PREDICTION OF KAP IN HYPOTHYROIDISM

1st Kapsa Sharvani
dept of computer science and Artificial Intelligence
SR University Warangal
sharvanikapsa@gmail.com

2nd Kadam Shiva kumar
dept of computer science and Artificial Intelligence
SR University Warangal
kadamshiva702@gmail.com

3rd Rupesh kumar Mishra
dept of computer science and Artificial Intelligence
Sr university,Warangal
rupeshmishra80@gmail.com

***Abstract*—Hypothyroidism is a widespread health issue impacting millions globally, requiring effective management strategies.This paper focuses on predicting the Knowledge, Attitudes, and Practices (KAP) regarding hypothyroidism.Through an in-depth review of existing literature and analysis of relevant studies, this paper investigates the factors influencing individual's knowledge, attitudes, and practices towards hypothyroidism.Understanding these predictors can aid healthcare providers in customizing interventions to enhance patient education, fostering positive attitudes, and promote healthy practices in managing hypothyroidism.Ultimately, this approach aims to improve overall management and outcomes of hypothyroidism..**

Keywords—KAP (Knowledge, Attitudes, Practices) Hypothyroidism, LSTM, BERT.

I. INTRODUCTION

Understanding how people perceive and manage hypothyroidism is crucial for effective healthcare interventions and public health strategies.Hypothyroidism, a common endocrine disorder characterized by insufficient thyroid hormone production, impacts millions worldwide. Knowledge, Attitudes, and Practices (KAP) regarding hypothyroidism play pivotal roles in disease management and prevention of complications. Therefore, investigating KAP levels among affected populations can provide valuable insights into knowledge gaps, attitudes, and behaviors influencing hypothyroidism management.This research aims to explore KAP related to hypothyroidism, utilizing advanced natural language processing techniques and diverse textual sources to gain a comprehensive understanding of this condition. Through this exploration, we endeavor to inform healthcare providers, policymakers, and public health practitioners to develop targeted interventions and strategies for better hypothyroidism management and improved patient outcomes.

II. RELATED WORDS:

The study investigated the effect of pharmacist-provided counselling on knowledge, attitude, and practice outcomes in hypothyroid patients in Nepal. Hypothyroidism is a common chronic endocrine disorder that affects about 30%

of the population in the eastern region of the country. Lack of awareness can lead to poor disease management. Education is one of the most important factors that determine awareness,hence, the present study aimed to investigate whether a person's education, age, and gender influence awareness about the thyroid gland. The study showed positive results for education and treatment.[1]

The study investigated that it's amazing how pharmacist provided counselling can have such a positive impact on education and treatment outcomes for hypothyroid patient la in Nepal with hypothyroidism being a common chronic endocrine disorder in the eastern region of the country it's crucial to raise awareness and improve disease management education plays a vital role in empowering individuals to understand their conditions better and make informed decisions about their health it's great to see that this study focused on investigating the influence of education age and and gender on awareness about the thyroid gland by identifying these factors we can work towards creating targeted educational programs that cater to specific demographic overall the study highlights the importance of education and it's positive impact on the management of hypothyroidism[2]

The study about Hypothyroidism in Moradabad north India it's fascinating to learn purpose of evaluating the knowledge attitude and practices of female in the region the findings indicating that hypothyroidism is more prevalent in women, age 30 to 45 is quite significant this age, groups seem to be particularly vulnerable to the condition, the study result highlight the need for better public education and awareness to improve knowledge about hypothyroidism and promote treatment adherence by increasing awareness individuals can better understand the importance of early detection and proper management of hypothyroidism this can lead to improved health outcomes for those affected it's crucial to Provide accessible and accurate information to empower women in the community to take charge of their health and seek appropriate medical care through education and awareness camping.[3] We work towards reducing the burden. Of hypothyroidism and ensuring better health for all

