金融・経済分析のためのテキストマイニング

# 金融テキストマイニングの最新動向

　機械学習，特に深層学習で極性分析をすると解釈が難しく，顧客への説明が．．．

⇒ 解釈可能なニューラルネットワークを紹介

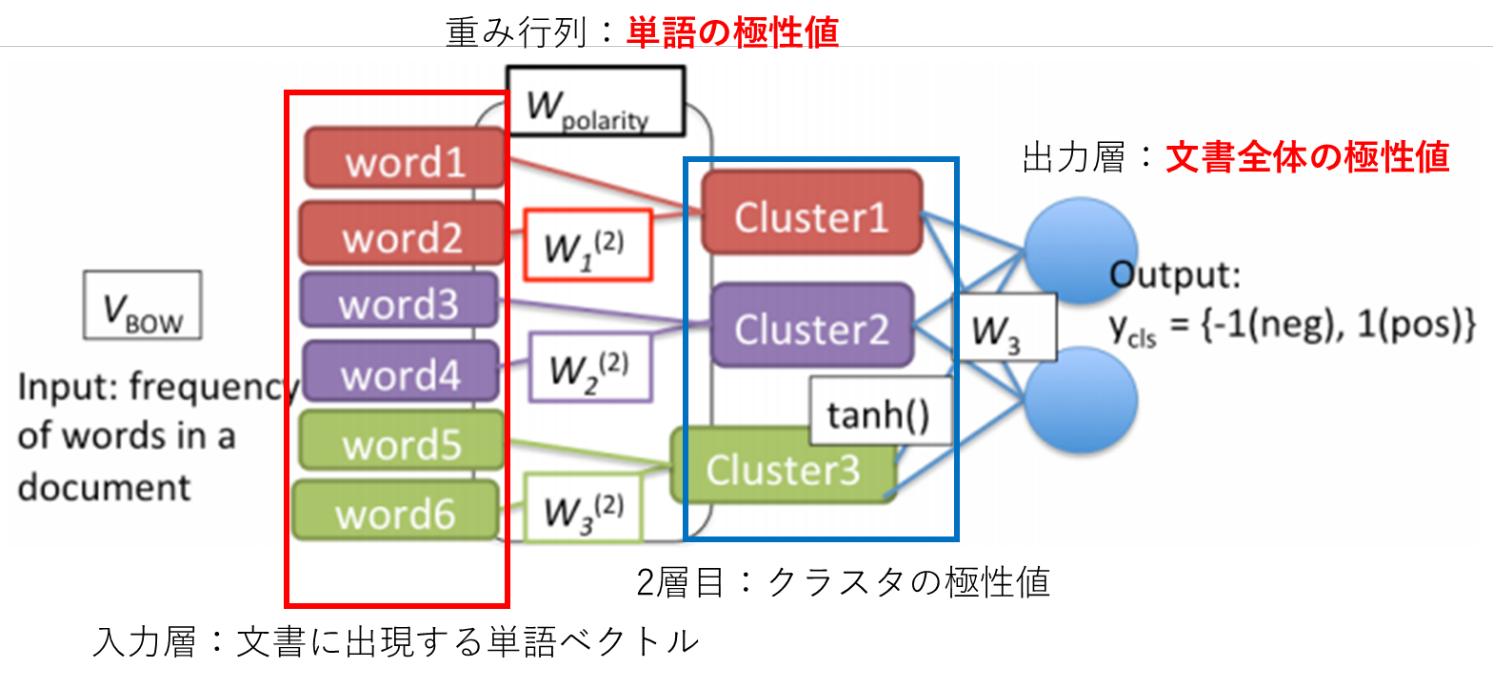
## 解釈可能なニューラルネットワーク

* **Importance Infiltration propagation (Ⅱ) algorithm（Ⅱ algorithm）**

目的：金融単語の極性情報を獲得

* 分析の流れ（目的箇所が**赤字**，工夫点が青字）

1. 金融に関連する大量の文書よりWord2vecによって，単語をベクトル化
2. Spherical k-means法によってクラスタリング
3. ポジネガタグを持つ金融文書によってニューラルネットワークを構築  
   入力層：文書に出現する単語ベクトル（おそらく頻度）．ただし，②のクラスタリング結果より，該当  
   するクラスタのみにエッジがつながるようにする．  
   2層目：クラスタの極性値  
   **出力層：文書全体の極性値**  
   **入力層と2層目の重み行列単語の：極性値**（極性辞書を初期値）  
   2層目と出力層の重み行列：言及なし，単なる重み行列
4. 誤差逆伝搬法によってモデルの学習  
   ⇒ 極性辞書外の単語を含めた極性値を獲得！！



その他では，Ⅱ algorithmによって，ポジ単語には正，ネガ単語には負の極性値が与えられる条件や人口データ・実データによる検証がなされていることが紹介されているが，書籍参照．