

Web内容マイニングによる個別知識の  
共通単語を介した相関ルールからの発想支援

1615052 山元悠貴

情報基盤工学講座 指導教員 奥原浩之

## 要約

本研究では、グーグル検索を用いた2つのキーワードからの発想支援について考える。既存の発想支援システムでは1つのキーワードから発想を広げていくものが多かった。そこで、複数のキーワードから関連する単語を用いることでより効率の良い発想支援につながると考えた。

キーワード：テキストマイニング、発想支援、  
共起ネットワーク

## 1 はじめに

現在、AIの発達に伴って発想支援についての研究も進んでいる。これから時代はよりアイデア発想が重要な要素になってくると考える。

そこで現状の発想支援について、複数のキーワードからの発想支援に着目した。

本研究では、複数のキーワードについて共起ネットワークからその関連語をランキング形式で表示する事を考える。

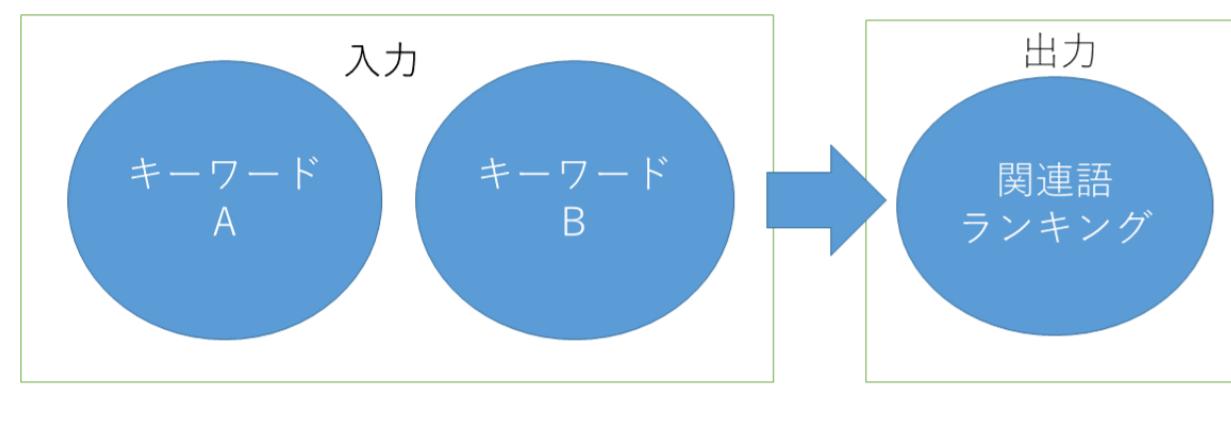


図1 イメージ図

## 2 発想支援システムとは

## 2.1 サイバー空間からのテキストデータの収集

発想支援において重要なことはキーワードからより関連度の高い単語をより多く表示させることである。

そこで、より良いデータを多く収集するためにサイバー空間からテキストデータを収集することとする。

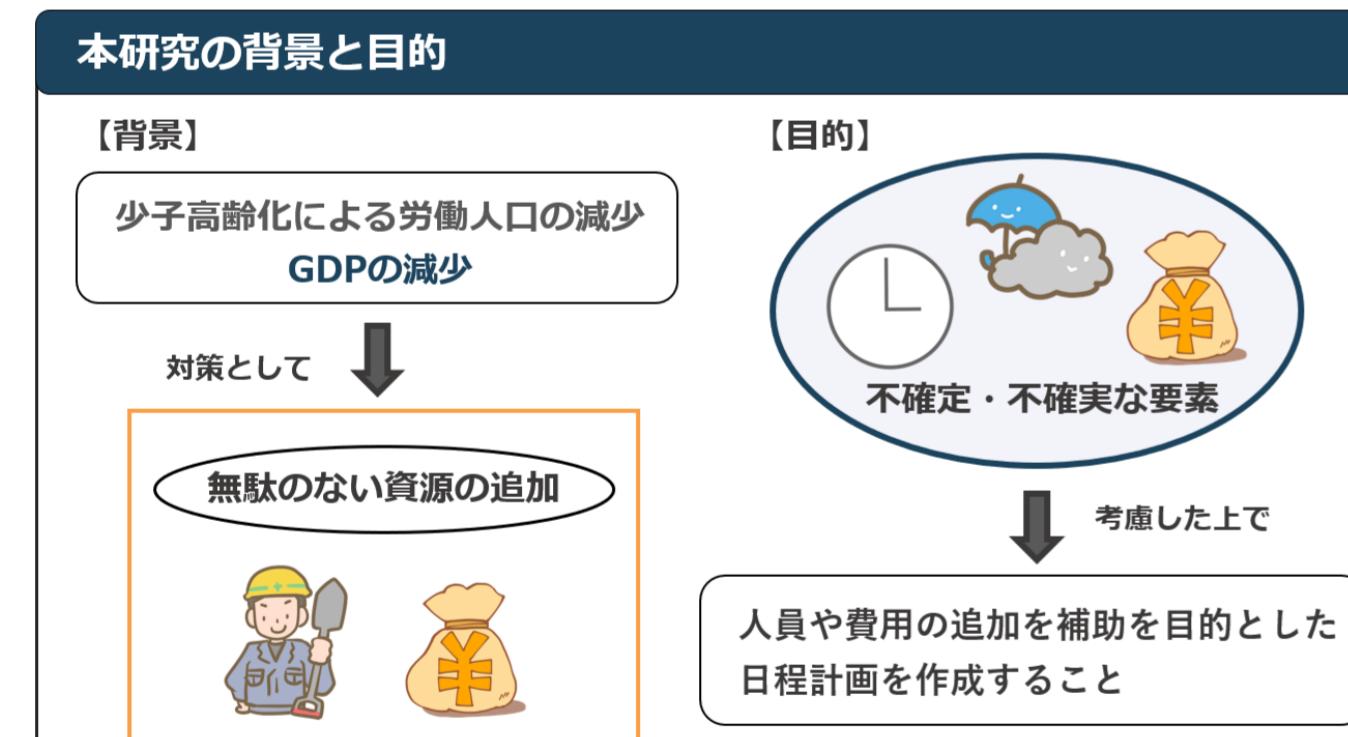


図2 クリティカルパスの最小化

## 2.2 テキストマイニング

テキストマイニングとは、大量かつ多量なデータを様々な観点から分析し、役に立つ情報を取り出そうとする技術である。

今回、インターネット上のテキストを用いることで大量のデータを活用することができる。



図3 テキストマイニングの流れ

## 2.3 既存の発想支援と課題

既存の発想支援システムにはAIブレーストスパーク[2]のひらめきマップがある。このツールはキーワードを入力するとその単語に共起する関連語を表示して発想支援に役立てるというものである。課題としては複数のキーワードに関して、それらのキーワードの関連語を表示することができないことがある。

