1. **Introduction**

**1.1 Research Background**

**1.1.1 Service design**

Experience economy is getting more important because more people are willing to spend more time and money on gaining special experiences in addition to goods consumption ([Bergen 1988](#_ENREF_3), [Pine and Gilmore 1998](#_ENREF_42)). In recent years, there have been many methods for designing customer experiences developed in service design communities. The emerging issues faced by service designers are how to connect customer value, which could be identified via service design process with business value that could be specified by business executives. The investment done by a service organization may be in vain if it cannot fulfill customers’ needs via the service delivery process which in general consists of numerous assets, processes, people, and materials ([Goldstein, Johnston et al. 2002](#_ENREF_16))

Experience perceived by customers in the constituent touch points is the key determents of customer satisfaction, and in turn, influencing customers’ loyalty to the service provider. A systematic approach is needed for a service provider to upgrade existing service or create new services in order to effectively create value for customers and also maintain its financial health in service innovation. For a service provider to design customer experience, it is essential to analyze every touch point to know what customers really experienced ([Spengler, Wirth et al. 2010](#_ENREF_46)) .

There are various service design methods nowadays. Different designers manage the design process by choosing various methods and tools for different stages. Among different methods, but the double diamond model capture the most commonly stages of a design process. They are discovery, definition, development, and delivery. In discovery stage, many tools, such as cultural probe, shadowing, customer experience journey, etc. have been used in specific situations. The precision of capturing customers’ needs in the discovery stage is essential for the success of follow-up service development. Therefore, the importance on enhancing the effectiveness of discovering customers’ needs cannot be over emphasized.

**1.1.2 EEG-based BCI wearable device**

The retrieval of signal directly from the brain has been researched for decades. In past years, a Brain Computer Interface (BCI) has been widely studied. A BCI is a non-muscular communication system that a person can use to directly communicate his/her intent from the brain to the environment ([Wolpaw, Birbaumer et al. 2002](#_ENREF_52)). The most commonly recognized application of BCI is for individuals who lose abilities to control their muscles to be able to operate machines in their daily lives without physically entering commands ([Vidal 1973](#_ENREF_50)).

There are many types of BCI device, e.g., invasive, partially invasive and non- invasive. What we choose for this study is a non-invasive method, which is safer and quicker to install on the head of a subject for developing a service which can be used in normal life. There are three kinds of non-invasive technology, EEG (Electroencephalography), fMRI (Functional magnetic resonance imaging) and EMG (Electromyography). Considering the cost and maturity, we choose EEG to be our research technology. Usually, experimental EEG is multi-electrode, medical grade, high price and needed be set on head with gel. But there are several new companies, such as Emotiv and Neurosky, making low-cost EEG devices to allow people to wear EEG while they can freely move. We chose NeuroSky Mindset™ (2009) as the devise in this study because there are several papers based on this device. It is proven to be a suitable experimental device for mobile EEG application ([Crowley, Sliney et al. 2010](#_ENREF_12), [Kwang-Ok, Jong-Bae et al. 2010](#_ENREF_29), [Larsen 2011](#_ENREF_30), [Mostow, Chang et al. 2011](#_ENREF_34), [Mak, Chan et al. 2013](#_ENREF_32), [Yoon 2013](#_ENREF_54)). NeuroSky has developed a non-invasive, dry, biosensor to read electrical activities in the brain to determine the states of attention and relaxation. The data is transmission via Bluetooth, which is been verified by NCC (National Communication Commission). Another reason is because single dry EEG sensor is more available for subject to wear to move freely, which can detect the brain activities while the subjects are moving. This device can filter muscle noise and provide data already decomposed frequency into 8 bands (high alpha, low alpha, high beta, low beta, high gamma, low gamma, delta, and theta) and each band stands for different meanings ([NEUROSKY 2009](#_ENREF_37)).

**1.2 Research Motivation**

In the double diamond model, discovery is the first stage that can affect the outcomes of the follow-up stages. There are many ways to research customer experience, such as cultural artifacts, field studying, shadowing, customer journey, interview, and so on. In the context that a researcher cannot participate in the scene with the studied subjects, an automatic recording method could be useful, such as audio/video recording. However, to capture a subject’s mental state beyond his/her facial/vocal expression needs the facilitation of BCI type of devise. Specifically, the following three situations may need the usage of BCI facilitation. First, when a subject needs to focus on something or very busy, *e.g*., driving or taking roller coaster, s/he can’t record the mental state at the same time. Second, when a subject has to spend times on recording their states, it may lead them to record what they think it’s important, and ignore other points which may important, too. Third, several discovery methods could be used to record a subject’s mental state via facial recognition, but this method needs specific knowledge and techniques for facial recognition, and is costly if the customer journey is very long.

In this research, we focused on the issue of e-learning because now technology allow a large group of learners, sometimes more than one hundred thousand learners, to register on one single class, how to ensure the high quality of learning outcomes is also concerned. Therefore, mental state is very important in this situation. The research is positioned to adopt BCI techniques to enhance service discovery stage of service design process by recording a subject’s mental states on his/her touch points of the customer journey. In order to effectively apply BCI for a customer’s mental detection, we have to validate the classification of a customer’s mental state. A machine learning approach for training a classifier will be developed in this study in order to supply a viable classifier for mental state detection.

**1.3 Research Objectives**

In this research, we build a service that can cost effectively detect a user’s mental state while emerging in the e-service context without interrupting the user’s interaction with the service delivery. In achieving this objective, we developed an automatic classification system, which is training by experiments set up in various e-service contexts to obtain interaction episodes with users’ reaction. A set of machine learning techniques in classification would be candidates to develop the classifier for mental state detection. After that, we build a website as the service interface for subjects experience English listening test materials and tag these materials. The output produced by subjects would be our input of classifier. After this, we evaluated this by recall and precision.

1. **Literature review**

**2.1 Service design process**

Design has already played an essential role in many success leading companies. Design council studied eleven world top design teams and developed the double diamond model (also called 4D model), which divides service design process into four distinct phases: *discover*, *define*, *develop,* and *deliver* (Design Council, 2005).

We describe each step as follows. The first quarter of the double diamond model marks the start of the design process. This begins with an initial idea or inspiration, often sourced from a discovery phase in which user needs are identified. Before inspiring ideas, we have to gather enough background knowledge and regularly updated information about a product or service. After gathering these types of information, a design team could use for defining the problem faced by the customers by analyzing obtained information. The second quarter of the double diamond model denotes the definition stage, in which interpretation and alignment of these needs to business objectives is achieved. A design team could identify the problem which will drive the design team to specify the potential value created from the generated product or service. The end of the definition phase is the time to make go or no-go decision with detailed understanding of the potential market for the new design, together with a good idea of the cost and complexity of realizing the idea. The third quarter marks a period of development where design-led solutions are developed, iterated and tested within the company. To solve a complicated problem, it is essential for the design team to equip the multidisciplinary working capability develop solutions from many perspectives. When developing new services, it is important to test concepts with prototypes, and to seek feedback from customers. The final quarter of the double diamond model is the delivery stage, where the resulting product or service is finalized and launched in the target market. In this step, a design team could identify constraints of making the product or delivering the service. Then, the new service or product can be launched to test the market and improve the service and product based on the feedback from customers.

