

【公報種別】特許法第17条の2の規定による補正の掲載  
【部門区分】第6部門第3区分  
【発行日】令和7年1月9日(2025.1.9)

【国際公開番号】WO2024/013991  
【出願番号】特願2024-533482(P2024-533482)  
【国際特許分類】  
G06F16/35(2025.01)  
【FI】  
G06F16/35

10

【手続補正書】  
【提出日】令和6年10月22日(2024.10.22)  
【手続補正1】  
【補正対象書類名】特許請求の範囲  
【補正対象項目名】全文  
【補正方法】変更  
【補正の内容】  
【特許請求の範囲】  
【請求項1】

20

複数の単語を有する複数の文を取得し、  
複数の単語を有する文の文ベクトルと、前記複数の単語の単語ベクトルとを基にして、  
前記複数の単語から特徴単語の組を特定する処理を、前記複数の文に対して実行し、  
同一の特徴単語の組を有する文が、同一のクラスタに含まれるように、前記複数の文を  
分類する  
処理をコンピュータに実行させることを特徴とする情報処理プログラム。

【請求項2】  
前記特徴単語の組を特定する処理は、前記文の文ベクトルと、前記複数の単語の単語ベクトルとのコサイン類似度を算出し、前記文ベクトルとのコサイン類似度が閾値以上となる単語ベクトルの単語を、特徴単語として特定することを特徴とする請求項1に記載の情報処理プログラム。

30

【請求項3】  
前記複数の文は記憶装置に登録され、  
前記分類する処理の分類結果を基にして、前記文の属するクラスタを識別する識別情報と、前記記憶装置における前記文の位置情報とを対応付けた転置インデックス情報を生成する処理を更にコンピュータに実行させることを特徴とする請求項1に記載の情報処理プログラム。

【請求項4】  
複数の単語を有する検索文を受け付けた場合に、前記検索文の文ベクトルと、前記検索文の単語の単語ベクトルとを基にして、前記検索文の複数の単語から特徴単語の組を特定し、特定した特徴単語の組に対するクラスタの識別情報と、前記転置インデックスとを基にして、前記検索文に対応する文を前記記憶装置から検索する処理を更にコンピュータに実行させることを特徴とする請求項3に記載の情報処理プログラム。

40

【請求項5】  
複数の検索文を受け付けた場合に、前記複数の検索文の文ベクトルを基にして、複数の特徴文を特定し、前記複数の特徴文から特徴単語の組をそれぞれ特定し、特徴文に対応する特徴単語の組を基にして、各特徴文のクラスタの識別情報を特定し、各特徴文のクラスタの識別情報と、前記転置インデックスとを基にして、各特徴文に対応する文であって、共通する文を前記記憶装置から検索する処理を更にコンピュータに実行させることを特徴とする請求項4に記載の情報処理プログラム。

50

## 【請求項 6】

前記複数の検索文から特定される複数の特徴文の数を増加または減少させる処理を更にコンピュータに実行させることを特徴とする請求項 5 に記載の情報処理プログラム。

## 【請求項 7】

複数の単語を有する複数の文を取得し、

複数の単語を有する文の文ベクトルと、前記複数の単語の単語ベクトルとを基にして、前記複数の単語から特徴単語の組を特定する処理を、前記複数の文に対して実行し、

同一の特徴単語の組を有する文が、同一のクラスタに含まれるように、前記複数の文を分類する

処理をコンピュータが実行することを特徴とする情報処理方法。

10

## 【請求項 8】

複数の単語を有する複数の文を取得し、

複数の単語を有する文の文ベクトルと、前記複数の単語の単語ベクトルとを基にして、前記複数の単語から特徴単語の組を特定する処理を、前記複数の文に対して実行し、

同一の特徴単語の組を有する文が、同一のクラスタに含まれるように、前記複数の文を分類する

処理を実行する制御部を有することを特徴とする情報処理装置。

20

30

40

50