CCRC Community College Research Center

TEACHERS COLLEGE, COLUMBIA UNIVERSITY

Starting with the End in Mind: Mapping Current Students' Program Pathways Using the Transcripts of Completing Students

Matthew Zeidenberg

Community College Research Center

Teachers College, Columbia University

Council for the Study of Community Colleges

Orlando, Florida

April 21, 2012

The Approach

Identifying the course-taking patterns of credential completing students can:

- Reveal what courses students are actually taking in college (as opposed to what we think they should be taking)
- Suggest pathways for current students

Overview of Presentation

- Identifying programs of study for students still in school
- Designing an electronic advising system
- Understanding student focus
- Performing a continuous degree audit

Identifying the Program of Study

- We employ a machine learning method to identify the relationship between:
  - the patterns of courses taken by completing students
  - the credential they earned
- E.g., the computer can learn that nursing program completers take a certain set of courses with specific probabilities
- Removes the need to manually enter all the course requirements for each program
- Reflects real course taking, not formal requirements, which can differ
- Joint work with Marc Scott of NYU

The Learning Algorithm

- Naïve Bayes learning algorithm:
  - given a student's transcript, indicates what program that transcript belongs in
- Algorithm correctly identifies the program of 66% of the completers not used in training
- Probably most of the incorrect identifications are courses that are assigned to similar program; investigating this …

Assigning a Program of Study to a Non-Completing Student

- Learning method used for:
  - Assigning program of study to every student
- This can be used to:
  - Understand activity at a college
  - Advise students

Designing an Electronic Advising System

- Unlike Degree Audit systems, an electronic advising system:
  - Is based only on transcript data
  - Has no rules
  - Uses completing students as models for non-completing students
- We are prototyping an advising system to illustrate these ideas

Aspects of an Electronic Advising System (1)

- Student can select one of a few suggested programs based on courses taken to date
- Then the system would suggest courses:
  - Taken by completers in same program
  - Taken by students with similar course-taking patterns
  - That are next in a sequence (e.g., History 202 follows History 201)
  - That are associated (e.g., students who take Math 201 also take Chem. 201)

Aspects of an Electronic Advising System (2)

- Shows progress in selected program of study
  - Program courses
  - General education courses
  - Gatekeeper courses
- Uses chi-squared method to determine type of course
- Grades may influence system behavior
- For each program, typical order of courses shown
- Displays level of focus of course-taking

Determining Whether a Course is Program-Specific

- Chi-squared statistic determines:
  - Whether a course is appearing more often than expected
- If this statistic is significant:
  - Then this is a program course
- Courses like Math 101 are not program-specific

System Interface Layout

• List of most likely programs based on courses •Allows student to select one program	• Completed courses listed by term • Color-coded as meeting general education or program-specific requirements • Progress bars	• List of suggested courses highlighted within display of entire program

Deducing the Order of Courses in a Program

- For each completer:
  - Assess in which semester each course was taken
- Then assign each course to:
  - The semester when it is most likely to be taken (based on all completers)
- Select the top six courses by semester to assemble the course ordering
- In advising, such a course ordering could help determine where a student is in the program

Example: Associate of Science in Business Admin. at One College

Semester	1	2	3	4
Pattern based on student top enrollments	ENG 111 BUS 100 ITE 115 SDV 100 HIS 121 STD 100	ENG 112 MTH 163 HIS 122 MTH 166 MTH 174 SPD 126	ACC 211 ECO 201 SPD 110 BIO 101 BUS 200 PED 135	ECO 202 ACC 212 PED 116 MTH 271 BIO 102 MTH 241
Program as listed on web site	ENG 111 ITE 115 SDV 100 HIS elective MTH elective Science elective	BUS 100 ENG 112 PED 116 Science elective Math elective	ACC 211 ECO 201 General elective Public speaking Humanities P.E. or Rec.	ACC212 ECO202 General elective Humanities

Continuous Degree Audit

- Every term:
  - Compare transcript with requirements
  - Can see if each student is on track
- Similar to electronic advising system for students, but instead it is used administratively by faculty and staff

