Machine Learning Approaches for Diabetes Classification: Perspectives to Artificial Intelligence Methods Updating

Giuseppe Mainenti[1], Lelio Campanile[2], Fiammetta Marulli[2,*], Carlo Ricciardi[3] and Antonio S. Valente[4]

[1]*DIEM, University of Salerno, Salerno, Italy*

[2]*Dep. di Matematica e Fisica, Università Della Campania "L. Vanvitelli", Caserta, Italy*

[3]*Department of Advanced Biomedical Sciences, University of Naples "Federico II", Naples, Italy*

[4]*Public Funding Accountable, San Nicola La Strada, Caserta, Italy*

Keywords: Diabetes Classification, Diabetes Management, Machine Learning, Artificial Intelligence, Big Data Analytics.

Abstract: In recent years the application of Machine Learning (ML) and Artificial Intelligence (AI) techniques in healthcare helped clinicians to improve the management of chronic patients. Diabetes is among the most common chronic illness in the world for which often is still challenging do an early detection and a correct classification of type of diabetes to an individual. In fact it often depends on the circumstances present at the time of diagnosis, and many diabetic individuals do not easily fit into a single class. The aim is this paper is the application of ML techniques in order to classify the occurrence of different mellitus diabetes on the base of clinical data obtained from diabetic patients during the daily hospitals activities.

1 INTRODUCTION

Recent estimation accounts about 460 million people worldwide affected by diabetes, and yet 1 in 2 persons remain untreated or undiscovered, causing blindness, fingers amputations, kidney failures and almost doubling the risk of heart attack and all-cause mortality, leading to hospitalization, long-term complications, and higher costs also for healthcare infrastructures. According to International Diabetes Federation previsions, for the next 20 years, the number of people affected by diabetes will reach about 700 million (IDF, 2017).

In (Patterson et al., 2019) methods, results and limitations of the 2019 International Diabetes Federation (IDF) Diabetes Atlas 9th edition are described, thus providing an estimation of worldwide numbers of cases of type 1 diabetes in children and adolescents. The performed research provided as insights that incidence rates were available for 45% of countries, ranging from 6% in the sub-Saharan Africa region to 77% in the European region), thus concluding that the worldwide estimation for the number of children and adolescents with type 1 diabetes continue to increase.

In this paper we experimentally analyze the adoption of AI in medicine with the aim to investigate how it could improve accuracy of diagnosis, making life easier to patients and clinicians. The application of AI will open the development of new care paradigms, aiding peoples in the administration of drugs, as well as in the implementation of personal medical assistants in every smartphone. The paper is organized as follows: Section 2 describes related works and the topic related literature; Section 3 explains the proposed research, describing the methods, the classification algorithms performed and the motivating case study. Section 4 describes the experimental design and setting as the data set and the evaluation metrics. Section 5 discusses the results from the performed experiments and Section 6 explains conclusions and future improvements for the presented research.

2 RELATED WORKS

Applications of ML, AI and cognitive computing (CC) offer effective promises in healthcare domain. The early detection of most of health diseases is cru-

*Corresponding Author.

Mainenti, G., Campanile, L., Marulli, F., Ricciardi, C. and Valente, A.
Machine Learning Approaches for Diabetes Classification: Perspectives to Artificial Intelligence Methods Updating.
DOI: 10.5220/0009839405330540
In *Proceedings of the 5th International Conference on Internet of Things, Big Data and Security (IoTBDS 2020)*, pages 533-540
ISBN: 978-989-758-426-8

cial to prevent tragic consequences and automatic systems and algorithms provide an effective support for their diagnosis. Diabets represents one of the most spread and ever increasing disease affecting the world population, as also for the carotid disease.

The study provided in (Verde and De Pietro, 2018) investigates several machine learning techniques able of detecting the presence of a carotid disease; such study compares performances of different machine learning methods by analysing the Heart Rate Variability (HRV) parameters of electrocardiographic signals selected from an on line available database on the Physionet website. The results obtained by the comparative study are provided in terms of accuracy, precision, recall and F-measure metrics.

