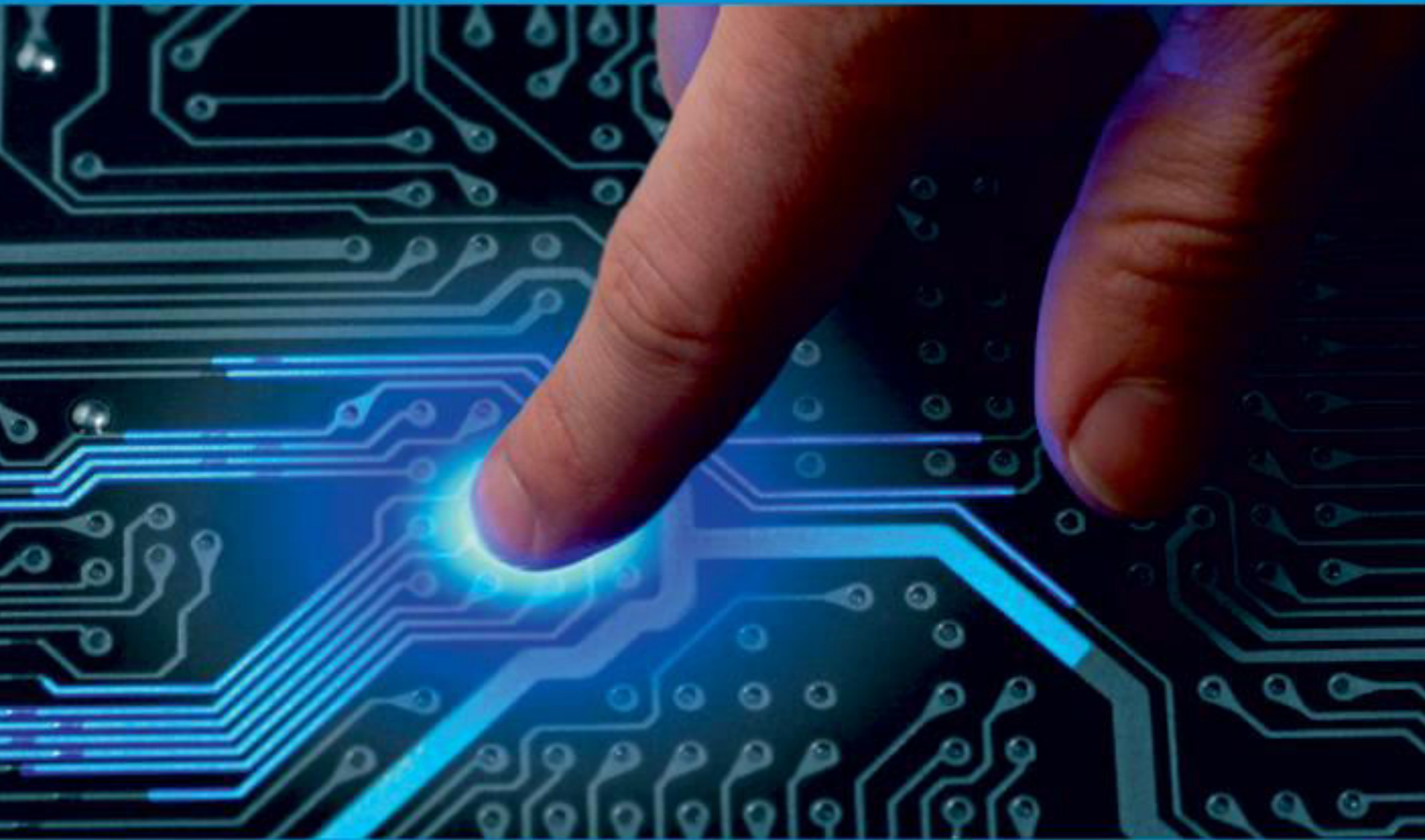
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 8, August 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008011 |

Impact of Hofstede's Cultural Dimensions on Intelligent Ethical Agent

Shreyas Bangalore Suresh, Tarun Raghunandan Kaushik, Varun Tarikere Shankarappa, Dr.Nagegowda KS

Graduate, Department of Computer Science, PES University, Bengaluru, India

Graduate, Department of Computer Science, PES University, Bengaluru, India

Graduate, Department of Computer Science, PES University, Bengaluru, India

Associate Professor, Department of Computer Science, PES University, Bengaluru, India

