A COMPUTER-ASSISTED TECHNIQUE FOR NERVE CONDUCTION STUDY IN EARLY DETECTION OF PERIPHERAL NEUROPATHY USING ANN

Mausumi Barthakur¹, Anil Hazarika²* and Manabendra Bhuyan²

¹Department of Neurophysiology, GNRC hospital, Six miles, Guwahati, Assam, India.
²Department of Electronics and Communication Engineering, Tezpur University, Napaam Sonitpur-784028, Assam, India.

ABSTRACT

This research is aimed at prospective use of Artificial Neural Network (ANN) for classification and diagnostic evaluation of neuropathy. The limitation over the traditional clinical study of the neuropathy necessitates a new statistical model that can be used as systematic practical tool for detection and classification of the neuropathy.

Nerve Conduction Study (NCS) is a detection protocol of the neurophysiologists for early evaluation of peripheral neuropathy (PN). The limitations over the traditional clinical study of the PN using NCS necessitate a new computer assisted model that should provide a reliable objective detection and classification of this class of neuropathy. This research is aimed at a prospective use of Artificial Neural Network (ANN) for classification and diagnostic evaluation of PN.

Electrophysiological results of nerve conduction studies (NCS) conducted from 2008 to 2009 on 420 suspected patients were evaluated and 5 NCS variables or features were determined for each patient in this study. The pre-processed feature variables were subjected to three different ANN models. The results were analyzed on different classification parameters such as accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). From the sensitivity analysis a new ‘score’ metric has been developed to re-train the ANN models to improve the detection process.

It was found that the best performing model is feed-forward back propagation (FFBP) which has higher sensitivity, specificity and accuracy than other two ANN models and also has higher classification speed. The receiver operating characteristics (ROC) curves were plotted for both training and testing set using FFBP model and the area under curves (AUC) were calculated. Further, ANN sensitivity analysis for NCS variables or features was performed to judge the most significant features and ranked them according to their sensitivity.

In a second approach, the sensitivity of each feature variables were used to determine a score metric for each patient, thereby reducing the feature for improvement of classification accuracy.
Classification accuracy has been improved from 94% to 96% by using score based feature instead of five NCS variables. It has been concluded that ANN analysis offers a promising implementation to established methods of statistical analysis of multivariable data on neuropathy patients for detection and classification.

**Keywords:** Artificial Neural Network, Nerve conduction study, Peripheral Neuropathy, Computer Assisted medical diagnosis

**1. INTRODUCTION**

Peripheral neuropathy (PN) is a common problem of damage to the peripheral nervous system. It is important to diagnose and evaluate PN as quickly as possible to reduce the risk of permanent nerve damage. A large number of patients register for primary detection of peripheral neuropathy due to complain of pain, tingling, burning sensation, pricking sensation, weakness of limbs, difficulties in walking, and imbalance of gait etc. However, in a large percentage of such cases, neuropathy is finally ruled out. Nerve conduction studies (NCS) can assist the neurophysiologists with the evaluation of PN [1-3] while needle electromyology (EMG) is a secondary investigation after NCS. NCS allow clinicians such as primary-care physicians to make timely and objective decisions about patients with neuromuscular symptoms. Deployment of NCS in the primary-care setting helps patient management by reducing the inconvenience, expense, and possible treatment delays incurred by patients who would otherwise have to obtain alternative treatment at a later date [4]. This is the reason why an advanced functionality of NCS is essential that can aid specialists with an accurate and reliable complement to their traditional manual analyses of NCS waveforms [4].

Peripheral neuropathy has numerous causes including hereditary, toxic, metabolic, infections, inflammatory, ischemic, Para-neoplastic disorder etc. The term PN is usually used to describe symmetric and universal damage to adjacent nerves. The damage and clinical manifestations are usually located distally with a proximal progression. Several disorders can damage peripheral nerves and cause PN, hence it is important to differentiate actual neuropathy from other disorders that can have a similar clinical presentation.

The peripheral nerves consist of bundles of long neuronal axons as they exit the central nervous system (CNS). Some peripheral nerves are wrapped in a myelin sheath generated by Schwann cells, whereas others are unmyelinated. Peripheral nerves serve different motor, sensory, and autonomic functions. Peripheral nerves include the cranial, spinal nerve roots, dorsal root ganglia, the peripheral nerve trunks with their trained branches and the peripheral autonomic nervous system. Neuropathies can be categorized accordingly to the fiber type that is primarily involved. Most toxic and metabolic neuropathies are initially sensory and later may involve the motor fibers. Pure sensory neuropathies can result from drug toxicity, paraneoplastic syndrome and nutritional deficiency. Primarily motor neuropathies include GullainBarre syndrome. Alcoholism and diabetes can both cause small fibre painful neuropathies. Autonomic involvement occurs in many small fiber neuropathies but can also occur in GuillainBarresyndrome and is sometimes life threatening.

Nerve conduction study (NCS) gives information on functioning of peripheral nervous system which may be used for diagnosis; description of disease state; longitudinal monitoring of disease with multiple studies and advice on prognosis management.

NCS may be diagnostically helpful in patients suspected of having almost any disorder of peripheral nervous system (PNS), including nerve roots, peripheral nerves, muscle and neuromuscular junction. NCS involve the application of a depolarizing square wave of electrical pulses to the skin over a peripheral nerve producing propagated sensory nerve action potential
(SNAP) recorded at a distal point over the same nerve, and a compound muscle action potential (CMAP) arising from the activation of muscle fibre in a target muscle supplied by the nerve. The different features or variables considered in NCS are explained with the waveform diagram of Fig 1 (a).

