An Enhanced Ride Sharing Model Based on Human Characteristics and Machine Learning Recommender System

Govind Yatnalkar, Husnu S. Narman, Haroon Malik

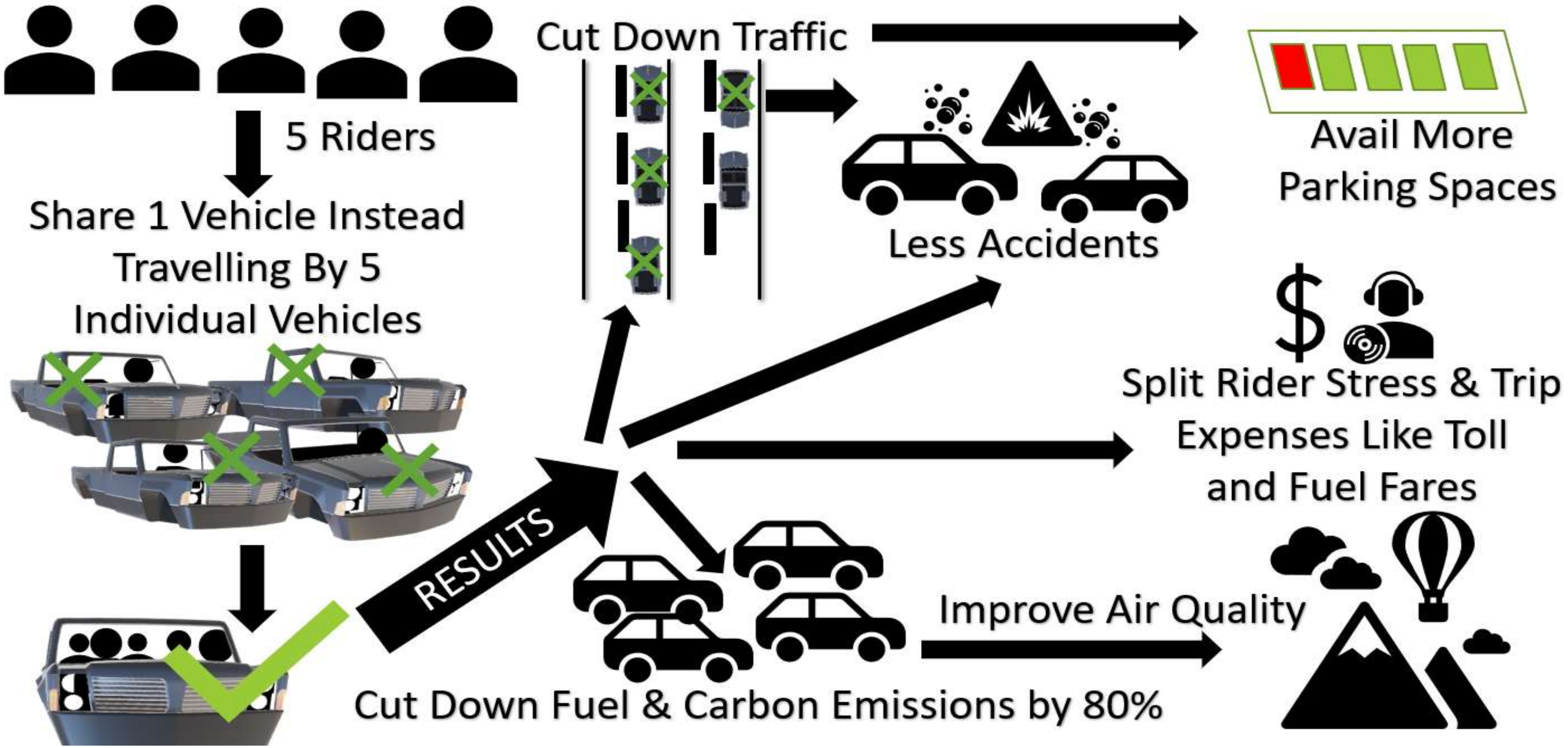
The 3rd International Conference on Emerging
Data and Industry 4.0 (EDI40)
April 6 - 9, 2020, Warsaw, Poland

Agenda for the Presentation

1. Motivation
2. Enhanced Ride Sharing Model (ERSM)
3. System Architecture
4. The Proposed Model
5. The Feedback System and the Machine Learning Models
6. Experimentations
7. Results and Analysis
8. Conclusion

1. Motivation

- Current rising population results in an increase in the number of vehicles. A higher number of vehicles results in the following issues:
  - Heavy traffic
  - Heavy consumption of oil and fuel resources
  - Large carbon emissions
  - Decreased air quality
  - Affects human health and other living beings on the planet
  - Overall results in Global Warming, profoundly affecting the environment



