

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

2026 年 1 月 18 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

修士論文概要書

Summary of Master's Thesis

Date of submission: 02/29/2026

専攻名(専門分野) Department	電気・情報生命	氏名 Name	渡邊 遼真	指導教員 Advisor	村田 昇
研究指導名 Research guidance	情報学習システム	学籍番号 Student ID number	5324E110-4		
研究題目 Title	マルチモーダルな GNN における説明可能な影響力予測				

研究背景

近年、スマートフォンおよびSNSの普及によりデジタルマーケティング市場は拡大しており、特にインフルエンサーマーケティングは企業の認知獲得や購買行動促進において重要な施策となっている。しかし、著名なトップインフルエンサーは起用コストが高騰し、競合との獲得競争も激しいため、早期に有望な人材を発掘し中長期で育成・契約する「青田買い」が実務上の鍵となる。

この意思決定を支えるため、SNS上の行動ログを用いて将来の影響力(例:エンゲージメント)を予測し、候補者をランクインする研究が進んでいる。一方で、多くの深層学習モデルは高精度でも予測プロセスがブラックボックスになりやすく、「なぜその人物が伸びると予測されたのか」という根拠が得られない。実務では、単に予測スコアが高いだけでは施策に結びつかず、投稿頻度・内容・関係性など、具体的に何が成長のドライバとなっているのかを説明可能な形で提示することが求められる。

そこで本研究は、時系列かつ異種混合(マルチモーダル)データに基づく影響力予測モデルに対して、予測と整合する説明を付与する枠組みを構築し、その妥当性と有用性を検証することを目的とする。

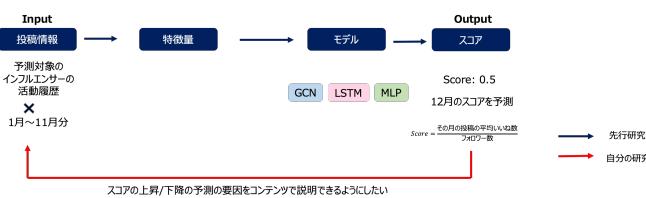


図1: 本研究の位置付け(予測モデルと説明手法の統合、および月次比較可能な寄与分析)

問題設定

本研究では、Instagramの投稿データ(画像・キャプション・ハッシュタグ等)とユーザー間の関係性(例:メンションやコメントの共起)を用い、月次の異種混合グラフ列を構築する。各時刻 t におけるグラフを

$$G_t = (V_t, E_t, \phi), \quad X_t \in \mathbb{R}^{|V_t| \times d}$$

とする。ここで V_t はノード集合(ユーザー、タグ、オブジェクト等)、 E_t はエッジ集合(例:ユーザー-タグ、ユーザー-メンションなど)、 ϕ はノード/エッジのタイプを表す。 X_t はノード特徴行列であり、数値特徴(投

稿数・投稿間隔・キャプションの長さ etc)に加え、画像・テキストから抽出した特徴(例:画像輝度やテキスト長)など、マルチモーダル特徴を含む。

タスクは、過去 T 期間における入力 $\{(G_{t-T+1}, X_{t-T+1}), \dots, (G_t, X_t)\}$ から、次期における各ユーザー u の影響力スコア \hat{y}_u を推定し、上位ユーザーをランクインすることである。本研究の目的は値の回帰精度よりも「有望ユーザーの上位抽出」であるため、学習は Learning to Rankとして定式化し、順位の整合性を直接最適化する。

なお、本データセットは完全なフォロー関係を含まないため、関係性はメンション等の代替指標により近似する。したがって、構築されるネットワークは「眞の社会関係」を完全に反映するものではなく、観測可能な相互作用に基づく近似である。本研究ではこの制約を明示した上で、予測と説明の枠組みの検証を行う。

提案手法

本研究は、(1) 時系列・異種混合グラフに基づく影響力予測(InfluencerRank型)と、(2) 最適化ベースのマスク学習による説明(MaskOpt)を統合する枠組みを提案する。目的は、単に「誰が伸びるか」を当てるだけでなく、「なぜ伸びると判断したのか」を特徴量および関係性(エッジ)レベルで提示し、月次比較可能な形に整えることである。

(1) 予測モデル: InfluencerRank型 各月の異種混合グラフ G_t に対してGNN(例:GCN)を適用し、ユーザー u の表現 $h_{u,t}$ を得る。

$$h_{u,t} = \text{GNN}(G_t, X_t; \theta_{\text{gnn}})$$

次に、過去 T 期間の表現列 $\{h_{u,t-T+1}, \dots, h_{u,t}\}$ をRNN(LSTM/GRU)で時系列統合し、隠れ状態 $s_{u,t}$ を更新する。

$$s_{u,\tau} = \text{RNN}(h_{u,\tau}, s_{u,\tau-1}; \theta_{\text{rnn}})$$

さらにTemporal Attentionにより各時刻の重要度 $\alpha_{u,\tau}$ を推定し、コンテキスト表現 $c_u = \sum_{\tau} \alpha_{u,\tau} s_{u,\tau}$ を得る。最後にMLPで将来スコア \hat{y}_u を出力し、このスコアでランクインを作る。この構成により、「どの月の状態が予測に効いたか」を注意重みとして解釈でき、後述の説明手法とも自然に接続できる。

(2) 説明手法: End-to-End マスク最適化 (MaskOpt) 予測モデルを固定した上で、説明対象ユーザー u と説

手法	NDCG@50	RBP(0.95)	Pearson
InfluencerRank	0.89	0.043	0.48
UA(投稿頻度)	0.49	0.024	0.09
DeepInf	0.56	0.031	0.23

明対象時刻（または時刻集合）を定め、入力要素に連続値マスクを導入する。具体的には、(i) ノード特徴量に対するマスク m^{feat} , (ii) 近傍エッジに対するマスク m^{edge} を学習し、マスク後の入力でも予測が大きく変化しない（忠実性）一方で、マスクが少数要素に集中する（簡潔性）ように最適化する。概念的には以下の目的関数を最小化する：

$$\mathcal{L} = \underbrace{\mathcal{L}_{\text{fidelity}}(\hat{y}_u^{\text{masked}}, \hat{y}_u^{\text{orig}})}_{\text{予測保持}} + \lambda_1 \|m^{\text{feat}}\|_1 + \lambda_2 \|m^{\text{edge}}\|_1$$

最適化後のマスク値を重要度 (importance) として用いる。重要度は「どの要素が選ばれたか」を表すが、符号（スコアを増やす/減らす方向）を含まないため、重要とされた要素を基準値へ置換したときの予測スコア差分を符号付き影響度 (score impact) として算出する。これにより、「何が重要か (importance)」と「どう効くか (score impact)」を分離して提示できる。なお、score impact は局所的な反実仮想量であり、因果効果の断定は行わない。

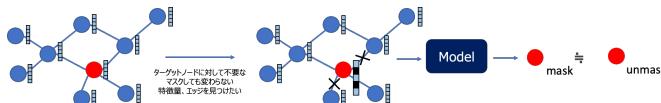


図 2: MaskOpt の概念図（予測を維持しつつ特徴量・エッジを疎に残すマスクを最適化）

解析例

ツールとしての信頼性を担保すべく、大きく 3 点の観点から解析を行った。

(1) 予測精度 (定量) 解析するにあたって、Instagram データセット（約 10 万ユーザー、2017 年 1 月～12 月）を用い、過去 11 ヶ月を学習用、12 月をテスト用とした。評価指標は NDCG@50 および RBP(0.95) を採用し、単純ベースライン（フォロワー数、投稿頻度）に基づく順位付け）と比較した。表 1 に主要指標を示す。InfluencerRank 型モデルは単純ベースラインを上回り、将来有望ユーザーの上位抽出に有効であることを確認した。また、図 3 に予測スコアのヒートマップ例を示す。

相関分析においても、予測値と真の値が、ずれてい る部分はあれど、一定の精度で予測できていることが 図からわかる。

この結果から、予測モデルが妥当なランキング性能を示すことを確認した。フォロワー数、投稿頻度等の基本的な順位づけに加えて、他の影響度を予測する手法と比較しても高い性能を示すことから、インフルエンサー選定において有効なツールとなり得ることが示唆された。特に、過去のエンゲージメントの値を使用せずにコンテンツの特徴量や関係性に基づいて将来の影響力を予測できている点が重要である。Pearson 相

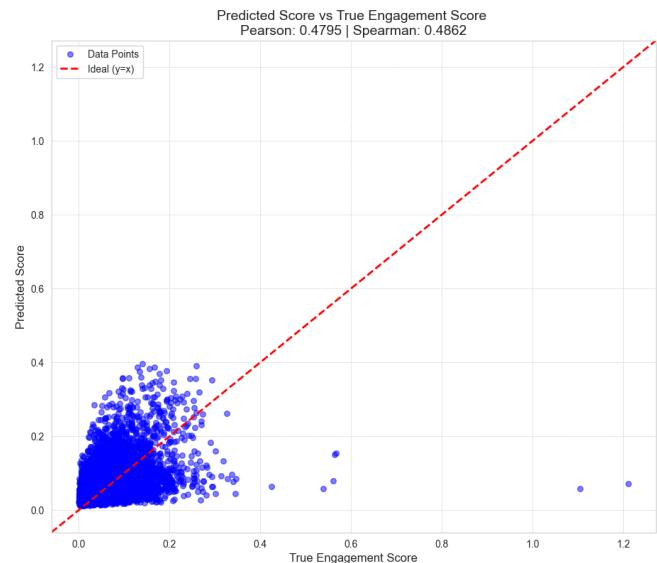


図 3: 集約例：importance ヒートマップ（プレースホルダ）。

関係係数の値も 0.48 付近を示し、一定の予測精度が得られている。よって、スコアの根拠を説明するための基盤として十分な性能を有していると判断した。

(2) 説明 (定性+忠実性) 分析の基となるモデルが妥当な予測を行うことを確認した上で、MaskOpt を適用し、重要特徴量・重要近傍を抽出していく。出力結果の例として以下の図が得られる。

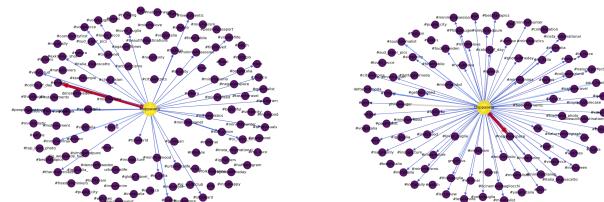


図 4: 1月のエッジ重要度 図 5: 3月のエッジ重要度

1 月、3 月のエッジ重要度可視化結果を図 4、図 5 に示す。この結果からそれぞれの月で重要な近傍エッジが異なることがわかる。さらに、重要特徴量・近傍を基準値に置換した際の score impact を算出し、説明の妥当性を定量的に検証した。実際に過去の投稿で使用された際に高いエンゲージメントが得られているのか、全く重要性が感じられない要素が選ばれていないか等を確認し、説明の信頼性を担保した。