Another popular service development process is design thinking, which is a solution-based methodology for creative resolution of problems. A design thinking process consists of five stages, *empathize, define, ideate, prototype*, and *test,* aiming to generate a prototype of product or service (Stanford d.school’s Design Thinking Resources Center). First, as a service designer, we empathize with users to recognize their physical and emotional needs by understanding the way they do things and why, what they think about the world, and what is meaningful to them. It could be viewed as a process of sensing and leading to the process of sense making mainly in define mode. Second, we all agree that a right solution comes from the framing of the right problem; thus, the goal of the definition stage is to draft a meaningful and actionable problem statement, which is also called a point-of-view. In the definition stage, the endeavor is to synthesize scattered findings into insights. Third, the ideation stage is aimed to generate ideas. Ideation supplies the fuel and source materials to build the prototypes and then create potential solutions to tackle the problems and satisfy users’ needs. In the ideation stage, it is very important to separate the generation of ideas from the evaluation of generated ideas. Fourth, quickly create a draft or prototype to present generated ideas. The prototype mode is an iterative generation of solutions including artifacts and processes intended to fulfill the users’ needs. Fifth, the test step is to solicit feedback from experiencing the prototypes built, which can gain the opportunity to empathize target user groups for whom we have been designing. These five stages are iterative until the tested prototyping products or processes can reach user groups’ satisfaction, and move toward the production stage of goods or services.

The first step of both methodologies mentioned before are essential for the follow-up steps. Since all steps are connected, the more accurate discovery of customers’ needs, the more possible that the generated products or services meeting their needs. In the initial stage of service design, among many methods existing for real world practice, the following approaches are commonly used for identifying customer behavior ([Kumar 2012](#_ENREF_28)):

* Cultural probe: Sometimes, it is infeasible to stay around subjects to get the information on customers’ contextual interaction. To conquer this difficulty, the cultural probe has been adopted for subjects to record their contextual information with cultural probe tool kit in the scenes they emerge. A cultural probe tool kit usually contains a camera, journal, video recorder for subjects to record things s/he sensed or thought and feeling they experienced. The tool kit also includes some tasks that they are asked to accomplish, sometimes.
* Field visit: Service designers could obtain the first hand information on customers’ behavior by spending time with these subjects in real world contexts. A field visit, unlike survey or interview, emphasizes on the observation and inquiry about what is being observed.
* Shadowing: This method requires researchers to immerse themselves in the lives of subjects in order to observe their behavior in interacting with the contexts they encountered.. Researcher can record subject journey and behavior with video, text or photographs. This method is also a useful technique for identifying those people may say one thing, and yet do another ([Stickdorn and Schneider 2010](#_ENREF_47)).
* Customer journey maps: This method provides a structured visualization of a service user’s experience upon their experience. All touch points are going to be identified first, and then we can record what customers do when they access the service. With this method, a researcher can easily identify the problems occurred in the customer journey and compare the service perceived by customers between with what obtained from competitors.

The aforementioned methods may be suitable for different situations. However, none of them can precisely record the mental states of subjects while interacting with contexts during the customer journey, which makes the efforts of acquiring real time mental states nontrivial.

**2.2 Non-invasive Brain Computer Interface**

There are three major and popular ways for recording human brain activities in both experimental and medical purposes using non-invasive brain computer interface. These are Electroencephalography (EEG), Magnetoencephalography (MEG) and Functional Magnetic Resonance Imaging (fMRI). All of these techniques have been researched for decades and many papers from Journal of Neuroscience or NeuroImage are based on these methods.

**2.2.1 Electroencephalography (EEG)**

Electroencephalography (EEG) is a way to record neurons electrical activity by setting electrodes on a person’s scalp to measure the voltage fluctuation over a period of time ([Niedermeyer and Silva 2004](#_ENREF_38)). When the brain wave, which is made by ions, reaches the electrodes, they can be captured by electrons on the metal on the electrodes. Because metal conducts the push and pull of electrons easily, the difference in such change of voltage can be measured by a voltmeter and then recorded over a period time, which muscle noise have been filtered, gives us the EEG ([Tatum, Husain et al. 2008](#_ENREF_49)).

The EEG could show brain activity in terms of rhythmic activities and transients. The rhythmic activity usually is decomposed into several specific bands by frequency. To some degree, these frequency bands have nomenclature. Frequency bands are usually extracted with eight bands via FFT (Fast Fourier Transform) as below:

1. Delta waves were defined as the frequency range up to 4 Hz. It is prominent frontal in adults and babies in slow wave when they are deeply sleeping. The wave is associated with seizure-like activity in the brain. [Harmony, Fernández et al. (1996)](#_ENREF_18) did a research to find the relation between delta band and mental loading. In the research, there are two experiments had been done. One was about calculation, delta band had increased obviously when subjects did calculation rather than read meaningless formula. Second one is about memory, delta band also been observed more easily when subjects try to remember five numbers rather than three numbers. Besides, this work also proved that all of delta, theta, alpha and beta bands are highly related to mental states.
2. Theta waves were defined as the frequency range from 4 Hz to 7 Hz. Theta is bound normally in young children or may be found drowsiness or arousal in older children and adults; it can also be seen in meditation ([Cahn and John 2006](#_ENREF_8)). Theta oscillations are very related to motor or cognitive task, no matter rodents or human. In the work of [Kahana, Sekuler et al. (1999)](#_ENREF_24), they test human with two type of maze; one had six junctions, the other one had nine junctions which is harder than first one. The experiments test the relation between short-term memory and theta oscillation, and they found that the harder maze is, the more episode of theta be observed. Also, the oscillation of theta band shows up when subjects in in problem solving, perceptual processing, learning, and memory ([Schacter 1977](#_ENREF_44), [Ertl 2013](#_ENREF_14)).
3. Alpha waves were defined as the frequency range from 7 Hz to 12 Hz. Hans Berger (1924) named rhythmic EEG activity he saw as the "alpha wave". Alpha waves are the "posterior basic rhythm". It predominantly emerges when subjects close eyes and with relaxation, and reduced when subjects eye opening or mental exertion ([Klimesch 1999](#_ENREF_25), [Vladimir, Ruth et al. 2001](#_ENREF_51)). Alpha wave in frontal lobe have strong relation with human relaxation and Beta wave has string relation with alert states ([Niemic 2002](#_ENREF_39), [Bos 2006](#_ENREF_6), [Bos 2007](#_ENREF_5)). High alpha means that brain has low activity, and low beta means brain has low alert state.
4. Beta waves were defined as the frequency range from 13 Hz to about 30 Hz. It is found usually on both sides in symmetrical distribution and is predominantly evident frontally, which is closely related to motor behavior and is generally minimized during active movements ([Pfurtscheller and Lopes Da Silva 1999](#_ENREF_41)). This wave is also related to active, anxious thinking, or active concentration.
5. Gamma waves were defined as the frequency range approximately 30–100 Hz. The rhythms are related to represent gathering different populations of neurons into a network for performing a certain cognitive or motor function, linguistic processing even associative learning based on the experiment, association between color and electric shock ([Miltner, Braun et al. 1999](#_ENREF_33)). Also, when these neurons clustered together, they will assist to bring up memories and associations from the visual perception to other notions.
6. Mu waves were defined as the frequency ranges 8–13 Hz. Sometimes this wave partly overlaps with other frequencies. The mu wave usually is found over the motor cortex which can control voluntary movement. The wave predominantly emerges not only when a subject is in active movement but also shows up when the subject observes the movement by others.

