**Literature Review**

1. **Multi Disease Prediction Using Data Mining Techniques.**  [1]

In this paper the authors used two different data mining techniques named Naïve Bayes and J48 decision tree for the prediction of various diseases such as heart disease, diabetes, and breast cancer and compared their performance for evaluating the best classifier. They found the highest accuracy for diabetes using J48. However, they didn’t provide the details of their dataset. They used WEKA tool but they did not show the results which were obtained by the percentage split and cross-validation separately.

1. **Survey on Data Mining Algorithms in Disease Prediction.** [2]

In this paper V. Kirubha and S. Manju Priya analyzed the application of the most popular data mining techniques in medical domain and they used some of the algorithms for disease prediction. In their study they showed that, by using various tools and techniques on different disease diagnosis a varieties number of results can be gained. However, they did not show the result which were obtained by using different data mining techniques.

1. **A Model for Predicting Ischemic Stroke Using Data Mining Algorithms.** [3]

The authors presented a study about the model of logistic regression and obtained the result with “XLSTAT” software. They showed several steps of the logistic regression model in their study. They have found the sensitivity rate was 77.58%, the specificity rate was 83.03% and the error rate was 19.7%.

1. **The development and implementation of stroke risk prediction model in National Health Insurance Service’s personal health record.** [4]

In this paper the authors presented the prediction of risk of stroke within 10 years.

And their dataset was within 1500000 men and 1200000 women had used data from the national health examination of the entire nation and they classified the total population in five ranges such as normal, slightly high, high, risky and very risky.

1. **Stroke Risk Prediction through Non-linear Support Vector Classification Models.** [5]

In this paper authors presented a study to find the possible risk of stroke by subjecting the risk factors to Support Vector Machines. They used 100 patient’s data with 8 attributes. They used Support Vector Classification model parameters through its kernel function named polynomial kernel and Gaussian (RBF) kernel. The authors evaluated the result through Confusion matrix and showed that the rate of the correctness of prediction by RBF was 98% whereby polynomial was 92%. So, the author told in this paper that the application of SVM models can be used for the processing of stoke-related risk factor data.

# Bibliography

|  |  |
| --- | --- |
| [1] | D. D. S. P. K.Gomathi Kamaraj, "ResearchGate," 17 September 2017. [Online]. Available: https://www.researchgate.net/publication/319851535\_Multi\_Disease\_Prediction\_using\_Data\_Mining\_Techniques. [Accessed 03 July 2019]. |
| [2] | S. M. P. V. Kirubha, "Survey on Data Mining Algorithms in Disease Prediction," *International Journal of Computer Trends and Technology (IJCTT) ,* vol. 38, pp. 124-128, 2016. |
| [3] | N. A. Balar Khalid, "A Model for Predicting Ishcemic Stroke Using Data Mining Algorithms," *IJISET- International Journal of Innovative Science, Engineering & Technology,* vol. 2, no. 11, pp. 18-23, 2015. |
| [4] | H.-s. L. Jae-woo-Lee, "The development and implementation of stroke risk prediciton model in National Health insurance service's personal health record.". |
| [5] | K. R. S. Sabibullah Mohamed Hanifa, "Stroke Risk Prediction Through Non-linear Support Vector Classification Models," *Internation Journal of Advanced Research in Computer Science,* vol. 1, pp. 47-53, 2010. |
| [6] |  |
| [7] | S. M. P. V. Kirubha, "Survey on Data Mining Algorithms in Disease Prediction," *International Journal of Computer Trends and Technology (IJCTT),* vol. 38, pp. 124-128, 2016. |
| [8] | H.-s. L. Jae-woo-Lee, "The development and implementation of stroke risk prediction model in National Health Insurance Service's personal health record," *Elsevier,* pp. 253-257, 2017. |