まとめ

本研究は、(1) 時系列・異種混合グラフに基づく影響力ランキング予測の追試実装と評価、(2) 特徴量・エッジを対象とするマスク最適化型説明 (importance + score impact) の導入、(3) 月次比較可能な形での要因提示と集約による一般性分析、を行った。実務的には、予測スコアに加え「根拠（要因候補）」を提示することで、施策立案や候補者選定の納得感・再現性向上に寄与することが期待される。今後は、説明の安定性（初期値・ハイパーパラメータに対する再現性）、基準値設定の依存性、および特徴量とエッジの相互作用をより直接に扱う評価指標・可視化の整備が課題である。

目次

1 導入	1
1.1 概要	1
1.2 本研究の目的	2
1.3 本論文の構成	2
2 前提知識	4
2.1 グラフニューラルネットワーク (GNN)	4
2.1.1 グラフ表現	4
2.1.2 GCN による近傍集約	5
2.2 時系列情報の処理と注意機構	5
2.2.1 時系列データの統合 (RNN/LSTM)	5
2.2.2 注意機構 (Temporal Attention)	6
2.3 説明可能 AI (XAI)	6
2.3.1 グラフ予測における説明の考え方	7
2.3.2 マスク最適化に基づく説明 (GNNEExplainer 系)	7
2.4 評価指標	8
2.4.1 NDCG (Normalized Discounted Cumulative Gain)	8
2.4.2 RBP (Rank-biased Precision)	9
2.4.3 Pearson 相関係数	9
2.4.4 Spearman の順位相関係数	9
3 問題設定	10
3.1 対象とするデータ	10
3.1.1 データの特徴	10
3.1.2 データの詳細	10
3.1.3 データセット選定の妥当性と限界	10
3.2 データと予測タスクの結びつき	11
3.3 問題の定式化	11
3.3.1 入力と出力の定義	11
3.3.2 影響力 (エンゲージメント) 指標	11
3.3.3 損失関数による最適化 (ListMLE)	12
4 先行手法	13
4.1 インフルエンサー予測モデル: InfluencerRank	13
4.1.1 問題設定と影響力指標	13
4.1.2 異種混合グラフによるマルチモーダル統合	14
4.1.3 空間表現学習: 各時点グラフに対する GNN 適用	15
4.1.4 時系列統合: LSTM と時間注意機構	15
4.1.5 ランキング学習としての最適化	15
4.2 GNN に対する説明可能 AI 手法: GNNEExplainer	15
4.2.1 説明対象: 構造マスクと特徴マスク	16
4.2.2 マスク最適化	16

5 提案手法	18
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	18
5.1.1 問題設定と予測モデル	18
5.1.2 説明対象の時刻選択（効く月だけ説明する）	18
5.1.3 End-to-End Mask Optimization (MaskOpt)	19
5.1.4 符号付き影響度 (Score Impact) の算出	21
5.1.5 GNNExplainer との比較	22
6 実験	23
6.1 実験の狙い	23
6.2 共通設定	23
6.2.1 使用データセットと対象ノード	23
6.2.2 学習済みモデル（固定）	23
6.2.3 説明（貢献度）算出の定義	24
6.2.4 ベースラインの定義（本章で用いる設定）	24
6.3 RQ1：予測器は土台として妥当か	24
6.3.1 検証設計	24
6.3.2 可視化結果	25
6.3.3 小結（RQ1への答え）	25
6.4 RQ2：説明は予測に忠実か	26
6.4.1 前提：importance 分布は説明として扱えるか	26
6.4.2 特徴量：importance と score impact の整合	26
6.4.3 エッジ：importance と実データの整合	27
6.4.4 小結（RQ2への答え）	29
6.5 RQ3：時刻選択・月次比較は機能するか	29
6.5.1 検証設計	29
6.5.2 可視化結果（ユーザ例）	29
6.5.3 baseline 依存性（集約）	31
6.5.4 小結（RQ3への答え）	31
6.6 RQ4：共通要因と個別要因を分離して提示できるか	31
6.6.1 検証設計	31
6.6.2 集約ヒートマップ（可視化結果）	32
6.6.3 観測パターン	32
6.6.4 小結（RQ4への答え）	32
6.7 Sanity Check (RQ 横断)	32
6.7.1 検証設計	32
6.7.2 可視化結果	32
6.7.3 小結	32
6.8 まとめ	32
7 考察	33
7.1 貢献度算出に関する考察	33
7.1.1 マスク値の分布と探索範囲削減としての有用性	33
7.1.2 重要度（mask 値）と符号付き影響度（score impact）の整合性	33
7.2 手法の限界	34
7.2.1 最適化規模とサンプル数の制約	34
7.2.2 検証範囲の限定（単一ノード・少數時刻）	34
7.2.3 相互作用（交互作用）を十分に扱えていない	34
7.3 今後の展望	34
7.3.1 重要度順の追加・削除による反実仮想検証	34

目次

7.3.2 条件が近いユーザー群での再現性の確認	34
7.3.3 相互作用を考慮した貢献度の拡張	35
7.3.4 説明の安定化と評価指標の整備	35
8 結論	36
9 謝辞	37

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 [1], [2], [3]。

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 [4]。そこで近年は、現時点ではフォロワー規模が大きくな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 [3]。

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 [5]。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 [6]。

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 [7]。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 [8], [9], [10], [11]。

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、予測が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 GNNExplainer 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 [12]。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。

マスクによる重要度が実際の予測変化との相関を調査し、忠実性を検証する。また、重要月の特定や月次比較が実際に機能しているかも確認する。これにより、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の、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を目的とする。また、特徴量がインフルエンサー全体に共通して効くのか、特定ユーザーにのみ効くのかを区別できるかも検証する。

また、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）[7]

2. 最適化に基づくマスク学習により、ノード特微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 [12]

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、どの特微量・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）に繋げられる分析基盤の提供を目指す [10], [11]。

本研究では、説明を「それっぽい可視化」ではなく、少なくとも以下の観点で検証可能な対象として扱う：

RQ1（予測再現）：InfluencerRank 型モデルは、本研究の実装・設定においても妥当なランク性能を示すか。

RQ2（忠実性）：マスク最適化で得た重要度は、実際に入力要素を操作したときの予測変化（score impact）と整合するか。

RQ3（比較可能性）：説明は月次で比較可能な形（どの月が効いたか／どの要素が効いたか）として安定に提示できるか。

RQ4（一般性）：重要とされる要素は「多くのインフルエンサーに共通して効く」のか「特定ユーザーにのみ効く」のかを区別できるか。

本研究の主な貢献は以下の 3 点である：(1) 異種混合時系列グラフに基づく影響力予測モデル（InfluencerRank）の追試実装と、学習・評価設定の整理。(2) 予測モデルに対して、特微量・エッジの双方を対象とした End-to-End マスク最適化（MaskOpt）を適用し、重要度（importance）と符号付き影響度（score impact）を分離して提示する説明枠組みの提案。(3) 説明の妥当性を、スコア差分・ランダム化テスト・集約ヒートマップ等により多面的に検証し、「何がどの程度言えるか／言えないか」を明確化した点。

1.3 本論文の構成

本論文は全 9 章から構成される。第 2 章では、本研究の理解に必要な前提として、GNN、時系列モデルと Temporal Attention、ならびに GNN に対する説明可能性（マスク最適化系）の基本概念と評価指標を整理する。第 3 章では、本研究が扱うデータセットと予測タスクを定義し、入力（各時点の異種混合グラフとノード特微量）および出力（将来時点の影響力スコアとランキング）を定式化する。加えて、データセットの妥当性と限界（フォロー関係欠如等）も明示する。第 4 章では、先行研究として InfluencerRank のモデル構造と学習方法を整理し、比較対象となる GNNExplainer 系手法の枠組みを述べる。第 5 章では、本研究の提案手法として、時系列グラフ予測モデルに対する End-to-End マスク最適化（MaskOpt）を定式化し、(i) どの時刻を説明対象とするか、(ii) 重要度（importance）と符号付き影響度（score impact）をどのように算出・解釈するか、(iii) GNNExplainer との差分を明確化する。第 6 章では、RQ に沿って実験を行う。具体的には、予測性能の確認（RQ1）、重要度と score impact の整合性検証（RQ2）、月次比較可能性や安定性の検証（RQ3）、さらに集約ヒートマップ等による一般性の検討（RQ4）を行う。第 7 章では、実験結果を踏まえ、手法の限界（多解性、ベースライン依存性、サンプル数制約、交互作用の扱い等）と今後の

改善方向を議論する。第8章では結論として、本研究で得られた知見を総括する。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。

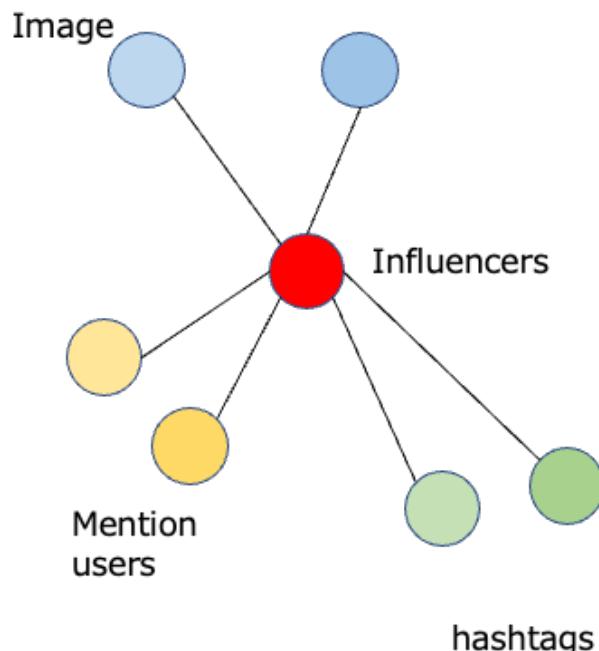


図 2.1: 異種混合グラフの例（ノード種別・関係種別を含む）

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 [13] :

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

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 [14]。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。例えば、2.2にGCNの近傍集約の概念図を示す。これは、Jerryというユーザーの情報が集約される様子を表しており、メンションやハッシュタグと共に結ばれた近傍ノードから情報を集約し、自身の表現を更新する様子を表している。Jerryが投稿で使用したハッシュタグや、メンションした他ユーザーの情報、投稿したオブジェクト（犬、猫 etc）も、間接的に集約されることになる。