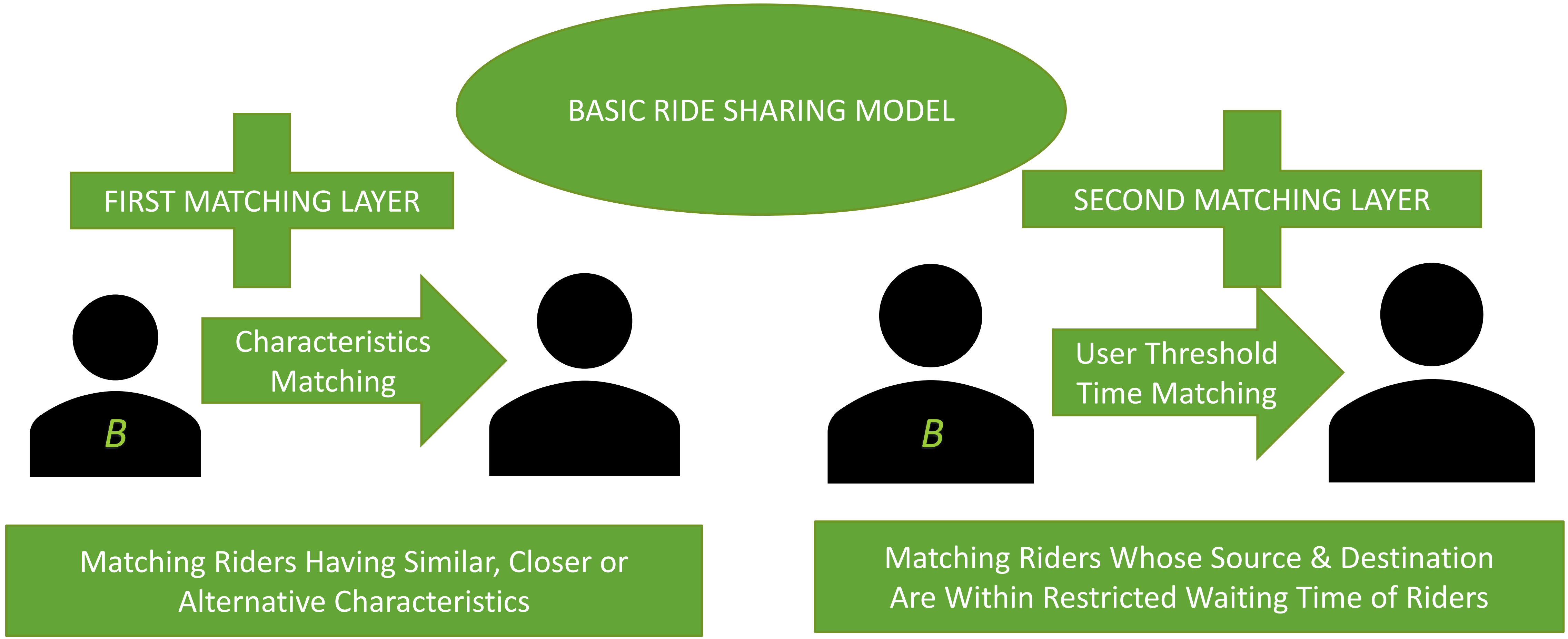
Basic Ride Sharing Model

DEFINITION - RIDERS TRAVEL THROUGH A COMMON PATH TO REACH THE SAME OR NEARBY DESTINATION.

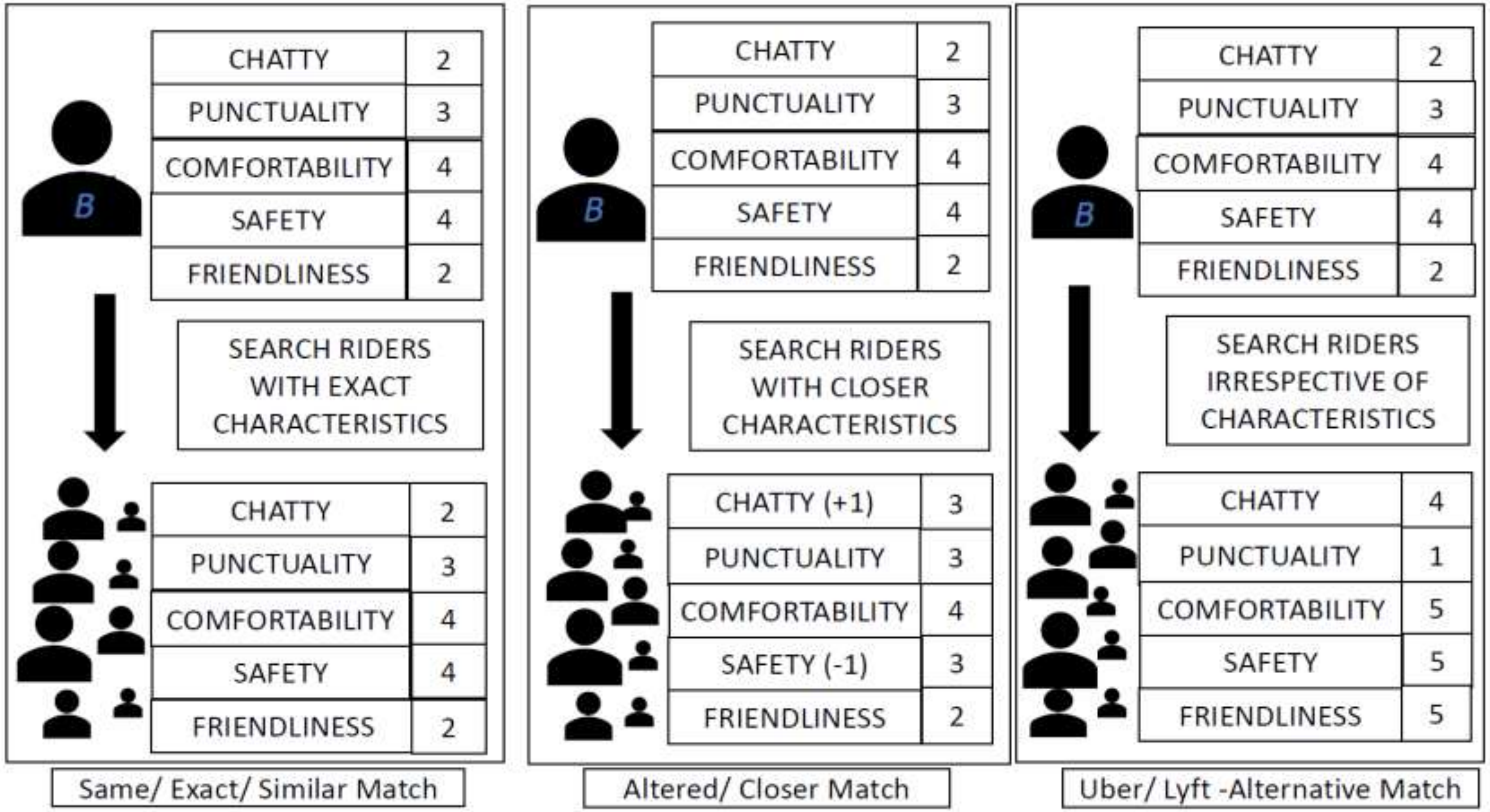
Limitations in Existing Ride Sharing Applications

