International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-2, July 2019

Diabot: A Predictive Medical Chatbot using Ensemble Learning

Manish Bali, Samahit Mohanty, Subarna Chatterjee, Manash Sarma, Rajesh Puravankara

Abstract*: Accessibility to medical knowledge and healthcare costs are the two major impediments for common man. Conversational agents like Medical chatbots, which are designed keeping in view medical applications can potentially address these issues. Chatbots can either be generic or disease-specific in nature. Diabetes is a non-communicable disease and early detection of the same can let people know about the serious consequences of this disorder and help save human lives. In this paper, we have developed a generic text-to-text 'Diabot' – a DIAgnostic chatBOT which engages patients in conversation using advanced Natural Language Understanding (NLU) techniques to provide personalized prediction using the general health dataset and based on the various symptoms sought from the patient. The design is further extended as a DIAbetes chatBOT for specialized Diabetes prediction using the Pima Indian diabetes dataset for suggesting proactive preventive measures to be taken. For prediction, there exists multiple classification algorithms in Machine Learning which can be used based on their accuracy. However, rather than considering only one model and hoping this model is the best or most accurate predictor we can make, the novelty in this paper lies in Ensemble learning, which is a meta-algorithm that combines a myriad of weaker models and averages them to produce one final balanced and accurate model. From literature reviews, it is observed that very little research has happened in ensemble methods to increase prediction accuracy. The paper presents a state-of-the art Diabot design with an undemanding front-end interface for common man using React UI, RASA NLU based text pre-processing, quantitative performance comparison of various machine learning algorithms as standalone classifiers and combining them all in a majority voting ensemble. It is observed that the chatbot is able to interact seamlessly with all patients based on the symptoms sought. The accuracy of Ensemble model is balanced for general health prediction and highest for diabetes prediction among all weak learners considered which provides motivation for further exploring ensemble techniques in this domain.

Index Terms*: Chatbot, Diabetes, Ensemble learning, Machine Learning, Ensemble model, Natural Language Understanding (NLU)

Revised Manuscript Received on 30 July 2019.

***** Correspondence Author
Manish Bali, PG Student, Ramaiah University, Bangalore.
Samahit Mohanty, PG Student, Ramaiah University, Bangalore.
Dr.Subarna Chatterjee, Assistant Professor Ramaiah University, Bangalore
Manash Sarma, PG Student, Ramaiah University, Bangalore.
Rajesh Puravankara, PG Student, Ramaiah University, Bangalore.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

I. INTRODUCTION

Clinical Big data is the next big challenge in healthcare systems and computational analytics is a new trend to mine medical intelligence and use it for future predictions. Mining medical intelligence from huge amounts of clinical data can help up a smart and proactive healthcare system focussing on patient care with added benefits of reduced medical cost and hospital revisits. A lot of diseases can be cured if diagnosed well in advance. A chatbot is a human-machine interface which interacts or communicates through textual or speech mechanism. Today Chatbots are growing at a pace what websites were in late 1990's and 2000 as shown in Fig 1. Medical chatbots can be used in many areas like diagnosing a generic disease, deep diagnosis of any particular disease or for patient care. Like in Medical parlance, we have general physician and a specialist, in this research, we mimic a similar state using NLU & Advanced ML algorithms to first diagnose a generic disease using a text-to-text conversational Diabot and then extend this study as a specialization into more deeper-level predictions of diabetes. Diabetes is a non-communicable disease and early detection of it can let people know of its serious consequences and help save human lives. The common causes of rise in this disease of late are obesity and sedentary lifestyle which has given this disease an epidemic nature with epic proportions. Glucose overdose in blood is the main reason for this disease. Any food that is taken gets converted into glucose which provides energy to the human body. When this glucose is not converted into energy by anybody, the sugar level in blood rises which leads to diabetes [1] [2]. Diabetes is one of the major healthcare epidemics being faced by Indians. According to diabetes atlas, in India there are close to 40 million people who suffer with diabetes and this number is estimated to touch 70 million people by 2025. Worldwide, among diabetics every fifth person is Indian. Worldwide, around 3.4 million deaths occur due to diabetes as per a WHO report [3]. Early prediction of this disease can help prevent 80% of deaths due to complications [2]. Diabetes also causes blindness, amputation and kidney failures. Hence, its early detection is of paramount importance. Normally to diagnose diabetes, a doctor has to study a person's past history, diagnostic reports, age, weight etc. A doctor's diagnosis can be individualistic and dependent on his experience, not to mention his availability if he is a specialist. Chatbots have seen an unprecedented growth over the years as shown in Fig 1 and they have found use in many verticals, including medical diagnostics. One of the ways to detect diabetes early can be through use of these devices.

