Emotion Modeling and Machine Learning in Affective Computing

Keeyoung Kim
Advisor: Pradipta De
Computer Science Department, Stony Brook University
(SUNY Korea, Song-do, South Korea)
kykim@sunykorea.ac.kr

ABSTRACT
Affective computing is a computing related to, arise from, or influences emotions. Various emotion modeling and machine learning methods are used in affective computing. To explain human emotional states, psychologists developed various emotion models. Their models and methods have been adapted to affective computing. Machine learning, a field of study that gives computers the ability to learn without being explicitly programmed, is also essential to make computers deal with uncertain object like emotion. In this survey, we first review several emotion modeling techniques. Then we review machine learning techniques which is used in affective computing.

Keywords
Emotion modeling, Machine learning, Affective computing

1. INTRODUCTION
Rosalind W. Picard, the pioneer of affective computing, stated that affective computing is a computing related to, arise from, or influences emotions [50]. When Picard started her affective computing research in 1990s and she insisted that computers may have the ability to express, recognize, and have emotions, the reaction from peers was a LOL (Laugh Out Loud) in the phrase of herself [51]. Now it is popular, and there is an IEEE transactions on Affective Computing. Nowadays in IT industries, companies more and more wants to provide personalized services to their customers using big data from web, e-mail, and market checkout. In 2000s, the number of smartphone users explosively increased, and it will give us another chance to get numerous, unobtrusive, and inexpensive sensor data from human. And even if wearable devices are increased, the amount of the data from human will take off. This increase of quantity and improvement of quality of human data will give us a great chance for affective computing.

Picard, in her paper [50], stated that affective computing will be used in various applications like assisted learning, perceptual information retrieval, arts and entertainment, - so-called "affective gaming", which is a good example of affective computing applications - health and human computer interaction. Izard also mentioned that emotional limbic brain influences plays a role in recognition function of our brain. All sensory inputs, external visceral, must pass through the limbic brain before being redistributed to the cortex for analysis [26]. Emotion recognition and feedback from teacher to students, from performer to audiences, from doctor to patients helps to make them more focus on, or more satisfy [50]. This is the reason that emotion recognition will help computers to do a better job in these applications. Emotion recognition and feedback from a computer system to students, audiences, patients, gamers will increase the achievement of the system.

To build a system to understand human emotion, it will be better to understand human emotion by ourselves. From the antiquity to the Second World War, philosophers, biologists, psychologists tried to explain emotional phenomena of human. In antiquity, Plato suggest that the soul has a tripartite structure, composed of the separate and opposing areas of cognition, emotion, and motivation. Aristotle argued for the impossibility of such a separation and for the assumption of an interaction between the different levels of psychological functioning. Descartes, the father of modern philosopher, insisted that mental and physiological processes have to deal with at the same time. Charles Darwin, in his book "The Expression of Emotion in Man and the Animals" (1872) emphasized the expression of emotion in face, body and voice with intercultural studies and developmental approaches. William James, an American psychologist who was also trained as a physician, suggested that the emotion is the perception of differentiated bodily changes, specific for each emotion. In contemporary, psychologists are interested in emotion elicitation and differentiation, patterning of reaction modalities, and proprioceptive emotional systems [56].

Today's affective computing research can be classified into two main areas according to direction of research. The first case is a research for AI agent. They are interested in building AI agent who can achieve emotional reaction against the change of environment. They are interested in emotion arousal process. Emotion arousal theories in psychology, like James-Lange Theory, Cannon-Bard Theory, Schachter-Singer's Two-Factor Theory, are used in these systems. Computationally, rule-based agent-environment model is commonly used for these systems. Though this is also interesting, but I will not cover this issue, because normally this research focused on artificial agent for simulations or games in specific situations, rather than a research in real world human affect. One thing I just want to remark about this case, reinforcement learning in machine learning will be naturally well suited for this kind of agent-based model, and often used. The second case is machine learning research as a tool of psychology. It means the case that psychologists used machine learning algorithm - generally supervised learning - as a tool for their research, or computer scientists used human emotion data as dataset to study. They were generally interested in building systems, which may have a webpages, smartphones or dedicated mobile devices to detect human conditions and provide 'proper' service in the context of user's emotional status, to make questions and get feedbacks, and so on. Machine learning is surely a proper tool to make computers study human emotion. ML enables that computer can deal with something that is not clearly programmed - like emotion. As the capacity of data grows, machine learning will be more and more emphasized. We will review machine learning techniques applied in the context of affective computing in the later chapter.

Besides these two directions, I want to talk about more fundamental issue about human affect and its research. Let's take the example of computational social science. Computational social science is an emerging research area, which rise about the same time with affective computing. Computational social science is a field which try to study social issues with the methodology of natural science. Gary King, in his article about computational social science in Science Magazine [30], talks about the scientific methodology to study social phenomenon and behavior of human. He mentioned about the success of science in the past centuries, and the limitation of existing humanities researches - to quote his sentence, "In sharp contrast, the (smaller number of) social scientists did not mention a single problem they thought might be addressed, much less solved, or any inventions or discoveries on the horizon." He insisted that it is time to use powerful methodology of science to study human more deeply, and concisely. Though the word, 'computational social science' has broader meaning than 'social computing' sometimes, but it is still true that they are targeting more fundamental issues. This can be equally applied for the affective computing. Though we can just apply ML techniques to data and models from psychological research, it would be much better, if we verify that the data, model, process of psychology suits for the scientific methodology. Building a reliable methodology to research about human (including observation, measuring, processing, feedback) is another important role to play for computer scientists who study affective computing. Galileo Galilei left a famous quote, "Measure what is measurable, and make measurable what is not so." We have to critically think that data is measured properly, even which is about emotions.

In this paper, I want to provide better understanding of state of the art emotion modeling and machine learning in affective computing, and investigate what is the challenges to get better understanding of human emotion from the perspective of computer science. The following sections are organized like this. In section 2, emotion modeling methods will be reviewed. In section 3, machine learning methods which are used in affective computing will be selectively reviewed. In section 4, we will propose issues of affective computing should be tackled, related to emotion modeling and machine learning. Then, conclusion will follow.

2. EMOTION MODELING IN AFFECTIVE COMPUTING

How can we make the emotion measurable? Let's briefly review major existing models for emotion research in psychology. In psychology, many models were developed and improved to explains affective states of human. Scherer, Rodriguez, Marsella each summarized emotional representation models which are used for affective computing. Scherer [56] classified emotion models into four categories, dimensional, discrete, meaning oriented, componential. Rodriguez [54] classified them into basic emotions, appraisal theory, core effect of emotions, PAD space model. Finally, Marsella [43] classified them into appraisal, dimensional, anatomical, rational. Among them, we will briefly cover five emotion models, discrete model, appraisal model, dimensional model, and circuit model, componential model.

