



## UTILIZING ARTIFICIAL INTELLIGENCE TO ANALYZE GENDER DIFFERENCES IN HYPERTENSION RISK FACTORS

Husnul Khuluq\*, Tri Cahyani, Lazuardi Fatahillah Hamdi, Tunjung Winarno

Universitas Muhammadiyah Gombong, Sangkalputung, Gombong, Kebumen, Jawa Tengah 54411, Indonesia

\*[husnulkhuluq@unimugo.ac.id](mailto:husnulkhuluq@unimugo.ac.id)

### ABSTRACT

Hypertension continues to pose a significant challenge to global health. Early identification of risk factors, particularly those influenced by gender differences, has the potential to markedly enhance treatment processes and outcomes. Artificial intelligence (AI), specifically machine learning (ML), offers a promising avenue for identifying and analysing these critical risk factors. Objective: This study aims to explore the influence of gender differences on risk factors affecting hypertensive patients by examining demographic, medication, clinical, and laboratory data. Method: The study utilized medical records of hospitalized hypertensive patients at PKU Muhammadiyah Hospital Yogyakarta, covering the years 2022 to 2023. Logistic regression analysis with Lasso penalty was applied to determine the most influential variables. Additionally, the Random Forest algorithm implemented in WEKA, combined with a 10-fold cross-validation approach, was employed to evaluate the model's diagnostic performance using metrics such as precision, sensitivity, specificity, accuracy, and the area under the receiver operating characteristic curve (ROC-AUC). Results: A total of 1,006 patients were included in the sample, comprising 504 males and 502 females. Among the 33 clinical variables analysed, 13 demonstrated non-zero coefficients. For female hypertensive patients, the five most significant risk factors, along with their coefficients, were Haemoglobin (0.03), Diabetes Mellitus (0.04), Lymphocytes (0.06), Anaemia (0.13), and Creatinine (0.15). In male hypertensive patients, the top five risk factors and their coefficients were Acute Kidney Injury (-0.32), Erythrocytes (-0.15), Congestive Heart Failure (-0.03), Leukocytes (-0.02), and Length of Stay (LOS) (-0.01). The model's overall performance, as reflected by a ROC-AUC score of 0.805, indicates a good level of predictive accuracy. Conclusions: The findings reveal a significant association between gender and hypertension risk factors. These results underscore the potential for gender-specific customization of hypertension treatments, paving the way for more individualized therapeutic strategies and improved patient outcomes.

Keywords: artificial intelligence; gender differences; hypertension; machine learning; risk factors

### How to cite (in APA style)

Khuluq, H., Widiastuti, T. C., Hamdi, L. F., & Winarno, T. (2024). Utilizing Artificial Intelligence to Analyze Gender Differences in Hypertension Risk Factors. *Indonesian Journal of Global Health Research*, 7(1), 589-598. <https://doi.org/10.37287/ijghr.v7i1.4929>.

## INTRODUCTION

The global prevalence of hypertension is estimated at 33% among adults aged 30-79 years, while the prevalence in Indonesia is 40%, with 36% of men and 45% of women affected. Only 5% of women and 4% of men are able to effectively control their blood pressure (Bin Zhou, et al, 2023). Globally, the prevalence of uncontrolled hypertension has decreased slightly, from 29% in 2010 to 26% in 2019. However, this reduction has not met the global target of 21% by 2025. Uncontrolled hypertension is associated with an increased long-term risk of serious conditions such as heart attacks, heart failure, strokes, kidney disease, and dementia (Ahuja et al, 2018). Research conducted by Kringeland et al., (2022) shows that hypertensive women are at greater risk of coronary heart disease compared to hypertensive men. Another study by Madsen et al. (2019) indicates that hypertensive women are at a higher risk of stroke compared to hypertensive men. Studies in the UK and Norway suggest that the incidence of myocardial infarction is higher among hypertensive men than hypertensive women (Millett et al 2018). However, research in Korea on young hypertensive patients shows that myocardial infarction and ischemic stroke are more likely to occur in women than

in men (Lee et al., 2020). In postmenopausal women, blood pressure is generally higher than in men, due to the decline in estrogen levels and increased salt sensitivity (Kim et al. 2014). Estrogen has a protective effect on blood pressure (Iorga et al., 2017), and salt sensitivity in women is influenced by genetic factors (Manosroi et al., 2017). A multi-ethnic study by Osude et al. (2021) showed that women are more frequently treated with diuretics, while men are more often prescribed ACE inhibitors. Similar findings were reported by Ljungman et al. (2014).