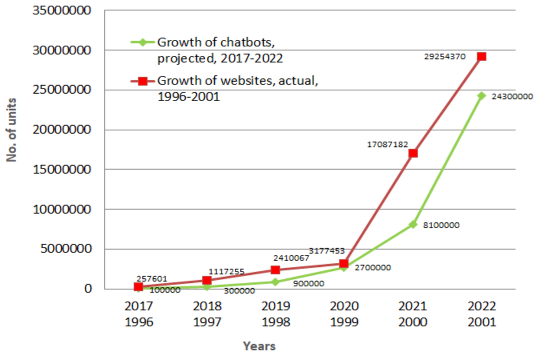

Fig 1: Growth of Chatbots

Therefore, in this work we present a NLU-based chatbot which engages with the patient and checks for various symptoms. It predicts various diseases generically as well as specifically. Specifically if the generic prediction is diabetes, it goes deeper into its prediction by considering diabetes-specific attributes. The work combines an ensemble of five classifiers – Multinomial Naïve Bayes (MNB), Decision Tree (DT), Random Forest (RF), Bernoulli Naïve Bayes (BNB) and Support vector machine (SVM) in the case of generic disease prediction using the General health dataset. In the case of Diabetes prediction, an ensemble of six classifiers - Naïve Bayes (NB), Decision tree (DT), Random forest(RF), K-nearest neighbour(KNN), Logistic regression (LR) and Gradient boosting (GB) are applied on "the Pima Indian diabetes (PID) data set" [13]. The balance sections of the presented manuscript are divided as follows: section II discusses the related work, which is followed by the discussion of the various existing classifiers for diabetes prediction. Section IV discusses the proposed system and its implementation. Section V presents the results and discussion, while the overall results and conclusion is presented in Section VI. References used are given at the end of the manuscript.

II. RELATED WORK

This section briefs some of the key research done in chatbot design and classification for disease diagnosis. An overview on chatbot Design strategies in Speech Conversation Systems by Sameera A et all [24] discovered that the advancement and improvement of chatbot configuration isn't developing at an anticipated rate because of the assortment of techniques and methodologies used to design a chatbot (no generalization). The systems for chatbot configuration are as yet an issue for discussion and no common methodology has yet been recognized. Investigators have so far worked in isolation with hesitance to reveal any improved methods that they have discovered, hindering the enhancements to chatbot design. Additionally universally acceptable chatbots need enhancements by crafting in-depth knowledge bases which are inadequate as on date. N. Jyothirmayi et all [21] found that chatbots are short of mining results as anticipated. Additionally, they don't contain adequate natural language characteristics. Monica Agrawal et all [27] in their work endeavoured to build a text-to-text diagnosis bot assemble that draws patients into discussion about their medical issues and gives a customized diagnosis depending based on their symptoms and past profile. Be that as it may, their algorithmic accuracy, recall and diagnosis percentage was low. Divya. S et all [23] in their chatbot design discovered its efficiency lacking and recommended enhancements by including more mix of words and expanding the utilization of database with the goal that the medical chabot could deal with all kind of diseases and become generic. Likewise including voice chat was proposed as future scope of research. Many individual classifier algorithms have been used for disease diagnosis. H.Temurtas et all [4], O.Karan et all [5] and T.Jayalaxhmi et all [6] used artificial neural network. N.H.Barakat et all used support vector machine [7]. Y.Huang et all used Naïve Bayes, Decision tree and Nearest-neighbour [8]. M.A.Chikh used only nearest neighbour [9], etc. Additionally, hybrid models that utilize the power of various classifiers have also been proposed by E.D.Ubeyli et all [10], D.Calisir [11] and M. Khashei et all [12]. However, diagnosis decision based on the classification result of a solitary classifier or a hybrid model alone may be weak. Various classifiers, though providing complementary information, offer contradictory classification results. Hence it makes sense to combine the results of different classifiers. In the event that the decision making depends on a group of classifiers which takes individual opinion of each classifier into account through some mechanism, the misclassified data - particularly the patients who went undiagnosed by a specific classifier may be accurately diagnosed because of the right choices of different classifiers. Therefore, there is a need for better algorithmic classification performance for diagnosis via alternative and emerging computing techniques. Also for chatbot design, a variety of methods have been proposed but there exists no generalization. There is a need for an easy user interface with sufficient Natural language understanding which a common man can be comfortable with. Moreover, not much research has been carried out to build a single system combining an undemanding user interface for a chatbot, an accurate generic disease prediction model and specialized diabetes prediction model using advanced machine learning techniques such as ensemble learning.

