

吉野貴晶のクオンツピックス：NO30 大規模言語モデル(BERT)を用いたアナリストレポート解析

BERTモデルとアナリストレポートの計量化について考察してみよう

- BERT、Transformerのモデル構造及び、ファインチューニングについて
- アナリストレポートの計量化及び、センチメントスコアの銘柄選択効果について

0. 予備知識、キーワード

基本的な自然言語処理、深層学習、金融工学の知識を仮定します。

キーワード：大規模言語モデル、Transformer、BERT、感情分析、シミュレーション

1. イントロダクション

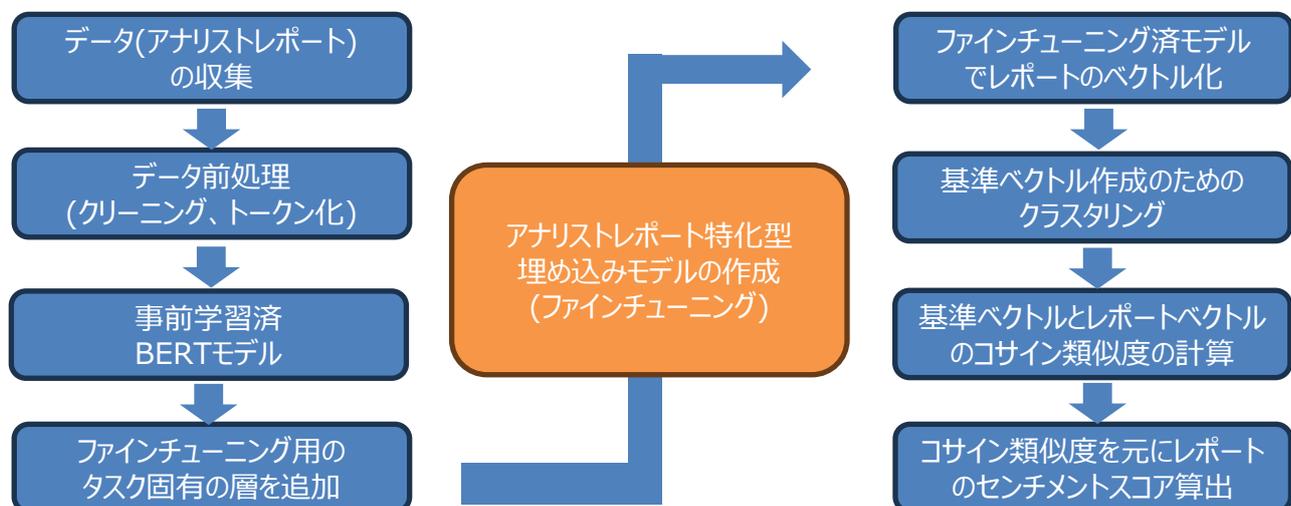
AI時代と呼ばれる昨今では、資産運用業界でもAIの実務活用が行われています。例えば、企業や投資家が意思決定を行う際に膨大な量のアナリストレポートを迅速かつ正確に分析することが求められており、自然言語処理（NLP）を用いた感情分析が注目されています。感情分析は、テキストデータからポジティブ、ネガティブ、中立といった感情の傾向を自動的に抽出・分類する技術であり、アナリストレポートの内容を定量的に評価することを可能にします。この定量評価された数値のことをセンチメントスコアと言います。今回は、BERTと呼ばれるモデルを用いて算出したアナリストレポートのベクトルをもとに、算出したセンチメントスコアの銘柄選択効果について紹介をしていきます。

2. 自然言語モデルを用いたアナリストレポートの計量化

今回は、Googleが開発したBERTと呼ばれる深層学習モデルを対象とします。このモデルの特徴の1つとして、文脈を考慮した分散表現を生成できることがあります。例えば、BERTから得られる単語の分散表現は、同じ単語でも文脈（周りの文章）が変わると、それに応じて異なる値をとります。文脈を考慮した分散表現を生成することは既存のモデルでも行われていましたが、BERTではAttentionという手法により、離れた位置にある情報も適切に取り入れることができるという特徴があります。この性質により、BERTは文脈を深く考慮したような処理が可能となります。以下の(図表1)は、BERTによるアナリストレポートの分散表現の獲得(ベクトル化)から、センチメントスコア算出までのフローチャートとなります。

