

Système d'analyse de la polarité de dépêches financières

Michel Génereux
Centro de Linguística da Universidade de Lisboa
Av. Prof. Gama Pinto, 2
1649-003 Lisboa - Portugal
genereux@clul.ul.pt

Résumé. Nous présentons un système pour la classification *en continu* de dépêches financières selon une polarité positive ou négative. La démonstration permettra ainsi d'observer quelles sont les dépêches les plus à même de faire varier la valeur d'actions cotées en bourse, au moment même de la démonstration. Le système traitera de dépêches écrites en anglais et en français.

Abstract. We present a system for classifying *on-line* financial news items into a positive or negative polarity. The demonstration will therefore allow us to observe which news are most likely to influence the price of shares traded at the time of the demonstration. The system will cover news items written in English and French.

Mots-clés : Analyse de Sentiments, Linguistique de Corpus, Dépêches Financières.

Keywords: Sentiment Analysis, Corpus Linguistics, Financial News Items.

Description

Nous avons mené des travaux sur l'analyse de sentiments en utilisant comme domaine d'investigation celui de la finance (Génereux *et al.*, 2011). Nous faisons l'hypothèse que la réaction du marché suite à la publication d'une dépêche reliée à une action particulière est un bon indicateur de la polarité de la nouvelle, et qu'un algorithme d'apprentissage à partir de ces dépêches permet de construire un système qui donne à l'investisseur une source d'information supplémentaire qui peut être exploitée de façon avantageuse dans une stratégie d'investissement. Dans cette démonstration, nous voulons donner un aperçu des possibilités offertes par ce genre d'analyse dans une situation concrète où des dépêches sont examinées et classées dès leur publication. Notre système effectue l'apprentissage d'un modèle pour la classification polaire (positif ou négatif) à partir de dépêches déjà annotées selon qu'elles précèdent une hausse ou une baisse de la valeur d'une action particulière. Selon cette perspective, toute dépêche précédant une hausse significative de la valeur d'une action sera annotée comme positive, négative dans le cas contraire. Ces dépêches annotées nous permettent d'établir un lexique de termes financiers discriminants servant à construire un classifieur SVM. Notre démonstrateur classe les dépêches *en continu*, i.e. qu'il analyse les flots de dépêches en format RSS telles que fournies par des sites spécialisés dans la finance (e.g. *Financial Times*, *Les Échos*). Lorsqu'une dépêche présente une orientation non-neutre, elle est mise en évidence par le système, permettant à un investisseur potentiel de se concentrer sur les dépêches ayant, en principe, une plus grande probabilité de faire varier la valeur de l'action visée par la dépêche. Certes, la taille restreinte de certaines dépêches que l'on retrouve sur les flux RSS limite la portée d'une approche comme la nôtre basée sur la sélection de traits lexicaux, et les coûts de transaction requièrent que les gains potentiels soient significativement *au-delà* de la mise de départ. Néanmoins, la démonstration permettra à l'utilisateur de se faire une idée du potentiel d'exploitation de l'analyse de sentiments comme stratégie (partielle) d'investissement financier.

Références

GÉNÉREUX M., POIBEAU T. & KOPPEL M. (2011). *Sentiment analysis using automatically labelled financial news items*, In *Affective Computing and Sentiment Analysis : Metaphor, Ontology, Affect and Terminology*, chapter 9, p. 111–126. Springer Verlag.