CHAPTER 2

THEORETICAL FOUNDATION

2.1 Affective Computing

The study of Pickard [2] concluded that human emotions are inseparable fragment from human intelligence. Human intelligence in the form of rational decision making, social interaction, perception, memory, learning process, and creativity that are necessary in day-to-day functioning is influenced by human emotion. In a process of decision-making and social interaction, for example, even tiny emotion may turn the decision and interaction to different ways. The importance of human emotion to human intelligence, therefore, has gained wide attention from various research communities.

The study of Ekman [1] on emotions in relations to facial expression concluded that facial expression is universal and can be broadly categorized into eleven categories namely: Amusement, Contempt, Contentment, Embarrassment, Excitement, Guilt, Pride in achievement, Relief, Satisfaction, Sensory pleasure, and Shame. Classifying human emotion is rather an easy problem for human. Regardless age and race, people can easily detect whether someone is happy, anger, or sad by looking into his or her face expression. However, classifying human face expression is a hard problem for computer.

Researcher has been trying to create a computer-based solution for rational decision making and problem solving, involving perception and learning process. Adding the fact that emotions has an important role in those expected abilities, Picard [2] has reported that to be genuinely intelligent and to be able to interact naturally with human, a computer has to have the ability to recognize, understand, have and even express emotions. Later, Pickard [2] coined Affective Computing as a research field that deals with emotion and computers. In this research field, it is necessary for researcher to have sufficient amount of knowledge on emotions and cognitive processes (judgments, evaluations, or thought), even though the emotion themselves has no clear theory and understanding [3].

2.1.1 Emotion and Intelligence

There is a plethora of definition about emotion ranging from the physiological changes caused in our body to purely intellectual thought-process [3]. In addition, there is another definition which views emotion as a mixed of complex physiological experience of human's state of mind in interacting with internal and external influences. Internal refers to biochemical and external refers to environmental factors [4].

Many published works on human emotion have proposed various emotion classification. According to Plutchik [5] human emotion can be broadly divided into eight basic emotions and eight advanced emotions. The eight basic emotions are joy, trust, fear, surprise, sadness, disgust,

anger, and, anticipation. The rest of eight advanced emotions, each are combination of two basic ones.

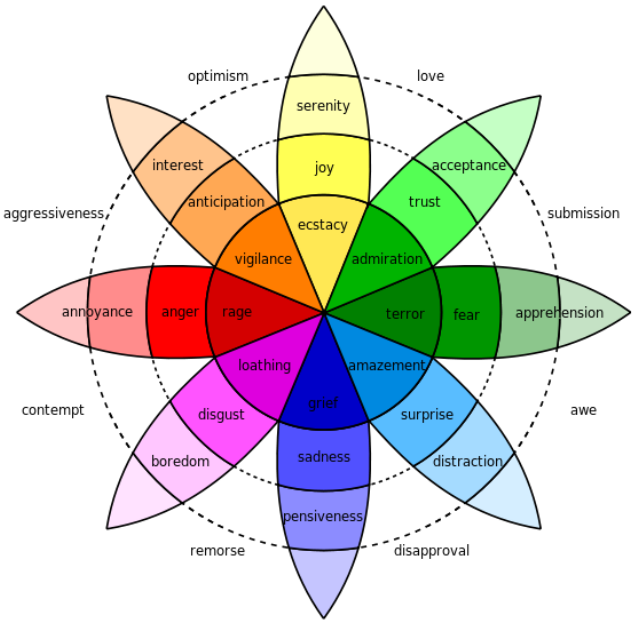

Figure 2.1 Plutchik's Wheel of Emotion (Source: upload.wikimedia.org)

Human recognizes emotions through visual sense, audio, and perception. This is proven by emotional misunderstanding that caused by non-face-to-face communication such as e-mail and texting. The main problem in the following kind of communication is the lack of sensory data; in e-mail and text message, information can be emotionally interpreted in many ways, the actual affective meanings may not be delivered to the intended recipient. Emotions are fundamental in human behavior; it can hijack human rationality in certain situation such as fear,

panic, or excitement [1]. It strongly affects human decision making, perception, learning process, and memory.