さらにアナリストレポートのセンチメントスコア化を知りたい方は[Kudo]を参照ください。

(図表1) BERTによる感情分析とセンチメントスコアの導出イメージ

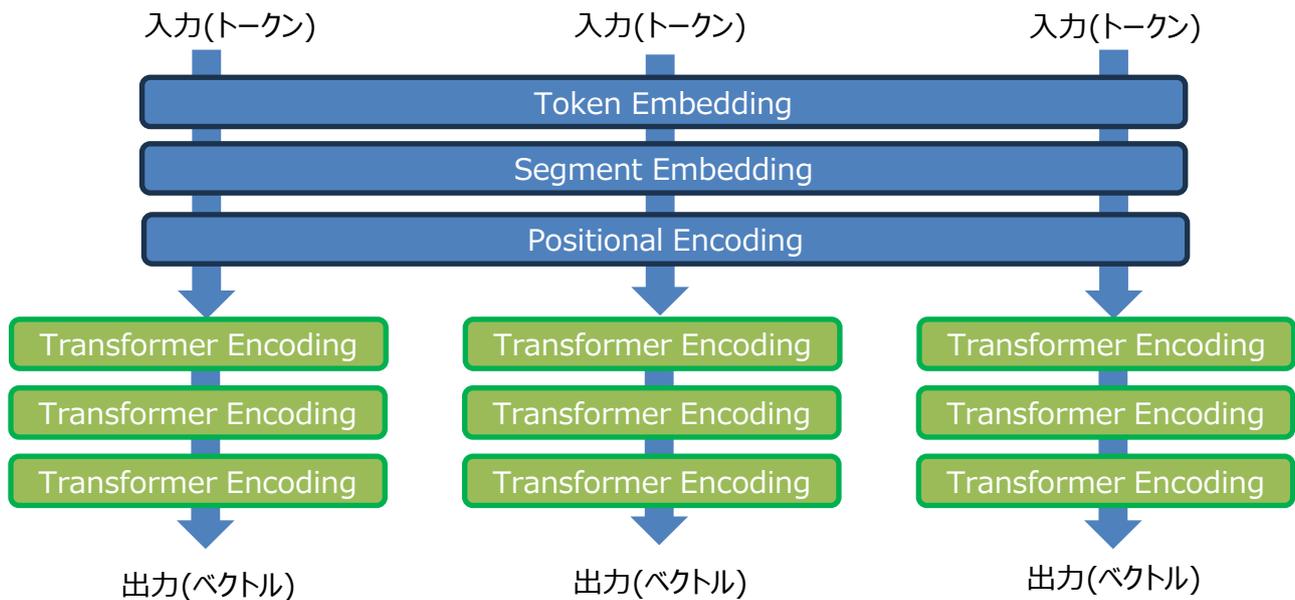


BERT、Transformerのモデル構造

3. Transformer EncoderとAttention機構

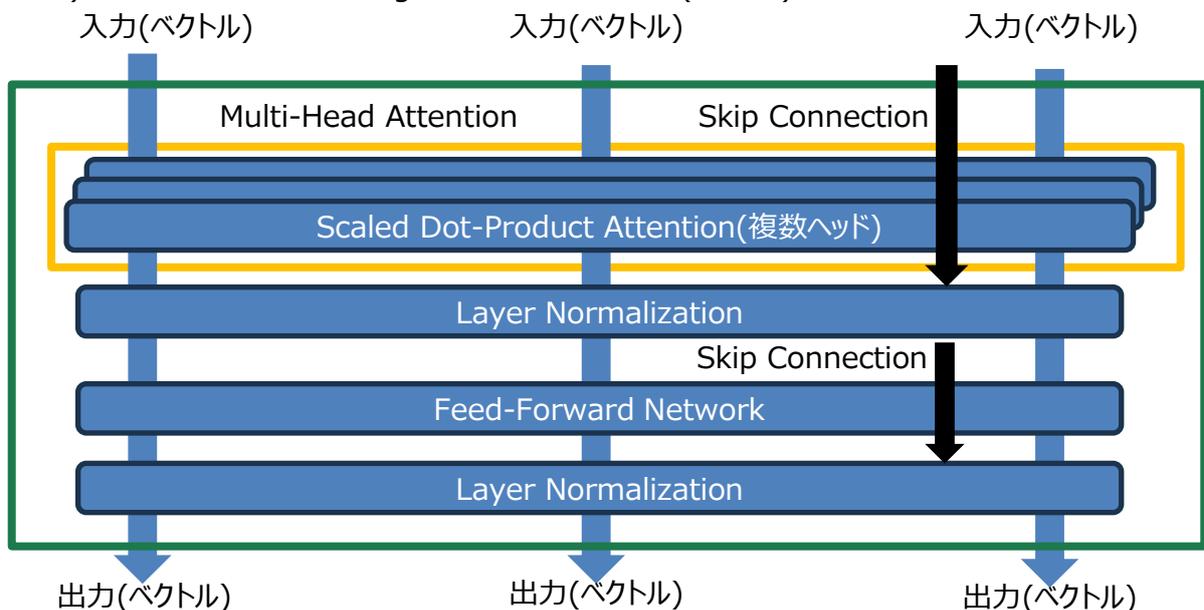
BERTのモデル構造は、Transformer Encodingと呼ばれるAttention機構を用いたニューラルネットワークモデルです。ここでいうEncodingとは、例えばトークン化したテキストをニューラルネットワークに入力して、特徴量（分散表現によって得られるベクトル）を出力することを指します。この分散表現によって、単語の意味という抽象的な概念を、ベクトルによる定量評価が可能となります。具体的なBERTの処理は、トークン化したテキストに対して、特徴空間に埋め込むToken Embeddingと文の区別を行うSegment Embedding、単語の位置情報（文脈情報）を与えるPositional Encodingを合わせ、複数層のTransformer Encodingを重ねてベクトルの出力を行います。（図表2）なお、実際にはトークンをまとめて並列的に処理を行っているため、RNNやLSTMと比較しても計算効率が高いという特徴があります。

（図表2）BERTのモデル構造のイメージ



各層のTransformer Encodingは、Multi-Head Attention、Layer Normalization、Skip Connection、Feed-Forward Networkの組み合わせで計算されます。

