



---

Skolkovo Institute of Science and Technology

MASTER'S THESIS

## Computational modeling of magnetic materials

Master's Educational Program: Materials Science

---

Student \_\_\_\_\_

Dmitry Volkov  
June 1, 2021

---

Research Advisor: \_\_\_\_\_

Artem R. Oganov  
Full Professor  
Habilitation, PhD

Moscow 2021

All rights reserved. ©

The author hereby grants to Skoltech permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole and in part in any medium now known or hereafter created.



Skolkovo Institute of Science and Technology

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

## Компьютерное моделирование магнитных материалов

Магистерская образовательная программа: Материаловедение

Студент

Дмитрий Волков

Июнь 1, 2021

Научный руководитель:

Артём Ромаевич Оганов

Профессор РАН

д-р физ.-мат. наук

Москва 2021

Все права защищены.©

Автор настоящим дает Сколковскому институту науки и технологий разрешение на воспроизведение и свободное распространение бумажных и электронных копий настоящей диссертации в целом или частично на любом ныне существующем или созданном в будущем носителе.

# **Computational modeling of magnetic materials**

Dmitry Volkov

Submitted to the Skolkovo Institute of Science and Technology  
on June 1, 2021

## **Abstract**

Ferromagnetic (FM) materials became an irreducible part of modern technology. Invariable usage of FM in power generators, engines as well as mass-market consumer products like cellphones, laptops, microphones, and hard drives have led to an increase in the demand for these materials hundreds of times over the past few decades. A vast majority of used now magnets contains rare earth elements (REEs) in its composition. Such structures proved to show extremely high magnetic properties at room temperature within a relatively low price per energy-density unit. But pricing for REEs starts to grow in the last few years rapidly. The reasoning for it is a strong monopoly of China in which, according to the World Trade Organization data, located more than 95% of world REEs mining. This situation is quite unsatisfying for many countries with high demand for these materials. And also raises a great interest of science and industry in discovering new REEs-free FM structures.

Nevertheless, despite several attempts to computationally predict new structures made in the last 25 years, this problem remains unsolved. One reason for it is a limitation of modern algorithms, which allows to exact only elementary properties like stability and magnetic moment, staying absolutely blind to FM critical temperature. Consequently, a huge amount of computational and experimental efforts results in materials with low technological potential due to unsatisfying values of Curie temperature. This problem indicates the need for the development of an effective and automated way for critical temperature estimation, which will empower the computational search of magnetic structures.

To sum up, in this work will be performed a computational search for the novel composition of ferromagnetic materials not containing REEs and estimation of their critical temperature by the newly developed method in order to find the most promising materials from the technological viewpoint.

Research Advisor:

Name: Artem R. Oganov

Degree: Habilitation, PhD

Title: Full Professor

# **Компьютерное моделирование магнитных материалов**

**Дмитрий Волков**

Представлено в Сколковский институт науки и технологий

Июнь 1, 2021

## **Реферат**

В наши дни сложно переоценить роль магнитных материалов в индустриальном и научном развитии. Повсеместное использование ферромагнетиков в тяжелой промышленности, цифровых приборах, накопителях (жесткие диски) и медицинском оборудовании (МРТ) делают этот класс материалов критически важным.

Несмотря на это, в последние 25 лет не было сделано никаких значительных открытий в области поиска новых ферромагнетиков. В то же время, наиболее широко используемые на сегодняшний день составы магнитных материалов неизменно включают в себя редкоземельные элементы (РЗЭ). Принимая во внимание, чрезвычайно неравномерное распределение добычи РЗЭ, на более чем 95% сконцентрированной в Китае, сложившаяся ситуация является не удовлетворительной для многих стран, экономика которых напрямую зависит от импорта этих материалов. Совокупность вышеперечисленных фактов ясно обозначает задачу поиска эффективной альтернативы уже существующим магнитам, состав которых не включал бы в себя критические или дорогие элементы.

Безусловно, в последние несколько лет было предпринято несколько попыток систематического поиска новых магнитов с использованием современных вычислительных алгоритмов, но, к сожалению, все разработанные на данный момент методы могут определять только некоторые элементарные свойства материала такие как: стабильность, магнитный момент на ячейку и т. д., при этом не давая данных о критической температуре. Как следствие, большая часть вычислительных ресурсов тратится не рационально: на материалы с неудовлетворительно низкой критической температурой и как следствие малым или отсутствующим технологическим потенциалом. Основываясь на вышесказанном, можно заключить, что на сегодняшний день температура Кюри является чрезвычайно важным параметром при оптимизации в вычислительном поиске новых магнитных материалов и объясняет актуальность разработки эффективного метода ее определения.

В представленной работе будет проведен вычислительный поиск новых структур ферромагнитных материалов, отвечающих обозначенным физическим и экономическим критериям с использованием эволюционного алгоритма. Для обнаруженных структур будет проведен расчет критической температуры с целью их ранжирования и определения наиболее перспективных с технологической точки зрения.

Научный руководитель:

Имя: Артём Ромаевич Оганов

Ученое звание, степень: д-р физ.-мат. наук

Должность: Профессор РАН

## Acknowledgments

I would like to express my gratitude to my research advisor, prof. Artem R. Oganov for the kind support and guidance in this master thesis.

I want to extend my thanks to colleagues from the laboratory Dr. Zahed Allahyari, Dr. Christian Tantardini, and Michele Galasso, who have readily helped me analyze difficult calculations and guide me through this complicated work.

I would like to express appreciation to prof. Andriy Zhugaevich and Dr. Dmitry Aksyonov for the guidance, support, and useful hints on the very first steps of this project.

Also, I want to thank prof. Sergey Levchenko for the guidance and advisory on the part of the project related to machine learning algorithms as well as for the productive discussion.

# Contents

<b>1</b>	<b>Introduction</b>	<b>10</b>
1.1	Background . . . . .	10
1.2	Aim and Goals . . . . .	12
1.3	Outline . . . . .	12
<b>2</b>	<b>Literature review</b>	<b>13</b>
2.1	General Introduction to Ferromagnetics . . . . .	13
2.1.1	Origin of Ferromagnetism . . . . .	13
2.1.2	Types of magnetism . . . . .	14
2.2	Magnetic anisotropy . . . . .	16
2.3	Magnetocrystalline anisotropy . . . . .	17
2.4	Domain and Domain Walls . . . . .	17
2.5	T-symmetry . . . . .	19
2.6	Magnetic models . . . . .	20
2.6.1	Ising model . . . . .	20
2.6.2	Heisenberg model . . . . .	21
2.6.3	Phase Transitions . . . . .	22
2.7	Monte Carlo Simulations . . . . .	23
2.7.1	Monte Carlo method . . . . .	23
2.7.2	Monte Carlo in Statistical Mechanics . . . . .	23
2.7.3	Metropolis algorithm . . . . .	24
<b>3</b>	<b>Materials and Methods</b>	<b>26</b>
3.1	Data driven estimation of critical temperature . . . . .	26
3.1.1	Data and predictors . . . . .	26
3.1.2	Features generation . . . . .	27
3.1.3	Data pre-processing . . . . .	29
3.1.4	Machine-learning models . . . . .	30
3.1.5	Physical understanding of the most important descriptors . . . . .	42
3.2	DFT + MC calculations of critical temperature . . . . .	43
3.2.1	Obtain initial geometry . . . . .	43

3.2.2	Generation of spin configurations . . . . .	43
3.2.3	Energy calculations . . . . .	43
3.2.4	Hamiltonian fitting . . . . .	45
3.2.5	Monte Carlo simulation . . . . .	48
3.2.6	End-to-end pipeline details . . . . .	50
<b>4</b>	<b>Results</b>	<b>52</b>
4.1	ML models results . . . . .	52
4.2	DFT calculations results . . . . .	53
4.3	Novel structures study . . . . .	55
<b>5</b>	<b>Discussion</b>	<b>57</b>
5.1	Results reasoning . . . . .	57
5.2	Future Work . . . . .	58
5.3	Conclusion . . . . .	59
<b>Bibliography</b>		<b>60</b>

# List of Figures

1.1	Development in energy product of permanent magnetic materials . . . . .	11
2.1	Diagram showing the magnetic moment associated with orbital and spin motion of an electron. . . . .	14
2.2	Schematic illustration of various types of magnetism. . . . .	14
2.3	Magnetic properties of pure elements . . . . .	15
2.4	Schematic densities of states (DOSs) for ferromagnetism. . . . .	15
2.5	Illustration of magnetization versus temperature. . . . .	16
2.6	Domain structures in cubic and FM materials. . . . .	19
2.7	Domain wall structures. . . . .	19
2.8	Illustration of time reversion symmetry breaking . . . . .	20
3.1	Critical temperature distribution histogram over 2557 samples . . . . .	30
3.2	Scheme of cross-validation process followed by final model evaluation on the test subset. . . . .	31
3.3	Comparison of experimental and predicted values of Curie temperature for different Kernels. . . . .	34
3.4	Single Tree model with depth 4 for the regression problem under consideration. . . . .	35
3.5	RF model hyper parameters tuning. . . . .	36
3.6	Optimized RF model performance, left comparison of predicted and experimental data. . . . .	36
3.7	XGBoost model hyperparameters tuning. . . . .	38
3.8	Optimized XGBoost performance . . . . .	38
3.9	Correlation matrix for the most valuable descriptors in RF. . . . .	40
3.10	SISSO model performance. . . . .	41
3.11	RF feature importance histogram. . . . .	42
3.12	Schematic determination of $J_1$ and $J_2$ coefficients in 2D case. . . . .	46
3.13	Three different spin configurations produced for EuO. . . . .	47
3.14	Three different spin configurations produced for MnO. . . . .	48
3.15	Time-dependent magnetisation caused by a stepwise increase in temperature. . . . .	49
3.16	MC magnetization temperature curve. . . . .	50

3.17	End-to-end pipeline scheme.	50
4.1	Comparison of calculated by GGA results with experimental.	53
4.2	Comparison of calculated by GGA+U results with experimental.	54
4.3	Histogram of estimated critical temperature using ML and DFT.	56

# List of Tables

2.1	Domain wall thickness . . . . .	18
3.1	Property statistics functions . . . . .	28
3.2	List of descriptors used in this work . . . . .	28
3.3	KRR models performance with respect to different Kernels . . . . .	33
3.4	RF models performance with respect to the features space . . . . .	36
3.5	XGBoost performance with respect to the features space . . . . .	39
4.1	Summary of the best ML models performance . . . . .	52
4.2	Comparison of calculated by GGA and GGA+U results with experimental. . . . .	55

# Chapter 1

## Introduction

### 1.1 Background

Ferromagnetic materials (FM) known to humanity for more than 2000 years. It's been a long time since the first samples of a mineral known as lodestone ( $\text{Fe}_3\text{O}_4$ ) amazed our ancestors by its unique properties. Since then, this class of materials was thoroughly analyzed and described by many generations of bright scientists. But yet even in the 21st century, we still have many questions to answer to fully understand this mysterious phenomenon.

The wide usage of FM materials in industrial applications started in the late 19th century when carbon steels were replaced with Alnico alloys. Further the discovery of the large magnetocrystalline anisotropy in the second part of the 20th century lead to the revolutionary introduction of rare-earth magnets. First produced samples of  $\text{Sm}_5\text{Gd}$  reported in 1960 lead to an industrial breakthrough and further rapid development in this field. From this moment, the remarkable increase in the energy product of magnetic materials was accompanied by the tremendous decrease in their volume, providing the same amount of magnetic energy as it is illustrated in figure 1.1. Progress in this direction ends up with the discovery of the strongest and most usable till now Nb-Fe-B composition [1].

Nowadays, FM are actively used not only in the design of complex technical devices such as engines, turbines, and generators but also as an irreplaceable part of widely used mass-market consumer products. Mobile phones, laptops, hard drives, audio devices, and an endless number of other applications have made this class of materials one of the most sought-after in the present times. As a result, the need for strong and cheap magnets has increased hundreds of times in just the last few decades with a special interest in rare-earth (REEs) magnets possessing excellent properties at room temperature within a relatively low price per energy-density unit.

But at this point, the modern world faced quite a risky situation. Thus, according to the recent data of the World Trade Organization, more than 95% of REEs mining concentrated in China. This monopoly has led to a rapid rise in prices for REEs over the past few years and is undoubtedly unsatisfactory for countries with a high demand for these materials. Moreover, the strong concentration of such a critical resource in one place in the event of unforeseen circumstances (such as COVID-19) can cause a complete interruption in supplies and multibillion money losses for man-

ufacturers around the globe. The possessed situation lead to a great interest of science and industry in discovering new ferromagnetic materials.

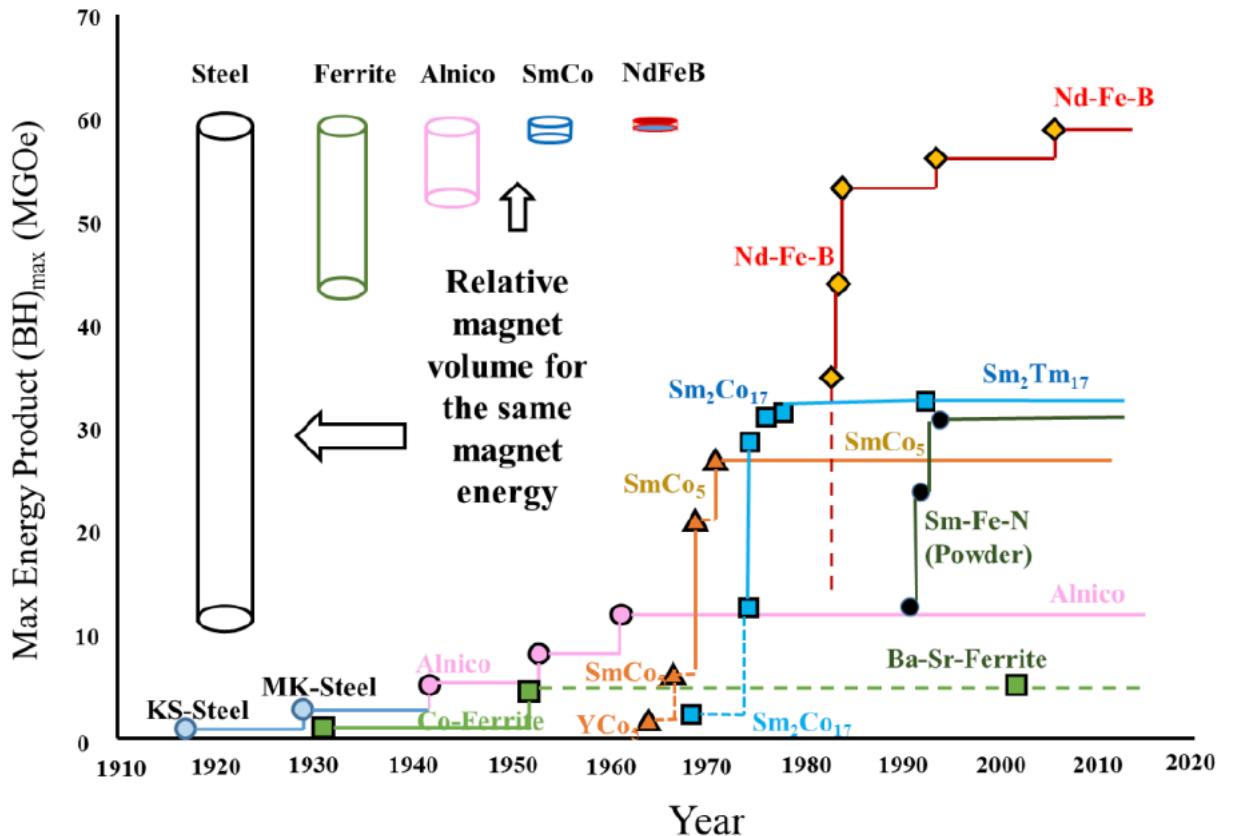


Figure 1.1: Development in energy product of permanent magnetic materials [1].

Nevertheless, despite several attempts to computationally predict new structures made in the last 25 years, this problem remains unsolved. One reason for it is a limitation of modern algorithms, which allows to exact only elementary properties like stability and magnetic moment, staying absolutely blind to FM critical temperature. As a consequence, a huge amount of computational and experimental efforts results in materials with low technological potential due to unsatisfying values of Curie temperature.

## 1.2 Aim and Goals

Based on a brief introduction, the ultimate goal of this study is to search for new technologically promising magnetic compositions. Considering the capabilities of modern evolutionary algorithms that allow identifying stable structures with a high magnetic moment, the problem of automated determination of the critical magnetic temperature (Curie point) remains unsolved. To solve the indicated problem and empower the end-user with a straightforward methodology for its calculation, the following research objectives are identified:

1. Development and implementation of the critical magnetic temperature calculation methodology based on DFT approaches followed by Monte Carlo simulations.
2. Training of several classical ML models to solve a regression problem of critical temperature estimation based on the available experimental data.
3. Testing of the developed methods on compositions with known experimental and theoretical values of Curie temperature.
4. Computational search for new compositions with high magnetic properties using evolutionary algorithm USPEX [2–4].
5. Benchmarking of the newly find structures concerning their critical temperature to determine the most promising from the technological viewpoint.

## 1.3 Outline

Chapter 1 contains the current section with a problem background and formulation of the project objectives. Chapter 2 contains a general overview of the ferromagnetic materials properties and theoretical models used for this study. Chapter 3 describes the methodology used in this work, namely the construction of several machine learning (ML) models over available experimental data and calculations based on density function theory (DFT) followed by Monte-Carlo (MC) simulations. Chapter 4 is dedicated to the discussion of the results obtained by both methods. Finally, chapter 5 summarizes the research outcome and outlines possible future studies.

# Chapter 2

## Literature review

### 2.1 General Introduction to Ferromagnetics

#### 2.1.1 Origin of Ferromagnetism

In all materials electrical, chemical and magnetic properties are determined by the electronic structure. Electrons are distributed in specific shells at definite distances from the nucleus, which is due to the different energy level in each shell. In the FM elements, including  $\alpha$ -Fe, Co, Ni and Gd, there are three important features:

1. There must be an unfilled inner electron shell within the atom.
2. There must be uncompensated electronic spins in the unfilled inner shell.
3. The atoms must form a crystal lattice having a lattice constant at least 3 times the radius of the unfilled electron shell.

Electrons have an intrinsic property called spin that contributes to their magnetic moment. The classical model of electron spin is the electron spinning on its axis, i.e. either spin-up or spin-down. Macroscopic magnetic properties of materials are a consequence of magnetic moments associated with individual electrons, which is the sum of spin moment Figure 2.1a and orbital moment Figure 2.1b.

The magnetic susceptibility  $\chi$  is a dimensionless proportionality constant that indicates the degree of magnetization of a material in response to an applied magnetic field, and is given by:

$$\chi = \frac{M}{H} \quad (2.1)$$

where  $\chi$  is the magnetic susceptibility,  $M$  is the magnetization of the material, and  $H$  is the applied field.

