TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

PULCHOWK CAMPUS

THESIS NO: T07/078

Prediction of Optimum Bitumen Content in Asphalt Mix Design Using Artificial Neural Network

by

Moti Ram Giri

A THESIS

SUBMITTED TO THE DEPARTMENT OF CIVIL ENGINEERING

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE IN TRANSPORTATION ENGINEERING

DEPARTMENT OF CIVIL ENGINEERING

LALITPUR, NEPAL

JULY, 2024

COPYRIGHT

TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

PULCHOWK CAMPUS

DEPARTMENT OF CIVIL ENGINEERING

The undersigned certify that they have read and recommended to Institute of Engineering for acceptance, a thesis entitled "**Prediction of Optimum Bitumen Content in Asphalt Mix Design Using Artificial Neural Network**" submitted by Moti Ram Giri in partial fulfillment of the requirement for degree of Master of Science in Transportation Engineering.

……

Supervisor: **Prof. Gautam Bir Singh Tamrakar**

Department of Civil Engineering

……

External Examiner: **Mr. Prabhat Kumar Jha**

Deputy Director General, Maintenance Branch,
Department of Roads, Government of Nepal

……

Committee Chairperson: **Mr. Anil Marsani**

Coordinator, M.Sc. in Transportation Engineering

Department of Civil Engineering

Date:

ABSTRACT

The Marshall design process, commonly employed for estimating Optimum Bitumen Content (OBC), is known for its designation as the asphalt mix design and quality control of asphalt concrete is often constrained by the conventional Marshall Mix Design methodology. Which is characterized by its time-intensive nature, labor requirements, and susceptibility to result variations. This study explores different predictive modeling techniques to enhance the efficiency and accuracy of determining Optimum Binder Content (OBC), Mechanical and Volumetric properties of hot mix asphalt. The study examines Multiple Linear Regression (MLR), Artificial Neural Networks (ANN), Support Vector Machine (SVM), and Random Forest (RF) models to predict mechanical (MST, MFV) and volumetric properties (AV, BSG, VFA, VMA) using variables such as aggregate gradation, specific gravity, and proportions of fine and coarse aggregates in the mix. A comprehensive dataset of 148 Marshall mix design forms was collected, and 141 valid sets were used after outlier analysis. Descriptive statistics revealed skewness, and compliance checks against SSRBW 2016 standards highlighted areas of concern in Marshall quotient and Filler-Binder Ratio. Centrality analysis showed significant deviations between mid-point gradation and sample mean for specific gradations.

The initial MLR model demonstrated moderate predictive capabilities with R^2 values ranging from 0.3017 to 0.6794 for various parameters, indicating limitations in accurately predicting complex asphalt properties. To address these limitations, the integration of ANNs significantly improved predictive accuracy, achieving higher R^2 values for all parameters compared to MLR.

Furthermore, the ANN model predicted OBC with R^2 of 0.7964 and other parameters with R^2 values between 0.7206 and 0.91, making it a robust tool for asphalt property estimation. ReLU - Linear and tanh - Linear activation functions as transfer function were identified as optimal for different parameters. SVM models also exhibited strong predictive capabilities with R^2 values up to 0.961 during training, although performance slightly declined during testing, suggesting potential overfitting. The RF model showed high predictive accuracy in training (R^2 up to 0.976) but had moderate generalization during testing.

Overall, ANN models emerged as the most effective, offering the highest accuracy and predictive performance across various parameters. The RF model demonstrated competitive performance in terms of Mean Absolute Error (MAE) for OBC and other properties. The research advocates for adopting ANN and RF models in asphalt mix design to reduce reliance on conventional methods, thereby saving time, resources, and labor while maintaining or enhancing the predictive reliability of asphalt mix properties.

Keywords: Hot Mix Asphalt, Marshall Mix Design, Optimum Binder Content, Artificial Neural Networks, Support Vector Machine, Random Forest, Predictive Modeling, Aggregate Gradation

ACKNOWLEDGEMENT

The significant progress of the research work is made possible through the unwavering motivation, guidance, and assistance of numerous individuals. I would like to express my sincere gratitude to my supervisor, Prof. Gautam Bir Singh Tamrakar, for their continuous support, motivation and invaluable guidance throughout the entire research period.

I would like to express my sincere gratitude to Asst. Professor Anil Marsani, Program Coordinator, M.Sc. in Transportation Engineering, Pulchowk Campus for his valuable support, guidance and coordination through the whole process to prepare this report. I would also like to express gratitude to Asst. Professor Dr. Pradeep Kumar Shrestha and Asst. Professor Dr. Rojee Pradhananga for their valuable suggestions and guidance throughout the work form proposal work to this point.

Moreover, I extend my appreciation to the Quality Research and Development Center, Department of Road for providing the research related data. Additionally, my heartfelt thanks go to Visow Lab Pvt. Ltd, Everest Lab Pvt. Ltd., Meh Geo Lab Pvt. Ltd., for providing the necessary Marshall Test Results required for this work. Their contributions are key in the successful execution of this research.

I am deeply thankful to all the classmates for their constant support. I extend my thanks to everyone who has been a part of this task including responders for literature review, data collection.

Name: **Moti Ram Giri**

Roll No.: **078/MSTrE/007**

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AC	Asphalt Concrete
ANN	Artificial Neural Network
AV	Percentage Air Void
BMs	Bituminous Mixes
BSG	Bulk Specific Gravity
C	Regression Coefficient
DBM	Dense Graded Bitumen Macadam
DNN	Deep Neural Network
Dof	Degree of Freedom
HMA	Hot Mix Asphalt
IRC	Indian Road Congress
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MFV	Marshall Flow Value
ML	Machine Learning
MLPNN	Multi-Layer Perceptron Neural Network
MLR	Multiple Linear Regression
MQ	Marshall Quotient

MSE	Mean Square Error
MST	Marshall Stability
NAS	Nominal Aggregate Size
NHL	Number of Hidden Layer
NLP	Natural Language Processing
NN	Number of Neuron
OBC/OAC	Optimum Bitumen Content /Optimum Asphalt Content
PD	Percentage Down
PP	Percentage Passing
R	Coefficient of Correlation
R^2/R^2	Coefficient of Determination
ReLU	Rectified Linear Unit
RF	Random Forest
RMSE	Root Mean Square Error
RSE	Relative Standard Error
SE	Standard Error
SG	Specific Gravity
SGB	Specific Gravity of Bitumen
SGF	Specific Gravity of Fine Aggregate
SSRBW	Standard Specification for Road and Bridge Work

SVM	Support Vector Machine
SVR	Support Vector Regressor
VFA	Void Filled with Asphalt
VMA	Voids in Mineral Aggregate
WC	Wearing Course

CHAPTER 1: INTRODUCTION

1.1 Background

Asphalt concrete stands as a widely employed pavement surface material globally, composed mainly of mineral aggregates, asphalt binder, and additives. Its porous structure, crafted at high temperatures, faces challenges related to the temperature sensitivity of the asphalt binder, leading to potential rutting and cracking in pavements. To enhance the longevity of bituminous roads, meticulous design, construction, and maintenance practices are imperative. The effectiveness of asphalt binder, a crucial adhesive component, dictates the overall performance of the asphalt pavement. Temperature fluctuations pose challenges, causing the binder to be soft in high temperatures and stiff in low temperatures, potentially leading to pavement distress. To mitigate such issues, precise mix design, specifically Marshall Mix Design, is recommended. Marshall Mix Design involves a labor-intensive process, with extensive testing to determine optimum bitumen content. The conventional method relies on physical testing of stability and flow values, followed by intricate calculations for various properties. This traditional approach is time-consuming and demands skilled labor.

Recognizing the inefficiencies of conventional methods, this research explores the application of Artificial Neural Networks (ANNs) in predicting crucial properties such as Marshall Stability and Flow values. ANNs, inspired by the human nervous system, excel in learning from data without making assumptions about the underlying mathematical representations. Machine Learning (ML), with their ability to establish empirical relationships, are employed to optimize bitumen content in asphalt mixtures. The complexity and non-linearity of Marshall Mix Design are addressed by using ANNs, SVMs, RFs MLR to predict key parameters, facilitating the optimization of bitumen content.

Machine learning, particularly ANNs, has found widespread applications in various fields, including transportation engineering. With its ability to formulate prediction models driven by historical data, machine learning is deemed essential in a data-driven field where relationships are often empirical. In transportation engineering, the empirical nature of relationships between stakeholders and decision parameters aligns well with the predictive capabilities of machine learning. The continuous evolution of machine learning techniques, coupled with the growing availability of big data in transportation, positions it as a promising tool for efficient and reliable predictions.

This research aims to harness the predictive power of ANNs to develop a robust model for determining the optimum bitumen content in asphalt concrete. By leveraging historical data and avoiding the need for extensive laboratory experiments, this approach seeks to streamline the mix design process, offering a more efficient and reliable solution for optimizing asphalt pavement properties.

1.2 Problem Statement

Asphalt concrete is made up of aggregates, asphalt binder, and air voids. Numerous studies have examined how various properties of asphalt concrete affect its performance. One of the key properties of asphalt mix that affect its performance is Optimum Bitumen Content (OBC). The Marshall Design method is widely used to estimate OBC for asphalt mixes, a process known as asphalt mix design. Despite its widespread use, this method is time-consuming, laborious as well as empirical, and its results can vary significantly (Othman, 2022). The Marshall Design method involves preparing three samples for each of the five different asphalt content levels to evaluate various strength and volumetric criteria. These tests help determine the best binder content (Asphalt Institute , 1988). Choosing the optimal binder content also relies on engineering judgment, considering factors like traffic, climate, and experience with local materials. As a result, there is a need for a more efficient and reliable method for predicting the optimum bitumen content in asphalt mixes.

The problem is that current methods are time-consuming and labor-intensive, and the goal is to develop a more efficient and accurate method that can predicting the optimum bitumen content in asphalt mix based on the historical data of Marshall Mix Design using machine learning algorithm. The stakeholders include construction companies, pavement engineers, and government agencies responsible for construction and maintaining roads and highways.

The potential impact is that a more efficient and accurate method for determining the optimum bitumen content could lead to fast and reliable determination of bitumen content saves time and cost of laboratory works for government agencies, construction companies and better road conditions for the public.

1.3 Objective of Study

The main objective of the study is to construct reliable prediction model for Marshall mix parameters in asphalt mix design. The specific objectives of the study are:

i. To develop the prediction model for the Optimum Bitumen Content in asphalt mix
ii. To develop the prediction model for volumetric properties of asphalt mix.
iii. To develop the prediction model for mechanical properties of asphalt mix.
iv. To test the effectiveness of ANN prediction model over Prediction model like MLR, RF and SVM

1.4 Scope of Study

In Nepal and south Asian region, almost all the contractor and government agencies use Marshall Mix Design Method for the determination of mix properties and it involves lot of labor, demand and consumes lot of resources. This research is done under the partial fulfillment of the requirement for the degree of Master of Science in transportation engineering. This research broadly includes following extent of work but not limited to following.

Literature review: An in-depth review of the existing research works on the use of ANN for the prediction of asphalt mixture properties like prediction of Marshall Flow value, Marshall Stability value etc., and research related to prediction model of optimum bitumen content of asphalt mix etc.

Data collection: Collection of data related to asphalt mixture such as bitumen content, aggregate properties, and performance parameters.

Model Development, Training and Testing: Development of models using various algorithms and activation functions such as support vector machine, random forest, random tree, Artificial Neurol Network etc. The developed models were evaluated using various performance metrics, including MAE, RMSE, R^2 and RSE.

Model Comparison: Comparison of the performance of the developed models with existing methods used for predicting optimum bitumen content

Optimization: Optimization of the best-performing model using various techniques such as feature selection, hyper parameter tuning, and assembling.

Model Implementation: Implementation of the optimized model on new data and prediction of optimum bitumen content.

Conclusion and Recommendations: Draw conclusion based on the results obtained from the study and provide recommendations for future research in the same field.

1.5 Organization of Report

This thesis work is structured into six chapters, as detailed below:

Chapter 1: Introduction chapter introduces the study, encompassing the background, problem statement, objectives, scope, and limitations of the study.

Chapter 2: Literature Review chapter offers an extensive overview of research on the Marshall mix design procedure, machine learning algorithms, and Artificial Neural Networks (ANNs). It explores the application of ANNs in addressing mix design problems.

Chapter 3: Methodology chapter outlines the methodology and approach used in the formation of the models, which form the basis for the analysis and evaluation of results. It describes the foundation of the proposed traditional prediction model and the Machine Learning model.

Chapter 4: Analysis, Results, and Discussion chapter presents the analysis of compliance, centrality, data processing, and both traditional and Machine Learning prediction models. It details the estimation procedures, calibration, and validation of the different models.

Chapter 5: Implementation Framework chapter discusses the development of the Graphical User Interface (GUI) and evaluates the error in the predictions made by the best-performing models.

Chapter 6: Conclusion and Recommendations chapter presents the key findings from the various models, drawing conclusions and making recommendations. It also suggests directions for future research.

1.6 Limitation of Study

i. The dataset comprises different Marshall mix designs, including DBM, AC NAS 19, AC NAS 13.2, and WC 20. This variability may lead to deviations in the desired outcomes due to differences in mix characteristics.

ii. All data utilized in this study were obtained from secondary sources, which may impact the reliability and validity of the findings.

iii. The Marshall Mix data were sourced from various laboratories, potentially leading to inconsistencies when used in a unified analytical context. The secondary nature of the data limits the assurance of consistency in the procedures, methodologies, and results related to the mix design.

iv. In instances of unavailability of the percentage passing of aggregate, values from preceding sieves were transferred to achieve continuous gradation. This adjustment may affect the gradation accuracy.

v. The mix combinations do not account for the compaction differences, with WC 20 mixes compacted for 50 blows and other mixes for 75 blows, potentially influencing the results.

vi. The viscosity grade or penetration grade of bitumen was not considered due to the limited availability of data, which could affect the mix design and performance outcomes.

vii. Key aggregate properties, such as water absorption, were not included in the analysis due to insufficient data, potentially impacting the study's comprehensive assessment of the asphalt mix properties.

CHAPTER 2: LITERATURE REVIEW

2.1 Background

Asphalt concrete is widely used worldwide as a common pavement material. It's made from mineral aggregates, asphalt binder, and additives, and is extensively used in road construction. It's essentially a blend of graded aggregates and hot asphalt applied and spread on roads. Asphalt concrete mixes is formulated at a very high temperature, around 180°C. It includes a mix of asphalt binder (bitumen), aggregate particles, and air spaces.–Asphalt binder, acting as an adhesive, significantly influences overall asphalt pavement performance. However, it reacts differently to temperature: it becomes soft in high temperatures and stiff in low temperatures. These reactions can lead to pavement distress such as rutting in hot weather and cracking in cold weather (Zhang, Chen, Zhu, & Wei, 2020). The life of a bituminous road can be prolonged through good design of pavement, construction, and maintenance practices. During design, engineers measure the traffic on a road, paying special attention to the number and types of trucks and recommend the pavement type. Different types of bituminous mixtures (BMs) used for road pavements must be well-designed as combinations of aggregates and bitumen to endure traffic loads and weather conditions. If these mixtures have inadequate mechanical and volumetric properties, they can lead to various pavement issues such as cracks from fatigue or low temperatures, permanent deformations, and stripping. These failures reduce the pavement's lifespan and pose safety risks for road users.

Consequently, it's important to understand how asphalt mixtures perform based on their ingredients to optimize their design. Currently, experimental methods are used to evaluate asphalt mix performance, which requires expensive lab tests and skilled labor. Any adjustment in mix composition, such as changing the bitumen content or type, or adjusting aggregate size, means more lab tests are needed, adding to costs and time in the design process. In Nepal, (Department of Road, 2016)has suggested Marshall Mix design for the evaluation of mix design.

Asphalt mix design typically uses the Marshall Mix Design method, which determines the composition of asphalt concrete. This method follows procedures outlined by the Asphalt Institute Marshall Mix Design Procedure. Bruce Marshall, a former engineer with the Mississippi State Highway Department, initially formulated the concepts behind the Marshall method for designing paving mixtures. Over time, the US Corps of Engineers enhanced and refined Marshall's test procedures through extensive research and correlation studies, ultimately establishing standardized mix design criteria. (Asphalt Institute , 1988). Since 1948, the test has been used by organizations and government departments in numerous countries, often with adjustments either to the procedure or how the results are interpreted (Tapkın, Çevik, & Uşar, 2010)

The Marshall test involves making cylindrical specimens, 102 mm in diameter and 64 mm height, using a standard compaction hammer and mold. Depending on road conditions, compaction is done with either 50 or 75 blows from a 4.53 kg hammer dropped from a height of 45.72 cm. These specimens are then tested for deformation resistance at 60°C, moving at a steady rate of 50.8 mm/min in a test rig. During testing, the specimen's stress distribution is complex because the jaws of the loading rig grip most, but not all, of its circumference, leaving the top and bottom unconfined (Asphalt Institute , 1988).

Two properties are determined during the Marshall test:

(a) Marshall stability, which indicates the maximum load the specimen can bear before failure.

(b) Marshall flow, which measures the amount of deformation the specimen undergoes before failure.

The Marshall Quotient, the ratio of stability to flow, serves as a measure of the material's resistance to permanent deformation, akin to a pseudo stiffness indicator. (Abdulhaq Hadi Abedali, 2014). The goal of the Marshall Design process is to find the best amount of asphalt binder. Designers achieve this by examining test curves to identify values that provide maximum stability, maximum specific gravity, and BC at 4% air voids. Then, the optimum asphalt binder content is determined by averaging these three values.

Determining the best amount of bitumen through the Marshall test takes a lot of time and skilled workers. After the test, we only get stability and flow values directly. Other important measurements like specific gravity, voids in mineral aggregate (VMA), voids filled with asphalt, and air voids need extra calculations. Using neural networks could make this process simpler by predicting these properties, which would make it easier to calculate the optimal bitumen content. Machine learning can also predict the bitumen content based on properties of mix ingredients and mix compositions.

Artificial Neural Networks (ANN) are models inspired by the human brain's neural network. They learn from past data that represents real-world phenomena or decision-making processes. ANNs consist of units organized into input, hidden, and output layers. One notable feature of ANNs is their ability to find relationships between input and output variables without needing a specific mathematical formula. They can extract complicated details and complex patterns from datasets. Unlike traditional regression models, ANNs excel in handling noisy data and do not rely on assumptions about how variables are related. (Sadiq, Rodriguez, & Mian, 2019). Neural Networks are versatile tools applicable to many tasks such as recognizing patterns, categorizing data, grouping similar items, simplifying data, interpreting images, understanding language, predicting outcomes, and analyzing trends (Yang, et al., 2021).

Artificial Neural Networks (ANNs) consist of input nodes, output nodes, and hidden nodes in between. Input nodes pass information to hidden nodes using activation functions, which determine whether hidden nodes activate based on the input. Hidden nodes then use weights to process this information, and when certain thresholds are met, outputs are generated. ANNs require extensive training with large datasets. However, they may not be suitable for rare or extreme events with limited data. ANNs also cannot incorporate human expertise to replace quantitative evidence Artificial Neural Networks (ANNs) can handle uncertainties by estimating the probability of each output, but how each output becomes most likely isn’t clear (ANNs are often seen as black boxes). The nodes in the hidden layers of ANNs lack physical interpretation, making it difficult to directly map outputs to processes. ANNs also require a predetermined path for making decisions on which information to process (inputs)

and how to classify it (outputs). To address nonlinearities in Marshall Mix Design, ANNs are used to create predictive models for Marshall Stability, Marshall Flow, and Optimum Bitumen Content. They also aid in optimizing bitumen content and improving decision-making in Marshal Mix design (Sadiq, Rodriguez, & Mian, 2019).

Artificial Neural Networks (ANNs) are designed to ensure that input data produces the desired output, either directly or through a gradual adjustment process. The connections, known as weights, within the network can be established based on existing knowledge or modified by exposing the network to learning patterns during training. Among the various training algorithms available, Reinforcement Learning has gained significant prominence in recent years due to its effectiveness. In this method, the network interacts with its environment, receives feedback based on its actions, and adjusts its parameters accordingly. This iterative process continues until the network achieves optimal performance, making it a valuable tool in modern neural network applications.

A Multi-layer Perceptron Neural Network (MLPNN) typically includes an input layer, at least one hidden layer, and an output layer. If it has more than two hidden layers, it's often referred to as a Deep Neural Network (DNN). Nodes in each layer are connected to nodes in adjacent layers through weights. These weights are adjusted during training to minimize model error. Various algorithms, like gradient descent, stochastic gradient descent, and Levenberg-Marquardt, are used for efficient training. (Shrestha & Mahmood, 2019).

The automation of features extractors is what differentiate a DNNs form earlier generation machine learning techniques (Shrestha & Mahmood, 2019). ANN is a type of neural network model consist of single hidden layer or multiple hidden layer that is trained with algorithms to learn representations form data sets without any manual design of features extractor. ANN with multi-layer is called Deep Neural Network. Deep learning may be of shallow learning model or deep learning model. Shallow is one with fewer layers of units. The shit form shallow to deep learning has allowed for more complex and non-linear functions to be mapped as they can't be efficiently mapped with shallow architectures.

As we all know, properties of asphalt concrete are unpredictable, and their relationships are empirical and hard to figure out what is it actually. Also, determination of different physical properties of asphalt concrete including Marshall Stability, Flow involves lot of effort, consumes time and money. In such context Artificial Neural Network Helps researcher to reduce effort and saves time by reliable and efficient prediction of such properties using historical data.

2.2 Development of Asphalt Concrete:

The use of asphalt in pavements dates back to ancient Babylon, where "Procession Street" was paved between 625 B.C. and 604 B.C. In the United States, the first asphalt pavements were constructed in Newark, New Jersey (1870), and Washington, D.C. (1876), along Pennsylvania Avenue.

In 1908, Clifford Richardson noted that asphalt mixtures were often haphazardly graded and proportioned without scientific basis. He authored the first asphalt pavement engineering textbook, "The Modern Asphalt Pavement," in 1905 while working at the Barber Asphalt Paving Company. Richardson emphasized the importance of air voids and voids in mineral aggregates (VMA). Meanwhile, Frederick Warren of the Warren Brothers Company introduced "bitulithic" mixes in 1901 and 1903, using larger stone aggregates tightly graded for stability and economic efficiency.

Prevost Hubbard and Frederick Field from the Asphalt Institute developed mix design procedures in the 1920s, focusing on fine mixes and establishing minimum stability criteria based on field performance against shoving. Francis Hveem introduced a mix design method in the 1930s that included the use of a kneading compactor to better simulate field conditions. Hveem also developed the Hveem Stabilimeter to assess mix performance under traffic shear forces. Bruce Marshall developed the Marshall Mix Design in the 1930s, advocating for asphalt content that maximizes density while meeting stability and flow

requirements. During World War II, the Marshall method was adapted by the U.S. Army Corps of Engineers for constructing durable runways under varying conditions.

2.3 Standard and Norms

Initially, Marshall Mix Design was regulated by (Department of Roads , 2001) and the related specifications for the gradation is given in Table 2-1 and limiting values for Marshall Variables is given in Table 2-2 Mix Mechanical and Volumetric Properties as per SSRBW 2001 Table 13.19**.**

Table 2-1 Gradation Requirements of the Mix SSRBW 2001

Sieve Size (mm)	Gradation I wearing course		Binder Course	
	Lower Limit	Upper Limit	Lower Limit	Upper Limit
	Percentage passing by weight			
20	-		100	100
16.000	100	100	96	100
10.000	76	82	75	100
4.750	54	59	48	90
2.360	33	37	24	62
0.600	16	21	12	24
0.075	6	9	3	8

Table 2-2 Mix Mechanical and Volumetric Properties as per SSRBW 2001 Table 13.19

Asphalt Concrete	Wearing course	Base Course
Properties	50 blows (N)	
Marshall Stability (N)	Min 8 000	Min 6 000
Flow Value (mm)	2-3	2-3
Voids in Total Mix (%)	2-4	3-7
Compression/Immersion Ratio	>0.75	>0.75

Later, on specification was reviewed and replaced by (Department of Road, 2016) which gives instructions on the necessary information that need to be collected before the preparation and execution of the Marshall Test for asphalt concrete. Since, our thesis is based on

the secondary data's, it plays vital role in the execution of prediction model based on the machine learning technique. The standard specification gives minimum standard of material to be used in the mix, their properties related to coarse aggregate (i.e. Cleanliness, particle shape, strength, durability, polishing, water absorption, stripping and Sensitivity), Fine aggregate, Bitumen and mineral Filler. The specification also suggested optimum range of aggregate grading and bitumen content for the mix based on the thickness of binder/surface course of pavement as specified in the Table 2-3.

Table 2-3 Gradation Requirement of Asphalt Concrete Mix as per SSRBW 2016

	Gradation I		Gradation II	
Nominal aggregate size*	19mm		13.2 mm	
Layer thickness	50mm		30-40mm	
Sieve Size (mm)	Lower Limit	Upper Limit	Lower Limit	Upper Limit
26.5	100	100		
19	90	100	100	100
13.2	59	79	90	100
9.5	52	72	70	88
4.75	35	55	53	71
2.36	28	44	42	58
1.18	20	34	34	48
0.6	15	27	26	38
0.3	10	20	18	28
1.18	5	13	12	20
0.075	2	8	4	10
Bitumen content % by mass of total mix *	Min 5.2*		Min 5.4*	

In this specification composition of Couse aggregate of Asphalt Concrete is well graded which is different form gap graded Dense Graded Bituminous Macadam. Also, the minimum asphalt content/Bitumen content % by mass of total mix is specified as 5.2%-5.4%. Corresponding to specific gravity of aggregate as 2.7 and Daily mean air temperature is 30 Degree Centigrade. Also, the limiting values for the mechanical and volumetric properties of the mix are highlighted in the specifications and shown in Table 2-4 and Table 2-5.

Table 2-4 Requirements for Dense Graded Bituminous Macadam

Properties	Viscosity Grade Paving Bitumen	Modified Bitumen		Test Method
		Hot climate	cold Climate	
Properties	75 blows on each face of the specimen			
Minimum stability (kN at 60 oC)	9	12	10	AASHTO T245
Marshall flow (mm)	2-4	2.5-4	3.5-5	AASHTO T245
Marshall Quotient	2-5	2.5-5		MS-2 and ASTM 02041
% air voids	3-5			
% Voids Filled with Asphalt (VFA)	65-75			

Table 2-5 Minimum Percentage Voids in Mineral Aggregate (VMA)

Nominal Maximum Particle Size (mm)	Minimum VMA Percent Related to Design Percentage Air voids		
	3	4	5
26.5	11	12	13
37.5	10	11	12

The bitumen content required shall be determined following the Marshall Mix design procedure contained in Asphalt Institute Manual MS-2. The Fines to Bitumen (F/B) ratio by weight of total mix shall range from 0.6 to 1.2.

Also, the specifications required the job mix formula to include the following details:

i. Source and location of all materials;

ii. Proportions of all materials specified clearly

iii. Type of binder and its percentage by weight in the total mix

iv. Percentage by weight of coarse aggregate, fine aggregate, and mineral filler in the total aggregate;

v. Specific percentages passing each sieve for the mixed aggregate

vi. Individual grading of each aggregate fraction and their proportions in the combined grading;

vii. Results from mix design tests including maximum specific gravity of loose mix (Gmm), densities of compacted specimens, Marshall stability and flow, VMA, VFA, along with relevant graphs, and AASHTO T 283 Moisture Susceptibility test results;

viii. For batch mixers, the weights of each aggregate type and binder per batch

ix. Test results for physical characteristics of aggregates used

x. Mixing and compacting temperatures specified

Manual for Dense Graded Bituminous Mixes (DBM/BC) published by (Maintenance Branch , 2018) is recommended to test five different asphalt contents with a single selected aggregate gradation to determine the optimum binder content. To ensure sufficient data, at least three test specimens should be prepared for each asphalt content. Therefore, a typical Marshall mix design using five asphalt contents will generally require at least 15 test specimens. It also recommended to record results as the average for three compacted, "identical" specimens. While selecting optimum binder content, it gives emphasis on engineering judgment depending on traffic, climate and experience with the local material used. This specification recommends using the Asphalt Institute Manual MS-2 to determine the asphalt content. It also states that the ratio of fines to bitumen by weight of the total mix should be between 0.6 and 1.2. Apart from the conformity with the grading and quality requirements the mixture must meet the different criteria as per (Department of Road, 2016)

2.4 Desired Mix properties

(Abdulhaq Hadi Abedali, 2014) describes different desired properties of the mix as mix design for Hot Mix Asphalt (HMA) involves selecting the right combination of aggregate and asphalt binder to achieve desired qualities, also suggested by (Roberts, Kandhal, Brown, Lee, & Kennedy, 1996) in the final product. These qualities includes:

2.4.1 Resistance to Deformation

HMA should maintain its shape under traffic loading. This is influenced by factors such as aggregate properties, gradation, and asphalt binder characteristics.

2.4.2 Resistance to Fatigue Cracking

HMA should withstand repeated loading without cracking. This depends on the amount and stiffness of the asphalt binder used.

2.4.3 Resistance to Low Temperature Cracking

HMA should resist cracking in cold temperatures, mainly determined by the stiffness of the asphalt binder.

2.4.4 Durability

HMA should not age prematurely during production and service life. This is affected by factors such as air voids and the thickness of the asphalt binder film around each aggregate particle.

2.4.5 Moisture Damage Resistance

HMA should withstand moisture penetration, which can degrade the mix. This relates to air voids and the properties of the aggregate.

2.4.6 Skid Resistance

The surface course of HMA should offer sufficient friction with vehicle tires. This can be compromised by certain aggregate characteristics or excessive asphalt binder content.

2.4.7 Workability

HMA must be easy to place and compact. This is influenced by factors such as aggregate texture, size, gradation, and the viscosity of the asphalt binder during mixing and placement.

2.5 Factor affecting different Behavior of Mix:

Dense-graded asphalt concrete (AC) consists of aggregates, mineral powders, and asphalt binders, with the aggregate particles uniformly graded and tightly interlocked. After compaction, dense-graded asphalt concrete typically has a void ratio of less than 10%.(Fang, Park, Singuranayo, Chen, & Li, 2019), (Ferreira, Babadopulos, Bastos, & Soares, 2020) (Khasawneh & Alsheyab, 2020). The mechanical and volumetric properties of Hot Mix Asphalt (HMA) are influenced by several factors, which include the type and properties of the asphalt binder, aggregate characteristics, mix design, and environmental conditions. Below are the key factors affecting the mechanical and volumetric properties of HMA

2.5.1 Asphalt Binder Properties

2.5.1.1 Viscosity and Penetration Grade

The viscosity of the asphalt binder affects the stiffness and flow characteristics of HMA. Binders with higher viscosity typically result in stiffer mixes. The penetration grade also

indicates the hardness of the binder, influencing the mix's resistance to deformation and cracking (Board., 2007).

2.5.1.2 Binder Content

The amount of binder in the mix influences the durability, flexibility, and resistance to moisture damage. Adequate binder content ensures proper coating of aggregates and affects the mix's ability to resist cracking and rutting (Huang, 2004)

2.5.2 Aggregate Characteristic

2.5.2.1 Gradation

The size distribution of aggregates impacts the density, stability, and void content of the mix. Well-graded aggregates tend to produce denser and more stable mixes, whereas poorly graded aggregates can lead to higher void content and reduced stability. The proportion of aggregates in the mix affects the optimum bitumen content. Finer gradations increase the total surface area of the aggregates, requiring more binder to coat them evenly. Conversely, coarser mixes with less surface area demand less asphalt. (Roberts, Kandhal, Brown, Lee, & Kennedy, 1996), (Park, Cocconcelli, & Chun, 2022).

2.5.2.2 Shape and Texture

Rough-textured angular aggregates provide better interlock and friction, leading to increased stability and resistance to deformation compared to rounded aggregates (Kandhal, Evaluation of baghouse fines in bituminous paving mixtures, 1980).

2.5.2.3 Moisture Content

The moisture content of aggregates during mixing affects the adhesion between the binder and aggregates. Excessive moisture can lead to stripping and reduced durability (Hicks, 1991).

2.5.3 Mix Design Parameters

2.5.3.1 Air Voids

The percentage of air voids in the mix influences its density and durability. Optimal air void content ensures sufficient compaction and longevity, reducing the potential for moisture damage and cracking (Board., 2007).

2.5.3.2 VMA (Voids in Mineral Aggregate) and VFA (Voids Filled with Asphalt):

Proper VMA and VFA levels ensure adequate space for binder in the mix, balancing durability and resistance to deformation ((Abdulhaq Hadi Abedali, 2014)).

2.5.4 Compaction

2.5.4.1 Compaction Effort

The level of compaction effort during construction affects the density and void structure of the HMA. Proper compaction ensures the mix achieves the desired mechanical properties, including stiffness and resistance to deformation (Board., 2007)

2.5.5 Temperature

2.5.5.1 Mixing and Compaction Temperature:

The temperatures at which mixing and compaction occur influence the workability and compaction of the mix. Proper temperature control is essential for achieving the desired density and mechanical properties (Abdulhaq Hadi Abedali, 2014).

2.5.5.2 Service Temperature:

The performance of HMA is temperature-dependent. High temperatures can lead to rutting and low temperatures can cause thermal cracking. The choice of binder grade is often based on the expected range of service temperatures Roberts

2.5.6 Additives and Modifiers

2.5.6.1 Polymers and Other Modifiers

The use of polymer-modified binders can enhance the mechanical properties of HMA, improving its resistance to deformation, cracking, and moisture damage (Huang, 2004).

2.5.6.2 Antistripping Agents

Additives like lime and liquid antistripping agents improve the adhesion between the binder and aggregates, enhancing the mix's resistance to moisture-induced damage (Hicks, 1991).

2.5.7 Environmental Factors

2.5.7.1 Moisture Exposure

HMA's resistance to moisture damage depends on its ability to prevent water infiltration and maintain binder-aggregate adhesion. Proper drainage and mix design can mitigate moisture-related issues (Hicks, 1991)., (Zhu, et al., 2020).

