

Márk MOLNÁR, Zsuzsanna NAÁR-TÓTH

### 3.7. APPLYING ALTERNATIVE METRICS IN QUANTIFICATION OF NEWS

#### Summary

Using a common definition, we can define news analysis as the measurement of the various qualitative and quantitative elements of textual news stories. These elements include sentiment, relevance and novelty. By quantifying news stories, we can gain a useful way to manipulate and use everyday information in a mathematically concise manner. In this article, a framework for news analytics techniques used in finance is provided. Various news analytic methods and software are discussed, and a set of metrics is given that may be applied to assess the performance of analytics. Various directions for this field are discussed. The proposed methods can help the valuation and trading of securities, facilitate investment decision making, meet regulatory requirements, or manage risk.

**Keywords:** textual news stories, finance, alternative metrics, software, investment decision making

#### A brief assessment of the acquisition of INA by MOL

##### *Corporate background*

The INA Privatization Act was accepted by the Croatian Parliament in 2002 to transform INA from a state-owned company to a privatized enterprise which would be controlled by independent private investors in order to boost its development and increase its operational efficiency and save the company from bankruptcy.

An open public tender was used to sell 25%+1 share of INA in May 2002. The winner was the Hungarian Oil Company (MOL) who offered a higher bid (\$505) than the competing Austrian oil company, ÖMV who offered (\$420M). Soon afterwards, MOL became an official strategic investor in INA. This is coupled by the Croatian government's veto power and special rights.

Following up on the potential offered by the new legislation the Hungarian Oil Company (MOL), built its equity in INA progressively, and in 8 years from 2003 it became the largest shareholder of the company with a share of 49.1% of the total equity. The Croatian government remained the second largest investor with a share of 44.84%. This increase in shares was undertaken through open market manoeuvres where a beneficially high price was offered for small shareholders.

From its entry onwards MOL always clearly stated its intentions to gain strategic control rights over INA. This was aided by its decisive position achieved by 2008 which allowed MOL to secure a majority in the Supervisory Board due to the terms laid down in the Shareholders Agreement in 2003 (which stated that the Croatian Government's stake in INA fell below 50%, it only had the right to nominate two members). The *de facto* acquisition of INA by MOL was approved by the Croatian Competition Agency and later by the European Commission.

The economic crisis of 2008 found INA in a quasi-bankrupt situation. In the following two years MOL provided a 330M € funding to help the Croatian company. This was augmented by the support from the domestic government which contributed to the stabilisation of INA's financial status.

The reasons for the severe impact of the crisis were multiple. A primary reason was the plunge in energy demand resulting in a direct decrease of revenues. For example, demand for diesel stayed 10% below the 2008 level even in 2012, for gasoline the drop was 14%. The aggregate effect was a 28% reduction in the size of the oil product market. Matching these figures with the refinery capacities, this would be equivalent to 70% of the Sisak plant's capacity.

Another impact was the change in gas price regulations which enforced a rate well under the market price generating a loss of approximately 325M € in the years 2007 and 2008.

As INA struggled with financing its investment projects due to the drying out of international financing sources. This was problematic because in the years preceding the crisis INA started significant investments designed to reach their maximum levels by 2010. A good indication for these investments would be that on average the average operating cash-flow was only about 60% of the investments. This practically meant that these investments could not be financed from the company's cash flow but funds had to be borrowed from external sources.

This was the reason why bankruptcy loomed over INA right when its investment programs were at their highest intensities. In the turmoil following the crisis the company had to satisfy its obligations towards banks, states hosting its projects and even to its suppliers. Key assets like the Syrian international oil and gas field development, the Adriatic fields, and the refinery modernisation all were at risk together with INA's role as a regional supplier.

In the followings, we give the dedicated steps taken to enhance liquidity and avoid default.

- Separation and sale of INA's gas business;
- Delayed tax payment: INA owed more than 204M € to the Croatian state at that time, and was given a temporary derogation.
- MOL's loan of approximately 330M € provided to INA during 2009 and 2010, at a time of frozen international financial markets. INA used part of this to cover salaries and oil purchases, and partially to repay debts towards the state.

