A Neural-based model to Predict the Future Natural Gas Market Price through Open-domain Event Extraction

Minh Triet Chau[1], Diego Esteves[1,2], and Jens Lehmann[1,3]

[1] University of Bonn, SDA Research, Bonn, Germany
s6michau@uni-bonn.de, jens.lehmann@cs.uni-bonn.de
[2] Farfetch, Porto, Portugal, diego.esteves@farfetch.com
[3] Enterprise Information Systems, Fraunhofer IAIS, Dresden, Germany
jens.lehmann@iais.fraunhofer.de

Abstract. We propose an approach to predict the natural gas price in several days using historical price data and events extracted from news headlines. While previous methods depend only on the appearance of verbs in the headlines, our event extraction detects not only the occurrence of phenomena but also the changes of attribution and characteristics. Moreover, instead of using sentence embedding as a feature, we use every word of the extracted events, encode and organize them before feeding to the learning models. Empirical results show favorable results, in terms of prediction performance, money saved and scalability.

1 Introduction

Accurate market forecasting is a major advantage in business. However, there have been controversies about its feasibility in the academic world. Examining the stock market, [15] proposes the *Efficient Market Hypothesis* (EMH) which states that all information is reflected through the price. Moreover, regardless of how precise a price prediction is, once one acts on it, the price would change, invalidating the original prediction. This theory is also supported by Burton Malkiel in [22]. Later on, his position had changed in [17], claiming that there are certain patterns of the market that investors may benefit from, albeit quickly volatile. Moreover, [21] states that while the arguments for or against EMH are far from over, it is beneficial to find a more useful theory and prediction method than its alternatives. In this view, devising market prediction methods can be seen as a race to outperform other methods. Unlike in the stock market, there are few attempts on commodities market prediction [40]. However, important commodities such as oil, gas, and gold are getting more sensitive to macroeconomic news and surprise interest rate changes [28]. Inspired by the sensitivity of the stock market to the mood of news, most methods use positiveness or negativeness of news as a pointer for prediction. We argue that the market is not only sentimental-driven but also event-driven. Furthermore, we aim to solve the scarcity of unannotated and annotated news data by using public data. Most researchers [1,4,20,27,33] have to either purchase or manually annotate their news datasets, which lead to difficulties in experimenting with long price series. To those ends, we rely on headlines from public news API and propose an approach to both filter irrelevant headlines and address the event extraction preliminary in [31]. Both price and text are fed to a 3D Convolution Neural Network [36] to learn the correlation between events and the market movement.

2 Related works

In this section, we review the news-market relationship and existing benchmarks of market prediction tasks. In Table 1, we highlight their temporal evolution and henceforth categorize them by their input features and architecture.

Table 1: Summary of market prediction models

Method	Year	Features	Architecture
[20]	1996	Price	Feedforward network
[33]	2002	Price	Feedforward network
[37]	2013	Price	Recurrent Neural Network
[40]	2013	BOW	GARCH [13]
[11]	2014	BOW, TF-IDF	SVM, Neural Network
[25]	2015	Price, feature from text	Bidirectional RNN
[6]	2017	Price	Hidden Markov Model
[5]	2017	Price	RNN and autoencoders
[35]	2017	Price	Bilinear layer and temporal attention
[19]	2017	Price, Word embedding	Bidirectional RNN
[32]	2018	Price	RNN
[3]	2018	Price	Autoregressive model
[34]	2018	Price	Autoregressive model

2.1 Effect of news to the market

[16] shows that (1) negative news affects the market more than positive news, and (2) the perception of positive or negative changes over time. Analogously, there has been a growing body of NLP works concerning sentimental analyzing [10, 26, 29, 38]. [7] used dictionary-based and phrase analysis to classify the

[3] The code repository of our work is at https://github.com/minhtriet/gas_market

sentiment of news. They observed that the stock market is more volatile on days with relevant news than days with irrelevant news or without news. Using data from financial news from Reuters, [40] filters by topic code and their manual BOW then employs [13] to calculate the volatility of the market. They confirm the effect of the news on the crude oil market.

2.2 Price prediction

Price as the only feature In the stock market, a common task is to predict and maximize the return by predicting the selling and buying time for a stock. Models being used come from the auto-regressive model [3, 34] to Feed-forward Neural Network [20,33]. The difference between them is that [33] uses the genetic algorithm, rather than the gradient method, to train the weight of the network. Another method is Hidden Markov Models [6]. [32,37] claim that RNN is superior to feed-forward network. [5] use autoencoder in combination with RNN. [35] proposes the use of bilinear layer and temporal attention mechanism.

News-based prediction The line of work above inspired the approach to use news headlines to predict the increment or decline of the market. All the methods in this section [11, 12, 19, 25] use the now unpublished financial news from Reuters and Bloomberg. [19] fuse news and prices to predict price increments or decrements. Their model is Bidirectional Recurrent Network with GRU gates with prebuilt word embedding. [11] used Reverb to split sentences into Subjects, Verb, Objects, and concatenate them in different ways and feed to an SVM and a Neural Network. [25] predict price delta in two consecutive days. They defined seed words, which may serve as indicators of market movements, then use word embedding to select the other 1000 words that are closest to these seed words. They also handcrafted features including TF-IDF score, polarity score and categorical-tag (e.g, `new-product`, `acquisition`, `price-rise`, and `price-drop`). [11] created a set of features by first getting the result `(Subject, Verb, Obj)`, casting the Verb to its class using Verbnet [30], then one-hot encode all subjects, objects, and verbs, then define a set of concatenations of objects and verbs as features. [12] follows the same approach, but use word embedding instead. [27] uses part of speech to extract events and classify events into 23 classes of events using [14] and further subclasses (e.g. `unveils - unveiled - announces` for class `Product`).

