Technical Correspondence

Blended Emotion Detection for Decision Support

Anuja Hariharan and Marc Thomas Philipp Adam

***Abstract*—Emotion elicitation and classification have been performed on standardized stimuli sets, such as international affective picture systems and international affective digital sound. However, the literature which elicits and classifies emotions in a financial decision making context is scarce. In this paper, we present an evaluation to detect emotions of private investors through a controlled trading experiment. Subjects reported their level of "rejoice" and "regret" based on trading outcomes, and physiological measurements of skin conductance response and heart rate were obtained. To detect emotions, three labeling methods, namely binary, tri-, and tetrastate "blended" models were compared by means of C4.5, CART, and random forest algorithms, across different window lengths for heart rate. Taking moving window lengths of 2.5 s prior to and 0.3 s postevent (parasympathetic phase) led to the highest accuracies. Comparing labeling methods, accuracies were 67% for binary rejoice, 44% for a tristate, and 45% for a tetrastate blended emotion models. The CART yielded the highest accuracies.**

***Index Terms*—Emotion recognition, multimodal sensors, user behavior.**

I. Introduction

Emotion detection in user systems serves three kinds of purposes: 1) enhancing user experience by means of empathetic systems that can support the user, 2) motivating the user toward a certain goal, and 3) building systems that model human-like behavior by mimicking the user state [1]. Affective systems use externally available user information, such as physiological changes, to approximate a possibly hidden user state of mind. Picard *et al.* [1] illustrate the use of complementary physiological information in building affective systems. To this end, datasets, such as international affective picture systems (IAPS) [2] and international affective digital sounds [3] were used to establish ground truth in eliciting emotions, whereas these datasets provide a standardized method in classifying physiological information, context-based real-life emotions can be harder to elicit as well as detect, as: they 1) can be ambiguous, 2) may originate at the unconscious level, or 3) may be difficult for the individual to ascertain and report. Moreover, a self-reported method in a real-life context suffers from the inherent bias that subjects are made to "think" about what they "feel." Building a user-independent physiological-based emotion detection system hence proves to be a challenging problem. A possible approach is to detect emotions in a controlled setting, wherein subjects are aware of the circumstances of measuring their emotions, and whether different reported emotion states could be suitably modeled by physiology.

This paper performs a proof-of-concept evaluation of detecting emotions in financial decision making, the role of which has been highlighted in [4] and [5]. Financial decision making has been found to be clouded with factors of risk, uncertainty, ambiguity [6], and how people respond to rewards and losses. The influence of emotions under these circumstances leads to biases and decisions, which are potentially regretted later, due to the financial cost of making these mistakes [7]–[9]. Damasio [10] distinguishes between *emotion* and *feeling an emotion*, i.e., the consciousness of experiencing an emotion. In the former case, one has less control over how emotions influence behavior, whereas the latter has shown to provide more options for decision making, by consciously choosing how he/she would like to factor in an emotion. Hence, an important aspect in aiding financial decision makers is to enable access to emotions, which occur at the unconscious level. If these unconscious processes were detected and presented to the trader, it could make the trader aware of his/her current emotional state, and where applicable, also detect an anomalous state. Traders could then regulate these emotions—downregulate potentially harmful emotions, or upregulate those positively correlated with decision performance—and, hence, be trained toward an emotionally aware and conscious financial decision process.

In this paper, we take the first step toward conscious decision making, i.e., detect emotions in trading decisions, by eliciting and evaluating self-reported emotions against physiologically measured emotions, acquired by means of skin conductance response (SCR) and heart rate (HR) data. Subjects decide whether to hold or sell an endowed stock based on the stock properties and the observed trend after one period. Since participants observe the change in price prior to and post the decision event, subjects may experience a combination of emotions depending on whether they made both gains and losses, or only gains, or only losses. Consequently, emotions are classified in a binary model (as "regret" or "rejoice"), a tristate mixed model with a blended emotion state (i.e., only regret, only rejoice, or blended), and a tetrastate model with two blended emotion states (only regret, only rejoice, both, and neither emotion) by use of C4.5, classification and regression trees (CART), and the random forest classification algorithm. We demonstrate that user emotions elicited and measured in a controlled trading environment can be classified using a blended emotion model. In this context, traders have exhibited emotions in making their decisions [11]. The questions addressed in this paper are as follows:

Manuscript received October 30, 2014; revised January 15, 2015. Date of publication April 21, 2015; date of current version July 11, 2015. This paper was recommended by Associate Editor C. P. Nemeth.

A. Hariharan is with the Institute of Information Systems and Marketing, Karlsruhe Institute of Technology, Karlsruhe 76131, Germany (e-mail: anuja.hariharan@kit.edu).

M. T. P. Adam is with the School of Design, Communication, and IT, The University of Newcastle, Callaghan, N.S.W. 2308, Australia (e-mail: marc.adam@newcastle.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/THMS.2015.2418231

1) What is the moving average time-window that leads to highest classification accuracies for information events for each frame of interest (3 s before, 3 s after, and 3–7 s after) to represent features of phasic HR differences?
2) What is the best algorithm between CART, C4.5, and random forest to predict emotions of regret and rejoice in a financial decision making context?
3) Which is the best annotation scheme between binary, tristate, and tetrastate blended emotion models to predict the above emotions?

