Chapter XXX
Machine Learning for Designing an Automated Medical Diagnostic System

Ahsan H. Khandoker
The University of Melbourne, Australia

Rezaul K. Begg
Victoria University, Australia

ABSTRACT

This chapter describes the application of machine learning techniques to solve biomedical problems in a variety of clinical domains. First, the concept of development and the main elements of a basic machine learning system for medical diagnostics are presented. This is followed by an introduction to the design of a diagnostic model for the identification of balance impairments in the elderly using human gait pattern, as well as a diagnostic model for predicating sleep apnoea syndrome from electrocardiogram recordings. Examples are presented using support vector machines (a machine learning technique) to build a reliable model that utilizes key indices of physiological measurements (gait/electrocardiography [ECG] signals). A number of recommendations have been proposed for choosing the right classifier model in designing a successful medical diagnostic system. The chapter concludes with a discussion of the importance of signal processing techniques and other future trends in enhancing the performance of a diagnostic system.

INTRODUCTION

Machine learning is the study of algorithms and techniques that allow computers to "learn." It refers to an intelligent system that makes decision based on the autonomous acquisition and integration of knowledge from accumulated experience contained in successfully solved cases, analytical observation, and other means. A machine learning system uses many different mathematical methods for exploiting the computational power of a computer. In many professional fields, good tests and measurements may be available, but methods of applying this information to solve a problem may be poorly understood. Physicians, for instance, are always searching for the best possible measures to

make a particular diagnosis at a very early stage. From a medical diagnostic system design perspective, there are several reasons why there has been a considerable interest in machine learning systems. The argument in favour of learning systems is that they have the potential to discover new relationships among concepts and hypotheses by examining the record of successfully solved cases. In biology and medicine, where expertise in understanding the complex physiological functions is limited, these learning systems may aggregate knowledge that has yet to be formalized. In this chapter, we confine our attention to the most prominent and basic learning task, that is, *classification*.

As illustrated in Figure 1, the fundamental goal of the machine learning system is to extract the generalized decision rules from sample data that will be applicable to new data. Then the learning system can be viewed as a classifier that produces a decision for new data to be classified. A typical learning system is designed to work with a classifier model such as a neural network, support vector machines, or a discriminant function. Learning helps to choose or adapt parameters within the model structure that work best on the samples at hand. For example, in medical diagnosis, the physician has observations and test results, and the objective is to pick the correct diagnosis. The objective of the machine learning algorithm is to customize

the classifier structure to the specific diagnosis by finding a general way of relating any particular pattern of symptoms to one of the specified diseases or conditions.

Machine learning systems have found many valuable applications in early diagnosis of diseases so that appropriate intervention can be exercised to achieve better outcomes (Begg & Palaniswami, 2006a; Ifeachor, Sperduti, & Starita, 1998; Teodorrescu, Kandel, & Jain, 1998). As biomedical diagnostic systems are becoming more specialized and complex, they adapt to new methods, instrumentation, and assay technologies that were originally developed for other applications; for example, defense, energy, and aerospace have found applications in the medical industry/environment. In this chapter, we will focus on two applications of machine learning approach in designing: (1) a diagnostic system for balance impairments based on human gait signals and (2) a diagnostic system for sleep apnoea syndrome based on electrocardiography (ECG) signals. Needless to say, such systems do not mean to replace the physician from being the decision maker but, rather, they attempt to enhance the physician's abilities to reach a correct decision.

Figure 1. A General overview of a machine learning system for classification
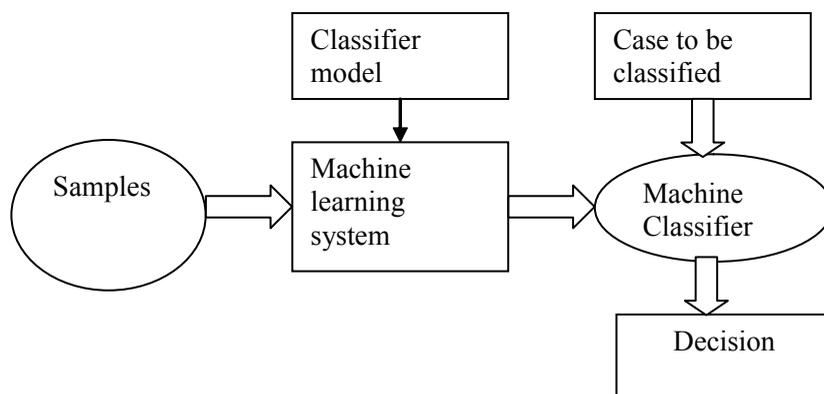
BACKGROUND

Case1: Modelling Human Gait

Gait analysis refers to the systematic recording and analysis of human walking patterns. This analysis is frequently undertaken in clinics and rehab centres to gauge the extent of abnormality in the lower limbs and also to evaluate overall walking capability. Besides disease, gait patterns change with age. There have been numerous studies undertaken in recent years focused on identifying gait measures that are affected as a result of the aging process and also that would indicate declines in the gait performance. Such declines threaten the balance control mechanisms of the locomotor system and have been reported in many gait measures (Begg, Sparrow, & Lythgo, 1998; Winter, 1991), and might be linked to reasons for falls in the older adults. Falls in the older population have been identified as a major health issue in Australia, costing the community $2.4 billion per annum (Fildes, 1994). While some research in aging gait has investigated time-distance variables (e.g., walking speed, stance/swing times, and step length) (Ostrosky, VanS-wearingen, Burdett, & Gee, 1994) to identify key variables of gait degeneration in the elderly, it has been suggested that more sensitive gait variables such as minimum foot clearance (MFC) during walking over the walking surface should be used to describe age-related declines in gait in an effort to find predictors of falls risk (Karst, Hageman, Jones, & Bunner, 1999).

Minimum foot clearance during walking, which occurs during the mid-swing phase of the gait cycle, is defined as the minimum vertical distance between the lowest point on the shoe and the ground. This has been regarded as an important gait gauge as this will allow successful negotiation of the environment in which we walk. At this event, the foot travels very close to the walking surface (mean MFC height has been reported to be ~1.29 cm) with a considerably high forward velocity (4.6 m/s) (Winter, 1991). This small mean MFC value combined with the vari-ability in MFC data (0.5-0.62cm) can potentially cause tripping during walking, especially for un-seen obstacles or obstructions. The literature also suggests a decrease in MFC height (1.11 cm) with aging (Winter, 1991), thereby providing a strong rationale for MFC being associated with tripping during walking, and implications for trip-related falls in older population. A model is therefore necessary that would associate MFC information with falls-risk individuals so that MFC features could be utilized to diagnose potential falls-prone individuals.

