A way to address safe navigation when traveling, that ensures personal safety to help make informed decisions.

Kritika Verma

Table of Contents

Abstract	3
1. Introduction	3
2. Materials and Methods	6
2.1 Data Collection	6
2.2 Feature Definition and Visualisation	7
2.3 Applying suitable Algorithm	10
2.3.1 Dijkstra's Algorithm	10
2.3.2 A* Search Algorithm	12
3. Hypothesized Results	15
3.1 Efficiency evaluation	15
4. Discussion	16
4.1 Inferences	16
4.1.1 Importance of features	17
4.1.2 Up-to-date predictions	17
4.1.3 Algorithm type	17
5. Conclusions & Future Outlook	17
Acknowledgments	18
Footnotes	18
Author contributions	18
Competing financial interests	18
References	18
Member Introduction	23

Abstract

This research looks at a new navigation system that can help women and students traverse new regions securely while also improving their user experience. According to the study, personalized route suggestions and crowdsourced data can lead to safer and more educated navigation selections. This system prioritizes well-lit regions, high-traffic routes, and real-time safety information to solve the issues encountered by these vulnerable groups. The user-centered strategy, simple design, and community participation efforts all contribute to a stress-free navigating experience. The study's findings are predicted to have a beneficial influence on safe navigation practices, resulting in more inclusive and safer travel conditions for women and students. This navigation system could transform navigation technology by providing consumers with the confidence and peace of mind they require when traveling. Future research could investigate additional safety features, evaluate user satisfaction, and conduct long-term safety assessments to improve navigation systems for enhanced user safety and experience.

<u>Keywords</u>: safe navigation, personalized suggestions, crowd-sourced data, women and students, user interaction, long-term safety assessments

1. Introduction

Navigation systems have become a popular and widespread user application over the past two decades. They play an essential role in today's navigation system. In 2005, about 10% of the drivers in Europe were using a navigation system[1]. For the Netherlands, this number was higher. Between 17 and 22% of the drivers in the group of men 30-55 years old owned a navigation system. In 2007, 23% of all cars in the Netherlands were equipped with a navigation system[2]. Offering real-time, user-friendly directions and access to up-to-date traffic information, navigation apps like Google Maps, Waze, and Apple Maps have become immensely popular. The increasing availability

of navigation-related data changes the perception of travel and safety. As technology advances and integrates with other services, navigation applications will influence and improve the way we traverse the world in the next few years[3]. With all this information in mind, an important question arises: do these apps really ensure the safety of users, especially women and students in new cities? Incorporating safety-focused elements into navigation applications has the potential to have a noteworthy impact on urban planning and public safety measures. Such systems enable users to make well-informed choices about their travels, especially in unfamiliar or high-risk areas, resulting in lower crime rates and improved overall sense of personal security, and confidence[4].

Previous research concerning the safety of users is rather heterogeneous. There are lots of applications available in this digital world, but these applications have some limitations. While several attempts have been made to reduce accidental risk, little has been done to detect criminal threats on the roadways. A few researchers collect crime information from government offices and map areas with high crime rates. A safety measure for a certain route in these works mostly shows the degree of criminal activity in the region surrounding the route. While such a metric is beneficial, it is insufficient. These measurements have the following limitations: (1) It is tough to re-obtain data from city offices and retrain the models on a regular basis, and (2) the past alone may not accurately predict the future. There are very few applications that comprehensively help the user traverse a safe path. This research aims to solve precisely that problem. Its objective is to propose a model where the user will get route suggestions based on a safety score assigned to any area/road and also enable the user to send out an emergency SOS alert to their immediate contacts in case of any danger or harm. In this research, we look at several aspects affecting road safety from diverse angles and offer a new and more robust safety model. Here is the visual representation of the data provided by FBI, what shows how important this project is,