2.5.7.2 Aging

The aging of asphalt binder over time affects the mix's stiffness and brittleness. Long-term aging can lead to increased susceptibility to cracking and reduced flexibility. (Huang, 2004)

2.6 MLR overview and used as prediction model

Multiple linear regression is a widely used statistical method for predicting the outcome of a dependent variable based on the values of two or more independent variables. It is an extension of linear regression and is commonly applied in various fields for prediction modeling and the analysis of complex relationships between variables. The technique is based on the assumption of a linear relationship between the dependent and independent variables, and it enables analysts to determine the variation of the model and the relative contribution of each independent variable in the total variance (Uyanık & Güler, 2013).

Tranmer and Elliot (2008) provide an overview of multiple linear regression analysis, presenting the necessary theory and examples of regression analysis. The course is intended to offer a practical outline, avoiding complicated algebra where possible and discussing important issues related to the assumptions and interpretation of results (Jobson, 2012)

A study by Sakarya University Education Faculty students (2013) utilized data for multilinear regression analysis to understand the relationship between variables. The study demonstrates the practical application of multiple linear regression in educational research. Furthermore, the concept of the multiple linear regression model and its applications are extensively outlined in a chapter by Springer Texts in Statistics The chapter by (Coker, 1995) describes the relationship between the dependent variable and a set of explanatory variables, and it presents a data set to be used as a multiple regression example. Multiple linear regression finds coefficients for a linear equation to predict a dependent variable's value. It identifies the best-fitting line or surface by minimizing the difference between predicted and actual values using the 'least squares' method. The goal is to estimate the dependent variable based on the independent variables.

2.7 Support Vector Machine

Support Vector Regression (SVR) is a powerful and flexible method within the family of Support Vector Machines (SVMs) used for regression tasks. It extends the principles of

SVM classification to regression problems, maintaining the core concept of maximizing the margin while accommodating regression's continuous output nature.

The roots of SVR trace back to the development of Support Vector Machines (SVMs), introduced by (Vapnik, 1964) in the 1960s and further developed in the 1990s. Their seminal work laid the foundation for what would become one of the most influential techniques in machine learning, particularly for classification tasks.

Furthermore, (Drucker, Burges, Kaufman, Smola, & Vapnik, 1996) introduced the concept of Support Vector Regression (SVR). Their work adapted the SVM framework for regression analysis, providing a robust method for predicting continuous outcomes while leveraging the advantages of SVM's margin maximization. Also, they have proposed the regression loss function and optimization objective function as:

$$\textit{minimize} \quad \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{\ell} (\xi_i + \xi_i^*)$$

$$\textit{subject to} \quad \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

SVR employs the ϵ-insensitive loss function, which ignores errors within a margin of ϵ around the actual value. This approach allows for a sparse solution, as many data points will not contribute to the loss function if they fall within the margin.

Kernel Trick was mentioned by (Smola & Schölkopf, 2004) stated SVR's strengths to handle non-linear relationships through the kernel trick. By applying a kernel function $K(x_i,x_j)=\phi(x_i)\cdot\phi(x_j)$ SVR can perform linear regression in a transformed feature space, enabling it to model complex, non-linear patterns without explicitly computing the transformation

Numerous improvements and extensions to SVR have been proposed over the years. Method proposed by (Suykens & Vandewalle, 1999) include the development of fast

training algorithms, integration with other machine learning techniques, and enhancements in kernel functions to improve performance and efficiency. For instance, Least Squares SVR (LS-SVR) suggested by simplifies the optimization problem, making it more computationally efficient while maintaining high accuracy.

In conclusion, Support Vector Regression has evolved from the foundational concepts of Support Vector Machines to become a vital tool in regression analysis. Its robustness, flexibility, and ability to handle non-linear data make it an indispensable method in the machine learning toolbox. Ongoing research and development continue to expand its applications and improve its efficiency, ensuring its relevance in the rapidly advancing field of data science.

2.8 Random Forest

Random Forest is an ensemble learning method for classification and regression that builds multiple decision trees during training and outputs the mode of the classes (classification) or mean prediction (regression) of the individual trees. It was introduced by Leo Breiman and Adele Cutler in the early 2000s and has since become a popular and effective machine learning technique due to its simplicity, robustness, and high performance.

The concept, presented by (Breiman, Bagging predictors, 1996) of ensemble methods, which combine multiple models to improve performance, predates Random Forest. The idea is rooted in statistical and machine learning research from the 1970s and 1980s, with early methods including Bayesian averaging, bagging, and boosting. In 2001, concept of Random Forest was introduced by *Leo Breiman* as an extension of the bagging method. By incorporating the concept of random feature selection in the construction of decision trees, *Breiman* enhanced the diversity among the individual trees in the ensemble, leading to improved predictive performance and reduced overfitting. Random Forest builds on the concept of decision trees, which are simple, interpretable models that partition the feature space into regions associated with different target values. However, individual decision trees are prone to overfitting, especially when they grow deep and capture noise in the training data.

In Random Forest training algorithm suggested by (Breiman, Random forests, 2001) each tree is trained on a bootstrap sample of the data, and at each split, a random subset of features is considered. This random feature selection introduces additional diversity into the ensemble and helps to decorrelate the trees, improving the overall performance. Several advancements have been made to improve the efficiency and performance of Random Forest. Techniques presented by (Liaw, Wiener, & others, 2002) techniques for handling missing data, scaling the algorithm to large datasets, and enhancing interpretability through variable importance measures and partial dependence plots.

Moreover, Numerous comparative studies conducted by (Cutler, Cutler, & Stevens, 2012)*;* (Fernández-Delgado, Cernadas, Barro, & Amorim, 2014) have demonstrated the effectiveness of Random Forest compared to other machine learning methods. It often outperforms traditional methods like linear regression and decision trees, particularly in situations with non-linear relationships and complex interactions among variables. In conclusion, Random Forest has become a staple in the machine learning community due to its robustness, flexibility, and ease of use. Its ability to handle high-dimensional data, provide accurate predictions, and offer insights into feature importance makes it an invaluable tool for both practitioners and researchers. Ongoing developments and applications continue to expand its capabilities and ensure its relevance in the rapidly evolving field of data science.

2.9 Previous Study on prediction model

(Haykin, 1998) in his work, provided a strong basis for understanding Neural Networks and introduced various methods for designing deep learning structures. Configuring deep networks with more than three layers has proven highly effective in training neural networks. Artificial Neural Networks (ANN) are algorithms that mimic the human brain's neural system. ANNs consist of neurons that communicate through connections called synapses. In recent years, ANNs have been increasingly used for prediction in pavement engineering. For instance, as early as 1995, Cal used an ANN to predict soil classification based

on factors like plasticity index, water content, and liquid limit. In 1998, Roberts and Attoh-Okine applied an ANN to forecast pavement condition. Similarly, (Attoh-Okine, 2001) proposed an ANN that considers pavement characteristics along with external factors such as weather, age, and traffic conditions to predict pavement condition.

(Ozgan, 2011) introduces a method using artificial neural networks (ANN) to predict Marshall Stability of asphalt concrete, revealing stability decreases with temperature and exposure time. The study utilizes ANN to predict Marshall Stability based on physical properties, exposure times, and environmental temperatures leveraging 65 asphalt core samples from a specific highway section in Turkey. The training and testing split involved 52 (80%) samples for training and 13 (20%) for testing, with a logarithmic sigmoid transfer function utilized in the hidden and output layers. The paper specifies optimum hyperparameters such as a learning rate of 0.001, momentum of 0.1, and a stopping criterion of 5,000 for training. It successfully correlates experimental results with the ANN model by R-Squared Value of 0.969 for training set and 0.933 for testing set, aiding in designing more resilient asphalt pavements. The study concludes that ANN effectively captures the nonlinear behavior of asphalt concrete stability, suggesting its potential applicability in hot mix asphalt designs,

In the study conducted by (Tapkın, Çevik, & Uşar, 2010) explored the application of neural networks (NN) to predict Marshall test results for polypropylene (PP) modified asphalt mixtures. The research demonstrates the improved properties achieved through the addition of PP fibers. Notably, a novel NN model is developed to predict stability, flow, and Marshall Quotient values, providing a practical tool for asphalt mixture design. The neural networks are employed to predict Marshall stability, flow, and Marshall Quotient values based on variables such as P type, PP%, bitumen%, specimen height (mm), unit weight (kg/m3), V.M.A. %, VFA % and AV%. The best training algorithm was Levenberg-Marquardt back-propagation. A hyperbolic tangent sigmoid function was used for the hidden layer, and a log sigmoid function was used for the output layer. The testing (20%) and training (80%) sets for NN training procedure are selected randomly form dataset. Furthermore, the study identifies optimal hyperparameters and performance parameters for the ANN models, such as the NN architecture (8-5-1 with 8 Input variables 5 hidden neurons and one output at a

time), the Levenberg-Marquardt backpropagation algorithm and Hyperbolic tangent sig-moid and log sig-moid transfer functions were found to be optimum training algorithm and transfer functions for the hidden and output layers. The statistical parameters used for the training of ANN is Mean Absolute Percentage Error (MAPE) or Root Mean Squared Error (RMSE) on and performance of model is evaluated based on the R-Squared Value. The R^2 values for stability, flow, and Marshall Quotient (MQ) on the training set are 0.99, 0.93, and 0.94, respectively, and on the testing, set are 0.94, 0.71, and 0.81, respectively. In conclusion, the study successfully predicts Marshall test results for PP modified asphalt mixtures using NNs.

The use of ANN models to predict the Marshall parameters of polyethylene-modified bitu-minous mixtures marks a significant advancement in asphalt mix design. The study con-ducted by (Khuntia, Das, Mohanty, & Panda, 2014), utilizes 90 laboratory observations, encompassing the percentages of CA, FA, filler material (F), waste polyethylene (P), and bitumen (B) to predict stability value(SV), flow value (FV), and air voids (AV). The dataset is divided into 63 training observations (70%) and 27 testing observations (30%). The best ANN setup for predicting stability was a 5-5-1 configuration. It used hyperbolic tangent and linear functions in the hidden and output layers, respectively. This achieved a high correlation of 0.936 in training and 0.93 in testing. For flow value prediction, the optimal architecture was 4-4-1 with log-sigmoid and linear transfer functions, yielding correlation coefficients of 0.96 for both training and testing. The prediction of air voids also utilized a 4-4-1 architecture with hyperbolic tangent sigmoid and linear transfer functions, achieving strong correlations ($R = 0.97$ for training and $R = 0.96$ for testing). The study set a stopping criterion of 2000 epoch for training stage across all networks, with a learning rate = 0.07 and a momentum = 0.1. The Levenberg-Marquardt backpropagation algorithm was identi-fied as the optimal training method, ensuring high accuracy and robust performance. The models were tested for reliability and accuracy using measures like RMSE and MAE. The Neural Network (NN) model was the most accurate, so it is recommended for practical use in the field.

The research by (Ozturk, Saglik, Demir, & Gungor, 2016) introduces an advanced Artificial Neural Network (ANN) model for predicting Hot Mix Asphalt (HMA) volumetrics properties in Marshall mix design, aiming to enhance the efficiency of mix design processes. Employing Levenberg-Marquardt backpropagation and with trial-and-error method model was optimized for one hidden layer and 20 neuron such that the model achieves impressive R^2 values of 0.94 for for Gmm (theoretical maximum specific gravity of the loose mixture) and 0.92 for Gmb(bulk specific gravity) for training dataset and R^2 values are 95% and 91 %for Gmm and Gmb, respectively for testing dataset. Initially, dataset of 835 mix was prepared and data set was then divided into 90% training and 10 % testing dataset designs, incorporating 12 factors such as gradation of mix, bulk specific gravity of aggregates (Gsb), and binder content of the mix. The study suggests that the developed ANN model can serve as a valuable tool for both design and quality control in road construction projects, streamlining mix design processes and contributing to overall efficiency in the industry.

(Zavrtanik, Prosen, Tušar, & Turk, 2016) utilize artificial neural networks for modeling air void content in aggregate mixture. This study uses artificial neural networks to model the relationship between different asphalt properties and air void content in aggregate mixtures. The approach involves feed-forward neural networks with error back-propagation and multiple linear regression. For the development of the models, seven different mixes are evaluated based on the asphalt mixes during the period from 1998 to 2009. The data collected was then studies based on binder content, sieve analysis, maximum density of aggregate and air void content in the mix. Analysis was carried out based on the combination of all seven mixes together such that 1634 input–output data sets were used for training and 180 input– output data sets for testing artificial neural networks. For the analysis NTR2003 was utilized and 6 different ANN model was employed with different hidden layer varying form 2-3 and neuron layer varied from 5 to 10 neurons per layer with learning rate varied from 0.5 to 0.1 and maximum number of iterations was varied from 10000 to 5000. Model configured with 7-10-15-20-1 produces R^2 value of 0.803 and RMSE of 0.41. Also, Multiple linear regression model was fitted with R^2= 717 and RMSE =48 for air void content concluding ANN model is more effective than the linear model

(Ivica & Ivan, 2017) had proposed ANN and MLR models for predicting Hot Mix Asphalt (HMA) properties based on the14 independent parameters namely 11 gradation parameter, binder content, density of the mix, density of the mineral mixture for air void and 13 parameters for binder content, aiming to reduce reliance on costly laboratory tests. Using a dataset of 386 HMA mixes such that 336 were used for training and validation and 50 for 186 evaluation purposes, the study focuses on predicting air void and binder content, essential indicators of asphalt quality. From the total number of mixes (386), 6.7% is of the 207 NAS of 32 mm, 25.5% of NAS 22 mm, 33.1% of the NAS of 16 208 mm, 20.7% - 11 mm, 0.8% - 8 mm, and 13.2% of discontinuous gradation emphasizing potential use of combined different mix gradation on single pot. For the development of ANN model, 1 to 3 hidden layers with neuron varying from 5 to 30 each layer were used with epochs varying from 200 to 800 and learning rate varying from 0.2 to 0.8. The R-Value= 0.918,0.74and R-squared = 0.842,0.54 for air void content ANN prediction model for training set data and testing set respectively. Also, R-Value = 0.981,0.92 and R-Squared = 0.962,0.84 soluble binder content ANN prediction model for training and testing data set respectively. Also, MLR was employed to draw the relationship between dependent and independent variable and this model resulted R- value for air voids are 0.815,0.59 and for binder content 0.943,92 for training and testing dataset emphasizing use of ANN as efficient prediction model

(Baldo, Manthos, Pasetto, & others, 2018) employs Artificial Neural Networks (ANNs) to predict the mechanical behavior of asphalt concretes, specifically Marshall stability, flow, and quotient. The independent variables include bitumen type, bitumen content, filler-bitumen ratio, air voids, voids in the mineral aggregates, voids filled with bitumen, and the type of production process. The models achieved high correlation coefficients. Optimized through trial-and-error, the ANNs utilized a feedforward structure with one hidden layer and 10 neurons. The hyperbolic tangent transfer function was used for the hidden layer and a linear function for the output layer. The Levenberg–Marquardt algorithm was adopted for training, with a learning rate of 0.001 and a batch size of 32. The dataset consisted of diabase aggregates from various quarries, two types of bitumen (conventional and polymer-modified), and mix properties like bitumen content and volumetric properties, allowing for comprehensive modeling of asphalt concrete behavior under different conditions. Of the 90 datasets, 70% were used for training, 15% for validation, and 15% for testing. The model's

performance was evaluated using the correlation coefficient (R) and mean square error (MSE). The R values for stability, flow, and quotient during training and testing were (0.90, 0.91), (0.938, 0.936), and (0.976, 0.936), respectively. The study concludes by highlighting the effective application of ANN modeling in predicting asphalt concrete performance, providing practical closed-form equations.

(Othman, 2022) used artificial neural networks (ANNs) to predict Marshall Test parameters (OBC, MST, MFV, AV, VMA) using aggregate gradation as input. They tested multiple ANNs to find the best hyperparameters, including activation functions, number of hidden layers, and neurons per layer. Heat maps were used to compare each ANN's performance. The results varied by parameter, showing that optimal hyperparameters differed. Deep neural networks achieved higher accuracy than shallow ones, with R values of 0.91 for OBC, 0.8 for stability, 0.53 for flow, 0.65 for density, 0.77 for air voids, and 0.66 for voids in mineral aggregate. The study concluded that linear activation functions were more effective than sigmoidal or Tanh functions.

(Cheng, Liu, Yaohui, Wang, & Wang, 2023) introduces a computational framework employing machine learning (ML) techniques to predict binder activity in reclaimed asphalt pavement (RAP) samples with high precision. It employs three ML methods—6th-degree multivariate polynomial regression with regularization, artificial neural network (ANN), and random forest regression (RFR). The study utilizes a dataset comprising 32 RAP samples tested by 17 laboratories, with each sample represented by a four-dimensional vector. ANN models are configured with parameters including layer count, units per layer, and a sigmoid activation function. On the other hand, RFR models are optimized with hyperparameters such as bootstrap = TRUE, max_depth = 40, max_features = sqrt, min_samples_leaf = 2, min_samples_split = 7, and n_estimators = 1000. Additionally, the paper evaluates model precision using statistical/performance parameters, employing MAE ranging from 12.2% to 12.8% on the test dataset. These parameters collectively contribute to assessing model effectiveness in predicting binder activity in RAP samples. Leveraging a diverse dataset and optimizing model parameters underscore the robustness and reliability of the predictive models developed in the study.

2.9.1 Conclusion of literature review on previous works

ANNs are effective in predicting different aspects of pavement engineering, including soil types, pavement condition, Marshall stability, and other properties of asphalt concrete. Studies by Cal (1995), (Roberts & Attoh-Okine, 1998) Roberts and Attoh-Okine (1998), and subsequent research have demonstrated the successful application of ANNs in this field. For instance, Ozgan (2011) and Tapkın et al. (2010) effectively utilized ANNs to predict Marshall stability and polypropylene-modified asphalt mix properties. Similarly, Khuntia et al. (2014) and Ozturk et al. (2016) confirmed the robustness of MLs in capturing the nonlinear behavior of asphalt mixtures. These studies underscore the potential of MLs to improve the design and performance evaluation of asphalt pavements, offering a reliable alternative to traditional empirical methods.

The main objective of this study is to assess how effectively machine learning (ML) can predict asphalt mix properties (both mechanical and volumetric) compared to traditional models like MLR. This aims to improve the laboratory mix design process. The Artificial Neural Networks (ANNs) created in this research will be trained and tested using laboratory data from Hot Mix Asphalts (HMAs) in Nepal.

2.10 Summary of Literature Review

The empirical nature, complexity, and variability in asphalt mix properties necessitate advanced predictive models to optimize pavement performance. Traditional methods such as Marshall Mix design to determine parameters like Marshall Stability (MST), Marshall Flow (MFV), Optimum Binder Content (OBC), and other volumetric properties, including Air Void (AV), Void in Mineral Aggregate (VMA), and Void filled with Asphalt (VFA), often fall short due to their labor-intensive and time-consuming nature. To address these limitations, recent research has focused on leveraging machine learning (ML) techniques and artificial intelligence (AI) to predict these parameters more accurately and efficiently.

Several studies have highlighted the utility of different ML algorithms in predicting asphalt mix properties. For instance, the development of hybrid AI approaches and the use of Support Vector Machine (SVM) algorithms have shown promise in predicting the Marshall parameters of Stone Matrix Asphalt. Key independent variables in these studies include aggregate gradation and types of bitumen. Artificial Neural Networks (ANNs), particularly feedforward backpropagation models with sigmoid, tanh, tangent sigmoid, ReLU activation functions, etc., have been extensively researched for their ability to predict the optimum bitumen content, stability, flow, density, air voids (AV), and voids in mineral aggregate (VMA). Studies have used various percentage passing sieve sizes and specific aggregate gradations as independent variables to predict different asphalt mix properties.

Further investigations into the prediction of hot asphalt mix properties using deep neural networks (DNNs) suggest that incorporating more hidden layers could enhance prediction performance. These studies, also utilizing gradation data from multiple sieve sizes, report promising results. Most of the ML models have predicted desired properties of the mix, and model performance was evaluated using different statistical matrices such as R, R Square, MAE, RMSE, MAE, and RSE. Moreover, in some of the studies, ANN models have shown superior performance over MLR in terms of lower RMSE and higher correlation coefficients, suggesting a robust approach for stability prediction in mixes.

Despite these advancements, there are still research gaps that need to be addressed in the context of Nepal. The major research gap is very few numbers of research contributing to the development of prediction models for different asphalt mix properties based on Marshall mix design in the context of Nepal. Another major gap is the need for further exploration of deep learning structures with more hidden layers to enhance predictive performance. Additionally, understanding the underlying physical processes through human expertise and engineering judgment remains crucial for the successful application of these advanced models. Therefore, this study contributes to this gap in the literature by developing a model to assess the different Marshall mix properties of the asphalt mix with the help of gradations of aggregate, specific gravity of aggregate and bitumen, and different percentages of course and fine aggregate in the mix.

CHAPTER 3: METHODOLOGY

3.1 Research Design

The proposed research design for this study is shown in the following Figure 3-1:

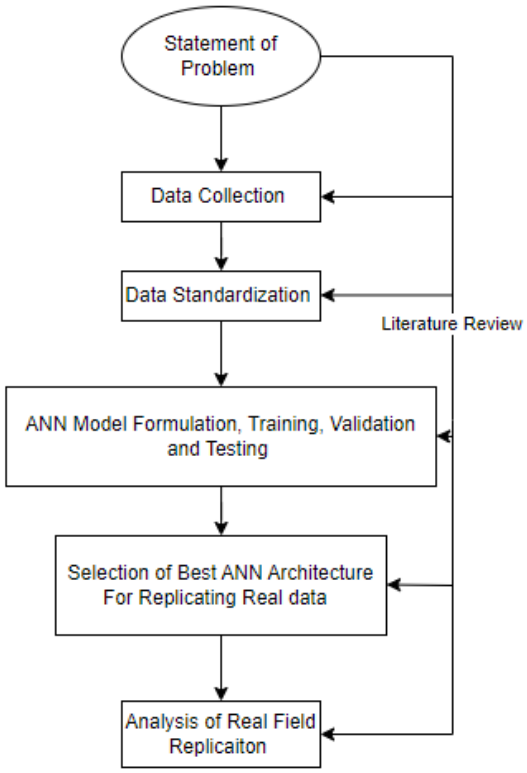


Figure 3-1 Research Design

The proposed research in Artificial Neural Network which is basically a quantitative approach and it is a computer algorithm that can accurately forecast and categorize data processing challenges. The above-mentioned research design if further elaborated in following flowchart in Figure 3-2. The major data required for the study are data related to Asphalt Mix Design as specified by the Standard Specification for Road and Bridge 2075, published by Ministry of Physical Infrastructure and Transport department of road. Data shall be

collected form government agencies, well certified laboratories, government projects, if required Marshall Test of mix design shall be performed in the laboratory after the collection of data, data shall be separated and standardized

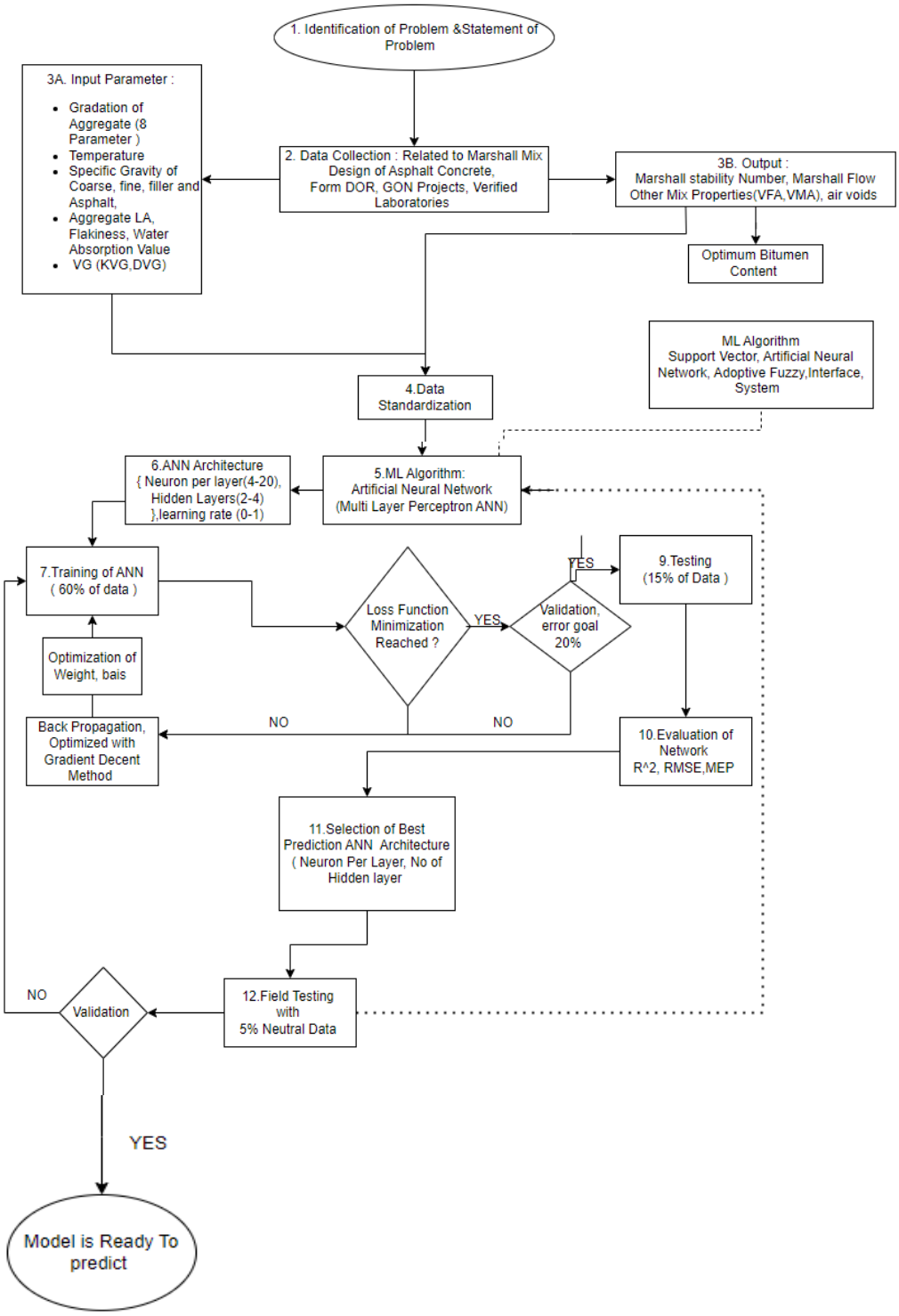
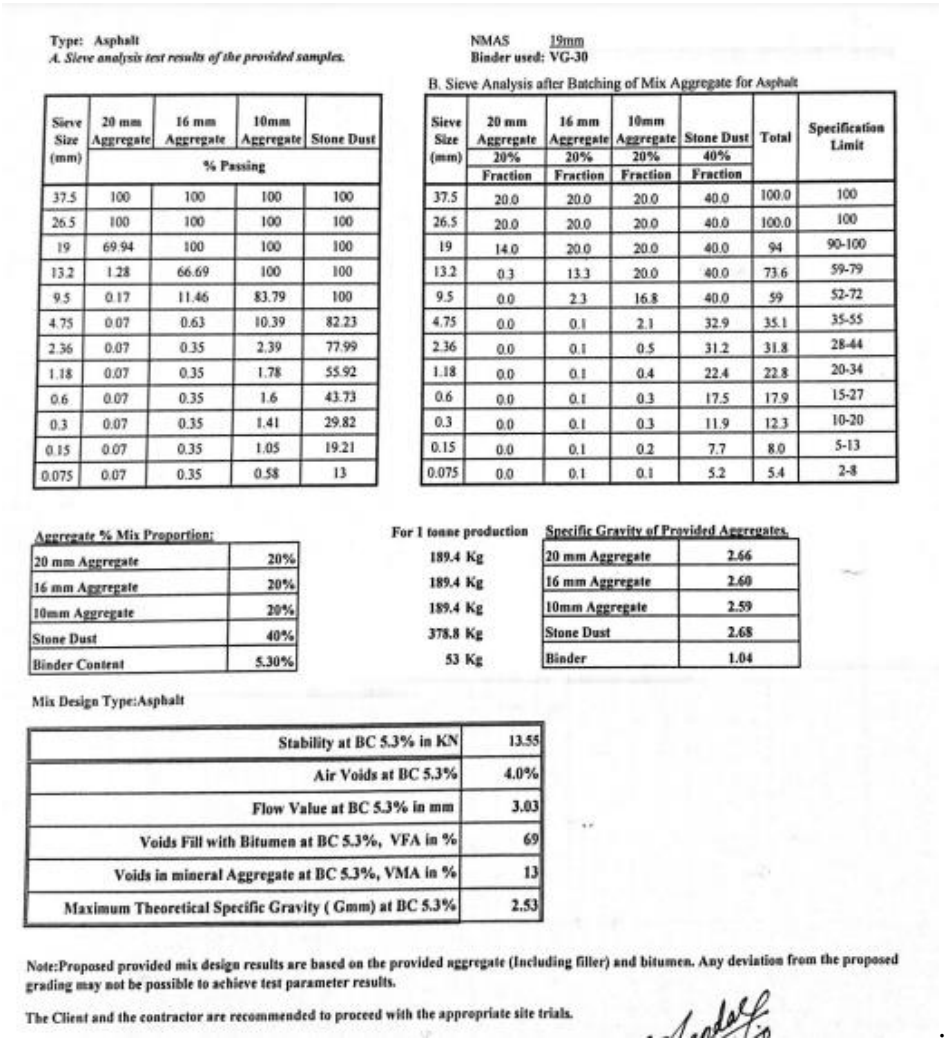


Figure 3-2 Elaborated form Research

3.2 Data Collection & Data Analysis

Data related to Marshall Test of mix design were collected from different laboratories all over the Nepal so that Prediction model works for whole Nepal is developed and represent diverse aggregate properties as much as gradation, specific gravity of materials confirming SSRBW-2001, SSRBW 2016(2075). Then, collected mix design example is shown in Figure 3-3 and Figure 3-4 and also properties and variables are summarized based on the variables presented in Table 3-1 with particular class.



Type: Asphalt
A. Sieve analysis test results of the provided samples.

NMAS 19mm
Binder used: VG-30

Sieve Size (mm)	20 mm Aggregate	16 mm Aggregate	10mm Aggregate	Stone Dust
	% Passing			
37.5	100	100	100	100
26.5	100	100	100	100
19	69.94	100	100	100
13.2	1.28	66.69	100	100
9.5	0.17	11.46	83.79	100
4.75	0.07	0.63	10.39	82.23
2.36	0.07	0.35	2.39	77.99
1.18	0.07	0.35	1.78	55.92
0.6	0.07	0.35	1.6	43.73
0.3	0.07	0.35	1.41	29.82
0.15	0.07	0.35	1.05	19.21
0.075	0.07	0.35	0.58	13

B. Sieve Analysis after Batching of Mix Aggregate for Asphalt

Sieve Size (mm)	20 mm Aggregate 20% Fraction	16 mm Aggregate 20% Fraction	10mm Aggregate 20% Fraction	Stone Dust 40% Fraction	Total	Specification Limit
37.5	20.0	20.0	20.0	40.0	100.0	100
26.5	20.0	20.0	20.0	40.0	100.0	100
19	14.0	20.0	20.0	40.0	94	90-100
13.2	0.3	13.3	20.0	40.0	73.6	59-79
9.5	0.0	2.3	16.8	40.0	59	52-72
4.75	0.0	0.1	2.1	32.9	35.1	35-55
2.36	0.0	0.1	0.5	31.2	31.8	28-44
1.18	0.0	0.1	0.4	22.4	22.8	20-34
0.6	0.0	0.1	0.3	17.5	17.9	15-27
0.3	0.0	0.1	0.3	11.9	12.3	10-20
0.15	0.0	0.1	0.2	7.7	8.0	5-13
0.075	0.0	0.1	0.1	5.2	5.4	2-8

Aggregate % Mix Proportion:

20 mm Aggregate	20%
16 mm Aggregate	20%
10mm Aggregate	20%
Stone Dust	40%
Binder Content	5.30%

For 1 tonne production

189.4 Kg
189.4 Kg
189.4 Kg
378.8 Kg
53 Kg

Specific Gravity of Provided Aggregates

20 mm Aggregate	2.66
16 mm Aggregate	2.60
10mm Aggregate	2.59
Stone Dust	2.68
Binder	1.04

Mix Design Type:Asphalt

Stability at BC 5.3% in KN	13.55
Air Voids at BC 5.3%	4.0%
Flow Value at BC 5.3% in mm	3.03
Voids Fill with Bitumen at BC 5.3%, VFA in %	69
Voids in mineral Aggregate at BC 5.3%, VMA in %	13
Maximum Theoretical Specific Gravity (Gmm) at BC 5.3%	2.53

Note:Proposed provided mix design results are based on the provided aggregate (Including filler) and bitumen. Any deviation from the proposed grading may not be possible to achieve test parameter results.

The Client and the contractor are recommended to proceed with the appropriate site trials.

