A Project report on

Machine Learning Techniques to Predict the Compressive Strength of Various Concrete Mixtures

Submitted in partial fulfillment of the requirements
for the award of the degree of

BACHELOR OF TECHNOLOGY

in

Civil Engineering

by

S. Shakeer Hussain	**184G1A0167**
P. Sravani	**184G1A0173**
G. L. Sri Shakthi	**184G1A0176**
Y. Vamsi Krishna	**184G1A0185**
B. Moses Lee Mories Raj	**194G5A0115**
S. Mahammad Shaluddin	**194G5A0125**

Under the Guidance of

Dr. Raghu Babu Uppara **M.Tech, Ph.D**
Associate Professor

DEPARTMENT OF CIVIL ENGINEERING
SRINIVASA RAMANUJAN INSTITUTE OF TECHNOLOGY
ANANTHAPURAMU
(Affiliated to JNTUA, Approved by AICTE, New Delhi)

2021-2022

SRINIVASA RAMANUJAN INSTITUTE OF TECHNOLOGY
ANANTHAPURAMU
(Affiliated to JNTUA, Approved by AICTE, New Delhi)

Certificate

This is to certify that the project report entitled Machine Learning Techniques To Predict The Compressive Strength Of Various Concrete Mixtures. is the bonafide work carried out by **G.L. Sri Shakthi** bearing Roll Number 184G1A0176 in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **Department of Civil Engineering** during the academic year 2021-2022.

Guide	**Head Of Department**
Dr. Raghu Babu Uppara M.Tech, Ph.D	D. Lakshmi Sireesha M.Tech
Associate Professor	Assistant Professor

Date:

Place: **EXTERNAL EXAMINER**

SRINIVASA RAMANUJAN INSTITUTE OF TECHNOLOGY
ANANTHAPURAMU

(Affiliated to JNTUA, Approved by AICTE, New Delhi)

SR IT
Empowering Knowledge

Certificate

This is to certify that the project report entitled Machine Learning Techniques To Predict The Compressive Strength Of Various Concrete Mixtures. is the bonafide work carried out by **S. Shakeer Hussain** bearing Roll Number 184G1A0167, **P. Sravani** bearing Roll Number 184G1A0173, **G.L. Sri Shakthi** bearing Roll Number 184G1A0176, **Y. Vamsi Krishna** bearing Roll Number 184G1A0185, **B. Moses Lee Mories Raj** bearing Roll Number 194G5A0115 and **S. Mahammad Shaluddin** bearing Roll Number 194G5A0125 in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **Department of Civil Engineering** during the academic year 2021-2022.

Guide	**Head Of Department**
Dr. Raghu Babu Uppara M.Tech, Ph.D	D. Lakshmi Sireesha M.Tech
Associate Professor	Assistant Professor

Date:

Place: **EXTERNAL EXAMINER**

DECLARATION

We S. Shakeer Hussain bearing reg no: 184G1A0167, P. Sravani bearing reg no: 184G1A0173, G. L. Sri Shakthi bearing reg no: 184G1A0176, Y. Vamsi Krishna bearing reg no: 184G1A0185, B. Moses Lee Mories Raj bearing reg no: 194G5A0115, S. Mahammad Shaluddin bearing reg no : 194G5A0125 , students of SRINIVASA RAMANUJAN INSTITUTE OF TECHNOLOGY, Rotarypuram, hereby declare that the dissertation entitled "MACHINE LEARNING TECHNIQUES TO PREDICT THE COMPRESSIVE STRENGTH OF VARIOUS CONCRETE MIXTURES" embodies the report of our project work carried out by us during IV Year Bachelor of Technology under the guidance of Dr. Raghu Babu Uppara, Department of Civil Engineering and this work has been submitted for the partial fulfilment of the requirements for the award of Bachelor of Technology degree.

The results embodied in this project report have not been submitted to any other Universities of Institute for the award of Degree.

S. Shakeer Hussain	Reg no: 184G1A0167
P. Sravani	Reg no: 184G1A0173
G.L. Sri Shakthi	Reg no: 184G1A0176
Y. Vamsi Krishna	Reg no: 184G1A0185
B. Moses Lee Mories Raj	Reg no: 194G5A0115
S. Mahammad Shaluddin	Reg no: 194G5A0125

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of people who made it possible, whose constant guidance and encouragement crowned our efforts with success. It is a pleasant aspect that we have now the opportunity to express my gratitude for all of them.

It is with immense pleasure that we would like to express my indebted gratitude to my Guide **Dr. Raghu Babu Uppara,** Associate Professor, Department of Civil Engineering who has guided me a lot and encouraged me in every step of the project work. We thank him for the stimulating guidance, constant encouragement and constructive criticism which have made possible to bring out this project work.

We are very much thankful to **Mrs. D. Lakshmi Sireesha,** Assistant Professor, Department of Civil Engineering**,** for her kind support and for providing necessary facilities to carry out the work.

We wish to convey my special thanks to **Dr. G. Bala Krishna,** Principal of Srinivasa Ramanujan Institute of Technology for giving the required information in doing my project work. Not to forget, we thank all other faculty and non-teaching staff, and my friends who had directly or indirectly helped and supported me in completing my project in time.

We also express our sincere thanks to the Management for providing excellent facilities**.**

Finally, we wish to convey our gratitude to our family who fostered all the requirements and facilities that we need.

Project Associates

Abstract

Although the mechanical properties of concrete are very important in the construction industry, it can take up to 28 days to test and is also labour-intensive. In addition, it costs a lot of money to test the mechanical properties of concrete. Furthermore, concrete made from pozzolanic materials such as metakaolin, fly ash, silica fume, etc., takes longer to gain full strength compared to conventional concrete made with ordinary Portland cement. The current work proposes prediction model for compressive strength to overcome problems due to such time consuming and laborious testing of concrete strength. Potential Machine learning technique such as classification and regression trees (CART), Artificial Neural Network (ANN) and Regression Models have been used to predict the compressive strength of supplementary Cementitious Materials and recycled aggregate concrete. For this purpose, seven predictors namely cement, blender, fine aggregate, coarse aggregate, water, super plasticizer and curing age of concrete are considered. The performance of the predicted model has been evaluated using R-squared value (R^2), root mean squared error (RMSE), mean squared error (MSE), and mean absolute percent error (MAPE). Further the performance of the tree with given predictors has been validated with 10-fold cross-validation technique. The current study confirms that the Machine Learning models are recommended for predicting the compressive strength of blended concrete and recycled aggregate concrete at a given curing period prior to conducting laboratory testing.

Contents

List of Tables

List of Figures

List of Graphs

List of Equation

Abbreviations

For the purpose of this standard project, the following letter symbols shall have the meaning indicated against each; where other symbols are used, they are explained at the appropriate place:

ANN	Artificial neural network
B	Binder
BFA	Bottom fly ash
C	Cement
CA	Coarse aggregates
CART	Classification and regression tree
D MAX	Maximum size of aggregates
D MIN	Minimum size of aggregates
DT	Decision trees
DWT	Discrete wavelet transforms
FA	Fine aggregates
FL	Fuzzy logic
GA	Genetic algorithm
GEP	Gene expression programming
GGBS	Ground granulated blast furnace slag
GOT	Genetic operation trees
GWPOT	Genetic weighted pyramid operation tree
HPC	High performance concrete
HSC	High strength concrete
LP	Limestone powder
LR	Linear regression
MK	Metakaolin
MLR	Multi-logistic model
MP	Marble powders
MS	Micro silica
MLR	Multiple nonlinear regression
NS	Nano silica

NSC	Non-slump concrete
OPC	Ordinary Portland concrete
P	Powder
PFA	Pulverized fuel ash
RF	Random forest
RHA	Rice husk ash
SA	Specific area
SCC	Self-consolidating concrete
SCM	Supplementary cementitious material
SF	Silica fume
SP	Super plasticizer
SR	Surface resistivity
SVM	Support vector machine
VMA	Viscosity modify admixture
W	Water
WRA	Water reducing agent
WSVM	Weighted support vector machines
Z	Zeolite
WC	Water Content
F.A	Flyash
T°	Temperature
W/B	water / Binder
NCA	Natural Coarse Aggregate
NFA	Natural Fine Aggregate
RCA	Recycled Coarse Aggregate
RFA	Recycled Fine Aggregate
FRCP	Fine Recycled Concrete Powder

1 Introduction

Determining the compressive strength of concrete is of the highest priority over other mechanical properties of concrete. It usually takes a lot of time and effort to test the compressive strength of concrete in a laboratory. It takes an average of 28 days to determine the compressive strength of the given concrete mixture in the laboratory.

Generally, compressive strength may vary depending on the type of cement, water to cement ratio, aggregate size, age of concrete, specimen size, shape, and chemical admixtures etc. Hence, the prediction of compressive strength of given concrete mixture prior to the construction is needed extensively.

Cement and cement concrete are common construction materials manufactured in large amounts. The demand for cement has increasing for many years which is an effect of global urbanization. In recent years above 90% increase in cement production was observed.

Unfortunately, production of cement clinker is not neutral for environment as it consumes large amount of energy and natural resources and causes emission of CO_2 which is estimated as 5–8% of the annual worldwide CO_2 emissions from anthropogenic sources (about 842 kg CO_2 per tonne of clinker, 600–700 kg CO2 per ton of cement). The greenhouse gas emission covers mainly CO2 arising from high-temperature decomposition processes taking place during cement clinker production. One of the ways, to reduce the CO2 emission from the manufacturing of cement is the utilization of Supplementary Cementitous Materials (SCMs).

1.1 Supplementary Cementitious materials

The sustainability aspects of cement is increasingly gaining interest as the demand for concrete infrastructure continues to rise on a global level. The high environmental impacts associated with cement production have led to the development of alternative binders involving waste materials and industrial byproducts. Some of the most frequently used SCMs to reduce the overall CO_2 emissions of cement-based construction materials include Fly Ash, ground granulated blast furnace slag (GGBS), Metakaolin (MK), Silica Fume (SF) Compared to Plain Cement, whose production leads to CO2 emissions of ~0.85 tonne/tonne (t/t), the production of SCM results in the emissions of ~0.009 and~0.02 t/t, respectively. The partial replacement of cement with these SCMs does not only present environmental advantages, but can also lead to strength gain via the reaction between the siliceous and aluminous phases in the SCMs with the calcium phases in cement. Furthermore, the use of SCMs can also contribute to strength development by enabling a reduction in the sample porosity through their micro filler effect. The inclusion of these SCMs in concrete could also greatly reduce the thermal conductivity of normal and lightweight concrete due to their lower densities, increasing the application areas in which these concrete formulations can be incorporated. In addition to the use of SCMs in cement-based mixes, alternative binders proposing lower environmental impacts have gained interest in recent years, both on local and global levels. For instance, in Singapore, the use of sustainable construction materials is introduced with the Green Mark scheme, which promotes green building design and technologies that improve energy efficiency and reduce the impact of buildings on the environment.

1.1.1 Metakaolin

Metakaolin is a pozzolanic additive/product which can provide many specific features. Metakaolin is available in many different varieties and qualities. Some of them also provide special reactivity. Metakaolin is a valuable admixture for concrete and or cement applications. Usually 8% - 20% (by weight) of Portland cement replaced by metakaolin. Such a concrete exhibits favourable engineering property. The pozzolanic reaction starts soon and continues between 7 to 28 days.

Metakaolin is not a by-product. It is obtained by the calcinations of pure or refined Kaolinite clay at a temperature between 650 °C and 850°C, followed by grinding to achieve a finesse of 700-900 m^2/kg.

1.1.2 Silica Fume

Silica fume is a byproduct of producing silicon metal or ferrosilicon alloys. One of the most beneficial uses for silica fume is in concrete. Because of its chemical and physical properties, it is a very reactive pozzolan. Concrete containing silica fume can have very high strength and can be very durable. Silica fume is available from suppliers of concrete admixtures and, when specified, is simply added during concrete production. Placing, finishing, and curing silica-fume concrete require special attention on the part of the concrete contractor.

Silica fume for use in concrete is available in wet or dry forms. It is usually added during concrete production at a concrete plant. Silica fume-concrete has been successfully produced in both central-mix and dry-batch plants. Assistance is readily available on all aspects of handling silica fume and using it to produce consistent, high-quality concrete.

1.1.3 Fly Ash

Fly Ash is the most widely used SCM in concrete and is a byproduct of coal combustion in electric power generating plants. The use of fly ash in concrete can contribute to LEED points through local materials, recycled contents and innovation credits. Fly ash can compensate for fine materials that may be lacking in sand quantities and can be very beneficial in improving the flowability and finish ability of concrete mixtures. The two designations for fly ash used in concrete are Class C and F and are described in ASTM C618.

Class C Ash: high calcium contents with low carbon and good pozzolanic and cementitious properties lend this material to use in higher performance mixtures where early age strength is important.

Class F Ash: low calcium ash effectively moderating heat gain during concrete curing and therefore ideal for mass placement conditions and high strength mixtures or use in hot weather climates; Also provides good sulfide and sulfate resistance to concrete through same capacity as Type V (CSA Type 50) cement.

1.1.4 Ground Granulated Blast Furnace Slag

Ground-granulated blast-furnace slag (GGBS or GGBFS) is obtained by quenching molten iron slag (a by-product of iron and steel-making) from a blast furnace in water or steam, to produce a glassy, granular product that is then dried and ground into a fine powder. Ground-granulated blast furnace slag is highly cementitious and high in calcium silicate hydrates (CSH) which is a strength enhancing compound which improves the strength, durability and appearance of the concrete.

Blast-furnaces operate at temperatures of about 1,500°C and are fed with a carefully controlled mixture of iron ore, coke and limestone. The iron ore is reduced to iron and the remaining materials form a slag that floats on top of the iron.

This slag is periodically tapped off as a molten liquid and if it is to be used for the manufacture of GGBS it has to be rapidly quenched in large volumes of water. The quenching optimizes the cementitious properties and produces granules similar to a coarse sand. This ‘granulated’ slag is then dried and ground to a fine powder.

1.2 Recycled Aggregates

For a variety of reasons, reuse of construction and demolition (C&D) materials by the construction industry has become more significant. In addition to environmental protection, conservation of natural aggregate resources, shortage of waste disposal land, and increasing cost of waste treatment prior to disposal are the main reasons for the growing interest in recycling C&D materials. Already many countries have introduced legislation and policy measures to encourage the use of recycled aggregates in civil engineering works. The potential benefits and drawbacks of using recycled aggregate in new concrete have been extensively studied. Although much information is already available on the effect of recycled aggregate on the mechanical properties of the concrete up to a curing age of 90 days, the literature contains only a limited number of research results on the long-term mechanical properties of recycled aggregate concrete. The porosity of recycled aggregate (RA) concrete is generally higher than that of natural aggregate (NA) concrete due to the adhered mortar present in recycled aggregates. The porosity and the pore size distribution are the most important characteristics of the pore system of concrete which influence the ingress of foreign substances to the interior of concrete. It is therefore important to understand the development on the pore system in order to assess the durability properties of the recycled aggregate concrete. The test results of recycled aggregate concrete are decreased compared to conventional concrete (35 MPa mix).

In recent years, some researchers have tried to evaluate the potential of using recycled concrete aggregates as a replacement for natural aggregates in the concrete. Self-compacting concrete (SCC), as one of the most significant advances in the concrete industry, exhibits a better performance than that of conventional concrete. This may be attributed to the association of supplementary cementitious materials (SCM) and filler materials that are considered at nuclear sites and to refine the porosity of the cement paste and reduce permeability. In fact, filler materials are commonly used as additives in SCC to enhance strength and long-term properties. Recycled aggregates have been successfully used and their performance was extensively investigated by several researchers to develop self-compacting concrete.

2 Literature Review

Many research studies have been conducted using artificial intelligence (AI) to determine the effect of SCMs such as fly ash, blast furnace slag, metakaolin, silica fume, and nano-silica on the mechanical properties of concrete. However, a comprehensive study is essential to evaluate the effect of the characteristics of the used SCM on the mechanical properties of concrete. Although it has been shown that the properties of SCMs influence their effects on various properties of concrete, limited research has been done to identify the extent to which each character is influential and their relative significance. The studies on mechanical properties of different concrete containing various SCMs using AI are summarized in Table 2. In this table, the input is the features that are selected to train the AI models. It can be seen that the majority of these researches developed the prediction models based on the amount of the main ingredients of concrete and replacement level of SCMs.

Classical statistical methods have been applied in industry for years. Recently, Neural Networks (NNs) methods have become tools of choice for a wide variety of applications across many disciplines. It has been recognized in the literature that regression and neural network methods have become competing model-building methods. For a large class of pattern-recognition processes, NNs is the preferred technique. NNs methods have also been used in the areas of prediction and classification. Since NNs was developed as generalizations of mathematical models of human cognition through biological neurons, it is regarded as an information processing system that has certain performance characteristics in common with human neural biology. The characteristics include ability for storing knowledge and making it available for use whenever necessary, propensity to identify patterns, even in the presence of noise, aptitude for taking past experiences into consideration and make inferences and judgments about new situations. Statistical methods such as regression analysis, multivariate analysis, Bayesian theory, pattern recognition and least square approximation models have been applied to a wide range of decisions in many disciplines. These models are attractive to decision makers because of their established methodology, long history of application, availability of software and deep-rooted acceptance among practitioners and academicians alike. NNs are data dependent and therefore, their performance improves with sample size. Statistical methods, such as Regression perform better for extremely small sample size, and also when theory or experience indicates an underlying relationship between dependent and predictor variables. Classification and Regression Tree (CART) models use tree-building algorithms, which are a set of if-then (split) conditions that permit prediction or classification of cases. A CART model that predicts the value of continuous variables from a set of continuous and/or categorical predictor variables is referred as regression-type model. For the prediction of the value of categorical variable from a set of continuous and/or categorical predictor variables, classification-type CART model is used. One noticeable advantage of decision tree-based models, such as CART, is that the decision tree-based models are scalable to large problems and can handle smaller data set than NNs models. Despite the apparent substantive and applied advantages of statistical models, Neural Networks (NNs) methods have also gained popularity in recent years. These methods are particularly valuable when the functional relationship between independent and dependent variables are unknown and there are ample training and test data available for the process. NNs models also have high tolerance for noise in the data and complexity. Moreover, the software technologies, such as, SPSS and Minitab that deploy neural networks algorithm have become extremely sophisticated and user-friendly in recent years.