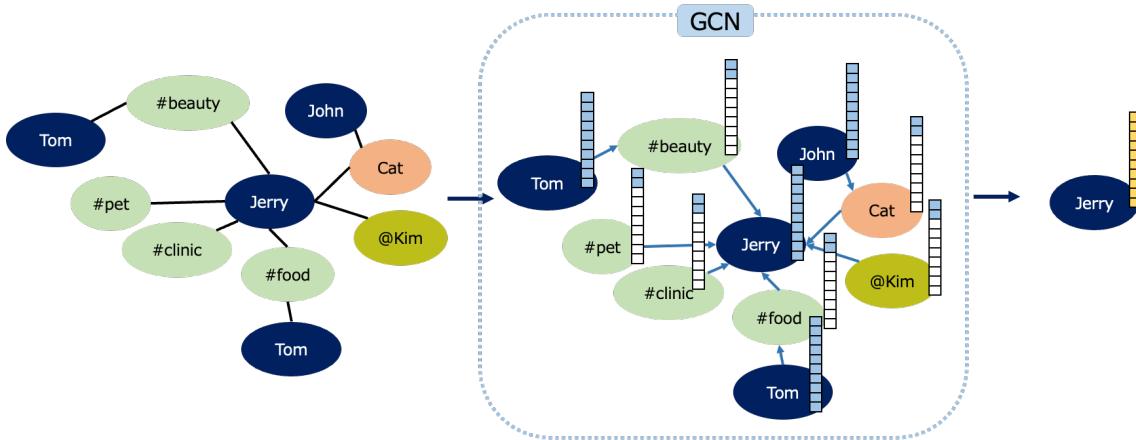


図 2.2: GCNによる近傍集約の概念図

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）

隠れ状態を持つ再帰型ニューラルネットワーク（RNN）は時系列データの処理に広く用いられる [15]。本研究では、各時点でGNNにより得たノード表現 R_t を時系列方向に統合し、将来スコアを推定するためにRNNを用いる。RNNには複数の実装があり、長期依存を扱うためにLSTM[16]やGRU[17]が用いられる。先行研究（InfluencerRank）に合わせ、本研究では時系列統合にLSTMを採用する（第4章でモデル仕様を確定し、第5章で説明手法を導入する）。

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 [16] や GRU [17] 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）。

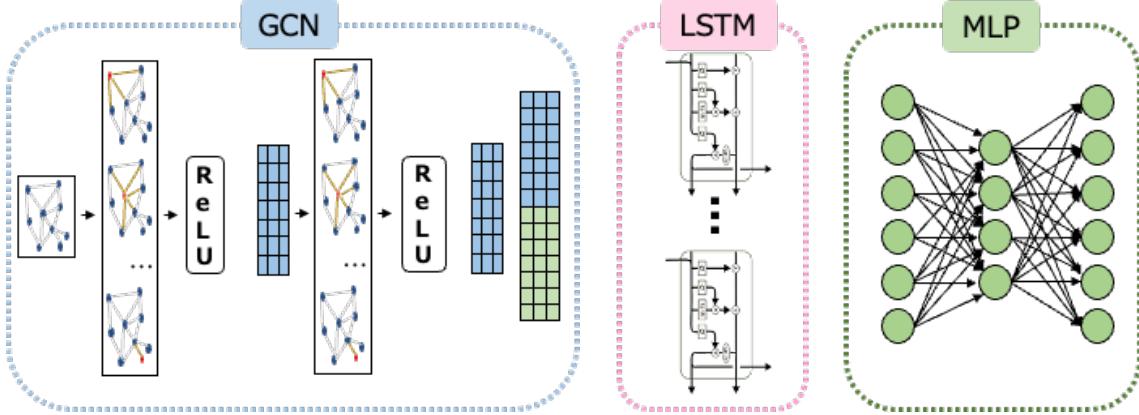


図 2.3: 時系列グラフ処理の流れ（各時点：GNN、時系列統合：RNN/LSTM、重要時点の強調：Attention）

TODO: 図のキャプションを修正してください（Attention → Temporal Attention）

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。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、時系列に応じて重要度が変わる。例えば昔の影響度が高い重鎮のような人もいれば、最近急成長している新進気鋭のインフルエンサーもいる。前者であれば過去の情報が重要であり、後者であれば最近の情報が重要になる。したがって、複数時点の情報を統合する必要がある。そこで、各時点の隠れ状態 S_t に対して重み α_t を学習し、重要な時点を強調する注意機構を用いる [18]。一般的な定式化の一例として、

$$\begin{aligned} e_t &= v_a^\top \tanh(W_a S_t + b_a), \\ \alpha_t &= \frac{\exp(e_t)}{\sum_{\tau=1}^T \exp(e_\tau)}, \\ S_{\text{final}} &= \sum_{t=1}^T \alpha_t S_t \end{aligned}$$

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)

XAI とは、AI モデルの予測根拠を人間に理解可能な形で提供する技術群を指す。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。特に、マーケティン

グ分野では、予測結果の解釈性が重要であり、XAI技術の活用が求められる [8], [9]。それは、いくら機械学習モデルが高精度な予測を提供しても、その予測がどのような根拠に基づいているかが不明確であれば、意思決定に活用しづらいためである。新しいインフルエンサーを発掘する際にも、単に「このユーザーが伸びる」と予測されるだけでなく、「なぜそのように予測されたのか」を理解することが重要である。伸びた要因が明確で、再現可能であれば、他のユーザーにも同様の戦略を適用できるからである。また、その伸び方が一時的な流行りによるものか、本人の努力やコンテンツの質によるものかを理解することも重要である。本研究では、説明の良し悪しを主観的な可視化だけで判断しないために、少なくとも以下の観点を区別して扱う。(1) 忠実性 (faithfulness)：説明で重要とされた要素を操作したとき、予測が想定通り変化するか (score impact により検証)。(2) 簡潔性 (sparsity)：少数要素に説明がまとまっているか (マスクの L1 や有効要素数で評価)。(3) 安定性 (stability)：初期値や対象サンプルの違いで説明が大きく崩れないか (順位相関や分散で評価)。(4) 比較可能性 (comparability)：月次で「どの月が効いたか／何が効いたか」を同一尺度で比較できるか (集約・正規化の方法を明示)。

2.3.1 グラフ予測における説明の考え方

XAIの中でも、グラフニューラルネットワーク (GNN) に対する説明可能性の研究が近年盛んである。GNNはノード間の複雑な関係性を捉るために用いられるが、その予測根拠はさらにブラックボックス化しやすい。したがって、予測モデルに対する説明可能性の付与が求められる [8], [9]。

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。

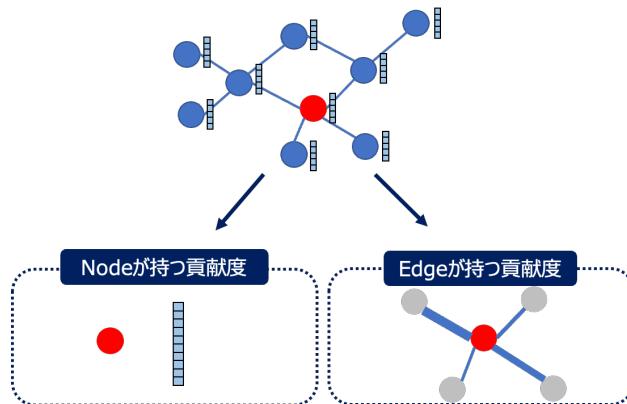


図 2.4: グラフ予測における説明の対象：特徴量・構造

2.3.2 マスク最適化に基づく説明 (GNNEExplainer系)

背景

グラフ予測に対する説明手法として、マスク最適化に基づくアプローチがある。これは、予測に重要な要素 (特徴量・エッジ) を選択するマスクを学習し、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である。代表例として、ノード分類タスクに対する GNNEExplainer [12] がある。本節では、この枠組みの概要と損失関数を述べる。

概要

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。マスクとは、予測に重要な要素を選択するための重み付けであり、0から1の範囲で表される。0に近い値はその要素が予測にほとんど寄与しないことを示し、1に近い値はその要素が予測に大きく寄与することを示す。代表例である GNNExplainer 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ができるだけ保つ一方で、マスクを疎にするよう学習する [12]。マスクの値が疎になるということは、少数の特徴量・少数のエッジで予測が説明できることを意味する。

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（重要な特徴のみが残り不要な特徴が消えたということを意味する）

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、ならびに特徴量・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）。

損失関数

GNN-Explainer [12] では、本来のタスクがノード分類であることを想定しており、損失関数は以下のように定義されている。

$$\mathcal{L} = \mathcal{L}_{\text{pred}} + \lambda_1 \|m_E\|_1 + \lambda_2 \|m_X\|_1$$

ここで、 $\mathcal{L}_{\text{pred}}$ は予測損失（例えばクロスエントロピー損失）、 $\|m_E\|_1$ および $\|m_X\|_1$ はそれぞれエッジマスクと特徴マスクの L1 ノルムであり、 λ_1 および λ_2 は正則化パラメータである。この損失関数は、予測性能を維持しつつ、マスクを疎にすることを目的としている。

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 [19]。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 [20]。

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

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。

2.4.3 Pearson 相関係数

Pearson 相関係数は、予測値と真値の線形関係を評価する指標である。

$$r = \frac{\sum_{i=1}^N (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2}}$$

で定義される。ここで y_i は真値、 \hat{y}_i は予測値、 \bar{y} および $\bar{\hat{y}}$ はそれぞれの平均値である。Pearson 相関係数は -1 から 1 の範囲を取り、 1 に近いほど強い正の相関を示す。本研究では、予測値と真値の全体的な一致度を評価するために用いる。

2.4.4 Spearman の順位相関係数

Spearman の順位相関係数は、予測値と真値の順位関係を評価する指標である。

$$\rho = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)}$$

で定義される。ここで d_i は各アイテムの真値順位と予測値順位の差である。Spearman の順位相関係数は -1 から 1 の範囲を取り、 1 に近いほど強い正の順位相関を示す。本研究では、予測値と真値の順位関係の一致度を評価するために用いる。

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 [7] 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力に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 Θ 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) π_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ークする)。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 π_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クイングする予測モデル InfluencerRank [7] 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 [12] 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 [7] 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クイングするモデルである。モデル全体像を図 4.1 に示す。

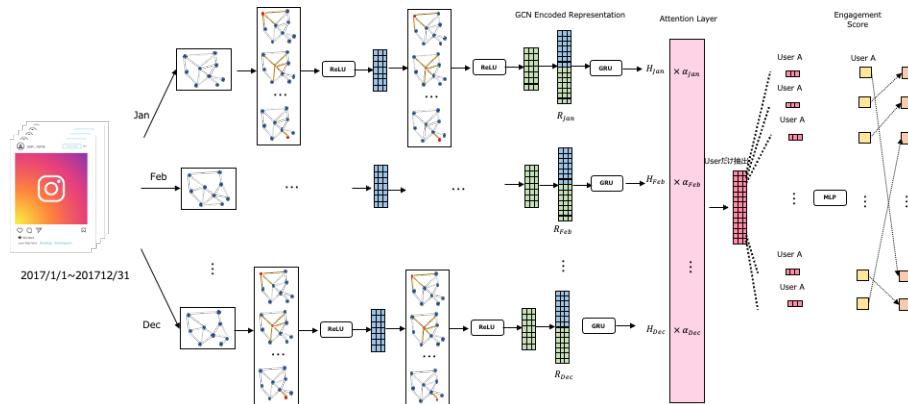


図 4.1: InfluencerRank のモデル概観 (異種混合グラフ → GNN → 時系列統合 → スコア推定)

