**Chapter 5**

**coding and implentation**

**Flask** is a micro [web](https://en.wikipedia.org/wiki/Web_framework) framework written in [Python](https://en.wikipedia.org/wiki/Python_(programming_language)). It is classified as a microframework because it does not require particular tools or libraries.It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions. However, Flask supports extensions that can add application features as if they were implemented in Flask itself. Extensions exist for object-relational mappers, form validation, upload handling, various open authentication technologies and several common framework related tools. Extensions are updated far more regularly than the core Flask program.we have used RESTful request dispatching for accessing classification model. We have stored model using pickles and flask supports to use model as an api.

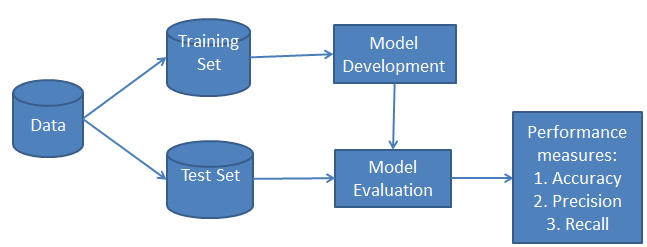
**Scikit-learn** is a [free software](https://en.wikipedia.org/wiki/Free_software) [machine learning](https://en.wikipedia.org/wiki/Machine_learning) [library](https://en.wikipedia.org/wiki/Library_(computing)) for the [Python](https://en.wikipedia.org/wiki/Python_(programming_language)) programming language. It features various [classification](https://en.wikipedia.org/wiki/Statistical_classification), [regression](https://en.wikipedia.org/wiki/Regression_analysis) and [clustering](https://en.wikipedia.org/wiki/Cluster_analysis) algorithms including [support vector machines](https://en.wikipedia.org/wiki/Support_vector_machine),[random forests](https://en.wikipedia.org/wiki/Random_forests), [gradient boosting](https://en.wikipedia.org/wiki/Gradient_boosting), [*k*-means](https://en.wikipedia.org/wiki/K-means_clustering) and [DBSCAN](https://en.wikipedia.org/wiki/DBSCAN), and is designed to interoperate with the Python numerical and scientific libraries [NumPy](https://en.wikipedia.org/wiki/NumPy) and [SciPy](https://en.wikipedia.org/wiki/SciPy). We are using Naive Bayes Algorithm it is a statistical classification technique based on Bayes Theorem. It is one of the simplest supervised learning algorithms. Naive Bayes classifier is the fast, accurate and reliable algorithm. Naive Bayes classifiers have high accuracy and speed on large datasets.Naive Bayes classifier assumes that the effect of a particular feature in a class is independent of other features.

**Algorithm**

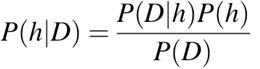
**Naive bayes algorithm**

## **Classification Work flow**

Whenever we perform classification, the first step is to understand the problem and identify potential features and label. Features are those characteristics or attributes which affect the results of the label. For example, in our system in case of assigning triage level, we have to specify patient-data Age**,** Gender**,** Pulse**,** B/P**,** Temperature**,**Alert**,** Voice Responsive**,** Pain Responsive**,**Unconscious**,** Airway Breathing**,** Oxysat **and** Triage\_level. These characteristics are known as features which help the model classify Triage levels.

The classification has two phases, a learning phase, and the evaluation phase. In the learning phase, classifier trains its model on a given dataset and in the evaluation phase, it tests the classifier performance. Performance is evaluated onthe basis of various parameters such as accuracy, error, precision, and recall.

This assumption simplifies computation, and that's why it is considered as naive. This assumption is called class conditional independence.



* P(h): the probability of hypothesis h being true (regardless of the data). This is known as the prior probability of h.
* P(D): the probability of the data (regardless of the hypothesis). This is known as the prior probability.
* P(h|D): the probability of hypothesis h given the data D. This is known as posterior probability.
* P(D|h): the probability of data d given that the hypothesis h was true. This is known as posterior probability.

