THE COMPUTERTECH
(An International Peer Review Journal)
VOLUME 4; ISSUE 2 (JULY-DEC); (2018)
WEBSITE: THE COMPUTERTECH

The Use of AI and Big Data in Health Care

Divya Sai Jaladi[1*], Sandeep Vutla[2]
[1]Senior Lead Application Developer, SCDMV, 10311 Wilson Boulevard, Blythewood, SC 29016, UNITED STATES
[2]Assistant Vice President, Senior-Data Engineer, Chubb, 202 Halls Mill Rd, Whitehouse Station, NJ 08889, UNITED STATES
*Corresponding Author: divyasaij26@gmail.com

Abstract

A lot of people across the world are talking about how AI and automation will affect the future of work, society, and the economy. We lay out the groundwork for artificial intelligence (AI) and big data (Big Data) and how they relate to public health in this article. We outline the possible consequences and difficulties that medical experts and diagnosticians may face, while also highlighting the concerns that are involved. In light of new findings, this article outlines the potential advantages of sophisticated data analytics and machine learning. The article identifies and discusses problems related to ethical concerns and the future of experts and professionals in the AI era.

Keywords: AI; Big Data; Healthcare; Medical; Public Health

Introduction

Big Data refers to the enormous computing resources that are required to manage the growing complexity and quantity of data from many sources, including the internet and distant sensor networks. There may be intricate syntactic, semantic, social, cultural, economic, and organisational links among the structured, semi-structured, and unstructured data that makes up Big Data. Data corruption due to noise and artefacts, data entry mistakes, duplicate records, missing records, and incomplete digital records pertaining to information on date, age, gender, and other variables are all examples of epistemic uncertainties that can enter the data processing pipeline and impede decision making by both humans and computers.

Industry 4.0, which includes cyber-physical systems, cloud computing, and the IoT, are all part of the Big Data culture. There is usually a high degree of process automation in these big data processing systems. The advent of artificial intelligence (AI) in diagnosis and decision making, spurred by advancements in computer technology, has coincided with the present worldwide emphasis on Big Data on the internet. Social media's meteoric rise and the price of microelectronics used in linked devices' precipitous decline were early triggers.

Using the field of public health as an example, this article presents and describes the basics of artificial intelligence and big data. Recent research published in several esteemed publications, including JAMA and Nature Reviews, provides examples of prospective advantages and critical remarks within the context of [1]. Public health issues including epidemiology, precision medicine, medical screening, visual enhancement, and psychology were represented by the examples given, which are meant to be illustrative rather than complete. Ethical concerns and what this means for the future of professionals and their opinions are up for debate. We provide some solutions, assess them, and then propose that public health experts should begin talking about how they might work with AI.

