Theme: Artificial Intelligence in Academic Peer Review

Title: Towards an AI-assisted Peer Review System

Tirthankar Ghosal
Supervisor(s): Dr. Asif Ekbal and Prof. Pushpak Bhattacharyya
Department of Computer Science and Engineering
Indian Institute of Technology Patna
tirthankar.pcs16@iitp.ac.in / +91-9832669697

Awarded the AWSAR-DST award 2018 (within top 100) in popular science story category (PhD)
https://www.awsar-dst.in/results

Declaration

I hereby declare that the following article is an original piece of work and not published anywhere.

THE AUTHOR

"Dear Mr Harsh, Thank you for submitting your manuscript 'Impact of Stress in Long Bars under Boundary Conditions' to the 'Journal of Industrial Machines'. After careful consideration by our editors, we regret to inform you that we must decline this submission on editorial grounds and subsequently have declined to send the paper out to external peer reviewers." ... blinked his email. Another Desk-Rejection!! Harsh let the email sink in slowly. It was his third rejection in a row and that too from the editor's desk. He could not believe that after three months of submission this is the response to his correspondence. The editor did not even consider sending his manuscript to the reviewers. It pained him to realise that not only did the editor dislike his toil but also hated it sufficiently to circumvent its progress through the review process. *What is going wrong with me?* He mumbled *At least he bothered to read it! My guide is not an expert in this field; he is finding very little time for me. What if ...?*

THE EDITOR

"Dear Sir, You would be pleased to know that I have been elevated to the position of Associate Editor in the highly reputed Journal of Industrial Machines" Dr Saha finished typing her email to Director IITZ and Professor-In-Charge Public Relations. *Another one ticked off my bucket list* .... she looked out of her cabin window, smiling. Little did she comprehend that this job could be so demanding? Regular classes, evaluations, meetings, PhD students, invited talks, Professor-In-Charge (Guest House), and now an added responsibility. To go through this huge number of submissions to this popular journal, check their plagiarism, decide relevance, find suitable reviewers, coordinate and write responses … an array of non-escaping tasks. She found most of the submissions are uninformed ones, not falling within the aims and scope of the *Journal of Industrial Machines* and even if they fall, most of them are of average-merit, not warranting further review. *Sometimes I wonder how people could write so much trash and send these irrelevant articles which almost add nothing to the existing body of work. Why don't they just consult the published papers and decide wisely about their work? What if ....?* she sighed in exasperation.

THE REVIEWER

Two deadlines already missed ... Dr Moorthy is having a hard time to manage everything. He recently came back to India after his second postdoc and joined a reputed private university. However, work pressure has taken a toll on him. Too many engagements! and to top it all, there are these pending paper reviews. He is struggling hard to find a spare time to go through this seemingly interesting paper. *What if there was an AI that could read the article and point out significant contributions?* Finally, after the third reminder, Dr Moorthy forced himself to review the paper.

All the three stories, although fictional, are true in the current context, converging to one deep question: *What if there was an Artificial Intelligence (AI) support to the peer review system? An AI to support research evaluation?*

Desk rejection is a common phenomenon in academia, a woe faced by most early career, sometimes even by seasoned researchers. It implies that the journal editor returns a prospective research article to the author without consulting its merit to the expert reviewers. Several reasons account for this activity: plagiarised content, the article not within the scope of the journal, below quality article with respect to the competitive benchmark of the journal, template mismatch, spelling, language and grammar, etc.

The current peer review system is mostly human-centric and is biased sometimes. With the exponential rise of article submissions (better known as the 'Publish or Perish' bubble in academia), it is increasingly becoming difficult for the journal editors to keep up the pace with the latest research, go through each submission and respond to the authors in reasonable time. What if there was an AI which could help the editors to take appropriate decisions by pointing out seemingly out-of-scope, below-quality submissions? What if the AI could relieve the editors from this "burden of science" to some extent?

Our current research is about investigating the role that AI could play in several aspects of the scholarly peer review process. We partnered with a reputed global scientific publishing house to pursue this very timely problem with the goal to ease the information overload on journal editors using Machine Learning and Natural Language Processing techniques. A system of such a kind could also help the authors to choose the journals wisely and retrospect on the quality of his/her paper according to the journal standards. An ambitious vision of this project is to help the reviewers identify the novel aspects of a proposed research. It is now somewhat impossible for a human to go through the massive volume of interdisciplinary research available. The need of the hour is to develop automated solutions for relevant literature discovery. We believe that the

progress of this investigation at any stage could lower down the average turnaround response time of a journal, thus speeding up the overall process of peer review.

We begin with investigating the general causes of desk rejection from the author-editor-reviewer interactions made available to us by our industry partner. Upon in-depth analysis of rejection comments and rejected papers, we found that more than 50% of desk-rejection accounts for the paper being *'out-of-scope'*. Inspite of having merit, sometimes an editor is left with no other choice than to reject because the paper won't find a reader among the audience of the particular journal. So we took up this problem and viewed it as classifying a paper as 'In-Scope' or 'Out-of-Scope' using Machine Learning techniques. Our seed idea was *"information contained in the accepted and published papers of a journal are the benchmark of reference which defines the domain of operation of the journal"*. We incorporate features extracted from almost all possible sections of a manuscript that may contribute to determining its belongingness to the journal concerned. We consider extracting keywords, topics from the full-text portions, clustering in-scope articles, author activity in the past five years, bibliographic information, etc. with respect to the published articles of the journal as our features. Our approach proved highly accurate and was able to outperform a popular state-of-the-art journal recommender by a relative margin of 37% for one journal. Thus with our method, a system could be developed to assist the editors and authors to identify out-of-scope submissions effectively.

Our further analysis of desk-rejection comments revealed that the editors are also concerned about the quality of research not matching to journal standards. We take a very simplistic view here: *'Good papers cite good papers'*. We look into the bibliography section and see how many influential papers are cited and where are the cited references published (reputed venues generally publish significant contributions). We take the citation counts of the references, reputation of venues (Impact Factor, CORE Rankings), the temporal distance of the citations (too many old citations may indicate that the authors are not aware of current state-of-the-art), the presence of mathematical content, etc. as our quality features. We also identify which citations

are influential to the current paper and which are[1] just incidental and appropriately adjust the corresponding feature weights. We found that the inclusion of quality features into our system greatly enhances its capability to identify desk-rejections. Although there are other factors which are basic checklists at the editor's desk like plagiarism, spelling, language and grammar, template mismatch, etc., we found proprietary state-of-the-art tools available to address them. However, these factors could make our proposed system more accurate.

Spending almost three years with the problem we are now pretty convinced that this is a problem of epic proportions, having multiple layers of investigation and require significant collaborative efforts. It is more like *doing science over science*. Designing an AI-based peer review system is just the surface form of this investigation. Connecting the ever-expanding human scientific knowledge (which manifests in the form of research papers) and translating it to a machine-understandable form is the grand vision. However, there are several practical challenges: to get hold of knowledge (research papers) that are behind paywalls, enforcing a more knowledge discoverable format of research papers rather than just PDFs, annotating scholarly artifacts, etc. With the rise of open access movements, significant technical efforts by AI companies like Chan Zuckerberg Initiative-Meta, Allen Institute of Artificial Intelligence, communities like FORCE11, we are hopeful that we are moving in the right direction.

Understanding novel scientific knowledge by an AI is a far stretched vision. However, with the ever-growing prowess of AI, in the near future, we could think of *helping THE REVIEWER to identify the NEW* as well as to validate it. Maybe someday we can have an AI scientist as our peer, *as a gatekeeper of scientific knowledge and wisdom*.

1