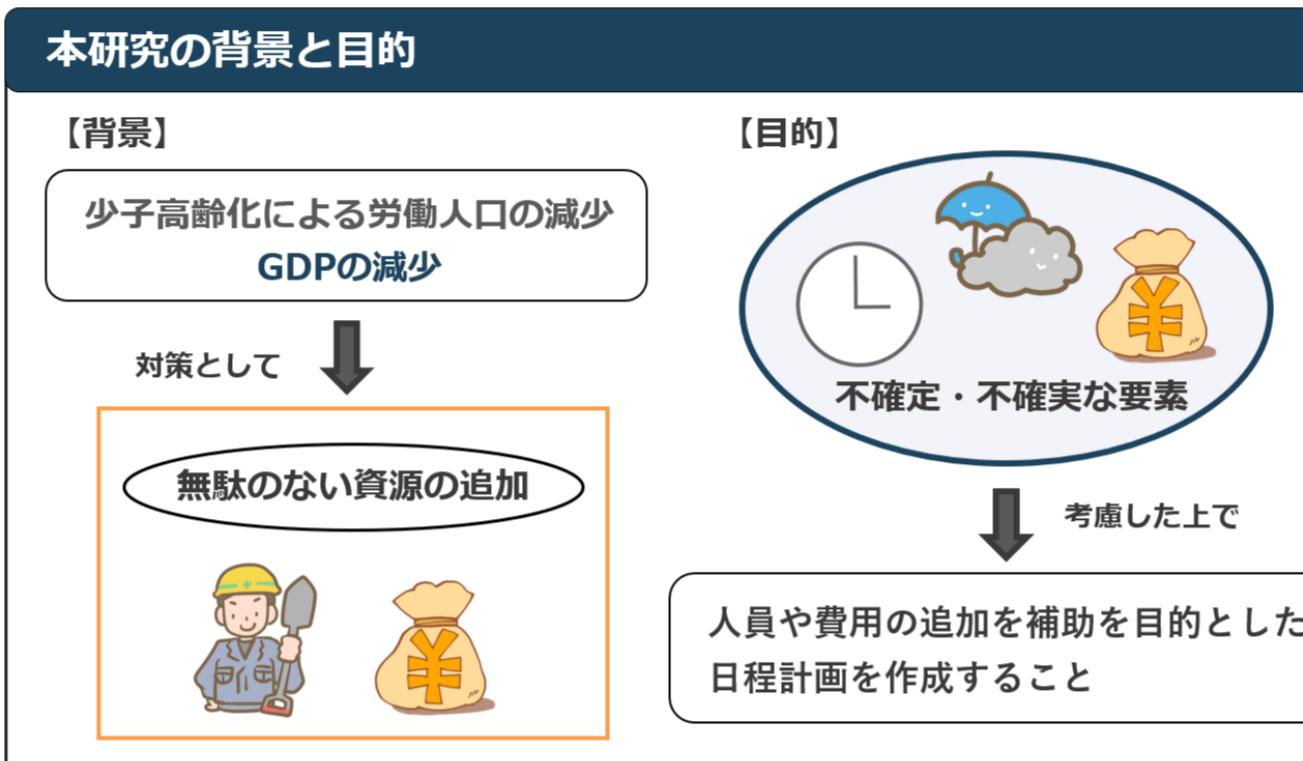


図4 作業履歴を用いたパラメータ設定

## 3 データドリブンによる最適化

## 3.1 単語の分散表現ベクトル

単語をベクトル表現する技術がある。(word2vec)

