Psychophysiological Modelling of Trust in Technology: Comparative Analysis of Psychophysiological Signals

Ighoyota Ben Ajenaghughrure, Sónia Cláudia Da Costa Sousa and David Lamas
School of Digital Technologies, Tallinn University, Narva Mnt 25, 10120, Tallinn, Estonia

Keywords: Trust, Machine Learning, Psychophysiology, Autonomous Vehicle, Artificial Intelligence.

Abstract: Measuring users trust with psychophysiological signals during interaction (real-time) with autonomous systems that incorporates artificial intelligence has been widely researched with several psychophysiological signals. However, it is unclear what psychophysiological is most reliable for real-time trust assessment during user's interaction with an autonomous system. This study investigates what psychophysiological signal is most suitable for assessing trust in real-time. A within-subject four condition experiment was implemented with a virtual reality autonomous vehicle driving game that involved 31 carefully selected participants, while electroencephalogram, electrodermal activity, eletrocardiogram, eye-tracking and facial electromyogram psychophysiological signals were acquired. We applied hybrid feature selection methods on the features extracted from the psychophysiological signals. Using training and testing datasets containing only the resulting features from the feature selection methods, for each individual and multi-modal (combined) psychophysiological signals, we trained and tested six stack ensemble trust classifier models. The results of the model's performance indicate that the EEG is most reliable, while the multimodal psychophysiological signals remain promising.

1 INTRODUCTION

1.1 Motivation

Artificial intelligence technologies are becoming more ubiquitous. As their applications and presence cuts across a broad spectrum of activities and task in modern societies (Siau, 2017).

For instance, autonomous vehicles (AV's) have been developed to transport people from one place to another without human driver intervention in the civil transportation industry. Besides, robot assisting surgery (RAS) has been developed in the medical sector to help surgeons carry out high precision surgical procedures.

The emergence of AI technologies makes it imperative to foster collaborative interaction between users and AI based systems. This is due to the fact that AI-based systems operate autonomously and user's delegates/take-over task/control to/from AI based systems during interaction. For instance, users' interactions with autonomous vehicles involve giving over navigational control to the vehicle AI controller. Also, doctors interact with RAS during surgical procedures by giving over control of processes (e.g., surgical incision) to the RAS.

Prior efforts aimed at fostering users-AI-based systems (e.g. AV) teaming utilized the principle of traded controls that requires the driver to take control in case of failure or limited capability over certain conditions (also referred to as to as disengagement) (Dixit et al., 2016). During this transition, user's timely, accurate and appropriate response is required. However, without trust, such human technology teaming is bound to fail. For instance, the Tesla AV crash which led to the death of its driver was blamed on the driver streaming video during the incident (Beer et al., 2014).

The importance of trust is further emphasized in the study conducted by Litman (2017), during which, data from eight AV companies suggests that there are more than one disengagement in every 5,600 miles an AV travelled in 2017. Therefore as AVs' disengagement is inevitable, so is the need for successful users-AVs' teaming, and this requires trust between users and AVs'. Furthermore, trust between users and AVs' is influenced by prior failure experience of AI algorithms that controls the AVs'. This is further exacerbated by the fact that user's lack understanding of how AI algorithms that controls the

Ben Ajenaghughrure, I., Sousa, S. and Lamas, D.
Psychophysiological Modelling of Trust in Technology: Comparative Analysis of Psychophysiological Signals.
DOI: 10.5220/0010237701610173
In *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2021) - Volume 2: HUCAPP*, pages 161-173
ISBN: 978-989-758-488-6

AV operates, due to its design complexity (e.g. how, when and why it decides to turn left or right) (Parasuraman and Riley, 1997).

Hence, as Hurlburt (2017) quotes that ”any tendency to put blind faith in what in effect remains largely untrusted technology can lead to misleading and sometimes dangerous conclusions”, there is no doubt that trust will play a significant role in users interactions with AI-based technologies (Gefen et al., 2003; Li et al., 2008; Saiu et al., 2004). As trust has been shown to influence users behaviour (e.g., reliance), perceived usefulness, pleasantness, and overall acceptance of AI-based technologies such as autonomous vehicles (Hergeth et al., 2016; Payre et al., 2016; Rajaonah et al., 2006; Sollner & Leimeister 2013).

In order to foster users trust in AI-based systems and enhance positive users experience, ensuring that both user’s and AI technologies (AVs’) can jointly plan, decide, or control a system (vehicle/device) by sharing control is imminent (Abbink, et al., 2018). Hence, some researchers suggest effective calibration of users trust to avoid overtrust[1] or under trust[2] (Fallon et al., 2010; Hoffman et al., 2013; Lee & See 2004; Mirnig et al., 2016; Pop et al., 2015). Other researchers suggests that making the AI-based system explain” what, why and how it operates” to users could enhance users trust (Glass et al., 2008; Pu & Chen 2006). Although, Pieters (2011) suggests that explanation should be provided until trust is established, these approaches fail to address when explanation should be provided.

However, since trust is dynamic and constantly changes over time, calibration or explanation would be most meaningful after effective assessment of users trust levels in these AI based technologies (e.g. AV’s) is achieved. However, measuring trust continues to remain a challenge (Hurlburt, 2017). We believe, this challenge should be first addressed before moving onto what next after trust level is accurately assessed.

The widely used self-reporting trust assessment tools such as those develop by Gulati et al., (2019) are not suitable in this context because they can only be administered after interaction, The use of behavioural data such as users decision to rely or not rely on AI-based system during interaction are highly dependent on the interaction, context and artefact. Hence leaving the use of psychophysiological signal a viable method for development of real-time trust assessment tools, provided that the psychophysiological correlates of trust is known.