The structure of this paper is as follows: Related work on emotions in financial decision making, emotion detection using HR and SCR, and classifiers for emotion detection is presented in Section II. The experimental methodology is explained in Section III. Results are presented in Section IV, followed by discussion and conclusion in Section V.

II. Related Work

A. Emotions in Financial Decision Making

Emotions in decision making have been categorized as "anticipatory" (occurring in anticipation of an outcome of different possible courses of action) and "immediate," at the moment of choice, which are further classified as "integral" and "incidental," depending on whether they occur due to the consequences of a decision, or from sources unrelated to the task at hand [9]. Immediate emotions consist of parasympathetic and sympathetic phases. These visceral changes have been observed when traders experience loss, regret their decisions, or are rejoiceful about the outcome of their decisions. In this paper, we focus on anticipatory and immediate emotions, which are integral and relevant to the decision at hand.

In financial decision making, two commonly studied emotions are "rejoice" and "regret." Rejoice is experienced if the action results in an outcome that is better than the alternative, whereas regret can be defined as an aversive emotion a decision maker experiences upon the discovery that he/she could have gained a higher level of utility for a different choice in the past [12]. Astor *et al.* [13] elicit and measure regret in an auction context, by means of SCR, and further classify it as a winner and loser regret. These physiological methods serve as a means to understand and quantify emotions in the financial decision making context. Fernandez *et al.* [14] demonstrate the necessity to provide ambient intelligent support to traders to detect the exact moment under which they are under stress, due to an outcome of their decision, and suggest the usage of SCR, HR, breathing rate, and brain waves to make traders stress aware. While traders are trained financial decision makers, financial decisions are also continuously made by less experienced private investors and consumers. In this paper, we focus on understanding the emotions in financial decisions at large, especially by people who are not adept in decision making. In order to measure emotions, the previous literature suggests that SCR is useful, easy to elicit, and reliable. Decrease in phasic HR as well as HRV features have been measured and studied in an economic context. The use of an arousal parameter with the above modalities hence suggests being a tangible way to understand the state of the person based on visceral changes.

B. Emotion Detection for Financial Decision Support

Research in neuropsychology and behavioral decision making indicates a dualistic role of emotions. Damasio [10] shows that emotions are essential for making advantageous decisions, as they play a communicative role of providing cognitive guidance to the task at hand. Finucane *et al.* [15] coined the term affect heuristic, which brings out the effort-reducing role of emotions in the case of complex decision making. On the other hand, behavioral economists argue that emotions could be the intrinsic explanation for biases, such as the disposition effect [7], loss aversion [16], and in the case of auctions, the phenomenon of auction fever [4]. These biases have been shown in laboratory and field studies, wherein adopting a strictly rational and cognitive method is predictable by utility models, whereas emotional decision making could lead to substantial financial losses for the stakeholders.

Damasio [10] provides what has been named as the "primary" and "secondary" emotion models, establishing the neural underpinnings for the physiological responses to environmental stimuli, to the emotion generative system. Of interest is the explanation with respect to the path emotions take, in addition to bodily response, namely, the *feeling of emotion*, or rather the consciousness that an emotion is being experienced. Consciousness revolves around an enlarged scheme, i.e., instead of leaving complete control to the automated response systems in human beings, consciousness enables to bring these automated systems under control. Extrapolating this, one could also generalize the knowledge of emotions and decide, for example, to be cautious about any circumstance that resembles a system which leads one to go out of control allowing the automated system to dominate in a potentially regrettable manner. One could also exploit this knowledge, and turn the course of a potentially vulnerable situation, into an individual's advantage. In short, consciousness of emotions offers the flexibility of response based on the history of interactions with the environment [10].

Several methods facilitate the consciousness of emotional state. Live biofeedback of HR or SCR parameters to construct an arousal meter or indicator is a one possible solution. By providing relevant information on physiological parameters, live biofeedback improves the consciousness of emotional state. Peira *et al.* [17] demonstrate by means of HR biofeedback that participants could efficiently downregulate their HR to negative and neutral pictures when asked to do so. Peira *et al.* [18] showed that HR biofeedback training enables regulation of bodily aspects of emotion also when feedback is not available.

Feedback provided to the user by most biofeedback systems is based on physiological features to determine the "arousal" level of a user, such as a thermometer whose temperature is positively correlated with the EDA level based on [19], or arousal meters based on cardiac parameters, such as decrease in HR and HRV features [20–24]. However, what is missing in these biofeedback mechanisms is the *interpretation* of what this increase in arousal is likely to mean by learning the user state from previous interactions. In other words, based on semantic and contextual interpretation of the signals and with the aid of annotation techniques, it should be feasible to determine which emotional state a person is likely to be in (whether it is regret or rejoice). Hence, a context-specific biofeedback system would be able to interpret these features, and detect the affective state of the user.

C. Emotion Detection Using Heart Rate and Skin Conductance Response

Of the modalities that communicate emotion, speech (word choice, voice tone), motion (facial expression, gesture, posture), and physiology (SCR, HR, muscle tension), have proven to be reliable modalities. Physiology is considered to be a robust indicator, since it is hard to control by the user and nearly always available. The downside is that there could be a large amount of signal variance across and within people. As suggested by Blair [25] and Tanaka *et al.* [26], building a subject-independent emotion detection system, especially addressing the requirements of different age groups or gender is a nontrivial problem.