Table 1. Literature Review

Name	Network type	Concrete type	Pozzolan	Input	Ref.
Kasperkiewicz et al.	ANN	HPC	SF	C, W, CA, FA, SF, SP	[1]
Sobhani et al.	ANN and ANFIS	NSC	SF	C, W, CA, FA, W/B, SF, Filler	[2]
Behnood and Golafshani	ANN	OPC	SF	B, W/B, SF/B, CA/(CA+FA), CA/B, SP/B, Dmax, Age	[3]
Serraye et al.	ANN	SCC	SF	B, W/B, SF, CA, FA, SP, Age	[4]
Sebastia et al.	ANN	OPC	FA	C, W, Fa, (CA+FA), SP	[5]
Yeh	ANN	OPC	FA	C, W, CA, FA, Fa, GGBS, SP	[6]
Topcu and Saridemir	ANN and FL	OPC	FA	C, W, CA, FA, WRA, Fa, CaO	[7]
Prasad et al.	ANN	SCC AND HPC	FA	C, W/C, W/P, CA/P, FA/P, WRA/P, VMA/P,Fa/B,MS/B	[8]
Baykasoglu et al.	GEP	HSC	FA	W/B, W, FA, Fa, A, SP	[9]
Siddique et al.	ANN	SCC	FA	C, W, CA, FA, Fa, BFa, W/P, SP	[10]
Cheng et al.	GWPOT	HPC	FA	C, W, CA, FA, Fa, GGBS, SP, Age	[11]
Aggarwal et al.	FL and ANN	HSC	FA	C, W, CA, FA, Fa, SF, SP, fiber, aspect ratio	[12]
Azimi-Pour et al.	SVM	SCC	FA	C, W/C, W/P, W/B, CA/P, FA/P, WRA/P, SP/P, Fa/B, MS/B	[13]
Roshani et al.	ANN	OPC	FA	C, W, CA, FA, Fa, SiO_2	[14]
Bilim et al.	ANN	OPC	GGBS	C, W, (CA+FA), GGBS, SP, Age	[15]
Sarıdemir	FL and ANN	OPC	GGBS	C, W, (CA+FA), GGBS, Age	[16]

Kandiri et al.	ANN	OPC	GGBS	C, W, CA, FA, GGBS, Activity index, Age	[17]
Gilan et al.	ANFIS	OPC	MK	C, W, MK, CA, FA, Age, SR	[18]
Yeh	ANN	HPC	FA & GGBS	C, W, CA, FA, Fa, GGBS, SP	[19]
Yeh and Lien	GOT	HPC	FA & GGBS	C, W, CA, FA, Fa, GGBS, SP, Age, W/C, W/B, W/(C+Fa+SL+CA+FA), (CA+FA)/B	[20]
Chou et al.	ANN	HPC	FA&GGBS	C, W, CA, FA, Fa, GGBS, SP, Age	[21]
Cheng et al.	wSVM	HPC	FA & GGBS	C, W, CA, FA, Fa, GGBS, SP, Age	[11]
Mousavi et al.	GEP	HPC	FA & GGBS	C, W, CA, FA, Fa, GGBS, SP, Age	[22]
Erdal	DT	HPC	FA & GGBS	C, W, CA, FA, Fa, GGBS, SP, Age	[23]
Erdal et al.	DWT	HPC	FA & GGBS	C, W, CA, FA, Fa, GGBS, SP, Age	[24]
Yu et al.	SVM	HPC	FA & GGBS	C, W, CA, FA, Fa, GGBS, SP, Age	[25]
Bui et al.	ANN	HPC	FA & GGBS	C, W, CA, FA, Fa, GGBS, SP, Age	[17]
Han et al.	RF	HPC	FA & GGBS	C, W, CA, FA, Fa, GGBS, SP, Age	[26]
Golafshani et al.	ANN and ANFIS	OPC & HPC	FA & GGBS	C, W, CA, FA, GGBS, Fa, SP, Age	[13]
Pala et al.	ANN	OPC	FA & SF	B, W, CA, FA, Fa, SF, WRA, Age	[27]
Ahmadi-Nedushan	ANN	HPC	FA & SF	W/B, W, FA/(CA+FA), Fa/B, A, SP	[28]
Zhang and Zhao	ANN	UHPC	FA & SF	C, W, CA, FA, Fa, SF, SP, W/C, (FA+CA)/C, FA/(FA+CA), Dmax-Dmin	[29]
Sarıdemir	ANN	OPC	MK & SF	C, W, CA, FA, MK, SF, SP, Age	[30]
Uysal and Tanyildizi	ANN	SCC	FA & LP	C, Fa, LP, MP, (CA+FA), SP, Unit weight, Water absorption	[31]
Khan	ANN	HPC	PFA & SF	C, W, CA, FA, PFA, Fa, SP, Age	[32]
Khan	ANN	HPC	MS & SF	C, W, CA, FA, MS, Fa, SP, Age	[33]
Chou et al.	ANN, SVM, CART, LR	HPC	GGBS, FA, SF, MK	(C, W, CA, FA, Fa, GGBS, SP, Age) (C, W, CA, FA, SF, SP, WRA, A, Age) (B, W, CA, FA, Fa, SF, WRA, Age) (B, W, FA, Fa, SP, A) (C, W, CA, FA, MK, Age)	[34]
Rebouh et al.	GA and ANN	OPC	NATURAL POZZOLAN	B, W/B, Pozzolan ratio, SP, CA, FA, Age	[35]
Nasr et al.	ANN	SCC	NS	C, Z, NS, slump flow, V-funnel flow, Age	[36]
Elemam et al.	ANN	SCC	FA, SF & LP	P, Fa, SF, LP, W/P, SP	[37]
Iqtidar et al.	ANN, ANFIS, NLR	OPC	RHA	C, W, CA, FA, RHA, SP, Age	[38]
Faraj et al.	LR, NLR, MLR	SCC	NS	B, W/B, NS, CA, FA, SP, Age	[39]

3 Methodology

3.1 Data Collection

Experimental investigations on concrete containing Metakaolin, Flyash, Silica Fume and Ground Granulated Blast Furnace Slag for cement replacement were carried out by several researchers to determine their properties, in the current study the experimental data gathered from earlier literatures and is used to develop prediction models at different age of concrete.

In this chapter the details of experimental data collected for developing the prediction models are discussed along with the methodology adopted in the present study.

3.1.1 Data for Metakaolin blended Concrete

Table 2. Metakaolin Data

Name	Concrete type	Pozzolan	Input	Output	Data points	Ref
O. O. Akin et al.	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	30 P	[40]
M. Saridemir et al.	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	64 P	[41]
C. S. Poon et al.	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	24 P	[42]
H. S. Wong et al.	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	56 P	[43]
P. Muthupriya et al.	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	40 P	[44]
R. M. Ferreira et al.	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	55 P	[45]
S. Wild et al.	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	35 P	[46]
K. B. Park et al.	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	24 P	[47]
S. Safarzadegan Gilan et al.	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	99 P	[18]
A. A Ramezanianpour et al.	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	1 P	[48]

3.1.2 Data for Silica Fume blended Concrete

Table 3. Silica Fume Data

Name	Concrete type	Pozzolan	Input	Output	Data points	Ref
W. Wongkeo et al.	OPC	SF	W, C, FA, CA, F.A, SP, SF, AGE, MK	CS	144 P	[58]
M. Saridemir et al.	OPC	SF	W, C, FA, CA, F.A, SP, SF, AGE, MK	CS	65 P	[41]
C. S. Poon et al.	OPC	SF	W, C, FA, CA, F.A, SP, SF, AGE, MK	CS	32 P	[42]
H. S. Wong et al.	OPC	SF	W, C, FA, CA, F.A, SP, SF, AGE, MK	CS	147 P	[43]
P. Muthupriya et al.	OPC	SF	W, C, FA, CA, F.A, SP, SF, AGE, MK	CS	70 P	[44]
T. C. Nwofor et al.	OPC	SF	W, C, FA, CA, F.A, SP, SF, AGE, MK	CS	21 P	[59]
M. Pala et al.	OPC	SF	W, C, FA, CA, F.A, SP, SF, AGE, MK	CS	144 P	[27]

3.1.3 Data for Fly ash blended Concrete

Table 4. Fly Ash Data

Name	Concrete type	Pozzolan	Input	Output	Data points	Ref
A. Oner et al.	OPC	FA	C, F.A, CA, FA, W, SP, AGE	CS	56 P	[60]
M. Pala et al.	OPC	FA	C, F.A, CA, FA, W, SP, AGE	CS	90P, 61P	[27]
W. Wongkeo et al.	OPC	FA	C, F.A, CA, FA, W, SP, AGE	CS	48 P	[58]
	OPC	FA	C, F.A, CA, FA, W, SP, AGE	CS	RAW DATA	

3.1.4 Data for Ground Granulated Blast Furnace Slag Blended Concrete

Table 5. GGBS Data

Name	Concrete type	Pozzolan	Input	Output	Data points	Ref
C. Bilim et al.	OPC	GGBS	T^0, C, B, W, W/B, GBBS/B, GGBS, FA, CA, SP, AGE, FA, SF	CS	225 P	[15]
I. J. Han et al.	OPC	GGBS	T^0, C, B, W, W/B, GBBS/B, GGBS,	CS	268 P, 194 P	[61]

			FA, CA, SP, AGE, FA, SF			
C. Duran Atis et al.	OPC	GGBS	T^0, C, B, W, W/B, GBBS/B, GGBS, FA, CA, SP, AGE, FA, SF	CS	45 P	[62]
	OPC	GGBS	T^0, C, B, W, W/B, GBBS/B, GGBS, FA, CA, SP, AGE, FA, SF	CS	83 P	
A. C. Ganesh et al.	OPC	GGBS	T^0, C, B, W, W/B, GBBS/B, GGBS, FA, CA, SP, AGE, FA, SF	CS	44 P	[63]
A. Oner et al	OPC	GGBS	T^0, C, B, W, W/B, GBBS/B, GGBS, FA, CA, SP, AGE, FA, SF	CS	224 P	[64]

3.1.5 Data for Recycled Aggregate Concrete Blended Concrete

Table 6. Recycled Aggregate Data

Name	Concrete type	Pozzolan	Input	Output	Data points	Ref
C. S. Poon et al.	OPC	RA	C, W, NCA, RCA, NFA, RFA, F.A, FRCP, SP, AGE	CS	36 P	[49]
S. C. Kou et al.	OPC	RA	C, W, NCA, RCA, NFA, RFA, F.A, FRCP, SP, AGE	CS	50 P	[50]
C. S. Poon et al.	OPC	RA	C, W, NCA, RCA, NFA, RFA, F.A, FRCP, SP, AGE	CS	28 P	[51]
R. Kumutha et al.	OPC	RA	C, W, NCA, RCA, NFA, RFA, F.A, FRCP, SP, AGE	CS	22 P	[52]
S. C. Kou et al.	OPC	RA	C, W, NCA, RCA, NFA, RFA, F.A, FRCP, SP, AGE	CS	9 P	[53]
F. Lopez Gayarre et al.	OPC	RA	C, W, NCA, RCA, NFA, RFA, F.A, FRCP, SP, AGE	CS	28 P	[54]
S. R. Salimbahrami et al.	OPC	RA	C, W, NCA, RCA, NFA, RFA, F.A, FRCP, SP, AGE	CS	12 P	[55]
G. Andreu et al.	OPC	RA	C, W, NCA, RCA, NFA, RFA, F.A,	CS	28 P	[56]

			FRCP, SP, AGE			
S. Baudali et al.	OPC	RA	C, W, NCA, RCA, NFA, RFA, F.A, FRCP, SP, AGE	CS	115 P	[57]

3.2 Machine Learning Techniques adopted to develop Prediction models

In order to develop prediction models for compressive strength the machine learning techniques namely artificial neural network, Decision tree and Regression models have been adopted. For the development of models, the following dependent and independent variables have been considered.

Table 7. List of variables for model development

S.No	System	Dependent Variable	Independent Variable
1.	Metakaolin blended Concrete	Compressive Strength	A, C, MK, FA, CA, WC, SP
2.	Silica Fume blended Concrete	Compressive Strength	W, C, FA, CA, FLYASH, SP, SF, A, MK
3.	Flyash blended Concrete	Compressive Strength	C, FLYASH, CA, FA, W, SP, DAYS
4.	Ground Granulated Blast Furnace Slag Blended Concrete	Compressive Strength	T, C, BC, W, W/B, G/B, G, FA, CA, SP, A, FLYASH, SF, SH, SS
5.	Recycled Aggregate Mixture	Compressive Strength	C, W, NCA, RCA, NFA, RFA, FLYASH, FRCP, SP, A

3.2.1 Artificial Neural Network

Artificial Neural Networks are similar computational models that mimic the human brain functions. Artificial neural networks are computing systems designed with simple, highly interconnected processing units that process information through the dynamic state of response to a given input. An artificial neural network is usually made up of layers of neurons, which are the core processing units of a network. First, we have the input layer that receives the input units, and the output layer predicts the final output. In the middle are hidden layers that perform most of the calculations required for the network.

Input data is provided as input to each neuron in the input layer, connected to the next layer by neuron channels in one layer, and each of these channels is assigned a numeric value called the weight (linear combinations plus bias coefficient). Each neuron in a layer other than the input layer first calculates the linear combination of the outputs of the neurons of the previous layer. The inputs are multiplied by the corresponding weights and sent as an input to the neuron in their total hidden layers. Each of these neurons is associated with a numerical value called bias. It is added to the input amount. This value goes through a threshold function called the activation function. The result of the activation function determines whether the specific neuron is activated and the activated neurons transmit the data through the channels to the neurons in the next layer. In this method the data is propagated to the network, which is called forward propagation, as shown in below figure(a).

The neuron with the highest value in the output layer determines the output, the value obtained from the output is basically probability. Sometimes the neural network misjudges, and the estimated output is compared to the actual output to eliminate this misconception.

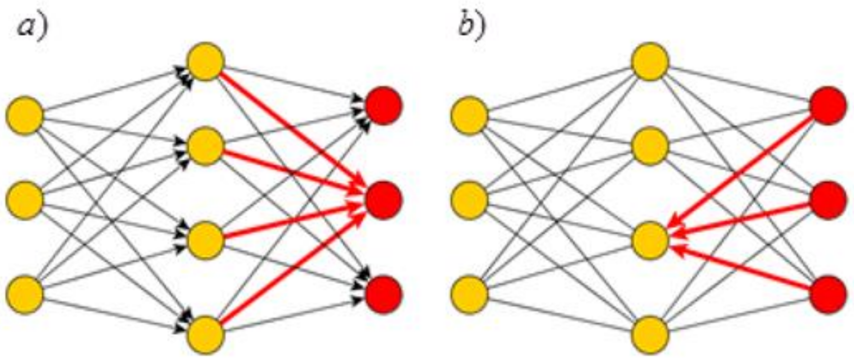


Figure 1. Neural Network Architecture

In order to detect the error in the prediction and reduce the magnitude of the error, the information is transferred backwards from the output layer to the neural network, which is called back propagation, as shown in above figure (b) Based on this information the weights are adjusted in the neural network. The cycle of forward campaigning and backward campaigning is repeated with multiple inputs to minimize error. This process will continue until our weights are assigned so that the network can accurately estimate the output in most cases. This will end our training. Generally, the selected "weight" for the training of input value to obtain the known output value may give error in result. so, the weight is modified accordingly to obtain the known output value. The modification of weight equation can be expressed as

$w_i = w_j + \Delta w_i$

Where, w_i = New Weight

w_j= Original weight

Δw_i= $n(t - o)x_i$

t = target output

o = actual output

x_i = input

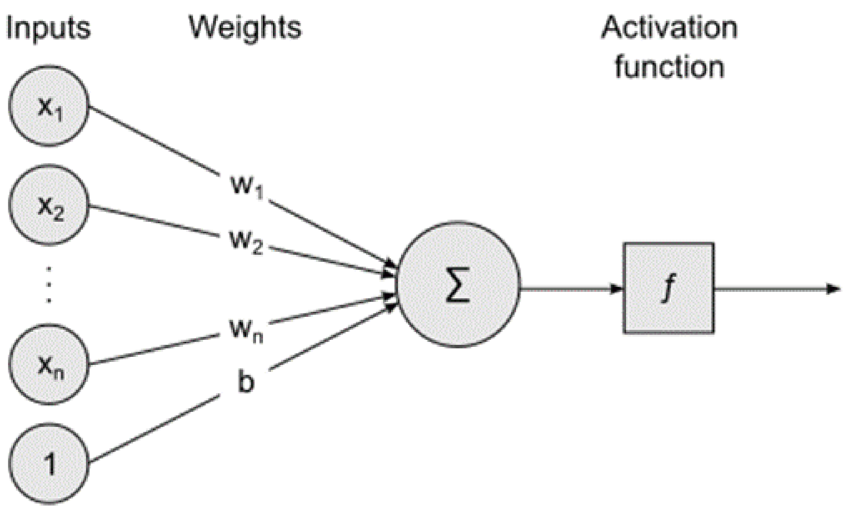


Figure 2. Function of ANN

In this research, the connection between input value and output value can be expressed as

$$Y_j^{a+1} = F\left(\sum_i W_{jb}^a \, X_b^a\right)$$

Equation 3.1

Here, F(x) is activation function. where, log sigmoid function is used. $F(x) = \frac{1}{1+e^{-x}}$. e.g., tan-sigmoid function, and Y_j^{a+1} is output of unit j in the ath layer, and W_{jb}^a is a weight function from unit b in ath layer to unit j in $(b+1)^{th}$. Network training is a procedure in which the related weights and biases of the ANN are embraced through a constant process of simulation by the environment in which the network is placed.

