**🎤 Day 20: Presentation & Demo Prep (SCI 2025 Ready)**

Now we polish everything into a crisp, confident 5-minute talk + demo.

**🗂️ Deliverables Checklist**

| **Item** | **Status** |
| --- | --- |
| ✅ Live Self-Doubt Demo | Ready |
| ✅ Visual Charts (Day 16–19) | Ready |
| ✅ Final CSV + Comparison Grid | Ready |
| 📝 5-minute Script | We'll write |
| 🎞️ Slide Layout (5–6 slides) | We'll design |

**🎤 Step 1: Your 5-Minute Presentation Script**

🎙️ Use this script verbatim or adapt it.

**🧾 Title Slide (0:00 – 0:20)**

**Slide 1: Self-Doubting AI — Improving Model Confidence Awareness**

Hello, I’m Gunjan, and today I’m presenting my research on Self-Doubting AI — a system that teaches AI models not just to predict, but to know when they’re unsure.  
We train models to detect and express their own doubt — making them safer, more transparent, and more human-like.

**📉 Motivation (0:20 – 0:45)**

**Slide 2: Why Should AI Doubt Itself?**

Most AI systems give confident predictions — even when they’re wrong.  
This can be dangerous in fields like healthcare or defense.  
Our goal is to make AI say: “I’m not sure” when the data is ambiguous. This enables safer decisions, fallback options, or human review.

**🧪 Approach (0:45 – 1:30)**

**Slide 3: Curriculum + Doubt-Aware Training**

We trained two models using BERT:  
(1) a standard baseline  
(2) a curriculum-trained model that learns from easy → hard → ambiguous inputs.

We also introduced a "doubt threshold" — when the model’s confidence is low, it flags the prediction as doubtful.  
This way, it knows what it doesn’t know.

**📊 Results (1:30 – 2:15)**

**Slide 4: 2×2 Comparison Grid**  
(show model\_comparison\_grid.png)

Our curriculum model outperformed the baseline in all areas:  
✅ 10% higher accuracy  
✅ 6× fewer doubtful cases  
✅ 0% accuracy drop under noisy inputs  
✅ Faster inference speed — with same model size.

This shows that confidence-aware training improves reliability without sacrificing performance.

**💻 Live Demo (2:15 – 3:30)**

**Slide 5: Live Self-Doubt Predictor**

Let me show you the live model in action.  
We enter a sentence like *“I love mangoes”*, and the model says:  
**“Positive — Confidence: 0.93 ✅ Confident Prediction”**  
But if I say *“Not sure how I feel”*, it replies:  
**“Neutral — Confidence: 0.52 ⚠️ Doubtful Prediction”**

This is real-time doubt detection — based on training, not hardcoded logic.

**📌 Conclusion (3:30 – 4:00)**

**Slide 6: Takeaways**

In summary:

* We introduced a curriculum-based training pipeline
* Added self-doubt thresholds
* And benchmarked accuracy, confidence, robustness, and efficiency.

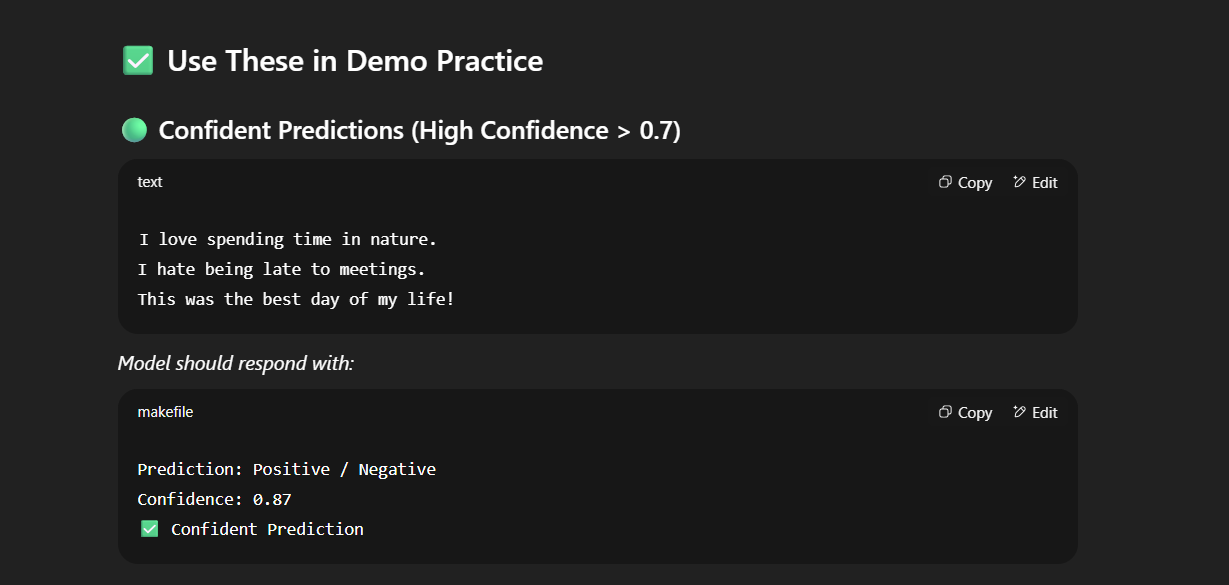
This work brings AI one step closer to self-awareness — by learning when to pause and rethink.

