

暗号通貨訓練のディープラーニング

この記事は、インテル® デベロッパー・ゾーンに公開されている「[Deep Learning for Cryptocurrency Trading](#)」の日本語参考訳です。

ディープラーニングの新しい潜在的なユースケースの 1 つに、暗号通貨トレーダーの心理検出の開発があります。私は現在、再帰型ニューラル・テンソル・ネットワーク (RNTN) を利用して、ニュースのヘッドライン、Reddit のポスト、Twitter のポストから全体的なトレーダーの心理に対する深い情報を提供する心理アナライザーを開発しています。トレーダーの心理は暗号通貨の価格変動を判断する際に重要な要因です。この記事では、暗号通貨トレーダーの心理分析の利点と RNTN の長所を説明します。このプロジェクトの長期的なビジョンは、トレーダーの心理を考慮して取引の意思決定を行うだけでなく、鞘取りのような機会を活用して資産の売買により価格差から利益を上げる人工知能 (AI) 暗号通貨トレーディング・ボットを開発することです。

アプリケーションの技術的なコンポーネントを理解するには、暗号通貨とブロックチェーン・アプリケーションの現在の状況を理解することが重要です。この記事には、暗号通貨の価値と重要な特性を引き出す方法に関する情報が含まれます。

暗号通貨は新興通貨でありデジタル資産としての交換媒体の役割を果たします。デジタル通貨およびデジタル資産はブロックチェーンの上に設計されています。暗号通貨の暗号的性質は、暗号通貨トークンの作成および複製を防ぐことです。暗号通貨の供給量は有限であり、供給量が有限で時間とともに需要が増加して価格が上昇すると投資家が考えることから、暗号通貨の価値の一部は得られています。



図 1: 暗号通貨

ブロックチェーンは、分散型の台帳内にトランザクションを永続的に記録します。ブロックチェーンを利用すると、中央機関なしで、あらゆる種類のトランザクション (通貨/データ/サービスの交換) を記録および実行できます。

トランザクションはネットワーク上のノード (マシン) により自動的に記録および追跡されるため、ブロックチェーン・プラットフォームは高度な安全性を備えています。トランザクションはネットワーク上の多くのノードにより検証された台帳でのみ処理および記録されます。トランザクションが発生すると、台帳の日付のハッシュ値およびタイムスタンプとともにトランザクションが台帳に記録されます。情報を操作または変更するには分散ネットワークの多くのノードのレコードを変更する必要があるため、情報の操作や変更はまず不可能です。ブロックチェーン上で行われるトランザクションは、高速、低コスト、高いセキュリティー、エラーの少なさ、攻撃や障害の中心ポイントの排除、など多くの利点があります。