2.1 Discrete model and Basic emotions

Basic emotions are widely used for emotion research in psychology or physiology. Ekman presented six basic emotions (anger, surprise, happiness, disgust, sadness, and fear) which are common regardless of cultural difference in the world in 1960s. He researches emotional expressions of face, and how people in different cultural environment recognize the facial expressions. And he found out that those six basic emotions are commonly recognizable in most of cultures [15]. In 1990s, he expanded it into 15 emotions like amusement, anger, contempt, contentment, disgust, embarrassment, excitement, fear, guilt, pride in achievement, relief, sadness (or distress), satisfaction, sensory pleausure, and shame [14]. Ekman's Basic emotions research has an important meaning that showed that emotions are cross-cultural, though anthropologist showed that some of emotions are cultural, and language has strong influence to categorize feeling and emotions into several (language-defined) classes of emotions later. But there are also limitations that it is hard to find out which emotions are basic, - fundamental or primary, and others are additional or subordinate, and even could be made of combinations of basic emotions. It is even hard to say what are qualifications for the basic emotions. Laird and Oatley(1989) [27] proposed five basic emotions (anger, despair, shame, anxiety, happiness). And there are also another version of six basic emotions (excited, tender, scared, angry, sad, happy), and 8 basic emotion categories (ecstasy, admiration, terror, amazement, grief, loathing, rage, vigilance) and 4 stage for each category. A recent biological study says that there are four basic emotions like happy, sad, afraid (or surprised), and angry (or disgusted) [35]. D'Mello argued that Ekman and other researcher's tried to decide so-called basic emotions, but basic emotions can varies over cases of environment and situation [12]. Though it has clear limitations, it still has a strong basis that our thought categorized by getting influences from words of language. Discrete model is widely used in psychology and affective computing research because of its simplicity. Almost all affective computing research uses dataset which has language-defined emotion class labels. Besides this, many affective computing researches used discrete model. Table 1 shows existing publications presenting research progress in affective computing and the emotion modeling methods applied in each of these works. Maxhuni, Hosub, Haque, Attabi, Chang, Schuller, Rachuri, Kolodyazhniy, Wu [7, 21, 22, 31, 37, 44, 52, 58, 63] all used discrete emotion models for their research. Though they used different number and kind of emotion classes, all of them has the same structure - fixed number of class labels, classification into one of them using machine learning. In addition, many of them have survey style data collection interfaces.

2.2 Appraisal model

The appraisal model is developed by Smith and Lazarus [59], which suggests that emotions arise from the dynamic interaction of appraisal and coping processes that depend on the agent's representation of its relationship with the environment. According to appraisal theories, the individual environment relationship is evaluated using a certain number of appraisal dimensions or variables, and connections between environmental changes, emotion changes, and cogni-

Table 1: Emotion models in affective computing papers

Paper	Emotion model	Continuous/discrete	How to collect
LiKamWa [39]	Circumplex model (dimensional)	discrete	choosing PA sliders in their apps 4 times a day
LiKamWa [40]	Circumplex model (dimensional)	discrete	choosing PA sliders in their apps 4 times a day
Maxhuni [44]	POMS (discrete)	discrete	EMA (survey)
Lee [37]	six basic emotions (discrete)	discrete	survey
Haque [25]	six basic emotions (discrete)	discrete	
Attabi [7]	five basic emotions (discrete)	discrete	FAU AIBO Emotion Corpus
Yang [64]	dimensional model	discrete	LJ40k, MER31k dataset
Chang [21]	3 emotional states*3 levels (discrete)	discrete	SUSAS, Belfast Naturalistic Database
Lee [36]	Uni-dimensional	continuous	call-center data 1187 calls
Hernandez [23]	Smile intensity (uni-dimensional)	continuous	JAFFE dataset
Schuller [58]	7 emotions + neutral (discrete)	discrete	EMO-CAR
Rachuri [52]	14 emotional states	discrete	Emotional Prosody Speech and Transcripts library
Kologyazhniy [31]	2 basic emotions + neutral (discrete)	discrete	sensors (facial and muscle activity)
Wu [63]	3 basic emotions + neutral (discrete)	discrete	

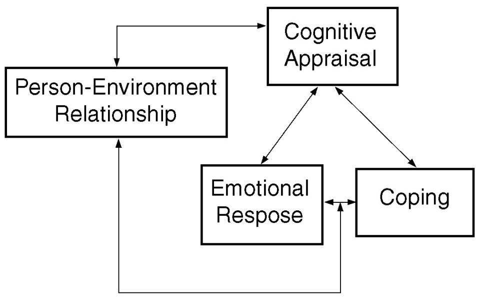

Figure 1: Appraisal model [1]: Appraisals which is caused by environment raise emotional responses.

tive responses.

- **Relevance** How relevant is the event for me? Does it directly affect me or my social reference group?
- **Implications** What are the implications or consequences of this event and how do these affect my well-being and my immediate or long-term goals?
- **Coping potential** How well can I cope with or adjust to these consequences?
- **Normative significance** What is the significance of this event with respect to my self-concept and to social norms and values?

Appraisal variables has "if-then" rules. As mentioned above, appraisal model is dominant theory about human emotion in computer science, especially it is using to build symbolic AI systems. AI agent for specific situation (like game or simulation) is a representative example [17, 18]. We will not cover this case here.

2.3 Dimensional model

Another most widely used emotion model is dimensional model, which denotes emotional states as points in a continuous dimensional space. Mehrabian and Russell presented and used this model for their research [45,55]. For simple classification, there are two kinds of dimensional models, uni-dimensional and multi-dimensional. Uni-dimensional model has just one dimension. PANAS (positive and negative affect scales) [62] is a popular uni-dimensional model. Multi-dimensional model has two or three dimensions. PAD space model is developed by Mehrabian and Russell (1974), which has three dimensions correspond to Pleasure (a measure of valence), Arousal (indicating the level of affective activation) and Dominance (a measure of power or control) [46]. Cowie concluded that that this kinds of continuous space is known to be has better performance in out-of-lab experiment than discrete models [19]. Though the continuous model shows good performance, it still has plenty of things left to research to improve the performance of affective computing. For example, though PAD (pleasure - arousal - dominance) or PA dimension is widely used and accepted by psychologists, Kaernbach showed that the euclidian space with P-A axis might not reflect the real world [28]. He analyzed data from Bradley & Lang's experiment [32], and found that they scattered like V-shape and showed some holes in space, even some obviously different type of emotion could be positioned very closely in PAD space. He claimed that this show the space could be distorted and non-euclidean space. And even six basic emotions could be mixed in PA space, some of them could not be differentiable with others. It is possible that PA space does not have enough power to differentiate emotions from each other. To explain this, he introduced "metamerism", a concept from colorimetry, in short, it means if a point is not indistinguishable with another point, then the two points have to be the same point, otherwise the dimensional space has not enough discriminative power. PA space does not follow metamerism. Lee [36] used uni-dimensional model which has from non-negative(0) to negative(1) continuous values. LiKamWa, and Yang [39, 64] used two-dimensional model with PA axis. In this context, there is an interesting recent research, which proposed a new dimensional model, "Lövheim Cube of emotion" [41]. It explains the relationship between the monoamine neurotransmitters and the emotions. It looks like three dimensional model using three neurotransmitters instead of P, A, D axis. Three monoamine neurotransmitters are serotonin, norepinephrine, dopamine. These monoamine neurotransmitters have close relationships with human emotion. Serotonin is related to obsession and compulsion. Norepinephrine is related to alertness and concentration. Dopamine is re-

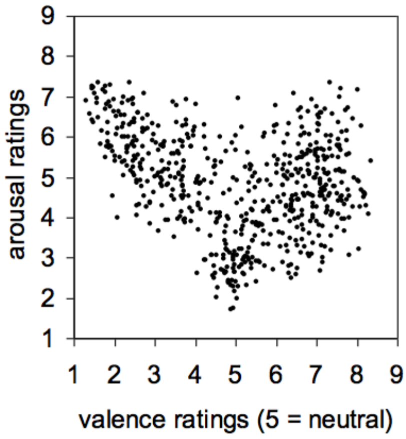

Figure 2: Plots of International Affective Picture System [28,32] in PA space: It is V-shaped.