Neural networks as a machine learning method have found widespread applications for gait pattern recognition and clustering of gait types, for example, to classify simulated gait patterns (Barton & Lees, 1997) or to identify normal and pathological gait patterns (Holzreiter & Kohle, 1993; Wu, Su, & Chou, 1998). Neural networks have also been useful in the automated recognition of aging individuals with balance disorders using gait measures (Begg, Kamruzzaman, & Sarker, 2006b). Further applications in gait and in other clinical biomechanical areas are found by Chau (2001) and Schollhorm (2004). Recently, support vector machines (SVMs), a machine learning technique, have been shown to be a powerful tool for learning from data and for solving classification and regression problems with superior classification performance (Chappelle, Haffner, & Vapnik, 1999; Vapnik, 1995). Recently, we have shown that MFC gait kinematics is use-ful in discriminating young/elderly adults as well as healthy/falls-risk elderly adults using SVMs (Begg, Palaniswami, & Owen, 2005a; Khandoker, Karmakar, & Palaniswami, 2007). These research outcomes also highlight that gait features carry useful information regarding the quality of gait and information relating to its functional status so that these characteristics can be used to detect declines in gait performance due to aging or pathology.

Case2: Sleep Apnoea Syndrome

Sleep apnoea syndrome is a medical condition caused by sleep apnoea which is defined as the ces-

sation of breathing for short periods during sleep. It is a common sleep related problem with a reported prevalence of 4% in adult men and 2% in adult women (Young, Palta, Dempsey, Skatrud, Weber, & Badr, 1993). When breathing does not stop but the volume of air entering the lungs with each breath is significantly reduced, then the respiratory event is called a hypopnea. Obstructive sleep apnoea (OSA) is characterized by intermittent pauses in breathing during sleep caused by the obstruction of the upper airway. The airway is blocked at the level of the tongue or soft palate, so that air cannot enter the lungs in spite of continued efforts to breathe. This is typically accompanied by a reduction in blood oxygen saturation and leads to wakening from sleep in order to breathe. Each apnoea event is defined as a respiratory pause lasting for at least 10 seconds.

Excessive day-time sleepiness is the most common complaint. An increased risk of accidents and a link between sleep apnoea and arterial hypertension have been proven in recent large-cohort studies (Nieto, Young, Lind, Shahar, Samet, Redline, et al., 2000). Sleep apnoea is now regarded as an important risk factor for the development of cardiovascular diseases (e.g. hypertension, stroke, congestive heart failure, and acute coronary syndromes) (Young, Peppard, Palta, Hla, Finn, Morgan, et al., 1997). It is successfully treated with continuous positive airway pressure (CPAP). If patients are treated at an early stage of the disease, their night-time and day-time blood pressure can be lowered, and the adverse health effects can be reduced (Dimsdale, Loredo, & Profant, 2000).

The traditional methods for the assessment of sleep-related breathing disorders are sleep studies (polysomnography), with the recording of electro-encephalography (EEG), electro-oculography (EOG), electromyography (EMG), ECG, oronasal airflow, respiratory effort, and oxygen saturation (AASM, 1999). Accurate identification of an apnoea or hypopnoea event requires direct measurement of upper airway airflow and of respiratory effort. Currently, a definitive diagnosis of sleep apnoea is made by counting the number of apnoea and hypopnoea events over a given period of time (e.g., a night's sleep). Averaging these counts on a per-hour basis leads to commonly used standards such as the apnoea/hypopnoea index (AHI) or the respiratory disturbance index (RDI) (AASM, 1999). An AHI up to 5 is regarded as normal, an AHI of 5 to 15 events per hour as mild sleep apnoea syndrome (SAS), an AHI of 15 to 30 events per hour as moderate SAS, and an AHI above 30 events per hour as severe SAS (AASM, 1999).

Sleep studies are expensive for patients, because they require overnight evaluation in sleep laboratories, with dedicated systems and attending personnel. Due to the scarcity of sleep laboratories, the vast majority of patients remain undiagnosed. Moreover, there is intra and interobserver difference in the identification of apnoea events from polysomnogram. Therefore, if SAS could be diagnosed using only ECG recordings, it would allow us to diagnose SAS simply and inexpensively from the patient's ECG recordings acquired, for example, in the patient's home.

Early in the investigation of obstructive sleep apnoea (Guilleminault, Connolly, Winkle, Melvin, & Tilkian, 1984) it was recognized that the events of apnoea and hypopnoea are accompanied by concomitant cyclic variations in heart rate (R-R intervals of ECG signals). Until now, this ordered variation in heart rate has been applied to the detection of sleep apnoea by only a few groups (Hilton, Bates, Godfrey, Chappell, & Cayton, 1999; Penzel, Amend, Meinzer, Peter, & Von Wichert, 1990; Roche, Gaspoz, Court-Fortune, Minni, Pichot, Duverney, et al., 1999). Thus, it seems possible to apply simplified ECG recording techniques combined with machine learning techniques to detect SAS because a machine learning technique has the ability to map dynamic measures like heart rate variability indexes nonlinearly onto resulting output as apnoea or no apnoea.

A number of studies during the past 15 years were accomplished for detecting SAS using features extracted from the electrocardiogram. Such

