**[0129]** In some non-limiting embodiments or aspects, applying mean centering may include determining a mean based on all embedding vectors of the set of embedding vectors. Additionally or alternatively, the mean may be subtracted from each embedding vector of the set of embedding vectors.

**[0130]** In some non-limiting embodiments or aspects, embedding normalization/alignment system 102a may apply geometric median normalization. For example, applying geometric median normalization may include determining (e.g., by embedding normalization/alignment system 102a) a geometric median (x*) based on the embedding set and/or normalizing (e.g., by embedding normalization/alignment system 102a) each embedding vector of the embedding set based on the geometric median (x*). For example, determining the geometric median may include inputting the embedding set into a Weiszfeld algorithm to determine the geometric mean. Additionally or alternatively, normalizing each embedding vector may include replacing each respective embedding vector with a respective modified embedding vector determined based on subtracting the geometric median (x*) from the respective embedding vector to provide a respective difference and dividing the difference by a magnitude (e.g., vector magnitude) of the difference. For the purpose of illustration, applying geometric median normalization may include applying the following algorithm, where A is an embedding set, $a_i$ is an ith embedding vector of the embedding set A, x* is a geometric median, and Weiszfeld() is a Weiszfeld algorithm:

**Algorithm 1**

1: $x^* \leftarrow \text{Weiszfeld}(A)$

2: **for all** $a_i \in A$ **do** $a_i \leftarrow \dfrac{a_i - x^*}{\| a_i - x^* \|}$

3: **return** $A$

For the purpose of illustration and not limitation, a Weiszfeld algorithm may include applying the following algorithm, where $a_i$ is an ith embedding vector of the embedding set A (e.g., $a_1$ through $a_n$), $x_0$ is a starting point, $x_k$ is the value of x for the kth iteration, T() is equation 1, and $x_{k+1}$ is determined based on equation 2:

**Algorithm 2**

**Input**: Anchor points, $(\mathbf{a}_1, \ldots \mathbf{a}_n), \mathbf{x}_0 \in \mathbb{R}^d$ and $\epsilon > 0$

1: $k \leftarrow 0$
2: **while** True **do**
3: $\quad \mathbf{x}_{k+1} \leftarrow T(\mathbf{x}_k)$
4: $\quad$ **if** $\|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2 < \epsilon$ **then**
5: $\quad\quad$ **return** $\mathbf{x}_{k+1}$
6: $\quad k \leftarrow k + 1$

**Equation 1**

$$T(\mathbf{x}) = \begin{cases} \tilde{T}(\mathbf{x}) = \dfrac{\sum_{i=1}^{n} \|\mathbf{a}_i - \mathbf{x}\|^{-1} \mathbf{a}_i}{\sum_{i=1}^{n} \|\mathbf{a}_i - \mathbf{x}\|^{-1}} & \text{if } \mathbf{x} \notin \{\mathbf{a}_1, \ldots \mathbf{a}_n\} \\ \mathbf{a}_i & \text{if } \mathbf{x} = \mathbf{a}_i, i = 1, \ldots, n \end{cases}$$

**Equation 2**

$$\mathbf{x}_{k+1} = T(\mathbf{x}_k) \ , \ k \in \mathbb{N}$$

In some non-limiting embodiments or aspects, applying geometric median normalization may be in addition to or in lieu of applying mean centering.

**[0131]** As shown in FIG. 3, at step 306, process 300 may include applying spectral normalization. For example, embedding normalization/alignment system 102a (e.g., a server, a part of transaction service provider system 102, a part of a third-party system, and/or the like) may apply spectral normalization to the at least one embedding set.

**[0132]** In some non-limiting embodiments or aspects, applying spectral normalization to the at least one embedding set may include decomposing the at least one embedding set to provide a left singular vector, a right singular vector, and a diagonal matrix. For example, decomposing the at least one embedding set may include performing singular value decomposition (SVD) on the at least one embedding set.

**[0133]** In some non-limiting embodiments or aspects, an average singular value of the at least one embedding set may be determined. For example, determining the average singular value may include determining a square root of an average squared singular value.

**[0134]** In some non-limiting embodiments or aspects, for each respective singular value of the diagonal matrix, whether the respective singular value is greater than a configurable multiple of the average singular value may be determined.