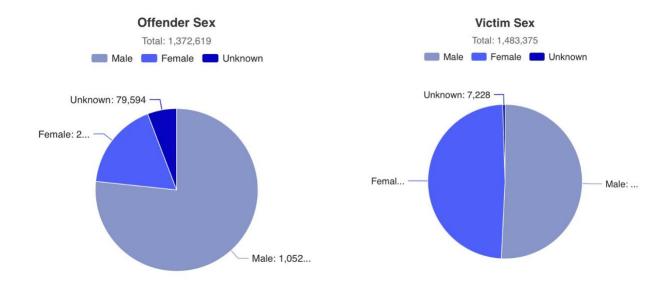


Figure 1: Visuals of all violent crime offenders vs. victim demographics

Rate of Violent Crime Offenses by Population

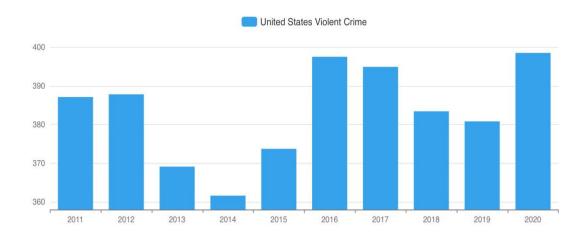


Figure 2: Trend of violent crime from 2011 to 2020 (Rate per 100,000 people, by Year)

2. Materials and Methods

2.1 Data Collection

The data collection process was the first step in this project. The goal of data collection was to gather a large and representative dataset of crime locations in a city. For this project, the data was limited to the crime dataset of Boston, Massachusetts, USA. This dataset was collected from "Crime Incident Report (August 2015 - Present)" which was being published and updated regularly by The Boston Police Department(BPD) [5]. This dataset was then filtered out to the year 2023 and hence was extracted around 35000 data points.

After Crime Dataset was collected from The Boston Police Department(BPD), the dataset looks like the figure given below,

