Script:

🎤 Final Presentation Script – *“Self-Taught Reasoner (STaR)”*

Slide 1: Technical Implementation

Time: 1.5 minutes

"Let’s start by unpacking how STaR actually works under the hood.

At its core, STaR is framed as a **latent variable model**, where the model learns to jointly generate rationales—denoted —and final answers .

There are two core equations here:

1. **Probability Decomposition**:
   This marginalizes over possible rationales—meaning the model learns to explore different paths to the right answer, just like how humans might consider multiple ways to solve a problem.
2. **Custom Loss Gradient**:
   The key here is the indicator function . This means the model **only updates its parameters** when both the answer and rationale are correct.
   That prevents it from reinforcing incorrect reasoning paths—making the training more **efficient and stable**.

For implementation, they fine-tune GPT-J (a 6B parameter model), using small batch sizes and a low learning rate on TPUs. This ensures smoother updates given the precision of the loss signal."*

Transition: *Let’s see how this theoretical advantage plays out in practice.*

Slide 2: CommonsenseQA Results

Time: 2 minutes

*"CommonsenseQA is where STaR truly shines.

First, it achieves **72.5% accuracy**, just shy of GPT-3's 73.0%—despite being **30 times smaller**. This shows the power of adding rationales, not just more parameters.

Second, it's impressively data-efficient. STaR uses only **86.7% of the training data**, including its self-generated rationales—meaning it reduces reliance on expensive human-labeled rationales.

And third, it even outperforms **much larger models** like LaMDA, which has 137B parameters. That's a strong indicator that structured reasoning can beat raw scale.

Visual explanation: This plot shows STaR sitting at a sweet spot in the accuracy vs. data usage tradeoff—it dominates both dimensions.

In short, rationales aren't just helpful—they can be a game-changer when it comes to scaling reasoning."*

Transition: *Now let's move on to a very different kind of reasoning—math.*

Slide 3: GSM8K Results

Time: 1.5 minutes

*"GSM8K is a benchmark of grade-school math word problems—and here, the gains are more modest but still important.

STaR improves upon few-shot baselines by **2 to 3×**, achieving **10.7% accuracy**. That may sound low, but it's a significant jump over direct fine-tuning, especially given the small size and limited supervision.

However, math requires more **symbolic and sequential reasoning**, so rationales help only slightly here—just **+0.6%** over the no-rationale variant.

The impressive part is that it does this using just **25–28%** of the training data.

Interestingly, in some examples, STaR finds **simpler or alternative solutions** compared to ground truth. That shows it's learning reasoning patterns—not just memorizing outputs."*

Transition: *So what makes STaR fundamentally different? Let's zoom into the "why."*

Slide 4: Why STaR is a Game-Changer

Time: 2 minutes

"STaR brings three major innovations to the table:

1. **Human-like Learning Cycle**:
   It mimics how we solve problems—**generate**, then **verify**, then **improve**. Crucially, it learns from **both correct and incorrect outputs**—unlike traditional fine-tuning, which learns mostly from correct pairs.
2. **Mathematical Edge**:
   The model explores —the rationale space **conditioned on both the question and answer**. This is different from the standard approach and allows for more targeted, useful reasoning paths.
   And that **custom gradient** we saw earlier acts like a filter. It only updates the model when the rationale leads to the correct answer—this gives **stronger learning signals** and helps avoid plateaus.
3. **AI Development Impact**:
   Because it learns to self-correct, we get **faster model improvement** and **less reliance on human labeling**.

In sum, STaR isn't just about generating explanations—it uses them to drive learning itself.

Visual support: The diagram here shows the bidirectional flow between rationale and answer, reinforcing this feedback loop."*

Transition: *But are these rationales actually useful to humans? Let's check.*

Slide 5: Human Evaluation

Time: 1.5 minutes

*"To answer whether STaR's rationales are truly useful, the authors conducted a **human evaluation**.

Participants were shown a QA task with either:

- No rationale,
- A rationale from few-shot prompting,
- Or a STaR-generated rationale.

They were asked to **judge the correctness** of the answer based on the explanation.

The result?
STaR rationales led to the **highest human agreement** with ground-truth answers. In other words, humans found STaR's reasoning more persuasive and clear.

This matters because rationales aren't just there to help the model—they're also there to help **us** interpret and trust the output. And STaR does both."*

Transition: *Let's wrap up with the broader picture.*

Slide 6: Conclusion & Broader Impact

Time: 2 minutes

🟦 **Key Contributions** *(<1> to <3>)*
*"To wrap up: what makes STaR significant?

First, it's the **first framework** to improve reasoning through rationale generation—a kind of **self-improvement loop** we hadn't seen before.

Second, it achieves performance competitive with models **30× larger**—while needing less data and fewer labels.

And third, it builds a path toward **transparent and explainable** AI—where we can see the 'why' behind the answers."*

🟦 **Transformative Potential** *(<4> to <6>)*
*"The broader impact here is exciting.

In sensitive domains like **healthcare, law, or education**, STaR could power **more explainable** and **auditable** decisions.

Its **rationale-first design** even helps with **bias detection**—a huge step forward for ethical AI.

And from a research angle, it opens the door to **new reinforcement learning strategies**, and applications that cross into domains like code, science, or planning."*

Quark STaR - ***Applied STaR to multimodal (text+image) reasoning.***

ReSTaR adds "memory" by storing high-quality self-generated rationales in a replay buffer, allowing the model to revisit and learn from its best reasoning traces across training iterations.

🔴 **Future Outlook** *(<8>)*
"In short, STaR gives us a **general framework** for building models that are not only more **capable** and **data-efficient**, but also more **trustworthy and interpretable**.

And that’s the kind of AI we need in the future.”
