**ADVERTISEMENT POPULARITY PREDICTION**

**1.1 INTRODUCTION**

Python is a widely used [interpreted](https://en.wikipedia.org/wiki/Interpreted_language), general-purpose, high level programming language. It was initially designed by Guido van Rossum in 1991 and developed by Python Software Foundation. It was mainly developed for emphasis on code readability, and its syntax allows programmers to express concepts in fewer lines of code. Python is a programming language that lets you work quickly and integrate systems more efficiently. Python is a general purpose programming language. Hence, you can use the programming language for developing both desktop and web applications. Also, you can use Python for developing complex scientific and numeric applications.

Python is designed with features to facilitate data analysis and visualization. Python's design philosophy emphasizes [code readability](https://en.wikipedia.org/wiki/Code_readability) with its notable use of [significant whitespace](https://en.wikipedia.org/wiki/Off-side_rule). Its language constructs and [object-oriented](https://en.wikipedia.org/wiki/Object-oriented_programming) approach aims to help programmers write clear, logical code for small and large-scale projects.Python is [dynamically typed](https://en.wikipedia.org/wiki/Dynamic_programming_language) and [garbage-collected](https://en.wikipedia.org/wiki/Garbage_collection_(computer_science)). It supports multiple [programming paradigms](https://en.wikipedia.org/wiki/Programming_paradigm), including [procedural](https://en.wikipedia.org/wiki/Procedural_programming), object-oriented, and [functional programming](https://en.wikipedia.org/wiki/Functional_programming). Python is often described as batteries included language due to its comprehensive [standard library](https://en.wikipedia.org/wiki/Standard_library).

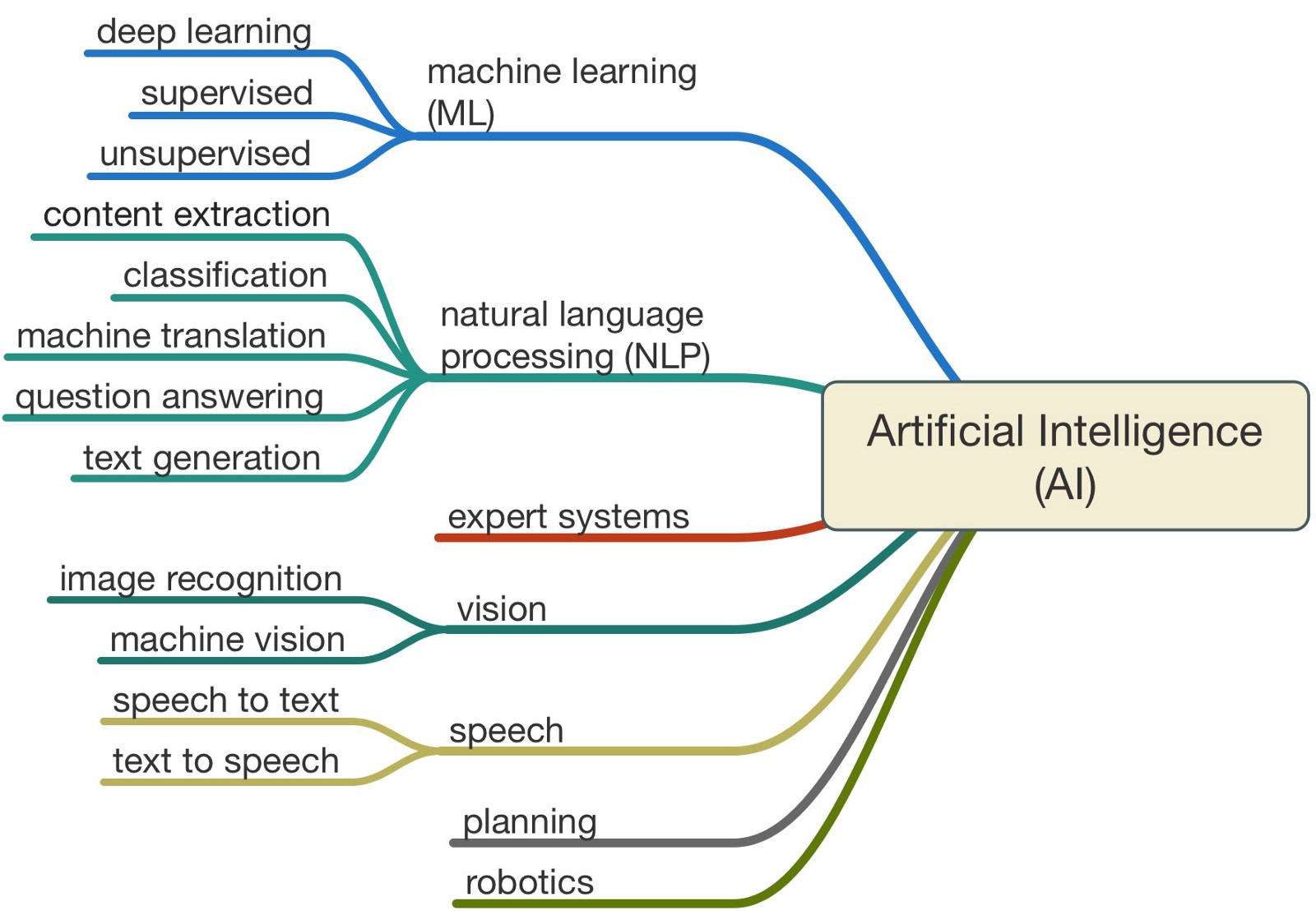
In [computer science](https://en.wikipedia.org/wiki/Computer_science), artificial intelligence (AI), sometimes called machine intelligence, is [intelligence](https://en.wikipedia.org/wiki/Intelligence) demonstrated by [machines](https://en.wikipedia.org/wiki/Machine), in contrast to the natural intelligence displayed by humans and animals. According to the father of Artificial Intelligence, John McCarthy, it is “The science and engineering of making intelligent machines, especially intelligent computer programs”.