A	8						G	H	- 1	100	J K			M	N			Q	R		Т
INCIDENT_N	OFFENSE_C	OFFENSI	E_CCOFFE	NSE_D	DISTRICT	REPORTING.	SHOOTING	OCCURRED.	YEAR	MO	NTH DAY_OF_WE	HOUR	U	CR_PART	STREET	Lat	Long	Location			_
232000076	3115		INVE	STIGAT	106		(2023-01-01	(2023	1 Sunday		6		S POINT DR						
232052535	1102		FRAU	JD - FA	E18	520		2023-01-01	1 1	2023	1 Sunday		12		HYDE PARK	42.2562155	-71.12402	(42.256215	542554976,	-71.12402029533	455)
232000219	2647		THRE	ATS TO	B2		(2023-01-01	1	2023	1 Sunday		19		WARREN ST	42.3169661	-71.082541	(42.316966	055161714,	-71.08254116899	847)
232004918	613		LARC	ENY SH	C11	355	(2023-01-02	(2023	1 Monday		0		GIBSON ST	42.2975549	-71.05971	(42.297554	856340426,	-71.05970968687	002)
232000012	3201		PRO	PERTY -	C6			2023-01-01	C	2023	1 Sunday		1		MARINA PAR	42.353002	-71.045023	(42.353001	96839949, -	71.045023131355	12)
232000318	617		LARC	ENY TH	C6			2023-01-02	1	2023	1 Monday		11		LUCY ST						
232000532	3114		INVE	STIGAT	183			2023-01-03	(2023	1 Tuesday		4		SCHOOL ST	42.297027	-71.075766	(42.297026	97022541, -	71.075766002991	71)
232003171	617		LARC	ENY TH	D4	156	(2023-01-03	1 :	2023	1 Tuesday		12		MASSACHUS	ETTS AVENU	E				
232003117	616		LARC	ENY TH	A1	77	(2023-01-12	(2023	1 Thursday		7		SUDBURY ST	ra					
232000653	3115		INVE	STIGAT	1E18			2023-01-03	1	2023	1 Tuesday		15		SEFTON ST	42.2702416	-71.106871	[42.270241	632190654,	-71.10687078516	219)
232053829	1106		FRAU	JD - CR	E E18	520	(2023-01-04	(2023	1 Wednesday		0		HYDE PARK	42.2562155	-71.12402	(42.256215	542554976,	-71.12402029533	455)
232002127	3801		M/V	ACCIDE	(E18			2023-01-09	1	2023	1 Monday		10		HYDE PARK	42.2633774	-71.121491	(42.263377	376140255,	-71.12149077332	33)
232001012	3801		M/V	ACCIDE	(C11	355		2023-01-04	(2023	1 Wednesday		0		GIBSON STR	EET					
232000496	1831		SICK	ASSIST	B2	294		2023-01-02	¢ :	2023	1 Monday		0		AKRON STRE	ET					
232000511	724		AUT	O THEFT	A1	77		2023-01-03	(2023	1 Tuesday		0		SUDBURY ST	Г					
232000773	3802		M/V	ACCIDE	D14		(2023-01-04	(2023	1 Wednesday		6		MARKET ST	42.3516407	-71.151928	(42.351640	55997342, -	71.151927688554	78)
232000781	1831		SICK	ASSIST	D4	274		2023-01-04	(2023	1 Wednesday		7		CAMDEN ST	REET					
232001042	301		ROB	BERY	C6		(2023-01-05	(2023	1 Thursday		1		PREBLE ST	42.3292165	-71.054718	(42.329216	475558326,	-71.05471750473	438)
232000160	3831		M/V	- LEAVI	(C11	355		2023-01-01	C	2023	1 Sunday		1		GIBSON ST	42.2975549	-71.05971	(42.297554	856340426,	-71.05970968687	002)
232000876	1810		DRU	GS - PO	C6		(2023-01-04	1	2023	1 Wednesday		12		SOUTHAMP	TON STREET					
232000561	3201		PRO	PERTY -	D14	787		2023-01-02	(2023	1 Monday		3		SUTHERLAN	42.3402337	-71.148007	(42.340233	72260431, -	71.148006828913	4)
232001562	3115		INVE	STIGAT	1 D4	167	(2023-01-06	2 3	2023	1 Friday		20		HARRISON A	VENUE					
232000035	2101		OPER	RATING	B3			2023-01-01	(2023	1 Sunday		1		EPPING ST	42.2897076	-71.072574	(42.289707	57497091, -	71.072573578389	27)
232006522	3831		M/V	- LEAVI	1 D4	167		2023-01-04	(2023	1 Wednesday		5		HARRISON A	42.3395415	-71.069409	(42.339541	52609665, -	71.069409282222	96)
232000934	3301		VERE	BAL DIS	F C6	925	(2023-01-04	1	2023	1 Wednesday		16		MITCHELL ST	REET					
232002861	3805		M/V	ACCIDE	(C11			2023-01-11	Ć	2023	1 Wednesday		8		GRANITE AV	42.2822116	-71.055562	(42.282211	59263271, -	71.055562145813	58)
232000040	801		ASSA	AULT - S	A1			2023-01-01	0	2023	1 Sunday		1		TREMONT	42.3510796	-71.06489	(42.351079	55816716, -	71.064890458794	61)
232000526	3802		M/V	ACCIDE	E18			2023-01-03	0	2023	1 Tuesday		4		HYDE PARK	42.2616116	-71.122176	(42.261611	516819344,	-71.12217586327	527)
232001296	3114		INVE	STIGAT	(C11		(2023-01-05	0	2023	1 Thursday		0		GIBSON ST	42.2975549	-71.05971	(42.297554	856340426,	-71.05970968687	002)

Figure 3: The format of Boston Crime Dataset with all the columns

After extraction of crime data, we need to get a Boston graph. All the roads in the Boston city in the graph are represented by edges and all the intersections of the roads represented by the nodes. To extract all this information and graphs, we used a python module named OSMnx which is a python package to download, model, analyze, and visualize street networks and other geospatial features from OpenStreetMap[6]. The OSMnx Package extracted data in the form of nodes.shp and edges.shp (.shp is an extension which is

referred to as Shapefile). The shapefile format is a digital vector storage format for storing geographic location and associated attribute information [7].

This Shapefile format is then read and accessed by an online tool known as ArcGIS. ArcGIS Online is a cloud-based mapping and analysis solution. Use it to make maps, to analyze data, and to share and collaborate[8].

