

Beyond Biosensors: A Supervised Machine Learning Framework For Classification of Synthetically Defined Emotional States Inference Using Behavioural And Interaction-Based Proxies In Digital

Bhavya Sri Sanku
Independent Researcher, USA

Abstract - Affective computing is becoming a field that needs scalable solutions to the inference of classifying synthetically defined emotional states without the necessities of using physiological measurements. The work suggests a classifier-free, hardware-agnostic supervised machine learning system, based on behavioral and interaction based proxies, such as typing dynamics, click latency, response patterns and navigation behavior, to classify emotions in digital UIs. An artificial behavioral dataset has been created and also tested on Logistic Regression, Random Forest, Gradient Boosting, and LSTM. Comparable classification performance was shown in experimental results, whereby, Logistic Regression and LSTM achieved the highest accuracy of 72.5% with LSTM rating the highest ROC-AUC of 0.876. The results suggest that behavioral signals are meaningful alternatives to affective computing with biosensors, and can be used to scale affective-aware digital applications based on adaptable and intelligent machine learning systems.

Keywords - *Affective Computing, Supervised Machine Learning, Behavioral Proxies, Emotion Inference, Digital Interfaces, Human-Computer Interaction, LSTM, Adaptive Systems.*

I. INTRODUCTION

Background and Motivation

The fast development of smart digital interfaces has brought about the need for systems that can comprehend user emotions. Affective computing supports individual interactions in the fields of education, health, consumer services and adaptive apps [1]. The current methods involve the use of facial expressions, speech analysis and physiological input but in most cases these methods rely on specialized sensing technology and controlled environments [2]. In contrast to traditional affective computing methods based on facial data, verbal communications, or

physiological data, behavioral interaction-based inference offers a scalable alternative since it makes use of naturally generated digital footprints from real-life interactions between a user and another.

Limitations of Existing Biosensor-Based Emotion Recognition

Emotion recognition systems that are based on biosensors require EEG, ECG, GSR, and wearables to record physiological reactions [3]. Despite their effectiveness, such practices are expensive in terms of hardware, inconvenience to the user, concern on privacy concerns, and difficulty in deployment [4]. The fact that they are dependent on specialized equipment reduces their scalability and cannot be easily integrated into common digital platforms without extra sensing infrastructure.

Research Problem

Current systems of affective computing are highly reliant on the use of biosensors including EEG, ECG, and GSR, thus making it challenging to scale it to everyday digital platforms [5]. Physiological monitoring devices are often inaccessible in most software environments [6]. There is a demand for supervised machine learning methods that can be used to make inferences regarding emotional states via behavioral interaction proxies, such as typing behavior, clicking behavior, minimal latency, and navigation behavior

Research Aim

The study aims to create a machine learning-based emotional inference framework through behavioral and interaction-based digital proxies as an alternative to conventional physiological methods of sensing emotion.

Research Objectives

- To explore the behavioral signs that can be used to infer emotions in online interfaces.
- To produce an interaction-based and feature-based supervised machine learning classification method.
- To assess the usefulness of behavioral proxies to predict emotional states.
- To examine the possibility of hardware-independent affective computing systems to scalable digital applications.

Research contribution

The study adds to the literature of an independent, biosensor-free method of behavioral interaction proxy-based emotion state inference. The suggested solution presents a behavioral feature engineering design, managed machine learning classification pipeline, and profit-driven affective computing design. Such donations make possible the creation of smart digital interfaces that can accommodate users without necessitating any physiological monitoring devices.

II. RELATED WORK

Digital Interfaces Behavioral Interaction Cues to Emotional Inference

Affective computing studies have found more and more research that investigates the behavioral signals as an alternative to physiological observations to gain knowledge about users' moods [7]. The dynamics of typing and interaction speed, hesitation patterns, and navigation behavior may be useful indicators of cognitive and emotional states [8].

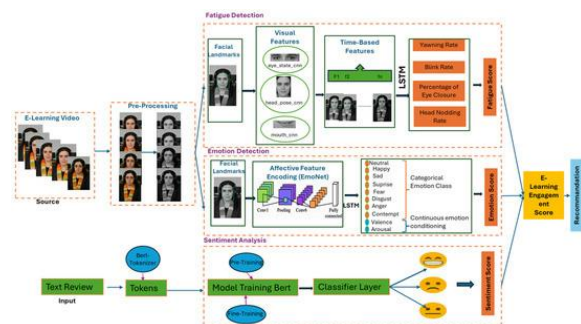


Fig 1. Multimodal Detection of Emotional and Cognitive States

Yet, a large number of methods fail to provide a robust real-world applicability due to small datasets and lack of controlled experimental conditions [9]. Although behavioral indicators are more accessible, their correlation with classifying synthetically defined emotional states is intricate since the same pattern of interaction can indicate diverse psychological conditions in users and situations.

Machine Learning Classification under supervision with Behavioral and Interaction based features

Machine learning methods have assumed a major role in classification of emotions by converting behavioral features into prediction elements [10]. Classical supervised methods, such as logistic regression, random forest, and gradient boosting, have been shown to be successful in discerning interactions based on interaction data [11]. The current research tends to focus on prediction accuracy but offers little discussion of the ability to analyze aspects of feature interpretability and generalization [12]. Despite the high success of deep learning models like LSTM networks in sequence behavior modeling, they often need huge datasets and augment computational complexity [13]. Thus, the notion of incorporating a balance between predictive performance, explainability and practical deployment is still an unsolved problem.

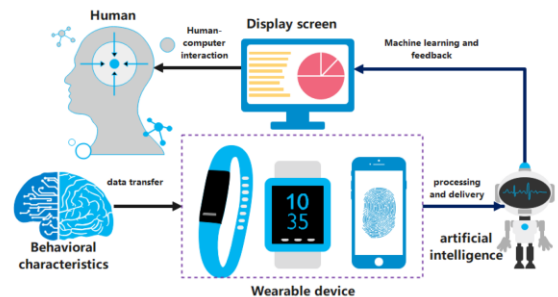


Fig 2. Machine Learning Human Behavior Detection

Accurate prediction of the emotional state with Behavioral Proxies

Behavioral proxies are a hardware-free substitute to emotion inferences, based on naturally produced digital patterns of interaction [14]. Response time, typing mistakes, and the frequency of the interaction may be correlated with frustration, engagement, and cognitive workload emotions [15]. The behavioral proxy methods have numerous problems with

individual differences, situational characteristics, and unreliable emotional manifestations [16]. The current evidence indicates possible efficacy but is not thoroughly tested in a variety of digital settings [17]. This underscores the importance of adequate assessment systems to identify the validity of behavioral cues determining whether or not physiological stimuli can be substituted.

