Can machine learning predict when STEM students switch majors?

John M. Aiken[1], Marcos (Danny) Caballero[1,2]
[1] Center for Computing in Science Education, Department of Physics, University of Oslo, Oslo, Norway
[2] Department of Physics and Astronomy & CREATE for STEM Institute, Michigan State University, East Lansing, MI

Research Question

Are there predictive initial conditions of students choice to finish a physics degree or switch to engineering? Initial conditions include:

1) Performance in courses
2) Time when course is taken
3) Prior preparation (as indicated by transfer credit)
4) Taking engineering courses or physics major courses
5) Demographic information (gender, ethnicity)

Data

Our data set includes data for all students beginning their degree programs at MSU from 1993 to 2013 (N=263412). We reduce this data set to include only students who registered as a physics major at some point and then either matriculated through physics or an engineering program (N=1422). The data includes all timestamped courses taken by students, demographic data (gender, ethnicity), course grades, and semester by semester information on major declaration. We prepare the data by reducing all variables to categorical: (high/low/transfer grades), taking courses before/after first year enrolled, whether they took any engineering course as a physics major, and whether they took modern physics.

Prior Work

Previously we have demonstrated that introductory physics and calculus performance and the time when courses are taken may be predictive of whether students will complete a degree in physics or transfer [1].

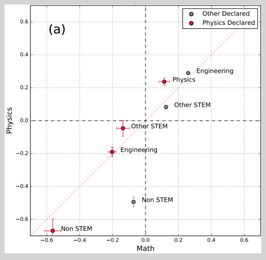


Figure 1: Physics students receiving BS/BA degrees in physics/astronomy are above average in introductory course performance in comparison to students who move to different programs before graduating. Course scores are grouped by graduating degree and z-scored then averaged [1].

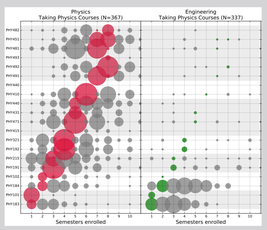
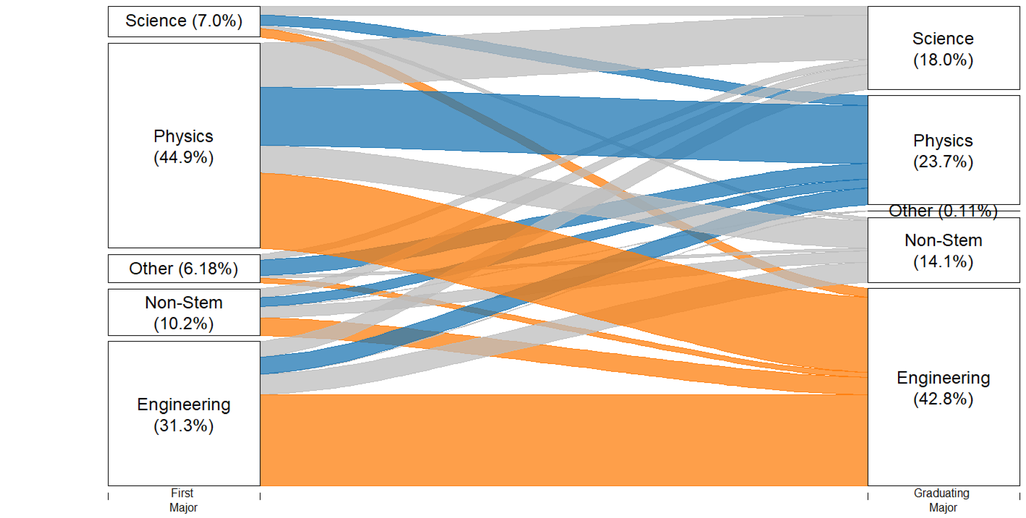


Figure 2: **:** Students who switch to engineering majors typically take physics after the recommended semester [1]. Bubble size indicates relative proportion, color indicates suggested semester to take course by dept. of physics at MSU.

This study was funded by the Norwegian Agency for Quality Assurance in Education (NOKUT) which supports the Centre for Computing in Science Education.

Understanding Students who switch to Engineering

Performance and course schedule may impact student persistence in a physics major [1]. Additionally, experiential reasons linked to gender and ethnicity may impact students switching away from STEM careers [2]. We explore the raw data for students declaring physics then receiving a bachelors in physics or an engineering program. Students who switched to other fields or did not complete MSU degree programs are not included in this data. This data analysis below demonstrates a time, performance, and demographic dependence that is not supported by the model.