III. OUR CONTRIBUTION

The key contributions of this research are:

- Development of a user-friendly text-to-text chat interface at the front-end to engage with the patient
- Generic disease prediction ensemble or meta-classifier combining all weak classifiers with quantitative performance assessments
- Diabetes prediction ensemble or meta-classifier combining all weak classifiers with quantitative performance assessments
- High-performance generic framework for future disease-specific chatbot designs.

Compared with other researches, the approach offers an end-to-end medical chatbot design with an easy UI, generic prediction and a disease-specific specialization implementation. With improved accuracy using ensemble learning, more diabetes patients (and other diseases in future designs)

can get to know about this disease and accordingly make proactive steps like changes in life-style, medications etc. to prevent or keep this disease and its complications at bay.

Retrieval Number: B2196078219/2019©BEIESP
DOI: 10.35940/ijrte.B2196.078219
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

IV. SYSTEM MODEL

At a high level, the system consists of a front-end User Interface (UI) for the patient to chat with the bot. It is developed using React UI platform which uses HTML and JavaScript. The chatbot communicates with the NLU engine at the backend via API calls. The NLU engine is developed using RASA NLU platform. Two models- one for generic health prediction and another for advanced diabetes prediction are trained at the backend using the general health dataset and the Pima Indian diabetes dataset which provide the necessary diagnosis decision to the NLU engine which is trained to provide output based on user queries. A brief schematic of the system is shown in Fig 2.

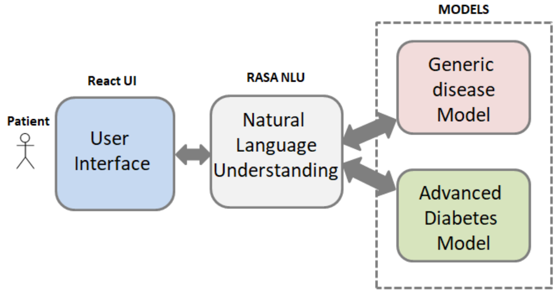

Fig 2: Brief schematic of the system

The detailed architecture of the advanced diabetes model – its training, testing & ensemble learning is shown in Fig 3. The same applies to the generic disease model also albeit different classifiers.

