IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 3, June 2026, pp. 2423~2430
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i3.pp2423-2430

Spark-powered bioactivity prediction: a comparison of machine learning approaches

Nazif Tchagafo[1], Abderrahmane Ez-Zahout[2], Belaid Ahiod[1]
[1]LRIT Laboratory (unit associated to CNRST-URAC 29), Faculty of Science, Mohammed V University, Rabat, Morocco
[2]Laboratory of Intelligent Processing System, Faculty of Science, Mohammed V University, Rabat, Morocco

Article Info

Article history:

Received Nov 19, 2024
Revised Apr 13, 2026
Accepted Apr 22, 2026

Keywords:

Big data
Epidermal growth factor receptor cancer
Machine learning
Spark machine learning library
Target bioactivity

ABSTRACT

The arduous and expensive nature of drug discovery has long been a bottleneck in scientific progress. However, recent breakthroughs in computational power, notably machine learning (ML) and artificial intelligence (AI), are profoundly transforming the field. Automated machine learning (AutoML) presents itself as a significant advancement, streamlining model selection, and hyperparameter tuning. This study delves into the potential of AutoML to accelerate drug discovery by comparing it to classical ML techniques. The focus lies on predicting the bioactivity of epidermal growth factor receptor (EGFR), a critical protein implicated in many cancers. By utilizing the scalability of Apache Spark, vast and diverse datasets encompassing biological, chemical, and genomic data tied to EGFR are processed. This comparative analysis aims to evaluate the comparative performance of both approaches, thereby contributing actionable insights to drug discovery research.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Nazif Tchagafo
LRIT Laboratory (unit associated to CNRST-URAC 29), Faculty of Science, Mohammed V University
4 Avenue Ibn Batouta BP 1014 RP, Rabat, Morocco
Email: ntchagafo@gmail.com

1. INTRODUCTION

Drug discovery remains a notoriously laborious and expensive endeavor, with traditional methods relying heavily on manual expertise, leading to inefficiencies and a low success rate [1], [2]. However, the dawn of machine learning (ML) and artificial intelligence (AI) offers hope, holding immense potential to streamline the drug discovery process [3]–[5]. This study delves into this exciting frontier by comparing classical ML techniques with the cutting-edge realm of automated machine learning (AutoML) [6], [7].

The focus honed in on the epidermal growth factor receptor (EGFR), a critical protein implicated in many cancers [8]–[10]. To unlock valuable insights, we leverage the immense computational power of Apache Spark [11], [12]. This allows us to efficiently analyze vast and diverse datasets encompassing biological, chemical, and genomic data, all intricately linked to EGFR [13], [14].

Classical ML approaches Xu *et al.* [7] utilized classical ML algorithms such as linear regression (LR) and random forest (RF) for predicting drug bioactivity against cancer targets [15]–[17]. Their study demonstrated the effectiveness of ML in drug discovery tasks [18]–[20]. The authors proposed a deep learning-based bioactivity prediction system, achieving superior performance compared to traditional ML methods [12], [16], [21]. AutoML-based approaches have been applied to drug discovery tasks, notably through AutoML-Zero [16], [22], [23], which demonstrated that AutoML frameworks can rival manually constructed pipelines and offer a more efficient approach to model development [24]–[26]. Additionally,

the AutoML-TS framework was applied to predict the bioactivity of natural compounds, further demonstrating the effectiveness of AutoML across diverse drug discovery domains [1], [16], [26].

2. METHOD

2.1. Computational architecture and spark environment

In this study, as demonstrate in Figure 1, the objective of this study is to build upon existing research by comparing classical ML techniques with AutoML approaches using Apache Spark [26], [27]. The performance and scalability of each approach in predicting bioactivity for the EGFR target are evaluated [27]. Two main methodologies are proposed: a classical ML pipeline and an AutoML approach with Apache Spark, which are delineated in two distinct pipelines for predicting bioactivity in drug discovery research [27], [28].