In recent years, artificial intelligence (AI) has advanced, particularly through machine learning (ML), which is capable of solving complex problems involving interactions among numerous variables (Rahul C. Deo, MD, 2017). Machine learning is a branch of computer science focused on applying algorithms to identify intricate patterns within data (Murphy, 2018). Supervised machine learning is a type of ML that predicts outcomes based on input variables, which can be in the form of numbers, text, or images (Yu, Beam, & Kohane, 2018). AI methods, including machine learning and deep learning, have shown superior accuracy over traditional statistical models. These advancements allow for better handling of nonlinear relationships and complex data structures that classical models often struggle with (Lotfi & Bouhadi, 2021). This study aims to address the following research questions: To what extent is machine learning effective in predicting the influence of gender differences on hypertensive patients receiving hospital treatment, considering medication, comorbidities, demographic characteristics, and laboratory data? Additionally, what are the key factors associated with gender differences among hypertensive patients based on these variables?

## **METHOD**

The research phases consist of Data Gathering, Data Cleaning and Feature Selection, Deployment of Machine Learning Models, and Model Assessment. Evaluation. The study's data collected from the inpatient medical records for individuals who received a hypertension diagnosis. Among the 1,006 patients in the sample, 504 (50.1%) were male and 502 (49.9%) were female. The data was collected between 2022 and 2023 from PKU Muhammadiyah Hospital in Yogyakarta. Feature selection was performed using logistic regression with Lasso penalty (L1 regularization) to identify the most significant clinical variables from a total of 33 variables. Lasso (Least Absolute Shrinkage and Selection Operator) is a regression technique used for feature selection and simultaneous coefficient estimation. The Lasso penalty uses the L1 norm, which adds a penalty to the absolute sum of the coefficient values. The primary goal of this penalty is to force some coefficients to become zero, effectively performing feature selection and generating a simpler, more interpretable model (Hastie, Tibshirani, & Wainwright, 2015). Lasso was executed using Python 3.9 and scikit-learn version 1.0.2.

WEKA 3.8.6 was used as the platform for running the Random Forest algorithm, and the 10-fold cross-validation method was employed to evaluate the model. WEKA is a machine learning (ML) software suite developed at the University of Waikato, New Zealand, using the Java programming language. This software includes various ML algorithms for data mining purposes. WEKA also provides several tools for data processing, including pre-processing, classification, association rules, and visualization (Witten, Frank, & Hall, 2016). Several studies on hypertension have used the Random Forest algorithm. For example, Fuente et al (2019) applied Random Forest to identify behavior patterns in electronic health records of patients with essential hypertension. Ray et al. (2017) conducted a study to evaluate symptoms of essential hypertension using Random Forest, predicting that left ventricular hypertrophy, neck stiffness, and sensitivity to sound were the most important symptoms for detecting hypertension. The 10-fold cross-validation method is commonly used in data mining research with ML, as it differs from traditional methods, such as percentage split. This method

helps reduce variation in prediction errors, optimizes the use of training and testing data, prevents overlap or overfitting during the training and testing process, and better facilitates data randomization (Bergman, Purucker, & Hutter, 2024). The model was evaluated using the following metrics: precision, sensitivity, specificity, accuracy, and ROC-AUC (Area Under the Receiver Operating Characteristic Curve) These metrics are summarized in the confusion matrix table(Saito & Rehmsmeier, 2015).

Table 1.  
Confusion matrix(Zhou & Valle, 2020)

		Actual Condition	
		Actual Positive Condition (Female patients) - True Positive (TP)	Actual Negative Condition (Male patients) - False Positive (FP)
Predicted condition	Predicted Positive Condition (Female patients)	$(TP / True Positive)$	$(FP / False positive)$
	Predicted Negative Condition (Male patients)	$(FN / False Negative)$	$(TN / True negative)$

From the confusion matrix table, several parameters are derived for evaluation, including:

$$accuracy = \frac{TP+TN}{(TP+TN+FP+FN)}, \text{ specificity} = \frac{TN}{(TN+FP)}, \text{ precision} = \frac{TP}{(TP+FPR_{recall})}, \text{ and}$$

$$\text{sensitivity/TPR} = \frac{TP}{(TP+FN)}.$$

The values of each parameter above range from 0 to 1, with the maximum value being 1. Subsequently, a ROC curve is created with the Y-axis representing Sensitivity / True Positive Rate (TPR), And the X-axis represents 1 - Specificity / False Positive Rate (FPR), calculated using the formula:  $FPR = \frac{FP}{(FP+TN)}$

## RESULT

The data collection for hypertensive patients was conducted from January 1, 2022, to December 31, 2023, involving a total of 701 patients, consisting of 504 males (53.07%) and 502 females (46.93%). The following are the data from the 33 parameters that were examined.