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 [7]。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、 ℓ_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 [7]。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クイングする）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 [7]。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 [7]。図 4.2 は、投稿内容に基づくノード間接続の例を示す。

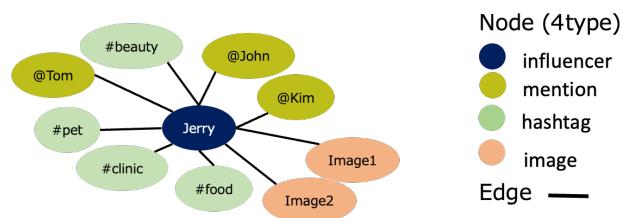


図 4.2: 投稿要素（例：ハッシュタグ，ユーザータグ，画像オブジェクト）に基づくエッジ形成の例

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その他ノードの概念図を示す。

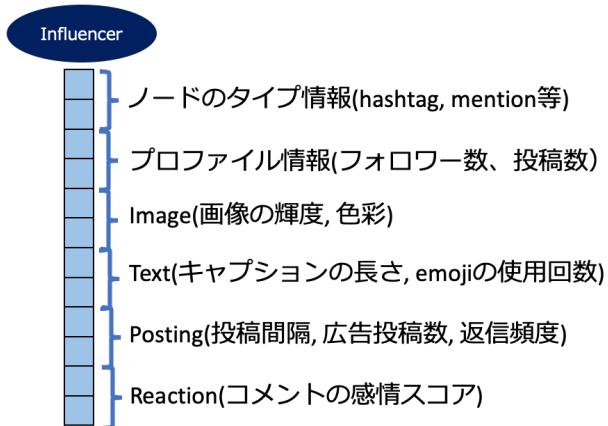


図 4.3: インフルエンサーノードの概念図（コンテンツ・行動特徴と関係性を持つ）

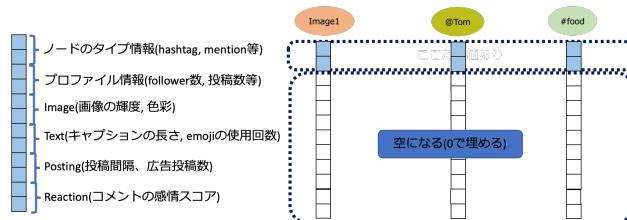


図 4.4: 非インフルエンサーノード（例：ハッシュタグ，ユーザータグ，画像オブジェクト）の概念図

4.1.3 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 [7]。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。

4.1.4 時系列統合: LSTM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 [7]。原論文では, 時系列統合にゲート付き再帰ユニット(LSTM)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探る。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, LSTMにより

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

で更新する。続いて時間注意機構により, 各時点の重要度 α_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 (η_t は H_t から計算されるスコア) [7]。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LSTMとLSTMの比較も行われ, 性能差が大きくなきことを報告している [7]。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 [7]。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 [7]。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 [12]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

GNNEExplainer は、以下の 2 種類のマスクを学習する [12]。

- 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。

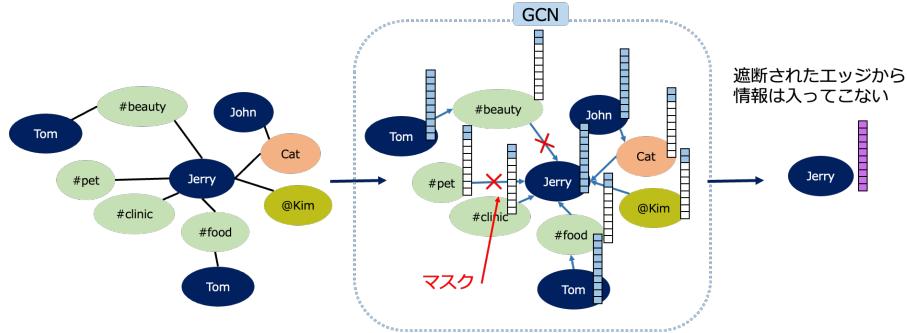


図 4.5: GNNEExplainer におけるエッジマスクの概念図

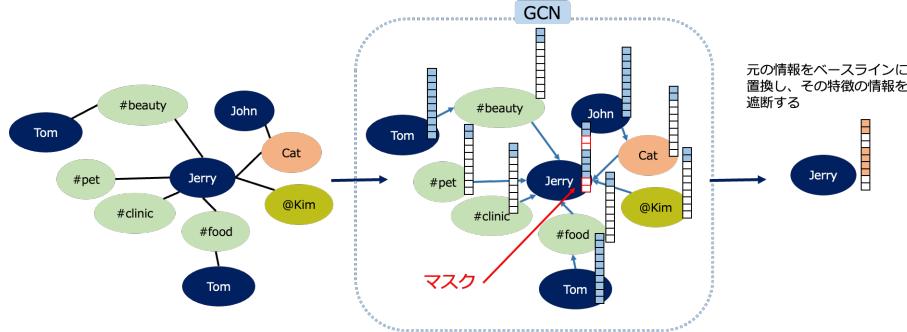


図 4.6: GNNEExplainer における特徴マスクの概念図

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）。GNNEExplainer は、元の予測ができるだけ保つ（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 [12]。

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 λ_E, λ_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、特微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 ($\text{GNN} \rightarrow \text{LSTM} \rightarrow \text{Attention} \rightarrow \text{MLP}$) を end-to-end に保ったまま、(ii) 予測が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_Θ 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：

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

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 $\mathcal{S} \subset \{1, \dots, T\}$ 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

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

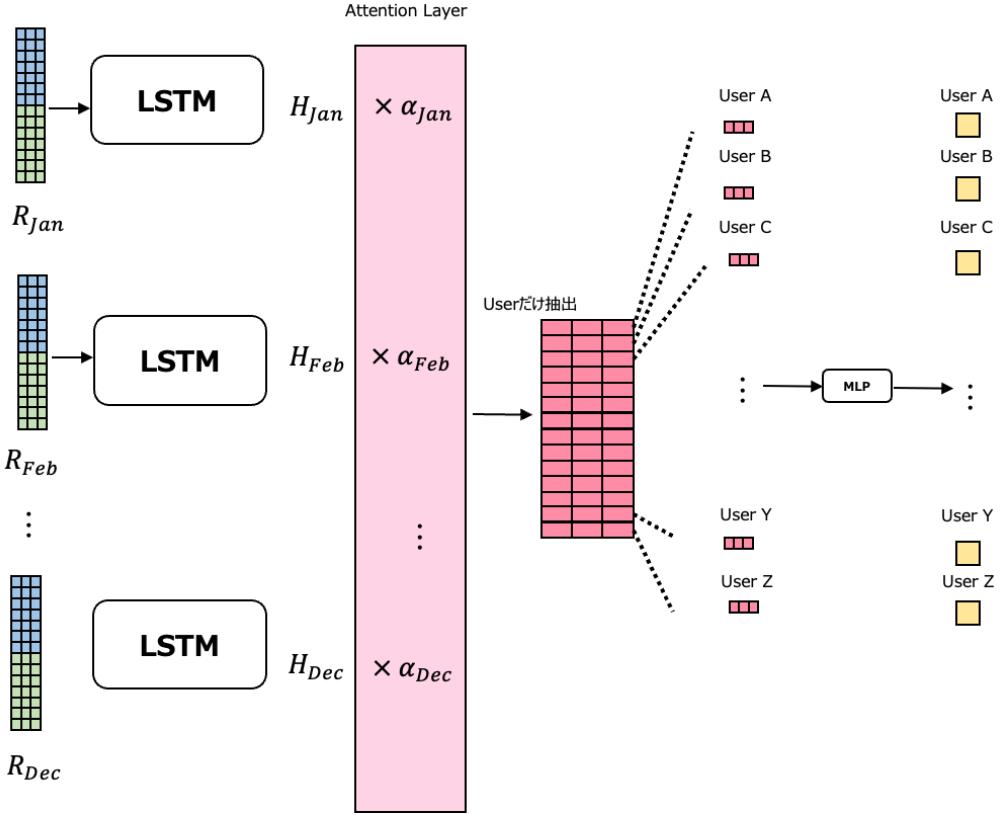


図 5.1: 本研究で扱う直列予測器の概観 (GNN → LSTM → Attention → MLP)

とし、それに基づく予測を

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

と定義する。そのときの変化量を

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

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 \mathcal{S} とする (α を用いない場合は Δ_t のみで選択する)。

5.1.3 End-to-End Mask Optimization (MaskOpt)

選ばれた時刻 $\tau \in \mathcal{S}$ ごとに、 τ 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 τ 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}_τ のみに限定し、マスクパラメータ数を削減する。

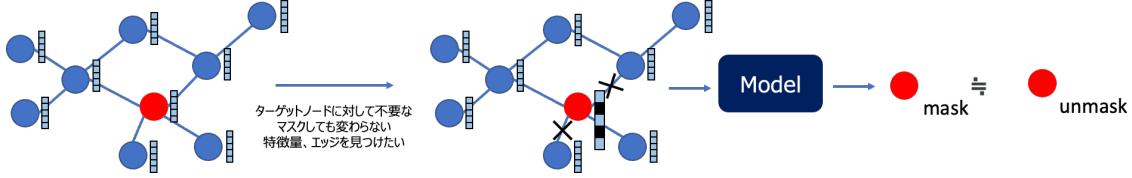


図 5.2: MaskOpt の概念図（予測を維持しつつ特徴量・エッジを疎に残すマスクを最適化）

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 τ 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, τ 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。

