MACHINE LEARNING – ENABLED LIQUEFACTION HAZARD ASSESSMENT OF SOME INDIAN SANDY SOIL DEPOSITS

TANMAY GUPTA

DEPARTMENT OF CIVIL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

FEBRUARY 2024

MACHINE LEARNING – ENABLED LIQUEFACTION HAZARD ASSESSMENT OF SOME INDIAN SANDY SOIL DEPOSITS

by

TANMAY GUPTA

DEPARTMENT OF CIVIL ENGINEERING

Submitted

in fulfilment of the requirements of the degree of Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

FEBRUARY 2024

Dedicated to Lord Mahakaleshwar

&

My Parents, Priyal, Bhavesh and Stavya

CERTIFICATE

This is to certify that the thesis entitled "**Machine Learning-Enabled Liquefaction Hazard Assessment of Some Indian Sandy Soil Deposits**", being submitted by **Tanmay Gupta** to the Indian Institute of Technology, Delhi, for the award of the degree of **Doctor of Philosophy,** is a bonafide record of research work carried out by him under our supervision and guidance. In our opinion, the thesis work has reached the requisite standard, fulfilling the requirements for the Doctor of Philosophy.

The results in this thesis have not been submitted, in part or full, to any other University or Institute for awarding any degree or diploma.

(**G. V. Ramana**)
Professor
Department of Civil Engineering
Indian Institute of Technology Delhi
Hauz Khas
New Delhi – 110016
India

(A. W. Elgamal)
Distinguished Professor
Department of Structural Engineering
University of California San Diego
Jacobs School of Engineering
California – 92093
The United States of America

Acknowledgements

Firstly, thanks to the almighty Lord Mahakaleshwar for His blessings and continued guidance throughout my educational journey.

At the onset, I would like to express my sincere gratitude to my supervisors, Prof. G. V. Ramana and Prof. Ahmed W. Elgamal, for their unconditional support and encouragement throughout this research work. Their guidance enabled me to achieve the objectives of this thesis successfully. I enjoyed the discussions we had over the research period. We had the most fruitful discussions during our coffee walks and discussing the physics behind geotechnical engineering. Prof. Elgamal's effective way of communicating his views is unique and exemplary. Prof. Ramana's friendly attitude made this journey remarkable. No number of words could do justice to their extraordinary guidance and inspiring nature.

I want to thank the SRC members, Prof. R Ayothiraman, Prof. B. J. Alappat and Prof. D. Ravi Kumar, for sparing their valuable time in reviewing this work. Their observations and guidance have helped me effectively communicate and frame my research findings.

I thank M/s Keller Ground Engineering India Pvt Ltd (my current employer) for showing interest and facilitating this journey. I want to mention the support of Mr Deepak Raj (MD – Keller ASEAN) and Mr Y. Hari Krishna (MD – Keller India) during this period. The technical discussions with Mr Valluri Sridhar (Head – BD – Keller India), Mr Madan K. Annam (Head – Engineering – Keller India) and Mr Jonathan Daramalinggam (Engineering Director – Keller AMEA) have encouraged me to come up with a simple but effective solution to the soil liquefaction. The routine discussions I had with Mr Saday Mishra (GM – N&W – Keller India), Mr Govind Raj (DGM – Geotechnical – Keller India), Mr P V S R Prasad (DGM – Geotechnical – Keller India) and Mrs Sangeen A Desai (Deputy Manager – BD) gave me a lateral width over CPT interpretation and routinely faced engineering issues at a project site.

Thanks to the operations, human resources, and safety team of Keller India for their support during the execution of the CPTs at various locations across India.

My time at IIT Delhi was joyous and adventurous with my friends Mr Sourabh Mhaski, Mr Pranjal Singh, and Mr Deepak Haritwal. The field visits and discussions provided significant value addition to this research work. I would also like to thank Dr. Sumeet Kumar Sinha for his critical remarks/inputs and editorial corrections, which helped improve the work's quality.

Last but not least, the love and support I received from my family could not be expressed in words. My father, Mr Y. K. Gupta and my mother, Mrs Rajni Gupta, supported me in every way imaginable to ensure I received the best education. My brother, Mr Bhavesh Gupta, has been a rock supporter. The unconditional love and support from my better half are the most important things which kept me going. The highlight of this journey is the relaxing time after a hard day's work with my son, Mr Stavya Gupta.

(Tanmay Gupta)

Abstract

Liquefaction of saturated clean sands caused by earthquake-induced shear stresses is responsible for severe damage to infrastructures worldwide. Structures built over marine, fluvial and alluvial deposits are susceptible to liquefaction hazards during earthquakes. Liquefaction could result in catastrophic failures, as evident from the infamous 1964 Niigata earthquake in the past and, more recently, during the 2021 Assam India and 2022 Java Indonesia earthquakes. The severe consequence of liquefaction necessitates the assessment of the triggering of liquefaction and estimating liquefaction-induced settlement as critical tasks for practising engineers.

Traditional deterministic approaches for assessing liquefaction hazards do not account for soil's lateral variability and uncertainties associated with earthquake loading. A machine learning-enabled method is developed for predicting the probability of liquefaction triggering and evaluating liquefaction-induced settlement of a 1 m thick layer at depth z (LIS_z) that accounts for these uncertainties. Cone penetration test (CPT) data from nine seismically vulnerable locations in India and scaled accelerograms are used to generate a big dataset. The CPTs are conducted at Mundra (Gujarat, 2 locations), Tezpur (Assam), Madanpur (Delhi), Madhepura (Bihar), Motihari (Bihar), Gorakhpur (Haryana), Rohtak (Haryana) and Polavaram (Andhra Pradesh). The selected locations reflect different earthquake zones per IS 1893 (Part 1 2016).

