Letter to the editor / List do redakcji

On the necessity for paradigm shift in psychoactive substances research: the implementation of machine learning and artificial intelligence

O konieczności zmiany paradygmatu w badaniach nad substancjami psychoaktywnymi przez wdrożenie uczenia maszynowego i sztucznej inteligencji

Ahmed Al-Imam[1,2] ID, **Marek A. Motyka**[3] ID

[1]Department of Anatomy and Cellular Biology, College of Medicine, University of Baghdad, Iraq
[2]CERVO Brain Research Centre, Faculty of Medicine, University of Laval, Quebec, Canada
[3]Institute of Sociology, University of Rzeszow, Poland

Alcohol Drug Addict 2019; 32 (3): 237-242
DOI: https://doi.org/10.5114/ain.2019.91004

Abstract

The twenty-first century is a time of spectacular scientific advances, especially in information technology and the ongoing scope of application of artificial intelligence. The use of the concepts of machine learning is revolutionising medical and paramedical research. These concepts also apply in the study of psychoactive substances, especially to search for information in the global databases of literature, the systematic review of current publications, the prediction of future trends for the spread of these substances all over

Streszczenie

Dwudziesty pierwszy wiek jest okresem spektakularnego postępu naukowego, zwłaszcza w dziedzinie informatyki i stale rozszerzających się zastosowań sztucznej inteligencji. Wykorzystanie koncepcji uczenia maszynowego rewolucjonizuje badania medyczne i paramedyczne. Tę koncepcję można również zastosować w badaniach nad substancjami psychoaktywnymi, zwłaszcza do wyszukiwania informacji ze światowych baz danych literatury przedmiotu, systematycznego przeglądu pojawiających się publikacji, przewidywania przyszłych trendów rozpowszech-

Correspondence to/Adres do korespondencji: Ahmed Al-Imam, CERVO Brain Research Centre, 2601, de la Canardière, Québec City (Québec), Canada G1J 2G3, phone:+1 (581) 700-0110, e-mail: tesla1452@gmail.com

Authors' contribution/Wkład pracy autorów: **Study design/Koncepcja badania:** A. Al-Imam; **Data collection/Zebranie danych:** A. Al-Imam; **Statistical analysis/Analiza statystyczna:** A. Al-Imam; **Data interpretation/Interpretacja danych:** A. Al-Imam; **Acceptance of final manuscript version/Akceptacja ostatecznej wersji pracy:** A. Al-Imam, M.A. Motyka; **Literature search/Przygotowanie literatury:** A. Al-Imam, M.A. Motyka

No ghostwriting and guest authorship declared./Nie występują zjawiska *ghostwriting* i *guest authorship*.

Submitted/Otrzymano: 23.09.2019 • **Accepted/Przyjęto do druku:** 17.10.2019

the world and also to determine patterns of their misuse and all activities linked to their use.

A systematic review of literature conducted in internet databases indicates apparent deficits in research activity as regards the matters concerning machine learning. It is the authors' opinion that filling this gap will allow a fuller analysis of the phenomenon of psychoactive substance use and identification of emerging tendencies.

The text presents a comprehensive research project applying the proposed concepts, the main aim of which is the evaluation of the prevalence of psychoactive substance use and abuse, including new psychoactive substances at the national and international level, and thus establishing adequate and effective preventive strategies, and developing guidelines for management in crisis situations.

Keywords: Artificial intelligence, Machine learning, Big data, Predictive Analytics, Real-Time Analytics

niania tych substancji na świecie, a także do ustalania wzoru niewłaściwego ich stosowania oraz wszelkich aktywności związanych z ich używaniem.

Systematyczny przegląd literatury przeprowadzany w internetowych bazach danych wskazuje na wyraźne deficyty działań badawczych w zakresie problematyki dotyczącej uczenia maszynowego. Zdaniem autorów wypełnienie tej luki pozwoli na pełniejszą analizę zjawiska używania substancji psychoaktywnych oraz ustalenie pojawiających się tendencji.

W pracy przedstawiono projekt szeroko zakrojonych badań z zastosowaniem proponowanej koncepcji, których głównym celem jest ocena rozpowszechnienia używania i nadużywania substancji psychoaktywnych, w tym nowych substancji psychoaktywnych, na poziomie krajowym i międzynarodowym, a przy tym ustalenie adekwatnych i rzetelnych strategii profilaktycznych oraz opracowanie wytycznych do zarządzania w sytuacjach kryzysowych.

