MACHINE LEARNING METHODS TO MAP STABILIZER EFFECTIVENESS BASED ON COMMON SOIL PROPERTIES

by

Amit Gajurel

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in Civil Engineering

Boise State University

December 2018

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

Amit Gajurel

Thesis Title: Machine Learning Methods to Map Stabilizer Effectiveness Based on Common Soil Properties

Date of Final Oral Examination: 26 October 2018

The following individuals read and discussed the thesis submitted by Amit Gajurel, and they evaluated the student's presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Bhaskar Chittoori, Ph.D.	Chair, Supervisory Committee
Partha S. Mukherjee, Ph.D.	Member, Supervisory Committee
Mojtaba Sadegh, Ph.D.	Member, Supervisory Committee

The final reading approval of the thesis was granted by Bhaskar Chittoori, Ph.D., Chair of the Supervisory Committee. The thesis was approved by the Graduate College.

DEDICATION

To my beloved parents who have spent their whole life in efforts to provide me with best possible education and showered me with unconditional love.

To my respected grandparents for teaching me the meaning of wisdom and one's purpose in life.

To my brother for his companionship.

To the love of my life – you have always been someone whom I look up to.

ACKNOWLEDGEMENTS

I would like to heartily thank my advisor Prof. Bhaskar Chittoori for making it possible for me to purse MS in Civil Engineering. I am very grateful for his encouragement and continued support for pursuing the research of my interest. I will forever be thankful for the guidance and support you provided me during my stay at Boise State University. I couldn't have asked for more.

I would like to sincerely thank Prof. Sadegh and Prof. Mukherjee for their unparalleled support and mentorship. My special thanks to Prof. Hastie and Prof. Tibshirani for their wonderful book and video lecture on *An Introduction to Statistical Learning*. I can't imagine completing my thesis work without it. I would like to thank Prof. Marshall for a wonderful class in *Data Analysis and Geostatistics* that encouraged me to pursue machine learning. I am very grateful to Lucas Spaete for introducing me to SSURGO database. I will also like to thank the Montana Department of Transportation (MDT) for financially supporting me through graduate school.

I would like to thank everyone in SuRGE laboratory for their support and companionship. My special thanks goes to Thomas Robbins, Md Touhidul Islam, Austin Berry, Adam Raibley, Aidin Golrokh, Paris Colton, Mariah Fowler, and Binaya Pudasaini. I feel very lucky to have chosen Boise State for my graduate studies. Go Broncos!!!.

ABSTRACT

Unconfined compressive strength (UCS) has been widely used as one of the primary criteria for the selection of optimum type and amount of chemical stabilizer for subgrade/base stabilization. Guidelines established by various state and federal agencies aid in selecting these optimum values by recommending an initial type and amount based on a wide range of soil index properties. A significant number of laboratory trials have to be done to establish the optimum type and amount of stabilizer for a given target strength. This process takes a copious amount of time, money, and the workforce. In addition to that, the finite number of samples brought to the laboratory for characterization of chemical stabilization might not be representative of the problematic area. This study proposes the use of machine learning models to minimize the number of trials and assist in sample collection strategies by spatial mapping of predicted stabilized strength. Supervised machine-learning approaches including regression and classification were used for predicting the quantitative and categorical (pass/fail for a given threshold strength) response respectively. The parameters that didn't have collinearity issues and are available in the Soil Survey Geographic Database (SSURGO) were chosen as input parameters for model development. An existing dataset from Australia was used to study the effectiveness of classification techniques in establishing optimum stabilizer type and amount. This analysis showed that classification methods performed well with a median correct-rate of 0.88 and median True Positive Rate (TPR) of 0.94. After this initial analysis, a database consisting of US soils and the corresponding stabilization data was

compiled. Regression models using this new dataset for US soils showed comparable or better performance than regression models reported by other researchers to predict UCS values with Root Mean Square Error (RMSE) of 0.50 MPa (72.52 psi) for lime treated soils and 0.53MPa (76.87 psi) for cement treated soils. The classification model for the US soils had a median correct-rate of 0.92 and TPR of 0.94 for lime treated soils, while the same for cement treated soils were 0.80 and 0.77. The carefully chosen model input parameters (soil properties from SSURGO) in this study not only assist in arriving at an optimal type and amount of stabilizer but also help visualize the spatial distribution of UCS values for any given area within the US thereby enhancing sample collection strategies.

TABLE OF CONTENTS

DEDICATION iv

ACKNOWLEDGEMENTS v

ABSTRACT vi

TABLE OF CONTENTS viii

LIST OF TABLES xi

LIST OF FIGURES xii

LIST OF ABBREVIATIONS xiv

CHAPTER ONE: INTRODUCTION AND BACKGROUND 1

1.1 Introduction and Research Problem 1

1.2 Research Objectives and Tasks 3

1.3 Organization of Thesis 5

1.4 References 5

CHAPTER TWO: ESTIMATING OPTIMAL ADDITIVE CONTENT FOR SOIL STABILIZATION USING MACHINE LEARNING METHODS 8

2.1 Abstract 8

2.2 Introduction 9

2.3 Statistical Classification Methods 11

2.3.1 Logistic Regression (LR) 11

2.3.2 Discriminant Analysis (DA) 12

2.3.3 K-Nearest Neighbor (KNN) 13

2.3.4 Support Vector Machine (SVM) 14

2.4 Model Assessment Measures 15

2.5 Database Preparation and Model Development 16

2.6 Results 18

2.7 Discussion 21

2.8 References 23

CHAPTER THREE: MACHINE LEARNING METHODS TO MAP STABILIZER EFFECTIVENESS BASED ON COMMON SOIL PROPERTIES 27

3.1 Abstract 27

3.2 Introduction 28

3.3 Literature Review 31

3.3.1 Current Practice for The Selection of Optimal Stabilizer Type and Amount 31

3.3.2 Current Practice in Selecting Sampling Locations for Laboratory Testing 33

3.4 Current Study Approach 37

3.5 Databases Used in This Study 38

3.5.1 Compilation of Database of Stabilized US-Soils 38

3.5.2 Soil Survey Geographic Database (SSURGO) 46

3.6 Model Development 46

3.6.1 Selection of predictors 47

3.6.2 Type of Machine Learning Models 48

3.6.3 Prediction Accuracy of Regression Models 54

3.6.4 Prediction Accuracy for Classification Models 56

3.6.5 Model Selection and Model Assessment 58

3.7 Results and spatial visualization 60
3.7.1 Model Selection 60
3.7.2 Model evaluation 61
3.7.3 Model Performance 72
3.7.4 Comparison of Model Predictions 74
3.8 Acknowledgment 75
3.9 References 75
SUMMARY, CONCLUSION, AND RECOMMENDATIONS 84
Summary and Conclusion 84
Recommendations for Future Research 85
APPENDIX A 87
Database for the US Soil 88
Spatial visualization of strength for Broadwater County 89

LIST OF TABLES

Table 2-1: Summary of predictors and the response considered in the study 17

Table 2-2: Mean and standard deviation of predictors and the response for 5 fold CV 17

Table 2-3: Hyper-parameters for models.....21

Table 3- 1: Highlights of Database for the US soils 41

Table 3- 2: Summary of parameters used on model development.....48

Table 3- 3: Tuning parameters for regression models.....60

Table 3- 4: Tuning parameters for classification models 61

Table 3- 5: Results for regression models 63

Table 3- 6: Results for classification models for lime treatment.....64

Table 3- 7: Results for classification models for cement treatment.....65

LIST OF FIGURES

Figure 1-1: Pictorial representation of research 5

Figure 2-1: Performance of LR 18

Figure 2-2: Performance of LDA 19

Figure 2-3: Performance of QDA.....20

Figure 2-4: Performance of KNN.....20

Figure 2-5: Performance of SVM-Polynomial.....20

Figure 2-6: Performance of SVM using radial kernel 21

Figure 3-1: Systematic pictorial representation machine learning application in UCS prediction 38

Figure 3-2: Location and number of soil samples 40

Figure 3-3: Graphical representation of correlation matrix for lime treated soils 45

Figure 3-4: Graphical representation of correlation matrix for cement treated soils ...45

Figure 3-5: Procedure adopted for resampling using k-fold CV and random sampling 60

Figure 3-6: Summary of test set performance for regression for lime 67

Figure 3-7: Summary of test set performance for regression for cement 68

Figure 3-8: Summary of test set performance for classification for lime 68

Figure 3-9: Summary of test set performance for classification for cement.....69

Figure 3-10: Spatial visualization regression result for UCS (MPa) of cement treated soil 70

Figure 3-11: Spatial visualization classification result for 2.06 MPa UCS of Cement treated soil.....71

Figure 3-12: Actual vs predicted strength for spatial samples.....71Figure A-1: .....Spatial visualization of input parameters for Broadwater County 89

Figure A-2: Spatial UCS visualization of regression results for 2 percent cement treatment 90

Figure A-3: Spatial UCS visualization of regression results for 4 percent cement treatment 91

Figure A-4: Spatial UCS visualization of regression results for 6 percent cement treatment 92

Figure A-5: Spatial UCS visualization of regression results for 8 percent cement treatment 93

Figure A-6: Spatial UCS visualization of classification results for 2 percent cement treatment 94

Figure A-7: Spatial UCS visualization of classification results for 4 percent cement treatment 95

Figure A-8: Spatial UCS visualization of classification results for 6 percent cement treatment 96

Figure A-9: Spatial UCS visualization of classification results for 8 percent cement treatment 97

LIST OF ABBREVIATIONS

DA	Discriminant Analysis
FPR	False Positive Rate
KNN	K – Nearest Neighbors
LDA	Linear Discriminant Analysis
QDA	Quadratic Discriminant Analysis
LR	Logistic Regression
OMC	Optimum Moisture Content
UCS	Unconfined Compressive Strength
SVM	Support Vector Machines
SSURGO	Soil Survey Geographic Database
TPR	True Positive Rate

CHAPTER ONE: INTRODUCTION AND BACKGROUND

1.1 Introduction and Research Problem

Soils along an alignment of road might not always be suitable for pavement construction and avoiding such areas might not always be feasible. In such cases, the problematic soils are often modified to improve their mechanical performance. Chemical stabilization using calcium based stabilizers are among widely used methods for soil stabilization and/or modification of pavement layers (Berry et al. 2007; Cole and Cepco 2006; Nelson and Miller 1997; Petry and Little 2002; Puppala 2016; Thompson 1972). Various state and federal agencies have published standard guidelines to establish an optimum stabilizer type and amount for a given soil. These guidelines outline a series of laboratory tests that need to be performed prior to establishing an optimum stabilizer. Since the initial stabilizer recommendations (by guidelines) are based on a wide range of soil properties, significant number of trials have to be done before reaching the optimum values – requiring substantial amount of time and effort. Even after establishing the optimum stabilizer by arduous laboratory testing, the stabilization design may not be representative of the whole site as a finite number of samples from randomly selected locations are tested in the laboratory. Sampling locations for laboratory tests as well as in-situ testing are established based on historical records, geological reports and publications, soil survey maps, aerial photography etc. (AASTHO 1988; CDOT 2017; IDOT 2015; WSDOT 2010). None of these guidelines use prior stabilization performance information for site delineation. Therefore, the main objective of this study was to

develop a model using machine learning algorithms, as a function of common soil properties, which can provide a preliminary estimate of the treated strength of the soil. The estimate will be very helpful in minimizing the number of laboratory trials. In addition, since the input parameters for the models are common soil properties whose spatial distribution can be obtained from public databases such as the Soil Survey Geographic Database (SSURGO), the models can also help visualize the spatial distribution of UCS values for any given location in the US. Sampling locations can be strategized based on this spatial distribution to make the laboratory tests representative to field conditions.

Several studies have been done to develop machine learning models that predicts the UCS value as function of various soil properties (Alavi et al. 2009; Das et al. 2011; Mozumder et al. 2017; Mozumder and Laskar 2015; Tinoco et al. 2011, 2014, 2016). Although, the models provide good insights into the relationship between selected input parameters (predictors) and the UCS values, these predictors for most of these models are soil properties that are not reported by public data bases which means additional laboratory testing is required to use these models. This dependence limits the application of these models in arriving at an optimum stabilizer (type and amount) that could be later tested. In addition to that, the regression models used in the current studies had Root Mean Squared Error (RMSE) ranging from 0.75 MPa (108.77 psi) to 1.99 MPa (288.62 psi),which is high compared to targeted subgrade strength that generally ranges from 0.34 MPa (50 psi) to 2.06 MPa (300 psi). Moreover, most of the existing models were developed using soils outside the United States which limit the applicability to US soils.

1.2 Research Objectives and Tasks

The hypothesis of this thesis is that a model developed using machine learning algorithms, using a selected set of parameters which are spatially available, can be used to estimate the strength of stabilized soil with reasonable accuracy and without bringing the soils to the laboratory. Such preliminary estimates can be used to reduce the amount of laboratory work and aid in sample collection strategies by developing a spatial map of stabilizers' performance. Research objectives to support the hypothesis are as follows:

1) Study the effectiveness of classification based machine learning algorithms in screening stabilizers as pass/fail at a specified threshold strength
2) Develop classification-based and regression-based machine learning algorithms to predict optimal stabilizers using commonly available soil properties
3) Verify model predictions using independent data sets generated from laboratory testing

Research tasks for accomplishing these objectives are:

1) *Review of the existing literature*: This was done to understand current state-of-practice of machine learning in stabilized soil strength prediction. It was found that most of the studies were done in Europe, Asia, and Australia. The most comprehensive database for stabilized soils was compiled by Burroughs (2001) for Australian soils.
2) *Selection of input parameters*: This was done by considering the collinearity issues and availability of parameters in SSURGO.

3) *Classification-based models*: Developed models using classification based machine learning algorithms to predict pass/fail for a given threshold, using the Australian database as a pilot study to assess the models applicability.
4) *Compiling US-soil data*: Developed a database for chemically stabilized soils in the US. The model developed from the data collected within the US will provide better confidence for real-world applications in the US than a model developed with the Australian database.
5) *Model development and assessment*: Developed models for quantitative and categorical (pass/fail) response using regression and classification machine learning algorithms respectively, using the US database.
6) *Mapping UCS*: Spatial visualization of UCS predictions was done using the best machine learning models.
7) *Verifying model prediction*: Predictions of the best performing model for treatment was compared to an independent set soil strength performed in the laboratory. Strength of treatment on three soils from two different counties in Montana was compared with the model's result.

A pictorial representation of the whole research is presented in Figure 1-1.

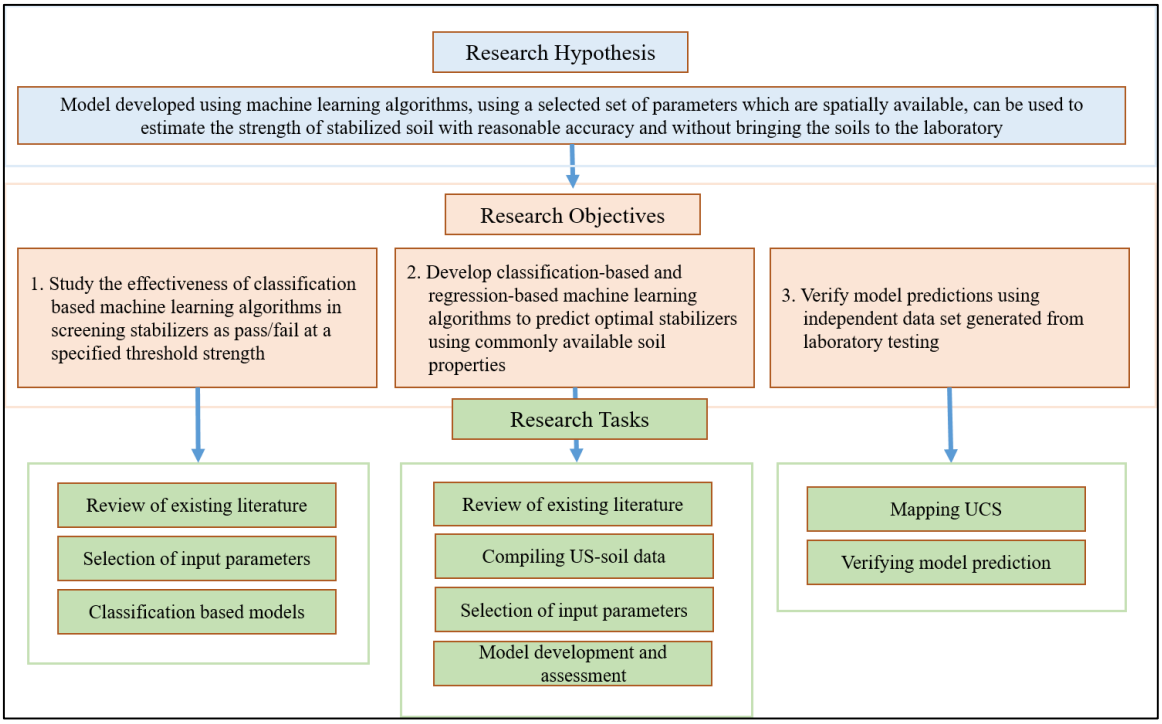

Figure 1-1: Pictorial representation of research

1.3 Organization of Thesis

The thesis starts with an introduction to the research problem, objectives, and tasks in chapter one. Chapter two and chapter three consists of manuscripts which are closely related to each other. The first manuscript in chapter two is on the use of classification machine learning algorithms to develop UCS prediction models using a selected set of parameters for the Australian database. This manuscript used a pilot study and was accepted to Geo-Congress 2019, Philadelphia, Pennsylvania. Chapter three consists of the second manuscript which is the extension of the machine learning algorithm to the new US database. The manuscript will be submitted to Elsevier's *Computers and Geotechnics.*

1.4 References

AASTHO. (1988). *Manual on Subsurface Investigations*. Washington, D.C.

Alavi, A. H., Gandomi, A. H., Gandomi, M., and Sadat Hosseini, S. S. (2009).

"Prediction of maximum dry density and optimum moisture content of stabilised soil using RBF neural networks." *IES Journal Part A: Civil and Structural Engineering*, 2(2), 98–106.

Berry, K. M., Hasan, A., and Cooling, T. L. (2007). "Geotechnical Aspects of the St. Louis-Lambert International Airport Expansions." *Aviation: A World of Growth. The 29th International Air Transport ConferenceAmerican Society of Civil Engineers*.

Burroughs, V. S. (2001). "Quantitative criteria for the selection and stabilisation of soils for rammed earth wall construction." University of New South Wales.

CDOT. (2017). *Geotechnical Design Manual*.

Cole, L. W., and Cepco, C. (2006). "PENNDOT'S Efforts to Improve Pavement Sub-grade During Highway Reconstruction–Airfield and Highway Pavements Specialty Conference." *Airfield and Highway Pavements Specialty Conference*, American Society of Civil Engineers, Atlanta, Georgia.

Das, S. K., Samui, P., and Sabat, A. K. (2011). "Application of Artificial Intelligence to Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilized Soil." *Geotechnical and Geological Engineering*, 29(3), 329–342.

IDOT. (2015). *Geotechnical Manual*.

Mozumder, R. A., and Laskar, A. I. (2015). "Prediction of unconfined compressive strength of geopolymer stabilized clayey soil using Artificial Neural Network." *Computers and Geotechnics*, Elsevier Ltd, 69, 291–300.

Mozumder, R. A., Laskar, A. I., and Hussain, M. (2017). "Empirical approach for strength prediction of geopolymer stabilized clayey soil using support vector

machines.” *Construction and Building Materials*, Elsevier Ltd, 132, 412–424.

Nelson, J. D., and Miller, D. J. (1997). *Expansive soils: Problems and Practice in Foundation and Pavement Engineering*. Wiley, New York, New York.

Petry, T. M., and Little, D. N. (2002). “Review of Stabilization of Clays and Expansive Soils in Pavements and Lightly Loaded Structures—History, Practice, and Future.” *Journal of Materials in Civil Engineering*, 14(6), 447–460.

Puppala, A. J. (2016). “Advances in ground modification with chemical additives: From theory to practice.” *Transportation Geotechnics*, 9, 123–138.

Thompson, M. R. (1972). “Deep-plow lime stabilization for pavement construction.” *Transportation Engineering Journal of ASCE*, ASCE, 98(2), 311–323.

Tinoco, J., António Alberto Santos, C., Da Venda, P., Correia, A. G., and Lemos, L. (2016). “A Data-driven Approach for qu Prediction of Laboratory Soil-cement Mixtures.” *Procedia Engineering*, Elsevier B.V., 143(Ictg), 566–573.

Tinoco, J., Gomes Correia, A., and Cortez, P. (2011). “Application of data mining techniques in the estimation of the uniaxial compressive strength of jet grouting columns over time.” *Construction and Building Materials*, Elsevier Ltd, 25(3), 1257–1262.

Tinoco, J., Gomes Correia, A., and Cortez, P. (2014). “Support vector machines applied to uniaxial compressive strength prediction of jet grouting columns.” *Computers and Geotechnics*, Elsevier Ltd, 55, 132–140.

WSDOT. (2010). *Geotechnical design manual*.

CHAPTER TWO: ESTIMATING OPTIMAL ADDITIVE CONTENT FOR SOIL STABILIZATION USING MACHINE LEARNING METHODS

2.1 Abstract

A majority of geotechnical guidelines for chemical stabilization of subgrade/base materials for pavements use unconfined compressive strength (UCS) in establishing the optimal amount of additive. Laboratory determination of UCS strengths for these stabilized soils involves multiple trials by varying amount of stabilizers to achieve target strength. This process takes copious amounts of time, energy, and workforce. In addition to that, these trials are generally made on few discrete field samples which may not be representative of the overall site. Therefore, this study is aimed towards minimizing the laboratory work along with aiding in improving the sample collection strategies by using machine learning models. For this study, statistical classification was chosen to estimate optimal additive type and content. This method was used to classify whether soil will pass or fail a target strength requirement for a given amount and type of treatment. Logistic Regression (LR), Discriminant Analysis (DA), K- Nearest Neighbors (KNN), and Support Vector Machines (SVM) were used for this purpose. Commonly measured soil properties such as Atterberg limits and gradation (reported in databases such as Soil Survey Geographic Database (SSURGO)) along with treatment amount and type were chosen as predictors and, treated UCS strength as a response. Prediction accuracy was calculated using the Area Under the Curve (AUC), Correct Prediction Rate, True Positive

Rate (TPR), and False Positive Rate (FPR). Optimal model was reported after model development using 5-fold cross-validation.

