Correlating ACT and Math Placement Test Scores to Performance in First Year Engineering and Math Courses

Jeff Kastner, Joni Torsella, Gregory Bucks, Kathleen Ossman,
Rodney Roseman, and F. James Boerio
University of Cincinnati, Cincinnati, OH 45221
Email: Jeffrey.Kastner@uc.edu

The College of Engineering and Applied Science (CEAS) at the University of Cincinnati (UC) recently enhanced the engineering curriculum by including three classes required by all first-year engineering students. Engineering Foundations and Engineering Models I are taken during the first semester while Engineering Models II is taken during the second semester. The Engineering Foundations class introduces students to the various fields of engineering through hands-on laboratory exercises while the Engineering Models I & II classes provides students with an introduction to MATLAB® and computing methods for solving calculus based engineering problems. All three classes take place in a new 10,000 square foot Learning Center dedicated to the education of first-year students in the College of Engineering and Applied Science (CEAS) and opened during the fall of 2012. This new facility features two state-of-the-art classrooms where 1 hour lecture sections are held and three “project rooms” where 2 hour recitation sections meet. In the project rooms, students work in small groups on problems involving applications of calculus in engineering.

Engineering Models I and II are taken concurrently with Calculus I and II and are intended to help students understand the applications of calculus to engineering problems. It has been well documented in the literature that success in engineering is closely related to success in calculus.[1] Our own research is consistent with the literature and has shown that students that earn a grade of C^+ or better in their first calculus course have about a 75% probability of graduating from CEAS; students that earn grades less than C^+ in their first calculus course have almost no chance of graduating from CEAS.

Prior to the beginning of each semester, we believe it is important for each instructor to understand the skill set of their students. This will ensure that the instructor does not expect more than his/her students can produce or reteach concepts and exercises that students have already mastered in high-school. Some useful metrics that provide some understanding of our students’ skill set are their ACT Math score, ACT composite score and the UC Math Placement test (MPT). This work describes an effort to correlate ACT and MPT scores to a student’s performance in first-year STEM classes. The advantage of the ACT scores is that these scores reflect a student’s competence in English, Math, and Reading. One challenge with the ACT metrics is that they are measured about a year before the student steps foot onto campus. Another challenge with the ACT math metric is that many engineering students perform well on that section of the ACT test. This leads to a smaller spread of that specific data. The MPT score may help remedy some of these challenges because it does not have the one year time delay and it tests more specific mathematical concepts which may lead to a larger data spread.

The metrics were specifically correlated to three courses during the fall semester (1) Calculus I, (2) Engineering Foundations, and (3) Engineering Models. The overall performance in Calculus

I was compared to ACT and MPT scores. For the Engineering Foundations class, the ACT and MPT scores were correlated to performance on the lab reports completed in Engineering Foundations. Four questions were asked of the data for this: (1) How did the students do on all 5 reports, (2) How did the students do on their first report, (3) How did the students do on the last 4 reports, and (4) How did the students do on the final report. For the Engineering Models class, the ACT and MPT scores were correlated to the performance in the class and the final exam.

Data based on human experiences often lack a high correlation due to the variability involved in human interactions. Certainly, data based on freshman engineering students is no different. While a student may be extremely academically talented, there are other factors affecting their success. For example, this may be their first time away from home, they may not have the discipline to structure their day for success, they may have more financial responsibilities, and they may be involved in relationships which are not conducive to success. Furthermore, we believe the time delay between taking the ACT and class performance may affect the strength of the correlation between these two variables.

