CREA\_library Methods

|  |  |  |  |
| --- | --- | --- | --- |
| Method | Description | Parameters | Output |
| CREA\_library() | Create an instance of CREA\_library | dict | None |
| get\_all\_vectors() | Get all vectors from dataset. | None | dict: All word vectors |
| get\_vector(word) | Gets the vector for a specific word. | word (str): Target word | list or None: Word vector if found |
| get\_vectors(words) | Gets the vectors for a list of words. | words (list): List of words | dict: Word-vector pairs for found words |
| select\_cols(words, columns) | Selects specific columns from word vectors. | words (list), columns (list) | dict: Words with selected columns |
| cosine\_similarity(vec1, vec2) | Computes cosine similarity between two words or vectors. | vec1 (str or list), vec2 (str or list) | float: Cosine similarity score |
| top\_n\_similar(word, n=5) | Finds the top n most similar words to a given word. Default 5 | word (str), n (int, default=5) | list: Top n most similar words with scores |
| strip(file\_name) | Removes category names from a JSON file, returning float-only vectors. | file\_name (str): JSON file path | list: Processed word vectors |
| get\_raw(file\_name) | Extracts raw CREA dataset into a CSV file. | file\_name (str): JSON file path | Writes to results.csv |
| load\_word\_from\_json(file\_name) | Loads a dictionary of word vectors from a JSON file. | file\_name (str): JSON file path | dict: Word vectors |
| calculate\_averages(file\_name) | Calculate the averages of the CREA dataset | file\_name (str): CSV file path | df: Averages calculated from raw scores |

strip is useful if the vectors have the category tags with them. For example:

{'molecule': {'Vision': 0.53125,

'Bright': 0.96875,

'Dark': 0.71875,

'Color': 1.1875,

'Pattern': 1.625,

'Large': 0.125},

‘automobile: {'Vision': 5.53,

'Bright': 1.9,

'Dark': 1.83,

'Color': 2.13,

'Pattern': 2.7,

'Large': 4.43 }}

Can be changed to:

{'molecule': [0.53125, 0.96875, 0.71875, 1.1875, 1.625, 0.125],

'automobile': [5.53, 1.9, 1.83, 2.13, 2.7, 4.43]}

This can then be initialized by CREA\_library. Now the remaining methods can be used.

Examples:

stripped\_dict = CREA\_library.strip('file\_path.json’)

word\_dict = CREA\_library(stripped\_dict)

word\_dict.get\_all\_vectors()

single\_word = word\_dict.get\_vector('word1')

multiple\_words = word\_dict.get\_vectors(['word1', 'word2', ... , 'wordn'])

specific\_cols = word\_dict.select\_cols(['word1'], [0, 1, 2])

similar = word\_dict.cosine\_similarity('word1', 'word2')

top\_similar = word\_dict.top\_n\_similar('word1', 5)

If getting data from psychopy and getting raw df or calculate averages

fname = 'results.json'

CREA\_library.get\_raw(fname)

with open('results.csv', 'r') as f:

df = pd.read\_csv(f)

CREA\_library.calculate\_averages('results.csv')

Exports data from psychopy named ‘results.json’ to a csv file that it updated with each new output of the experiment