approaches are minimally intrusive, relatively inexpensive, and may be particularly well-suited for medical screening tasks. Therefore, a challenge was offered to the biomedical research community to demonstrate the efficacy of ECG-based methods for apnoea detection using a large, well-characterized, and representative set of data (Penzel, Moody, Mark, Goldberger, & Peter, 2000). That competition was jointly conducted, between February and September 2000, by Computers in Cardiology (CINC) and PhysioNet. Computers in Cardiology is an annual IEEE sponsored conference that provides publicity for the event and a venue for meetings and discussion of the competition entries (Moody, Mark, Goldberger, & Penzel, 2000; Penzel et al., 2000). PhysioNet is a Web-based library of physiological data and analytic software sponsored by the US National Institutes of Health's National Center for Research Resources (NIH NCRR) (Goldberger, Amaral, Glass, Hausdorff, Ivanov, Mark, et al., 2000; Moody, Mark, & Goldberger, 2001). PhysioNet provides free access to the database of ECG recordings. The data for that challenge was provided by Dr Thomas Penzel and is made publicly available through the PhysioNet database (Penzel et al., 2000). The dataset consists of a training and a test set, each of which contains the ECG for a full night from 35 subjects. Of the 35 subjects, 20 suffered from apnoea, 10 were normal, and 5 were borderline. Borderline subjects were excluded for classification. In the challenge event, eight participants provided fully automatic analysis (De Chazal, Heneghan, Sheridan, Reilly, Nolan, & O'Malley, 2000; Jarvis & Mitra, 2000; Maier, Bauch, & Dickhaus, 2000; Marchesi, Paoletti, & Digaetano, 2000; Mietus, Peng, Ivanov, & Goldberger, 2000; Ng, Garcia, Gomis, La Cruz, Passariello, & Mora, 2000; Schrader, Zywietz, Voneinem, Widiger, & Joseph, 2000; Shinar, Baharav, & Akselrod, 2000), and five provided a visual classification stage (Ballora, 2000; Drinnan, 2000; McNames, 2000; Raymond, 2000; Stein & Domitrovich, 2000). Thirteen algorithms were compared for the apnoea recording identification. Four algorithms (De Chazal et al., 2000; Jarvis &

Mitra, 2000; McNames & Fraser, 2000; Raymond, Cayton, Bates, & Chappell, 2000) achieved the maximum possible score for the apnoea screening (30 out of 30). Only two of them (De Chazal et al., 2000; Jarvis & Mitra, 2000) were able to utilize the automatic analysis. Correctly classified test subjects scored one point each and the borderline test subjects were not scored so the maximum score obtainable was 30.
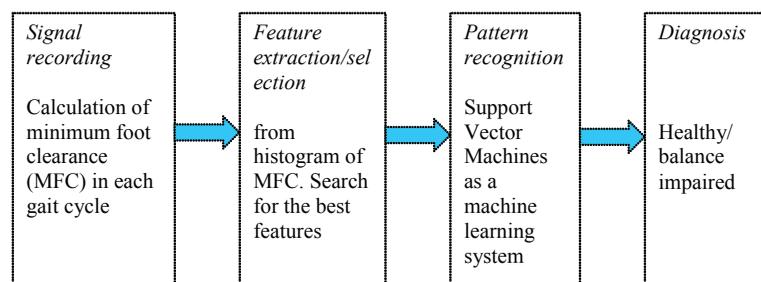
DESIGN OF AUTOMATED MEDICAL DIAGNOSTIC SYSTEM

Case1: Detection of Balance Impairments in the Elderly

We applied SVMs, for automated diagnosis of gait patterns due to balance impairments from MFC gait features, and compare their suitability as a gait classifier.

There are four main steps, as shown in Figure 2, in designing a diagnostic system for automated diagnosis of balance impairments:

1. **Signal recording:** Foot clearance (FC) data for these subjects were collected during their steady state, self-selected walking on a treadmill using a PEAK MOTUS 2D motion analysis system (Peak Technologies Inc, USA). MFC was calculated by subtracting ground reference from the minimum vertical coordinate during the swing phase through a 2D geometric model (Begg, Best, Taylor, & Dell'Oro, 2007).

2. **Feature extraction:** Each subject's MFC data was plotted as histograms showing individual MFC data and their respective frequencies (Begg, Palaniswami, & Owen, 2005a). Features included mean, median, min, max, standard deviation (SD), skewness, and so forth. In young/old gait pattern recognition, the geometry of Poincaré plot of MFC data was used to extract the reliable features

Figure 2. Schematic diagram of SVM-based gait diagnostic model

| *Signal recording* Calculation of minimum foot clearance (MFC) in each gait cycle	*Feature extraction/selection* from histogram of MFC. Search for the best features	*Pattern recognition* Support Vector Machines as a machine learning system	*Diagnosis* Healthy/ balance impaired

(Begg, Palaniswami, & Owen 2005a; Begg, Lai, Taylor, & Palaniswami, 2005b). A feature selection can be subsequently included in this process, by which feature vector is reduced in dimension, which includes only the most relevant features necessary for discrimination and sometimes assisted by a priori knowledge or rules. In this regard, a hill-climbing feature selection algorithm was found to be useful to identify features that provide the most contribution in separating the two gait classes (Begg, Palaniswami, & Owen, 2005a).

3. **Pattern recognition:** SVMs as a machine learning tool were used to realize an input-output mapping of the exact relationship between MFC features and falls/no-falls category. For classification, SVMs operate by finding a hypersurface in the space of possible inputs. As shown in Figure 3, this hypersurface attempts to split the positive examples from the negative examples. The split is chosen to have the largest distance from the hypersurface to the nearest of the positive and negative examples. The most important characteristic of SVMs is to condense information from a large training set into a very small number of points (i.e., the support vectors). These support vectors contain all the necessary information to define the decision hypersurface. The detailed theory behind SVM has appeared in many textbooks and journal articles (e.g., Kecman, 2002;

Vapnik, 1995). Performance of the classifier can be evaluated using accuracy rates and measures of receiver operating characteristics (ROC) curves (Chan, Lee, Sample, Goldbaum, Weinreb, & Sejnowski, 2002). The area under the ROC curve provides a measure of overall performance of the classifier, that is, the larger the ROC area the better is the classification accuracy over a range of thresholds. Many studies (e.g., Begg, Palaniswami, & Owen, 2005a; Chan et al., 2002) have used a ROC plot and its area as an index for evaluating classifier performance. Cross validation procedure is used to determine the region of optimal SVM parameters (Begg, Palaniswami, & Owen, 2005a). First a subset is used to train the SVM model while the remaining data examples are used for testing. The process is repeated for the other subsets so that in the end each example had been tested. Results of each test are combined to obtain an average result for the measure of accuracy.

4. **Diagnosis:** After the classifier is trained to set up an optimized model structure during the pattern recognition step, this step detects the balance-impaired subjects from the test samples to be detected. Our SVM model trained using a Gaussian kernel gave a high accuracy of 95% in classifying healthy and balance-impaired elderly subjects (Begg, Lai, Taylor, & Palaniswami, 2005b).