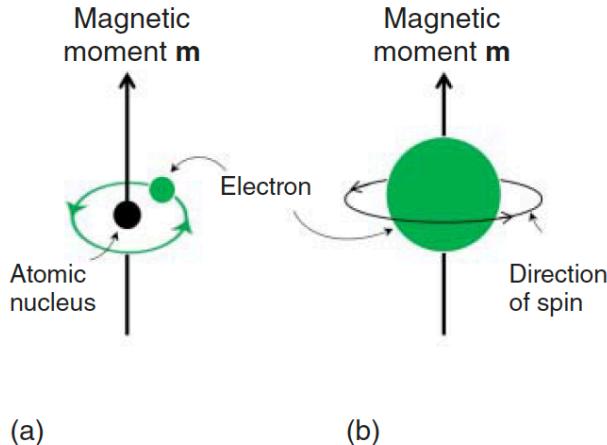


Figure 2.1: Diagram showing the magnetic moment associated with (a) orbital motion and (b) spin motion of an electron.

Depending on the magnetic ordering and the sign, magnitude and temperature dependence of the magnetic susceptibility, the magnetic materials are classified into diamagnetic, paramagnetic (PM), ferromagnetic (FM), antiferromagnetic (AFM) and ferrimagnetic (FiM) materials.

### 2.1.2 Types of magnetism

As it was stated in terms of magnetization, materials can be classified into diamagnetic, PM, FiM, FM and AFM as it is shown in Figure 2.2.

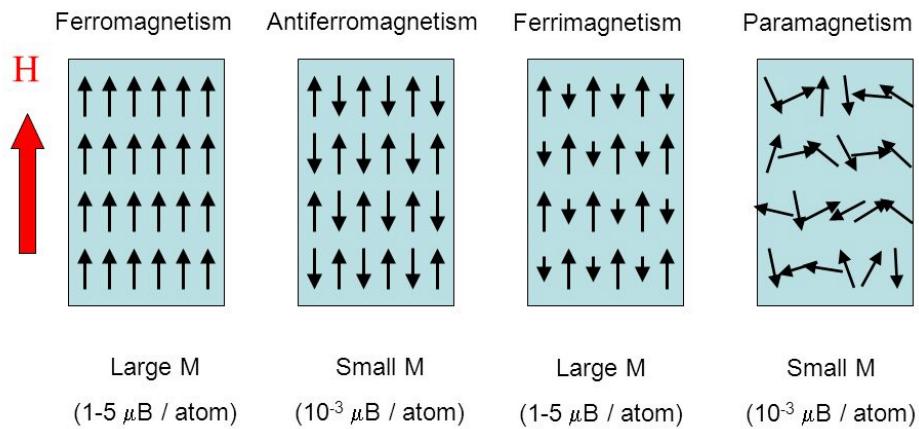


Figure 2.2: Schematic illustration of various types of magnetism.

A FM material usually forms permanent magnet or is attracted to magnets, and it undergoes phase change to PM above Curie temperature ( $T_C$ ) through a second-order phase transition. FM (including FiM) is the strongest type; it is the only type that creates forces strong enough to be felt and is responsible for the common phenomena of magnetism encountered in everyday life. An AFM has magnetization microscopically but has no macroscopic magnetization due to the can-

celation of antiparallel spins of alternative layers. The temperature from an AFM to PM phase transition is called Néel temperature. FiM is a relatively weak FM, such as  $\text{CoFe}_2\text{O}_4$  spinel. A diamagnetic generates weak but negative magnetization when a magnetic field is applied.

Figure 2.3 houses characteristic types of magnetization of pure elements in a solid-state and low temperature.

The periodic table displays the magnetic properties of elements. Elements are categorized into four groups based on their behavior at low temperatures:

- para**: Para-magnetic elements include Li, Be, Na, Mg, K, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, As, Se, Br, Kr, Cd, In, Sn, Sb, Te, I, Xe, Hg, Tl, Pb, Bi, Po, At, Rn, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, Tb, Pa, U, Np, Pu, Am, Cm, Bk, Cf, Es, Fm, Md, No, Lw.
- dia**: Dia-magnetic elements include Be, O, Al, Si, P, S, Cl, Ar.
- AF**: Antiferromagnetic elements include Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, As, Se, Br, Kr, Cd, In, Sn, Sb, Te, I, Xe, Hg, Tl, Pb, Bi, Po, At, Rn, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, Tb, Pa, U, Np, Pu, Am, Cm, Bk, Cf, Es, Fm, Md, No, Lw.
- He**: Helium is listed as a diamagnetic element.

Figure 2.3: The magnetic properties of pure elements in a solid state at low temperature.

A magnetic solid, made up of atoms with magnetic moment, has quantum exchange interactions that tend to align the magnetic moments at low temperature. When  $T < T_C$ , the macroscopic magnetization arises and retains even in the absence of a magnetic field. The magnetic moments tend to align in the same direction without the aid of an external magnetic field. This is known as the FM phase. In the FM category, materials are divided into strong and weak ferromagnets. By introducing the total particle number  $N = N \uparrow + N \downarrow$  and the spin polarization  $s = N \uparrow - N \downarrow$ , a strong ferromagnet has almost  $s = 100\%$  at the Fermi energy as shown in Figure 2.4a, while a weak ferromagnet gets a smaller spin polarization and paramagnet has no net spin polarization Figure 2.4 (b, c).

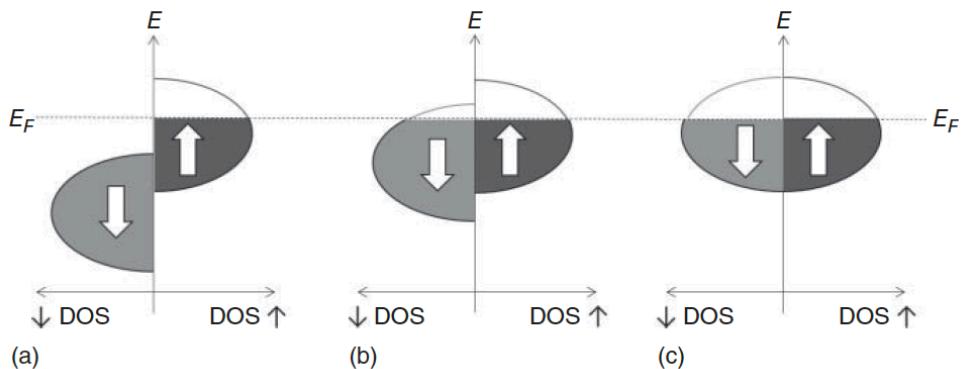


Figure 2.4: Schematic densities of states (DOSS) for FM: (a) strong FM, (b) weak FM, and (c) PM.

Meanwhile, there is another type of magnets that possess magnetization microscopically,

but have no macroscopic magnetization due to the cancelation of antiparallel spins of neighboring pairs – this is known as AFM phase. A magnet that exhibits no macroscopic magnetization at high temperature (when  $H = 0$ ) is known as the PM phase, in which the magnetic moment induced by the applied H-field is rather weak. Figure 2.5 shows the relationship between magnetization and temperature of a FM, where one can see that the magnetization starts decreasing close to  $T_C$ , and when  $T > T_C$ , magnetic moments align randomly resulting in zero macroscopic magnetization.

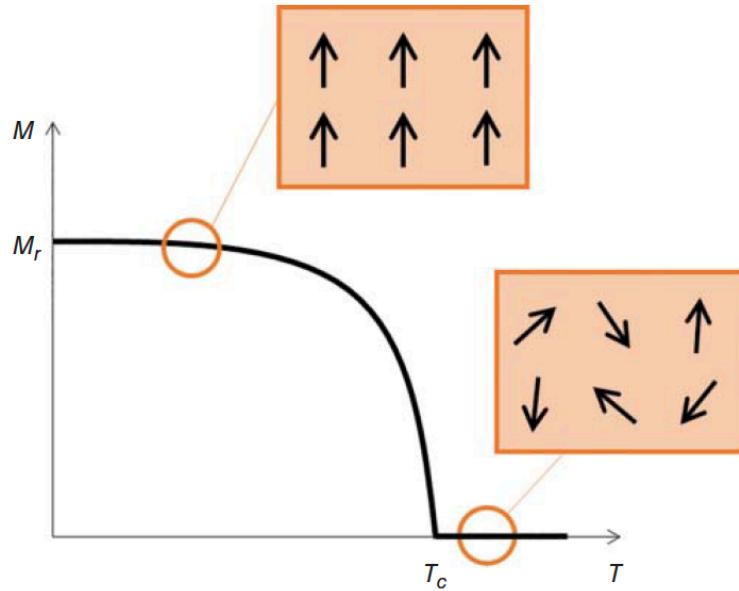


Figure 2.5: Illustration of magnetization versus temperature.

Any atoms or ions with existing unpaired electrons exhibit ferromagnetism such as Fe, Co, and Ni. Magnetization can exist, even very weak, in non-FM materials such as graphene as long as their defects or edge structures induce dangling bonds with unpaired electrons [6, 7]. However, those FM materials with known practical application potentials are usually compounds containing Fe, Co, or Ni.

## 2.2 Magnetic anisotropy

Magnetic anisotropy is defined as the directional dependence of the magnetic properties for materials. Specifically, the preferential direction for its magnetic moment in the absence of an applied magnetic field. Strong easy-axis anisotropy is a prerequisite for hard magnetism, while near-zero anisotropy is desirable for soft magnets. Generally, the tendency for magnetization to lie along an easy axis is represented by the energy density term:

$$E_a = K_1 \sin^2 \theta \quad (2.2)$$

where  $\theta$  is the angle between the magnetic field and the anisotropy axis, and  $K_1$  is the anisotropy constant, which ranges from  $1kJ/m^3$  to more than  $20MJ/m^3$ . There are several sources of magnetic anisotropy:

- Magnetocrystalline anisotropy: intrinsic property due mainly to spin-orbit coupling;
- Shape anisotropy: induced by the nonspherical shape of the grains;
- Stress anisotropy: created by applied mechanical stress due to the existence of magnetostriction, which could alter the domain structure;
- Exchange anisotropy: occurs when the interaction between antiferromagnet and a ferromagnet occurs at their interface;
- Anisotropy induced by grain alignment and stress through magnetic annealing, irradiation, and plastic deformation.

## 2.3 Magnetocrystalline anisotropy

The magnetocrystalline anisotropy primarily arises from spin-orbit coupling. When an external field tries to reorient the spin of an electron, the orbit of that electron also tends to be reoriented. But the orbit is strongly coupled to the lattice and therefore resists the attempt to rotate the spin axis. The energy required to rotate the spin system of a domain away from the easy direction, anisotropy energy, is the energy required to overcome the spin-orbit coupling. The strength of the magnetocrystalline anisotropy in any particular crystal is measured by the magnitude of the anisotropy constant  $K_1$ ,  $K_2$ , etc.

The magnitude of the magnetocrystalline anisotropy generally decreases with temperature more rapidly than the magnetization vanishes at the Curie temperature. Since the anisotropy contributes strongly to the coercive field, it has a great influence on industrial uses of FM materials. Materials with high magnetic anisotropy usually have high coercivity; that is they are hard to demagnetize. The high anisotropy of REEs metals is mainly responsible for the strength of widely used magnets containing them.

## 2.4 Domain and Domain Walls

In FM, materials, exchange, and dipolar couplings are competing. The equality between both interactions leads to a state in which the local macroscopic magnetization is extremely small. Still, an almost parallel alignment of the moments is preserved at short distances. In the absence of magnetocrystalline anisotropy, a gradual rotation of the moments takes place. Overall, the material splits

into a number of isolated zones, called magnetic domains. Within each domain, the moments are parallel as required by coupling, the magnetization is directed along a precise so-called easy direction, as required by anisotropy. From one zone to another, the direction of magnetization changes so that the total magnetization disappears. Between two neighboring domains, a gradual rotation of the moments takes place which defines a Bloch wall. A Bloch wall has a special parameter characteristic thickness,  $\delta$ , for which its value of energy per unit area,  $\gamma$ , is minimum.  $\delta$  and  $\gamma$  are given by :

$$\delta = \pi \sqrt{A/K} \quad (2.3)$$

$$\gamma = 4\sqrt{AK} \quad (2.4)$$

here  $A$ , the exchange constant expressed in  $J/m$ , characterises the strength of the exchange coupling and  $K$  is the anisotropy constant expressed in  $J/m^3$ . Most typically  $\delta$  varies from 50 nm in the 3d metals to 3 nm in high-anisotropy compounds ( $\text{SmCo}_5$ ,  $\text{Nd}_2\text{Fe}_{14}\text{B}$ ).

Table 2.1: Domain wall thickness

Magnetic material	Domain wall width (nm)
Co	14
$\text{Nd}_2\text{Fe}_{14}\text{B}$	3.9
$\text{Sm}_2\text{Fe}_{17}\text{N}_3$	3.6
$\text{Sm}_2\text{Co}_{17}$	8.6
FePd	11.4

In the general case, splitting of a FM material into domains separated by walls occurs spontaneously. In a homogeneous material, the wall energy does not depend on the wall position. Under an applied external field  $H$ , the initial magnetization variation is to minimize the Zeeman energy and the dipolar energy. In a system of cubic symmetry, the susceptibility is equal to the inverse of the demagnetizing field slope; it is defined by the sample shape and not connected to the intrinsic material properties. During this process, the magnetization variation occurs by growing the domains aligned by the applied field at the expense of others. The material is in the single domain state at a larger field, and magnetization variation occurs by moment rotation.

FM materials also exhibit domain structure where each domain has its own magnetization direction that can be switched by external magnetic field. There are easy axes in FM materials that are the directions magnetic moments should follow. For example, for a cubic structured FM material such as Fe, the  $\langle 100 \rangle$  directions are usually the easy axes, while for a hexagonal-structured FM material such as Co, the  $\langle 0001 \rangle$  are the easy axes. Due to this difference, in a cubic structure, both  $90^\circ$  and  $180^\circ$  magnetic domains can be formed (Figure 2.6a), while in a hexagonal structured

FM material, the domains are usually aligned with an angle of  $180^\circ$  (Figure 2.6b).

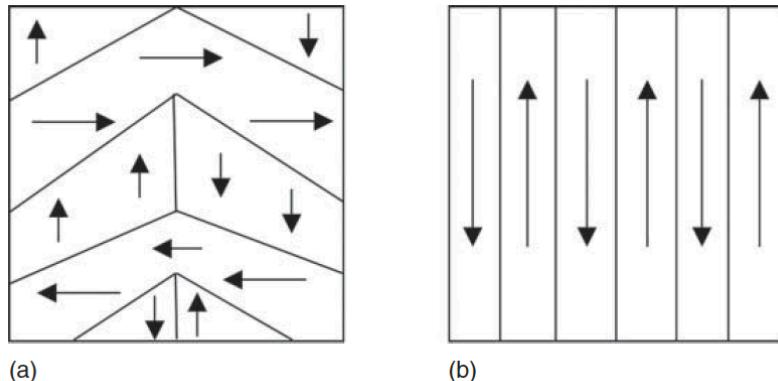


Figure 2.6: Domain structures in cubic (a) and hexagonal (b) structured FM materials.

The easy axes of cubic-structured Ni are  $\langle 111 \rangle$ , so it can form  $180^\circ$ ,  $71^\circ$ , and  $109^\circ$  magnetic domains; these are possible angles between all  $\langle 111 \rangle$  easy axes. The magnetization switching can be realized by not only external magnetic field but also mechanical stress.

FM domain wall is only one to a few atomic layers. Figure 2.7a illustrates the  $180^\circ$  domain wall structure in FM materials, where the more common one is the Bloch wall, but in thinner films a Néel wall is often favored Figure 2.7b. By contrast, Figure 2.7c demonstrates a narrow domain wall (Ising type) structure usually as the case of a ferroelectric domain structure.

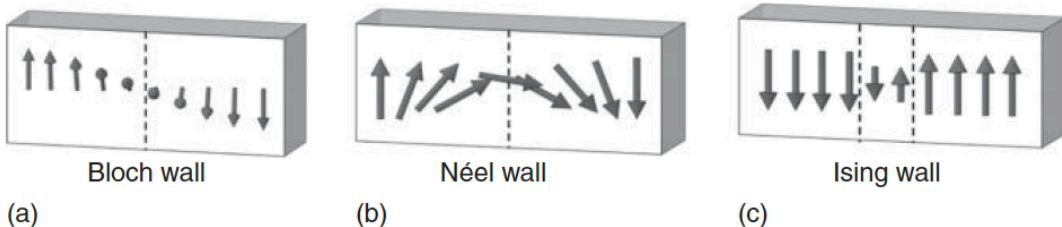


Figure 2.7: Domain wall structures of (a) Bloch-type FM domain walls, (b) Néel-type FM domain walls, and (c) Ising-type ferroelectric domain walls.

## 2.5 T-symmetry

Generally T-symmetry or Time Reversal symmetry is the symmetry of physical laws under a time reversal transformation ( $\tau$ ).

$$\tau : t \rightarrow -t \quad (2.5)$$

The origin of ferromagnetism (FM) is time reversal symmetry breaking due to the existence of unpaired spin of electrons and the associated current [5]. This statement can be understood by

considering that if time is reversed, the spin or current flow direction will thus be reversed and this results in magnetization direction reversion, i.e. time reversal symmetry breaking.

$$\tau : \mu = -\mu \quad (2.6)$$

In other words as it is illustrated on Figure 2.8 the local magnetic moment  $\mathbf{m}$  may be represented classically by a charge that dynamically traces an orbit, as indicated by the arrowheads. A spatial inversion produces no change, but time reversal switches the orbit and thus  $\mathbf{m}$ .

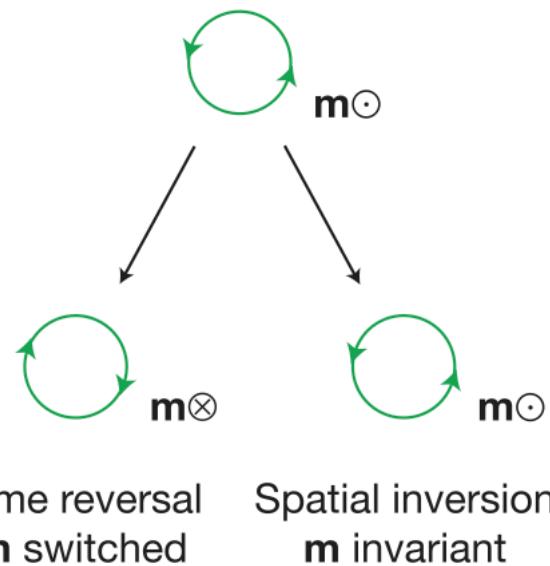


Figure 2.8: Illustration of time reversion symmetry breaking [5]

Thus, magnetic ordering always breaks at least one symmetry of the crystal, the invariance under time inversion. Invariance under rotations around axes not coinciding with the ordered moments also disappears. This symmetry braking is not related to asymmetries in the spin Hamiltonian. Spin Hamiltonians are constructed to describe the behaviour of the degrees of freedom of atoms or ions occupying lattice sites. Therefore, they have the full symmetry of the lattice.

## 2.6 Magnetic models

### 2.6.1 Ising model

Ising or 1/2-spin model is known as one of the most famous and simplest theoretical representations of magnetic systems. Despite this fact, this model has proven to be convenient for the descriptions of rather complicated interactions giving quite accurate results and being pretty useful as a test case for new approximations in systems of interacting particles even not related to magnetism. For

instance, presently, it is applied to study liquids freezing and evaporation, ordered-disordered transformation in alloys, the behavior of glassy substances, and even active form of protein molecules.

In the framework of original model, every discrete spin of the lattice is presented as +1 or -1. For the case of a one-dimension, it is a chain of spins, every of which has only nearest-neighbor interaction. Ising himself rigorously solved this model in 1924 in his PhD work.

