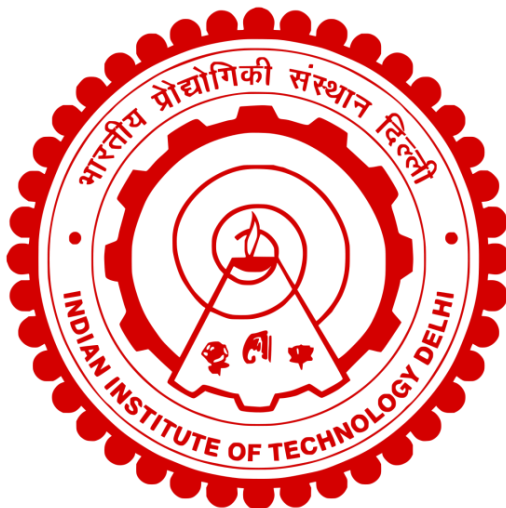


DATA-DRIVEN MODELING AND PHYSICS-INFORMED MACHINE LEARNING FOR GLASS DISCOVERY

RAVINDER



**DEPARTMENT OF CIVIL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI
JANUARY 2023**

© Indian Institute of Technology (IITD), New Delhi, 2023

Data-driven Modeling and Physics-informed Machine Learning for Glass Discovery

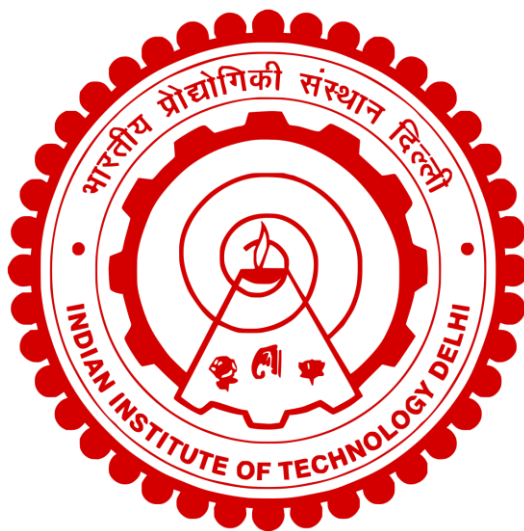
by

RAVINDER
Department of Civil Engineering

Submitted

in fulfillment of requirements of degree of *Doctor of Philosophy*

to the



Indian Institute of Technology Delhi
January 2023

Dedicated to my family

Certificate

This is to certify that the thesis entitled *Data-driven Modeling and Physics-informed Machine Learning for Glass Discovery* is being submitted by Mr. Ravinder to the Indian Institute of Technology Delhi for the award of the degree of Doctor of Philosophy. This is a record of the research work and is entirely carried out by him under my supervision and guidance. The research report presented in this thesis has not been submitted for the award of any other degree or diploma.



Prof. N. M. Anoop Krishnan

Department of Civil Engineering
Indian Institute of Technology Delhi

Declaration

I hereby declare that the thesis contains the work carried out in fulfillment of the requirements for the award of the degree of Doctor of Philosophy under the guidance and supervision of Prof. N. M. Anoop Krishnan, Associate Professor, Department of Civil Engineering, IIT Delhi. The contents of the dissertation, including text, tables, figures, etc., have not been reproduced from other reports, theses, etc. Wherever reproduction has been made from books, journals, reports, manuals, websites, etc., the sources have been duly acknowledged and referenced appropriately.

Ravinder

RAVINDER
2017CEZ7518

Acknowledgements

First and foremost, I would like to express my great gratitude to my parents for their unconditional love, encouragement, and support throughout my life. I am also thankful to my sister and wife to give me stability, affection and unquestionable support.

It is a pleasure to express sincere thanks to my supervisor, Prof. N. M. Anoop Krishnan, for his constant support, availability, dedication, patience, and passion for research. Prof. N. M. Anoop Krishnan has always been very keen on finding new opportunities for me to explore new training sessions, workshops, seminars, and conferences to improve my skills and knowledge. He allowed me to enjoy my academic freedom to the fullest and taught me how to read, write, think, and present confidently in academic research. I would like to thank my research collaborators especially Prof. Sayan Ranu. I would like to express my gratitude to my research committee: internal expert Prof. D. R. Sahoo, external expert Prof. Ajeet Kumar, and chairman Prof. B. R. Chahar for their valuable input at different stages of my research work. I also acknowledge the Civil Engineering Department and Ministry of Education (Prime Minister Research Fellowship) for the financial support for lab equipment, workshops, and conferences. I also acknowledge the computational resources and support provided by HPC, IIT Delhi. I would like to thank the members of the M3RG lab, especially Sir Rajesh, Tanu, and Suresh, for their love and support. Finally, I thank the almighty, who has granted me countless blessings, opportunities, knowledge, and energy to write this thesis.

Abstract

The progress of human civilizations has been dependent on material discovery throughout history. With the discovery of iron, it replaced most farming tools such as sickle, plow tips, etc., which were earlier made from bronze and stones, leading to efficient farming. Also, the early use of iron in weapons made some civilizations dominant in warfare. Similarly, glasses have been significantly affecting human civilizations since their first use as an ornament. Today, inorganic glasses are being used in numerous industries as they provide unique properties with an astronomical number of possible glass compositions (*glasses* and *inorganic glasses* are being used interchangeably throughout the whole work). Glasses find their applications in computer screens, optical lenses, bio-active implants, optical fiber, the automotive industry, nuclear waste management, laboratory glassware, electrolytes, etc. The glass composition is not fixed, unlike crystalline materials, and can be formed in almost any stoichiometry. The glass property can significantly change with a slight change in glass composition or the addition and removal of glass components. Hence, finding an optimum glass composition for a targeted property is a challenging task. A naive trial and error method or heuristic approach with experiments can get easily overwhelmed with the number of possibilities of a glass composition. Therefore, we use machine learning (ML) methods to model the composition–property relationship for glasses. These relationships can be used with optimization algorithms to find the optimum glass composition given some constraints.

