iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XI **Month of publication:** November 2022

DOI: https://doi.org/10.22214/ijraset.2022.47774

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

Run Your Health Profile -A Digital Health Profile to Store Personal Health History

A. G. K. P. Wickramasinghe[1], S. P. N. L. Perera[2], I. T. Jayarathna[3], R. A. K. C. Mario[4]
[1, 2, 3, 4]*Sri Lanka Institute of Information Technology*

Abstract: ***Communication between a doctor and a patient is a significant key to enhancing the doctor-patient relationship and providing high-quality healthcare. Over the years, health support has been evolving, and, in the present, it has taken a major step with e-health support. Over the coming years, doctor-patient face-to-face communication will become more limited and will be more supported by communication mediated by technology, such as electronic and personal health records (EHRs and PHRs) and Internet applications. E-health replaces paper-based medical records and documents with digital health records, which let the doctor review basic information about a patient, a patient's past visits, previous results of lab tests, the medications the patient is or was on, and an overall review of the health issue and its current status. The main goal is to provide insight into e-health tools which is essential for healthcare systems globally to adapt and embrace the rapidly evolving digital technologies.***
Keywords: Electronic medical records, Doctor-patient communication, OCR

I. INTRODUCTION

Information and communication technology (ICT) has resulted in enhanced efficiency and reliability in the day-to-day operations of the healthcare industry. The ubiquitous presence of ICT solutions has also resulted in the healthcare industry catering to more patients and hence a higher number of related requests such as tests, admissions, etc [1]. According to this study [2], the digitization process of healthcare services has been introduced in Western countries during the last ten years. However, the situation may be different in developing countries, where paper documents are still common for health reports and records in hospitals. The medical laboratory report is one kind of important clinical data, that helps healthcare professionals with patient assessment, diagnosis, and long-term monitoring. Nowadays, many techniques and products are developed for the paperless healthcare industry to protect our environment, but paper documents are still common in healthcare services like medical laboratory reports, especially in developing countries. For effective medical care and accurate diagnoses, the reports are required to be up-to-date and complete, but the paper files are neither portable nor convenient for long-term preservation [3]. To solve this problem, there should be an electronic laboratory report processing system that can extract information from medical laboratory reports. Income is the earning an individual receive in a givethatiod of time (generally a month)[4]. Income is an essential aspect of the daily life where an individual gets paid for the service or work, they perform. Doctors being the forefront of healthcare ensures to manage a sustainable health which is an incomparable service to all. Therefore, it is also crucial that a doctor receives an accurate income based on their experience and skilled qualifications. Machine Learning is a type of artificial intelligence which outputs patterns from raw data sets utilizing an algorithm rather than with the intervention of direct programming[5]. The most important feature of machine learning methodologies is the attempt to predict the final output as precisely as possible. A single methodology will not provide the best output, as different models will result in different outputs. Therefore, it requires testing each model to find the most desirable model[6]. For this research, linear, regression, LASSO, Ridge, Elastic net, Random Forest and Support vector regression models were tested to be used in predicting the consultation fee of doctors based on their experience and skilled qualification. Based on the results it was shown that linear regression outperformed among the other models. Linear Regression is an algorithm of machine learning where the data is inputted to predict the value of a dependent variable based on independent variables. The metrics that were used to evaluate the performance of the models were Root Mean Squared Error (RMSE) and R-Squared (R2). R2 known as the co-efficient of determination is the proportion of variation in the dependent variable, which is determined by the independent variable[7]. R-squared value ranges between 0 to 1. Thus, the model fits better with the algorithm, if the R2 is higher or closer to 1 and lesser RMSE error values get[8]. Communication between a doctor and a patient is a significant key to enhance the doctor patient relationship to provide a high-quality healthcare. Modern healthcare services require teamwork. To manage health issues doctors must be able to cooperate with patients and their families as well. At any stage of treatment, doctors play a vital role in planning, organizing, implementing and evaluating[9].

