Automated or Algorithmic - replace all "AI" with algorithm(ic) or automated

Project Proposal

An AI-based Teaching Practices and Classroom Activities Tool to Improve Education

Students:

Jinghan Zhang, Joshua Blazek, Leo Garcia, and Naiqi Yao

Sponsor:

Christof Teuscher

teuscher.:Lab
teuscher-lab.com

Professor:

John Lipor

ECE Capstone - Winter 2021

Version 0.8 – 1/21/2021

Table of Contents

Overview

is the primary goal improving education research? or is it helping teachers improve their teaching?

really?

Research done on teaching practices is currently a time consuming process involving significant time investment with human observers each of whom have their own subjective methods of observations. "An AI-based Teaching Practices and Classroom Activities Tool to Improve Education" is an automated tool that seeks to overcome some of these limitations.

example of unnecessary caps

Our project will use AI techniques to analyze student and teaching behaviors captured in recorded videos. One recording will capture the Professor and their teaching materials. The other recording will capture the student audience.

? no need for ellipsis here

The analysis of the Professor's recording will include teaching pace, materials pace (e.g. slides…), text word count, and spoken word count. The student analysis will include how many questions were asked, on which slides or materials these questions were asked, and the general attentiveness of the student audience. Classroom activities should also be classified as either lecture time or interactive time and may be disambiguated further into either lecture, discussions, individual work, exercises, or quiz time.

be sure these align with your requirements after updating

This analysis will be visualized in a form where a non-technical user can glean informative feedback related to their changes in teaching behavior and the commensurate effects of those changes on the student audience.

We will deliver a functional prototype that accepts two recordings of full-length class sessions and analyze these sessions. We also will deliver relevant documentation. We expect to encounter COVID related collaboration issues and we are concerned about accidental feature creeping technological overreach.

hyphenate compound adjectives when they come before the noun they describe

Product Design Specification

Concept of Operations

This project will allow a user, such as a professor, to record their lecture and their student's response to different moments within that lecture. These recordings will be uploaded and processed by a remote server. The user will receive a report back from this server that describes to them (description detailed in "Requirements") how their students responded to different moments throughout that lecture. The user can then use this feedback to modify their future teaching behaviors to maximize student interaction or student attentiveness in accordance with their goals and repeat the analysis.

program?

not necessarily - could be an app they run locally

Stakeholders

Industry Sponsor:
- Professor Christof Teuscher

Faculty Advisor:
- Professor John Lipor

Project Developer:
- Team 9:
    - Jinghan Zhang
    - Joshua Blazek
    - Leo Garcia
    - Naiqi Yao

The User of the Project:
- University
- Professor
- Teacher's Assistant

Requirements

Our project must have the ability to receive and analyze one full-length class lecture recording from two perspectives. These perspectives must include the Professor with their presented materials and the student audience. These recordings will be processed on a database that ? must be able to analyze those two recordings and produce a visualization, containing information outlined below, in a form where a non-technical user can glean informative feedback. This feedback will describe the user's teaching behavior across time and the commensurate effects of those changes on the student audience.

("slides" here also refers to non-slide presented materials such as an overhead projector)

The analysis of the Professor's recording **must** include the following:
- Total number of slides
- Average text per slide
- Slide number across time ? number of slides
- ? Words spoken per slide

The analysis of the student audience recording **must** include the following:
- Total number of questions asked
- Questions across each slide by whom?
- Student attentiveness across time
    - Attentive = student gaze forward facing or recently forward facing.
- Lecture time over interactive time across time

p ok

The analysis of the student audience recording **should** include:

- Which specific students ask questions
- Further breakdown of “interactive time” into the subcategories:
    - Individual work
    - Discussions
    - Exercises
    - Quiz/Exam

The analysis of the student audience recording **may** include:

- Rate of student audience note taking across time
- Student questioning across attentiveness
- Ratio of the purpose for downward facing gaze of the student audience across the following categories:
    - Note taking
    - Book reading
    - Phone checking
    - Laptop/tablet checking

Specifications

Recorded Video Codec:

- AV1, VP9, H.263 (MPEG-2), H.264 (MPEG-4), or H.265 (HVAC)
- 720p, 1080p, 1440p, or 2160p resolution

Recorded Audio Codec:

- FLAC, Opus, MP3, PCM, Vorbis, or AAC and its derivatives
- 56 Kbps to 2 Mbps bitrate

