Biomedical Data Mining: Healthcare Data

Sharath Cholleti
Health Informatics

• Clinical
  – improve outcome
  – Transplant success

• Quality control
  – Length of stay
  – Cost of stay
  – Mortality
  – 30 day readmission
Flow Cytometry
Data: 16 dimensions
Manual Gating

Defining CD8+ Cells

Lymph/Mono → Dump/CD3 → Singlet → CD4/CD8

CM → Naive → Total CD8+

CD45RA → PD-1
Transplant outcome
Overall Analysis

• Removed the figure as it is part of a proposal under review. Talk to me if you want to see this again.
Healthcare Projects

• Prediction
  – Readmission
  – Length of stay
  – Mortality rate
Healthcare Data

• Patient information
  – Age, sex, height, weight
  – Primary diagnosis
  – Secondary diagnosis: comorbidities
  – Lab tests and procedures codes
  – Medicines used
  – Socioeconomic information: salary, home neighborhood
Temporal Patterns

• Longitudinal study
• Flow cytometry
Visualization

- Ben Schneiderman
- Treemaps
- http://www.cs.umd.edu/hcil/
Interactive Information Visualization of a Million Items

One Million Items Treemap

40,000 items population of the USA scatter plot (Census 2000 data)

12,134 items Linux Kernel 2.5.33 source

Scroll to:
- Project description
- Participants
- Publications
- Related Sites
- Downloads