Figure 3-3 Example of Marshall Mix Design Summary Sheet

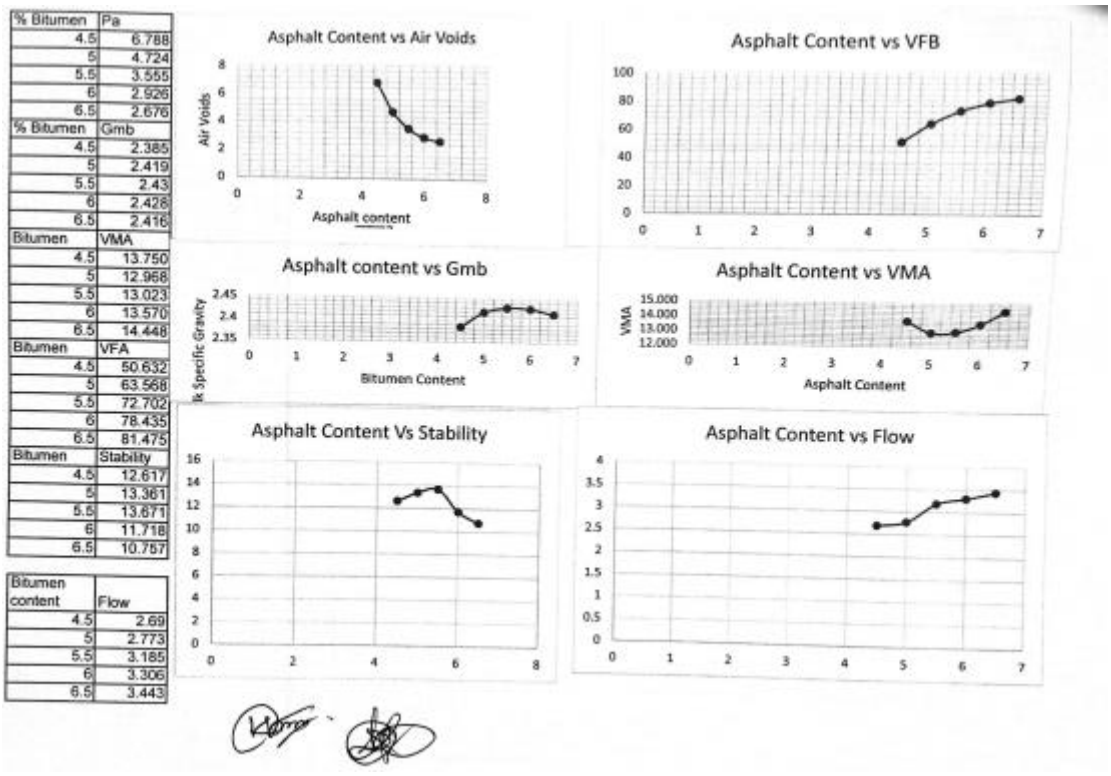


Figure 3-4 Marshall Mix Properties at different Bitumen Content

Table 3-1 Considered variables for this study

Designation	Name of variables	Class of Variables
PD 20	Percentage 26/20/19mm Down	Percentage of Coarse and Fine Aggregate
PD16	Percentage 16 mm Down	
PD 10	Percentage 13/10/9.5 mm Down	
PD 4.75	Percentage 4.75mm Down	
SG 20	Specific Gravity of 26/20/19 Down Aggregate	Specific Gravity Criteria
SG 16	Specific Gravity of 16 Down Aggregate	
SG 10	Specific Gravity of 13/10/9.5mm Down Aggregate	
SGF	Specific Gravity of 4.75 Down Aggregate	
SGB	Specific Gravity of Bitumen	
PP 26	Percentage Passing through 26mm Sieve	Gradation Criteria
PP 19	Percentage Passing through 19mm Sieve	
PP13.2	Percentage Passing through 13.2mm Sieve	
PP 9.5	Percentage Passing through 9.5mm Sieve	
PP 4.75	Percentage Passing through 4.75mm Sieve	
PP 2.36	Percentage Passing through 2.36mm Sieve	
PP 1.18	Percentage Passing through 1.18mm Sieve	
PP 0.6	Percentage Passing through 0.6mm Sieve	
PP 0.3	Percentage Passing through 0.3mm Sieve	
PP 0.15	Percentage Passing through 0.15mm Sieve	
PP 0.075	Percentage Passing through 0.075mm Sieve	

For the Analysis of the data, centrality analysis of gradation related to different mixes was evaluated and outlier analysis was conducted and extremely deviated values of dependent and dependent related independent variable, were removed from the data set (training and testing set). It also helps to avoid over fitting problem of the network during training and validation process. Also, descriptive statistics of data set was developed and correlation between different variables was established during this study.

3.3 Model Preparation / Model Formulation

3.3.1 Multiple Linear Regression

Multiple Linear Regression (MLR) is a statistical method used to evaluate the relationship between a dependent variable and multiple independent variables by fitting a linear equation to observed data. The general form of an MLR model is

$$Y = a\,Xi + b \qquad \textit{Equation } 1$$

where Y represents the predicted output, X_i are the independent variables, a is the matrix of regression coefficients, and b is the intercept.

The detailed equation is expressed as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon, \qquad \textit{Equation } 2$$

where y is the dependent variable,

x_1, x_2, ..., x_n are the independent variables,

β_0, β_1, ..., β_n are the regression coefficients, and ε is the residual error.

The model parameters are optimized to minimize the sum of the squares of the vertical deviations from each data point to the regression line, achieving maximum similarity between the predicted and actual values. MLR's popularity stems from its ease of interpretation and construction.

The effectiveness of MLR models is assessed using several statistical metrics, including the R, R^2, Adjusted R^2, MAE, and MAPE, which provide insights into the model's accuracy and predictive power.

3.4 Machine learning model

3.4.1 Artificial Neural Network

3.4.1.1 Mathematical Formulation for training and validation algorithm

Artificial Neural Networks (ANNs) can be mathematically formulated through a combination of linear algebra and calculus, particularly in the context of forward propagation and backpropagation. Here is an overview of the mathematical formulation of ANNs:

3.4.1.1.1 Network Architecture

An ANN consists of multiple layers of neurons, including the input layer, hidden layers, and output layer. Each neuron in a layer is connected to every neuron in the subsequent layer and is given in Figure 3-5.

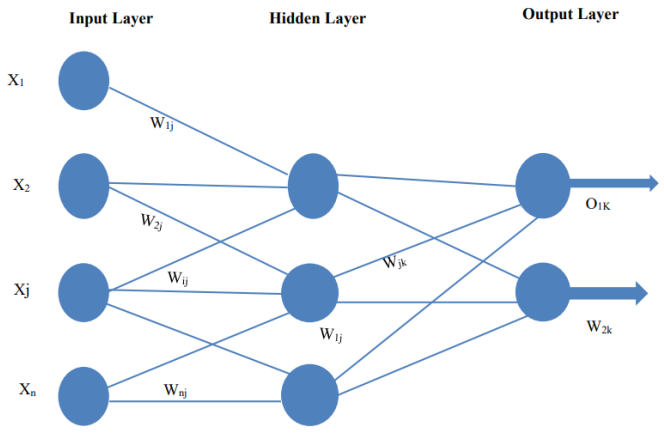


Figure 3-5 System Diagram of ANN

3.4.1.1.2 Forward Propagation

Forward propagation is the process of passing the input data through the network to obtain the output. It involves matrix multiplications and the application of activation functions.

3.4.1.1.3 Input Layer

Let $X \in R^{n \times m}$ be the input matrix where n is the number of features and m is the number of samples:

$$X = [x\_1, x\_2, \ldots, x\_m]\text{^}\top \qquad \textit{Equation 3}$$

Hidden Layers

For the l-th layer:

- Weights Matrix: $W^{(l)} \in R^{n_l \times n_{l-1}}$
- Bias Vector: $b^{(l)} \in R^{n_l}$

The output of the l-th layer $\left(a^{(l)}\right)$ is computed as:

$$z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)} \qquad \textit{Equation 4}$$

$$a^{(l)} = \phi\left(z^{(l)}\right) \qquad \textit{Equation 5}$$

where ϕ is an activation function (e.g., ReLU, Siqmoid).

3.4.1.1.4 Output Layer

The output layer applies a similar transformation:

$$z^{(L)} = W^{(L)} a^{(L-1)} + b^{(L)} \qquad \textit{Equation 6}$$

$$a^{(L)} = \phi\left(z^{(L)}\right) \qquad \textit{Equation 7}$$

3.4.1.1.5 Loss Function

The loss function quantifies the error between the predicted output and the actual target values. The loss function used in this regression model is Mean Squared Error (MSE). The loss function quantifies the error between the predicted output ($\hat{y}$) and the actual target values (y.

For Mean Squared Error (MSE):

$$L(y, \hat{y}) = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2 \qquad \textit{Equation 8}$$

3.4.1.1.6 Backpropagation and Gradient Calculation

Backpropagation is the process of computing gradients of the loss function with respect to the weights, using the chain rule of calculus. These gradients are used to update the weights to minimize the loss.

Gradient for Output Layer

$$\delta^{(L)} = \frac{\partial L}{\partial z^{(L)}} = \left(a^{(L)} - y\right) \circ \phi'\left(z^{(L)}\right) \qquad \textit{Equation 9}$$

Gradient for Hidden Layers

For $l = L-1, L-2, \ldots, 1$:

$$\delta^{(l)} = \left(W^{(l+1)}\right)^{\top}\delta^{(l+1)} \circ \phi'\left(z^{(l)}\right) \qquad \textit{Equation 10}$$

where $\circ$ denotes the element-wise multiplication (Hadamard product).

3.4.1.1.7 Weight Updates

The weights are updated using gradient descent or an optimization algorithm like Adam

$$W^{(l)} = W^{(l)} - \eta\frac{\partial L}{\partial W^{(l)}} \qquad \textit{Equation 11}$$

$$b^{(l)} = b^{(l)} - \eta \frac{\partial L}{\partial b^{(l)}} \qquad \textit{Equation 12}$$

Where, η is the learning rate

3.4.1.1.8 Optimization of Model

In this study, the Adam optimizer is employed for optimization, diverging from the Levenberg–Marquardt (LM) optimizer commonly used in previous research. The prior preference for LM was due to its effectiveness in single-layer or shallow multi-layer perceptron neural networks, where computing the Jacobian matrix is straightforward. However, given the focus on deep neural networks in this work, the Adam optimizer is more suitable due to its superior performance and adaptability in training complex models. Thus, Adam was chosen for its efficiency and robustness in handling the optimization of deep neural networks. A simple algorithm for the optimization process used is given below

Initialization:

- Initialize the time step t to 0.
- Initialize the exponential decay rates for the moment estimates β_1 and β_2 (typically set to 0.9 and 0.999, respectively).
- Initialize small constant ϵ to 10^{-8} to prevent division by zero.
- Initialize the first moment vector m_0 and the second moment vector v_0 to 0.
- Initialize the parameter vector θ_0.

Algorithm:

For each iteration:

- Increment the time step: t=t+1
- Compute the gradient $g(t)$ of the objective function with respect to the parameters θ.

- Compute gradients:

$$g_t = \frac{\partial L}{\partial \theta} \qquad \textit{Equation 13}$$

- Update biased first moment estimate:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \qquad \textit{Equation 14}$$

- Update biased second moment estimate:

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \qquad \textit{Equation 15}$$

- Compute bias-corrected first moment estimate:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \qquad \textit{Equation 16}$$

- Compute bias-corrected second moment estimate:

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \qquad \textit{Equation 17}$$

- Update parameters:

$$\theta_t = \theta_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \qquad \textit{Equation 18}$$

where:

θ represents the parameters (weights and biases).

η is the learning rate.

β_1 and β_2 are exponential decay rates for the moment estimate

ϵ is a small constant to prevent division by zero.

- Termination:

The algorithm terminates when the stopping criteria are met, which could be a specified number of iterations, convergence tolerance, or any other predefined condition.

3.4.1.2 Activation Functions

Activation function is a very important component of Artificial Neural Network. Activation function is inspired by activity in our brain. An Activation Function determines if a neuron should be activated. It decides if the input to the neuron is significant for predicting outcomes using basic mathematical operations. The Activation Function's job is to produce an output based on the input values provided to a node. Sigmoid, Tanh and ReLU are some of Activation Function. In this research, Sigmoid, Tanh and RelU activation function are used for the activation of Neural Network. Different activation function is used for selecting best ANN architecture and are explained as follow:

3.4.1.2.1 Sigmoidal Activation Function

his function accepts any real number as input and produces an output ranging from 0 to 1. Larger input values (more positive) result in outputs closer to 1.0, while smaller input values (more negative) yield outputs closer to 0.0. Mathematically, it can be described as,

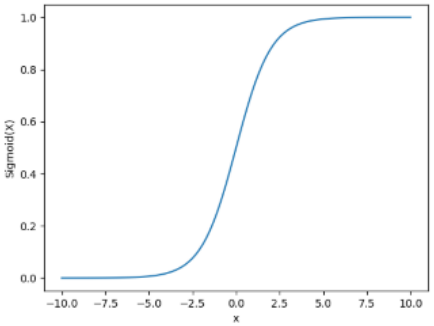


Figure 3-6 Sigmoidal Activation function

$$f(x) = \frac{1}{1 + e^{-x}} \qquad \textit{Equation 19}$$

3.4.1.2.2 Tanh Function

The Tanh function is similar to the sigmoid or logistic activation function, sharing the same S-shape shown in Figure 3-7 but with outputs ranging from -1 to 1. In Tanh, higher inputs (more positive) result in outputs closer to 1.0, while lower inputs (more negative) yield outputs closer to -1.0

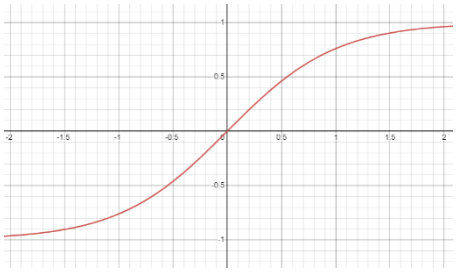


Figure 3-7 Tanh Activation Function

Mathematically it can be represented as:

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$ *Equation 20*

3.4.1.2.3 ReLU

It is not always true that activation function will always calculate output between 0 and 1, In fact ReLU transforms input in to 0 either input itself.So if the input is less than or equal to 0, then ReLU will output 0. If the input is greater than 0, ReLU will then just output the given input. ReLU is one of the most widely used activation functions.

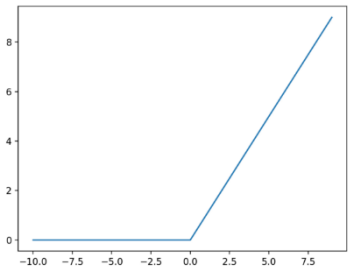


Figure 3-8 ReLU Activation Fucntion

3.4.2 Artificial Neural Network Model

To formulate the ANN model, the structure of the model and relationship between different variables were studied and analyzed. The causal relationships in the form of equation or graphical functions were formulated by data analysis. Based on the relevancy to the task and influence on the prediction model, variable was omitted or selected in the model framework shown in Figure 3-9. ANN model with different architecture were used. It was defined in terms of hidden layers, neuron per hidden layer, method of optimization of hyper parameter, weight and bias of the link and nodes, activation function and learning speed of the neural network.

The model of a Supervised Feed Forward Artificial Neural Network (ANN) for predicting the Marshall Properties is developed in python 10.8. The whole methodology involves data preparation, model training, and evaluation, with a specific focus on optimizing the

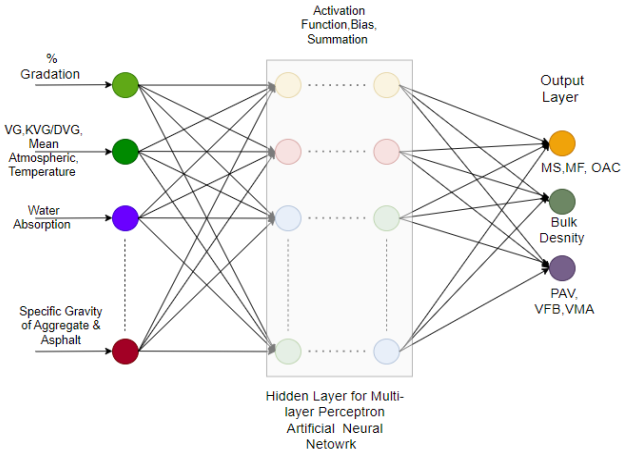


Figure 3-9 Feedforward ANN Architecture

model's performance through hyperparameter tuning. Typical architect of Multilayer Feed Forward Artificial Neural Network was presented as,

In the data preparation phase, independent variables are extracted from all columns except the dependent variable. Several target variables, including "Marshall Stability," "Marshall Flow Value," "Air Voids," "Void Filled with Asphalt," "Void in Mineral Aggregate," and "Bulk specific Gravity," are also extracted and their relationship is functionalized as below:

OBC = f (PD 20, PD 16, PD 10, PD 4.75, SG20, SG16, SG10, SGF, SGB, PP 26.50, PP 19, PP 13.20, PP 9.50, PP 4.75, PP 2.36, PP 1.18, PP 0.60. PP 0.30, PP 0.15, PP 0.075)	*Equation 21*
Mechanical Properties (Stability, Flow Value) = f (PD 20, PD 16, PD 10, PD 4.75, SG20, SG16, SG10, SGF, SGB, PP 26.50, PP 19, PP 13.20, PP 9.50, PP 4.75, PP 2.36, PP 1.18, PP 0.60. PP 0.30, PP 0.15, PP 0.075, BC)	*Equation 22*
Volumetric Properties (BSG, VMA, AV, VFA) = f (PD 20, PD 16, PD 10, PD 4.75, SG20, SG16, SG10, SGF, SGB, PP 26.50, PP 19, PP 13.20, PP 9.50, PP 4.75, PP 2.36, PP 1.18, PP 0.60. PP 0.30, PP 0.15, PP 0.075, BC)	*Equation 23*

The dataset is then split into training (80%) and testing (20%) sets using the *train_test_split* function from *sklearn.* This ensures that the model is trained on one portion of the data and evaluated on another, unseen portion to gauge its performance. Standardization is applied to the features using StandardScaler from sklearn. This process ensures that all features have a mean of 0 and a standard deviation of 1, which helps in faster convergence of the neural network during training.

For model training and validation, the Sequential class is utilized from tensor-flow.keras.models to build a neural network model. The input layer with number of neuron

units with different activation function such as ReLU (Rectified Linear Unit), tanh and sigmoid activation function as transfer function, is evaluated in order to evaluate their efficiency in handling the vanishing gradient problem. A loop structure is employed to varying the number of hidden layers (Nol) and neurons per layer (NN), with the hidden layers also using the different activation function as transfer function. The output layer has one unit with a linear activation function, as it is suitable for regression tasks where the output is continuous. The model is formulated using the Adam optimizer at learning rate of 0.001 which tries to minimize the loss function by adjusting the weights between the neurons in each iteration. The loss function used is MSE, which measures the average squared difference between the predicted and actual values. The model also tracks MAE as an additional performance metric. During training, 20% of the training data is set aside for validation using the validation_split parameter. Typical flowchart of Multilayer Feed Forward Back Propagating Artificial Neural Network is presented in Figure 3-10,

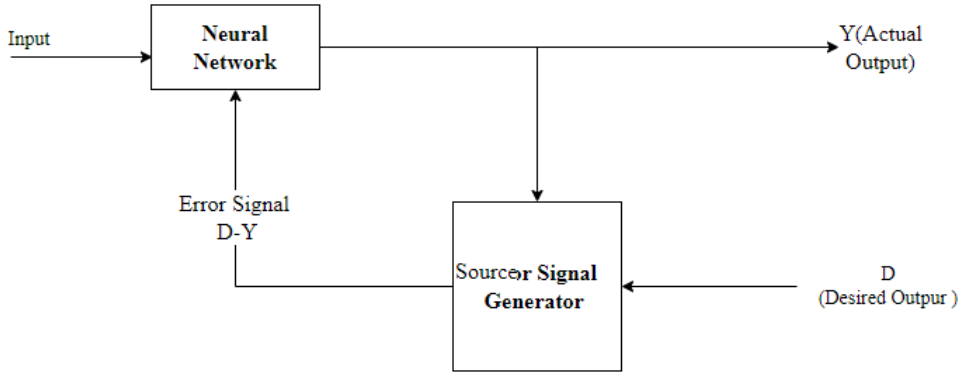


Figure 3-10 Training Process of Feedforward Back propagation ANN

For the selection of best fit ANN architect, thousands number of ANN architect with number of HL and NoN per each layer is used in rectangular architect and the optimum number of hidden layer and neurons in the hidden layer is determined with the help of trails and error method based on the least Mean absolute error. This helps in monitoring the model's performance on unseen data during training and prevents overfitting. The model is trained for 250 epochs with a batch size of 64, and the training history is recorded, which includes the loss and MAE for both training and validation sets.

For model evaluation, the model's performance is first evaluated on the training data to calculate loss and MAE. Predictions are made on the training data, and metrics such as R^2, R, RMSE, and RSE are calculated to assess the model's accuracy and fit. The model is then evaluated on the testing data to calculate loss and MAE. Predictions are made on the testing data, and similar metrics (R^2, R, RMSE, RSE) are calculated to evaluate the model's generalized performance.

Finally, the trained model for each configuration is saved to disk using the .h5 format. Performance metrics for both training and testing phases are saved to text files for the further analysis. The developed model implements a comprehensive approach to building and optimizing an ANN model for predicting the target variable.

3.4.3 Support vector Regressor

In this work, a comprehensive Support Vector Regression (SVR) methodology is employed for predicting multiple Marshall Properties, focusing particularly on minimizing the Mean Absolute Percentage Error (MAPE) for the Marshall. The process begins with data preprocessing, where independent and target variables are extracted from the dataset. The features include all input variables, while the target variables include stability, flow value, air voids, VFA, VMA, and bulk specific gravity.

The dataset is then split into training and testing sets, with 80% of the data splited for training and 20% for testing. Standardization of features is performed using StandardScaler to ensure that all input data is scaled properly, facilitating better model performance. The proposed model initializes an SVR model and defines a range of hyperparameters to search, including different kernel types (RBF, polynomial, sigmoid, and linear), various C values, and epsilon values. RBF kernel, which maps input features into an infinite-dimensional space to handle non-linear relationships; the Polynomial kernel, which models non-linear relationships by representing the similarity of vectors over polynomials of the original variables; the Sigmoid kernel, related to neural network activation functions and useful for

specific non-linear relationships; and the Linear kernel, which is the simplest and used for linear relationships between input features.

A grid search is conducted to identify the optimal hyperparameters. This involves iterating through all combinations of kernels, C values, and epsilon values. For each combination, an SVR model is created and fitted to the training data within a pipeline that includes feature scaling. The model's effectiveness is assessed using test data by predicting the target variable and calculating MAPE.

The process aims to find the combination of hyperparameters that results in the lowest MAPE. The best model parameters are stored, and the model is saved for future use. Additionally, the grid sear includes an early stopping mechanism, which terminates the search if the MAPE=0.005 falls below a predefined threshold, saving computational resources.

The methodology employed ensures a thorough exploration of different kernel functions and their respective hyperparameters, ultimately identifying the SVR model configuration that best captures the underlying patterns in the data. This approach leads to the most accurate predictions for the Marshall Properties included in the dataset.

3.4.4 Random Forest Method

The Random Forest Regressor. method employed here aims to develop a robust predictive model for the Marshall mix properties. The process is structured into several key phases to ensure the development of an accurate and reliable model. Initially, the data preparation phase involves extracting the independent variable and dependent variables from the dataset. Following data extraction, the dataset undergoes a train-test split using the train_test_split function from scikit-learn. The split ratio is set to 80:20, with 80% of the data allocated for model training and 20% reserved for testing. This partitioning strategy ensures an adequate balance between model training and evaluation on unseen data.

To standardize the features and enhance model performance, feature scaling is applied using the StandardScaler from scikit-learn. Standardization normalizes the feature values, ensuring they have a mean of 0 and a standard deviation of 1, which aids in algorithm

convergence and model.The model is initialized with default hyperparameters and a random state of 57. Hyperparameter tuning is conducted using a hyperparameter grid, specifying various values for parameters such as 'n_estimators,' 'max_depth,' 'min_samples_split,' and 'min_samples_leaf.' This grid search process systematically explores the hyperparameter space to identify the optimal configuration.To evaluate model performance and enhance generalization, K-Fold Cross-Validation with 5 folds is employed. The GridSearchCV technique is then utilized to perform an exhaustive search over the hyperparameter grid, selecting the combination that minimizes the negative mean absolute percentage error ('neg_mean_absolute_percentage_error').Upon identifying the optimal hyperparameters, the best model is fitted to the training data, and its performance is evaluated using metrics such as Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R2) on the test set. This rigorous evaluation ensures the model's ability to accurately predict the Marshall properties while maintaining reliability and generalization capabilities.

3.4.5 Testing & Performance Evaluation

The performance evaluation of the developed MLR and ML models, in a subset of training and validation, testing for prediction of OBC. MST, MFV, AV, VMA, VFA and BSG was assessed using number of standard analytical measuring tools. Includes such as mean absolute error (MAE), mean square error (MSE), mean absolute percentage error (MAPE), root mean square error (RMSE), coefficient of correlation (R), and coefficient of determination (R^2) were calculated as respectively.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} OV_i - PV_i \qquad \textit{Equation 24}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (OV_i - PV_i)^2 \qquad \textit{Equation 25}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{(OV_i - PV_i)^2}{OV_i} \qquad \textit{Equation 26}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(OV_i - PV_i)^2} \qquad \textit{Equation 27}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(OV_i - PV_i)^2}{\sum_{i=1}^{n}(OV_i - PV_{mean})^2} \qquad \textit{Equation 28}$$

Where, OV_i and PV_i represent the observed and predicted values of the variable, respectively. "n" is the total number of data points and OV_{mean} and PV_{mean} refer to the average model output value representing the observed and predicted values, respectively.

CHAPTER 4: ANALYSIS, RESULTS AND DISCUSSION

4.1 Overview

This chapter presents various data analysis techniques, including outlier analysis and correlation analysis. Additionally, the collected data are evaluated to check their compliance with SSRBW 2001 and SSRBW 2016 standards. Descriptive statistical parameters for the training and testing datasets are also provided. The chapter then introduces different models related to Marshall Properties developed in this study and assesses their performances. The evaluation models for OBC, MST, MFV, BSG, VMA, VFA, and AV are constructed using multiple linear regression and machine learning techniques such as Artificial Neural Networks (ANN), Support Vector Techniques, and Random Forests, as discussed in the previous chapter. Subsequent subsections detail the model development, training, validation, and testing results. Various sample sizes were used for the estimation models to establish the relationships between Marshall Properties and different variables. Comparisons between the multiple regression models and machine learning models are conducted to identify the most effective model for estimating the dependent variables. Model performance is evaluated using Mean Square Error (MSE), Root Mean Square Error (RMSE), the coefficient of determination (R^2), and the Coefficient of Correlation (R Value).

4.2 Data Processing

4.2.1 Data Composition

The Table 4-1 presented summarizes the various types of mix designs collected during a thesis work focused on road construction materials. It details the number of mix designs for each type, reflecting the extensive research and data collection undertaken. The types of mix designs included are Dense Bituminous Macadam (DBM), Airfield, Gradation I,

Gradation II, Wearing Course, and Binder Course. The number of designs varies significantly across these categories, with Gradation II having the highest number of mix designs at 58, indicating perhaps a more detailed exploration or a greater variation in this type. Gradation I follow closely with 52 mix designs, while the Wearing Course and DBM have 29 and 19 designs respectively. The Binder Course and Airfield mix designs are less frequent, with only 5 and 4 designs, suggesting either a narrower scope of study or less variability in those mixes. This comprehensive collection and categorization of mix designs form a crucial part of the thesis, providing a robust foundation for analyzing the properties and performance of different road construction materials.

Table 4-1 Composition of Dataset

Sn	Type of Mix	Number of Mix Design	Percentage of Mix
1	DBM	19	13%
2	Airfield	4	3%
3	Gradation I	52	35%
4	Gradation II	38	26%
5	Wearing Course	29	20%
6	Binder Course	5	3%
Total		147	100%

4.2.2 Outlier determination and box plotting

Outliers are extreme values that differ from most other data points in a dataset. They can have a significant impact on statistical analyses, making it essential to carefully identify potential outliers and handle them appropriately to ensure accurate results.

In the context of the thesis work, statistical outlier detection was employed to ensure the accuracy and reliability of the collected data on various mix design properties. This process involved converting data points into z-scores, which measure how many standard deviations away a value is from the mean of the dataset. Typically, values with a z-score greater than 3 or less than -3 are considered outliers. However, for this analysis, a more stringent criterion was applied, using a z-score threshold of 2 or -2 to identify outliers. This approach aimed to detect and exclude more subtle anomalies in the data.

The properties analyzed included OBC, MST, MFV, AV, BSG, VMA, VFA. The data collected, the identified outliers at the 95th percentile, and the results after applying the z-value method with the specified threshold are summarized in the Table 4-2 below. Additionally, the overall representation of the data was done using box plots in Figure 4-1, Figure 4-2, Figure 4-3, and Figure 4-4.

Table 4-2 Collected Data Summary and Data taken after outlier analysis

Properties	Data Collected	Outliers at 95 percentiles	Total Data	Test Set 20%	Training Set 80%
OBC	147	6	141	29	112
MST	572	28	544	109	435
MFV	573	18	555	111	444
AV	633	33	600	120	480
BSG	623	47	576	116	460
VMA	628	38	590	118	472
VFA	441	15	426	86	340

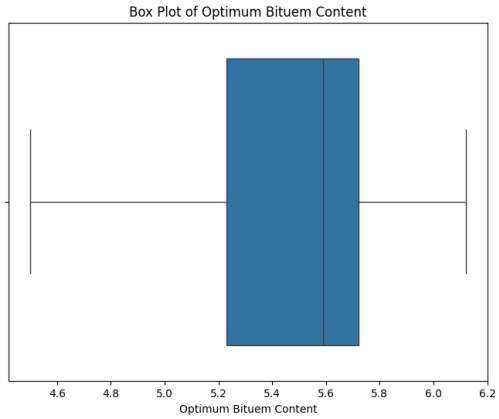


Figure 4-1 Box Plot of OBC dataset and outliers

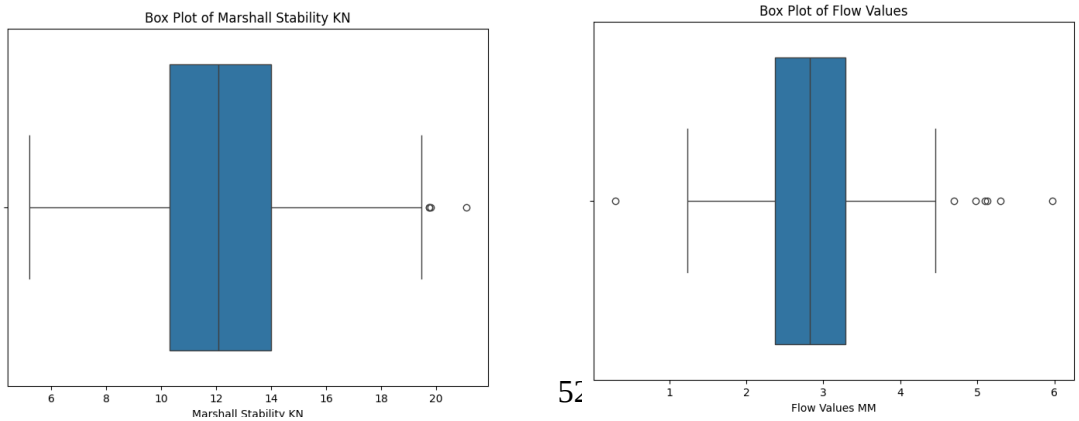


Figure 4-2 Box Plot of Marshall Stability and Marshall Flow value dataset and outliers

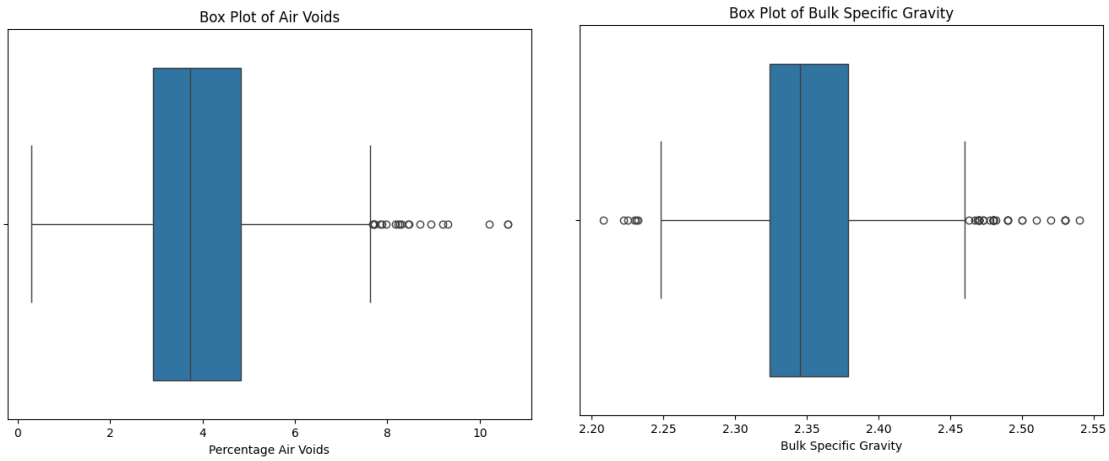


Figure 4-3 Box Plot of Percentage Air Voids and Bulk Specific Gravity dataset and Outliers

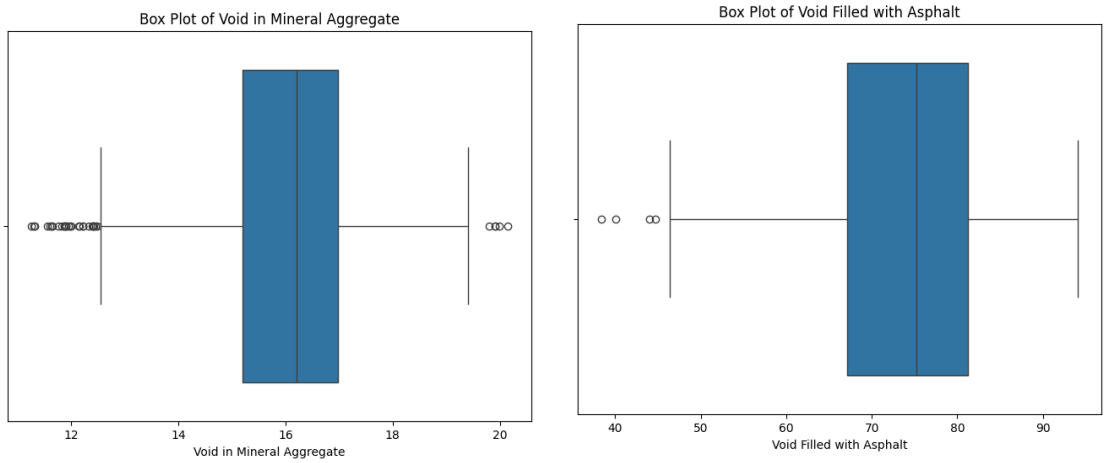


Figure 4-4 Box Plot of Void in Mineral Aggregate and Void Filled by Aggregate dataset Outliers

4.2.3 Descriptive Statistics

Descriptive analysis was conducted for datasets related to mixes based on SSRBW 2016 DBM, SSRBW 2016 NAS 26.5 mm, SSRBW 2016 NAS 19 mm, and SSRBW 2001 Wearing Course. The analysis was initially performed separately and presented in APPENDIX A: DESCRIPTIVE STATISTICS. Subsequently, the data were merged to create an overall dataset, and descriptive statistics for optimum bitumen content (OBC), as well as mechanical and volumetric properties, were calculated. The results for OBC for the combined dataset are presented in Table 4-6, while the results for mechanical and volumetric properties

are provided in APPENDIX A: DESCRIPTIVE STATISTICS, acknowledging that the sample sizes for each property differ.