The research work use machine learning methods to model numerous glass properties (25 glass properties, namely (a) Abbe number, (b) annealing point, (c) bulk modulus, (d) compressive strength, (e) crystallization temperature, (f) density, (g) dielectric constant, (h) electrical conductivity, (i) Vickers hardness, (j) refractive index, (k) shear modulus, (l) softening point, (m) strain point, (n) liquidus temperature, (o) Littleton temperature, (p)

thermal expansion coefficient (TEC), (q) tensile strength, (r) glass transition temperature, (s) thermal conductivity, (t) thermal shock resistance, (u) infrared (IR) transmittance, (v) solar transmittance, (w) visible transmittance, (x) ultraviolet (UV) transmittance, and (y) Young's modulus). The research work introduces glass selection charts which can be used to identify optimum glass compositions by optimizing two-glass properties simultaneously.

Further, these ML models are then used with explainable ML techniques (SHAP) to understand the compositional control of particular glass properties. The research captures the already known composition–property relationships such as Lowenstein's rule, boron anomaly, etc. It also elucidates the known phenomenon, such as the increase in thermal expansion coefficient for low network-former glasses, decrease in liquidus temperature for high network-modifier glass, etc. Finally, the models were tested on a new glass family $0.5\text{P}_2\text{O}_5-(1-x)\text{CaO}-x\text{MgO}$, for which ML models predict an accurate trend for density, hardness, and refractive index. The effect of MgO mol% is captured by SHAP values elucidating the effectiveness of SHAP analysis in understanding the composition–property relationship.

The composition–property relationships in inorganic glasses are complex, which ML methods can capture and explain with significant accuracy. The mechanisms behind these relationships can not be captured with black-box ML models, and therefore molecular dynamics or DFT simulations are needed to reveal these mechanisms. The DFT simulations are costly, and interatomic potential for MD simulations is not always available. Therefore, the research introduces two frameworks, namely MCLNN, and LGNN, which can be used to approximate and machine learning interatomic potentials from DFT simulation trajectories. MCLNN can learn simple systems with pairwise interactions such as LJ systems, harmonic spring systems, etc., while LGNN can learn systems with complex interactions such as liquid silica systems. LGNN is able to capture structural features such as pair distribution function, bond angle distribution, etc., with high accuracy. The MD simulations can only simulate systems with nanometer size therefore, an upscaling of material property is required to

simulate the system at mesoscale (such as peridynamics simulations) with a large system size. The research introduces a software package, PeriDyn, which provides an interface to write arbitrary peridynamics workflow with machine learning frameworks. The package can be used to learn the peridynamics material model parameters from MD trajectories to upscaling the material property to mesoscale.

Overall, the research explores the use of machine learning methods in modeling composition–property relationships for inorganic glasses, explaining composition dependence of inorganic glass properties, understanding mechanisms with MD simulations, and up-scaling of material response to mesoscale using MD or DFT simulation trajectories.

Keywords: Inorganic glass, glass, machine learning, explainable ML, PINN, Physics-informed ML, ML interatomic potential, upscaling material model, peridynamics

सार

मानव सभ्यताओं की प्रगति पूरे इतिहास में नए पदार्थ की खोज पर निर्भर रही है। लोहे की खोज के साथ अधिकांश खेती के औजार जैसे हंसिया, हल की नोक आदि बदल गए, जो पहले कांसे और पत्थरों से बनाए जाते थे, जिससे कुशल खेती हुई। हथियारों में लोहे के शुरुआती उपयोग ने कुछ सभ्यताओं को युद्ध में प्रभावी बना दिया जिससे लोहे को प्राप्त करना आवश्यक हो गया। इसी तरह, शीशा मानव सभ्यताओं को एक आभूषण के रूप में पहली बार उपयोग करने के बाद से महत्वपूर्ण रूप से प्रभावित कर रहा है। आज, कई उद्योगों में शीशे का उपयोग किया जा रहा है क्योंकि शीशा अशंख्य संरचनाओं के साथ प्राप्त किया जा सकता है एवं हर संरचना अलग गुण प्रदान करती है। शीशा कंप्यूटर स्क्रीन, ऑप्टिकल लेंस, बायो-एक्टिव इम्प्लांट्स, ऑप्टिकल फाइबर, ऑटोमोटिव उद्योग, परमाणु अपशिष्ट प्रबंधन, प्रयोगशाला परखनली, इलेक्ट्रोलाइट्स इत्यादि में प्रयोग किया जाता है। शीशे की संरचना तय नहीं होती है जिससे यह लगभग किसी भी अनुपात में बनाया जा सकता है। शीशे की संरचना में मामूली बदलाव या शीशे के घटकों को जोड़ने और हटाने पर शीशे के गुण महत्वपूर्ण रूप से बदल सकते हैं। इसलिए, लक्षित गुण के लिए शीशे की इष्टतम संरचना खोजना एक चुनौतीपूर्ण कार्य है। शीशे की इष्टतम संरचना खोजने के लिए एक सहज ट्रायल-एंड-एरर विधि के साथ प्रयोग शीशे की संरचना की संभावनाओं की संख्या से आसानी से अभिभूत हो सकते हैं। इसलिए, हम शीशे के लिए रचना-गुण-संबंध को मॉडल करने के लिए मशीन लर्निंग विधियों का उपयोग करते हैं। शीशे की इष्टतम संरचना को खोजने के लिए इन मॉडल्स का उपयोग एक अनुकूलन एल्गोरिदम के साथ उपयोग किया जा सकता है।

अनुसंधान कार्य कई शीशे के गुणों (25 शीशे के गुणों (a) एब्बे नंबर, (b) एनीलिंग पॉइंट, (c) बल्क मॉड्यूलस, (d) कंप्रेसिव स्ट्रेंथ, (e) क्रिस्टलिजेशन तापमान, (f) घनत्व, (g) परावैद्युत स्थिरांक, (h) विद्युत चालकता, (i) विकर्स हार्डनेस, (j) अपवर्तक सूचकांक, (k) शियर मॉड्यूलस, (l) सॉफ्टनिंग पॉइंट, (m) स्ट्रेन पॉइंट, (n) लिक्विडस तापमान, (o) लिटलटन तापमान, (p) ताप विस्तार गुणांक (TEC), (q) टेंसिल स्ट्रेंथ, (r) ग्लास ट्रांजीशन तापमान, (s) थर्मल चालकता, (t) थर्मल शॉक प्रतिरोध, (u) अवरोध (IR) संप्रेषण, (v) सौर संप्रेषण, (w) विज़िबल संप्रेषण, (x) पराबैंगनी (UV) संप्रेषण, और (y) यंग का मापांक) को मॉडल करने के लिए मशीन लर्निंग के तरीकों का उपयोग करता है। यह शोध कार्य शीशा-चयन-चार्ट पेश करता है जिसका उपयोग शीशे के दो गुणों को एक साथ अनुकूलित करके इष्टतम शीशे की पहचान करने के लिए किया जा सकता है।

