

# Prediction of Prediabetes, No Diabetes and Diabetes Mellitus-2 Using Pattern Recognition



Anupriya Kamble, Shaikh Abdul Hannan, Anubha Jain,  
and Ramesh Manza

**Abstract** As of now everyone knows that the glucose or the sugar level when increases or decreases in the blood it leads to diabetes disorder. The doctors do have a need of an automated system to detect diabetes disorder by which their time for writing down the prescription will be minimized. The automated detection of the disorder will help doctors to reduce the time spent on the test and also it will be helpful for the patient to take precautions much early if in case a subject is suffering from the disorder. The present analysis concludes that if a subject is possessing PD, NODB or DM2 and if it has to be detected then NaiveBayesMultinomialText classification can be used if and only if it has to be used in comparison with CVParameterSelection or Stacking classification as NaiveBayesMultinomialText gave highest result 48.837% compared to CVParameterSelection and Stacking.

**Keywords** PD · NODB · DM2 · NaiveBayesMultinomialText · CVParameterSelection · Stacking · Diabetes

---

A. Kamble (✉) · R. Manza

Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar,  
Marathwada University, Aurangabad, Maharashtra, India  
e-mail: anupriya.k.145@gmail.com

R. Manza

e-mail: manzaramesh@gmail.com

S. A. Hannan

Department of Computer Science, Al-Baha University, Albaha, Saudi Arabia  
e-mail: abdulhannan05@gmail.com

A. Jain

IIS (Deemed to Be University), Jaipur, India  
e-mail: anubha.jain@iisuniv.ac.in

© Springer Nature Singapore Pte Ltd. 2021

V. S. Rathore et al. (eds.), *Rising Threats in Expert Applications and Solutions*,  
Advances in Intelligent Systems and Computing 1187,  
[https://doi.org/10.1007/978-981-15-6014-9\\_90](https://doi.org/10.1007/978-981-15-6014-9_90)

749

## 1 Introduction

### 1.1 What Is Diabetes?

As of now, everyone knows that the glucose or the sugar level when increases or decreases in the blood it leads to diabetes disorder, which may further cause different diseases, for example, diabetic retinopathy, glaucoma, and so on [1].

### 1.2 Types of Diabetes

**Prediabetes.** A condition in which blood sugar levels are higher than the normal blood sugar level but not high enough to be Type 2 diabetes (Type 2 because of the lifestyle of people, they are much likely to suffer from Type 2) and by precautions and medications can be normalized is known as prediabetes [2]. Sugar plays an important role of fueling up of the body cells. When it fails to do so, it automatically, instead of fueling up, builds up in the bloodstream. These are the steps of occurrence of prediabetes.

Type 1. When a person's body fails to produce insulin, the state is known as Type 1 diabetes. Insulin helps the glucose to pass from blood to the cells of the body [3].

Type 2. When the sugar level or glucose in the blood is higher than normal, it is known as Type 2 diabetes. Type 2 diabetes is also known as hyperglycemia. Mostly people suffer from Type 2 diabetes [4].

Gestational diabetes. If a woman suffers from high blood sugar levels during pregnancy, then it is known as gestational diabetes [5].

### 1.3 Why Automated Diabetes Diagnosis Detection?

Today in 4G digitalized lifestyle almost everything is digitalized. Only a few things may be there which have to be digitalized. Even the kitchen, washing facility, most of the things are made automatically operated. Then why not the doctors have an expert system for them? By which their time for writing down the prescription will be minimized. As of today, number of patients suffering from diabetes is increasing day-by-day, so is the queue in the hospital too. The automated detection of the disorder will help the doctor to reduce the time spent on the test and also it will be helpful for patient to take precautions much earlier if in case a subject is suffering from the disorder.

