

# 感情分析と時系列モデルを統合した データ駆動型予測システムの開発 ー PBL 科目実践紹介ー

○下沢 亮太郎<sup>1</sup>

<sup>1</sup> 東京都立産業技術高等専門学校, ものづくり工学科 情報通信工学コース

キーワード: 機械学習 時系列予測 感情分析 BERT LSTM

## 1 緒言

都立産技高専荒川キャンパスでは, 選抜された 3~5 年生を対象に, IoT/AI 技術を学習する未来工学教育プログラムを実施している. 5 年生の PBL(Project Based Learning) 科目では, 学生たちの自由な発想を基に, 学習した知見や技術を利用してアイデア実装に挑戦している. 本報告では, PBL 科目で取り組んだ内容について紹介する.

社会的な事象に影響を与えることが多いのは文字列や映像なことが多いという事実に対して有名な時系列予測モデルは入力に数値しかとらないことが多い.

そこで, 本研究ではニュースの文章を BERT という自然言語に特化した分類モデルが出力した値を「感情らしさ」として採用し, よく使われる指標とともに時系列予測に特化した LSTM という機械学習モデルで学習し, 予測した.

## 2 BERT の転移学習

BERT(Bidirectional Encoder Representations from Transformers) は Google によって事前学習されたモデルをクライアント側で用意したデータセットで転移学習 (Fine-tuning) することで任意の形式で出力するモデルを作ることができる.

ChatGPT で作成した存在しない企業のニュースとそれに応じた失望, 楽観, 懸念, 興奮, 安定の 5 つのパラメータに点数を振ったデータセットを作成し, 転移学習を行った.

転移学習を行った BERT モデルの推論例を表 1 に示す. プラスの文章として, 「【速報】世界が注目する MVIDIA が決算発表「最終的な利益 前年比 7.3 倍 2 兆 3300 億円」勢い止まらず」, マイナスの文章として「UU スチール買収計画が窮地に鉄鉄, 訴訟も視野」という架空のニュースを推論した. プラスの文章では楽観と興奮, マイナスの文章では懸念と失望が高く出ていることから目標とした傾向を持っているモデルを作ることができたと考えられる.

プラスの文章				
失望	楽観	懸念	興奮	安定
0.0534	0.395	0.109	0.231	0.212
マイナスの文章				
失望	楽観	懸念	興奮	安定
0.254	0.110	0.351	0.135	0.150

表 1: 推論結果例

## 3 結言

ご質問がありましたら以下にご連絡ください。

お問い合わせ先

氏名: 高田 拓

E-mail: takada@metro-cit.ac.jp