Progressive Visual Analytics User-Driven Visual Exploration of In-Progress Analytics

Chad Stolper Georgia Tech

Adam Perer

IBM T.J. Watson Research Center David Gotz UNC Chapel Hill

Select Dataset

























https://www.flickr.com/photos/veggiefrog/3435380297/

1. Increasingly large quantities of data



https://www.flickr.com/photos/veggiefrog/3435380297/

1. Increasingly large quantities of data

2. Increasingly complex analytics



1. Increasingly large quantities of data

2. Increasingly complex in analytics

https://www.flickr.com/photos/veggiefrog/3435380297/

Batch Visual Analytics Workflow



Batch Visual Analytics Workflow



Batch Visual Analytics Workflow











D. Fisher, I. Popov, S. Drucker, and m. c. schraefel, "Trust me, i'm partially right: incremental visualization lets analysts explore large datasets faster," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2012, pp. 1673–1682.





Progressive Visual Analytics

Progressive, User-Driven Analytics Progressive, Interactive Visualization

Progressive Visual Analytics

Progressive, User-Driven Analytics

Progressive, Interactive Visualization

Semantically-Meaningful Partial Results

Semantically-Meaningful Partial Results

(Thomas-- Results Feedback)

Semantically-Meaningful Partial Results

The partial results are of the same form as the final results

Semantically-Meaningful Partial Results

http://en.wikipedia.org/wiki/K-means_clustering

Semantically-Meaningful Partial Results



Semantically-Meaningful Partial Results



Semantically-Meaningful Partial Results



Semantically-Meaningful Partial Results



Semantically-Meaningful Partial Results

K-Means

Logistic Regression

Progressive Visual Analytics

Progressive, User-Driven Analytics

Progressive, Interactive Visualization
Incorporate Analyst Knowledge

Incorporate Analyst Knowledge

Incorporate Analyst Knowledge

Incorporate Analyst Knowledge

(Thomas- Result Control)

Progressive, User-Driven Analytics

Progressive, User-Driven Analytics

Progressive, User-Driven Analytics

Visualize the Partial Results

Visualize the Partial Results (Thomas– Results Feedback)

Analytic Visualization Analyst

Up-to-Date Information

Analytic Visualization Analyst





http://info.jmu.edu/dux/files/2013/01/chaos.jpg





Analytic Visualization Analyst

http://info.jmu.edu/dux/files/2013/01/chaos.jpg

http://www.historyforkids.org/learn/medieval/history/byzantine/justinian.jpg

Progressive, User-Driven Analytics







Progressive, User-Driven Analytics Progressive, Interactive Visualization

So what might such a system look like?





Procedures

Lab Tests

Diagnoses

Medications













Frequent Patterns





Frequent Patterns





Frequent Patterns

Event Series

Correlate with Outcomes

What Works?



Frequent Patterns

Event Series

Correlate with Outcomes

What Works?

(What Doesn't Work?)



Frequent Patterns

Event Series

Correlate with Outcomes

J.Ayres, J.Flannick, J.Gehrke, and T.Yiu. Sequential Pattern mining using a bitmap representation. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '02, pages 429–435, New York, NY, USA, 2002. ACM.

Sequential Pattern Mining Algorithm (Thomas– Dependent Subdivision)

J.Ayres, J.Flannick, J.Gehrke, and T.Yiu. Sequential Pattern mining using a bitmap representation. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '02, pages 429–435, New York, NY, USA, 2002. ACM.

- Build a tree of every possible pattern
 - (up to a maximum length)
- Prune the tree
 - (using a support threshold)



- Build a tree of every possible pattern
 - (up to a maximum length)

Prune the tree

• (using a support threshold)



- Build a tree of every possible pattern
 - (up to a maximum length)

Prune the tree

• (using a support threshold)



- Build a tree of every possible pattern
 - (up to a maximum length)

Prune the tree

• (using a support threshold)

Frequent Patterns

- **Event Series**
- Correlate with Outcomes

- Depth-First Search
- Outputs frequent patterns as each is found
Sequential Pattern Mining Algorithm