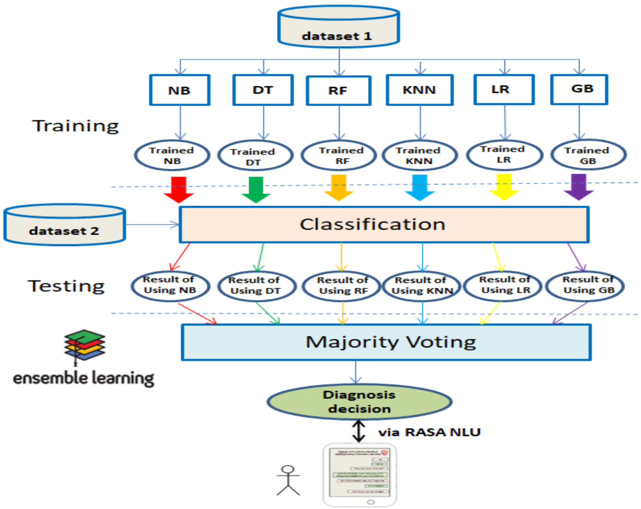

Fig 3: Procedure of training, testing & Ensemble learning for Diabetes dataset

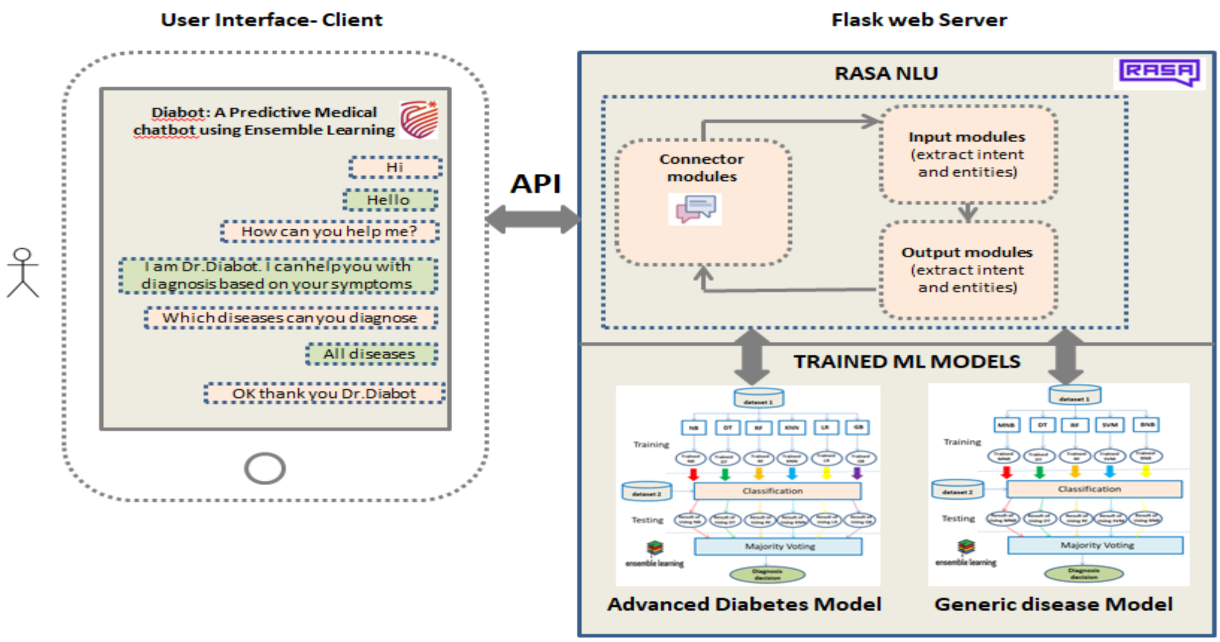

Fig 4: Detailed Chatbot architecture

The detailed architecture of the Chatbot is shown in Fig 4. It uses a client-server technology. The client is developed using the semantic UI React library, which is an open-source JavaScript library. It uses HTML for development framework. It communicates with a trained NLU engine at the backend.

The NLU engine is connected to both the generic health and the advanced diabetes models and their outputs (Model data) are saved in 'Pickle' file format. On receiving an API call, the model data in the pickle file format is loaded and made available to the UI via the Flask webserver. The actual image of the chatbot UI developed is shown in Fig 5.

Ensemble Learning

One of the key contributions in this work is Ensemble learning. Ensembling depends on the presumption that diverse models trained autonomously are probably going to be useful for various reasons: each model looks at marginally different parts of the data to make predictions, getting some portion of reality however not every last bit of it [25]. The popular methods of combining the classifiers in ensemble learning are mixture of experts, majority voting ensemble, boosting, bagging and stacking. Majority voting ensemble is actually a combiner method that can be used alongside stacking based ensemble learning. Stacking is based on a heterogeneous set of weak learners. Every classifier is trained autonomously and final choice is made by a majority vote, averaging the result [26]. Current work uses stacking-based ensemble learning with majority voting ensemble as combiner. There are five (5) weak learners for generic disease model considered and six (06) in case of the diabetes model.

The dataset is split into 80:20 ratio, 80% for training and 20% for testing. To combine the performance of each classifier, the following majority ensemble voting rule is used. It works on labels. Below, $d_{t,j}$ can be 1 or 0. If classifier 't' chooses 'j', the value is 1. Then the ensemble selects class 'J' which gets the highest total number of votes:

$$\sum_{t=1}^{T} d_{t,j}(X) = max_{j=1,\dots,c} \sum_{t=1}^{T} d_{t,j}$$

For generic disease prediction, the general health dataset is considered. It is a table with many symptoms/features and an associated disease for a set of above symptoms. In processing, we have created a transpose of this matrix and assigned values to all features. There are about 40 features and have been used in training and testing our algorithms.

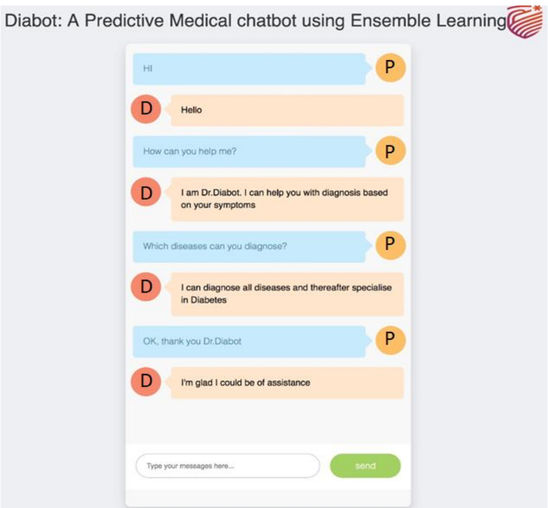

Fig 5: Actual image of the user interface developed

For diabetes prediction, the most popular Pima Indian diabetes data set [19] was used for prediction of diabetes. The data set consists of 2 sets of subjects, but all females with minimum age of 21 years – one are diabetes patients which are 268 in number and second normal subjects totalling 500. All subjects belong to Pima Indian heritage. The samples consist of 8 attributes as listed below and one predictor/class variable which have values 0 and 1. Here, output 1 indicates tested positive for diabetes and 0 indicates the contradictory i.e. no diabetes. The eight attributes are:

