

2026 年度修士論文

# マルチモーダルな GNN における 説明可能な影響力予測

Influencer Rank アーキテクチャへの mask 適用による成長要因の分析手法の検討

渡邊 遼真

2025 年 12 月 20 日

早稲田大学 先進理工学研究科  
電気・情報生命専攻  
情報学習システム研究室  
指導教員 村田昇

# 目次

<b>1</b>	<b>導入</b>	<b>1</b>
1.1	概要	1
1.2	本研究の目的	2
1.3	本論文の構成	2
<b>2</b>	<b>前提知識</b>	<b>3</b>
2.1	グラフニューラルネットワーク (GNN)	3
2.1.1	グラフ表現	3
2.1.2	GCN による近傍集約	4
2.2	時系列情報の処理と注意機構	4
2.2.1	時系列データの統合 (RNN/LSTM)	4
2.2.2	注意機構 (Temporal Attention)	4
2.3	説明可能 AI (XAI)	5
2.3.1	グラフ予測における説明の考え方	5
2.3.2	マスク最適化に基づく説明 (GNNExplainer 系)	5
2.4	評価指標	6
2.4.1	NDCG (Normalized Discounted Cumulative Gain)	6
2.4.2	RBP (Rank-biased Precision)	6
<b>3</b>	<b>問題設定</b>	<b>7</b>
3.1	対象とするデータ	7
3.1.1	データの特徴	7
3.1.2	データの詳細	7
3.1.3	データセット選定の妥当性と限界	7
3.2	データと予測タスクの結びつき	8
3.3	問題の定式化	8
3.3.1	入力と出力の定義	8
3.3.2	影響力 (エンゲージメント) 指標	8
3.3.3	損失関数による最適化 (ListMLE)	9
<b>4</b>	<b>先行手法</b>	<b>10</b>
4.1	インフルエンサー予測モデル: InfluencerRank	10
4.1.1	問題設定と影響力指標	10
4.1.2	異種混合グラフによるマルチモーダル統合	11
4.1.3	空間表現学習: 各時点グラフに対する GNN 適用	11
4.1.4	時系列統合: GRU と時間注意機構	12
4.1.5	ランキング学習としての最適化	12
4.2	GNN に対する説明可能 AI 手法: GNNExplainer	12
4.2.1	説明対象: 構造マスクと特徴マスク	13
4.2.2	マスク最適化	13

---

<b>5</b>	<b>提案手法</b>	<b>14</b>
5.1	End-to-End マスク最適化による時系列グラフ予測の説明可能 AI	14
5.1.1	問題設定と予測モデル	14
5.1.2	説明対象の時刻選択（効く月だけ説明する）	14
5.1.3	End-to-End Mask Optimization (MaskOpt)	15
5.1.4	符号付き影響度 (Score Impact) の算出	17
5.1.5	GNNExplainer との比較	18
<b>6</b>	<b>実験</b>	<b>19</b>
6.1	実験の目的	19
6.2	実験設定	19
6.2.1	使用データセットと対象ノード	19
6.2.2	学習済みモデル（固定）	19
6.2.3	説明（貢献度）算出の定義	19
6.3	説明の妥当性検証：学習なしで可能な検証	20
6.4	実験1：マスク値の確認（1月，2月，3月）	21
6.4.1	目的	21
6.4.2	手順	21
6.4.3	結果	21
6.5	実験2：重要度と実影響の整合性チェック（1月，2月，3月）	21
6.5.1	目的	21
6.5.2	手順	21
6.5.3	結果	23
6.6	まとめ	23
<b>7</b>	<b>考察</b>	<b>25</b>
7.1	貢献度算出に関する考察	25
7.1.1	マスク値の分布と探索範囲削減としての有用性	25
7.1.2	重要度（mask 値）と符号付き影響度（score impact）の整合性	25
7.2	手法の限界	26
7.2.1	最適化規模とサンプル数の制約	26
7.2.2	検証範囲の限定（単一ノード・少数時刻）	26
7.2.3	相互作用（交互作用）を十分に扱えていない	26
7.3	今後の展望	26
7.3.1	重要度順の追加・削除による反実仮想検証	26
7.3.2	条件が近いユーザー群での再現性の確認	26
7.3.3	相互作用を考慮した貢献度の拡張	26
7.3.4	説明の安定化と評価指標の整備	27
<b>8</b>	<b>結論</b>	<b>28</b>
<b>9</b>	<b>謝辞</b>	<b>29</b>



# 第1章 導入

## 1.1 概要

近年、スマートフォンと SNS の普及に伴い、デジタルマーケティングは急速に拡大している。中でも、SNS 上で影響力を持つ個人（インフルエンサー）を介して情報を届けるインフルエンサーマーケティングは、企業の認知獲得やブランディング、購買行動の喚起において重要な施策として位置付けられている **Lou2019, Campbell2020, Hudders2021**。

マーケティング実務における重要課題の一つは、限られた予算の中で「誰に依頼すべきか」を意思決定することである。しかし、著名なトップインフルエンサーは依頼費用が高額であり、競合との獲得競争も激しい **Haenlein2020**。そこで近年は、現時点ではフォロワー規模が大きいくとも将来的に伸長が見込まれる人材を早期に発掘し、長期的関係を構築する選定（以下、青田買い）が注目されている **Hudders2021**。

将来伸長するインフルエンサーの発掘には、投稿頻度、投稿内容（テキストや画像）、ハッシュタグ選択、他者との関係性（メンションや共起）など、多様な要因が複合的に関与する **Casalo2020**。さらに、これらは時系列に沿って変化し、SNS 上の関係構造も動的に推移する。そのため、単純な静的指標（フォロワー数等）だけでは将来の影響力を捉えにくい。また、不正アカウントや購入フォロワー等により指標が歪む可能性も指摘されている **Cresci2015**。

このような背景のもと、異種混合グラフと時系列学習を組み合わせ、将来伸長が見込まれる有望なインフルエンサーをランキングする手法が提案されている。例えば **InfluencerRank** は、各期間の異種グラフをグラフ畳み込み (GCN) で表現し、その推移を再帰モデルと注意機構で統合することで、将来の影響力を推定する **Kim2023InfluencerRank**。一方で、深層学習に基づく予測はブラックボックスになりやすく、実務上は「誰が伸びるか」だけでなく、「なぜ伸びると判断したか」という根拠提示が、投資判断・説明責任・施策設計の観点から重要である **Gunning2019, Arrieta2020XAI, Liao2020QuestioningAI, Rai2020GlassBox**。

本研究では、**InfluencerRank** 型のマルチモーダル時系列グラフ予測モデルに対して、最適化ベースのマスク学習により説明を与える枠組みを検討する。具体的には、予測をできるだけ維持しつつ入力（ノード特徴量およびエッジ）を疎に残すマスクを学習し、予測に寄与する要素をコンパクトに抽出する。これは **GNNEExplainer** に代表される、予測と説明の両立（忠実性と簡潔性のトレードオフ）に基づく考え方である **Ying2019GNNEExplainer**。本枠組みにより、画像・テキスト・数値特徴量とグラフ構造の双方について、どの要素が将来スコアに寄与したかを月次で比較可能な形で提示することを目指す。

また、実務では「いいね数の非表示」等のように可観測な評価指標が変化し得る。これに対し本研究の枠組みは、過去のエンゲージメント値そのものを入力として用いず、投稿コンテンツ（画像・テキスト）や行動・関係性といった情報から将来指標を推定することを重視する。この点は、観測可能な指標や表示仕様が変化しても、コンテンツとネットワーク形成に基づく評価・予測へ接続しやすいという実務上の利点に繋がると考えられる。

本研究のアプローチは以下の 2 点の統合からなる。

1. 異種混合グラフと時系列モデル（本研究では LSTM）による将来エンゲージメント推定と、インフルエンサーノードのランキング (**InfluencerRank** 型モデル) **Kim2023InfluencerRank**
2. 最適化に基づくマスク学習により、ノード特徴量およびエッジの寄与を抽出し、さらに予測スコアの変化に基づく符号付き影響度を定義して月次比較可能な形で提示する説明手法 **Ying2019GNNEExplainer**

## 1.2 本研究の目的

本研究の目的は、インフルエンサーマーケティングにおける青田買い意思決定を支援するために、将来予測の精度と説明可能性を両立する枠組みを構築することである。具体的には、(i) マルチモーダルかつ時系列・グラフ構造を含む入力から将来の影響力スコアを推定し、(ii) その推定に対して、どの特徴量・どの関係性がどの時点で効いたのかを、予測に整合する形で提示する。本研究では影響力（エンゲージメント）を、いいね数とコメント数の合計をフォロワー数で正規化した指標として扱い、ランキング対象はインフルエンサーノードに限定する。これにより、候補者選定の納得感を高めるとともに、施策（投稿頻度の改善、ハッシュタグ方針、コラボ関係の設計等）に繋がられる分析基盤の提供を目指す **Rai2020GlassBox**, **Liao2020QuestioningAI**。

## 1.3 本論文の構成

本論文は全9章から構成される。本文中では、章・節・小節の番号をそれぞれ「章番号」、「章番号. 節番号」、「章番号. 節番号. 小節番号」（例：2.1, 2.3.1）の形式で表記する。

第2章では、インフルエンサーマーケティングにおけるインフルエンサー選定の背景と課題を整理した上で、「将来伸長するインフルエンサー」を予測する既存研究を概観する。特に、**InfluencerRank** 型モデルについて、異種混合グラフの表現、GCN による各期間グラフの表現学習、時系列モデルと注意機構による統合、ランキング学習および評価指標の位置づけを述べる **Kim2023InfluencerRank**。

第3章では、本研究の問題設定を与える。入力として用いるデータ（各時点の異種混合グラフ、ノード特徴量、エッジ集合、対象とするインフルエンサーノード、観測窓長など）と、出力として扱う量（将来時点の影響力スコアおよびランキング）を定式化する。また、説明の対象（どの時点のどの要素を説明するか）と、説明に求める要件（忠実性・簡潔性・比較可能性）を整理する。

第4章では、先行研究および本研究のベースとなる説明手法を整理し、時系列グラフ予測モデルに適用可能な最適化ベースのマスク学習手法を述べる。具体的には、時点ごとのマスク付与、直列モデル（GCN → 時系列モデル/Attention → MLP）に対する説明の設計、月次比較を行うための集約や正規化の方針を示す。さらに、特徴量とエッジの両方について、予測スコアの変化に基づく符号付き影響度を定義し、その計算手順と解釈を述べる。

第5章では、実データを用いた追試実験を行い、予測性能（ランキング指標）と説明結果を評価する。どの特徴量群およびどの関係性が予測に寄与したか、それらが月次でどのように変化したか、重要度と符号付き影響度が一致・乖離する状況は何か、といった観点から結果を整理する。加えて、反実仮想的な検証や安定性の観点から説明の妥当性を議論する。

第6章では、本研究の限界と今後の展望を述べる。マスク最適化に内在する不安定性や多解性、ベースライン設定や正則化の影響、時系列比較におけるスケーリングの課題、因果性との混同リスクなどを整理し、改善の方向性を議論する。

第7章では考察として、本研究で得られた知見を実務上の意思決定へ接続する観点からまとめる。

第8章では結論として、本研究の貢献を総括する。

第9章では謝辞を述べる。

## 第2章 前提知識

本研究で扱う課題は、(i) グラフ構造（ユーザー間関係や投稿内容の関係性）、(ii) 時系列変化（成長・バズ・関係性の変動）、(iii) 深層学習モデルの説明可能性（XAI）という複数要素が同時に現れる点に特徴がある。本章では、提案手法の理解に必要な基礎として、グラフニューラルネットワーク、時系列モデルと注意機構、説明可能 AI の代表的枠組み、ならびに評価指標を整理する。

