Featureless approach for predicting Critical Temperature of Superconductors

Madhava Gaikwad
Cisco Systems, Bengaluru
India
madgaikw@cisco.com

Ashwini R Doke
Department of MCA
CMR Institute Of Technology,Bengaluru
India
yavani@outlook.com

Abstract— Superconductive materials holds key for fully functional extreme ultra reliable low latency communication. The cost effectiveness of latter technology depends upon ability to create and manufacture materials exhibiting superconductivity at high temperature. In this paper we estimate a data driven machine learning approach to predict critical temperature of superconductor without extracting the feature from superconductors chemical formula. Yet All other research is focusing on generating features from molecular formula and then using them for prediction. our approach is unique, straightforward, featureless and simple. Our approach demonstrates the accuracy of different machine learning algorithms applied on dataset which contains chemical formulas of superconductor. This approach only predicts the critical temperature of superconductor it will not give information about material is superconductor or not. We have provided source code so further experiments can be carried out.

***Index Terms*—Superconductivity, Critical temperature of super conductor, Machine learning, Material Science.**

I. INTRODUCTION

The near future wireless networks(6G) have to enable many-fold increase in the network services capacity and performance. The expected delay for communication channel is sub-1 ms[11]. The very stringent requirements for performance will lead to development of completely new paradigms and technologies and widespread usage. We expect quantum computing will be used to deploy such high performance low latency services[14]. The very high quantum computing capabilities of superconductor based quantum computing platforms are in final stage of research and developments[12]. Designing new materials exhibiting superconductive behaviour at high temperature or room temperature is highly desired.

Superconductor is a material that conduct an electricity from one atom to another atom with no resistance that mean no any other form of energy such as heat sound would be released from material ,when material riches such state is known as critical temperature and that point material becomes superconductive. Such superconductive material with zero resistance has significant application such as in healthcare MRI(magnetic resonance systems) systems are used to get details body imaging, also in superconducting coils are used to maintain high magnetic fields in the Large Hadron Collider at CERN, we can use superconductor in many applications, but it has two issues as superconductor conducts current with no resistance only at below superconducting critical temperature. And to guess the critical temperature is an open problem since the innovation of superconductivity. This paper uses data driven machine learning approach to predict the critical temperature of super conductor only by use of Mendeleev's atomic table. The critical temperature is measured by Appling different machine learning algorithms such as SVM, PCA, decision tree, Bayes classification, random forest … etc. and finally demonstrates the performance of such algorithm to measure the critical temperature.

basically, superconductor are of two types, Type1 superconductors uses basic elements to conduct electricity e.g. electrical wire to microchip of computer. Whereas type II superconductor requires higher temperature to reach at superconductivity state often these types of conductor are composed of metallic compounds such as copper or lead.

The study demonstrates how to find the critical temperature of superconductor using data driven featureless machine learning approach.

II LITERATURE REVIEW

Valentine [1] used machine learning modeling to guess critical temperature of superconductor, his research focus on machine learning approach to find the critical temperature of superconductor. As research on superconductivity is carried since more than century there are few features of superconductivity has not been clearly defined or understood, so to overcome this gap it has been demonstrated that machine learning can be used to predict critical temperature of super conductor. Research on superconductivity has been proposed from century as superconductivity is the subject of intense physics, chemistry and one of the most important topics of material science. There is some statistical model [2] has been developed to predict critical temperature of super conductor by extracting features from superconductors chemical formula such model is reasonably predicting well. To generate material science data is complex procedure so it will be expensive as well much research these days are focusing on understanding and extracting important features from chemical formula.