Figure 3. 2D-SVM decision surface for Polynomial kernel (d=4) plotted for Feature 1 (first-quartile) against Feature 9 (kurtosis) (Begg, 2005b)
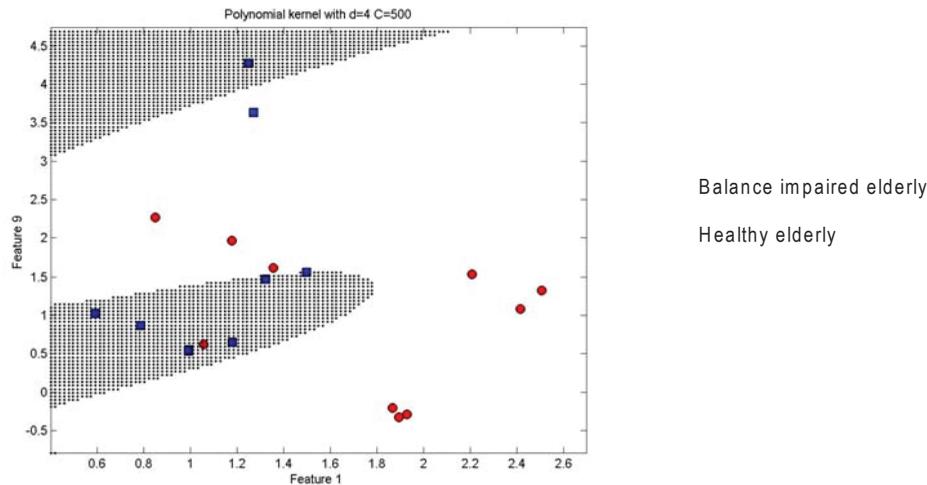
The following paragraph describes steps for designing a diagnostic system for sleep apnoea prediction.

Case 2: Automated Prediction of Sleep Apnoea Syndrome

There are five main steps as shown in Figure 4 for designing a diagnostic system for predicting sleep apnoea syndrome from ECG signals.

1. **Signal recording:** In which patient's ECG signals are recorded and filtered. A commercially available ECG recording system is always used.
2. **Signal processing:** This step processes the acquired signals to produce a relevant description of the signals. QRS complex of ECG signal detection times are calculated and the length between each QRS (R to R peak interval time) are then calculated (Engelse & Zeelenberg, 1979). Heart rate is the reciprocal of RR interval. Due to poor signal quality and errors in the automatically generated QRS detections, the RR-interval sequences generated from both sets of QRS detection

times contained physiologically unreasonable time (De Chazal, Heneghan, Sheridan, Reilly, Nolan, & O'Malley, 2003).

3. **Feature extraction:** Heart rate variability indexes (Task force, 1996) are used as features. Morphology of ECG can also be used. Among all participants in the Apena challenge, six participants used spectral analysis of heart rate variability (De Chazal et al., 2000; Drinnan, Allen, Langley, & Murray, 2000; Jarvis & Mitra, 2000; McNames & Fraser, 2000; Schrader et al., 2000; Shinar et al., 2000). Two studies used the Hilbert transform to extract frequency information from the heart rate signal (Mietus et al., 2000; Schrader et al., 2000). Three algorithms used time-frequency maps for the presentation of the heart rate variability (Jarvis & Mitra, 2000; McNames & Fraser, 2000; Schrader et al., 2000). One of the participants used a threshold for the ratio of the spectral power of the heart rate in two fixed frequency bands (0.01–0.05 cycles per beat and 0.005–0.010 cycles per beat) (Drinnan et al., 2000). Another participant combined spectral analysis, Hilbert transform frequencies, and discrete wavelet analysis to use more

parameters for a subsequent feature selection (Schrader et al., 2000). Several studies used different ECG-derived parameters in addition to heart rate variability, that is, ECG pulse energy (McNames & Fraser, 2000), R-wave duration (Shinar et al., 2000), and amplitude of the S component of each QRS complex (McNames & Fraser, 2000), and two used the ECG-derived respiration (EDR) technique (Moody, Mark, Zoccola, & Mantero, 1985) to measure the amplitude modulation of the ECG signal to estimate respiratory activity. These were based on spectral analysis of the R-wave amplitude using power spectral density (PSD) (De Chazal et al., 2000) and of the T-wave amplitude using the discrete harmonic wavelet transform (Raymond et al., 2000). The algorithms that performed best used frequency-domain parameters of heart rate variability or the ECG-derived respiration signal with R-wave morphology (De Chazal et al., 2000; McNames & Fraser, 2000; Raymond et al., 2000; Shinar et al., 2000).

4. **Pattern recognition:** This step involves the development of intelligent machine learning systems dealing with the extracted and selected information in the previous step. Several machine learning techniques have been used to r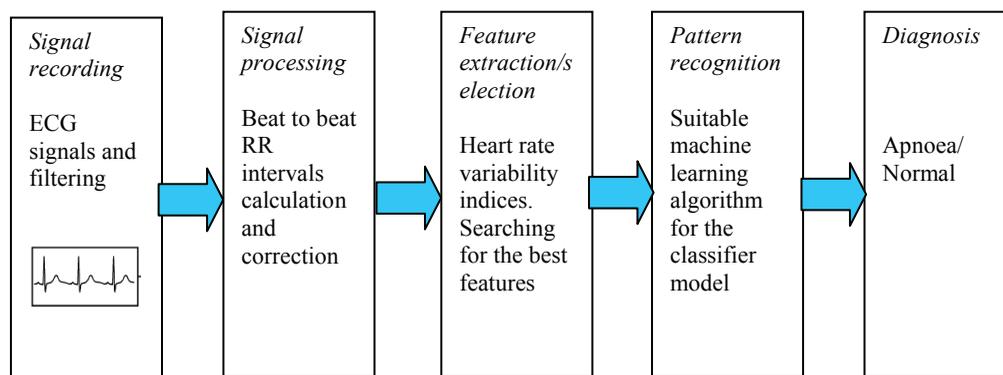ecognize sleep apnoea syndrome based on the selected heart rate variability **(**HRV) indexes. De Chazal et al. (2003) used linear and quadratic discrimant models. Roche, Pichot, Sforza, et al. (2003) used the classification and regression tree (CART) method. Gracia, Gomis, La Cruz, Passeriello, and Mora (2000) used the Bayesian hierarchical model. We used SVM. All SVM architectures were trained and tested on the D2CSVM software (Lai, Palaniswami, & Mani, 2003, 2005).

