

## 吉野貴晶 のクオンツ トピックス : NO4 AI (人工知能) を利用した投資手法の開発

### 投資分野へのAI (人工知能) の進出

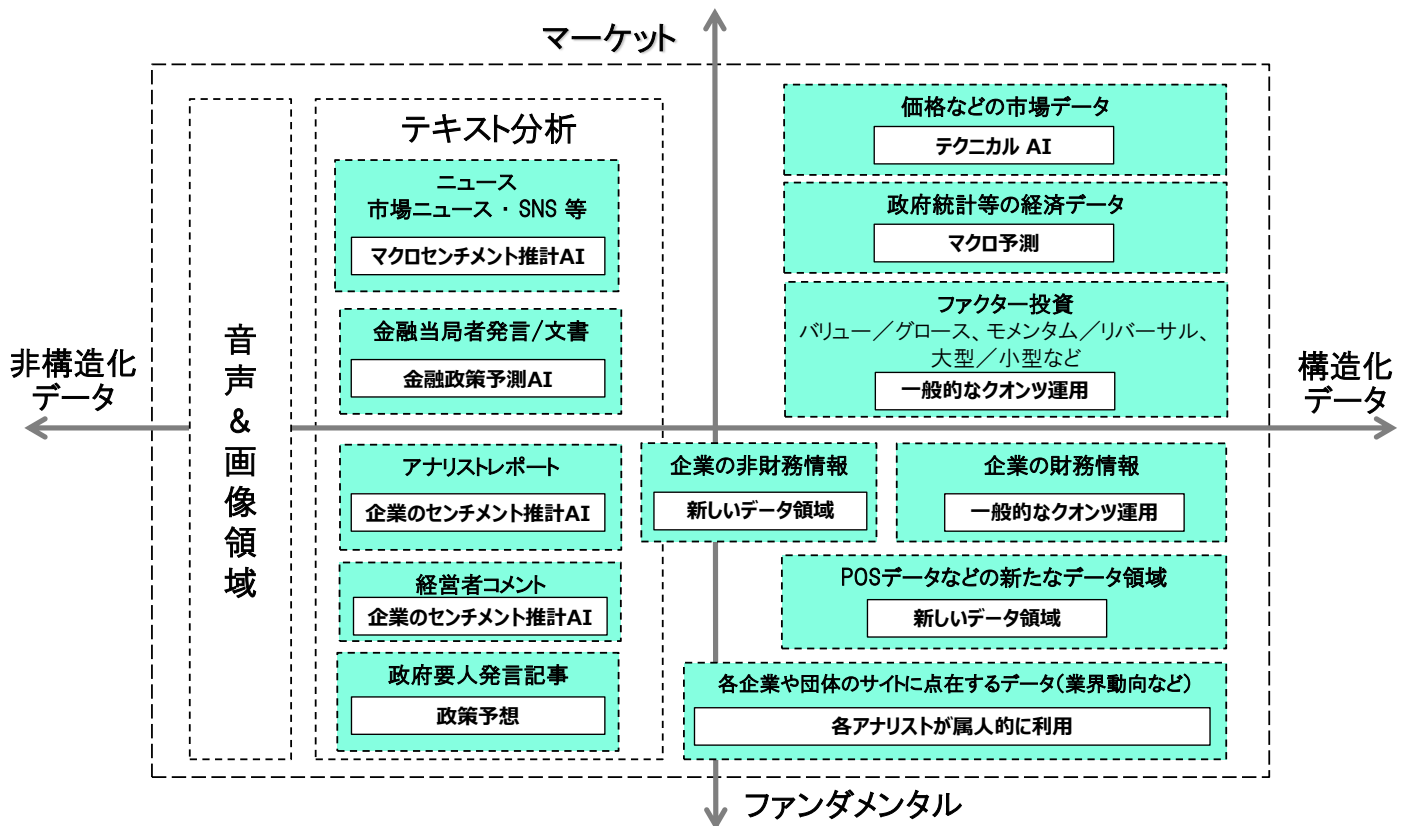
- AI (人工知能) と投資手法の関係性について。
- データの特性や構造を理解した上で、AIに何をさせたいか？を考えることが重要。

最近、AI (人工知能) に関連するニュースが増えています。投資の分野でも研究開発が盛んに行われており、実際に投資手法として利用可能な段階まで進展しています。本レポートでは、AI (人工知能、以下AI) と投資手法の関係性をご紹介します。

#### 1. データの特性とAIモデルの関係性

AIを投資手法に導入すると言っても、いきなりではイメージが湧きづらいと思います。まず考えなければならない事は、どのデータを使って、どのような結果を目指すのか、という大まかな設計を考える事になります。図1は、データの特性を切り口に、データ領域 (どのデータを使うか) とAIモデル (どのような結果を目指すか) をマッピングしてみたものになります。