2.2 Introduction

The stabilization of soft and problematic ground facilitates the construction of civil infrastructures in such challenging situations. Among many stabilization methods, chemical modification is widely used in civil engineering infrastructures due to its established history of performance, a wide range of application, ease of field mixing, and controlled alteration of soil properties to meet the specifications (Chittoori et al. 2011; Nelson and Miller 1992; Puppala 2016). Amidst civil infrastructures, a considerable amount of problematic pavement has resorted to chemical stabilization (Berry et al. 2007; Cole and Cepco 2006; Petry and Little 2002). Pavements are designed to distribute the traffic load to subgrade. Depending on the project requirements and existing soil conditions, modification of subgrade might be necessary to provide adequate support during its construction and design life. Chemical stabilization of subgrade has been practiced by following standard guidelines such as Jones et al. (2012); NLA (2006); PCA (1992); TxDOT (2005); U.S. Army TM 5-882-14/AFM 32-1019 (1994). Unconfined compressive strength (UCS) is the primary criteria for selecting the type and amount of stabilizer for treating problem soils.

In deciding the type and amount of stabilizer to be used in the treatment, most of the standard stabilization guidelines use a range of Atterberg limits and soil gradation for initial selection of type and amount of stabilizer. After the initial selection, Optimum Moisture Content (OMC) and Maximum Dry Unit Weight (MDUW) is established and UCS samples are prepared and tested. The type and/or amount of stabilizer are updated to

meet the required strength criteria. For a single type and amount of stabilizer, UCS tests generally take 7 to 28 days excluding the time for sample preparation. Therefore, multiple sessions of such UCS tests need to be carried out for selecting an optimal type and amount of stabilizer for a given soil. Such numerous repetitions grosses significant amount of time for conducting tests on a single soil. Even after such arduous laboratory testing, stabilizer type and amount can be established for the finite amount of field sample collected at discrete locations. Stabilization strategies based on such discrete results might not be representative of the whole surrounding area. Therefore, machine learning can be used to overcome these challenges by developing models that can predict optimal stabilizer content using readily available soil properties in databases such as SSURGO. These models can then be used to develop a map that shows the locations best suited for this stabilizer and help strategize sampling locations for laboratory testing.

Machine learning, which is a set of tools in statistics for modeling and understanding complex datasets, has been used considerably in geotechnical engineering for quantitative as well as qualitative prediction (Bhattacharya and Solomatine 2006; Chou et al. 2016; Das et al. 2011; Lai et al. 2006; Marjanović et al. 2011; Mozumder and Laskar 2015; Suman et al. 2016; Tinoco et al. 2016). Various supervised learning techniques like Multiple Linear Regression (MLR), Generalized Additive Model (GAM), and Classification and Regression Trees (CART), Support Vector Machines (SVM), Artificial Neural Networks (ANN), Logistic Regression (LR), discriminant analysis, K–Nearest Neighbor (KNN), etc. are being used for solving these real-life problems in geotechnical engineering. Das et al. (2011), Mozumder and Laskar (2015), and Tinoco et al. (2016) have used various machine learning tools to predict the UCS strength of

stabilized soils. Some predictors used in these models include mixture properties which warrant the soil to be brought to the laboratory. This impedes the very purpose of the machine learning models to provide an initial estimate of the UCS strength without having to bring the soil to the laboratory. In this study, authors have chosen predictors that are readily available in databases (such as SSURGO) to avoid preliminary sample collection.

Statistical classification methods were employed to demonstrate the use of machine learning techniques to classify whether a particular amount of additive would pass/fail a target UCS value. In this paper, authors have used the data compiled by Burroughs (2001). Classification methods including Logistic regression, Discriminant analysis, KNN, and SVM were used in this study. The classification was based on strength cutoff of 300 psi – strength limit for treated subgrade to be considered as base as per TxDOT guideline (Veisi et al. 2010a).

The following section discusses the basics of the machine learning algorithms used for classification. Strengths and limitations of the learning methods are described. Performance measurement, as well as resampling techniques used for classification schemes, are briefly discussed.

2.3 Statistical Classification Methods

2.3.1 Logistic Regression (LR)

LR is the simplest way of classification, i.e., using regression method for a qualitative response. Unlike regression, the value of the response variable is the probability that response belongs to a particular class. LR model for ‘p’ distinct predictors is given by Eq. (2-1).

$$p(X) = \Pr(Y = 1|X) = \frac{e^{\beta_0+\beta_1 X_1+\cdots+\beta_p X_p}}{1 + e^{\beta_0+\beta_1 X_1+\cdots+\beta_p X_p}} \quad (2\text{-}1)$$

where, Pr ($Y = 1|X$) = probability that Y=1 (1/0 response) given X; $X_1, X_2, \ldots, X_p$ = predictors; $\beta_1, \beta_2, \ldots, \beta_p$ = regression coefficients; β_0 is the intercept. These coefficients are estimated using maximum likelihood function. For multiclass classification, logistic regression is not preferred as discriminant analysis is widely used for that purpose and use of LR for well separated class can result in unstable parameters (James et al. 2013).

2.3.2 Discriminant Analysis (DA)

Classification using discriminant analysis uses Bayes theorem to calculate the probability (posterior) of given observation 'x' for each 'k' class as given by Eq. (2-2). Distribution of 'x' in each class (prior) and class membership probabilities are used for the calculation of posterior probability. 'x' is classified to that class with the highest probability. Discriminant analysis is superior to logistic regression when the classes are distinctly separated, responses have more than two classes, and 'x' is approximately normal in each class.

$$\Pr(Y = k|X = x) = \frac{\pi_k f_k(x)}{\sum_{l=1}^{k} \pi_l f_l(x)} \quad (2\text{-}2)$$

where, $\Pr(Y = k|X = x)$ = probability of x belonging to class k; $f_k(x)$ = probability distribution function evaluated at x for class k; π_k = class membership probabilities. In linear discriminant analysis (LDA), for a p-dimensional problem, x is assumed to be drawn from $f_k(x)$ i.e., multivariate normal distribution with a class specific mean vector (μ_k) and common covariance matrix ($\sum^{-1}$). Posterior proability of 'x', assuming a Gaussian distribution for prior, is given by the simplied equation in Eq. (2-3) which is linear in 'X'.

In quadratic discriminant analysis (QDA), for a p-dimensional problem, x is assumed to be drawn from $f_k(x)$ i.e. multivariate normal distribution with a class specific mean vector (μ_k) and class specific covariance matrix ($\sum^{-1}{}_k$). Posterior proability of 'x', assuming a Gaussian distribution for prior, is given by the simplied equation in Eq. (2-4) which is quadratic in 'X'.

$$\delta_k(k) = X^T\sum^{-1}\mu_k - \frac{1}{2}\mu_k{}^T\sum^{-1}\mu_k + \log\pi_k \tag{2-3}$$

$$\delta_k(k) = -\frac{1}{2}X^T\sum^{-1}{}_k X + X^T\sum^{-1}{}_k\mu_k - \frac{1}{2}\mu_k{}^T\sum^{-1}{}_k\mu_k + \log\pi_k \tag{2-4}$$

2.3.3 K-Nearest Neighbor (KNN)

KNN is a fully non-parametric method for classification which does not assume the distribution of predictors and shape of decision boundary. Such classification scheme is effective when the Bayes decision boundary is highly non-linear. In this method, the probability of Y belonging to class 'j' given 'x' is calculated using Eq. (2-5). The observation 'x' is then classified into the class 'j' which has the highest probability. K controls the bias-variance tradeoff and is estimated using cross-validation.

$$\Pr(Y = j|X = x) = \frac{1}{K}\sum_{i\epsilon N_0} I(y_i = j) \tag{2-5}$$

where, $\Pr(Y = j|X = x)$ = probability of x belonging to class j; K = number of nearest points to 'x' that is to be considered; N_0 = nearest 'K' number of points to observation 'x'; y_i = response (the class where x_i, the i-th nearest point of x belongs to); I( ) – indicator function.

2.3.4 Support Vector Machine (SVM)

SVM was introduced in computer science by Cortes and Vapnik (1995) which is a generalization of maximum margin classifiers intended for binary classification. SVM classification is based on separating hyperplane – a flat affine subspace of p-1 dimension which separates the two categories. Substituting values of each predictor of a data point in the hyperplane equation gives a signed distance from the hyperplane to that data point. Sign of that output is used to designate the class. The absolute value of the signed distance represents the confidence of classification. The shape of the hyperplane is defined by a small subset of training observation which is known as support vectors. Therefore, the decision rule is not affected by data points far away from the decision boundary, unlike LDA and QDA. SVM is considered as one of the best "out of box" classifiers (James et al. 2013). The simplified representation of SVM classifier is given by Eq. (2-6).

$$f(x) = \beta_0 + \sum_{i \in S} \alpha_i K(x, x_i) \tag{2-6}$$

For polynomial kernel

$$K(x_i, x_i') = \left(1 + \sum_{j=1}^{p} x_{ij} x'_{ij}\right)^d$$

For radial kernel

$$K(x_i, x_i') = e^{\left(-\gamma \sum_{j=1}^{p} (x_{ij} - x'_{ij})^2\right)}$$

where, f(x) = signed distance from hyperplane ; $K(x, x_i)$ = kernel (complexity of hyperplane equation) ; α_i = parameter associated with data point 'i'; S = support vectors., d = degree of polynomial; γ = positive constant

2.4 Model Assessment Measures

K-fold cross-validation was used as the resampling method which involves randomly dividing the dataset into 'k' fold of approximately equal sizes. The first fold was treated as the test set and remaining 'k-1' fold was treated as a training set for fitting the model, and the accuracy of prediction of that model was evaluated using the test set. This process was repeated k times; each time a different fold was used as the test set and accuracy was averaged. K is generally taken as 5 or 10 (James et al. 2013; Kohavi 1995). In this study, 5-folds cross-validation was used.

Prediction accuracy of classification model was accessed using the AUC of Receiving Operating Characteristic (ROC) and correct rate as given by Eq. (2-7). ROC is a plot between True Positive Rate (TPR) and False Positive Rate (FPR) for different threshold values. TPR and FPR were calculated as per Eq. (2-8) and Eq. (2-9) respectively.

$$Correct\ rate = = \frac{1}{n}\sum_{i=1}^{n} I(y_i = \hat{y}_i) \quad (2\text{-}7) \qquad TPR = \frac{TP}{TP + FN} \quad (2\text{-}8)$$

$$FPR = \frac{FP}{FP + TN} \quad (2\text{-}9)$$

where, n = total number of samples; I( ) – indicator function; y_i= actual response; $\hat{y}$ = model's response TP = True positive; FN = False negative; FP = False positive and TN = True negative

AUC of a model can have a value between 0 and 1. A model with AUC of less than 0.5 is not considered useful whereas the model with AUC of 1 is considered as a perfect classifier. Generally, the AUC for a good model lies between 0.5 and 1. The correct rate, whose value lies in between 0 and 1, is the ratio of correctly predicted data to

the total amount of data. A model with a higher correct rate is considered a better model. TPR of the model is the ratio of correctly predicted positive responses to total positive responses which also ranges from 0 to 1. A model with a higher (closer to 1) TPR is considered a better model. FPR of a model is the ratio of incorrectly predicted negative responses to total negative responses which can lie between 0 and 1. A model with a lower value (closer to 0) of FPR is considered a better model.

2.5 Database Preparation and Model Development

The compiled database has information on the liquid limit (LL), plastic limit (PL), plasticity index (PI), linear shrinkage (LS), fines content, sand content, gravel content, the percentage of treatment (cement, lime, and asphalt) and UCS value for 193 samples. The information on soil properties that is common in Burroughs (2001) and SSURGO database was selected as a predictor. Data points with missing predictor values in the Burroughs (2001) database were removed. Collinearity check was used to select predictors that are independent of one another in predicting UCS values. The check was performed using variance inflation factor (VIF) and condition number. Predictors such as plasticity index and % gravel were removed from the model since high correlation is evident with other predictors. The final selected predictors were LL, PL, %Fines, %Sand, % Lime, % Cement and % Asphalt. Response variable "y" was set to 0 for fail and 1 for pass for all models except for SVM. For SVM, response variables were set to -1 for fail and 1 for pass. In addition to that, to remove the effect of scale, input parameters for SVM and KNN were standardized with mean of 0 and standard deviation of 1. The selected classification methods LR, DA, K-NN, and SVM were implemented in the R platform. The LR, DA, and K-NN were performed using a base package of R in RStudio

(R Core Team 2017; RStudio Team 2016). The package developed by Meyer et al. (2017) was used for SVM, Sing et al. (2005) for AUC and ROC determination, and Wickham (2009) for plotting results. 5 fold cross-validation (CV) was used for evaluation of each model assessment parameters. Subsequently, 200 simulations of such 5 fold CV was performed and the results are discussed in the next section.

Summary of minimum, maximum, median and mean values of the selected predictors in the study is given in Table 2-1. Each of these soils was treated with a mixture of lime, cement, and asphalt with concentrations ranging from 0 to 6 percent. Mean and standard deviation (sd) of predictors and responses for train sets and test sets for 5 fold CV is given in Table 2-2.

Table 2-1: Summary of predictors and the response considered in the study

	LL	PL	%Fines	%Sand	UCS (psi)	% Cement	% Lime	%Asphalt
Min.	18	12	5	30	145	0*	0*	0*
Max.	73	18	53	94	783	6	6	3
Median	31	36	25	63	380	4	2	0
Mean	33.01	18.81	25.79	63.11	399	4.14	1.80	0.83

*Treated with a stabilizer other than that specific stabilizer

Table 2-2: Mean and standard deviation of predictors and the response for 5 fold CV

	LL	PL	% Fines	% Sand	UCS (psi)	% Cement	% Lime	% Asphalt
Train -mean	36.59	19.07	25.36	64.98	380.55	4.01	1.92	0.72
Test-mean	36.77	19.08	25.23	64.96	378.82	3.98	1.94	0.71
Train-sd	0.50	00.12	0.42	0.42	2.39	0.06	0.01	0.02
Test-sd	2.20	0.47	1.57	1.70	8.80	0.29	0.10	0.10

2.6 Results

Distribution of AUC, correct rate, TPR, and FPR for the model developed using various classification methods are shown in Figure 2-1 to Figure 2-6. For LR, it can be noted that TPR of the test sets range from 0.62 to 1 with a median value of 0.81 and correct prediction rate of test sets ranged from 0.55 to 0.88 with a median value of 0.75 as shown in Figure 2-1. Performance of LDA and QDA was similar to that of LR. TPR of test sets for LDA ranged from 0.63 to 1 with the median value of 0.9, and correct prediction ranged from 0.62 to 0.88 with a median of 0.75 which is shown in Figure 2-2. As for QDA, TPR of the test sets ranged from 0.55 to 0.98 with a median of 0.78, and correct prediction ranged from 0.55 to 0.88 with a median value of 0.72 which is shown in Figure 2-3.

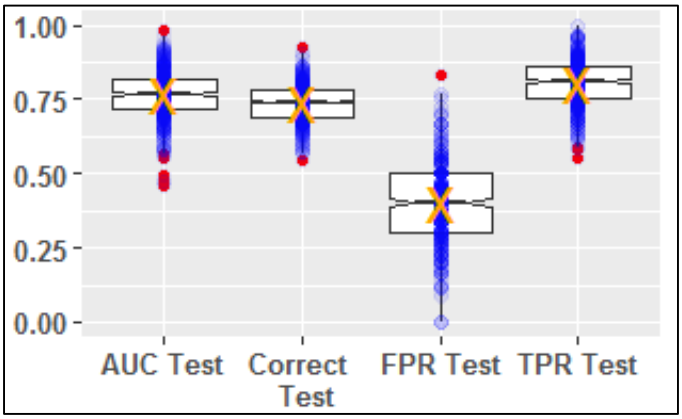

Figure 2-1: Performance of LR

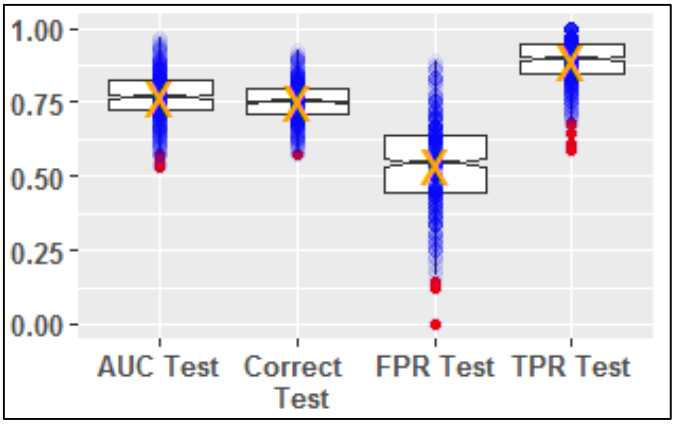

Figure 2-2: Performance of LDA

KNN showed excellent performance in comparison to LR, LDA, and QDA which is shown in Figure 2-4. The number of nearest points to be considered (K) was chosen using 5-fold cross-validation which was found to be 11. TPR for KNN ranged from 0.81 to 1.00 with the median at 0.94, and correct prediction rate ranged from 0.75 to 1 with the median value of 0.88 for the test set. SVM was performed using both radial and polynomial kernel functions. SVM with radial kernel showed better performance than SVM with the polynomial kernel as seen in Figure 2-5 and Figure 2-6. TPR for SVM with polynomial kernel ranged from 0.52 to 0.90 with a median value of 0.70, and correct prediction rate ranges from 0.55 to 0.88 with a median of 0.70. SVM with radial kernel had TPR that ranged from 0.8 to 1 with median value 0.95, and correct prediction rate ranges from 0.63 to 0.88 with a median value of 0.75.

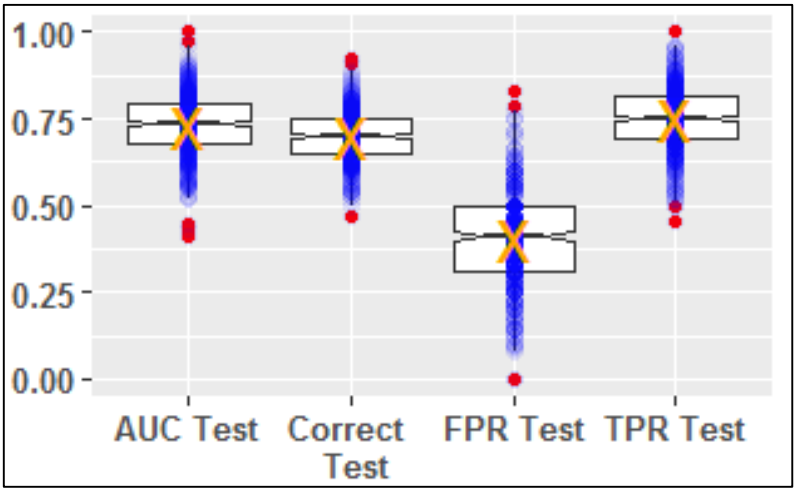

Figure 2-3: Performance of QDA

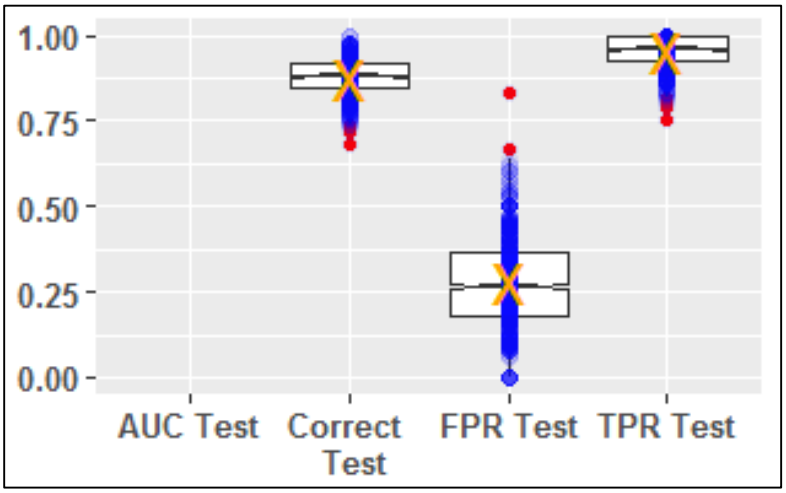

Figure 2-4: Performance of KNN

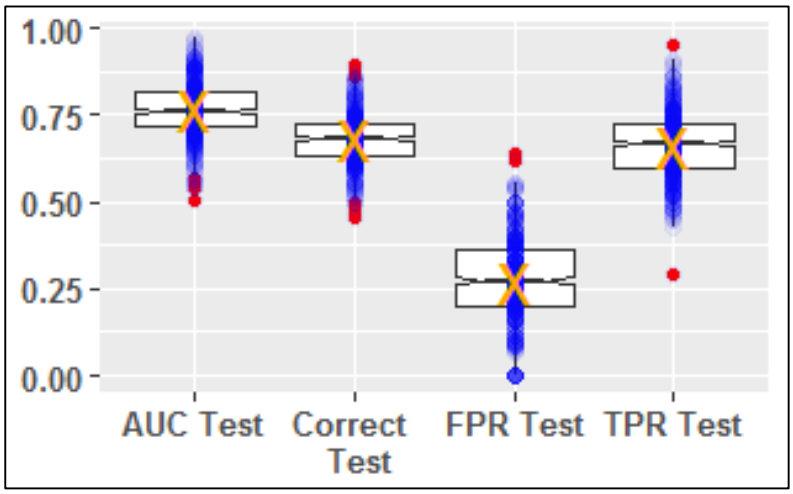

Figure 2-5: Performance of SVM-Polynomial

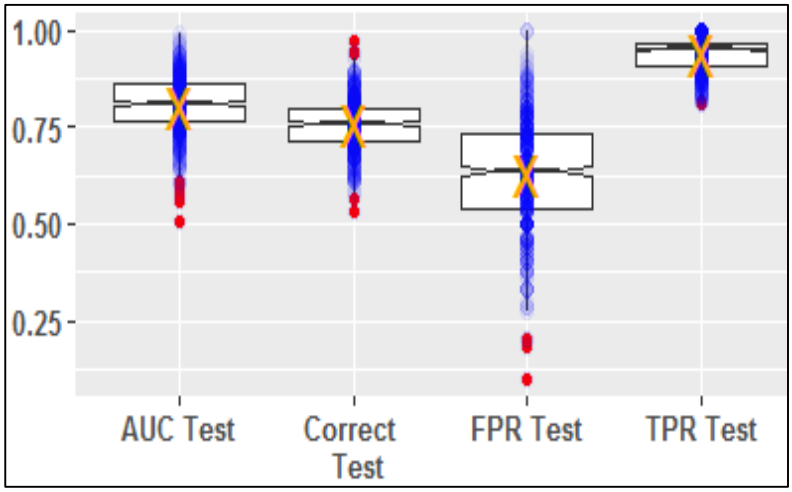

Figure 2-6: Performance of SVM using radial kernel

Details regarding the formulation of the kernel can be found in James et al. (2013). SVM with radial kernel's accuracy lied in between LR and KNN for the test sets. Hyper-parameters calculated using 5 fold CV in different models are given in Table 2-3.

