

ABSTRACT

Polycystic Ovary Syndrome or PCOS is an endocrine disorder that occurs in women of reproductive age. The condition once detected cannot be cured but treatment can help relieve its effects.

The exact cause of PCOS is still unknown but there are certain factors that portray the risk of getting PCOS. The factors that result in this syndrome are obesity, insulin resistance, blood pressure, depression, inflammation. The symptoms include: hirsutism, Oligo-ovulation, acne, heavy bleeding, skin darkening. Using the causes and symptoms, a model is prepared in order to accept them as features and outputs the presence or absence of this condition. The machine learning models used for supervised classification are Logistic Regression. The reason behind building multiple models is to find out the best one for the given dataset, in the known scope of knowledge

- web- flask server
- access to all public
- refer the nearest gynecologist

TABLE OF CONTENTS

Chapter No.	TITLE	Page No.
	ABSTRACT	v
	LIST OF FIGURES	viii
	LIST OF ABBREVIATIONS	ix
1	INTRODUCTION	1
	1.1. OVERVIEW	2
	. MACHINE LEARNING	3
	MACHINE LEARNING STRATEGIES	3
	SUPERVISED LEARNING	3
	UNSUPERVISED LEARNING	4
	DEEP LEARNING	
2	LITERATURE SURVEY	6
	RELATED WORK	6
3	METHODOLOGY	8
	EXISTING SYSTEM	8
	PROPOSED SYSTEM	8
	OBJECTIVE	8
	SOFTWARE AND HARDWARE REQUIREMENTS	9
	SOFTWARE REQUIREMENTS	9
	HARDWARE REQUIREMENTS	9
	LIBRARIES	9
	PROGRAMMING LANGUAGES	10
	PYTHON	10
	HTML	
	CSS	
	3.5.4. DOMAIN	10
	SYSTEM ARCHITECTURE	11
	ALGORITHMS USED	11

	MULTI LAYER PERCEPTRON	11
	LOGISTIC REGRESSION	12
	MODULES	13
	DATASET COLLECTION	13
	TRAIN AND TEST THE MODELS	13
	DEPLOY THE MODELS	13
4.	RESULTS AND DISCUSSION	14
	WORKING	14
5.	CONCLUSION	15
	CONCLUSION	15
	REFERENCES	16
	APPENDICES	17
	A. SOURCE CODE	17
	B. SCREENSHOTS	20
	C. PLAGIARISM REPORT	24
	D. JOURNAL PAPER	25

LIST OF ABBREVIATIONS

ABBREVIATIONS	EXPANSION
ML	Machine Learning
DL	Deep Learning
MLP	Multi-Layer Perceptron
LG	Logistic Regression

CHAPTER 1

INTRODUCTION

PCOS (Polycystic ovary syndrome) is one of the most common hormonal disorders which most women suffer once in a lifetime during their reproductive age. This paper considers the issue when it affects more than 70% of the population of females, it is not that known among the people. Many times females ignore the symptoms and only recognize it when it gets very serious when one has to take drastic measures against it. PCOS is a hormonal disorder where women develop some follicles in their ovaries called cysts and these follicles cause ovaries to fail in releasing eggs. Women suffering from PCOS can release excessive androgen (male hormone) and can suffer from extended or infrequent menstrual periods .

This paper focuses on the female population to make them aware about this disorder so that they can take the symptoms into consideration as not everyone goes to screenings as this disorder can be diagnosed through it, considering the symptoms can be seen as a cause to go for screenings. Up until 2019, all that has been done regarding PCOS are always based on ultrasound images and image recognition techniques to diagnose the disorder using particle swarm optimization, automated detection using follicle recognition, segmentation using poly cystic ovary using ultrasound images, data mining techniques or speed gradient method.

