sustainability

Article

Route Optimization of Mobile Medical Unit with Reinforcement Learning

Shruti Maheshwari [1,*], Pramod Kumar Jain [2] and Ketan Kotecha [1]

[1] Symbiosis Institute of Technology, Symbiosis International (Deemed) University, Pune 412115, Maharashtra, India
[2] Indian Institute of Technology, Banaras Hindu University Campus, Varanasi 221005, Uttar Pradesh, India
* Correspondence: shruti.maheshwari@sitpune.edu.in

Abstract: In this paper, we propose a solution for optimizing the routes of Mobile Medical Units (MMUs) in the domain of vehicle routing and scheduling. The generic objective is to optimize the distance traveled by the MMUs as well as optimizing the associated cost. These MMUs are located at a central depot. The idea is to provide improved healthcare to the rural people of India. The solution is obtained in two stages: preparing a mathematical model with the most suitable parameters, and then in the second phase, implementing an algorithm to obtain an optimized solution. The solution is focused on multiple parameters, including the number of vans, number of specialists, total distance, total travel time, and others. The solution is further supported by Reinforcement Learning, explaining the best possible optimized route and total distance traveled.

Keywords: vehicle routing problem; healthcare; mobile medical unit; reinforcement learning; heuristics

check for updates

Citation: Maheshwari, S.; Jain, P.K.; Kotecha, K. Route Optimization of Mobile Medical Unit with Reinforcement Learning. *Sustainability* **2023**, *15*, 3937. https://doi.org/10.3390/su15053937

Academic Editors: Juan M. Corchado, Alfonso González-Briones and Amit Kumar Mishra

Received: 18 January 2023
Revised: 9 February 2023
Accepted: 20 February 2023
Published: 21 February 2023

1. Introduction

In supply chain management, the vehicle routing problem (VRP) is a well-known issue, where customers with established demands are satisfied by one or more depots. The bed capacity in private clinics and hospitals has dropped during the last 20 years. The increasingly elderly population has increased the rates of chronic degenerative diseases, which has doubled the number of patients in rural India. Patients receiving care or therapy for severe chronic conditions desire treatment that keeps them as close to their families as possible for comfort reasons. The same was witnessed in the country during COVID-19; in these scenarios, mobile medical units have proved a boon to society, delivering better and improved healthcare to the rural population of India [1].

Many advanced and emerging countries have investigated the use of mobile medical units (MMUs) to remove obstacles to address the widening gaps between patients and services. Due to MMUs' adaptability, it is possible to deliver primary care locally and according to demand in rural areas. A MMU is placed in the center of a major city (where they do not serve) and departs every day to deliver medical care in nearby rural villages [2]. The use of MMUs is not a recent development in the world. Many states around the globe had previously used MMUs. The ongoing problem of unreachable populations, though, necessitates novel solutions and an analysis of previous tactics. The task is difficult, and every situation calls for a different reaction. The problem can be overcome, though, with careful planning, the use of the right human resources, and the creation of service packages tailored to specific needs. Mobile Medical Units have given mixed results in previous years. These operational guidelines, which are based on field experience, give a general framework and recommendations for reforming the current MMU implementation design to deliver the best service possible while making efficient use of available resources. The MMU's main goal is to offer services up to a range of preventative, motivating, and better services and to facilitate referrals [3].

It is very necessary to decide the routes of these mobile medical units so that they cover maximum areas serving the maximum number of patients. To choose the optimum routes of MMUs considering multiple constraints and parameters, the vehicle routing problem (VRP) is a very promising means to synchronize them. VRP is generally focused on finding the optimal route for a vehicle, with the constraints such as window, route and length. The main objectives are to maximize the efficiency of the vehicle, minimize the last-mile delivery cost and save time. It also requires massive computational efforts, and must handle the unpredictable factors, such as customer demand, roadblocks due to maintenance, and so on [4].

Researchers have recently discussed hospital care organizations in the scientific literature. Problems dealing with MMUs scheduling and routing are currently being worked on. Among recent scenarios, many MMUs throughout the globe are permitted to offer home healthcare (HHC) services, as seen during COVID-19. Though many MMUs are growing, the summary of the literature review reveals that no studies have been performed so far to describe the logistical requirements and obstacles faced by MMUs. Such evidence demonstrates the necessity of characterizing and defining logistics management as to services provided by MMUs to reduce the gap between the state of the art and the actual delivery of healthcare services.

To optimize the routes of these mobile medical units, artificial intelligence (AI) can help in changing the game. AI is constantly seeking data, acclimatizing according to it, and evaluating new approaches to ensure that vehicles take the most efficient route in real time. AI route planning algorithms foresee impediments such as weather conditions and consider them when rerouting after learning about traffic conditions. They enable planners to take proactive measures to prevent clogging up particular channels. By anticipating resource allocation constraints, AI route planning systems improve the transportation ecosystem over time [1].

Artificial Intelligence in Healthcare Supply Chain

Computing innovations using artificial intelligence (AI) mimic natural intelligence's supporting systems, such as cognition, deep learning, engagement, adaptation, and sensory perception. A variety of sectors, including health and medicine, can use these methods, as they are interdisciplinary. The supply chain uses artificial intelligence (AI) capabilities to transport commodities between points more efficiently and safely [5]. More specifically, AI is changing the supply chain by giving end to end visibility with almost real-time data, reducing human work and effort, assisting in pattern detection at scale, finding analytical insights at scale, and using machine learning to make the most efficient decisions [6].

Machine learning (ML) is a subset of AI. ML aids the machine's ability to learn from data and develop based on prior knowledge or work accomplished on it. In ML, the systems are already taught to look for patterns and relationships across a variety of datasets to select the optimal option. As it gets access to more and more data, it becomes increasingly accurate [7]. In this paper, supervised machine learning is used to achieve the objective of providing vehicle routing and scheduling for MMUs. This output is tested on the data provided by Symbiosis University Hospital and Research Centre (SUHRC), Pune, and is suggested for their implementation.

The rest of the paper is described as a detailed literature review as to work completed so far in the area of healthcare supply chains with artificial intelligence. The next section describes the mathematical model for calculating the distance between two places and the routes to be followed by MMU. In Section 4, an approach to a solution is discussed, and in the next section, the results of the experiment conducted are discussed, together with its accuracy. The final section concludes the paper, suggesting the future scope of the technology.

2. Literature Review

MMUs, acting as a link between hospitals and people, can target both clinical and social factors of health, and they are playing an important role in improving healthcare in developing countries. Clinical efficacy must be demonstrated by further research, both in terms of service quality and expenditures. To maximize the effectiveness that MMUs can provide to various consumer groups and patient care as a whole, more measures in both qualitative and quantitative dimensions will need to be investigated. MMUs carry an important role in the healthcare system in many situations, providing accessible, long-term care of a quality comparable to that of standard healthcare settings. Not many research papers have been published in this field. However, few authors have worked on the scope and effectiveness of mobile medical units, and few authors have even calculated the cost and assessed the functioning of mobile medical units.

As the area has not been much researched, very little literature can be found on mobile medical units. Some data have been taken from health policy data available on the national health policy of India and other websites. The data has provided a detailed view of the population of India facing challenges in primary healthcare services. Khanna and Narula (2016) expressed the descriptive knowledge of mobile healthcare units, which provided insight into finding new lines of work on mobile medical units [8]. Abbasi et. al. (2016) investigated the various challenges faced by mobile medical units. The authors concluded that the beast way to improve the effectiveness of the medical units would be reducing the number of weak units [9].

Khanna and Narula (2017) in their paper have highlighted the importance of a mobile medical unit. The paper focuses on the various shreds of evidence that may help in making mobile health units more effective. They have used various databases to explore the implications of the mobile medical unit [10]. Raikwar et. al. (2021) has provided a unit cost estimation of providing mobile medical unit services from a healthcare service provider's point of view in the local and tribal areas of Andhra Pradesh [11]. Stephanie et al. (2017) in their research have stressed the consequences of mobile health units in the US. Using the database of PubMed, the authors have analyzed various cases and their impact on healthcare services in the USA [12]. Kumar et al. (2009) in their work have presented the ways of working of mobile medical units in Jharkhand. They have represented the frequency of mobile medical units in the districts of Jharkhand concluding the factors that can express ideas for better utilization of the units [13,14]. Singh et. al. (2020) [15] have formulated the problem of mixed integer programming models to solve the generalized covering salesman problem and solved it using two different heuristics, resulting in the efficiency of the algorithm.