Many studies sugested that human tends to associate judgments of value and feelings in making an important decision. Before making decision, human will consider the fact and then involves the feelings. The same goes with perception. In learning process, human first creates an emotion for learning, and then the stimulus response [1]. The emotion that influences decision making and learning also affects human memory retrieval. It is proven that good feelings will bring positive outcomes, while bad feeling such as under pressured, unfamiliarity, risk will bring negative outcomes.

2.1.2 Affective Computing

Affective computing is an emerging research field that studies and develops a system that recognizes, interprets processes, and stimulates human emotions [7]. It consists of four relate areas: Recognizing emotions, expressing emotions, having emotions, and emotion intelligence.

Recognizing emotions is the foundation of affective computing. The system should be able to conclude an emotional state from the emotional expressions observation. In order to achieve this ability, the system needs to gather information such as audio, video, facial expression, vocal

intonations, even information that cannot be generated by human sense such as reading temperature. Using this data, the system will be able to discover situations that probably lead to the expression. Then it will be able to sensed and recognize human emotions, that will be used in emotion generation. This ability alone seems to be enough in enabling the affective computing. However, affective computing does not only rely on recognition in interpreting emotions, which is why other three abilities need to be added in order to make the system emotionally capable.

Picard [7] argued that in order to communicate with human effectively, a computer system should be able to recognize human emotion and express it emotions. Therefore, a system has to fulfill the basic requirement: have channels of communication in visual and audio, and be able to communicate affectively over those channels [1]. A system is then considered has the affective communication ability if the system able to express emotions that makes the user felt emotional, such as feeling happy. However, a system can express emotions without having real emotions, as it will do as it's programmed. Having emotions and expressing emotion are inseparable; having emotions affects the system ability to express them.

Picard [2] further proposed a five components model that should be present in a system that exploits human emotions. The first component is emergent emotions. It is attributes to the system based on observing the emotional behavior. The second component is fast primary emotions,

following by cognitive emotions as the third component. This cognitive emotion could be a satisfying feeling after completing a difficult task. Emotional experience is the fourth component and the body mind interaction is the last one. Emotion influences cognitive process that also makes it influence intelligence.

A system that possesses emotional intelligence will be able to understand, express, and recognize emotions. It should be able to use the emotions to motivate positive behavior. By recognizing emotions, a system can reasoning emotion based on the situation, it's able to understand other person. These forms of emotional intelligence are depending on the previous three areas of affective computing.

The goal of affective computing is to have a system that capable of communicating affectively, also to create a problem solving intelligent computers that will be able to do rational decision making and problem solving, involving perception and learning process. Picard [7] has stated that the goals are to reduce user frustration in using technology, enabling comfortable communication between system and user, develop infrastructure and application that able to handle affective information and lastly to build a tool that help the development of social-emotional skills.

2.1.3 Facial Affect Detection

Facial affect detection is one of many areas in affective computing. It detects and process human facial expressions, recognizing the expressions as emotions. It uses pictures or videos that contains human facial expression as the input, process and classify the input and produce the output in a form of specified emotions.

Figure 2.2 Facial Affect Detection (Source: www.doc.ic.ac.uk & cameronsublett.files.wordpress.com)

Facial expression is believed to be closely related to emotions [1]. Facial expression is the movement of one or more muscles in the face skin [8]. These motions of muscle(s) convey the emotional state of the person. Facial expression is involuntary action that is nearly impossible to be avoided.

The prominent approaches to classify human emotions can be divided broadly into to categories: (i) supervised training classifier, and (ii) unsupervised training classifier. Among those proposed methods form supervised training classifiers are: Artificial Neural Network and Support Vector Machine.

2.2 E-Learning

The proliferation of internet access and the advance of computer technologies have promoted E-Learning to be adopted in education institution worldwide. Further, various intelligent devices have made E-Learning facilities to be easily accessed anytime and anywhere.