Haag *et al.* [27] describe a procedure to train computers to recognize emotions from the IAPS set using multiple signals from different biosensors. Wiens *et al.* [28] attempt to classify emotions of different valence (amusement, anger, and fear). Their results suggest that better self-perception is correlated with a higher experience of intense emotions. Using a database recorded from an actress, Sharma *et al.* [29] classified sadness, anger, happiness, and neutral state. They applied feature-level and decision-level fusion of auditory and visual data achieving up to an 89% classification rate for four emotions. Applying these findings, we attempt a fusion technique consisting of SCR and phasic HR features to classify two emotions associated with the consequences of trading decisions, namely "rejoice" and "regret."

D. Classifiers and Models for Emotion Detection

Emotion detection is typically performed using a range of machine learning algorithms, such as neural networks, support vector machines (SVM), linear discriminant analysis, quadratic discriminant analysis classifiers, and decision trees [30]. Lee *et al.* [31] employ a binary-tree-based classifier, where consecutive emotions are extracted at each node, yielding recognition rates of up to 72% for speaker-independent recognition of five emotions (anger, boredom, fear, sadness, joy, and a neutral state). While both SVM and decision trees would be suitable to achieve high accuracy, decision trees are less computationally and resource intensive than SVM's, and better suited for real-time analyses.

Regarding the annotation scheme, Devillers *et al.* [32] survey classifier algorithms employed, and develop a novel annotation scheme for labeling five emotions. In place of a singular emotion, they adopt a major/minor labeling method and propose a proportion of each emotion in this annotation scheme, resulting in blended emotions. The best detection score is obtained with SVM, where around 72% of correct detection between neutral and negative emotions was achieved. Based on this method, in addition to binary emotion models, we propose to employ the tri- and tetrastate model, which mainly allows for more expressiveness of emotions. The tristate allows three states: namely rejoice, regret, and a mixed state (which can be either both, or none of these emotions), while the tetrastate model allows for clear expressiveness of emotions when participants feel both regret and rejoice, or neither of these emotions. Hence, in comparing binary versus blended emotion models, the focus of this paper is primarily on decision trees and comparing accuracies across different types of decision trees.

III. Methodology

A. Participants

About 100 university students (Bachelors and Masters) participated, of which 78 were males and 22 females. Their ages ranged between 17–32 years and they were recruited using ORSEE [34]. The study is not restricted to expert financial traders, but focusses rather on private investors, who might or might not have sufficient knowledge and training, and hence, students were recruited to capture behavior of these (potentially) nonexpert traders [33].

B. Procedure

The study focuses on event-elicited emotions, wherein emotions are not self-imagined, but rather evoked explicitly due to the circumstance in the lab. Subjects know that their physiological data are being recorded, although they are not informed about the purposes of data collection, whereas there was no explicit instruction asking subjects to act like a trader, the instructions of the task were clearly explained, and they were told to trade the stocks endowed to them.

Two quantities of information are provided to the subject: the probability of price increase, and the stakes associated with the stock; with two levels for each variable (0.45 or 0.55 for probability, and €2 or €10 for stakes). Similar to the experimental design of Weber and Welfens [11], each subject is initially endowed a stock worth €100, and information about the stock's value after one trading round (whether it gained/lost). The subjects' task is to then decide whether to hold/sell this stock. After their decision, the participant is shown how the stock performed in the next trading round. If the participant held (sold) the stock, their final profit would be the loss (gain) after two (one) round(s) of trading. Since this profit is dependent on their decision, subjects experience either counterfactual regret or rejoice. At the end of each trial, participants reported their strength of emotions, namely rejoice and regret about the current trading decision, on a seven-point Likert scale (see Fig. 1). Stimuli consisted of the following: 1) stock information, 2) first stock trend, 3) decision screen, 4) second stock trend, 5) final payoff with emotion scales, and 6) a next button, to move to next period. Hence, emotions elicited were purely based on the information provided, the decisions made, and reactions to the outcomes of these decisions. Four practice rounds and thereafter 60 trading rounds were played, with no pauses in between. Hence, these six stimuli were each presented 60 times, without randomization, in the aforementioned order. The final profit from a chosen round (by roll of die) is paid out as real money at the end of the experiment in order to rule out cash balance effects.

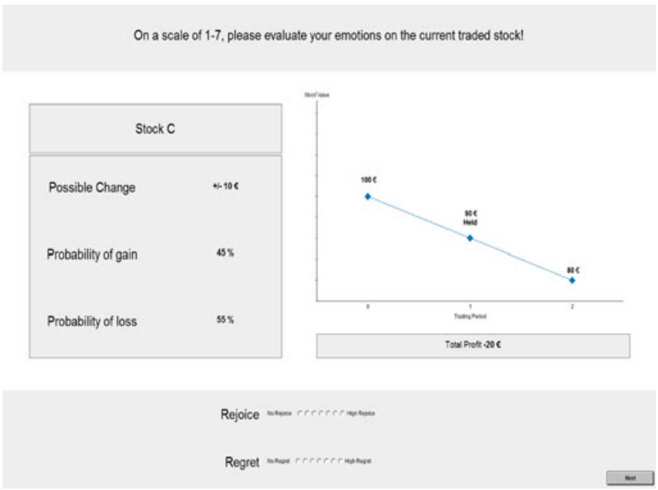

Fig. 1. Experiment screenshot.