To the state of the art, Hodgson, Laporte, and Semet's (1998) study of a location-routing challenge for only one MMU, in which the authors decide on the vehicle stops and route simultaneously, is one of the early publications on the planning of MMUs. A feasible solution must have an appropriate number of population centers located adjacent to a vehicle's stop along the intended route, with the route's overall length kept to a minimum [16]. Hachicha, Hodgson, Laporte, and Semet (2000) expand on this issue by adding multiple cars and vehicle stops that need to be serviced in a subsequent piece. Each MMU route begins and ends at a central depot, and both the number of stops per route and the total route length is restricted to ensure the proper balance of workload among MMUs [17]. According to Ozbaygin, Yaman, and Karasan (2016), the partial coverage objective for the one-vehicle setup is to cover only those populations served by an MMU, not the population centers in reach of an MMU stop [18]. Recently, the concepts of partial coverage and numerous vehicles were further generalized by Yücel, Salman, Bozkaya, and Gökalp (2018), additionally, data-driven optimization models based on credit card transactions were integrated with their MIP concept [19]. We also cite the references for additional writings on set-covering issues with uncertainty from Büsing et al. (2021) [2].

Few research papers have been studied on the vehicle routing problem attempting to help in finding optimal routes for the problem formulated. Euchi, J., Zidi, S., and Laouamer, L. (2020) in their paper proposed a technique for routing home healthcare and worked on scheduling, also using artificial intelligence techniques to provide better services to the needy. In the paper, the authors have used the Ant colony system with cluster algorithm to achieve their results. They concluded that, through a number of benchmarks, and compared to several existing metaheuristic methodologies, a promising conclusion has been attained. In order to avoid having bad local optima, the local search technique approach is quite useful [20].

Also, to our best information, there are very few papers describing the way machine learning has helped healthcare systems. Additionally, none of them has explained the way machine learning can be utilized in depicting MMUs' routes and augmenting decisions made by healthcare experts. This paper is the first contribution presenting the use of machine learning to identify the optimized route for MMUs. The paper considers the positive and dynamic demands at different centers. The model presented in the following section takes into consideration the maximum distance an MMU can travel on a particular day. Machine learning can help us judge the optimized route for a particular MMU on a particular day.

3. Problem Definition and Formulation

3.1. Problem Definition

Now, utilizing the vehicle flow concept, we offer a mathematical formulation for the capacitated VRP. The generalized covering multi-traveling salesman problem is explained on a directed graph G = (V, A) where the set of medical centers with a single depot D, a set of villages P, and a set of location centers (where patients are to be treated) Q. A set of arcs are defined for the links between depot and location centers and are given by $A = \{(q, r) \mid q, r \in Q \cup \{D\}\}$. Each arc (q, r) is linked with a distance dqr. In addition, auxiliary sets are defined between villages and location centers, such as $B = \{(p, q) \mid p, q \in Q \cup P\}$ and it is having the distance dpq. It is assumed in the problem that the coverage of the village by a location center is dependent on the distance between them. Location center q can serve village p only if the distance dpq between them is less than or equal to a pre-specified coverage distance C. As an alternative solution, a location center Q can "cover" a village P if it is within the coverage radius Q of the location center. On a similar note, if a village p is located within a coverage radius R of a location center q then it can be 'covered' by the facility. As well, in the problem, a set of specialists' S has been considered who can fulfill the demand of location centers Q. Each village P has a demand of Jps units for a specialist S that is believed to be fulfilled only if at least one location center is within the coverage distance of the village and is visited by a mobile medical unit (MMU) M. It is assumed that a set of MMU M has a depot as its starting and end point. Hence, the objective of the problem is to determine the optimized route of the MMU with a specialist S such that the demands at each location center are maximized with a constraint that an MMU can travel a particular distance in a day. The pictorial representation of the problem is presented in Figure 1. The following notation is used to formulate a strategic programming model for the problem:

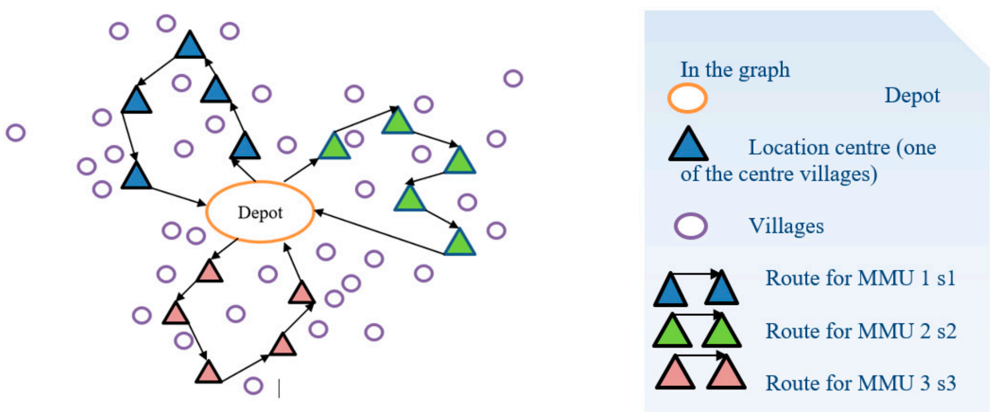

Figure 1. Vehicle Routing of Mobile Medical Units with specialist.

3.2. Problem Formulation

3.2.1. Sets and Indices

V	*Set of all vertices*
D	*Depot*
P = {1, 2, ...,m}	*Set of villages indexed by p*
Q = {1, 2, ...,n}	*Set of location centers by q, r*
S = {1, 2, ..., s}	*Set of specialists indexed by s*
M = {1, 2, . . . , o}	*Set of mobile medical units*

3.2.2. Parameters

Jps	*Demand of the patients in a village at a particular location center to be served by a specialist*
Dpq	*Distance between the villages and the location center*
Dqr	*Distance between the location center q and r*
D	*Maximum distance that a mobile medical unit can travel/reach*

$$\alpha_{pq} = \begin{cases} 1, & if\ d_{pq} \le R \\ 0, & Otherwise \end{cases}$$

The problem is modelled using the following variables, as:

$$Z_{pqs} = \begin{cases} 1, & if\ village\ p\ is\ covered\ by\ location\ center\ q\ by\ a\ particular\ specialist\ s \\ 0, & Otherwise \end{cases}$$

$$Y_{qms} = \begin{cases} 1, & if\ location\ center\ q\ is\ visited\ by\ a\ specialist\ s\ mobile\ medical\ unit\ m \\ 0, & Otherwise \end{cases}$$

$$X_{qrms} = \begin{cases} 1, & if\ mobile\ medical\ unit\ m\ travels\ from\ facility\ q\ to\ r\ by\ specialist\ s \\ 0, & Otherwise \end{cases}$$

$$v_{ms} = \begin{cases} 1, & if\ specialist\ s\ is\ assigned\ to\ mobile\ unit\ m \\ 0, & Otherwise \end{cases}$$

Uq and Us: variables to define sub-tour elimination condition.