3 Event extraction and embedding

Event extraction and semantic relationships are closely related. [24] proposes leveraging known relationships from databases (Freebase, DBPedia, YAGO) to classify a new relationship. However, the same entities can have different, even opposite relationship in news data. Another approach is using off-the-shelf IE frameworks (OpenIE, Reverb) for relation extraction as seen in [12]. Most methods rely on defined classes of events [9], which may not guarantee to cover every possible future event. Note that it is tricky to measure the accuracy of an open domain relation extraction method due to the high expense of manual annotation. One attempt is [27], who annotate on a few hundred tweets or Wikipedia sentences.

As a motivation example, we use two news headlines, in which the events are underlined.

Cuadrilla pauses mining operations after <u>tremor in Lancashire site</u>. (1)

With <u>natural gas plentiful and cheap</u>, carbon capture projects stumble. (2)

Although the two events above do not contain any verbs, they convey an occurrence of a phenomenon in (1) or an attribute in (2). However, both verb-based methods and Reverb could not extract any relation from these headlines. Conclusively, it is instinctive for humans to understand events, but elusive to obtain the same level of understanding with a machine. [18] classifies three different methods for event extraction (1) Data-driven which applies statistics to extract patterns, (2) Knowledge-driven which applies syntactic and schema and (3) Hybrid. According to their taxonomy, ours is a hybrid method, which leans towards the data-driven approach. For the sake of generalization, we define an *event* as a clause or phrase that conveys the occurrence of a phenomenon, an act or a change of an attribute.

Inspired by [2], we define a pipeline (Fig. 1a) to identify an event indicator using linguistic features, WordNet and a word sense disambiguation tool [41], which classifies lexical meaning of words from a sentence according to WordNet taxonomy. We depict the amount comparison of different methods in Fig. 1b.

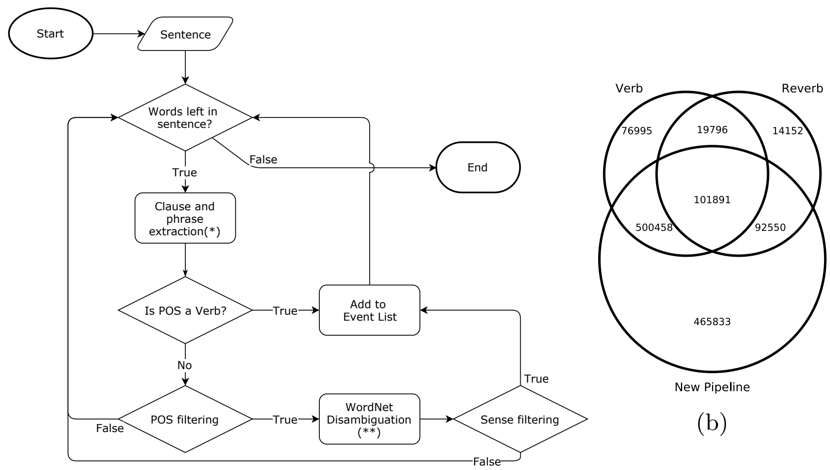
A common method to embed a sentence is using Sentence embedding. `spaCy` and `fasttext` treat an embedding of a sentence as a normalized or unnormalized average of its words' embedding. While it helps in some cases, two sentences with opposite meanings can have a small distance in the embedding space for just sharing a large number of similar words. We fix that by leveraging the even extraction pipeline above and concatenating the embedding of every word to form a representation of an event.

4 Experiments and Evaluation

In this section, we aim to test the predictive power of different models as well as applying them to a mock trading scenario to measure the amount of money saved. Before getting to the details, it may be beneficial to understand the structure of the natural gas market. It consists of the weekday-only *future market* in which an order is delivered from three months to three years, and the daily *spot market* in which an order is delivered on the very next day.

4.1 Data description

Our training data includes price series from Bayer AG suppliers (Fig. 2a). The future prices and spot price series are from 2 July 2007 to 12 October 2018 and



(a) (b)

Fig. 1: **(a)** Event extraction pipeline. (*) We take all the words whose POS in {`ADP`, `Verb`} or dependency in {`acl`, `advcl`, `ccomp`, `rcmod`, `xcomp`}. (**) If a phrase contains a word whose Wordnet sense is `noun.phenomenon` (e.g. death, birth), `noun.act` (e.g. acquisition, construction), `noun.event` (e.g. the **rise** and **fall**) or `adj.all`, `adv.all`, `noun.attribute` (which implies the change of attribute of that noun), we consider that phrase contains an event. **(b)** Performance comparison of different event extraction methods. Out of 1,271,675 sentences in headlines, our pipeline discover events from 1,160,732 headlines (91.27%), in comparison to 699,140 headlines (54.98%) of Verb-based only method and 228,389 headlines (17.96%) of Reverb.

from 2 June 2011 to 18 October 2018, respectively. We use the oldest 60% of the future price as the training data. The rest 40% and Spot Market price series are test data.

