



A report submitted in partial fulfilment of the requirements for the degree  
of Bachelor of Science (BSc) in

**COMPUTER SCIENCE**

**UNIVERSITY OF THE WEST OF ENGLAND**

**FACIAL EMOTION RECOGNITION  
FOR MUSIC RECOMMENDATION SYSTEM**

By

YIE NIAN CHU

Supervisor: Craig Duffy

School of Computing  
and Creative Technologies

UNIVERSITY OF THE WEST OF ENGLAND

Date of submission: April 11, 2024  
Word count: WORD COUNT

# **DECLARATION**

I, Yie Nian Chu confirm that the work presented in this report is my own. Where information has been derived from other sources, I confirm that this has been indicated in the report.

---

Yie Nian Chu

# **ABSTRACT**

## **ACKNOWLEDGEMENTS**

# ACRONYMS

**AI** Artificial Intelligence

**CNNs** Convolutional Neural Networks

**FER** Facial Emotion Recognition

**GD** Gradient Descent

**k-NN** K-Nearest Neighbors

**LR** Linear Regression

**ML** Machine Learning

**MLP** Multi-Layer Perceptron

**MSE** Mean Squared Error

**RF** Random Forest

**SVM** Support Vector Machine

# CONTENTS

<b>Declaration</b>	<b>1</b>
<b>Abstract</b>	<b>2</b>
<b>Acknowledgement</b>	<b>3</b>
<b>Acronyms</b>	<b>4</b>
<b>List of Figures</b>	<b>9</b>
<b>List of Tables</b>	<b>10</b>
<b>1 Introduction</b>	<b>11</b>
1.1 Background .....	11
<b>2 Literature Review</b>	<b>12</b>
2.0.1 Introduction of Music .....	12
2.0.2 Evolution and Diversity of Music .....	13
2.0.3 Music and Emotion.....	15
2.1 Artificial Intelligence and Machine Learning .....	15
2.1.1 Artificial Intelligence.....	15
2.1.2 Machine Learning .....	16
2.1.3 Machine Learning Algorithms .....	16
2.1.3.1 Linear Regression.....	16
2.1.3.2 Support Vector Machine.....	18
2.1.3.3 Random Forest.....	19
2.1.3.4 K-Nearest Neighbors.....	20
2.1.3.5 Neural Networks .....	21
2.1.3.5.1 Multi-Layer Perceptron.....	21

2.1.3.5.2	Convolutional Neural Networks .....	21
2.2	Facial Emotion Recognition .....	23
2.3	Music Recommendation based on FER .....	28
<b>3</b>	<b>Requirements</b>	<b>29</b>
3.1	Functional Requirements and Non-functional Requirements.....	29
3.1.1	Functional Requirements .....	29
3.1.2	Non-Functional Requirements .....	31
<b>4</b>	<b>Methodology</b>	<b>35</b>
4.1	Introduction .....	35
4.2	Research Methodology .....	35
4.2.1	Waterfall methodology .....	35
4.2.2	Spiral methodology .....	36
4.2.3	Agile methodology.....	37
4.3	Comparison and Selection.....	37
4.4	Justification for Choosing Agile .....	39
<b>5</b>	<b>Design</b>	<b>41</b>
5.1	Introduction .....	41
5.2	UML Diagrams.....	41
5.2.1	Block Diagram .....	41
5.2.2	Use Case Diagram.....	44
5.2.3	Sequence Diagrams.....	44
5.2.4	Class Diagram.....	44
5.2.5	Flowchart .....	44
5.2.6	Entity-relationship Diagram .....	44
5.3	Logo Design .....	44
5.4	Interface Design.....	44
5.5	Artificial Intelligence and Machine Learning .....	44
5.5.1	Model Architecture .....	44
<b>6</b>	<b>Implementation</b>	<b>48</b>
<b>7</b>	<b>Project Evaluation</b>	<b>49</b>

<b>8 Conclusion and Future Outlook</b>	<b>50</b>
8.1 Conclusion .....	50
8.2 Future Outlook .....	50
<b>References</b>	<b>55</b>
<b>Appendix</b>	<b>56</b>
Appendix A .....	56
Appendix B .....	57

# LIST OF FIGURES

2.2	Linear Regression .....	17
2.3	Gradient Descent .....	17
2.4	Support Vector Machine .....	18
2.5	Random Forest .....	19
2.6	K-Nearest Neighbors .....	20
2.7	Multi-Layer Perceptron.....	21
2.8	Convolutional Neural Networks .....	22
2.9	Facial Landmarks .....	23
2.13	Linear Regression Classification Accuracies Table .....	26
2.14	Random Forest Classification Accuracy .....	27
2.15	Support Vector Machine Accuracy .....	27
4.1	Waterfall Methodology (Communitcation Team, 2022).....	36
4.2	Spiral Methodology (Kumar Pal, 2018).....	36
4.3	Agile Methodology (Laoyan, 2022) .....	37
4.4	Kanban from Notion .....	39
5.1	Block Diagram .....	41
5.2	Use Case Diagram .....	42
5.3	Login Sequence Diagram .....	42
5.4	Emotion Recognition Sequence Diagram .....	43
5.5	Playlist Generation Sequence Diagram .....	43
5.6	Flowchart .....	43
5.9	Login Page .....	45
5.10	Sign Up Page .....	45
5.11	Dashboard Page .....	46
5.12	Emotion Music Page .....	46

5.13 Emotion Recognition Page .....	46
5.14 User Settings Page.....	47

# **LIST OF TABLES**

3.1 Functional Requirements.....	30
3.2 Non-Functional Requirements.....	34
4.1 Comparison of Methodologies.....	38

# **1. INTRODUCTION**

## **1.1. BACKGROUND**

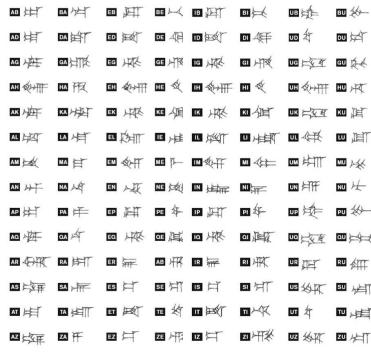
## **2. LITERATURE REVIEW**

### **2.0.1. Introduction of Music**

The development of human communication and social cohesiveness is closely linked to the origins of music. Protohumans probably created the first musical sounds through rhythmic awareness and vocalizations with variable pitch before musical instruments were invented. Primitive humans showed the ability to compose music even before they learned to speak. The development of early human civilizations may have benefited from this musical talent by fostering social cohesion and collaboration.

The comparison of early speech or the pitched sounds of animals to historical music demonstrates how primitive music was as a medium of expression before sophisticated language. A tight relationship exists between the growth of rhythmic awareness and the capacity to produce percussion sounds and the formation of music in the evolutionary timeline. The ability to perceive and produce music was probably important for social interactions when *Homo neanderthalensis* and *Homo sapiens* evolved. One major explanation for the evolution of music is the innate human desire to collaborate and form social contexts for doing so. Early human society appears to have relied heavily on music, judging from its role in fostering social cohesiveness, entertaining, enabling dance, and supporting ritualistic behaviors.

Essentially, the origins of music can be found in the basic human urge for collaboration, socialization, and the development of cultural contexts that support these endeavors. Even in its most primitive incarnations, music influenced human communities and may have helped *Homo sapiens* outperform other hominid species in terms of cultural development.



(a) Cuneiform Alphabet  
Image from (Finkel and Taylor, 2021)



(b) World's Earliest Music Composition  
Image from (Porter and Molana-Allen, 2018)

## 2.0.2. Evolution and Diversity of Music

The history of music is extensive and ancient, and it is an essential part of human culture. In Syria, a cuneiform "alphabet" (a) containing the earliest known written composition of music was found (b) (Porter and Molana-Allen, 2018). It is thought to have been composed some 3400 years ago. The oldest musical instruments, according to archaeological evidence, date back approximately 40,000 years, but the history of music is far older still (Killin, 2018). Despite not yet being documented in the archaeological record, these instruments offer a window into far older musical endeavors. Proto-musical evolution most likely started about 400,000 years ago, based on the social brain hypothesis (Dunbar, 1998). Musical traditions developed further as people left Africa and spread over the world, reaching a turning point during the Holocene. This historical voyage demonstrates the persistent and varied impact that music has played in human history. Moving forward, there are so many types of music, each representing unique expression of culture, emotion and creativity:

- **Classical Music:** a diverse and evolving tradition, extending beyond the commonly associated period of 1750 to 1820 and encompassing composers from Bach to contemporary artists. It serves as a living and influential force in the world of music, shaping compositions for orchestras, chamber ensembles, solo performers, and even finding unexpected expressions in various genres, from video-game scores to popular music. (Gabler, 2013)
- **Jazz:** originating in early 20th-century New Orleans, is characterized by complex harmony, syncopated rhythms, and a focus on improvisation. Emerging from a rich tradition of ragtime and blues, jazz evolved into a versatile genre that expanded

its influence globally, encompassing popular music standards, modal music, and even avant-garde compositions.(Beek, 2021)

- **Blues:** both a musical form and genre, initially associated with melancholy themes, has evolved to encompass a broader range of subjects and emotions, aiming to uplift through music.(Chaudhuri, 2022) Characterized by specific chord progressions, a walking bass, call and response, and unique features like microtonality and flattened 'blue' notes, blues is known for its distinct sound and expressive style. (BBC, 2023)
- **Electronic:** utilizes diverse sound sources, ranging from recorded sounds captured by microphones to those generated by electronic oscillators and complex computer installations. Typically played back through loudspeakers, electronic music can be created using various technologies, with the exception of "live electronic music," which involves real-time performance. (Hiller, 2019)
- **Country:** originating in rural Southern and Western areas in the early 20th century, was initially labeled "hillbilly music" before adopting the official term "country and western music" in 1949. Rooted in the ballads, folk songs, and popular tunes of English, Scots, and Irish settlers, country music gained commercial recognition in the early 1920s, marked by its realistic portrayal of rural life contrasting with the sentimental tone prevalent in popular music of that era. (Britannica, 2019)
- **Hip-Hop:** a cultural movement that gained widespread popularity in the 1980s and '90s, serves as the foundational music for rap—a style featuring rhythmic and/or rhyming speech, which became the movement's enduring and influential art form. Hip-hop, beyond its musical dimension, encompasses diverse elements such as graffiti art, breakdancing, and social activism, reflecting a multifaceted expression of urban culture and creativity. (Tate and Light, 2019)
- **Rock:** a form of popular music that emerged in the 1950s and by the end of the 20th century became the dominant global music genre, influencing the recording industry, international retail, and radio and television playlists. While dictionary definitions often focus on its strong beat and instrumentation, the cultural significance of rock lies in its social and ideological distinctions from other music

genres, particularly its development as a term to distinguish certain attitudes and practices from those associated with pop music.(Frith, 2018)

### **2.0.3. Music and Emotion**

Since ancient times, people have been fascinated by the paradoxical connection that exists between music and emotion. Even though music is an abstract art form that appears to be removed from everyday life, it has a profound potential to evoke strong emotional responses. This ability of music to trigger strong emotions is also demonstrated in other social circumstances, such as advertising. This encounter is further enhanced by the relationship that exists between music and our individual life experiences. Emotions, influenced by these encounters, provide our perception and thought processes a personalized meaning that connects the abstract quality of music to the concrete events of our everyday life. This complex tapestry that highlights the profound influence of music on our emotional environment is created by the blending of music, emotion, and human experience (Juslin and Sloboda, 2013).

