

SIGKDD FY'06 Annual Report
July 2005-June 2006
Submitted by: Gregory Piatetsky, SIGKDD Chair

1. Annual Awards

The 2005 ACM SIGKDD Service Award was given to the Weka team for their development of the freely available Weka Data Mining Software, including the accompanying book Data Mining: Practical Machine Learning Tools and Techniques (now in second edition) and much other documentation.

The 2006 ACM SIGKDD Service Award was given to Won Kim, for his key role in founding and growing ACM SIGKDD.

The 2005 ACM SIGKDD Innovation Award was given to Leo Breiman for being one of the founding fathers of modern machine learning and data mining. He has been actively contributing to these fields, as well as to statistics, for more than 30 years. His best known contribution is his landmark work on decision trees (Classification and Regression Trees, 1984, known as CART(R)), written with Jerome Friedman, Richard Olshen, and Charles Stone.

The 2006 ACM SIGKDD Innovation Award was given to Ramakrishnan Srikant for his seminal work on mining association rules and privacy preserving data mining.

2. Significant Publications

The SIGKDD 2005 annual conference is maintaining its position as the leading conference on data mining and knowledge discovery. The KDD-2005 conference published 76 high quality papers on important research areas including Temporal Mining, Privacy, Cost Sensitive Learning, Streaming Data, Ensemble Learning, Graph Mining, Clustering and Grouping, Support Vector Machines, Text and Web Mining, Associations, and Novel Learning Algorithms.

An excellent Industry Track had papers on E-Commerce, Sequence Mining, Anomaly Detection, and Document Analysis.

There were also 9 KDD-05 workshops which produced proceeding in these important areas:

- * Data Mining Methods for Anomaly Detection
- * OSDM 2005: Open Source Data Mining
- * UBDM 2005: Utility-Based Data Mining
- * MRDM 2005: Multi-Relational Data Mining
- * BIODDD 2005: Data Mining in Bioinformatics
- * DM-SSP 2005: Data Mining Standards, Services, and Platforms
- * WebKDD 2005: Taming Evolving, Expanding and Multi-faceted Web Clickstreams
- * LinkKDD 2005: Link Discovery: Issues, Approaches and Applications
- * Multimedia Data Mining: "Mining Integrated Media and Complex Data"

SIGKDD Explorations published two excellent issues:

Dec 2005 7(2) Issue with the focus on Link Mining, and June 2006 8(1) Issue with the focus on Successful Real-World Data Mining Applications.

3. Significant programs that provided a springboard for further technical efforts

SIGKDD Curriculum committee, lead by Jiawei Han program has published its recommended curriculum for teaching data mining on SIGKDD website - see <http://www.acm.org/sigs/sigkdd/curriculum.php> .

The KDD chapters program did not see much formal action, but there is a promising local chapter of ACM in San Francisco Bay area which organized several data mining related talks.

4. Innovative programs which provide service to some part of your technical community

In addition to previously mention programs (see question 3), SIGKDD is considering a program to launch a series of webcasts on topics relating to data mining.

SIGKDD is also considering creation of SIGKDD Best Dissertation Award.

5 A very brief summary for the key issues that the membership of that SIG will have to deal with in the next 2-3 years.

SIGKDD had an election in 2005 which lead to replacement of most of the Executive Committee, but the new EC has adapted well.

Due to rising industry demand for data miners (as indicated. e.g. by the record number of job ads in KDnuggets job section in 2006 Q1 and Q2), it is getting harder to find willing volunteers.

A major challenge continues to be finding a good location for upcoming KDD conferences with the local willing and able general chair.