Therefore, making it imperative to develop tools that can assess users trust level in AI technologies (AVs’) in real-time using psychophysiological signals. A real-time trust assessment tool could enable algorithms that AI controls technologies such as AVs’ learn about users trust state and adapt its operations accordingly (Ajenaghughrure et al., 2019). As cognitive states (such as trust) can be used as feedback to the system in order to correct mistakes or inform the refinement of a learned control policy (Perrin et al., 2011). A potential application of real-time trust assessment tool is presented in Fig. 1

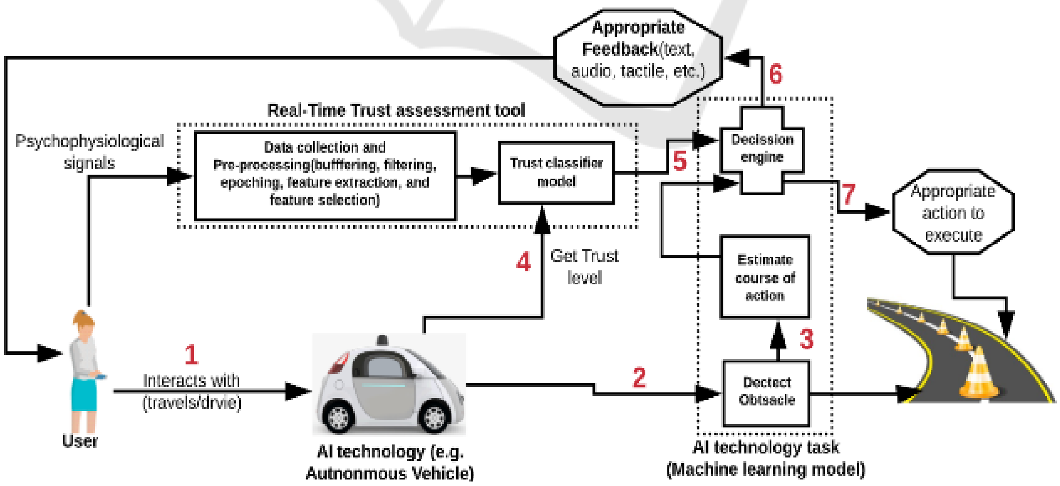

Figure 1: Typical use-case of real-time trust assessment.

[1] when a user trust a faulty or unreliable automated system

[2] when a user does not trust a reliable or non-faulty automated system [3]i.e biofeedback e.g. brain computer interface applications

below, a user interacting with an autonomous vehicle during a road trip-(1) the car detects an obstacle ahead using its sensor data. (2) The car uses its inbuilt machine learning model to determine its best navigational strategy. (3) The users trust state is assessed with the help of the trust classifier model that received as input the users' physiological signal data (EEG) pre-processed in real-time. The car provides the user with appropriate feedback–"*e.g., when trust is low: I understand that you are concerned about my ability to drive you through the obstacle ahead without involving in any crash, however I am 100% capable of navigating the obstacle ahead without any crash, kindly sit back and enjoy the ride*".

In addition, the same could be applied in the context of e-commerce where users trust could be measured during checkout and if found to be low, appropriate feedback such as *"hello we understand that you are concerned about purchasing product xyz, hence the merchant has agreed that you will not be charge until you receive and use the product for six months. If satisfied, then you will be charged."*

Further, in the context of doctors-RAS interaction, a realtime trust assessment tool could help foster cooperation between doctors and RAS during surgical procedure (Shafiei, et al., 2018).

1.2 Problem Statement

Although the use of psychophysiological signals for assessing users trust has been equally investigated by quiet a number of researchers, the question of what psychophysiological signal could be most reliable or should multi-modal psychophysiological signals be used to assess trust remains unattended.

Consequently, it is unclear which psychophysiological signal is most reliable for assessing users trust.

For instance, given that the psychophysiological correlates of trust were found in multi-modal psychophysiological signals such as the combination of eye-tracking combined with ECG by Leichtenstern et al., (2011), the psychophysiological correlates of trust in single psychophysiological signals has equally been found. For example, EEG was used by Oh et al., (2017) and Wang et al., (2018). Audio/voice and ECG was used by Elkins & Derrick (2010) and Watz et al., (2014). Eye tracking was used by Hergeth et al., (2016). However, it remains unclear which psychophysiological signals correlates better with users varying trust levels.

Furthermore, researchers investigating users trust assessment in real-time (i.e., during interaction) using single (electroencephalogram (EEG), functional near infrared spectroscopy (FNIRS), and electrodermal activity (EDA)) and multi-modal (EEG+EDA, audio/speech+ photoplethysmography + video) psychophysiological signals has developed fairly accurate classifier models that are capable of detecting users trust state from psychophysiological signals during interaction with AI-based systems. (Ajenaghughrure et al., 2019; Hirshfield et al., 2011; Shafiei et al., 2018; Lochner et al., 2019; Akash et al., 2018; Hu et al., 2016;).

It also remains unclear what psychophysiological signal is most suitable for developing real-time trust assessment tools? Further reinforced by the fact that there is dominance of features from one signal over the other(s) in studies where multimodal psychophysiological signals were utilized. For instance, Hu et al., (2016), despite extracting 108 features from the psychophysiological signals (EEG 105, EDA 3), the model utilized more EEG features (8) and less EDA features (2). Also, Akash et al., (2018), despite extracting 147 EEG features and 2 EDA features, both models (general and customized) used more EEG features (11 and 10) than EDA features (1 and 2). Furthermore, though the resulting model developed by Khalid et al., (2018) utilized features extracted (facial action code units, photoplethysmography (video-heart rate), audio/speech) from video and audio/speech psychophysiological signals, no details of the numbers of selected features per signal was provided.

1.3 Goals and Contribution

The goal of this study is to investigate what psychophysiological signal is most suitable for assessing users trust in real-time through developing and comparing stack ensemble trust classifier models, taking into account five psychophysiological signals (EEG: electroencephalogram, ECG: electrocardiogram, eye tracking, EDA: electrocardiogram, and facial EMG: electromyogram). These signals were considered because they are have been used in prior studies. In addition, we demonstrate the effectiveness of virtual reality technique for eliciting users trust dynamics during user's interactions with AI technologies that are otherwise expensive to acquire for conducting user experience studies.

2 METHODOLOGY

Virtual reality offers both the opportunity to immerse users in virtual environment where they experience products synonymously to real-world and the ability to assess user's experience (e.g., cognitive states such as trust and/or effective states such as emotions) (Rebelo et al., 2012).

Therefore, following game theoretic approach similar to prior research investigating trust (Ajenaghughrure et al., 2019), we developed an autonomous vehicle (AV) driving game. The game affords participants the opportunity to experience an AV under four categories of risk conditions that are directly mapped onto the automotive safety integrity levels (ASIL), also known as ISO-26262.

Elicitation of varying levels of risk through the game was motivated by the fact that risk is one (1) of the main factors that influences users' trust in technology (Gulati et al., 2019). ASIL classifies the inherent safety in automotive systems into four categories (A,B,C,D) based on the combination of severity of accident, likelihood of accident and exposure to accident (i.e. $ASIL = Severity * (Exposure * Likelihood)$) (Kinney and Wiruth, 1976).

Hence, a within subject 4 condition (very-high risk, high risk, low risk, no risk) experiment design was implemented as a game that tasked participants to stay safe. During the game, we captured participants trust dynamics through recording participants psychophysiological responses (EEG, EDA, ECG, facial EMG and eye tracking signals) during interactions with the AI technology (a simulated AV game) under various risk conditions.