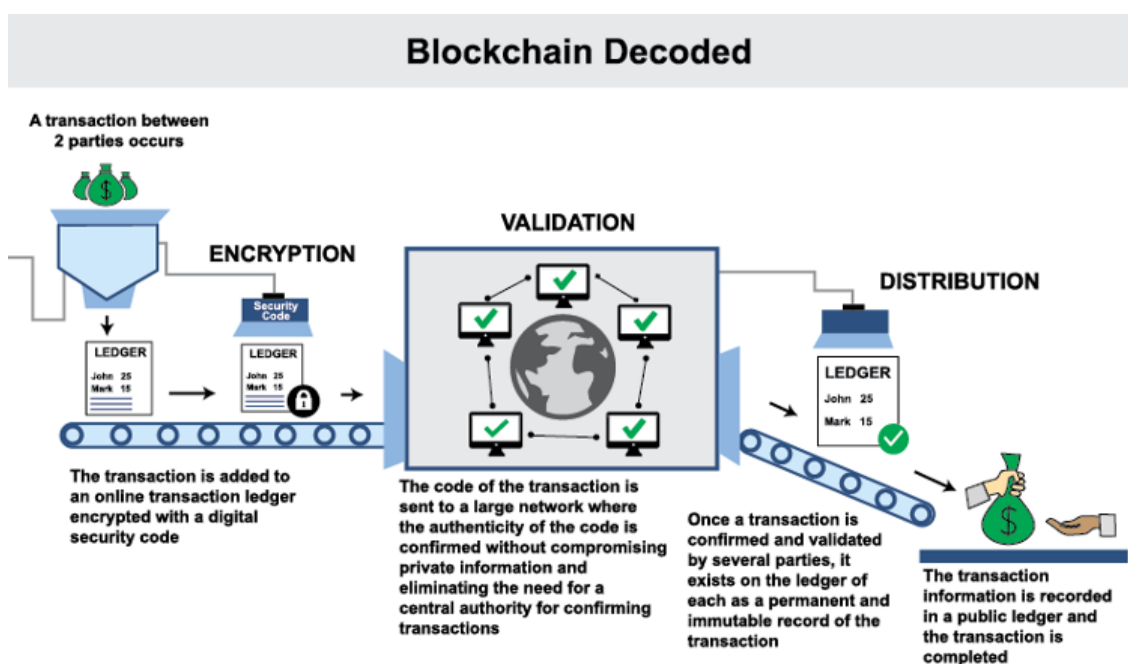


図 2: ブロックチェーン

ブロックチェーンに記録された情報とトランザクションの不変性により、スマート・コントラクトをプログラムとして実行できます。スマート・コントラクトは、特定の基準が満たされたときにトランザクションを促進、検証、実行するコンピューター・プロトコルです。スマート・コントラクトを利用すると、ほかの契約とやりとりし、意思決定を行い、データを格納して、ユーザー間で通貨を転送できます。スマート・コントラクトは、ダウンタイム、検閲、不正行為、妨害なしで、プログラムされたとおりに実行できます。スマート・コントラクトは、ユーザー間で何らかのトランザクションが必要とされるすべての産業、ビジネス、システムのビジネスプロセスを改善するために使用できます。

ブロックチェーンは、トラストプロトコルと呼ばれます。分散型台帳というブロックチェーンの独創的な設計とスマート・コントラクトにより、信頼性が求められるトランザクションを実行することが可能になります。ブロッ

ブロックチェーンの投資家は、社会に存在する従来のトランザクション・プラットフォームを分散型のプラットフォームに置き換え可能であると信じています。

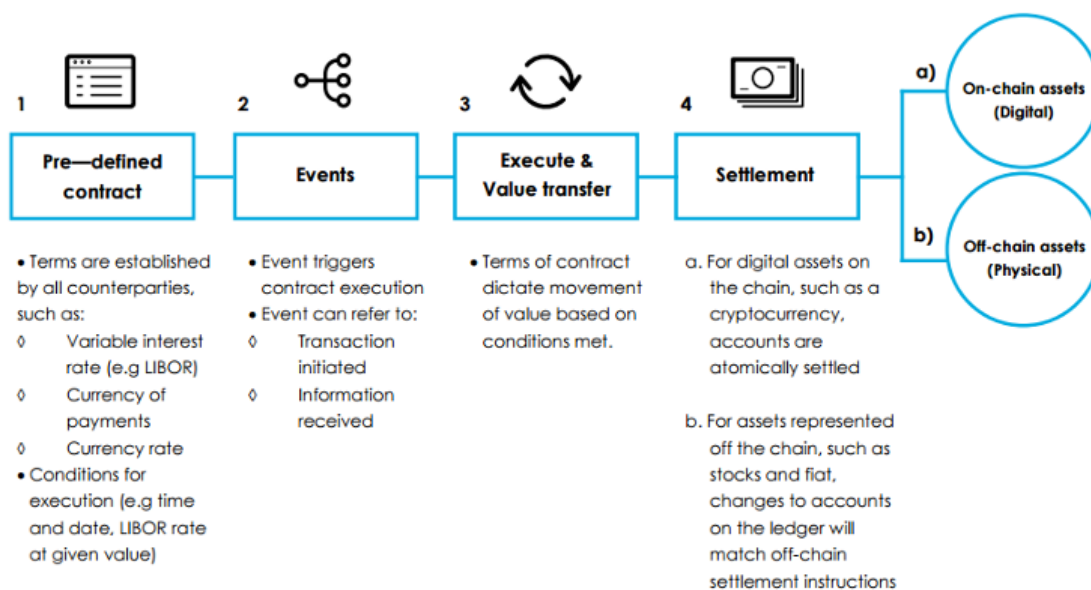


図 3: スマート・コントラクト

多くの革新的な暗号通貨がありますが、中でも有名で長く続いているものは、Bitcoin (ビットコイン)、Litecoin (ライトコイン)、Ethereum (イーサリアム)、Golem (ゴーレム)、Siacoin (シアコイン) です。ビットコインとライトコインは、どちらもピア・ツー・ピア型の支払いプロトコルです。イーサリアムは、スマート・コントラクトを実行できる分散型アプリケーション (DApps) を開発するためのプラットフォームです。ゴーレムは、AI アプリケーションのテストや高い計算能力が必要なタスクに利用できる、分散型スーパーコンピューター・ネットワークです。シアコインは、分散型クラウド・ストレージ・ネットワークです。