5. **Diagnosis:** It consists of detecting potential disorders and characterizing the case by means of a pattern recognition step. This step exploits the classification provided by the pattern recognition step. Our SVM trained model showed 100% accuracy over apnoea/healthy classification of 30 ECG data sets from Physionet (see Figure 5). The novelty we introduced is the estimation of relative degree of apnoea in all subjects. As a result, the borderline group can easily be isolated (Khandoker, Lai, Palaniswami, & Begg, in press).

Some Remarks on Selecting Machine Learning Techniques for Medical Diagnostics

Machine learning techniques can be viewed as an attempt to automate parts of the diagnostic system.

Figure 4. Schematic representation of a diagnostic system for predicting sleep apnoea syndrome based on ECG signals

| *Signal recording*	*Signal processing*	*Feature extraction/selection*	*Pattern recognition*	*Diagnosis*
ECG signals and filtering	Beat to beat RR intervals calculation and correction	Heart rate variability indices. Searching for the best features	Suitable machine learning algorithm for the classifier model	Apnoea/ Normal

The goal of machine learning in designing a medical diagnostic system is that it can be successful in diagnosing a disease or a medical condition. We should design a diagnostic system to minimize the misdiagnosis. While much attention is often paid to the classification method, the evaluation of the performance of the system deserves equal importance. The most important question for the designer of a learning system is: Which of the well-known methods will work best on a real problem? For most problems, where the characteristics of the data are not well-known, it is impossible to answer such a question without actually trying out a variety of methods on the available samples and empirically comparing the results. There are two criteria in choosing a right classifier. First, the accuracy of diagnosis will always remain the primary criterion in choosing a classifier. Independent of any particular learning method, general principles for applying a learning method are to yield the best performance in terms of the best or smallest predictive error rate. Second, the speed of computation, which can

have a strong influence on the choice of learning method, should be considered.

The problems of the true error rate estimation arise with smaller sample sizes. Typical real-world samples run in the hundreds, not the thousands. In general, 10-fold cross-validation is adequate and sufficient for obtaining reliable error rate estimators. When the sample size is small (less than 100) and particularly when the sample is less than 50, leave-one-out or possibly bootstrapping and 2-fold cross-validation should be used (Weiss & Kulikowski, 1991). For medical diagnosis, distinctions among different types of errors turn out to be important. For example, the error committed in tentatively diagnosing someone as healthy when one has life-threatening illness (known as a false negative decision) is usually considered far more serious than the opposite type of error, that is, of diagnosing someone as ill when one is in fact healthy (known as a false positive).

Figure 5. Posterior class (Apnic) probability estimate $P(Apnic|SVM_{output})$, for 20 subjects with apnoea (1-20), 5 subjects with borderline apnoea (21-25) and 10 subjects with no apnoea (26-35), calculated from SVM output values for each class (Platt, 2000, pp. 1-10). ECG data were collected from Computers in Cardiology Challenge 2000. (http://www.physionet.org/physiobank/database/apnea-ecg/)

Posterior class probabilities of SVM outputs
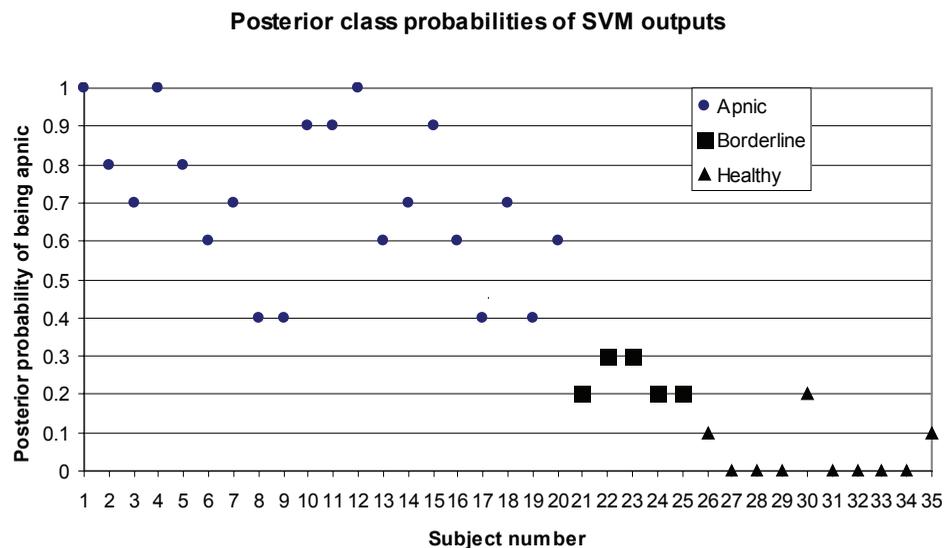
Some Remarks on Extracting and Selecting the Best Features for Machine Learning Technique in Medical Diagnostics

We hope to learn from samples that good predictions can be made for new cases. To a large extent, we are at the mercy of the features. These are the fundamental measurements or tests that we are given, and upon which we have to make decisions. In attempting to learn from the data, if all methods do badly, we may eventually just have to conclude that better features are needed to make reliable predictions. However, in some instances, it may be possible to combine and transform (mathematically and logically) the original features so that we obtain new features which enhance the maximum discriminatory information from the available data. Wavelet analysis as a sophisticated signal processing and feature enhancement technique was applied in clinical diagnosis (Ivanov, Rosenblum, Peng, Mietus, Havlin, Eugene, et al., 1996). The integration of a number of additional approaches has also been proposed for the further improvement in the reliability of the diagnostic performance. While interscale dependencies in the signal characteristics of healthy subjects has been observed (Ivanov et al., 1996), the identification of a similar multiscale relationship in the HRV data for SAS patients can be achieved by multiscale signal enhancement applied in the wavelet domain (Sendur & Selesnick, 2002). Wavelet-based signal enhancement can be customized for the HRV data in order to exploit relevant signal features, while seeking to incorporate the shift invariant properties of a family of complex wavelet transforms (Kingsbury, 2001). A shift invariant signal decomposition approach will greatly aid in the identification of the time and frequency dependent signal features upon which the diagnosis is ultimately based.

If the features have all good predictive capabilities, any classification method should do well. Otherwise, the situation is much less predictable. In practice, many features in an application are often poor, noisy, and redundant. Adding new information, in the form of weak features, can actually degrade performance (Begg, Lai, Taylor, & Palaniswami, 2005b; Begg, Palaniswami, & Owen, 2005a). That is why the primary approach to minimize the effects of feature noise and redundancy is feature selection. There are several techniques for feature selection available in the literature (e.g., Aha & Bankert, 1996; Begg, Palaniswami, & Owen, 2005a; Chang, 1990).

