UNDER REVIEW

RESCIENCE C

Editorial

[RE] MEDIRL Maximum Entropy Deep Inverse Reinforcement Learning for Human Social Navigation

Unknown Individual[1,2] and Unknown Individual[1,2]

[1]Both authors were equal contributors – [2]Purdue University, Computer Science

Edited by
(Editor)

Reviewed by
(Reviewer 1)
(Reviewer 2)

Received
07 November 2023

Published
–

DOI
–

1 Reproducibility Summary

Our paper focuses on the reproducibility of studies within the domain of human-robot interaction (HRI) by revisiting and expanding upon the groundbreaking work of Muhammad Fahad, Zhuo Chen, and Yi Guo in their study on maximum entropy deep inverse reinforcement learning (MEDIRL) [1] for understanding human navigation behaviors in crowded environments. Our objective is to rigorously retest and augment their findings, emphasizing the need for robust and socially intelligent navigation systems in HRI scenarios.
Our re-experimentation process involves:

1. **Comprehensive Replication and Validation**: We aim to replicate the original methodology while conducting a thorough validation process, ensuring the reliability and applicability of the MEDIRL model in real-world HRI scenarios.
2. **In-Depth Component Analysis**: Our focus is on dissecting and analyzing the individual components of the MEDIRL model through ablation studies. These studies involve the selective removal or alteration of critical elements, such as learning rate, state dimensions, network layers, and the loss function, to understand their impact on the model's performance.
3. **Refinement and Enhancement**: We seek to refine the MEDIRL model by optimizing critical parameters, learning strategies, and eliminating biases. Our goal is to improve the model's robustness and adaptability, ensuring its deployment in diverse HRI scenarios while adhering to social norms and safety protocols.
4. **Deeper Insights**: The results of our ablation studies will provide deeper insights into the model's performance dynamics, shedding light on the intricate mechanisms at play within the MEDIRL framework.

Ultimately, our experimentation serves as a testament to the pursuit of knowledge, with the ambition to redefine and fortify the pathways to socially intelligent navigation.

Scope of Reproducibility – Recreating the original MEDIRL framework, as outlined in the research paper, proved challenging due to the lack of comprehensive documentation. Additionally, the absence of a publicly available GitHub repository with the necessary data required us to independently develop the algorithm, using the limited pseudocode

Copyright © 2019 Unknown Individual, released under a Creative Commons Attribution 4.0 International license.
Correspondence should be addressed to Unknown Individual (UnknownIndividual@purdue.edu)
The authors have declared that no competing interests exists.
Code is available at https://github.com/rescience-c/template.

provided in the paper as guidance. The lack of substantial information about the Social Affinity Map (SAM) feature map added to the complexity of our replication efforts. Unfortunately, the paper did not provide a reference or access to the dataset used, which further complicated our task.

Methodology — To reproduce the research paper's results, we employed a stepwise approach. Initially, we independently generated the MEDIRL model, relying on our interpretation of its implementation. Following this, we conducted ablation studies to break down its individual components and functions. We additionally optimized our efforts by subsetting the provided data and reducing the number of epochs, enabling us to execute the code on standard computing resources (we used a MacBook Pro 2018 i7 chip). To ensure future reproducibility and enhanced accessibility, we seamlessly integrated our code into Dags Hub (https://dagshub.com/ML-Purdue/hackathonf23-Stacks), along with data versioning via DVC and metrics tracked via MLFlow. It is important to note that we chose to omit the presented SAM Feature Map to focus solely on the capabilities of the Maximum Entropy Deep Inverse Reinforcement Learning Model. As such the comparisons we provide will be between the metrics that we gather, as to account for the differing manner of data processing.

Results — We prioritized the consideration of the average displacement from the model's predicted trajectory to the trajectory that the human in the testing data takes. The ranking from lowest displacement to highest displacement is as follows: Removed State Dimension, Original, Removed Discount Factor, Removed Hidden Layer, Removed Max Entropy and replaced with Mean Squared Error, Leaky ReLu instead of ReLU for activation.

What was easy — During the reproducibility process, setting up our environment and understanding what was entailed in the various experiments by the original authors was fairly straightforward. Additionally, it was simple to get started with the algorithm logic as the pseudo-code was provided by the paper along with the outlines for the Deep Learning Model.

What was hard — One significant challenge we faced was implementing the model with inadequate computational resources to handle the substantial 15 million-line dataset effectively. Furthermore, finding the dataset and implementing the model without any attached source code as a reference was difficult at times.