Methods of Determining Student Focus

- Students are less focused if:
  - They take courses from more departments
  - They take courses outside their programs
- Weak program classification implies weak program focus
- High vector entropy (disorder) indicates less focus:
  - Departmental course-taking vector
  - Program classification vector
- Some programs are more focused
  - Occupational programs are generally more focused than liberal arts programs

Completion Statistics: CBD College, AA in Liberal Arts & Sciences

Course Name	Enrollment Rank	Percentage of Earners Taking Course	Share of Total Course Enrollments by Earners	Cumulative Share of Enrollments by Earners
ENGLISH COMP 2	1	91.5%	3.3%	3.3%
ENGLISH COMP 1	2	86.6%	3.1%	6.4%
COLLEGE ALGEBRA	3	83.8%	3.0%	9.4%
INTRO/MICRO USAGE	4	83.8%	3.0%	12.5%
FUND OF SPEECH COMM	5	80.1%	2.9%	15.4%
INTERMEDIATE ALGEBRA	6	63.9%	2.3%	17.7%
INTRO TO PSYCHOLOGY	7	52.2%	1.9%	19.5%
STATISTICAL METHODS	8	48.7%	1.8%	21.3%
HUMAN GROWTH & DEV	9	47.6%	1.7%	23.0%
HUMANITIES	10	47.3%	1.7%	24.7%
INTRO TO PHILOSOPHY	11	46.1%	1.7%	26.4%
GENERAL EDU BIOLOGY	12	45.6%	1.6%	28.0%
PRIN/ECONOMICS-MACRO	13	44.8%	1.6%	29.6%
ESENTALS OF HUM NUTR	14	39.5%	1.4%	31.1%
SOCIAL ENVIRONMENT	15	35.4%	1.3%	32.4%

Completion Statistics: CBD College, AS in Registered Nursing

Course Name	Enrollment Rank	Percentage of Earners Taking Course	Share of Total Course Enrollments by Earners	Cumulative Share of Enrollments by Earners
PSYCHIATRIC NURSING	1	100.0%	2.6%	2.6%
OBSTETRICAL NURSING	2	100.0%	2.6%	5.1%
COMM HLTH NURSNG LAB	3	100.0%	2.6%	7.7%
PEDIATRIC NUR CLIN L	4	100.0%	2.6%	10.2%
ADV MED-SUR NUR CLIN	5	100.0%	2.6%	12.8%
PSYCHIATRIC NUR CL L	6	100.0%	2.6%	15.3%
OBSTETRICAL NUR CL L	7	100.0%	2.6%	17.9%
PEDIATRIC NURSING	8	100.0%	2.6%	20.4%
INTRO NUR MATH&PHARM	9	98.7%	2.5%	23.0%
HUM ANAT & PHY 2	10	86.8%	2.2%	25.2%
HUM ANAT & PHY LAB 2	11	86.3%	2.2%	27.4%
FNDMTLS NUR SKLL LAB	12	82.4%	2.1%	29.5%
FNDMTLS NUR CLIN LAB	13	82.4%	2.1%	31.6%
HUMAN GROWTH & DEV	14	82.4%	2.1%	33.7%
MEDICAL-SURGICAL NUR	15	81.9%	2.1%	35.8%

CCRC Community College Research Center
TEACHERS COLLEGE, COLUMBIA UNIVERSITY

Take-aways

- The following are feasible and could be useful for student advising and in understanding student activity:
  - Assigning programs of study to current students
  - Identifying pairs of courses found in sequence or often together in transcripts
  - Identifying courses often taken by completing students
  - Identifying the sequences of courses of completing students
- Programs and students within them vary in focus
  - To boost completion, boost focus?
- Electronic student advising could boost student outcomes

For more information:

Please visit us on the web at
http://ccrc.tc.columbia.edu,
where you can download presentations, reports, and briefs, and sign-up for news announcements.
We're also on Facebook and Twitter.

Community College Research Center
Institute on Education and the Economy, Teachers College, Columbia University
525 West 120th Street, Box 174, New York, NY 10027
E-mail: ccrc@columbia.edu
Telephone: 212.678.3091

CCRC Community College Research Center
TEACHERS COLLEGE, COLUMBIA UNIVERSITY