



Grade-Based Performance Prediction at the University of Michigan Benjamin P. Koester, William Murdock, James Fogel, and Timothy A. McKay

Abstract

Many of the goals of learning analytics are rooted in the desire to predict student outcomes. These predictions can be used as inputs to student advising, the identification of outlier groups of students, or as a baseline to evaluate the effect of pedagogical "treatments" employed by instructors. Grades are one metric of outcome that is abundant and meticulously recorded. They contain rich information about performance and learning in a course, but also depend on the subject, term, instructor, and a student's peers. Here, we examine the predictive power of grades and show how accounting for the sources of variation underlying their distribution can be leveraged to greatly improve prediction.

Predicting Outcomes as Measured by Grades

Brute force linear regression models (not shown) that relate grades to various potential predictors (GPA, gender, ethnicity, total credits, college, and ACT/SAT composite) provide one window into the power of different predictors. They provide strong evidence that grade-point average in other classes (GPAO) is the strongest predictor of grades in future classes. This has been leveraged in studies of male and female outcomes in intro STEM classes (e.g. Physics 140). GPAO captures much but not all of the variation in grade among students: a large gap in performance between genders is observed (Figure 1).



Figure 1. GPA in other courses (GPAO) used to predict course grades. Using GPAO as a predictor, a 'grade penalty' was identified that varies with GPAO: mean grades for male and female are plotted in bins of GPAO.

- <u>SIMPLE GRADE PENALTY</u>: the average male and average female difference between their actual grade and expected grade (GPAO).
- MATCHED MEAN GRADE: Is the gender gap really due to gender? We match (Hansen, 2009) males and females on GPAO, ACT composite, total credits, college, and ethnicity and compare the mean grades of these matched samples and find that the gap persists.

Grades at the University of Michigan

Students' GPAs are determined by grades awarded by many different departments whose content, course structure, and grading policies differ. The mean grades for the top 10 courses by enrollment are shown for many departments, ordered by enrollment-weighted mean, in Figure 2. In addition to strong variations in mean grade between departments, some show clear grading patterns that depend on course level.

										7
104	105	106	107	150	151	254	256	270	280	LS&A First Year Seminars
118	304	310	362	391	392	401	402	406	490	School Of Education
139	139	140	149	344	345	346	347	348	349	School of Music, Theatre and Dance
211	221	231	321	331	418	419	450	458	499	Biomedical Engineering
230	240	350	351	354	363	368	453	459	468	Office of International Programs
100	101	101	102	102	122	201	202	281	331	Near Eastern Studies Department
101	102	221	231	232	243	322	325	326	386	Germanic Languages & Lit Dept
100	201	204	205	206	209	240	301	374	399	American Culture Program
102	111	200	209	210	211	272	315	370	375	Department of Linguistics
220	240	253	270	295	300	324	375	400	483	Women's Studies Department
101	102	125	126	201	202	220	225	226	230	Asian Languages And Cultures
100	110	120	121	130	150	151	220	231	300	School Of Art And Design
111	120	230	240	250	270	280	303	370	401	Psychology Department
122	210	245	252	254	354	356	358	454	456	School Of Nursing
101	101	102	191	192	222	231	232	372	385	Classical Studies Department
102	105	110	111	139	201	211	232	302	360	Program in the Environment
124	125	223	225	239	240	313	317	325	367	English Language & Literature Dept
100	101	103	110	151	195	280	390	455	490	Engineering Undergraduate Educ
101	111	140	160	300	314	353	389	489	496	Political Science Department
122	201	202	230	280	296	310	312	381	481	Studies In Religion
200	236	236	272	290	350	360	366	366	370	Screen Arts and Cultures
101	161	272	285	298	330	344	364	365	368	Anthropology Department
103	111	111	340	358	450	451	458	490	495	Department of Afro-American and African Stud
101	102	112	212	222	250	251	271	272	394	History Of Art Department
101	101	110	111	111	230	241	320	330	340	School Of Kinesiology
110	160	161	201	218	241	266	318	322	396	History Department
300	306	310	400	418	422	427	428	429	436	Molecular, Cellular, and Developmental Biology
211	212	260	303	325	351	360	402	421	431	Civil & Environmental Engr
101	102	103	104	106	111	112	115	127	142	Astronomy Department
409	410	411	412	431	432	434	462	485	486	College Of Pharmacy
100	210	239	256	301	306	337	375	418	438	
101	102	111	211	351	361	371	381	439	458	Sch Of Nat Resources & Environ
100	101	102	210	303	305	310	344	345	368	Communication Studies
215	225	245	285	305	315	325	335	345	405	Sociology Department
271	272	300	300	300	300	301	312	350	403	Aerospace Engineering
										School of Business Administration
180	181	196	201	202	232	303	355	359	361	Philosophy Department
220	242	250	330	350	360	412	420	480	489	Materials Science & Engineering
230	330	341	342	343	344	360	460	466	487	Chemical Engineering Department
312	313	314	315	316	317	322	323	326	425	College of Architecture & Urban Planning
125	126	127	128	135	136	140	141	240	241	Physics Department
211	235	240	250	350	360	382	395	450	495	Mech Eng & Applied Mech Dept
125	126	130	210	211	215	216	230	241	260	Chemistry Department
201	202	265	310	316	333	334	366	373	425	Industrial-Operations Engr Dep
100	250	350	401	402	408	412	425	426	470	Statistics Department
183	203	215	270	280	281	314	370	482	496	Electrical Engr & Computer Sci
101	102	103	103	231	231	232	232	275	276	Romance Languages Department
118	162	171	172	173	207	225	226	305	310	Biology Department
101	102	310	340	395	398	401	402	404	435	Economics Department
100	181	183	198	270	280	303	370	380	482	Program In Computer Science
100	102	103	105	106	107	110	111	113	222	Earth and Environmental Sciences
	105	115	116	156	215	216	217	417	425	