PDMY03 constitutive model is selected to capture the liquefaction and post-liquefaction behaviour of sands. A numerical calibration study using the cyclic simple shear test in OpenSEES determines the PDMY03 model parameters based on relative density interpreted from CPTs. Subsequently, the numerical soil column with PDMY03 constitutive model is verified against the VELACS experiment number 1.

The CPTs from nine locations are interpreted to quantify the soil's lateral variability in terms of relative density. Based on the relative density variations obtained from CPT, numerical soil columns are modelled in OpenSEES with Lysmer-Kuhlemeyer dashpot at the base. The dashpot accounts for the finite rigidity of the soil below the modelled/explored depth. The numerical soil column is subjected to selected 12 earthquake motions that are scaled to reflect the seismic vulnerability of the location. A fully coupled (u-p) formulation captures the triggering of liquefaction and associated liquefaction-induced settlement. For liquefaction triggering, site response analyses with permeability 10^{-4} m/s (SRA_a) are conducted, and liquefaction is said to have been triggered if the excess pore pressure ratio reaches 0.98 or the single amplitude cyclic shear strain reaches 3 per cent. Similarly, site response analyses with permeability 10^{-2} m/s (SRA_b) are conducted to compute LIS_z.

The big dataset generated from numerical analyses (1,944 analyses) is used in machine learning algorithms to develop predictive models for the probability of liquefaction triggering (classification algorithms) and estimation of LIS_z (regression algorithms). Machine learning algorithms, namely logistic regression (for classification), multiple linear regression (for regression), k-nearest neighbours (for classification and regression), decision tree (for classification and regression), random forest (for classification and regression), support vector machines (for classification and regression) and XGBoost (for classification and regression), are employed, and their efficiency is evaluated using performance scores. The system/influencing variables considered in the current study are relative density at the depth of interest, relative density at the base of the numerical soil column, initial effective vertical stress at the depth of interest, peak acceleration and Arias intensity of the earthquake motion.

It is observed that the XGBoost algorithm performs better than other considered algorithms and is adopted in the current study. Simple user interfaces are developed which predict the probability of liquefaction triggering and LIS_z based on these five system/influencing variables. The developed methodology is validated by plotting the predicted values against the actual values for the test set data.

These developed and calibrated algorithms can be used in professional practice for preliminary assessment before resorting to detailed investigation as long as the relative density and earthquake intensity fall within the range considered in this thesis.

सारांश

भूकंप-प्रेरित कतरनी तनाव के कारण संतृप्त रेत का द्रवीकरण दुनिया भर में इमारतों को गंभीर नुकसान पहुंचाने के लिए जिम्मेदार है। समुद्री, नदी और जलोढ़ निक्षेपों पर बनी संरचनाएँ भूकंप के दौरान द्रवीकरण के खतरों के प्रति संवेदनशील होती हैं। द्रवीकरण के परिणामस्वरूप विनाशकारी विफलताएं हो सकती हैं, जैसा कि अतीत में कुख्यात 1964 निगाटा भूकंप और हाल ही में 2021 असम भारत और 2022 जावा इंडोनेशिया भूकंप के दौरान स्पष्ट है। द्रवीकरण के गंभीर परिणाम के लिए द्रवीकरण की शुरुआत का मूल्यांकन और द्रवीकरण-प्रेरित सेटलमेंट का आकलन करना अभ्यासरत अभियंता के लिए आवश्यक हो जाता है।

द्रवीकरण खतरों के आकलन के लिए पारंपरिक नियतात्मक दृष्टिकोण मिट्टी की पार्श्व परिवर्तनशीलता और भूकंप बल से जुड़ी अनिश्चितताओं को ध्यान में नहीं रखते हैं। गहराई z पर 1 मीटर मोटी परत के द्रवीकरण ट्रिगर की संभावना और द्रवीकरण-प्रेरित सेटलमेंट (LIS_z) का मूल्यांकन करने के लिए एक मशीन लर्निंग-सक्षम विधि विकसित की गई है जो इन अनिश्चितताओं का विचार करती है। भारत में भूकंप की दृष्टि से संवेदनशील नौ स्थानों से शंकु प्रवेश परीक्षण (सीपीटी) डेटा और स्केल किए गए एक्सेलेरोग्राम का उपयोग एक बड़ा डेटासेट तैयार करने के लिए गया है । सीपीटी मुंद्रा (गुजरात, 2 स्थानों), तेजपुर (असम), मदनपुर (दिल्ली), मधेपुरा (बिहार), मोतिहारी (बिहार), गोरखपुर (हरियाणा), रोहतक (हरियाणा) और पोलावरम (आंध्र प्रदेश) में करे गए है । चयनित स्थान आईएस 1893 (भाग 1 2016) के अनुसार विभिन्न भूकंप क्षेत्रों को दर्शाते हैं।

रेत के द्रवीकरण और द्रवीकरण के बाद के व्यवहार को पकड़ने के लिए PDMY03 संवैधानिक मॉडल का चयन किया गया है। सीपीटी के उपयोग से सापेक्ष घनत्व के आधार पर PDMY03 मॉडल पैरामीटर निर्धारित किए गए है। इसके लिए OpenSEES में चक्रीय सरल कतरनी परीक्षण के माध्यम से एक संख्यात्मक अंशांकन अध्ययन किया गया है। इसके बाद, PDMY03 संवैधानिक मॉडल के साथ संख्यात्मक मिट्टी स्तंभ को VELACS प्रयोग संख्या 1 के विरुद्ध सत्यापित किया गया है।

