Chapter 2

Materials and Methodology

2.1 General

This chapter describes a brief overview of the materials, planning, experimental procedures, apparatus and basic methodologies used in the dissertation which is outlined in brief, as follows:

(a) Selection of base materials.

(b) Factors affecting strength and durability of stabilized material.

(c) Planning of experiment

(d) Tests for index and mechanical properties.

(e) Testing procedure.

(f) Evaluation of mineralogy and morphology of red mud before and after treatment.

(g) Method of estimation of mechanical behavior using experimental designed approach and

(h) Working procedure of artificial neural network.

2.2 Material Selection

Industrial waste drawn from alumina plant (red mud) is used as the base material in this study. A brief discussion of red mud is presented herein.

2.2.1 Red Mud

Red mud used in the study was collected from Hindalco Industries Limited, Renukoot, Uttar Pradesh, (N24°13.33′ and E83°1.50′), India. It is dumped/discharged after extraction of alumina from bauxite ore as loose cake in the pond situated in vicinity of the plant (Fig. 2.1). It was oven-dried at 105 ± 5°C for moisture removal and screened by 425-micron screen before use in the entire study.

2.2.1.1 Physical and Chemical Properties

It appears reddish in color and contains mostly fine particles on visual examination. It is classified as silt with low compressibility (ML) according to unified soil classification system ($USCS$). Grain size distribution of red mud is shown in Fig. 2.2.

The physical and chemical properties of red mud are summarized in Tables 2.1 and 2.2 respectively. Commercial grade dry hydrated lime ($CaOH_2$) is used as the cementing agent and it consists mainly of oxides of Ca and Mg. Detailed chemical

FIGURE 2.1: Location of alumina plant (source: Google Earth, accessed on 04/12/2014,14:00)

properties are summarized in Table 2.2. The physical and chemical properties of red mud are summarized in Tables 2.1 and 2.2 respectively.

TABLE 2.1: Physical and geotechnical properties of red mud

Properties	Value
Maximum dry density (kN/m^3)	15.0
Optimum moisture content (%)	30
pH	10.2
Specific gravity	3.1
Sand content (%)	28
Silt size particle (%)	60
Clay size particle (%)	12
Effective diameter , D_{10} (mm)	0.007
Coefficient of uniformity, C_u	8.57
Coefficient of curvature, C_c	0.69
Unified soil classification system ($USCS$)	ML
Unconfined compressive strength (kPa) (Corresponding to maximum dry density = $15kN/m^3$)	138.6

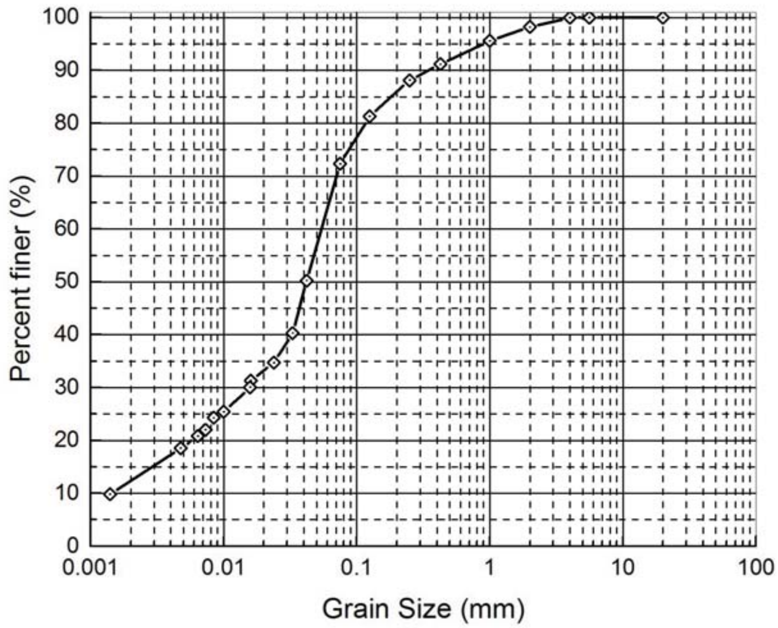

FIGURE 2.2: Grain size distribution of red mud.

TABLE 2.2: Chemical properties of red mud and lime

Compound (Wt. %)	Redmud	Lime
Al_2O_3	32.26	1.48
SiO_2	25.84	4.29
Fe_2O_3	25.00	0.274
TiO_2	7.09	0.109
Na_2O	5.42	0.47
P_2O_5	1.64	0.653
SO_3	0.634	4.56
ZrO_2	0.630	-
CaO	0.386	71.46
Cr_2O_3	0.0988	-
CuO	0.0852	0.206
NiO	0.0335	0.0898
MnO	0.0308	0.0286
PbO	0.0245	0.037
CO_3O_4	0.0197	-
ZnO	0.0128	-
MgO	-	15.48