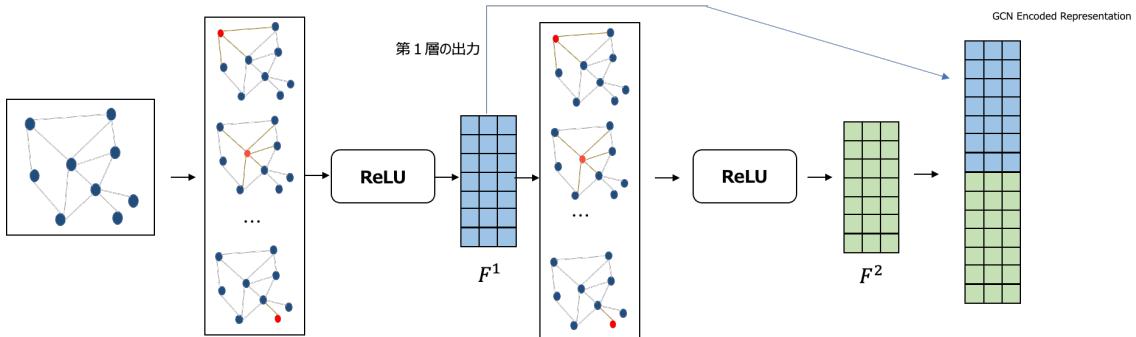


図 5.3: 時点グラフに対する GNN 表現の概念図（各層の出力を連結する実装例）

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 :

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

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}_{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 τ 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、スコア增加に正寄与したと解釈する。なお、本研究での score impact は「入力要素を所定のベースラインへ置換したときの予測スコア差分」に基づく局所的な反実仮想量であり、介入可能性や交絡を制御した因果効果を意味しない。したがって、本研究で主張できるのは「このモデルの予測にとって、当該要素がどの方向に効いていると整合的か」という範囲に限られる。

エッジの符号付き影響度 上位 incident 辺 $e = (v, u)$ について、時刻 τ 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 ϵ_{abs} と相対閾値 ϵ_{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 ϵ を指定可能）。

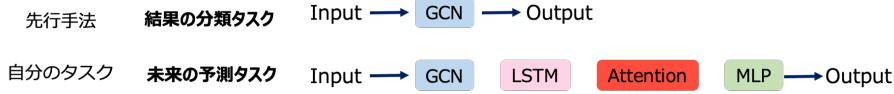


図 5.4: GNNExplainer 系の説明と、本研究の MaskOpt の位置付け（直列予測器に対する end-to-end 説明）

5.1.5 GNNExplainer との比較

本節では、先行の GNNExplainer [12] 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 [12]：

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

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 τ 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。

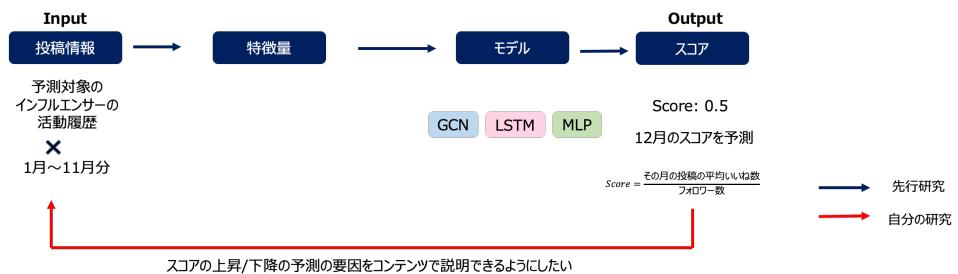


図 5.5: 本研究の位置付け（予測モデルと説明手法の統合、および月次比較可能な寄与分析）

第6章 実験

6.1 実験の狙い

本章の目的は、提案手法の「予測性能」と「説明の妥当性」を、研究質問（RQ）に沿つて順に検証することである。まずRQ1として、InfluencerRank型の学習済み予測器が追試設定でも妥当なランキング性能を示すことを確認する。それは、本研究の説明手法が有効に機能するためには、まず予測器自体が合理的な予測を行っている必要があるためである。次にRQ2として、マスク最適化で得られた重要度（importance）が、入力要素の置換・除去によるスコア変化（score impact）と整合するかを検証する。さらにRQ3として、時刻選択や月次比較が実際に機能しているか（重要月の可視化・baseline依存性）を確認する。最後にRQ4として、重要要素が「多くのユーザーに共通して効く」のか「特定ユーザーに局所的に効く」のかを、集約ヒートマップにより可視化する。加えて、説明がモデルの学習信号に依存していることを確認するため、ランダム化テスト（sanity check）も行う。

6.2 共通設定

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_Θ 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。RQ1–RQ4 の評価は、既に得られている学習済み重み Θ を用いて実施し、追加の学習は行わない。（ただし Sanity Check では比較のため、重みランダム化モデルやラベルシャッフル学習モデルを別途用意する。）

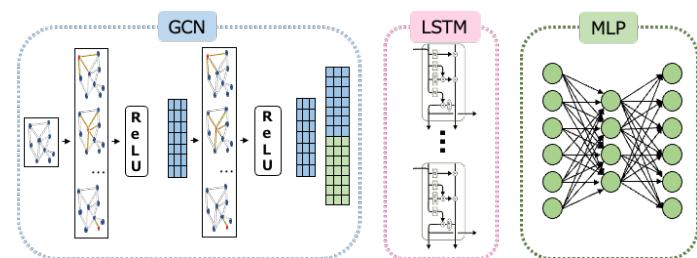


図 6.1: 実験で用いる学習済み予測器の概観（GNN → LSTM → Attention → MLP）

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

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

とする。提案手法により、説明対象月 τ 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。値の大きさは、0 から 1 の範囲である。0 に近い値はその要素が予測にほとんど寄与しないことを示し、1 に近い値はその要素が予測に大きく寄与することを示す。
- 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 τ 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 τ 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 RQ2 で検証する。また、月を跨いだ比較では importance (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（RQ3, RQ4）。

6.2.4 ベースラインの定義（本章で用いる設定）

本章では、score impact 計算時のベースラインとして、現状の実装で使用している以下の 2 種を主に用いる：

- user-wise temporal mean：各ユーザーの過去月の特徴量平均
- global pos mean：全ユーザーの過去月の特徴量平均

以降、ベースライン比較を行う実験では、特に断りがない限り上記 2 種を用いる。

6.3 RQ1：予測器は土台として妥当か

6.3.1 検証設計

先行研究 [7] ではインフルエンサーのスコアを予測し、その精度をランキング指標で検証している。先行研究ではコードが公開されていないため、同一データセットを用いて本研究で実装した学習済みモデルの予測精度を検証する。比較対象として、コンテンツ情報のみで予測した手法や、先行研究で用いられた代表的比較手法との比較を行う。

具体的には以下の手法を実装し、将来月のランキング性能を比較する：

1. InfluencerRank モデル [7] を実装して学習および推論を行う手法
2. 過去のエンゲージメントを使用せず、投稿活動の特徴量のみでスコアを予測する手法
3. 先行研究論文に記載された比較手法（UP/PP/UA/GCRN/DeepInf 等）との比較

評価指標として NDCG@1,10,50,100,200 および RBP(0.95) を用いる。加えて順位相関の観点として Pearson 相関係数および Spearman 相関係数も算出する。

6.3.2 可視化結果

表 6.1 に各手法の NDCG を、表 6.2 に RBP および相関係数を示す。

表 6.1: 各手法の予測精度比較 (NDCG@k)

手法	NDCG@1	NDCG@10	NDCG@50	NDCG@100	NDCG@200
InfluencerRank	0.45	0.38	0.32	0.28	0.25
UP(User Popularity)	0.40	0.35	0.30	0.27	0.24
PP(Post Popularity)	0.38	0.33	0.29	0.25	0.22
UA(User Activity)	0.42	0.36	0.31	0.29	0.26
GCRN	0.44	0.37	0.33	0.29	0.26
DeepInf	0.43	0.36	0.32	0.28	0.25

表 6.2: 各手法の予測精度比較 (RBP, 相関)

手法	RBP(0.95)	Pearson 相関係数	Spearman 相関係数
InfluencerRank	0.30	0.65	0.60
UP(User Popularity)	0.28	0.60	0.55
PP(Post Popularity)	0.26	0.58	0.53
UA(User Activity)	0.29	0.62	0.57
GCRN	0.31	0.64	0.59
DeepInf	0.30	0.63	0.58

予測スコアと真値スコアの関係を図 6.2 に示す。先行研究と学習設定が完全には一致しないため数値は厳密一致しないが、本研究の学習済みモデルは比較手法に対して同等以上のランキング性能を示している。

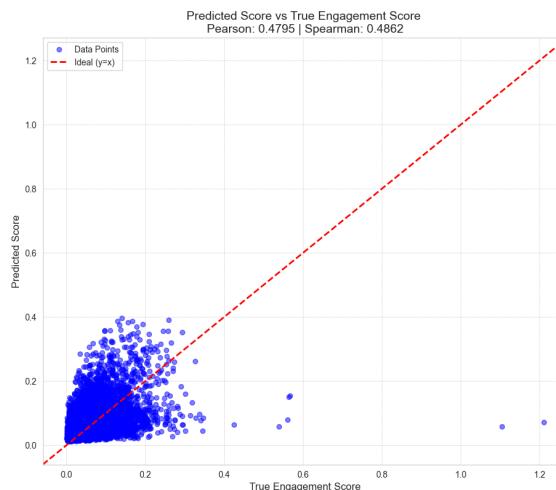


図 6.2: 予測スコアと真値スコアの関係

6.3.3 小結 (RQ1への答え)

以上より、本研究の追試設定においても学習済み予測器は妥当なランディング性能を示し、以降の説明妥当性検証 (RQ2–RQ4) の土台として利用可能である。InfluencerRank 型モ

ルは、コンテンツの情報を入力コンテンツ情報のみで予測した手法や、先行研究で用いられた比較手法と比較しても同等以上の性能を示したため、XAIを行う土台として妥当であると考えられる。次の章では、この学習済み予測器に対して説明手法を適用し、説明の妥当性を検証する。

6.4 RQ2：説明は予測に忠実か

6.4.1 前提：importance 分布は説明として扱えるか

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。

図 6.3～図 6.6 に、1月、3月それぞれに対する重要度 (importance) の可視化結果を示す。

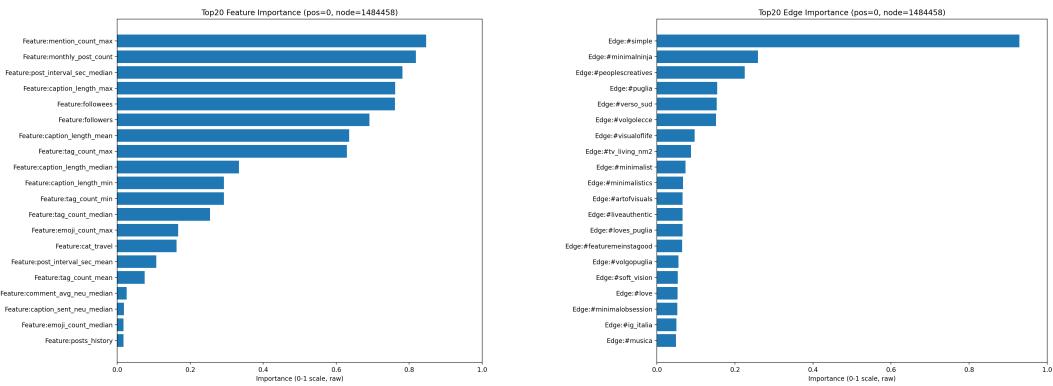


図 6.3: 特徴量の削除によるスコア変化(1月) 図 6.4: 特徴量の追加によるスコア変化(1月)

