as 'approved' if it was above a threshold value and 'not Approved' if it was below the threshold value. The threshold value for each algorithm was determined using the criteria described in the Model evaluation criteria section. The predicted outcomes were compared with the actual outcomes for each of the five algorithms. The training of the data using each of the five algorithms is described in online supplemental appendix 1.

Model evaluation criteria

All the models were compared using the following metrics: accuracy, specificity, sensitivity and F1.[10] The evaluation criteria were as follows:

1. True positive: The candidate was promoted and was predicted to be promoted (1).
2. False positive: The candidate was not promoted but was predicted to be promoted (1).
3. True negative: The candidate was not promoted and was predicted as not promoted (0).
4. False negative: The candidate was promoted and predicted as not promoted (0).

For clarity, table 3 lists the confusion matrix used.

It is difficult for the leadership to manage when faculty members have been advised that they would likely be promoted when they were eventually not, that is, a false positive outcome. Therefore, we optimised our model thresholds by reducing the number of false positives to significantly low, and concurrently, the sensitivity or number of true negatives to significantly high. Figure 2 outlines the steps involved in this process.

RESULTS

Table 4 compares the metrics of the five models.

The random forest was the best model, with the highest sensitivity and F1 scores. F1 is the harmonic mean of precision and recall, with a score between 0 and 1, where 1 is a perfect score. A good F1 score indicates a model with low false positives and false negatives, that is, it correctly identifies real threats and not false alarms.

For the best random forest model, we generated predicted probabilities of promotion for each candidate. Thereafter, we grouped the candidates into ranges of predicted probabilities that could be useful as potential candidates, as listed in table 5.

The actual proven rate of 37/(14+37)=0.73 fell within the predicted probability range of 0.7–0.8. Similarly, the actual

Table 3 Confusion matrix for evaluation of model outcome

		Actual result	
		Not promoted (0)	Promoted (1)
Predicted result	Not promoted (0)	True negative	False negative
	Promoted (1)	False positive	True positive

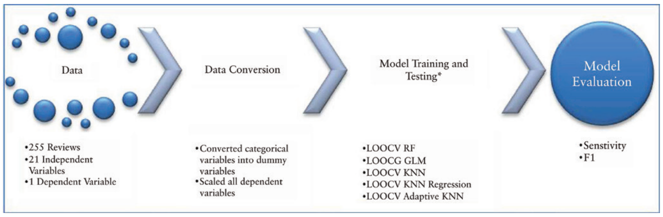

Figure 2 Using machine learning to develop a promotion calculator *Models: LOOCV RF, leave-one-out cross-validation with random forest; LOOCV GLM, leave-one-out cross-validation with logistic regression; LOOCV KNN, Leave-one-out cross-validation with K nearest neighbour classification; LOOCV KNN Regression, Leave-one-out cross validation with K nearest neighbour regression; LOOCV Adaptive KNN, Leave-one-out cross validation with adaptive K nearest neighbour.

proved result of 53/(9+53)=0.85 fell within the predicted approval probability range of 0.8–0.9. Similar results are obtained for the other two ranges. We decided that any prediction under 0.7 is not sufficiently reliable to provide a candidate with a specific probability. Empirically, we find that candidates with predicted probabilities under 0.7 have been promoted 57% of the time in the past. All these probabilities can be communicated transparently to the candidates.

The random forest model yielded the highest sensitivity and F1. Out of 10 candidates who were not promoted, our 'best' model could predict about eight of them. This was important from the school's leadership perspective, as support and guidance could be provided to as many future faculty candidates as possible who were predicted to be unsuccessful in promotion application. Simultaneously, our model minimises the number of false positives, which in turn minimises the encouragement of candidates whose portfolios are not sufficiently robust for promotion.

We further evaluated the model by disaggregating the predictions into four ranges and comparing the predictions in each range with the actual probability. The predicted probability ranges are consistent with the actual results, indicating a high level of accuracy. For example, when a faculty candidate received a predicted probability of promotion from the model, they could understand that historically, others receiving the same prediction had obtained promotions at a similar rate to the predicted probability. This information could facilitate faculties in deciding whether to apply for promotion. Conversely, another faculty member whose probability of being promoted was in a range less than 0.7 could decide to arrange for further guidance with the leadership and faculty affairs department.

Table 4 Evaluation of the five machine learning models in predicting probability of faculty promotion

Model	Accuracy	Sensitivity	Specificity	F1
LOOCV random forest	0.64	0.79	0.60	0.50
LOOCV logistic regression	0.70	0.52	0.76	0.44
LOOCV K nearest neighbours classification	0.53	0.74	0.47	0.42
LOOCV K nearest neighbours regression	0.54	0.60	0.52	0.37
LOOCV adaptive KNN	0.75	0.41	0.85	0.43

KNN, K nearest neighbours; LOOCV, leave-one-out cross-validation.

Table 5 Evaluation of the disaggregated confusion matrix of the best model, that is, random forest

		Actual result		
		Not approved	Approved	Actual rate of approval
Predicted probability of approval	<0.7	32	42	0.57
	0.7–0.8	14	37	0.73
	0.8–0.9	9	53	0.85
	>0.9	3	65	0.96

DISCUSSION

In this study, we created a tool to predict the probability of faculty promotion using data on faculty research, education, leadership and service contributions. We identified the basis for a calculator that could be adopted in practice to aid leadership, tenure and promotion committees, as well as to improve academic progression guidance for the faculty of the AMC.