With all information on the health issues and the health history of patients and their families, it gets easier for a doctor to bring positive effects and a quality service, as a doctor is more likely to understand and response with a proper treatment. Furthermore, with the use of current technology it is much convenient and beneficial to maintain the doctor-patient relationship

II. LITERATURE REVIEW

The 6th International Conference on Digital Arts, Media, and Technology (DAMT) and 4th ECTI Northern Section Conference on Electrical, Electronics, Computer, and Telecommunications Engineering (NCON) introduce a health management platform based on a seamless electronic medical record (sEMR) approach. The system is designed to exchange emergency patients' data between the community hospital and the health-promoting hospital. This system uses electronic medical records for emergency patients. They present real-time visualization and data exchange between community hospitals and health-promoting hospitals to help medical staff with real-time health management [10].

In September 1998, a new National Health Service (NHS) strategy was launched in the U.K. A key goal is to create electronic health records (EHRs) by 2005.

The realization of the strategy will have several advantages, such as improving patient booking and appointment systems and providing speedier diagnoses, test results, online advice, and information to patients and caregivers. They used a framework called Ariadne to overcome challenges in developing a telematic system for process support, properly capturing and supporting the processes and the interactions between the various parties involved. They used data mining techniques to find patterns that could have implications for future healthcare requirements and recommendations [11].

The paper [12] "Disease Prediction using Machine Learning" used KNN, Naïve Bayes, Logistic Regression and Decision Tree algorithms to make a disease prediction system that can predict the disease on the basis of symptoms and implemented using grail framework.

As per the above researchers, articles, and findings the need of the person's medical history record in digitally are increasing and the functionalities of them are increasing with the developing technologies and according to the demand of society.

According to the aforementioned researchers, articles, and findings, it is clear that the need for a person's medical history record to be digitally stored is growing, and the functionalities of these records are growing in tandem with the advancement of technology and societal demand.

The process of identifying text content on scanned materials (such as images) and converting that text into easily understandable encoded text for computer readers is known as optical character recognition (OCR). The document can be scanned and processed, and the text retrieved and stored in an editable form through OCR, making the digitization process more pleasant. One type of important clinical data is the medical laboratory report, which is typically stored in pdf or image format. There have been some studies done on text detection using OCR in clinical laboratory reports.

The study done by the Beijing Key Lab of Transportation Data Analysis and Mining at Beijing Jiaotong University [13], followed two steps to extract text from medical laboratory reports. which are text localization and recognition. This study used convolutional neural networks (CNN) and Long-Short Term Memory (LSTM) to make significant progress in text recognition. For the inhomogeneous illumination conditions, they used gray histograms of sub-areas to enhance the contrast between the texts and the background. Due to the limited data, the CRNN model is trained on one million synthetic data points and tested on all real data points.

A study was done by Caraga State University in Caraga [14] extracted text from paper-based documents by converting them into a digital format using the pytesseract library. Then the recognized text is populated in the metadata form for record indexing.

Several machine learning methods and algorithms have been used over the years to optimize the performance of data organization and income prediction. These algorithms have helped predict the income using different datasets and data models best as possible[15]. In this research a suitable consultation fee of a doctor is predicted based on their specialized skills and the experience utilizing the linear regression algorithm.

Some previous research has been conducted to Matbouli and Alghamdi (2022) used machine learning to predict salary of labor holistically among major occupational groups[16]. Bjelland, J. et al., 2016 used the deep learning methodology to predict individual income data based mobile phone call detail records[17]. Ramadhani et al., 2018 conducted research to predict and classify income based on historical data using a DCNN (Deep Convolutional Neural Network) other machine learning techniques such as SVMs (Support Vector Machines), Gaussian, and decision tree.

III. METHODOLOGY

There are four main parts in the application, Namely, Health profile management, Test management, Doctor management, and medicine management. The application provides an opportunity to store health history digitally, Extract text from medical lab reports and store, predict doctor's consultation fees, and extract texts from medical prescriptions.

