Course Friction
Explorer

Marie Salomon, Noa Heyl, Shizuko Akamoto, ToTo Tokaeo

December 15, 2021
Overview

1. **Why?** Background, Motivation, Problem

2. **What?** Data, Usage Scenario

3. **How?** Idea, Implementation, Results

4. Demo
Why?
CPSC 310

- Introductory Software Engineering course
- More than 300 students registered
- Involves a term-long project
- Different resources for help available
Understanding the problems

01. No easy access to information to determine the cause of students struggling

02. Low scores on assessments as the only reliable signal so far

03. Limited possibility for interventions
Motivation

- Target User: CPSC 310 staff
- Explore students’ experiences
- Identification of possible causes of struggle
- Identification of patterns of struggle or friction
- Creating intervention possibilities
What?
Input Data

CPSC 310 data from 2020W2

- 409 people (students + TAs)
- 2366 Office Hours visits
- 5801 Piazza contributions
- 25488 Autotest results
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Attribute name</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contributions</td>
<td>Piazza contributions</td>
<td>kind</td>
<td>Contribution kind</td>
<td>“follow-up”</td>
</tr>
<tr>
<td>autotest_results</td>
<td>Automated test results</td>
<td>is_project</td>
<td>Whether the contribution was project related</td>
<td></td>
</tr>
<tr>
<td>queue_visits</td>
<td>TA-held office hours</td>
<td>anon_id</td>
<td>Deidentified user hash of the committer</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>committer</td>
<td>Deidentified user ID</td>
<td></td>
</tr>
</tbody>
</table>
## Ordered Attributes

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Attribute name</th>
<th>Description</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contributions</td>
<td>Piazza contributions</td>
<td>created_at</td>
<td>Timestamp of contribution</td>
<td></td>
</tr>
<tr>
<td>autotest_results</td>
<td></td>
<td>score</td>
<td>Test score of Autotest run</td>
<td>0.00 to 100.00</td>
</tr>
<tr>
<td>queue_visits</td>
<td>TA-held office hours</td>
<td>answer_start</td>
<td>Timestamp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>answer_finish</td>
<td>Timestamp</td>
<td></td>
</tr>
</tbody>
</table>
Usage Scenario

- Identifying outliers
- Identifying causes of struggle
- Comparing causes of struggle
- Identifying struggling students
- Identifying patterns
How?
Idea - Solution

- Create a dashboard that
  - allows for groupings of students by different **indicators**
  - comparisons between these groupings and the outcome group
  - relative to time
- Reveals how well an indicator predicts some outcome

Definitions:
- **Attributes**: property of the students (derived)

- **Indicator**: combination of attributes
  - using operators (e.g. logic, arithmetic, comparison, etc.)
  - **Outcome**: another indicator (to predict)
Attribute - Example

Attributes are derived data from the underlying dataset attributes. They abstract away time.

- num_commits
- num_piazza_questions
- avg_delta
- ...and 15 more

Student #1234  student.num_commits

0 1 5 7 7
Indicator - Example

Indicator

- combination of attributes
- with operators (e.g. logic, arithmetic, comparison, etc.)

Examples

\[(\text{min(weekly(student.num_commits)) > 0})\]
- Students who make at least 1 commit each week

\[(\text{(student.num_office_hours - student.num_piazza_questions) > 0})\]
- Students who go to more OH more often than asking Piazza questions
Implementation

- **Frontend**: JS, TS
  - D3 data visualization
  - React: UI
  - Redux: states

- **Backend**: Python
  - DSL: Antlr
  - Server: Uvicorn, FastAPI
  - SQLite: database
Results - Visualizations

01 Histograms
02 Circular packing
03 Stacked bar chart
04 Histogram widgets

Overview

Indicators Board
Results - Histograms

01

Histograms with attributes and total amount of students

Idioms:
- small multiples
- change over time
- selection (hover)

Channels:
- length, color, vertical position
Results - Circular Packing

02

Outcome and indicator circles linked by similarity (F-Score)

Idioms:
- zooming, panning
- selection (click, hover)
- change with time

Channels:
- circle size, color, link length, 2D position
Bar chart showing the ratio of true/false positives and negatives overall

- linked views
Results - Histogram Widgets

04

Juxtapose selected indicator students to the outcome

Idioms:
- superimposed views

Channels:
- color, length, vertical position
Demo
Demo link

Short demo version (2:30 min):
https://www.youtube.com/watch?v=-yWqeSw0s-k

More detailed version (4:30 min):
https://www.youtube.com/watch?v=ly-if4Qrb2k
Course Friction Explorer: A tool for CPSC 310

Allows staff to.....

- identify causes for struggle
- identify patterns
- identify struggling students
- create interventions
THANKS!

Any questions?
Future Work

- Scalability of circular packing for larger indicators
  - Introduce additional hierarchy level, group the students
- Sparsity of viz, uses distance between circles to encode similarity
  - Explore other channels, like link width for encoding similarity (F-Score)
- Difficult to identify recurring students in different indicators
  - Visualize overlap between indicators when an indicator is selected
- No UI for defining indicators (rely on DSL)
  - Buttons, selections, sliders to express query without needing the DSL explicitly