Artificial Intelligence is a way of making a computer, a computer-controlled robot, or a software think intelligently, in the similar manner the intelligent humans think. AI is accomplished by studying how human brain thinks and how humans learn, decide, and work while trying to solve a problem, and then using the outcomes of this study as a basis of developing intelligent software and systems. While exploiting the power of the computer systems, the curiosity of human, lead him to wonder, “Can a machine think and behave like humans do?”.Thus, the development of AI started with the intention of creating similar intelligence in machines that we find and regard high in humans.



**Fig 1.1.1 Artificial Intelligence**

Artificial Intelligence can be broadly classified into:



**Fig 1.1.2 Classification of Machine Learning**

**1.2 OBJECTIVES OF RESEARCH**

Advertising is any paid form of non-personal presentation & promotion of ideas, goods, or services by an identified sponsor. Generally Advertisement can be defined in such a way that It is a paid communication message intended to inform people about something or to influence them to buy or try something.

The precise and timely prediction of product popularity is of great value for content providers, advertisers, and broadcast TV operators. This information can be beneficial for operators in TV program purchasing decisions and can help advertisers formulate reasonable advertisement investment plans. Moreover, in terms of technical matters, a precise program popularity prediction method can optimize the whole broadcasting system, such as the content delivery network strategy and cache strategy. Advertising has a great importance when launching a new product, service or even an idea to the market. When an advertisement of a product or service is done appropriately through the right media channels and at the right time, it draws customers’ attention and directs them to take an action. Therefore, it contributes to marketing success in the whole picture. Advertising is a must to arouse an interest among the audience about the products and services. Without arousing that interest, creating demand in the market is nearly impossible. Eventually, growing demand results in a remarkable increase in sales. In this context, it is of great importance to predict the popularity of Advertisements.

**1.3 PROBLEM STATEMENT**

Problem statement is Advertising Popularity Prediction. Here we are predicting the popularity of the advertisement by considering the properties the sales before advertisement and sales after advertisement. If Sales after Advertisement is greater than sales before advertisement then we can say that advertisement is popular.

**2. REVIEW OF LITERATURE**

On the Internet, the popularity of Advertisement can have different connotations. If by content we refer to the subject of the content, such as a person or an organization, then popularity could be expressed by a greater web presence or activity. Social media platforms have democratized the process of Advertising creation allowing mere consumers to become creators and distributors of content. But this has also contributed to an explosive growth of information and has intensified the online competition for user’s attention, since only a small number of items become popular while the rest remain unknown. Understanding what makes one item more popular than another, observing its popularity dynamics, and being able to predict its popularity has thus attracted a lot of interest in the past few years.

Predicting the popularity of Advertisement is useful in many areas such as online marketing, or real-world outcome prediction. In this survey, we review the current findings on Advertisement popularity. We describe the different features that have shown good predictive capabilities, and reveal factors known to influence Advertisement popularity. Predicting the popularity of Advertisement has become an active area of research and, while still in an incipient phase, a large number of prediction methods for different types of Advertisement have been proposed in the latest years. In this project we review the current state of research in this field, identify trends, and suggest domains that can benefit from these studies.