図1. 各情報の整理学、AIによるモデル化のイメージ



※マップ上の各項目とも複数の領域に跨る可能性があるが、分類案の一例として図示

## データの特性と構造

### 1-1. マーケットデータとファンダメンタルデータ

図1ですが、縦軸は、データの特性がマーケット寄りか、それともファンダメンタル寄りか、を表しています。マーケット寄りのデータとは、株価情報や出来高など、一般に市場から取得されるデータになります。対してファンダメンタル寄りのデータとは、企業の財務状況など、その企業の状態を表すデータになります。

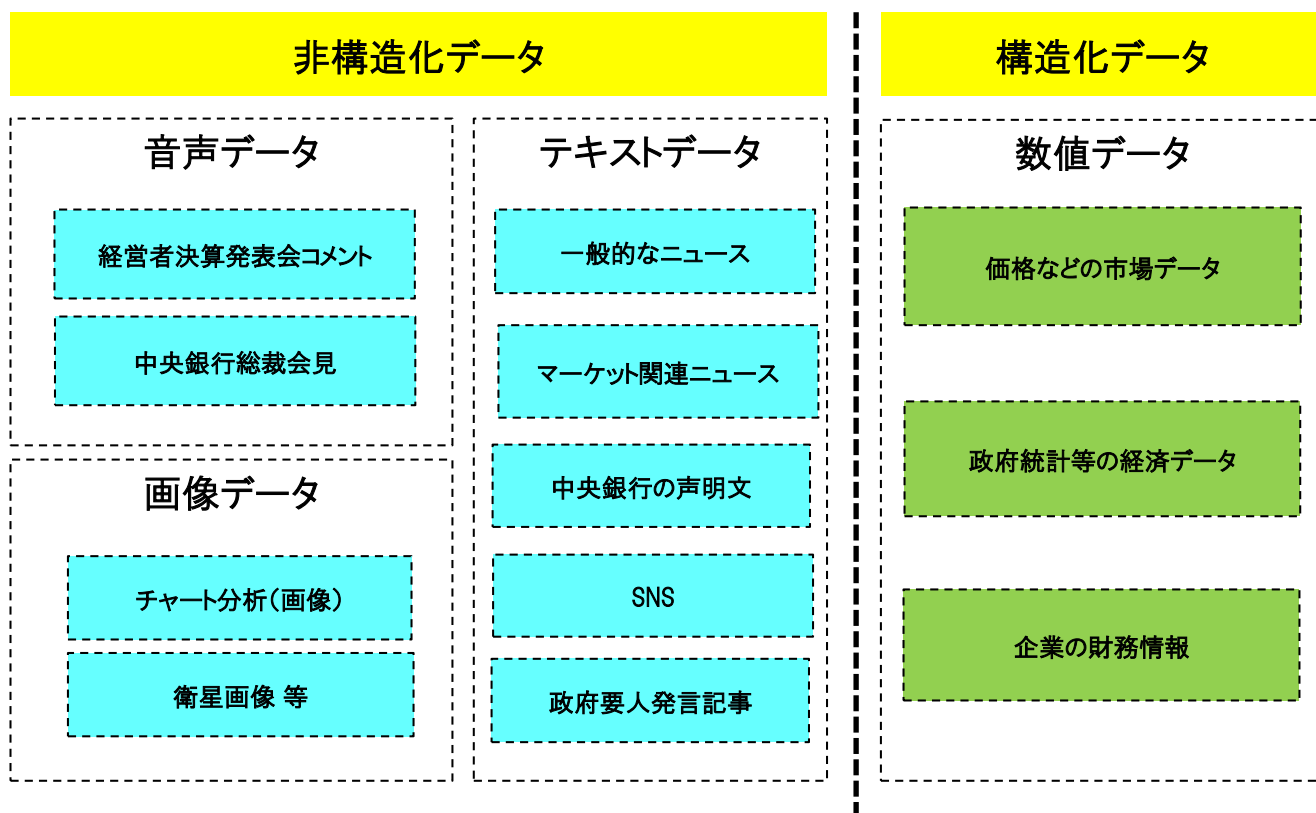
### 1-2. 構造化データと非構造化データ

一方、横軸はデータの構造を表しています。あまり聞きなれないかもしれませんが、そのデータが構造化データか、非構造化データか、を切り口にしています。構造化、非構造化データの定義は様々ですが、本レポートでは、株価のような最初から扱いやすい数字データになっているものは構造化データとし、テキストや画像、音声データは非構造化データ、と括っています。

### 1-3. AIを用いた非構造化データの活用

構造化データはクオント運用をはじめとして、今までも一般的に利用されてきました。一方、非構造化データはそのままでは活用しづらい点がネックとなっていました。しかし、昨今のAIにおける発展として、今まで扱いつらかった非構造化データの利用が進んでおり、扱えるデータの裾野が広がってきています。この点も投資手法の開発にAIを使う事例が増え始めた理由と言えるかと思います。