E-Learning is a term which refers to the entire forms of the learning and teaching process using electronic devices. This term also refers to out-of-and-in classroom education process that experienced through technology. The lecture can be delivered in a form of CD, video tape, audio, or via internet; students do not have to physically attend class. According to Holmes and Gardner [9], E-Learning process which uses online access to learning resources to experienced learning process, anywhere and anytime via electronic devices. E-Learning

provides opportunities for learners and educators to improve the learning and teaching experiences.

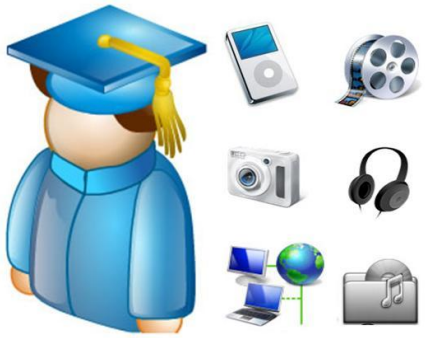

Figure 2.3 E-Learning media (Source: ticonsultor.net)

Historically, Suppes and Atkinsons [10] who are both psychology professors of Stanford University, researched and experimented in teaching mathematics and reading to students in East Palo Alto elementary schools using computers. The early E-Learning system was based on computer-based learning. This study was followed by Graziadei in 1993 that delivered an online computer-delivered lecture using e-mail and later, a strategy of course development and management for an educational system using technology-based was conceived. Nowadays E-Learning is using combination of many technologies from collaborative software to blogs.

In E-Learning, any physical limitations in accessing knowledge are unconcerned. Learners will be able to choose the location that supports their comfort and at the time that suits them. E-learning also contributes to the learning process evolution [10]. Learners are encourages to actively seeking

information and involved in learning process. It will extend and enrich the learning and teaching experience for both the students and the teachers.

E-learning services is a promising industry with great business opportunity. The worldwide industry of E-Learning is estimated to be worth over $48,000,000,000 [9]. E-Learning is easy to use, maintain, portable, and long-term cost effective. The development of technologies enables E-Learning to grow.

E-Learning environment has been widely claimed as supporting learning experience delivery, exploration, and application of the knowledge. However, managing E-learning environment involves some level of complexity. Not only the environment itself such as online lecture, discussion group, online tutorials, online quizzes and assignments, but also the learner's details, examinations, and assessment profiles should be observed [9]. Interaction in E-Learning activity is considered as asynchronous communication as the communication is not happening in the same time frame, such as email and discussion board. Despite some positive claims of E-learning system, many questions have been arised in related to effectiveness of the system. There are still many open problems which need to be solved in order to make e-learning effective. First, it has been widely accepted that not all subjects are appropriate to be delivered by e-learning system. Delivery of many subjects required some close interaction between teacher and learners in class such as those subjects which teach some manual or technical skills. Second, limitation on the tutor and students interaction. The lack of face-to-face interaction that can be found on the in-class learning process might be a limitation of E-Learning. This face-to-face interaction provides

guidance to the tutors on how well the students understand the subject. Direct communication such as asking questions to test how deep the student understanding is also not possible. Lack of automatic response to learner progress in learning process has questioned the role of e-learning system in replacing the role of class teacher. These concerns has become motivation of a large body of research in combining Affective Computing and E-learning. Finally, Assesment format in E-Learning system. The assessment should be able to show how well the learners mastery on the given subject. It also preferably machine-marked-able based on the correct answers provided by the tutor.

2.3 Artificial Neural Networks

It has been widely accepted that solving a complex problem by divide and conquer method promises better performance than brute-force methods. The divide-and-conquer method divides the complex problem into number of simpler elements that makes it easier to solve, later the partial results are combined to form the whole results.

Artificial Neural Networks (ANN) is one of the approaches in this divide and conquer approach. ANN is a computational model that consists of interconnected group of artificial neurons. It is inspired by the natural neurons. It consists of inputs that are multiplied by weights or the strength of the respective signals. The inputs then will be computed by mathematical function that will determine the neuron activation. It will also produce output.