इसके अलावा, शीशे के विशेष गुणों के संरचनागत नियंत्रण को समझने के लिए इन मशीन लर्निंग मॉडलों का उपयोग SHAP जैसे व्याख्यात्मक मशीन लर्निंग तकनीकों के साथ किया जाता है। अनुसंधान पहले से ही ज्ञात संरचना-गुण संबंधों जैसे की लोवेस्टीन के नियम, बोरोन विसंगति आदि को पकड़ने में सक्षम है। यह ज्ञात नियमों को भी प्रकट करता है, उधारणतः कम नेटवर्क-फॉर्मर के होने पर ताप विस्तार गुणांक में वृद्धि, उच्च नेटवर्क-मॉडिफ़िएर के लिए लिक्विडस तापमान में कमी आदि। अंततः, मशीन लर्निंग मॉडलों का परीक्षण एक नए शीशे ($0.5\text{P}_2\text{O}_5-(1-x)\text{CaO}-x\text{MgO}$) पर किया गया, जिसके लिए मशीन लर्निंग मॉडल्स घनत्व, कठोरता और अपवर्तक सूचकांक के लिए सटीक अनुमान लगाते हैं। इसमें MgO का प्रभाव SHAP मूल्यों द्वारा सटीक दर्शाया जाता है, जो संरचना-गुण सम्बन्ध को समझने में SHAP विश्लेषण की प्रभावशीलता को स्पष्ट करता है।

शीशे में संरचना-गुण सम्बन्ध जटिल होता है, जिन्हें मशीन लर्निंग विधियां सटीकता के साथ पकड़ और समझ सकती हैं। इन संबंधों के पीछे के तंत्र को ब्लैक-बॉक्स मशीन लर्निंग विधियों के साथ नहीं समझा जा सकता है। इसलिए इन तंत्रों को समझने के लिए MD सिमुलेशन या DFT सिमुलेशन की आवश्यकता होती है। DFT सिमुलेशन काफी समय लेते हैं और MD सिमुलेशन के लिए इंटरटोमिक पोटेंशियल हमेशा उपलब्ध नहीं होता है। इसलिए, अनुसंधान दो रूपरेखाओं को पेश करता है, MCLNN एवं LGNN, जिसका उपयोग DFT सिमुलेशन प्रक्षेपवक्र से मशीन लर्निंग इंटरटोमिक पोटेंशियल को अनुमानित करने के लिए किया जा सकता है। MCLNN सरल प्रणालियों जैसे युग्मानुसार स्थितिज ऊर्जा, LJ प्रणाली, हार्मोनिक स्प्रिंग प्रणाली आदि सीख सकता है, जबकि

LGNN तरल सिलिका प्रणाली जैसे जटिल इंटरैक्शन वाली प्रणाली सीख सकता है। LGNN संरचनात्मक विशेषताएं जैसे युग्म बंटन-फलन, कोण बंटन-फलन इत्यादि को उच्च सटीकता के साथ प्राप्त करने में सक्षम है। MD सिमुलेशन केवल नैनोमीटर आकार की प्रणालियों का ही अनुकरण कर सकते हैं, इसलिए बड़े आकार (मेसोस्केल) की प्रणालियों (उधारणतः peridynamics सिमुलेशन) का अनुकरण करने के लिए पदार्थ की गुणता की अपस्केलिंग की आवश्यकता होती है। अनुसंधान एक सॉफ्टवेयर पैकेज, PeriDyn पेश करता है, जो मशीन लर्निंग फ्रेमवर्क के साथ मनमानी कार्य प्रणाली लिखने के लिए एक इंटरफेस प्रदान करता है। सॉफ्टवेयर पैकेज का उपयोग एमडी प्रक्षेपवक्र से peridynamics मटेरियल मोडल मापदंडों को सीखने के लिए किया जा सकता है, जिससे पदार्थ की गुणता की अपस्केलिंग की जा सकती है।

कुल मिलाकर, अनुसंधान शीशे के संरचना-गुण संबंध, शीशे के गुणों की संरचना निर्भरता की व्याख्या करना, MD सिमुलेशन से तंत्र को समझना और पदार्थ के गुणों की अपस्केलिंग करने इत्यादि में मशीन लर्निंग विधियों के उपयोग की पड़ताल करता है।

Table of contents

Certificate	i
Declaration	iii
Acknowledgements	v
Abstract	vii
List of figures	xxi
List of tables	xxxv
Nomenclature	xxxvii
List of publications	xli
I Introduction	1
1 Introduction	3
1.1 Inorganic glasses	5
1.2 Discovering novel glass composition	6
1.3 Organization of the Thesis	7

2	Literature Review, Methodology, and Datasets	11
2.1	Literature Review	11
2.1.1	Data driven modeling of inorganic glasses	11
2.1.2	Machine learning methods in material simulation	14
2.1.3	Up-scaling MD simulation using peridynamics	24
2.1.4	Gaps in the literature	25
2.1.5	Research objective	26
2.2	Methodology and Datasets	27
2.2.1	Machine Learning	27
2.2.2	Datasets	33
2.2.3	Software packages	33
2.2.4	Molecular dynamics	35
2.2.5	Miscellaneous	40
II	Predicting and Interpreting of Glass Property from Experimental Data	43
3	Machine Learning for Glass Design and Optimization	45
3.1	Introduction	46
3.2	Methods	48
3.2.1	Data preparation	48
3.2.2	Model development	49
3.3	Results	51
3.3.1	Deep learning model development	51
3.3.2	Exploring the compositional domain	53
3.3.3	Capturing the underlying physics: The mixed-modifier effect	53
3.3.4	Glass selection charts	55