Kurtosis measures the "tailedness" of the distribution, with positive kurtosis (leptokurtic) indicating more data in the tails and a sharper peak, and negative kurtosis (platykurtic) indicating less data in the tails and a flatter distribution. In our OBC dataset, variables such as SG16, SG10, SGF, SGB, and PP 26.50 exhibited high positive kurtosis, suggesting heavy tails and potential outliers. Conversely, variables like PD 20, PD 16, PP 13.20, and PP 4.75 displayed negative kurtosis, indicating lighter tails and flatter distributions.

Skewness measures the asymmetry of the distribution. Positive skewness indicates a right-skewed distribution, while negative skewness indicates a left-skewed distribution. In our dataset, variables such as SG16, SG10, SGF, and PP 26.50 exhibited high negative skewness, indicating a left-skewed distribution. On the other hand, variables such as SGB, PP 2.36, and PP 0.60 showed moderate to high positive skewness, indicating a right-skewed distribution.

The descriptive statistics of the OBC dataset reveal significant skewness and kurtosis, underscoring the necessity for preprocessing to ensure effective model training and robust performance. In our thesis work, we focused on addressing the skewed and flattened distributions observed in our dataset by employing normalization techniques, specifically using the StandardScaler. Normalization, achieved through StandardScaler, standardized the data to have a mean of zero and a standard deviation of one. This transformation ensured that all features contributed equally to the model training process, enhancing the stability and convergence of the ML models such as ANN, SVM, RF etc. By addressing the issues of scale and distribution asymmetry, we improved the suitability of the data for ANN modeling, leading to better performance and generalization capabilities.

Table 4-3 Statistical range of input and output data of OBC

Variables	Mean	Standard Error	Median	Mode	Standard Deviation	Sample Variance	Kurtosis	Skewness	Range	Minimum	Maximum	Count
PD 20	17.08	1.32	20.00	0.00	15.68	245.82	-1.22	0.29	51.00	0.00	51.00	141.00
PD 16	12.77	1.03	15.00	0.00	12.24	149.81	-1.45	0.20	35.00	0.00	35.00	141.00
PD10	26.73	0.83	30.00	30.00	9.80	96.12	1.37	0.04	65.00	0.00	65.00	141.00
PD 4.75	42.37	0.58	43.00	45.00	6.83	46.70	1.72	-0.60	45.00	15.00	60.00	141.00
SG20	2.65	0.00	2.66	2.63	0.04	0.00	2.02	0.50	0.22	2.56	2.78	141.00
SG16	2.64	0.00	2.64	2.63	0.06	0.00	16.31	-2.86	0.50	2.26	2.76	141.00
SG10	2.64	0.01	2.64	2.68	0.06	0.00	10.19	-2.22	0.50	2.26	2.76	141.00
SGF	2.66	0.01	2.66	2.69	0.06	0.00	6.55	-1.33	0.51	2.32	2.84	141.00
SGB	1.03	0.00	1.03	1.03	0.01	0.00	8.11	1.38	0.08	1.01	1.09	141.00
PP 26.50	99.82	0.07	100.00	100.00	0.84	0.71	70.54	-7.77	8.50	91.50	100.00	141.00
PP 19	97.00	0.43	100.00	100.00	5.08	25.77	2.53	-1.77	24.70	75.30	100.00	141.00
PP 13.20	80.84	0.82	78.69	91.10	9.72	94.43	-0.67	0.29	39.01	60.80	99.81	141.00
PP 9.50	69.60	1.11	70.90	81.97	13.24	175.18	-0.47	-0.67	49.73	39.10	88.83	141.00
PP 4.75	48.57	0.73	46.80	59.68	8.63	74.42	-1.02	0.28	34.60	35.10	69.70	141.00
PP 2.36	37.35	0.54	34.79	45.11	6.46	41.71	-0.43	0.79	30.00	24.10	54.10	141.00
PP 1.18	25.50	0.69	23.62	23.62	8.23	67.71	-0.78	0.32	33.19	8.60	41.79	141.00
PP 0.60	21.52	0.51	19.94	19.10	6.03	36.33	-0.13	0.62	26.30	8.60	34.90	141.00
PP 0.30	14.33	0.46	13.54	13.16	5.42	29.36	-0.83	0.22	19.67	5.40	25.07	141.00
PP 0.15	9.64	0.25	8.80	7.68	2.92	8.54	-1.04	0.40	11.34	4.60	15.94	141.00
PP 0.075	6.67	0.11	6.70	6.40	1.31	1.71	1.15	-0.52	7.95	2.02	9.97	141.00
OBC	5.50	0.03	5.60	5.70	0.32	0.10	-0.38	-0.69	1.33	4.79	6.12	141.00

4.3 Correlation analysis

This correlation table provides a comprehensive overview of the linear relationships between various properties within the dataset. Each cell in the table represents the correlation coefficient between two properties, ranging from -1 to 1. If a correlation coefficient is 1, it indicates a perfect positive correlation, meaning as one variable increases, the other variable increases in a perfectly linear fashion. A correlation coefficient of -1 indicates a perfect negative correlation, implying that as one variable increases, the other variable decreases in a perfectly linear fashion and correlation coefficient of 0 suggests no linear relationship between the variables. Table 4-4 and Table 4-5 show the correlation coefficients for different variables considered in the analysis.

4.3.1 Correlation for Optimum Bitumen Content

Table 4-4 Correlation of input parameters for OBC

	PD 20	PD 16	PD10	PD 4.75	SG20	SG16	SG10	SGF	SGB	PP 26.50	PP 19	PP 3.20	PP 9.50	PP 4.75	PP 2.36	PP 1.18	PP 0.60	PP 0.30	PP 0.15	PP 0.075	OBC
PD 20	1.0																				
PD 16	-0.8	1.0																			
PD10	-0.5	0.1	1.0																		
PD 4.75	-0.2	0.0	-0.3	1.0																	
SG20	0.1	-0.2	0.0	0.1	1.0																
SG16	-0.1	0.0	0.1	0.1	0.7	1.0															
SG10	-0.1	0.0	0.1	0.1	0.5	0.8	1.0														
SGF	0.1	-0.2	0.1	0.0	0.2	0.1	0.2	1.0													
SGB	0.3	-0.2	-0.1	0.0	0.2	0.1	0.0	0.1	1.0												
PP 26.50	-0.2	0.2	0.2	0.1	0.0	0.3	0.3	0.0	-0.2	1.0											
PP 19	-0.6	0.3	0.4	0.2	-0.1	0.2	0.3	0.0	-0.1	0.6	1.0										
PP 13.20	-0.5	0.2	0.3	0.4	-0.1	0.1	0.2	0.0	-0.2	0.3	0.6	1.0									
PP 9.50	-0.7	0.3	0.5	0.2	-0.2	0.1	0.2	0.0	-0.3	0.4	0.7	0.8	1.0								
PP 4.75	-0.5	0.1	0.3	0.5	0.0	0.1	0.2	0.0	-0.1	0.2	0.5	0.7	0.7	1.0							
PP 2.36	-0.5	0.2	0.2	0.5	-0.1	0.0	0.1	-0.1	-0.1	0.1	0.4	0.7	0.6	0.8	1.0						
PP 1.18	-0.4	0.1	0.1	0.5	-0.1	0.0	0.1	-0.1	-0.1	0.3	0.4	0.7	0.6	0.6	0.8	1.0					
PP 0.60	-0.5	0.1	0.2	0.5	-0.1	0.0	0.2	0.0	-0.1	0.3	0.5	0.7	0.6	0.7	0.8	0.9	1.0				
PP 0.30	-0.1	-0.1	-0.2	0.4	0.0	0.0	0.1	-0.1	0.0	0.1	0.1	0.5	0.2	0.5	0.6	0.8	0.8	1.0			
PP 0.15	-0.2	0.1	0.0	0.4	0.0	0.1	0.2	-0.1	-0.1	0.2	0.3	0.7	0.4	0.5	0.7	0.8	0.8	0.8	1.0		
PP 0.075	-0.2	0.1	0.1	0.2	0.1	0.1	0.1	0.0	-0.3	0.1	0.2	0.2	0.2	0.3	0.2	0.1	0.2	0.1	0.4	1.0	
OBC	-0.5	0.3	0.4	0.3	0.0	0.3	0.3	0.2	-0.1	0.4	0.7	0.5	0.7	0.5	0.4	0.4	0.5	0.1	0.3	0.2	1.0

4.3.2 Correlation For Mechanical and Volumetric properties.

Table 4-5 Correlation of input parameters for Mechanical (Marshall Stability and Marshall Flow Value) and Volumetric Properties (AV, BSG, VMA, VFA)

	BC	PD 20	PD 16	PD10	PD 4.75	SG20	SG16	SG10	SG Fe	SGB	PP 26.5	PP 19	PP 13.2	PP50	PP 4.75	PP 2.36	PP 1.18	PP 0.60	PP 0.30	PP 0.15	PP 0.075
BC	1.0																				
PD 20	-0.2	1.0																			
PD 16	0.1	-0.8	1.0																		
PD10	0.1	-0.5	0.1	1.0																	
PD 4.75	0.1	-0.2	0.0	-0.3	1.0																
SG20	-0.1	0.2	-0.2	0.1	-0.1	1.0															
SG16	-0.1	-0.1	-0.1	0.2	0.0	0.8	1.0														
SG10	0.0	-0.1	-0.1	0.2	0.1	0.6	0.8	1.0													
SGF	0.0	0.0	-0.2	0.2	0.1	0.2	0.3	0.3	1.0												
SGB	0.0	0.2	-0.2	-0.1	0.0	0.2	0.1	0.0	0.2	1.0											
PP 26.50	0.2	-0.2	0.2	0.0	0.1	-0.2	0.0	0.0	0.0	-0.1	1.0										
PP 19	0.3	-0.6	0.5	0.3	0.3	-0.2	0.0	0.1	0.1	-0.1	0.5	1.0									
PP 13.20	0.2	-0.7	0.3	0.4	0.5	-0.2	0.0	0.1	0.1	-0.1	0.3	0.6	1.0								
PP 9.50	0.3	-0.7	0.4	0.5	0.3	-0.3	-0.1	0.0	0.1	-0.2	0.4	0.7	0.8	1.0							
PP 4.75	0.2	-0.7	0.4	0.3	0.5	-0.3	-0.1	0.0	0.1	-0.1	0.3	0.7	0.8	0.8	1.0						
PP 2.36	0.2	-0.6	0.3	0.3	0.5	-0.2	-0.1	0.0	0.0	-0.2	0.2	0.5	0.8	0.7	0.8	1.0					
PP 1.18	0.2	-0.5	0.2	0.1	0.5	-0.3	-0.2	0.0	0.0	-0.1	0.3	0.5	0.8	0.6	0.7	0.9	1.0				
PP 0.60	0.2	-0.6	0.3	0.2	0.5	-0.3	-0.2	0.1	0.1	-0.1	0.2	0.5	0.8	0.7	0.7	0.8	0.9	1.0			
PP 0.30	0.1	-0.2	0.0	0.0	0.4	-0.1	-0.1	0.1	0.0	-0.1	0.1	0.2	0.6	0.3	0.5	0.7	0.9	0.8	1.0		
PP 0.15	0.1	-0.3	0.2	0.1	0.4	-0.1	-0.1	0.1	0.0	-0.2	0.2	0.3	0.7	0.5	0.5	0.7	0.8	0.8	0.8	1.0	
PP 0.075	0.0	-0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.0	-0.4	0.0	0.1	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.4	1.0

4.4 Compliance to SSRBW 2016

4.4.1 Compliance to SSRBW DBM, NAS 19mm and NAS 13.2mm mixes

Marshall mix design that are designed according to section 1309 of SSRBW 2016, in Nepal were examined based on parameter specified in (Department of Road, 2016). The Table 4-6 provides a summary of the Marshall Mix Design properties for DBM and different mixes based on gradations (I and II) against the standard values specified in the SSRBW 2016 standard for Viscosity Grade Paving Bitumen. Compliance percentages indicate how well each mix with different gradation meets the specified criteria.

Table 4-6 Compliance Summary to SSRBW 2016 DBM, Asphalt Mix NAS 19mm and Asphalt Mix 13.2mm

Marshall Mix Design Properties	Properties values for Viscosity Grade Paving Bitumen	Status As per SSRBW 2016		
		DBM	Gradation I	Gradation II
Number of Samples		19	52	38
Minimum Stability (KN at 60 D Centigrade)	9	100%	98%	100%
Marshall Flow (mm	2-4	100%	94%	97%
Marshall Quotient	2-5	68%	67%	71%
% Air Voids	3-5	95%	92%	84%
% VFA	65-75	84%	62%	55%
minimum % VMA	12	100%	100%	100%
Coating of aggregate Particle	95% Minimum	NA	NA	NA
Tensile Strength Ration	80% Minimum	NA	NA	NA
Filler -Binder Ratio, SSRBW 2016	0.6 - 1.2	32%	50%	32%

As per the standards outlined in SSRBW 2016, the Asphalt Concrete pavements must meet a minimum stability requirement of 9 KN at 60°C. It is crucial to highlight that all the samples included in the analysis have demonstrated exceptional stability, exceeding the stipulated minimum value.

Additionally, the flow values, indicating the horizontal deformation of the asphalt mix at the maximum load, fall within the desired range of 2-4 mm for 94% of the samples, suggesting that the majority of the samples possess optimal deformation characteristics.

However, when examining the Marshall Quotient (MQ), a parameter reflecting the stiffness of the mix or resistance to share, it becomes evident that only 68%, 67% and 71% of the samples related to DBM, Gradation I and Gradation II mixes respectively, meet the specified requirements. The MQ is calculated by dividing the stability (KN) by the flow value (mm), and it serves as an indicator of the mixture's ability to withstand load (Putri, Kasyafi, & Ahmad, 2023) [23]. Higher MQ values suggest a stiffer or more brittle mix, while lower values indicate a mix that may fail to withstand the load, potentially leading to the development of ruts (Tong, Ma, Shen, Zhang, & Wu, 2022). In this context, the 68%, 67%, and 71% compliance rates highlight that a significant number of the samples demonstrate satisfactory stiffness. However, there is room for improvement in meeting the specified standards for this particular parameter.

Modifying any factor or mix design procedure may lead to a decline in performance or service life. Research has shown that mixtures consolidating to less than 2 percent air voids are prone to rutting and shoving in heavy traffic locations (Hafeez, 2009). Problems may arise if, over time, the final air void content exceeds 5 percent or if the initial construction involves over 8 percent air voids, leading to issues like brittleness, premature cracking, raveling, and stripping (Abdulhaq Hadi Abedali, 2014). In this context, 92% and 84% of the Gradation I AC and Gradation II AC samples meeting specified requirements suggest areas for improvement in adhering to the specified standards.

Voids in Mineral Aggregate (VMA) denote the total volume of voids in compacted aggregate, critical for asphalt mix durability. An inadequate VMA may compromise durability, while excessive VMA poses stability challenges and results in uneconomical binder consumption, leading to rapid binder oxidation if voids are inadequately filled (Chadbourn, Skok Jr, Newcomb, Crow, & Spindle, 1999) (Pouranian & Haddock, 2018) (Kandhal, Chakraborty, & others, Evaluation of voids in the mineral aggregate for HMA paving

mixtures, 1996). The objective is to provide ample space for asphalt, ensuring adhesion to aggregate without bleeding during temperature fluctuations. Design bitumen contents within a specific range may exhibit bleeding or plastic flow, and extra compaction from traffic can cause rutting in high-traffic areas (Park T. , 2007). Optimal design bitumen content should be slightly left of the low point on the VMA curve. If the mix is on the left-hand side, it would be scorched, susceptible to segregation, and likely to have elevated air voids (Balitsaris, 2012). In our study, all the mixes meet the requirement by 100%.

The interrelation of VFA, VMA, and Percentage Air Voids in asphalt mix design requires attention. While any two values can solve for the third, incorporating VFA criteria is vital to avoid mixes with marginally acceptable VMA. VFA primarily limits maximum VMA levels and bitumen content, restricting allowable air void (DoR, 2016) (Balitsaris, 2012). Mixes designed for lower traffic volumes struggle to meet VFA criteria. In contrast, those for heavy traffic with low air voids fail, indicating increased susceptibility to top-down rutting (Zhang, et al., 2019), with only 62% and 55% of job mix related to NAS 19 and NAS 13.2 meeting standards.

The filler-binder ratio is a pivotal determinant of Hot Mix Asphalt (HMA) workability, notably affecting resistance to plastic deformations (Vale, do, & Grecco, 2016). Higher filler content imparts stiffness to the mix, on the other hand asphalt become more fragile and consequently more susceptible to crack under fatigue and low temperatures. Due to these antagonistic effects on performance, selecting the right amount of mineral filler to compound asphalt is a task of great importance (Vale, do, & Grecco, 2016). IRC 135, 2022 prescribes a ratio of 0.8 to 1 (SP:135., 2022), and SSRBW 2016 recommends 0.6 to 1.2 (Department of Road, 2016). This property significantly influences the workability of asphalt mixtures, as a low Filler-Binder ratio can render a mix challenging to compact. However, compliance issues arise, with only 32%,50% and 32% of mixes meeting the requirement, posing potential quality concerns in the field. Additionally, considerations for aggregate composition must address potential inadequacies in voids for optimal bitumen coverage within Gradations in Marshall Design

4.4.2 Compliance to SSRBW 2001 NAS 20mm

Marshall mix designs, designed according to SSRBW 2001, in Nepal were examined based on parameters specified in Table 4-7. A summary of compliance percentages indicates how well each mix meets the specified criteria.

Table 4-7 Compliance Summary to SSRBW 2001 NAS 20mm Asphalt Mix

Marshall Mix Design Properties	Wearing Course	Status As per SSRBW 2001
		Compliance on Wearing Course
Number of Samples		30
Marshall Stability (KN)	Min 8 KN	100%
Flow Value (mm)	2-3	60%
Voids in Total Mix (%)	2-7	100%
Compression / Immersion Ratio	0.75	NA

4.5 Centrality Analysis

In order to evaluate how we are choosing gradation while performing Marshall mix design is crucial for its optimal design in the field for maximum density centrality analysis is carried based on the hypothesis testing and t- statistics is used to evaluate hypothesis. Hypothesis is setup as follows:

H0: Mid-point grade and Sample mean of gradation have no significant differences.

H1: Mid-point grade and sample means of gradation have significant differences.

And, centrality for respected mixes is evaluated and presented here:

4.5.1 Dense Graded Bituminous Macadam (DBM)

Critical value for Right tailed T-Statistics (*dof=18*, significance level =5%) = 1.729

Critical value for two tailed T-Statistics (*dof=18*, significance level =5%) = 2.093

Table 4-8 Summary of Centrality analysis for DBM gradation

Sieve Size	Mid - Grade	Mean	t-stat (absolute)	Verdict on One Tailed	Verdict on Two Tailed
PP 26.50mm	95.00	98.68	0.45	Reject	Reject
PP 19mm	83.00	86.15	1.01	Reject	Reject
PP 13.20mm	68.00	68.13	1.08	Failed To Reject	Failed To Reject
PP 4.75 mm	46.00	38.64	1.31	Reject	Reject
PP 2.36 mm	35.00	32.87	0.62	Reject	Reject
PP 0.30 mm	14.00	13.25	0.52	Failed To Reject	Failed To Reject
PP 0.075 mm	5.00	6.45	0.26	Reject	Reject

Based on the given Table 4-8, we compare the t-statistics to the critical t-value of 1.675 to determine the significance of differences between sample means and mid-grade percentages for each sieve size. For the PP 26.50mm sieve, with a mid-grade of 95.00 and a sample mean of 98.68, the absolute t-statistic is 0.45, which does not exceed the critical value, so we fail to reject the null hypothesis. Similarly, for the PP 19mm sieve (mid-grade 83.00, mean 86.15) and the PP 13.20mm sieve (mid-grade 68.00, mean 68.13), the t-statistics are 1.01 and 1.08, respectively, both below the critical value, leading us to fail to reject the null hypothesis.

For the PP 4.75mm sieve (mid-grade 46.00, mean 38.64), the t-statistic is 1.31, also not exceeding the critical value, so we fail to reject the null hypothesis. The PP 2.36mm sieve (mid-grade 35.00, mean 32.87) has a t-statistic of 0.62, and the PP 0.30mm sieve (mid-grade 14.00, mean 13.25) has a t-statistic of 0.52, both failing to exceed the critical value, resulting in a failure to reject the null hypothesis. Lastly, for the PP 0.075mm sieve (mid-grade 5.00, mean 6.45), the t-statistic is 0.26, also below the critical value, so we fail to reject the null hypothesis. In summary, all sieve sizes have t-statistics that do not exceed the critical value of 1.675, indicating no significant differences between the sample means and the mid-grade percentages.

4.5.2 Asphalt Concrete with NAS 19MM

Critical value for Right tailed T-Statistics (*dof=51*, significance level =5%) = 1.675

Critical value for two tailed T-Statistics (*dof=51*, significance level =5%) = 2.008

Table 4-9 Centrality analysis for Asphalt Mix with NAS 19mm

Sieve Size	Mid - Grade	Mean	t-stat (absolute)	Verdict on One Tailed
PP 26.50mm	100.00	99.97	1.00	Failed To Reject
PP 19mm	95.00	95.66	0.94	Failed To Reject
PP 13.20mm	69.00	76.37	8.66	Reject
PP 9.50mm	62.00	64.07	2.35	Reject
PP 4.75 mm	45.00	42.93	2.43	Reject
PP 2.36 mm	36.00	33.71	4.31	Reject
PP 1.18 mm	27.00	24.08	7.30	Reject
PP 0.60 mm	21.00	19.21	5.16	Reject
PP 0.30 mm	15.00	14.11	2.98	Reject
PP 0.15 mm	9.00	9.07	0.27	Failed To Reject
PP 0.075 mm	5.00	6.02	5.61	Reject

Based on the Table 4-9, for t-statistics -1.00, 0.94, 0.27, and 0.27, none of them exceed the critical t-value of 1.675, so we fail to reject the null hypothesis for these percentage passing through respective sieves. For t-statistics 8.66, 5.61, 2.35, all of them exceed the critical t-value of 1.675, so we reject the null hypothesis for these particle sizes, indicating there is a significant difference between the mean of these percentage passing through these sizes and the mid-grade percentage passing. For t-statistics -2.43, -4.31, -7.30, -5.16, -2.98, these absolute values exceed the critical t-value of 1.675, so we reject the null hypothesis for these particle sizes, indicating there is a significant difference between the sample mean and the mid-grade gradation for respective sieves.

4.5.3 Asphalt Concrete with NAS 13.2MM

Critical value for Right tailed T-Statistics (*dof=38*, significance level =5%) = 1.687

Critical value for two tailed T-Statistics (*dof=38*, significance level =5%) = 2.026

Table 4-10 Centrality analysis for Asphalt Mix with NAS 13.2mm

Parameters	Mid-Grade	Mean	Standard Error	T-Statistics	T critical (one tailed)	Verdict on One Tailed
PP 19mm	100.00	99.52	0.33	1.44	1.687	Failed To Reject
PP 13.20mm	95.00	92.12	1.13	2.55	1.687	Reject
PP 9.50mm	79.00	80.83	1.11	1.65	1.687	Failed To Reject
PP 4.75 mm	62.00	58.25	0.90	4.18	1.687	Reject
PP 2.36 mm	50.00	45.73	0.74	5.79	1.687	Reject
PP 1.18 mm	41.00	36.18	0.81	5.94	1.687	Reject
PP 0.60 mm	32.00	29.18	0.90	3.13	1.687	Reject
PP 0.30 mm	23.00	21.00	0.53	3.74	1.687	Reject
PP 0.15 mm	16.00	13.43	0.21	12.09	1.687	Reject
PP 0.075 mm	7.00	7.34	0.22	1.54	1.687	Failed To Reject

Based on the Table 4-10, we compare the t-statistics to the critical values (1.687 for one-tailed and 2.026 for two-tailed tests). For PP 19mm, PP 9.50mm, and PP 0.075mm sieves, the t-statistics (1.44, 1.65, and 1.54, respectively) do not exceed the critical values, so we fail to reject the null hypothesis, indicating no significant differences. For the other sieves (PP 13.20mm, PP 4.75mm, PP 2.36mm, PP 1.18mm, PP 0.60mm, PP 0.30mm, and PP 0.15mm), the t-statistics exceed the critical values, so we reject the null hypothesis, indicating significant differences between the sample means and the mid-grade percentages.

4.5.4 SSRBW 2001 Wearing Course

Critical value for Right tailed T-Statistics (*dof=29*, significance level =5%) = 1.687

Critical value for two tailed T-Statistics (*dof=29*, significance level =5%) = 2.026

Table 4-11 Centrality Analysis for SSRBW 2001 Wearing Couse

Parameters	Mid-Grade	Mean	Standard Error	T-Statistics	T critical (one tailed)	Verdict on One Tailed
PP 19mm	100.00	99.79	0.21	1.00	1.699	Failed To Reject
PP 13.20mm	79.00	79.66	0.67	0.99	1.699	Failed To Reject
PP 4.75 mm	56.50	49.08	1.13	6.55	1.699	Reject
PP 2.36 mm	35.00	34.63	0.70	0.52	1.699	Failed To Reject
PP 0.60 mm	18.50	18.85	0.30	1.16	1.699	Failed To Reject
PP 0.075 mm	7.50	7.26	0.16	1.48	1.699	Failed To Reject

In the given Table 4-11, we compare the t-statistics to the critical values (1.699 for one-tailed and 2.045 for two-tailed tests). For PP 19mm (t-statistic 1.00), PP 13.20mm (t-statistic 0.99), PP 2.36mm (t-statistic 0.52), PP 0.60mm (t-statistic 1.16), and PP 0.075mm (t-statistic 1.48), the t-statistics are below the critical values, leading us to fail to reject the null hypothesis, indicating no significant differences. Conversely, for PP 4.75mm (t-statistic 6.55), the t-statistic surpasses both critical values, resulting in the rejection of the null hypothesis and indicating a significant difference between the sample mean and the mid-grade percentage.

4.6 Multiple Linear Regression Model

4.6.1 Multiple Linear Regression Analysis OBC Model Results

4.6.1.1 Multiple Linear Regression Analysis Model of Optimum Bitumen in Percentage Weight

The MLR analysis is performed in Excel- Stat plus add-ins based on the 20 Independent parameters such as Percentage Coarse and Fine Aggregate represented by Percentage 20 mm down, Percentage 16mm down aggregate, Percentage 13 mm down aggregate, Percentage 4.75mm down, Specific gravity representation by SG20 mm, SG16mm, SG10mm,

SGF, SGB, and Gradation of Mix represented by PP 26.50, PP 19, PP 13.20, PP 9.50, PP 4.75, PP 2.36, PP 1.18, PP 0.60, PP 0.30, PP 0.15, PP 0.075 with total number of **141 samples (75% for training and 25% for testing).** And it can be seen from following Table 4-12 that the performance of MLR shows moderate fitting of datasets in training stage and low to moderate fitting in testing stages for the prediction of OBC, with Coefficient of Correlation as (0.835, 0.7857), Coefficient of Determination R-Square as (0.698,0.0.617), Adjusted R-Square as (0.626, 0.6061), for both stages. Also, F-statistic is greater than probability of (F-statistic) indicates regression is significant.

Table 4-12 Statistical Results of MLR model for Optimum Bitumen Content (OBC)

Parameter	At Training Stage	Testing Stage
N	106	35
R	0.8354	0.7857
R-Squared	0.698	0.6173
Adjusted R-Square	0.626	0.6061
conditional No	44600	
F-statistic:	9.703	
Prob (F-statistic):	1.92E-14	

The following equation, which indicates the sign and the magnitude of each feature's contribution to the modeled asphalt property, is obtained by using the MLR model as

$$\textit{OPTIMUM BITUMEN CONTENT} = -0.4642 + 0.0055 * PD\ 20 + 0.0080 * PD\ 16 + 0.0051 * PD\ 10 + 0.0076 * PD\ 4.75 + -1.0295 * SG\ 20 + 0.3765 * SG\ 16 + 0.5471 * SG\ 10 + 1.0810 * SGF + 1.3391 * SGB + -0.0069 * PP\ 26.50 + 0.0220 * PP\ 19 + -0.0129 * PP\ 13.20 + 0.0094 * PP\ 9.50 + 0.0098 * PP\ 4.75 + -0.0141 * PP\ 2.36 + -0.0088 * PP\ 1.18 + 0.0250 * PP\ 0.60 + -0.0085 * PP\ 0.30 + 0.0311 * PP\ 0.15 + -0.0201 * PP\ 0.075 \qquad \text{Equation 29}$$

Some of the coefficients seem to have negligible effect on the Optimum bitumen content but these variables are not neglected because of SSRBW 2016. The condition number is large, 4.46e+04. This might indicate that there is strong multicollinearity or other numerical problems.

The MLR OBC model was evaluated using a 25% independent dataset. With an R^2 value of 0.617, the model demonstrates moderate predictive ability, explaining some portion of the variability in OBC based on the given independent variables, as illustrated in Figure 4-5. Nevertheless, there could be additional factors influencing the variation in OBC that the model does not account for.

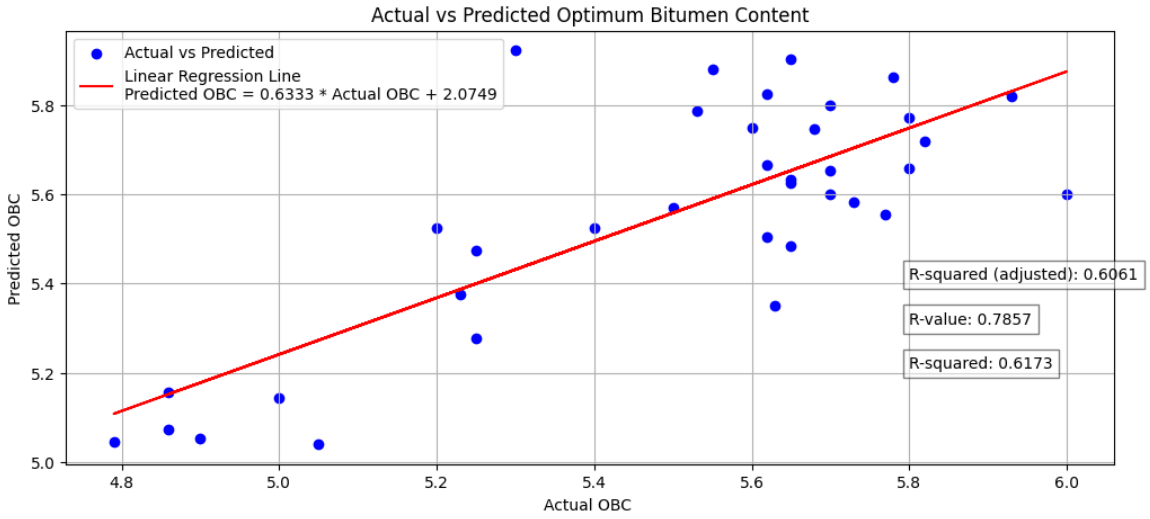


Figure 4-5 Relationship between Actual OBC and MLR Predicted OBC for Testing Data Set

The comparison between actual and predicted values of Optimum Bitumen Content, based on MLR model at testing stage, is shown in the following Figure 4-6 Actual OBC and MLR model predicted OBC Figure 4-6. Also, these figures illustrate the distribution of errors between the predicted and actual values over different observations.

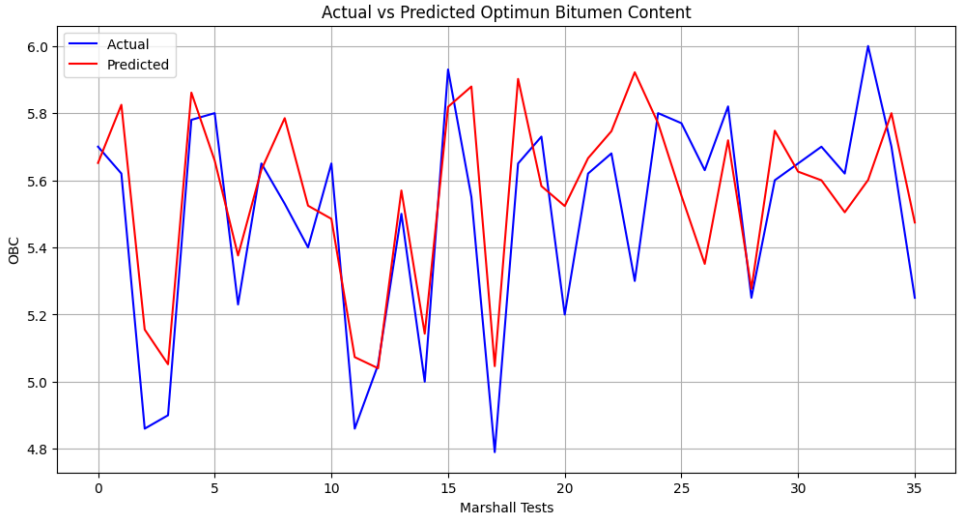


Figure 4-6 Actual OBC and MLR model predicted OBC for Testing Set

4.6.1.2 Multiple Linear Regression Analysis Model of Marshall Stability

For development of the linear multiple regression model for Marshall Stability, data were divided into test set and testing set by dividing 544 samples into 75% and 25 % respectively and performing a multiple regression analysis with 21 independent parameters. Following table shows dependent and independent variables.

