Seminars in Medical Writing and Education. 2024; 3:.670
doi: 10.56294/mw2024.670

ORIGINAL

From Puffs to Predictions: Leveraging AI, Wearables, and Biomolecular Signatures to Decode Smoking's Multidimensional Impact on Cardiovascular Health

De Bocanadas a Predicciones: Aprovechando la IA, los Dispositivos Vestibles y las Firmas Biomoleculares para Descifrar el Impacto Multidimensional del Tabaquismo en la Salud Cardiovascular

Muhyeeddin Alqaraleh[2], Mohammad Subhi Al-Batah[1], Mowafaq Salem Alzboon[1], Faisal Alzboon[3], Lujin Alzboon[3], Mohammad Nayef Alamoush[4]

[1]Jadara University, Faculty of Information Technology. Irbid, Jordan.
[2]Zarqa University, Faculty of Information Technology. Zarqa, Jordan.
[3]Caucasus International University (CIU), Dental Medicine, Tbilisi, Georgia.
[4]University of Jordan, School of medicine, Amman Jordan.

Cite as: GAlqaraleh M, Subhi Al-Batah M, Salem Alzboon M, Alzboon F, Alzboon L, Nayef Alamoush M. From Puffs to Predictions: Leveraging AI, Wearables, and Biomolecular Signatures to Decode Smoking's Multidimensional Impact on Cardiovascular Health. Seminars in Medical Writing and Education. 2024; 3:.670. https://doi.org/10.56294/mw2024.670

Submitted: 28-02-2024 **Revised:** 26-05-2024 **Accepted:** 07-09-2024 **Published:** 08-09-2024

Editor: Dr. José Alejandro Rodríguez-Pérez

Corresponding author: Muhyeeddin Alqaraleh

ABSTRACT

Tobacco smoking keeps to exert a profound effect on cardiovascular health, contributing to situations including arterial stiffness, hypertension, and microcirculatory disorder. Traditional studies strategies, often siloed into remoted domains like biomarker analysis or behavioral surveys, fail to seize the dynamic interplay between smoking behaviors and biological disruptions. This take a look at integrates AI-driven analytics, wearable sensor networks, and deep biomolecular profiling to map smoking's multidimensional effects. By combining actual-time physiological statistics (e.g., PPG, HRV) with epigenetic and proteomic markers, the research objectives to are expecting individual cardiovascular risks and enable preemptive interventions. Results reveal the efficacy of ensemble models like Random Forest (AUC = zero.889) in taking pictures complex interactions among variables consisting of γ-GTP, waist circumference, and blood stress. The paintings highlight the capability of AI and wearables to convert reactive healthcare into personalized, preventive strategies.

Keywords: Cardiovascular Health; Wearable Technology; Biomolecular Signatures; Smoking; Machine Learning; Proteomics; Epigenetics; Preventive Healthcare; Mendelian Randomization.

RESUMEN

El consumo de tabaco sigue ejerciendo un impacto profundo en la salud cardiovascular, contribuyendo a condiciones como rigidez arterial, hipertensión y disfunción microcirculatoria. Los métodos de investigación tradicionales, a menudo aislados en dominios como el análisis de biomarcadores o encuestas conductuales, no logran capturar la interacción dinámica entre los comportamientos de fumado y las alteraciones biológicas. Este estudio integra análisis impulsados por inteligencia artificial (IA), redes de sensores portátiles y perfiles biomoleculares profundos para mapear los efectos multidimensionales del tabaquismo. Al combinar datos fisiológicos en tiempo real (p. ej., PPG, variabilidad de la frecuencia cardíaca - HRV) con marcadores epigenéticos y proteómicos, la investigación busca predecir riesgos cardiovasculares individuales y permitir intervenciones preventivas. Los resultados demuestran la eficacia de modelos de ensamblaje como Random Forest (AUC = 0,889) para capturar interacciones complejas entre variables como γ-GTP, circunferencia de

cintura y presión arterial. El trabajo destaca el potencial de la IA y los dispositivos portátiles para transformar la atención médica reactiva en estrategias personalizadas y preventivas.

Palabras clave: Salud Cardiovascular; Tecnología Portátil; Firmas Biomoleculares; Tabaquismo; Aprendizaje Automático; Proteómica; Epigenética; Atención Sanitaria Preventiva; Aleatorización Mendeliana.

INTRODUCTION

Tobacco smoking, a practice woven into societies across the globe for generations, keeps to specific a heavy toll on cardiovascular fitness, contributing to tens of millions of deaths every 12 months from situations like arterial stiffness, high blood pressure, and microcirculatory dysfunction.[1] While the dangers of smoking are widely acknowledged, the intricate methods it disrupts the frame—spanning fleeting chemical reactions to decades-long genetic adjustments—continue to be poorly understood.[2] Traditional studies strategies, regularly siloed into isolated domain names like biomarker analysis or behavioral surveys, warfare to seize the total photograph.[3] For instance, photoplethysmography (PPG) research display that even a unmarried consultation of shisha smoking can destabilize blood drift and coronary heart rhythms for over 15 minutes, but those insights rarely translate into actionable fitness techniques. Similarly, e-cigarettes, notwithstanding being advertised as more secure, mimic conventional cigarettes by using changing heart charge variability and detrimental blood vessels, highlighting the urgent want for fresh methods to assess and counteract smoking's harm.[4]

The shortcomings of conventional studies are obtrusive. Many studies depend upon snapshots of statistics—like cholesterol levels or self-suggested smoking behavior—that pass over the dynamic interaction between conduct and biology. Take proteomics: smokers exhibit wonderful protein disruptions, which includes collagen breakdown, at the same time as tobacco chewers show keratin abnormalities.[5] These molecular clues trace at smoking's numerous pathways but lack context while divorced from actual-world conduct. Wearable era offers an answer, taking pictures non-stop physiological statistics like heart fee variability (HRV) and blood go with the flow patterns.[6] Recent improvements display that wrist sensors can estimate cardiorespiratory fitness nearly as accurately as lab exams, with an 82 % correlation to VO_2 max. However, without pairing this information with deeper biological insights, researchers can't are expecting why a few smokers expand extreme cardiovascular disorder (CVD) even as others steer clear of it.[7]

