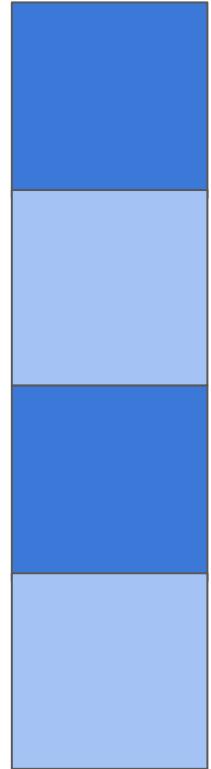


# Keep-Me-Engaged

Customized Recommendations to Improve e-Learner Engagement

*Lindsay Warrenburg*



# New Tool for Premium Subscribers: *Customized Recommendations*

zoom

GoToTraining

Cisco  
webex

Segment



stripe

zapier

salesforce

Adobe Connect

INSIGHT

# New Tool for Premium Subscribers: *Customized Recommendations*

7,000+  
Companies



200,000+  
Courses



Amazon Redshift



PostgreSQL

## Access and Download Data

*9.9B data points*



Amazon Redshift

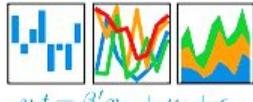


PostgreSQL

Access and Download Data  
*9.9B data points*



python



pandas

$$y_i t = \beta^T x_{it} + \mu_i + \epsilon_{it}$$



Feature Engineering  
Random Forest Regression



Amazon Redshift

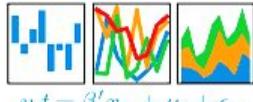


PostgreSQL

Access and Download Data  
*9.9B data points*



python



$y_i t = \beta^T x_{it} + \mu_i + \epsilon_{it}$   
pandas



Feature Engineering  
Random Forest Regression



Streamlit



HEROKU

Web App Creation

THOUGHT

INDUSTRIES®



## Learner Engagement

INSIGHT



## Learner Engagement



### Admin. Features:

Price of Course

Length of Course



## Learner Engagement



### Admin. Features:

Price of Course

Length of Course



### Learner Features:

Time Spent on Course

Average Quiz Grade



## Learner Engagement



### Admin. Features:

Price of Course  