In two dimensional case, this model is a square lattice with an equal number of spins in two directions. Each spin interacts with two nearest neighbors in both lateral axes (4 in total). This model was intensively studied theoretically by various methods like Green functions, mean-field theory, high and low-temperature series expansion, transfer matrixes but so far was solved only for the case of an absent external magnetic field. Cases of two-dimensional models with applied external field or simplest examples of the three-dimensional magnetic model remain unsolved.

The most common representation of Ising model for ferromagnetic material in applied external field  $H$  given by following Hamiltonian:

$$\mathcal{H}_{Ising} = - \sum_{i,j} J_{ij} S_i S_j - H \sum_i S_i \quad (2.7)$$

Here  $J_{ij}$  is exchange interaction between spins in different lattice sites  $S_i$  and  $S_j$  respectively.

In a framework of this model positive value of nearest-neighbors exchange coupling  $J_{ij} > 0$  represents a ferromagnetic system, negative  $J_{ij} < 0$  antiferromagnetic, and zero  $J_{ij} = 0$  nonmagnetic as spins don't interact with each other.

Typically this model might be additionally modified to represent more complex systems with an addition of coupling with next nearest-neighbors or dipole interaction.

## 2.6.2 Heisenberg model

More complex Heisenberg spin model related to prototypical Ising theory has a form of following Hamiltonian:

$$\mathcal{H}_{Heisenberg} = - \sum_{i,j} J_{ij} S_i S_j - H \sum_i S_i - K \sum_i (S_i^z)^2 \quad (2.8)$$

In this case variable  $S_i$  related to spin already represents 3D vectors with  $|S_i| = S$ . While  $K$  denotes a magnetocrystalline anisotropy of a single ion in the system.

In the case of extreme anisotropy, we can find the Hamiltonians which define two important models.

In a case of extremely high anisotropy when the limit of infinite single-ion anisotropy  $K \rightarrow \infty$ . All the magnetic moments tend to align to the point in one direction and Heisenberg model reduces to already known to us, Ising model.

On the opposite case when  $K \rightarrow -\infty$ . All the spins tend to lie in one plane and Heisenberg model reduces to XY model.

Hence Ising model represents the spin variety in a single dimension, XY model allows spin rotation in 2D case, and the Heisenberg model denotes spin variety in tree dimensions. Therefore varying value of anisotropy in a Heisenberg spin model we may allow system to approach behavior typical for XY or Ising model without any contradictions.

As the Ising and XY models exclude some components of the spin vectors, the entities they describe can be looked upon as one- and two-dimensional moments, respectively. However, this does not automatically make them low-dimensional models. The dimensionality is determined by the meaning of the indices  $i$  and  $j$ . Those indicate sites in a lattice of well defined dimension. Thus one can study three-dimensional Ising models or one-dimensional XY models; there is also no contradiction in terms here.

### 2.6.3 Phase Transitions

In thermodynamics phase transition defined as a physical process at which system order parameter undergoes changes from zero in one state to non-zero in another. The behavior of order parameter as a function of some variables like temperature determine a certain point in which systems transits from one phase to another.

The first and second types of phase transition are distinguished by the type of change in the order parameter. In the first type of transition, the order parameter will intermittently change from zero to a certain value, while the system can be in both states simultaneously. In the second type, also known as a continuous phase transition, the order parameter will gradually change, reaching a certain value.

The order parameter for ferromagnetic systems is the magnetization which is continuously falls to zero while the system reaches a its critical temperature ( $T_C$ ) also known as Curie temperature.

Continuous phase transition under consideration is also might be characterized by a series of quantities called critical exponents, denoting their relations with a critical point. In magnetic systems these exponents difend for zero field specific heat, magnetization and susceptibility by following equation:

$$C \propto |T - T_C|^{-\alpha}, \quad (2.9)$$

which characterize the behavior of a ferromagnetic system specific heat  $C$  near the Curie point. This exponent are interesting because of universality: systems that appear different but have a few essential properties in common possess the same critical exponent. A universality class

comprises systems which exhibit this behavior, e.g. certain ferromagnetic systems and the liquid-gas phase transition for fluids. The thermodynamic properties of the system would seem to depend only on the few parameters, such as dimensionality and symmetry

## 2.7 Monte Carlo Simulations

### 2.7.1 Monte Carlo method

Monte Carlo method have proven effective in a wide variety of areas. For example, today, this method is used to predict financial markets and risk analysis, modeling biological processes, studying black holes, and solving differential equations, artificial intelligence, chemistry, and so on. Despite the complexity of the stated problems, this method is based on fairly simple and intuitive principles, which explains its high popularity.

The method can be summarized as follows: a set of random values is generated for a target random variable, and then the required values are calculated on its basis. The modern version of the method was formed in the framework of the Manhattan Project, where it was used to simulate the distances that neutrons can travel in various materials. The idea of modeling based on the generation of a set of random values has already existed for some time before, but it was specially developed with the creation of the atomic bomb and then spread to many other luckily more peaceful areas of knowledge.

The big advantage of Monte Carlo method is that it allows for the element of randomness and complexity of the real world to be taken into account in the model. In addition, the method is robust concerning changes in various parameters, such as the distribution of a random variable.

### 2.7.2 Monte Carlo in Statistical Mechanics

From the viewpoint of statistical mechanics, Monte Carlo method allows computing the internal energy of the system based on its particle states. Hence from the Maxwell-Boltzman statistics, the probability of the system to be in a specific state  $a$  can be written as follows:

$$p_a = \frac{1}{Z} e^{-\beta E_a}. \quad (2.10)$$

Here  $\beta = \frac{1}{kT}$ ,  $T$  -temperature,  $E_a$  - internal system energy,  $Z$  - partition function defined as:

$$Z = \sum_a e^{-\beta E_a}. \quad (2.11)$$

In a specific case of ferromagnetic system study the expectation value (avarage over different system states) of its magnetization might be written as:

$$\langle m \rangle = \frac{1}{Z} \sum_a m_a e^{-\beta E_a}. \quad (2.12)$$

In the same fashion, we can also write an equation for the energy of the system.

$$\langle E \rangle = \frac{1}{Z} \sum_a E_a e^{-\beta E_a}. \quad (2.13)$$

### 2.7.3 Metropolis algorithm

Despite the apparent simplicity of the Ising spin model described in section 2.6 in practice, it is often difficult to simulate numerically for the systems with a high number of states. For instance considering Ising model for the system with  $N$  sites of the lattice and 2 possible individual spin values  $S \in \{-1, +1\}$  its ends up with  $\{-1, +1\}^N$  or simply  $2^N$  possible states. Shown level of complexity motivates the application of Monte Carlo techniques in the simulation in a framework of Ising model.

Used in this case Hamiltonian has a slightly simplified form:

$$\mathcal{H}(a) = -J \sum_{i,j} S_i S_j - H \sum_i S_i \quad (2.14)$$

Which might be simplified even more assuming zero contribution of the applied external field  $H$ . Such simplification is justified since many questions described by the considered model might be solved in the absence of an external field. Hereby the following equation denotes the energy of the systems in state  $a$ .

$$\mathcal{H}(a) = -J \sum_{i,j} S_i S_j \quad (2.15)$$

Given simplified version of Hamiltonian allows calculating several properties of magnetic materials such as specific heat or magnetization at a given temperature.

The Metropolis-Hastings algorithm is the most frequently used algorithm for the Ising model simulations. As an initial stage algorithm chooses the probabilities  $p(a, b)$  which denotes the possibility of the algorithm to choose state  $b$  when the system is in a state  $a$ . It then uses acceptance probabilities  $A(a, b)$  so that demand of detailed balance is satisfied. If the new state  $b$  is accepted, then we move to that state and repeat with selecting a new state and deciding to accept it. If  $b$  is not accepted then we stay in  $a$ . This process is repeated until some stopping criterion is met, which for the Ising model is often when the lattice becomes ferromagnetic.

Given algoritm follows the concept of single-spin-flip dynamics, which states that in each transition, we will only change one of the spin sites on the lattice. Furthermore, by using single-spin-flip dynamics, one can get from any state to any other state by flipping each site that differs between the two states one at a time.

The derivation of the Metropolis algorithm follows from a few simple steps. From state  $a$ , there are  $N$  possible states that can be reached after one flip to create different states. The probability to create a specific state  $b$  from  $a$  is thus  $g(a, b) = \frac{1}{N}$ , as they are all equally favored; the probability of creating state  $a$  from  $b$  is also the same. The condition of detailed balance, which follows from, can then be stated as:

$$\frac{P(a, b)}{P(b, a)} = \frac{A(a, b)g(a, b)}{P(b, a)g(b, a)} = \frac{A(a, b)}{A(b, a)} = \frac{\frac{1}{Z}e^{\beta E_b}}{\frac{1}{Z}e^{\beta E_a}} = e^{-\beta(E_b - E_a)} \quad (2.16)$$

The choice of acceptance ratio can be made in almost any fashion, so long as this equation is obeyed; however, a low acceptance ratio would lead to many wasted calculations, and as such a large acceptance ratio is therefore generally more efficient. In general the acceptance probabilities chosed to satisfy:

$$\frac{A(a, b)}{A(b, a)} = e^{-\beta(E_b - E_a)} \quad (2.17)$$

if  $E_b > E_a$ , then  $A(b, a) > A(a, b)$ . Metropolis sets the larger of  $A(a, b)$  or  $A(b, a)$  to be 1. By this reasoning the acceptance algorithm is:

$$A(a, b) = \begin{cases} e^{-\beta(E_b - E_a)}, & \text{if } E_b - E_a > 0 \\ 1, & \text{else} \end{cases} \quad (2.18)$$

As a summary to the usage of the Metropolis algorithm in Monte Carlo simulations, the following steps are done for one application of the algorithm:

1. Calculate the energy of the system.
2. Pick a random spin on the lattice.
3. Calculate the energy if the spin is flipped and  $\Delta E = e^{-\beta(E_b - E_a)}$ , the difference in energy between this energy and the previous one.
4. If  $\Delta E < 0$ , the spin is flipped. Otherwise, the next step is followed.
5. Calculate the Boltzmann weight  $w = e^{-\beta(E_b - E_a)}$ .
6. Generate a random number  $0 \leq r < 1$ .
7. If  $r < w$ , flip the spin. If not, no change is made.

## Chapter 3

# Materials and Methods

### 3.1 Data driven estimation of critical temperature

This part of the work was mainly inspired by the publications related to the ML-driven estimation of the critical temperature for the superconducting [9] and ferromagnetic materials [10]. In the first work [9] were described several ML models for the estimation of the superconducting critical temperature over the 18000 samples dataset. The ML pipeline presented in this work consists of two models working sequentially. The first model classifies superconductors into three groups (Fe-based, cuprates, and low  $T_C$  superconductors). While the second one solves the regression problem for the particular class, in other words predicting the  $T_C$ . As it was reported classification models that worked solely on chemical composition information showed strong predictive power with an accuracy score of 92%. Regression Random Forest model worked already with compositional and structural descriptors. It showed  $R^2$  score of 0.85, 0.80 and 0.74 for these classes respectively

The second work [10] reported the construction of several ML models in a similar way to predict the critical temperature of FM with descriptors based only on chemical composition. The construction of proposed models did not involve any electronic structure calculations and has been entirely trained over available 2500 records of experimental data. The Random Forest algorithm gave the best-reported result with a mean absolute error (MAE) of 57 K and  $R^2$  value of 0.81 over the test set containing 767 compounds. Also, the authors report that they made several attempts to include structural information into the descriptors. Still, such an approach did not outperform the model's predictive power containing data obtained from chemical composition only.

#### 3.1.1 Data and predictors

The success of any ML method ultimately depends on access to reliable and plentiful data. Curie temperature data used in this work was extracted from the database provided by the authors of the described previously publication [10]. It houses 5092 experimental values of the critical temperature for ferromagnetic materials from various sources [11, 12].

### 3.1.2 Features generation

The problem which we are working on is an estimation of Curie temperature ( $T_C$ ) from the sample chemical composition. It might be classified as supervised ML for the regression. The goal of supervised ML is to find a function  $f(X)$  that maps the feature vector  $X$ , to the target variable  $y(T_C)$  and allows to interpolate a set of available data to the target property in the most accurate way. In our case feature vector ( $X$ ) contains only the information which might be extracted from the chemical composition of a given sample. Typically, ML employs an intermediate step between compiling raw data (chemical composition) and applying a machine learning algorithm. This step converts data from a raw format into a numerical representation which is more convenient for ML algorithms and visualization. There are no rigorous rules on how it should be done, but the performance of any ML models ultimately depends on the ability to choose the most relevant descriptors. In this work features were constructed according to the following criteria:

- Generation of features should not affect the size of the dataset (one should avoid the cases when descriptor could not be constructed for all the samples in dataset and missing values appears)
- They should be easy to compute, ideally some already known properties of compounds
- Features that allow distinguishing different materials with different chemical composition.

To accomplish this criterion were tested several existing algorithms from references [13–16], and finally made a decision to use ones described in ref. [17, 18] employing the Materials Agnostic Platform for Informatics and Exploration (Magpie). All produced descriptors might be divided into three groups:

1. Attributes which might be determined from the periodic law: Number ( $N$ ), Period ( $P$ ), Group ( $G$ ), Atomic weight ( $AW$ ) and at some sense Mendeleev Number ( $MN$ ).
2. Properties of a particular elements in a composition: Magnetic moment ( $\mu$ ), Bandgap ( $KS$ ), Melting Temperature ( $T_M$ ), Electronegativity ( $\varepsilon$ ), Covalent radius ( $R$ ), lattice volume ( $V$ )
3. Attributes of electronic structure, namely the fraction of electrons from the  $s$ ,  $p$ ,  $d$ , and  $f$  shells

For all categories of descriptors listed above was applied statistical functions defined below: For all categories of descriptors listed above was applied statistical functions defined below:

Table 3.1: Property statistics functions

N	Function	Symbol
1	Minimum	$x_{min}$
2	Maximum	$x_{max}$
3	Range	$\Delta x$
4	Mode	$\tilde{x}$
5	Mean	$\bar{x}$
6	Mean absolute deviation	$\langle x \rangle$

For instance, property Atomic Weight ( $AW$ ) has features: Minimum Atomic Weight ( $AW_{min}$ ), maximum Atomic Weight ( $AW_{max}$ ), range Atomic Weight ( $\Delta AW$ ), mode Atomic Weight ( $\widetilde{AW}$ ), mean Atomic Weight ( $\overline{AW}$ ), mean absolute deviation Atomic Weight  $\langle AW \rangle$ .

High dimensional descriptor  $v_{chem} = \{x_H, x_{He}, x_{Li}, x_{Be}\dots\}$  describes the atomic fraction of every element in a particular sample. For example,  $Mn_2O_3$  might be represented by the vector  $\{\dots, 0, 2/5, 0, \dots, 0, 3/5, 0, \dots\}$  with 3/5 assigned to the element in 8th position (Oxygen) and 2/5 to 25th (Manganese) other positions are equal to zero. The final dimensionality of generated descriptors is 215. All the constructed descriptors presented in Table 2.

Table 3.2: List of descriptors used in this work

N	Features	Symbol	Dim.
1	Atomic fraction vector	$v_{chem}$	90
2	Number	$AW_{min}, AW_{max}, \Delta AW, \widetilde{AW}, \overline{AW}, \langle AW \rangle$	6
3	Period	$P_{min}, P_{max}, \Delta P, \widetilde{P}, \overline{P}, \langle P \rangle$	6
4	Group	$G_{min}, G_{max}, \Delta G, \widetilde{G}, \overline{G}, \langle G \rangle$	6
5	Mendeleev Number	$MN_{min}, MN_{max}, \Delta MN, \widetilde{MN}, \overline{MN}, \langle MN \rangle$	6
6	Atomic weight	$AW_{min}, AW_{max}, \Delta AW, \widetilde{AW}, \overline{AW}, \langle AW \rangle$	6
7	Lattice Volume	$V_{min}, V_{max}, \Delta V, \widetilde{V}, \overline{V}, \langle V \rangle$	6
8	Magnetic moment	$\mu_{min}, \mu_{max}, \Delta \mu, \widetilde{\mu}, \overline{\mu}, \langle \mu \rangle$	6
9	Bandgap	$KS_{min}, KS_{max}, \Delta KS, \widetilde{KS}, \overline{KS}, \langle KS \rangle$	6
10	Melting Temperature	$T_{min}^M, T_{max}^M, \Delta T^M, \widetilde{T^M}, \overline{T^M}, \langle T^M \rangle$	6
11	Electronegativity	$\varepsilon_{min}, \varepsilon_{max}, \Delta \varepsilon, \widetilde{\varepsilon}, \overline{\varepsilon}, \langle \varepsilon \rangle$	6
12	Covalent radius	$R_{min}, R_{max}, \Delta R, \widetilde{R}, \overline{R}, \langle R \rangle$	6
13	s-shell electrons (filled/unfilled)	$N_{min}^s, N_{max}^s, \Delta N^s, \widetilde{N^s}, \overline{N^s}, \langle N^s \rangle$	12
14	p-shell electrons (filled/unfilled)	$N_{min}^p, N_{max}^p, \Delta N^p, \widetilde{N^p}, \overline{N^p}, \langle N^p \rangle$	12
15	d-shell electrons (filled/unfilled)	$N_{min}^d, N_{max}^d, \Delta N^d, \widetilde{N^d}, \overline{N^d}, \langle N^d \rangle$	12
16	f-shell electrons (filled/unfilled)	$N_{min}^f, N_{max}^f, \Delta N^f, \widetilde{N^f}, \overline{N^f}, \langle N^f \rangle$	12

### 3.1.3 Data pre-processing

Initially, the dataset contained 5092 samples. However, this data needed pre-processing since it included many compositions for which in literature was reported different values of  $T_C$ . There are several reasons for these occurrences. In some compounds magnetism is related to the quality of the sample, so the experiments performed by different groups and over different periods of time may report sufficiently different values of  $T_C$ . A second reason related to the polymorphism, namely to the existence of compounds with the same stoichiometry but different crystal structure and hence different values of  $T_C$  (e.g Fe-bcc and Fe-fcc). Since at this stage the aim of this project was to construct the ML model based only on the chemical composition, structural information or details of the experiment were not included in the descriptors. In other words, feature vectors ( $X$ ) of two samples with the same stoichiometry are indistinguishable even thou the target property ( $y$ ) differs and it is not possible to extract any information about polymorphism or sample quality. As such it was made a decision to establish criteria to exact a single value of  $T_C$  for the samples with the same stoichiometry. In this work was used the median of the distribution instead of mean or maximum value since it is less dependent on the possible outliers. For example, if for the same stoichiometry was reported 3 different values of  $T_C$  305, 330, and 600K they were replaced with a single median value of  $T_C = 330K$ . Of course, such an approach has drawbacks. In cases when we have an even number of samples with the same stoichiometries (e.g. 4 or 6)  $T_C$  also depends on a half sum of the distribution central values. But after all, it allows us to operate with more reliable values since most of them are associated with real experimental measurements but not with a mean value of several experiments. After this operation, the size of the dataset decreased almost by half and became 2557 samples, which might be considered as relatively small from the point of view of ML algorithms. Distribution of  $T_C$  temperatures in the pre-processed data shown on Figure 3.1.