\- Not Detected

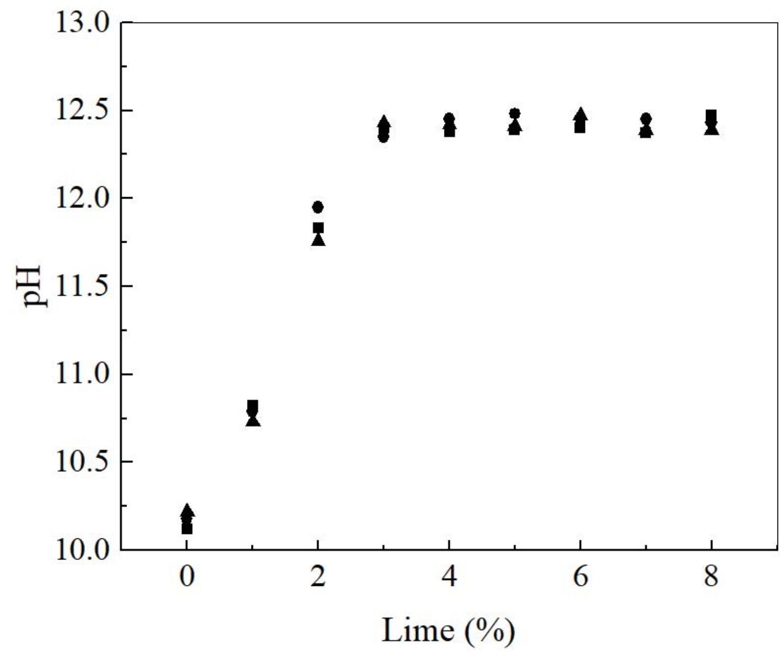

FIGURE 2.3: Variation of pH of red mud with lime

2.3 Initial Dose of Lime

Estimation of the soil-lime proportion requirement is an important factor for soil stabilization [44, 210]. In this study, initial requirement of lime has been chosen as per $ASTMD$6276 [211]. In this process soil is thoroughly mixed with different lime quantities and put in glass bottles, then the distilled water added into the liquid at solid ratio 4 and the soil-lime water is shaken for at least 30 s or until the specimens are thoroughly mixed after recording and drawing pH of all the samples. The minimum percentage of lime in soil that gives a pH of 12.4 is the approximate lime percentage for stabilizing the soil and also termed as lime fixation point. The variation of pH of red mud with addition of lime is shown in Fig.2.3.

From the Fig. 2.3, it appears that about 3% of lime stops the pH variation (corresponding to pH of 12.4) and hence chosen as minimum amount of lime for the stabilization of red mud. The concept is well known to increase the strength of the mixture soil-lime by an increase in lime. In addition, literature also supports that this increase in strength depends on time and can last for weeks and years .

2.4 Sample Preparation

Firstly, the required amount of red mud and lime were mixed in a dry state to a uniform consistency. Water was then added while continuing the mixing process until a uniform, homogeneous mixture was obtained. After mixing, it was compacted in three layers in a split mold. After the molding process, the sample was extracted from the mold and placed in an airtight polythene bag. The sample was allowed to cure for desired curing periods in desiccators at room temperature ($23 \pm 3\,^{\circ}$C) by maintaining the relative humidity of more than 95%. After the molding and curing processes, the sample was tested in an automatic load compression device of 50 kN capacity with a proving ring of 10 kN capacity.

2.5 Selection of Factors

Five factors were selected for the investigation of evaluation of strength and durability of red mud on the basis of the literature review and were included the experiment. Table 2.3 summarizes the factors and their range used in the present study. Strength behavior is investigated by the factors (i, ii, iii and iv), while factors (i, ii, iii, iv and v) are considered for the durability characteristics of the stabilized red mud.

TABLE 2.3: Summaries of factors and their ranges selected for the study

Factor	Designation	Variable Type	Purpose
Additive content (%)	i	Quantitative	Influence of cementing agent
Curing Period (days)	ii	Quantitative	Effect of curing time
Molding moisture content (%)	iii	Quantitative	Effect of molding moisture
Dry unit weight, (kN/m3)	iv	Quantitative	Effect of packing of particle
Wetting drying cycle (N)	v	Quantitative	Weathering effect

N= Number of cycle

2.6 Planning of Experiment

Two different approaches, conventional and design of experiment, are used in the present study.

2.6.1 Conventional Approach

In conventional method, one factor at time ($OFAT$) approach is used to plan the experiment. In this approach, one factor is varied by keeping other factors constant and observing their influence on the response (output). Furthermore, this process is repeated by varying other factors one by one until all the factors have been treated for the range of experiment. Thus, it results in more number of experiments. Furthermore, it gives no idea of the effect of interactions of more than one factor on the output of the experiment. So, experimental designed (design of experiment) approach is being used to plan the experiment to overcome these limitations.

2.6.2 Design of Experiment/ Experimental Designed Approach

It uses a multi factor at time ($MFAT$) approach to design the experiment, that results in fewer number of experiments compared to the conventional approach. Replication, randomization and blocking are the basic steps in the design of the experiment. Replication means repeating the experiment in order to achieve a more precise result (mean value) and estimate the experimental error. Randomization is the random order in which the experiment is conducted. In this way, the conditions in one run are not dependent on the previous run conditions or on the following runs. This is achieved by organizing experiments in groups similar to each other in order to reduce the sources of variability and improve precision. The methodology of design of experiment is available in detail in reference books and literature [190–194, 212] and hence, are not described in detail herein.

The choice of a suitable design of experiment technique is an important step and solely depends on the nature, goal, and aim and complexity of the experimentation. However, certain guidelines for the selection of appropriate experimental methods are available in the literature [212, 213]. For example, randomized complete block design and Plackett-Burman method would be preferable if the objective is to focus only on the primary factor for estimating the output. If the main variable and their interaction effects need to be incorporated, it is better to choose a fractional or a full factorial method. When noise variables can have a significant influence on the problem, Taguchi method is proposed. In the case of surface response methodology (when the aim is to obtain an optimum response), complete factorial, central composite, Box-Behnken or a space filling technique can be chosen for design of experiment. The details of various design of experiment techniques, their cost in term of number of experiments, and their aims are summarized in Table 2.4.

Based on the suitability of methods summarized in Table 2.4, central composite and Box-Behnken method are chosen in the present study.

TABLE 2.4: Summary of various design of experiment methods

Method	Number of experiments (N)	Suitability
Randomized complete block design	$N(M_i) = \prod_{i=1}^{Z} M_i$	Estimating the main effects
Latin squares	$N(M) = M^2$	Focusing on a primary factor
Full factorial	$N(M, Z) = M^Z$	Computing the main and the interaction effects, building response surfaces
Fractional factorial	$N(M, Z, P) = M^{Z-P}$	Estimating the main and the interaction effects
Central composite	$N(Z) = 2^Z + 2Z + 6$	Building response surfaces
Box-Behnken	$N(Z) = 2Z(Z-1) + 6$	Building response surfaces
Note: M= Level of independent variables	Z= Number of independent variables.	

2.6.2.1 Central Composite Design

The central composite design is essentially a full factorial design method with two levels with additional central and star points. The star points are the sample points where all but one parameter is set at the mean level. The remaining parameter value is given by the distance from the center point. For example, if the distance between the center point and each full factor point is normalized to 1 and the distance of the star points from the center point is set to 1, all samples are placed on a hypersphere with respect to the central point and this is known as the circumcised central composite. Likewise, if the distance of the star points from the center point is set to $\frac{\sqrt{L}}{L}$, the central composite design is called. The graphic representation of various central experimental composite designs is shown in Fig 2.4.