Length of Course



### Instructor Features:

% of Non-Graded Quizzes  
Number of Collaborations

### Learner Features:

Time Spent on Course  
Average Quiz Grade



## Learner Engagement



### Admin. Features:

Price of Course  
Length of Course



### Learner Features:

Time Spent on Course  
Average Quiz Grade



### Instructor Features:

% of Non-Graded Quizzes  
Number of Collaborations



### Website Features:

% Interactive Content  
% Videos



## Learner Engagement



### Admin. Features:

Price of Course  
Length of Course



### Learner Features:

Time Spent on Course  
Average Quiz Grade



### Instructor Features:

% of Non-Graded Quizzes  
Number of Collaborations

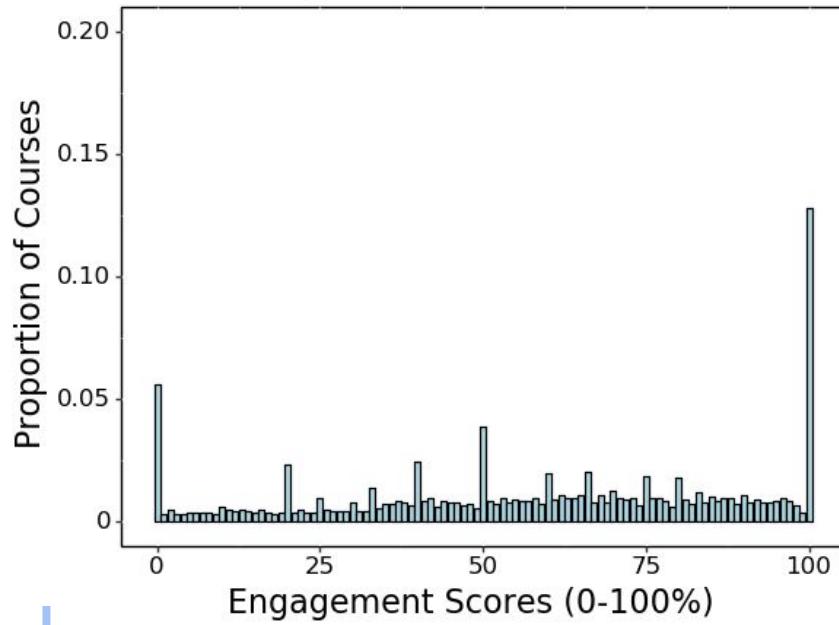


### Website Features:

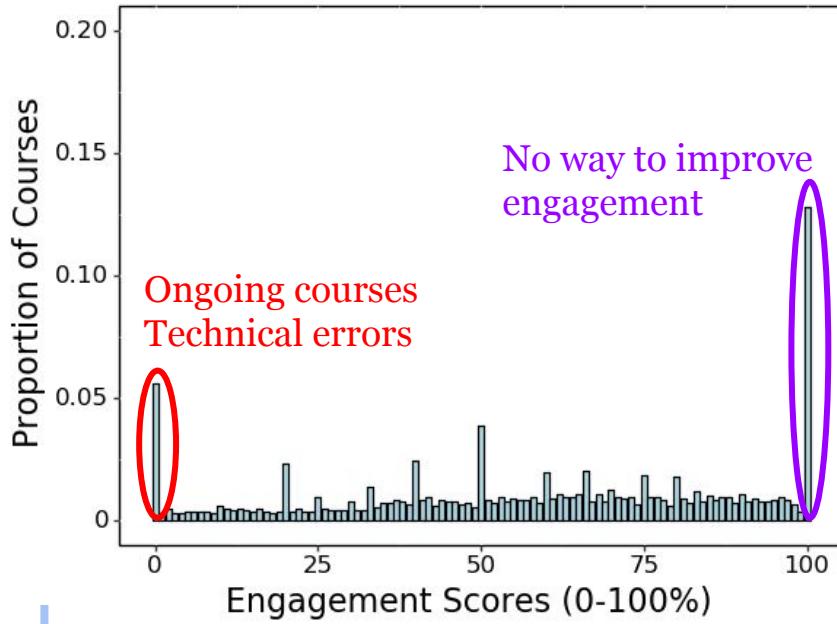
% Interactive Content  
% Videos

# The distribution of Engagement Scores is highly irregular

Original Distribution



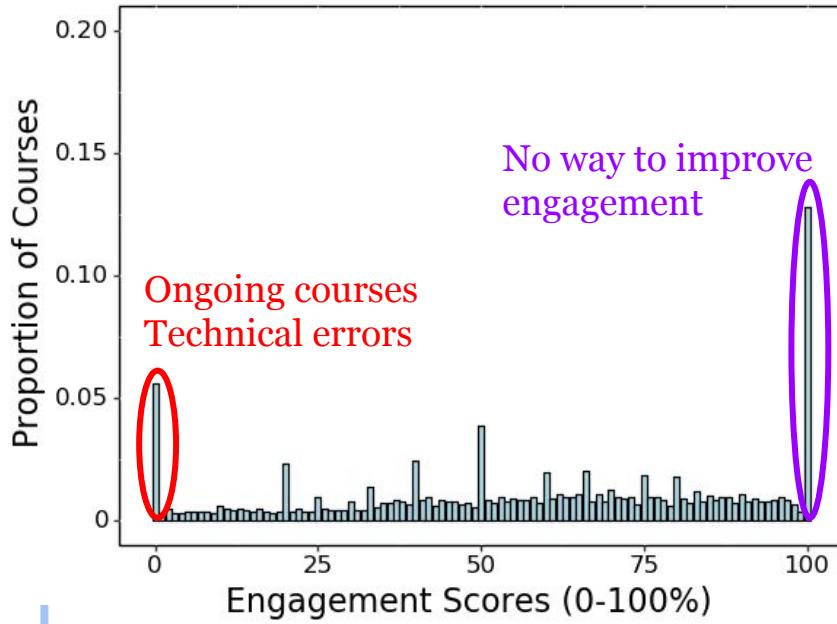
## Original Distribution



## Scope of Keep-Me-Engaged

- Help course developers with Engagement Scores of 1-99% understand current engagement scores

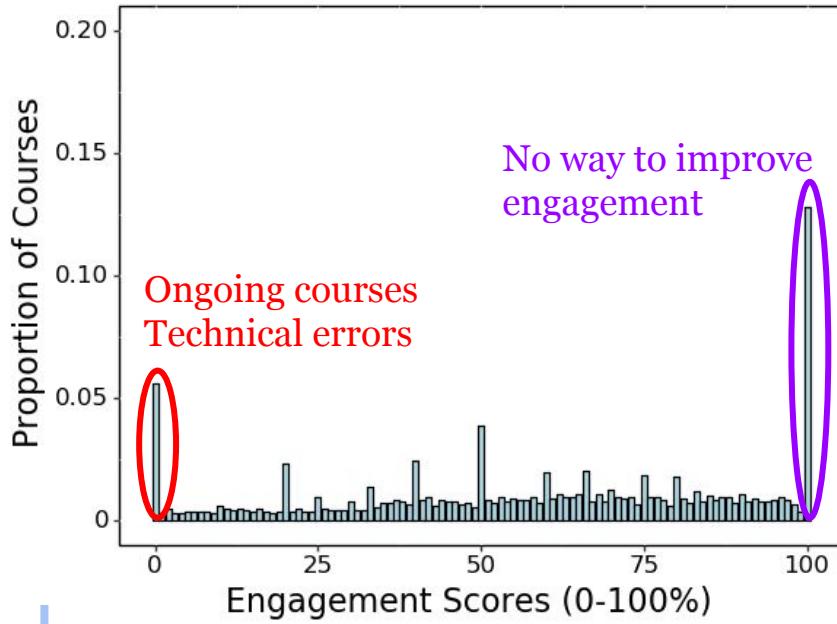
### Original Distribution



## Scope of Keep-Me-Engaged

- Help course developers with Engagement Scores of 1-99% understand current engagement scores
- Learn how to improve engagement scores for future iterations of the course (e.g., next semester)