複数のキーワードの関連語がなかった場合にも発想支援できるように単語をベクトル表現することで近いベクトルの単語同士を関連語として見れるのではないかと考える。

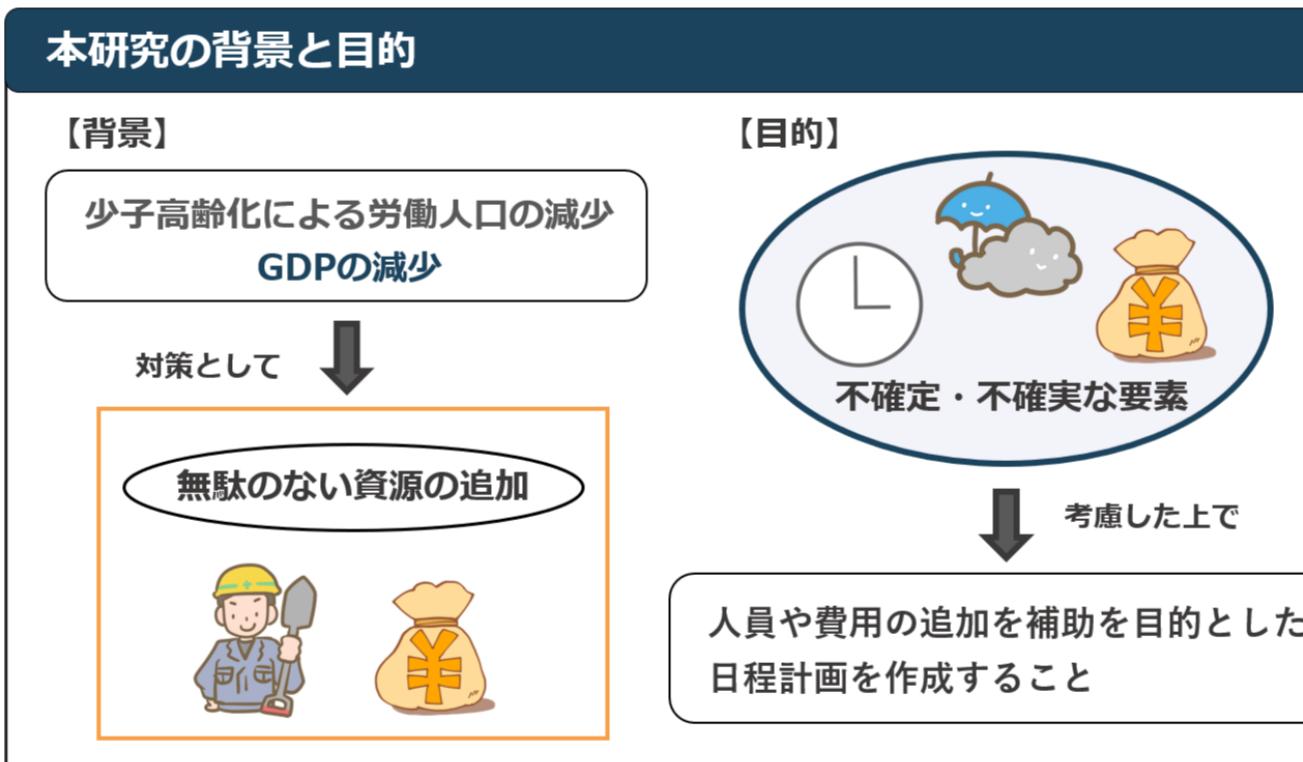


図5 ファジィ・ランダム

## 3.2 複数のキーワード間の共起ネットワーク

jaccard係数を使って共起ネットワークを作成する  