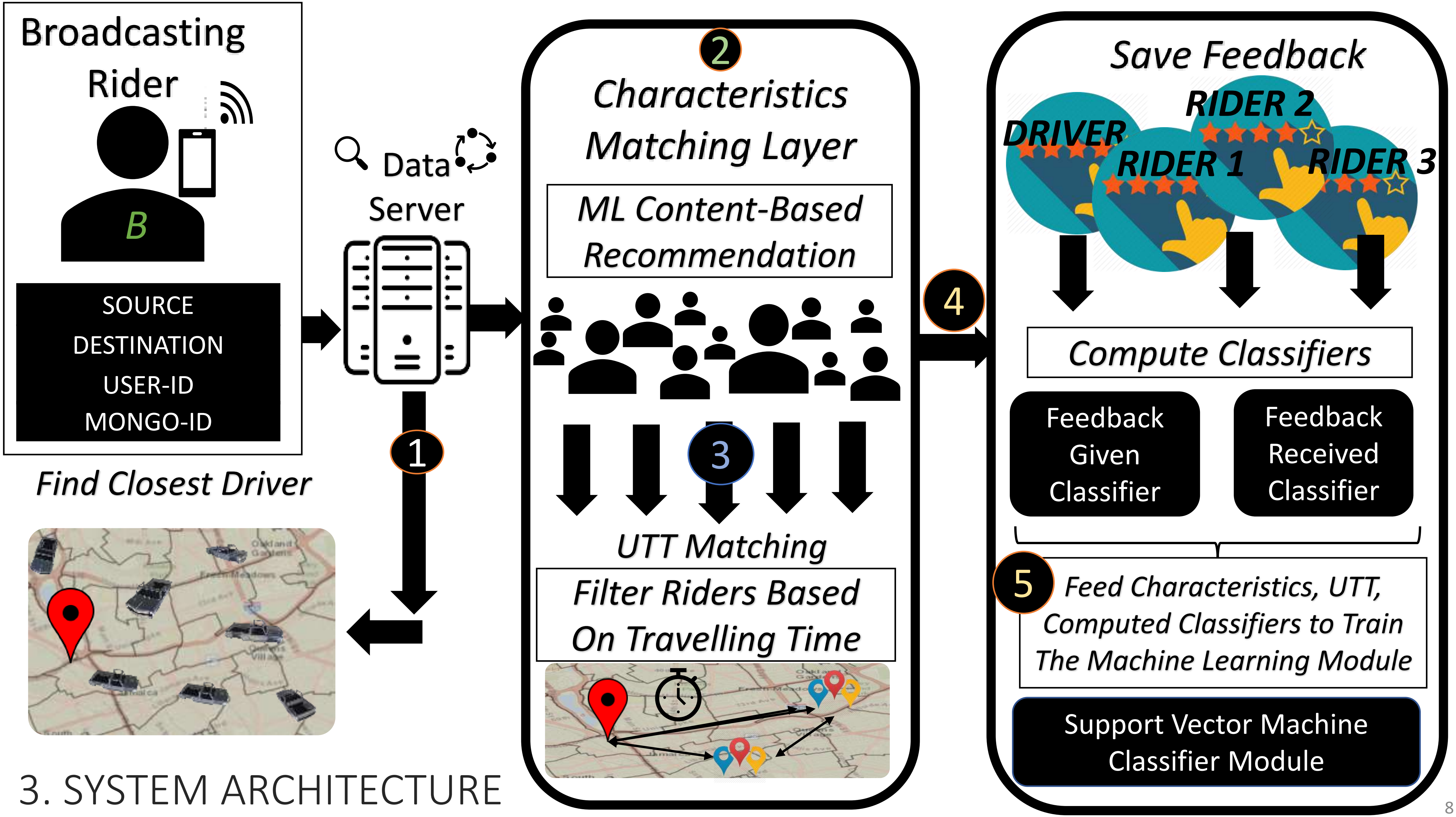
- Ride Sharing only efficient when the pool of the trip is completed.
- Car-Pooling discouraged due to social barriers.
- Sudden elongation of trips due to unexpected addition of riders.
- Absence of the rider-to-rider feedback system.
- Unfair pricing or billing models.

2. Enhanced Ride Sharing Model (ERSM)



Introduction to Characteristics

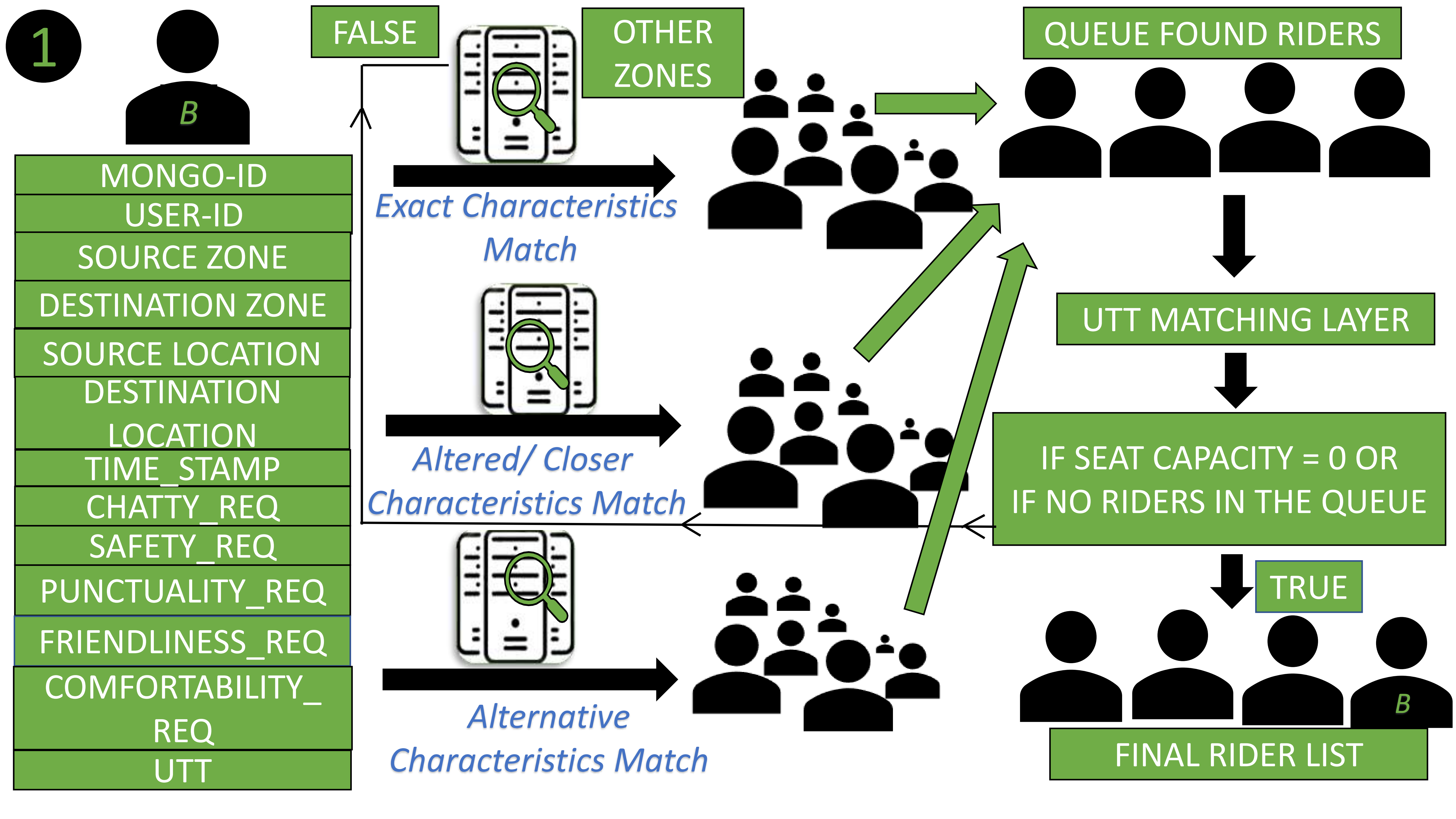




3. SYSTEM ARCHITECTURE

4. The Proposed Model

THE CHARACTERISTICS MATCHING


1
B
MONGO-ID
USER-ID
SOURCE ZONE
DESTINATION ZONE
SOURCE LOCATION
DESTINATION LOCATION
TIME_STAMP
CHATTY_REQ
SAFETY_REQ
PUNCTUALITY_REQ
FRIENDLINESS_REQ
COMFORTABILITY_REQ
UTT
FALSE
OTHER ZONES
Exact Characteristics Match
Altered/ Closer Characteristics Match
Alternative Characteristics Match
QUEUE FOUND RIDERS
UTT MATCHING LAYER
IF SEAT CAPACITY = 0 OR
IF NO RIDERS IN THE QUEUE
TRUE
B
FINAL RIDER LIST

