Extending the Goals of Peer-Assessment
Predicting Student Progress using Peer-Graded Responses

Michael Mogessie    Marco Ronchetti    Giuseppe Riccardi

University of Trento
Department of Information Engineering and Computer Science

02 July 2016
Outline

1. Introduction
2. A Semi-Automated Peer-Assessment Platform
3. Continuous Prediction for Monitoring Student Progress
4. Discussion and Conclusion
Student Performance Prediction

The process of predicting student performance

- At any point during the course of learning
- At any level of education

Earlier performance prediction studies used:

- standard test results + high school grades $\rightarrow$ success in college
- statistical measures of correlation

Later studies used:

- more data: demographic, assignment results, project grades
- Linear Regression, Neural Nets
- Data from online learning environments
Student Performance Prediction

The process of predicting student performance

• At any point during the course of learning
• At any level of education

Earlier performance prediction studies used:

• standard test results + high school grades → success in college
• statistical measures of correlation

Later studies used:

• more data: demographic, assignment results, project grades
• Linear Regression, Neural Nets
• Data from online learning environments
Student Performance Prediction

The process of predicting student performance

- At any point during the course of learning
- At any level of education

Earlier performance prediction studies used:

- standard test results + high school grades → success in college
- statistical measures of correlation

Later studies used:

- more data: demographic, assignment results, project grades
- Linear Regression, Neural Nets
- Data from online learning environments
Course levels and subject areas

- Most at the undergraduate level
- The majority: Computer science and engineering courses

Predicting overall success vs. specific outcome

- Pass or Fail - Classification
- More granular predictions - Grades, exact scores

One-off vs. Continuous predictions

- The majority: one-off studies
- Mature online learning platforms → continuous predictions
Student Performance Prediction (cont’d)

Course levels and subject areas
- Most at the undergraduate level
- The majority: Computer science and engineering courses

Predicting overall success vs. specific outcome
- Pass or Fail - Classification
- More granular predictions - Grades, exact scores

One-off vs. Continuous predictions
- The majority: one-off studies
- Mature online learning platforms → continuous predictions
Student Performance Prediction (cont’d)

Course levels and subject areas
- Most at the undergraduate level
- The majority: Computer science and engineering courses

Predicting overall success vs. specific outcome
- Pass or Fail - Classification
- More granular predictions - Grades, exact scores

One-off vs. Continuous predictions
- The majority: one-off studies
- Mature online learning platforms → continuous predictions
Student Performance Prediction - Stats about 53 studies

- Publication year distribution
- Accuracy vs. F1 score vs. RMSE vs. MSE vs. MAE vs. R
- Categories: pass/fail, grade/gpa/performance level, final score
- Data sources: demographic, high school, previous semesters, current semester, online platform
Peer-Assessment

“... an arrangement in which individuals consider the amount, level, value, worth, quality, or success of the products or outcomes of learning of peers of similar status.”

— Topping (1998)
Peer-Assessment - Research Areas of Interest

500+ studies in over half a century:
Peer-Assessment - Research Areas of Interest

500+ studies in over half a century:

- Reliability and validity of PA
- Student involvement
- Variables of PA
- Quality and design
- Peer-feedback
Peer-Assessment - Unexplored Potentials

- Its informative power about students
- Automated PA may facilitate labelling of data
- Such data may help with student performance prediction
The Problems

• Current prediction relies on data that needs to be labelled by experts
Problem Statement and Research Questions

The Problems

- Current prediction relies on data that needs to be labelled by experts
- Making timely and continuous predictions is challenging because collecting enough data takes a long time
Problem Statement and Research Questions

The Problems

- Current prediction relies on data that needs to be labelled by experts
- Making timely and continuous predictions is challenging because collecting enough data takes a long time

Research Questions:

- Can we predict student performance using peer-assessment data?
- Can we model students and test items using peer-rated responses?
- Can we make continuous predictions using peer-assessment data to track student progress?
Problem Statement and Research Questions

The Problems

- Current prediction relies on data that needs to be labelled by experts
- Making timely and continuous predictions is challenging because collecting enough data takes a long time

Research Questions:

- Can we predict student performance using peer-assessment data?
- Can we model students and test items using peer-rated responses?
- Can we make continuous predictions using peer-assessment data to track student progress?
Problem Statement and Research Questions

The Problems

- Current prediction relies on data that needs to be labelled by experts
- Making timely and continuous predictions is challenging because collecting enough data takes a long time

Research Questions:

- Can we predict student performance using peer-assessment data?
- Can we model students and test items using peer-rated responses?
- Can we make continuous predictions using peer-assessment data to track student progress?
Problem Statement and Research Questions