3.4	Discussion	59
3.5	Conclusion	59
4	Understanding the Compositional Control on Glass Properties with Interpretable Machine Learning	61
4.1	Introduction	62
4.2	Methods	64
4.2.1	Dataset Preparation	64
4.2.2	Dataset visualization	65
4.2.3	Model development and hyperparameter optimization	66
4.3	Results	67
4.3.1	Dataset visualization	67
4.3.2	Machine learning models	70
4.3.3	Interpreting the composition-property relationships	75
4.3.4	Interaction between input features	82
4.4	Discussion	88
4.5	Conclusion	89
III	Physics Informed Machine Learning for Glass Simulation	93
5	Momentum Conserving Lagrangian Neural Network	95
5.1	Introduction	96
5.2	Methodology	100
5.2.1	Preliminaries and Problem Formulation	100
5.2.2	The MCLNN Architecture	103
5.3	Results	107
5.3.1	Experimental Setup	107

5.3.2	Accuracy	111
5.3.3	Generalizability to unseen system sizes	112
5.3.4	Interpretability	114
5.4	Learning interatomic potentials with MCLNN	115
5.5	Discussion	118
5.6	Conclusion	119
6	Lagrangian Graph Neural Network	121
6.1	Introduction	122
6.2	Methodology	124
6.2.1	Neural Architecture of LGNN	124
6.2.2	Theoretical Characterization	128
6.3	Results	132
6.3.1	Experimental setup	132
6.3.2	Spring system comparison with baseline	134
6.4	Learning interatomic potentials with LGNN	138
6.5	Discussion	139
6.6	Conclusion	141
7	Learning constitutive relationship from simulation trajectory	143
7.1	Introduction	144
7.2	Theory of Peridynamics	144
7.2.1	Classical local theory	145
7.2.2	Modelling fracture in LEFM	145
7.2.3	Peridynamics	146
7.3	Peridynamics software packages and PeriDyn	151
7.4	Overview of PeriDyn	152

7.5	Uniaxial tensile simulation with PeriDyn	153
7.6	Machine learned constitutive relationship	159
7.7	Conclusion	163
8	Conclusion and Future Outlook	165
8.1	Introduction	165
8.2	Conclusion	165
8.3	Research contribution	167
8.4	Future outlook	168
	References	171
IV	Appendices	199
	Appendix A	201
A.1	Dataset preparation	201
A.2	Deep Neural Network	202
A.3	Benchmarking the DNN	203
A.4	Training curves for the DNN	209
	Appendix B	211
B.1	Optimum hyperparameter of the XGBoost model for each property	211
B.2	Glass component list for each property	236
B.3	Range of hyperparameters used by Optuna to train the XGBoost models	244
B.4	Train and test R^2 , RMSE, and MAE values for the all the properties . . .	245
B.5	t-SNE embedding of the compositions plotted for each property	247
B.6	SHAP violin and river flow plots for each property	260
B.7	SHAP interaction plots for each property	273

Appendix C	287
C.1 Lagrangian Mechanics and Machine Learning	287
C.1.1 Preliminaries on Learning Dynamical Systems	287
C.1.2 Lagrangian mechanics with Pfaffian Constraints, Dissipation, and External Forces	288
C.2 Lagrangian Neural Network	290
C.2.1 Baseline LNN	290
C.2.2 Generalizability to unseen system sizes	290
C.2.3 Computing acceleration from Euler-Lagrange equations	292
C.2.4 Translational Symmetry	293
C.2.5 Rotational Symmetry	294
C.3 Initial Conditions and Task Configuration	295
C.4 Hardware and Software packages	298
Appendix D	299
D.1 Energy violation of LNN	299
D.2 Momentum conservation of LGNN	300
D.3 Experimental systems	302
D.3.1 n -spring system	302
D.4 Implementation details	302
D.4.1 Dataset generation	302
D.4.2 Architecture	304
D.4.3 Training details	304
D.5 Comparison of LGNN with baseline	305
D.6 Generalizability with drag	306
D.7 Interpretability	307

Appendix E	309
E.1 PeriDyn	309
E.1.1 PeriDyn Environment	309
E.1.2 Material Discretization	310
E.1.3 Material Models	311
E.1.4 Contact Models	313
E.1.5 Boundary Conditions	314
E.1.6 Solvers	316
E.1.7 Simulation	316
E.1.8 Parallel Computing	317
E.1.9 Input/Output	318
E.1.10 Visualization	319
E.2 Benchmarking PeriDyn models	320
E.2.1 Tensile simulation	320
E.2.2 Flexure simulation	323
E.2.3 Wave propagation simulation	324
E.2.4 Impact simulation	329
E.3 Machine learned constitutive relationship	330
E.3.1 Peridynamics simulation with bond based material	330
E.3.2 Machine learning with PeriDyn	334
E.3.3 Peridynamics simulation with machine learned material	338
CV	341

List of figures

1.1	The material timeline (reprinted with permission from Ashby (2008)). . .	4
1.2	The process of development of a model and experimental validation (reprinted with permission from (Valavala and Odegard)).	5
2.1	Time and length scales of a material simulation.	15
2.2	Material modeling techniques (reprinted with permission from (Valavala and Odegard)).	16
2.3	Machine learned interatomic potentials.	17
2.4	Schematic for descriptor construction for deep potentials (reprinted with permission from Wen et al. (2022)).	19
2.5	Three main components for GAP: (1) a reference database of quantum- mechanical data for suitably chosen structural models, (2) a representa- tion of the atomic environments (typically using combinations of 2-body and SOAP descriptors), and (3) the GP regression itself (reprinted with permission from Deringer et al. (2019)). Copyright 2019 WILEY-VCH. .	20
2.6	Principal Component Analysis. a) Evolution of variance with respect to input dimension. b) The variation of R^2 with respect to input dimension.	29
2.7	Flow of MD calculation.	36
2.8	Typical LJ interatomic potential.	39

3.1	Schematic representation of the methodology for the rational design of glasses. Glass dataset comprising of the input compositions and the corresponding properties are collected from the literature and glass databases. Deep learning (DL) models developed based on the processed dataset are used to populate the composition-property space and develop glass selection charts. These charts act as the road-map to selecting compositions satisfying multiple constraints, thus, enabling the design of novel glasses for targeted applications.	47
3.2	Property predictions based on deep learning (DL) models. Predicted values of (a) density, (b) Young's modulus, (c) hardness, (d) shear modulus, (e) thermal expansion coefficient (TEC), (f) glass transition temperature (T_g), (g) liquidus temperature, and (h) refractive index of oxide glasses using optimized DL models with respect to the experimental values. Note that the color represents the number of points per unit area for associated with each property following the respective coloring scheme. The inset shows the error in the predicted values as a probability density function (PDF) with the shaded region representing the 90% confidence interval. R^2 values of the training and test set are also provided.	52

