Imagine/project how A.I. will improve people's lives by 2035. Propose the role of Natural Scientists in making this happen.

The major problems faced by natural scientists today are complex, thorny ones, with solutions often too obscure for the capability of conventional computational methods. These complexities are rooted in the combinatorial explosion of solutions. Imagine the solution to finding what the perfect game of chess looks like: there are numerous possible moves at each stage of the game, and each of these moves creates a new arrangement of pieces that then offers a different set of next possible moves. Each step of the game exponentially increases the size of the problem; there are too many possible combinations to analyze exhaustively by humans or conventional machines before the solution can be found.

In science, comparably complex problems are everywhere: how do we predict a protein's 3D structure given the sequence of amino acids that it is made up of?; how do we know how susceptible a patient is to a certain genetic disease?; how do we analyze the impact of climate change on a particular species? While these questions can in theory be answered by exhaustively assessing every possible outcome, AI offers a practical and efficient method to tackling the combinatorial explosion of these problems.

Primary implications – how AI benefits natural science

An unintelligent machine depends entirely on the commands and rules inputted by users to recognize patterns in data. In other words, it can only search for trends that we program it to look for. The fundamental advantage AI has is its ability to process large amounts of data and discover subtle patterns, and more importantly, *its constantly evolving and adapting memory that allows it to learn*. AI systems are able to derive their own rules from the data it analyses, it can identify patterns that we had no prior knowledge of and consequently did not think to look for. AI therefore has progressive performance and value as the program becomes more experienced and as the amount of data it has processed increases.

This is particularly helpful in tasks such as DNA sequencing. We may feed a machine large amounts of DNA sequences derived from healthy and sick patients in the hopes of understanding a disease caused by more than one possible gene mutation. An unintelligent machine can only identify the mutation patterns and their locations where we program it to, and only do so with DNA sequences it is trained with. An AI program, such as an artificial neural network, can learn what possible harmful mutations are present in a sequence, and then *apply this knowledge to previously unseen DNA sequences* to determine the susceptibility of the patient from which the DNA was derived is to the disease in question. We can further feed the AI with data of the mutated gene for it to learn how to predict the impact of the mutated gene on the individual.

This method can be applied to any procedure requiring recognition, for instance, image analysis of graphical data such as spectroscopy. For such graphs, one of the things the program may be trained to identify is characteristic absorption peaks of a particular molecule.

Furthermore, artificial neural networks can incorporate "fuzzy logic", meaning it can provide answers to questions that have a level of subjectivity. *How* susceptible is the patient to the disease? The unintelligent machine cannot analyze such a judgement, whereas the AI can be fed subjective descriptions by assigning numerically what "very" or "unlikely" may look like.

Similarly, this method can be used for the identification of molecules, or to predict the interaction between two molecules of known structures. But, what if the problem isn't finding a solution, but determining *which solution out of many possible correct ones* is most likely?

The structure of a specific protein, given the sequence of amino acid that comprise it, is notoriously impossible to predict; although many conformations are possible in theory, only one specific conformation is adopted, and it is this *precise* 3D structure that governs its biological function. This impossibility is due to combinatorial explosion of: the angle of rotation between every adjacent amino acid, the potential interaction between one amino acid *and every other amino acid in the sequence* (one protein can consist of a chain of thousands of amino acids). An AI program, such as a genetic algorithm, can act as an optimization technique that attempts to discover this precise structure through evolutionary principles. The program operates in a cyclic manner by selecting random conformations, assesses their fitness, then favorable solutions are reproduced and merged whilst unfavorable solutions are manipulated. The cycles repeat, driving the solution towards an optimum conformation, in a manner similar to natural selection. The machine can be fed data of known protein structures and sequences for it to "learn" which solutions are more or less favored. An unintelligent machine would approach the issue of protein structure in a much less efficient manner, by exhaustively testing every possible combination without any logical direction.

The ways in which AI can be used in natural sciences is innumerable and far extends the examples listed above. The specific methods, however, are similar to the ones described, but adapted for its specific use. Other examples of usage may include selecting the cheapest, most efficient conditions and route of synthesizing a desired chemical.

Secondary implications – what can we do with these improvements in natural science?

When viewed individually, the impact of these applications on human life may not be immediately clear. It is instrumental to bear in mind that the natural sciences do not exist in isolation. The advancement of one area of natural science can certainly help the advancement in a different area, thus amplifying the positive impact of AI. It seems helpful to give specific examples of how natural science applications of AI may be merged and utilized in areas that are extensions rather than direct branches of natural science, such as environmental science and pharmacology.

Accuracy and depth in the analysis of DNA genome can be used by microbiologists to understand the biochemical processes of particular bacteria species. AI will allow this kind of DNA screening to occur with time- and cost-efficiency, which is certainly useful to our increasingly urgent battle against antibiotic resistance, and eventually decrease the number of deaths by infectious disease. Once the organism is sufficiently understood, we may then use prediction of protein structure and interaction to develop necessary drugs. A large impediment to current drug development is the lengthy, costly, and highly uncertain development process, and the difficulty in identifying drug targets. Using AI to develop possible proteins and to predict interactions with potential targets reduces the number of trial-and-error experiments greatly. This may also be used by pharmacologists to develop more personalized medicine for patients with specific genetic disease, thereby improving the quality of healthcare through tailored treatment.

Identification of specific chemicals by artificial neural networks can aid environmentalists in recognizing toxins and pollutants, either by directly recognizing the molecular structure or their assays. Predicted interactions can then be used to determine how certain pollutants affect our environment, visualize how they might disperse, and methods to prevent this dispersion can then be developed. This may aid the provision of clean drinking water, for instance.

There are, of course, countless possible benefits beyond these examples, as a greater understanding of the genome and of protein structure underlies much of life science advancements, and the overall logical methodology used by AI can be used to approach any computationally intractable problem in any of the natural sciences.

Tertiary implications – the role of these developments on wider society

To fully consider how AI, and its natural science applications, will improve our lives in 15 years, we should also briefly consider its impact on a wider scale, beyond science.

The quality of AI output is inextricably linked to the quality of the data we feed it. Presently, information and data flow has and will unmistakably continue to progress exponentially, and this will no doubt lead to enhanced AI performance. In the years approaching 2035, AI will advance inwards (more sophisticated AI methods, more data available, more talent) and outwards (increased application in natural science and other sectors) in a complementary fashion. The advancement in one direction promotes advancement in the other: for example, with improving AI techniques in structure recognition, more protein structures are made known to improve the quality of unknown protein structure prediction.

In order for these AI applications to have deep and lasting impact, however, a structural reorganization of society is perhaps also necessary. Funding and investment into developing AI techniques as well as the natural science methods they are used for are essential; a transparent information system that will encourage trust in AI methods and higher data availability; and tighter restrictions on the problems that AI has been used to tackle so that we may progress from developing solutions to existing problems (antibiotic resistance) to creating preventative measures before the problems arise (stricter use of antibiotics). Together, these will allow AI and natural sciences to achieve maximum positive effects on our future.