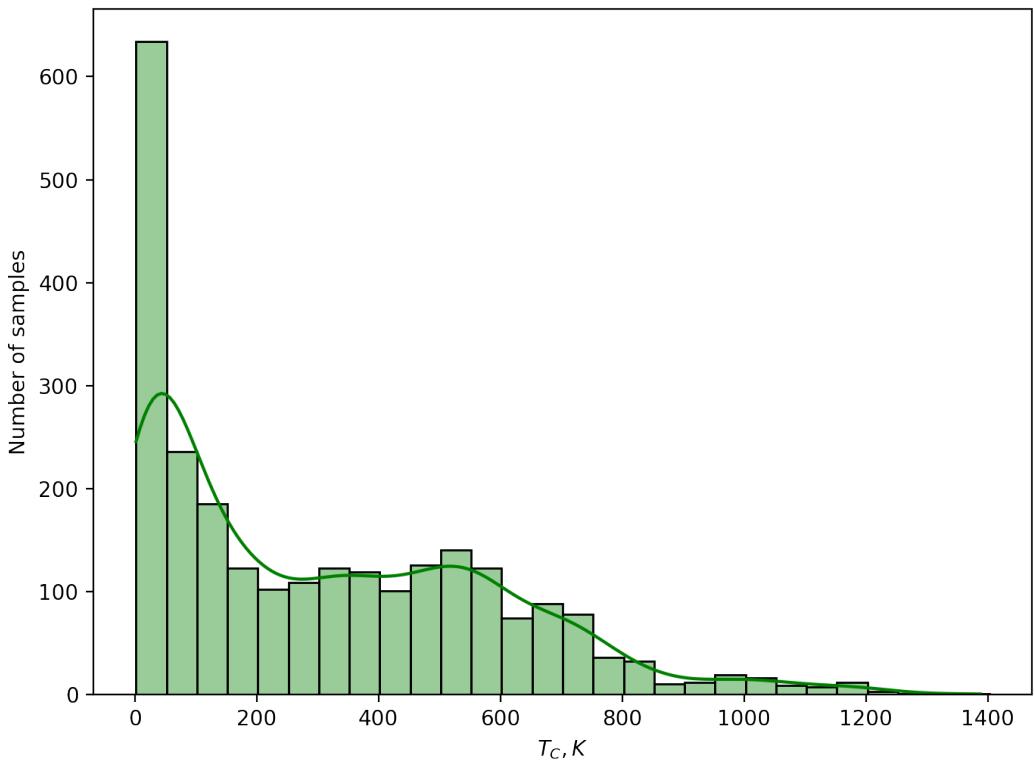


Figure 3.1: Distribution of critical temperature over 2557 samples.

### 3.1.4 Machine-learning models

#### Models evaluation

As it was already mentioned the size of our data is relatively small. This fact means that we were not able to follow the standard procedure of the model training and evaluation which requires splitting of the data into three equal and mutually exclusive partitions namely training, testing, and validation sets. In case when the dataset is not so big such split makes each of the subsets too small decreasing the overall ability to learn and thus accuracy. That's why the entire dataset was divided only into 2 not equal parts: 1st is a training set which included 2/3 of the data (1714 samples) and 2nd is testing set consisting of the remaining 834 samples. Schematically this process shown in Figure 3.2.

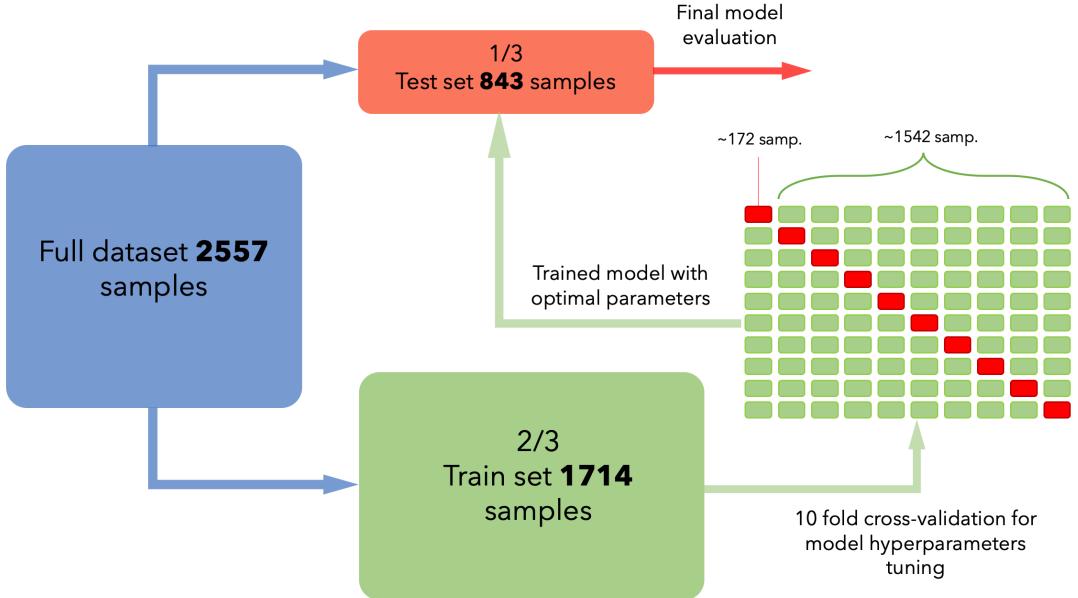


Figure 3.2: Scheme of cross-validation process followed by final model evaluation on the test subset.

Hyperparameters optimization for all models was performed through the standard K-fold cross-validation procedure. The training set is split into K subsets (in this work K=10) then the model is trained over the K-1 sets and evaluated on the remaining one, thus the cross-validation error is an average error of the K evaluation rounds. The best model (model with optimum hyperparameters) is a model with the lowest cross-validation error. Finally, the test set was used for estimation the generalization error of the model, namely for assessing how well the given model performs on never-seen-before data.

Quantitatively model's performance were measured by the coefficient of determination  $R^2$ :

$$R^2 = 1 - \frac{\sum_i (y_i - f(x_i))^2}{\sum_i (y_i - \mu)^2} \quad (3.1)$$

and root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{\sum_i (y_i - f(x_i))^2}{n}} \quad (3.2)$$

Here:  $y_i$  - target property;  $f(x_i)$  - predicted value;  $\mu$  - mean

For the perfect predictor, namely  $y_i = f(x_i)$ , these values should be equal to 1 and 0 respectively.  $RMSE$  and  $R^2$  were specified as the main optimization parameters for all trained models. Also, for the all models were estimated and reported maximum error but this parameter was not chosen as an optimization one. In the construction of ML models were compared four

different algorithms:

1. Kernel Ridge Regression (KRR)
2. Random Forest (RF)
3. Extreme Gradient Boosting (XGBoost)
4. Sure Independence Screening plus Sparsifying Operator (SISSO)

The first three models come from Python distributions Scikit-Learn [19] and XGBoost [20] respectively. SISSO is a separate algorithm on Fortran 90 [21, 22]. KRR, RF, and XGBoost were trained and evaluated following the 10-fold cross-validation and testing procedure described previously. SISSO was trained on entire dataset without accurate validation, due to the lack of computational resources. Hyperparameters tuning for every model will be described explicitly.

### **Kernel ridge regression.**

The first model which was chosen as a baseline for particular work is a Kernel Ridge Regression (KRR). Simple Ridge Regression was not considered since the sufficiently large dimensionality of the descriptors might significantly affect the performance of this model. Kernel ridge regression combines ridge regression (linear least squares with  $l_2$ -norm regularization) with the kernel trick. Which in compare to Ridge Regression might be written as follows: Ridge regression:

$$y = w^T \phi(x) \quad (3.3)$$

The idea is to minimize loss function (estimate the best weights):

$$J(w) = \sum_{i=1}^n (y_i - w^T x_i)^2 \quad (3.4)$$

Not kernilized solution of this equation for the nonlinear regression might be written as follows:

$$w = (\phi^T \phi + \lambda I)^{-1} \phi^T y \quad (3.5)$$

Kernel Trick allows as to sufficiently decrease the complexity of this solution, since the dot product  $\phi \phi^T$  can be found just as  $K(x_m, x_n)$  function:

$$w = (\phi^T (\phi \phi^T + \lambda I))^{-1} y \quad (3.6)$$

$$K(x_m, x_n) = \phi \phi^T \quad (3.7)$$

It thus learns a linear function in the space induced by the respective kernel and the data. For non-linear kernels, this corresponds to a non-linear function in the original space. In this work KRR algorithm was tuned with respect to the regularization parameter  $\lambda$  and different Kernels, namely:

$$K(x_m, x_n) = x_m^T x_n \quad (3.8)$$

$$K(x_m, x_n) = (x_m^T x_n + r)^d \quad (3.9)$$

$$K(x_m, x_n) = \exp\left(-\frac{\|x_m - x_n\|^2}{2\sigma}\right) \quad (3.10)$$

$$K(x_m, x_n) = \exp(-\alpha \|x_m - x_n\|^2) \quad (3.11)$$

It turns out that for all non-linear Kernels optimal value of optimization parameter  $\lambda$  is equal to 0.10 (tested in range 0.05 - 1). Best  $R^2$  and  $RMSE$  scores were achieved by the Laplacian Kernel which is also the most computationally demanding one (required almost 5 times more computational resources than others). The lowest value of maximum error was given by the simplest linear Kernel. Performance of different KRR models shown on scatter plots in Figure 3.3, and also in table 3.3.

Table 3.3: KRR models performance with respect to different Kernels

Metrics/Kernel	Linear	Polynomial	RBF	Laplacian
R2 (test)	0.616	0.759	0.756	0.793
R2 (cross.val)	0.601	0.768	0.754	0.787
RMSE (K)	161.9	131.4	132.3	121.7
Max Error (K)	598.1	783.4	795.6	642.05

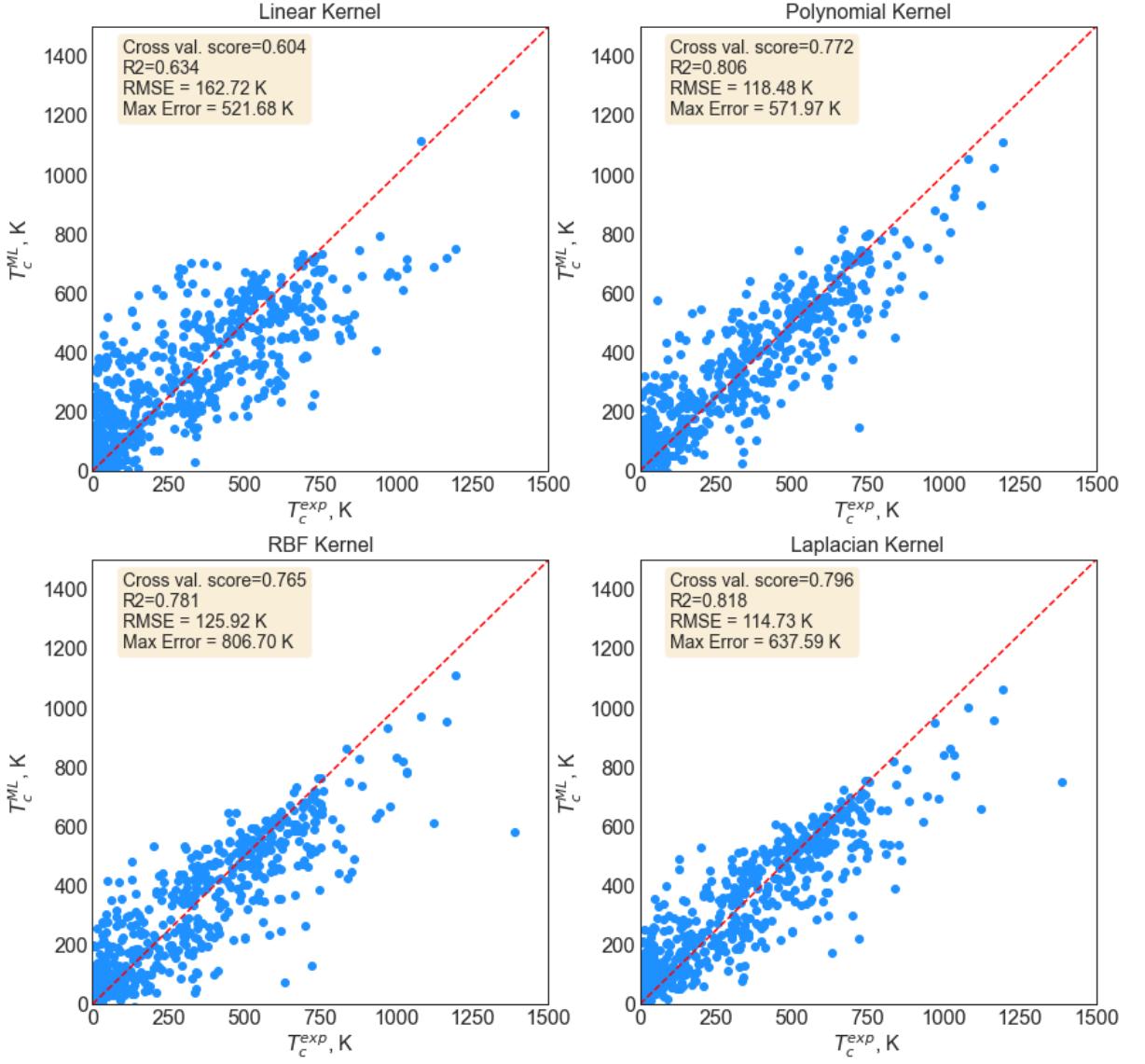


Figure 3.3: Comparison of experimental ( $T_c^{exp}$ ) and predicted values ( $T_c^{ML}$ ) for different Kernels.

### Random forest algorithm

Random forest (RF) is one of the most powerful, versatile, and widely used ML methods. RF, like its name implies, consists of a large number of individual decision trees (see Figure 3.4) that operate as an ensemble. Each individual tree in the RF trains only on a random part of the training set samples and also only on the randomly selected subspace of descriptors (i.e. every tree in RF learns on different samples and features). This approach allows to build a large number of relatively uncorrelated models (trees) which operating as a committee will outperform any of the individual constituent models.

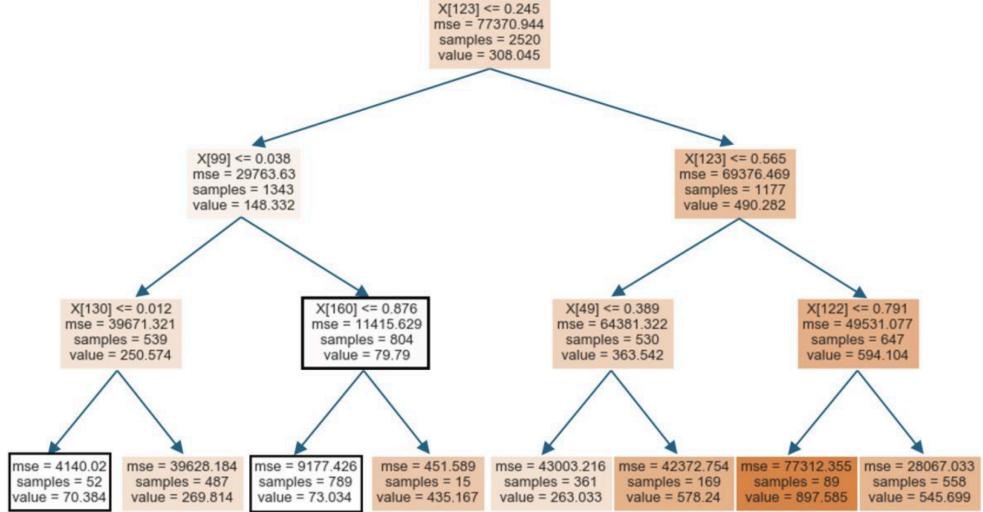


Figure 3.4: Single Tree model with depth 4 for the regression problem under consideration.

There are several important advantages that make it especially suitable for the problem under consideration. First, it can find complicated non-linear dependencies from the data. Compare to the number of other methods (e.g. linear regression), it does not make any assumptions or fitting about the functional form of the relationship between the predictors and the target variable. The second important advantage of this method is that, by combining information from individual independent predictors (trees), it can determine the importance of each particular feature, thus making the whole model more interpretable, and allows us to extract useful insights from the trained model. For example, predictor importance was employed for the decreasing of the feature space dimensionality. Predictors were sorted according to their importance and firstly were removed 25% of less important features (Q75) and then another 25% (Q50). New models were trained and evaluated on the reduced features spaces (Q50, Q75) as it is shown in Table 3.4. Hyperparameters that were optimized are maximum tree depth as shown in Figure 3.5 (a) and the number of trees Figure 3.5(b). Convergence was achieved after 100 trees in an ensemble with a depth equal to 10.

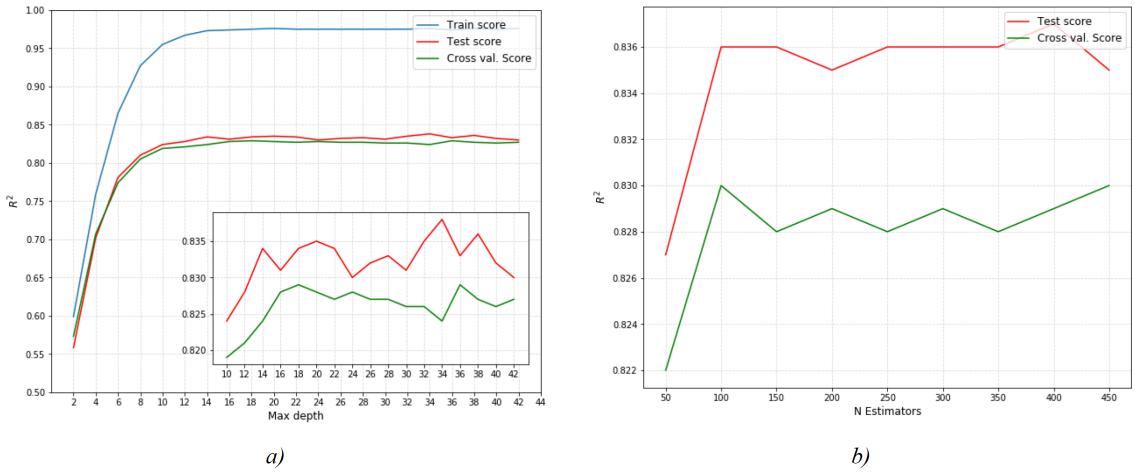


Figure 3.5: RF model hyper parameters tuning a) Convergence with respect to the tree depth b) Convergence with respect to the number of trees.

