Introduction

Superconductors are a material that under a certain critical temperature, the material is able to conduct current with zero resistance. This makes superconductors very favorable in certain applications. The most common application is the use of them in fMRI machines used by the healthcare industry, as well as various research facilities, such as the Large Hadron Collider.

Since the discovery of superconductors in 1911 by Heike Kamerlingh Onnes, there have been two main limitations that prevent more abundant use of superconductors [1]. One being that superconductors favorable properties emerge only at very low temperatures making it challenging to keep the material under its critical temperature. The other issue is that there is no current model to predict the critical temperature of superconductors.

The goal of this project is to attempt to apply machine learning methods in order to predict the critical temperature of a superconductor given its chemical formula.

The data will be accessed from the Superconducting Material Database maintained by Japan’s National Institute for Materials Science (NIMS) [2]. There are 21,263 superconductor entries with 81 variables per entry. These variables will be features of chemical properties of the material such as atomic mass, thermal conductivity, or number of valance electrons.

Initially, dimensionality reduction techniques will be used to determine if all 81 inputs are required to accurately predict the critical temperature, and if not, how many will be sufficient. Neural Networks and Regression will be utilized to attempt to predict the critical temperatures of superconductors, and the two models will be compared in accuracy, efficiency, and reliability.

Dataset

The dataset used for this project was borrowed from 2018 paper titled, “A Data-Driven Statistical Model for Predicting the Critical Temperature of a Superconductor” [3]. It contains 21,263 rows and 81 input columns. The 81 input columns are made up of various feature extractions of each semiconductor material. Eight element properties were used to create features. The eight properties are atomic mass, first ionization energy, atomic radius, density, electron affinity, fusion heat, thermal conductivity, and number of valence. For each of the eight properties, ten methods were applied to extract features.

The following are the ten ways that the element data is manipulated.

*Mean:* (1)

*Weighted Mean:* (2)

*Geometric Mean:* (3)

*Weighted Geometric Mean:* (4)

*Entropy:* (5)

*Weighted Entropy:* (6)

*Range:* (7)

*Weighted Range:* (8)

*Standard Deviation:* (9)

*Weighted Standard Deviation:* (10)

Where,

t­i is an element’s given property,

pi is the ratio of an element in a material,

is the fraction of a given property of a material,

A

B =

This is how 80 input columns were generated with the last one being the number of elements in the material. The element data was gathered using ElementData from Mathematica, and the superconductors material information was obtained from the Superconducting Material Database supported by NIMS.

**Problem Statement**

**(Optional) Data Source**

**Methodology**

**Evaluation and Final Results**

R^2, MAE, RMSE

**References**

[1] ‘Superconductivity’ CERN. Accessed on: https://home.cern/science/engineering/superconductivity

[2] MDR SuperCon Datasheet. National Institute for Materials Science Available: https://mdr.nims.go.jp/collections/5712mb227

[3] Hamidieh, Kim. (2018). ‘A Data-Driven Statistical Model for Predicting the Critical Temperature of a Superconductor’. Available: https://arxiv.org/abs/1803.10260