

A Statistical Prediction Model of Students' Finishing Date on Job Hunting Based on Probabilistic Latent Class Model

YAMAGAMI Kan

1 研究背景・目的

近年、学生の就職活動を Web 上で支援する就職ポータルサイトが広く利用されている。学生はこのサービスを利用して企業情報の検索や企業の採用活動への申し込み（以下、エントリー）を行うことができるため、就職活動において欠かせないツールとなっている。このサービスの活用により学生の就職活動の利便性が高まる一方、その長期化は社会的な問題となっている。このため、就職活動が長期化する可能性がある学生を早期に発見し、早い段階から支援を行うことは重要であり、就職活動終了日の予測はこの解決の一助となる可能性がある。

この問題に対して、早川ら [1] により就職ポータルサイト上に蓄積されたデータを用いた就職活動終了日予測モデルが提案されている。早川らは実データの分析を通じて、就職活動終了日の分布が学生の属性情報に強く依存することを指摘し、属性情報による層別と混合ワイブル分布に基づく就職活動終了日の予測モデルを構築している。しかし、就職ポータルサイト上には属性情報以外にも学生が就職ポータルサイト上でとった様々な行動の履歴データが蓄積されている。これらの情報もまた就職活動終了日を決める一つの要素であることが考えられるため、その活用は就職活動終了日の予測精度向上に寄与する可能性がある。それに加え、就職活動終了日と行動情報の統計的な関係を明らかにすることができれば、サイト運営者は学生への適切な支援施策を立案することができる。

そこで本研究では、就職活動を行う学生全体は様々な特徴の異なるグループが混在した集合体であるという仮説に基づき、学生の属性情報、行動情報、就職活動終了日の三者の関係を表現する潜在クラスモデルを提案し、パラメータ推定アルゴリズムを導出する。提案する潜在クラスモデルにより、学生の属性情報と行動情報から就職活動終了日の予測を行うことが可能となる。提案手法の有効性を示すために、大手就職ポータルサイトの実データを用いたシミュレーション実験を行う。

2 準備

2.1 就職活動終了日に関する統計モデル [1]

早川らは実データを分析し、学生の就職活動終了時期の傾向は属性情報に強く依存しているということを示した。図 1 に、ある就職ポータルサイトから抽出した 2014 年 3 月卒の学生の就職活動終了割合の推移を示す。3 つの曲線はそれぞれ A 大学の学生の推移、B 大学の学生の推移、全学生の推移を表す。図 1 より、A 大学の学生は全学生と比べて早く就職活動を終えており、一方で B 大学の学生には就職活動が長期化する傾向がみられる。したがって、就職活動終了日は学生が所属する学校名などの属性情報に依存する可能性があるといえる。

一方、図 1 の曲線は信頼性工学で用いられる累積故障割合で描かれるものと類似しており、早川らは学校ごとの就職活動終了者数の推移データがワイブル分布で近似できるか否かについて検討を行っている。ワイブル分布

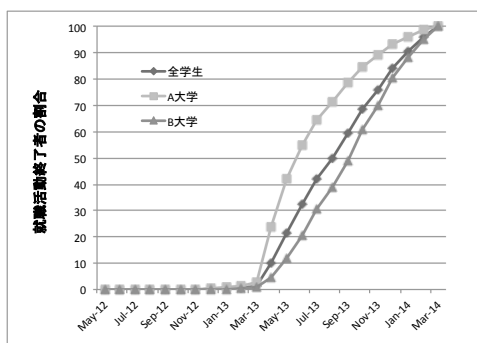


図 1: 就職活動終了者の割合の推移

の確率密度関数は時間データ x に対して、形状パラメータ m 、尺度パラメータ η を用いて式 (1) により表される。

$$p(x|m, \eta) = \frac{m}{\eta} \left(\frac{x}{\eta}\right)^{m-1} \exp\left\{-\left(\frac{x}{\eta}\right)^m\right\} \quad (1)$$

