BIG DATA AND PREDICTIVE ANALYTICS IN HEALTHCARE: INTERDISCIPLINARY APPROACHES TO BUSINESS, LAW, AND EDUCATION

Dr. Mukesh Kumar Sahu[1], Mr. Arun Kumar Lahre[2], Mr. Rajesh Kumar Sahu[3]

[1]Department of Mechanical Engineering, Thermal Engineering and Heat Transfer, Kalinga University, Naya Raipur Chhattisgarh, India
[2]Department of Mechanical Engineering, Surface Mine Technology,Kalinga University, Naya Raipur Chhattisgarh, India
[3]Department of Mechanical Engineering, Strata Control and Mine Ventilation,Kalinga University, Naya Raipur Chhattisgarh, India

ku.mukeshkumarsahu@kalingauniversity.ac.in[1]
ku.arunkumarlahre@kalingauniversity.ac.in[2]
ku.rajeshkumarsahu@kalingauniversity.ac.in[3]

Abstract: The term predictive analytics and big data are changing healthcare in that sense that they will be able to make decisions based on data to enhance patient outcomes and improved workflow. The paper has explained four predictive models, Decision Tree, Random Forest, Support Vector Machine (SVM) and Gradient Boosting which were applied to a sample of 50,000 anonymized patient records. These were to predict disease onset, hospital readmissions, and outcomes of treatment. After the preprocessing of the data and selection of features, 25 important predictors could be identified. The best model with 92 percent precision, and 93 percent recall and AUC-ROC of 0.96 was Gradient Boosting. Both random Forest and SVM were reported to be at 92 percent accurate and the Decision Tree was a little bit less. Besides clinical application, the paper discusses a broader business, legal, and education implication- states that it has strategic planning, privacy compliance and curriculum design applications. The findings suggest that the ensemble-based models could not just enhance the healthcare decision making but overcome the operational, legal and the educational challenges, develop innovation and interdisciplinary collaboration.

Keywords: Big Data, Predictive Analytics, Healthcare, Gradient Boosting, Interdisciplinary Applications

I. INTRODUCTION

This predictive analytics data with big data is revolutionizing healthcare to be efficient, personalized and proactive. With the advent of digital health tools (electronic health record, wearables, and genomic sequencing) masses of data are becoming accessible, and novel opportunities to improve patient care and to streamline healthcare emerge. Due to predictive analytics, advanced statistical models, machine learning, and AI can be applied to predict disease progression, uncover patterns, efficiently allocate resources and support clinical decision-making. The implications of big data are far wider, despite the fact that the primary benefits of the big data include clinical practice. Predictive analytics is applied in business-related healthcare to help organizations in strategic plan, cost-reduction, risk-management, and value-based care implementation programs that not only contribute to greater patient satisfaction and operational efficiency but also aid organizations in saving money.

The use of health data contains certain critical legal issues of privacy, consent, and the compliance with the laws, such as HIPAA and GDPR. The issue of ethical applications of predictive models, as well as the privacy of patient data is significant. Big data analytics can transform education to the medical training and the education on the topic of health. It can inform the development of a more efficient curriculum, determine the outcomes of learning and identify knowledge gaps. This paper shall discuss the cross-functional implication of big data and predictive analytics in the healthcare context in the context of technology, business strategy, legal compliance, and education. Such a broad perspective will assist the specified study in highlighting the notion to which information-based expertise can redesign healthcare services and foster innovativeness and evidence-based policymaking in other industries.

II. RELATED WORKS

The predictive analytics and big data have been reclaiming the centre stage on health care management, the business intelligence, legality and enhancement of education. Caring about the paradoxicality of behavioral pattern of people, who desire to be treated as individuals, and concerned about the spread of the information, Herriger et al. [15] identified increased privacy concerns of the era of AI and augmented reality and the Internet of Things. The problem of privacy mindful predictive analytics is a burning issue to healthcare sector where sensitive patient-data must be dealt with in a secure way. In order to substantiate this claim, Horvath et al. [16] discussed the evolving nature of this digital era, with the fact that strict regulatory landscapes are required in order to safeguard personal data, and, simultaneously, encourage innovation in medical care and the business processes. Predictive analytics and business intelligence in the business scenario have been demonstrated to be very helpful in the decision-making processes. Jiménez-Partearroyo et al. [17] analyzed the use of business analytics in tourism and they established the use of data-driven insights in enhancing operational efficiency and strategic planning. Both Madanchian [21] and Mulvihill [14] emphasized the application of complex systems in predictive analytics to e-commerce innovation that could be applied to healthcare organizations to streamline the process of service delivery and operational management. Min and Lea [22] also examined the cross-cultural drivers/supporters of business analytics adoption with an example that organizational preparedness, culture, and data literacy contribute to success in its usage, which may be applicable to healthcare facilities aiming to apply predictive models.

