ISSN: 2995-3286

Review Article

International Journal of Media and Networks

Predict Unknown Properties of Elements of Periodic Table with Machine Learning

Yash K. Bhatia, Maxwell Gao, Sruthi Ainapurapu, Abhilash Akula, Elizabeth A. Christophy, Payam Norouzzadeh, Steven Buckner and Bahareh Rahmani*

Saint Louis University, St. Louis, MO

*Corresponding Author
Bahareh Rahmani, Saint Louis University, St. Louis, MO.

Submitted: 2024, Mar 21 **Accepted**: 2024, Apr 12 **Published**: 2024, Apr 24

Citation: Bhatia, Y. K., Gao, M., Ainapurapu, S., Akula, A., Christophy, E. A. et al. (2024). Predict Unknown Properties of Elements of Periodic Table with Machine Learning. *Int J Med Net, 2*(4), 01-04.

Abstract

The periodic table of elements includes 92 elements with many unknown properties like melting point, boiling point, heat of vaporization, and molar heat capacity of some specific elements such as Curium, Berkelium, Californium, and Einsteinium. Physicists, chemists, and other scientists have done many successful experiments to predict these mysterious features using the first principal methods. But still many properties have been unclear.

In this project we apply Machine Learning models such as linear and logistic regression to predict this unknow values. The known values split to train and test data to find and confirm the model. Then the model will be run over unknown variables.

1. Introduction

In recent past, there have been several applications of machine learning to the furthering of scientific understanding. Specifically, there is interest in using the predictive power of machine learning toward predicting qualities of elements in the periodic table. Mainly, these studies attempt to predict the stability of various elements or identify and utilize the trends that define the structuring of the elements in the periodic table to reconstruct the table using machine learning.

Dr. Zhang and his research group predicted the structure of the periodic table using convolutional neural networks to forecast chemical information [1]. Other researchers recreated the periodic table and prediction of elemental stability using an unsupervised machine learning algorithm or a SOAP kernel [2,3]. Dr. Shayue investigated the potential energy surfaces of atoms using a graph neural network [4]. On the other project, Dr. Tavazza and his research group analyzed the crystal structures using unified atomistic line graph neural network-based force fields [5]. All of which have important applications to the scientific community in the development of chemical reaction efficiency and yield, materials discovery, and catalyst improvement [6].

Predicting chemical information can be extremely valuable to scientists in laboratory research attempting to create compounds for pharmaceuticals, efficient fuels, or any number of applications. Knowing the stability and properties of elements can better inform decisions and directions of scientific inquiry. And, as stated above, while such things as stability and some properties have been assessed using machine learning techniques, what has yet to be well studied or modeled is the use of periodic table trends to predict the magnetic properties of elements.

Magnetism is a categorical property where elements are diamagnetic (no magnetic moments in atoms), Paramagnetic (Random magnetic moment orientation), ferromagnetic (parallel and aligned magnetic moments), or anti-ferromagnetic (antiparallel and aligned magnetic moments). But this property has generally only been tested in laboratory settings which limits the availability or usability of such a property to elements that are easily handled or stable enough to be tested. Such requirements become unreachable in newer elements that are unstable. So, to predict the magnetic properties of elements using machine learning models, other known characteristics can be used based on their predictive value such as melting point, boiling point, density at standard temperature and pressure, triple point, heat of fusion and vaporization, and molar heat capacity.

Dr. Wang and his research group applied ML-guided framework to explore magnetic properties of P-table elements [7]. In another study, Dr. Kioseoglou, et all, applied machine learning methods to Density Functional Theory simulation data to find the relationship between elements' structure and magnetization [8]. Mohanty, et all, used regression and neural networks to find magnetic elements of elements from microstructures [9].

In this paper, more accessible techniques of machine learning such as linear regression and logistic modeling have been used to fill in the missing values of periodic table characteristics and predict the magnetic properties of elements.

2. Data Description

This data from GitHub (https://gist.github.com/

GoodmanSciences/c2dd862cd38f21b0ad36b8f96b4bf1ee) is the most fundamental information about the elements. It includes atomic numbers, element's name, symbol, the number of protons and electrons in a neutral atom. The number of neutrons is for the most abundant isotope, and the number of naturally occurring isotopes is also noted.

