

ACM Data Mining SIG
10/25/2010

Email Filtering:
A View from the Inside

Tom Fawcett

Association for Computing Machinery

Announcements

- About the Association of Computing Machinery
- Future Data Mining SIG Meetings & Events
- Future ACM Meetings
- Other Events
- Other Announcements (Companies Hiring)
- Introduction of Tonight's Talk

Association for Computing Machinery

www.SFBayACM.org

- SF Bay ACM Founded 1957
 - - Promote Knowledge Modern Computing
 - - Create Community
 - - Cooperate with Professional Groups
- Monthly Evening Seminars
 - - ACM General Meetings at HP on 3rd Wed
 - - Seeking new Location ... Volunteers?
 - - Data Mining SIG Meetings LinkedIn 4th Monday
- National ACM journals & conferences

Association for Computing Machinery

Data Mining SIG

Data Mining Camp: November 13, 2010

Day Long UnConference

Seeking Volunteers

A Smarter Process for Sensing the Information Space

W. Scott Spangler, IBM Almaden Research Center

November 22, 2010

Ensemble Methods

Giovanni Seni, PhD Elder Research

January 24, 2010

Association for Computing Machinery

The Man Who Lied to His Laptop: The Key Social
Rules that Every Engineer Needs to Know
November 17, 2010

Other Events

www.SDForum.com

Software Architecture & Modeling SIG

Scalable Analytical Processing with LinkedIn's Avatara

October 27, 2010 - LinkedIn

Java SIG

Java Persistence API (JPA)

November 2, 2010 - Cubberley Center

Semantic Web SIG

Make Machines Learn for Us

November 3, 2010 - Cubberley Center

Other Events Users Groups

R User Group

R in Agricultural Research

Hurricane Electric Internet Services

Nov 9 – 6:30PM, Fremont

<http://www.meetup.com/R-Users/calendar/14944208/>

SAS User Group

Nov 3-5: Western Users of SAS Software

San Diego

University of California Santa Cruz Extension

Fall Class Enrollment Still Open
Course Catalogs Available

Special Announcements

Data Mining SIG

- Companies Hiring
- Brief Announcements from those Hiring
- Other Announcements

Tom Fawcett

- Proofpoint, Inc email management Company
- International Conference on Machine Learning ICML
 - 2003 Chair
- Machine Learning Journal
 - Editorial Board
 - Past Action Editor
- Published Research
 - Journal Papers
 - Numerous Conferences
- University of Massachusetts, Ph.D.

Email Filtering: A View from the Inside

Past Approaches - Research

- Basic Text Classification**
- Classification Algorithms**
- Bayesian, SVM, Ensemble Methods**

Industry Approaches

- learning and adaptation**
- multiple distributed sources**
- quickly changing data streams**
- errors, privacy issues**
- clever devious adversaries**
- infinite representation space**
- real-time processing constraints**