（図表3）Transformer Encodingのモデル構造のイメージ(緑枠内)



●当資料は、市場環境に関する情報の提供を目的として、ニッセイアセットマネジメントが作成したものであり、特定の有価証券等の勧誘を目的とするものではありません。●当資料は、信頼できると考えられる情報に基づいて作成しておりますが、情報の正確性、完全性を保証するものではありません。●当資料のグラフ・数値等はあくまでも過去の実績であり、将来の投資収益を示唆あるいは保証するものではありません。また税金・手数料等を考慮しておりませんので、実質的な投資成果を示すものではありません。●当資料のいかなる内容も将来の市場環境の変動等を保証するものではありません。

Attention機構、ファインチューニング

まず最初に、Multi-Head Attentionについて説明します。直前の層で出力した、第*i*番目のトークンに対応したベクトル x_i （横ベクトル）に対し、以下の3つの行列を用いて線形変換を行うことで、それぞれクエリ q_i 、キー k_i 、バリュー v_i と呼ばれる*d*次元ベクトルを用意します。Attention機構では、以下（2）のベクトル w_i が第*i*番目のトークンに対する重み（スコア）とします。*n*はテキスト内のトークン個数とします。

$$q_i = x_i W^Q, \quad k_i = x_i W^K, \quad v_i = x_i W^V \quad (1)$$

$$w_i = (w_{i,1}, \dots, w_{i,n}) = \text{Softmax}\left(\frac{q_i \cdot k_1}{\sqrt{d}}, \dots, \frac{q_i \cdot k_n}{\sqrt{d}}\right) \quad (2)$$