Table 2-3: Hyper-parameters for models

	Hyper- Parameters	
KNN	No. of Nearest Points = 11	
SVM-Poly	Cost = 1	Degree = 2
SVM - Radial	Cost = 10	Gamma = 0.1

2.7 Discussion

The classification scheme was performed with a threshold of 300 psi. Median correct prediction rate for test sets had a comparatively high value which ranged from 0.70 to 0.88. Similarly, median TPR for test set ranged from 0.70 to 0.95. Best performance was observed in case of KNN, followed by SVM – radial kernel which had non-linear decision boundaries. This suggests that decision boundary in case of these parameters is non-linear. Since the median TPR of test sets is 0.70 to 0.95, the "pass" prediction of the classification models can be made with higher confidence for any given

set of parameters. Although, the same cannot be said for the “fail” prediction since the median FPR for test sets ranged from 0.25 to 0.60. One of the reasons for the lower accuracy of “fail” prediction could be the lower number of training data points for the failed case. The inclusion of additional data that does not pass the given threshold to the database can improve the accuracy of “fail” prediction by decreasing the FPR.

The classification models developed by the authors are based on predictors available in web soil surveys of the USDA (2017). These models can be used to select stabilizer type/content based on the decision whether the strength of treatment will reach the threshold value. Such prediction can be made for any given area available in USDA (2017) when type and amount of treatment are given. The authors believe that this would undoubtedly help agencies and practitioners that deal with chemical stabilization of subgrade in prioritizing sampling locations, intensity and get an overview of the effectiveness of chemical stabilization even before any sample collection with the aid of the map that can be developed using this model. In addition to that, selection of optimal stabilizer amount while doing UCS tests in the laboratory using this model can help us reduced the repetitions in testing. In moving further, additional parameters that affect the soil strengths - organic content, cation exchange capacity, specific surface area, etc , can be incorporated in development of a model with improved accuracy. But in doing so, training data for model development shall have all the additional parameter information as well as the parameters used in the current study. Scarcity of such training data with all possible parameters has limited the current study to the presented set of input parameters.

2.8 References

Berry, K. M., Hasan, A., and Cooling, T. L. (2007). "Geotechnical Aspects of the St. Louis-Lambert International Airport Expansions." *Aviation: A World of Growth. The 29th International Air Transport ConferenceAmerican Society of Civil Engineers*.

Bhattacharya, B., and Solomatine, D. P. (2006). "Machine learning in soil classification." *Neural Networks*, 19(2), 186–195.

Burroughs, V. S. (2001). "Quantitative criteria for the selection and stabilisation of soils for rammed earth wall construction." University of New South Wales.

Chittoori, S., Pedarla, A., Puppala, A. J., Hoyos, L. R., Nazarian, S., and Saride, S. (2011). "Leachate studies on lime and portland cement treated expansive clays." *Geo-Frontiers 2011: Advances in Geotechnical Engineering*, 4479–4488.

Chou, J.-S., Yang, K.-H., and Lin, J.-Y. (2016). "Peak Shear Strength of Discrete Fiber-Reinforced Soils Computed by Machine Learning and Metaensemble Methods." *Journal of Computing in Civil Engineering*, 30(6), 4016036.

Cole, L. W., and Cepco, C. (2006). "PENNDOT'S Efforts to Improve Pavement Sub-grade During Highway Reconstruction–Airfield and Highway Pavements Specialty Conference." *Airfield and Highway Pavements Specialty Conference*, American Society of Civil Engineers, Atlanta, Georgia.

Cortes, C., and Vapnik, V. (1995). "Support-vector networks." *Machine learning*, Springer, 20(3), 273–297.

Das, S. K., Samui, P., and Sabat, A. K. (2011). "Application of Artificial Intelligence to Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilized Soil." *Geotechnical and Geological Engineering*, 29(3), 329–342.

Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The elements of statistical learning*. Springer series in statistics New York.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *Introduction to Statistical Learning*. Springer.

Jones, D., Rahim, A., Saadeh, S., and Harvey, J. T. (2012). "Guidelines for the Stabilization of Subgrade Soils In California." *California Department of Transportation*.

Kohavi, R. (1995). "A study of cross-validation and bootstrap for accuracy estimation and model selection." *International Joint Conference on Artificial Intelligence*, Montreal, Canada, 1137–1145.

Lai, S., Chang, W., and Lin, P. (2006). "Logistic Regression Model for Evaluating Soil Liquefaction Probability Using CPT data." *Journal of Geotechnical and Geoenvironmental Engineering*, 132(June), 694–704.

Little, N. (1999). Evaluation of structural properties of lime stabilized soils and aggregates.

Marjanović, M., Kovačević, M., Bajat, B., and Voženílek, V. (2011). "Landslide susceptibility assessment using SVM machine learning algorithm." *Engineering Geology*, 123(3), 225–234.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., and Leisch, F. (2017). "e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071)." TU Wien.

Mozumder, R. A., and Laskar, A. I. (2015). "Prediction of unconfined compressive strength of geopolymer stabilized clayey soil using Artificial Neural Network." *Computers and Geotechnics*, Elsevier Ltd, 69, 291–300.

Mozumder, R. A., Laskar, A. I., and Hussain, M. (2017). "Empirical approach for strength prediction of geopolymer stabilized clayey soil using support vector machines." *Construction and Building Materials*, Elsevier Ltd, 132, 412–424.

Nelson, J. D., and Miller, D. J. (1992). "Expansive soils." *Willy, New York*.

NLA. (2006). "Mixture Design and Testing Procedures for Lime Stabilized Soil." *National Lime Association*.

PCA. (1992). “Soil Cement Laboratory Handbook.” *Portland Cement Association*.

Petry, T. M., and Little, D. N. (2002). “Review of Stabilization of Clays and Expansive Soils in Pavements and Lightly Loaded Structures—History, Practice, and Future.” *Journal of Materials in Civil Engineering*, 14(6), 447–460.

Puppala, A. J. (2016). “Advances in ground modification with chemical additives: From theory to practice.” *Transportation Geotechnics*, 9, 123–138.

R Core Team. (2017). “R: A language and environment for statistical computing. R Foundation for Statistical Computing.” Vienna, Austria.

RStudio Team. (2016). “RStudio: Integrated Development for R.” Rstudio, Inc., Boston,MA.

Samui, P. (2008). “Support vector machine applied to settlement of shallow foundations on cohesionless soils.” *Computers and Geotechnics*, 35, 419–427.

Sing, T., Sander, O., Beerenwinkel, N., and Lengauer, T. (2005). “ROCR: visualizing classifier performance in R.” *Bioinformatics*, 21(20), 7881.

Suman, S., Mahamaya, M., and Das, S. K. (2016). “Prediction of Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilised Soil Using Artificial Intelligence Techniques.” *International Journal of Geosynthetics and Ground Engineering*, Springer International Publishing, 2(2), 11.

Tinoco, J., António Alberto Santos, C., Da Venda, P., Correia, A. G., and Lemos, L. (2016). “A Data-driven Approach for qu Prediction of Laboratory Soil-cement Mixtures.” *Procedia Engineering*, Elsevier B.V., 143(Ictg), 566–573.

TxDOT. (2005). “Guidelines for Modification and Stabilization of Soils and Base for Use in Pavement Structures.” *Texas Deparment of Transportation*.

U.S. Army TM 5-882-14/AFM 32-1019. (1994). Soil Stabilization for Pavements. Technical Manual No. 5-822-14; Air Force Manual No. 32-1019, Joint Departments of the Army and the Air Force, USA.

USDA. (2017). “Web Soil Survey (WSS).” *Natural Resources Conservation Service*.

Veisi, M., Chittoori, B. S., Celaya, M., Nazarian, S., Puppala, A. J., and Solis, C. (2010). Accelerated Stabilization Design of Subgrade Soils. El Paso.

Wickham, H. (2009). *ggplot2: Elegant* Graphics *for Data Analysis*. Springer-Verlag New York.

CHAPTER THREE: MACHINE LEARNING METHODS TO MAP STABILIZER EFFECTIVENESS BASED ON COMMON SOIL PROPERTIES

3.1 Abstract

Most of the chemical stabilization guidelines for subgrade/base use unconfined compressive strength (UCS) of treated soil as the primary acceptance criteria in laboratory testing. Establishing optimal additive content involves a trial-and-error procedure which is resource intensive. Also, samples collected from discrete sample locations for the laboratory trials may not be representative of the overall site. Therefore, this study is aimed towards minimizing the number of laboratory trials and help strategize sampling locations by developing spatial maps of UCS values at different treatment levels for both lime and cement. These spatial maps were developed with the help of machine-learning algorithms that were trained and tested on a database compiled from various reported studies on lime and cement stabilization of soils in the United States. Popular supervised learning methods under regression and classification categories were used to quantify and classify UCS values after treatments, respectively. Commonly available soil properties like Atterberg limits, gradation, and organic contents along with treatment type and amount were used as predictors (input parameters) and UCS value as a response (output). The r-squared value for the regression models ranged from 0.75 for lime to 0.82 for cement, while the True Prediction Rate for the classification models ranged between 0.77 for cement and 0.94 for lime. The results show that good predictions can be made regarding stabilizer effectiveness using simple soil information available in

most databases. Best performing models under each category were selected for generating the spatial maps for two counties in Montana. Soil samples collected from these counties were tested with different lime and cement contents to verify the predictions. The results indicate that the models have 92.31% of prediction accuracy. The authors hope that this study and future studies like these will increase data-driven-decision-making in geotechnical engineering practices.

3.2 Introduction

Excessive urbanization has led to the construction of civil infrastructures on challenging and problematic ground. These problematic soils are often combined with chemical additives to modify and/or enhance their mechanical performance and make them suitable for construction. Soil properties alteration using calcium-based additivities has been widely used in stabilization and/or modification of subgrade due to its established history of performance, a wide range of application, ease of field mixing, and controlled alteration of soil properties to meet the specifications (Chittoori et al. 2011; Nelson and Miller 1997; Puppala 2016). Substantial stabilization work for pavement subgrades was performed using calcium-based additives such as lime and cement (Berry et al. 2007; Cole and Cepco 2006; Petry and Little 2002; Thompson 1972). Standard guidelines for stabilization and/or modification of subgrade and base layers within a pavement have been established by several agencies such as the California Department of Transportation (Jones et al. 2012), the National Lime Association (NLA 2006), the Portland Cement Association (PCA 1992), the Texas Department of Transportation (TxDOT 2005), and the Department of the Army, the Navy, and the Air Force (U.S. Army TM 5-882-14/AFM 32-1019 1994). As per these guidelines, initial estimates of the

type and amount of stabilizers are based on Atterberg limits, soil gradation, organic contents, and soluble sulfates. Soil samples are prepared using the initial estimates and tested for unconfined compressive strengths. This process is iterated for different types and amounts of the stabilizers until a target strength is achieved. For obtaining a single UCS value for a given type and amount of stabilizer, the testing procedure generally takes 7-28 days excluding the time for sample preparation. In order to obtain the optimum type and amount of stabilizer, several repetitions of such tests are needed – requiring a significant amount of time and effort. In addition, since randomly selected sampling locations are put through this process, the optimized stabilizer type and amount may not be representative of the entire site. As a solution to these problems, the current study focused on developing machine learning models that could predict optimal stabilizer amounts using simple soil properties. This predicted amount can then be verified quickly in the laboratory thereby minimizing the amount of time and effort required for establishing optimal stabilizer amounts. Also, these models can be used to generate spatial maps indicating stabilizer effectiveness which can help to strategize sampling locations for laboratory testing.

Machine learning methods have been widely used in recent years in geotechnical engineering and have shown potential for solving real engineering problems (Bhattacharya and Solomatine 2006; Chou et al. 2016; Das 2012; Das et al. 2011; Goh 1994; Goh et al. 2005; Lai et al. 2006; Marjanović et al. 2011; Mozumder et al. 2017; Mozumder and Laskar 2015; Suman et al. 2016; Tinoco et al. 2011, 2014, 2016). Some of the widely used supervised learning methods in geotechnical engineering are support vector machines (SVM), classification and regression trees (CART), and artificial neural

networks (ANN). These methods were used for various applications including soil classifications using cone penetration test (CPT) data, assessing landslide susceptibility, predicting settlement of shallow foundations, predicting peak shear strength of fiber reinforced soil, and predicting undrained side resistance for drilled shafts.

In this study, a database consisting of UCS values corresponding to treatment along with the Atterberg limits, particle size distribution, and organic content of the soil was compiled from various research studies. This database was used to train and test different supervised machine learning models. Different models under both regression and classification categories were studied. The regression models studied were multiple linear regression (MLR), generalized additive models (GAM), and K-nearest neighbors (KNN), and SVM, while the classification models studied were logistic regression (LR), discriminant analysis, KNN, and SVM. The regression models generate quantitative predictions (actual value of UCS) while the classification models generate categorical predictions (pass/fail for a given threshold UCS value). Model performance was assessed using performance measures such as coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) for regression models, and true positive rate (TPR), false positive rate (FPR), correct prediction rate (CPR), Cohen's Kappa (K), area under the curve (AUC) for classification models. Both quantitative and categorical model responses were used to generate spatial distributions of UCS values and pass/fail potential at a stated threshold, respectively, for different types and amounts of treatments. Such distribution facilitates the development of scientific sampling strategies, minimize laboratory testing, and give an overview of suitability of a type of stabilization without having to collect samples in the preliminary stage.

Moving further in the paper, literature review on the current state-of-practice for selection of optimal stabilizer type and amount, delineation of sampling locations, and use of machine learning in geotechnical engineering is discussed. The compilation of the US-soil database used for model development, various training algorithms for regression and classification categories along with their performance metrics, model selection, and model evaluation are briefly described. Additionally, the results for the model selection and evaluation were discussed. The quantitative and categorical response of the best performing models is compared with the laboratory data for three soils in Montana with varying stabilizer contents. An example spatial map of stabilization performance is plotted for Broadwater County in Montana. Finally, the concluding remarks and recommendations for future research are provided in the last section.

3.3 Literature Review

3.3.1 Current Practice for The Selection of Optimal Stabilizer Type and Amount

Lime, cement, and fly ash, are the most common chemical stabilizers in practice (Puppala 2016; Tastan et al. 2011). When these stabilizers are mixed with soil in the presence of water, several processes like cation exchange, flocculation & agglomeration, cementitious hydration, and pozzolanic reactions occur (Chittoori and Puppala 2011; Mitchell and Soga 2005; Prusinski and Bhattacharja 1999; Thompson 1967). These processes result in the reduction of plasticity and increase in workability and compressive strength. Performance of these stabilized mixes is contingent on several factors such as plasticity of natural soil, types and amount of stabilizer, mixing and compaction methods, curing conditions, organic matter, gradation and pulverization, clay fraction, mineralogy and presence of soluble sulfates (Bhattacharya and Solomatine 2006; Croft 1967;

Hampton and Edil 1998; Pedarla et al. 2011; Prusinski and Bhattacharja 1999; Thompson 1967; Tremblay et al. 2002). Considering these factors, several federal and state agencies have developed guidelines for improving engineering properties of soils (Jones et al. 2012; Little and Nair 2009; NLA 2004; PCA 1992; TxDOT 2005; U.S. Army TM 5-882-14/AFM 32-1019 1994). The majority of these guidelines specify the initial selection of stabilizer type based on the plasticity index (PI). For instance, TxDOT (2005) recommends using cement as the first choice for soil with PI < 15, whereas lime for soils with PI >=15. Similarly, CALTRANS (Jones et al. 2012) recommends cement for soil with PI<15, cement or lime for 15 >= PI <35, and lime for PI >= 35. The initial amount of stabilizer is determined based on soil classification (in case of cement) or the Eades and Grim test (1966) (U.S. Army TM 5-882-14/AFM 32-1019 1994) (in case of lime). For example, U.S. Army TM 5-882-14/AFM 32-1019 (1994) recommends initial cement content from 5% to 11% depending on the USCS classification of soil. A similar process is followed by other agencies like PCA (1992), NCHRP (Little and Nair 2009), and FHWA (1992). After the initial selection, a Proctor test is conducted to determine Maximum Dry Unit Weight (MDUW) and Optimum Moisture Content (OMC) for the selected type and amount of stabilizer. Samples were prepared at the corresponding MDUW and OMC to determine UCS. If the test UCS values exceeds a preset threshold value, durability studies are conducted for this combination of treatment. Once both UCS and durability criteria are satisfied, the selected type and amount of stabilizer is termed as optimum. The process is repeated until the optimum values are determined for a given soil sample. Strength criterion in most of the guidelines is established based on UCS values hence, UCS has been used by many researchers as a measure of stabilization

performance (Chittoori et al. 2013; Davidson et al. 1962; Ladd et al. 1960; Little 1999; Pedarla et al. 2011; Sariosseiri and Muhunthan 2009; Thompson 1967; Veisi et al. 2010b; Wen et al. 2014).

3.3.2 Current Practice in Selecting Sampling Locations for Laboratory Testing

Selecting the right locations for in-situ testing and sampling is of utmost importance for any geotechnical investigation. The samples brought to the lab dictate many of the decisions made regarding the superstructure. For example, AASTHO (1988) reported that, for roadway and airfield pavements, subgrade conditions are critical as weak subgrades warrant thickened pavement section, removal, and replacement of poor subgrade, or soil stabilization/improvement. Therefore, the identification of site sampling locations must be performed very carefully. These locations are typically selected by doing a preliminary desk study/literature search where useful information that may be present at a site is gathered. Several studies have revealed that the common source of information for desk studies are a) historical records from prior site investigations, b) performance records from nearby facilities, c) geologic reports and publications, d) geologic maps (bedrock and surficial geology maps), soil survey maps, topographic and LiDAR maps, utility maps, insurance maps, etc., e) aerial photographs, satellite/remote sensing imagery, and f) review of nearby or adjacent projects (AASTHO 1988; CDOT 2017; IDOT 2015; Leohr et al. 2016; WSDOT 2010). The majority of geologic maps are published by the United States Geological Survey (USGS) which has information on depth of rock, the location of rock outcrop, engineering properties of various soil types, and geologic history and groundwater. Soil survey maps, which are compiled by the efforts of the United States Department of Agriculture (USDA), has data on

physical/chemical properties and suitability/limitation for use - at each soil parcels of the mapped area. AASTHO (1988) reported that the soil survey data are the most suitable in the preliminary planning of any geotechnical project. Furthermore, TxDOT's test procedure for surveying and sampling soil for highways establishes the scope of the soil survey to determine the extent and location of each type of soil, location of suitable fill, base and aggregate material, and evaluate the need for stabilization of sub-grade, sub-base, and base material (Tex-100-E 1999).

However, none of the current studies have included prior stabilization performance knowledge of various soils for reconnaissance of project area and delineation of soil sampling sites. Therefore, authors believe that a spatial map of stabilization performance, along with currently available spatial information, will aid in the preliminary investigation process for projects considering stabilization of problematic soils.

3.3.2 Machine Learning in Geotechnical Engineering

The complexity of environmental phenomena along with far from a complete understanding of the underlying process in geotechnical engineering has resulted in favor of statistically derived empirical and semi-empirical relations in contrast to constitutive models (Das 2012). Use of experimental data to derive and/or update correlation is widely practiced since traditional geotechnical engineering started – which is the core idea in machine learning. For instance, the studies by Ching and Phoon (2014a; b) presented an excellent example of how the existing correlations between measurement and design parameters can be derived or updated by using a global database. Machine learning is a set of tools for modeling and understanding complex datasets which have

been extensively used in geotechnical engineering. For instance, the study by Goh (1994) assessed the possibility of using artificial neural networks (ANN) to predict liquefaction potential by modeling its complex relationship with soil parameters such as standard penetration test (SPT) values, mean grain size, equivalent dynamic shear stress, total and effective stress along with earthquake magnitude, and maximum horizontal acceleration at ground surface. Furthermore, Goh (1994) argued that prediction of ANN is more reliable than conventional methods. In addition to that, Lai et al. (2006) compiled a database of liquefied and non-liquefied soil after several earthquakes and demonstrated the use of logistic regression for evaluating the liquefaction potential using only cone penetration test (CPT) data. Pal (2006) showed the similar performance of SVM in predicting liquefaction potential. Goh and Goh (2007) did a similar study on 226 field records for liquefaction potential assessments using CPT data using SVM and reported a classification success rate of 98%. The research study by Bhattacharya and Solomatine (2006) for automated classification of soil using CPT data also found that the predictive accuracy of decision trees (DT), ANN, and SVM was high (83%). In a study on landslide susceptibility assessment based on various geological, morphological, and environmental parameters, Marjanović et al. (2011) showed that SVM outperformed ANN, decision trees as well as assessments made by experts. Samui (2008) studied the application of SVM in settlement prediction of shallow foundation on cohesionless soil and argued its superiority against the existing empirical methods. The above finding is consistent with the study by Chou et al. (2016) in which accuracy of various machine learning and meta-ensemble techniques was compared with theoretical and empirical models for predicting

the shear strength of fiber-reinforced soil. The results showed that the machine learning models outperformed the theoretical and empirical models.