Table 2.  
Descriptive Data of Hypertensive Patients

Parameter	Male (%)	Male (f)	Female (%)	Female (f)
Treatment				
Amlodipine	48.61	245.0	52.59	264.0
Bisoprolol	2.58	13.0	2.99	15.0
Captopril	2.78	14.0	2.99	15.0
Candesartan	3.37	17.0	3.39	17.0
Clonidine	1.59	8.0	2.99	15.0
Dopamet	0.00	0.0	0.20	1.0
Furosemide	12.50	63.0	13.35	67.0
Nicardipine	4.17	21.0	4.18	21.0
Nifedipine	0.99	5.0	0.40	2.0
Propranolol	0.00	0.0	0.10	1.0
Ramipril	1.59	8.0	1.99	10.0
Spirolactone	2.38	12.0	2.79	14.0
Comorbidities				

Parameter	Male (%)	Male (f)	Female (%)	Female (f)
Age >65 years	41.87	211.0	44.02	221.0
AKI	4.17	21.0	0.80	4.0
Anaemia	5.95	30.0	10.36	52.0
Asthma	0.79	4.0	1.39	7.0
CHF	2.58	13.0	3.19	16.0
CI	5.95	30.0	7.57	38.0
CKD	4.17	21.0	3.99	20.0
DM	15.08	76.0	16.73	84.0
Pneumonia	3.17	16.0	3.99	20.0
Laboratory Indicators				
Erythrocytes	6.75	34.0	7.57	38.0
Glucose	4.37	22.0	5.58	28.0
Haemoglobin	7.14	36.0	8.37	42.0
Creatinine	5.16	26.0	5.98	30.0
Leukocytes	38.69	195.0	38.65	194.0
Lymphocytes	50.00	252.0	57.17	287.0
Neutrophils	41.27	208.0	48.41	243.0
Thrombocytes	8.13	41.0	9.76	49.0
Vital Signs and LOS				
LOS	67.46	340.0	65.94	331.0
D (IN)	91.07	459.0	93.03	467.0
D (OUT)	27.58	139.0	28.88	145.0
S (IN)	89.29	450.0	89.84	451.0
S (OUT)	83.73	422.0	86.25	433.0

AKI = Acute Kidney Injury, CHF = Chronic Heart Failure, CKD = Chronic Kidney Disease, DM = Diabetes Mellitus, LOS = Length of Stay, D = diastolic, S = systolic.

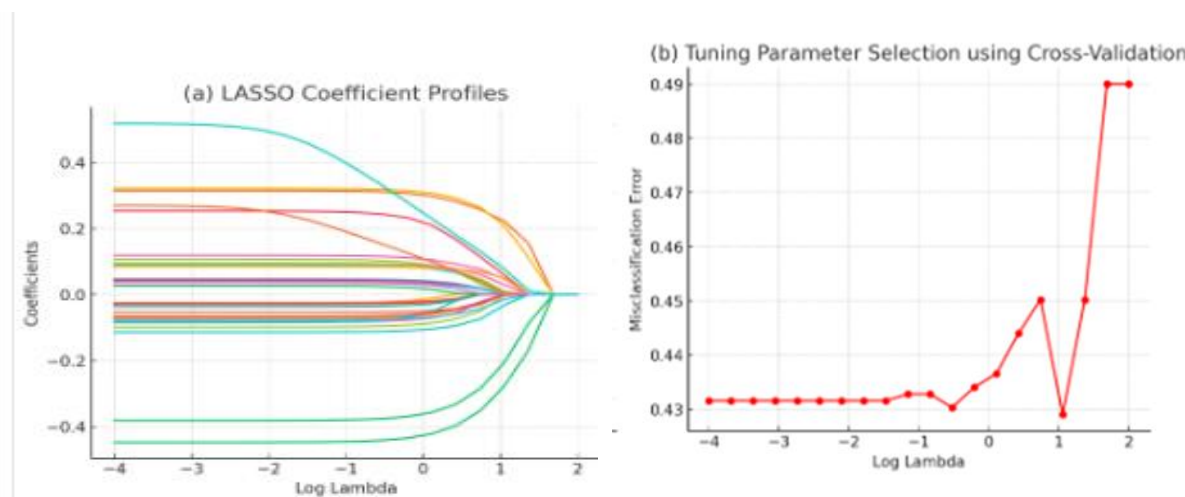


Figure 1 (a) - LASSO Coefficient Profile: This graph illustrates the changes in the coefficients of each variable as the regularization parameter, lambda, varies. On the left side (where log(lambda) is smaller), more coefficients are active with higher values. As log(lambda) increases, many coefficients shrink to zero, indicating that stronger regularization reduces model complexity by eliminating certain variables.

