Connecting the Dots in Student Dropout Data

MeasurEDU™ Case Study

Drew Ceccato and Jeremy Hodgson

wgu labs

Contents

Executive Summary

Universities are swimming in data, yet often struggle to connect the dots to timely student support. **Early Alert Systems (EAS)** such as Dropout Detective and Civitas are expensive, require manual data uploads, result in counselors sifting through large data sets, and often aren't built to meet the unique needs of small and midsize institutions.

The primary function of an EAS is to provide timely notifications for both faculty and staff when it is determined that a student may be at risk of failing a course or dropping out entirely. Such systems may utilize a variety of data inputs, including demographic information, financial aid status, assessment scores, course norms, and interaction data. In July 2020, The College Innovation Network found that of its member institutions, many expressed a desire to explore the potential of utilizing Early Alert Systems (EAS) in service to their students. Multiple members indicated dissatisfaction with their current EAS and a desire to promote greater collaboration between instructors and campus services. With cost being identified as

a major barrier for institutions and programs that would benefit most from the utilization of an effective EAS, exploration for a better solution was undertaken in partnership between WGU Labs and a large online institution of higher education (IHE). The goal of this partnership was to develop a solution that could be more effective in detecting students at risk of dropping out and provide that valuable solution at a price point accessible to the institutions that would benefit the most.

The resultant product is a software solution called **MeasurEDU™**, which was designed specifically to simplify the journey from data to actionable insights. Under the hood is a proprietary API that automates the data extraction process, displays insights via easy-to-use dashboards, and provides timely curated data for counselors and advisors to easily identify students in need. In doing so, MeasurEDU brings the promise and potential to be an effective alternative solution to the industry standard.

THE GOAL

Build a system capable of identifying students at risk of failing a course or dropping out

Provide this service at an accessible price point to institutions that would benefit the most

Key features developed and showcased within MeasurEDU not found in other EAS are as follows:

- **Flexible risk score cut-offs** that allow college staff to set thresholds as they tune their insights to specific parameters to best identify persistence issues and focus their efforts where they are needed most.
- **Summary list of factors** that contribute to a student's risk score that provides valuable insight into what is causing the flag and offers a starting point for intervention.
- **Weekly updates** to student risk scores provide the opportunity to monitor intervention effectiveness while a term is in progress rather than waiting for end-of-term results.
- **Sophisticated Machine Learning model** that identifies the most impactful measures from across the vast expanse of data in the college's LMS and SIS.

In partnering with an online IHE, testing of MeasurEDU was planned as a pilot to go head-to-head with Dropout Detective, an industry-standard solution currently in use at the college. The pilot allowed for a direct comparison of predictions from Dropout Detective and MeasurEDU — and real-life student outcomes. The goal was to better understand how much more effective MeasurEDU was in capturing dropouts while also narrowing the pool of students exhibiting risk factor indicators that would require attention from college staff. Identical live student data was run through both systems within a controlled timeframe in order to see which platform was better equipped to predict dropout rates and risk scores.

The data used for this analysis included the following:

Initial Student Status

A complete list of active students who began at the IHE in the fall term on October 2, 2023, which included 10,576 unique student IDs, was filtered down to only those students with relevant active statuses at the start of the term. This list came directly from the IHE's data sources, and the data in MeasurEDU was confirmed to be in alignment. This list represented the initial status of students for the pilot analysis.

MeasurEDU Predictions

From the MeasurEDU Student Journey module within our solution, a unique risk score was calculated for each student. Risk Scores are based on 194 unique data points that cover all aspects of a student's journey. This data describes a student's academic history in terms of things like cumulative GPA and cumulative credits earned. A student's more recent academic performance and engagement are also represented in the data with scores on assignments and quizzes as well as attendance and clicks in online learning tools. A Random Forest Classifier consumes this vast collection of data to determine which are the most relevant and impactful measures, then a sophisticated machine learning model calculates a unique score for each student based on these curated and highly personalized data points. Students with an above-average risk score were labeled as predicted drops.