ABSTRACT: Since the advent of Artificial Intelligence, many autonomous machines are making their way into the society. With the burgeoning development of autonomous systems like self-driving cars have come concerns about how machines will make moral decisions and thus a new field called Machine Ethics has emerged. Machine ethics deals with moral dilemmas in machines while interacting with humans, or possibly other machines as well, and ensures the decisions taken by the algorithm are morally acceptable. This is in contrast to computer ethics, which solely focuses on ethical problems and protocol surrounding humans' use of technology. In this article, we have explored the moral dilemmas faced by autonomous vehicles and have tried to train an artificial intelligence model that makes ethically acceptable decisions based on the data collected by the famous moral machine experiment. Here, we describe the results obtained from the model. Firstly, we summarize the accuracies obtained upon training multiple models with different techniques. Later, we document the variation of accuraciesin the model upon using the Hofstede model of six dimensions of national cultures as a factor when pre-processing the data.

KEYWORDS: Artificial Intelligence;Moral Machine; Ethics; AI ethics;Machine Ethics; Hofstede's Cultural Dimensions; Autonomous Vehicles; Ethical Agent; Support Vector Machine; Logistic Regressor; Neural Networks; K-fold cross validation;

I. INTRODUCTION

The rise of Artificial Intelligence in recent history is being hailed as the Fourth Industrial Revolution as its applications are increasing in many fields from Industrial robots to personal assistants. The machines or robots which were previously 'dumb' are being made smarter by AI. Although these machines are a product of perfectionism, they have to work in an imperfect world of humans where there is always room for error. Many have criticized these machines because of their capabilities and whether their choices in cases of accidents or unprecedented situations are ethical. This has given birth to new field of ethics namely, Machine Ethics. Lots of research and survey has been done on moral dilemma. Multiple models have been proposed. However, each agent is trained by various survey datasets that might be biased because of a few cultural dimensions. This article presents our results of creating an Intelligent Ethical Agent based on one such experimental data. We try to enhance our results by clustering our data based on Hofstede's six cultural dimensions and discuss on our observations.

II. LITERATURE SURVEY

The roots of ethical dilemma in the automotive industry can be traced backed to the trolley problem [1], where there is a conundrum of choosing between saving a life or multiple lives. This problem is also an ideological clash between schools of utilitarianism and deontological ethics. Prior to the 21st century, machine ethics had been the subject of science fiction and popular movie culture due to the limitation of hardware and AI. The Term Machine Ethics was coined by Mitchell Waldrop in 1987 in his article [2] where he proposes application in practice of the fictional three laws of robotics of Asimov. After the official theoretical foundations were laid out in 2004 in AAAI Workshop, A breakthrough in the field was the Machine Ethics: Creating an Ethical Intelligent Agent [3] article in 2007 where the

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008011 |

importance of machine ethics, the idea of explicit ethical machines, challenges in the field and possible steps to create an ethical agent were proposed and discussed. J. Goodall in his article [4] has clearly documented the criticism stating the need for ethical agents in the Automated Vehicle industry where he stresses why humans have to implement such agents in an automated vehicle, one such being the machines can't be held accountable for any accident where lives of people are involved.

Whereas many have proposed an AI solution to the problem caused by the AI, one article [5] has taken the path of logical reasoning and a set of rules i.e. a rational agent to the solve problem of ethical dilemma where one such choice is not available. Questions were raised about ethical agents deciding the fate of human life without the consent of society since an experiment [6] was conducted where it depicted many ethical dilemmas in the field of automated vehicles to gather the people's decisions in such scenarios by million respondents in multiple countries and culture. Using the data of the survey many ML models for Ethical agents have been proposed. Hofstede's Cultural Dimensions [7] have been a guiding light for many psychological types of research between cultures. We use the same 6 dimensions for creating our ML-based Ethical agent and observe the results on how culture influences the ethical choices of people.

III. METHODOLOGY

To create our ethical agents [shown in Fig. 4], we used the following steps.

A. Data Collection

a. Moral Machine Experiment

The Data [8] to train our ethical agent is extracted from Moral Machine Experiment [9]. It consists of a scenario involving an ethical dilemma where the respondent has to make a choice between two options. These scenarios have a range of contexts from age to social status. One such scenario is depicted in Fig.1 where the respondent has to choose between passengers and pedestrians. Here the context includes age, lawfulness, and the number of people.