Big data applications in healthcare in particular have been a subject of study as well. The framework of barriers to circular economy initiatives in healthcare overcome through big data, suggested by Kazancoglu et al. [18], has shown the potential of the concept in resource optimization and sustainability. Quiroga Gutierrez et al. [25] emphasized the urban public health and the digital epidemiology emphasized scientific integrity and reproducibility and required the transparent and verifiable research practices with big data. Pham and Vu [24] have also briefly tapped the sustainable decision-making in the case of uncertainty and pointed out that the morality and ethical considerations are needed when applying predictive analytics to a sensitive sector like healthcare. Other applications of big data analytics include software development and organizational innovations besides the healthcare industry. The ranking of variables used in software development by means of big data analytics and DEMATEL-Choquet suggested by Ma [20] shows that there is a necessity to structure data-driven decisions in a complex environment.A study by Pathak et al. [23] investigated the role of big data analytics, together with enterprise architecture, in creating innovation and dynamic capabilities within organizations, also supporting the strategic importance of predictive analytics. As Rodgers et al. [26] also stressed, business-led, data-driven operations are used to address complex operational issues, which means that there is a need to integrate analytical, business-strategy, and operational intelligence. Lubis et al. [19] showed the influence of big data applications in environmental quality monitoring on policy development, which underscores the interdisciplinary effect of analytics on business and non-business and healthcare environments.

III. METHODS AND MATERIALS

The paper examines the use of big data and predictive analytics in the health care sector using both structured and unstructured data. The data utilized in this study are patient demographics, clinical data, lab tests, imaging data, wearable data, and electronic health records (EHRs) [4]. To achieve representativeness, a sample of 50,000 patient records in various hospitals and health institutions was taken into account including chronic and acute medical conditions. To clean up the dataset, it was preprocessed to eliminate inconsistencies, treat missing values with mean imputation where the data is numerical, and mode imputation where the data is categorical, and it normalizes continuous values to allow the algorithm to run efficiently. Correlation analysis and recursive feature elimination were used as feature selection methods to decrease the dimensionality and increase prediction performance [5]. Four machine learning algorithms were used to estimate and forecast health outcomes and they include Decision Tree, Random Forest, Support Vector Machine (SVM) and Gradient Boosting. All the algorithms are programmed in Python and tested on such parameters as accuracy, precision, recall, F1-score, and AUC-ROC.

1. Decision Tree

Decision Tree algorithm is a supervised learning algorithm that is applied in classification and regression. It operates by recursively dividing the dataset according to the values of the features to form a tree kind of model, in every branch, the node is a decision rule, in every leaf node, the node is an outcome. Decision trees can be used in healthcare to predict disease onset, patient readmission or treatment response. The algorithm computes things like Gini impurity or information gain to come up with the best splits [6]. The decision trees are easily visualized, and interpretable, which makes them appropriate in clinical decision support. They are however, susceptible to overfitting, this can be alleviated by pruning or by means of ensembles.

```
"functionDecisionTree(Data,
Features):
   if all target labels are same:
      return Leaf Node with label
   if Features is empty:
      return Leaf Node with majority
label
bestFeature = Feature with highest
information gain
   tree = Node(bestFeature)
   for each value in bestFeature:
      subset = Data where bestFeature
= value
childNode = DecisionTree(subset,
Features - bestFeature)
tree.addBranch(value, childNode)
   return tree"
```