3.3. Model

$$Objective\ 1 \quad Maximize \sum\nolimits_{p\in \boldsymbol{P}}\sum\nolimits_{q\in \boldsymbol{Q}}\sum\nolimits_{s\in \boldsymbol{S}} \boldsymbol{z_{pqs}J_{ps}} \tag{1}$$

subject to

$$\sum_{r\in Q} x_{qrms} = y_{qms} \qquad \forall q \in Q \cup \{D\}, \quad \forall m \in M,\ s \in S \tag{2}$$

$$\sum_{r\in Q} x_{rqms} = y_{qms} \qquad \forall q \in Q \cup \{D\}, \quad \forall m \in M,\ s \in S \tag{3}$$

$$x_{qrms} = v_{ms},\ \ \forall q \in Q,\ \forall m \in M,\ s \in S \tag{4}$$

$$y_{qms} \le\ v_{ms} \quad \forall q \in Q,\ \forall m \in M,\ s \in S \tag{5}$$

$$\sum_{m} v_{ms}\ \le 1, \quad \forall s \in S \tag{6}$$

$$\sum_{s} v_{ms}\ \le 1, \quad \forall m \in M \tag{7}$$

$$Z_{pqs} \le \sum_{m\in M} \alpha_{pq} y_{qms} \qquad \forall p \in D,\ \forall q \in Q,\ s \in S \tag{8}$$

$$\sum_{m}\sum_{s} y_{qms}\ \le 1,\ \forall q \in Q \tag{9}$$

$$\sum_{q\in Q}\sum_{s\in S} Z_{pqs} \le 1 \qquad\qquad \forall p \in P \tag{10}$$

$$\sum_{q\in Q\cup\{D\}}\ \sum_{r\in Q\cup\{D\}} d_{qr} X_{qrms} \le TD \qquad \forall m \in M,\ \ s \in S \tag{11}$$

$$Uq - Ur + \mathrm{n}Xqrms \leq (\mathrm{n}-1) \quad \forall q, r \in Q, \forall m \in M, \; s \in S \tag{12}$$

$$Zpq \in \{0, 1\} \quad \forall p \in P, \forall q \in Q \tag{13}$$

$$Yqm \in \{0, 1\} \quad \forall q \in Q, \; m \in M \tag{14}$$

$$Xqrm \in \{0, 1\} \quad \forall q, \mathrm{r} \in Q, \forall m \in M \tag{15}$$

$$Uq, Ur \geq 0 \, \forall q, \mathrm{r} \in Q \cup \{D\} \tag{16}$$

where

In objective function (1), the goal is to maximize patient demand in all villages that can be covered from location centers by all available specialists and mobile medical units.

Constraints (2) and (3) define moving in-arc and out-arc to and from a location center q by mobile medical unit m by specialist s, respectively, thus ensuring one location center is served by only one mobile medical unit and a specialist is required at that location center.

Constraint (4) explains that every specialist should be assigned a mobile medical unit.

Constraint (5) explains that the number of locations to be visited by a specialist should be less than equal to the number of specialists assigned in the mobile medical unit.

Constraints (6) and (7) say that one specialist is assigned to only one mobile medical unit and vice-versa.

Constraint (8) explains that a village can be served by a particular specialist and mobile medical unit only and only if the location center is visited by at least one specialist and mobile medical unit.

Constraint (9) states that the location center is visited only by a particular specialist of a mobile medical unit.

Constraint (10) ensures that patients in a village are treated only by a particular specialist of a mobile medical unit, in case of a village lying within the coverage distance of two different location centers.

Constraint (11) ensures that the mobile medical unit covers a distance that is less than or equal to the maximum distance allowed.

Constraint (12) is an elimination constraint.

The remaining constraints (13) to (16) explain the nature of different variables.

This model, also known as a distance-constrained covering multi-traveling salesman problem, can be thought of as a variation of the generalized covering traveling salesman problem. While adhering to the constraint on the maximum driving distance for each salesman, we are trying to increase the amount of demand that can be satisfied by m-salesmen visiting the facilities. However, for any given problem, there is a chance to have additional optimal solutions utilizing the same optimal value (maximum covered demand) but by varying salesmen's routes or cumulative journey distances.

To represent and justify our model, we first validated the results using the random data generated, and then with the information provided by Symbiosis University Hospital & Research Center.

4. Solution Approach

In this section, we explain our strategy for dealing with the multi-objective optimization of the distance-constrained covering multi-traveling salesman problem and outline its level of complexity as a whole. We provide the timeline technique for the time frames and halt intervals and also the cost function with which we approach the multi-objective problem as well as our two-stage strategy. The two-stage technique divides the solution space into two distinct issues, hence reducing the complexity of the optimization mechanisms as a whole. Although we cannot be certain, we assume that our technique does not have a negative effect on optimality, particularly since the used meta-heuristics are already suboptimal.

Machine learning has been categorized into three different categories. One of them is reinforcement learning (RL) which is based on goal-directed learning. When the machine must make decisions based on an experimental assumption and learns from its actions and prior experiences, RL is used [1]. In this paper, the multi-traveling salesman problem belongs to the nondeterministic polynomial hard (NP-hard) problem. RL helps in searching for the route that has maximum reward over time.

RL is the decision-making of science to learn the optimal behavior required to obtain maximum outcome rewards. It is the optimal behavior learned through interaction with the environment and observing how to respond to it.

Some important elements of the RL system are the learner, the environment, and the results, or the outcomes that the agent will get upon taking action. The RL algorithm can be classified as either model-free or model-based. An explicit environment is not created by the model-free method. and can it be treated as a trial-and-error method in which the algorithm runs with the help of the environment and the actions taken and gives the most optimal policy out of it.

The action critic algorithm, which essentially combines two approaches—value-based and policy-based—is the most efficient RL algorithm. This helps to enable the efficient use of data with consistent convergence through the employment of both algorithms.

One example of RL: autonomous driving. In an uncertain environment, an autonomous vehicle must execute numerous tasks including perception, decision-making, and planning. Vehicle route planning and motion prediction are some of the specific tasks in which RL is applied. To make judgments at various temporal and geographical scales, vehicle route planning involves several low-level and high-level policies.

Think about a situation where the environment is kept in a certain state but changes are based on the actions made. In RL, the environment is kept in a certain state but changes are based on the actions made. We have an agent that interacts with this environment; it sequentially chooses actions and receives input on how well or poorly the new state is after each action is made. RL aims at finding a method for the agent to choose actions based on the present state that, on average, results in a favorable state.

4.1. Dataset

We thank SUHRC for providing data for one month for our analysis. In the paper, the current implementation divides the villages into ones. Each zone covers 4–5 villages which are in a 5 km area proximity. Here:

V1...V4	Villages covered
	Data is of the distance from the base village (we will call it zone)
S	Specialist
P	Number of patients attended (as per the specialist 1...6)
Z	Zone

The data in Table 1 should be limited by date to assess the outcome. If we cover 20 villages in a day, then we would require 4 vans, as one van can cover 5 villages (or possibly more). The sample data would be presented as shown in Table 1.

Table 1. Sample data for execution of algorithm.

Village	**X**	**Y**	**Distance from Depot**	**V1**	**V2**	**V3**	**V4**	**Van Sent**
1	**20**	**30**	**5**	**4**	**3**	**5**	**2**	**1**
S1	S2	S3	S4	S5	S6	Dispatch time	In time	Out Time
Eye Specialist	Gynecologist	Physician	Pediatrician	ENT	Skin Specialist	10:00	10:15	17:00
P1	P2	P3	P4	P5	P6	Total	Zone	
85.00	25.00	85.00	35.00	25.00	52.00	307.00	Z1	

In the sample data table, X, Y are the coordinates of village 1. The distance of the village is 5 KM from the base station. The villages around it are V1 . . . V4. Van V1 is sent to village 1. Each van would have specialists (S1 . . . S6), and not all the villages need to have all the specialists (this sample is just as an example). In implementation, they will have binary values 0 or 1, with 1 for true and 0 for false. Then we have patients attended P1 . . . P4 (P1 is the number of patients attended by each specialist). Then we total the number of patients who attended and then we allot it a zone number.

