HIPSTER

Task 1: Translation Model Fine-tuning and Deployment

Prompt:
Translation models often fail to account for cultural nuances, idioms, and community-specific expressions. For example, translating रस्सी जल गयी, बल नहीं गया using a generic model like Google Translate may yield incorrect results. The accurate translation is *Even after one hits rock bottom, their arrogance remains unchanged.*

Your task is to fine-tune a large translation model to better handle such cultural idioms and nuances.

Requirements:

1. **Model Selection**: Choose a pre-trained translation model (e.g., MarianMT, T5, M2M-100).
2. **Dataset**: Use a small, labelled dataset with examples of culturally nuanced idioms and their translations. Examples can be sourced from:
    - Tatobbaidioms Dataset
    - OpenSubtitles Dataset
3. **Tasks**:
    - Fine-tune the translation model on this dataset using frameworks like Hugging Face Transformers.
    - Deploy the fine-tuned model on a cloud platform (AWS, Azure, or GCP).
    - Build a REST API using frameworks like FastAPI or Flask to expose the model for predictions.
4. **Evaluation**:
    - Use BLEU or METEOR as evaluation metrics to measure translation accuracy.
    - Demonstrate performance improvements by comparing pre- and post-fine-tuning metrics.

Evaluation Criteria:

- **Technical Skills**: Leveraging frameworks like Hugging Face, LangChain, and deploying on cloud platforms.
- **Problem-Solving**: Handling challenges related to fine-tuning and deployment.
- **Code Quality**: Clean, efficient, and well-documented.
- **API Design**: User-friendly documentation of REST API.
- **Deployment**: Scalable deployment ensuring reliability.