Deliverables

- Project Proposal
- Weekly Progress Reports
- Final ~~R~~r Report
- Final ~~P~~p Presentation
- ECE Capstone Poster Session poster
- Github Repository Including:
    - Code base
    - Revision history
    - All documentation
- Fully functional prototype that meets, at minimum, the “must” requirements
- End User manual
- Detailed design documentation which will include the “what”, “how”, and “why” of decisions.
- Discussion on servicing, upgrade paths, and future plans
- Open ~~S~~s Source license (MIT)

look for unnecessary caps throughout

Initial Product Designs

split

probably not in this capstone

We propose
~~We're proposing~~ to use two video/audio recordings from a class lecture to analyze student and teacher behavior. These videos will be uploaded to an online database through a web portal. ~~Once uploaded~~, the server will demux the videos into separated audio and video files for classification. These files will be fed through our models.

analysis?

careful about use of "models" throughout - may want to replace with algorithm

There will be four models in use (see below for potential model candidates), an audio and video model for classifying the Professor facing camera and an audio and video model for classifying the student audience facing camera. The classified data that comes out of these 4 models will be passed into analyzer programs that will parse and combine the data to create the datasets described in the "Requirements" section of this document. These datasets will be visualized into a graph format and emailed to the user.

four - write out numbers below 10

Software Options:

- Python (code base), TensorFlow (feature tracking), and OpenCV (feature isolation)

Potential Video Classifier Options:

- [OpenFace] Constrained Local Model (CLM) for facial feature isolation and Discriminative Response Map Fitting (DRMF) for facial landmark detection (3).
- [Eye-Tracking] Random Forest framework with bootstrapped tree classifiers (4).
- [DeepGaze] Convolution Neural Network (CNN) for feature isolation and Haar-Cascade Classification for eye tracking (5).

Potential Audio Classifier Options:

- [GMM w/PLDA and i-vectors] Gaussian Mixture Models (GMM) with factor analysis for i-vector and Probabilistic Linear Discriminant Analysis (PLDA) to differentiate speakers.
- [DNN and x-vectors] Deep Neural Networks (DNN) embeddings and clustering using x-vectors to differentiate speakers.
- [DART] Decibel Analysis for Research in Teaching using an ensemble of binary decision trees (1).

Potential Student Classification Difficulties:

- If a student in the second row is looking down at their table to take notes, is the student obscured by the row in front of them going to reduce the accuracy of classification?
- If there is no stadium seating how much will this impact our classification accuracy?
- Is there a distinct enough visual difference between a student looking at their phone, their laptop, taking notes, or reading a textbook to accurately classify those categories?
- How much constraint will we place on the placement of the video cameras and how much will camera placement change classification accuracy?
- If a student gets up and moves or leaves what happens to their classifications?

Potential Professor Classification Difficulties:

- If a Professor has animated slides will that impact classification accuracy?
- If a Professor writes a stream of notes on a projector how will that be classified?
- If there is non-text data (e.g. map, drawing…) on a slide how is that counted?