Numerous musical elements that have been thoroughly explored are combined to convey emotions through music. Pace, mode, harmony, interval, rhythm, sound level, timbre, timing, articulation, accents, tone attacks and decays, and vibrato are some of these characteristics. Emotion in music is expressed through compositional elements as well as performance elements. Still, it's not an easy task to express feelings through music. Certain musical elements can be employed to convey a range of moods, showing that certain elements are not always reliable predictors of a certain emotion. The Lens Model (Juslin and Sloboda, 2013), which characterizes emotional expression in music as involving probabilistic and partially redundant auditory cues, sheds more light on this complexity. Listeners combine several cues for successful emotion recognition, and the redundancy of cues allows for a high level of emotion recognition through different combinations, offering room for creativity and personal expression (Pereira et al., 2011).

## **2.1. ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

### **2.1.1. Artificial Intelligence**

Artificial Intelligence (AI) is a broad field that includes using technology to build machines and computers that can replicate cognitive abilities connected to human

intellect. These abilities include making recommendations, language understanding, data processing, and visual perception. AI should be viewed as a group of technologies incorporated into systems rather than as a stand-alone system that can understand, learn, and respond to complex problems.

### **2.1.2. Machine Learning**

Machine Learning (ML) is a branch of AI focused on enabling machines and systems to learn and enhance their performance through experience. Instead of relying on explicit programming, machine learning employs algorithms to analyze vast datasets, derive insights, and subsequently make informed decisions. These algorithms continually improve their performance as they are exposed to more data. The outcomes of this learning process are the machine learning models, which become more proficient with increased exposure to data.

### **2.1.3. Machine Learning Algorithms**

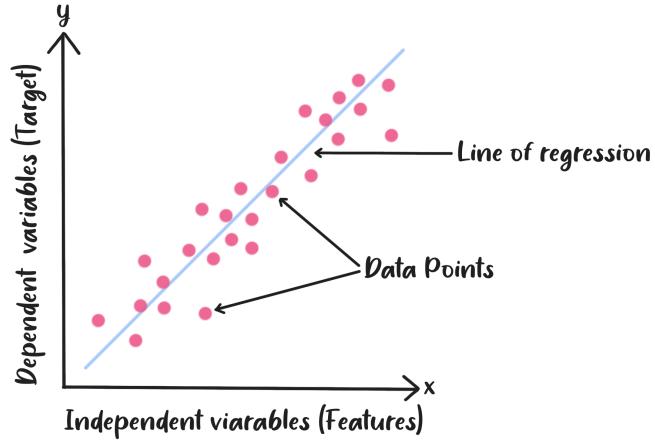
#### **2.1.3.1. Linear Regression**

Linear Regression (LR) is a supervised learning algorithm used to model the relationship between a dependent variable (target) and one or more independent variables (features). The fundamental assumption of LR is that there exists a linear relationship between the input variables and output. (IBM, 2022)

$$y = mx + b \quad (2.1)$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon \quad (2.2)$$

A LR model can be represented by the equation 2.2 where  $Y$  represented the dependent variable.  $\beta_0$  is the y-intercept,  $\beta_1, \beta_2, \dots, \beta_3$  are the coefficients and the  $\epsilon$  is an error term, representing the unobserved factors that affect  $Y$  but are not accounted for by the model. The logic of it is same as Linear Equation (2.1) where using the gradient of the line ( $m$ ), the value of  $x$  and the y-intercept ( $b$ ) to get the value of  $y$ , which is what we are trying to predict.

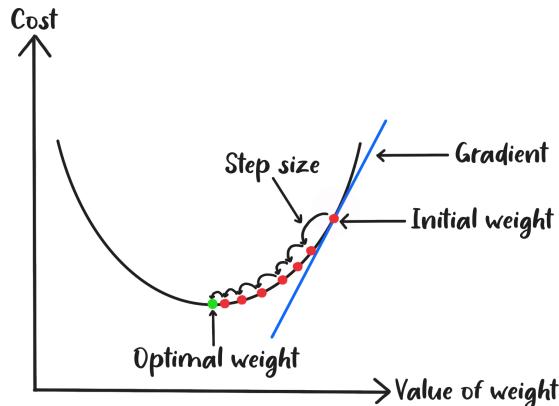


**Figure 2.2:** Linear Regression

While using LR, the main objective is to find the values of  $\beta_0, \beta_1, \dots, \beta_n$  that minimize the error between the predicted value ( $\hat{Y}$ ) and the actual value ( $Y$ ) in the training data.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2.3)$$

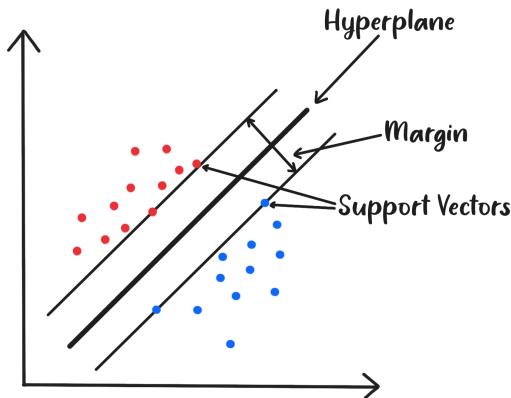
Therefore, Mean Squared Error (MSE), a cost function to measures the average squared difference between  $\hat{Y}$  and  $Y$ , is introduced to quantify the goodness of fit of the model to the training data. If the current result is not optimized, Gradient Descent (GD), an optimization algorithm, will be used to adjust the coefficients,  $\beta_0, \beta_1, \dots, \beta_n$ , towards the direction that minimizes the MSE until it reaches the convergence (Figure 2.3).



**Figure 2.3:** Gradient Descent

### 2.1.3.2. Support Vector Machine

Support Vector Machine (SVM) is a supervised ML algorithm where we used for classification, regression and outliers detection. The main goal of it is to find the hyperplane that best separates the data into different classes based on statistical approaches. (Géron, 2019)



**Figure 2.4:** Support Vector Machine

While using SVM, the greater the margin, the better the result would be as it has better generalization to new or unseen data. There are two types of SVM for classification, which are Linear SVM and Non-linear SVM. A linear SVM finds the optimal hyperplane that maximizes the margin between classes for linearly separable data as Figure 2.4.

$$f(x) = \text{sign}(\mathbf{w} \cdot \mathbf{x} + b) \quad (2.4)$$

The decision function of linear SVM (Equation 2.4) is used to defines the ability to classify data points into different classes. When the result is greater than or equal to zero, the prediction would be positive. If  $f(x)$  is less than zero, the decision function predicts the negative class.

For non-linearly separable data, SVM uses kernel functions such as polynomial, sigmoid and radial basis function (RBF), to map the data into a higher-dimensional space where hyperplane can separate the classes. The equation 2.5 is the decision

$$f(x) = \text{sign}\left(\sum_{i=1}^n \alpha_i y_i K(x, x_i) + b\right) \quad (2.5)$$

$$K(x, x_i) = (x \cdot x_i + c)^d \quad (2.6)$$

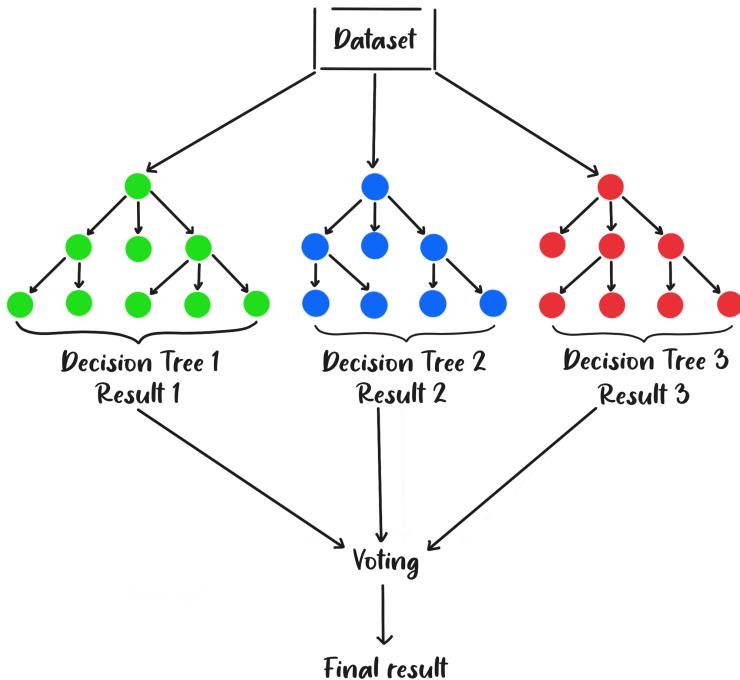
$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (2.7)$$

$$K(x, x_i) = \tanh(\alpha x \cdot x_i + c) \quad (2.8)$$

function of non-linear SVM. The kernel function, denoted by  $K(x, x_i)$  in the equation, would be replaced by equation 2.6 if a polynomial kernel were used. Similar with the other kernels, if the RBF kernel is employed, it would be exchanged with equation 2.7, and for sigmoid kernel, it would be replaced with equation 2.8.

### 2.1.3.3. Random Forest

Random Forest (RF) is an ensemble learning algorithm that belongs to the family of decision tree-based methods. A group of decision trees that have been trained on various dataset subsets make up the forest of RF, and the final result is derived from averaging the predictions of each individual tree. (IBM, 2023b)



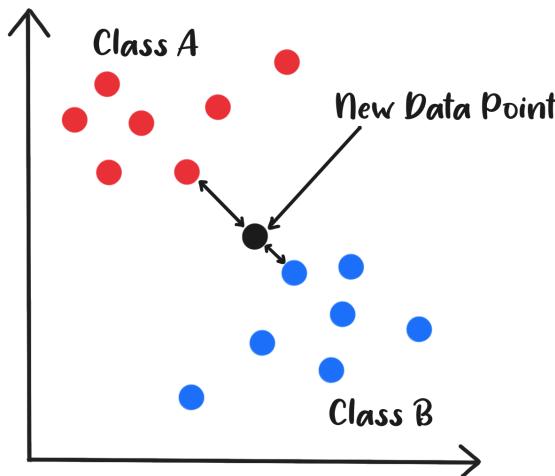
**Figure 2.5:** Random Forest

As Figure 2.5 shown, RF builds multiple Decision Trees and combines them to get a

more accurate and stable prediction than any individual model. This process is called bagging which is one of a type of ensemble learning. In the process, the entire dataset is separated into subsets and each decision tree is trained individually on a subset that is selected at random. This adds variety and unpredictability to the trees. The training process will then generate results for each model, and the final output is determined by the "votes" for a class from each tree. The class with the majority of votes is chosen as the final prediction.

