Measures of Generalisation for Deep Reinforcement Learning Models

Literature Review - Group 28

Mark Flowers, Arpita Goyal, Patrick Leask, Akshay Shrinivas Munde, Amol Sahebrao Rathod, Kshitij Singh, Saeth Wannasuphoprasit

Structure of the Presentation

- The importance of generalisation and why it is difficult to achieve
- Introduction to Deep Reinforcement Learning
- Measuring generalisability
  - Empirical analysis
  - Theoretical analysis
    - Formal methods
- Discussion of the state of the art
- Directions for further research
- Q&A

Generalisation

Generalisation matters

- Adaptability of a model to a new environment.
- Hallmark of intelligent systems (Lake et al. 2017).
- Essential for safety critical tasks, driving

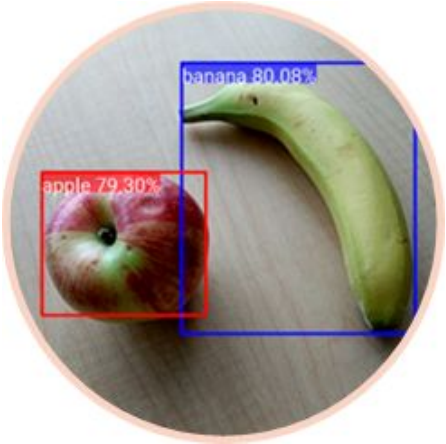
Generalisation is hard

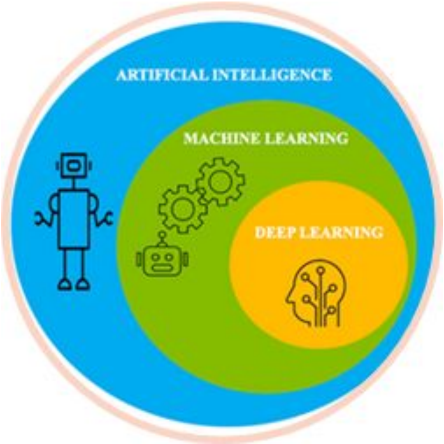
- Sit!
- Deep learning performance is highly susceptible to small changes in the environment

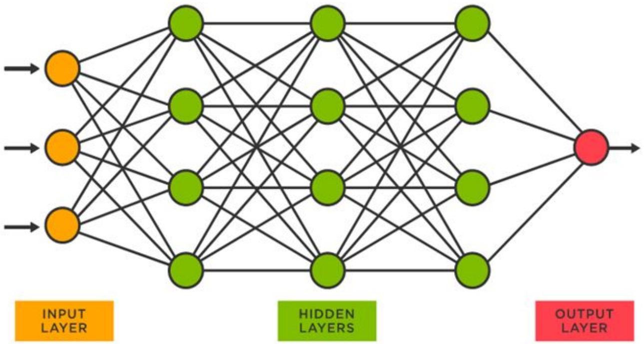
Deep Reinforcement Learning

Deep Learning

- Models mimic the neural structure of the brain.
- Excel on high dimensionality, sparse datasets.
- Example use cases
  - Image recognition
  - Translation