クエリを「注目する単語」、キーを「参照対象の単語」と考えると、式（2）の内積は「どの単語がどの単語に注目すべきかのスコア」であり、それをSoftmax関数によって確率化しています。言い換えると、このスコアが高いほど、トークン間の関連度が高いことを意味しています。また、学習が進みやすくするために、次元数*d*に関してスケールリングを行っています。得られたスコアによってバリュー v_i を重み付けしたものが（3）であり、離れたトークンの情報も考慮できることが分かります。これら一連の処理をScaled Dot-Product Attentionといい、まとめて行列表示すると（4）のようになります。

$$a_i = w_{i,1} v_1 + \dots + w_{i,n} v_n \quad (3)$$

$$Q = XW^Q, \quad K = XW^K, \quad V = XW^V, \quad A = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (4)$$

Multi-Head Attentionは、クエリ、キー、バリューの組(4)を複数用意しておき、それぞれの組に対してScaled Dot-Product Attentionを適用し、最後に出力を一つに集約するような方法です。（複雑になるので本記事では式展開を割愛します。）このことは、例えばあるヘッドが主語述語の関係に注目し、また別のヘッドが形容詞の関係を見つけるといったことができることを意味しています。特に、クエリ、キー、バリューが全て同じものをSelf-Attentionと呼び、系列内（トークン同士）の依存関係を学習するため、文脈理解が得意とされています。

Skip Connectionは、Multi-Head Attentionの入力と出力を足し算することで深い層を持つモデルに対する学習が適切に行えるようにする処理、Layer Normalizationでは、学習の収束を早めるために平均と標準偏差によって正規化をしています。Feed-Forward Networkは、GELU関数と呼ばれる活性化関数に入力する処理で、その後再び足し算と正規化を行ったものがTransformer Encodingの出力となります。基本的なBERTモデルでは、Transformer Encodingの層数が12層、Multi-Head Attentionの組数（ヘッド数）が12個になっています。さらにBERTの技術的な内容を知りたい方は[Tunstall] や[Ohmi] を参照ください。

4. アナリストレポートの感情分析に特化したベクトルの獲得

前節では、BERT（Transformer）のモデル構造について解説しましたが、実務上は毎回一から学習を行うのは効率的でなく、事前学習済みのモデルを利用することが一般的です。また、パラメータ調整（ファインチューニング）を行うことも特定タスクに対して有効です。ここでは、Sentence BERTという手法で、例えばポジティブな内容のアナリストレポートに対して、ポジティブ(ネガティブ)な内容のレポートは類似度が高く(低く)なるように学習を行います。2012年から2024年までのアナリストレポートの感情分析(2値分類)を行った際の評価指標が以下で、やはりファインチューニングを行うと精度が向上することが確認できました。

(図表4) 感情分析に関する評価指標(月次平均)

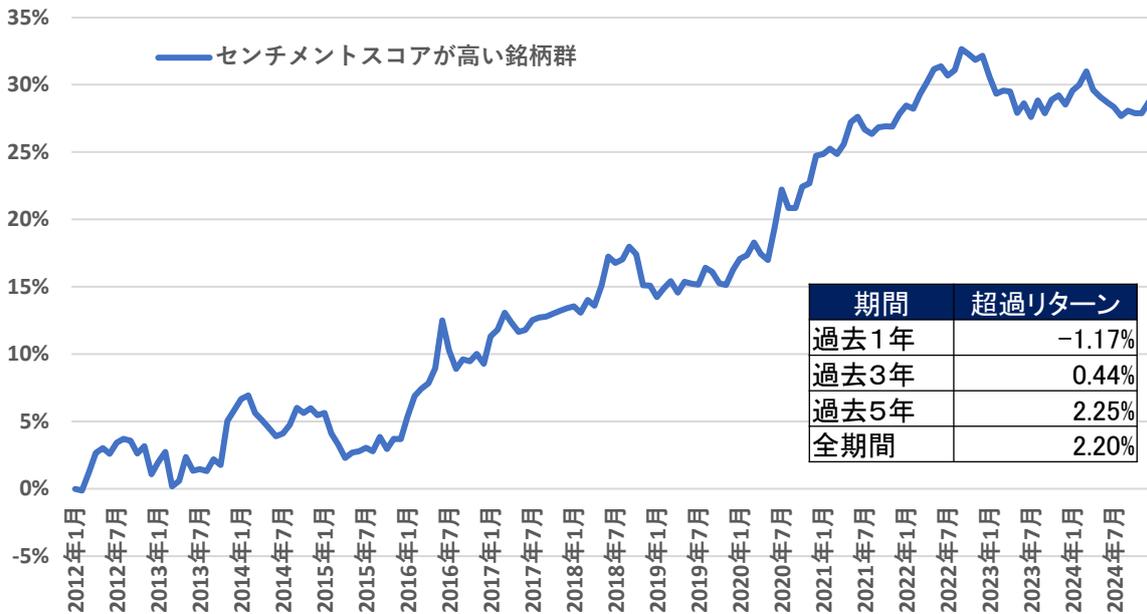
	Accuracy(正解率)	F1(適合率と再現率の調和平均)
事前学習モデルのみ	0.52	0.50
ファインチューニング済モデル	0.69	0.72