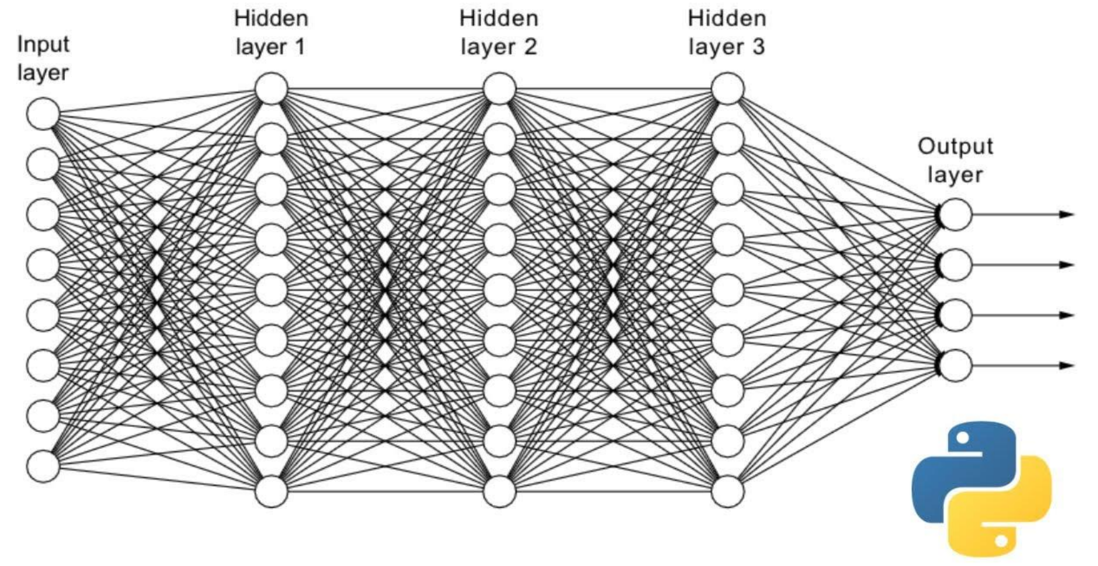


Figure 3. Multi-Layer ANN Architecture

3.2.2 Classification And Regression Tree

In recent times, there has been increasing interest in the use of classification and CART analysis. CART analysis is a tree-building technique, which is different from traditional data analysis methods. Breiman et al. (1984) developed CART, which is a sophisticated program forfitting trees to data. Depending on the value of the predictor variables, a decision tree partitions the data set into regions, so that the response variable is roughly constant in that region. The attractive feature of the CART methodology is a sequence of hierarchical questions that the algorithm asks. This method is relatively simple to understand and interpret. The questions can be answered as ‘yes’ or ‘no’ and depending on the answer, it either proceeds to another question, or arrives at a fitted response values from data.

. This decision tree method constructs classification trees or regression trees depending on the variable type, which may be categorical or numerical [43,44]. Breiman et al. showed that a learning tree can be optimized by using a learning data set to prune the saturated tree and select among the obtained sequence of nested trees [43]. This process helps to retain a simple tree, which ensures robustness. Depending on the target field, several impurity measures can be used to locate splits for CART models. For instance, Gini is usually applied to symbolic target fields while the least-squared deviation method is used for automatically selecting continuous targets without explaining the selections. For node t in a CART, Gini index g(t) is defined as.

$$g(t) = \sum_{j \neq i} p\left(\frac{j}{t}\right) p\left(\frac{i}{t}\right)$$

Equation 3.2

Where i and j are target field categories

A classification tree is developed when the target variable has a discrete value while the regression tree is developed when the target variable is continuous. Root node is the starting point of the classification tree and it possess entire learning dataset at the top of the tree. A node is mainly classified into two types one is terminal nide and another is non terminal node, and it acts as a subset for the set of variables. The node which splits into two child nodes are called non terminal nodes, on the other hand which does not split are called terminal nodes.

CART analysis consists of four basic steps [2]. The first step consists of tree building, during which a tree is built using recursive splitting of nodes. Each resulting node is assigned a predicted class based on the distribution of classes in the learning dataset which would occur in that node and the decision cost matrix. The assignment of a predict class to each node occurs whether or not that node is subsequently split into child nodes. The second step consists of stopping the tree building process. At this point a maximal tree has been produced which probably greatly overfits the information contained within the learning data set. The third step consists of tree pruning which results in the creation of a sequence of simpler and simpler trees through the cutting off of increasingly important nodes. The fourth step consists of optimal tree selection during which the tree which fits the information in the learning dataset but does not overfit the information, is selected from among the sequence of pruned trees.

3.2.3 Regression

Regression analysis is a statistical method to model the relationship between a dependent (target) and independent (predictor) variables with one or more independent variables. More specifically, Regression analysis helps us to understand how the value of the dependent variable is changing corresponding to an independent variable when other independent variables are held fixed. It predicts continuous/real values such as Age, Compressive strength, Flexure Strength, modulus of elasticity, etc.

Regression is a supervised learning technique which helps in finding the correlation between variables and enables us to predict the continuous output variable based on the one or more predictor variables. It is mainly used for prediction, forecasting, time series modeling, and determining the causal-effect relationship between variables. The regression tree is primarily used for the regression task, which is used to estimate continuous value outputs instead of discrete outputs. regression tree is constructed by a procedure known as binary recursive partitioning.

"Regression shows a line or curve that passes through all the datapoints on target-predictor graph in such a way that the vertical distance between the datapoints and the regression line is minimum".

Regression is employed to spot patterns and relationships among a dataset, which may then be applied to new and unseen information. This makes regression a key part of machine learning in finance, and is commonly leveraged to assist forecast portfolio performance or stock prices and trends. Models will be trained to grasp the link between a range of various options and a desired outcome. In most cases, machine learning regression provides organizations with insight into explicit outcomes. however, as a result of this approach will influence AN organization's decision-making method, the explain ability of machine learning is a vital thought.

There are a range of various approaches utilized in machine learning to perform regression. different popular algorithms are used to achieve machine learning regression. The various techniques might include different numbers of independent variables or method different types of information. Distinct forms of machine learning regression models can also assume a distinct relationship between the independent and dependent variables. for example, linear regression techniques assume that the connection is linear, therefore wouldn't be effective with datasets with nonlinear relationships.

Some of the foremost common regression techniques in machine learning is sorted into the following kinds of regression analysis: Simple linear regression, Multiple linear regression, Logistic regression. In this study, out of all regression models, the data used for this study has goodness-for-fit for linear regression model. So, the Linear regression is a statistical regression method which is used for predictive analysis. Linear regression shows the linear relationship between the independent variable and the dependent variable, hence called linear regression. If there is only one input variable, then such linear regression is called **simple linear regression**. And if there is more than one input variable, then such linear regression is called **multiple linear regression**.

A stepwise regression procedure was conducted using SPSS. In the process, some of the variables and nonlinear interactions were thrown away by the procedure due to lack of significant contributions towards the prediction of the value of the dependent variable, Y. Multicollinearity among independent variables was also a factor in the final selection of the model. The final example nonlinear regression model is as follows:

$$Y = 1.474 + 3.536X_2 + 5.856X_4 - 1.734X_1X_2 + 1.505X_1X_3 - 2.564X_2X_3 - 3.438X_3X_4$$

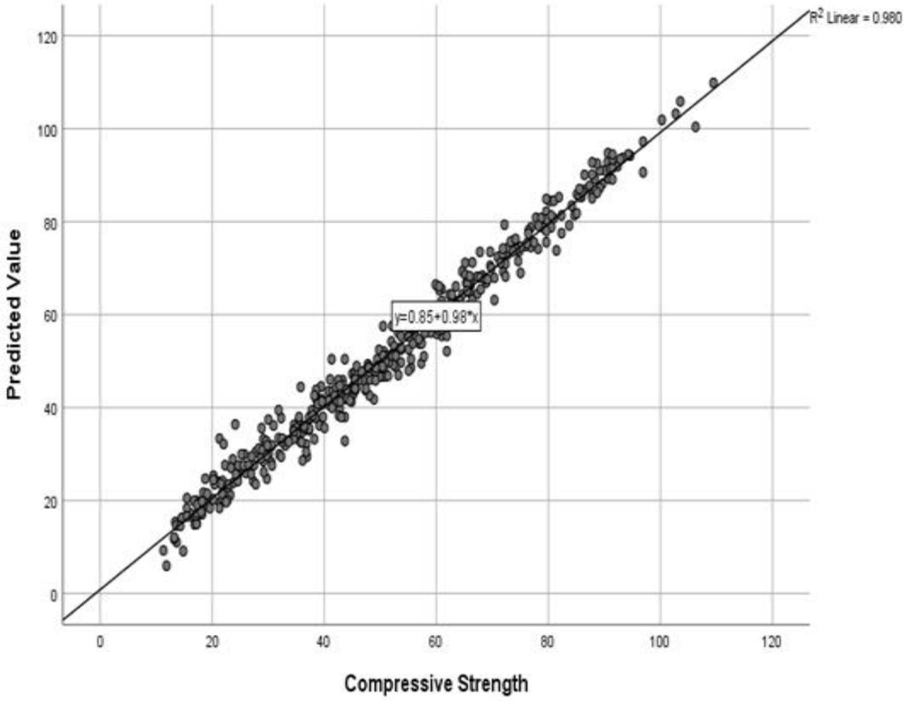
Equation 3.3

4 Experimental Results

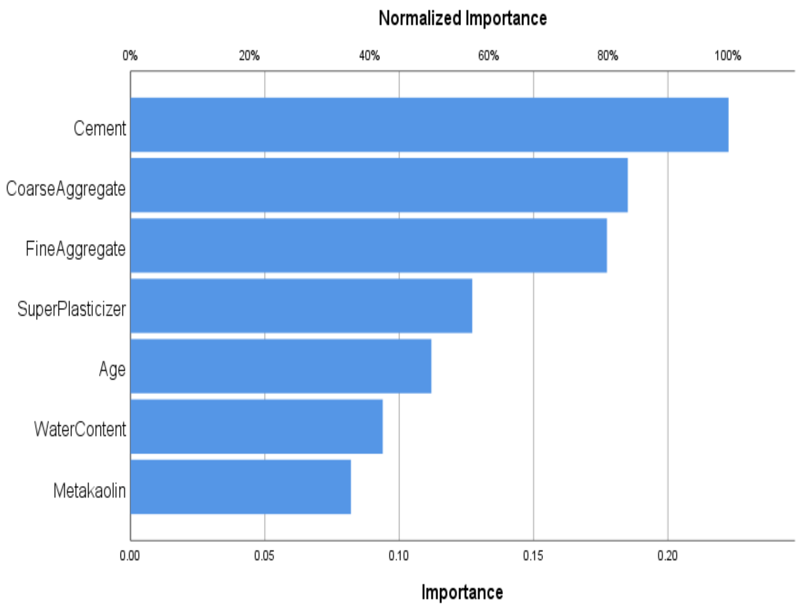
4.1 Metakaolin

4.1.1 Artificial Neural Network

In order to evaluate the efficiency of training and tested model. The R^2 value is considered to indicate how the predicted values varies with original values.



Graphs 1. Experimental Values Vs Predicted Values



Graphs 2. Importance of variables

The graph describes which input variable have more impact on final predicted model.

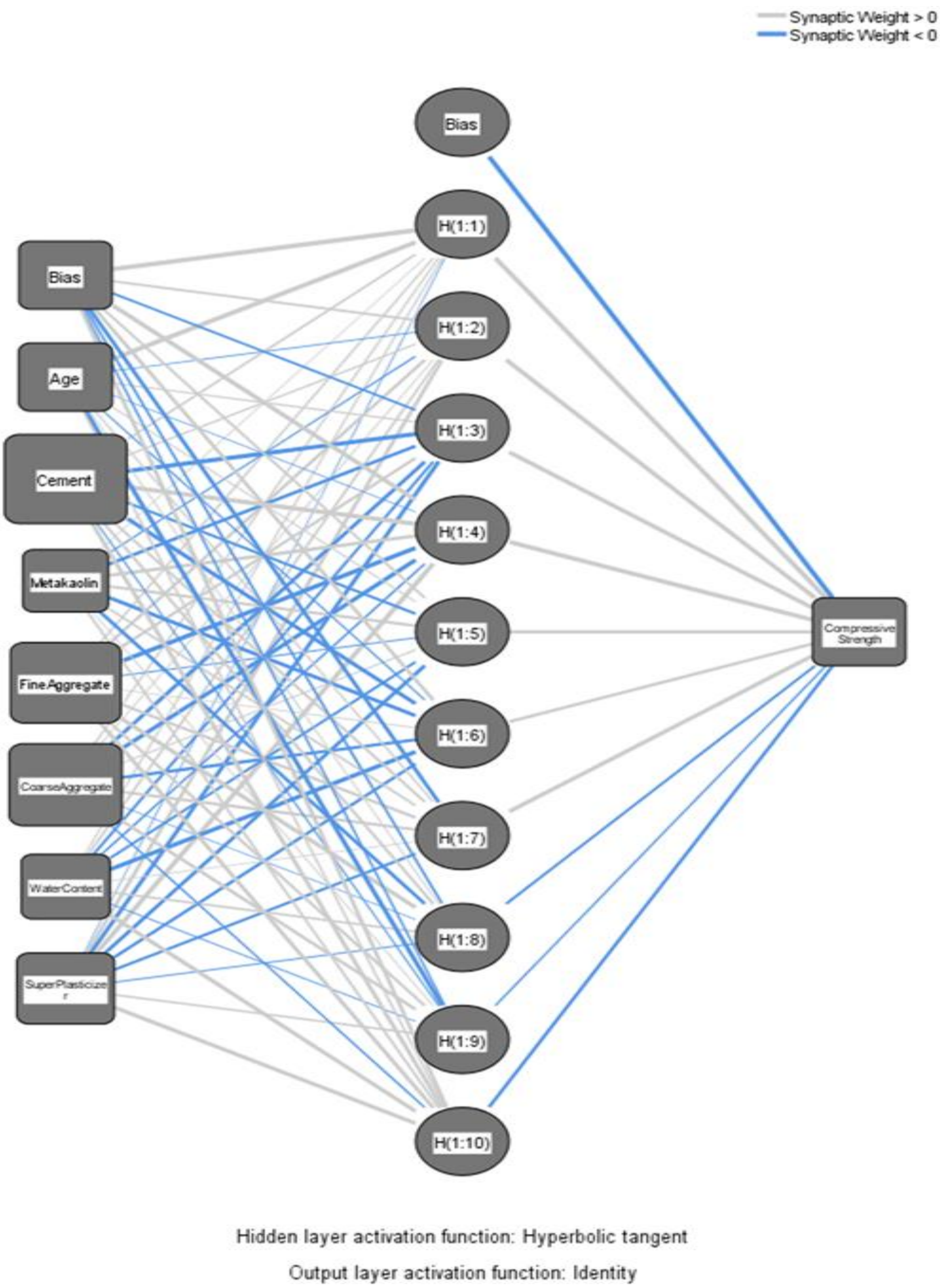


Figure 4. ANN Model Architecture

The Neural Network shows how the input variables and output variables interconnected by the formation of hidden layers. The interconnection between layers are synaptic weights.

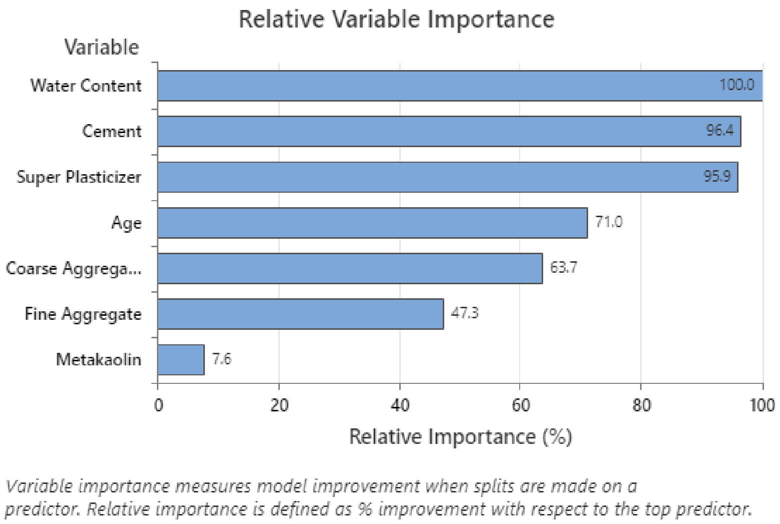
4.1.2 Classification And Regression Tree

From the available data a decision tree is formed in classification and regression tree analysis.

Table 8. Observations from the CART analysis

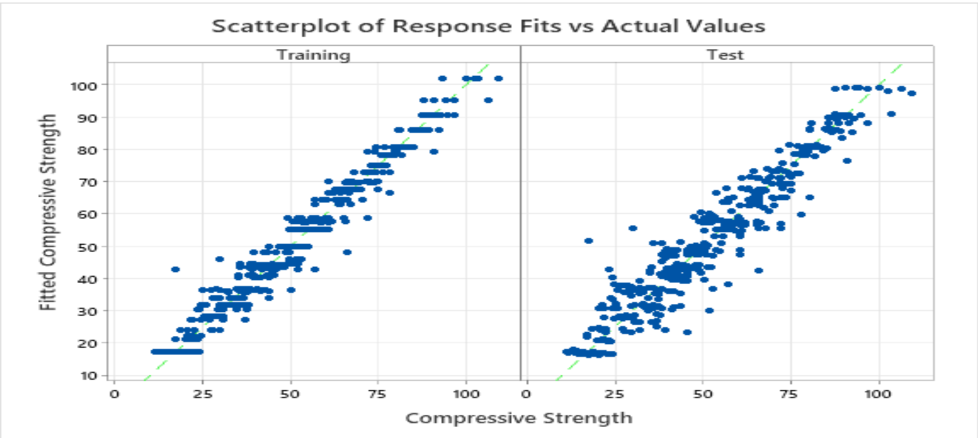
Statistics	Training	Test
R-squared	0.9641	0.9220
Root mean squared error (RMSE)	4.2120	6.2105
Mean squared error (MSE)	17.7412	38.5699
Mean absolute deviation (MAD)	3.0647	4.3930
Mean absolute percent error (MAPE)	0.0783	0.1105

The R-Squared value from the CART analysis is 0.92 (> 0.9). The CART analysis shows the goodness of fit for the present study and the model created in the CART analysis use the nearer values of compressive strength to actual values obtained from manual testing.