## Original Distribution

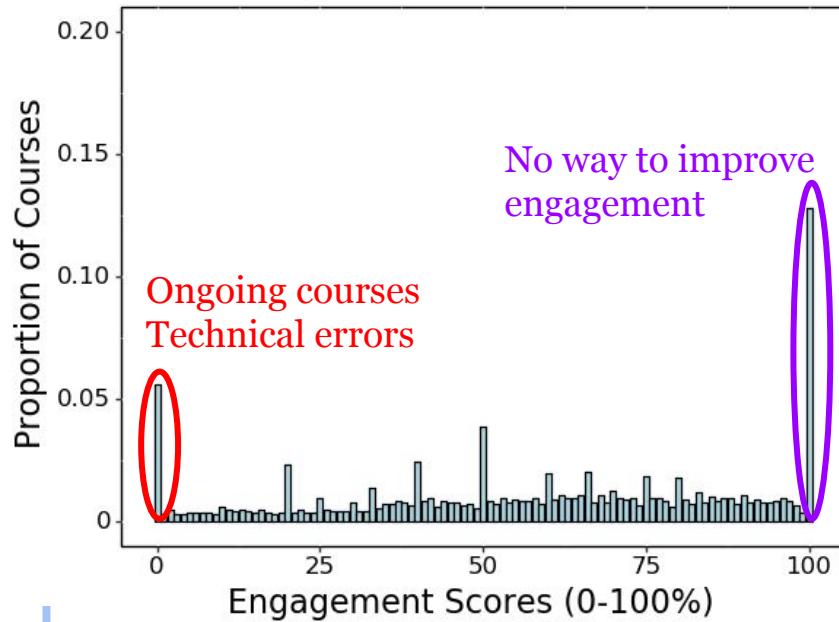


## Scope of Keep-Me-Engaged

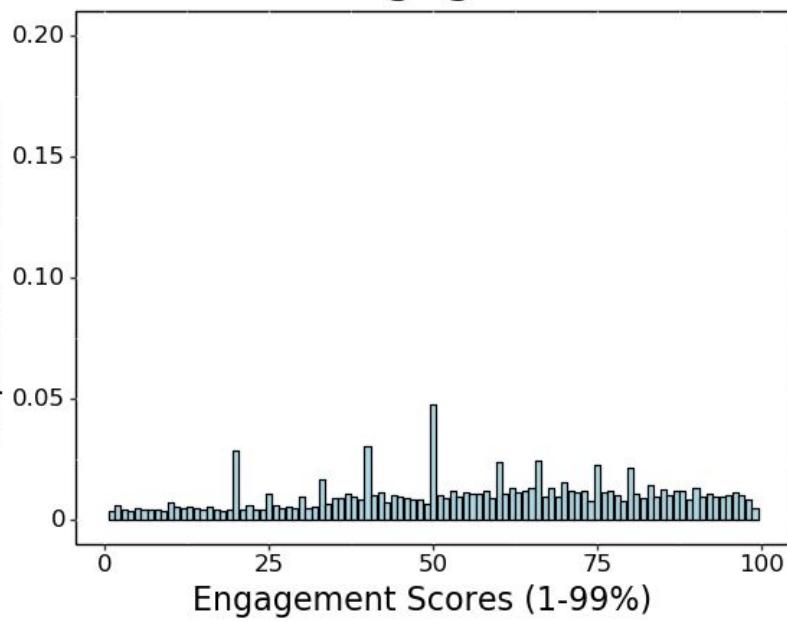
- Help course developers with Engagement Scores of 1-99% understand current engagement scores
- Learn how to improve engagement scores for future iterations of the course (e.g., next semester)
- *Not aimed at predicting engagement scores of a brand new course*

