

criteria

Harito ID

2025-09-18

Bộ tiêu chí đánh giá

Code chạy đúng và không bị lỗi (50%)

- ☒ Interface cho Tokenizer và Vectorizer
- ☒ Triển khai SimpleTokenizer
- ☒ Triển khai RegexTokenizer
 - ☐ Test SimpleTokenizer và RegexTokenizer cho 1 số ví dụ cơ bản
 - ☐ Test SimpleTokenizer và RegexTokenizer cho dataset UD English EWT
 - ☐ Triển khai CountVectorizer
 - ☐ Test CountVectorizer cho 1 số ví dụ cơ bản
 - ☐ Test CountVectorizer cho dataset UD English EWT

Eg: 3 / 8

Viết báo cáo (50%)

- ☒ Nêu rõ các bước triển khai
- ☒ Cách chạy code và ghi log kết quả
 - ☐ Giải thích các kết quả thu được
 - ☐ Nêu rõ các khó khăn gặp phải và cách giải quyết
 - ☐ Nếu tham khảo các nguồn bên ngoài, cần ghi rõ nguồn tham khảo
 - ☐ Nếu sử dụng các model tạo sẵn, cần ghi rõ model nào, từ đâu, prompt

Eg: 2 / 6

Tổng điểm

Trong ví dụ trên tính là:

$$Score = \left(0.5 \times \frac{3}{8} + 0.5 \times \frac{2}{6}\right) \times 10$$

$$Score = (0.5 \times 0.375 + 0.5 \times 0.3333...) \times 10$$

$$Score = (0.1875 + 0.1666...) \times 10$$

$$Score = 0.354166... \times 10$$

$$Score \approx 3.54 \approx 3.5$$

Điểm đơn vị: 0.25