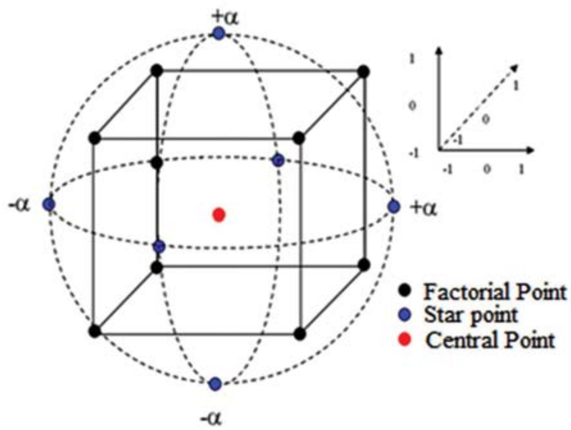

FIGURE 2.4: Schematic representation of central composite experimental designs

In central composite design, $2 * M$ factorial, $2 * M$ star and six (minimum one) central points is required to design the experiment for M variables. Thus, it can be classified into three types based on the position of the star points and can be presented as follows

- Circumscribed central composite (CCC) - needs five levels (-α, -1, 0, 1, α).
- Inscribed central composite (ICC) - needs five levels (-1, -1/α, -0, 1/α, 1).
- Face-centered composite (FCC) needs three levels (-1, 0, 1); $\alpha = \pm 1$.
  where, $\alpha = [2^M]^{\frac{1}{4}}$ = Position of star points.

2.6.2.2 Box-Behnken Design

Box-Behnken design (BBD) is a three - tier factorial method used in experiment design. It does not use the extreme combinations of all input parameters but compensates for the points in the factor space center. It requires factor points $(2{*}M{*}(M{-}1))$ and five central points for the number of independent variables in L.However, this

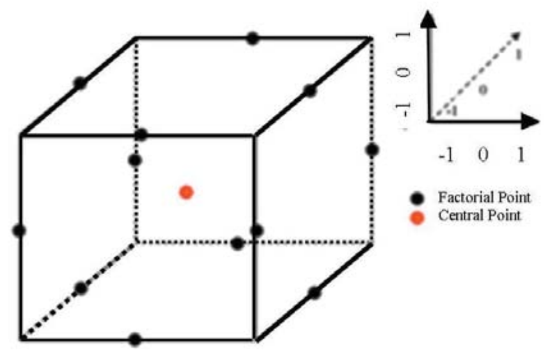

Figure 2.5: Schematic diagram of Box-Behnken experimental design

method has a limitation that it works better for 3 or more factors. Fig 2.5 shows the schematic representations of the central composite design and Box-Behnken design.

2.7 Test Procedure and Equipment

Characterization is the first and most important step in understanding the soil's behavior. To this end, it is necessary to understand the mechanism of changes in the characteristics of red mud due to treatment. The experimental procedure and equipment used to perform physical, chemical and microstructural characteristics of red mud with and without treatment are presented and summarized systematically in the following sub - sections.

2.7.1 Index properties

Index properties are the physical properties which help to identify and classify the soil for various engineering purposes. It also indicates a qualitative soil behavior under the loads.

2.7.1.1 Grain Size Distribution

Firstly, the soil samples were washed through 75 µm sieve and then the fraction coarser and finer than 75 µm were collected and oven dried . In addition, mechanical sieving [214] and hydrometer methods [215] are used to analyze coarser fractions and finer fractions. Finally, the results are plotted between % finer (passing) and log of particle size.

2.7.1.2 Specific Gravity of Solids

Specific gravity tests were performed using density bottle in accordance with IS 2720 (Part 3) [216]. Firstly, weigh the empty bottle of 50 ml with a stopper and then take approximately 10 to 20 g of oven dry air cooled soil sample and transfer it to the bottle and record the weight of the bottle and soil. After that, 10 ml of distilled water is placed in the bottle for 2 hours to completely soak the soil. Fill the bottle with distilled water again, place the stopper and keep the bottle under constant temperature water baths, and then the bottle is taken out and washed clean and dry, with the weight of the bottle and its contents (surface and water). The density bottle is finally filled with distilled water and weighed. The same procedure is repeated 2 to 3 times, and the specific gravity of the soil is reported as the average of three readings.

2.7.1.3 Atterberg Limits

It provides information about different states such as solid, semi - solid, plastic and liquid that is also linked to physical and mechanical behavior (Fig 2.6). ASTM

D4318-17e1 [217] was used to determine the liquid limit, plastic limit and shrinkage limit.

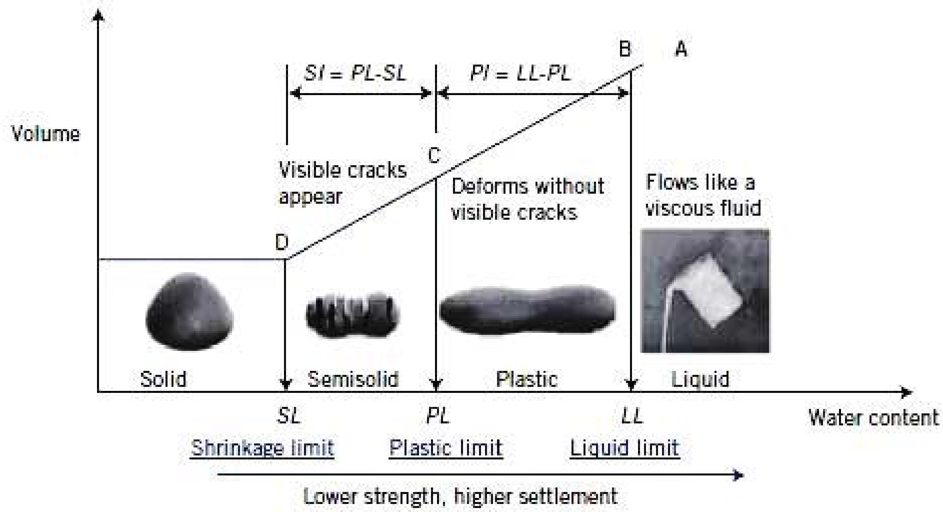

FIGURE 2.6: Volume-moisture content relationship [218]