सापेक्ष घनत्व के संदर्भ में मिट्टी की पार्श्व परिवर्तनशीलता को मापने के लिए नौ स्थानों से सीपीटी की व्याख्या की गयी है। सीपीटी से प्राप्त सापेक्ष घनत्व भिन्नताओं के आधार पर, लिस्मर-कुहलेमेयर डैशपॉट के साथ OpenSEES में संख्यात्मक मिट्टी के स्तंभों को मॉडल किया किया गया है। डैशपॉट मॉडल की गई/अन्वेषित गहराई के नीचे की मिट्टी की सीमित कठोरता का हिसाब लगाता है। संख्यात्मक मृदा स्तंभ को चयनित 12 भूकंप गतियों के अधीन किया जाता है जिन्हें चयनित स्थान की भूकंपीय संवेदनशीलता को प्रतिबिंबित करने के लिए मापा जाता है। एक पूरी तरह से युग्मित (u-p) फॉर्मूलेशन द्रवीकरण और संबंधित द्रवीकरण-प्रेरित सेटलमेंट की ट्रिगरिंग को पकड़ता है। द्रवीकरण ट्रिगरिंग के लिए, पारगम्यता 10^{-4} मीटर/सेकेंड (SRA_a) के साथ साइट प्रतिक्रिया विश्लेषण किया जाता है, और कहा जाता है कि यदि अतिरिक्त छिद्र दबाव अनुपात 0.98 तक पहुंच जाता है या एकल आयाम चक्रीय कतरनी तनाव 3 प्रतिशत तक पहुंच जाता है तो द्रवीकरण शुरू हो गया है। इसी तरह, LIS_z की गणना करने के लिए पारगम्यता 10^{-2} m/s (SRA_b) के साथ साइट प्रतिक्रिया विश्लेषण किया जाता है।

संख्यात्मक विश्लेषण (1,944 विश्लेषण) से उत्पन्न बड़े डेटासेट का उपयोग मशीन लर्निंग कलनविधि में द्रवीकरण ट्रिगरिंग (वर्गीकरण कलनविधि) की संभावना और LIS_z (प्रतिगमन कलनविधि) के अनुमान के लिए पूर्वानुमानित मॉडल विकसित करने के लिए किया जाता है। मशीन लर्निंग कलनविधि, अर्थात् लॉजिस्टिक रिग्रेशन (वर्गीकरण के लिए), मल्टीपल लीनियर रिग्रेशन (प्रतिगमन के लिए), के-निकटतम पड़ोसी (वर्गीकरण और प्रतिगमन के लिए), निर्णय वृक्ष (वर्गीकरण और प्रतिगमन के लिए), यादृच्छिक वन (वर्गीकरण और प्रतिगमन के लिए), समर्थन वेक्टर मशीनें (वर्गीकरण और प्रतिगमन के लिए) और XGBoost (वर्गीकरण और प्रतिगमन के लिए) कार्यरत हैं, और उनकी दक्षता का मूल्यांकन प्रदर्शन स्कोर का उपयोग करके किया जाता है। वर्तमान अध्ययन में विचार किए गए सिस्टम/प्रभावी चर हैं: कोई गहराई पर सापेक्ष घनत्व, संख्यात्मक मिट्टी स्तंभ के आधार पर सापेक्ष घनत्व, उसी गहराई पर प्रारंभिक प्रभावी ऊर्ध्वाधर तनाव, भूकंप गति का शिखर त्वरण और एरियस तीव्रता।

यह देखा गया है कि XGBoost कलनविधि अन्य चुने गए कलनविधियो की तुलना में बेहतर प्रदर्शन करता है और वर्तमान अध्ययन में इसे अपनाया गया है। सरल उपयोगकर्ता इंटरफ़ेस विकसित किए गए हैं जो इन पांच सिस्टम/प्रभावी चर के आधार पर द्रवीकरण ट्रिगरिंग और LIS_z की संभावना की पूर्वसूचना देते हैं। विकसित कार्यप्रणाली को परीक्षण सेट डेटा के वास्तविक मूल्यों के विरुद्ध अनुमानित मूल्यों को प्लॉट करके मान्य किया जाता है।

इन विकसित और कैलिब्रेटेड कलनविधि का उपयोग विस्तृत जांच का सहारा लेने से पहले प्रारंभिक मूल्यांकन के लिए पेशेवर अभ्यास में किया जा सकता है, जब तक कि सापेक्ष घनत्व और भूकंप की तीव्रता इस थीसिस में विचार की गई सीमा के भीतर आती है।

