International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 08 | Aug 2022 www.irjet.net p-ISSN: 2395-0072

Machine learning Models for classification of cushing's syndrome using retrospective Data

Harsha R Rao[1]

Dept. of MCA, Vidya Vikas Institute Of Engineering And Technology, Karnataka, India

---

Abstract – Treatments for Cushing syndrome can heal your body's cortisol levels to required healthy ratio and improve your chances of treating the disease completely. If the earlier treatment begins, there are better chances for recovery.

Accurate classification of Cushing's Syndrome (CS) plays a very important role in providing the early and correct diagnosis of CS that may make the treatment and improve patient health.

According to the algorithm/model comparison and selection process, we found that RF algorithm is best suitable for the classification of CS. Also, RF-based models were developed using training and validation (T&V) dataset with chosen features derived from measurements of Stage-1 biochemical screening test, Stage-2 biochemical later tests, and image findings. Trained models were tested using separate independent dataset to show their predictive performance on data that is not seen in a real-world environment. The developed models were used for further analysis and classification study of CS.

We can show that using ML approach, it is possible to classify subjects with CS with an average accuracy of 92% and an average F1 score of 91.5%, depending on the class comparison strategy and also the features that we have selected so far.

***Key Words*: Cushing's syndrome (CS), Cortisol levels, Hormones, Adrenocorticotropic hormone (ACTH), Pituitary gland.**

1. INTRODUCTION

Cushing disease can show problems and mood disorders such as anxiety, irritability, and depression. This condition can also make effects on a person's concentration and memory. People with Cushing disease have more likely chance of getting high blood pressure (hypertension) and diabetes.

Cushing's Syndrome (CS) is a potentially lethal disorder caused by abnormally high levels of cortisol hormone, first described in 1912 by Harvey Cushing . Early diagnosis plays a crucial role in reducing mortality and improving the prognosis of this syndrome. The treatment of CS can be little difficult, for instance, due to slow development of symptoms, and due to overlap with features of metabolic syndrome like increased blood pressure, high blood sugar, excess body fat around the waist, and abnormal cholesterol or triglyceride levels. Also, many of these features are common in many people in the general population .

Diagnosis of CS is a complex process, which will require careful and back to back interpretation of signs and symptoms, results of multiple biochemical test measurements, and findings of medical imaging by physicians with a high level of specialty and knowledge to do correct decisions. Variations in accuracy of biochemical tests, changes in test and decision criteria, and overlapping characteristics of CS subtypes make it even more difficult.

Treatment of cushing's syndrome in the modern era,there are majorly Four reasons to make it difficult to identify Cushing syndrome. Cushing asperger phenotype, which is caused by the increase use of exogenous cortisone; the confusion caused by no pathophysiological hyperparathyroidism, which can be present with symptoms consistent with Cushing syndrome;

2.Existing System

Excessive level of cortisol (i.e., hypercortisolism) shows clinical features like a large rounded face, abdominal obesity, and reducing of muscle tissue or skin thickness. Early diagnosis plays a crucial role in reducing mortality and improving the prognosis of this syndrome.

The difficulty may be solved by using rough down samples and oversampling techniques. Down sampling the classification algorithm reduces the amount of useful information since the dataset is already too small. It's possible that new training dataset created with a few old samples will bring mingle with the accuracy since they can't implement much more changes to the dataset.

Disadvantage:

Diagnosing CS and classifying its subtypes Wants required number of interpretation of medical testing, which is complicated by the limited capacity and experience of clinicians to handle more and complicated data.

Cortisol may give effects like hypertension, obesity, metabolic syndrome or diabetes, hyperglycemia, and osteoporosis if it is secreted in little amounts.A

asymptomatic CS, however, cannot be directly blamed for these negative effects, since they are common in today 's society.

3.Proposed System

The proposed approach is suitable for use in both screening testing stage for the diagnosis of CS, and also for the follow-up testing stage for the identification of the cause of CS.

Based on the classification comparison technique and the data that is selected, the suggested approach is able to define CS with a classification performance of 92%.

Alternatively, the body may also produce more cortisol due to endogenic factors. Adrenal hormones manufacturing too much endogenous cortisol are most regularly caused by Cushing's disease, a CS subtype.

Advantage:

Machine learning techniques are analyzed to show their potential as clinical practice guidelines to know and categorise CS to help in the evaluation, prognosis, and treatment of CS.