これらの分散型アプリケーションでは、暗号通貨が各ネットワークを利用するエントリーポイントとしての役割を果たします。イーサリアムでプログラムされたアプリケーションは、ネットワークへの支払いとしてイーサリアムが提示されない限り実行できません。ゴーレムのスーパーコンピューター・ネットワークは、ゴーレム暗号通貨が引き換えに提示されない限り利用できません。シアコインのクラウドストレージは、シアコインが引き換えに提示されない限りネットワークに実装できません。

暗号通貨トークンは、ネットワーク上で動作しているノードに対する報酬や懸賞金として提示されることもあります。例えば、ビットコインのノードのオーナーは、ネットワークを維持するために計算能力を提供していることに対する報酬としてビットコインを受け取ります。ライトコインのノードのオーナーは、ライトコイン・ネットワークの動作を支援していることに対する報酬としてライトコインを受け取ります。イーサリアム・ネットワークも、ネットワークの維持をノードに依存しています。イーサリアムのノードは、イーサリアム・ブロックチェーンの情報を作成、検証、公開、伝播します。イーサリアム・ネットワークのノードとして各自のコンピューターを提供することと引き換えに、ノードのオーナーはイーサリアムを受け取ります。ゴーレムとシアコインも同様に、ネットワークの維持に対する報酬として、ノードのオーナーはゴーレムまたはシアコインをそれぞれ受け取ります。

ゴーレムのオーナーは、ブロックチェーンを活用したスーパーコンピューターを動作させるために余っている計算能力を提供します。シアコインのノードのオーナーは、分散型ストレージ・ブロックチェーンを維持するために余っている計算ストレージを提供します。ノードのオーナーに対する報酬システムは、オーナーがネットワークの維持を続けるインセンティブとして機能しています。ネットワークを維持しているノードは、暗号通貨とその価値の継続的な存在を保証します。ノードのオーナーが多いほど、暗号通貨の価値も高くなります。



図 4: ブロックチェーンのノード

独自の分散型アプリケーションを開始しようとしているチームは、ICO (Initial Coin Offerings) を行うことができます。ICO 中に、分散型アプリケーションを開発したチームは、販売可能なトークンの一部を安価で初期投資家に提示します。また、投資家がプロジェクトで利用しているブロックチェーン・テクノロジーを理解できるように、プラットフォームと暗号通貨のすべての特徴を漏れなく説明した技術論文 (ホワイトペーパー) を公開します。ホワイトペーパーには、ブロックチェーンの利点と、投資家や将来のノードオーナーのコミュニティに提供される経済的なインセンティブがリストされます。チームとプロジェクトを信用した投資家はトークンを購入します。ICO の参加者の投資額により、暗号通貨の初期価格が決まります。ICO を行ったチームは、調達された資金をプロジェクトの開発およびプロジェクトに関連するほかの費用に利用します。

プロジェクトの開発が続行されると、さらに多くの投資家がプロジェクトに参加し、投資と引き換えにトークンを受け取ります。人々は、暗号通貨を受け取る報酬として、ネットワーク上のノードを実行します。ノードオーナーの数が多いほどネットワークは強力になります。プロジェクトがすべてうまく行った場合、その暗号通貨は、非常に高速で、コスト効率の高い、ほかのトランザクション・プラットフォームよりも優れたピア・ツー・ピア型のトランザクション・プラットフォームになります。暗号通貨の開始時点では、現在 100 万ドル、10 億ドル、あるいは 1 兆ドルのブロックチェーン・プラットフォームで利用されている暗号通貨の開始時点と比較して妥当な価格が設定されます。ブロックチェーン・プラットフォームがその規模に成長するかどうかは、時間が経過しないと分かりません。次の図は、この記事で紹介した暗号通貨の成長率と価格を示したものです。