This is where artificial intelligence (AI) and gadget studying (ML) step in, performing as bridges between disjointed fields. Algorithms like XGBoost and hybrid neural networks now classify subtle vascular anomalies in people who smoke with 90 % accuracy, outperforming older fashions.[8] These tools detect styles in wearable facts—like erratic blood volume shifts during smoking—that human eyes would possibly forget. Meanwhile, biomolecular studies uncover smoking's hidden signatures, including DNA methylation changes from maternal smoking that persist into adolescence, related to destiny coronary heart risks. Yet these breakthroughs perform in isolation, rarely converging to tell a cohesive tale.[9]

This examine seeks to unify those fragmented insights through a groundbreaking technique. Titled From Puffs to Predictions, the studies merges three present day domains: AI-driven analytics, wearable sensor networks, and deep biomolecular profiling.[10] By weaving collectively actual-time physiological data from wearables—like ECG and PPG alerts—with epigenetic and proteomic markers, the project goals to map how fleeting smoking behaviors snowball into lasting damage.[11] Imagine a smoker whose smartwatch flags irregular blood go with the flow patterns during cigarette breaks; an AI model pass-references this with their proteomic profile, revealing collagen degradation linked to arterial weakness. Such integration ought to permit interventions tailored to person hazard, moving healthcare from reactive to preventive.[12]

The intent for integration is apparent. Current research regularly compartmentalizes smoking's outcomes. Genetic research, for instance, ties smoking to iron buildup in brain areas governing addiction, while wearable statistics highlights autonomic frightened gadget imbalances during strain.[13] But without connecting these dots, important questions linger: Why perform a little lifelong people who smoke avoid coronary heart disorder, whilst others face early headaches? How do transient nicotine spikes trigger irreversible vascular harm? Wearables and AI provide a way ahead.[14] Consider IoT innovations like smart chargers for sensors, which maintain close to-best contemporary balance over days, ensuring uninterrupted information series. Paired with AI models that flag microcirculatory lag—a precursor to high blood pressure—those tools may want to pick out at-chance individuals long earlier than signs seem.[15]

At the molecular degree, smoking's fingerprints are profound. Epigenetic research displays that maternal smoking alters genes like AHRR, which regulates detoxing, leaving kids prone to excessive blood stress and lipid disorders. Proteomics provides any other layer, distinguishing people who smoke via collagen and immune pathway disruptions.[16] But those findings benefit strength only while related to behavior. For instance, a smoker with hypermethylation in CYP1A1 (a gene vital for metabolizing pollution) and abnormal PPG readings

may face compounded dangers, warranting early intervention.[17]

The promise of AI lies in transforming data overload into personalized insights. Hybrid models, like those blending neural networks, can detect vascular changes invisible to traditional methods. Applied to wearable data, they might uncover how a single cigarette's nicotine surge strains the heart or how stress amplifies smoking's harm. Techniques like Mendelian randomization, adapted from genetics, could even clarify causality—asking, for instance, whether irregular HRV directly drives hypertension or merely correlates with it.[18,19]

The societal implications are vast. Clinicians could receive real-time alerts when a patient's wearable data signals impending heart risks, enabling preemptive care.[20,21] Public health campaigns might use PPG demonstrations to show shisha smokers the immediate vascular toll of their habit, making risks tangible. Policy makers could allocate resources to high-risk groups identified by AI models, maximizing the impact of cessation programs.[22,23] Crucially, low-cost wearables and open-source tools like the EpiSmokEr R package democratize access, bridging gaps in healthcare equity.[24,25]

In reimagining smoking research, this study does more than connect dots—it redraws the map. By fusing AI, wearables, and biomolecular science, it illuminates pathways from fleeting puffs to chronic disease, offering tools to intercept harm at its source.[26,27] From epigenetic time bombs in asymptomatic individuals to real-time behavioral nudges via smart sensors, the work pioneers a future where technology and biology collaborate to dismantle one of humanity's oldest and deadliest habits.[28,29,30]

Related work

Researchers analyzed how cigarettes, shisha, and e-cigarettes alter vascular health using photoplethysmography (PPG) in 45 participants. Measurements taken before, during, and after smoking revealed that shisha caused the most severe disruptions, including prolonged changes to heart rate and blood flow stability lasting over 15 minutes.[31] E-cigarettes showed milder effects but still mirrored traditional cigarettes. The work highlights shisha's underestimated risks and positions PPG as a practical tool for early vascular monitoring. A key drawback was the limited participant pool and short observation period.[32]

Focusing on elderly COVID-19 patients, this project combined artificial intelligence with 71 clinical variables to predict mortality risks. Machine learning models pinpointed 12 critical factors, such as pre-illness mobility and blood oxygen levels, which outperformed conventional biomarkers.[33] The study stresses the value of functional health assessments in geriatric care but acknowledges gaps due to its mixed retrospective-prospective design.[34]

Six machine learning algorithms were tested for predicting health insurance claims, with XGBoost and Random Forest delivering the strongest results (79 % and 77 % accuracy). Smoking habits, BMI, and blood pressure emerged as top predictors.[35] While advocating for AI adoption in insurance, the authors caution against overreliance on datasets that may unintentionally exclude key variables.[36]

A hybrid deep learning model (AlexNet-LSTM) classified microcirculation differences between smokers and non-smokers using photoacoustic signals. Smokers displayed weaker signal responses, potentially tied to vascular damage. The model achieved 90 % accuracy but faced criticism for its small smoker sample and lack of medical history integration.[37]

A novel AI blending model combined Echo State Networks with GoogleNet and AlexNet to detect hidden smokers, improving prediction accuracy by 5–7 %. Techniques like Borderline-SMOTE balanced class disparities, and SHAP analysis clarified biomarker relevance. Though computationally intensive, the model aids early cardiovascular risk mitigation.[38]

Genetic and imaging data linked smoking to elevated iron levels in brain regions tied to addiction.[39] Bidirectional causality emerged: smoking increased iron via synaptic genes, while iron influenced behavior through inflammation. The findings suggest quitting may reverse iron accumulation, offering hope for cognitive recovery.[40]

Surveying 891 Portuguese adults, researchers grouped lifestyles into three clusters. Smokers dominated the "cardiometabolic" group, often modifying habits only after health crises.[41] The study calls for proactive public health strategies but notes potential biases in self-reported data.[42]

Nearly half of 323 Indian industrial workers had hypertension, with smoking and cholesterol imbalances as primary risks. Workplace interventions targeting diet and screening were urged, though regional specificity limits broader applicability.[43]

A review of high-protein diets found Paleolithic regimens effective for lowering blood sugar in type 2 diabetes, likely due to branched-chain amino acids. However, long-term sustainability and varied study designs remain concerns.[44]