![Waveform Diagram](image)

**Fig 1** NCV signals (a) Typical (b) Screen shot of NCV machine for demylinating *median nerve*

The Amplitude (CMAP) is measured from the baseline to peak (CD), expressed in millivolt (mV) in case of motor conduction study. CMAP amplitudes are indicative of the efficiency of neuromuscular transmission and the number of muscle fibers composing the recorded muscle that can generate action potentials. Latency is a time measurement expressed in milliseconds (ms). The distal motor latency (DML) is the time interval between the moment of nerve stimulation point (A) and the onset of the resulting CMAP represented by AB in Fig 1(a). The latency obtained on distal stimulation is one of the reported components of the nerve conduction study; whereas the latency obtained on proximal stimulation (proximal motor latency (PML)) is used to calculate a conduction velocity along the nerve segment between the two stimulation points. Axonal loss leads to lower CMAP amplitudes, and demyelination causes prolonged latency and slow conduction velocity. The motor latencies reflect time required for conduction of impulses along motor nerves, neuromuscular transmission, and initiation of muscle action potentials. Fig.1 (b) shows screen shot of NCV signal for a demyelinating *median nerve*. The five vital NCS variables used in our work are shown in Table 1.

<table>
<thead>
<tr>
<th>NCS Variables considered in this study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Features or variables</strong></td>
</tr>
<tr>
<td>Distal motor latency (DML)</td>
</tr>
<tr>
<td>Proximal motor latency(PML)</td>
</tr>
<tr>
<td>Compound muscle action potential (CMAP)</td>
</tr>
<tr>
<td>Motor nerve conduction velocity (MNCV)</td>
</tr>
<tr>
<td>Minimal F-response (MFR)</td>
</tr>
</tbody>
</table>

49
The MNCV (m/s) is calculated using Equation (1) with reference to Fig 1(b) given as-

\[ MNCV (m/s) = \frac{l}{(GD - AB)} \]

Where \( l \) is the proximal to distal length in metre, \( AB \) is the distal latency and \( GD \) is the proximal latency in seconds. Minimal F-response (MFR) is the time taken by an impulse to traverse through the peripheral motor nerves and roots. It is expressed in millisecond (ms). Type of recording electrode during nerve conduction study is important. Different types of electrodes like, needle electrode, surface electrode are used. Convenience and non-invasiveness is the two reasons, why the clinical electro-physiologists prefer to use the surface recording electrodes.

Fig 2 shows a median nerve NCS for detection of peripheral neuropathy performed in an NCS instrument (Keypoint Medtronic Functional Diagnostics), at GRNC Hospitals, Guwahati, India. Diagnosis, detection and patient management of Neuropathy has been carried out by researchers by digital signal processing of bio-neuro signals as discussed in various research articles. Those can be classified mainly under two categories- Statistical signal processing based classification and knowledge based classification.

1.1 STATISTICAL BASED CLASSIFICATION

Chuang-Chien Chiu et al [5] have investigated the feasibility of using power spectral density (PSD) analysis to continuous cerebral blood flow velocity (CBFV), function of cerebral auto regulation (CA) and continuous arterial blood pressure (ABP) with diabetic autonomic neuropathy (DAN). Frequency domain parameter of Heart Rate Variability (HRV) signal has been also used in [6] to detect Diabetic Cardiac Autonomic Neuropathy (DCAN). However, these methods cannot detect peripheral neuropathy since the techniques tries to relate cardiac abnormalities with diabetic autonomic neuropathy. In [7], Chen Haifeng et al have developed a new instrument on non-invasive measurement for the early diagnosis of the Diabetic peripheral neuropathy (DPN) by nerve conduction studies. Signal averaging and cross correlation technique have been adopted to classify four types of neuropathies among 60 patients - functional neuropathies, symptomatic neuropathies, Asymptomatic neuropathies, and normal. Although the NCS was adopted in this
method, the technique is limited to only diabetic peripheral neuropathy and cannot cover a wide range of neuropathy developed due various causes apart from diabetic mellitus.

1.2 KNOWLEDGE BASED CLASSIFICATION

In [8] three different muscle types were classified using ANN for diagnosis of Neuropathy. For this purpose, the electromyogram (EMG) signals were recorded from biceps, frontallis, abductorpollisibrevis muscles. For the modeling of EMG signals, autoregressive models were used to train and test several ANNs. The results of experiments show that Radial Basis Function (RBF) neural network has 93.3% accuracy to classify the muscles. In another work [9] various features like root mean square, spectrogram, Kurtosis, entropy and power of EMG signals of isometric contraction of one muscle abnormality- Amyotrophic Lateral Sclerosis were used. The classification was done by a knowledge-based expert system and disease diagnosis classifier. In [10] real time recordings of motor unit action potential (MUAP) signals from myopathy (MYO), neuropathy (NEU), and normal (NOR) subjects, using intramuscular electromyography (needle EMG) are treated and processed in Feedforward-backpropagation (FFBP) neural network.