### 2.1 グラフニューラルネットワーク（GNN）

#### 2.1.1 グラフ表現

グラフは  $G = (V, E)$  と表し、 $V$  はノード集合、 $E$  はエッジ集合である。ノード数を  $n = |V|$  とすると、隣接行列  $A \in \{0, 1\}^{n \times n}$  は  $(i, j) \in E$  のとき  $A_{ij} = 1$ （それ以外は 0）で定義される。各ノード  $v$  は特徴ベクトルを持ち、特徴行列を  $X \in \mathbb{R}^{n \times d}$  とする。

本研究の対象では、インフルエンサー・一般ユーザー・ハッシュタグ・画像オブジェクト等の異種ノードが存在し、またフォロー・メンション・共起など複数種類の関係が現れるため、グラフは一般に異種（heterogeneous）である。前提知識として本節では、まず標準的な同種グラフ（単一関係種別）を想定した GNN を述べ、異種性や複数関係種別の扱いは後述（第 4 章および第 5 章）で整理する。なお SNS では有向関係が現れることが多いが、本章では説明の簡潔化のため隣接行列  $A$  による表現を用い、有向性の影響は後段で補足する。

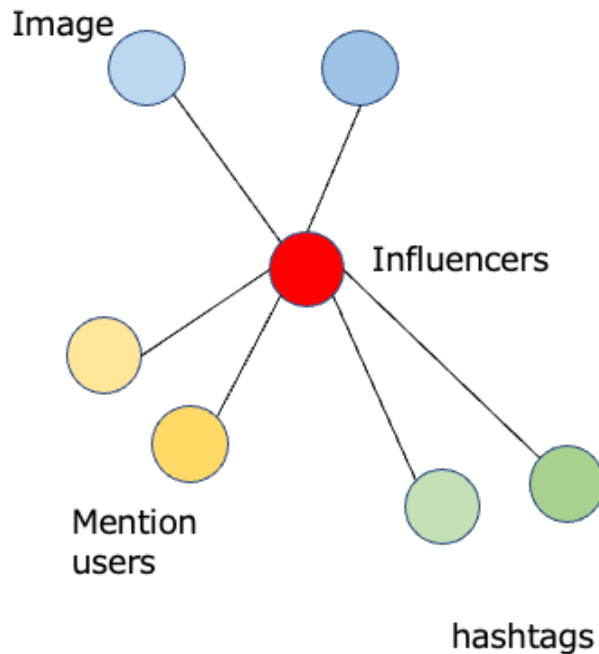


図 2.1: 異種混合グラフの例（ノード種別・関係種別を含む）

### 2.1.2 GCN による近傍集約

GNN の基本的な考え方は、各層で近傍ノードの情報を集約し、変換して次層の表現を得ることである。代表例として、Kipf & Welling による Graph Convolutional Network (GCN) は、層  $l$  のノード表現  $H^{(l)} \in \mathbb{R}^{n \times d_l}$  から次層  $H^{(l+1)}$  を次式で更新する **kipf2017gcn** :

$$\begin{aligned}\tilde{A} &= A + I, \quad \tilde{D}_{ii} = \sum_j \tilde{A}_{ij}, \\ \hat{A} &= \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}}, \\ H^{(l+1)} &= \sigma(\hat{A} H^{(l)} W^{(l)}).\end{aligned}$$

ここで、 $I$  は単位行列、 $\tilde{D}$  は次数行列、 $W^{(l)}$  は学習可能パラメータ、 $\sigma(\cdot)$  は非線形活性化関数（例：ReLU）である。通常、入力特徴を  $H^{(0)} = X$  とおくことで、各層の更新によりグラフ構造を反映した表現を得る。この更新により、ノードは自身と近傍から情報を取り込み、グラフ構造を反映した表現を獲得できる。

なお、GCN は GNN の一例であり、より一般的にはメッセージパッシングとして統一的に理解できる **gilmer2017mpnn**。本研究では、各時点で構築したグラフに対して GNN (GCN) を適用し、時点ごとの空間的（構造的）表現を得る。

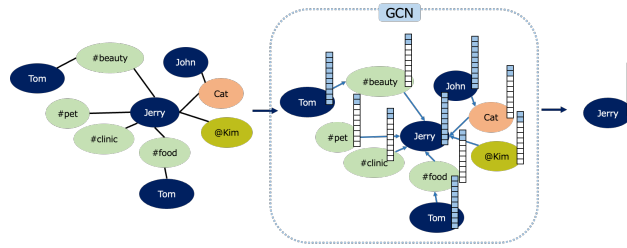


図 2.2: GCN による近傍集約の概念図

## 2.2 時系列情報の処理と注意機構

### 2.2.1 時系列データの統合 (RNN/LSTM)

インフルエンサーの影響力は時間とともに変化するため、単一時点のグラフ表現だけでなく、複数時点の情報を統合する必要がある。時点  $t$  における GNN の出力（ノード表現）を  $R_t$  とすると、再帰型ニューラルネットワーク (RNN) により時系列方向へ統合し、隠れ状態  $S_t$  を更新する：

$$S_t = \text{RNN}(R_t, S_{t-1}).$$

RNN には複数の実装があり、長期依存を扱うために LSTM **hochreiter1997lstm** や GRU **cho2014gru** が用いられることが多い。本研究では時系列統合に LSTM を採用する。また、InfluencerRank 型モデルは GNN で得た各時点表現を再帰モデルと注意機構で統合し、最終的に将来スコアを推定する（詳細は第 4 章および第 5 章で述べる）。

### 2.2.2 注意機構 (Temporal Attention)

全ての時点が予測へ等しく寄与するとは限らない。例えば、特定月の急成長やバズ、コラボ関係の変化が将来スコアに強く影響することがある。そこで、各時点の隠れ状態  $S_t$  に対し



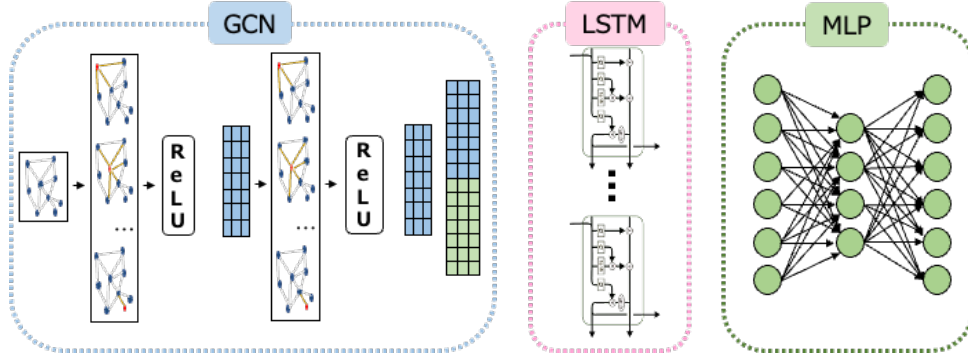


図 2.3: 時系列グラフ処理の流れ（各時点：GNN、時系列統合：RNN/LSTM、重要時点の強調：Attention）

て重み  $\alpha_t$  を学習し、重要な時点を強調する注意機構を用いる **bahdanau2015attention**。一般的な定式化の一例として、

$$e_t = v_a^\top \tanh(W_a S_t + b_a),$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{\tau=1}^T \exp(e_\tau)},$$

$$S_{\text{final}} = \sum_{t=1}^T \alpha_t S_t$$

を用いることで、時系列全体から予測に有用な情報を集約できる。ここで  $W_a, v_a, b_a$  は学習可能パラメータである。

## 2.3 説明可能 AI (XAI)

### 2.3.1 グラフ予測における説明の考え方

深層学習は高い表現能力を持つ一方で、予測根拠がブラックボックス化しやすい。マーケティングの意思決定では「誰が伸びるか」だけでなく「なぜそう判断したか」が重要であるため、予測モデルに対する説明可能性の付与が求められる **Gunning2019, Arrieta2020XAI**。

グラフ予測に対する説明は、大きく (1) 特徴量（ノード属性）の寄与と、(2) 構造（どのエッジ／どの部分グラフが重要か）の二面から捉えられる。本研究はこの両者を同一枠組みで扱い、さらに時系列（どの月が効いたか）まで比較可能にすることを目指す。

### 2.3.2 マスク最適化に基づく説明 (GNExplainer 系)

本研究では、最適化に基づいて予測をできるだけ維持しつつ、入力（特徴量・エッジ）を疎に残すマスクを学習し、予測に寄与する要素をコンパクトに抽出するタイプの説明を採用する。代表例である GNExplainer は、エッジマスク  $m_E \in [0, 1]^{|E|}$  と特徴マスク  $m_X \in [0, 1]^d$  を導入し、マスク後の入力でも予測をできるだけ保つ一方で、マスクを疎にするよう学習する **Ying2019GNExplainer**。

直観的には、

- 忠実性 (fidelity) : マスク後も予測が大きく変わらない（説明が予測に整合する）
- 簡潔性 (sparsity) : 少数の特徴・少数のエッジで説明できる

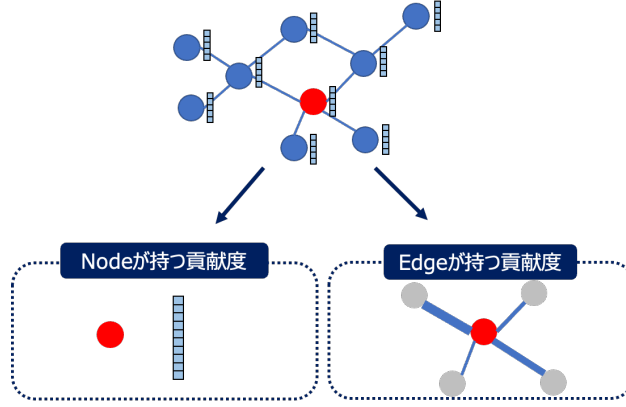


図 2.4: グラフ予測における説明の対象：特微量・構造・時系列

を両立するように、マスクを最適化する。時系列グラフモデルへの適用方法、ならびに特微量・エッジの符号付き影響度の定義と計算は第5章で詳述する（関連する先行枠組みの整理は第4章で述べる）。

## 2.4 評価指標

本研究は将来有望なインフルエンサーの順位付けを目的とするため、回帰誤差だけでなくランキング指標で性能を評価する。本節では代表的な指標としてNDCGとRBPを述べる。

### 2.4.1 NDCG (Normalized Discounted Cumulative Gain)

NDCGは、上位に関連度の高いアイテムが配置されるほど高く評価する指標である `jarvelin2002ndcg`。順位  $i$  の関連度を  $rel(i)$  とし、ゲイン関数  $g(\cdot)$  を用いると、

$$DCG@k = \sum_{i=1}^k \frac{g(rel(i))}{\log_2(i+1)},$$

$$NDCG@k = \frac{DCG@k}{IDCG@k}$$

と定義される。ここで  $IDCG@k$  は理想順位における  $DCG@k$  である。 $g(rel) = rel$  あるいは  $g(rel) = 2^{rel} - 1$  など複数の流儀があるため、本研究で採用する関連度設計とゲイン関数は実験設定（第6章）で明記する。