Overall, borrowing the loan from MOL allowed INA to cover its operating expenses, and fully pay its taxes to the state by the Autumn of 2010

### ***Investment details***

The appearance of MOL as a major investment started a period which was earmarked by a period of intensive investments, this is underlined by the doubling of the annual investments. This beneficial change improved INA's core competences (exploration-production, refinery and distribution) and boosted its performance.

It is interesting to note that the average annual investment of INA after its privatisation surpassed pre-privatisation values with a factor of more than two. From an annual average value of 135M € it more than doubled to reach an average value of 365M€ per

annum. This was facilitated by a higher-than-the-industry's average reinvestment rate of approximately 80%.

Investment in upstream oil and gas production reached 2.35 billion euros, while targeted field investments in the northern Adriatic topped at 525 million euros, this brought about a historical high in gas production from this region in 2010-2011.

The proven reserves of natural gas increased with 3 billion m<sup>3</sup> in the past ten years due to increased exploration and development activities in the Adriatic areas. This was according to the initial agreement made and approved by the Croatian government which also focused on international projects. The overwhelming part of the allocated funds went to international targets (cca. 207M€), while only a minor share was assigned to Croatian plays.

Exploration activities were supported by MOL after its appearance as an owner of INA, this is confirmed by the discovery of three quite significant sites with one which turned out to be significant on the country level.

Joining forces also resulted in a cross-border geological research initiative between Croatia and Hungary in 2006-2007. Both of these drillings were successful and highlighted the operative potential in future cooperation. The research and exploitation of national hydrocarbon resources are also beneficial for improving regional energy security.

As of today upstream operations range from inland sites (land and off-shore in Croatia) to outland sites in Egypt and Angola. The Syrian location is blocked due to the conflict and sanctions placed on Syria's energy sector by the EU.

The downstream industry (refining and retail) faced a severe challenge with the upsurge of the economic crisis. This is highlighted by the plunge of diesel oil demand (-10%) in Croatia from 2008 to 2012, for gasoline the respective contraction was 14%. With the disappearance of fuel oil demand the total market loss reached 28%. With the parallel reduction of regional demand in Slovenia and Bosnia, the drop amounted to 70% of the Sisak refinery's capacity.

This however did not prevent management to successfully tackle the decline in the regional market share and penetrate other markets. The domestic share reached 70% by 2014 and even the crisis could not bring down motor fuel sales which showed continuous growth. To further ensure its strong market presence and improve competitiveness in the region, INA has taken the following steps:

- Modernisation of retail network, with 160 gas updated stations surpassing all competitors
- Investments targeted at improving fuel quality (EURO V)
- Increased investment in gas stations, reaching 30M€ in 2012
- Introduction of premium fuels

By the end of 2012, the retail network boasted with almost 400 gas stations in Croatia, 45 in Bosnia, 6 in Slovenia and one in Montenegro.

The total direct investment undertaken by MOL in Croatia reached 1.4 billion € making MOL the second largest foreign investor. With its entry on the Croatian market, MOL promotes economic development and transformation and aids the evolution of an environment which ensures INA's leading position in the South

Eastern European market. This was also supported by the construction of the Croatian-Hungarian gas interconnector.

### ***Strategic investments***

Numerous projects of strategic importance were undertaken and investment projects were launched in upstream, midstream and downstream. The initial agreement set during the privatisation process were surpassed in all areas with respect to investment levels, reaching an annual average of 365M€ which is more than double of the respective figure before the privatisation.

Regarding the upstream segment investments reached 2.25 billion € which is multiple times higher than a decade earlier. At the time of takeover no upstream projects were ready to launch so the new executive board decided to revise earlier project ideas and design new projects. A good example is INA's EOR project which targeted the improvement of oil and gas recovery from existing locations.

Important projects realised include the overwhelming share of investments in the Northern Adriatic and the Syrian plays, the cross-border exploration (in Zalata-Dravica and Novi Gradac-Potony), onshore projects and seismic scanning of the Southern Adriatic region.

