

BrainEx

Ver 1.0.3

Preprocess Dataset

Working dataset:

Select Sequence

Selected Sequence:

Find Similar Sequences

Find Similar
Sequences Result:

< Select New Dataset

Clean "menu screen" so
you can start at different
points in the flow

BrainEx Menu

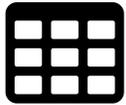
Query Finder



Dataset Explorer



Cluster Explorer



Load Another
Dataset

BrainEx

Ver 1.0.3

Preprocess Dataset

Preprocessing Progress...51%



Working dataset: dataset4.csv

Select Sequence

Selected Sequence:

Find Similar Sequences

Find Similar
Sequences Result:

< Select New Dataset

Clean "menu screen" so
you can start at different
points in the flow

BrainEx

Ver 1.0.3

Preprocess Dataset

Preprocessing Progress...100%



Working dataset: dataset4.csv

Select Sequence

Selected Sequence: dataset4.csv

Find Similar Sequences

Find Similar
Sequences Result:

< Select New Dataset

Clean "menu screen" so
you can start at different
points in the flow

BrainEx

Ver 1.0.3

Preprocess Dataset

Preprocessing Progress...100%



Working dataset: dataset4.csv

Select Sequence

Selected Sequence: dataset4.csv

Find Similar Sequences

Find Similar
Sequences Result: queryresults.csv

< Select New Dataset

Clean "menu screen" so
you can start at different
points in the flow

Choose an existing dataset to begin analysis:

SART1 x

SART2 x

SART3 x

SART4 x

+ Add file

BrainEx

Ver 1.0.3

 Use existing dataset

 Select new dataset from my machine

BrainEx | Select Dataset



Unpreprocessed
CSV files

Dataset1

Dataset2

Dataset3

< Back

Upload new file

Next



For simplicity, click on Dataset2 if you want
to click on an existing dataset on the side
bar

Navigation icons: back, forward, computer icon, address bar: > Computer, refresh icon, search icon

Organize System Properties Uninstall or change a program Map network drive Open Control Panel

- ★ Favorites
 - Recently Changed
 - Desktop
- Libraries
 - Documents
 - Music
 - Pictures
 - Movies

Name	Date Modified	File Type
Loreem ipsum dolor	1/13/2012 9:00	File Folder
Loreem ipsum dolor	1/13/2012 9:00	Text File
Loreem ipsum dolor	1/13/2012 9:00	Text File
Loreem ipsum dolor	1/13/2012 9:00	Text File
Loreem ipsum dolor	1/13/2012 9:00	Text File
Loreem ipsum dolor	1/13/2012 9:00	Text File
Loreem ipsum dolor	1/13/2012 9:00	Text File

Upload a CSV file:

Dataset4.csv

Upload

Cancel



BrainEx I Preprocessing Options

② Similarity Threshold: 0.1 0.2 Custom: ____

② Type of Distance:

Warped Euclidean ▼

Warped Euclidean
Warped City Block (manhattan)
Warped Minkowski
Warped Chebyshev

② Sequence Length of Interest: between and

< Back to Menu

Start Preprocessing

Question: what are the true names of the distances?

Length of interest means the number of seconds of how long the sequence lasted that the user wants to see

BrainEx I Loading Clusters



Preprocessing Progress...51%



Data preprocessed: 102/200

Cancel

Next

Notes:

This page is intended to automatically load and display the finished loading clusters page. However, you need to click next because of Balsamiq's (prototype website that we used for creating this mockup) limitations on transitions.

Cluster loading is complete!

Clusters Processed: 200/200

Menu

Explore Loaded Data

Find Similar Sequences

Pick an option to explore the loaded data:

Dataset Explorer ⓘ

Cluster Explorer ⓘ

< Back to Menu



Pick an option to explore the
loaded data:

Dataset Explorer ?

Cluster Explorer ?