### 2.4.2 RBP (Rank-biased Precision)

RBPは、ユーザが上位から確率  $p$  で閲覧を継続するという行動モデルに基づく指標である `moffat2008rbp`。

$$RBP = (1-p) \sum_{i=1}^N rel_i p^{i-1}$$

で定義される。ここで  $N$  は評価対象のランキング長である。 $p$  が大きいほど下位順位までを広く評価し、 $p$  が小さいほど上位を重視する。マーケティングでは上位候補の抽出が重要であるため、本研究でもRBPを用いて上位重視の評価を行う。

## 第3章 問題設定

本章では、本研究が扱うデータと予測タスクを整理し、時系列異種混合グラフに基づくランキング学習問題として定式化する。特に、(i) マルチモーダルな投稿コンテンツ、(ii) ユーザー間関係を表すグラフ構造、(iii) 時系列変化を同時に扱う点を明確にする。

### 3.1 対象とするデータ

本研究で扱うのは、SNS プラットフォームから収集された大規模なマルチモーダルデータである。インフルエンサーマーケティングに関わるデータは、概ね以下の要素を含む。

#### 3.1.1 データの特徴

1. 画像、テキスト、数値など異なるモダリティの情報が混在している
2. ユーザー間のフォロー関係やメンションなど、複雑なネットワーク構造を持っている
3. 時間の経過とともに投稿や関係性が変化する時系列データである
4. 過去のデータであるため、未来時点の正解（エンゲージメント）が確定している

本研究では特に、上記 1 および 2 の複合的な情報がインフルエンサーの影響力を形成する状況を想定する。このようなデータが得られる例としては、Instagram の投稿データ、Twitter の拡散ログ、YouTube の視聴者推移データなどが挙げられる。本研究では、既存研究で用いられている Instagram データセットを用いて追試と分析を行うため、以下でデータの内容と本研究目的との適合を述べる。

#### 3.1.2 データの詳細

本研究で扱う Instagram データは、インフルエンサーの活動記録であり、マルチモーダルデータの一種である。このデータは、ユーザー、投稿、タグなどが相互に関連し合う異種混合グラフとして表現できる。データには、投稿画像の輝度や色温度などの画像特徴、キャプションのテキスト長などのテキスト特徴、ハッシュタグ数やメンション情報といった行動・関係性の情報が含まれる。また、「いいね」数やコメント数などの反応量も含まれるが、本研究ではこれらを将来予測の正解ラベルとして扱い、モデル入力には直接用いない（詳細は 3.2 節および 3.3 節で述べる）。以上より、本データはコンテンツの内容とユーザー間のつながりを同時に含むため、「誰がなぜ人気になるのか」を分析するという目的に適合する。

#### 3.1.3 データセット選定の妥当性と限界

本研究で使用する Instagram Influencer Dataset **Kim2023InfluencerRank** は、本来インフルエンサーの興味・関心に基づくカテゴリ分類を目的として構築されたものである。そのため、本研究が目指す時系列での影響力予測（Influencer Rank）および青田買いシステムの構築を行う上では、以下の制約が存在する。

1. 完全なソーシャルグラフの欠如: 本データセットにはフォロー・フォロワー関係の完全なグラフが含まれていない。これに対し本研究では、投稿内のメンション (@usertags) やコメントの共起関係などから擬似的なネットワークを構築し、これをユーザー間の関係性の代替指標として利用することで対処する（具体的な構築方法は第6章で述べる）。
2. データの適時性とトレンドの乖離: 収集期間が2017年であり、現在のトレンド（短尺動画の台頭など）とは乖離がある。ただし、本研究では「投稿コンテンツとネットワーク形成がエンゲージメントに影響する」という基本的なメカニズムに注目し、その有効性が一定程度保たれると仮定する。

これらの制約はあるものの、本データセットは画像・テキスト・数値が紐づいた大規模なマルチモーダルデータであり、かつ過去データであるため未来時点の正解ラベル（実際の反応量）が確定している。この点において、本研究で扱う予測モデルおよびXAIによる要因分析手法を定量的に評価するための検証環境として適していると判断し採用する。

### 3.2 データと予測タスクの結びつき

本研究では、インフルエンサーの成長予測において、マルチモーダル性とグラフ構造が将来の影響に与える影響を与えると仮定する。単一の静的指標（例：フォロワー数）だけに注目すると、購入フォロワーや不正アカウント等により指標が歪む可能性がある。そこで本研究は、投稿コンテンツ（画像・テキスト）や行動・関係性（ハッシュタグ、メンション、共起等）といった複合的な情報に基づき、翌月の影響力をどの程度予測できるかを検証する。

また、実務上は評価指標の仕様（可視化される反応指標など）が変化し得る。本研究の枠組みでは、将来予測の入力として過去のエンゲージメント値そのものを用いず、コンテンツおよび関係性に基づく情報から予測することを重視する。これにより、可観測な指標の変化があっても、コンテンツの力に基づく評価・予測へ接続しやすいことが期待される。

### 3.3 問題の定式化

本節では、インフルエンサーの影響力予測問題を数理的に定義する。問題を、時系列異種混合グラフデータを用いたランキング学習タスクとして定式化する。

#### 3.3.1 入力と出力の定義

時点  $t$ （例：月）における異種混合グラフを  $G_t = (V_t, E_t)$  とし、各ノードの特徴行列を  $X_t$  とする。ここで  $V_t$  はノード集合（インフルエンサー、一般ユーザー、ハッシュタグ、画像オブジェクト等）、 $E_t$  はエッジ集合（メンション、共起等）を表す。インフルエンサーノード集合を  $I_t \subseteq V_t$  とし、本研究のランキング対象は  $I_t$  に限定する。

観測窓長を  $T$  とし、予測対象月を  $t+1$  とする。予測モデル  $F$  は過去  $T$  期間の系列  $\{(G_{t-T+1}, X_{t-T+1}), \dots, (G_t, X_t)\}$  を入力として、各インフルエンサー  $u \in I_t$  に対する翌月の予測スコア  $\hat{s}_{u,t+1}$  を出力する：

$$\hat{s}_{u,t+1} = F((G_{t-T+1}, X_{t-T+1}), \dots, (G_t, X_t); \Theta),$$

ここで  $\Theta$  はモデルの学習可能パラメータである。

#### 3.3.2 影響力（エンゲージメント）指標

本研究では、時点  $t$  におけるインフルエンサー  $u$  の影響力（エンゲージメント）を、いいね数とコメント数の合計をフォロワー数で正規化した値として定義する：

$$y_{u,t} = \frac{\text{likes}_{u,t} + \text{comments}_{u,t}}{\text{followers}_{u,t}}.$$

この  $y_{u,t}$  は予測の正解ラベルとして用いる。一方で、入力特徴  $X_t$  には  $y_{u,t}$  を直接含めず、投稿コンテンツや関係性に由来する特徴に基づいて将来スコアを推定する。

### 3.3.3 損失関数による最適化 (ListMLE)

本研究の目的は、将来の影響力が高いインフルエンサーを相対順位として正確に推定することである。したがって回帰誤差ではなく、ランキング学習として損失関数を設計する。

学習では、対象月  $t$  を固定したときのインフルエンサー集合  $I_t$  から、長さ  $L$  のリスト (クエリ)  $Z_i$  をサンプリングする：

$$Z_i = \{u_{i1}, \dots, u_{iL}\} \subseteq I_t.$$

モデルは各  $u \in Z_i$  に対し翌月の予測スコア  $\hat{s}_{u,t+1}$  を出力する。以降、記号を簡略化して  $\hat{s}_{iu} = \hat{s}_{u,t+1}$  と書く。

翌月の正解ラベル  $y_{u,t+1}$  に基づき、リスト  $Z_i$  上の正解順序 (置換)  $\pi_i$  を

$$y_{i\pi_i(1)} \geq y_{i\pi_i(2)} \geq \dots \geq y_{i\pi_i(L)}$$

を満たすように定める (同値はランダムにタイブレイクする)。ここで  $y_{i\pi_i(k)}$  は  $Z_i$  内の要素  $u_{i\pi_i(k)}$  の翌月ラベル  $y_{u_{i\pi_i(k)},t+1}$  を表す。

ListMLE は、予測スコア  $\hat{s}_i = \{\hat{s}_{iu}\}_{u \in Z_i}$  に対して、Plackett-Luce モデルで正解順序  $\pi_i$  が生成される確率を

$$P(\pi_i | \hat{s}_i) = \prod_{k=1}^L \frac{\exp(\hat{s}_{i\pi_i(k)})}{\sum_{j=k}^L \exp(\hat{s}_{i\pi_i(j)})}$$

と定義し、その負の対数尤度を損失として最小化する：

$$\mathcal{L}_{\text{ListMLE}}(\Theta) = \frac{1}{m} \sum_{i=1}^m \left( -\log P(\pi_i | \hat{s}_i) \right) = \frac{1}{m} \sum_{i=1}^m \sum_{k=1}^L \left[ \log \left( \sum_{j=k}^L \exp(\hat{s}_{i\pi_i(j)}) \right) - \hat{s}_{i\pi_i(k)} \right].$$

ここで  $m$  は学習で用いるリストの総数である。本研究では、ランキング最適化を採用する先行研究に倣い list-wise learning-to-rank の枠組みで学習するが、0-1 損失は微分不可能であるため、実装可能な滑らかな代理損失として ListMLE を用いる。また、リスト  $Z_i$  は実装上ランダムサンプリングで構成する (例： $L = 10$ )。

## 第4章 先行手法

本章では、本研究の基盤となる2つの先行手法を整理する。第1に、将来有望なインフルエンサーを時系列異種混合グラフからランキングする予測モデル **InfluencerRank Kim2023InfluencerRank** を述べる。第2に、GNN の予測根拠をグラフ構造と特徴量の両面から説明する **GNNExplainer Ying2019GNNExplainer** を述べる。本研究は、InfluencerRank 型の直列モデル (GNN → RNN/Attention → MLP) に対して、GNNExplainer 系の最適化ベース説明を拡張適用する立場である。

### 4.1 インフルエンサー予測モデル: InfluencerRank

InfluencerRank **Kim2023InfluencerRank** は、インフルエンサーマーケティングにおける「将来伸長が見込まれるインフルエンサー」を発見することを目的に、投稿コンテンツと関係性を異種混合グラフとして表現し、その時系列変化を学習してランキングするモデルである。モデル全体像を図 4.1 に示す。

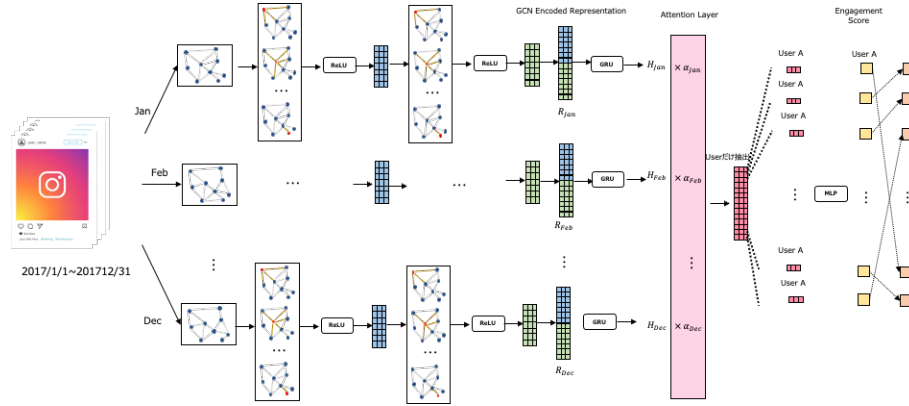


図 4.1: InfluencerRank のモデル概観 (異種混合グラフ → GNN → 時系列統合 → スコア推定)