Corresponding news headlines are from The New York Times[4] (NYT), The Guardian[5] (TG) and The Financial Times[6] (FT) published at the corresponding time with the aforementioned price data. All the news providers allow filtering news within a time-range. TG and FT require a keyword (we chose "gas") and return filtered results while NYT requires downloading the whole dataset. Note that FT and TG return a headline if the keyword is in the article's body. Consequently, not every headline in the corpus contains the word "gas" (Table 2). We use the same keyword to filter the NYT dataset and name it NYTf (NYT filtered), the unfiltered dataset is NYTu (NYT unfiltered). An overview of the news dataset is in Fig. 2b.

[4] https://developer.nytimes.com, Accessed: 2020-03-30
[5] https://open-platform.theguardian.com, Accessed: 2020-03-30
[6] https://developer.ft.com/portal, Accessed: 2020-03-30

Table 2: Headlines we deem hard to discern their effect on the market

Date	Headline	Date	Headline
2007-04-27	Energy vs environment?	2007-05-03	Shell on a roll
2007-05-16	Big cap oil and mining	2007-05-17	Alternative energy
2007-05-24	Stress testing the hedge fund sector	2007-05-27	Darfur syndrome and Burma's grief
2007-08-17	Soil mates	2007-09-22	Master of the Universe (Rtd)
2007-09-23	Eni in Kazakhstan	2007-10-30	Texas Gold
2007-11-06	A Map of the Oil World	2010-07-19	For Cajuns, What Now?
2013-09-26	An Indian Tribe's Battle	2015-04-23	New Balance of Power
2015-08-04	Qatar's Liquid Gold	2015-12-08	Clean Sailing
2016-07-13	Report on China's Coal Power Projects	2018-04-14	Grand National 2018: horse-by-horse betting guide

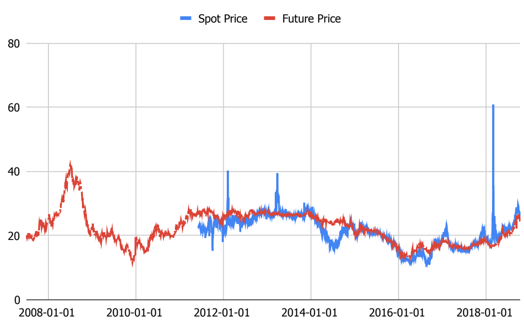
4.2 Baselines

Weak baselines Let i, j be two dates, $i < j$, p_k the price of gas on day k, $Y_i^j \in \{0, 1\}$ in which 0 means $p_i \geq p_j$ and 1 otherwise. We use chained CRF with the GloVe embedding of filtered news on day i to find Y_i^j. We first reimplement and compare [11] (See Table 1) with CRF and with ARIMA [8] without seasonal. The results are in Table 3a. [11] has worse result than in the original. Our hypothesis is that they used financial news dataset, while we just used a simple keyword filter.

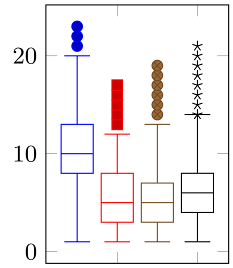
Strong baseline We feed the price and sentence embedding of filtered news using `spaCy` small English (Context tensor trained on [39], 300-d embedding vector) and large English model (trained on both [39] and Common Crawl, 300-d embedding vector, 685,000 vocabulary) of `spaCy` to a stacked LSTM structure as a strong baseline. Learning rate is 1×10^{-4}, dropout rate is 0.5, the LSTM layers have [128, 32] neurons. The overview of the structure is depicted in Fig. 3.

4.3 Event embedding with 3D Convolution (C3D)

We apply C3D [36] to a sequence of tensors, each of them being an embedding of the price and events of each day (Fig. 4). The event extraction pipeline (Fig. 1a) returns a list of event strings. For each string, we remove the stop words, then convert the rest to their stemming. Words that appear in more than 90% or less than three headlines are removed. In total, we have a vocabulary size of



(a) Overview of price data(€/m^3)



(b) Number of words in headlines distribution. Left to right: TG, FT, NYTf, NYTu

Fig. 2: Overview of price and headlines data. Best viewed in color.

2394 words + 1 OOV symbol for the training set. The next step is to find the tensor dimension. For our dataset, we find that limiting the number of events each day to 5 and words for each event to 15 covers the majority of our dataset. However, one could experiment more with these hyper-parameters. If a day has less than 5 events, an OOV vector is inserted into a random position to ensure homogeneous dimensions. If an event is shorter than 15 words, we OOV right pad it. Otherwise, its 15 first words are taken as input. We first fit a standard scaler on the price of the training set, then use the same scaler to transform the price of the test set. The size of the kernel is $3 \times 3 \times (300 + 1)$. We use SGD, learning rate 1×10^{-6}, with Nesterov Momentum, decay rate 1×10^{-7}. The experiment results are in Table 4. The noises from the unfiltered dataset contribute to the huge performance margin. Even within the filtered dataset, using only events instead of averaging the whole headlines helps to bring down the MSE for the C3D method.