Advertising effectiveness can be defined as the extent to which advertising generates a certain desired effect. Measuring the effects of advertising is very important, given the amount of investments needed for advertising. While it is not possible to obtain a global measure of the advertising effectiveness, we should seek to develop and apply methods and measures for a partial verification of results. Regarding the difficulty of measuring the overall effectiveness, we believe that it is due essentially to the following considerations.

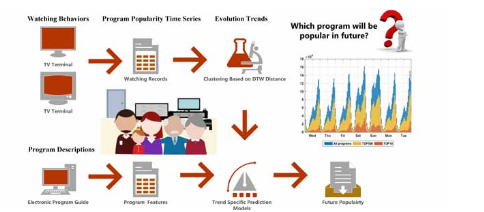
**3. DATA COLLECTION**

In order to predict the popularity of an Advertisement ,considering the properties like Gender, Age, SalesBeforeAdvertisement,Radio,TV,Newspaper,SalesAfterAdvertisement based on these parameters predicting the output label as Popularity Prediction.

* **Gender:** This field is considered the advertisement of a product is related to male or female, based on it sales also get affected.
* **Age:** This field takes the value of age based on the group of people that product belongs to.
* **SalesBeforeAdvertisement:** Sales before advertisement is considered based on the previous year.
* **Types of Advertisement:**

There are three main types of news media:

* Print media.
* Broadcast media.
* Internet.



**Fig 3.1 Flow of Advertisement Prediction**

We considered the fields Radio, TV, Newspaper i.e., based on the percentage of Advertisement given in that media, values are considered.

* **SalesAfterAdvertisement:** After advertisement the sales percentage is considered, based on that prediction of an advertisement is done.



**Fig 3.2 Abstract for Advertisement**

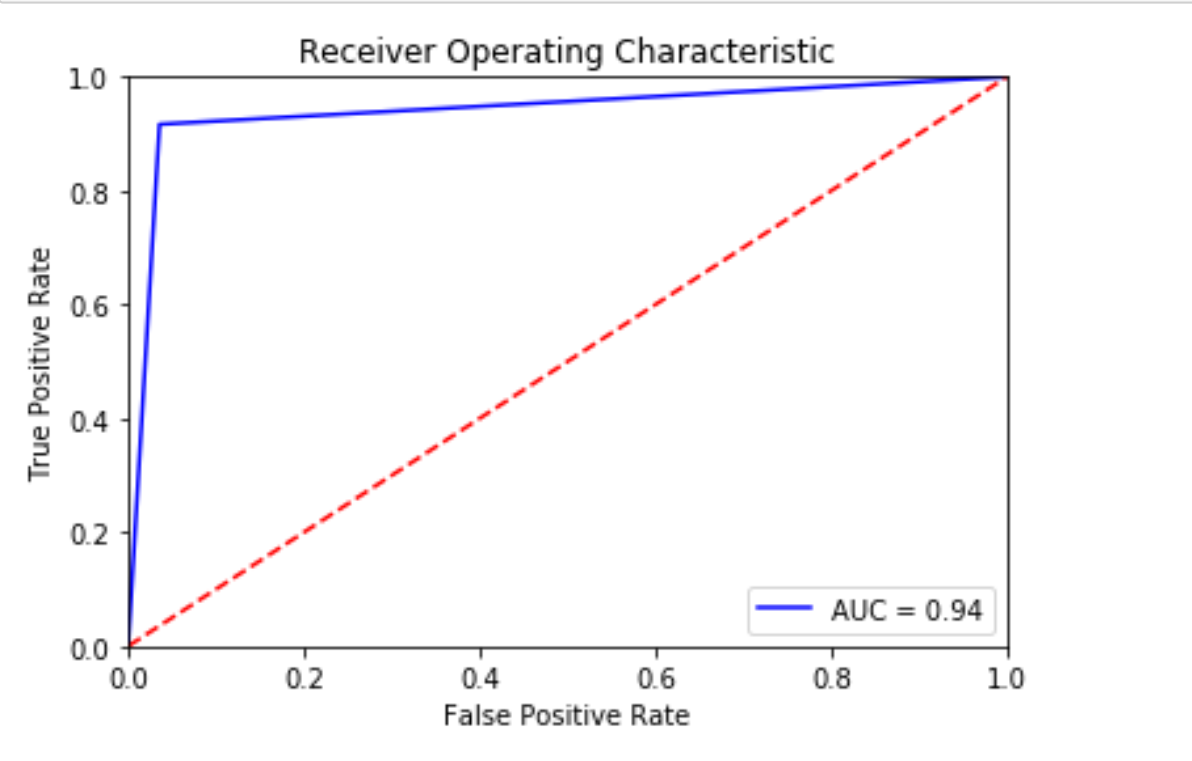
While designing the model, for predicting the popularity of an advertisement by considering the two parameters SalesBeforeAdvertisement and SalesBeforeAdverisement. Based on the condition if salesafteradvertisement is greater than salesbeforeadvertisement then advertisement is effective. Then,we can say that advertisement is popular.