Figure 2 (b) - Misclassification Error Based on Cross-Validation: This graph presents the misclassification error for each lambda value during cross-validation. The red points and error lines demonstrate how the error changes across different lambda values. The error tends to remain stable at smaller lambda values but increases when lambda becomes too large, suggesting that the model becomes overly simplistic and loses accuracy. The data was analyzed using logistic regression with the Lasso (L1) regularization approach,

which facilitates feature selection and identifies the most significant variables in predicting the influence of gender on hypertension patients. This analysis was performed using Python version 3.9 and scikit-learn version 1.0.2. The results are presented in Table 3 and Figure 1. The optimal lambda value selected was approximately 0.0036.

Table 3.  
Results of LASSO Analysis

Parameter	Coefficient
AKI	-0.32
Erythrocytes	-0.15
CHF	-0.03
Leukocytes	-0.02
Length of Stay (LOS)	-0.01
Clonidine	0.01
Neutrophil	0.01
Amlodipine	0.02
Haemoglobin	0.03
DM	0.04
Lymphocyte	0.06
Anemia	0.13
Creatinine	0.15

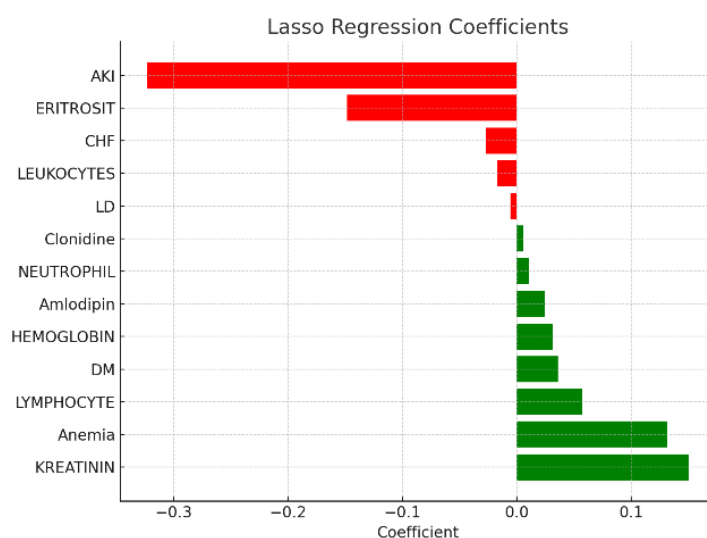


Figure 3: LASSO analysis results, where green bars represent positive coefficients correlated with females, and red bars indicate negative coefficients correlated with males. Following feature selection with LASSO, the Random Forest algorithm was employed for classification. The model was implemented using WEKA version 3.8.6 and evaluated through 10-fold cross-validation. The results include the confusion matrix (Table 3), followed by calculations of evaluation metrics such as precision, accuracy, sensitivity, specificity, and ROC-AUC (Receiver Operating Characteristic Area Under the Curve), as shown in Table 4. Finally, the sensitivity and false positive rate were plotted to visualize the ROC curve (Figure 4)

Tabel 4.  
Confusion matrix

<i>True Positive (TP)</i>	<i>False Positive (FP)</i>	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>
587	201	218	284

Table 5.  
Evaluation Results Using the Random Forest Algorithm

Metric	Sensitivity	Specificity	Precision	Accuracy	ROC AUC
Value	75.3%	33.9%	74.7%	75.33%	80.5%