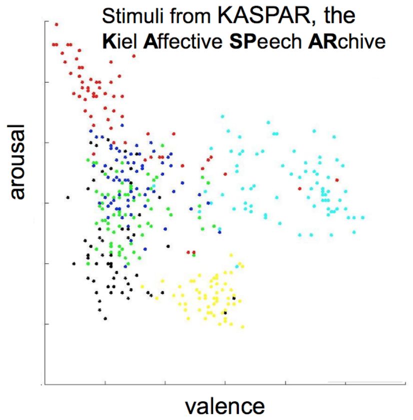

Figure 3: KASPAR in PA space [28]: Except anger:red, joy:cyan, and neutral:yellow, other emotions are not differentiated. (sad:black, fear:blue, disgust:green, suprise:grey)

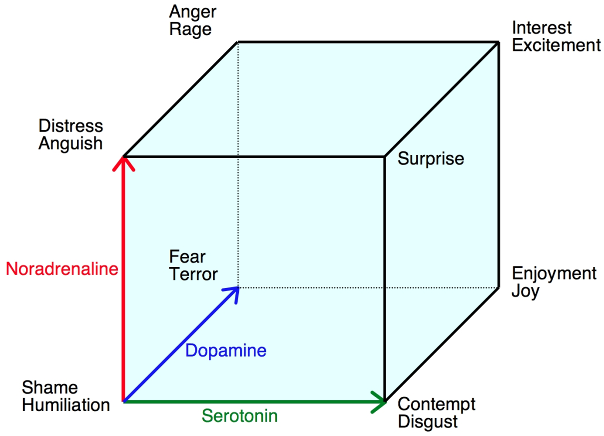

Figure 4: Lövheim Cube of emotion [41]: a dimensional model using monoamine neurotransmitters as its dimensions.

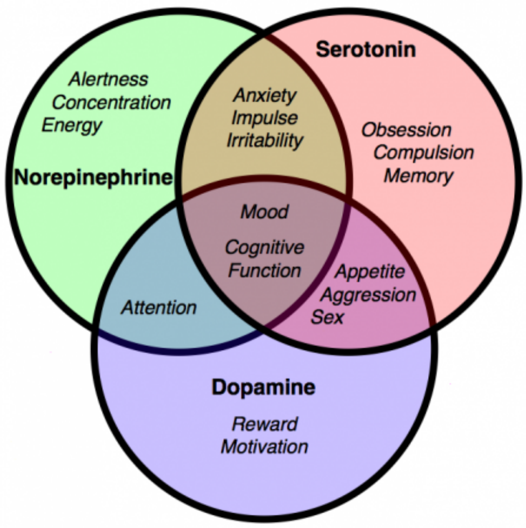

Figure 5: Effect of three monoamine neurotransmitters, serotonin-norepinephrine-dopamine [5]

lated to motivation. Detailed description about these neurotransmitters are below. This is very promising that these neurotransmitters will be measurable [8]. This will be almost direct measurements of human emotion.

2.4 Circuit model

Circuit model, which is also called anatomical model, is proposed by a neuroscientist, Joseph E. Ledoux. He suggested that each emotion is considered to processing in a different, discrete neural circuits and emphasized processes or systems associated with these circuits. Neuropsychologists believe that fundamental emotions and their differentiation are determined by evolutionary neural circuits in a brain. They found several fundamental circuits (so-called survival circuits) of primitive emotions like rage, fear, expectancy, and panic. Ledoux even insisted that the differentiation of emotion by language, could be unimportant. The important thing is that there are real different circuits (path) to compute different emotions in our brain [34]. It seems that this different processing circuits for each emotions are like labels or classes of each emotions. Though circuit model is less known than above models and still limited to target only primitive emotions, it seems that circuit model has a great potential, because it is based on objective observations. When it started, the research relies on signals of the microelectrode which is pinned to mouse brain. For now, medical technology, especial-

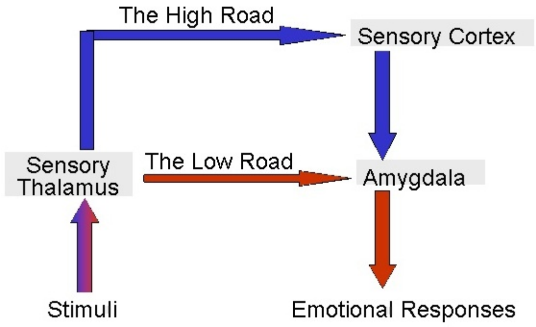

Figure 6: Survival circuit is belong to the low road, which is for faster and instinctive response [11].

ly medical imaging field, is developed dramatically, so that we can use fMRI to see functioning of the brain. Though it has a limitation that only a few primitive emotions related circuits are examined, we may apply circuit model just to primitive emotions apart from higher-level emotions where recognition process and memory are engaged. Otherwise, we may use this to study emotions of animal using stimulation, physiological signals, and fMRI images.

2.5 Componential model

Componential models assumed that emotions are elicited by a cognitive (but not necessarily conscious or controlled) evaluation of situations and events and that the patterning of the reactions in the different physiological responses, face expression, gestures, postures, and feelings [56]. There are as many different emotional states as there are different cognitive appraisals in the processing of body reactions. Lazarus also argued that each emotion probably has its own unique somatic response pattern [33]. This can be a basis for or a bridge to physiological studies that emotion appraisals and explicit physiological reaction patterns. Many physiologists [9,31,42] were interested in research about the emotional state and physiological reactions. Up to now, this studies are performed only in limited experimental conditions. However, in near future, if wearable devices with physiological sensors like pedometer, cardiometer, or thermometer will be popular, we can expect to get numerous reliable dataset related to human emotion and their physiological responses. We can expect that this sensor data will be much more objective than traditional psychological experiment results, a survey based on the memory of subjects. If we can get data from the wearable devices, plotting them in a proper dimensional spaces (like Lövheim Cube of emotion) and clustering them can be a good research topic. As Christian Kaernbach [28] presented, current PAD dimensional spaces of emotion does not have discriminative performance of each emotional clusters. V-shape distribution in PAD space could be more uniform in Lövheim Cube of emotion or in a totally different dimensional space with measurable dimensional variables, for example, eye movement - pulse - body temperature, something like this. Or finding additional dimensions will be possible to make it have more discriminative power. To achieve this, we will have to prepare lots of objectively measurable dataset of emotions.

3. MACHINE LEARNING IN AFFECTIVE COMPUTING

Machine learning, a branch of artificial intelligence, field of study that gives computers the ability to learn without being explicitly programmed (by Arthur Samuel in 1959), is widely used nowadays in the most fields of computer science, like bioinformatics, natural language processing, speech recognition, computer vision, and even in other academic fields like psychology, medicine , economics , sociology. In affective computing, machine learning also plays a key role. It has been used to analyze data from various sensors, and draw the knowledge about human activity and emotion. Machine learning can be divided into three types, supervised learning, unsupervised learning, and reinforcement learning.

Supervised learning is a machine learning category using labeled training data. Training examples in the training data are pairs of an input vector and desired outputs. A machine learning algorithm for supervised learning tries to build a function to explain the given training data. After learning stage is finished, the model is used to predict a most-likely output from an input feature vector. Supervised learning is generally used for classification or regression problem. Classification is a task to generate a 'label' for input objects to divide them into several sets which have the same characteristics. Spam filtering is a typical example. There is two cases, 'spam' and 'non-spam'. The characteristics of spam mail and non-spam mail is learned from a given training dataset. When a new mail has arrived, it will be classified into spam or non-spam class using the model learned. Regression is a task to find a function which meets the tendency of observable training examples, then unobserved state or value can be predicted using that function. It is often used for time variant variables, like stock price, network traffic, and chemical concentration. Though both of classification and regression are supervised learning, they are different task. It depends on each algorithms which can solve classification or regression problem, or both of them. Decision tree, artificial neural network (ANN), Bayesian network (BN), naive bayes (NB), k-nearest neighbor (k-NN), support vector machine (SVM), random forest (RF) are representative algorithm of supervised learning. We will also talk about deep learning which comes up recently for big data analysis.