	1 week	1 month	3 month	6 months	1 year	2 year
Bitcoin Growth Rate	32%	78%	157%	385%	803%	2235%
Ethereum Growth Rate	11%	49%	623%	589%	2776%	56134%
Litecoin Growth Rate	26%	43%	57%	542%	1571%	1945%
Golem Growth Rate	-4%	1%	-18%	179%	1333%	N/A
Siacoin Growth Rate	-6%	8%	-43%	411%	892%	10418%

	Price as of 10/14/2017 in \$	Price 1 week	Price 1 month	Price 3 months	Price 6 months	Price 1 year ago	Price 2 years ago
Bitcoin	5735.05	4350.51	3226.41	2232.65	1183.44	635.08	245.59
Ethereum	344.29	308.92	230.84	47.65	49.97	11.97	0.61224
Litecoin	65.02	51.8	45.59	41.53	10.13	3.89	3.18
Golem	0.215729	0.225016	0.2142176	0.26199	0.077259	0.015051	0
Siacoin	0.004207	0.004483	0.003902	0.00735	0.000824	0.000424	0.00004

図 5: 暗号通貨の成長率

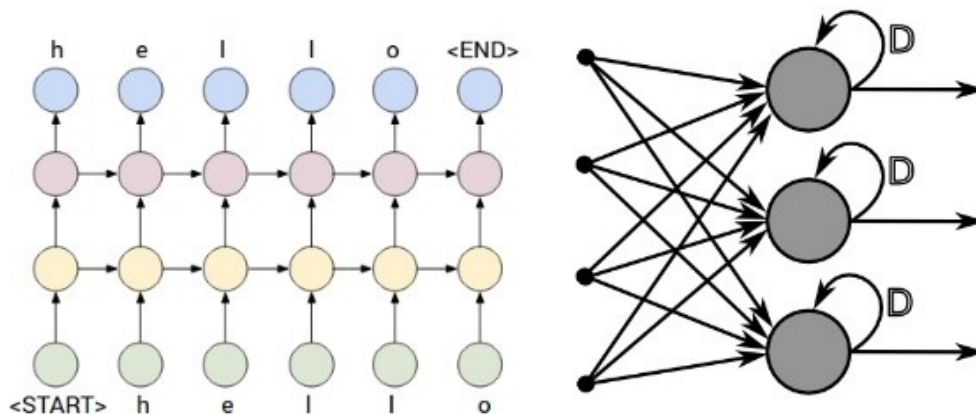
これらの成長率は、大きな利益が得られることを示しています。重要な問題は、最新の予測テクノロジーをどのように利用すれば価格変動を予測できるかということです。価格変動を予測する標準化された手法はありません。多くのブロックチェーン・プラットフォームは、トランザクション・プラットフォームの既存の非効率性に対処する技術的なソリューションを提供することを目標としています。プロジェクトの展望を測定する基本的な方法は、プロジェクトのホワイトペーパーを徹底的に評価することです。ホワイトペーパーは、提案されているブロックチェーンの技術的側面とプロジェクトの全体的な規模についての優れた観点を提供します。しかし、価格追跡に標準的な手法はありません。株式市場と比較して観察されるファンダメンタルもなく、評価の基となる四半期レポートもありません。テクニカル分析は、取引で最良のスプレッドを達成するために役立つ手法であり、鞘取りの状況を活用するためにも使用されます。テクニカル分析は価格変動の予測にも利用できますが、この記事では（ディープレニングを効率的に適用できる）市場心理に焦点を当てます。

市場心理は、暗号通貨の価格変動を検出する重要な要素です。暗号通貨およびそのブロックチェーン・プロトコルが初期の開発段階を過ぎると、大きな事件、政治的な出来事、ネットワークのアップグレード、会議、パートナーシップ、Segwit、トレーダーの心理などが暗号通貨の価格変動を引き起こす傾向があります。これらの暗号通貨の価格変動を測定する 1 つの方法は、自然言語処理 (NLP) のサブセットである心理分析を利用することです。これらの暗号通貨の価格変動の多くは、集団本能により決定されます。行動ファイナンス理論における集団本能とは、個々の意思決定の欠如により、人々が周りの大多数の人と同じ方法で考えて行動するという心理的な特徴のことです。価格変動は市場心理および暗号通貨を取り巻くコミュニティの見解に基づく傾向があります。これらの理由に基づき、ニュースのヘッドライン、Reddit のポスト、Twitter のポストの心理分析が暗号通貨の価格変動の動向を示す最良の指標であると信じられています。