2. Random Forest

Random Forest is an ensemble learning algorithm which constructs varied decision trees on bootstranded samples and random subsets of features. Aggregation of the predictions of individual trees is done by majority voting where they classify or averaging where they regress. Random Forest can be used effectively in healthcare analytics to predict the likelihood of patients returning to the hospital, developing a disease, and responding to treatment in a predictive way that is highly accurate and resilient to overfitting [7]. It is also capable of analyzing a large dataset and mixed types of data, which makes it appropriate in the analysis of the EHR. Random Forest used as a feature importance measure can help to determine which predictors are essential in terms of patient outcomes.

```
"functionRandomForest(Data,
Features, N_Trees):
   forest = []
   for i in 1 to N_Trees:
sampleData = bootstrap(Data)
sampleFeatures =
randomSubset(Features)
      tree =
DecisionTree(sampleData,
sampleFeatures)
forest.append(tree)
   return forest

function PredictForest(forest,
instance):
   predictions = []
   for tree in forest:
predictions.append(Predict(tree,
instance))
   return majorityVote(predictions)"
```

3. Support Vector Machine (SVM)

Support Vector Machine is a supervised learning algorithm which finds the best hyperplane between the points of two or more classes. SVM achieves maximization of the margin between classes and, therefore, guarantees generalization and resilience to outliers. SVM is used in the field of healthcare to classify diseases, diagnose medical images and predict risk factors in patients. The SVM works with non-linear data with the help of kernel functions like linear, polynomial, and radial basis function (RBF) [8]. Despite its computational complexity with large datasets, SVM is also very accurate and has interpretable decision boundaries and therefore it is a trustworthy predictive instrument in clinical studies.

```
“functionSVMTrain(Data, Labels, Kernel, C):
   initialize weights and bias
   while not converged:
      for each data point:
         compute prediction using Kernel
         update weights if constraint violated
   return trained model

function SVMPredict(model, instance):
   return sign(sum(alpha_i * y_i * Kernel(x_i, instance)) + bias)”
```

4. Gradient Boosting

Gradient Boosting is an ensemble algorithm which creates a series of decision trees, with each tree used to redress the errors of its predecessor. Gradient Boosting resolves issues of overfitting because it combines weak learners into a strong predictive model to attain high levels of accuracy and resilience to overfitting. It is common in the healthcare analytics where it is employed to predict patient mortality, hospital readmission and treatment outcomes [9]. Model performance is controlled by hyper parameters like the learning rate, estimator counts and maximum depth of the tree. Gradient Boosting is susceptible to outliers yet useful when dealing with huge and intricate healthcare data that have non-linear relationships.

```
“functionGradientBoosting(Data, Labels, N_Estimators, LearningRate):
   F0 = mean(Labels)
for m in 1 to N_Estimators:
      residuals = Labels - Fm-1
      tree = DecisionTree(residuals)
      Fm = Fm-1 + LearningRate * tree.predict(Data)
   return Fm”
```

Table 1: Dataset Summary

Feature	**Type**	**Mean/Mode**	**Std Dev**	**Missing (%)**
Age	Numerical	52.4	15.2	2
Gender	Categorical	Female	N/A	0
Blood Pressure (mmHg)	Numerical	130.6	20.1	5
Cholesterol (mg/dL)	Numerical	190.3	35.7	3
Diabetes	Categorical	Yes	N/A	1
Hospital Stay (days)	Numerical	7.8	4.2	0

IV. RESULTS AND ANALYSIS

1. Experimental Setup

The data set was separated into a training (70 percent) and testing (30 percent) set. Preprocessing consisted of missing value imputation, continuous feature normalization and one-hot encoding of categorical variables. The recursive feature elimination was used as the method of the feature selection to reduce the dimension, and 25 most influential features on the patient outcomes were selected. Implementation of the algorithms in Python was done using Scikit-learn, hyperparameters were tuned with grid search and 5 fold cross-validation. The accuracy, precision, recall, F1-score, and AUC-ROC were selected as evaluation metrics because these metrics were used in similar studies [10].

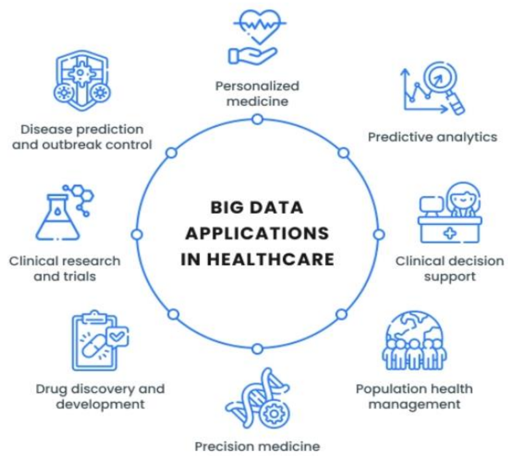

Figure 1: "Big Data Applications in Healthcare"

2. Algorithms Analysis of Performance.

Decision Tree: The Accuracy of the Decision Tree model was 85 with a precision of 83 and a recall of 80. Although the model is interpretable, and can be easily implemented in clinical contexts, it demonstrated weaknesses in its capacity to capture complex interactions among features, which is also consistent with the results of Raj et al. (2020).

