Mountain NER Project: Improvement

Areas for Improvement

1. Model Architecture and Training

- Hyperparameter Optimization:

    - Use automated hyperparameter tuning (e.g., Optuna, Ray Tune)
    - Optimize focal loss parameters
    - Experiment with different learning rates and scheduling strategies

- Model Architecture Enhancements:

- Evaluate newer transformer architectures (DeBERTa, RoBERTa)

- Implement model distillation for faster inference

- Consider multi-task learning to jointly predict mountain properties

2. Data Quality and Quantity

- Data Augmentation:

    - Implement synonym replacement for non-mountain words
    - Use back-translation for sentence variation
    - Create synthetic examples using templates

- Dataset Expansion:

    - Include multilingual mountain names
    - Add geographical context (elevation, location, range)
    - Incorporate different naming conventions and local variations
    - Simply make it bigger

3. Feature Additions

1. Extended Entity Information:

- Mountain height classification
- Mountain range association
- Geographical location tagging

2. Advanced Recognition:

- Multi-language support
- Handling of abbreviations and local names
- Recognition of mountain-related terminology