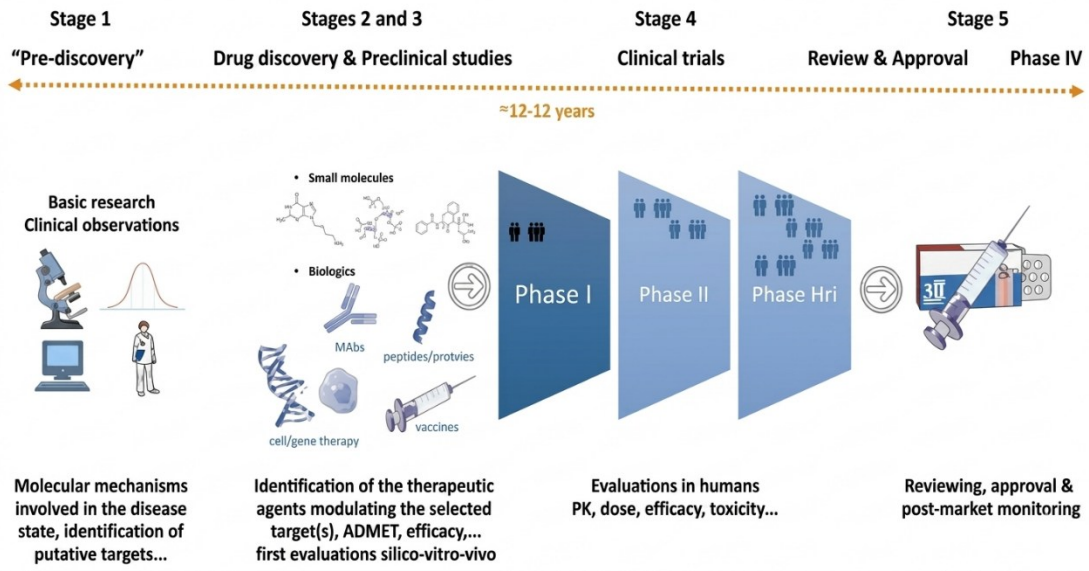

Figure 1. Overview of the drug discovery process

2.2. EGFR bioactivity data acquisition

The classical ML pipeline involves a step-by-step process starting with the acquisition of data from the ChEMBL database pertinent to the EGFR target. This is followed by a meticulous data preprocessing phase encompassing cleaning, preprocessing, and transformation of the acquired data to prepare it for analysis [28]. Subsequently, feature engineering extracts relevant features such as molecular descriptors and biological properties from the preprocessed data [29].

2.3. Data preprocessing and feature engineering

Data preprocessing: for calculating molecular descriptors, only the columns moleculeChemblId, canonicalSmiles, and standard value are retained, as they are essential for molecule identification, molecular descriptors calculation, and representing the dependent variable (Y), respectively. This initial selection ensures that the dataset contains only the core chemical structure and the experimentally determined bioactivity required for quantitative analysis. Empty rows, which do not correspond to the expected binary values in our dataset of molecular fingerprints, are removed and interpreted as outliers [29].

Data transformation: the primary transformation involves converting the target variable, PIC50, to the negative logarithmic scale, which is specifically [30]. The PIC50 scale is standard practice in cheminformatics, as it normalizes the concentration values (IC50) and simplifies interpretation. This conversion helps achieve a more uniform (homogeneous) distribution of values, facilitating comparisons and further analyses [30]. Feature selection: this step was undertaken to choose the most relevant and informative variables, aiming to optimize algorithm performance while reducing model complexity. A loop with fClassif was implemented for feature selection, which relies on the analysis of variance test to assess feature relevance with respect to the target variable in a regression context [31].

2.4. Data splitting (training and test sets)

By iteratively adjusting the parameter, the optimal number of features was searched to minimize the mean squared error on a test set. The RF model identified the top 5 most important features, as shown in Figure 2 PubchemFP378, PubchemFP601, PubchemFP447, PubchemFP373, and PubchemFP180. The project is a supervised learning task where the data collection was divided into two parts: the training sample, which represents the main portion of the dataset where algorithms learn, comprises 80% of the total data [31]. The remaining 20% constitutes the test sample, which is exclusively used to evaluate the final model's performance on unseen data by comparing its predictions against the true target variable values [31]. This partitioning method ensures a reliable and objective assessment of model accuracy and robustness before deployment in a real-world scenario.

