*Tendance de marché issue des réseaux sociaux pour l'amélioration du modèle Generative Adversarial Network dans la prédiction de cours d'une action*

Table des matières

[Introduction 3](#_Toc73709466)

[Motivation 3](#_Toc73709467)

[Problématique 3](#_Toc73709468)

[Approche 3](#_Toc73709469)

[Etat de l'art 4](#_Toc73709470)

[IA pour la finance 4](#_Toc73709471)

[Prédiction de tendance 4](#_Toc73709472)

[Generative Adversarial Network pour la finance 7](#_Toc73709473)

[Analyse des réseaux sociaux pour la finance 9](#_Toc73709474)

[Prédiction du cours de l'action X grâce à un GAN 10](#_Toc73709475)

[Données (Type, Sources, Utilité) 11](#_Toc73709476)

[Méthode (possible fusion avec la section précédente) 11](#_Toc73709477)

[Application 11](#_Toc73709478)

[Analyse des résultats 11](#_Toc73709479)

[Mesure du "sentiment de marché" grâce à l'activité sur Twitter 11](#_Toc73709480)

[Données (Type, Sources, Utilité) 11](#_Toc73709481)

[Méthode 11](#_Toc73709482)

[Application (possible fusion avec la section précédente) 11](#_Toc73709483)

[Analyse des résultats 11](#_Toc73709484)

[Intégration du sentiment de marché au modèle GAN 11](#_Toc73709485)

[Méthode 11](#_Toc73709486)

[Application (possible fusion avec la section précédente) 11](#_Toc73709487)

[Analyse des résultats 11](#_Toc73709488)

# Introduction

## Motivation

Intrigue ?

Le Machine Learning, né dans les années 80, est aujourd’hui largement, et de plus en plus utilisé dans des domaines divers et variés. Son application aux domaine financier date de <date>. Avec l’expansion des volumes de données et l’apparition de méthodes nouvelles, l’intérêt de cette pratique m’est apparu naturellement.

## Problématique

La question qui a par la suite découlé de cette réflexion est la suivante : **Comment utiliser les réseaux sociaux et les réseaux de neurones dans la prédiction de cours d'action ?**

A qui s’adresserait la solution ?

## Approche

* Déduction d’une tendance à partir du lexique employé sur les réseaux sociaux associé à une compagnie cotée en bourse
* Utilisation d’un modèle GAN avec différents indicateurs en entrée (closing prices, volume, etc., tendance issue des réseaux)
* Comparaison avec le modèle GAN classique

GAN ?

Generative adversarial network (GAN) was introduced by Goodfellow et al. [13], where images patches are generated from random noise using two networks trained simultaneously. Specifically, in GAN a discriminative net 𝐷 learns to distinguish whether a given data instance is real or not, and a generative net 𝐺 learns to confuse 𝐷 by generating high quality data. (Although this approach has been successful and applied to a wide range of fields, such as image inpainting, semantic segmentation, and video prediction, as far as we know, it has not been used for stock forecasting. 🡪 dans un article de 2018 !)

Why to use an adversarial loss is that it can simulate the operating habits of financial traders. An experienced trader usually predicts stock price through the available indicator data, which is the work of the generative model 𝐺, and then judges the correct probability of his own forecast with the previous stock price, as the discriminative model 𝐷 does.

# Etat de l'art

## IA pour la finance

* Vincent Bouchet, 2017. ***Machine learning en finance : vers de nouvelles stratégies ?***[<https://dumas.ccsd.cnrs.fr/dumas-01706572/document>]

Ce travail se décompose en trois parties. Premièrement, la revue du fonctionnement du Machine Learning, ainsi qu’une présentation des principaux algorithmes. En deuxième, l’étude des applications de ces algorithmes dans les domaines de l’assurance, de la gestion de crédit, de la détection de fraudes et de la gestion de portefeuille. Enfin, la dernière partie repose sur l’étude de 42 publications académiques, pour l’analyse des rôles des nouveaux algorithmes, sur la question de la prédictibilité des cours et des stratégies de trading.

En conclusion, l’article précise que pour les acteurs financiers, les principes et applications de l’apprentissage automatique devient primordial. Non pas que les algorithmes utilisés soient nouveaux, mais les capacités de calculs et les sources de données ayant augmentées, la précision des modèles prédictifs vont également dans ce sens-là, et permettent de jouir d’un avantage compétitif sur le marché.