Each node of an ANN can be considered as computational units. The nodes will receive inputs, process the inputs and produce an output. The processing could be very simple such as summing or subtracting the inputs, or contain more complex mathematical computation. The connections between the nodes define the flow of information. The flow can be unidirectional or bidirectional. Unidirectional when the flow of information only in one sense, bidirectional when the in either sense. These interactions between nodes will lead the networks to a global behavior, that is said to be emergent [11]. Emergent means that the network has the abilities to replace one of its elements. This makes networks a powerful tool. Many systems can be considered as networks, that is why networks are often used to model physic, computer science, mathematics, economics, and biology system.

The strength of the input will be multiplied by the height of artificial neuron weight. The computation of the neuron will depend on the weights. In order to achieve a desire output specified by the inputs, the weight of artificial neuron should be adjusted. This process is called learning or training process. Learning or training process could be a problem when the ANN involves hundreds or thousands of neurons. That is why learning algorithm is needed. This algorithm is mathematical computation that will adjust the weights of ANN to obtain desired output.

Historically, the first neural model is developed by McCulloh and Pitts in 1943 [11]. Since then, hundreds different models of ANN has been developed. The differences are located on their function, inputs values, learning algorithm,

and the topology itself. There is also number of hybrid ANN models that combines the previous model. The most common model of ANN is the one that is developed by Rumelhart and McClellan in 1986. The ANN uses back-propagation algorithm.

The ANN below is a network that consists of four artificial neurons. The neurons are divided into two which are two input neurons that will receiver inputs and two output neurons that will give outputs. Each arrow represents the information flow which is also assigned by weights.

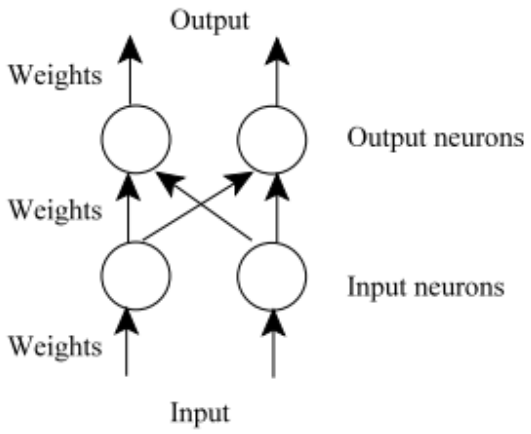

Figure 2.4 Simple Artificial Neural Network [11]

The weight will be multiplied by the values that go through each arrow. These network neurons only sum their inputs. The input neurons only have one input, therefore the output will be the input received multiplied by weight of the arrow. The output neurons have 2 inputs from the inputs neurons, the output of the input neuron will be multiplied by the weight of the arrow and then summed

by output neurons. Then the output of the output neurons will be multiplied by the weight of the arrow and produce the final output.

The weights of the network below are all set to be equal to one and the inputs are (1, 1).