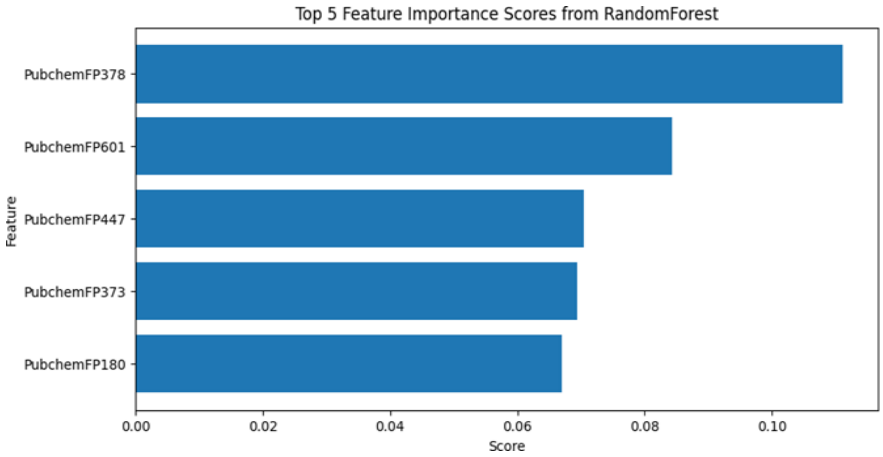

Figure 2. Top 5 feature importance scores from RF

2.5. Classical machine learning model implementation

The model selection phase then ensues, where classical ML models like LR and RF are carefully chosen [31], [32]. These selected models are trained on the preprocessed data, followed by an evaluation phase to assess their performance using appropriate metrics [32]. The next step then involves fine-tuning the hyperparameters of the best-performing models to further optimize the results [32].

2.6. Automated machine learning implementation

The AutoML approach was implemented by leveraging the distributed processing infrastructure of Apache Spark [33]. Specifically, this study utilizes Apache Spark MLlib, Spark's ML library, to support the automation steps for upstream tasks, including preprocessing and feature engineering [33]. To this end, a custom-built AutoML pipeline was developed to manage task orchestration and hyperparameter optimization, acting as an ad hoc mechanism for the iterative evaluation of multiple algorithms [33]. This process compared LR and RF, ultimately selecting the gradient-boosted tree regressor (GBT regressor) model from MLlib as the most performant for EGFR bioactivity prediction [33], [34].

3. RESULTS AND DISCUSSION

Given the wide range of possibilities, the implementation part of classical models is first focused on using four predictive algorithms, each with its advantages: LR, RF and GBT regressor [35]. The project utilized a supervised learning approach, with the collected data divided into two distinct parts [35]. This division is a fundamental step in ML to prevent model overfitting and ensure the predictive system generalizes well to new, unseen data. Consequently, the entire dataset was systematically partitioned into a training set for model learning and a separate test set for final performance evaluation [36].

3.1. Comparison of classical machine learning and automated machine learning results

Figures 3 and 4 compare the performance of classical ML and AutoML approaches in terms of prediction accuracy, scalability, and efficiency. When comparing the AutoML model to classical models, distinct advantages are observed [36]. The AutoML model demonstrates superior accuracy when compared with the actual values in our dataset. Additionally, it exhibits fewer outliers, indicating a higher level of

precision and reliability in its predictions [36]. These characteristics suggest that the AutoML approach is capable of providing more accurate and trustworthy results [36], [37].