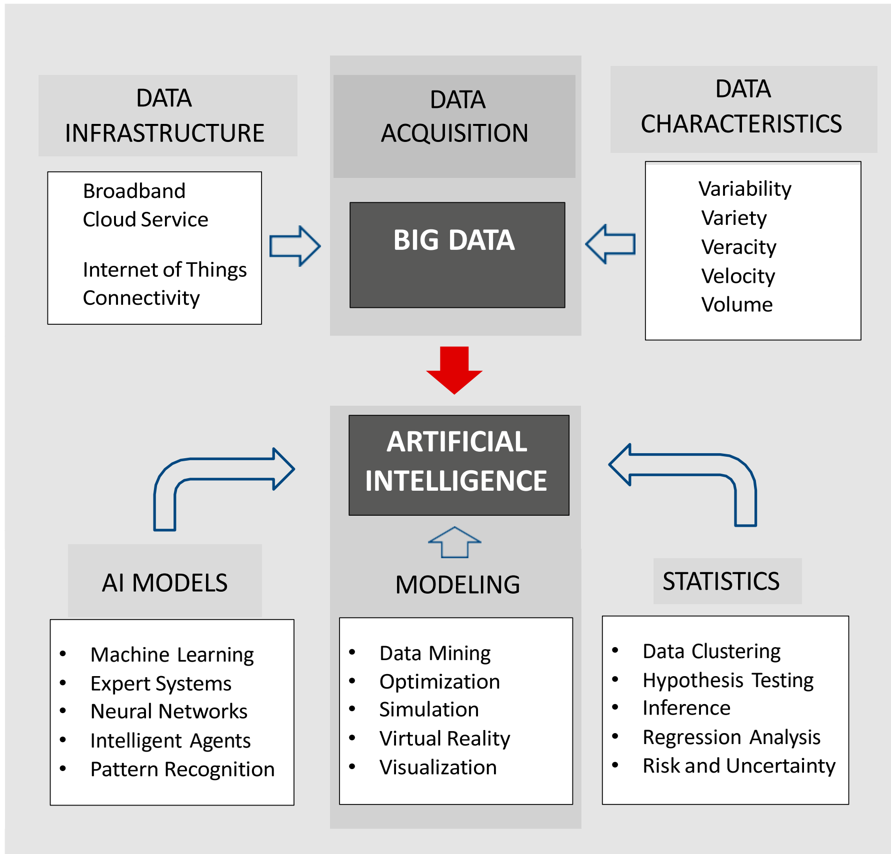

Figure 1. Artificial intelligence (AI) and Big Data.

Big Data

Big Data is commonly linked to massive macrosystems that process data in a distributed fashion, which is frequently more complex and time-consuming than what local desktop computers and conventional software can handle. Streaming media assets such as photos, videos, and audio files are all possible outcomes of information processing. Some previous sources have mentioned five features of Big Data, which can exist alone or in various combinations [1-3]):

• Veracity (data accuracy, noise, and uncertainty); • Volume (very huge data sets); • Variety (inclusive of audio files, images, numerical data, and text data); • Speed (real-time processing and extremely fast transmission); and • Variability (absence of structure, consistency, and context).

Epistemic uncertainties and statistical replications in large-scale scientific investigations impact the amount and quality of Big Data. When Big Data and AI are utilised together, they are expected to have the greatest societal impact.

An innovative study published by Deiner showcased the immense potential of Big Data in the field of epidemiology. The study showed that by monitoring online enquiries on illness symptoms on social media platforms like Twitter and Google Search, early detection of epidemics may be achieved. To identify, categorise, and detect illness patterns associated with conjunctivitis incidence, data analytics and pattern recognition were utilised. According to the study's findings, monitoring internet searches might help spot biosecurity concerns and influenza epidemics early on. Electronic medical records, pharmaceutical sales, and hospital admissions are examples of concrete data that may be used to validate the apparent benefits of early intervention and timely identification.

Minority groups' restricted internet access and engagement by age, gender, and ethnicity impact the veracity of social media statistics. Mistakes in Twitter enquiries might lead to inaccurate disease incidence forecasts due to ambiguities in Boolean logic employing If-Then-Else rules, just like in Google search mistakes.

Problems with internet-based methods, such as questionable data validity, slow update rates, and inaccurate search engines, may necessitate more investigation in the future [4]. "Fake news" and "fake" Twitter accounts have been exposed as a means by which certain governments and data analytics firms have attempted to sway public opinion. This has been a "hot topic" in the mainstream media. Applying screening methods based on AI to online Big Data might solve this problem. One application case is the correlation of social media data with news articles, candidates' names, media outlets, and locations based on patterns identified by a neural network or other machine learning algorithm trained to identify specific recurrent words, phrases, and opinions expressed on social media. This method can uncover clusters and connections that may not be entirely benign.

In precision medicine, another use of Big Data is to replace the traditional population-based method with clinical trials that choose patients based on DNA profile, which yields biomarkers for focused therapy [2]. Clinical studies including people with the same genetic defect improve the accuracy of both medication development and therapy. Without the requirement for a placebo, further benefits include smaller sample sizes and less unpredictability in clinical studies. Genome sequencing may potentially lead to improved diagnostic sensitivity and targeted therapy [4]. To bolster enquiries into causation, larger-scale population studies may be conducted and Big Data analytics applied to obtain more precise genotypic and phenotypic data.

Whether genetic profiling is more beneficial than other factors, including changes in lifestyle and exposure to environmental stresses, is one of several questions in the precision medicine approach [14]. Factors such as food, physical activity, medicine, alcohol, tobacco, and the gut flora all influence how well a patient responds to pharmacological therapy [5]. It is necessary to furnish details on these elements inside in all-encompassing Big Data structure. To find answers to many of these questions, large-scale population studies are necessary.