The profile management component plays an important role in our research. It outputs the user (health profile owner) to a profile to the user's personal records that the user has input to the system. This component fills the gap between the current solutions. Most current solutions focus only on keeping medical records or only to predict disease. Due to the capabilities of this component, we are giving users to have a profile to keep and maintain their medical records, can predict their health by using medical history records or by giving the symptoms, and can share their medical records with a dependent by linking the profiles. The output of the user profile will be generated by using Node.js. The main reason to be chosen Node.js is, it is a free and open-source server environment. And uses JavaScript on the server.

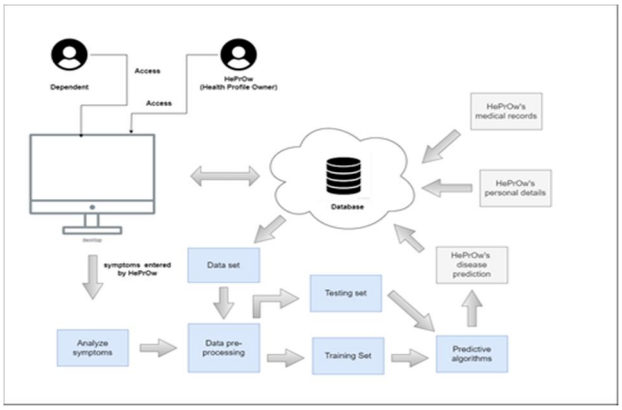


Figure 1 System diagram of Health profile management

The procedure for profile management could be listed as follows.

A. Create Health Profile

The web application gives users the chance to enter their personal details and profile picture to create a healthy profile by registering into the application. This component reads the user's personal details which are required to create the health profile.

B. Enter Medical Reports

After creating a health profile users can enter their medical records into the profile. They can add conditions by entering it. After the user enter conditions, if that condition started many years ago and now the user does not have that condition, they can remove that condition from the history by moving it. Mainly users need to enter their height, weight, blood group, blood sugar level, and blood pressure. Users can also add medicine prescriptions and laboratory reports as pdf or images to their profile. After entering records users can maintain records by updating or deleting them.

C. Health Prediction

In this phase, users can enter the symptoms that they have, and the system predicts the disease according to the symptoms entered by users, or the system predicts the disease by using users' medical records that are stored in their profile for several diseases. To identify the diseases, we have created a model using python. We used Python as a platform to run our Machine Learning algorithms for this model which will employ algorithms such as Decision Tree, Random Forest, Naive Bayes, and KNN to accurately forecast disease for early detection and better patient treatment. and we have collected a dataset from a Kaggle website. After collecting the dataset for this process, the model was trained by a sample dataset and tested the dataset.

D. Dependent Access

Here users can share their medical records with a dependent. Users can add dependents by adding their email address to the user profile using settings. Then dependents get access to view user medical records as well as a user also can see the dependent list from their profile.

The prediction model for predicting the doctor's consultation fees was built by focusing on two parameters doctors' education qualification and doctors' experience. A regression algorithm was used for the prediction model.

To extract the texts from laboratory reports, laboratory reports are converted into grayscale. After converting to grayscale images become less complex to process as they will have only two values 0 and 1. Then the converted grayscale image is cropped and sent through the pytesseract library. Pytesseract library extracts the texts from cropped images. The extracting of text is based on one selected lab report type "Complete blood picture".

Figure 2 Example lab report used for extracting text

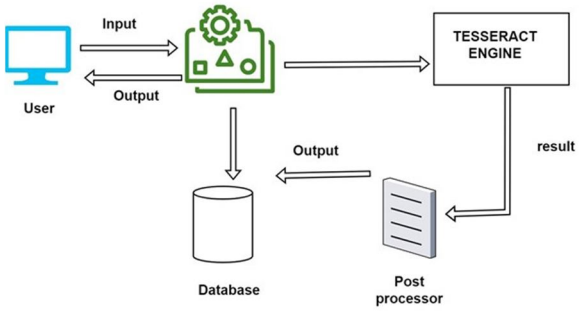


Figure 3 Text extraction method from laboratory reports

Furthermore, the other features of this application are sending laboratory prescriptions to nearby pharmacies and allowing users to book appointments with doctors. The Run Your health profile is an application programming interface that follows the REST architectural style and interacts with Restful web services