Nowadays, the adoption of mobile technologies in the healthcare sector is also increasing significantly, since they provide promising solutions for people who desire the detection, monitoring, and treatment of their health conditions anywhere and at any time, beyond offering ways for communicating multimedia content (e.g. clinical audio-visual notes and medical records). This aspect is investigated in (Verde et al., 2018) where several machine learning techniques are compared for supporting investigations on voice pathology and disorders detection, referring specifically to the dysphonia, an alteration of the voice quality affecting about one person in three at least once in his/her lifetime. The results of the performed study are provided in terms of accuracy, sensitivity, specificity, and receiver operating characteristic area. Alternative algorithms, as support vector machine or decision tree, are finally suggested, depending on the features evaluated and feature selection methods.

As for diabetes, ML and AI applications and researches are quickly growing even more, as it is evidenced by the estimation in the number of related articles indexed in the Google Scholar database(Contreras and Vehí, 2018).

In (Dankwa-Mullan et al., 2019) is performed an accurate literature analysis, addressing currently AI advances relevant to diabetes ecosystem, made of patients affected by diabetes, families, caregivers and clinicians. Such literature recognition was performed by consulting the online database PubMed (Fiorini et al., 2017). Key terms adopted for research were “diabetes” and “artificial intelligence”, by which they met 450 publicly available articles and sources of information from 2009 to 2019.

Furthermore, global market insights related to the 2019 are analyzed in (Zhou and Myrzashova, 2020), and research results reveal that an explosive growth of 40% is expected from 2017 to 2024 in the use of AI in the health care field. The use of AI in medicine today can significantly improve the accuracy of diagnosis, make life easier for patients with various diseases, and with the development of technology it will make the emergence of highly effective personal medicines real, as well as a personal medical assistant in every smartphone. Nowadays, another clinical challenge is the realization of personalized treatments, which falls into the more general paradigm of the precision medicine.

In (Coronato and Naeem, 2019) a Reinforcement Learning paradigm is adopted to build an intelligent system, able to self-adapt to users' skills aiming at assisting them in the healthcare treatment. A method to classify patients affected by diabetes using a set of characteristic selected in according to World Health Organization criteria is described in (Mercaldo et al., 2017), where the state of the art ML algorithms were applied for valuating real-world data. Experiments obtained a 0.770 value for the precision metric and a recall equal to 0.775 using the HoeffdingTree algorithm. (Brunese et al., 2019) proposes a neural network-based method aimed to discriminate between different lung cancer types, to assist medics and radiologists in the diagnosis formulation. In this work, they adopt a set of 30 radiomic feature directly obtained from magnetic resonance; the neural network model is tuned during the variation of the momentum and the loss functions and was evaluated on a data set consisting of 2000 MRI labelled through medical reports.

Researches in (Osman and Aljahdali, 2017) proposed an integration approach between the SVM technique and K-means clustering algorithms to diagnose diabetes disease. The focus of the method was adjusted so that only the most important features received attention. They performed the T-Test order to quantify the improvements achieved by their approach before and after combination process between K-means and SVM algorithms. Authors in (Nanda et al., 2011) develop a model for the prediction of gestational diabetes mellitus (GDM) from maternal characteristics and biochemical markers at 11 to 13 weeks' gestation. They demonstrate that in the screening study, maternal age, body mass index, racial origin, previous history of GDM and macrosomic neonate were significant independent predictors of future. The detection rate was 61.6% at a false-positive rate of 20% and the detection increased to 74.1% by the addition of adiponectin and sex hormone-binding globulin. A data informed framework for identifying subjects with and without Type 2 Diabetes Mellitus (T2DM) from Electronic Health Records via feature engineering and ML is proposed in (Zheng et al., 2017). K-means clustering analysis is adopted in (Bennetts et al., 2013) in order to identify typical regional peak plantar pressure distributions in a group of 819 diabetic feet. The number of analyzed clusters ranged from 2 to 10 to examine the effect on the differentiation and classification of regional peak plantar pressure distributions. Such analysis aimed to provide an understanding of the variability of the regional peak plantar pressure distributions seen within the diabetic population and serves as a guide for the preemptive assessment and prevention of diabetic foot ulcers. Finally, in (Alssema et al., 2011) authors employed the Finnish diabetes risk questionnaire to identify subjects exposed at risk for drug-treated Type 2 diabetes. in this research, additional predictors were added and the risk questionnaire was updated by using clinically diagnosed and screen-detected Type 2 diabetes instead of drug-treated diabetes. The conclusion of the study was that the predictive value of the original Finnish risk questionnaire could be improved by adding information on sex, smoking and family history of diabetes.

