

課題ID: jh221002

課題名:

医療・介護領域の人材マッチングに最適化された大規模グラフニューラルネットワーク

課題代表者名: 鈴木豊太郎^{1,2}

課題参加者:

脇聡志¹、小林秀³、金刺宏樹²、華井雅俊²、小貝洸希³、長谷川正彦³

¹ 東京大学・大学院情報理工学系研究科

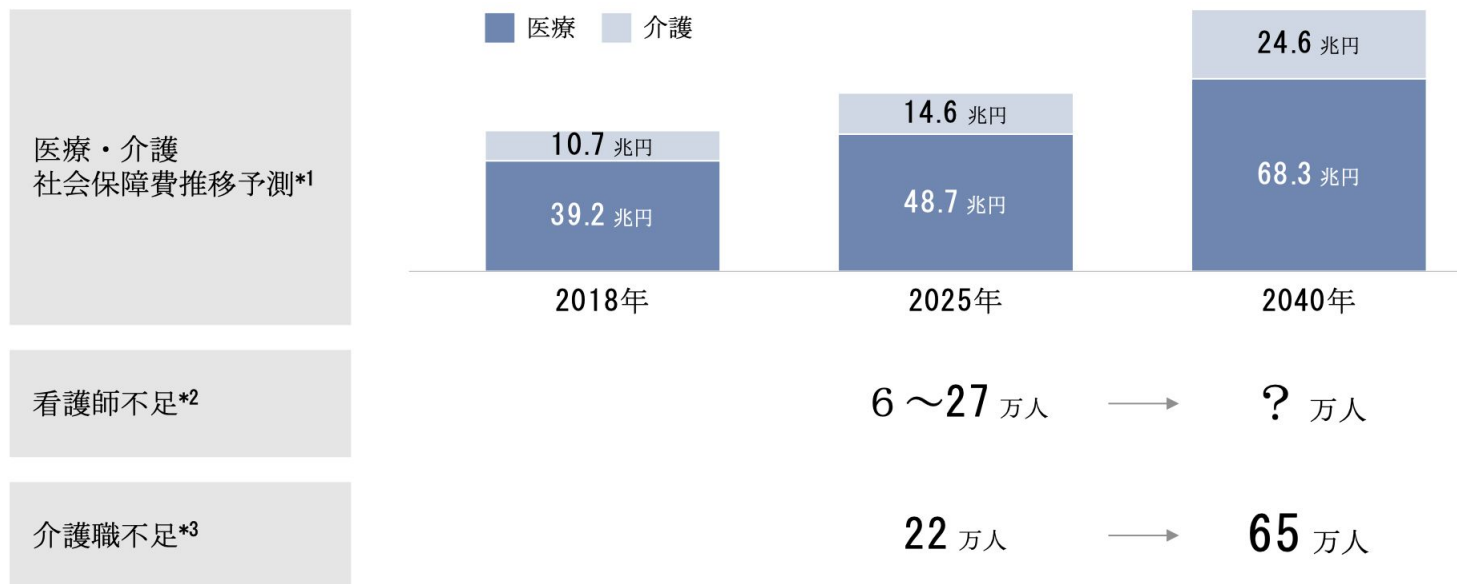
² 東京大学・情報基盤センター

³ 株式会社エス・エム・エス



事業課題

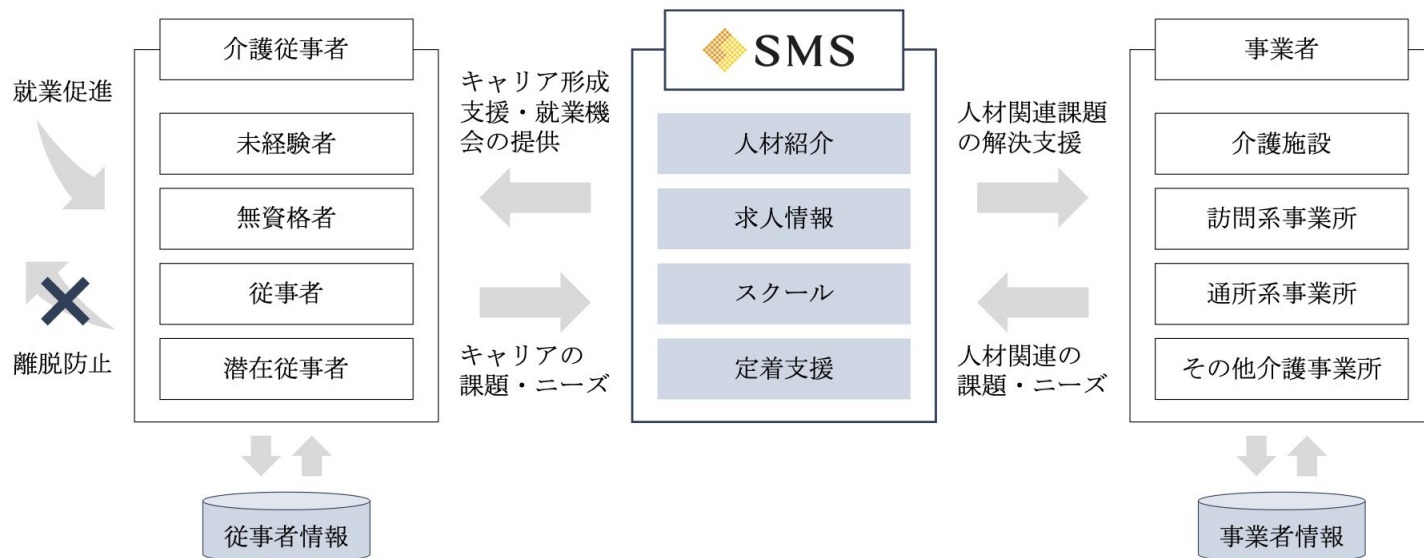
医療・介護の需要が増大する中、生産年齢人口の減少により、医療・介護従事者の不足が深刻化し、質の高い医療・介護サービスの提供が困難に



1. 内閣官房・内閣府・財務省・厚生労働省「2040年を見据えた社会保障の将来の見通し」
2. 厚生労働省「医療従事者の需給に関する検討会 看護職員需給分科会」
3. 厚生労働省「第8期介護保険事業計画に基づく介護職員の必要数について」

