ABSTRACT

As we all know that there is a larger population, patients are in line at the hospital's front for care. Smart health prediction is a framework in which this problem can be resolved by using machine learning algorithm. This is an informed health prediction system. Once the standard of medical information is incomplete, the accuracy of the study will be reduced. Moreover, completely different regions have a distinctive appearance of bound regional diseases, which can lead to a weakening of the prediction of unwellness outbreaks. The patient provides symptoms in this model after that decision tree algorithm is used to determine the possible illness. This framework is implemented by using flask for interface and decision tree algorithm. The analysis accuracy is increased by using Machine Learning algorithm. This framework also displays the specialist name after entering specialty and area. This framework help to patients to identify the disease as well as locate the doctors near-by.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AI	-	Artificial Intelligence
OCR	-	Optical Character Recognition
ML	-	Machine Learning
API	-	Application Programming Interface
WSGI	-	Web Server Gateway Interface
GPL	-	General Public License
GUI	-	Graphical User Interface
KNN	-	K-nearest neighbor
SVM	-	Support Vector Machine
ANN	-	Artificial Neural Network
ID3	-	Iterative Dichotomiser 3
CSV	-	Comma-Separated Values

CHAPTER 1

INTRODUCTION

1.1 DOMAIN INTRODUCTION

Machine learning is a subfield of artificial intelligence (AI). The goal of machine learning generally is to understand the structure of data and fit that data into models that can be understood and utilized by people. Although machine learning is a field within computer science, it differs from traditional computational approaches. In traditional computing, algorithms are sets of explicitly programmed instructions used by computers to calculate or problem solve. Machine learning algorithms instead allow for computers to train on data inputs and use statistical analysis in order to output values that fall within a specific range. Because of this, machine learning facilitates computers in building models from sample data in order to automate decision-making processes based on data inputs.

The analysis accuracy is reduced when the quality of medical data in incomplete. Moreover, different regions exhibit unique characteristics of certain regional diseases, which may weaken the prediction of disease outbreaks. However, those existing work mostly considered structured data. There is no proper methods to handle semi structured and unstructured. The proposed system will consider both structured and unstructured data. The analysis accuracy is increased by using Machine Learning algorithm.

Any technology user today has benefitted from machine learning. Facial recognition technology allows social media platforms to help users tag and share photos of friends. Optical character recognition (OCR) technology converts images of text into movable type. Recommendation engines, powered by machine learning, suggest what movies or television shows to watch next based on user preferences. Self-driving cars that rely on machine learning to navigate may soon be available to consumers. The analysis accuracy is reduced when the quality of medical data in incomplete. Moreover, different regions exhibit unique characteristics of certain regional diseases, which may weaken the prediction of disease outbreaks. However, those existing work mostly considered structured data. There is no proper methods to handle semi structured and unstructured. The proposed system will consider both

structured and unstructured data. The analysis accuracy is increased by using Machine Learning algorithm.

1.2 MACHINE LEARNING METHODS

In machine learning, tasks are generally classified into broad categories. Two of the most widely adopted machine learning methods are supervised learning which trains algorithms based on example input and output data that is labeled by humans, and unsupervised learning which provides the algorithm with no labeled data in order to allow it to find structure within its input data. The following are the detailed explanation of the methods.

1.2.1 Supervised Learning

In supervised learning, the computer is provided with example inputs that are labeled with their desired outputs. The purpose of this method is for the algorithm to be able to "learn" by comparing its actual output with the "taught" outputs to find errors, and modify the model accordingly. Supervised learning therefore uses patterns to predict label values on additional unlabeled data.

1.2.2 Unsupervised Learning

In unsupervised learning, data is unlabeled, so the learning algorithm is left to find commonalities among its input data. As unlabeled data are more abundant than labeled data, machine learning methods that facilitate unsupervised learning are particularly valuable. So, the goal of unsupervised learning may be as straightforward as discovering hidden patterns within a dataset, but it may also have a goal of feature learning, which allows the computational machine to automatically discover the representations that are needed to classify raw data.