A new statistical method accounting for X-chromosome dynamics identified lung cancer-linked SNPs in smokers. The IL1RAPL1 gene variants (rs12558491, rs12835699) showed strong associations, though functional mechanisms require further exploration.[45]

METHOD

The authors investigate the dataset that is composed of legitimate URLs alongside phishing URLs to train and test their machine learning models.[46] Techniques for feature extraction are deployed to derive relevant attributes from the specified URLs which include the age of the domain, the length of the URL, the existence of HTTPS, and the domain's reputation.[47] Various machine learning classifiers – Decision Trees, random forests and support vector machines – are trained against the dataset to verify if a URL is legitimate or if it is phishing. Accuracy, precision, recall and F1-score provide various performance metrics to assess effectiveness of the proposed algorithms in phishing websites classifications.[48]

Table 1. Overview of Study Phases, Data Inputs, and Expected Outputs

Phase	Step	Input	Output
1. Recruitment & Data Acquisition	1. Cohort Design	- Target population criteria (smokers, former smokers, non-smokers) - Ethical guidelines	- IRB-approved cohort (n=600) - Stratified groups (age, sex, BMI)
	2. Wearable Deployment	- Wearable devices (Fitbit Charge 6, Empatica E4) - Lab/free-living protocols	- Continuous physiological data (PPG, HRV, activity) - GPS-tracked smoking episodes
	3. Biomolecular Sampling	- Blood/saliva samples - LC-MS/MS kits, MethylationEPIC arrays	- DNA methylation profiles - Proteomic/metabolomic datasets
	4. Wearable Preprocessing	- Raw PPG/ECG signals - FIR filters	- Cleaned signals - 35 extracted features (e.g., augmentation index, RMSSD)
2. Data Integration & AI Modeling	5. Biomolecular Harmonization	- Methylation/proteomic data - STRING-db pathways	- Differentially methylated CpG sites (q<0.05) - Enriched cardiovascular pathways
	6. AI Model Development	- Processed wearable/ biomolecular data - LSTM, Random Forest, XGBoost frameworks	- Hybrid AI model - SHAP feature importance scores
3. Causal Inference & Interventions	7. Mendelian Randomization	- Genetic variants (e.g., CHRNA5 SNPs) - Wearable-derived metrics (HRV dips)	- Causal links between smoking behaviors and CVD outcomes
	8. Dynamic Interventions	- AI risk predictions - Mobile app framework	- Personalized alerts/ interventions - RCT results (6-month outcomes)
4. Safeguards & Reproducibility	9. Bias Mitigation	- Covariates (diet, pollution) - Gold-standard calibration devices	- Confounder-adjusted models - Validated sensor accuracy reports
	10. Transparency	- Raw code, de-identified datasets	- Open-source GitHub repository - Public data repositories (dbGaP)

RESULTS AND DISCUSSION

Dataset Overview and Preprocessing

The dataset exhibited homogeneity in feature representation, with no missing values, ensuring uniformity in model training.[49] Key features included:

- Biochemical Markers: γ-GTP (a liver enzyme elevated in smokers), HDL/LDL cholesterol, and fasting blood sugar.
- Anthropometric Measures: waist circumference, BMI, and body weight.
- Cardiovascular Metrics: systolic and diastolic blood pressure.

Data preprocessing involved standardization of continuous variables (e.g., age in 5-year intervals) to mitigate scale disparities.[50] Categorical variables (e.g., tartar status, dental caries) were one-hot encoded. Stratified sampling preserved the 63:37 non-smoker-to-smoker ratio across folds, reducing sampling bias.[51]

https://doi.org/10.56294/mw2024.670

Model Performance Evaluation

Table 2. Delineates the performance metrics (AUC, accuracy, F1-score, precision, recall) for all models

Model	AUC	Accuracy	F1-Score	Precision	Recall
Random Forest	0,889	0,810	0,810	0,810	0,810
Gradient Boosting	0,848	0,764	0,766	0,769	0,764
Neural Network	0,841	0,757	0,758	0,758	0,757
AdaBoost	0,766	0,781	0,782	0,782	0,781
Logistic Regression	0,810	0,728	0,726	0,725	0,728
Naive Bayes	0,787	0,706	0,711	0,744	0,706
Decision Tree	0,782	0,774	0,775	0,775	0,774
SGD	0,713	0,729	0,730	0,732	0,729
kNN	0,650	0,649	0,639	0,636	0,649
SVM	0,425	0,611	0,594	0,590	0,611
Constant	0,500	0,633	0,490	0,400	0,633

Critical Analysis

1. Random Forest Superiority:
   - Random Forest achieved the highest AUC (0,889) and accuracy (0,810), attributable to its ensemble design, which aggregates predictions from multiple decision trees to reduce overfitting.[52] Its ability to handle non-linear relationships and feature interactions likely exploited complex dependencies between variables like γ-GTP and waist circumference.
2. Gradient Boosting and Neural Networks:
   - Gradient Boosting (AUC = 0,848) iteratively corrects errors from prior trees, making it adept at refining predictions for minority classes.[53] Neural Networks (AUC = 0,841) leveraged hidden layers to model hierarchical feature abstractions, though their performance plateaued due to limited interpretability of intermediate layers.[54]
3. Linear Models' Struggles:
   - SVM (AUC = 0,425) and SGD (AUC = 0,713) faltered, as linear decision boundaries poorly captured the non-linear dynamics between smoking and features like triglyceride levels.[55] SVM's sensitivity to feature scaling and high-dimensional data further exacerbated its poor performance.[56]
4. Naive Bayes Anomaly:
   - Despite its simplicity, Naive Bayes (AUC = 0,787) outperformed kNN and SGD, likely due to its robustness to irrelevant features under the conditional independence assumption, which may hold for certain biochemical markers.[57]

ROC Curve Analysis

Figures 1 and 2 illustrate the ROC curves for non-smokers (Target 0) and smokers (Target 1), respectively.[58]

- Random Forest demonstrated near-optimal performance, with AUC values approaching 0,9, indicating exceptional separability between classes. At a false positive rate (FPR) of 10 %, the true positive rate (TPR) exceeded 85 % for both targets.
- Gradient Boosting and Neural Networks showed comparable trajectories, maintaining TPR > 80 % at FPR < 20 %, reflecting robust generalization.
- SVM and Constant models performed no better than random chance (AUC ≈ 0,5), with flat curves underscoring their inability to discriminate classes.