Dropout Detective Predictions

A list of students that Dropout Detective had identified as likely drops for the October 2 term was used to evaluate the accuracy of this tool's predictions.

Final Student Status

Based on the initial list of active students for the October 2 term, students were identified who switched from an active status to a dropped status, on leave status, or completed status at any point between the term start and the census of the following term. This list served as a source of truth, to which the predictions from MeasurEDU and Dropout Detective could be compared.

The performance of both tools was evaluated using confusion matrices and several recognized performance metrics. Ultimately, MeasurEDU was able to accurately predict more drops with a smaller number of students assigned to the risk group - 136 drops out of 972 at-risk students. Comparatively, Dropout Detective accurately predicted 105 drops out of 1496 at-risk students.

At the conclusion of the head-to-head comparison, our solution outperformed Dropout Detective in four out of four model performance metrics, highlighted by a 6.23 percentage point improvement in overall accuracy as well as a more refined risk pool, 35% smaller than that which Dropout Detective predicted. This smaller and more precise risk pool means that intervention efforts can be more accurately targeted at those students who are most likely to benefit from them.

It is important to note that this was the first test of the MeasurEDU system in a live environment, and we fully expect these numbers to improve as the system is tuned and aspects of the machine learning engine are refined. Even with no tuning of the model on this initial launch, the findings showed that MeasurEDU was able to outperform Dropout Detective in its ability to accurately measure student dropout risk. It is expected that with further usage and refinement, MeasurEDU will continue to improve its accuracy and facilitate more effective support for at-risk students.

Comparative Analysis

Data Quality Check

The accuracy of data in MeasurEDU was assessed by recreating the IHE's "Total Student Population by Status" dashboard, employing MeasurEDU's SQL query environment and BI dashboard tool for successful data validation and alignment of key student details.

To assess the accuracy and quality of the data that is extracted from the IHE's sources and imported into the MeasurEDU environment, an analytics dashboard was created in the MeasurEDU BI Dashboards tool. This dashboard is based on the "Total Student Population by Status" dashboard that exists in the online college's PowerBI environment. The intent was to recreate their dashboard and confirm that the pertinent values aligned in the MeasurEDU environment. A SQL query was created in the MeasurEDU SQL Lab, which generates a list of all students at the online insitution and includes their current status, program and degree enrollment details, the label of the vertical or school in which the student is enrolled, and other relevant details to identify current active students. This list was then used to generate visualizations representing the number of students grouped by their various active statuses, degree types, program assignments, and verticals.

A key focus was on aligning the Total Student Population by Status. There are many status labels that students may have, with several options that indicate various states of active status, several that may be assigned to students at different points in their enrollment and matriculation path, and others that are assigned as students leave the school, suspend their studies temporarily or indefinitely, or complete their programs. With so many status variables, this was a critical data label to confirm in the new system. Some assignments of specific degree types remained unresolved, as many Bachelor, Certificate, and Masters students were grouped into the Non-Degree category based on the intricate case logic within the SQL query. In all cases, however, the total student counts and their active statuses were confirmed within an acceptable tolerance, given the possibility that statuses may change daily and the report ETL was limited to once daily.

Predictions Analysis: Dropout Detective & MeasurEDU

MeasurEDU's Student Risk Score, derived from a neural network analyzing 180+ student variables, predicts potential dropouts with greater efficiency and accuracy than Dropout Detective, as measured by several performance metrics.

A key feature of the MeasurEDU platform is the unique Student Risk Score available for active students in the Student Journey module. Calculating this risk score begins with a Random Forest Classifier model, which analyzes 180+ pieces of unique student data from CampusNexus and Canvas source tables and calculates the top 50 most impactful variables. Then, a Random Neural Network is employed, using a Long Short Term Memory strategy to return a continuous value between 0-100 that represents a student's risk of dropping.

The Student Journey module will display all active students and their risk scores, as well as allow users to identify a cut-off score that they would like to use to identify selected at-risk students. For example, a user can adjust their own Risk Profile Setting so that only students with a Risk Score of 10 and above will be highlighted. In this way, the lists can be highly customized to capture curated pools of predicted at-risk students.