TABLE OF CONTENTS

CERTIFICATE .... **i**

Acknowledgements .... **ii**

Abstract .... **iv**

List of Figures .... **xv**

List of Tables .... **xxiii**

List of Notations (English Alphabet) .... **xxv**

List of Notations (Greek Alphabet) .... **xxviii**

Abbreviations .... **xxix**

Chapter 1 Introduction **1**

1.1 Background 1

1.2 Motivation for the Present Study 5

1.2.1 Emerging Trends in Industry 4.0 7

1.3 Objectives 8

1.4 Methodology 8

1.5 Limitations of the proposed approach 9

1.6 Organization of Thesis 11

Chapter 2 Literature Review **13**

2.1 Liquefaction Revisited 13

2.2 Constitutive Modelling 16

2.2.1 PDMY Models 18

2.2.2 Critical State-Based Bounding Surface Models for Sands 19

2.2.3 NorSand Model 23

2.2.4 Comparison of Constitutive Models 24

2.2.5 Performance of Constitutive Models with Centrifuge Tests 30

2.3 State-of-Practice in the Industry 32

2.3.1 Flow Liquefaction 34

2.3.2 Cyclic Liquefaction 35

2.4 Machine Learning in Geotechnical Engineering 46

2.5 Concluding Remarks 49

2.5.1 Deterministic Approach or Probabilistic Approach? 50

Chapter 3 Site Investigations using CPT 52

3.1 Cone Penetration Tests 52

3.1.1 Why Cone Penetration Tests? 53

3.1.2 Corrections 54

3.1.3 Graphical Presentation of Results 55

3.2 Seismic Vulnerability 56

3.2.1 Earthquake Uncertainty 60

Chapter 4 Interpretation of CPT 74

4.1 Some Basic Correlations 74

4.2 Determining Relative Density from CPTs 78

4.3 Lateral Variation in Relative Density 80

4.4 How to Use for Liquefaction Hazard Assessment? 81

Chapter 5 Numerical Modelling: Verification and Generation of Dataset for Machine Learning 91

5.1 Calibration of PDMY03 Model 91

5.2 Modelling Level Ground 94

5.3 Numerical Simulation of VELACS Experiment 96

5.4 Dataset Generation for Machine Learning 100

5.4.1 Dataset Generation for Liquefaction Triggering 101

5.4.2 Dataset Generation for Prediction of LISz 103

5.5 Summary 105

Chapter 6 Machine Learning 107

6.1 Introduction 107

6.2 Machine Learning 107

6.3 Machine Learning Algorithms 108

6.3.1 Logistic Regression 109

6.3.2 Multiple Linear Regression 110

6.3.3 k-Nearest Neighbours Classification and Regression 110

6.3.4 Decision Tree Classification and Regression (DTR) 112

6.3.5 Random Forest Classification and Regression 113

6.3.6 Support Vector Machines Classification and Regression 114

6.3.7 XGBoost Classification and Regression 117

6.4 Sampling Algorithms 119

6.4.1 Stratified k-fold Cross-Validation 119

6.4.2 Repeated k-fold cross-validation 119

6.5 Performance Evaluation for Classification MLA 120

6.5.1 Confusion Matrix 120

6.5.2 Accuracy (ACY) 121

6.5.3 Cohen's Kappa (κ) 121

6.5.4 Precision, Recall and f1 score 121

6.5.5 ROC and AUC 122

6.6 Performance Evaluation for Regression MLA 123

6.6.1 R-squared (Coefficient of determination) 123

6.6.2 Mean squared error (MSE) 124

6.6.3 Explained Variance (EV) 124

Chapter 7 Assessment of Triggering of Liquefaction and LIS_z 125

7.1 Numerical Analysis and Machine Learning 125

7.1.1 Classification Framework 130

7.1.2 Regression Framework 133

7.1.3 Interpretation of Results 136

7.2 Validation of Proposed Method for LISz 143

Chapter 8 Summary and Suggestions for Future Work 145

8.1 Summary 145

8.1.1 Earthquake Uncertainty 146

8.1.2 Soil Lateral Variation 146

8.1.3 Numerical Model and Analyses 147

8.1.4 Implementing Machine Learning 147

8.2 Developed Methodology 148

8.3 Suggested Future Work 149

References 151

Appendix A: Location Plan and CPT Records 169

Appendix B: PDMY03 parameters for all relative densities 255

Appendix C: Statistical distribution of system/influencing variables and results from finite element analyses 261

Appendix D: Results from Classification MLA 283

Appendix E: Results from Regression MLA 292

Appendix F: Sample Calculation for Assessment of Factor of Safety Against Liquefaction based on IS Code (Annexure F, IS 1893, Part 1, 2016) 302

List of Publications 304

Biodata 305

List of Figures

Figure 2.1 Typical response of sands in the laboratory under undrained monotonic triaxial compression. a) Stress path, and b) Stress-strain curve (Modified from Robertson and Wride 1998) 14

Figure 2.2 Concept of cyclic and flow liquefaction 15

Figure 2.3 Description of associative and non-associative flow rule in stress space 17

Figure 2.4 Grain size distribution curve for Yamuna and Nevada sand 29

Figure 2.5 Flow chart for evaluation of liquefaction (Robertson 1994)........ 33

Figure 2.6 Flow chart for problems associated with liquefaction (Ishihara 1993)........ 34

Figure 2.7 Plasticity-based soil classification chart (IS 1498) 35

Figure 2.8 Seismic zones of India per IS1893 (Part 1 2016)........ 37

Figure 2.9 Soils that were observed to be liquefied (Wang 1979) 38

Figure 2.10 Empirical graphs between penetration resistance (or shear wave velocity) and CSR (or CRR) based on case histories (Youd et al. 2001) 40

Figure 3.1 A typical representation of a cone penetrometer (Lunne et al. 2002)........ 55

Figure 3.2 Typical representation of CPT data (CPeT-IT)........ 57

Figure 3.3 Site locations presented on the map of India 58

Figure 3.4 Acceleration time history of earthquakes considered in the current study 62

Figure 3.5 Response spectra of earthquakes considered in the current study 62

Figure 3.6 Representation of assumed normally distributed probability distribution curves for Zone III, Zone IV, and Zone V in the current study 63

Figure 3.7 Measured CPT data, friction ratio and SBT (I_c) for Mundra I Gujarat. 65

Figure 3.8 Measured CPT data, friction ratio and SBT (I_c) for Mundra II Gujarat........ 66

Figure 3.9 Measured CPT data, friction ratio and SBT (I_c) for Tezpur Assam........ 67