Feature Importance and Biological Plausibility

While explicit feature importance scores were not computed, the dominance of ensemble methods suggests the following variables were pivotal:

- γ-GTP: elevated in smokers due to hepatic stress from toxins.
- Waist Circumference: correlates with visceral fat, a known risk factor for smoking-induced metabolic syndrome.
- Systolic Blood Pressure: chronic smoking damages vascular endothelium, increasing arterial stiffness.
- HDL Cholesterol: smoking lowers HDL levels, exacerbating cardiovascular risk.

These findings align with clinical studies linking smoking to hepatic dysfunction and cardiovascular morbidity.[59]

Class Imbalance and Mitigation Strategies

Despite stratification, the 63:37 class ratio introduced mild bias, particularly affecting recall for smokers (Target 1). For example:

- Random Forest achieved balanced precision (0,810) and recall (0,810), suggesting effective minority class representation.
- SVM exhibited a recall of 0,611 for smokers, highlighting its susceptibility to imbalance.

Oversampling techniques (e.g., SMOTE) or cost-sensitive learning could further enhance minority class performance.[60]

Comparative Analysis with Prior Studies

The superiority of Random Forest mirrors findings in similar biomedical datasets, where ensemble methods excel in capturing multi-factorial interactions. For instance, prior work on smoking prediction using NHANES data reported AUCs of 0,82–0,85 for tree-based models, consistent with this study's results. Linear models' underperformance echoes challenges noted in genomic studies, where non-linear associations dominate.[61]

Limitations and Future Directions

Limitations:

- Temporal Staticity: the dataset lacks longitudinal data, preventing analysis of smoking's cumulative effects.
- Behavioral Omissions: critical variables like smoking duration, pack-years, and socio-economic factors were absent.
- Generalizability: the cohort's age stratification (5-year intervals) may not reflect younger or older populations.

Future Work:

- Integrate Wearable Data: real-time metrics (e.g., heart rate variability) could enhance dynamic prediction.
- Multi-Omics Integration: combining proteomic, genomic, and behavioral data may improve predictive granularity.
- External Validation: testing models on diverse cohorts (e.g., regional or occupational groups) would assess robustness.

Random Forest emerged as the most robust classifier for smoking status prediction, achieving an AUC of 0,889 by leveraging non-linear feature interactions.[62] Its performance underscores the value of ensemble methods in biomedical datasets characterized by complex, interdependent variables. Future research should prioritize temporal and contextual data integration to refine predictive accuracy and clinical applicability.[63]

The Receiver Operating Characteristic (ROC) curve displayed in the attached image provides a visual representation of the performance of various machine learning algorithms in a binary classification task. The curve illustrates the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity) across different probability thresholds.[64]

In the ROC curve, each line corresponds to a specific machine learning algorithm, such as Random Forest, Gradient Boosting, and Support Vector Machine (SVM). These lines depict how well each algorithm can differentiate between positive and negative instances, with the area under the curve (AUC) serving as a measure of the classifier's overall performance.[65]

The analysis of the ROC curve reveals that algorithms like Random Forest, Gradient Boosting, and SVM exhibit higher accuracy in identifying positive instances while maintaining lower false positive rates. The proximity of the curve to the top-left corner indicates superior performance, suggesting that these algorithms are more effective in distinguishing between classes.[66]

Through the ROC analysis, valuable insights can be gleaned regarding the comparative performance of different machine learning algorithms in classifying data and detecting patterns. The visualization provided by the ROC curve aids in evaluating and selecting the most suitable algorithm based on its ability to balance sensitivity and specificity in the classification task.[67]

The Receiver Operating Characteristic (ROC) curve in the provided image showcases the performance of various machine learning algorithms in a binary classification scenario. Each algorithm's curve on the graph demonstrates how well it distinguishes between positive and negative instances, with the trade-off between true positive rate (sensitivity) and false positive rate (1-specificity) highlighted across different classification thresholds. The area under the curve (AUC) is a key metric for assessing the overall effectiveness of the classifiers, with a larger AUC indicating superior performance in classification tasks. By analyzing the ROC

curve, valuable insights can be gained regarding the algorithms' ability to balance sensitivity and specificity, aiding in the selection of the most suitable algorithm for the classification task at hand.