4.2. Importing Libraries

Figure 2 shows the libraries imported in python to gain access to many functions so that we can analyse the data. In the figure it is described, that first we import panda followed by random and then calling function to solve our data using OR tools.

```
In [15]: pip install deap

         Collecting deap
           Downloading deap-1.3.3-cp39-cp39-win_amd64.whl (114 kB)
              ------------------------------------ 114.3/114.3 kB 317.3 kB/s eta 0:00:00
         Requirement already satisfied: numpy in c:\users\predator\anaconda3\lib\site-packages (from deap) (1.21.5)
         Installing collected packages: deap
         Successfully installed deap-1.3.3
         Note: you may need to restart the kernel to use updated packages.

In [43]: import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt

In [44]: import random
         from deap import base, creator, tools, algorithms, benchmarks
         from deap.benchmarks.tools import diversity, convergence, hypervolume
         import matplotlib.pyplot as plt

In [2]:  from ortools.constraint_solver import routing_enums_pb2
         from ortools.constraint_solver import pywrapcp
```

Figure 2. Importing libraries.

4.3. Importing Dataset

Firstly we import the dataset using pd.read_csv(). Then we call the data using the head() function. This function shows the first 5 columns of the dataset (Figure 3).

```
dataset = pandas.read_csv("C:\\Users\\Sai\\Downloads\\Book1.csv")
x=dataset
```

Figure 3. Importing dataset.

4.4. Data Cleaning

For data cleaning, we first see if the data is duplicated or if there are any null or NA values to be replaced or removed. We use the df.isnull() function for finding the null values and since there are no null values, we proceed to data exploration (Figure 4).

```
In [52]: x.isnull().sum()

Out[52]: village                                    0
         lat                                        0
         lng                                        0
         Distance (symbiosis medical college)       0
         van                                        0
         Patient of weakness and general issues     0
         fever                                      0
         Cardio Vascular Sytem                      0
         Respiratory org system                     0
         ENT                                        0
         eye                                        0
         git                                        0
         Genito Urinary System                      0
         Injury / Burns / Trauma                    0
         Musculo Skeleton / Bones, Joints           0
         skin                                       0
         dental                                     0
         mental                                     0
         Neurological                               0
         Obstetric                                  0
         Other                                      0
         dtype: int64
```

Figure 4. Data cleaning.

4.5. Data Exploration

For the data exploration part, we use df.info() to get the summary of the dataset. In addition to column numbers, column labels, column data types, and memory usage, it also includes the number of cells in each column (non-null values) (Figure 5).

```
In [6]: dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 32 entries, 0 to 31
Data columns (total 21 columns):
 #   Column                                  Non-Null Count  Dtype
---  ------                                  --------------  -----
 0   village                                 32 non-null     object
 1   lat                                     32 non-null     float64
 2   lng                                     32 non-null     float64
 3   Distance (symbiosis medical college)    32 non-null     int64
 4   van                                     32 non-null     int64
 5   Patient of weakness and general issues  32 non-null     int64
 6   fever                                   32 non-null     int64
 7   Cardio Vascular Sytem                   32 non-null     int64
 8   Respiratory org system                  32 non-null     int64
 9   ENT                                     32 non-null     int64
 10  eye                                     32 non-null     int64
 11  git                                     32 non-null     int64
 12  Genito Urinary System                   32 non-null     int64
 13  Injury / Burns / Trauma                 32 non-null     int64
 14  Musculo Skeleton / Bones, Joints        32 non-null     int64
 15  skin                                    32 non-null     int64
 16  dental                                  32 non-null     int64
 17  mental                                  32 non-null     int64
 18  Neurological                            32 non-null     int64
 19  Obstetric                               32 non-null     int64
 20  Other                                   32 non-null     int64
dtypes: float64(2), int64(18), object(1)
memory usage: 5.4+ KB
```

Figure 5. Data Exploration.

We also use df.describe() for the statistical description of the dataset. For any numerical data, it includes count, mean, standard deviation, minimum, 1st quartile, 2nd quartile, 3rd quartile, and maximum (Figure 6).

```
In [5]: dataset.describe()
Out[5]:
```

	lat	lng	Distance (symbiosis medical college)	van	Patient of weakness and general issues	fever	Cardio Vascular Sytem	Respiratory org system	ENT	eye	git	Genito Urinary System	Injury / Burns / Trauma	Musculo Skeleton / Bones, Joints
count	32.000000	32.000000	32.000000	32.0	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000
mean	18.517866	73.606059	22.656250	1.0	12.031250	1.062500	0.031250	24.562500	0.218750	0.281250	5.531250	0.218750	0.906250	6.375000
std	0.137473	0.117475	7.554573	0.0	14.523583	1.162242	0.176777	28.238543	0.490844	0.581121	6.010659	0.659148	1.279097	8.023152
min	18.248400	73.270600	5.000000	1.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	18.455500	73.561375	17.750000	1.0	1.750000	0.000000	0.000000	5.000000	0.000000	0.000000	1.000000	0.000000	0.000000	1.000000
50%	18.508100	73.598400	23.000000	1.0	7.000000	1.000000	0.000000	14.500000	0.000000	0.000000	4.000000	0.000000	0.000000	3.000000
75%	18.551825	73.676725	28.000000	1.0	14.250000	1.000000	0.000000	31.500000	0.000000	0.000000	8.000000	0.000000	1.000000	10.000000
max	19.114600	73.877300	33.000000	1.0	55.000000	5.000000	1.000000	95.000000	2.000000	2.000000	27.000000	3.000000	5.000000	28.000000

Figure 6. Statistical Description of Dataset.

We also used different data visualization and data analyzing tools such as Anaconda, R studio and Power BI for a better understanding of our data and to get a clear idea of our objectives. To illustrate the mathematical model discussed above in Section 3.3, we present a case problem of SUHRC with 5 different specialists in different villages around SUHRC. The results of different routes taken and total distance travelled are summarized in different graphs, as below:

The Figure 7 bar graph shows the number of patients for each category of health issue. We can see that the number of patients with respiratory issues is around 786 and for general health issues is 385.

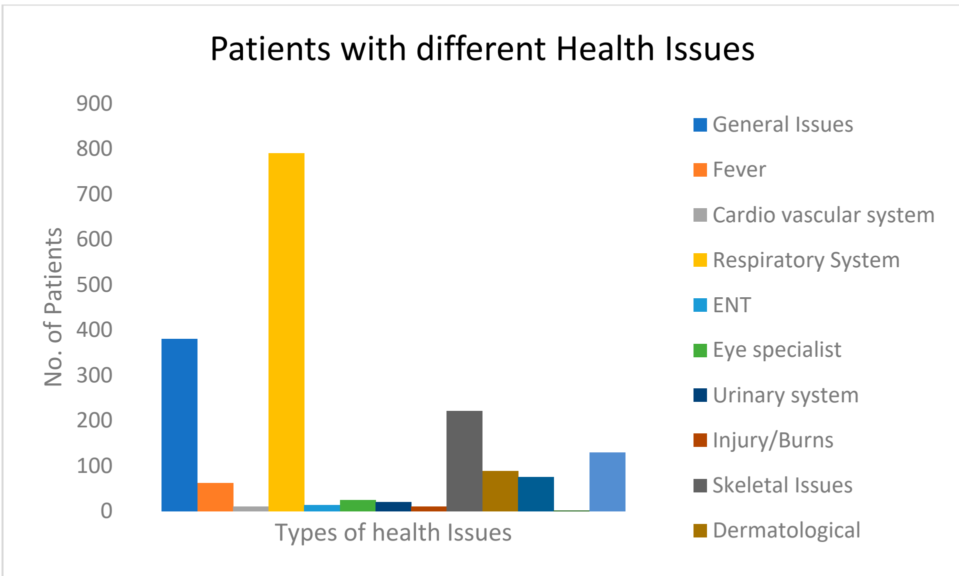

Figure 7. Health issues of patients.

Figure 8 shows the number of patients in each village. The highest number of patients are in the village of Lavale (208 patients), followed by Mulkhed (201 patients).