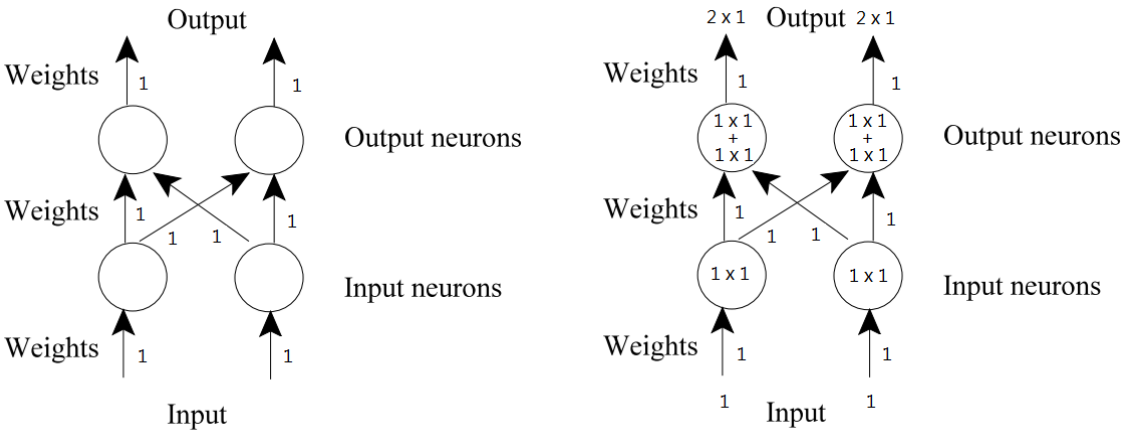

Figure 2.5 Networks with identical digit weights

The inputs (1, 1) will be first multiplied by the weight of the arrow: 1 x 1 = 1. This number will be inputted to the input neurons. Since there is only one input of input neurons, the output of the input neurons is still 1. This output then will be multiplied by the weight of the arrow: 1 x 1. The output neurons however have two inputs; the two inputs will be summed by the output neurons: 1 x 1 + 1 x 1 = 2. The output of the output neurons will be multiplied again by the weight of the arrow that produces the final output: 2 x 1 = 2. Therefore, the output of the following network is (2, 2).

The following networks will be slightly modified. One of the weights of the arrow will be changed to 0. It then will compute the same inputs as the previous network: (1, 1).

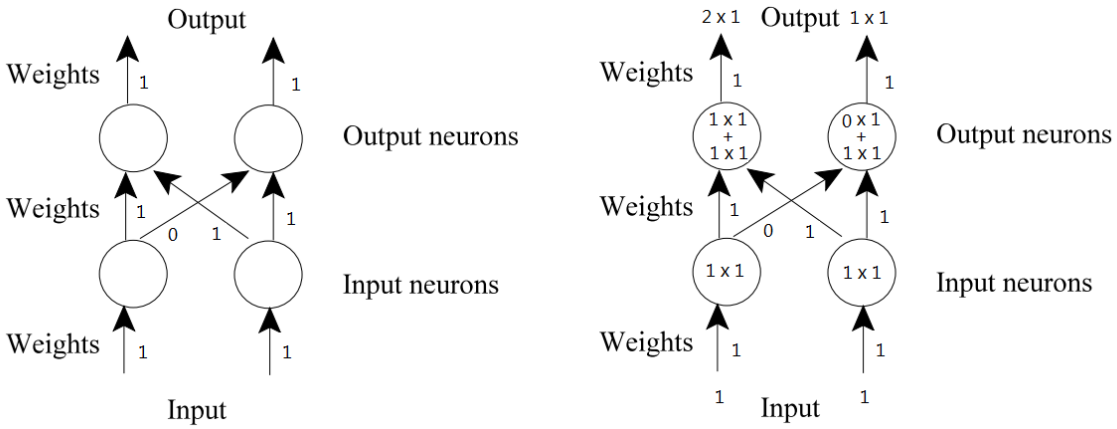

Figure 2.6 Networks with different digit weights

Following the previous simple computation, the network produces different outputs: (2, 1). It can be seen that by slightly changing the weight, the output will be completely different. The output of the network can be controlled, for example to produce inverse output from the given input; it can be achieved by only changing the weight. The next thing to do is setting the thresholds to the neurons. If the previous output of the neuron is greater than the threshold of the neuron the output of the neuron will be one, and otherwise zero [11]. The thresholds also can be adjusted to achieve desired outputs.

2.3.1 The Back-propagation Algorithm

Although adjusting the weights and thresholds of a small ANN can be easily performed manually, the computation burden will be less effective to be done manually when the number of nodes is large. In dealing with

the larger and more complicated network, where the adjustment capabilities are limited, the back-propagation algorithm can be used. This algorithm is used in layered feed forward artificial neural networks that mean the ANNs are organized in layers. It sends the signal forward and the errors will be propagated backwards. There are three layers in this ANN, input layer which receives inputs, output layer that produces output, and hidden layers [11]. This algorithm also uses supervised learning. It means that in order to make this network works, example of inputs and desired outputs should be provided for the network to learn. The learning process or usually called the training process begins with random weights assigned to the networks. Then the network will compute the example inputs and desired outputs. If the inputs do not come with desired output, it will be counted as error. The network will continue by adjusting the weight and computes the example inputs and desired output again until the error is minimized.