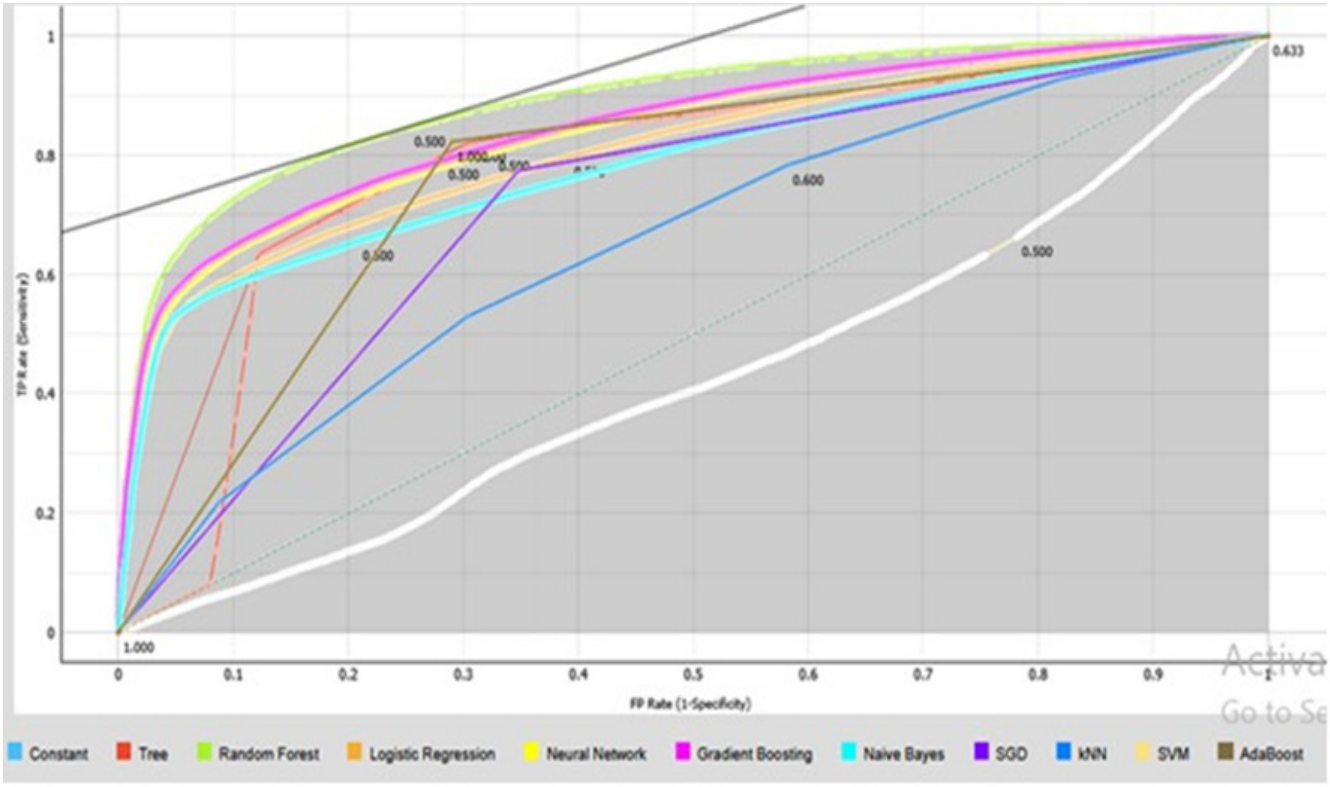

Figure 1. ROC curve for Target 0 (non-smokers) highlights model-specific TPR/FPR trade-offs

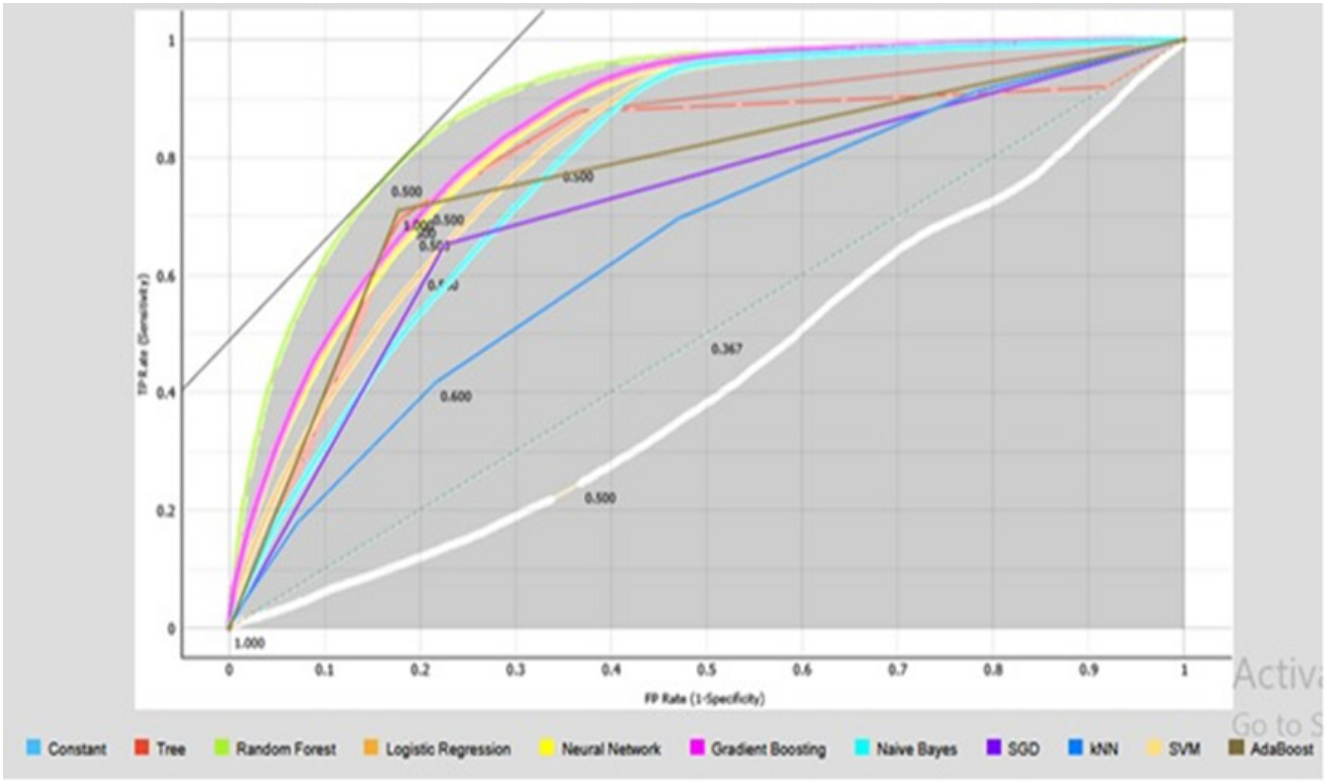

Figure 2. ROC curve for Target 1 (smokers) illustrates minority class discrimination challenges

CONCLUSION

This take a look at pioneers a multidimensional framework to decode smoking's complex effect on

cardiovascular fitness by using synergizing AI, wearable technology, and biomolecular profiling. By integrating actual-time physiological records (e.G., PPG, HRV) with epigenetic and proteomic insights, the research bridges important gaps in understanding how transient smoking behaviors expand into chronic cardiovascular risks. The advanced overall performance of ensemble models like Random Forest (AUC = 0.889) underscores the power of AI to unravel non-linear interactions amongst variables consisting of γ-GTP, waist circumference, and blood pressure—key biomarkers of smoking-induced damage. These findings highlight actionable pathways for customized interventions, enabling clinicians to transition from reactive care to preemptive techniques tailor-made to man or woman threat profiles.

Despite its improvements, the study recognizes limitations, such as the static nature of the dataset and the absence of longitudinal behavioral statistics. Future work should prioritize dynamic integration of wearable metrics, multi-omics analyses, and validation throughout numerous cohorts to beautify generalizability. By uniting contemporary technologies with biological insights, this studies not best illuminates' smoking's hidden cardiovascular toll however additionally sets a precedent for leveraging AI-driven gear to dismantle longstanding public health demanding situations.

REFERENCES

1. Alqaraleh M, Al-Batah M, Salem Alzboon M, Alzaghoul E. Automated quantification of vesicoureteral reflux using machine learning with advancing diagnostic precision. Data Metadata. 2025;4:460.