Tumour excision: Tumors causing the condition may be present in the pituitary, adrenal gland, lungs, or pancreas and can be surgically removed.

Radiation therapy:If the tumour cannot be completely removed by surgery, radiation is used to destroy the remaining tumour cells.

One should also limit the intake of sodium (salty foods) and fat.

4. SYSTEM DESIGN

In order to understand from a problem to a solution, the first step in the process is to design the system. Manage to begin the process of moving from the issue domain to the solution making plans, the problem must be defined.

There are a few parts of the system that must be taken into account. As a result of the research conducted in this section, new forms for presenting the findings will be devised.

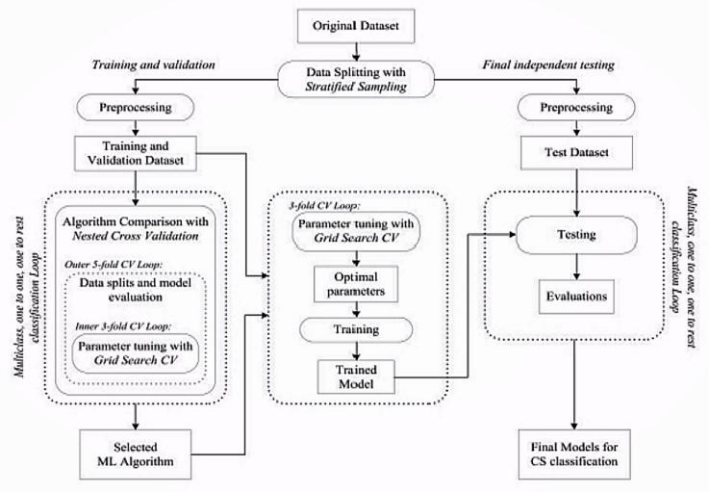

Fig 4 : Architectural Design

4.1 Dataset Profile

The CS dataset will be having 241 samples and 11 features. The target variable shows 3 classes of subtypes for patients with CS and 1 class for patients without CS. The other properties (i.e. mean, standard deviation, ratio, and counts (n)) of the dataset features according to the type of treatment process to be done.

4.2 Dataset Splitting and Subset Preparation

The input dataset was randomly split into T&V dataset (n = 168, 70%) and FIT dataset (n = 73, 30%) by stratisfied sampling knowing that ratios of classes are shown in the newly created datasets. Due to the limited dataset size, two-sample Kolmogorov-Smirnov (K-S) test was performed for each feature to check whether both T&V and FIT datasets are good and responsive for the same distribution.

4.3 DETAILED DESIGN:

Use Case Diagram

The use-case analysis in the Unified Modeling Language is what's discussed and constructed in the use-case diagram, which is a behavioural diagram (UML).