Figure 3.10 Measured CPT data, friction ratio and SBT (I_c) for Madanpur Delhi........ 68

Figure 3.11 Measured CPT data, friction ratio and SBT (I_c) for Madhepura Bihar. 69
Figure 3.12 Measured CPT data, friction ratio and SBT (I_c) for Motihari Bihar. 70
Figure 3.13 Measured CPT data, friction ratio and SBT (I_c) for Gorakhpur Haryana. 71
Figure 3.14 Measured CPT data, friction ratio and SBT (I_c) for Rohtak Haryana. 72
Figure 3.15 Measured CPT data, friction ratio and SBT (I_c) for Polavaram Andhra Pradesh. 73
Figure 4.1 Normalised CPT soil behaviour type charts $Q_{tn} – F_r$ and $Q_t – B_q$ (Robertson 1990, updated by Robertson 2010) 76
Figure 4.2 Flowchart for interpretation of clean sand equivalent cone resistance, q_{c1Ncs} (Youd et al. 2001 and Robertson and Cabal 2022)................................... 78
Figure 4.3 Generated relative density variations at Mundra I Gujarat. 82
Figure 4.4 Generated relative density variations at Mundra II Gujarat................................ 83
Figure 4.5 Generated relative density variations at Tezpur Assam...................................... 84
Figure 4.6 Generated relative density variations at Madanpur Delhi................................... 85
Figure 4.7 Generated relative density variations at Madhepura Bihar................................. 86
Figure 4.8 Generated relative density variations at Motihari Bihar 87
Figure 4.9 Generated relative density variations at Gorakhpur Haryana 88
Figure 4.10 Generated relative density variations at Rohtak Haryana.................................. 89
Figure 4.11 Generated relative density variations at Polavaram Andhra Pradesh 90
Figure 5.1 Verification and calibration of the PDMY03 model used in the current study 92
Figure 5.2 Nine Four Node Quad u-p element (Mazzoni et al. 2006) having translational degrees of freedom for all nine nodes and pore pressure degree of freedom for four corner-hatched nodes. 93

Figure 5.3 A conceptual model for single element cyclic simple shear test employed for calibration of PDMY03 material parameters................................93
Figure 5.4 Numerical model employed in the current study with Lysmer-Kuhlemeyer dashpot 95
Figure 5.5 a) Conceptual layout of VELACS No. 1 experiment (Taboada 1995); b) Applied acceleration time history at the base 97
Figure 5.6 Comparison of excess pore water pressure recorded in VELACS #1 experiment with numerical simulation...99
Figure 5.7 Comparison of vertical settlement recorded at LVDT1 in VELACS #1 experiment with numerical simulation on a prototype scale.............................100
Figure 5.8 Excess pore pressure ratio response with time for a) Darfield and b) Chi Chi earthquake at different depths (Tezpur Assam) 102
Figure 5.9 Depth-wise maximum excess pore pressure ratio response for different numerical soil columns subjected to the Darfield and the Chi Chi earthquake (Tezpur Assam) 103
Figure 5.10 Result from finite element analysis considering a) Darfield earthquake and b) Chi-Chi earthquake for different numerical soil columns 104
Figure 5.11 Flow chart for numerical analyses 105
Figure 6.1 A typical visualization of the k-NN machine learning algorithm (Taunk et al. 2019)...111
Figure 6.2 A typical visualization of a decision tree (Awad and Khanna 2015)................113
Figure 6.3 Conceptual flowchart for random forest classification model (Pant and Ramana 2022) 114
Figure 6.4 A typical representation of one-dimensional SVR (Awad and Khanna 2015) 116

Figure 6.5 Types of the loss functions in SVR: (a) Linear, (b) quadratic, and (c) Huber. Loss is the penalty imposed on the predictions farther away from the desired output (Awad and Khanna 2015).......117

Figure 6.6 ROC_AUC for an ideal curve with AUC = 1.0.......122

Figure 7.1 Typical output from finite element analysis for use in machine learning (Boolean variable 0: Non-Liquefiable, 1 Liquefiable).......127

Figure 7.2 Typical output from finite element analysis for use in machine learning (for regression of LIS_z).......128

Figure 7.3 Statistical distribution of system/influencing variables for machine learning (Unified Dataset).......129

Figure 7.4 Statistical distribution of results from FEM (Unified Dataset).......130

Figure 7.5 Variation of performance metrics for XGBoost classification of liquefaction triggering (Unified dataset).......131

Figure 7.6 Visual representation of stratified k-fold cross validation.......132

Figure 7.7 Prediction of triggering of liquefaction based on the numerical analyses.......133

Figure 7.8 Variation of performance metrics for XGBoost regression of LIS_z (Unified dataset).......135

Figure 7.9 Prediction of LIS_z based on the result of numerical analyses.......135

Figure 7.10 Heatmap for system/influencing variables (Unified dataset).......141

Figure 7.11 User interface for predicting the probability of liquefaction triggering: a) Input Frame, and b) Disclaimer.......142

Figure 7.12 User interface for prediction of LIS_z.......143

Figure 7.13 Plot of predicted vs. actual settlement values for validation of machine learning algorithm (XGBoost).......144

Figure A.1 CPT location plan for Mundra I Gujarat.......170

Figure A.2 Measured CPT data, friction ratio and SBT (I_c) at Mundra I Gujarat 184

Figure A.3 CPT location plan for Mundra II Gujarat 185

Figure A.4 Measured CPT data, friction ratio and SBT (I_c) at Mundra II Gujarat 189

