

Uses and opportunities for machine learning in hypertension research



Dharmika Amaratunga^a, Javier Cabrera^b, Davit Sargsyan^a, John B. Kostis^{a,*}, Stavros Zinonos^a, William J. Kostis^a

^a Cardiovascular Institute, Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ 08901, USA

^b Department of Statistics, Rutgers University, Piscataway, NJ 08854, USA

ARTICLE INFO

Keywords:

Machine learning
Deep neural networks
Hypertension
Disease management
Personalized disease network

ABSTRACT

Background: Artificial intelligence (AI) promises to provide useful information to clinicians specializing in hypertension. Already, there are some significant AI applications on large validated data sets.

Methods and results: This review presents the use of AI to predict clinical outcomes in big data i.e. data with high volume, variety, veracity, velocity and value. Four examples are included in this review. In the first example, deep learning and support vector machine (SVM) predicted the occurrence of cardiovascular events with 56%–57% accuracy. In the second example, in a data base of 378,256 patients, a neural network algorithm predicted the occurrence of cardiovascular events during 10 year follow up with sensitivity (68%) and specificity (71%). In the third example, a machine learning algorithm classified 1,504,437 patients on the presence or absence of hypertension with 51% sensitivity, 99% specificity and area under the curve 87%. In example four, wearable biosensors and portable devices were used in assessing a person's risk of developing hypertension using photoplethysmography to separate persons who were at risk of developing hypertension with sensitivity higher than 80% and positive predictive value higher than 90%. The results of the above studies were adjusted for demographics and the traditional risk factors for atherosclerotic disease.

Conclusion: These examples describe the use of artificial intelligence methods in the field of hypertension.

1. Introduction

The growing availability of large volumes of biomedical data, some derived from novel sources, is starting to help clinical researchers gain valuable insights into medical conditions. For researchers in hypertension, these insights will offer opportunities to assess hypertension prevalence and risk, to diagnose and gauge the severity of hypertension, and to estimate the risks of subsequent complications, thus offering the opportunity for timely treatment [1–4]. Overall, these developments offer great opportunities to significantly improve clinical management and patient care.

As an example, many studies have been conducted using the

Myocardial Infarction Data Acquisition System (MIDAS). MIDAS is a statewide database containing de-identified records of all cardiovascular disease hospitalizations in New Jersey since 1986. It contains over 17 million records corresponding to over 4 million patients with cardiovascular disease. Mining this database has led to important publications that has changed the way cardiovascular medicine is practiced as follows. Wellings et al. reported that the rate of admission for heart failure (HF) after discharge for a first myocardial infarction as well as all-cause death decreased markedly from 2000 to 2015 [5]. Using a Personalized Disease Network (PDN), Cabrera et al. described the development of complications in these patients and observed the prevalence of hypertension prior to the onset of cardiovascular diseases such as HF [6]. This is consistent

Abbreviations: MIDAS, Myocardial Infarction Data Acquisition System; PDN, Personalized Disease Network; ICD, International Classification of Diseases; CS/E, Computer Sciences/Engineering; CART, Classification and Regression Trees; SVM, Support Vector Machine; CNN, Convolution Neural Net; AMI, Acute Myocardial Infarction; HF, Heart Failure; NPV, Negative Predictive Value; PPV, Positive Predictive Value; EHR, Electronic Health Record; PPG, photoplethysmography; SBP, Systolic Blood Pressure; DBP, Diastolic Blood Pressure.

* Corresponding author. Cardiovascular Institute Rutgers Robert Wood Johnson Medical School 125 Paterson Street, CAB-4180A New Brunswick, NJ. 08901, USA.

E-mail address: kostis@rwjms.rutgers.edu (J.B. Kostis).

<https://doi.org/10.1016/j.ijchy.2020.100027>

Received 7 February 2020; Received in revised form 9 March 2020; Accepted 12 March 2020

Available online 19 March 2020

2590-0862/© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

with the known fact that hypertension is a major risk factor for cardiovascular disease.