4.4 Apply to mock trading

Settings The goal is to buy 1200m^3 of natural gas within D days. A daily goal is $\frac{1200}{D}$ m^3, on day d the algorithm should have bought $\frac{1200}{D}d$ m^3. If the algorithm does not buy on day d', it must buy the neglected amount in the next purchase. Given day d and prediction $Y = \{y_{d+1}, y_{d+2}..., y_{d+10}\}$ from the model trained with NYTf + TG + FT, if $\forall y \in Y : p_d < y$, buy immediately. The experiments in different markets and time frames are in Fig. 5 and Table 5. To see if the event extraction pipeline chooses the relevant words, we rank the words with the highest TF-IDF score in Table 6. Due to their high loss in Table 3b, we exclude [11] and ARIMA in this experiment.

Result analysis Both methods decide to buy on 07 February 2012 (Fig. 5e and Figure 5f) when the market reaches its peak at 40.27 €/m^3. A query for

Table 3: Performance comparison using data from previous ten days.

(a) Predict h-th day *away* with accuracy as the metric.

h	1	2	3	4	5
CRF	0.50	**0.55**	0.44	**0.54**	**0.54**
[11]	**0.54**	0.51	**0.51**	0.50	0.50

(b) Predict h *consecutive* days with MSE as the metric.

h	1	2	3	4	5
ARIMA	29.03	26.81	26.20	28.93	31.08
[11]	27.10	37.14	37.14	46.82	44.82

Table 4: Our comparison between difference prediction method using information in ten days to predict the price of the next five days. We use MSE as the metric.

	Small English model		Large English model		[11]	ARIMA
	LSTM	C3D	LSTM	C3D	(Table 3b)	(Table 3b)
NYTf+FT+TG	5.162	**2.862**	4.89	**2.858**	-	-
NYTu+FT+TG	25.513	**22.862**	25.189	**22.158**	44.82	31.08

"natural gas" from 06 February 2012 to 08 February 2012[7] returns a handful of results and does not show any news covering the shocking increment of this market. We conclude that this movement went under the radar. In the case of the sharp increment on 01 March 2018, there was news related to the matter, but not in both of our filtered and unfiltered news dataset.

On a brighter note, in Fig. 5b and 5f, C3D is always able to buy when the market is at the lowest peak (12 September 2018 in Future Market and 11 March 2012 in Spot Market). News headlines include "Energy price cap could be a muddle that satisfies no one", "Trump Administration Wants to Make It Easier to Release Methane Into Air", "Republicans' tired remedy for rising gas prices won't fix anything", "California drivers are using a lot less gas than they did in 2005". These decisions, however, do not save much money due to their small volumes. It is also evident in the small amount the third last purchase in Fig. 5b. Therefore, the amount of money saved may not be a strong performance indicator. Approaches using reinforcement learning are surveyed in [23], which

[7] https://www.google.com/search?q=%22natural+gas%22+%2B+news&tbs=cdr:1,cd_min:2/6/2012,cd_max:2/8/2012

Table 5: Performance comparison of buying all markets and time frames. The average prices are weighted by purchase volume. Our baseline is to buy the same amount every day.

			Average price (€/m^3)	
	Volume (m^3)	Cost (€)	Weighted	Unweighted
Future Market 2018				
Baseline	1,200	24,320.40	20.27	20.27
LSTM with Sentence embedding (Fig. 5a)	1,200	23,895.28	19.91	19.84
C3D with Event embedding (Fig. 5b)	1,187	23,600.87	**19.88**	**19.74**
Spot Market 2018				
Baseline	1,200	26,707.00	22.26	22.26
LSTM with Sentence embedding (Fig. 5c)	1,186	26,361.74	**19.70**	**19.66**
C3D with Event embedding (Fig. 5d)	1,191	26,659.71	22.38	22.18
Spot Market 2012				
Baseline	1,200	31207.31	26.01	26.01
LSTM with Sentence embedding (Fig. 5e)	1,198	31,262.27	26.09	25.34
C3D with Event embedding (Fig. 5f)	1,196	30,124.03	**25.19**	**25.01**

Table 6: Words with highest TF-IDF score from (a) Raw headlines, (b) Events after extraction pipeline, (c) Events from 10 days before a purchase in Fig. 5

No.	1 Jan 2012 - 1 Jan 2013			1 Jan 2018 - 1 Oct 2018		
	(a)	(b)	(c)	(a)	(b)	(c)
1	Sudan	energy	oil	nature	energy	energy
2	price	price	energy	week	oil	gas
3	deal	nature	price	change	China	oil
4	drill	fall	FTSE	US	Trump	China
5	nature	shale	fall	China	trade	Trump
6	energy	hit	shale	trade	plan	trade
7	approve	say	power	UK	rise	price
8	state	over	coal	supply	LNG	LNG
9	give	new	deal	regulation	plan	UK
10	reach	low	Shell	sell	demand	raise

