Data Mining Initiative @ Minnesota
- A University-Industry Partnership -

Jaideep Srivastava, Professor
Computer Science & Engineering
srivasta@cs.umn.edu
Outline

- Minnesota data mining group
  - Composition
  - Visibility and leadership
- Areas of expertise
- Why partner with Minnesota
- Data Mining Initiative
  - Benefits
  - Interaction models
Minnesota Data Mining Group

- ~12 faculty
  - across 5 different departments
  - from 3 different colleges
- 4+ post doctoral researchers
- 30+ Ph.D. students
- 30+ MS and BS students
Minnesota’s Data Mining Leadership

- Among the very top data mining programs worldwide
- Visible & recognized leadership
  - Research – publications, grants, awards & recognition, invited keynotes & tutorials
  - Technology transfer – industry & government
  - Professional service – conference organization, journal editorial boards
- Education – graduates in faculty positions, industry, and government research labs
Visibility & Recognition

- Multiple papers at every major data mining conference consistently over the past many years
- 8+ National Science Foundation grants in the past three years
  - the most competitive forum for federally funded research
- Grants from various federal and state agencies, and industry
- Faculty delivered Keynote addresses and tutorials regularly at major conferences
- Faculty recognition
  - Distinguished University Professor
  - IEEE Fellows
  - ACM Fellows
Technology Transfer

- Minnesota Intrusion Detection System (MINDS) – being used by Army Research Lab and University of Minnesota
- Sales Opportunity Miner – IBM is building a full-fledged tool based on this research
- Global Climate Modeling for NASA
- Many others …
Data Mining Areas of Activity

**Science & Engineering Applications**
- Climate modeling (NASA)
- High energy physics (Fermi)
- Penetration mechanics (ARL)
- Simulation (LLNL, DOE)

**Bio-Medical Applications**
- Medical informatics (Mayo)
- Bioinformatics (NIH)
- Behavioral ecology (NSF)
- Drug discovery (NIH)

**Government Applications**
- Cyber security (ARDA, ARL)
- Transportation (MnDoT, FHWA)
- Physical security (United Tech)

**Business Applications**
- Sales & marketing (IBM)
- Automotive (Daimler/Chrysler)
- CRM (Intel)

**Core Research**
- Models, algorithms
tools
- NSF, ARL
Why partner with Minnesota?

- Recruitment
- Training
- Joint projects
- Expert consultation
- Joint 3rd party proposals
- Being “connected” with a top data mining program
**UMN Data Mining Initiative**

- **Concept**
  - Faculty from various UMN departments have gotten together to form a focal point for data mining activity
  - Single point for industry to collaborate with UMN DM research

- **Key activities**
  - Industry-University information exchange
  - Research & problem solving interactions
    - Model 1: Data mining laboratory
    - Model 2: Collaborative partnership in data mining
Industry-University information exchange

- Newsletters, reports, etc.
  - On mailing list for DTC newsletter
  - Access to technical reports published by the DM group
  - Mailing list for seminars – option available to subscribe through streaming media
- Invitation to annual DM open house
  - Annual ‘technical-get-to-know’ meeting with DM faculty
- Access to point person for leads regarding recruitment and expert consulting
- For projects
  - Access to University facilities (m/cs, space, etc. for joint projects)
  - Access to university’s contract management expertise
Model 1: Data Mining Laboratory (DML)

- **Concept**
  - A ‘loose collaboration’ of DM faculty which acts a single point for accessing their expertise by the industry

- **Interaction process**
  - Industrial partner approaches DML with a problem
  - DML identifies one or more faculty with relevant expertise and sets up meetings
  - Once a project is identified, DML helps with students, coordination, etc.
  - If it is a 3rd party project, e.g. federal government proposal, University’s contract management expertise is available
Model 2: Collaborative partnership in data mining

- **Concept**
  - A ‘meeting place’ for companies that have problems, companies that have tools, companies that have hardware, and university researchers to get together and advance the state of practice and research

- **Example**
  - Ameriprise, Best Buy, Carlson, Target, Thomson, Yahoo! (and others) are all interested in Customer Relationship Management, marketing, etc.
  - KXEN has high end DM tools; Unisys has high speed hardware
  - UMN has significant R&D expertise in the area,

- **Interaction process**
  - Collaboration model needs to be worked out in discussion with organizational representatives
Next Steps

- Become a Data Mining Consortium partner
  - Identify a ‘UMN representative’ in your organization
- Identify potential projects for interaction Model 1
  - You might already have such projects
  - A meeting between your ‘UMN rep’ and DM faculty can identify such problems
- Interaction Model 2
  - We will provide a list of areas in which such possibilities exist; will ask for your interest
  - If you have a specific interest, inform us
A good ‘UMN representative’ has

- Recognition in the organization as “the UMN person”
- Access to senior leadership
- Understanding of organization’s needs in terms of data mining/related technologies
- Interest, enthusiasm, commitment, etc. in being associated with the University
Thanks!

Questions, Discussion