This research about the hypothyroidism in Chennai. It sounds like the researchers wanted to assess the knowledge, practice and treatment adherence of patients with hypothyroidism in an endocrinology department.[4] They recruited 120 patients, both male and female, and used questioners are interviews to gather data. The results showed that 61.7 % of patients had moderately adequate knowledge about hypothyroidism, 76.7 % had adequate practice, but only 49.2% had low adherence to treatment.[5] The mean score for knowledge on hypothyroidism was 16.69, for practice it was 8.28, and for treatment adherence it was 2.09. this gives us some insights into the overall understanding and behavior of patients in the managing their condition. Hypothyroidism is a condition where the thyroid hormone doesn't function optimally, affecting various body functions. In India, the prevalence of hypothyroidism is around 10.95 %, with sub-clinical hypothyroidism(SCH)

being present in 8.5 % of cases.[5]According to a report by the world health organization , around 83.2% of urban and 66.1 % of rural Indians were found to have adequate iodine intake, which is important for thyroid health. The study also found that 16.8 % of hypothyroid patients felt that treatment could be discontinued once their reports normalized, with 6.55 % continued treatment until their symptoms subsided. This highlights the need for clear communication between patients and healthcare providers to ensure proper understanding and adherence to treatment protocols.[7]

The research on the use of machine learning and deep learning techniques in predicting thyroid disease literature survey using peer-reviewed article databases is great way to gather relevant information.[10] It's interesting to see that different methods have been employed for thyroid cancer detection and prediction, each with varying levels of accuracy. One study used the least absolute shrinkage and selection operator(LASSO) and logistic regression (LR) models to select malignant thyroid nodules, while another study utilized logistic Lasso regression(LLP) with random forest(RF). these methods seem to be effective in identifying potential cancerous nodules.[12,13,14,15] It's fascinating that machine learning algorithms such as random forest, logistic regression, gradient boosting machines, support vector machine and deep neural networks have been employed to predict high probable molecules in thyroid hormone homeostasis. This early prediction of molecules can be incredibly valuable for further testing and inter venting in the early stages of thyroid disease. The classification of thyroid disease in indeed a significant challenge in the healthcare industry.[16,17] It's great to see that the authors obtained datasets from the university of California machine learning library to assess the performance of different algorithms in selective feature optimization. This approach can help in finding the most effective algorithms for accurate classification.[18,19,20,21]

The research on the use of machine learning and deep learning techniques in predicting thyroid disease literature survey using peer-reviewed article databases is great way to gather relevant information.[10] It's interesting to see that different methods have been employed for thyroid cancer detection and prediction, each with varying levels of accuracy. One study used the least absolute shrinkage and selection operator(LASSO) and logistic regression (LR) models to select malignant thyroid nodules, while another study utilized logistic Lasso regression(LLP) with random forest(RF). these methods seem to be effective in identifying potential cancerous nodules.[12,13,14,15] It's fascinating that machine learning algorithms such as random forest, logistic regression, gradient boosting machines, support vector machine and deep neural networks have been employed to predict high probable molecules in thyroid hormone homeostasis. This early prediction of molecules can be incredibly valuable for further testing and inter venting in the early stages of thyroid disease. The classification of thyroid disease in indeed a significant challenge in the healthcare industry.[16,17] It's great to see that the authors obtained datasets from the university of California machine learning library to assess the performance of different algorithms in selective feature optimization. This approach can help in finding the most effective algorithms for accurate classification.[18,19,20,21]

Few studies have used about ML and DL methods to predict the thyroid illness the recent developments in data processing and computing.[31] the toxicity forecaster database research publications provided training data sets. Along with statistical methodologies and ml methods including random forest and svm (support vector machine) artificial neural networks, three data balancing process were tested [31]. the classifiers had 83% and 81% of F1-scores on the hold out datasets and they were used two thyroid datasets one from the knowledge extracting based on evaluatory learning(KEEL) one from the Pakistani hospital using the various distance measures KNN model was performed [31] and it utilized naive Bayes, KNN, SVM and LR to classify the data by comparing the prediction performance of KNN, decision tree and logistic regression where shown the highest accuracy with XGB [32] by this research paper we came to conclude that classifiers with L1 based features selections had great efficiency. The highest results of thyroid prediction was ANN with 98% and accuracy with 95.7% and F1 score was 95.7% [32].