Only a privileged few may profit from the creation of pricey genetically-tailored medications. Medical research that helps the most people would be a better use of scarce funds. In many instances, environmental variables take precedence over a disease's genetic foundation, and dietary or exercise-related behaviour change might be equally effective—and cheaper—than the alternative. Furthermore, it is possible that a combination of already-existing, affordable medications would be more beneficial than a costly new treatment created from a clinical study. With research budgets being cut, it's more important than ever to examine the costs and advantages of various fields of study and use data analytics to make informed decisions.

Concerns about consumer use or abuse, genetic data ownership, privacy, and control over its distribution are major obstacles to the creation of ethics for precision medicine. When it comes to staff selection or credit ratings, for instance, consumers' perceptions of risk might be impacted by unauthorised access to genetic information. An individual's DNA cannot be altered like a password. Discussions on genetic privacy should be more robust in order to address issues like possible enhancements to regulatory oversight, the appropriate balance between public rights to privacy and evidence needed for law enforcement, and other related topics [6].

Artificial intelligence (AI)

Simply said, artificial intelligence refers to processes or machines that can learn from their surroundings and adjust their behaviour accordingly to achieve a predetermined goal, such as improving efficiency or accuracy in a given task [7]. Applications of logic, statistics, mathematics, and computer programming enable the learning process to be put into practice. Iteratively training the AI model on data with parameter modification through trial and error utilising reinforcement rules is made possible by the learning process. It is possible that the performance index is the reduction of the discrepancy between the model's predictions and the experimental results.

The classical scientific method, which began with Aristotle in antiquity and was codified by Francis Bacon about 1600, forms the basis of the training phase. The AI model may be taught and then used to make judgements using new data. Identifying, evaluating, and categorising are all steps in the decision-making process. Because it resembles an adaptive system that learns with the help of a feedback loop and predetermined objectives, this method is called supervised learning. Correctly naming labelled pictures of various skin malignancies is one example, as is making a medical diagnosis based on pathology data. Due to the enormous resources required for predictive analytics, researchers have proposed that AI might play a role in processing Big Data in precision medicine [8]. Looking at Big Data as an AI database really brings forth its potential (Figure 1).

The goal of unsupervised learning in artificial intelligence models like the k-means clustering algorithm is to find significant groups or defining characteristics in the data [9]. Exploiting Big Data to create high-level abstractions that can mimic a human subjective reaction is a big problem for future AI research. It is possible to create a digital expert that can automatically translate pathology findings into a report or an explanation for use in clinical settings.

Machine learning, expert systems, pattern recognition, and fuzzy logic are some examples of AI techniques. The statistical foundations and learning principles of many AI models are built on nonlinear optimisation processes. Predefined architectures with weight matrices that are iteratively modified according to the specified performance index are common components of artificial neural networks (ANNs) [6].

Data Mining

While some AI techniques, such expert systems, necessitate the explicit programming of new rules, machine learning relies on training on new data using an adaptive methodology, like a neural network [7]. Iteratively presenting the data and adjusting the model's coefficients (weights) leads to an optimised performance index. A "deep learning" approach for detecting diabetic retinopathy and diabetic macular oedema (DME) was validated in a recent study by Gulshan et al. [6]. In comparison to human professionals, the study's results showed higher sensitivity (87%) and specificity (98%). In contrast to conventional pattern recognition, the deep learning method did not rely on linear or edge-based explicit feature measurements. The model was validated using test photos from a database after being trained on a massive number of tagged images; discrimination was based on utilising the full image as a pattern, without explicit feature extraction. To reduce the impact of picture variability, very big training sets can be used instead of canonical transformations to normalise for different geometrical aberrations.

A convolutional neural network with several hidden layers that were optimised using a multistart gradient descent process made up the deep learning algorithm [1]. Through the use of graphics cards and parallel processing with hundreds of photos, extremely fast processing rates were attained during the learning phase. The remarkable categorisation findings have been mirrored in comparable research that has been examined in other places, prompting concerns over the potential future responsibilities of human experts in medical diagnosis [10].

Because of the importance of the findings, JAMA published a perspective piece, two invited editorials, and a commentary on the machine learning technique as a digital diagnostician [7-9]. Based on a majority decision from medical professionals, future study is still needed to refine the labelling method used for training data and assess performance in clinical situations. Use of more objective measurements for identification of DME, including optical coherence tomography, might increase the performance index.