TABLE I
FEATURE SPACE

Feature	Description
hr_preX	Normalized HR 0.x s before event, where x [1, 30] represents anticipatory emotions, hence 30 features
hr_postY	Normalized HR 0.y s after event, where y [1, 30] representing parasympathetic activity of immediate emotion
hr_postZ	Normalized HR 0.z s after event, where z [31, 70] representing sympathetic activity of immediate emotion
icd_subject	HR during the initial cool down phase
cda_tonic	Mean tonic activity within response window (wrw) of 3 s after the event
cda_nscr	Number of significant (= above – threshold) SCRs wrw
cda_ampsum	Sum of SCR-amplitudes of significant SCRs wrw [muS]
cda_scr	Average phasic driver wrw. Represents phasic activity wrw most accurately, not correlated to SCR amplitudes [muS]
cda_iscr	Area (i.e., time integral) of phasic driver wrw [muS∗s]
ttp_latency	Response latency of first significant SCR wrw [s]
global_mean	Mean SC value wrw
global_max_deflection	Maximum positive deflection wrw

C. Apparatus

The interface allows subjects to report blended emotions (or multiple emotions) at the same time. bioPLUX measurement [35] system (Version 2.0) was used, which enables simultaneous collection of both SCR and HR data. Data storage was carried out using an in-house plugin, which accesses the bioPLUX Java API, and stores the sensor output into a csv file.

D. Affect Recognition System

Analogous to Picard *et al.* [1], the building blocks in the affected recognition system are as follows: 1) preprocessing of signals, 2) feature extraction, 3) classification, 4) selection strategy, and 5) output generation. For preprocessing, both the HR and SCR signals were bandpass filtered and transformed to their appropriate ranges using MATLAB toolboxes. Features were extracted (see Table I) using Ledalab [36] for SCR and ECG MATLAB Toolbox for HR. HR was represented by 101 features, with one for the initial cool down, and 30 for the

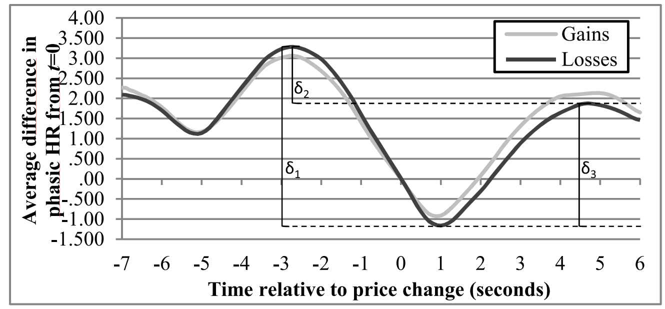

Fig. 2. Average difference in phasic HR from $t = 0$, denoting event of second price change.

average difference in HR, 3 s before and 7 s after the event of interest in the experiment, taken at every 0.1 s (rows 1–4 of Table I). We focus on the point in time when the subject observes the first price change prior to decision making and the second price change after deciding to hold/sell the stock.

Classification analyses were performed using the Weka toolbox [37]. Tenfold cross validation was used as a standard method for preventing overfitting and for pruning the decision trees. For the emotion detection and classification stage, we focus upon three types of decision trees, namely C4.5, CART, and random forest. The main distinction between the C4.5 and CART lies in the pruning strategy. C4.5 uses subtree rising to grow the tree if at least a predefined percentage of instances (10% is the default) are explained by addition of a node [37]. CART uses subtree replacement by means of *k*-fold cross validation to determine subtrees that lead to the smallest increase in error among the remaining subtrees in the current version of the tree and to prune them. The third algorithm is random forest that differs in the tree-growing method. Instead of an algorithm that greedily picks the best option at every tree iteration, it is randomized by picking one of the *N* best options instead of a single winner, or by choosing a random subset of options and picking the best from that [37].

E. Data Analysis

Lang *et al.* [38] use the R-wave pulses, to derive what is referred to as the triphasic cardiac waveform, consisting of a small initial deceleration, a larger midinterval acceleration, and a final deceleration. With the second price change event denoted at $t = 0$, Fig. 2 represents this triphasic waveform from the data in this experiment, by computing the average difference in HR for the second price change event. Prior to the price change event, participants experience a steady difference in HR, starting from approximately at $t = -3$. Due to the impact of the information, there is a point of minima in the window of $t = [0, 3\,\mathrm{s}]$, signifying the parasympathetic activity of the nervous system. The final phase is the sympathetic phase with a point of maxima in the interval $t = [3, 7]$. These findings are in tune with [39]–[41] which elucidate that all physiological reactions are considered relevant that occur up to 5 s before and up to 5 s after an event. The challenge that we currently address is to determine the most appropriate moving average window size for each of these three phases of the information event that best

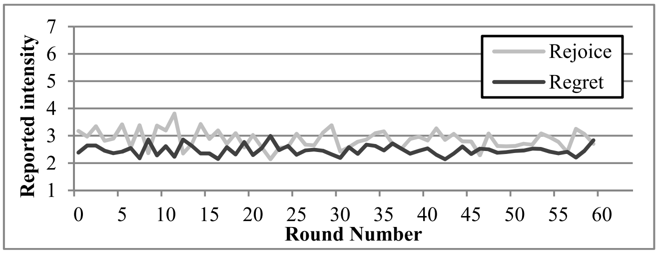

Fig. 3. Breakdown of reported emotions of rejoice and regret (on a scale of 1–7), in each trading round (from 1 to 60) averaged across 100 participants.