**4.METHODOLOGY**

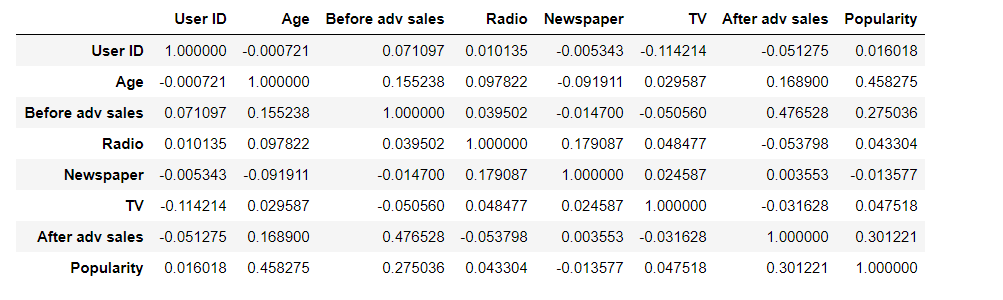
**4.1 EXPLORATORY DATA ANALYSIS**

4.1.1 FIGURES AND TABLES

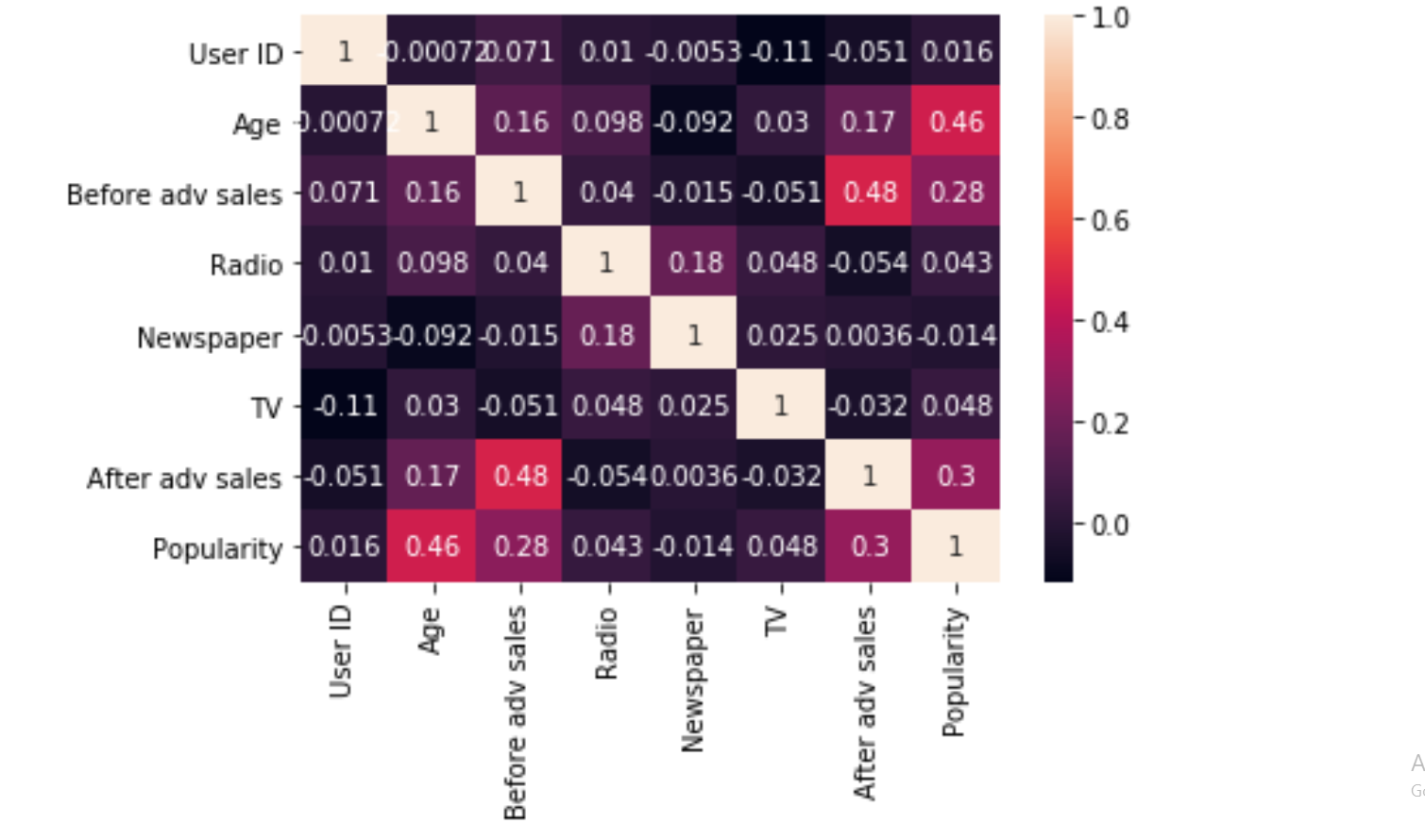
* This figure represents the ROC curve of our model.