 $J(A,B) = \frac{|A \cap B|}{|A \cup B|}$



図6 共起ネットワーク作成の流れ

しかし、

図7 目的関数のメンバシップ関数とファジィ目標の導入  
また、

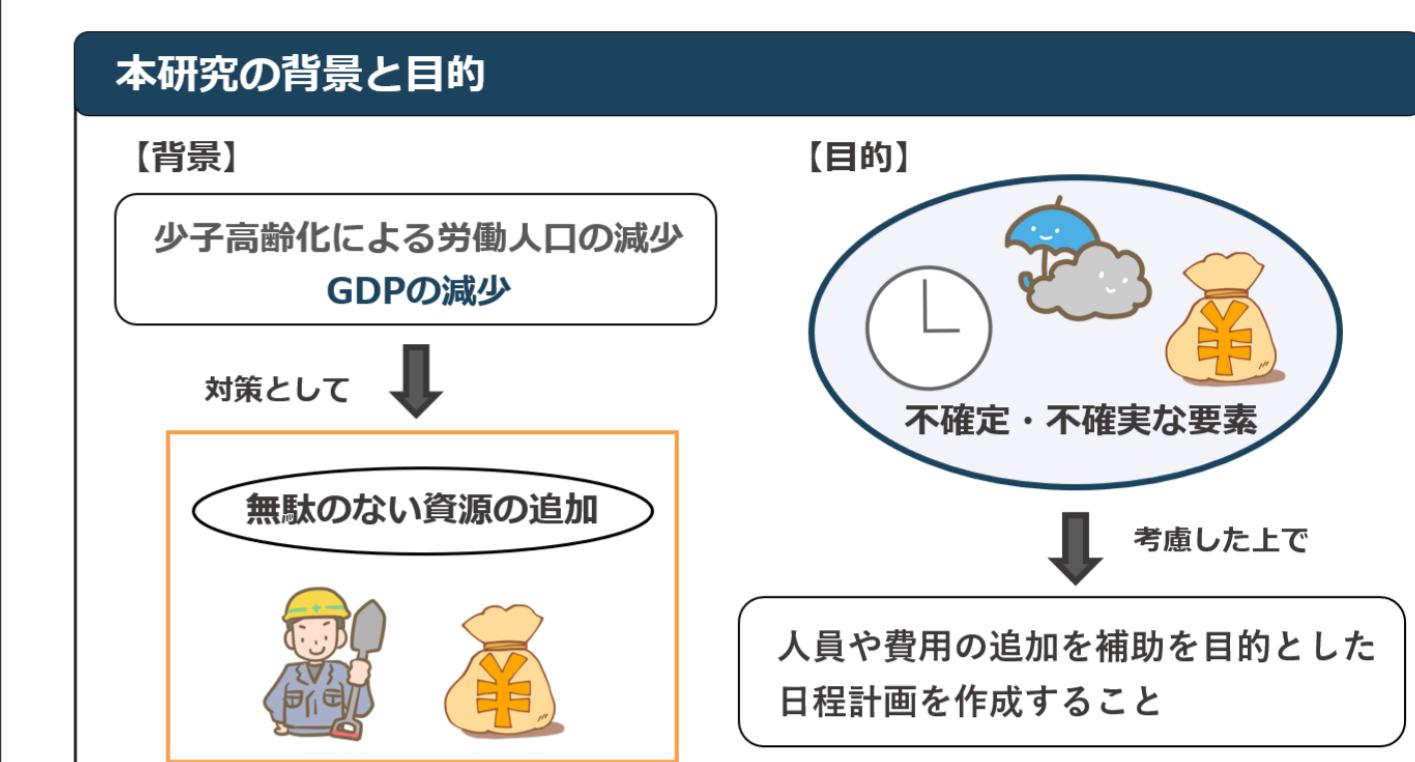


図8 確率最大化モデルに基づいた変換

## 3.3 nodeの強さを考慮したランキング

今回の目的のnodeの強さを考慮したランキングを作成するため