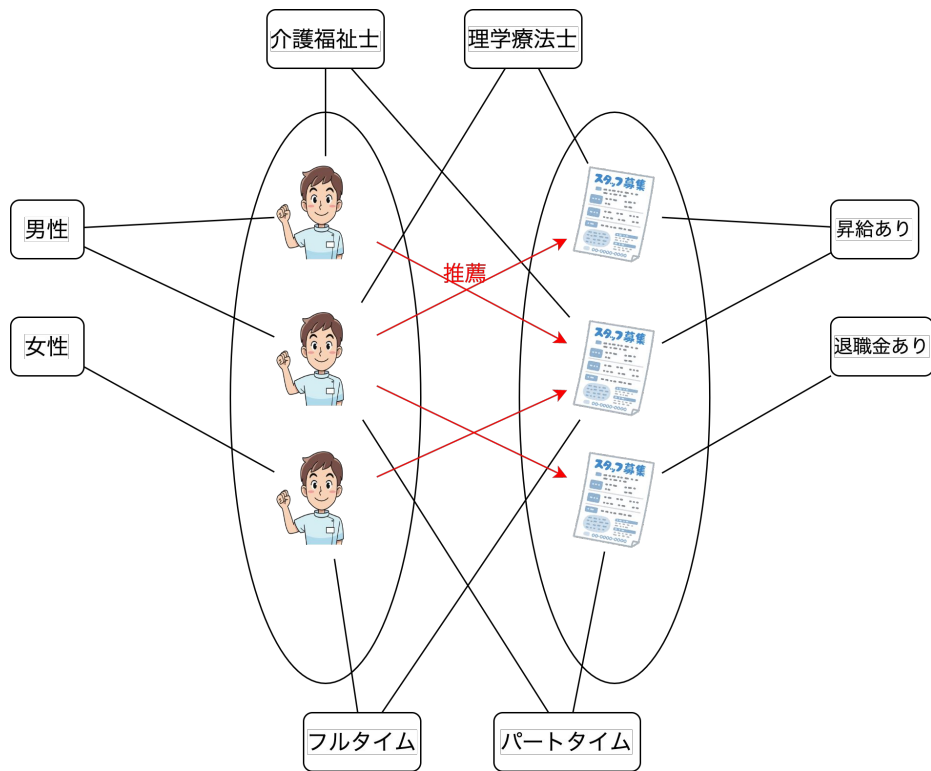
介護キャリア事業の戦略

採用や労働環境の改善などの事業者の人材関連課題の解決支援を通じてより良い職場環境を実現するとともに、従事者との最適なマッチングにより業界からの離脱も防止



医療・介護領域の人材マッチングに最適化された大規模グラフニューラルネットワーク