3.3 Ternary plots of calcium aluminosilicate glasses from DL models.

Ternary plots of (a) density, (b) Young's modulus, (c) hardness, (d) shear modulus, (e) thermal expansion coefficient (TEC), (f) glass transition temperature (T_g), (g) liquidus temperature, and (h) refractive index of calcium aluminosilicate glasses. Square markers represent experimental compositions with the color representing the value of the property from the experiments. The underlying heatmap represents the predictions of the model. The scale bar corresponding to each subfigure shows the property range along with the coloring scheme. 54

3.4 Composition-dependent variations of properties highlighting the mixed-modifier effect.

Evolution of (a) density, (b) Young's modulus, (c) hardness, (d) shear modulus, (e) thermal expansion coefficient (TEC), (f) glass transition temperature (T_g), (g) liquidus temperature, and (h) refractive index with respect to the ratio of modifiers of four and five component glasses. The selected compositions of glasses are shown in (i) with x and y representing the appropriate modifier fraction. Note that alkaline, alkaline earth, and alkali-alkaline earth modifiers are considered here. Both negative and positive deviations from linearity are observed for the properties with respect to the modifier fraction, a signature of the *mixed-modifier* effect. 56

3.5	Glass selection charts. Variation of two independent properties for a large series of commonly used glasses, namely, aluminosilicate, silicate, borate, borosilicate, phosphosilicate, aluminoborate, and tellurite. (a) Young's modulus with density, (b) Young's modulus with glass transition temperature, (c) shear modulus with hardness, (d) Young's modulus with thermal expansion coefficient, (e) Glass transition temperature with liquidus temperature, and (f) density with refractive index. Dotted lines in the figures correspond to design lines with $X/Y = \text{constant}$ for subfigures (a), (b), (c), (e), and (f) and $XY = \text{constant}$ for subfigure (d). These glass selection charts can enable the rational design of glasses by allowing the choice of glasses subjected to multiple property and compositional constraints.	58
4.1	An ML pipeline for interpretable machine learning.	63
4.2	Periodic table with the elements present in our dataset highlighted as network formers, modifiers, intermediate, and others. (b) Two-dimensional t-SNE embedding of all the glass compositions used in the present work comprises 275,516 compositions and 221 components. The prominent families are highlighted as per the scheme shown in the legend. (c) Bar plot of the total number of glass compositions (upper subplot) and components (lower subplot) for the properties in the dataset.	68
4.3	Predicted values with respect to measured values for (a) annealing point, (b) crystallization temperature, (c) softening point, (d) strain point, (e) liquidus temperature, (f) Littleton temperature, (g) thermal expansion coefficient (TEC), (h) glass transition temperature, and (i) thermal conductivity for the test dataset.	71

4.4	Predicted values with respect to measured values for (a) bulk modulus, (b) compressive strength, (c) density, (d) Vickers hardness, (e) shear modulus, (f) tensile strength, (g) thermal shock resistance, and (h) Young's modulus for the test dataset.	72
4.5	Predicted values with respect to measured values for (a) Abbe number, (b) refractive index, (c) infrared (IR) transmittance, (d) solar transmittance, (e) visible transmittance, and (f) ultraviolet (UV) transmittance for the test dataset.	74
4.6	Predicted values with respect to measured values for (a) dielectric constant and (b) electrical conductivity for the test dataset.	75
4.7	SHAP violin and river flow plots for (a) annealing point, (b) crystallization temperature, (c) softening point, (d) strain point, (e) liquidus temperature, (f) Littleton temperature, (g) thermal expansion coefficient (TEC), (h) glass transition temperature, and (i) thermal conductivity.	78
4.8	SHAP violin and river flow plots for (a) bulk modulus, (b) compressive strength, (c) density, (d) Vickers hardness, (e) shear modulus, (f) tensile strength, (g) thermal shock resistance, and (h) Young's modulus.	80
4.9	SHAP violin and river flow plots for (a) Abbe number, (b) refractive index, (c) infrared (IR) transmittance, (d) solar transmittance, (e) visible transmittance, and (f) ultraviolet (UV) transmittance.	81
4.10	SHAP violin and river flow plots for (a) dielectric constant and (b) electrical conductivity.	82

4.11	Heat map of normalized SHAP interaction values for top 20 glass components for (a) annealing point, (b) crystallization temperature, (c) softening point, (d) strain point, (e) liquidus temperature, (f) Littleton temperature, (g) thermal expansion coefficient (TEC), (h) glass transition temperature, and (i) thermal conductivity.	85
4.12	Heat map of normalized SHAP interaction values for atmost top 20 glass components for (a) bulk modulus, (b) compressive strength, (c) density, (d) Vickers hardness, (e) shear modulus, (f) tensile strength, (g) thermal shock resistance, and (h) Young's modulus.	86
4.13	Heat map of normalized SHAP interaction values for atmost top 20 glass components for (a) Abbe number, (b) refractive index, (c) infrared (IR) transmittance, (d) solar transmittance, (e) visible transmittance, and (f) ultraviolet (UV) transmittance.	87
4.14	Heat map of normalized SHAP interaction values for top 20 glass components for (a) dielectric constant and (b) electrical conductivity.	88
4.15	Predicted and experimental values of (a) density, (b) hardness and (c) refractive index with respect to mol % of MgO. SHAP values of (c) density, (d) hardness and (e) refractive index with respect to mol % of MgO. . . .	90
5.1	MCLNN framework.	102
5.2	Results of the forward simulation of linear spring system from the same initial configuration using ground truth (continuous line), baseline LNN (dashed and dotted line), and MCLNN (dashed lines). (a) The evolution of Lagrangian, Hamiltonian (total energy), linear, and momentum for all the systems. (b) The error in Lagrangian, Hamiltonian (total energy), linear, and momentum predicted by the baseline LNN and MCLNN with respect to the ground truth.	111