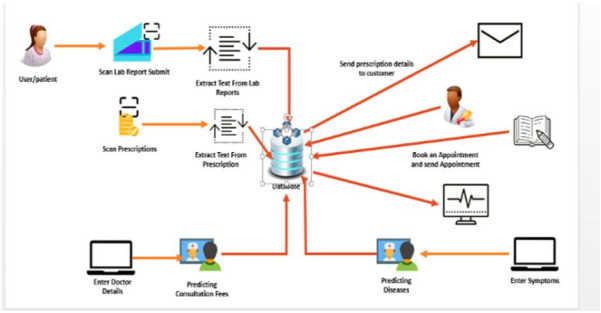


Figure 4 System diagram of the application

IV. RESULTS AND DISCUSSION

In the Health profile management component, we were able to store our medical records and we can predict our health by adding symptoms. When it comes to predicting disease there are two ways to predict. The first one is by adding symptoms and the other one is taking the medical records that are stored in the profile. Users need to enter symptoms into the system, a minimum of 3 symptoms, and a maximum of 17 symptoms that users can enter then they need to enter the prediction button then they will get the disease that they have.

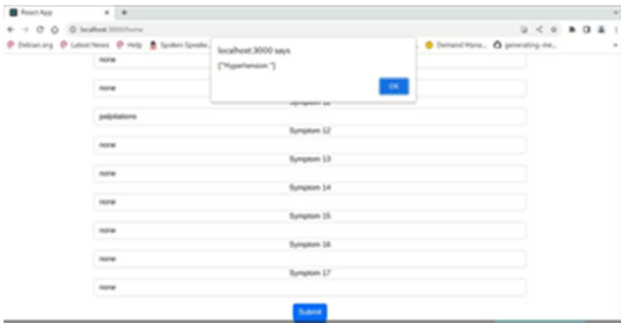


Figure 5 Disease prediction

This application successfully extracts the texts in electronic medical laboratory reports, by detecting the characters using the proposed image processing methodology and then by processing them to recognize characters for machine-readable versions using easyOCR.

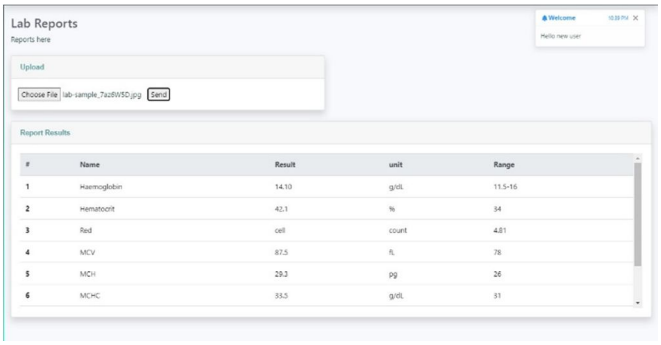


Figure 6. Text extraction of medical laboratory reports

For predicting the consultation fees, Dataset training was done with a data set of 1000 records taken from the data flair website. The accuracy of the dataset was done through removing unnecessary records. And The other purpose of this research is to provide an interface to send laboratory appointments to the nearby laboratories.

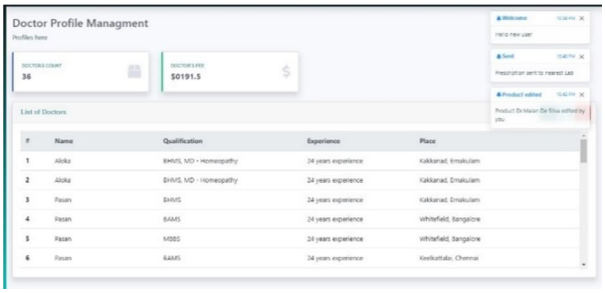


Figure 7 Prediction of Doctors' consultation fees

V. CONCLUSION

In Sri Lanka, there is no application available to store individual health records digitally. RUN YOUR HEALTH provide its users to store their health records digitally. As future tasks, It is expected to develop the application to extract texts from many types of lab reports and predict consultation fees of doctors by considering many parameters.