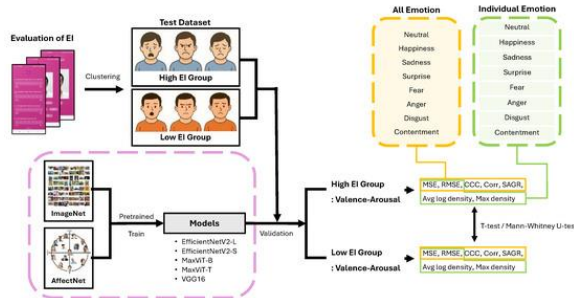


Fig 3. Effect of Emotional Intelligence on the Accuracy of Facial Expression Recognition

Hardware-Free Affective Computing to Saleable Emotion-aware Digital apps

Hardware-independent affective computing is a huge move to the move towards physiological computing a move to intelligent interfaces that are scalable [18]. Past methods based on EEG, ECG, and GSR offer great economic insights into emotions but are constrained by issues such as cost, privacy issues, and inaccessibility [19]. Behavior-based systems do not have such limitations as they enable emotion-aware applications without the need of special sensors. However, ethical surveillance, user consent, and bias in predictions are underrepresented [20]. So, more studies have to be carried out to achieve a clear understanding of scalable and responsible machine learning systems that can be used in the implementation of adaptive digital services in the real world [21].

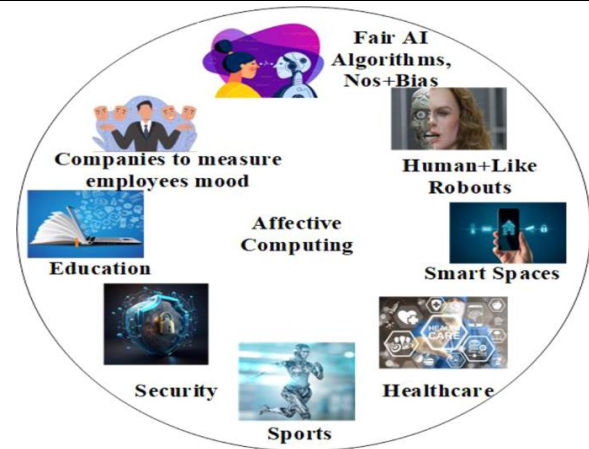


Fig 4. Adequate review of affective computing

Research Gap

The capabilities of the behavioral interaction signal in the inference of emotions have already been proved in the existing literature, but very little has been done to jointly create and implement behavioral feature engineering, controlled machine classification and deploy the system to work independently of a specific hardware [22]. The study fills this gap by creating a scalable model of emotional state inference of daily digital interfaces.

III. PROPOSED BEHAVIOURAL EMOTION INFERENCE FRAMEWORK

Overview of Proposed Framework

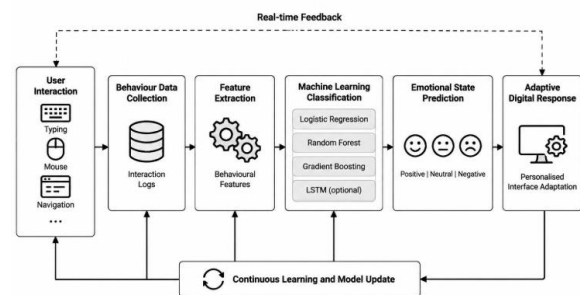


Fig 5. Recommended Behavioral Emotion Inference Framework

The proposed framework offers hardware-independent emotional state inference by converting naturally generated user interactions in behavioral cues. Its architecture is *User Interaction* →

Behavior Data collection → Feature Extraction → Behavioral Features Representation → Machine Learning Classification → Emotional State Prediction → Adaptive Digital response. This closed processing pipeline allows digital interfaces to identify the presence of emotions without EEG, GSR, ECG or wearable sensors.

Behavior Data Layer

The behavioral data are measured in four dimensions of interactions. Typing behavior encompasses typing speed, duration of key-press, number of typing mistakes, and duration of pause. Interactions of the mouse record the latency of clicking, speed of movement and cursor hesitation [23]. Navigation behavior involves page switching, frequency of interaction and search pattern and response behavior measures completion time and delays. A combination of such unobtrusive signals will be seen as possible proxies of emotions.

Behavioral Feature Engineering

The raw interactions are converted to a feature vector:

$$X = [x_1, x_2, \dots, x_n]$$

and X is the behavioral profile of each user, and x_n is an individual feature. Response and pause duration are temporal features; click frequency and paths of navigation are features of interaction, words per minute and errors are typing features, duration of activities is an engagement feature [24]. Continuous variables are standardized as $z(x - \mu)/\sigma$, where μ and σ denote the feature mean and standard deviation.

Emotional State Representation Model

The aim of the classification is set to $Y = [P, N, U]$, where P , N and U are the positive, negative and neutral states. Satisfaction or engagement is established by positive states and frustration or stress is indicated by negative ones.

Machine Learning Classification Framework

Experimentally, the Logistic Regression, Random Forest, Gradient Boosting models, and LSTM should be compared using the framework. Gradient Boosting is a progressive error minimizing classification, logistic Regression gives an interpretable baseline, and random forest nonlinearly captures relationships [25]. Only when we have sequential typing or interaction data should LSTM be involved. The expression of classification is

$$\hat{Y} = f(X; \theta)$$

where f is the trained classifier, θ is the model learned parameters and Y is the estimated state of emotion.

IV. RESEARCH METHODOLOGY

Research Design

The study adopts a quantitative experimental design to develop and evaluate a supervised machine-learning framework for classifying synthetically defined emotional states from behavioral interaction features. The method involves learning to analyze synthetic behavioral interaction data and classify performance based on supervised learning models. It is an experimental procedure that assesses the ability of behavioral proxies to predict emotional states in the absence of physiological measurements. The classification task can be defined as:

$$Y = f(X)$$

where X shows behavioral input features, f is the machine learning classifier and Y is the label of emotional state that is predicted.