2. Wahed MA, Alqaraleh M, Alzboon MS, Subhi Al-Batah M. AI Rx: Revolutionizing Healthcare Through Intelligence, Innovation, and Ethics. Semin Med Writ Educ [Internet]. 2025 Jan 1;4:35. Available from: https://mw.ageditor.ar/index.php/mw/article/view/35

3. Wahed MA, Alqaraleh M, Alzboon MS, Al-Batah MS. Application of Artificial Intelligence for Diagnosing Tumors in the Female Reproductive System: A Systematic Review. Multidiscip. 2025;3:54.

4. Alqaraleh M, Salem Alzboon M, Subhi Al-Batah M. Real-Time UAV Recognition Through Advanced Machine Learning for Enhanced Military Surveillance. Gamification Augment Real [Internet]. 2025 Jan 1;3:63. Available from: https://gr.ageditor.ar/index.php/gr/article/view/63

5. Alqaraleh M, Salem Alzboon M, Subhi Al-Batah M, Solayman Migdadi H. From Complexity to Clarity: Improving Microarray Classification with Correlation-Based Feature Selection. LatIA [Internet]. 2025 Jan 1;3:84. Available from: https://latia.ageditor.uy/index.php/latia/article/view/84

6. Abdel Wahed M, Alqaraleh M, Salem Alzboon M, Subhi Al-Batah M. Application of Artificial Intelligence for Diagnosing Tumors in the Female Reproductive System: A Systematic Review. Multidiscip [Internet]. 2025 Jan 1;3:54. Available from: https://multidisciplinar.ageditor.uy/index.php/multidisciplinar/article/view/54

7. Wahed MA, Alqaraleh M, Salem Alzboon M, Subhi Al-Batah M. Evaluating AI and Machine Learning Models in Breast Cancer Detection: A Review of Convolutional Neural Networks (CNN) and Global Research Trends. LatIA [Internet]. 2025 Jan 1;3:117. Available from: https://latia.ageditor.uy/index.php/latia/article/view/117

8. Salem Alzboon M, Subhi Al-Batah M, Alqaraleh M, Alzboon F, Alzboon L. Guardians of the Web: Harnessing Machine Learning to Combat Phishing Attacks. Gamification Augment Real [Internet]. 2025 Jan;3:91. Available from: http://dx.doi.org/10.56294/gr202591

9. Alzboon MS, Subhi Al-Batah M, Alqaraleh M, Alzboon F, Alzboon L. Phishing Website Detection Using Machine Learning. Gamification Augment Real [Internet]. 2025 Jan 16;3:81. Available from: http://dx.doi.org/10.56294/gr202581

10. Alqaraleh M, Salem Alzboon M, Mohammad SA-B. Optimizing Resource Discovery in Grid Computing: A Hierarchical and Weighted Approach with Behavioral Modeling. LatIA [Internet]. 2025 Jan 1;3:97. Available from: https://latia.ageditor.uy/index.php/latia/article/view/97

11. Wahed MA, Alzboon MS, Alqaraleh M, Al-Batah M, Bader AF, Wahed SA. Enhancing Diagnostic Precision in Pediatric Urology: Machine Learning Models for Automated Grading of Vesicoureteral Reflux. In: 2024 7th International Conference on Internet Applications, Protocols, and Services (NETAPPS) [Internet]. IEEE; 2024. p. 1–7. Available from: http://dx.doi.org/10.1109/netapps63333.2024.10823509

12. Alqaraleh M. Enhanced Resource Discovery Algorithm for Efficient Grid Computing. In: Proceedings of the 3rd International Conference on Applied Artificial Intelligence and Computing, ICAAIC 2024. 2024. p. 925–31.

13. Alqaraleh M, Alzboon MS, Al-Batah MS. Skywatch: Advanced Machine Learning Techniques for Distinguishing UAVs from Birds in Airspace Security. Int J Adv Comput Sci Appl [Internet]. 2024;15(11). Available from: http://dx.doi.org/10.14569/IJACSA.2024.01511104

14. Abdel Wahed M, Al-Batah M, Salem Alzboon M, Fuad Bader A, Alqaraleh M. Technological Innovations in Autonomous Vehicles: A Focus on Sensor Fusion and Environmental Perception [Internet]. 2024 7th International Conference on Internet Applications, Protocols, and Services (NETAPPS). IEEE; 2024 Nov. Available from: http://dx.doi.org/10.1109/netapps63333.2024.10823624

15. Mowafaq SA, Alqaraleh M, Al-Batah MS. AI in the Sky: Developing Real-Time UAV Recognition Systems to Enhance Military Security. Data Metadata. 2024;3:417.

16. Al-Batah MS, Alzboon MS, Alzyoud M, Al-Shanableh N. Enhancing Image Cryptography Performance with Block Left Rotation Operations. Appl Comput Intell Soft Comput. 2024;2024(1):3641927.

17. Abuashour A, Salem Alzboon M, Kamel Alqaraleh M, Abuashour A. Comparative Study of Classification Mechanisms of Machine Learning on Multiple Data Mining Tool Kits. Am J Biomed Sci Res 2024 [Internet]. 2024;22(1):1. Available from: www.biomedgrid.com

18. Wahed MA, Alzboon MS, Alqaraleh M, Halasa A, Al-Batah M, Bader AF. Comprehensive Assessment of Cybersecurity Measures: Evaluating Incident Response, AI Integration, and Emerging Threats. In: 2024 7th International Conference on Internet Applications, Protocols, and Services (NETAPPS) [Internet]. IEEE; 2024. p. 1–8. Available from: http://dx.doi.org/10.1109/netapps63333.2024.10823603

19. Muhyeeddin A, Mowafaq SA, Al-Batah MS, Mutaz AW. Advancing Medical Image Analysis: The Role of Adaptive Optimization Techniques in Enhancing COVID-19 Detection, Lung Infection, and Tumor Segmentation. LatIA [Internet]. 2024 Sep 29;2(74):74. Available from: https://latia.ageditor.uy/index.php/latia/article/view/74

20. Alzboon MS, Al-Shorman HM, Alka'awneh SMN, Saatchi SG, Alqaraleh MKS, Samara EIM, et al. The Role of Perceived Trust in Embracing Artificial Intelligence Technologies: Insights from Jordan's SME Sector. In: Studies in Computational Intelligence [Internet]. Springer Nature Switzerland; 2024. p. 1–15. Available from: http://dx.doi.org/10.1007/978-3-031-74220-0_1

21. Alqaraleh M. Enhancing Internet-based Resource Discovery: The Efficacy of Distributed Quadtree Overlay. In: Proceedings of the 3rd International Conference on Applied Artificial Intelligence and Computing, ICAAIC 2024. 2024. p. 1619–28.