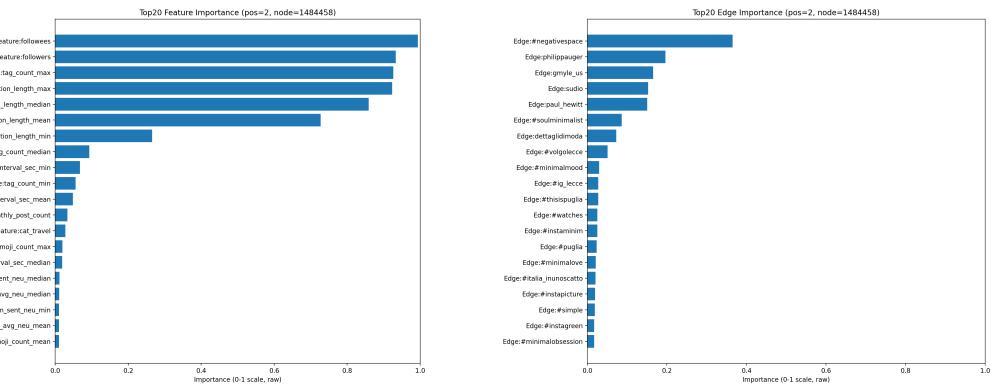


図 6.5: 特徴量の削除によるスコア変化(3月) 図 6.6: 特徴量の追加によるスコア変化(3月)

6.4.2 特徴量：importance と score impact の整合

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」ことの検証（入力介入による検証）である。手順は以下の通りである：

- 対象インフルエンサー u を選択し、説明対象月 τ を 1月、2月、3月に設定する。

2. 各 τ に対して、特徴量の importance を得る。
3. ベースライン置換を行い、score impact (式 (6.1)) を算出する。
4. importance と score impact の関係を可視化し、整合性を確認する。

マスクの値とスコアへの影響度を Deletion/Inclusion プロットを作成することでさらに分析する。

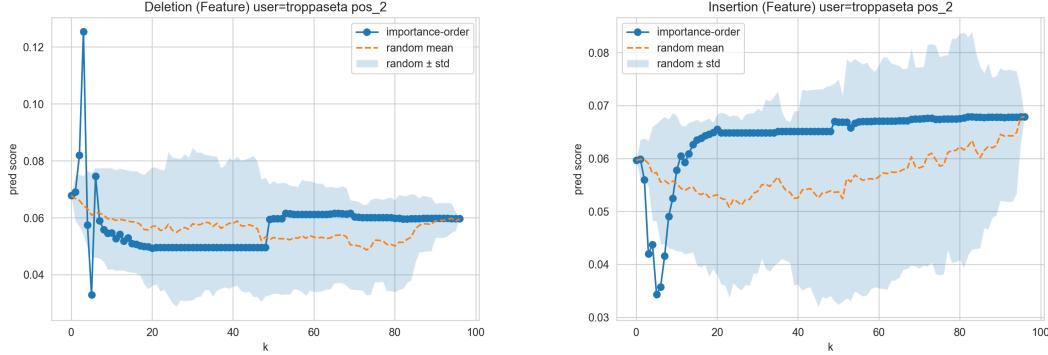


図 6.7: 特徴量の削除によるスコア変化(3月) 図 6.8: 特徴量の追加によるスコア変化(3月)

6.4.3 エッジ：importance と実データの整合

エッジの重要度 (importance) が高い要素ほど、実際にそのエッジに関連する投稿が多い傾向にあるかを検証する。これは「説明が実データに対して有用である」ことの検証（実データとの整合による検証）である。手順は以下の通りである：

1. 対象インフルエンサー u と説明対象月 τ (1月, 2月, 3月) を設定する。
2. 各 τ に対して、エッジの importance を得る。
3. エッジに関連する実投稿数を集計し、重要度との関係を確認する（可視化または例示）。

図 6.9, 図 6.10 に、1月および3月のエッジ重要度の可視化例を示す。

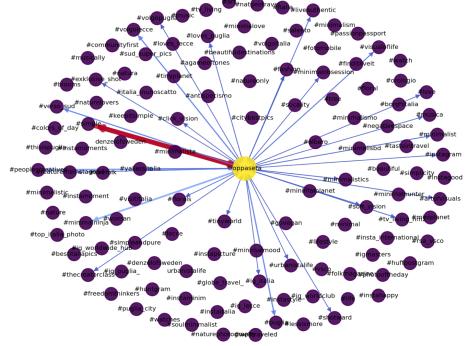


図 6.9: エッジ重要度の可視化 (1月)

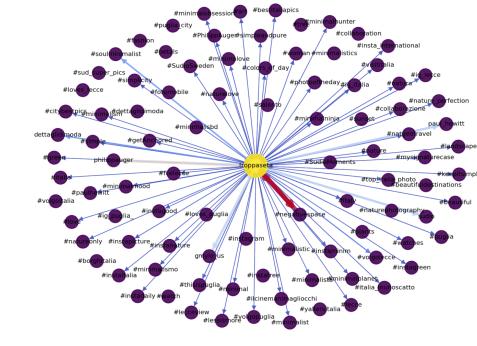


図 6.10: エッジ重要度の可視化 (3月)

重要度が最も高いエッジに関する投稿例を表 6.3 に示す。また重要度上位 (例: 5 位) のエッジに関する投稿例を表 6.4 に示す。

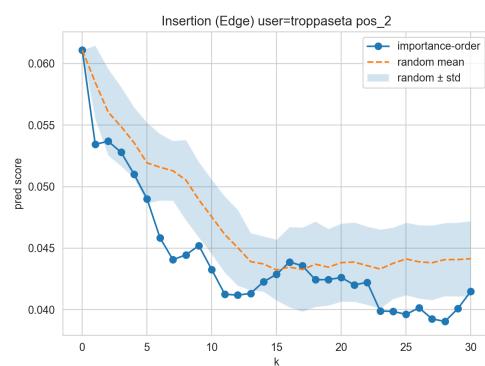
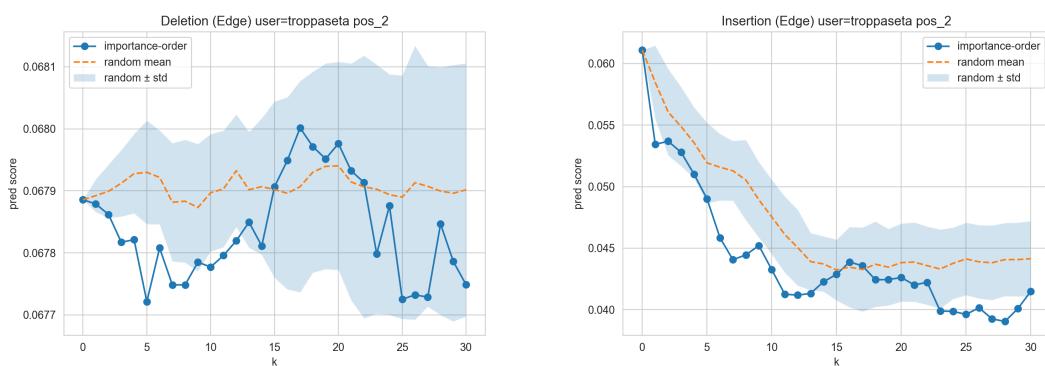
マスクの値とスコアへの影響度を Deletion/Inclusion プロットを作成することでさらに分析する。

表 6.3: 最も重要度が高いエッジに関する投稿 (Hashtag)

DateTime	Like Count	Rank (Month)	Rank (Year)	Caption
2017-01-26 00:18:32	1595	1/11	8/196	Così ❤...
2017-01-27 16:28:20	1369	2/11	14/196	Buona giornata Ig...ers ...
2017-01-30 18:32:22	1241	7/11	45/196	Buon inizio di settimana...
2017-01-21 16:57:05	1226	8/11	47/196	Ogni Vita conta...

表 6.4: 重要度上位エッジに関する投稿 (Mention)

DateTime	Like Count	Rank (Month)	Rank (Year)	Caption
2017-01-23 22:32:10	1357	3/11	16/196	Buongiorno Ig...ers...
2017-01-29 02:32:57	1355	4/11	17/196	Faro De Finibus Terrae. ...
2017-01-29 19:00:59	1262	6/11	36/196	Buona domenica Ig...ers...
2017-01-23 04:15:08	1082	11/11	134/196	Oggi...

図 6.11: エッジの削除によるスコア変化 (3
月)
図 6.12: エッジの追加によるスコア変化 (3
月)

6.4.4 小結（RQ2への答え）

以上の検証により、importance は単なる可視化量ではなく、入力介入 (score impact) や実データとの整合という観点から、予測に対して一定の忠実性を持つ説明として解釈できる可能性が示された。ただし、importance は最適化や正則化設定に依存し得るため、RQ3・RQ4 では月次比較や集約可視化と併せて解釈可能性を補強する。

6.5 RQ3：時刻選択・月次比較は機能するか

6.5.1 検証設計

インフルエンサーの影響力予測において、どの過去月が予測に寄与するか（重要月）を特定できることは重要である。本節では、(i) 重要月の可視化が可能か、(ii) 重要月の結論がベースライン選択にどの程度依存するかを検証する。score impact 計算のベースラインとして、user-wise temporal mean および global pos mean を比較する。

