The Impact of Artificial Intelligence on Medicinal Applications

Karan Chawla

Abstract: ***Explainability is a long - standing issue in artificial intelligence, because traditional AI methods were easily understood and reproducible. Their inability to handle the uncertainties of the actual world was a drawback, though. Applications grew more and more successful when probabilistic learning was introduced, but they also became more and more opaque. The introduction of traceability and transparency in statistical black - box machine learning techniques, especially deep learning (DL), is the focus of explainable AI. We contend that explainable AI is not sufficient. Causability is necessary to bring medicine to a level of understandability. Causability includes metrics for the quality of explanations, just as usability includes measurements for the quality of usage.***

Keywords: explainability, artificial intelligence, probabilistic learning, transparency, causability

1. Introduction

Perhaps the oldest branch of computer science, artificial intelligence (AI) covers a wide range of topics, including creating systems that can learn and think like people and simulating cognitive processes for the purpose of solving problems in the real world. As a result, to distinguish it from human intelligence, it is frequently referred to as machine intelligence. The field was centered on the nexus between computer science and cognitive science. Nowadays, AI is a very popular topic because of machine learning (ML) practical achievements.

With the goal of creating software that can automatically learn from past data to gain knowledge from experience and continuously improve its learning behavior to make predictions based on fresh data, machine learning (ML) is a very useful area of artificial intelligence. Making sense of the world, comprehending context, and making decisions in the face of ambiguity are the three great problems. As the backbone of artificial intelligence, machine learning (ML) is becoming widely used in business, engineering, and science. This is resulting in a greater emphasis on making evidence - based decisions. The creation of novel statistical learning algorithms, along with the accessibility of massive data sets and inexpensive computing, has propelled the tremendous advancement in machine learning.

A family of machine learning models with a lengthy history, DL is based on deep convolutional neural networks. Because they are producing incredible outcomes even at human performance levels, DL is highly popular these days. A recent study by the Thrun group serves as an example of optimal practice; using a deep learning methodology, they were able to perform on par with physicians, showing that these methods can accurately classify skin cancer at a level of proficiency that is comparable to that of human dermatologists.

An further instance would be the encouraging outcomes in detecting diabetic retinopathy and associated ocular conditions. All of these are excellent illustrations of the advancement and utility of artificial intelligence (AI), but even the most well - known proponents of these (automatic) methods have recently stressed how challenging it is to achieve usable intelligence because in addition to learning from previous data, extracting knowledge, generalizing, and overcoming the curse of dimensionality, we also need to separate the underlying explanatory factors of the data in order to comprehend the context in an application domain, where a doctor - in - the - loop is still necessary.

Explainability and Causability

One of the biggest applications of AI, ML, and DL is medicine. We deal with probabilistic, unknown, incomplete, imbalanced, heterogeneous, noisy, unclean, erroneous, inaccurate, and missing data sets in arbitrarily high - dimensional domains when we assist medical decisions. Frequently, we just lack vast data sets. Future medical research aims to model patient complexity in order to customize treatments, medical decisions, and healthcare procedures for each patient. This presents difficulties, especially when it comes to the mapping, integration, and fusion of diverse distributed and heterogeneous data, as well as the visual analysis of this heterogeneous data. Explainable AI in the medical domain must therefore consider that a variety of data sources may contribute to an important outcome.

Explainability is not so much a product of AI as it is a challenge that has existed for at least as long. The early AI employed symbolic and logical thinking techniques. These methods worked well, but only in a very narrow range of domains and with very little real - world application. MYCIN, an expert system designed in Lisp to recognize germs causing serious diseases and suggest antibiotics, is a common example. Mycin was never employed in a therapeutic setting, possibly as a result of its stand - alone nature and the significant work required to keep up its knowledge base. But these early artificial intelligence systems were able to produce a record of their reasoning processes by using some type of logical inference on symbols that were accessible by humans.

This served as the foundation for the explanation, and some early related work is available, for instance. It is important to note that there are three different kinds of explanations here: A scientific explanation in the strict meaning of science theory; (2) an educational explanation as it is carried out between teachers and students; and (3) a peer - to - peer explanation as it is carried out among physicians during medical reporting. We stress that we are referring to the first kind of explanation in this text.

The need for AI techniques in the medical field that not only work effectively but also are transparent, dependable, understandable, and comprehensible to human experts—for instance, natural language sentences in the field—is expanding. To replicate and understand the learning and knowledge extraction process, as well as to reenact the machine decision - making process, methods and models are required. This is significant because comprehending the causal relationship between learned representations is essential for decision support.

Moreover, medical practitioners' confidence in AI systems in the future may be strengthened by explainability in AI. In order to develop explainable AI systems for use in medicine, research must continue to achieve excellent learning performance for a variety of machine learning and human - computer interaction methodologies. Explainability and ML performance (predictive accuracy) are inherently at odds. The most effective techniques, like deep learning (DL), are frequently the least visible, and the ones that give a clear explanation, like decision trees, are less precise.