医療・介護領域での求職者に対する求人の推薦問題を解く。



医療・介護求人データ

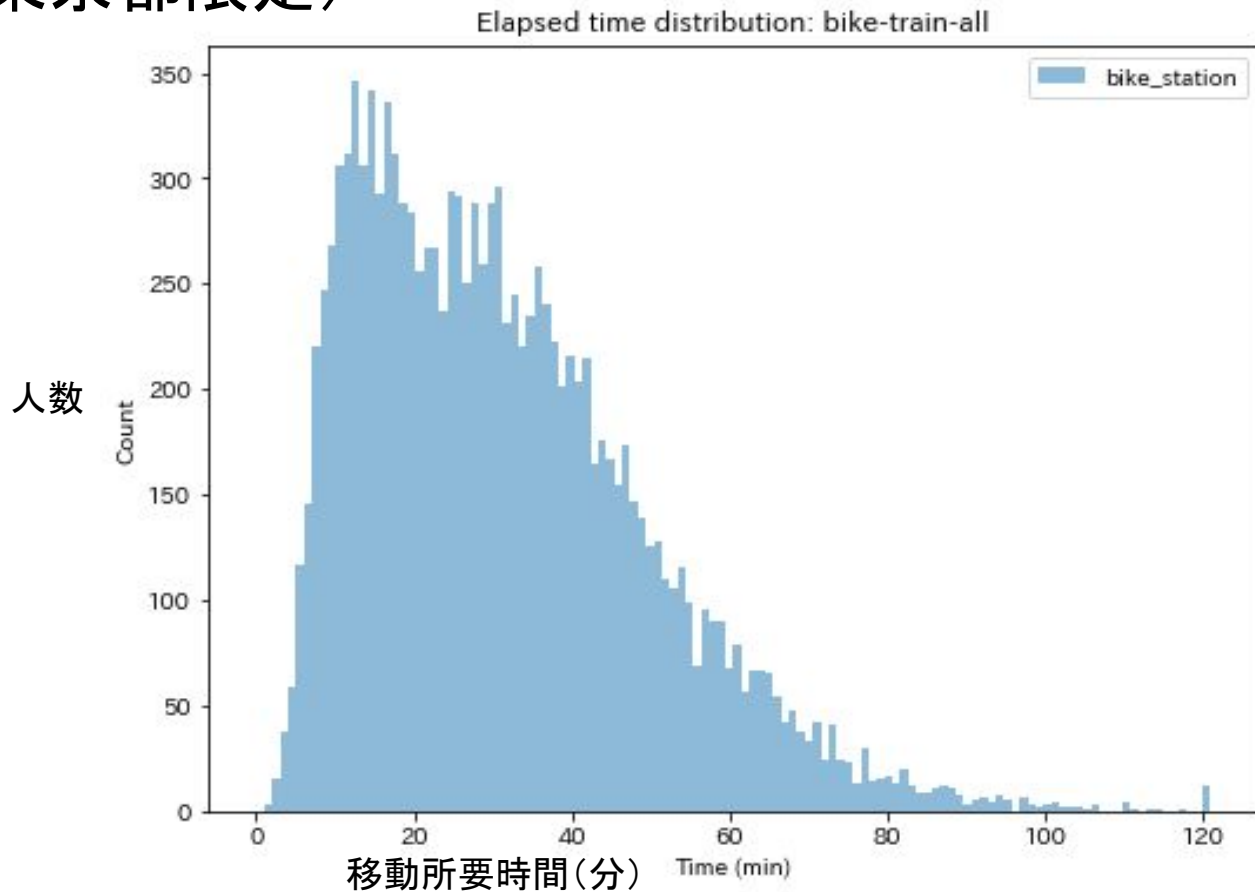
株式会社エス・エム・エスが運用する介護・福祉職向け求人サイト「カイゴジョブ」での情報を匿名化し提供されたデータを用いる。