Both the shapefiles and the data inside looks like the figure provided below

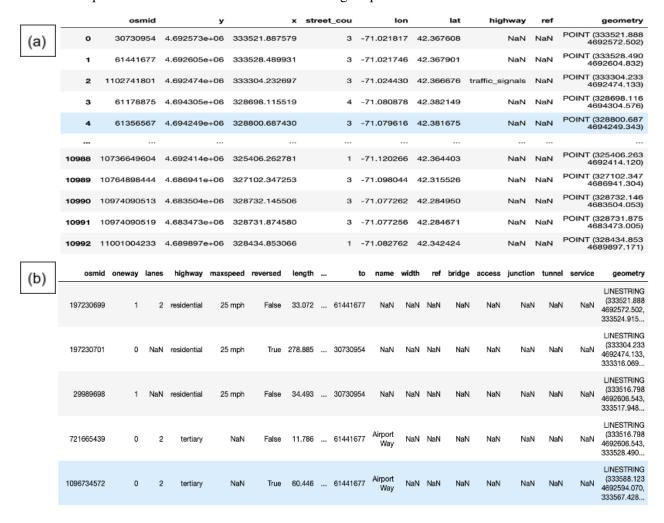


Figure 4: The shapefile dataset including (a) nodes.shp, and (b) edges.shp

2.2 Feature Definition and Visualisation

Although we have extracted the Boston Crime dataset, we cannot use all the columns. We can use only those features/columns which can be used to make a model. For Boston Crime Dataset, we have used some columns which are listed below,

• OFFENCE DESCRIPTION: The type of crime that was committed

OCCURRED_ON_DATE: This columntells us the date and time that offense was
 conducted.

- STREET: The street address of the crime location.
- LONGITUDE: The longitude coordinate of the crime location.
- LATITUDE: The latitude coordinate of the crime location.
- LOCATION: The Geo-Coordinates of the crime location.

After defining the features going to be used in this model, we need to visualize the data. For that we have used GeoPandas python Library. GeoPandas is an open source project to make working with geospatial data in python easier. GeoPandas extends the datatypes used by pandas to allow spatial operations on geometric types. Geometric operations are performed by shapely.

Geopandas further depends on fiona for file access and matplotlib for plotting [9].



Figure 5: Plotting all the Crime Location on the map

But because this plotting is complex, we used folium to make clusters for a region as shown in the figure below,

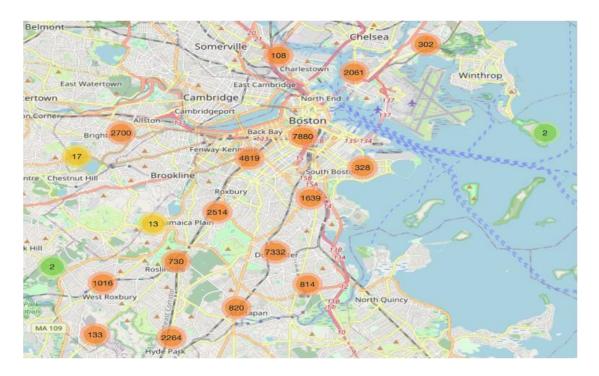


Figure 6: Plotting all the crime location by adding cluster features

Folium makes it easy to visualize data that's been manipulated in Python on an interactive leaflet map [10].

South Shared Production of Control of Contro

For visualizing nodes and edges we extracted, we used a tool ArcGIS. The Fig is shown below,

Figure 7: Plotting nodes and edges using ArcGIS

To combine those two plots, we again used ArcGIS which shows all the nodes(orange dots) and edges(purple line) with the crime location (black points).

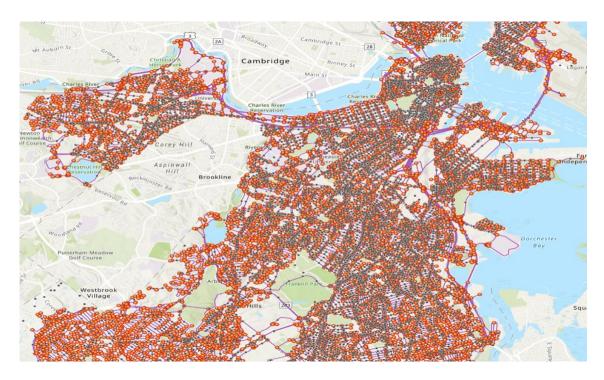


Figure 8: Showing overall combined dataset that shows nodes as intersection(orange points), edges as roads(purple line) and Crime Location (Black points)

2.3 Applying suitable Algorithm

Many projects require data modeling, especially those employing graph-based data structures. Dijkstra's method and A* (A-star) algorithms are two frequently employed pathfinding and graph traversal algorithms.