Neural network described in [3] has address a generative machine learning framework to predict critical temperature which also shows the generality of machine learning techniques used in field of material science.
As described in [4] there are lot many challenges in material science to predict the critical temperature of superconductor. Review [4] suggest what classification and regression methods can be used on material problems.
Data mining can also be used to identify hidden relationship between materials and also gives new opportunity for material design and understanding. As amount of data grows day by day such techniques can able to extract scientific principles and design rules that couldn't be determine through conventional methods [5]
As described in [6] there are some guidance and perspective for how apply machine learning in field of material science to discover hidden patterns or trends in data which can be helpful for material scientist in their research.
Further in [7] as stated use of machine learning, deep learning, artificial intelligence in field of material, processes, and structure engineering can bring innovative change in field of MPSE.
In different fields of material science its common now to generated large volume of data, to analyze such data with conventional method is challenging task. In order to overcome such challenges, we can use data science, machine learning models and tools. So that it will be easy to understand and extract important features from such huge data to make better prediction [8]
The recent study applies deep learning models to predict critical temperature of superconductor, the author [9] demonstrated such approach are useful in field of material science. It represents how deep learning model learns automatically by Appling atomic table once.
The study recently published in nature[10] states that lot of machine learning model requires manual effort in order to create features from chemical formula, but here they have address this challenge by using CNN convolution neural network which only need component information model will automatically learn the properties and construct the features.
Although all studies above show significant advancement in field of material science by using machine learning, deep learning, AI, to add more advancement our research demonstrate the featureless approach to predict critical temperature of superconductor. By measuring accuracy of different machine learning algorithm. Which is unique approach in field of material science.

III EXPERIMENTAL RESULTS:

In order to predict critical temperature of Superconductor we used molecule chemical composition formula to train, learn and predict critical temperature.
Below supervised machine learning techniques are used:
Decision Tree which is one of the widely used supervised machine learning algorithm which has been used to predict critical temperature.by giving molecule chemical composition formula as input the decision tree algorithm is calculating the critical temperature. Random forest is also one of the supervised learning technique which create the multiple decision tree.by using random forest algorithm we have predicted the critical temperature of superconductor, another algorithm used is Bayes classification algorithm which uses supervised machine learning approach for classification, then we have used linear classification algorithm to predict the critical temperature of superconductor. Then we have also applied PCA and decision PCA in order to predict the critical temperature. XGboost is also ensemble machine learning technique that uses gradient boosting framework. Then finally we used SVM (support vector machine algorithm which is also one of the important machine learning technique.
To guess critical temperature of superconductor we have used molecule chemical composition formula from Mendeleev's atomic table. The input to above machine learning algorithm is chemical formula from atomic table and the output will be the critical temperature of that particular chemical formula. The data set has been collected from japans national institute of material science. The 70% of the dataset has been used to train machine learning algorithm and 30% dataset is used to test the algorithm with new given formula. Final analysis has been done by measuring different parameters such as max_error, mean square error, explained variance, mean absolute error, Mean squared log error, R2 error of different algorithms. Below is the results of all algorithms

1 Mean square error – It measures the average squared error of prediction, It calculate the difference between the actual and expected results at each point and then average those values. It defined as

$$\text{Mean square error} = 1/N \sum (a_i - b_i)^2$$

Where ai is expected result and b_i is predicted result of model.The result of mean squared error score is as shown in table1.

Table 1. Mean squared error of evaluated algorithm

Evaluated algorithm	score
Random forest(RF)	370
Decision Tree	400
Bayes	470
Linear	470
Decision Tree PCA	470
SVM	320
XGBoost	700
SVMRBF	450

Fig 1 shows the performance of algorithms by using mean squared error. It has been seen that XG boost perform well as the score of XG boost is 700 while RF performance is not up to mark.

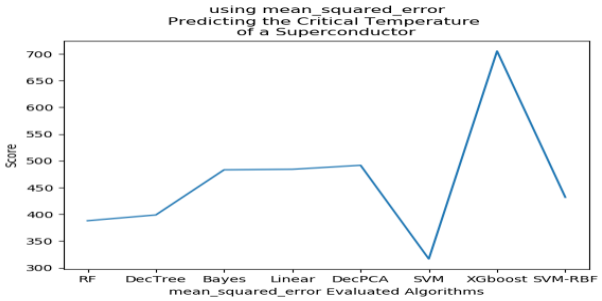

Fig1. Using mean_squared_error.

