The Advanced Machine Learning for Innovative Drug Discovery (AIDD) project

Newsletter 4, June 2022

The **Advanced Machine Learning for Innovative Drug Discovery** (AIDD) project is a Marie-Skłodowska-Curie Innovative Training Network (ITN) for Early Stage Researchers (ESRs) funded by the European Commission under the Horizon 2020 Programme, Marie Skłodowska-Curie grant agreement No 956832. The project brings together fifteen academic and industry beneficiaries from ten European countries and the University of British Columbia (Canada) to train sixteen PhD students in close collaboration with partners from the USA, Australia, China, Israel, and other countries.

Project development

The **AIDD project started on 01.01.2022**.

Second AIDD School - Spring School on Advanced Machine Learning in Lugano

During the past months, the Project Management Team has been busy with preparing the Second Spring School of the AIDD project, which took place from May 9 to May 18 in Lugano, hosted by IDSIA USI-SUPSI. Since the IDSIA is a research institute for Artificial Intelligence (AI), the school focused on theory and application of methods of AI, with many high-profile experts both local and from abroad giving talks. Among speakers who accepted our invitation were scientists from academic institutions such as EPFL, MIT, University of Bern, University of Linz, as well as large industries such as Bayer, Firmenich, Dompé as well as SMEs, e.g. Chelonia and Ascenion. We were even able to secure a talk by Dr. Silvio Giancola from KAUST, the Saudi-Arabian Research University. The school was co-organized with the AIDD project check (which took place remotely) as well as with the meeting of the General Assembly and the Supervisory Board.

The school took place in the new "Campus East" building which is shared by the IDSIA as well as by several departments of the SUPSI (University of Applied Sciences of Southern Switzerland) and the USI (University of Southern Switzerland). The AIDD fellows thus had the chance to get to know the “university landscape” of Ticino, where several high-profile institutions successfully work together while still retaining their own characteristics. Practically all speakers and fellows, with few exceptions, managed to come to Lugano, and also most of the talks were given in person. Nonetheless we also offered Zoom streaming to the general public in order to best disseminate knowledge, and to encourage collaboration and networking across Europe and the world. Beyond some members of the local academic

population, we were happy to welcome five external participants in person, namely from the H2020 project VIRTUOUS (in which IDSIA is likewise a partner), as well as up to 30 remote participants.

Figure 1. The Campus Est, home of IDSIA USI-SUPSI.

The School started on the 9th of May with a lecture by Oleg Szehr: the first of three parts of an introductory course to sequential decision making and Markov decision processes. The remainder of the day was dedicated to chemical applications: Philippe Schwaller (EPFL Lausanne) presented approaches to use machine learning for synthesis planning, and Florian Häse from Bayer presented SELFIES, a novel method to textually describe organic molecules. Finally, Silvano Coletti and Carmine Talarico (Chelonia Applied Science) presented results on COVID-related drug discovery from Exscalate4CoV, the private-public consortium supported by the EU H2020 program.

On 10th of May, the day started with two very different talks: Jürgen Schmidhuber, PI at IDSIA, presented his groundbreaking work on Artificial Curiosity, and Leonardo Medrano, AIDD representative from University of Luxemburg, talked about Machine Learning based construction of Molecular Force Fields. In the afternoon, the AIDD consortium virtually hosted our EU project officer for the first part of the project check, during which we formally presented the progress of our work to the European Commission as our funding agency, thus demonstrating that we fulfil our research and training duties, and receiving valuable feedback. Almost all partner organisations had sent a representative to Lugano, and we ended the day with a formal business dinner at the Restaurant Canvetto Luganese, close to University.

We continued the official meetings of the AIDD consortium on the following day (May 11th). In the afternoon the AIDD School resumed both physically and virtually, with talks by Günter

Klambauer (University of Linz) about Few and zero-shot learning, which has recently received great interest in a variety of fields, from Mike Preuss (University of Leiden) about experimental computational work, and by Peter Ertl, the renowned Novartis Scientist, on navigating the space of bioactive rings.

On 12th of May, we saw another full day of talks, mostly dedicated to topics of foundational artificial intelligence. The first speaker of the day was Silvio Giancola from KAUST (Geometric Deep Learning), followed by Jean-Louis Reymond from University of Bern (Artificial Intelligence and the Chemical Space), Samuel Genheden from AstraZeneca (Synthetic Route Prediction), and finally by two IDSIA scientists, Alessandro Antonucci and Alessandro Facchini, giving a two-hour introduction on Explainable Artificial Intelligence, which is a very current topic not only due to its scientific implications, but also due to its high societal relevance. This day finished with an informal group dinner at the Biggio Brewery Restaurant, a well-known local landmark.