In these works neuropathy detection has been performed by ANN where features are derived from EMG signals. EMG study is supplementary to NCS findings in detection of neuropathy where neuropathy is secondary to diseases of the spinal cord and anterior horn cells, with or without involvement of peripheral nerves. Moreover, EMG is done using needle electrodes, however needle EMG is painful and neurophysiologists prefer surface NCS than needle EMG. Moreover, motor NCS assesses the entire PNS because their endpoint is not a motor nerve action potential but rather a CMAP. Thus, the motor axons are evaluated by stimulating them and then recording the response in NCS [11]. The advantage of this arrangement is the signal magnification effect. Activation of a single motor axon causes the near simultaneous initiation of impulses in most individual muscle fiber (up to several hundred), the number depending upon the innervations ratio of the recorded muscle. The resulting CMAP amplitudes are of sufficient magnitude which can be measured in millvolts (mV). This is the principal reason why motor NCS became a diagnostic tool before sensory nerve conduction study. Motor nerve conduction study is a valuable diagnostic aid for several reasons. In 1961, Lambert [12] listed nine reasons for using motor nerve conduction study including the following-

i) Motor nerve conduction study provides objective evidence of motor unit abnormalities in patients suspected hysteria, malingering or upper motor neuron lesions

ii) Identify and localize focal lesions along individual nerves

iii) Separate polyneuropathies from both myopathies and motor neuron diseases

iv) Detect various disorders in neuromuscular transmission and distinguish them from one another

v) Reveal some peripheral nerves anomalies (e.g. MartinGruber anastomosis).

In detection of peripheral neuropathy, traditionally normal values of NCS are compared to matched “normal” values for NCS parameters, which are derived from studies of groups of neurologically normal subjects. The neurophysiologists assess a number of parameters together to make judgment whether a clinically relevant abnormality should be emphasized in the report or not. While doing so the clinical neurophysiologist faces problem in assessing and analyzing all the parameters together and then clinically correlating them. It needs time, accuracy and there is a possibility of subjective variation in interpretation of the data. So the neurophysiologists need an intelligent tool that should be able to help them to make a good decision.
The essentials and pitfalls of NCS have been discussed in [11] where the author regretfully stated that the most frequent statistics used are limits of 95% or less frequently 99% confidence limits of a normal group to indicate abnormality of a single parameter. This approach may mislead as a crude separation between ‘normal’ and ‘abnormal’ and dilutes the information. Instead a ‘score metric’, for example, indicating the separation between a single value and the group mean expressed in standard deviation, may be more informative. Alternatively, a number of electrophysiological parameters may be taken together either as an ‘index’ or ‘score’, or the neurophysiologist assesses a number of parameters together to make a judgment as to whether a clinically relevant numerical abnormality should be emphasized in the report interpretation or not[11].

The measurement and correlating of the amplitude and duration parameters is still a complicated task for the neurophysiologist and/or the computer-aided method used. The description of an extensively accepted criterion that will allocate the computer-aided measurement of neuropathy is still absent [11]. On the other hand, frequency domain features of NCS parameters like the mean or median frequency, bandwidth and other quality factor give supplementary information for the assessment of neuromuscular disorders only.

The motivation behind this work is to remove the traditional crude separation between ‘normal’ and ‘abnormal’ neuropathic condition indicated by the separation between a single NCS parameter value. The aim of this work is to model a score metric in the form of Mean Full Score (MFS) and standard deviation (SD) derived from the five vital NCS parameters. First the rank and sensitivity of the five NCS parameters in a terrain of 420 patients are determined using the ANN models. A weight equal to the sensitivity value is assigned to each parameter to get the score metric which are used to re-train the ANNs for detection of neuropathy.

2. MATERIALS

From March 2008 to December 2010, five NCS features of a total of 420 patients registered in the Department of Neurophysiology, GRNC Hospital suspected of peripheral neuropathy were taken and were evaluated by a group of neurologists for each patient for overall judgment on neuropathy. A total of 190 (45.08%) males and 230 (54.91%) females were included in the study. The inclusion criteria of patients were- pain, tingling, burning sensation, pricking sensation, weakness of limbs, difficulties in walking, and imbalance of gait. The patients’ characteristics are shown in the Table 2.

An NCS system (Keypoint; Medtronic Functional Diagnostics, Skovlunde, Denmark) at GRNC Hospitals, Guwahati, India was used for NCS study in this research. The NCS was conducted for the median,(MED) ulner(ULN), common peroneal(CPN) and posterior tibial nerves (PTN) on Abductor Pollicis brevis(ABP),Abductor Digiti Minimi(ADM), Exterior Digits,Bravis(EDB) and Hallucis Longus muscles respectively using surface electrodes of 10 mm diameter with a recording gel diameter 16 mm; impedance at 20 Hz below 200 kOhms. The single triggering pulse was for 0.1ms with a maximum current of 100mA. The signal was filtered with a lower cutoff frequency of 20Hz and an upper cutoff frequency of 10kHz. The sweep speed of the signal display was 5ms/div and a sensitivity of 5mV/div. The signal was sampled at 20kHz with a 12-bit resolution.