2.7.2 Compaction Characteristics

Compaction test (standard and modified effort) was carried out in accordance with ASTM D698 [219] and ASTM D1557 [220]. A cylindrical metal mould with an inner diameter of 101.6 mm and a length of 116.5 mm with detachable collars and base plate was used for standard and modified effort. In the compaction of the prepared soil sample, 25 blows from a rammer of 2.5 kg in weight fall freely through a height of 305 mm are carried out in 3 layers whereas in the case of a modified proctor test, the soil is compacted in 5 layers with 25 blows per layer in the standard Proctor mold with a hammer of 4.54 kg in weight and falls freely from 457 mm. For each compaction test, approximately 2.5 kg of soil is used. The required amount of water is added to the ground and mixed thoroughly to balance the humidity. The mold is cleaned, dried and greased to reduce friction on the sidewall and extrude the compacted sample after the test. The mold is then fixed to the base plate. The soil is compacted in the mold in three layers. Approximate quantity of the soil is placed in the mold for the first layer and the required number of blows is then applied to the soil by dropping the hammer on the energy transferred by the foot of the frame. When the hammer strikes the energy that transmits the foot that the frame (top rod) is not in touch with the hand, attention should be paid. The surface of the ground is scarified before the second layer is applied after the required number of blows. For the second layer, the mold is filled with the soil and compacted again. The top collar is placed on the mold when the second layer is compacted and the third layer is placed and compacted. After compaction, the collar is removed and excess soil is cut off to form the level surface.

The weight of compacted soil and mold is noted and the compacted soil weight is determined. In addition, a small amount of soil sample from the middle layer is taken to determine the sample's moisture content. The results of the compaction test are presented as dry density compared vs water content and from the plot, the maximum dry density and the optimal water content are determined.

2.7.3 Strength Characteristics

The safety of the structure depends on many factors including the soil's strength. If the ground fails, a structure on or inside it can collapse, endanger lives and cause economic damage. For any located structure, therefore, an assessment of soil strength is necessary.

2.7.3.1 Unconfined compressive strength Test

Unconfined compressive strength is quick, reliable and widely accepted laboratory test for determining the shear strength. Unconfined compressive strength test (UCS) were conducted in accordance with ASTM-D2166 [221]. A metallic mould, having size 50 mm inner diameter and 100 mm long, with additional detachable collars at both ends were used to prepare cylindrical specimens. Firstly, the required quantity of red mud and additives corresponding to the chosen dry weight of mix was then mixed thoroughly and then required quantity of water was added to the mix. After mixing, it was compacted in three layers in a split mould into 50 mm diameter and 100 mm high sample to a targeted dry unit weight of the mix. To ensure uniform compaction, specimen was compressed statically from both ends it just reached the dimensions of the required specimen size. After the molding process, the sample was extracted from the mould with universal hydraulic sample extrusion facility (Fig 2.7) and placed in an airtight polythene bag.

It was then cured for desired curing periods in a desiccator at room temperature ($23 \pm 3\,°C$) by maintaining the relative humidity of more than 95%. After the molding and curing processes, the sample was tested at a strain rate of 1.2 mm per minute in an automatic load compression device of 50 kN capacity with a load cell of 10 kN capacity (Fig 2.8). The mean of three test is reported as UCS value of the corresponding mix.

2.7.3.2 Split Tensile Strength Test

Tensile strength is an important mechanical parameter associated with predicting the possibility of tensile cracks commonly found in earth structures such as dams, slopes, sub grades of the runway, road and railway. Split strength test was carried

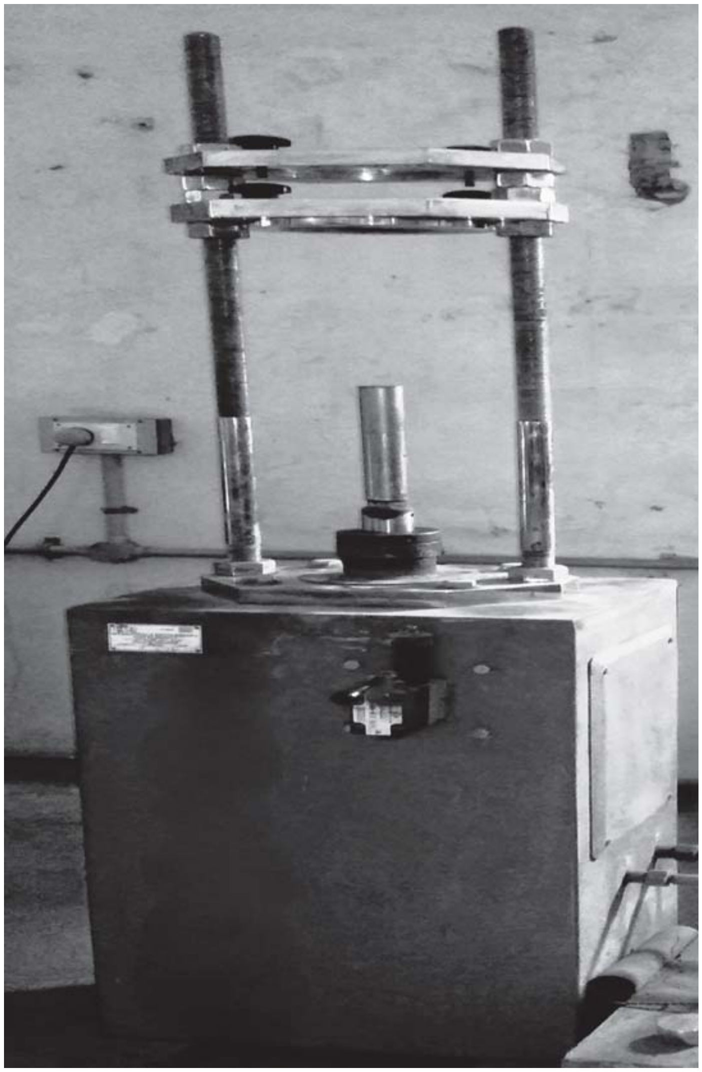

FIGURE 2.7: Universal hydraulic sample extruder(Departmental laboratory, IIT(BHU), Varanasi)

out in accordance with ASTM D3967-08 [222]. The samples were prepared using the same method as the unconfined compressive strength test with an aspect ratio of 0.5, i.e., 50 mm in diameter and 25 mm in length. A cylindrical specimen is placed between the platens of a split tensile strength test apparatus (Fig 2.9).