< Back to Menu



[← Back to Menu](#)

Current Selection



Filter

Channel

- Channel1
 Channel2
 Channel3
 Channel4
 Channel5

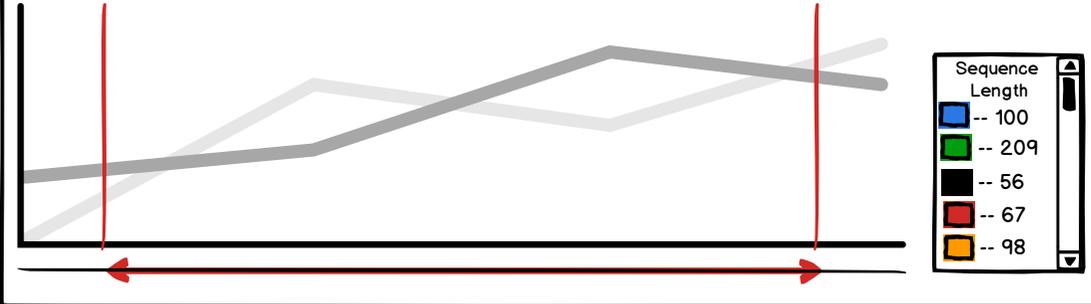
SubjectID

Labels

Statistics

Total Number of Data
Records: 2000
 Average length of time: 121
seconds
 Minimum: 222 seconds
 Maximum: 800 seconds

Data Visualizer



Data Viewer

SubjectID	Event Name	Channel Name	Start Time	End Time
101 HART	target correct	Channel-1 HBO	300	332
93 SART	target incorrect	Channel-4 VBO	789	800
34 TART	target no response	Channel-18 JVO	210	222

Save selected sequence

View Selected Sequence

Dataset Explorer

Cluster Explorer

Notes:

Sort the data here in descending order of (end time - start time)

Do not show all data at once, let the user filter first and then show the data

Design Justification:

We wanted to make the screen straightforward, simple to look at and easy to navigate. Keeping these goals in mind, we put all the user parameter selections on one side (the left) and visualizations on the other (the right).

After performing necessary functions, the user can save the file and then proceed to the next step which is querying the sequence.

At the top, there is a navigation bar that allows the user to go to the main page as well as each of the explorer/query pages.

We chose a slider instead of brushing on chart region so it does not interfere with selecting an actual sequence from the chart.

[← Back to Menu](#)

Current Selection



Filter

Channel

- Channel1
- Channel2
- Channel3
- Channel4
- Channel5

SubjectID

Labels

Statistics

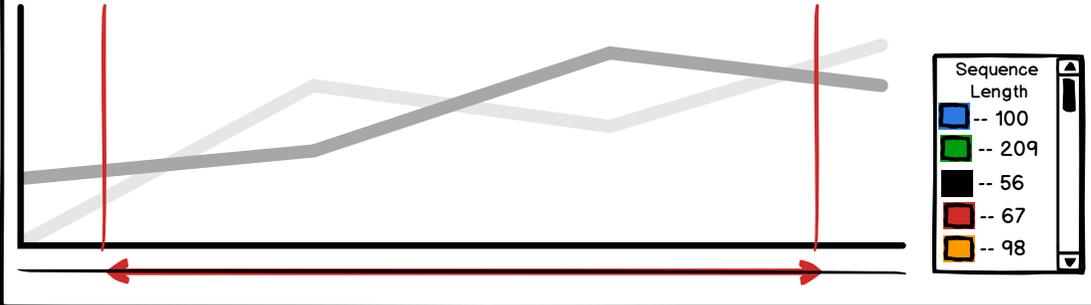
Total Number of Data
Records: 2000

Average length of time: 121
seconds

Minimum: 222 seconds

Maximum: 800 seconds

Data Visualizer



Data Viewer

SubjectID	Event Name	Channel Name	Start Time	End Time
101 HART	target correct	Channel-1 HBO	300	332
93 SART	target incorrect	Channel-4 VBO	789	800
34 TART	target no response	Channel-18 JVO	210	222

Save selected sequence

View Selected Sequence

Dataset Explorer

Cluster Explorer

Notes:

Sort the data here in descending order of (end time - start time)

Do not show all data at once, let the user filter first and then show the data

Design Justification:

We wanted to make the screen straightforward, simple to look at and easy to navigate. Keeping these goals in mind, we put all the user parameter selections on one side (the left) and visualizations on the other (the right).