The elements are further classified in this data by position on the periodic table. The three general categories of elements are metals, nonmetals, and metalloids. Metals have high melting and boiling points. They are malleable, ductile and conduct electricity and heat. Nonmetals with low melting and boiling points, are brittle and do not conduct electricity and heat. Metalloids are a transition between the two. Elements are classified by their position on the periodic table. The alkali metals are the first column, the alkaline earth metals are the second column. The Noble gases are the last (18th) column, and the halogens are in column 17. Elements in the same column have similar properties. For example, alkali metals and halogens are known to have a high reactivity and are never found free in nature. By contrast, Noble gases characteristically do not react with other elements. These reactivity patterns relate to the date of discovery that is also noted in the table. The elements found in prehistory tend to be low reacting elements, such as transition metals, which are often found to be pure elements. The noble gases were among the last elements to be found, due to their inability to react. Most were found in the 18th and 19th centuries, when active searches for elements were common. The artificial elements were created starting in the atomic era and this experimentation is on-going, with the latest elements being synthesized early in the 21st century.

It is well known that the data in a periodic table have patterns. The atomic radius, electronegativity and first ionization energy vary periodically with atomic number. Atomic radius refers to the size of the atom. The size decreases across a period and increases down a group due to the increasing number of protons in the nucleus and the increasing number of rings of electrons. Electronegativity is a measure of how easy it is for an atom to gain an electron from another element. It increases across a period from a more positive nucleus, and decreases down a group, due to the atomic radius. First ionization energy takes to pull a valence electron from a neutral atom. First ionization energy also increases across a period, due to increasing positive charge from the nucleus, and decreases down a group due to the increasing distance between the valence electrons and the positive charge of the nucleus.

The number of valence electrons is the number of electrons in the outermost shell of the atom and varies from 1 to 8. The number of valence electrons is also a periodic property. It increases from 1 to 8 across a period, then repeats for each one. It is not listed for transition metals, lanthanides, or actinides because the valence electrons for those elements are variable. The "number of shells" column refers to the number of "rings" of electrons in the solar system model of the atom and corresponds to the period of the element.

3. Methodology and Results

We apply 2 regression models to predict the unknown magnetic field of elements. For this aim, we must fill in the blanks in our dataset. We split the known variables to train and test data. Linear regression finds the model of train set. Then we apply the model over test data to confirm the high accuracy of model. We apply Logistic Regression model over completed data to predict unknown magnetic properties of element.

3.1 Linear Regression

In this section we try to fill in missing values in each dataset using a regression model, specifically a Linear Regression model. Regression is a statistical method used in machine learning that attempts to predict a dependent variable based on the values of one or more independent variables. In other words, it is used to understand the relationship between variables. Linear Regression aims to model the relationship between two variables by fitting a linear equation to observed data.

The model separates the data with existing values and fits a linear regression model on it. The trained model is then used to predict the missing values. These predicted values are rounded to three decimal places for precision and are used to replace the null values in the original data. Unknown values of melting point, boiling point, heat of vaporization, molar heat capacity and so on will be predicted with this method.

In summary, this script uses a linear regression model to predict and fill missing values for specific columns in a dataset of element properties. This allows for a complete and more useful dataset, as missing values can often lead to inaccuracies or difficulties in further analysis or machine learning tasks. Figure 1 shows the variation of predicted values.

3.2 Logistic Regression

The features mentioned in Fig 1 to be used for training and prediction of unknown magnetic properties of elements. We split the data into training and test sets. The training set consists of rows where the 'magnetic' column has non-null values, while the test set contains rows with null 'magnetic' values.

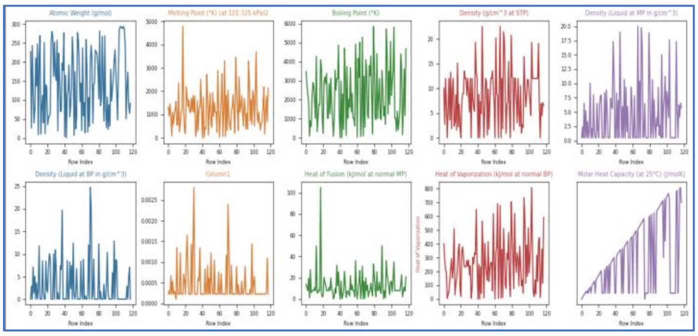

Figure 1: Predicted Values for Different Features

The data is preprocessed using a StandardScaler to standardize the features. Afterward, we predict the missing values in the 'magnetic' attribute using the trained. Table 1 shows the predicted values of some unknown variables using logistic regression. Figure 2 shows the number of predicted magnetic properties. The magnetic properties of all elements have been shown in table 1.