Dependent variable	Marshall Stability, KN
Independent variables	PD 20, PD 16, PD 13, PD 4.75, SG 20, SG 16, SG 10, SGF, SGB, PP 26.50, PP 19, PP 13.20, PP 9.50, PP 4.75, PP 2.36, PP 1.18, PP 0.60, PP 0.30, PP 0.15, PP 0.075, BC
N	544 (75% training and 25% Testing) Data

The developed model and its statistical parameters are presented in following Table 4-13 at training and testing stage.

Table 4-13 Statistical Results of MLR Model for Marshall Stability

Parameter	At Training Stage	Testing Stage
N	408	136
R	0.550	0.5363
R-Squared	0.303	0.2876
Adjusted R-Square	0.265	0.2823
conditional No	9.01E+04	
F-statistic:	7.979	
Prob (F-statistic):	4.68E-20	

Based on regression analysis, the model for multiple linear regression of Stability was developed and presented as the following equation.

$$\begin{aligned} &\textit{Stability in KN} = 44.0040 + 0.0314 * PD\ 20 + 0.0222 * PD\ 16 + \\ &0.0863 * PD\ 13 + 0.1633 * PD\ 4.75 + -21.0539 * SG\ 20 + 14.2203 * \\ &SG\ 16 + -14.7201 * SG\ 10 + 0.7384 * SGF + -8.1431 * SGB + 0.2354 \\ &* \ PP\ 26.50 + 0.0041 * \ PP\ 19 + 0.0298 * \ PP\ 13.20 + -0.0469 * \\ &PP\ 9.50 + -0.0431 * \ PP\ 4.75 + 0.1714 * \ PP\ 2.36 + -0.0525 * \ PP \\ &1.18 + -0.1308 * PP\ 0.60 + -0.0156 * \ PP\ 0.30 + 0.1274 * \ PP\ 0.15 \\ &+ -0.3607 * \ PP\ 0.075 + -0.0060 * BC \end{aligned} \qquad \textit{Equation 30}$$

The regression coefficients for each variable are summarized in Table 4-14. Variables with a P-value greater than 0.05 do not significantly affect the predicted Marshall Stability and can be excluded from the model. However, due to practical reasons, the percentages of coarse and fine aggregate are retained in the equation to monitor gradation limits as specified in SSRBW 2016.

Table 4-14 Test statistics and regression coefficients for MLR Marshall Stability Model

Variables	Coefficient	SE	t	P>|t|
Intercept	64.961	25.718	2.526	0.012
PD 20	0.007	0.034	0.205	0.838

Variables	Coefficient	SE	t	P>|t|
PD 16	-0.0075	0.035	-0.216	0.829
PD 10	0.0594	0.03	1.962	0.05
PD 4.75	0.1598	0.033	4.903	0
SG 20	-24.9354	5.155	-4.837	0
SG 16	21.4128	5.063	4.229	0
SG 10	-17.4205	3.114	-5.595	0
SGF	-2.8645	1.796	-1.595	0.112
SGB	0.5471	12.802	0.043	0.966
PP 26.50	0.0108	0.211	0.051	0.959
PP 19	0.0551	0.029	1.888	0.06
PP 13.20	-0.0074	0.026	-0.29	0.772
PP 9.50	-0.0418	0.017	-2.461	0.014
PP 4.75	-0.0426	0.025	-1.717	0.087
PP 2.36	0.2093	0.041	5.055	0
PP 1.18	-0.0992	0.068	-1.45	0.148
PP 0.60	-0.1543	0.061	-2.529	0.012
PP 0.30	0.0811	0.062	1.299	0.195
PP 0.15	0.1533	0.11	1.394	0.164
PP 0.075	-0.4615	0.119	-3.885	0
BC	-0.1382	0.147	-0.94	0.348

Testing of the MLR MST model was performed using a 25% independent set of data. Overall, the model appears to have low to moderate predictive ability, with R^2=0.2876, explaining only a small portion of the variation in stability using the given independent variables and relation between Actual and Predicted Marshall stability is shown in Figure 4-7. However, there may be other factors not accounted for in the model that also influence the variability of Marshall Stability.

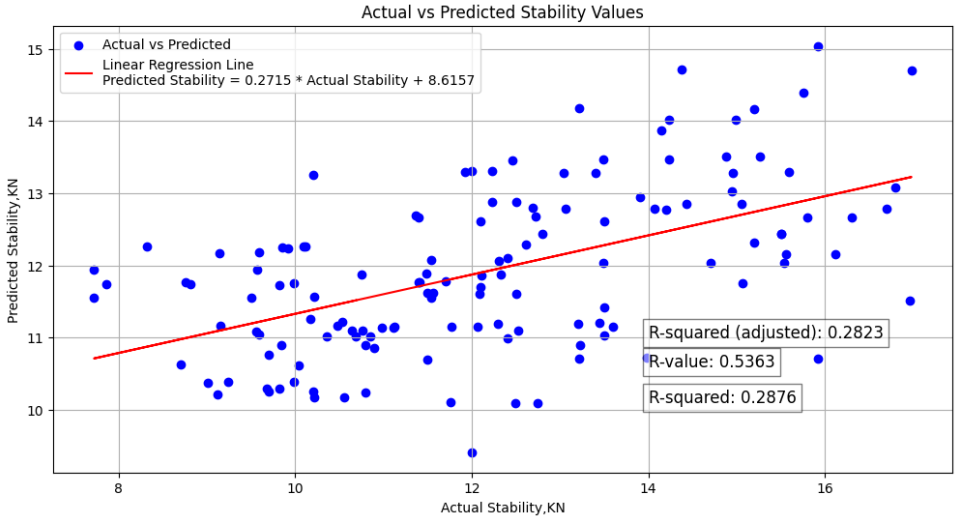


Figure 4-7 Relation between Actual and Predicted Marshall Stability for Testing Set

Figure 4-8 compares the actual Marshall Stability values to those predicted by the MLR model during the testing phase. It also shows the error distribution between predicted and actual values across various observations.

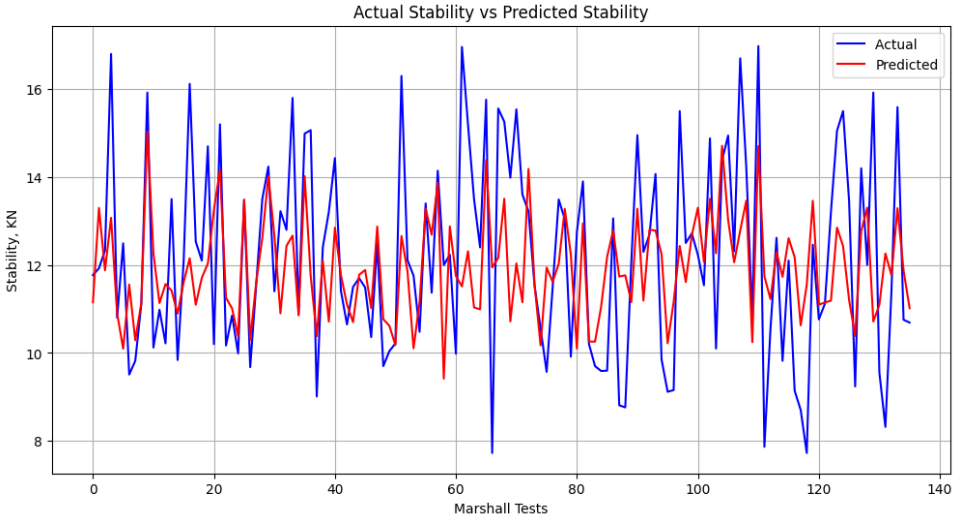


Figure 4-8 Actual MST and MLR model predicted MST for Testing Set

4.6.1.3 Multiple Linear Regression Analysis Model of Flow Value

For constructing the multiple linear regression model, the dataset of 574 samples was divided into a training set comprising 75% of the data and a validation set comprising 25%. A multiple regression analysis was performed using 21 independent variables. The dependent and independent variables are presented in the following table.

Dependent variable	Marshall Flow Value, MM
Independent variables	PD 20, PD 16, PD 13, PD 4.75, SG 20, SG 16, SG 10, SGF, SGB, PP 26.50, PP 19, PP 13.20, PP 9.50, PP 4.75, PP 2.36, PP 1.18, PP 0.60, PP 0.30, PP 0.15, PP 0.075, BC
N	555 (75% for training and 25% for testing)

With the help of Excel Add-ins, a MLR for Flow value was developed and its statistical parameters are presented in Table 4-15 Statistical Results of MLR model for Marshall Flow Value.

Table 4-15 Statistical Results of MLR model for Marshall Flow Value

Parameter	At Training Stage	Testing Stage
N	416	139
R	0.665	0.657
R-Squared	0.442	0.4321
Adjusted R-Square	0.412	0.4279
conditional No	5.79E+04	
F-statistic:	14.84	
Prob (F-statistic):	4.50E-38	

Above table indicates a multiple correlation coefficient (Multiple R) of 0.665, which represents the strength and direction of the linear relationship between the Marshall Flow and the set of independent variables. The coefficient of determination (R^2) is 0.442, meaning that approximately 44.2% of the variability in the Marshall Flow can be explained by the independent variables.

Based on regression analysis, The model for multiple linear regression of Flow Value was developed and presented as following equation and regression coefficients for each variable are summarized in the

Table 4-16.

$$\begin{aligned} Flow\ Value = & -3.2054 + -0.0206 * PD\ 20 + -0.0166 * PD\ 16 + \\ & 0.0029 * PD\ 10 + 0.0227 * PD\ 4.75 + -0.0720 * SG\ 20 + 0.2134 * \\ & SG\ 16 + -2.4455 * SG\ 10 + 0.0676 * SGF + 2.6038 * SGB + 0.0644 \\ & * PP\ 26.50 + 0.0171 * PP\ 19 + -0.0024 * PP\ 13.20 + -0.0042 * \\ & PP\ 9.50 + -0.0132 * PP\ 4.75 + 0.0140 * PP\ 2.36 + 0.0079 * PP \end{aligned} \qquad \text{Equation 31}$$

*1.18 + -0.0400 * PP 0.60 + 0.0649 * PP 0.30 + -0.1173 * PP 0.15 + -0.0155 * PP 0.075 + 0.4240 * BC*

Table 4-16 Test Statistics and Regression Coefficients for MLR Marshall Flow Model l

Parameters	C	SE	t	P>|t|
Intercept	-3.2054	6.398	-0.501	0.617
PD 20	-0.0206	0.008	-2.452	0.015
PD 16	-0.0166	0.009	-1.94	0.053
PD 10	0.0029	0.008	0.372	0.71
PD 4.75	0.0227	0.008	2.776	0.006
SG 20	-0.072	1.19	-0.06	0.952
SG 16	0.2134	1.176	0.181	0.856
SG 10	-2.4455	0.711	-3.442	0.001
SGF	0.0676	0.382	0.177	0.86
SGB	2.6038	2.994	0.87	0.385
PP 26.50	0.0644	0.053	1.211	0.227
PP 19	0.0171	0.007	2.499	0.013
PP 13.20	-0.0024	0.006	-0.403	0.687
PP 9.50	-0.0042	0.004	-1.011	0.312
PP 4.75	-0.0132	0.006	-2.097	0.037
PP 2.36	0.014	0.01	1.415	0.158
PP 1.18	0.0079	0.015	0.511	0.61
PP 0.60	-0.04	0.014	-2.841	0.005
PP 0.30	0.0649	0.014	4.699	0
PP 0.15	-0.1173	0.023	-5.124	0
PP 0.075	-0.0155	0.027	-0.572	0.568
BC	0.424	0.036	11.849	0

Testing of the MLR MFV model was performed using a 25% independent set of data. Overall, the model appears to have low to moderate predictive ability, with R^2=0.4321, explaining only a limited portion of the variation in Flow value using the given independent and relation between Actual and Predicted MFV shown in Figure 4-9. However, there may be other factors not accounted for in the model that also influence the variability of Marshall Flow value.

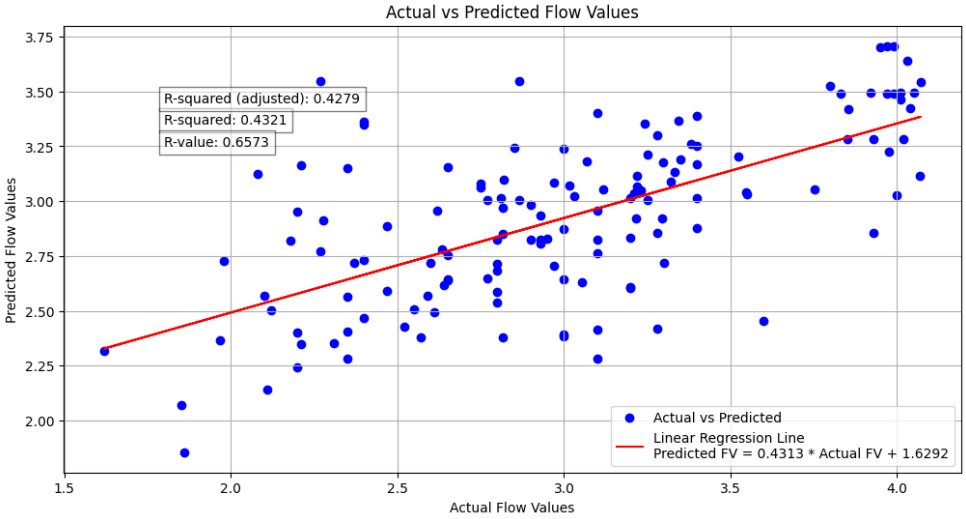


Figure 4-9 Relation between Actual and Predicted MFV for Testing Set

The comparison between actual and predicted values of Marshall Flow Value, mm, based on the MLR model at testing stage, is shown in the following Figure 4-10. Also, this figure illustrates the distribution of errors between the predicted and actual values over different observations.

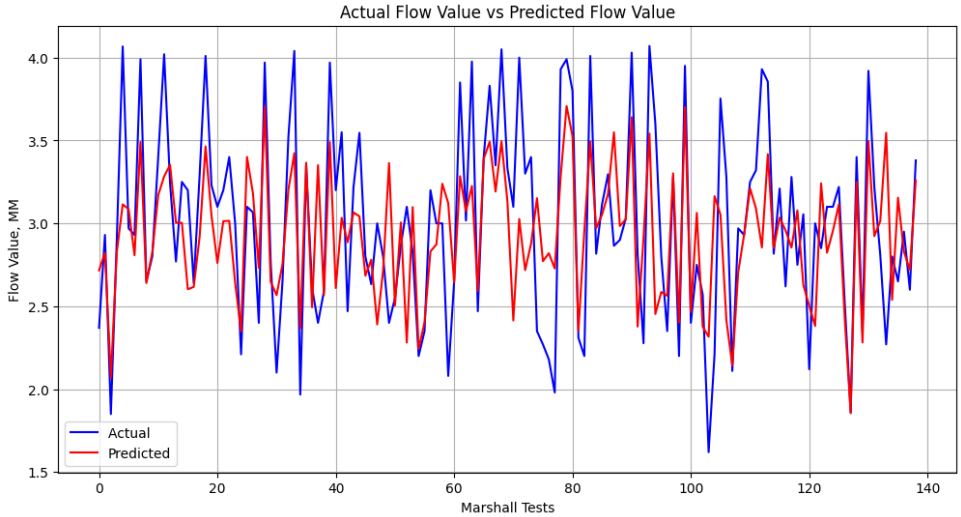


Figure 4-10 Actual MVF and MLR model predicted MVF for Testing Set

4.6.1.4 Multiple Linear Regression Analysis Model of Air Voids

For development of the linear multiple regression model, data were divided into test sets and validation set into 80% and 20 % ratios and performed a multiple regression analysis. With 21 independent parameters. As shown below with 633 input and output data set.

Dependent variable:	Air Voids
Independent variables:	PD 20, PD 16, PD 13, PD 4.75, SG 20, SG 16, SG 10, SGF, SGB, PP 26.50, PP 19, PP 13.20, PP 9.50, PP 4.75, PP 2.36, PP 1.18, PP 0.60, PP 0.30, PP 0.15, PP 0.075, BC
N:	600 (75% for training and 25% for testing)

The developed model and its statistical parameters are presented in following Table 4-17 at training and testing stage.

Table 4-17 Statical Results of MLR model for Percentage Air Aoids

Parameter	At Training Stage	Testing Stage
N	450	150
R	0.812	0.806
R-Squared	0.66	0.6494
Adjusted R-Square	0.643	0.647
conditional No	8.45E+04	
F-statistic:	1.53E-86	
Prob (F-statistic):	4.50E-38	

Based on regression analysis. The model for multiple linear regression of Stability was developed and presented as following equation (), regression coefficients for each variable are summarized in the Table 4-19.

$$\textit{Air voids} = -9.8363 + -0.0070 * PD\ 20 + 0.0058 * PD\ 16 + 0.0268 * PD\ 10 + -0.0035 * PD\ 4.75 + -7.3799 * SG\ 20 + 5.5638 * SG\ 16 + -1.0053 * SG\ 10 + 1.4850 * SGF + 23.0471 * SGB + 0.0027 * PP \quad \textit{Equation 32}$$

*26.50 + 0.0189 * PP 19 + -0.0195 * PP 13.20 + 0.0090 * PP 9.50 + 0.0154 * PP 4.75 + 0.0001 * PP 2.36 + 0.0211 * PP 1.18 + 0.0072 * PP 0.60 + -0.0011 * PP 0.30 + -0.0380 * PP 0.15 + -0.0580 * PP 0.075 + -1.6125 * BC*

Table 4-18 Test Statistics and Regression Coefficients for MLR Air Voids Model

Parameters	C	SE	t	P>|t|
Intercept	-9.8363	15.644	-0.629	0.53
PD 20	-0.007	0.016	-0.436	0.663
PD 16	0.0058	0.016	0.359	0.72
PD 10	0.0268	0.015	1.782	0.075
PD 4.75	-0.0035	0.015	-0.231	0.818
SG 20	-7.3799	2.161	-3.415	0.001
SG 16	5.5638	2.149	2.588	0.01
SG 10	-1.0053	1.245	-0.807	0.42
SGF	1.485	0.677	2.195	0.029
SGB	23.0471	5.109	4.511	0
PP 26.50	0.0027	0.142	0.019	0.985
PP 19	0.0189	0.011	1.701	0.09
PP 13.20	-0.0195	0.009	-2.181	0.03
PP 9.50	0.009	0.007	1.351	0.177
PP 4.75	0.0154	0.01	1.578	0.115
PP 2.36	0.0001	0.013	0.009	0.993
PP 1.18	0.0211	0.024	0.898	0.369
PP 0.60	0.0072	0.023	0.312	0.755
PP 0.30	-0.0011	0.022	-0.05	0.96
PP 0.15	-0.038	0.039	-0.963	0.336
PP 0.075	-0.058	0.043	-1.342	0.18
BC	-1.6125	0.06	-26.983	0

Testing of the MLR AV model was performed using a 25% independent set of data. Overall, the model appears to have low to moderate predictive ability, with R^2=0.6495, explaining only the model explains about 65% of the variation in Air Voids and relation between Actual and Predicted AV is shown in Figure 4-11. However, there may be other factors not accounted for in the model that also influence the variability of Air Voids.

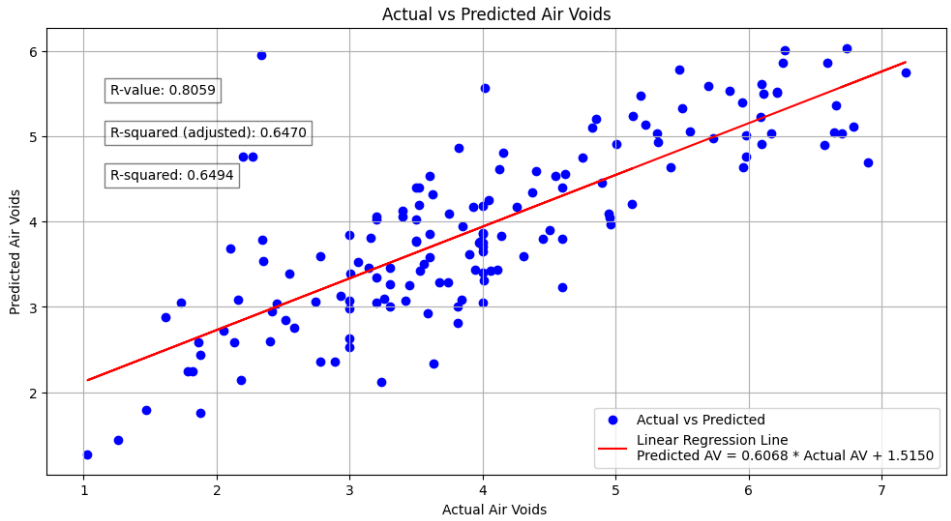


Figure 4-11 Relation Between Actual and Predicted Air Voids for Testing Set

Actual and predicted Air Voids during Testing Stage and error distribution can be seen in Figure 4-12:

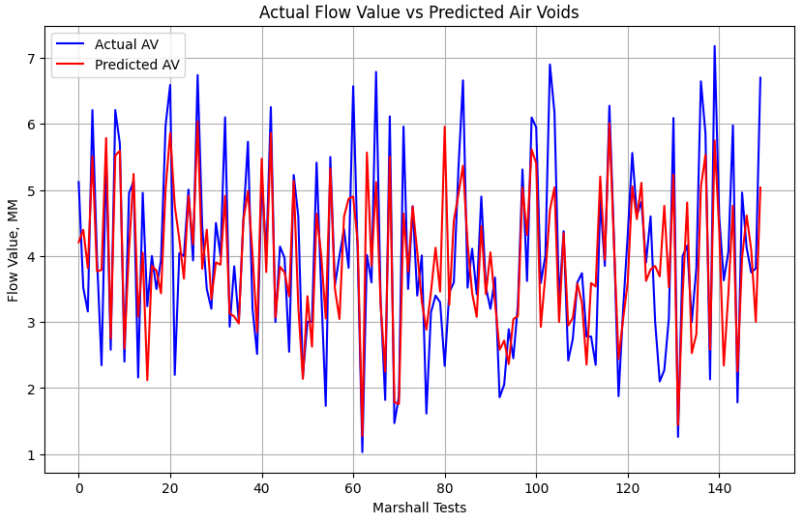


Figure 4-12 Actual PAV and MLR model predicted PAV for Testing Set

4.6.1.5 Multiple Linear Regression Analysis Model of VMA

For development of the linear multiple regression model, data were divided into test sets and validation sets into 80% and 20 % ratios and performed a multiple regression analysis. With 21 independent parameters. As shown below with 621 input and output data set.

Dependent variable:	VMA,
Independent variables:	PD 20, PD 16, PD 13, PD 4.75, SG 20, SG 16, SG 10, SGF, SGB, PP 26.50, PP 19, PP 13.20, PP 9.50, PP 4.75, PP 2.36, PP 1.18, PP 0.60, PP 0.30, PP 0.15, PP 0.075, BC
N:	628 (75% training and 25% testing data set)

The developed model and its statistical parameters are presented in following Table 4-19 at training and testing stage.

Table 4-19 Statical Results of MLR Model for Void in Mineral Aggregate

Parameter	At Training Stage	Testing Stage
N	471.000	157
R	0.580	0.5488
R-Squared	0.336	0.3011
Adjusted R-Square	0.302	0.2962
conditional No	8.99E+04	
F-statistic:	9.87E+00	
Prob (F-statistic):	1.31E-25	

Based on regression analysis, the model for multiple linear regression of Stability was developed and presented as following equation, regression coefficients for each variable are summarized in the Table 4-20.

$$VMA = 5.2120 + -0.0002 * PD\ 20 + -0.0000 * PD\ 16 + -0.0001 * PD\ 10 + 0.0021 * PD\ 4.75 + -0.1403 * SG\ 20 + 0.0348 * SG\ 16 + 0.1095 * SG\ 10 + -0.0874 * SGF + 0.0683 * SGB + -0.0276 * PP\ 26.50 + -0.0000 * PP\ 19 + 0.0011 * PP\ 13.20 + 0.0003 * PP\ 9.50 + - \quad \text{Equation 33}$$

*0.0035 * PP 4.75 + 0.0008 * PP 2.36 + 0.0005 * PP 1.18 + -0.0013 * PP 0.60 + -0.0013 * PP 0.30 + -0.0000 * PP 0.15 + -0.0012 * PP 0.075 + 0.0070 * BC*

Table 4-20 Test Statistics and Regression Coefficients for MLR VMA Model

Parameters	C	SE	t	P>|t|
Intercept	5.212	0.657	7.934	0
PD 20	-0.0002	0.001	-0.397	0.692
PD 16	-3.62E-05	0.001	-0.068	0.945
PD 10	-8.77E-05	0	-0.187	0.852
PD 4.75	0.0021	0	4.256	0
SG 20	-0.1403	0.087	-1.611	0.108
SG 16	0.0348	0.088	0.393	0.694
SG 10	0.1095	0.049	2.218	0.027
SGF	-0.0874	0.026	-3.366	0.001
SGB	0.0683	0.2	0.341	0.733
PP 26.50	-0.0276	0.006	-4.563	0
PP 19	-1.15E-05	0	-0.026	0.979
PP 13.20	0.0011	0	2.825	0.005
PP 9.50	0.0003	0	1.093	0.275
PP 4.75	-0.0035	0	-8.627	0
PP 2.36	0.0008	0.001	1.453	0.147
PP 1.18	0.0005	0.001	0.623	0.533
PP 0.60	-0.0013	0.001	-1.37	0.171
PP 0.30	-0.0013	0.001	-1.594	0.112
PP 0.15	-1.61E-05	0.001	-0.011	0.991
PP 0.075	-0.0012	0.002	-0.772	0.441
BC	0.007	0.002	3.11	0.002

Testing of the MLR VMA model was performed using a 25% independent set of data. Overall, the model appears to have low to moderate predictive ability, with R^2=0.4373, explaining only a small portion of the variation in Void in Mineral Aggregate and relation between Actual and Predicted VMA is shown in Figure 4-13. However, there may be other factors not accounted for in the model that also influence the variability of void in mineral aggregate.

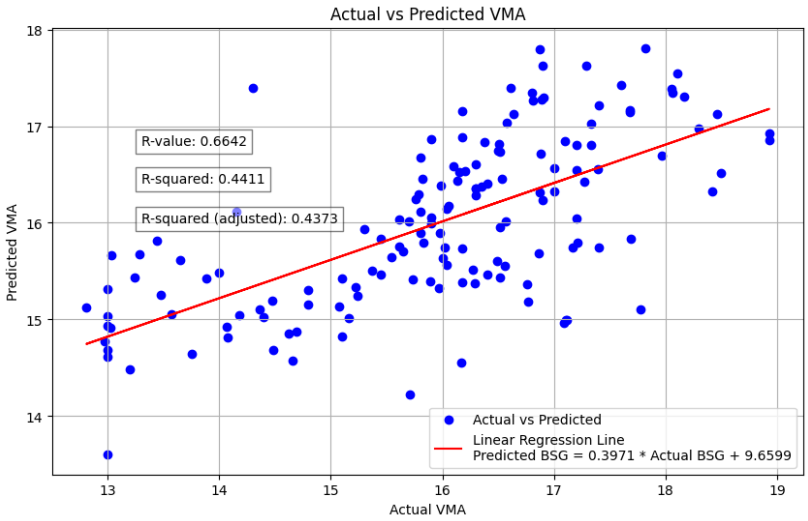


Figure 4-13 Relation Between Actual and Predicted VMA For Testing Set

Figure 4-14 compares the actual Marshall Stability values to those predicted by the MLR model during the testing phase. It also shows the error distribution between predicted and actual values across various observations.

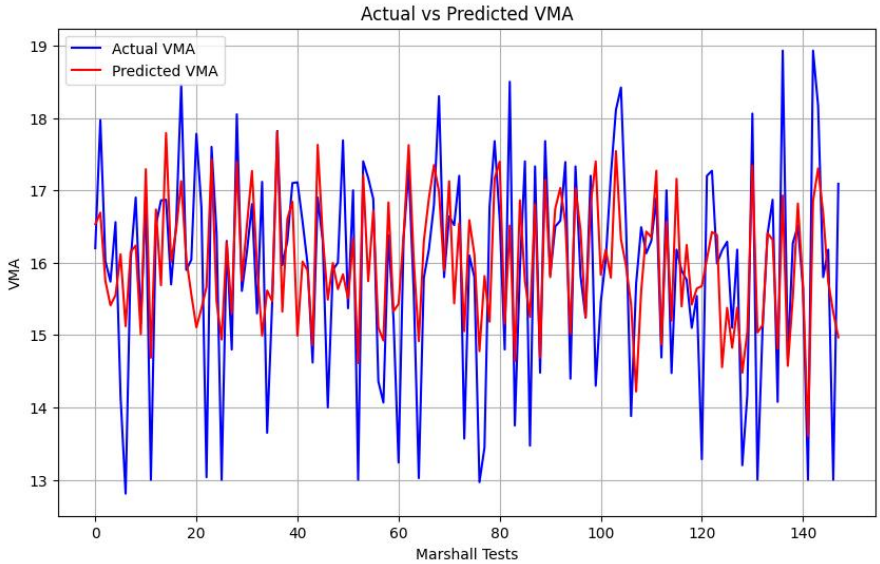


Figure 4-14 Actual VMA and MLR model predicted VMA for Testing set

4.6.1.6 Multiple Liner Regression Analysis Model of VFA

For development of the linear multiple regression model, data were divided into test set and validation set into 80% and 20 % ratio and perform a multiple regression analysis. With 21 independent Parameters. As shown below with 621 input and output data set.

Dependent variable	VFA,
Independent variables	PD 20, PD 16, PD 13, PD 4.75, SG 20, SG 16, SG 10, SGF, SGB, PP 26.50, PP 19, PP 13.20, PP 9.50, PP 4.75, PP 2.36, PP 1.18, PP 0.60, PP 0.30, PP 0.15, PP 0.075, BC
N	441

The developed model and its statistical Parameters are presented in following Table 4-21 at training and testing stage.

Table 4-21 **Statical Results of MLR Model for Percentage Void Filled with Asphalt**

Parameter	At Training Stage	Testing Stage
N	331.000	110
R	0.918	0.9079
R-Squared	0.842	0.8243
Adjusted R-Square	0.831	0.8226
conditional No	9.38E+04	
F-statistic:	7.52E+01	
Prob (F-statistic):	1.33E-105	

Based on regression analysis, the model for multiple linear regression of Stability was developed and presented as following equation and overall, all regression coefficient and test statistics are shown in Table 4-22.

$$VFA = -159.8002 + -0.1976 * PD\ 20 + -0.3585 * PD\ 16 + -0.5571 * PD\ 10 + -0.1490 * PD\ 4.75 + 75.5966 * SG\ 20 + -53.7394 * SG\ 16 + 25.7464 * SG\ 10 + -3.0644 * SGF + -47.7422 * SGB + 1.4114 * PP \quad \text{Equation 34}$$

*26.50 + -0.1230 * PP 19 + -0.2148 * PP 13.20 + 0.1344 * PP 9.50 + -0.0958 * PP 4.75 + 0.0516 * PP 2.36 + -0.2883 * PP 1.18 + 0.3084 * PP 0.60 + -0.1579 * PP 0.30 + 0.8793 * PP 0.15 + 0.4412 * PP 0.075 + 12.1251 * BC*

Table 4-22 Test Statistics and Regression Coefficients for MLR VFA Model

Parameters	C	SE	t	P>|t|
Intercept	-159.8	102.247	-1.563	0.119
PD 20	-0.1976	0.105	-1.88	0.061
PD 16	-3.59E-01	0.098	-3.672	0
PD 10	-5.57E-01	0.092	-6.072	0
PD 4.75	-0.149	0.085	-1.747	0.082
SG 20	75.5966	18.251	4.142	0
SG 16	-53.7394	17.698	-3.036	0.003
SG 10	25.7464	6.682	3.853	0
SGF	-3.0644	4.269	-0.718	0.473
SGB	-47.7422	38.292	-1.247	0.213
PP 26.50	1.4114	0.966	1.46	0.145
PP 19	-1.23E-01	0.065	-1.903	0.058
PP 13.20	-0.2148	0.066	-3.235	0.001
PP 9.50	0.1344	0.063	2.144	0.033
PP 4.75	-0.0958	0.083	-1.155	0.249
PP 2.36	0.0516	0.1	0.517	0.606
PP 1.18	-0.2883	0.173	-1.666	0.097
PP 0.60	0.3084	0.301	1.025	0.306
PP 0.30	-0.1579	0.203	-0.778	0.437
PP 0.15	8.79E-01	0.241	3.654	0
PP 0.075	0.4412	0.275	1.602	0.11
BC	12.1251	0.335	36.169	0

Testing of the MLR VFA model was performed using a 25% independent set of data. Overall, the model appears to have low to moderate predictive ability, with R^2=0.8243 explaining only a small portion of the variation in Void Filled with Asphalt using the given independent variables shown Figure 4-15. However, there may be other factors not accounted for in the model that also influence the variability of Marshall Air Voids.

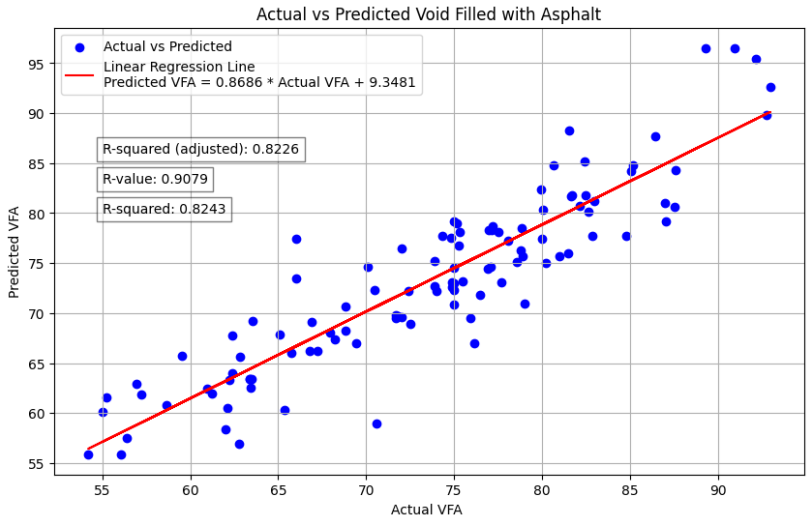


Figure 4-15 Relation between Actual and Predicted VFA for Testing Set

Figure 4-16 compares the actual Marshall Stability values to those predicted by the MLR model during the testing phase. It also shows the error distribution between predicted and actual values across various observations.