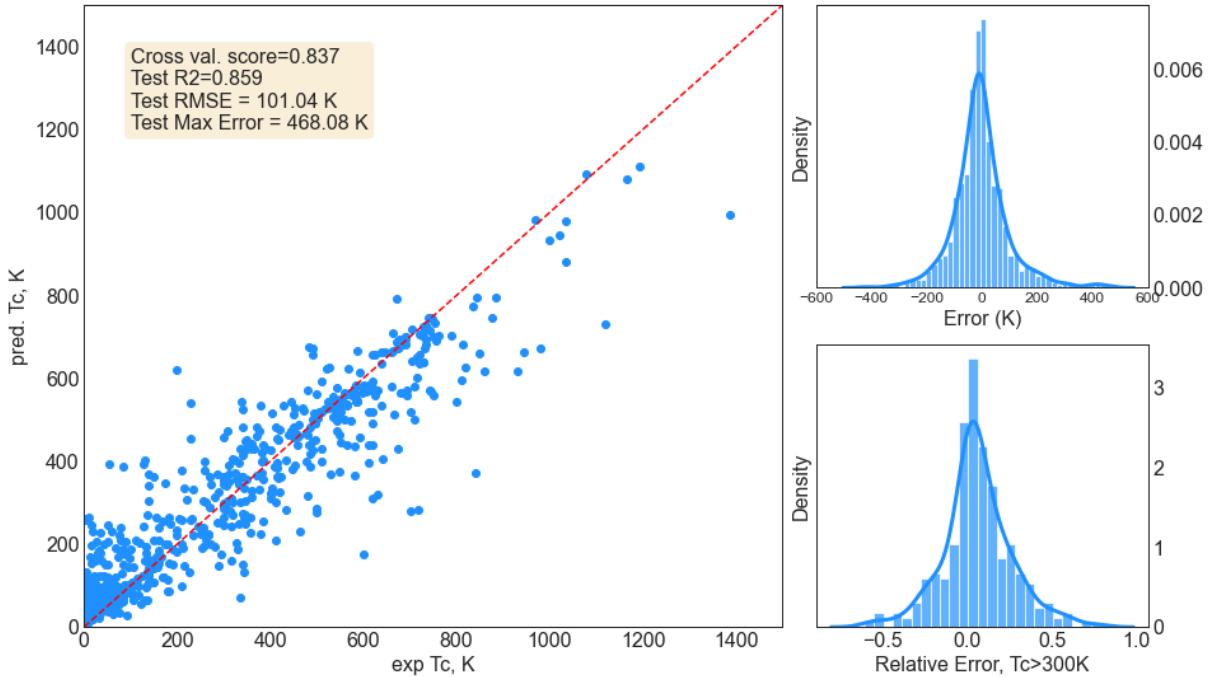


Figure 3.6: Optimized RF model performance, left comparison of predicted and experimental data, right errors and relative errors for the compounds with  $T_c > 300K$ .

Table 3.4: RF models performance with respect to the features space

Metrics	All	Q75	Q50
$R^2$ (test)	0.833	0.819	0.810
$R^2$ (cross.val)	0.828	0.816	0.807
RMSE (K)	109.4	113.6	131.3
Max Error (K)	516.3	540.5	601.3

As it might be seen from the table 3.4 in general RF outperform the best KRR model and shows good predictive power. Error distribution (Figure 3.6 right top) shows the relation between experimental and predicted values of critical temperature ( $T_C^{exp} - T_C^{pred}$ ). In our case It is highly symmetrical around zero and might be considered as Gaussian type. From this fact we can conclude that RF model has no systematic bias towards overestimation or underestimation. The distribution of relative errors  $(T_C^{exp} - T_C^{pred})/T_C^{exp}$  showed only for the compounds with  $T_C > 300K$  since for the compounds with low TC relatively small absolute error may lead to a huge relative one (i.e. for  $T_C^{exp} = 1K ; T_C^{pred} = 15K$ , 1400% error). Relative error distribution already not so symmetrical but it might be explained due to the low number (136) of samples in it.

## XGBoost algoritm

XGBoost is another ensemble algorithm based on decision trees, which utilizes a gradient boosting (GB) framework. Initially, it was designed for small-to-medium structured or tabular datasets, which is exactly the case in this work. Gradient boosting (GB) is an ML technique very similar to RF. It produces the prediction model as an ensemble of weak prediction models. The algorithm starts with a single leaf which represents an initial guess for the target variable ( $y$ ) for the samples, if it is not specified explicitly initial guess is always the average value of ( $y$ ). The general difference from the simple RF algorithm that GB builds a new tree, based on the errors made by the previous regressor and combines it with the results of the original leaf to make a new prediction. Trees depth in this algorithm controlled by the number of leaves (for regression problems and relatively small datasets  $<5000$  the empirical value of leaves between 8 and 20). To scale the contribution of each predictor GB use a special parameter called learning rate which is the value between 0 and 1. This parameter is a coefficient that shows the contribution (in RF we couldn't affect it) of each individual tree to the overall result and also controls the total amount of estimators. Empirical evidence shows that taking lots of small steps in the right direction (use small learning rate) results in a better prediction with testing dataset i.e. lower the variance but also require additional resources. In this work learning rate was set to 0.1 to keep the balance between accuracy and a reasonable amount of computational time. XGBoost parameter tuning was done according to the general recommended approach:

1. Firstly, at fixed learning rate was estimated the optimal number of trees as it is shown in Figure 3.7 (a).
2. Secondly was tuned tree-specific parameters in our case only trees depth see Figure 3.7 (b)
3. Were determined optimal regularization parameters ( $\gamma, \alpha$ ). These values were set to 1 and 0.5 respectively.

#### 4. Optimization of the learning rate

Therefore, final hyperparameters were as follows: Max depth=8, Number of trees=250, learning rate=0.1,  $\gamma = 1$ ,  $\alpha = 0.5$ .

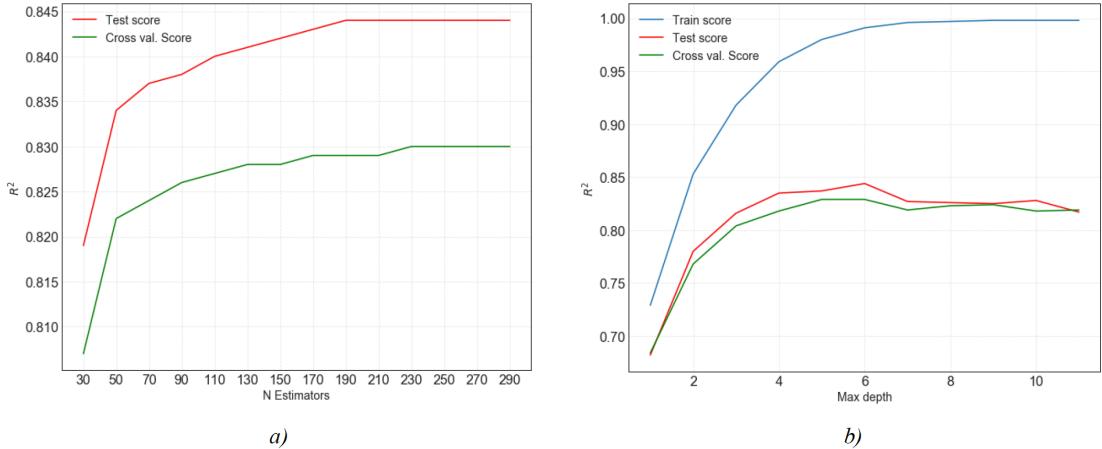


Figure 3.7: XGBoost model hyperparameters tuning: a) Convergence with respect to the number of trees b) Convergence with respect to the tree depth.

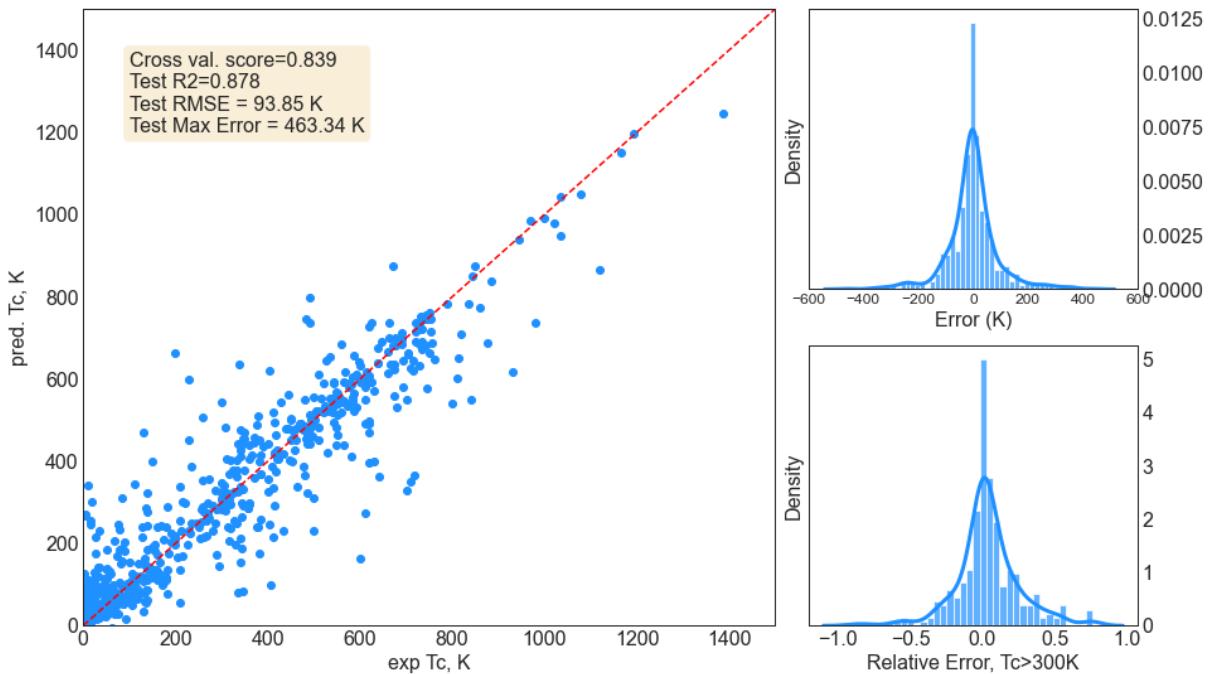


Figure 3.8: Optimized XGBoost performance, left comparison of predicted and experimental data, right errors and relative errors for the compounds with  $T_C > 300\text{K}$ .

Table 3.5: XGBoost performance with respect to the features space

Metrics	All	Q75	Q50
$R^2$ (test)	0.845	0.833	0.823
$R^2$ (cross.val)	0.837	0.825	0.817
RMSE (K)	105.6	111.8	123.6
Max Error (K)	595.6	623.5	677.3

From the Table 3.5 we can conclude that XGBoost in general shows higher accuracy than the RF and KRR. Errors distribution (Figure 3.8 top right) is a Gaussian type and highly symmetrical around zero, so this model also has no biases towards the overestimation or underestimation of the  $T_C$ .

Surprisingly maximum error of this model is sufficiently larger than that one estimated from RF. Finally, we can conclude that XGBoost is the most accurate model out of all considered in this work with respect to the  $R^2$  and  $RMSE$  optimization parameters. The mean absolute error ( $MAE$ ) for this model is 69K.

### Sure independence screening and sparsifying operator (SISSO)

The sure independence screening and sparsifying operator (SISSO) is a systematic approach for discovering descriptors for materials' properties, within the framework of compressed-sensing-based dimensionality reduction. It gives an ability to handle huge and correlated features spaces and converges to the optimal solution building a combination of features relevant to the materials' target property. There are several works [21–23] which report the ability of this algorithm to work efficiently both with regression and classification problems and also to achieve a stable results with relatively small training sets. The output of SISSO is a mathematical model, in the form of explicit, analytic functions of the input physical quantities. This aspect makes this model highly interpretable and gives an opportunity to inspect the equations and obtain some useful physical insights about the problem under consideration. An input 20-dimensional features space for SISSO were build-out of the 60 most important descriptors which were determined from RF. Were tested different combinations of these features in order to minimize the overall correlation between them. Finally, was constructed a 20-dimensional feature space as it is shown in the correlation matrix (Figure 3.9).

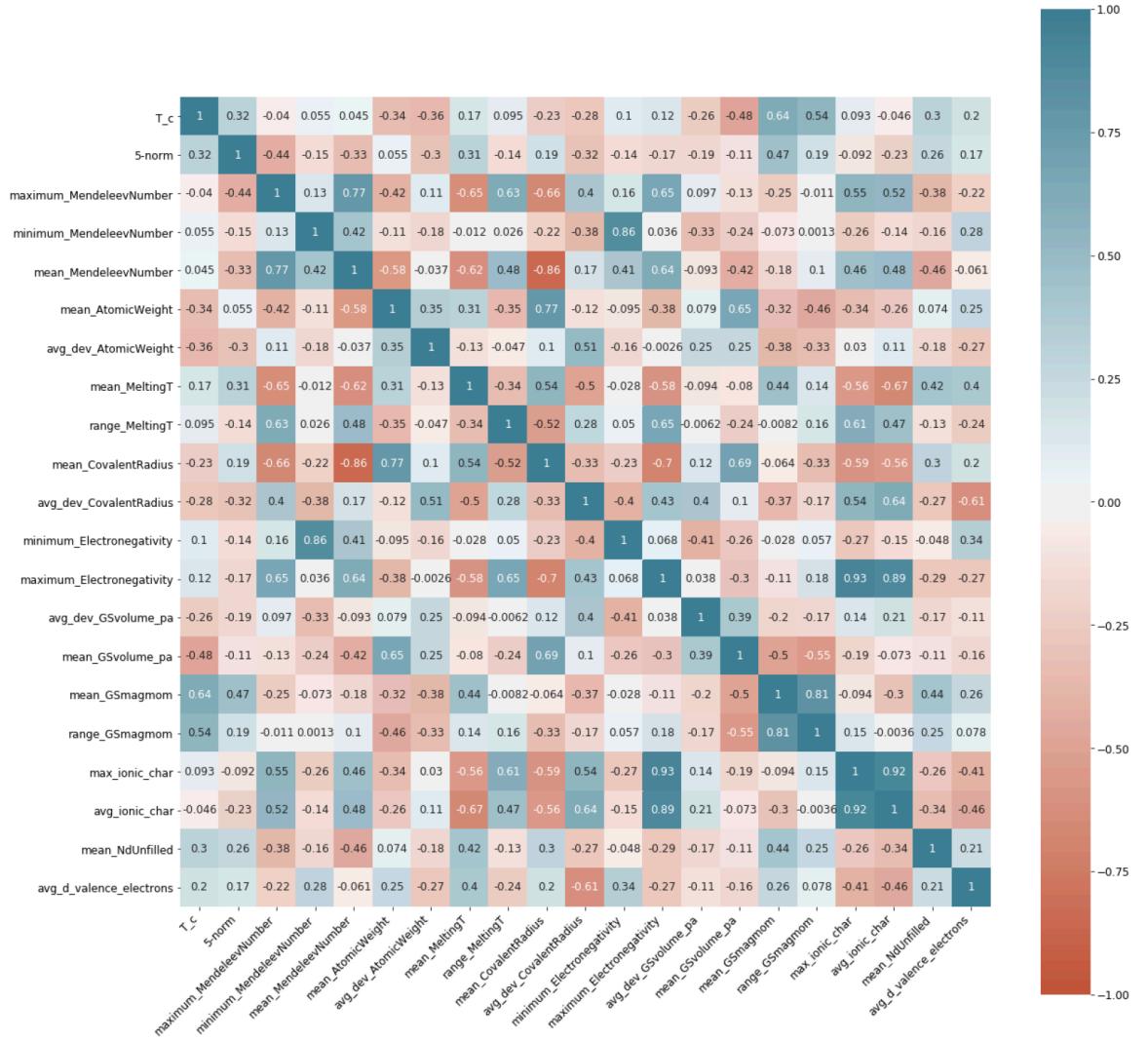


Figure 3.9: Correlation matrix for the most valuable descriptors in RF.  
These descriptors were used as input for the SISSO algorithm.

After that SISSO was trained on the entire dataset. Initially, for a baseline model, feature space rung parameter were set to 2, and calculations were performed up to the 7-dimensional descriptor (see Figure 3.10a). As it was expected RMSE and maximum error has a tendency to decrease with the increased size of produced descriptors (see Figure 3.10c). Then for the second run rung on feature space were increased by one and calculations were also started up to the 7-dimensional descriptor. Unfortunately, only a 3-dimensional (see Figure 3.10b) descriptor was calculated.

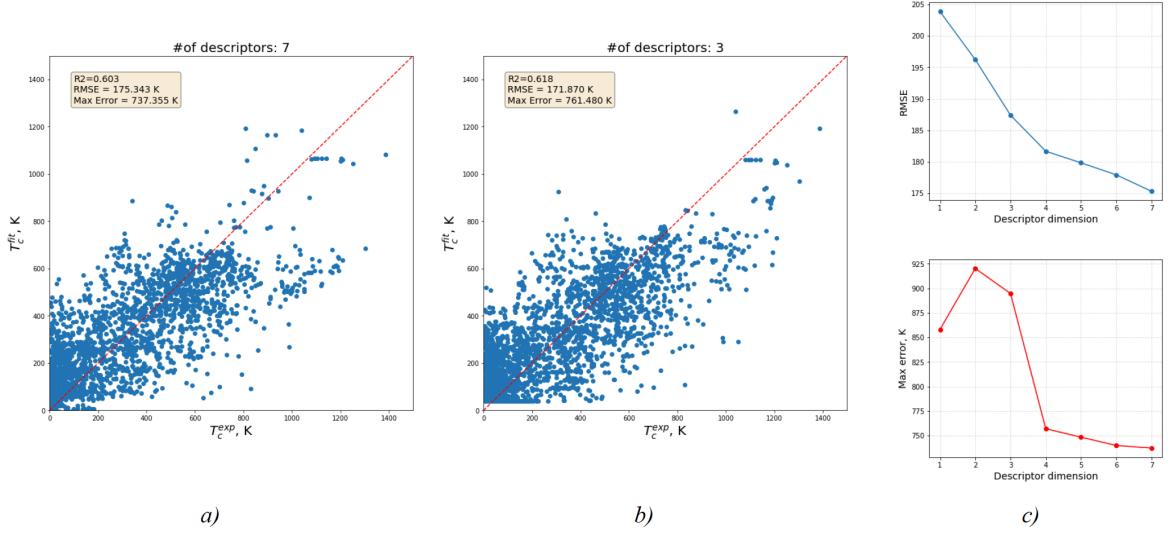


Figure 3.10: SISSO model performance: a) Rung of feature space=3, subspace size=100, b) Rung of feature space=2, subspace size=10, c) RMSE and Max error metrics with respect to the descriptors dimensionality for the feature space with rung 2.

As it is shown  $R^2$  value and  $RMSE$  metrics for SISSO worse than for the other algorithms. But probably the potential of SISSO is still not fully disclosed and it is possible to achieve a way better result. The only problem is that this algorithm is rather computationally demanding and requires a lot more resources than its available now for the author of this work. Also, it is interesting to check which features were produced. Here is shown 3 features which were generated during the second run (rung of features space equal to 3):

$$X_1 = \frac{\bar{\mu} \cdot \varepsilon_{max} \cdot MN_{max}}{\bar{V}^2 \cdot \bar{T}_M \cdot \Delta T_M} \quad (3.12)$$

$$X_2 = \frac{\bar{N}^d \cdot MN_{min}}{\exp(\Delta\mu)} - \frac{\bar{N}^d \cdot \bar{MN}}{\exp(\varepsilon_{min})} \quad (3.13)$$

$$X_3 = \frac{\left( N^V - \bar{N}^d \right)^2 \bar{\mu}^2}{\varepsilon_{min}^3} \quad (3.14)$$

As it might be seen the most frequently used primary features is a magnetic moment ( $\bar{\mu}, \Delta\mu$ ) and electronegativity ( $\varepsilon_{min}, \varepsilon_{max}$ ), these features were included in all tree generated descriptors. Features related to the electronic structure namely number of valence electrons  $N^V$  and number of electrons on d-shell ( $\bar{N}^d$ ) as well as Mendeleev number ( $MN_{min}, MN_{max}, \bar{MN}$ ), were included in 2 out of three descriptors. Properties like melting temperature ( $\bar{T}_M, \Delta T_M$ ) and lattice volume ( $\bar{V}$ ) were used only for one descriptor.