Out of all the methods one of the problems is they are considering the symptoms but focusing on the ultrasound images or recognizing follicles through any other screening method and then giving the result that determines whether a person is suffering from PCOS or not. The researcher is predicting the percentage of chances that PCOS can occur in a female only through the symptoms using back propagation algorithm in neural network and genetic algorithm as they are known to provide a good result. These two methods can also help us recognize the disorder in the early reproductive age of a female (9-15 years) to warn them about their future.

Polycystic Ovary Syndrome is a disease that occurs in women during their child-bearing years. The reproductive organs of women called ovaries that produce progesterone and estrogen-hormones that regulate the menstrual cycle, are affected.

Ovaries also produce little amounts of androgen also called as male hormones. The basic features of PCOS are:

- Cysts in ovaries.
- High levels of hormone: androgen.
- Irregular Periods
- Excessive body hair growth

Since the condition is a syndrome, it has a collection of symptoms that suggest its presence. These symptoms play a vital role in detection of this condition. Along with these symptoms, causes which may lead to the risks of the syndrome can also be considered. It is necessary to detect PCOS as early as possible since it holds the risk of infertility, diabetes, endometrial cancer and cardiovascular disease at a later stage of the condition. Here, a couple of machine learning models are built in order to determine the presence of PCOS. Since the dataset does classify if the condition is present or not, supervised machine learning algorithms are used called: K-Nearest Neighbor (K-NN) and Logistic Regression. The former is a distance based technique and the latter is probability based, this is why these two techniques, that are poles apart, are used and their accuracy is compared. Logistic Regression is more accurate with accuracy 92% while K-NN's is 90.74%. But we tried to get more accuracy than above two techniques. so we tried to use another algorithm which is Multi layer Perceptron from DL. which gave us the accuracy of 98%.

OVERVIEW

This report discusses the result of the work done in development of "YUNI-PCOS DETECTOR" using machine learning and python. It aims at the development of an application that assists females to check whether they are prone to pcos or not, and can get some food tips and gynecologist nearby .

There are also some remote villages which lack medical facilities. The dataset was processed in DL models Multi Layer Perceptron. While processing the data, symptoms are given as input and the disease was received as an output. This project helps to get the idea about the disease of an individual based on the symptoms he/she have, and get the treatment easily by contacting the concerned doctor.

1.2.MACHINE LEARNING

Machine learning could be a subfield of computer science (AI). The goal of machine learning typically is to know the structure of information and match that data into models which will be understood and used by folks. Although machine learning could be a field inside technology, it differs from ancient process approaches.

In ancient computing, algorithms are sets of expressly programmed directions employed by computers to calculate or downside solve. Machine learning algorithms instead give computers to coach on knowledge inputs and use applied math analysis so as to output values that fall inside a particular vary. thanks to this, machine learning facilitates computers in building models from sample knowledge so as to modify decision-making processes supported knowledge inputs.

1.3 Machine Learning Strategies

In machine learning, tasks square measure typically classified into broad classes. These classes square measure supported however learning is received or however feedback on the educational is given to the system developed. Two of the foremost wide adopted machine learning strategies square measure supervised learning that trains algorithms supported example input and output information that's tagged by humans, and unattended learning that provides the algorithmic program with no tagged

information so as to permit it to search out structure at intervals its computer file.

Supervised Learning

In supervised learning, the pc is given example inputs that square measure labeled with their desired outputs. The aim of this technique is for the algorithmic program to be ready to “learn” by comparison its actual output with the “taught” outputs to search out errors, and modify the model consequently. Supervised learning thus uses patterns to predict label values on extra unlabelled information. For example, with supervised learning, an algorithm may be fed data with images of sharks labeled as fish and images of oceans labelled as water. By being trained on this data, the supervised learning algorithm should be able to later identify unlabelled shark images as fish and unlabelled ocean images as water.