# The distribution of Engagement Scores is highly irregular

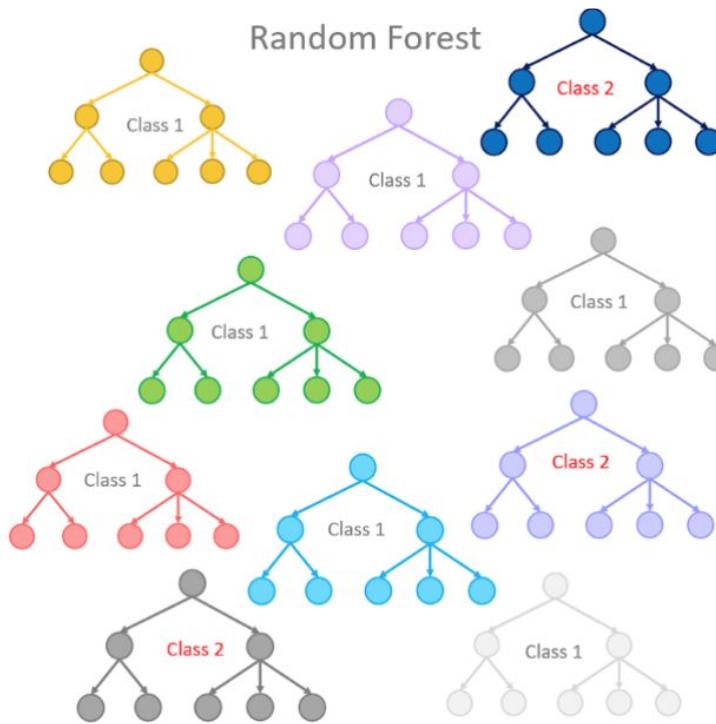
Original Distribution



Continuous Engagement Scores



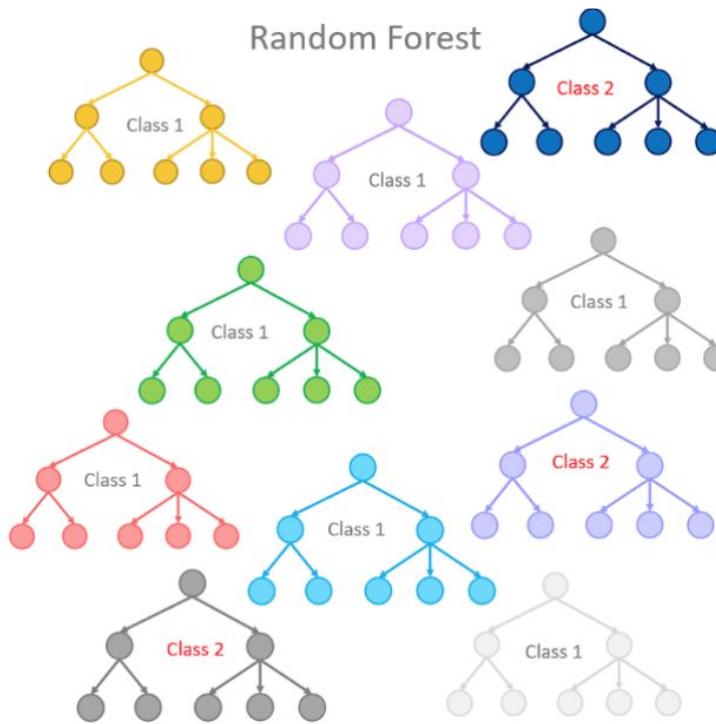
# Predicting User Engagement Scores



**Random Forest Regression**  
 $R^2 = 0.62$

*Train:Test split = 80:20*

# Predicting User Engagement Scores



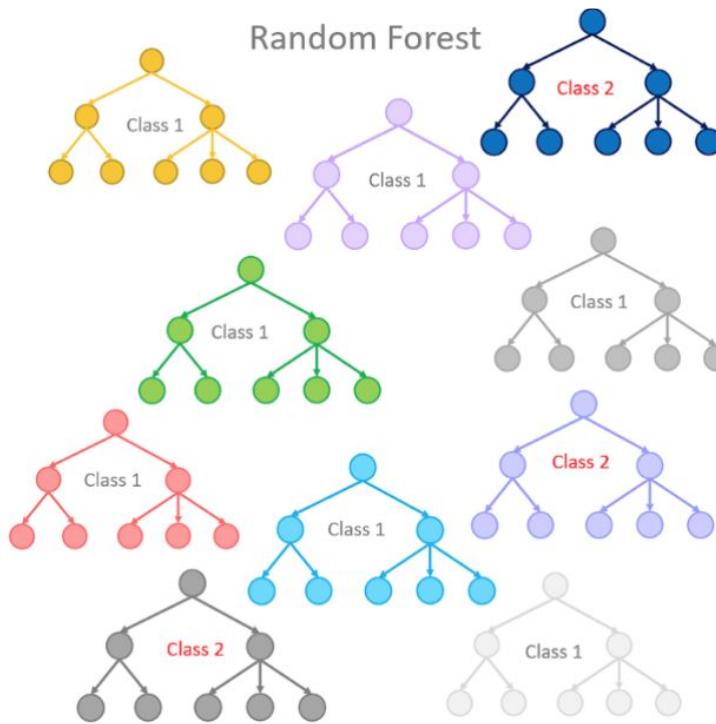
## Random Forest Regression

$R^2 = 0.62$

## Linear Regression (log)

$R^2 = 0.16$