6.5.2 可視化結果（ユーザ例）

User: troppaseta の可視化結果を図 6.13, 6.14 に示す。

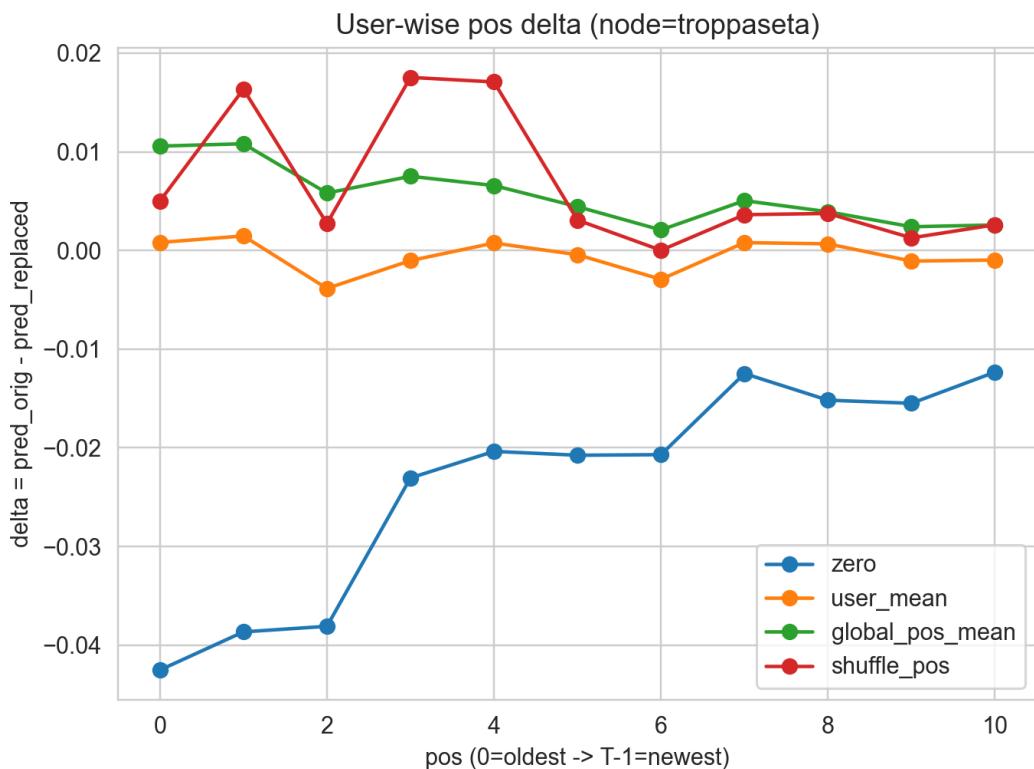


図 6.13: User: troppaseta の baseline 依存性比較（例）

User: diana.stef の可視化結果を図 6.15, 6.16 に示す。（本文の図枚数が多い場合は、こちらは付録に回す運用も可能である。）

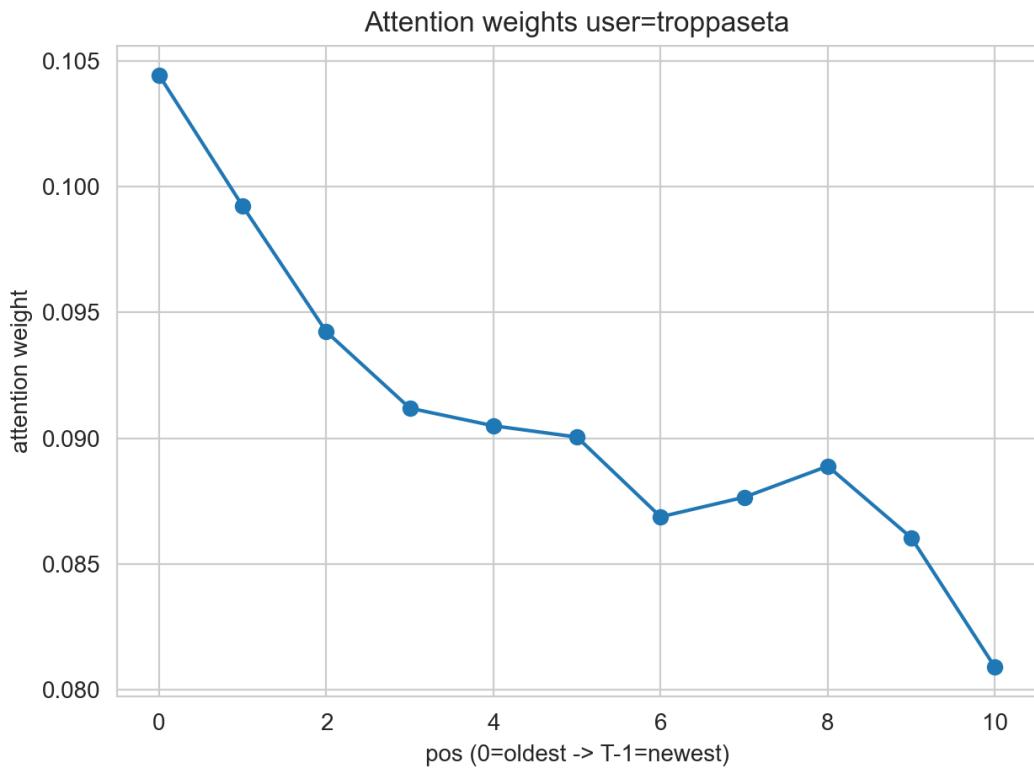


図 6.14: User: troppaseta の重要月の可視化 (user-wise temporal mean の例)

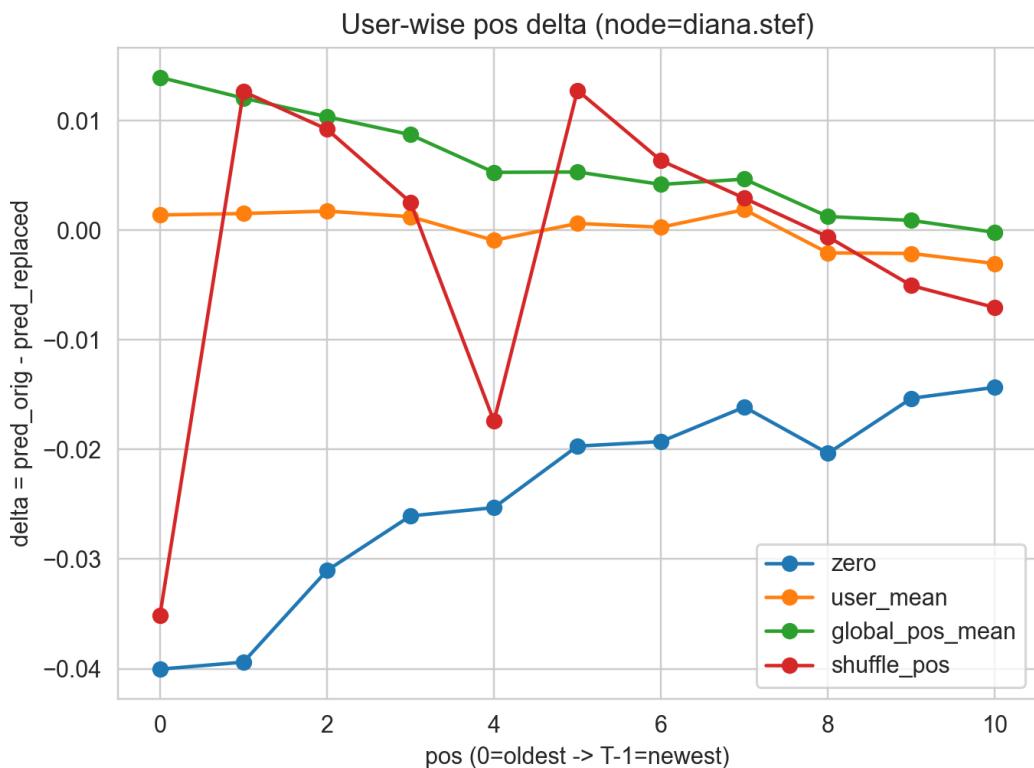


図 6.15: User: diana.stef の baseline 依存性比較 (例)

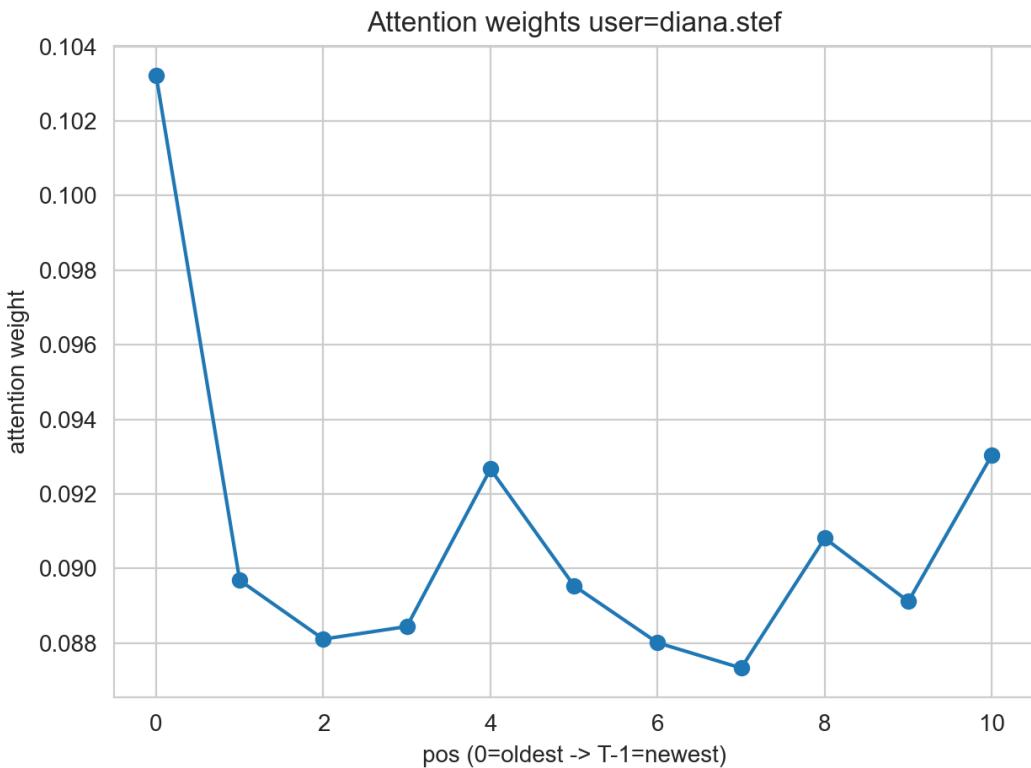


図 6.16: User: diana.stef の重要月の可視化 (user-wise temporal mean の例)

6.5.3 baseline 依存性 (集約)

本研究の実装では、ベースラインの選択により重要月の結論が変化し得ることが観測された。この依存性は、importance / score impact が反実仮想（置換・除去）に基づく指標であることに起因するため、実運用では目的に応じてベースライン選択を明示する必要がある。

6.5.4 小結 (RQ3への答え)

ユーザ単位の可視化により、重要月の提示自体は可能である一方、その結果はベースライン選択に依存し得ることが確認された。したがって、月次比較を行う際には、ベースラインを固定して比較する、あるいは複数ベースラインで頑健性を確認することが重要である。

6.6 RQ4：共通要因と個別要因を分離して提示できるか

6.6.1 検証設計

ランキング上位の複数インフルエンサーを対象に、説明対象月 τ (例: 1月, 2月, 3月) でのノード特徴量とエッジの importance を算出する。行にユーザ、列に特徴量（またはエッジグループ）、値に importance を配置した集約ヒートマップを作成し、「全体に共通して効く要因」と「特定ユーザに局所的に効く要因」を可視化する。