Machine Learning Recommendation System

Broadcasting

Registered Rider

Characteristics

B

CHATTY: 3
SAFETY: 4
PUNCTUALITY: 3
FRIENDLINESS: 3
COMFORTABILITY : 4

[chatty, safety, punctuality, friendliness, comfortability]

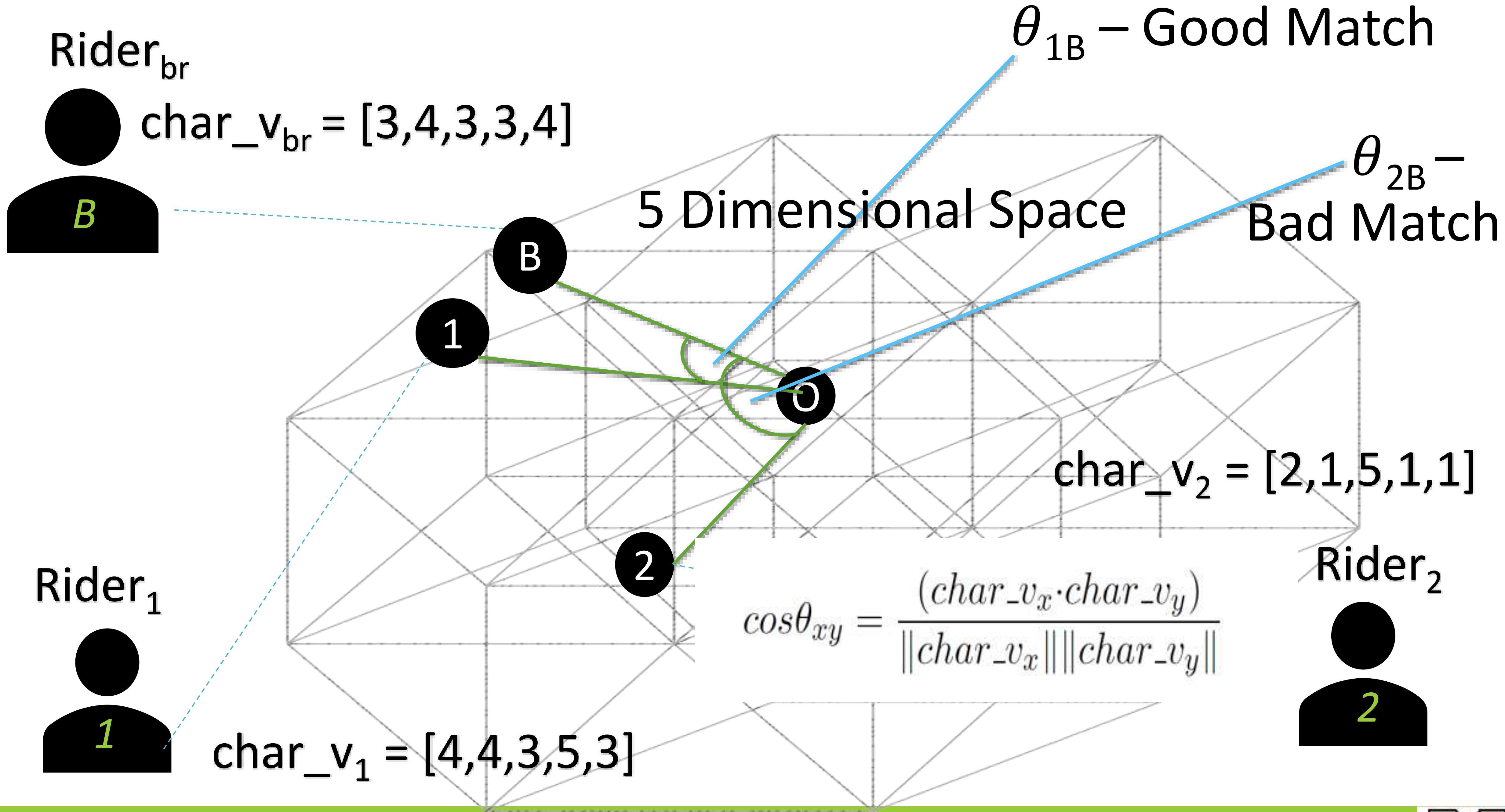
$char\_v_{br} = [3,4,3,3,4]$

1

$char\_v_1 = [4,4,3,5,3]$

2

$char\_v_2 = [2,1,5,1,1]$

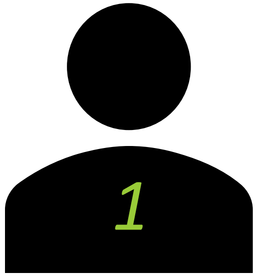


$$cos\theta_{xy} = \frac{(char\_v_x \cdot char\_v_y)}{||char\_v_x||\,||char\_v_y||}$$

5. The Feedback System and the Machine Learning Models

The Rider Feedback System

- The feedback system is designed for tracking the rider characteristics and generation of classifiers.
- The feedback consists of rating the drivers plus riders in terms of the five characteristics.



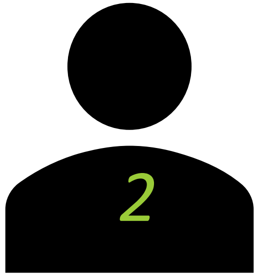
$\text{chatty}_{\text{Rider12}}$: 5

$\text{safety}_{\text{Rider12}}$: 3

$\text{punctuality}_{\text{rider12}}$: 3

$\text{friendliness}_{\text{Rider12}}$: 0

$\text{comfortability}_{\text{Rider12}}$: 0



Computing Feedback Based Classifiers

- The search criteria for the users is redefined using the computed classifiers.
- Classifiers are computed using the equation for variance.

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x - x_i)^2}{N}$$

$L_1 = [1, 0, 5, 4, 0]$ Variance of L1 = 5.5

$L_2 = [0, 0, 0, 0, 2]$ Variance of L2 = 0.8

$L_3 = [4, 4, 4, 4, 4]$ Variance of L3 = 0.0

The Feedback-Given-Classifier

Let the feedback given by $Rider_1$ to $Rider_2$, $Rider_3$, and $Rider_4$ be as follows:

Riders	Chatty	Safety	Punctuality	Friendliness	Comfortability
$Rider_2$	0	2	1	4	0
$Rider_3$	0	3	0	4	0
$Rider_4$	1	5	0	4	0

- Generate Sample sets for every characteristic and compute variance for $Rider_1$:

$chatty_{Rider1}$ = [0,0,1] $safety_{Rider1}$ = [2,3,5] $punctuality_{Rider1}$ = [1,0,0]

$friendliness_{Rider1}$ = [4,4,4] $comfortability_{Rider1}$ = [0,0,0].