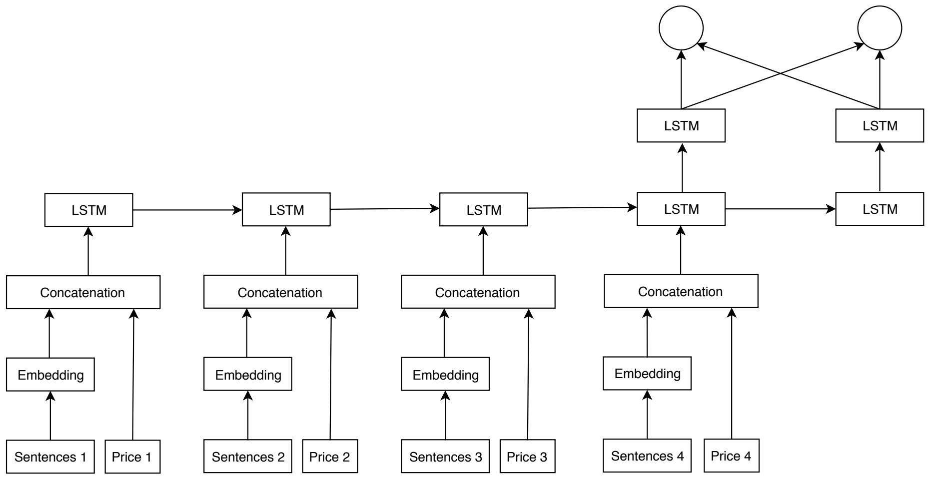


Fig. 3: A demonstration of the stacked LSTM structure. In this example, it uses data of four days to predict the gas price of the next two days

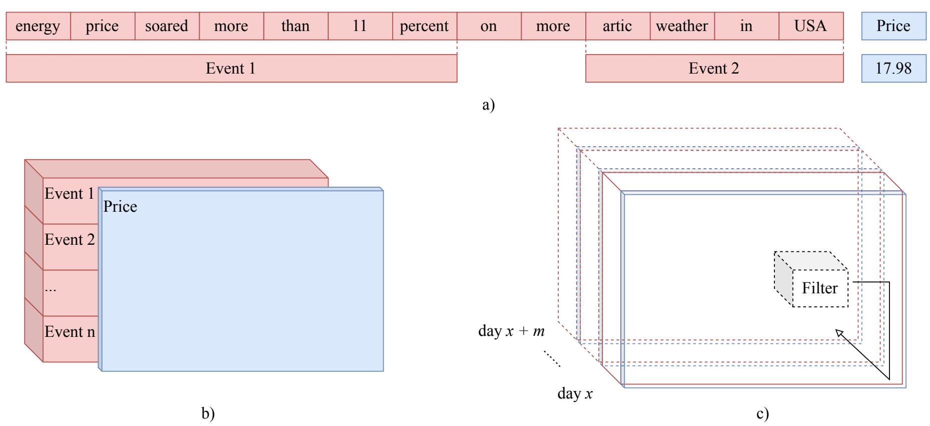
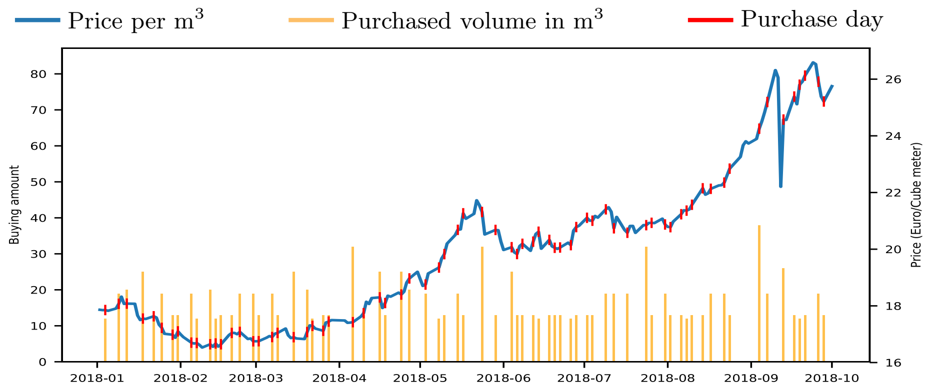
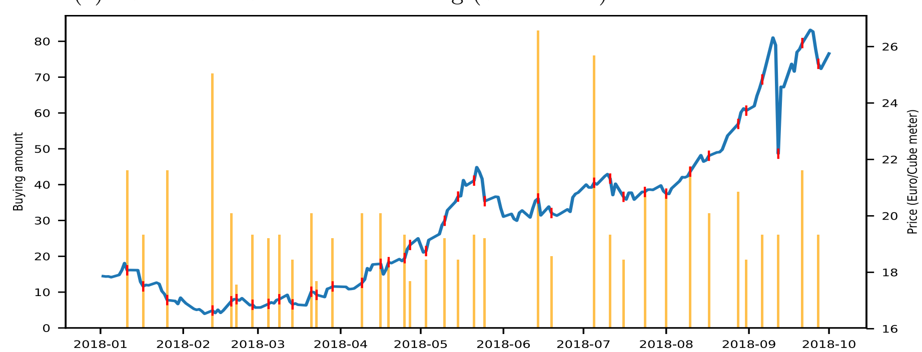


Fig. 4: **a)** Original data and the events segmentation **b)** The tensor representation of a day. Each word in events is embedded and appended. Price becomes another dimension on top of events. We form a tensor of $15 \times 5 \times (k + 1)$, in which k the dimension of the word embedding **c)** Consecutive m days are stacked together. The depth of a kernel is equal to the depth of one day's embedding (price + word embedding). Best viewed in color.