Aussi, pour la prédiction de cours, l’auteur nous dit qu’en 2017, les algorithmes utilisés sont principalement des algorithmes de classifications, visant à prévoir une hausse ou une baisse des titres. Concernant, les réseaux de neurones, avec leurs dérivés, ce sont les plus présents dans la littérature. Cependant, les modèles hybrides (exemple : algorithme génétique pour sélectionner les variables d’entrées + arbre de décision) sont plus performants, et permettraient dans certains cas de palier au problème de la « boite noire » rencontré avec les réseaux de neurones.

On pointe également du doigt l’importance de la qualité des données par rapport à l’algorithme utilisé, et l’analyse de texte (notamment sur les réseaux sociaux) présente un champ de recherche récent, dont les résultats sont mitigés.

On retient également que la précision (de l’ordre de 65%) des modèles de prédiction de cours n’a pas augmenté significativement sur les 25 dernières années. Néanmoins, l’étude soulève plusieurs approches qui seraient pertinentes pour des études futures. La première est l’utilisation d’un lexique spécifique afin de mieux mesurer les sentiments du web. Notamment en ciblant un secteur ou un titre en particulier. Une autre piste, serait l’analyse du lien entre volatilité des sentiments extraits du web et volatilité des titres.

* Weiwei Jiang, 2020. ***Applications of deep learning in stock market prediction: recent progress*** [<https://arxiv.org/pdf/2003.01859v1.pdf>]

Depuis longtemps, la question de prédictibilité des cours est un problème qui challenge les économistes et les informaticiens. Plus tardivement, les modèles de Deep Learning sont apparus et connaissent un développement rapide. L’objectif de cet article est de répertorier les principales avancées récentes, sur les modèles de Deep Learning appliqués à la prédiction de cours. On présente à chaque fois, les sources de données, les structures de réseaux de neurones, les métriques d’évaluation utilisées, l’implémentation, et enfin, la reproductibilité.

Je me suis surtout aidé de cet article pour me rediriger vers les travaux en lien avec mon sujet.

### Prédiction de tendance

* Shunrong Shen, Haomiao Jiang et Tongda Zhang, 2012. ***Stock Market Forecasting Using Machine Learning Algorithms*** [<http://cs229.stanford.edu/proj2012/ShenJiangZhang-StockMarketForecastingusingMachineLearningAlgorithms.pdf>]

Ce travail élabore une stratégie de prédiction de la tendance (hausse/baisse) du cours d’indices boursiers au jour suivant, à l’aide d’un algorithme de Machine Learning basé sur la méthode SVM (Support Vector Machine). Il propose un nouvel algorithme de prédiction, qui exploite la corrélation temporelle entre les marchés boursiers mondiaux et des produits financiers variés. En particulier, on s’intéresse à la corrélation entre les prix de fermeture des marchés qui ferment juste avant, ou au moment de l’ouverture des marchés américains.

Ici, les auteurs ont analysé les corrélations entre la tendance de l’indice NASDAQ et celle d’autres indices et instruments mondiaux, notamment en cherchant des valeurs temporellement corrélées, à un ou quelques jours près, qui seraient disponibles avant l’ouverture du marché US. Ils observent par la suite que plus la période d’analyse est grande, plus la corrélation entre les différents instruments sera élevée. En effet, ce phénomène est d’une part dû à l’équation utilisée, mais de plus, l'opération effectue implicitement un moyennage des données dans l'intervalle observé, qui supprime effectivement le bruit et ainsi rend la corrélation sous-jacente entre les marchés plus visible. En combinant tendance au jour près et mouvement à long terme, sur les 4 instruments les plus corrélés avec le NASDAQ, avec un algorithme SVM on obtient une précision de 74.4% sur la prédiction de la tendance de cet indice.

* Rohit Choudhry et Kumkum Garg, 2008. ***A Hybrid Machine Learning System for Stock Market Forecasting*** [<https://www.researchgate.net/publication/238747905>]

Ici, les auteurs choisissent de combiner deux méthodes de Machine Learning dans la prédiction du cours d’une action, pour dépasser, en termes de précision, les modèles basés sur un simple SVM. Un algorithme génétique (GA) est utilisé en amont pour sélectionner les données les plus pertinentes à passer en entrée du SVM prédicteur. L’étude porte sur le marché indien, car a cette époque, relativement peu de travaux traitaient ce marché. Aussi, les auteurs indiquent que le choix du marché étudié va influer sur les résultats, car chaque place boursière a ses caractéristiques. On en conclut alors que des résultats obtenus pour un instrument, ne seront pas forcément vrais pour un autre.