2 Introduction

In the realm of human-robot interaction (HRI), the confluence of humans and autonomous entities within shared spaces marks a paradigm shift in technological advancements [2]. This coexistence necessitates the development of robust and socially intelligent navigation systems, ensuring not just efficient movement but also safety, user acceptance, and the seamless integration of robots into human spaces. Within this dynamic landscape, the study Learning How Pedestrians Navigate: A Deep Inverse Reinforcement Learning Approach, by Fahad, Chen, and Guo [1] presents a pioneering methodology that harnesses maximum entropy deep inverse reinforcement learning (MEDIRL) to understand and replicate socially acceptable human navigation behaviors.

This groundbreaking research underscores the essential need for robots to navigate human-centric environments while adhering to social norms and conventions, thus fostering a natural and intuitive human-robot interaction [3]. The Fahad, Chen, and Guo study, which initially introduced the MEDIRL framework, serves as a cornerstone in

this transformative domain. Their work, focusing on capturing and modeling human navigation behaviors in crowded settings, laid the foundation for leveraging intricate datasets of human pedestrian trajectories, a nonlinear reward function facilitated by deep neural networks, and the integration of social affinity maps (SAM) for nuanced navigation decision-making.

Maximum Entropy Deep Inverse Reinforcement Learning (MEDIRL) holds a central position as a crucial machine learning and reinforcement learning framework in the field of human-robot interaction (HRI). It specifically focuses on the advancement of socially intelligent navigation. Within this multifaceted framework, the primary objective revolves around endowing robots with the ability to extract valuable insights from human behavior. This involves discerning the latent reward functions that underlie these behaviors and subsequently enabling the robots to make navigation decisions that go beyond mere efficiency, as described in reference [4].

Building on this pivotal research, the objective of our re-experimentation is to delve deeper into Fahad, Chen, and Guo's work, rigorously retesting, and expanding their findings. We aim not only to replicate their methodology but to significantly augment their research through nuanced re-analysis, additional experimentation, and a comprehensive validation process. By scrutinizing and extending the boundaries of their groundbreaking model, our goal is to further reinforce the reliability and applicability of MEDIRL within real-world human-robot interaction scenarios.

A critical aspect of our re-experimentation involves not just replicating the findings of the initial study but expanding its horizons. Through comprehensive evaluation against real-world pedestrian trajectories and rigorous comparisons against established methodologies, we aim to showcase a deeper understanding and validation of the MEDIRL model. Our mission is to advance this model to generate pedestrian trajectories that mirror human-like behaviors more accurately, encompassing vital aspects such as collision avoidance strategies, leader-follower dynamics, and intricate split-and-rejoin patterns.[5]

Additionally, the emphasis in our re-experimentation will be on reinforcing the reliability of the MEDIRL model. By employing strategic refinements, such as fine-tuning critical parameters, optimizing learning strategies, and meticulously eliminating biases, we aim to ensure the robust deployment of this technology in varied real-world HRI scenarios. This rigorous refinement process is pivotal in not only upholding social norms but also adhering to stringent safety protocols.

Crucially, our re-experimentation will systematically deconstruct and analyze the individual components constituting the MEDIRL model introduced by the Original Study. By employing meticulous ablation studies, we aim to dissect and comprehend the impact of each component on the overall performance of the model. Ablation studies play a pivotal role in dissecting and comprehending the individual contributions of distinct components within the MEDIRL framework [6].These studies involve selective removal or alteration of critical elements to gauge their influence on the overall performance of the model.

1. **Removal of Hidden Layer**:
    (a) The hidden layer in the MEDIRL model serves as an essential component in deep learning architectures. It plays a critical role in capturing and representing complex relationships within the data [7]
    (b) Ablating the hidden layer involves eliminating one or more hidden neural network layers from the MEDIRL model. This modification seeks to understand

how the depth of the network impacts the model's capacity to learn intricate features and non-linear relationships. [8]

(c) The ablation aims to evaluate whether a shallower network can still adequately capture the nuances of human navigation behaviors, or if a deeper network is essential for modeling the complexity of real-world scenarios.

2. Removal of State Dimension:

   (a) The state dimension in the MEDIRL model typically represents the environmental states and conditions that the robot and pedestrians navigate in. It encapsulates critical information about the surroundings. For our study, we removed the height component by modifying the model to account for x and y directions solely.

