SLIDE 1 OF 8 / 0:00-0:40

DAgger and HG-DAgger

Main idea: A robot can learn from the moments when a person takes over.

EXPLANATION CUES

Open with the photo of the expert's driving station in the HG-DAgger car: a separate steering wheel and pedals. Ask students to imagine teaching a new driver: watching a good drive helps, but correcting a mistake teaches something different.

Name the papers newest first: HG-DAgger, from 2019, builds on DAgger, from 2011. The shared question is how to collect examples a learning robot actually needs.

Define the central term once: a policy is the robot's learned controller, the part that chooses its next movement. Use "controller" for the rest of the talk.

POINT OR DEMONSTRATION

Point to the expert's steering wheel. Turning it takes control from the robot, and a button hands control back. The teacher acts instead of only describing what should happen.

TRANSITION

First, what goes wrong when a robot only copies successful examples?

IF ASKED

HG means human-gated. DAgger abbreviates Dataset Aggregation: keeping old examples while adding new ones.

Source: https://arxiv.org/pdf/1810.02890; https://proceedings.mlr.press/v15/ross11a/ross11a.pdf

SLIDE 2 OF 8 / 0:40-2:00

Why copying a good example can fail

Main idea: A small mistake can create a situation the demonstrations never showed.

EXPLANATION CUES

Use the can example: a demonstration shows reaching, grasping, then placing the can. Training pairs the camera view with the movement a person made.

Now imagine the robot misses the grasp. In the demonstration the hand already held the can. In this run the hand is empty, so blindly continuing the placement motion cannot finish the job.

This is why another perfect demonstration may be less useful than one example of recovering a missed grasp. The robot needs instruction in the situation its own error created.

POINT OR DEMONSTRATION

Use the household-robot image only as an illustration from our project. It is not an experiment from the two reviewed papers.

TRANSITION

DAgger changes where the teacher supplies examples.

IF ASKED

The technical name is distribution shift: training and practice contain different situations. Ross et al. give a worst-case compounding-error bound; it is not a prediction of the exact failure rate of every robot.

Source: https://proceedings.mlr.press/v15/ross11a/ross11a.pdf (Section 2); Project illustration: team report, 3 November 2026, Figure 2a.

SLIDE 3 OF 8 / 2:00-3:30

Teaching from the robot’s mistakes

Main idea: DAgger teaches on the robot’s situations; HG-DAgger gives the human control during corrections.

EXPLANATION CUES

Walk the table one row at a time. Copying demonstrations records a person doing the task well. The robot later tries to copy those actions.

With DAgger, the robot practices and the expert supplies the action it should take at each visited step. Add those examples to the earlier data and train again. The expert may provide an answer even when that answer is not the action executed.

With HG-DAgger, the person watches, chooses when help is needed, takes full control, corrects, and hands back. The human-control intervals supply the new labels.

A label here simply means the action the teacher wants associated with an image. It does not mean a word such as “can” or “table”.

POINT OR DEMONSTRATION

Trace the difference between the second and third rows. Who chooses the handover, and whose action enters the recording?

TRANSITION

The papers contribute related but different ideas.

IF ASKED

Original DAgger can mix expert and learner control using a probability. Repeating collection and training is essential to its guarantee; one batch of corrections is not the full procedure.

Source: https://proceedings.mlr.press/v15/ross11a/ross11a.pdf (Algorithm 3.1); https://arxiv.org/pdf/1810.02890 (Algorithm 1)

SLIDE 4 OF 8 / 3:30-4:50

What each paper adds

Main idea: DAgger changes which examples we collect; HG-DAgger changes how a person can teach.

EXPLANATION CUES

DAgger keeps the examples collected so far and adds expert answers from the robot's latest practice. Its theory explains conditions under which this reduces the compounding-error problem.

HG-DAgger asks a human-factors question: how can someone provide a good correction if their controls do not reliably move the system? It gives the person uninterrupted control during a takeover.

The takeover also marks a moment the person considered risky. The paper trains several controllers and looks at how much they disagree. That disagreement is a warning signal, not proof of danger.

POINT OR DEMONSTRATION

Point left for where to collect examples; point right for who holds the controls. Keep the mathematical detail for questions.

TRANSITION

What evidence did the authors provide?

IF ASKED

The human still chooses training takeovers. HG-DAgger learns a risk threshold from the last quarter of takeover-disagreement values; this does not make the gate automatic. Uncertainty is an estimate, not a safety guarantee.

Source: https://proceedings.mlr.press/v15/ross11a/ross11a.pdf (Theorems 3.1-3.2); https://arxiv.org/pdf/1810.02890 (Section II, Eq. (4); Section IV-D, Fig. 6)

SLIDE 5 OF 8 / 4:50-6:20

Games and obstacle avoidance

Main idea: The methods helped in the reported games and driving experiments, within limited tests.

EXPLANATION CUES

For DAgger, describe the racing game: steering errors can push the kart off track. After 15 training rounds, the reported evaluation had zero falls per lap.

For HG-DAgger, explain the table before reading numbers. Rows are training methods; columns count obstacle collisions and road departures. The system uses a physical car with simulated obstacles.

HG-DAgger has zero in both columns, versus one collision and six departures for copying demonstrations, and one of each for DAgger.

Qualify the denominator: the first 5,000 control samples are a short sequence of measurements, not 5,000 independent drives. These results are evidence in the tested setting, not proof of general driving safety.