現時点で有名な心理分析の手法は、再帰型ニューラル・ネットワーク (RNN) を使用することです。RNN は指定されたテキストを解析して単語をトークンにします。単語の頻度が識別され、単語表現のバッグが生成されます。次に、各単語の主観性が既存の辞書から検索されます。辞書は研究者により事前に記録された心理的な値のデータベースです。辞書に基づいてテキストが分類され、全体的な心理が計算されます。RNN は長いテキストは問題なく処理できますが、ニュースのヘッドライン、Reddit のポスト、Twitter のポストのような短いテキストの心理分析には役立ちません。RNN は、合成性（語順）を考慮できないため、言語学のセマンティクスを考慮できず、心理変化の識別および否定の範囲の理解には有用でないためです。

What are Recurrent Neural Networks

1. In their simplest form (RNNs), they are just Neural Networks with a feedback loop
2. The previous time step's hidden layer and final outputs are fed back into the network as part of the input to the next time step's hidden layers.



@QuantixResearch

図 6: 循環ニューラル・ネットワーク

RNTN はテキストの意味的な合成性を考慮できるため、この種のプロジェクトに非常に適しています。ツイートのような短いテキストに取り組む場合、心理を判断する情報が少ないため、テキストの合成性を検出できることが非常に重要になります。

前述したように、RNTN は構文的な順序を考慮するという点で非常に優れています。RNTN は、親 (根) グループ、子 (葉) グループ、スコアを含む、複数の部分で構成されます。葉グループは入力を受け取り、根グループは分類器を使用してクラスとスコアを決定します。

データが心理アナライザーに渡されると、データは二分木で解析されます。特定のベクトル表現がすべての単語から形成され、葉として表されます。下から上へ、ベクトルは最適化するパラメーターおよびソフトマックス分類器の特徴入力として使用されます。ベクトルはスコアとともに 5 つのクラスに分類されます。

次のステップは再帰が発生する場所です。2 つの単語間の類似性がコード化されると、2 つのベクトルは次の根に移動します。スコアとクラスが出力されます。スコアは解析のポジティブ性またはネガティブ性を表し、クラスは現在の解析の構造をコード化します。最初の葉グループは解析を受け取り、2 目目の葉は次の単語を受け取ります。3 つの単語をすべて含む解析のスコアが出力され、次の根グループに移動します。再帰プロセスは、すべての単一の単語を含むすべての入力を使用されるまで続けられます。実際のアプリケーションでは、RNTN はより複雑になります。RNTN は、次の葉グループの文の次の 1 つの単語を使用するのではなく、次のすべての

単語を使用します。つまり、解析全体のベクトルを確認します。再帰プロセスのすべてのステップでこの処理を実行することで、RNTN は構文解析のすべてのスコアを分析しています。