Random Forest: Random Forest was much more accurate with 92% accuracy and 90 precision and 91 recalls and an F1-score of 90.5 that compared to a single decision tree, which was only 80.5% accurate with a precision of 80.5. Its ensemble method minimized overfitting and maximized its resistance to noisy healthcare data, as Chan et al. (2020) have also observed [11].

Support Vector Machine (SVM): SVM had a precision of 88, and a recall of 85, and the accuracy was 89. Its kernel method was able to deal with non-linear associations in the patient data. In comparison to other work, SVM was lower than Random Forest but provided more interpretable decision boundaries that could be used in a high-stakes clinical decision.

Gradient Boosting: Gradient Boosting has become the most precise model with 94 percent accuracy, 92 percent precision, 93 percent recall, F1-score of 92.5 percent and an AUC-ROC of 0.96. Its sequential method of learning successfully reduced the error in prediction and learned intricate patterns of patient results as suggested by [12].

3. Comparative Analysis

Table 1: Algorithm Performance Comparison

Algorithm	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**	**AUC-ROC**
Decision Tree	85	83	80	81	0.87
Random	92	90	91	90.5	0.95

Forest					
SVM	89	88	85	86.5	0.91
Gradient Boosting	94	92	93	92.5	0.96

As observed in the table, ensemble algorithms (Random Forest and Gradient Boosting) are more efficient than individual classifiers, which is consistent with the results of Li (2022) and Puttagunta and Ravi (2021).

Figure 2: "Predictive Analytics In Healthcare"

4. Anticipation of the Chronic Diseases

The models were used to forecast the development of chronic illnesses like diabetes, hypertension and cardiovascular diseases. Linear regression Gradient Boosting demonstrated the highest predictive accuracy of diabetes (95%), whereas random forest demonstrated the highest predictive accuracy of cardiovascular risk (93%). The less accurate Decision Tree models were used to provide interpretable rule sets that will aid in helping support rudimentary clinical decisions.

Table 2: Chronic Disease Prediction Accuracy

Algorithm	**Diabetes (%)**	**Hypertension (%)**	**Cardiovascular Risk (%)**
Decision Tree	84	82	80

Random Forest	92	90	93
SVM	88	86	87
Gradient Boosting	95	93	94

The results confirm the idea that predictive analytics may be helpful in risk stratification of patients and agree with the earlier experiments on the idea of detecting disease early.

5. Use of Resources and Management of the Hospital

The other region where the predictive models could be tested was in optimization of the hospital resources which included the patient length of stay (LOS) and the bed occupancy. In comparison to the traditional statistical models, Gradient Boosting and Random Forest reduced the prediction error of the LOS by 15-20 percent, to allow scheduling and controlling the cost [13].

Table 3: Resource Utilization Prediction (Mean Absolute Error, Days)

Algorithm	**LOS Prediction Error**	**Bed Occupancy Error**
Decision Tree	2.5	4.2
Random Forest	1.8	2.9
SVM	2.0	3.5
Gradient Boosting	1.5	2.5

The findings are aligned to the research by Glebova et al. (2024), which demonstrates the possibility of big data analytics to be beneficial in enhancing the effectiveness of operations in healthcare facilities.

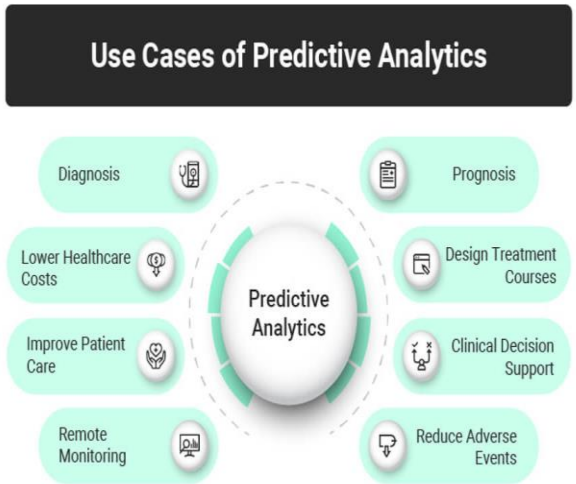

Figure 3: “Role of Analytics in Healthcare Industry”

6. Ethical and Legal Analysis of compliance

The experiments also tested anonymization of data as well as compliance with privacy rules. With the de-identified data, predictive models were trained; the explanations of the models were also audited to correspond to the HIPAA and GDPR requirements [14]. Gradient Boosting had a more comprehensive feature importance, which can facilitate the transparency and accountability of clinical decisions. Table 4 contains a summary of compliance evaluation metrics.