2 **Explained variance** it is a measured used to find discrepancy between expected and actual data. The result of explained variance score is shown in table2.

Table2. explained variance of evaluated algorithm.

Evaluated algorithm	score
Random forest(RF)	0.66
Decision Tree	0.66
Bayes	0.58
Linear	0.58
Decision Tree PCA	0.56
SVM	0.72
XGBoost	0.58
SVMRBF	0.62

As shown in figure2 SVM outperforms well as explained variance of SVM is 0.72 which is higher than all other algorithms

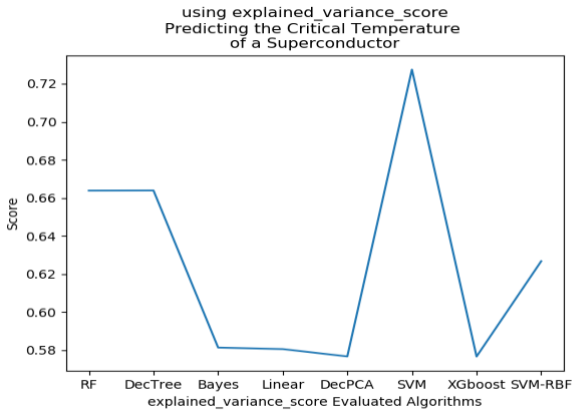

Fig2 using explained variance score predicting critical temperature

3. **Mean Absolute Error (MAE):** In MAE is an average of absolute differences between the expected and actual values. which can be calculated as

Mean Absolute Error = $1/N \sum |a_i - b_i|$The result of mean absolute error score is shown in table3

Table 3. mean_absolute_error score

Evaluated algorithm	score
Random forest(RF)	14
Decision Tree	11
Bayes	15
Linear	15
Decision Tree PCA	13
SVM	12
XGBoost	19
SVMRBF	6

Performance of XGBOOST is more as compared to all other algorithms as shown in fig3 below.

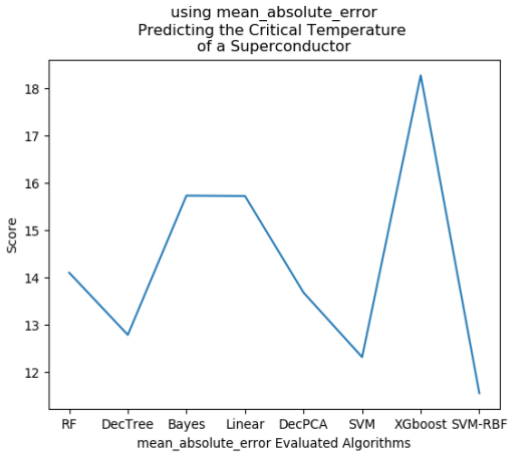

Fig.3 score using mean_absolute_error

4 **R2 error:** R2 error means sum of the differences between the actual values and the expected ones. Figure shows r2 error calculation. It is an important measure to analyze accuracy of any machine learning algorithm. The result obtained by using r2 error are shown in table4.

Table 4. score of R2 error

Evaluated algorithm	score
Random forest(RF)	0.66
Decision Tree	0.65
Bayes	0.59
Linear	0.59

Decision Tree PCA	0.57
SVM	0.72
XGBoost	0.20
SVMRBF	0.63

As shown in fig.4 SVM has highest R2 error while XGboost has lowest r2 error so svm can be used to predict critical temperature.as it outperform well as compared to other algorithms.

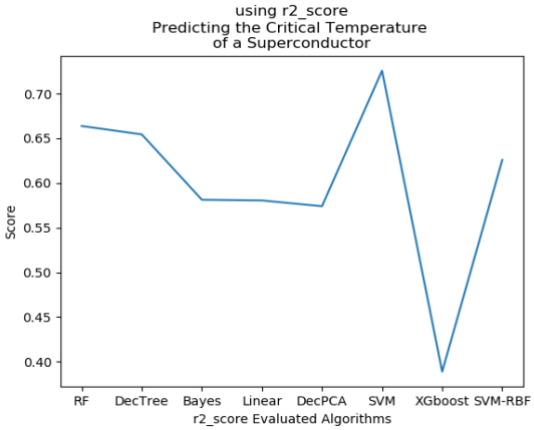

Fig.4. Score of r2 error.

