

Frequently Asked Questions

COVID-19: Impact on Predictions

As educational institutions quickly adjust to safely support their students during these unprecedented events, such as moving all services and coursework online, Civitas Learning® understands that administrators will have questions about how our predictive models and scores will be affected given these sudden changes.

How will our current student predictions be affected and can I trust prediction scores during this time?

Civitas Learning does not expect that the disruption will have a major impact on current predictive models and scores, and the prediction scores should still be used to identify potentially at-risk students during this uncertain and challenging time.

Our models take into account more than 100 variables across student enrollment, academic performance, pathway progress, demographics, LMS engagement, financial aid, and more. Most of the variables that do measure online engagement are calculated relative to the enrolled section or course. That is, for example, a student's engagement in MATH101 is only compared to other students' engagement in the same MATH101 section. Only a few of the model variables, such as the Average Count of LMS Gradebook Activities, are directly affected by increased online engagement. While we do not anticipate that alterations in these variables will significantly change prediction scores, we will closely monitor the affected variables, along with scores, as the situation evolves.

How will our model performance be affected by these sudden changes? How will this impact student predictions long term, e.g. Fall-to-Fall persistence predictions?

While we don't expect the current prediction scores to change significantly and we consider them as accurate as can be expected given the most recent historic information known at the time the current models were trained, we must acknowledge the unknown effects that this unprecedented situation will likely have on student behaviors and outcomes for this term.

Students will not only be grappling with sudden changes in how they engage with your institution and their coursework, but may also be suddenly facing additional life challenges with their health, employment, food and housing insecurity, child care, and technology access just to name a few examples. These factors, along with new ways institutions are reaching out and supporting their students in response to the crisis, will likely affect student outcomes. Current prediction models are not aware of these new factors. Because of this, we are prepared to potentially see large amounts of data changes (non-stationarity) and unexpected student outcomes. Here,

non-stationarity refers to changes in data distributions, availability, and importance in predicting student persistence.

Specifically, these are some factors that we expect may influence model performance during this uncertain period:

1. *Increased LMS usage by moving instruction online:* While many of our LMS model variables are based on relative activity levels of students in the same section or course, their predictiveness still rely on having students with a larger spectrum of low through high LMS engagement. If this spectrum of LMS engagement becomes less varied (i.e. students within the same section are engaging with the LMS about an equal amount), which might happen in the instance of mass migrating all coursework online and enforcing strict and consistent usage of the LMS, there can be a noticeable impact on model performance. In addition, increases in absolute LMS activity-based features like Average Count of LMS Gradebook Activities could lead to optimistic predictions.
2. *Impact on social and psychological factors:* While we infer social and psychological factors through various derived model features, we do not yet know a priori how the amplified feelings of isolation, stress, and uncertainty can manifest in these derived features that encompass enrollment behaviors, engagement, and consistencies in academic performance and pathway progress.
3. *Heightened vulnerabilities for certain socioeconomic statuses or those who suffer from health-related issues:* The social and economic costs of disease progression are not yet well understood given current strategies to suppress transmission at extreme costs. Since institutions typically do not capture dynamically-changing economic dislocations that can affect students, there is a likely gap in the current financial aid variables used in our models during this situation. Furthermore, students might encounter sudden job loss or reduced income, sudden increase in childcare responsibilities, sudden health crisis (student or family), and other unexpected life changes not directly captured by institutions and used in our models.
4. *New policies and interventions at an institutional and governmental level:* These are unknown environmental changes that may not be captured in our current model data and can have a material impact on student outcomes and model performance.

Given this evolving situation, our Data Science team will be closely monitoring your model variables, prediction scores, and student persistence outcomes to ensure there is immediate awareness on any factors that require attention. Using this information, the Data Science team will determine a plan for how to address these changes in future model retrains and improvements. We will be communicating in the near future about how we use the data from these affected terms in our models and other dependent application features going forward.

How will COVID-19 affect Impact analysis?

Since current predictive models are not able to accommodate the many unknowns specific to the effects on student persistence from governmental and institutional responses to COVID-19, any Impact analysis of the student success programs from this turbulent time period (e.g. Spring 2020) should be carefully and cautiously reviewed. The impact results for this time period will likely only be valid if:

- Both participant and eligible comparison groups are from the same time period (i.e. no “pre-post” matching across different terms)

AND

- The unknown confounders are equally distributed to both participant and eligible comparison groups

Impact analyses that use pre-post matching (when the participant and eligible comparison groups are from different terms) may be inaccurate since current and previous conditions are radically different. Therefore, we do not recommend pre-post matching for current Spring term student success programs.

However, we also anticipate that many institutions will run COVID-19-related interventions to all affected students, and that can cause challenges with having a large enough eligible comparison group for baseline matching (matching within the same term). Where information and data are provided, our Data Science team will be reviewing across- and within-institutional Impact analyses that include this Spring time period to determine the potential impacts of COVID-19 interventions on student success, and will share those insights as soon as they are available.

How will this term be addressed in future model training?

If through our Data Science team’s monitoring we see that student-level term-day features and/or persistence outcomes are quite different from recent historical patterns, we will throw out this Spring 2020 term data from future model training.

If we move to a Pass/Fail for all classes rather than grades, will that impact models?

Moving to a Pass/Fail for all classes will impact several model variables such as prior term GPA, prior year GPA, and cumulative GPA as we progress into the Fall term. Since these are commonly important features for predicting persistence, such a move will likely have a significant impact on model performance in future terms. In these cases, please alert your Customer Success Manager as soon as you plan to implement major changes and our Data Science team will evaluate how to best proceed and account for these changes from a modeling and scoring perspective.

What happens to the model if we cancel the rest of the term (i.e. call it done when we went on quarantine)?

Assuming business as usual in the upcoming Fall term, we can retrain the models by removing the current Spring term. However, this removal would assume that credits earned and course grades for the COVID-19 affected terms will be null. If students are allowed to earn partial credits, please alert your Customer Success Manager as soon as possible and our Data Science team will evaluate how to best proceed and account for these changes from a modeling and scoring perspective.

How will increased LMS adoption affect engagement scores for Inspire for Faculty (IFF) customers? Can we add on-ground courses being converted into online courses?

Student engagement score models are based on *relative* LMS activity patterns in each section or course, so we do not foresee significant changes in the engagement scores for current courses in IFF.