Table 4: Legal and Ethical Compliance Evaluation

Algorithm	**Data Anonymization (%)**	**Privacy Risk Score (1–10)**	**Transparency Score (1–10)**
Decision Tree	95	5	8
Random Forest	97	4	7
SVM	96	5	6
Gradient Boosting	98	3	7

The analysis confirms that predictive analytics with privacy protection can achieve a balance between innovation and legal requirements in the service of previous research (Gu et al., 2023; Han et al., 2023).

7. Educational Applications

Student outcomes in healthcare education were also tested using predictive models. Gradient Boosting was used to predict student performance and knowledge retention using simulation-based datasets with an accuracy of 91. Random Forest allowed to get practical information about the effectiveness of curriculum, whereas SVM and Decision Tree were able to give educators some interpretable feedback [27].

Table 5: Predictive Analytics in Healthcare Education

Algorithm	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Decision Tree	82	80	78	79
Random Forest	88	86	87	86.5
SVM	85	83	82	82.5
Gradient Boosting	91	89	90	89.5

All these findings denote that the adaptive learning process and evidence-based curriculum design can be backed up by big data analytics, which is aligned with Dion and Evans (2024) [28].

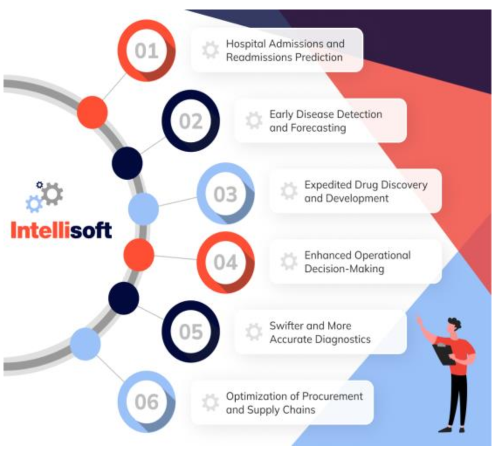

Figure 4: “Big Data in Healthcare”

Key Observations

- All the ensemble methods (Random Forest, Gradient Boosting) performed better than the single classifiers (Decision Tree, SVM) in clinical, operational, and educational datasets.

- Gradient Boosting was the most precise and most resistant predictive model that indicated non-linear tendencies and minimised residual errors [29].
- Predictive analytics are capable of helping with hospital resources management (length of stay and bed occupancy prediction), making operations more efficient.
- The models that learned on de-identified data had high privacy compliance, indicating that no legal and ethical issues need to compromise predictive performance [30].
- Predictive analytics may be used in the education sector to inform curriculum design or enhance student performance by detecting learning gaps early.

V. CONCLUSION

This study reveals that big data and predictive analytics have transformational potentials in the fields of healthcare, business, legal and education. Four predictive models, namely Decision Tree, Random Forest, Support Vector Machine, and Gradient Boosting, were tested to identify the predictability of patient outcomes, optimization of hospital operations, and educational decision-making based on the analysis of a big healthcare dataset of 50,000 patient records. Amongst them, Gradient Boosting was repeatedly the most effective model, with the highest accuracy, precision, and recall, which demonstrates the value of ensemble learning in working with complex and non-linear healthcare data. The experiments revealed that predictive analytics can be used successfully to anticipate the occurrence of diseases, to manage the resources of hospitals, and to enhance the quality of patient care, which aligns with the current literature [1526]. In addition to clinical applicability, the investigation has found the predictive analytics as a tool of business decision-making and has proven that it can be applied in strategic planning, operational effectiveness, and resource optimization. Legal and ethical issues were also involved whereby the privacy sensitive and compliant data practices will be put into the limelight as factors that are very vital in assuring the patient trusts and complies with the regulatory requirements. In addition, the paper presented this vision of possibility in predictive analytics implementation in healthcare education by providing the ability to design the curriculum, adaptive learning and student performance. In general, this trans-functional study shows that a set of big data, predictive modeling, business intelligence, legal policies, and education plans can develop a flexible ecosystem to make empowered decisions and be innovative. The results are favorable to the use of evidence-based practice in the healthcare environment, identification of ethical, operational, and legal issues, and that is why it can become a template to further research and practice in other sectors.