## 2 Dataset

The dataset of the present study consists of 215 subjects, both male and female, which have been numbered from P1 to P215. The dataset contains subjects that are in the first stage of diabetes, that is, prediabetes (PD), who are not suffering from diabetes (NODB) and who are suffering from diabetes mellitus-2 (DM2). The dataset has been collected for the research study from Dr. Varsha Apte, Ameya Hospital, Aurangabad, Maharashtra (India) and Dr. Vinit Rao, Borse Hospital, Pune, Maharashtra (India), in which the parameters considered are age, sex, height in cm, weight in kg, body mass index (BMI), waist in inches, hip in inches, waist to hip ratio, systolic blood pressure, hypertension (HTN), obesity, gestational diabetes mellitus (GDM), heart disease, smoking, tobacco, blurry vision, shoe probe, exercise/physical activity, lifestyle, alcohol, weight gain, weight loss, polydipsia (excessive thirst), polyurea, polyphagia (excessive eating), fatigue, drowsiness, nausea/vomiting, leg pain, dysphoea, angina, palpitation, breathlessness, jaw/throat pain, delayed wound healing, tingling in hands or legs, fruity smell of breath and sweat, diabetic foot, blood sugar level after fasting of 8 or 12 or 14 h, blood sugar level taken post meal, blood sugar level taken randomly, serum creatinine level (Sr. creatinine), low-density lipoprotein [6], high-density lipoprotein [7], triglycerides, family history which contains diabetes mellitus, HTN, obesity, heart disease [8]. In the present study the data should be classified into three parameters, that is, PD, NODB and DM2.

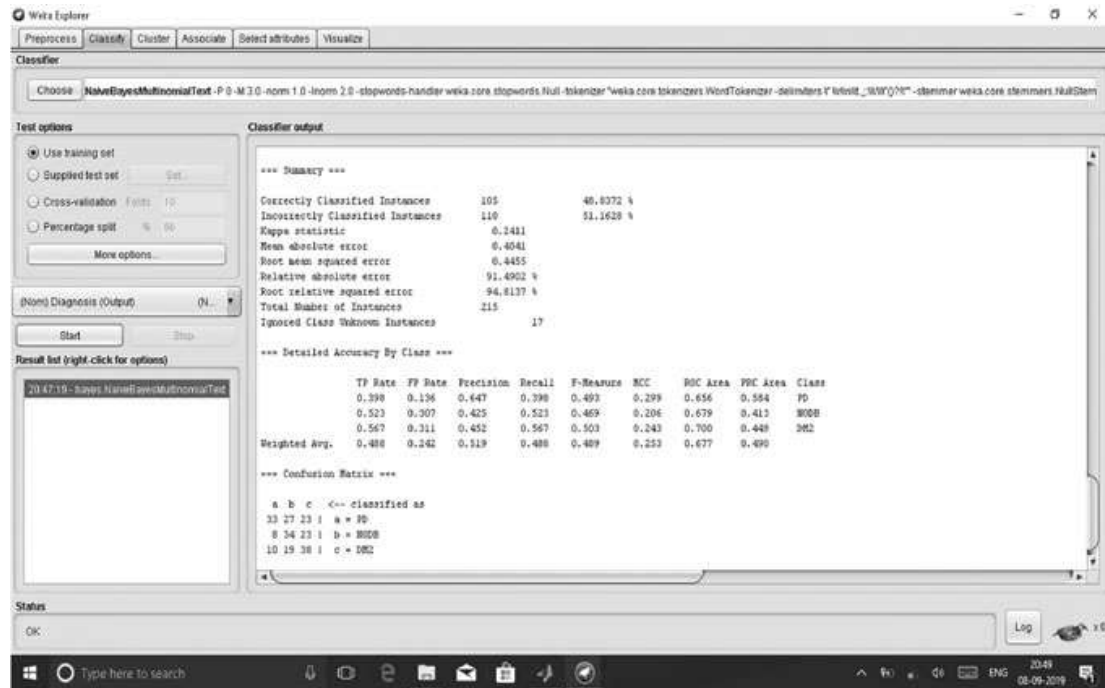
## 3 Methodology

### 3.1 *NaiveBayesMultinomialText*

It is a type of Bayes classification. The NaiveBayesMultinomialText was selected as it is applied only on string. If any other input attributes are present other than string, then they are ignored but accepted during the training as well as classification [9]. Here, in this classification technique different objects are contained into one object, which is also known as aggregate object [10].

### 3.2 *CVParameterSelection (Cross-Validation Parameter Selection)*

It is a meta type of classifier. In this classifier parameter selection is performed by using cross-validation for any of the classifier [11].



**Fig. 1** Result of NaiveBayesMultinomialText classification

### 3.3 Stacking

It is also a meta type of classifier. This classification mainly combines different type of maximum classifiers by using stacking method. Regression also can be performed by this class [12].

## 4 Results

### 4.1 Result of NaiveBayesMultinomialText Classification

The NaiveBayesMultinomialText gave result as correctly classified instances as 48.8372% and confusion matrix had value in all three parameters mixed with each other as shown in Fig. 1.

### 4.2 Result of CVParameterSelection Classification

The CVParameterSelection gave result as correctly classified instances as 38.6047% and confusion matrix had value in only one parameter, that is, first parameter namely prediabetes (PD), as shown in Fig. 2.

