Critical Temperature Prediction for Superconducting Materials

Oussama. A
Université Laval
Quebec, Canada

Ayoub. E
Université Laval
Quebec, Canada

Abstract — This project article presents a data-driven approach to the task of critical temperature (Tc) prediction for superconducting materials. Deep neural network models were used as surrogate models to learn the mapping between the features of known superconducting materials and their respective critical temperature (Tc) observed in laboratory. A dataset of 21 263 superconductors was used with 82 attributes per superconductors, including characteristics of the materials such as the number of elements that compose them, their average atomic weights and the entropy of their atomic masses.

I. INTRODUCTION

A. The Problem of discovering new superconducting materials

In recent decades, the search for superconductors at room temperature has been a top priority for the global scientific community. This is because the heavy cooling devices required to operate today's materials make them impractical for widespread adoption in modern engineering solutions.

The lack of understanding related to the theory has often been credited as one of the main factors slowing down the search for materials with superconductivity at high temperatures. With the emergence of large amounts of data related to the phenomenon of superconductivity, the problem has gradually turned to statisticians and computer scientists specializing in learning models in order to be able to predict the Tc of certain materials.

B. Overview of the Theory

Since the discovery of the resistanceless electrical conduction of Mercury at 4.2K in 1911 by Kamerlingh Onnes, several theories of superconductivity have succeeded. It is often said that the most popular are the London equations, the theory of Ginzberg and Landau, and finally the BCS theory, which is considered the most important with its explanation of phenomena via the formation of Cooper pairs through an atomic lattice. Although considered successful, the BCS theory fails to explain some phenomena of superconductivity at high temperatures such as those observed on materials called Cuprate (copper oxide). Therefore, no precise fundamental theory has been developed to explain the overall behavior of superconductors.

This is where large, rough models can help. If trained on enough data, these can potentially "map" atomic characteristics to the critical temperature of superconductors, providing a tool for the scientist to predict with reasonable accuracy the Tc of a promising material.

II. STATE OF THE ART

In order to solve the task of predicting the critical temperature of superconducting hardware, many authors have used empirical approaches based on a large number of statistical models. In the case of Kam Hamidieh [1], he trained an XG Boost statistical model based on characteristics taken from the chemical formula of the material. These characteristics are mainly based on the

thermal conductivity, atomic radius, electronic structure and atomic mass of the elements present in the superconductor. After training the model on a database of 21,263 materials from the Japanese National Institute of Materials Science (NIMS), the author obtained an out-of-sample RMSE of 9.4 K and an out-of-sample R2 is 0.92.

As for Thanh Dung Le & all [2], they opted for a model called Bayesian neural network using again the characteristics taken from superconducting chemical elements and their respective formula to predict the critical temperature Tc. After training, their R2 was 0.94 and their RMSE was 3.83.

III. DESCRIPTION OF THE DATASET

In order to accomplish the task, a dataset of 21 263 superconductors is made available. It also contains 82 attributes of these superconductors, including critical temperature. Other attributes are characteristics of materials such as the number of elements that compose them, their average atomic weights and the entropy of their atomic masses. These characteristics are available in the csv file located in the code that accompanies this article.

IV. DESCRIPTION OF THE APPROACH USED

To solve this regression problem, the proposed solution is to use a multilayer perceptron with reread activation functions. Indeed, the attributes of the given game do not depend on space (e.g. images) or time (e.g. financial data). Thus, the best deep learning approach to use is the multilayer perceptron. In addition, the reread activation function is very useful to prevent certain problems such as the *vanishing gradient problem*.

Also the dimensionality reduction approach will be studied, a principal component analysis (PCA) will be applied to the dataset. This analysis allows the dataset to move to a reduced dimensional space by combining the attributes to obtain linear combinations of them and potentially increase the accuracy of the model. The architecture of the model is shown in the figure below.

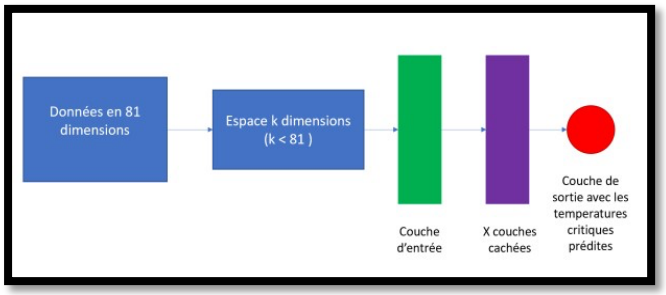

Fig. 1. Architecture of the model used.