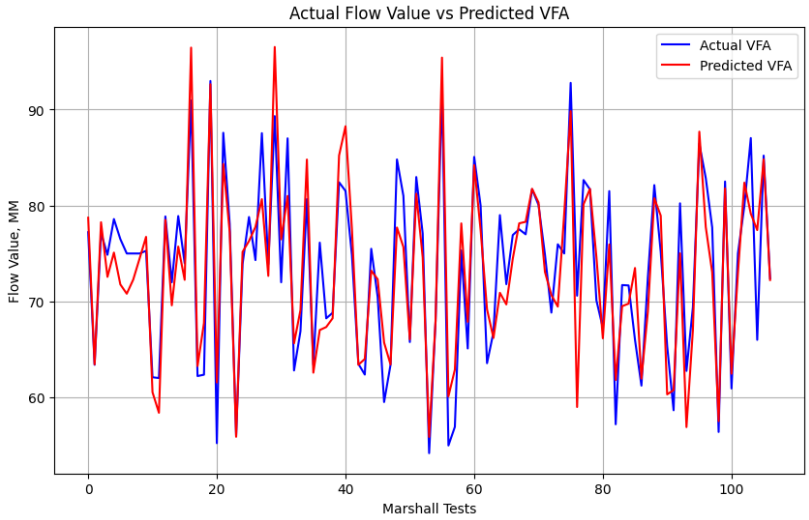


Figure 4-16 Actual VFA and MLR Model Predicted VFA for Testing set

4.6.1.7 Multiple Linear Regression Analysis Model of BSG

The MLR analysis is performed in Excel- Stat plus add-ins based on the 21 Independent parameters such as PD 20, PD 16, PD 13, PD 4.75, SG 20, SG 16, SG 10, SGF, SGB, PP 26.50, PP 19, PP 13.20, PP 9.50, PP 4.75, PP 2.36, PP 1.18, PP 0.60, PP 0.30, PP 0.15, PP 0.075, and BC with total number of **623 samples**. And it can be seen from following Table 4-23 that the performance of MLR shows overfitting of datasets in training and testing stages for the prediction of OBC, with CC as (0.6485, 0.6738), R2 as (0.4205, 0.4496) for both stages.

Table 4-23 Statical Results of MLR model for Bulk Specific Gravity

Parameter	At Training Stage	Testing Stage
N	467	156
R	0.580	0.5488
R-Squared	0.336	0.3011
Adjusted R-Square	0.302	0.2962
conditional No	8.99E+04	
F-statistic:	9.87E+00	
Prob (F-statistic):	1.31E-25	

Based on regression analysis, the model for multiple linear regression of Bulk Specific Gravity was developed and presented as the following equation, and MLR regression coefficients are summarized in Table 4-24.

$$\textit{Bulk Specific Gravity} = 5.2120 + -0.0002 * PD\ 20 + -0.0000 * PD\ 16 + -0.0001 * PD\ 10 + 0.0021 * PD\ 4.75 + -0.1403 * SG\ 20 + 0.0348 * SG\ 16 + 0.1095 * SG\ 10 + -0.0874 * SGF + 0.0683 * SGB + -0.0276 * PP\ 26.50 + -0.0000 * PP\ 19 + 0.0011 * PP\ 13.20 + 0.0003 * PP\ 9.50 + -0.0035 * PP\ 4.75 + 0.0008 * PP\ 2.36 + 0.0005 * PP\ 1.18 + -0.0013 * PP\ 0.60 + -0.0013 * PP\ 0.30 + -0.0000 * PP\ 0.15 + -0.0012 * PP\ 0.075 + 0.0070 * BC \qquad \textit{Equation 35}$$

Table 4-24 Test statistics and regression coefficients for MLR BSG Model

Parameters	C	SE	t	P>|t|
Intercept	5.212	0.657	7.934	0
PD 20	-0.0002	0.001	-0.397	0.692
PD 16	-3.62E-05	0.001	-0.068	0.945

Parameters	C	SE	t	P>|t|
PD 10	-8.77E-05	0	-0.187	0.852
PD 4.75	0.0021	0	4.256	0
SG 20	-0.1403	0.087	-1.611	0.108
SG 16	0.0348	0.088	0.393	0.694
SG 10	0.1095	0.049	2.218	0.027
SGF	-0.0874	0.026	-3.366	0.001
SGB	0.0683	0.2	0.341	0.733
PP 26.50	-0.0276	0.006	-4.563	0
PP 19	-1.15E-05	0	-0.026	0.979
PP 13.20	0.0011	0	2.825	0.005
PP 9.50	0.0003	0	1.093	0.275
PP 4.75	-0.0035	0	-8.627	0
PP 2.36	0.0008	0.001	1.453	0.147
PP 1.18	0.0005	0.001	0.623	0.533
PP 0.60	-0.0013	0.001	-1.37	0.171
PP 0.30	-0.0013	0.001	-1.594	0.112
PP 0.15	-1.61E-05	0.001	-0.011	0.991
PP 0.075	-0.0012	0.002	-0.772	0.441
BC	0.007	0.002	3.11	0.002

Testing of the MLR BSG model was performed using a 25% independent set of data. Overall, the model appears to have low to moderate predictive ability, with R^2=0.2962, explaining only a small portion of the variation in bulk specific gravity using the given independent variables shown in Figure 4-17. However, there may be other factors not accounted for in the model that also influence the variability of bulk specific gravity of Marshall mix. And Figure 4-18 shows the error distribution related to BSG.

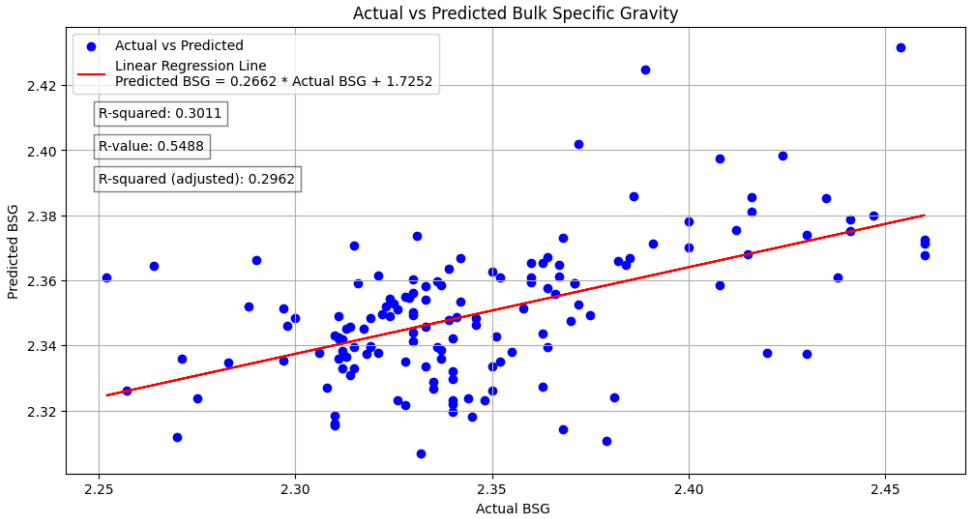


Figure 4-17 Relation between Actual and Predicted BSG For Testing Set

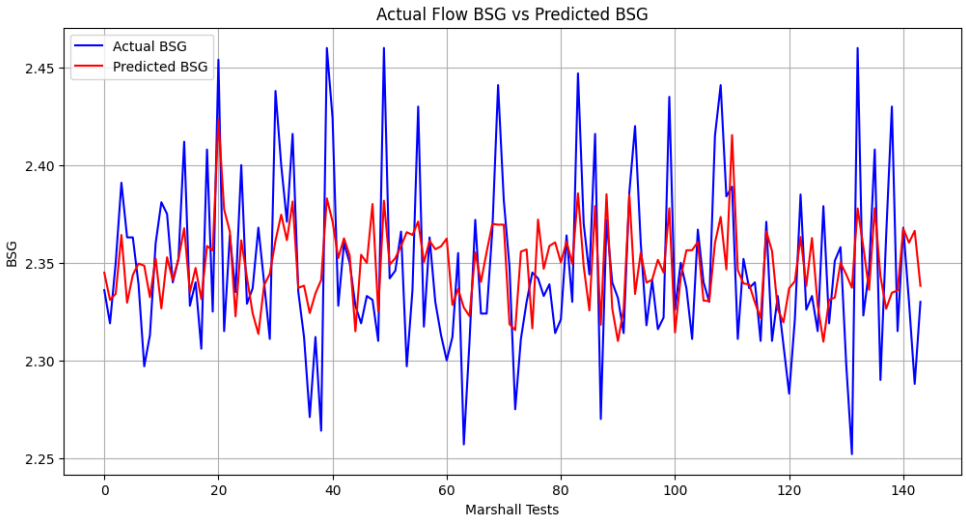


Figure 4-18 Actual BSG and MLR Model Predicted BSG for Testing Set

4.6.2 Statistical result summary of different MLR models related to Asphalt mix Properties

During the development and evaluation of various predictive models for Marshall properties, the data was split into training and testing stages. For the Optimum Binder Content (OBC) model, the correlation coefficient (R) decreased slightly from 0.8354 during training to 0.7857 during testing, with the coefficient of determination (R^2) dropping from 0.698 to 0.6173, indicating a moderate reduction in model accuracy on the testing set. The Stability model showed low predictive power, with R values of 0.55 (training) and 0.5363 (testing) and corresponding R^2 values of 0.303 and 0.2876, highlighting limited explanation of variance. The Flow Value model performed moderately, with R declining from 0.665 to 0.657 and R^2 from 0.442 to 0.4321. The Air Voids model maintained high performance, with R values of 0.812 and 0.806, and R^2 values of 0.66 and 0.6494. The Voids in Mineral Aggregate (VMA) model and Bulk Specific Gravity (BSG) model showed moderate results, while the Voids Filled with Asphalt (VFA) model had the highest predictive ability, with R values of 0.918 and 0.9079 and R^2 values of 0.842 and 0.8243. These metrics suggest variability in model accuracy, with the VFA model being the most reliable. The overall summary is given in Table 4-25.

Table 4-25 Statistical Calculation Results Summary of different MLR models

Marshall Properties	Parameter	At Training Stage	Testing Stage
OBC Model	R-Squared	0.698	0.6173
Stability Model	R-Squared	0.303	0.2876
Flow Value Model	R-Squared	0.442	0.4321
Air Voids Model	R-Squared	0.66	0.6494
VMA Model	R-Squared	0.336	0.3011
VFA Model	R-Squared	0.842	0.8243
BSG Model	R-Squared	0.336	0.3011

4.7 Artificial Neural Network Model

4.7.1 Results Summary for Optimum Bitumen Content ANN model

The model of OBC is developed using data sets containing 20 parameters as independent variables in Python 10.8. For the development of OBC model, 141 data Marshall tests data collected from a different labs and thesis papers are used. Out of 141 Marshall test, some of the samples are presented in APPENDIX B: SAMPLE TRAINING TESTING DATASET FOR MLR, ANN, SVM, & RF 80 % of data set (106 data set) is used for training the models and 20% of data set (35 records) are used for testing of the model. Among the 106 records used for training, 20 % is used for validating the model along with batch size of 64. The decision on the network architecture of the ANN is made based on trial-and-error method by conducting several trials runs as discussed in section 3.4.2 and sample python code is provided in APPENDIX C: SAMPLE CODES FOR ANN. During the training phase of ANN, the learning rate is set as 0.001 and the number of epochs required to deliver the best model is found out by comparing the MSE for each number of iterations performed on the same data. For the training of the model, the number of epochs is set as 250 and optimized by radam & adam optimizer.

For training and testing the dataset, ANN models with different architectures activated by various activation functions such as Rectified Linear Unit (ReLU), Hyperbolic Tangent (tanh), and Logistic Sigmoid (sigmoid) were evaluated. The best model was identified

using metrics including the correlation coefficient (R), coefficient of determination (R^2), mean absolute error (MAE), residual standard error (RSE), and root mean square error (RMSE

4.7.1.1 Rectified Linear Unit (ReLU) Activated ANN Model for OBC

For the ReLU-activated OBC model, hyperparameters were tuned using a trial-and-error approach, varying hidden layers from 30 to 49 and neurons from 35 to 55, evaluating 400 models. The model with 35 hidden layers and 44 neurons had the best performance, with R^2 values of 0.93 (training) and 0.75 (testing), indicating strong predictive accuracy and generalization. It also achieved low MAE (0.04 training, 0.12 testing) and RMSE (0.09 training, 0.16 testing), demonstrating superior performance and lower prediction errors, as shown in Table 4-26.

Table 4-26 Summary of Statistical Calculations Related to ReLU activated ANN Prediction Model for OBC

Training Stage Statistical Results						Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R^2	R	Loss	MAE	R^2	R
30	48	0.01	0.05	0.92	0.96	0.04	0.15	0.64	0.8
31	48	0.01	0.05	0.9	0.95	0.04	0.16	0.61	0.79
32	47	0.01	0.05	0.91	0.96	0.05	0.16	0.56	0.79
33	52	0.01	0.05	0.88	0.94	0.03	0.15	0.68	0.85
34	53	0.01	0.05	0.89	0.95	0.03	0.13	0.68	0.83
35	44	0.01	0.04	0.93	0.97	0.03	0.12	0.75	0.87
36	46	0.01	0.05	0.92	0.96	0.04	0.14	0.65	0.83
37	49	0.01	0.05	0.89	0.94	0.04	0.14	0.65	0.82
38	52	0.02	0.05	0.87	0.94	0.03	0.14	0.68	0.83
39	45	0.01	0.06	0.9	0.95	0.04	0.14	0.63	0.8
40	37	0.01	0.07	0.89	0.95	0.04	0.15	0.63	0.8
41	54	0.01	0.05	0.91	0.96	0.03	0.14	0.74	0.87
42	53	0.01	0.06	0.89	0.95	0.04	0.15	0.63	0.81
43	44	0.01	0.04	0.91	0.96	0.03	0.12	0.73	0.86
44	35	0.01	0.04	0.92	0.96	0.03	0.13	0.69	0.83
45	44	0.02	0.06	0.87	0.94	0.04	0.15	0.62	0.81
46	36	0.02	0.05	0.88	0.94	0.04	0.14	0.62	0.79
47	48	0.01	0.06	0.91	0.95	0.04	0.15	0.58	0.77
48	41	0.02	0.06	0.87	0.94	0.03	0.13	0.69	0.84
49	36	0.02	0.07	0.81	0.91	0.03	0.15	0.69	0.84

4.7.1.2 Hyperbolic Tangent (Tanh) Activated ANN Model for OBC

For the Tanh-activated OBC model, hyperparameters were tuned using a trial-and-error approach, varying hidden layers from 8 to 30 and neurons from 80 to 128, evaluating 1056 models. During the training stage, models show moderate to high predictive abilities, with R^2 values ranging from 0.78 to 0.92 and R values from 0.88 to 0.97. For the testing stage, performance is generally lower, with R^2 values between 0.65 and 0.80 and R values from 0.81 to 0.89. Based on the highest R^2 value in the testing set, the model with 21 hidden layers and 126 neurons per layer achieves the best performance, exhibiting R^2 of 0.80 and R of 0.89, indicating a stronger predictive ability compared to other configurations. The Table 4-27 shows overall summary of the work and example of visualization is provided in APPENDIX D: SAMPLE TRAINING TESTING STAGE VISUILIZATION OF ANN MODELS.

Table 4-27 Summary of Statistical Calculations Related to Tanh Activated ANN Prediction Model for OBC

Training Stage Statistical Results						Testing Stage Statistical Results					
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R	RSE	RMSE
8	85	0.02	0.1	0.84	0.92	0.03	0.13	0.7	0.86	0.18	0.18
9	92	0.02	0.08	0.87	0.93	0.03	0.13	0.72	0.85	0.18	0.17
10	87	0.02	0.1	0.81	0.9	0.03	0.14	0.71	0.86	0.18	0.18
11	127	0.03	0.09	0.8	0.9	0.03	0.13	0.69	0.86	0.19	0.18
12	88	0.01	0.08	0.89	0.94	0.04	0.14	0.66	0.82	0.2	0.19
13	112	0.02	0.1	0.82	0.91	0.03	0.13	0.73	0.86	0.17	0.17
14	82	0.02	0.09	0.85	0.93	0.03	0.13	0.75	0.89	0.17	0.16
15	104	0.02	0.1	0.86	0.93	0.02	0.1	0.78	0.89	0.16	0.15
16	94	0.02	0.09	0.81	0.91	0.03	0.13	0.76	0.88	0.16	0.16
17	115	0.02	0.09	0.84	0.92	0.03	0.11	0.75	0.87	0.17	0.16
18	119	0.01	0.07	0.89	0.94	0.02	0.12	0.78	0.89	0.16	0.15
19	117	0.02	0.08	0.86	0.93	0.03	0.13	0.72	0.85	0.18	0.17
20	121	0.01	0.07	0.92	0.97	0.03	0.14	0.71	0.85	0.18	0.18
21	126	0.02	0.1	0.87	0.94	0.02	0.1	0.8	0.89	0.15	0.15
22	92	0.03	0.11	0.78	0.89	0.03	0.11	0.76	0.87	0.16	0.16
23	115	0.02	0.09	0.87	0.94	0.03	0.13	0.73	0.86	0.17	0.17
24	122	0.02	0.09	0.86	0.93	0.03	0.13	0.7	0.84	0.18	0.18
25	123	0.02	0.09	0.86	0.93	0.03	0.13	0.73	0.85	0.17	0.17
26	106	0.02	0.12	0.82	0.92	0.03	0.14	0.7	0.84	0.18	0.18
27	123	0.02	0.09	0.84	0.92	0.03	0.13	0.71	0.85	0.18	0.18
28	123	0.01	0.08	0.89	0.95	0.03	0.12	0.72	0.86	0.18	0.17
29	80	0.03	0.12	0.79	0.89	0.04	0.14	0.65	0.81	0.2	0.19
30	80	0.03	0.12	0.78	0.88	0.03	0.13	0.7	0.84	0.18	0.18

4.7.1.3 Sigmoid Activated ANN Model for OBC.

For the Sigmoid-activated OBC model, Table 4-28 shows various configurations were tested during the hyperparameter tuning phase, exploring different combinations of hidden layers and neurons. Across the training stage, the models exhibited a range of predictive capabilities, with R^2 values spanning from 0.59 to 0.72 and R values from 0.74 to 0.85. Upon moving to the testing stage, the performance generally declined, with R^2 values ranging from 0.531 to 0.573 and R values from 0.73 to 0.76. Among the tested configurations, the model with 4 hidden layer and 185 neurons achieved the highest R^2 value of 0.573 in the testing set, indicating a moderate predictive ability compared to other configurations.

Table 4-28 Summary of Statistical Calculations Related to Sigmoid Activated ANN Prediction Model for OBC

Sigmoid Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	211	0.04	0.14	0.72	0.85	0.05	0.15	0.541	0.74
2	203	0.04	0.14	0.71	0.84	0.05	0.16	0.531	0.73
3	185	0.04	0.16	0.66	0.81	0.05	0.15	0.561	0.75
4	185	0.05	0.17	0.59	0.77	0.05	0.17	0.573	0.76
5	190	0.05	0.17	0.59	0.77	0.05	0.16	0.543	0.74
6	190	0.06	0.19	0.51	0.74	0.06	0.19	0.478	0.72
7	213	0.13	0.3	0	0.72	0.11	0.29	-0.047	0.77
8	204	0.13	0.3	0	0.49	0.11	0.29	-0.044	0.64

4.7.2 Results for Marshall Stability ANN model

The model of MST is developed using data sets containing 21 parameters as independent variables in Python 10.8. For the development of MST model, 544 data Marshall tests data collected from a different labs and thesis papers are used. Out of 544 Marshall tests, 80 % of data set (435 data set) is used for training the models and 20% of data set (109 records) are used for testing of the model. Among the 435 records used for training, 20 % is used for validating the model along with batch size of 64. The decision on the network architecture of the ANN is made based on literature and by conducting several trials runs as

discussed in section 3.4.2. During the training phase of ANN, the learning rate is set as 0.001 and the number of epochs required to deliver the best model is found out by comparing the MSE for each number of iterations performed on the same data. For the training of the model, the number of epochs is set as 250 and optimized by Adam optimizer.

For training and testing the dataset, ANN models with different architectures activated by various activation functions such as Rectified Linear Unit (ReLU), Hyperbolic Tangent (tanh), and Logistic Sigmoid (sigmoid) were evaluated. The best model was identified using metrics including the correlation coefficient (R), coefficient of determination (R^2), mean absolute error (MAE), residual standard error (RSE), and root mean square error (RMSE.

4.7.2.1 Rectified Linear Unit (ReLU) Activated ANN Model for Marshall ST Model

For the ReLU-activated Marshall Stability ANN model, the hyperparameters were adjusted by testing various configurations, varying the number of hidden layers (NHL) from 15 to 35 and number of Neurons from 20 to 208. Throughout this exploration presented in Table 4-29, a total of 21 independent parameters were utilized. During the training phase, the models displayed a range of predictive capabilities, with R^2 values spanning from 0.89 to 0.95 and R values from 0.94 to 0.98. Conversely, in the testing phase, the performance generally diminished, with R^2 values varying between 0.61 and 0.67, and R values from 0.79 to 0.86. Notably, the model with 29 hidden layers and 134 neurons per layer exhibited the most reliable performance, achieving R^2 of 0.72 and R of 0.86, indicating a moderate to stronger predictive capability compared to other configurations.

Table 4-29 Summary of Statistical Calculations Related to ReLU activated ANN Prediction Model for Marshall ST

Rectified Linear Unit (ReLU) Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
15	145	0.29	0.29	0.94	0.97	1.55	0.9	0.61	0.79
16	146	0.35	0.26	0.93	0.96	1.41	0.87	0.65	0.81
17	150	0.56	0.31	0.89	0.94	1.34	0.85	0.67	0.83

Rectified Linear Unit (ReLU) Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
18	140	0.35	0.27	0.93	0.96	1.41	0.82	0.65	0.81
19	129	0.3	0.24	0.94	0.97	1.33	0.85	0.67	0.82
20	138	0.31	0.24	0.94	0.97	1.35	0.81	0.66	0.82
21	124	0.41	0.25	0.92	0.96	1.12	0.76	0.72	0.85
22	129	0.26	0.24	0.95	0.97	1.23	0.8	0.69	0.83
23	149	0.46	0.42	0.91	0.96	1.37	0.84	0.66	0.83
24	122	0.29	0.25	0.94	0.97	1.26	0.8	0.69	0.84
25	140	0.34	0.36	0.93	0.97	1.34	0.8	0.66	0.82
26	144	0.44	0.41	0.91	0.96	1.36	0.83	0.66	0.81
27	135	0.36	0.32	0.93	0.96	1.33	0.87	0.67	0.82
28	145	0.65	0.56	0.87	0.95	1.48	0.86	0.63	0.81
29	134	0.37	0.3	0.92	0.96	1.12	0.76	0.72	0.86
30	142	0.33	0.37	0.93	0.97	1.54	0.9	0.62	0.82
31	149	0.42	0.33	0.92	0.96	1.5	0.9	0.63	0.8
32	129	0.36	0.29	0.93	0.96	1.12	0.75	0.72	0.85
33	143	0.71	0.61	0.86	0.95	1.42	0.91	0.65	0.81
34	137	0.46	0.37	0.91	0.95	1.49	0.85	0.63	0.8
35	138	0.45	0.33	0.91	0.95	1.38	0.85	0.66	0.82

4.7.2.2 Hyperbolic Tangent (Tanh) Activated ANN Model for Marshall ST

In the exploration of the hyperbolic tangent (Tanh)-activated Marshall Stability ANN model, adjustments to hyperparameters involved varying the number of hidden layers (NHL) from 1 to 15 and neurons from 20 to 128. Across this experimentation, a comprehensive set of 21 independent parameters were employed. During the training stage, diverse predictive capacities were observed and presented in Table 4-30, with R^2 values ranging from 0.85 to 0.93 and R values from 0.92 to 0.96. However, during the testing phase, the model's performance showed a decline, with R^2 values fluctuating between 0.55 and 0.67, and R values from 0.75 to 0.82. Noteworthy was the model configuration comprising 3 hidden layers and 110 neurons per layer, demonstrating the most accurate prediction, yielding R^2 of 0.672 and R of 0.83. This suggests a moderate predictive ability compared to other configurations.

Table 4-30 Summary of statistical calculations related to Tanh activated ANN prediction model for Marshall ST

Hyperbolic Tangent (Tanh) Activated Function									
Training Stage Statistical Results						Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	118	0.75	0.59	0.85	0.92	1.91	1.03	0.583	0.77
2	102	0.65	0.57	0.87	0.93	1.53	0.98	0.666	0.82
3	110	0.42	0.47	0.91	0.96	1.5	0.94	0.672	0.83
4	73	0.5	0.49	0.9	0.95	1.59	0.97	0.653	0.82
5	107	0.48	0.4	0.9	0.95	1.77	0.97	0.613	0.8
6	114	0.42	0.39	0.92	0.96	1.57	0.86	0.658	0.81
7	84	0.52	0.45	0.89	0.95	1.72	1	0.624	0.8
8	63	0.51	0.51	0.9	0.95	1.83	0.97	0.6	0.79
9	42	0.75	0.61	0.85	0.92	1.6	0.97	0.65	0.81
10	125	0.42	0.4	0.91	0.96	1.59	0.93	0.65	0.82
11	117	0.36	0.36	0.93	0.96	1.88	0.98	0.59	0.78
12	70	1.09	0.78	0.78	0.88	1.78	1.01	0.61	0.78
13	114	0.47	0.47	0.9	0.95	2.07	1.08	0.55	0.76
14	54	0.93	0.71	0.81	0.91	2.05	1.07	0.55	0.75
15	99	0.4	0.38	0.92	0.96	1.94	1.02	0.58	0.77

4.7.2.3 Sigmoid Activated ANN Model for Marshall ST

Table 4-31 detailing the performance of the Marshall Stability ANN model activated by the Sigmoid function, the exploration of hyperparameters involved testing various configurations, including adjusting the number of hidden layers (NHL) from 1 to 5 and the number of neurons from 12 to 191. Across this spectrum, a comprehensive array of 21 independent parameters was incorporated. Throughout the training phase, the models showcased diverse predictive abilities, with R^2 values spanning from 0.31 to 0.44 and R values from 0.56 to 0.67. Conversely, during the testing phase, there was a general decline in performance, with R^2 values fluctuating between 0.352 and 0.405, and R values from 0.62 to 0.65. Remarkably, the model with 185 hidden layers and 185 neurons per layer emerged as the best model, demonstrating R^2 of 0.62 and R of 0.65, indicating moderate predictive capability compared to other neurol architect.

Table 4-31 Summary of Statistical Calculations Related to Sigmoid Activated ANN Prediction Model for Marshall ST

Stability Sigmoid									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	190	3.06	1.41	0.38	0.63	2.96	1.35	0.352	0.62
2	185	2.76	1.32	0.44	0.67	2.72	1.29	0.405	0.65
3	177	3.27	1.42	0.33	0.58	2.9	1.33	0.366	0.62
4	176	3.27	1.41	0.33	0.58	2.95	1.36	0.354	0.61
5	181	3.39	1.48	0.31	0.56	3.19	1.42	0.303	0.58

4.7.3 Results for Marshall Flow Value ANN model

The model of MFV is developed using data sets containing 21 parameters as independent variables in Python 10.8. For the development of MFV model, 555 data Marshall tests data collected from different labs and thesis papers are used. Out of 555 Marshall tests, 80 % of data set (444 data set) is used for training the models and 20% of data set (111 records) are used for testing of the model. Among the 444 records used for training, 20 % is used for validating the model along with batch size of 64. The decision on the network architecture of the ANN is made based on literature and by conducting several trials runs as discussed in section 3.4.2.

4.7.3.1 Rectified Linear Unit (ReLU) Activated ANN Model for Marshall FV Model

In search of the optimum ReLU-activated Marshall Flow Value ANN model, with number of hyperparameters arrangement given in Table 4-32, exploring different setups. We played around with the number of hidden layers (NHL), testing the hidden layer from 8 to 17 layers, and the number of neurons from 32 to 132. And we found that the setup with 16 hidden layers and 110 neurons per layer stood out with its stellar performance. It predicted an impressive R^2 of 0.852 and R value of 0.9. This indicates a moderate to stronger predictive ability when compared to the other setups we experimented with.

Table 4-32 Summary of Statistical Calculations Related to ReLU activated ANN Prediction Model for Marshall FV

Rectified Linear Unit (ReLU) Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
8	63	0.04	0.08	0.92	0.96	0.08	0.2	0.81	0.9
9	127	0.02	0.07	0.96	0.98	0.08	0.19	0.818	0.91
10	109	0.02	0.06	0.97	0.98	0.07	0.19	0.826	0.91
11	79	0.02	0.07	0.96	0.98	0.08	0.19	0.817	0.91
12	67	0.02	0.07	0.95	0.97	0.07	0.19	0.832	0.91
13	122	0.02	0.08	0.96	0.98	0.08	0.2	0.822	0.91
14	83	0.02	0.06	0.96	0.98	0.07	0.19	0.826	0.91
15	122	0.02	0.06	0.96	0.98	0.07	0.19	0.832	0.92
16	110	0.02	0.07	0.954	0.977	0.06	0.18	0.8523	0.9
17	130	0.02	0.07	0.962	0.982	0.07	0.19	0.8313	0.9

4.7.3.2 Hyperbolic Tangent (Tanh) Activated ANN Model for Marshall FV

In constructing the Tanh-activated Flow Value ANN model, various hyperparameters were fine-tuned and tested to optimize performance. The exploration spanned different configurations presented in Table 4-33, including adjustments to the number of hidden layers (NHL) ranging from 1 to 15 and the number of neurons per layer (NN) from 30 to 140. Interestingly, the model exhibited notable variation in its performance across different set-ups. Among the configurations tested, the one featuring 6 hidden layers with 140 neurons per layer emerged as particularly noteworthy. This architecture demonstrates R^2 of 0.839 and R value of 0.922. These results signify a robust predictive capacity compared to other configurations explored during experimentation.

Table 4-33 Summary of Statistical Calculations Related to Tanh activated ANN Prediction Model for Marshall FV

Hyperbolic Tangent (Tanh) Activated Function									
Training Stage Statistical Results						Testing Stage Statistical Results			
NHL	NoN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	101	63	0.04	0.08	0.92	0.08	0.2	0.807	0.908
2	107	0.09	0.22	0.79	0.9	0.09	0.22	0.79	0.897
3	99	0.03	0.13	0.93	0.96	0.08	0.19	0.822	0.909
4	43	0.03	0.11	0.93	0.97	0.08	0.21	0.807	0.905
5	128	0.02	0.11	0.95	0.98	0.09	0.21	0.802	0.908

Hyperbolic Tangent (Tanh) Activated Function									
Training Stage Statistical Results						Testing Stage Statistical Results			
NHL	NoN	Loss	MAE	R2	R	Loss	MAE	R2	R
6	140	0.02	0.09	0.96	0.98	0.07	0.18	0.839	0.922
7	54	0.04	0.14	0.92	0.96	0.08	0.21	0.806	0.906
8	105	0.03	0.11	0.94	0.97	0.08	0.2	0.816	0.911
9	140	0.03	0.11	0.94	0.97	0.08	0.18	0.817	0.908
10	109	0.02	0.09	0.97	0.98	0.07	0.19	0.836	0.924
11	125	0.02	0.09	0.97	0.98	0.07	0.18	0.83	0.915
12	104	0.02	0.09	0.96	0.98	0.07	0.18	0.832	0.913
13	89	0.02	0.09	0.96	0.98	0.07	0.18	0.837	0.917
14	99	0.02	0.1	0.95	0.98	0.08	0.2	0.821	0.915
15	131	0.02	0.1	0.95	0.98	0.07	0.19	0.834	0.919

4.7.3.3 Sigmoid Activated ANN model for Marshall FV

In developing the Sigmoid-activated Flow Value ANN model, we tried to refine various hyperparameters to enhance its effectiveness. We tested with different setups given in Table 4-34, adjusting the number of hidden layers (NHL) from 1 to 10 and the number of neurons per layer (NN) from 16 to 127. Such that the model's performance showed significant diversity across these configurations. Model setup with 4 hidden layers and 88 neurons per layer showcased R^2 of 0.839 and R value of 0.922. These outcomes underscore its robust predictive capability compared to the array of configurations we experimented with.

Table 4-34 Summary of Statistical Calculations Related to Sigmoid Activated ANN Prediction Model for Marshall FV

Sigmoid Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	75	0.13	0.26	0.73	0.85	0.12	0.26	0.726	0.86
2	74	0.12	0.24	0.75	0.87	0.12	0.26	0.723	0.86
3	89	0.11	0.22	0.78	0.88	0.11	0.24	0.742	0.87
4	88	0.11	0.24	0.77	0.88	0.1	0.22	0.763	0.88
5	81	0.11	0.22	0.78	0.88	0.12	0.24	0.723	0.85
6	63	0.11	0.23	0.76	0.87	0.11	0.24	0.754	0.87
7	56	0.13	0.26	0.73	0.85	0.14	0.28	0.68	0.83
8	66	0.17	0.31	0.64	0.8	0.17	0.31	0.598	0.78
9	85	0.25	0.37	0.5	0.7	0.22	0.36	0.479	0.7
10	46	0.34	0.43	0.3	0.5	0.29	0.44	0.315	0.6

4.7.4 Results for Percentage Air Voids ANN model

The model of AV is developed using data sets containing 21 parameters as independent variables in Python 10.8. For the development of OBC model, 600 data Marshall tests data collected from different labs and thesis papers are used. Out of 600 Marshall tests, 80 % of data set (480 data set) is used for training the models and 20% of data set (120 records) are used for testing of the model. Among the 480 records used for training, 20 % is used for validating the model along with a batch size of 64. The decision on the network architecture of the ANN is made based on literature and by conducting several trials runs as discussed in section 3.4.2.