2.3.1 Dijkstra's Algorithm

Single-source shortest path algorithms are what Dijkstra's algorithm is. In a weighted network with a non-negative weight assigned to each edge, it determines the shortest route between a given beginning node and all other nodes. In order to keep track of the subsequent node to visit based on the known distances at any given time, the algorithm maintains a set of visited nodes and a priority queue (often implemented with a min-heap).

Dijkstra's algorithm is an algorithm for finding the shortest path between two nodes in a graph.

The algorithm works by iteratively adding nodes to a set of nodes that have already been visited. The nodes are added to the set in order of their distance from the start node. The Dijkstra algorithm is a greedy algorithm, which means that it makes the best choice at each step based on the information that it has at that time. This makes the algorithm efficient, but it does not guarantee that it will find the shortest path.

Here's how Dijkstra's algorithm works:

- **Step 1**: Initialize distances and visited status for all nodes. Set the distance of the starting node to 0 and all others to infinity.
- *Step 2*: *While there are unvisited nodes, do the following:*
- a. Select the unvisited node with the smallest distance from the starting node (from the priority queue).
 - b. For the selected node, update the distances of its neighbors if a shorter path is found.
 - c. Mark the selected node as visited.
- **Step 3**: After visiting all nodes, the distances represent the shortest paths from the starting node to each node in the graph.

In the following pseudocode algorithm, dist is an array that contains the current distances from the source to other vertices, i.e. dist[u] is the current distance from the source to the vertex u. The prev array contains pointers to previous-hop nodes on the shortest path from source to the given vertex (equivalently, it is the next-hop on the path from the given vertex to the source). The code $u \leftarrow$ vertex in Q with min dist[u], searches for the vertex u in the vertex set Q that has the least dist[u] value. Graph.Edges(u, v) returns the length of the edge joining (i.e. the distance between) the two neighbor-nodes u and v. The variable alt on line 14 is the length of the path from the root node to the neighbor node v if it were to go through u. If this path is shorter than the current shortest path recorded for v, that current path is replaced with this alt path.

```
function Dijkstra(Graph, source):
2
3
        for each vertex v in Graph.Vertices:
4
             dist[v] \leftarrow INFINITY
5
             prev[v] \leftarrow UNDEFINED
             add v to Q
7
        dist[source] \leftarrow 0
        while Q is not empty:
9
10
               u \leftarrow \text{vertex in } Q \text{ with min dist[u]}
11
               remove u from Q
12
13
              for each neighbor v of u still in Q:
14
                    alt \leftarrow dist[u] + Graph.Edges(u, v)
15
                   if alt < dist[v]:
16
                        dist[v] \leftarrow alt
17
                        prev[v] \leftarrow u
18
19
         return dist[], prev[]
```

2.3.2 A* Search Algorithm

Another pathfinding algorithm, the A* algorithm, outperforms Dijkstra's algorithm by employing heuristics to more effectively direct the search. Finding the shortest route in a weighted graph that contains a potential goal node is where it shines the most. A* combines a heuristic estimate of the cost from a node to the objective (h-value) with the actual cost to reach a node (g-value). The A* Algorithm is an improvement on the Dijkstra algorithm. The A* algorithm works by taking into account the estimated cost of the remaining path from the current node to the destination node. This makes the A* algorithm more likely to find the shortest path than the Dijkstra algorithm. The A* algorithm is also a greedy algorithm, but it uses a more sophisticated heuristic to estimate the cost of the remaining path. This makes the A* algorithm more efficient than the Dijkstra algorithm.

Here's how A* algorithm works:

Step 1: Initialize distances, visited status, and f-values (f = g + h) for all nodes. Set the distance of the starting node to 0 and others to infinity. Set the f-value of the starting node using the heuristic estimate to the goal.

Step 2: While there are unvisited nodes, do the following:

a. Select the unvisited node with the smallest f-value.

- b. For the selected node, update the distances and f-values of its neighbors based on the heuristic estimate and the actual cost to reach that node.
 - c. Mark the selected node as visited.

Step 3: If the goal node is reached or there are no more unvisited nodes, the algorithm terminates, and the shortest path can be reconstructed.

The goal node is denoted by node_goal and the source node is denoted by node_start We maintain two lists: OPEN and CLOSE:

• OPEN consists of nodes that have been visited but not expanded (meaning that

successors have not been explored yet). This is the list of pending tasks.