#### 2.1.3.4. K-Nearest Neighbors

K-Nearest Neighbors (k-NN) is a intuitive supervised machine learning method used for both classification and regression tasks. The main idea of k-NN is to predict the label of new data point based on its k-nearest data points in the feature space. (IBM, n.d.b)



**Figure 2.6:** K-Nearest Neighbors

k-NN uses distance metric (Equation 2.9) to calculate how similar two data points are to one another. k-NN finds the k training data points that, according to the selected distance metric, are closest to a given data point. In classification tasks, the majority class among a new data point's k-nearest neighbors predicts the class that the data point will fall into.

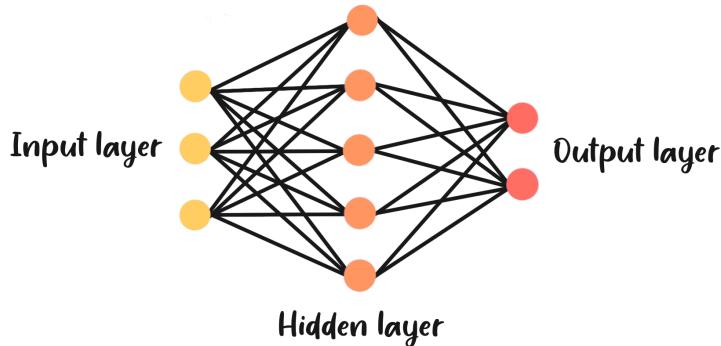
$$d(P, Q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (2.9)$$

$$\hat{Y} = \operatorname{argmax}_y \left( \sum_{i=1}^k I(y_i = y) \right) \quad (2.10)$$

The key hyperparameter of k-NN is the value of  $k$  (Equation 2.10), representing the number of nearest neighbors to consider. The choice of  $k$  can significantly impact the performance of the algorithm, and it is often selected through cross-validation.

### 2.1.3.5. Neural Networks

#### 2.1.3.5.1 Multi-Layer Perceptron

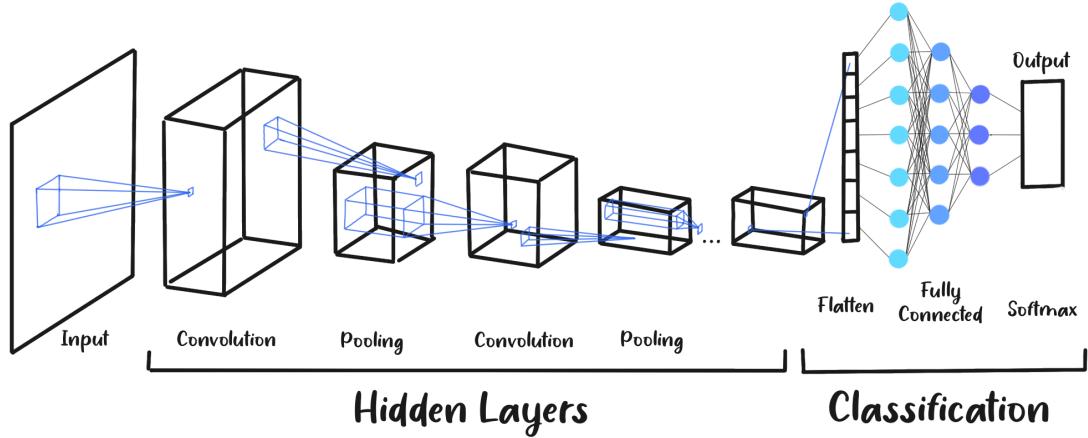


**Figure 2.7:** Multi-Layer Perceptron

An input layer, one or more hidden layers, and an output layer are the minimum number of nodes that make up a Multi-Layer Perceptron (MLP) feedforward artificial neural network type (Figure 2.7). All nodes in these layers—aside from those in the input layer—are linked using a specific weight and employ a nonlinear activation function. Because of its nonlinearity, the network may learn and carry out more complicated tasks as well as represent intricate connections between the input and output. A MLP model is trained by comparing its output to the expected output and propagating errors back through the network to alter the weights. This process is known as backpropagation. (Haykin et al., 2014)

#### 2.1.3.5.2 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a family of deep learning algorithms that are mostly utilized for processing input that has a grid structure, like images. (Yamashita et al., 2018)



**Figure 2.8:** Convolutional Neural Networks

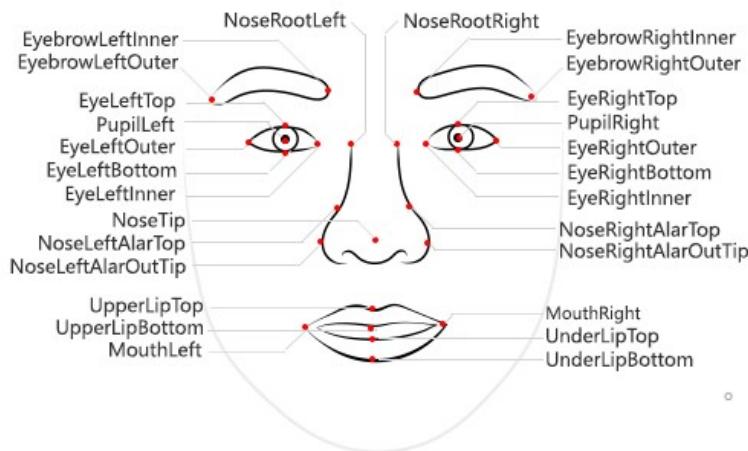
CNNs are designed to adaptively learn spatial hierarchies of features from the data. This learning process includes convolution layers, pooling layers, flatten layer, and fully connected layers. By applying filters, also known as kernels, to the input, these layers carry out the convolution process and produce feature maps. Local elements like textures and edges are captured throughout this procedure. Pooling layers, which come after convolutional layers, help to reduce the number of parameters and computation in the network by reducing the spatial dimensions (width and height) of the input volume.

The features from the input image are retrieved by the convolutional and pooling layers, and the next stage is to categorize the features which is done in flatten layer. The feature maps are converted into a one-dimensional vector in the flatten layer, which is necessary for fully connected layers. The flattened vector is then fed into the fully connected layers (which resemble the standard neural network layers with fully connected nodes) for the classification task. These fully connected layers divide the image into discrete groups based on the high-level characteristics found in the preceding levels.

The last layer of layers in the network architecture play the crucial job of generating the output. The output layer typically uses a softmax activation function in multi-class classification settings to translate the network's raw output into probabilities given to each class. The output node with the highest probability is then chosen to determine the anticipated class.

## 2.2. FACIAL EMOTION RECOGNITION

Human emotions can be inferred from facial expressions. Deciphering these signs of emotion has become a popular research topic in the fields of Human Computer Interaction and Psychology. (Vemou et al., 2021) The development of Facial Emotion Recognition (FER) technology has been significantly aided by technological advancements, particularly with the introduction of ML and Pattern Recognition.



**Figure 2.9:** Facial Landmarks  
Image from (Farley et al., 2023)

FER is a broad field that intersects with Computer Science, AI, Psychology and other fields. It involves analyzing a person's facial expressions in still images and videos in order to determine their emotional state. A three-step approach is used in the methodology: face detection, facial expression identification, and categorization of the expression into a certain emotional state. Facial landmark (Figure 2.9) detection and analysis of changes in their positions are key components of this complex process. FER attempts to offer insights into people's emotional experiences by identifying muscular contractions linked to various emotions from visual clues found in facial expressions.

As stated by Tarnowski et al. (2017), the creative feature extraction from facial expressions using coefficients that detailed aspects of emotional states is what makes the research successful. They distinguished between seven different emotional states: happiness, sorrow, surprise, wrath, fear, and contempt, as well as the more subdued

Emotions	neutral	joy	surprise	anger	sadness	fear	disgust
neutral	881	125	2	155	340	48	101
joy	106	922	7	238	115	1	154
surprise	13	1	1135	8	8	390	13
anger	130	151	6	862	81	2	120
sadness	229	130	33	101	823	104	88
fear	41	0	220	5	88	871	4
disgust	76	147	73	107	21	60	996

(a) Confusion Matrix for k-NN Classifier

Emotions	neutral	joy	surprise	anger	sadness	fear	disgust
neutral	1160	75	9	29	644	61	41
joy	81	1178	0	57	141	0	129
surprise	3	0	1153	4	2	426	10
anger	58	137	0	1346	44	0	88
sadness	122	13	5	0	561	75	2
fear	8	1	308	1	74	910	2
disgust	44	72	1	39	10	4	1204

(b) Confusion Matrix for MLP Classifier

Tables from (Tarnowski et al., 2017)

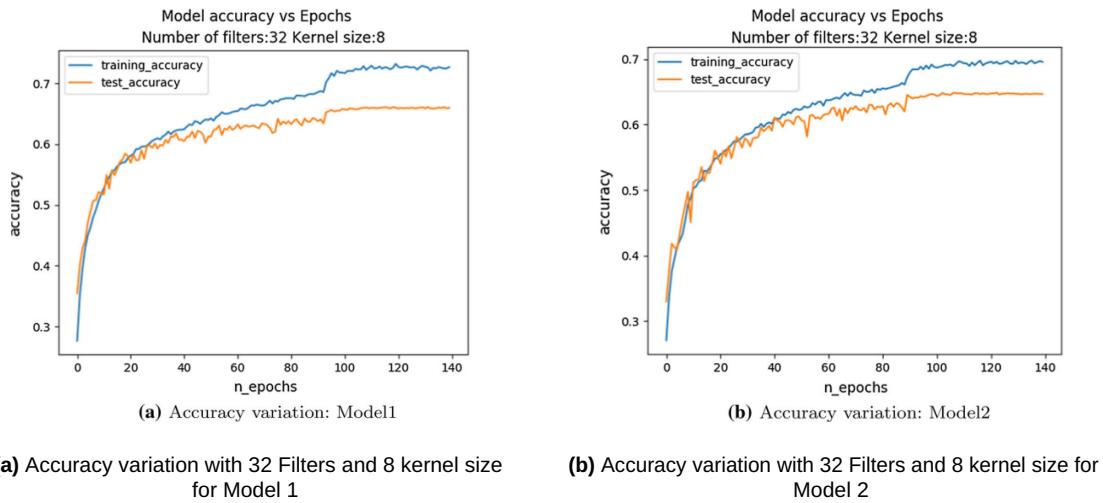
displays of neutrality. For feature computation, they employed a three-dimensional facial model as an alternative to conventional two-dimensional approaches. This enables them to collect more detailed and subtle data, which may improve the accuracy of identifying emotions. On top of that, they used the MLP neural network and the k-NN classifier to classify each emotional state. According to their research and comparative analysis, the MLP is more accurate than the k-NN in classifying emotional states, with a 73% classification accuracy compared to a 63% accuracy for k-NN (Tarnowski et al., 2017).