From the prior research, the normal electroencephalography varies by ages and individuals. An individual shows different shapes of electroencephalography even in the same context. Thus, the same band between different individuals cannot be compared directly. Besides, EEG-based mental recognition has been researched for many years. Each part of brain may present specific rhythm at different frequency ranges ([Bradley and Lang 1994](#_ENREF_7), [Arslan, Brouse et al. 2006](#_ENREF_2)). For different hemispheric, left and right part have different majors in emotion. The left one is more major in processing positive emotions and active behavior, whereas the right hemisphere is more major in processing negative emotions and withdrawal behavior ([Coan and Allen 2004](#_ENREF_10)). [Henriques and Davidson (1991)](#_ENREF_20) observed that depressed subjects had less left-sided activation than did normal control subjects. The pattern of excessive left-sided frontal alpha activation provides an effective basis for the detection of depression. Besides, for frontal lobe, there are several findings. The frontal cortex is particularly critical in mental processing ([Allen, Harmon-Jones et al. 2001](#_ENREF_1), [Seo, Gil et al. 2008](#_ENREF_45)).

After eliciting specific mental state of a subjet via EEG-based BCI wearable devise, the mental state can be assessed via Self-Assessment Manikin (SAM)([Bradley and Lang 1994](#_ENREF_7)). This technique can directly measure the pleasure, arousal, and dominance from a subject’s affective reaction with non-berbal pictorial assessment techique. This technique classifies a subject’s mental states more systemicly, and very popular in neuro science.

**2.2.2 Magnetoencephalography (MEG)**

Different from the EEG, MEG is a functional neuroimaging technique based on detection of the weak magnetic fields, which is induced by neural oscillation. There are several advantages here: First, because MEG detects the magnetics, which is lesser affected by skull and scalp, this will have higher resolution than EEG and subject’s comfort will improved without lots of sensors on skin. Second MEG’s time resolution can be resolved events with 10 milliseconds or faster, is also higher then EEG even fMRI. Third, preparation time is lesser than EEG considerably. With these advantages, several experimental BCI researches hybrid EEG and MEG.

The brain's magnetic field is usually smaller than the ambient magnetic noise from environment, thus, the weakness of the signal relative to the sensitivity of the detectors, and to the competing environmental noise. Appropriate magnetic shielding can be obtained by constructing rooms made of aluminum and mu-metal for reducing high frequency and low-frequency noise. In the room, there is a large helmet enclosed all region of subject’s head. Usually, the scientists will combine EEG and MEG to read brainwave from overall and partial area ([Bergen 1988](#_ENREF_3)).

**2.2.3 Functional Magnetic Resonance Imaging (fMRI)**

Functional Magnetic Resonance Imaging (fMRI) is a functional neuroimaging procedure using to measure brain activities. With MRI technology, we can get high-resolution pictures by detecting the changes of brain while blood flowing. This technique relies on the fact that blood flow is strongly related to neuronal activation. Blood flow increases in a specific part of brain when this area is in use.

The fundamental principle of detecting is that the fMRI uses the blood-oxygen-level dependent (BOLD) contrast (Ogawa and Lee 1990). The change in the MR signal from neuronal activities is called the hemodynamic response. This method lags the neuronal events triggered about 2 seconds, because it needs to take a time for the vascular system to respond to the procedure that brain gets glucose, so that a peak which could be detected needs about 5 seconds after the stimulus. The peak will spread to a flat plateau while the neurons stay active ([Huettel, McCarthy et al. 2009](#_ENREF_21)).

However, fMRI has some risks and backwards. An fMRI test may lead participants to have Claustrophobia and the radioactive is risky for pregnant women when they are in the scanning process ([Sahito and Wolfgang 2012](#_ENREF_43)). Also, in the scanning process, the high-pitched noises and gradient switching could lead to uncomfortable feeling.

**2.2.4 Conclusion of BCI**

From the information presented in the above subsections, we can compare the difference among these three brain data acquisition methods. EMG needs to set several sensors on facial muscle to detect facial emotion which is not suitable for outdoor context. fMRI can provide high resolution pictures of brain activities with high precision, but the output would delay 5 second before we can formulate the results. Moreover, fMRI device is very expensive. Compared to these techniques, EEG is relatively more suitable for our work aiming to detect mental states in mobile settings because it is portable and affordable devise, for example, NeuroSky and Emotiv are more acceptable for the general public.

**2.3 EEG-based Brain State Recognition using Classification Techniques**

Data mining is a set of techniques to discover knowledge from data or find rules in a huge database. There are two basic methods in data mining: classification and clustering. For classification, a model is trained via feeding a set of observed examples with tagged classes to identify patterns from the data. Clustering different from classification does not need to tag classes on data first. Clustering is the process of grouping a set of data objects into multiple groups or clusters, so that objects within a cluster have high similarity, and at the same time, it is very dissimilar between objects between different clusters ([Han 2005](#_ENREF_17)).

Before feeding data into classifier, we should consider about the imbalanced problem, not to mention that the data from “real world”. Predictive accuracy or others modern evaluation assessment is not a useful when dataset is imbalanced([Chawla 2005](#_ENREF_9), [Blagus and Lusa 2010](#_ENREF_4)). There are three major ways to balance data: down-sampling, over-sampling and Multiple down-sampling. The basic idea of multiple down-sampling is that Neglecting information from the majority class as in simple down-sizing is intuitively unappealing, therefore we considered repeatedly down-sampling with randomly selecting samples from majority class and all samples from minority. Then developed a classifiers on each training set, and evaluated the final result by voting by all classifiers.

This work focuses on the classification task. Among various classification techniques, we identified classifiers which are popular in neuron science field as our candidate classification techniques for building EEG data classifier. Table 1 summarizes the prior works done by researchers in building EEG data classifiers.