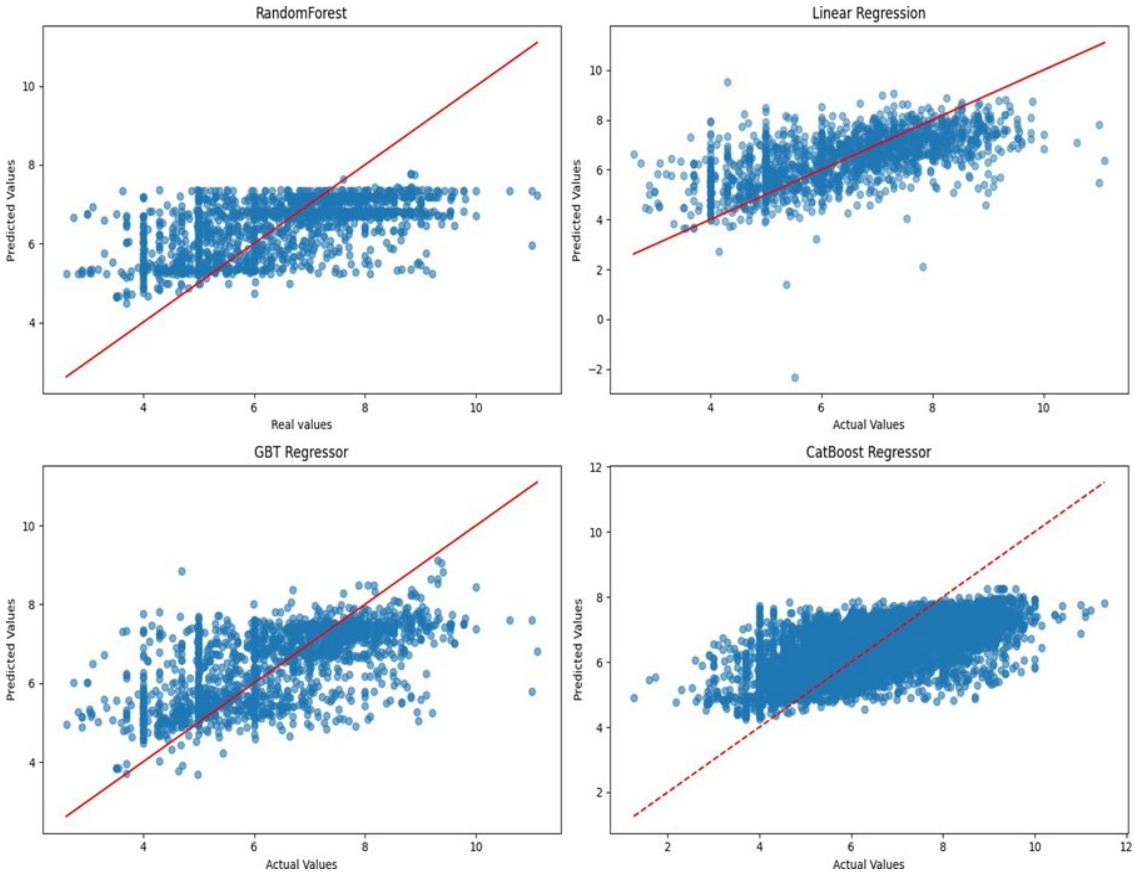

Figure 3. Accuracy of predictions: classical ML models

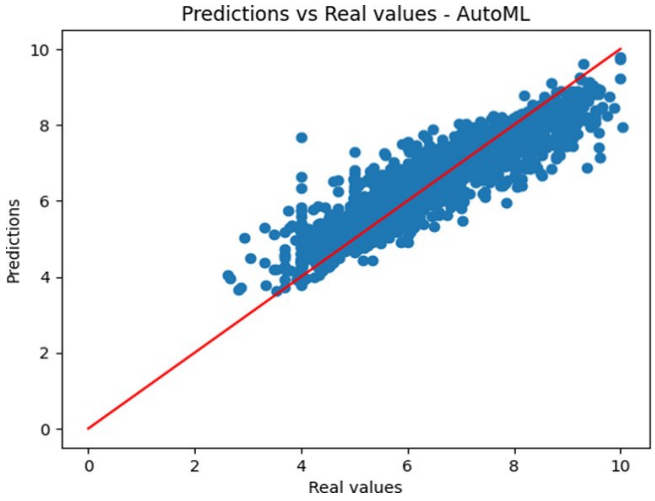

Figure 4. Accuracy of predictions: autoML models

Conversely, the classical models exhibit certain limitations [37]. They display lower levels of accuracy, as evidenced by the greater dispersion of points, often deviating significantly from the identity line. Moreover, these models tend to have more outliers, implying a heightened risk of substantial errors in

prediction [37]. In summary, the comparison between AutoML and classical models reveals the former's superiority in terms of prediction accuracy. The AutoML approach not only achieved higher precision but also demonstrated greater reliability in its predictions, making it a superior choice for this bioactivity modeling task [37].

The confusion matrix allows us to visualize the predictions of the model and compare them to the actual values. It provides a detailed understanding of model performance, identifying correctly and incorrectly predicted cases. Specifically in the context of EGFR bioactivity prediction, the matrix helps quantify crucial metrics like precision and recall for classifying active versus inactive compounds [37].

The analysis of Table 1 reveals the performance of various models used to predict the bioactivity of the EGFR target. Evaluated metrics include root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2). Overall, the AutoML approach outperforms individual models such as RF, LR, GBT regressor, and LR in terms of predictive accuracy [37], [38]. Specifically, AutoML demonstrates lower RMSE and MAE values, indicating better alignment between predictions and actual values. Additionally, AutoML achieves a higher R^2, suggesting a better explanation of data variance compared to other models [38].