Comme précédemment, l’article pointe du doigt la non-isolation des actifs, et le fait que les mouvements de l’un peuvent être étudiés pour prédire un autre. Ici, comme ce qu’on pouvait supposer, ce sont les valeurs du même groupe ou du même secteur qui sont le plus corrélés à l’actif à prédire.

Comme valeurs d’entrée, ce sont 35 indicateurs techniques sur chacune des valeurs corrélées et sur la cible elle-même qui sont fournis, avec par exemple le Momentum, le Williams %R, le ROC, la disparité, le %K stochastique et le PVT. L’algorithme génétique va donc choisir les caractéristiques les plus pertinentes parmi ces 35\*m indicateurs.

Les résultats de l’étude montrent que le traitement des données grâce à un algorithme génétique en amont améliore la précision obtenue de 3% environ par rapport à un simple modèle SVM qui obtenait des résultats autour de 56%. Enfin, ils conseillent d’intégrer des données politiques et économiques en valeurs d’entrée, ainsi que des connaissances du domaine spécifique au marché.

* Sidra Mehtab et Jaydip Sen, 2020. ***Stock Price Prediction Using Convolutional Neural Networks on a Multivariate Timeseries*** [<https://www.researchgate.net/publication/338477393>]

Dans ce travail, l’objectif est d’utiliser une méthode hybride de Machine Learning et Deep Learning pour la prédiction des prix de fermeture de l’indice NIFTY 50, sur le marché indien. Des modèles prédictifs de Machine Learning ont été construit sur les données de l’indice sur la période 2015-2018, pour à partir de ces modèles prédire les prix de clôture de l’année suivante, à un horizon de prédiction de 1 semaine (5 jours). Pour prédire les mouvements de l’indice, les auteurs ont utilisés 8 méthodes de classification, tandis que pour prédire les prix, ils ont construit 8 modèles de régression. C’est ensuite, pour améliorer la performance de ces modèles qu’ils ont usé d’un modèle de régression basé sur du Deep Learning, notamment sur les CNN (Convolutionnal Neural Network).

L’analyse du passé de l’indice, de janvier 2015 à décembre 2018, s’est portée sur 9 variables issues de ces données brutes: (i) Date, (ii) Open, (iii) High, (iv) Low, (v) Close, et (vi) Volume. Ainsi, les auteurs en ont dérivé : (a) mois, (b) jour du mois, (c) jour de la semaine, (d) variations de Close sur deux jours successifs en pourcentage, (e) variations de Low (comme (d)), (f) variations de High (comme (d)), (g) variations de Open (comme (d)), (h) variations de Volume (comme (d)), (j) différence d’amplitude sur deux jours consécutifs.

L’objectif de la technique de régression est alors de prédire valeur de (d) pour chaque jour de la semaine suivante, en se basant sur les données passées, jusqu’à la semaine actuelle. Pour la méthode de classification, la variable (d) est binarisé pour en déduire : « 1 » une hausse ou « 0 » une baisse de l’indice, selon le signe de la variable.

L’évaluation des 8 approches de classification se fait sur 5 critères : le rappel (sensibilité) qui est la proportion de classés positifs parmi tous les items positifs, la spécificité qui mesure la capacité du modèle à donner un résultat négatif lorsque l’hypothèse n’est pas vérifiée, la précision (*positive predictive value)* ou la proportion de classifications positives effectivement correctes, la valeur prédictive négative (*negative predictive value*), et l’exactitude de classification soit la proportion de prédictions correctes. Pour les méthodes de régression, les mesures se font sur : la Root Mean Square Error (RMSE), et la corrélation entre les valeurs actuelles et prédites de la variable observées, c’est-à-dire des variations de Close (d).