After placing the sample, it was compressed by loading at a strain rate of 1.2 mm/min that causes a tensile deformation in the sample perpendicular to

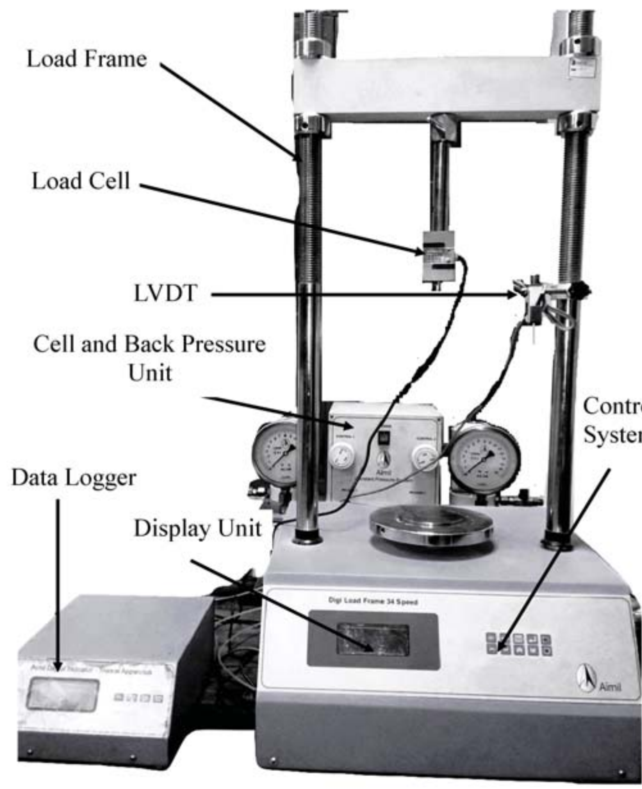

FIGURE 2.8: Automatic unconfined compression soil testing machine setup (Departmental laboratory, IIT(BHU), Varanasi)

the loading direction, resulting in a tensile failure. By recording the corresponding ultimate load and knowing the size of the specimen, the split tensile strength of the material is calculated.

2.7.4 Durability Study

For stabilized compacted materials, durability is an important aspect. It demonstrates the weather resistance of such materials. Due to climate change, weather-drying cycles can lead to tension and surface cracks in compacted stabilized materials

FIGURE 2.9: Split tensile mold assembly(Departmental laboratory, IIT(BHU), Varanasi)

reducing their resistance. Wetting-drying and freezing-thawing are two commonly used techniques for stabilized material durability study [223]. The freezing-thawing study is preferred in cold regions, while the wetting-drying study is preferred in warm regions. Wetting-drying tests better simulate conditions in India. The wetting-drying test was therefore considered in the present study.

2.7.4.1 Wetting Drying Test

Wetting-drying tests were performed according to ASTM D559/D559M-15 [224]. Following the procedure described in section 2.7.3.1, the samples were prepared and cured. After the curing process, the specimens were submerged in tap water at 23°C for 5 hours and then placed in an oven at 71°C for 42 hours and then removed. Afterwards, 18 to 20 vertical brush strokes were applied to the entire height and width of the specimen and the mass loss of the specimen was recorded.

The procedure described above form a cycle of wetting and drying (48 h). Samples were again submerged in water and the whole process was repeated for 12 cycles.

2.7.5 Micro structural Analysis

Micro structural observation is very important for cemented soil. The EVO 18, Zeiss 4.0 high resolution scanning electron microscope (SEM) with energy-dispersive X-ray spectroscopy (EDS) detector was employed for characterizing and examining the morphology of red mud specimen before and after treatment (Fig. 2.10).

FIGURE 2.10: Scanning electron microscopy with energy-dispersive X-ray spectroscopy detector setup (Departmental laboratory, IIT(BHU), Varanasi)

The failed specimens, after performing unconfined compressive strength were preserved for SEM-EDX test. Prior to the scanning process, a small amount of oven-dried and fine powdered sample is mounted on the tape glued to the flat surface of the gold - coated SEM stub and sputter. Several images are recorded at different magnification (20X to approximately 30,000X) for examining the variation in the crystal structures and spatial changes in chemical compositions (acquiring elemental maps using EDS) of red mud before and after treatment.

2.7.6 Preparation of Leachate Solution and Testing for Trace Metals

In the present study, toxicity leachate characteristics procedure ($TCLP$) by USEPA is adopted for preparation of leachate solution of raw red mud with and without treatment. In this procedure, the failed specimen, after unconfined compressive strength test was immersed in the leachate solution of pH 4.93, prepared with glacial acetic acid and 1N Sodium hydroxide ($NaOH$)in liquid to solid ratio 20:1. Further, the solution was agitated in a rotary extractor for a period of 18 hours at 30 rpm at room temperature. After the agitation, the solution was loaded into close air-tight zero head space extractor. In order to extract leachate fluid, a pressure of 50 psi is applied through the extractor and 50 ml of expelled fluid was collected. The superannuated leachate solution was filtered through Whatman 40 filter paper for analyzing trace elements using ECIL- AAS4141, atomic absorption spectrophotometer (AAS)(Fig. 2.11) .

FIGURE 2.11: Atomic absorption spectrophotometer setup (Departmental laboratory, IIT(BHU), Varanasi)

2.8 Artificial Neural Network

2.8.1 Introduction

Artificial neural network is a soft computing technique inspired by the biological nervous system. Human brain having millions of biological neurons connected with each other, which make it superior to distinguish infinite range of input patterns [225].