Limited studies were done in developing statistical models that relate stabilization parameters such as UCS, MDUW, and OMC with soil and additive properties. Alavi et al. (2009) used ANN to predict the values of MDUW and OMC of soil-stabilizer mix based on liquid limit (LL), plastic limit (PL), linear shrinkage (LS), sand, fines, lime content, cement content, and asphalt content by training on the dataset compiled by Burroughs (2001). The reported result showed a MAE of 0.38 kN/m^3 (2.42 pcf) for MDUW and 0.717% for OMC. A similar study to predict the UCS value and MDUW of cement stabilized soil was done by Das et al. (2011) by using different classes of ANN and SVM on dataset compiled by Burroughs (2001). SVM models performed better than ANN with an RMSE of 1.26 MPa (182.74 psi) for UCS value and 0.80 kN/m^3 (5.10 pcf) for MDUW. Furthermore, Mozumder and Laskar (2015) reported the efficacy of ANN over MLR in predicting the UCS value of geopolymer stabilized clayey soil based on the predictors such as LL, plasticity index (PI), % stabilizers, molar concentration of alkali activator, ratio of alkali to binder, ratio of Na/Al, and ratio of Si/Al. In another study, Mozumder et al. (2017) demonstrated the performance of SVM in predicting UCS value using the same set of parameters and data as Mozumder and Laskar (2015) that reported RMSE of 0.75 MPa (108.77 psi) and MAPE of 4.5. In similar studies by Tinoco et al. (2011, 2014), the applicability of various machine learning models in prediction of UCS value for jet grouting material was presented. Among multiple learners, SVM's prediction was superior with RMSE of 1.99 MPa (288.62 psi) and R^2 value of 0.51.

Although the machine learning models discussed above give a good insight into the sensitivity of the parameters to the UCS values, these models have some limitations. For example, the models proposed by Das et al. (2011), requires the use of parameters such as OMC along with other soil properties as inputs to the model. Estimating OMC requires laboratory testing which means field sampling is a requirement to use this model. On the other hand, Mozumder et al. (2017) and Mozumder and Laskar (2015), use parameters such as sodium to aluminum ratio among other parameters which are specific to alkali activation studies. So far, the only literature that highlighted this issue was Tinoco et al. (2016). However, in their model, they used % of clay and organic content as the only two input soil parameters for capturing the effect of soil in the soil-cement mix. Also,in most of these models the RMSE reported in the literature ranges from 0.75 MPa (108.77 psi) to 1.99 MPa (288.62 psi). This value is comparatively high while dealing with subgrade modification/stabilization in which treated strength generally ranges from 0.34 MPa (50 psi) to 2.06 MPa (300 psi) (Jones et al. 2012; U.S. Army TM 5-882-14/AFM 32-1019 1994; Veisi et al. 2010b).

3.4 Current Study Approach

To address the issues highlighted in the literature review section of this paper, this study used soil properties in publicly available databases (such as soil survey geographic database - SSURGO) as input parameters to predict UCS values and generate a spatial distribution of UCS values at various treatment levels. First, a database comprising of soil properties, stabilizer type and amount and the corresponding UCS values were compiled from the published literature. This database was then used to train and test various models developed using machine learning algorithms. The best performing models were

used to generate spatial stabilization performance for various treatments. This data was verified by conducting laboratory tests on samples collected from locations within the spatially mapped data. A systematic pictorial representation of this approach is summarized in Figure 3-1. Each of the following sections are thoroughly discussed in the paper.

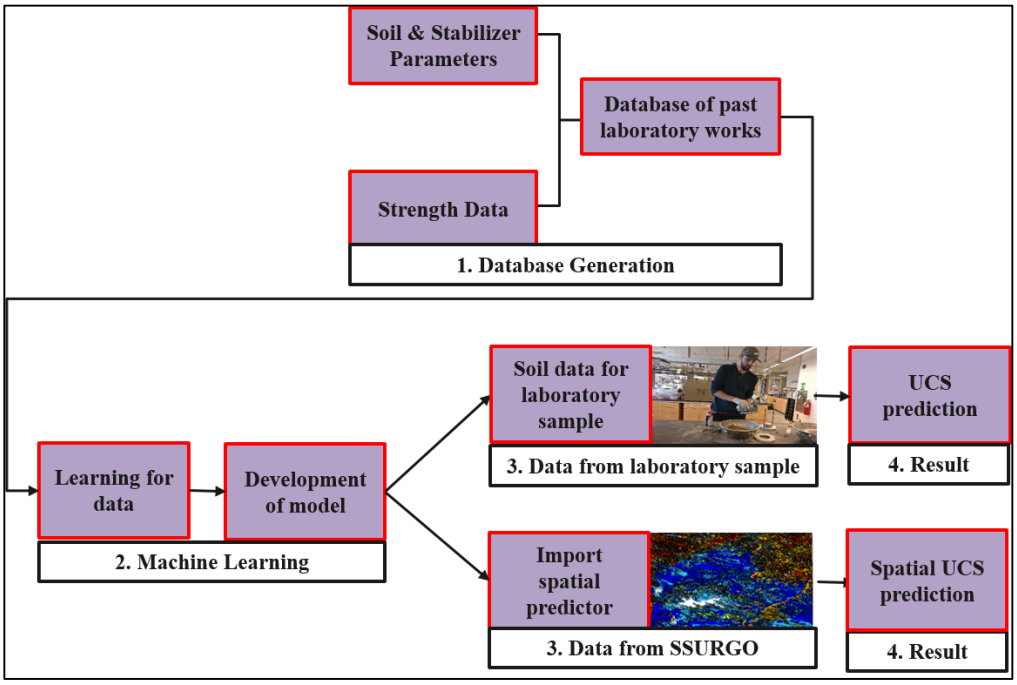

Figure 3-1: Systematic pictorial representation machine learning application in UCS prediction

3.5 Databases Used in This Study

3.5.1 Compilation of Database of Stabilized US-Soils

The scarcity of existing database for physical and chemical properties and strength of chemically stabilized US soils was evident in the literature review process. Therefore, one of the primary objectives of this study was to create a database of various properties of chemically stabilized soils' in the US. As shown in Figure 3-1, this is the first step in the application of machine learning. Authors collected data from digitally

available journals, thesis and dissertations, and technical reports (Davidson et al. 1960, 1962; Davidson and Bruns 1960; Pietsch and Davidson 1961; Remus and Davidson 1961;Whitehurst 1955). The major challenge faced in dataset generation was finding a consistent a) set of reported soil parameters, b) compaction energy for sample preparation, c) standard curing time and procedure, d) sample aspect ratio. In considering all these constraints, the majority of the data in the database was from the laboratory works at Iowa State University in the 1960's. All the samples in the database are cylindrical samples with height and diameter of 50.8 mm (2 in.) and are compacted with standard proctor energy. The attributes in the collected data includes name and location of soil, unified soil classification system (USCS), organic content, % of sand, % silt, % clay, % gravel, LL, plasticity index (PI), % stabilizer (i.e. lime or cement), UCS values (7 days soaked/un-soaked and 28 days soaked/un-soaked). The summary statistics of each of data source is presented in Table 3-1. Locations and number of different soil samples collected within the US is presented in Figure 3-2. Graphical representation of co-relation matrix for input parameter for lime and cement treatment are shown in Figure 3-3 and Figure 3-4 respectively.

Figure 3-2: Location and number of soil samples

Table 3-1: Highlights of Database for the US soils

References	No. of Soils	USCS of each soil type	No. of Data Points	Location	% Sand	% Silt	% Clay
Davidson et al. (1962)	4	SP-SM, CH -3	12	Sand collected from east central Iowa, clays collected from north central Florida, Illinois, and Iowa	0-94-30	2-19-12	6-100-71
Pietsch and Davidson (1961)	20	CH-4, CL-8, OL - 4, OH-1, ML-2, SC-1	101	Soils were collected from all over Iowa considering the parent material.	0-41-15	15-81-48	13-69-35
Davidson and Bruns (1960)	3	SM,ML,CL	15	Collected from Iowa (Typical widely spread material for stabilized road construction)	0-71-34	22-80-44	7-39-22
Pietsch and Davidson (1961)	5	ML,CH-2, CL-2	19	Collected from Iowa - South eastern, northwestern, deep loess bordering the Missouri river.	1-37-18	15-81-40	18-63-42
Remus and Davidson (1961)	9	CL-4, CH-3, ML, SM	45	Collected from all over US (IA, TX, MI, IL VA, NC)	0-45-16	18-61-36	7-75-45

References	No. of Soils	USCS of each soil type	No. of Data Points	Location	% Sand	% Silt	% Clay
Whitehurst (1955)	4	GC, GM, SC,SM	28	Gravels and cherts collected from Benton county and western Tennessee	15-65-33	5-15-9	0-15-10
Davidson et al. (1960)	3	CL, CH-2	21	Typical subgrade found in southern Iowa	0-32-15	14-61-37	39-71-48

Note: e.g. 2-19-12 = Minimum – Maximum – Mean

References	LL	PI	Organic Content	% Stabilizer	UCS – 7-day Soaked (MPa)	UCS – 7-day Un-soaked (MPa)	UCS – 28-day Soaked (MPa)	UCS – 28-day Un-soaked (MPa)
Davidson et al. (1962)	19-87-53	0-53-26	0.00-1.50-0.50	Cement 8-16-12	1.10-7.14-3.06	1.72-7.72-3.69	N/A	N/A
Pietsch and Davidson (1961)	21-76-44	4-21-54	0.00-5.00-1.00	Lime 0-12-5	0 - 2.14 - 0.58	0.23-2.36-0.90	0 - 4.21-0.94	0.26-5-1.30
Davidson and Bruns (1960)	19-42-31	3-22-10	0.16-0.17-0.17	Cement: 0-10-5.6	0-4.56-1.96	-	0-7.22-2.65	-
Pietsch and Davidson (1961)	33-76-48	5-50-28	0.10-0.50-0.25	Cement: 0-10-5	0-3.28-1.37	-	0-4.21-1.81	-
Remus and Davidson (1961)	36-65-48	0-47-24	0.02-2.62-0.67	Lime: 4-12-8	0.37-2.33-1.13	-	0.41-3.76-1.94	-
Whitehurst (1955)	20-32-25	0-13-5	0.00-0.00-0.00	Lime : 1-4-2	0.08-0.47-0.26	-	-	-

References	LL	PI	Organic Content	% Stabilizer	UCS – 7-day Soaked (MPa)	UCS – 7-day Un-soaked (MPa)	UCS – 28-day Soaked (MPa)	UCS – 28-day Un-soaked (MPa)
				Cement: 1-4-2	0.4-2.31-1.21	-	-	-
Davidson et al. (1960)	42-76-56	22-50-34	0.17-0.20-0.19	Lime 0-14-7	0-1.32-0.77	-	0-2.28-1.25	N-

Note: e.g. 2-19-12 = Minimum – Maximum – Mean

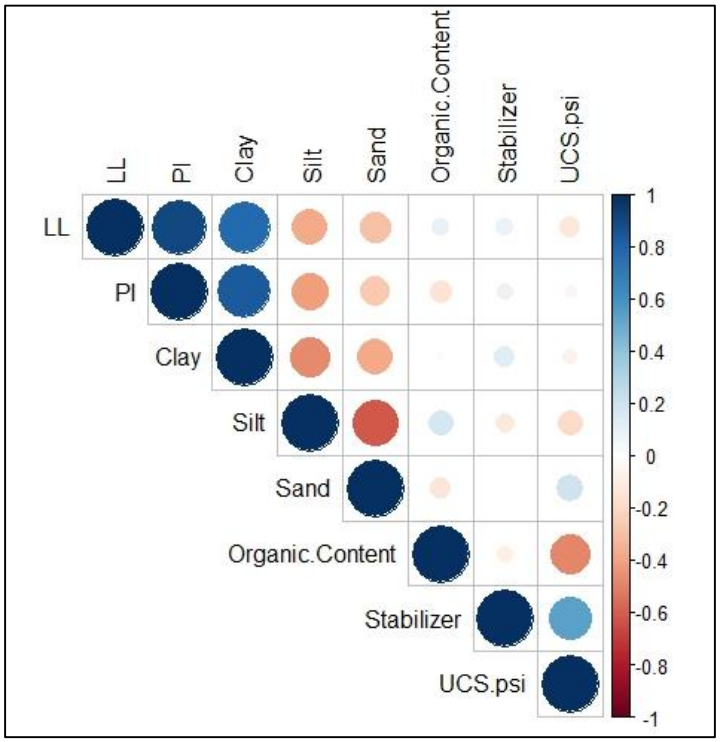

Figure 3-3: Graphical representation of correlation matrix for lime treated soils

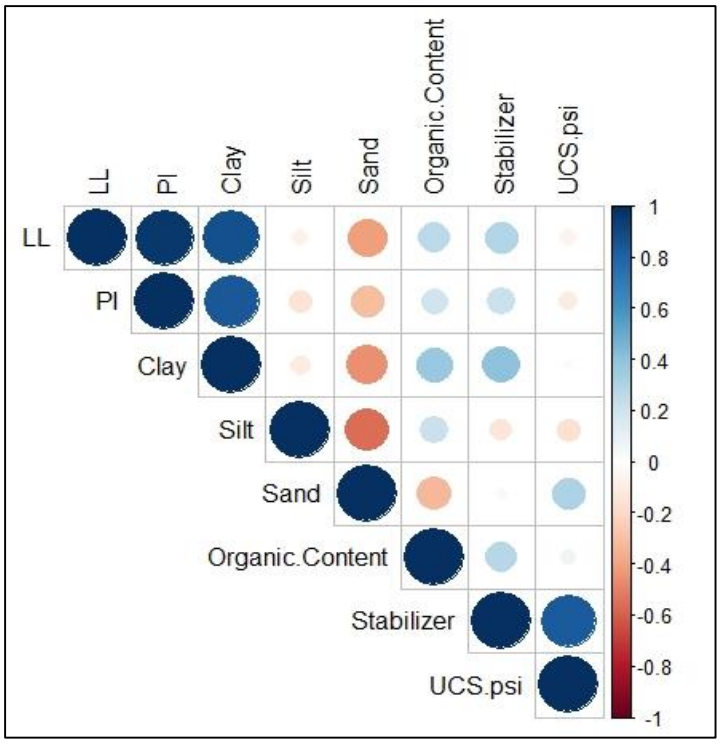

Figure 3-4: Graphical representation of correlation matrix for cement treated soils

3.5.2 Soil Survey Geographic Database (SSURGO)

Most of the geotechnical manuals use the soil survey data by the United States Department of Agriculture (USDA) in desk study phase for planning the preliminary in-situ and laboratory studies and is also considered the third most useful available information for highway design. (AASTHO 1988; CDOT 2017; IDOT 2015; Leohr et al. 2016; WSDOT 2010). SSURGO is a database of the digitally produced version of soil survey data over the course of 100 years in the US. Some examples of available data in SSURGO are a) chemical soil properties – cation exchange capacities, gypsum, etc., b) soil health properties – available water capacity, organic matter, etc., c) soil physical properties – water capacity, bulk density, Atterberg limits, gradation, etc., and d) soil qualities and features – USCS classification, parent material, etc.

3.6 Model Development

After the data compilation in first stage, the second stage was to develop a predictive model using machine learning as shown in Figure 3-1. Preprocessing of the data was done to select appropriate parameters and remove the data with missing parameters before feeding the data for model development. Several parametric and non-parametric machine learning algorithms were implemented on the compiled database for regression and classification model development. The primary intent of the regression model is to predict the UCS of a treated soil characterized by its Atterberg limits, gradation, organic content, and type and amount of stabilizer. The primary goal of a classification model is to predict whether a soil, characterized by Atterberg limits, gradation, organic content, and type and amount of stabilizer meets a certain threshold of

strength. The model selection and assessment were in done in R and scripting was done using RStudio (R Core Team 2017; RStudio Team 2016).

3.6.1 Selection of predictors

Literature reviews of machine learning in stabilized soil's UCS prediction has indicated that the primary focus of most of those studies leaned towards making inferences about the parameter – i.e. understanding relationship in the data. The focus of this study is more inclined towards predictive modeling – i.e., focus on accurate prediction as much as possible. Therefore, for selecting the predictors, the authors considered a) availability of the parameter in compiled database – models are developed based on data of these parameters, b) availability of spatial distribution of parameter in SSURGO –spatial prediction of UCS is only possible if the spatially distributed parameters are available, c) parameters' effect on stabilized soil strength from past studies (Das et al. 2011; Prusinski and Bhattacharja 1999; Thompson 1967; Tinoco et al. 2011, 2014, 2016), and d) multi-collinearity – ensure independence of each parameter with one-another. Considering all these conditions, the soil parameter that qualified were LL, PI, % sand, % silt, % clay, organic content, and % stabilizer for UCS prediction. Separate models were developed for lime and cement treatment. For this particular study, 7-day soaked strength for cement treatment, and 28-day soaked strength for lime was selected as UCS value for the development of the model. Similar curing protocols were reported by agencies (PCA 1992; U.S. Army TM 5-882-14/AFM 32-1019 1994). The database has 167 complete training examples for lime and 60 complete training examples for cement treated soils. Interested readers can further utilize the generated database and

suggested algorithms by the authors to develop models using UCS values for different curing protocols. Model parameters used in this study are summarized in Table 3-2.

Table 3-2: Summary of parameters used on model development

		Predictors							Response
		LL (%)	PI (%)	Clay (%)	Silt (%)	Sand (%)	Organic Content (%)	Stabilizer (%)	UCS (MPa)
Model for lime treatment	Min	21	0	7	14	0	0.00	0.00	0.00
	Q1	38	17	30	30	1	0.00	2.00	0.27
	Median	44	21	39	39	10	0.20	6.00	1.28
	Mean	47	23	40	43	15	1.06	6.20	1.25
	Q3	52	30	46	61	30	1.62	10.00	1.89
	Max	76	54	75	81	45	4.77	14.00	4.22
Model for cement treatment	Min	19	0	0	2	0	0.00	0.00	0.00
	Q1	25	5	15	14	6	0.03	2.00	0.68
	Median	42	22	39	22	24	0.16	6.00	1.63
	Mean	40	18	34	29	27	0.21	5.87	1.78
	Q3	51	30	48	30	32	0.17	10.00	2.49
	Max	88	53	82	81	94	1.5	16.00	7.14

3.6.2 Type of Machine Learning Models

Several parametric (MLR, LR, DA), semiparametric (GAM), and non-parametric (KNN, SVM) machine learning models were chosen for this study. Such a wide range of models gives a good insight into the advantage of increasing a model's flexibility, the expense of its interpretability, and for improved predictive performance. Several literatures have pointed out there is no specific algorithm which performs well for all data types (Friedman 1995; Hastie et al. 2001; James et al. 2013). A brief introduction to

several machine learning algorithms used in the study for model development is presented in the following section. Details regarding these models can be found in Cortes and Vapnik (1995), Hastie et al. (2001), James et al. (2013); Kuhn and Johnson (2013), and Vapnik (1998, 2013).

Multiple Linear Regression (MLR)

MLR predicts a quantitative response (Y) based on linear combinations of predictors (X) as shown in Eq.(3-1) The regression coefficients (β), which relates the response with predictors, are calculated by minimizing the sum-of-squared errors (SSE) which is given by Eq. (3-2).

$$Y_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_p X_{ip} + \varepsilon_i \tag{3-1}$$

where, Y = quantitative response variable; $X_1, X_2, \ldots, X_p$ = predictors; $\beta_1, \beta_2, \ldots, \beta_p$ = regression coefficients; β_0 is the intercept, ε = independent error term

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3-2}$$

where, y_i the actual value and $\hat{y}_i$ is the model prediction

Significant advantage of MLR is its interpretability. The coefficients β represents change in the response with respect to a unit change in that specific predictor, holding all other constant. In addition to that, the statistical significance of a predictor can be computed, without complex simulations, which provides good inferential insight of the model. Limitation of this model is evident when the relationship between predictors and responses are not linear.

Logistic Regression (LR)

The output of LR model is the probability of association of a given sample to a particular categorical response. For instance, the probability of Y being 1 given a set of parameters X is given by Eq.(3-3). The coefficients β is calculated by maximum likelihood i.e., choosing β in such a way that the value of $p(X)$ in Eq. (3-3) is close to 1 for those samples belonging to category 1 and is close to 0 for those samples that do not belong to category 1. Such output are used to classify a given sample into binary classes.

$$p(X) = \Pr(Y = 1|X) = \frac{e^{\beta_0+\beta_1 X_1+\cdots+\beta_p X_p}}{1 + e^{\beta_0+\beta_1 X_1+\cdots+\beta_p X_p}} \quad (3\text{-}3)$$

where, Pr ($Y = 1|X$) = probability that Y=1 (1/0 response) given X; $X_1, X_2, \ldots, X_p$ = predictors; $\beta_1, \beta_2, \ldots, \beta_p$ = regression coefficients; β_0 is the intercept

LR is a powerful model due to its simple formulation and good inferential capabilities. Although the equation for $p(X)$ is non-linear, the decision boundary produced by this method is linear.