Comme cité plus haut, les auteurs se sont servis des CNN pour améliorer la prévision, les avantages qu’offre un CNN sont l'apprentissage automatique des features les plus pertinents et la capacité du modèle à générer directement un vecteur « multi-step » (vecteur multi-step?). Pour exploiter cela on utilisera une stratégie de prévision soit récursive, soit directe. De manière récursive, le modèle effectue des prédictions étape par étape et les sorties sont fournies en tant qu'entrées pour les prévisions suivantes. Dans l'autre approche, le CNN prédit la séquence de sortie entière en une seule fois. En utilisant ces deux approches, on peut construire trois différents types de modèles CNN pour la prévision en plusieurs étapes de séries chronologiques des cours d’actions :

1. Prévision avec 1 variable d’entrée (d)
2. Prévision avec plusieurs variables d’entrée, chacune ayant son propre canal, et traitées séparément (Ex : image avec canal bleu, rouge et vert)
3. Prévision avec plusieurs variables d’entrée, chacune traitée dans un sous-modèle CNN, puis combinées pour être interprétées et donner la prédiction finale

Enfin, les résultats de cette étude montrent qu’en terme de classification, l’algorithme de Boosting a été le plus efficace en phase de test. Concernant les modèles de régression, les algorithmes de Machine Learning ont largement été surperformé par les CNN, et parmi eux, c’est la seconde approche avec multi-variables traitées séparément par canaux, qui a montré la meilleure performance. Pour finir, ils partagent leur intérêt pour l’exploitation d’un modèle GAN pour la prédiction de mouvements de marché.

### Generative Adversarial Network pour la finance

* Ricardo Alberto Carrillo Romero, 2019. ***Generative Adversarial Network for Stock Market price Prediction*** [<https://cs230.stanford.edu/projects_fall_2019/reports/26259829.pdf>]

L’objectif de cette étude est de comparer le modèle GAN (Generative Adversarial Network) aux modèles de Deep Learning suivants, dans la prédiction de la tendance, à la hausse ou à la baisse, un jour après la période analysée : le modèle ARIMA model, un modèle LSTM (Long Short-Term Memory), un modèle deep LSTM. Ici, le modèle GAN est basé sur un CNN (Convolutionnal Neural Network) à 3 couches comme discriminateur, et un perceptron multicouche (3 couches) comme générateur de prédiction de prix de fermeture pour 500 compagnies de Standard and Poor’s. Comme entrée de ce modèle on a pris 20 jours de prix ainsi que des indicateurs financiers avec parmi eux la *value over earnings*, le *price-to-book* (PB), ou le *price-to-earnings* (PE), additionnellement des tests sont effectués avec des données fondamentales sur les compagnies.

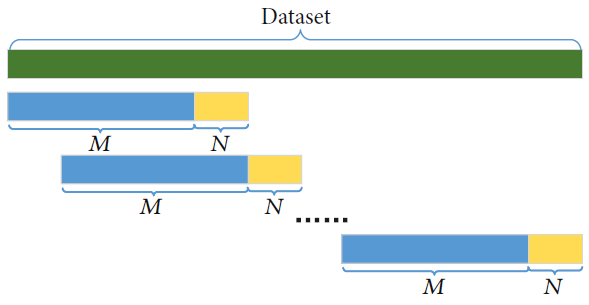
Dans les résultats observés pour le modèle Deep LSTM, l’auteur s’est aperçu que l’algorithme à une légère tendance à prédire une hausse plutôt qu’une baisse, il explique cela en supposant que généralement, les valeurs ont tendances à monter au fil du temps, en particulier ces dernières années (avant Covid), ce qui biaise le modèle en faveur d’une prédiction à la hausse. Pour le modèle GAN, la précision mesurée est de 73% en cas de hausse de l’actif, et 62% en cas de baisse. On peut alors supposer que l’hypothèse citée précédemment pourrait s’appliquer ici, et s’attendre, dans notre étude à une la même observation.

Enfin, en fin d’article, on nous suggère d’orienter nos recherches vers d’autres architectures de GAN pour simuler les séries temporelles, spécialement celles traditionnellement utilisées dans les modèles LSTM. Aussi, il serait intéressant d’explorer des fonctions de perte non conventionnelles pour le modèle GAN, comme celle des Wasserstein GAN (WGAN) qui use de la distance de Wasserstein. Il nous est également supposé le fait d’améliorer le paramétrage de notre réseau, avec un algorithme de renforcement. Finalement, on nous propose d’explorer la piste des Metropolis-Hastings GAN (MHGAN), pour une meilleure sélection de la distribution du discriminateur, qui va se servir de celle créée par le générateur.