5 **Mean squared log error (MSLE)** - It is the ratio between the expected and actual values. The result of

Mean squared log error is shown in table5

Table5: Mean squared log error score.

Evaluated algorithm	Score
Random forest(RF)	0.71
Decision Tree	0.74
Decision Tree PCA	0.67
XGBoost	0.30

As shown in fig mean squared log error of decision tree is highest while XGboost is very less so if we consider performance of RF, Decision Tree, DecPCA, XGboost we can see the decision tree perform best.

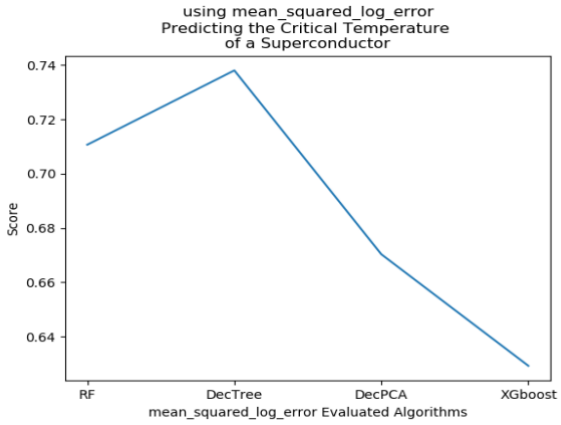

Fig.5 score using mean_squared_log_error.

6 **Median absolute error**: it is defined as median of expected and actual result.The score of evaluated algorithm is shown in table6

Table 6. score of **Median absolute error**

Evaluated algorithm	score
Random forest(RF)	8.5
Decision Tree	7
Bayes	12
Linear	12
Decision Tree PCA	6.5
SVM	7.5
XGBoost	9
SVMRBF	5

Fig.6 shows the score of median absolute error of different algorithms. As we can analyze the performance of Bayes and Linear is very high as compared to all other algorithms

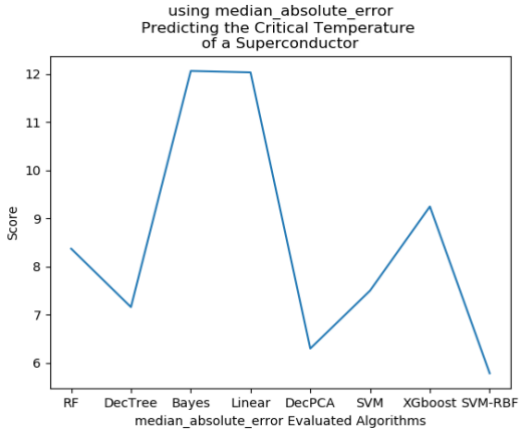

Fig.6. score using median absolute error

7 **Max error**: it is maximum difference between the point estimate and the actual result. The score obtained by using max error is shown in table7.

Table 7. Max error score.

Evaluated algorithm	score
Random forest (RF)	100
Decision Tree	100
Bayes	400
Linear	>400
Decision Tree PCA	175
SVM	100
XGBoost	100
SVMRBF	>150

Figure7 shows the max error calculations of different algorithms. here we can see max difference of Bayes and linear is high as compared to other algorithms so we can consider these two algorithm for better analysis.

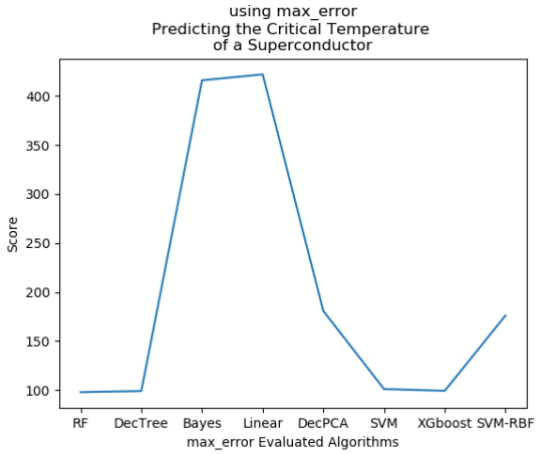

Fig.7.score using max_error.