1. Pregnancies
2. Glucose
3. Blood pressure
4. Skin thickness
5. Insulin
6. Body mass index
7. Diabetes pedigree function
8. Age
9. Outcome or class variable (0/1)

We carry out data pre-processing and remove for example values like zero which are biologically impossible (e.g. blood pressure). Next, feature selection is done and we found out that all attributes are mandatory as there exists strong correlation among them and for the sake of better accuracy as shown in Fig 6. It shows that darker the colour, the more dependency between

the variables exist. We also plot a histogram of probabilities between each of the attribute being a 0 or 1 as shown in Fig 7. Five (05) weak learners or algorithms in case of generic disease model and six (06) weak learners or algorithms in case of diabetes model are run on the data to measure accuracy. Ensemble model combines the performance of all weak learners and accuracy measured.

V. RESULTS AND DISCUSSION

The matrix of correlations between the variables is shown in Fig 6 and the histogram of probabilities between each of the attributes being 0 or 1 is shown in Fig 7.

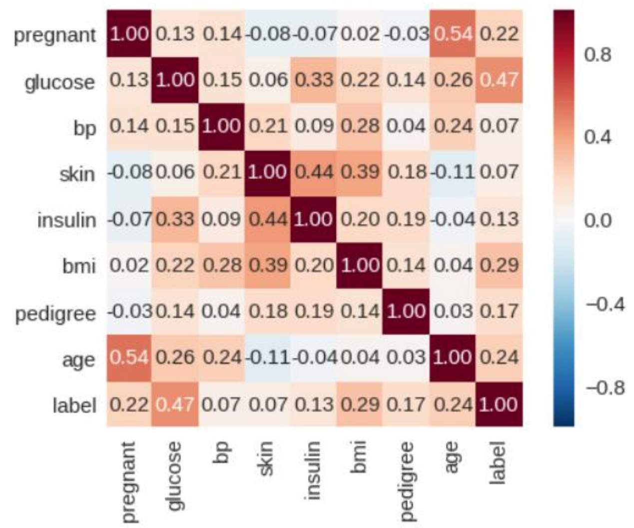

Fig 6: Correlation matrix between variables

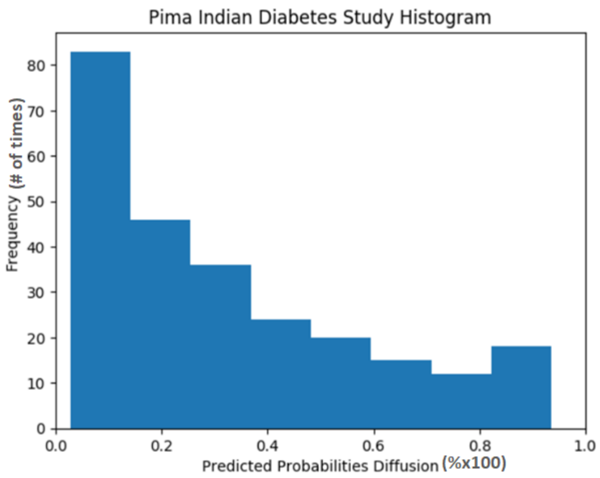

Fig 7: Histogram study of probabilities between attributes

Table 1 showcases the performance of all the classification algorithms separately and also the ensemble of classifiers using majority voting for the generic disease dataset. Even though the accuracy of Ensemble classifier is lower than Multinomial and Bernoulli's Naïve Bayes, our intension is not to only improve accuracy but to combine all classifiers via the ensemble model and get a performance which is more balanced by taking along all the weak learners which is achieved. Fig 8 compares the accuracy of all 05 models. It is observed that least accuracy is obtained on the Decision Tree model. Overall the accuracy difference between all the five classifiers is not significant. Table 2 showcases the performance comparison of each of the classification algorithm separately and also the ensemble using majority voting for the diabetes dataset. The performance of the Ensemble classifier is higher than all other classifiers.

