SIGNATE × TECH OCEAN Student Cup 2025

２位入賞解法

Team: 2%
Member: 種村玲音　kou1218

本アプローチの結論

本アプローチのカギは、**「高性能API（Gemini Pro / NotebookLM）の活用」**と**「アンサンブルによる出力の安定化」**にある。



推論エンジンの最適化

Web版Gemini ProやNotebookLMを使用し、高度な読解・推論能力を引き出す



ロバスト性の確保

複数の推論結果を用いたアンサンブルと、信頼度の低い予測に対する再推論による、極限のブレ制御

Public Score：0.6764
Private Score：0.6441

初期アプローチの失敗と得られた気づき

ベクトル類似度検索
ローカルLLM
疑似テストデータ作成

→

✕ 正解率0.2未満



表面上の単語一致（キーワードマッチング）や、
パラメータ数の少ないモデルでは、深いテーマ性や文章のねじれを到底解決できない。

→

推論エンジンをWeb版APIに切り替え、
文脈理解力と精度が大幅上昇

Public Score：0.5294

スコア分析から導いた仮説と戦略

仮説１：Baseデータの情報不足

・既存のあらすじのみでは、キー要素が欠落しており特定困難

例）「君の名は」

Testデータの多くに「結び」という単語があるが、

元のBaseデータには存在していない。

→ あらすじ拡張

（情報不足の補完）

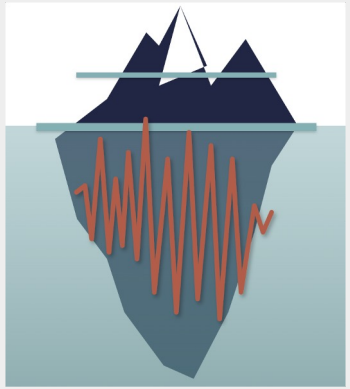