After performing necessary functions, the user can save the file and then proceed to the next step which is querying the sequence.

At the top, there is a navigation bar that allows the user to go to the main page as well as each of the explorer/query pages.

We chose a slider instead of brushing on chart region so it does not interfere with selecting an actual sequence from the chart.

BrainEx | Cluster Explorer

< Back to Menu

Clusters

Number: 5



Number: 60



Number: 45



Number: 16



Filter

Number of Clusters to Display:

1 20

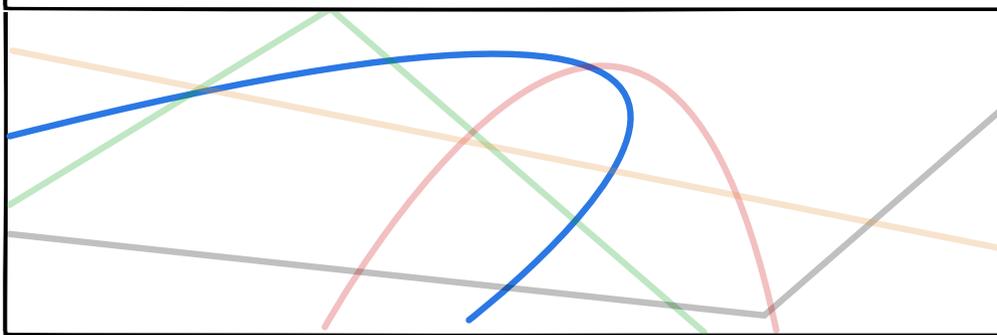
Range of sequence length in each cluster:

0 100

Statistics

Number of clusters: 3456
Sequences
Mean: 354
Minimum: 5
Maximum: 500

Data Visualizer



Data Viewer

Sequence Length (in seconds)	Number of Sequences
100	5
209	60
56	45
67	16
98	78
509	109
290	32
876	12
487	9
1067	10

Save selected sequence

View Selected Sequence

Dataset Explorer

Cluster Explorer

Notes:

Data Visualizer will display the user specified amount of cluster representatives and number of sequences
Data Table will have length (time) of sequences and number of sequences
There will never be a cluster with the same length so can just use length as unique cluster ID -> this is false
There can be multiple clusters with the same length as long as they are not similar in shape
All sequences in a cluster will have the same length
Selecting cluster and representative will highlight it on the data visualizer graph

Filters:

Number of clusters (top 15 or 20)
Range of length (slider)

Statistics:

Show the cluster with the most/least sequences
Show the top 3/least 3 with the greatest/least length of sequences

Question:

How do we decide how many representatives to display to the user (maybe ask to input a number from 1-10) and on what basis (show representatives of the top 10 clusters with most data points)?
If we have 30,000 clusters, how do we filter them down to approx 20? Should we ask for user input or another screen?
How to identify each cluster? Should we assign an id to each?

Design Justification:

To keep the UI simple, this cluster explorer screen as well as the rest of the explorer/query finder screens have the same layout with user inputs on the left and visualizations on the right
The view data shapes area displays the data shape of the representative of each cluster which would make it simpler for the user to view and easier to pick which area to investigate more
The number of sequences in each cluster is displayed as the label for the view data shapes area to make it easier for users to view that information. The filters are meant to allow the user to narrow down the number of clusters to view at a time

< Back to Menu



Clusters

- Number: 5
- Number: 60
- Number: 45
- Number: 16

Filter

Number of Sequences from this Cluster to Display:

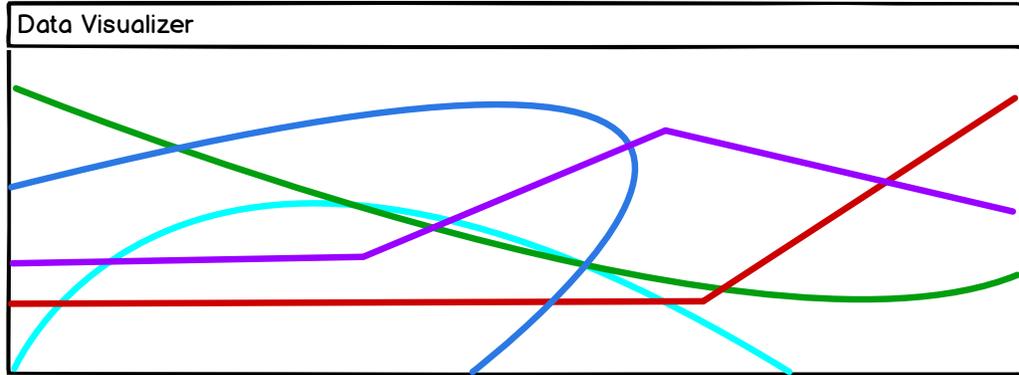
1 ————— 20

Range of sequence length:

0 ————— 500

Statistics

Number of clusters: 3456
 Sequences
 Mean: 354
 Minimum: 5
 Maximum: 500



Legend

- █ -- 100
- █ -- 100
- █ -- 100
- █ -- 100
- █ -- 100

Data Viewer

SubjectID	Event Name	Channel Name	Start Time	End Time	Sequence Length (in seconds)
101 HART	target correct	Channel-1 HBO	300	400	100
93 SART	target incorrect	Channel-4 VBO	789	889	100
34 TART	target no response	Channel-18 JVO	210	310	100
101 HART	target correct	Channel-1 HBO	350	450	100
93 SART	target incorrect	Channel-4 VBO	675	775	100

Back

Save selected sequence

View Selected Sequence

Notes:
 Data Visualizer will display the sequences in the cluster selected
 Data Table will have data information for each sequence in the cluster selected

Design Justification:
 In order to make the cluster explorer more understandable and easy to use, we have a graph that displays the current user selection at the top of the screen.
 The user can filter some options and select the number of sequences to view in the cluster as well as the range of the sequence length. This part is to allow the user to narrow down the number of sequences to view at a time

Question: If each sequence in one cluster has the same length, how do we identify?

Is sequence length end-start time?

Doesn't the graph need a slider?

[← Back to Menu](#)

Current Selection



1. Sequence Selection

Upload a sequence file

2. Enter Parameters

Number of best sequence matches:

Exclude data from selected sequence

Overlap allowed:

Length of Interest:

Show Query Results

Statistics

Notes:

Show query result is disabled on this page because parameters have not been entered. Gray out selection if selection not made in cluster explorer/dataset explorer.

Design Justification:

The input required for query is setup as a menu with a scrollbar because we are trying to make sure users put in the information in order, but user testing showed that accordion menus are ineffective.

< Back to Menu

Current Selection



1. Sequence Selection

- Upload a sequence file
- Use selection from dataset explorer

2. Enter Parameters

- Number of best sequence matches:
- Exclude data from selected sequence
- Overlap allowed: 10%

Show Query Results

Statistics

Notes:

Show query result is disabled on this page because parameters have not been entered. Gray out selection if selection not made in cluster explorer/dataset explorer.

Design Justification:

The input required for query is setup as a menu with a scrollbar because we are trying to make sure users put in the information in order, but user testing showed that accordion menus are ineffective. The length of interest field is hidden on this page for space reasons.

[← Back to Menu](#)

Current Selection

1. Sequence Selection

- Upload a sequence file
- Use selection from dataset explorer
- Use selection from cluster explorer

2. Enter Parameters

Number of best sequence matches:

5

Show Query Results

Statistics

Notes:

Show query result is disabled on this page because parameters have not been entered. Gray out selection if selection not made in cluster explorer/dataset explorer.

Design Justification:

The input required for query is setup as a menu with a scrollbar because we are trying to make sure users put in the information in order, but user testing showed that accordion menus are ineffective. The exclude queried sequence, overlap, and length of interest fields are hidden on this page for space reasons.