2.1 Apparatus

Hardware: An MSI core i7 high performance gaming computer was used for the experiment. In addition, a 30inch LCD monitor was used to enhance visual display. Also, a Keyboard and mouse was provided to allow participants complete the trust in technology questionnaire (Gulati et al., 2019). In addition, a joystick was provided to participants to enable them to control the car when needed.
Software: Lab-stream layer software was used for aggregated recording of event markers from the game and all other psycho-physiological signals (EEG, ECG, EMG (facial) and EDA) into a single file in xdf format. In addition, using unity and C# programming language, we developed a hybrid fully autonomous vehicle (AV) driving game. More details about the game is described in (Ajenaghughrure et al., 2020). Also, Google hangout video call session running on a computer equipped with high definition camera installed in the experiment room was used to enable remote monitoring of participants during the experiment.

2.2 Participants

Invitation was sent through university mailing list, and printed handbills, with the help of an assistant.

Upon acceptance of the invitation, participants were asked to complete a google form to help us ascertain that each participant are right handed, free from any health condition that prevents them from driving, and are at least 18years and above. All participants that satisfied the above criteria were administered the driving habit questionnaire (DHQ) and behavioural inhibition / behavioural activation system questionnaire (BIS/BAS). Finally, only thirty one (31) healthy and right-handed participants (26.7% female, 73.3% male) aged 18 and above (M= 27.93333333, SD=5.607466287) participated in this study. This age range was considered based on prior studies which did not find any significant difference in psychophysiological responses when user aged 18 and above exhibit varying trust behaviour (Lemmers-Jansen et al., 2017). Furthermore, based on the responses recorded from the DHQ and BIS/BAS questionnaire (Owsley et al., 1999; Carver and White 1994), all participants had prior driving experience and symmetric personality traits with high BIS and BAS score (mean BIS>=2.5, mean BAS>=2.5, BIS score>=19, BAS score >= 40). In addition, order effect was avoided by grouping participants into two equal groups, each group is assigned to the four main game condition in reverse order.

2.2.1 Experiment Procedure

Upon arrival, participants were introduced to the experiment as a game involving test riding a prototype fully autonomous SUV vehicle intended for the future. Thereafter, participants completed and sign the informed consent form.

After that, an 8-channel wireless EEG recorder (G.tech Gmbh Austria.) was affixed to participants scalp. In addition, using bitalino wireless bio-signal acquisition systems, we affixed EDA sensor electrodes (2) to participants left hand palm area, EMG sensor electrodes (3) were placed on participants left and right eye sides to obtain horizontal EOG (Electrooculography), ECG sensor electrodes were placed on participant chest (left and right collar bone, and below the left chest area). Also, eye tracking data calibration with Miramatrix eye tracker was performed.
Thereafter, participants played the test game session to acquaint themselves to the available joystick (Logitech 3D Pro) controls that applies to the autonomous vehicle without any obstacles. At the end

of the demo game session, participants completed the trust in technology questionnaire adopted from Gulati et al., (2019) to obtain participants initial trust levels. This is followed with a 45seconds (game sessions loading time) relaxation acting as a baseline correction for the psychophysiological signals being recorded. After that, the experimenter exits the experiment room as participants began the main game session. After completing a game session (i.e. 13 trials), participants completed the trust in technology questionnaire adapted from Gulati et al., (2019) to obtain participants trust perception. In addition, the game logs consist of participants trust related behaviour (number of times AI was relied upon vs number of times joystick was relied upon). After completing the four game sessions, all psychophysiological sensors were removed following vendor guidelines. Finally, participants were debriefed and thanked with a gift card voucher worth 10EUR irrespective of the final score obtained at the end of the game (<75 or >=75points).

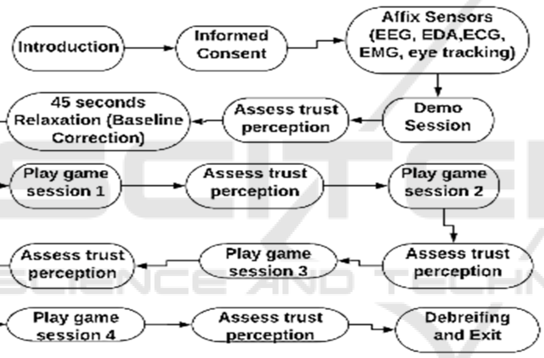

Figure 2: Experiment Procedure.

2.3 Data Collection and Pre-processing

Multimodal psychophysiological signals were recorded using labstream layer software and API for the respective physiological sensors.

The continuous EEG data was recorded using a wireless 8channel (Cz, Fz, C3, C4, F3, F4, P7 and P8 based on 10-20system) electrode amplifier from G.Tech Gmbh Austria. The sampling rate was 250Hz and impedance was <20kohm. Electrolyte gel was applied to each electrode to ensure proper conductivity and data quality. In addition, we used 75% metabolic spirit fluid to wipe the right ear lobe before affixing the ground electrode. Low pass filter of 120hz, high pass filter of 0.10Hz and notch filter of 50hz were used to remove sharp spikes, low-frequency drift noise and high-frequency sinusoidal power line noise respectively. The ground reference electrode was placed on the right earlobe, in addition to common ground.

Figure 3: Participant during experiment.

Also, the continuous ECG, EDA and Facial-EMG signals were recorded at a sampling rate of 1000hz. The EDA signals were acquired with two (2) gel prefilled electrodes that were placed on the left palm area of participant's. Using ledalad software, the EDA signals were: down-sampled to 50hz to reduce the computation cost (time) and denoised using adaptive smoothing to remove noise related with movements (Benedek and Kaernbach 2010).

In addition, the facial EMG signals were acquired with three gel prefilled electrodes attached to the left and right eye sides, and above the left eye brow, to obtain horizontal EOG signals. In addition, hand sanitizer applied to wipes were first used to wipe the areas before affixing the facial EMG prefilled gel electrodes.

Further, the ECG signals were acquired with three gel prefilled electrodes that were placed on the left (black electrode) shoulder, right shoulder (white electrode) and below the left chest (red electrode) area. In addition, hand sanitizer applied to wipes were first used to wipe the areas before affixing the ECG prefilled gel electrodes. Also, the ECG signals were downsampled to 50hz to reduce the computation cost (time) and filtered using neurokit python library (Makowski, 2016).

Furthermore, participants trust perception was measured subjectively using the trust in technology questionnaire adopted from Gulati et al., (2019). It consists of fourteen (14) items (question measuring risk perception, general trust, benevolence, reciprocity, and competence) measured on a scale of one (1) to five (5). Participants trust score was obtained by summing up the total response. This instrument was chosen because of its empirical nature.

