**Lr research Gap**

In here we are talking about what are the other researches which have already done similar to our project, and what are the similarities, differences between them and our project. Also, we are talking about the methods they used, methods we are going to use, how they collected the Data, how we are collecting the Data likewise.

About our project

In here we are collecting Data from the hospitals and Diabetics centres. We have made a form to gather Data from the patients and asking questions according to the parameters we are using Blood Pressure, Age, Diabetes Pedigree, Skin Thickness, Insulin, BMI, Gestational Diabetes, Glucose Level, Prediabetes. We are using Random forest classification, Xgboosting, Gaussian classification as methods and using Naive base algorithms.

Similar projects/research papers

1. ***Prediction on Diabetes Using Data mining Approach Pardha Repalli, Oklahoma State University***

They gathered Data From National Diabetes Fact Sheet on 2011. They are using data mining approach, general predictive algorithms,Cross Industry Standard Process for Data Mining methodology,Decision trees and Regression models as methods. Some of their findings are,

* The number of diabetic people in dataset is around 5% and the remaining people are nondiabetic.
* The interesting fact observed in the analysis is that majority of the people affected by diabetes are people whose age is above 45 years which is 4.2% out of 5.1%.
* 34% of the population whose age is below 20 years is not affected by diabetes
* 33.9% of the population whose age is above 20 and below 45 years is not affected by diabetes.
* 26.8% of the population whose age is above 45 years is not diabetic

Their parameters are the variables considered for lifestyle activities are Last Dental checkups, Last checkup, Last Cholest\_Check, Last PSA test, Last PAP test, Last Breast Exam, Last Mammogram and Wears Seat Belts. They have studied these research papers,

1. ***International Journal of Computer Trends and Technology (IJCTT) – volume 11 number 2 – May 2014 ISSN: 2231-2803 http://www.ijcttjournal.org Page94 Prognosis of Diabetes Using Data mining Approach-Fuzzy C Means Clustering and Support Vector Machine***

Their data set is donated by Vincent Sigillito from the Applied Physics Laboratory at the Johns Hopkins University, is one of the most well-known datasets for testing classification algorithms. Their methods are ALGORITHMS & FLOW DIAGRAM( Fuzzy C-means clustering (FCM), Support Vector Machine (SVM)). Some of their findings are,

* 26.9% of the population affected by diabetes are people whose age is greater than 65
* 11.8% of all men aged 20 years or older are affected by diabetes
* 10.8% of all women aged 20 years or older are affected by diabetes

Their parameters are Number of times pregnant, Diastolic blood pressure (mm/Hg), Triceps skin fold thickness (mm), Diabetes Pedigree function, Age (years) etc.

1. ***A Tool for Diabetes Prediction and Monitoring Using Data Mining Technique***

They were gathered data from Srinivas diagnostic laboratory, Gayatripuram 1st stage, Mysuru,Balaji Diagnostic laboratory, Hospet, Annapurna multi-speciality hospital, Gangavathi Rotaract club of NMIT, Bangalore. Their method is ID3 (Iterative Dichotomiser 3). They did not find any findings because this is about a predicting app. This is very similar to our one but not exactly but partially. Their parameters are age, weight, tiredness, wound healing, Sleepy/drowsy.

*The research papers which are referred by us and them are included in the References.*

**1)**[**https://www.lexjansen.com/scsug/2011/repalli1/Prediction%20on%20Diabetes%20Using%20Data%20mining%20Approach.pdf**](https://www.lexjansen.com/scsug/2011/repalli1/Prediction%20on%20Diabetes%20Using%20Data%20mining%20Approach.pdf)

* Centers for Disease Control and Prevention. National diabetes factsheet: national estimates and general information on Diabetes and prediabetes in the United States, 2011. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2011. <http://www.cdc.gov/diabetes/pubs/pdf/ndfs_2011.pdf>
* <http://www.diabetes.org/diabetes-basics/diabetes-statistics/>
* Kattamuri S.Sarma: ‘Predictive Modeling with SAS Enterprise Miner’- Practical Solutions for Business Application SAS Press.

**2)**[**https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1072.5839&rep=rep1&type=pdf**](https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1072.5839&rep=rep1&type=pdf)

[1] UCI Machine Learning Repository- Center for Machine Learning and Intelligent System, <http://archive.ics.uci.edu>

[2]Pardha Repalli, “Prediction on Diabetes Using Data mining Approach”.

[3] Joseph L. Breault., “Data Mining Diabetic Databases:Are Rough Sets a Useful Addition”.

[4] G. Parthiban, A. Rajesh, S.K.Srivatsa, “Diagnosis of Heart Disease for Diabetic Patients using Naive Bayes Method “, International Journal of Computer Applications (0975 – 8887) Volume 24– No.3, June 2011.

[5] P. Padmaja, “Characteristic evaluation of diabetes data using clustering techniques”, IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.11, November 2008.

[6] Lin, C., Lee, C., “Neural Fuzzy Systems,” PrenticeHall, NJ, 1996.

[7] Tsoukalas, L., Uhrig, R., “Fuzzy and Neural Approaches in Engineering,” John Wiley & Sons, Inc., NY, 1997.

[8] De Oliveira, J. V., & Pedrycz, W. (2007). Advances in fuzzy clustering and its applications. (1 ed., pp. 4-69). London: Wiley.

[9] Everitt, S., Landau, S., Leese, M. (2011). Cluster Analysis. (5 ed., pp. 76- 80). London: Wiley.