In this analysis, we use various Risk Score values to generate lists of predicted student drops. These predictions were then compared to the actual outcomes for the IHE students in the term starting Oct. 2, 2023. The same comparison and analysis were done with a list of predicted student drops generated by the Dropout Detective tool currently in use at the online college.

The data collected indicated that 479 students at the online institution either dropped in the Oct. 2 term or did not maintain an active status as of the census date of the following term. We have termed this group "drops" and all others "not-drops." Dropout Detective predicted 1496 drops, and out of those predictions, 105 actually dropped. The MeasurEDU predictions were analyzed at various Risk Scores in order to determine the optimal and most accurate Risk Score to use as the predictor value. Notably, when considering students with a Risk Score of 10+, the model excelled by outperforming Dropout Detective, predicting 416 drops, of which 61 were actual drops. At this Risk Score threshold, the model demonstrates superior accuracy, precision, and F1-score than the Dropout Detective predictions.

By refining the Risk Score threshold to 4+, the model's performance is optimized for all metrics, including accuracy, recall, precision, and F1-Score. Detailed outcomes, including confusion matrices and performance metric charts, are presented below.
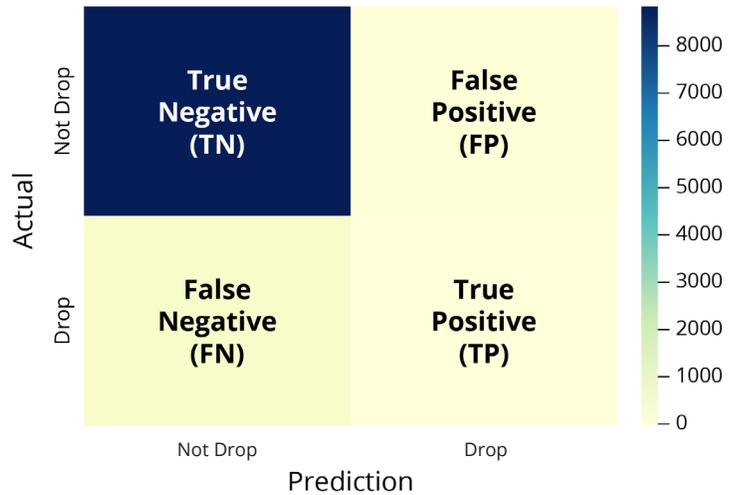
Confusion Matrix

The performance of the models is evaluated by constructing a confusion matrix for each set of predictions. Relevant performance metrics are calculated from the matrix values.

The confusion matrix provides an overview of all predictions made by the model and relates them to the actual outcomes in the dataset.

In this case, a predicted drop is a Positive prediction, and a not-drop would be a Negative prediction.

Confusion Matrix - Risk Score _____

Performance Metrics

Various statistical calculations can be made with the values collected in the confusion matrix. This analysis used the following four metrics to evaluate the performance of the models:

$$Accuracy \ = \ \frac{TN + TP}{TN + FP + TP + FN}$$

The total number of correctly classified instances over the total number of observed instances.

$$Precision \ = \ \frac{TP}{TP + FP}$$

Precision improves and nears 1 when the number of false positives is low. Higher precision indicates a model that makes fewer incorrect positive predictions.