The first step in the development of any classification solution is to identify the independent input variable that contributes the classification decision. The five NCS variables or features included in this study are listed in the Table 1. Normalization of feature data is a vital first step in any ANN operation. In this method, each feature value was normalized in the scale [0 1] by vector normalization method defined as-
Table 2 Patient Characteristic ranges

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients [n]</td>
<td>420</td>
</tr>
<tr>
<td>Feature/Variable</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age [year]</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>39</td>
</tr>
<tr>
<td>Median</td>
<td>45</td>
</tr>
<tr>
<td>Range</td>
<td>20-60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CMAP:[Min: Max][mV]</td>
<td>MED [0.50-32.85]</td>
<td>ULN [1.59-24.29]</td>
<td>CPN [0.07-18.54]</td>
<td>PTN [0.21-41.48]</td>
</tr>
<tr>
<td>MNCV:[Min: Max][m/s]</td>
<td>MED [23.33-84.94]</td>
<td>ULN [0.67-104.96]</td>
<td>CPN [26.00-74.00]</td>
<td>PTN [29.18-90.00]</td>
</tr>
<tr>
<td>MFR:[Min: Max][ms]</td>
<td>MED [20.00-98.25]</td>
<td>ULN [20.00-77.75]</td>
<td>CPN [4.75-79.50]</td>
<td>PTN [20.50-71.00]</td>
</tr>
<tr>
<td>Temperature[°C]</td>
<td>Mean 27.66</td>
<td>Median 28</td>
<td>Range 26-30</td>
<td></td>
</tr>
</tbody>
</table>

\[ x_{n,i} = \frac{x_i}{x_m} \]  \hspace{1cm} (2)

Where \( x_i \) is the feature data value and \( x_m \) is the maximum value of the feature vector data. The pre-processed data is then partitioned into three subsets: a training set, a validation set and a testing set in the ratio of 60:20:20. The total data set consists of \( N_p \times N_f \times N_n \) i.e. 420 x 5 x 4,
where $N_p$ is the number of patients, $N_f$ is the number of features and $N_n$ is the number of nerves tested.

The patient characteristics with neurophysiologist diagnosis were plotted as histograms for different variables such as age (year), DML (ms), PML (ms), CMAP (mV), MNCV (m/s), FMR (ms) (Fig. 3(a-f)).

**Fig.3** Histograms of different patient data (a) Age (year); (b) DML (ms); (c) PML (ms); (d) CMAP (mV); (e) MNCV (m/s); (f) MFR (ms); (NP: Normal patient; AP: Abnormal patient)

### 3. METHODS

ANNs are mathematical models that can be defined as structures composed of a large number of densely interconnected, adaptive, simple processing elements (neurons) working in unison to perform massively parallel computations for data processing and knowledge representation [13]. The advantages of ANNs over other multi-factorial analysis techniques include their ability to model non-linear functions, robustness to noise in data, their capacity to learn and adapt to new data and capability to handle imprecise and fuzzy information [14]. Network of ANNs contain mainly an input layer, hidden layers and an output layer. In this work three ANN models were used-Feed-forward back-propagation (FFBP), Cascade feed-forward back-propagation (CFFBP) and Learning Vector Quantization (LVQ). Feed forward neural network (FFBP) is the simplest model, which consists of layers where the subsequent layer has a connection from the preceding layer. FFBP is trained using the BP algorithm according to the following equations-

$$U_k(t) = \sum_{j=1}^{n} w_{j,k}(t) x_j(t) + b_{o,k}(t)$$  \hspace{1cm} (3)

$$Y_k(t) = \varphi(U_k(t))$$  \hspace{1cm} (4)
where, $x_j(t)$ is input value of $j$ at time $t$, $w_{jk}(t)$ is the weight assigned by neuron $k$ to input value of $j$ at time $t$, $\phi$ is a nonlinear activation function, $b_k(t)$ is the bias of $k$-neuron at time $t$, and $y_k(t)$ is output from neuron $k$ at time $t$.

We have chosen Gradient Descent with Adaptive Learning Rate Back Propagation (GDA) which is a learning function that updates the weight and bias values according to the gradient descent with adaptive learning rate. An adaptive learning rate attempts to keep the learning step size as large as possible while keeping learning stable. The learning rate is made responsive to the complexity of the local error surface. First, the initial network output and error are calculated and at each epoch new weights and biases are calculated using the current learning rate. Using updated weights and biases new outputs and errors are then calculated. If the new error exceeds the old error by more than a predefined ratio, typically 1.04, the new weights and biases are discarded. In addition, the learning rate is decreased (typically by multiplying by 0.7), otherwise, the new weights, etc., are kept. If the new error is less than the old error, the learning rate is increased (typically by multiplying by 1.05). This procedure increases the learning rate, but only to the extent that the network can learn without large error increases. Thus, a near-optimal learning rate is obtained for the local terrain. When a larger learning rate could result in stable learning, the learning rate is increased. When the learning rate is too high to guarantee a decrease in error, it is decreased until stable learning resumes. The training and learning parameters of the ANNs are shown in Table 3.

The learning process or weight adjustments to minimize the error $(\epsilon_k)$ between the network’s desired and actual output using GDA iterative procedure can be written as-

$$e_k = (y_k - \hat{y}_k)y_k(1-y_k)$$

$$w_{j,k}(t+1) = w_{j,k}(t) - \mu(t)\frac{\partial e_k}{\partial w_{j,k}}$$

The CFFBP is similar to FFBP but CFFBP includes a connection from input and every previous layer to following layers. Additional connections improve the speed at which ANN learns the desired relationship. On the other hand LVQ consists of two layers, where the first layer maps input vectors into clusters that are found by the network during training. The second layer maps merge groups of first layer clusters into the classes defined by the target data. The total number of first layer clusters is determined by the number of hidden neurons. The larger the hidden layer, the more clusters the first layer can learn and the more complex mappings of input to target classes can be made. The ANN structure and different steps of the detection technique is shown in Fig 4.