2.8.2 Biological Model of Neuron

Neuron(cell)is the fundamental processing unit of the biological nervous system which receive and process the signal (input)in the form of electrical stimulation through different dendrites. It consists of a nucleus, dendrites and a tubular axon [226]. The information reaching the dendrites of the neuron is sum up and then transmitted to its output path called axon, where the information is transmitted to other neurons though a junction called synapses if the combined signal is strong but if the combined signal is too low it will not transmitted to the next neuron . In former case, the neuron is said to be activated while in later case, it is said to be inhibited. The basic biological model of a neuron is shown in Fig.2.12.

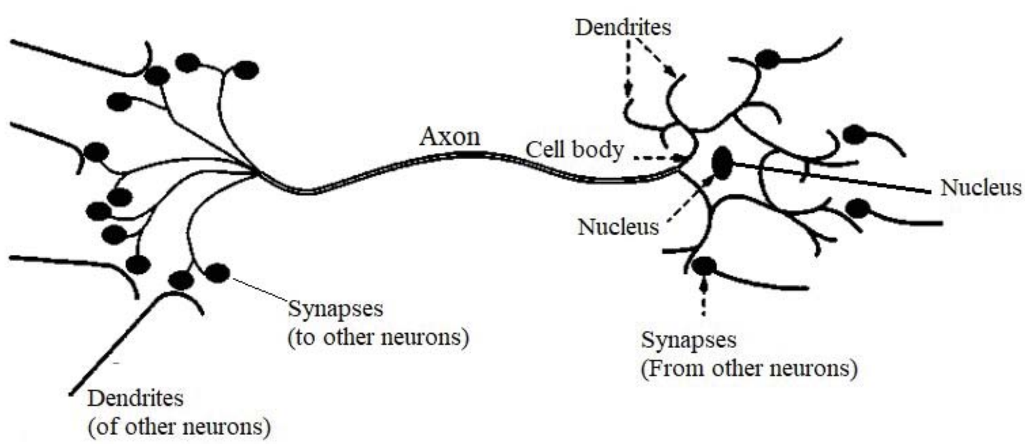

FIGURE 2.12: Biological model of a neuron

2.8.3 Mathematical Model of Neuron

In artificial neural network model, a computer program is framed to replicate to process the information in a manner that human brain performs, which improves its skill in identifying the pattern/ relations among the input and output parameters. It contains basic element such as input layer, hidden layer and its neuronal connections and output layer. Input layer receive information from the external environment. These information in the form of data, samples, signals, measurements or patterns are normally normalized within the limit values produced by activation/ transfer functions for better precision for the mathematical operations performed by the artificial neural network. Hidden layer are composed with neurons and extract the various pattern associated with the system to be analyzed. It also perform most of the internal processing from the network. Output layer is also composed of neurons and is responsible for producing final network outputs. Neuron is described as the processing element in the mathematical form of the artificial neural network model. It interacts with various input through weights and pass the information (as weighted input) to the next layer through transfer/ activation function. In artificial neural network input unit is passed/transmitted to the hidden neuron present in the next layer through connections. Further, the output of the hidden neuron is passed through a layer of activation function and repeats the same procedure and become the input of the next layer and finally it is passed through another transfer function to get the output of the network. Fig.2.13 shows the basic element of the artificial neural network. It contain n input units, k neurons in a hidden layers and p neurons

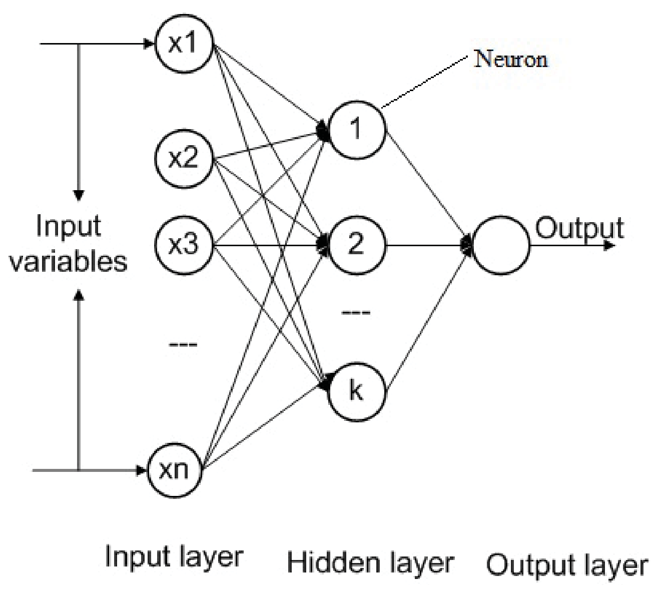

FIGURE 2.13: Basic artificial neural network architecture

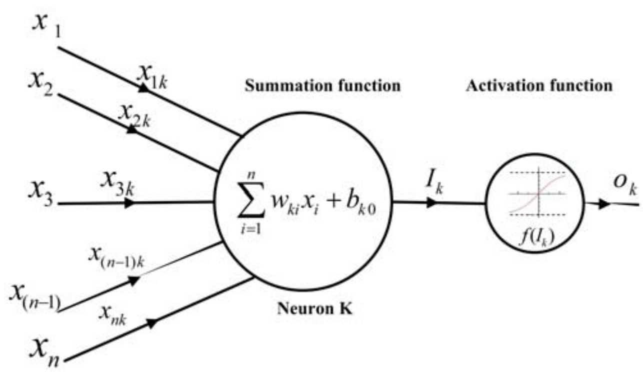

FIGURE 2.14: Mathematical model of neuron

in output layer.

Fig.2.14 represent the basic mathematical operations of a neuron present in the artificial neural network. Artificial neuron performs various mathematical task during the analysis which is presented as follows:

Step 1: The input unit is transmitted through connection that is multiplied by the Weight so the input value (S_m) of the hidden unit is given as:

$$S_m = \sum_{i=1}^{n} w_{mi}x_i + b_{m0} \tag{2.1}$$

Step 2: The input of hidden unit passed through properly selected transfer functions and becomes output (T_m) of the hidden units and produce:

$$T_m = f(S_m) = f(\sum_{i=1}^{n} w_{mi}x_i + b_{m0}) \tag{2.2}$$

Where, x_i (i = 1, 2, 3, – n) are the input units and w_{mi} and b_{m0} (m = 1, 2, 3, – k) are the weights and biases of the neurons in the hidden layer based on trained network.

