

Analyzing the Risk factors of Polycystic Ovary Syndrome (PCOS) with Insulin Resistance (IR)

R. Bakkiyalakshmi¹, T. Jenifer², Gayathri Premanand³

Assistant Professor, Department of Computer Science, Sri Saradha College for Women, Perambalur, India^{1,2}

Abstract: Polycystic Ovary Syndrome (PCOS) is the most common endocrine disorder in women of reproductive age, yet debate over appropriate diagnostic criteria and design limitations with sampling methodology have left some doubt as to the actual prevalence in the community. Polycystic Ovary Syndrome (PCOS) is not only the problem but it also associated with many others like causes of infertility, metabolic aberrations and manifests with hirsutism, acne, and obesity. Here we have to investigate its impact on health related issues, we have to analysis the risk factors; predict the reasons, and what is the best treatment for PCOS with Insulin Resistance. So the major goal of this research is to find the best solution by comparing patients with PCOS with and without Insulin Resistance using Data Mining Techniques.

Keywords: PCOS, IR, CART, K-means, Obesity, Type 2 Diabetes

I. INTRODUCTION

In the last 25 years, it has become clear that Polycystic Ovary Syndrome (PCOS) is an important metabolic disorder. Women with PCOS have profound peripheral Insulin Resistance. Because of this, both premenopausal women and adolescent girls with PCOS are at markedly increased for Type 2 Diabetes mellitus (DM). Although obesity and age significantly increase this risk, both non obese and young PCOS women can also affected.

In the Rotterdam classification, should be used to define Polycystic Ovary Syndrome (PCOS) in the event of:

- Menstrual cycle anomalies: amenorrhea, oligomenorrhoea or long cycles;
- Clinical and/or biochemical hyperandrogenism
- Ultrasound appearance of polycystic ovaries.

The presence of two of these three criteria is sufficient once all other diagnosis have been ruled out. Some of the most common signs of PCOS include:

- Irregular periods that come every few month, not at all, or too frequently
- Extra hair on the face or other parts of the body, called hirsutism
- Acne
- Patches of dark skin on the back of the neck and other areas, called acanthosis nigricans.
- Weight gain and/or trouble losing weight. In the case of obesity (BMI greater than 30 kg/m²), oral glucose tolerance testing (OGTT) is recommended to figure out whether you have a high risk of developing diabetes.

In the medical history, gathers enormous information that is not to be “mined” for successful decision making. In that case we have to use Data Mining and Knowledge Discovery to found the information. In this Study, the potential utilization of Classification and Clustering to monstrous volume of medical service information are inspected. Utilizing medicinal profile, for example, age, sex, pulse, blood hormones and glucose, it can anticipate the probability of patients getting a Polycystic Ovary Syndrome (PCOS) with Insulin Resistance. It empowers huge information, e.g. relationships between medical factors, patterns identified with Polycystic Ovary Syndrome (PCOS) and Insulin Resistance (IR) to be set up.

II. LITERATURE REVIEW

[1] S. Rathnavalli, Dr. M. Manimekalai, “ a Hypothesis analysis on the proposed methodology for Prediction of Polycystic Ovary Syndrome (PCOS)”, In this paper, they uses the chi-square test to know whether the methods will give effect when it is implemented for predicting Polycystic Ovary Syndrome (PCOS). They also uses the Natural Fuzzy Rough Set (NFRS) for comparing PCOS patients with and without obese. The proposed NFRS is compared with

the information gain by considering the type of menstrual cycle when the patient with and without obese. And it is found that the NFRS remove the null hypothesis i.e the value of chi square is greater than critical value. Dr. K. Meena, Dr. M. Manimekalai and S. Rathnavalli, "Correlation of Artificial Neural Network Classification and NRFs attribute filtering Algorithm for PCOS Data", the proposed method, the algorithm combines Neural Fuzzy Rough Set evaluation and Artificial Neuran Network which reduces the task of classification and Feature selection. The data set contains 26 attributes and 303 instances. They use the Orange tool for mining the given data. From this paper, they conclude that instead of using Classification and Feature Selection separately, We can do it together to provide the better result for predicting the PCOS disease.