TABLE II
CLASSIFICATION RESULTS USING BINARY EMOTION CLASSES

Description of Classes (High/Low)	Feature Space (best window size)	Avg. Classification Rate (Weighted precision, Weighted F-Measure)
Rejoice	hr_preX(2.5 s)	65.93% (0.652, 0.634)
Regret	hr_preX(2.5 s)	60.55% (0.605, 0.605)
Rejoice	hr_postY(0.3 s)	66.91% (0.663, 0.648)
Regret	hr_postY(0.3 s)	61.16% (0.611, 0.609)
Rejoice	hr_postZ (0.1, 0.3, 0.5, 0.9 s)	67.43% (0.670, 0.652)
Regret	hr_postZ (0.1, 0.3, 0.5, 0.9 s)	60.66% (0.606, 0.605)

N = 5482 decisions using CAR, and tenfold cross validation.

represents the HR activity while experiencing gains and losses. Fig. 3 represents the breakdown of the range of emotion ratings, averaged across participants in the 60 rounds, on the scale of 1 to 7.

In order to represent this information, the average of the differences in HR is computed by assuming an appropriate moving average window length.

Based on the feature space outlined in Table I, we analyze if classifiers can be trained to recognize a binary class emotion problem, between high (scale >1) and low (scale $= 1$) rejoice, respectively, and similarly between high and low regret. The split was based on the distribution of the self-reported values in this particular class, and the median was taken to decide the best split applicable to both emotions. To select the moving average window lengths suitable for building the feature sets, phasic differences in HR were taken prior to and postprice change information events (in both parasympathetic and sympathetic phases). Specifically, for X seconds prior to the second price change event (hr_preX) window lengths of 1, 1.5, 2, 2.5, and 3 s were taken, for Y seconds after the second price change event in the parasympathetic phase (hr_postY), 0.1, 0.3, 0.5, and 0.7 s were compared, and for Z seconds after the second price change event in the sympathetic state (hr_postZ), 0.1, 0.3, 0.5, and 0.9 s were compared.

IV. RESULTS

A. Binary Emotion Detection in Financial Decision Making

For each event, the best classification accuracy with respect to the individual rejoice and regret event is reported in Table II.

TABLE III
CLASSIFICATION RESULTS USING BLENDED EMOTION CLASSES

Decision Tree Variant	Blended Emotion Model	Avg. Classification Rate (Weighted Precision, Weighted F-Measure)	Benchmark for n-class problem
CART	Tristate	44.00% (0.387, 0.384)	33%
C4.5	Tristate	41.96% (0.420, 0.420)	33%
Random Forest	Tristate	39.13% (0.371, 0.381)	33%
CART	Tetrastate	45.68% (0.447, 0.422)	25%
C4.5	Tetrastate	40.86% (0.410, 0.410)	25%
Random Forest	Tetrastate	32.76% (0.300, 0.300)	25%

N = 5482 decisions, tenfold cross validation.

The best accuracies for hr_preX were obtained with a window length of 2.5 s, hr_postY with 0.3 s, and for hr_postZ, all window lengths were classified with almost similar accuracy, precision and F-measures. Other window lengths predicted the binary emotions with up to 5% lesser accuracy than the reported values, and with a lower weighted precision or F-measure.

B. Blended Emotion Detection in Financial Decision Making

Several participants were found to report a high/low intensity of both emotions (regret and rejoice) in the same trial. About 36% of emotions were indicated as rejoice, 23% as regret, and 40% are reported conflictingly (as both low, and both high).

Table III reports overlaps of self-perception with measured physiology for the tri- and tetrastate blended emotion. The best obtained average accuracies for the tristate mode were up to 44% (denoting 11% improvement over the chance average accuracy of 33%) using the CART algorithm, and by a fusion of HR and SCR features with a weighted precision of 0.387, and a weighted F-measure of 0.384. Not shown in the table is that rejoice was predicted with higher precision (0.437) than regret (0.261), and the blended class had the highest precision of (0.496). Since this blended state contains both the states of "two high intensity emotions," as well as "two low intensity emotions," we break them further into a tetrastate model by annotating four classes, "rejoice," "regret," "both," and "none." The best obtained accuracies for the tetrastate mode were up to 45.68% (denoting 20% improvement over the chance average accuracy of 25%) using the CART algorithm, with a precision of 0.447 and an F-measure of 0.422. A detailed breakup of the obtained true positive and false positive rates are summarized in Table IV. Finally, the emotion "rejoice" is consistently classified with higher accuracy than "regret" in the final blended model, as indicated by the precision and F-measures in Table IV. The states of "none" and "both" are classified with high precision (0.619 and 0.49, respectively).