Table 1. Prior works on EEG data classification

|  |  |  |  |
| --- | --- | --- | --- |
|  | Classifier | Attributes | Description |
| 1 | Naive Bayes binary classifier | Attention and mediation from Neurosky MindBuilder-EM with 20s frame with 10s overlapping. | Subjects were asked to fill SAM, which is assessed by valence and arousal after playing a brain-controlled video game. After experiment, the serial of mediation and attention were transformed into the mean, standard deviation, maximum, minimum, and slope then become features. The most recognized emotion were engaged (69.59%) and neutral (64%) ([Yoon 2013](#_ENREF_54)). |
| 2 | Neural network with back propagation and Levenberg- Margardt algorithm | Wavelet transformation of Delta, Theta, Alpha and Beta with 3s frame. | Classification of EEG via using epileptic seizure or non-epileptic seizure data and retrieve several bands coefficients by wavelet transform, which considers both frequency and time, and train this ANN with back propagation algorithm. Also, they found the optimum number of nodes in hidden layer is found as 21 In this work, they assess the performance by sensitive and specificity, both are above 90%. Besides, this work also compared the ANN with logistic regression, and ANN obtained higher accuracy in classification ([Subasia and Ercelebi 2005](#_ENREF_48)). |
| 3 | k- Nearest Neighbor classifier | Beta/Alpha ratio with 4s frame. | Assess subjects’ depression degree via EEG signal with two electrodes set in left and right frontal lobe, because in previous studies show that the left hemisphere is good at processing positive emotions, whereas the right hemisphere is more related to process negative emotions. The total number of correct claimed client achieves 100% ([Peng, Hu et al. 2011](#_ENREF_40)). |
| 4 | Neural network with back propagation | Alpha and Beta band | Classify two motor images from two electrodes set in left brain and right brain. In the neural network, this work add he classifier contain finite impulse response filter between the nodes to consideration time series with dynamic weight adjust. This filter could process different data with different via given different time coefficients. In this way, the time variable could be process without process before feeding into classifier ([Haselsteiner and Pfurtscheller 2000](#_ENREF_19)). |
| 5 | Neural network with mutual information | EEG raw data from 16- channel with 80s frame | This work uses the difference of time delay between epilepsy and non-epilepsy EEG data which is produced via mutual information, then fed into probabilistic neural network (PNN) classifier. The accuracy of this work is 100% ([Yuan, Li et al. 2008](#_ENREF_55)). |
| 6 | Neural network and k Nearest Neighbor approach | Spectral entropy and spectral centroid from Alpha, Beta and Gamma which were retrieved by FFT with 5s frame. | Five emotions were defined before, and elicit emotion via video clips. Final result is that KNN performs well than PNN with almost same accuracy but lesser computation time ([Murugappan and Murugappan 2013](#_ENREF_35)). |
| 7 | k Nearest Neighbor classifier | Alpha and Beta with time domain feature (*i.e*., statistical-based features and higher order crossing) and spectral domain features(*i.e*., wavelet transform) | In this work, they use Emotive epoch this device to record EEG data with 14 channels. Eliciting subjects’ emotion with IAPS system and assess arousal degree of emotion via SAM to quality training data. They train KNN classifier with 5-fold cross validation. The emotion recognition rate of 5NN is 77.44% ([Xu and Plataniotis 2012](#_ENREF_53)). |
| 8 | Neural Network | EEG original raw data with sampling rate once per second. | This work records signals from positions P3, P4, C3, C4, O1, and O2 electrodes defined by the international 10-20 systems. They used ANN by virtue of its principal characteristics of learning by example and simplicity. ([Kottaimalai, Pallikonda Rajasekaran et al. 2013](#_ENREF_27)). |
| 9 | Neural Network | The signals were sampled at the rate of 256 Hz. 64 channels | Five emotion: anger, happy, sadness, neutral, surprise. The classification of only 2  types of emotions (neutral and sadness)  produced the highest percentage of correct  classification at 97.50%. The percentage  is then followed by the classification of  3 emotions, 97.20%, 4 emotions, 95.42%  and 5 emotions, the lowest performance at  95.00% ([Yuen, San et al. 2009](#_ENREF_56)). |
| 10 | Fuzzy K-Means | A set of conventional features from db4 of wavelet transform  (power, standard deviation and variance) | They played emotion clips to elicit five emotion which are Disgust, Happy, Surprise, Fear and Neutral, and average classification rate are 80% ([Murugappan, Ramachandran et al. 2010](#_ENREF_36)). |
| 11 | SOM-based SSP classifier | Power of bands from 8-30Hz and gives a frequency resolution of 2 Hz. | In this work, they classified imagination of hand move, calculation of math and cube rotation, and EEG sample rate are once per second. Correct recognition is around 70%([del R Millan, Mourino et al. 2002](#_ENREF_13)). |

2.3.1 Bayes Classifier

Bayesian classifiers are statistical classifiers. They can predict probabilities between different class members, such as the probability that a given tuple belongs to a particular class. Bayesian classification is based on Bayes’ theorem, which is based on mathematical conditional probabilities. It is denoted as given an incident X occurs.

2.3.2 Supervised Artificial Neural Network

Artificial Neural Network (ANN) is an information processing method which is inspired by a neuron interaction in a nervous system. The pattern of ANN consists of three kinds of layer, input layer, hidden layer and output layer. Each layer made by several nodes embeds weights and threshold to connect to the nodes of its neighboring layer. Figure 1 illustrates the structure of an ANN.