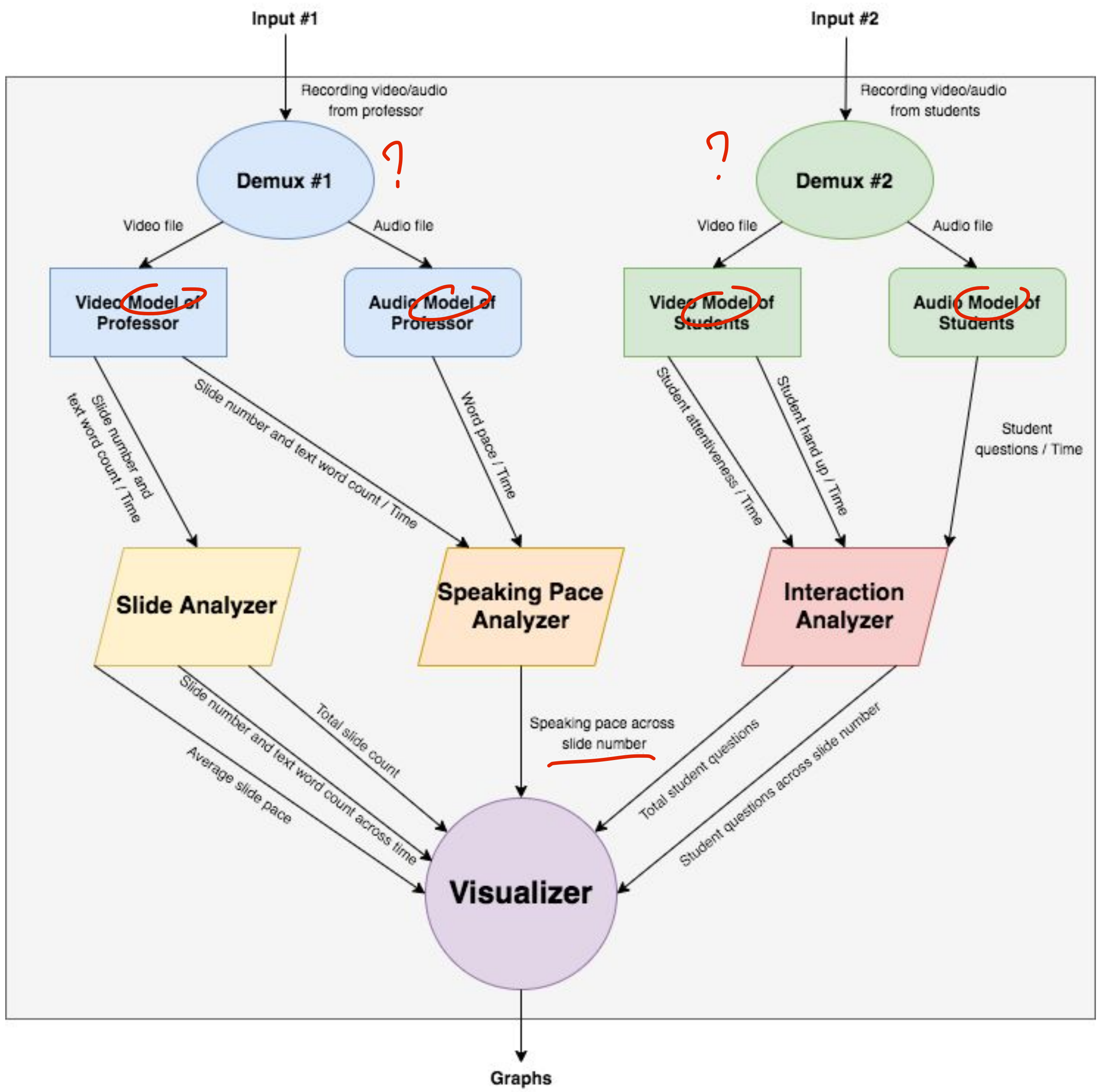
*output figures as eps or pdf if your program allows - Lvl 1 diagram is grainy

Block Diagrams

Level 0 Block Diagram

Repeat

Level 1 Block Diagram



User Experience

- Camera One positioned to capture professor and teaching materials
- Camera Two positioned to capture the student audience
- After the lecture, both recordings are stopped
- Both videos are uploaded via web interface and an email address is provided
- Within 24 hours, the user will receive an email with a graphical representation of the analysis.

Verification Plans

We will process at least 10 hours of class lectures across a minimum of 3 subjects. The uploaded material will be analyzed by our program and the uploaded material will be reviewed by a human auditor using the same criteria as our program. The human auditor will be assumed to have a 100% classification accuracy and this audit will be compared against the AI results to determine AI efficacy and classification accuracy.

Tests (Human and Machine): Should list each mst req.

Professor's recording:
- Total number of slides during the whole class.
- Numbers of slides per minute/how long does professor take on each slide
- Professor's words spoken during each slide
- Number of text words per slide

Student audience recording:
- Percent of students gazing forward across the duration of the recording
- Total number of questions during the whole class
- Number of questions across each slide
- Lecture time and interactive time across time

Project Management Plan

Timeline

?

(notes: milestones. Design so we're working backwards. Realistic. More doable. Fine grained. Block diagram to guide. Include timeline implementation for algorithm choices. What's available? Assign resources. Most important - terms of design cycles. 3 or 4. Research, Implement, test, …. Decision points? 3 basic algorithms. Which wins?.... Audio model 3 cycles -> DART, time analysis, frequency, (research, analysis, decision matrix) -> decision point to choose one of 3 methods… video model)

What eq.?