The Problems

- Current prediction relies on data that needs to be labelled by experts
- Making timely and continuous predictions is challenging because collecting enough data takes a long time

Research Questions:

- Can we predict student performance using peer-assessment data?
- Can we model students and test items using peer-rated responses?
- Can we make continuous predictions using peer-assessment data to track student progress?
Outline

1. Introduction
2. A Semi-Automated Peer-Assessment Platform
3. Continuous Prediction for Monitoring Student Progress
4. Discussion and Conclusion
Main Goals

- Develop a prototype Peer-Assessment platform
Main Goals

- Develop a prototype Peer-Assessment platform
- To be used in some CS courses at Uni Trento
Main Goals

- Develop a prototype Peer-Assessment platform
  - To be used in some CS courses at Uni Trento
  - To collect semester-wide student performance data
Design

- Weekly activities - asking, answering and evaluating answers
  - students ask questions about selected topics
  - a selected number of questions are randomly assigned to students
  - Students submit answers
  - Q&A sets randomly distributed to students
  - students vote for the best answer

- participation not mandatory but has bonus points
- anonymous and random PA activities
- Q&A sets made available to students every week
Design

- Weekly activities - asking, answering and evaluating answers
  - students ask questions about selected topics
  - a selected number of questions are randomly assigned to students
  - Students submit answers
  - Q&A sets randomly distributed to students
  - students vote for the best answer

- participation not mandatory but has bonus points
- anonymous and random PA activities
- Q&A sets made available to students every week
Outline

1. Introduction
2. A Semi-Automated Peer-Assessment Platform
3. Continuous Prediction for Monitoring Student Progress
4. Discussion and Conclusion
Some Interpretations of Progress

We identify three ways of measuring progress

- At a specific point in the course, how does the student’s performance fare against those of others? *(Type A)*
- How far is the student from achieving objectives of the entire course? *(Type B)*
- How far is the student from achieving objectives of the course modules? *(Type C)*
Some Interpretations of Progress

We identify three ways of measuring progress

- At a specific point in the course, how does the student’s performance fare against those of others? *(Type A)*
- How far is the student from achieving objectives of the entire course? *(Type B)*
- How far is the student from achieving objectives of the course modules? *(Type C)*
Some Interpretations of Progress

We identify three ways of measuring progress

- At a specific point in the course, how does the student’s performance fare against those of others? *(Type A)*
- How far is the student from achieving objectives of the entire course? *(Type B)*
- How far is the student from achieving objectives of the course modules? *(Type C)*
Some Interpretations of Progress

We identify three ways of measuring progress

- At a specific point in the course, how does the student’s performance fare against those of others? *(Type A)*
- How far is the student from achieving objectives of the entire course? *(Type B)*
- How far is the student from achieving objectives of the course modules? *(Type C)*
Some Interpretations of Progress

We identify three ways of measuring progress

- At a specific point in the course, how does the student’s performance fare against those of others? (Type A)
- How far is the student from achieving objectives of the entire course? (Type B)
- How far is the student from achieving objectives of the course modules? (Type C)
Progress Type A

- compares a student's standing at any point in the course to those of students from previous years.
Progress Type A

- compares a student's standing at any point in the course to those of students from previous years.
- Compared to how other students were doing at this stage, how well is this student doing now?
Progress Type A

- compares a student's standing at any point in the course to those of students from previous years.
- Compared to how other students were doing at this stage, how well is this student doing now?
- Prediction using data from previous editions of the course may provide the answer.
Progress Type B

- How far is a student from achieving course objectives?
Progress Type B

- How far is a student from achieving course objectives?
- E.g. Is the student’s expected final grade at any point a desirable outcome?
Progress Type B

- How far is a student from achieving course objectives?
- E.g. Is the student’s **expected final grade at any point** a desirable outcome?
- Good measurements at several intervals should provide reliable progress information.
Designing Prediction Models Accordingly

Procedure:

- Linear regression using numerical grades
- Data came from two courses - IG1 and PR2
- 8 weeks of PA activities
- 8 sets of data, 1 for each week
Data and Models

- Data for two courses
  - IG1 - Training (2012-2013), Test (2013-14)
## Data and Models (Cont’d)

<table>
<thead>
<tr>
<th>Progress Type</th>
<th>Number of Models</th>
<th>Training Set</th>
<th>Test Set</th>
<th>Test Set Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>8 (1 per week)</td>
<td>IG1=115, PR2=114</td>
<td>IG1=88, PR2=81</td>
<td>Previous edition of course</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>IG1=115, PR2=114</td>
<td>IG1=115, PR2=114</td>
<td>Data from previous weeks</td>
</tr>
</tbody>
</table>
Data and Models - Regression Variables

- Tasks Assigned
- Tasks Completed
- Questions Asked
- Questions Answered
- Votes Cast
- Questions picked
- Votes Earned
- Votes Earned Total Difficulty
- Votes Earned Total Relevance
- Votes Earned Total Interestingness
- Selected Q total difficulty
- Selected Q total relevance
- Selected Q total interestingness
Experiments and Results

Focus when evaluating student performance models on:

- How many of the students the model predicted not to be at-risk were actually at-risk and eventually performed poorly (False Positive Rates)
- How many of the students that the model predicted to be at-risk of failing were indeed at-risk (True Negative Rates).