In addition, participants non reliance (i.e., take-over: disengagement of AI control to manual control) on the AV was measured by aggregating the total joystick activation (0=not moved, 1=moved) beginning from the onset of an obstacle until an obstacle is past for all 52 trials.

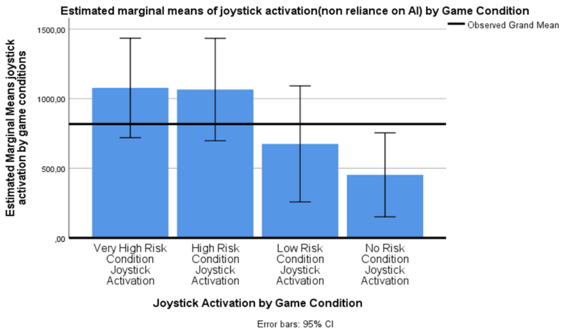

Figure 4: User non-reliance (joystick activation).

3 DATA ANALYSIS

3.1 Subjective Trust Perception and Objective Behavioural Trust Assessment

The result of the one way repeated measure ANOVA performed on the trust scores obtained from the participants before the playing the game and after playing each game sessions revealed that users trust before beginning the game (initial trust) was higher with statistical significant difference when compared to users trust during the very high risk and high risk game session were lower (difference in mean trust score 14,581 and 15,355 respectively, sig (0.00) <0.05). Further, although users trust before beginning the game (initial trust) was higher but was not statistically significant when compared to users trust during the low risk and no risk game session (difference in mean trust score 4,355 and 2,226 respectively, sig (0.220 and 1.00)>0.05 respectively).

In addition, users trust during the high risk game session is lower with statistical significant difference when compared to users trust during the low risk and no risk game session (Mean difference -11,000 and -13,129 respectively, sig (0.001) <0.05). Also, users trust during the very high risk game session is lower with statistical significant difference when compared to users trust during the low risk and no risk game session (difference in mean trust score -12,355 and -13,129 respectively, sig (0.001) <0.05).

However, there was no statistical significant difference between users trust during the very high risk and high risk game session. Same applies to the low risk and no risk game session.

Furthermore, users non-reliance (joystick usage) during the very high risk and high risk game session were higher with statistical significant difference when compared to user non-reliance (joystick usage) during the no risk game session (difference in mean trust score 624,258 and 612,742 respectively, sig (0.002 and 0.000 respectively) <0.05). Also, though users non-reliance (joystick usage) during the very high risk and high risk game session are higher when compared to users non-reliance (joystick usage) during the low risk game session (difference in mean trust score 402,129 and 390,613 respectively, sig (0.320 and 0.138) <0.05), it was not statistically significance, probably because users do not differentiate risk as low or high but present or absent.

These results suggest perceived risk during interaction with autonomous technologies influences users trust and overall reliance on autonomous technologies. In particular as risk increases trust and overall reliance decreases. Thereby reinforcing the need for real-time trust assessment tools.

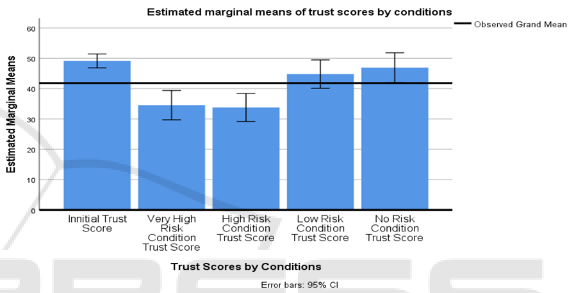

Figure 5: Users trust by game sessions.

3.2 Feature Extraction

The continuous EEG, EDA, ECG, eye tracking and facial EMG data were first divided into 4s epoch. Each epoch begins from the obstacle onset and ends 4s after. This time window was chosen because the average response time (i.e. the time from obstacle onset until first joystick movement) in cases where participants trust was low was four (4) seconds. Each epoch was labelled as high trust (coded as 2, if the joystick was not used during a trial) or low trust (coded as 1: if the joystick was used during a trial).

3.2.1 EEG

Using customised python script implementing python libraries from Python MNE (Gramfort et al., 2013) and MNE-feature extraction (Schiratti, et al., 2018), we extracted 160 exhaustive features from both time and frequency domain. The time domain features extracted for each EEG channel(i.e. 8 times 10) are the mean, variance, kurtosis, peak to peak amplitude (ptp amp), skewness, standard deviation (std), spectral entropy (spect entropy), singular value decomposition fisher information (svd fisher info), singular value decomposition entropy (svd entropy) and decorrelation time (decorr time). Further, the frequency domain features extracted from five frequency bands (alpha. beta, theta, gamma and delta) and each channel (i.e. 5 times 8 times 2) are the power spectrum (pow freq bands) and the band energy (energy freq bands). However, Only 30 participants data were included for further analysis, as one participants EEG epoch data were too noisy rendering all its epoch data invalid.

3.2.2 Facial EMG

Mean and peak to peak amplitude features were extracted from all 31 participants epoch facial-emg data using a customized python script implemented with python MNE libraries and MNE-feature libraries (Gramfort et al. 2013; Schiratti, et al., 2018). Therefore only two features were extracted from the facial EMG signal.

3.2.3 EDA

Using mathlab and ledlab software (Beer & Kaernbach, 2014), we extracted 12 EDA features from all 31 participants epoch and pre-processed EDA signals. Amongst which includes seven continuous phasic/tonic features using continuous decomposition analysis (CDA) based on standard deconvolution, three standard trough-to-peak (TTP) features, and two global measures (see Beer & Kaernbach, (2014), for detailed description of the features)

3.2.4 Eye Tracking

All 31 participants epoch eye tracking data were further pre-processed for feature extraction by computing the mean of each default features provided by the open-eye api (Hennessey & Duchowski, 2010). Therefore, the mean of each of the forty features outlined provided by the openeye api were computed (Hennessey & Duchowski, 2010).

3.2.5 ECG

Using customized python script implementing neurokit library (Makowski, 2016), we extracted three features (clean raw ecg, ecg rate, and ecg peak) from the epoch ECG psychophysiological signal data (aggregated from all 31 participants).

