Experimental approach for the usage of Neural Network architectures in predicting Hypothyroidism.

Shreya Pande, Vikram Ghanote

Abstract:

Just as a small spark can ignite a large flame, a malfunctioning thyroid can trigger a series of health complications. A medical condition characterized by inadequate hormone production by the thyroid gland, can affect individuals across all age groups and demographics. Understanding the risks of thyroid and its detection relies primarily on blood tests, which pose challenges due to the complex nature of the data and the need for accurate predictions.

To address this, we aim to examine the efficacy of a few neural network architectures in predicting the risk of thyroid by some distinct hyperparameters.

Drawing upon the Hypothyroid dataset from the University of California, Irvine's machine learning repository we experimented with neural network models LSTM and sequential deep neural network architecture to identify the algorithm that yields the highest accuracy in forecasting hypothyroid outcomes. Given the dataset's predominantly imbalanced target variable classes, relying solely on accuracy scores becomes inadequate to gauge prediction performance. Consequently, our evaluation criteria encompass both accuracy and recall ratings.

We observed that the dense neural architecture having a accuracy of 0.97 and F1 score 0.97 was performing the best. LSTM and sequential models were also implemented with the F1 of 0.97 and 0.56 respectively.

Introduction:

The thyroid gland is the most conspicuous pure endocrine glands, located at the front of the neck and encircling the trachea (Beynon & Pinneri, 2016).

Hypothyroidism is a medical condition characterized by the underproduction of thyroid hormones by the thyroid gland, located in the neck. Thyroid hormones play a crucial role in regulating various bodily functions, including metabolism, heart rate, and temperature. When the thyroid doesn't create and release enough thyroid hormone into your bloodstream. This makes your metabolism slow down. Also called underactive thyroid, hypothyroidism can make you feel tired, gain weight and be unable to tolerate cold temperatures..

According to the American Cancer Society's most current forecasts, 44,280 new instances of thyroid cancer (12,150 men and 32,130 women) have occurred in the United States since 2021, and 1,050 men and 1,150 women have died of thyroid cancer. Between 2009 and 2018, the thyroid cancer fatality rate increased dramatically (0.6 percent per year) (Akbas et al., 2021)

Traditionally, diagnosis of hypothyroidism includes:

- Medical history and physical examination.
- Blood tests measure thyroid hormone levels (T3, T4) and thyroid-stimulating hormone (TSH).
- Radioactive iodine uptake (RAIU) test to measure the thyroid's ability to absorb iodine.
- Thyroid scintigraphy assesses thyroid gland structure and function.
- Ultrasound imaging to visualize the thyroid gland and detect nodules.

Developing neural network architectures for detecting hypothyroidism can offer several advantages in the diagnostic process. A carefully designed Neural network model might excel at learning patterns from data and making predictions.

Several aspects required in the detection of hypothyroidism can be trained in a good architecture, such as ‘Complex Data Patterns’, ‘Feature Learning’ and if required, Integration with Imaging.

Early detection plays an important role in all diseases. The 5-year survival rate for thyroid cancer is 97.8%, thanks to early diagnosis and improved treatments (Quang T. Nguyen, 2015)

These can help support medical decision making and reduce the risk of human error and burden on the medical system.

Background:

We found various studies being performed on a lot of thyroid datasets. There are currently 10 datasets in the UCI machine learning repository (one of which is used in this study). We used the hypothyroid dataset to perform our analysis as we found it to be a rare dataset of the 10 on which studies were conducted. Parikh et al. (2015) discuss the two prediction models they developed to solve their multiclass classification problem. They used artificial neural networks and support vector machines and achieved an accuracy of 97.17 percent with the ANN. Ozyilmaz & Yildirim (2002) pioneered the use of an artificial neural network to diagnose thyroid disease, and they examined many neural network models, including backpropagation (MLP with backpropagation).

PeerJ Comput Sci. 2022; 8: e898, used ANN to get the best results when implemented on Sick-euthyroid UCI dataset.

Approach:

The step-by-step approach to the problem set is described below:

Dataset Overview:

Of the ten datasets available on Thyroid in UCI machine learning repository(Dua & Graff, 2017), we are using the 'Hypothyroid Dataset'. It has 3163 rows and 26 columns available. We also observed that not many research papers and insights by code or algorithms were available where this particular dataset of the ten was used, hence, we decided to go ahead and explore this one.