Discriminant Analysis (DA)

Discriminant analysis is a classification technique which is stable for well-separated classes in comparison to LR. Bayes theorem is used in the discriminant analysis for calculation of the posterior probability of a given observation "x" for each class "k". This classifier is mathematically represented by Eq. (3-4). For a given observation "x", the probability of it belonging to each of "k" classes is calculated using Eq.(3-4). The observation is assigned to the class which has the highest probability.

$$\Pr(Y = k|X = x) = \frac{\pi_k f_k(x)}{\sum_{l=1}^{k} \pi_l f_l(x)} \quad (3\text{-}4)$$

where, $\Pr(Y = k|X = x)$ = probability of x belonging to class k – posterior probability ; $f_k(x)$ = probability distribution function evaluated at x for class k; π_k = class membership probabilities – prior probability.

In case of linear discriminant analysis (LDA) for more than one predictor and k classes, $f_k(x)$ is assumed to have multivariate normal distribution with class specific mean vector (μ_k) and common covariance matrix ($\sum^{-1}$). This formulation results in a linear decision boundary. In case of quadratic discriminant analysis (QDA), $f_k(x)$ is also assumed to have multivariate normal distribution with class specific mean vector (μ_k) and class specific covariance matrix ($\sum^{-1}{}_k$). The resulting decision boundary in case of QDA is quadratic. Simplified formula for calculating the posterior class probability for LDA and QDA is given by Eq. (3-5) and Eq. (3-6) respectively.

$$\delta_k(k) = X^T\sum^{-1}\mu_k - \frac{1}{2}\mu_k{}^T\sum^{-1}\mu_k + \log\pi_k \tag{3-5}$$

$$\delta_k(k) = -\frac{1}{2}X^T\sum^{-1}{}_k X + X^T\sum^{-1}{}_k\mu_k - \frac{1}{2}\mu_k{}^T\sum^{-1}{}_k\mu_k + \log\pi_k \tag{3-6}$$

Generalized Additive Model (GAM)

GAM provides an excellent framework for linear models to extend their capability in the non-linear realm (Hastie 2017; Hastie and Tibshirani 1987; James et al. 2013). The model is set in such a way that each predictor is modeled by its non-linear function and maintains additivity among predictors. Formulation for GAM is given by Eq.(3-7).

$$Y_i = \beta_0 + f_1(x_{i1}) + f_2(x_{i2}) + \cdots + f_p(x_{ip}) + \varepsilon_i \tag{3-7}$$

where, Y_i = quantitative response; $f_1(x_{i1}), f_2(x_{i2}), \ldots, f_p(x_{ip})$ = non-linear function for individual predictor; β_0 is the intercept, ε = independent error term

The non-linear functions can be modeled using polynomials, splines, local regression, etc. – thus incorporating the non-linearity. Moreover, the additivity nature of the GAM retains its inference capability. In this study natural splines (NS) and smoothing splines (SS) were used

K-Nearest Neighbor (KNN)

Unlike the other methods discussed – where an equation is developed from data, KNN just utilizes the sample's location in feature space of entire training data to predict the response. This is known as instance based learning or lazy learning. Prediction is done by using only the "K" number of nearest training examples to the given sample. "K" is a tuning parameter which can be estimated using various model selection techniques. The measurement scale difference between different predictors greatly affects the distance calculation. Therefore, each predictor is generally scaled and centered prior to feeding the data into the model. The quantitative response is the mean or median of the K training example as given by Eq.(3-8). For categorical response, the probability of "K" nearest training data that belong to each class is calculated as per Eq.(3-9). The sample is assigned to the class with the highest probability.

$$Y(for\ a\ given\ X) = \frac{1}{K}\sum_{i \epsilon N_0} y_i \tag{3-8}$$

where, K = number of nearest points to 'X' that is to be considered; N_0 = nearest 'K' number of points to observation 'X'; y_i = response of each training data

$$\Pr(Y = j|X = x) = \frac{1}{K}\sum_{i \epsilon N_0} I(y_i = j) \tag{3-9}$$

where, $\Pr(Y = j|X = x)$ = probability of x belonging to class j; K = number of nearest points to 'x' that is to be considered; N_0 = nearest 'K' number of points to

observation 'x'; y_i = response (the class where x_i, the i-th nearest point of x belongs to); I( ) – indicator function.

KNN is non-parametric approach as it doesn't make assumptions on the distribution of predictors and decision boundary. Classification using this method is effective when the Bayes' decision boundary is highly non-linear (James et al. 2013).

Support Vector Machines (SVM)

SVMs are popular and are considered as one of the best out-of-box machine learning algorithms (James et al. 2013; Kuhn and Johnson 2013; Tinoco et al. 2014). SVM are based on generalizations of linear maximum margin classifier i.e. creating a linear maximum margin classifier in high-dimensional feature space – formed by projecting the input data using kernels (Cortes and Vapnik 1995; Kuhn and Johnson 2013). Maximum margin classifiers are based on separating hyperplanes- a flat affine surface of p-1 dimension, which perfectly separates the data into two classes. The distance between the hyperplane and nearest training data is called margin and the intent is to select such a hyperplane that maximizes the margin. Formulation for currently used SVMs are based on almost separating hyperplane, also called soft margin, which allows misclassification in the training phase. Provision of soft margin greatly increases the robustness of the classifier. The sign of the result, obtained by substituting the value of predictors in the hyperplane equation, is used to separate the class. The absolute value of such result reflects the confidence of classification. A standard representation of such hyperplane equation, as referred to as classification decision function, is given by Eq.(3-10). The coefficients α_i are obtained by maximizing the width of the margin under inequality constrains. These constrains control the tradeoff between a decision function's

complexity and amount of misclassification. The shape of the hyperplane is defined by a small subset of training data known as support vectors. Support vectors are those training examples which either a) lie of the margin, b) lie on the wrong side of the margin, or c) lie on the wrong side of the hyperplane. The classifier is not sensitive to extreme training data since the decision boundary is established by only support vectors.

$$f(x) = \beta_0 + \sum_{i \in S} \alpha_i y_i K(x, x_i) \tag{3-10}$$

For polynomial kernel

$$K(x_i, x_i') = \left(1 + \sum_{j=1}^{p} x_{ij} x'_{ij}\right)^d$$

For radial kernel

$$K(x_i, x_i') = e^{\left(-\gamma \sum_{j=1}^{p} (x_{ij} - x'_{ij})^2\right)}$$

where, f(x) = signed distance from hyperplane ; $K(x, x_i)$ = kernel (complexity of hyperplane equation) ; α_i = parameter associated with data point 'i'; S = support vectors., d = degree of polynomial; γ = positive constant.

Extension of SVM to regression setting has also become popular. The concept which was initiated by Smola (1996) and Drucker et al. (1997) . In a regression setting, the input vectors are mapped to high dimensional space using kernels after which a linear model fitting is done (Cherkassky and Ma 2004; Vapnik 1998, 2013). The coefficients of regression is estimated by minimizing the ϵ – insensitive loss function and models complexity.

3.6.3 Prediction Accuracy of Regression Models

The output of regression model is quantitative. Some of the standard metrics used by authors for accessing the models' accuracy are discussed in the following section (Chou et al. 2016; James et al. 2013; Kuhn and Johnson 2013; Tinoco et al. 2014).

Root Mean Square Error (RMSE)

RMSE is calculated by taking the square root of the mean of squared residuals as given by Eq. (3-11). The unit of RMSE is the same as the unit of predictor and the value is generally interpreted as the measure of average distance between the actual and model predicted values. This value is more sensitive to large error as it square the error term and give higher weight to large errors.

$$RMSE = \sqrt{\frac{\sum_{i=0}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{3-11}$$

Where, y_i = actual value of response; $\hat{y}_i$ = model predicted value of response; n = number of sample

Mean Absolute Error (MAE)

MAE is calculated by taking mean of the sum of absolute value of residuals as given by Eq.(3-12). The unit of MAE is the same as the unit of the predictor. The magnitude of MAE is less than or equal to RMSE and is less sensitive to large error as it doesn't square error term.

$$MAE = \frac{\sum_{i=0}^{n}|y_i - \hat{y}_i|}{n} \tag{3-12}$$

Where, y_i = actual value of response; $\hat{y}_i$ = model predicted value of response; n = number of sample

Coefficient of determination (R^2)

R^2 is a measure of percentage variance in the response explained by the model. This metric is the measure of the correlation between actual and predicted value, not the accuracy of the model. There are multiple formulations for R^2, but authors in this text used the formulation in James et al. (2013) which is given by Eq.(3-13).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i))^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (3\text{-}13)$$

Where, y_i = actual response value; $\hat{y}_i$ = model predicted value of response; $\bar{y}$ = mean of actual values; n = number of sample

3.6.4 Prediction Accuracy for Classification Models

Classification models output discrete classes as well as a sort of continuous probability of class association for a given set of test data. Although, the formulation of regression and classification models are similar, assessment of models' performance is distinctly different that from the regression model. Some of the standard model assessment metrics used by authors are discussed below in the following section (Ferri et al. 2009; Gneiting 2011; Kuhn and Johnson 2013; Majnik and Bosni 2013; Tofallis 2015).

Correct Prediction Rate (CPR)

CPR is the most straightforward representation of classifiers performance which is the ratio of total correct prediction by the total number of dataset as given by Eq.(3-14). The confusion matrix is a cross-tabulation of predicted and actual class for a given dataset. The sum of diagonal of that matrix represents total correct class prediction. Further computation of Cohen's kappa s provides information on whether the accuracy of CPR was due to chance (i.e., the relative frequency of each class) or efficacy of the classification model. Cohen's kappa is calculated by Eq.(3-15). The value of K=0 infers to accuracy by chance whereas K= 1 represents perfect agreement between model's prediction and actual classes.

$$CPR = \frac{sum\ of\ diagonal\ of\ confusion\ matrx}{Total\ samples} \quad (3\text{-}14)$$

$$Kappa(K) = \frac{O - E}{1 - E} \tag{3-15}$$

where, O = observed accuracy (from model's output i.e. CPR); E = expected accuracy (relative class frequency)

Area Under Curve (AUC)

AUC, the value of area of Receiver Operating Characteristics (ROC) curve, is a widely used measure of a classifier's performance (Bradley 1997; Majnik and Bosni 2013). Its values ranges from 0 to 1. A useful models' AUC shall exceed 0.5. ROC curve for binary classification is a 2-D plot with False Positive Rate (FPR) in x-axis against True Positive Rate (TPR) in y-axis for every threshold. TPR and FPR are calculated as per Eq. (3-16) and Eq.(3-17) respectively, by using the values from the confusion matrix. TPR represents the probability of actual true prediction when model predicts true which is also known as sensitivity. FPR represents probability of wrongly predicting the actual negative class as positive. AUC is particularly useful when comparing several classifiers without having to select a decision threshold.

$$TPR = \frac{TP}{TP + FN}$$

where, TP = True Positive – Number of correctly predicted positive result; FN = False Negative – Number of wrongly predicted positive result (3-16)

$$FPR = \frac{FP}{FP + TN}$$

where, FP = False Positive – Number of wrongly predicted negative result; TN = True Negative – Number of correctly predicted negative result (3-17)

3.6.5 Model Selection and Model Assessment

Model selection refers to finding the best performing model, for a given machine learning algorithm, by adjusting its hyperparameters. Model assessment refers to accessing the generalized performance of the selected model. The model assessment result is used to compare the performance of different machine learning algorithms. Model selection and model evaluations are most commonly done using a k-fold cross validation. This technique involves randomly dividing the dataset into k folds. The first fold is stored as a validation set – i.e., independent test set and rest of the fold is used for model fitting. An appropriate model evaluation metric discussed in the previous section is computed using the data on the validation set. This procedure is repeated for 'k' times where a different fold is treated as a validation set. The average value of the evaluation metric is reported as the k-fold cross validation estimate. For this study the value of k = 5 is used for cross-validation which is recommended by several studies (Chou et al. 2016; Hastie et al. 2001; Kohavi 1995).

For parametric models like MLR, LDA, QDA, and LR, the model selection is redundant since the structure of the model is already predefined and only requires calculation of model parameters from the training data. Unlike the parametric models, the semi-parametric and non-parametric models have specific parameters that has to be established prior to model assessmmment, called hyper-parameters or tuning parameters. For instance, the degree of freedom in case of splines in GAM is estimated by gradient decent in R using base R package (R Core Team 2017) coupled with Latin hypercube sampling to minimize the RMSE obtained by 5-fold cross-validation. A range of degrees of freedom was established for each parameter and Latin hypercube sampling was done

to choose a random combination of the degree of freedom. The random value was then used as a starting point for gradient descent algorithm for ensuring the algorithm doesn't get stuck in local minima. In the case of KNN, the value of "k" was optimized using grid search algorithm. A range of value of K was selected and performance metric was optimized using 5-fold cross-validation. In the case of regression, the value of "k" with minimum RMSE was chosen, while for the classification, the value of "k" giving maximum CPR was chosen. SVM regression with ϵ- insensitivity loss function was used for this study. ϵ was calculated as per Eq.(3-18) (Cherkassky and Ma 2004) while the regularization parameter (C), degree of polynomial, and kernel width (γ) were estimated by grid search algorithm using 5-fold cross validation. The values of the parameter that gave lowest RMSE was chosen for the model. In case of SVM classification, similar grid search approach was followed by maximizing the correct prediction rate for calculation of model parameters.

$$\epsilon = 3\sigma\sqrt{\frac{\ln n}{n}} \tag{3-18}$$

where, σ is the square root of noise variance calculated from k-nearest neighbor regression

After the hyper-parameters of models were estimated, selected model's generalized performance was accessed by 200 simulations of 5-fold cross-validation. The distribution of the average value of the evaluation metric for train and test set was done for all regression and classification models. The best model among various learning algorithms was chosen based on the median performance in the majority of the metrics. A pictorial representation the processes with 5-fold cross-validation and simulation used in

this study is shown in Figure 3-5. The results of the model selection and assessment are presented in the following section.

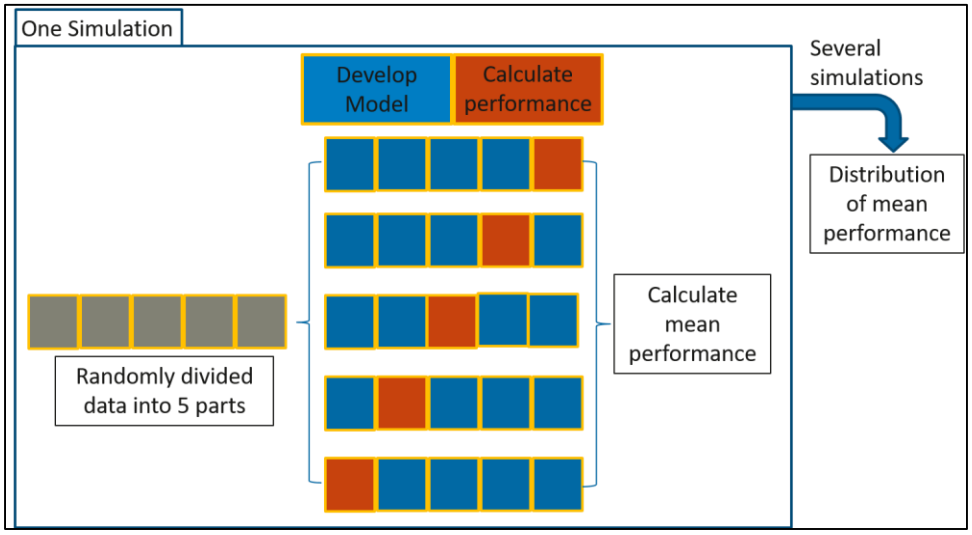

Figure 3-5: Procedure adopted for resampling using k-fold CV and random sampling

3.7 Results and spatial visualization

3.7.1 Model Selection

For GAM using natural cubic splines, the lack of data for fitting eight parameters resulted in a rank deficient matrix and results were removed from consideration. The model parameters for regression are given in Table 3-3 and for classification, with a threshold of 1.03 MPa (150 psi) for lime treated soil and 2.06 MPa (300 psi) for cement-treated soil, are given in Table 3-4.

Table 3-3: Tuning parameters for regression models

Model	Parameters	
	Lime Models	**Cement Models**
GAM-SS –	Degrees of freedom: LL = 3,PI =2 ,%Clay = 1,% Silt = 2, % Sand = 1 , Organic	Degrees of freedom: LL = 4,PI =3 ,%Clay = 2,% Silt = 3, % Sand = 2 , Organic

	Content =3 , Stabilizer = 4, LL.PI = 1	Content =2 , Stabilizer = 3, LL.PI = 3
KNN	K = 4	K = 3
SVM - Linear	C = 3.16, ϵ = 0.36	C = 0.56, ϵ = 0.91
SVM - Polynomial	C = 3.16,degree = 1, and ϵ = 0.31	C = 3.16, degree = 1, and ϵ = 0.91
SVM - Radial	C = 100, γ = 0.1, and ϵ = 0.31	C = 100, γ = 0.01, and ϵ = 0.91

Table 3-4: Tuning parameters for classification models

Model	**Parameters**	
	Lime Models	**Cement Models**
KNN	K = 5	K = 3
SVM - Linear	C= 236.43	C = 1
SVM - Polynomial	C = 177.82 and Degree = 1	C = 1.62 and Degree = 1
SVM - Radial	C = 31.62 and γ = 0.02	C = 56.23 and γ = 0.01

3.7.2 Model evaluation

The summary of model evaluations results are presented in Table 3-5, Table 3-6, and Table 3-7. These results are calculated for train and test sets by 200 random simulations of 5-fold cross validation (2-fold cross validation – for cement classification models). The Q1, Q2, and Q3 value of the evaluation metrics for various regression and classification models are presented in Figure 3-6 to Figure 3-9. The error bars in these figures represent the Q1 and Q3 while the marker represents the Q2 value. A star is given to each best performing metric based on the Q2 value. The model with the higher number of stars is selected as the best performing model.

Among regression models with lime treatment, the best performing model was SVM radial closely followed by KNN. The performance of all other remaining models

were identical. SVM with radial kernel has median R^2 value of 0.83 (train) and 0.75 (test), median MAE of 0.37 MPa (train) and 0.44 MPa (test), and median RMSE of 0.42MPa (train) and 0.50 MPa (test). In the case of regression models with cement treatment, the best performing model was MLR followed by GAM-SS and SVM-radial. MLR had median R^2 value of 0.88 (train) and 0.82 (test), MAE of 0.36 MPa (train) and 0.45MPa (test), and RMSE of 0.50Mpa (train) and 0.53 MPa (train).

Similarly, for classification models with lime treatment that has cutoff strength of 1.034 MPa (150 psi), SVM with radial kernel had the best performance which was closely followed by KNN and LR. SVM with radial kernel had median correct rate was 0.95 (train) and 0.92 (test), median Cohen's Kappa was 0.66 (train) and 0.60 (test), median AUC was 0.99 (train) and 0.98 (test), median TPR of 0.97 (train) and 0.94 (test), and median FPR of 0.06 (train) and 0.10 (test). As for the classification model for cement treatment that has cutoff strength of 2.068 MPa (300 psi), LDA has the best overall stable performance. LDA had median correct rate of 0.92 (train) and 0.80 (test), median Cohen's Kappa of 0.8 (train) and 0.52 (test), median AUC of 0.99 (train) and 0.90 (test), median TPR of 0.94 (train) and 0.77 (test), and median FPR of 0.05 (train) and 0.17 (test). Although SVMs' performed well in every other metric, AUC values of the SVMs' showed a bi-modal distribution where it was expected to have a normal distribution. The parameters for LR became unstable. Lack of data hindered the execution of QDA.