### 3.1.5 Physical understanding of the most important descriptors

Building an efficient model that could make an accurate prediction is an important part of the ML pipeline. But what is even more important is to interpret your model properly and find what kind of physical insights and understanding of a particular problem might be extracted from it. That's why in some cases simpler models like decision trees or Random Forest are more preferable due to their high interpretability and transparency. Figure 3.11 shows the histogram of the top 15 most valuable features which were extracted from our trained RF model. Actually, this feature is not strongly different between 3 of our models (KRR, RF and XGBoost), that's why it's enough to consider only one such histogram.

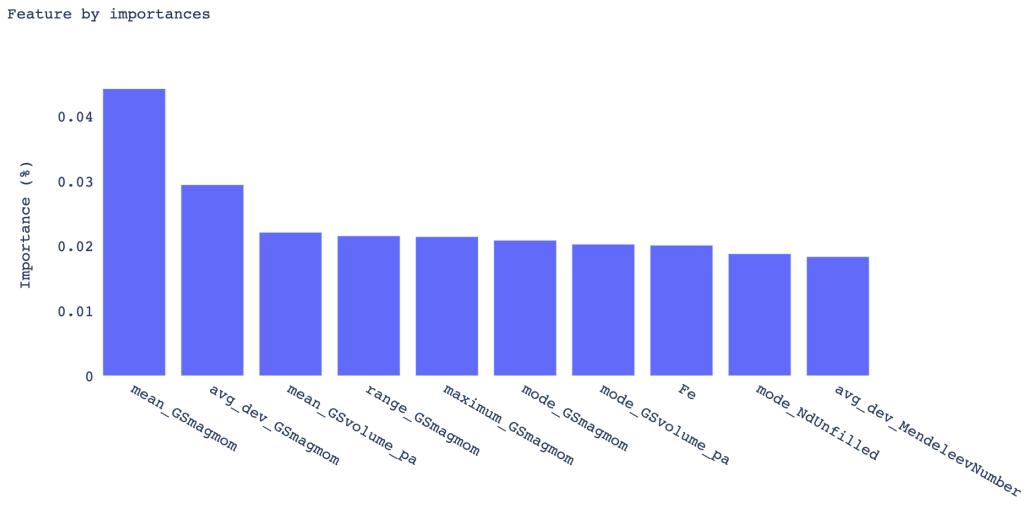


Figure 3.11: RF feature importance histogram.

Five descriptors out of fifteen related to the magnetic moment, namely:  $\mu_{max}$ ,  $\Delta\mu$ ,  $\bar{\mu}$ ,  $\tilde{\mu}$ ,  $\langle \mu \rangle$ . This fact is quite understandable since by definition target property ( $T_C$ ) is a point of sharp changes in the magnetic properties. There are 2 descriptors corresponding to the atomic fraction of particular elements, namely Iron (Fe) and Cobalt (Co) it is also might be explained due to their strong ferromagnetic nature, other important elements like Nickel (Ni) and Manganese (Mn) are 25th and 40th respectively. There 2 features related to the number of electrons on d-shell  $\overline{N^d}$ ,  $\widetilde{N^d}$  and again it's a well-known fact that partially filled d-shell usually correlated with nonzero total spin and thus magnetism. Also, there 2 descriptors related to lattice volume  $\overline{V}$ ,  $\widetilde{V}$ , this already not so trivial, but probably related to the phenomena of the exchange coupling which is strongly dependent on the number of nearest neighbours in the crystal structure. Comparing the most valuable features extracted from KRR, RF, and XGBoost with the features computed from SISSO we may say that they have a lot in common despite the fact that all these algorithms use different mathematical approaches and approximations.

## 3.2 DFT + MC calculations of critical temperature

One of the foundation stones of this particular work is developing a stable and efficient method for estimating ferromagnetics critical temperature (Curie temperature). In a framework of described in section 2.6 spin model, it is straightforward that the main parameter which should be found to conduct this calculation is exchanged interaction  $J_{ij}$ . Finding this parameter allows to carry out Monte Carlo simulation as it is described in section 2.7 for the precise description of the system's thermodynamic properties or apply less computational demanding but yet quite accurate Curie-Weiss approximation. Steps required for these purposes are schematically illustrated in figure 3.17 and described in more detail in upcoming sections.

### 3.2.1 Obtain initial geometry

On a stage of development and validation of algorithm, this work used FM structures with known magnetic properties, and critical temperature preferably obtained both from experimental and theoretical studies. In particular was used structural information provided by the Materials Project and NOVOMAG databases, which have already proven as a reliable tool for computational materials exploring.

### 3.2.2 Generation of spin configurations

From obtained structural information of a parent FM cell, one may generate various derivative structures with specified AFM properties (i.e., ratio of spin-up and spin-down atoms is 1:1) utilizing existing algorithms for derivative structure generation. For these purposes, currently used open-source python library, *pymatgen* [24] with integrated derivative structure generator *enumlib* [25–28]. Based on symmetry analysis its generates different AFM spin configurations from the parental FM. One of the big advantages of the current derivative structure enumerator is the possibility of eliminating the heuristics-based approach. On the contrary of producing a huge number of randomly oriented spin configurations with the hope of getting the one that is close enough to FM ground state, it makes a rigorous and thus reliable search for all possible spin states. The number of obtained spin configurations we may consider almost exhaustive, which is also proven in a series of studies by comparing with experimental data. [29, 30].

### 3.2.3 Energy calculations

Initial FM cell, as well as various generated AFM structures, are then relaxed using collinear density functional theory, and their energies are accurately calculated.

Spin-polarized DFT calculations are carried out using generalized gradient approximation (GGA) as implemented in the Vienna Ab initio Simulation Package (VASP) [31] with projector augmented-wave (PAW) [32] method applying the Perdew–Burke–Ernzenhof (PBE) [33] exchange-correlation functional. For all calculations was used sufficiently large value of cut-off energy up to 1.5 of the maximum provided in pseudopotential file.

Because the energy differences between various close to ground state spin configurations are usually on several milli-electron volts (meV), it's essential to conduct computations with sufficiently high accuracy but also keep computational demanding on a reasonable limit. To keep the balance between efficient usage of computational resources and accuracy requirements, stages of relaxation and static energy calculations use a step-by-step approach (with increasing accuracy). This way allows to quickly determine unsuitable for the future calculations structures and exclude them saving resources.

For the relaxation stage was employed gamma-centered grid with the default specified in VASP reciprocal density of  $64 \text{ \AA}^{-3}$  while the static calculations utilize significantly higher reciprocal density of  $300 \text{ \AA}^{-3}$ . This type of mesh is known to be a pretty flexible (and probably the most commonly used) flavor of the automatic grid generation mode and is shown to achieve faster energy convergence compared to standard Monkhorst-Pack grids.

The parameter of electronic convergence (energy difference between successive electronic steps) is varying from  $10^{-4} \text{ eV}$  for the initial relaxation stages up to  $10^{-6} \text{ eV}$  for the finalizing static runs.

Worth mentioning that at this step, it is essential to ensure that among all produced spin configurations, the ground state is found to be FM. Otherwise (AFM structure is ground state), it is needed to treat this material differently, and the workflow cannot be continued. Another important step is to check if all the structures kept their local magnetization unchanged (or at least these changes are not significant). Quite often, especially structures with low initial magnetic moments exhibit total disappearance of magnetization in AFM state. Such behavior makes them absolutely nonmagnetic and thus not representative for further Hamiltonian fitting.

For the case of structures with strongly correlated d-shell electrons, it was also tested Hubbard correction approach (GGA + U) using the Dudarev [34] formulation which is proven to give accurate results for such systems.

The effective U values for transition metals calibrated by performing a fitting to experimental binary formation enthalpies have been taken from the Materials Project website [35]. This is a common practice and has already shown reasonable results in a high-throughput screening study performed in ref. [29]. Of course, usage of Huburd correction values without proper calibration, for instance, by linear response method [36], may lead to unsatisfying results, especially for structures containing elements with relatively low magnetization. Unfortunately, the computational demand

of such calibration is sufficiently large and not applicable in high-throughput calculations.

### 3.2.4 Hamiltonian fitting

#### Building the system

The energies calculated at the previous stage for the relaxed parental FM and derivative AFM configurations make it possible to determine the exchange coupling parameters ( $J_{ij}$ ). For these purposes initially carried out mapping of magnetic neighbors in every generated spin configuration utilizing *pymatgen* [24] and *siman* [37] libraries . This stage performed not only mapping of magnetic nearest (mNN) and next-nearest magnetic neighbors (mNNN), located in the first and second coordination spheres respectively, the search is extended to the higher-order coordination spheres up to the number of generated spin configurations. Hence for any FM system with a number of generated AFM spin configurations equal to  $N$ , mapping was done up to the  $N - th$  coordination sphere in every structure (e.g., for 6 generated AFM structures, mapping of magnetic neighbors performed up to the 6-th coordination spheres).

Mapping procedure returns a series of coefficients ( $n_{ij}$ ) which, with the addition of calculated energies ( $E_{FM/AFM}$ ), allows building a system of equations. Produced system includes  $N + 1$  equations (1 FM +  $N$  AFM configurations) and  $N + 1$  variables ( $E_g, J_1, \dots, J_n$ ).

$$\begin{cases} E_g + n_{11}J_1 + \dots + n_{1N}J_n = E_{FM} \\ E_g + n_{21}J_1 + \dots + n_{2N}J_n = E_{AFM1} \\ \vdots \\ E_g + n_{N1}J_1 + \dots + n_{NN}J_n = E_{AFM_N} \end{cases} \quad (3.15)$$

Its might be illustrated on the toy example from figure 3.12. Suppose we have a 2D structure with a spin-up red atoms and spin-down blue atoms. As a 1st nearest neighbors of our spin-up central atom, we have 4 atoms with an antiparallel spin which means that sign of their interaction will be negative. Building the second coordination sphere we can see that all 2nd neighbors have the same spin with the central one, therefore, the sign of their interaction will be positive.

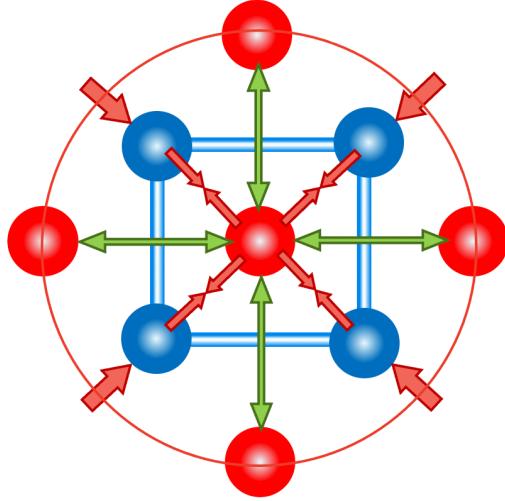


Figure 3.12: Schematic determination of  $J_1$  and  $J_2$  coefficients in 2D case.

Hence the total energy of the given ferromagnetic system can be written as:

$$E_{AFM} = E_g - 4J_1 + 4J_2 \quad (3.16)$$

Where  $E_g$  – corresponds for the part of energy given by the geometry itself.

### Solving the system

Extracting values of exchange couplings from the produced systems were done in two different methodologies.

In the case of a nonzero determinant, it is possible to obtain an exact solution directly solving the system.

$$\det \begin{pmatrix} n_{11} & \dots & n_{1N} \\ \dots & & \dots \\ n_{21} & \dots & n_{NN} \end{pmatrix} \neq 0 \quad (3.17)$$

Otherwise, the less stable (with the highest energy) AFM configuration is excluded from the calculations until the equations system becomes solvable or runs out of the structures. This method has obvious disadvantages. Firstly quite frequently from the mapped coefficients, it's impossible to build any solvable system, and the calculations fail. Secondly, excluding less favorable spin configuration from the solution, one by one, frequently leads to a minimal possible system  $2 \times 2$ , which returns only the  $J_1$  parameter and thus makes the results poorly representative for the further  $T_C$  estimation.

The second methodology utilizes a common strategy applicable for overdetermined systems, namely the least-squares method. In this study, fitting was done starting from the  $N \times N$

system, continually decreasing the number of taken into account variables up to the  $N \times 2$  ( $E_g, J_1$ ). Such an approach allows to better distinguish strong and weak exchange correlations in the materials showing how every group of magnetic atoms located on a particular coordination sphere contribute to the final exchange values. Worth mentioning that contrary to the direct system solving this approach is extremely sensitive to the quality of all the produced AFM structures. In other words, any poorly relaxed or disordered physically meaningless structure that might appear in the system won't be excluded at any stage and strongly affect all the produced results. For clarity let's consider a couple of examples, namely FM Mn and Eu oxides.

### Example 1

For the case of EuO were generated 3 unique AFM cells (see figure 3.13). Based on the energies of given structures obtained from the spin-polarized DFT calculations, one may build a system of 3 equations that has 3 variables ( $E_g, J_1, J_2$ ). Since the coefficients obtained for a given system form a nonzero determinant, this system can be solved exactly.

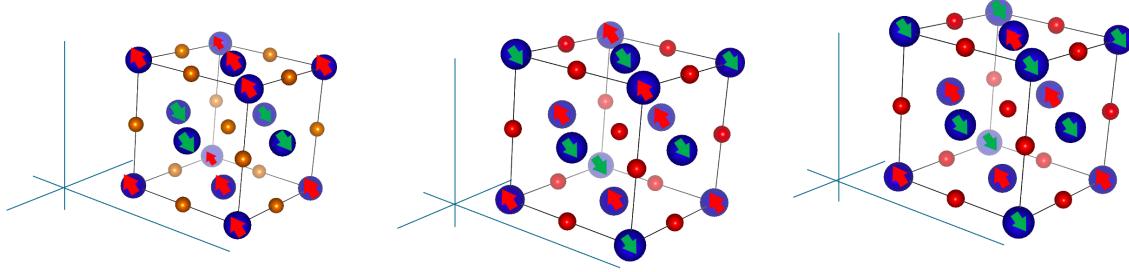


Figure 3.13: Three different spin configurations produced for EuO that allows to calculate the  $J_1$  and  $J_2$  exchange couplings.

$$\begin{cases} E_g + 4J_1 - 6J_2 = E_{AFM1} \\ E_g + 0J_1 + 6J_2 = E_{AFM2} \\ E_g - 4J_1 - 2J_2 = E_{AFM3} \end{cases} \quad (3.18)$$

$$\begin{aligned} E_g &= -41.63324 \text{ eV} \\ J_1 &= 1.18 \text{ meV} \\ J_2 &= 0.42 \text{ meV} \\ T_C^{GGA} &= 69 \text{ K} \\ T_C^{\text{exp.}} &= 67 \text{ K} \end{aligned} \quad (3.19)$$

Solution of the system allows calculating exchange coupling parameters for the first and second coordination spheres ( $J_1, J_2$ ). As it might be seen from the calculated values, in this case, exchange couplings have the order of meV. This fact once again points out the critical importance of extremely accurate relaxation and energy estimation steps. The calculated value of critical temperature 67K is in perfect agreement with the experimental value of 69K.

## Example 2

For the case of MnO were considered 3 spin configurations (parental FM and two generated AFM). In the same manner, for the given structures were performed accurate relaxation followed by single-point energy calculations.

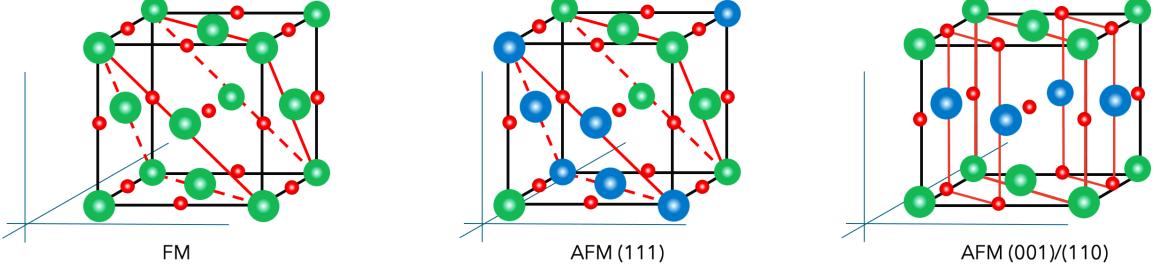


Figure 3.14: Three different spin configurations produced for MnO that allows to calculate the  $J_1$  and  $J_2$  exchange couplings.

$$\begin{cases} E_g + 6J_1 + 3J_2 = E_{FM} \\ E_g - 2J_1 + 3J_2 = E_{AFM1} \\ E_g + 0J_1 - 3J_2 = E_{AFM2} \end{cases} \quad (3.20)$$

$$\begin{aligned} J_1 &= 9.50 \text{ meV} \\ J_2 &= 14.91 \text{ meV} \\ T_C^{GGA+U} &= 249 \text{ K} \\ T_C^{exp.} &= 118 \text{ K} \end{aligned} \quad (3.21)$$

Obtained values of exchange coupling for the first ( $J_1$ ) and second coordination spheres ( $J_2$ ) on the order of several meV. In this case, the estimated value of critical temperature  $T_C^{GGA+U} = 249K$  overestimates the experimental one  $T_C^{exp.} = 118K$  more than two times. Nevertheless, results still might be considered valuable since they allow to distinguish structures with doubtful technological potential.

One more important point here is the quite a noticeable difference between the experimental values obtained for the system under consideration by different groups. For instance  $90K$  in ref. [38], while ref. [39] reported  $118K$  and the most recent results of  $85K$  in ref. [40]. This variability also raises how it is correct to compare theoretically estimated values with experimentally obtained. Since samples of the same material might have a different quality and even small defects might strongly affect the magnetic properties. While varying experimental techniques and differ in equipment might lead to even bigger divergence. This fact was also briefly discussed during the description of the data pre-processing in section 3.1.3

### 3.2.5 Monte Carlo simulation

As a final step of the workflow values of exchange coupling parameters ( $J_{ij}$ ) used to study the material transition from ferromagnetic (FM) to paramagnetic (PM) state. It is done by the source

of constrained Monte Carlo (MC) [41] metropolis algorithm as implemented in Vampire [42, 43] software which allows to carry out calculations of the equilibrium temperature-dependent magnetic properties. Vampire contains a predefined functionality to estimate critical temperature by performing a temperature sweep and calculating the average magnetization, returning the classic M-T curve.

For all the systems, exchange energy was considered isotropic (equal in all directions). To neglect finite-size and surface effects in a bulk material was chosen sufficiently large dimensionality of  $10 \times 10 \times 10 \text{ nm}^3$  and defined periodic boundary conditions in all three spatial directions. Temperature range of modeling was specified from 0 to  $1500 \text{ K}$  with an increment of  $25 \text{ K}$  for each temperature done  $10^5$  equilibration time steps which is needed to reach thermal equilibrium in the system followed by the  $10^5$  loop-time steps, which is already counted to the average of magnetization as it is illustrated in figure 3.15. Quite a large number of equilibration steps explained due to known equilibration-temperature dependence which tends to drastically slow down close to the Curie point. Specified numbers chosen heuristically but yet yields reasonable results for the most of systems. The bigger number of averaging steps, in general, lead to more smooth data but obviously requests more computational resources.

