# Teachers Are The Best Teacher

Alok Singh

**Please record**

# What

How to (make a computer) learn to do things from an expert.

Formally, given a set of *states*, pick the best *actions* to *maximize sum of rewards* over time.

# Why Learn From An Expert?

Experience is expensive. Same reason we read books and listen to advice.

# Definitions

Policy Function from states to actions

Expert Policy Function from states to *good* actions

**How**

# Behavior Cloning

Use data gathered by expert to train policy to copy expert.
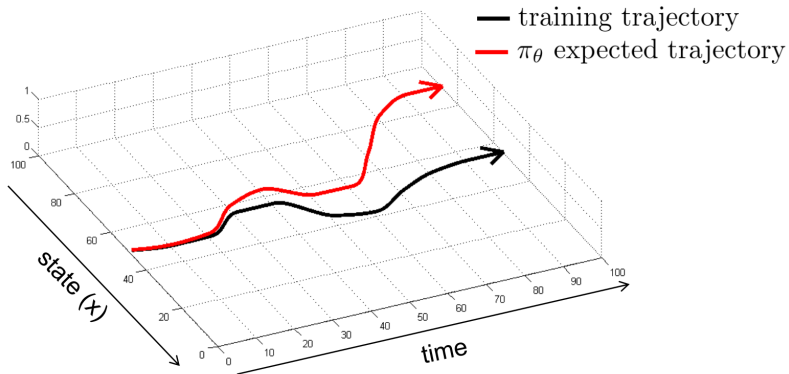
# Issue

This doesn't work.



Figure 1: New states lead to bad actions

**How To Improve?**

# DAgger (Dataset AGGregration)

Idea: Use the expert to point out what *to* do while you're still learning.

Why learn from experience when you can have the expert tell you what to do?

- Use expert policy to recommend actions for new, suboptimal states, but *don't do them*.
- *Store* expert action in training data and *do* regular action
- Use that training data to train policy
- Repeat until policy converges to match expert policy

Using expert's *actions* to train leads to good *decisions* and using the learner's *states* leads to *exploration* of new states.

Used in self-driving cars

# Feedback

alok.blog/about