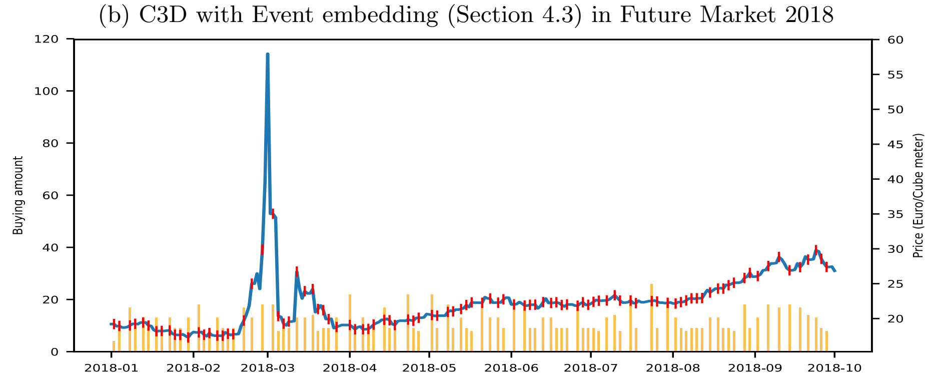
claims that RL delivers a substantive improvement on profitability and forecast accuracy. They also advocate for a combination of RL and deep neural networks.



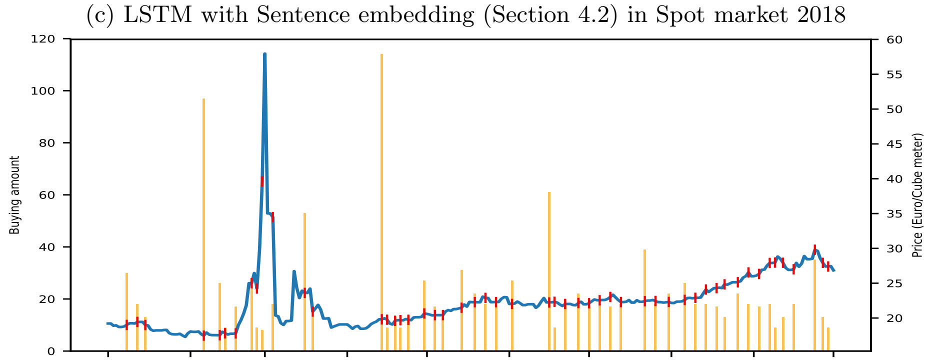
(a) LSTM with Sentence embedding (Section 4.2) in Future Market 2018



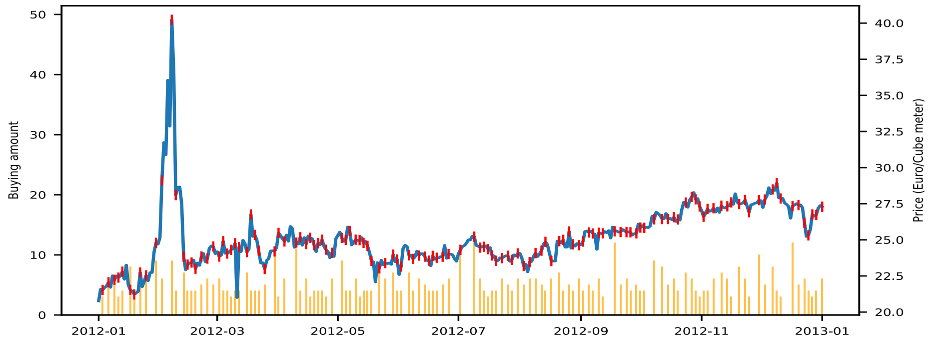
(b) C3D with Event embedding (Section 4.3) in Future Market 2018



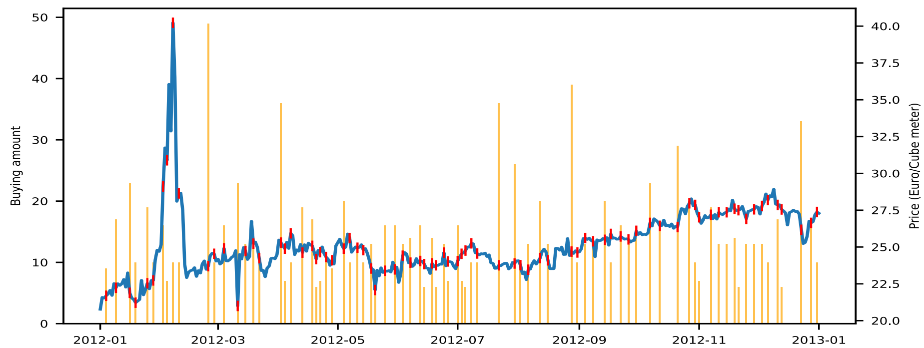
(c) LSTM with Sentence embedding (Section 4.2) in Spot market 2018



(d) C3D with Event embedding (Section 4.3) in Spot Market 2018



(e) Stacked LSTM with Sentence embedding (Section 4.2) in Spot Market 2012



(f) C3D with Event embedding (Section 4.3)in Spot Market 2012

Fig. 5: Comparison between buying methods in different time-frames and markets. Best viewed in color.