In (Brunese et al., 2020) a research aiming to recognize the different brain cancer grades by analysing brain magnetic resonance images, since the brain cancer is one of the most aggressive tumour. The proposed method aims to identify the components of an ensemble learner. The ensemble learner is focused on the discrimination between different brain cancer grades using non invasive radiomic features, belonging to five different groups: First Order, Shape, Gray Level Co-occurrence Matrix, Gray Level Run Length Matrix and Gray Level Size Zone Matrix. As research results, authors evaluated the features effectiveness through hypothesis testing and through decision boundaries, performance analysis and calibration plots thus we select the best candidate classifiers for the ensemble learner.

3 THE PROPOSED RESEARCH

Computerization in healthcare is on the rise thus leading to large patient databases, with specific properties. ML techniques are able to examine and to extract knowledge from large databases in an automatic way (Meyfroidt et al., 2009). Several examples can be found in literature as regards the classification of diagnosis or prognosis of different pathologies in a wide range of medical fields (Ricciardi et al., 2020a), (Ricciardi et al., 2020b), (Romeo et al., 2020).

In this work a number of predictive models was implemented including Decision Tree, Random Forest, k-Nearest Neighbors (KNN), Logistic Regression and Multilayer Perceptron to classify the type of diabetes of the patients. Similar approach have been investigated in literature in order to classify cardiotocogram data (Sahin and Subasi, 2015). To conduct our analysis we exploited data of patients contained in a clinical database that has been made available within the TablHealth Project.

After analyzing the data and selecting the most suitable subset, a Data Cleaning strategy has been applied to manage the imbalance between classes and missing data. Based on the dataset obtained, a web platform was then implemented using Flask, a web micro-framework written in Python, to classify five different types of diabetes: type 1 diabetes, type 2 diabetes, gestational diabetes, diabetes Maturity-Onset Diabetes of the Young (MODY) and diabetes Latent Autoimmune Diabetes of Adult (LDA). The categorical variables considered were: age, diet ([kcal]), fasting blood sugar, diastolic blood pressure, systolic blood pressure, Body Mass Index (BMI), glycated hemoglobin ($HbA1c$), reduced glucose tolerance, syndrome metabolic disease, macrosomy, microalbuminuria, ischemic heart disease, high blood pressure and cerebral vasculopathy.

3.1 Automatic Classification of Diabetes Types

The idea of the present work is that to exploit ML techniques to investigate and build classification models that can predict with high accuracy the type of diabetes. The search for a better prediction and therefore a faster diagnosis by the specialized medical staff can be vital for the health of the diabetic patient, forced to face numerous complications that can become lethal if not managed in a timely manner.

3.2 The Methodology

In Python, the following algorithms have been implemented through the use of the Scikit-Learn framework:

- Decision Tree: a structure made up of leaves and nodes that represent the attributes and the predicted classes;
- Random Forest: it is a results of applying randomization and bagging on the decision tree to improve its accuracy;
- KNN: it is an instance-based model;
- Logistic Regression: it is a statistical-based model, usually employed for binary classification;

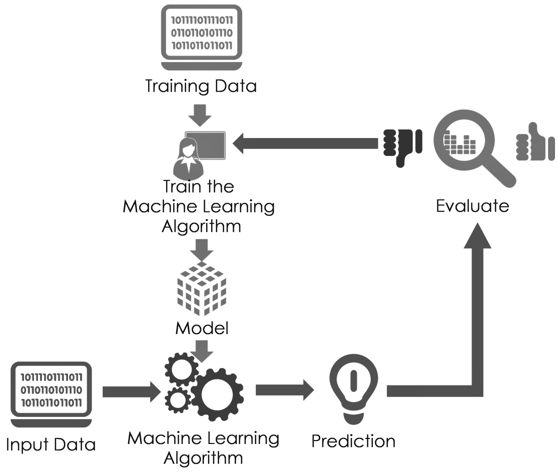

Figure 1: Classification process.