22. Al-Shanableh N, Alzyoud M, Al-Husban RY, Alshanableh NM, Al-Oun A, Al-Batah MS, et al. Advanced ensemble machine learning techniques for optimizing diabetes mellitus prognostication: A detailed examination of hospital data. Data Metadata. 2024;3:363.

23. Alzboon MS, Alqaraleh M, Wahed MA, Alourani A, Bader AF, Al-Batah M. AI-Driven UAV Distinction: Leveraging Advanced Machine Learning. In: 2024 7th International Conference on Internet Applications, Protocols, and Services (NETAPPS) [Internet]. IEEE; 2024. p. 1–7. Available from: http://dx.doi.org/10.1109/netapps63333.2024.10823488

24. Al-Batah MS, Salem Alzboon M, Solayman Migdadi H, Alkhasawneh M, Alqaraleh M. Advanced Landslide Detection Using Machine Learning and Remote Sensing Data. Data Metadata [Internet]. 2024 Oct 7;1. Available from: https://dm.ageditor.ar/index.php/dm/article/view/419/782

25. Wahed MA, Alzboon MS, Alqaraleh M, Ayman J, Al-Batah M, Bader AF. Automating Web Data Collection: Challenges, Solutions, and Python-Based Strategies for Effective Web Scraping. In: 2024 7th International Conference on Internet Applications, Protocols, and Services (NETAPPS) [Internet]. IEEE; 2024. p. 1–6. Available from: http://dx.doi.org/10.1109/netapps63333.2024.10823528

26. Alqaraleh M, Alzboon MS, Al-Batah MS, Wahed MA, Abuashour A, Alsmadi FH. Harnessing Machine Learning for Quantifying Vesicoureteral Reflux: A Promising Approach for Objective Assessment. Int J online Biomed Eng. 2024;20(11):123–45.

27. Al-Batah M, Salem Alzboon M, Alqaraleh M, Ahmad Alzaghoul F. Comparative Analysis of Advanced Data Mining Methods for Enhancing Medical Diagnosis and Prognosis. Data Metadata. 2024;3:465.

28. Alzboon MS, Qawasmeh S, Alqaraleh M, Abuashour A, Bader AF, Al-Batah M. Pushing the Envelope: Investigating the Potential and Limitations of ChatGPT and Artificial Intelligence in Advancing Computer Science Research. In: 2023 3rd International Conference on Emerging Smart Technologies and Applications, eSmarTA 2023. 2023.

29. Alzboon MS, Qawasmeh S, Alqaraleh M, Abuashour A, Bader AF, Al-Batah M. Machine Learning Classification Algorithms for Accurate Breast Cancer Diagnosis. In: 2023 3rd International Conference on Emerging Smart Technologies and Applications, eSmarTA 2023. 2023.

30. Alzboon MS, Bader AF, Abuashour A, Alqaraleh MK, Zaqaibeh B, Al-Batah M. The Two Sides of AI in Cybersecurity: Opportunities and Challenges. In: Proceedings of 2023 2nd International Conference on Intelligent Computing and Next Generation Networks, ICNGN 2023. 2023.

31. Putri AK, Alzboon MS. Doctor Adam Talib's Public Relations Strategy in Improving the Quality of Patient Service. Sinergi Int J Commun Sci. 2023;1(1):42–54.

32. Qananwah Q, Khader A, Al-Hashem M, Mumani A, Dagamseh A. Investigating the impact of smoking habits through photoplethysmography analysis. Physiol Meas. 2024 Jan;45(1):15003.

33. Al-Batah MS, Alzboon MS, Alazaidah R. Intelligent Heart Disease Prediction System with Applications in Jordanian Hospitals. Int J Adv Comput Sci Appl. 2023;14(9):508–17.

34. Fedecostante M, Sabbatinelli J, Dell'Aquila G, Salvi F, Bonfigli AR, Volpato S, et al. Prediction of COVID-19 in-hospital mortality in older patients using artificial intelligence: a multicenter study. Front Aging. 2024 Oct;5.

35. Alzboon MS, Al-Batah MS. Prostate Cancer Detection and Analysis using Advanced Machine Learning. Int J Adv Comput Sci Appl. 2023;14(8):388–96.

36. Alam A, Prybutok VR. Use of responsible artificial intelligence to predict health insurance claims in the USA using machine learning algorithms. Explor Digit Heal Technol. 2024;30–45.

37. Chua HL, Huong A. Deep Learning and Photoacoustic Technology for Microcirculation Classification: Comparison Between Smoking and Nonsmoking Groups. Glob Clin Eng J. 2024;6(3):37–45.

38. Ammar M, Javaid N, Alrajeh N, Shafiq M, Aslam M. A Novel Blending Approach for Smoking Status Prediction in Hidden Smokers to Reduce Cardiovascular Disease Risk. IEEE Access. 2024;

39. Alzboon MS, Al-Batah M, Alqaraleh M, Abuashour A, Bader AF. A Comparative Study of Machine Learning Techniques for Early Prediction of Diabetes. In: 2023 IEEE 10th International Conference on Communications and Networking, ComNet 2023 - Proceedings. 2023. p. 1–12.

40. Trofimova O, Iuliani I, Bergmann S. Bidirectional genetic and phenotypic links between smoking and striatal iron content involving dopaminergic and inflammatory pathways. medRxiv. 2024;

41. Alzboon MS, Al-Batah MS, Alqaraleh M, Abuashour A, Bader AFH. Early Diagnosis of Diabetes: A Comparison of Machine Learning Methods. Int J online Biomed Eng. 2023;19(15):144–65.

42. Teixeira A, Páscoa R, Martins C, Rodrigues PP. Finding Groups in the Portuguese Population Through Lifestyle: A Cross-Sectional Study. Stud Health Technol Inform. 2024;316:1659–63.

43. Goyal P, Goyal GK, Yadav K, Bhatt A, Nassa K, Raushan SK, et al. Attributes of hypertension among industrial workers in Northern India - An alarming signal. J Fam Med Prim Care. 2024;13(1):330–5.