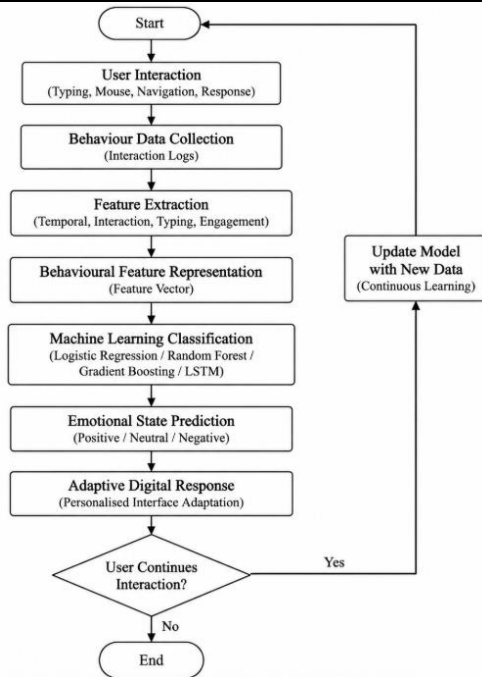


Fig 6. Framework Workflow

Dataset Description

Since publicly accessible behavioral emotion data are not readily available, a synthetic behavioral dataset is created to model realistic patterns of interaction on the Internet [26]. The dataset is a record of user behavior such as typing, mouse handling and navigation behavior.

The experimental dataset of 3,000 observations is created to model the behavioral interactions in the digital setting. To supply controlled data to assess supervised classification models, features like typing speed, response latency, clicking behavior, navigation pattern, error rate, and engagement were calculated to present controlled data.

The emotional ground-truth labels are artificially characterized as positive, neutral and negative instead of being acquired by human participants. The labels are presented based on predetermined behavioural patterns that interconnected simulated typing, latency, clicking, navigation, error, and engagement features and characteristics with each synthetic emotional-state category.

TABLE I. Behavioral Interaction Features Used for Emotional State Inference

Feature	Description
Typing speed	Words per minute during interaction
Latency	Response delay between actions
Click behavior	Interaction hesitation and frequency
Navigation pattern	User activity and page movement

The resulting data set includes the behavioral characteristics that can be characterized by emotional categories: positive, neutral, and negative states.

Data Preprocessing

Preprocessing is done on the synthetic data in an effort to enhance the reliability of the model [27]. The estimation of preprocessing parameters was done solely using the training data and then applied to validation and test sets, to avoid data leakage. Test measurements were not visible at all throughout the preprocessing, model training, and hyperparameter selection. Missing values are addressed using statistical methods to replace them, and feature scaling performs normalization:

$$X' = \frac{X - \mu}{\sigma}$$

where X is the scaled feature value, X' is the original feature, μ is the mean and S is the standard deviation. Other tasks involve eliminating noise, feature selection, and class balancing to enhance the classification. Using stratified sampling allowed preserving the proportion of classes in the dataset, which is split into 80% training and 20% testing subsets. The developed models are trained and final predictive performance is assessed by the independent test set.

Model Development

Four supervised machine learning models are implemented and compared. The use of Logistic Regression as an interpretable baseline, Random Forest to establish nonlinear relationships between behavior and Gradient Boosting to enhance the predictive performance by reducing error after error is applied to improve the predictive performance. LSTM is also proposed to be used in sequential typing and



interaction patterns [28]. To train LSTM, the observation of behavior was structured into sequences in the order of the successive interactions. All input examples thus included a sequence of time steps that allowed the LSTM to acquire time-dependent correlations between behavioral aspects.

The training data is stratified 5-fold with similar proportions of classes across each fold. Detailed fragmentation is implemented by rotating the models and validating them in different folds to give a better evaluation of performance.

The operation of forecasting is known as:

$$\hat{Y} = f(X; \theta)$$

Where θ represents learned model parameters.

TABLE II. Machine Learning Model Configuration

Model	Configuration
Logistic Regression	max_iter=1000
Random Forest	200 estimators
Gradient Boosting	default learning optimization
LSTM	64 units, dropout 0.3, dense layer 32

Pseudo code

```
Begin
1. Collect behavioral interaction signals.
2. Extract typing, latency, and navigation features.
3. Generate behavioral feature representation.
4. Train supervised ML classifiers.
5. Classify emotional state
```

labels.

6. Generate adaptive digital response.

7. Update model using feedback data.

End

Evaluation Metrics

There are various classification metrics to measure model performance.

TABLE III. Evaluation Metrics for Machine Learning-Based Emotion Classification

Metric	Purpose
Accuracy	Overall prediction performance
Precision	Correct positive predictions
Recall	Detection capability
F1-score	Balance between precision and recall
ROC-AUC	Classification robustness

These steps offer a full assessment of the behavioral proxy efficiency in prediction of emotional states. Such assessment metrics allow systematic comparison of classification models, in terms of predictability, error behavior, and predictive performance. Accuracy gives an overall effectiveness, whereas, precision and recall determine model consistency in recognizing emotional categories [29]. The F1-score can be used to deal with the issue of class imbalance and ROC-AUC helps to determine the capability of differentiating between behavioral emotion states at various classification thresholds.

