Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery

Machine Learning for Materials Science
6th - 10th May 2019
Aalto University
Aalto Design Factory, Aalto University, Finland

Dr James Cumby
University of Edinburgh

07/06/2019

Network: Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery

Contents

1   Event Details

Title	Machine Learning for Materials Science
Organisers	Aalto University
Dates	6th - 10th May 2019
Programme	Programme
No. Participants	90
Location	Aalto Design Factory, Aalto University, Finland
Organisation / Local Chairs	Dr. Milica Todorović, Prof. Patrick Rinke & Prof. Adam Foster - Aalto Unviersity Claudio Zeni & Aldo Glielmo - King's College London Dr. Kevin Rossi - Ecole Polytechnique Federale de Lausanne
Sponsors	CECAM, CDC, Aalto University School of Science, International Journal of Quantum Chemistry

2   Event Summary and Format

This five-day workshop consisted of a two-day school introducing machine learning techniques for materials science with a mixture of presentations and hands-on tutorials, followed by a three-day conference made up of invited presentations, contributed talks, a poster session and a number of panel discussions. The event drew participants from primarily chemistry, physics and computer science backgrounds, all with a focus on applying machine learning to materials discovery or understanding. Some of the invited talks were given by industrial participants, complementing the academic research presenting, and offering a different view. Coffee breaks and a conference dinner/poster session allowed plenty of time for networking amongst this diverse group.

3   Workshops

The first two days of the event were dedicated to a hands-on introduction to machine learning

methods and tools, with a particular focus on their application to materials science problems. This consisted of introductory lectures giving an overview of methods (such as ridge regression or neural networks) and their application to specific problems (such as learning interatomic potentials), followed by tutorial sessions where attendees were guided through a range of Jupyter notebooks introducing the main tools and techniques. Examples included generating atomic descriptors of materials (such as SOAP, smooth overlap of atomic positions) through to training a neural network to predict electron density of chemical functional groups based on their atomic environment. The tutorial sessions were structured such that there was time to explore more advanced problems in areas of interest, with guidance from local experts.
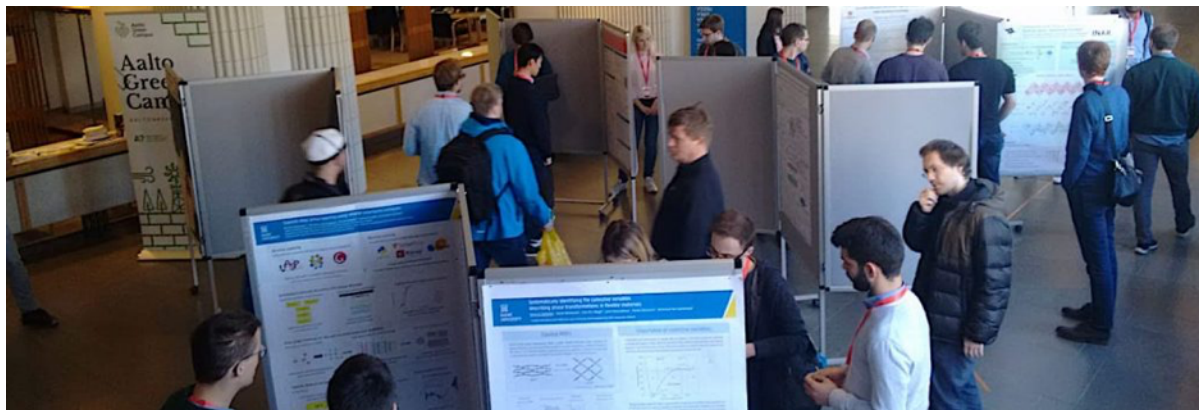
4 Talks

Talks were loosely clustered by a common theme, with individual sessions focussing on Bayesian methods, surface science, materials descriptors (including my own talk) and neural network approaches. A common theme throughout was the use of machine learning in generating accurate interatomic potentials, to achieve similar accuracy as ab initio methods but for much less computational cost. Dotted throughout were industry talks ranging from improving self-driving vehicle technology (T. Roman, NVIDIA) to predicting chemical reactions (P. Schwaller, IBM).

