

THE ROUGH GUIDE TOWARDS BECOMING
A GOOD SCIENTIST

BART GOETHALS

TAKE MY ADVICE

I'M NOT USING IT ANYWAY

MY BACKGROUND

Frequent Itemset Mining



WHY START A PHD?

- Become famous
- Become rich
- For world peace
- Stay off the street

“I’m in it for the chicks ...”



WHY START A PHD?

- Because it's fun!
- It makes your mom proud!
- We get paid to study the things we're excited about!
- We can travel the world
- We can work anywhere and anytime

WHAT IS A GOOD TOPIC?

On private scalar product computation for privacy-preserving data mining

B Goethals, S Laur, H Lipmaa, T Mielikäinen

Information Security and Cryptology-ICISC 2004, 104-120

Mining all non-derivable frequent itemsets

T Calders, B Goethals

Principles of Data Mining and Knowledge Discovery, 74-86

Survey on frequent pattern mining

B Goethals

Univ. of Helsinki

Advances in frequent itemset mining implementations: report on FIMI'03

B Goethals, MJ Zaki

ACM SIGKDD Explorations Newsletter 6 (1), 109-117

Tiling databases

F Geerts, B Goethals, T Mielikäinen

Discovery science, 278-289

Frequent set mining

B Goethals

Data Mining and Knowledge Discovery Handbook, 377-397

FIMI'03: Workshop on frequent itemset mining implementations

B Goethals, MJ Zaki

Third IEEE International Conference on Data Mining Workshop on Frequent ...

Non-derivable itemset mining

T Calders, B Goethals

Data Mining and Knowledge Discovery 14 (1), 171-206

Data Mining and Knowledge Discovery 14 (1), 171-206

T Calders, B Goethals

Non-derivable itemset mining

Third IEEE International Conference on Data Mining Workshop on Frequent ...

B Goethals, MJ Zaki

CHOOSE A GOOD ADVISOR



THE KEY TO SUCCESS

Good Science!

GOOD SCIENCE?

- Try to solve real problems
- Chase your own ideas
- A good mix of low hanging fruit and not so low hanging fruit
- Question related work

GOOD SCIENCE?

- Prove all claims
- Good experiments (and implementation)

LOOKING FOR HIGH IMPACT?

- If impact = nr of citations
- the best and coolest scientific result does not necessarily get highest impact
- When do people cite?
 - if your result is being improved
 - if your paper is (among) the first in a specific domain

DON'T PANIC

- A paper being rejected three times can still receive a best paper award
- The reviewers obviously didn't understand it
- But who's fault is that?

A GOOD PAPER IS IN THE WRITING

- Make the reviewers' job easy
 - clearly enumerate the contributions (don't exaggerate)
 - give the strong points in your paper
 - share your enthusiasm

YOUR REVIEWS SHOULD BE BETTER!

- review as if it could be made public
- do NOT tell the reviewers they missed YOUR work

BE OPEN

- Open source code!
- Open data!

GIVE GOOD PRESENTATIONS

- the audience knows the background: WRONG
- focus on the general idea of the contributions
- slides should not contain much text
- avoid formulas
- rule of thumb: one slide per minute

NETWORKING

- The most interesting discussions happen in the corridors or in a bar after 8pm (the best lasting networking results!)



NETWORKING

- Read papers before the conference (if possible)
- Ask questions and give your blunt opinion
- Collaborate!

BE INVOLVED

- As part of the organisation (e.g. as volunteer)
- for every paper you submit, you should review 3!

FOR EVEN MORE ADVICE

*How to do good research,
get it published in
SIGKDD and get it cited!*

Eamonn Keogh

Computer Science & Engineering Department
University of California - Riverside
Riverside, CA 92521
eamonn@cs.ucr.edu