- Feedback-Given-Classifier = (In this example) safety class

The Feedback-Received-Classifier

Let the feedback provided to $Rider_1$ by $Rider_2$, $Rider_3$, and $Rider_4$ be as follows:

Riders	Chatty	Safety	Punctuality	Friendliness	Comfortability
$Rider_2$	4*0.32	2*4.31	0*2.10	2*0.1	4*1.73
$Rider_3$	3*3.45	1*0.15	1*0.55	0*5.72	3*3.34
$Rider_4$	3*9.21	0*3.21	3*0.02	0*0.21	0*1.32
$\sum$ Total	39.26	8.77	0.61	0.2	16.92

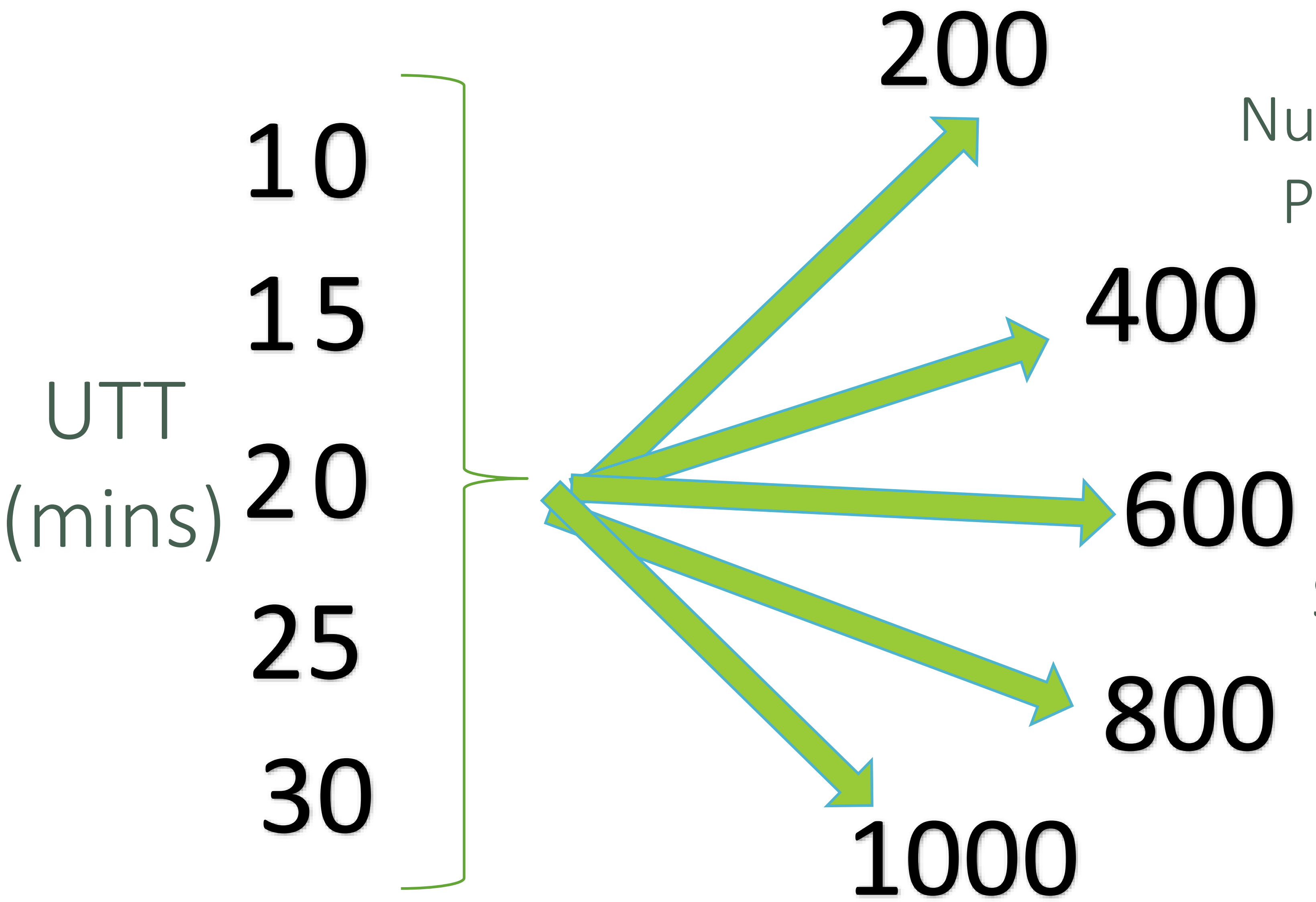
- Initially, fetch every characteristic variance of every rider.
- Multiply by the fetched variance by respective rated value.
- Integrate all ratings characteristic wise.
- Feedback-Received-Classifier = (In this example) chatty class for $Rider_1$

The Support Vector Machines (SVM)

- The function of the SVM is Classifiers prediction.
- Input to the SVM are the registered characteristics and UTT.
- The output is the computed classifier.
- For two classifiers, we have 2 distinct SVM modules.
- The prediction by the SVMs marks the last step of the proposed architecture.

Feedback-Given-Classifier Data-set						
Class_Given	**Chatty**	**Safety**	**Punctuality**	**Friendliness**	**Comfortability**	**UTT**
Comfortability	3	5	4	1	4	20
Chatty	1	2	4	3	5	10