* Xingyu Zhou, Zhisong Pan, Guyu Hu, Siqi Tang, et Cheng Zhao, 2018. ***Stock Market Prediction on High-Frequency Data Using Generative Adversarial Nets*** [<https://doi.org/10.1155/2018/4907423>]

Dans cette étude, on implémente une méthode usant des LSTM (Long Short-Term Memory) et CNN (Convolutional Neural Network) dans l’apprentissage d’un réseau GAN pour la prédiction de mouvement d’actions isolées ou d’indice à l’horizon d’une minute, sur la base de leurs données historiques. On rappelle pour commencer, que dans la plupart des cas, les résultats prédictifs sont jugés selon deux critères principaux : Premièrement l’erreur de prédiction, autrement dit la RMSE (ou RMSRE) entre le prix réel et la valeur prédite. Deuxièmement, la précision de la prédiction de direction (DPA), c’est-à-dire le pourcentage de prédictions correctes de direction (ou tendance) de séries de prix. Selon les algorithmes utilisés et ces critères d’évaluation, les auteurs nomment leur modèle GAN-FD (GAN for minimizing forecast error loss and direction prediction loss). D’une manière générale, ici le travail est double. On va d’abord adapter le modèle GAN à des fins de prévision des prix, ce qui constitue, à leur connaissance, la première application de ce réseau en bourse. Enfin, on effectue une sélection glissante sur les sets de données brutes d’entrainement et de test, pour étudier l’effet d’un cycle de mise à jour des paramètres du modèle sur les performances, tandis que les résultats expérimentaux montrent qu’un cycle plus petit peut améliorer les résultats de prévision.

Dans l’article, on prend les données historiques de 42 valeurs du CSI 300, soit 42x242 valeurs par journée, pour 242 minutes du China Stock Exchange. Aussi, l’expérience s’appuie sur 13 indicateurs techniques en plus des prix de fermeture historiques, qui sont : Prix d’ouverture, Prix maximum, Prix minimum, Volume, Turnover, Biais, Bandes de Bollinger, Indice de mouvement directionnel, Moyennes mobiles exponentielles, Indice stochastique, Moyennes mobiles, MACD (Convergence et divergence des moyennes mobiles) et Indice de force relative (RSI). Les auteurs précisent que la plupart des travaux connexes utilisent la méthode de partitionnement des données traditionnelle ; qui est, la séparation directe du dataset entier en set d’entraînement et set de test. Ici, pour faire part du changement fréquent de stratégie de trading, on décide de faire une répartition mobile des 2 sets à partir des données expérimentales, comme ceci :

 Avec en vert le dataset complet, en bleu la partie dédiée à l’entraînement du modèle, et en jaune celle dédiée au test du modèle. En prenant à chaque fois M jours de données pour l’entraînement et N jours pour le test, puis en avançant la fenêtre temporelle de N jours jusqu’à ce que toutes les données aient été observées.

Après expérimentation sur une période d’un an, les résultats sont les suivants : la moyenne maximale du DPA observée (0.6956) et la moyenne minimum du RMSRE (0.0079) sont obtenus pour M=20 et N=5. Même pour les autres modèles de comparaison, les résultats sont meilleurs lorsque N=5 plutôt que 10 ou 20. Cela démontre le fait que la tendance à très court-terme semble être la plus adaptée pour la prédiction des minutes futures. Donc un modèle avec un plus petit cycle de glissement est préférable ici. On note aussi que les pires résultats obtenus sont lorsqu’une portion faible et une portion grande des données sont attribuées respectivement à la phase d’entraînement et de test.

* Kang Zhanga, Guoqiang Zhonga,­, Junyu Donga, Shengke Wanga, Yong Wang, 2018. ***Stock Market Prediction Based on Generative Adversarial Network*** [<https://www.sciencedirect.com/science/article/pii/S1877050919302789>]

Ici, l’objectifs des auteurs est de prédire le prix de fermeture d’indices boursiers mondiaux, en générant la même distribution que les données journalières grâce à un modèle GAN construit avec un LSTM en tant que générateur, et un perceptron multicouche (MLP) comme discriminateur. Ils citent une hypothèse, nommée *Mean Reversion*, qui énonce le fait que les prix sont temporaires et tendent à bouger vers un prix moyen à travers le temps. De plus, il découle de cette hypothèse, celle de *Moving Average Reversion* (MAR), qui suppose que la moyenne du prix est la moyenne des prix d’une fenêtre temporelle passée (5 jours par exemple).

