Real-time Emotion Recognition

Project Initiation Document

Iordache Razvan - Eduard - 3699757

PROJECT PLAN

FONTYS UNIVERSITY OF APPLIED SCIENCES

HBO-ICT: English Stream

Data student:	
Family name, initials:	Iordache R.E.
Student number:	3699757
assignment period: (from – till)	29 August 2022 - 5 February 2023
Data company:	
Name company/institution:	Fontys University of Applied Sciences
Department:	Innovation Lab
Address:	Achtseweg Zuid 151C, 5651 GW Eindhoven
Company tutor:	
Family name, initials:	Li, L.
Position:	Lecturer and researcher
University tutor:	
Family name, initials:	Van Vucht, M.
Project plan:	
Title:	Real-time Emotion Recognition
Version:	1.6
Date:	03.10.2022

Approved and signed by the company tutor:

Date:
Signature: Li,Li L. Digitally signed by Li,Li L. Date: 2022.10.10 14:14:25 +02'00'

Approved and signed by the university tutor:

Date:
Signature: Vucht,Mieke C.M. van Digitally signed by Vucht,Mieke C.M. van Date: 2022.10.10 09:17:12 +02'00'

Agreed and signed by the student:

Date:
Signature:

Version history

Version	Date	Changes
0.1	07.09.2022	• First draft
0.2	08.09.2022	• Risks, constraints, method of working
0.3	09.09.2022	• Added university tutor, research questions
0.4	12.09.2022	• Added communication planning, updated research questions, updated client company
0.5	13.09.2022	• Updated research activity, • added project timeline
1.0	14.09.2022	• Updated project timeline
1.1	16.09.2022	• Updated organization, about the client
1.2	20.09.2022	• Updated goals, deliverables
1.3	21.09.2022	• Added in-depth introduction • Updated current situation, goals, product • Updated research questions, phasing, risks
1.4	26.09.2022	• Updated timeline and phasing
1.5	02.10.2022	• Updated current situation / problem • Added glossary
1.6	03.10.2022	• Updated references

Table of contents

Introduction **5**

The Document 7

The Company 7

The Client 7

The Stakeholders 8

The Project **9**

Current situation / problem 9

The Goals 11

The Product 11

Research methods **12**

Research questions **13**

Main Question 13

Sub Questions 13

Project scope **15**

Project deliverables and non-deliverables **15**

Deliverables 15

Non-deliverables 15

Project constraints **16**

Project risks **16**

Project management **18**

Method of working 18

Project timeline **19**

Phase 1 - Initiation 19

Phase 2 - Research, prototyping and implementation 19

Phase 3 - Final testing and development of prototype 20

Phase 4 - Finalization 21

Work and communication schedule **22**

Organization **23**

References **23**

Introduction

Emotion plays an important role in life. It alone can positively influence the driving force of a person, or deeply affect it negatively. It is as important to take care of it as with a person's own body, and this can be done with the help of mental therapy.

In these troubling times, full of uncertainties and unfortunate events, it is especially important for people to take care of their mental health. Thankfully, there exist therapists that aid people with such problems. Mental therapy can help patients to regain their strength in the face of challenges, heal pains from their past and even teach them to overcome and handle strong emotions such as anger, fear, or grief.

Although this type of therapy has become popular over time, some people still have problem getting that courage to try and meet a therapist, as they feel uneasy expressing their feelings and hardships to a totally new stranger. Despite this, therapists have come with a possible solution to this problem, which hopes to accommodate new and existing patients even more.

Teletherapy is a relatively new way for people to get in touch with their therapists and participate in therapy sessions remotely, from the comfort of their own home. There are multiple benefits that comes with teletherapy, mainly more privacy, which was the primary concern of potential patients, and greater access for people that may be living in different cities or are suffering from physical disabilities. On top of that, with the recent COVID-19 health crisis, it proves to be better for public health, as it eliminates the possibility of spreading an illness.

This type of therapy, unfortunately, has its drawbacks as well. On top of the usual possible technical difficulties that can arise while in a session, communication through a digital channel greatly diminishes the readability of subtle emotional cues and certain body language. This project will focus on researching and developing a solution that can assist therapists in catching these emotional cues during their teletherapy sessions.

Glossary

Term	Description
D3	Double 3
ML	Machine Learning

The Document

This document introduces the project plan of the Emoticall Robot graduation project. The project will take place in the period of September 2022 until February 2023. The student will work 20 weeks on the graduation project.

