Test NLP Processor

Tests whether the NLP Processor matches words with similar meaning correctly, and handles edge cases effectively.

Blackbox Tests:

Expected:

T01. comparison of empty lists of keywords returns score of 0

T02. comparison with/against an empty list also returns score 0

T03. comparison of same words in different tense should result

high score

T04. comparison of same words in different case should result in

in a score of 100%

T05. comparison of same words surrounded by different puntuation

should result in a score of 100%

T06. comparison of same words should result in a score of 100%

T07. comparison of words with whitespaces should result in a

score of 0%

Whitebox and Mocking Tests:

Expected:

T08. successful creation of nlp processor

T09. using mock tokens and simplified similarity logic

test if the comparison score is calculated

appropriately (mocked to isolate scoring logic)

a. comparison of source keywords with target

b. comparison of target keywords with source

(checking for source subset of target, so

scores will change)

T10. using mock tokens and simplified similarity to check

if only tokens with similarity more than or equal to match are

considered for scoring (mock to isolate token factoring logic)

->BVA on match percentage

Boundary = match percentage value (3-way BVA)

BVA points =>

a. before match percentage value

b. at match percentage value

c. after match percentage value

T11. test in case comparison raises error

Tests Implementation: https://github.com/SGTK06/Job-Application-Tracker/blob/f39cf96de8ea966d38b4f4f4a5d0883a3db42b8e/tests/test_nlp_processor.py

Test Report:

Test ID	Input	Expected Output	Actual Output	Pass/Fail
T01	empty lists	Score 0%	Score 0%	PASS
T02a	empty source keyword list	Score 0%	Score 0%	PASS
T02b	empty target keyword list	Score 0%	Score 0%	PASS
T03	Words with different tenses	Score > 70%	Score > 70%	PASS
T04	Words with different case	Score 100%	Score 100%	PASS
T05	Words with punctuation	Score 100%	Score 100%	PASS
T06	Same words	Score > 90%	Score > 90%	PASS
T07	Only spaces and blanks	Score 0%	Score 0%	PASS
T08	Create NlpProcessor object	processor created successfully	processor created successfully	PASS
T09a	Mock isolated scoring (from source to target)	Score 100%	Score 100%	PASS
T09b	Mock isolated scoring (switch to from toarget to source to check if score percentage changes correctly)	Score 50%	Score 50%	PASS
T10a	Mock tokens to get predictable match percentage and BVA before boundary (comparison score 69, threshold 70)	Score 0%	Score 0%	PASS

T10b	Mock tokens to get predictable match percentage and BVA at boundary (comparison score 70, threshold 70)	Score 100%	Score 100%	PASS
T10c	Mock tokens to get predictable match percentage and BVA after boundary (comparison score 71, threshold 70)	Score 100%	Score 100%	PASS
T11	Mock comparison to raise error	Score 0%	Score 0%	PASS

Mocking Usage Documentation:

The actual spacy NLP model has been mocked with a different function to convert words into fake tokens, or erroneous.

The fake tokens return a fixed value on comparison which makes comparison logic predictable and allows checking using BVA

The errored tokens raise error on comparison allowing checking of the error handling logic of NLP processor.