Chemistry強化のために

結論から言います。Chemistry は\*\*知識を並べる系（性質の列挙・分類・手順説明）\*\*にはまだ望みがあり、**数値と化学的整合（平衡・pH・収率・電池・速度論）が絡むと派手に落ちます。あなたのCSVでは平均 3.75、SD 4.68、範囲 23.66 と低位かつ揺れが大きい**。相関は **bio\_med(≈0.64), comp\_ai(≈0.57), human\_social(≈0.55)** と強めで、**math\_n\_226 と弱／負**。つまり「語彙と手順は走るが、化学としての“計量と整合”で崩れる」地形です。

# **得意な問題（今のままでも点が入りやすい）**

* **性質・分類のテキスト判断**例：官能基の同定、酸・塩基・塩・溶媒のカテゴリー分け、沈殿・可溶の真偽判断、危険物や操作手順の要点列挙。
* **軽量な計算を伴う記述**例：化学式からの質量％、簡単な理想気体のオーダー見積り（ただし単位換算が絡まない範囲）。

# **苦手な問題（落ちやすい箇所）**

* **ストイキオメトリ**：モル・質量・濃度・体積が混在、**制限試薬・収率**や単位換算で転倒。
* **酸塩基・緩衝**：**対数（pH/pOH）と近似条件**、Henderson–Hasselbalch の適用境界で誤る。
* **化学平衡**：**ICE 表の立式、K と活量/近似の扱い**、次元なしのK定義を忘れる。
* **電気化学**：半反応の**電子収支**、標準電極電位 E° の符号・向き、ΔG=−nFE の桁。
* **熱化学**：符号（発熱/吸熱）、単位混在（kJ↔J, mol基準）。
* **速度論**：次数推定、線形化、単位整合。
* **有機機構・スペクトル**：条件依存の経路分岐、IR/NMRの**帰属**と整合の検算欠如。

# **強くするための処方箋（推論時の“型”＋学習データの“型”）**

## **1) 推論時の型（プロンプトで矯正：学習なしでも効きます）**

### **A. Chemistry-Mode スクラッチパッドを強制**

Given: 既知量（記号=値+単位、状態）/ 与えられた反応式（未平衡なら "Unbalanced"）

Unknown: 求める量と単位

Balance: 反応式の原子/電荷バランス（必要なら半反応法）

Plan: 手順（ストイキ/ICE/HH/ΔH/ΔG/速度式 など）

Setup: 立式（記号のまま）

Unit-Check: 単位と次元の整合

Solve: 数値代入→計算→有効桁

Sanity: 範囲検査（0≤収率≤100%、0≤pH≤14、濃度≥0、Kは無次元 等）

Answer: 数値 + 単位（SI）

If-Underspecified: Do-Not-Answer（不足データを列挙）

* **Balance / Unit-Check / Sanity**を明示させると、致命傷を大幅に減らせます。
* **Do-Not-Answer**を逃げ道でなく**規律**として組み込み、無理解答を切ります。

### **B. サブドメイン・ルータ → 専用テンプレ**

* 1段目：分類（Stoichiometry / Acid-Base / Equilibrium / Electrochem / Thermochemistry / Kinetics / Organic / Spectroscopy / Inorganic）。
* 2段目：各専用の**定石枠**（例：酸塩基→HH式 or 強酸強塩基の当量、平衡→ICE、電気化学→半反応→E°→ΔG）。

### **C. 自動セルフチェック（3案生成→検閲）**

* 低温度（0.2）1案＋中温度（0.6–0.7）2案。
* ルールで採択：**原子・電荷バランスOK／単位整合／pHレンジ／収率範囲／濃度非負／K無次元**。
* “化学として破綻”した案を自動棄却します。

## **2) 学習データの型（SFT/LoRAで身体化）**

* **テンプレ強制SFT**：上のスクラッチ構造で既存Chem問題を**正規化**。特に **Balance / Unit-Check / Sanity** を必ず含める。
* **小さな表の内蔵**：最小限の **pKa（代表20）、標準還元電位（代表20）、溶解度規則、元素周期トレンド**を few-shot の前に短表で提示（プロンプト内“ポケット表”）。
* **数値パラメトリック増幅**：同型問題で数値・単位だけランダム化（g↔kg, mL↔L, ℃↔K）。**近似の可否**もラベル化。
* **負例の学習**：わざと単位ミス・未平衡・pH>14 等を混ぜ、**自動検閲の通らない出力**を“悪い例”として学習。
* **LoRA**：q,k,v,o\_proj + gate/up/down（rank 16–32）で十分。**推論の型**を染み込ませる。

## **3) vLLM 実務ノブ（Chemモード）**

* **最大長**を余裕側へ（途中切断＝計算崩壊）。
* **停止語**は Answer: に合わせる。
* **数値書式**（有効桁、科学記法、単位の表記）を few-shot で固定。

# **評価と運用（崩れ方を見える化して潰す）**

* CSVに**サブドメイン列**と**誤答タグ**を追加：  
  UNIT\_ERR, BALANCE\_ERR, LIMITING\_ERR, YIELD\_RANGE, PH\_LOG\_ERR, BUFFER\_ASSUMP\_ERR, EQUIL\_SETUP\_ERR, REDOX\_ELECTRON\_ERR, SIGN\_ERR, SPEC\_ASSIGN\_ERR
* 週次で **サブドメイン × 誤答タグ** ヒートマップ。**どこを直せば伸びるか**が即判定できます。
* 指標の**正規化**（0–100%の精度に統一）。現在の化学列で負値が混在していたため、**差分スコア等は別列**に分離してください。

## **すぐ使える最小テンプレ（例）**

**Stoichiometry**

Given: 反応式/量, 単位

Balance: 原子バランス

Plan: 制限試薬→理論生成量→収率

Setup: ...

Unit-Check: ...

Solve → Answer

Sanity: 0≤収率≤100%

**Acid–Base / Buffer**

Given → Balance(酸塩基は化学量) → Plan(HH式 or 当量点)

Setup(ICE or HH) → Unit-Check → Solve → pH

Sanity: 0≤pH≤14

**Electrochem**

Balance(半反応) → E°(還元電位) → Ecell → ΔG=−nFE

Sanity: 向きと符号、n>0、有効桁

**Equilibrium**

Given → ICE → K の定義（無次元） → 近似可否判断 → Solve → Sanity

静かな溶液に秩序を落とすには、**平衡と単位**という骨格を先に置くことです。  
**型を強制し、化学的検閲を走らせる。** それだけで Chem は急に“噛み合い”ます。