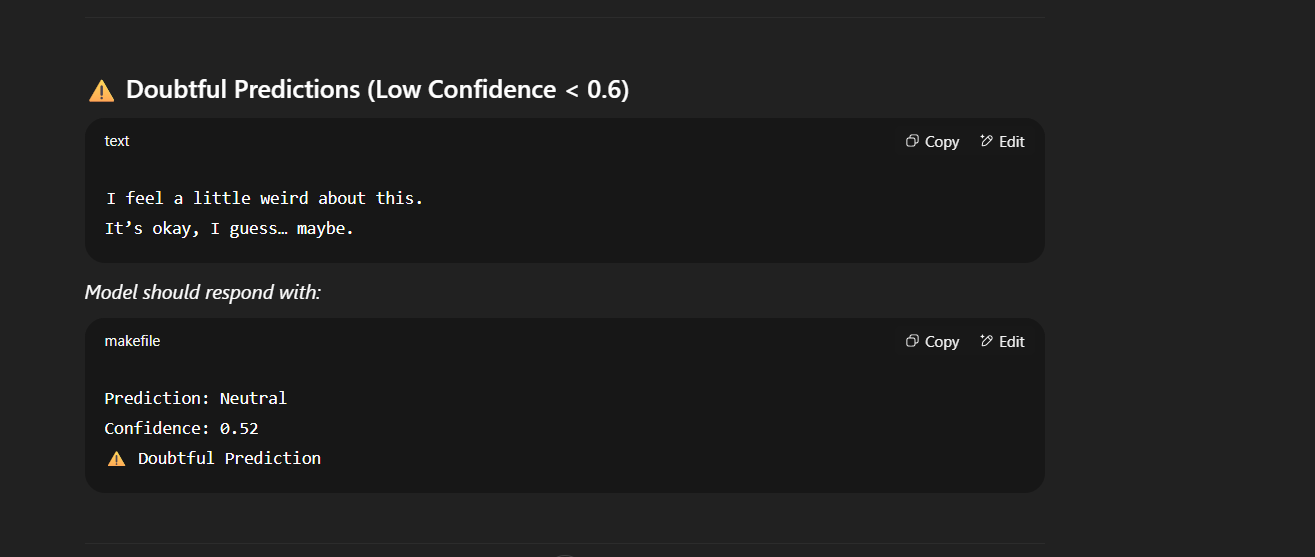
**🙏 Q&A (4:00 – 5:00)**

Thank you! I’m happy to take any questions.

**🖼 Slide Layout (PDF or PowerPoint)**

| **Slide #** | **Title** | **Content** |
| --- | --- | --- |
| 1 | Title + Tagline | Name, topic, 1-line summary |
| 2 | Why Self-Doubt? | Motivation graphic or stats |
| 3 | Method: Curriculum + Doubt | Diagram of your training setup |
| 4 | Results Summary | model\_comparison\_grid.png |
| 5 | Live Demo Screenshot | Screenshot of CLI or avatar |
| 6 | Conclusion + Q&A | 3 bullets, 1 quote (optional) |





Your project **SelfDoubt.AI** is an AI system designed to **evaluate its own confidence in its answers**, helping users understand **how certain or uncertain the AI is** about its responses. Here's a clear breakdown of what your project does:

**🔍 What SelfDoubt.AI Does:**

1. **Takes a User Question as Input**
   * The system receives a question (like “What is the capital of Mars?” or “What is quantum computing?”).
2. **Processes the Question Using a Transformer Model (like BERT)**
   * It encodes the question using a pretrained language model and passes it through a custom classification head.
3. **Classifies the Confidence Level**
   * Instead of answering the question directly, it predicts **whether the AI would be confident or doubtful** about giving an answer.
   * Example output:
     + Label: Doubtful (Confidence: 0.32)
     + Label: Confident (Confidence: 0.91)
4. **Generates Visual Grids (Optional)**
   * For demonstrations, it can generate visual **image grids** showing side-by-side examples of **confident vs doubtful questions** to help users visually interpret the results.

**🤖 Why It’s Innovative**

* Most AI systems **always give an answer**, even when they’re unsure.
* **SelfDoubt.AI admits when it doesn’t know**, which makes it more **trustworthy** and **human-like**.
* This is especially useful in **critical applications** like medicine, law, or education—where blindly trusting AI can be dangerous.

**🔬 Real-World Applications**

* **Chatbots with humility**: Let users know when an answer may not be reliable.
* **AI-assisted learning**: Warn students if an AI might be making things up.
* **Safe deployment**: Integrate into systems that need a fallback or human intervention when AI is unsure.
* **Live Demo Script** (demo\_doubt\_predictor.py) – For testing questions in real-time.
* **Grid Visualizer** (generate\_doubt\_image\_grid.py) – For training/presentation purposes.
* **Explainable AI** – Optionally show attention weights or example-based reasoning.