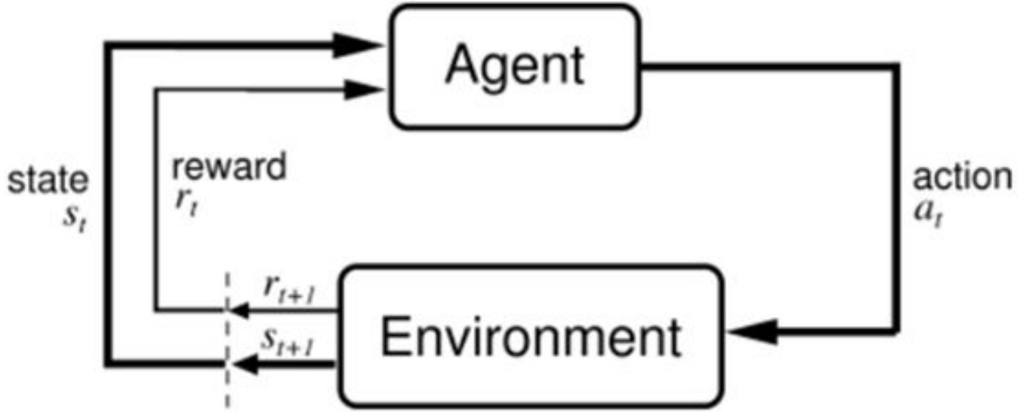






Reinforcement Learning

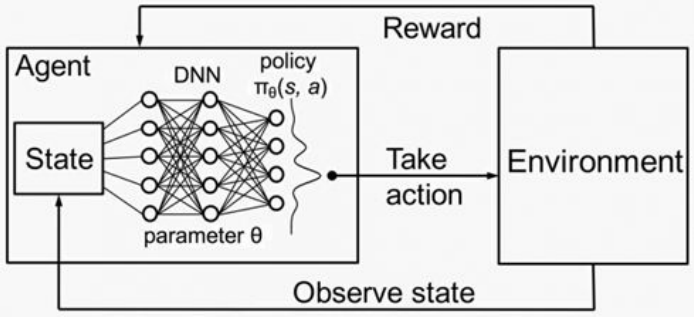
- The agent performs an action based on the state of the environment.
- The action updates the state of the environment and solicits some level of reward from the environment.
- The new state of the environment and level of reward form the new input to the model



Applications of reinforcement learning

Deep Reinforcement Learning

- Uses deep learning models to find rewarding policies.
- Better handles the complexity of high dimensionality state spaces.
- Deep RL models have mastered Go (Silver et al. 2016) and StarCraft II (Vinyals et al. 2019)
- Expectations for DRL in the future are very high; Silver et al. (2021) claim that reward maximisation is sufficient for intelligence.



Measuring Generalisation

Overview

- Empirical methods - test a given model on new data
- Theoretical methods - predict the performance of a model given the training data
- Reachability analysis - given a trained model, what behaviour is possible

Empirical Methods

- Procedurally generating new levels of a game, eg. CoinRun (Cobbe et al. 2019).
- Variations in the parameters of an environment (Packer et al. 2019).

A procedurally generated level in CoinRun

Empirical Methods - Discussion

- Test environments are overly simplistic, and include only similar tasks (e.g. varying one element of the environment).
- Empirical tests do not explain why models are successful at generalising.
- Danger of researchers overfitting to test environments (cf. ImageNet).
- Test environments are model agnostic.

Theoretical Measures - Generalisation Bounds

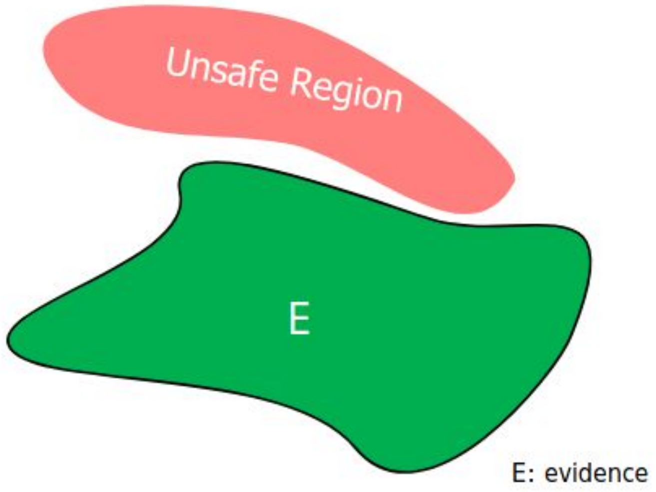
- Attempt to establish a bound for the generalisation gap of a model on a dataset.
- DL models generally can achieve zero training error - eg. random ImageNet categories
- Prove a bound on the generalisation gap based on the training data and model (not test data) - hard to get narrow bounds

Theoretical Measures - Correlated measures

- Correlate generalizability with measures, eg. sharpness of minima
- Correlation != causation, need careful experiments

Reachability Analysis

- Provide guarantees on model behaviour.
- Given a range of inputs, what are the possible outputs of a model.
- Derive properties such as the robustness of the model.
- Useful for understanding behaviour where testing is not appropriate.