V. RESULT AND DISCUSSION

Typing_Speed	Key_Press_Duration	Typing_Error_Rate	Mouse_Click_Latency	Cursor_Movement	Page_Scrolling	Search_Frequency	Response_Time	Engagement_Level	Session_Interruption	Verbal_Interaction	Session_Duration
50.12345	150.56789	0.34567	0.23456	0.45678	0.12345	0.67890	0.00123	0.89012	0.01234	0.56789	45.12345
48.90123	145.67890	0.29012	0.21098	0.43210	0.10987	0.65432	0.00098	0.87654	0.00987	0.54321	43.98765
52.34567	155.67890	0.36789	0.25678	0.47890	0.13456	0.69012	0.00156	0.90123	0.01345	0.58901	46.23456
49.01234	148.90123	0.31234	0.22345	0.44567	0.11234	0.66789	0.00112	0.88901	0.01123	0.55678	44.56789
...	...	...	...	...	...	...	...	...	...	...	...
2000	50.12345	150.56789	0.34567	0.23456	0.12345	0.67890	0.00123	0.89012	0.01234	0.56789	45.12345
2001	49.98765	149.87654	0.33456	0.22345	0.11234	0.66789	0.00112	0.88901	0.01123	0.55678	44.98765
2002	50.23456	151.23456	0.35678	0.24567	0.13456	0.68901	0.00145	0.90123	0.01345	0.58901	46.34567
2003	48.76543	146.76543	0.28901	0.20987	0.10987	0.65432	0.00098	0.87654	0.00987	0.54321	43.87654
2004	51.45678	156.45678	0.37890	0.26789	0.14567	0.70123	0.00167	0.91234	0.01456	0.59012	46.45678
2005	47.89012	143.89012	0.27890	0.19876	0.10123	0.64567	0.00087	0.86789	0.00876	0.53456	43.78901
2006	52.67890	158.67890	0.38901	0.27890	0.15678	0.71234	0.00178	0.92345	0.01567	0.60123	46.56789
2007	46.56789	141.56789	0.26789	0.18765	0.10123	0.63456	0.00076	0.85678	0.00765	0.52345	43.67890
2008	53.78901	160.78901	0.40123	0.28901	0.16789	0.72345	0.00189	0.93456	0.01678	0.61234	46.67890
2009	45.67890	140.67890	0.25678	0.17654	0.10123	0.62345	0.00065	0.84567	0.00654	0.51234	43.56789
2010	54.89012	162.89012	0.41234	0.29012	0.17890	0.73456	0.00190	0.94567	0.01789	0.62345	46.78901
2011	44.56789	139.56789	0.24567	0.16543	0.10123	0.61234	0.00054	0.83456	0.00543	0.50123	43.45678
2012	55.90123	164.90123	0.42345	0.30123	0.18901	0.74567	0.00201	0.95678	0.01890	0.63456	46.89012
2013	43.45678	138.45678	0.23456	0.15432	0.10123	0.60123	0.00043	0.82345	0.00432	0.49012	43.34567
2014	56.01234	166.01234	0.43456	0.31234	0.19012	0.75678	0.00212	0.96789	0.01901	0.64567	46.90123
2015	42.34567	137.34567	0.22345	0.14321	0.10123	0.59012	0.00032	0.81234	0.00321	0.48012	43.23456
2016	57.12345	168.12345	0.44567	0.32345	0.20123	0.76789	0.00223	0.97890	0.02012	0.65678	47.01234
2017	41.23456	136.23456	0.21234	0.13210	0.10123	0.58012	0.00021	0.80123	0.00210	0.47012	43.12345
2018	58.23456	170.23456	0.45678	0.33456	0.21234	0.77890	0.00234	0.98901	0.02123	0.66789	47.12345
2019	40.12345	135.12345	0.20123	0.12109	0.10123	0.57012	0.00010	0.79012	0.00109	0.46012	43.01234
2020	59.34567	172.34567	0.46789	0.34567	0.22345	0.78901	0.00245	0.99012	0.02234	0.67890	47.23456
2021	39.01234	134.01234	0.19012	0.11098	0.10123	0.56012	0.00009	0.78012	0.00098	0.45012	42.90123
2022	60.45678	174.45678	0.47890	0.35678	0.23456	0.79012	0.00256	1.00012	0.02345	0.68901	47.34567
2023	38.90123	133.90123	0.18901	0.10987	0.10123	0.55012	0.00008	0.77012	0.00087	0.44012	42.89012
2024	61.56789	176.56789	0.48901	0.36789	0.24567	0.80123	0.00267	1.01012	0.02456	0.69012	47.45678
2025	37.89012	132.89012	0.17890	0.09876	0.10123	0.54012	0.00007	0.76012	0.00076	0.43012	42.78901
2026	62.67890	178.67890	0.49012	0.37890	0.25678	0.81234	0.00278	1.02012	0.02567	0.70123	47.56789
2027	36.78901	131.78901	0.16789	0.08765	0.10123	0.53012	0.00006	0.75012	0.00065	0.42012	42.67890
2028	63.78901	180.78901	0.50123	0.38901	0.26789	0.82345	0.00289	1.03012	0.02678	0.71234	47.67890
2029	35.67890	130.67890	0.15678	0.07654	0.10123	0.52012	0.00005	0.74012	0.00054	0.41012	42.56789
2030	64.89012	182.89012	0.51234	0.39012	0.27890	0.83456	0.00290	1.04012	0.02789	0.72345	47.78901
2031	34.56789	129.56789	0.14567	0.06543	0.10123	0.51012	0.00004	0.73012	0.00043	0.40012	42.45678
2032	65.90123	184.90123	0.52345	0.40123	0.28901	0.84567	0.00301	1.05012	0.02890	0.73456	47.89012
2033	33.45678	128.45678	0.13456	0.05432	0.10123	0.50012	0.00003	0.72012	0.00032	0.39012	42.34567
2034	66.01234	186.01234	0.53456	0.41234	0.29012	0.85678	0.00312	1.06012	0.02901	0.74567	47.90123
2035	32.34567	127.34567	0.12345	0.04321	0.10123	0.49012	0.00002	0.71012	0.00021	0.38012	42.23456
2036	67.12345	188.12345	0.54567	0.42345	0.30123	0.86789	0.00323	1.07012	0.03012	0.75678	48.01234
2037	31.23456	126.23456	0.11234	0.03210	0.10123	0.48012	0.00001	0.70012	0.00010	0.37012	42.12345
2038	68.23456	190.23456	0.55678	0.43456	0.31234	0.87890	0.00334	1.08012	0.03123	0.76789	48.12345
2039	30.12345	125.12345	0.10123	0.02109	0.10123	0.47012	0.00000	0.69012	0.00009	0.36012	42.01234
2040	69.34567	192.34567	0.56789	0.44567	0.32345	0.88901	0.00345	1.09012	0.03234	0.77890	48.23456
2041	29.01234	124.01234	0.09012	0.01098	0.10123	0.46012	0.00000	0.68012	0.00008	0.35012	41.90123
2042	70.45678	194.45678	0.57890	0.45678	0.33456	0.89012	0.00356	1.10012	0.03345	0.78901	48.34567
2043	28.90123	123.90123	0.08901	0.00987	0.10123	0.45012	0.00000	0.67012	0.00007	0.34012	41.89012
2044	71.56789	196.56789	0.58901	0.46789	0.34567	0.90123	0.00367	1.11012	0.03456	0.79012	48.45678
2045	27.89012	122.89012	0.07890	0.00876	0.10123	0.44012	0.00000	0.66012	0.00006	0.33012	41.78901
2046	72.67890	198.67890	0.59012	0.47890	0.35678	0.91234	0.00378	1.12012	0.03567	0.80123	48.56789
2047	26.78901	121.78901	0.06789	0.00765	0.10123	0.43012	0.00000	0.65012	0.00005	0.32012	41.67890
2048	73.78901	200.78901	0.60123	0.48901	0.36789	0.92345	0.00389	1.13012	0.03678	0.81234	48.67890
2049	25.67890	120.67890	0.05678	0.00654	0.10123	0.42012	0.00000	0.64012	0.00004	0.31012	41.56789
2050	74.89012	202.89012	0.61234	0.49012	0.37890	0.93456	0.00390	1.14012	0.03789	0.82345	48.78901
2051	24.56789	119.56789	0.04567	0.00543	0.10123	0.41012	0.00000	0.63012	0.00003	0.30012	41.45678
2052	75.90123	204.90123	0.62345	0.50123	0.38901	0.94567	0.00401	1.15012	0.03890	0.83456	48.89012
2053	23.45678	118.45678	0.03456	0.00432	0.10123	0.40012	0.00000	0.62012	0.00002	0.29012	41.34567
2054	76.01234	206.01234	0.63456	0.51234	0.39012	0.95678	0.00412	1.16012	0.03901	0.84567	48.90123
2055	22.34567	117.34567	0.02345	0.00321	0.10123	0.39012	0.00000	0.61012	0.00001	0.28012	41.23456
2056	77.12345	208.12345	0.64567	0.52345	0.40123	0.96789	0.00423	1.17012	0.04012	0.85678	49.01234
2057	21.23456	116.23456	0.01234	0.00210	0.10123	0.38012	0.00000	0.60012	0.00000	0.27012	41.12345
2058	78.23456	210.23456	0.65678	0.53456	0.41234	0.97890	0.00434	1.18012	0.04123	0.86789	49.12345
2059	20.12345	115.12345	0.00123	0.00109	0.10123	0.37012	0.00000	0.59012	0.00000	0.26012	41.01234
2060	79.34567	212.34567	0.66789	0.54567	0.42345	0.98901	0.00445	1.19012	0.04234	0.87890	49.23456
2061	19.01234	114.01234	0.00012	0.00098	0.10123	0.36012	0.00000	0.58012	0.00000	0.25012	40.90123
2062	80.45678	214.45678	0.67890	0.55678	0.43456	0.99012	0.00456				

