<u>Project Proposal</u>

An Algorithmic Teaching Practices and Classroom Activities Tool to Improve Education

Students:

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

Sponsor:

Christof Teuscher

teuscher.:Lab
teuscher-lab.com

Professor:

John Lipor

ECE Capstone - Winter 2021

Version 0.9 – 1/29/2021

Table of Contents

Overview

Research done on teaching practices is currently conducted with human observers, each of whom have their own subjective methods of observations and non-uniform methods of record keeping. "An Algorithmic Teaching Practices and Classroom Activities Tool to Improve Education" is an automated tool that seeks to overcome some of these limitations.

Our project will use algorithmic 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.

The analysis of the professor's recording will include teaching pace, materials pace, 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.

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.

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. The user will receive a report back from this program that describes to them how their students responded to different moments throughout that lecture (description detailed in "Requirements"). 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.

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 by a program 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)

Sponsor request	Proposed "must" requirements
How many slides are shown	Total number of slides
	Slides per minute ❏ Slides per minute = Total slides over total lecture time
How much text is on each slide	Average text per slide
At what pace are slides advanced	Slide number across time
How much explanations are provided by the instructor per slide	Words spoken per slide
How many questions were asked by the students	Total number of questions asked
How many questions are asked for each slide	Questions from all students across each slide

What percent of the time do students faces face the instructor	Student attentiveness across time ❏ Attentive = student gaze forward facing or recently forward facing.
What percent of the time do students face down	
Classroom activities over time	Lecture time over interactive time across time
Sponsor request	**Proposed “should” requirements**
Who asked how any questions	Which specific students ask questions
Do students who focus on the instructor ask more questions	Student questioning across attentiveness
Sponsor request	**Proposed “may” requirements**
What percent of the time do students write with a pen	Further breakdown of “interactive time” into the subcategories: ❏ Individual work ❏ Discussions ❏ Exercises ❏ Quiz/Exam
What percent of the time do students face down and engage with a laptop, tablet, or phone	Rate of student audience note taking across time ❏ Rate = percentage of student audience
What percent of the time do students face down and engage with a book	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
Sponsor request	**Proposed beyond scope of requirements**
When does the attention span start to decrease for a given educational practice and instructor	Possible stretch goal for the future.
A basic AI-based recommendation and feedback engine that provides direct feedback to the instructor based on the key measures observed	Possible stretch goal for the future.

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 report
- Final 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 source license (MIT)

Initial Product Designs

We propose to use two video/audio recordings from a class lecture to analyze student and teacher behavior. These videos will be uploaded and the program will split the videos into separate audio and video files for analysis. These files will be fed through our video or audio classifiers.

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 four video or audio classifiers 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.

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?