**[0135]** In some non-limiting embodiments or aspects, if a respective singular value is greater than the configurable multiple of the average singular value, a respective substitute singular value may be determined based on a quotient of the respective singular value divided by the configurable multiple of the average singular value. Additionally or alternatively, if a respective singular value is not greater than the configurable multiple of the average singular value, the respective substitute singular value may be determined to be a configurable value (e.g., 1, a predetermined integer, a predetermined value, and/or the like). In some non-limiting embodiments or aspects, a substitute diagonal matrix may include the respective substitute singular value for each respective singular value of the diagonal matrix. In some non-limiting embodiments or aspects, the at least one embedding set may be replaced with a product of the at least one embedding set, the right singular vector, and an inverse of the substitute diagonal matrix.

**[0136]** In some non-limiting embodiments or aspects, for the purpose of illustration, applying spectral normalization may include applying the following algorithm, where A is an embedding set, svd() is a singular value decomposition function, U is a left singular vector, V is a right singular vector, Σ is a diagonal matrix, T is the transpose operator, η is an average singular value, D is a (substitute) diagonal matrix, d is the dimension of the embedding vectors, $||A||_F$ is the Frobenius norm of embedding set A, and β is a parameter (e.g., hyperparameter, selectable parameter, and/or the like) used to determine the configurable multiple of the average singular value:

**Algorithm 3**

1: Compute $\text{svd}(A) = U\Sigma V^{\mathsf{T}}$; Let $D \in \mathbb{R}^d$ be a diagonal matrix.
2: Compute $\eta = \sqrt{\|A\|_F^2/d}$, where $d$ is the dimension of the word embedding
3: **for** $i = 1, \ldots, d$ **do**
4: **if** $(\Sigma_{ii} > \beta\eta)$ **then** $D_{ii} \leftarrow \Sigma_{ii}/(\beta\eta)$
5: **else** $D_{ii} = 1$
6: **return** $AVD^{-1}$

For example, if a respective singular value (e.g., $\Sigma_{ii}$) is greater than the configurable multiple of the average singular value (e.g., βη), a respective substitute singular value (e.g., $D_{ii}$) may be determined based on a quotient of the respective singular value

divided by the configurable multiple of the average singular value (e.g., $\Sigma_{ii}/(\beta\eta)$). Additionally or alternatively, if a respective singular value (e.g., $\Sigma_{ii}$) is not greater than the configurable multiple of the average singular value (e.g., $\beta\eta$), the respective substitute singular value (e.g., $D_{ii}$) may be determined to be a configurable value (e.g., 1, a predetermined integer, a predetermined value, and/or the like).

**[0137]** As shown in FIG. 3, at step 308, process 300 may include applying length normalization. For example, embedding normalization/alignment system 102a (e.g., a server, a part of transaction service provider system 102, a part of a third-party system, and/or the like) may apply length normalization to the at least one embedding set.

**[0138]** In some non-limiting embodiments or aspects, applying length normalization may include adjusting each embedding vector of the set of embedding vectors to have a 2-norm (e.g., Euclidean norm) of 1.

**[0139]** In some non-limiting embodiments or aspects, as shown in FIG. 3, steps 304, 306, and 308 may be repeated for a configurable number of iterations. For example, embedding normalization/alignment system 102a (e.g., a server, a part of transaction service provider system 102, a part of a third-party system, and/or the like) may iteratively repeat applying mean centering, applying spectral normalization, and applying length normalization to the at least one embedding set for a configurable number of iterations.

**[0140]** In some non-limiting embodiments or aspects, for the purpose of illustration, iteratively repeating may include applying the following algorithm, where m is a configurable number of iterations, A is an embedding set, Center is mean centering, SpecNorm is spectral normalization, and Unit Length Normalization is length normalization, as described herein:

**Algorithm 4**

1: **for** $m$ steps **do**
2: $A \leftarrow \text{Center } A$
3: $A \leftarrow \text{SpecNorm } (A)$
4: $A \leftarrow \text{Unit length normalization of } A$
5: **return** $A$

**[0141]** In some non-limiting embodiments or aspects, the parameters may be tuned. For example, the parameter β used to determine the configurable multiple of the average singular value and/or the parameter m for the configurable number of