Name	Duration	Start	Finish
Write Project Proposal	**6 days?**	**01/15/21 08:00 AM**	**01/20/21 05:00 PM**
Concept of operations	3 days?	01/15/21 08:00 AM	01/17/21 05:00 PM
Requirements	3 days?	01/17/21 08:00 AM	01/19/21 05:00 PM
Design Specifications	1 day?	01/20/21 08:00 AM	01/20/21 05:00 PM
Stakeholders	3 days?	01/15/21 08:00 AM	01/17/21 05:00 PM
Timeline with Milestones	**124 days?**	**01/23/21 08:00 AM**	**05/26/21 05:00 PM**
Research phase	9 days?	01/23/21 08:00 AM	01/31/21 05:00 PM
research audio	9 days?	01/23/21 08:00 AM	01/31/21 05:00 PM
GMM and I-vectors	2 days	01/23/21 08:00 AM	01/24/21 05:00 PM
decision matrix	1 day?	01/25/21 08:00 AM	01/25/21 05:00 PM
DNN x-vectors	2 days	01/26/21 08:00 AM	01/27/21 05:00 PM
decision matrix	1 day?	01/28/21 08:00 AM	01/28/21 05:00 PM
DART	2 days	01/29/21 08:00 AM	01/30/21 05:00 PM
decision matrix	1 day?	01/31/21 08:00 AM	01/31/21 05:00 PM
research video	8 days?	01/23/21 08:00 AM	01/30/21 05:00 PM
open face	2 days	01/23/21 08:00 AM	01/24/21 05:00 PM
decision matrix	1 day?	01/25/21 08:00 AM	01/25/21 05:00 PM
eye tracking	2 days	01/26/21 08:00 AM	01/27/21 05:00 PM
decision matrix	1 day?	01/28/21 08:00 AM	01/28/21 05:00 PM
deep gaze	2 days	01/29/21 08:00 AM	01/30/21 05:00 PM
Initial Design	7 days	02/01/21 08:00 AM	02/07/21 05:00 PM
Writing code	35 days	02/08/21 08:00 AM	03/14/21 05:00 PM
first test	7 days?	03/15/21 08:00 AM	03/21/21 05:00 PM
code review	1 day?	03/15/21 08:00 AM	03/15/21 05:00 PM
revision	30 days	03/16/21 08:00 AM	04/14/21 05:00 PM
final test	14 days	04/15/21 08:00 AM	04/28/21 05:00 PM

deployment phase research	14 days	04/29/21 08:00 AM	05/12/21 05:00 PM
deployment	14 days	05/13/21 08:00 AM	05/26/21 05:00 PM
Writing Final Report	20 days	05/01/21 07:00 AM	05/20/21 05:00 PM
Writing/Presenting Other Documents	1 day?	06/01/21 08:00 AM	06/01/21 05:00 PM
Creating Final Poster	1 day?	06/02/21 08:00 AM	06/02/21 05:00 PM

Budget and Resources

- Current Resources
    - Professor Lipor's Compute resources
    - PSU Compute Cluster (COEUS)
    - Cell Phones
    - Webcams
    - Laptops
- Needed Resources
    - None

Team and Development Process

Professor Christof Teuscher: Sponsor. Industry expert. ~~Alan Turing enthusiast.~~
Professor John Lipor: Faculty advisor. Industry expert.
Jinghan Zhang: Team member. Proficient in AuthorCAD, art.
Joshua Blazek: Team member. Proficient in C, Python, OpenCV, TensorFlow.
Leo Garcia: Team member. Proficient in Python, full-stack.
Naiqi Yao: Team member. Proficient in EagleCAD, Draw.IO.

Citations

Format in IEEE style, complete

Author?

1. Proceedings of the National Academy of Sciences Mar 2017, 114 (12) 3085-3090; DOI: 10.1073/pnas.1618693114
2. Wieman, C., Address correspondence to: Carl Wieman (E-mail Address: cwieman@stanford.edu).*Department of Physics and Graduate School of Education, Gilbert, S., Initiative, C., F, A., WK, A., . . . Smith, M. (2017, October 13). The Teaching Practices Inventory: A New Tool for Characterizing College and University Teaching in ?

Mathematics and Science. Retrieved January 22, 2021, from https://www.lifescied.org/doi/full/10.1187/cbe.14-02-0023

3. Baltrušaitis, T., Robinson, P., and Morency, L.-P. (2016). “OpenFace: an open source facial behavior analysis toolkit,” in IEEE Winter Conference on Applications of Computer Vision (Lake Placid, NY) 1–10.
4. Shojaeizadeh, Mina et al. “Detecting task demand via an eye tracking machine learning system.” Decis. Support Syst. 116 (2019): 91-101.
5. DeepGaze: ML Powered Eye Tracking & Gaze Prediction. 27 Apr. 2020, towards.ai/deepgaze-ml-powered-eye-tracking-gaze-prediction/

Revision History

Version 0.5 - 01/10/2021 - Initial document.
Version 0.6 - 01/17/2021 - Added “Requirements” and “Overview.”
Version 0.7 - 01/19/2021 - Added block diagrams.
Version 0.8 - 01/21/2021 - Heavily modified all sections after sponsor/advisor feedback.

Sponsor Sign Off

I, Christof Teuscher, as the sponsor of this project, agree that this proposal describes an acceptable minimum viable product.

Signed ____________________