   (b) By removing a state dimension, we aim to assess the model's adaptability to changes in the state space. This ablation examines whether the model can generalize well and make robust navigation decisions when a part of the state information is missing.

   (c) Understanding the impact of this ablation is vital for assessing the model's capacity to adapt to variations in the environment.

3. Removal of Discount Factor:

   (a) The discount factor in reinforcement learning models influences the importance of future rewards in the decision-making process. It determines the model's preference for immediate rewards over long-term goals.

   (b) Removing the discount factor helps evaluate the model's ability to make decisions solely based on immediate consequences. This ablation assesses whether the model can adapt to scenarios where long-term planning and future rewards are not considered.

   (c) The results of this ablation will shed light on the role of discount factors in modeling navigation decisions and their impact on the balance between short-term and long-term considerations.

4. Removal of ReLU activation (replaced with Leaky ReLU):

   (a) Leaky Rectified Linear Unit (ReLu) is an activation function in neural networks. It allows a small gradient for negative input values, making it suitable for capturing non-linear relationships in the data [9].

   (b) Using Leaky ReLu as an activation function replaces the standard ReLu activation in the model. This change explores how the choice of activation function affects the model's ability to capture non-linear patterns in human behavior [10].

   (c) This ablation aims to assess whether the Leaky ReLu activation function enhances the model's capability to represent complex and non-linear features in the data, potentially improving its performance in modeling human navigation behaviors.

5. Removal of Max Entropy (replaced with mean squared):

   (a) Maximum entropy reinforcement learning encourages exploration by maximizing the entropy of the policy. It promotes diversity in the model's actions and adaptability to different scenarios. [11]

   (b) The removal of the max entropy component assesses the impact on the model's exploration-exploitation trade-off. Without it, the model may become less exploratory and may exhibit more deterministic behavior.

(c) This ablation will provide insights into the role of entropy in shaping the model's navigation decisions and whether reducing exploration influences its performance in diverse situations.

Each of these ablation studies plays a critical role in understanding the individual contributions and significance of specific components within the MEDIRL framework. The results from these detailed investigations will not only provide valuable insights into the model's performance but also guide further refinements and enhancements to create a more robust and adaptable model for socially intelligent navigation in human-robot interaction scenarios.

The analysis stemming from these ablation studies will not only provide deeper insights into the model's performance dynamics but also enable a refined understanding of the intricate mechanisms at play within the MEDIRL framework. [12] Ultimately, this meticulous approach to dissection and analysis will pave the way for an enhanced and fortified MEDIRL model, offering unparalleled advancements in socially intelligent navigation within the domain of human-robot interaction [13].

Our goal is to delineate the critical components significantly contributing to the model's effectiveness in replicating human navigation behaviors and fostering a deeper understanding of the intricate mechanisms at play. Through the meticulously conducted re-experimentation, our ambition is to unveil deeper insights and refined conclusions about the reliability and efficacy of MEDIRL within the realm of social affinity and its implications on navigation within the ambit of HRI [14]. This re-experimentation stands as a testament to the relentless pursuit of knowledge, aiming not just to replicate but to redefine and fortify the pathways to socially intelligent navigation within human-robot interaction.

3 Scope of reproducibility

The MEDIRL paper provides a series of information regarding the algorithm they developed, a Maximum Inverse Reinforcement Learning Model integrated with a Deep Learning Neural Network with differing levels of detail.
The original paper begins by outlining the Markov Decision Making Process elements.
Given MDP Elements:

- **States S:** The original paper uses states to represent all possible positions or situations the mobile robot could find itself it. It was denoted as the following set $S = \{s_1...s_n\}$, where 'n' is the total number of such possible states.
- **Actions A:** The original paper uses actions to represent all the possible decisions the mobile robot could make. This was denoted by the following set $A = \{a_1...a_p\}$, where 'p' denotes the total number of possible actions.
- **Discount Factor, γ:** This was denoted by the original paper as a number between 0 and 1 that outlined the impact a reward would have on the mobile robot based on its distance from the mobile robot.
- **Reward Function $R(s_i)$:** This was outlined as the function that the mobile robot would come up with on how it should operate within a state action space.