5.3	Results of the forward simulation of non-linear spring and gravitational system from the same initial configuration using ground truth (continuous line) and MCLNN (dashed lines). (a) The evolution of Lagrangian, Hamiltonian (total energy), linear, and momentum for all the systems. (b) The error in Lagrangian, Hamiltonian (total energy), linear, and momentum predicted by the baseline LNN and MCLNN with respect to the ground truth.	113
5.4	Results of the forward simulation of (a) linear spring with six particles, (b) non-linear spring with six particles, and (b) gravitational system with eight particles from the same initial configuration using ground truth (continuous line) and MCLNN (dashed lines). The evolution of Lagrangian for all the systems is plotted.	113
5.5	Variation of V_{ij} with respect to q_{ij} as learned by the MCLNN (dashed line) in comparison to the ground truth (continuous line) for (a) linear spring, (b) non-linear spring, and (c) gravitational potential. The shaded region represents the range of q_{ij} values in the training set. Non-shaded region between two shaded region corresponds to interpolation; otherwise, it represents extrapolation.	115
5.6	Top: Potential energy with respect to interatomic distance for a unary LJ system. The dotted green line shows the analytical equation obtained using symbolic regression with the MCLNN model. Bottom: The error in MCLNN and analytical equation with respect to LJ equation.	116

5.7	Top: Potential energy with respect to interatomic distance for a) type A – type A, b) type A – type B, and c) type B – type B interaction in a binary LJ system. The dotted green line shows the analytical equation obtained using symbolic regression with the MCLNN model. Bottom: The error in MCLNN and analytical equation with respect to LJ equation.	117
5.8	Potential energy with respect to interatomic distance for a) O–O, b) Si–O, and c) Si–Si interaction in a liquid silica system.	117
6.1	General architecture of Lagrangian Graph Neural Network. Note that the drag predictions are similar to the kinetic energy computations. However, the drag is directly featured in <i>EL</i> equation and hence is not shown here.	125
6.2	Visualization of a 5-spring system and a liquid silica system.	132
6.3	Energy violation (top) and rollout error (bottom) of LGNN (red) and LNN (blue) for spring systems with 3, 4, and 5 links and without drag. Note that the LGNNs for springs are trained on 3 and 5 links, respectively, and tested on all the systems. LNNs are trained and tested separately on each of the systems. The shaded region shows the 95% confidence interval from 100 initial conditions.	133
6.4	Energy violation and rollout error for 5, 50, and 500 spring systems with LGNN, GNS, and LGN trained on five spring systems. The shaded region shows the 95% confidence interval from 100 initial conditions.	136
6.5	Data efficiency of LNN and LGNN on the 3-spring system.	137
6.6	(a) Pair distribution function, (b) structure factor for DFT and LGNN simulation for a liquid silica system.	138
6.7	Bond angle distribution for a) O–Si–O, b) Si–O–Si, and c) Si–Si–Si for DFT and LGNN simulation for a liquid silica system.	140

7.1	Peridynamics material points and interaction of points at x_k and x_j (reprinted with permission from Madenci and Oterkus (2014)).	146
7.2	Deformed state of peridynamics material points (reprinted with permission from Madenci and Oterkus (2014)).	147
7.3	Blocks showing a condensed view of different parts of <code>PeriDyn.Env struct</code>	153
7.4	Schematic showing major steps during the simulation.	154
7.5	Bar showing evolution of damage during tensile load simulation.	157
7.6	Normal stress vs. strain curve for bar simulation under constant strain rate for the bond-based model.	158
7.7	a) Stress-strain relationship obtained from the MD simulation, peridynamics model, and analytical equation for uniaxial tension. b) Stress-strain relationship obtained from the machine learning model, peridynamics model, and analytical equation for uniaxial tension.	161
7.8	Comparison of bond force density of bond-based peridynamics model and machine-learned model.	162
A.1	The schematic of the dataset preparation methodology	202
A.2	Feature selection for DL models. R^2 and mean squared error (MSE) with respect to number of features for (a) test dataset and (b) training dataset of Young's modulus using LARS.	202
A.3	Distribution of number of glass components for oxide glasses. Histogram of number of glass components for (a) density, (b) Young's modulus, (c) hardness, (d) shear modulus, (e) thermal expansion coefficient (TEC), (f) glass transition temperature (T_g), (g) liquidus temperature, and (h) refractive index dataset after the application of LARS.	204

A.4	Distribution of properities for oxide glasses. Histogram of values for (a) density, (b) Young's modulus, (c) hardness, (d) shear modulus, (e) thermal expansion coefficient (TEC), (f) glass transition temperature (T_g), (g) liquidus temperature, and (h) refractive index of oxide glasses.	205
A.5	Predicted values of (a) density, (b) Young's modulus, (c) hardness, (d) shear modulus, (e) thermal expansion coefficient (TEC), (f) glass transition temperature (T_g), (g) liquidus temperature, and (h) refractive index of oxide glasses using linear regression with respect to the experimental values. Note that the color represents the number of points per unit area associated with each property following the respective coloring scheme. The inset shows the error in predicted values as a probability density function (PDF), with the shaded region representing the 90% confidence interval. R^2 values of the training and test set are also provided.	206
A.6	Predicted values of (a) density, (b) Young's modulus, (c) hardness, (d) shear modulus, (e) thermal expansion coefficient (TEC), (f) glass transition temperature (T_g), (g) liquidus temperature, and (h) refractive index of oxide glasses using xgboost with respect to the experimental values. Note that the color represents the number of points per unit area associated with each property following the respective coloring scheme. The inset shows the error in the predicted values as a probability density function (PDF), with the shaded region representing the 90% confidence interval. R^2 values of the training and test set are also provided.	207