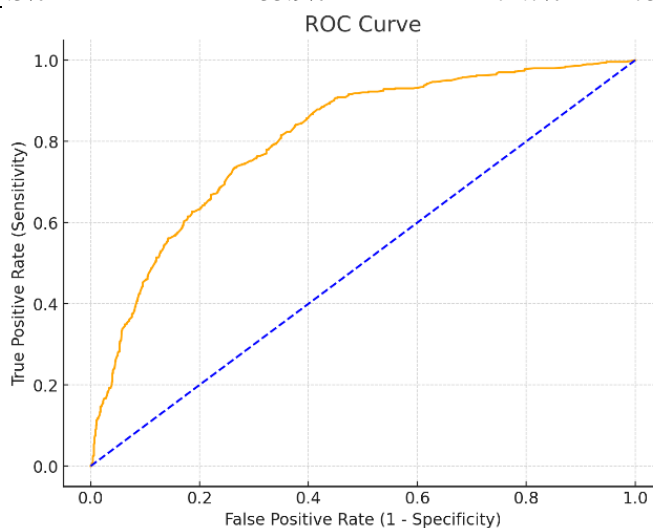


Figure 4. ROC Curve Using the Random Forest Algorithm

**DISCUSSION**

The ROC-AUC value in this study is 80.5% (or 0.805), which falls under the "good" category (Nahm, 2022). The ROC curve above (Figure 4) provides a visual representation of the Random Forest algorithm's performance in distinguishing between two classes (male and female). Approaching the Top Left Corner: If the ROC curve approaches the top left corner, it indicates a high True Positive Rate (TPR) and a low False Positive Rate (FPR) (Das, Gupta, & Ghosh, 2023). This demonstrates that the model can identify more positive cases (female) with fewer errors. Diagonal Line (Random Guessing): The diagonal line from (0,0) to (1,1) represents a random model with no discriminative ability. If the ROC curve approaches the diagonal line, it suggests that the model is barely better than random guessing (Team, 2024). Several studies in the field of healthcare have demonstrated that the random forest algorithm performs good, with values exceeding 0.8. For instance, the research by Suleiman, Luka, & Ibrahim (2023) on predicting cardiovascular disease risk using random forest reported an AUC value of 0.9. Another study by Windy et al (2024) on predicting stroke patients using random forest achieved an AUC value of 0.98. Additionally, the research by Khuluq *et al* (2024) on predicting COVID-19 patient mortality reported a random forest AUC value of 0.94. The Random Forest is a tree-based learning technique that generates robust predictions, making it one of the most popular algorithms. Additionally, this algorithm possesses the advantage of effectively handling complex data.(Manzali *et al*, 2024)

*Lasso coefficient (Table 3)*

Creatinine(Coefficient:+0.15)

Abnormal (high) creatinine levels are more frequently observed in female hypertension patients. Sydorчук et al. ( 2020) found that the CYP11B2 gene polymorphism in hypertensive women is associated with increased creatinine and aldosterone levels, which

contribute to elevated blood pressure and reduced kidney function Rekhviashvili et al . (2015) further demonstrated that in hypertensive patients, creatinine levels positively correlate with carotid and coronary artery damage, particularly in hypertensive women.

*Anemia(Coefficient:+0.13)*

Anemia is strongly associated with hypertension in women. Park et al.(2023) identified a link between anemia and hyperuricemia in hypertensive patients, especially in women, suggesting that anemia may contribute to the metabolic risks associated with hypertension.

*Lymphocyte(Coefficient: +0.06)*

High lymphocyte levels are closely linked to hypertension in women. Tipton et al. (2012) demonstrated that hypertensive female mice exhibit greater anti-inflammatory T cell infiltration compared to males, indicating that lymphocytes contribute to hypertension in females and suggest sex-specific sensitivities.

*Diabetes(Coefficient:+0.04)*

Hypertension can exacerbate insulin resistance, a significant factor in developing type 2 diabetes. This decreased insulin sensitivity is more common in women than men, leading to the body's inability to effectively use insulin, ultimately contributing to higher blood sugar levels in hypertensive women ((Ciarambino et al, 2023)

*Hemoglobin(Coefficient:+0.03)*

Anemia and low hemoglobin levels are associated with increased blood pressure, particularly in women. Research indicates that women with low hemoglobin levels have a higher risk of developing hypertension (Tahe et al, 2024).

*AcuteKidneyInjury(AKI)(Coefficient:-0.31)*

Hypertensive men are an independent predictor of severe acute kidney injury (Szczech et al., 2010).

*Erythrocytes(Coefficient:-0.15)*

While hypertensive men and women both experience changes in erythrocyte function and deformability, men tend to show more significant alterations ((Dylewska et al, 2019)

*CongestiveHeartFailure(CHF)(Coefficient:-0.03)*

Studies show that men are more likely to develop eccentric LVH due to hypertension compared to women, who tend to have concentric LVH. This distinction may be influenced by genetic factors and hormonal differences(Aronow, 2017) (Dudink et al., 2018)

*Leucocyte (coefficient= -0.02)*

Elevated leukocyte counts are associated with higher blood pressure and an increased risk of hypertension, suggesting inflammation as a contributing factor. Men often show a stronger association due to higher inflammation markers linked with lifestyle factors(Xi et al., 2015). Men demonstrate higher leukocyte and neutrophil counts than women in cardiovascular and hypertensive conditions, which might be linked to genetic and hormonal influences(Rajkumari et al, 2013)

*Length of Stay (LOS) (coefficient= -0.01)*

Longer hospital stays for hypertensive patients are often linked to heart failure severity, with male patients more prone to longer durations due to disease progression and the need for intensive management(Reynolds et al., 2015).