In the context of explainable AI, understanding, interpreting, and explaining are frequently used interchangeably, and a variety of interpretation methodologies have been utilized in the past. When discussing explainable artificial intelligence, the term "understanding" typically refers to a functional understanding of the model rather than a low - level algorithmic understanding, which aims to characterize the model's "black - box" behavior without delving into its internal mechanisms or representations.

We contend that explainable AI is desperately needed in medicine for a variety of applications, such as clinical decision - making, research, and education. Even in situations where advanced artificial intelligence (AI) systems supplement medical professionals and, in the future, even significantly influence the decision - making process, human experts must possess the ability to comprehend and reverse - engineer the machine's decision - making process as needed.

In addition, it's fascinating to realize that, contrary to popular belief, people are not always able to justify their choices! Because there are so many diverse and large - scale sources of knowledge, experts are occasionally unable to offer an answer. Therefore, explainable - AI places a new and significant emphasis on usability and human - AI interaction while also requiring confidence, safety, security, privacy, ethics, fairness, and trust. Collectively, these elements are essential to application in medicine in general and to customized medicine in particular in the future.

Before moving on to an example and a medical use - case from histopathology, we first provide some definitions to clarify what kind of explainability we mean. This will lead us to use the term "Causability" instead of the more common "Causality. " After that, we briefly discuss the state - of - the - art of some current explainable models. In conclusion, we highlight the critical need for a systems causability scale that incorporates social dimensions of human communication in order to assess the quality of an explanation.

Explainable AI Models

Glass - box methods can be used to understand ad hoc systems; examples of these systems often include fuzzy inference systems, decision trees, and linear regression. These have a long history, can be created using expert knowledge or data, and offer a solid foundation for the interaction of human expert knowledge with hidden information found in data, all from the perspective of human - AI interaction.

Another example was given, in which high - performance generalized additive models with pairwise interactions (GAMs) were applied to medical domain problems, producing understandable models. This revealed unexpected patterns in the data that had previously prevented complex learned models from being used in this field; what's important is that they showed that these methods could be scaled to large data sets with thousands of attributes and hundreds of thousands of patients, all while maintaining intelligibility and accuracy comparable to the best (unintelligible) machine learning techniques.

It has been shown that deep neural networks (DNN), in particular convolutional neural networks (CNN) and recurrent neural networks (RNN), are useful for a variety of real - world issues, including movement identification and picture recognition and classification. Furthermore, because these methods take into account human processes, they are noteworthy from a scientific perspective as well. Human thought processes, for example, are hierarchical, and new research has shown similarities between learnt models in CNNs and the human visual ventral pathway. People have attempted to describe artificial neural networks ever since the field's inception. The use of gradients in the form of sensitivity analysis was one of the initial methods.

An artificial neural network (ANN) is a group of neurons arranged in a series of layers. The neurons in an ANN take in the activations of the neurons in the layer before them as input and use that information to carry out a basic calculation. A sophisticated nonlinear mapping from the input to the output is collectively implemented by the network's neurons. Backpropagation, which continuously modifies the weights of the connections in the network to minimize the difference between the current output vector and the desired output vector, is how this mapping is learned from the data by adjusting the weights of each individual neuron.

Internal hidden units that are not included in the input or output indicate significant task domain features as a result of the weight modifications, and the interactions between these units capture task regularities.

There are numerous methods for examining and deciphering deep neural networks. A measure of uncertainty can be found in the so - called model uncertainty or epistemic uncertainty, which is the effect of small perturbations of training data on model parameters, or in the predictive uncertainty or aleatoric variability, which is the effect of changes in input parameters on the prediction for a single example. Through variational approaches, the model parameters are approximated in a Bayesian Deep Learning manner, yielding model weight uncertainty information and a way to extract predictive uncertainty from the model outputs. When there is ambiguity, it is easier to apply model predictions appropriately in situations where several sources of information are merged, as is frequently the case in medicine. Additionally, we can distinguish between heteroscedatic and homoscedatic uncertainty, which is independent of a specific input.

Attribution techniques aim to establish a connection between certain deep neural network outputs and input variables. This, in a way, links the uncertainty of the prediction to the elements of a multivariate input. Activation maps are used by researchers to pinpoint image segments that are important for a network prediction. Approaches to attribution for generative models have been introduced recently.

2. Conclusion

In the medical field, supervised learning is highly costly due to the difficulty in obtaining comprehensive ground - truth labels and robust supervision information. In particular, classifying a histological image is a crucial and time - consuming task for the diagnosis of cancer, since it is clinically significant to divide the cancerous tissues into several groups. The very large image size (and the associated problems for DL), the inadequately labeled images (the sparse training data available), the pathologist's time required (expensive labeling), the insufficient labels (region of interest), the various magnification levels (resulting in varying levels of information), color variation, and artifacts (sliced and placed on glass slides) are some of the general issues to be taken into consideration when analyzing digital pathological images.

We suggest classifying entire slide images in the context of weakly supervised learning using commonly - used scoring systems based on correlation with histomorphological features and an overall predictive score. Additionally, we provide a relevance map created by watching the human expert during diagnosis making. The human causal model can be expanded using established human features in conjunction with novel multiscale morphological classifiers. Additionally, the CNN model can be clarified using established histomorphological features. Our proposal involves extracting from single cell nuclei that are malignant or benign, and then classifying the chromatin arrangement within the nuclei to establish a correlation with molecular markers and histopathological traits.