According to Jha and Topol [9], medical specialists might be able to adapt to the challenge of machine learning by becoming information professionals and generalists. Artificial intelligence (AI) automated picture categorisation might be used for screening and detection; nevertheless, humans would be responsible for managing and integrating the data in a clinical setting, suggesting further testing, and guiding physicians. This possible alliance, however, presupposes that artificial intelligence is limited to recognising patterns and classifying images using machine learning techniques.

Systems capable of logic, inference, strategy, and data integration are also possible in various AI models. Software like "Deep Blue" and "Watson" from IBM and "DeepMind" from Google are examples of this; they've used AI algorithms to beat human experts at chess and GO and even win the general knowledge quiz show Jeopardy by consulting Wikipedia's full database. With the rise of software personal assistants like "Siri" from Apple and "Cortana" from Microsoft, which offer conversational response and guidance, it seems like specialised AI algorithms are getting smarter by the day. The next generation of personal assistants will include voice-activated platforms for interactive phone calls that mimic people. Reports have surfaced recently that internet titans Amazon (with "Alexa") and Google (with "Assistant") are seeking commercial partners to further develop this technology.

Babies learn to crawl, walk, and communicate, among other specialised abilities, by a combination of experience and trial and error. Along these lines, generalised AI may be possible in the future thanks to collaboration amongst specialised AI algorithms (such as machine learning) that mimic human cognition. A master AI program having a standard set of rules and axioms (like encoded DNA) may, for instance, exert top-down control over a suite of specialised algorithms. A number of businesses have proven that AI and ALP can substitute human experts by analysing diagnostic data, writing a final report, and providing verbal counselling (e.g., Arria NLG, Retresco, etc.).

According to the well-known Turing test, the only real difference between artificial intelligence and human intelligence is a question of degree. Given the present state of software and hardware advancements, feedback loops including multispectral IoT sensors, and complexity theory, AI has the potential to eventually catch up to, or perhaps surpass, human judgement. Artificial intelligence (AI) has been making exponential, rather than linear, strides in recent years, permeating every aspect of our lives.

A lot of people need to talk about and debate what human experts, diagnosticians, and doctors will do in the future. For artificial intelligence (AI) to find a home in healthcare, rules governing user safety, algorithmic openness, and the preservation of human rights must be established. If at all feasible, algorithms should provide explanations rather than only predictions, and they should not contain any bias in their code or data. Managing the rate of automation would guarantee gradual implementation, if needed, with ongoing process assessment [8].

Augmenting One's Vision

The OrCam is a portable artificial vision system that uses a microprocessor to recognise optical characters using a small television camera that is put on the frame of spectacles [1]. An program can transform any pattern, including faces, into audible words that can be heard through an earphone. This is a wearable gadget that uses artificial intelligence [2]. Positive findings were obtained in a recent study evaluating the OrCam for visually impaired individuals [11-15].

A 10-item test was used to evaluate the device, and it consisted of reading a menu, a faraway sign, and a newspaper headline. Scores were determined by the number of test items that were vocalised properly. It would be much easier to compare with other devices like NuEyes if the elements were weighted according to importance, which was not the case before [3]. As a starting point, you may utilise a standard eye chart.

It is impossible to overstate the promise of this noninvasive AI gadget for replacing lost senses. It is possible to digitally represent even colour and texture as words over the earpiece. An outdoor navigator who is blind might get situational awareness and bionic eyesight with the help of GPS data and bionic vision [14]. Adding location-specific lettering on city pavements, such "road," "crossing," or "bus stop," might enhance the OrCam device's performance. To fit in with the aesthetics of the surrounding area, text information might be shown with low contrast. Putting words on the walls and floors of buildings would make it easier to navigate unfamiliar places while simultaneously giving constant commentary, information, and environmental awareness.

Mining Data

Passive data sets kept in archives don't do much use for data science unless they can be analysed for analytics, forecasting, and decision making. Finding interesting characteristics, trends, patterns, correlations, or anomalies in a database is the goal of data mining, a knowledge discovery process [6]. Decisions, projections, and quantitative research can all benefit from the outcomes. According to some, data mining is a field that combines aspects of statistics, computer science, and database administration. It is also becoming more and more common to use artificial intelligence (AI) methods, such as machine learning, and visualisation, which makes use of complex visuals, particularly in Big Data analytics [12].