図 1 の経験分布から最尤推定法によりパラメータ推定を行ったところ、単独のワイブル分布では当てはまりが悪いことが判明したため、複数のワイブル分布を混合した混合ワイブル分布の導入が必要であるとしている。以上より、早川らは学生の属性情報による層別と混合ワイブル分布に基づく就職活動終了日の予測モデルを構築した。

2.2 事前分析

属性情報に加え、行動情報も就職活動終了日に影響を与えている可能性があるため、それらの関係について 2014 年 3 月卒の学生を対象に分析を行った。以下では、就職活動終了日に影響を与えている可能性のある行動情報として、早期エントリー数とインターンシップエントリーの有無について就職活動終了日との関係を考える。

まず、早期から活発に就職活動を行っている学生は早い段階で内々定が得られると仮定し、早期エントリー数と就職活動終了日との関係について分析を行った。いま、2014 年 3 月卒の学生が 2013 年 3 月 1 日以前に行ったエントリーを早期エントリーと定義する。

表 1 に就職活動終了日の早さによって層別した学生の早期エントリー数の平均を示す。終了日が 2012/12/1 から 2013/5/31 の間の学生を「早期終了学生」、2013/6/1 から 2013/10/31 の間の学生を「平均的な学生」、2013/11/1 から 2014/3/21 の間の学生を「長期化学生」とラベル付けを行った。表 1 より、早期エントリー数と就職活動終了日は完全に無関係とはいえないことがわかる。

表 1: ラベル付けされた学生の早期エントリー数の平均

早期終了学生	平均的な学生	長期化学生	全学生
25.505	24.189	20.291	23.363

また、インターンシップエントリーの有無と就職活動終了日の関係についても分析を行った。図 2 にインターンシップのエントリーを 1 件以上行った学生と行わなかった学生の就職活動終了者の割合の推移を示す。

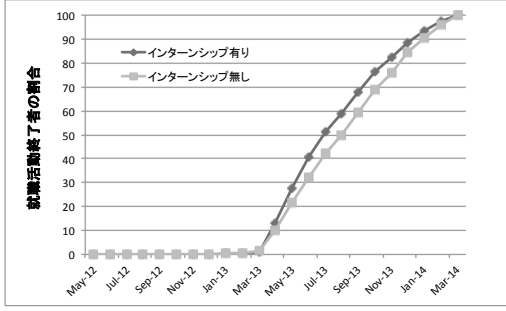


図 2: インターンシップエントリーの有無別就職活動終了者の割合の推移

図 2 より、インターンシップのエントリーを行った学生は行わなかった学生と比べて若干ではあるものの、早く就職活動が終了する傾向にあることがわかる。したがって、インターンシップのエントリーは就職活動終了日に影響を与えていることが示唆される。

3 単一潜在クラスモデルを用いた就職活動終了日予測法

3.1 概要

上述の通り、早川らの予測モデルは属性情報のみを用いている。しかし、事前分析により学生の行動情報もまた就職活動終了日に影響を与えていることがわかり、このデータを適切に活用することで予測精度の向上が期待される。さらに、学生の就職ポータルサイト上での行動が就職活動終了日にどのように影響を与えるかを明らかにすることができれば、サイト運営者が学生への支援施策を決める際の一助となると考えられる。本研究では、学生の属性情報と行動情報の双方を考慮した就職活動終了日予測モデルの提案を行う。

早川らにより就職活動終了日は混合ワイブル分布により近似されることが示されている。そのため、就職活動終了日の傾向について、同じ属性情報を持つ学生が一樣に同じではなく、いくつかの互いに異質なグループの集合であると仮定することが自然である。そこで、学生の就職活動終了日、属性情報、行動情報が潜在クラスにより規定される確率分布に従うと仮定し、その関係を表現するような確率的潜在クラスモデルを提案する。