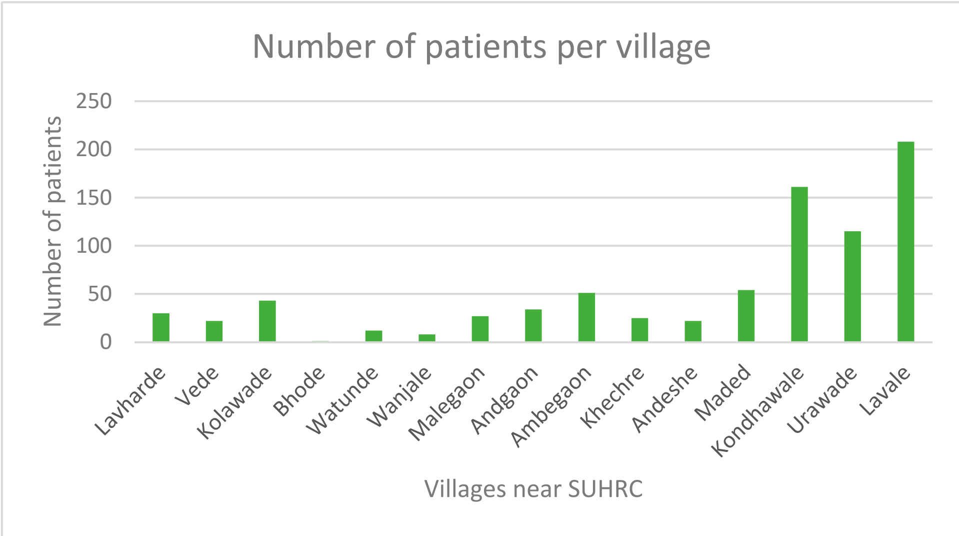

Figure 8. Number of patients per village.

Figure 9 shows the scatter plotting of all the villages the MHUs travel, using the village coordinates.

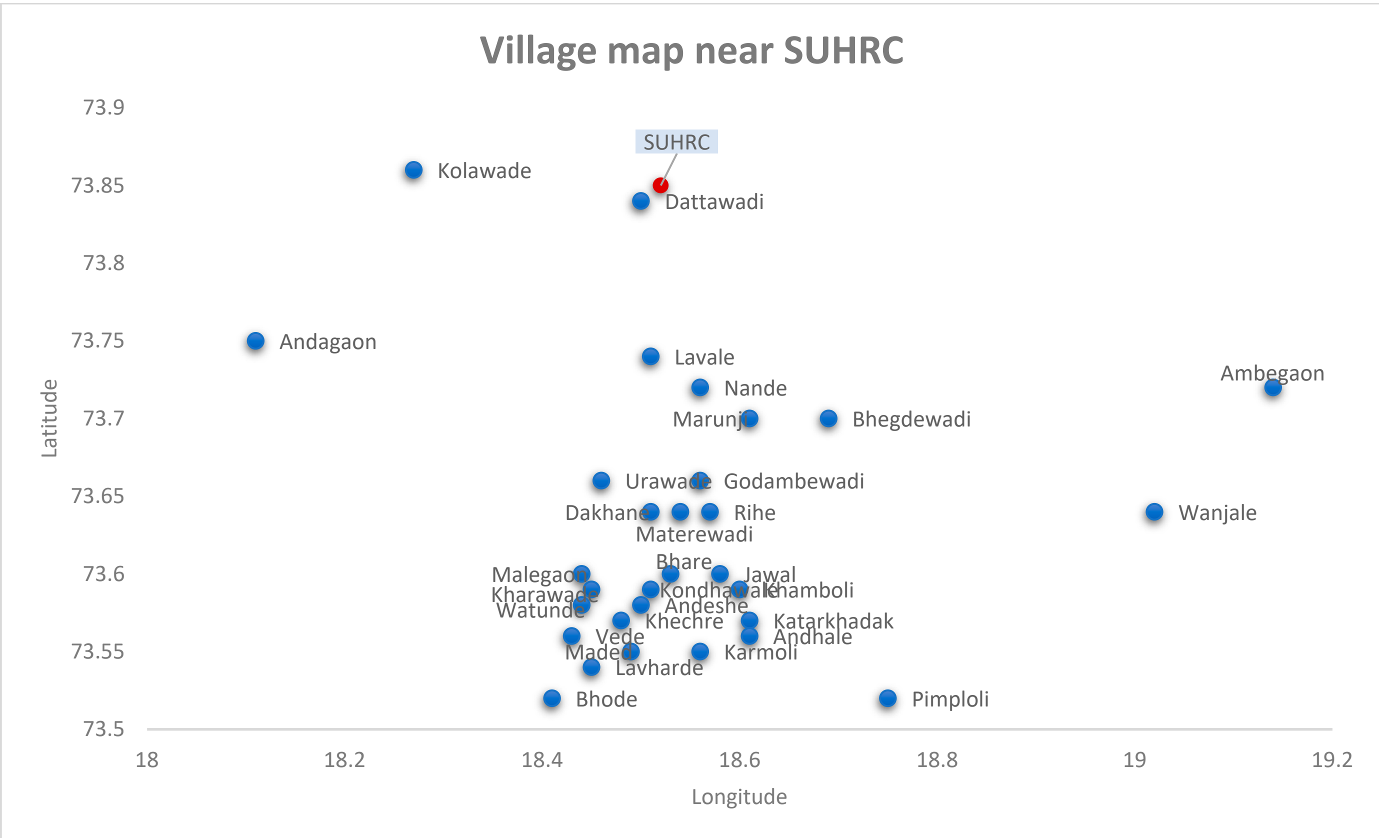

Figure 9. Scatter plot for different villages around SUHRC.

Figure 10 shows the optimal route for the number of villages we have been provided with, where 0 is the depot, i.e., SUHRC, and 1,2,3, and so on are the villages represented in Table 2, where the MMUs travels for regular checkups. The optimized diagram for different routes in Figure 10 is plotted below:

Route for Day 0:
0 → 9 → 4 → 2 → 5 → 6 → 3 → 1 → 0
Distance of the route: 73 km

Route for Day 1:
0 → 30 → 10 → 7 → 8 → 15 → 0
Distance of the route: 68 km

Route for Day 2:
0 → 18 → 27 → 28 → 0
Distance of the route: 62 km

Route for Day 3:
0 → 32 → 31 → 26 → 23 → 21 → 22 → 25 → 24 → 20 → 29 → 16 → 0
Distance of the route: 71 km

Route for Day 4:
0 → 19 → 14 → 13 → 12 → 11 → 17 → 0
Distance of the route: 72 km

Figure 10. Optimal Routes of Villages.

Table 2. Numbers indicating the villages in Figure 10.

Village	Number	Village	Number
Andagaon	1	Bhegdewadi	17
Watunde	2	Dakhane	18
Kharawade	3	Karmoli	19
Kolawade	4	Bhare	20
Bhode	5	Khamboli	21
Vede	6	Katarkhadak	22
Wanjale	7	Rihe	23
Malegaon	8	Jawal	24
Lavharde	9	Andhale	25
Ambegaon	10	Pimploli	26
Khechre	11	Godambewadi	27
Andeshe	12	Marunji	28
Maded	13	Lavale	29
Kondhawade	14	Dattawadi	30
Urawade	15	Materewadi	31
Pirangut	16	Nande	32

The route for one vehicle for day 0, as represented in Figure 11:

SUHRC → Lavharde → Kolawade → Watunde → Bhode → Vede → Kharawade → Andgaon → SUHRC

Figure 12 shows the Network diagram for the MMU for Day 1, according to the coordinates taken from Google Maps.