The use of neural networks and multivariate regression analysis are two popular methods for building prediction models. Figure 1 provides a list of other ways, and literature studies [7] detail many more strategies. Scanning electronic medical records for risk assessment is one use of machine learning, as is training a neural network to identify new patterns in patient data and associations between drugs and negative side effects [8]. To eliminate common epistemic uncertainties like erroneous, redundant, incomplete, or corrupt records, it may be required to "clean" the stored records using suitable data cleaning tools before data mining [9]. Data management solutions built for Big Data frameworks are among the many software tools available to assist with data quality and auditing; for more information, check out Hadoop, Cloudera, MongDB, and Taklend [13].

Data visualisation approaches, which rely on human perception to unearth information in complicated datasets devoid of apparent patterns, have the potential to improve data mining. As a springboard for in-depth numerical research, visualisation allows for the discovery and investigation of patterns, clusters, and artefacts [12]. Some examples of visualisation techniques are maps, bar graphs, stereo images, video data, polygons of varying sizes, colours, and forms, and two- and three-dimensional colour graphics. When it comes to bioinformatics, one example is the Microbial Genome Viewer. This web-based application allows users to combine genomic data, such as chromosomal wheels and linear genome maps, from a MySQL database in an interactive and visually appealing way [3].

Here is a small-scale example from the study of psychology that illustrates basic ideas in data visualisation and mining [4]. A pilot research examined the relationship between first-year university students' subjective assessments of anxiety and the perceived and actual danger they faced. The participants ranged in age from 16 to 43 years old and were of mixed gender ($n = 31$, $F = 24$, $M = 7$). The participants were asked to rate the look and perceived danger of 29 insect and animal pictures on a three-point scale: 1 for not at all, 2 for somewhat, and 3 for extremely. When the original study's data was fitted to a typical multiple linear regression (MLR) model using fear as the dependent variable, it was found that the weighting for perception of danger was 0.85 vs. 0.15 for appearance of threat. This study used the results of training a regular feed-forward backpropagation ANN on the statistics using commercial software, with the understanding that the small sample size would yield acceptable suggestive findings for illustrative purposes [5]. On training data, both MLR and ANN generated correlation coefficients $r > 0.90$, indicating that this method has predictive analytics potential. In order to boost confidence, future research should conduct larger-scale studies and make better use of Big Data for statistical power.

Figure 2 shows the result of employing commercial software to create a response surface based on a hybrid 2D polynomial function z = f(x,y) using regression analysis, which was used to simulate the original tabular data for data visualisation purposes [6]. The tiny sample size prevented the use of ANN's alternative function approximation, which is conceivable. A least-squares error regression study with gradually displayed model predictions along each axis would have been required for this. Two hidden areas of zero-fear response under certain combinations of appearance and perception of threat were revealed by a 90-degree clockwise rotation and colour coding, which allowed for data visualisation in this case (e.g., animals that were perceived as dangerous but were actually dead or disabled).

Therefore, data visualisation using colour and 3D geometry can increase the likelihood of discovering valuable information in a database or spreadsheet. There are greater chances for data mining with AI automation and DV for early trend discovery because to Big Data's enormous data volumes. Inherent in human vision are percepts and gestalts, which may be used as supplementary tools to improve data mining processes.

It all started with presenting and explaining the outcomes of data analytics in public health and the medical sciences in connection to the usage of AI and Big Data. The use of artificial intelligence and big data may cause modifications to screening techniques in regular pathology, general practice, and specialist diagnostics [7]. There has to be a greater understanding of these new technologies among lawmakers, bureaucrats, and business executives in order to construct policies and public health. The critical ethical issues and the need for a thorough regulatory framework surrounding precision medicine and artificial intelligence have not been sufficiently addressed by public health officials. These are the kinds of worries that will surface again and again whenever the topic of autonomous cars or military robots comes up.