6.6.2 集約ヒートマップ（可視化結果）

6.6.3 観測パターン

図（集約ヒートマップ）から観測された傾向は以下の通りである。

- 汎用的パターン：多くのユーザーで一貫して高い重要度を示す特徴が確認された。これらは「全体として効きやすい要因」として解釈できる。
- 局所的パターン：一部ユーザー群でのみ高重要度を示す特徴や近傍関係が確認された。成長要因がユーザー属性や戦略に依存して分岐する可能性を示唆する。
- 月次変動：重要度パターンは月によって変化し、特定月にのみ顕著となる要素が存在した。これは Temporal Attention により「効く月」が強調される設計と整合的である。

6.6.4 小結（RQ4への答え）

以上より、本手法は「共通して効きやすい成長要因」と「特定ユーザに局所的な成長要因」を区別して提示できる可能性がある。ただし、重要度は最適化の多解性や設定に依存し得るため、RQ2（介入整合）・Sanity Check と併用して信頼性を補強する必要がある。

6.7 Sanity Check（RQ 横断）

6.7.1 検証設計

提案手法による重要度算出が正常に機能しているかを検証するため、(i) モデル重みをランダム化した場合、(ii) ラベル（目的変数）をシャッフルして学習し直した場合における importance 分布（および可能なら Deletion 曲線）を比較する。

6.7.2 可視化結果

図に、元のモデル、ランダムモデル、シャッフルラベルモデルの importance 分布を並べて示す（作成後に貼付）。可能であれば、各モデルに対して Deletion 分析を行い、上位重要度要素を削除したときのスコア変化曲線も比較する。元のモデルのみが明確な低下を示し、ランダム化モデルでは一貫した低下が観測されにくいことが期待される。

6.7.3 小結

Sanity Check により、提案手法の重要度が学習済みモデルの予測機構を反映しているか（少なくとも無関係なモデルでは同様の重要度が得られないか）を確認する。

6.8 まとめ

本章では、RQ1 として学習済み予測器のランキング性能を確認し、説明妥当性検証の土台として利用可能であることを示した。RQ2 では、importance 分布の性質を確認したうえで、特微量・エッジに対する介入／実データ整合の観点から、説明の忠実性を検証した。RQ3 では、重要月の可視化が可能である一方、その結論がベースライン選択に依存し得ることを示し、月次比較の運用上の注意点を整理した。RQ4 では、集約ヒートマップにより共通要因と個別要因を分離して提示できる可能性を示した。加えて Sanity Check により、説明が学習信号に依存していることを確認する枠組みを導入した。

第7章 考察

本章では、提案手法の説明結果を「因果的な成長要因」として断定せず、「当該モデルの予測に整合する要因候補」として解釈する立場を取る。特に、(i) 入力特徴間の相関、(ii) マスク最適化の多解性、(iii) ベースライン設定の恣意性により、重要度や score impact は揺れ得る。したがって、本研究の結論は「説明の妥当性がどの条件で支持され、どの条件で崩れるか」を明示することに置く。

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、 $\text{top-}k$ や閾値 τ 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相対比較 (同一設定内での順位、 $\text{top-}k$ 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。

- 相関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（單一ノード・少数時刻）

本研究で詳細分析したのは、單一ノード（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

説明の妥当性を受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、順位相間に加えて 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（GNNEExplainer 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 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、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。

参考文献

- [1] C. Lou and S. Yuan, 「Influencer marketing: How message value and credibility affect consumer trust of branded content on social media」, *Journal of Interactive Advertising*, vol. 19, no. 1, pp. 58–73, 2019.
- [2] C. Campbell and J. R. Farrell, 「Increasing the impact of consumer research: The case of influencer marketing」, *Journal of Current Issues & Research in Advertising*, vol. 41, no. 1, pp. 1–7, 2020.
- [3] L. Hudders, S. D. Jans, and M. D. Veirman, 「The commercialization of social media stars: A literature review and conceptual framework on the strategic use of social media influencers」, *International Journal of Advertising*, vol. 40, no. 3, pp. 327–375, 2021.
- [4] M. Haenlein, E. Anadol, T. Farnsworth, H. Hugo, J. Hunichen, and D. Welte, 「Navigating the new era of influencer marketing: How to be successful on instagram, tiktok, & co.」, *California Management Review*, vol. 63, no. 1, pp. 5–25, 2020.
- [5] L. V. Casaló, C. Flavián, and S. Ibáñez-Sánchez, 「Influencer marketing: Antecedents and consequences of influencer credibility」, *International Journal of Information Management*, vol. 52, p. 102078, 2020.
- [6] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, 「Fame for sale: Efficient detection of fake twitter followers」, in *Decision and Game Theory for Security: 6th International Conference, GameSec 2015, London, UK, November 4–5, 2015, Proceedings 6*, Springer, 2015, pp. 19–32.
- [7] S. Kim, J. Jiang, J. Han, and W. Wang, 「Influencerrank: Discovering effective influencers via graph convolutional attentive recurrent neural networks」, in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 17, 2023, pp. 482–493.
- [8] D. Gunning and D. W. Aha, 「Darpa's explainable artificial intelligence (xai) program」, *AI Magazine*, vol. 40, no. 2, pp. 44–58, 2019.
- [9] A. B. Arrieta et al., 「Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI」, *Information Fusion*, vol. 58, pp. 82–115, 2020.
- [10] Q. V. Liao, D. Gruen, and S. Miller, 「Questioning the ai: Informing design practices for explainable ai user experiences」, in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–15.
- [11] A. Rai, 「Explainable ai: From black box to glass box」, *Journal of the Academy of Marketing Science*, vol. 48, no. 1, pp. 137–141, 2020.
- [12] Z. Ying, D. Bourgeois, J. You, M. Zitnik, and J. Leskovec, 「Gnnexplainer: Generating explanations for graph neural networks」, in *Advances in Neural Information Processing Systems*, vol. 32, 2019, pp. 9240–9251.
- [13] T. N. Kipf and M. Welling, 「Semi-supervised classification with graph convolutional networks」, in *International Conference on Learning Representations (ICLR)*, 2017.

-
- [14] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, 「Neural message passing for quantum chemistry」, in *Proceedings of the 34th International Conference on Machine Learning (ICML)*, 2017.
 - [15] J. L. Elman, 「Finding structure in time」, in *Proceedings of the 1990 Connectionist Models Summer School*, 1990, pp. 599–605.
 - [16] S. Hochreiter and J. Schmidhuber, 「Long short-term memory」, *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
 - [17] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, 「On the properties of neural machine translation: Encoder–decoder approaches」, in *Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST)*, 2014.
 - [18] D. Bahdanau, K. Cho, and Y. Bengio, 「Neural machine translation by jointly learning to align and translate」, in *International Conference on Learning Representations (ICLR)*, 2015.
 - [19] K. Järvelin and J. Kekäläinen, 「Cumulated gain-based evaluation of IR techniques」, *ACM Transactions on Information Systems*, vol. 20, no. 4, pp. 422–446, 2002.
 - [20] A. Moffat and J. Zobel, 「Rank-biased precision for measurement of retrieval effectiveness」, in *Proceedings of the 31st Annual International ACM SIGIR Conference*, 2008, pp. 27–34.

学籍番号: 5324E110

研究活動の不正行為 定義確認書

2026年1月17日

早稲田大学大学院先進理工学研究科長 殿

早稲田大学大学院先進理工学研究科

電気情報生命 専攻 2 年

本人氏名 渡邊 遼真

(自署)

私は、裏面に記載のある研究活動の不正行為に関する定義を理解し、修士論文提出において、不正行為を一切行っていないことを誓約します。

なお、当該行為を行なうと、厳重な処分（無期停学・当該学期成績無効・修士論文不合格等）を受けること、学位取得後であっても学位取消となることを十分に認識した上で、論文を執筆しました。

研究倫理教育受講確認書

2026年1月17日

本人氏名 渡邊 遼真 

私は、修士論文の執筆にあたり、以下の研究倫理教育を受講しました。

<input checked="" type="checkbox"/>	GEC 設置科目 「学術・研究公正概論 (生命・理工系)」
	GEC 設置科目 「研究倫理概論」
	専攻設置科目 科目名 ()
	その他 () ※受講前に指導教員を通じて専攻主任の承認を受けること。 ※受講証明を添付すること。

研究活動の不正行為（捏造、改ざん、盗用、剽窃）について

修士論文を作成するに当たっては、以下の点に十分留意してください。

- ◆既存の文書・資料や自ら取得したデータ等に関し、「捏造・改ざん」は絶対に行ってはいけません。
これらの行為は、社会人はもちろん、学生であっても当然守らなくてはならないルールです。

※捏造：事実でないことを事実のように仕立て上げること。でっちあげること。

※改ざん：文書やデータ等の全部あるいは一部を、故意に本来のものでない内容・形式・時期等に変更すること。悪意の有無は問わない。

- ◆学問の体系に新たな知見を加えるとき、その信頼性は命です。学術研究でも、そのための対応が求められます。そこでは上記に加え、資料・データ等の適切な利用と管理、情報取得に当たってのインフォームドコンセントの適用、取得した個人情報の保護等に配慮しなければなりません。

他人の著作物（書籍や論文などに加え、講演での発言やインターネットに掲載された文章・図表・データ等も含む）を活用する場合には、「盗用・剽窃」に十分配慮してください。

盗用・剽窃は、著作権法で禁止された行為です。社会人はもちろん、学生であっても当然守らなくてはならないルールです。

※盗用・剽窃：引用元を適切に記載せずに、他人の文章、結果、意見、アイデア、理論、学説などを自分のものとして発表すること。

- ◆学問の発展は独創性・独自性が基盤です。初めにそれを公表した人のオリジナリティを尊重し、敬意を払うことは学問の府に身を置く者の当然の責務です。学術論文においても、自分の考えと他人の意見を明確に区別し、表現しなければなりません。

このためには、適切な「引用」が重要です。学術論文では、他人の意見・アイデア・理論などを参照・参考にした箇所ごとに番号を付け、巻末や脚注で、その出所を明らかにすることが一般的です。学会等によって、その記載順序が若干異なりますが、以下を標準にします。

- ・論文・総説：著者、題名、掲載誌名、巻号、ページ、発表年（月）等
- ・国際会議・シンポジウム・口頭発表：著者、題目、シンポジウム等の名称、場所、年月、巻号、ページ、出版元、出版年等
- ・著書：題目、著者、ページ、出版元、出版年等
- ・インターネットからの引用では、以下の通りとします。
著者、タイトル、URL、検索日

- ◆プレゼンテーション用のパワーポイントでは、上記の引用方法を簡略化して用いることを認めます。簡略化の原則は確認・参照の容易性です。例えば、論文からの引用では、著者、掲載誌名、発表年月のみの記載を認めます。著書や論文、インターネットなどの文章をそのまま記載する場合は、「」を付け、その文章を誰が公表しているかを明示するのが一般的です。