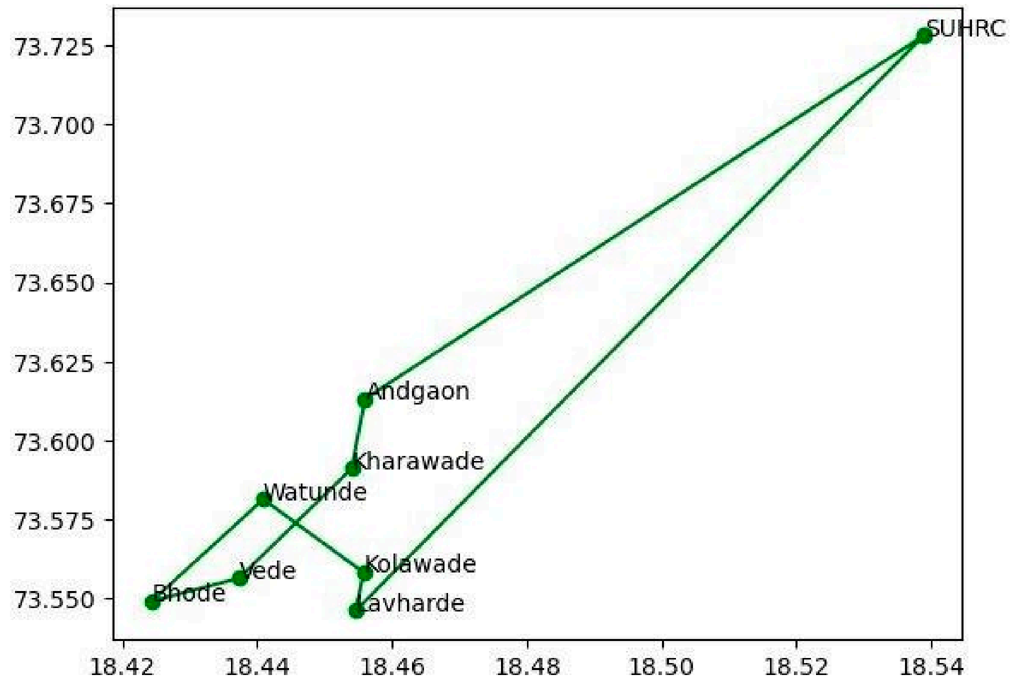

Figure 11. Route for MMU for Day 0.

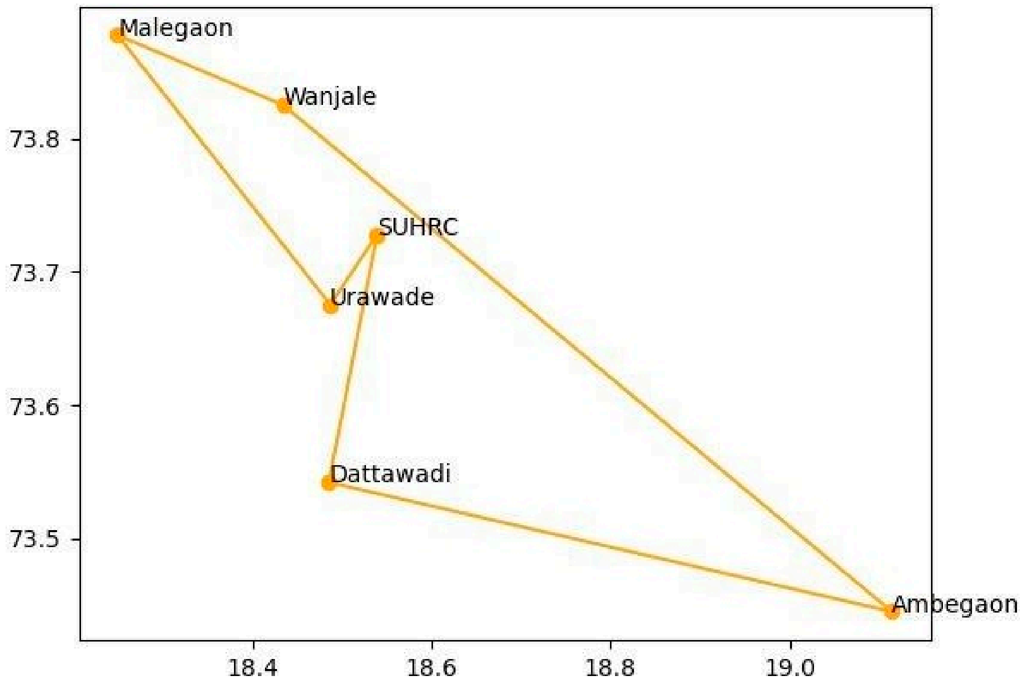

Figure 12. Route for MMU for Day 1.

Route for Day 1:

SUHRC →Dattawadi →Ambegaon →Wanjale →Malegaon →Urawade →SUHRC

Figure 13 shows the network diagram of the Route for the MMU for Day 2, according to the coordinates taken from Google Maps.

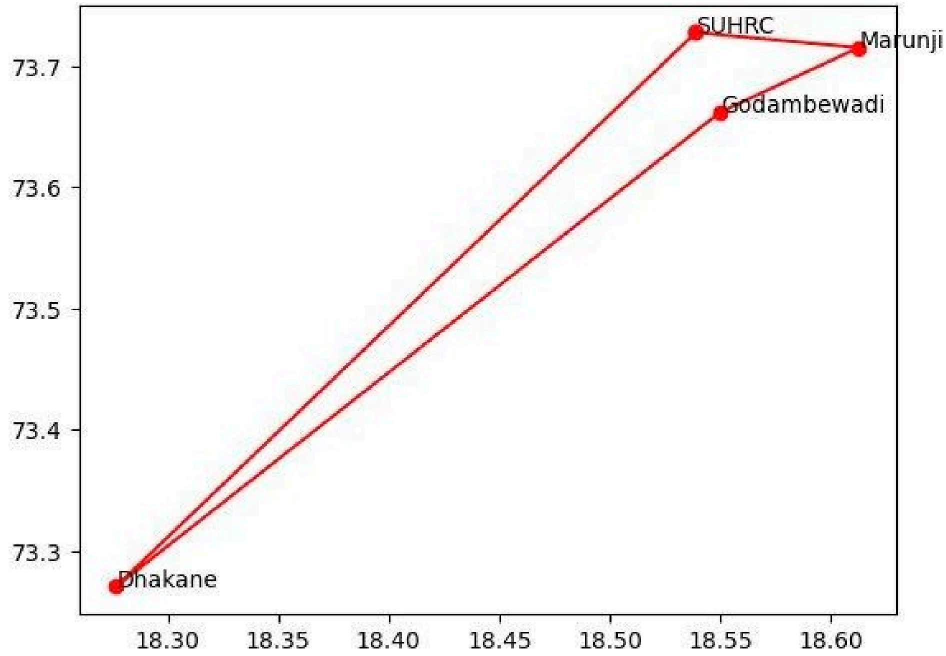

Figure 13. Route for MMU for Day 2.

Route for Day 2:
SUHRC →Dakhane →Godambewadi →Marunji →SUHRC
Figure 14 shows the network diagram of the route for the MMU for Day 3, according to the coordinates taken from Google Maps.

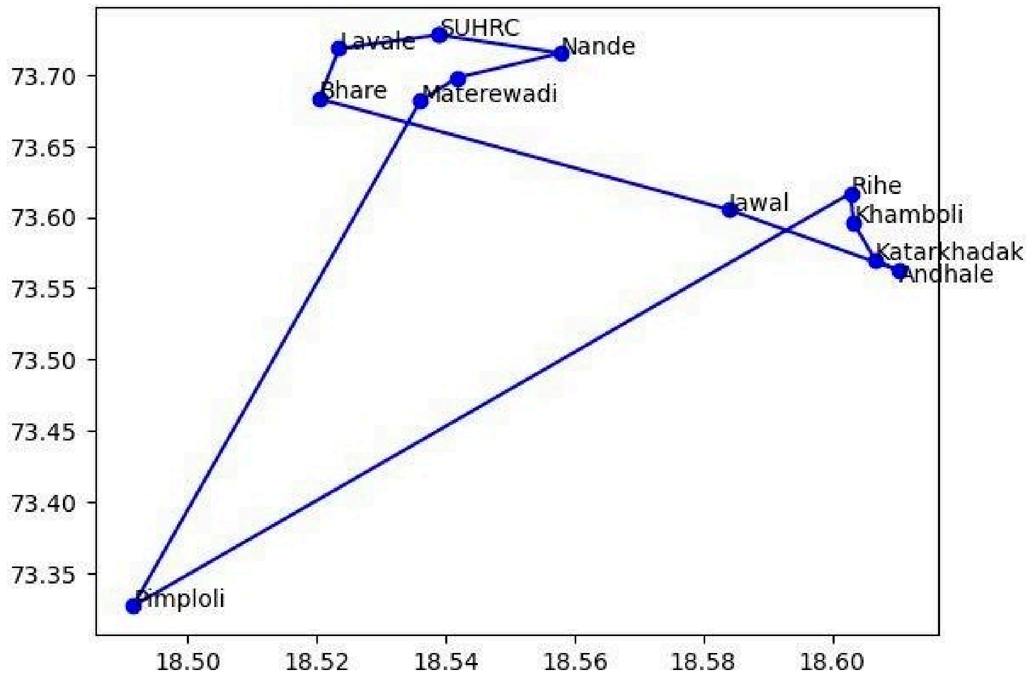

Figure 14. Route for MMU for Day 3.