Graphs 3. Relative Variable Importance

The graph describes which input variable have more impact on final predicted model.



Graphs 4. Scatter plot Response Fits Vs Actual Values

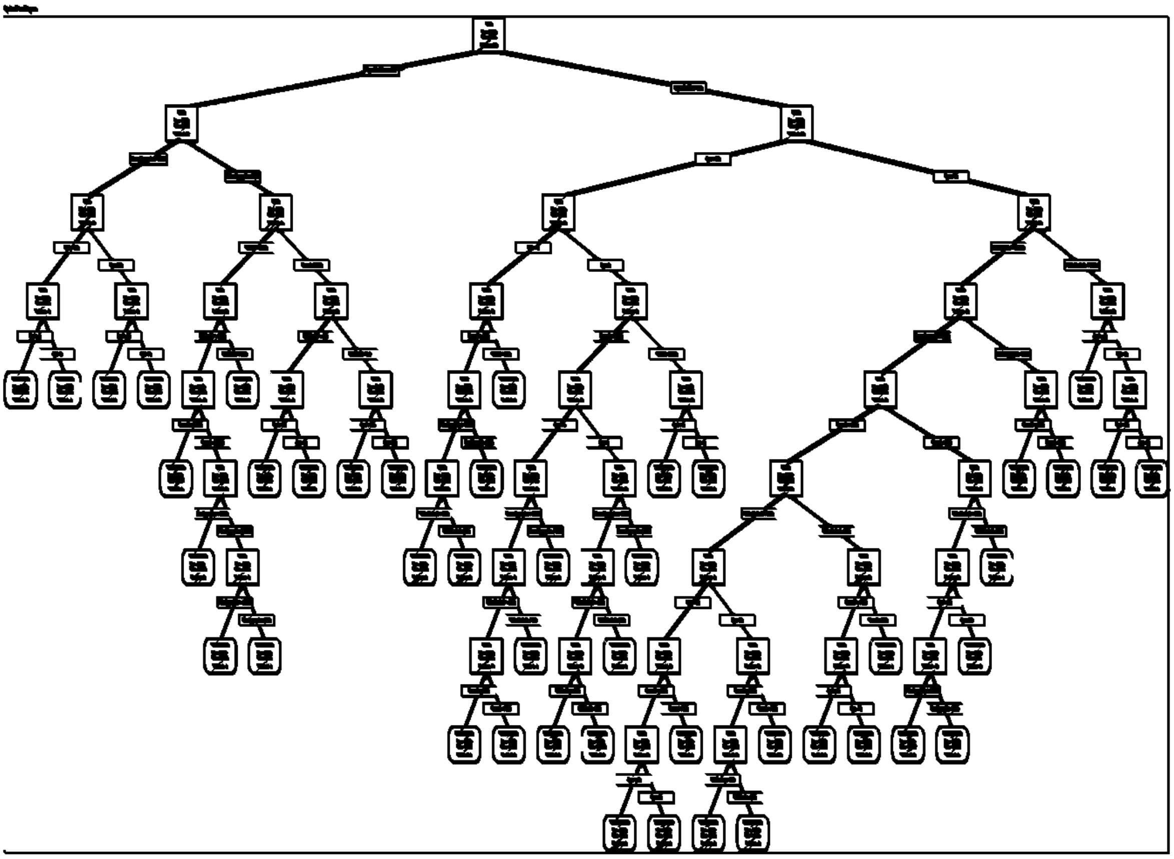


Figure 5. Decision Tree Model

4.1.3 Regression

The regression model proposes an equation to determine the compressive strength. The equation indicates which input variable have more impact on final predicted model.

Compressive Strength = 0.1413 Age + 0.1309Cement + 0.1530 Metakaolin+ 0.03356 Fine Aggregate + 0.01692 Coarse Aggregate - 0.3403 Water Content + 0.1933 Super Plasticizer

Equation 4.1

From the above equation we can conclude that age of concrete, cement, quantity of metakaolin, fine aggregate, coarse aggregate, superplasticizers have positive correlation (Two variables tends to decrease or increase) with compressive strength of harden concrete. However, on the other hand water content has negative correlation (one variable increases and other variable decreases) with concrete which affects the final compressive strength.

Table 9. Summary of Regression Model

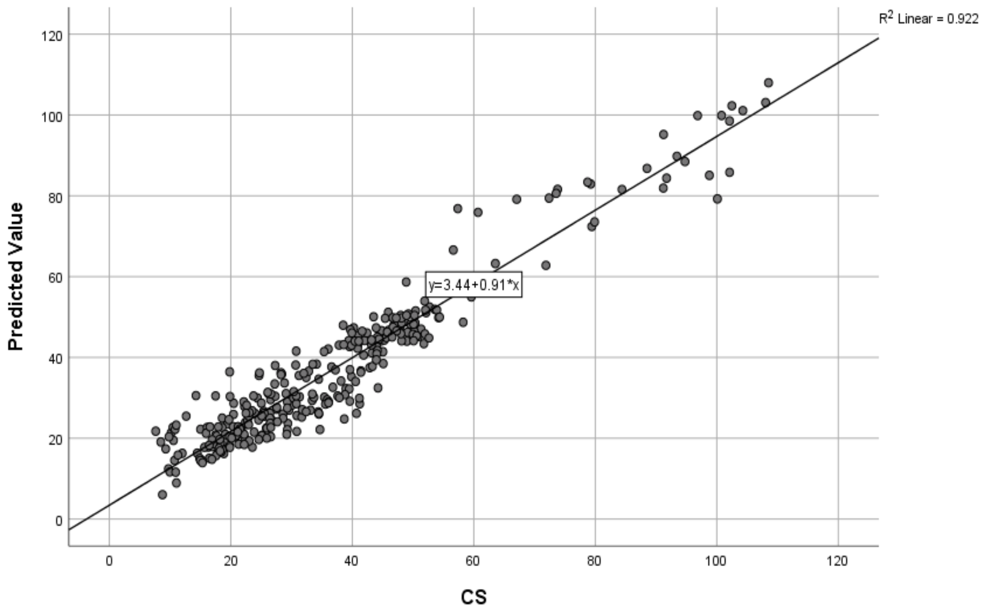
S	R-sq	R-sq (adj)	R-sq (pred)	10-fold S	10-fold R-sq
15.1250	95.40%	95.29%	**95.16%**	15.3720	95.13%

The R-Squared value from the Regression analysis is 0.95 (> 0.9). The Regression analysis shows the goodness of fit for the present study and the model created in the Regression analysis use the nearer values of compressive strength to actual values obtained from manual testing.

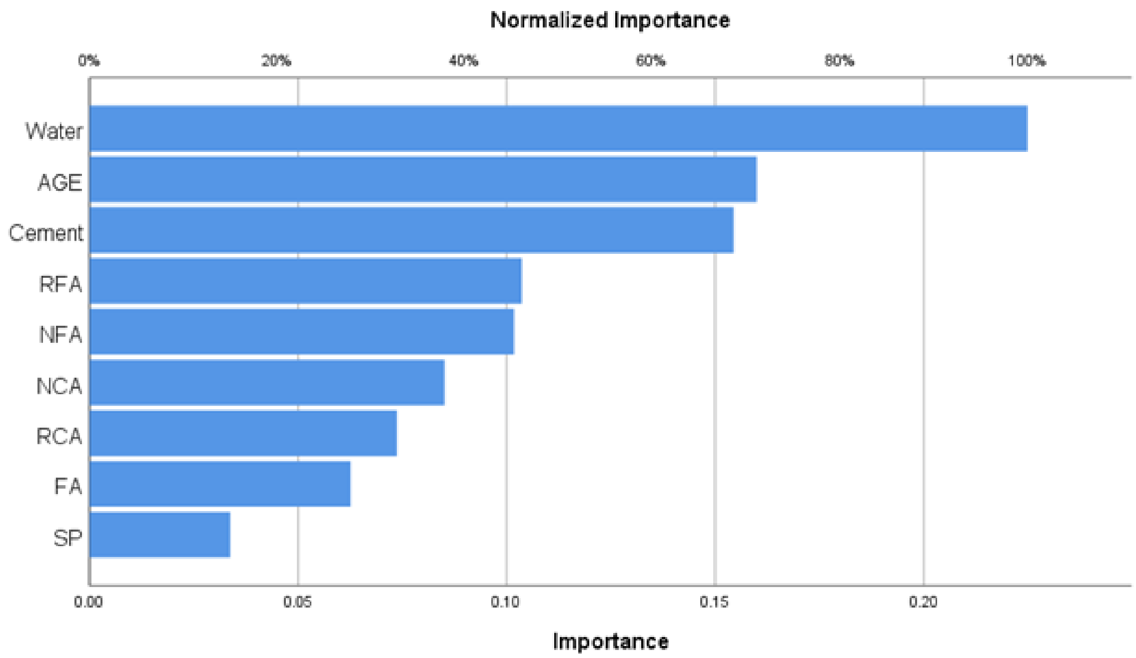
4.2 Recycled Aggregates

4.2.1 Artificial Neural Network

In order to evaluate the efficiency of training and tested model. The R^2 value is considered to indicate how the predicted values varies with original values



Graphs 5. Experimental Compressive Strength Vs Predicted Values



Graphs 6. Relative Importance of Variables

The graph describes which input variable have more impact on final predicted model.

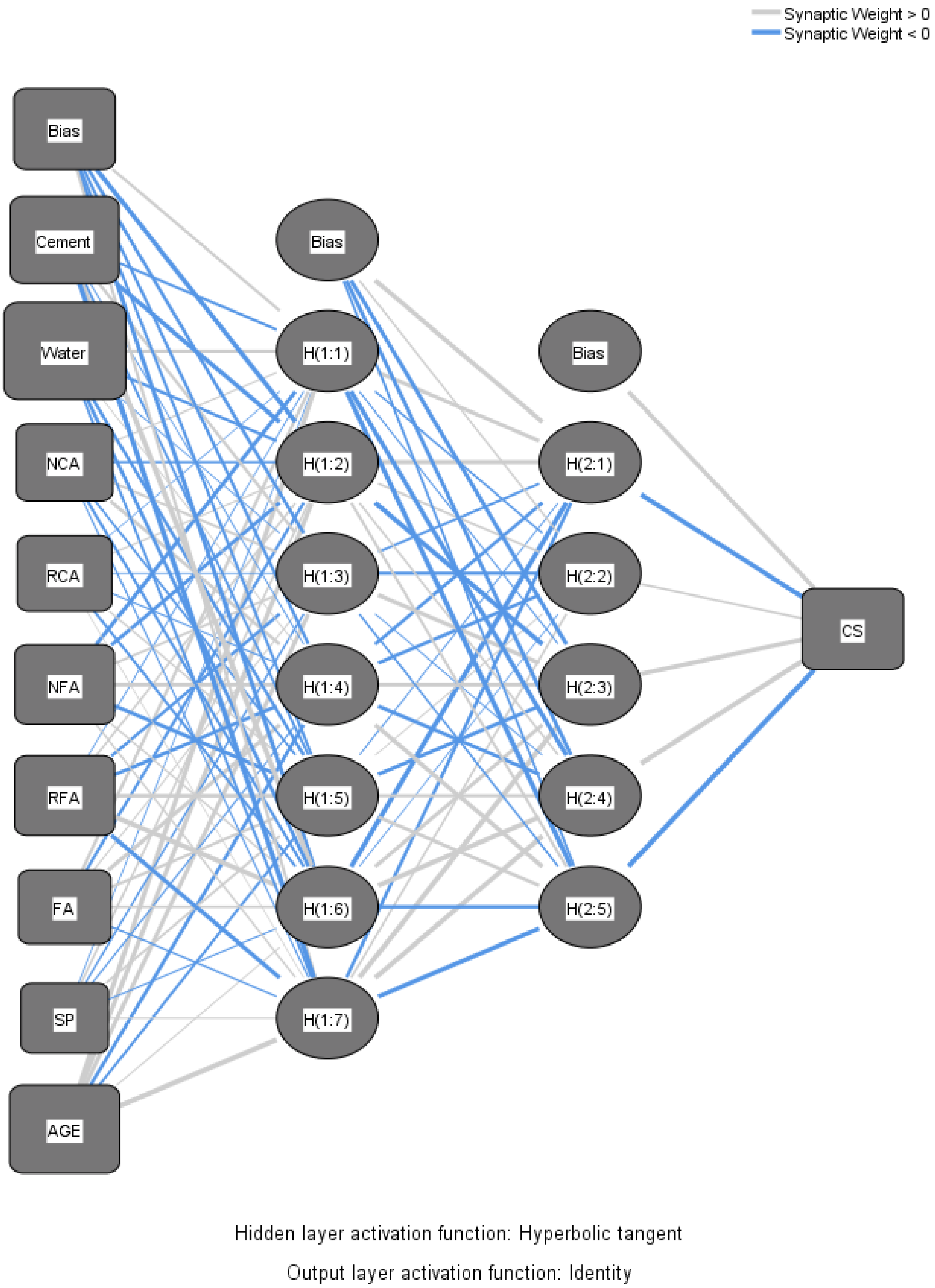


Figure 6. ANN Model Architecture

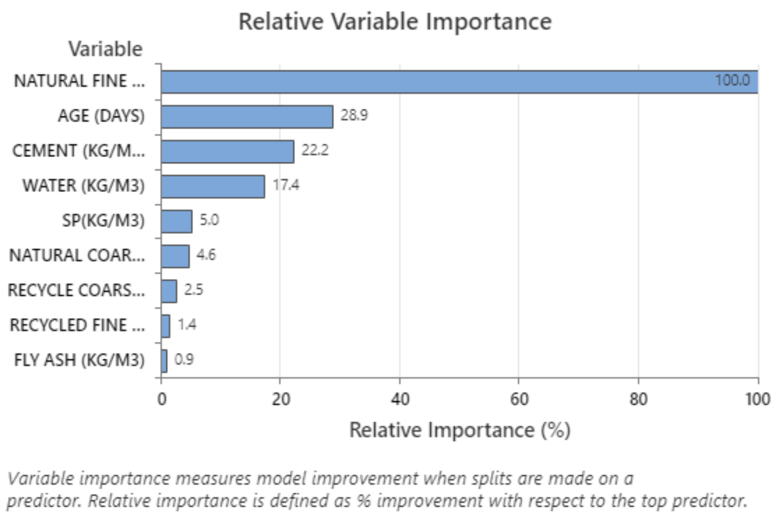
Neural Network shows the relation between input variables and output variables interconnected by formation of hidden layer. The connection line between input layer, hidden and output layer are synaptic weights.

4.2.2 Classification And Regression Tree

Table 10. CART Statistics

Statistics	Training	Test
R-squared	0.9472	0.9237

The R-Squared value from the CART analysis is 0.92(> 0.9). The CART analysis shows the goodness of fit for the present study and the model created in the CART analysis use the nearer values of compressive strength to actual values obtained from manual testing.



Graphs 7. Relative Importance of Variables

The graph describes which input variable have more impact on final predicted model

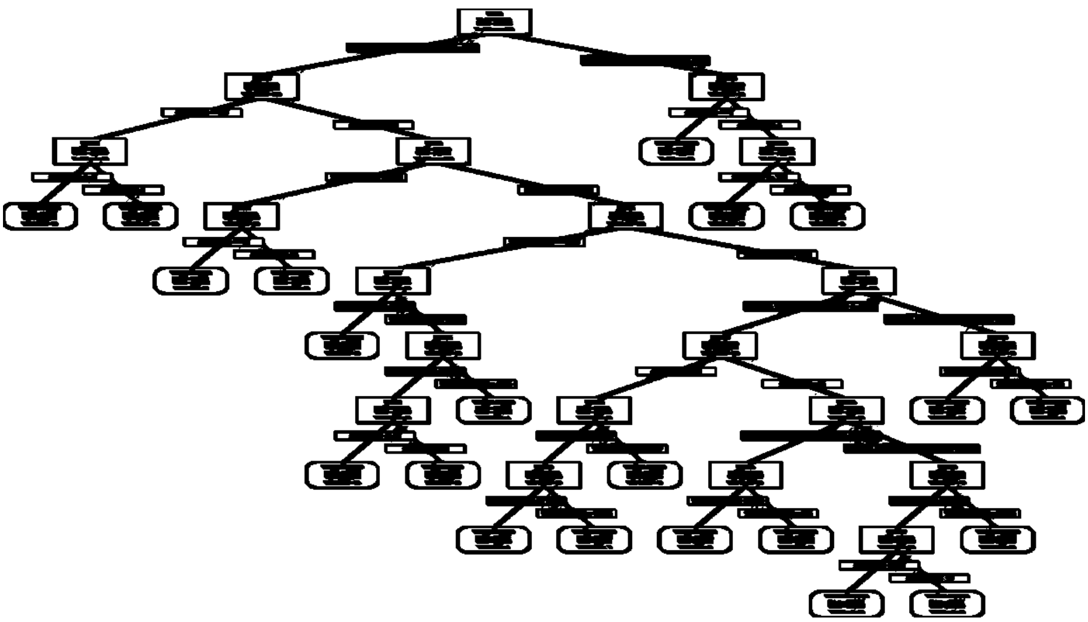


Figure 7. Decision Tree Model

4.2.3 Regression

The regression model proposes an equation to determine the compressive strength. The equation indicates which input variable have more impact on final predicted model.