**Naive Bayes classifier working:**

* Step 1: Calculate the prior probability for given class labels
* Step 2: Find Likelihood probability with each attribute for each class
* Step 3: Put these value in Bayes Formula and calculate posterior probability.
* Step 4: See which class has a higher probability, given the input belongs to the higher probability class.

**Building** **Naive Bayes classifier Model:**

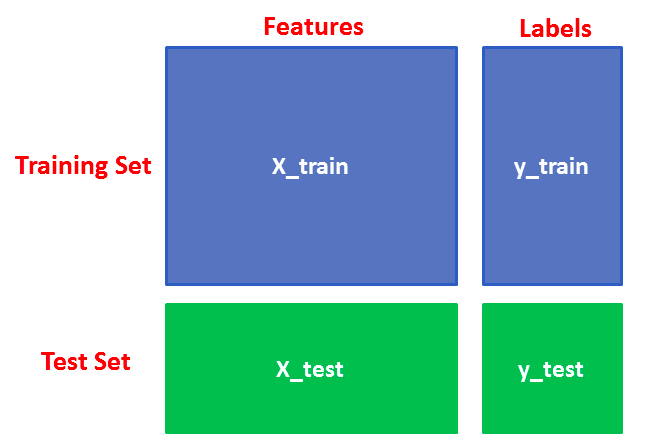
Step 1: Load dataset from csv file and convert it into dataframe using pandas.

* Step 2: we need to convert these string labels into numbers .This is known as label encoding. Scikit-learn provides LabelEncoder library for encoding labels with a value between 0 and one less than the number of discrete classes.
* Step 3: Clean and standardise the data.
* Step 4: split data into training and testing data.

Step 5: Create naive bayes classifier and Fit the dataset on classifier

to Perform prediction.

* Step 6: store model for predicting triage using pickles.



**Server using flask and hospital searching**

* Step 1: Input will be received from client side in JSON format.
* Step 2: Load stored model using pickles at server side.
* Step 3: Return calculated triage to client side in JSON format.
* Step 4: Foursquare places api is used for finding nearby hospitals according to our provided parameters and endpoints.
* Step 5: List of nearby hospitals and triage levels will be displayed at client side.

**TOOLS**

**PyCharm**

is an integrated development environment (IDE) used in [computer programming](https://en.wikipedia.org/wiki/Computer_programming), specifically for the [Python](https://en.wikipedia.org/wiki/Python_(programming_language))language. It is developed by the Czech company [JetBrains](https://en.wikipedia.org/wiki/JetBrains).It provides code analysis, a graphical debugger, an integrated unit tester, integration with version cotrol system (VCSes), and supports web development with Django as well as Data science with [Anaconda](https://en.wikipedia.org/wiki/Anaconda_(Python_distribution)).PyCharm is [cross-platform](https://en.wikipedia.org/wiki/Cross-platform), with Windows ,[ma](https://en.wikipedia.org/wiki/MacOS)cOS and [Linu](https://en.wikipedia.org/wiki/Linux)x ersions. The Community Edition is released under the [Apache License](https://en.wikipedia.org/wiki/Apache_License),and there is also Professional Edition with extra features – released under a [proprietary license](https://en.wikipedia.org/wiki/Proprietary_software).

It provides coding assistance and [analysis](https://en.wikipedia.org/wiki/Code_analysis), with [code completion](https://en.wikipedia.org/wiki/Autocomplete), syntax and error highlighting, [linter integration](https://en.wikipedia.org/wiki/Lint_(software)), and quick fixes. It has Project and code navigation which provides specialized project views, file structure views and quick jumping between files, classes, methods and usages.It supports python [refactorin](https://en.wikipedia.org/wiki/Refactoring)g including rename, extract method, introduce variable, introduce constant, pull up, push down and others.It has integrated [unit testing](https://en.wikipedia.org/wiki/Unit_testing), with line-by-line [code coverag](https://en.wikipedia.org/wiki/Code_coverage)e and [Google App Engin](https://en.wikipedia.org/wiki/Google_App_Engine)e Python development.PyCharm has a Version control integration: unified user interface for [Mercurial](https://en.wikipedia.org/wiki/Mercurial), [Git](https://en.wikipedia.org/wiki/Git_(software)), [Subversion](https://en.wikipedia.org/wiki/Apache_Subversion), [Perforce](https://en.wikipedia.org/wiki/Perforce)e and [CV](https://en.wikipedia.org/wiki/Concurrent_Versions_System)S with change lists and merge