#### 4.1.1 問題設定と影響力指標

InfluencerRank は、インフルエンサー  $u$  の時点  $t$  における影響力をエンゲージメント率として定義し、この将来値が高い順にインフルエンサーを順位付けする問題として定式化する **Kim2023InfluencerRank**。原論文では、時点  $t$  のエンゲージメント率を

$$E_u^t = \frac{\ell_u^t}{f_u^t}$$

とし、 $f_u^t$  はフォロワー数、 $\ell_u^t$  は当該期間における平均エンゲージメント (原論文では平均いいね数) である **Kim2023InfluencerRank**。一方、本研究では問題設定 (第3章) にて、いいね数とコメント数の双方を反映した指標を採用している。以降の説明では、原論文の設計思想 (エンゲージメント率を目的変数としてランキングする) を中心に述べ、指標の具体は第3章の定義に従う。

### 4.1.2 異種混合グラフによるマルチモーダル統合

SNS 上の投稿は、数値特徴だけでなく画像・テキストなどのコンテンツ情報を含み、さらにハッシュタグやメンションにより他者・概念と結びつく。InfluencerRank は、これら多様な情報を統合的に扱うため、異種混合グラフ (heterogeneous graph) としてモデリングする **Kim2023InfluencerRank**。一般に、異種混合グラフではノード集合  $V$  とエッジ集合  $E$  が、ノード種別・関係種別により区別される。

原論文では、インフルエンサー (Influencer)、画像オブジェクト (Image Object)、ハッシュタグ (Hashtag)、ユーザータグ (User Tag) などをノードとして扱い、インフルエンサーと各要素の共起や付与関係をエッジとして表現することで、投稿内容と行動・関係性を同一のグラフ上に埋め込む **Kim2023InfluencerRank**。図 4.2 は、投稿内容に基づくノード間接続の例を示す。

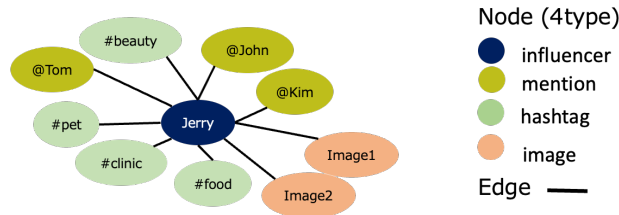


図 4.2: 投稿要素 (例: ハッシュタグ, ユーザータグ, 画像オブジェクト) に基づくエッジ形成の例

本研究においても、対象ノードはインフルエンサーノードであり (第 3 章), それ以外のノードはインフルエンサーのコンテンツや関係性を表す媒介として位置付ける。図 4.3 および図 4.4 は、インフルエンサーと他のノードの概念図を示す。

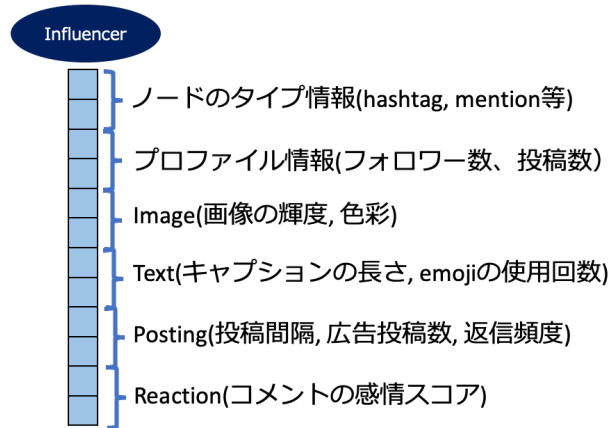


図 4.3: インフルエンサーノードの概念図 (コンテンツ・行動特徴と関係性を持つ)

### 4.1.3 空間表現学習: 各時点グラフに対する GNN 適用

時点  $t$  (例: 月) ごとに構築した異種混合グラフ  $G_t = (V_t, E_t)$  と特徴行列  $X_t$  に対し、GNN を適用してノード表現を得る。同種グラフに対する標準的 GCN の更新式は第 2 章で述べた通りであり、InfluencerRank では、各時点グラフを GNN でエンコードして得たノード表現を時系列モデルへ入力する **Kim2023InfluencerRank**。記号として、時点  $t$  の GNN 出力 (各ノードの埋め込み) を  $R_t$  と書く。



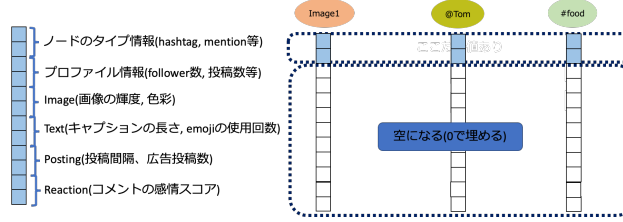


図 4.4: 非インフルエンサーノード（例：ハッシュタグ、ユーザータグ、画像オブジェクト）の概念図

#### 4.1.4 時系列統合: GRU と時間注意機構

InfluencerRank は、複数時点にわたるノード表現列  $\{R_{t-T+1}, \dots, R_t\}$  を再帰モデルで統合し、将来の影響力を推定する **Kim2023InfluencerRank**。原論文では、時系列統合にゲート付き再帰ユニット（GRU）を採用し、さらに時間注意機構により重要な時点を強調する構成を採る。直観的には、バズやコラボなどが生じた特定月が将来の影響力に強く影響するため、全時点を等価に扱うのではなく重み付けして集約する。

具体的には、各時点の隠れ状態を  $H_t$  とし、GRU により

$$H_t = \text{GRU}(R_t, H_{t-1})$$

で更新する。続いて時間注意機構により、各時点の重要度  $\alpha_t$  を推定し、

$$\alpha_t = \frac{\exp(\eta_t)}{\sum_{\tau} \exp(\eta_{\tau})}, \quad C = \sum_t \alpha_t H_t$$

のように重み付き和  $C$  を得る（ $\eta_t$  は  $H_t$  から計算されるスコア）**Kim2023InfluencerRank**。最後に、全結合層（MLP）により将来スコアを出力し、このスコアに基づいてインフルエンサーをランキングする。

なお、原論文では GRU と LSTM の比較も行われ、性能差が大きいことを報告している **Kim2023InfluencerRank**。本研究では、時系列統合モジュールの選択は実装上の設計として扱い、提案手法章で採用設定を明示する。

#### 4.1.5 ランキング学習としての最適化

InfluencerRank は、回帰ではなくランキングとして問題を捉え、list-wise learning-to-rank の枠組みで最適化する **Kim2023InfluencerRank**。原論文では、ランキング予測  $\hat{y}(Z_i)$  と正解順位  $y_i$  の不一致を表す 0-1 損失を導入し、

$$L_S(\hat{y}) = \frac{1}{m} \sum_{i=1}^m l(\hat{y}(Z_i), y_i), \quad l(\hat{y}(Z_i), y) = \begin{cases} 1, & \hat{y}(Z_i) \neq y \\ 0, & \hat{y}(Z_i) = y \end{cases}$$

の形でランキング学習を記述している **Kim2023InfluencerRank**。一方で 0-1 損失は微分不可能であるため、実装では滑らかな代理損失が必要となる。本研究では第 3 章で定義した ListMLE を学習損失として採用し、ランキング最適化を実装可能な形で行う。

## 4.2 GNN に対する説明可能 AI 手法: GNNExplainer

GNNExplainer **Ying2019GNNExplainer** は、GNN の予測を説明するための代表的手法である。GNN はグラフ構造データを扱うため、予測にはノード特徴量とエッジ構造の両方が寄与し得る。GNNExplainer は、特定ノードの予測を説明するため、その周囲のサブグラフと特徴量に注目し、説明に必要な要素をマスク最適化で抽出する。



### 4.2.1 説明対象: 構造マスクと特徴マスク

GNNExplainer は、以下の 2 種類のマスクを学習する **Ying2019GNNExplainer**。

- 構造マスク（エッジマスク） $m_E \in [0, 1]^{|E|}$ : 各エッジが予測に寄与する度合いを表す。値が 1 に近いほど重要と解釈する。
- 特徴マスク（特徴量マスク） $m_X \in [0, 1]^d$ : ノード特徴の各次元が予測に寄与する度合いを表す。値が 1 に近いほど重要と解釈する。

図 4.5 および図 4.6 は、それぞれエッジマスクと特徴マスクの概念を示す。

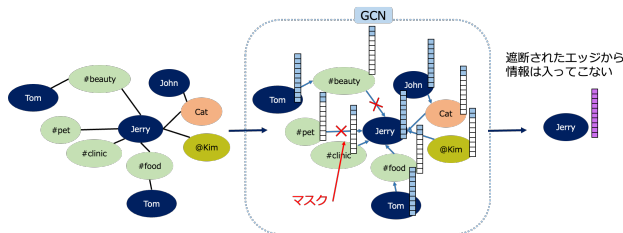


図 4.5: GNNExplainer におけるエッジマスクの概念図

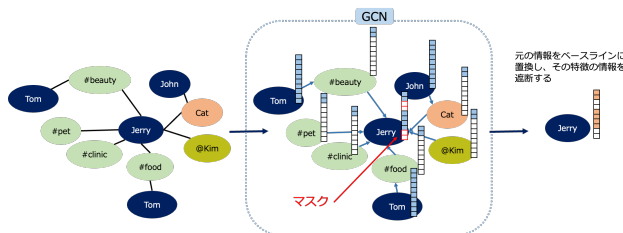


図 4.6: GNNExplainer における特徴マスクの概念図

### 4.2.2 マスク最適化

マスクを用いると、入力グラフ構造と特徴は要素ごとに抑制される。例えば隣接行列（またはエッジ重み行列）を  $A$  とすると、マスク後の構造を  $A \odot m_E$  と表せる。同様に特徴行列  $X$  に対しては  $X \odot m_X$  と表せる（ $\odot$  は要素ごとの積）。GNNExplainer は、元の予測をできるだけ保つ（忠実性）一方で、マスクを疎にする（簡潔性）ことを目的としてマスクを最適化する **Ying2019GNNExplainer**。

目的関数の典型形は、予測保持の損失  $\mathcal{L}_{\text{pred}}$  と疎性正則化を組み合わせた

$$\mathcal{L} = \mathcal{L}_{\text{pred}} + \lambda_E \|m_E\|_1 + \lambda_X \|m_X\|_1$$

である。ここで  $\lambda_E, \lambda_X$  は疎性の強さを制御するハイパーパラメータであり、勾配降下法により  $\mathcal{L}$  を最小化してマスクを得る。得られたマスク値は、どのエッジ・どの特徴が予測に必要なだったかを表す重要度として解釈される。

ただし、GNNExplainer が直接与えるのは主に重要度（マスクの大きさ）であり、どの要素がスコアを上げる方向に働いたか／下げる方向に働いたかという符号付き影響の解釈や、複数時点にまたがる比較可能な集約は自明ではない。本研究では、時系列グラフを入力とする直列モデルに対してマスク学習を適用し、特徴量とエッジの双方について符号付き影響度を定義して月次で比較可能にする点を提案手法として述べる（第 5 章）。