4.7.4.1 Rectified Linear Unit (ReLU) Activated ANN Model for Percentage AV Model

In constructing the ReLU-activated Air Voids ANN model, different hyperparameters were adjusted and evaluated to enhance performance. The experiments included varying the number of hidden layers (NHL) from 1 to 15 and the number of neurons per layer (NN) from 12 to 128. The model's performance varied significantly across this architecture. For instance, with a 6 hidden layer and 115 neurons, the model achieved R ^2 of 0.877 and R value of 0.937 in the testing stage. These findings given in Table 4-35 suggest that certain architectures, particularly those with more hidden layers and specific neuron counts, offer improved predictive performance compared to others.

Table 4-35 Summary of Statistical Calculations Related To Relu Activated ANN Prediction Model for Percentage AV

Rectified Linear Unit (ReLU) Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NtN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	75	0.25	0.29	0.91	0.954	0.43	0.43	0.832	0.914
2	113	0.18	0.2	0.933	0.966	0.41	0.4	0.839	0.917
3	63	0.16	0.2	0.942	0.971	0.34	0.39	0.865	0.931
4	116	0.14	0.17	0.95	0.975	0.38	0.39	0.851	0.924
5	63	0.2	0.2	0.926	0.963	0.33	0.37	0.87	0.933
6	115	0.16	0.2	0.94	0.97	0.31	0.37	0.877	0.937
7	107	0.15	0.19	0.946	0.974	0.37	0.41	0.854	0.926

Rectified Linear Unit (ReLU) Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NtN	Loss	MAE	R2	R	Loss	MAE	R2	R
8	109	0.18	0.2	0.932	0.967	0.43	0.43	0.831	0.913
9	83	0.14	0.17	0.948	0.975	0.38	0.39	0.851	0.926
10	49	0.15	0.21	0.945	0.974	0.39	0.38	0.846	0.921
11	120	0.19	0.2	0.929	0.967	0.4	0.41	0.843	0.919
12	61	0.15	0.22	0.945	0.973	0.36	0.41	0.859	0.929
13	52	0.22	0.25	0.921	0.963	0.66	0.51	0.743	0.871
14	86	0.11	0.2	0.96	0.983	0.41	0.42	0.841	0.922
15	68	0.12	0.17	0.955	0.977	0.41	0.41	0.839	0.918

4.7.4.2 Hyperbolic Tangent (Tanh) Activated ANN Model for Percentage AV

In developing the Tanh-activated Flow Value ANN model, a range of hyperparameters were systematically fine-tuned to enhance performance. This involved varying the ANN architecture, with the number of hidden layers (NHL) ranging from 1 to 15 and the number of neurons per layer (NN) ranging from 12 to 128. Performance varied significantly across different architectures. The architecture featuring 3 hidden layers and 59 neurons per layer was particularly notable, achieving R ^2 of 0.854 and R value of 0.925, indicating a superior predictive capability compared to other architectures tested given in Table 4-36.

Table 4-36 Summary of Statistical Calculations Related to Tanh Activated ANN Prediction Model for Percentage AV

Hyperbolic Tangent (Tanh) Activated Function									
Training Stage Statistical Results						Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	65	0.25	0.37	0.909	0.954	0.56	0.49	0.782	0.892
2	79	0.23	0.33	0.915	0.958	0.49	0.44	0.809	0.907
3	59	0.17	0.28	0.939	0.969	0.37	0.39	0.854	0.925
4	53	0.19	0.29	0.93	0.965	0.41	0.42	0.84	0.917
5	78	0.17	0.28	0.938	0.972	0.39	0.41	0.848	0.927
6	77	0.17	0.26	0.939	0.969	0.4	0.43	0.843	0.918
7	118	0.15	0.23	0.946	0.973	0.43	0.42	0.832	0.913
8	16	0.17	0.26	0.938	0.969	0.43	0.46	0.832	0.913
9	34	0.16	0.27	0.942	0.974	0.42	0.42	0.836	0.919
10	63	0.13	0.22	0.954	0.977	0.44	0.42	0.829	0.913
11	18	0.21	0.29	0.923	0.961	0.42	0.43	0.835	0.918
12	31	0.16	0.29	0.941	0.972	0.42	0.44	0.834	0.914
13	128	0.13	0.21	0.952	0.977	0.42	0.44	0.835	0.916
14	128	0.14	0.24	0.949	0.977	0.4	0.43	0.843	0.922
15	66	0.14	0.25	0.948	0.974	0.43	0.45	0.831	0.912

4.7.4.3 Sigmoid Activated ANN Model for Percentage AV

In constructing the ANN model with a sigmoid activation function for AV, various architectures were fine-tuned and evaluated for optimal performance and then presented in Table 4-37. And Model with an architecture with 4 hidden layers and 94 neurons per layer achieved particularly strong results, with R ^2 of 0.771 and R value of 0.879, indicating a robust predictive capability relative to other tested architectures.

Table 4-37 Summary of Statistical Calculations Related to Sigmoid Activated ANN Prediction Model for Percentage AV

Sigmoid Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	96	0.49	0.52	0.822	0.907	0.67	0.57	0.739	0.863
2	118	0.39	0.48	0.856	0.925	0.63	0.55	0.751	0.869
3	69	0.4	0.47	0.855	0.925	0.63	0.54	0.754	0.869
4	94	0.35	0.44	0.872	0.934	0.58	0.52	0.771	0.879
5	120	0.36	0.45	0.87	0.933	0.62	0.55	0.757	0.871
6	125	0.38	0.46	0.859	0.929	0.66	0.57	0.742	0.865
7	93	0.42	0.49	0.845	0.925	0.64	0.57	0.749	0.871
8	112	0.44	0.51	0.838	0.916	0.71	0.59	0.722	0.85
9	98	0.56	0.58	0.794	0.895	0.82	0.65	0.677	0.827
10	79	0.69	0.63	0.747	0.869	0.91	0.69	0.641	0.802
11	29	1.31	0.87	0.52	0.722	1.36	0.94	0.465	0.685

4.7.5 Results for Bulk Specific Gravity ANN Model

The model of BSG is developed using data sets containing 21 parameters as independent variables in Python 10.8. For the development of the BSG model, data Marshall tests data collected from different labs and thesis papers are used. Out of 141 Marshall tests, 80 % of data set (106 data set) is used for training the models and 20% of data set (35 records) are used for testing of the model. Among the 106 records used for training, 20 % is used for validating the model along with a batch size of 64. The decision on the network architecture of the ANN is made based on literature and by conducting several trials runs as discussed in section 3.4.2.

4.7.5.1 Rectified Linear Unit (ReLU) Activated ANN Model for BSG Model

In constructing the ANN model with a ReLU activation function for BSG, various architectures were fine-tuned and evaluated as given in Table 4-38 for optimal performance. The architecture with 44 hidden layers and 102 neurons per layer achieved particularly strong results, with R ^2 of 0.854 and R value of 0.93, indicating robust predictive capability compared to other tested architectures

Table 4-38 Summary of Statistical Calculations Related to ReLU activated ANN Prediction Model for BSG

Rectified Linear Unit (ReLU) Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NoN	Loss	MAE	R2	R	Loss	MAE	R2	R
26	31	0	0.01	0.88	0.95	0	0.02	0.803	0.9
27	100	0	0.01	0.92	0.97	0	0.02	0.82	0.91
28	96	0	0.01	0.94	0.97	0	0.01	0.848	0.92
29	90	0	0.01	0.94	0.97	0	0.01	0.834	0.91
30	98	0	0.01	0.94	0.97	0	0.01	0.821	0.91
31	91	0	0	0.95	0.97	0	0.01	0.827	0.91
32	109	0	0.01	0.95	0.97	0	0.01	0.832	0.92
33	96	0	0.01	0.94	0.97	0	0.01	0.831	0.91
34	91	0	0.01	0.95	0.98	0	0.01	0.845	0.92
40	83	0	0.01	0.93	0.97	0	0.01	0.838	0.92
41	81	0	0	0.95	0.98	0	0.01	0.845	0.92
42	96	0	0.01	0.92	0.97	0	0.01	0.85	0.93
43	100	0	0.01	0.93	0.97	0	0.01	0.844	0.92
44	102	0	0.01	0.94	0.98	0	0.01	0.854	0.93
45	51	0	0.01	0.95	0.97	0	0.01	0.838	0.92
46	107	0	0	0.96	0.98	0	0.01	0.846	0.92
47	86	0	0.01	0.94	0.97	0	0.01	0.834	0.92
48	110	0	0.01	0.95	0.97	0	0.01	0.832	0.91
49	103	0	0.01	0.94	0.97	0	0.01	0.817	0.91

4.7.5.2 Hyperbolic Tangent (Tanh) Activated ANN Model for BSG

In developing the ANN model with a hyperbolic tangent (Tanh) activation function for BSG, different architectures were explored and assessed to determine the most effective configuration. For instance, architectures with 8 hidden layers and 179 neurons per layer

achieved R^2 of 0.837 during testing, suggesting a robust ability to explain variability in the output as provided in Table 4-39.

Table 4-39 Summary of Statistical Calculations Related to Tanh activated ANN Prediction Model for BSG

Hyperbolic Tangent (Tanh) Activated Function									
Training Stage Statistical Results						Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
6	184	0.0003	0.013	0.87	0.94	0.001	0.017	0.759	0.88
7	185	0.0003	0.012	0.87	0.93	0.001	0.0156	0.781	0.88
8	179	0.0002	0.01	0.9	0.95	0	0.0136	0.837	0.92
9	134	0.0003	0.012	0.86	0.93	0	0.0143	0.818	0.9
10	132	0.0003	0.012	0.87	0.94	0	0.014	0.828	0.92
11	152	0.0003	0.012	0.87	0.94	0	0.0152	0.818	0.91
12	142	0.0003	0.013	0.87	0.94	0	0.0149	0.834	0.92
13	175	0.0002	0.01	0.9	0.95	0	0.0139	0.827	0.91
14	171	0.0003	0.012	0.86	0.93	0.001	0.0156	0.807	0.9
15	192	0.0003	0.012	0.87	0.94	0	0.0148	0.82	0.91
16	130	0.0003	0.013	0.85	0.95	0	0.0165	0.812	0.91
17	161	0.0003	0.013	0.86	0.94	0	0.0156	0.828	0.92
18	130	0.0002	0.01	0.89	0.95	0	0.0137	0.832	0.92
19	188	0.0003	0.012	0.87	0.94	0	0.0145	0.834	0.91
20	131	0.0003	0.013	0.85	0.94	0	0.0168	0.812	0.91
21	173	0.0003	0.013	0.85	0.92	0.001	0.0158	0.804	0.9
22	136	0.0003	0.013	0.86	0.93	0.001	0.0147	0.808	0.9
23	139	0.0002	0.011	0.89	0.94	0.001	0.0145	0.804	0.9

4.7.5.3 Sigmoid Activated ANN Model for BSG

In constructing the ANN model with a sigmoid activation function, various architectures were explored and assessed for their effectiveness in predicting BSG. Among the configurations tested, the model with 4 hidden layers and 198 neurons per layer stood out for its moderate performance. This architecture yielded R^2 value of 0.692 and R value of 0.84 during the testing stage, suggesting a robust predictive capability compared to other configurations evaluated as given in Table 4-40.

Table 4-40 Summary of Statistical Calculations Related To Sigmoid Activated ANN Prediction Model for BSG

Sigmoid Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	172	0.001	0.018	0.74	0.87	0.001	0.021	0.682	0.83
2	191	0.001	0.017	0.77	0.88	0.001	0.019	0.688	0.84
3	256	0.001	0.019	0.73	0.86	0.001	0.022	0.66	0.82
4	198	0.001	0.018	0.74	0.88	0.001	0.021	0.692	0.84
5	142	0.001	0.019	0.72	0.86	0.001	0.022	0.652	0.81
6	223	0.001	0.021	0.66	0.82	0.001	0.023	0.621	0.79
7	175	0.001	0.021	0.65	0.81	0.001	0.024	0.571	0.76

4.7.6 Results for Void in Mineral Aggregate ANN model

The model of VMA is developed using data sets containing 21 parameters as independent variables in Python 10.8. For the development of the OBC model, 590 data Marshall tests data collected from different labs and thesis papers are used. Out of 590 Marshall tests, 80 % of data set (472 data set) is used for training the models and 20% of data set (118 records) are used for testing of the model. Among the 472 records used for training, 20 % is used for validating the model along with a batch size of 64. The decision on the network architecture of the ANN is made based on literature and by conducting several trials runs as discussed in section 3.4.2.

4.7.6.1 Rectified Linear Unit (ReLU) Activated ANN Model for VMA

In developing the Artificial Neural Network (ANN) model employing Rectified Linear Unit (ReLU) activation function for predicting Void in Mineral Aggregate (VMA), a thorough exploration of various architectures was undertaken and presented in Table 4-41. Moreover, the ANN architecture with 7 hidden layers and 112 neurons per layer demonstrated superior predictive ability, yielding R ^2 of 0.864 and R value of 0.93 during the testing stage. These results suggest that this particular model configuration exhibits strong predictive capabilities in estimating VMA, outperforming other architectures tested.

Table 4-41 Summary of Statistical Calculations Related to ReLU Aactivated ANN Prediction Model for VMA

Rectified Linear Unit (ReLU) Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	127	0.51	0.46	0.8	0.9	0.77	0.66	0.749	0.87
2	56	0.27	0.33	0.89	0.95	0.45	0.48	0.853	0.92
3	79	0.16	0.23	0.94	0.97	0.49	0.51	0.84	0.92
4	81	0.16	0.22	0.94	0.97	0.48	0.47	0.843	0.92
5	61	0.2	0.24	0.92	0.96	0.46	0.49	0.85	0.92
6	127	0.18	0.26	0.93	0.97	0.49	0.5	0.842	0.92
7	112	0.13	0.19	0.95	0.97	0.42	0.44	0.864	0.93
8	70	0.17	0.25	0.93	0.97	0.44	0.48	0.857	0.93
9	49	0.19	0.24	0.925	0.962	0.51	0.5	0.8353	0.9
10	119	0.18	0.27	0.929	0.968	0.49	0.52	0.8411	0.9
11	115	0.15	0.21	0.94	0.97	0.46	0.48	0.85	0.92
12	80	0.18	0.24	0.93	0.97	0.49	0.51	0.84	0.93
13	56	0.13	0.22	0.95	0.97	0.47	0.48	0.85	0.92
14	76	0.18	0.3	0.93	0.97	0.46	0.5	0.85	0.93
15	80	0.14	0.23	0.94	0.97	0.42	0.49	0.86	0.93

4.7.6.2 Hyperbolic Tangent (Tanh) Activated ANN Model for VMA

In developing the Artificial Neural Network (ANN) model utilizing the Hyperbolic Tangent (Tanh) activation function, a systematic exploration of different architectures was conducted to ascertain the most effective configuration. Among the configurations tested presented in Table 4-42, the model with 6 hidden layers and 106 neurons per layer consistently demonstrated notable performance across both training and testing stages. It exhibited strong statistical results, including R^2 value of 0.87 and R value of 0.93, indicating strong predictive capability regarding the variation in VMA explained by the independent variables considered.

Table 4-42 Summary of Statistical Calculations Related to Tanh activated ANN Prediction Model for VMA

Hyperbolic Tangent (Tanh) Activated Function									
Training Stage Statistical Results						Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	98	0.41	0.46	0.84	0.92	0.58	0.56	0.81	0.9
2	101	0.32	0.42	0.87	0.94	0.63	0.53	0.8	0.89
3	84	0.27	0.38	0.89	0.95	0.63	0.57	0.8	0.9

Hyperbolic Tangent (Tanh) Activated Function									
Training Stage Statistical Results						Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
4	109	0.26	0.35	0.9	0.95	0.53	0.54	0.83	0.91
5	128	0.21	0.33	0.92	0.96	0.49	0.51	0.84	0.92
6	106	0.21	0.32	0.92	0.96	0.42	0.46	0.87	0.93
7	88	0.3	0.38	0.88	0.94	0.55	0.54	0.82	0.91
8	124	0.2	0.31	0.92	0.96	0.55	0.53	0.82	0.91
9	112	0.2	0.32	0.92	0.96	0.58	0.55	0.81	0.9
10	128	0.29	0.38	0.89	0.94	0.56	0.53	0.82	0.91
11	118	0.32	0.44	0.87	0.95	0.64	0.63	0.79	0.91
12	112	0.35	0.44	0.86	0.93	0.73	0.64	0.76	0.88
13	112	0.29	0.41	0.88	0.94	0.65	0.59	0.79	0.89
14	120	0.24	0.34	0.9	0.95	0.65	0.6	0.79	0.9
15	116	0.39	0.42	0.84	0.92	0.7	0.57	0.77	0.88

4.7.6.3 Sigmoid Activated ANN Model for Marshall VMA

In constructing the ANN model with a sigmoid activation function, various architectures presented in Table 4-43 were explored and assessed for their effectiveness in predicting BSG. Among the configurations tested, the model with 3 hidden layers and 297 neurons per layer stood out for its moderate performance. This architecture yielded R^2 value of 0.641 and R value of 0.81 during the testing stage, suggesting a moderate predictive capability compared to other configurations evaluated.

Table 4-43 Summary of Statistical Calculations Related to Sigmoid activated ANN Prediction Model for VMA

Sigmoid Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	286	0.76	0.66	0.7	0.84	1.2	0.82	0.611	0.79
2	279	0.68	0.62	0.73	0.85	1.11	0.81	0.639	0.8
3	297	0.63	0.59	0.75	0.87	1.11	0.8	0.641	0.81
4	237	1.04	0.75	0.59	0.77	1.5	0.94	0.513	0.72

4.7.7 Results for Void Filled with Asphalt ANN model

The model of OBC is developed using data sets containing 21 parameters as independent variables in Python 10.8. For the development of OBC model, 426 data Marshall tests data

collected from different labs and thesis papers are used. Out of 426 Marshall tests, 80 % of data set (340 data set) is used for training the models and 20% of data set (86 records) are used for testing of the model. Among the 340 records used for training, 20 % is used for validating the model along with a batch size of 64. The decision on the network architecture of the ANN is made based on literature and by conducting several trials runs as discussed in section 3.4.2.

4.7.7.1 Rectified Linear Unit (ReLU) Activated ANN Model for VFA

In constructing the ANN model with the Rectified Linear Unit (ReLU) activation function, different architectures were explored and assessed for their effectiveness. Overall summary of work is presented in Table 4-44 among the architecture tested, the model with 14 hidden layers and 56 neurons per layer exhibited good performance, demonstrating R^2 of 0.91 and R value of 0.96 during the testing stage. These results suggest a robust predictive ability of the model, surpassing other architectures evaluated.

Table 4-44 Summary of Statistical Calculations Related to ReLU activated ANN Prediction Model for VMA

Rectified Linear Unit (ReLU) Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	126	21.91	3.25	0.8	0.9	21.97	3.66	0.81	0.9
2	121	12.18	2.35	0.89	0.94	16.01	3.1	0.87	0.93
3	104	9.43	1.69	0.91	0.96	12.68	2.67	0.89	0.95
4	95	8.93	1.68	0.92	0.96	13.15	2.66	0.89	0.95
5	91	8.8	1.55	0.92	0.96	11.97	2.63	0.9	0.95
6	26	9.48	1.91	0.91	0.96	12.03	2.6	0.9	0.95
7	87	7.33	1.36	0.93	0.97	12.32	2.55	0.9	0.95
8	82	7.28	1.35	0.93	0.97	11.64	2.54	0.9	0.95
9	32	8.48	1.52	0.92	0.96	11.27	2.5	0.9	0.96
10	82	7.95	1.33	0.93	0.96	12.44	2.53	0.9	0.95
11	57	8	1.38	0.93	0.96	11.06	2.35	0.91	0.95
12	54	8.23	1.42	0.92	0.96	10.7	2.34	0.91	0.96
13	115	6.86	1.35	0.94	0.97	10.8	2.42	0.91	0.95
14	56	8.51	1.51	0.92	0.96	10.11	2.31	0.91	0.96
15	13	9.17	1.69	0.91	0.96	12.38	2.54	0.9	0.95

4.7.7.2 Hyperbolic Tangent (Tanh) Activated ANN Model for VFA

In constructing the ANN model with a hyperbolic tangent (Tanh) activation function, a range of architectures presented in Table 4-45 were explored and analyzed to determine the most effective configuration. For instance, the architecture with 2 hidden layers and 148 neurons per layer consistently outperforms other models in terms of predictive ability, with higher R^2 values approaching 0.91 and strong correlation coefficients (R) exceeding 0.96 during the testing stage. This indicates robust predictive capability relative to other tested architectures,

Table 4-45 Summary of Statistical Calculations Related to Tanh activated ANN Prediction Model for VFA

Hyperbolic Tangent (Tanh) Activated Function									
Training Stage Statistical Results						Testing Stage Statistical Results			
NHL	NN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	156	8.27	1.67	0.92	0.96	11.02	2.35	0.91	0.95
2	148	6.25	1.2	0.94	0.97	10.83	2.23	0.91	0.96
3	135	6.02	1.14	0.94	0.97	12.88	2.36	0.89	0.95
4	138	6.61	1.27	0.94	0.97	11.85	2.33	0.9	0.95
5	148	6.79	1.32	0.94	0.97	12.3	2.47	0.9	0.95
6	107	10.55	2.18	0.9	0.96	13.14	2.8	0.89	0.94
6	125	8.65	1.75	0.92	0.96	13.19	2.71	0.89	0.94
7	126	9.24	1.86	0.91	0.96	14.97	2.89	0.87	0.94
8	145	40.76	4.73	0.62	0.79	43.53	5.08	0.63	0.8

4.7.7.3 Sigmoid Activated ANN Model for FVA

In developing the ANN model with a sigmoid activation function, different architectures were explored and assessed to ascertain the most effective hyperparameters as given in Table 4-46. Among these, the model featuring 2 hidden layers, each comprising 126 neurons, predicted more accurately. This architecture demonstrated good statistical outcomes during both the training and testing stages. Specifically, it achieved R^2 value of 0.841, indicating a high level of explained variance, along with R value of 0.92, suggesting strong linear correlation for testing stage dataset.

Table 4-46 Summary of Statistical Calculations Related to Sigmoid Activated ANN Prediction Model for VFA

Sigmoid Activation Function									
		Training Stage Statistical Results				Testing Stage Statistical Results			
NHL	NoN	Loss	MAE	R2	R	Loss	MAE	R2	R
1	126	16.48	2.83	0.85	0.92	22.31	3.89	0.812	0.9
2	126	13.14	2.41	0.88	0.94	18.86	3.57	0.841	0.92
3	120	107.05	8.35	0	0.85	118.69	8.69	-0.001	0.86

4.7.8 ANN Results Summary

In developing the ANN model for predicting Marshall parameters, different activation functions were assessed to determine their effectiveness. For the **Optimum Binder Content (OBC)**, the **tanh** activation function exhibited the highest predictive accuracy, with R ^2 value of 0.796. When modeling **Marshall Stability (MST)**, **ReLU** proved to be the most effective activation function, achieving R ^2 value of 0.7206. For predicting the **Marshall Flow Value (MFV)**, the **ReLU** function again showed superior performance with R ^2 of 0.8523. The **Air Voids (AV)** parameter was best modeled using **ReLU**, which attained R ^2 value of 0.877. Similarly, for **Bulk Specific Gravity (BSG)**, **ReLU** was the optimal activation function with R ^2 of 0.8542. In the case of **Voids in Mineral Aggregate (VMA)**, the **tanh** activation function demonstrated the highest R^2, measuring 0.8650. Lastly, for **Voids Filled with Asphalt (VFA)**, **ReLU** once again provided the best predictive performance, with R ^2 of 0.91. Therefore, **ReLU** generally emerges as the most effective activation function for most Marshall parameters, though **tanh** also shows competitive performance in specific instances. Overall statistical calculation summary is shown in Table 4-47.

Table 4-47 Summary of Best ANN Architect Based on Different Activation Function

Properties	Activation	HL	NN	Test State Statistical Result			
				Loss	MAE	R2	R
OBC	ReLU	35	44	0.03	0.12	0.75	0.87
	tanh	21	126	0.02	0.1	0.796	0.89
	sigmoid	4	185	0.05	0.17	0.573	0.76
MST	ReLU	21	124	1.12	0.76	0.7206	0.85
	tanh	3	110	1.5	0.94	0.672	0.83

Properties	Activation	HL	NN	Test State Statistical Result			
				Loss	MAE	R2	R
	sigmoid	2	185	2.72	1.29	0.405	0.65
MFV	ReLU	16	110	0.0633	0.183	0.8523	0.9247
	tanh	6	140	0.07	0.18	0.839	0.922
	sigmoid	4	88	0.1	0.22	0.763	0.88
AV	ReLU	6	115	0.31	0.37	0.877	0.937
	tanh	3	59	0.37	0.39	0.854	0.925
	sigmoid	4	94	0.58	0.52	0.771	0.879
BSG	ReLU	44	102	0.0003	0.0135	0.8542	0.9279
	tanh	8	179	0.0004	0.01	0.837	0.92
	sigmoid	4	198	0.001	0.021	0.692	0.84
VMA	ReLU	7	112	0.42	0.44	0.864	0.93
	tanh	6	106	0.42	0.46	0.865	0.93
	sigmoid	3	297	1.11	0.8	0.641	0.81
VFA	ReLU	14	56	10.11	2.31	0.91	0.96
	tanh	2	148	10.83	2.23	0.909	0.96
	sigmoid	2	126	18.86	3.57	0.841	0.92

4.8 Support Vector Machine Model

In the analysis of Marshall parameters using a Support Vector Machine model, the training and testing stages provided insights into the model's performance for each parameter. These parameters include Optimum Binder Content (OBC), Stability (MST), Flow Value(MFV), Air Voids (AV), Bulk Specific Gravity (BSG), Voids in Mineral Aggregate (VMA), and Voids Filled with Asphalt (VFA). For each parameter, we can identify the optimal SVM model hyperparameters presented in Table 4-48, based on the least MAPE statistical matrix using 80% data set for training stage and 20% dataset for testing stage. The python code used in Random Forest Regression is shown in Table 4-48**.**

Table 4-48 Summary of Optimum Hyperparameters for different SVM Regression Models for Marshall Parameters

Marshall Parameters	Optimum Hyper Parameters for Support Vector Machine Regression		
	Kernel	C	Epsilon
OBC	rbf	0.8	0.007
Stability	rbf	10	0.007
Flow Value	rbf	10	0.007
Air Voids	rbf	50	0.1
BSG	rbf	10	0.007
VMA	rbf	50	0.1

Marshall Parameters	Optimum Hyper Parameters for Support Vector Machine Regression		
	Kernel	C	Epsilon
VFA	rbf	100	0.1

The Table 4-49 presents the statistical evaluation of a Support Vector Machine (SVM) model applied to Marshall parameters during both training and testing stages. For each Marshall parameter (OBC, MST, MFV, VMA, AV, BSG, VFA), key performance metrics are provided. During the training stage, the SVM model demonstrates high predictive accuracy for most parameters, with R^2 values ranging from 0.872 to 0.961, indicating strong correlation between predicted and actual values. However, during testing, the model's performance varies. While some parameters maintain strong predictive ability ($R^2 > 0.8$), others exhibit lower accuracy ($R^2 < 0.75$), suggesting potential limitations in generalization to unseen data. Notably, VFA shows substantial discrepancies between training and testing stages, with high RMSE and RSE during testing, indicating significant prediction errors. Overall, the table highlights the SVM model's efficacy in modeling Marshall parameters during training, but underscores the importance of robustness testing to assess its reliability across diverse datasets.

Table 4-49 Summary of Statistical Calculations Related to SVM Prediction Model For Marshall Parameters

Marshall Parameters	Training Stage Statistical Results				Testing Stage Statistical Results			
	MAPE	MAE	R^2	R	MAPE	MAE	R^2	R
OBC	0.008	0.042	0.931	0.965	0.026	0.137	0.642	0.803
MST	0.041	0.49	0.872	0.935	0.062	0.724	0.739	0.861
MFV	0.024	0.067	0.951	0.976	0.061	0.163	0.813	0.904
VMA	0.01	0.16	0.954	0.977	0.023	0.372	0.824	0.911
AV	0.051	0.164	0.953	0.976	0.108	0.372	0.814	0.904
BSG	0.003	0.007	0.925	0.962	0.005	0.011	0.805	0.901
VFA	0.014	0.972	0.961	0.98	0.024	1.742	0.923	0.962

The fitting performance of the Support Vector regression model for OBC is shown in Figure 4-19 for training and Figure 4-20 for testing stage. The predicted value of OBC is marked by orange points/line and actual value of OBC is marked by blue points/line. The vertical difference between the orange and blue points represents absolute error.

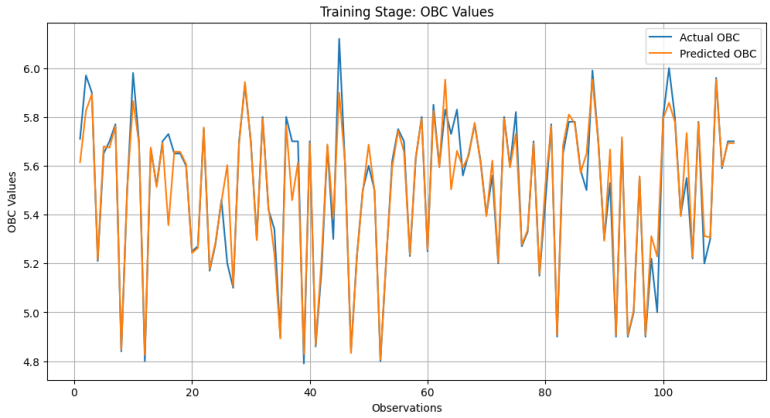


Figure 4-19 Training Stage SVM model fitting for OBC

Figure 4-20 Testing Stage SVM model fitting for OBC

4.9 Random Forest Model

In the analysis of Marshall parameters using a Random Forest model, the training and testing stages provided insights into the model's performance for each parameter. These parameters include Optimum Binder Content (OBC), Stability Flow (MST), Flow Value (MFV), Air Voids (AV), Bulk Specific Gravity (BSG), Voids in Mineral Aggregate (VMA), and Voids Filled with Asphalt (VFA). For each parameter, we can identify the optimal RF model function presented in, based on the least MAPE statistical matrix using 80% data set for training stage and 20% dataset for testing stage. The python code used in Random Forest Regression is shown in **APPENDIX.**

Table 4-50 Summary of Optimum Hyperparameters for Different RF Regression Models for Marshall Parameters

Marshall Parameters	Optimum Hyper parameters		
	Maximum Depth	Min Sample Leaf	Estimators
VMA	None	2	100
VFA	20	2	250
BSG	None	2	250
AV	None	2	50
MFV	20	2	150
MST	None	2	300
OBC	None	2	200

The Table 4-51 presents statistical results from the training and testing stages of a Random Forest model applied to various Marshall parameters, including OBC, MST, MFV, AV, BSG, VMA, and VFA. For each parameter, performance metrics are presented. During training, the model shows high predictive accuracy across all parameters, with R^2 values ranging from 0.957 to 0.976 and low error metrics. However, in the testing stage, the performance declines, with R^2 values between 0.605 and 0.844, indicating moderate predictive ability. The discrepancies between training and testing results suggest some overfitting, where the model performs well on training data but less effectively on unseen data, highlighting areas for model refinement.

Table 4-51 Summary of Statistical Calculations Related to RF Prediction Model for Marshall Parameters

Marshall Parameters	Parameter Evaluation							
	Training Stage Statistical Results				Testing Stage Statistical Results			
	MAE	R^2	R	RMSE	MAE	R^2	R	RMSE
0BC	0.051	0.957	0.982	0.068	0.051	0.605	0.798	0.18
MST	0.311	0.962	0.985	0.435	0.311	0.716	0.859	1.066
MFV	0.072	0.969	0.988	0.105	0.072	0.809	0.904	0.264
AV	0.174	0.964	0.984	0.254	0.174	0.761	0.874	0.667
BSG	0.004	0.965	0.984	0.007	0.004	0.812	0.903	0.019
VMA	0.158	0.959	0.981	0.266	0.158	0.81	0.901	0.607
VFA	1.003	0.976	0.988	1.462	1.003	0.844	0.92	3.634

The fitting performance of the RF model for OBC is shown in Figure 4-21for training and Figure 4-22 for testing stage. The predicted value of OBC is marked by orange points/line and actual value of OBC is marked by blue points/line. The vertical difference between the orange and blue points represents absolute error.

