

※本ニュースリリースは株式会社NTTデータと株式会社NTTデータ数理システムが共同で配信しています。重複して配信されることがありますが、ご了承願います。

2014年3月7日

## Twitter データを用いた金融マーケット向け「Twitter センチメント指標」を開発 ～NTTデータとNTTデータ数理システムが共同で Twitter 全量データを用いた金融市場センチメント分析を実施～

株式会社NTTデータ  
株式会社NTTデータ数理システム

株式会社NTTデータ(本社:東京都江東区、代表取締役社長:岩本 敏男、以下NTTデータ)と株式会社NTTデータ数理システム(本社:東京都新宿区、代表取締役社長:山本 二郎、以下NTTデータ数理システム)は、共同で Twitter データを用いた金融マーケット向け「Twitter センチメント指標」を開発しました。本指標は、Twitter ユーザーの感情を定量化し、株式指標との関係性を分析したもので、実際の金融取引でも有効な判断指標となる可能性があることを実証しました。

### 【背景】

NTTデータとNTTデータ数理システムの両社は Twitter データのさらなる用途拡大に向けて金融分野でのリアルタイム活用の検討を行ってきました。

両社はNTTデータの持つ「Twitter 全量データ」および「日本語解析技術」とNTTデータ数理システムが得意とする「データ解析技術」を活用し、Twitter データを用いた金融マーケット向け「Twitter センチメント指標」の開発とその検証を実施し、本指標が実際の金融取引における有効な判断指標となる可能性があることを実証しました。

### 【概要】

実証にあたっては分単位から週単位までさまざまなパターンで数値化した「Twitter センチメント指標」を算出しました。また、検証では Intel 社の協力のもと 2011 年 1 月から 2013 年 11 月までの合計 35 か月間の日本語の「Twitter 全量データ」から株式に関する Tweet を抽出。それに対しセンチメント分析<sup>(注1)</sup>を行い「Twitter センチメント指標」と金融指標値との関連性を評価しました。

その結果、いくつかの「Twitter センチメント指標」が「日経平均 VI」<sup>(注2)</sup>に対し統計的に有意な関係性を持ち、実際の金融取引でも有効な判断指標となる可能性がある事を実証しました。

### 【特長(検証の概要)】

#### ① データ抽出方法について

NTTデータが持つ Twitter の全量データから、独自の辞書に基づき株式市場に関連するツイートを抽出しました。抽出した株式関連ツイートにNTTデータ独自の日本語解析技術を適用し、得られた市場センチメントから「Twitter センチメント指標」を構築しました。

#### ② 検証に用いた指標について

金融指標値として株式会社日本経済新聞社が算出する日経平均 VI を採用しました。日経平均 VI は投資家が将来の日経平均株価の変動の大きさをどのように想定しているのかを表した指数です。

#### ③ 経済指標との相関分析について

データ分析技術を保有するNTTデータ数理システムが「Twitter センチメント指標」と「日経平均 VI」の関連性を評価し、以下の結果が得られました。

- Twitter センチメント指標と日経平均 VI に統計的に有意な相関を確認
- Twitter センチメント指標を用いることで日経平均 VI の予測精度が統計的に有意に向上

- Twitter センチメント指標を取引モデルに組み込むことで投資成績が改善

なお、検証にあたっては Intel 社のエンタープライズ向け最新モデル インテル® Xeon® プロセッサーE7 v2 製品ファミリー搭載機を用いました。

#### 【今後について】

今後、両社は「指標値提供サービス」の商用化に向けて、分析精度の向上と合わせて膨大な Twitter データから高速に指標値計算をおこなえるようアルゴリズムを開発していきます。さらに、NTTが開発したオープンソースのリアルタイム機械学習フレームワーク「Jubatus®」<sup>(注 3)</sup>等の技術を活用し、リアルタイムでの指標値提供がおこなえるよう技術開発を行っていきます。