As these results are obtained using a large subset of features, we consider the following values to represent the features to reduce the feature size

$$\delta 1 = \max(\mathrm{avg}(\Theta\mathrm{HR}[-3,0]) - \min(\mathrm{avg}(\Theta\mathrm{HR}[0,3])) \quad (1)$$

$$\delta 2 = \max(\mathrm{avg}(\Theta\mathrm{HR}[3,7]) - \max(\mathrm{avg}(\Theta\mathrm{HR}[-3,0])) \quad (2)$$

$$\delta 3 = \max(\mathrm{avg}(\Theta\mathrm{HR}[3,7]) - \min(\mathrm{avg}(\Theta\mathrm{HR}[0,3])) \quad (3)$$

TABLE IV
CLASSIFICATION RESULTS USING TRI- AND TETRASTATE MODEL

Annotation scheme	Benchmark accuracy	Simple CART	Random Forest	C4.5
Binary Rejoice (Low, High)	50%	65.35%	61.28%	64.49%
Binary Regret (Low, High)	50%	60.35%	56.32%	61.10%
Tristate model (Rejoice, Regret, Blended)	33%	44.84%	42.03%	42.54%
Tetrastate model (Rejoice, Regret, Both, None)	25%	44.59%	37.38%	41.74%

N = 5482 decisions, tenfold cross validation.

TABLE V
CLASSIFICATION RESULTS USING THREE DELTA FEATURES

Emotion	TP Rate	FP Rate	Precision	F-Measure	ROC Area
Rejoice	0.747	0.546	0.437	0.552	0.632
Regret	0.119	0.086	0.301	0.171	0.567
Both	0.309	0.102	0.490	0.379	0.695
None	0.522	0.062	0.619	0.567	0.832
Overall	0.457	0.253	0.447	0.422	0.664

CART, N = 5482 decisions, No. of Classes = 4; TP Rate = True Positive Rate; FP Rate = False Positive Rate, Precision = TP Rate/(TP Rate + FP Rate); ROC Area = TP Rate versus FP Rate.

where Θ denotes the difference in HR of every 10th s beginning with the start range to the end range (see Fig. 2); $\delta 1$ and $\delta 3$ represent the difference between the maxima and minima,$\delta 2$ between two maxima. As shown in Table V, accuracies close to those obtained by choosing best moving average window lengths can be obtained using this feature selection method as well, hence, reducing the feature size as well as the computation cost for obtaining these accuracies.

V. DISCUSSION

An experiment was designed to build and evaluate emotion detection systems in a financial decision making context. The uniqueness of this study lies in that it allows the expression of emotions in binary and blended states based on their physiological changes, especially when conflicting emotions occur due to consecutive gains and losses. Since emotions have been found to impact decision making positively as well as negatively [41], [42], such an evaluation is necessary for understanding decision behavior. The combinations of best window size, best annotation method, and best algorithm are derived in terms of classification accuracy. As an addendum, instead of using 100 features, the results are replicated by three proposed delta features based on the triphasic cardiac form. While the accuracies of the tri- and tetrastate models are higher than the benchmark models, the results indicate that the training data are optimistic compared with what might be obtained from an independent test set from the same source. Hence, it would be possible to improve the training dataset by including a balanced number of gains and losses for each treatment.

The findings of this study could be useful in designing biofeedback or ambient systems, since it is computationally less demanding to derive a single parameter for a subject to be fed into the decision tree. These results are however based on a fixed number of trading interactions per subject and, hence, will have to be extended by increasing the number of trades, either by increasing the interaction with the interface, or by moving to a context outside the laboratory. Additionally, the classifier is developed based on the self-reported emotions of participants in earlier trading round. Self-reported scales are desirable, since they are less costly, easier to administer, and take considerably less time to complete than performance tests [39]. However, they are problematic, since respondents can provide socially desirable responses rather than truthful ones, or are more susceptible to faking than performance tests. While ground truth is difficult to elicit and obtain, we employ the method of self-reports in this context to train the blended emotion detector after the emotion has been experienced, but not during the moment of experience. This is based on the assumption that subjects have better interoceptive skills postexperience than during emotion experience. The proposed emotion detector on the other hand is expected to detect emotions during emotional experience, which a person might not be able to perceive well, due to his/her involvement in the task, and would involve obtrusion of the emotion. In addition, analyzing and reporting blended emotions, or to understand how emotional reactions change over time, has been referred to as a more advanced emotional intelligence skill than a basic one [40]. Hence, in order to detect complex blended emotions, we utilize the discrete emotion labels of participants, to infer and learn blended states, wherever possible. It cannot be assumed that the discrete states are labeled accurately; however, this observed lack of clarity in the discrete states is what justifies building a blended detector.

In implementing emotion detection in financial decision systems, several challenges remain to be addressed: 1) incorporating suitable real-time methodologies for integrating emotion detection and biofeedback, 2) calibrating systems to make the system both person-independent yet sufficiently capable to learn and detect blended emotions for an individual (which might also need to factor in individual differences in personality, age, gender, or the type of emotion regulation strategy a person employs), 3) external validity in trading systems, since the speed of transactions are very high in trading systems and online markets, maintaining affective data and determining the right point in time to display the detected emotion might pose a challenge, 4) in relation to the previous point, it remains to be addressed whether subjects would have sufficient time and mental bandwidth to factor in the affective information, and how far this technology will be accepted, remains a question to be empirically investigated, and 5) while traders might be definitely more experienced in regulating their emotions, financial decisions are also made by private investors and by consumers in several contexts. As observed in the works of Dane and Pratt [43] and Dane *et al.* [44], experts have shown to have the specialized ability to leverage intuitive decision making for achieving managerial effectiveness. The distinction between traders and nontrained financial decision makers is an interesting line of research, and would have to be implemented as a field experiment.