2.3.2 Advantages and Limitations of Artificial Neural Network

Artificial Neural Network is used to recognize and extract pattern that are complex and un-noticeable by human or computer techniques [12]. ANN is considered to be the expert in solving patterns and classification problem. The adaptive learning of ANN enables the network to learn based on the training data given and complete the further classification. It is also a self-organization network as it creates its own organization that

represents the training data. The computations in ANN are carried out in parallel and real time operation. ANN is also fault tolerance via the redundant information coding [12].

ANN has a lot numbers of advantages; however it is also has limitations. The back-propagation algorithm that is used by ANN functions as the network "black box" because of its nature to operate in completely unknown rules [12]. The architecture of the network initially input the network with random values, it makes the user has no other role that feed the input to the network. The algorithm is also tend to be slower in training the networks and might requires thousands of epochs before meeting the desired result [13]. Using this algorithm requires a large amount of computational powers and might take more time to complete depending on the machine used and the amount of the training data.

2.4 Support Vector Machine

Support Vector Machine or SVM is a statistical based classifier to classify data that is trained using supervised learning approach [14]. SVM is used to classify data into one of object classes by exploiting the data pattern learned from the trained dataset.

During the supervised training process, each training data is marked based on its class and SVM will build a model based on the given training data set. This SVM model is represented as points in space, divided by a gap that separates

each class. The new data then will be mapped on the same space and predicted belong to which class based on their location on which side of the gap on the space.

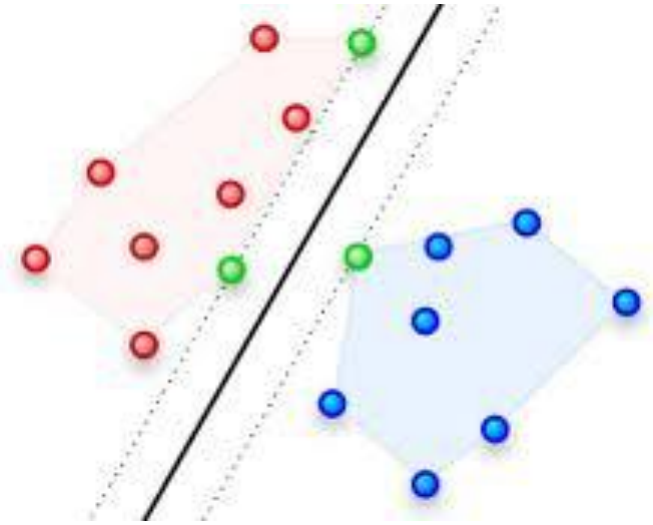

Figure 2.7 Support Vector Machine (Source: www.m8j.net)

SVM was introduced by Boser and Vapnik in Conference on Computational Learning Theory (COLT) in 1992 [15]. SVM was developed from Statistical Learning Theory also known as Vapnik–Chervonenkis theory. This theory is developed by Vapnik and Alexey Chervonenkis during 1960 – 1990 [16]. SVM was introduced as a concept in pattern recognition method and knows for its good performance. The current standard of SVM was suggested in 1995 by Vapnik and Corinna Corter [15].

The main goal of SVM training is to find the best hyperplane that separates the two classes on the space [17]. The Figure 9 below shows pattern from 2 different classes. Red squares as Class -1 and Yellow circles as Class +1.