Unsupervised learning, on the other hand, is targeting to find hidden structure from unlabeled data. Because training data does not have any desired outputs, it tries to find similarity and dissimilarity of set of data. The output of unsupervised learning is a set of clusters of similar data points. Unsupervised learning is generally used for clustering problem. It is widely used, like pattern recognition, bioinformatics, computer vision, and so on. K-means clustering and Gaussian mixture model (GMM) are representative algorithm of unsupervised learning.

The other category, reinforcement learning, which is originally inspired by psychology, has three parts in its system - environment, agent, and action. An agent in given environmental condition can take an action to react the environment. Based on the action, the agent will get rewarded. Reinforcement learning is trying to find a sequence of actions to get the greatest accumulated reward. Though reinforcement learning is now widely used, for example, in game theory of economics, it is naturally designed to train behavior of an artificial agent to choose their action. It has been studied in affective computing area for a long time, especially building an AI agent for the appraisal theory research. However, we will not cover the AI-agent-based appraisal model here.

3.1 Classification algorithm

In affective computing research except AI agent, they are trying to predict emotional states by analyzing the data which is captured from facial expression, gesture, posture, motion, activity, location and physiological signals in laboratory or in the wild. This set of researches tend to use supervised learning algorithm as its tool.

Table 2: Machine learning in affective computing papers

Author	ML algorithm	Subjects	Amount data(duration)	Kinds data	Features	Output type
LiKamWa [39] A2	DBSCAN	25 iphone users smartphone usage	over one month	app, phone calls, email, SMS, web browsing, calendar, location	frequency count	histogram
LiKamWa [40]	Multi-linear regression	32 participants smartphone usage	over two month	app, phone calls, email, web browsing, calendar, location	normalized frequency	histogram
Maxhuni [44]	GMM	3 participants mobile phone usage	over 5 working days	voice	MFCC	p value
Epp [16]	Decision Tree	12 people keyboard usage	for 4 weeks	typing rhythm	key duration, key latency, key down-to-down, special characters	from 5 to 7 basic emotion classes
Lee [37]	Bayesian network	one person	314 text messages	text messages + sensor input	typing speed, function freq, special symbol, text length, and so on. + illuminance, location, time, weather	7 emotion classes
Lee [38]	Multi-class SVM, GMM, and k-means	focus group study (upto five interactants) smartphone usage	-	voice, bluetooth	MFCC, duration of speech, bluetooth pairing information	meta-linguistic information graph (turn-taking, dominance)
Attabi [7]	SVM	WCCN dataset (voice)	9,959 chunk to learn 8,257 for testing	voice	MFCC	5 emotion classes
Yang [64]	Multi-class SVM (linear, RBF-kernel)	LJ40k / MER31k dataset	28000 training, 6000 testing / 19427 training, 6000 testing	music-mood tags	music-mood tags	20-40 mood classes
Chang [21]	SVM	SUSAS dataset / Belfast Naturalistic database	298 audiovisual clips from 125 speakers	physical property of voice	450 features including MFCC, glottal timings	positive / negative stressed / neutral
Lee [36]	Linear discriminant classifier (LDC), and k-NN	call-center application with a machine agent	1187 calls (about 7200 utterances)	frequency, energy, duration, formants	mean, median, standard deviation, range, bandwidth	salience level (0-1 real value)
Hernandez [23]	Shore framework	deployed in campus	178 images of facial expressions	CCTV images	smile intensity	5 basic emotions
Schuller [58]	SVM, ANN, Decision Tree	EMO-CAR dataset	2829 dialogue	voice	Acoustic and linguistic features	7 emotion classes
Rachuri [52]	GMM	18 users mobile phone usage	for 10 days	voice	frame rate 10ms, windows size 30ms sampling	5 emotion classes
Kolodyazhniy [31]	LDA, QDA, ANN, k-NN	34 participants	response for two-sets of three 10-min film clips	physiological signals	physiological signals	fearful, sad, neutral
Petrantonakis [49]	QDA, k-NN, SVM, Mahalanobis Distance	-	EEG recording during image projection	Electroenchphalogram (EEG)	frequence domain	5 emotion classes
Wu [63]	ensemble of GMM, SVM, MLP	8 people	2033 utterances	acoustic prosodic and semantic lables of speech	pitch, intensity, formants, bandwidth, jitter, shimmer, hormonicity, MFCC	4 emotion classes
Kapur [29]	logistic regression, NB, Decision Tree, ANN, SVM	2 dancers	body motion capture using 14 markers, 6 cameras, 25 * 10 sec for each emotion	marker balls for body movement	velocity, acceleration, and their standard deviation	4 emotions
Bailenson [9]	linear regression, ANN, SVM	41 participants	19625 instances for amusement	face expression, cardio/somatic activity physiological signals	-	real value: amusement(<3), neutral(>0.5>
Nicolau [48]	ANN, SVM	sensitive artificial listener database	-	audiovisual data, 20 facial, neck, and shoulder feature points	MFCC and MFCC-delta,	prediction in PA space
Rigas [53]	RF, k-NN, PCA	9 subjects	118 instances	facial electromyograms (EMGs) electrocardiogram (ECG), respiration effort and electrofermal activity (EDA)	mean, median, standard deviation for each time frame	fear, disgust, happiness

Table 2 shows the list of classification algorithm of each papers. Unfortunately, lots of researches have just tried machine learning like a toolbox, in other words, just used several famous machine learning algorithm without any further investigation. However, Tarasov [61] did an interesting experiment. She benchmarked classification algorithms and compared performances of them using emotional database of natural speech. Following her research, SVM was the best classification algorithm for database, but multi-layer perceptron (MLP) and k-NN hold the first rank, as good as SVM. Artificial neural network-radial basis function (RBF), NB, decision tree (C4.5) had lower performance than them, holding the second rank. In this section, we will briefly & selectively survey machine learning algorithms in affective computing, especially, related to its compatible applications.

3.1.1 Artificial Neural Network

Artificial neural network (ANN) is the machine learning algorithm by mimicking the mechanism of neurons. ANN is one of the machine learning techniques which is most widely used. It is often used in affective computing. There is several kinds of ANN, representatively multi-layer perceptron (MLP), radial basis function (RBF). Deep learning, which is also a type of ANN - we will review this in the later subsection. Each specific type of ANN have different functions to generate an output for each node, and different learning algorithm. ANN consist of many nodes in several layers. A node in the input layer gets an element in an input vector. A node in the output layer produces an output. A node between the input layer and the output layer is called hidden node which passes signal from the input layer, to the output layer direction. A layer of hidden nodes is called hidden layer. There are edges from nodes in a layer to nodes in the next layer. ANN is used for various purpose like classification, clustering, regression of a function, forecasting, content addressed memory, and so on.