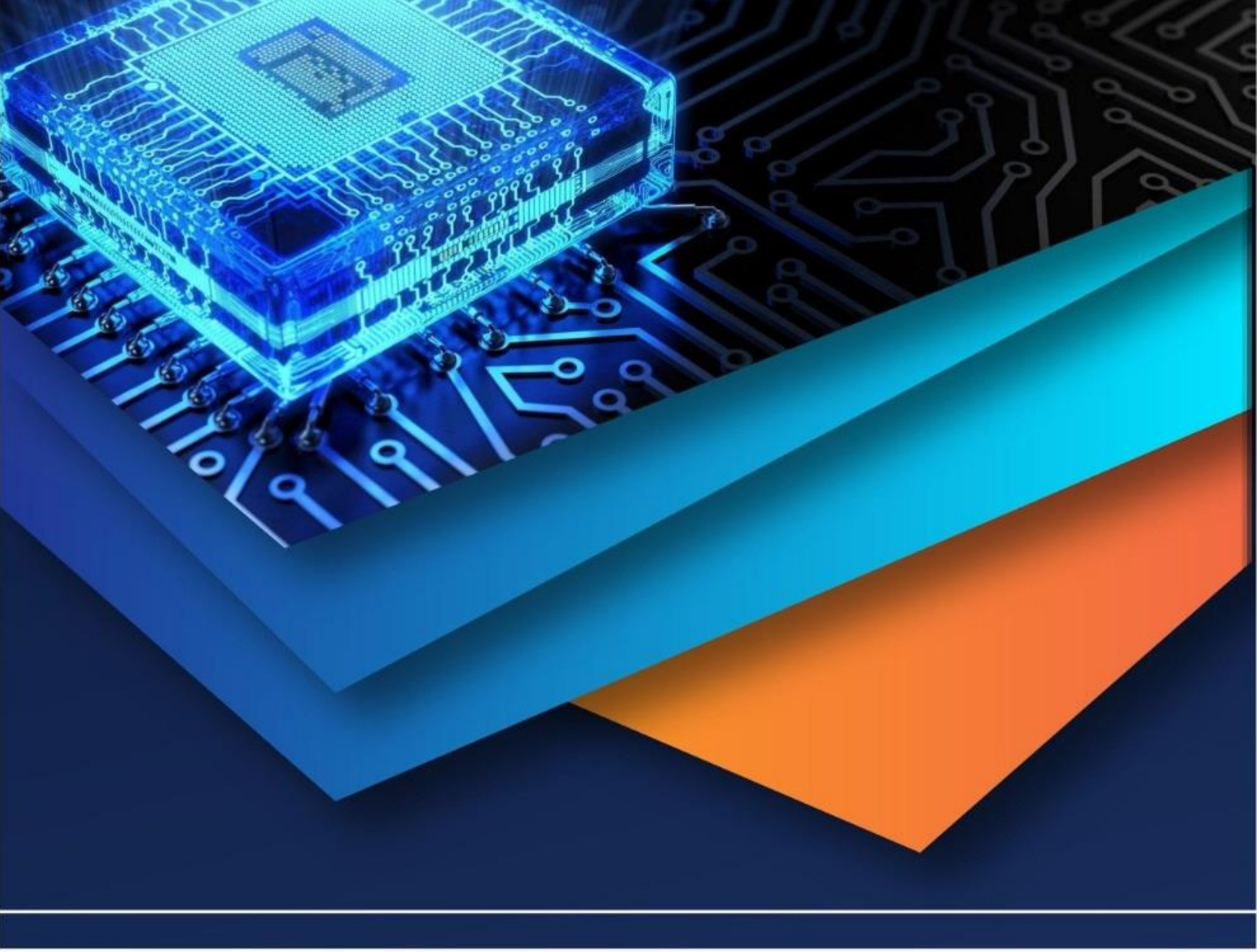
The medication and the dosage of that medication are primarily sent to the patient via the digitalized health profile. The patient can then readily identify the medications their doctor has prescribed for their specific ailment. Additionally, their health profile was simultaneously updated to reflect both the disease they had and the recommended medication. The patient does not have to go into detail about their past medical history and the nature of their illness when they next visit the doctor. because a doctor can see the patient's past medical history and all the medications, they were prescribed for each and every disease they had by entering their name. With this digitalized health profile, the patient can receive remedies to their health difficulties without any hassle even if they see a doctor other than their regular medical consultant.