In regards to the Deep Learning Neural Network Backbone, we are told that it consists of one input layer, two hidden layers, and one output layer. The two hidden layers respectively have 4096 and 2048 nodes. Equation 1 displays the reward function formula the original paper gives us and Equation 2 represents the Bayesian inference that the original paper uses.

$$R^* = g(\phi, \theta_1, \theta_2, \theta_3, \ldots, \theta_j), \; = g_1(g_2(\ldots(g_j(\phi, \theta_j), \ldots), \theta_2), \theta_1). \tag{1}$$

$$L(\theta) = \log P(D, \theta | R^*) = \log P(D|R^*) + \log P(\theta) \tag{2}$$

The original paper also outlines equation 3 to represent the gradient descent taking place for the neural network optimization with respect to the network parameters θ and equation 4 outlines the gradient descent with respect to the reward function.

$$\frac{\partial L}{\partial \theta} = \frac{\partial LD}{\partial \theta} + \frac{\partial L\theta}{\partial \theta}. \tag{3}$$

$$\frac{\partial LD}{\partial \theta} = \frac{\partial LD}{\partial R^*} \cdot \frac{\partial R^*}{\partial \theta}. \tag{4}$$

No further information is provided about the Maximum Entropy Inverse Reinforcement Learning Model embedded into the Deep Learning Network beyond its formulas shown in Equations 5 and 6, where $\mu_D - E[\mu_m]$ is the state visitation matching feature.

$$L_m D = \log(\pi_m) \cdot \mu_a \tag{5}$$

$$\frac{\partial LD}{\partial R_m^*} = \mu_D - E[\mu_m] \tag{6}$$

The MEDIRL paper captures the pedestrian behavior and evaluates it as such:

- Accuracy of 96.6%
- Average Displacement Error of 0.40 meters
- Final Displacement Error of 0.81 meters
- Average Non-Linear Displacement Error of 0.41 meters

It also compares it to another state-of-the-art algorithm to indicate that its model should be the new state of the art.
Given the missing information, we had the make the following assumptions about the model:

- Used our own discount factor of 0.01.
- Epoch number of 3.
- Used a standard number of nodes for the input and output layers given out data set.
- Used a standard maximum entropy inverse reinforcement optimization method for the deep learning network.

It is also important to note that from the data set provided by the original paper, we subseted 100 lines for training and 40 lines for testing. We did this to adjust the dataset to be suited for the lack of computational power we had available to us for this study. As we had 6 ablation studies with no access to an intel core Macbook, we subseted the data. The original paper claims that given a 1080ti with dual Xeon processors, it would take 20 hours to run the code.

From the provided metrics of the original paper, we intend to focus on the Average displacement error of the model, as it is the most consistent metric considering the difference in training data size (due to computational restrictions).

4 Methodology

We began this study by re-creating the algorithm shown in the original paper as displayed in the Algorithm 1.

Algorithm 1 Maximum Entropy Deep Inverse Reinforcement Learning (MEDIRL)

```
Require:
    num_trajectories: Number of human-like trajectories
    trajectory_length: Length of each trajectory
    state_dim: State-space dimension
    lr: Learning rate
    epochs: Training epochs
Ensure:
    irlModel: Trained IRL model
function MaxEntIRL(state_dim)
    irl ← Initialize MaxEntIRL model
    return irl
end function
function TrainMaxEntIRL(num_trajectories, trajectory_length, file_path, lr, epochs)
    irl ← MaxEntIRL(state_dim)
    data ← LoadDataset(file_path)
    irl.train_irl(data, use_dataset = True, lr, epochs)
    model_path ← '/path/to/save/model.pkl'
    irl.save_model(model_path)
    return model_path
end function
function TrainIRLWithDataset(data, lr, epochs)
    optimizer ← Initialize Adam optimizer with lr
    for epoch ← 1 to epochs do
        totalLoss ← 0
        state_frequencies ← Calculate state frequencies from data
        for idx ← 1 to len(data) do
            state, velocity ← data[idx]
            Using GradientTape:
            preferences ← irl.model(state)
            prob_human ← Softmax(preferences)
            maxent_irl_objective ← Calculate MaxEnt IRL objective
            grads ← Compute gradients
            Apply gradients using optimizer
            totalLoss ← totalLoss + ∑(maxent_irl_objective)
        end for
        avg_loss ← totalLoss/len(data)
        Log loss metric in MLflow
        Print "Epoch epoch/epochs, MaxEnt IRL Loss: avg_loss"
    end for
end function
```

We then proceeded with creating the code for our ablation studies. We:

- removed a Hidden Layer consisting of 2048 nodes, keeping the bigger one of 4028 nodes as the sole hidden layer. We hypothesize this will lead to a far more inaccurate model due to the reduction of neurons.
- removed a State Dimension making the state space 2 instead of 3. We hypothesize this will lead to an increase in model accuracy, as removing the height dimension

for traversing space will reduce the dimensionality of the issue making it easier for the model to understand.