<table>
<thead>
<tr>
<th>ANN</th>
<th>TF</th>
<th>AF</th>
<th>PF</th>
<th>$n$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFBP</td>
<td>Gradient descent</td>
<td>Logsig</td>
<td>Mean square error (MSE)</td>
<td>10,15,20</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>with adaptive learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFFBP</td>
<td>Gradient descent</td>
<td>Logsig</td>
<td>Train-sig</td>
<td>10,15,20</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>with adaptive learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LVQ</td>
<td>[0.5,0.5]</td>
<td>-0.01</td>
<td>Learn-lv1</td>
<td>10,15,20</td>
<td>5</td>
</tr>
</tbody>
</table>

Note: TF: Training function; AF: Activation function; PF: Performance function; OC: Output class; LF: Learning Function; LR: Learning Rate; $n$: Nos. of hidden neurons; $N$: Nos of layers
NCS database: 420 patients × 5 NCS parameters × 4 nerves

Neurophysiologist decision:
For General ANN: Neuropathy and Normal

Preprocessing:
For general ANN: Normalization of NCS features

Generate target matrix:
For ANN:

\[
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
\end{bmatrix}
\]

= normal (NOR)

= neuro (NEU)

General ANN:
Classification, sensitivity, Specificity and rank NOR and NEU

Pre-processing:
For score based ANN:
Evaluation of scores and Standard deviation and its normalization

Score based ANN:
Classification: NOR and NEU

**Fig 4** (a) FFBP ANN structure (b) stages of ANN processing for neuropathy detection
In disease diagnosis, detection or classification algorithm predictive power of the result is very important. The decision surface generated by the ANNs described must be tested on an independent data set to determine their effectiveness in reaching the accurate decisions. Five metrics are often used in medical applications to measure accuracy and predictive power—given by-

\[
\text{Sensitivity (} S \text{)}: \quad S = \frac{P_C}{P_T} \\
\text{Specificity (} S_p \text{)}: \quad S_p = \frac{N_C}{N_T} \\
\text{Accuracy (} A_c \text{)}: \quad A_c = \frac{P_C + N_C}{P_T + N_T} \\
\text{Positive predictive value (} PPV \text{)}: \quad PPV = \frac{P_T}{P_C + N_T} \\
\text{Negative predictive value (} NPV \text{)}: \quad NPV = \frac{N_T}{N_C + P_T}
\]

Where the measures are described as—Positive cases correctly classified \((P_C)\); Positive case incorrectly classified \((P_I)\); Total no. of positive cases \((P_T)\); Negative case correctly classified \((N_C)\); Negative cases incorrectly classified \((N_I)\) and total no. of negative cases \((N_T)\).

**Sensitivity** and **specificity** are statistical performance metrics of a binary classification test. **Sensitivity** of a test measure the probability or proportion of true positive cases correctly classified while **specificity** of a test measures the proportion or probability of negative cases which are correctly identified. The **accuracy** is the probability or proportion of true cases (true positive and true negative) correctly classified. The **PPV** of a test is the probability that a patient is detected as positive when a positive test result is observed. The **NPV** of a test is the probability that a case detected as negative when a negative test result is observed.

Further, the binary classifications are reflected in certain characteristics curves—Receiver operating characteristic (ROC) and area under the receiver operating characteristic curve (AUC). ROC curve display the relationship between **sensitivity** (true positive rate) and false positive rate (1-specificity) on unit square across all possible threshold values that define the positivity of a disease or condition [15-16]. Often ROCs are used to analyze the balance between sensitivity and specificity. A convenient method of consolidating both sensitivity and specificity into a single summary statistic is to use AUC [15]. The sensitivity is plotted along y-axis and 1-specificity is plotted along the x-axis. The goal is to try to find a combination that is as close as possible to the upper hand corner of unit square graph. This is measured in terms of AUC value. The higher value of AUC represents the higher predictive power of the method.

**Sensitivity** analysis for the input variables or features was performed to judge what parameters are the most and the least significant during generation of the satisfactory ANN. The sensitivity analysis provides insight into the usefulness of the individual variable or feature. For the **sensitivity** analysis the following **sensitivity** values are defined as weights—

- **DML**: Sensitivity \((W_1)\)
- **PML**: Sensitivity \((W_2)\)
- **CMAP**: Sensitivity \((W_3)\)
- **MNCV**: Sensitivity \((W_4)\)
- **MFR**: Sensitivity \((W_5)\)

Using these sensitivities as weights \((W)\) the full-scale (FS) for each patient had been calculated as discussed in section 3.1.
3.1 SCORE MATRIC EVALUATION

The clinical interpretation of NCS data relies on comparing an individual patient’s measurements with a reference range obtained from a healthy population. NCS reference ranges determine a normal limit above or below which a given NCS parameter is considered abnormal. Traditionally, NCS reference ranges are developed in individual laboratories, as each specialist has his own preferences in data-acquisition setups, such as filter settings and electrode placements that impact the NCS parameter values [4].