FUTURE TRENDS

- **Case 1:** For the diagnosis of pathological gait, like balance impairments, processing, and extraction of gait features that correlate with that particular pathology can be thought of as an important step in designing a diagnostic model of such category. Investigation into combining MFC data with other types of gait features (e.g., stride-to-stride time and distance, foot-ground reaction forces, joint/muscle moments, and electrical activity of lower limb muscular contractions) is particularly important (Begg & Palaniswami, 2006a, pp. 259-260). Furthermore, other machine learning techniques (e.g., neural networks, fuzzy logic, Bayesian approach, or genetic algorithms) can be tried to work in cascade with the SVM technique in a hybrid design mode to further improve the performance of the classification. A further benefit that could be obtained from the diagnostic system is through the categorization of the classification outcome results in various falls risk scales (e.g., high-risk, medium-risk, low/no-risk).

- **Case 2:** The posterior class probability outputs of a classifier can be calculated to recognize the borderline subjects by estimating a relative degree of sleep apnoea on individual subjects. The estimated values can easily be validated by investigating the effect of CPAP therapy on probability outputs of a classifier. A sleep

apnoea syndrome patient undergoing CPAP therapy was reported to improve (Belozeroff, Berry, Sassoon, & Khoo, 2002). Surrogate parameters such as the HRV indexes can never replace the target variables (nasal airflow with oxygen saturation) that are derived from the direct recording of respiration. As a consequence, the recognition of sleep apnoea syndrome based on HRV indexes or parameters should always be named differently to make the derived nature clear. Penzel et al. (2002) suggests that when disordered breathing in sleep is determined based on surrogate parameter such as heart rate and ECG, it would be more appropriate to make the derived nature clear. Therefore, classification labels with recommendations have the potential to make the diagnostic system more informative (Table 1).

CONCLUSION

In this chapter, we described the step-by-step methodology of designing the automated medical diagnostic system for two pathological conditions: (1) detection of balance impairments in the elderly based on gait signals and (2) prediction of sleep apnoea syndrome based on ECG signals. General illustration of applying machine learning approach in the perspective of diagnostic system design may help readers understand how a classifier model can be used to make decision. Our emphasis is on practical methods that have proven successful in building a machine learning system providing useful information to the end user. The design strategies in Case 1 can help build a SVM gait recognition model in the detection of gait changes in older adults due to balance impairments and falling behaviour. The proposed improvements in machine learning techniques and feature extraction could be encouraging not only in the falls risk diagnostic applications but also for evaluating the need for referral for fall prevention/intervention programs (e.g., exercise prescriptions to improve balance and rehabilitation).

The design strategies (described in Case 2) for predicting sleep apnoea syndrome (SAS) from ECG recording may provide essential information for introducing a novel screening device that can aid sleep specialist or other physicians in the initial assessment of patients with suspected SAS and estimate the relative risk of a sleep related breathing disorder, thereby indicating the need for referral for overnight sleep studies (i.e., polysomnogram [PSG] recording). This in turn may help prioritize patients, so that those in greatest need of treatment will undergo full PSG recordings in a timely man-

Table 1. Proposed apnoea classification labels based on HRV

Class label	Recommendation
Healthy	*Not significant SAS_{HRV}*
Healthy (borderline)	*Very slight symptom of SAS_{HRV}. Sleep physician consultation recommended*
Apnic (borderline)	Potential symptom of SAS_{HRV}. Overnight sleep studies recommended
Apnic	*Typical SAS_{HRV}. Treatment by CPAP recommended (otherwise wait for possible heart failure in a couple of years!!)*

ner, while those without apnoea will be able to avoid this tedious procedure.

REFERENCES

Aha, D.W., & Bankert, R. L. (1996). A comparative evaluation of sequential feature selection algorithms. In D. Fisher & J.-H. Lenz (Eds.), *Artificial intelligence and statistics V.* New York: Springer-Verlag.

American Academy of Sleep Medicine (AASM) Task Force. (1999). Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. *Sleep*, 22, 667-689.

Ballora, M., Pennycook, B., Ivanov, P. C., Goldberger, A., & Glass, L. (2000). Detection of obstructive sleep apnea through auditory display of heart rate variability. *Comput. Cardiol.*, *27*, 739-740.

Barton, J. G., & Lees, A. (1997). An application of neural networks for distinguishing gait patterns on the basis of hip-knee joint angle diagrams. *Gait and Posture*, *5*, 28-33.

Begg R.K., Best R.J., Taylor S., & Dell'Oro L. (2007). Minimum foot clearance during walking: Strategies for the minimization of trip-related falls. *Gait and Posture, 25*(2), 191-8.

Begg R. K., & Palaniswami M. (Eds.) (2006a). *Computational intelligence for movement sciences: Neural networks and other emerging techniques.* IGI Publishing. ISBN: 1-59140-836-9.

Begg R. K., Kamruzzaman J., & Sarker R. (Eds.) (2006b). *Neural networks in healthcare: Potentials and challenges.* IGI Publishing. ISBN: 1-59140-848-2.

Begg R. K., Palaniswami M., & Owen B. (2005a). Support vector machines for automated gait classification. *IEEE Trans Biomed Eng*, *52*, 828-838.

Begg R.K., Lai D., Taylor S., & Palaniswami M. (2005b). SVM-based models in the assessment of balance impairments. *The Third International Conference on Intelligent Sensing and Information Processing* (pp. 248-253). Banglore, India. IEEE Press.

Begg R.K., Sparrow W.A., Lythgo N.D. (1998). Time-domain analysis of foot-ground reaction forces in negotiating obstacles. Gait Posture ;7:99-109.

Belozeroff, V., Berry, R. B., Sassoon, C. S. H, & Khoo, M. C. K. (2002). Effects of CPAP therapy on cardiovascular variability in obstructive sleep apnea: A closed-loop analysis. *AJP – Heart, 282*, 110-121.

Chan, K., Lee, T. W., Sample, P. A., Goldbaum, M. H., Weinreb, R. N., & Sejnowski, T. J. (2002). Comparison of machine learning and traditional classifiers in glaucoma diagnosis. *IEEE Trans Biomed Eng*, *49*, 963-74.