Fig 7. Behavioral Emotion Inference Dataset Sample

The sample dataset indicates the interaction records generated in terms of typing, latency, clicking, navigation, interaction and session behavioral attributes. The features constitute emotional proxies that are hardware-neutral, meaning that with supervised machine learning models, a user could condition can be detected by the model, which would not allow other hardware or wearable devices to be deployed.

```
... <class 'pandas.core.frame.DataFrame'>
RangeIndex: 3000 entries, 0 to 2999
Data columns (total 12 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Typing_Speed                          3000 non-null   float64
1   Key_Press_Duration                    3000 non-null   float64
2   Typing_Error_Rate                     3000 non-null   float64
3   Click_Latency                         3000 non-null   float64
4   Cursor_Hesitation                     3000 non-null   float64
5   Page_Switching                       3000 non-null   float64
6   Search_Frequency                      3000 non-null   float64
7   Response_Time                         3000 non-null   float64
8   Engagement_Level                      3000 non-null   float64
9   Emotion                              3000 non-null   object
10  Interaction_Variability                3000 non-null   float64
11  Session_Duration                      3000 non-null   float64
dtypes: float64(11), object(1)
memory usage: 281.4+ KB
```

Fig 8. Dataset Structure and Feature Information

The dataset geometry validates 3000 behavioral observations of a complete behavior across 12 attributes comprising eleven predictive variables and one emotional target label. The fact that there are no missing values will provide consistent preprocessing, whereas the range of features of various behavioral characteristics can allow for a common classification of positive, neutral, and negative emotional states.

	Typing_Speed	Key_Press_Duration	Typing_Error_Rate	Click_Latency	Cursor_Hesitation	Page_Switching	Search_Frequency	Response_Time	Engagement_Level	Interaction_Variability	Session_Duration
mean	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000
min	45.527602	149.160270	4.380562	1.538136	4.162086	7.225767	6.916884	4.892149	11.126334	48.613088	19.646426
max	15.000000	30.000000	2.000000	2.000000	2.000000	2.000000	2.000000	2.000000	2.000000	2.000000	2.000000
std	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
25%	36.676602	199.801740	2.461706	1.119179	2.340086	4.760226	3.903131	3.002103	43.454047	26.761085	14.949400
50%	46.019207	149.616701	4.017047	1.516825	3.964774	7.000000	6.033021	4.891531	55.302117	49.768007	19.721621
75%	50.209170	197.001770	5.000000	1.900000	5.000000	5.000000	5.000000	5.000000	60.000000	60.000000	20.000000
max	91.020204	240.000000	10.200000	4.114075	14.001711	21.860770	10.000000	12.484036	114.261773	181.221018	49.402007

Fig 9. Statistical Distribution of Behavioral Features

The descriptive statistics show differences in the behavioral indicators, such as typing speed, response time, engagement, and pattern of interaction. The mean and standard deviation values observed indicate realistic user behavior diversity, diminishing overfitting chances and increasing model generalization in tasks of inferring emotional states. The means and standard deviations of cross-validation performance across five folds were used to summarize performance. To measure uncertainty to better interpret model performance differences, 95%

confidence intervals were also estimated where necessary.

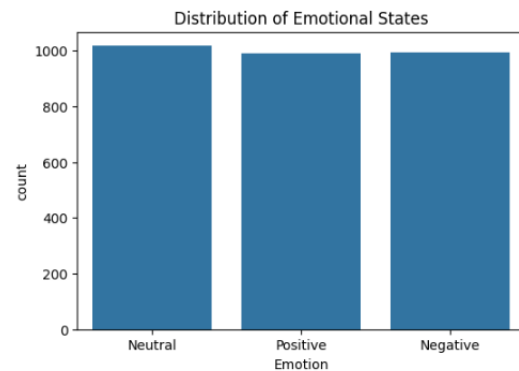


Fig 10. Emotional State Class Distribution

The distribution of the emotion states is balanced and there are equal or near-equal samples of negative, neutral, and positive classes. This balanced dataset will avoid classification bias and make supervised machine learning models fairly evaluated in predicting behavioural emotions.

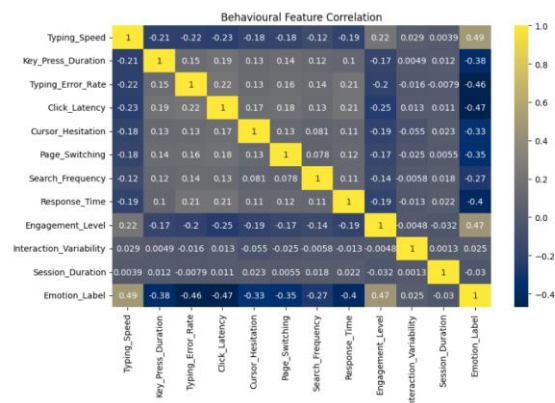


Fig 11. Behavioral Feature Correlation Matrix

The correlation matrix determines that behavioral characteristics have significant correlations with emotional labels. Speed of typing and involvement are positively related, but there is a negative relationship between latency, errors and hesitation. These trends support feature selection by warranting that behavioral predictors hold valuable emotional data to the classification models.