$$Recall \ = \ \frac{TP}{TP + FN}$$

Recall improves and nears 1 when the number of false negatives is low. Higher recall indicates a model that makes fewer incorrect negative predictions.
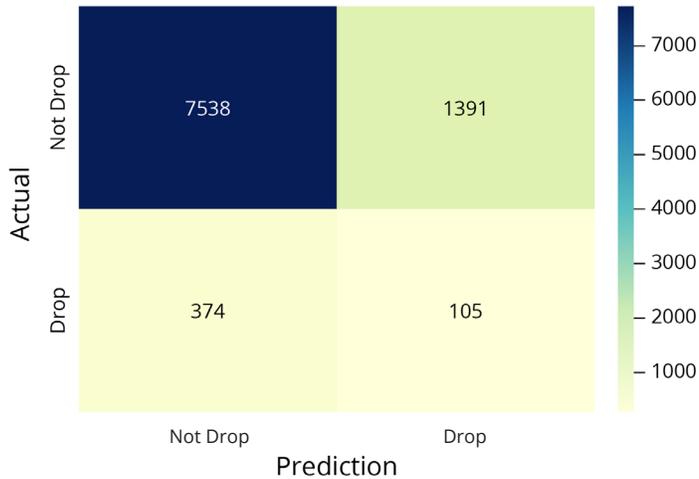
$$F1 \ Score \ = \ \frac{Precision \times Recall}{Precision + Recall}$$

On their own, accuracy, precision, and recall provide limited context for a model's performance. The F1 Score is the harmonic mean of precision and recall and offers a holistic metric to evaluate a model's performance, especially when comparing models to each other.

Dropout Detective Predictions

The Dropout Detective list contained 1496 student names, all of which were counted as positive predictions; put differently, these were all predicted drops. Of those positive predictions, 105 were True Positives, while 1391 were False Positives. There were 374 False Negatives — or students that the model predicted would persist — but, in reality, ended up dropping.

Confusion Matrix - Dropout Detective

	Dropout Detective
Accuracy	81.24%
Precision	7.02%
Recall	21.92%
F1-Score	10.63%
At-Risk Identified	1496

MeasurEDU Predictions

Predictions from MeasurEDU were collected at various Risk Score thresholds, and the results are detailed below. Where Risk Scores of 10 or greater were counted as Positive predictions, the model had higher overall performance metrics than the Dropout Detective predictions, with an F1-Score of 13.63%.

Confusion Matrix - Risk Score 10+

	MeasurEDU Risk Score 10+
Accuracy	91.78%
Precision	14.66%
Recall	12.73%
F1-Score	13.63%
At-Risk Identified	416

Performance metrics for at-risk groups based on other Risk Scores are seen here:

| | Risk Score | | | | | | | | | Dropout Detective |
	90+	80+	70+	60+	50+	40+	30+	20+	10+	
Accuracy	93.85%	93.68%	93.63%	93.56%	93.44%	93.29%	93.04%	92.81%	91.78%	81.24%
Precision	0.00%	0.00%	0.00%	0.78%	0.71%	1.9%	3.68%	10.76%	14.66%	7.02%
Recall	0.00%	0.00%	0.00%	0.21%	0.21%	0.63%	1.46%	5.64%	12.73%	21.92%
F1-Score	0.00%	0.00%	0.00%	0.33%	0.32%	0.94%	2.09%	7.40%	13.63%	10.63%
At-Risk Identified	100	116	120	129	140	158	190	251	416	1496

Additional tuning of the MeasurEDU predictions can be made, with incremental improvement and eventual decline in the F1-Score. In this way, the peak performance can be found at Risk Scores of 4 and above.

| | Risk Score | | | | | Dropout Detective |
	5+	4+	3+	2+	1+	
Accuracy	89.06%	87.47%	84.78%	78.58%	60.46%	81.24%
Precision	14.83%	13.99%	12.57%	9.71%	7.10%	7.02%
Recall	24.22%	28.39%	33.40%	38.62%	55.95%	21.92%
F1-Score	18.40%	18.75%	18.26%	15.51%	12.59%	10.63%
At-Risk Identified	782	972	1273	1906	3777	1496