本研究では、一般的な確率的潜在クラスモデルである Aspect Model (以下、AM) [2], [3] の考え方を応用する。AM は共起データに対する潜在クラスモデルであり、各変数が潜在クラスの下で独立な確率分布に従うと仮定している。提案モデルでは一人の学生の属性情報、行動情報、終了日を共起データとしてそれらの間に潜在クラスを仮定する。このデータは多様な変数を持つため、各変数に合った分布を仮定し、パラメータの推定を行う。

いま、 \mathbf{x}_n は行動情報の変数のベクトル、 y_n は属性情報の変数、 r_n は就職活動終了日を表すものとし、 n 番目の学生のデータを (\mathbf{x}_n, y_n, r_n) 、 N 人の学生の学習データの集合を $\mathcal{S} = \{(\mathbf{x}_n, y_n, r_n)\}_{n=1}^N$ と表現する。 K 個の潜在クラスの集合を $\{z_1, z_2, \dots, z_K\}$ とし、提案する潜在クラスモデルのグラフィカルモデルを図 3 に示す。

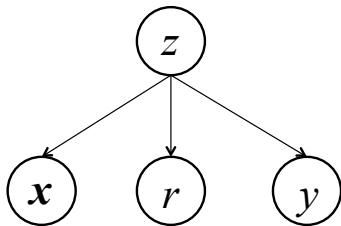


図 3: 単一潜在クラスモデルのグラフィカルモデル

いま、モデルに用いる学生の行動情報を早期エントリー数、インターンシップエントリーの有無とする。 n 番目の学生の早期エントリー数を $x_n^1 \in \mathcal{Z}^+$ 、インターンシップエントリーの有無を $x_n^2 \in \{0, 1\}$ とすると、行動情報のベクトルは $\mathbf{x}_n = (x_n^1, x_n^2)$ となる。ただし、 \mathcal{Z}^+ は非負整数の集合を表す。また、属性情報はその学生が所属する学校名とし、 n 番目の学生の所属する学校を $y_n \in \mathcal{C}$ とする。ただし、 \mathcal{C} は大学や大学院、高等専門学校などを含む学校の集合を表す。このとき、これらの同時確率は式 (2) で表される。

$$P(\mathbf{x}_n, y_n, r_n) = \sum_{k=1}^K P(z_k) P(x_n^1 | z_k) P(x_n^2 | z_k) P(y_n | z_k) P(r_n | z_k) \quad (2)$$

式 (2) において、 $P(r_n | z_k)$ は従来研究同様にワイブル分布を仮定し、 $P(y_n | z_k)$ には多項分布を仮定する。いま、 L 個の学校の集合を $\mathcal{C} = \{c_1, c_2, \dots, c_L\}$ とすると、この分布のパラメータは $P(c_l | z_k)$ ($l = 1, \dots, L$) となる。また、早期エントリー数の分布 $P(x_n^1 | z_k)$ は平均 μ_k と分散 σ_k^2 の正規分布に従うと仮定する。インターンシップエントリーの有無の分布 $P(x_n^2 | z_k)$ にはベルヌーイ分布を仮定する。変数 e を学生がインターンシップのエントリーを行っていたら 1、行っていなかったら 0 をとる確率変数とすると、この分布のパラメータは $P(e = 1 | z_k)$ となる。

3.2 パラメータ推定法

提案モデルのパラメータを推定するため、本研究では EM アルゴリズム [4] を適用する。学習データセットに対する対数尤度関数 LL は式 (3) のように示される。

$$LL = \sum_{n=1}^N \log P(\mathbf{x}_n, y_n, r_n) \quad (3)$$

EM アルゴリズムは対数尤度関数 LL を最大化するようなパラメータを E-step と M-step の繰り返し計算により求める。

E-step では、学習データ集合 \mathcal{S} と与えられたパラメータを用いて潜在クラス z_k の事後確率 $P(z_k | \mathbf{x}_n, y_n, r_n)$ を以下の式 (4) で計算する。

$$P(z_k | \mathbf{x}_n, y_n, r_n) = \frac{P(\mathbf{x}_n, y_n, r_n)}{\sum_k P(\mathbf{x}_n, y_n, r_n)} \quad (4)$$