...	precision	recall	f1-score	support
0	0.82	0.80	0.81	199
1	0.62	0.61	0.61	203
2	0.74	0.77	0.75	198
accuracy			0.72	600
macro avg	0.73	0.73	0.73	600
weighted avg	0.72	0.72	0.72	600

Fig 12. Logistic Regression Classification Report

The Logistic Regression Classification Report shows the performance of the initial model with an overall accuracy of 72% on 600 test samples. Class 0(0.82 preciseness and 0.80 recall) performed best and Class 1 (0.62 preciseness and 0.61 recall) performed worse this is because it is hard to differentiate neutral emotion with other classes.



Fig 13. Logistic Regression Confusion Matrix for Emotional State Prediction

The Logistic Regression Confusion Matrix demonstrates that there is a distribution in the predictions across three emotional classes. The model accurately recognized 160, 123 and 152 samples in classes 0, 1 and 2 respectively. The highest misclassifications were between classes 1 and 2 with 47 neutral samples being predicted to be positive, indicating that behavior within an emotional state holds similarity.

...	precision	recall	f1-score	support
0	0.82	0.80	0.81	199
1	0.62	0.62	0.62	203
2	0.72	0.74	0.73	198
accuracy			0.72	600
macro avg	0.72	0.72	0.72	600
weighted avg	0.72	0.72	0.72	600

Fig 14. Random Forest Classification Report

Random Forest Classification Report depicts an overall classification of 72% over 600 test observations. Class 0 gives the highest performance with 0.82 precision, 0.80 recall and 0.81 F1-score. Class 1 has an F1-score of 0.62 and Class 2 attains 0.73 indicating moderate dissimilarity amidst smeared behavioral emotional design.

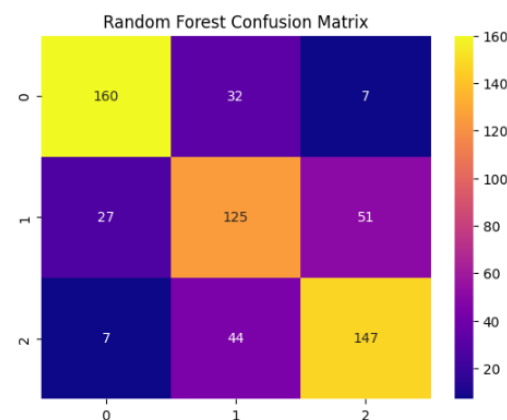


Fig 15. Random Forest Confusion Matrix

Random Forest Confusion Matrix has 160 correct predictions, 125 correct predictions of 0, 1 and 2, respectively. The greatest error is 51 Class 1 cases that are assumed to be Class 2, and 44 Class 2 cases that convert to Class 1 which means that there is a significant overlap in behavior associated with the neutral and positive emotional state.

...	precision	recall	f1-score	support
0	0.82	0.78	0.80	199
1	0.59	0.60	0.59	203
2	0.72	0.75	0.73	198
accuracy			0.71	600
macro avg	0.71	0.71	0.71	600
weighted avg	0.71	0.71	0.71	600

Fig 16. Gradient Boosting Classification Report for Behavioral Emotion Inference

The Gradient Boosting Classification Report demonstrates the overall model performance of 71% on 600 test samples. Class 0 had the best performance with 0.82 precision and 0.78 recall, Class 1 had worse performance (0.59 precision, 0.60 recall). The lower Class 1 performance suggests that there is a problem in isolating neutral emotional patterns and other neighboring states.

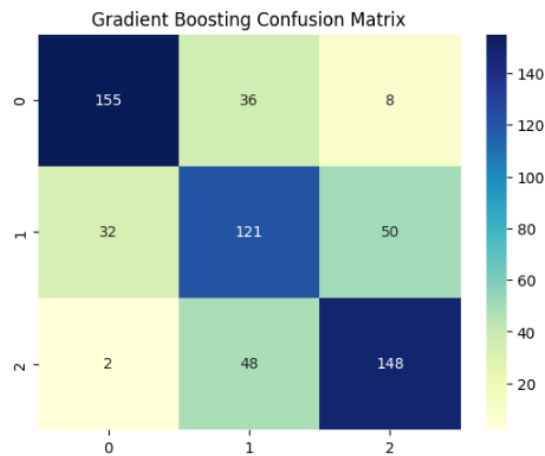


Fig 17. Gradient Boosting Confusion Matrix for Emotional State Prediction

Gradient Boosting Confusion Matrix shows the results of prediction of three emotional categories. The model accurately diagnosed 155, 121, and 148 Class 0, Class 1, and Class 2 samples respectively. The greatest misunderstanding is in class 1 and 2 where the number of neutral cases classified as positive was 50 indicating a similarity in behavioral features between the situations in these two emotional states.

Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 64)	19,456
dropout (Dropout)	(None, 64)	0
dense (Dense)	(None, 32)	2,080
dense_1 (Dense)	(None, 3)	99

Total params: 21,635 (84.51 KB)
Trainable params: 21,635 (84.51 KB)
Non-trainable params: 0 (0.00 B)

Fig 18. LSTM Model Architecture

The aim of the LSTM Model Architecture is a sequential learning network that provides the ability to recognize temporal relationships among interaction patterns of behavior. The model has 64 units and dropout regularization in the LSTM layer, 32 neurons in the dense layer and a 3-class output layer which yields 21,635 trainable parameters to classify the emotional states.

	precision	recall	f1-score	support
0	0.81	0.82	0.81	199
1	0.64	0.58	0.60	203
2	0.72	0.78	0.75	198
accuracy			0.72	600
macro avg	0.72	0.73	0.72	600
weighted avg	0.72	0.72	0.72	600

Fig 19. LSTM Classification Report for Behavioral Emotion Prediction

The LSTM Classification Report indicates an approximate accuracy rate of 72% in a test data of 600 samples. Class 0 had the best results of 0.81 precision, 0.82 recall and 0.81 F1-score whereas Class 1 had less recall (0.58) which shows difficulty in identifying neutral emotion based on sequencing behavioral patterns.