Finally, by considering the triphasic cardiac waveform, we propose three combinations, namely, the difference between the maxima and minima in the anticipatory–parasympathetic, anticipatory–sympathetic, and parasympathetic–sympathetic phases. Emotion detection accuracies obtained earlier were replicated with these three delta values, and it was observed that accuracy obtained using the complete HR feature set (with 100 features) can be replicated by using only these three computed values. These results could go toward building a suitable affect-based real-time decision support system to assist users to understand their emotions better before making their decisions, as well as in understanding hidden user states.

This paper has employed the modalities of ECG and EDA to approximate the user state. Multimodal fusion techniques including other sensors, such as electromyography, electroencephalography, and blood flow, could help in improving the classification accuracy. Further, the self-evaluation method needs to be extended to include the possibility to report an explicit "blended" emotion state in order to enable better classification of these user states using physiology. In addition, decision trees have been employed to model possible emotional states of a subject. An extension would be to apply the SVM algorithm, which has been used in the affective computing domain, and compare accuracies obtained with decision trees.

References

[1] R. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 10, pp. 1175–1191, Oct. 2001.

[2] M. M. Bradley, P. J. Lang, and B. N. Cuthbert, "International affective picture system (IAPS): Digitized photographs, instruction manual and affective ratings," Tech. Rep. A-6, Center for Research in Psychophysiology, Univ. Florida, Gainesville, FL, USA, 2005.

[3] M. M. Bradley and P. J. Lang, "The International affective digitized sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual," Tech. Rep. A-8, Center for Research in Psychophysiology, Univ. Florida, Gainesville, FL, USA, 2007.

[4] M. T. P. Adam, J. Krämer, C. Jähnig, S. Seifert, and C. Weinhardt, "Understanding auction fever: A framework for emotional bidding," *EM Int. J. Electron. Markets*, vol. 21, no. 3, pp. 197–207, 2011.

[5] Bechara, "Deciding advantageously before knowing the advantageous strategy," *Science*, vol. 275, no. 5304, pp. 1293–1295, 1997.

[6] L. Krain, A. M. Wilson, R. Arbuckle, F. X. Castellanos, and M. P. Milham, "Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making," *NeuroImage*, vol. 32, no. 1, pp. 477–484, 2006.

[7] H. M. Shefrin and M. Statman, "The disposition to sell winners too early and ride losers too long: Theory and evidence," *J. Finance*, vol. 40, no. 3, pp. 777–790, 1985.

[8] G. Ku, D. Malhotra, and J. K. Murnighan, "Towards a competitive arousal model of decision-making: A study of auction fever in live and Internet auctions," *Org. Behav. Human Decision Processes*, vol. 96, no. 2, pp. 89–103, 2005.

[9] S. Rick and G. F. Loewenstein, "The role of emotion in economic behavior," in *Handbook of Emotions*, 3rd ed. New York, NY, USA: Guilford, 2007.

[10] R. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain*, 17th ed. New York, NY, USA: Quill, 2003.

[11] M. Weber and F. Welfens, "Splitting the disposition effect: Asymmetric reactions towards 'selling winners' and 'holding losers'," *SSRN Electron. J.*, 2008, doi: http://dx.doi.org/10.2139/ssrn.1176422

[12] S. J. Humphrey, "Feedback-conditional regret theory and testing regret-aversion in risky choice," *J. Econ. Psychol.*, vol. 25, no. 6, pp. 839–857, 2004.

[13] P. J. Astor, M. T. P. Adam, C. Jähnig, and S. Seifert, "Measuring regret: Emotional aspects of auction design," presented at the Eur. Conf. Inform. Syst., Helsinki, Finland, 2011, p. 88.

[14] J. M. Fernandez, J. C. Augusto, R. Seepold, and N. M. Madrid, "A sensor technology survey for a stress-aware trading process," *IEEE Trans. Syst, Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 809–824, Nov. 2012.

[15] M. L. Finucane, A. Alhakami, P. Slovic, and S. M. Johnson, "The affect heuristic in judgments of risks and benefits," *J. Behav. Decision Making*, vol. 13, no. 1, pp. 1–17, 2000.

[16] P. Sokol-Hessner, M. Hsu, N. G. Curley, M. R. Delgado, C. F. Camerer, and E. A. Phelps, "Thinking like a trader selectively reduces individuals' loss aversion," *Proc. Nat. Acad. Sci.*, vol. 106, no. 13, pp. 5035–5040, 2009.

[17] N. Peira, M. Fredrikson, and G. Pourtois, "Controlling the emotional heart: Heart rate biofeedback improves cardiac control during emotional reactions," *Int. J. Psychophysiol.*, vol. 91, no. 3, pp. 225–231, 2014.

[18] N. Peira, G. Pourtois, M. Fredrikson, and M. Gray, "Learned cardiac control with heart rate biofeedback transfers to emotional reactions," *PLoS ONE*, vol. 8, no. 7, p. e70004, 2013.

[19] W. Boucsein, *Electrodermal Activity*. Boston, MA, USA: Springer Science+Business Media LLC, 2012.