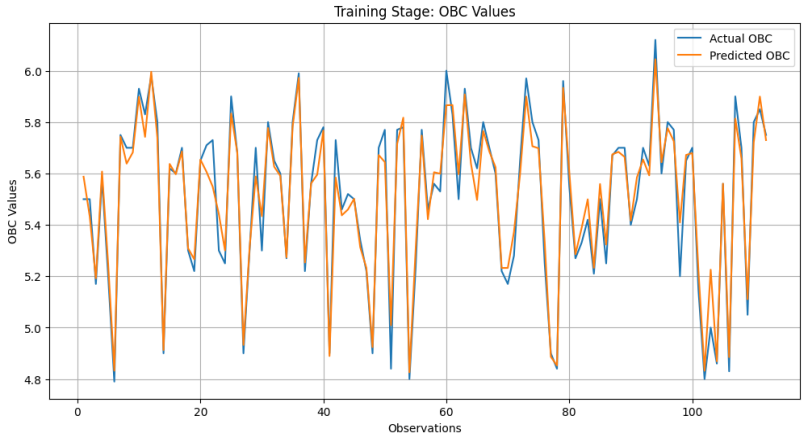


Figure 4-21 Training Stage RF model fitting for OBC

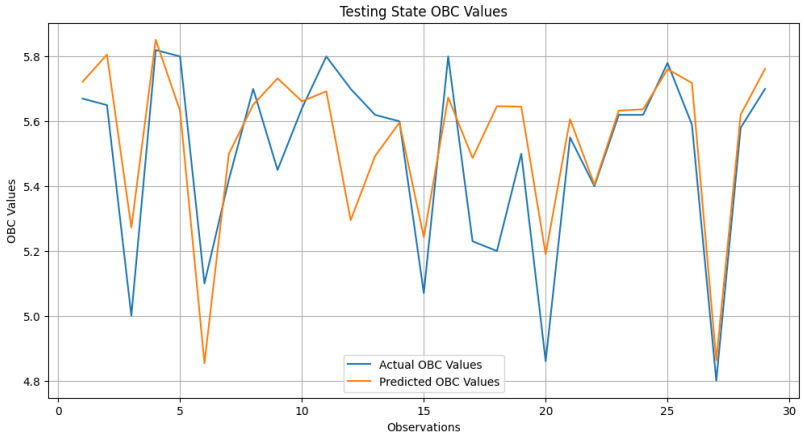


Figure 4-22 Testing Stage RF Model Fitting for OBC

4.10 Summary of Model Comparison and selection of best prediction model

In comparing the performance of Multiple Linear Regression (MLR) and Artificial Neural Network (ANN) models across various asphalt properties, the ANN consistently demonstrates superior predictive capabilities. Across all Asphalt Mix Properties, the ANN

achieves higher coefficients of determination (R-Square) and lower mean squared errors (MSE) on the validation sets compared to MLR. When comparing the predictive capabilities of MLR, ANN, SVM, and Random Forest models across various asphalt mix properties, the R-squared values and MAE metrics offer insights into their relative performance.

For Optimum Binder Content (OBC), ANN models consistently demonstrate superior performance, with R -squared value of 0.796 indicating 79% variability of OBC is explained by independent variables. In predicting Marshall Stability, the SVM model yields the highest R-squared value of 0.739, while ANN and Random Forest models perform best, with R -squared value of 0.724. Similarly, for Marshall Flow Value, ANN shows the highest R-squared value of 0. 8523.Moving to Air Voids, ANN models again show superior performance, with R -squared value of 0.876 showcasing its superior predictive accuracy among other models. When considering Voids in Mineral Aggregate, ANN models exhibit the highest R-squared value of 0.865. Lastly, for Voids Filled with Asphalt, SVM model shows the highest R-squared value of 0.92. Overall, and ANN models consistently demonstrate superior predictive accuracy across all Marshall Mix properties, as evidenced by their competitive lower MAE values compared to SVM and RF models.

When comparing the predictive capabilities of MLR, ANN, SVM, and Random Forest models across various asphalt mix properties, SVM and Random Forest consistently demonstrate superior performance based on the Mean Absolute Error (MAE). For Optimum Binder Content (OBC), both SVM and Random Forest models achieve the lowest MAE of 0.051, indicating the highest accuracy. In predicting Marshall Stability, these models again perform best, with an MAE of 0.311. Similarly, for Marshall Flow Value (MFV), SVM and Random Forest models show the lowest MAE of 0.072. This trend continues across other properties: Air Voids (AV) with an MAE of 0.174, Bulk Specific Gravity (BSG) with an MAE of 0.004, Voids in Mineral Aggregate (VMA) with an MAE of 0.158, and Voids Filled with Asphalt (VFA) with an MAE of 1.003. The summary of the comparison is presented in Table 4-52. In conclusion, the consistent performance of SVM and Random Forest models in achieving the lowest MAE across all Marshall Mix properties highlights their superior predictive accuracy compared to MLR and ANN.

Table 4-52 Summary of Model Comparison and Selection of Best Prediction Model for Respective Method.

Properties	MLR	ANN	SVM	RF
	R^2	R^2	R^2	R^2
OBC	0.6173	0.7964	0.642	0.605
MS	0.2876	0.7206	0.739	0.716
MFV	0.4321	0.8523	0.813	0.809
AV	0.6494	0.8765	0.814	0.761
BSG	0.3011	0.8542	0.805	0.812
VMA	0.3011	0.865	0.824	0.81
VFA	0.8243	0.9147	0.923	0.844

CHAPTER 5: IMPLEMENTATION FRAMEWORK

The implementation framework, developed in python 3.10 and sample code and GUI interface is provided in APPENDIX E: SAMPLE CODES FOR GUI WITH VISUZILATION, focuses on predicting Marshall Parameters using advanced machine learning models and initially, the framework determines the Optimum Bitumen Content based on 20 independent variables. Subsequently, Marshall's Mechanical and Volumetric properties are predicted using the Optimum Bitumen Content along with 21 independent variables, as detailed in Table 5-1. The predictive methodology employs the best-performing Artificial Neural Network (ANN), Random Forest (RF), and Support Vector Machine (SVM) models. These models are utilized sequentially to predict the Optimum Bitumen Content and then to forecast Marshall's Mechanical and Volumetric properties.

Table 5-1 Parameters Used in the Model

Independent variables	Values		Independent variables	Values	
	NAS 13.2	NAS 19		NAS 13.2	NAS 19
PD 20	0	25	PP 19	100	100.00
PD16	15	18	PP13.2	97.15	78.48
PD 10	35	12	PP 9.5	83.55	70.74
PD 4.75	50	45	PP 4.75	59.47	45.95
SG 20	2.648	2.706	PP 2.36	47.3	30.70
SG 16	2.648	2.693	PP 1.18	35.57	21.31
SG 10	2.623	2.680	PP 0.6	26.06	15.58
SGF	2.58	2.528	PP 0.3	18.19	11.29
SGB	1.03	1.019	PP 0.15	13.35	9.08
PP 26	100	100	PP 0.075	9.93	7.12

Based on the independent variables related to NAS 13.2 mix, predictions were made for Marshall mix parameters, and the deviations in these parameters were presented in Table

5-2. Figure 5-1 and Figure 5-2 provide representations of the predicted values for OBC and asphalt mix properties in the GUI.

From the Table 5-2, it is evident that Artificial Neural Network (ANN) models are more reliable in predicting mechanical and volumetric properties, with the exception of stability. On the other hand, Random Forest (RF) models showed accurate predictions for OBC (Optimum Bitumen Content), while Support Vector Machine (SVM) models excelled in predicting stability. This indicates that different machine learning models have varying strengths in predicting specific aspects of NAS 13.2 mix properties, highlighting the importance of choosing the right model depending on the desired prediction outcomes.

Table 5-2 Predicted Mechanical and Volumetric Properties of Asphalt Mix of NAS 13.2mm

Marshall Mix Parameters for Mix		ANN		SVR		RF	
	Actual value	PV	Percentage Deviation	PV	Percentage Deviation	PV	Percentage Deviation
OBC, % weight	5.7	5.63	1.23%	5.3	7.02%	5.65	0.88%
MST, KN	11.34	14.4	26.98%	11.78	3.88%	11.99	5.73%
MFV, mm	3.22	2.81	12.73%	2.61	18.94%	2.88	10.56%
AV, %	4	3.86	3.50%	4.95	23.75%	2.48	38.00%
BSG, g/cc	2.32	2.34	0.86%	2.33	0.43%	2.36	1.72%
VMA, %	16.8	15.16	9.76%	16.56	1.43%	16.46	2.02%
VFA		79.28		67.83		82.99	
Average absolute Error			9.18%		9.24%		9.82%

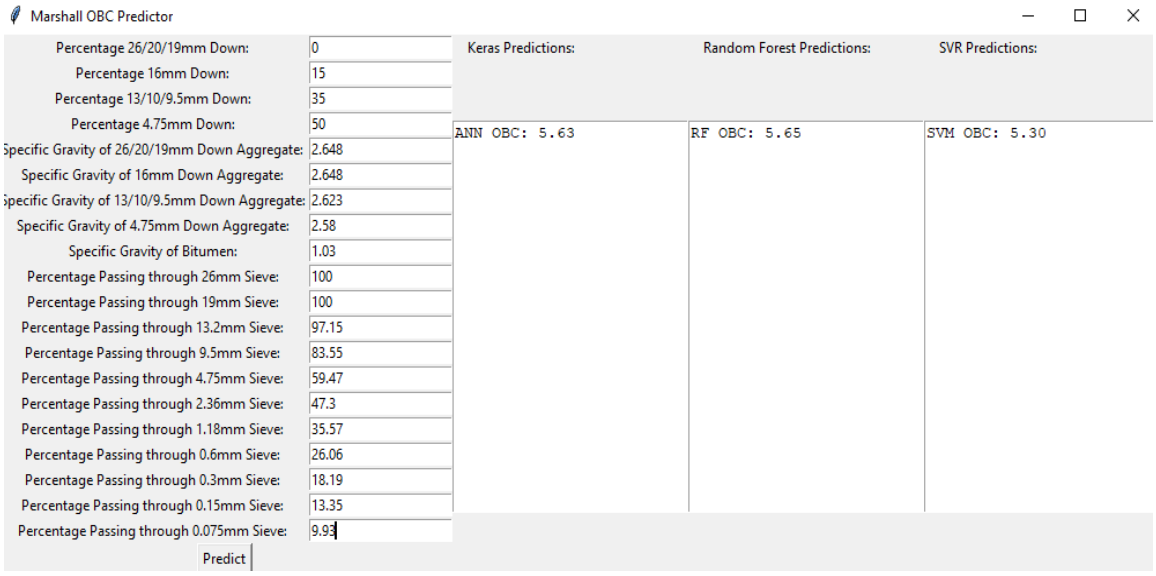


Figure 5-1 OBC Prediction Graphic User Interface and Predicted Values for NAS 13.2 mix

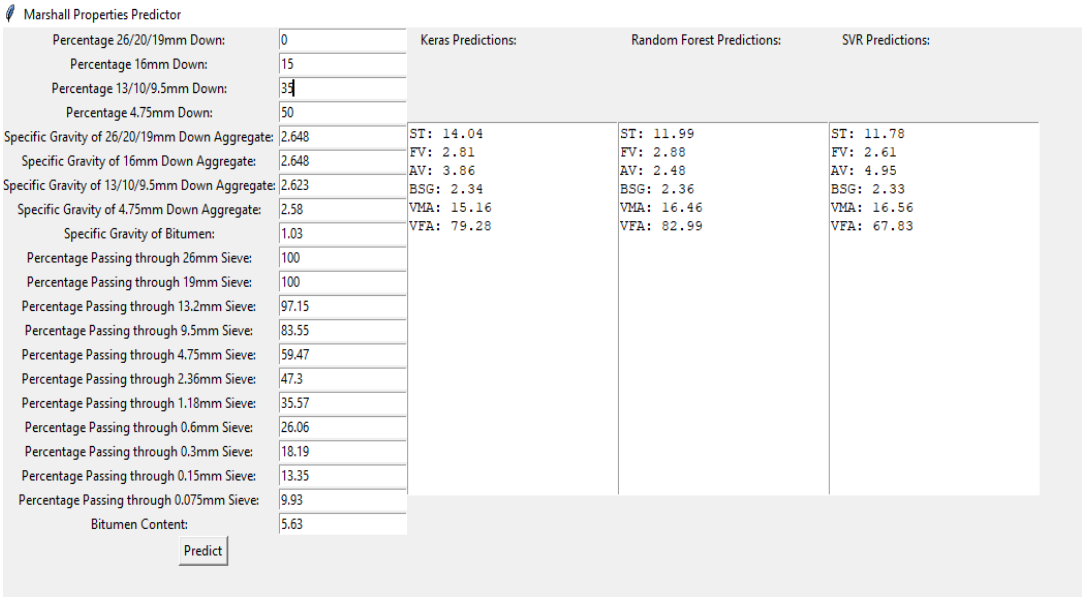


Figure 5-2 Marshall Parameters Prediction Graphic User Interface and Predicted Values for NAS 13.2 mix

Table 5-3 presents the asphalt mix parameters predicted based on the independent variables related to NAS 19mm. The table demonstrates that the Artificial Neural Network (ANN) effectively captures the variability in input parameters. In contrast, the prediction models developed using Support Vector Machine (SVM) and Random Forest (RF) algorithms, collectively referred to as the OBC model, appear to overfit to specific training data. This

characteristic makes them less sensitive to new sets of input data, resulting in similar predicted results, for NAS 13.2 mixes. Specifically, it is suggested that the Deep Neural Network (DNN) model outperforms its SVM and RF counterparts in generating accurate predictions. Figure 5-3 and Figure 5-4 shows the predicted values for OBC and Marshall parameters, respectively.

Table 5-3 Predicted Mechanical and Volumetric Properties of Asphalt Mix of NAS 13.2mm

Marshall Mix Parameters	NMS 19	ANN		SVR		RF	
	Actual value	PV	Percentage Deviation	PV	Predicted Value	PV	Actual value
OBC, % weight	5.43	5.66	4.24%	5.3	2.39%	5.65	4.05%
MST, KN	12.37	11.63	5.98%	11.78	4.77%	11.9	3.80%
MFV, mm	3.44	3.33	3.20%	2.61	24.13%	2.87	16.57%
AV, %	4.00	2.44	39.00%	4.95	23.75%	2.48	38.00%
BSG, g/cc	2.33	2.35	0.86%	2.33	0.00%	2.36	1.29%
VMA, %	16.38	17.59	7.39%	16.56	1.10%	16.43	0.31%
VFA		83.36		67.83		83.05	
Average abso-lute Error			10.11%		9.36%		10.67%

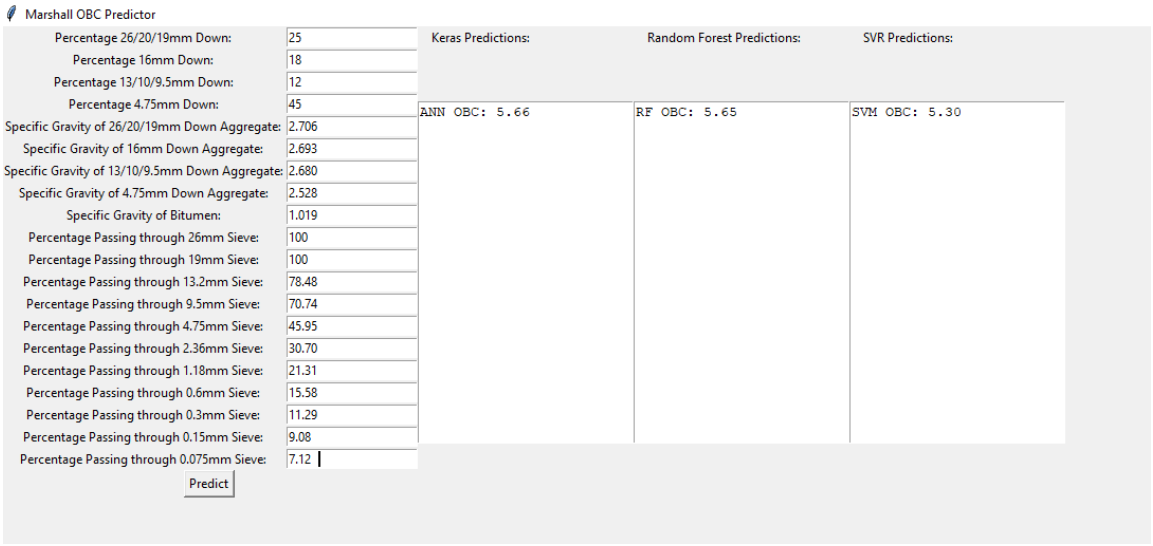


Figure 5-3 OBC Prediction Graphic User Interface and Predicted Values for NAS 19 mix

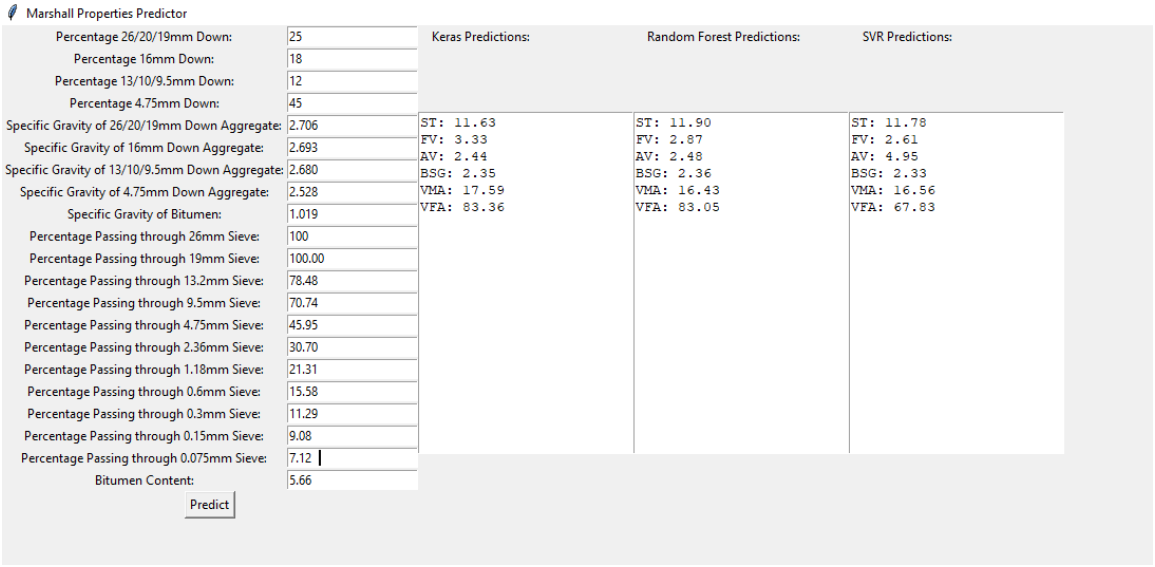


Figure 5-4 Marshall Parameters Prediction Graphic User Interface and Predicted Values for NAS 19 mix

CHAPTER 6: CONCLUSION AND RECOMMENDATION

6.1 Conclusion

Asphalt concrete, a widely used pavement material and maintaining its quality is often challenging due to the nature of Marshall Mix Design methodology. As the conventional Marshall Mix Design method is time-consuming and requires skilled labor for physical testing, perform reliable calculations and draw conclusions. This research is designed to overcome some of the limitations of Marshall Mix design and optimize the resources and time required for mix design. Before developing perdition models, different data processing techniques namely Outlier analysis, compliance and centrality of the gradation is carried out. Since, the research work primarily focuses on the application of Artificial Neural Networks (ANNs) for predicting hot mix asphalt properties, without explicitly mentioning the utilization of Support Vector Machine (SVM), Multiple Linear Regression (MLR), and Random Forest (RF) models.

The study focuses on the prediction of the mechanical (stability, flow), and volumetric properties (VMA, VFA, AV, Bulk Specific Gravity) of the hot asphalt mix from the percentage coarse and fine aggregate, aggregate gradation and specific gravity of material used in bitumen mix using different techniques. Different prediction models are trained and tested to choose the best prediction model that optimizes the prediction accuracy based on results of 141 asphalt mix designs. The main outcomes of this study can be summarized as follows:

i. Descriptive statistics suggested skewness of the dataset and outlier analysis is carried out to remove outliers using z-score method with threshold value of 2 to detect mild outlier that disproportionately affected the mean and standard deviation of the dataset, leading to potential biases in the statistical models

ii. Compliance checks against SSRBW 2016 standards indicated satisfactory performance in stability and flow but highlighted concerns in Marshall quotient and Filler-Binder Ratio.

iii. For Marshall Quotient only 68%,67% and 71% DBM, NMS 19mm, NMS 13.2mm mix satisfies specifications respectively and for Filler-binder ratio only 32%,50% and 32% of mixes meeting the requirement of mix designs. Suggesting need for cross examination of mix properties before approving mix design.

iv. Centrality analysis based on Hypothesis testing revealed significant differences between mid-point gradation and sample mean of DBM, NMS 19mm, NMS 132mm of SSRBW 2016 gradation.

v. There is a significant correlation between Marshall Parameters (OBC, MST, MFV, BSG, AV, VMA, VFA) and the independent variables, which included Proportion of fine and coarse aggregate, specific gravity of different sized aggregate and gradation.

vi. The multiple regression model demonstrated a moderate to strong relationship in the variance of OBC and the independent variables, with R 2 of 0.698 for training stage and R2 of 0.617 for testing stage. MLR model for Mechanical and Volumetric Properties, exhibiting reasonable performance, particularly struggles with the prediction of Stability, as indicated by a relatively low R-Square of 0.28

vii. In developing the ANN model for predicting Marshall parameters, different activation functions were assessed to determine their effectiveness. For the Optimum Binder Content, the tanh activation function exhibited the highest predictive accuracy, with R ^2 value of 0.796 . When modeling Marshall Stability, ReLU proved to be the most effective activation function, achieving R ^2 value of 0.7206. For predicting the Marshall Flow Value, the ReLU function again showed superior performance with R ^2 of 0.8523. The Air Voids parameter was best modeled using ReLU, which attained R ^2 value of 0.877. Similarly, for Bulk Specific Gravity, ReLU was the optimal activation function with R ^2 of 0.8542. In the case of Voids in Mineral Aggregate, the tanh activation function demonstrated the highest R^2, measuring 0.8650. Lastly, for Voids Filled with Asphalt, ReLU once again provided the best predictive performance, with R ^2 of 0.91. Therefore, ReLU generally

emerges as the most effective activation function for most Marshall parameters, though tanh also shows competitive performance in specific instances. Overall statistical calculation summary is shown in Table 4-47

viii. The SVM model demonstrates high predictive accuracy for most parameters, with R^2 values ranging from 0.872 to 0.961 at training and testing stage, the overall summary of the result is shown in the Table 4-49.

ix. The RF model shows high predictive accuracy across all parameters as shown in Table 4-51, with R^2 values ranging from 0.957 to 0.976 and low error metrics in training stage. However, in the testing stage, the performance declines, with R^2 values between 0.605 and 0.844, indicating moderate predictive ability of the Random Forest Models.

x. All ML models indicate strong correlation between predicted and actual values in the training stage. However, during testing, the model's performance varies. While some parameters maintain strong predictive ability ($R^2 > 0.8$), others exhibit lower accuracy ($R^2 < 0.75$), suggesting potential overfitting and limitations in generalization to unseen data.

xi. In comparing the performance of Multiple Linear Regression (MLR) and Artificial Neural Network (ANN) models across various asphalt properties, the ANN consistently demonstrates superior predictive capabilities. Across all Asphalt Mix Properties, the ANN achieves higher coefficients of determination (R-Square) and lower mean squared errors (MSE) on the validation sets compared to MLR.

xii. In comparing the performance of ANN, SVM and RF models, the ANN consistently demonstrates superior predictive capabilities across most Marshall properties with the highest R^2 values and competitive MAE values, making it the best model for predicting OBC, Mechanical and Volumetric properties of the mix. For example, the ANN model for OBC shows the highest R^2 at 0.7964, indicating superior predictive accuracy, while SVM and RF have lower R^2 values at 0.642 and 0.605, respectively.

xiii. When comparing the predictive capabilities of MLR, ANN, SVM, and Random Forest models across various asphalt mix properties, and Random Forest consistently demonstrate superior performance based on the Mean Absolute Error (MAE).

For example, Optimum Binder Content (OBC) RF Random Forest models achieve the lowest MAE of 0.051, indicating the highest accuracy.

xiv. The implementation framework employs ANN, RF, and SVM models sequentially to predict Optimum Bitumen Content and Marshall's properties. ANN proves reliable for most properties, while RF excels in predicting OBC and SVM in predicting stability. However, the OBC model (SVM and RF) tends to overfit, making it less sensitive and responsive to new input data compared to ANN, which demonstrates superior adaptability and accuracy in predicting asphalt mix parameters across diverse datasets.

6.2 Recommendations

To enhance the effectiveness of developed model and optimize its performance following recommendations were made.

i. The use of the ANN can be useful during the phase of the design of the asphalt mix process because of its ability to predict variables with high accuracy. In this case, the ANN can be used for the prediction of the Marshall Parameters instead of the standard Marshall test procedures at the design and estimation stage that require the preparation of fifteen samples. Even during the standard testing, only three samples at 3 bounding BCs are required to prepare and test, to estimate the design parameters and make sure they match the design criteria. This approach saves time, resources, and the required effort to estimate.

ii. Expand research by incorporating additional variables such as aggregate water absorption, flakiness, angularity properties, and bitumen viscosity to boost predictive accuracy.

iii. Explore advanced technologies like ANFIS, MEP, etc., to further improve predictive accuracy in Marshall Mix properties.

iv. Investigate the feasibility of integrating real-time data acquisition systems for continuous optimization of the ML model in predicting Marshall Parameters.

v. Collaborate with industry stakeholders to implement ML models in practical pavement engineering projects, bridging the gap between research and real-world applications.

vi. Conduct long-term monitoring and validation studies to evaluate the reliability and performance of the ML model in actual pavement construction scenarios.

CHAPTER 7: REFERENCES

Abdulhaq Hadi Abedali. (2014). *MS-2 Asphalt Mix Design Methods.* Asphalt Institute.

Akhtar, M., & Moridpour, S. (2021). A review of traffic congestion prediction using artificial intelligence. *Journal of Advanced Transportation, 2021*, 8878011.

Asphalt Institute . (1988). *The Asphalt Handbook, Manual Series No. 4(MS-2).* Lexinton,KY: Asphalt Institute.

Attoh-Okine, N. O. (2001). Grouping Pavement Condition Variables for Performance Modeling Using Self-Organizing Maps. *Computer-Aided Civil and Infrastructure Engineering, 16*, 112–125.

Baldo, N., Manthos, E., Pasetto, M., & others. (2018). Analysis of the mechanical behaviour of asphalt concretes using artificial neural networks. *Advances in Civil Engineering, 2018*.

Balitsaris, M. (2012). Deviations in standard aggregate gradation and its affects on the properties of Portland cement concrete.

Board., T. R. (2007). Practical Approaches to Hot-Mix Asphalt Mix Design and Production Quality Control Testing. *TRANSPORTATION RESEARCH CIRCULAR E-C124*.

Breiman, L. (1996). Bagging predictors. *Machine learning*, *24*, 123–140.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

Chadbourn, B. A., Skok Jr, E. L., Newcomb, D. E., Crow, B. L., & Spindle, S. (1999). The effect of voids in mineral aggregate (VMA) on hot-mix asphalt pavements.

Cheng, C., Liu, J., Yaohui, S., Wang, L., & Wang, X. (2023). Optimizing asphalt mix design using machine learning methods based on RIOCHTrack data.

Coker, A. K. (1995). *Fortran programs for chemical process design, analysis, and simulation.* Elsevier.

Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. *Ensemble machine learning: Methods and applications*, 157–175.

Department of Road. (2016). *Standard Specification for Road and Bridge Works.* Kathmandu: Department of Government of Nepal.

Department of Roads . (2001). *Standard Specification of Road and Bridge Works* . Kathmandu: Department of Roads, Government of Nepal.

DoR. (2016). *Standard Specification for Road and Bridge Work.* Government of Nepal.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. *Advances in neural information processing systems, 9*.

Fang, M., Park, D., Singuranayo, J. L., Chen, H., & Li, Y. (2019). Aggregate gradation theory, design and its impact on asphalt pavement performance: a review. *International Journal of Pavement Engineering, 20*, 1408–1424.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *The journal of machine learning research, 15*, 3133–3181.

Ferreira, J. L., Babadopulos, L. F., Bastos, J. B., & Soares, J. B. (2020). A tool to design rutting resistant asphalt mixes through aggregate gradation selection. *Construction and Building Materials, 236*, 117531.

Hafeez, I. (2009). Impact of hot mix asphalt properties on its permanent deformation behaviour. *Taxila, Pakistan: Department of Civil Engineering Faculty of Civil & Environmental Engineering University of Engineering and Technology Taxila*.

Haykin, S. (1998). *Neural networks: a comprehensive foundation.* Prentice Hall PTR.

Hicks, R. G. (1991). *Moisture damage in asphalt concrete.* Transportation Research Board.

Huang, Y. H. (2004). *Pavement analysis and design* (Vol. 2). Pearson Prentice Hall Upper Saddle River, NJ.

Ivica, A., & Ivan, M. (2017). Development of ANN and MLR models in the prediction process of the hot mix asphalt (HMA) properties. *Can J Civil Eng. https://doi. org/10, 1139*.

Jobson, J. D. (2012). *Applied multivariate data analysis: regression and experimental design.* Springer Science & Business Media.

Kandhal, P. S. (1980). Evaluation of baghouse fines in bituminous paving mixtures. *Association of Asphalt Paving Technologists Proceedings*, (pp. 150-210).

Kandhal, P. S., Chakraborty, S., & others. (1996). *Evaluation of voids in the mineral aggregate for HMA paving mixtures.* Tech. rep., National Center for Asphalt Technology.

Khasawneh, M. A., & Alsheyab, M. A. (2020). Effect of nominal maximum aggregate size and aggregate gradation on the surface frictional properties of hot mix asphalt mixtures. *Construction and Building Materials, 244*, 118355.

Khuntia, S., Das, A. K., Mohanty, M., & Panda, M. (2014). Prediction of Marshall parameters of modified bituminous mixtures using artificial intelligence techniques. *International Journal of Transportation Science and Technology, 3*, 211–227.

Liaw, A., Wiener, M., & others. (2002). Classification and regression by randomForest. *R news, 2*, 18–22.

Maintenance Branch . (2018). *Manual for Dense Graded Bituminous Mixes (DBM/BC).* Department of Roads, Government of Nepal .

Othman, K. (2022). Prediction of the hot asphalt mix properties using deep neural networks. *Beni-Suef University Journal of Basic and Applied Sciences, 11*, 40.

Ozgan, E. (2011). Artificial neural network based modelling of the Marshall Stability of asphalt concrete. *Expert Systems with Applications, 38*, 6025–6030.

Ozturk, H. I., Saglik, A., Demir, B., & Gungor, A. G. (2016). An artificial neural network base prediction model and sensitivity analysis for marshall mix design. *Proceedings of the 6th Eurasphalt & Eurobitume Congress, Prague, Czech Republic.*

Park, B., Cocconcelli, C., & Chun, S. (2022). Gradation characteristics-based interlayer mixture design method for enhanced rutting resistance of asphalt pavements. *Canadian Journal of Civil Engineering, 49*, 607–616.

Park, T. (2007). Causes of bleeding in a hot-in-place asphalt pavement. *Construction and Building Materials, 21*, 2023–2030.

Pouranian, M. R., & Haddock, J. E. (2018). Determination of voids in the mineral aggregate and aggregate skeleton characteristics of asphalt mixtures using a linear-mixture packing model. *Construction and Building Materials, 188*, 292–304.

Putri, E. E., Kasyafi, F. M., & Ahmad, F. (2023). Performance of rubber asphalt in split mastic asphalt mixture. *E3S Web of Conferences, 464*, p. 11009.

Roberts, C. A., & Attoh-Okine, N. O. (1998). A comparative analysis of two artificial neural networks using pavement performance prediction. *Computer-Aided Civil and Infrastructure Engineering, 13*, 339–348.

Roberts, F. L., Kandhal, P. S., Brown, E. R., Lee, D.-Y., & Kennedy, T. W. (1996). Hot mix asphalt materials, mixture design and construction.

Sadiq, R., Rodriguez, M. J., & Mian, H. R. (2019). Empirical models to predict disinfection by-products (DBPs) in drinking water: An updated review.

Shrestha, A., & Mahmood, A. (2019). Review of deep learning algorithms and architectures. *IEEE access, 7*, 53040–53065.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing, 14*, 199–222.

SP:135., I. (2022). Manual for the Design of Hot Bituminous Mixes.

Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural processing letters, 9*, 293–300.

Tapkın, S., Çevik, A., & Uşar, Ü. (2010). Prediction of Marshall test results for polypropylene modified dense bituminous mixtures using neural networks. *Expert Systems with Applications, 37*, 4660–4670.

Tong, J., Ma, T., Shen, K., Zhang, H., & Wu, S. (2022). A criterion of asphalt pavement rutting based on the thermal-visco-elastic-plastic model. *International Journal of Pavement Engineering, 23*, 1134–1144.

Uyanık, G. K., & Güler, N. (2013). A study on multiple linear regression analysis. *Procedia-Social and Behavioral Sciences, 106*, 234–240.

Vale, A. C., do, F. A., & Grecco, F. (2016). Effects of filler/bitumen ratio and bitumen grade on rutting and fatigue characteristics of bituminous mastics. In *Proceedings*

of 6th Eurasphalt & Eurobitume Congress. doi:https://doi.org/10.14311/EE.2016.298

Vapnik, V. N. (1964). A note on one class of perceptrons. *Automat. Rem. Control, 25*, 821–837.

Yang, X., Guan, J., Ding, L., You, Z., Lee, V. C., Hasan, M. R., & Cheng, X. (2021). Research and applications of artificial neural network in pavement engineering: a state-of-the-art review. *Journal of Traffic and Transportation Engineering (English Edition), 8*, 1000–1021.

Zavrtanik, N., Prosen, J., Tušar, M., & Turk, G. (2016). The use of artificial neural networks for modeling air void content in aggregate mixture. *Automation in Construction, 63*, 155–161.

Zhang, H., Chen, Z., Zhu, C., & Wei, C. (2020). An innovative and smart road construction material: Thermochromic asphalt binder. In *New materials in civil engineering* (pp. 691–716). Elsevier.

Zhang, W., Shen, S., Wu, S., Chen, X., Xue, J., & Mohammad, L. N. (2019). Effects of in-place volumetric properties on field rutting and cracking performance of asphalt pavement. American Society of Civil Engineers.

Zhu, J., Ma, T., Fan, J., Fang, Z., Chen, T., & Zhou, Y. (2020). Experimental study of high modulus asphalt mixture containing reclaimed asphalt pavement. *Journal of Cleaner Production, 263*, 121447.

APPENDIX A: DESCRIPTIVE STATISTICS

APPENDIX B: SAMPLE TRAINING TESTING DATASET FOR MLR, ANN, SVM, & RF

APPENDIX C: SAMPLE CODES FOR ANN

APPENDIX D: SAMPLE TRAINING TESTING STAGE VISUILIZATION OF ANN MODELS

APPENDIX E: SAMPLE CODES FOR GUI WITH VISUZILATION