Table 3-5: Results for regression models

Model	Algorithm	R^2		MAE(MPa)		RMSE(MPa)	
		Train	Test	Train	Test	Train	Test
Lime	MLR	0.54-0.55-0.57	0.46-0.50-0.54	0.53-0.54-0.55	0.55-0.59-0.63	0.67-0.69-0.71	0.63-0.67-0.71
	GAM-SS	0.52-0.53-0.55	0.43-0.48-0.53	0.55-0.56-0.57	0.57-0.60-0.64	0.69-0.71-0.72	0.65-0.71-0.72
	KNN	0.80-0.81-0.82	0.65-0.71-0.78	0.30-0.31-0.32	0.37-0.41-0.46	0.43-0.45-0.47	0.47-0.53-0.57
	SVM - Linear	0.52-0.54-55	0.45-0.49-0.52	0.53-0.54-0.56	0.56-0.59-0.63	0.69-0.71-0.72	0.64-0.68-0.72
	SVM -	0.52-0.54-55	0.45-0.49-0.52	0.53-0.54-0.56	0.56-0.59-0.63	0.69-0.71-0.72	0.64-0.68-0.72
	SVM - Radial	0.83-0.83-0.84	0.71-0.75-0.79	0.37-0.37-0.38	0.41-0.44-0.47	0.41-0.42-0.43	0.45-0.50-0.53
Cement	MLR	0.87-0.88-0.89	0.75-0.82-0.87	0.34-0.36-0.37	0.39-0.45-0.53	0.47-0.50-0.53	0.44-0.53-0.61
	GAM-SS	0.86-0.87-0.88	0.66-0.78-0.85	0.37-0.39-0.40	0.43-0.50-0.59	0.51-0.53-0.54	0.45-0.53-0.61
	KNN	0.76-0.78-0.80	0.50-0.59-0.67	0.45-0.47-0.49	0.57-0.67-0.79	0.62-0.68-0.71	0.65-0.78-0.90
	SVM -Linear	0.77-0.79-0.81	0.69-0.75-0.80	0.48-0.53-0.56	0.51-0.59-0.68	0.64-0.66-0.68	0.54-0.62-0.74
	SVM -	0.71-0.74-0.77	0.63-0.70-0.80	0.56-0.59-0.63	0.58-0.64-0.71	0.68-0.73-0.78	0.62-0.70-0.78
	SVM - Radial	0.79-0.81-0.83	0.67-0.75-0.82	0.45-0.48-0.53	0.48-0.54-0.64	0.61-0.63-0.65	0.53-0.63-0.70

Note : The values in the table are Q1-Q2-Q3

Table 3-6: Results for classification models for lime treatment

Model	Algorithm	Correct Rate		Cohen's Kappa		AUC		TPR		FPR	
		Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
Lime (Cutoff)= 1.03 MPa (150psi)	LR	0.95-0.95-0.95	0.85-0.90-0.95	0.65-0.67-0.70	0.55-0.62-0.70	0.97-0.97-0.97	0.95-0.96-0.97	0.95-0.95-0.95	0.85-0.90-0.95	0.07-0.08-0.09	0.05-0.11-0.15
	LDA	0.90-0.91-0.91	0.85-0.88-0.93	0.61-0.62-0.65	0.52-0.58-0.68	0.98-0.98-0.98	0.94-0.96-0.97	0.88-0.90-0.91	0.82-0.86-0.93	0.15-0.17-0.20	0.15-0.20-0.26
	QDA	0.90-0.91-0.92	0.80-0.85-0.88	0.62-0.63-0.65	0.45-0.55-0.60	0.95-0.96-0.97	0.90-0.92-0.95	0.88-0.90-0.92	0.80-0.87-0.92	0.07-0.09-0.10	0.08-0.15-0.22
	KNN	0.95-0.95-0.95	0.85-0.90-0.93	0.70-0.71-0.71	0.54-0.62-0.70	-	-	0.94-0.95-0.96	0.86-0.91-0.95	0.04-0.05-0.06	0.08-0.10-0.17
	SVM - Lin	0.90-0.91-0.91	0.86-0.89-0.91	0.59-0.60-0.61	0.49-0.55-0.61	0.97-0.97-0.97	0.95-0.96-0.97	0.93-0.94-0.95	0.86-0.90-0.95	0.09-0.10-0.11	0.06-0.13-0.17
	SVM - Poly	0.90-0.91-0.91	0.86-0.89-0.91	0.59-0.60-0.61	0.49-0.55-0.61	0.97-0.97-0.97	0.95-0.96-0.97	0.93-0.94-0.95	0.86-0.90-0.95	0.09-0.10-0.11	0.06-0.13-0.17
	SVM - Rad	0.95-0.95-0.95	0.88-0.92-0.95	0.65-0.66-0.67	0.54-0.61-0.67	0.99-0.99-0.99	0.98-0.99-0.99	0.96-0.97-0.97	0.90-0.94-1.00	0.05-0.06-0.09	0.05-0.10-0.15

Note : The values in the table are Q1-Q2-Q3

Table 3-7: Results for classification models for cement treatment

Model	Algorithm	Correct Rate		Cohen's Kappa		AUC		TPR		FPR	
		Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
Cement (cutoff) = 2.06 MPa (300 psi)	LR	Unstable parameters									
	LDA	0.90-0.92-0.95	0.75-0.80-0.85	0.75-0.8-0.88	0.45-0.52-0.60	0.98-0.99-1.00	0.85-0.90-0.95	0.90-0.94-1	0.70-0.77-0.85	0.00-0.05-0.10	0.11-0.17-0.25
	QD	Lack of data									
	KNN	0.85-0.87-0.90	0.65-0.71-0.76	0.55-0.62-0.71	0.20-0.30-0.41	-	-	0.70-0.80-0.86	0.38-0.50-0.70	0.05-0.06-0.12	0.10-0.15-0.25
	SVM – Lin	0.90-0.93-0.95	0.78-0.82-0.86	0.70-0.78-0.86	0.42-0.52-0.62	0.00-0.50-1.00	0.10-0.60-0.95	0.83-0.90-0.95	0.70-0.77-0.82	0.00-0.04-0.05	0.08-0.14-0.20
	SVM - Poly	0.90-0.93-0.95	0.78-0.82-0.86	0.70-0.78-0.86	0.42-0.52-0.62	0.00-0.50-1.00	0.10-0.60-0.95	0.83-0.90-1.00	0.68-0.77-0.82	0.00-0.04-0.05	0.09-0.14-0.20
	SVM - Rad	0.94-0.95-0.96	0.78-0.82-0.85	0.74-0.80-0.88	0.44-0.50-0.60	0.00-0.99-1.00	0.1-0.88-0.92	0.85-0.90-1.00	0.68-0.75-0.86	0.00-0.00-0.05	0.10-0.12-0.18

Note : The values in the table are Q1-Q2-Q3

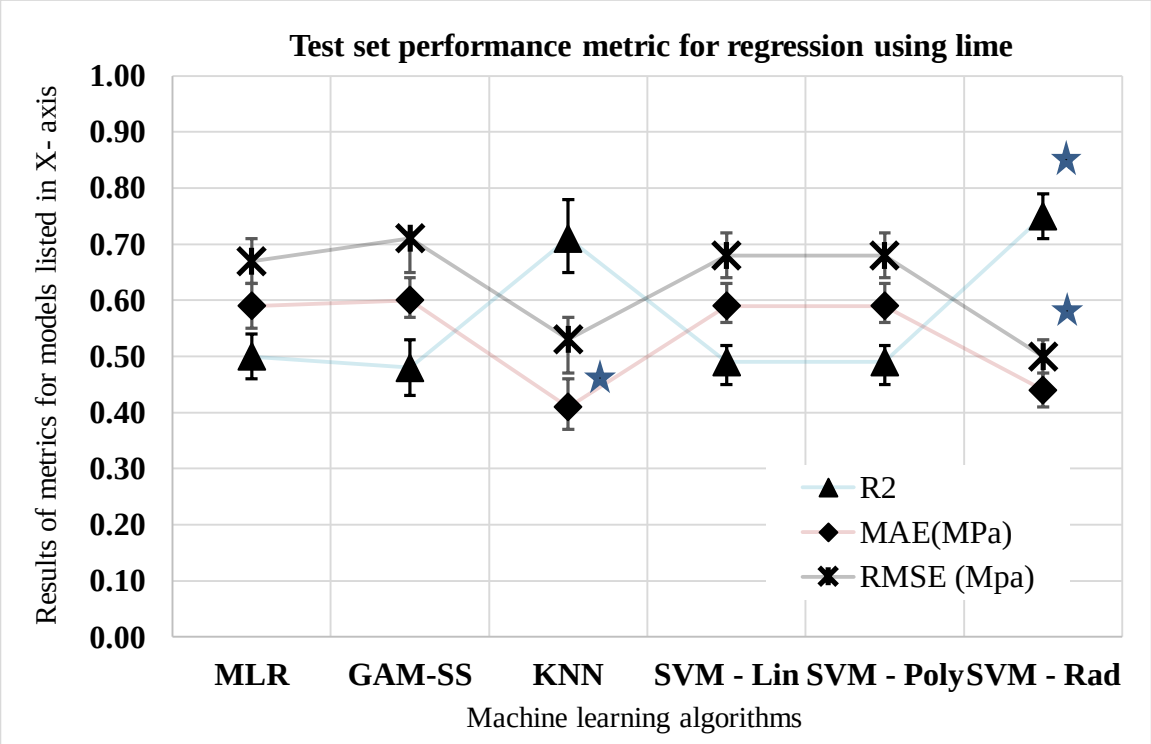

Figure 3-6: Summary of test set performance for regression for lime

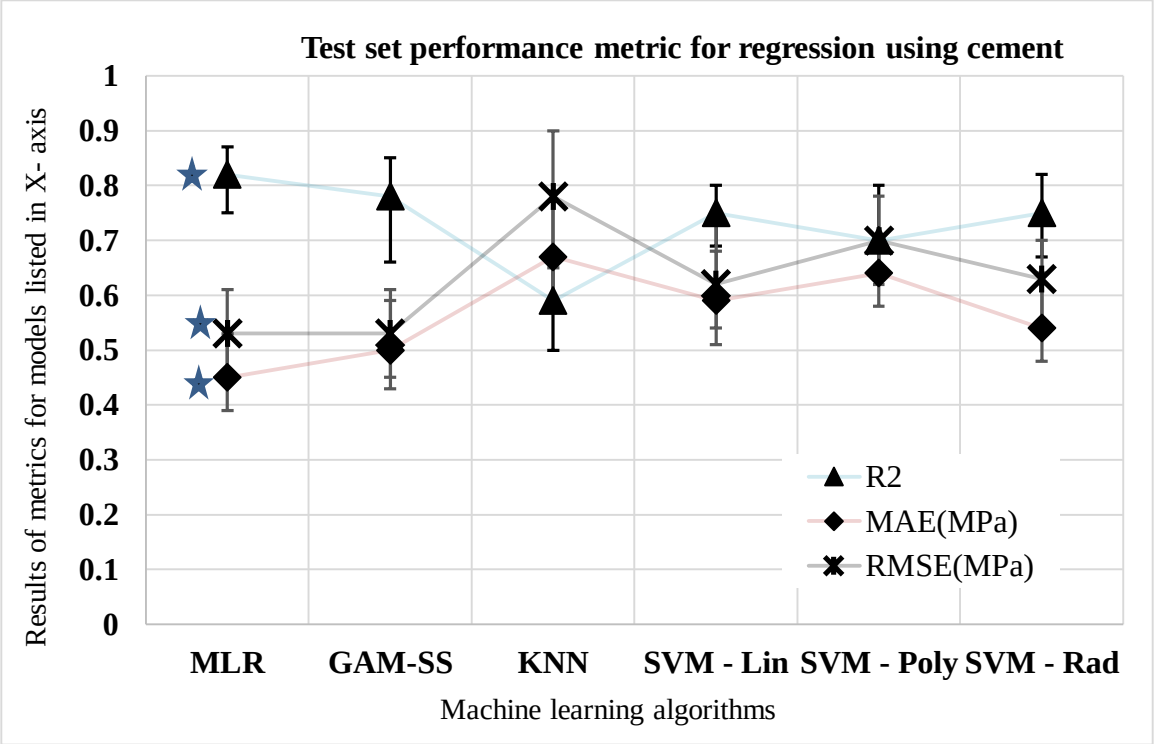

Figure 3-7: Summary of test set performance for regression for cement

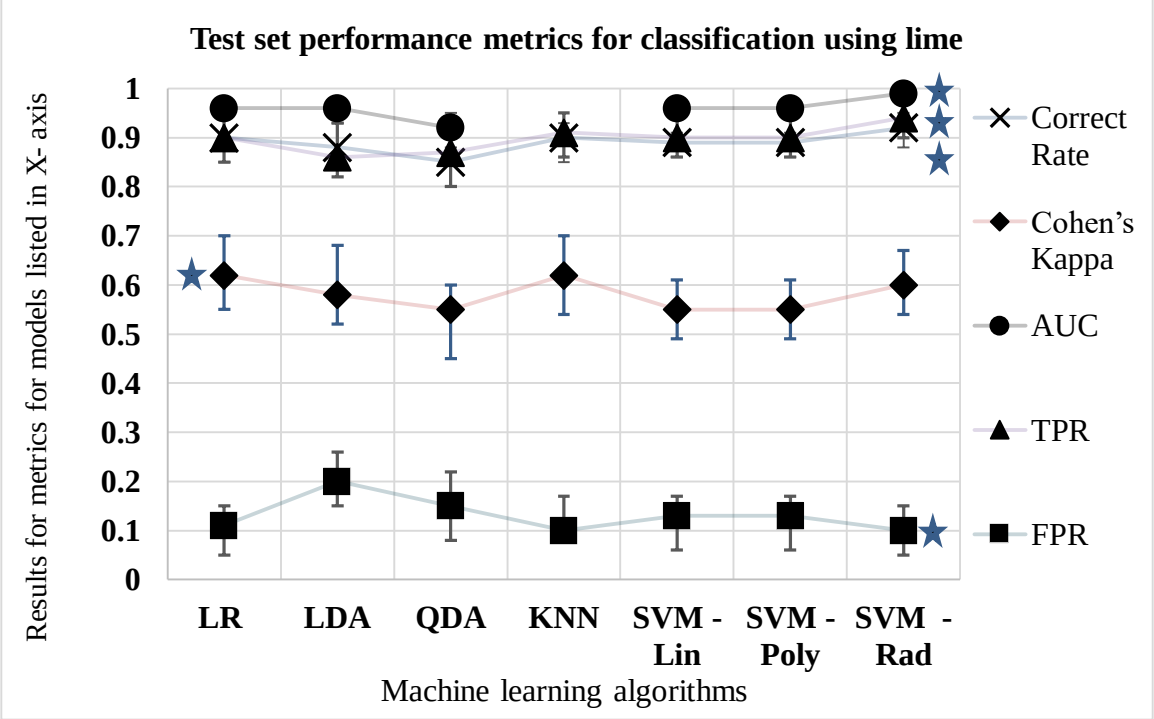

Figure 3-8: Summary of test set performance for classification for lime

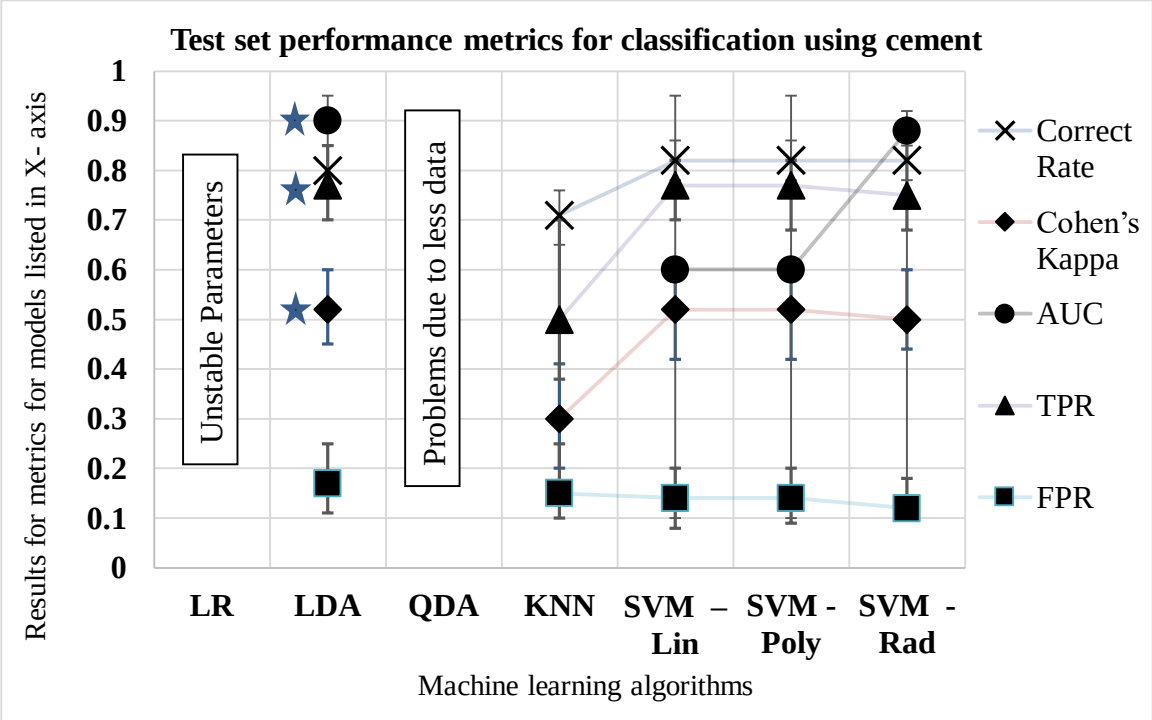

Figure 3-9: Summary of test set performance for classification for cement

3.5.1 Spatial visualization and comparison

An example application of the best performing models with their optimized hyperparameter are used for spatial visualization of quantitative and categorical results. For this purpose, Broadwater County in Montana was chosen. Soil parameters for the model were obtained from USDA's SSURGO database. The database was accessed by ArcGIS using the Soil Thematic Map Tool extension from USDA. The required parameters were imported from the database and individually rasterized. The rasters were then imported to R and were stacked as a raster brick using "raster" package by Hijmans and Etten (2017). As the parameters in the model were chosen to be same as the ones in the spatial data by USDA, regression and classification model was applied to the raster brick at each cell which resulted in a raster with predicted UCS value and binary response for pass/fail respectively for a given type and amount of stabilizer. For instance, the

regression result for application of 6% cement for Broadwater County is shown in Figure 3-10 and classification result for the threshold of 2.06 MPa (300 psi) is shown Figure 3-11 by using MLR and LDA respectively. The authors also collected two samples from Broadwater County and one sample from Garfield County from cement treatment in the laboratory. The treated UCS specimens in the laboratory were made using 2:1 aspect ratio and were cured at 100% humidity for seven days. The predicted UCS from the model using the parameters from the SSURGO data at that sampling location was plotted against the laboratory strength in Figure 3-12. The solid line in that figure represents the 45-degree line and the dotted line represents the 15% variation from the 45-degree line. In addition to the regression values, the CPR for the classification model was 92% for a total of 13 samples.

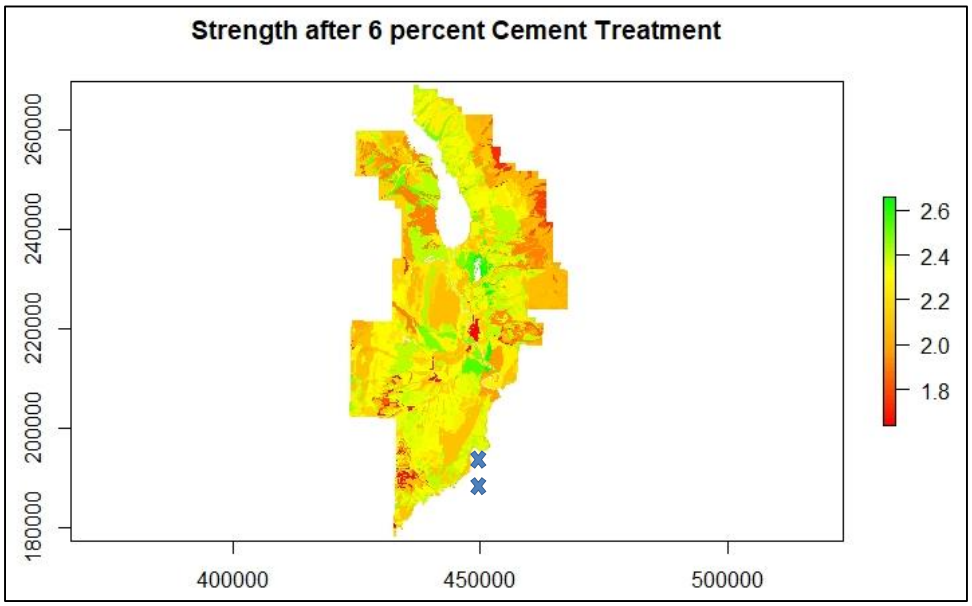

Figure 3-10: Spatial visualization regression result for UCS (MPa) of cement treated soil

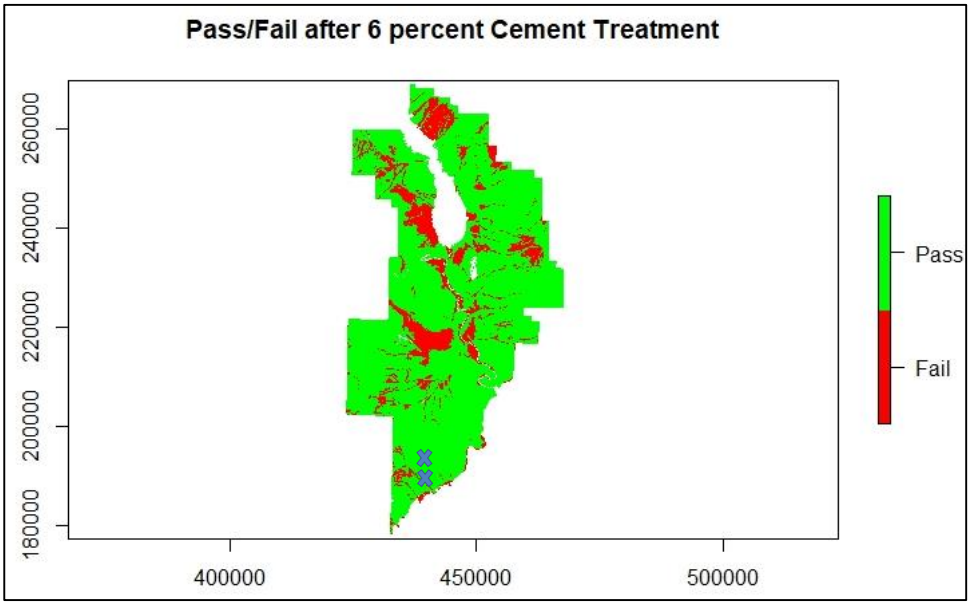

Figure 3-11: Spatial visualization classification result for 2.06 MPa UCS of Cement treated soil

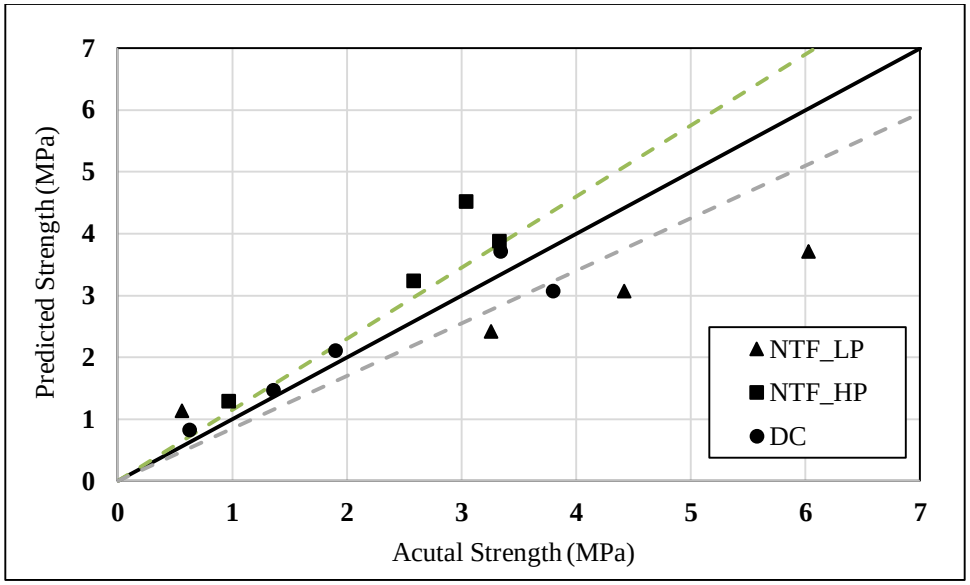

Figure 3-12: Actual vs predicted strength for spatial samples

3.6 Discussion

The results for this study indicate that for an end user who is interested in preliminary quantitative predictions, the regression model provides a decent estimation of UCS for any given type and amount of stabilization. Moreover, if the end user is more

concerned about a treatment passing a threshold, the classification model provides an excellent alternative. The performances of these regression and classification models strongly support the argument that there isn't a single model that works best for all the datasets. Therefore, the authors strongly suggest running a series of machine learning models before deciding on the specific model. Discussion regarding the model performance and comparison of model prediction with laboratory values are discussed in the following sections.

