

**Doctoral School of the University of Torino**

**PhD Program in Chemical and Materials Sciences - XXXV cycle**

**Title of the thesis: Applications of Machine learning in Materials Science**

**II** **year PhD Progress Report**

**PhD Student: Udaykumar Gajera**

**Supervisor: Dr. Silvia Picozzi**

Insert here a graphical abstract to capture the reader’s attention and give the reader a quick visual impression of the essence your PhD research, without providing specific results.



**Head of the Doctoral School:**

Prof. Alberto Rizzuti

**Coordinator of the Program:**

Prof. Bartolomeo Civalleri

**Part A: Training**

**A1: Courses and Schools**

**Summer/Winter schools**

Insert here the schools attended. **1 day = 1 CFU**

* Name of the school, location (Place, Country), date. CFU =
* “Setting gateway to academia” organized online by the University of Cambridge, 7-9 April 2021, CFU=3
* MaX School on “Advanced Materials and Molecular Modeling with Quantum ESPRESSO” organized online by ICTP, 17 - 28 May 2021, CFU=10
* School “From fundamental properties of matter to magnetic materials and applications” organized online by EMA (European magnetism Association) on Sep 6–17 September 2021, CFU=10
* “Magnetoelectricity in biomedicine: Healthcare for the 21st century” organized by ETH Zurich, Switzerland, 8-12 November 2021, CFU=4

**TOTAL CFU = 27**

**PhD Courses**

Insert here the list of classes, seminars and conferences attended. **4 h = 1 CFU**

| **Course title** | **Professor** | **University/ Department** | **Hours** | **CFU** |
| --- | --- | --- | --- | --- |
| From biomass to chemicals and biopolymers towards a sustainable future | Silvia Tabasso | chimica industriale | 12 | 3 |
| Introduction to scientific programming | Alessandro Erba | chimica fisica | 20 | 5 |
| Chemical Sensors for Scientific Research and Everyday Life | Ornella Abollino | chimica analitica | 8 | 2 |
| Electrochemical Energy Storage and Conversion Systems | Mauro Sgroi | Materials Sciences | 12 | 3 |
|  |  |  |  |  |

**TOTAL CFU = 13**

**TOTAL CFU(A1) =20**

(max = 20 CFU)

**A2: Participation to seminars and congresses/workshops**

**Seminars**

Insert here the list of all the seminars attended.

**4 h = 1 CFU**

* “Machine Learning Meets Chemistry” organized by Department of Chemistry, University of Torino, 17-18 February 2020, CFU=2
* “From Stimuli Responsive and Shape Memory Materials to Algorithmic Materials that Mimic Behaviorists Classical Conditioning “, prof. Olli Ikkala, Virtual DCMIC Seminar, 22 May 2020. CFU= 1
* Raman Days-2021, Prof. Anna Maria Ferrero, PhD doctorate in Earth Science, 28-29 January, Torino, CFU=2

**TOTAL CFU = 5**

**Congresses/Workshops**

Insert here the list of all the congresses/workshops attended.

**1 day = 1 CFU, max = 5 CFU**

* Title, speaker, location, date, duration. CFU =
* “Fundamentals of magnetoelectricity and how it can boost energy efficiency” organized by CNR, 9-10 September 2020, CFU=3.5
* “Training on synthetic and Magnetoelectric characterization methodologies” organized by the Leibniz Institute for Solid State and Materials Research Dresden (IFW), 7-11 December 2020, CFU=4

**TOTAL CFU = 7.5**

**TOTAL CFU(A2) =10**

(max = 10 CFU)

**A3: Assistance for teaching**

Insert here the experiences in teaching assistance.

**10 h = 1 CFU, max = 40 h/year**

* Period, course, Degree, year, professor, number of hours. CFU =

**TOTAL CFU(A3) =0**

**Part B: Research outcomes**

**B1: Second year presentation**

Evaluation A (3 CFU) ☐; B (2 CFU) ☐; C (1 CFU) ☐

**TOTAL CFU(B1) =**

**B2: Presentations at congresses/workshops**

**Presentations at Conferences and Congresses/Workshops**

Insert here the list of all the contributions presented at congresses/workshops.

I = International, N = national, O = oral, P = poster; bold: presenting author.

**[I,O] = 3 CFU; [I,P] = 1 CFU; [N, O] = 2 CFU; [N, P] = 0.5 CFU**

* [I; O] **N. Surname**, N. Surname,…, Title of the paper, *Journal*, **year**, pages, DOI. CFU = 3
* **U. Gajera**, H. Qin, H. Tan, S. Picozzi, S. Dijken, “Magnonic Neural Network”- Magnetoelectricity in biomedicine: healthcare for the 21st century, November 12, 2021, ETH-Zurich, Switzerland, CFU=3

**TOTAL CFU(B2) =3**

**B3: Period abroad/workshops**

Insert here the periods spent abroad.

**3 CFU each consecutive month, max = 18 CFU**

* Period, Institution/Department, Supervisor
* From 16 September 2021 to 31October 2021 at the Department of Applied Physics, Aalto University under supervision of Prof. Sebastiaan van Dijken.