## CONCLUSION

This study highlights the significant role of gender differences in influencing risk factors for hypertension, which can be more accurately identified using LASSO and Random Forest algorithms. The analysis identified 13 clinical variables that differentiate between male and female hypertension patients. The model demonstrated good predictive accuracy with an ROC-AUC of 0.805, indicating the effectiveness of machine learning in predicting hypertension risk factors, thereby supporting more personalized and targeted treatment for hypertension patients. Recommendations for Future Research: Utilize larger datasets, incorporating patients from multiple hospitals (multi-centre studies), to improve model generalization

## REFERENCES

- Ahuja, R., Ayala, C., Tong, X., Wall, H. K., & Fang, J. (2018). Public awareness of health-related risks from uncontrolled hypertension. *Preventing Chronic Disease*, 15(4), 6–11. <https://doi.org/10.5888/pcd15.170362>
- Aronow, W. (2017). Hypertension and left ventricular hypertrophy. *Annals of Translational Medicine*, 5(15), 310. <https://doi.org/10.21037/atm.2017.06.14>
- Bergman, E., Purucker, L., & Hutter, F. (2024). Don't Waste Your Time: Early Stopping Cross-Validation. *ArXiv*, 2405.03389. Retrieved from <https://arxiv.org/abs/2405.03389>
- Bin Zhou, Rodrigo M Carrillo-Larco, Goodarz Danaei, Leanne M Riley, Christopher J Paciorek, Gretchen A Stevens, Edward W Gregg, James E Bennett, Bethlehem Solomon, Rosie K Singleton, Marisa K Sophiea, Maria Lc Iurilli, Victor Pf Lhoste, Melanie J Cowan, S, R. C. (2023). *Global Report on Hypertension*.
- Ciarambino, T., Crispino, P., Guarisco, G., & Giordano, M. (2023). Gender Differences in Insulin Resistance: New Knowledge and Perspectives. *Current Issues in Molecular Biology*. <https://doi.org/10.3390/cimb45100496>
- Das, S., Gupta, A., & Ghosh, S. (2023). Analysing the Role of AHP Model to Identify Flood Hazard Prone Areas of Sagar Island, India. *Journal of the Indian Society of Remote Sensing*, 51(1), 1–14. <https://doi.org/10.1007/s12524-023-01697-x>
- Dudink, E., Nieuwlaat, R., Rienstra, M., PhDD, M., Gelder, I., Camm, A., ... Luermans, J. (2018). Effect of Systemic Hypertension With Versus Without Left Ventricular Hypertrophy on the Progression of Atrial Fibrillation (from the Euro Heart Survey). *The American Journal of Cardiology*, 122(4), 578–583. <https://doi.org/10.1016/j.amjcard.2018.04.053>
- Dylewska, M., Chomiczka, I., & Małyszko, J. (2019). Hypertension in patients with acute kidney injury. *Wiadomosci Lekarskie (Warsaw, Poland : 1960)*, 72(11 cz 2), 2199–2201.
- Fuente, C. de la, Urrutia, A., & Chávez, E. (2019). Using the Random Forest Algorithm for Searching Behavior Patterns in Electronic Health Records. *IEEE Latin America Transactions*, 17(05), 875–881. <https://doi.org/10.1109/TLA.2019.8891957>
- Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*. (1st, Ed.). Boca Raton, FL: Chapman and Hall/CRC. <https://doi.org/10.1201/b18401>
- Iorga, A., Cunningham, C. M., Moazeni, S., Ruffenach, G., Umar, S., & Eghbali, M. (2017). The protective role of estrogen and estrogen receptors in cardiovascular disease and the controversial use of estrogen therapy. *Biology of Sex Differences*, 8(1), 33. <https://doi.org/10.1186/s13293-017-0152-8>
- Khuluq, H., Yusuf, P. A., Perwitasari, D. A., & Nguyen, T. (2024). Mortality prediction of COVID-19 patients using supervised machine learning. *IAES International Journal of Artificial Intelligence*, 13(4), 4472–4479.