Compressive Strength = 0.01880 CEMENT (KG/M3)
0.02093 WATER (KG/M3)
+ 0.004290 NATURAL FINE AGGREGATE (KG/M3)
+ 0.000683 RECYCLED FINE AGGREGATE (KG/M3)
+ 0.01581 FLY ASH (KG/M3) + 0.743 SP(KG/M3) + 0.001285 AGE (DAYS)

Equation 4.2

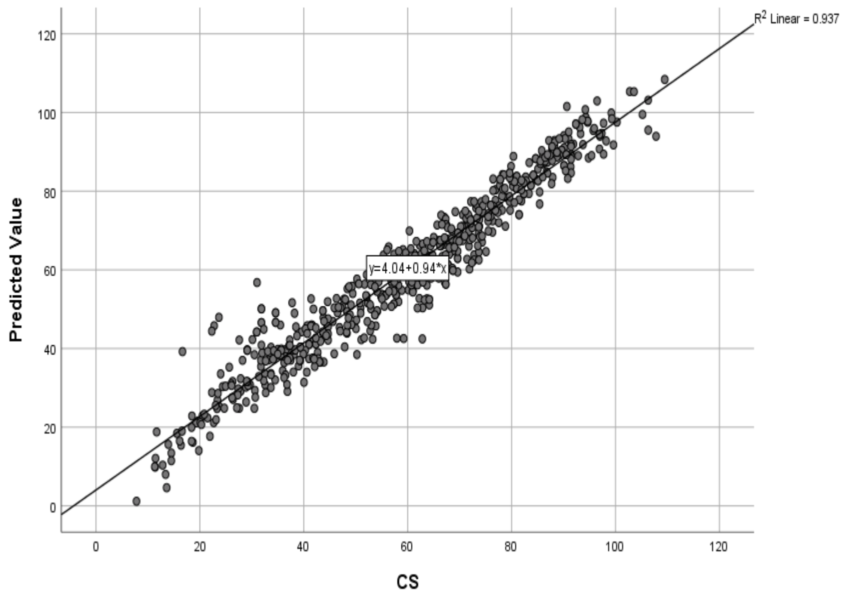
The R-Squared value from the Regression analysis is 0.95 (> 0.9). The Regression analysis shows the goodness of fit for the present study and the model created in the Regression analysis use the nearer values of compressive strength to actual values obtained from manual testing.

S	R-sq	R-sq(adj)	R-sq(pred)	10-fold S	10-fold R-sq
1.11764	96.73%	96.66%	96.57%	1.13984	96.52%

4.3 Silica Fume

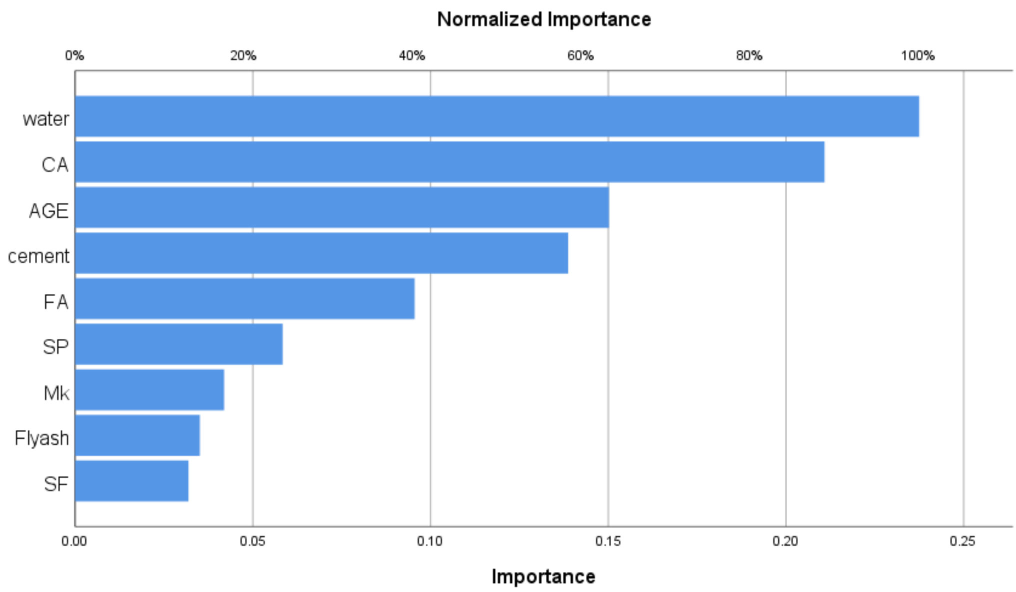
4.3.1 Artificial Neural Network

In order to evaluate the efficiency of training and tested model. The R^2 value is considered to indicate how the predicted values varies with original values.



Graphs 8. Experimental values Vs Predicted Values

The graph describes which input variable have more impact on final predicted model.



Graphs 9. Relative Variable Importance

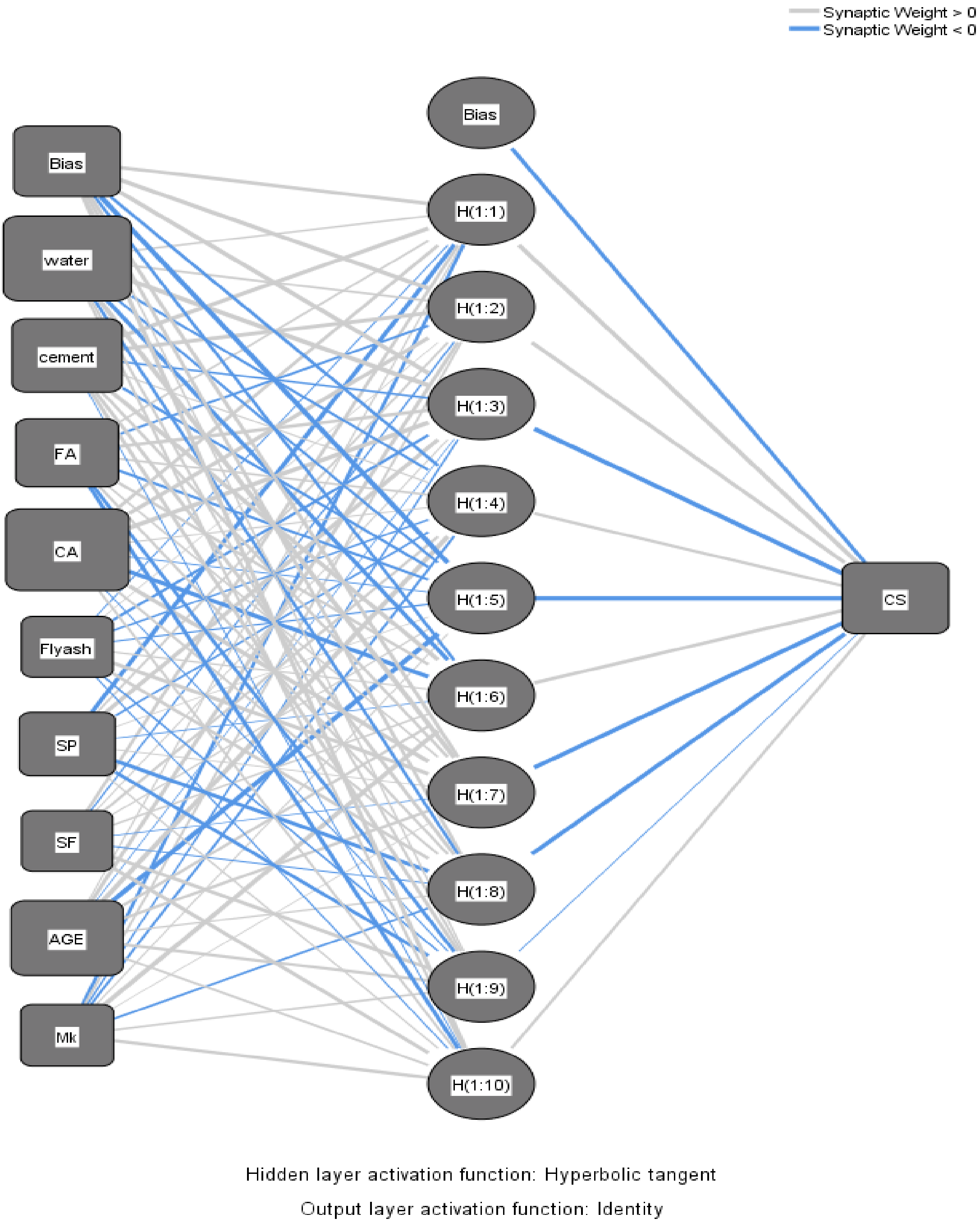


Figure 8. ANN Model Architecture

The Neural Network shows how the input variables and output variables interconnected by the formation of hidden layers. The interconnection between layers are synaptic weights

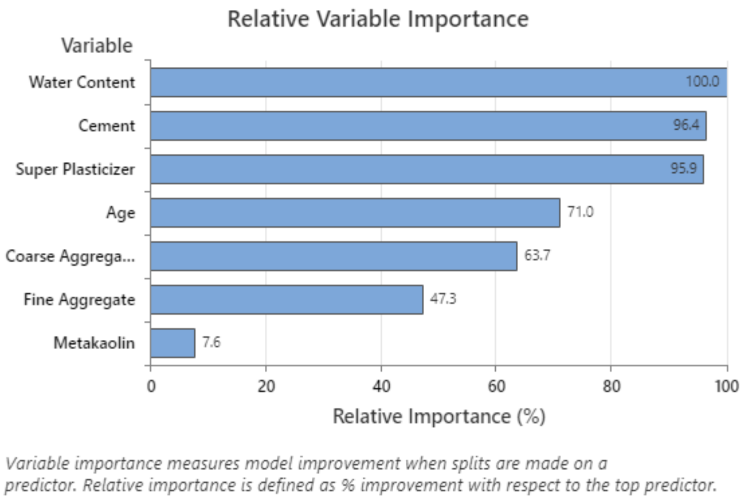
4.3.2 Classification And Regression Tree

From the available data a decision tree is formed in classification and regression tree analysis.

Table 11. CART Analysis Statistics

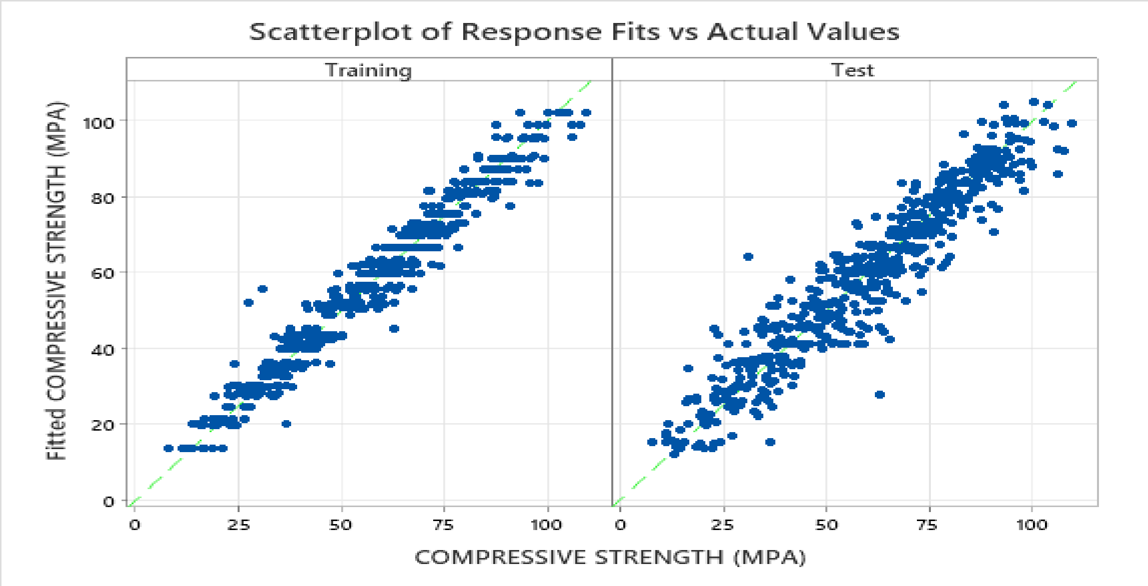
Statistics	Training	Test
R-squared	0.9608	0.9046
Root mean squared error (RMSE)	4.4165	6.8942
Mean squared error (MSE)	19.5057	47.5299
Mean absolute deviation (MAD)	3.3076	5.1227
Mean absolute percent error (MAPE)	0.0683	0.1088

The R-Squared value from the CART analysis is 0.904 (> 0.9). The CART analysis shows the goodness of fit for the present study and the model created in the CART analysis use the nearer values of compressive strength to actual values obtained from manual testing.



Graphs 10. Relative Variable Importance

The graph describes which input variable have more impact on final predicted model.



Graphs 11. Scatter plot of Response Fits Vs Actual Values

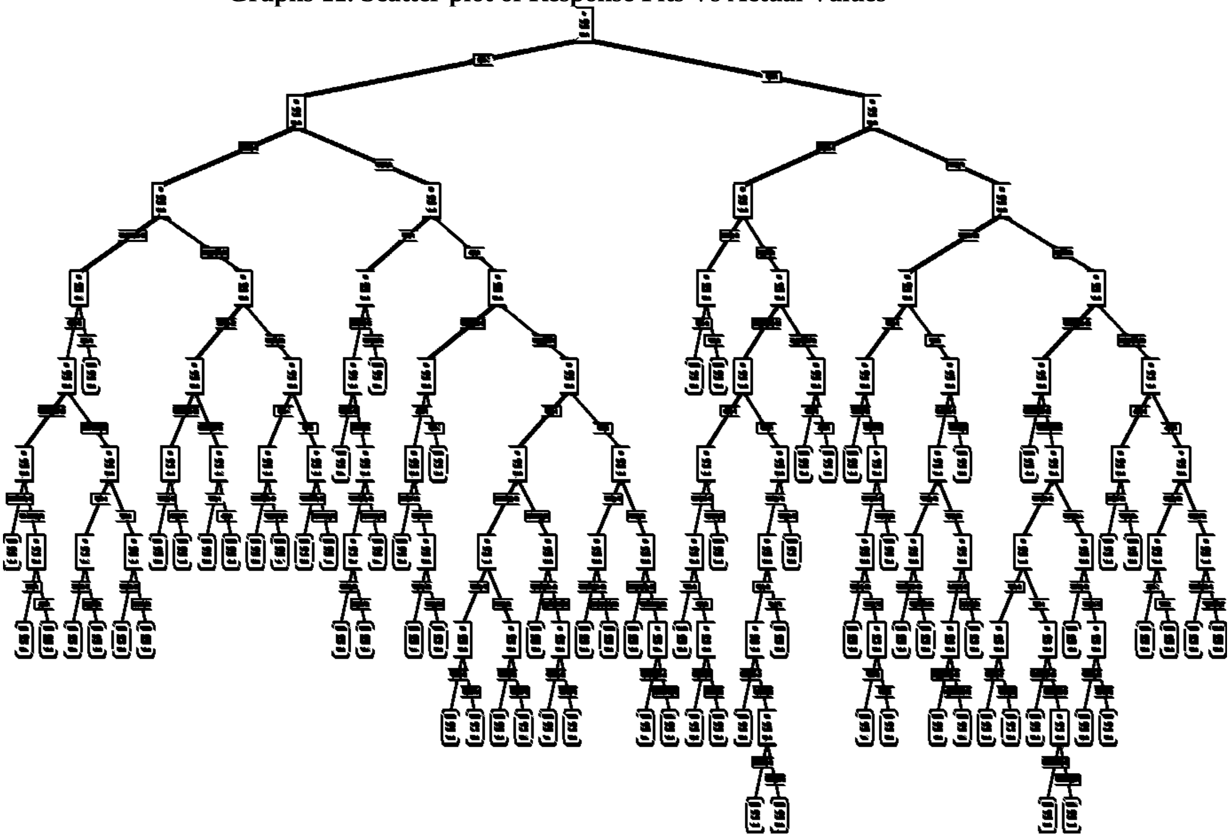


Figure 9. Decision Tree Model

4.3.3 Regression

The regression model proposes an equation to determine the compressive strength. The equation indicates which input variable have more impact on final predicted model.

Compressive Strength = 0.02207 WATER + 0.00896 CEMENT + 0.004658 FA + 0.003111 COARSE AGGREGATE + 0.00385 FLYASH - 0.02325 SUPER PLASTICIZER + 0.01293 SILICA FUME+ 0.01800 AGE + 0.01619 MK

From the above equation we can conclude that age of concrete, cement, quantity of metakaolin, fine aggregate, coarse aggregate, superplasticizers have positive correlation (Two variables tends to decrease or increase) with compressive strength of harden concrete. However, on the other hand water content has negative correlation (one variable increases and other variable decreases) with concrete which affects the final compressive strength.