## 第5章 提案手法

### 5.1 End-to-End マスク最適化による時系列グラフ予測の説明可能 AI

本章では、時系列異種混合グラフに基づく予測モデルに対して、最適化ベースのマスク学習により説明を与える枠組みを提案する。提案法は、(i) 直列構造を持つ予測器（GNN → LSTM → Attention → MLP）を end-to-end に保ったまま、(ii) 予測をできるだけ維持しつつ入力（特徴量・エッジ）を疎に残すマスクを学習し、(iii) 得られたマスクの大きさ（重要度）に加えて符号付き影響度（スコアを上げる／下げる方向）を算出し、(iv) 月次で比較可能な形で要因を提示することを目的とする。

以降、説明対象はインフルエンサーノード  $v$  とし（第3章）、「どの月のどの特徴量・どのエッジが予測に寄与したか」を抽出する。

#### 5.1.1 問題設定と予測モデル

時刻（本研究では月） $t \in \{1, \dots, T\}$  におけるグラフを  $G_t = (V, E_t)$ 、ノード特徴を  $X_t \in \mathbb{R}^{|V| \times F}$  とする。対象ノード  $v \in V$  に対して、モデルは将来のスコア（影響力など） $\hat{y}_v$  を出力する。

本研究の予測器  $f_\Theta$  は、(i) 特徴射影、(ii) GCN による構造混合、(iii) LSTM による時系列混合、(iv) Attention による重み付け集約、(v) MLP によるスカラー出力からなる直列構造である：

$$\begin{aligned} P_t &= \phi(X_t) \in \mathbb{R}^{|V| \times d_p}, \\ H_t &= \text{GCN}_\theta(P_t, E_t) \in \mathbb{R}^{|V| \times d_g}, \\ s_{v,1:T} &= [H_1[v], \dots, H_T[v]] \in \mathbb{R}^{T \times d_g}, \\ \{h_{v,t}\}_{t=1}^T &= \text{LSTM}_\psi(s_{v,1:T}), \\ \alpha_{v,t} &= \text{softmax}_t(a^\top \tanh(W_a h_{v,t} + b_a)), \quad c_v = \sum_{t=1}^T \alpha_{v,t} h_{v,t}, \\ \hat{y}_v &= \text{softplus}(\text{MLP}_\omega(c_v)). \end{aligned}$$

ここで  $\Theta = \{\theta, \psi, \omega, \phi, a, W_a, b_a\}$  は学習可能パラメータである。以降、元の予測を

$$\hat{y}_v^{(0)} := f_\Theta(\{(G_t, X_t)\}_{t=1}^T; v)$$

と書く。

#### 5.1.2 説明対象の時刻選択（効く月だけ説明する）

全時刻に対して説明（マスク最適化）を実行すると計算コストが大きい。そこで本研究では、説明対象時刻  $S \subset \{1, \dots, T\}$  を事前に選別する。

対象ノード  $v$  の時系列埋め込み  $s_{v,1:T}$  に対して、「時刻  $t$  をドロップしたときの予測変化」を測る。具体的に、 $t$  の埋め込みのみをゼロ化した系列  $s'_{v,1:T}$  を

$$s'_{v,\tau} = \begin{cases} 0 & (\tau = t), \\ s_{v,\tau} & (\tau \neq t), \end{cases}$$

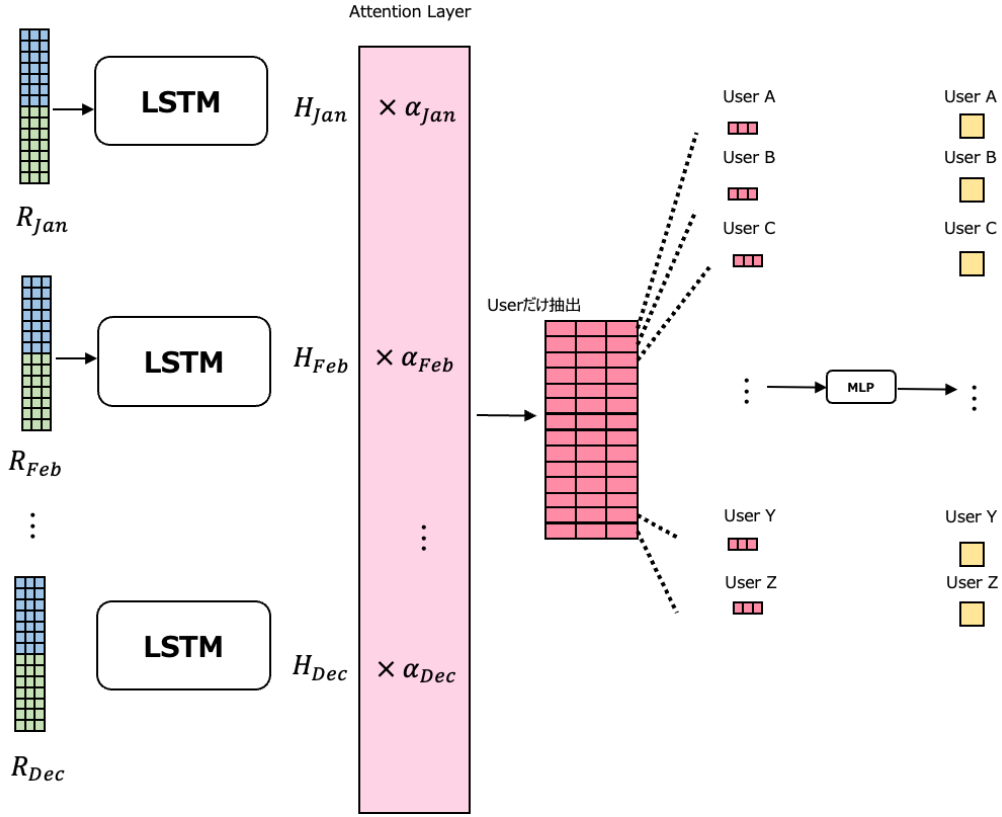


図 5.1: 本研究で扱う直列予測器の概観 (GNN → LSTM → Attention → MLP)

とし、それに基づく予測を

$$\hat{y}_{v,\backslash t} := f_{\Theta}(s'_{v,1:T}; v)$$

と定義する。そのときの変化量を

$$\Delta_t := |\hat{y}_v^{(0)} - \hat{y}_{v,\backslash t}|$$

とする。Attention 重み  $\alpha_{v,t}$  が得られる場合は、

$$\text{score}_t := \alpha_{v,t} \cdot \Delta_t$$

として上位  $K$  個を  $S$  とする ( $\alpha$  を用いない場合は  $\Delta_t$  のみで選択する)。

### 5.1.3 End-to-End Mask Optimization (MaskOpt)

選ばれた時刻  $\tau \in S$  ごとに、 $\tau$  の入力グラフおよび特徴にのみマスクを掛け、最終出力  $\hat{y}_v$  を end-to-end に維持するようにマスクを最適化する。本研究の立場では、教師ラベルを当てることよりも、元の予測  $\hat{y}_v^{(0)}$  をできるだけ保ちながら入力を簡潔化することを優先し、「予測に寄与する要素をコンパクトに抽出する」ことを説明の目的とする。

**局所説明のための部分グラフ化** 計算量を抑えるため、時刻  $\tau$  における  $k$ -hop 部分グラフ  $G_{\tau}^{(k)}$  を抽出する (ノード集合  $V_{\tau}^{(k)}$ )。さらにエッジマスクは対象ノード  $v$  に incident なエッジ集合  $\mathcal{I}_{\tau}$  のみに限定し、マスクパラメータ数を削減する。

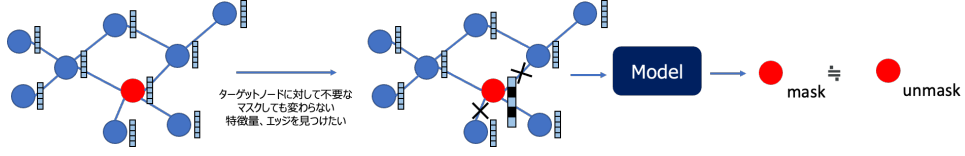


図 5.2: MaskOpt の概念図（予測を維持しつつ特徴量・エッジを疎に残すマスクを最適化する）

**マスクと適用** 特徴マスク  $m_x \in (0, 1)^F$ , エッジマスク  $m_e \in (0, 1)^{|\mathcal{I}_\tau|}$  を導入し, 時刻  $\tau$  にのみ適用する。

特徴に関しては, 対象ノード  $v$  の行のみをゲートする (target-only):

$$X'_\tau[u, :] = \begin{cases} X_\tau[v, :] \odot m_x & (u = v), \\ X_\tau[u, :] & (u \neq v), \end{cases}$$

ここで  $\odot$  は要素ごとの積を表す。

エッジに関しては, incident 辺のみをゲートし, それ以外は固定で 1 とする (incident-only):

$$w_\tau(e) = \begin{cases} m_e(e) & (e \in \mathcal{I}_\tau), \\ 1 & (e \notin \mathcal{I}_\tau). \end{cases}$$

GCN は重み付きメッセージパッシングとして実装されるとし,  $\tau$  以外の時刻はキャッシュした埋め込みを用いて再計算を省略する。

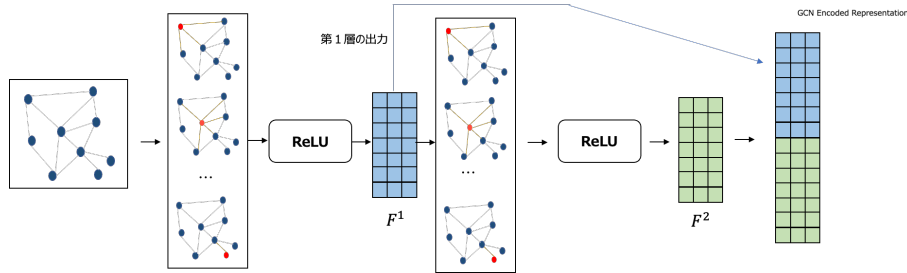


図 5.3: 時点グラフに対する GNN 表現の概念図（各層の出力を連結する実装例）

**最適化目的** ゲートはロジット  $l_x, l_e$  を用いて  $m_x = \sigma(l_x)$ ,  $m_e = \sigma(l_e)$  とパラメータ化し, Adam により最適化する。本研究では, (i) 予測の維持 (忠実性), (ii) マスクの疎性 (簡潔性), (iii) 離散化の促進 (0/1 に近づける), および (任意として) (iv) 補集合側が予測を十分変えることを促す項を組み合わせ, 次の損失を最小化する:

$$\begin{aligned} \mathcal{L}(m_x, m_e) = & \underbrace{\lambda_{\text{fid}} (\hat{y}_v(m_x, m_e) - \hat{y}_v^{(0)})^2}_{\mathcal{L}_{\text{fid}}} + \underbrace{\lambda_x \frac{1}{F} \|m_x\|_1}_{\mathcal{L}_{x, \text{size}}} + \underbrace{\lambda_e \frac{1}{|\mathcal{I}_\tau|} \|m_e\|_1}_{\mathcal{L}_{e, \text{size}}} \\ (5.1) \quad & + \underbrace{\beta_x \frac{1}{F} \sum_{j=1}^F H(m_{x,j})}_{\mathcal{L}_{x, \text{ent}}} + \underbrace{\beta_e \frac{1}{|\mathcal{I}_\tau|} \sum_{i=1}^{|\mathcal{I}_\tau|} H(m_{e,i})}_{\mathcal{L}_{e, \text{ent}}} + \underbrace{\lambda_{\text{ctr}} \max(0, \gamma - |\hat{y}_v(1 - m_x, 1 - m_e) - \hat{y}_v^{(0)}|)}_{\mathcal{L}_{\text{contrast}}}. \end{aligned}$$