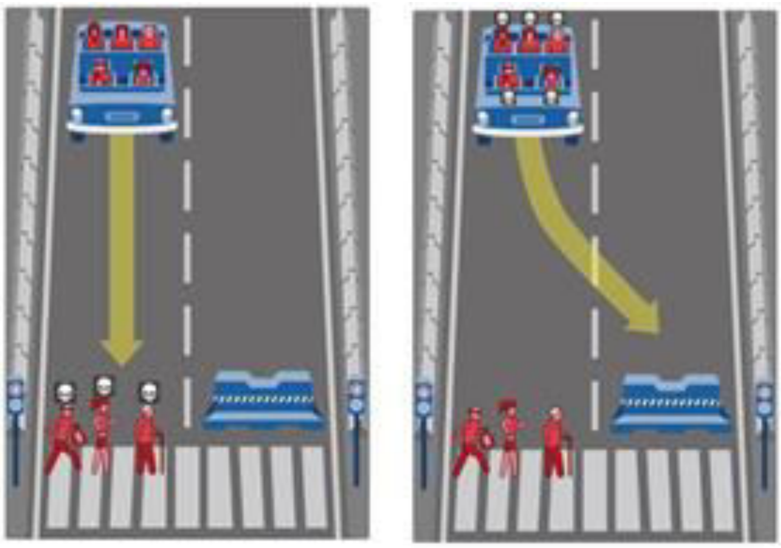


Fig.1. Moral Machine Experiment sample scenario

b. Hofstede's Cultural Dimensions

Hofstede's cultural dimensiondata [10] consist of countries' Power Distance Index (high versus low), Individualism Versus Collectivism, Masculinity Versus Femininity, Uncertainty Avoidance Index (high versus low), Long- Versus Short-Term Orientation, Indulgence Versus Restraint. Each of these parameters gives a relative insight into countries' social and economic cultural behavior. Countries with comparable scores are said to have similar cultural behavior and

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008011 |

take akin decisions. Hofstede's six cultural dimensions' data set was available on a few websites and many missing countries' data had to be extracted from its official site [11]. The official data from the site was scrapped using Selenium web drivers and the data was accumulated as needed.

B. Pre-Processing

The data from the Moral Machine Experiment had a lot of unprocessed junk and had to be excluded. And many textual data had to be quantified so these variables can be ML-model ready. Countries common in both the Hofstede data set and the Moral Machine data set had to be considered for further part of the algorithm. So a join between the two data sets was performed keeping the country or their ISO 3 country code as keys. Many countries had multiple names and there was a mismatch between the two data sets. So, a manual correction between the country names and ISO 3 country code had to be done before performing join operations.

C. Clustering

The Cultural dimensional Hofstede pattern data are taken as input to this module. Each of the countries is grouped together based on Hofstede's six cultural dimensional scores. The algorithm used to cluster the countries is the K-Means clustering algorithm.

a. Elbow Method

Applying the elbow method [12], we found the most optimal number of clusters to be 6. Used two metrics to plot the elbow curve distortion [shown in Fig. 2], and inertia [shown in Fig. 3]. Distortion is the average of the Euclidean squared distance from each cluster's centroid. Inertia is the sum of squared distances of samples to their closest cluster center [8]. In the elbow method, after plotting the values 1 to 20, the point of intersection of the diagonal tangent with the curve was found at value 6 and it is said to be the point of optimality. Both the plots pointed to an optimum number of clusters being 6 for the Hofstede country vectors.