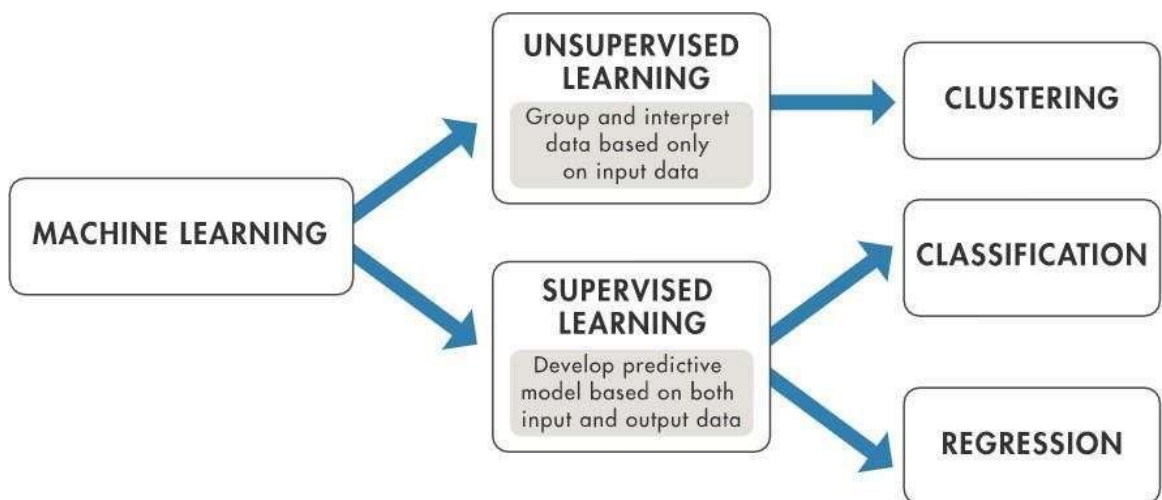
A common use case of supervised learning is to use historical information to predict statistically probable future events. It's going to use historical stock exchange info to anticipate approaching fluctuations, or be used to filter spam emails. In supervised learning, labeled photos of dogs are often used as input files to classify unlabeled photos of dogs.

Unattended Learning

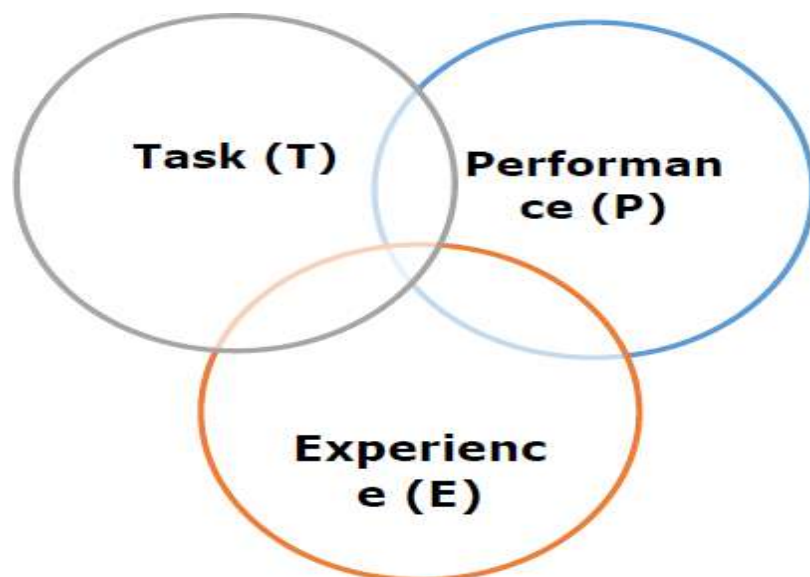
In unattended learning, information is unlabeled, that the learning rule is left to seek out commonalities among its input file. The goal of unattended learning is also as easy as discovering hidden patterns at intervals a dataset, however it should even have a goal of feature learning, that permits the procedure machine to mechanically discover the representations that square measure required to classify data.

Unsupervised learning is usually used for transactional information. You will have an oversized dataset of consumers and their purchases, however as a person's you'll probably not be able to add up of what similar attributes will be drawn from client profiles and their styles of purchases.