Chang, E. I. (1990). *Using genetic algorithms to select and create features for pattern classification.* Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, MA.

Chapelle, O, Haffner, P., & Vapnik, V. N. (1999). Support vector machines for histogram-based classification. *IEEE Trans. Neural Net., 10*(5),1055-1064.

Chau, T. (2001). A review of analytical techniques for gait data. Part 2: Neural network and wavelet methods. *Gait and Posture*, *13*, 102-120.

De Chazal, P., Heneghan, C., Sheridan, E., Reilly, R., Nolan, P., & O'Malley, M. (2000). Automatic classification of sleep apnea epochs using the electrocardiogram. *Comput. Cardiol.*, *27*, 745-748.

De Chazal, P., Heneghan, C., Sheridan, E., Reilly, R., Nolan, P., & O'Malley, M. (2003). Automated processing of the single lead electrocardiogram for the detection of obstructive sleep apnoea. *IEEE Trans. of Biomedical Engineering*, *50*(6) 686 -696.

Dimsdale, J. E., Loredo, J. S., & Profant, J. (2000). Effect of continuous airway pressure on blood pressure. *Hypertension*, *35,* 144-147.

Drinnan, M. J., Allen, J., Langley, P., & Murray, A. (2000). Detection of sleep apnoea from frequency analysis of heart rate variability. *Comput. Cardiol.*, *27,* 259-262.

Engelse, W. A. H., & Zeelenberg, C. (1979). A single scan algorithm for QRS detection and feature extraction. *Computers in Cardiology, 6,* 37-42.

Fildes, B. (1994). *Injuries among older people: Falls at home and pedestrian accidents*. Melbourne, FL: Dove Publications.

Garcia, N. F. I., Gomis, P., La Cruz, A., Passeriello, G., & Mora, F. (2000). Bayesian hierarchical model with wavelet transform coefficients of the ECG in obstructive sleep apnea screening. *Comput. Cardiol.*, *27,* 275-278.

Goldberger, A. L., Amaral, A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al. (2000). Physiobank, physiotoolkit, and physionet. *Circulation*, *101,* 215-220.

Guillemiault, C., Connolly, S. J., Winkle, R., Melvin, K., & Tilkian, A. (1984). Cyclical variation of the heart rate in sleep apnoea syndrome. Mechanisms and usefulness of 24 h electrocardiography as a screening technique. *The Lancet, I,* 126-131.

Hilton, M. F., Bates, R. A., Godfrey, K. R., Chappell, M. J., & Cayton, R. M. (1999). Evaluation of frequency and time–frequency spectral analysis of heart rate variability as a diagnostic marker or the sleep apnoea syndrome. *Med. Biol. Eng. Comput.*, *37,* 760-769.

Holzreiter, S. H., & Kohle, M. E. (1993). Assessment of gait pattern using neural networks. *Journal of Biomechanics*, *26*, 645-651.

Ifeachor, E. C. Sperduti, A., & Starita, A. (1998). *Neural networks and expert system in medicine and health care*. Singapore: World Scientific Publishing.

Ivanov, P. C., Rosenblum, M. G., Peng, C.-K., Mietus, J., Havlin, S. Eugene, H., et al. (1996). Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis. *Nature, 383,* 323-32.

Jarvis, M. R., & Mitra, P. P. (2000). Apnea patients characterized by 0.02 Hz peak in the multitaper spectrogram of electrocardiogram signals. *Comput. Cardiol.*, *27,* 769-772.

Karst, M. G., Hageman, A. P., Jones, F. T., & Bunner, S. H. (1999). Reliability of foot trajectory measures within and between testing sessions. *J. Gerontol.: Med. Sci.*, *54,* 343-347.

Kecman, V. (2002). *Learning and soft computing: Support vector machines, neural networks and fuzzy logic models*. Cambridge, MA: MIT Press.

Khandoker, A. H., Karmakar, C. K., & Palaniswami, M. (2007, October 1-3). *Screening obstructive sleep apnoea syndrome from electrocardiogram recordings using support vector machines*. Paper presented at the 34th Annual Conference on Computers in Cardiology, Durham, USA.

Khandoker, A. H., Lai, DTH, Palaniswami, M., & Begg, R. K. (2007). Wavelet-based feature extraction for support vector machines for screening balance impairments in the elderly. *IEEE transaction of Neural and Rehabilitation Engg. 15*(4), 587-597.

Kingsbury, N. G. (2001). Complex wavelets for shift invariant analysis and filtering of signals. *Applied and Computational Harmonic Analysis*, *10*(3), 234-253.

Lai, D., Palaniswami, M., & Mani, N. (2003). *A new method to select working sets for decomposition methods solving support vector machines* (Tech. Rep. MECE-30-2003). Monash University, Australia.

Lai, D., Palaniswami, M., & Mani, N. (2005). *A basic heuristic decomposition framework for training Support Vector Machines* (Tech. Rep. MECSE-26-2005)*.* Monash University, Australia.

Maier, C., Bauch, M., & Dickhaus, H. (2000). Recognition and quantification of sleep apnea by

analysis of heart rate variability parameters. *Comput. Cardiol.*, *27,* 741-744.

Marchesi, C., Paoletti, M., & Digaetano, S. (2000). Global waveform delineation for RR series estimation: Detecting the sleep apnea pattern. *Comput. Cardiol., 27,* 71-74.

Mcnames, J. N., & Fraser, A. M. (2000). Obstructive sleep apnea classification based on spectrogram patterns in the electrocardiogram. *Comput. Cardiol.*, *27,* 749-752.

Mietus, J. E., Peng, C. K., Ivanov, P. C., & Goldberger, A. L. (2000). Detection of obstructive sleep apnea from cardiac interbeat interval time series. *Comput. Cardiol.*, *27,* 753-756.

Moody, G. B., Mark, R. G., & Goldberger, A. L. (2001). PhysioNet: A Web-based resource for the study of physiologic signals. *IEEE Eng. Med. Biol.*, *20,* 70-75.

Moody, G. B., Mark, R. G., Goldberger, A. L., & Penzel, T. (2000). Stimulating rapid research advances via focused competition: The computers in cardiology challenge 2000. *Comput. Cardiol.*, *27,* 207-210.

Moody, G. B., Mark, R. G., Zoccola, A., & Mantero, S. (1985). Derivation of respiratory signals from multi-lead ECGs. *Comput. Cardiol.*, *12,* 113-116.