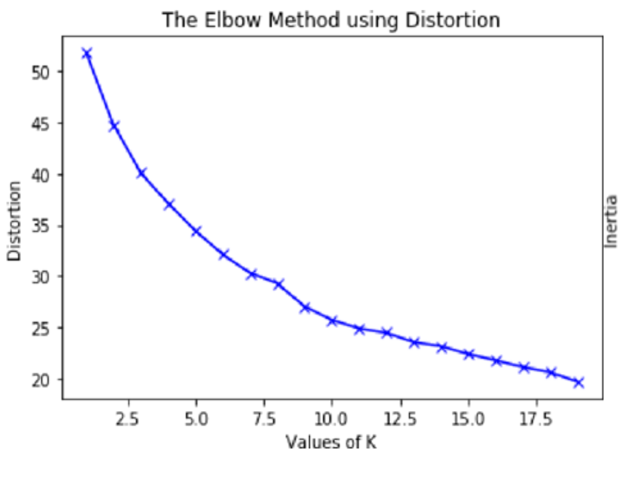


Fig. 2. Elbow plot using distortion

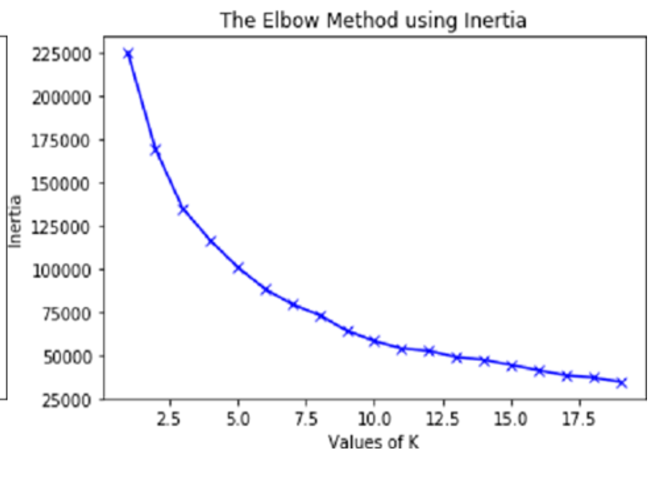


Fig.3. Elbow plot using inertia

b. K-Means Clustering

K-means clustering is a type of unsupervised learning, which is used to find groups that have not been explicitly labeled in the data. The algorithm works iteratively to assign each data point to one of the K groups based on the features that are provided. Data points are clustered based on feature similarity. K-means chooses the initial centroid point randomly, each centroid of a cluster is a collection of feature values that define the resulting groups. Examining the centroid feature weights can be used to qualitatively interpret what kind of group each cluster represents [13]. The centroids of the K clusters are used to label new data. Running the algorithm multiple times with a constant k value set to 6 and taking the minimum inertia out of the results will yield a better-clustered result. Countries with similar Hofstede scores were clustered into 6 groups by executing K means algorithm 100 times. The clusters obtained with the

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165
|| Volume 10, Issue 8, August 2022 ||
| DOI: 10.15680/IJIRCCE.2022.1008011 |

merest inertia were the final list of clusters obtained from the algorithm. The Moral Machine data set was categorized based on these clusters and used in the further machine learning models.

D. ML Models

An ethical agent was built to take decisions analogous to ethical decisions taken by humans in the Moral Machine experiment. Various machine learning models were used to make the ethical agent. Usage of multiple machine learning models circumvents bias produced by any single machine learning model. Each of the ML models has been trained over the universal Moral Machine data sets. Later the same experiment is repeated with clustered Moral Machine data. Usage of the following algorithms produces a better result for binary output datasets like Moral Machine data sets.

a. Logistic Regression

The first choice of machine learning algorithms is a logistic regressor because the regression model is suitable for estimating based on parameters and is a binary classifier (decisions are binary). The Moral Machine data consists of a multitude of parameters and a regressor-based logistic regressor performs well for such data. The results also help us as a benchmark and for verification.

b. Artificial Neural Networks

A vanilla ANN was trained with a variable number of hidden layers to verify whether there exists a non-linear relation between the features. ANN with 3 hidden layers with 50 neurons each with ‘RELU’ activation- running 100 epochs gave a better result with faster convergence.

c. Support Vector Machine

The last of the machine learning models for ethical agents is an SVM. The SVMs are experimentally proven to work better than other models when data is high dimensional and sparse. SVM with the ‘RBF’ kernel gave better results compared to SVMs with other kernels. Consequently, ‘RBF’ kernel SVM with auto gamma factor was used to build the ethical agent.

