

reward

Authors

Alexander Jung^{*}, Department of Computer Science, Aalto University, Espoo, Finland

^{*}Corresponding author: alex.jung@aalto.fi

Abstract

A reward is an observed or measured quantity that allows a machine learning (ML) method to estimate the loss incurred by the prediction or action of a hypothesis $h(\mathbf{x})$. A reward replaces the label: it does not reveal the correct prediction, only how good the chosen one was, and it may be delayed or noisy. Reinforcement learning (RL) methods construct a loss from the rewards received and choose actions that maximize the cumulative reward. For example, a self-driving vehicle whose steering direction is $h(\mathbf{x})$ can derive a reward from a collision sensor, assigning a low reward to a direction that moves the vehicle toward an obstacle.

Definition

A reward refers to some observed (or measured) quantity that allows estimation of the loss incurred by the prediction (or decision) of a hypothesis $h(\mathbf{x})$. For example, in a machine learning (ML) application to self-driving vehicles, $h(\mathbf{x})$ could represent the current steering direction of a vehicle. A reward could be constructed from the measurements of a collision sensor that indicate if the vehicle is moving toward an obstacle. A low reward is defined for the steering direction $h(\mathbf{x})$ if the vehicle moves dangerously toward an obstacle.

Cross References

loss, MAB, RL

References

- 1.