The Company

This assignment has been provided by the Fontys ICT InnovationLab, an initiative of Fontys School of ICT and its Partners in Innovation, united in SPARC.
Located in Eindhoven, the Lab is a collaboration between students, professors, and business and research professionals, with the focus on developing next-generation ICT solutions. This is where ideas are created for the ICT industry.
Partners of the InnovationLab provide students and researchers with necessary tools and resources in order to conduct applied research projects.

The Client

Professor Dr. Mark de Graaf has been Professor of Interaction Design at Fontys Hogeschool ICT since 2015, together with Dr. Gerard Schouten (Big Data research group) on research into Data-Driven Service Design for motivation and behavioral change. In view of the enormous impact that technological developments have on the society, the ethical aspects of innovations are also explicitly included. This project is funded by lector Dr. de Graaf, and it will be the extension of KIEM project Emoticall, for which Li Li has conducted applied research work.

The Stakeholders

Stakeholder analysis approach is a research method used to gather knowledge about the main actors of this project who have high interest, influence or can be influenced by the results of this project. A stakeholder analysis was conducted with the main objective of gaining early alignment on all the stakeholder goals and plans and expectations. Secondly, the student wanted to have a clear understanding of how to communicate with each stakeholder throughout the entire project process.

Who	Role	E-Mail
Li Li	Product owner, company mentor	li.li@fontys.nl
Mieke van Vucht	University mentor	m.vanvucht@fontys.nl
Iordache Razvan	Project Leader, intern	r.iordache@student.fontys.nl
Nienke Gerrits	Psychology student	n.gerrits@student.fontys.nl
Dr. Mark de Graaf	Professor, benefactor	m.degraaf@fontys.nl

The Project

Current situation / problem

The Emoticall Therapy app was developed with the goal of aiding therapists during their remote sessions. Lead by a previous graduate student at Fontys, the prototype allows people to call their therapists remotely through a website link, without the need to install any third-party software beforehand. Once the call is initiated, a video feed of both the patient and therapist are displayed.

The web application makes use of an API to detect the current emotion of the patient. This API is built on top of TensorFlow.js, a library for machine learning in JavaScript, which allows developers to train and deploy machine learning models in the web browser.
The API in question, called “face-api.js”, uses a ML model that has been trained on a collection of images from publicly available datasets and images from the web[1]. It analyzes the facial features, such as eyebrows, lips, and contour of the face, to determine the most probable emotion they are displaying. Seven emotions can be detected: angry, disgusted, scared, happy, sad, surprised, neutral. The results from the face emotion detection API are shown in real-time above the patient’s video feed during the call.

The prototype currently features a speech emotion recognition script that transcribes the spoken words of the patient into text, and then determines whether a sentence is positive or negative by the sum of the valence of the words. This is done with the help of a JavaScript library called “sentiment”[2].
The valence level comes from psychology, which refers to the degree to which an emotional state can be positive or negative. The results from this script are not shown in real-time.

The graduation project primarily focuses on upgrading the Emoticall application with a real-time speech emotion detection algorithm.

Furthermore, The Fontys ICT InnovationLab proposed an idea that can further benefit patients and therapists, and now students and teachers.

Fontys houses a “Double 3” robot. This robot consists of a tablet-like screen at the top, connected to a pair of wheels at the bottom, resembling a segway. The robot features a webcam at the top of the screen, speakers, and a microphone, which represent the “eyes”, “mouth” and “ears” of the robot, respectively. These robots can be controlled and moved through a provided software application.

The robot has been bought by Student+, as a solution for students who, due to certain medical issues, have not been able to attend school to their desired extent. While video communication platforms like Microsoft Teams can be used to follow courses from a distance, this only allows students to see the provided material, and rarely their classmates. This leads to less interaction with other students, which is an important aspect in student life.

A group of students from Fontys created a remote-control application, as part of their ProEP project. This application allows students stuck at home to make use of the Double 3 robots very easily by accessing a website, which gives them full control over the robot, being able to see the surroundings through the robot’s webcam and maneuver it with the keyboard.