- ユーザーのデータ
- 求人のデータ
- 応募データ
- 求人の閲覧データ
- 求人のキープ(一時保存)データ



The screenshot displays the KaiGoJob website interface. At the top, there are navigation links for 'ログイン' (Login), 'キープ' (Keep), and 'メニュー' (Menu). Below the header, there are search filters for 'エリア' (Area), 'こだわり条件' (Preferred conditions), '職種' (Job type), and 'もっと詳しい条件' (More detailed conditions). A job listing is shown with details such as '住所' (Address), '募集職種' (Recruitment job type), '雇用形態' (Employment type), and '給与' (Salary). A '★ どりあえずキープ' (★ Keep) button is visible. On the right side, there is a section for '求人検索を依頼したい方はこちら' (Click here if you want to request a job search), featuring a phone number '0120-301-916' and a '電話で相談' (Consult by phone) button. Below this, there is a '勤続支援金' (Loyalty support fund) section with a '3万円' (30,000 yen) benefit. At the bottom, there are buttons for '会員登録' (Register) and 'ログイン' (Login), and a link to '転職サポートコラム' (Job transition support column).

求職者・求人間の公共交通機関を考慮した移動所要時間分布 (東京都限定)



平均	31.76分
標準偏差	24.91分
最小値	1.36分
第一四分位数	16.73分
中央値	28.48分
第三四分位数	42.18分

データセットの特性

他の推薦データセットと比べてユーザーとアイテム(求人)間の相互作用が極端に少ない。

(正解の辺数) / (ユーザー数 × アイテム数) =

- SMS (ユーザー数: 1250581、求人数: 96465): 6.334e-7(apply), 5.514e-06(apply+pv)
- 一件も応募していない/されていないユーザー/求人除去 & 東京都のみの SMSデータ (ユーザー数: 6358、求人数: 5365): 3.866e-4(apply), 1.067e-3(apply+pv)
- [Amazon](#) (ユーザー数: 20.98e6、アイテム数: 9.35e6): 4.223e-1
- [Delicious Bookmarks](#) (ユーザー数: 1867、アイテム数: 53388): 8.109e-4
- [Last.FM](#) (ユーザー数: 1892、アイテム数: 17632): 2.783e-3
- [MovieLens + IMDb/Rotten Tomatoes](#) (ユーザー数: 2113、アイテム数: 10197): 3.971e-2

グラフモデリング

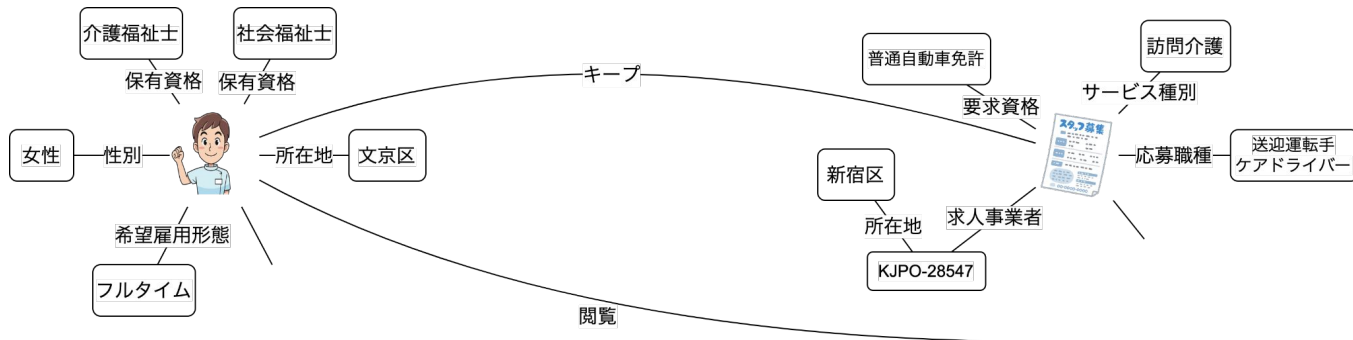
ユーザーや求人属性情報をすべて頂点として表す「知識グラフ」として扱う。

ユーザーの知識グラフ:

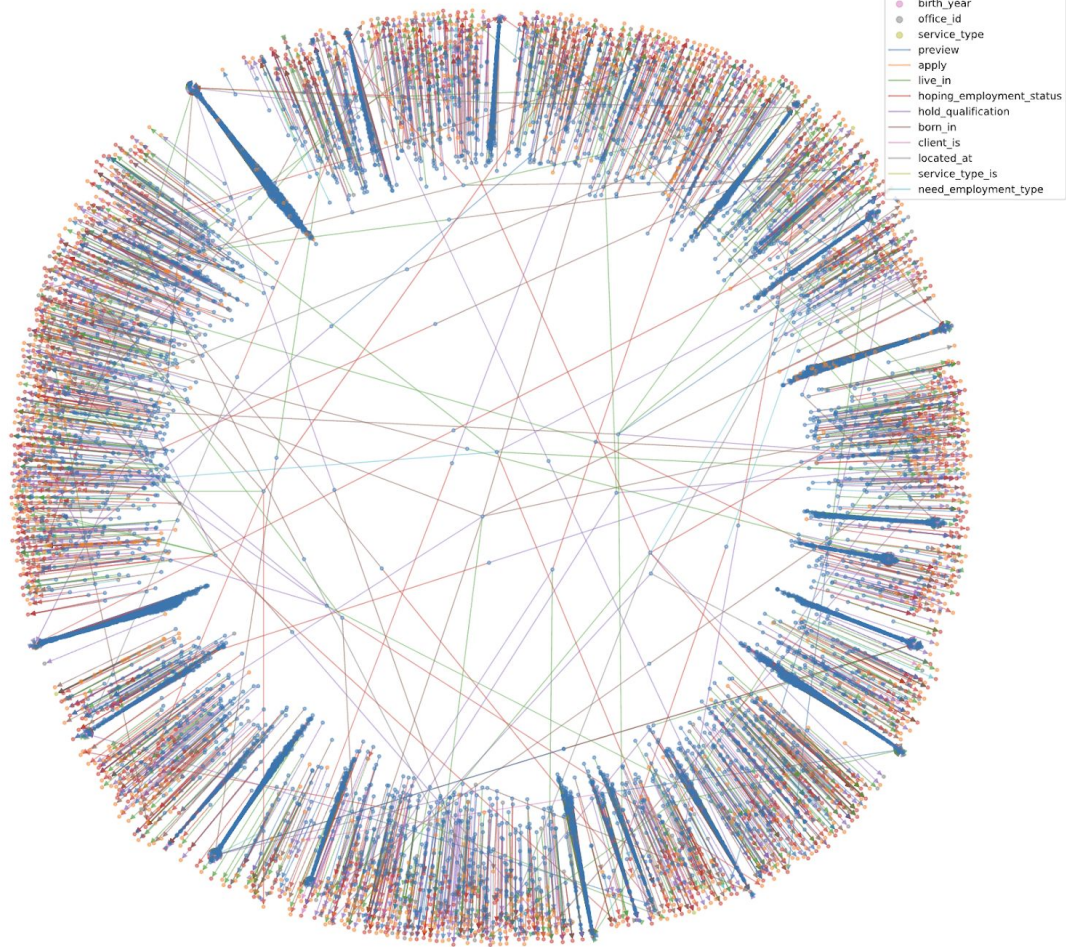
- (ユーザーID, “所在地”, 市区町村ID)
- (ユーザーID, “希望雇用形態”, 雇用形態ID)
- (ユーザーID, “保有資格”, 資格ID)
- ...

求人の知識グラフ:

- (求人ID, “求人事業者”, 事業者ID)
- (事業者ID, “所在地”, 市区町村ID)
- (求人ID, “要求資格”, 資格ID)
- ...