3.3 Ensemble Trust Classifier Model

Based on previous study (Ajenaghughrure et al., 2020), we selected five most promising algorithms (multi-layer perception (MLP), linear support vector machine (LSVM), regularised support vector machine algorithm (RBF-SVM), linear discriminant analysis algorithm (LDA), quadratic discriminant analysis algorithm (QDA)). These algorithms offers diverse characteristics that compliments the limitation of one another, thereby reducing the resulting classifier model biases and increasing its generalizability. Also, these algorithms has been successfully applied in brain computer interface research previously (Lotte et al., 2007). Our implementation of the ensemble trust classifier model was therefore achieved by combining all five algorithms through a technique known as classifier stack ensemble method (Lotte et al., 2007; Pedregosa et al., 2011). Ensemble of several algorithms aims to reduce classification error as suggested by prior research (Ajenaghughrure et al., 2019; Hu et al., 2016). Also, stack ensemble method was preferred over all other method because prior study has demonstrated that it is most superior when compared to other ensemble methods (e.g. voting, bagging, boosting) and unsupervised method such as deep neural network (Ajenaghughrure et al., 2020).

3.4 Feature Selection

We used hybrid feature selection method to select features from each individual and combined (multi-modal) psychophysiological signal epoch data-sets (i.e. EEG, EDA, ECG, EMG, and eye tracking). The choice of hybrid feature selection method was informed by prior study which has demonstrated that the resulting features from such method yields the most optimum ensemble trust classifier model performance (Ajenaghughrure et al., 2020). Hybrid feature selection method entails the combination of different feature selection method (e.g. filter and wrapper method).

The hybrid feature selection process applied to each individual and combined psychophysiological signal is detailed as follows: (1) Divide the epoch data samples into training and test samples (80% and 20% respectively). (2) Apply relieff filter feature selection method on subset of the training data sample, to identify model independent features. Relieff is an automated process that has been successfully applied in previous trust studies (Hu et al., 2016). Our implementation of the relieff feature selection method was achieved through a customised python script that implemented the relieff algorithm python library (Urbanowicz, et al., 2018). (3) Obtain model dependent features that promises optimum performance of the trust classifier model by applying wrapper feature selection (sequential forward floating feature selection method (SFFFS)) method on the subsets of the training samples containing only features obtained from step2. Our implementation of the wrapper feature selection method was achieved through a customised python script that implemented mlxtend python library (Raschka et al., 2018). This method evaluates our stack ensemble trust classifier model performance on various combinations of the model independent features to identify the most relevant feature for the specific model.

3.4.1 Multi-modal Psychophysiological Signal Feature Selection

The entire epoch multi-modal psychophysiological data (aggregated from 30 participants one participant EEG data epochs was corrupted.) containing 217 feature vector was first subjected to step1. Thereafter, step2 was applied on subset of the training epoch data (multi-modal psychophysiological signals) samples, and this process identified 30 model independent feature vectors (Urbanowicz, et al., 2018). Furthermore, applying step3 to subsets of the training epoch data (multi-modal psychophysiological signals) samples containing only features selected in step2 identified 14 relevant model dependent features that promises the utmost performance of the trust classifier model. Amongst which include: the global mean of the EDA signal, svd entropy from four EEG channels (c3, c4, f3, cz), svd fisher info from four EEG channels (c3, p7, f3 and cz), skewness from two EEG channels (p7 and cz), gamma power frequency band from EEG channel f3 and gamma energy frequency bands from two EEG channels (p8, and c3).

3.4.2 EEG

After excluding one participant data due to bad epochs, we applied step1 to the epoch EEG data samples containing 160 features. Thereafter, we applied step2 to subsets of the training epoch data (EEG) samples which resulted to top 15 model independent features being selected. Furthermore, we applied step3 to subsets of the training epoch data (EEG) samples containing 15 features selected in step2. The result of step3 is 10 model dependent features that promises optimum performance of the trust classifier model. The feature selected are the gamma energy frequency bands from two EEG channels (cz and c4), svd fisher info from six EEG channel (p7, p8, f3, f4, c3, c4), and svd entropy from two EEG channels (p6 and c4).

3.4.3 ECG

The epoch ECG psychophysiological signal data (aggregated from all 31 participants) samples containing all three features were first subjected to step1. Thereafter, we applied step2 to subset of the training epoch data (ECG) samples. The result of step2 is the selection of the top 2 model independent features. Furthermore, step3 was applied to subset of the training epoch data (ECG) samples containing only the two features selected in step2. This resulted to selecting only one model dependent feature (i.e., the clean raw epoch ECG signal) that promises utmost model performance.

3.4.4 EDA

Step1 was first applied to the epoch EDA psychophysiological signal data samples (aggregated from all 31 participants) containing the twelve features we extracted, and subset of the training epoch data (EDA) samples were further subjected to step2 process. The result of step2 is five model independent features selected. Furthermore, we applied step3 on subset of the training epoch data (EDA) samples containing only the five features sleeted in step2. The result of step3 is four model dependent features, amongst which includes: two CDA features (CDA.nSCR: Number of significant skin conductance response within response window (wrw), and CDA.Tonic: Mean tonic activity wrw of decomposed tonic component), one standard trough to peak feature (TTP.nSCR:Number of significant skin conductance response within response window (wrw)) and one global measure feature (Global.MaxDeflection: Maximum positive deflection wrw).

3.4.5 EMG

Step1 was applied to the epoch EMG psychophysiological signal data (aggregated from all 31 participants) containing the two feature extracted and further subjected to step2 which utilizes subset of the training epoch data (EMG) samples. The result of step3 are two model dependent features (mean, and peak-to-peak amplitude). Here we skipped step2 because we had extracted only two features.

3.4.6 Eye Tracking

Step1 was applied to epoch eye tracking psychophysiological signals data (aggregated from all 31 participants) samples containing the forty feature vectors and subset of the training epoch data(eye tracking) samples were further subjected to step2 process which selected seven model independent features. Furthermore, we applied step3 to subset of the training epoch data samples containing only the seven model independent features and this resulted to six model dependent features (RPUPILD: float right eye pupil diameter (mm), RPV: right eye pupil image valid, FPOGID: fixation number, REYEX: right eye position in X -left/+right (cm), CS: cursor button state, RPOGV: right point-of-gaze valid) that promises optimum performance of the trust classifier model.

3.5 Model Training and Validation

Using each psychophysiological signals (individual signals and multimodal signal) training data-sets (80%) containing the final features selected with SFFS method, we trained six stack ensemble trust classifier model outlined below:

- The first stack ensemble trust classifier model (V1) was trained with training data sets that consists of only multi-modal psychophysiological signals selected features.
- The second stack ensemble trust classifier model (V2) was trained with training data sets that consists of only EEG psychophysiological signal selected features.
- The third stack ensemble trust classifier model (V3) was trained with training data sets that consists of only eye tracking psychophysiological signal selected features.
- The fourth stack ensemble trust classifier model (V4) was trained with training data sets that consists of only EDA psychophysiological signals selected features.
- The five stack ensemble trust classifier model (V5) was trained with training data sets that consists of only ECG psychophysiological signals selected features.
- The sixth stack ensemble trust classifier model (V6) was trained with training data sets that consists of only facial-EMG psychophysiological signal selected features.