The authors used the wavelet support vector machine and generalized discriminant analysis to develop an expert system for thyroid disease diagnosis and achieved the classification accuracy of about 91.86%. Another study combined the artificial immune recognition system with fuzzy weighted preprocessing and achieved an accuracy of 85 %. then there was a study that used neural networks and got accuracy rates of 93.19 %, 94.81% and 90.05 % using different types of neural networks. And finally , a multi layer perceptron trained with back propagation algorithm achieved an impression accuracy of 96.8 %. [36]

III. METHODOLOGY

The research analysis took place in python using google co-lab notebook as python is more adaptable and can be used for wider range of tasks. Here text preprocessing is done and text preprocessing is a crucial step in natural language processing (NLP) that entails cleaning and preparing text data for analysis or modeling.

a. Data Collection: conduct the survey with the chosen sample when the survey instrument complete. This might entail making phone calls, delivering paper surveys, sending out emails with survey links. Give respondents precise instructions on how to complete the survey and explain the goal and privacy of the study.

b. Data Analysis: descriptive data for summarizing patient attributes and the frequency of renal abnormalities will be included in the statistical analysis. After controlling for possible con-founders, uni variate and multivariate logistic regression analysis will be performed to determine predictive variables for kidney abnormalities in hypothyroid individuals. We will analyze the receiver operating characteristic(ROC) curve to evaluate the selected predictors' discriminating power. To investigate relationships

stratified by age, gender and length of hypothyroidism subgroup analysis will be carried out.

c. Text preprocessing:In natural language processing(nlp)text preprocessing describes the operations used to clean and alter unprocessed text input prior to analysis or modeling. Tokenizing the text into individual words or tokens, deleting stop words, changing the text to lowercase, and removing punctuation are some of the activities involved. The text input is more standardized and more suited for NLP tasks like sentiment analysis,text categorization, and language development thanks to these preparation procedures.

d. Import Libraries: the code imports necessary libraries such as torch,data-loader,berttokenizer,Bert for sequence classification, adamW,train test split from sklearn and accuracy score from sklearn.metrics. These libraries provide the tools needed for loading and training the BERT model

e. Load pre-trained BERT Tokenizier: the pre-trained BERT Tokenizier(Bert-base-uncased) is loaded from the hugging face model hub. This Tokenizier is responsible for converting text inputs into tokens that the BERT model can understand.

f. Create attention masks: Attention masks are created to indicate which tokens are actual words and which are padding tokens. This helps the model focus on the relevant parts of the input sequence during training.

g. Split data: the datasets is split into training and validation sets using the train test split function from sklearn.Model selection. This allows for evaluating the model's performance on unseen data .

h. Create data loaders: PyTorch data loader objects are created for both the training and validation sets. These data loaders enable efficient loading of data in batches during training, which helps speed up the training process.

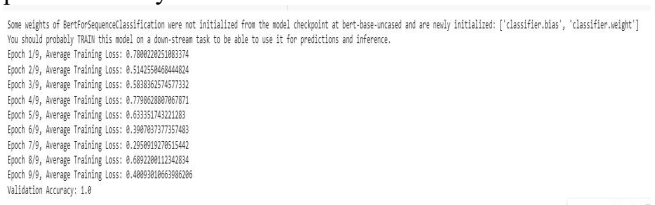
i. Load pre-trained BERT model: the pre-trained BERT model("Bert base uncased") for sequence classification is loaded using the BertForSequence classification class for the hugging face model hub. This model has been pre-trained on a large corpus of text data and can be fine tuned for specific classification tasks

j. Validation: After each epoch the models performance is evaluated on the validation.

