Predicting Superconducting Critical Temperature Using Regression Analysis

Miguel Ramirez
Department of Statistics and Data Science
University of Central Florida
Orlando, United States
miramirez@knights.ucf.edu

***Abstract*—This project develops a regression model to predict the superconducting critical temperature based on features extracted from the chemical formula. We evaluate linear regression models, apply stepwise feature selection, and compare predictive performance.**

***Index Terms*—Superconductor, regression, machine learning, UCI dataset**

I. Introduction

A superconductor is a material that allows electricity to flow without resistance. The temperature at which this occurs is called the critical temperature (T_c). Predicting T_c from a compound's chemical composition is a long-standing challenge in materials science.

II. Data

We used the Superconductivity Data Set from the UCI Machine Learning Repository (https://archive.ics.uci.edu/dataset/464/superconductivty+data), which contains 21,263 superconductors, 81 predictor variables, and one target variable (T_c).

III. Methodology

We fit a linear regression model:

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \epsilon \tag{1}$$

where Y is the transformed T_c and X_i are the predictor variables.

A. Model Selection

We applied stepwise regression using BIC to reduce the number of predictors.

IV. Results

Table I shows performance metrics for the final regression model.

TABLE I
Model Performance

Metric	Train	Test
R^2	0.78	0.77
MSE	2.51	2.56

V. Conclusion

The regression model explained about 78% of the variation in T_c. Key predictors included atomic mass, valence, and thermal conductivity.

References

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