Many types of data are becoming available for biomedical research including data from (a) clinical trials, (b) databases such as MIDAS that have large numbers of observations, (c) diagnostic tests such as electrocardiograms and imaging, (d) wearable biosensors which have the potential to generate real time data, (e) genomics, transcriptomics and proteomics, which may have considerably more variables than observations. Also, there is the potential to integrate data from multiple possibly very different sources to gain broader insights.

Nevertheless, there are many challenges to generating truly meaningful and generalizable insights from these large amounts of diverse and complex data. These include issues of data quality and data volume; or, to use a term common in big data analysis, the “4Vs”: Volume, Variety, Velocity, and Veracity. Overcoming these issues is not trivial.

Many steps are involved when dealing with big data. The first step should always be an appraisal of whether the data are appropriate for the research objective, followed by an examination of the data via exploration and visualization techniques to assess heterogeneities in the data, imbalances of key variables, outlier detection, and other issues that could affect the validity and generalizability of any subsequent findings. For example, the information in MIDAS for vital status, age, gender and race was checked and found to be over 98.8% accurate [7]. This is a major consideration as unstructured electronic health records (EHR) and data drawn from multiple sources for other purposes could have quality issues as well as inconsistencies in variable definitions (e.g., changes in International Classification of Diseases (ICD) codes). In addition, prior to analysis, some structuring of the data may be necessary taking into consideration concepts such as statistical experimental design, matching, propensity scoring, calibration, normalization, and data projection. This first step is essential.

The data can then be analyzed. One approach is to apply familiar statistical models, such as those used in multiple linear regression, logistic regression and survival analysis. These methods can generate predictions together with a performance measure and an explanation of which factors most contribute to the prediction. However, the volume and complexity of the data may be such that alternative approaches designed for analysis of big data and artificial intelligence could be considered. This includes methods such as artificial neural networks, support vector machines (SVM) and deep learning. These methods employ complex algorithms which include ensembles of models interacting with each other and enable the study of complex associations which cannot easily be reduced to an equation. Since the algorithms are set to seek solutions over a broad domain without being provided explicit instructions of where to search, they are often able to provide better predictions than simpler schemes.

In the following sections, we will briefly describe the machine learning approach and give some examples in which it has been applied for hypertension research.

2. Methods: machine learning and deep learning

Machine learning is devoted to the methodology and algorithms for predictive modeling with the objective of making the best possible linear or non-linear predictions. In classical statistics, the emphasis was on statistical inference and predictive modeling and was just another element of the statistical analysis. Fig. 1 shows the intersections between machine learning, statistics and computer sciences/engineering (CS/E) in the context of data science. Statistics and biostatistics intersect with CS/E in the areas of machine learning and big data and only in part in bioinformatics. However, there are major parts of CS/E that are not part of data science.

Machine learning was initiated in the late 1970s by Breiman et al. who proposed the methodology of Classification and Regression Trees (CART) and later proposed the methods of Bagging and Random Forests [8–10]. Contributions from Computer Science came later in the 1990s,

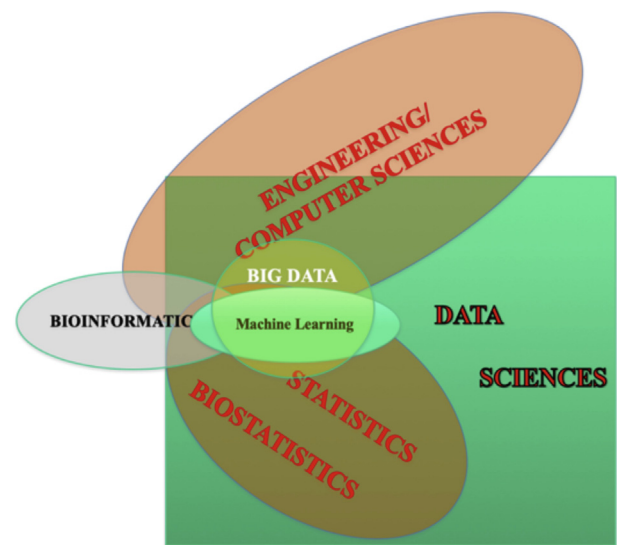


Fig. 1. Machine learning and statistics in the context of data science. Machine learning and statistics are at the core of data science. Other disciplines including engineering/computer sciences, data science, statistics/biostatistics, big data and bioinformatics intersect with machine learning in clinical medicine.