Mellouk et al. (2020) showed in their study that FER may be achieved with great accuracy and effectiveness by utilizing deep learning techniques. They provided a thorough analysis of multiple FER databases, emphasizing their diversity in terms of picture, video content, lightning circumstances, and demographic variances—all of which are important determinants of FER performance—in order to guarantee the credibility of the results.

Traditional facial recognition techniques included manually defining and extracting features from facial photos, a procedure that was frequently less flexible and efficient. Examples of these techniques include Local Binary Patterns (LBP), Facial Action Coding System (FACS), Local Directional Patterns (LDA), and Gabor wavelet. CNNs and LSTMs can automatically extract and learn complex patterns from facial data, according to Mellouk et al., which will improve the reliability and accuracy of emotion recognition. Furthermore, the difficulties that traditional approaches faced—such as variances in facial characteristics due to diverse demographics, occlusions, and data diversity—are resolved with the use of deep learning, increasing their versatility and effectiveness. The study also examines preprocessing methods including image scaling, cropping, normalization, and data augmentation that are crucial for improving

the accuracy of these deep learning models.

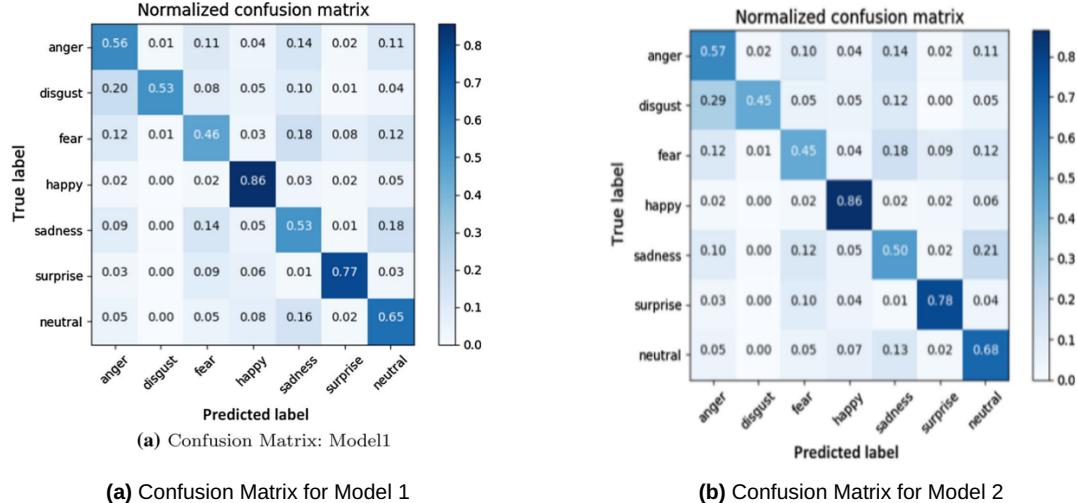
With the help of deep learning and the preprocessing techniques they found, the outcome demonstrates proficiency in accurately classifying the fundamental emotions, with some models reaching over 90% accuracy under specific circumstances (Mellouk and Handouzi, 2020). It indicates that as machines improve at deciphering human emotions, interactions between humans and machines may become more intuitive and natural.



Graphs from (Agrawal and Mittal, 2019)

Based on the findings of Agrawal et al.'s work, the kernel size and the number of filters significantly impact CNNs accuracy. Using the FER-2013 dataset as their primary emphasis, two CNN architectures are put forth after a thorough analysis of various kernel sizes and filter counts. To find the optimal set of parameters that could yield the best convergence and accuracy, Agrawal et al. ran tests with the combination of 6 different kernel sizes (2, 4, 8, ..., 64) and 8 different number of filters (2, 4, 8, ..., 256). They discovered that when network depth increased, a network with 32 filters and an 8 kernel size demonstrated a discernible gain in accuracy (Figure 2.11a, Figure 2.11b).

Even while Model 2 is simpler due to its constant kernel size, lack of dropout layers, and fully connected layers, it was nevertheless able to achieve an accuracy of 65% on the FER-2013 dataset, which is comparable to human performance (Agrawal and Mittal, 2019). Furthermore, in comparison with other emotions, the proposed models were



Graphs from (Agrawal and Mittal, 2019)

able to categorize happiness and surprise with a higher degree of accuracy, which is consistent with people's challenges in picking out distinct emotions (Figure 2.12a, Figure 2.12b).

Gestures	Recognition Accuracy
Neutral	99.00%
Smile	98.50%
Anger	98.50%
Scream	99.50%
Overall	98.88%

**Figure 2.13:** Linear Regression Classification Accuracies Table  
Table from (Naseem et al., 2010)

A pivotal study by Naseem et al. (2010) incorporates the analysis of facial expressions, recognizing them as crucial variations in appearance induced by internal emotions or social communications. In order to evaluate their LR Classification approach, they therefore took into account occlusion modes, brightness changes, and expressions such as scream, smile, rage, and neutral. Notably, the LR Classification algorithm showed an excellent recognition accuracy for all facial expressions tested, averaging 98.88% in a 100D feature space (Figure 2.13). For the screaming expression, the algorithm outperformed other accuracies by achieving an accuracy of 99.5% (Figure 2.13). This great accuracy demonstrates the reliability and efficacy of the LRC approach in handling a wide range of facial emotions.

	Facial Expression				Total
	Anger	Happiness	Sadness	Surprise	
Training (70%)	31	48	19	58	156
Testing (30%)	14	21	9	25	69
Total	45	69	28	83	225
Success	79%	95%	89%	96%	90%

**Figure 2.14:** Random Forest Classification Accuracy  
Table from (Munasinghe, 2018)

According to Munasinghe (2018), RF Classifier are capable of handling facial expression variability well and without overfitting. Also, the researcher asserts that facial landmarks (Figure 2.9) provide an accurate feature extraction capability that capture subtle changes in facial emotions. A facial feature vector obtained from these landmarks and normalized to reduce variance in face size is used to discern emotions with a RF Classifier. With the aid of feature vector, the RF Classifier achieved an average success rate of 90% in classifying four different emotions: anger, happiness, sadness, and surprise (Figure 2.14).

	happy	surprise	fear	angry	sad	disguise
h	144	7	5	10	11	23
s	21	147	6	10	1	15
f	1	5	143	8	19	24
a	11	2	21	132	26	8
s	6	14	18	27	119	16
d	12	13	24	21	6	124

**Figure 2.15:** Support Vector Machine Accuracy  
Table from (Xia, 2014)

Li Xia (2014) presents a unique method of facial emotion detection that employs multi-classification SVM. The study proposes a two-on-two classification method which is an innovative approach to overcome the limitations of traditional classification methods like one-against-one (classifier is trained for each pair of classes) and one-against-the-rest (classifier is trained against all other classes combined). With this novel method, the classification process is faster with fewer sub-classifiers and reduced classification errors. The results of this study were impressive, showing the classifier in this investigation demonstrated a high average recognition rate of 92.7% when six distinct emotions were considered, including happiness, surprise, anger, fear, disgust,

and sadness.

### **2.3. MUSIC RECOMMENDATION BASED ON FER**

From Chakrapani et al.'s approach, music recommendation system with deep learning algorithm could enhance the listening experience by accurately detect and interpret the user's emotions. This is achieved by using CNNs to analyze the user's age, gender, and facial emotion. Based on these data, the system would cater to the user's preferences and present mood. To ascertain the user's emotional state, they used the webcam to take pictures of the user and then processed the image using the CNNs models. The system then provided tailored music recommendations based on the predictions generated by the CNNs models. This approach offers a creative and user-centric alternative for music selection based on emotional cues while streamlining playlist construction and management.(S et al., 2023)

Additionally, Athavle et al. discovered that using CNNs model helps a music recommendation system to accurately detect emotions and subsequently recommend music that aligns with the user's mood. They train a CNNs model for emotion detection in their work. While maintaining great precision, this method lowers total system costs and computing time. The system uses real-time emotion detection to work, and then sends the data to the CNNs model to classify the user's emotions. An appropriate playlist will be recommended as soon as the technology determines the user's current feeling, making the user experience engaging and responsive. In order to guarantee optimal classification accuracy and efficacy, they employed categorial cross-entropy as a loss function to manage missing and anomalous values inside the FER2013 dataset. Despite their result being less accurate than Chakrapani et al.'s work (71%), it nevertheless shows that the model is effective and trustworthy in identifying emotions from facial expressions. (Athavle, 2021)

# **3. REQUIREMENTS**

## **3.1. FUNCTIONAL REQUIREMENTS AND NON-FUNCTIONAL REQUIREMENTS**

### **3.1.1. Functional Requirements**

<b>Req. No.</b>	<b>Categories</b>	<b>Requirements</b>	<b>Priority</b>
FR1	User Registration and Account Management	The system must allow user to register by providing a unique username, user's actual name, date of birth, email, and password.	High
FR2		The system must verify user accounts through an email verification process.	High
FR3		Users must be able to login with their email or username and password. A "Remember Me" option should allow users to stay logged in for 7 days.	High
FR4		Users can access a settings page to update their name, date of birth, email, password, and profile picture. Usernames cannot be changed.	Medium
FR5		Users must be able to reset their passwords through a password reset feature on the login page.	Medium

Req. No.	Categories	Requirements	Priority
FR6	Facial Emotion Recognition	The application integrates a machine learning model to recognize user's facial emotions via their device's camera.	High
FR7	Spotify Web Playback Integration	The system integrates with Spotify Web Playback SDK to play music within the web application.	High
FR8		The application must allow users to connect their Spotify account before accessing music playback services. This integration should facilitate authentication and authorization seamlessly within the web application.	High
FR9	Music Recommendation System	The application must generate playlists based on the user's recognized emotion using an algorithm.	High
FR10	User Interface and Experience	The web application supports a toggle between light and dark themes, automatically detecting and applying the user's device theme upon first use.	Medium
FR11		The application supports multiple languages: English, Japanese, Chinese, Korean, and Malay.	Low