**Fig 4.1.1 Receiver Operating Characteristic Curve**

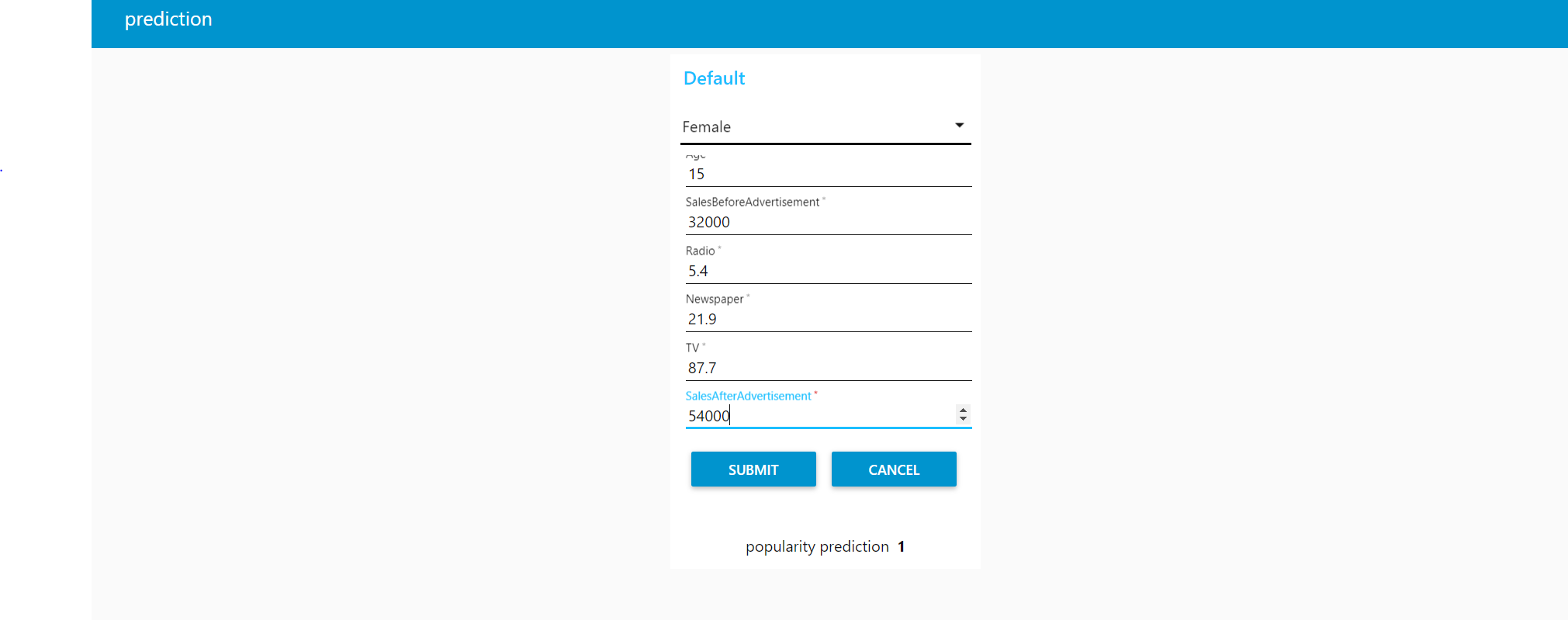
* This table represents the correlation of the dataset used in our model.

**Table 4.1.1 Correlation table**

* This figure represents the heat map of our model. 