The dataset is composed of two files: hypothyroid.names and hypothyroid.data. hypothyroid.names is a file that contains the names of the attributes and any associated data. The other data file contains the database's metadata, as well as the classes, attribute names, and any other data that may be present. The following table outlines the data set's general structure as available in the 'names' file.

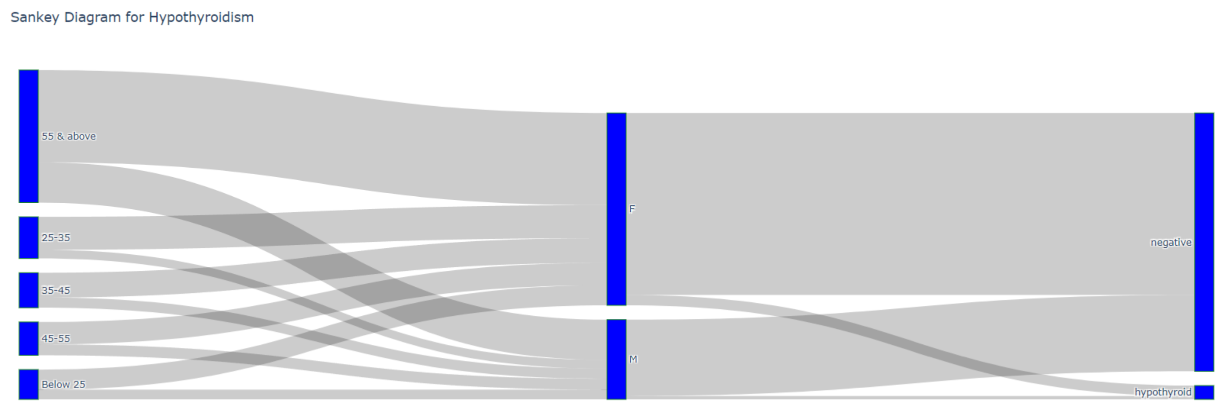
Table1:

```
hypothyroid, negative.

age:                               continuous,?.
sex:                               M,F,?.
on_thyroxine:                      f,t.
query_on_thyroxine:                f,t.
on_antithyroid_medication:         f,t.
thyroid_surgery:                   f,t.
query_hypothyroid:                 f,t.
query_hyperthyroid:                f,t.
pregnant:                          f,t.
sick:                              f,t.
tumor:                             f,t.
lithium:                           f,t.
goitre:                            f,t.
TSH_measured:                      f,t.
TSH:                               continuous,?.
T3_measured:                       f,t.
T3:                                continuous,?.
TT4_measured:                      f,t.
TT4:                               continuous,?.
T4U_measured:                      f,t.
T4U:                               continuous,?.
FTI_measured:                      f,t.
FTI:                               continuous,?.
TBG_measured:                      f,t.
TBG:                               continuous,?.
```

In order to closely observe the easier patterns of our dataset, we applied some visualization techniques to get a bird's eye view. The segregation of people with 'Hypothyroid' and "Negative' seemed quite biased as shown in the figure below:

Figure1:



Therefore, up sampling is required to make the unbalanced dataset a balanced one.

The new examples can be synthesized from current ones (Han, Wang & Mao, 2005). This technique is referred to as the Synthetic Minority Oversampling Technique or SMOTE for short, and it is used to augment data for the minority class (Chawla et al., 2002). This research makes use of the SMOTE class implementation provided by the imbalanced-learn Python library. Like a scikit-learn data transform object, the SMOTE class must be defined and configured prior to being fitted to a dataset and applied to create a new modified version of the dataset.

Architecture Implementation:

<u>Deep Sequential Neural Networks[Denoyer, Ludovic, 2014]:</u>

A Sequential Neural Network is a type of neural network architecture that's organized in a sequential manner, where each layer feeds its output as input to the next layer in a linear fashion

<u>MLP (Multi-Layer Perceptron)/ Dense Neural Networks[Li, Qian, et al]:</u>

The term "Dense Neural Network" typically refers to a neural network architecture where each neuron in a layer is connected to all neurons in the following layer. Also known as a fully connected neural network, it's a common configuration in neural networks like MLPs. A Dense Neural Network and a Multi-Layer Perceptron (MLP) are often used interchangeably to refer to the same type of neural network architecture. Both terms describe a neural network structure that consists of multiple layers of neurons, where each neuron is densely connected to the neurons in the adjacent layers. When implementing these architectures using libraries like Keras or TensorFlow, the terms "Dense Neural Network" and "MLP" are often used interchangeably.