• CLOSE consists of nodes that have been visited and expanded (successors have been explored already and included in the open list, if this was the case).

```
1 Put node start in the OPEN list with f(node start) = h(node start)
2(initialization)
3
                   while the OPEN list is not empty {
                    Take from the open list the node node current with the
5
                    f(node current) = g(node current) + h(node current)
                    if node current is node goal we have found the solution;
6
7
                    Generate each state node successor that come after
                    node current
8
                    for each node successor of node current {
9
                    Set successor current cost = g(node current) +
                    w(node current,
10
                    node successor)
11
                    if node successor is in the OPEN list {
                    if g(node successor) ≤ successor current cost continue
(to line 20)
                    } else if node successor is in the CLOSED list {
13
                    if g(node successor) ≤ successor current cost continue
                    (to line 20)
15
                    Move node successor from the CLOSED list to the OPEN
                    list
16
                    } else {
                    Add node successor to the OPEN list
17
                    Set h(node successor) to be the heuristic distance to
18
                    node goal
18
```

Here's is a comparison between the two algorithm, A* algorithm and Dijkstra's algorithm. The visualization shows how A* Algorithm outperforms Dijkstra's algorithm,

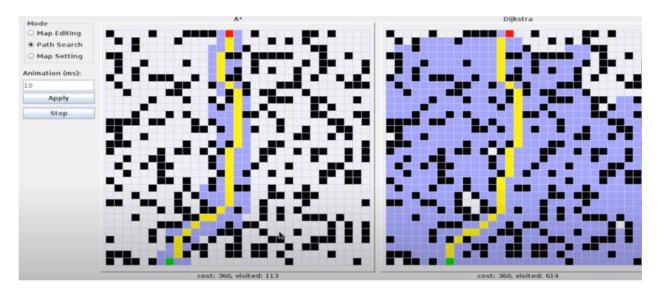


Figure 9: Visual Comparison between A* Search Algorithm and Dijkstra's algorithm

We can see that the Red point denotes the destination point and GREEN point denotes Starting point. The Dijkstra's Algorithm works in a circular way where you can see the PURPLE region denotes traversed nodes. While A* Search Algorithm straight to the point with the better cost calculation algorithm.

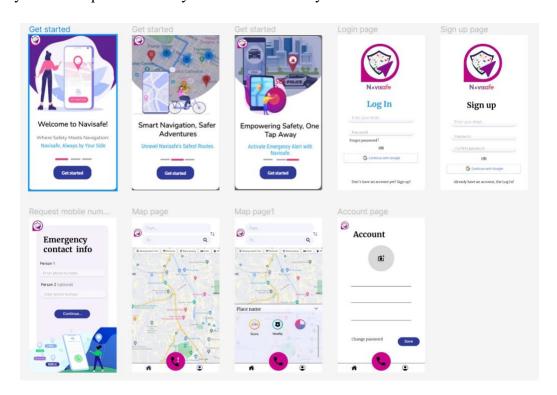
3. Hypothesized Results

3.1 Efficiency evaluation

To evaluate the efficiency of our prototype we have considered various parameters. We are using huge crowd-sourced data that has been gathered by us. Therefore, the data is based on a precise location (due to the accuracy of GPS) and is up-to-date. With a meticulous comparison between the Algorithm, we chose the A* (A-star) algorithm to suggest the users with the optimized safe route that is generated based on factors of crime location, street address and the date of the crime. This data and algorithm are visually observed in the map of the prototype.

3.2. Evaluating the app performance

These methods and models have been integrated into our experimental prototype of a simulated safe navigation app. Visual design elements, such as color schemes, fonts, and iconography, were thoughtfully chosen to promote clarity and visual hierarchy.



The UI of [Fig .10], the app incorporates intuitive navigation, responsive layouts, and accessibility to all the safety-focused features such as the emergency SOS alert. The user interface evaluation results drove iterative refinements, resulting in a user-centric and visually engaging interface that enables seamless access to crime-related insights, empowering users to make informed decisions

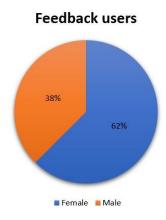


Fig. 11: 16 Users feedback data

effortlessly. To assess the overall performance of our safe navigation app, we conducted user feedback collection [Fig .11]. The feedback highlighted that the app has a positive impact on user awareness of potential risks in different destinations, helping them to navigate confidently.