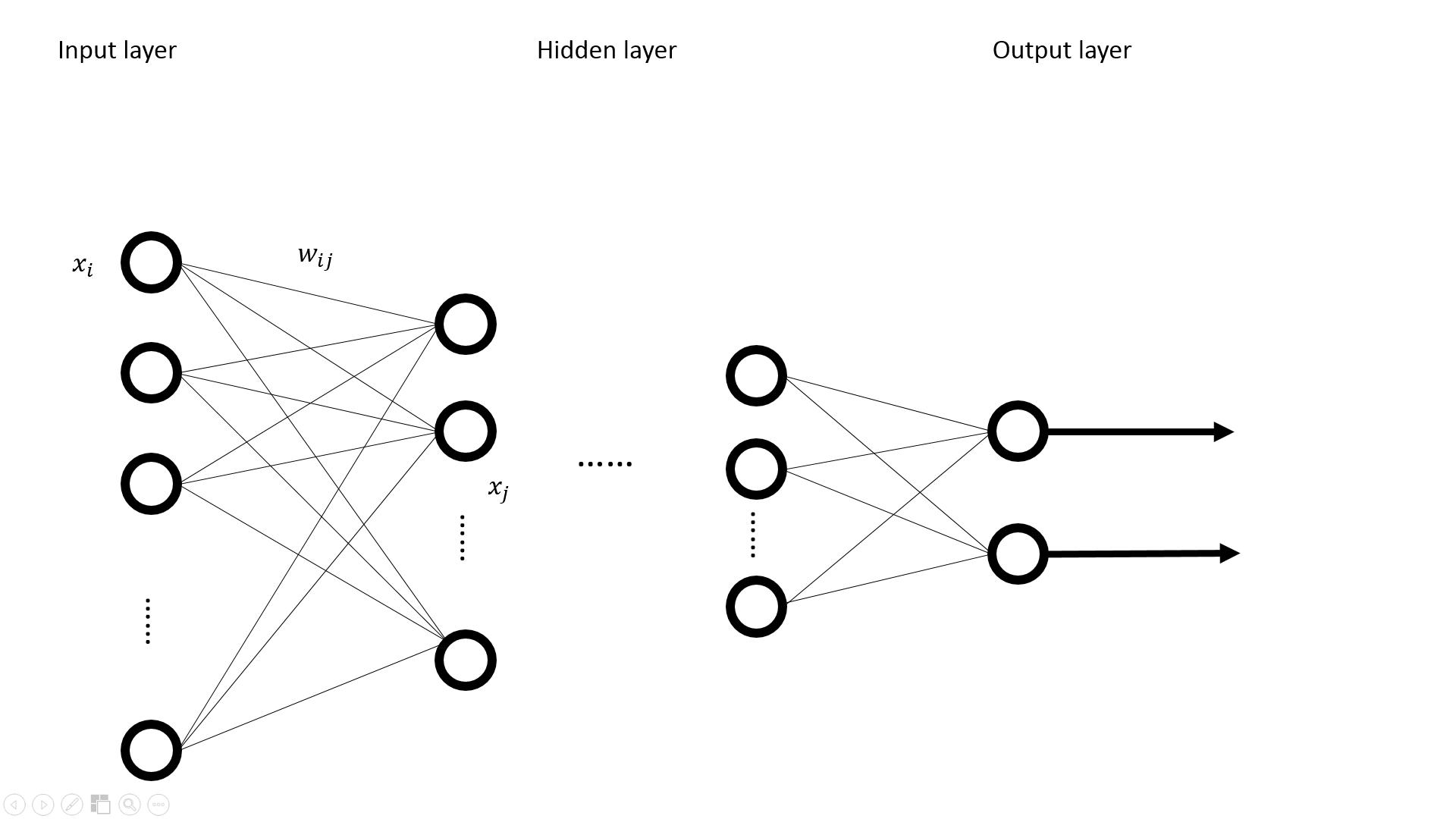


Figure 1. The general structure of an artificial neural network

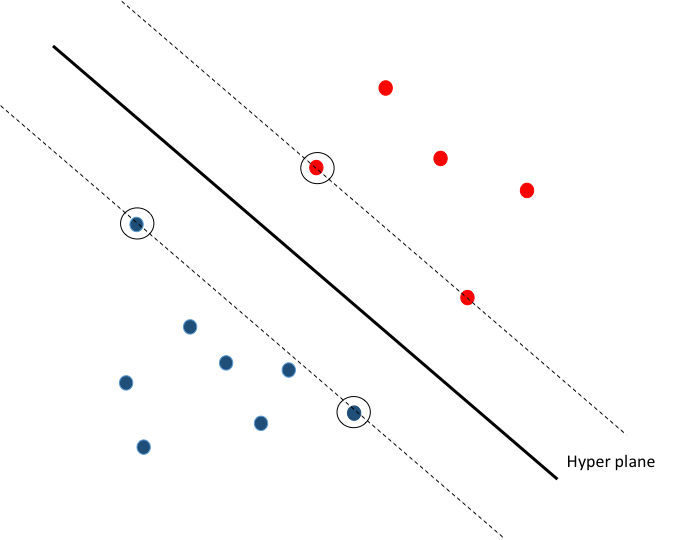
* The input layer: The input data simultaneously is fed into the nodes of this layer and counted via weight.
* The hidden layer: The input data from the input layer have been reconstructed and passed value to the next hidden layer, and so on. A neural network usually contains one or two hidden layers and then feeds the values to the output layer.
* The output layer: This layer is a tuple that output the classes the input tuple is classified to.

The efficiency of ANN is highly related to the learning algorithm. Back propagation (BP) ANN is one of the learning algorithms based on gradient descent as weights change.

2.3.3 K-Nearest Neighbors

This classifier is trained by labelling a point with a class which is the majority of class its k nearest neighboring points belongs to by comparing the distance between points, which is data in an *n*-dimensional space. Then, if an un-labelled data object in the space can be classified into the class where its k nearest neighboring objects. This classifier is the simplest technique in classification.

2.3.4 Support Vector Machine

 Given a set of training dataset, linear SVM uses a discriminant hyperplane to identify classes by maximizes the margins, the distance from nearest training points([Lotte, Congedo et al. 2007](#_ENREF_31)). Besides, the linear SVM was extend to a nonlinear classifier by using the ‘kernel trick’([Jrad, Congedo et al. 2011](#_ENREF_23)). For the dataset which with high dimension and nonlinear, SVM provide a way which mapping data to another space, different kernel would map data into different space. Also, SVM has been proved that such mechanism will have good output for nonlinear and massive datasets.

The circled points positioned on the dashed lines are called support vectors (SV).

2.3.5 The Event Related Potential

Besides, neuron reflection from brain, which is elicited by event, has strong relation with period ([Coles, G.H. et al. 1996](#_ENREF_11)). This method is called EPR (Event-Related Potential) ([Zhang, Kong et al. 2011](#_ENREF_57)). Attractive faces elicited more enhanced ERP amplitudes than did unattractive faces in judgment (N300 and P350–550 msec) and recognition (P160 and N250–400 msec and P400–700 msec) tasks on anterior locations. Moreover, longer reaction times and higher accuracy rate were observed in identifying attractive faces than unattractive faces.

2.3.6 The assessment of Neurosky headset

Here are some emotion classifications via Neurosky Mindset. Crowley(2010) compared the output from the Neurosky Mobile and observers based on Stroop Test. 32 times in 41 times tests (Neurosky device and Observers) in the same categorization, shows that we are able to assess the suitability of the Neurosky headset for measuring the meditation and attention level of an individual. Also, Yoon(2013) used mediation and attention from Neurosky device via Naive Bayes classifier and average recognition accuracy 66.04% ([Yoon 2013](#_ENREF_54)).

**3. Research Methodology**

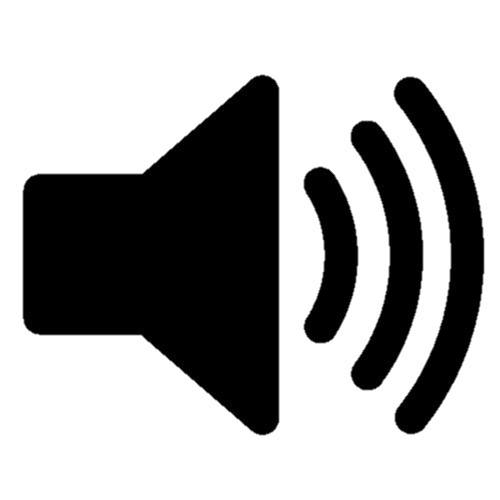
To achieve the research objectives, we need to accomplish two major tasks: (1) to build a English listening system, which can atomically adjust difficulty of tests, in order to capture a subject’s mental states via EEG-based BCI wearable device; (2) to build a mental state classifier based on the data from English listening system and evaluate the efficiency. The following subsections will describe the tentative methods and tools used for achieving these objectives.

**3.1 The Development of the Mental State Recognition System**

3.1.1 English listening system

We built a system for EEG data acquisition; the overview of the architecture is shown in Figure 3. A good database for developing mental recognition system is essential for achieving the mental of classification. We collected data from a series of experiments to build a database with well-defined mental tags.

1. Send EEG signal from Neurosky Mindwave Mobile to smartphone via bluetooth



2. Play audio and receive EEG and user entered data

3. Save all datasets

Figure 3. Architecture of Mental Recognization System

To get EEG data, a subject wear a NeuroSky MindWave Mobile, a wireless Bluetooth headset with only single dry electrode that was placed on the subject’s forehead. This devise provides EEG-related data, *i.e*., the strength of delta, theta, high alpha, low alpha, high beta, low beta, high gamma, low gamma per second, and attention, mediation and strength of blink value can be read from SDK developed by NeuroSky. Figure 4 denotes an example of screenshot of current NeuroSky MindWave, which captures brain waves for additional data processing.

We developed an application providing two different mental state for receiving and mapping these data to different mental states for marking EEG data. Because of the experimental goal is to classify mental state, we recorded the power of all different frequency waves which have a strong relation with mental state according to literature review.