5 Conclusion

We proposed a new method to predict the natural gas price. Instead of averaging the embedding vectors, we extract and organize events from news and reshape them into 3D tensors. A limitation of our method is the reliance on the window approach for prediction. It is tricky to determine the length of a window that includes all events that have effects on the price of a specific day. An alternative is using a chain of linked events, proposed in [31]. Furthermore, our method cannot take events that happen on a non-trading day into account due to the absence of price data leading to the wrong dimension of input data. The news headlines curation needs minimum collecting efforts. Transfer learning only requires retraining on the last layers. Overall, our approach allows easier adaption to different domains prediction with minimal changes. We compare the money saved using our method and the average market price and prove its efficiency as well as the importance of a better purchase strategy.

Acknowledgement We are immensely grateful to Dr. Bernard Sonnenschein for his comments on an earlier version of the manuscript.

References

1. Aiello, L.M., Petkos, G., Martin, C., Corney, D., Papadopoulos, S., Skraba, R., Goker, A., Kompatsiaris, I., Jaimes, A.: Sensing trending topics in twitter. Trans. Multi. **15**(6), 1268–1282 (Oct 2013). https://doi.org/10.1109/TMM.2013.2265080, https://doi.org/10.1109/TMM.2013.2265080
2. Araki, J., Mitamura, T.: Open-domain event detection using distant supervision. In: Proceedings of the 27th International Conference on Computational Linguistics. pp. 878–891. Association for Computational Linguistics (2018), http://aclweb.org/anthology/C18-1075
3. Ariyo, A.A., Adewumi, A.O., Ayo, C.K.: Stock price prediction using the arima model. In: 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation. pp. 106–112 (March 2014). https://doi.org/10.1109/UKSim.2014.67
4. Atzeni, M., Dridi, A., Reforgiato Recupero, D.: Fine-grained sentiment analysis on financial microblogs and news headlines. In: Dragoni, M., Solanki, M., Blomqvist, E. (eds.) Semantic Web Challenges. pp. 124–128. Springer International Publishing, Cham (2017)
5. Bao, W., Yue, J., Rao, Y.: A deep learning framework for financial time series using stacked autoencoders and long-short term memory. In: PloS one (2017)
6. Betancourt, B., Rodríguez, A., Boyd, N.: Modelling and prediction of financial trading networks: An application to the NYMEX natural gas futures market. ArXiv e-prints (Oct 2017)
7. Boudoukh, J., Feldman, R., Kogan, S., Richardson, M.: Which news moves stock prices? a textual analysis. Working Paper 18725, National Bureau of Economic Research (January 2013). https://doi.org/10.3386/w18725, http://www.nber.org/papers/w18725
8. Box, G.E.P., Jenkins, G.: Time Series Analysis, Forecasting and Control. Holden-Day, Inc., USA (1990)
9. Christopher Walker, Stephanie Strassel, J.M.K.M.: Ace 2005 multilingual training corpus (2015), https://catalog.ldc.upenn.edu/LDC2006T06
10. Denil, M., Demiraj, A., Kalchbrenner, N., Blunsom, P., de Freitas, N.: Modelling, visualising and summarising documents with a single convolutional neural network. CoRR **abs/1406.3830** (2014), http://arxiv.org/abs/1406.3830
11. Ding, X., Zhang, Y., Liu, T., Duan, J.: Using structured events to predict stock price movement: An empirical investigation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 1415–1425. Association for Computational Linguistics (2014). https://doi.org/10.3115/v1/D14-1148, http://aclweb.org/anthology/D14-1148
12. Ding, X., Zhang, Y., Liu, T., Duan, J.: Deep learning for event-driven stock prediction. In: Proceedings of the 24th International Conference on Artificial Intelligence. pp. 2327–2333. IJCAI'15, AAAI Press (2015), http://dl.acm.org/citation.cfm?id=2832415.2832572
13. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica **50**(4), 987–1007 (1982), http://www.jstor.org/stable/1912773