Figure A.5 CPT location plan for Tezpur Assam 190

Figure A.6 Measured CPT data, friction ratio and SBT (I_c) at Tezpur Assam 194

Figure A.7 CPT location plan for Madanpur Delhi 195

Figure A.8 Measured CPT data, friction ratio and SBT (I_c) at Madanpur Delhi 199

Figure A.9 CPT location plan for Madhepura Bihar 200

Figure A.10 Measured CPT data, friction ratio and SBT (I_c) at Madhepura Bihar 215

Figure A.11 CPT location plan for Motihari Bihar 216

Figure A.12 Measured CPT data, friction ratio and SBT (I_c) at Motihari Bihar 223

Figure A.13 CPT location plan for Gorakhpur Haryana 224

Figure A.14 Measured CPT data, friction ratio and SBT (I_c) at Gorakhpur Haryana 235

Figure A.15 CPT location plan for Rohtak Haryana 236

Figure A.16 Measured CPT data, friction ratio and SBT (I_c) at Rohtak Haryana 242

Figure A.17 CPT location plan for Polavaram Andhra Pradesh 243

Figure A.18 Measured CPT data, friction ratio and SBT (I_c) at Polavaram Andhra Pradesh 254

Figure B.1 Representative results from single-element cyclic simple shear tests for 57 % relative density sand and stress ratio of 0.13 (Section 5.1) 260

Figure C.1 System/influencing variables and results from numerical analyses (Mundra I) 262

Figure C.2 Heatmap of correlation coefficients (Mundra I Gujarat) 263

Figure C.3 System/influencing variables and results from numerical analyses (Mundra II Gujarat) 264

Figure C.4 Heatmap of correlation coefficients (Mundra II Gujarat) 265

Figure C.5 System/influencing variables and results from numerical analyses (Tezpur Assam) 266

Figure C.6 Heatmap of correlation coefficients (Tezpur Assam) 267

Figure C.7 System/influencing variables and results from numerical analyses (Madanpur Delhi) 268

Figure C.8 Heatmap of correlation coefficients (Madanpur Delhi) 269

Figure C.9 System/influencing variables and results from numerical analyses (Madhepura Bihar) 270

Figure C.10 Heatmap of correlation coefficients (Madhepura Bihar) 271

Figure C.11 System/influencing variables and results from numerical analyses (Motihari Bihar) 272

Figure C.12 Heatmap of correlation coefficients (Motihari Bihar) 273

Figure C.13 System input variables and results from numerical analyses (Gorakhpur Haryana) 274

Figure C.14 Heatmap of correlation coefficients (Gorakhpur Haryana) 275

Figure C.15 System/influencing variables and results from numerical analyses (Rohtak Haryana) 276

Figure C.16 Heatmap of correlation coefficients (Rohtak Haryana) 277

Figure C.17 System/influencing variables and results from numerical analyses (Polavaram Andhra Pradesh) 278

Figure C.18 Heatmap of correlation coefficients (Polavaram Andhra Pradesh) 279

Figure C.19 Depth-wise maximum excess pore pressure ratio response for different numerical soil columns subjected to different earthquakes (Tezpur Assam) 282

Figure D.1 Variation of performance metrics for XGBoost classification of liquefaction triggering (Mundra I Gujarat) 284

Figure D.2 Variation of performance metrics for XGBoost classification of liquefaction triggering (Mundra II Gujarat)..285

Figure D.3 Variation of performance metrics for XGBoost classification of liquefaction triggering (Tezpur Assam)..286

Figure D.4 Variation of performance metrics for XGBoost classification of liquefaction triggering (Madanpur Delhi)...287

Figure D.5 Variation of performance metrics for XGBoost classification of liquefaction triggering (Madhepura Bihar) 288

Figure D.6 Variation of performance metrics for XGBoost classification of liquefaction triggering (Motihari Bihar) 289

Figure D.7 Variation of performance metrics for XGBoost classification of liquefaction triggering (Gorakhpur Haryana) 290

Figure D.8 Variation of performance metrics for XGBoost classification of liquefaction triggering (Rohtak Haryana)...291

Figure E.1 Variation of performance metrics for XGBoost regression of LIS_z (Mundra I Gujarat) 293

Figure E.2 Variation of performance metrics for XGBoost regression of LIS_z (Mundra II Gujarat)...294

Figure E.3 Variation of performance metrics for XGBoost regression of LIS_z (Tezpur Assam)..295

Figure E.4 Variation of performance metrics for XGBoost regression of LIS_z (Madanpur Delhi)..296

Figure E.5 Variation of performance metrics for XGBoost regression of LIS_z (Madhepura Bihar) 297
Figure E.6 Variation of performance metrics for XGBoost regression of LIS_z (Motihari Bihar) 298
Figure E.7 Variation of performance metrics for XGBoost regression of LIS_z (Gorakhpur Haryana) 299
Figure E.8 Variation of performance metrics for XGBoost regression of LIS_z (Rohtak Haryana) 300
Figure E.9 Variation of performance metrics for XGBoost regression of LIS_z (Polavaram Andhra Pradesh) 301

List of Tables

Table 1.1 Flow liquefaction: some case histories 2

Table 1.2 Cyclic liquefaction: some case histories 4

Table 2.1 Features of some constitutive models 18

Table 2.2 Comparison of e_{max}, e_{min}, specific gravity and uniformity coefficient (c_u) for Yamuna sand and Nevada sand 29

Table 2.3 Proposed liquefaction susceptibility criteria for fine-grained soils by Andrews and Martin (2000) 38

Table 2.4 A summary of methods for computation of liquefaction-induced settlement 45