ここで  $H(p) = -p \log p - (1-p) \log(1-p)$  は 2 値エントロピーである。 $\mathcal{L}_{\text{size}}$  はスパース性を促し、 $\mathcal{L}_{\text{ent}}$  はマスクの 0/1 化を促す。 $\mathcal{L}_{\text{contrast}}$  は補集合マスクが予測を十分変える ( $\geq \gamma$ ) ことを促すことで、自明解 (全マスクが 1 付近) を避ける目的で任意に導入する。

**出力 (重要度)** 最適化後のゲート  $m_x, m_e$  をそれぞれ特徴重要度・エッジ重要度として用いる。ただし  $m$  は選択の強さ (重要度) であり、スコアを増加させる／減少させる方向の情報 (符号) は持たない。そこで次節で、重要要素に対して符号付き影響度を別途推定する。

#### 5.1.4 符号付き影響度 (Score Impact) の算出

本研究では、最適化で得られた重要度 (マスク値) とは別に、one-at-a-time ablation により符号付き影響度を算出する。上位重要特徴 (または上位重要エッジ) について、1 つずつ置換・除去したときのスコア差を計測し、増減方向と大きさを与える。

**特徴の符号付き影響度** 時刻  $\tau$  の説明サンプルに対し、基準ベクトル  $b \in \mathbb{R}^F$  を定める。本研究では以下のいずれかを採用する (実装では切替可能) :

$$b = \begin{cases} \text{mean}_{u \in V} X_\tau[u, :] & (\text{full\_graph\_month}), \\ \text{mean}_{u \in V_\tau^{(k)}} X_\tau[u, :] & (\text{explain\_subgraph}), \\ 0 & (\text{target\_only}). \end{cases}$$

特徴  $j$  のアブレーション入力  $X_\tau^{(\setminus j)}$  を

$$X_\tau^{(\setminus j)}[v, j] = (1 - \rho)X_\tau[v, j] + \rho b_j$$

で作る ( $\rho \in (0, 1]$  は強さ、 $\rho = 1$  で完全置換)。そのときの符号付き影響度を

$$(5.2) \quad \text{Impact}_x(j) := \hat{y}_v^{(0)} - f_\Theta(\dots, X_\tau^{(\setminus j)}, \dots; v)$$

と定義する。 $\text{Impact}_x(j) > 0$  は「その特徴を潰すとスコアが下がる」ことを意味し、スコア増加に正寄与したと解釈する。

**エッジの符号付き影響度** 上位 incident 辺  $e = (v, u)$  について、時刻  $\tau$  の隣接を両方向まとめてドロップし、

$$(5.3) \quad \text{Impact}_e(e) := \hat{y}_v^{(0)} - f_\Theta(\dots, G_\tau \setminus \{(v, u), (u, v)\}, \dots; v)$$

で影響を測る。 $\text{Impact}_e(e) > 0$  は、当該接続がスコアを押し上げたことを意味する。

**Zero 判定 (数値安定化)** 数値丸めにより 0 に見える問題を避けるため、絶対閾値  $\epsilon_{\text{abs}}$  と相対閾値  $\epsilon_{\text{rel}}$  を用い、

$$|\text{Impact}| \leq \max(\epsilon_{\text{abs}}, \epsilon_{\text{rel}} |\hat{y}_v^{(0)}|)$$

なら Zero と判定する (実装では特徴・エッジで別の  $\epsilon$  を指定可能)。

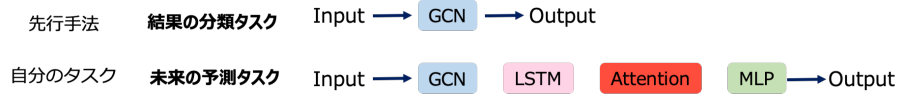


図 5.4: GNNExplainer 系の説明と、本研究の MaskOpt の位置付け（直列予測器に対する end-to-end 説明）

### 5.1.5 GNNExplainer との比較

本節では、先行の GNNExplainer **Ying2019GNNExplainer** と、本研究の MaskOpt の相違点を整理する。比較図を図 5.4 に示す。

GNNExplainer は、説明対象インスタンスごとにサブグラフ（および特徴）を選び、予測と説明の相互情報量を最大化する形で定式化される：

$$\max_{G_S, X_S} I(Y; (G_S, X_S)).$$

実装上は、連続マスク  $M$  を学習し、(タスクに応じて) 負の対数尤度 (分類なら交差エントロピー等) にスパース・エントロピー正則化を加えて最適化する形となる **Ying2019GNNExplainer**：

$$\min_M \mathcal{L}_{\text{task}}(f_{\Theta}(G \odot M, X); y) + \lambda \|M\|_1 + \beta \sum_i H(M_i).$$

これに対し本研究の MaskOpt は、(i) 直列の時系列モデル (GCN→LSTM→Attention→MLP) を end-to-end に保ったまま、(ii) ラベルを当てるのではなく元の予測  $\hat{y}_v^{(0)}$  を維持する（忠実性）ことを主目的として式 (5.1) を最適化する点が異なる。また (iii) 時刻  $\tau$  のみにマスクを掛け、他時刻はキャッシュを活用することで、「どの月の情報が効いたか」を月別に分解しやすい。さらに (iv) 重要度（ゲート値）とは別に、式 (5.2), 式 (5.3) により符号付き影響度（増減方向）を推定し、特徴量とエッジの双方について月次比較を可能にする点が本研究の狙いである。

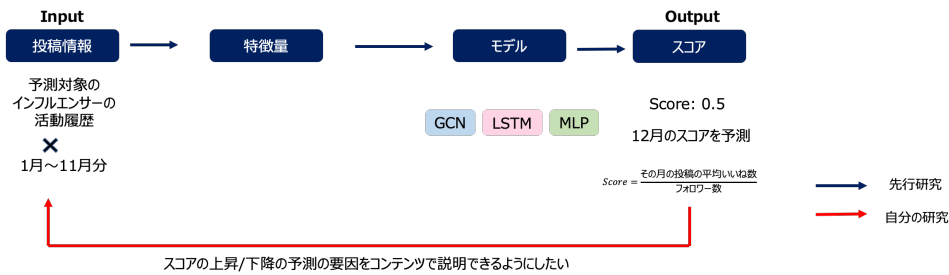


図 5.5: 本研究の位置付け（予測モデルと説明手法の統合、および月次比較可能な寄与分析）

## 第6章 実験

### 6.1 実験の目的

本研究では、学習済みの時系列グラフ予測モデルに対して、最適化ベースのマスク学習により得られる説明が妥当な振る舞いをするかを検証する。ここで妥当性とは、重要とされた要素（特徴量・エッジ）を操作すると予測が想定通りに変化すること、および説明が少数要素にまとまり月次で比較可能な形で提示できることを指す。

本章の目的は以下の2点である。

1. 予測性能の定量評価（学習済みモデルの確認）：学習済みモデルが、将来月の影響力ランキングをどの程度統合的に出力できているかを、ランキング指標（NDCG, RBP等）で評価する。
2. 説明（貢献度）の妥当性検証（学習なしで可能な検証）：マスク最適化で得られた重要度（importance）と、アブレーションで得る符号付き影響度（score impact）が、予測スコアの変化と整合するかを検証する。

なお、本章では再学習を行わない。すなわち、学習済みモデル  $f_\theta$  を固定し、推論・説明計算・入力操作（アブレーション）により評価を行う。

### 6.2 実験設定

#### 6.2.1 使用データセットと対象ノード

Instagram データセットを用い、各月  $t$  ごとに異種混合グラフ  $G_t = (V, E_t)$  とノード特徴  $X_t$  を構築する。ランキング対象はインフルエンサーノードに限定する（第3章の問題設定に従う）。

影響力（正解指標）は、対象インフルエンサー  $u$  の月  $t$  におけるエンゲージメント率として

$$y_u^t = \frac{\text{likes}_u^t + \text{comments}_u^t}{\text{followers}_u^t}$$

を用いる（実装では月内集計値から算出する）。

#### 6.2.2 学習済みモデル（固定）

学習済みモデル  $f_\theta$  は、月次グラフに対する GNN と、時系列統合 (LSTM) および Attention を直列に接続し、最終的にスコアを出力する。本章では、既に得られている学習済み重み  $\theta$  を用い、追加の学習は行わない。

#### 6.2.3 説明（貢献度）算出の定義

対象インフルエンサー  $u$  の予測スコアを

$$\hat{y}(u) = f_\theta(\{(G_t, X_t)\}_{t=1}^T; u)$$

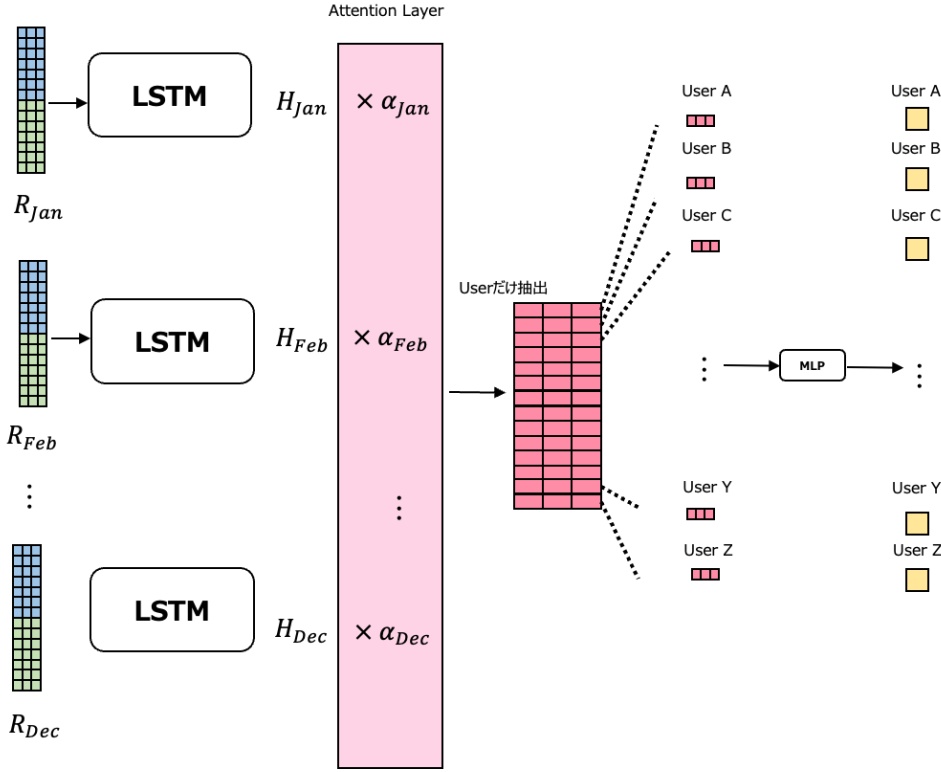


図 6.1: 実験で用いる学習済み予測器の概観 (GNN → LSTM → Attention → MLP)

とする。提案手法により，説明対象月  $\tau$  に対して特徴量マスク  $m^X$  とエッジマスク  $m^E$  を学習し，次を算出する：

- importance (重要度)：マスク値そのもの（例： $m_j^X$ ,  $m_e^E$ ）を重要度として用いる．値が大きいほど，予測を維持するために残したい度合いが高いと解釈する．
- score impact (符号付き影響度)：ある要素をベースライン置換（特徴）または除去（エッジ）したときのスコア差で定義する．特徴量  $j$  について，

$$(6.1) \quad \Delta_j^X(u; \tau) = \hat{y}(u) - f_{\Theta}(\dots, X_{\tau}^{(j \leftarrow b)}, \dots; u),$$