Route for Day 3:
SUHRC → Nande → Materewadi → Pirangut → Pimploli → Rihe → Khamboli→ Katarkhadak → Andhale → Jawal →Bhare → Lavale → SUHRC
Figure 15 shows the network diagram of the route of the MMU for Day 4, according to the coordinates taken from Google Maps.

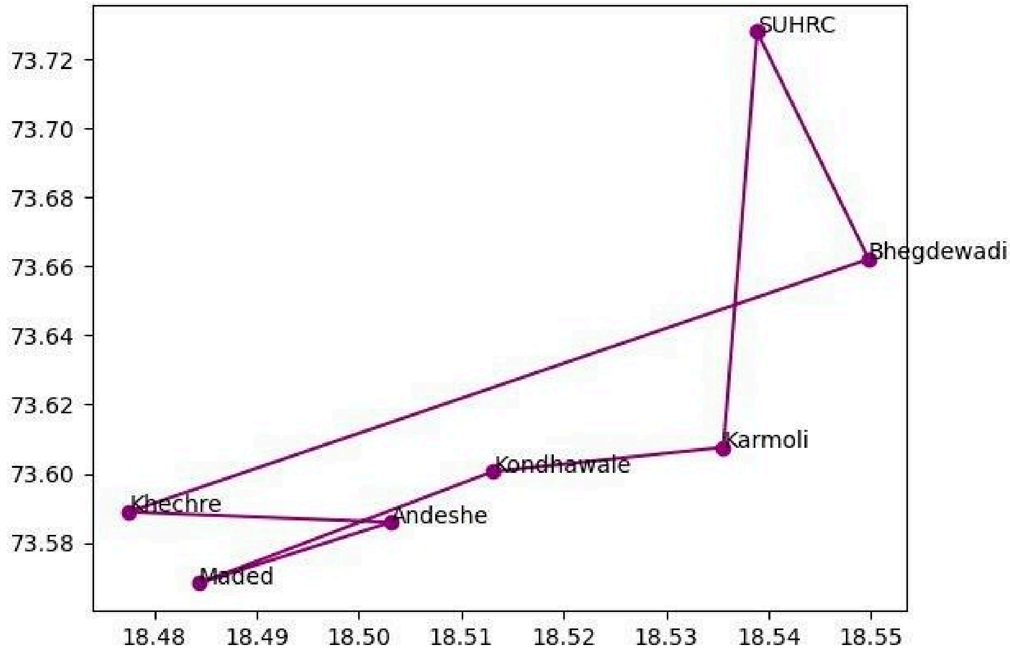

Figure 15. Route of MMU for Day 4.

Route for Day 4:

SUHRC → Karmoli → Kondhawale → Maded → Andeshe → Khechre → Bhegdewadi → SUHRC

The above Figure 16 shows the complete network diagram for the optimal route which we found by using the algorithm, according to the coordinates taken from Google Maps.

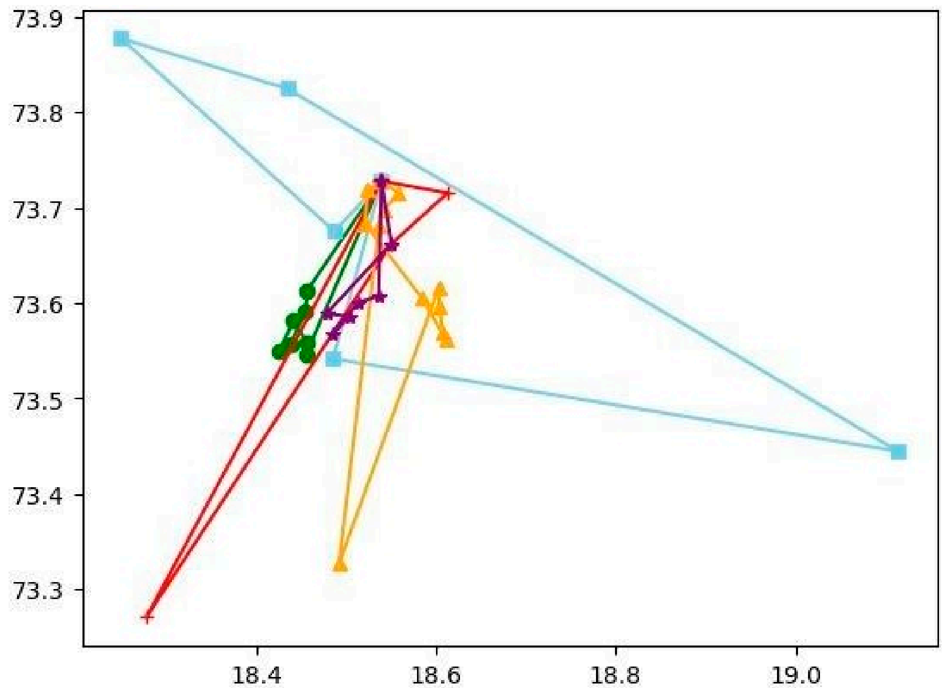

Figure 16. Optimal route for the MMU (Different colors shows the different routes for different days).

To justify the results obtained from the algorithm, a classification algorithm of machine learning is used to determine the category of a given dataset, and these algorithms are primarily employed to forecast the results for categorical data.

The basic condition of each of these algorithms is that they work after learning from past data, which is presented in the date-determined format. Additionally, machine learning works on huge dataset; it should learn from a dataset of at least few months' data. In the dataset, we treat every column as a feature of data, as explained in Table 1. The different machine learning algorithms learn from the data we feed. For example, Random Forest builds decision trees on different samples and takes their majority vote for classification and average in case of regression. Additionally, it can handle continuous as well as categorial data, which is one of its important features. The working of the whole process is explained below:

Step 1: Select random x points from the training data set.

Step 2: Create the decision trees linked to the chosen data points (subsets).

Step 3: Choose the appropriate number "N" for building the required decision tree.

Step 4: Repeat Step 1 and Step 2.

Step 5: Find each decision tree's forecasts for any new data points, then arrange them in the category that receives the most votes.

The accuracy obtained by using these steps is almost 95% for different machine learning algorithms for randomly generated data.

The different algorithms used in Figure 17 are:

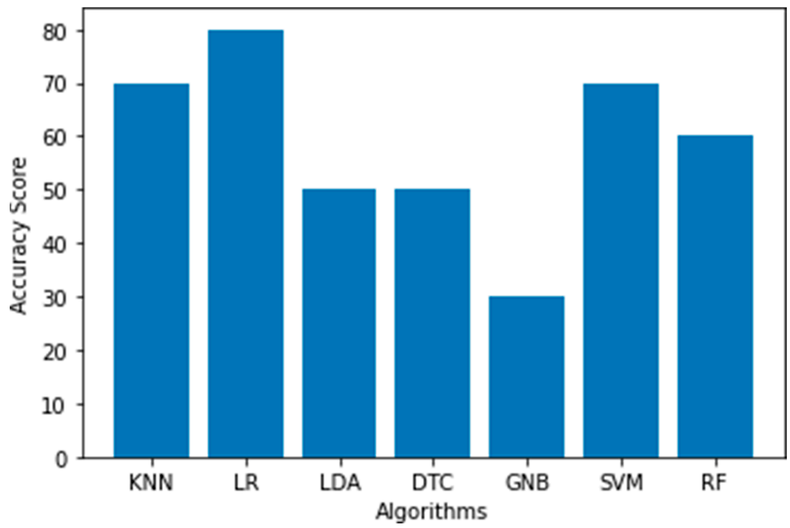

Figure 17. Accuracy score of different classification algorithm heuristics.