Kolodyazhniy [31] used ANN to classify human physiological data by emotion difference. He prepared 10min video clips to arouse fearful, and sad. He collected physiological signals, autonomic, respiratory, and facial muscle activity from audiences when they are watching the movie clips. He used physiological features. In this paper, he compared various feature selection and learning algorithm. He used two feature selection method, sequential forward selection and sequential backward selection, then it is learned with five machine learning models ANN (MLP, RBF), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), k-NN. As a result, k-NN showed the best accuracy. Shuller [58] used 4 kinds of machine learning algorithm to classify six basic human emotions (anger, disgust, fear, joy, sad, surprise), and neutral. He applied it to the corpus from FERMUS III project [57], which consists of 2,829 emotional utterance samples and scripts from 13 actors for one year. Using PCA and SVM based sequential forward floating selection (SVM-SFFS), he derived 20 statistical acoustic features from pitch, energy, and duration of the utterances, and nine linguistic features from selected 20 emotional words. SVM and ensemble classifiers are shown the best result, but MLP was also shown comparable accuracy. Kapur [29] used ANN to study gesture-based affect classification. She also used multiple machine learning algorithms, which are logistic regression, Naive Bayes, decision tree(C4.5), ANN, and support vector machine (SVM). She captured human gesture using 14 body-attachable ball markers and 6 cameras 25times * 10 seconds for each four emotion states (sadness, joy, anger, fear). From the motion capture data, she extracted four kinds (velocity, acceleration, and standard deviation of velocity and acceleration) of features from the 14 markers. ANN classifier showed over 90% accuracy almost equal to SVM. Bailenson [9] performed regression using ANN to predict emotion intensity (in continuous value). He gathered facial expression data from 22 track points, and cardio activity and somatic activity data from 41 participants, 19,625 instances. Nicolaou [48] used a kind of Recurrent Neural Network to predict Valence-Arousal (VA, which is the same to PA) changes over time. He used Sensitive Artificial Listener Database [13], which is audiovisual data (audio, face, shoulder). 20 facial feature points, neck+two shoulder points, and MFCC is used for features.

Many examples show that ANN is quite good. Normally, accuracy is not better than SVM, but comparable. However, ANN has advantages that it is flexible. It can solve both of regression and classification. ANN can deal with high-dimensional input vectors regardless of discrete or continuous. ANN can also have outputs in discrete or continuous. It is robust to noise. In addition, it is powerful to be capable of non-linear classification, though its algorithm is simple. And it also has an particular advantage to be able to multiple outputs in a model. There are two main disadvantages of ANN. The one is that there is no guidelines to decide the number of nodes and hidden layers. If it is too many, performance will degrade significantly. If it is too small, accuracy of prediction will be low. Another is that a classifier which is generated by ANN is not understandable to human.

3.1.2 Support vector machine

Support vector machine (SVM) is one of the most popular machine learning algorithm for classification nowadays. SVM is theoretically well motivated, and it has been empirically very successful in various fields like bioinformatics, text analysis, computer vision, and so on. SVM is proposed by Vlamir Vapnik in late 1970s. It could not attract an attention at that time. In 1990s, SVM is probed to have outstanding performance in writing recognition, and it is propagated to various fields. In typical machine learning algorithm, objective function tries to minimize error, on the other hand, SVM tries to maximize this 'margin' as well as to minimize error. Margin is distance from a hyperplane, which divides two class instances, to a closest instance. By maximizing margin, it can have better classification performance for unseen instances. Though SVM is basically linear classification method, it can also permit non-linear classification using special kernels.

SVM is the most popular machine learning algorithm nowadays. Total 9 papers over 25 used SVM solely or together with other algorithm [7, 9, 21, 29, 38, 48, 49, 63, 64]. Lee [38] tried to infer metalinguistic information using smartphones' microphone and bluetooth neighbor discovery function. He uses SVM to identify every durations, which interactants are talking, or all of them are silent, from speech data in real time. Once SVM is trained, it is very fast to classify, so that it seems that a good choice to use SVM for the system. However, SVM can classify just two classes at once. So, he used multi-class SVM, which has n SVM to divide n+1 classes, turn by turn. Attabi [7] and Chang [21] also used SVM for voice data. Attabi with a kernel to infer 5 emotional states from voice dataset, FAU AIBO Emotion Corpus [60]. Features are extracted using very common Mel-Frequency Cepstral Coefficients (MFCC). He tested SVM, SVM with RBF kernel, logistic regression, MLP and RF. Among them, SVM and MLP showed the best result. Chang applied SVM to SUSAS dataset [20], which has 298 audiovisual clips from 125 speakers. As features, in addition to MFCC, he used several physical properties like glottal vibrational cycle.

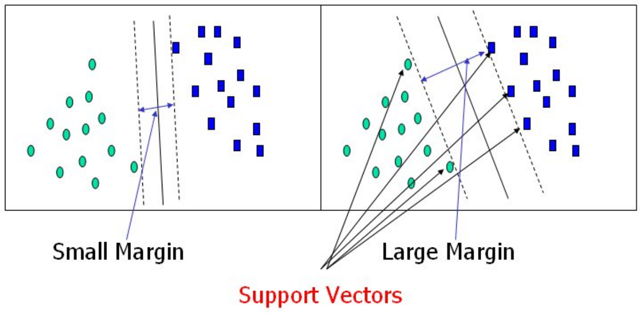

Figure 7: Support vector machine [2]: Instead of just finding a hyperplane to divide two classes, SVM focuses on maximizing 'margin' of the hyperplane.

SVM has many advantages. It can avoid over-fitting, because it has regularization parameter. SVM can permit non-linear classification when it uses a kernel, especially, when it uses a special types of kernels, like Gaussian kernel, training SVM can be reduced to convex optimization problem, so that it can be solved efficiently without worrying about local minima. And once training is finished, classification speed is very fast. However, SVM also has many disadvantages, it is hard to deal with discrete data. SVM is basically supports only two classes. Except several special kernels, optimization is very hard. Escpecially, because complexity of its algorithm is high and is not memory-efficient, SVM is not suitable for big data.

3.1.3 K-nearest neighbor

k-nearest neighbor(k-NN) is one of the simplest machine learning algorithm. K-NN algorithm is very simple. K-NN just saves all the training examples. When an instance comes, it calculates distance from the new instance to training examples. Then it finds k closest training examples to the new instance. The output is a majority voting or an average of the k closest training examples for classification or regression, respectively. K, the number of training examples which will participate to the majority voting, is the one and only parameter of the algorithm.

Rigas [53] constructed bio-signal based fast emotion recognition systems for three emotions (fear, disgust, happiness). He used random forest (RF) and k-NN to analyze 118 instances of facial electromyograms (EMGs), electrocardiogram (ECG), respiration effort and electrodermal activity (EDA) data from 9 subjects. K-NN showed better performance than RF. Lee [36] and Petetrantonakis [49] also used k-NN to study emotion. Lee compared k-NN and linear discriminant classifier (LDC) to classify valence of emotion from 1187 calling database. K-NN is significantly better than LDC in their experiment. Petrantonakis compared several machine learning algorithms like quadratic discriminant analysis (QDA), k-NN, Mehalanobis Distance (MD), SVM. He transformed Electroencephalogram (EEG) information into frequency domain using Fast Fourier Transform (FFT), and classified into 6 basic emotions (happiness, surprise, anger, fear, disgust, sadness). Kolodyazhniy also used k-NN and it showed the better performance than ANN, LDA, QDA.

K-NN is simple to implement and easy to understand, but quite powerful to be able to be able to deal with non-linear classification. It can be used for classification and regression. Typical disadvantages

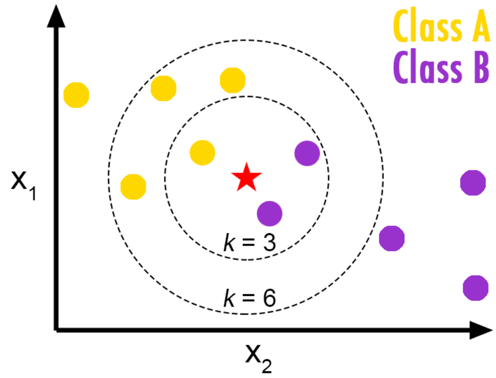

Figure 8: K-nearest neighbor [3]: When k=3, a new instance(star) belongs to class B, when k=6, then class A.

of k-NN are huge memory consumption to store all training examples and slow classification speed to find k nearest neighbors for all instances. In learning process, k-NN simply store all the training instances. The thing that learning stage didn't do nothing and all things are done in classification stage. This characteristic is called 'lazy' learning. Because of this lazy learning property, classification process is slow. And in many real problems, choosing appropriate similarity (or distance) metric can be tricky. As we see above, if you use PAD dimensional model, the emotional space might be distorted.