The previously mentioned idea from InnovationLab aims to combine these two projects, the Emoticall Therapy app and the Double 3 remote control application, into a single product. Certain medical issues can have a big impact on a person's mental state and can even lead to depression. If left untreated, this can strain a person's ability to carry on with their daily life and damage relationships between family and friends.
Students in this situation may benefit from this product because they will be able to interact more naturally with their university classmates and teachers while continuing with their student life, helping to alleviate the psychological effects of their condition.
Therapists would also greatly benefit, as they could further help them gradually heal through aided teletherapy sessions with the robot, keeping track of the evolution of the student's mental health.

The Goals

For this project, the final goals are the following:

1. Integrate real-time speech emotion detection algorithm into the Emoticall project
2. If proven successful, integrate the updated Emoticall project with the Double 3 remote control application

The Product

The expected result and deliverables of the project are:

1. Fully functioning multi-modal real-time emotion recognition app (facial + speech)
2. Good code quality and clear design documents
3. Performance measurements and improvements
4. Integration with the Double 3 remote control application

Research methods

Multiple strategies and methods are utilized to gather and collect data and information relevant to the project. A short summary of each method is present in the following table. For creating the research, the DOT, or The Development Oriented Triangulation Framework is used. This type of ICT research is solution-oriented, and the goal is not to create new knowledge, such as in the case of scientific research.

The following five strategies can be employed to gather results and information:

Strategy	Strategy name	Explanation
	Library research	Focuses on exploring the information and knowledge that already exist. Accessing information from an existing source: literature, documents.
	Field research	Focuses on exploring the context of the project. Gaining information on stakeholders. Interviews and focus groups.
	Lab research	Focuses on testing the prototype (parts) to see if everything works the way intended.
	Showroom research	Focuses on showcasing and gathering feedback from stakeholders, gaining insights from observing others' interactions with the project.
	Workshop research	Focuses on exploring the opportunities and possibilities by prototyping and brainstorming, to gain new insights.

Research questions

Main Question

How can Emoticall provide real-time speech emotion recognition?

Sub Questions

1. **What types of emotion recognition solutions through speech are presently available?**
2. **What is the level of reliability of these algorithms?**
3. **How would the implementation of speech emotion recognition be achieved in real-time?**
4. **How would the integration of the Emoticall app with the Double 3 remote-control application be achieved?**

The table below shows the chosen research strategies and relevant actions to be taken to reach a conclusion for each sub question.

Research question	Strategy	Methods & description
1. What types of emotion recognition solutions through speech are presently available?	 	**Available product analysis:** Search online for examples of available algorithms that manage to collect information about emotion from speech. **Literature study:** Read articles of interest about applied solutions. **Community research:** Ask various communities and forums about possible libraries that could help. **Document analysis:** Read the previous intern's documentation on the matter.
2. What is the level of reliability of these algorithms?	 	**Literature study:** Research about how the algorithms were implemented and analyze their success rate. **Community research:** Ask relevant communities and forums about the accuracy of their preferred solutions. **Available product analysis, Best good and bad practices** **Component test:** Test available code snippets and determine the rate of success.

3. How would the implementation of speech emotion recognition be achieved in real-time?	 	**Prototyping:** Try and implement multiple options in small prototypes until the most optimal solution is reached. **Unit test:** While implementing the most optimal way to analyze speech emotion, test for any bugs that may occur. **Usability testing:** Test the prototype with psychology students while developing.
4. How would the integration of the Emoticall app with the Double 3 remote-control application be achieved?	  	**Code review:** Analyze the code of the remote-control application in detail. **Component test:** Test parts of the prototype while in the process of integration of the two applications. **Usability testing:** Further test the prototype with psychology students. **Peer review:** Test the prototype with other students.

Project scope

The project scope defines what is within and outside the confines of this project.

Within the project scope
Research an improved way of emotion detection through speech
Restructure and refactor the code of the Emoticall app
Implement the real-time voice emotion recognition feature into Emoticall
Integrate the emotion detection features of the updated Emoticall with the D3 remote control application

Project deliverables and non-deliverables

Deliverables

- ∉ Project plan
- ∉ Project final report
- ∉ Working prototype of the Emoticall project with real-time speech emotion detection
- ∉ Working prototype of the Double 3 remote control application integrated with the updated Emoticall application
- ∉ Final presentation
- ∉ User (and technical) manual
- ∉ Day declaration

Non-deliverables

- Maintenance of the final product

Project constraints

Constraint 1: Time

The project should be completed within 5 months, starting from 29.08.2022 until 05.02.2023. All final deliverables must be submitted by the deadline.