[20] P. J. Astor, M. T. P. Adam, P. Jerčić, K. Schaaff, and C. Weinhardt, "Integrating biosignals into information systems: A NeuroIS tool for improving emotion regulation," *J. Manage. Inform. Syst.*, vol. 30, no. 3, pp. 247–278, 2013.

[21] B. Wansink, *Mindless Eating: Why We Eat More Than We Think*. New York, NY, USA: Bantam Books, 2007.

[22] R. H. Thaler and S. Benartzi, "Save more tomorrow TM: Using behavioral economics to increase employee saving," *J. Political. Econ.*, vol. 112, nos. 1/2, pp. S164–S187, 2004.

[23] L. Chittaro and R. Sioni, "Affective computing vs. affective placebo: Study of a biofeedback-controlled game for relaxation training," *Int. J. Human-Comput. Stud.*, vol. 72, nos. 8/9, pp. 663–673, 2014.

[24] H. C. Lo, P. C. Lin, and K. P. Lin, "Development of a wearable biofeedback emotional control vest for older adult," *Adv. Mater*, vol. 680, pp. 466–472, 2013.

[25] S. N. Blair, "Physical fitness and all-cause mortality," *J. Amer. Med. Assoc.*, vol. 262, no. 17, pp. 2395–2401, 1989.

[26] H. Tanaka, K. D. Monahan, and D. R. Seals, "Age-predicted maximal heart rate revisited," *J. Amer. College Cardiol.*, vol. 37, no. 1, pp. 153–156, 2001.

[27] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion recognition using bio-sensors: First step toward an automatic system," presented at the Affective Dialog Syst., Tutorial Res. Workshop, Kloster Irsee, Germany, 2004.

[28] S. Wiens, E. S. Mezzacappa, and E. S. Katkin, "Heartbeat detection and the experience of emotions," *Cognit. Emotion*, vol. 14, no. 3, pp. 417–427, 2000.

[29] R. Sharma, T. Darrell, M. Harper, G. Lazzari, M. Turk, C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, "Analysis of emotion recognition using facial expressions, speech and multimodal information," in *Proc. 6th Int. Conf. Multimodal Interfaces*, 2004, pp. 205–211.

[30] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal Process. Mag.*, vol. 18, no. 1, pp. 32–80, Jan. 2001.

[31] C.-C. Lee, E. Mower, C. Busso, S. Lee, and S. Narayanan, "Emotion recognition using a hierarchical binary decision tree approach," *Speech Commun.*, vol. 53, nos. 9/10, pp. 1162–1171, 2011.

[32] L. Devillers, L. Vidrascu, and L. Lamel, "Challenges in real-life emotion annotation and machine learning based detection," *Neural Netw.*, vol. 18, no. 4, pp. 407–422, 2005.

[33] D. Friedman, *Experimental methods: A Primer for Economists*. Cambridge, U.K.: Cambridge Univ. Press, 1994.

[34] B. Greiner, "An online recruitment system for economic experiments," in *Research and Scientific Computing* (*Forschung und wissenschaftliches Rechnen)*, K. Kremer and V. Macho Eds. Göttingen, Germany, 2004, vol. 63, pp. 79–93.

[35] Q. Yang, P. Tang, R. Gu, W. Luo, and Y.-j. Luo, "Implicit emotion regulation affects outcome evaluation," *Social Cognit. Affective Neurosci.*, to be published.

[36] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *J. Neurosci, Methods*, vol. 190, no. 1, pp. 80–91, 2010.

[37] E. Frank and I. H. Witten, *Data Mining: Practical Machine Learning Tools and Techniques With Java Implementations*. San Francisco, CA, USA: Morgan Kaufmann, 1999.

[38] M. M. Bradley, L. Miccoli, M. A. Escrig, and P. J. Lang, “The pupil as a measure of emotional arousal and autonomic activation,” *Psychophysiology*, vol. 45, no. 4, pp. 602–607, 2008.

[39] M. A. Brackett, S. E. Rivers, and P. Salovey, “Emotional intelligence: Implications for personal, social, academic, and workplace success,” *Social Personality Psychol. Compass*, vol. 5, no. 1, pp. 88–103, 2011.

[40] J. D. Mayer, P. Salovey, D. R. Caruso, and G. Sitarenios, “Measuring emotional intelligence with the MSCEIT V2.0,” *Emotion*, vol. 3, no. 1, pp. 97–105, 2003.

[41] B. Shiv, G. Loewenstein, A. Bechara, H. Damasio, and A. R. Damasio, “Investment behavior and the negative side of emotion,” *Psychol. Sci.*, vol. 16, no. 6, pp. 435–439, 2005.

[42] M. T. P. Adam, J. Krämer, and C. Weinhardt, “Excitement up! price down! measuring emotions in dutch auctions,” *Int. J. Electron. Commun.*, vol. 17, no. 2, pp. 7–39, 2012.

[43] E. Dane and M. G. Pratt, “Exploring intuition and its role in managerial decision making,” *Acad. Manage. Rev.*, vol. 32, no. 1, pp. 33–54, 2007.

[44] E. Dane, K. W. Rockmann, and M. G. Pratt, “When should I trust my gut? Linking domain expertise to intuitive decision-making effectiveness,” *Org. Behav. Human Decision Processes*, vol. 119, no. 2, pp. 187–194, 2012.