- Kim, J.-M., Kim, T.-H., Lee, H.-H., Lee, S. H., & Wang, T. (2014). Postmenopausal Hypertension and Sodium Sensitivity. *J Menopausal Med*, 20(1), 1–6. Retrieved from <https://doi.org/10.6118/jmm.2014.20.1.1>
- Kringeland, E., Tell, G. S., Midtbø, H., Iglund, J., Haugsgjerd, T. R., & Gerds, E. (2022). Stage 1 hypertension, sex, and acute coronary syndromes during midlife: the Hordaland Health Study. *European Journal of Preventive Cardiology*, 29(1), 147–154. <https://doi.org/10.1093/eurjpc/zwab068>
- Lee, H., Yano, Y., Cho, S. M. J., Park, J. H., Park, S., Lloyd-Jones, D. M., & Kim, H. C. (2020). Cardiovascular Risk of Isolated Systolic or Diastolic Hypertension in Young Adults. *Circulation*, 141(22), 1778–1786. <https://doi.org/10.1161/CIRCULATIONAHA.119.044838>
- Ljungman, C., Kahan, T., Schiöler, L., Hjerpe, P., Hasselström, J., Wettermark, B., ... Manhem, K. (2014). Gender differences in antihypertensive drug treatment: results from the Swedish Primary Care Cardiovascular Database (SPCCD). *Journal of the American Society of Hypertension*, 8(12), 882–890. <https://doi.org/https://doi.org/10.1016/j.jash.2014.08.015>
- Lotfi, I., & Bouhadi, A. (2021). Artificial Intelligence Methods: Toward a New Decision Making Tool. *Applied Artificial Intelligence*, 36. <https://doi.org/10.1080/08839514.2021.1992141>
- Madsen, T. E., Howard, G., Kleindorfer, D. O., Furie, K. L., Oparil, S., Manson, J. E., ... Howard, V. J. (2019). Sex Differences in Hypertension and Stroke Risk in the REGARDS Study: A Longitudinal Cohort Study. *Hypertension (Dallas, Tex. : 1979)*, 74(4), 749–755. <https://doi.org/10.1161/HYPERTENSIONAHA.119.12729>
- Manosroi, W., Tan, J. W., Rariy, C. M., Sun, B., Goodarzi, M. O., Saxena, A. R., ... Seely, E. W. (2017). The Association of Estrogen Receptor- $\beta$  Gene Variation With Salt-Sensitive Blood Pressure. *The Journal of Clinical Endocrinology and Metabolism*, 102(11), 4124–4135. <https://doi.org/10.1210/jc.2017-00957>
- Manzali, Y., Akhiat, Y., Abdoulaye Barry, K., Akachar, E., & El Far, M. (2024). Prediction of Student Performance Using Random Forest Combined With Naïve Bayes. *The Computer Journal*, 67(8), 2677–2689. <https://doi.org/10.1093/comjnl/bxae036>
- Millett, E. R. C., Peters, S. A. E., & Woodward, M. (2018). Sex differences in risk factors for myocardial infarction: cohort study of UK Biobank participants. *BMJ (Clinical Research Ed.)*, 363, k4247. <https://doi.org/10.1136/bmj.k4247>
- Murphy, K. P. & B. F. (2018). *Machine Learning: A Probabilistic Perspective* (MIT Press, Cambridge, 2012). Acknowledgements. New computer technology (Vol. 16). cambridge. <https://doi.org/10.55056/nocote.v16i0.830>
- Nahm, F. S. (2022). ROC Curve: overview and practical use for clinicians. *Korean Journal of Anesthesiology*, 75(1), 25–36.
- Osude, N., Durazo-Arvizu, R., Markossian, T., Liu, K., Michos, E. D., Rakotz, M., ... Kramer, H. (2021). Age and sex disparities in hypertension control: The multi-ethnic study of atherosclerosis (MESA). *American Journal of Preventive Cardiology*, 8, 100230. <https://doi.org/10.1016/j.ajpc.2021.100230>
- Park, S. M., Gi, M. Y., Cha, J. A., Sung, H. H., Park, S. Y., Park, C. H., & Yoon, H. (2023). The relationship between anemia and hyperuricemia and hypertension in Korean adults: the Korea National Health and Nutrition Examination Survey 2016–2018. *Current Medical Research and Opinion*, 39(6), 819–825. <https://doi.org/10.1080/03007995.2023.2213004>
- Rahul C. Deo, MD, P. (2017). Machine Learning in Medicine Rahul. *Physiology & Behavior*, 176(3), 139–148. <https://doi.org/10.1161/CIRCULATIONAHA.115.001593>
- Rajkumari, R., Laishram, D., Thiyam, J., & Javan, N. (2013). Hypertensive leucocytosis. *Journal of the Indian Medical Association*, 111(4), 226–229.