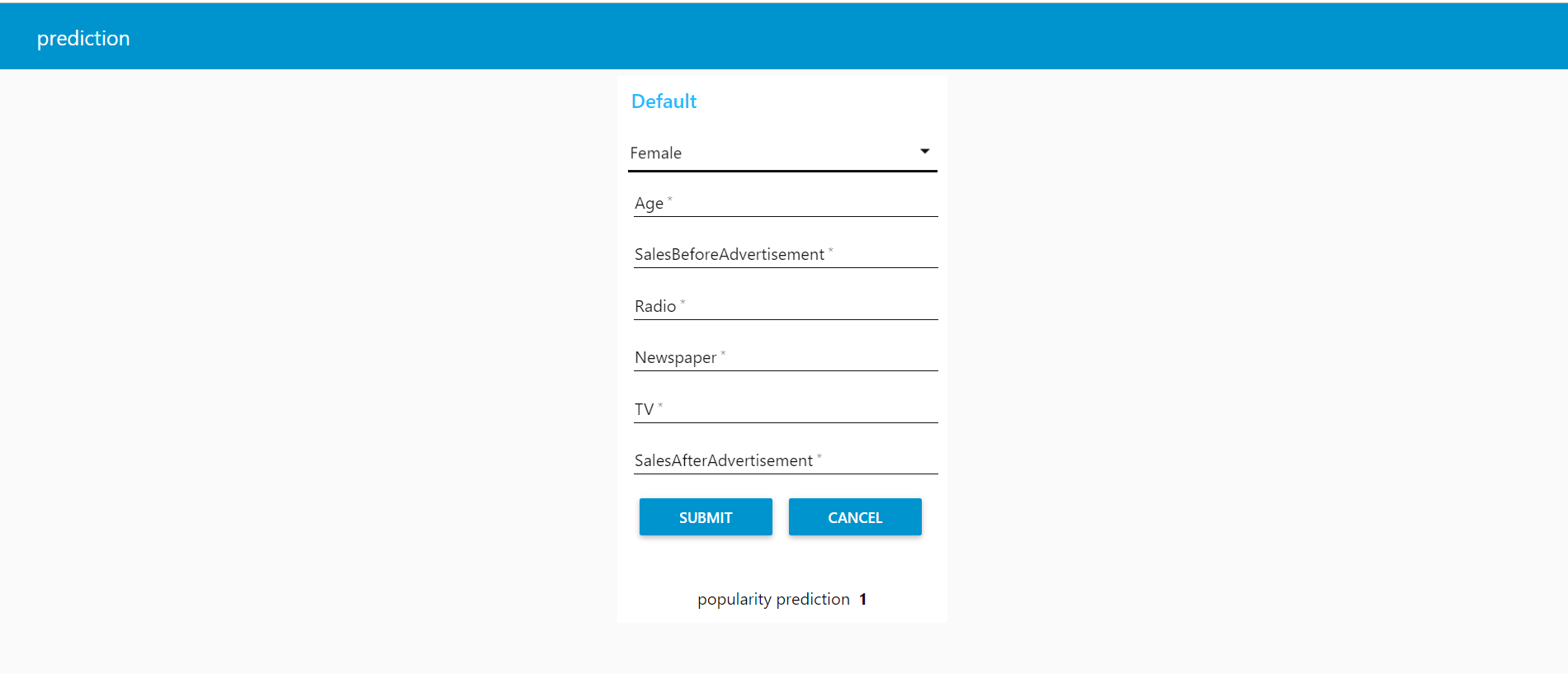
**Fig 4.1.2 Heat Map**

* It represents the form to give input to the flow.