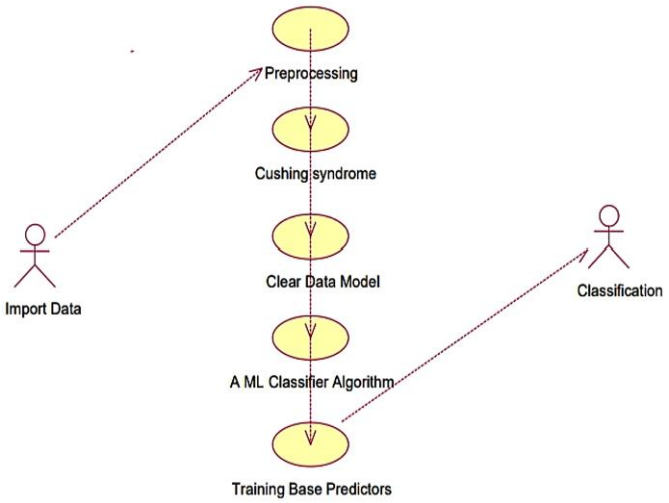

Fig 4.3 Use Case Diagram

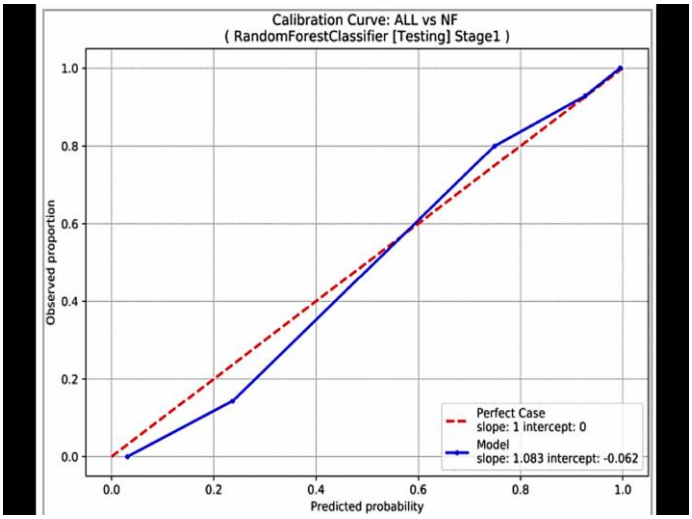

Fig 5.1: Calibration curve for the ALL vs NF Model at Stage1.

5. IMPLEMENTATION

Modular description and methodology

Cushing's syndrome:

It is essential and important to categorise Cushing's disorder (CS) so that early and accurate diagnosis may be made, which can lead to a good management and improved patient outcomes.

Technique:

It is well-known that a model's prediction performance suffers when a dataset has an imbalance of classes. Down sampling and oversampling technologies may be used to address this issue. Reverse sampling the majority class is ineffective due to the short dataset.

Methodology:

In the case of little data, this might result in misunderstanding owing to parameter tuning failures and, as a result, a diminished capacity to generalise the results. The more data we gather, otherwise we'll be able to compare and contrast imputation strategies.

5.1Clinical Performance Evaluation

The calibration curve for the ALL vs NF Model at Stage1.Predictions are on the x-axis. The observed proportions are on the y-axis.

Perfect case is represented by 45◦ line with a slope of 1 and intercept of 0 on the x-axis.

6. SYSTEM TESTING

ML algorithm comparison results indicate that most of the compared algorithms get well estimates of predictive performances. But still, RF algorithm is found to be the best algorithm in overall performance according to the scores achieved for different class comparison schemes and feature sets that we got by the implementation.

Test Cases

TC No	Positive scenario	Required Input	Expected output	Actual output	Test Result
1	Upload datasets	Upload video	Should successfully upload	uploaded	Pass
2	Pre-processing	Process dataset	Remove unwanted datasets	Unwanted datasets are removed	Pass
3	train image	Image processing	Identify object	Object detected	Pass
4	Classification	Objects are classify	Identify the object and classify which type of object it is	Object classified	Pass
5	Performance analysis	Find Accuracy	Display Accuracy information	Accuracy information displayed	Pass

7. CONCLUSIONS

Despite our sample's class imbalance and the dataset's modest size, it was proven that RF-based models could accurately predict CS interpretation. Because the RF algorithm can work well with short sample sizes and unbalanced CS data and understand complicated relationships also, as well as reduce variance, it has proven to be a great success. Data scaling or standardisation isn't required in advance of it.

ML approach can help improve physicians' judgment in diagnosing CS subtypes with limited biochemical tests which are cumbersome and stressful for patients.

REFERENCES

[1] H. Cushing, The pituitary body and its disorders: clinical states produced by disorders of the hypophysis cerebri. jB Lippincott, 1912.

[2] A. Lacroix, R. A. Feelders, C. A. Stratakis, and L. K. Nieman, "Cushing's syndrome," The Lancet, vol. 386, no. 9996, pp. 913 – 927, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140673614613751

[3] L. Vilar, M. d. C. A. A. Freitas, M. Faria, R. Montenegro, L. A. Casulari, L. Naves, and O. D. Bruno, "Pitfalls in the diagnosis of Cushing's syndrome," Arquivos Brasileiros de Endocrinologia e Metabologia, vol. 51, pp. 1207 – 1216, 11 2007. [Online]. Available: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-27302007000800006&nrm=iso

[4] R. Alwani, L. Schmit Jongbloed, F. De Jong, A.-J. van der Lely, W. De Herder, and R. Feelders, "Differentiating between cushing's disease and pseudo-cushing's syndrome: comparison of four tests," Eur J Endocrinol, vol. 170, no. 4, pp. 477–86, 2014.

[5] Cushing Syndrome: Treatment, Home Remedies, Symptoms (epainassist.com)