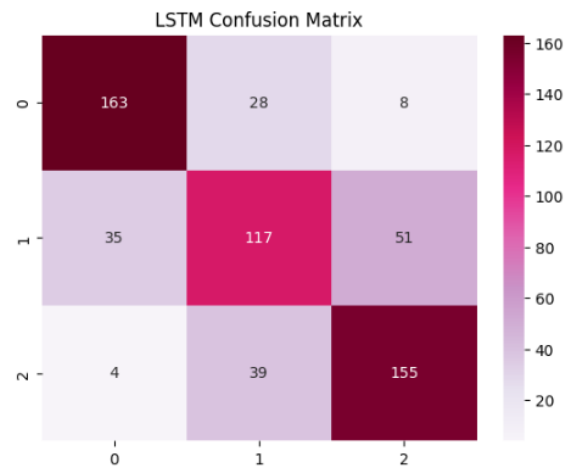


Fig 20. LSTM Confusion Matrix

The LSTM Confusion Matrix shows correct classifications of 163 Class 0, 117 Class 1, and 155 Class 2 samples. Major misclassification was between Class 1 and Class 2 where the prediction was Class 51 neutral samples. This shows that serial behavioral patterns enhance recognition and it does not deny the overlap amongst emotional categories.

	Model	Accuracy
0	Logistic Regression	0.725000
1	Random Forest	0.720000
2	Gradient Boosting	0.706667
3	LSTM	0.725000

Fig 21. Comparative Accuracy Evaluation of Machine Learning Models

The Comparative Accuracy Evaluation of Machine Learning Models shows that there are differences in the performance of four behavioral emotion

classification methods. Logistic Regression and LSTM had highest accuracy of 72.5 with a difference of 72 -Random Forest with 72 and Gradient Boosting with 70.67. The findings show that the predictive ability of the supervised learning models is similar.

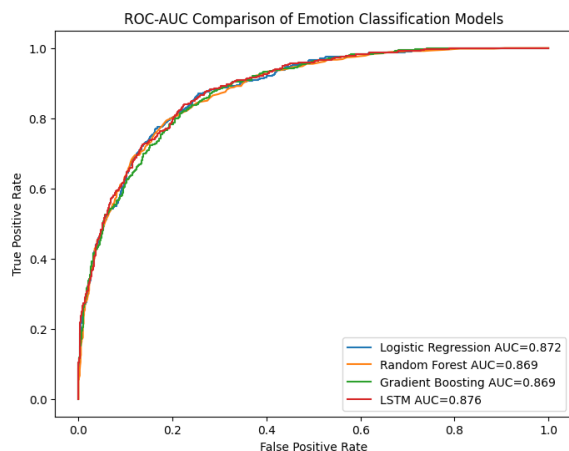


Fig 22. ROC-AUC Comparison of Emotion Classification Models

The ROC-AUC Comparison of Emotion Classification Models compares the strength of classification models based on the values of area under the curve. The highest AUC of 0.876 is achieved with LSTM and then Logistic Regression (0.872), Random Forest (0.869) and Gradient Boosting (0.869). In the three-class classification problem, the one-versus-rest approach has been used to predict the class probabilities, and ROC-AUC was calculated. The class-specific values of AUC were then computed and macro-averaged to give the sample multiclass ROC-AUC.

Discussion

The findings demonstrate the feasibility of classifying synthetically defined emotional states from simulated behavioral interaction features. The tested models performed similarly, with the highest accuracy of 72.5% being achieved by Logistic Regression and LSTM, and the ROC-AUC of 0.876 being the best that LSTM can achieve, showing a reliable ability to separate classes. The findings verify that typing patterns, latency, engagement and navigation behaviors are meaningful with emotional indicators that do not require physiological sensors. The state of being is also problematic when it comes to elusive classification between neutral and positive states as behavior might be similar. The paper notes that

supervised machine learning may offer a scalable alternative to biosensor based affective computing but larger real world data sets and adaptive learning are also needed to enhance generalization, strength and effectiveness in real world deployment.

Limitation

- The research is completely based on synthetic data and the usage of predetermined emotional categories, which restrict ecological validity and generalizability.
- It is possible that behavioral-emotional relations are based on generation assumptions and not on human behavior, thus, it is crucial to validate them with actual participants.
- The external behavioral proxies include typing speed, latency, and navigation patterns which can be affected by some external factors lowering the direct interpretation of emotions.
- The measured machine learning models should have increased diverse data size, explicative methodology and ongoing validation to enhance scalability, dependability, and successful deployment of distinct digital applications.

VI. CONCLUSION AND FUTURE DIRECTIONS

Conclusion

The paper illustrates that emotional state inference that is not hardware dependent is possible using interaction-based and behavioral digital proxies. The dynamics of typing, latency of click, pattern of responses, navigation behavior, and indicators of engaged yielded valuable information to be fed into machine learning classification with supervision. Logistic Regression and LSTM are shown the strongest ROC-AUC with good accuracy. These results from behavioral cues as a viable measurement of alternative are important to the biosensor-reliant affective computing. Future studies are to be informed with real-world longitudinal data, personalized behavioral baseline, explainable AI, multimodal interaction cues, and privacy preserving learning. These developments have the potential to enhance the generalization, emotional differentiation and

implementation in adaptive digital application in a variety of places.

Future Directions

Future studies are necessary to test behavioral emotion inference models on large-scale interaction data from real-world contexts through numerous user groups and digital contexts. Recognition of complex behavioral patterns can be enhanced with highly developed deep learning structures, such as transformer-based sequential models [30]. The incorporation of explainable AI techniques will improve trust and transparency in emotionally predicting users. Federated learning is a privacy-protective method that merits consideration in a bid to promote secure deployment. Moreover, integration of multiple modalities of behavior, user adaptation and continuous learning can enhance resilience, scalability and functioning of the emotion-sensitive smart digital interfaces of the future.