[6] Fernando Ovalle, M.D., and Ricardo Azziz, M.D., M.P.H., M.B.A., "Insulin Resistance, Polycystic Ovary Syndrome (PCOS) and Type 2 Diabetes Mellitus", They defined IR (Insulin Resistance) as a common pathology state in which target cells fails to response the ordinary levels of circulating insulin. And also state two major problems associated with IR and PCOS and diabetes. They use four techniques for measuring insulin sensitivity are Glucose Tolerance Test, and Homeostatic Model Assessment Calculation. In their experience, more than 600 patients with PCOs, about 60% are obese. Although unique IR of PCOS occur independent of obesity, the hyperinsulinemia of PCO appears to be aggravated by the presence of obesity. They also states insulin is consider a fundamental defect in patient with Type 2 Diabetes Mellitus. Finally they conclude that PCOS and Type 2 Diabetes Mellitus are common abnormalities of IR among reproductive aged women and the patient with one disorder are at the great risk of another.

[7] The Rotterdam ESHRE/ ASRM – Sponsored PCOS consensus Workshop Group "Revised 2003 Consensus on Diagnostic Criteria and Long – term Health Risk related to Polycystic Ovary Syndrome (PCOS)", The Rotterdam (2003) consensus workshop conclude that PCOS is a Syndrome of ovarian dysfunction along with the cardinal feature hyperandrogenism and Polycystic Ovaries Morphology. Its clinical manifestations may include menstrual irregularities, signs of androgen excess, and obesity. Insulin Resistance and evaluated serum LH levels are also common features in PCOS. PCOS is associated with increased risk of Type 2 Diabetes and Cardiovascular Events.

The following are the syndrome encompasses a broader spectrum of signs and symptoms of ovarian dysfunction, then those defined by the original diagnostic criteria.

1. Oligo – or anovulation,
2. Clinical and /or biochemical signs of hyperandrogenism,
3. Polycystic ovaries and exclusion of other etiologies (congenital adrenal hyperplasia, androgen-secreting tumors, Cushing's Syndrome)

Finally, women with PCOS have multiple risk factors for diabetes including obesity, a family history of Type 2 diabetes and abnormalities in insulin action (both insulin resistance and beta cell dysfunction). There are several lines of evidence suggesting that women with PCOS are also at increased risk of cardiovascular disease. But the lifestyle changes (diet and exercise) should be strongly encouraged to reduce the risk of both Type 2 Diabetes and Cardiovascular disease.

III. DATA MINING

Data mining is going to gain its importance due to the exponential growth and use of large amounts of data. Data mining is the process of analyzing and applying algorithms on data that produce a particular enumeration of pattern over that data. Simply stated, Data mining refers to extracting or "mining" knowledge from large amount of data. Many other terms carry a similar or slightly different meaning to data mining, such as knowledge mining from data, knowledge extraction, data/pattern analysis, data archaeology, and data dredging. Knowledge discover as a process of consists of an iterative sequence of the following steps:

1. Data cleaning
2. Data integration
3. Data selection
4. Data transformation
5. Data mining
6. Pattern evaluation
7. Knowledge presentation

In the above steps, First 4 step are used for pre-processing the given data. 5th step is the major one which defines what are the data mining tool are applied on that data such as classification, prediction, regression, clustering, outlier analysis, and mining frequent pattern. 6th step is used for predicting the interesting pattern on the data using pre-defined constraints. Step 7 is to visualize the mined data to the user.

IV. CLASSIFICATION**CART Algorithm**

By applying the data mining algorithm on data set in recommendation system predict the data according to the user preferences. Prediction can be categorized into: classification, density estimation and regression. In classification, the predicted variable is a binary or categorical variable. We work on CART the decision tree learning algorithm used for classify data. The benefits using CART algorithm are fast and it is a supervised learning method. The output of the algorithm can be either decision tree or regression tree which means Classification trees output classes, regression trees output numbers, so the output is in human readable form.

V. CLUSTERING**K-means Algorithm**

The process of grouping a set of physical or abstract objects into class of similar objects is called clustering. A cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the objects in other clusters. Here we have to use K-means algorithm for cluster the data. K-means creates k groups from a set of objects so that the members of a group are more similar to each other. The benefits of using K-means algorithm are faster and more efficient than other algorithms, especially over large data sets.

