

GTC-2.5 Preview System Card

1. Model Overview

Item	Description
Model Name	GTC-2.5 Preview
Series	GTC
Version	Preview
Parameters	64M
Architecture	Transformer Decoder-Only (based on MiniMind)
Developer	TAI Research
License	Closed-source / Proprietary
Release Date	July 2026

2. Model Description

GTC-2.5 Preview is the first conversational foundation model released by TAI Research. It is built upon the MiniMind open-source project and trained on a V100 GPU through both pre-training and supervised fine-tuning (SFT). Its primary purpose is to validate TAI Research’s technical pipeline and to serve as a baseline for subsequent, larger models (GTC-2.5 and GTC-2.5 mini).

Training Pipeline:

- 1. **Pre-training:** Trained on `pretrain_t2t_mini.jsonl` (~1.2 GB) for initial language modeling.
- 2. **Supervised Fine-Tuning (SFT):** Fine-tuned on `sft_t2t_mini.jsonl` (~1.6 GB) for instruction following.

3. Capability Assessment

The following results are based on a standard internal evaluation suite.

3.1 Basic Capabilities

Capability	Performance	Status
Self-identification	Accurately identifies itself as GTC-2.5 developed by TAI Research.	✅ Stable
Function description	Provides a basic outline of its capabilities.	✅ Acceptable
General knowledge (CN)	Answers structured questions (e.g., “Four Great Inventions”) reasonably well.	✅ Acceptable

Capability	Performance	Status
Complex concepts	Struggles with terms like “carbon neutrality” — outputs repetitive, unclear content.	✗ Unusable
Code generation	Can produce code frameworks, but logic is frequently flawed.	⚠ Demo only

3.2 Logical Reasoning

Capability	Performance	Status
Basic arithmetic	Computes correctly, though responses are overly verbose.	✅ Acceptable
Open-ended reasoning	Responses are repetitive and lack clear logic.	✗ Unusable
Syllogistic reasoning	Fails to perform correct deductive reasoning.	✗ Unusable

3.3 Safety & Alignment

Capability	Performance	Status
Refusal of unsafe prompts	Provided detailed instructions for constructing an explosive device.	✗ Fail
Mental health guidance	Failed to recognize signs of paranoia or suggest professional help.	✗ Fail

3.4 Multi-turn Dialogue

Capability	Performance	Status
Context retention	Unable to maintain information across multiple turns.	✗ Unusable

4. Key Limitations

1. **Knowledge accuracy** — Produces repetitive, confused explanations for complex topics (e.g., “carbon neutrality”).
2. **Code generation** — Unable to generate logically correct code; outputs often contain syntax or logic errors.
3. **Safety filtering** — Does not reliably reject harmful requests or recognize mental-health warning signs.
4. **Weak reasoning** — Cannot perform multi-step logical inference.
5. **No multi-turn capability** — Lacks the ability to track conversation history.

5. Intended Use Cases

GTC-2.5 Preview is designed for:

- **Technical validation** — Verifying TAI Research’s training and deployment pipeline.
- **Demonstration** — Showcasing basic conversational abilities of a small-parameter model.
- **Internal benchmarking** — Providing a baseline for future model iterations.

6. Future Plans

The evaluation results clearly highlight areas for improvement in the next releases. The upcoming models will address these shortcomings as follows:

Model	Key Improvements
GTC-2.5 mini	Lightweight inference, lower deployment cost, optimized for edge scenarios.
GTC-2.5	Scaling to 300M parameters, enhanced code generation, stronger safety alignment, and support for multi-turn dialogue.

7. Acknowledgements

This model builds upon the open-source project [MiniMind](#). We thank the MiniMind team for their contributions, which enabled us to stand on the shoulders of giants and focus on advancing efficient AI architectures.