Input: Emotion labeled NRC Lexicon, regional word embeddings

Select E as the subset of the labeled NRC lexicon where each word w_i is marked with at least one emotion, and $\{w_i\}$ are unique after our text normalization and processing.

	anger	fear	anticip.	trust	surprise	sadness	joy	disgust
w_1	1	0	0	0	1	0	0	0
$\vdots$								
w_n	0	1	0	0	0	1	0	0

Each region is then represented by a projection induced by the word embedding subset $\{w_i\}$ as follows. Let $w \in \{w_i\}$, and π_w^{-1} be the inverse of the permutation that sorts $\cos(w_1, w), \cdots, \cos(w_n, w)$; π_w^{-1} is then the representation of w.
Finally, a region embedding X is composed by all word representations $\pi_{w_i}^{-1}$

$$\cos(u, v) = \frac{\sum_{i=1}^{n} u_i \cdot v_i}{\sqrt{\sum_{i=1}^{n} u_i^2}\sqrt{\sum_{i=1}^{n} v_i^2}}$$

The distance between two regions word embeddings X, Y is then determined as follows;

$$d(X, Y) = \sum_{j=1}^{n} \sqrt{\sum_{i=1}^{n} (X_{i,j} - Y_{i,j})^2}$$

$$X_{n \times n} = \begin{pmatrix} \pi_{w_1}^{-1} & \cdots & \pi_{w_n}^{-1} \end{pmatrix}$$

The i-th coordinate of π^{-1} contains the index where the i appears in π. We use each π^{-1} as a vector to create the regional matrix.

Computes $d(X, Y)$ for all regions to obtain the affinity matrix. Normalize to 1 dividing all distances by the maximum distance in the matrix.

Output: Affinity matrix of regional embeddings.