POINT OR DEMONSTRATION

Read the table across one row, then compare rows. Explain “off-road” as leaving the road.

TRANSITION

The practical lesson concerns the interaction between the teacher and the learner.

IF ASKED

The original kart demonstration video is in the sources. HG-DAgger also tests simulated driving. Do not turn the zero collision count into a claim of guaranteed safety.

Source: https://proceedings.mlr.press/v15/ross11a/ross11a.pdf (Section 5.1, Fig. 2); https://arxiv.org/pdf/1810.02890 (Section IV-C, Table II); https://www.youtube.com/watch?v=V00npNnWzSU

SLIDE 6 OF 8 / 6:20-7:20

Useful corrections need clear control

Main idea: The robot needs a useful correction, and the teacher needs predictable control.

EXPLANATION CUES

Return to the missed grasp. From the robot's perspective the missing example is what to do with an empty hand after the attempted pick-up.

From the teacher's perspective the issue is whether moving the controller causes the expected movement. If control switches unpredictably, the teacher may keep correcting something the robot does not execute.

HG-DAgger's design gives the person full authority during a correction, with a clear start and end. For VR, this makes the handover and feedback part of data collection.

POINT OR DEMONSTRATION

Read the two example questions as two viewpoints on the same teaching interaction.

TRANSITION

That is a useful design idea, but the evidence still has limits.

IF ASKED

The authors propose perceived lag as an explanation. They do not isolate it in a controlled user study, so do not state that latency was proven to cause the performance difference.

Source: https://arxiv.org/pdf/1810.02890 (Sections I-III)

SLIDE 7 OF 8 / 7:20-8:40

Limits of the evidence

Main idea: The methods still depend on good experts, enough practice, and appropriate evaluation.

EXPLANATION CUES

DAgger needs many expert answers and repeated retraining. On a physical robot, allowing a learner to make mistakes can create safety problems.

HG-DAgger depends on a person seeing trouble early enough to take over. A poor view, a fast event, or communication delay can make that difficult.

The real-car evaluation is small and task-specific. Neither paper establishes how different VR interfaces affect correction quality across people.

Separate these limitations from the contribution: learning from the robot's own situations remains useful even though no method removes all teaching or safety costs.

POINT OR DEMONSTRATION

Use one example for each limitation rather than listing technical assumptions.

TRANSITION

Those limits suggest a study we could run in VR.

IF ASKED

Avoid saying "HG-DAgger solves safety". Its risk estimate and human intervention depend on assumptions that may fail.

Source: https://arxiv.org/pdf/1810.02890 (Sections III and IV); https://proceedings.mlr.press/v15/ross11a/ross11a.pdf (Section 3)

SLIDE 8 OF 8 / 8:40-10:00

A VR experiment on taking control

Main idea: Test whether handover design and delay change the corrections people produce.

EXPLANATION CUES

Propose restoring the same saved robot failures for each interface condition. This helps compare interfaces on similar problems.

Compare a person taking full control with a person sharing control with the robot. Separately compare a responsive interface with added delay. Explain delay as a visible pause between hand movement and robot response.

Measure whether the correction works, how long it takes, and how hard it feels. Then test whether robots trained on those recordings improve.

Make clear that this is our proposed study. The reviewed papers did not run it. The separate project talk explains the VR system we built.

POINT OR DEMONSTRATION

Point to the question students can answer without knowing machine learning: can people teach better with a clearer handover?

TRANSITION

Invite questions about the teaching loop, the evidence, or the proposed interface study.

IF ASKED

Give conditions in different orders across participants and keep starting failures comparable. AI tools helped draft and construct the presentation; sources support the reported claims.

Source: https://arxiv.org/pdf/1810.02890 (motivation); https://proceedings.mlr.press/v15/ross11a/ross11a.pdf (learner-state sampling); Harvesting Failures for Imitation, team report, 3 November 2026 (Section 4.3)

Plain-language glossary

Use the explanation before introducing the technical term.

Imitation learning

Teaching a controller using examples of another agent's actions.

Behavioral cloning

Learning to copy the actions in recorded demonstrations.

Policy

The robot's learned rule or model for choosing its next action.

DAgger

Let the learner practice, collect expert answers on the situations it visits, add them to the data, and train again.

Human-gated

A person chooses when to take control and when to hand it back.

Uncertainty

An estimate of how unsure a model is. It does not prove whether an action is safe.

Rehearsal

Speak from the cues in your own words. Use the running-time markers to stay near ten minutes. In the project talk, the VR clip is included in slide 3's time.

The same slide notes are available in the Google Slides speaker-notes pane and Presenter view.

Open the Google Slides presentation

Questions and short answers

Why not collect more perfect demonstrations?

They can help, but may omit the situations created by the learner's own mistakes. DAgger targets those situations.

Does DAgger require a human?

No. An expert can be a person, a script, or a stronger controller, provided it can supply useful actions on learner states.

What does the human gate change?

The person chooses when to intervene and has full control during the correction. Only those human intervals add new action labels.

Does controller disagreement prove danger?

No. It is a warning estimate. The paper calibrates it using human takeovers, and evaluates it in specific driving settings.

What is the main weakness for a VR audience?

The papers do not isolate interface authority or delay in a controlled user study. That is an opportunity for follow-up work.