仮説２：Public スコアの罠と大幅下落（Shake）

Publicデータ（全体の30%）

プロンプトや条件の変化によって、
推論結果が100件以上変化しても
Public Scoreは変動しなかった

Privateデータ（残りの70%）

単一ファイル提出は見えないデータで
致命的にブレる危険性が極めて高い

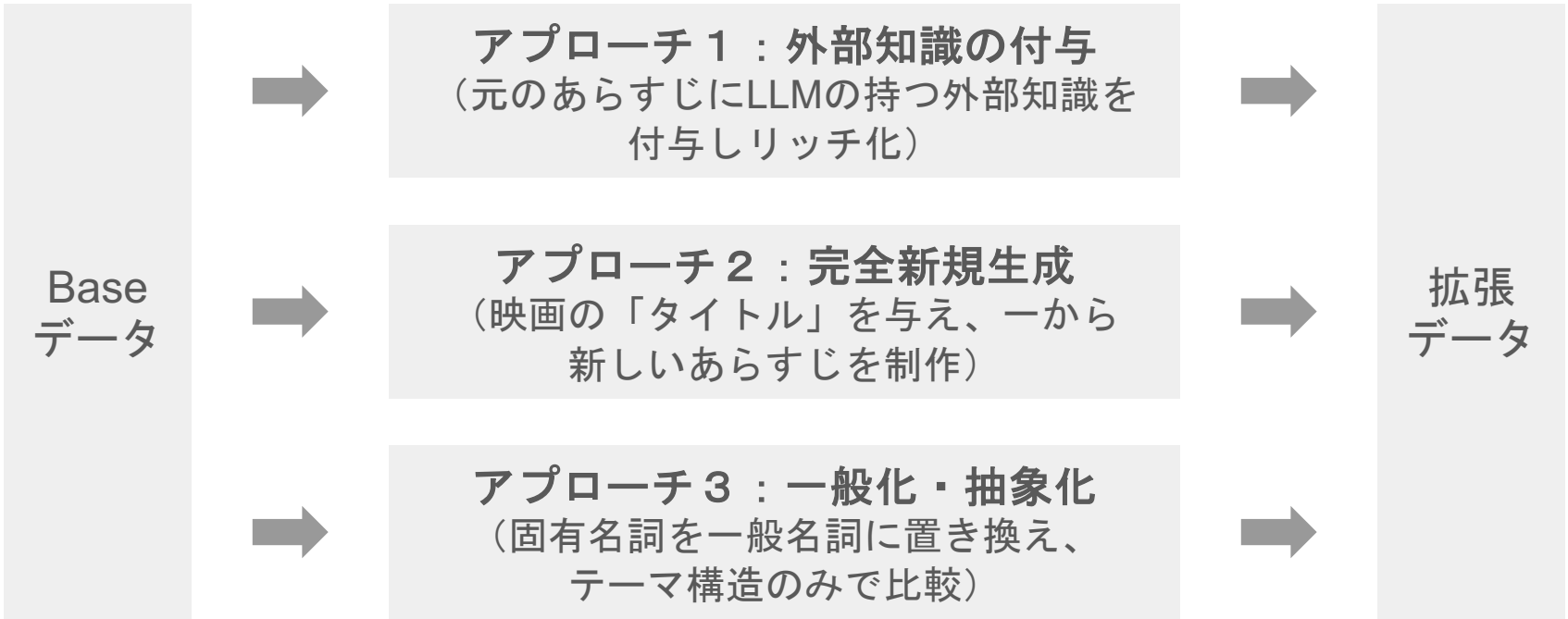


→ アンサンブル

（揺らぎの吸収・制御）

【工夫①】あらすじ拡張による情報不足の補完

目的：Baseデータに欠落している推論のヒント（知識）を補う



【工夫②】NotebookLMを活用した「知識の遮断」と「論理推論」

目的：LLM特有のハルシネーションや不要な事前知識によるノイズを完全遮断

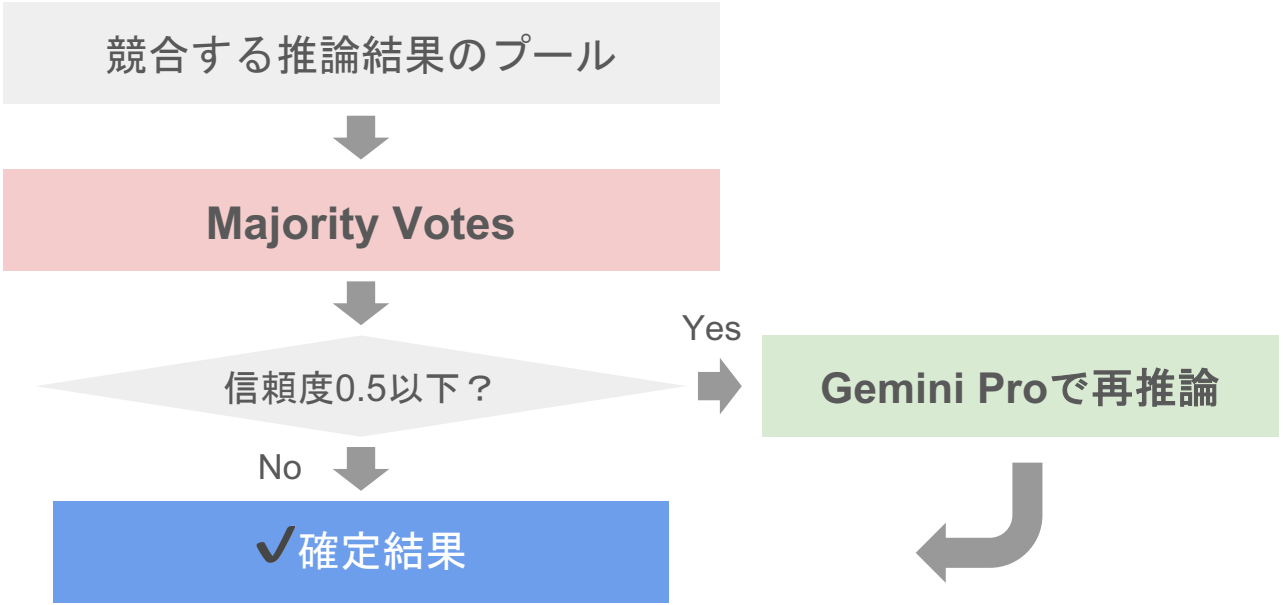
「ソースのみを参照する」NotebookLMの特性を利用



【工夫③】アンサンブルと再推論による揺らぎの制御

目的：単一モデルによるPrivateスコア下落リスクを回避

10ファイル以上の推論結果を統合（Gemini Pro, ChatGPT, NotebookLM　＋　拡張あらすじ数種類）



全体まとめとスコアの推移

DATA
（文脈の補完）

データ拡張により推論に必要な情報の欠損を充足

MODEL
（外部ノイズの遮断）

NotebookLMでハルシネーションを防ぎソースに集中

OUTPUT
（揺らぎの制御）

10種以上のファイルアンサンブルでロバスト性を確立