6. Experimentations

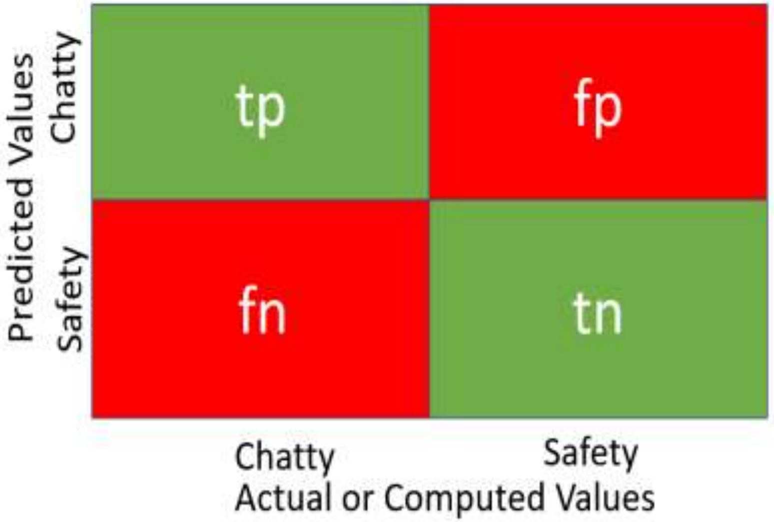


Number of Riders
Per Simulation

Simulations Performed
10 Times – Phase 1
5 Times – Phase 2

7. Results

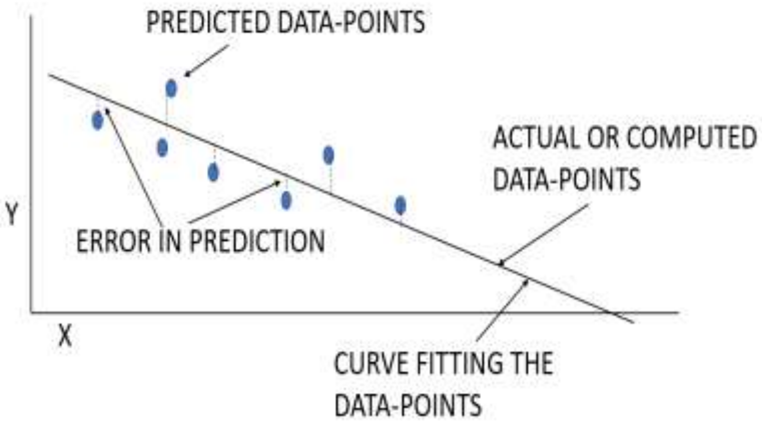
Performance Measures of a Machine Learning Classification Model



$$model\_accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

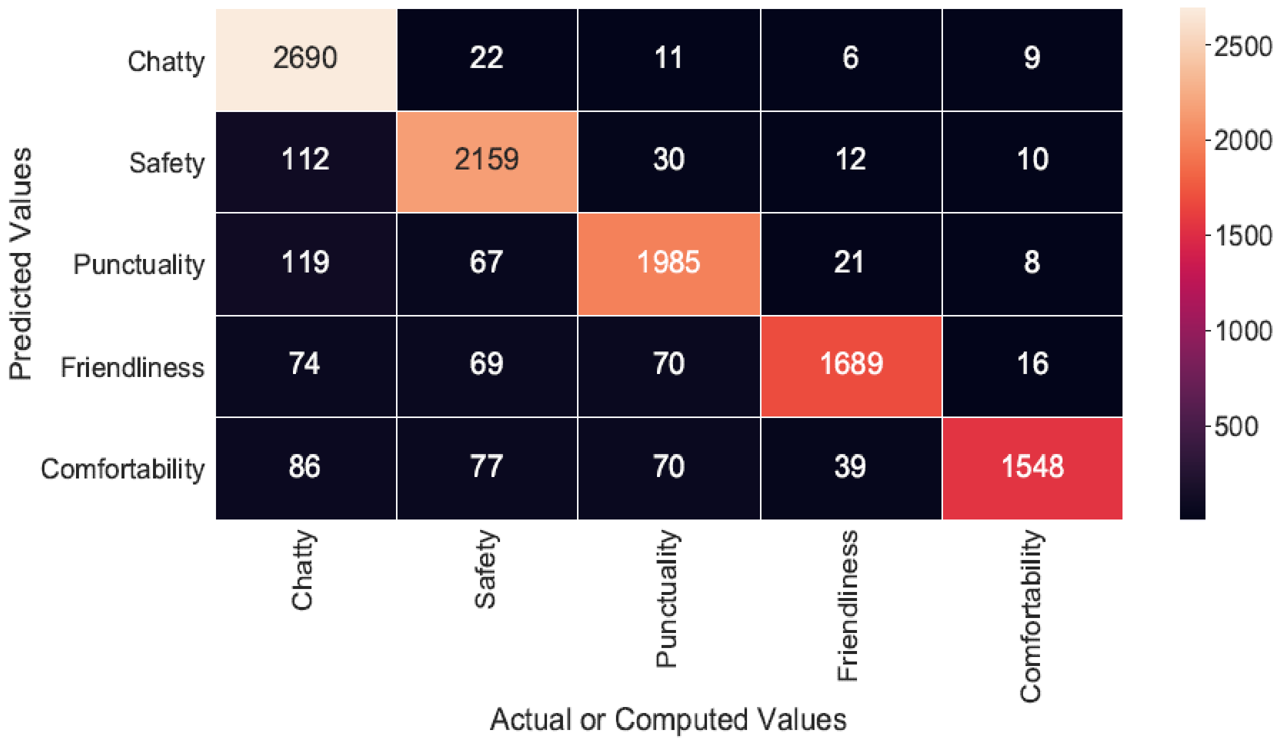
$$model\_precision = \frac{tp}{tp + fp} \qquad model\_recall = \frac{tp}{tp + fn}$$

$$F1\_Score = 2 * \frac{recall * precision}{recall + precision}$$



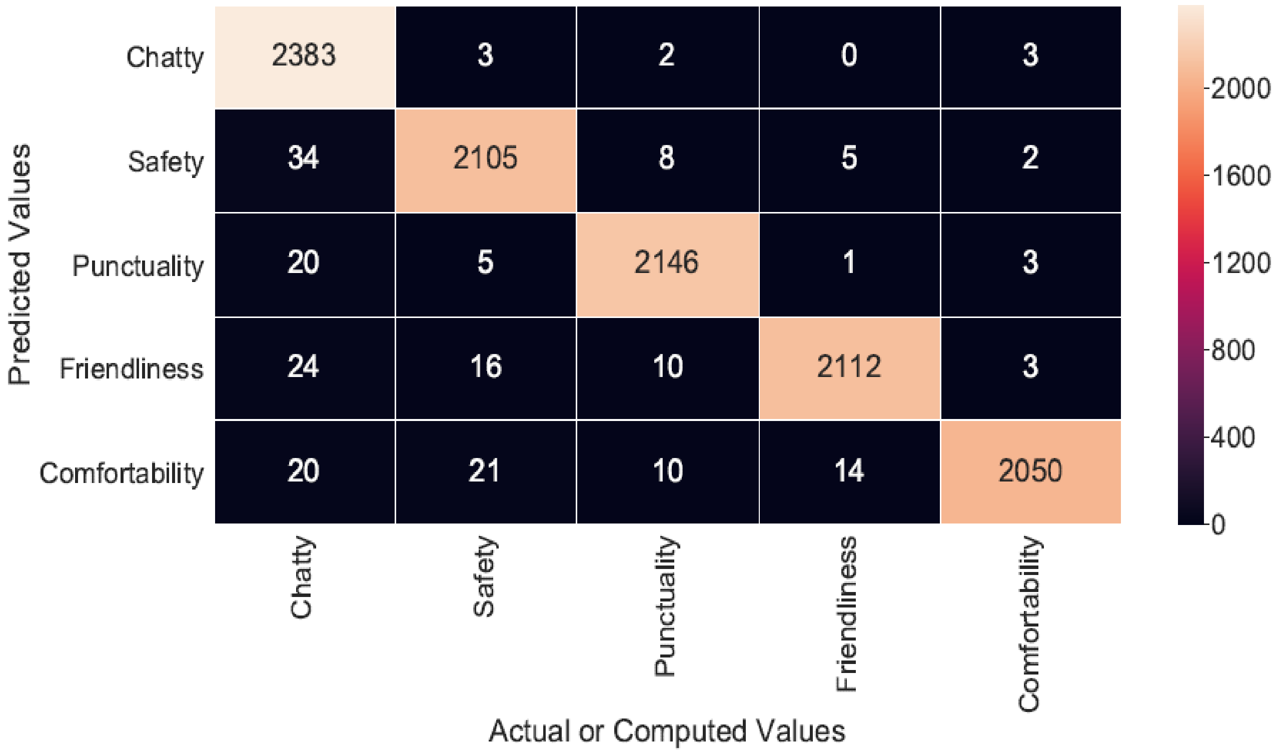
$$RMSE = \sqrt{\frac{(error)^2}{total\_sample\_points}} = \sqrt{\frac{(y_{computed} - y_{predicted})^2}{total\_sample\_points}}$$