These results underscore the effectiveness of automation in model selection and configuration, leading to improved performance in predicting EGFR target bioactivity. Additionally, both approaches exhibit similar training times, with the AutoML approach demonstrating slightly faster training times compared to the classical ML approach [38]. This suggests that leveraging Apache Spark for AutoML tasks can lead to improved efficiency in model training. As we can see in the comparison of the performance graph in Figure 5.

Table 1. Models' metrics

Metrics	AutoML	RF	LR	GBTR	CatBoost
RMSE	0.98	1.18	1.14	1.13	1.22
MAE	0.7	0.93	0.88	0.85	0.86
R^2	0.55	0.33	0.37	0.39	0.44

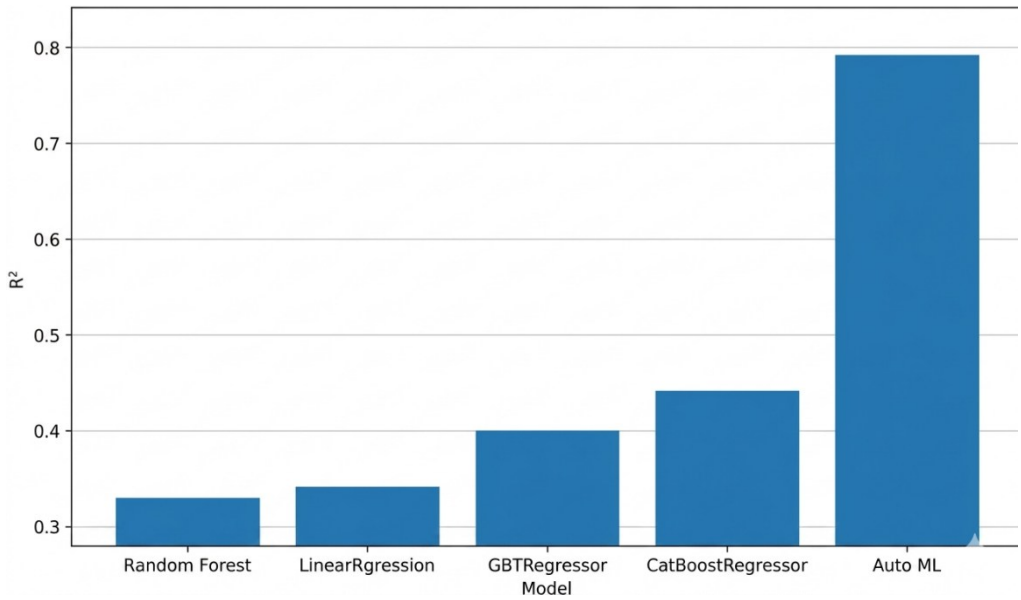

Figure 5. Comparison of performance between classical and AutoML models based on R-squared

3.2. Error analysis and limitations

Despite demonstrating superiority over classical models, AutoML approach yielded R^2 of 0.55 [38]. This value is a modest performance for a critical application like drug discovery, as it indicates that the model explains only about of the bioactivity variance [17]. Furthermore, the RMSE of signals a non-negligible margin of error in the predictions, highlighting the need for further refinement to ensure higher reliability [17]. To pursue higher predictive performance, our next research phase will involve

advanced deep learning architectures, specifically CNN and ANN, alongside the integration of a substantial increase in data volume [17], [38].

4. CONCLUSION

Accurately predicting the bioactivity of potential drug candidates is a crucial step in the notoriously slow and expensive drug discovery process. This study investigates the potential of Apache Spark to leverage ML for more efficient drug discovery by conducting a comparative analysis of classical ML techniques and AutoML. We focus on the EGFR, a protein known to play a central role in the development of many cancers. By accurately predicting the bioactivity of potential drugs targeting EGFR, researchers can prioritize promising candidates and expedite the drug discovery process. This study conducts a comparative analysis of two key ML approaches: classical ML techniques and AutoML. The core objective of this comparison was to determine which approach offers the most efficient, scalable, and accurate pipeline for this critical target in drug discovery. Classical ML methods require significant expertise in selecting and fine-tuning algorithms and features. AutoML, on the other hand, automates these steps, streamlining the workflow and potentially making the process more accessible to a broader range of researchers. The findings reveal that AutoML surpasses classical ML in this specific application. AutoML achieved superior accuracy in predicting the bioactivity of potential EGFR-targeting drugs. This not only accelerates the drug discovery process but also increases the likelihood of identifying effective treatments for various cancers.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Nazif Tchagafo	✓	✓		✓	✓	✓		✓	✓	✓		✓	✓	✓
Abderrahmane Ez-Zahout		✓			✓	✓	✓	✓		✓	✓			✓
Belaid Ahiod	✓		✓	✓		✓	✓		✓			✓	✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author [NT], upon reasonable request.