Pour entraîner le générateur, on lui fournit les données journalières sur 20 ans, de 7 indicateurs : Prix maximum, Prix minimum, Prix d’ouverture, Prix de fermeture, Volume, Taux de Turnover, Moyenne des prix de fermeture des 5 derniers jours (Ma5). Ces indicateurs semblent alors être les plus pertinents pour la prédiction dans les hypothèses citées précédemment. Les inputs sont représentés par X = {x0, x1, …, xt}, avec chaque xk un vecteur composé des 7 indicateurs à la date k. On normalise les données comme ceci : , avec et les moyenne et écart-type de X; on sélectionne t=5 car les auteurs souhaitent prédire le prix avec les données de la semaine passée. Dans l’article on choisit d’utiliser les premiers 90%-95% des données pour la phase d’entraînement puis les 5%-10% restant pour tester.

Pour évaluer le modèle, l’article utilise les indicateurs statistiques suivant : Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) et Average Return (AR).

## Analyse des réseaux sociaux pour la finance

* Tharsis T. P. Souza, Olga Klochyna, Philip Treleaven et Tomaso Aste, 2015. ***Twitter Sentiment Analysis Applied to Finance: A Case Study in the Retail Industry*** [<https://www.researchgate.net/publication/279864932>]

Cet article étudie le lien entre les sentiments issus de Twitter, concernant certaines sociétés de commerce de détail cotées, et les rendements, et volatilité des actions financières. Il vérifie s'il existe des informations statistiquement significatives dans cette relation, et les compare également à une analyse correspondante en utilisant les sentiments issus des sources de presse traditionnelles (Dow Jones Newswires, le Wall Street Journal, et Barron's).

Il conclut en démontrant l’existence d’une relation, statistiquement pertinente, entre les sentiments issus de Twitter et les valeurs boursières des marques observées. Les résultats suggèrent donc que l'activité sur les réseaux sociaux, joue un rôle significatif dans la dynamique du secteur du commerce de détail sur les marchés financiers.

* Johan Bollen, Huina Mao, Xiaojun Zeng, 2010. ***Twitter mood predicts the stock market*** [<https://doi.org/10.1016/j.jocs.2010.12.007>]

Ici, l’article estime la corrélation entre les états d'humeur collectifs, dérivés du flux Twitter (à grande échelle), et la valeur du Dow Jones Industrial Average (DJIA), au fil du temps. Il analyse le contenu textuel du flux Twitter quotidiens avec deux outils d’analyse : OpinionFinder, qui mesure l'humeur positive par rapport à l'humeur négative, et Google-Profile of Mood States (GPOMS) qui mesure l'humeur suivant 6 dimensions (calme, alerte, serein, « vital » , gentil et heureux). Enfin, l’article cross-valide la série chronologique de l'humeur résultante, en mesurant sa capacité à détecter la réponse du public à l'élection présidentielle et au jour de Thanksgiving en 2008.

Les résultats indiquent que la précision de la prédiction du DJIA peut être augmenté significativement en incluant des dimensions spécifiques de sentiments du public.

# Prédiction du cours de l'action X grâce à un GAN

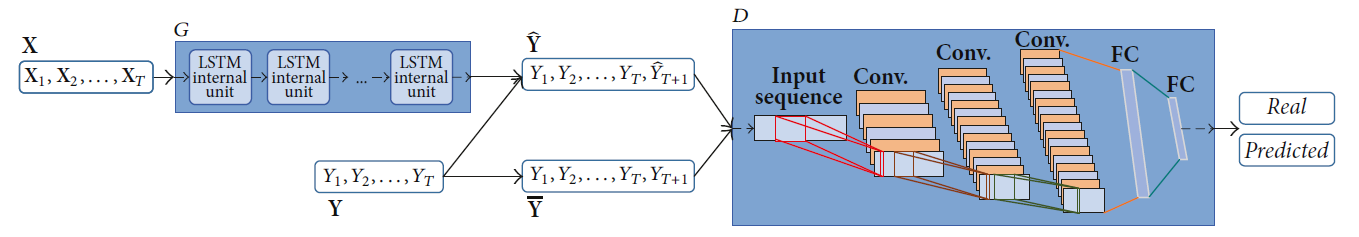
Efficient market hypothesis (EMT) 🡪 maintains that technical analysis or fundamental analysis (or any analysis) would not yield any consistent over-average profit to investors. However, many researchers disagree with EMH. EMH = prices follow a random walk and cannot be predicted based on their past behavior. According to EMH, all information that enters the market affects the prices instantaneously. If the EMH were true, it would not be possible to use AI techniques to predict the market. However, due to the success of technical analysts in the financial world and a number of studies appearing in academic literature successfully using AI techniques to predict the market, EMH is widely believed to be a null hypothesis now. Technical analysts make use of technical indicators, which are mathematical formulations which give us clues about the trend of the market. An example of a technical indicator is the famous stochastic oscillator %K.

