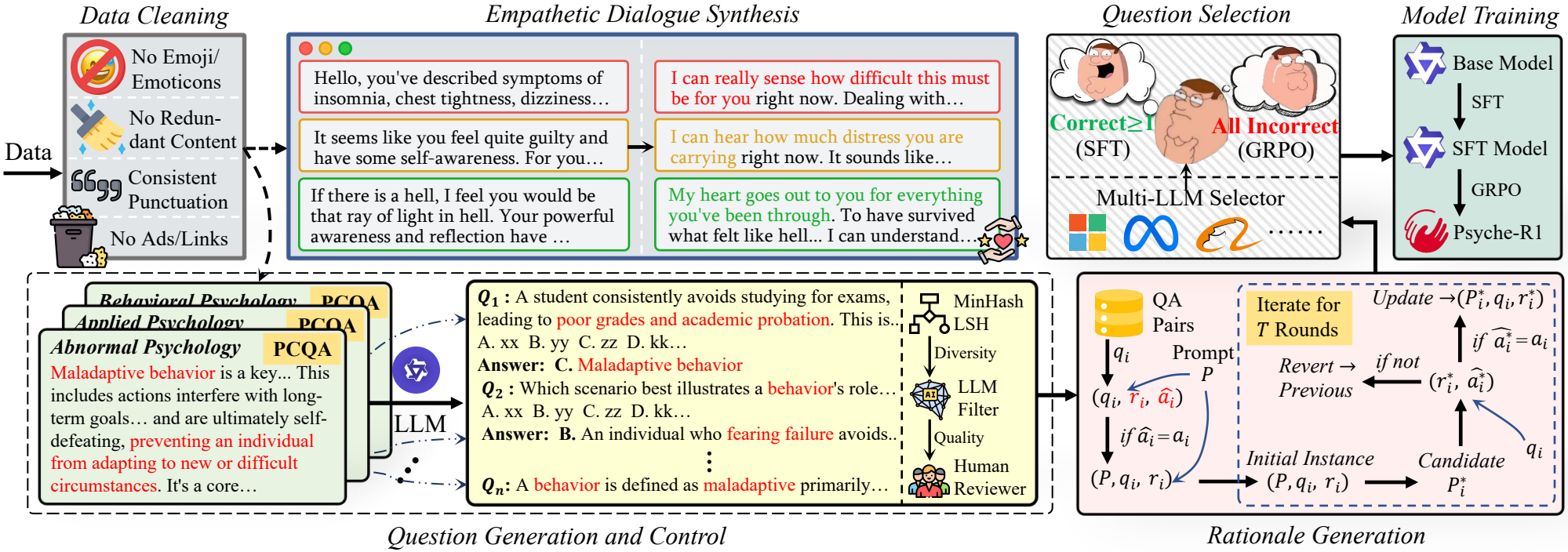
Data Cleaning
Empathetic Dialogue Synthesis
Question Selection
Model Training
Data
No Emoji/ Emoticons
No Redundant Content
Consistent Punctuation
No Ads/Links
Hello, you've described symptoms of insomnia, chest tightness, dizziness…
I can really sense how difficult this must be for you right now. Dealing with…
It seems like you feel quite guilty and have some self-awareness. For you…
I can hear how much distress you are carrying right now. It sounds like…
If there is a hell, I feel you would be that ray of light in hell. Your powerful awareness and reflection have …
My heart goes out to you for everything you've been through. To have survived what felt like hell... I can understand…
Correct≥1 (SFT)
All Incorrect (GRPO)
Multi-LLM Selector
Base Model
SFT
SFT Model
GRPO
Psyche-R1
Behavioral Psychology PCQA
Applied Psychology PCQA
Abnormal Psychology PCQA
Maladaptive behavior is a key... This includes actions interfere with long-term goals… and are ultimately self-defeating, preventing an individual from adapting to new or difficult circumstances. It's a core…
LLM
Q_1 : A student consistently avoids studying for exams, leading to poor grades and academic probation. This is..
A. xx B. yy C. zz D. kk…
Answer: C. Maladaptive behavior
Q_2 : Which scenario best illustrates a behavior's role…
A. xx B. yy C. zz D. kk…
Answer: B. An individual who fearing failure avoids..
Q_n: A behavior is defined as maladaptive primarily…
MinHash LSH
Diversity
LLM Filter
Quality
Human Reviewer
QA Pairs
q_i
Prompt P
$(q_i, r_i, \hat{a}_i)$
if $\hat{a}_i = a_i$
(P, q_i, r_i)
Iterate for T Rounds
Update $\rightarrow (P_i^*, q_i, r_i^*)$
if $\hat{a}_i^* = a_i$
Revert → Previous
if not
$(r_i^*, \hat{a}_i^*)$
Initial Instance (P, q_i, r_i)
Candidate P_i^*
q_i
Question Generation and Control
Rationale Generation