Each model was trained using the stratified three-fold cross validation method. This method first divides the training data (80% of the entire data samples) into specified partitions (three in this case) containing equal percentage of each class samples, then trains the given model on some data partition(given partition minus one, i.e. two) and evaluates the given model on the reserved data partition

The results of each model performance (accuracy minimum, maximum, and mean) based on the cross validation is outlined in table 1.

The stack ensemble trust classifier model V1 achieved an accuracy of 78.4% (minimum) for some samples, while for other samples, it achieved an accuracy of 82.0%. Also, its mean accuracy is 80.0%.

Also, the stack ensemble trust classifier model V2 achieved an accuracy of 80.4% (minimum) for some samples, while for other samples, it achieved an accuracy of 87.8%. Also, its mean accuracy is 83.4%.

In addition, the stack ensemble trust classifier model V3 achieved an accuracy of 50.2% (minimum) for some samples, while for other samples, it achieved an accuracy of 57.6%. Also, its mean accuracy is 53.9%. Also, the stack ensemble trust classifier model V2 achieved an accuracy of 51.0% (minimum) for some samples, while for other samples, it achieved an accuracy of 58.8%. Also, its mean accuracy is 54.8%.

Furthermore, the stack ensemble trust classifier model V2 achieved an accuracy of 51.4% (minimum) for some samples, while for other samples, it achieved an accuracy of 53.4%. Also, its mean accuracy is 52.0%. Also, the stack ensemble trust classifier model V2 achieved an accuracy of 59.2% (minimum) for some samples, while for other samples, it achieved an accuracy of 64.9%. Also, its mean accuracy is 61.8%.

Therefor these results suggest that all the ensemble trust classifier models, irrespective of the psychophysiological signal utilized during their development, are stable. Considering that the minimum accuracy's ranges from 50.2% to 80.4%, the maximum accuracy's ranges between 53.9% to 82.9%, and the mean accuracy's ranges from 53.4% to 87.8%. Also, no model had accuracy below 50% for any given sample.

However, with regards to performance, the stacked ensemble trust classifier model (V2) developed with EEG psychophysiological signal attained the most performance. The stack ensemble trust classifier model (V1) developed with multi-modal psychophysiological signals attained the second most optimum performance.

With regards to all other stacked ensemble trust classifier models (V3, V4, V5, V6), the model (V6) developed with facial-EMG psychophysiological signal is the next most optimum model, followed by the model (V4) developed with EDA psychophysiological signal, and next is the model (V3) developed with eye tracking psychophysiological signal. The least optimum is the model (V5) developed with ECG psychophysiological signal.

The implication of these results is that EEG is the most relevant psychophysiological signals for assessing trust. While multimodal psychophysiological signal is equally promising, but more research is still required. In addition, facial EMG is equally a promising psychophysiological signal for assessing trust. However, the performance of both EDA, ECG, and eye tracking psychophysiological signals were not too encouraging.

Table 1: Models CV performance (Accuracy(%) minimum. maximum, mean).

SN	Model	Mean	Min	Max	Stability
1	Multimodal	0.800	0.784	0.820	0.036
2	EEG	0.834	0.804	0.878	0.074
3	Eye-Tracking	0.539	0.502	0.576	0.074
4	EDA	0.548	0.510	0.588	0.078
5	ECG	0.520	0.514	0.534	0.02
6	Facial EMG	0.618	0.592	0.649	0.057

3.6 Model Validation/ Evaluation

Considering that the validation during cross validation and training could have some leaked data samples present in both the validation and training data partitions, and consequently results to model over-fitting as argued by some scholars (Lotte et al., 2007). Therefore, we further tested each ensemble trust classifier model with reserved test data (i.e. 20% of the entire data samples).

As outlined in table 2 below, the stack ensemble trust classifier model (V1 and V2) developed with multi-modal psychophysiological signals and EEG psychophysiological signal yielded the most optimum performance (accuracy 80.5% and 79.8% respectively). However, the performance difference (0.7%) between both models (V1 and V2) is quiet low. Furthermore, the stack ensemble trust classifier model (V6) developed with Facial EMG psychophysiological signals is the next most performing model with an accuracy of 61.6%.

In addition, the stack ensemble trust classifier models (V4 and V5) developed with EDA and ECG psychophysiological signals performance(accuracy 56.7% and 56.5% respectively) were below the performance of the stack ensemble trust classifier models(V1, and V2) developed with EEG and multimodal psychophysiological signal. Although, the stack ensemble trust classifier models(V4 and V5) developed with ECG and EDA psychophysiological signals appears to be more promising than the stack ensemble trust classifier model developed with eye tracking psychophysiological signal which attained 55.4%, all three models performance are poor in comparison to the models developed with EEG and multi-modal psychophysiological signals.

Therefore these results implies that EEG and multimodal psychophysiological signals are the most reliable psychophysiological signals for developing stacking ensemble models for assessing users trust during interaction with technology. Although, facial EMG seems promising, there is still room for more research using facial EMG, in order to understand its scope better. Also,

Table 2: Models test performance.

SN	Models	Accuracy	Recall	Precision	ROC-AUC
1	Multimodal	0.805	0.805	0.843	0.805
2	EEG	0.798	0.787	0.846	0.800
3	Eye tracking	0.554	0.948	0.563	0.493
4	EDA	0.567	1.000	0.567	0.500
5	ECG	0.565	1.000	0.565	0.500
6	Facial EMG	0.616	0.954	0.601	0.563

3.7 Discussion and Implication for HCI Researchers Investigating Trust

The results of this study clearly identified EEG psychophysiological signal as the most reliable psychophysiological signal for assessing users trust in technology. Although this result is reinforced by the fact that the trust classifier model (V2) developed with EEG psychophysiological signal outperformed the other models (v3, v4, v5, v6) developed with other psychophysiological signals, the comprehensive review by the authors in (Ajenaghughrure et al, 2020) identified EEG as the most frequently used psychophysiological signals in studies assessing trust with psychophysiological signals. One reason for this result could be because EEG has high temporal resolution, compared to the other psychophysiological signals.