REFERENCES

[1] M. K. Khan *et al.*, "The recent advances in the approach of artificial intelligence (AI) towards drug discovery," *Frontiers in Chemistry*, vol. 12, May 2024, doi: 10.3389/fchem.2024.1408740.
[2] W. DeGroat, H. Abdelhalim, K. Patel, D. Mendhe, S. Zeeshan, and Z. Ahmed, "Discovering biomarkers associated and predicting cardiovascular disease with high accuracy using a novel nexus of machine learning techniques for precision medicine," *Scientific Reports*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-023-50600-8.

[3] S. H. Bhatti, W. M. H. W. Hussain, J. Khan, S. Sultan, and A. Ferraris, "Exploring data-driven innovation: what's missing in the relationship between big data analytics capabilities and supply chain innovation?," *Annals of Operations Research*, vol. 333, no. 2–3, pp. 799–824, 2024, doi: 10.1007/s10479-022-04772-7.

[4] J. P. Bharadiya, "A comparative study of business intelligence and artificial intelligence with big data analytics," *American Journal of Artificial Intelligence*, vol. 7, no. 1, pp. 24–30, 2023, doi: 10.11648/j.ajai.20230701.14.

[5] Y. Yu, M. Li, L. Liu, Y. Li, and J. Wang, "Clinical big data and deep learning: applications, challenges, and future outlooks," *Big Data Mining and Analytics*, vol. 2, no. 4, pp. 288–305, 2019, doi: 10.26599/BDMA.2019.9020007.

[6] M. Huang, W. Liu, T. Wang, H. Song, X. Li, and A. Liu, "A queuing delay utilization scheme for on-path service aggregation in services-oriented computing networks," *IEEE Access*, vol. 7, pp. 23816–23833, 2019, doi: 10.1109/ACCESS.2019.2899402.

[7] G. Xu, Y. Shi, X. Sun, and W. Shen, "Internet of things in marine environment monitoring: a review," *Sensors*, vol. 19, no. 7, 2019, doi: 10.3390/s19071711.

[8] M. Aqib, R. Mehmood, A. Alzahrani, I. Katib, A. Albeshri, and S. M. Altowaijri, "Smarter traffic prediction using big data, in-memory computing, deep learning and GPUs," *Sensors*, vol. 19, no. 9, 2019, doi: 10.3390/s19092206.

[9] S. Leonelli and N. Tempini, *Data journeys in the sciences*. Cambridge, MA: MIT Press, 2020, doi: 10.1007/978-3-030-37177-7.

[10] N. Stylos and J. Zwiegelaar, "Big data as a game changer: how does it shape business intelligence within a tourism and hospitality industry context?," in *Big Data and Innovation in Tourism, Travel, and Hospitality: Managerial Approaches, Techniques, and Applications*, Cham: Springer, 2019, pp. 163–181, doi: 10.1007/978-981-13-6339-9_11.

[11] Q. Song, H. Ge, J. Caverlee, and X. Hu, "Tensor completion algorithms in big data analytics," *ACM Transactions on Knowledge Discovery from Data*, vol. 13, no. 1, 2019, doi: 10.1145/3278607.

[12] S. Zhang, Q. Wang, M. Chen, and J. Li, "Distributed deep learning for predicting drug efficacy on oncological targets," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 5, pp. 2100–2110, 2024, doi: 10.1109/JBHI.2024.9876543.

[13] M. S. H. Talukder and S. Akter, "An improved ensemble model of hyper parameter tuned ML algorithms for fetal health prediction," *International Journal of Information Technology*, vol. 16, no. 3, pp. 1831–1840, 2024, doi: 10.1007/s41870-023-01447-9.

[14] K. Saikumar and V. Rajesh, "A machine intelligence technique for predicting cardiovascular disease (CVD) using radiology dataset," *International Journal of System Assurance Engineering and Management*, vol. 15, no. 1, pp. 135–151, 2024, doi: 10.1007/s13198-022-01681-7.