- Multilayer Perceptron: this is an example of neural network, a complex structure simulating the one of the human brain.

3.3 The Case Study

The idea behind it was to provide the specialized medical technical staff with an interface that would allow all features that were considered to be in input when creating the dataset and then in the training of the models, choose The model to be used and output is shown the resulting prediction made on the input data with the chosen model.

4 EXPERIMENT DESIGN

A ML classification algorithm to make the best needs a good dataset. The creation of a well-structured dataset suitable for the algorithm is far from trivial. In this paper work the data was made available under courtesy from the company where I did the internship. They were collected in an Excel file by medical staff from patients who made outpatient visits. We can divide the process of implementing and developing the web application into three phases:

1. Creating the dataset;
2. Implementing modeling algorithms;
3. Web application development with Flask.

4.1 Data Set

The manipulation of the dataset used for the algorithms, has been done in Jupyter Notebook environment by means the Pandas library.

Table 1: Features.

Parameter [Unit]	Type	Range
Age [Year]	Integer	[11,92]
Diet [Kcal]	Float	[1200, 2000]
Fasting blood sugar [mg/dL]	Float	[52, 434]
Systolic blood pressure [mmHg]	Float	[90, 180]
Diastolic blood pressure [mmHg]	Float	[60, 95]
BMI [kg/m^2]	Float	[18.25, 63.19]
Glycated hemoglobin [HbA1c%]	Float	[4.71, 14.40]
Reduced glucose tolerance	Integer	$\{0,1\}$
Syndrome metabolic disease	Integer	$\{0,1\}$
Macrosomy	Integer	$\{0,1\}$
Microalbuminuria [mg/L]	Float	[10, 300]
Ischemic heart disease	Integer	$\{0,1\}$
High blood pressure	Integer	$\{0,1\}$
Cerebral vasculopathy	Integer	$\{0,1\}$

Step 1. A data cleaning and analysis has been executed on the information contained into the Excel sheet in which were registered all patient data and output classes. To this aim data have been imported into a dataframe and analyzed thanks to command *dataframe.info()* of Pandas. The result of the analysis showed several errors, such as missing, noisy and inconsistent data. Therefore, we proceeded to clean and normalize data by exploiting the Python's Pandas library in order to get a good dataset. At the end of this first step we obtained a dataset of 3378 elemnts related to patients affected by diabetes.
Step 2. First of all we identified all the relevant features such as BMI, fasting blood glucose, etc. We found 14 valid features, reported in Table 1 with the type and the range for the sake of clarity.

Then, five output classes distinguishing the type of diabetes have been identified:

- Type 1 diabetes;
- Type 2 diabetes;
- Gestational diabetes;
- MODY diabetes;
- LADA diabetes;

Step 3. Finally a graphical analysis has been done to emphasize the characteristics of dataset with regarding the most significative clinical features. In Figure 2 is depicted the boxplot of the BMI with respect the Hb1c.
In 3 we reported the scatter-plot of fasting glucose with respect BMI using a gradient scale of colors

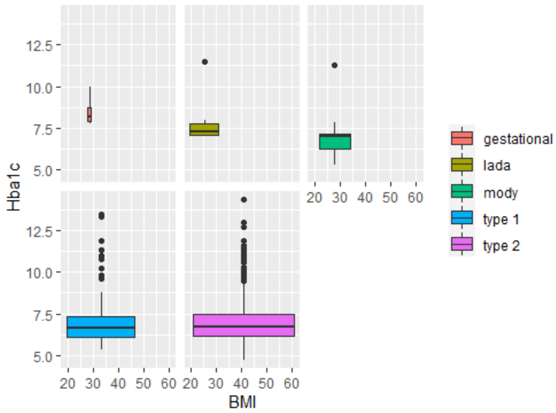

Figure 2: Boxplot for each outcome. On the x axis Bmi and on the y axis Hb1c.