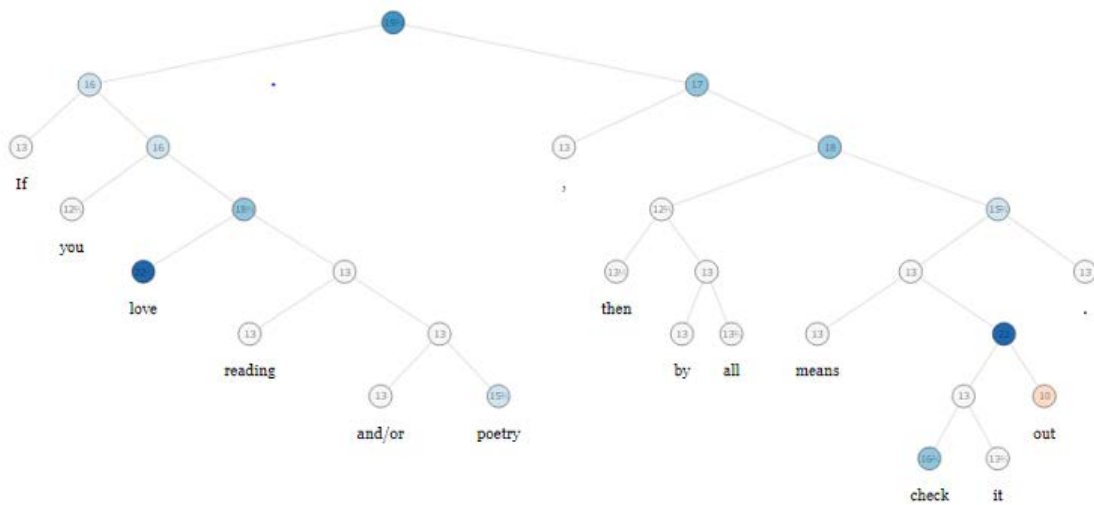


図 7: ラベル付けされた心理ステートメント

RNTN は、各再帰プロセスで最良のサブ構造を選択するため、根グループにより生成されたスコア値を使用します。最終的な構造が決定されたら、元に戻り、文の文法的構造が分かるようにデータにラベルを付けます。RNTN はバックプロパゲーションにより訓練されます。訓練は、予測された文構造と (ラベル付けされた訓練データのセットから取得された) 適切な文構造を比較して行われます。データが訓練されると、訓練に使用したデータを解析する RNTN の優れた可能性が明らかになります。再帰型ニューラル・ネットワークと RNTN の 1 つの大きな違いは、RNTN では同じ合成関数がテンソルとして使用されるため、学習するパラメーターが少ないことです。この違いにより、同様の単語は同様のベクトルになるように、同様の単語は同様の合成的な動作になります。

使用されるリソースは、多くの NLP ツールを含む [StanfordCoreNLP](#) (英語) です。心理アナライザーは [Stanford Sentiment Treebank](#) (英語) も使用します。Stanford Sentiment Treebank は、言語における心理の合成效果の完全な分析を考慮して、解析木のすべてにラベルを付けた大規模なデータのコーパスです。再帰プロセス中に、RNTN はこのデータセットを参照して指定された解析のクラスとスコアを決定します。Stanford Sentiment Treebank には、3 人の研究者により注釈が付けられた解析木の 10,662 の文から合計 215,154 の固有のフレーズが含まれています。Stanford Sentiment Treebank のデータセットは [rottentomatoes.com](#) (英語) の映画レビューに基づいています。

プロジェクト中に直面する課題の 1 つは、Stanford Sentiment Treebank を使用した RNTN モデルには暗号通貨の心理を判断できる十分なデータが含まれないことです。Stanford Sentiment Treebank は映画のレビューデータに基づいているため、暗号通貨に関連する新しいすべての用語に対する心理を認識できません。例えば、誰かが "Ethereum is going to the moon this weekend" (イーサリアムはこの週末に急上昇している) とポストした場合、この問題が発生します。これは人々が暗号通貨の価格が急上昇している、あるいはまも

なく急上昇すると考えるときに使用される一般的なフレーズです。これはとてもポジティブなフレーズの例ですが、心理アナライザーは暗号通貨のコンテキスト/合成性における "going to the moon" のポジティブな意味を理解できないため、このフレーズをニュートラルとして分類してしまいます。