3.1.4 Decision tree

Decision tree is one of the oldest machine learning technique. However, decision tree is still effective technique in many problems. Representative algorithms are Classification And Regression Tree (CART), ID3 and C4.5. CART performs greedy and recursive partitioning. 'Greedy' means that it chooses a feature which gives most information gain (half:half division) at this stage. Decision tree is easy to understand the result, even good to identify a few critical features from the result. But decision tree is very slow to train, and has risk of overfitting and limitation of representation powers, which can't represent XOR. Though many affective computing research has still used decision tree [16, 29, 58], random forest (RF) could be better choice. RF, which is developed by improving decision tree, has overwhelming accuracy and performance than decision tree. We will cover RF in the following subsection.

3.1.5 Random forest

Random forest (RF) is very popular and widely used algorithm, which is comparable to SVM. RF is an improved variance of the decision tree, which comes from statistics. RF is consisted of many small decision trees. Output is decided by the majority voting of these trees. In learning stage, RF divides training sets into many small subsets, and then train a decision tree for each subset. Because RF is consisted of many decision trees, RF can be considered as an ensemble classifier. Generally, ensemble classifier usually has much higher accuracy and robustness than its element classifiers, and it needs much more memory and computing power. But training RF is much faster than decision tree in both cases of training and predicting, because RF divides an original problems into subproblems. It can be even faster, if it solved using parallel processing. RF has highest accuracy - comparable to SVM - among state of the art machine learning algorithm. RF can deal with huge dataset and thousands of features. RF is robust to outliers, so that it can

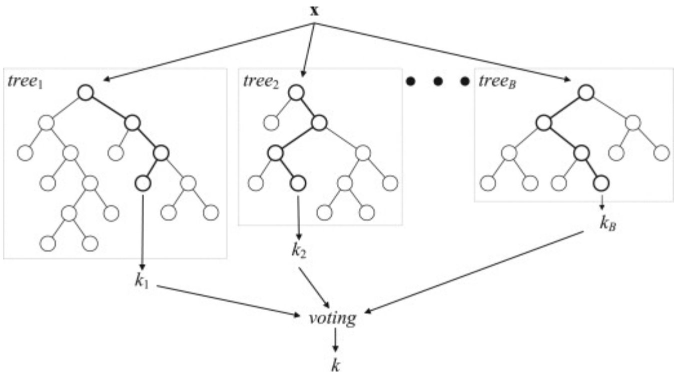

Figure 9: Random forest [4]: It is a set of decision tree of the subsets. The output is decided by majority voting of those trees.

deal with missing values well, or even can be used to detect unsupervised clustering problem or outlier detection. RF has unbiased property, so that it does not need cross validation.

Though RF is fast, accurate, and robust, it was rarely used in affective computing research. Rigas [53] used RF, but it did not show impressive performance. We will be able to study why RF is not often used for affective computing. RF will be able to help us to improve performance of many affective computing research. Otherwise, it is possible that RF is not suitable for emotional dataset. RF is said to have tendency to overfit for some datasets with noisy classification or regression tasks.

3.1.6 Naive Bayes

Naive Bayes (NB) classifier is a classifier based on Bayes theorem. Naive Bayes is also widely used, because it is very simple and fast, and has pretty good accuracy. It is very effective in storage, and versatile to deal with real or discrete either. Especially, it is used in case that very fast processing speed is needed, like computer vision data processing. Applications of NB are text classification, spam filtering, system performance management, and medical diagnosis. In affective computing, Kapur [29] used NB as well as other machine learning algorithms like logistic regression, decision tree(C4.5), ANN, and SVM. Hoque [25] also compared Naive Bayes to other various classifiers like Random Forest(RF), Bagging, MLP, RBF to classify emotions into positive and negative using acoustic and prosodic feature like pitch, duration, intensity, formant, rhythm. NB was one of the simplest algorithm, but it showed performance close to average. Schuller [58] also used NB with other classifiers like ANN, decision tree, SVM, and so on, but NB showed the lowest accuracy among them. Though NB has clear advantages that it is very fast, the limit of its performance seems also quite clear.

3.1.7 Deep learning

Deep learning is a kind of ANN with deep architecture. Deep architecture means an algorithm has many hidden layers. Traditional neural network with many hidden layers had limitations like that [47],

1. Lacks the ability to train the unlabeled data while in practice most data is unlabeled.
2. The correcting signal will be weakened when it passes back via multiple layers.
3. Learning is too slow across multiple hidden layers.

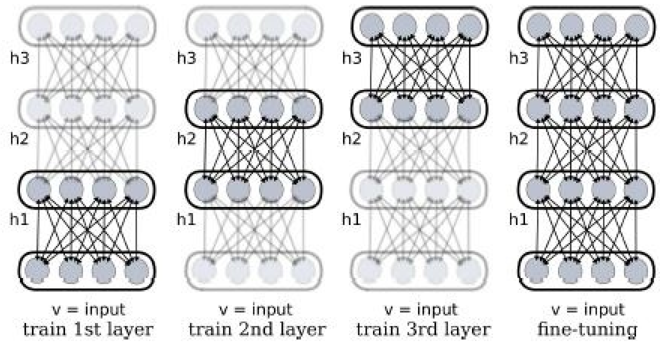

Figure 10: The deep learning scheme [6]: a greedy unsupervised layerwise pretraining stage followed by a supervised finetuning stage affecting all layers.

4. It can get stuck in poor local optima.

So, hidden layers could not exceed one or two. During neural networks couldn't improve its performance, shallow architecture learning algorithm (SLP, SVM) has been popular. Up to now, SVM is the popular machine learning algorithm. But SVM has naturally lack of learning features due do its shallow architecture itself. Bengio and Lecun [10] have proved that shallow architectures can be very inefficient in terms of required number of computational elements and examples, and kernel machines have limitations with a local kernel. To solve this, they proposed deep learning. Deep architectures help deep learning by trading a more complicated space for better performance, in some cases, even for less computation time [47]. With experimental results, deep learning algorithm like deep belief network was more accurate and faster than SVM. Like ANN, deep learning has several specific models, the representative algorithm is deep belief network (DBN) [24]. As a node of DBN, restricted Boltzmann machines(RBM) [6] are used. RBM is a kind of energy function. DBN has two training steps, unsupervised layer-wise pre-training step and supervised fine-tuning. In pre-training steps, from bottom to top layer, each layers are trained. The pre-training is a greedy layer-by-layer learning procedure to get optimum weight vector in RBM. In fine-tuning steps, backpropagation is used.

Though deep learning is rising up, especially related to big data, we have to be careful to apply it. Though deep learning works well - it outperforms the SVM with Gaussian kernel for many problems, it does not mean that it work well for all kinds of problems. Deep learning does not work for some specific problems like natural language processing(NLP). It will be needed to validate deep learning for affective computing.

3.2 Clustering algorithm

Representative algorithm of unsupervised learning are K-means and Gaussian mixture model. K-means is an algorithm to find mean value of each clusters in dataset. It can be calculated using EM-algorithm. To resolve the issue that k-Means only has the mean values of each clusters, Gaussian mixture model is developed, which can get standard deviation values as well as mean values.

