How to do Good Data Mining Research and get it Published in Top Venues

Eamonn Keogh
University of California, Riverside
Riverside, California
eamonn@cs.ucr.edu

***Abstract*—While ICDM has traditionally enjoyed an unusually high quality of reviewing, there is no doubt that publishing in ICDM is very challenging. In this tutorial Dr. Keogh will demonstrate some simple ideas to enhance the probability of success in getting your paper published in a top data mining conference; and after the work is published, getting it highly cited.**

Keywords-component; Research, Experimenation;

I. INTRODUCTION

While ICDM has traditionally enjoyed an unusually high quality of reviewing, there is no doubt that publishing in ICDM (and other high quality data mining conferences) is very challenging. This is especially true for young faculty, grad students whose primary advisor is not an experienced ICDM author, or people from outside the community (i.e. a biologist or mathematician who has a result that might greatly interest the database/data mining community).

In this tutorial Dr. Keogh will demonstrate some simple ideas to enhance the probability of success in getting your paper published in a top data mining conference; and after the work is published, getting it highly cited.

These tips and tricks are based on 12 years experience as a prolific author and reviewer, and wisdom solicited from many of the most prolific data mining researchers/reviewers.

Topics covered in the tutorial include:

- Finding the right problems to work on (80% of the battle).
- Don't summarize, sell! Writing abstracts that put the reviewer on your side from the start.
- Getting or creating the perfect dataset.
- Experiments that tell a story.
- Making effective and interesting figures.
- Getting the reviewers on your side.
- The top-ten avoidable reasons why papers get rejected from ICDM.
- Three simple tricks to increase the number of citations to your work.

While Dr. Keogh does not claim to have a "magic bullet" for publishing in database/data mining, his significant track record of publishing in top data mining venues, combined with extensive (and deliberately uncredited) experience in helping younger researchers "break-in" to ICDE/VLDB/SIGKDD have placed him in a unique position to share useful and actionable advice.

While writing this tutorial Dr. Keogh, sought and received advice from many respected data mining researchers, their advice is incorporated into this tutorial. Their advice is framed both positively as researchers ("*A good idea is the get data that shows…*") and negatively as reviewers ("*I hate it when authors don't list the…")*..

II. PRESENTER BIO

Dr. Keogh is a prolific author in data mining conferences, as of June 2009, he is one of only three people to have at least ten papers in each of the top three data mining conferences, ACM SIGKDD, IEEE ICDM and SIAM SDM (The other two are Philip Yu and Jiawei Han). While he is only 9 years out from his PhD he has already obtained an H-index of 40. He has given well-received tutorials at SIGKDD (Three times), ICDM (twice), VLDB, SDM, ACM Multimedia and CIKM.

Several of his papers have won "best paper" awards (including at this conference). In addition he has won several teaching awards, and he was the sole recipient of the University of California Riverside University Scholar for 2008. He is the recipient of a 5-year NSF Career Award for "Efficient Discovery of Previously Unknown Patterns and Relationships in Massive Time Series Databases" and two additional large NSF awards for data mining.

ACKNOWLEDGMENT

Dr. Keogh's tutorial are funded by NSF awards 0803410 and 0808770.

1550-4786/10 $26.00 © 2010 IEEE
DOI 10.1109/ICDM.2010.165