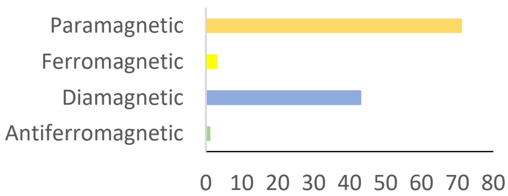

Figure 2: Number of Predicted Magnetism Values

Magnetic Property	Atomic Number	Magnetic Property	Atomic Number	Magnetic Property	Atomic Number	Magnetic Property	Atomic Number	Magnetic Property
Diamagnetic	25	Paramagnetic	49	Diamagnetic	73	Paramagnetic	97	Paramagnetic
Diamagnetic	26	Ferromagnetic	50	Diamagnetic	74	Paramagnetic	98	Paramagnetic
Paramagnetic	27	Ferromagnetic	51	Diamagnetic	75	Paramagnetic	99	Paramagnetic
Diamagnetic	28	Ferromagnetic	52	Diamagnetic	76	Paramagnetic	100	Paramagnetic
Diamagnetic	29	Diamagnetic	53	Diamagnetic	77	Paramagnetic	101	Paramagnetic
Diamagnetic	30	Diamagnetic	54	Diamagnetic	78	Paramagnetic	102	Paramagnetic
Diamagnetic	31	Diamagnetic	55	Paramagnetic	79	Diamagnetic	103	Paramagnetic
Paramagnetic	32	Diamagnetic	56	Paramagnetic	80	Diamagnetic	104	Paramagnetic
Diamagnetic	33	Diamagnetic	57	Paramagnetic	81	Diamagnetic	105	Paramagnetic
Diamagnetic	34	Diamagnetic	58	Paramagnetic	82	Diamagnetic	106	Paramagnetic
Paramagnetic	35	Diamagnetic	59	Paramagnetic	83	Diamagnetic	107	Diamagnetic
Paramagnetic	36	Diamagnetic	60	Paramagnetic	84	Diamagnetic	108	Paramagnetic
Paramagnetic	37	Paramagnetic	61	Paramagnetic	85	Diamagnetic	109	Paramagnetic
Diamagnetic	38	Paramagnetic	62	Paramagnetic	86	Diamagnetic	110	Paramagnetic
Diamagnetic	39	Paramagnetic	63	Paramagnetic	87	Paramagnetic	111	Paramagnetic
Diamagnetic	40	Paramagnetic	64	Paramagnetic	88	Diamagnetic	112	Paramagnetic
Diamagnetic	41	Paramagnetic	65	Paramagnetic	89	Paramagnetic	113	Paramagnetic
Diamagnetic	42	Paramagnetic	66	Paramagnetic	90	Paramagnetic	114	Diamagnetic
Paramagnetic	43	Paramagnetic	67	Paramagnetic	91	Paramagnetic	115	Diamagnetic

Table 1: Predict the unknown magnetic properties of elements including Diamagnetic and Paramagnetic

References

1. Zheng, X., Zheng, P., & Zhang, R. Z. (2018). Machine learning material properties from the periodic table using convolutional neural networks. *Chemical science, 9*(44), 8426-8432.
2. Kusaba, M., Liu, C., Koyama, Y., Terakura, K., & Yoshida, R. (2021). Recreation of the periodic table with an unsupervised machine learning algorithm. *Scientific reports, 11*(1), 4780.
3. Willatt, M. J., Musil, F., & Ceriotti, M. (2018). Feature optimization for atomistic machine learning yields a data-driven construction of the periodic table of the elements. *Physical Chemistry Chemical Physics, 20*(47), 29661-29668.
4. Chen, C., & Ong, S. P. (2022). A universal graph deep learning interatomic potential for the periodic table. *Nature*

Computational Science, 2(11), 718-728.
5. Choudhary, K. (2023). Unified Graph Neural Network Force-field for the Periodic Table. In *APS March Meeting Abstracts* (Vol. 2023, pp. K53-001).
6. Janet, J.P.; Kulik, H.J. Machine Learning in Chemistry. American Chemical Society. 2020, *1*.
7. Xia, W., Sakurai, M., Balasubramanian, B., Liao, T., Wang, R., Zhang, C., ... & Wang, C. Z. (2022). Accelerating the discovery of novel magnetic materials using machine learning–guided adaptive feedback. *Proceedings of the National Academy of Sciences, 119*(47), e2204485119.
8. Katsikas, G., Sarafidis, C., & Kioseoglou, J. (2021). Machine learning in magnetic materials. *physica status solidi (b), 258*(8), 2000600.
9. Talapatra, A., Gajera, U., P, S. P., Arout Chelvane, J., & Mohanty, J. R. (2022). Understanding the magnetic microstructure through experiments and machine learning algorithms. *ACS applied materials & interfaces, 14*(44), 50318-50330.