Table 1: Performance comparison of classifier(s) for generic health dataset

Sl. No.	Classifier	Accuracy (%)
1	Multinomial Naïve Bayes	91
2	Random Forest	87
3	Decision Tree	80
4	SVM	86
5	Bernoulli Naïve Bayes	89
6	**Ensemble classifier**	**86**

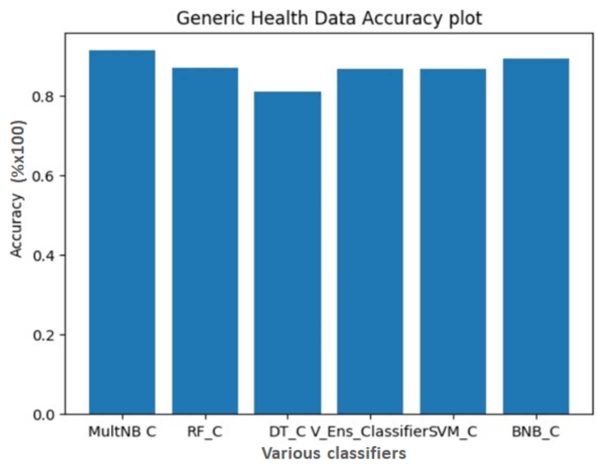

Fig 8: Comparison of Model accuracy (Generic Health data)

Fig 9 compares the accuracy of all 6 models. It is observed that The Naïve Bayes and Decision Tree models have the least accuracy. Also the accuracy of Logistic regression and Ensemble learning is very close with very little difference.
Fig 10 plots the Receiver operating curve (ROC) curve for the Diabetes classifier. Is shows perfect discrimination, which goes to prove higher overall accuracy of our test.

Table 2: Performance comparison of classifier(s) for Pima Indian Diabetes dataset

Sl. No.	Classifier	Accuracy (%)
1	Naïve Bayes	52
2	Decision tree	64
3	Random Forest	80
4	K-Nearest neighbour	73
5	Logistic Regression	84
6	Gradient Boost	82
7	**Ensemble classifier**	**84.2**

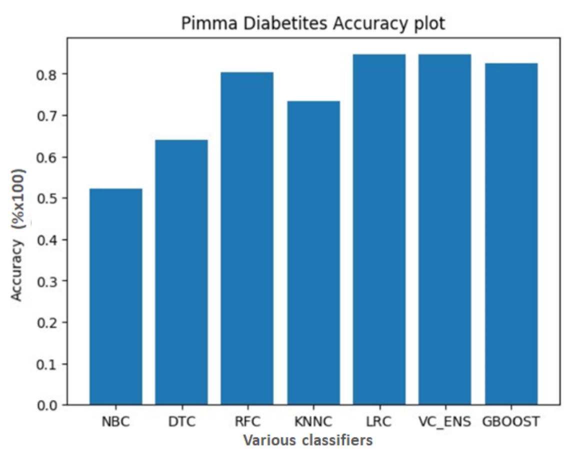

Fig 9: Comparison of Model accuracy (Pima Indian Diabetes data)