with Neural Networks, Boosting and Support Vector Machine (SVM) [11, 12]. These methods can be used to study the relationship between an outcome variable and several features that could potentially affect the outcome. For example, they could be used to predict the likelihood of a person with certain clinical features developing hypertension. Depending on the methods, it may be possible to also identify the main features that affect the outcome.

The main considerations in the analysis are balancing bias, variance and model complexity. Resampling or the use of a set-aside test set can be used to prevent overfitting which could be a major concern when fitting complex models such as these; overfitting occurs when a model is found to fit only the training data very well, but does not generalize to other data. Deep learning is a technique that was introduced recently as an updated version of neural networks. Deep learning is expected to become one of the most successful methods in terms of performance in predicting

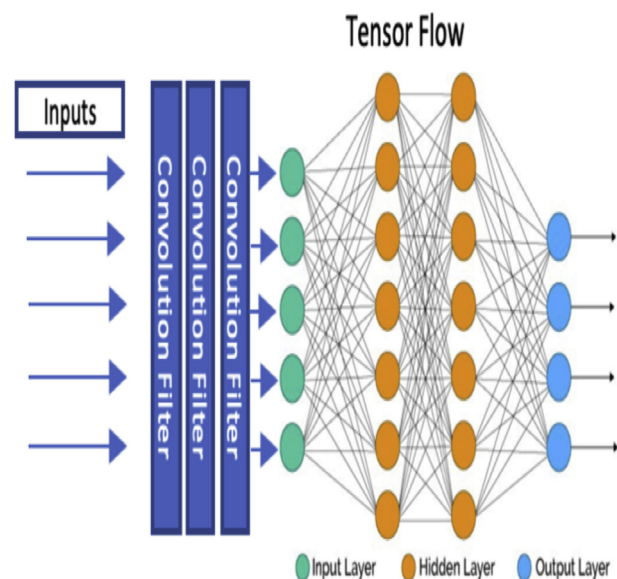


Fig. 2. Conceptual schema of the workflow for convolution neural networks (CNN). CNNs process the data by layers of convolution filters that are followed by a second set of layers of a fully CNN also known as a multilayer perceptron.

outcomes.

One of the most successful algorithms for deep learning is the Convolution Neural Net (CNN). The idea of a CNN is to apply a sequence of linear convolution filters to the set of predictors and to use the outcomes as inputs to Tensor Flow, the main neural net algorithm with several hidden layers (Fig. 2). To improve performance, the tensor flow algorithm will use a training set, a validation set, and a testing set and it will be configured with three hidden layers and it will be trained using thousands of optimizations called "epochs". The outcomes of tensor flow will be class predictions if doing classification or predicted values if the response is continuous. The importance of each predictor can be assessed by comparing the performance of the model by including and excluding the predictor. These techniques could be used to identify subsets of patients with particular constellations of risk factors that lead to a higher risk of events.

Both support vector machines and deep learning methods can also be helpful in identifying risk factors leading to cardiovascular death or serious morbidity. The main idea is to use the concept of variable importance combined with penalized selection which produces an ordered list of important predictors [13].

Machine learning also includes methods for unsupervised learning. Thus, cluster analysis can be used to group or segment datasets into clusters of observations which are similar to one another based on a set of attributes. Cluster analysis algorithms include hierarchical clustering and neural networks.