**Table 3.1:** Functional Requirements

### 3.1.2. Non-Functional Requirements

Req. No.	Categories	Requirements	Priority
NFR1	Performance and Scalability	The application shall load within 3 seconds for 95% of its users under standard network conditions.	High
NFR2		The system must be scalable to support up to 100 concurrent users without significant degradation in performance.	High
NFR3	Compliance and Security	All user data, including passwords and personal information, must be encrypted.	High
NFR4		The application must implement secure authentication mechanisms to prevent unauthorized access.	High
NFR5		User data must be stored in a secure database with access strictly limited to the backend server. The database shall not be directly accessible from any public network (0.0.0.0/0).	High
NFR6		User passwords must be encrypted using a secure hashing algorithm (e.g., bcrypt) to ensure their safety even in the event of a data breach.	High
NFR7		All forms of data transmission involved in user authentication and registration must be over HTTPS, and sensitive information shall not appear in URLs or any part of the HTTP request visible to the client side.	High

<b>Req. No.</b>	<b>Categories</b>	<b>Requirements</b>	<b>Priority</b>
NFR8		The application must comply with relevant data protection and privacy regulations, including GDPR where applicable, ensuring user's rights to privacy and data security are upheld.	High
NFR9	Usability	The application shall be designed with a user-friendly interface, ensuring ease of navigation and accessibility.	Medium
NFR10		User input fields should provide immediate feedback to correct errors or invalid data.	Medium
NFR11	Compatibility and Interoperability	The web application must be compatible with the latest versions of Chrome, Firefox, Safari and Edge browsers.	High
NFR12		The system must ensure seamless integration with the Spotify API and maintain compatibility with Spotify's update.	High
NFR13	Localization and Internationalization	The application must support multi-language interfaces, allowing users to switch languages easily.	Medium
NFR14		Date and time formats should adapt to the user's selected language and region preferences.	Low
NFR15	Maintenance and Support	The system should be designed to allow easy updates and maintenance without significant downtime.	Medium

<b>Req. No.</b>	<b>Categories</b>	<b>Requirements</b>	<b>Priority</b>
NFR16		Documentation must be provided for end-users and developers, detailing usage, integration features, and troubleshooting steps.	Medium
NFR17	Application Performance	The facial emotion recognition feature must provide a response within 5 seconds from the time of user's request under standard network conditions.	High
NFR18		The system should ensure a Spotify playback start time of less than 3 seconds after user selection or playlist generation.	High
NFR19		The web application's overall time to interactive (TTI) should not exceed 5 seconds for 90% of its users under standard network conditions.	High
NFR20	User Interface Design	The application must adhere to WCAG 2.1 AA standards for color contrast, navigability, and text size to ensure accessibility for users with disabilities.	High
NFR21		All user interface components (buttons, links, form elements) must be navigable using a keyboard in a logical order to support users with mobility or visual impairments.	High

<b>Req. No.</b>	<b>Categories</b>	<b>Requirements</b>	<b>Priority</b>
NFR22	Data Handling and Authentication	Implement OAuth 2.0 for secure authentication with Spotify, ensuring that user credentials are handled safely and in line with best security practices.	High
NFR23		Apply secure session management practices, including the generation of unique session tokens for users during login and their secure storage on the client side.	High

**Table 3.2:** Non-Functional Requirements

# **4. METHODOLOGY**

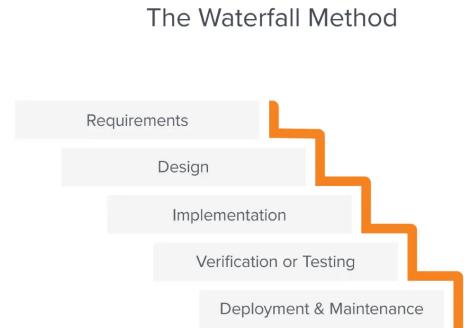
## **4.1. INTRODUCTION**

To effectively navigate the intricacies of software development and guarantee the project's success, choosing an appropriate approach is essential. The concepts, procedures, and practices that guide a project's development, implementation, and completion are collectively referred to as its methodology. The variety of alternative methods, each with specific advantages and applicability to various project types, means that selecting a methodology should be done with careful consideration. This section explores the reasoning behind the choice of an Agile-based methodology, with a particular emphasis on the Kanban methodology, made by the use of Notion for project management. This decision was driven by the project's requirement for flexibility, continuous improvement, and a visual workflow management system. The following sections will outline the comparative comparison between various approaches, along with the reasoning behind choosing Agile due to its alignment with the project's objectives and task specifications.

## **4.2. RESEARCH METHODOLOGY**

### **4.2.1. Waterfall methodology**

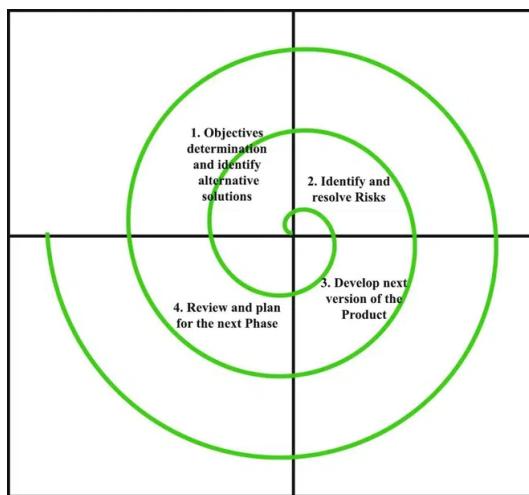
Waterfall methodology, with its linear and sequential approach, stands as a traditional yet relevant framework for software development projects that require a clear, phased progression. Requirements, design, implementation, testing, development and maintenance are the steps in this process. They guarantee that one phase must be finished before moving on to the next, which makes them especially appropriate for projects with stable, well-defined requirements that are unlikely to change. Crespo-Santiago & de la Cruz Dávila-Cosme (Crespo-Santiago and Dávila-Cosme,



**Figure 4.1:** Waterfall Methodology (Communication Team, 2022)

2022) highlight the Waterfall methodology helps maintain the scope of their library project within the requirements, establishing cost and time control, and documenting evidence of project governance.

#### 4.2.2. Spiral methodology



**Figure 4.2:** Spiral Methodology (Kumar Pal, 2018)

Spiral methodology, an evolutionary software development process introduced by Barry Boehm in 1986, is a model for process flexibility and risk control in software development. (Boehm, 1986) The methodology is distinguished by its four-phase cycle approach, which includes planning, risk analysis, implementation, evaluation. This allows for ongoing iterations that involve setting project goals, identifying potential risks, carrying out development, and incorporating stakeholder feedback. Its iterative design ensures a flexible and adaptive development process by permitting incremental product

improvements based on changing needs and stakeholder input. The Spiral methodology provides a methodical approach to managing the complexities and uncertainties inherent in software development. It shines in contexts where project needs are ambiguous or subject to change because of its heavy emphasis on risk management.

#### 4.2.3. Agile methodology



Figure 4.3: Agile Methodology (Laoyan, 2022)

Agile methodology, an approach originated from the Agile Manifesto, published in 2001 by a group of software developers, prioritizes adaptability, collaboration, customer satisfaction and timely delivery of high-quality software. While implementing Agile methodology, project is broken into small, manageable pieces, known as iterations or sprints. Each sprint involves cross-functional teams working on various aspects like planning, design, coding, and testing, with a working iteration of the product delivered at the end of each cycle. This methodology works effectively for projects whose requirements are changing or unclear since it allows ongoing feedback and adjustment.

### 4.3. COMPARISON AND SELECTION

Software development can be approached differently using the Agile, Waterfall, and Spiral techniques, each with its own set of benefits and difficulties. From Table 4.1, Agile is highly flexible and adaptable, ideal for projects with evolving requirements, but may lead to unpredictable costs. Waterfall is straightforward and orderly, perfect for projects with well-defined requirements, but inflexible to changes. Spiral combines iterative development with focusing on risk management, but it could be costly. The project's scale and the clarity of its needs determine which technique is best: Waterfall for its

**Table 4.1:** Comparison of Methodologies

Aspect	Agile	Waterfall	Spiral
<b>Pros</b>	<ul style="list-style-type: none"> <li>• High flexibility and adaptability to changes.</li> <li>• Frequent releases and feedback.</li> <li>• Enhanced customer satisfaction.</li> <li>• Reduced time to market.</li> </ul>	<ul style="list-style-type: none"> <li>• Simple and easy to understand and use.</li> <li>• Clear project milestones and deliverables.</li> <li>• Well-suited for projects with defined requirements.</li> </ul>	<ul style="list-style-type: none"> <li>• Focus on risk management.</li> <li>• Flexibility in design and development.</li> <li>• Suitable for large, complex projects with uncertain risks.</li> </ul>
<b>Cons</b>	<ul style="list-style-type: none"> <li>• Less predictable budget and timeline.</li> <li>• Requires close collaboration and customer involvement.</li> <li>• Not ideal for low-change projects or those with fixed requirements.</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to incorporate changes once the project has started.</li> <li>• Potential for late discovery of problems or errors.</li> <li>• Not suitable for projects where requirements may evolve.</li> </ul>	<ul style="list-style-type: none"> <li>• Can be complex and costly to implement.</li> <li>• Requires significant risk assessment expertise.</li> <li>• May lead to prolonged project duration due to iterative nature.</li> </ul>

structure, Agile for its flexibility, or Spiral for its risk emphasis.

#### 4.4. JUSTIFICATION FOR CHOOSING AGILE

The Agile methodology, particularly the Kanban variant, was chosen in order to satisfy the demand for an adaptable, graphical, and iterative developments process. Given the dynamic nature of the project and its ever-changing objectives, Kanban, which places a strong focus on continuous delivery and workflow efficiency, is an excellent fit.