M-step では式 (4) で計算した $P(z_k | \mathbf{x}_n, y_n, r_n)$ を用いてパラメータの更新を以下の式 (5)-(13) で行う。ただし、簡単のため $P(z_k | \mathbf{x}_n, y_n, r_n) = q_{k,n}$ とおく。

$$P(z_k) = \frac{1}{N} \sum_{n=1}^N q_{k,n} \quad (5)$$

$$\mu_k = \frac{\sum_n q_{k,n} x_n^1}{\sum_n q_{k,n}} \quad (6)$$

$$\sigma_k^2 = \frac{\sum_n q_{k,n} (x_n^1 - \mu_k)^2}{\sum_n q_{k,n}} \quad (7)$$

$$P(e = 1 | z_k) = \frac{\sum_n \delta(x_n^2, 1) q_{k,n}}{\sum_n q_{k,n}} \quad (8)$$

$$P(c_l | z_k) = \frac{\sum_n \delta(y_n, c_l) q_{k,n}}{\sum_n q_{k,n}} \quad (9)$$

$$a_k = \frac{\sum_n q_{k,n}}{\sum_n q_{k,n} (r_n)^{b_k}} \quad (10)$$

$$b_k = \frac{\sum_n q_{k,n}}{\sum_n q_{k,n} \log(r_n) \{a_k (r_n)^{b_k} - 1\}} \quad (11)$$

$$m_k = b_k \quad (12)$$

$$\eta_k = \left(\frac{1}{a_k}\right)^{\frac{1}{b_k}} \quad (13)$$

ただし、 $\delta(a, b)$ は $a = b$ のとき 1, $a \neq b$ のとき 0 をとるインジケータ関数とする。また、ワイブル分布のパラメータは式 (10), (11) を繰り返し計算し、 a_k, b_k の値が収束したとき式 (12), (13) により推定する。

3.3 就職活動終了日の予測

パラメータの推定後、学習したモデルを新たな学生のデータに対する就職活動終了日の予測に用いる。ここで、 t 番目の学生データを (\mathbf{x}_t, y_t) とすると、このデータに対する就職活動終了日の予測値 \hat{r}_t を本研究では計算の便宜上、以下の式 (14), (15) を用いて算出する。

$$\hat{r}_t = \sum_{k=1}^K h_k \hat{P}(z_k | \mathbf{x}_t, y_t) \quad (14)$$

$$\hat{P}(z_k | \mathbf{x}_t, y_t) = \frac{P(z_k)P(x_t^1 | z_k)P(x_t^2 | z_k)P(y_t | z_k)}{\sum_k P(z_k)P(x_t^1 | z_k)P(x_t^2 | z_k)P(y_t | z_k)} \quad (15)$$

ここで、 h_k は潜在クラス z_k における終了日の予測値であり、本研究では潜在クラス z_k のワイブル分布の中央値を用いる。 $\hat{P}(z_k | \mathbf{x}_t, y_t)$ は t 番目の学生の潜在クラスへの所属確率を表す。したがって、式 (14) は潜在クラスの終了日の予測値をその所属確率により重みづけ和をとる計算を行っている。

4 単一潜在クラスモデルの実験

4.1 実験条件

提案手法の有効性を確認するため、ある就職ポータルサイトのデータを用いたシミュレーション実験を行った。実験では、2013 年 3 月卒の学生のデータを学習データ、2014 年 3 月卒の学生のデータをテストデータとして用いる。学習データの学生数、学校数はそれぞれ約 100,000、約 1,500 である。テストデータの学生数は約 100,000 である。比較手法として、行動情報を考慮せず属性情報と就職活動終了日のみを考慮した潜在クラスモデルを用いる。