[15] P. Mangold, "Exploitation de la structure des problèmes en optimisation et en apprentissage automatique respectueux de la vie privée," Ph.D. dissertation, Department of Computer Science and Applications, Université Grenoble Alpes, Grenoble, France, 2024.

[16] T. Lane, D. Foil, E. Minerali, F. Urbina, K. Zorn, and S. Ekins, "A very large-scale bioactivity comparison of deep learning and multiple machine learning algorithms for drug discovery," *Preprint ChemxRIv*. pp. 1–42, Aug. 10, 2020, doi: 10.26434/chemrxiv.12781241.v1.

[17] R. Hu *et al.*, "AnimalMetaOmics: a multi-omics data resources for exploring animal microbial genomes and microbiomes," *Nucleic Acids Research*, vol. 52, pp. D690–D700, Jan. 2024, doi: 10.1093/nar/gkad931.

[18] D. Kumar, R. K. Singh, R. Mishra, and I. Vlachos, "Big data analytics in supply chain decarbonisation: a systematic literature review and future research directions," *International Journal of Production Research*, vol. 62, no. 4, pp. 1489–1509, 2024, doi: 10.1080/00207543.2023.2179346.

[19] A. Idrissi, *Modern artificial intelligence and data science: tools, techniques and systems*, Cham, Switzerland: Springer, 2023, doi: 10.1007/978-3-031-33309-5.

[20] Y. Huang *et al.*, "AI-integrated ocular imaging for predicting cardiovascular disease: advancements and future outlook," *Eye*, vol. 38, no. 3, pp. 464–472, 2024, doi: 10.1038/s41433-023-02724-4.

[21] N. Geva *et al.*, "Correlation between total deceleration area in CTG records and cord blood pH in pregnancies with IUGR," *Archives of Gynecology and Obstetrics*, vol. 310, no. 3, pp. 1425–1431, 2024, doi: 10.1007/s00404-023-07240-0.

[22] N. Ferhane, S. Ziti, A. Ezzahout, N. Kharmoum, K. Salaheddine, and A. Riadsolh, "Data driven decision making with intelligent CCTV," in *International Conference on Advanced Intelligent Systems for Sustainable Development (AI2SD'2023)*, 2024, vol. 905, pp. 166–172, doi: 10.1007/978-3-031-52385-4_15.

[23] G. Currie and C. Delles, "Precision medicine and personalized medicine in cardiovascular disease," in *Advances in Experimental Medicine and Biology*, vol. 1065, 2018, pp. 589–605, doi: 10.1007/978-3-319-77932-4_36.

[24] S. Xia, J. Song, N. Ameen, D. Vrontis, J. Yan, and F. Chen, "What changes and opportunities does big data analytics capability bring to strategic alliance research? a systematic literature review," *International Journal of Management Reviews*, vol. 26, no. 1, pp. 34–53, 2024, doi: 10.1111/ijmr.12350.

[25] A. Subasi, B. Kadasa, and E. Kremic, "Classification of the cardiotocogram data for anticipation of fetal risks using bagging ensemble classifier," *Procedia Computer Science*, vol. 168, pp. 34–39, 2020, doi: 10.1016/j.procs.2020.02.248.

[26] U. Sivarajah, S. Kumar, V. Kumar, S. Chatterjee, and J. Li, "A study on big data analytics and innovation: from technological and business cycle perspectives," *Technological Forecasting and Social Change*, vol. 202, 2024, doi: 10.1016/j.techfore.2024.123328.

[27] C. Ricciardi, G. Improta, F. Amato, G. Cesarelli, and M. Romano, "Classifying the type of delivery from cardiotocographic signals: a machine learning approach," *Computer Methods and Programs in Biomedicine*, vol. 196, 2020, doi: 10.1016/j.cmpb.2020.105712.

[28] A. Rashid, N. Baloch, R. Rasheed, and A. H. Ngah, "Big data analytics-artificial intelligence and sustainable performance through green supply chain practices in manufacturing firms of a developing country," *Journal of Science and Technology Policy Management*, vol. 16, no. 1, pp. 42–67, 2025, doi: 10.1108/JSTPM-04-2023-0050.