Navigation icons: back, forward, computer icon, address bar: > Computer, refresh icon, search icon

Organize System Properties Uninstall or change a program Map network drive Open Control Panel

- ★ Favorites
 - Recently Changed
 - Desktop
- Libraries
 - Documents
 - Music
 - Pictures
 - Movies

Name	Date Modified	File Type
Loreem ipsum dolor	1/13/2012 9:00	File Folder
Loreem ipsum dolor	1/13/2012 9:00	Text File
Loreem ipsum dolor	1/13/2012 9:00	Text File
Loreem ipsum dolor	1/13/2012 9:00	Text File
Loreem ipsum dolor	1/13/2012 9:00	Text File
Loreem ipsum dolor	1/13/2012 9:00	Text File
Loreem ipsum dolor	1/13/2012 9:00	Text File

Upload a CSV file:

queryseq1.csv

Upload

Cancel



< Back to Menu

Current Selection



1. Sequence Selection

- Upload a sequence file
- Use selection from dataset explorer
- Use selection from cluster explorer

2. Enter Parameters

? Number of best sequence matches:

Show Query Results

Statistics

Notes:

Show query result is disabled on this page because parameters have not been entered. Gray out selection if selection not made in cluster explorer/dataset explorer.

Design Justification:

The input required for query is setup as a menu with a scrollbar because we are trying to make sure users put in the information in order, but user testing showed that accordion menus are ineffective. The exclude queried sequence, overlap, and length of interest fields are hidden on this page for space reasons.

< Back to Menu

Current Selection



1. Sequence Selection

- Upload a sequence file
- Use selection from dataset explorer
- Use selection from cluster explorer

2. Enter Parameters

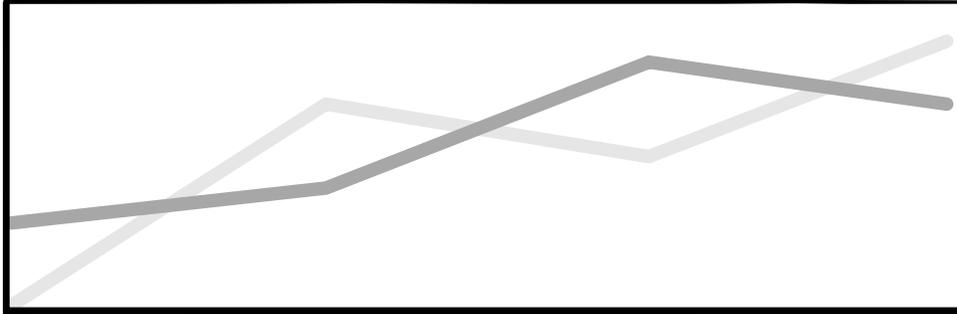
Number of best sequence matches:

Show Query Results

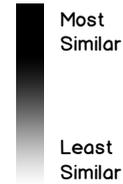
Statistics

Similarity
 Mean: 67%
 Standard Deviation: 3.2
 Variance: 10.24

Data Visualizer



Similarity



Data Viewer

SubjectID	Best	Overla	Similarit	Event Name	Channel	Start	End
101 HART	5	20%	65%	target correct	Channel-1 HBO	300	9000
93 SART	6	40%	75%	target incorrect	Channel-4 VBO	789	45756
34 TART	4	50%	85%	target no	Channel-18 JVO	210	45900
101 HART	2	36%	95%	target correct	Channel-1 HBO	300	9000
93 SART	3	48%	55%	target incorrect	Channel-4 VBO	789	45756

Save sequence results

Notes:

Data Table: Best matches, overlap, similarity (95% similar), subject ID, channel, start, end
 Visualization: data points in each explorer are displayed in graph (same as graph in cluster explorer copy)

Design Justification:

The input required for query is setup as a menu with a scrollbar because we are trying to make sure users put in the information in order, but user testing showed that accordion menus are ineffective. The exclude queried sequence, overlap, and length of interest fields are hidden on this page for space reasons.

Question:

Takes time to extract things from results and require more functionality on backend. What if i want to see top 5, middle 3, bottom 3? (talk this over with Leo about feasibility of our current idea, but Rodica brought up a good point on being able to select specific sequences that are not adjacent in the table/ranking)