-----------------------

Studies have also shown that predicting direction as compared to value can generate higher profits [Chen, A.S., Leung, M.T., and Daouk, H. Application of Neural Networks to an Emerging Financial Market: Forecasting and Trading the Taiwan Stock Index. Computers and Operations Research 30, 2003, 901-923.]

----------------------

Generative adversarial network (GAN) is a framework for estimating generative models via an adversarial process, in which we simultaneously train twomodels: a generativemodel𝐺that captures the data distribution and a discriminative model 𝐷 that estimates the probability that a sample came from the training data rather than 𝐺. The training procedure for 𝐺 is to maximize the probability of 𝐷 making a mistake. This framework corresponds to a minimax two-player game. In the space of arbitrary functions 𝐺 and D, a unique solution exists, with 𝐺 recovering the training data distribution and 𝐷 equal to 0.5 everywhere in [I. J. Goodfellow, J. Pouget-Abadie, M. Mirza et al., “Generative adversarial nets“]. Most researches recently constructed 𝐺 and 𝐷 on the basis of Long Short-Term Memory (LSTM) or convolutional neural network (CNN) for a large variety of application.



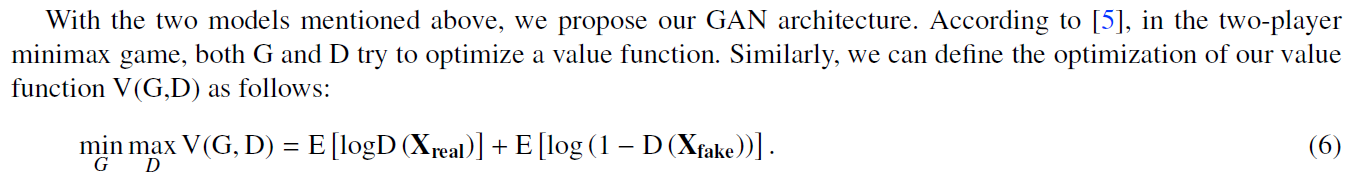
GAN-FD architecture.The generator (𝐺) is founded on LSTM, which applies to predicting Y^(𝑇+1).The discriminator (𝐷) is based on CNN for the purpose of estimating the probability whether a sequence is real (**Y**) or being predicted (**Y-**). *Conv.* means convolutional layer, *FC* is an abbreviation for fully connected layer.The structure of 𝐺 and 𝐷 can be adjusted according to the specific application.

Y^(T+1) = G(X)

LSTM is a basic deep learning model and capable of learning long-term dependencies. A LSTM internal unit is composed of a cell, an input gate, an output gate, and a forget gate. LSTMinternal units have hidden state augmented with nonlinear mechanisms to allow state to propagate without modification, be updated, or be reset, using simple learned gating functions. LSTM work tremendously well on various problems, such as natural language text compression, handwriting recognition, and electric load forecasting.

CNN is a class of deep, feed-forward artificial neural networks that has successfully been applied to analyzing visual imagery. A CNN consists of an input layer and an output layer, as well as multiple hidden layers.The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers, and normalization layers. CNN also has many applications such as image and video recognition, recommender systems, and natural language processing.

----------



## Données (Type, Sources, Utilité)

## Méthode (possible fusion avec la section précédente)

Voir [***Stock Market Prediction on High-Frequency Data Using Generative Adversarial Nets, Section 3]***

## Application

## Analyse des résultats

# Mesure du "sentiment de marché" grâce à l'activité sur Twitter

## Données (Type, Sources, Utilité)

## Méthode

* Quantifier les sentiments
  + identify two types of mood: (i) polarity (positive vs. negative) and (ii) emotions (calm, alert, sure, vital, kind, and happy) ["C:\Users\mthib\Documents\M2IF\Memoire\Social Network\2011\_Twitter\_mood\_predicts\_the\_stock\_market.pdf"]
* Mesurer l’attention donnée (le volume)

## Application (possible fusion avec la section précédente)

## Analyse des résultats

# Intégration du sentiment de marché au modèle GAN

## Méthode

## Application (possible fusion avec la section précédente)

## Analyse des résultats