3.7.3 Model Performance

During the model selection process for regression models, the lack of sufficient data for the number of parameters to fit the GAM model with natural cubic splines created an issue of rank deficient matrix in both lime and cement models. Therefore, the results of the model were excluded from further analysis. Although the development of GAM model using smoothing splines was successful, its performance improvement against MLR was not significant (maximum of 6% in median values) for lime while the performance was worse than MLR in case of cement models. This suggests that the residuals that exists in the MLR model is not due its inability to model non-linear behavior and forcing non-linearity in the data, even with a model that has good regularization -smoothing splines, it doesn't increase the prediction performance of the model. Among non-parametric models, KNN for lime performed significantly better than MLR i.e. maximum of 47% increase in median performance. On the contrary, results for cement model using KNN was worse than MLR. Since KNN is an instance based learner, lack of sufficient data in case of cement treatment might have hindered its performance. Model selection results for SVM-linear and SVM –polynomial for lime and cement

models revealed similar model parameters. This suggests that projecting the dataset into kernel space of higher polynomial dimension didn't increase the models prediction performance. The performance of SVM-linear and SVM-polynomial models for lime was very much similar to that of MLR. As expected, SVM- radial showed improvement in performance in comparison to MLR for lime models. But in the case of cement model, the performance of all the SVM's were worse than MLR, although in this study, regularization parameter (C) and the degree of the polynomial (n) were optimized, the ϵ – insensitivity loss was taken as constant suggested by Cherkassky and Ma (2004). Authors believe in future models developments, tuning the ϵ – insensitivity loss might improve the performance as SVM have been known to be very sensitive to hyper parameters (Chapelle and Vapnik 2000; Vapnik 1998). Therefore, among all the models for regression for lime, SVM-radial's performance was the best with a median test set MAE and RMSE of 0.44 MPa and 0.50 MPa respectively. Whereas for cement model, MLR performance was the best with a median test set MAE and RMSE of 0.45 and 0.53 respectively.

In the case of classification for lime model with threshold of 1.03 MPa, best performing model was SVM radial which was closely followed by KNN and LR. The "γ" for SVM – radial model was also 0.02 and performance for all other models were comparatively similar. This suggests that the presence of a somewhat linear separating decision boundary between the two class of samples. In the case of cement samples, the coefficients for LR became unstable, AUC values for SVMs' showed a bimodal distribution, and small sample size barred the development of QDA model. Therefore, LDA performance was optimum for cement treatment models. The performance of best

model for lime treatment has median correct rate of 0.95 (train) and 0.80 (test), Cohen's kappa of 0.66 (train) and 0.61 (test), AUC of 0.99 (train) and 0.90 (test), TPR of 0.97 (train) and 0.94 (test), and FPR of 0.06 (train) and 0.10 (test). Similarly, the best performing cement treatment models has median correct rate of 0.92 (train) and 0.80 (test), Cohen's Kappa of 0.8 (train) and 0.52 (test), AUC of 0.99 (train) and 0.90 (test), TPR of 0.94 (train) and 0.77 (test), and FPR of 0.05 (train) and 0.11 (test).

The performance of regression model in this study using a limited set of the parameters was comparable to those of previous studies and but provides a wide range of application possibility of the model as the data for the parameters in the authors' model are spatially available.

3.7.4 Comparison of Model Predictions

The prediction of stabilized strength using the parameters from SSURGO database at each of three locations for the different amount of cement treatment showed good performance when compared to the laboratory data as shown in Figure 3-12. Even though the aspect ratio for the soil in the laboratory and the aspect ratio of the sample used in the model development was different, most of the prediction was within/near the ± 15 % of the 45-degree line. Moreover, the classification success rate for the threshold of 2.06 MPa (300 psi) was 92.31% i.e. only one sample out of the 13 samples was mispredicted. The failing sample which was predicted as passed had UCS of 1.91 MPa. This value is very near to the cutoff point and the difference between the predicted and actual strength was 0.15 MPa which is 7.24% of the cutoff value. Therefore, the results of regression as well as the classification model showed good prediction performance and further supports the application of these models for real-world applications.

3.8 Acknowledgment

The authors would like to heartily thank the Montana Department of Transportation (MDT) for financially supporting the graduate student involved in this research.

3.9 References

AASTHO. (1988). *Manual on Subsurface Investigations*. Washington, D.C.

Alavi, A. H., Gandomi, A. H., Gandomi, M., and Sadat Hosseini, S. S. (2009). "Prediction of maximum dry density and optimum moisture content of stabilised soil using RBF neural networks." *IES Journal Part A: Civil and Structural Engineering*, 2(2), 98–106.

Berry, K. M., Hasan, A., and Cooling, T. L. (2007). "Geotechnical Aspects of the St. Louis-Lambert International Airport Expansions." *Aviation: A World of Growth. The 29th International Air Transport ConferenceAmerican Society of Civil Engineers*.

Bhattacharya, B., and Solomatine, D. P. (2006). "Machine learning in soil classification." *Neural Networks*, 19(2), 186–195.

Bradley, A. (1997). "The use of the area under the ROC curve in the evaluation of machine learning algorithms." *Pattern Recognition*, 30(7), 1145–1159.

Burroughs, V. S. (2001). "Quantitative criteria for the selection and stabilisation of soils for rammed earth wall construction." University of New South Wales.

CDOT. (2017). Geotechnical Design Manual.

Chapelle, O., and Vapnik, V. (2000). "Model Selection for Support Vector Machines." *Advances in Neural Information Processing Systems 12*, (1), 230–236.

Cherkassky, V., and Ma, Y. (2004). "Practical selection of SVM parameters and noise estimation for SVM regression." *Neural networks*, Elsevier, 17(1), 113–126.

Ching, J., and Phoon, K.-K. (2014a). "Transformations and correlations among some clay parameters—the global database." *Canadian Geotechnical Journal*, NRC Research Press, 51(6), 663–685.

Ching, J., and Phoon, K.-K. (2014b). "Correlations among some clay parameters—the multivariate distribution." *Canadian Geotechnical Journal*, NRC Research Press, 51(6), 686–704.

Chittoori, B. C. S., Puppala, A. J., Wejrungsikul, T., and Hoyos, L. R. (2013). "Experimental Studies on Stabilized Clays at Various Leaching Cycles." *Journal of Geotechnical and Geoenvironmental Engineering*, 139(10), 1665–1675.

Chittoori, B., and Puppala, A. J. (2011). "Quantitative Estimation of Clay Mineralogy in Fine-Grained Soils." *Journal of Geotechnical and Geoenvironmental Engineering*, 137(11), 997–1008.

Chittoori, S., Pedarla, A., Puppala, A. J., Hoyos, L. R., Nazarian, S., and Saride, S. (2011). "Leachate studies on lime and portland cement treated expansive clays." *Geo-Frontiers 2011: Advances in Geotechnical Engineering*, 4479–4488.

Chou, J.-S., Yang, K.-H., and Lin, J.-Y. (2016). "Peak Shear Strength of Discrete Fiber-Reinforced Soils Computed by Machine Learning and Metaensemble Methods." *Journal of Computing in Civil Engineering*, 30(6), 4016036.

Cole, L. W., and Cepco, C. (2006). "PENNDOT'S Efforts to Improve Pavement Sub-grade During Highway Reconstruction–Airfield and Highway Pavements Specialty Conference." *Airfield and Highway Pavements Specialty Conference*, American Society of Civil Engineers, Atlanta, Georgia.

Cortes, C., and Vapnik, V. (1995). "Support-vector networks." *Machine learning*, Springer, 20(3), 273–297.

Croft, J. B. (1967). "The influence of soil mineralogical composition on cement stabilization." *Geotechnique*, Thomas Telford Ltd, 17(2), 119–135.

Das, S. K. (2012). "Artificial Neural Networks in Geotechnical Engineering: Modeling and Application Issues." *Metaheuristics in Water, Geotechnical and Transport Engineering*, Elsevier, Oxford,U.K., 231–270.

Das, S. K., Samui, P., and Sabat, A. K. (2011). "Application of Artificial Intelligence to Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilized Soil." *Geotechnical and Geological Engineering*, 29(3), 329–342.

Davidson, D. T., and Bruns, E. W. (1960). "Comparison of type I and type III Portland cements for soil stabilization." *Highway Research Board Bulletin*, (267).

Davidson, D. T., George, P. L., Mateos, M., and George, K. P. (1962). "Mositure-Density, Moisture-Strength and Compaction Charateristrics of Cement-Treated Soil Mixtures." *Bulletin*, 353.

Davidson, D. T., Mateos, M., and Barnes, H. F. (1960). "Improvement of Lime Stabilization of Montmorillonitic Clay Soils with Chemical Additives." *Highway Research Board Bulletin*, (262).

Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A. J., and Vapnik, V. (1997). "Support vector regression machines." *Advances in neural information processing systems*, 155–161.

Eades, J. L., and Grim, R. E. (1966). "A quick test to determine lime requirements for lime stabilization." *Highway Research Record*, (139), 61.

Ferri, C., Hernández-Orallo, J., and Modroiu, R. (2009). "An experimental comparison of performance measures for classification." *Pattern Recognition Letters*, Elsevier B.V., 30(1), 27–38.

FHWA. (1992). "Soil And Base Stabilization And Associated Drainage Considerations Volume II, Mixture Design Consideration." *Federal Highway Administration*.

Friedman, J. (1995). *Introduction to computational learning and statistical prediction*. Technical report, Department of Statistics, Stanford University.

Gneiting, T. (2011). "Making and evaluating point forecasts." *Journal of the American Statistical Association*, 106(494), 746–762.

Goh, A. T. C. (1994). "Seismic Liquefaction Potential Assessed by Neural Networks." *Journal of Geotechnical Engineering*, 120(9), 1467–1480.

Goh, A. T. C., and Goh, S. H. (2007). "Support vector machines: Their use in geotechnical engineering as illustrated using seismic liquefaction data." *Computers and Geotechnics*, 34(5), 410–421.

Goh, A. T. C., Kulhawy, F. H., and Chua, C. G. (2005). "Bayesian neural network analysis of undrained side resistance of drilled shafts." *Journal of Geotechnical and Geoenvironmental Engineering*, 131(1), 84–93.

Hampton, M. B., and Edil, T. B. (1998). "Strength gain of organic ground with cement-type binders, Soil Improvement for Big Digs." *Geotechnical Special Publication No*, 81, 135–148.

Hastie, T. J. (2017). "Generalized additive models." *Statistical models in S*, Routledge, 249–307.

Hastie, T., and Tibshirani, R. (1987). "Generalized additive models: some applications." *Journal of the American Statistical Association*, Taylor & Francis, 82(398), 371–386.

Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The elements of statistical learning*. Springer series in statistics New York.

Hijmans, R. J., and Etten, J. van. (2017). "raster: Geographic analysis and modeling with raster data."

IDOT. (2015). Geotechnical Manual.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *Introduction to Statistical Learning*. Springer.

Jones, D., Rahim, A., Saadeh, S., and Harvey, J. T. (2012). "Guidelines for the Stabilization of Subgrade Soils In California." *California Department of Transportation*.

Kohavi, R. (1995). "A study of cross-validation and bootstrap for accuracy estimation and model selection." *International Joint Conference on Artificial Intelligence*, Montreal, Canada, 1137–1145.

Kuhn, M., and Johnson, K. (2013). *Applied predictive modeling*. Springer.

Ladd, C. C., Moh, Z. C., and Lambe, T. W. (1960). "Recent Soil-Lime Research at the Massachusetts Institute of Technology." *Highway Research Board Bulletin*, 262.

Lai, S., Chang, W., and Lin, P. (2006). "Logistic Regression Model for Evaluating Soil Liquefaction Probability Using CPT data." *Journal of Geotechnical and Geoenvironmental Engineering*, 132(June), 694–704.

Leohr, J. E., Lutenegger, A., Rosenblad, B., and Boeckmann, A. (2016). *Geotechnical Site Charaterization*. Washington, D.C.

Little, D. N., and Nair, S. (2009). "Recommended practice for stabilization of subgrade soils and base materials." National Cooperative Highway Research Program, Transportation Research Board of the National Academies Washington, DC.

Little, N. . (1999). Evaluation of structural properties of lime stabilized soils and aggregates.

Majnik, M., and Bosni, Z. (2013). "ROC analysis of classifiers in machine learning : A survey." 17, 531–558.

Marjanović, M., Kovačević, M., Bajat, B., and Voženílek, V. (2011). "Landslide susceptibility assessment using SVM machine learning algorithm." *Engineering Geology*, 123(3), 225–234.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., and Leisch, F. (2017). "e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071)." TU Wien.

Mitchell, J. K., and Soga, K. (2005). "Fundamentals of soil behavior." John Wiley & Sons Hoboken, NJ.

Mozumder, R. A., and Laskar, A. I. (2015). "Prediction of unconfined compressive strength of geopolymer stabilized clayey soil using Artificial Neural Network." *Computers and Geotechnics*, Elsevier Ltd, 69, 291–300.

Mozumder, R. A., Laskar, A. I., and Hussain, M. (2017). "Empirical approach for strength prediction of geopolymer stabilized clayey soil using support vector machines." *Construction and Building Materials*, Elsevier Ltd, 132, 412–424.

Nelson, J. D., and Miller, D. J. (1997). Expansive soils: Problems and Practice in Foundation and Pavement Engineering. Wiley, New York, New York.

NLA. (2004). “Lime-Treated Soil Construction Manual Lime Stabilization & Lime Modification.” *National Lime Association*.

NLA. (2006). “Mixture Design and Testing Procedures for Lime Stabilized Soil.” *National Lime Association*.

Pal, M. (2006). “Support vector machines-based modelling of seismic liquefaction potential.” *International Journal for Numerical and Analytical Methods in Geomechanics*, Wiley Online Library, 30(10), 983–996.

PCA. (1992). “Soil Cement Laboratory Handbook.” *Portland Cement Association*.

Pedarla, A., Chittoori, S., and Puppala, A. J. (2011). “Influence of mineralogy and plasticity index on the stabilization effectiveness of expansive clays.” *Transportation Research Record*, SAGE Publications Sage CA: Los Angeles, CA, 2212(1), 91–99.

Petry, T. M., and Little, D. N. (2002). “Review of Stabilization of Clays and Expansive Soils in Pavements and Lightly Loaded Structures—History, Practice, and Future.” *Journal of Materials in Civil Engineering*, 14(6), 447–460.

Pietsch, P. E., and Davidson, D. T. (1961). “Effects of lime on plasticity and compressive strength of representative Iowa soils.” Iowa State University.

Prusinski, J., and Bhattacharja, S. (1999). “Effectiveness of Portland cement and lime in stabilizing clay soils.” *Transportation Research Record: Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, (1652), 215–227.

Puppala, A. J. (2016). “Advances in ground modification with chemical additives: From theory to practice.” *Transportation Geotechnics*, 9, 123–138.

R Core Team. (2017). “R: A language and environment for statistical computing. R Foundation for Statistical Computing.” Vienna, Austria.

RStudio Team. (2016). “RStudio: Integrated Development for R.” Rstudio, Inc., Boston,MA.

Remus, M. D., and Davidson, D. T. (1961). “Relation of Strength to composition and density of lime-treated clayey soils.” *Highway Research Board Bulletin*, (304).

Samui, P. (2008). “Support vector machine applied to settlement of shallow foundations on cohesionless soils.” *Computers and Geotechnics*, 35, 419–427.

Sariosseiri, F., and Muhunthan, B. (2009). “Effect of cement treatment on geotechnical properties of some Washington State soils.” *Engineering Geology*, Elsevier B.V., 104(1–2), 119–125.

Sing, T., Sander, O., Beerenwinkel, N., and Lengauer, T. (2005). “ROCR: visualizing classifier performance in R.” *Bioinformatics*, 21(20), 7881.

Smola, A. J. (1996). “Regression estimation with support vector learning machines.” Master’s thesis, Technische Universität München.

Soil Survey Staff. (n.d.). “Natural Resources Conservation Service, United States Department of Agriculture,Web Soil Survey.” *Web Soil Survey*.

Suman, S., Mahamaya, M., and Das, S. K. (2016). “Prediction of Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilised Soil Using Artificial Intelligence Techniques.” *International Journal of Geosynthetics and Ground Engineering*, Springer International Publishing, 2(2), 11.

Tastan, E. O., Edil, T. B., Benson, C. H., and Aydilek, A. H. (2011). “Stabilization of organic soils with fly ash.” *Journal of geotechnical and Geoenvironmental Engineering*, American Society of Civil Engineers, 137(9), 819–833.

Tex-100-E. (1999). Surveying and Sampling Soils for Highways.

Thompson, M. R. (1967). *Factors influencing the plasticity and strength of lime-soil mixtures*. University of Illinois at Urbana Champaign, College of Engineering. Engineering Experiment Station.

Thompson, M. R. (1972). “Deep-plow lime stabilization for pavement construction.” *Transportation Engineering Journal of ASCE*, ASCE, 98(2), 311–323.

Tinoco, J., António Alberto Santos, C., Da Venda, P., Correia, A. G., and Lemos, L. (2016). "A Data-driven Approach for qu Prediction of Laboratory Soil-cement Mixtures." *Procedia Engineering*, Elsevier B.V., 143(Ictg), 566–573.

Tinoco, J., Gomes Correia, A., and Cortez, P. (2011). "Application of data mining techniques in the estimation of the uniaxial compressive strength of jet grouting columns over time." *Construction and Building Materials*, Elsevier Ltd, 25(3), 1257–1262.

Tinoco, J., Gomes Correia, A., and Cortez, P. (2014). "Support vector machines applied to uniaxial compressive strength prediction of jet grouting columns." *Computers and Geotechnics*, Elsevier Ltd, 55, 132–140.

Tofallis, C. (2015). "A better measure of relative prediction accuracy for model selection and model estimation." *Journal of the Operational Research Society*, 66(8), 1352–1362.

Tremblay, H., Duchesne, J., Locat, J., and Leroueil, S. (2002). "Influence of the nature of organic compounds on fine soil stabilization with cement." *Canadian Geotechnical Journal*, NRC Research Press, 39(3), 535–546.

TxDOT. (2005). "Guidelines for Modification and Stabilization of Soils and Base for Use in Pavement Structures." *Texas Deparment of Transportation*.

U.S. Army TM 5-882-14/AFM 32-1019. (1994). Soil Stabilization for Pavements. Technical Manual No. 5-822-14; Air Force Manual No. 32-1019, Joint Departments of the Army and the Air Force, USA.

Vapnik, V. (1998). *Statistical learning theory. 1998*. Wiley, New York.

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.

Veisi, M., Chittoori, B. S., Celaya, M., Nazarian, S., Puppala, A. J., and Solis, C. (2010a). *Accelerated Stabilization Design of Subgrade Soils*. El Paso.

Veisi, M., Chittoori, B. S., Celaya, M., Nazarian, S., Puppala, A. J., and Solis, C. (2010b). *Accelerated Stabilization Design of Subgrade Soils*. El Paso.

Wen, H., Muhunthan, B., Wang, J., Li, X., Edil, T., and Tinjum, J. M. (2014). Characterization of Cementitiously Stabilized Layers for Use in Pavement Design and Analysis. National Cooperative Highway Research Program.

Whitehurst, E. A. (1955). "Stabilization of Tennessee Gravel and Chert Bases." *Highway Research Board Bulletin*, (108).

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

WSDOT. (2010). Geotechnical design manual.

SUMMARY, CONCLUSION, AND RECOMMENDATIONS

Summary and Conclusion

The purpose of this thesis was to assess applicability of models for prediction of UCS value of stabilized soil that are developed using different machine learning algorithms and using only the parameters that are spatially available in databases like SSURGO. Such model would be very beneficial in the selection of optimum initial stabilizer in the laboratory works, aiding preliminary strength assessment for a given spatial location, delineation of laboratory sample collection, etc. without having to bring the soil to the laboratory. In addition to that, this model makes use of the vast amount of spatial soil data that is readily available. With this intent, classification models (using LR, KNN, DA, SVMs) were developed using existing Australian database as a pilot study. Then similar database for the US soil were compiled. Data from the US soils were used for regression models (using MLR, GAM, KNN, SVMs') and classification models (using LR, DA, KNN, SVMs') development. The major findings from this study are:

1) Classification models, developed using Australian database, showed good generalized performance and encouraged a compilation of similar database for the US soils.
2) The median performance of regression models developed using the US database are similar or better than the performance report in the literature.

3) The performance of classification models developed using the US database was found to be satisfactory, with respect the chosen threshold values.
4) The comparison of the best performing regression and classification models' outputs with the actual laboratory strength further supports the applicability of these models in the US.
5) The output of the model can be used to select an optimum stabilizer content, get an overview of stabilization performance for any spatial area within the US, and strategize the sampling operations for chemical stabilization projects.
6) Analysis of the models results strongly suggest that a single model cannot perform well for all the type of data. This result is in perfect agreement with many research in machine learning community.
7) As the models were derived with the US soils database with parameters also available in SSURGO, the results of the model better represent the US soils.

Recommendations for Future Research

1) The quantity of test results available in published journals and dissertations, which were used in the development of the current US database, were restricted by aspect ratio, compaction energy, and type of stabilizer. Addition of test results from the additional unpublished literature will result in a diverse database for better generalized performance over a wide range of soil.