Ng, F., Garcia, I., Gomis, P., La Cruz, A., Passariello, G., & Mora, F. (2000). Bayesian hierarchical model with wavelet transform coefficients of the ECG in obstructive sleep apnea screening. *Comput. Cardiol.*, *27,* 275-278.

Nieto, F. J., Young, T. B., Lind, B. K., Shahar, E., Samet, J. M., Redline, S., et al. (2000). Association of sleep disordered breathing, sleep apnea, and hypertension in a large community-based study. *J. Am. Med. Assoc., 283,* 1829-1836.

Ostrosky, K. M., VanSwearingen, J. M., Burdett, R. G., & Gee, Z. (1994). A comparison of gait characteristics in young and old subjects. *Phys. Ther., 74,* 637-646.

Penzel, T., Amend, G., Meinzer, K., Peter, J. H., & Von Wichert, P. (1990). Mesam: A heart rate and snoring recorder for detection of obstructive sleep apnea. *Sleep, 13,* 175-182.

Penzel, T., Moody, G. B., Mark, R. G., Goldberger, A. L., & Peter, J. H. (2000). The Apnea-ECG database. *Comput. Cardiol.*, *27,* 255-258.

Platt, J. C. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In A.J. Smola, P. Bartlett, B. Schölkopf, & D. Schuurmans (Eds.), *Advances in Large Marg in Classifiers* (pp. 61-74). Cambridge, MA: MIT Press.

Raymond, B., Cayton, R. M., Bates, R. A., & Chappell, M. J. (2000). Screening for obstructive sleep apnoea based on the electrocardiogram. Computers in cardiology challenge. *Comput. Cardiol.*, *27,* 267-270.

Roche, F., Gaspoz, J. M., Court-Fortune, I., Minni, P., Pichot, V., Duverney, D., et al. (1999). Screening of obstructive sleep apnea syndrome by heart rate variability analysis. *Circulation, 100,* 1411-1415.

Roche, F., Pichot, V., Sforza, E., et al. (2003). Predicting sleep apnoea syndrome from heart period: A time-frequency wavelet analysis. *Eur Respir J, 22,* 937-942.

Schollohorn, W. I. (2004). Applications of arificial neural nets in clinical biomechanics. *Clinical Biomechanics, 19,* 876-98.

Schrader, M., Zywietz, C., Voneinem, V., Widiger, B., & Joseph, G. (2000). Detection of sleep apnea in single channel ECGs from the PhysioNet data base. *Comput. Cardiol.*, *27,* 263-266.

Sendur, L., & Selesnick, I. W. (2001). Bivariate shrinkage with local variance estimation. *IEEE Signal Processing Letters*, *9*(12), 438-441.

Shinar, Z., Baharav, A., & Akselrod, S. (2000). Obstructive sleep apnea detection based on electrocardiogram analysis. *Comput. Cardiol., 27,* 757-760.

Stein, P. K., & Domitrovich, P. P. (2000). Detecting OSAHS from patterns seen on heart-rate tachograms. *Comput. Cardiol.*, *27,* 271-274.

Task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. *Circulation*, *93,* 1043-1065.

Teodorrescu, T., Kandel, A., & Jain L.C (1998). *Fuzzy and neuro-fuzzy systems in medicine.* Boca Raton, FL: CRC Press.

Vapnik, V. N. (1995). *The nature of statistical learning theory.* New York: Springer.

Weiss, S. M., & Kulikowski, C. A. (1991). *Computer systems that learn*. San Mateo, CA: Morgan Kaufmann Publishers, Inc.

Winter, D. A. (1991). *The biomechanics and motor control of human gait: Normal, elderly and pathological*. Waterloo, Canada: University of Waterloo Press.

Wu, W. L., Su, F. C., & Chou, C. K. (1998). Potential of the back propagation neural networks in the assessment of gait pattern in ankle arthrodesis. In E. C. Ifeachor, A. Sperduit, & A. Starita (Eds.), *Neural networks and expert systems in medicine and health care* (pp. 92-100). Singapore: World Scientific Publishing.

Young, T., Palta, M., Dempsey, J., Skatrud, J., Weber, S., & Badr, S. (1993). The occurence of sleep-disordered breathing among middle-aged adults. *New Engl. J. Med.*, *328,* 1230-1235.

Young, T., Peppard, P., Palta, M., Hla, K. M., Finn, L., Morgan, B., et al. (1997). Population-based study of sleep-disordered breathing as a risk factor for hypertension. *Arch. Intern. Med.*, *157,* 1746-1752.

KEY TERMS

Beat to Beat Heart Rate: It is calculated from the time interval between R to R peaks of ECG signals. Fluctuations in beat to beat heart rate indicate the activities of the autonomous nervous system or cardiovascular system.

Electrocardiogram (ECG): The ECG signal is a representation of the bioelectrical activity of the heart's pumping action. This signal is recorded via electrodes placed on the patient's chest. The physician routinely uses ECG time history plots and the associated characteristic features of P, QRS, and T waveforms to study and diagnose the heart's overall function.

Gait: Walking or running pattern. Gait analysis is routinely used to diagnose musculoskeletal problems in the lower limb and also for evaluation of treatment outcomes.

Gait Diagnostics: A process that uses objective methods to describe a patient's walking function with a view to understand the underlying pathomechanisms and to inform the clinical decision-making process which aims to improve the patient's gait through various form of interventions.

Heart Rate Variability (HRV): HRV has become the conventionally accepted term to describe variations of both beat-to-beat heart rate and RR intervals

Minimum Foot Clearance: It is defined as the minimum vertical distance between the lowest point under the front part of the shoe/foot and the ground during the mid-swing phase of the gait cycle. It has been identified as a potential gait parameter associated with especially trip-related falls in the older population.

Sleep Apnoea: It is defined as the cessation of breathing for short periods during sleep. It is successfully treated with home ventilation using nasal continuous positive airway pressure (CPAP).

Support Vector Machines (SVM): A new generation supervisor learning system based on recent advances in statistical learning theory. SVM maps input vectors to a higher dimensional space where a maximal separating hyperplane is constructed.

Wavelet Transform: Signal transformation that transforms the time signal to the frequency domain. The wavelet transform is based on a decomposition of a signal using an orthogonal family of basis functions derived from the so-called mother wavelet function.