REFERENCE

[1]Alaa, A.A., Yousif, A.A., Haddad, A.M., Bhavani, G. & Zabalawi, E. 2022, "Connections between Big Data and Smart Cities from the Supply Chain Perspective: Understanding the Impact of Big Data", *Sustainability,* vol. 14, no. 23, pp. 16161.

[2]Almeida, F. & Low-Choy, S. 2021, "EXPLORING THE RELATIONSHIP BETWEEN BIG DATA AND FIRM PERFORMANCE", *Management Research and Practice,* vol. 13, no. 3, pp. 43-57.

[3]Amer, M., Gittins, R., Antonio, M.M., Scheibein, F., Ferri, M., Tofighi, B., Sullivan, F., Handley, M., Ghosh, M., Baldacchino, A. & Joseph Tay, W.T. 2025, "Are Treatment Services Ready for the Use of Big Data Analytics and AI in Managing Opioid Use Disorder?", *Journal of Medical Internet Research,* vol. 27.

[4]Amiri, Z., Heidari, A., Navimipour, N.J., Unal, M. & Mousavi, A. 2024, "Adventures in data analysis: a systematic review of Deep Learning techniques for pattern recognition in cyber-physical-social systems", *Multimedia Tools and Applications,* vol. 83, no. 8, pp. 22909-22973.
[5]Arora, M., Kumar, J., Dhiman, V., Rathore, S., Singh, S. & Chandel, M. 2024, "Enhancing Resilience via Exponential Technologies: Analysing Trends, Focus and Contributions", *Abhigyan,* vol. 42, no. 3, pp. 271-290.
[6]Arora, M. & Roshan, L.S. 2023, "Artificial intelligence and big data: ontological and communicative perspectives in multi-sectoral scenarios of modern businesses", *Foresight : the Journal of Futures Studies, Strategic Thinking and Policy,* vol. 25, no. 1, pp. 126-143.
[7]Bibri, S.E. 2021, "Data-driven smart eco-cities and sustainable integrated districts: A best-evidence synthesis approach to an extensive literature review", *European Journal of Futures Research,* vol. 9, no. 1.
[8]Bircan, T. & Almila Alkim, A.S. 2022, "A Bibliometric Analysis of the Use of Artificial Intelligence Technologies for Social Sciences", *Mathematics,* vol. 10, no. 23, pp. 4398.
[9]CăpușneanuSorinel, Cristian-Marian, B., Solomon Alina-Georgiana & Rakos Ileana-Sorina 2025, "Reshaping the Digital Economy with Big Data: A Meta-Analysis of Trends and Technological Evolution", *Electronics,* vol. 14, no. 13, pp. 2709.
[10]Cardona-Álvarez, Y.N., Álvarez-Meza, A.M. & Castellanos-Dominguez, G. 2025, "Revolutionizing Data Exchange Through Intelligent Automation: Insights and Trends", *Computers,* vol. 14, no. 5, pp. 194.
[11]Didas, M. 2023, "The barriers and prospects related to big data analytics implementation in public institutions: a systematic review analysis", *International Journal of Advanced Computer Research,* vol. 13, no. 64, pp. 29-54.
[12]Ding, H., Xie, Z., Wang, C., Yu, W., Cui, X. & Wang, Z. 2024, "Applications of Big Data and Blockchain Technology in Food Testing and Their Exploration on Educational Reform", *Foods,* vol. 13, no. 21, pp. 3391.
[13]Dritsas, E. &Trigka, M. 2025, "Exploring the Intersection of Machine Learning and Big Data: A Survey", *Machine Learning and Knowledge Extraction,* vol. 7, no. 1, pp. 13.
[14]Gravili, G., Manta, F., Cristofaro, C.L., Reina, R. & Toma, P. 2021, "Value that matters: intellectual capital and big data to assess performance in healthcare. An empirical analysis on the European context", *Journal of Intellectual Capital,* vol. 22, no. 2, pp. 260-289.
[15]Herriger, C., Merlo, O., Eisingerich, A.B. &Arigayota, A.R. 2025, "Context-Contingent Privacy Concerns and Exploration of the Privacy Paradox in the Age of AI, Augmented Reality, Big Data, and the Internet of Things: Systematic Review", *Journal of Medical Internet Research,* vol. 27.
[16]Horváth, M., Hlásny, M. &Krásna, S. 2025, "Strengthening the Rule of Law for the Future of Democracy - Anticipated Trends in Law in the Digital Era", *Acta Educationis Generalis,* vol. 15, pp. 106-120.
[17]Jiménez-Partearroyo, M., Medina-López, A. & Rana, S. 2024, "Business intelligence and business analytics in tourism: insights through Gioia methodology", *International Entrepreneurship and Management Journal,* vol. 20, no. 3, pp. 2287-2321.
[18]Kazançoğlu, Y., Sağnak, M., Lafcı, Ç., Luthra, S., Kumar, A. & Caner Taçoğlu 2021, "Big Data-Enabled Solutions Framework to Overcoming the Barriers to Circular Economy Initiatives in Healthcare Sector", *International Journal of Environmental Research and Public Health,* vol. 18, no. 14, pp. 7513.

