THYROID PREDICTION - SUMMARY

Paper 1

Title of the paper	Godara, Sunila, and Sanjeev Kumar. "Prediction of thyroid disease using machine learning techniques." *International Journal of Electronics Engineering* 10.2 (2018): 787-793.
Area of work	Prediction of thyroid disease
Dataset	Dataset was taken from UCI repository. The Thyroid dataset has 30 attributes and 3772 records.
Methodology / Strategy	Logistic regression and Support Vector Machine are compared on basis of Precision, Recall, F measure, ROC and RMS Error.
Algorithm	Logistic regression, Support Vector Machine
Result/Accuracy	Logistic Regression - 96.8452 SVM - 93.6108

Paper 2

Title of the paper	Kashyap, Suresh Kumar, and Neelam Sahu. "An Analysis of Predictive Models For Thyroid Disease Using Machine Learning Techniques. "*International Research Journal of Engineering and Technology (IRJET)* 8.03 (2021).

Area of work	Build prediction modeling of the given medical data of patients with and without thyroid
Dataset	Data set from UCI machine learning Repository that was used for implementation with 3772 instances of 23 independent attribute and 1 dependent attribute.
Methodology / Strategy	Describes the algorithm, language and software used for prediction
Algorithm	Random Forests, Naive Bayes
Result/Accuracy	The results were compared and it was seen that Random Forest classifiers is better than naïve bayes. Random forest classifiers could be successfully used to help the diagnosis of thyroid disease.

Paper 3

Title of the paper	Krishna. "Classification of Thyroid Disease using Machine Learning." *International Research Journal of Engineering and Technology* 8.02│ Feb (2021).
Area of work	compare different classification algorithms used in machine learning
Dataset	The dataset contains 3090 instances. In these 149 data comes under hypothyroid and 2941 data is negative cases. In total it has 25 features
Methodology / Strategy	Compare and study different classification algorithms used in machine learning

Algorithm	Naive Bayes, Support Vector Machine, k-Nearest Neighbors, Random Forest Classifier, Logistic Regression
Result/Accuracy	Logistic Regression - 96.1929% KNN - 95.5584% Random Forest Classifier - 95.5584% SVM - 95.4315% Naïve Bayes - 36.2944%

CONCLUSION

From the above three papers, we get to know that different approaches are used for thyroid prediction. From the observation it is found out that Logistic Regression is more accurate than the other algorithms with approximate accuracy of 96%.