Performance Measures of Feedback-Given-Classifier SVM



Overall SVM Accuracy: 91.65%					
Root Mean Square Error: 0.64					
Accuracy Measure By Class					
Measurement(%)	Chatty	Safety	Punctuality	Friendliness	Comfortability
F1 Score	92.34	91.65	91.07	91.92	90.90
Precision	87.04	90.40	91.97	95.84	97.35
Recall	98.31	92.94	90.20	88.32	85.25

Performance Measures of Feedback-Received-Classifier SVM



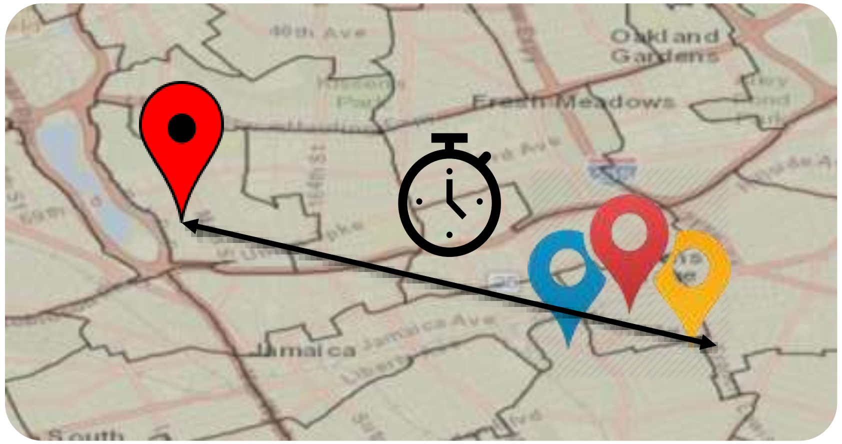
Overall SVM Accuracy: 91.33%					
Root Mean Square Error: 0.42					
Accuracy Measure By Class					
Measurement(%)	Chatty	Safety	Punctuality	Friendliness	Comfortability
F1 Score	87.85	89.02	90.63	93.22	93.21
Precision	86.13	87.52	92.58	91.97	95.48
Recall	89.21	88.82	89.67	94.49	96.96

Observations

Total Riders Traversed in

Complete Simulation

Phase1: 276400 | Phase 2: 90800



Average Trip Formation Time (mins)

Phase 1: 0.80 | Phase 2: 1.02

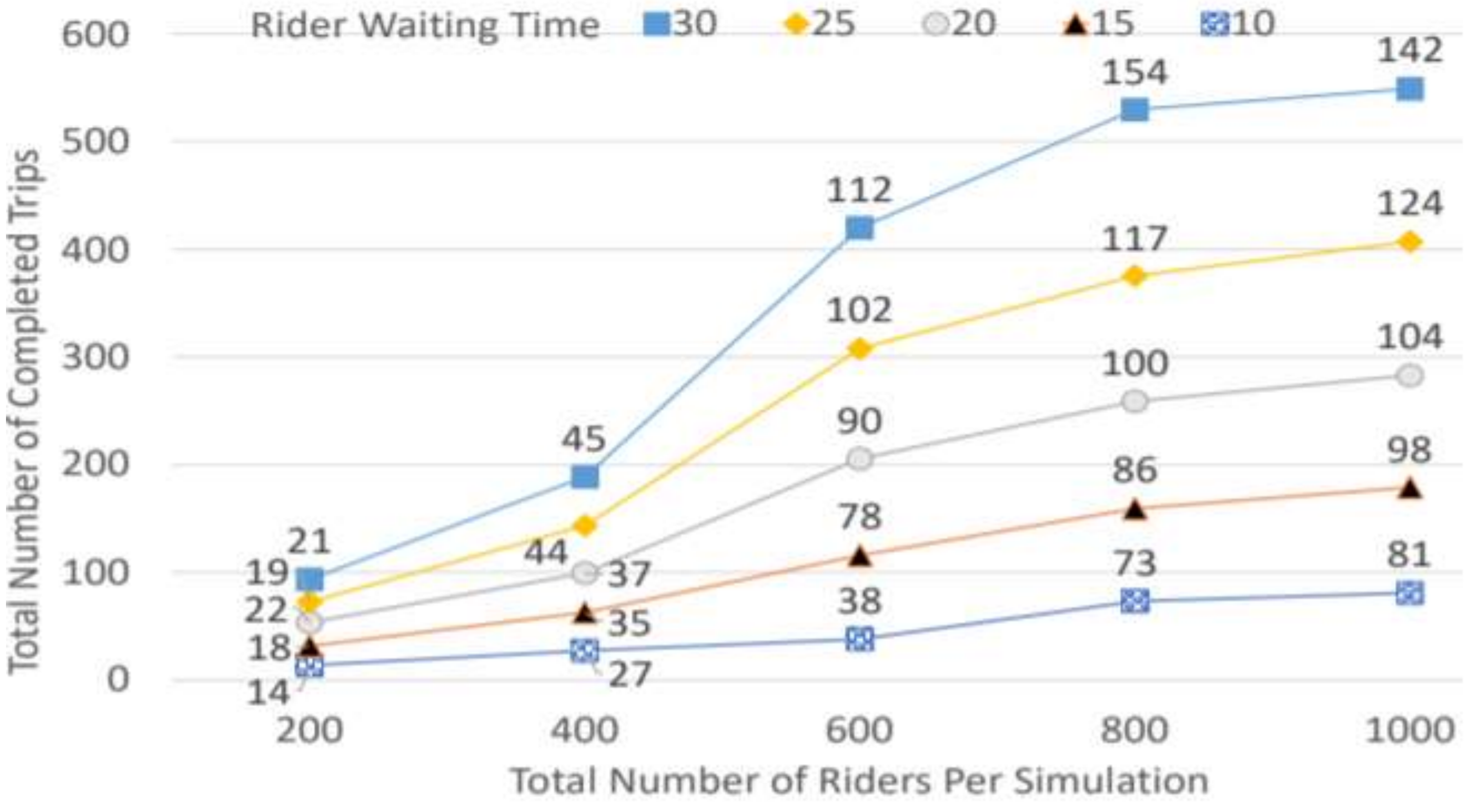
Total Number of Computed Trips

Phase 1 : 7159 | Phase 2: 10921

TOTAL NUMBER OF COMPLETED TRIPS

Objective: Observe the effects on the completed trips.