In addition, the models (v3,v4,v5,v6) developed with the other psychophysiological signals (eye tracking, ECG, EDA, facial-EMG) performing poorly could be as a result of the data epoch time window (4s) that was chosen based on the average response time in this study, and the context being a time sensitive context. Probably when longer epoch time window is used in other context (e.g. e-commerce) that are not time sensitive, these other psychophysiological signals could perform better. Therefore, future research could examine epoch duration. Hence the result of this study is not entirely applicably across all technical artefact context, but subject to further investigation.

Furthermore, though the model (v1) developed with multimodal psychophysiological signal outperforms the model (v2) developed with EEG signal during validation, it is worth pointing out the majority of the selected features in the multimodal model (v1) are EEG psychophysiological signal features, and just a single EDA psychophysiological signal feature was selected. This leaves an important question on why such occurrence, and how best to perform feature selection for multi-modal psychophysiological signals. Hence, the subject of multimodal psychophysiological for assessing trust remain largely unclear and requires further investigation.

Further, like prior studies, we have performed an offline model development and evaluation. However, it is unclear how these models will perform when applied in real-time context.

In addition, though the maximum accuracy reported in this study and most prior studies _+/- 80% or more, therefore, when these models are deployed in realtime context with adaptive feedback based on users trust important questions about how wrongly estimated trust level and corresponding feedback would affect users trust and overall experience would emerge.

Although, the current study result suggests EEG is most optimum, the implementation of AV's has neither explored the current concept of real-time trust assessment with adaptive feedback. Hence, another important dimension that future research could examine is the application of real-time trust assessment with adaptive feedback in the wild. Though EEG systems are available in various form factors with cost ranging from less a 100USD to several thousands, dealing with noise and other physical activity that could impair the signal quality is another issue that future research must address. In addition to exploring non-invasive signals such as voice/audio.

The significance of this study result for future HCI researchers and designers investigating realtime trust assessment in AI-based systems is in the aspect of eliciting and informing the choice of psychophysiological signal to utilize during the development of a trust state classifier model that can automatically classify users trust state (users experience) based on psychophysiological signals. Our result generally shows that it is feasible to assess users trust state during interaction (real-time) with an autonomous system and the most reliable signal to use at the moment is EEG.

4 CONCLUSION

In conclusion, a user study involving autonomous vehicle in virtual. In addition, as we transition into the era of AI technologies, creating a symbiotic interactive atmosphere that guarantees successful user's technologies (e.g. AV's) teaming is imperative. Hence, trust researchers have attempted the development of ensemble trust classifier models that can assess users trust in technology during interaction from psychophysiology.

However, due to the fact that there are plethora of psychophysiological signals, the choice of what psychophysiological signals to employ when developing real-time trust assessment tools and its dependent components such as trust classifier models, is solely researchers discretion. Consequently, it is unclear what psychophysiological signal is most reliable for assessing users trust during interaction with AI-based systems (e.g., AV).

Hence motivating this study which investigated what psychophysiological signal is most reliable for assessing trust. The results of six ensemble trust classifier models we developed with individual (i.e. EEG. ECG, EDA, facial EMG, eye tracking) and multimodal psychophysiological signal features extracted through hybrid feature selection methods. The result indicates that the EEG and multimodal psychophysiological signal led to the most optimum ensemble trust classifier models (V2, and V1).

Although these results are obtained in offline model development and evaluation mode, future research will examine the model performance in real-time mode. Depending on how successful this becomes a new research line inquiring into the direction of identifying relevant feedback modalities.

ACKNOWLEDGEMENTS

This research was supported by Tallinn University research fund project-F/5019, Estonia Research Council grant PUT1518and the European Union Horizon 2020 research and innovation program under the NGITRUST grant agreement no 825618.

REFERENCES

Abbink, D. A., Carlson, T., Mulder, M., de Winter, J. C., Aminravan, F., Gibo, T. L., & Boer, E. R. (2018). A topology of shared control systems—finding common ground in diversity. IEEE Transactions on Human-Machine Systems, 48(5), 509-525.

Ajenaghughrure, I. B., da Costa Sousa, S. C., & Lamas, D. (2020, June). Risk and Trust in artificial intelligence technologies: A case study of Autonomous Vehicles. In 2020 13th International Conference on Human System Interaction (HSI) (pp. 118-123). IEEE.

Ajenaghughrure, I. B., Sousa, S. C., Kosunen, I. J., & Lamas, D. (2019, November). Predictive model to assess user trust: a psycho-physiological approach. In Proceedings of the 10th Indian Conference on Human-Computer Interaction (pp. 1-10).

Ajenaghughrure, I. B., Sousa, S. D. C., & Lamas, D. (2020). Measuring Trust with Psychophysiological Signals: A Systematic Mapping Study of Approaches Used. Multimodal Technologies and Interaction, 4(3), 63.

Ajenaghughrure, I. B., Sousa, S., and Lamas, D. (2020) Psychphysiological modelling of trust in technology: Comparative analysis of algorithm ensemble methods, IEEE SAMI 2021 (accepted)

Ajenaghughrure, I. B., Sousa, S., and Lamas. D. (2020) Psychophysiological modeling of trust in technology: The influence of feature selection methods. 13TH EICS PACM journal (accepted)

Akash, K., Hu, W. L., Jain, N., & Reid, T. (2018). A classification model for sensing human trust in machines using EEG and GSR. ACM Transactions on Interactive Intelligent Systems (TiiS), 8(4), 1-20.

Beer, J. M., Fisk, A. D., & Rogers, W. A. (2014). Toward a framework for levels of robot autonomy in human-robot interaction. Journal of human-robot interaction, 3(2), 74.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. Journal of neuroscience methods, 190(1), 80-91.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. Journal of personality and social psychology, 67(2), 319.

Elkins, A. C., & Derrick, D. C. (2013). The sound of trust: voice as a measurement of trust during interactions with embodied conversational agents. Group decision and negotiation, 22(5), 897-913.

Fallon, C. K., Murphy, A. K., Zimmerman, L., & Mueller, S. T. (2010, May). The calibration of trust in an automated system: A sensemaking process. In 2010 International Symposium on Collaborative Technologies and Systems (pp. 390-395). IEEE.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. MIS quarterly, 27(1), 51-90.

Glass, A., McGuinness, D. L., & Wolverton, M. (2008, January). Toward establishing trust in adaptive agents. In Proceedings of the 13th international conference on Intelligent user interfaces (pp. 227-236).

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. Frontiers in neuroscience, 7, 267.

Gulati, S., Sousa, S., and Lamas, D., (2019) Design, development and evaluation of a human-computer trust scale. Behaviour & Information Technology, pp. 1–12.

