Health Diagnosis Cross-Platform Application Development

Richa Sharma, Sujoy Mitra, Meher Singh, Monaami Pal

Computer Engineering Department, Vivekananda Education Society of Institute of Technology, Mumbai, India.

Abstract

In our society today, people all around the globe are having major health problems like Diabetes and Heart Problems or both, and there is a great need for risk prediction these problems people are having. Most of the time we are hacked with hassles of traveling and time from reaching the health practitioner. But our mobile phone can now help us bridge the gap where we can personally diagnose ourselves based on the risk of the above-mentioned problems. Today, mobile devices have changed how we conceive software. There is a great range of development alternatives. In this paper, we propose a health diagnostic system to monitor and predict the health parameters of a patient using machine learning algorithms and PhoneGap to implement a cross-platform mobile application. Supervised machine learning is the construction of algorithms that finds the mapping function between the input parameters and output values. Data Categorization from prior information is the aim of supervised machine learning classification algorithms. Classification is administered very frequently in data science problems. This paper will give us a clear understanding of PhoneGap, cross-platform mobile application, Prediction Analysis and also how this health diagnosis mobile application works in this context.

1. INTRODUCTION

The objective of the project is to make a cross-platform mobile web application that allows users to bridge the gap between the user and the health practitioner where we can personally diagnose ourselves in a less difficult situation as most at times we are cut up with the hassles of traveling and time from reaching the health practitioner. For more severe conditions like Diabetes and Thrombosis (Heart blockage - can result in cardiac arrest) as these are one of the most common major ailments any person has today, this project helps the user to check whether one has to visit the doctor or not as per the requirements of the patient. This app is available in four different OSs, namely, Android OS, Windows OS, and iOS.

For prediction, we are going to apply Supervised Learning Algorithms and train the model with the help of training data to accept some test data and predict the corresponding output accordingly.

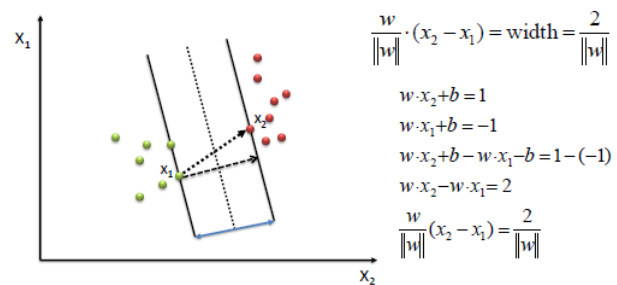## 2. METHODS AND TECHNOLOGY

2.1. Machine Learning

Machine learning (ML), a subset of AI, is the scientific study of algorithms and statistical models used by computer systems to perform selected tasks without using explicit instructions, calculating patterns, and inferring rules instead. In its application to business problems, machine learning is also referred to as predictive analytics.

2.2. Supervised Learning

Supervised learning algorithms build a mathematical model of a group of knowledge that contains both the inputs and, therefore, the desired outputs. Data were understood as training data and consists of a group of coaching examples. Each training example has one or more inputs, and therefore the desired output also referred to as a supervisory signal. There are many different types of Supervised Learning. Some of them are listed below.[1]:

2.2.1. Support Vector Machine:

In machine learning, support vector machines (SVMs, and support- vector networks) are standard learning models with related learning algorithms that analyze the data used for classification and multivariate analysis.During a high- or infinite-dimensional space, SVM builds a hyperplane or set

$$\frac{w}{\|w\|} \cdot (x_2 - x_1) = \text{width} = \frac{2}{\|w\|}$$

$$w \cdot x_2 + b = 1$$
$$w \cdot x_1 + b = -1$$
$$w \cdot x_2 + b - w \cdot x_1 - b = 1 - (-1)$$
$$w \cdot x_2 - w \cdot x_1 = 2$$
$$\frac{w}{\|w\|}(x_2 - x_1) = \frac{2}{\|w\|}$$

Equation of S.V.M

[12]of hyperplanes, which is useful for classification, regression, or other tasks like outliers detection. [10]
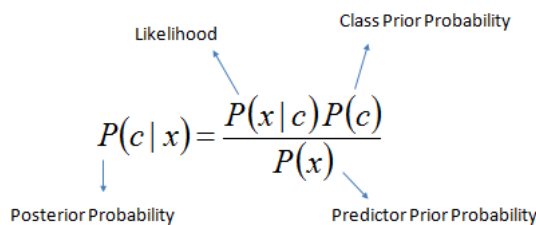
2.2.2. Naïve Bayes Classifier:

The naïve Bayes classifiers belong to a family of simple "probabilistic classifiers "with strong (naïve) independence assumptions between the features. They are among the simplest forms of the Bayesian network. But they can be integrated with the Kernel density estimation and achieve high accuracy levels.[1]

The Naïve Bayes has been extensively studied since the 1960s. It was introduced (albeit under that name) to the literary community in the early 1960s and remains a popular (basic) method of text categorization, the problem to judge documents as belonging to one category or to another (document categorization) (such as spam or legal, sports or political, etc.) by word frequencies as symbols. With correct preprocessing, it gives better results as compared to more advanced methods, including SVM. It is also highly used in automatic medical diagnosis.[2]

The Naïve Bayes classifiers are very efficient, requiring a number of parameters linear with the number of variables (factors/predictors) in the learning problem. Better data training can be achieved by evaluating a closed-form expression taking linear time, rather than by expensive iterative approximation as used for many other types of classifiers.

Naive Bayes models are known under a variety of names, including simple Bayes and independence Bayes in the statistics and computer science literature. All these names may reference the use of Bayes' theorem in the classifier's decision rule, however naïve Bayes is not (necessarily) a Bayesian method. [10]

K-Nearest Neighbours: The k-nearest neighbor's algorithm (k-NN) is a non-parametric method used for classification and regression in pattern recognition with input consisting of k closest training examples in feature space and output decided upon whether k-NN is used classification or regression

[13][14].In the k-NN classification, the result is a class member. A plurality vote of its neighbor is used to classify the object, the class most common amongst its k nearest neighbor (k is a positive integer, typically small) owns the object. If k=1, the class of that single nearest neighbor owns the object. In k-NN regression, the result is the property value for the object. It is given by the average of the values of k nearest neighbors. k-NN can be instance-based learning or lazy learning, where the function is only approximated locally, and all computation is deferred until function evaluation.[2]

Both for classification and regression, weights are assigned to the contribution of each neighbor, leading to the contribution of nearest neighbor much more to the average than the distant one. For example, a weight of 1/d is given to each neighbor, where d is the distance to the neighbor. The set of objects for which the class (for k-NN classification) or the object property value (for k-NN regression) is known are considered as neighbors. It is also treated as the training set for the algorithm, though no explicit training step is required. However, it should be known that the k-NN algorithm is sensitive to the local structure of the data. [10]

2.2.3. Random Forest Classifier:

Random forest is a supervised learning algorithm. It builds a group of decision trees called "forest" which is usually trained with the "bagging" method. The general idea of the bagging method is that a mixture of learning models increases the general result. In a decision tree, split points are chosen by finding the attribute and the value of that attribute that results in the lowest cost.

For classification problems, the Gini index refers to the cost function, which calculates the purity of the groups

Distance functions

$$\text{Euclidean} \quad \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$$

$$\text{Manhattan} \quad \sum_{i=1}^{k}|x_i - y_i|$$

$$\text{Minkowski} \quad \left(\sum_{i=1}^{k}\left(|x_i - y_i|\right)^q\right)^{1/q}$$

Fig .3: Equation for Random Forest Classifier

$$P(c\,|\,x) = \frac{P(x\,|\,c)P(c)}{P(x)}$$

Likelihood — Class Prior Probability

Posterior Probability — Predictor Prior Probability

$$P(c\,|\,X) = P(x_1\,|\,c) \times P(x_2\,|\,c) \times \cdots \times P(x_n\,|\,c) \times P(c)$$

Fig .2: Equation for Naïve Bayes Classifiers

$$K_k^{cc}(\mathbf{x}, \mathbf{z}) = \sum_{k_1,\ldots,k_d,\sum_{j=1}^{d}k_j=k} \frac{k!}{k_1!\cdots k_d!}\left(\frac{1}{d}\right)^k \prod_{j=1}^{d} \mathbf{1}_{\lceil 2^{k_j}x_j\rceil = \lceil 2^{k_j}z_j\rceil},$$

$$\text{for all } \mathbf{x}, \mathbf{z} \in [0,1]^d.$$

Fig .4: Equation for Naïve Bayes Classifiers

[15]of data created by the split point. A Gini index of 0 is perfect purity where class values are perfectly separated into two groups, in the case of a two-class classification problem. Finding the best split point in a decision tree involves evaluating the cost of each value in the training dataset for each input variable.

Random forests or random decision forests are an ensemble learning method for classification, regression, and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests overcome the overfitting of the training set for decision trees. [10]

2.3. Cross-Platform Application

Cross-platform mobile software development is creating software applications that can execute on any mobile operating system. Originally, the difficulty of making mobile applications was compounded by the difficulty of building a back-end that works across multiple platforms such as iOS, Android, and the Universal Windows Platform. Beyond mobile, it's the method of making software, applications, or services that will run on quite one platform or OS.[3]

2.3.1. Hardware And Software Tools Used For CPMAD.

- Apache Cordova: Apache Cordova (formerly PhoneGap) may be a mobile application development framework that was originally created by Nitobi. Adobe Systems purchased Nitobi in 2011, rebranded it as PhoneGap, and later released an open-source version of the software called Apache Cordova. Apache Cordova enables software programmers to make applications for mobile devices using CSS3, HTML5, and JavaScript instead of relying on platform-specific APIs like those in Android, iOS, or Windows Phone.[16]

- Web Development Tools: Web development tools allow developers to work with a selection of web technologies, including HTML, CSS, the DOM, JavaScript, PHP and other components that are handled by the online browser. PhoneGap enables wrapping from CSS, HTML, and JavaScript code depending upon the platform of the device. It extends the features of HTML, CSS, and JavaScript to figure with the device. The resulting applications are hybrid, meaning that they are neither truly native mobile application (because all layout rendering is completed via Web views instead of the platform's native UI

Table 1: Advantages and Disadvantages of ML Algorithms [11]

Techniques	Advantages	Disadvantages
Random Forest	Need 90% numerical data for model training. Missing data is estimated effectively and shows accuracy even when a large portion of the data set is missing. It can balance errors in data sets for imbalanced classes; this leads to unsupervised clustering, data views, and outlier detection for unlabeled data.	Algorithms complexity is high and the training is time-consuming It is a black-box approach for statistical modelers; we have very little control over what the model does. In the case of regression,it may lead to overfitting noisy data sets as it fails to predict beyond the range in the training data set.
SVM	Needs 90% numerical data or model training . SVM works relatively well when there is a clear margin of separation between classes. SVM is more-effective in high dimensional spaces. SVM is more effective, where the number of dimensions is greater than the number of samples. SVM is relatively memory efficient	Not suitable for large data sets. SVM does not perform very well when target classes are overlapping, i.e., when the data set is noisy. In cases where the number of features for each data point exceeds the number of training data samples, the SVM will underperform.
kNN	Simple to implement. Can learn non-linear boundaries, robust to noise in the input data.	Inefficient ad for every prediction, the entire data set is processed Time complexity is $O(dMNlog(k))$ where d =dimension of the data M = size of training data and N = size of test data.
Naive Bayes	Implementation is easy, and execution is fast.	It is mostly used in text classification. It requires that the features must be independent, which is unlikely in real-life applications.

framework) nor purely Web-based (because they are not just Web apps, but are packaged as apps for distribution and have access to native device APIs). [16]

• Operating System: Windows 7 or above OS.
• Hardware Requirements: Minimum 8GB RAM

2.3.2. Apache Plugins

A plugin could also be a little touch bit of an add-on code that provides JavaScript interface to native components. They permit your app to use native device capabilities beyond what's available to pure web apps. A plugin uses a Javascript API for native SDK functionality. Plugins are typically hosted on npm, and you will search for them on the plugin search page. Some key APIs are provided by the Apache Cordova open source project, and these are mentioned as Core Plugin APIs[4].

To add a plugin in HTML, you can use the CLI to launch the search page:

$ Cordova plugin search camera

To add and save the camera plugin to config.xml and package.json, we'll specify the npm package name for the camera plugin.[5]

The CLI adds a plugin code as appropriate for each platform. If you'd wish to develop with lower-level shell tools or platform SDKs, you'd wish to run the Plugman utility to feature plugins separately for every platform.

Use plugin ls (or plugin list, or plugin by itself) to seem at currently installed plugins. Each display by its identifier:

If you'd wish to develop with lower-level shell tools or platform SDKs, you'd wish to run the Plugman utility to feature plugins separately for every platform.[6]

Use plugin ls (or plugin list, or plugin by itself) to seem at currently installed plugins. Each display by its identifier:[7]

If you'd wish to develop with lower-level shell tools or platform SDKs, you'd wish to run the Plugman utility to feature plugins separately for every platform. Use plugin ls (or plugin list, or plugin by itself) to seem at currently installed plugins.[16]

2.3.3. Installing Cordova

The Cordova command-line tool is distributed as an npm package.

To install the Cordova command-line tool, follow these steps:[8]

• Download and install Node.js. On installation, you need to be able to invoke node and npm on your instruction.
• (Optional) Download and install a git client, if you are doing not have already got one. Following installation, you need to be able to invoke git on your instruction. The CLI uses it to download assets once they're referenced employing a URL to a git repo.
• Install the Cordova module using the npm utility of Node.js. The Cordova module will automatically be downloaded by the npm utility.

On Windows the syntax for installing Cordova is:

C:\>npm install -g Cordova

The -g flag above tells npm to put in Cordova globally. Otherwise, it'll be installed within the node_modules subdirectory of the present working directory.

Following installation, you ought to be ready to run Cordova on the instruction with no arguments, and it should print help text.[17]

3. PROPOSED SYSTEM

The proposed system consists of 3 modules: Health Diagnosis app, predictive system, and web application. Patients are required to provide information by filling details in the UI of the App. They can also communicate with clinicians and health practitioners. Mobile Services are used by the app to send and receive data.

Predictive systems perform the following five main tasks: data acquisition, data preprocessing (filtering, feature extraction, selection), data integration, data processing, and forecasting using machine learning supervised algorithms. Cross-platform mobile app and desktop web applications have minor differences; the former is specifically designed for the use of the supervisors or the clinicians, whereas the latter is designed for the use of patients.
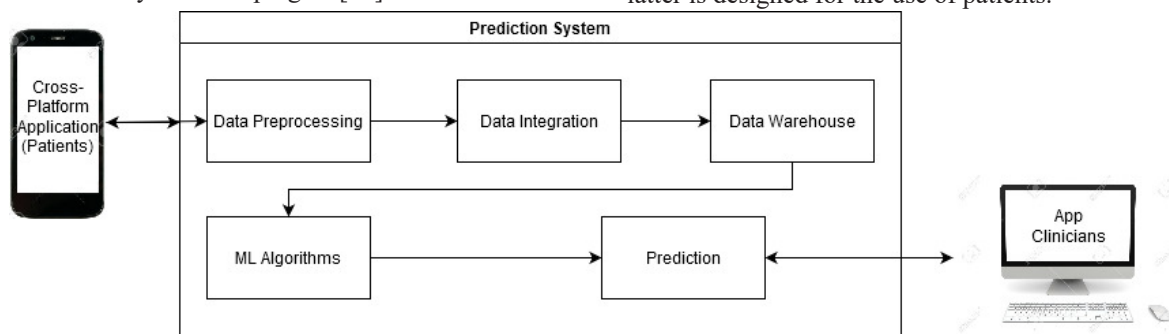

Fig .5: Health diagnosis cross-platform application

The proposed system will help the general public to monitor their daily health parameters like blood pressure, sugar levels etc and will generate an alarm if parameters degrade to the patient as well as to the health practitioner so that timely medical assistance can be procured. The system will also generate weekly and monthly reports which will be shared by the health practitioner.

4. CONCLUSION

Medical diagnosis and prognosis may be a topic closely associated with machine learning. In machine learning, when a dataset is trained with an appropriate learning algorithm, the prediction of a specific feature is often obtained. This sort of skill can help professionals to require precautionary measures or analyze a further problem. However, the physical body is extremely complex, and medical decisions might not be based solely on speculation, the concept of machine learning are often an excellent tool for doctors and practitioners.[4]

The main contribution of this cross-platform health diagnosis app for patients is that it helps to watch health parameters and also alert them if there's a sudden change in them so that timely medical assistance is often procured. The proposed app aims to supply services to patients and physicians through a web patient-assisted web application, with the assistance of this technique, physicians are ready to evaluate each patient's case more closely with the parameters completed by the corresponding patient.[9]

5. ACKNOWLEDGMENT

6. ABBREVIATIONS AND ACRONYMS

CPMAD: Cross-Platform Mobile Application Development
ML: Machine Learning[]

7. REFERENCES

7.1. Journal Article:

[1] Mario L. Bernardi, & Marta Cimitile. Dept. of Engineering-RCOST, University of Sannio, Italy, Unitelma Sapienza University, Italy. Model Driven Development Of Cross-Platform Mobile Applications.

[2] Sebastian Raschka (November 2018). University of Wisconsin–Madison, Department of Statistics. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning.

[3] Dr. V. Karthikeyani, & I. Parvin Begum (2015). Comparison a Performance of Data Mining Algorithms (CPDMA) in Prediction of Diabetes Disease.

[4] Avong Alexander Eric (April 2016). An Investigation about Cross-Platform Mobile Application Development Technologies and Their Application in Developing Health Diagnosis Mobile Application.

[5] Lisandro Delia, Nicolas Galdamez, Pablo Thomas, Leonardo Corbalan, & Patricia Pesado, Institute of Research in Computer Science III-LIDI. School of Computer Science. National University of La Plata. La Plata, Argentina. Multi-Platform Mobile Application Development Analysis.

[6] Anirudh Nagesh, MS Student School of Information Studies, Syracuse University, & Carlos E. Caicedo, Assistant Professor School of Information Studies, Syracuse University. Cross-Platform Mobile Application Development

[7] Apurva P. Pawar, Vandana S. Jagtap, & Mamta S. Bhamare (October 2014). Citation: Survey on Techniques for Cross-Platform Mobile Application Development. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 3(10)

[8] Lisandro Delia, Nicolas Galdamez, Pablo Thomas, Leonardo Corbalan, & Patricia Pesado, Institute of Research in Computer Science III-LIDI. School of Computer Science. Multi-Platform Mobile Application Development Analysis

[9] Mohammad Hajarian. M.Sc. Software Engineering, Staffordshire University. Toward The Development Of Cross-Platform Mobile Applications.

7.2. Book

[10] Peter Harrington. Machine Learning in Action: Manning Publications.

[11] Jiawei Han, Micheline Kamber, & Jian Pei Morgan. Data Mining Concepts and Techniques: Kaufmann Publications

7.3. Website Links

[12] Support Vector Machine – Classification (SVM). Retrieved from: https://www.saedsayad.com/support_vector_machine.htm

[13] Naïve Bayesian. Retrieved from: https://www.saedsayad.com/naive_bayesian.htm

[14] K-Nearest Neighbors Classification. Retrieved from: https://www.saedsayad.com/k_nearest_neighbors.htm

[15] Random forest. Retrieved from: https://en.wikipedia.org/wiki/Random_forest

[16] Apache Cordova. Retrieved from: https://en.wikipedia.org/wiki/Apache_Cordova

[17] Create your first Cordova App. Retrieved from: https://cordova.apache.org/docs/en/latest/guide/cli/