K Nearest Neighbor (KNN) is a supervised machine learning algorithm that classifies data based on similarity, assigns new data points, and saves existing data.

Logistic Regression (LR) is a supervised learning model, a well-known machine learning technique that forecasts categorical dependent variables using a predetermined set of independent factors.

Linear Discriminant Analysis (LDA) is one of the frequently-used dimensionality reduction techniques in machine learning used to address problems involving more than two classes.

A Decision Tree Classifier (DTC) consists of two nodes: a decision node and a leaf node. In contrast to leaf nodes, which are the results of decisions, decision nodes are created by making decisions, and have many branches.

In a Gaussian NB (GNB), two nodes—the decision node and the leaf node—make up a decision tree. A choice is made using a decision node, which has several branches, whereas a leaf node represents the result of that decision and does not have any more branches.

Support Vector Machine (SVM) operates by defining the optimal line or boundary in n-dimensional space to split n-dimensional space into classes in the future, and the SVM algorithm can then categorize fresh data points quickly. Using a hyperplane as a boundary is the best choice.

Random Forest (RF) is the most often-used algorithm. This method's goal is to merge the findings from many decision trees to get a single conclusion.

The accuracy of the different classification algorithms here is sufficient but not good, as the dataset we had was very small (just one month). Our generated accuracy, though, is much higher in the case of data generated randomly, which is approximately 95% across the months of data.

5. Conclusions and Future Scope

5.1. Conclusions

The purpose of this project was to create an optimized route using AI to minimize the distance traveled by MMUs to different underdeveloped village areas for regular check-ups of patients who are not able to get proper health facilities nearby. This project is helpful for society, as it is a facility that can solve a lot of health-related problems for people who live in underdeveloped areas. The algorithm, as can be seen, has achieved almost 95% accuracy by the random data generated over a few months.

This result also helps society, in a way that would increase the overall health of the people who live in rural places, and it will also save their expenses. The main advantage of working on this project is that we get to learn about various algorithms used to optimize routes and how they are implemented using AI. The results from this project can be provided to the hospitals or NGOs, where they can implement the algorithm. We are thankful to SUHRC for sharing their data and helping us to validate the algorithm. The algorithm can act as a base of various vehicle routing problems in home service providers such as repairing services, tutoring services, etc.

5.2. Future Scope

The application of vehicle routing in real life is more difficult. For instance, consumer requirements are frequently ambiguous. Further study on the method's application to these more challenging issues is warranted. This model can be further improved using more data and can be implemented using different optimization algorithms for even better and more accurate results.

Author Contributions: All authors were involved in conceptualization, visualization, investigation, methodology, writing—review and editing. Formal analysis, S.M.; writing—original manuscript and software implementation, K.K. and P.K.J.-validation, resource, supervision, project administration contributed to implementation. All authors have substantially contributed for the development of this manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was waived as no personal information was collected from respondents and only a general opinion survey was conducted.

Data Availability Statement: Any dataset or material used to prepare the manuscript is available from the corresponding author on reasonable request.

Acknowledgments: We would like to acknowledge the working of Mobile Medical Unit by healthcare workers in different organizations, especially Symbiosis University Hospital Research Centre.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Coronato, A.; Naeem, M.; De Pietro, G.; Paragliola, G. Reinforcement learning for intelligent healthcare applications: A survey. *Artif. Intell. Med.* **2020**, *109*, 101964. [CrossRef] [PubMed]
2. Büsing, C.; Comis, M.; Schmidt, E.; Streicher, M. Robust strategic planning for mobile medical units with steerable and unsteerable demands. *Eur. J. Oper. Res.* **2021**, *295*, 34–50. [CrossRef]
3. Mobile Medical Unit (MMUs): National Health Mission. Available online: https://nhm.gov.in/index1.php?lang=1&level=2&sublinkid=1221&lid=188 (accessed on 5 December 2022).
4. Alves, F.; Alvelos, F.P.; Rocha, A.M.A.; Pereira, A.I.; Leitao, P. *Periodic Vehicle Routing Problem in a Health Unit*; University of Minho: Braga, Portugal, 2019.
5. Jayatilake, S.M.D.A.C.; Ganegoda, G.U. Involvement of machine learning tools in healthcare decision making. *J. Healthc. Eng.* **2021**, *2021*, 6679512. [CrossRef] [PubMed]
6. Secinaro, S.; Calandra, D.; Secinaro, A.; Muthurangu, V.; Biancone, P. The role of artificial intelligence in healthcare: A structured literature review. *BMC Med. Inform. Decis. Mak.* **2021**, *21*, 125. [CrossRef] [PubMed]
7. Sarker, I.H. Machine learning: Algorithms, real-world applications and research directions. *SN Comput. Sci.* **2021**, *2*, 160. [CrossRef] [PubMed]
8. Khanna, A.B.; Narula, S.A. Mobile health units: Mobilizing healthcare to reach unreachable. *Int. J. Healthc. Manag.* **2016**, *9*, 58–66. [CrossRef]
9. Abbasi, S.; Mohajer, H.; Samouei, R. Investigation of mobile clinics and their challenges. *Int. J. Health Syst. Disaster Manag.* **2016**, *4*, 1.
10. Khanna, A.B.; Narula, S.A. Mobile Medical Units—Can They Improve the Quality of Health Services in Developing Countries? *J. Health Manag.* **2017**, *19*, 508–521. [CrossRef]
11. Raikwar, A.A.; Dogra, V.; Giri, A.; Rathnam, N.; Hegde, S.K. Cost analysis of a mobile medical unit programme in Andhra Pradesh: A microcosting study protocol. *BMJ Open* **2021**, *11*, e038191. [CrossRef] [PubMed]
12. Yu, S.W.; Hill, C.; Ricks, M.L.; Bennet, J.; Oriol, N.E. The scope and impact of mobile health clinics in the United States: A literature review. *Int. J. Equity Health* **2017**, *16*, 178. [CrossRef] [PubMed]
13. Kumar, A.; Khattar, P.; Tiwari, V.K.; Shivdasani, J.P.; Dhar, N.; Nandan, D. An assessment of functioning of mobile medical units in Jharkhand. *Indian J. Public Health* **2009**, *53*, 157–160. [CrossRef] [PubMed]
14. Kumar, R. Academic institutionalization of community health services: Way ahead in medical education reforms. *J. Fam. Med. Prim. Care* **2012**, *1*, 10. [CrossRef] [PubMed]
15. Singh, P.; Kamthane, A.R.; Tanksale, A.N. Metaheuristics for the distance constrained generalized covering traveling salesman problem. *Opsearch* **2021**, *58*, 575–609. [CrossRef]
16. Hodgson, M.J.; Laporte, G.; Semet, F. A covering tour model for planning mobile health care facilities in SuhumDistrict, Ghama. *J. Reg. Sci.* **1998**, *38*, 621–638. [CrossRef]
17. Hachicha, M.; Hodgson, M.J.; Laporte, G.; Semet, F. Heuristics for the multi-vehicle covering tour problem. *Comput. Oper. Res.* **2000**, *27*, 29–42. [CrossRef]
18. Ozbaygin, G.; Yaman, H.; Karasan, O.E. Time constrained maximal covering salesman problem with weighted demands and partial coverage. *Comput. Oper. Res.* **2016**, *76*, 226–237. [CrossRef]
19. Yücel, E.; Salman, F.S.; Bozkaya, B.; Gökalp, C. A data-driven optimization framework for routing mobile medical facilities. *Ann. Oper. Res.* **2020**, *291*, 1077–1102. [CrossRef]
20. Euchi, J.; Zidi, S.; Laouamer, L. A hybrid approach to solve the vehicle routing problem with time windows and synchronized visits in-home health care. *Arab. J. Sci. Eng.* **2020**, *45*, 10637–10652. [CrossRef] [PubMed]