14. Erosheva, E., Fienberg, S., Lafferty, J.: Mixed-membership models of scientific publications. Proceedings of the National Academy of Sciences **101**(suppl. 1), 5220–5227 (2004). https://doi.org/10.1073/pnas.0307760101, http://www.pnas.org/content/101/suppl_1/5220.abstract
15. Fama, E.F.: Efficient capital markets: A review of theory and empirical work. The Journal of Finance **25**(2), 383–417 (1970), http://www.jstor.org/stable/2325486
16. Feuerriegel, S., Neumann, D.: News or noise? how news drives commodity prices. In: ICIS (2013)
17. G. Malkiel, B.: The efficient market hypothesis and its critics. Journal of Economic Perspectives **17**, 59–82 (02 2003). https://doi.org/10.1257/089533003321164958
18. Hogenboom, F., Frasincar, F., Kaymak, U., Jong, F.D.: An overview of event extraction from text. In: Workshop on Detection, Representation, and Exploitation of Events in the Semantic Web (DeRiVE 2011) at Tenth International Semantic Web Conference (ISWC 2011). Volume 779 of CEUR Workshop Proceedings., CEURWS.org (2011) 48–57 (2011)
19. Huynh, H., Dang, L.M., Duong, D.: A new model for stock price movements prediction using deep neural network (2017). https://doi.org/10.1145/3155133.3155202
20. Kaastra, I., Boyd, M.: Designing a neural network for forecasting financial and economic time series (1996)
21. Lekovic, M.: Evidence for and against the validity of efficient market hypothesis. Economic Themes **56**, 369–387 (11 2018)
22. Malkiel, B.G.: A Random Walk Down Wall Street. Norton, New York (1973)
23. Meng, T.L., Khushi, M.: Reinforcement learning in financial markets. Data **4**(3), 110 (Jul 2019). https://doi.org/10.3390/data4030110, https://doi.org/10.3390/data4030110
24. Mintz, M., Bills, S., Snow, R., Jurafsky, D.: Distant supervision for relation extraction without labeled data. In: Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2 - Volume 2. pp. 1003–1011. ACL '09, Association for Computational Linguistics, Stroudsburg, PA, USA (2009), http://dl.acm.org/citation.cfm?id=1690219.1690287
25. Peng, Y., Jiang, H.: Leverage financial news to predict stock price movements using word embeddings and deep neural networks. CoRR **abs/1506.07220** (2015), http://arxiv.org/abs/1506.07220
26. Poria, S., Cambria, E., Hazarika, D., Vij, P.: A deeper look into sarcastic tweets using deep convolutional neural networks. CoRR **abs/1610.08815** (2016), http://arxiv.org/abs/1610.08815
27. Ritter, A., Mausam, Etzioni, O., Clark, S.: Open domain event extraction from twitter. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 1104–1112. KDD '12, ACM, New York, NY, USA (2012). https://doi.org/10.1145/2339530.2339704, http://doi.acm.org/10.1145/2339530.2339704
28. Roache, S.K., Rossi, M.: The effects of economic news on commodity prices. The Quarterly Review of Economics and Finance **50**(3), 377–385 (2010), https://EconPapers.repec.org/RePEc:eee:quaeco:v:50:y:2010:i:3:p:377-385
29. Ruder, S., Ghaffari, P., Breslin, J.G.: INSIGHT-1 at semeval-2016 task 5: Deep learning for multilingual aspect-based sentiment analysis. CoRR **abs/1609.02748** (2016), http://arxiv.org/abs/1609.02748
30. Schuler, K.K.: Verbnet: A Broad-coverage, Comprehensive Verb Lexicon. Ph.D. thesis, Philadelphia, PA, USA (2005), aAI3179808

31. Shekarpour, S., Shalin, V.L., Thirunarayan, K., Sheth, A.P.: CEVO: comprehensive event ontology enhancing cognitive annotation. CoRR **abs/1701.05625** (2017), http://arxiv.org/abs/1701.05625
32. Siami-Namini, S., Namin, A.S.: Forecasting economics and financial time series: ARIMA vs. LSTM. CoRR **abs/1803.06386** (2018), http://arxiv.org/abs/1803.06386
33. Skabar, A., Cloete, I.: Neural networks, financial trading and the efficient markets hypothesis. Aust. Comput. Sci. Commun. **24**(1), 241–249 (Jan 2002), http://dl.acm.org/citation.cfm?id=563857.563829
34. Taylor, S.J., Letham, B.: Forecasting at scale. The American Statistician **72**(1), 37–45 (2018). https://doi.org/10.1080/00031305.2017.1380080, https://doi.org/10.1080/00031305.2017.1380080
35. Tran, D.T., Iosifidis, A., Kanniainen, J., Gabbouj, M.: Temporal attention augmented bilinear network for financial time-series data analysis. CoRR **abs/1712.00975** (2017), http://arxiv.org/abs/1712.00975
36. Tran, D., Bourdev, L.D., Fergus, R., Torresani, L., Paluri, M.: C3D: generic features for video analysis. CoRR **abs/1412.0767** (2014), http://arxiv.org/abs/1412.0767
37. Valipour, M., Banihabib, M.E., Behbahani, S.M.R.: Comparison of the arma, arima, and the autoregressive artificial neural network models in forecasting the monthly inflow of dez dam reservoir. Journal of Hydrology **476**(Complete), 433–441 (2013). https://doi.org/10.1016/j.jhydrol.2012.11.017
38. Wang, P., Xu, J., Xu, B., Liu, C., Zhang, H., Wang, F., Hao, H.: Semantic clustering and convolutional neural network for short text categorization. pp. 352–357 (01 2015). https://doi.org/10.3115/v1/P15-2058
39. Weischedel, R., Palmer, M., Marcus, M., Hovy, E., Pradhan, S., Ramshaw, L., Xue, N., Taylor, A., Kaufman, J., Franchini, M., et al.: Ontonotes release 5.0 ldc2013t19. Linguistic Data Consortium, Philadelphia, PA (2013)
40. Wex, F., Widder, N., Liebmann, M., Neumann, D.: Early warning of impending oil crises using the predictive power of online news stories. In: 2013 46th Hawaii International Conference on System Sciences. pp. 1512–1521 (Jan 2013). https://doi.org/10.1109/HICSS.2013.186
41. Zhong, Z., Ng, H.T.: It makes sense: A wide-coverage word sense disambiguation system for free text. In: Proceedings of the ACL 2010 System Demonstrations. pp. 78–83. ACLDemos '10, Association for Computational Linguistics, Stroudsburg, PA, USA (2010), http://dl.acm.org/citation.cfm?id=1858933.1858947