2) Development of similar models using machine learning algorithms for durability assessment of chemical stabilization would be very beneficial for the practitioners
3) Development of graphical user interface (GUI) for the current source code.

APPENDIX A

Database for the US Soil

The database for complied for the US soils and digitized for the Australian soils are uploaded in the GitHub. Please follow the link

https://github.com/amitgajurel/MS_Thesis

Spatial visualization of strength for Broadwater County

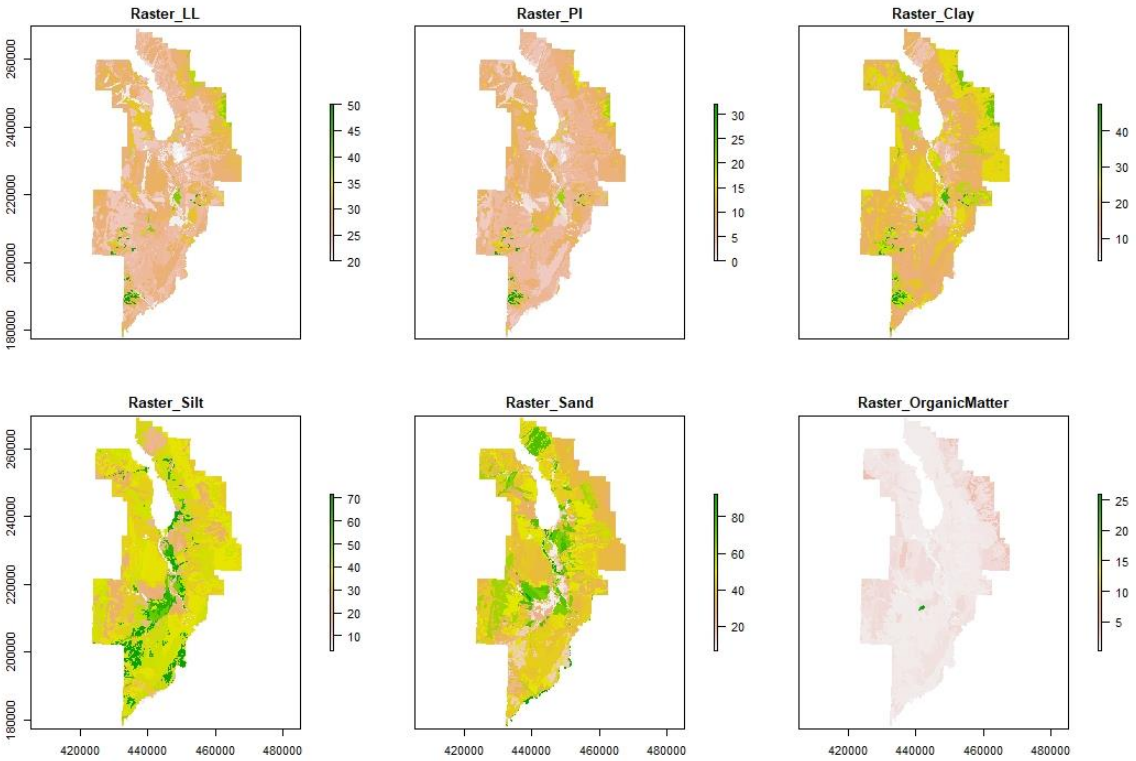

Figure A-1: Spatial visualization of input parameters for Broadwater County

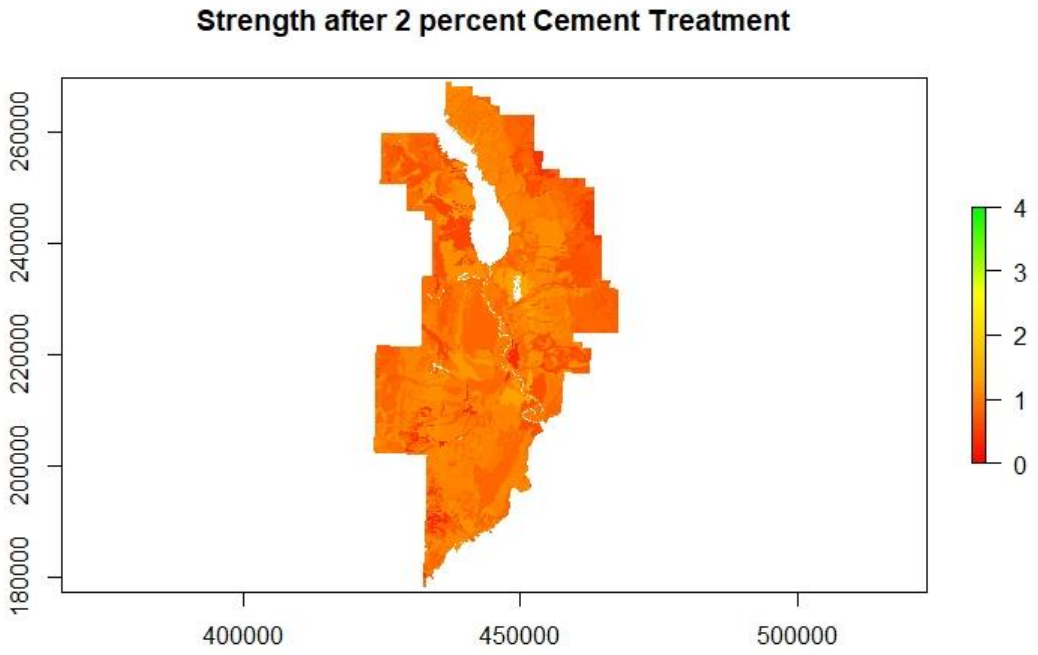

Figure A-2: Spatial UCS visualization of regression results for 2 percent cement treatment

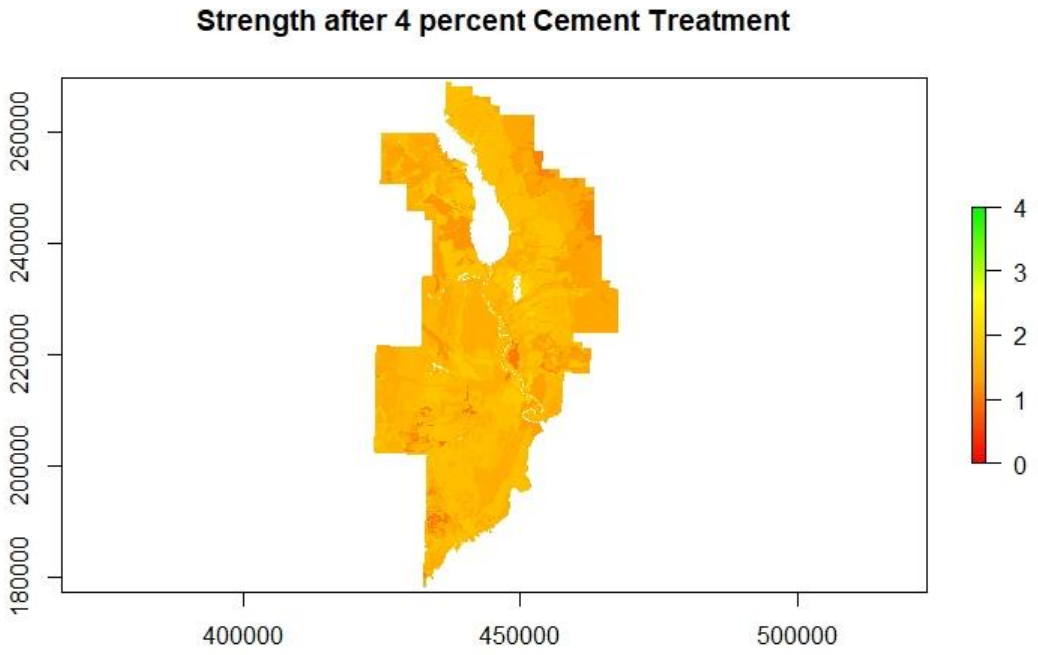

Figure A-3: Spatial UCS visualization of regression results for 4 percent cement treatment

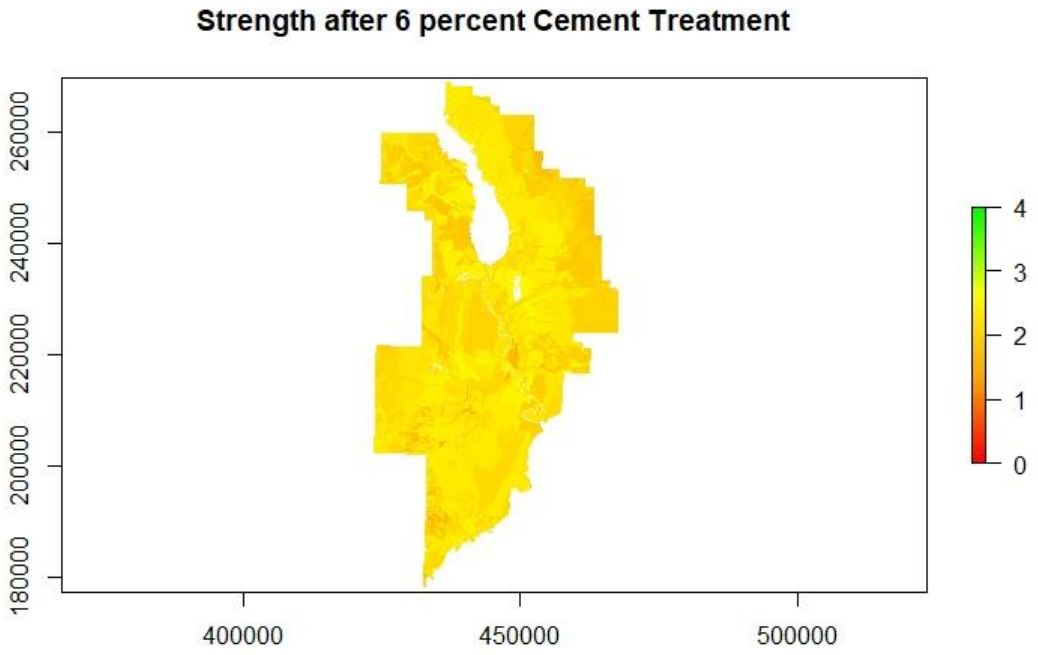

Figure A-4: Spatial UCS visualization of regression results for 6 percent cement treatment

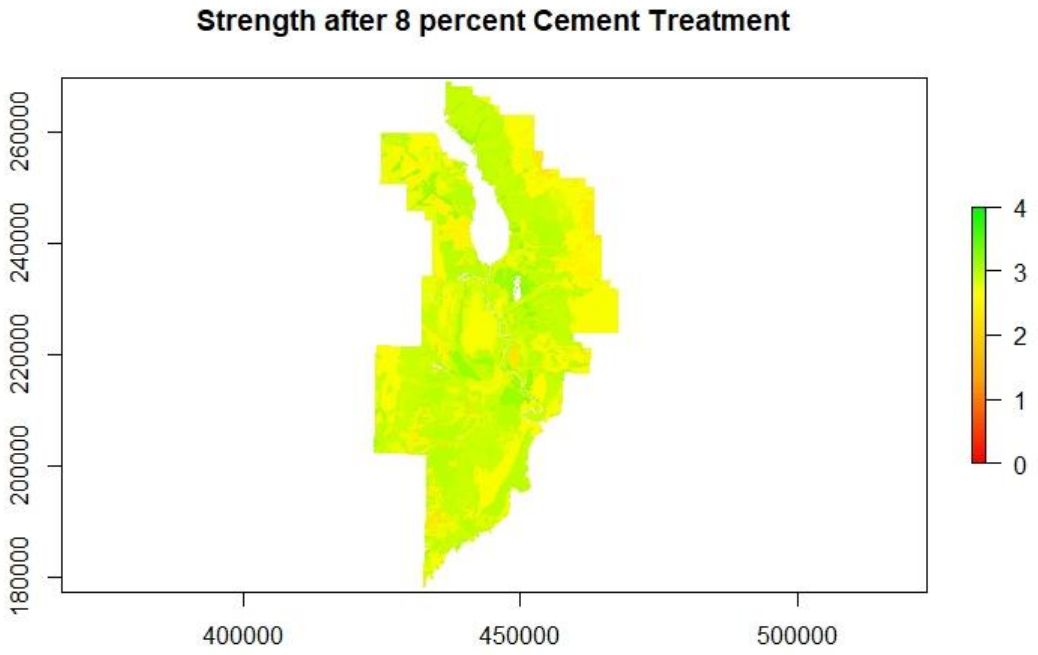

Figure A-5: Spatial UCS visualization of regression results for 8 percent cement treatment

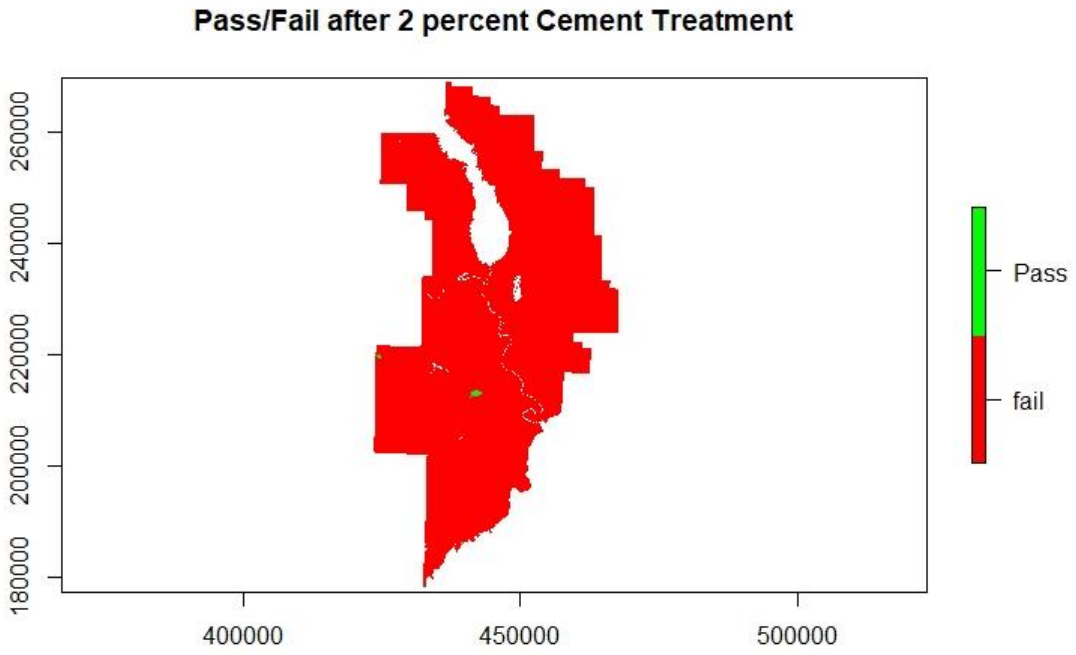

Figure A-6: Spatial UCS visualization of classification results for 2 percent cement treatment

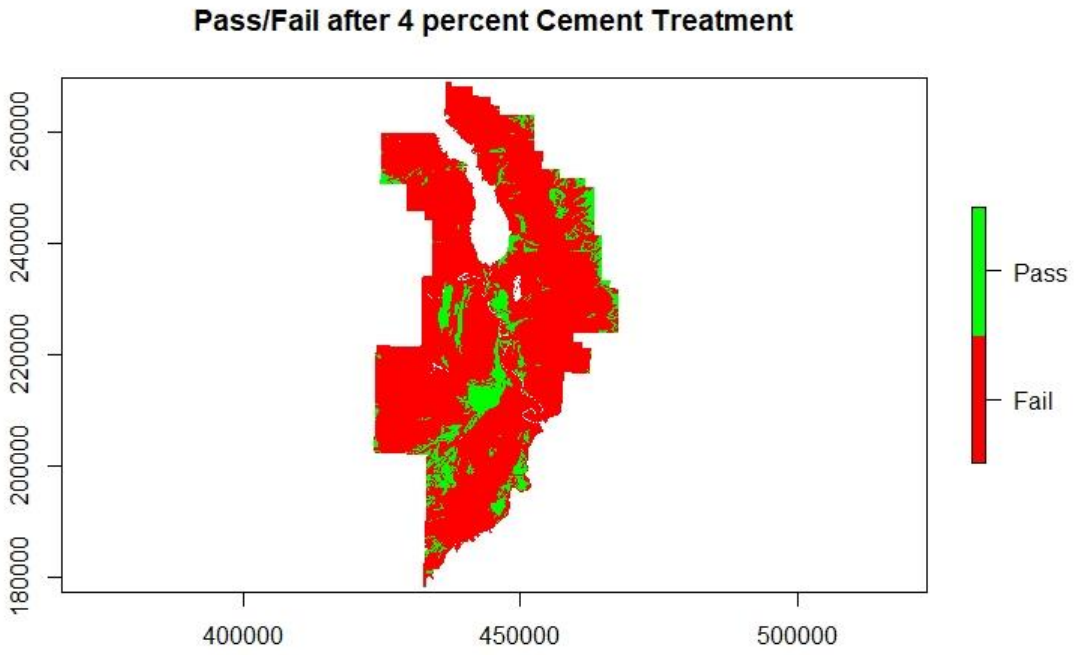

Figure A-7: Spatial UCS visualization of classification results for 4 percent cement treatment

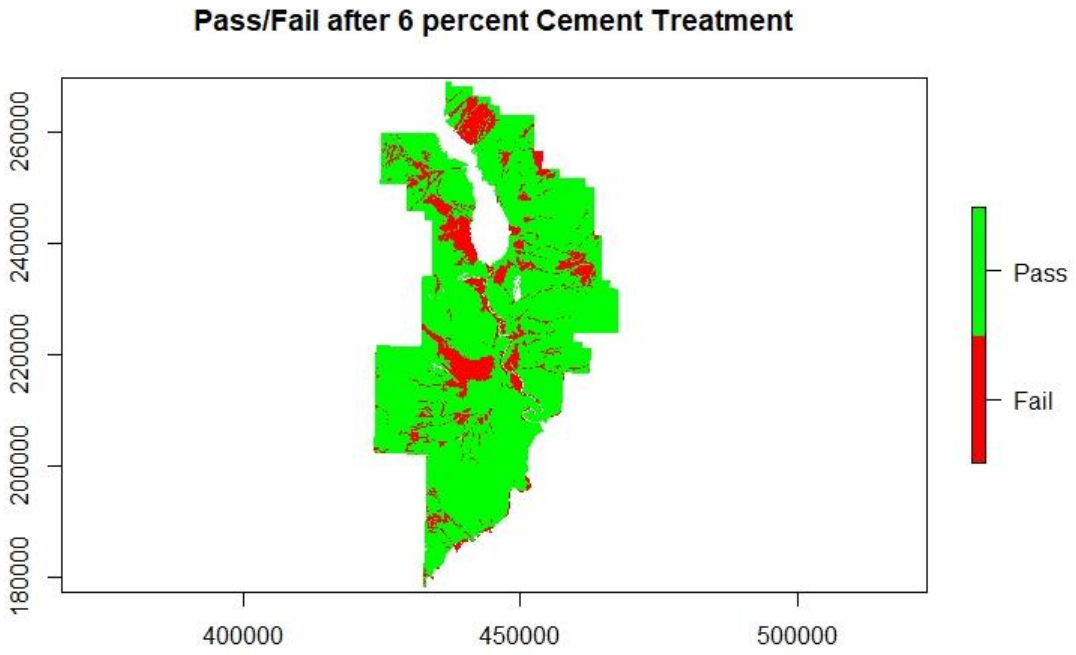

Figure A-8: Spatial UCS visualization of classification results for 6 percent cement treatment

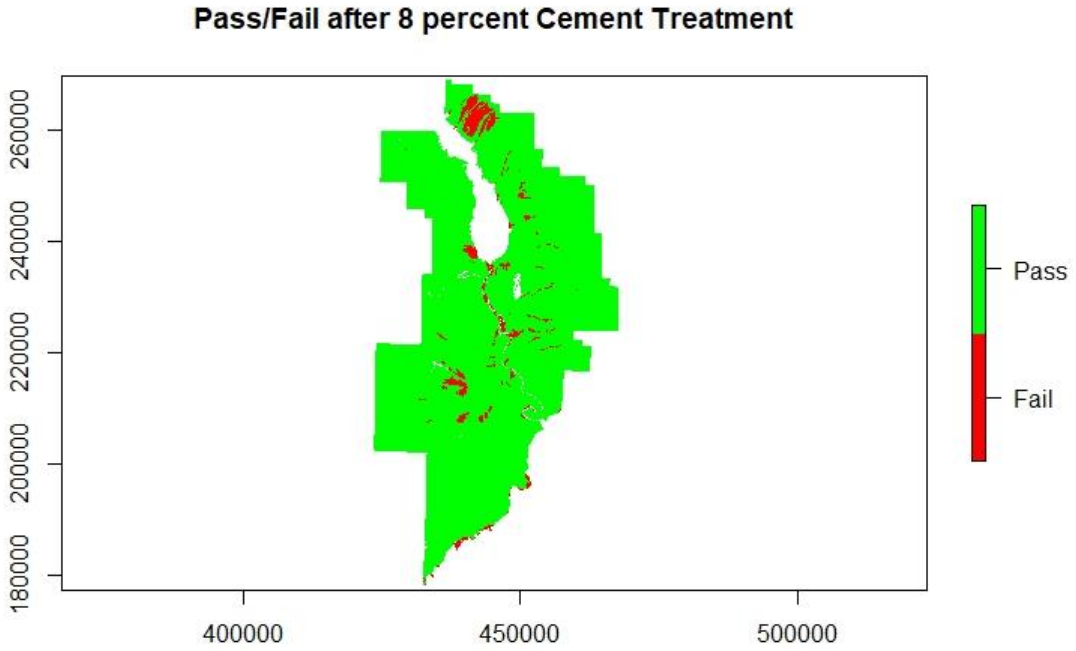

Figure A-9: Spatial UCS visualization of classification results for 8 percent cement treatment