Słowa kluczowe: sztuczna inteligencja, uczenie maszynowe, *Big data*, prognostyczna analiza danych, analiza danych w czasie rzeczywistym

■ Psychoactive substances research: the status quo

The twenty-first century is witnessing phenomenal growth in all disciplines of science, particularly in data science and the ever-expanding applications of artificial intelligence [1]. The exploitation of concepts of machine learning is revolutionising medical and paramedical research [2]. Concomitantly, the investigation of psychoactive substances and novel psychoactive substances (NPS) is evolving at an exponential scale, while the inclusion of artificial intelligence technologies is lagging [3]. A lot of psychoactive substances are labelled as high-risk due to the adverse effects on an individual's health as well as being a burden on the national economy and the healthcare system [4, 5]. Psychoactive chemicals can be classified using a plethora of systems based on their chemical properties, molecular structure, biochemistry, pharmacodynamics and pharmacokinetics [6]. Dargan and Wood developed a structural classification of NPS, based on analyses of substances seized by the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) [7]. The architectural scheme allocated seven categories of NPS including phenethylamines (principally CNS stimulants), tryptamines (hallucinogens), piperazines (CNS stimulants), cathinones (CNS stimulants), cannabimimetics (CB_1 receptor agonists), pipradrol derivatives (CNS stimulants) and miscellaneous substances (mainly CNS stimulants) [7].

■ Machine learning for psychoactive substances research

Although researchers have implemented the use of real-time analytics and predictive models in various specialisations of mathematical and natural sciences, the analog implementation of those techniques is still lagging in describing, assessing, and anticipating phenomena connected to psychoactive substances [2]. Accordingly, we are assuming, i.e., questioning, that there were no attempts to deploy machine learning tools for psychoactive substances research. This assumption applies to data retrieval from literature databases of interest, the systematic review of the literature, and the prediction of future trends on the global diffusion of those chemicals and the patterns of substance misuse and addiction behaviours [8]. Therefore, we shall test our valiant assumption via an evidence-based review of databases of the published literature to weight our opinion on the assumed significant deficit of machine learning-based research activities.

The literature

During the second half of August 2019, we conducted a systematic review of the literature via the Cochrane Library (the Cochrane Database of Systematic Reviews/the Cochrane Collaboration), PubMed (the United States National Library of Medicine), and Embase (Elsevier Database/Scopus). We pragmatically "dissected" the indexed publications by deploying an elaborate set of generic terms and MeSH-based keywords, in combination with Boolean operators as well as truncations, to retrieve potential papers that implemented artificial intelligence technologies including data mining, machine learning, real-time analytics and predictive modelling for big data in connection with psychoactive substances research. Contrary to our expectation, the review strategy generated an impressive total count of 1,496,833 papers distributed into United States National Library of Medicine (1,279,287, 85.47%), Embase (216,655, 14.47%) and the Cochrane Library of Systematic Reviews (891, 0.06%) (Figures 1 and 2). Following a full-text retrieval of the papers, we were able to find that only three peer-reviewed arti-

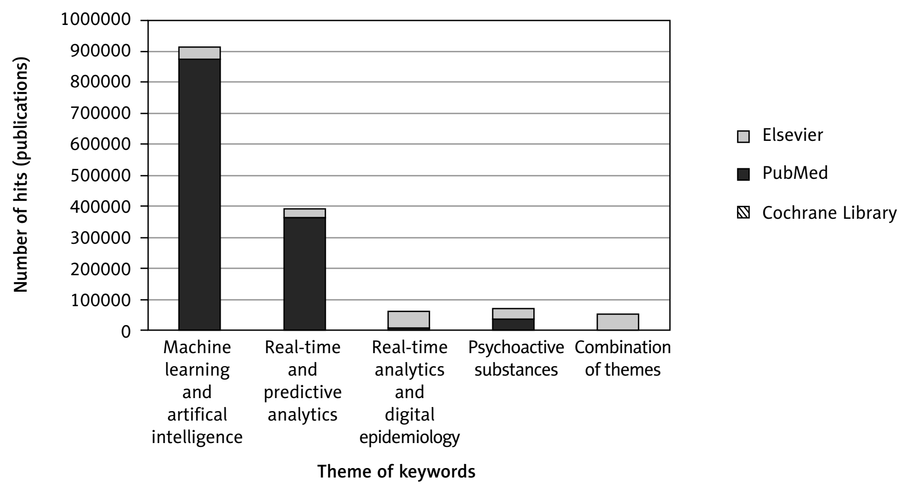

Figure 1. The differential contribution of databases of literature based on thematic keywords search