In developing the calculator, the scenario in which candidates were informed that they would be successful but were eventually not promoted was one that was difficult for the leadership to manage and had to be significantly minimised. This necessitated the use of sensitivity as our main evaluation criterion for identifying the best model, that is, minimising the number of False positives or encouraging candidates whose portfolios were not sufficiently robust and would not be promoted.

Jarrahi called for a new 'human–machine collaboration', where the superior analytical capabilities of artificial intelligence complement human decision makers who possess subconscious decision heuristics to evaluate and facilitate decisions and outcomes.[11] Accordingly, we suggest that our tool can be used to guide faculty seeking promotions. It can help the leadership and faculty affairs departments by providing advice to those who are in need. Tenure and promotion committees can use this as a reference for their evaluation decisions.

Furthermore, it can be used to identify potential cases of bias or discrimination. If a faculty member was not promoted, but the model predicted that he/she had a greater than 90% chance of being promoted, this could raise a flag for the promotion committee to review the application again and the reasons for not promoting the candidate.

Limitations

The algorithm was limited by the nature of the data on which it was trained. The dataset in this study consisted of clinicians from the AMC. Different outcomes would occur when using other datasets, such as those from AMC biomedical researchers. Ongoing refinement of the model is required through the constant input of fresh tenure and promotion review outcomes.

Although the data in the study were collected over 3.5 years, the time variable was excluded to avoid the identification of individual faculty candidates during this period. Thus, it is impossible to associate any changes with the dependent variable (promotional approval) associated with time.

Additionally, the arms-length reference letters required for promotion to associate professor, professor and clinical professor were excluded in the calculator, as they could not be adequately quantified.

Xu *et al* advocated that '… there is no machine substitute for higher-level interactions, critical thinking and ambiguity …'.[12] Similarly, our findings complement but are not substitutes for careful and informed evaluations by candidates seeking academic progression and academic committees in tenure and promotion decisions.

CONCLUSIONS

Through predictive modelling that employed machine learning algorithms, we established the groundwork for a 'promotion calculator' that provides evidence-informed guidance to candidates seeking academic promotion. It aids the leadership, tenure and promotion committee members within an AMC in reviewing and advising aspiring candidates for promotion.

Contributors MMY developed the research idea, performed the analysis and wrote the manuscript. S-HL reviewed and edited the manuscript. AK verified the methods used and critically reviewed the manuscript. AWT provided guidance and support throughout the study. MMY accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval The National University of Singapore Institutional Review Board approved this study (reference code: NUS-IRB-2020–569).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

Author note Within the healthcare system of Singapore, associate consultants and consultants are specialists who have completed residency/fellowship or basic and advanced specialist training and fulfil several conditions outlined in the Singapore Ministry of Health Holdings website, https://www.physician.mohh.com.sg/medicine/medical-service-career-path.

ORCID iD

May May Yeo http://orcid.org/0000-0002-8823-1799

REFERENCES

1 Raghavan M, Barocas S, Kleinberg J, *et al*. Mitigating bias in algorithmic employment screening: evaluating claims and practices. *SSRN Journal* 2020:469–81.
2 Chalfin A, Danieli O, Hillis A, *et al*. Productivity and selection of human capital with machine learning. *American Economic Review* 2016;106:124.
3 Punnoose R, Ajit P. Prediction of employee turnover in organizations using machine learning algorithms. *Ijarai* 2016;5:22–6.
4 Gennatas ED, Chen JH. Artificial intelligence in medicine: past, present, and future. In: *In Artificial intelligence in medicine*. Academic Press, 2021: 3–18.
5 Fox S, Bunton S. *The case for strategic talent management in academic medicine*. Portl State Univ PDXScholar, 2011.
6 R Core Team. *R: A language and environment for statistical computing*. Vienna: R foundation for statistical computing, 2021. Available: https://www.r-project.org/
7 Kumar A, Edwards RA. Adaptivelearnalytics: adaptive predictive learning analytic tools. 2021. Available: https://github.com/readcreate/AdaptiveLearnalytics
8 Kumar A, DiJohnson T, Edwards RA, *et al*. The application of adaptive minimum match K-nearest neighbors to identify at-risk students in health professions education. *J Physician Assist Educ* 2023;34:171–7.
9 Browne MW. Cross-validation methods. *J Math Psychol* 2000;44:108–32.
10 Agarwal R. The 5 classification evaluation metrics every data scientist must know. Towards data science. 2019. Available: https://towardsdatascience.com/the-5-classification-evaluation-metrics-you-must-know-aa97784ff226 [Accessed 15 Mar 2022].
11 Jarrahi MH. Artificial intelligence and the future of work: human-AI symbiosis in organizational decision making. *Business Horizons* 2018;61:577–86.
12 Xu L, Sanders L, Li K, *et al*. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. *JMIR Cancer* 2021;7:e27850.

APPENDIX 1 – FIVE MACHINE LEARNING ALGORITHMS USED TO TRAIN DATASET

Leave-One-out Cross Validation with Random Forest:

We issued a command to run random sampling without replacement such that each observation in the old data was only selected once in the new data. Next, we created a loop to run the random forest model 255 times.

Leave-One-out Cross Validation with Logistic Regression:

We create a loop to run the logistic regression model 255 times.

Leave-One-out Cross Validation with K Nearest Neighbour Classification:

We tested several values of k (number of nearest neighbours) and obtained k=3, which yielded the highest level of sensitivity.

Leave-One-out Cross Validation with K Nearest Neighbour Regression:

We tested several values of k and arrived at k=3, which yielded the highest level of sensitivity.

Leave-One-out Cross Validation with Adaptive KNN:

We used a package developed at the Institute of Health Professions of Massachusetts General Hospital (Boston, MA, and USA).