3. Examples

Example 1: Mining MIDAS: Each record in the MIDAS database mentioned in the introduction contains diagnosis and procedure information coded according to ICD-9th revision standards, as well as admission, discharge and procedure dates, and patient demographics. Normally, the ICD-9 codes are mapped to clinical outcomes (e.g. heart failure, myocardial infarction and stroke), and comorbidities (hypertension, atherosclerosis, other organ system diseases, etc.) using a hierarchically structured list of ICD-9 codes in conjunction with the knowledge of clinicians. The expertise of the latter is necessary due to the specificity of a medical condition being defined, e.g. ICD-9 code 410 specifies acute myocardial infarction (AMI) but it has 30 subcategories with billable codes for each type of AMI (410.00 through 410.92). The researcher might only be interested in a specific type of AMI such as inferior and lateral wall AMI (410.4x and 410.5x), or only the initial episode of care (410.x1) which makes automation challenging. The definitions can also be derived from codes in different categories. This problem will become even more complicated in data from 2016 onwards when data will be coded using ICD 10th revision. For comparison, ICD-9 contains approximately 13,000 unique codes but that number grows to around 68,000 in ICD-10. When applying machine learning to the ICD codes, patterns can emerge that help clustering individual billable codes into larger categories thus enriching the definitions of medical conditions of interest. Additionally, the ICD-9/ICD-10 mapping is not one-to-one but rather many-to-many. Machine learning techniques must be applied in many instances to cluster and map the codes from different revisions minimizing the methodology drift risk between the two ICD epochs.

Several reports have been published based on these data. In a recent study, data corresponding to 93,436 patients who had been discharged alive with a first diagnosis of HF were analyzed via machine learning and deep learning to examine the possibility of predicting HF-specific and all-cause readmission and mortality up to one year using covariates such as hospital (including whether or not it is a teaching hospital), age, sex, race, primary insurance, and comorbidities at first hospitalization (e.g., atrial fibrillation, AMI, anemia, chronic kidney disease, chronic obstructive pulmonary disease, diabetes, hyperlipidemia, sleep apnea, Parkinson's disease, stroke, transient ischemic attack). Of the patients, 40,000 were used for training and the remaining 53,436 were assigned to the testing set. The performance (the percent of correct predictions) in the testing set

was 57% for deep learning and 56% for SVM. This performance is not unreasonable given the diversity and high variability of the data.

Example 2: Weng et al. conducted a prospective cohort study using routine clinical data of 378,256 patients from UK family practices [14]. The patients were free from cardiovascular disease at study outset and this data was used to predict whether or not the individuals in the study would have a cardiovascular event during the next 10 years. Among the patients, 24,970 had experienced cardiovascular events over the 10 years. Several different machine learning methods were used to analyze the data. Of these, a neural network algorithm had decent performance: high negative predictive value (NPV) (95.7%), good sensitivity (67.5%), good specificity (70.7%), but low positive predictive value (PPV) (18.4%). Age, gender, ethnicity, chronic kidney disease, mental illness, and atrial fibrillation were among those identified as major risk factors by this algorithm.

Example 3: Ye et al. used data from individual patient EHR extracted from the Maine Health Information Exchange network to develop a risk prediction model for incident essential hypertension within one year [15]. Retrospective (823,627, calendar year 2013) and prospective (680,810, calendar year 2014) cohorts were formed. A machine learning algorithm was used to generate an ensemble of classification trees and assign a predictive risk score to each individual. The model had good performance with 82.3% PPV, 94.9% NPV, 50.9% sensitivity, 98.8% specificity, 0.917 AUC in the retrospective cohort, and 0.870 AUC in an independent prospective cohort. Diabetes, lipid disorders, cardiovascular diseases, mental illness, clinical utilization indicators, and socioeconomic determinants were identified as major risk factors, with the very high-risk population comprised mainly of elderly individuals with multiple chronic conditions. They report that Maine has already deployed their real-time predictive analytic model.