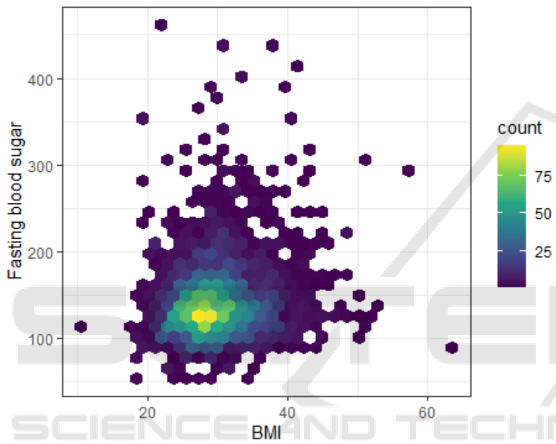

Figure 3: Frequency and fasting glucose features.

varying according to numbers of samples.

Note that the resulting dataset is characterized by a strong imbalance, which fully reflects the state of the art in literature. For proper classification, the imbalance of the dataset is a major obstacle. In these cases, the classifier used could pay special attention to the majority class at the expense of the minority class. Packages such as smoote are created to solve these cases using under/over sampling techniques, however they will not be used for this work. It is true that the imbalance can affect the performance of the algorithm, in terms of accuracy, precision and recall but not timing performance, however if the imbalance affects all the classifiers in the same way then the absolute performance of the classifier is affected but not the relative one; Based on this, getting a good result in these conditions comes from the robustness of the classifier compared to the imbalance.

4.2 Evaluation Metrics

The following evaluation metrics, well-known in literature for assessing the performance of ML algorithms (Hossin and Sulaiman, 2015), were used:

- Accuracy: the number of correct predictions over the total;
- Precision: the ratio between true positives and the overall positives;
- Recall: the ratio between true positives and the sum of true positive and false negatives;
- F1-score: an harmonic mean of precision and recall.

Each model has been evaluated with the metrics discussed above, logically related to the multiclass case, in this case scikit-learn to evaluate non-binary classification models uses a normalized or macro-media. The latter is simply calculated as the average scores of different systems and weighs all classes equally to evaluate a classifier's overall performance against the most frequent class labels. The weighted macro-average is calculated by weighing the score of each label of the class by the number of true instances at the time of the average calculation. It is useful if we have a problem of poor class balance.

4.3 Implementing Classification Algorithms

Even at this stage the use of Jupyter Notebook was decisive, because from time to time the code was modified and executed in real time. In this section, we create the models with the classification algorithms listed in Paragraph 3.2. The first step taken at this stage is to divide it into two parts, loading the dataset, and dividing the dataset into two parts, training datasets and test datasets.

1. Loading the dataset with the pandas ***read_csv()*** method;
2. Assign features to a dataframe named x and outputs to a y-named dataframe using the ***loc()*** method of the Pandas library;
3. Separate the dataset into training datasets (80% of the data) and test datasets (20% of the data), this we can do with the ***train_test_split()*** method.

Scikit-Learn provides a class and various methods for each template you want to use. The methods used to perform the training and perform the prediction are ***fit()*** and ***predict()***, respectively.

4.3.1 Decision Tree

After importing the DecisionTreeClassifier class, an instance of the class itself was created and the ***fit()*** and ***predict()*** methods were executed. For this algorithm, two models were tested, one with all the parameters set by default and the other with the addition of the parameter ***class_weight*** = "**balanced**". This parameter has been tested to cope with the imbalance of the dataset, and as can be seen from the results obtained (section 5) it is much better performing than the first model tested.

4.3.2 Random Forest

A random forest model can be considered as a set of decision trees. The idea of putting more decision trees together is to combine multiple simpler models to build a more robust one that offers better classification performance and is less susceptible to overfitting issues. The advantage of using random forest as a classifier is that it doesn't need to modify as many hyper-parameters to achieve great performance. After importing the RandomForestClassifier class, an instance of the class itself was created and the ***fit()*** and ***predict()*** methods were executed. As can be seen from the result, the model is able to classify all outcomes very well excluding the diabetes category *lada* presumably because it is the category with the least samples.

4.3.3 K-Nearest-Neighbors

For this model was created an object called "knn" that knows how to do KNN classification once the data is provided. All parameters of the KNeighborsClassifier class are set to the default values excluding the "n_neighbours" parameter, which is the optimization/hyper-parameter (k) parameter. The choice of the k value is of paramount importance if you want to have the maximum accuracy of the model, for this reason a method has been written that does the train and test on the model with k ranging from 0 to 25 and saves the result of accuracy, both test and train, in a list and then plots the result with the help of the *matplotlib* library. The choice of k = 1 is precisely the output of the function, where the accuracy of the model was higher than the other values of k. By evaluating the model with the selection of all the parameters set by default and then with the optimization of the hyper-parameter k you can see how all the evaluation metrics have improved.

4.3.4 Logistic Regression

Again, an instance of the LogisticRegression class was created and the ***fit()*** and ***predict()*** methods were invoked on it. All parameters of the Logistic Regression class are set to their default values. The logistic regression model has a fairly high accuracy. The most logical explanation for this behavior may be the problem of the unbalanced dataset.

4.3.5 Multilayer Peceptron

Before training the MLPerceptron model, a method was used to standardize observations with the StandardScaler integrated tool. This standardization was performed on the train set and the test set with the use of the fit_transform method of the StandardScaler class. The process of standardizing the train and test set is done because the MultiLayer Perceptron algorithm is sensitive to feature scaling. After standardization, you create an instance of the imported model. For the MLP model there are many parameters that can be changed such as the number of hidden layers(hidden_layer_sizes), the choice of activation function for hidden layers (activation), the optimization of weights between nodes (solver), and the number of eras (max_iter). In this research work, only the number of eras has changed, i.e. max_iter, which is traditionally a large number, often hundreds or thousands, and allows the algorithm to run until the model error has been sufficiently reduced to the minimum.

5 RESULTS AND DISCUSSION

Different algorithms were available for ML. This project used linear models and only scratched the surface of deep learning with multi-layer perceptron. During background study, were found to have used deep learning models. Deep learning tends to perform better than linear models. This project should have included deep learning methods than linear models. The accuracy score achieved by optimizing the models is satisfactory. In KNN model with the optimization of the hyperparameter k you can see how all the evaluation metrics have improved.The logistic regression model has a fairly high accuracy. The most logical explanation for this behavior may be the problem of the unbalanced dataset. As can be seen from the result the Random Forest model is able to classify all outcomes very well excluding the diabetes category *lada* presumably because it is the category with the least samples.Finally seeing the results of MLP model you can see that in support of an accu-

Table 2: Type 2 diabetes.

Model/Metrics	Precision	Recall	F1-Score
KNN	0.90	0.91	0.90
KNN(k=1)	0.99	0.95	0.97
LR	0.93	0.99	0.96
DT(optimized)	1.00	0.97	0.98
RF	0.99	0.99	0.99
MLP	1.00	0.99	0.99

Table 3: Type 1 diabetes.

Model/Metrics	Precision	Recall	F1-Score
KNN	0.31	0.19	0.24
kNN(k=1)	0.76	0.92	0.83
LR	0.93	0.70	0.80
DT(optimized)	0.85	1.00	0.92
RF	0.97	0.98	0.98
MLP	0.94	1.00	0.97

Table 4: Gestational diabetes.

Model/Metrics	Precision	Recall	F1-Score
KNN	0.37	0.80	0.49
KNN(k=1)	0.83	1.00	0.91
LR	0.20	0.07	0.10
DT(optimized)	0.94	1.00	0.97
RF	0.94	1.00	0.97
MLP	0.94	1.00	0.97

racy of 98.9% there are also very high precision and recall values.This data makes us understand that for the dataset used to classify the type of diabetes the MultiLayer Perceptron model is the best in terms of rating metrics, but perhaps not in terms of the speed of execution of the algorithm, which contrary to the metrics takes a significant time to run the algorithm. Below will be the results of the classifiers developed for each type of diabetes. The tables show the results of precision, recall and f1-score in a way One vs All, that is, for each individual type of diabetes are considered positive samples belonging to that type and all other negatives. Accuracy was not included in the result comparison tables because it would lead to a foreclosing due to the imbalance of the dataset.

Table 5: MODY diabetes.

Model/Metrics	Precision	Recall	F1-Score
KNN	0.68	0.91	0.79
KNN(k=1)	0.89	1.00	0.94
LR	0.78	0.58	0.67
DT(optimized)	0.89	1.00	0.84
RF	0.96	1.00	0.98
MLP	0.96	1.00	0.98

Table 6: LADA diabetes.

Model/Metrics	Precision	Recall	F1-Score
KNN	0.32	0.55	0.40
KNN(k=1)	0.79	1.0	0.88
LR	0	0	0
DT(optimized)	0.73	1.00	0.85
RF	0.88	0.64	0.74
MLP	0.85	1.00	0.92

6 CONCLUSIONS AND FUTURE WORKS

The idea underlying this work was to investigate and build, through ML techniques, classification models that can accurately predict the type of diabetes. The search for a better prediction and therefore a faster diagnosis by the specialized medical staff, can be vital for the health of the diabetic patient, forced to face numerous complications that can become lethal if not managed in a timely manner. To complete this research some further work has to be done, since the system is already in a prototyping stage and it's not ready yet to be made available to the public, since it lacks in some aspects concerning the safety and validation of the models performing the prediction. Anyway, the results we obtained can be considered satisfactory and motivate us to go further in our investigations beyond completing the robustness and safety aspects of the prototyping system. Future developments of this research carried out include:

- Refine the techniques of extracting features from raw data, in particular, integrate data also from CGM, insulin pump, Artificial pancreas System like OpenAPS and platform like Tidepool;
- Solve the problem of dataset imbalance with undersampling and/or oversampling algorithms for more accurate classification of classes with fewer samples;

- Broaden the work done for the classification of type of diabetes in other medical areas as well;
- Consider using a NO-SQL database(like MongoDb) to manage the many data and features made available from an electronic medical record;

ACKNOWLEDGEMENTS

The research presented in this paper was co-funded by the activities of Gesan s.r.l. within the TablHealth Project (CUP: B49J17000710008). This research was also partially funded by the project "Attrazione e Mobilità dei Ricercatori" Italian PON Programme (PON_AIM 2018 num. AIM1878214-2).

REFERENCES

Alssema, M., Vistisen, D., Heymans, M., Nijpels, G., Glümer, C., Zimmet, P. Z., Shaw, J. E., Eliasson, M., Stehouwer, C. D., Tabák, A. G., et al. (2011). The evaluation of screening and early detection strategies for type 2 diabetes and impaired glucose tolerance (detect-2) update of the finnish diabetes risk score for prediction of incident type 2 diabetes. *Diabetologia*, 54(5):1004–1012.

Bennetts, C. J., Owings, T. M., Erdemir, A., Botek, G., and Cavanagh, P. R. (2013). Clustering and classification of regional peak plantar pressures of diabetic feet. *Journal of biomechanics*, 46(1):19–25.

Brunese, L., Mercaldo, F., Reginelli, A., and Santone, A. (2019). Neural networks for lung cancer detection through radiomic features. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–10. IEEE.

Brunese, L., Mercaldo, F., Reginelli, A., and Santone, A. (2020). An ensemble learning approach for brain cancer detection exploiting radiomic features. *Computer methods and programs in biomedicine*, 185:105134.

Contreras, I. and Vehí, J. (2018). Artificial intelligence for diabetes management and decision support: Literature review. *Journal of Medical Internet Research*, 20.

Coronato, A. and Naeem, M. (2019). A reinforcement learning based intelligent system for the healthcare treatment assistance of patients with disabilities. In *International Symposium on Pervasive Systems, Algorithms and Networks*, pages 15–28. Springer.

Dankwa-Mullan, I., Rivo, M., Sepulveda, M., Park, Y., Snowdon, J., and Rhee, K. (2019). Transforming diabetes care through artificial intelligence: the future is here. *Population health management*, 22(3):229–242.

Fiorini, N., Lipman, D. J., and Lu, Z. (2017). Cutting edge: towards pubmed 2.0. *Elife*, 6:e28801.

Hossin, M. and Sulaiman, M. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2):1.

IDF, I. (2017). International diabetes federation. idf diabetes atlas, 8th edn. brussels, belgium: International diabetes federation, 2017.

Mercaldo, F., Nardone, V., and Santone, A. (2017). Diabetes mellitus affected patients classification and diagnosis through machine learning techniques. *Procedia computer science*, 112:2519–2528.

Meyfroidt, G., Güiza, F., Ramon, J., and Bruynooghe, M. (2009). Machine learning techniques to examine large patient databases. *Best Practice & Research Clinical Anaesthesiology*, 23(1):127–143.

Nanda, S., Savvidou, M., Syngelaki, A., Akolekar, R., and Nicolaides, K. H. (2011). Prediction of gestational diabetes mellitus by maternal factors and biomarkers at 11 to 13 weeks. *Prenatal diagnosis*, 31(2):135–141.

Osman, A. H. and Aljahdali, H. M. (2017). Diabetes disease diagnosis method based on feature extraction using k-svm. *Int J Adv Comput Sci Appl*, 8(1).

Patterson, C. C., Karuranga, S., Salpea, P., Saeedi, P., Dahlquist, G., Soltesz, G., and Ogle, G. D. (2019). Worldwide estimates of incidence, prevalence and mortality of type 1 diabetes in children and adolescents: Results from the international diabetes federation diabetes atlas. *Diabetes Research and Clinical Practice*, 157:107842.

Ricciardi, C., Cantoni, V., Improta, G., Iuppariello, L., Latessa, I., Cesarelli, M., Triassi, M., and Cuocolo, A. (2020a). Application of data mining in a cohort of italian subjects undergoing myocardial perfusion imaging at an academic medical center. *Computer Methods and Programs in Biomedicine*, page 105343.

Ricciardi, C., Edmunds, K. J., Recenti, M., Sigurdsson, S., Gudnason, V., Carraro, U., and Gargiulo, P. (2020b). Assessing cardiovascular risks from a mid-thigh ct image: a tree-based machine learning approach using radiodensitometric distributions. *Scientific Reports*, 10(1):1–13.

Romeo, V., Cuocolo, R., Ricciardi, C., Ugga, L., Cocozza, S., Verde, F., Stanzione, A., Napolitano, V., Russo, D., Improta, G., et al. (2020). Prediction of tumor grade and nodal status in oropharyngeal and oral cavity squamous-cell carcinoma using a radiomic approach. *Anticancer Research*, 40(1):271–280.

Sahin, H. and Subasi, A. (2015). Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques. *Applied Soft Computing*, 33:231–238.

Verde, L. and De Pietro, G. (2018). A machine learning approach for carotid diseases using heart rate variability features. In *HEALTHINF*, pages 658–664.

Verde, L., De Pietro, G., and Sannino, G. (2018). Voice disorder identification by using machine learning techniques. *IEEE Access*, 6:16246–16255.

Zheng, T., Xie, W., Xu, L., He, X., Zhang, Y., You, M., Yang, G., and Chen, Y. (2017). A machine learning-based framework to identify type 2 diabetes through electronic health records. *International journal of medical informatics*, 97:120–127.

Zhou, H. and Myrzashova, R. (2020). Ai based systems for diabetes treatment: a brief overview of the past and plans for the future. *Journal of Physics: Conference Series*, 1453:012063.