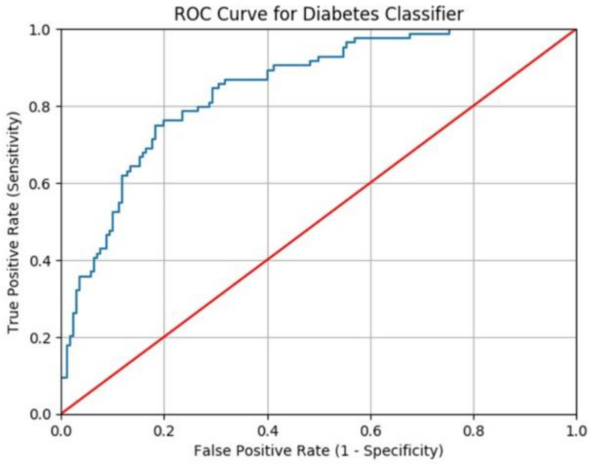

Fig 10: ROC curve for the Diabetes classifier test

VI. DISCUSSION AND CONCLUSION

In summary, two models have been proposed - one is for diagnosis of a generic disease and second is a more specific diabetes prediction model. The challenges in a real-time implementation are mainly related to accuracy. Here, we have used the static general health and Pima Indian diabetes datasets respectively. The larger and complex the dataset, the better the accuracy to implement it in a real-time scenario. Also, the chatbot needs to be trained for more specialized disease prediction (a disease-specific BOT) beyond just the generic prediction if we have to make use of this in a real-time implementation. The advantages of this research though are manifold. Firstly, the generic framework can be used for any disease prediction. This will help cut-down the visit of patients to clinics or hospitals for basic self-diagnosis and medication. Secondly, the generic framework can be extended to develop more complex disease-specific chatbots. This has been demonstrated with the DIAbetes chatBOT model, which is a specialized implementation for advanced diabetes prediction. This will help predict ailments which are becoming epidemics in early stages for taking proactive preventive measures and for any life-style changes. Also, since the results are derived using ensemble learning of all classifiers and not a single classifier that could possibly dominate, it is a simple and efficient approach to combine weak and/or dominant classifiers while providing a good balanced output. As part of the future work, the work would concentrate on ingesting patient vitals directly into a diabot server using IoT sensors and building a recommendation engine using the generic framework suggested here. Also smartphones are becoming powerful by the day and are seen as personal assistants by their users. Integrating health parameter APIs dynamically like fitness data from wearable's, diet, water intake etc. into future chatbot designs both on Android and iOS platform can be potentially disrupting technology. Mobile bots can revolutionize the healthcare Industry and can potentially replace the mundane thermometers that are usually found at homes. Increasing the size and quality of ensemble i.e. including more weak classifiers and better performing classifiers is in future scope of research.

REFERENCES

1. K. Kaul, J. M. Tarr, S. I. Ahmad, E. M. Kohner, R. Chibber, 2012. Introduction to Diabetes Mellitus, S.I Ahmad, Diabetes: An Old disease, a New Insight, Landes Bioscience and Springer Science and Business Media. 771:1-11.
2. Prableen Kaur, Manik Sharma. 2017. "A Survey on Using Nature Inspired Computing For Fatal Disease Diagnosis". International Journal of Information System Modelling and Design. 8.2.: 70-91.
3. S.A. Kaveeshwar, J.Cornwall, 2014. The current state of diabetes mellitus in India. Australia Med J. 7(1): 45-48.
4. H. Temurtas, N. Yumusak, and F. Temurtas, "A comparative study on diabetes diagnosis using neural networks", Expert Systems with Applications, vol. 36, pp. 8610 – 8615, 2009.
5. O. Karan, C. Bayraktar, H. Gumuskaya, and B. Karlik, "Diagnosing diabetes using neural networks on small mobile devices", Expert Systems with Applications, vol. 39, pp. 54 – 60, 2012.
6. T. Jayalakshmi, and A. Santhakumaran, "A novel classification method for diagnosis of diabetes mellitus using artificial neural networks", the Proceedings of 2010 International Conference on Data Storage and Data Engineering, Bangalore, pp. 159 – 163, 2010.
7. N. H. Barakat and A. P. Bradley, "Intelligible Support vector machines for diagnosis of diabetes mellitus", IEEE Transactions on Information Technology in Biomedicine, vol. 14 (4), pp. 1114 –1120, 2010.
8. Y. Huang, P. McCullagh, N. Black, and R. Harper, "Feature Selection and classification model construction on type 2 diabetic patients' data", Artificial Intelligence in Medicine, vol. 41, pp. 251 – 262, 2007.
9. M. A. Chikh, M. Saidi, and N. Settouti, "Diagnosis of diabetes diseases using an artificial immune recognition systems2 (AIRS2) with fuzzy knearest neighbour", Journal of Medical Systems, vol. 36 (5), 2721 –2729, 2012.
10. E. D. Ubeyli, "Automatic diagnosis of diabetes using adaptive neurofuzzy inference systems", Expert Systems, vol. 27 (4), pp. 259 – 266, 2010.
11. D. Calisir, and E. Dogantekin, "An automatic diabetes diagnosis system based on LDA-wavelet support vector machine classifier", Expert Systems with Applications, vol. 38, pp. 8311 – 8315, 2011.
12. M. Khashei, S. Eftekhari, J. Parvizian, "Diagnosing diabetes type 2 using a soft intelligent binary classification model", Review of Bioinformatics and Biometrics, vol. 1 (1), pp. 9 – 23, 2012.
13. Pima Indians Diabetes Data Set, [Online], Available: http://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes
14. Dharwadkar Rashmi and Deshpande A. Neeta (2018) 'Pharmabot - A Recommendation on General Medicines– Survey' in International Journal of Innovative Research in Computer and Communication Engineering Vol. 6, Issue 6
15. N. Jyothirmayi, A. Soniya, Y. Grace, Kishor C. Kumar Reddy and B.V. Ramana Murthy, 2019. 'Survey on Chatbot conversational system' in Journal of Applied Science and Computations Volume VI, Issue I
16. M. Sharma, G. Singh and R. Singh, 2018. 'An Advanced Conceptual Diagnostic Healthcare Framework for Diabetes and Cardiovascular Disorders 'in EAI Endorsed Transactions on Scalable Information Systems
17. S Divya, V Indumathi, S Ishwarya, M Priyasankari and Devi Kalpana (2018) 'A Self-Diagnosis Medical Chatbot Using Artificial Intelligence' in Journal of Web Development and Web Designing Volume 3 Issue 1
18. Abdul-Kader A. Sameera and Woods John (2015) 'Survey on Chatbot Design Techniques in Speech Conversation Systems' in International Journal of Advanced Computer Science and Applications, Vol. 6, No. 7
19. K. Bache and M. Lichman, UCI Machine Learning Repository, [Online], available: http://archive.ics.uci.edu/ml
20. Dharwadkar Rashmi., Deshpande A., Neeta, "Pharmabot - A Recommendation on General Medicines – Survey", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 6, Issue 6, 2018.
21. Jyothirmayi N., Soniya A., Grace Y., Reddy Kumar Kishor Murthy Ramana B.V., "Survey on Chatbot conversational system", Journal of Applied Science and Computations, Volume VI, Issue I, 2019.
22. Sharma M., Singh G., Singh R., "An Advanced Conceptual Diagnostic Healthcare Framework for Diabetes and Cardiovascular Disorders", EAI Endorsed Transactions on Scalable Information Systems, 2018.
23.