A typical statistical analysis was done between datasets where correlation values and trend lines were studied. This data alone tended to have low correlation values ($R^2<0.2$). The data was split into statistical clusters to improve the statistical significance of the correlation values. The kmeans cluster analysis[2,3] was done using MATLAB®. The initial analysis began by forming kmeans clusters on data from the Engineering Models and Calculus I class. The Engineering Models I class used the final grade and two separate cluster analysis was done for the first exam in Calculus I and the final grade in Calculus. The two different Calculus subsets were formed based on implementing this analysis into the classroom. The first test in Calculus is early in the semester and the data from this could be used immediately in the Engineering Models I class. The final Calculus I grade was used since it had a very significant correlation to the Engineering Models I final grade ($R^2 \sim 0.4$) and the data could be used to understand the student population in Engineering Models II.

One challenge with a cluster analysis is the ability for the results to be repeatable. The analysis can be done to any number of clusters, but as the numbers of clusters increases, the repeatability of the cluster groups reduces. This is because as the number of clusters increases, the number of boundaries between clusters increases where the repeatability is weakest. Three clusters of the Models I and Calculus I were initially formed and found to 100% repeatable. The three clusters were for students who did very good, average, and below average in Calculus I. The cluster with the best Calculus grade also had the highest average of any cluster for the Models I grade. The other two clusters followed this trend by the average Calculus I grade was in the same cluster as the average Models I grade. The cluster with the below average Models I grade also had the lowest Calculus I grade.

All three clusters were further broken into sub-clusters. This begins to form a hierarchy of clusters which allows us to further analyze each cluster. For example, the below average Calculus I cluster had a very large spread in the Engineering Models I grade. The standard deviation for this cluster was 10% while the other two clusters had a spread of only 5% for the average group and 3% for the above average. Breaking the clusters into a few more sub-clusters might help us understand why one student did well in Models and another did not. The challenge with sub-clusters is that they require more data. Currently, we are analyzing 200 students and each cluster has between 45 and 80 students. Thus breaking a cluster down further into sub-clusters will begin to remove the statistical significance. However, this part of the analysis will continue as more data becomes available.

Even with the large variability, a couple important results appear. First of all, for a given cluster the average of the composite ACT score can show some validity to the final grade. It is also shown that in the Engineering Models class (MATLAB® programming) that ACT English and ACT Reading score is 3.3 and 2.4 points lower for students who struggled with both Calculus I and Models I when compared to the students who did above average in both. This cluster of students who struggled in both classes also had an ACT Math score 2 points lower. The cluster of students with the average Calculus I grade and the average Models I grade had ACT scores between the two other clusters.

The best way to further the data conclusions will be to have more students. The current dataset was formed from half of the first year students in the College of Engineering, who took the ACT and took Calculus I the fall semester. The other half of first year students will be analyzed to increase the number of samples. This should provide a total of 400 students per year for this analysis. Other clusters not included in the analysis include the students enrolled in Calculus II during the fall semester. The classes are also taken by non-engineering majors, so the ACT scores for this cluster have not been accesses but will in the near future. :

The results from this analysis will be used to generate a questionnaire for first-year engineering students and then use the data to form teams of students that are most likely to succeed together. The questionnaire/survey will help us better understand the tools our students have going into the class and how to actively steer the class so they are properly taught. All our first-year engineering classes emphasize team work. However, our formation of teams is very ad hoc. The questionnaire and understanding of their Calculus Test I, ACT and MPT scores will help us form teams that can enhance the active learning environments in the first-year engineering classes. The results and efforts made will then be statistically tracked to see if there is an increased retention and increased performance in future classes.

(1) Klingbeil, N., Rattan, K., Raymer, N., Reynolds, D., Mercer, R., Kukreti, A., and Randolph, D. (2008). The WSU Model for Engineering Mathematics Education: A Multiyear Assessment and Expansion to Collaborating Institutions. In Proceedings of ASEE Annual Meeting, Pittsburgh, PA, June 22-25, 2008.
(2) Lloyd, S., (1982). Least squares quantization in PCM, IEEE Transactions on Information Theory, Vol. 28, No. 2, pp 129-137.
(3) Estivill-Castro, V. (2002), Why so many clustering algorithms, ACM SIGKDD Explorations Newsletter, Vol 4, No. 65.