<u>LSTM (Long Short-Term Memory):</u>

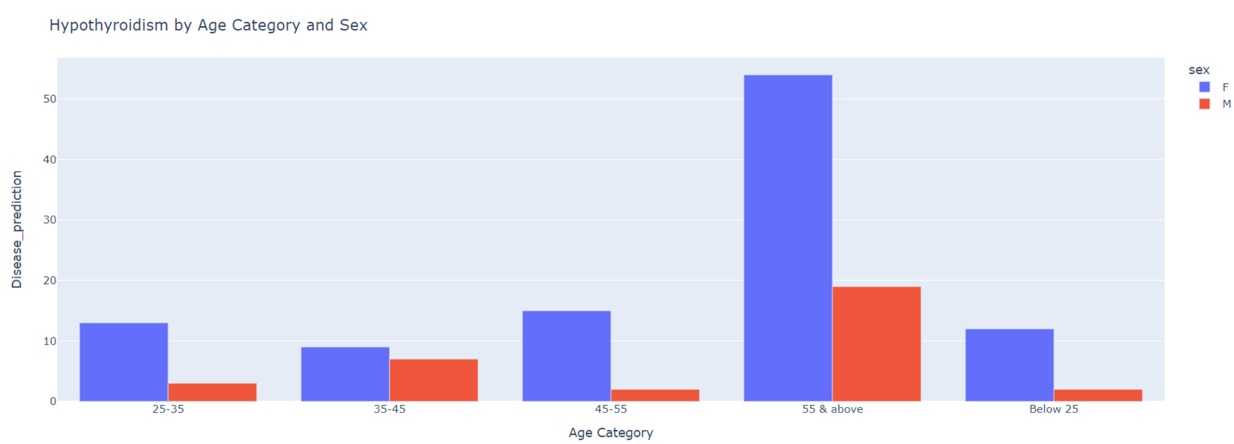
It's a type of recurrent neural network (RNN) architecture that's designed to handle sequence data and overcome some of the limitations of traditional RNNs. LSTMs are designed to work with sequential or time-series data where maintaining context over longer sequences is important. They are capable of learning and remembering patterns in sequences, making them effective for tasks like natural language processing, speech recognition, and stock price prediction.

The above architectures were used with a combination of various depts, dropout rates, activation, loss and optimization combination to observe different results.

Results:

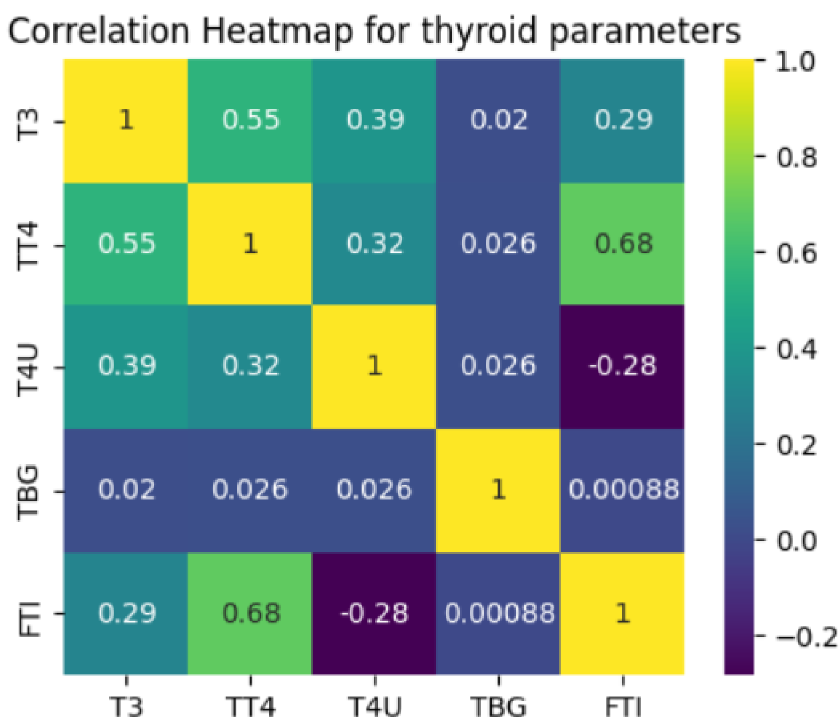
We viewed our data at glance and found that, more females are affected by hypothyroidism than males. Women in the age group 55 and above are most affected.