Constraint 2: Working days

The number of working days must consist of a minimum of 90 days.

Project risks

Risk 1: Time mismanagement

Probability: Low

Impact on Project: Medium

Mitigation: Make a project plan, follow the planned timeline, and complete tasks for each iteration and respect deadlines. Make different timelines for different case scenarios. In the case the timeline is not strictly followed, the planning should be changed accordingly.

Risk 2: Lack of communication

Probability: Low

Impact on Project: Medium

Mitigation: Have a communication plan which includes frequency, goal, and audience of each communication. Identify stakeholders early and make sure they are considered in the plan. Use the most appropriate channel for the audience.

Risk 3: Improper problem definition

Probability: Medium

Impact on Project: High

Mitigation: Review the scope of the project with the company tutor and the stakeholders in detail before moving on to the next phase. Make necessary changes if the need arises.

Risk 4: Lack of technical knowledge

Probability: Medium

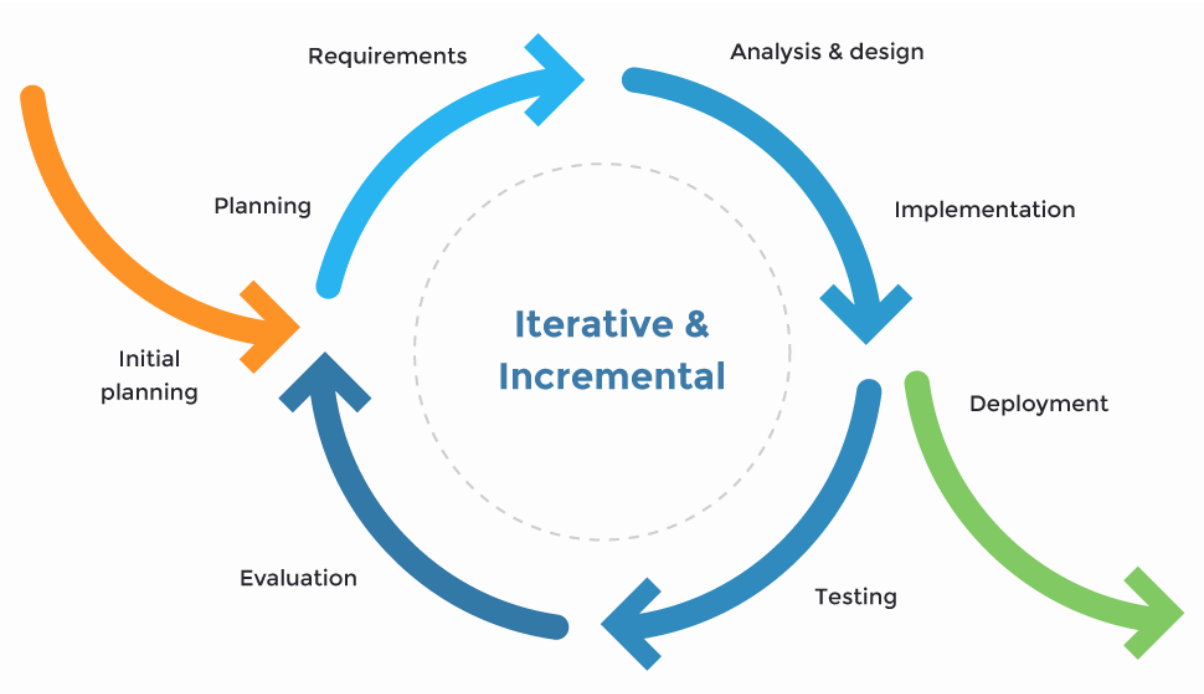
Impact on Project: High

Mitigation: Research in full before moving on to implementation. Keep in touch with the company mentor whenever there are issues.

Project management

Method of working

For this project, the Agile Iterative and Incremental method of development has been chosen.



Firstly, an incremental process consists of delivering features of a software in parts. Each part, or increment, represents a complete subset of functionality.
Secondly, the iterative process represents making progress through subsequent refinement. With each iteration, the developer enhances the initial features until the product is satisfactory.
The chosen method of development combines these two concepts. Using this allows the developer to gradually build up their product and refine and adjust any of the features through multiple iterations, according to the stakeholders' feedback.

Project timeline

The timeframe of the project is 20 weeks. Each iteration will last approximately three weeks. Milestones are objectives planned to be reached by the end of each iteration.