Invited talks covered a range of topics; on the first day, Antonietta Mira (Università della Svizzera italiana) talked about using approximate Bayesian methods to determine model parameters without knowledge of a likelihood function (and a new parallel implementation with ABCpy) and Teemu Roos (University of Helsinki) discussed using a combination of two neural networks to predict energy barriers and the associated error for vacancy migration. The invited speakers for day two began with Volker Roth (University of Basel) who talked about using encoding-decoding schemes to simplify the complexity of chemical space and Rafael Gomex-Bombarelli (MIT) who discussed coarse-graining methods for overcoming complexity in organic molecules.

In the afternoon sessions, we heard Karsten Wedel Jacobsen (Technical University of Denmark) talk about graph-based methods for representing materials for machine prediction of properties, and finally Thomas Hammerschmidt (University Bochum) discuss ways to accurately represent electronic structure on a coarse-grained basis. Invited speakers for the final day of the workshop consisted of Michele Allegra (Institut de Neurosciences de la Timone) who talked about methods for finding the intrinsic dimensionality of protein configurations, Simon Olsson (Freie Universitat Berlin) on machine-learned force fields for protein molecular dynamics and Olexandr Isayev (University of North Carolina) who spoke about using deep neural networks to predict the energies and properties of materials to the same accuracy as DFT or coupled-cluster methods.

A number of contributed talks featured throughout the workshop. Of particular interest to me were talks on JARVIS-ML for predicting materials properties (Kamal Choudhary, NIST) and the PANNA software to interface between density functional theory and neural networks in order to predict materials properties (Franco Pellegrini, Scuola Internazionale Superiore di Studi Avanzati). All talks were live-streamed online, and are available on YouTube through http://ml4ms2019.aalto.fi/workshop-materials/.

5   Posters & Demos

A poster session of almost 50 posters was held in conjunction with the conference dinner, and provided an opportunity for networking with participants from a wide range of backgrounds. Topics ranged from the realism of chemical oxidation states from Bader analysis of ab initio calculations through to the clustering of atomic aerosols, via coarse-grained models of protein dynamics and the growth of graphene on a liquid copper surface. The wide range of topics led to some lively discussions, and the transfer of new ideas from lots of different perspectives.

6   Expert Panel

The final two days of the event included discussion panels, consisting of invited speakers. In addition to in-person questions, an online forum also allowed questions to be asked remotely. Topics covered included the interpretability of machine learning algorithms, the reproducibility of results (including best methods to disseminate data and algorithms), the difficulties for researchers new to machine learning and the future of machine learning for materials science. The experts (from different disciplines) gave thought-provoking responses, and demonstrated that there are still a lot of unanswered questions in this area.

7   Conclusions

This event brought together researchers with a wide range of expertise from different disciplines, enabling a broad discussion of approaches and methods in machine learning as applied to materials science. The conference showed the current strength of research in using machine learning for generating models of atomic potentials, but more importantly demonstrated that

approaches can be applicable between disciplines. The preceding school was particularly beneficial for someone getting started with machine learning techniques in this area, and the overall format of the meeting allowed lots of time for interactions and helpful discussions.

8   Related Events

For those who are interested in getting involved with machine learning for materials science and attending related events there are some additional events that cover similar areas of interest.

3rd International Conference on Materials Sciences and Nanomaterials, The Lane, Oxford (22nd-24th July 2019)
2nd RSC-BMCS / RSC-CICAG Artificial Intelligence in Chemistry, Fitzwilliam College Cambridge (2nd-3rd September 2019)

Upcoming events of interest can be found on the AI3SD website events page.
http://www.ai3sd.org/events/ai3sd-events
http://www.ai3sd.org/events/events-of-interest