Figure2:



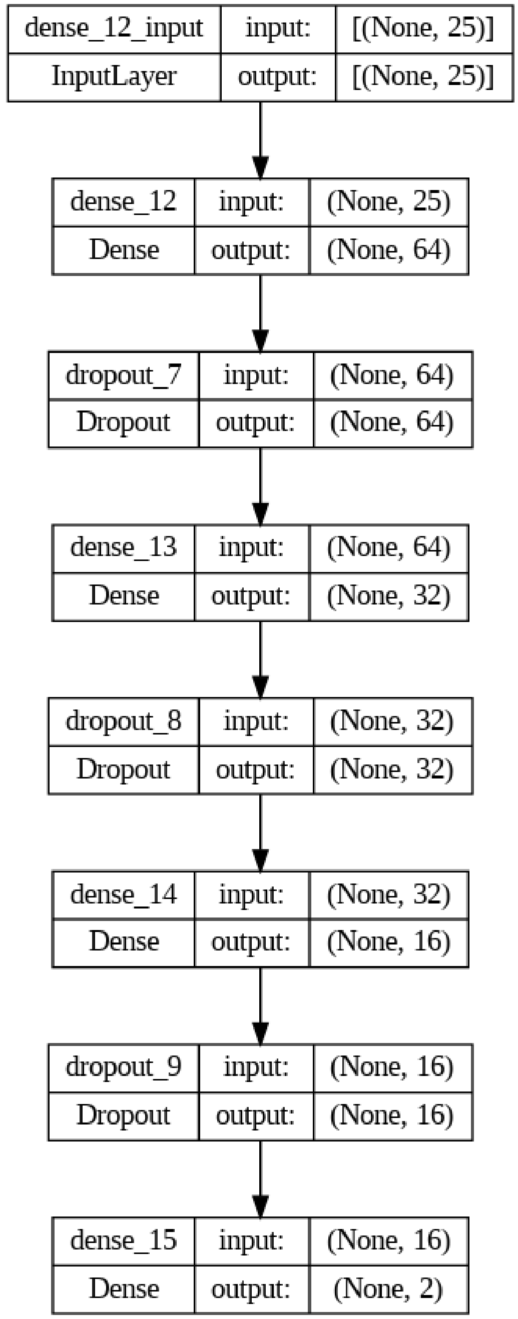
The parameters determining hypothyroidism, such as TBG, TT4, FTI are correlated in the dataset used as below:

Figure 3:

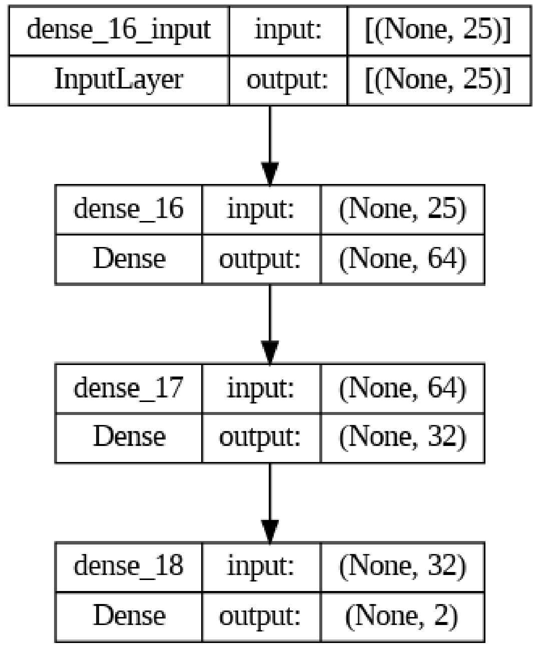


The model structure used can be viewed below:

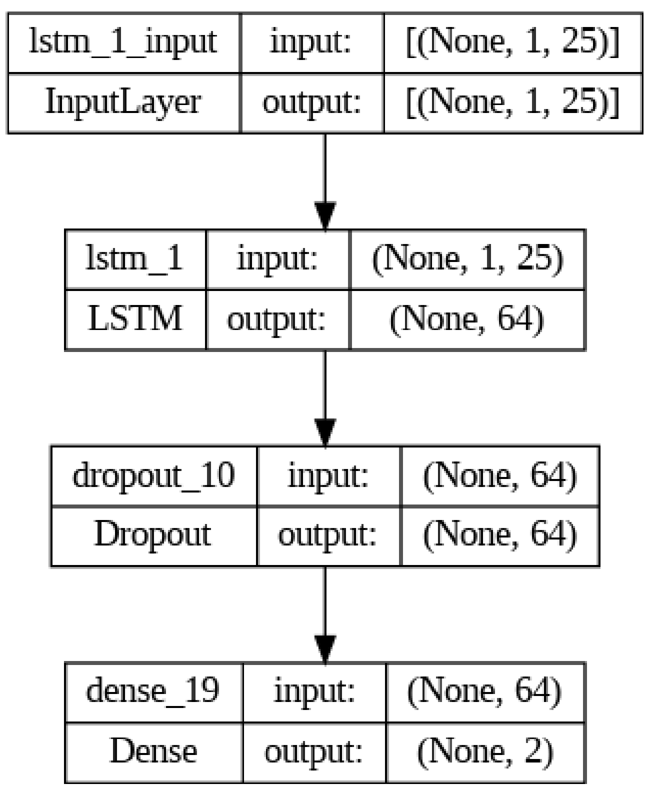
Dense Neural Network [Figure 4]:



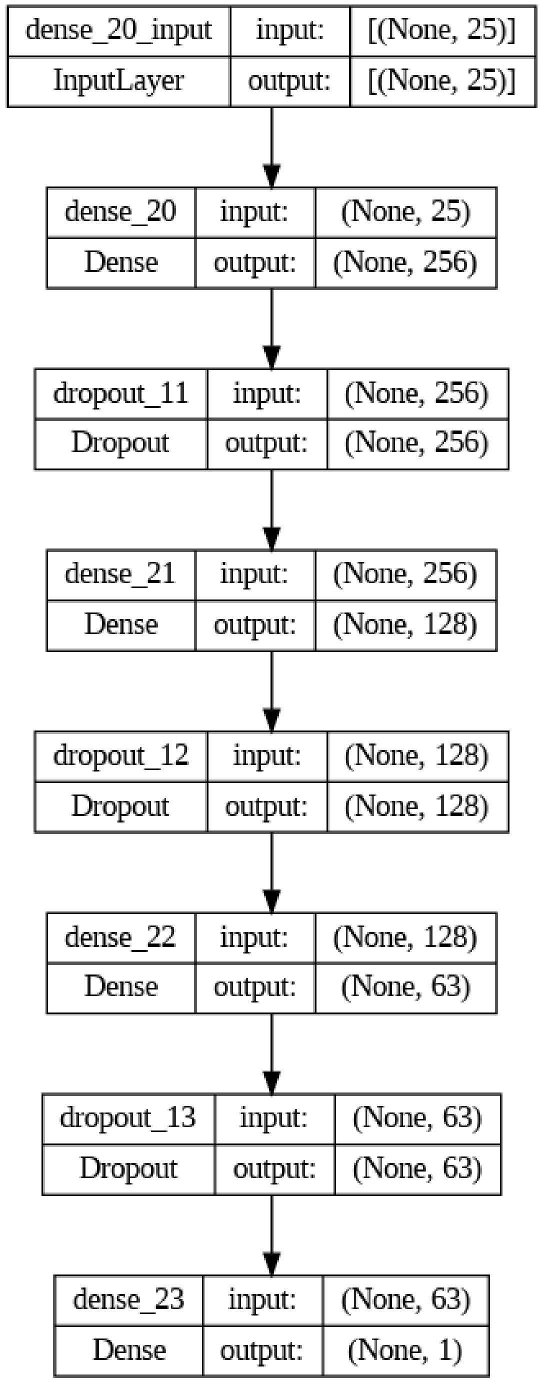
MLP [Figure 5]:



LSTM [Figure 6]:



Sequential Neural Network [Figure 7]:



The performance parameters for these were observed as below [Figure 8]:

Model	Loss	Accuracy	F1 Score
DNN	0.056	0.98	0.98
MLP	0.049	0.986	0.986
LSTM	0.056	0.981	0.981
SNN	0.081	0.97	0.95

Discussion and Conclusion:

The results of the hypothyroid classification/detection task reveal intriguing insights into the performance of different neural network models. The evaluation metrics, including Loss, Accuracy, and F1 Score, shed light on how each architecture handles the intricacies of the dataset.