Figure 2. Mean grades in courses at Michigan since 2005. Color shading ranges from red (mean grade = 2.65) to light yellow/white (mean grade = 3.85). The top 10 courses by enrollment (left to right) are shown for departments with at least 10 courses with enrollments typically > 50. Departments are ordered top-to-bottom by enrollment-weighted mean grade. Department names are color-coded by division: science/engineering (red), social science (orange), humanities (blue), other (black).

Department of Physics, University of Michigan

Grade based performance prediction: GPA and beyond

<u>GPA</u> provides our baseline grade predictor. It is a credit hour weighted-mean of grades that is agnostic to subject, student performance, term, etc. **<u>GPAR</u>** or grade points above replacement (e.g. Caulkins et al., 1996) provides a first refinement to performance prediction. To calculate GPAR, we form the credit weighted average of the difference between student grade and course mean grade. GPAR values center on zero, with positive scores for above average students. GPAR accounts for variations in grading practice across departments, but does not account for variation in the academic strength of student cohorts in each class. **<u>SFE</u>** (Murdock, Fogel, et al. *in prep*) This model captures the course/term variability as well as the strength of the peer group. Each grade in a course is a linear combination of a course-term invariant student-effect and a student-invariant course-term effect:

In practice, a LS solution requires manipulating a sparse ~1.5 x 10⁵ by 5 x 10⁶ matrix. We use the sparse-matrix approximation of Arcidiacono (2012) to solve for the coefficients.

Testing performance predictors: GPA, GPAR, or SFE?

To compare these three performance predictors, we calculate them for all students based on grade data up to and including Fall 2013, then use them to predict Winter 2014 course grades. To compare, we compute the Pearson correlation coefficient (-1 < ρ_{P} < 1) between the predicted and actual grades for Winter 2014 and find that the SFE model performs substantially the better:

In individual courses, we can examine this in more detail. Each model effectively



Figure 3. Single course examinations of GPA and SFE as predictors. For selected courses in Chemistry in Fall 2013, we compare GPA and SFE for their ability to predict rank, using the Spearman rank correlation between predicted and actual rank.



 $Grade_{ict} = StudentFE_i + ClassFE_{ct} + \epsilon_{ict}$

 $\rho^{\rm SFE}{}_{\rm P} = 0.734$

 $\rho^{\text{GPAR}}_{P} = 0.423$

 $\rho^{\text{GPA}}_{P} = 0.403$

predicts a rank that we can correlate with the actual grade (or rank) in the course.