1.3 MACHINE LEARNING ARCHITECTURE

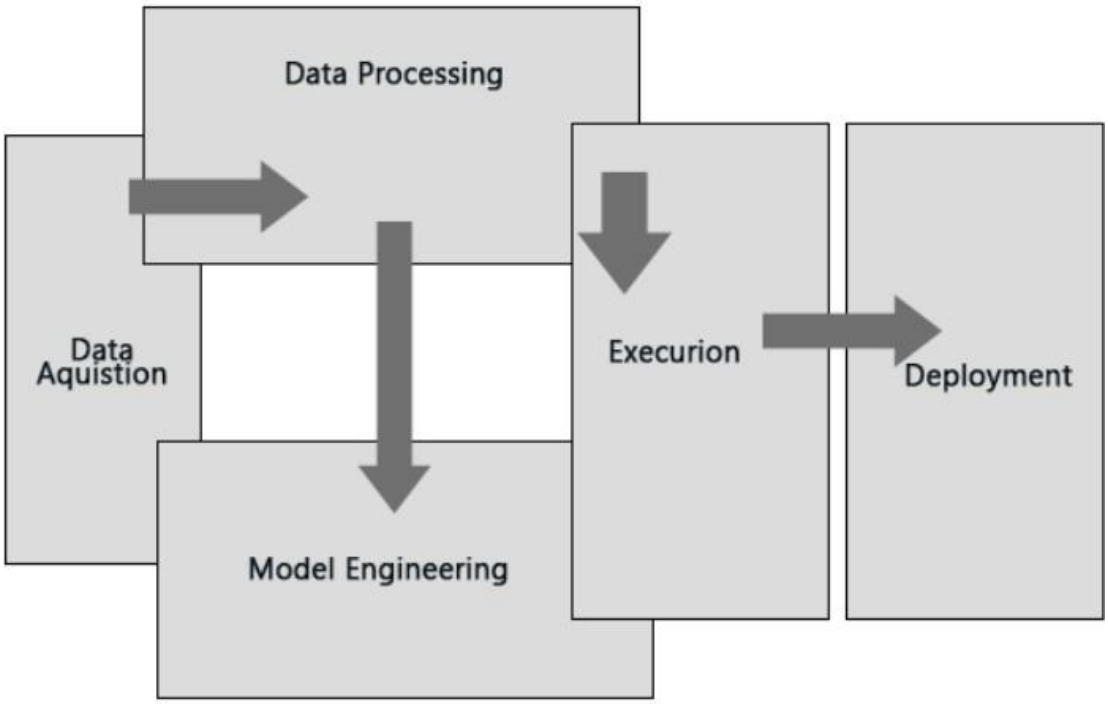

Fig.1.1: Machine learning Architecture

Architecture of Machine Learning involves a series of steps before predicting the accurate outcome for the inputs given. The following is the brief explanation about the steps involved for the prediction process.

1.3.1 Data Acquisition

As machine learning is based on available data for the system to make a decision hence the first step defined in the architecture is data acquisition. This involves data collection, preparing and segregating the case scenarios based on certain features involved with the decision making cycle and forwarding the data to the processing unit for carrying out further categorization. This stage is sometimes called the data preprocessing stage. The data model expects reliable, fast and elastic data which may be discrete or continuous in nature. The data is then passed into stream processing systems (for continuous data) and stored in batch data warehouses (for discrete data) before being passed on to data modeling or processing stages.

1.3.2 Data Processing

The received data in the data acquisition layer is then sent forward to the data processing layer where it is subjected to advanced integration and processing and involves normalization of the data, data cleaning, transformation, and encoding. The data processing is also dependent on the type of learning being used. For e.g., if supervised learning is being used the data shall be needed to be segregated into multiple steps of sample data required for training of the system and the data thus created is called training sample data or simply training data. Also, the data processing is dependent upon the kind of processing required and may involve choices ranging from action upon continuous data which will involve the use of specific function based architecture, for example, lambda architecture, Also it might involve action upon discrete data which may require memory bound processing. The data processing layer defines if the memory processing shall be done to data in transit or in rest.

1.3.3 Data Modeling

This layer of the architecture involves the selection of different algorithms that might adapt the system to address the problem for which the learning is being devised, these algorithms are being evolved or being inherited from a set of libraries. The algorithms are used to model the data accordingly, this makes the system ready for execution step.

1.3.4 Execution

This stage in machine learning is where the experimentation is done, testing is involved and tunings are performed. The general goal behind being to optimize the algorithm in order to extract the required machine outcome and maximize the system performance, The output of the step is a refined solution capable of providing the required data for the machine to make decisions.

1.3.5 Deployment

Like any other software output, ML outputs need to be operationalized or be forwarded for further exploratory processing. The output can be considered as a non-deterministic query which needs to be further deployed into the decision-making system.

1.4 FLASK

Flask is an API of Python that allows to build up web-applications. It was developed by Armin Ronacher. Flask's framework is more explicit than Django's framework and is also easier to learn because it have less base code to implement a simple web-Application. A Web-Application Framework or Web Framework is the collection of modules and libraries that helps the developer to write applications without writing the low-level codes such as protocols, thread management, etc. Flask is based on WSGI (Web Server Gateway Interface) toolkit and Jinja2 template engine.This framework uses flask for web Interface.

Flask is a lightweight web application framework of python and baseband on the WSGI toolkit and jinja2 template engine. Flask takes the versatile python programming language and offers a simple web creation framework. When loaded into python, Flask can be used to save time for web applications that build up. It leaves the essence basic and stretchable. It has no abstraction layer of the database, type validation, or any other elements. Flask allows prolongations. Extensions exist for object-related mappers, type validation, upload management, different techniques for flexible authentication.

Flask is a web application framework written in Python. Armin Ronacher, who leads an international group of Python enthusiasts named Pocco, develops it. Flask is based on Werkzeug WSGI toolkit and Jinja2 template engine. Both are Pocco projects. Http protocol is the foundation of data communication in world wide web. Different methods of data retrieval from specified URL are defined in this protocol.

The following table summarizes different http methods −