In affective computing, few authors tried to apply unsupervised learning to their research. Lee [38] uses GMM and k-means to identify speakers. Wu [63] also used GMM He clusters users' time sequence into activity segments, which has consistent context. Later, users can review their activity segments and label them. Using these users' labeling, the system can label activity segments automatically in the future. This can make users easy to remind their past activities, and the system improve its performance gradually. This is very interesting approach. For now, labeling of emotional data is very hard except for surveying. Only facial expression may be close to the feedback from users. But getting feedbacks from a user is expensive operations itself, as well as it is hard to believe its accuracy. In this context, clustering signals (from smartphone sensors to pictures or even physiological signals) and finding groups will be foundation to get an exact emotional states later.

4. CHALLENGES

We have reviewed emotion modeling methods and machine learning techniques in affective computing, and several existing researches in section 2 and section 3. Though, as an interdisciplinary study, affective computing will have various points to improve, three issues have to be considered from the perspective of computer science.

- Emotion modeling is a key to understand human emotion. However, existing emotion models mostly come from Psychology. Though Psychology has studied human emotion for a long time and draw enormous knowledge about it, it is originally designed for surveying. It has to be existing improve emotion models or developing new emotion models to get proper models for scientific measurement. In addition, as we can see in Kaernbach's research, PAD space model did not meet metamerism. It means that existing emotional models may not be enough to distinguish emotional states well. We have to consider how can we validate each emotion models using data, or how can we translate an existing emotional space into another space, based on above experimental result. For an example, we will be able to measure monoamine neurotransmitters with various stimulus. ploting this data into new emotional space, for example Serotonin - Norepinephrine - Dopamine, will show us new kinds of clusters to give us new insight. For another example, we can get data from the wearable devices, plotting them into a totally new dimensional spaces, like mobility-location-tone of voice-temperature-pulse, to find clusters them to draw new knowledge. This data-driven validation can be a major contribution of computer science to affective computing research.

- In case of machine learning method, it will be needed to find a suitable method for each area, and also be needed to try recent machine learning techniques for it. For example, random forest is not widely used in affective computing research yet. Rigas [53] used RF, but it did not show impressive result. Normally, random forest shows outstanding accuracy, which is comparable to support vector machine. We may apply random forest to other databases to see it is working. On the other hand, recent developed machine learning algorithm has to be applied. Deep learning is getting popular nowadays. We may apply deep learning to affective computing. Because deep learning is fast to deal with huge database, and it can use unlabeled datasets and labeled datasets together, it will be suitable for emotion dataset.

- Finally, the interdisciplinary research has to be promoted. Affective computing is, like cognitive science and social computing, an interdisplinary research area. Affective computing is done by many computer scientist, psychologists, and physiologists. However, it seems that it is rare that those scholars worked together. Conferences and Journals are also separated except "Psychophysiology". If we have interdisciplinary conference for affective computing, it will be able to promote many scholars in different fields work together and deal with affective computing from various point of view in a conference. Physiologists and electrical engineers will be able to design sensors for wearable devices, and computer sciencetists and psychologists will be able to analyze the data to draw new knowledge and validate emotional models for better understanding of human emotion.

5. CONCLUSION

We have reviewed representative emotion modeling and machine learning method of affective computing. Various emotion models from psychology is used for affective computing. But many of them are designed for psychological research, so that we may need to validate that it is suitable for scientific experiments, and even it is needed to find a new model for it. Collecting data, and converting existing data into these scientific emotional space will be the most urgent and significant issue. In case of machine learning, we found that many researchers uses popular classification algorithm like ANN, k-NN, SVM, Decision tree, NB, RF, and so on. Applying brand new classification algorithm, like deep learning, or even applying clustering algorithm will give us new insight. Nowadays, smartphones are used more than 1.75 billion users worldwide. Even wearable devices are actively developing. Affective computing, one of data-driven research about human, is getting more and more important. In computer science, it can be applied to build better AI, which understands owner's or customer's emotion and mood. Affective gaming is a good example. On the other hand, it will take an effect to humanities and sociology. For example, affective computing and big data will be able to use webpages, smartphones or wearable devices everywhere to measure happiness of the society. This will provide a new change to human being to focus on 'happier' life than 'richer' life in the future.

6. REFERENCES

[1] Cognitive appraisal theory, http://psychlopedia.wikispaces.com/cognitive+appraisal+theory.

[2] Introduction to support vector machine (svm) models, http://www.dtreg.com/svm.htm.

[3] k-nearest neighbor, http://datasciencerules.blogspot.kr.

[4] The random enchanted forest, http://blogs.sas.com/content/analitika/2012/04/05/az-elvarazsolt-veletlen-erdo/.

[5] The three biogenic amine systems, http://countyourculture.com/2010/08/30/the-three-biogenic-amine-systems/.

[6] L. Arnold, S. Rebecchi, S. Chevallier, and H. Paugam-Moisy. An introduction to deep-learning. In *Advances in Computational Intelligence and Machine Learning, ESANN'2011*, pages 477–488, Apr. 2011.

[7] Y. Attabi and P. Dumouchel. Anchor models for emotion recognition from speech. *T. Affective Computing*, 4(3):280–290, 2013.

[8] R. D. Badgaiyan. *Neurotransmitter Imaging*. Springer New York, 2006.

[9] J. N. Bailenson, E. D. Pontikakis, I. B. Mauss, J. J. Gross, M. E. Jabon, C. A. C. Hutcherson, C. Nass, and O. John. Real-time classification of evoked emotions using facial

feature tracking and physiological responses. *Int. J. Hum.-Comput. Stud.*, 66(5):303–317, 2008.

[10] Y. Bengio and Y. LeCun. Scaling learning algorithms towards ai. In L. Bottou, O. Chapelle, D. DeCoste, and J. Weston, editors, *Large-Scale Kernel Machines*. MIT Press, 2007.

[11] S. Blakeslee. Dr. joseph ledoux: Using rats to trace routes of fear, Nov. 1996.

[12] S. D'Mello and R. A. Calvo. Beyond the basic emotions: What should affective computing compute? In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 2287–2294, New York, NY, USA, 2013. ACM.

[13] E. Douglas-Cowie, R. Cowie, I. Sneddon, C. Cox, O. Lowry, M. Mcrorie, J. claude Martin, L. Devillers, S. Abrilian, A. Batliner, N. Amir, and K. Karpouzis. The humaine database: Addressing the collection and annotation of naturalistic and. In *Induced Emotional Data, Affective Computing and Intelligent Interaction*, pages 488–500, 2007.

[14] P. Ekman. *Basic Emotions*. Wiley, New York, 1999.

[15] P. Ekman and W. V. Friesen. Constants across cultures in the face and emotion., 1971.

[16] C. Epp, M. Lippold, and R. Mandryk. Identifying emotional states using keystroke dynamics. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems (CHI 2011)*, pages 715–724, Vancouver, BC, Canada, 2011.

[17] N. Ferreira, S. Mascarenhas, A. Paiva, G. di Tosto, F. Dignum, J. McBreen, N. Degens, G. J. Hofstede, G. Andrighetto, and R. Conte. An agent model for the appraisal of normative events based in in-group and out-group relations. In *AAAI*. AAAI Press, 2013.

[18] J. Gratch and S. Marsella. Tears and fears: modeling emotions and emotional behaviors in synthetic agents. In *Agents*, pages 278–285, 2001.

[19] H. Gunes, B. Schuller, M. Pantic, and R. Cowie. Emotion representation, analysis and synthesis in continuous space: A survey. In *FG*, pages 827–834. IEEE, 2011.

[20] J. H. L. Hansen. Susas ldc99s78. web download. philadelphia: Linguistic data consortium, 1999.