References

[1] Acid, S. and De Campos. L. M., Searching for bayesian network structures in the space of restricted acyclic partially directed graphs, Journal of Artificial Intelligence Research, Vol.18, pp.445 - 490, 2003.
[2] Alai, M., AI, scientific discovery and realism, Minds and Machines, Vol.14, No.1, pp.21 - 42, 2004.
[3] Al - Ani, A. and Deriche, M., A new technique for combining multiple classifiers using the dempstershafer theory of evidence, Journal of Artificial Intelligence Research, Vol.17, pp.333 - 361, 2002.
[4] Ambite, J. L. and Knoblock, C. A., Planning by rewriting, Journal of Artificial Intelligence Research, Vol.15, pp.207 - 261, 2001
[5] Andrew, A. M., Short paper: lifemanship and AI. Kybernetes: The International Journal of Systems & Cybernetics, Vol.30, No.2., 2001.
[6] Argamon - Engelson, S. and Dagan, I., Committee - based sample selection for probabilistic classifiers, Journal of Artificial Intelligence Research, Vol.11, pp.335 - 360, 1999.
[7] Atkinson K. and Bench - Capon T., Practical reasoning as presumptive argumentation using action based alternating transition systems, Artificial Intelligence, Vol.171, No.10 - 15, pp.855 - 874, 2007.
[8] Baader, F., Lutz, C., Sturm, H. and Wolter, F., Fusions of description logics and abstract description systems, Journal of Artificial Intelligence Research, Vol.16, pp.1 - 58, 2002
[9] Bacchus, F., The power of modeling - a response to PDDL2.1., Journal of Artificial Intelligence Research, Vol.20, pp.125 - 132, 2003.
[10] Baget, J. F. and Mugnier, M. L., Extensions of simple conceptual graphs: the complexity of rules and constraints, Journal of Artificial Intelligence Research, Vol.16, pp.425 - 465, 2002.
[11] Balazinski, M., Czogala, E., Jemielniak, K. and Leslie J., Tool condition monitoring using artificial intelligence methods, Engineering Applications of Artificial Intelligence, Vol.15, Issue 1, pp.73 - 80, 2002.
[12] Barber, F., Reasoning on Interval and Point - based Disjunctive Metric Constraints in temporal contexts, Journal of Artificial Intelligence Research, Vol.12, pp.35 - 86, 2000.
[13] Barzilay, R., Elhadad, N. and McKeown K. R., Inferring strategies for sentence ordering in multidocument news summarization, Journal of Artificial Intelligence Research, Vol.17, pp. s 35 - 55, 2002.
[14] Baxter, J., A model of inductive bias learning, Journal of Artificial Intelligence Research, Vol.12, pp.149 - 198, 2000.
[15] Baxter, J. and Bartlett, P. L., Infinite - horizon policy - gradient estimation, Journal of Artificial Intelligence Research, Vol.15, pp.319 - 350, 2001
[16] Baxter, J., Bartlett, P. L. and Weaver, L., Experiments with infinite - horizon, policy - gradient Estimation, Journal of Artificial Intelligence Research, Vol.15, pp.351 - 381, 2001.
[17] Becker, A., Bar - Yehuda, R. and Geiger, D., Randomised algorithms for the loop cutset problem, Journal of Artificial Intelligence Research, Vol.12, pp.219 - 234, 2000.
[18] Bhattacharyya, C. and Keerthi, S. S., Mean field methods for a special class of belief networks, Journal of Artificial Intelligence Research, Vol.15, pp.91 - 114, 2001
[19] Bhuyanb, Tea quality prediction using a tin oxide - based electronic nose: an artificial intelligence approach, Sensors and Actuator B: Chemical, Vol.94, Issue 2, pp.228 - 237, 2003
[20] Birnbaum, E. and Lozinskii, E. L., The good old Davis - Putnam procedure helps counting models, Journal of Artificial Intelligence Research, Vol.10, pp.457 - 477, 1999.

[21] Blockeel, H., Dehaspe, L., Demoen, B., Janssens, G., Ramon, J. and Vandecasteele, H., Improving the efficiency of inductive logic programming through the use of query packs, Journal of Artificial Intelligence Research, Vol.16, pp.135 - 166, 2002
[22] Bod, R., A unified model of structural organisation in language and music, Journal of Artificial Intelligence Research, Vol.17, pp.289 - 308, 2002.
[23] Boddy, M. S., Imperfect match: PDDL 2.1 and real applications, Journal of Artificial Intelligence Research, Vol.20, pp.133 - 137, 2003
[24] Boutilier, C. and Brafman, R. I., Partial - order planning with concurrent interacting actions, Journal of Artificial Intelligence Research, Vol.14, pp.105 - 136, 2001.
[25] Boutilier, C., Dean, T., Hanks, S., Decision - theoretic planning: Structural assumptions and computational leverage, Journal of Artificial Intelligence Research, Vol.11, pp.1 - 94, 1999.