# Predicting User Engagement Scores



## Random Forest Regression

$R^2 = 0.62$

## Linear Regression (log)

$R^2 = 0.16$

## Ridge Regression (log)

$R^2 = 0.16$

## Interactive *Keep-Me-Engaged* Tool

<https://keep-me-engaged.herokuapp.com/>

# Thought Industries will use the Keep-Me-Engaged tool as part of their Premium Package

Customized recommendations for

30,000+ courses

550+ companies

# Lindsay Warrenburg



Market Research • Data Visualization • B2B Data

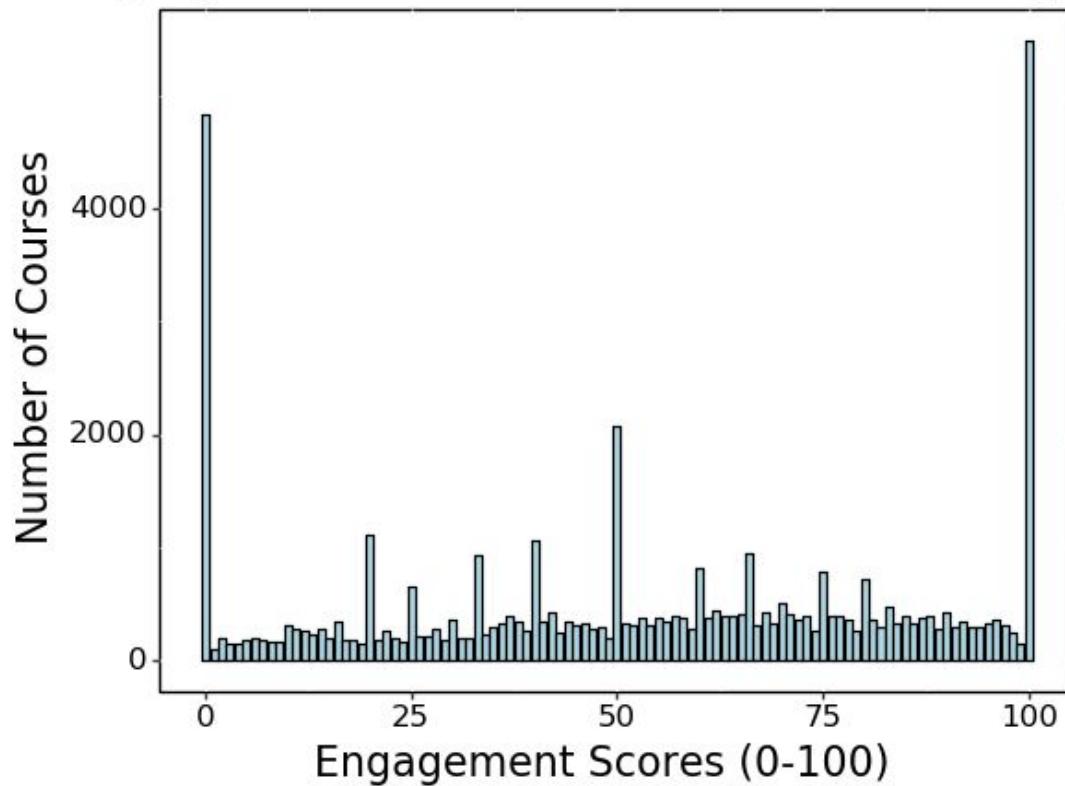


**Penn**  
UNIVERSITY *of* PENNSYLVANIA



