# **Retention and Student Success** Beyond Retention - An Integrated Student Success Model

#### Charge:

Engage students in assessment process

#### Timeline:

Academic year - Fall 2023 through Spring 2024.

#### Role:

Support creation of reports from early alert data. I practices to analyze large data sets.

#### Goals

Enhance the assessment and use of the early alert sy and related outcomes

- Better understand the relation between the s flags and grade outcome
- 2. Qualitative analysis of notes data from flags and

#### Questions

- How will notes entered by faculty and advisors support improvement of retention and student success?
- Can notes data be used as a leading indicator of s success?

		S	en	time	ent	t da
	Attendance Concern - Grade at Risk		F			
	Attendance Concern - Informational			I		
	I Need An Advisor's Help					
	I Need Help In A Course					
	In Danger of Failing		I			
	Low Participation - Grade at Risk	•				
	Low Participation - Informational			ŀ		
FLAG	Low Quiz/Test Scores		F			
	Marked Absent					
	Missed First Class Session					
	Missed Four Live Class Sessions					
	Missed Three Live Class Sessions					
	Missed Two Live Class Sessions					
	Never Attended				I	
	Stopped Attending			•	$\vdash$	
	Estimated midterm grade is passing					
	Excellent leadership skills					
(UDO	Great critical thinking skills					
	Great participation					
	Keep Up the Good Work					
	Outstanding Academic Performance					
	Showing Improvement				•	•
		-1.0	-C	).8	-0	.6

Prabin Raj Shrestha '24 Hope Smalling, Ed.S.

### **Assessment of Future Practice**

<ul> <li>Tools used:</li> <li>Python: data processing, data modeling</li> <li>R: data modeling</li> <li>Tableau: visualization, data explore</li> <li>Models: distilbert-base-multilingual sentiments-student, association rule</li> <li>New processes</li> <li>Utilize extensive note data from flwithin a machine learning opera</li> <li>Capture more diverse scenarios an more comprehensive insights.</li> <li>Train models effectively to yield his accuracy.</li> <li>Use down sampling techniques to be distribution.</li> </ul>	Use best	Scope and Tools Timeline: late Fall 2023 to Spring 2024 Scope of data • Early alert data (flags/kudos) [Fall • Grades and enrollment data [Fall '2
d kudosNew processes• Utilize extensive note data from fl within a machine learning opera• Capture more diverse scenarios an more comprehensive insights.• Train models effectively to yield his accuracy.• Use down sampling techniques to k distribution.	ystem (EAS) selected	<ul> <li>Tools used:</li> <li>Python: data processing, data mod</li> <li>R: data modeling</li> <li>Tableau: visualization, data exploration</li> <li>Models: distilbert-base-multilinguation</li> <li>sentiments-student, association rule</li> </ul>
<ul> <li>Train models effectively to yield hist</li> <li>accuracy.</li> <li>Use down sampling techniques to k</li> <li>distribution.</li> </ul>	d kudos Ipport	<ul> <li>New processes</li> <li>Utilize extensive note data from fla within a machine learning operation</li> <li>Capture more diverse scenarios and more comprehensive insights.</li> </ul>
	student	<ul> <li>Train models effectively to yield his accuracy.</li> <li>Use down sampling techniques to b distribution.</li> </ul>

#### ata from flag/kudo comments



#### '22 to Spring '23] '22 to Spring '23]

#### leling

- ation I-casede mining
- ags and kudos tion. d patterns for

#### igher predictive

#### balance class

#### Key text analytic metrics

## • Sentiments

#### • Emotions

#### • Lexical Diversity

Lexical diversity is a measure of how many different words appear in a text.

- Lexical Density
  - of words.

## Sentence specificity

The level of detail or precision contained within a sentence.



### Model results • Random forest exhibits better overall

- mance, making it the most suitable mo this dataset.
- Decision tree show similar performance
- Support vector machine learning (SVM ly outperforming Random Forest and De Tree. However, as this data set contains noise, such as overlapping target classe will not perform as well when scaled.

Kal Srinivas, Ph.D.; Radell Roberts; ShawnMarie Parry; Samantha Trumble; Hopeton Smalling, Ed. S.

Taylor, L. (2012, July). Beyond Retention: Using Targeted Analytics to Improve Student Success. https://www.academicimpressions.com/how-early-alert-and-student-success-initiatives-fail/



The overall tone conveyed in a text, speech, or piece of writing.

The overall emotion expressed in a text, speech, or piece of writing

The proportion of lexical words in a text compared to the total number

				Cross Validation (3-Folds)			
		Class Imbalace	Accuracy	Precision (Weighted)	Recall (Weighted)	FI-Score (Weighted)	
perfor-		None	0.59	0.67	0.59	0.56	
odel for	Decision Tree	Up Sampleing	0.63	0.67	0.63	0.60	
		Down Sampling	0.63	0.66	0.63	0.60	
e levels.		SMOTE	0.60	0.68	0.60	0.57	
M) slight-	Random	Up Sampleing	0.63	0.66	0.63	0.61	
ecision	Forest	Down Sampling	0.63	0.66	0.63	0.61	
s a lot of	SVM	Down Sampling	0.64	0.67	0.64	0.61	
		Up Sampleing	0.68	0.69	0.68	0.66	
	KNN	Down Sampling	0.62	0.65	0.62	0.57	
	L						