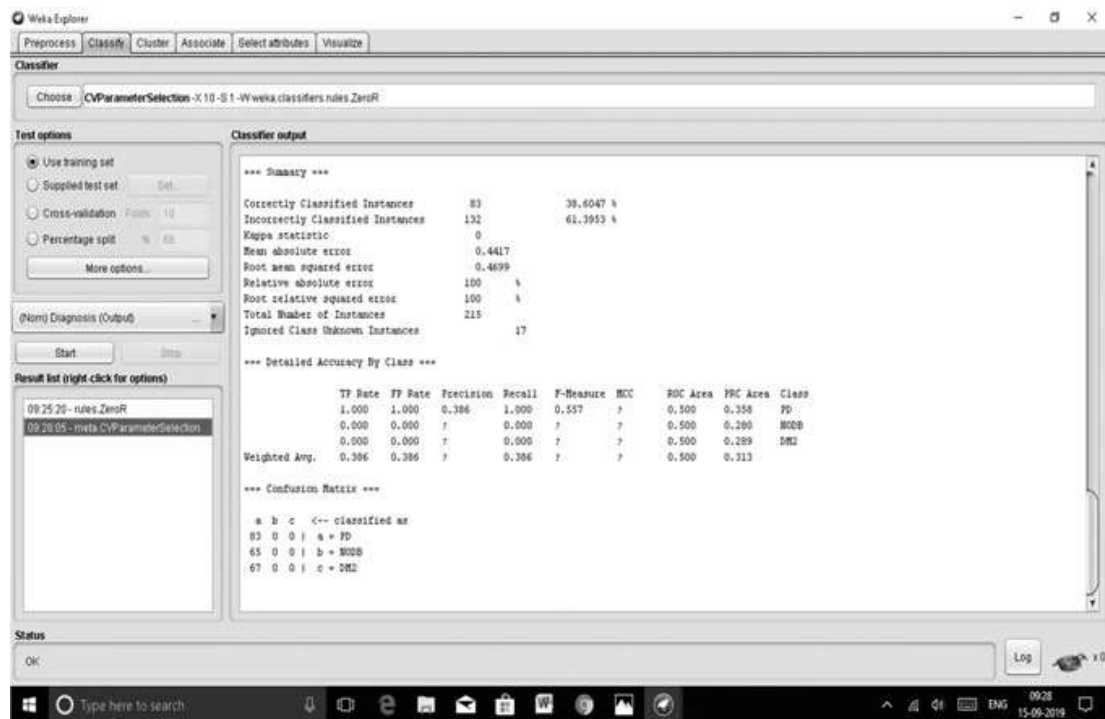


Fig. 2 Result of CVParameterSelection classification

### 4.3 Result of Stacking

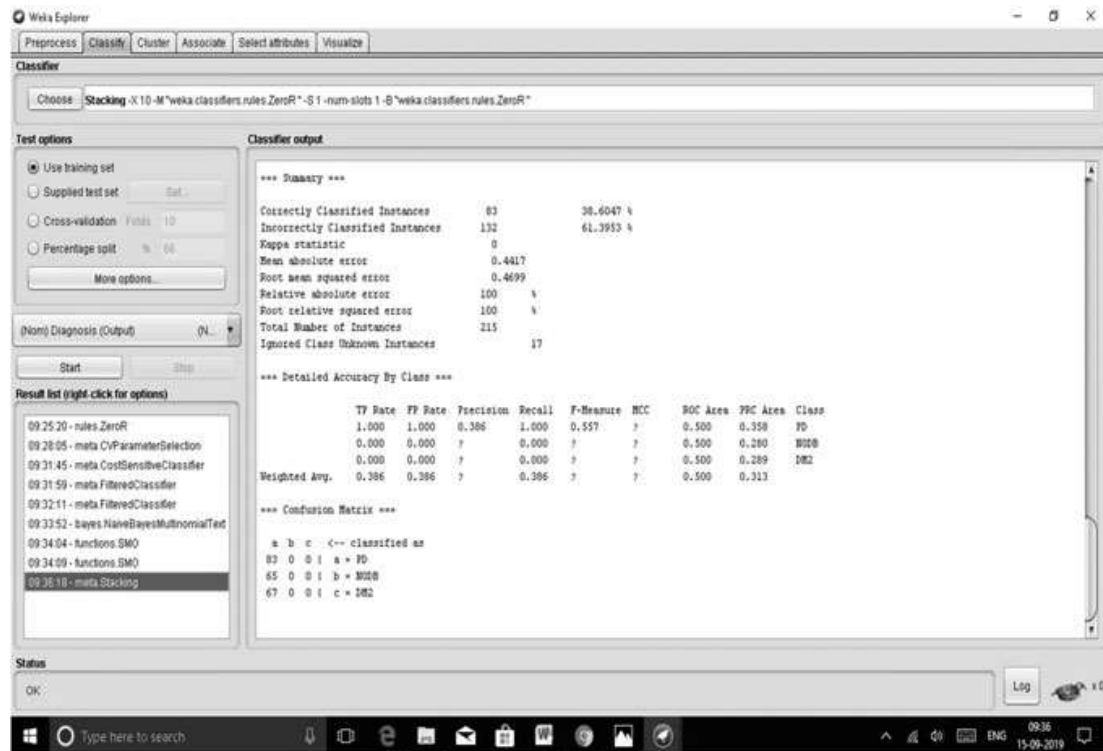
The Stacking classification also gave result as CVParameterSelection as correctly classified instances as 38.6047% and confusion matrix also had value in only one parameter, that is, first parameter namely PD, as shown in Fig. 3.