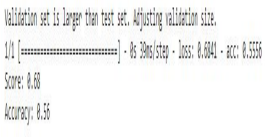
k. The process of preparing data, building, training, and evaluating an Lstm model for this series forecasting. However, it's important to note that this is just a basic example, and further tuning and optimization may be required depending on the specific problem and datasets. Additionally evaluating the model's performance on the test set and fine-tuning hyper parameters are essential steps for model improvement.

IV. RESULTS

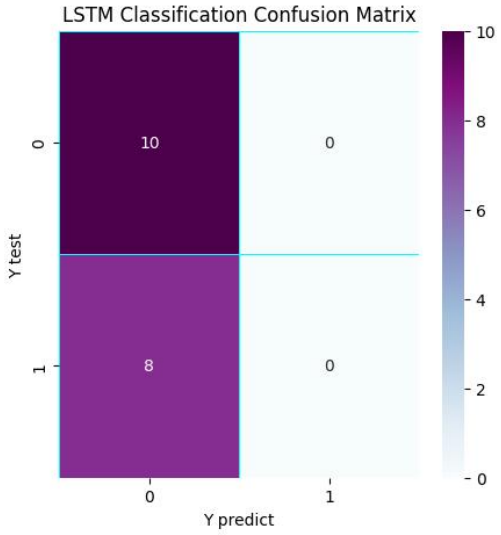
The code loads a dateset which containing the information about the age and gender from the CVS file and performs some data preprocessing steps like data cleaning, tokenization, and splitting into training and testing sets.
The results of Bert process show the model is able to achieve a validation accuracy of 100 %, indicating that it performed very well on the validation data.



And Lstm process show the model is able to achieve the accuracy of 56% indicating that it performed well on the dateset we provided.



Confusion matrix tells us about the visualization depicts the performance of the lstm model in predicting gender based on age-related textual data. It showcase the distribution of predicted and actual gender labels, highlighting the correct predictions. The heat-map illustrates and the model's accuracy and errors, providing insights into the classification performance.



V. Conclusion

After using some cool tools like BERT and LSTM .we have learned a lot about the hypothyroidism and how people feel and know about it. We collected data by survey the people and we analyzes the data to see what people knew about it and how they felt about hypothyroidism. We preprocessed all the text we collected to make easier for the computer to understand. We turned the words into tokens,removed unnecessary stuff like punctuation,and made everything the same lowercase. After that, we imported some libraries and loaded a pre-trained BERT Tokenizier. This helped us understand the text even better. We split our data into training and validation sets and created data loaders to help the computer learn faster. Then we loaded a pre-trained BERT model and trained it on our data. After each training session, we checked how well it was doing on our validation set to make sure it was learning correctly. Lastly we had using about an LSTM model for time series forecasting, which is like predicting the future based on past patterns. By analyzing people's knowledge and attitudes and using advanced techniques, we can help healthcare providers make better plans for managing hypothyroidism and improving patient outcomes .

VI. References

1 Unnikrishnan AG, Kalra S, Sahay RK, Bantwal G, John M, Tewari N. Prevalence of hypothyroidism in adults: An epidemiological study in eight cities of India. Indian J Endocrinol Metab. 2013;17(4):647-652.

2 Maskey R, Shakya DR, Baranwal JK, Lavaju P, Karki P, Poudel SK. Hypothyroidism in diabetes mellitus patients in Eastern Nepal. Indian J Endocrinol Metab. 2015;19(3):411-415.

3 Ross DS, Mana DL, Bruno OD, Wartofsky L. Coma mixedematoso [Myxedema coma] Medicina (B Aires). 2012:77.

4 Kumar P, Khandelwal D, Mittal S, Dutta D, Kalra S, Katiyar P, Aggarwal V. Knowledge, awareness, practices and adherence to treatment of patients with primary hypothyroidism in Delhi. Indian J Endocrinol Metab. 2017;21(3):429-433.