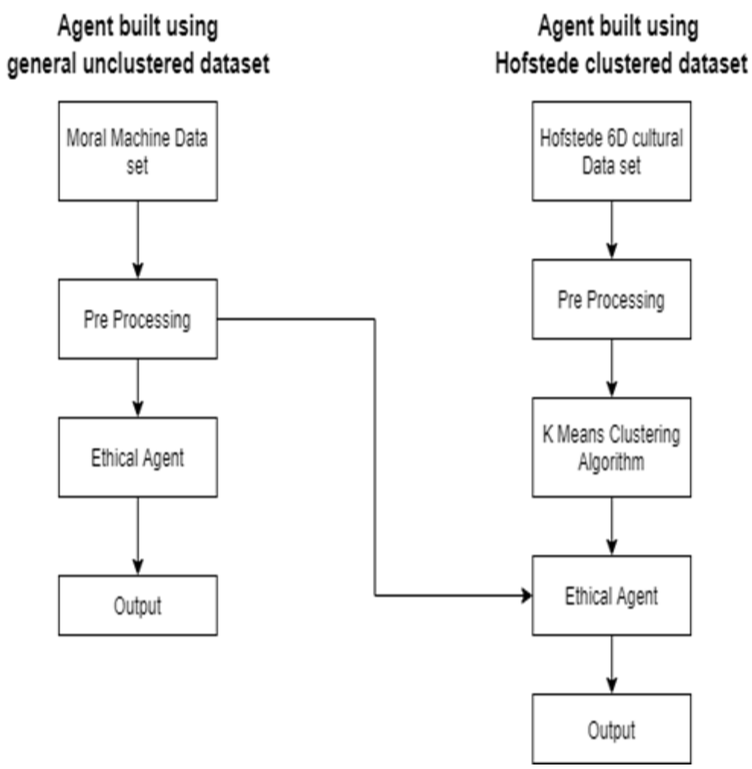


Fig. 4. Ethical agent flow diagram

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008011 |

IV. EVALUATIONS

For evaluating the models without any clustering we split the data into training and testing data with a 90-10 split and used 90% of the data for training the models and the remaining 10% for testing the data. This provided a wide variance in accuracies every time the model ran especially for some of the models. Thus we used the k-foldcross validation algorithm [14] to split the data into equal parts and use one of the parts for testing while the remaining parts are used for training the models. We used k=5 in our runs where we split the data into training and testing data in an 80-20 ratio and used 80% of the data for training the models and the remaining 20% for testing the data. After re-training the models with each of the different portions of the k-fold split being used as testing data, the average of all the accuracies was taken. Later individual cluster and overall accuracies was compared with the un-clustered data accuracy [as shown in Fig. 5].

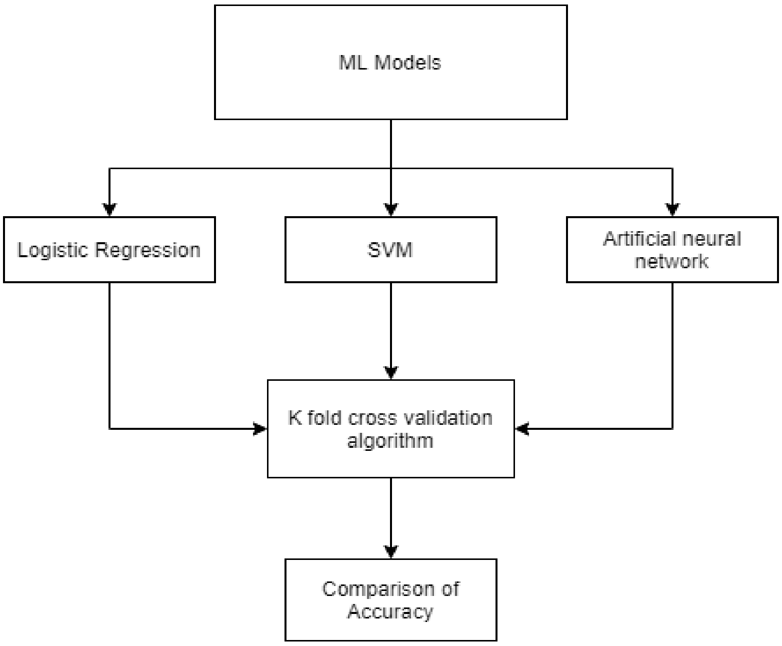


Fig. 5. Ethical evaluation flow diagram

V. RESULT

The results obtained from testing our models is documented in the below table.