With this information fed into Associate in Nursing unattended learning rule, it should be determined that ladies of a definite age vary UN agency obtain unscented soaps square measure probably to be pregnant, and so a promoting campaign associated with physiological condition and baby will be merchandised.



MACHINE LEARNING CLASSIFICATION

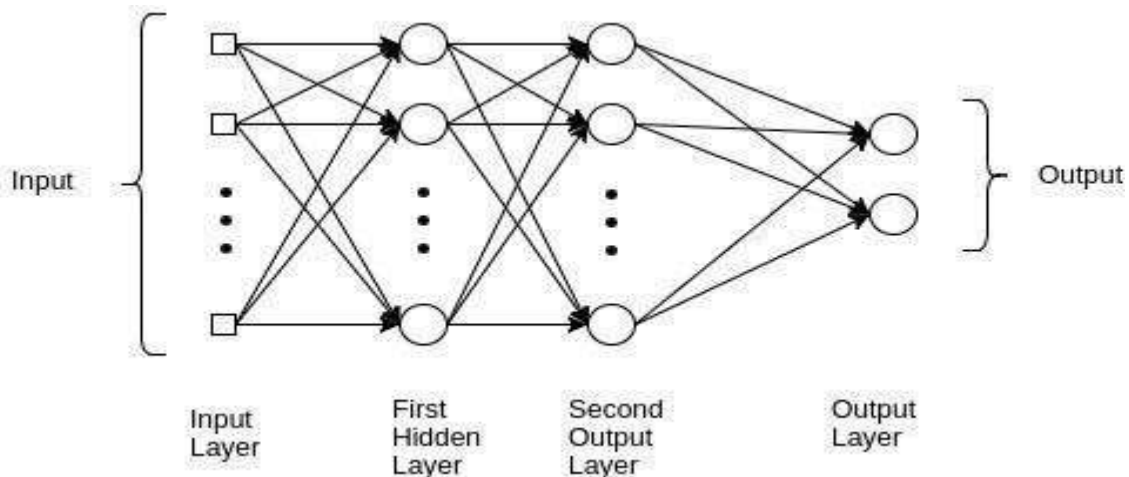


MACHINE LEARNING TASK

DEEP LEARNING

Deep learning is a machine learning technique that teaches computers to do what comes naturally to humans: learn by example. Deep learning is a key technology behind driverless cars, enabling them to recognize a stop sign, or to distinguish a pedestrian from a lamppost. It is the key to voice control in consumer devices like phones, tablets, TVs, and hands-free speakers. Deep learning is getting lots of attention lately and for good reason. It's achieving results that were not possible before.

In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers.



DEEP LEARNING

CHAPTER 2

LITERATURE SURVEY

2.1 RELATED WORK

Zhang,2018 ,Tried to predict and identify the risk of pcos through Hyperprolactinemia, Cushing's syndrome and non-classic congenital adrenal hyperplasia algorithm.

Norman,2007,paper determines the data analysis on the Risk of Obesity and Overweight in a Women study on the disorder and its three diagnostic criteria in depth giving us insights on not just pcos.

P. Mehrotra, 2011, used machine learning algorithms like Bayes and Logistic Regression .Performance of both the models was compared and the best classifier is used to predict the above diseases.

B.Cahyono Adiwijaya, M.s. Automated Detection of PCOS using follicle recognition(fr) has tried to recognize PCOS based on the FR through pre-processing to increase the quality of image.

SharvariS.Deshpande ,Asmita's implementations of Convolutional Neural Network on Pcos classifies patients into Pcos and non-Pcos class by using convolutional neural network algorithms.

Honglei Jing,Xiang Xu, Jinzhu Wang, provided a prospect that Pcos can be prevented at an early age .

Xiangyang Liu, Juan Bao, Yan, used various types of machine learning techniques. In that Rf algorithm is found superior.

Akinrotimi,Akinyemi Omololu, paper determined that patient suffering from obese are more likely to get hypertension and more.so by controlling obesity in patients makes them more responsive towards clomiphene.