Confusion Matrix - Risk Score 4+
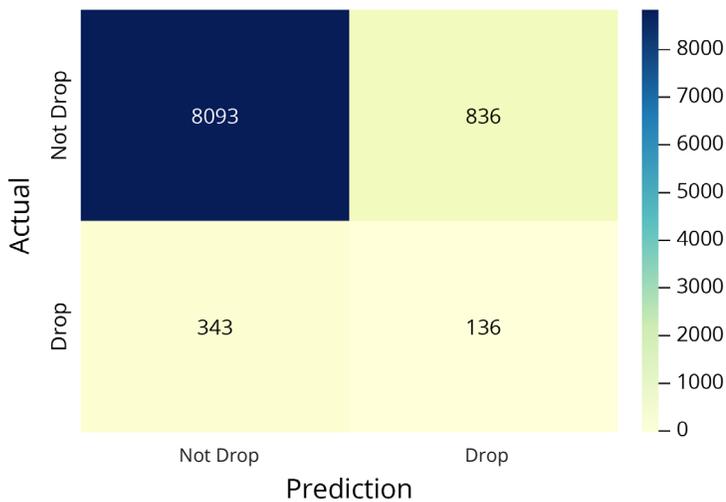
	Risk Score 4+	Dropout Detective
Accuracy	87.47%	81.24%
Precision	13.99%	7.02%
Recall	28.39%	21.92%
F1-Score	18.75%	10.63%
At-Risk Identified	972	1496

When MeasurEDU operates at its optimal performance level, identifying at-risk students with Risk Scores of 4+, it generates a list that encompasses 28.39%, or 136 of the actual 479 student drops. This level of accuracy signifies a notable improvement of nearly 30% compared to Dropout Detective. In simpler terms, MeasurEDU can generate a smaller list with more actual drops than Dropout Detective's larger list, which contains fewer actual drops.

Adjusting for Parity in Risk Pool Size

While Risk Scores below 4 result in lower performance metrics for the MeasurEDU predictions, it is possible to handicap the model and adjust to overpredict a number of at-risk students equal to the number of predictions from Dropout Detective. With Risk Scores 2.55+, the pool of predicted drops is inflated to 1495 students, nearing the 1496 predicted drops from Dropout Detective. Even with this large number of over-predicted drops, the MeasurEDU model maintains better performance metrics than Dropout Detective. This approach could be used to identify a larger pool of students with whom to implement interventions aimed at improving student retention. However, as the reduced accuracy demonstrates, much of this expanded effort would be misdirected at students who are not actually likely to drop out.

	Risk Score 2.55+	Dropout Detective
Accuracy	82.61%	81.24%
Precision	11.30%	7.02%
Recall	35.28%	21.92%
F1-Score	17.12%	10.63%
At-Risk Identified	1495	1496

Additional Analysis Recommendation

For the term in review, which began Oct.2, 2023, 9408 students at the online IHE listed as having an active status were included in the analysis. Most students either continued into the next term or completed their programs. 479 students dropped or did not continue into the following term. Dropout Detective accurately predicted 105 of these drops, and MeasurEDU accurately predicted 136 drops when the Risk Score threshold was 4 and above. This means that the Dropout Detective and MeasurEDU missed 374 and 343 student drops, respectively.

With such a large number of False Negatives surfacing in both platforms, additional investigation was conducted to understand which factors, if any, could be tuned in the model to help improve its ability to distinguish between low-risk drops and low-risk not-drops.

The additional analysis included a Random Forest model, which used normalized data from the Oct. 2, 2023 term, representing the 50 predictive features identified in the MeasurEDU model. The model evaluated the feature importance in distinguishing between students with a 4+ Risk Score who did drop and those with a 4+ Risk Score who did not drop. The Gini index was computed to represent each feature's importance in distinguishing between these two groups. The output of this model surfaced the following features as those that could be weighted differently in future model iterations in order to improve the model's ability to distinguish between low-risk drops and low-risk not drops.

- Cumulative Credits Earned
- Previous Cumulative GPA
- Previous Term Credits Earned
- Previous Term GPA
- Total Quizzes Open for the Student
- Cumulative GPA
- Cumulative Quartile Click amount
- Previous Term GPA points
- Previous Term Credits Attempted
- Total Required Credits

Further investigation into tuning the model's predictive ability could result in improved performance, namely an increase in predicted real drops and a reduction in falsely predicted non-drops.

wgu labs