Therefore to assign scores to the NCS data, we have developed the pdf of the entire data set from which the normal ranges can be derived. In this work, each NCS parameter is found to have Gaussian distributions in their native domains of ‘normal’ and ‘abnormal’ patients. The PDF distribution of the five NCS variables for the 420 patients and the four tested nerves were plotted and the plot for the median nerve is shown in Fig.4. In each pdf of the NCS variables, there are two distinct Gaussian distributions (except that for CAMP)–for the normal and abnormal condition. The individual distributions were smoothed to a single distribution using a normal kernel function (NKF).

A kernel smoother is a statistical technique for estimating a real valued function \( f(X)(X \in \mathbb{R}^p) \) by using its noisy observations, when no parametric model for this function is known. The normal kernel algorithm is given as-

\[
X_0 \in \mathbb{R}^p
\]

Let \( \hat{Y}(X) : \mathbb{R}^p \to \mathbb{R} \) be a continuous function of \( X \). For each \( X_0 \in \mathbb{R}^p \), the Nadaraya-Watson kernel-weighted average (smooth \( Y(X) \) estimation) is defined by-

\[
Y(X_0) = \frac{\sum_{i=1}^{N} K_{h_0}(X_0, X_i)Y(X_i)}{\sum_{i=1}^{N} K_{h_0}(X_0, X_i)} \quad (12)
\]

The kernel function is given by-

\[
K_{h_0}(X_0, X_i) = D(\frac{\|X - X_0\|}{h_0(X_0)}) \quad (13)
\]

where:
- \( X, X_0 \in \mathbb{R}^p \)
- \( \| \| \) is the Euclidean norm
- \( h_0(X_0) \) is a parameter (kernel radius)
- \( D(t) \) typically is a positive real valued function, value of which is decreasing for the increasing distance between the \( X \) and \( X_0 \).

From the smoothed pdf distribution the mean of normal and abnormal values were evaluated as shown in Table 4.

The principle behind this first step classification solution is to identify the rank of the input features that contribute to the classification decision. In this method, each patient data of five NCS features for all the four nerves were scored according to the normal and abnormal scale. For example DML is considered normal if it falls within the range of 1.5-3.2ms otherwise it is considered as abnormal. The normal and abnormal scales of the five features are shown in Fig.6 for the median nerve. Each feature or variable are scaled in between 0 and 10. The scales shows that the parameters DML, PML and MFR score higher with lower values (Score-I) while CMAP and MNCV score higher with higher values (Score-II).
Fig 5 PDF of five NCS variables for *median nerve*  

Fig 6 Scores against the *normal* and *abnormal* ranges of the five NCS variables for *median nerve*

### Table 4 Statistical parameters of the patient PDF distribution

<table>
<thead>
<tr>
<th>NCS variable</th>
<th>MED</th>
<th>ULN</th>
<th>CPN</th>
<th>PTN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Total</td>
<td>Normal</td>
<td>Total</td>
</tr>
<tr>
<td>DML</td>
<td>$m_{nor}$</td>
<td>$m_T$</td>
<td>$SD_T$</td>
<td>$m_{nor}$</td>
</tr>
<tr>
<td>PML</td>
<td>9.0</td>
<td>12.0</td>
<td>3.3</td>
<td>8.9</td>
</tr>
<tr>
<td>CMAP</td>
<td>12.6</td>
<td>12.6</td>
<td>5.4</td>
<td>15.2</td>
</tr>
<tr>
<td>MNCV</td>
<td>56.7</td>
<td>53.6</td>
<td>7.3</td>
<td>55.9</td>
</tr>
<tr>
<td>MFR</td>
<td>27.3</td>
<td>34.6</td>
<td>11.1</td>
<td>27.79</td>
</tr>
</tbody>
</table>

Feature reduction is another important reason for converting the feature values to a score in this method. We have assigned a Full Score ($FS$) to a single tested nerve of the patients and then the mean full score ($MFS$) is calculated for the four tested nerves using the following equation.

$$FS = \sum_{i=1}^{5} W_i S_i$$  \hfill (14)$$

Where $W_i$ is the weight and $S_i$ is the score value of the $i^{th}$ feature variable.

Traditionally neurophysiologist performs NCS to all the suspected nerves and then takes a judgment based on the NCS data. In our work, it has been observed that there is a mixed correlation between the individual nerve condition and overall judgment on ‘normal’ or ‘abnormal’ on peripheral neuropathy. The confusion matrix in Table 5 shows that the individual nerve condition has very least sensitivity to the overall judgment of the neurophysiologist. For example the classification for normal median nerve has a sensitivity of only 0.38 while that for the abnormal median nerve falls behind the final decision. Therefore we have taken the sensitivity of each normal nerve to determine a mean value of $FS$. 


Table 5 Confusion matrix of patient diagnosis

<table>
<thead>
<tr>
<th>Nerve condition</th>
<th>Normal (N)</th>
<th>Abnormal (A)</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MED</td>
<td>117</td>
<td>185</td>
<td>0.38</td>
</tr>
<tr>
<td>CPN</td>
<td>117</td>
<td>44</td>
<td>0.72</td>
</tr>
<tr>
<td>PTN</td>
<td>117</td>
<td>67</td>
<td>0.63</td>
</tr>
<tr>
<td>ULN</td>
<td>117</td>
<td>133</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Note: FB: Fall behind

After calculating the FS of each patient, the mean value of FS (MFS), variance (\(\sigma\)) and standard deviation (SD) were calculated for each patient using the following equations

\[
MFS = \frac{0.38FS_{med} + 0.46FS_{ulner} + 0.72FS_{CPN} + 0.63FS_{PTN}}{4}
\]  
(15)

The variance and standard deviation (\(\sigma^2\)) are calculated by-

\[
\sigma^2 = (X - MFS)^2
\]  
(16)

\[
\sigma = \sqrt{(X - MFS)^2} = \pm m
\]  
(17)

Where \(m\) is the standard deviation.