It is strongly thought that the upstream segment is the core competence and key asset of INA and thus MOL is quite eager to develop this segment.

### ***Investment in refineries***

Modernisation of outdated refinery capacities were necessary as the competition also upgraded its facilities and Croatia accessed the EU. With the acquisition of the community legislation, the new fuel standards also required the refurbishment of refining. The first phase was finished in 2011, until 2013 altogether around 1.1 billion € was invested in the refineries. The investments were slowed down by slow administration, this delayed a coker investment in Rijeka resulting in 300-350M € investment lost. The coker would have allowed for a large scale of improvement, however the permits were not issued by the end of 2013. The new EU standards require a low sulphur content in fuels. INA could not by itself reach the required 10 PPM level thus requiring the addition of a significant desulphurisation capacity in order to remain on the market.

The desulphurisation units in Sisak were contracted by the end of 2005 to decrease sulphur in end-of-pipe emissions significantly, moreover to reduce benzene emissions and CO<sub>2</sub> emissions (flaring) from the refinery. This investment also enabled the sales of sulphur to the chemical industry. After the conclusion of the investment, a large isomerization unit was launched. The Sisak refinery was further refurbished with a combined Hydrodesulphurization & Fluid Catalytic Cracker (FCC) which permitted the production of low sulphur fuels.

Other investments targeted the Rijeka refinery's tank farm, sewage system and gas leakage. Further contracts involved a hydrocracking complex, a sulphur recovery unit and a hydrogen generation unit.

By 2009 INA has already spent 350M€ on the modernisation programme, which grew to 520M€ by the end of 2010. All the above-mentioned modernisation investments were operational and online by January 2012.

### ***Development of the retail segment***

INA's retail network is presently the most update in the Croatian market and among the leading ones in the region. This is due to the Blue Concept, a modernisation program undertaken by INA with the support of MOL starting from 2010. Beforehand only minor upgrades were accomplished thus jeopardising competitiveness and falling behind other market participants.

With the introduction of premium brand fuels which incorporated biofuel components a new line of gas stations became necessary to ensure the appropriate market reception, this project was named Blue Concept and targeted the rejuvenation of INA's retail. This was coupled with a doubling of capital investment in this segment by the company and resulted in the modernisation of more than 130 gas stations by the end of 2012 with a new design and a facelift. Infrastructure, equipment, catering and services were improved and by the end of 2014, already 160 stations boasted with the new image. It can be stated without hesitation that this rate of improvement cannot be matched with earlier periods' developments.

All these abovementioned developments helped INA to retake its leadership in its markets and successfully increase its overall operational efficiency. It has to be noted that these processes were achieved in a market with an ongoing recession and the contemporary accession to the EU unified internal market where no trader barriers (therefore no protectionist measures) are allowed thus rendering INA suddenly one of the many participants in a competitive environment.

Later in 2013 following the European trends a pilot project was launched which basically is renting the gas stations management rights to the entrepreneurs who seems to be most fit managers, and maintains INA as the owner of capital assets like concessions, buildings, and infrastructure.

The franchise allow for the highest quality service and exploitation of the potential turnover of the station ensuring profitable operation for both INA and the individual holder of the franchise rights.

### ***Oil and gas exploration in Croatia***

Starting from 2003 approximately 2.35 billion € was invested into oil and gas extraction, and about 525 million € in the Northern Adriatic fields allowing for a record production from 2010. The exploration activities added an additional proven reserve of more than 3 billion m<sup>3</sup> to INA's reserve base assets.

This was in line with the Initial Business Plan which shifted the focus of exploration to offshore Adriatic and allocated more than 200M € but less than 10M € to inland exploration.

As mentioned earlier, cross border explorations with Hungary were also very successful and helped INA's proven and probable reserves to reach 267 MMboe, and a daily production of 48.6Mboe. Extensive upstream operations in Croatia are still a key to the country's energy security.

### ***Operational development of INA***

As the strengthening and improved profitability of INA is aligned with MOL's long term objectives, MOL initiated the restructuring of INA's operative profile by starting rationalisation programmes and efficiency improvement which resulted in significant improvement, doubling of operational profit and a decrease in its debts. This helps INA further play a vital role in the country's energy security.