図2. 構造化データと非構造化データに見る該当データの分布



## なぜAIを使うのか

### 2. AIに何ができるか

さて、マッピングされた図1に話を戻します。この図は、それぞれのデータを使った場合、どのような投資判断用のAIモデルが作れそうか、を表現しています。例えば、右上のマーケットデータかつ構造化データである価格情報を使うと、株価の動きをチャート等から予測するテクニカルモデルが出来るかもしれません。一方、非構造化かつファンダメンタルの情報である経営者のコメントを使えば、経営者（企業）のセンチメントを推計するAIモデルが出来るかもしれません。

#### 2-1. ボトムアップアプローチとトップダウンアプローチ

このように、まずは世界にどのようなデータがあるか、そのデータの特性を確認してAIモデルを組み立てる場合があります。これはボトムアップ方式と言えそうです。逆に、特定の目的のためのAIモデルを作りたい、そのためにはこのようなデータが必要だ、といった開発のアプローチも勿論ありえます。こちらは対照的にトップダウン方式になるかと思えます。このように、いざAIを使って投資手法を作ろうとしても、その領域は多岐に渡ります。今後のレポートでは、図1のいくつかの領域において、どのようなアプローチをしていくのか、具体例を交えて紹介していく予定です。

### 3. テキスト情報を用いた投資手法の開発

今回は、テキスト情報の利用についてご紹介したいと思います。先述の通り、最近のAIにおける発展として、今まで扱いつらかった非構造化データの利用が進んだ点があります。この構造化データの大きな部分を占めるテキストデータの活用方法ですが、どのようにデータを投資手法の開発まで繋げるのか、というフローを考える必要があります。様々なアプローチがありますが、一例としては図3の手順が考えられます。

#### 3-1. テキスト情報をAIに解析させる意味

テキスト情報の利用について、簡単な例を挙げたいと思います。日々大量のマーケット関連ニュースが投資家に読まれています。もし景気に関するニュースがポジティブ（ネガティブ）だった場合、日本株が上昇（下落）しやすい、と言ったような関係があることに、誰かが気づいたとします。しかし、その気づいた人も体調等の影響から、毎回、同じ基準でこの文章がどの程度ポジティブか、ネガティブかを判定することが難しいことがあるかもしれません。あるいは、大量のニュースが作成される中で、全てを読んで一つずつ判断するのは難しいかもしれません。上記のような場合に、AIが活躍するチャンスが生まれます。学習済みのAIに、文章がどれくらいポジティブか、ネガティブかを判定させればいいのです。機械なので体調は関係なく、毎回一定のルールで判断できます。また、AIは昼夜休まずニュースを読み続けられるので、大量のニュースも逐次読んで行くことができます。このように、今まで扱いつらかった非構造化データであるテキスト情報を利用した投資手法の研究開発が実際に行われています。

図3. テキストデータの活用アプローチ（例）

- ①：テキストデータを取得
- ②：テキストデータを綺麗な状態に整形（クレンジング、分かち書き）
- ③：AIが読み取れるようにデータを加工（数値情報に変換）
- ④：単語 または 文章単位でスコア化（AI：機械学習）
- ⑤：算出したスコアと投資対象との関連性を確認（スコアとTOPIXとの関係等）
- ⑥：実際に投資してリターン獲得

## 超過収益獲得への挑戦

### 3-2. 超過収益獲得への挑戦

ここまでで、AI（機械学習）におけるテキスト情報の利用に興味を持っていただけましたでしょうか？今まで扱われていた通常の構造化データとは違い、まだまだ定まった手法は存在しない分野だと思いますが、それだけに超過収益を獲得する余地があると考えています。

### 4. 投資工学開発室での取り組み

投資工学開発室では投資判断用のAIモデル作成や、AIを利用した有効な投資情報の提供について、日々、研究開発を実施しています。今回ご紹介した、テキスト情報を利用したAIモデルについても取り組んでいます。図4は、株式市場のセンチメント把握を目的とし、AI手法を用いて作成した指数値の一例です。例えば、この指標で株価の方向感を捉え、そもそも株を買うべきか否か、株を買うにしてもどの業種を買えば良いのか、これらの判断材料にならないかを検証しています。このように、様々なデータとAI手法を組み合わせることで、経済や市場の動きを捉える試みが、世界中で研究されています。

### 5. 今後のレポートについて

本レポートでは、今後もAIによる投資手法への応用に関する記事を発信していく予定です。今後は投資手法に応用するための具体的な作業、検証ポイントもご紹介できたらと考えています。

図4. AI手法を利用した株式市場のセンチメント把握を目的とした指数値例



#### ～執筆者の紹介～



吉野貴晶 (写真：左)

「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。ビックデータや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。

高野幸太 (写真：右)

ニッセイアセット入社後、ファンドのリスク管理、マクロリサーチ及びアセットアロケーション業務に従事。17年4月に投資工学開発室に異動後は、主に計量的手法やAIを応用した新たな投資戦略の開発を担当する。