13th of May, the last day of an intense, but also very successful week, was again dedicated to applications of Machine Learning in Chemistry. The first talk was given by Vittorio Limongelli, a local USI professor of the Faculty of Biomedical Sciences, who presented both a theoretical talk and a practical demonstration entitled “Graph Neural Networks at the Service of Molecular Simulations”. The afternoon talks were given by external speakers: Michael Sattler (Helmholtz Zentrum München) lectured about Structure-based Drug Discovery, Guillaume Godin (Firmenich) presented their in-house AI formula generator, which got Swiss Digital innovation award 2021, and Clara Wong-Fannjiang (UC Berkeley), connected via Zoom, instructed us on Conformal prediction for the protein design problem.

Saturday 14th of May saw a well-deserved break. Michael Wand, local organiser of the AIDD school, took the fellows on a scenic hike from Morcote, one of the most beautiful villages of Switzerland, to San Salvatore, the iconic mountain towering over Lugano. The following Sunday was free of any official activity and students had a possibility to self organise visits to Lugano and its neighbourhood.

The AIDD school recommenced on 16th of May, with a day split between foundational machine learning and chemical applications. In the morning, Dario Azzimonti from IDSIA gave an introductory lecture on Gaussian processes, followed by AIDD colleague Adam Arany (KU Leuven) giving a talk on Bayesian inference. The late afternoon talks were held by Axel Pahl (MPI Dortmund) on Cell Painting Assay, and by Emilio Benfenati (Mario Negri Institute, Milan) on Toxicity Prediction methods.

On the last full day of the school, May 17th, we had a talk which was very different from the ones so far: Sigrid Scheek from Ascenion instructed us in both theory and practice on dealing with Intellectual Property rights, requirements, and pitfalls. This talk led us beyond the foundational instruction which was covered in the main part of the school, and towards advanced topics of managing proper research output, along the lines of the EID program. The theoretical talks of the day were given by Floriane Montanari and Marco Bertolini (both Bayer) on the topics of Graph Neural Networks and their explainability, as well by Oleg Szehr, who concluded his three-lecture series on Markov Decision processes.

Figure 2. The AIDD Fellows during the social hike on Saturday May 14.

On noon of May 18th, we finally adjourned the school and said goodbye to each other, after many exciting talks and fruitful discussions with our colleagues. The final talks of this day were intended to give a perspective beyond the AIDD project: IDSIA scientists Dario Piga and Gianvito Grasso presented the H2020 project VIRTUOUS, in which IDSIA likewise is a partner, and three further IDSIA researchers (Omar Chavez Garcia, Oscar Lithgow, and Michael Wand) gave pep talks about their research challenges far away from the field of chemistry, which again led to interesting discussions with the AIDD fellows and the other school participants.

The detailed list of all lecturers, as well as all lecture slides, can be found at: https://ai-dd.eu/lectures.

Voices about the school:

"The AIDD School at SUPSI was a great opportunity to hear about recent research in the cross-section between drug discovery and deep learning. We had a number of interesting speakers from the hosting institute as well as from the other partners and external sources. The speakers were diverse

and came from both the industrial and academic side of the field. Building on the previous school, we went into more detail on some core machine learning concepts and discussed how they can be utilised throughout the drug discovery process. As always, the event was a valuable networking opportunity with speakers and colleagues alike and the location of Lugano only added to the experience." - Emma Svensson, AIDD fellow

"The increasing speed of technological innovation and the rise of machine learning, in particular, have made the creation and fostering of innovation capabilities vital for the chemical industry.

The interdisciplinary AIDD PhD program in chemo-informatics bridges the gap between academia and industry and makes an ideal platform for the transition of know-how in both directions.

Through a dedicated series of lectures, students learn about the state-of-the-art of machine learning but also about important practical topics, such as the patent laws.

The depths in machine learning expertise and the focus on the chemical industry bring the AIDD school into a unique position in Europe making an ideal ground for the growth of ground-breaking know-how.

The school's organisation has given plenty of opportunity for building new contacts, innovative thoughts and stimulating discussions.

We are looking forward to many collaborations and projects originating from this memorable event! Great thanks to the organisers for the excellent planning!" - Oleg Szehr, IDSIA researcher

Next AIDD School in Leuven, October 17-25 will be organised as a hybrid meeting. A Zoom access of external participants to selected presentations is currently planned and will be announced on the web site of the project.