評価指標として、予測問題で一般的に用いられる指標であり、解釈が容易である平均絶対誤差 (MAE) を用いる。テストデータの数を T とし、 t 番目の学生の終了日を r_t とすると、MAE は以下の式 (16) で定義される。

$$MAE = \frac{1}{T} \sum_{t=1}^T |r_t - \hat{r}_t| \quad (16)$$

4.2 実験結果・考察

表 2 に属性情報のみを用いたモデルによる予測と全変数を用いたモデルによる予測の MAE を示す。

表 2: MAE の比較

K	属性情報のみ	全変数
5	78.782	80.035
10	78.670	79.532
15	78.630	79.282
20	78.634	79.438

表 2 より、属性情報のみを用いた手法の方が予測精度が高いことがわかる。これは属性情報が行動情報よりも大きく終了日に影響を与えているためであると考えられる。また、行動情報を加えたモデルの精度が低いのは行動情報のパラメータが過学習を起こしているためであると想定される。推定されたパラメータの値を見ると、行動情報のパラメータは潜在クラス間で大きく異なっているが、終了日の予測値 h_k はいくつかの潜在クラスで似た値をとっていることが明らかとなった。これにより、潜在クラスは終了日より行動履歴の差異を強く考慮して構築されたと考えられる。

5 複合潜在クラスモデルを用いた就職活動終了日

予測法

5.1 概要

前節の実験により、属性情報と終了日のモデルは精度の良い予測ができるが、行動情報を加えた場合に過度に適応するモデルになってしまうという問題が明らかとなった。そこで、ユーザとアイテムそれぞれに潜在クラスを仮定することで AM を拡張した Flexible Mixture Model [5] の考え方を応用し、学生の属性情報と行動情報それぞれに潜在クラスを仮定し、その潜在クラスの組合せにより終了日の分布が推定されるようなモデルを考える。

いま、行動情報と属性情報の潜在クラス集合をそれぞれ $\{u_1, u_2, \dots, u_I\}$, $\{v_1, v_2, \dots, v_J\}$ と定義する。ただし、 I と J はそれぞれの潜在クラス数を表す。このとき、モデルは式 (17)、そのグラフィカルモデルは図 4 で表される。

$$P(\mathbf{x}_n, y_n, r_n) = \sum_{i=1}^I \sum_{j=1}^J P(u_i)P(v_j)P(x_n^1 | u_i)P(x_n^2 | u_i) \times P(y_n | v_j)P(r_n | u_i, v_j) \quad (17)$$

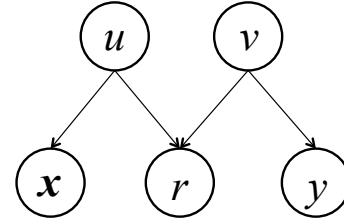


図 4: 複合潜在クラスモデルのグラフィカルモデル

5.2 パラメータ推定法と終了日の予測

3.2 節と同様に EM アルゴリズムによりパラメータの推定を行う。E-step, M-step はそれぞれ以下ようになる。ただし、 $P(u_i, v_j | \mathbf{x}_n, y_n, r_n) = q'_{i,j,n}$ とおく。

[E-step]

$$P(u_i, v_j | \mathbf{x}_n, y_n, r_n) = \frac{P(\mathbf{x}_n, y_n, r_n)}{\sum_{i,j} P(\mathbf{x}_n, y_n, r_n)} \quad (18)$$

[M-step]

$$P(u_i) = \frac{1}{N} \sum_{n=1}^N \sum_{j=1}^J q'_{i,j,n} \quad (19)$$

$$P(v_j) = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^I q'_{i,j,n} \quad (20)$$

$$\mu_i = \frac{\sum_n \sum_j q'_{i,j,n} x_n^1}{\sum_n \sum_j q'_{i,j,n}} \quad (21)$$

$$\sigma_i^2 = \frac{\sum_n \sum_j q'_{i,j,n} (x_n^1 - \mu_i)^2}{\sum_n \sum_j q'_{i,j,n}} \quad (22)$$

$$P(e = 1 | u_i) = \frac{\sum_n \sum_j \delta(x_n^2, 1) q'_{i,j,n}}{\sum_n \sum_j q'_{i,j,n}} \quad (23)$$