Step 3: The output of hidden unit repeats the same process as in Step 1 and is given as and becomes input of the output layer (U_r) :

$$U_r = \sum_{m=1}^{k} w_{rm} T_m + b_{r0} \tag{2.3}$$

Step 4: The input of output unit is again passed though another transfer function and the final output is given as:

$$Y_r = f(U_r) = f(\sum_{m=1}^{k} w_{rm} T_m + b_{r0}) \tag{2.4}$$

Where, (Y_r) $(r = 1, 2, 3, - p)$

The above mathematical expression of the neural network will be used for developing the predictive equation in subsequent parts of the thesis.

2.8.4 Overview of Working Procedure of Artificial Neural Network

The successful execution of artificial neural network needs fundamental knowledge of all aspects related to the model development and depends on many factors thus a very brief and concise discussion has been presented herein.

(a) **Database preparation :** Artificial neural network based model requires sufficient numbers of data set for good training, testing and validation of network. It does not perform better in case of insufficient data set .There is still lack of concept and thumb rule regarding the minimal quantity of data sets required for training the neural networks, and it mainly depends upon the types of problem and quality data set [227].

(b) **Selection of input parameters :** Proper selection of input variables are the key steps in artificial neural network analysis . It affects the size and run time of the network. Large number of inputs increase the size of the network and thus increases the run time or decreases the processing speed of the computer. Various techniques such as cross-correlation analysis, Garson's approach, principal component analysis etc are used for selection of important input parameters.However,in civil engineering problem in general, trial and error approach is used to select the important or significant input parameters to the neural network analysis. In this approach different combinations of input variables are used for training and testing of the neural network and error of the corresponding testing data set as well as different performance indices between measured and predicted value are measured simultaneously. The combination of inputs with minimum error and better correlation are considered as significant inputs for the neural network.

(c) **Data proportioning and preprocessing :** Division of data set is the key steps for the successful execution of artificial neural network. Generally, it is divided into three subsets such as training, testing and validation set. Training set is used to train/learn the network based on the input and target supplied to the neural network. Once the training is completed, testing set is used to check the performance of the trained network. However, in some cases validation set is also used to avoid the over fitting of the network i.e., the ANN perform very well with training data set but not with other data set.

Preprocessing of data set is recommended for ANN application. It is also required to accommodate the different activation/transfer functions used in the ANN model. Basically, activation/transfer function used to conciliate non-linearity in the input output relationship. Generally, Sigmoid (logical and hyperbolic tangents) and non-sigmoids (polynomial, rational and Fourier series) are used as transfer functions in the artificial neural network. It has been observed that non-sigmoid function performs better when data sets were noiseless and contained highly non-linearity relationship whereas its performance is inferior when data sets have noise and contained mildly non-linearity relationship. However, sigmoid function performs always better in any cases either the data sets were noisy/noiseless or having mild or high non-linearity variations. If logistic sigmoid transfer function (Eq. 3.3) is used then data set is scaled in the range [0,1] whereas data set is scaled between -1 to 1 as hyperbolic tangent sigmoid (Eq. 3.7) types of activation function is used. Thus normalization/ scaling of data set primarily depends on the types of the transfer functions used .

$$x_{norm} = (\frac{x - x_{min}}{x_{max} - x_{min}}) \tag{2.5}$$

$$x_{norm} = 2(\frac{x - x_{min}}{x_{max} - x_{min}}) - 1 \tag{2.6}$$

Where, x, x_{min}, x_{max} and x_{norm} are the actual, minimum, maximum and normalized values of the data set respectively.

In geotechnical engineering generally sigmoid function: logistic sigmoid and hyperbolic tangent sigmoid and pure linear transfer function is used. Generally, same transfer function can be used at all layers. However, sigmoid transfer function between input and hidden layer and pure linear transfer function between hidden and output layer can be an advantage in case when it is necessary to extrapolate the value of data set outside the input range. Figure 2.15 shows the commonly used transfer functions used in the artificial neural network.

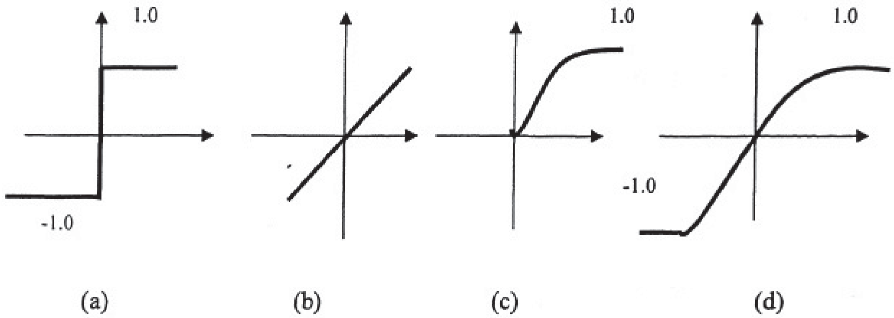

FIGURE 2.15: Different transfer function (a) stepped (b) linear (c) logistic sigmoid and (d) hyperbolic tangent sigmoid

(d) **Selection of network architecture :** The architecture (geometry) of artificial neural network defines how neurons are arranged/located, with regards to each other. These arrangements are essentially structured by managing the synaptic neuronal connections. As per the difference in architecture, ANN can be categorized as follows: (i) single-layer feedforward network, (ii) multilayer feedforward networks and (iii) recurrent/ feedback networks. In single-layer feedforward network, the information always flows in one direction i.e from input layer to output layer. These types of networks are primarily used in pattern classification and linear filtering problems. Multilayer feedforward networks (Multilayer Perceptron (MLP) and the Radial Basis Function (RBF)) has also similar architecture as single-layer feedforward network but it has one or more hidden neural layers. It is commonly applied for mapping nonlinear behavior of static processes such as pattern classification, system identification, optimization and robotics problems. In feedback networks, the outputs of the neurons are propagated backward and used as feedback inputs for neuron in hidden layers which updates the weights of the layers connections as per backpropagation learning algorithm. In this network architecture, first input and target data are fed into the ANN model, then the network process the inputs as per learning algorithms and generate a predicted output. After that error is calculated between its predicted and target outputs, then the outputs of the neurons are propagated backward and used as feedback inputs for neuron in hidden layers which updates the weights of the layers connections in ways that will reduce the error and the process continues until the network reaches a certain specified error.

