

## Peut-on prédire les cours des cryptomonnaies ?

En cryptomonnaie, l'analyse technique consiste en l'étude des cours et la reconnaissance de tendances afin de prévoir l'évolution future des marchés. Toutefois, celle-ci est-elle totalement insensible aux chocs exogènes ? La mise en place de méthodes de reconnaissance de formes peuvent nous aider à valider ou non certaines de ces théories.

Dans le cadre de notre étude, la reconnaissance de forme a pour but de prévoir l'évolution d'un cours. De manière générale, celle-ci servirait à prévoir la tendance d'une loi dont les mêmes causes produiraient les mêmes effets, donc peu sensible à des chocs exogènes, et donc de prévenir ses conséquences.

### Positionnement thématique (ETAPE 1)

*INFORMATIQUE (Informatique pratique), INFORMATIQUE (Informatique Théorique), MATHEMATIQUES (Algèbre).*

### Mots-clés (ETAPE 1)

Mots-Clés (en français)	Mots-Clés (en anglais)
<i>Cryptomonnaies</i>	<i>Cryptocurrencies</i>
<i>Reconnaissance de motifs</i>	<i>Pattern recognition</i>
<i>Mesures de similarités</i>	<i>String metric</i>
<i>Analyse technique</i>	<i>Technical analysis</i>
<i>Distance d'édition</i>	<i>Edit distance</i>

### Mots-clés (ETAPE 2)

Mots-Clés (en français)	Mots-Clés (en anglais)
<i>Cryptomonnaies</i>	<i>Cryptocurrencies</i>
<i>Reconnaissance de motifs</i>	<i>Pattern recognition</i>
<i>Mesures de similarités</i>	<i>String metric</i>
<i>Analyse technique</i>	<i>Technical analysis</i>
<i>Trading algorithmique</i>	<i>Algorithmic trading</i>

### Bibliographie commentée

L'analyse technique, née par l'analyse des cours de bourses, est aujourd'hui un pan important de l'étude des marchés des cryptomonnaies. Celle-ci, basée sur l'idée que les mêmes causes produisent les mêmes effets, permettrait de prévoir l'évolution d'un cours de bourse par la simple connaissance de l'historique de ce dernier, reflétant les comportements psychologiques humains qui, en présence d'une certaine situation, tendraient à réagir de manière instinctive et donc prévisible [7].

La mise en place d'un algorithme analysant ces données s'est effectuée en plusieurs étapes.

Tout d'abord, il s'agit de collecter un grand nombre de données relatives aux cours de différentes cryptomonnaies afin de constituer une base de données importante reflétant les évolutions des marchés ces dernières années. La première étape est donc de relever les valeurs de différents actifs au cours des dernières années (par minute), et donc d'implémenter un algorithme les collectant. Le module Python CCXT permet de relever ces valeurs depuis les principales plateformes d'échange pour un grand nombre de devises [5]. En particulier, les données ont été relevées depuis l'interface Binance, permettant de collecter des données plus complètes telles que les valeurs des bougies d'ouvertures et de fermeture, et ainsi d'avoir une analyse plus complète [6]. Ainsi, plus d'une dizaine de millions de valeurs ont été relevées et stockées, formant ainsi la base de données à laquelle les courbes seront comparées.

La deuxième étape consiste à s'approprier l'idée de reconnaissance de formes sur une chaîne de caractères afin d'en cerner les spécificités et les adapter à la problématique [1]. Il s'agit également d'explorer la notion de distance sur un corps fini, deux vecteurs similaires étant ceux dont la distance sera la plus faible [2]. Plusieurs distances intéressantes ont été étudiées.

La distance de Hamming, réputée plus adéquate à la reconnaissance de motifs que d'autres distances naturelles telles que la distance euclidienne [3] semble être aisée à implémenter. Toutefois, celle-ci, prenant en compte la différence de valeurs bit par bit, ne tient pas compte de la structure globale de la chaîne de caractère. La forme de la courbe, prenant notamment en compte les supports et les résistances, est un élément primordial de l'analyse technique [7] donc une analyse trop locale semble être vouée à l'échec.

Ainsi, une autre distance, prenant en compte la forme globale de la chaîne et moins influencée par des perturbations locales (le cours est très volatile donc les petites variations sont très erratiques) doit être implémentée. La distance de Levenshtein, également appelée distance d'édition, puisqu'elle peut supprimer et insérer des éléments (contrairement à la distance de Hamming où les bits sont inamovibles), semble mieux correspondre à la structure d'un cours de bourse [4]. Celle-ci est d'ailleurs utilisée en biologie computationnelle dans la reconnaissance de similitudes entre différents brins d'ADN [4].