$$P(c_l | v_j) = \frac{\sum_n \sum_i \delta(y_n, c_l) q'_{i,j,n}}{\sum_n \sum_i q'_{i,j,n}} \quad (24)$$

$$a_{ij} = \frac{\sum_n q'_{i,j,n}}{\sum_n q'_{i,j,n} (r_n)^{b_{ij}}} \quad (25)$$

$$b_{ij} = \frac{\sum_n q'_{i,j,n}}{\sum_n q'_{i,j,n} \log(r_n) \{a_{ij} (r_n)^{b_{ij}} - 1\}} \quad (26)$$

$$m_{ij} = b_{ij} \quad (27)$$

$$\eta_{ij} = \left(\frac{1}{a_{ij}} \right)^{\frac{1}{b_{ij}}} \quad (28)$$

終了日の予測は3.3節と同様の計算を行う。ただし、式(14)、(15)の代わりに以下の式(29)、(30)を用いる。

$$\hat{r}_t = \sum_{i=1}^I \sum_{j=1}^J h_{ij} \hat{P}(u_i, v_j | \mathbf{x}_t, \mathbf{y}_t) \quad (29)$$

$$\begin{aligned} \hat{P}(u_i, v_j | \mathbf{x}_t, \mathbf{y}_t) \\ = \frac{P(u_i)P(v_j)P(x_t^1|u_i)P(x_t^2|u_i)P(y_t|v_j)}{\sum_{i,j} P(u_i)P(v_j)P(x_t^1|u_i)P(x_t^2|u_i)P(y_t|v_j)} \end{aligned} \quad (30)$$

6 複合潜在クラスモデルの実験

6.1 実験条件

行動情報と属性情報それぞれに潜在クラスを仮定した提案モデルの有効性を示すために4節の実験と同様にシミュレーション実験を行う。実験条件は4節の実験と同様であり、行動情報と属性情報それぞれの潜在クラス数を変化させた実験を行う。評価指標は式(16)で表されるMAEを用いる。

6.2 実験結果・考察

表3に行動情報と属性情報のそれぞれの潜在クラス数におけるMAEの比較を示す。

表3: MAEの比較2

$I \backslash J$	2	3	4	5	6	7	8
2	78.788	78.594	78.476	78.420	78.381	78.386	78.424
3	78.652	78.498	78.387	78.333	78.337	78.346	78.338
4	78.655	78.480	78.410	78.324	78.341	78.299	78.321
5	78.668	78.499	78.417	78.323	78.357	78.338	78.359
6	78.671	78.499	78.402	78.322	78.312	78.369	78.335
7	78.696	78.465	78.405	78.324	78.326	78.363	78.374
8	78.726	78.451	78.393	78.377	78.310	78.443	78.369

表3より、 $(I, J) = (4, 7)$ が最適であり、4節の実験のどの手法よりも精度が高いことがわかる。したがって、属性情報に加えて行動情報を考慮することは終了日の予測精度向上に有用であるといえる。行動情報と属性情報それぞれに対して潜在クラスを仮定することで、属性情報と終了日の統計的性質を取り込みつつ、それに加えて行動情報を考慮するような潜在クラスモデルを構築できたと考えられる。また、最適な潜在クラス数から潜在クラス数を増やしていくと精度が低下しているが、提案手法は $I \times J$ 個の終了日の分布を仮定するため、過学習が起きやすいと考えられる。

ある程度予測精度が高い潜在クラス数の組み合わせに対しその詳細を検討するため、潜在クラス数 $(I, J) = (4, 5)$ としたときの潜在クラスの組合せによる終了日の予測値 h_{ij} を表4に示す。また、行動情報の潜在クラスごとのパラメータの推定値を表5に示す。