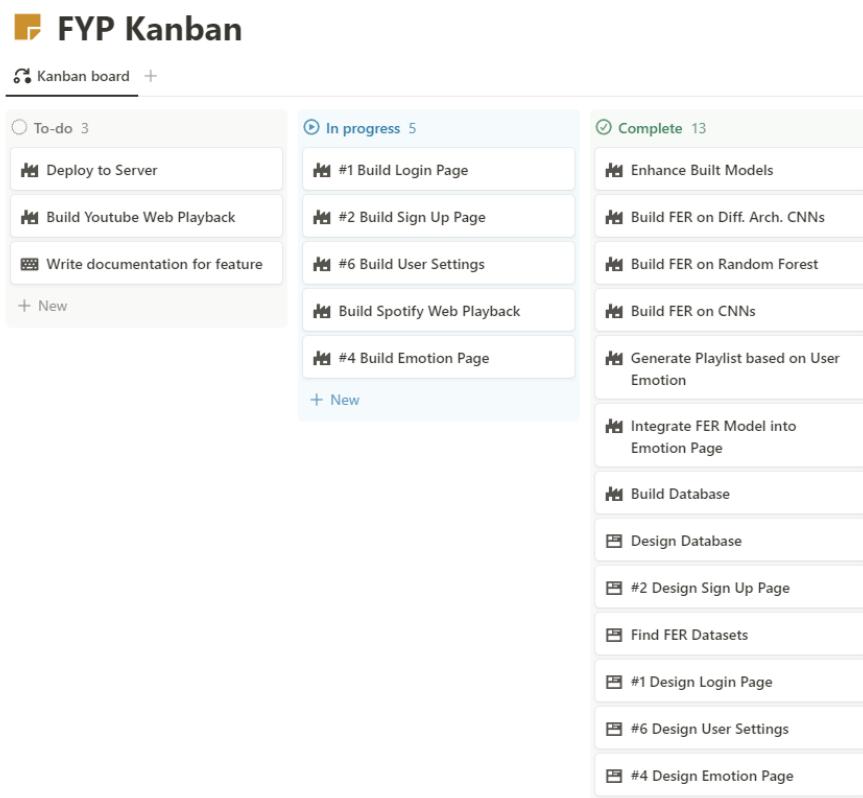


Figure 4.4: Kanban from Notion

In this project, Kanban is implementing using Notion. It gives tasks a visual representation, making it easier to organize and keep track of them as they go through various stages of development. Also, with the adaptability of Notion's platform, project plans could be updated easily which is crucial for keeping the plan responsive to changing project dynamics.

Kanban method's inherent simplicity and its focus on delivering work just-in-time are particularly beneficial for projects demanding flexibility and time efficiency. Kanban with Notion provides a clear picture of the project's progress, enabling developer to identify

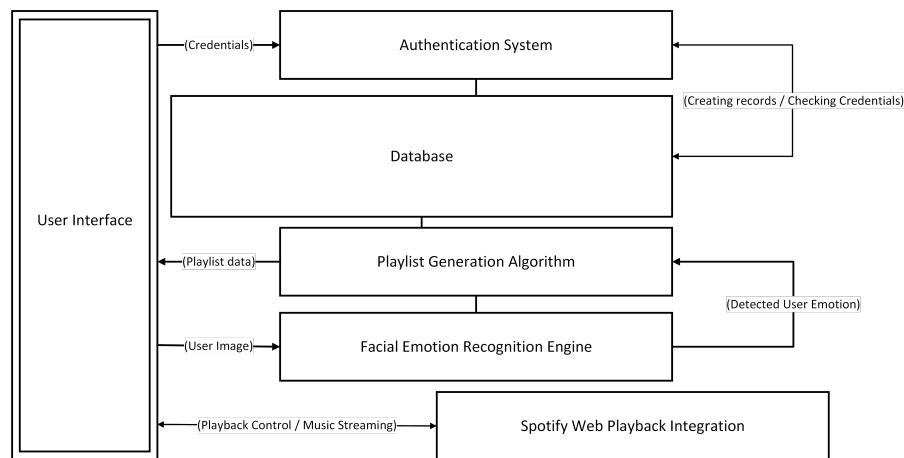
and resolve bottlenecks quickly, and efficiently manage task prioritization. Therefore, the combination of Notion's features with Agile Kanban creates a strong foundation for project management by fusing Kanban's visual clarity and streamlined efficiency with Agile's flexibility.