Hennessey, C., & Duchowski, A. T. (2010, March). An open source eye-gaze interface: Expanding the adoption of eye-gaze in everyday applications. In Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications (pp. 81-84).

Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. F. (2016). Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving. Human factors, 58(3), 509-519.

Hirshfield, L. M., Hirshfield, S. H., Hincks, S., Russell, M., Ward, R., & Williams, T. (2011, July). Trust in human-computer interactions as measured by frustration, surprise, and workload. In International Conference on Foundations of Augmented Cognition (pp. 507-516). Springer, Berlin, Heidelberg.

Hoffman, R. R., Johnson, M., Bradshaw, J. M., & Underbrink, A. (2013). Trust in automation. IEEE Intelligent Systems, 28(1), 84-88.

Huckle, T., & Neckel, T. (2019). Bits and Bugs: A Scientific and Historical Review on Software Failures in Computational Science. Society for Industrial and Applied Mathematics.

Hurlburt, G. (2017). How much to trust artificial intelligence?. IT Professional, 19(4), 7-11.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human factors, 46(1), 50-80.

Leichtenstern, K., Bee, N., André, E., Berkmüller, U., & Wagner, J. (2011, June). Physiological measurement of trust-related behavior in trust-neutral and trust-critical situations. In IFIP International Conference on Trust Management (pp. 165-172). Springer, Berlin, Heidelberg.

Lemmers-Jansen, I. L., Krabbendam, L., Veltman, D. J., & Fett, A. K. J. (2017). Boys vs. girls: gender differences in the neural development of trust and reciprocity depend on social context. Developmental Cognitive Neuroscience, 25, 235-245.

Litman, T. (2014). Autonomous vehicle implementation predictions. Victoria Transport Policy Institute, 28(2014).

Lochner, M., Duenser, A., & Sarker, S. (2019, December). Trust and Cognitive Load in semi-automated UAV operation. In Proceedings of the 31st Australian Conference on Human-Computer-Interaction (pp. 437-441).

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain–computer interfaces. Journal of neural engineering, 4(2), R1.

Makowski, D. (2016). Neurokit: A python toolbox for statistics and neurophysiological signal processing (eeg eda ecg emg...). Memory and Cognition Lab'Day, 1.

Mirnig, A. G., Wintersberger, P., Sutter, C., & Ziegler, J. (2016, October). A framework for analyzing and calibrating trust in automated vehicles. In Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 33-38).

Oh, S., Seong, Y., & Yi, S. (2017). Preliminary study on neurological measure of human trust in autonomous systems. In IIE Annual Conference. Proceedings (pp. 1066-1072). Institute of Industrial and Systems Engineers (IISE).

Owsley, C., Stalvey, B., Wells, J., & Sloane, M. E. (1999). Older drivers and cataract: driving habits and crash risk. Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences, 54(4), M203-M211.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human factors, 39(2), 230-253.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011).

Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.
Pu, P., & Chen, L. (2006, January). Trust building with explanation interfaces. In Proceedings of the 11th international conference on Intelligent user interfaces (pp. 93-100).
Rajaonah, B., Anceaux, F., & Vienne, F. (2006). Trust and the use of adaptive cruise control: a study of a cut-in situation. Cognition, Technology & Work, 8(2), 146-155.
Raschka, S. (2018). MLxtend: providing machine learning and data science utilities and extensions to Python's scientific computing stack. Journal of open source software, 3(24), 638.
Rebelo, F., Noriega, P., Duarte, E., & Soares, M. (2012). Using virtual reality to assess user experience. Human Factors, 54(6), 964-982.
Rebelo, F., Noriega, P., Duarte, E., & Soares, M. (2012). Using virtual reality to assess user experience. Human Factors, 54(6), 964-982.
Schiratti, J. B., Le Douget, J. E., Le van Quyen, M., Essid, S., & Gramfort, A. (2018, April). An ensemble learning approach to detect epileptic seizures from long intracranial EEG recordings. In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 856-860). IEEE.
Shafiei, S. B., Hussein, A. A., Muldoon, S. F., & Guru, K. A. (2018). Functional brain states measure mentor-trainee trust during robot-assisted surgery. Scientific reports, 8(1), 1-12.
Siau, K. (2017). Impact of artificial intelligence, robotics, and automation on higher education.
Siau, K., Sheng, H., Nah, F., & Davis, S. (2004). A qualitative investigation on consumer trust in mobile commerce. International Journal of Electronic Business, 2(3), 283-300.
Söllner, M., & Leimeister, J. M. (2013). What we really know about antecedents of trust: A critical review of the empirical information systems literature on trust. Psychology of Trust: New Research, D. Gefen, Verlag/Publisher: Nova Science Publishers.
Urbanowicz, R. J., Olson, R. S., Schmitt, P., Meeker, M., & Moore, J. H. (2018). Benchmarking relief-based feature selection methods for bioinformatics data mining. Journal of biomedical informatics, 85, 168-188.
V. L. Pop, A. Shrewsbury, and F. T. Durso. Individual differences in the calibration of trust in automation. Human factors, 57(4):545–556, 2015.
V. V. Dixit, S. Chand, and D. J. Nair. Autonomous vehicles: disengagements, accidents and reaction times. PLoS one, 11(12):e0168054, 2016.
W. Payre, J. Cestac, and P. Delhomme. Fully automated driving: Impact of trust and practice on manual control recovery. Human factors, 58(2):229– 241, 2016.
W. Pieters. Explanation and trust: what to tell the user in security and ai? Ethics and information technology, 13(1):53–64, 2011.
W.-L. Hu, K. Akash, N. Jain, and T. Reid. Real-time sensing of trust in human-machine interactions. IFAC-PapersOnLine, 49(32):48–53, 2016.
Wang, M., Hussein, A., Rojas, R. F., Shafi, K., & Abbass, H. A. (2018, November). EEG-based neural correlates of trust in human-autonomy interaction. In 2018 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 350-357). IEEE.
Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. Journal of Experimental Social Psychology, 52, 113-117.
X. Li, T. J. Hess, and J. S. Valacich. Why do we trust new technology? a study of initial trust formation with organizational information systems.The Journal of Strategic Information Systems, 17(1):39–71, 2008.
X. Perrin, F. Colas, C. Pradalier, R. Siegwart, R. Chavarriaga, and J. d. R. Millan. Learning user habits for semi-autonomous navigation using low´ throughput interfaces. In 2011 IEEE International Conference on Systems, Man, and Cybernetics, pp. 1–6. IEEE, 2011.
G Fine Kinney and AD Wiruth. 1976. Practical risk analysis for safety management. Technical Report. Naval Weapons Center China Lake Ca.