Source Code: Source code for experiment is available from GitHub [13]. It provides source code and dataset used for experiment.

IV CONCLUSION

We have applied different supervised machine learning algorithm to guess the critical temperature of superconductor. As per results advance machine learning algorithms like SVM, SVM with RFB and XGBoost are useful to predict critical temperature as they have best learning ability, we can also say that PCA is also equally important as it reduces data dimension and converse fast with applied data. Finally, we can conclude that such featureless approach to predict critical temperature of superconductor will help many material scientists to find out critical temperature of any chemical formula. Which brings quick innovation in field of material science.

We expect more experiments to be carried out and provides ability to predict new types of superconductive materials to generate low cost quantum computing needed for development of 6G and get quickly deployed at mass

V. REFRENCES

[1] Stanev, V., Oses, C., Kusne, A.G., Rodriguez, E., Paglione, J., Curtarolo, S. and Takeuchi, I., 2018. Machine learning modeling of superconducting critical temperature. *npj Computational Materials*, *4*(1), pp.1-14.

[2] Hamidieh, K., 2018. A data-driven statistical model for predicting the critical temperature of a superconductor. *Computational Materials Science*, *154*, pp.346-354.

[3] Le, T.D., Noumeir, R., Quach, H.L., Kim, J.H., Kim, J.H. and Kim, H.M., 2020. Critical temperature prediction for a superconductor: A variational bayesian neural network approach. *IEEE Transactions on Applied Superconductivity*, *30*(4), pp.1-5.

[4] Lookman, T., Balachandran, P.V., Xue, D., Pilania, G., Shearman, T., Theiler, J., Gubernatis, J.E., Hogden, J., Barros, K., BenNaim, E. and Alexander, F.J., 2016. A perspective on materials informatics: state-of-the-art and challenges. In *Information science for materials discovery and design* (pp. 3-12). Springer, Cham.

[5] Jain, A., Hautier, G., Ong, S.P. and Persson, K., 2016. New opportunities for materials informatics: Resources and data mining techniques for uncovering hidden relationships. *Journal of Materials Research*, *31*(8), pp.977-994.

[6] Wagner, N. and Rondinelli, J.M., 2016. Theory-guided machine learning in materials science. *Frontiers in Materials*, *3*, p.28.

[7] Dimiduk, D.M., Holm, E.A. and Niezgoda, S.R., 2018. Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering. *Integrating Materials and Manufacturing Innovation*, *7*(3), pp.157-172

[8] Ferguson, A.L., 2017. Machine learning and data science in soft materials engineering. *Journal of Physics: Condensed Matter*, *30*(4), p.043002.

[9] Konno, T., Kurokawa, H., Nabeshima, F., Sakishita, Y., Ogawa, R., Hosako, I. and Maeda, A., 2018. Deep Learning Model for Finding New Superconductors. *arXiv preprint arXiv:1812.01995*.

[10] JAIN, A,HAUTIER,G,ONG,SP and person K.2016 New opportunities for material informatics resources and data mining techniques for uncovering hidden relationships journal of material research,31(8),pp.977-994

[11] Park, Jihong, et al. "Extreme URLLC: Vision, Challenges, and Key Enablers." *arXiv preprint arXiv:2001.09683* (2020).

[12] E. Connover, "Google moves toward quantum supremacy with 72-qubit computer", *Sci. News*, vol. 193, no. 6, pp. 13, 2018.

[13] https://github.com/krimler/predict_critical_temp_of_superconductors

[14] Nawaz, Syed Junaid, et al. "Quantum machine learning for 6G communication networks: State-of-the-art and vision for the future." *IEEE Access* 7 (2019): 46317-46350