

criteria

Harito ID

2025-11-22

Lab 5 Evaluation Criteria: RNNs for Text and Token Classification

Part 1: Tìm hiểu về RNNs và Phân loại Token (15%)

- Nghiên cứu lý thuyết:**
 - Đọc và nắm vững các khái niệm về RNNs, LSTMs, và GRUs được trình bày trong lab5_pytorch_introduction.pdf.
 - Hiểu rõ kiến trúc và cách hoạt động của mô hình RNN cho bài toán phân loại token (Token Classification) qua file lecture5_rnn_token_classification.pdf.
- Thực hành Pytorch:**
 - Hoàn thành các bài tập trong lab5_pytorch_introduction.md để làm quen với Pytorch.

Example: 0 / 3 items completed.

Part 2: RNNs cho Phân loại Văn bản (35%)

- Task 1: Baseline Model 1 (TF-IDF + Logistic Regression)**
 - Xây dựng và huấn luyện thành công mô hình.
 - Ghi nhận lại kết quả Accuracy/F1-score.
- Task 2: Baseline Model 2 (Word2Vec + Dense Layer)**
 - Xây dựng và huấn luyện thành công mô hình.
 - Ghi nhận lại kết quả Accuracy/F1-score.
- Task 3: LSTM Model with Pre-trained Embeddings**
 - Xây dựng và huấn luyện thành công mô hình.
 - Sử dụng word embeddings đã được huấn luyện trước.
- Task 4: LSTM Model with End-to-End Training**
 - Xây dựng và huấn luyện thành công mô hình với Embedding Layer được học từ đầu.
- Task 5: So sánh và Phân tích**
 - So sánh hiệu năng của 4 mô hình trên.
 - Viết báo cáo phân tích ngắn gọn về ưu và nhược điểm của từng phương pháp, đặc biệt là sức mạnh của mô hình chuỗi (LSTM) trong việc “hiểu” ngữ cảnh.

Example: 0 / 9 items completed.

Part 3: Part-of-Speech Tagging với RNN (25%)

- Task 1: Chuẩn bị dữ liệu**

- Tải và tiền xử lý bộ dữ liệu UD-English.
- Xây dựng vocabulary cho từ và cho nhãn POS.
- Task 2: Xây dựng mô hình RNN**
 - Implement pipeline Embedding, RNN/LSTM, và Linear cho bài toán POS tagging.
- Task 3: Huấn luyện và Đánh giá**
 - Huấn luyện mô hình trên tập train.
 - Đánh giá hiệu năng trên tập test (sử dụng accuracy hoặc F1-score).

Example: 0 / 5 items completed.

Part 4: Named Entity Recognition với RNN (25%)

- Task 1: Chuẩn bị dữ liệu**
 - Sử dụng một bộ dữ liệu NER phổ biến (ví dụ: CoNLL-2003).
 - Tiền xử lý và tạo vocabulary cho từ và nhãn NER.
- Task 2: Xây dựng mô hình Bi-LSTM**
 - Implement mô hình Bi-LSTM (hoặc Bi-LSTM-CRF nếu có thể).
- Task 3: Huấn luyện và Đánh giá**
 - Huấn luyện mô hình.
 - Đánh giá và báo cáo kết quả trên tập test.

Example: 0 / 5 items completed.

Yêu cầu chung về Báo cáo

- Cấu trúc rõ ràng:** Báo cáo cần có các phần tương ứng với 4 phần của bài lab.
- Giải thích code:** Mô tả ngắn gọn các bước implement quan trọng.
- Hướng dẫn chạy code:** Cung cấp hướng dẫn để có thể chạy lại các thí nghiệm.
- Phân tích kết quả:** Đưa ra các bảng kết quả, biểu đồ (nếu có) và nhận xét, phân tích sâu sắc.
- Nêu khó khăn và giải pháp:** Ghi lại những thách thức gặp phải và cách bạn đã vượt qua.

Example: 0 / 5 items completed.

Scoring Formula

The final score is calculated as follows:

$$Score = \left(0.5 \times \frac{\text{CodeTasks}}{22} + 0.5 \times \frac{\text{ReportTasks}}{5} \right) \times 10$$

Where:

- `CodeTasks` is the number of completed items in Part 1.
- `ReportTasks` is the number of completed items in Part 2.

The score will be rounded to the nearest 0.25.

Note:

- The above criteria are mandatory.
- Late submissions will be penalized.
- You must cite any external sources, models, libraries, or tools used.