Table 3.1 Summary of CPTs conducted 59

Table 3.2 Summary of earthquake records used in numerical analysis 61

Table 3.3 Summary of mean peak acceleration to reflect seismic vulnerability 63

Table 3.4 Summary of earthquake records used in numerical simulations for different zones 64

Table 4.1 Applicability of CPT/CPTU for parameter estimation (Robertson and Cabal 2022) 74

Table 4.2 Empirical equations for estimation of various soil parameters from CPT data 77

Table 4.3 Correlations used for interpretation of relative density from CPT data 79

Table 4.4 Correlations for interpretation of relative density from SPT data 79

Table 5.1 Summary of numerical analyses conducted 106

Table 6.1 A general representation of two class confusion matrix 121

Table 7.1 Performance of different MLA employed (Unified Data) 132

Table 7.2 Performance of different regression MLA employed (Unified dataset) 134

Table B.1 Result from the calibration study for the PDMY03 constitutive model for different relative densities 258
Table D.1 Performance of different classification MLA (Mundra I Gujarat) 284
Table D.2 Performance of different classification MLA (Mundra II Gujarat) 285
Table D.3 Performance of different classification MLA (Tezpur Assam) 286
Table D.4 Performance of different classification MLA (Madanpur Delhi) 287
Table D.5 Performance of different classification MLA (Madhepura Bihar) 288
Table D.6 Performance of different classification MLA (Motihari Bihar) 289
Table D.7 Performance of different classification MLA (Gorakhpur Haryana) 290
Table D.8 Performance of different classification MLA (Rohtak Haryana) 291
Table E.1 Performance of different regression MLA (Mundra I Gujarat) 293
Table E.2 Performance of different regression MLA (Mundra II Gujarat) 294
Table E.3 Performance of different regression MLA (Tezpur Assam) 295
Table E.4 Performance of different regression MLA (Madanpur Delhi) 296
Table E.5 Performance of different regression MLA (Madhepura Bihar) 297
Table E.6 Performance of different regression MLA (Motihari Bihar) 298
Table E.7 Performance of different regression MLA (Gorakhpur Haryana) 299
Table E.8 Performance of different regression MLA (Rohtak Haryana) 300
Table E.9 Performance of different regression MLA (Polavaram Andhra Pradesh) 301

List of Notations (English Alphabet)

$(N_1)_{60}$	SPT N value corrected for 1 atm pressure and 60 % energy ratio.
a	Net area ratio of the cone in CPT
A_c	Cross-section area of the cone in CPT
a_{max}	Peak acceleration of the earthquake motion
A_n	The cross-section area of the cone load cell behind the cone in CPT
A_s	The surface area of the sleeve in CPT
A_{sb}	The cross-section area of the sleeve at the base in CPT
A_{st}	The cross-section area of the sleeve at the top in CPT
B	Bulk modulus of soil
B_q	Normalised excess pore pressure.
c_h	Coefficient of consolidation
C_N	Normalizing factor for cone resistance
D_{50}	The grain size, D_{50}, is the size for which 50% of the particle mass consists of finer particles.
D_r	Relative density of soil
e	Void ratio of soil
E	Young's modulus of soil
e_0	Initial void ratio
e_{cs}	Critical state void ratio
e_{max}	maximum void ratio of soil
e_{min}	minimum void ratio of soil
$f1$	Performance score for classification MLA
F_r	Normalised friction ratio.
f_s	Measured sleeve friction in CPT
f_t	Corrected sleeve friction in CPT
G	Shear modulus of soil
g	The acceleration due to gravity (9.8 m/s^2).
G_0	Small strain shear modulus
I_1	First invariant of stress tensor (sum of principal stresses)

I_a	Arias intensity
I_c	Soil behaviour type index
I_p	Plasticity index of soil
J_2	Second invariant of deviatoric stress tensor (sum of the product of deviatoric stresses in principal stress system taken two at a time)
k	The coefficient of permeability of the soil
k'	number of folds in k'-fold cross-validation
K_0	In-situ stress ratio
K_c	Correction factor for grain characteristics in CPT interpretation
k-NN	k-nearest neighbours
M	Earthquake magnitude on the Richter scale
M_f	Failure surface in PDMY models
M_s	Surface wave magnitude of the earthquake
M_v	The coefficient of compressibility
M_w	Moment magnitude of the earthquake.
N	Number of blow counts measured during SPT
n	Stress exponent factor in CPT interpretation
N_{60}	SPT N corrected for a 60 % energy ratio
N_{78}	SPT N value corrected for 78 % energy ratio
N_{kt}	Factor evaluated as (10.5+ 7 log(F_r) for interpretation of undrained shear strength in CPT
OCR	Over Consolidation Ratio
OH	High plasticity organic soil
OI	Intermediate plasticity organic soil
OL	Low plasticity organic soil
P	Precision (performance score for classificaiotn MLA)
p'	Mean effective stress
p_a	Atmospheric pressure (100 kPa or 1 atm).
$P_L (X)$	Probability of triggering of liquefaction
Q	Cone resistance normalised by atmospheric pressure
q	Deviatoric stress

q_c	Measured cone resistance in CPT
q_{c1Ncs}	Clean sand (<5 % FC) equivalent cone resistance that would have been measured if the soil was clean sand under vertical stress of 1 atm
$Q_L (X)$	Transformation of $P_L (X)$ using the logit function
Q_p	Cone resistance normalised by mean effective stress
Q_t	Cone resistance normalised by vertical effective stress
q_t	Corrected cone resistance
R	Recall (performance score for classificaiotn MLA)
R^2	R-squared (Performance score for regression MLA)
R_f	Friction ratio
r_s	Spearman's correlation coefficient
S_t	Sensitivity
s_u	Undrained shear strength of soil
u	Pore water pressure
u_0	Ambient pore water pressure around the cone (generally hydrostatic)
u_2	Measured pore water pressure behind the cone in CPT
u_3	pore water pressure at the top of the sleeve in CPT
V_s	Shear wave velocity
V_{s1}	Shear wave velocity normalised with vertical stress of 1 atm
z	Depth of interest