[10] T. J. Ross, Fuzzy Logic with Engineering Applications, Third Edition, ISBN: 978-0-470-74376-8, John Wiley & Sons, 2010

[11] <http://en.wikipedia.org/wiki/Diabetes_mellitus>

[12].Yamaguchi M, Kaseda C, Yamazaki K, Kobayashi M. “Prediction of blood glucose level of type 1 diabetics using response surface methodology and data mining”. Med Biol Eng Comput.2006; 44(6):451–7.

[13] Centers for Disease Control and Prevention. National diabetes factsheet: national estimates and general information on Diabetes and prediabetes in the United States, 2011. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2011. <http://www.cdc.gov/diabetes/pubs/pdf/ndfs_2011.pdf>

[14] Vapnik, V., Estimation of Dependences Based on Empirical Data, Springer-Verlag, (1982).

[15] Cortes, C., Vapnik, V., "Support Vector Networks," Machine Learning, 20:273-297, (1995).

[16] Burges, C. J. C., "A Tutorial on Support Vector Machines for Pattern Recognition," submitted to Data Mining and Knowledge Discovery,http://svm.research.bell-labs.com/SVMdoc.html, (1998).

**3)** [**http://j.mecs-press.net/ijitcs/ijitcs-v8-n11/IJITCS-V8-N11-4.pdf**](http://j.mecs-press.net/ijitcs/ijitcs-v8-n11/IJITCS-V8-N11-4.pdf)

[1] M. Durairaj, V. Ranjani, “Data Mining Applications In Healthcare Sector: A Study”, International journal of scientific & technology research volume 2, issue 10, October 2013, ISSN 2277-8616.

[2] P.Yasodha, M. Kannan, “Analysis of a Population of Diabetic Patients Databases in WEKA Tool”, International Journal of Scientific & Engineering Research Volume 2, Issue 5, May-2011, ISSN 2229-5518. [3] Rashedur M. Rahman, Farhana Afroz, “Comparison of Various Classification Techniques Using Different Data Mining Tools for Diabetes Diagnosis”, Journal of Software Engineering and Applications, 2013, 6, 85-97

[4] D.Lavanya, K.Usha Rani, “Performance Evaluation of Decision Tree Classifiers on Medical Datasets”, International Journal of Computer Applications, July 2011 (0975 – 8887) Volume 26– No.4.

[5] K. R. Lakshmi, S.Prem Kumar, “Utilization of Data Mining Techniques for Prediction of Diabetes Disease Survivability”, International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June-2013, 933 ISSN 2229-5518.

[6] Akash Rajak, “A Temporal Reasoning System for Diagnosis and Therapy Planning”, I.J. Information Technology and Computer Science, 2015, 12, 23-29 Published Online November 2015 in MECS (http://www.mecs-press.org/) DOI:10.5815/ijitcs.2015.12.03

[7] Vaishali Jain , Supriya Raheja “Improving the Prediction Rate of Diabetes using Fuzzy Expert System” I.J. Information Technology and Computer Science, 2015, 10, 84-91 Published Online September 2015 in MECS

[8] Nida Chammas, Radmila Juric, Nigel Koay, Varadraj Gurupur, Sang C. Suh, “Towards a Software Tool for Raising Awareness of Diabetic Foot in Diabetic Patients”, 46th Hawaii International Conference on System Sciences, 2013, 1530-1605.

[9] Nouf Almutairi, Riyad Alshammari, “Diabetes Early Warning System”, College of Public of Health and Health Informatics.

[10] Diabetes Dataset: Srinivasa diagnostic laboratory, Gayatripuram 1st stage, Mysuru, Balaji Diagnostic laboratory, Hospet, Annapurna multi-speciality hospital, Gangavathi, Rotaract club of NMIT, Bangalore. [11] Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. Diabet Med. 1998; 13:539–553.

[12] [http://knowledgetablet.blogspot.in/2011/12/bloodsugarblood glucose-tests-rbs-fbs.html](http://knowledgetablet.blogspot.in/2011/12/bloodsugarblood%20glucose-tests-rbs-fbs.html).

[13] <http://www.merckmanuals.com/home/hormonal-andmetabolic-disorders/diabetes-mellitus-dm/diabetesmellitus>

[14] Mwangi MW, Githinji GG, Githinji FW. Knowledge and awareness of diabetic retinopathy amongst diabetic patients in kenyatta national hospital, kenya. International Journal of Humanities and Social Science. 2011; 13(21):140–146.

[15] World Health Organization. Diabetes Programme. Country and Regional Data on Diabetes. WHO African Region; 2012.Available from: http://www.who.int/diabetes/facts/ world\_figures/en/ (accessed 13 Feb 2013)

[16] Mafomekong Ayuk Foma, Yauba Saidu,corresponding, Semeeh Akinwale Omoleke, and James Jafali "Awareness of diabetes mellitus among diabetic patients in the Gambia: a strong case for health education and promotion", BMC Public Health. 2013; 13: 1124. Published online 2013 Dec 5.

[17] Danquah I, Bedu-Addo G, Terpe K-J, Micah F, Amoako Y, Awuku Y, Dietz E, van der Giet M, Spranger J, Mockenhaupt F. Diabetes mellitus type 2 in urban Ghana: characteristics and associated factors. BMC Public Health. 2012; 13(1):210. doi: 10.1186/1471-2458-12-210.