# 5. DESIGN

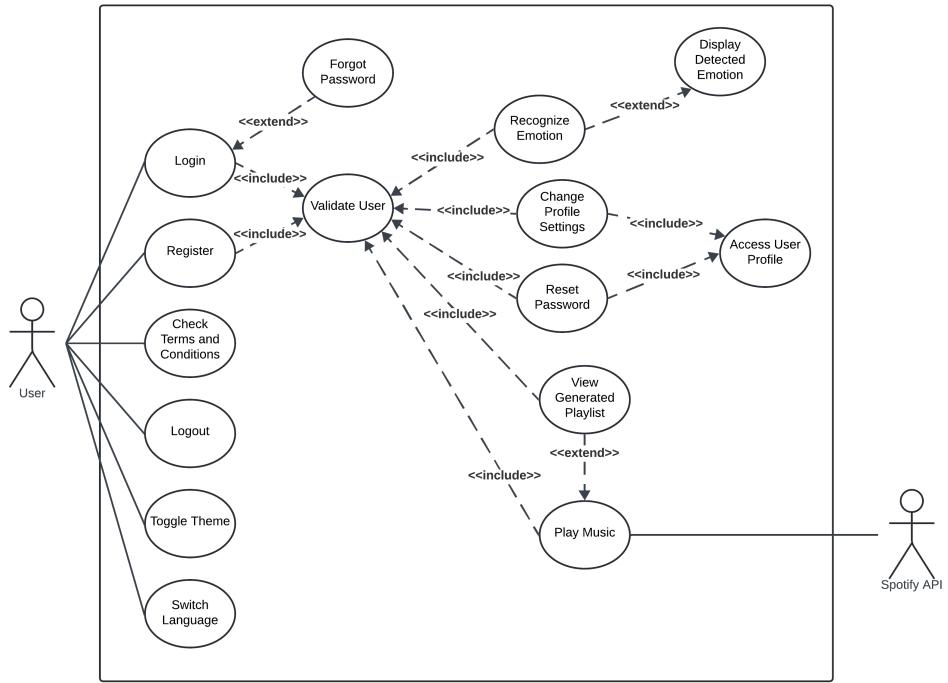
## 5.1. INTRODUCTION

## 5.2. UML DIAGRAMS

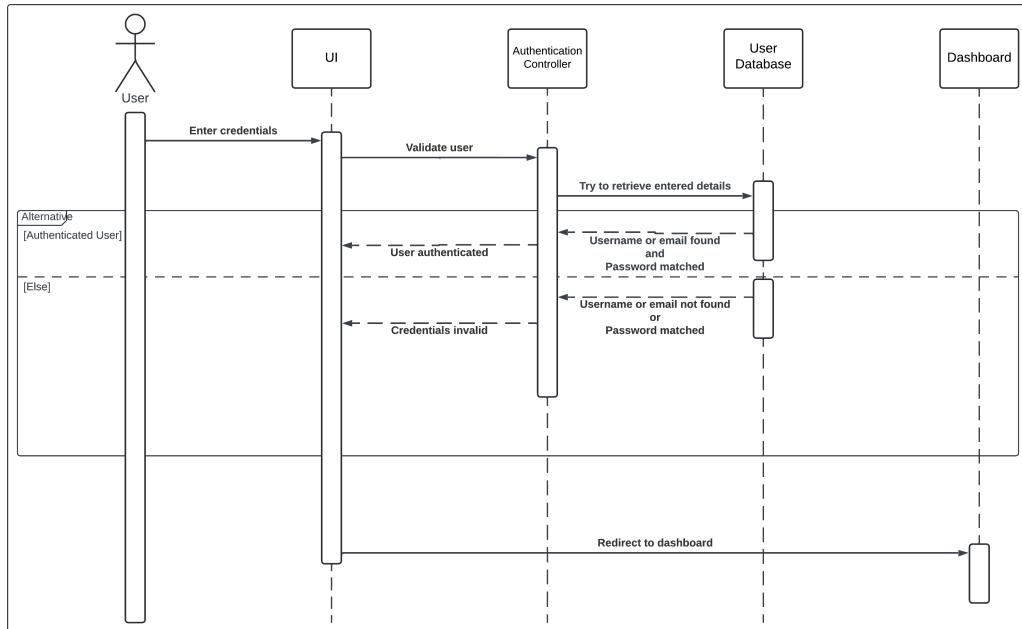
### 5.2.1. Block Diagram



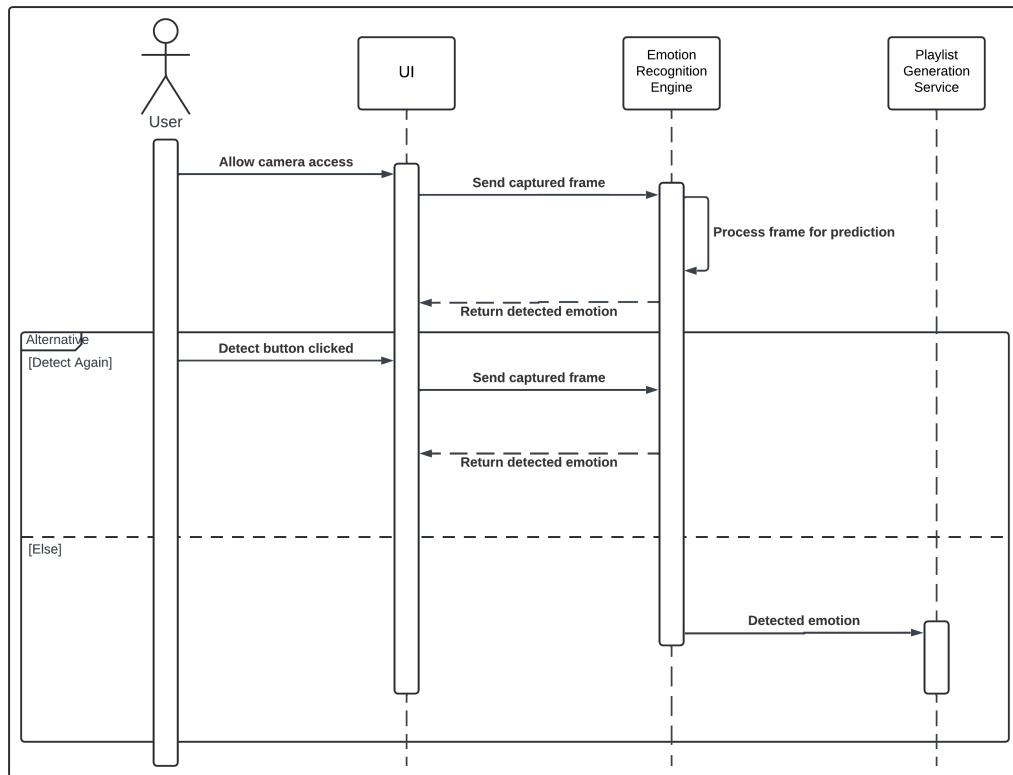
**Figure 5.1:** Block Diagram



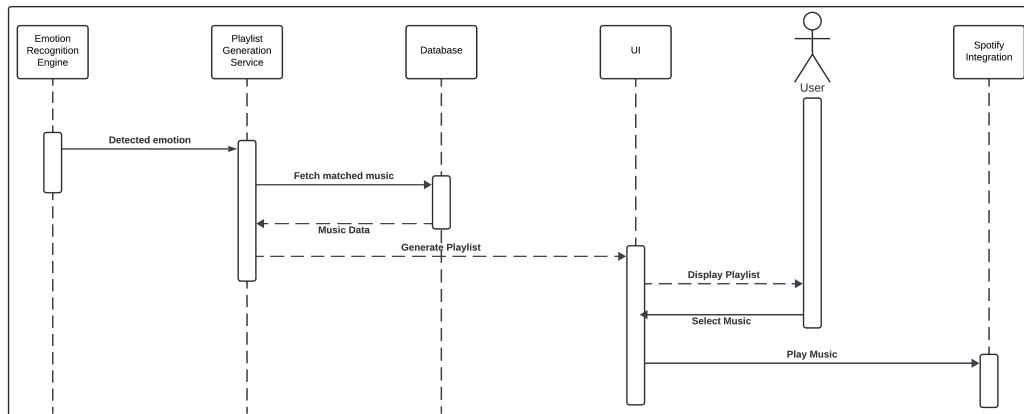
**Figure 5.2:** Use Case Diagram



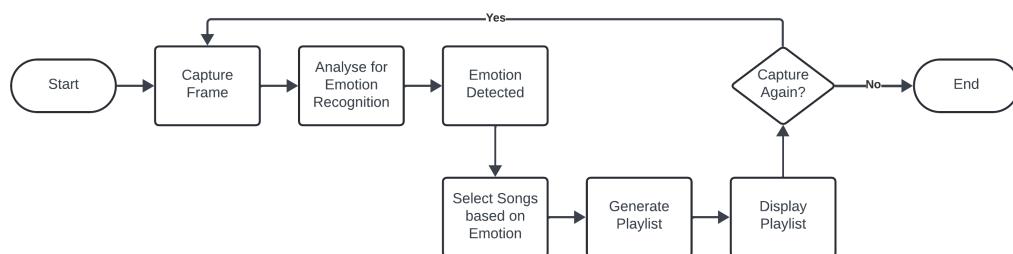
**Figure 5.3:** Login Sequence Diagram



**Figure 5.4:** Emotion Recognition Sequence Diagram



**Figure 5.5:** Playlist Generation Sequence Diagram



**Figure 5.6:** Flowchart

## **5.2.2. Use Case Diagram**

## **5.2.3. Sequence Diagrams**

## **5.2.4. Class Diagram**

## **5.2.5. Flowchart**

## **5.2.6. Entity-relationship Diagram**

## **5.3. LOGO DESIGN**

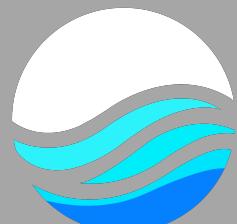


**(a) Logo**



**(b) Banner**

Light Theme Logo



**(a) Logo**



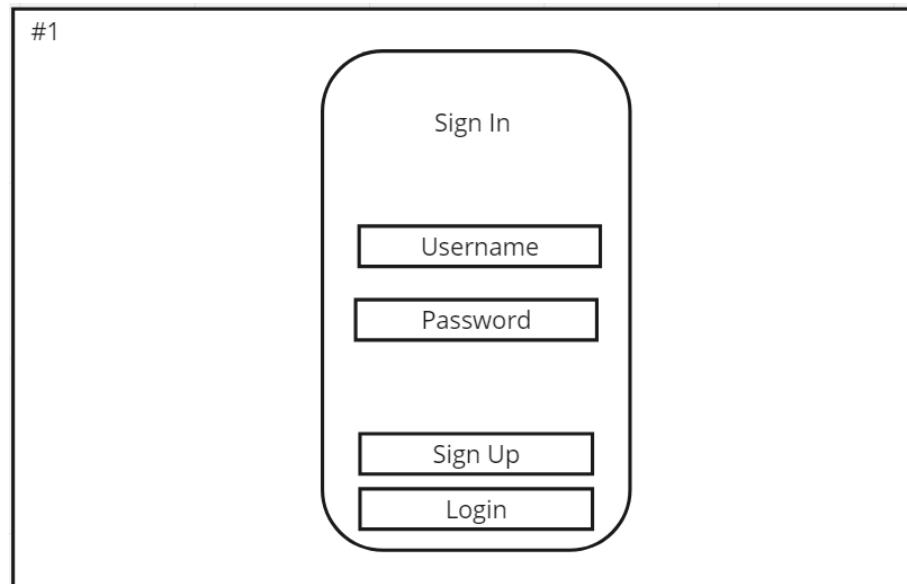
**(b) Banner**

Dark Theme Logo

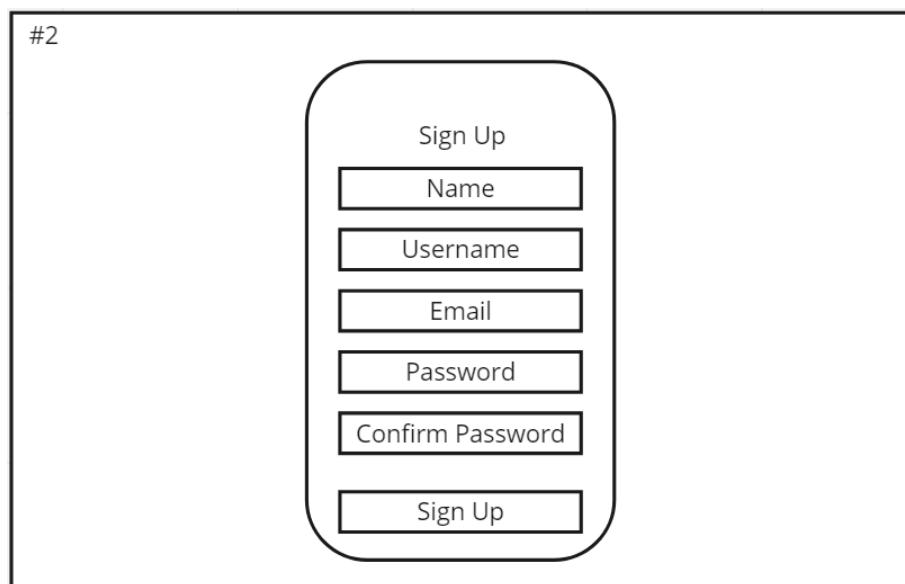
## **5.4. INTERFACE DESIGN**

## **5.5. ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

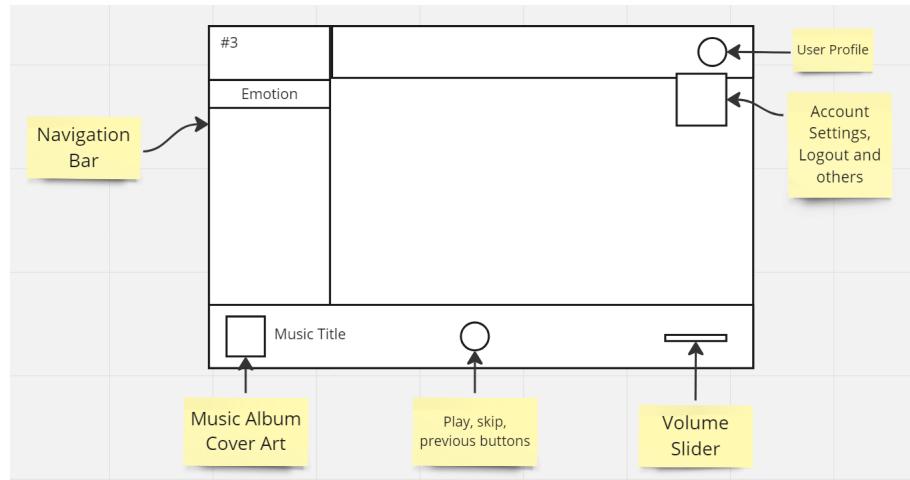
### **5.5.1. Model Architecture**



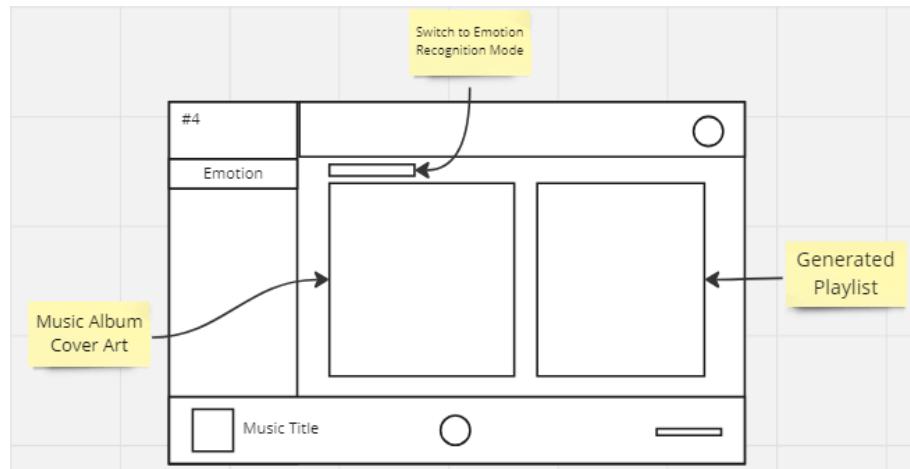
**Figure 5.9:** Login Page



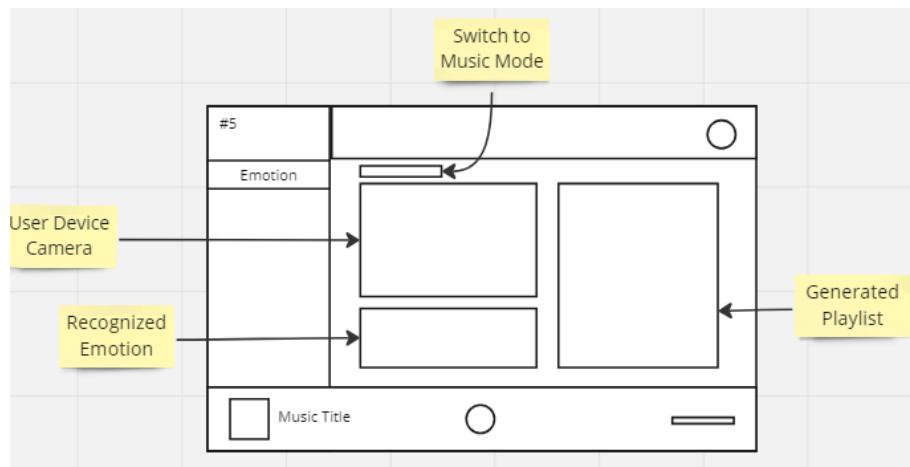
**Figure 5.10:** Sign Up Page



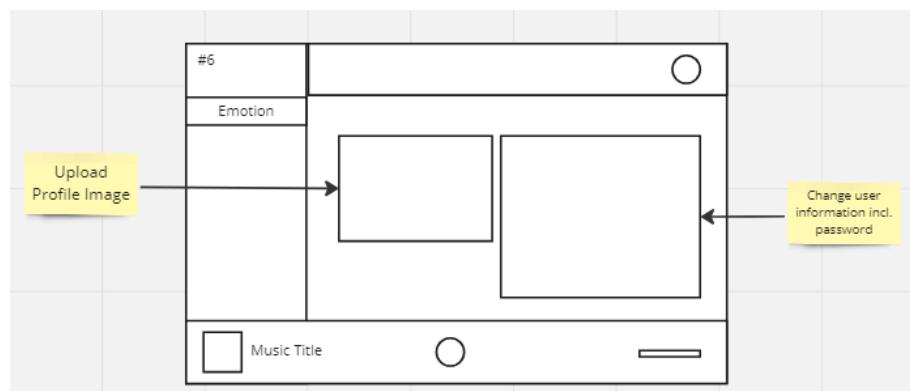
**Figure 5.11:** Dashboard Page



**Figure 5.12:** Emotion Music Page



**Figure 5.13:** Emotion Recognition Page



**Figure 5.14:** User Settings Page

## **6. IMPLEMENTATION**

## **7. PROJECT EVALUATION**

## **8. CONCLUSION AND FUTURE OUTLOOK**

### **8.1. CONCLUSION**

### **8.2. FUTURE OUTLOOK**

# BIBLIOGRAPHY

- , G. (2019), 'Descending into ml: Training and loss | machine learning crash course', <https://developers.google.com/machine-learning/crash-course/descending-into-ml/training-and-loss>.
- , M. (2021), 'Country music guide: History and sounds of country music', <https://www.masterclass.com/articles/country-music-guide>.
- , s.-l. (n.d.), '1.4. support vector machines', [https://scikit-learn.org/stable/modules/svm.html#:~:text=Support%20vector%20machines%20\(SVMs\)%20are](https://scikit-learn.org/stable/modules/svm.html#:~:text=Support%20vector%20machines%20(SVMs)%20are).
- Agrawal, A. and Mittal, N. (2019), 'Using cnn for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy', *The Visual Computer* **36**, 405–412.
- Athavle, M. (2021), 'Music recommendation based on face emotion recognition', *Journal of Informatics Electrical and Electronics Engineering (JIEEE)* **2**, 1–11.
- BBC (2023), 'History of the blues', <https://www.bbc.co.uk/bitesize/articles/zkbh2v4>.
- Beek, M. (2021), 'What is jazz music?', <https://www.classical-music.com/articles/jazz-music-what-it-is-and-how-it-evolved>.
- Berwick, R. (n.d.), 'An idiot's guide to support vector machines (svms) r. berwick, village idiot svms: A new generation of learning algorithms', <https://web.mit.edu/6.034/wwwbob/svm-notes-long-08.pdf>.
- Boehm, B. (1986), 'A spiral model of software development and enhancement', *ACM SIGSOFT Software Engineering Notes* **11**, 22–42.
- Britannica (2019), *country music | Definition, Artists, History, & Facts*.

- Chaudhuri, L. (2022), 'What is blues music?', <https://www.classical-music.com/articles/blues-music>.
- Communication Team, A. (2022), 'Waterfall methodology a complete guide', <https://business.adobe.com/blog/basics/waterfall>.
- Crespo-Santiago, C. A. and Dávila-Cosme, S. d. I. C. (2022), 'Waterfall method: A necessary tool for implementing library projects', *HETS Online Journal* 1(2), 81–92.
- Dunbar, R. I. M. (1998), 'The social brain hypothesis', *Evolutionary Anthropology: Issues, News, and Reviews* 6(5), 178–190.
- E R, S. (2021), 'Random forest | introduction to random forest algorithm', <https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/>.
- Farley, P., Urban, E. and Mehrotra, N. (2023), 'Face detection, attributes, and input data - face - azure ai services', <https://learn.microsoft.com/en-us/azure/ai-services/computer-vision/concept-face-detection>.
- Finkel, I. and Taylor, J. (2021), 'How to write cuneiform', <https://www.britishmuseum.org/blog/how-write-cuneiform>.
- Frith, S. (2018), *Rock | music*.
- Gabler, J. (2013), 'What is classical music?', <https://www.yourclassical.org/story/2013/10/15/what-is-classical-music>.
- Gandhi, R. (2018), 'Introduction to machine learning algorithms: Linear regression', <https://towardsdatascience.com/introduction-to-machine-learning-algorithms-linear-regression-14c4e325882a>.
- GeeksforGeeks (2018), 'K-nearest neighbours - geeksforgeeks', <https://www.geeksforgeeks.org/k-nearest-neighbours/>.
- Géron, A. (2019), *Hands-on machine learning with Scikit-Learn and TensorFlow concepts, tools, and techniques to build intelligent systems*, 2 edn, O'Reilly Media, Inc.

- Harrison, O. (2018), 'Machine learning basics with the k-nearest neighbors algorithm', <https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761>.
- Haykin, S., York, N., San, B., London, F., Sydney, T., Singapore, T., Mexico, M. and Munich, C. (2014), 'Neural networks and learning machines third edition', [https://cours.etsmtl.ca/sys843/REFS/Books/ebook\\_Haykin09.pdf](https://cours.etsmtl.ca/sys843/REFS/Books/ebook_Haykin09.pdf).