The features were then normalized between \([0, 1]\) using the vector normalization method of Equation (1). To distinguish between patients with equal value of FS we have considered the \(\pm m\) values.

In this score based method, the five features for four nerves of a patient (total 20 features) are reduced to only three features- \(MFS, (MFS + m)\) and \((MFS - m)\). This score based feature reduction technique is the key to the improvement in classification rate of the ANN classifiers.

The input matrix having dimension 3x420 was subdivided into three matrices (3x252, 3x84, and 3x84) for training, validation and testing respectively for ANN analysis. Fig 7 shows the flowchart of the score based ANN classification algorithm.
4. RESULTS

First we have trained the ANN models with the normalized five features for all the four nerves to analyze the classification accuracy, sensitivity, specificity and predictive power for three different ANN paradigms with different hidden neurons.

4.1 ANN CLASSIFIERS WITH FIVE NCS FEATURES

The classification accuracy of three different ANN models with different number of neurons \((n=10, 15, 20\) and \(35\)) is listed in Table 6 and it was found that the FFBP model has higher training accuracy of 94.04\% \((n=20)\) compared to CFFBP and LVQ.

The sensitivity, PPV, specificity and NPV were calculated with three models (Table 6). It was found that the FFBP model has higher sensitivity, PPV, NPV and specificity than the other two models.

In Table 6 the classification speed, epoch (complete cycle of iteration) and \(MSE\) of the three models are shown. It was found that the classification speeds of FFBP and CFFBP have higher than the LVQ model and the \(MSE\) of these two model are \(4.50x10^{-7}(n=20)\) and \(0.000242\) \((n=20)\) respectively which is again higher than the LVQ model. So, for classification decision FFBP and CFFBP can be more widely used.

In this study of sensitivity analysis to judge the significance of each NCS parameters we have found that CMAP (Rank-1) is the most significant variable. Other highly significant variables are DML (Rank-2) and PML (Rank-3) while the least significant one is MFR (Rank-5) (Table 7).

In this method, the ROC curves of training and testing set with the best performing model (FFBP; \(n=20\)) were plotted in the Fig.8 (d) and (e). The AUCs of training and testing sets are found to be 0.976 \((n=20)\) and 0.965 \((n=20)\) respectively. The histogram plot of the ANN output for both normal patients (NP) and abnormal patients (AP) are shown in Fig.8 (a-c).
### Table 6 Result of ANN Classification with five NCS features

<table>
<thead>
<tr>
<th>Data size</th>
<th>ANN</th>
<th>n</th>
<th>ST (A:N:CT)</th>
<th>Ac (%) (T)</th>
<th>SP (%)</th>
<th>PPV</th>
<th>NPV</th>
<th>Epoch</th>
<th>CPU time(s)</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FFBP</td>
<td>10</td>
<td>49:25:74</td>
<td>88.10</td>
<td>83.33</td>
<td>1.00</td>
<td>1.00</td>
<td>567</td>
<td>5.19</td>
<td>4.42x10^-3</td>
</tr>
<tr>
<td></td>
<td>FFBP</td>
<td>15</td>
<td>50:27:77</td>
<td>91.67</td>
<td>90.00</td>
<td>1.02</td>
<td>0.96</td>
<td>714</td>
<td>7.19</td>
<td>4.42x10^-3</td>
</tr>
<tr>
<td></td>
<td>FFBP</td>
<td>20</td>
<td>51:28:79</td>
<td>94.04</td>
<td>93.33</td>
<td>1.02</td>
<td>0.97</td>
<td>795</td>
<td>7.39</td>
<td>4.42x10^-7</td>
</tr>
<tr>
<td>251:85:84</td>
<td>FFBP</td>
<td>35</td>
<td>43:22:65</td>
<td>77.38</td>
<td>73.33</td>
<td>1.07</td>
<td>0.93</td>
<td>925</td>
<td>16.45</td>
<td>5.72x10^-2</td>
</tr>
<tr>
<td></td>
<td>LVQ</td>
<td>10</td>
<td>48:26:74</td>
<td>88.09</td>
<td>83.33</td>
<td>1.02</td>
<td>0.96</td>
<td>25</td>
<td>20.06</td>
<td>1.76x10^-2</td>
</tr>
<tr>
<td></td>
<td>LVQ</td>
<td>15</td>
<td>48:28:76</td>
<td>90.47</td>
<td>93.33</td>
<td>1.08</td>
<td>0.88</td>
<td>75</td>
<td>64.27</td>
<td>1.32x10^-2</td>
</tr>
<tr>
<td></td>
<td>LVQ</td>
<td>20</td>
<td>49:27:76</td>
<td>90.47</td>
<td>90.00</td>
<td>1.03</td>
<td>0.93</td>
<td>100</td>
<td>86.04</td>
<td>8.84x10^-3</td>
</tr>
<tr>
<td></td>
<td>CFFBP</td>
<td>10</td>
<td>45:28:73</td>
<td>86.90</td>
<td>93.33</td>
<td>1.49</td>
<td>0.81</td>
<td>377</td>
<td>5.07</td>
<td>4.25x10^-3</td>
</tr>
<tr>
<td></td>
<td>CFFBP</td>
<td>15</td>
<td>48:29:77</td>
<td>91.67</td>
<td>96.67</td>
<td>1.10</td>
<td>0.85</td>
<td>408</td>
<td>6.19</td>
<td>4.25x10^-3</td>
</tr>
<tr>
<td></td>
<td>CFFBP</td>
<td>20</td>
<td>50:28:78</td>
<td>92.85</td>
<td>93.33</td>
<td>1.03</td>
<td>0.93</td>
<td>443</td>
<td>7.09</td>
<td>4.20x10^-3</td>
</tr>
</tbody>
</table>