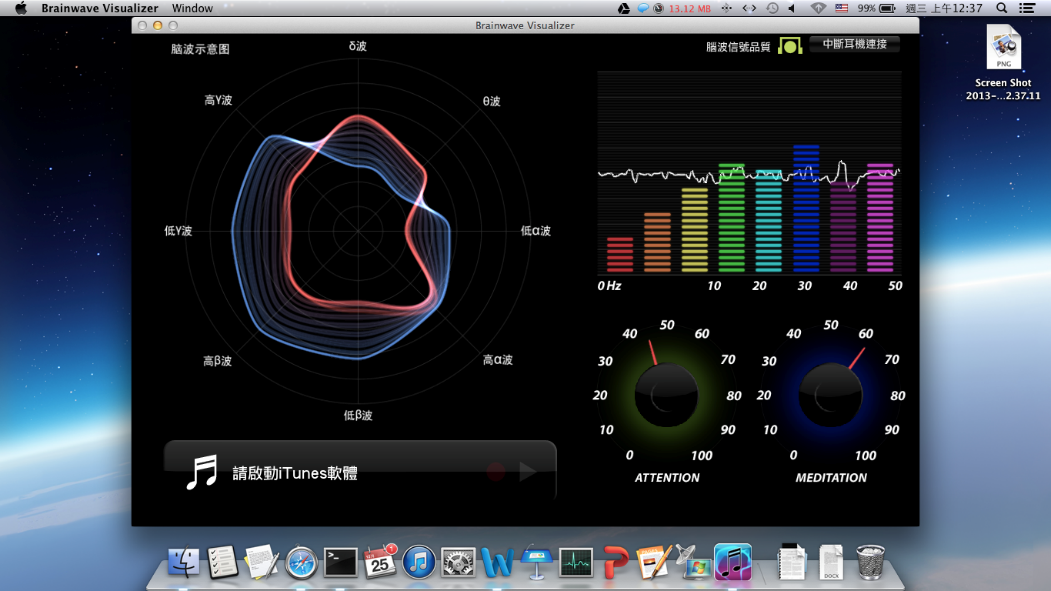


Figure 4. A screenshot of NeuroSky MindWave

3.1.2 The training of a mental state classification model

The first thing to generate a mental state recognition model is to build a classifier. In the work done by [Subasia and Ercelebi (2005)](#_ENREF_48), the accuracy of the classification using neural networks is better than that using logistic regression. However the work done by [Murugappan and Murugappan (2013)](#_ENREF_35) claimed that the KNN has better performance than probabilistic neural network, but the input of KNN is delay time of alpha and beta, which is totally different from original data. [Kottaimalai (2013)](#_ENREF_26) conducted the classification work using neural networks by virtue of its principal characteristics of ANN in learning by doing and simplicity. Among these endeavors, the most impressive work is time-dependent neural network ([Haselsteiner and Pfurtscheller 2000](#_ENREF_19)). It built time filter between inside nodes because there are no rule for find patterns before feeding into the neural network. There are many works about EEG classification based on neural networks, but very few studies have been done based on a novel portable EEG devise. Besides, neural network is the most popular classification methodology in BCI research ([Larsen 2011](#_ENREF_30)). Also, [Garrett, Peterson et al. (2003)](#_ENREF_15) found that nonlinear classifier is better than linear classifier in EEG classification, and so did [Bergen (1988)](#_ENREF_3), who found that nonlinear classifier have better accuracy in mental task differentiation. [Jain, Jianchang et al. (1996)](#_ENREF_22) mentioned that exhibit patterns of ANN similar to those exhibited by humans, so this algorithm is very popular in cognitive sciences. Considering all potential situations encountered in this EEG data classification, we adopt neural networks as the mental state classification model.

All data frame will be sent to the classification model that we were going to develop. The serial of features collected in each seven seconds could turn into a sequent of features, as inputs to the neural network. Then, we will create ten input nodes of the neural network as all waves are recorded. From literature review, we knew that the features of mental states from EEG are highly related to both frequency and time. Thus, this work will also consider both two variables. Because frequency features are processed via FFT, which is a filter embedded in NeuroSky MindWave headset, we only have to focus on time feature. We plan to use two methods to process the data captured before feeding data into a neural network.

* Use mutual information to remove time delay: Before inputting data into a neural network, we could calculate the Mutual Information (MI) of a prior labeled frame data to remove the noise feature when mental states have not been elicited.
* Add time coefficient between nodes: We proposed a time-based approach to process features in a time frame without giving the same weights to the features collected from different time frames. For example, after we retrieve alpha waves from subjects in a seven-second time frame, we set the weights for individual time frames as the multipliers for input nodes as elaborated in Figure 5. Node *O1* receives the sum of the values from input nodes multiplied by time weights and output to the next layer via transformation function, *e.g*., squash function.

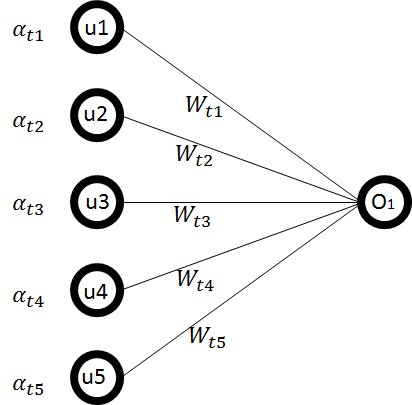


Figure 5. Time-based Neural Network

By taking the value inputted from input nodes, each node on the hidden layer summed up the multiplications of the value of an input node with its weight. Then, the output of each hidden node is generated through a transformation function, e.g., squash function. The output nodes are four possible mental states by taking the sources of value from hidden layer. The architecture of the neural network is shown in Figure 6.

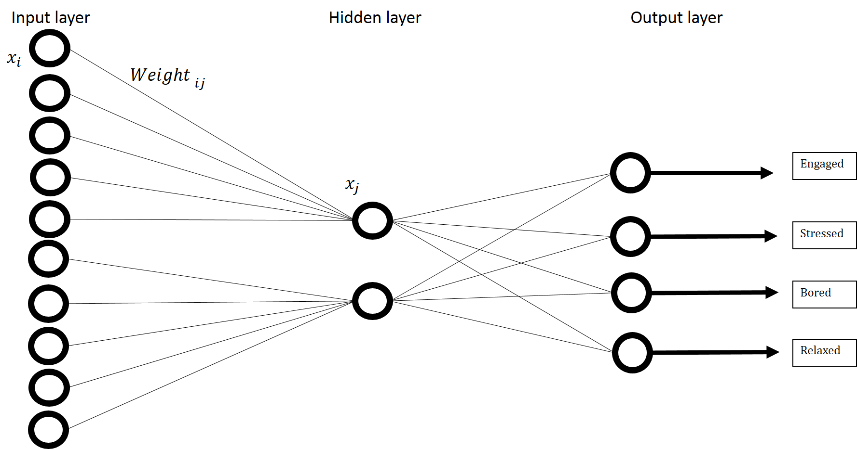


Figure 6. The time-based ANN classifier

In the experiments, the output from ANN classifier has be measured by accuracy, recall rate and precision rate; Precision rate means a fraction of retrieval instances which are relevant instances from previous defined, Recall rate means a fraction of relevant instances that are retrieved. The confusion matrix for the measurement is listed in Table 2, and we will explain our efficiency in result part.

Table 2. Confusion matrix

|  |  |  |
| --- | --- | --- |
|  | P’ (predicted) | N’ (predicted) |
| P (actual) | A | B |
| N (actual) | C | D |

**4. Experimental design**

The main task of the project is to build a mental state classifier in order to capture a subject’s mental states via EEG-based BCI wearable devise. The performance of the system is highly relied on EEG data retrieved from participants. The experimental design for collecting EEG data is described as below:

1. Experimental Setting

Many literatures we studied show that elicitor, setting, focus, and subject awareness (open recording versus hidden recording) are the factors that can affect the mental results. So, we designed a stimulus via playing one hundred audios from English listening test, which provide translations, answers and without mature contents and one CNN student news. To elicit a subject’s mental state without direct or indirect body contact, we played with Stereo Sets and set volume level into a comfortable range to subjects before the process started. The question’s level of difficulty is wired; sometimes subjects may feel very easy, sometimes not for eliciting both easy and difficult mental states. We do this experiment under a controllable environment: meeting room in National Tsing Hua University Library, the spacious room about six square meters with comfortable setting, no noise. Before the subject entered this room, we clean the sets with alcohol for avoiding subjects from both mental worrying and infectious disease. Then we recorded data in the other room to avoid the potential intervention of the subject’s mental state. After elicitations, each subject was asked to express whether they feel the test is easy or not by clicking buttons. The laboratory technicians stayed around the room for assisting subjects if there were any emergency accidents.