VI. CONCLUSION

In this paper, we have to analysis the PCOS patients with and without Insulin resistance using data mining techniques. Using Classification, we must predict what are the risk factors of PCOS patients with Insulin resistance. Using clustering, we have to group the data set for predicting the similar reasons for PCOS with insulin resistance.

ACKNOWLEDGEMENT

Any achievement, be it scholastic or otherwise does not depend solely on the individual efforts but on the guidance, encouragement and cooperation of intellectuals, elders and friends. A number of personalities, in their own capacities have helped us in carrying out this paper work. We would like to take this opportunity to thank them all. We thank our Parents, and all the Faculty members of Department of Computer Science, Applications and Information Technology for their constant support and encouragement. We conveying our sincere thanks towards the following Papers are help us while preparing this paper. Last, but not the least, I would like to thank my peers and friends who provided me with valuable suggestions to improve my paper.

REFERENCES

- [1]. Dr. M. Manimekalai and S. Rethinavalli, "A Hypothesis on the proposed Methodology for Predicting of Polycystic Ovarian Syndrome ", IJCSET, 2016.
- [2]. Dr. K. Meena, Dr. M. Manimekalai and S. Rethinavalli, "A Novel Framework for Filtering the PCOS Attributes using Data Mining Techniques", International Journal of Engineering Research & Technology (IJERT), Vol. 4 Issue 01, January-2015, pp.no 702-706.
- [3]. Dr. K. Meena, Dr. M. Manimekalai and S. Rethinavalli, "Correlation of Artificial Neural Network Classification and NFRS Attribute Filtering Algorithm for PCOS Data", International Journal of Research in Engineering and Technology, Volume 4, Issue 3, pp. 519-524, March 2015.
- [4]. Dr. K. Meena, Dr. M. Manimekalai and S. Rethinavalli, "Implementing Neural Fuzzy Rough Set and Artificial Neural Network for Predicting PCOS", International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 3 Issue: 12, pp.6722-6727, December 2015
- [5]. PCOs resource for healthier You - <http://www.youngwomensheath.org/>
- [6]. Xindong Wu, Vipin Kumar, "The Top 10 algorithms in Data mining"
- [7]. The Rotterdam ESHRE/ ASRM – Sponsored PCOS consensus Workshop Group "Revised 2003 Consensus on Diagnostic Criteria and Long – term Health Risk related to Polycystic Ovary Syndrome (PCOS)", Fertility and Sterility, Vol. 81, No. 1, January 2004.

BIOGRAPHIES

Bakkiyalakshmi was born in 1994. Now she is working as a Assistant Professor in Department of Computer Science Sri Saradha College for Women, Perambalur. She Complete her research (M.Phil), Department of Computer Science, Applications and Information Technology in Shrimati Indira Gandhi College, Trichy. She obtained B.Sc and M.Sc (Computer Science) degree in 2014 and 2016 respectively, from Seethalakshmi Ramasami College, Trichy. She is an avid coder who always looks to develop something new. Now she is looking forward to improve her skill in software developing. Her areas of working interest are Network Security, Data Mining and Data Structures.



Jenifer was born in 1990, now she is working as a Assistant Professor in Department of Computer Science Sri Saradha College for Women, Perambalur. She has done her M.Phil at Department of Computer Science, Applications and Information technology in Shrimati Indira Gandhi College, Trichy. She finished her Masters in Computer Application from M.A.M Engineering College, Trichy in 2013. She has done Bachelor Degree of Computer Science from Shrimati Indira Gandhi College, Trichy in 2010. She has worked as a teacher at CSC Computer Education Center and worked as a Software Developer at Techlife Global. Her current research interests include Data Mining and Network Security.



Gayathri Premanand was born in 1993. She completed her Master of Philosophy in Computer Science from Department of Computer Science, Applications and Information technology in Shrimati Indira Gandhi College, Trichy. She has completed her MCA in Cauvery College for Women, Trichy during 2015 and she completed her Bachelor Degree in Computer Science at the same college during 2013. She loves to work on Software related problems. Her research interests cover Network Security and Computer Graphics.