Type of data	Accuracy		
	Logistic Regression	ANN	SVM
Cluster 1	63.50%	60.23%	70.17%
Cluster 2	63.67%	60.48%	69.93%
Cluster 3	63.68%	58.97%	71.24%
Cluster 4	64.29%	58.74%	72.21%
Cluster 5	65.02%	60.52%	71.87%
Cluster 6	63.71%	59.43%	70.33%
Average of 6 clusters	**63.98%**	**59.73%**	**70.96%**
Un-clustered	**61.67%**	**58.00%**	**67.80%**

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008011 |

VI. CONCLUSION AND FUTURE WORK

The concept of morality is not constant across the world. People in different parts of the world have decided differently on the moral machine experiment and neither one can be proved to be right or wrong. However, we have proved that this can also be factored in when designing an AI. We have used the Geert Hofstede scores that were established by a study a long time ago and proved that the machine is able to decide better when taking into account only the countries with closer Hofstede score vectors. The weightage of each dimension may vary across cultures and that can also be further studied. Also on the availability of larger diverse datasets along with advanced ML algorithms a better ethical agent can be modeled.

Out of the many challenges faced in the area of machine ethics, the foremost is the need for understanding and dialogue between ethicists, researchers, and the general public. People fear artificial intelligence and are threatened by the idea of them making ethical decisions, especially in autonomous systems such as self-driving cars. There is a need for people to understand that the artificial agent is trained using human decisions and data. Thus, the decision taken by the autonomous system should be a collective reflection of the decision of the general public. Doing so will bring confidence in humans additionally it persuades them to embrace and bolster the development ofintelligent machines to improve our lives.

REFERENCES

[1]J. J. Thomson, “The Trolley Problem,” The Yale Law Journal, vol. 94, no. 6, pp. 1395–1415, 1985, Accessed: Aug. 24, 2022. [Online]. Available: http://www.jstor.org/stable/796133.
[2]M. M. Waldrop, “A Question of Responsibility,” AI Magazine, vol. 8, no. 1, p. 28, 1987, doi: 10.1609/aimag.v8i1.572.
[3]M. Anderson and S. L. Anderson, “Machine Ethics: Creating an Ethical Intelligent Agent,” AI Magazine, vol. 28, no. 4, p. 15, 2007, doi: 10.1609/aimag.v28i4.2065.
[4]N. J. Goodall, “Machine Ethics and Automated Vehicles,” in Road Vehicle Automation, Springer International Publishing, 2014, pp. 93–102. doi: 10.1007/978-3-319-05990-7_9.
[5]L. Dennis, M. Fisher, M. Slavkovik, and M. Webster, “Formal verification of ethical choices in autonomous systems”Robotics and Autonomous Systems, vol. 77, pp. 1–14, 2016, doi: https://doi.org/10.1016/j.robot.2015.11.012.
[6]E. Awad et al., “The Moral Machine experiment,” Nature, vol. 563, no. 7729, pp. 59–64, Nov. 2018, doi: 10.1038/s41586-018-0637-6.
[7]G. Hofstede, “Dimensionalizing Cultures: The Hofstede Model in Context,” International Journal of Behavioral Medicine - INT J BEHAVIORAL MEDICINE, vol. 2, Jan. 2007, doi: 10.9707/2307-0919.1014.
[8]E. Awad, “Moral machine,” OSF, 22-Jul-2018. [Online]. Available: https://goo.gl/JXRrBP.
[9] “Moral machine,” Moral Machine. [Online]. Available: https://www.moralmachine.net/.
[10] “Dimension data matrix,” Geert Hofstede. [Online]. Available: https://geerthofstede.com/research-and-vsm/dimension-data-matrix.
[11] “Country comparison,” Hofstede Insights. [Online]. Available: https://www.hofstede-insights.com/country-comparison/.
[12] A. Gupta, “Elbow method for optimal value of k in Kmeans,” GeeksforGeeks. [Online]. Available: https://www.geeksforgeeks.org/elbow-method-for-optimal-value-of-k-in-kmeans/.
[13] A. Trevino, “Introduction to K-means Clustering,” Blogs.oracle.com. [Online]. Available: https://blogs.oracle.com/ai-and-datascience/post/introduction-to-k-means-clustering.
[14] J. D. Rodriguez, A. Perez, and J. A. Lozano, “Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 3, pp. 569–575, 2010, doi: 10.1109/TPAMI.2009.187.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details