1. Experimental Process

There are two processes here: One is for English listening materials, another is for CNN student news. Subjects wear a NeuroSky headset, with a signal dry sensor settled on the frontal scalp, to record each band, from alpha to low gamma, with sampling once per second. EEG data are collected for from a stimulus of seven seconds and subjects were been asked to tag data with difficulty and answer the question by clicking button on the screen with mouse.

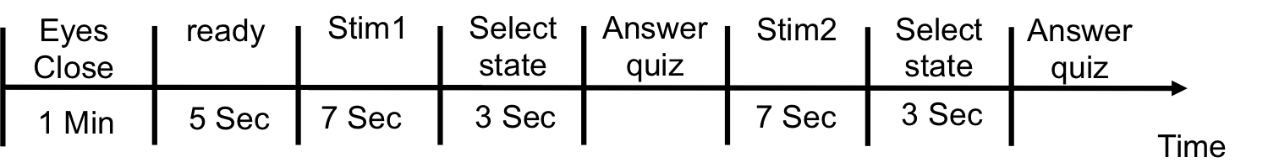
For English listening, there are one hundred elicitation subjects had gone through. Subjects started the experiment by closing his/her eyes for one minute for resting subjects’ states to bring them back to a similar mental state, and then they had five second to get ready, the whole process would be describe in figure 6. For CNN student news, we played it for 10 minutes. Once the subjects don’t understand the content, they can press the button on the screen. 

Figure 6. Experiment procedure

1. Subjects

We recruited 20 individuals (10 male and 10 female) healthy without Epilepsy or others brain damage, right-handed, aged ranging from 20 to 25, to participate in the experiments. To get enough difficult and easy state data for balanced class size, we limited the English ability between:

TOEIC: 600-800

TOEFL CBT: 173-135

We posted experiment information, which includes requirements and experiment process on Internet and social media. After confirming the experimental information, readers registered by providing their email.

1. System evaluation

We evaluated the output by 10-fold cross validation assessing how the results of statistical analysis will generalize to an independent data set. This method divided the dataset into 10 folds and randomly choice one of ten for testing data, rest of these become training data. And this procedure would repeat for 10 times.

1. Experimental scenarios

Our trained mental state classification system will be embedded in the English learning system. Thus, the individualized EEG data can be used for formulated corresponding users’ mental state recognizer when users are conducting learning activities. Then, subjects as learners will be asked to conduct the subject learning until they finish their works with NeuroSky MindWave headset on their heads in library, where has good contexts for learning. In the same time, learner will be asked to record which sentence is understood or not. Finally, we can compare our classifier with user thoughts.

After the progress mentioned here, learner could practice English listening everywhere even in moving. With this service, learner could listening without any actively mark; the parts which is difficult to learner would be marked automatically. Then after learner arrive office or home, they can review the difficult parts directly on the web platform.

**5. Result**

We collected data from 34 subjects, 16 males and 18 females. All subjects listened same materials, but the class sets is very different because the level of listening of each subjects is very different, and the datasets are imbalanced.

**6. Discussion**

**7. Challenges, Limitations and Future work**

In this work, the objective is building a cost-effective service to provide user a better way learning English listening. For this reason, we utilized the commercialized EEG devices and real world setting, which unlike traditional EEG research conducted under carefully controlled laboratory conditions, involved several limitations and challenges.

First, the English Listening materials of this work is from TOEIC test, which could test perception, parsing and utilization ability of subjects, but we could not understand why they feel difficult with such experimental process without collecting the level of understanding. Even subject thought that this question is very easy, but the truth is that they misunderstood this meaning (e.g. the sentence “What’s the language like that” would be misunderstand meaning as “what’s the language you are talking now”, but the two sentences have totally different meaning.). Also, with such English materials, we could not experiment the relation between the length of paragraph and the human difficulty.

The second one is that the class size is imbalanced because the individual English level is very different. Although we built a audio playing system, which could adjust the difficulty automatically based on the current accuracy rate, these classes cannot be the exactly same size. To solve this problem, we down sampled the datasets, but this process lead to smaller datasets.

The third one is that we did experiments for each subject at different time, there are several subjects felt sleepy and can not very focused on the materials. This induced that some instances we can not sure if it is outliers. Maybe next time, such experiment should set eyes moved trackers.

**8. Reference**

Allen, J. J. B., et al. (2001). "Manipulation of frontal EEG asymmetry through biofeedback alters self-reported emotional responses and facial EMG." Psychophysiology **38**: 685-693.

Arslan, B., et al. (2006). A real time music synthesis environment driven with biological signals. ICASSP.

Bergen, D. (1988). "ENcyclopedia of neuroscience." JAMA **260**(1): 104-104.

Blagus, R. and L. Lusa (2010). "Class prediction for high-dimensional class-imbalanced data." BMC Bioinformatics **11**(1): 1-17.

Bos, D. (2007). EEG-based emotion recognition. The Inﬂuence of Visual and Auditory Stimuli. Computer Science, University of Twente.

Bos, D. O. (2006). "EEG-based Emotion Recognition - The Influence of Visual and Auditory Stimuli." Emotion **57**(7): 1798-1806.

Bradley, M. M. and P. J. Lang (1994). "Measuring Emotion: The Self-assessment Manikin and the Semantic Differential " Journal of Behavior Therapy and Experimental Psychiatry **25**: 49-59.

Cahn, B. R. and P. John (2006). "Meditation states and traits: EEG, ERP, and neuroimaging studies." Psychological Bulletin **132 (2)**: 180-211.

Chawla, N. (2005). Data Mining for Imbalanced Datasets: An Overview. Data Mining and Knowledge Discovery Handbook. O. Maimon and L. Rokach, Springer US**:** 853-867.

Coan, J. and J. J. B. Allen (2004). "Frontal EEG asymmetry as a moderator and mediator of emotion." Biological Psychology **67**: 7-49.

Coles, et al. (1996) Event-related brain potentials: an introduction. 1-27

Crowley, K., et al. (2010). "Evaluating a Brain-Computer Interface to Categorise Human Emotional Response." IEEE International Conference on Advanced Learning Technologies **10**: 276-278.

del R Millan, J., et al. (2002). "A local neural classifier for the recognition of EEG patterns associated to mental tasks." Neural Networks, IEEE Transactions on **13**(3): 678-686.

Ertl, M. (2013). "Emotion regulation by cognitive reappraisal — The role of frontal theta oscillations." NeuroImage **81**: 412-421.

Garrett, D., et al. (2003). "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification." Neural Systems and Rehabilitation Engineering, IEEE Transactions on **11**(2): 141-144.

Goldstein, S. M., et al. (2002). "The service concept: the missing link in service design research?" Journal of Operations Management **20**(2): 121-134.

Han, J. W. (2005). Data mining: concepts and techniques. Singapore, Morgan kaufmann.