Example 4: The potential of employing wearable biosensors and portable devices as a means of continuous monitoring of assessing a subject's risk of developing hypertension is being explored. Liang et al. studied the possibility of using wearable biosensors which use photoplethysmography (PPG) to assess hypertension [16,17]. They first applied statistical and machine learning tools to select 10 characteristics of the PPG signal, which were then used in four machine learning algorithms to separate hypertensive patients (systolic blood pressure (SBP) over 140 or diastolic blood pressure (DBP) over 100) from non-hypertensive subjects (SBP below 120 or DBP below 80) with high sensitivity (over 80%) and high PPV (over 90%). Separately, Tison et al. used a deep neural network to predict hypertension using heart rate and step count data obtained from off-the-shelf wearables with a PPG heart rate sensor and accelerometer; they found that they could predict hypertension in a validation set with high sensitivity (over 80%) and good specificity (over 60%) [18]. These studies indicate the potential of employing wearable biosensors and portable devices for regular monitoring of patients at risk for hypertension or complications thereof.

4. Discussion

Machine learning is beginning to have an impact on clinical research and practice in various ways. One way is by aiding effective clinical decision making through insights obtained from mining massive amounts of data, such as the MIDAS data mentioned above, and identifying complex patterns associated with hypertension and other clinical conditions. This was not possible to be done at a scale and speed now achieved by machine learning. Some of these findings could allow researchers to study the broad heterogeneity of pathophysiologic factors and processes that interact and contribute to disease and thus could conceivably pave the way to a form of personalized treatment of diseases. At the other end of the spectrum, it could also allow epidemiologists and other researchers to study patterns of disease occurrence and gain a better understanding of what factors in patients' lives influence their health and susceptibility to disease. This should, in turn, enable physicians as well as government bodies and community health organizations

to recommend preventive measures and improve disease management and control.

Another way is by having a means to monitor patient behavior, such as through wearable sensors and other mobile communication devices. This will help clinicians track patient movement and behavioral patterns (perhaps even remotely in real time) and recognize early warning signs of the onset of potentially dangerous conditions. Also, in the case of patients at risk, this will allow physicians to assess compliance with clinical recommendations [17,19–21]. The information obtained from these devices could also be used for the design of clinical research. They could also be useful as a way to develop personalized intervention plans tailored to suit the complex clinical needs unique to individuals. Overall, these developments have the potential to radically change health care. A clinician practicing hypertension may benefit from reading this manuscript and the use of artificial intelligence methods by learning new associations that were unknown before the use of this technique and by learning the opportunity of using such techniques in future research. Artificial intelligence explores big data with high volume, variety, veracity, velocity and value. However, at the present time, there has been much enthusiasm, but little benefit in clinical practice mainly because clinical validation of the findings of artificial intelligence in different data sets is lacking. Also, predictive models change from year to year over time and that requires calibration of model performance. Data quality is critical.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no competing interests.

CRediT authorship contribution statement

Dharmika Amaratunga: Conceptualization, Formal analysis, Software, Writing - original draft. **Javier Cabrera:** Data curation, Formal analysis, Methodology, Software. **Davit Sargsyan:** Data curation, Software, Validation. **John B. Kostis:** Project administration, Validation, Writing - review & editing. **Stavros Zinonos:** Methodology, Validation, Software. **William J. Kostis:** Validation, Visualization, Writing - review & editing.

References

- [1] C. Kritanawong, A.S. Bomback, U. Baber, S. Bangalore, F.H. Messerli, W.H. Wilson Tang, Future direction for using artificial intelligence to predict and manage