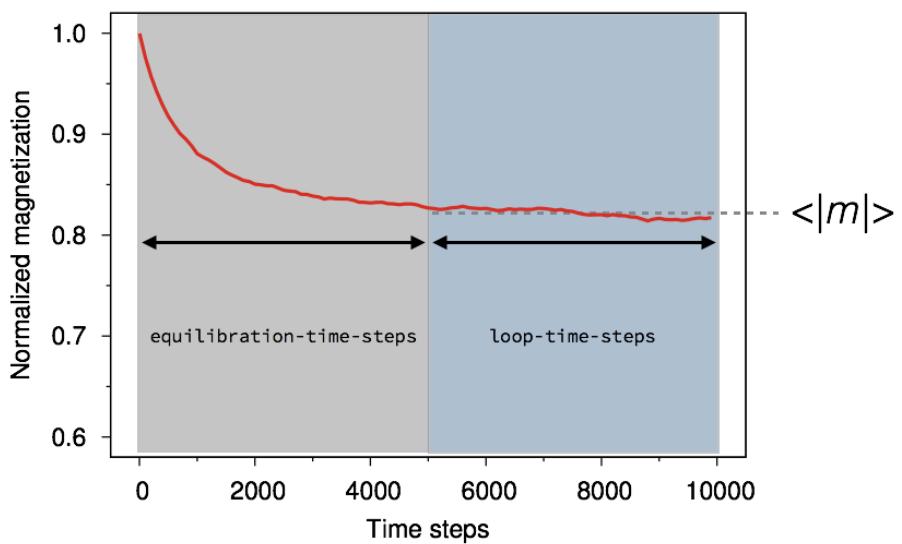


Figure 3.15: Time-dependent magnetisation caused by a stepwise increase in temperature.

As a result of running MC simulation returns data for the magnetization-temperature curve as it is shown in figure 3.16, from which one may determine the critical temperature as point with a maximum gradient.

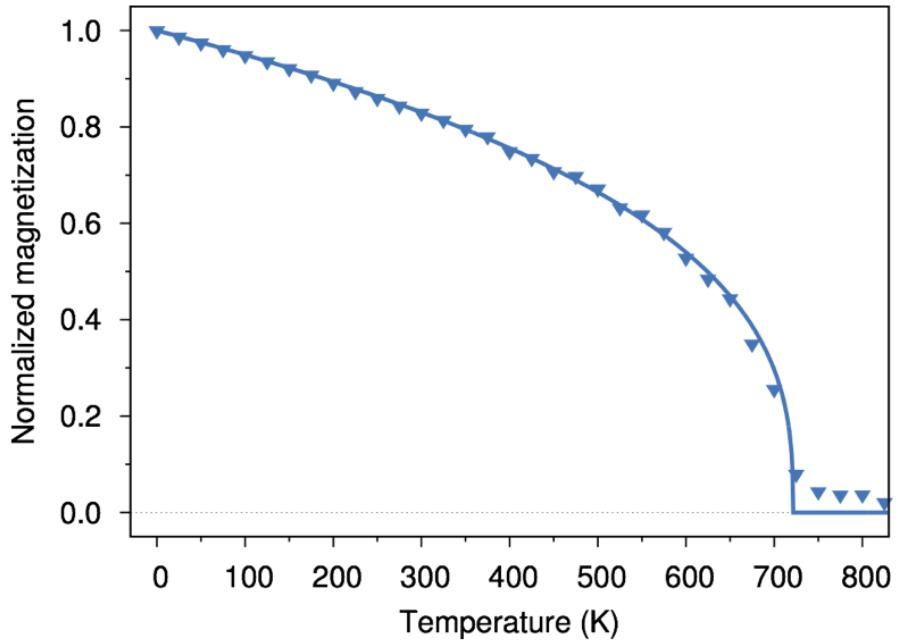


Figure 3.16: MC magnetization temperature curve.

### 3.2.6 End-to-end pipeline details

All pieces of the code for integration with side packages and data pre/post-processing produced during the work were formed into a complete end-to-end Python-based package. The main goal of the prepared package was to automate all the stages of the critical temperature calculations, eliminating the need for user control and making it usable for any FM material. Schematically algorithm shown in figure 3.17.

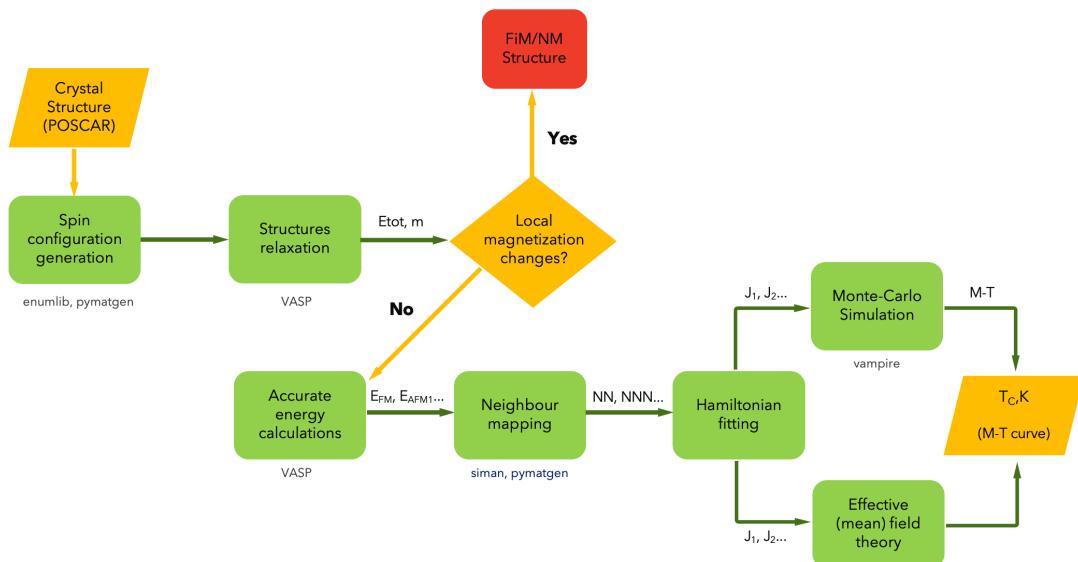


Figure 3.17: End-to-end pipeline scheme.

The package was built with the philosophy to be suitable even for people without special knowledge in electronic structure calculations keeping the entry barrier as low as possible. In the architecture of the code, we tried to save a balance between being simple for the novice yet maintaining flexibility for advanced users who know exactly what it does. As minimum input information its only needs to provide crystal structure data in a form of a commonly used POSCAR type file. All the required parameters for calculations by the DFT method are set by default. If necessary, the user can always change them by editing a JSON-type file using the keywords specified by the VASP manual (for the INCAR and KPOINTS files). This adjustment is more than welcomed if the user has any specific knowledge about the particular system, which may speed up the convergence or, in general, lead to more accurate results, but not mandatory.

At the moment, for the code to work correctly, user need installed *python 3.6* or later with a series of easily accessible through the Python Package Index (PyPI) repository libraries (*pymatgen*, *ase*, *siman*, *numpy*, *scipy*). Structure enumeration process required additional compilation of *enumlib* and its submodule *symlib* using any available fortran compiler (i.e., gfortran). Also, it is necessary to have installed program for *ab-initio* calculation *vasp-5.4.4* with a possibility to run in several threads.

The built package currently did not include the stage of MC simulation since they become excessively slow in a python implementation due to its interpretable nature. Therefore at the final stage, the critical temperature is estimated by the Curie–Weiss law, and for more accurate simulations, it's still necessary to run *Vampire* software explicitly. In the future iteration of the project, this disadvantage expected to be fixed.

All the described code available at the GitHub repository ([https://github.com/Volkov-da/curie\\_calculator](https://github.com/Volkov-da/curie_calculator)), which is planned to be maintained and updated.

## Chapter 4

# Results

### 4.1 ML models results

In section 3.1 was described the construction of several ML models aimed at predicting ferromagnetic materials critical temperature based exclusively on their chemical composition. The final summary of the best models' performance in comparison with reference work is presented in table 4.1.

Table 4.1: Summary of the best ML models performance

Metrics	SISSO	KRR	RF	XGB	RF (ref.)	KRR (ref.)
$R^2(test)$	0.618	0.793	0.833	0.845	0.87	0.72
$R^2(cross.val)$	-	0.787	0.828	0.837	0.81	-
MAE (K)	138	84	71	69	57	-
RMSE (K)	171	121	109	105	-	-
Max Error (K)	761	642	516	595	-	-

The highest accuracy demonstrated by the XGBoost model with  $R^2$  and mean absolute error of 0.845 and 69 K respectively on the test set. But the lowest maximum error is shown by the RF. It is worth to mention that descriptors used in this work are different from the one used in the reference. A serious drawback of considered features is the absence of any structural information which makes it impossible to distinguish between two compositions with similar stoichiometry. But after all this work might be considered as a second small step (after ref. [10] toward the ML-driven computational search for novel magnetic materials.

All the generic python code produced in this part of the work stored at GitHub repository ([https://github.com/Volkov-da/curie\\_calculator/tree/master/ML\\_model](https://github.com/Volkov-da/curie_calculator/tree/master/ML_model)). The whole process of data cleaning, preprocessing as well as model training and validation stored in a form of briefly commented jupyter-notebook (file *ML\_training.ipynb*).

The pipeline showed the best performance (MinMaxScaler + XGBoost) was additionally re-trained over the entire dataset and stored in easy reproducible JSON format.

## 4.2 DFT calculations results

Using the framework described in section 3.2 were calculated values of critical temperature for 20 binary structures containing pure metals, oxides, borides, etc., and 2 gadolinium containing ternary structures. Obtained values were compared with experimental ones as well as with the results from the theoretical calculations of other groups. Relative error was calculated as follows:

$$\Delta = \frac{|T_{exp} - T_{DFT}|}{T_{exp}} \quad (4.1)$$

As it might be seen from the table 4.2 in general GGA + U approach outperforms GGA. Mean relative error for the unary structures in the case of GGA is 39.5% for GGA + U is 30.9%, for oxides 40.4% and 33.5%, for other binary structures (borides, etc.) it is 42.4% and 25.1% respectively.

From the plot 4.1 we may conclude that in general GGA tends to underestimate critical temperature with the exclusions of 4 structures, namely: *MnO*, *NiO*, *MnAs*, *EuS*.

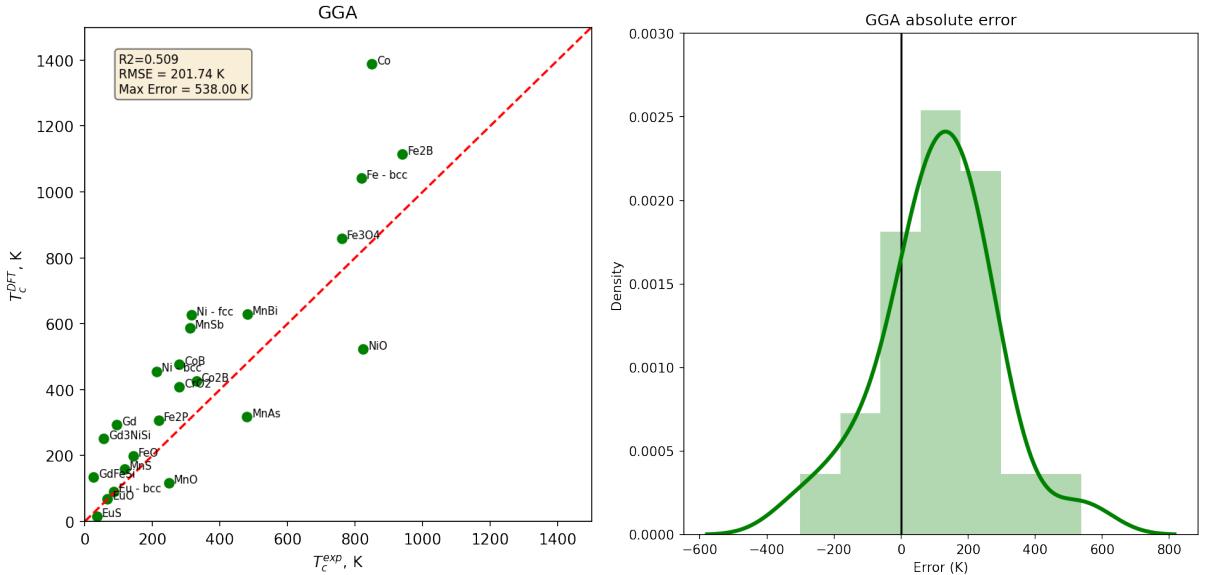


Figure 4.1: Left: Comparison of calculated by GGA results with experimental. Right: Absolute errors distribution.

Maximum relative errors of 125 % are shown in the case of *EuS*, but anyway, the resulting value we may consider pretty reasonable due to the small absolute difference (16 K and 36 K respectively). A maximum absolute error of 538 K appears in the case of cobalt, the FM material with the highest known critical temperature. The lowest absolute difference in calculated and experimental values was shown by three Eu containing structures its 5 K for *Eu-bcc*, 20 K for *EuS* and 2 K for *EuO* which is so far also the material with the lowest relative error of only 5.5%.

Results obtained during the GGA+U calculations presented in figure 4.2. The distribution of

errors is closer to zero compared to GGA results, and there is no clear tendency towards over/under-estimation of Curie point. Maximum relative error of 53.4% in this case shown by  $MnO$ , while a bit bigger maximum absolute error of 608 K once again by cobalt. Lowest absolute difference of 15 K and 23 K as well as the lowest relative error of 9.4% and 11.6% were shown by  $MnS$  and  $FeO$  respectively. Eu-containing structures were not considered in GGA+U calculations due to the absence of a predefined U value. Once again worth mentioning that due to the automated way of calculating, none of Hubbard correction values was calibrated explicitly. Hence we may expect even better results in the case of proper calibration with a linear response method for each particular structure.

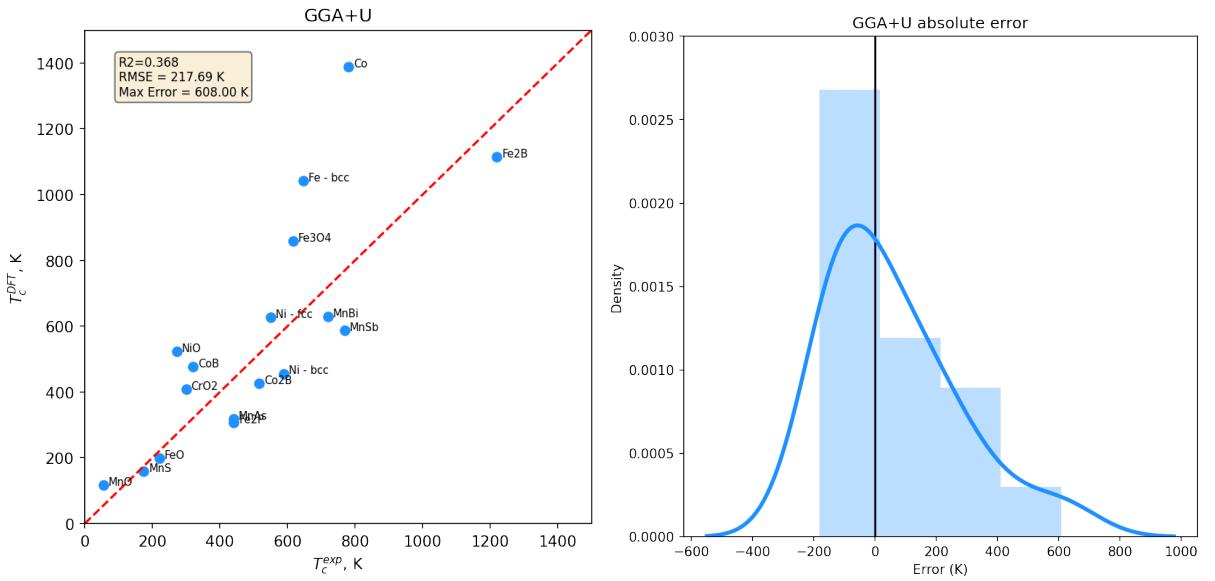


Figure 4.2: Left: Comparison of calculated by GGA+U results with experimental. Right: Absolute errors distribution.

Results obtained during GGA and GGA + U calculations with relative errors compared to the experimental data presented in the table 4.2.

Table 4.2: Comparison of calculated by GGA and GGA+U results with experimental.

Material	$T_C, K$				$\Delta, \%$	
	ref.	exp.	GGA	GGA+U	GGA	GGA+U
Fe-bcc [44]	950	1043	820	648	21.4	37.9
Co-fcc [44]	1311	1388	850	780	38.8	43.8
Ni-fcc [44]	300	627	315	550	49.8	12.3
Ni-bcc [45]	250	456	212	590	53.5	29.4
Gd [46]	293	294	94	-	68.0	-
Eu-bcc	111	91	86	-	5.5	-
EuO [47]	35	69	67	-	2.9	-
$Fe_3O_4$	-	858	760	616	11.4	28.2
$CrO_2$ [48]	305	408	280	300	31.4	26.5
MnO [39, 49]	240	118	249	55	111.0	53.4
NiO [39, 49]	393	523	824	272	57.6	48.0
$Fe_2B$ [50]	1000	1115	940	1220	15.7	9.4
$Co_2B$ [50]	450	426	330	516	22.5	21.1
CoB	411	477	280	320	41.3	32.9
$Fe_2P$	-	306	219	441	28.4	44.1
MnBi	-	630	481	720	23.7	14.3
MnSb	-	587	311	769	47.0	31
MnAs	-	318	480	440	50.9	38.4
EuS [51]	-	16	36	-	125.0	-
$Gd_3NiSi_2$ [52]	215	251	56	-	77.7	-
GdFeSi	145	135	25	-	81.5	-

### 4.3 Novel structures study

At a final stage of all the performed work develop methods for critical temperature calculations we applied in a direct exploration of newly predicted magnetic structures. As mentioned before, nowadays, it's already possible to perform a computational search of such structures but optimizing them only concerning stability and magnetic moment. So as a final step of this work, a developed in section 3.2 algorithm based on DFT + MC calculations and the best-trained ML model from section 3.1.4 was applied to benchmark new structures with respect to their Curie temperature. Namely were tested 55 promising magnetic compositions obtained during coevolutionary (Mendelevian) search as implemented in MendS alghoritm [53].

Distribution of estimated critical temperatures presented in the figure 4.3. As it is shown, the correlation between the results produced by the two described methodology is quite weak. Generally, it might be explained due to the nature of the descriptors used in ML model training. Since prepared features were exclusively based on chemical information without considering structural data, the whole predictive model tends to predict high values of critical temperature for the compositions with the close to known in model stoichiometry. Thus even for previously unknown Fe and Co-containing structures ML model always expect values of critical temperature higher than 400 K in the same fashion as for the known from training data compositions containing these elements.