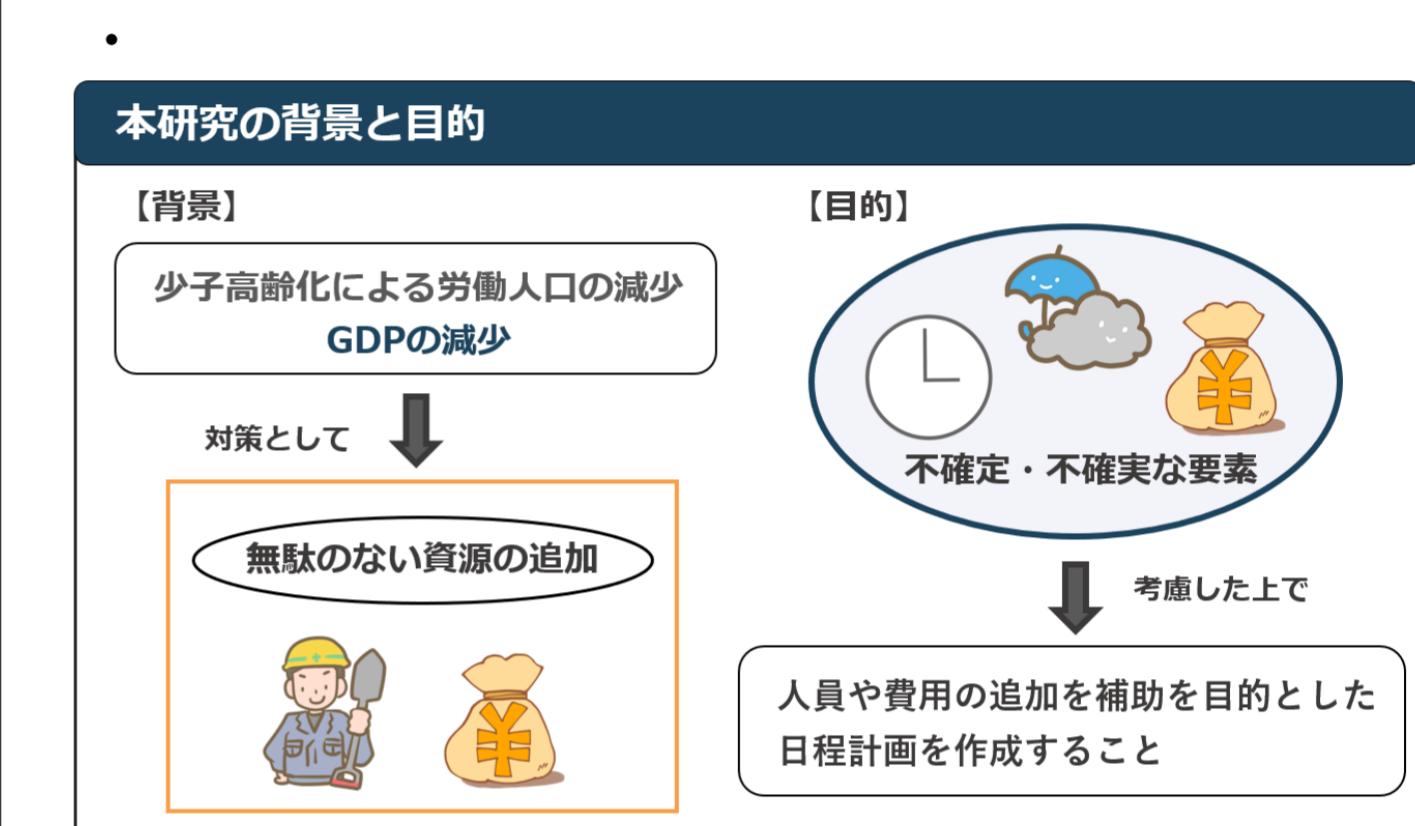


図3 時間費用関数と多目的最適化問題

## 4 数値実験ならびに考察

今回は、

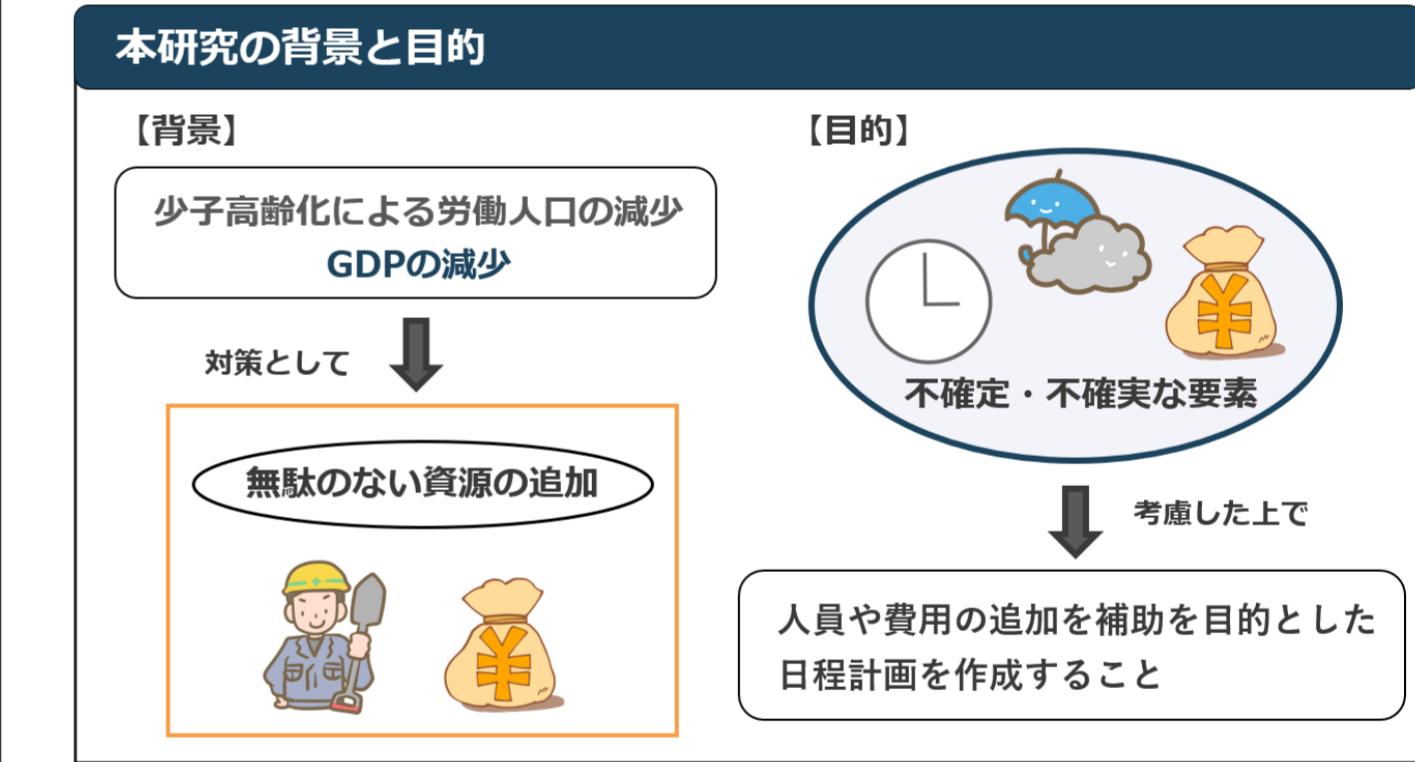


図9 GAを用いた数値実験例

## 5 おわりに

本研究では、

## 図10 今後の課題

ペイジアンネットワークなど他の方法も検討する。  
システムの評価方法についての考察をする。

## 参考文献

- [1] “国内人口推移が2030年の「働く」にどのような影響を及ぼすか”  
<https://www.reruit-ms.co.jp/research/2030/report/trend1.html>
- [2] “AIブレーストスパーク”  
<https://www.ai-b-spark.com/>