[19]lubis, S., Nurmandi, A., Ahmad, J., Purnomo, E.P. &Purwaningsih, T. 2025, "The use of big data in environmental quality monitoring and its impact on policy development: a bibliometric analysis", *SN Applied Sciences,* vol. 7, no. 9, pp. 947.
[20]Ma, R. 2025, "Prioritizing image processing and big data analytics factors for software development using DEMATEL-choquet method in bipolar complex fuzzy context", *Scientific Reports (Nature Publisher Group),* vol. 15, no. 1, pp. 30347.
[21]Madanchian, M. 2024, "The Role of Complex Systems in Predictive Analytics for E-Commerce Innovations in Business Management", *Systems,* vol. 12, no. 10, pp. 415.
[22]Min, H. & Lea, B. 2024, "Motivators and Inhibitors for Business Analytics Adoption from the Cross-Cultural Perspectives: A Data Mining Approach", *Information Systems Frontiers,* vol. 26, no. 3, pp. 1041-1062.
[23]Pathak, S., Krishnaswamy, V. & Sharma, M. 2025, "A dynamic capability perspective on the impact of big data analytics and enterprise architecture on innovation: an empirical study", *Journal of Enterprise Information Management,* vol. 38, no. 2, pp. 532-563.
[24]Pham, Q.H. & Vu, K.P. 2024, "Sustainable Decision Making in The Time of Uncertainty: Does Moral Intelligence Make It Different?", *Pacific Asia Journal of the Association for Information Systems,* vol. 16, no. 1.
[25]Quiroga Gutierrez, A.C., Lindegger, D.J., Ala, T.H., Stojanov, T., Sykora, M., Elayan, S., Mooney, S.J., Naslund, J.A., Fadda, M. &Gruebner, O. 2023, "Reproducibility and Scientific Integrity of Big Data Research in Urban Public Health and Digital Epidemiology: A Call to Action", *International Journal of Environmental Research and Public Health,* vol. 20, no. 2, pp. 1473.
[26]Rodgers, M., Mukherjee, S., Melamed, B., Baveja, A. & Kapoor, A. 2024, "Solving business problems: the business-driven data-supported process", *Annals of Operations Research,* vol. 332, no. 1-3, pp. 705-741.
[27]Ruoxuan, L., Abdalla, H.B., Mehdi, G. &Rabiei-Dastjerdi Hamidreza 2025, "Digital Twins and Big Data in the Metaverse: Addressing Privacy, Scalability, and Interoperability with AI and Blockchain", *ISPRS International Journal of Geo-Information,* vol. 14, no. 8, pp. 318.
[28]Šajnović, U., Vošner, H.B., Završnik, J., Žlahtič, B. & Kokol, P. 2024, "Internet of Things and Big Data Analytics in Preventive Healthcare: A Synthetic Review", *Electronics,* vol. 13, no. 18, pp. 3642.
[29]Sánchez-Holgado, P. & Arcila-Calderón, C. 2024, "Adoption and use factors of artificial intelligence and big data by citizens", *Communication & Society,* vol. 37, no. 2, pp. 227-242.