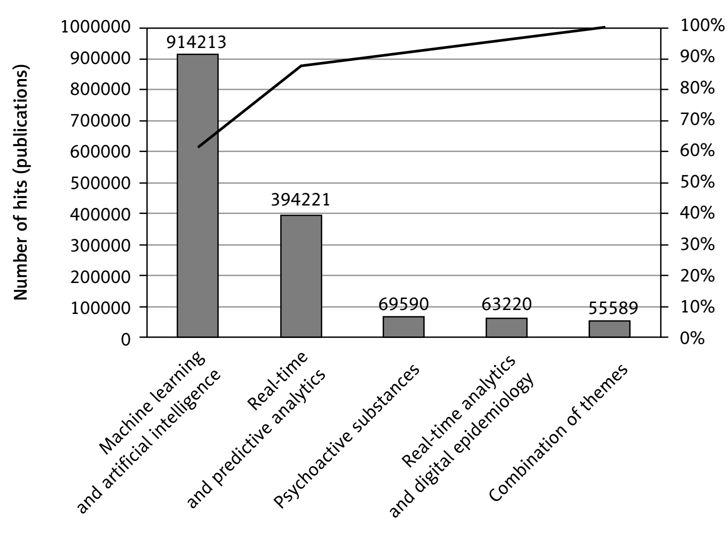

Figure 2. The cumulative contribution of databases of literature based on thematic keywords search

cles, including two original studies and one review article, that deployed real-time analyses [9-11]. However, these were epidemiological studies that implemented waste-water analyses. In conclusion, what we retrieved represents only false-positive data signals originating from the databases of literature, reflecting the specificity of the keywords-based search strategy. Hence, we conclude that there was no implementation of artificial intelligence technologies in any of the publications addressing our primary research question.

■ What is machine learning?

Machine learning relies upon the analyses of big data using a plethora of well-established techniques of mathematical and data science models, including non-biological computational neural networks, diverse modalities of regression analytics, and decision tree classifications [12]. Artificial intelligence attempts to reach the lowest achievable error rates of mathematically-interpreted predictions for causality associations [13]. The exposure-outcome interaction includes the main effect as well as the interaction effects between explanatory variables (predictors) and cofactors (covariables) to forecast the futuristic probability of event occurrences (outcomes). Machine learning is mandatory for unwitnessed benefits when it comes to applications related to the spatio-temporal description and prediction of phenomena of interest, including epidemiological and digital epidemiological investigations of psychoactive substances [2, 14]. The infrastructure of big data upon which machine learning algorithms operate is the same as those designated for the classical epidemiology and digital epidemiological research [3, 14]. Researchers can retrieve data from the databases using survey tools and internet snapshots, longitudinal studies and cross-sectional studies, analyses of web-based social networks, and electronic commerce websites analytics of the surface as well as the deep web including the infamous darknet hypermarket.

■ Anticipated high-impact research: integrated machine learning

High-impact researchers should dedicate to two-pathway informed communications with regulating international organisations including the World Health Organization (WHO), the United Nations Office on Drugs and Crime (UNODC), the Global Public Health Intelligence Network (GPHIN), the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), the World Anti-Doping Agency (WADA), and the Eastern Mediterranean Public Health Network (EMPHNET). For futuristics research projects on psychoactive substances, researchers should endeavour their scientific stamina for implementing a hybrid of the conventional study designs in combination with the real-time analytics in addition to predictive modelling and trends projection for time-series analyses. The challenging and innovating aspect of such projects will be the tendency to incorporate the use of machine learning technologies, neural networks, as well as linear and logistic regression models [15]. Speaking for ourselves, we are aiming to consolidate our prospective research with parallel interdisciplinary doctoral projects while conducting cross-country collaborations among researchers and validated experts. The primary objective is to assess the prevalence of substance use and misuse of specific psychoactive and novel psychoactive chemicals among populations of interests at national and international levels. The data will be retrieved from virtual open-source big data repositories, classical offline records and databases of public health services, and the private psychiatry-mental health clinics specialised in managing patients with addictions and other psychiatric comorbidities. In the chronological order of events, we shall strive to:

1. Conduct a rigorous systematic review of the existing databases of literature in connection with the primary objective of this letter.
2. Fine-tune the systematic retrieval and appraisal of the existing literature by deploying real-time and automated systems (i.e., non-human based retrieval of data) [8, 16]. This step will require the implementation of concepts of automation codes using high-level programming languages (HLL), including Python, R, MatLab and Octave programming languages, in addition to spreadsheet templates and statistical packages for social sciences [17].
3. Assess the prevalence of substance misuse and related phenomena via cross-sectional as well as longitudinal studies.

4. Compare the results on prevalence with collateral data from online resources of interest, including Google Trends and Google Analytics open-source deposits of big data to assess the digital epidemiology and geographic mapping of the phenomena at a global scale [18].

5. Build an exhaustive predictive model to anticipate which individuals possess a tendency to develop substance misuse disorder in their lifetime based on multivariate analyses and regression models.

6. Attempt to enhance the predictive power by refining the mathematical basis and the computational methods for real-time and predictive models via using reliable data transformers to boost the accuracy of robust statistics of the results while reducing the computational processing demands that are critical to analysing exhaustive sets of big data and potentially in real-time [19, 20].

7. Experiment with various non-Bayesian statistics to extrapolate data on the patterns of addiction for between-subjects comparative analytics. We shall achieve this by comparing the consistency of different statistical analyses, including Point Biserial Correlation, Independent Student's t-test and Fisher's One-Factorial ANOVA [21, 22].

8. Evaluate the predictive tools (models) in connection with their sensitivity and specificity, predictive values (positive and negative) and the magnitude of error (statistical noise) [23, 24].

9. Execute concepts of integral calculus in receiver operating characteristics (ROC) for an assessment of the prognostic precision of predictive analytics.

10. Enact and actualise our models to foresee an impending and ominous geo-specific addiction crisis and future substance misuse epidemics.

11. Create pre-disaster prophylactic measures by creating a customisable computer-based application based on the already established tools for use by diagnosticians and interventional healthcare professionals.

12. Formulate practical and externally-valid guidelines for managing the worst-case scenario of disaster situations through realistic collaboration with public and private health authorities. The guidelines represent data-driven and evidence-based consultations on how to prognosticate and control a catastrophic substance misuse epidemics and by engaging community services to assess, analyse and manage the underlying predisposition in individuals and populations towards addictive behaviours.

13. Finally, we shall disseminate our knowledge by communicating our collaborative and evidence-based expertise with global regulatory agencies and health organisations, thus aiming for global dissemination of unbiased high-impact applicable science [25].

Conflict of interest/Konflikt interesów

None declared./Nie występuje.

Financial support/Finansowanie

None declared./Nie zadeklarowano.

Ethics/Etyka

The work described in this article has been carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) on medical research involving human subjects, EU Directive (210/63/EU) on protection of animals used for scientific purposes, Uniform Requirements for manuscripts submitted to biomedical journals and the ethical principles defined in the Farmington Consensus of 1997.

Treści przedstawione w pracy są zgodne z zasadami Deklaracji Helsińskiej odnoszącymi się do badań z udziałem ludzi, dyrektywami UE dotyczącymi ochrony zwierząt używanych do celów naukowych, ujednoliconymi wymaganiami dla czasopism biomedycznych oraz z zasadami etycznymi określonymi w Porozumieniu z Farmington w 1997 roku.