表4: 各潜在クラスの終了日の予測値

$u_i \backslash v_j$	1	2	3	4	5
1	2013/4/19	2013/6/7	2013/9/17	2014/1/3	2014/3/7
2	2013/4/22	2013/6/2	2013/8/30	2013/12/17	2014/3/1
3	2013/5/3	2013/6/10	2013/8/24	2013/12/7	2014/2/27
4	2013/4/23	2013/5/27	2013/8/3	2013/11/15	2014/2/19

表5: 行動情報のパラメータ推定値

u_i	μ_i	σ_i	$P(e=1 u_i)$	$P(u_i)$
1	3.47	2.01	0.0429	0.1576
2	13.92	6.63	0.1183	0.3401
3	38.20	16.87	0.2082	0.3914
4	86.93	46.91	0.2931	0.1108

表4の列に着目すると、ある属性情報の潜在クラスにおける行動情報の潜在クラスの変化に対する就職活動終了日の変化を読み取ることができる。反対に行に着目すると、同様に行動情報の潜在クラスについて読み取ることができる。列方向でみた場合、行方向でみた場合と比べて終了日のばらつきが小さいことから、提案手法は行動情報よりも属性情報を重視して終了日を予測するモデルになっていることがわかる。また、属性情報の潜在クラスごとに行動情報の潜在クラスによる終了日のばらつきが異なることもわかる。これは属性情報によって行動情報が終了日に与える影響の度合いが異なることを示す。

属性情報の潜在クラス4に着目すると、行動情報の潜在クラスによるばらつきが大きい。他に比べ、行動情報による影響が大きいことがわかる。この潜在クラスは就職活動終了日が比較的遅い学校群である。表5と併せてみると、この潜在クラスでは早期エンタリ数の多さとインターンシップエンタリが就職活動の早期終了につながっていることがわかる。多くの企業に早期からエンタリする学生やインターンシップのエンタリを行う学生は就職活動に意欲的であり、その終了日も早くなると考えられる。一方、属性情報の潜在クラス1は就職活動終了日が早い学校群であり、行動情報による影響が小さい。これは就職活動終了日が早い学校群においては行動情報によって大きな差異が生まれないことを示している。

このように、属性情報によって行動情報が終了日に与える影響は異なるという複雑な関係性を持つため、支援施策を決める際には属性情報を考慮し、それぞれに適した施策を考える必要があると考えられる。

7 まとめと今後の課題

本研究では、学生の属性情報と行動情報、就職活動終了日の三変数の関係を表現する潜在クラスモデルによる就職活動終了日の予測法を提案した。提案モデルの有効性を実際の就職ポータルサイトのデータを用いたシミュレーション実験により示した。また、分析により学生の属性情報と行動情報、就職活動終了日の関係性を明らかにした。

今後の課題としては、行動情報として用いる変数の追加が考えられる。今回用いた早期エンタリ数やインターンシップエンタリの有無の他に、人気企業へのエンタリや特定の業種へのエンタリも就職活動終了日に影響を与えている可能性がある。そのため、これらの変数と就職活動終了日の関係の分析やモデルへの追加による予測精度への影響の確認などを行う必要がある。

参考文献

- [1] 早川真央, 三川健太, 石田崇, 後藤正幸, 小川晋一郎, “層別木と混合ワイブル分布に基づく就職活動終了時期の予測モデル,” 第36回情報理論とその応用シンポジウム予稿集, 2013.
- [2] T. Hofmann, “Probabilistic Latent Semantic Analysis,” *Proc. UAI*, pp. 289–296, 1999.
- [3] T. Hofmann, “Collaborative Filtering via Gaussian Probabilistic Latent Semantic Analysis,” *Proc. of 26th Ann. International ACM SIGIR Conf.*, pp. 259–266, 2003.
- [4] A. P. Dempster, N. M. Laird, D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of Royal Statistical Society. Series B*, vol. 39, no. 1, pp. 1–38, 1977.
- [5] L. Si, R. Jin, “Flexible Mixture Model for Collaborative Filtering,” *Proc. International Conference on Machine Learning*, vol. 2, pp. 704–711, 2003.