Music • Cognition & Perception • Emotion

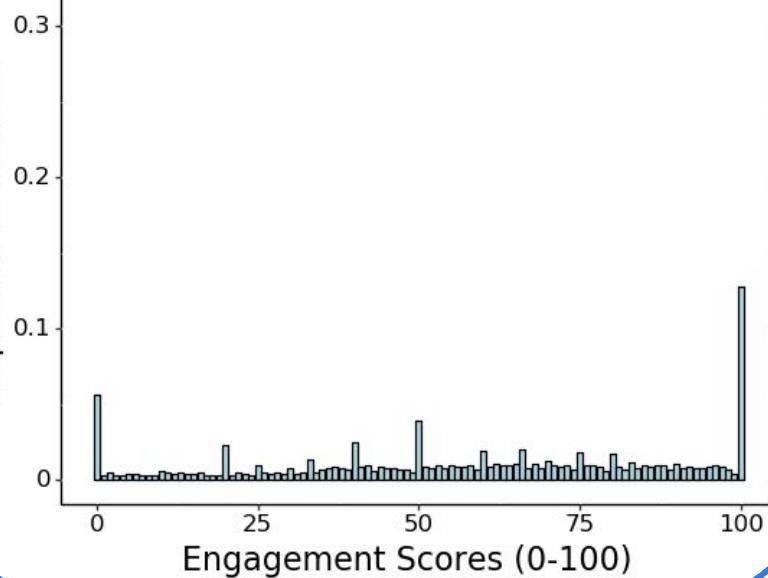
# Engagement Across All Course Types



INSIGHT

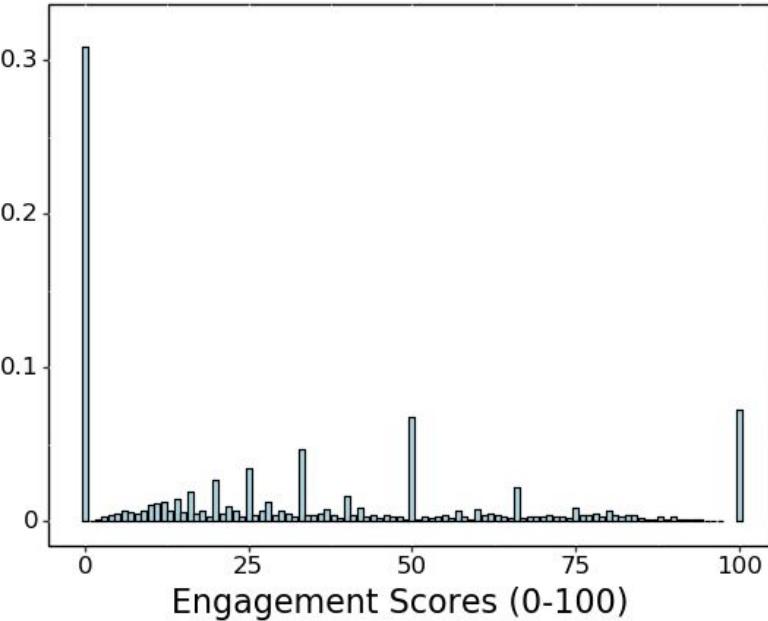
## Online Courses

Proportion of Courses



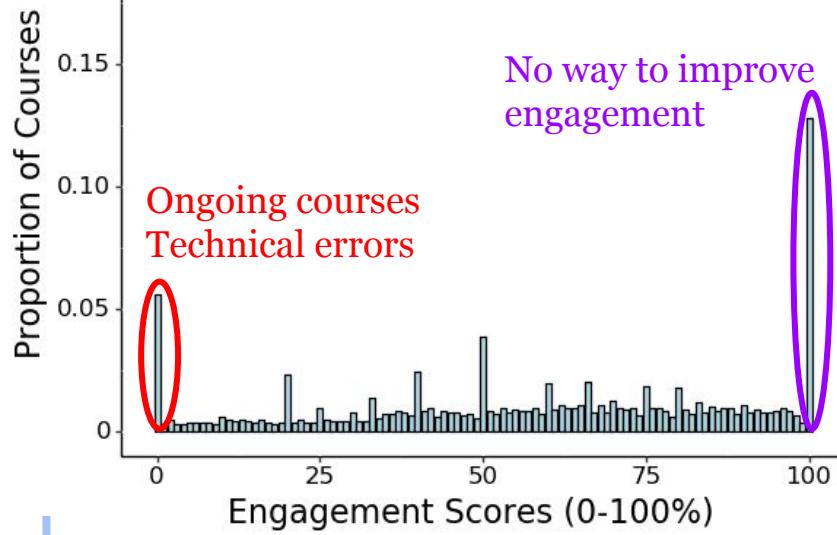
## Articles & Videos

Proportion of Courses

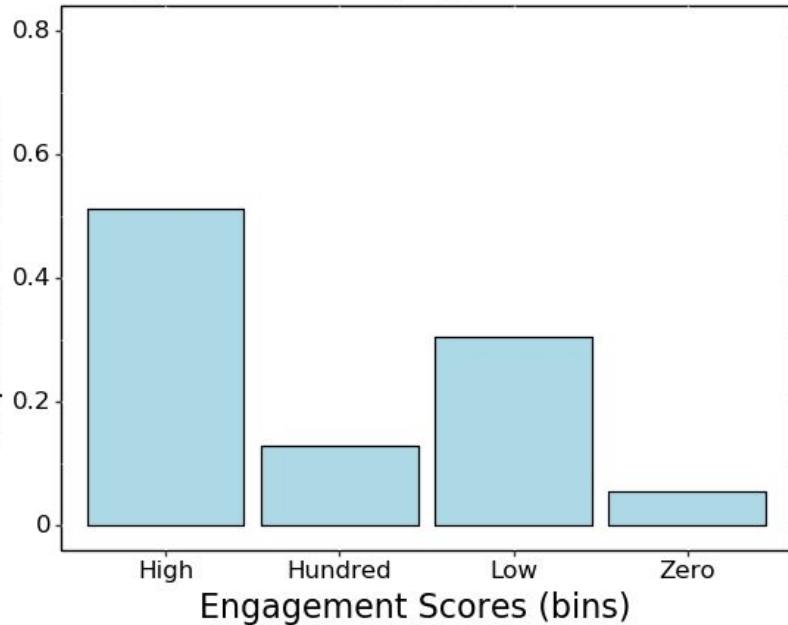


INSIGHT

Original Distribution



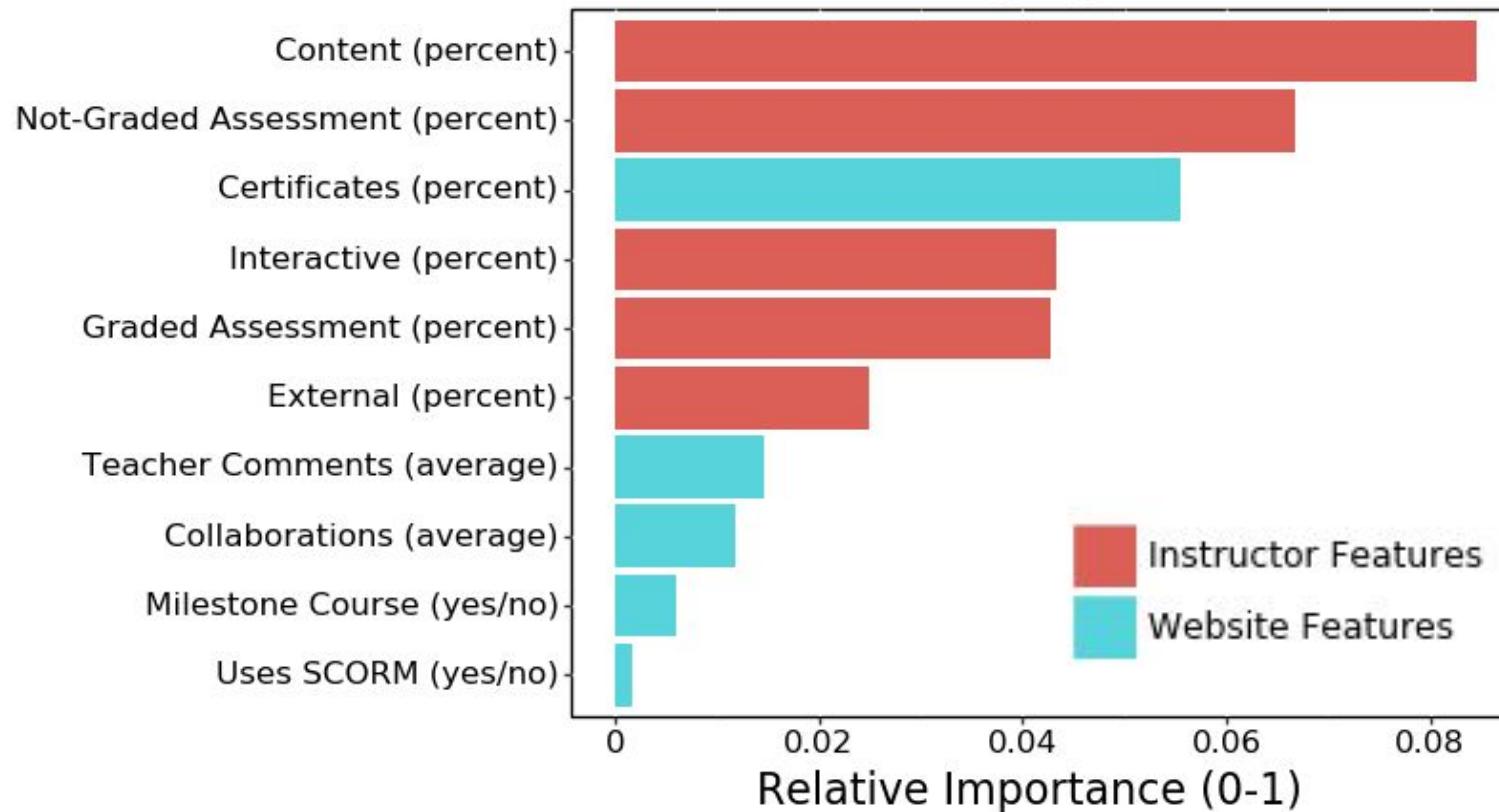
Categorical Engagement Scores



# Missing Values

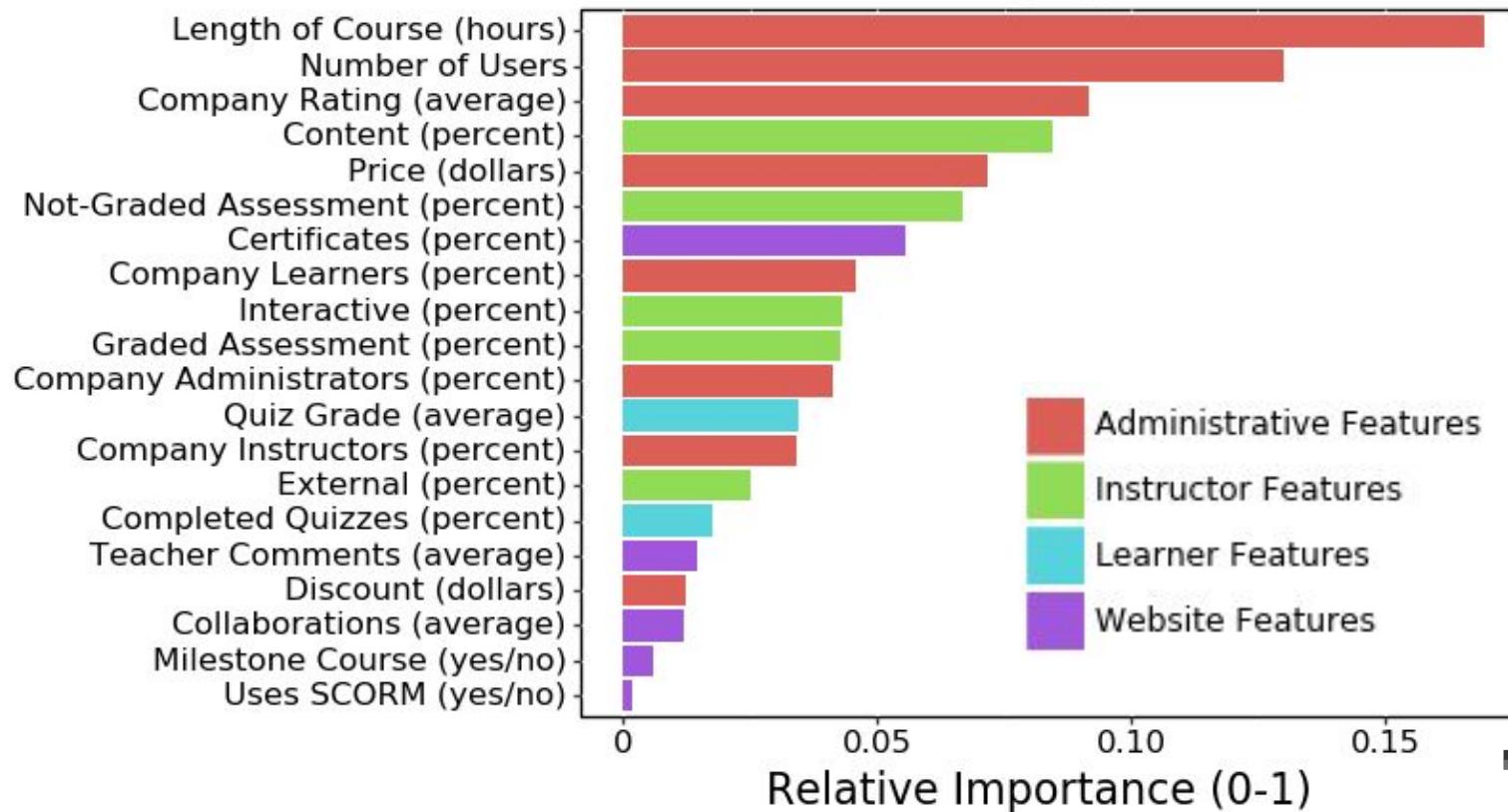
- **Delete some courses with missing values**
  - *Example:* Courses that are missing our target (Engagement Scores)
- **Replace missing values with 0**
  - *Example:* No value for Number of Collaborations = no collaborations
- **Combine multiple features into one summary feature when only half the courses use SCORM at all**
  - *Example:* Average SCORM score, % SCORMs completed, % API SCORM, % Shareable SCORM ⇒ “Uses SCORM yes/no”
- **Fill in missing values with the median value**
  - *Example:* Average quiz grade