References/Piśmiennictwo

1. Jackson PC, Jr. *Introduction to artificial intelligence*. 2nd Enlarged Edition, Series: Dover Books on Mathematics. New York: Dover Publications Inc.; 1985. [Reprint of the Petrocelli/Charter, New York, 1974 ed.].
2. Lin SY, Mahoney MR, Sinsky CA. Ten ways artificial intelligence will transform primary care. *J Gen Intern Med* 2019; 34(8): 1626-30.
3. Al-Imam A. *Monitoring and Analysis of Novel Psychoactive Substances in Trends Databases, Surface Web and the Deep Web, with Special Interest and Geo-Mapping of the Middle East*. eu-repo/semantics/masterThesis [dissertation on the Internet]. United Kingdom: University of Hertfordshire; 2017. DOI: 10.13140/RG.2.2.27636.24961.
4. Al-Imam A. Retrospective Analyses of High-risk NPS: Integrative Analyses of PubMed, Drug Fora, and the Surface Web. *Glob J Health Sci* 2017; 9(11): 40-50.
5. Motyka MA, Al-Imam A. Musical Preference and Drug Use among Youth: An Empirical Study. *Res Adv Psychiatry* 2019; 6(2): 50-7.
6. Catalani V, Prilutskaya M, Al-Imam A, Marrinan S, Elgharably Y, Zloh M, et al. Octodrine: New Questions and Challenges in Sport Supplements. *Brain Sci* 2018; 8(2): 1-13.
7. Dargan P, Wood D (eds.). *Novel psychoactive substances: classification, pharmacology and toxicology*. London: Academic Publishing; 2013.
8. Al-Imam A, Khalid U, Al-Hadithi N, Kaouche D. Real-time Inferential Analytics Based on Online Databases of Trends: A Breakthrough Within the Discipline of Digital Epidemiology of Dentistry and Oral-Maxillofacial Surgery. *Mod Appl Sci* 2019; 13(2): 81-94.
9. Leung KS, Cottler LB. Ecstasy and other club drugs: a review of recent epidemiologic studies. *Curr Opin Psychiatry* 2008; 21(3): 234-41.
10. Postigo C, de Alda ML, Barceló D. Evaluation of drugs of abuse use and trends in a prison through wastewater analysis. *Environ Int* 2011; 37(1): 49-55.
11. Zuccato E, Chiabrando C, Castiglioni S, Bagnati R, Fanelli R. Estimating community drug abuse by wastewater analysis. *Environ Health Perspect* 2008; 116(8): 1027-32.
12. Michie D, Spiegelhalter DJ, Taylor CC. *Machine Learning: Neural and Statistical Classification*. London, UK: Ellis Horwood Ltd.; 1994.
13. Witten IH, Frank E, Hall MA, Pal CJ. *Data Mining: Practical machine learning tools and techniques*. Burlington: Morgan Kaufmann; 2016.
14. Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clin Infect Dis* 2017; 66(1):149-53.
15. Durstewitz D, Koppe G, Meyer-Lindenberg A. Deep neural networks in psychiatry. *Mol Psychiatry* 2019; 24, 11: 1583-98.
16. Al-Imam A. Inferential Analysis of Big Data in Real-Time: One Giant Leap for Spatiotemporal Digital Epidemiology in Dentistry. *ORAL Implantol* 2019; 12(1): 1-14.
17. Al-Imam A. Novel Psychoactive Substances Research: On the Necessity of Real-time Analytics and Predictive Modelling. *Res Adv Psychiatry* 2019; 7(1) [In press].
18. Al-Imam A, Abdul Majeed BA. The Most Popular Chemical Categories of NPS in Four Leading Countries of the Developed World: An Integrative Analysis of Trends Databases, Surface Web, and the Deep Web. *Glob J Health Sci* 2017; 9(11): 27-39.
19. Rahman MA, Dash PK, Downton ER. Digital protection of power transformer based on weighted least square algorithm. *IEEE Trans Power App Syst* 1982: 4204-10.
20. Razali NM, Wah YB. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *J Stat Modelling and Anal* 2011; 2(1): 21-33.
21. Cohen J. Statistical power analysis. *Curr Dir Psychol Sci* 1992; 1(3): 98-101.
22. Gelman A. Objections to Bayesian statistics. *Bayesian Anal* 2008; 3(3): 445-9.
23. Akobeng AK. Understanding type I and type II errors, statistical power and sample size. *Acta Paediatr* 2016; 105(6): 605-9.
24. Kraemer HC, Blasey C. *How many subjects?: Statistical power analysis in research. Thousand Oaks*, CA: Sage Publications; 2015.
25. UNODC. World drug report 2019. New York, NY: United Nations Office on Drug and Crime; 2019.