VII. REFERENCES

- [1] McStay, A., 2023. *Automating empathy: Decoding technologies that gauge intimate life* (p. 304). Oxford University Press.
- [2] Maeda, T. and Quan-Haase, A., 2024, June. When human-AI interactions become parasocial: Agency and anthropomorphism in affective design. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1068-1077).
- [3] Hibbeln, M., Jenkins, J.L., Schneider, C., Valacich, J.S. and Weinmann, M., 2017. How is your user feeling? Inferring emotion through human-computer interaction devices. *Mis Quarterly*, 41(1), pp.1-21.
- [4] Likforman-Sulem, L., Esposito, A., Faundez-Zanuy, M., Cl  men  on, S. and Cordasco, G., 2017. EMOTHAW: A novel database for emotional state recognition from handwriting and drawing. *IEEE Transactions on Human-Machine Systems*, 47(2), pp.273-284.
- [5] Alipour, M., Moghaddam, M.T., Vaidhyathan, K. and Kj  rgaard, M.B., 2023. Emoticontrol: emotions-based control of user-interfaces adaptations. *Proceedings of the ACM on Human-Computer Interaction*, 7(EICS), pp.1-29.
- [6] Li, R., Wu, Z., Jia, J., Li, J., Chen, W. and Meng, H., 2018, October. Inferring user emotive state changes in realistic human-computer conversational dialogs. In *Proceedings of the 26th ACM international conference on Multimedia* (pp. 136-144).
- [7] Ash, J., Anderson, B., Gordon, R. and Langley, P., 2018. Digital interface design and power: Friction, threshold, transition. *Environment and planning D: society and space*, 36(6), pp.1136-1153.
- [8] Kaklauskas, A., Abraham, A., Ubarte, I., Kliukas, R., Luksaite, V., Binkyte-Veliene, A., Vetloviene, I. and Kaklauskienė, L., 2022. A review of AI cloud and edge sensors, methods, and applications for the recognition of emotional, affective and physiological states. *Sensors*, 22(20), p.7824.
- [9] Ienca, M. and Malgieri, G., 2022. Mental data protection and the GDPR. *Journal of Law and the Biosciences*, 9(1), p.lsa006.
- [10] Mehta, D., Siddiqui, M.F.H. and Javaid, A.Y., 2018. Facial emotion recognition: A survey and real-world user experiences in mixed reality. *Sensors*, 18(2), p.416.
- [11] Schneider, C., Weinmann, M. and Vom Brocke, J., 2018. Digital nudging: guiding online user choices through interface design. *Communications of the ACM*, 61(7), pp.67-73.
- [12] He, Q., Li, L., Li, D., Peng, T., Zhang, X., Cai, Y., Zhang, X. and Tang, R., 2024. From digital human modeling to human digital twin: Framework and perspectives in human factors. *Chinese Journal of Mechanical Engineering*, 37(1), pp.1-14.
- [13] Wang, F., 2021, July. Study on dynamic emotional design expression in interface vision of digital media art. In *International Conference on Human-Computer Interaction* (pp. 497-504). Cham: Springer International Publishing.
- [14] Bond, R.R., Mulvenna, M.D., Potts, C., O'Neill, S., Ennis, E. and Torous, J., 2023. Digital transformation of mental health services. *Npj mental health research*, 2(1), p.13.
- [15] Braun, M., Weber, F. and Alt, F., 2021. Affective automotive user interfaces—reviewing the state of driver affect research and emotion regulation in the car. *ACM Computing Surveys (CSUR)*, 54(7), pp.1-26.



- [16] Makanadar, A., 2024. Neuro-adaptive architecture: Buildings and city design that respond to human emotions, cognitive states. *Research in Globalization*, 8, p.100222.
- [17] Greco, D., Barra, P., D'Errico, L. and Staffa, M., 2024, June. Multimodal interfaces for emotion recognition: models, challenges and opportunities. In *International Conference on Human-Computer Interaction* (pp. 152-162). Cham: Springer Nature Switzerland.
- [18] Shin, D., Shin, D. and Shin, D., 2017. Development of emotion recognition interface using complex EEG/ECG bio-signal for interactive contents. *Multimedia Tools and Applications*, 76(9), pp.11449-11470.
- [19] AlEisa, H.N., Alrowais, F., Negm, N., Almalki, N., Khalid, M., Marzouk, R., Alnfai, M.M., Mohammed, G.P. and Alneil, A.A., 2023. Henry gas solubility optimization with deep learning based facial emotion recognition for human computer interface. *Ieee Access*, 11, pp.62233-62241.
- [20] Lu, J., Liu, Y., Lv, T. and Meng, L., 2024. An emotional-aware mobile terminal accessibility-assisted recommendation system for the elderly based on haptic recognition. *International Journal of Human-Computer Interaction*, 40(22), pp.7593-7609.
- [21] Medjden, S., Ahmed, N. and Lataifeh, M., 2020. Adaptive user interface design and analysis using emotion recognition through facial expressions and body posture from an RGB-D sensor. *PloS one*, 15(7), p.e0235908.
- [22] Muhammad, G. and Hossain, M.S., 2021. Emotion recognition for cognitive edge computing using deep learning. *IEEE Internet of Things Journal*, 8(23), pp.16894-16901.
- [23] Ma, Y., Du, R., Abdelraouf, A., Han, K., Gupta, R. and Wang, Z., 2024. Driver digital twin for online recognition of distracted driving behaviors. *IEEE Transactions on Intelligent Vehicles*, 9(2), pp.3168-3180.
- [24] Lian, H., Lu, C., Li, S., Zhao, Y., Tang, C. and Zong, Y., 2023. A survey of deep learning-based multimodal emotion recognition: Speech, text, and face. *Entropy*, 25(10), p.1440.
- [25] Khan, U.A., Xu, Q., Liu, Y., Lagstedt, A., Alamäki, A. and Kauttonen, J., 2024. Exploring contactless techniques in multimodal emotion recognition: insights into diverse applications, challenges, solutions, and prospects. *Multimedia Systems*, 30(3), p.115.
- [26] Grill, G. and Andalibi, N., 2022. Attitudes and folk theories of data subjects on transparency and accuracy in emotion recognition. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW1), pp.1-35.
- [27] Mehta, D., Siddiqui, M.F.H. and Javaid, A.Y., 2018. Facial emotion recognition: A survey and real-world user experiences in mixed reality. *Sensors*, 18(2), p.416.
- [28] Wei, Q., Huang, X. and Zhang, Y., 2022. FV2ES: A fully end2end multimodal system for fast yet effective video emotion recognition inference. *IEEE Transactions on Broadcasting*, 69(1), pp.10-20.
- [29] Ciralo, D., Fazio, M., Calabrò, R.S., Villari, M. and Celesti, A., 2024. Facial expression recognition based on emotional artificial intelligence for tele-rehabilitation. *Biomedical Signal Processing and Control*, 92, p.106096.
- [30] Erat, K., Şahin, E.B., Doğan, F., Merdanoğlu, N., Akcakaya, A. and Durdu, P.O., 2024. Emotion recognition with EEG-based brain-computer interfaces: a systematic literature review. *Multimedia tools and applications*, 83(33), pp.79647-79694.