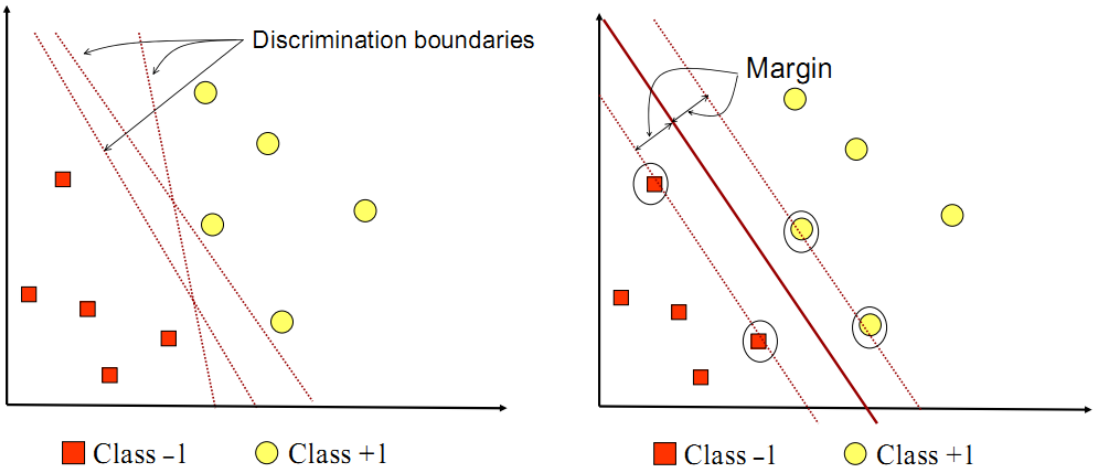

Figure 2.8 Discrimination Boundaries and Margin on SVM [14]

The discrimination boundaries on the left figure are alternatives hyperplanes produces by SVM. The hyperplane can be found by calculating the hyperplane margin and find the maximal points. Margin is the distance between hyperplane with the nearest pattern from each class. The nearest pattern on the space are called support vector. The right figure shows the best hyperplane of the SVM, and the patterns that circled are the support vector. This process is called SVM training.

2.4.1 The Soft Margin

Vapnik and Cortes proposed a modified maximum margin concept to improve training process [14]. This soft margin is used when there is no hyperplane that able to separate the two classes. It will choose a hyperplane that separates the classes as clean as possible and at the same time still maximizing the distance of the nearest patterns or the support

vectors [15]. This method makes use of slack variables that measures the degree of miscalculation.

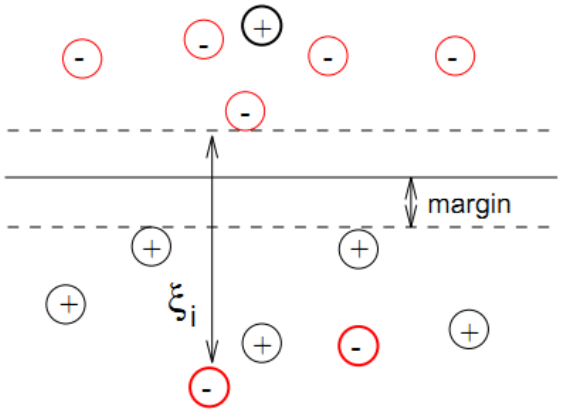

Figure 2.9 Slack Variables (ξi) (Source: www.nec-labs.com)

2.4.2 Advantages and Limitations of Support Vector Machine

Support Vector Machine uses a selection of kernel function that makes the computational load reasonable [14]. The kernel choices make the SVM flexible and accurate [18]. SVM provides a good out-of sample generalization and robustness. The absence of local minima enables SVM to deliver unique solution while ANN delivers multiple solutions that might not be robust over different data sample [18]. SVM also obtained the capacity control through optimizing the margin [19].

There are some limitations to SVM methods; one of them is although SVM has good generalization, the test phase might be slow [20]. SVM also uses high algorithmic complexity and require large amount of memory to

compute. The selection of kernel that makes SVM flexible might also be its limitation as the wrong selection might bring un-desired result [20].

2.5 Comparison between ANN and SVM

ANN development process follows a heuristic path, starts with applications and extensive experimentation of previous theory while SVM involved the theory and then the implementation and experiments [21]. Multiple local minima in ANN resulted in the multiple solutions that might affect the robustness of the method.

Table 2.1 Comparison Table of ANN and SVM (Source: www.svms.org)

	Artificial Neural Network	**Support Vector Machine**
Development Process	Using Neural Network and Back-propagation Algorithm	Using Hyperplane and Soft-Margin
Multiple Local Minima	Yes	No
Solution(s)	Multiple Solutions	Unique Solution
Risk Minimization	Empirical	Structural
Computational Complexity	Depends on input space	Does not depend on input space

Computational Time	Slower	Faster

From Table 2.1 it can be seen that SVM might outperform ANN in practice, however studies prove that both methods performance will depends on the size of dataset [22] and the class type of data (linear or multi-class) [23].