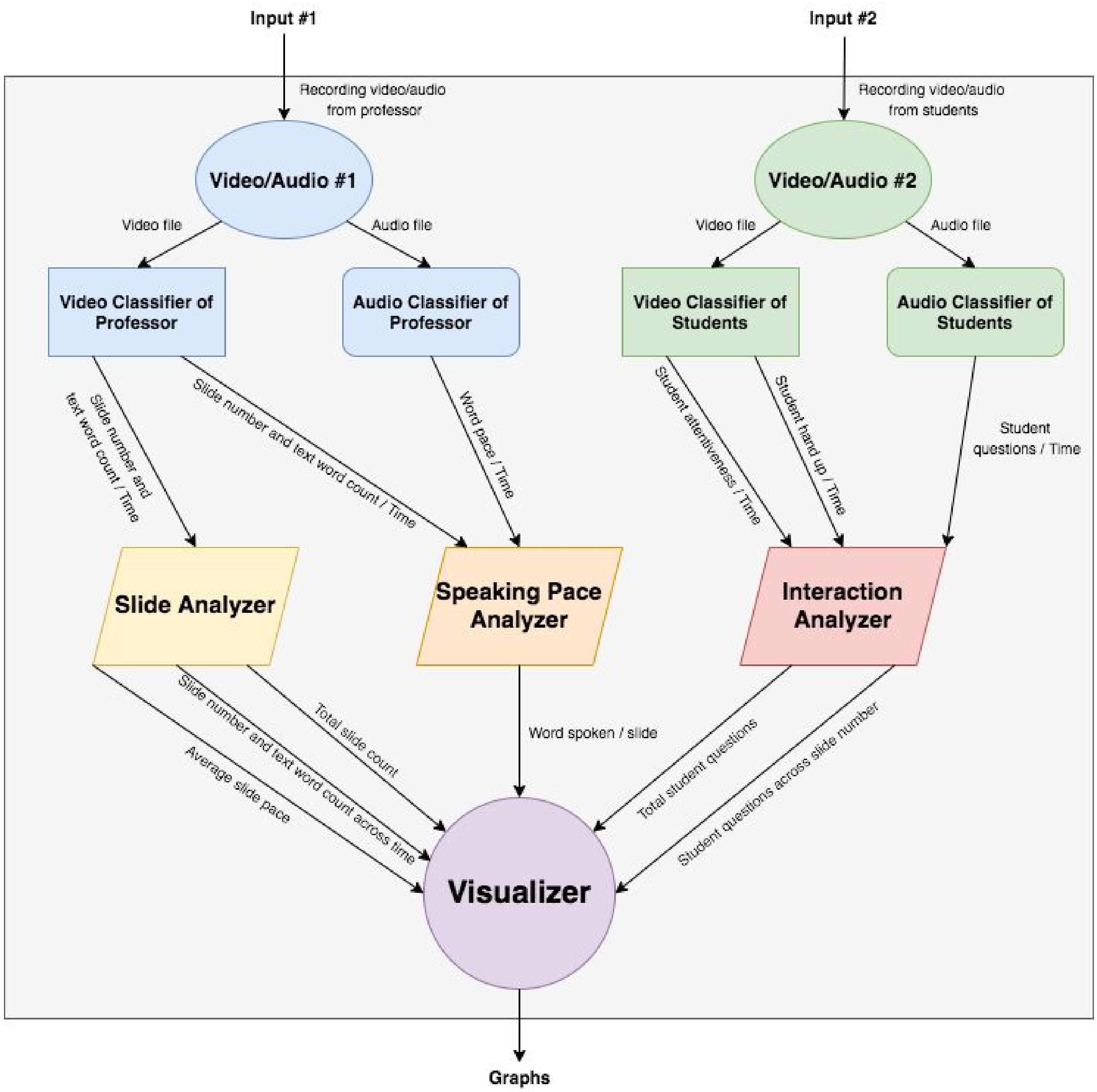
Potential Professor Classification Difficulties:

- 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?

Block Diagrams

Level 0 Block Diagram

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 algorithmic results to determine algorithmic efficacy and classification accuracy.