VI. ACKNOWLEDGMENT

The authors of this research paper would like to thank the Sri Lanka Institute of Information Technology (SLIIT) for the encouragement and for providing better guidance to pursue this research.

REFERENCES

[1] S. Srivastava, P. Khurana, A. Rai, A. Cheema and . P. Srivastava, "High performance and adaptive lab report generation in hospital management information systems," in 2016 IEEE Annual India Conference (INDICON), Vancouver, 2016.
[2] W. Xue, Q. Li and . Q. Xue, "Text detection and recognition for images of medical laboratory reports with a deep learning approach," in IEEE Access, vancouver, 2019 .
[3] W. Xue , Q. Li , Z. Zhang , Y. Zhao and H. Wang , "Table analysis and information extraction for medical laboratory reports.," in 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech, Vancouver, 2018.
[4] Z. Wang, S. Sugaya and D. Nguyen, "Salary prediction using bidirectional-GRU-CNN model," in Assoc. Nat. Lang. Process, Vancouver, 2019
[5] S. Das, R. Barik and A. Mukherjee, "Salary prediction using regression techniques," in Industry Interactive Innovations in Science, Engineering & Technology, Vancouver, 2020.
[6] Matkowski and Michael, "Prediction of Individual Income: A Machine Learning Approach," Chicago, 2021.
[7] D. Chicco, M. Warrens and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," in PeerJ Computer Science, Vancouver, 2021.
[8] A. Varshini, K. Kumari, D. Janani, and S. Soundariya, " Comparative analysis of Machine learning and Deep learning algorithms for Software Effort Estimation.," Vancouver, 2021
[9] R. Werdhani, "Leadership in the doctor-patient relationship: Implementation on patient's case management in primary care.," in Medical Journal of Indonesia, Vancouver, 2017.
[10] Intawong, Kannikar, Parichat Ong-artborirak, and Waraporn Boonchieng. "Seamless Electronic Medical Record for Health Management System in Emergency Patients." 2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer, and Telecommunication Engineering. IEEE, 2021.
[11] James, A., Y. Wilcox, and R. N. G. Naguib. "A telematic system for oncology based on electronic health and patient records." IEEE Transactions on Information Technology in Biomedicine 5.1 (2001): 16-17.
[12] Pingale, Kedar., Surwase, Sushant., Kulkarni, Vaibhav., Sarage, Saurabh., &Karve, Prof. Abhijeet.(2019). Disease Prediction using Machine Learning.International Research Journal of Engineering and Technology, 6(12), 2810-2813.
[13] W. Xue , Q. Li , Z. Zhang , Y. Zhao and H. Wang , "Table analysis and information extraction for medical laboratory reports.," in 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech, Vancouver, 2018.
[14] J. Jayoma , E. Moyon and E. Morales , "OCR based document archiving and indexing using PyTesseract: A record management system for dswd caraga, Philippines," in 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM) , Vancouver, 2020.
[15] A. Ramadhani, N. Kim and H. Choi , "Predicting Employment Earning using Deep Convolutional Neural Networks," Journal of Digital Convergence, vol. 16(6), pp. 151-61, 2018.
[16] Y. Matbouli and S. Alghamdi , "Statistical Machine Learning Regression Models for Salary Prediction," Economy Wide Activities and Occupations. Information. 2022 Oct 12, vol. 13(10), p. 495, 2022.
[17] P. Sundsøy, J. Bjelland B. Reme, A. Iqbal, and E. Jahani, "Deep learning applied to mobile phone data for individual income classification," in Artificial Intelligence: Technologies and Applications, Vancouver, 2016.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY