αiM

AI & Mathematics

The importance of mathematics for Dutch AI research

www.αiMath.nl

AI and Mathematics (AIM)
The importance of mathematics for Dutch AI research

Over the past 10 years, AI has changed our lives, and it will change our lives much more in the years to come - AI is well regarded as the most important technology for the coming decades. AI is pre-eminently a multidisciplinary technology that connects scientists from a wide variety of research fields, from behavioral science and ethics to mathematics and com-puter science. Without wanting to diminish the importance of that variety, with this pamphlet we want to highlight the contribution of mathematics to Dutch AI research. With a clear vision of where mathematics can contribute to AI, we want to facilitate mathematicians to engage in the conversation with research partners from other fields of science.

Mathematical AI in the Netherlands is very successful. Several themes in which mathematics plays a major role are mentioned in recent AI research agendas and investigated in large research programs (responsibility, transparancy, generalizability, explainability). We fully endorse the importance of these themes and want to contribute from mathematics to consor-tia that address these research challenges. In addition to research in multidisciplinary con-sortia, fundamental mathematical research remains important in AI, as a basis for new deve-lopments in the longer term. In the early days of AI, much emphasis was on theoretical com-puter science and logic; in the meantime, many more types of mathematics are important: including mathematical statistics, game theory, graph theory, and dynamical systems. Often the greatest challenges and opportunities lie at the intersection of different fields, and these intersections are often unexplored territory in mathematics as well. Excellent mathematical research can and should play a major role in the Dutch AI research of the future.

We illustrate this by means of three important roles of mathematics in AI: as a basis for the design of AI methods, as a basis for analysis and understanding of AI methods, and directly in the application of AI

Mathematics and the design of AI methods

Today, successful AI often involves deep learning, learning from large amounts of data using huge neural networks, with suc-cessful applications in speech and image recognition. Deep learning has emerged from brilliant intuitions and grown through intensive trial-and-error engineering and the use of huge amounts of data; mathema-tics played a modest role until recently. This perhaps underplayed the fact that mathe-matical concepts and methods are the basis of just about all other successful methods and innovations in AI: support vector ma-chines, kernel methods boosting (state-of-the-art methods for machine learning when there is little data), Bayesian learning, causal learning and reasoning, graphical models (crucial for explainable AI): all equally successful and mathematical in nature. Also very recent develop-ments in the important field of machine learning and privacy lean heavily on mathematics, where the concept of differential privacy was introduced as early as 2011.

Geometric Deep Learning

This field focuses on bringing geometric structure to learning systems to improve them or, conversely, to enable new data structures. For example, groups and representation theory can be used to exploit symmetries in data and problems that promote generalization and reduce the search space for usefulfunctions (deep networks).

Bekkers, E. J. (2019). B-Spline CNNs on Lie groups. International Conference on Learning Representations.

Indispensable mathematical topics at the base of all these methods are probability, statistics, optimization and approximation theory, and logic. But very different forms of mathematics also play a role: for example, we see mathematics in the form of differential geometry and PDEs recurring in the design of new, advanced machine learning methods. Directed graphs are indispensable in research on reasoning, argumentation and uncertainty. The study of logic, game theory, and dynamical systems is useful for the design of knowledge, social procedures, and interaction.

Mathematics and the analysis of AI methods

Major applications of deep learning include speech recognition (Siri, Alexa, Google Assistant), machine translation and image recognition - with AI in medical image recognition often outperforming human experts. But there is something strange going on: key AI methods produce "black box" algorithms that cannot be explained in domain-relevant concepts. They are based on extremely large neural networks that are found through a labor-intensive process: all kinds of networks and parameter values are tried out, and in applications where very large amounts of data are available one eventually gets something that works very well. But to the obvious and important question "why?", the system has no answer.

Bayesian networks for explainability

Bayesian networks (BNs) use directed graphs to describe probability distributions compactly and intuitively. For example, the NWO Forensic Science research program investigated how and to what extent BNs can be designed and explained for the analysis of homicide cases. An important question for AI is how to use BNs to learn understandable causal structures from data. This is a field in which essential contributions are made from many different disciplines, and in which the underlying mathematics is the common denominator that creates synergy and helps to avoid speech confusion.

Mooij, J.M., Magliacane, S., Claassen, T. (2020). Joint Causal Inference from Multiple Contexts. Journal of Machine Learning Research 21(99):1-108

There are paradoxical examples where one changes some pixels in a picture of, say, a dog just a little bit (a human does not see the difference), but the computer suddenly thinks it is a cat - and we do not understand why the computer changes its mind. Without a fundamental understanding of these issues, applicability ultimately remains limited - in applications where, for example, little data is available, where an estimate of reliability is essential ('uncertainty quantification') or where data is strongly 'biased', truly successful AI applications are therefore often left waiting for the time being. Mathematics provides the pre-eminent means for studying and understanding how AI methods work. In particular, with methods from analysis, statistics, and logic, it is possible to gain insight into fundamental capabilities and limitations, theoretical performance guarantees, optimality, and uncertainty quantification, explainability, accountability, and social behavior. It is precisely through the interaction of different forms of mathematics and their application that progress can be made.

For several years now, mathematical research on "why does deep learning work so well?" has begun to pick up steam. There is also increasing insight into the fundamental limitations, for example for explainable algorithms and responsible behavior of AI systems. With its strong tradition in mathematical statistics and logic, the Netherlands can play a major role in this.

Another important question is why AI systems sometimes use such huge amounts of energy compared to the human brain. For example, a fundamental theoretical question involved is whether a continuous stream of data requires alternatives to digital computing (e.g., neuromorphic computing).

Mathematics and the application of AI methods

In applied mathematics, there is a long tra-dition of transferring mathematical methods to concrete applications. AI methods are increasingly used in this context, in a multi-tude of applications. This often involves combining ideas and methods from multiple fields. One example is causal reasoning, often based on Bayesian networks. The theory development is grounded in mathe-matical AI for correctly modeling reasoning with uncertainty and later proved applicable in many other fields. For example, causal reasoning is indispensable for efficient com-putational models for genetic data and is intensively used in research on new medical treatments.

Why does Deep Learning work?

Using techniques from statistics and approximation theory, it is possible to better understand the situations in which Deep Learning works and the optimal architecture of deep networks. For example, it turns out that, despite the huge number of parameters, the 'effective complexity' of neural networks is often limited, which explains why there is not too much 'overfitting'.

Schmidt-Hieber (2020). Nonparametric regression using deep neural networks with ReLU activation function. Annals of Statistics, Volume 48, Number 4, 1875-1897, 2020 (with discussion).

AIM for the best

In the coming years there will be a lot of investment in the development and application of AI in the Netherlands. Mathematical AI research can play a major role in many applications, and is essential for laying a solid foundation for the future of AI.

The AIM network

The AIM network unites Dutch research groups doing research on AI in which mathematics plays a major role. A list of all AIM rrespresentatives can be found at www.AIMath.nl

www.aiMath.nl