Tests (Human and Machine):
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

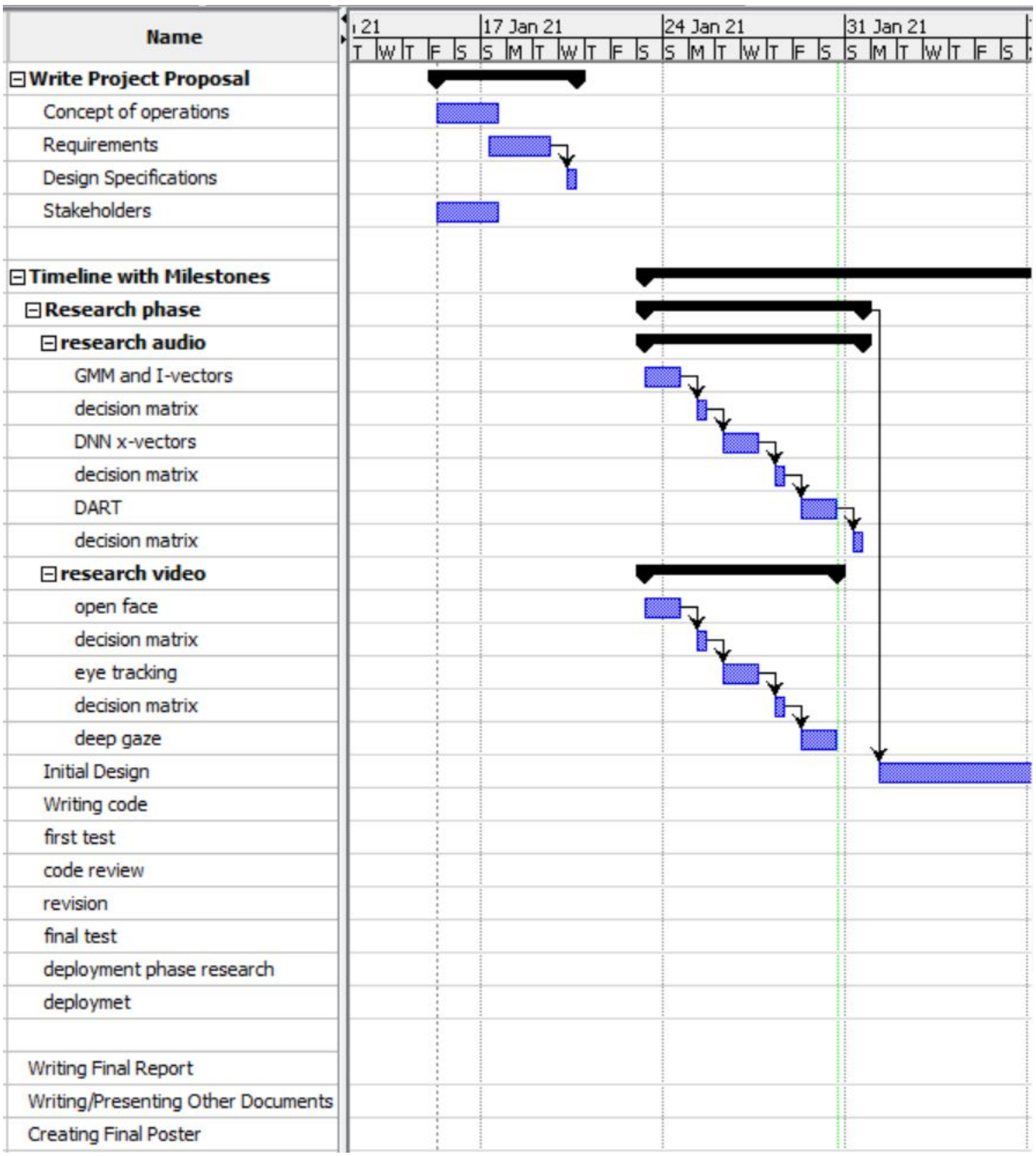
Student audience recording:

- 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

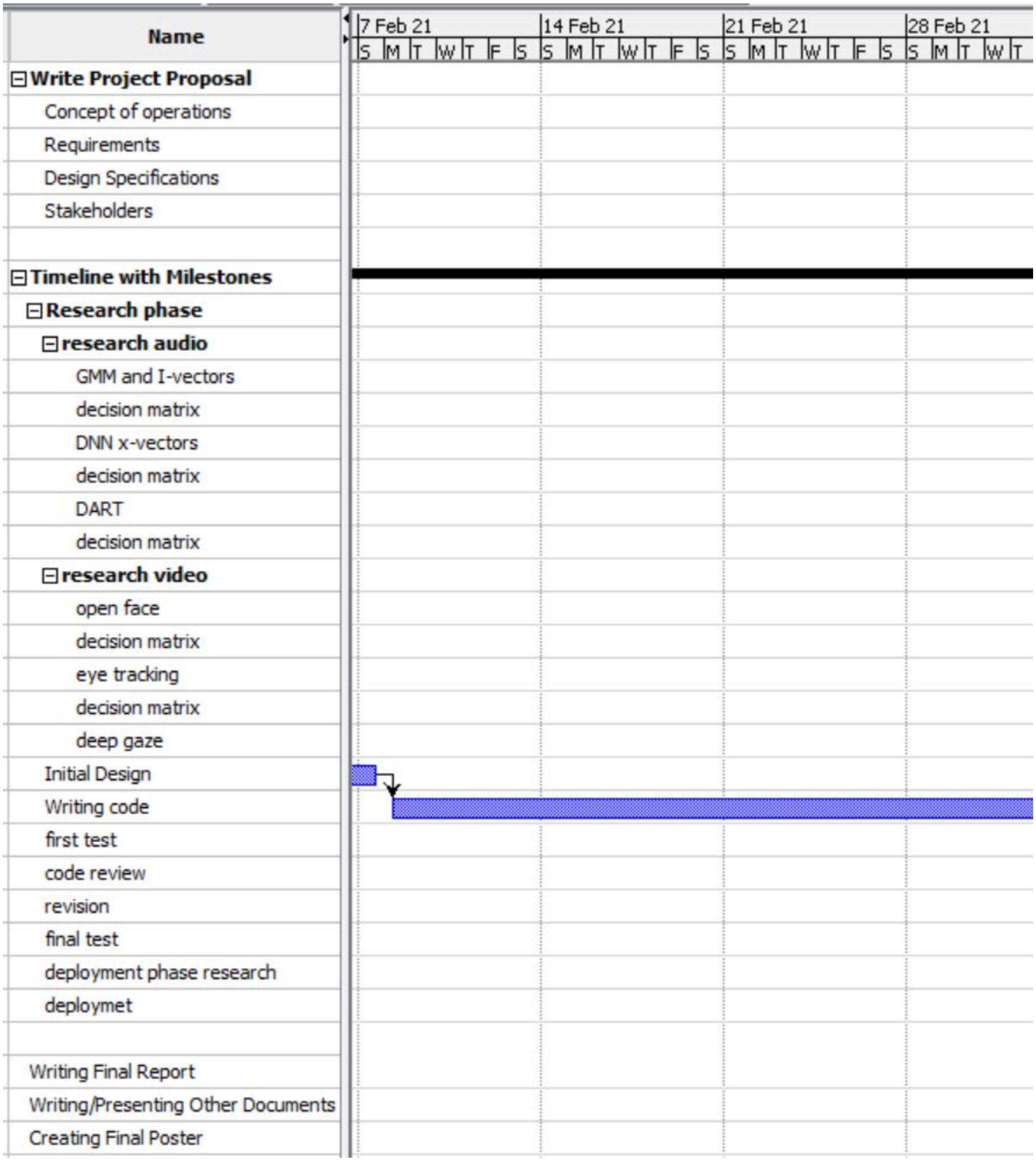
Project Management Plan

Timeline

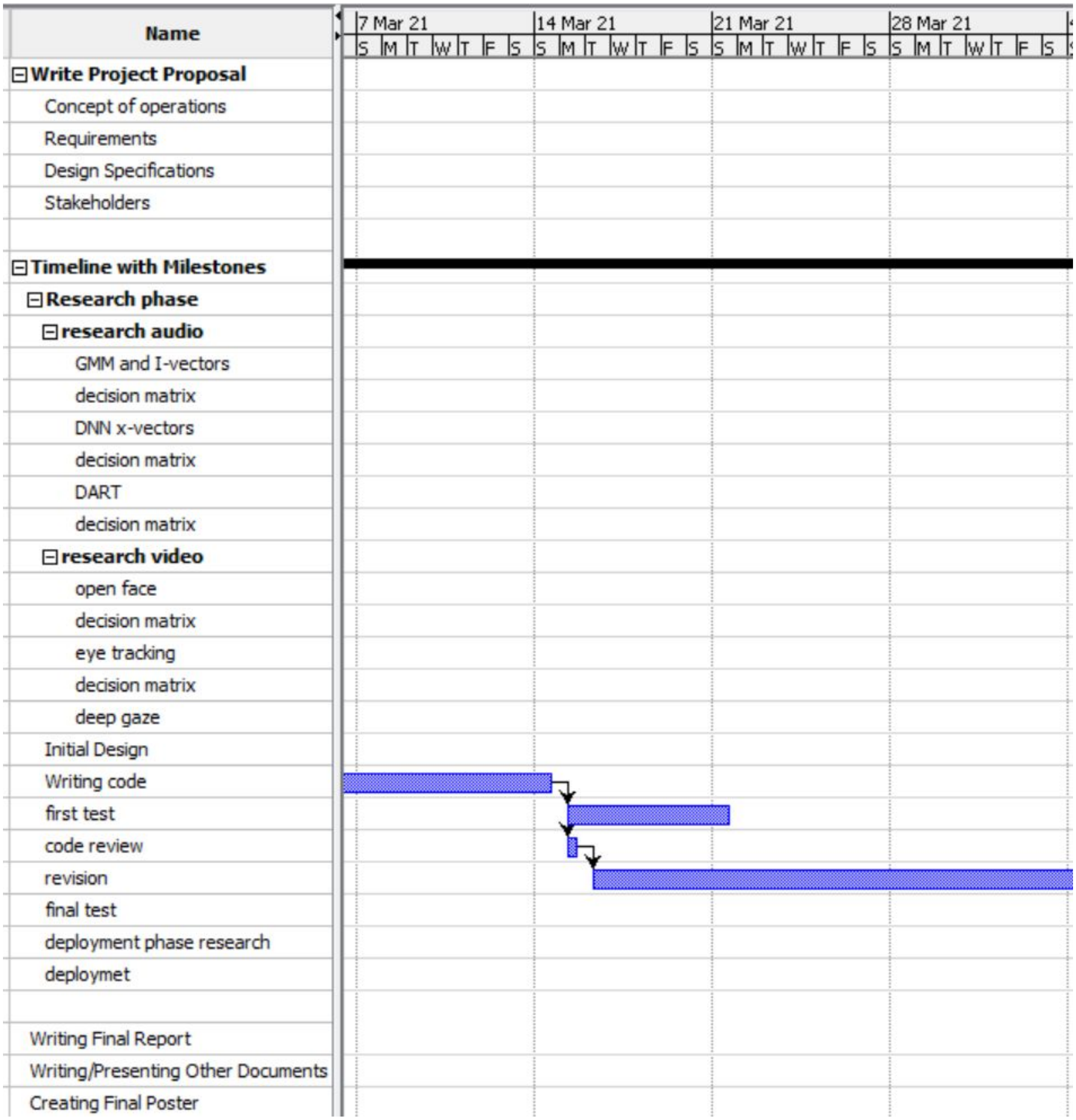
January Gantt Chart



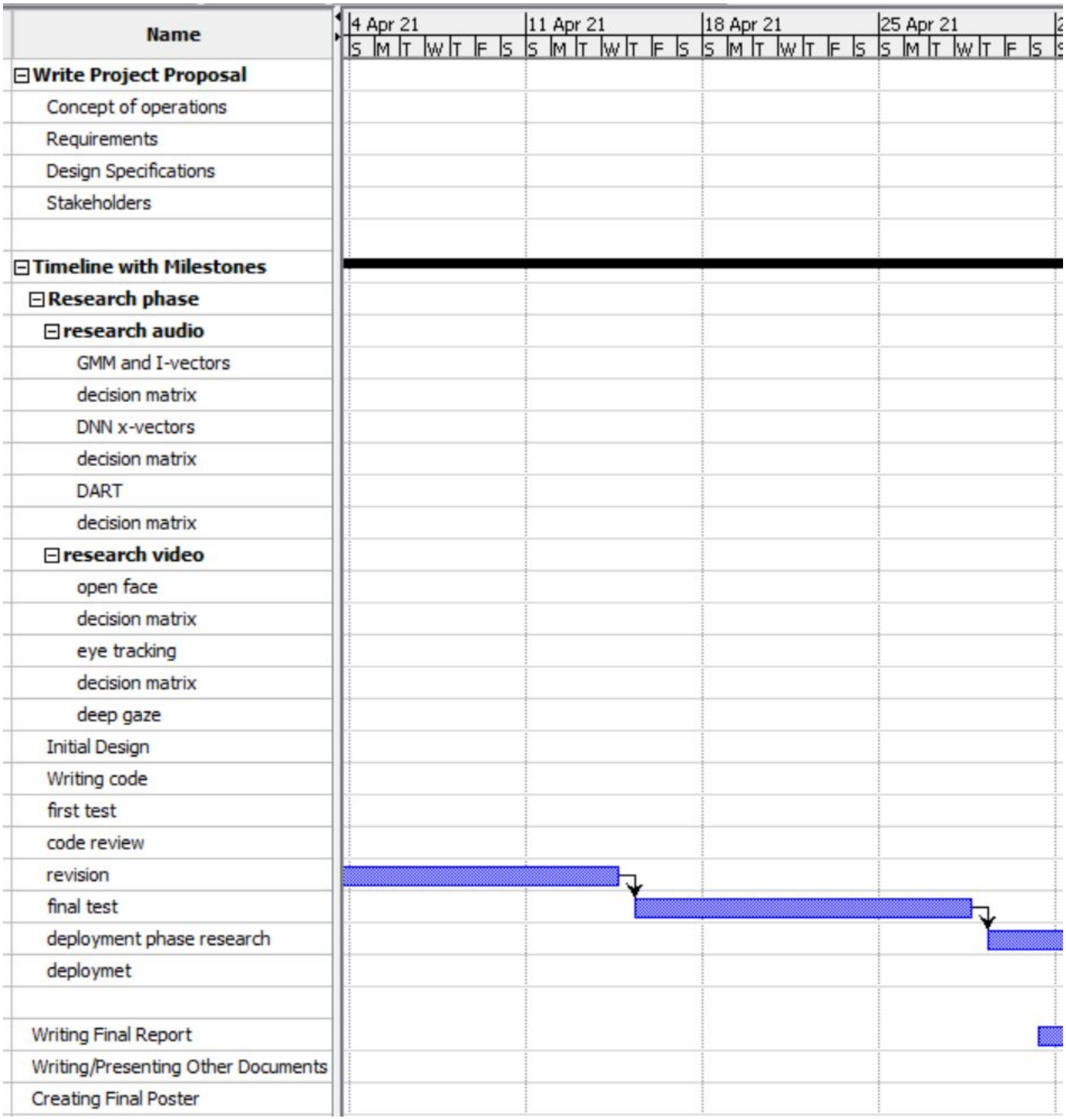
February Gantt Chart



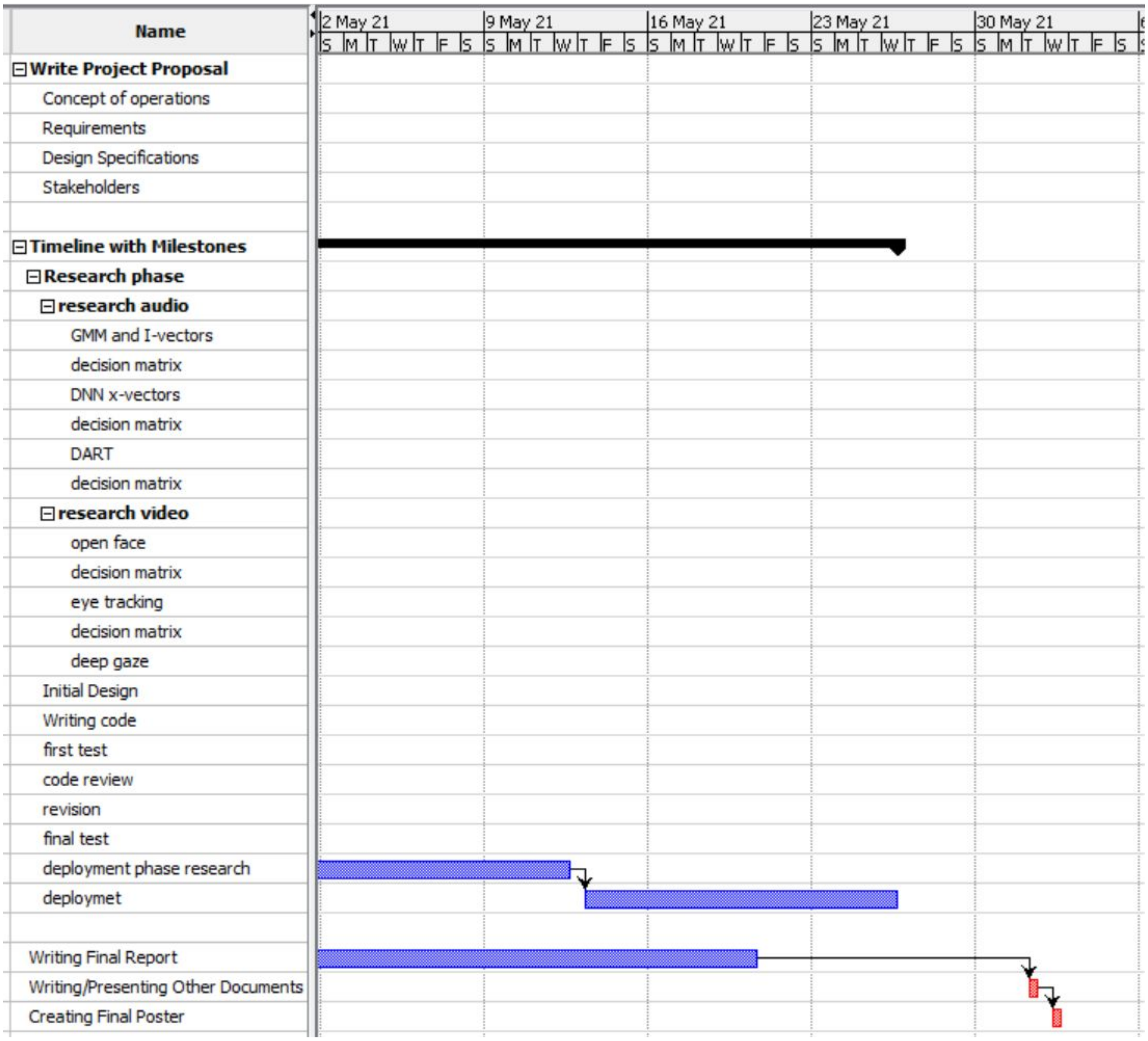
March Gantt Chart



April Gantt Chart



May Gantt Chart



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.
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

1. B.Alberts, “Classroom sound can be used to classify teaching practices in college science courses,” PNAS, vol. 114 no. 12 3085-3090, DOI: 10.1073/pnas.1618693114

2. C. Wieman and S. Gilbert, “The Teaching Practices Inventory: A New Tool for Characterizing College and University Teaching in Mathematics and Science,” CBE—Life Sciences Education, vol. 13, no. 3, pp. 552–569, 2014, DOI: 10.1187/cbe.14-02-0023

3. T. Baltrušaitis et al., "OpenFace: An open source facial behavior analysis toolkit," 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Placid, NY, 2016, pp. 1-10, DOI: 10.1109/WACV.2016.7477553.

4. M. Shojaeizadeh et al., “Detecting task demand via an eye tracking machine learning system,” Decision Support Systems, vol. 116, pp. 91–101, 2019, DOI: 10.1016/j.dss.2018.10.012.

5. H. Mokrane, “DeepGaze: ML Powered Eye Tracking & Gaze Prediction,” Towards.AI, 27-Apr-2020. [Online]. Available: https://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.
Version 0.9 - 01/28/2021 - Updated timeline with Gannt. Incorporated additional feedback.

Sponsor Sign Off

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

Signed ______________________