5 Dew R, King K, Okosieme OE, Pearce SH, Donovan G, Taylor PN, Hickey J, Dayan CM, Leese G, Razvi S, Wilkes S. Attitudes and perceptions of health professionals towards management of hypothyroidism in general practice: a qualitative interview study. BMJ Open. 2018;8(2):1-9.

6 Ma C, Xie J, Huang X, Wang G, Wang Y, Wang X, Zuo S. Thyroxine alone or thyroxine plus triiodothyronine replacement therapy for hypothyroidism. Nucl Med Commun. 2009;30(8):586-593. 7. Sanii Y, Torkamandi H, Gholami K, Hadavand N, Javadi M. Role of pharmacist counseling in pharmacotherapy quality improvement. J Res Pharm Pract. 2016;5(2):132-137.

7 Maharjan S, Chhetri HP. Assessment of interventions by pharmacist in improving knowledge, attitude and practice towards hypothyroidism among the patients attending at an endocrine clinic in Nepal. Indian Journal of Pharmacy Practice. 2015;8(2):67-71.

8 Charan J, Biswas T. How to calculate sample size for different study designs in medical research?. Indian J Psychol Med. 2013;35(2):121-126.

9 Perumal SS, Prasad S, Surapaneni KM, Joshi A. Health ınformation-seeking behavior among hypothyroid patients at saveetha medical college and hospital. Ethiop J Health Sci. 2015;25(2):147-154.

10 Maheshwari P, Mohan R, Shanmugarajan TS. KAP Study on thyroid disorders (hypothyroidism and hyperthyroidism) in a tertiary care hospital. J Res Pharm Tech. 2017;10(1):41-43.

11 Goel A, Shivaprasad C, Kolly A, Pulikkal AA, Boppana R, Dwarakanath CS. Frequent occurrence of faulty practices, misconceptions and lack of knowledge among hypothyroid patients. J Clin Diagn Res. 2017;11(7):15-20.

12 Sethi B, Khandelwal D, Vyas U. A cross-sectional survey to assess knowledge, attitude, and practices in patients with hypothyroidism in India. Thyroid Res Pract 2018;5:15-22

13 Farling, P. A. (2000). Thyroid disease. British Journal of Anaesthesia, 85(1), 15-28.

14 Patil, N., Rehman, A., Jialal, I., & Saathoff, A. D. (2021). Hypothyroidism (Nursing).

15 Akgül, G., Çelik, A. A., AYDIN, Z. E., & ÖZTÜRK, Z. K. (2020). Hipotiroidi Hastalığı Teşhisinde Sınıflandırma Algoritmalarının Kullanımı. Bilişim Teknolojileri Dergisi, 13(3), 255-268.

16 Raju, K. B., Lakineni, P. K., Indrani, K. S., Latha, G. M. S., & Saikumar, K. (2021, October). Optimized building of machine learning technique for thyroid monitoring and analysis. In 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC) (pp. 1-6). IEEE

17 Dietterich, T. G. (2000, June). Ensemble methods in machine learning. In International workshop on multiple classifier systems.Springer, Berlin, Heidelberg. pp. 1-15.

18 Tyagi, A., Mehra, R., & Saxena, A. (2018, December). Interactive thyroid disease prediction system using machine learning technique. In 2018 Fifth international conference on parallel, distributed and grid computing (PDGC) (pp. 689-693). IEEE

19 Knudsen, N., Bulow, I., Jorgensen, T., Laurberg, P., Ovesen, L., & Perrild, H. (2000). Comparative study of thyroid function and types of thyroid dysfunction in two areas in Denmark with slightly different iodine status. European journal of endocrinology, 143(4), 485-491.