**Postman**

API endpoint testing is one of the most important things we do as web developers. If the routes to reach our data are incorrect, the required parameters are not included, the authorization is missing, or a host of other things arent correctly hooked up, your app doesnt work. If your app doesnt work, your customers cant use it. If your customers cant use it, its going to be a really bad day for you, your team, your boss, and so on and so forth.

Postman is an API(application programming interface) development tool which helps to build, test and modify APIs. Almost any functionality that could be needed by any developer is encapsulated in this tool. It is used by over 5 million developers every month to make their API development easy and simple. It has the ability to make various types of HTTP requests(GET, POST, PUT, PATCH), saving environments for later use, converting the API to code for various languages(like JavaScript, Python).We have used postman to test our model.We are sending input in JSON format through pstman to test whether it is working properly and we are getting output back correctly.

**Technologies**

**Flask** is a micro [web](https://en.wikipedia.org/wiki/Web_framework) framework written in [Python](https://en.wikipedia.org/wiki/Python_(programming_language)). It is classified as a microframework because it does not require particular tools or libraries.It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions.

Flask supports extensions that can add application features as if they were implemented in Flask itself. Extensions exist for object-relational mappers, form validation, upload handling, various open authentication technologies and several common framework related tools. Extensions are updated far more regularly than the core Flask program.we have used RESTful request dispatching for accessing classification model. We have stored model using pickles and flask supports to use model as an api.

**Scikit-learn** is a [free software](https://en.wikipedia.org/wiki/Free_software) [machine learning](https://en.wikipedia.org/wiki/Machine_learning) [library](https://en.wikipedia.org/wiki/Library_(computing)) for the [Python](https://en.wikipedia.org/wiki/Python_(programming_language)) programming language. It features various [classification](https://en.wikipedia.org/wiki/Statistical_classification), [regression](https://en.wikipedia.org/wiki/Regression_analysis) and [clustering](https://en.wikipedia.org/wiki/Cluster_analysis) algorithms including [support vector machines](https://en.wikipedia.org/wiki/Support_vector_machine),[random forests](https://en.wikipedia.org/wiki/Random_forests), [gradient boosting](https://en.wikipedia.org/wiki/Gradient_boosting), [*k*-means](https://en.wikipedia.org/wiki/K-means_clustering) and [DBSCAN](https://en.wikipedia.org/wiki/DBSCAN), and is designed to interoperate with the Python numerical and scientific libraries [NumPy](https://en.wikipedia.org/wiki/NumPy) and [SciPy](https://en.wikipedia.org/wiki/SciPy).

We are using Naive Bayes Algorithm it is a statistical classification technique based on Bayes Theorem. It is one of the simplest supervised learning algorithms. Naive Bayes classifier is the fast, accurate and reliable algorithm. Naive Bayes classifiers have high accuracy and speed on large datasets.Naive Bayes classifier assumes that the effect of a particular feature in a class is independent of other features.

**Dependancies**

**PIP**

**pip** is a package-management system used to install and manage software packages written in **Python**. Many packages can be found in the default source for packages and their dependencies —**Python** Package Index (PyPI). **Pip** is a recursive acronym for "**Pip**Installs Packages".Most importantly **pip** has a feature to manage full lists of packages and corresponding version numbers, possible through a "requirements" file.This permits the efficient re-creation of an entire group of packages in a separate environment (e.g. another computer) or [virtual environment](https://en.wikipedia.org/wiki/Virtualization).