- Ray, S., Mondal, A. C., Neogi, A., & Dey, K. (2017). A Study to Evaluate Symptoms in Essential Hypertension Using Random Forest Decision Tree Algorithm. *Indian Science Cruiser*, 31(4 SE-Search & Survey), 28–35. <https://doi.org/10.24906/isc/2017/v31/i4/158408>
- Rekhviashvili, A., Godziashvili, M., & Archvadze, N. (2015). PP.13.14: Association Between Blood Creatinine Concentration And Subclinical Atherosclerosis In Hypertensive And Normotensive Patients. *Journal of Hypertension*, 33. Retrieved from [https://journals.lww.com/jhypertension/fulltext/2015/06001/pp\\_13\\_14\\_\\_association\\_between\\_blood\\_creatinine.736.aspx](https://journals.lww.com/jhypertension/fulltext/2015/06001/pp_13_14__association_between_blood_creatinine.736.aspx)
- Reynolds, K., Butler, M., Kimes, T., Rosales, A., Chan, W., & Nichols, G. (2015). Relation of Acute Heart Failure Hospital Length of Stay to Subsequent Readmission and All-Cause Mortality. *The American Journal of Cardiology*, 116(3), 400–405. <https://doi.org/10.1016/j.amjcard.2015.04.052>
- Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Sari, W. J., Melyani, N. A., Arrazak, F., Anahar, M. A. Bin, Addini, E., Al-Sawaff, Z. H., & Manickam, S. (2024). Performance Comparison of Random Forest, Support Vector Machine and Neural Network in Health Classification of Stroke Patients. *Public Research Journal of Engineering, Data Technology and Computer Science*, 2(1), 34–43. <https://doi.org/10.57152/predatecs.v2i1.1119>
- Suleiman, A., Luka, S., & Ibrahim, M. (2023). Cardiovascular Disease Prediction Using Random Forest Machine Learning Algorithm. *Fudma Journal Of Sciences*, 7, 282–289. <https://doi.org/10.33003/fjs-2023-0706-2128>
- Sydorchuk, L., Dzhuryak, V., Sydorchuk, A., Levytska, S., Petrynych, V., Knut, R., ... Sydorchuk, R. (2020). The cytochrome 11B2 aldosterone synthase gene rs1799998 single nucleotide polymorphism determines elevated aldosterone, higher blood pressure, and reduced glomerular filtration, especially in diabetic female patients. *Endocrine Regulations*, 54(3), 217–226. <https://doi.org/10.2478/enr-2020-0024>
- Taha, Z., Elhag, H., Hassan, A. A., & Adam, I. (2024). Association Between Hemoglobin Level, Anemia, and Hypertension Among Adults in Northern Sudan: A Community-Based Cross-Sectional Study. *Vascular Health and Risk Management*, 20, 323–331. <https://doi.org/10.2147/VHRM.S472480>
- Team, S. (2024). How to Interpret the ROC Curve. *Statology*. Retrieved from <https://www.statology.org/interpret-roc-curve/>
- Tipton, A. J., Baban, B., & Sullivan, J. C. (2012). Female spontaneously hypertensive rats have greater renal anti-inflammatory T lymphocyte infiltration than males. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 303(4), R359–R367. <https://doi.org/10.1152/ajpregu.00246.2012>
- Witten, I. H., Frank, E., & Hall, M. A. (2016). *WEKA: Data Mining Software in Java*. (4th, Ed.). Burlington, MA: Elsevier. Retrieved from <https://www.cs.waikato.ac.nz/ml/weka/>
- Xi, L., Hao, Y., Liu, J., Wang, W., Wang, M., Qi, Y., ... Zhao, D. (2015). Relationship between leukocyte count and risk of hypertension. *Zhonghua Xin Xue Guan Bing Za Zhi*, 43(4), 312–318. <https://doi.org/10.3760/CMA.J.ISSN.0253-3758.2015.04.007>
- Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, 2(10), 719–731. <https://doi.org/10.1038/s41551-018-0305-z>
- Zhou, X., & Valle, A. (2020). Range Based Confusion Matrix for Imbalanced Time Series Classification. In *2020 6th Conference on Data Science and Machine Learning Applications (CDMA)* (pp. 1–6). <https://doi.org/10.1109/CDMA47397.2020.00006>