**Note:** n-No. of hidden neuron; ST:-size of testing set, A:-Abnormal data, N:-Normal data and CT:-correct total data , Ac:-Accuracy.

**Fig.8.** Histograms of (a) ANN output; (b) Normal patient (zoomed fig.1) ; (c) abnormal patient (zoomed fig.2)  (d) ROC of training set and (e) ROC of testing set );(NP: Normal patient; AP: Abnormal patient)

### 4.2 RESULTS OF SCORE BASED ANN

The FFBP ANN models were used for improvement of classification accuracy using score based features. The classification accuracy of score based ANN model with different number of neurons (n=25, 30, 35 and 40) are listed in Table 8. In case of score based ANN classification it was
found that the classification accuracy of the model increases and attains a training accuracy of 96% \((n=35)\). If the number of hidden neuron is increased further \((n=40)\) the training accuracy again falls which indicates the over fitting of the model.

The sensitivity, PPV, specificity and NPV were calculated for the ANN classification model (Table 8). It was found that the FFBP model with \(n=40\) has the highest sensitivity, specificity, PPV and NPV.

In Table 8 the classification speed, complete cycles of iteration (epochs) and \(MSE\) are shown. It was found that the classification speed of FFBP with \(n=35\) is highest \((3.1408s)\) while \(MSE\) is lowest \((2.31x10^-3)\) with \(n=30\). So, for classification of peripheral neuropathy with the proposed score based method FFBP with \(n=35\) is more suitable.

In this method, the ROC curves of training and testing set with the best performing model (FFBP; \(n=35\)) were plotted as shown in Fig.9. The AUCs of training and testing set are found to be \(0.987\) \((n=35)\) and \(0.985\) \((n=35)\) respectively.

<table>
<thead>
<tr>
<th>Table 7 Sensitivity analysis of feature variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>DML</td>
</tr>
<tr>
<td>PML</td>
</tr>
<tr>
<td>CMAP</td>
</tr>
<tr>
<td>MNCV</td>
</tr>
<tr>
<td>MFR</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 8 Accuracy, Sensitivities, positive predictive values, specificities and negative predictive values, classification speed of FFBP models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of data set</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>FFBP</td>
</tr>
<tr>
<td>251:85:84</td>
</tr>
<tr>
<td>FFBP</td>
</tr>
<tr>
<td>FFBP</td>
</tr>
</tbody>
</table>

Note: \(n\)-No. of hidden neuron; \(ST\):-size of testing set, \(A\):-Abnormal data, \(N\):-Normal data and \(CT\):-correct total data, \(Ac\):-Accuracy, \(SN\) sensitivity \(SP\)-specificity

![Fig.9 ROC (a) training set and (b) ROC of testing set](image-url)
5. CONCLUSION AND DISCUSSION

From this study it has been demonstrated that a ‘core metric’ based on NCS variables can train the ANN in a better way than only taking the variables as features in early detection of peripheral neuropathy. This will eliminate the problem of subjective judgment and complexity for the neurophysiologists. In the traditional method neurophysiologists perform a crude separation between ‘normal’ and ‘abnormal’ patients based on abnormality of single parameter with a certain confidence limit, however in this proposed method predictivity has been improved by assessing a number of parameters together by the ANN to make a judgment.

Our classifier demonstrated an **accuracy** of 96.42%, **sensitivity** of 96.29% and **specificity** of 96.66%. Although three models have been used, ROC curve of only FFBP models has been analysed as this model has higher training accuracy in comparison to the other two models. The major factor in quantifying the discriminating ability of ANNs is the choice of numbers of hidden neurons \( n \) since the **accuracy**, **sensitivity** and **specificity** depend on it. No theoretical guidelines exist to determine how an ideal value of \( n \) could be chosen. One possible method for selecting the optimum \( n \) would be to increase the number of \( n \) in the ROC curve in order to get more **sensitivity-specificity** pairs. An optimal selection of \( n \) where both **specificity** and **sensitivity** were maximized could therefore be determined.

In this study the potential application of ANN has been proved based on clinical studies on peripheral neuropathy. The application of ANN has been validated in the neuropathy detection study that represents one of the best methods, offering the possibility of early and reliable detection of normal and abnormality of nerve. In this study, three models have been used and in comparison to CFFBP and LVQ, FFBP has higher predicting power and processing speed. Our result shows that ANN model is helpful to neurophysiologists for patient management by reducing the inconvenience, expense and possible treatment delay.

6. ACKNOWLEDGEMENTS

The authors would like to thank Dr. N. C. Borah, GNRC Hospital, Guwahati for his assistance in obtaining the data from the patient databases which have formed the basis for this research.

7. REFERENCES