センチメントスコアの銘柄選択効果、まとめ

5. BERTベースのセンチメントスコアの銘柄選択効果

最後に、センチメントスコアの投資尺度としての効果を見てみましょう。具体的な検証方法は次の通りです。まずは、2012年を起点とし、毎月末に対象銘柄(おおよそTOPIX500程度)に関して、各月末時点で取得できるアナリストレポート情報を用いて、センチメントスコアを計算し、スコアの高い上位銘柄を抽出します。こうして選んだ銘柄に等金額投資したポートフォリオの翌月のリターンを求め、分析対象とした銘柄全体に等金額投資した場合のリターンを引いて超過部分を計算します。

(図表5) センチメントスコアが高い銘柄の累積株式リターン



出所：東京証券取引所のデータを基に、ニッセイアセットマネジメント作成

5. まとめ

BERTをファインチューニングして得られたアナリストレポートのベクトルからセンチメントスコアを算出し、当スコアが高い銘柄に投資をすると、右肩上がりの超過リターンになることが分かりました。(センチメントスコアの銘柄選択効果が高い。) 一方、直近1年間のようにBERT(深層学習)ベースのセンチメントスコアが効きにくい時期があります。実務上は、これら深層学習による手法と、極性辞書等のカウントベースの手法を上手く組み合わせしていく必要があります。

参考文献

- [Tunstall] Lewis Tunstall,他,機械学習エンジニアのためのTransformers 最先端の自然言語処理ライブラリによるモデル開発,オライリー・ジャパン,(2022)
- [Ohmi]近江崇宏,他, BERTによる自然言語処理入門 Transformersを使った実践プログラミング,ストックマーク株式会社,(2021)
- [Kudo]工藤秀明,他,自然言語処理技術を用いたアナリストレポートの実証分析 センチメントの変化と株式市場の反応について,証券アナリストジャーナル,(2017)

～執筆者の紹介～

吉野貴晶 (写真：右)

「日経ヴェリタス」アナリストランキングのクオンツ部門で16年連続で1位を獲得。ビッグデータやAIを使った運用モデルの開発から、身の回りの意外なデータを使った経済や株価予測まで、幅広く計量手法を駆使した分析や予測を行う。



塚本恵 (写真：左)

ニッセイアセット入社後、投資工学開発部にて主に機械学習を含む数理的な定量的手法、オルタナティブデータを活用した新たな投資戦略の研究開発を担当。

●当資料は、市場環境に関する情報の提供を目的として、ニッセイアセットマネジメントが作成したものであり、特定の有価証券等の勧誘を目的とするものではありません。●当資料は、信頼できると考えられる情報に基づいて作成しておりますが、情報の正確性、完全性を保証するものではありません。●当資料のグラフ・数値等はあくまでも過去の実績であり、将来の投資収益を示唆あるいは保証するものではありません。また税金・手数料等を考慮しておりませんので、実質的な投資成果を示すものではありません。●当資料のいかなる内容も将来の市場環境の変動等を保証するものではありません。