CSE 643 Artificial Intelligence
Paper Critique

Kaustav Vats (2016048)

B.Tech, Computer Science and Engineering

Indraprastha Institute of Information
Technology

Okhla Industrial Estate, Phase III, New Delhi

I. PAPER INFORMATION

1) Title: A Voting-Based System for Ethical Decision Making [1]
2) Authors:
   - Ritesh Noothigattu
   - Snehalkumar Neil S. Gaikwad
   - Edmond Awad
   - Sohan Dsouza
   - Iyad Rahwan
   - Pradeep Ravikumar
   - Ariel D. Procaccia
3) Year Published: 2018
4) Conference Name: AAAI [2]
5) Link: https://arxiv.org/pdf/1709.06692.pdf

II. SUMMARY

This paper discussed the problem of ethical decision making in autonomous vehicle domain. An autonomous vehicle having brake failure including backup brakes, leading to inevitable death. Then a vehicle with enough computation power should be able to make an informed decision. Paper presents an approach to automating ethical decisions, using machine learning and computational social choices. Basically, a model based on the decision made by society on the same events. Paper provided an algorithm which evaluates the system for ethical decision making in the autonomous vehicle domain, using the data collected from 1.3 million people through Moral Machine website. Their approach to solving this problem consists of these steps:

- **Data:** For the set of questions asking human beings to make a decision that is least catastrophic. Where the choices consist of saving people of different age, gender, species etc.
- **Learning:** Pairwise comparison to learn a model based on the decision made by each voter for all possible alternatives.
- **Summarization:** Combining model of each voter into a single model.
- **Aggregation:** At runtime, applying a voting rule on decisions of each voter into a collective decision.

An autonomous vehicle can have infinite possibilities at runtime. Since the model has many features like age, species, passenger etc, it requires a finite set of options to choose from, these generalized models are called the permutation process. Paper used a voting rule to aggregate the preferences. It has defined a notion of swap dominance between alternative and shown that the permutation process satisfies the natural property. Their algorithm uses a specific permutation process called the Thurstone-Mosteller Process(TM Process) with a specific linear parametrization. TM is consistent permutation process with Gaussian utility process which gives a finite set of alternatives.

Paper also discussed different anonymous social choices and explained that there are many different social choice axioms that give completely different outcome for same preference profile. It also explained how they are applying in the computationally efficient way. In short, they explained that they need to explicitly construct an anonymous profile and then apply SCC to it. It has listed some classic social choice axioms considering that an Anonymous SCC f is monotonic if below condition hold:

- If $a \in f(\pi)$, and π' is obtained by pushing a upward in the ranking, then $a \in f(\pi')$.
- If $a \in f(\pi)$, and $b \notin f(\pi')$, and π' is obtained by pushing a upward in the rankings, then $b \notin f(\pi')$.

Paper has further proved various definitions and theorems used in the paper. After applying TM process and finding suitable computational parameter values, it has aggregated to a single TM process to work on the runtime. Paper has experimented using synthetic data and the Moral Machine dataset. For Synthetic data they have represented their preference of each voter in a TM process, they have computed anonymous preference profile of each voter and then taken a mean across all voters to obtain the desired profile. they estimated by sampling ranking for each voter.

Moral Machine Dataset Platform to gather data regarding the ethical decision taken of real-life situations. It lists every session user with some random questions. Each question has a few common things like its related to an autonomous car, brake failure in the car etc. every choice is given consider some features like relations, legality. explicit legal crossing etc. This dataset consists of about 1.3 million voters decision.

REFERENCES

[1] Paper used for the summary: https://arxiv.org/pdf/1709.06692.pdf
[2] https://aaai.org/Conferences/AAAI-18/wp-content/uploads/2017/12/AAAI-18-Accepted-Paper-List.Web.pdf