Harmony, T., et al. (1996). "EEG delta activity: an indicator of attention to internal processing during performance of mental tasks." International Journal of Psychophysiology **24**(1–2): 161-171.

Haselsteiner, E. and G. Pfurtscheller (2000). "Using Time-Dependent Neural Networks for EEG Classification." IEEE TRANSACTIONS ON REHABILITATION ENGINEERING **8**: 457-463.

Henriques, J. B. and R. J. Davidson (1991). "Left Frontal Hypoactivation in Depression " Journal of Abnormal Psychology **100**: 535-545.

Huettel, S. A., et al. (2009). Functional Magnetic Resonance Imaging. Massachusetts, Sinauer.

Jain, A. K., et al. (1996). "Artificial neural networks: a tutorial." Computer **29**(3): 31-44.

Jrad, N., et al. (2011). "sw-SVM: sensor weighting support vector machines for EEG-based brain–computer interfaces." Journal of Neural Engineering **8**(5): 056004.

In many machine learning applications, like brain–computer interfaces (BCI), high-dimensional sensor array data are available. Sensor measurements are often highly correlated and signal-to-noise ratio is not homogeneously spread across sensors. Thus, collected data are highly variable and discrimination tasks are challenging. In this work, we focus on sensor weighting as an efficient tool to improve the classification procedure. We present an approach integrating sensor weighting in the classification framework. Sensor weights are considered as hyper-parameters to be learned by a support vector machine (SVM). The resulting sensor weighting SVM (sw-SVM) is designed to satisfy a margin criterion, that is, the generalization error. Experimental studies on two data sets are presented, a P300 data set and an error-related potential (ErrP) data set. For the P300 data set (BCI competition III), for which a large number of trials is available, the sw-SVM proves to perform equivalently with respect to the ensemble SVM strategy that won the competition. For the ErrP data set, for which a small number of trials are available, the sw-SVM shows superior performances as compared to three state-of-the art approaches. Results suggest that the sw-SVM promises to be useful in event-related potentials classification, even with a small number of training trials.

Kahana, M. J., et al. (1999). "Human theta oscillations exhibit task dependence during virtual maze navigation." Nature **399**(6738): 781-784.

Klimesch, W. (1999). "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis." Brain Research Reviews **29**(2–3): 169-195.

Kottaimalai, R. (2013). EEG Signal Classification using Principal Component Analysis with Neural Network in Brain Computer Interface Applications. IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology**:** 227-231.

Kottaimalai, R., et al. (2013). EEG Signal Classification using Principal Component Analysis with Neural Network in Brain Computer Interface Applications. IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology**:** 227-231.

Kumar, V. (2012). Design Method: A STRUCTURED APPROACH FOR DRIVING INNOVATION IN YOUR ORGANIZATION. New Jersey, John Willy & Sons.

Kwang-Ok, A., et al. (2010). Development of an emergency call system using a brain computer interface (BCI). Biomedical Robotics and Biomechatronics (BioRob), 2010 3rd IEEE RAS and EMBS International Conference on.

Larsen, E. A. (2011). Classification of EEG Signals in a Brain-Computer Interface System. Master of Science in Computer Science, Norwegian University of Science and Technology.

Lotte, F., et al. (2007). "TOPICAL REVIEW: A review of classification algorithms for EEG-based brain computer interfaces." Journal of Neural Engineering **4**(2).

Mak, J. N., et al. (2013). Evaluation of mental workload in visual-motor task: Spectral analysis of single-channel frontal EEG. Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE.

Miltner, W. H. R., et al. (1999). "Coherence of gamma-band EEG activity as a basis for associative learning." Nature **397**(6718): 434-436.

Mostow, J., et al. (2011). Toward exploiting EEG input in a reading tutor. Proceedings of the 15th international conference on Artificial intelligence in education. Auckland, New Zealand, Springer-Verlag**:** 230-237.

Murugappan, M. and S. Murugappan (2013). Human emotion recognition through short time Electroencephalogram (EEG) signals using Fast Fourier Transform (FFT). IEEE International Colloquium on Signal Processing and its Applications 9th**:** 289-294.

Murugappan, M., et al. (2010). "Classification of human emotion from EEG using discrete wavelet transform." Journal of Biomedical Science and Engineering **3**: 390-396.

NEUROSKY (2009) Brain Wave Signal (EEG)

Niedermeyer, E. and F. L. d. Silva (2004). Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, Lippincot Williams & Wilkins.

Niemic, C. P. (2002). "Studies of emotion: A theoretical and empirical review of psychophysiological studies of emotion." Journal of Undergraduate Research **1**: 15-18.

Peng, H., et al. (2011). User-centered depression prevention: An EEG approach to pervasive healthcare. Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on.

Pfurtscheller, G. and F. H. Lopes Da Silva (1999). "Event-related EEG/MEG synchronization and desynchronization: Basic principles." Clinical Neurophysiology **110 (11)**: 1842-1857.

Pine, I. I. B. J. and J. H. Gilmore (1998). "WELCOME TO THE EXPERIENCE ECONOMY." Harvard Business Review **76**(4): 97-105.

Sahito, F. and S. Wolfgang (2012). "Functional Magnetic Resonance Imaging and the Challenge of Balancing Human Security with State Security." Human Security Perspectives **1**: 38-66.

Schacter, D. L. (1977). "EEG theta waves and psychological phenomena: A review and analysis." Biological Psychology **5**(1): 47-82.

Seo, S., et al. (2008). The relation between affective style of stressor on EEG asymmetry and stress scale during multi-modal task. Third 2008 International Conference on Convergence and Hybrid Information Technology**:** 461-466.

Spengler, C., et al. (2010) 360-Degree-Touchpoint-Management - How important is twitter for our brand? Content published in: Marketing Review 14-20

Stickdorn, M. and J. Schneider (2010). THIS IS SERVICE DESIGN THINKING. Netherland, BIS Publishers.

Subasia, A. and E. Ercelebi (2005). "Classification of EEG signals using neural network and logistic regression." Computer Methods and Programs in Biomedicine **78**: 87-99.

Tatum, W. O., et al. (2008). Handbook of EEG Interpretation, Demos Medical.

Vidal, J. (1973). "Toward direct brain-computer communication." Annual review of biophysics and bioengineering **2**: 157-180.

Vladimir, F., et al. (2001). "Multiplicity of the α Rhythm in Normal Humans." Journal of Clinical Neurophysiology **18 (4)**: 331-344.

Wolpaw, J. R., et al. (2002). "Brain–computer interfaces for communication and control." Clinical Neurophysiology **113**(6): 767-791.

Xu, H. and K. N. Plataniotis (2012). Affect recognition using EEG signal. Multimedia Signal Processing (MMSP). Banff, AB**:** 299-304.

Yoon, H. (2013). "Emotion Recognition of Serious Game Players Using a Simple Brain Computer Interface." IEEE ICTC: 783-786.

Yuan, Y., et al. (2008). Delay Time-Based Epileptic EEG Detection Using Artificial Neural Network. Bioinformatics and Biomedical Engineering**:** 502-505.

Yuen, C. T., et al. (2009). "Classification of Human Emotions from EEG Signals using Statistical Features and Neural Network." International Journal of Integrated Engineering **1**(3): 71 - 79.

Zhang, Y., et al. (2011). "Identifying Cognitive Preferences for Attractive Female Faces: An Event-Related Potential Experiment Using a Study-Test Paradigm." Journal of Neuroscience Research **89**: 1887-1893.