Table 12. Regression Statistics

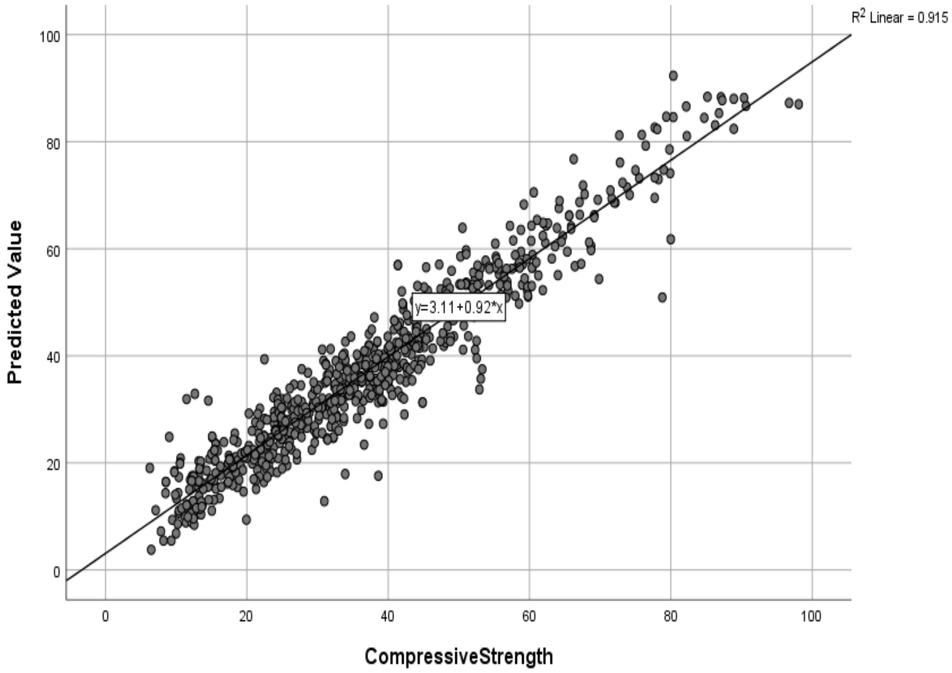
S	R-sq	R-sq(adj)	R-sq(pred)	10-fold S	10-fold R-sq
1.46901	96.45%	96.39%	**94.68%**	96.40%	2258.49

The R-Squared value from the Regression analysis is 0.94 (> 0.9). The Regression analysis shows the goodness of fit for the present study and the model created in the Regression analysis use the nearer values of compressive strength to actual values obtained from manual testing.

4.4 Fly Ash

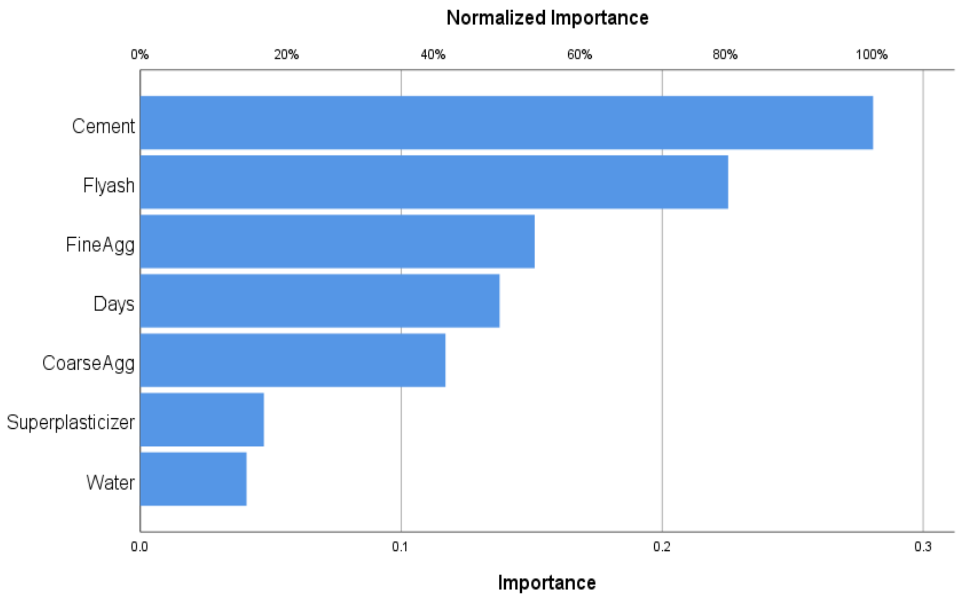
4.4.1 Artificial Neural Network

In order to evaluate the efficiency of training and tested model. The R^2 value is considered to indicate how the predicted values varies with original values.



Graphs 12. Predicted Values Vs Actual Values

The graph describes which input variable have more impact on final predicted model.



Graphs 13. Relative Variable Importance

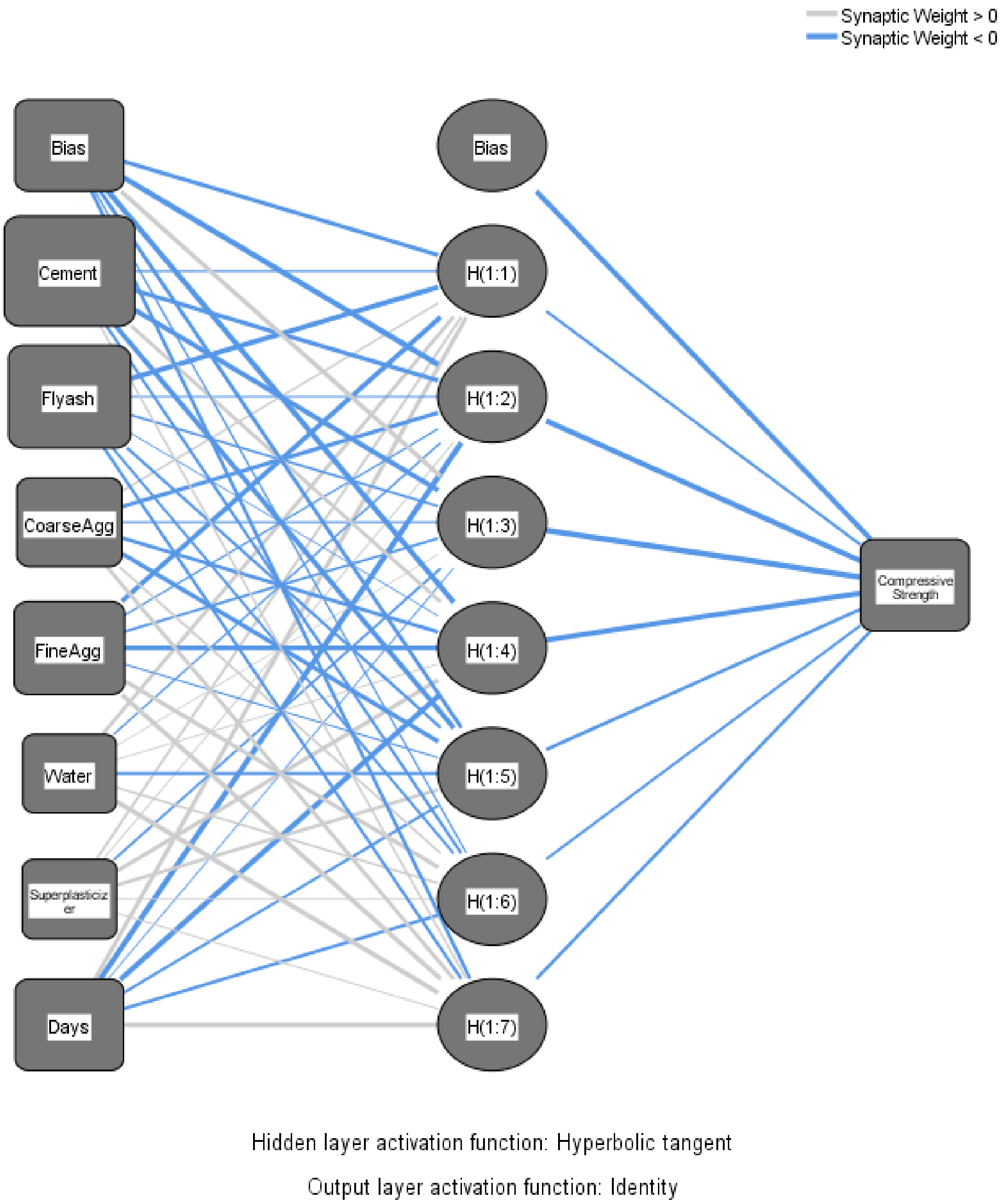


Figure 10. ANN Model Architecture

The Neural Network shows how the input variables and output variables interconnected by the formation of hidden layers. The interconnection between layers are synaptic weights.

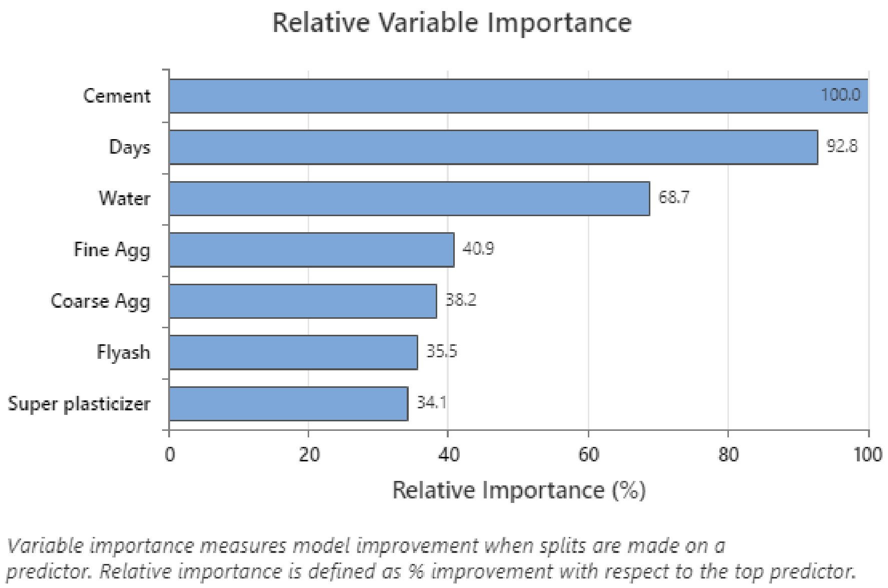
4.4.2 Classification And Regression Tree

From the available data a decision tree is formed in classification and regression tree analysis.

Table 13. CART Analysis Statistics

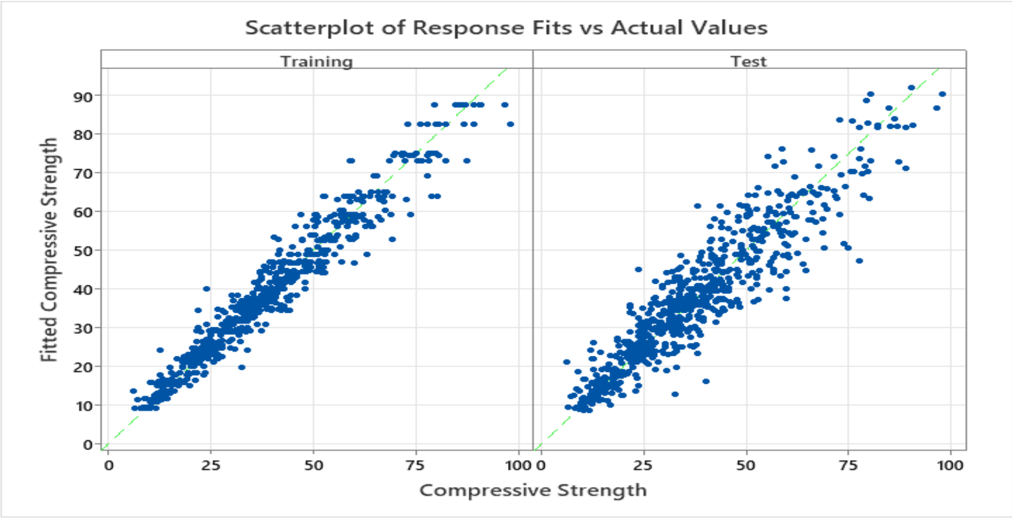
Statistics	Training	Testing
R – Squared	0.945	0.862

The R-Squared value from the CART analysis is 0.863. The CART analysis shows the goodness of fit for the present study and the model created in the CART analysis use the nearer values of compressive strength to actual values obtained from manual testing.



Graphs 14. Relative Variable Importance

The graph describes which input variable have more impact on final predicted model.



Graphs 15. Scatterplot of Response fits Vs Actual Values

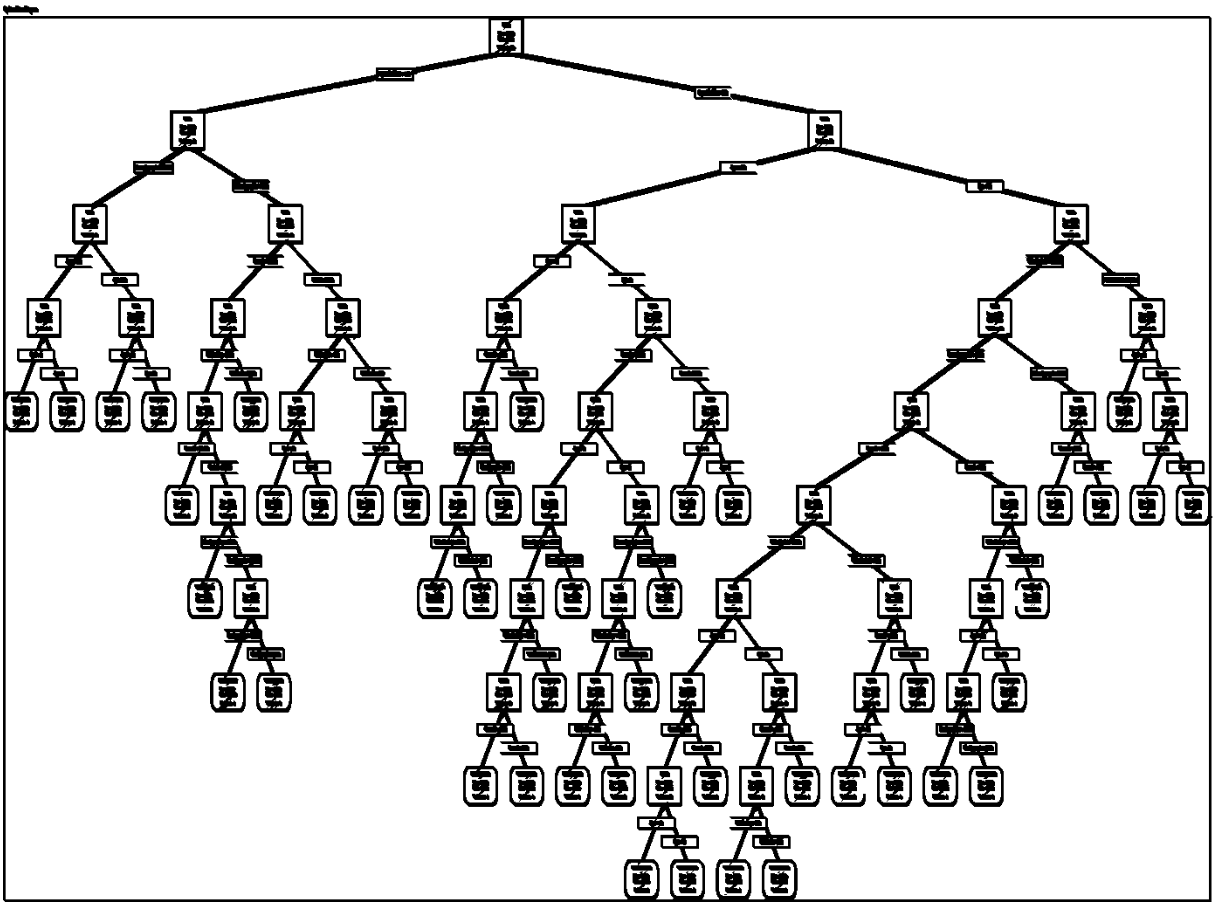


Figure 11. Decision Tree Model

4.4.3 Regression

The regression model proposes an equation to determine the compressive strength. The equation indicates which input variable have more impact on final predicted model.

Compressive Strength = 0.14121 Cement + 0.08100 Flyash - 0.00191 Coarse Agg
+ 0.02133 Fine Agg - 0.2060 Water + 0.4029 Super plasticizer
+ 0.11404 Days

Equation 4.3

From the above equation we can conclude that age of concrete, cement, quantity of metakaolin, fine aggregate, coarse aggregate, superplasticizers have positive correlation (Two variables tends to decrease or increase) with compressive strength of harden concrete. However, on the other hand water content has negative correlation (one variable increases and other variable decreases) with concrete which affects the final compressive strength.

Table 14. Regression Model Statistics

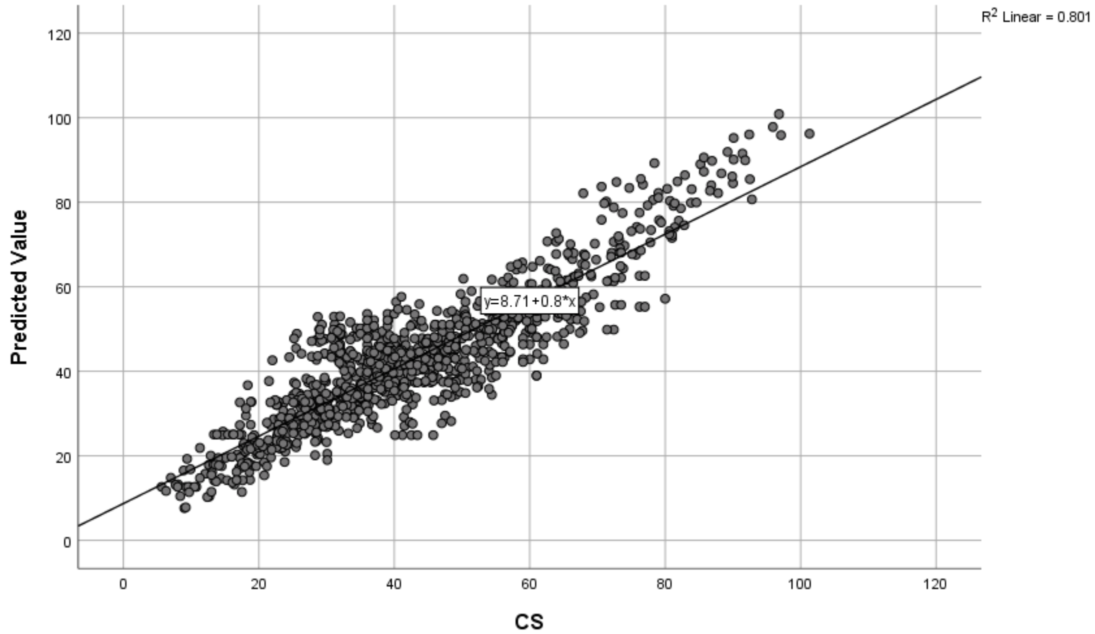
S	R-sq	R-sq(adj)	R-sq(pred)	10-fold S	10-fold R-sq
9.66553	94.53%	94.48%	**94.37%**	9.78386	94.34%

The R-Squared value from the Regression analysis is 0.94 (> 0.9). The Regression analysis shows the goodness of fit for the present study and the model created in the Regression analysis use the nearer values of compressive strength to actual values obtained from manual testing.