Protecting against unsafe behaviour (DARPA)

Reachability Analysis - Discussion

- Scaling is a problem - currently <60k nodes.
- Determining the unsafe zone and adjusting model behaviour.
- Heuristics like robustness need interpretation.

Analysis

Findings

- No consensus on measures for generalisability in the literature.
- Interpretation of the results is difficult.
- Lower research intensity vs. model building.

Recommendations

- Establish a standard empirical test set
  - Hold back test data to prevent overfitting
  - Increase complexity and realism as models improve
- Provide more education on these methods
  - How to interpret results of theoretical methods
  - Applicability of empirical results to new environments

- Silver, D., Schrittwieser, J., Simonyan, K. et al. (2017) Mastering the game of Go without human knowledge. Nature 550, 354–359. https://doi.org/10.1038/nature24270
- Vinyals, O., Babuschkin, I., Czarnecki, W.M. et al. (2019) Grandmaster level in StarCraft II using multi-agent reinforcement learning. Nature 575, 350–354. https://doi.org/10.1038/s41586-019-1724-z
- Levenson RM, Krupinski EA, Navarro VM, Wasserman EA (2015) Pigeons (Columba livia) as Trainable Observers of Pathology and Radiology Breast Cancer Images. PLoS ONE 10(11): e0141357. doi:10.1371/journal.pone.0141357
- A. T. Simin, S. Mohsen Ghorabi Baygi and A. Noori, (2020), "Cancer Diagnosis Based on Combination of Artificial Neural Networks and Reinforcement Learning," 2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), pp. 1-4, doi: 10.1109/ICSPIS51611.2020.9349530.
- Zhuo Liu, Chenhui Yao, Hang Yu, Taihua Wu,(2019), Deep Reinforcement learning with its application for lung cancer detection in medical Internet of Things,Future Generation Computer Systems, Volume 97, Pages 1-9, ISSN 0167-739X, https://doi.org/10.1016/j.future.2019.02.068.
- Vincent François-Lavet, Peter Henderson, Riashat Islam, Marc G. Bellemare and Joelle Pineau (2018), “An Introduction to Deep Reinforcement Learning”, Foundations and Trends in Machine Learning: Vol. 11, No. 3-4. DOI: 10.1561/2200000071.
- Naeem, Muddasar & Rizvi, Syed & Coronato, Antonio. (2020). A Gentle Introduction to Reinforcement Learning and its Application in Different Fields. IEEE Access. 8. 209320-209344. 10.1109/ACCESS.2020.3038605.
- SIlver, David, Singh, Santinder, Precup, Doina, Sutton, Richard, (2021), Reward is enough,Artificial Intelligence,Volume 299, 103535, ISSN 0004-3702, https://doi.org/10.1016/j.artint.2021.103535.
- Whiteson, S., Tanner, B., Taylor, M. E., and Stone, P. (2011) Protecting against evaluation overfitting in empirical reinforcement learning. In IEEE Symposium on Adaptive Dynamic Programming And Reinforcement Learning,
- Zhang, A., Ballas, N., and Pineau, J. 2018 A dissection of overfitting and generalization in continuous reinforcement learning. arXiv:1806.07937
- Sutton, R. S. and Barto, A. G. (2017) Reinforcement Learning: An Introduction. MIT Press, 2nd edition
- Rosenblatt, Frank (1957). "The Perceptron—a perceiving and recognizing automaton". Report 85-460-1. Cornell Aeronautical Laboratory.
- Haider, T., Roza, F. S., Eilers, D., Roscher, K., Günnemann, S. (2021) Domain Shifts in Reinforcement Learning: Identifying Disturbances in Environments http://ceur-ws.org/Vol-2916/paper_11.pdf