- Frequent Patterns
 - **Event Series**
 - Correlate with Outcomes

- Depth-First Search
 - Outputs frequent patterns as each is found

73

Sequential Pattern Mining Algorithm

- Frequent Patterns
 - **Event Series**
 - Correlate with Outcomes

- Depth-First Search
- Outputs frequent patterns as each is found

Breadth-First Search

Frequent Patterns

- **Event Series**
- Correlate with Outcomes

 Outputs frequent patterns as each is found

Breadth-First Search

Frequent Patterns

- **Event Series**
- Correlate with Outcomes

 Outputs frequent patterns as each is found



- **Frequent Patterns**
 - **Event Series**
 - Correlate with Outcomes

- Breadth-First Search
 - Shorter patterns first

 Outputs frequent patterns as each is found





http://thenextweb.com/wp-content/blogs.dir/1/files/2013/02/queue.jpg

- Breadth-First Search
 - Shorter patterns first

 Outputs frequent patterns as each is found



http://thenextweb.com/wp-content/blogs.dir/1/files/2013/02/queue.jpg

- Breadth-First Search
 - Shorter patterns first
 - Analyst prioritization
- Outputs frequent patterns as each is found

- Breadth-First Search
 - Shorter patterns first
 - Analyst prioritization
- Outputs frequent patterns as each is found





- Breadth-First Search
 - Shorter patterns first
 - Analyst prioritization
- Outputs frequent patterns as each is found
- Manual Pruning

- Breadth-First Search
 - Shorter patterns first
 - Analyst prioritization
- Outputs frequent patterns as each is found
- Manual Pruning





- Breadth-First Search
 - Shorter patterns first
 - Analyst prioritization
- Outputs frequent patterns as each is found
- Manual Pruning

Progressive Visual Analytics

Progressive, User-Driven Analytics

Progressive, Interactive Visualization

Progressive Visual Analytics

Progressive, User-Driven Analytics

Progressive, Interactive Visualization













Support \$	Correlation +
heart failure diuretics3	 angiotensin-converting enzyme (ace) inhibitors diuretics3>heart failure>diuretics3
chronic airways obstruction, not elsewhere classified heart failure>diuretics3	 heart failure>angiotensin-converting enzyme (ace) inhibitors diuretics3>heart failure>glucocorticoids>diuretics3
heart failure>chronic airways obstruction, not elsewhere clas diuretics3>heart failure	dyslipidemics, hmg coa reductase inhibitors
beta blockers2 glucocorticoids	 diuretics3>heart failure>diuretics3>chronic airways obstr diuretics3
chronic airways obstruction, not elsewhere classified>heart t heart failure>beta blockers2	 dyslipidemics, hmg coa reductase inhibitors>angiotensin-co diuretics3>heart failure>chronic airways obstruction, not el

Correlation \$
angiotensin-converting enzyme (ace) inhibitors
 duretics3>neart failure>duretics3 heart failure>angiotensin-converting enzyme (ace) inhibitors
 diuretics3>heart failure>glucocorticoids>diuretics3 dyslipidemics, hmg coa reductase inhibitors
beta blockers2
diuretics3>heart failure>diuretics3>chronic airways obstr diuretics3
 dyslipidemics, hmg coa reductase inhibitors>angiotensin-co diuretics3>heart failure>chronic airways obstruction, not el

Support +	Correlation ÷
 heart failure diuretics3 chronic airways obstruction, not elsewhere classified heart failure>diuretics3 heart failure>chronic airways obstruction, not elsewhere clas diuretics3>heart failure beta blockers2 glucocorticoids chronic airways obstruction, not elsewhere classified>heart 1 	 angiotensin-converting enzyme (ace) inhibitors diuretics3>heart failure>diuretics3 heart failure>angiotensin-converting enzyme (ace) inhibitors diuretics3>heart failure>glucocorticoids>diuretics3 dyslipidemics, hmg coa reductase inhibitors beta blockers2 diuretics3>heart failure>diuretics3>chronic airways obstr diuretics3 dyslipidemics, hmg coa reductase inhibitors>angiotensin-co
heart failure>beta blockers2	diuretics3>heart failure>chronic airways obstruction, not el