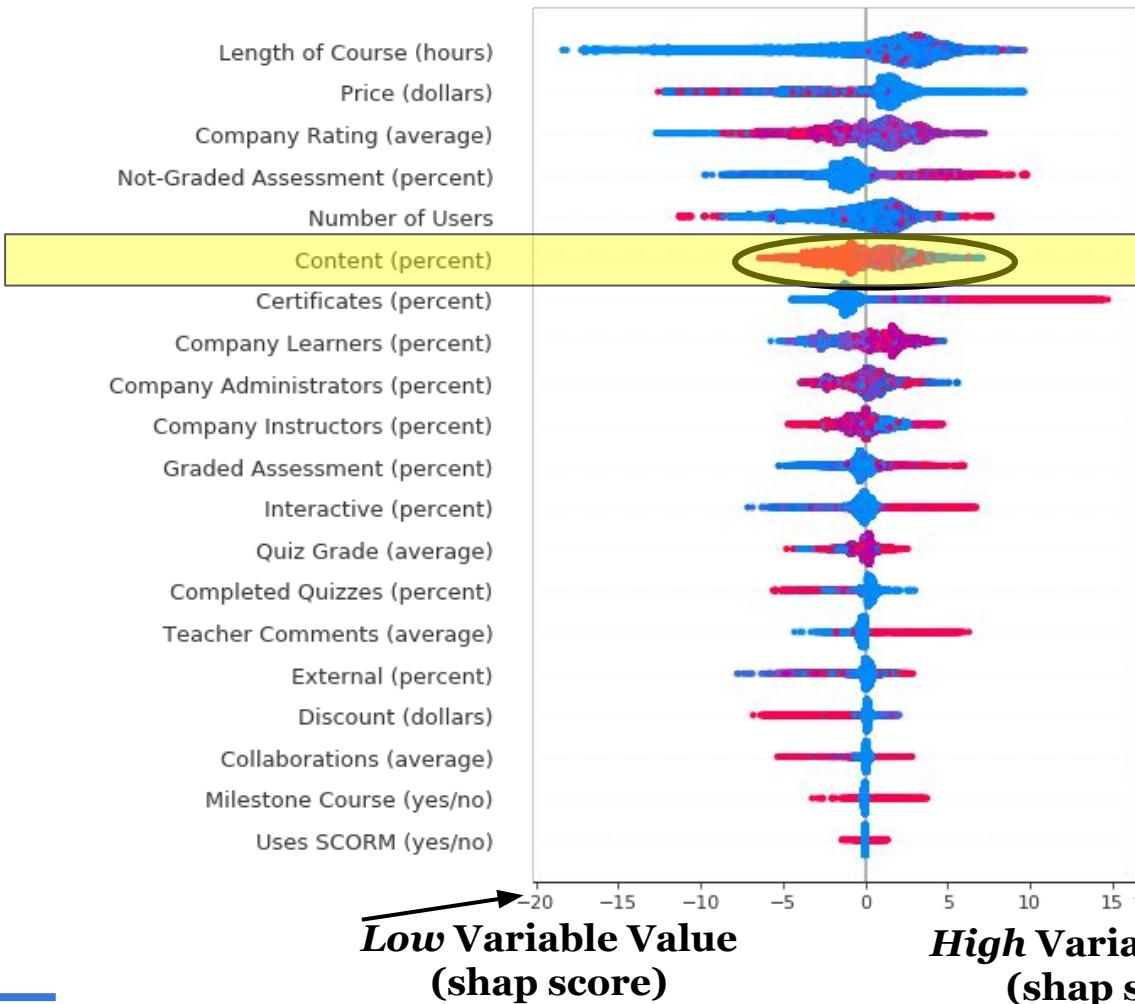
# Feature Importance for User Engagement



INSIGHT

# Feature Importance for User Engagement





**High  
Engagement  
Scores**

Decreasing content results in *higher* engagement scores

**Low  
Engagement  
Scores**