A.7 Training curves of XGBoost models. Training curves showing the evolution of the root mean squared error (RMSE) loss with respect to number of epochs, for the training and validation set for: (a) density, (b) Young's modulus, (c) hardness, (d) shear modulus, (e) thermal expansion coefficient (TEC), (f) glass transition temperature (T_g), (g) liquidus temperature, and (h) refractive index.	208
A.8 Training curves of the deep learning (DL) models. Training curves showing the evolution of the mean squared error (MSE) loss with respect to the epoch for the training and validation set for: (a) density, (b) Young's modulus, (c) hardness, (d) shear modulus, (e) thermal expansion coefficient (TEC), (f) glass transition temperature (T_g), (g) liquidus temperature, and (h) refractive index.	210
B.1 Two-dimensional t-SNE embedding of all the glass compositions for (a) Abbe number, (b) annealing point, (c) bulk modulus, (d) compressive strength, (e) crystallization temperature, (f) density, (g) dielectric constant, (h) electrical conductivity, (i) Vickers hardness, (j) refractive index, (k) shear modulus, (l) softening point, (m) strain point, (n) liquidus temperature, (o) Littleton temperature, (p) thermal expansion coefficient (TEC), (q) tensile strength, (r) glass transition temperature, (s) thermal conductivity, (t) thermal shock resistance, (u) infrared (IR) transmittance, (v) solar transmittance, (w) visible transmittance, (x) ultraviolet (UV) transmittance, and (y) Young's modulus. The major families are highlighted as per the scheme shown in the legend.	259

B.2	SHAP violin and river flow plots for (a) Abbe number, (b) annealing point, (c) bulk modulus, (d) compressive strength, (e) crystallization temperature, (f) density, (g) dielectric constant, (h) electrical conductivity, (i) Vickers hardness, (j) refractive index, (k) shear modulus, (l) softening point, (m) strain point, (n) liquidus temperature, (o) Littleton temperature, (p) thermal expansion coefficient (TEC), (q) tensile strength, (r) glass transition temperature, (s) thermal conductivity, (t) thermal shock resistance, (u) infrared (IR) transmittance, (v) solar transmittance, (w) visible transmittance, (x) ultraviolet (UV) transmittance, and (y) Young's modulus.	272
B.3	Heat map of normalized SHAP interaction values for top 20 glass components for (a) Abbe number, (b) annealing point, (c) bulk modulus, (d) compressive strength, (e) crystallization temperature, (f) density, (g) dielectric constant, (h) electrical conductivity, (i) Vickers hardness, (j) refractive index, (k) shear modulus, (l) softening point, (m) strain point, (n) liquidus temperature, (o) Littleton temperature, (p) thermal expansion coefficient (TEC), (q) tensile strength, (r) glass transition temperature, (s) thermal conductivity, (t) thermal shock resistance, (u) infrared (IR) transmittance, (v) solar transmittance, (w) visible transmittance, (x) ultraviolet (UV) transmittance, and (y) Young's modulus.	285
C.1	Predicted force for (a) linear spring, (b) non-linear spring, and (c) gravitational system with respect to the measured forces on the test dataset of baseline LNN.	291

C.2	Results of the forward simulation of (a) linear spring, (b) non-linear spring, and (b) gravitational system from the same initial configuration using ground truth (continuous line) and MCLNN (dashed lines). The evolution of Lagrangian, Hamiltonian (total energy), linear and angular momenta for all the systems with different number of particles than that was present in the training set are plotted.	291
D.1	Energy violation with respect to rollout error (right) of LGNN (red) and LNN (blue) for pendulum and spring systems with 3, 4, and 5 links. LNNs clearly exhibit a larger energy violation error.	306
D.2	Energy violation and rollout error for 5, 50, and 500 spring system with LGNN trained on 5 spring system with linear drag. Shaded region shows the 95% confidence interval from 100 initial conditions.	306
D.3	Learned node-level drag force.	307
E.1	Time taken vs. the number of threads with different system sizes.	318
E.2	A truncated sample data file from PeriDyn simulation.	319
E.3	Visualization of a data file from PeriDyn simulation.	320
E.4	Bar showing evolution of damage during tensile load simulation.	322
E.5	Normal stress vs strain curve for bar simulation under constant strain rate for a) bond based model and b) ordinary state based model.	323
E.6	Damage evolution of a bar under flexure load.	325
E.7	Lateral force vs displacement curve for bar under flexure load.	325
E.8	Heatmap of pressure wave propagation for a bar under impulse load.	327
E.9	Time evolution of normal stress at fixed end of a bar after impulse load.	328
E.10	Damage evolution of a disk being hit by a projectile.	332
E.11	Damage evolution of a disk being hit by a projectile.	333

List of tables

7.1	Comparison of Young's modulus.	156
7.2	Comparison of Young's modulus.	160
A.1	Architecture of the deep neural network for each of the properties. The hidden layers and the number of units per hidden layer along with the activation function is shown.	203
B.1	XGBoost tree ensemble specific hyperparameters	244
B.2	Booster (DART and GBtree) specific hyperparameters	244
B.3	DART specific hyperparameters	244
B.4	The train and test R^2 , RMSE, and MAE values for the 25 properties trained using XGBoost	245
C.1	Train and validation loss for each NN architecture for the linear spring task for LNN. Note that for hyperparameter search (i.e. number of hidden units), minimum train loss was set to $1.0e-8$ as stopping criteria.	296
E.1	List of parameters used for simulation of bar under tensile loading.	321
E.2	Comparison of Young's modulus.	321
E.3	List of parameters used for simulation of bar under flexure loading.	324
E.4	Comparison of Young's modulus.	324
E.5	List of parameters used for simulation of bar under impulse loading.	326

E.6	Comparison of Young's modulus.	326
E.7	List of parameters used for simulation of disk being hit by a projectile. . .	329

Nomenclature

Roman Symbols

F interatomic force

Acronyms / Abbreviations

ACE atomic cluster expansion

ALE accumulated local effects

CART classification and regression tree

CGCNN crystal graph convolution neural network

CNN covariant compositional networks

DeepLIFT Deep Learning Important Features

DFT density function theory

DNN deep neural network

FFNN feed forward neural network

FSDP first sharp diffraction peak

GAP gaussian approximation potential

GNNFF graph neural network force field

GNN graph neural network

GPR gaussian process regression

LGNN lagrangian graph neural network

LIME local interpretable model-agnostic explanations

LJ Lennard-Jones

MCLNN momentum conserving lagrangian neural network

MD molecular dynamics

MEGNet MatErials graph network

ML machine learning

MPGNN message passing graph neural network

MTP moment tensor potential

NN neural network

NNP neural network potential

PCA principal component analysis

PDF pair distribution function

PES potential energy surface

PINN physically informed neural network

RF random forest

SHAP shapely additive explanations

SNAP spectral neighbor analysis potential

SOAP smooth overlapping atomic positions

SVM support vector machine

List of publications

Papers featuring in the thesis work:

1. Ravinder, R; Sridhara, Karthikeya H; Bishnoi, Suresh; Grover, Hargun Singh; Bauchy, Mathieu; Jayadeva, J; Kodamana, Hariprasad; Krishnan, NM Anoop; Deep Learning Aided Rational Design of Oxide Glasses, Materials Horizons, 2020, Royal Society of Chemistry
2. Ravinder, R; Venugopal, Vineeth; Bishnoi, Suresh; Singh, Sourabh; Zaki, Mohd; Grover, Hargun Singh; Bauchy, Mathieu; Agarwal, Manish; Krishnan, NM Anoop; Artificial intelligence and machine learning in glass science and technology: 21 challenges for the 21st century, International Journal of Applied Glass Science, 2021
3. Ravinder, R; Bishnoi, Suresh; Zaki, Mohd; Krishnan, NM; Revealing the compositional control of electrical, mechanical, optical, and physical properties of inorganic glassesarXiv preprint arXiv:2103.12050, 2021 (under review)
4. Bhattoo, Ravinder; Ranu, Sayan; Krishnan, NM; Lagrangian neural network with differentiable symmetries and relational inductive bias, arXiv preprint arXiv:2110.03266, 2021 (pre-print)
5. Bhattoo, Ravinder; Ranu, Sayan; Krishnan, NM; Lagrangian graph neural network (Submitted)
6. Bhattoo, Ravinder; Ranu, Sayan; Krishnan, NM; Learning Rigid Body Dynamics with Lagrangian Graph Neural Network (Submitted)

Papers published related to the thesis work:

7. Rivera, Jared; Berjikian, Jonathan; Ravinder, R; Kodamana, Hariprasad; Das, Sumanta; Bhatnagar, Naresh; Bauchy, Mathieu; Krishnan, NM Anoop; Glass fracture upon ballistic impact: new insights from peridynamics simulations, *Frontiers in Materials*, 6, 239, 2019, Frontiers Media SA
8. Bishnoi, Suresh; Singh, Sourabh; Ravinder, R; Bauchy, Mathieu; Gosvami, Nitya Nand; Kodamana, Hariprasad; Krishnan, NM Anoop; Predicting Young's modulus of oxide glasses with sparse datasets using machine learning, *Journal of Non-Crystalline Solids*, 524, 119643, 2019, North-Holland
9. Nayak, Sumeru; Ravinder, R; Krishnan, NM; Das, Sumanta; A Peridynamics-Based Micromechanical Modeling Approach for Random Heterogeneous Structural Materials, *Materials*, 13, 6, 1298, 2020, Multidisciplinary Digital Publishing Institute
10. Bishnoi, Suresh; Ravinder, R; Grover, Hargun Singh; Kodamana, Hariprasad; Krishnan, NM Anoop; Scalable Gaussian processes for predicting the optical, physical, thermal, and mechanical properties of inorganic glasses with large datasets, *Materials Advances*, 2, 1, 477-487, 2021, Royal Society of Chemistry
11. Zaki, Mohd; Venugopal, Vineeth; Bhattoo, Ravinder; Bishnoi, Suresh; Singh, Sourabh Kumar; Allu, Amarnath R; Krishnan, NM Anoop; Interpreting the optical properties of oxide glasses with machine learning and Shapely additive explanations, *Journal of the American Ceramic Society*, 105, 6, 4046-4057, 2022
12. Bishnoi, Suresh; Bhattoo, Ravinder; Krishnan, Ranu, Sayan; NM Anoop; Enhancing the Inductive Biases of Graph Neural ODE for Modeling Dynamical Systems (Submitted)
13. Thangamuthu, Abishek; Kumar, Gunjan; Bishnoi, Suresh; Bhattoo, Ravinder; Krishnan, NM Anoop; Ranu, Sayan; Unravelling the Performance of Physics-informed Graph Neural Networks for Dynamical Systems (Submitted)

Other papers published during the Ph.D. course:

14. Krishnan, NM Anoop; Ravinder, R; Kumar, Rajesh; Le Pape, Yann; Sant, Gaurav; Bauchy, Mathieu; Density-stiffness scaling in minerals upon disordering: Irradiation vs. vitrification, *Acta Materialia*, 166, 611-617, 2019, Pergamon
15. Dhawan, Sameer; Ghosh, Sukanya; Ravinder, R; Bais, Sachendra S; Basak, Soumen; Krishnan, NM Anoop; Agarwal, Manish; Banerjee, Manidipa; Haridas, V; Redox sensitive self-assembling dipeptide for sustained intracellular drug delivery, *Bioconjugate chemistry*, 30, 9, 2458-2468, 2019, American Chemical Society
16. Ravinder, R; Kumar, Rajesh; Agarwal, Manish; Krishnan, NM; Evidence of a two-dimensional glass transition in graphene: Insights from molecular simulations, *Scientific reports*, 9, 1, 1-9, 2019, Nature Publishing Group
17. Ravinder, R; Garg, Prateet; Krishnan, NM Anoop; Glass Transition and Crystallization in Hexagonal Boron Nitride: Crucial Role of Orientational Order, *Advanced Theory and Simulations*, 1900174, 2019
18. Ravinder, R; Kumar, Abhishek; Kumar, Rajesh; Vangla, Prashanth; Krishnan, NM Anoop; Irradiation-induced brittle-to-ductile transition in α -quartz, *Journal of the American Ceramic Society*, 103, 7, 3962-3970, 2020
19. Ravinder, R; Singh, Sourabh; Bishnoi, Suresh; Jan, Amreen; Sharma, Amit; Kodamana, Hariprasad; Krishnan, NM Anoop; An adaptive, interacting, cluster-based model for predicting the transmission dynamics of COVID-19, *Heliyon*, 6, 12, e05722, 2020, Elsevier
20. Bhaskar, Pratik; Kumar, Rajesh; Maurya, Yashasvi; Ravinder, R; Allu, Amarnath R; Das, Sumanta; Gosvami, Nitya Nand; Youngman, Randall E; Bødker, Mikkel S; Mascaraque, Nerea; Cooling rate effects on the structure of 45S5 bioglass: Insights from experiments and simulations, *Journal of Non-Crystalline Solids*, 534, 119952, 2020, Elsevier