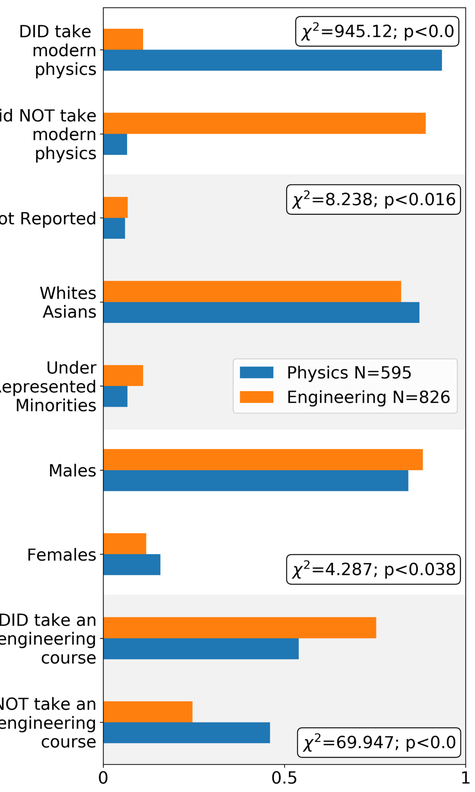
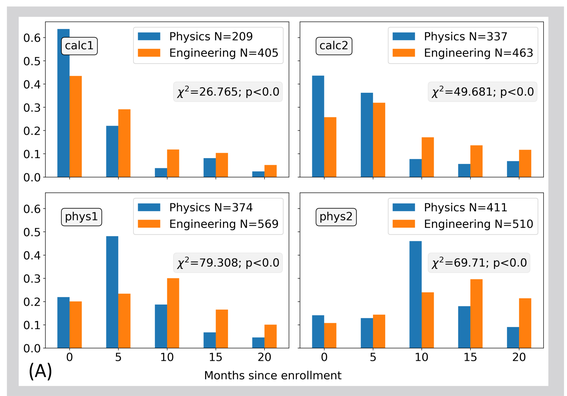


Figure 3: In comparison to students who switch to engineering, students who get physics degrees are more likely to have taken modern physics, be white or Asian, be female, and have not taken an engineering course. Chi-squared statistics are reported in the plot, all are significant (p<0.05).



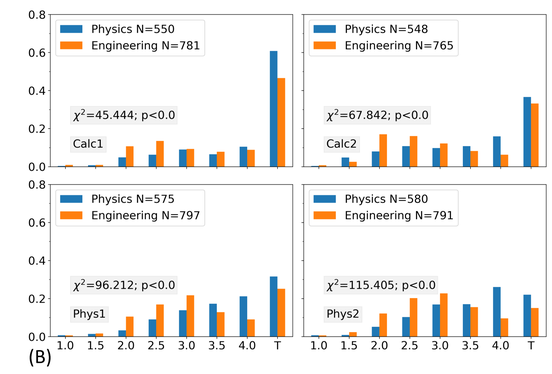
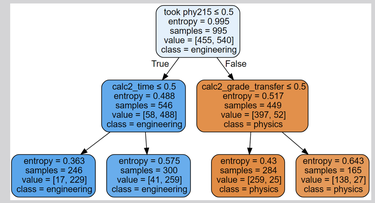
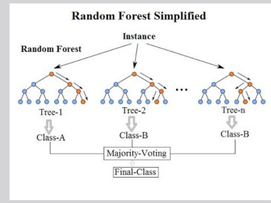


Figure 4: Students who receive physics degrees are more likely to take intro physics and calculus courses earlier (A) and have higher grades in those courses than students who switch into engineering programs (B). Students who receive physics degrees are also more likely to have transferred grades from another institution (e.g., Advanced Placement).

Random Forest Classifiers

Random forest classifiers use a forest of decision trees to produce majority vote decisions to label a sample according to the class it belongs. The model feature importances are then derived from average reduction in entropy for each node.





Model Results

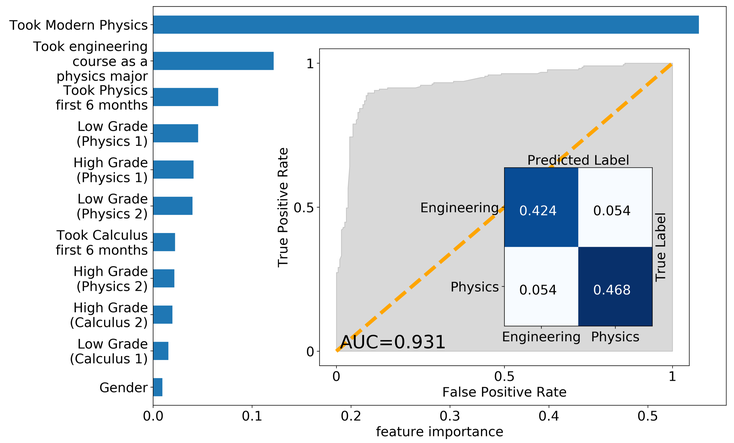


Figure 5: The model accurately predicts students ultimate degree if they end up in as a physics major or an engineering major (ACC=0.892, AUC=0.931). The two most important features of the model are whether the students took courses related to their ultimate degree program.

Key Takeaways

- Performance and timing of intro physics and calculus courses was not as important for students to remain as physics majors as suggested in Aiken and Caballero 2016.
- Taking the first major course (a course in modern physics) was very important for determining whether a student would leave for engineering or not, this is a complementary result to Rodriguez et al. 2016
- Other factors such as transfer credits (which may indicate prior preparation), gender and ethnicity were not important indicators in our model

References

1. JM Aiken and MD Caballero, Methods for analyzing pathways through a physics major, PERC 2016.
2. Seymour, Elaine, et al. Talking about leaving: Why undergraduates leave the sciences. Vol. 12. Boulder, CO: Westview press, 1997.

Have a comment or want to get your institution's data involved in this project? Leave your ideas or contact info here!