Phase 1 - Initiation

Duration: Weeks 1 - 3

This first phase serves as an introduction to the project. Here, the student gets to understand more about the assignment and studies the code and documentation of the projects from the previous groups.

Activities:

- Project research – understanding past projects (Emoticall real-time emotion recognition and Hybrid Learning Robot)
- Create first versions of the project plan

Phase 2 - Research, prototyping and implementation

Duration: Weeks 4 - 13

The second phase involves delving into research about speech emotion recognition and prototyping and implementing real-time speech emotion recognition. In this phase, there will be three iterations or sprints, each with their own progress goal or milestones.

Iteration 1: Weeks 4 – 6

Activities:

- Research about how to analyze speech emotion recognition
- Check existing solutions for speech emotion recognition from the Internet and test them to analyze their accuracy and reliability levels
- Share project progress with company tutor weekly and receive feedback
- Create software architecture design document

Milestone: Reach the most suitable solution for speech emotion recognition.

Iteration 2: Weeks 7 - 10

Activities:

- Improve prototypes according to the stakeholders' feedback
- Research about upgrading the speech emotion recognition to be real-time
- Implement and test real-time speech emotion recognition in a working prototype
- Refactor code of Emoticall for clear structure after implementation of real-time speech emotion recognition
- Share project progress with company tutor weekly and receive feedback
- Start working on the final project report

Milestone: Working prototype of Emoticall with real-time speech emotion recognition.

Iteration 3: Weeks 11 - 13

Activities:

- Improve prototype according to the stakeholders' feedback
- Continue testing of the real-time speech emotion recognition
- Analyze the Hybrid Learning Robot project code in order to prepare the integration with the real-time emotion recognition project.
- Start integration of the real-time emotion recognition project with the Hybrid Learning Robot project
- Continue working on the final project report
- Share project progress with company tutor weekly and receive feedback

Milestone: First prototype of integrated Hybrid Learning Robot project with real-time emotion recognition.

Phase 3 - Final testing and development of prototype

Duration: Weeks 14 - 16

During this phase, the final steps towards full integration of real-time emotion recognition with the Hybrid Learning Robot project will be made.
Final testing of the prototype will be done with psychology students and other users.

Activities:

- Finalize a working prototype of the fully integrated Hybrid Learning Robot project with real-time emotion recognition
- Functional and user testing
- Perform any bug fixes found in testing
- Continue working on the final project report
- Share project progress with company tutor weekly and receive feedback

Milestone: Finished prototype of integrated Hybrid Learning Robot project with real-time emotion recognition, with additional bug fixes.

Phase 4 - Finalization

Duration: Weeks 17 - 20

Towards the end of the project, the intern should have the finished prototype of the Hybrid Learning Robot project with the multi-modal real-time emotion recognition features. In terms of documentation, the final report and a user manual will be available, and an end presentation will be held, showing the progress and final product of the project.

Activities:

- Final testing of the product
- Create user and technical manual
- Prepare final presentation
- Deliver the final product and finished project report

Milestone: Final product delivery, final report delivery, final presentation

Work and communication schedule

The table below lists the various meetings that the student will participate in, showing the members that are required to be present, as well as the frequency of the meetings.

Type of communication	Objective	Medium	Frequency	Who is involved
Weekly meeting with company mentor	Discuss what has been done, and any challenges the student came across.	Via MS teams	Monday	Company mentor Student
Weekly update to university tutor	Shortly present what has been done and achieved during the week.	Via email	Every Friday	University mentor Student
Knowledge Sharing Sessions	Showcase of progress from the previous week. Share and receive feedback amongst peers.	On-site at TQ4.2	Every Tuesday	All Fontys interns

Organization

Name	Role
Iordache Razvan	Project leader
Li Li	Company tutor
Mieke van Vucht	University tutor
Gennip Bart	Company colleague

References

1. Mühler, V. [justadudewhohacks]. (n.d.). JavaScript API for face detection and face recognition in the browser and nodejs with tensorflow.js. Retrieved September 30, 2022, from https://github.com/justadudewhohacks/face-api.js#models-face-expression-recognition
2. Sliwinski, A. [thisandagain]. (n.d.). AFINN-based sentiment analysis for Node.js. Retrieved October 4, 2022, from https://github.com/thisandagain/sentiment