Conclusion

The merging of artificial intelligence and big data has the potential to profoundly impact our future. The role of doctors and other medical experts will be threatened by the proliferation and integration of these technologies. No one has yet gotten too worked up about how diagnosticians and specialists will be responsible in the future due to the possibility of more sensitive and specialised performance from specialised AI systems. Will artificial intelligence eventually supplant humans in medical diagnostics? Does case management represent the natural progression for human professionals, and will screening, detection, and diagnosis be replaced by AI in the future? An approach that reclassifies human experts as information specialists and generalists might hold the key to solving these problems, according to some. The diagnostics industry may benefit from the use of AI and automation, particularly in the interpretation of pathology results and the classification of images. After that, individuals may freely manage the data in a clinical context, determine the necessary tests, and provide continuous assistance to both physicians and patients. This approach is effective for short- and medium-term use of speciality AI, but will it be sufficient for the long-term when generic AI becomes the norm?

References

[1] Pasham, S.D. (2017) AI-Driven Cloud Cost Optimization for Small and Medium Enterprises (SMEs). The Computertech. 1-24.

[2] Pasham, S.D. (2018) Dynamic Resource Provisioning in Cloud Environments Using Predictive Analytics. The Computertech. 1-28.

[3] Sai, K.M.V., M. Ramineni, M.V. Chowdary, and L. Deepthi. *Data Hiding Scheme in Quad Channel Images using Square Block Algorithm*. in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. 2018. IEEE..

[4] Alam, K., Mostakim, M. A., & Khan, M. S. I. (2017). Design and Optimization of MicroSolar Grid for Off-Grid Rural Communities. Distributed Learning and Broad Applications in Scientific Research, 3.

[5] Agarwal, A. V., & Kumar, S. (2017, November). Unsupervised data responsive based monitoring of fields. In 2017 International Conference on Inventive Computing and Informatics (ICICI) (pp. 184-188). IEEE.

[6] Agarwal, A. V., Verma, N., Saha, S., & Kumar, S. (2018). Dynamic Detection and Prevention of Denial of Service and Peer Attacks with IPAddress Processing. Recent Findings in Intelligent Computing Techniques: Proceedings of the 5th ICACNI 2017, Volume 1, 707, 139.

[7] Mishra, M. (2017). Reliability-based Life Cycle Management of Corroding Pipelines via Optimization under Uncertainty (Doctoral dissertation).

[8] Agarwal, A. V., Verma, N., & Kumar, S. (2018). Intelligent Decision Making Real-Time Automated System for Toll Payments. In Proceedings of International Conference on Recent Advancement on Computer and Communication: ICRAC 2017 (pp. 223-232). Springer Singapore.

[9] Agarwal, A. V., & Kumar, S. (2017, October). Intelligent multi-level mechanism of secure data handling of vehicular information for post-accident protocols. In 2017 2nd International Conference on Communication and Electronics Systems (ICCES) (pp. 902-906). IEEE.

[10] Malhotra, I., Gopinath, S., Janga, K. C., Greenberg, S., Sharma, S. K., & Tarkovsky, R. (2014). Unpredictable nature of tolvaptan in treatment of hypervolemic hyponatremia: case review on role of vaptans. Case reports in endocrinology, 2014(1), 807054.

[11] Shakibaie-M, B. (2013). Comparison of the effectiveness of two different bone substitute materials for socket preservation after tooth extraction: a controlled clinical study. International Journal of Periodontics & Restorative Dentistry, 33(2).

[12] Gopinath, S., Janga, K. C., Greenberg, S., & Sharma, S. K. (2013). Tolvaptan in the treatment of acute hyponatremia associated with acute kidney injury. Case reports in nephrology, 2013(1), 801575.

[13] Shilpa, Lalitha, Prakash, A., & Rao, S. (2009). BFHI in a tertiary care hospital: Does being Baby friendly affect lactation success?. The Indian Journal of Pediatrics, 76, 655-657.

[14] Singh, V. K., Mishra, A., Gupta, K. K., Misra, R., & Patel, M. L. (2015). Reduction of microalbuminuria in type-2 diabetes mellitus with angiotensin-converting enzyme inhibitor alone and with cilnidipine. Indian Journal of Nephrology, 25(6), 334-339.

[15] Gopinath, S., Giambarberi, L., Patil, S., & Chamberlain, R. S. (2016). Characteristics and survival of patients with eccrine carcinoma: a cohort study. Journal of the American Academy of Dermatology, 75(1), 215-217.