可視化



グラフニューラルネットワーク

既存手法のRelational Graph Convolutional Networks (R-GCNs) [Michael Schlichtkrull et al., 2018]を用いる。

R-GCNsは、グラフに対して畳み込み演算をおこなうGraph Convolutional Networks (GCNs)と呼ばれる手法に辺の関係性を追加したモデルである。

具体的にはステップ目での更新式は以下:

$$h_i^{(l+1)} = \sigma \left(\sum_{r \in \mathcal{R}} \sum_{j \in \mathcal{N}_i^r} \frac{1}{c_{i,r}} W_r^{(l)} h_j^{(l)} + W_0^{(l)} h_i^{(l)} \right)$$

ただし $h_i^l \in \mathbb{R}^d$: 頂点 v_i の l 層目での特徴量,

$\sigma: \mathbb{R} \rightarrow \mathbb{R}$: 活性化関数,

\mathcal{R} : 関係 (辺のラベル) の集合,

$\mathcal{N}_i^r \in \{j \mid (v_i, r, v_j) \in \mathcal{E}\}$: 頂点 v_i と関係 r でつながった頂点集合,

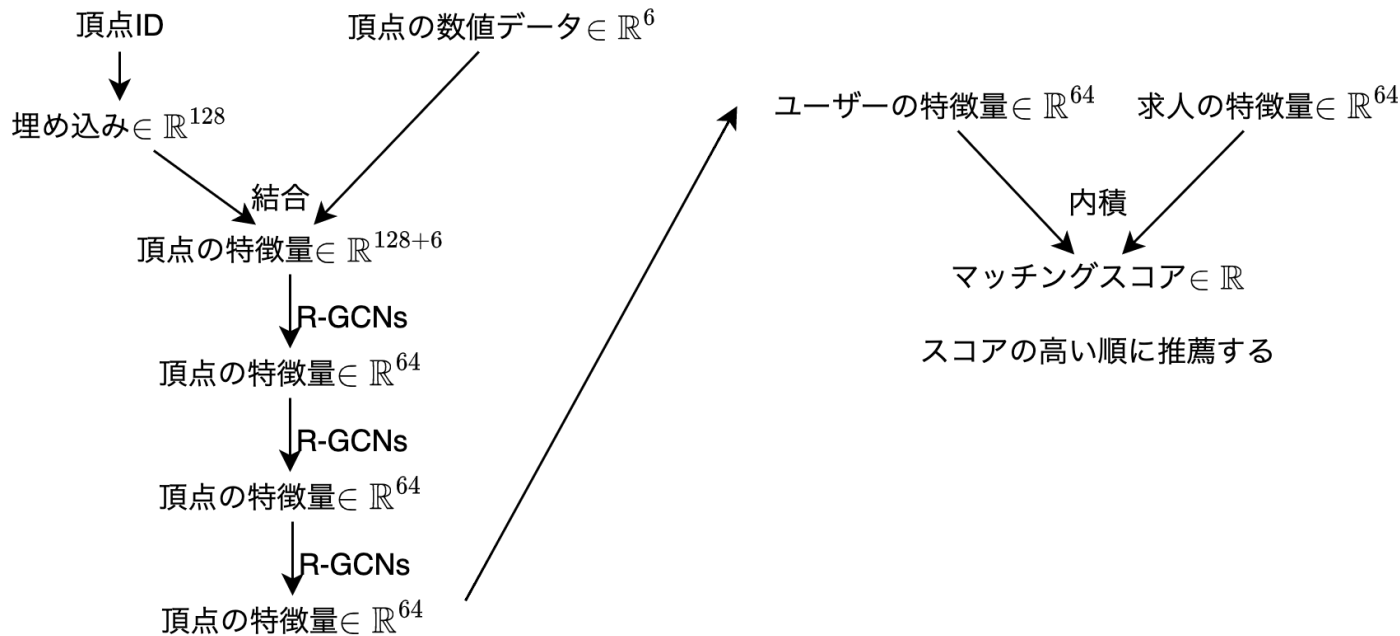
$c_{i,r} \in \mathbb{R}$: 正規化定数,

$W_r^{(l)}, W_0^{(l)} \in \mathbb{R}^{d \times d}$: l 層目の学習する行列.

計算環境と推薦システム

計算環境: mdx (1GPU, 56GB DRAM, 18 vCPUs, 1.6GB disk usage)

推薦システムの全体像:



今後の展望

- 地域や資格等によってユーザーや求人フィルタリングし、より相互作用を多くしたグラフで解く。
- アンバランスなノード分類問題のモデルを参考に、正解の辺数が少ないデータにうまく対応したモデルを提案する。