Meet AIDD partners at (see continuously updated list at AIDD web site):

- *Strasbourg Summer School in Cheminformatics, 27 June - 1 July, Strasbourg, France,* ***AIDD participant****: Johannes Kirchmair, Ola Engkvist*
- *WATOC 2020: 12th Triennial Congress of the World Association of Theoretical and Computational Chemists, 3-8 July, 2022, Vancouver, Canada,* ***AIDD participant****: Alexandre Tkatchenko*
- *Conference on Games (CoG), the Foundations of Digital Games (FDG) Conference and the Parallel Problem Solving from Nature (PPSN), 9-13 July, Boston, USA,* ***AIDD participant****: Mike Preuss*
- *Methods in Molecular Simulations and Machine Learning, 14-16 July, Barcelona, Spain,* ***AIDD Poster:*** *Non-Locality in Machine Learning Force Fields and Equivariant Interatomic Potentials, Gianni De Fabritiis, Julian Cremer,* ***AIDD participant****: Frank Noe*
- *Multiscale Modeling of Complex Systems: Methods and Applications, 17-22 July, Castelldefels, Spain,* ***AIDD participant****: Frank Noe, Gianni De Fabritiis*
- *2nd International Conference on Noncovalent Interactions, 18-22 July, 2022, Strasbourg, France.* ***AIDD participant****: Alexandre Tkatchenko*
- *Thirty-ninth International Conference on Machine Learning, 19-22 July, Baltimore Convention Center, USA,* ***AIDD posters****: Contrastive learning of image and structure-based representations in drug discovery, Ana Sanchez-Fernandez; Robust task-specific adaption of drug-target interaction models, Emma Svensson,* ***AIDD participant****: Jürgen Schmidhuber, Frank Noe, Adam Arany, Markus Heinonen*
- *The 28th Conference on Uncertainty in Artificial Intelligence, 1-5 August, Eindhoven, The Netherlands,* ***AIDD participant****: Markus Heinonen*
- *Psi-k conference 2022, 22-25 August, Lausanne, Switzerland.* ***AIDD lecture****: Many-body DFTB repulsive potentials for organic molecules from a physics-inspired neural network, Alexandre Tkatchenko, Leonardo Medrano*
- *70th International Congress and Annual Meeting of the Society for Medicinal Plant and Natural Product Research, 28-31 August, Thessaloniki, Greece,* ***AIDD participant****: Johannes Kirchmair*
- *1st Nordic Conference on Computational Chemistry 2022, 30-31 August 2022, Gothenburg, Sweden,* ***AIDD participant:*** *Yasmine Nahal*
- *19th International Conference on Density Functional Theory and Applications, 28 August - 2 September, Brussels, Belgium.* ***AIDD participant****: Alexandre Tkatchenko*
- *5th RSC-BMCS / RSC-CICAG Artificial Intelligence in Chemistry. 1-2 September, Cambridge, United Kingdom*
- *21st European Conference on Computational Biology (ECCB2022), 12-21 September, Sitges, Barcelona,* ***AIDD presentation:*** *Rosa Friesacher*
- *XVIth International Congress of Toxicology, 18-21 September, Maastricht, the Netherlands,* ***AIDD lecture:*** *Predictive Toxicology: Explainable AI, Igor Tetko*
- *Drug Discovery Africa (DDA2022), 19-23 September, Afe Babalola University in Ado Ekiti, Nigeria,* ***AIDD lecture:*** *Development of ADMETox and QSAR/QSPR models using open OCHEM http://ochem.eu platform, Igor Tetko*
- *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), 19-23 September, Grenoble, France*
- *23rd European Symposium on Quantitative Structure-Activity Relationship (EuroQSAR), 26-30 September , Heidelberg,* ***AIDD participant****: Johannes Kirchmair, Floriane Montanari*
- *Lab of the Future Europe, 4-5 October, The Beurs van Berlage, Amsterdam, Netherlands,* ***AIDD participant****: Djork-Arné Clevert*
- *The 21st International Conference on Systems Biology (ICSB 2022), 8-12 October, Berlin, Germany,* ***AIDD participant****: Markus Heinonen*
- *RDKit UGM, 12-14 October, Berlin, Germany,* ***AIDD participant****: Djork-Arné Clevert, Floriane Montanari, Paul Czodrowski*

- *Thirty-sixth Conference on Neural Information Processing Systems (NEUrips), 28 November - 9 December, New Orleans,* ***AIDD lecture:*** *Critical assessment of molecular machine learning, (to be confirmed),* ***AIDD participant****: Jürgen Schmidhuber, Günter Klambauer, Markus Heinonen, Ana Sanchez-Fernandez, Emma Svensson*
- *Critical assessment of molecular machine learning workshop, 2 or 3 December, New Orleans - within NEUrips (submitted),* ***AIDD participant****: Günter Klambauer, Ana Sanchez-Fernandez, Emma Svensson*

Project publications with participation of AIDD fellows:

- **Sanchez-Fernandez, A.**; Rumetshofer, E.; Hochreiter, S.; Klambauer, G. Contrastive learning of image- and structure-based representations in drug discovery, ICLR2022 Machine Learning for Drug Discovery, 2022.

Additional information

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions, grant agreement No 956832.

Follow us at Twitter: https://twitter.com/AiddOne