Néanmoins, les petites variations (le "bruit blanc") des cours restent une entrave à une analyse trop précise des similitudes des motifs [7]. Une troisième idée serait donc de lisser les courbes afin de, cette fois ci, ne se concentrer que sur la tendance globale des motifs. La mise en place de méthodes de régressions polynomiales telles que la méthode des moindres carrés permet d'avoir une photographie plus globale des tendances et ainsi d'obtenir des comparaisons plus précises [8].

Ainsi, comparer les cours aux courbes lissées de manière polynomiale semble être la comparaison la plus fidèle.

Mais cela suffira-t-il pour prévoir avec précision les tendances des marchés ?

## **Problématique retenue**

Comment implémenter un algorithme d'analyse technique basé sur la reconnaissance de formes à l'aide d'une base de données ? Quelle technique est-elle la plus adaptée à cette fin ? L'étude des cours historiques des cryptomonnaies permet-elle réellement d'en prévoir les tendances futures ?

## Objectifs du TIPE

En adéquation avec le thème de l'année, l'objectif principal de ce TIPE est d'implémenter une méthode permettant de prévenir l'évolution d'une loi dont les mêmes causes produiraient les mêmes effets à partir d'une base de données "historique" et par comparaison avec les données temporelles actuelles.

En particulier, ce TIPE s'oriente sur la reconnaissance des motifs des cours des cryptomonnaies, permettant ainsi également d'affirmer ou d'infirmer des théories économiques telles que, par exemple, la psychologie humaine permettrait de prédire des tendances en raison d'un supposé caractère "aveugle" de l'investisseur, qui serait insensible aux chocs exogènes.

## Références bibliographiques (ETAPE 1)

- [1] J. P. MARQUES DE SÁ : Pattern Recognition: Concepts, Methods, and Applications : *Springer-Verlag Berlin Heidelberg New York, 2001*
- [2] H. SALEHI & R. BURGUEÑO : Structural Health Monitoring from Discrete Binary Data through Pattern Recognition : [https://www.researchgate.net/profile/Hadi-Salehi-4/publication/314245426\\_Structural\\_Health\\_Monitoring\\_from\\_Discrete\\_Binary\\_Data\\_through\\_Pattern\\_Recognition/links/58bdaa6692851c471d5bede1/Structural-Health-Monitoring-from-Discrete-Binary-Data-through-Pattern-Recognition.pdf](https://www.researchgate.net/profile/Hadi-Salehi-4/publication/314245426_Structural_Health_Monitoring_from_Discrete_Binary_Data_through_Pattern_Recognition/links/58bdaa6692851c471d5bede1/Structural-Health-Monitoring-from-Discrete-Binary-Data-through-Pattern-Recognition.pdf)
- [3] MOHAMMAD NOROUZI, DAVID J. FLEET, RUSLAN SALAKHUTDINOV : Hamming Distance Metric Learning : *Departments of Computer Science and Statistics, University of Toronto*
- [4] GONZALO NAVARRO : A Guided Tour To Approximate String Matching : *Dept. of Computer Sciences, University of Chile*
- [5] CCXT Github : <https://docs.ccxt.com/en/latest/index.html>
- [6] PART TIME LARRY : CCXT - Cryptocurrency Exchange Trading Library Tutorial : <https://www.youtube.com/watch?v=2Zdm2ISdm1Q>
- [7] STEVEN B. ACHELIS : Technical Analysis from A to Z : *Equis International, 2003*
- [8] EVA OSTERTAGOVÁ : Modelling using polynomial regression : *Department of Mathematics and Theoretical Informatics, Technical University of Košice*

## DOT

- [1] *Septembre - octobre : Choix du TIPE, spécialisation sur la reconnaissance de formes, cas pratique des cours de cryptomonnaies.*
- [2] *Novembre - décembre : Documentation sur les algorithmes de reconnaissance de formes, en particulier sur les distances sur les corps finis (distance de Hamming, distance de Levenshtein). Documentation sur les théories économiques de l'analyse technique.*
- [3] *Janvier : Constitution des bases de données grâce au module Python CCXT et à la plateforme*

*d'échange Binance, traitement des données brutes et organisation sous sous formes de listes.*

**[4]** *Février : Implémentation des premiers algorithmes, en particulier les modélisations avec les distances de Hamming et de Levenshtein. Abandon de la distance de Levenshtein, algorithme inapplicable en pratique pour de larges bases de données. Rédaction de la MCOT.*

**[5]** *Mars : Implémentation de la modélisation 8 bits. Étude de la corrélation des cours, restriction de l'étude au seul cours du Bitcoin afin d'alléger le temps d'exécution des algorithmes.*

**[6]** *Avril : Divers essais concernant la structure des cours, implémentation de la troisième modélisation.*

**[7]** *Mai - juin : Derniers essais et amélioration de la troisième modélisation. Conception de la présentation.*