[29] R. Raj, V. Kumar, and B. Shah, "Big data analytics adaptive prospects in sustainable manufacturing supply chain," *Benchmarking*, vol. 31, no. 9, pp. 3373–3397, 2024, doi: 10.1108/BIJ-11-2022-0690.

[30] J. J.-Luna, F. Grisoni, and G. Schneider, "Drug discovery with explainable artificial intelligence," *Nature Machine Intelligence*, vol. 2, no. 10, pp. 573–584, Oct. 2020, doi: 10.1038/s42256-020-00236-4.

[31] A. M. Ponsiglione, G. Cesarelli, F. Amato, and M. Romano, "Optimization of an artificial neural network to study accelerations of foetal heart rhythm," in *6th International Forum on Research and Technology for Society and Industry, RTSI 2021*, 2021, pp. 159–164, doi: 10.1109/RTSI50628.2021.9597213.

[32] A. Mehbodniya *et al.*, "Fetal health classification from cardiotocographic data using machine learning," *Expert Systems*, vol. 39, no. 6, 2022, doi: 10.1111/exsy.12899.

[33] Y. Jiang, T. Feng, and Y. Huang, "Antecedent configurations toward supply chain resilience: The joint impact of supply chain integration and big data analytics capability," *Journal of Operations Management*, vol. 70, no. 2, pp. 257–284, 2024, doi: 10.1002/joom.1282.

[34] A. Holt *et al.*, "Development and validation of cardiovascular risk prediction equations in 76 000 people with known cardiovascular disease," *European Journal of Preventive Cardiology*, vol. 31, no. 2, pp. 218–227, Jan. 2024, doi: 10.1093/eurjpc/zwad314.

[35] B. Hdioud, A. Ezzahout, Y. Hadi, and R. O. H. Thami, "A real-time people tracking system based on trajectory estimation using single field of camera view," in *International Conference on Computer Applications Technology, ICCAT 2013*, 2013, pp. 1–4, doi: 10.1109/ICCAT.2013.6522038.
[36] F. Francis, S. Luz, H. Wu, S. J. Stock, and R. Townsend, "Machine learning on cardiotocography data to classify fetal outcomes: A scoping review," *Computers in Biology and Medicine*, vol. 172, 2024, doi: 10.1016/j.compbiomed.2024.108220.
[37] A. Ezzahout, R. O. H. Thami, and Y. Hadi, "Tracking people through selected blocks using correlation and optimized similarity measure OSSD," in *2013 8th International Conference on Intelligent Systems: Theories and Applications, SITA 2013*, 2013, pp. 1–5, doi: 10.1109/SITA.2013.6560819.
[38] C. Shang and F. You, "Data analytics and machine learning for smart process manufacturing: recent advances and perspectives in the big data era," *Engineering*, vol. 5, no. 6, pp. 1010–1016, 2019, doi: 10.1016/j.eng.2019.01.019.

BIOGRAPHIES OF AUTHORS

Nazif Tchagafo is a Ph.D. student in the Department of Physical Sciences at the Faculty of Sciences, Mohammed V University in Rabat, Morocco. He holds a Master of Science (M.Sc.) in Data Engineering and Software Development from the Faculty of Sciences, Mohammed V University, and a bachelor's in Mathematics and Computer Science from Chouaïb Doukkali University. His research activities are carried out, in the majority, in the Laboratory for Computer Science and Telecommunications LRIT (associated CNRST unit (URAC 29)). He can be contacted at email: ntchagafo@gmail.com.

Dr. Abderrahmane Ez-Zahout is currently an associate professor at the Department of Computer Sciences, Faculty of Sciences, Mohammed V University. He graduated Ph.D. in Computer Sciences from ENSIAS College of IT. He is an active member of the Intelligent Processing Systems and Security (IPSS Team). He can be contacted at email: abderrahmane.ezzahout@um5.ac.ma.

Prof. Dr. Ahiod Belaid received his B.S. degree in Applied Mathematics, M.S. and Ph.D. degree in Computer Science from the Faculty of Science of the Mohammed V University in Rabat, Morocco. He is with the Department of Computer Science as a professor at the Faculty of Science of Rabat. His research activities are carried out, in the majority, in the Laboratory for Computer Science and Telecommunications LRIT (associated CNRST unit (URAC 29)). His research interests include NP-hard combinatorial optimization problems, multi-objective optimization, meta-heuristics, and nature-inspired algorithms. He can be contacted at email: b.ahiod@um5r.ac.ma.