とする．ここで  $X_{\tau}^{(j \leftarrow b)}$  は，月  $\tau$  の対象ノードの特徴  $j$  をベースライン  $b_j$  に置換した入力である．エッジ  $e$  についても同様に，月  $\tau$  のグラフから  $e$  を除去した入力を用いて  $\Delta_e^E(u; \tau)$  を定義する．

重要度は選択強度，影響度は実際にスコアがどれだけ動くか（符号付き）であり，両者がどの程度整合するかを検証する．また，月を跨いだ比較では重要度（mask 値）の絶対値比較は慎重に扱い，主に影響度（スコア差）に基づいて月次比較を行う．

### 6.3 説明の妥当性検証：学習なしで可能な検証

本節では，学習済みモデルを固定したまま，説明（importance / score impact）が妥当な振る舞いをするかを検証する．検証は，実験 1（マスク値の確認）と実験 2（重要度と実影響の整合性チェック）からなる．



## 6.4 実験 1：マスク値の確認（1月，2月，3月）

### 6.4.1 目的

マスク最適化で得られるマスク値が，説明として利用可能な分布とスケールを持つかを確認する．具体的には，(i) 上位要素が少数に集中しているか，(ii) 多数の要素が 0 付近に張り付く（飽和する）傾向がないか，(iii) 月ごとに分布がどの程度変動するかを観察し，以降の整合性評価の前提を明確にする．

### 6.4.2 手順

1. 対象インフルエンサー  $u$  を選択する（例：将来月の予測ランキング上位）．
2. 説明対象月  $\tau$  を 1 月，2 月，3 月に設定し，各  $\tau$  に対して特徴量マスク  $m^X$  とエッジマスク  $m^E$  を得る．
3. 各マスクを降順ソートし，top 要素を抽出して可視化する（ノード特徴量とエッジを別々に表示する）．

### 6.4.3 結果

図 6.2～図 6.4 に，1 月，2 月，3 月それぞれに対する重要度（importance）の可視化結果を示す．これらの図は，マスク最適化により得られた重要度を上位からプロットし，値の分布がどの程度広がるかを確認する目的で作成したものである．

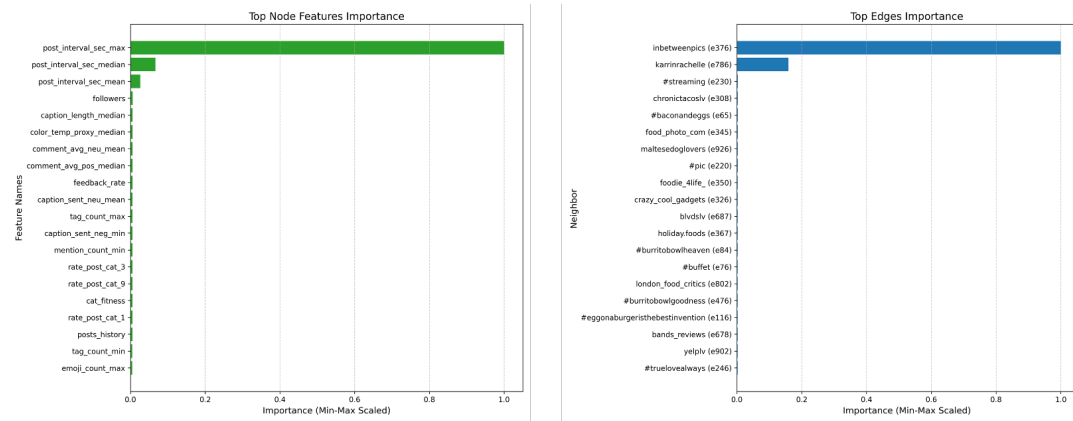


図 6.2: 実験 1 の結果（1 月）：ノード特徴量とエッジの importance の上位プロット

## 6.5 実験 2：重要度と実影響の整合性チェック（1月，2月，3月）

### 6.5.1 目的

重要度（importance）が高い要素ほど，その要素を操作したときのスコア変化（score impact）が大きい傾向にあるかを検証する．これは「説明が予測に忠実である」ことの，学習不要な検証（入力介入による検証）である．

### 6.5.2 手順

1. 実験 1 と同一の対象インフルエンサー  $u$  を用い，説明対象月  $\tau$  を 1 月，2 月，3 月に設定する．

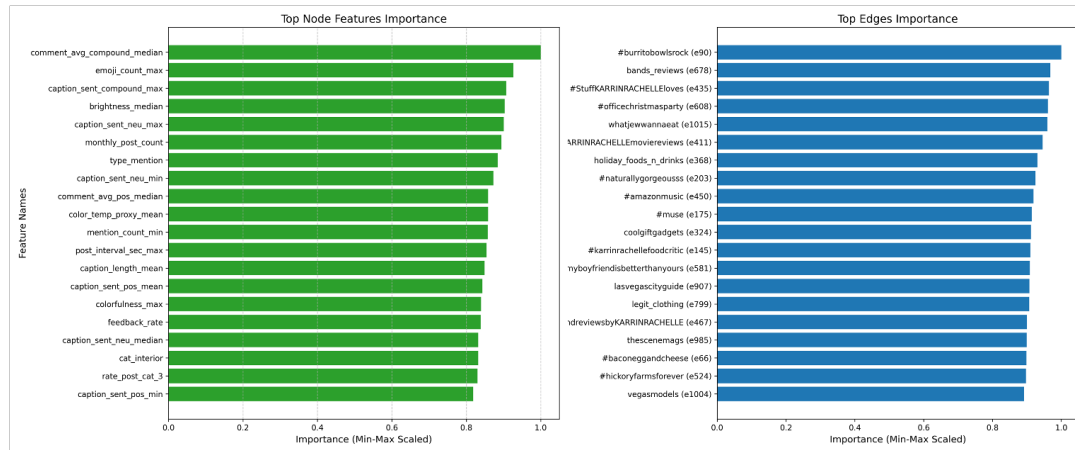


図 6.3: 実験 1 の結果 (2 月): ノード特徴量とエッジの importance の上位プロット

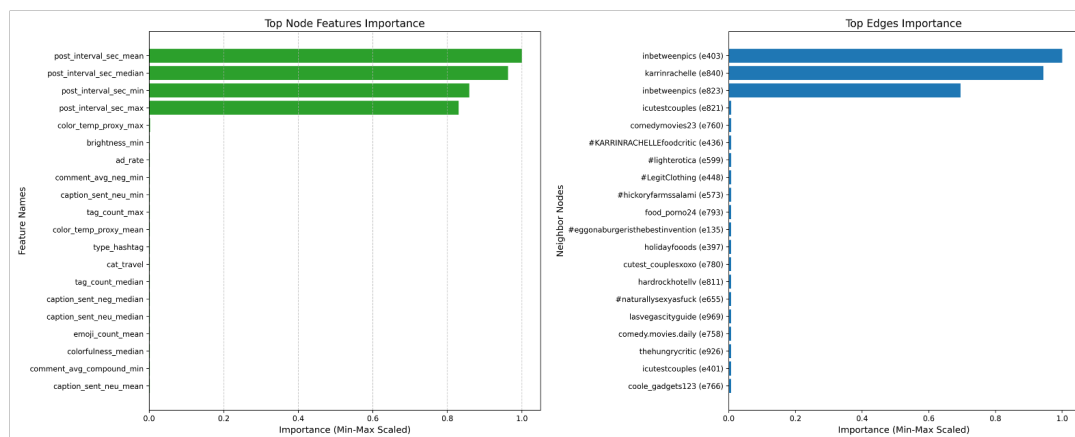


図 6.4: 実験 1 の結果 (3 月): ノード特徴量とエッジの importance の上位プロット

2. 各  $\tau$  に対して、特徴量とエッジの importance を得る.
3. 特徴量はベースライン置換, エッジは除去を行い, score impact (符号付き影響度) を算出する (式 (6.1)).
4. importance と score impact の関係を可視化し, 整合性を確認する.

### 6.5.3 結果

図 6.5～図 6.7 に, 1 月, 2 月, 3 月それぞれにおける importance と score impact の関係を示す. 本図は, 重要度が高いとされた要素が, 実際にスコアに影響する要素として振る舞っているかを確認する目的で作成したものである.

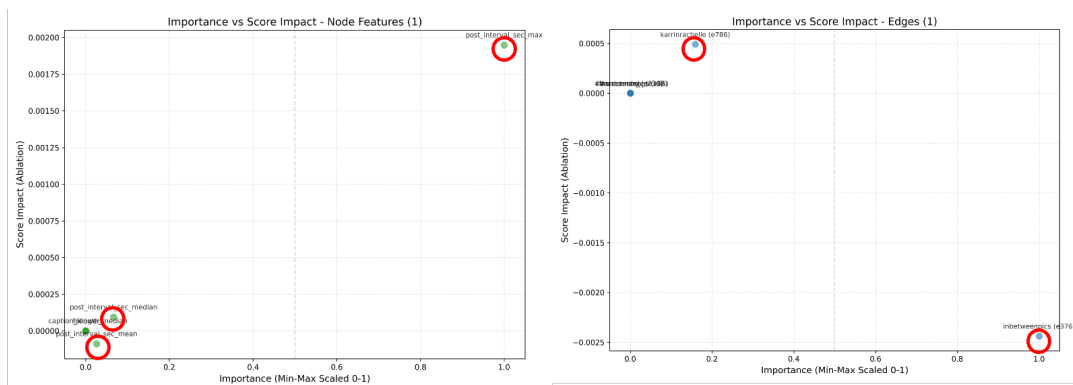


図 6.5: 実験 2 の結果 (1 月): importance と score impact の整合性の可視化

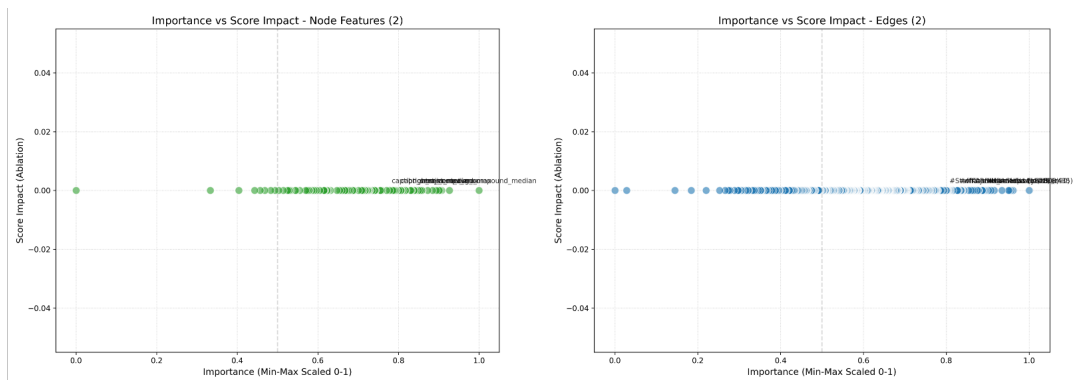


図 6.6: 実験 2 の結果 (2 月): importance と score impact の整合性の可視化

## 6.6 まとめ

本章では, 学習済みモデルを固定したまま実行可能な評価として, (1) ランキング指標による予測性能の確認, (2) 実験 1 としてマスク値 (importance) の分布確認 (1 月, 2 月, 3 月), (3) 実験 2 として重要度と実影響 (score impact) の整合性チェック (1 月, 2 月, 3 月), を実施し, 説明が利用可能かどうかを検証するための結果を整理した. 次章では, これらの結果を踏まえて考察と限界, および今後の展望を述べる.