In fact, FPR and TNR provide two interpretations of the same outcome. FPR = \frac{-TNR}{TNR}.
Experiments and Results

Focus when evaluating student performance models on:

- How many of the students the model predicted not to be at-risk were actually at-risk and eventually performed poorly (False Positive Rates)
Experiments and Results

Focus when evaluating student performance models on:

- How many of the students the model predicted not to be at-risk were actually at-risk and eventually performed poorly (False Positive Rates)
- How many of the students that the model predicted to be at-risk of failing were indeed at-risk (True Negative Rates).
Experiments and Results

Focus when evaluating student performance models on:

- How many of the students the model predicted not to be at-risk were actually at-risk and eventually performed poorly (False Positive Rates)
- How many of the students that the model predicted to be at-risk of failing were indeed at-risk (True Negative Rates).
- In fact, FPR and TNR provide two interpretations of the same outcome. $FPR = 1 - TNR$
Evaluation Metrics and Labels

- performance in making a prediction that is within a one grade-point range of the actual grade.
Evaluation Metrics and Labels

- performance in making a prediction that is within a one grade-point range of the actual grade.
- Positive - A prediction that is either A or B
- Negative - A prediction that is either C or D
• performance in making a prediction that is within a one grade-point range of the actual grade.

• Positive - A prediction that is either A or B

• Negative - A prediction that is either C or D

• Metrics: Precision (P), Recall (R), F1 scores, TNR, FPR
Progress Type A for Course PR2

At a specific point in the course, how does the student’s performance fare against those of others in the past?
Progress Type A for Course IG1

At a specific point in the course, how does the student’s performance fare against those of others in the past?
Progress Type B for Course PR2

How far is the student from achieving objectives of the entire course?

![Graph showing progress type B over weeks]
Progress Type B for Course IG1

How far is the student from achieving objectives of the entire course?
Outline

1. Introduction
2. A Semi-Automated Peer-Assessment Platform
3. Continuous Prediction for Monitoring Student Progress
4. Discussion and Conclusion
Discussion and Conclusions

- Studies using PA data to build prediction models are sparse.
- Many PA experiments are localised, a number of basic PA datasets available.
- How to monitor student progress using PA data:
  - PA data helps monitor progress using data from previous editions of courses
  - Using PA activities, we can measure how far a student is from achieving goals
- High levels of performance for both progress types
- Prediction results for Progress Type B better than Progress Type A
- High Progress Type B prediction results achieved early in the course
Discussion and Conclusions

- Studies using PA data to build prediction models are sparse.
- Many PA experiments are localised, a number of basic PA datasets available.
- How to monitor student progress using PA data:
  - PA data helps monitor progress using data from previous editions of courses
  - Using PA activities, we can measure how far a student is from achieving goals
- High levels of performance for both progress types
- Prediction results for Progress Type B better than Progress Type A
- High Progress Type B prediction results achieved early in the course
Discussion and Conclusions

- Studies using PA data to build prediction models are sparse.
- Many PA experiments are localised, a number of basic PA datasets available.
- How to monitor student progress using PA data:
  - PA data helps monitor progress using data from previous editions of courses
  - Using PA activities, we can measure how far a student is from achieving goals
- High levels of performance for both progress types
- Prediction results for Progress Type B better than Progress Type A
- High Progress Type B prediction results achieved early in the course
Challenges and Future Work

Challenges

- Task incompletion and increasing attrition rates towards the end of courses
- Need to wait for a semester to collect complete data

Future Work

- Better algorithms addressing task incompletion
- Integrating prediction models into the PA platform
- Introducing game-like competition features
- Automation of tasks → Question selection, detection of potential dishonest behaviour
Challenges and Future Work

Challenges

- Task incompletion and increasing attrition rates towards the end of courses
- Need to wait for a semester to collect complete data

Future Work

- Better algorithms addressing task incompletion
- Integrating prediction models into the PA platform
- Introducing game-like competition features
- Automation of tasks → Question selection, detection of potential dishonest behaviour
Thank you! Questions?