## 5 Comparison

The first classification applied, that is, NaiveBayesMultinomialText gave result as correctly classified instances as 48.8372%. The second classifier, that is, the CVParameterSelection gave less result than the first classifier as correctly classified instances as 38.6047%. And the third, Stacking classification classifier gave result the same as the second classifier, that is, correctly classified instances as 38.6047%, which means it is also less than the NaiveBayesMultinomialText.

## 6 Conclusion

The above analysis concludes that if a subject is possessing PD, NODB or DM2, and it if has to be detected then NaiveBayesMultinomialText Classification can be used if and only if it has to be used in comparison with CVParameterSelection or



**Fig. 3** Result of stacking classification

Stacking classification as NaiveBayesMultinomialText gave highest result 48.837% compared to CVPParameterSelection and Stacking. But as it clearly shows the value obtained does not reach its higher level of achievement, and the present classifiers cannot be used for the detection of PD, NODB or DM2. The study therefore concludes that NaiveBayesMultinomialText classification can be trusted for classification only 48.357%.

## 7 Future Scope

As in this analysis new algorithms used, that is, NaiveBayesMultinomialText, CVPParameterSelection and Stacking did not give good results, in future more new algorithms are present which can be applied to get a good sort of result that can help in the detection of PD, NODB or DM2.

**Acknowledgements** The authors are also thankful to University Grant Commission (UGC) for providing financial support in the form of Rajiv Gandhi National Fellowship.

## References

1. About diabetes. World Health Organization. [Internet]. <http://www.who.int/mediacentre/factsheets/fs312/en/>
2. What is Prediabetes? American Diabetes Association Clinical DIABETES [Internet]. <http://clinical.diabetesjournals.org/content/31/2/95>
3. Type 1 Diabetes Symptoms. Diabetes.co.uk the global diabetes community [Internet] <https://www.diabetes.co.uk/type1-diabetes-symptoms.html>
4. Type 2 Diabetes Symptoms. Diabetes.co.uk the global diabetes community [Internet]. <https://www.diabetes.co.uk/type2-diabetes-symptoms.html>
5. Gestational Diabetes. MedicineNet.com [Internet]. [https://www.medicinenet.com/gestational\\_diabetes/article.htm#gestationaldiabetesdefinitionandfacts](https://www.medicinenet.com/gestational_diabetes/article.htm#gestationaldiabetesdefinitionandfacts)
6. Java T point. <https://www.javatpoint.com/ldl-full-form>
7. HDL: The “Good” Cholesterol. <https://medlineplus.gov/hdlthegoodcholesterol.html>
8. NaiveBayesMultinomialText. <http://weka.sourceforge.net/doc.dev/weka/classifiers/bayes/NaiveBayesMultinomialText.html>
9. Lecture Notes On Object-Oriented Programming. <https://atomicobject.com/resources/oo-programming/object-oriented-aggregation>
10. LWL. <http://weka.sourceforge.net/doc.dev/weka/classifiers/lazy/LWL.html>
11. weka.classifiers.meta <http://weka.sourceforge.net/doc.dev/weka/classifiers/meta/CVParameterSelection.html>
12. Class stacking. <http://weka.sourceforge.net/doc.dev/weka/classifiers/meta/Stacking.html>