- removed the Discount Factor entirely such that the distance of the reward would have no effect on the model. We hypothesize this will lead to the model prioritizing farther but larger rewards than closer and easier to achieve ones leading to it operating worse than before.
- removed the RelU activation and replaced it with the Leaky ReLU shown in equation 7. We hypothesize this will lead to the activation function being more robust changing the overall decisions of the model and its consideration of negative weights.

$$f(x) = \begin{cases} x, & \text{if } x > 0, \\ \alpha x, & \text{if } x \leq 0. \end{cases} \tag{7}$$

- removed the maximum entropy loss calculation and replaced it with mean squared shown in equation 8. We hypothesize this will lead to less exploration within the model and make it more imitative of the behaviors of the demonstrators which would change the mobile robots decisions.

$$\text{MSE}(\theta) = \frac{1}{N}\sum_{i=1}^{N}(y_i - f(x_i, \theta))^2 \tag{8}$$

We then save these models as a pickle file locally so that we can run it against test data. The run time of all these models is O(N) and the space complexity is also O(N).The key metric we aim to take note of is the difference between the trajectory the model would take and the trajectory the human actually takes.

In conducting these ablation studies we aim to identify which features of the MEDIRL model are necessary and the impact it has on the overall performance of the model. We do this by cross-referencing the data against the "standard" that we establish with the MEDIRL model's performance.

5 Results

After conducting our reproducibility study as per the method outlined above we noted the following results.
The model's Epoch Training Loss and Average Displacement were as follows:

- Original, Epoch Training Loss shown in figure 1. Average Displacement: 1.12 m shown in Figure 2.

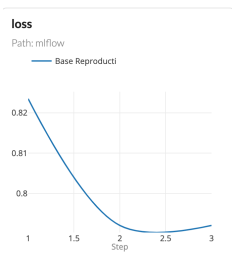

Figure 1. *Figure displays the epochs of the original model.*

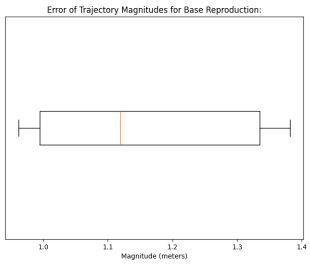

Figure 2. *Figure displays the displacement of the predictions made of the original model from the actual decisions made by the pedestrians.*

- Removed a Hidden Layer, Epoch Training Loss shown in Figure 3. Average Displacement: 1.14 m as shown in Figure 4.

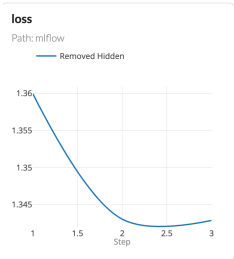

Figure 3. *Figure displays the epochs of the original model without a hidden layer.*

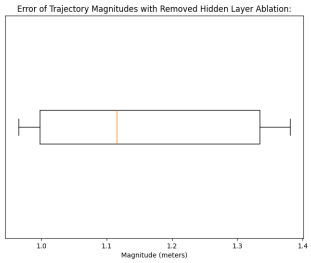

Figure 4. *Figure displays the displacement of the predictions made of the model without a hidden layer from the actual decisions made by the pedestrians.*

- Removed the vertical State Dimension, Epoch Training Loss shown in Figure 5. Average Displacement: 0.91 m Figure 6.

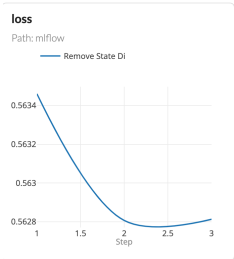

Figure 5. *Figure displays the epochs of the original model without a vertical state dimension.*

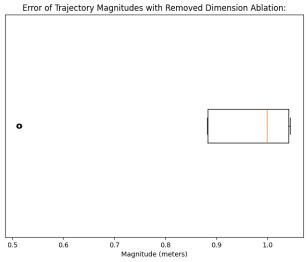

Figure 6. *Figure displays the displacement of the predictions made of the model without a vertical state dimension from the actual decisions made by the pedestrians.*

- Removed the Discount Factor, Epoch Training Loss shown in Figure 7, Average Displacement: 1.13 m as shown in Figure 8.