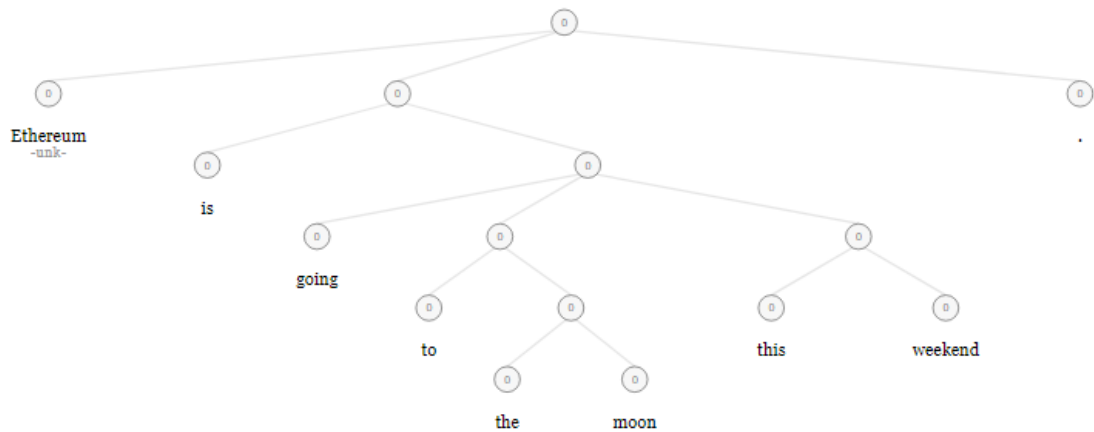


図 8: 文で未検出の心理

この問題に対する回避策は、暗号通貨のツイートや Reddit のポストのデータセットにラベルおよび注釈を付けて、暗号通貨の価格における心理を検出できることを RNTN に教えることです。次の図は、この方法でラベルを付けた同じ文の例です。心理スコアはさまざまで、5 つのレベル (とてもポジティブ、ポジティブ、ニュートラル、ネガティブ、とてもネガティブ) があります。"moon" は通常ポジティブな意味があるため、ポジティブに含めました。"To", "The", "going" およびほかの多くの単語は、暗号通貨との関連付けではポジティブまたはネガティブな意味はないため、ニュートラルに含めました。ただし、Stanford Sentiment Treebank は同じ解析で "going" と "moon" を関連付けているため、"going to the moon" の解析結果の値は高くなります。この解析では、"going" の暗号通貨の価格変動の向きがポジティブな価格変動の向きである "moon" と関連付けられているため、とてもポジティブ (濃い青色) になっています。

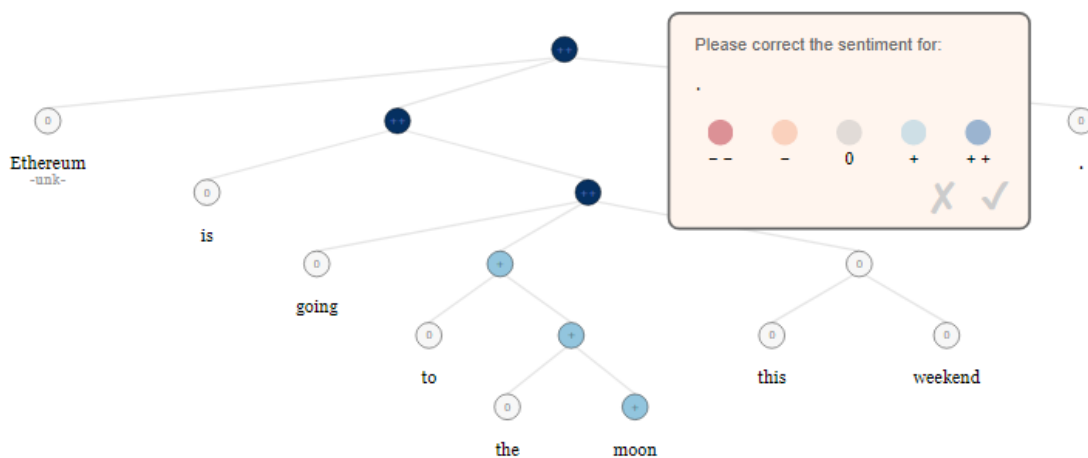


図 9: ブロックチェーン・ステートメントのラベル付けされた心理

暗号通貨の価格変動はトレーダーの心理によって引き起こされるため、ソーシャルメディアのポストやニュースのヘッドラインの心理を検出できれば、価値のある情報を得ることができます。心理アナライザーは、価格変動を検出する鍵となる可能性を秘めています。プロジェクトの長期的なビジョンは、AI 暗号通貨トレーディング・ボットを開発することです。ブロックチェーンとスマート・コントラクトが発展を続けるにつれて、デジタル通貨、デジタル資産、データおよびサービスの交換を含む多くのプロセスが自動化され、ブロックチェーン・ベースのトランザクション・プラットフォームも増加するでしょう。この移行がこの数年間で進むにつれて、ソーシャルメディアのポストやニュースのヘッドラインの心理アナライザーは暗号通貨とその価格変動に関する正確で価値のある情報を提供し、暗号通貨の価値は上昇を続けるでしょう。

画像の出典:

図 1: [Cryptocurrency](#)

図 2: [Blockchains](#) (英語)

図 3: [Smart Contracts](#) (英語)

図 4: [Blockchain Nodes](#) (英語)

図 6: [Recurrent Neural Networks](#) (英語)

図 7: [Stanford Sentiment Treebank](#) (英語)

コンパイラーの最適化に関する詳細は、[最適化に関する注意事項](#)を参照してください。