44. Flores-Hernández MN, Martínez-Coria H, López-Valdés HE, Arteaga-Silva M, Arrieta-Cruz I, Gutiérrez-Juárez R. Efficacy of a High-Protein Diet to Lower Glycemic Levels in Type 2 Diabetes Mellitus: A Systematic Review. Vol. 25, International Journal of Molecular Sciences. 2024.

45. Jantzen R, Camilleri-Broët S, Ezer N, Broët P. A Statistical Testing Strategy Accounting for Random and Nonrandom (Skewed) X-Chromosome Inactivation Identifies Lung Cancer Susceptibility Loci among Smokers. Hum Hered. 2024 Jun;89(1):71–83.

46. Alzboon MS, Al-Batah M, Alqaraleh M, Abuashour A, Bader AF. A Comparative Study of Machine Learning Techniques for Early Prediction of Prostate Cancer. In: 2023 IEEE 10th International Conference on Communications and Networking, ComNet 2023 - Proceedings. 2023. p. 1–12.

47. Alzboon MS. Survey on Patient Health Monitoring System Based on Internet of Things. Inf Sci Lett. 2022;11(4):1183–90.

48. Alzboon M. Semantic Text Analysis on Social Networks and Data Processing: Review and Future Directions. Inf Sci Lett. 2022;11(5):1371–84.

49. Alzboon MS, Aljarrah E, Alqaraleh M, Alomari SA. Nodexl Tool for Social Network Analysis. Turkish J Comput Math Educ. 2021;12(14):202–16.

50. Alomari SA, Alzboon MS, Al-Batah MS, Zaqaibeh B. A novel adaptive schema to facilitates playback switching technique for video delivery in dense LTE cellular heterogeneous network environments. Int J Electr Comput Eng [Internet]. 2020 Oct;10(5):5347. Available from: http://ijece.iaescore.com/index.php/IJECE/article/view/16563

51. Alomari SA, Alqaraleh M, Aljarrah E, Alzboon MS. Toward achieving self-resource discovery in distributed systems based on distributed quadtree. J Theor Appl Inf Technol. 2020;98(20):3088–99.

52. Alzboon MS, Mahmuddin M, Arif S. Resource discovery mechanisms in shared computing infrastructure: A survey. In: Advances in Intelligent Systems and Computing. 2020. p. 545–56.

53. Alomari SA, Salaimeh S Al, Jarrah E Al, Alzboon MS. Enhanced logistics information service systems performance: using theoretical model and cybernetics' principles. WSEAS Trans Bus Econ [Internet]. 2020 Apr;17:278–87. Available from: https://wseas.com/journals/bae/2020/a585107-896.pdf

54. Banikhalaf M, Alomari SA, Alzboon MS. An advanced emergency warning message scheme based on vehicles speed and traffic densities. Int J Adv Comput Sci Appl. 2019;10(5):201–5.

55. Alzboon MS. Internet of things between reality or a wishing - list : a survey. Int J Eng \& Technol. 2019;7(June):956–61.

56. Shawawreh S, Alomari SA, Alzboon MS, Al Salaimeh S. Evaluation of knowledge quality in the E -learning system. Int J Eng Res Technol. 2019;12(4):548–53.

57. Al Tal S, Al Salaimeh S, Ali Alomari S, Alqaraleh M. The modern hosting computing systems for small and medium businesses. Acad Entrep J. 2019;25(4):1–7.

58. Alomari, Alzboon, Zaqaibeh, Al-Batah, Saleh Ali, Mowafaq Salem, Belal MS. An Effective Self-Adaptive Policy for Optimal Video Quality over Heterogeneous Mobile Devices and Network Discovery Services. Appl Math Inf Sci [Internet]. 2019 May;13(3):489–505. Available from: http://www.naturalspublishing.com/Article.asp?ArtcID=19739

59. Al-Batah M, Zaqaibeh B, Alomari SA, Alzboon MS. Gene Microarray Cancer classification using correlation based feature selection algorithm and rules classifiers. Int J online Biomed Eng. 2019;15(8):62–73.

60. Alzboon MS, Alomari S, Al-Batah MS, Alomari SA, Banikhalaf M. The characteristics of the green internet of things and big data in building safer, smarter, and sustainable cities Vehicle Detection and Tracking for Aerial Surveillance Videos View project Evaluation of Knowledge Quality in the E-Learning System View pr [Internet].

Vol. 6, Article in International Journal of Engineering and Technology. 2017. p. 83–92. Available from: https://www.researchgate.net/publication/333808921

61. Mahmuddin M, Alzboon MS, Arif S. Dynamic network topology for resource discovery in shared computing infrastructure. Adv Sci Lett. 2017;23(6):5402–5.

62. Arif S, Alzboon MS, Mahmuddin M. Distributed quadtree overlay for resource discovery in shared computing infrastructure. Adv Sci Lett. 2017;23(6):5397–401.

63. Alzboon MS, Arif AS, Mahmuddin M. Towards self-resource discovery and selection models in grid computing. ARPN J Eng Appl Sci. 2016;11(10):6269–74.

64. Mowafaq Salem Alzboon M. Mahmuddin ASCA. Challenges and Mitigation Techniques of Grid Resource Management System. In: National Workshop on FUTURE INTERNET RESEARCH (FIRES2016). 2016. p. 1–6.

65. SalemAlzboon, Mowafaq and Arif, Suki and Mahmuddin, M and Dakkak O. Peer to Peer Resource Discovery Mechanisms in Grid Computing : A Critical Review. In: The 4th International Conference on Internet Applications, Protocols and Services (NETAPPS2015). 2015. p. 48–54.

66. Alzboon MS, Sintok UUM, Sintok UUM, Arif S. Towards Self-Organizing Infrastructure : A New Architecture for Autonomic Green Cloud Data Centers. ARPN J Eng Appl Sci. 2015;1–7.

67. Al-Oqily I, Alzboon M, Al-Shemery H, Alsarhan A. Towards autonomic overlay self-load balancing. In: 2013 10th International Multi-Conference on Systems, Signals and Devices, SSD 2013. Ieee; 2013. p. 1–6.

FINANCING

None.

CONFLICT OF INTEREST

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Data curation: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Formal analysis: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Research: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Methodology: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Project management: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Resources: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Software: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Supervision: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Validation: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Display: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Drafting - original draft: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.

Writing: Muhyeeddin Alqaraleh, Mohammad Subhi Al-Batah, Mowafaq Salem Alzboon, Faisal Alzboon, Lujin Alzboon, Mohammad Nayef Alamoush.