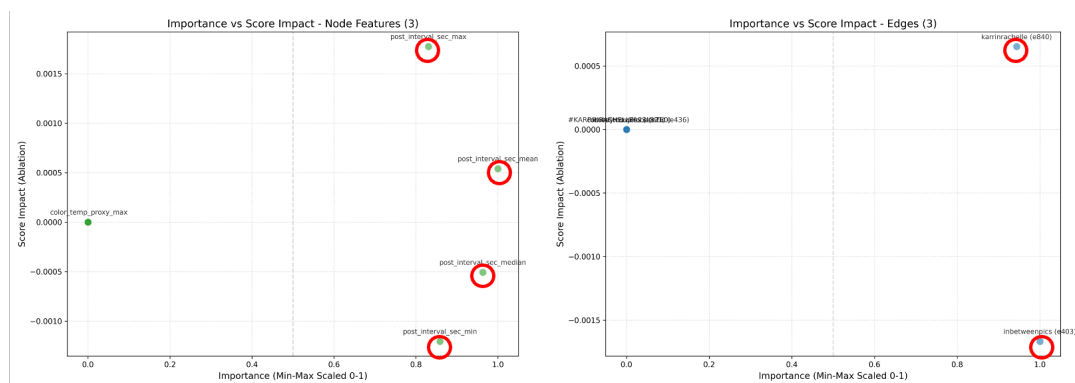


図 6.7: 実験 2 の結果 (3 月): importance と score impact の整合性の可視化

## 第7章 考察

本研究では、InfluencerRank 型の時系列グラフ予測モデルの追試を行い、学習済み予測器を固定したまま最適化ベースのマスク学習 (GNNEExplainer 系) を適用することで、予測根拠 (特徴量およびエッジ) を抽出し、その妥当性を検証した。本章では、実験結果から得られた知見を整理し、本手法の限界と今後の展望を述べる。

### 7.1 貢献度算出に関する考察

#### 7.1.1 マスク値の分布と探索範囲削減としての有用性

実験1では、月ごとに得られる特徴量マスク  $m^X$  とエッジマスク  $m^E$  を可視化し、マスク値の分布が説明として利用可能かを確認した。1月および3月の結果では、上位要素 (特徴量・エッジ) が相対的に上位へ集中し、top- $k$  や閾値  $\tau$  により候補要素を限定できる傾向が観測された。このことは、特徴量次元やエッジ数が大きい設定において、説明を重要候補の絞り込みとして用いる際に有用であることを示唆する。

一方で、1月および3月では、スコアへ影響を与える上位要素は検出できたものの、それ以外の多数要素が0付近に張り付く傾向が強く観測された。この現象は、最小構造を追求する損失の重み (疎性正則化) が強すぎる場合や、探索空間が大きい状況での最適化の不安定性により、説明が過度に疎になっている可能性を示す。また2月の結果では、説明自体が安定して学習できていない可能性があり、月によって最適化が成功しない条件が存在することが示唆された。

以上より、本研究の設定では、マスク値を精密な連続量として解釈するよりも、条件を固定した上での相対比較 (同一設定内での順位, top- $k$  の一致, 閾値後の残存数) に寄せて用いるのが安全である。特に月次比較では、重要度 (mask 値) の絶対値を直接比較するのではなく、月内の順位や、後述のスコア差分 (影響度) に基づく比較を主とする必要がある。

#### 7.1.2 重要度 (mask 値) と符号付き影響度 (score impact) の整合性

実験2では、重要度に基づく順位と、実際に要素を置換・除去したときのスコア変化 (score impact) に基づく順位の整合性を調べた。説明が妥当に機能していれば、重要度が高い要素ほど  $|\Delta|$  が大きい傾向が期待される。この傾向が観測される場合、マスク最適化が予測を維持するために必要な要素を優先して残していることと整合する。

ただし、重要度と影響度が一致しないケースも生じうる。主な要因として以下が考えられる。

- ・ 関連した特徴量が多数存在する場合、どれか一つを残せば予測が保持されるため、重要度が分散または入れ替わる。
- ・ マスク最適化は忠実性と疎性のトレードオフであり、影響度が最大の要素を必ずしも選ぶ最適化ではない。
- ・ 局所解・初期値依存により、似た性能の別マスクへ収束する (多解性)。

本研究の結果においても、月によっては重要度が0付近に集中しすぎるなど、重要度だけでは説明の信頼性を担保しにくい状況が示唆された。したがって、重要度のみで結論を出すのではなく、スコア差分に基づく検証 (重要度上位の要素を操作したときに予測がどの程度変化するか) を併用し、さらに複数ユーザー・複数月で再現性を確認する必要がある。

## 7.2 手法の限界

### 7.2.1 最適化規模とサンプル数の制約

本研究では、説明のための最適化においてサンプル数が 30000 程度である一方、最適化対象となるパラメータ（マスクの自由度）が 150000 程度に達する場合があった。パラメータ数がサンプル数に比べて大きい状況では、最適化が不安定になったり、多解性が強くなったりして、得られたマスクの数値としての信頼性が低下する可能性がある。その結果、探索範囲を絞る目的には有効でも、マスク値を精密な定量量として扱うことには注意が必要である。特に月によって最適化が失敗する可能性（2 月の挙動）は、この制約が現実的に影響した可能性がある。

### 7.2.2 検証範囲の限定（単一ノード・少数時刻）

本研究で詳細分析したのは、単一ノード（あるいは少数ノード）と少数の月に対する説明が中心であり、一般性（他ノード・他期間・他条件でも同様に成立するか）は十分に検証できていない。グラフ構造や投稿特性が異なるユーザー群では、マスクの疎性や安定性、重要度と影響度の整合性が変化しうる。よって、現時点の結論は限定された条件下での観察結果であることを明記する必要がある。

### 7.2.3 相互作用（交互作用）を十分に扱えていない

本研究では、一つの特徴量／エッジを置換・除去したときのスコア差分として影響度を定義した。しかし、実際には複数要素の組合せによる相互作用（例：投稿頻度と特定関係性の同時変化）がスコアに影響する可能性がある。単一要素の差分だけでは、相互作用に起因する要因を十分に捉えられない場合がある。また、相互作用が強い状況では、重要度が分散しやすく、説明の再現性が低下する可能性がある。

## 7.3 今後の展望

### 7.3.1 重要度順の追加・削除による反実仮想検証

今後は、重要度順に要素を段階的に削除（または追加）し、予測スコアがどのように変化するかを評価する実験を行う。これにより、top- $k$  の要素が予測保持にどの程度寄与しているかを、単発の置換よりも安定に検証できる。具体的には  $k = 1, 2, \dots$  に対して、上位  $k$  を削除した場合の予測低下と、同数をランダムに削除した場合の予測低下を比較し、説明がランダムよりも有効であることを確認する。この評価は、0 付近への過度な集中が観測された場合にも、上位要素が本当に意味を持つかを直接検証できる。

### 7.3.2 条件が近いユーザー群での再現性の確認

説明の妥当性を妥当性として主張するためには、同様の条件を持つ複数ユーザーに対して、類似の説明パターンが再現されることの確認が重要である。例えば、投稿頻度・フォローワー規模・カテゴリなどが近いユーザー群を抽出し、同一設定で説明を生成して、top- $k$  の一致率や重要度と影響度の相関係数の分布を比較する。これにより、説明の一般性と実務的な再利用可能性（パターン化）を評価できる。

### 7.3.3 相互作用を考慮した貢献度の拡張

本研究では単一要素の差分に基づく影響度を扱ったが、今後は相互作用を捉える貢献度の設計が課題となる。例えば、複数要素を同時に操作したときのスコア変化を用いる、特徴量

群・エッジ種別単位で集約した上で影響度を定義する，または相互作用を評価するための反実仮想設計を導入することで，単独では弱いが組合せで効く要因の抽出が可能になると考えられる．

#### 7.3.4 説明の安定化と評価指標の整備

最適化に基づく説明は初期値依存・多解性が避けられないため，複数初期値の平均・アンサンブルや，得られた解の分散を併記するなど，安定性を前提とした報告形式が必要である．また，説明の評価指標として，順位相関に加えて top- $k$  一致率，閾値に対する残存要素数（疎性），重要度順の削除曲線（ランダム比較），fidelity と疎性のトレードオフ曲線など，説明の目的（絞り込みか定量解釈か）に対応した指標の整備が重要である．

## 第8章 結論

本研究では、将来有望なインフルエンサーを早期に発掘する青田買いタスクを対象に、マルチモーダルな情報（ノード特徴量とグラフ構造）を入力とする時系列グラフ予測モデルにより、将来月のエンゲージメント（影響力）を予測する枠組みを検討した。さらに、学習済み予測器を固定したまま最適化ベースのマスク学習（GNNExplainer 系）を適用し、予測根拠を特徴量およびエッジの観点から抽出・検証する手順を提示した。

実験の結果、マスク最適化により得られる重要度（importance）を可視化することで、月ごとに上位へ集中する要素が現れ、候補要素を  $\text{top-}k$  や閾値で限定できる場合があることを確認した。このことは、説明を重要候補の絞り込みとして用いる観点で有用であり、高次元な特徴量や多数のエッジを含む設定において、人手分析の探索範囲を削減できる可能性を示す。一方で、月によってはマスク値が0付近に集中しすぎる傾向や、最適化が安定しない可能性も観測され、マスク値を精密な定量量として解釈するには注意が必要であることが示唆された。そこで本研究では、重要度のみで結論を出すのではなく、要素の置換・除去によるスコア差分（score impact）を併用し、重要度と実際の影響の整合性を確認することの必要性を整理した。

以上より、本研究の貢献は、インフルエンサー予測モデルに対して、学習不要な検証（入力介入）に基づく説明の評価手順を与え、予測結果に対する説明を候補絞り込みと整合性チェックの形で提示可能であることを示した点にある。これにより、意思決定者は単なるランキング出力だけでなく、どの要素が予測に関連しているかを参照しつつ分析を進めるための足場を得られる。

今後の課題としては、第一に、説明の安定性と一般性の検証が挙げられる。複数ユーザー・複数月への適用を拡大し、初期値や正則化条件に対する頑健性、 $\text{top-}k$  の一致率や削除曲線などの評価指標に基づき、再現性を確認する必要がある。第二に、単一要素の差分では扱いにくい相互作用（交互作用）を考慮した貢献度設計が課題である。第三に、グラフ構造由来の要因をより信頼できる形で扱うため、エッジ（関係性）に対する影響度の検証手順を拡張し、ノード特徴量とネットワーク構造の双方から成長メカニズムを説明できる枠組みへ発展させる必要がある。



## 第9章 謝辞

本論文の作成にあたり、村田 昇教授には、ゼミのみならずお忙しい合間を縫って、研究についての助言をいただきました。特に、発表に関しては、自分の研究に興味を持って聴いてもらうための術など多くのことをご指導いただきました。心より感謝いたします。研究室の先輩方には、ゼミ内での議論や実験環境の構築など様々な場面でアドバイスをいただきました。

また、研究室の同期の皆様には、研究を進める上で精神的なサポートをしていただきました。お互い切磋琢磨しながら研究を進めることができたことに感謝いたします。

最後に、これまでのご指導やご支援に心より感謝申し上げます。