(e) **Learning Process :** It is one of the important steps in the neural network model process. The learning process in artificial neural network refers to the ability of network to learn from their sample patterns and further improve the performance. In other words, one can say that learning is a process of non-linear optimization of error function or optimization of connection weight of the network. Supervised, unsupervised and reinforcement learning are the two commonly used learning techniques in ANN modeling. Supervised learning requires a table with input/output data whereas output is unknown in unsupervised learning process. In case of supervised learning, weight is optimized to minimized the error between predicted output and known target whereas unsupervised learning process doesn't requires any knowledge of any output sample and thus weight is optimized based on the other learning methods such as clustering, hebbian Learning,Self-organizing map adaptive resonance theory etc. In reinforcement learning technique, system is trained for a particular job, learns based on its previous experiences and determine the ideal behavior to maximize their performance. At the end of the training phase,optimized weight of the trained network are stored in the memory of the ANN model. In next phase (testing phase), the trained network is nourished with new set of data and the predicted value is compared with the target value to assess the capability of the ANN model for the range of the data used for the training process. Finally, once training, testing and validation process complete, the corresponding neural network model can be used for practical application in that particular area of study.

(f) **Choice of Performance Criteria :** It's an important step and used to check the efficacy of the proposed ANN model. Generally, coefficient of correlation is used to compare the predictability of the ANN model. However, it is not sufficient to check the efficacy of the predictive model as it gives biased estimate sometimes. So in addition of this, other statistical measure need to used as performance indices in order to test the effectiveness of the predictive model. Details of various performance indices are discussed in the next sub section.

2.9 Performance Analysis

Several statistical parameters such as maximum absolute error (MAE), average, absolute error (AAE), mean square error (MSE), root mean square error ($RMSE$), mean absolute percentage error ($MAPE$) and coefficient of determination (R^2) are used to evaluate the prediction capability of the proposed models and are expressed as follows (Eqs. 3.15-3.17).

$$R^2 = 1 - \frac{\sum_i (o_{pi} - o_i)^2}{\sum_i (o_{pi} - \overline{o}_i)^2} \tag{2.7}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (o_i - o_{pi})^2 \tag{2.8}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - o_{pi})^2} \tag{2.9}$$

$$AAE = \frac{1}{n}\sum_{i=1}^{n}|o_i - o_{pi}| \tag{2.10}$$

$$MAE = max\left(|o_i - o_{pi}|\right) \tag{2.11}$$

$$MAPE = \left[\frac{1}{n}\sum_{i=1}^{n}\left|\frac{o_i - o_{pi}}{o_i}\right|\right] \times 100 \tag{2.12}$$

Where, o and o_p : measured and predicted output respectively, $\overline{o_p}$: average of the predicted output and n : number of observations.

Additionally, scaled percent error versus cumulative frequency and Taylor's plot can also be considered to verify the predictive models.

2.9.1 Scaled Percent Error and Cumulative Frequency

It gives the idea of the frequency of error in a particular class with the help of scaled percent error versus cumulative frequency plot. Merits of scaled percent error (SPE) versus cumulative frequency (CF) plot as performance index has been discussed by researchers ([228, 229]). It can be plotted by adopting simple steps as follows:

- Calculate the scaled percent error (SPE) of measured and predicted data set using the following expression and plot on x-axis.

$$SPE = \left[\frac{(o_p - o_m)}{([o_m]_{\max} - [o_m]_{\min})}\right] \times 100 \tag{2.13}$$

  Where, o_p and o_m are the predicted and measured values and $[o_m]_{\max}$ and $[o_m]_{\min}$ are the maximum and the minimum measured value respectively.

- Arrange in increasing order and find out the maximum, the minimum value and frequency of SPE.

- Determine the value of cumulative frequency and plot on y-axis. Frequency of an error in a data set refers to the number of times that error appears in the particular selected range; whereas, the sum of all previous frequencies of error up to that point is taken as the cumulative frequency and expressed in percentage.

2.9.2 Taylor's Plot

Taylor's plot is also an alternate way to observe the goodness of the predictive model [230]. It uses three performance indices namely correlation coefficient (R^2) as black line, the centered root error mean square ($RMSE$) (green lines contours) and the standard deviation (in blue line contour) to draw the plot (Fig. 2.16). The point (square marker) marks on the x-axis of the plot represents the reference point, whereas circular and triangular marker represents the model point based on predictive model data set obtained from observed and predicted data set respectively. The marker close to the reference point depicts the ideal fit of that model. It also

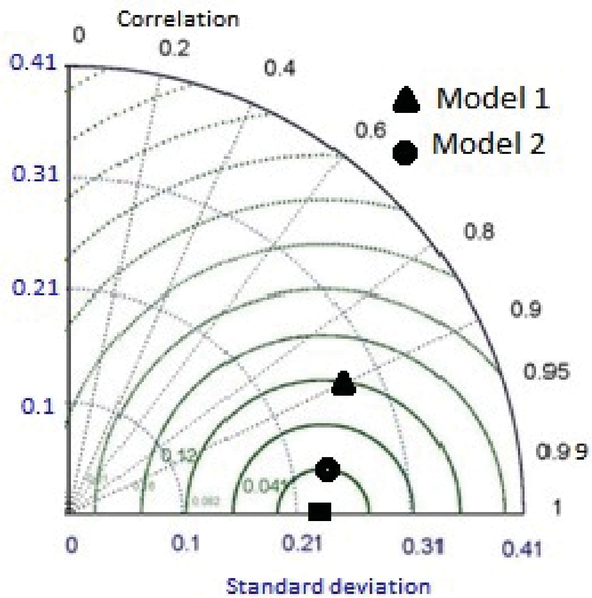

FIGURE 2.16: Taylor's plot

gives insight about the over or under estimation of the predictive model. If the standard deviation (SD) of the predicted values is found to be more or less than that of computed values then it shows over or under estimation of the predictive model.

2.10 Concluding Remarks

The chapter presents the details of materials, experimental procedures, apparatus used and the basic of different methodologies adopted. The results of the above studies have been presented and discussed in detail in **Chapters 3** and **4**.