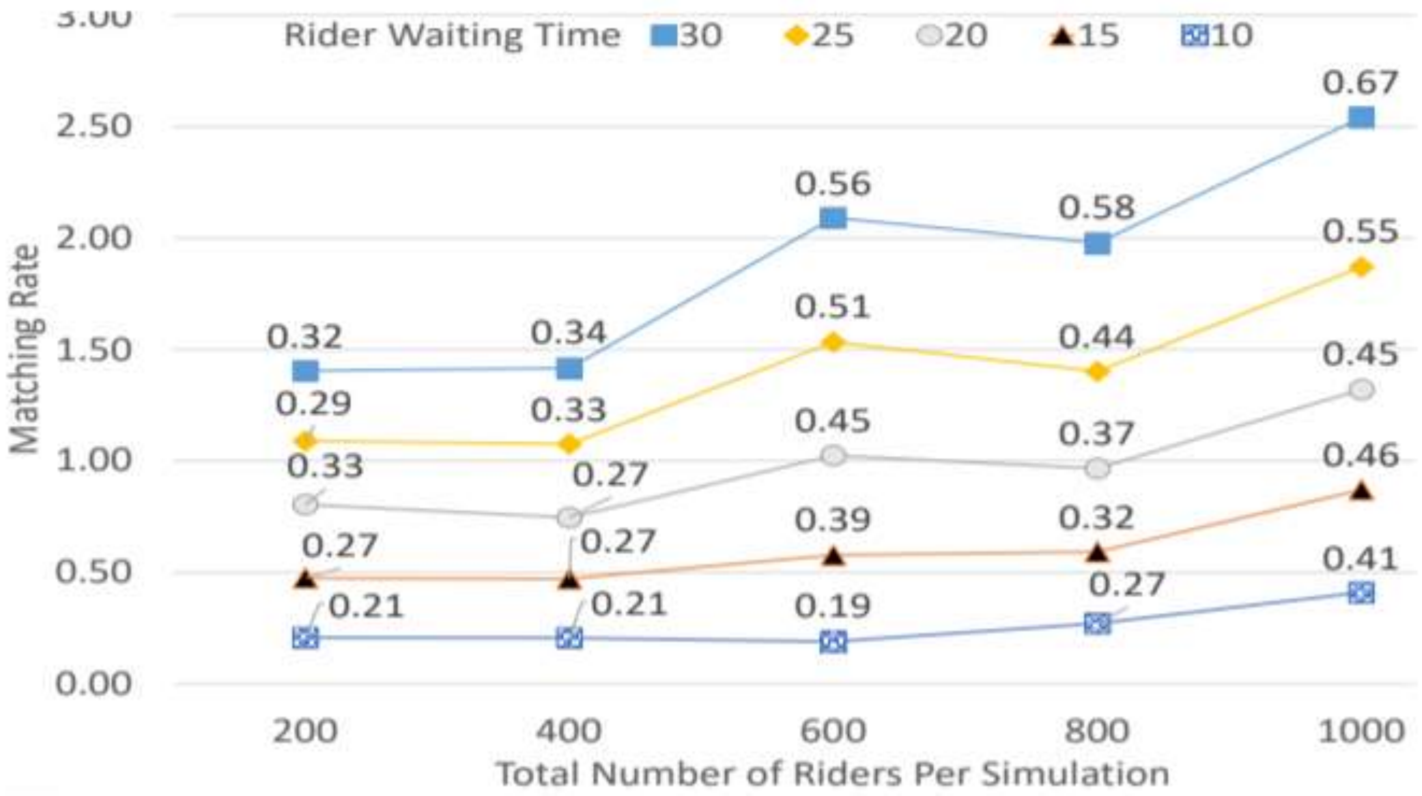
Results: The number of completed trips increases as the number of riders increases.



NUMBER OF MATCHES BY MATCHING TYPE

Objective: Observe the effects on number of rider matches by the characteristics matching types.

Results: High percentage of matching achieved for Exact or Closer characteristics matching.



8. Conclusion

We implemented the proposed Enhanced Ride Sharing Model based on rider characteristics addressing the current user expectations and discovered issues in the existing systems.

The average trip formation time in both phases rounds up to a minute, which promotes in providing a timely response to the passengers.

The goal of the pool completion for a maximum number of trips achieved. The goal of pairing maximum riders with similar characteristics achieved in Phase 2.

Machine Learning SVM modules run with an accuracy of 90% and provides a quality prediction of classifiers. Also, the recommendation system eliminates large computations and assists in tuning up the model performance during matching of riders.

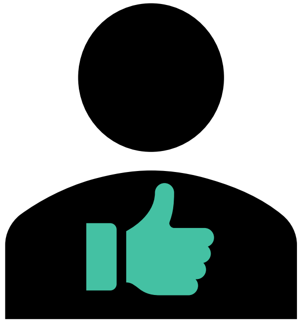
The overall system efficiency is tested by subjecting the model to an extensive simulation. The parameters, matching rate, completed trip count and trip simulation time keeps increasing with the increasing number of riders, which proves that the model performance is consistent as the rider count keeps scaling up.

Shortcomings

1. The limitation of zones – The Ride Sharing model currently performs matching on the basis of zones
2. The limitations of Google Map Keys – System ceases to function if a Google Map API Key is completely utilized.
3. Allocation a rider with Exact characteristics for every trip is difficult.

Future Work

Mobile Application as an User Interface



Recommend “Favorites” in Future Trips

Virtual “Badges” in Form of Points

A Sophisticated Billing Model for Handling Transactions

An Enhanced Ride Sharing Model Based on Human Characteristics and Machine Learning Recommender System

THANK YOU

Q&A

Govind Yatnalkar, Husnu S. Narman, Haroon Malik

The 3rd International Conference on Emerging Data and Industry 4.0 (EDI40)
April 6 - 9, 2020, Warsaw, Poland