

**PREDICTING THE BAND GAP VALUES OF ABX_3
TYPE PEROVSKITE MATERIALS BY USING
MACHINE LEARNING**

Warusa Hennadige Kasuni Nisansala

179441 M

Masters of Science in Materials Science

Department of Materials Science and Engineering

University of Moratuwa

Sri Lanka

August 2023

**PREDICTING THE BAND GAP VALUES OF ABX_3
TYPE PEROVSKITE MATERIALS BY USING
MACHINE LEARNING**

Warusa Hennadige Kasuni Nisansala

179441 M

**Thesis was submitted in partial fulfilment of the
requirements for the Masters of Science in Materials Science**

Department of Materials Science and Engineering

University of Moratuwa

Sri Lanka

August 2023

DECLARATION

To the best of my knowledge and belief, this thesis doesn't contain any prior printed or written material by anyone other than where an acknowledgment is made in the text. I hereby declare that this is entirely my own work and that no material previously submitted for a degree or certification at another university or higher education institution is included.

In addition, I hereby grant the University of Moratuwa the non-exclusive right to publish, reproduce, and disseminate my thesis or dissertation in whole or in part in print, electronic, or other formats. I reserve the right to incorporate all or a portion of this content into my future works.

Signature:

Date: 23.08.2023

The higher than candidate has allotted analysis for the master thesis underneath our supervision.

Name of the supervisors:

1.Prof. (Mrs.) Galhenage A. Sewwandi

2.Prof (Mr.) P. K. W. Abeygunawaradhana

Signature of the supervisors:

Date :

1.

.....

2.....

ACKNOWLEDGEMENT

I would prefer to express my sincere gratitude to everyone who gave me the support to finish this report. I would especially like to express my gratitude to my supervisors, Prof. Galhenage, A.Sewwandi, and Prof. P. K. W. Abeygunawaradhana, whose contributions in the form of stimulating ideas and encouragement helped me to finish this report.

Additionally, I would like to express my sincere gratitude to the department's head Mr. V. Sivarhar, as well as to all of the both academic and non-academic employees who work at the Department of Materials Science and Engineering.

ABSTRACT

Perovskite material has recently attracted a lot of attention due to the increased demand for renewable energy. Metal halide perovskite material is useful in a variety of sectors. Nonetheless, there remains a requirement for the implementation of a machine learning approach to guide the enhancement of perovskite materials' performance. This study employs machine learning to build design methodologies, anticipate perovskite material performance, and improve material composition.

By employing a genetic algorithm and accounting for both the tolerance factor and the octahedral factor, ten stable perovskite structures were determined by evaluating the substitution effects of various alternatives among the seventy-five combinations. Crossover and mutation techniques were employed to ascertain the ten stable perovskite compounds identified at room temperature.

With regard to previously evaluated papers, 120 data points were carefully chosen to create the machine-learning models. These models direct the design of contemporary perovskite structures. Utilizing the Support Vector Regression (SVR) and Extreme Gradient Boosting Regression (XG boosting) techniques.

Novel compositions of perovskite materials are created using machine learning techniques, and these compositions are subsequently synthesized to assess the effectiveness of the developed models.. It shows a high-stability perovskite composition in iodide halide with Pb^{2+} . Furthermore, NH_4GeCl_3 gives the band gap value as 1.6eV, and NH_4GeBr_3 gives 1.8eV. These two combinations are newly introduced by this research.

A common validation split would allocate 60% of the data for training, 20% for validation, and 20% for testing. Every model's prediction accuracy on the test set is assessed using the explained variance (R^2 coefficient) and root mean square error (RMSE). The XG boosting model provides higher values of RMSE and R^2 coefficients of 0.48 and 0.88, respectively, than the SVR model. The SVR model exhibited accuracy rates of 59%, whereas the XG Boosting model exhibited accuracy rates of 87%. Therefore, according to the results, the SVR model can be considered relatively weak in comparison to the XG Boosting regression model.

TABLE OF CONTENTS

Page No

Declaration	i
Acknowledgements	ii
Abstract	iii
Table of contents	iv
List of Figures	v
List of Tables	vi
1. Introduction	1
2. Literature review.....	3
3. Dataset and feature engineering.....	13
3.1 Dataset	13
3.2 Feature Engineering.....	14
4. Methodology.....	16
4.1 Procedure of the checking stability in perovskite solar cell	16
4.2 Procedure of the train model for predicting band gap values using the machine learning	12
4.3 Criteria of evaluation	13
5. Results and Discussion	25
5.1 Part I – finding the most stable perovskites in room temperature by using the genetic algorithm	25
5.2 Part II – Predicting band gap values in ABX ₃ type Perovskite by using machine learning.....	27
5.3 Evaluation of the machine learning mode.....	33
6. Conclusion	34
References.....	35

LIST OF FIGURES

Figure	Page No
Figure 3.2: The 11 selected features are ranked with its percentage.....	15
Figure 4.1.1 : Flow chart of the methodology for finding most stable perovskite material using the mutation and crossover in the genetic algorithm, along with the tolerance factor and the octahedral factor.....	17
Figure 4.1.2 : Flow chart illustrating the steps involved in training a machine learning model to forecast band gap values using SVR method.	19
Figure 5.2.1 : The interface of the first model trained to use previous experimental data by using SVR algorithm	27
Figure 5.2.2 : The interface of the Google Colab note book used to predict band gap values by using XG Boost algorithm.	28
Figure 5.2.3 : Graphical representation of the predicted bandgap values of the perovskites compounds which selected as the most stable structure by using the SVR algorithm.....	30
Figure 5.2.4 : Graphical representation of the predicted band gap values of perovskites compounds which selected as the most stable structure by using XG Boost algorithm.....	30
Figure 5.2.5 : Graphical representation and comparison between the predicted band gap values of the perovskite compounds by using both SVR and XG Boost algorithm.....	31

LIST OF TABLE

Table	Page No
Table 3.1 Schematic illustration of the perovskite of type ABX_3 , chosen	
Material design space	14
Table 5.1: Data on the structures of several methylammonium lead	
trihalides at various temperatures	25
Table 5.2 : Predicted band gap values of the ABX_3 type perovskite	
material using models trained and introduced in this research.....	29
Table 5.3: Results of R^2 coefficient, RMSE, and prediction accuracy	
values in against the two models.....	33