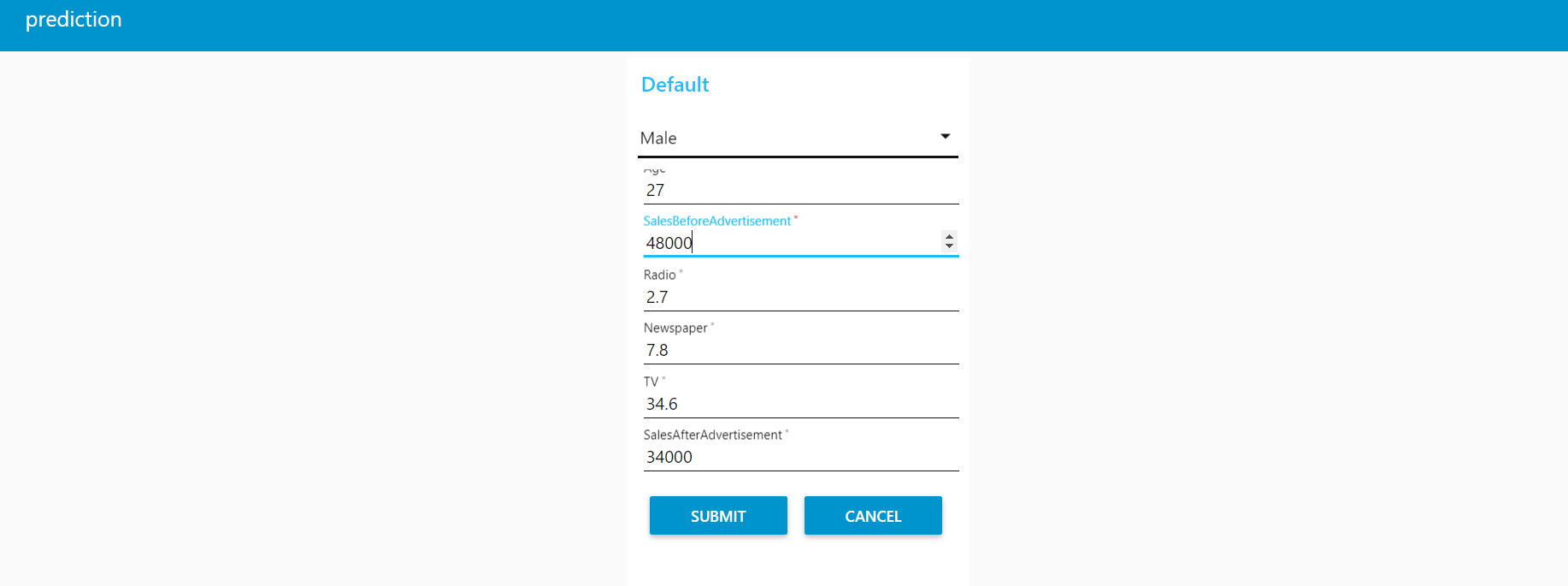
**Fig 4.1.3 Input Form**

* It is the output displayed when above details is submitted.

****

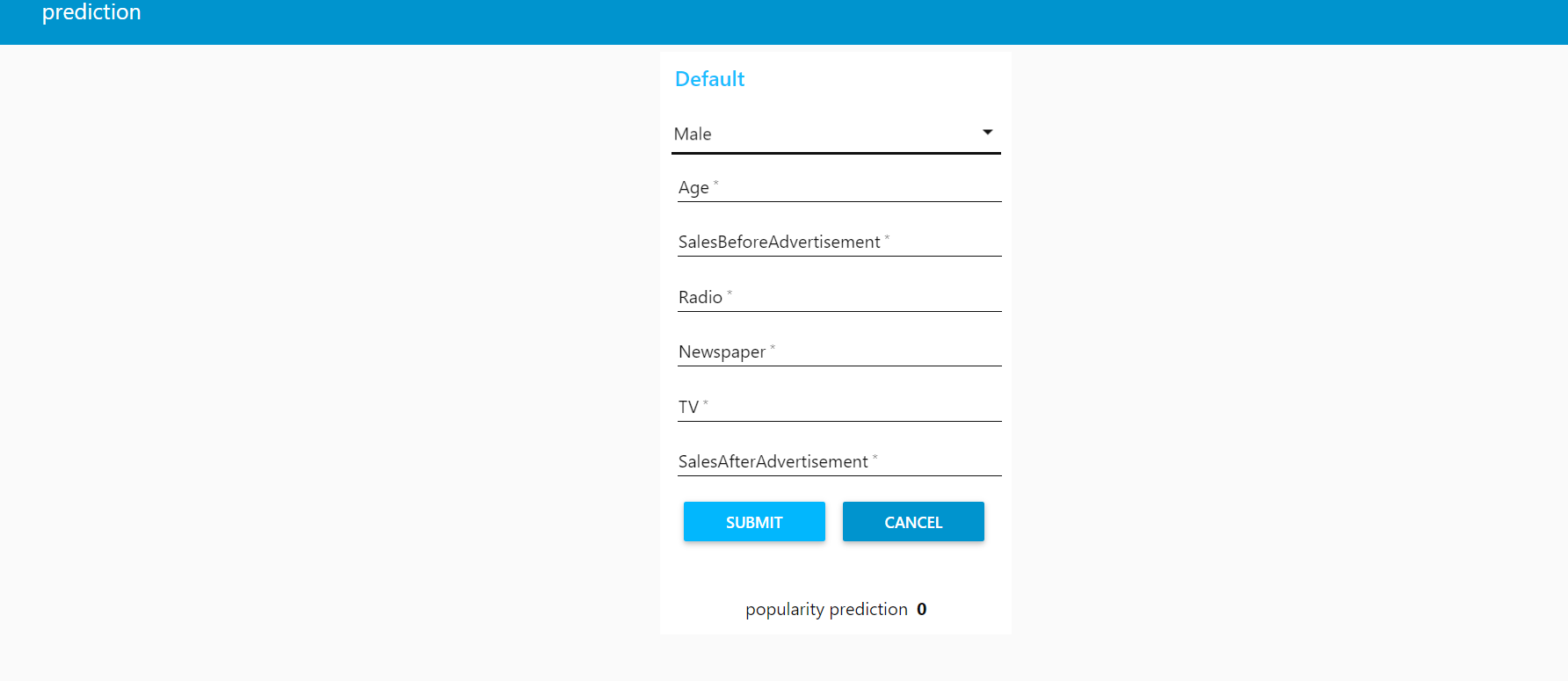
**Fig 4.1.4 Output Value of Prediction**