4.5 Ground Granulated Blast Furnace Slag

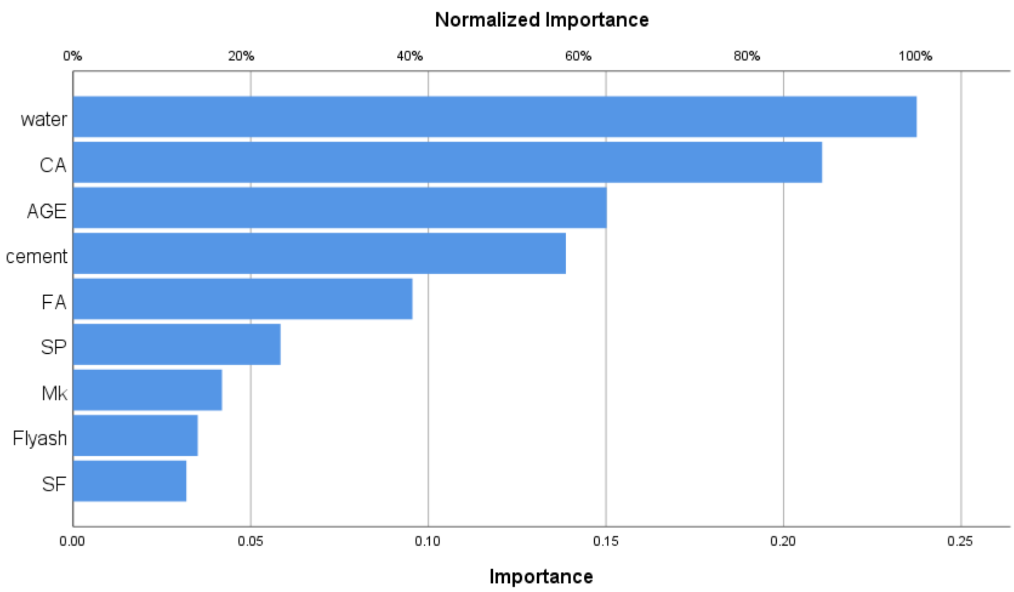
4.5.1 Artificial Neural Network

In order to evaluate the efficiency of training and tested model. The R^2 value is considered to indicate how the predicted values varies with original values.



Graphs 16. Predicted Values Vs Actual Values

The graph describes which input variable have more impact on final predicted model.



Graphs 17. Relative Importance of Variables

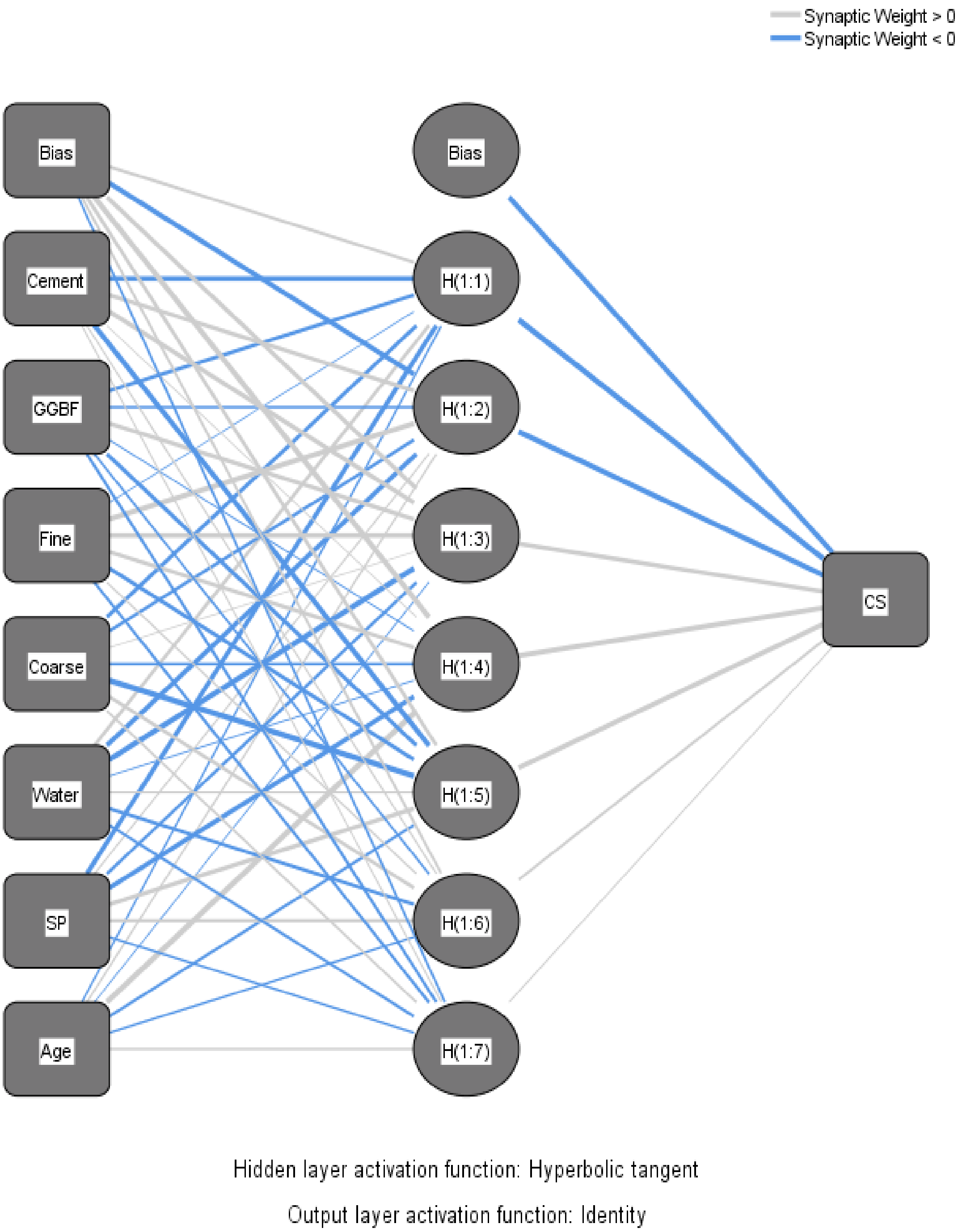


Figure 12. ANN Model Architecture

The Neural Network shows how the input variables and output variables interconnected by the formation of hidden layers. The interconnection between layers are synaptic weights.

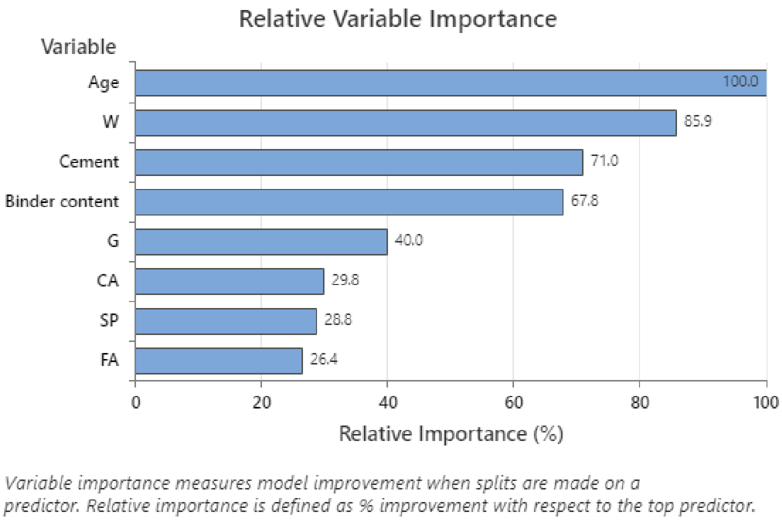
4.5.2 Classification And Regression Tree

From the available data a decision tree is formed in classification and regression tree analysis.

Table 15. CART Model Statistics

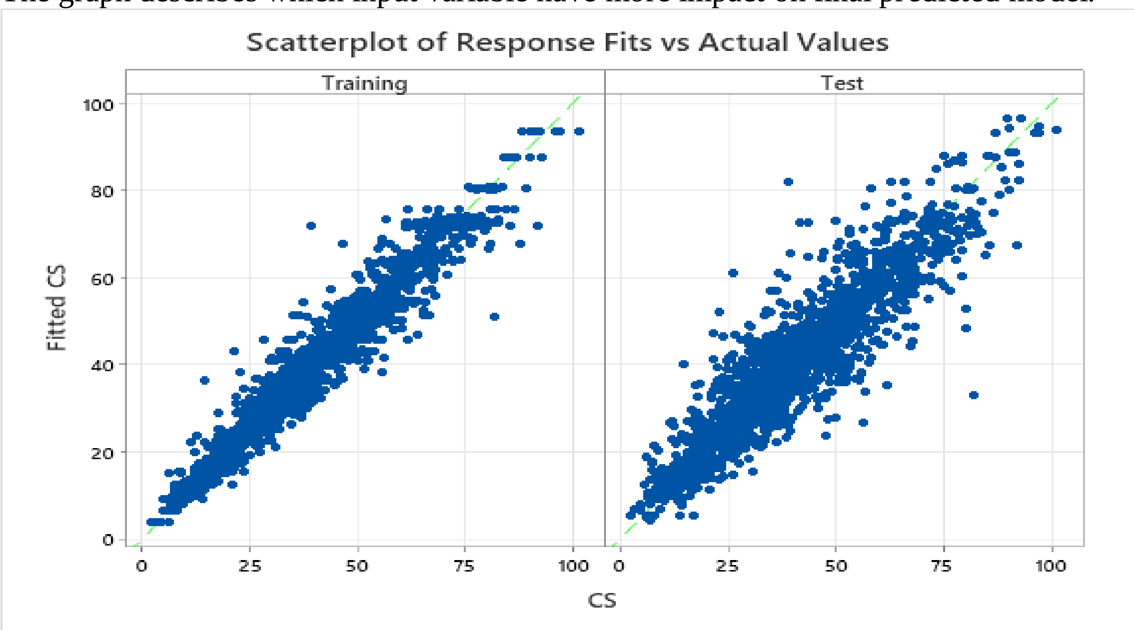
Statistics	Training	Testing
R – Squared	0.94	0.86

The R-Squared value from the CART analysis is 0.86. The CART analysis shows the goodness of fit for the present study and the model created in the CART analysis use the nearer values of compressive strength to actual values obtained from manual testing.



Graphs 18. Relative Variable Importance

The graph describes which input variable have more impact on final predicted model.



Graphs 19. Scatter Plot of Response Fits Vs Actual Values

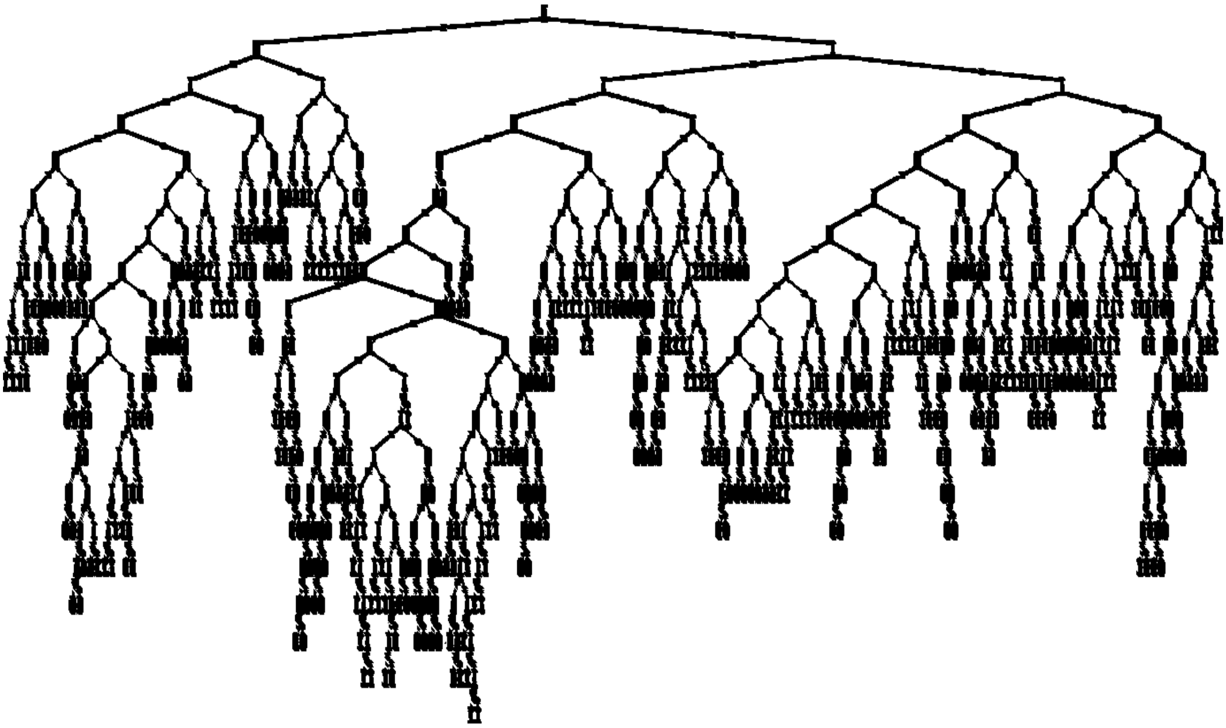


Figure 13. Decision Tree Model

4.5.3 Regression

The regression model proposes an equation to determine the compressive strength. The equation indicates which input variable have more impact on final predicted model.

Compressive Strength = 0.05921 Cement + 0.05542 Binder content + 0.02583 W + 0.01957 G + 0.02006 FA + 0.00216 CA + 1.3544 SP + 0.07695 Age

Equation 4.4

From the above equation we can conclude that age of concrete, cement, quantity of metakaolin, fine aggregate, coarse aggregate, superplasticizers have positive correlation (Two variables tends to decrease or increase) with compressive strength of harden concrete. However, on the other hand water content has negative correlation (one variable increases and other variable decreases) with concrete which affects the final compressive strength.

Table 16. Regression Model Statistics

S	R-sq	R-sq (adj)	R-sq (pred)	10-fold S	10-fold R-sq
15.6360	86.93%	86.88%	**86.82%**	15.6682	86.83%

The R-Squared value from the Regression analysis is 0.87. The Regression analysis shows the goodness of fit for the present study and the model created in the Regression analysis use the nearer values of compressive strength to actual values obtained from manual testing.

5 Conclusion

• A compressive strength prediction models have been developed by considering the independent variables namely Age of the concrete, Content of SCMs, Water Content, Fine aggregate (Kg/m^3), coarse aggregate (Kg/m^3), super plasticizers (Kg/m^3), by adopting ANN, CART, and Regression models.

• Using Machine learning Approach, Predictive model for compressive strength of recycled aggregate mixture has been developed by considering the input variables, Cement (Kg/M^3), Water, Natural Coarse Aggregate (Kg/M^3), Recycle Coarse Aggregate (Kg/M^3), Natural Fine Aggregate (Kg/M^3), Recycled Fine Aggregate (Kg/M^3), Flyash (Kg/M^3), Super plasticizer (Kg/M^3), Age (Days).

Table 17. Comparison of Various Models

Supplementary cementitious materials	R^2 -Value		
	ANN	CART	Regression
Metakaolin (MK)	0.908	**0.909**	0.95
Recycled aggregate (RA)	0.922	0.923	**0.967**
Silica fume (SF)	0.937	0.904	**0.96**
Flyash (FA)	0.915	0.868	**0.94**
Ground granulated blast furnace slag (GGBFS)	0.8	0.85	**0.86**

References

[1] J. Kasperkiewicz and A. Dubrawskp, “BPe STRENGTH PREDICTION USING ARTIFICIAL NEURAL NETWORK By Janusz Kasperkiewicz ,] Janusz Racz , 2 and Artur DubrawskP,” *J. Comput. Civ. Eng.*, vol. 9, no. 4, pp. 279–284, 1996.

[2] J. Sobhani, M. Najimi, A. R. Pourkhorshidi, and T. Parhizkar, “Prediction of the compressive strength of no-slump concrete: A comparative study of regression, neural network and ANFIS models,” *Constr. Build. Mater.*, vol. 24, no. 5, pp. 709–718, 2010, doi: 10.1016/j.conbuildmat.2009.10.037.

[3] E. M. Golafshani, A. Behnood, and M. Arashpour, “Predicting the compressive strength of normal and High-Performance Concretes using ANN and ANFIS hybridized with Grey Wolf Optimizer,” *Constr. Build. Mater.*, vol. 232, p. 117266, 2020, doi: 10.1016/j.conbuildmat.2019.117266.

[4] M. Serraye, S. Kenai, and B. Boukhatem, “Prediction of compressive strength of self-compacting concrete (SCC) with silica fume using neural networks models,” *Civ. Eng. J.*, vol. 7, no. 1, pp. 118–139, 2021, doi: 10.28991/cej-2021-03091642.