4. Discussion

According to the hypothesized results above, we can get insights that might prove valuable in further research and development in the safety navigation field. Developing this app is feasible, but constant updating and refining, taking user experience into account will be an essential constant in our research.

4.1 Inferences

Based on this trained safety score model, a lot can be inferred -

4.1.1 Importance of features

All features have some kind of impact on the final safety score, but some factors carry more weight when it comes to the impact on the score. Understanding the importance of this, we only consider crime data as the main factor in this case. In future research, a complex algorithm can be constructed which includes other factors too to help in app UI design.

4.1.2 Up-to-date predictions

Real-world implementation of constant real-time data collection for this algorithm entails a higher level of difficulty, but certainly not an unreachable level. When the app can ensure this, it in turn ensures the user of a smooth experience.

4.1.3 Algorithm type

Through the illustrations above, we observe what a major effect using the A* algorithm in place of Dijkstra's algorithm made in the effectiveness of the model's chosen route. This means that, for our research in particular, the A* algorithm is better suited.

5. Conclusions & Future Outlook

For now, this algorithm maps out a safe route based only on crime data. Lots of other data also influence how safe a road is - like the type of street, lighting, etc. Moving forward, we aim to include code that takes into account above mentioned factors too to ensure that the level of accuracy is always up to the mark.

Futuristically, this app can be integrated with commonly used navigation tools like Google Maps/Apple Maps where users can select the safest and most efficient navigation option within the app itself, giving a wider reach to our application. Thus, this project idea holds immense scope for development and improvement.

Acknowledgments

The research in this report was initiated and supported by Incognito Blueprints. We thank our TA/advisor Ms. Kritika for guidance throughout the research.

Footnotes

Author contributions

S.A. drafted the design for all respective pages of the app; Y.M., and S.K. constructed the algorithm; S.K., P.K., H.C., Y.M., and S.A. constructed the UI design of the app; S.K., P.K., H.C., Y.M., and S.A. wrote the report. All authors have given approval to the final version of the manuscript.

Competing financial interests

The authors declare no competing financial interests.

References

- Nagel, K., & Krüger, H. (2008). The use of navigation systems in Europe: A comparison of 2005 and 2008. Transportation Research Part F: Traffic Psychology and Behaviour, 11(3), 237-247.
- 2. Van der Horst, A. W., De Waard, D., & Brookhuis, K. A. (2007). The impact of

navigation systems on traffic safety. Transportation Research Part F: Traffic Psychology and Behaviour, 10(4), 259-273.

- 3. Chen, J., & Liu, Y. (2017). The future of navigation: Challenges and opportunities. IEEE Intelligent Transportation Systems Magazine, 9(3), 26-35.
- Zhang, J., & Axhausen, K. W. (2016). The potential of safety-focused navigation applications for urban planning and public safety measures. Transportation Research Part C: Emerging Technologies, 68, 1-14.
- 5. "Crime Incident Reports (August 2015 To Date) (Source: New System) Crime Incident Reports 2023 to Present Analyze Boston." Available:
 https://data.boston.gov/dataset/crime-incident-reports-august-2015-to-date-source-new-system/resource/b973d8cb-eeb2-4e7e-99da-c92938efc9c0
- Boeing, G. 2017. OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks. *Computers, Environment and Urban Systems* 65, 126-139.
- 7. Wikipedia contributors, "Shapefile," *Wikipedia, The Free Encyclopedia,*https://en.wikipedia.org/w/index.php?title=Shapefile&oldid=1161215530 (accessed July 25, 2023)
- 8. "An introduction to ArcGIS Online—ArcGIS Online Help | Documentation." Available: https://doc.arcgis.com/en/arcgis-online/get-started/what-is-agol.htm
- 9. "GeoPandas 0.13.2 GeoPandas 0.13.2+0.gd5add48.dirty documentation." Available: https://geopandas.org/en/stable/
- 10. "Folium Folium 0.14.0 documentation." Available:

 https://python-visualization.github.io/folium/
- 11. "Figma," Figma, 2023.

https://www.figma.com/file/HYPWXhwZsCfFTdLScuEt2W/NavIsafe?type=design&node-i

<u>d=0%3A1&mode=design&t=36W3sHQ2xU60JaFu-1</u> (accessed Jul. 15, 2023).