20 Magri, F., Buonocore, M., Oliviero, A., Rotondi, M., Gatti, A., Accornero, S., ... & Chiovato, L. (2010). Intraepidermal nerve fiber density reduction as a marker of preclinical asymptomatic small-fiber sensory neuropathy in hypothyroid patients. European journal of endocrinology, 163(2), 279

21 UCI Machine Learning Repository: Thyroid Disease Data Set. (n.d.). Archive.ics.uci.edu. Retrieved July 27, 2023

22 Setian, N. (2007). Hypothyroidism in children: diagnosis and treatment. Jornal de pediatria, 83, S209-S216.

23 Bloice, M. D., & Holzinger, A. (2016). A tutorial on machine learning and data science tools with python. Machine Learning for Health Informatics: State-of-the-Art and Future Challenges, 435-480

24 Pam. (2016, September 26). 5 Pioneers in the Field of Artificial Intelligence Computer Science Degree Hub. Computer Science Degree Hub; Paul Csogi

25 Almahshi, H. M., Almasri, E. A., Alquran, H., Mustafa, W. A., & Alkhayyat, A. (2022, May). Hypothyroidism prediction and detection using machine learning. In 2022 5th international conference on engineering technology and its applications (IICETA) (pp. 159-163). IEEE

26 Alyas, T., Hamid, M., Alissa, K., Faiz, T.,Tabassum, N., & Ahmad, A. (2022). Empirical method for thyroid disease classification using a machine learning approach. BioMed Research International, 2022 Benvenga, S., Tuccari, G., Ieni, A., & Vita, R. (2018). Thyroid gland: anatomy and physiology. Encyclopedia of Endocrine Diseases, 4,

27 A. Shrivas and P. Ambastha, "An ensemble approach for classification of thyroid disease with feature optimization,"Inter-national Education and Research Journal, vol. 3, no. 5, pp. 1–4, 2019.

28 G. Chaubey, D. Bisen, S. Arjaria, and V. Yadav, "Thyroid disease prediction using machine learning approaches,"National Academy Science Letters, vol. 3, pp. 128–133, 2021

29 A. Dewangan, A. Shrivas, and P. Kumar, "Classification of thyroid disease with feature selection technique,"International Journal of Engineering & Technology, vol. 2, no. 3, pp. 128–133, 2016.

30 A. Begum and A. Parkavi, "Prediction of thyroid disease usingdata mining techniques,"in International Conference onAdvanced Computing & Communication Systems (ICACCS),pp. 342–345, Coimbatore, India, 2019.

31 J. C. Naafs et al., "Cognitive and motor outcome in patients with early-detected central congenitalhypothyroidism compared with siblings,"J. Clin. Endocrinol. Metab., vol. 106, no. 3, pp. e1231–e1239, Mar.2021. doi: 10.1210/clinem/dgaa901.

32 P. Duggal and S. Shukla, "Prediction of thyroid disorders using advanced machine learning techniques,"in 2020 10th Int. Conf. Clo. Comp. Data Sci. Eng., Noida, India, Jan. 29–31, 2020, pp. 29–31. doi:10.1109/Confluence47617.2020.9058102.

33 K. Polat, S. Şahan, and S. Güneş, "A novel hybrid method based on artificial immune recognition system (AIRS) with fuzzy weighted pre-processing for thyroid disease diagnosis," Expert Systems with Applications, vol. 32, no. 4, pp. 1141-1147, 2007

34 F. Temurtas, "A comparative study on thyroid disease diagnosis using neural networks," Expert Systems with Applications, vol. 36, no. 1, pp. 944-949, 2009.

35 B. M. Casey and K. J. Leveno, "Thyroid disease in pregnancy," Obstetrics & Gynecology, vol. 108, no. 5, pp. 1283-1292, 2006.

36 A. H. Shahid and M. P. Singh, "Computational intelligence techniques for medical diagnosis and prognosis: Problems and current developments,"Biocybernetics and Biomedical Engineering, 2019.