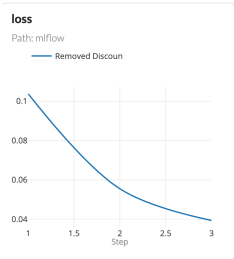

Figure 7. *Figure displays the epochs of the original model without a discount factor.*

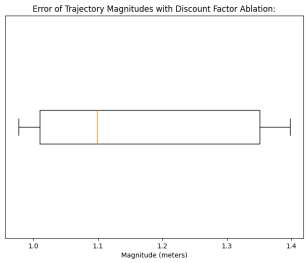

Figure 8. *Figure displays the displacement of the predictions made of the model without a discount factor from the actual decisions made by the pedestrians.*

- Removed the ReLU activation in favor of Leaky ReLU, Epoch Training Loss shown in Figure 9, Average Displacement: 1.15 m as shown in Figure 10.

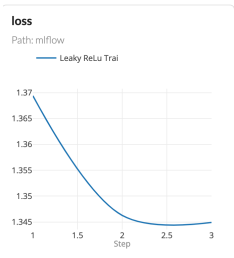

Figure 9. *Figure displays the epochs of the original model without a discount factor.*

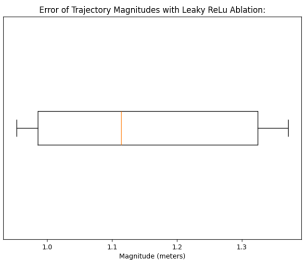

Figure 10. *Figure displays the displacement of the predictions made of the model with a leaky ReLU activation instead of a ReLU activation from the actual decisions made by the pedestrians.*

- Removed the maximum entropy loss in favor of Mean Squared Loss Calculation, Epoch Training Loss shown in Figure 11, Average Displacement: 1.15 m as shown in Figure 12.

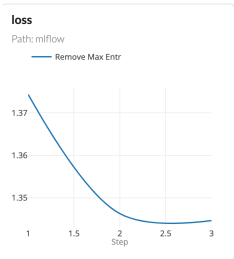

Figure 11. *Figure displays the epochs of the original model without a discount factor.*

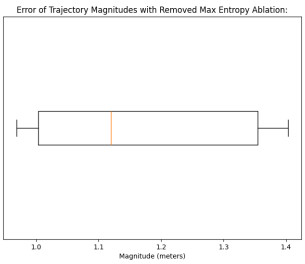

Figure 12. *Figure displays the displacement of the predictions made of the model with mean squared instead of maximum entropy from the actual decisions made by the pedestrians.*

A full comparison of the Epochs between the models can be displayed in Figure 13.

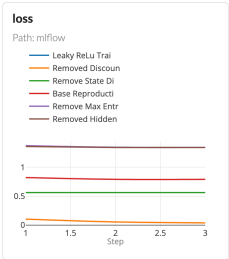

Figure 13. *Figure displays the displacement of the predictions made of the model with mean squared instead of maximum entropy from the actual decisions made by the pedestrians.*

The ranking from lowest displacement to highest displacement is as follows: Removed State Dimension, Original, Removed Discount Factor, Removed Hidden Layer, Removed Max Entropy, Leaky ReLU.

6 Results reproducing original paper

Our replication of the original model netted us an average displacement of 1.12 m in comparison to the .5 m that the original paper's model was able to get. This difference is likely because we trained our data set with a significantly smaller subset of the data given our lack of computational power. It is also important to note that while a majority of the ablation studies did worse than the original study replication, removing the vertical state dimension seems have increased its accuracy. This is likely because the humans within the environment are not vertically moving and this additional dimension just leads to excess unnecessary error.

7 Discussion

Based on our reproducibility attempt alongside our ablation studies we can clearly see how each component of the machine learning model had an effect on its capabilities to replicate human behavior in social navigation settings. The ablation study indicates that future research within the human social navigation context should establish their Markov Decision Making Framework within the two-dimensional space if no vertical movement is present, to mitigate any error that could occur based on the height of the individual in question. By doing this within our ablation study we were able to reduce the average displacement of the model. Another thing to note from our ablation study is the importance of the Maximum Entropy Component that was presented in the Original Paper. Once that component was removed from the model the average displacement increased significantly making the model substantially worse when using Mean Squared Error loss calculation instead. It is also important to note that swapping the ReLU activation with Leakly ReLU is the ablation study that did the worst and likely not something that should be done for future research in human social navigation settings. Something else to consider is that given the discount factor that we used was so small 0.01 is it likely that removing it all together in the ablation study had minimal effects hence its results being similar to the original study. And as one would expect removing a hidden layer made it model worse and increased its displacement.