Comparing the models, the Multi-Layer Perceptron (MLP) and the Long Short-Term Memory (LSTM) stand out with remarkable precision. Both models demonstrate consistently low Loss values of 0.049 and 0.056, respectively. The high Accuracies of 0.986 and 0.981, combined with F1 Scores of 0.986 and 0.981, underscore their proficiency in effectively identifying hypothyroid cases. This robust performance can be attributed to the MLP's ability to capture complex relationships in structured data and the LSTM's aptitude for capturing temporal dependencies in sequential data.

The Dense Neural Network (DNN) also delivers commendable results with a Loss of 0.056, Accuracy of 0.98, and F1 Score of 0.98. Its effectiveness in handling this task reinforces the power of densely connected layers in learning intricate patterns within the dataset.

On the other hand, the Sequential Neural Network (SNN) exhibits relatively lower performance, with a Loss of 0.081, Accuracy of 0.97, and F1 Score of 0.95. This outcome suggests that the sequential structure might not be as well-suited for capturing the underlying complexities of hypothyroid detection as compared to other architectures.

In conclusion, the study highlights the significance of choosing the right neural network architecture for specific tasks. The MLP and LSTM models shine as robust options for hypothyroid classification, demonstrating superior accuracy and F1 Score. The DNN also proves effective, while the SNN's comparatively lower results prompt consideration of its compatibility with the dataset's characteristics. These findings provide valuable guidance for selecting optimal models in medical classification scenarios and emphasize the significance of adapting architecture to the nature of the data.

References:

1. Akbas et al. (2021) Akbas, et al. American Cancer Society; 2021. Key statistics for thyroid cancer.
2. Beynon & Pinneri (2016) Beynon ME, Pinneri K. An overview of the thyroid gland and thyroid-related deaths for the forensic pathologist. Academic Forensic Pathology.
3. Nguyen QT, Lee EJ, Huang MG, Park YI, Khullar A, Plodkowski RA. Diagnosis and treatment of patients with thyroid cancer.
4. Dua & Graff (2017) Dua D, Graff C. UCI machine learning repository. 2017. http://archive.ics.uci.edu/ml
5. Han, Wang & Mao (2005) Han H, Wang W-Y, Mao B-H. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. International conference on intelligent computing; Berlin, Heidelberg. 2005. pp. 878–887.
6. Chawla et al. (2002) Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research. 2002;16:321–357. doi: 10.1613/jair.953.
7. Denoyer, Ludovic, and Patrick Gallinari. "Deep sequential neural network." arXiv preprint arXiv:1410.0510 (2014).
8. Li, Qian, et al. "A survey on text classification: From shallow to deep learning." arXiv preprint arXiv:2008.00364 (2020).
9. Islam SS, Haque MS, Miah MSU, Sarwar TB, Nugraha R. Application of machine learning algorithms to predict the thyroid disease risk: an experimental comparative study. PeerJ Comput Sci. 2022 Mar 3;8:e898. doi: 10.7717/peerj-cs.898. PMID: 35494828; PMCID: PMC9044232.
10. Ozyilmaz & Yildirim (2002) Ozyilmaz L, Yildirim T. Diagnosis of thyroid disease using artificial neural network methods. Proceedings of the 9th international conference on neural information processing, 2002. ICONIP'02, vol. 4; Piscataway. 2002. pp. 2033–2036.
11. Ioniţă & Ioniţă (2016) Ioniţă I, Ioniţă L. Prediction of thyroid disease using data mining techniques. BRAIN. Broad Research in Artificial Intelligence and Neuroscience. 2016;7(3):115–124.
12. Parikh et al. (2015) Parikh KS, Shah TP, Kota R, Vora R. Diagnosing common skin diseases using soft computing techniques. International Journal of Bio-Science and Bio-Technology. 2015;7(6):275–286.