Support \$	Correlation \$
hear, failura	angiotensin-converting enzyme (ace) inhibitors
and stics3	diuretics3>heart failure>diuretics3
chronic airways obstruction, not elsewhere classified	heart failure>angiotensin-converting enzyme (ace) inhibitors
heart failure>diuretics3	diuretics3>heart failure>glucocorticoids>diuretics3
heart failure>chronic airways obstruction, not elsewhere clas	dyslipidemics, hmg coa reductase inhibitors
diuretics3>heart failure	beta blockers2
beta blockers2	diuretics3>heart failure>diuretics3>chronic airways obstr
glucocorticoids	diuretics3
chronic airways obstruction, not elsewhere classified>heart t	dyslipidemics, hmg coa reductase inhibitors>angiotensin-co
heart failure>beta blockers2	diuretics3>heart failure>chronic airways obstruction, not el

Support \$	Correlation +
heart failure	angiotensin-converting enzyme (ace) inhibitors
chronic airways obstruction, not elsewhere classified	 diuretics3>heart failure>diuretics3 heart failure>angiotensin-converting enzyme (ace) inhibitors diuretics3>heart failure>diuretics3
heart failure>chronic airways obstruction, not elsewhere clas	dyslipidemics, hmg coa reductase inhibitors
beta blockers2	diuretics3>heart failure>diuretics3>chronic airways obstr
Glucocorricolds Chronic airways obstruction, not elsewhere classified>heart 1 heart failure>beta blockers2	 duretics3 dyslipidemics, hmg coa reductase inhibitors>angiotensin-co diuretics3>heart failure>chronic airways obstruction, not el







Support

-
Support



-

Support



-











Selected Patte
Support:
Correlation:
∆Correlation:











Selected Patte Support: Correlation: ACorrelation:























27

ЭM

So Does it Work?

So Does it Work?

Medical Researchers, University Hospital of North Norway

So Does it Work?

AMIA 2014

Skrovseth, Perer, Delaney, Revaug, Lindsetmo, Augestad

"Detecting Novel Associations for Surgical Hospital Readmissions in Large Datasets by Interactive Visual Analytics"



- A precise definition and design guidelines for Progressive Visual Analytics
- An example Progressive Visual Analytics system, Progressive Insights
- (In the paper) A case study of using Progressive Insights for exploring frequent patterns in EHR

- A precise definition and design guidelines for Progressive Visual Analytics
- An example Progressive Visual Analytics system, Progressive Insights
- (In the paper) A case study of using Progressive Insights for exploring frequent patterns in EHR

- A precise definition and design guidelines for Progressive Visual Analytics
- An example Progressive Visual Analytics system, Progressive Insights
- (In the paper) A case study of using Progressive Insights for exploring frequent patterns in EHR

- A precise definition and design guidelines for Progressive Visual Analytics
- An example Progressive Visual Analytics system, Progressive Insights
- (In the paper) A case study of using Progressive Insights for exploring frequent patterns in EHR

Chad Stolper

Georgia Tech chadstolper@gatech.edu





Adam Perer IBM T.J. Watson Research Center



Thank You!

Chad Stolper Georgia Tech

chadstolper@gatech.edu





Adam Perer IBM T.J. Watson Research Center



Questions?

Chad Stolper Georgia Tech

chadstolper@gatech.edu





Adam Perer IBM T.J. Watson Research Center



Questions?

Chad Stolper Georgia Tech

chadstolper@gatech.edu





Adam Perer IBM T.J. Watson Research Center



Progressive Visual Analytics

Seven Design Guidelines

Progressive Visual Analytics Systems Analytics should...

- 1. Provide increasingly meaningful partial results as the algorithm executes
- 2. Allow users to focus the algorithm to subspaces of interest
- 3. Allow users to ignore irrelevant subspaces

Progressive Visual Analytics Systems Visualizations should...

- 4. Minimize distractions by not changing views excessively
- 5. Provide cues to indicate where new results have been found
- 6. Support on-demand refresh
- 7. Provide an interface to specify where the analytics should *focus* and ignore