The key takeaways from our ablation study are as follows for future human social navigation research: The importance of utilizing a two-dimensional Markov decision-making framework when no vertical movement is involved, using a ReLU activation function over a Leaky ReLU activation function and proper documentation through the presenting of a model as to make it for future researchers to replicate.

Editorials, letters and special issues — ReScience C remains dedicated to the publication of computational replications but we (i.e., the editorial team) would like to have the opportunity to publish *editorials* when deemed necessary and to give anyone the opportunity to write *letters* to the community on a specific topic related to reproducibility. Both editorials and letters are expected to be 1 or 2 pages long (but no hard limit will be enforced), will be (quickly) peer reviewed, and will be assigned a DOI. Furthermore, with the advent of reproducibility hackatons worldwide, we will host *special issues* with guest editors (such as, for example, the organizers of a hackaton) in order to publish the results and to enhance their discoverability. Each entry will have to go through the regular open peer-reviewed pipeline.

We hope that most readers will agree on the proposed changes such that we can commit to them in the next few weeks. The review for this editorial is open (as usual) and anyone can comment on and/or oppose any of the proposed changes. New ideas are also welcome.

References

1. M. Fahad, Z. Chen, and Y. Guo. "Learning How Pedestrians Navigate: A Deep Inverse Reinforcement Learning Approach." In: **2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)** (2018).
2. K. Kosuge and Y. Hirata. "Human-Robot Interaction." In: **2004 IEEE International Conference on Robotics and Biomimetics** (Apr. 2004), pp. 8–11.
3. S. Yao, G. Chen, Q. Qiu, J. Ma, X. Chen, and J. Ji. "Crowd-Aware Robot Navigation for Pedestrians with Multiple Collision Avoidance Strategies via Map-based Deep Reinforcement Learning." In: **2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)** (2021), pp. 8144–8150.
4. M. Wulfmeier, P. Ondruska, and I. Posner. "Maximum entropy deep inverse reinforcement learning." In: **arXiv preprint arXiv:1507.04888** (2015).
5. D. Helbing and P. Molnar. "Social force model for pedestrian dynamics." In: **Physical Review E** 51.5 (1995), p. 4282.
6. R. Meyes, M. Lu, C. W. de Puiseau, and T. Meisen. "Ablation studies in artificial neural networks." In: **arXiv preprint arXiv:1901.08644** (2019).
7. T. Haarnoja. "Acquiring diverse robot skills via maximum entropy deep reinforcement learning." University of California, Berkeley, 2018.
8. Y. Bengio, I. Goodfellow, and A. Courville. **Deep Learning**. MIT Press, 2017.
9. B. Xu, N. Wang, T. Chen, and M. Li. "Empirical evaluation of rectified activations in convolutional networks." In: **arXiv preprint arXiv:1505.00853** (2015).
10. A. Almeida and G. Azkune. "Predicting Human Behaviour with Recurrent Neural Networks." In: **Applied Sciences** 8.2 (2018), p. 305.
11. Y. Zhou, R. Fu, and C. Wang. "Learning the car-following behavior of drivers using maximum entropy deep inverse reinforcement learning." In: **Journal of Advanced Transportation** 2020 (2020), pp. 1–13.
12. S. Sheikholeslami, M. Meister, T. Wang, A. H. Payberah, V. Vlassov, and J. Dowling. "Autoablation: Automated parallel ablation studies for deep learning." In: **Proceedings of the 1st Workshop on Machine Learning and Systems**. 2021, pp. 55–61.
13. M. Muffoletto, A. Qureshi, A. Zeidan, L. Muizniece, X. Fu, J. Zhao, and O. Aslanidi. "Toward patient-specific prediction of ablation strategies for atrial fibrillation using deep learning." In: **Frontiers in Physiology** 12 (2021).
14. R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, and J. Wang. "Designing robots for long-term social interaction." In: **2005 IEEE/RSJ International Conference on Intelligent Robots and Systems**. 2005, pp. 1338–1343.