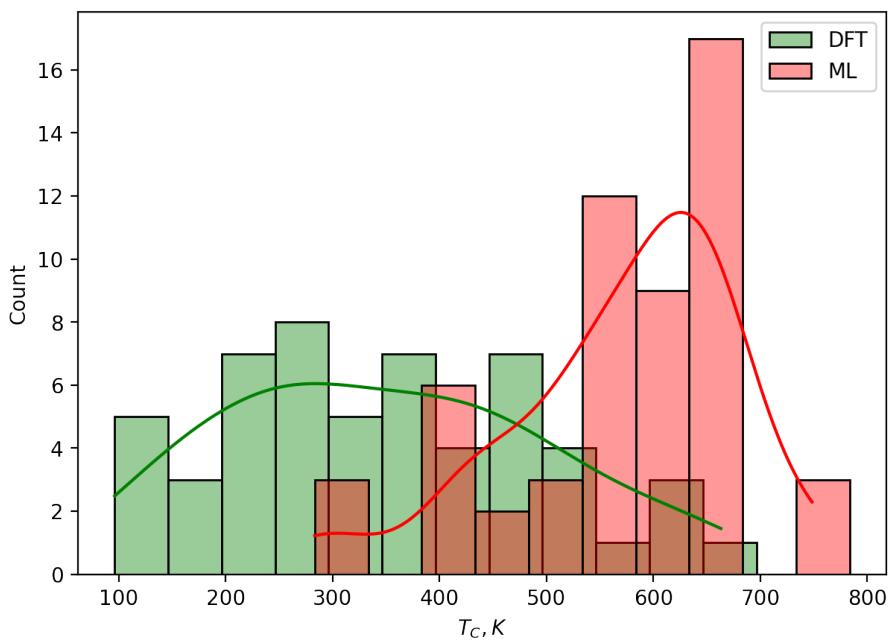


Figure 4.3: Histogram of estimated critical temperature using ML and DFT.

Estimation given for the new structures by DFT + MC approach (green histogram) has a more uniform distribution and closer in shape to the experimental critical temperature distribution obtained for known FM materials from figure 3.1. Of course, such a comparison might seem incorrect but intuitively its looks like the more new structures will be tested the more these distributions should have in common.

According to the produced results tree newly predicted structures were classified as a materials with high critical temperature namely  $C2/m - Fe_6Mo$  with  $T_C^{DFT} = 630K$  and  $T_C^{ML} = 479K$ , stable  $Cmmm - HfCo_5$  with  $T_C^{DFT} = 633K$  and  $T_C^{ML} = 425K$  and metastable  $F\bar{4}3m - HfCo_5$  with  $T_C^{DFT} = 663K$  and  $T_C^{ML} = 425K$ .

# Chapter 5

## Discussion

### 5.1 Results reasoning

Results obtained in this work were quite expected and, in general, showed a comparable level of accuracy with existing theoretical studies. Estimated values of critical temperature obtained by GGA and GGA + U on average differed from the experimental on 40% and 30%, respectively. The reasoning for such deviation comes from the following sources of errors. First of all, despite the fact that DFT theoretically is exact in practice accuracy of calculations is strongly dependent on the choice of exchange-correlation functional. General gradient approximation (GGA) used in this work tends to show good accuracy for the majority of materials properties still has known shortages in magnetic calculations [54]. The difficulties in accurate estimation of magnetic exchange coupling was discussed in several theoretical studies [55–57]. For instance, in ref. [58] was reported statistical error of 40% for the number of transitional metals complexes containing Fe, Co, Mn. At the same time, other work reported an even bigger divergence of almost 11 times compared to experimental values [57]. Secondly used in this work Heisenberg spin model still leave out a lot of complexity of real material due to a number of simplifications and approximations. Thirdly inaccuracies might appear on a stage of Monte Carlo simulations. All the materials were treated as purely isotropic, which might lead to a certain level of errors. Also, we cannot fully deny the possible contribution of size effects since the finite size of modeled cube. Finally, it is worth admitting that taking into account all possible factors by theoretical methods is beyond our control. Magnetism is an extremely complex phenomenon in which properties should be modeled on all size scales starting from electronic and atomistic, ending up on a macroscale level. Even a small amount of unaccounted impurities or defects can cause a significant scatter in the experimental results, but at the same time not be reflected in the existing theoretical models.

The ML model developed in this work showed good predictive power with the  $R^2$  value of 0.87 and MAE of 57  $K$  despite the pretty limited descriptors space. The prepared model complements a rigorous and demanding DFT + MC approach and when it becomes possible to include structural features eventually can replace it. Currently this model applicable for rapid prototyping which might be performed even in extremely limited computational resources.

## 5.2 Future Work

The project presented in this master's thesis is planned to be continued. Part of the work related to DFT + MC is planned to be additionally tested concerning exchange-correlation energy functionals. The nearest future planned to run tests using Local-density approximations LDA, LDA + U and possibly hybrid functionals. Additionally, it is planned to develop an end-to-end methodology for magnetic anisotropy energy calculations which further might be considered for more accurate Monte Carlo simulations. The prepared Python package is also planned to be maintained and improved with respect to parallelization and memory usage.

Part of the work related to the data-driven estimation of critical temperature also seems to be extremely promising. Despite the absence of structural information in descriptors, it already shows good predictive power. Inclusion of such information into the descriptors is already well-established practice by the source of various fingerprints. At the same time, the usage of modern architectures of Neural Networks (i.e., Graph Convolutional NN) may lead to stunning results frequently outperforming classical electronic structures calculations and obviously using much fewer computational resources. The only stopping factor here is the absence of any reliable and sufficiently big magnetic structures database at the moment. Fortunately, this only a matter of time, and sooner or later author of this work or one of many groups worldwide will create it.

## 5.3 Conclusion

1. In this work, a methodology for calculating the critical magnetic temperature using DFT and subsequent Monte Carlo simulations based on the Ising spin model was developed and tested. Results obtained during these calculations showed an average divergence of 30% with experimental values but yet in the majority of cases allows to distinguish materials with high critical temperature.
2. As a competitive approach, several regression machine learning models have been trained on the descriptors solely based on the chemical composition of the material. Nevertheless, the best-trained model showed high predictive power, making this approach comparable in accuracy to demanding electronic structure calculations.
3. Both approaches have been applied to determine the critical temperature of 55 newly predicted structures with promising magnetic properties. For three of the studied structures, a high value of the critical temperature was estimated, which allows us to assume their possible technological potential.
4. Two declared methodologies are implemented as a Python library. A simple interface and setting most of the required numeric parameters by default minimizes the need for end-user management and, in most cases, fully automates the calculations.

# Bibliography

- [1] O. Gutfleisch, M. A. Willard, E. Brück, C. H. Chen, S. G. Sankar, and J. P. Liu, “Magnetic materials and devices for the 21st century: Stronger, lighter, and more energy efficient,” *Advanced Materials*, vol. 23, no. 7, pp. 821–842, 2011.
- [2] A. R. Oganov and C. W. Glass, “Crystal structure prediction using ab initio evolutionary techniques: Principles and applications,” *The Journal of Chemical Physics*, vol. 124, p. 244704, 2021/05/31 2006.
- [3] A. R. Oganov, A. O. Lyakhov, and M. Valle, “How evolutionary crystal structure prediction works—and why,” *Accounts of Chemical Research*, vol. 44, pp. 227–237, 03 2011.
- [4] A. O. Lyakhov, A. R. Oganov, H. T. Stokes, and Q. Zhu, “New developments in evolutionary structure prediction algorithm uspex,” *Computer Physics Communications*, vol. 184, no. 4, pp. 1172–1182, 2013.
- [5] W. Eerenstein, N. Mathur, and J. F. Scott, “Multiferroic and magnetoelectric materials,” *nature*, vol. 442, no. 7104, pp. 759–765, 2006.
- [6] Y. W. Ma, Y. H. Lu, J. B. Yi, Y. P. Feng, T. S. Herring, X. Liu, D. Q. Gao, D. S. Xue, J. M. Xue, J. Y. Ouyang, and J. Ding, “Room temperature ferromagnetism in teflon due to carbon dangling bonds,” *Nature Communications*, vol. 3, no. 1, p. 727, 2012.
- [7] Y. Liu, N. Tang, X. Wan, Q. Feng, M. Li, Q. Xu, F. Liu, and Y. Du, “Realization of ferromagnetic graphene oxide with high magnetization by doping graphene oxide with nitrogen,” *Scientific Reports*, vol. 3, no. 1, p. 2566, 2013.
- [8] M. Shahjahan, I. Razzakul, and M. Rahman, “First-principles calculation of stable magnetic state and curie temperature in transition metal doped iii-v semiconductors,” *Computational Condensed Matter*, vol. 9, pp. 67 – 71, 2016.
- [9] V. Stanev, C. Osés, A. G. Kusne, E. Rodriguez, J. Paglione, S. Curtarolo, and I. Takeuchi, “Machine learning modeling of superconducting critical temperature,” *npj Computational Materials*, vol. 4, no. 1, p. 29, 2018.
- [10] J. Nelson and S. Sanvito, “Predicting the curie temperature of ferromagnets using machine learning,” *Physical Review Materials*, vol. 3, pp. 104405–, 10 2019.

- [11] Y. Xu, M. Yamazaki, and P. Villars, “Inorganic materials database for exploring the nature of material,” *Japanese Journal of Applied Physics*, vol. 50, 2011.
- [12] “Bibliography of magnetic materials and tabulation of magnetic transition temperatures.,” 1968.
- [13] K. Rajan, “Materials informatics,” *Materials Today*, vol. 8, no. 10, pp. 38–45, 2005.
- [14] K. Rajan, “Materials informatics: The materials “gene”and big data,” *Annual Review of Materials Research*, vol. 45, pp. 153–169, 2021/04/27 2015.
- [15] T. Lookman, F. J. Alexander, and A. R. Bishop, “Perspective: Codesign for materials science: An optimal learning approach,” *APL Materials*, vol. 4, p. 053501, 2021/04/27 2016.
- [16] A. Agrawal and A. Choudhary, “Perspective: Materials informatics and big data: Realization of the “fourth paradigm”of science in materials science,” *APL Materials*, vol. 4, p. 053208, 2021/04/27 2016.
- [17] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, “A general-purpose machine learning framework for predicting properties of inorganic materials,” *npj Computational Materials*, vol. 2, no. 1, p. 16028, 2016.
- [18] L. Ward, A. Dunn, A. Faghaninia, N. E. R. Zimmermann, S. Bajaj, Q. Wang, J. Montoya, J. Chen, K. Bystrom, M. Dylla, K. Chard, M. Asta, K. A. Persson, G. J. Snyder, I. Foster, and A. Jain, “Matminer: An open source toolkit for materials data mining,” *Computational Materials Science*, vol. 152, pp. 60–69, 2018.
- [19] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, and G. Louppe, “Scikit-learn: Machine learning in python,” *Journal of Machine Learning Research*, vol. 12, 01 2012.
- [20] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD ’16, (New York, NY, USA), pp. 785–794, Association for Computing Machinery, 2016.
- [21] R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, and L. M. Ghiringhelli, “Sisso: A compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates,” *Phys. Rev. Materials*, vol. 2, p. 083802, Aug 2018.
- [22] R. Ouyang, E. Ahmetcik, C. Carbogno, M. Scheffler, and L. M. Ghiringhelli, “Simultaneous learning of several materials properties from incomplete databases with multi-task SISSO,” *Journal of Physics: Materials*, vol. 2, p. 024002, mar 2019.

- [23] L. M. Ghiringhelli, J. Vybiral, S. V. Levchenko, C. Draxl, and M. Scheffler, “Big data of materials science: Critical role of the descriptor,” *Phys. Rev. Lett.*, vol. 114, p. 105503, Mar 2015.
- [24] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, “Python materials genomics ( pymatgen): A robust, open-source python library for materials analysis,” *Computational Materials Science*, vol. 68, pp. 314–319, 2013.
- [25] G. L. W. Hart and R. W. Forcade, “Algorithm for generating derivative structures,” *Phys. Rev. B*, vol. 77, p. 224115, Jun 2008.
- [26] G. L. W. Hart and R. W. Forcade, “Generating derivative structures from multilattices: Algorithm and application to hcp alloys,” *Phys. Rev. B*, vol. 80, p. 014120, Jul 2009.
- [27] G. L. Hart, L. J. Nelson, and R. W. Forcade, “Generating derivative structures at a fixed concentration,” *Computational Materials Science*, vol. 59, pp. 101–107, 2012.
- [28] W. S. Morgan, G. L. W. Hart, and R. W. Forcade, “Generating derivative superstructures for systems with high configurational freedom,” *Computational Materials Science*, vol. 136, pp. 144–149, 2017.
- [29] Y. Zhu, X. Kong, T. D. Rhone, and H. Guo, “Systematic search for two-dimensional ferromagnetic materials,” *Phys. Rev. Materials*, vol. 2, p. 081001, Aug 2018.
- [30] Z. Jiang, P. Wang, J. Xing, X. Jiang, and J. Zhao, “Screening and design of novel 2d ferromagnetic materials with high curie temperature above room temperature.,” *ACS Appl Mater Interfaces*, vol. 10, pp. 39032–39039, Nov 2018.
- [31] G. Kresse and J. Furthmüller, “Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set,” *Computational Materials Science*, vol. 6, no. 1, pp. 15–50, 1996.
- [32] G. Kresse and D. Joubert, “From ultrasoft pseudopotentials to the projector augmented-wave method,” *Phys. Rev. B*, vol. 59, pp. 1758–1775, Jan 1999.
- [33] J. P. Perdew, K. Burke, and M. Ernzerhof, “Generalized gradient approximation made simple,” *Phys. Rev. Lett.*, vol. 77, pp. 3865–3868, Oct 1996.
- [34] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, “Electron-energy-loss spectra and the structural stability of nickel oxide: An lsda+u study,” *Phys. Rev. B*, vol. 57, pp. 1505–1509, Jan 1998.

- [35] L. Wang, T. Maxisch, and G. Ceder, “Oxidation energies of transition metal oxides within the GGA + U framework,” *Phys. Rev. B*, vol. 73, p. 195107, May 2006.
- [36] M. Cococcioni and S. de Gironcoli, “Linear response approach to the calculation of the effective interaction parameters in the LDA + U method,” *Phys. Rev. B*, vol. 71, p. 035105, Jan 2005.
- [37] D. Aksyonov, S. Fedotov, K. Stevenson, and A. Zhugayevych, “Understanding migration barriers for monovalent ion insertion in transition metal oxide and phosphate based cathode materials: A dft study,” *Computational Materials Science*, vol. 154, pp. 449–458, 2018.
- [38] M. E. Lines and E. D. Jones, “Antiferromagnetism in the face-centered cubic lattice. ii. magnetic properties of mno,” *Phys. Rev.*, vol. 139, pp. A1313–A1327, Aug 1965.
- [39] W. L. Roth, “Magnetic structures of mno, feo, coo, and nio,” *Phys. Rev.*, vol. 110, pp. 1333–1341, Jun 1958.
- [40] G. Pepy, “Spin waves in MnO; from 4 °k to temperatures close to  $t_n$ ,” *Journal of Physics and Chemistry of Solids*, vol. 35, no. 3, pp. 433–444, 1974.
- [41] P. Asselin, R. F. L. Evans, J. Barker, R. W. Chantrell, R. Yanes, O. Chubykalo-Fesenko, D. Hinzke, and U. Nowak, “Constrained monte carlo method and calculation of the temperature dependence of magnetic anisotropy,” *Phys. Rev. B*, vol. 82, p. 054415, Aug 2010.
- [42] R. F. L. Evans, W. J. Fan, P. Chureemart, T. A. Ostler, M. O. A. Ellis, and R. W. Chantrell, “Atomistic spin model simulations of magnetic nanomaterials,” *Journal of Physics: Condensed Matter*, vol. 26, p. 103202, feb 2014.
- [43] R. F. L. Evans, U. Atxitia, and R. W. Chantrell, “Quantitative simulation of temperature-dependent magnetization dynamics and equilibrium properties of elemental ferromagnets,” *Phys. Rev. B*, vol. 91, p. 144425, Apr 2015.
- [44] M. Pajda, J. Kudrnovský, I. Turek, V. Drchal, and P. Bruno, “Ab initio calculations of exchange interactions, spin-wave stiffness constants, and curie temperatures of fe, co, and ni,” *Phys. Rev. B*, vol. 64, p. 174402, Oct 2001.
- [45] C. S. Tian, D. Qian, D. Wu, R. H. He, Y. Z. Wu, W. X. Tang, L. F. Yin, Y. S. Shi, G. S. Dong, X. F. Jin, X. M. Jiang, F. Q. Liu, H. J. Qian, K. Sun, L. M. Wang, G. Rossi, Z. Q. Qiu, and J. Shi, “Body-centered-cubic Ni and its magnetic properties.,” *Phys Rev Lett*, vol. 94, p. 137210, Apr 2005.

- [46] E. A. Tereshina, S. Khmelevskyi, G. Politova, T. Kaminskaya, H. Drulis, and I. S. Tereshina, “Magnetic ordering temperature of nanocrystalline Gd: enhancement of magnetic interactions via hydrogenation-induced “negative” pressure,” *Scientific Reports*, vol. 6, no. 1, p. 22553, 2016.
- [47] N. Jutong, U. Eckern, T. Mairosen, and U. Schwingenschlögl, “Effect of Gd doping and O deficiency on the curie temperature of EuO.,” *Scientific Reports*, vol. 5, no. 1, p. 8038, 2015.
- [48] S. Huang, X. Wu, J. Niu, and S. Qin, “Structural, magnetic and electronic properties of CrO<sub>2</sub> at multimegarbar pressures,” *RSC Adv.*, vol. 8, pp. 24561–24570, 2018.
- [49] T. Archer, C. D. Pemmaraju, S. Sanvito, C. Franchini, J. He, A. Filippetti, P. Delugas, D. Pugnioni, V. Fiorentini, R. Tiwari, and P. Majumdar, “Exchange interactions and magnetic phases of transition metal oxides: Benchmarking advanced ab initio methods,” *Phys. Rev. B*, vol. 84, p. 115114, Sep 2011.
- [50] M. Däne, S. Kyung Kim, M. P. Surh, D. Åberg, and L. X. Benedict, “Density functional theory calculations of magnetocrystalline anisotropy energies for ( fe1-xcox)2b,” vol. 27, no. 26, p. 266002, 2015.
- [51] S. V. Houten, “Magnetic interaction in EuS, EuSe, and EuTe,” 1962.
- [52] X. Liu and Z. Altounian, “First-principles calculation on the curie temperature of Gd<sub>3</sub>NiSi<sub>2</sub>,” *Journal of Physics: Condensed Matter*, vol. 21, p. 416002, 09 2009.
- [53] Z. Allahyari and A. R. Oganov, “Coevolutionary search for optimal materials in the space of all possible compounds,” *npj Computational Materials*, vol. 6, no. 1, p. 55, 2020.
- [54] A. D. Becke, “Density-functional exchange-energy approximation with correct asymptotic behavior,” *Phys. Rev. A*, vol. 38, pp. 3098–3100, Sep 1988.
- [55] E. Ruiz, S. Alvarez, J. Cano, and V. Polo, “About the calculation of exchange coupling constants using density-functional theory: The role of the self-interaction error,” *The Journal of Chemical Physics*, vol. 123, no. 16, p. 164110, 2005.
- [56] E. Ruiz, “Exchange coupling constants using density functional theory: Long-range corrected functionals,” *Journal of Computational Chemistry*, vol. 32, no. 9, pp. 1998–2004, 2011.
- [57] R. P. Joshi, J. J. Phillips, K. J. Mitchell, G. Christou, K. A. Jackson, and J. E. Peralta, “Accuracy of density functional theory methods for the calculation of magnetic exchange couplings in binuclear iron(iii) complexes,” *Polyhedron*, vol. 176, p. 114194, 2020.

- [58] J. J. Phillips and J. E. Peralta, “Magnetic exchange couplings from semilocal functionals evaluated nonself-consistently on hybrid densities: Insights on relative importance of exchange, correlation, and delocalization,” *Journal of Chemical Theory and Computation*, vol. 8, pp. 3147–3158, 09 2012.