List of Notations (Greek Alphabet)

α	Factor to convert q_t to N_{60} ($N_{60} = q_t/\alpha$)
α_{vs}	Factor based on soil behaviour type index (Ic) used for interpretation of shear wave velocity
$\beta'_0, \beta'_1, \beta'_2, \ldots\beta'_n$	Regression coefficient for MLR
$\beta_0, \beta_1, \beta_2, \ldots\beta_n$	Regression coefficient for LR
γ_w	Unit weight of water
γ_{xy}	Shear strain
$d\varepsilon^p$	Incremental plastic strain tensor
$d\sigma$	Incremental stress tensor
ε	Support vector/Allowable error in SVM
κ	Cohen's kappa (performance score for classification MLA)
μ	Mean of the normally distributed relative densities for 1 m thick layers
ν	Poisson's ratio of soil
ξ_i^*	Distance of datapoints from ε-sensitive tube
σ	Standard deviation of the normally distributed relative densities for 1 m thick layers
$\boldsymbol{\sigma}$	Stress tensor (**bold** face)
σ_1	Major principal stress
σ_2	Minor principal stress (in-plane strain)
σ_v'	Current effective vertical stress.
σ_{vo}	Total vertical stress.
σ_{vo}'	Initial effective vertical stress.
ψ	State parameter ($\psi = e_0 - e_{cs}$) of soil
ϕ	Friction angle of soil
ϕ'	Effective peak friction angle of soil

Abbreviations

ACY	Accuracy (performance score for classification MLA)
AH	Accelerometer measuring in the horizontal direction
AI	Artificial Intelligence
ANN	Artificial Neural Network
ASTM	American Society for Testing and Materials
AV	Accelerometer measuring in the vertical direction
BI	Business Intelligence
Boolean variable 0	Liquefaction is not triggered
Boolean variable 1	Liquefaction is triggered
CanRF	Canonical Random Forest
CH	High plasticity clay
CI	Intermediate plasticity clay
CL	Low plasticity clay
CPT	Cone Penetration Test
CPTU	Cone Penetration Test with pore water pressure measurements
CRR	Cyclic resistance ratio
CSL	Critical State Line
CSR	Cyclic stress ratio
CSSM	Critical State Soil Mechanics
DNN	Deep Neural Network
D_r_base	Relative density of soil at the base of the numerical soil column
DTR	Decision Tree Regression
EPPR	Excess Pore Pressure Ratio
EPWP	Excess Pore Water Pressure
EV	Explained Variance (Performance score for regression MLA)
FC	Fines Content (percentage of particles finer than 75 microns)
FN	False Negative
FP	False Positive
FPR	False Positive Rate

GSD	Grain Size Distribution
GWT	Ground Water Table
IoT	Internet of Things
IRC	Indian Road Congress code of practice
IS Code	Indian Standard Code of Practice
ISSMGE	International Society for Soil Mechanics and Geotechnical Engineering
LIS	Liquefaction-induced settlement. In the context of current research work, LIS is the vertical settlement of soil measured either at the surface (LIS_0) or for any 1 m thick layer at depth *z* (LIS_z) due to liquefaction.
LIS_0	Liquefaction-induced settlement at the surface. Summed-up result of LIS_z for considered depths.
LIS_z	Liquefaction Induced Settlement for 1 m thick layer at depth *z*.
LL	Liquid limit of soil
LR	Logistic Regression
LSS	Limited Strain Softening soil
LVDT	Linear Variable Differential Transducer
MASW	Multichannel Analysis of Surface Waves
MH	High plasticity silt
MI	Intermediate plasticity silt
ML	Low plasticity silt
ML	Machine Learning
MLA	Machine Learning Algorithms
MLR	Multiple Linear Regression
MSE	Mean Squared Error (Performance score for regression MLA)
NCEER	National Centre for Earthquake Engineering and Research
NCL	Normal Compression Line
NSF	National Science Foundation
PDMY/02/03	Pressure Dependent Multi-Yield surface soil constitutive models and their subsequent developments
PEER-NGA	Pacific Earthquake Engineering Research Centre - Next Generation Attenuation

PGA	Peak Ground Acceleration of the earthquake motion (in units of *g*).
PI	Plasticity Index of soil
PSO-KELM	Particle Swarm Optimization coupled with an extreme Kernel Learning Machine
RF	Random Forest
ROC_AUC	Receiver Operating Characteristics Area Under the Curve
RotRF	Rotational Random Forest
SBT	Soil Behaviour Type
SCPT	Seismic Cone Penetration Test
SH	Strain Hardening soil
SPT	Standard Penetration Test
SRA_a	Site response analysis with permeability 10^{-4} m/s
SRA_b	Site response analysis with permeability 10^{-2} m/s
SS	Strain Softening soil
SSL	Steady State Line
Stress ratio	Principal shear stress (σ_{12}) divided by initial vertical effective stress (σ_{vo}')
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
TPR	True Positive Rate
Unified dataset	The combined dataset from nine project sites considered in the current study
USCS	Unified Soil Classification System
VELACS	Verification of Liquefaction Analysis by Centrifuge Studies
XGBoost	eXtreme Gradient Boosting MLA