- hypertension, *Curr. Hypertens. Rep.* 20 (9) (2018 Jul6) 75, <https://doi.org/10.1007/s11906-018-0875-x>.
- [2] V.J. Dzau, C.A. Balatbat, Future of hypertension, *Hypertension* 74 (3) (2019 Sep) 450–457.
- [3] K.W. Johnson, J.T. Soto, B.S. Glicksberg, K. Shameer, R. Miotto, M. Ali, E. Ashley, J.T. Dudley, Artificial intelligence in cardiology, *JACC (J. Am. Coll. Cardiol.)* 71 (2018) 2668–2679.
- [4] A.A. Miyazawa, Artificial intelligence: the future of cardiology, *Heart* 105 (2019) 1214.
- [5] J. Wellings, J.B. Kostis, D. Sargsyan, J. Cabrera, W.J. Kostis, For the Myocardial Infarction Data Acquisition System (MIDAS 31) Study Group, Risk factors and trends in incidence of heart failure following acute myocardial infarction: risk factors, *Am. J. Cardiol.* 122 (1) (2018 Jul 1) 1–5.
- [6] J. Cabrera, D. Amaratunga, W. Kostis, J. Kostis, Precision Disease Networks (PDN).
- [7] N. Al Falluji, J. Lawrence-Nelson, J.B. Kostis, C.R. Lacy, R. Ranjan, A.C. Wilson AC, Myocardial infarction data acquisition system (MIDAS #8) study group, effect of anemia on 1-year mortality in patients with acute myocardial infarction, *Am. Heart J.* 144 (4) (2002 Oct) 636–641.
- [8] L. Breiman, C.J. Stone, J.D. Gins, New Methods for Estimating Tail Probabilities and Extreme Value Distributions (No. TSC-PD-A226-1), Technology Service Corp, Santa Monica, California, 1979.
- [9] L. Breiman, Bagging predictors, *Mach. Learn.* 24 (2) (1996 Aug) 123–140.
- [10] L. Breiman, Random forests, *Mach. Learn.* 45 (1) (2001 Oct) 5–32.
- [11] R.E. Schapire, The strength of weak learnability, *Mach. Learn.* 5 (2) (1990 Jun) 197–227.
- [12] C. Cortes, V. Vapnik, Support-vector network, *Mach. Learn.* 20 (3) (1995 Sep) 273–297.
- [13] P. Cortez, M.J. Embrechts, Using sensitivity analysis and visualization techniques to open black box data mining models, *Inf. Sci.* 225 (2013) 1–17.
- [14] S.F. Weng, J. Reys, J. Kai, J.M. Garibaldi, N. Qureshi, Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One* 12 (4) (2017 Apr 4), e0174944 <https://doi.org/10.1371/journal.pone.0174944>.
- [15] C. Ye, T. Fu, S. Hao, Y. Zhang, O. Wang, B. Jin, et al., Prediction of incident hypertension within the next year: prospective study using statewide electronic health records and machine learning, *J. Med. Internet Res.* 20 (1) (2018 Jan 30), <https://doi.org/10.2196/jmir.9268> e22.
- [16] Y. Liang, Z. Chen, R. Ward, M. Elgendi, Hypertension assessment using Photoplethysmography: a risk stratification approach, *J. Clin. Med.* 8 (1) (2018 Dec 21), <https://doi.org/10.3390/jcm8010012> pii: E12.
- [17] M. Elgendi, Y. Liang, R. Ward, Toward generating more diagnostic features from photoplethysmogram waveforms, *Diseases* 6 (1) (2018 Mar) pii: E20.
- [18] G.H. Tison, A.C. Singh, D.A. Ohashi, et al., Cardiovascular risk stratification using off-the-shelf wearables and a multi-task deep learning algorithm, *Circulation* 136 (suppl 1) (2017 Nov). Abstract 21042.
- [19] J. Kvedar, M.J. Coye, W. Everett, Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth, *Health Aff.* 33 (2) (2014 Feb) 194–199.
- [20] M. Elgendi, R. Fletcher, Y. Liang, N. Howard, N.H. Lovell, D. Abbott, et al., The use of photoplethysmography for assessing hypertension, *NPJ. Digit. Media* 2 (60) (2019 Jun 26), <https://doi.org/10.1038/s41746-019-0136-7>.
- [21] Y. Wu P. Zhao, W. Li, M.Q. Cao, L. Du, J.C. Chen, The effect of remote health intervention based on internet or mobile communication network on hypertension patients: protocol for a systematic review and meta-analysis of randomized controlled trials, *Medicine* 98 (9) (2019 Mar), <https://doi.org/10.1097/MD.00000000000014707> e14707.