**­**

**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 Multidisciplinary Material Science**

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

**PhD Student: Udaykumar Gajera**

**Supervisor: Prof. 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 Modelling with Quantum ESPRESSO” organized online by ICTP, 17 - 28 May 2021, CFU=10
* 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 ME 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 contribution 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
* **U. Gajera**, S.Picozzi, “High-throughput and Machine Learning for First Principal Simulation of Multi-ferroics”, December 8, 2020, CFU=3
* **U. Gajera**, S.Picozzi, “Use of machine learning methods on ferroelectrics” – TWEET (ToWards fErroElectricity in Two-dimensions) – Italy, November 27, 2020. CFU=3

**TOTAL CFU(B2) =**

**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 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**, pages, DOI. CFU = 4

**TOTAL CFU(B4) =4**

(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**

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 benefit from artificial intelligence dramatically. The report shows a brief introduction of implementing three different machine learning approaches for three different material scales for three different material science applications: 1). Construction of the descriptors to predict the energy difference between Rock-salt (RS) and Zinc-blend (ZB) structures using Ab-initio data. 2). Designing microscopic magnonic devices using Genetic Algorithm (GA), 3). Prediction of the 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 energy difference between of 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. 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), Kernel Ridge (KRR), and LASSO regression) for descriptor selection. The descriptor has been modeled as a combination of simple functions (power, exponential) made of atomic quantities (electronegativity, electronic energy levels, atomic shells radii, etc.).

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

There are two significant problems with traditional silicon device: 1). Energy consumption and 2) Heat dissipation. However, it is hard to replace these devices from average daily lives until we don’t find other accurate, logical device. Therefore, scientists are trying to make devices that can reduce the dependence on traditional silicon devices. For example, GPUs can take care of intense graphic computation, MRAM can store more data with lower switching time, spin valves can work in neuromorphic computing, etc.

In the present research, we propose a spin wave-based method to construct a magnonic device than can work as a physical Neural Network (NN). A spin wave can define as propagating disturbance in ordering a magnetic material and it can only propagate through the magnetic moments and therefore no motion of electrical charge, which theoretically require lower energy to regulate and dissipate no heat. We can imagine a magnonic devices that can perform a particular task of NN, like filters of the images that can reduce the computation load on the CPU. In order to build magnonic device, we start from the one example from literature, wang at el. where the authors successfully created a demultiplexer to separate two different spin-wave frequencies using a magnonic device, and from that, we try to build the more complex physical classifier using a Genetic algorithm and micromechanical magnetic simulation code, MuMax.

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

With advanced instruments, the characterization of materials is becoming more popular. These characterization techniques can now find intrinsic properties like crystal structure, metals' concentration, and magnetism. Advance techniques like Magnetic Force Microscopy can even detect magnetic properties at the nanoscale. However, we need different instruments to characterize different properties. Therefore, we propose a method where we combine magnetic images of material from a single spectrophotometer and predict other properties from the data or image created by that spectrophotometer. We showed an example of predicting intrinsic and extrinsic properties like exchange length (), damping parameter (α), anisotropy constant (), and temperature (T) using 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 searching for accurate descriptors and used LR to identify a final “best formula” on our test case of RS-ZB phases of 82 binary compounds for verification. 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 related to the p shell of the cation) play an essential role in deciding the stable crystalline phase. 3). We further studied 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).**

For designing the magnonic device, we have adopted the methodology of GA to find an optimized structure that can separate two frequencies in two output terminals from one input terminal. Furthermore, we used MuMax to simulate spin-wave propagation in YIG material. In search of optimized structure, we found that GA can optimize structure more efficiently than the method used by wang et al. Therefore, we started simulating with 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 bigger structure. Therefore, we are further trying to optimize the method to optimize structure even more quickly.

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

Advance machine learning techniques like NN have already been to everyone’s life part. Image classification in photos to face recognition on smartphones has become the new normal. Therefore, we also used these advanced NN methods to predict the intrinsic material properties of magnetic materials. 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 to predict four intrinsic values:, α, and T. We verified our trained model on 100 other images to check the accuracy where we found that, despite having a magnitude of difference among four properties, NN can identify all four properties with more than 90% accuracy.

**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 our in-house ferroelectric dataset, the error between predicted output and the actual output of selected formulas was relatively high. Here, we found that the basic properties we provide to build the descriptors are insufficient to describe the higher dimension features. Therefore, now we are expanding our atomic properties database, and we are hopeful that the new extended atomic dataset can also help predict higher-order properties.

2). Though we successfully separate spin waves between two frequencies, our main objective for this project is to build and fabricate a device that can work like physical NN or NN filters. 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 further or find a superior method.

3). Smart characterization techniques are an old concept. For our project, we would like to show that simple ML techniques can dramatically improve the usefulness of these techniques. Here, we only showed one example to predict the four intrinsic properties 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, especially for ion-beam radiation, 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