[21] K. hao Chang, D. Fisher, J. Canny, and B. Hartmann. How's my mood and stress?: An efficient speech analysis library for unobtrusive monitoring on mobile phones. In *Proceedings of the 6th International Conference on Body Area Networks*, BodyNets '11, pages 71–77, ICST, Brussels, Belgium, Belgium, 2011. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[22] M. M. Haque, M. Adibuzzaman, D. Polyak, N. Jain, S. I. Ahamed, and L. Liu. Irene: Context aware mood sharing for social network. In K.-J. Lin, C. Huemer, M. B. Blake, and B. Benatallah, editors, *SOCA*, pages 1–8. IEEE, 2011.

[23] J. Hernandez, M. E. Hoque, W. Drevo, and R. W. Picard. Mood meter: Counting smiles in the wild. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 301–310, New York, NY, USA, 2012. ACM.

[24] G. E. Hinton, S. Osindero, and Y.-W. Teh. A fast learning algorithm for deep belief nets. *Neural Comput.*, 18(7):1527–1554, July 2006.

[25] M. E. Hoque, M. Yeasin, and M. M. Louwerse. Robust recognition of emotion from speech. In J. Gratch, M. Young, R. Aylett, D. Ballin, and P. Olivier, editors, *IVA*, volume 4133 of *Lecture Notes in Computer Science*, pages 42–53. Springer, 2006.

[26] C. E. Izard. Four systems for emotion activation: Cognitive and noncognitive process. *Psychological Review100*, 1:68–90, 1993.

[27] P. N. Johnson-Laird and K. Oatley. The language of emotions: An analysis of a semantic field. *Cognition and Emotion*, 3:81–123, 1989.

[28] C. Kaernbach. On dimensions in emotion psychology. In *2011 IEEE International Conference on Automatic Face and Gesture & Recognition and Workshops (FG 2011)*, pages 792–796. IEEE, Mar. 2011.

[29] A. Kapur, A. Kapur, N. Virji-Babul, G. Tzanetakis, and P. F. Driessen. Gesture-based affective computing on motion capture data. In *Proceedings of the First International Conference on Affective Computing and Intelligent Interaction*, ACII'05, pages 1–7, Berlin, Heidelberg, 2005. Springer-Verlag.

[30] G. King. Ensuring the data rich future of the social sciences. *Science*, 331:719–721, 2011 2011.

[31] V. Kolodyazhniy. An affective computing approach to physiological emotion specificity: toward subject-independent and stimulus-independent classification of film-induced emotions. *Psychophysiology*, 48(7), July 2011.

[32] P. J. Lang and M. M. Bradley. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30:261–273, 1993.

[33] R. S. Lazarus. *Emotion and Adaptation*. Oxford University Press, 1991.

[34] J. E. LeDoux. Rethinking the emotional brain. *Neuron*, 73:653–676, Feb 2012.

[35] J. E. LeDoux. Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology*, 24:187–192, Jan 2014.

[36] C. M. Lee and S. S. Narayanan. Toward detecting emotions in spoken dialogs. *IEEE Transactions on Speech and Audio Processing*, 13(2):293–303, 2005.

[37] H. Lee, Y. S. Choi, S. Lee, and I. P. Park. Towards unobtrusive emotion recognition for affective social communication. In *CCNC*, 2012.

[38] Y. Lee, C. Min, C. Hwang, J. Lee, I. Hwang, Y. Ju, C. Yoo, M. Moon, U. Lee, and J. Song. Sociophone: everyday face-to-face interaction monitoring platform using multi-phone sensor fusion. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, MobiSys '13, pages 375–388, New York, NY, USA, 2013. ACM.

[39] R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong. Can your smartphone infer your mood? In *PhoneSense workshop*, 2011.

[40] R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong. Moodscope: building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, MobiSys '13, pages 389–402, New York, NY, USA, 2013. ACM.

[41] H. Lövheim. A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical Hypotheses*, 78(2):341–348, 2012.

[42] R. L. Mandryk and M. S. Atkins. A fuzzy physiological approach for continuously modeling emotion during

interaction with play technologies. *Int. J. Hum.-Comput. Stud.*, 65(4):329–347, Apr. 2007.

[43] S. C. Marsella, J. Gratch, and P. Petta. Computational models of emotion. In *A blueprint for an affectively competent agent: Cross-fertilization between Emotion Psychology, Affective Neuroscience, and Affective Computing*. Oxford University Press, Oxford, 2010.

[44] A. Maxhuni, A. Matic, V. Osmani, and O. Mayora-Ibarra. Correlation between self-reported mood states and objectively measured social interactions at work: A pilot study. In *PervasiveHealth*, pages 308–311. IEEE, 2011.

[45] A. Mehrabian. *Basic dimensions for a general psychological theory: implications for personality, social, environmental, and developmental studies*. Social Environmental and Developmental Studies. Oelgeschlager, Gunn & Hain, 1980.

[46] A. Mehrabian. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14(1):261–292, Dec. 1996.

[47] D. Mo. A survey on deep learning: one small step toward ai, 1971.

[48] M. A. Nicolaou, H. Gunes, and M. Pantic. Continuous prediction of spontaneous affect from multiple cues and modalities in valence-arousal space. *T. Affective Computing*, 2(2):92–105, 2011.

[49] P. Petrantonakis and L. J. Hadjileontiadis. Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis. *T. Affective Computing*, 1(2):81–97, 2010.

[50] R. W. Picard. *Affective computing*. MIT Press, Cambridge, MA, USA, 1997.

[51] R. W. Picard. Affective computing: From laughter to ieee. *IEEE Trans. Affect. Comput.*, 1(1):11–17, Jan. 2010. Fellow-Picard, Rosalind W.

[52] K. K. Rachuri, M. Musolesi, C. Mascolo, P. J. Rentfrow, C. Longworth, and A. Aucinas. Emotionsense: A mobile phones based adaptive platform for experimental social psychology research. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, Ubicomp '10, pages 281–290, New York, NY, USA, 2010. ACM.

[53] G. Rigas, C. D. Katsis, G. Ganiatsas, and D. I. Fotiadis. A user independent, biosignal based, emotion recognition method. In *User Modeling*, volume 4511 of *Lecture Notes in Computer Science*, pages 314–318. Springer, 2007.

[54] L.-F. Rodríguez, F. Ramos, and Y. Wang. Cognitive computational models of emotions. In *IEEE ICCI*CC*, pages 75–84. IEEE, 2011.

[55] J. A. Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161–1178, 1980.

[56] K. R. Scherer. *The Neuropsychology of Emotion (pp. 137-162)*. New York: Oxford University Press, 198 Madison Avenue, New York, NY, 10016, USA, 2000.

[57] B. Schuller. Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture. volume 1, pages 577–580. IEEE, May 2004.

[58] B. Schuller, R. J. Villar, G. Rigoll, M. Lang, and T. U. MÃijnchen. Metaclassifiers in acoustic and linguistic feature fusion-based affect recognition. In *in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '05*, pages 325–328, 2005.

[59] C. A. Smith and R. S. Lazarus. *Emotion and adaptation*. Guildford, 1990.

[60] S. Steidl. *Automatic classification of emotion related user states in spontaneous children's speech.* PhD thesis, University of Erlangen-Nuremberg, 2009. http://d-nb.info/992551641.

[61] A. Tarasov and S. J. Delany. Benchmarking classification models for emotion recognition in natural speech: A multi-corporal study. In *FG*, pages 841–846. IEEE, 2011.

[62] D. Watson, L. A. Clark, and A. Tellegen. Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of Personality and Social Psychology*, 54:1063–1070, 1988.

[63] P. Wu, J. Zhu, and J. Y. Zhang. Mobisens: A versatile mobile sensing platform for real-world applications. *Mob. Netw. Appl.*, 18(1):60–80, Feb. 2013.

[64] Y.-H. Yang and J.-Y. Liu. Quantitative study of music listening behavior in a social and affective context. *IEEE Transactions on Multimedia*, 15(6):1304–1315, 2013.