[5] M. Sebastiá, I. F. Olmo, and A. Irabien, “Neural network prediction of unconfined compressive strength of coal fly ash-cement mixtures,” *Cem. Concr. Res.*, vol. 33, no. 8, pp. 1137–1146, 2003, doi: 10.1016/S0008-8846(03)00019-X.

[6] I.-C. Yeh, “Analysis of Strength of Concrete Using Design of Experiments and Neural Networks,” *J. Mater. Civ. Eng.*, vol. 18, no. 4, pp. 597–604, 2006, doi: 10.1061/(asce)0899-1561(2006)18:4(597).

[7] I. B. Topçu and M. Saridemir, “Prediction of compressive strength of concrete containing fly ash using artificial neural networks and fuzzy logic,” *Comput. Mater. Sci.*, vol. 41, no. 3, pp. 305–311, 2008, doi: 10.1016/j.commatsci.2007.04.009.

[8] B. K. R. Prasad, H. Eskandari, and B. V. V. Reddy, “Prediction of compressive strength of SCC and HPC with high volume fly ash using ANN,” *Constr. Build. Mater.*, vol. 23, no. 1, pp. 117–128, 2009, doi: 10.1016/j.conbuildmat.2008.01.014.

[9] A. Baykasoğlu, A. Öztaş, and E. Özbay, “Prediction and multi-objective optimization of high-strength concrete parameters via soft computing approaches,” *Expert Syst. Appl.*, vol. 36, no. 3 PART 2, pp. 6145–6155, 2009, doi: 10.1016/j.eswa.2008.07.017.

[10] R. Siddique, P. Aggarwal, and Y. Aggarwal, “Prediction of compressive strength of self-compacting concrete containing bottom ash using artificial neural networks,” *Adv. Eng. Softw.*, vol. 42, no. 10, pp. 780–786, 2011, doi: 10.1016/j.advengsoft.2011.05.016.

[11] M. Y. Cheng, J. S. Chou, A. F. V. Roy, and Y. W. Wu, “High-performance Concrete Compressive Strength Prediction using Time-Weighted Evolutionary Fuzzy Support Vector Machines Inference Model,” *Autom. Constr.*, vol. 28, pp. 106–115, 2012, doi: 10.1016/j.autcon.2012.07.004.

[12] P. Aggarwal, Y. Aggarwal, R. Siddique, S. Gupta, and H. Garg, “Fuzzy logic modeling of compressive strength of high-strength concrete (Hsc) with supplementary cementitious material,” *J. Sustain. Cem. Mater.*, vol. 2, no. 2, pp. 128–143, 2013, doi: 10.1080/21650373.2013.801800.

[13] M. Azimi-Pour, H. Eskandari-Naddaf, and A. Pakzad, “Linear and non-linear

SVM prediction for fresh properties and compressive strength of high volume fly ash self-compacting concrete," *Constr. Build. Mater.*, vol. 230, p. 117021, 2020, doi: 10.1016/j.conbuildmat.2019.117021.

[14] M. M. Roshani, S. H. Kargar, V. Farhangi, and M. Karakouzian, "Predicting the effect of fly ash on concrete's mechanical properties by ann," *Sustain.*, vol. 13, no. 3, pp. 1–16, 2021, doi: 10.3390/su13031469.

[15] C. Bilim, C. D. Atiş, H. Tanyildizi, and O. Karahan, "Predicting the compressive strength of ground granulated blast furnace slag concrete using artificial neural network," *Adv. Eng. Softw.*, vol. 40, no. 5, pp. 334–340, 2009, doi: 10.1016/j.advengsoft.2008.05.005.

[16] M. Saridemir, "Predicting the compressive strength of mortars containing metakaolin by artificial neural networks and fuzzy logic," *Adv. Eng. Softw.*, vol. 40, no. 9, pp. 920–927, 2009, doi: 10.1016/j.advengsoft.2008.12.008.

[17] A. Kandiri, E. Mohammadi Golafshani, and A. Behnood, "Estimation of the compressive strength of concretes containing ground granulated blast furnace slag using hybridized multi-objective ANN and salp swarm algorithm," *Constr. Build. Mater.*, vol. 248, p. 118676, 2020, doi: 10.1016/j.conbuildmat.2020.118676.

[18] S. Safarzadegan Gilan, H. Bahrami Jovein, and A. A. Ramezanianpour, "Hybrid support vector regression - Particle swarm optimization for prediction of compressive strength and RCPT of concretes containing metakaolin," *Constr. Build. Mater.*, vol. 34, pp. 321–329, 2012, doi: 10.1016/j.conbuildmat.2012.02.038.

[19] I.-C. Yeh, "Modeling Concrete Strength with Augment-Neuron Networks," *J. Mater. Civ. Eng.*, vol. 10, no. 4, pp. 263–268, 1998, doi: 10.1061/(asce)0899-1561(1998)10:4(263).

[20] I. C. Yeh and L. C. Lien, "Knowledge discovery of concrete material using Genetic Operation Trees," *Expert Syst. Appl.*, vol. 36, no. 3 PART 2, pp. 5807–5812, 2009, doi: 10.1016/j.eswa.2008.07.004.

[21] J.-S. Chou, C.-K. Chiu, M. Farfoura, and I. Al-Taharwa, "Optimizing the Prediction Accuracy of Concrete Compressive Strength Based on a Comparison of Data-Mining Techniques," *J. Comput. Civ. Eng.*, vol. 25, no. 3, pp. 242–253, 2011, doi: 10.1061/(asce)cp.1943-5487.0000088.

[22] S. M. Mousavi, P. Aminian, A. H. Gandomi, A. H. Alavi, and H. Bolandi, "A new predictive model for compressive strength of HPC using gene expression programming," *Adv. Eng. Softw.*, vol. 45, no. 1, pp. 105–114, 2012, doi: 10.1016/j.advengsoft.2011.09.014.

[23] H. I. Erdal, "Two-level and hybrid ensembles of decision trees for high performance concrete compressive strength prediction," *Eng. Appl. Artif. Intell.*, vol. 26, no. 7, pp. 1689–1697, 2013, doi: 10.1016/j.engappai.2013.03.014.

[24] H. I. Erdal, O. Karakurt, and E. Namli, "High performance concrete compressive strength forecasting using ensemble models based on discrete wavelet transform," *Eng. Appl. Artif. Intell.*, vol. 26, no. 4, pp. 1246–1254, 2013, doi: 10.1016/j.engappai.2012.10.014.

[25] Y. Yu, W. Li, J. Li, and T. N. Nguyen, "A novel optimised self-learning method for compressive strength prediction of high performance concrete," *Constr. Build. Mater.*, vol. 184, pp. 229–247, 2018, doi: 10.1016/j.conbuildmat.2018.06.219.

[26] Q. Han, C. Gui, J. Xu, and G. Lacidogna, "A generalized method to predict the

compressive strength of high-performance concrete by improved random forest algorithm," *Constr. Build. Mater.*, vol. 226, pp. 734–742, 2019, doi: 10.1016/j.conbuildmat.2019.07.315.
[27] M. Pala, E. Özbay, A. Öztaş, and M. I. Yuce, "Appraisal of long-term effects of fly ash and silica fume on compressive strength of concrete by neural networks," *Constr. Build. Mater.*, vol. 21, no. 2, pp. 384–394, 2007, doi: 10.1016/j.conbuildmat.2005.08.009.
[28] B. Ahmadi-Nedushan, "An optimized instance based learning algorithm for estimation of compressive strength of concrete," *Eng. Appl. Artif. Intell.*, vol. 25, no. 5, pp. 1073–1081, 2012, doi: 10.1016/j.engappai.2012.01.012.
[29] J. Zhang and Y. Zhao, "Prediction of compressive strength of ultra-high performance concrete (UHPC) containing supplementary cementitious materials," *Proc. - 2017 Int. Conf. Smart Grid Electr. Autom. ICSGEA 2017*, vol. 2017-Janua, pp. 522–525, 2017, doi: 10.1109/ICSGEA.2017.150.
[30] M. Saridemir, I. B. Topçu, F. Özcan, and M. H. Severcan, "Prediction of long-term effects of GGBFS on compressive strength of concrete by artificial neural networks and fuzzy logic," *Constr. Build. Mater.*, vol. 23, no. 3, pp. 1279–1286, 2009, doi: 10.1016/j.conbuildmat.2008.07.021.
[31] M. Uysal and H. Tanyildizi, "Predicting the core compressive strength of self-compacting concrete (SCC) mixtures with mineral additives using artificial neural network," *Constr. Build. Mater.*, vol. 25, no. 11, pp. 4105–4111, 2011, doi: 10.1016/j.conbuildmat.2010.11.108.
[32] M. I. Khan, "Predicting properties of High Performance Concrete containing composite cementitious materials using Artificial Neural Networks," *Autom. Constr.*, vol. 22, pp. 516–524, 2012, doi: 10.1016/j.autcon.2011.11.011.
[33] M. I. Khan, "Mix proportions for HPC incorporating multi-cementitious composites using artificial neural networks," *Constr. Build. Mater.*, vol. 28, no. 1, pp. 14–20, 2012, doi: 10.1016/j.conbuildmat.2011.08.021.
[34] J. S. Chou, C. F. Tsai, A. D. Pham, and Y. H. Lu, "Machine learning in concrete strength simulations: Multi-nation data analytics," *Constr. Build. Mater.*, vol. 73, pp. 771–780, 2014, doi: 10.1016/j.conbuildmat.2014.09.054.
[35] R. Rebouh, B. Boukhatem, M. Ghrici, and A. Tagnit-Hamou, "A practical hybrid NNGA system for predicting the compressive strength of concrete containing natural pozzolan using an evolutionary structure," *Constr. Build. Mater.*, vol. 149, pp. 778–789, 2017, doi: 10.1016/j.conbuildmat.2017.05.165.
[36] D. Nasr, B. Behforouz, P. R. Borujeni, S. A. Borujeni, and B. Zehtab, "Effect of nano-silica on mechanical properties and durability of self-compacting mortar containing natural zeolite: Experimental investigations and artificial neural network modeling," *Constr. Build. Mater.*, vol. 229, p. 116888, 2019, doi: 10.1016/j.conbuildmat.2019.116888.
[37] W. E. Elemam, A. H. Abdelraheem, M. G. Mahdy, and A. M. Tahwia, "Optimizing fresh properties and compressive strength of self-consolidating concrete," *Constr. Build. Mater.*, vol. 249, p. 118781, 2020, doi: 10.1016/j.conbuildmat.2020.118781.
[38] A. Iqtidar *et al.*, "Prediction of compressive strength of rice husk ash concrete through different machine learning processes," *Crystals*, vol. 11, no. 4, 2021, doi: 10.3390/cryst11040352.
[39] R. H. Faraj, A. A. Mohammed, A. Mohammed, K. M. Omer, and H. U. Ahmed, "Systematic multiscale models to predict the compressive strength of self-compacting concretes modified with nanosilica at different curing ages,"

Eng. Comput., 2021, doi: 10.1007/s00366-021-01385-9.
[40] O. O. Akin, A. Ocholi, O. S. Abejide, and J. A. Obari, "Prediction of the Compressive Strength of Concrete Admixed with Metakaolin Using Gene Expression Programming," *Adv. Civ. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/8883412.
[41] M. Saridemir, "Prediction of compressive strength of concretes containing metakaolin and silica fume by artificial neural networks," *Adv. Eng. Softw.*, vol. 40, no. 5, pp. 350–355, 2009, doi: 10.1016/j.advengsoft.2008.05.002.
[42] C. S. Poon, S. C. Kou, and L. Lam, "Compressive strength, chloride diffusivity and pore structure of high performance metakaolin and silica fume concrete," *Constr. Build. Mater.*, vol. 20, no. 10, pp. 858–865, 2006, doi: 10.1016/j.conbuildmat.2005.07.001.
[43] H. S. Wong and H. A. Razak, "Efficiency of calcined kaolin and silica fume as cement replacement material for strength performance," *Cem. Concr. Res.*, vol. 35, no. 4, pp. 696–702, 2005, doi: 10.1016/j.cemconres.2004.05.051.
[44] P. Muthupriya, K. Subramanian, and B. G. Vishnuram, "Prediction of Compressive Strength and Durability of High Performance Concrete By Artificial Neural Networks," *Int. J. Optim. Civ. Eng.*, vol. 1, pp. 189–209, 2011.
[45] R. M. Ferreira, J. P. Castro-Gomes, P. Costa, and R. Malheiro, "Effect of metakaolin on the chloride ingress properties of concrete," *KSCE J. Civ. Eng.*, vol. 20, no. 4, pp. 1375–1384, 2016, doi: 10.1007/s12205-015-0131-8.
[46] S. Wild, J. M. Khatib, and A. Jones, "Relative strength, pozzolanic activity and cement hydration in superplasticised metakaolin concrete," *Cem. Concr. Res.*, vol. 26, no. 10, pp. 1537–1544, 1996, doi: 10.1016/0008-8846(96)00148-2.
[47] K. B. Park, R. S. Lin, Y. Han, and X. Y. Wang, "Model-based methods to produce greener metakaolin composite concrete," *Appl. Sci.*, vol. 11, no. 22, 2021, doi: 10.3390/app112210704.
[48] A. A. Ramezanianpour and H. Bahrami Jovein, "Influence of metakaolin as supplementary cementing material on strength and durability of concretes," *Constr. Build. Mater.*, vol. 30, pp. 470–479, 2012, doi: 10.1016/j.conbuildmat.2011.12.050.
[49] C. S. Poon, S. C. Kou, and L. Lam, "Influence of recycled aggregate on slump and bleeding of fresh concrete," *Mater. Struct. Constr.*, vol. 40, no. 9, pp. 981–988, 2007, doi: 10.1617/s11527-006-9192-y.
[50] S. C. Kou, C. S. Poon, and H. W. Wan, "Properties of concrete prepared with low-grade recycled aggregates," *Constr. Build. Mater.*, vol. 36, pp. 881–889, 2012, doi: 10.1016/j.conbuildmat.2012.06.060.
[51] C. S. Poon, Z. H. Shui, L. Lam, H. Fok, and S. C. Kou, "Influence of moisture states of natural and recycled aggregates on the slump and compressive strength of concrete," *Cem. Concr. Res.*, vol. 34, no. 1, pp. 31–36, 2004, doi: 10.1016/S0008-8846(03)00186-8.
[52] R. Kumutha and K. Vijai, "Strength of concrete incorporating aggregates recycled from demolition waste," *J. Eng. Appl. Sci.*, vol. 5, no. 5, pp. 64–71, 2010.
[53] S. C. Kou, C. S. Poon, and M. Etxeberria, "Influence of recycled aggregates on long term mechanical properties and pore size distribution of concrete," *Cem. Concr. Compos.*, vol. 33, no. 2, pp. 286–291, 2011, doi: 10.1016/j.cemconcomp.2010.10.003.
[54] F. López Gayarre, C. López-Colina Pérez, M. A. Serrano López, and A.

Domingo Cabo, "The effect of curing conditions on the compressive strength of recycled aggregate concrete," *Constr. Build. Mater.*, vol. 53, pp. 260–266, 2014, doi: 10.1016/j.conbuildmat.2013.11.112.

[55] S. R. Salimbahrami and R. Shakeri, "Experimental investigation and comparative machine-learning prediction of compressive strength of recycled aggregate concrete," *Soft Comput.*, vol. 25, no. 2, pp. 919–932, 2021, doi: 10.1007/s00500-021-05571-1.

[56] G. Andreu and E. Miren, "Experimental analysis of properties of high performance recycled aggregate concrete," *Constr. Build. Mater.*, vol. 52, pp. 227–235, 2014, doi: 10.1016/j.conbuildmat.2013.11.054.

[57] S. Boudali, B. Abdulsalam, A. H. Rafiean, S. Poncet, A. Soliman, and A. Elsafty, "Influence of fine recycled concrete powder on the compressive strength of self-compacting concrete (Scc) using artificial neural network," *Sustain.*, vol. 13, no. 6, 2021, doi: 10.3390/su13063111.

[58] W. Wongkeo, P. Thongsanitgarn, A. Ngamjarurojana, and A. Chaipanich, "Compressive strength and chloride resistance of self-compacting concrete containing high level fly ash and silica fume," *Mater. Des.*, vol. 64, pp. 261–269, 2014, doi: 10.1016/j.matdes.2014.07.042.

[59] T. C. Nwofor and C. Ukpaka, "A Mathematical Model for Prediction of Metakaolin-Silica Fume High Strength Concrete," vol. 7, no. 5, pp. 137–142, 2017, doi: 10.5923/j.jce.20170705.02.

[60] A. Oner, S. Akyuz, and R. Yildiz, "An experimental study on strength development of concrete containing fly ash and optimum usage of fly ash in concrete," *Cem. Concr. Res.*, vol. 35, no. 6, pp. 1165–1171, 2005, doi: 10.1016/j.cemconres.2004.09.031.

[61] I. J. Han, T. F. Yuan, J. Y. Lee, Y. S. Yoon, and J. H. Kim, "Learned prediction of compressive strength of GGBFS concrete using hybrid artificial neural network models," *Materials (Basel).*, vol. 12, no. 22, 2019, doi: 10.3390/ma12223708.

[62] C. Duran Atiş and C. Bilim, "Wet and dry cured compressive strength of concrete containing ground granulated blast-furnace slag," *Build. Environ.*, vol. 42, no. 8, pp. 3060–3065, 2007, doi: 10.1016/j.buildenv.2006.07.027.

[63] A. C. Ganesh and M. Muthukannan, "Development of high performance sustainable optimized fiber reinforced geopolymer concrete and prediction of compressive strength," *J. Clean. Prod.*, vol. 282, p. 124543, 2021, doi: 10.1016/j.jclepro.2020.124543.

[64] A. Oner and S. Akyuz, "An experimental study on optimum usage of GGBS for the compressive strength of concrete," *Cem. Concr. Compos.*, vol. 29, no. 6, pp. 505–514, 2007, doi: 10.1016/j.cemconcomp.2007.01.001.


SRIT