**TOTAL CFU(B3) = 4.5**

(max = 18 CFU)

**B4: Papers**

Insert here the list of the papers/patents published (or accepted) so far.

**First author = 4 CFU; co-author = 2 CFU; max = 16 CFU.**

* [I] **U. Gajera**, L. Storchi, D. Amoroso, F. Delodovici, S. Picozzi, “Towards machine learning for microscopic mechanisms: a formula search for crystal structure stability based on atomic properties”, *AIP Journal of Applied Physics(submitted)*, **2022**, CFU = 0

**TOTAL CFU(B4) =0**

(max = 16 CFU)

**B5: Patents**

Insert here the list of filed patents.

**international = 2 CFU; national = 1 CFU.**

* N. Surname, N. Surname,…, Patent number, **year**. CFU = Select
* N. Surname, N. Surname,…, Patent number, **year**. CFU = Select

**TOTAL CFU(B5) =**

**C. SUMMARY OF SCIENTIFIC ACTIVITY** (max 3 pages)

**Research Question / Subject presentation**

*Give a short explanation of the main objectives of your PhD, in relation to a more general context. No more than 15 lines.*

Materials design is currently undergoing a revolution, driven by the increasing importance of machine learning (ML) approaches to boost materials discovery. While ML for material simulation is rather unshaped at present, it is likely to develop in the next decade, with all materials' classes (including functional compounds such as magnetoelectrics and multiferroics) expected to dramatically benefit from artificial intelligence. The report shows a brief introduction on the implementation of three different machine learning approaches for three different material science applications: 1) Construction of descriptors to predict the energetics of different crystallographic phases (i.e. total energy difference between Rock-salt (RS) and Zinc-blend (ZB) structures in binary semiconductors using Ab-initio data). 2) Designing microscopic magnonic devices using Genetic Algorithm (GA), 3) Prediction of intrinsic magnetic properties of materials from images using Convolutional Neural Network (CNN).

**Short review of the state-of-the-art**

*Briefly summarize the main assessments in the literature and the still open question. How is your work inserted in the existing literature? No more than 25 lines.*

**1) Construction of the descriptors to predict the total energy difference between Rock-salt (RS) and Zinc-blend (ZB) structures using Ab-initio data**

Combining first-principles simulations based on density functional theory (DFT) and machine learning approaches using regression methods seems a powerful tool for future materials design. Ghiringhelli et al. show a prototypical case in material science where the author uses the ML approach to classify the most stable crystal structure between RS and ZB for semiconductor AB binary compounds. In addition to that Park, et al. (Mach. learn.: sci. technol., 2, 025030, 2021) predicted cubic phase stability of perovskites using ML and DFT calculations.

STATE OF THE ART: add one or two sentences on Ghiringhelli work and other’s in the same field Within this context, our general ML approach has been to model a specific property by finding the more accurate descriptor, correlating different fundamental atomic quantities, by using regression methods such as Linear (LR) and combinatorial approaches, for descriptor selection. The descriptor has been modeled as a combination of simple functions (power, exponential) built of atomic quantities (electronegativity, electronic energy levels, atomic shells radii, etc.).

**2) Designing microscopic magnonic devices using Genetic Algorithm (GA).**

Energy consumption and heat dissipation represent significant problems with traditional silicon devices. Therefore, scientists are trying to overcome these drawbacks by proposing novel devices, i.e. GPUs can take care of intense graphic computation, MRAM can store more data with lower switching time, spin valves can be used for neuromorphic computing, etc.

In the present research, we propose a spin wave-based method to construct a magnonic device that can work as a physical Neural Network (NN). A spin wave can be defined as propagating disturbance in an ordered magnetic material; it propagates via the magnetic moments and, as such, no motion of electrical charge is involved, therefore theoretically dissipating no heat. We can imagine a magnonic device that performs a particular task of NN (like filtering of the images), reducing the computation load on the CPU. In order to build a magnonic device, we start from the one example from literature: Wang et al. (Nat. Comm. 12, 2636, 2021) successfully created a demultiplexer to separate two different spin-wave frequencies using a magnonic device. From that, we try to build a more complex physical classifier using a Genetic algorithm and a micromagnetic simulation code, MuMax.

**3) Prediction of intrinsic magnetic properties of materials from images using Convolutional Neural Network (CNN).**

Advanced machine learning techniques, like Neural Networks (NN), are currently applied in a wide variety of fields, including image classification in photos or face recognition on smartphones. Courtin et al.(EPL, 50, 94, 2000) analized magnetic domain patterns using the perceptron neural network model. Furthermore, Talpara et al (J. Magn. Magn. Mater, 448, 360) simulate magnetic microstructure in thin films. Therefore, in our case, we try to use advanced NN methods to predict intrinsic properties of these types of magnetic thin films.