In addition, batch normalizations are applied to the outputs of the input layer and hidden layers of the model to allow the model to train faster and achieve better results. At the exit of the last hidden layer, a Dropout layer is added to prevent overfitting of the model. The latter assigns a 20% probability to each neuron of being deactivated. Finally, only the parameters of the multilayer perceptron are trainable, parameters such as the number of dimensions obtained with PCR are hyperparameters that should be determined.

V. METHODOLOGY AND EXPERIMENTATION

The first step is to load the dataset contained in a csv file and isolate the column that corresponds to the critical temperatures of the materials. The dataset is then separated into 3, a training, validation and testing set. For each of these sets, a principal component analysis in k dimensions is performed followed by a standard normalization. This standard normalization is intended to allow all attributes in the dataset to contribute equally to model training.

For model training, the Adam optimizer is used with a learning rate of 0.002. In the first tests, this optimizer performed better than SGD for the same learning rates. The error function used is the average square of errors. This choice of error aims to penalize predictions with a large error thanks to the power of 2 in its formula. The next step is to optimize some of the chosen hyperparameters. These are shown in the following table:

Description hyperparameters	Interval tested
The number of hidden layers in the model.	1 to 5 hidden layers.
Number of input and output neurons from each of the hidden layers of the model.	10 to 200 neurons.
The number of dimensions obtained after principal component analysis of the dataset.	10 to 60 dimensions.

Fig. 2. Table of the names of the hyperparameters to be optimized as well as the chosen test interval.

The optimization of these hyperparameters is done using python's open-source Optuna library. This library uses algorithms such as Tree-structured Parzen Estimator and Grid Search to optimize the search for the best hyperparameters [1]. This hyperparameter search process seeks to reduce the validation error of the last training period. Each training performed during the iterations, has a fixed number of eras and a fixed batch size that are respectively 10 and 32. Finally, after obtaining the optimal hyperparameters, the model is tested on the test dataset while calculating the R-squared coefficient useful for measuring the performance of the regression obtained.

VI. DISCUSSION OF RESULTS

The hyperparameter configuration determined after optimization with Optuna is shown in the following table:

Description of the hyperparameter	Value obtained
The number of hidden layers in the model.	5 hidden layers
Number of input and output neurons from each of the hidden layers of the model.	Layer 1：(input: 20, output: 99) Layer 2：(input: 99, output: 128) Layer 3：(input: 128, output: 43) Layer 4：(input: 43, output: 168) Layer 5 :(input 168：, output: 44)
The number of dimensions obtained after principal component analysis of the dataset.	20 dimensions

Fig. 3. Table of the names of the hyperparameters to be optimized and their values obtained.

This configuration made it possible to achieve a validation error of 256.95 at the last training period. The mean squares of the errors, MSE and the R-squared coefficient of the model in the training and test dataset are shown in the following table:

MSE Train	MSE test	R^2 Train	R^2 Tis
222.13	343.57	0.81	0.71

Fig. 4. Table of average squares of errors and coefficients of determination in training and test set.

An MSE of 343.57 represents an average deviation of 18 Kelvin from the model's true critical temperature. This discrepancy may seem very significant. However, despite this uncertainty, the model makes it possible to identify interesting materials with which scientists can carry out more rigorous experimental tests.

The R-squared coefficient obtained shows that there is a correlation between the model's predictions and the true critical temperatures. It also shows that there is no significant variance between true and predicted values. The model error is mainly due to bias in the bias-variance dilemma. Ways to improve the model can be explored. First, it is possible to reduce the dimensions of the dataset using other methods such as the t-SNE algorithm. Also, more combinations of hyperparameters can be studied if computing power is available.

VII. CONCLUSION

In conclusion, the model obtained allows an estimate of the critical temperature of materials with an average uncertainty of about 18 Kelvin. Accelerating the discovery and production of superconducting materials can lead to great advances, which add great value to society, such as the installation of lossless electrical transmission cables or the production of strong magnetic fields used for nuclear fusion, particle accelerators, or simply the "MRI machines" used in hospitals.

VIII. REFERENCES

[1] Kam Hamidieh, "A Data-Driven Statistical Model for Predicting the Critical Temperature of a Superconductor", July 2019.

[2] Thanh Dung Le, Rita Noumeir, Huu Luong Quach, Ji Hyung Kim, Jung Ho Kim, and Ho Min Kim, "Critical Temperature Prediction for a Superconductor: A Variational Bayesian Neural Network Approach", Jan 2020.

[3] Takuya Akiba1, Shotaro Sano, Toshihiko Yanase1, Takeru Ohta1, and Masanori Koyama1, "Optuna: A Next-generation Hyperparameter Optimization Framework", July 2001