また、両社は「指標値提供サービス」だけでなく、さまざまな金融機関の個別ニーズに基づいた以下のような金融マーケット向け「Twitter 指標」を提供する予定です。

#### 「今後サービス予定の金融マーケット向け「Twitter センチメント指標」の特長」

##### ① 日本語全量 Twitter データを利用

日本語全量 Twitter データを用いることで取り漏らすことなく精緻にセンチメントを捉えられることが期待されます。

##### ② 日本語解析技術により市場センチメントを算出

NTTデータ独自の日本語解析技術を活用し、株式関連ツイートから株式市場のセンチメントを定量化します。

##### ③ デイリーで算出した指標値の提供を予定

今後、日本語解析アルゴリズムの高速化を行い週次のみならず日次での提供を予定しています。

#### ■インテル株式会社からのエンドースメント

「ビジネス改革を推進するエンタープライズの新たな基盤となるインテルの最新プロセッサー、インテル Xeon プロセッサー E7 v2 製品ファミリーは、NTTデータとNTTデータ数理システムが共同で開発した金融マーケット向け「Twitter センチメント指標」について、大量データに対する検証において優れたコンピューティング環境を提供しました。今日ビッグデータを効率的に分析し、活用する為には、先進のサーバー基盤が不可欠となっています。このたびの両社の取り組みに協力できたことを大変光栄に思います。」

インテル株式会社 ビジネスデベロップメント プラットフォーム事業開発 ディレクター 友真 衛

#### ■NTTソフトウェアイノベーションセンタからのエンドースメント

NTTソフトウェアイノベーションセンタ(SIC)は、この度のNTTデータとNTTデータ数理システムによる「Twitter センチメント指標」の開発について心より歓迎いたします。

インターネット上の大量データや多種多様なデバイスから刻一刻と発生するセンサーデータをリアルタイムに分析することが、ビジネスを推進する上で、ますます重要になってきております。この分野において、Jubatus をコアテクノロジーとして活用いただく事例も増えてまいりました。

今後の「指標値提供サービス」の商用化においても、Jubatus が活用されるべく、支援をしてまいりたいと思います。

SIC では、今回の両社の取り組み等を通じて、世の中に新しい価値を提供すべくこれからも、Jubatus をはじめとしたビッグデータ関連の R&D を推進してまいります。

NTT ソフトウェアイノベーションセンタ

所長 藤田 敏昭

(注 1) センチメント分析とは、市場参加者のマーケットに対する強気や弱気などの市場心理を調査してそれを元に行なう相場の分析のことです。

(注 2) 日経平均VIとは、市場が期待する日経平均株価の将来1か月間の変動の大きさ(ボラティリティ)を表す数値で、株式収益率の標準偏差のようにパーセン

ト単位で表示され、計算には大阪証券取引所の日経 225 オプション価格が用いられ株式会社日本経済新聞社によりリアルタイムで算出・公表されています。

(注3) Jubatusとは、NTTソフトウェアイノベーションセンタと株式会社Preferred Infrastructureが共同開発しオープンソースソフトウェアとして公開しているビッ  
グ

データ向けのリアルタイム分散機械学習基盤です。

\*Jubatus は日本電信電話株式会社の登録商標です。

\*Intel、インテル、Xeon は、米国および他の国におけるインテルコーポレーションの商標です。

\*その他の商品名、会社名、団体名は、各社の商標または登録商標です。

**【本件に関するお問い合わせ先】**

■ 製品・サービスに関するお問い合わせ先

株式会社NTTデータ

第三法人事業本部 ソーシャルビジネス推進担当

富田、佐藤

Tel:050-5546-9092

■ 検証内容に関するお問い合わせ先

株式会社NTTデータ数理システム

ビジネスインテリジェンス推進センタ

中川、鍋谷

[TEL:03-3358-1809\(内線:2222\)](tel:03-3358-1809)