ALSO HERE, TRY TO CITE SOME PREVIOUS WORKS TO GIVE AN IDEA OF THE STATE OF THE ART

In closer detail, we propose a method where we combine magnetic images of material from a single spectrophotometer (are you sure ?) and try to infer crucial spin Hamiltonian parameters affecting the magnetic behavior of a material. We showed an example related to the prediction of intrinsic parameters (like exchange length (), damping parameter (α), anisotropy constant ()) and external conditions (such as temperature, T), using the Convolutional Neural Network (CNN).

**Methodology and progress**

*Give a short description of the experimental and/or theoretical methods you have adopted so far for the achievement of the objectives. Briefly summarize the main achievements so far and explains any obstacle, dilemmas, or setbacks. No more than 1 page.*

**1) Construction of the descriptors to predict the energy difference between of Rock-salt (RS) and Zinc-blend (ZB) structures using Ab-initio data**

We systematically constructed a set of possible features by searching for accurate descriptors and used LR to identify a final “best formula” on our test case: the energetic stabilization of Rocksalt-Zincblende (RS-ZB) phases for 82 binary compounds. From this approach, we achieved; 1) Formulas identified by our approach can successfully classify the stable phase between RS and ZB and quantitatively predict the energy difference with good accuracy. 2) From the selected formulas by LR, we also learn that atomic shell radii (in particular the p shell of the cation) play an essential role in deciding the stable crystalline phase. 3) We further analyzed the importance of individual atomic properties inside the ML-suggested formulas by using a procedure called “formula optimization”. This procedure improves the performances of the formulas at a relatively small computational cost.

**2) Designing microscopic magnonic devices using Genetic Algorithm (GA).**

To design the magnonic device, we adopted the methodology of GA to find an optimized structure that can separate two frequencies in two output terminals from one input terminal. For separating two frequencies, we created a number of defects inside the structure using square holes. GA will optimize the location of these defects in order to find optimum separation between two frequencies. Furthermore, we used MuMax to simulate spin-wave propagation in YIG material. In the search for optimized structures, we found that GA can optimize the structure more efficiently than the method used by Wang et al. Therefore, we started simulating more complex devices, where - instead of separating two different frequencies - we use the same frequency with nine inputs representing 3x3 pixels and three outputs for three classes. In this case, we can differentiate between three different classes of information. Here, the major problem we faced is that the time to simulate the spin-wave for different structures is significantly higher because of the larger structure. Therefore, we are further trying to optimize the method to speed-up the simulations.

**3) Prediction of intrinsic magnetic properties of materials from images using Convolutional Neural Network (CNN).**

In order to do that, first, we created 800 artificial images of 2µmX2µm specimen of CoPd magnetic material using MuMax and used these images to train the NN model, so as to predict four parameters:, α, and T. We verified our trained model on 100 other images to check the accuracy; we found that, despite having a large difference in units among four properties, NN can identify all four properties with an accuracy higher than 90%.

**Remaining questions**

*Honestly assess the work that must still be completed. Think right on the page in this section, posing questions, speculating meaningfully, exploring your options. No more than 20 lines*

1) The code for the descriptor generation using the combinatorial approach and all the tests we performed is publicly accessible from the https://github.com/lstorchi/matinformatics website. We have shown only one, though paradigmatic, example. However, the same approach can be applied to any physical problem in which a compound’s property of interest can be described in terms of basic properties (i.e., atomic properties and not only) of the constituents. As such, the methodology and the related code that we have developed look very general and can pave the way to applications in many different fields. However, when we checked this code on another dataset (related to the stability of ferroelectricity vs antiferroelectricity in a class of ternary compounds), the error between predicted output and the actual output of selected formulas was relatively high. Here, we found that the basic properties we provided to build the descriptors are insufficient to describe the higher dimension features. Therefore, we are now expanding our atomic properties database, so that hopefully even complex properties can be successfully predicted.

2) Though we successfully separated spin waves between two frequencies, our main objective for this project is to build and fabricate a device that can work like a physical NN or NN filter. However, the time for micro-mechanical magnetic stimulation increases with complexity. The present method takes around a week of simulation time of simple optimization. Therefore, we need to optimize the ongoing method or find a more efficient method.

3) So far, we only showed one example to predict the four parameters from simulation images. The next step would be to verify the accuracy of the same model on experimental images. Furthermore, we would like to add the influence of the external magnetic field in the images and try to predict the different properties. This model will be a huge step and can be applied to study the domain wall motion in a magnetic material.

**Expected results**

*Shortly summarize the expected results. Commit to a schedule for obtaining those results if possible. No more than 15 lines*

1. So far, we have successfully analyzed the energetics between RS and ZB structures which is first-order transformation. Shortly, we plan to use our current approach to solve more complex second-order transformation problems, e.g., predicting stable phases among Ferro-electrics, Paraelectric, and Antiferro electric materials.
2. As we have mentioned in the previous section, first, we will try to optimize the GA to shorten the simulation time, and after that, we will model the device which can work like physical NN or physical NN filters. After successfully modeling the device, our collaborator will fabricate the device for further experimentation.
3. For this research, we are preparing a manuscript for publication to show our NN approach for predicting intrinsic properties, and after that, if time allows, we would like to include for prediction further higher dimensional properties like an external field.