- Hiller, L. (2019), *electronic music | Definition, History, & Facts*.
- IBM (2022), 'About linear regression | ibm', <https://www.ibm.com/topics/linear-regression#:~:text=Resources->.
- IBM (2023a), 'What are convolutional neural networks? | ibm', <https://www.ibm.com/topics/convolutional-neural-networks>.
- IBM (2023b), 'What is random forest? | ibm', <https://www.ibm.com/topics/random-forest#:~:text=Random%20forest%20is%20a%20commonly>.
- IBM (n.d.a), 'What is gradient descent? | ibm', <https://www.ibm.com/topics/gradient-descent>.
- IBM (n.d.b), 'What is the k-nearest neighbors algorithm? | ibm', <https://www.ibm.com/topics/knn#:~:text=Next%20steps->.
- Juslin, P. N. and Sloboda, J. A. (2013), 'Music and emotion', *The Psychology of Music* pp. 583–645.
- Killin, A. (2018), 'The origins of music', *Music & Science* 1, 205920431775197.
- Kumar Pal, S. (2018), 'Software engineering | spiral model', <https://www.geeksforgeeks.org/software-engineering-spiral-model/>.
- Kwiatkowski, R. (2021), 'Gradient descent algorithm — a deep dive', <https://towardsdatascience.com/gradient-descent-algorithm-a-deep-dive-cf04e8115f21>.
- Laoyan, S. (2022), 'What is agile methodology? (a beginner's guide)', <https://asana.com/resources/agile-methodology>. Accessed: 2023-04-09.

- Mali, K. (2021), ‘Linear regression | everything you need to know about linear regression’, <https://www.analyticsvidhya.com/blog/2021/10/everything-you-need-to-know-about-linear-regression/>.
- Mandal, M. (2021), ‘Cnn for deep learning | convolutional neural networks (cnn)’, <https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/>.
- McCollum, S. (2019), ‘Hip-hop: A culture of vision and voice’, <https://www.kennedy-center.org/education/resources-for-educators/classroom-resources/media-and-interactives/media/hip-hop/hip-hop-a-culture-of-vision-and-voice/>.
- Mellouk, W. and Handouzi, W. (2020), ‘Facial emotion recognition using deep learning: review and insights’, *Procedia Computer Science* **175**, 689–694.
- Mishra, M. (2020), ‘Convolutional neural networks, explained’, <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>.
- Munasinghe, M. I. N. P. (2018), ‘Facial expression recognition using facial landmarks and random forest classifier’, *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)* .
- Naseem, I., Togneri, R. and Bennamoun, M. (2010), ‘Linear regression for face recognition’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**, 2106–2112.
- Pentreath, R. (2022), ‘Why do we call classical music ‘classical music?’’, <https://www.classicfm.com/discover-music/music-theory/why-do-we-call-it-classical-music/>.
- Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L. and Brattico, E. (2011), ‘Music and emotions in the brain: Familiarity matters’, *PLoS ONE* **6**, e27241.
- Porter, L. and Molana-Allen, L. (2018), ‘Did syria create the world’s first song?’, <https://www.bbc.com/travel/article/20180424-did-syria-create-the-worlds-first-song>.

- S, M. C. D., Iram, S., Bhat, R., Supritha, L. and Leelavathi, S. (2023), 'Music recommendation based on facial emotion recognition', *international journal of advanced research in computer and communication engineering* **12**.
- Saha, S. (2018), 'A comprehensive guide to convolutional neural networks - the eli5 way', <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>.
- Saini, A. (2021), 'Support vector machine(svm): A complete guide for beginners', <https://www.analyticsvidhya.com/blog/2021/10/support-vector-machinessvm-a-complete-guide-for-beginners/>.
- Scikit-learn (2018), 'sklearn.ensemble.randomforestclassifier — scikit-learn 0.20.3 documentation', <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>.
- Srivastava, T. (2019), 'Introduction to knn, k-nearest neighbors : Simplified', <https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/>.
- Tarnowski, P., Kołodziej, M., Majkowski, A. and Rak, R. J. (2017), 'Emotion recognition using facial expressions', *Procedia Computer Science* **108**, 1175–1184.
- Tate, G. and Light, A. (2019), *hip-hop | Definition, History, Culture, & Facts*.
- Vemou, K., Horvath, A. and Zerdick, T. (2021), 'Facial emotion recognition', [https://edps.europa.eu/system/files/2021-05/21-05-26\\_techdispatch-facial-emotion-recognition\\_ref\\_en.pdf](https://edps.europa.eu/system/files/2021-05/21-05-26_techdispatch-facial-emotion-recognition_ref_en.pdf).
- Xia, L. (2014), 'Facial expression recognition based on svm', *2014 7th International Conference on Intelligent Computation Technology and Automation* .
- Yamashita, R., Nishio, M., Do, R. K. G. and Togashi, K. (2018), 'Convolutional neural networks: an overview and application in radiology', *Insights into Imaging* **9**, 611–629.
- Yiu, T. (2019), 'Understanding random forest', <https://towardsdatascience.com/understanding-random-forest-58381e0602d2>.

# **APPENDIX**

## **APPENDIX A**

Lore ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lore ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

## **APPENDIX B**

  Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.