* It represents the form to give input to the flow



**Fig 4.1.5 Input Form**

* It is the output displayed when above details is submitted.



**Fig 4.1.6 Output Value of Prediction**

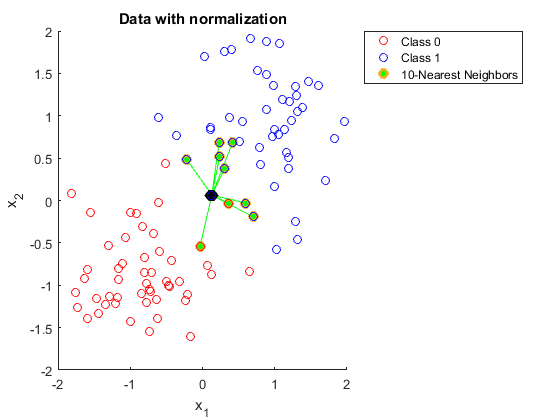
**4.2 DATA MODELLING**

For this prediction, we are using Classification algorithm called “KNN”.The k-nearest neighbours (KNN) algorithm is a simple, easy-to-implement supervised machine learning algorithm that can be used to solve both classification and regression problems. A supervised machine learning algorithm (as opposed to an unsupervised machine learning algorithm) is one that relies on labeled input data to learn a function that produces an appropriate output when given new unlabeled data.

Imagine a computer is a child, we are its supervisor (e.g. parent, guardian) and we want the child (computer) to learn what a dog looks like. We will show the child several different pictures, some of which are pigs and the rest could be pictures of anything (cats, rats, etc).When we see a dog, we shout dog!” When it’s not a dog, we shout “no, not dog!” After doing this several times with the child, we show them a picture and ask “dog?” and they will correctly (most of the time) say “dog!” or “no, not dog!” depending on what the picture is. That is supervised machine learning.

A classification problem has a discrete value as its output. For example, “likes pineapple on pizza” and “does not like pineapple on pizza” are discrete. There is no middle ground. The analogy above of teaching a child to identify a dog is another example of a classification problem.

The KNN algorithm assumes that similar things exist in close proximity. In other words, similar things are near to each other. There are other ways of calculating distance, and one way might be preferable depending on the problem we are solving. However, the straight-line distance (also called the Euclidean distance) is a popular and familiar choice.



**Fig 4.2.1 Data with normalization**

****

**Fig 4.2.2 Distance Functions**

**Advantages**

1. The algorithm is simple and easy to implement.
2. There’s no need to build a model, tune several parameters, or make additional assumptions.
3. The algorithm is versatile. It can be used for classification, regression, and search (as we will see in the next section).

**Disadvantages**

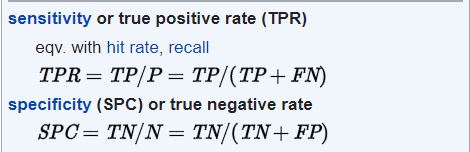
1. The algorithm gets significantly slower as the number of examples and/or predictors/independent variables increase.

**Finding the accuracy of the model:**

For finding the accuracy of a classification model, ROC curve is to be plotted. For plotting ROC curve fpr and tpr values must be calculated using confusion matrix.



**Fig 4.2.2 Classification of Model**



**Fig 4.2.3 Formulas**

**5. REFERENCES**

**1. https://ieeexplore.ieee.org/document/8086153**

**2. https://medium.com/coinmonks/predicting-product-sales-through-ads-delivered-on-social-networking-sites-using-k-n-n-in-python-b9f955d184af**

**3. https://machinelearningmastery.com/k-nearest-neighbors-for-machine-learning/**

**6. CONCLUSION**

With the rapid growth in hyper-connected consumers, AI is assuming an increasingly prominent role in marketing and advertising efforts as a means to provide more authentic and personalized engagement opportunities. In such a way that we can use machine learning algorithm for predicting the popularity of an advertisement that can be published through any kind of media by using this model which helps in finding out whether it reaches the right people in the right context or not.