24. Divya,S., Indumathi,V., Ishwarya, S., Priyasankari M., Devi, Kalpana, "A Self-Diagnosis Medical Chatbot Using Artificial Intelligence", Journal of Web Development and Web Designing, Volume 3 Issue 1, 2018.
25. Kader,Abdul, Sameera, A., John, Woods, "Survey on Chatbot Design Techniques in Speech Conversation Systems", International Journal of Advanced Computer Science and Applications, Vol. 6, No. 7, 2015.
26. Chollet, Francois., (2018). Deep Learning with Python. Manning Publications.
27. Bonaccorso, Giuseppe., (2018). Machine Learning Algorithms (2nd ed.). Packt. Publishing.
28. Agrawal, Monica, Cheng, Janette, Tran, Caelin, "What's Up, Doc? A Medical Diagnosis Bot", Spoken Language Processing (CS224S), Spring 2017, Stanford University

AUTHORS PROFILE

Manish Bali has 27+ years of experience in Enterprise Computing, IT and technology services domain having held various senior level positions in leading Indian and MNC cos. He holds a B.E. in Electronics from Bangalore University and an MBA. He is currently pursuing M.Tech in Machine Learning & Intelligent Systems from RUAS. He is an Adjunct Professor in Computer Science at MIT-ADT University Pune, Chitkara University Chandigarh and Member-Board of Studies at various Engineering colleges across the country helping to bridge the Industry-Academia divide. He is on Advisory board of start-ups in the field of HPC/AI/ML/DL/IoT. He is an active member of various Industry & Academic bodies like ACM and HiPC.

Samahit is a software solution architect at Intel Bangalore with 22+ years of experience almost in all areas of databases, strong experience in architecting mission critical database system and data warehousing systems. He is doing part time M Tech course in Machine Learning and Intelligent Systems from RUAS and currently focusing on genomic research using Machine Learning algorithms.

Dr.Subarna Chatterjee has a PhD in Medical Imaging from IIEST Shibpur, M.E. (CSE specialization in Embedded System) from WBUT (Gold medal for 1st rank) and B.E. (IT) from MCKV Institute of Engg., (V.U.). She is currently working as an Asst. Prof. in Dept. of CSE, RUAS Bangalore. She has many papers in National and International Journals and was awarded INSPIRE Fellowship by DST Govt. of India and selected JRF in SMST at IIT KGP, sponsored by TI(I).

Manash Sarma is currently working as Advisory Engineer in IBM India Pvt. Ltd. Prior to IBM he worked in Canon, Draeger, NIIT Technologies etc. He has 24 years of experience in software development in various technologies and domain areas like power systems virtualization, power systems firmware, digital video, printing, image processing, route and pricing in ticket, health and medical. He is doing M.Tech course in Machine Learning and Intelligent Systems from RUAS and currently focusing on genomic research using Machine Learning algorithms.

Rajesh Puravankara is currently pursuing his final year M Tech in Machine Learning and Intelligent Systems from RUAS. He received his Bachelor's degree in Computer Engineering from Mangalore University, Karnataka, India. He works as a Manager in a Technology firm in the field of Digital Marketing and Analytics. His interest is predictive analytics and natural language processing.