Concerning energy security a major step was the construction of the gas interconnector with Hungary. The construction of the Croatian-Hungarian gas connection began in 2009 by FGSZ, a daughter company of MOL. This significantly improved the security of energy supply in Croatia. Financing was partially supported by EU, and Hungary also invested 275M€ on their side. The pipeline is 205 kms long in Hungary and 88kms long in Croatia, carrying a maximum 19.2M m<sup>3</sup> per day.

The interconnection supports the priorities of the EU energy policy as it strongly promotes CEE region's connectivity to mobilise energy supplies and contributes to the North-South gas corridor. It is considered as an important step towards the unified transparent community gas market.

Practical life brought a real stress test to this enhanced connectivity, as in February 2012 a very cold weather was coupled with a unscheduled disruption of gas production in the Northern Adriatic gas fields. The interconnector securely allowed for the resupplying of the Croatian gas system from Hungary amounting to one-third of the daily gas consumption. In this case, households and industrial consumers remained undisturbed in their consumption and the supply was secured.

Despite the intergovernmental agreement between Croatia and Hungary, this interconnector is still unidirectional, from Hungary towards Croatia. As this was a significant investment this limitation strongly jeopardises MOL's investments return and undermines the security of energy supply in the broader region.

### ***Conflicts with the Government of Croatia***

As the Government's accordance outlined in the Gas Master Agreement the state owned company Plinacro purchased the gas storage business subsidiary (PSP Okoli) from INA at a price which was set by an internationally renowned accountancy. This was coupled with an agreement which stated that INA rents this facility for five years for an annual fee of approximately 20M Euros thus ensuring a return of the investment in 5-6 years. This agreement also contained a clause that the Croatian government would take over Prirodni Plin, the gas retail company of INA.

There was no disaccord on the leasing back of the gas storage facility but the part on the purchase of Prirodni Plin was not respected by Croatia.

Additional unfriendly steps included the Croatian Government adding important amendments to the Gas Market Act without allowing the market actors to prepare for the changes. It was not declared who would be the supplying wholesales gas to households from 2014 Spring, although this obligation was of INA for the past 30 years. With the approach of the end of the five-year agreement for the gas storage plant (PSP Okoli) it became obvious that a complete change was about to happen on the Croatian gas market. This was problematic because the in order to ensure the security

of gas supply Prirodni Plin has store a significant amount of gas in its storage facility (e.g. to be able to provide gas in the winter period, etc.)

As it was published, HEP, the power provider of Croatia was appointed as the official gas wholesales agency and was granted access to 70% of the storage capacity of PSP Okoli, all this happened without any prior notice on denouncing Prirodni Plin from its role.

This was harmful not only for the gas company but for INA, as INA's production and sales portfolio relies on being able to sell its gas, if this is strictly limited than a large amount of gas will be either not produced or will be flared.

This step was taken without prior consultation and resulted in INA being forced to sell a large amount of gas stored in the facility in a very short period of time, notably one month. This was not possible due to low seasonal demand and the limited daily pumping capacity of the storage containers.

Besides a prohibitive (and probably non EU-compatible) regulation disallowing the export of domestic gas, a large amount of gas was locked in the storage facility. Although the company tried to reason with the state regulators and stakeholders to be permitted to remove gas gradually over a timespan of several months, the requests were denied and the possibility of a forced auction was raised for the unsold and unwithdrawn amount.

Despite numerous attempts to resolve the situation and an open letter published by the management board of INA to stop the sales, the gas content was auctioned and sold for about one quarter of the market price.

The losses in terms of total revenue might be as high as 20% of INA's investment spent in Croatia, currently the figures show a loss of approximately 25M € for the 109 m<sup>3</sup> gas sold below market price. The buyers were HEP and Plinacro, both state owned companies.

The restructuring of the gas market resulted in direct losses to Prirodni Plin since it trusted that the regulatory changes would be neutral and it tried to maintain its responsibility as a gas supplier. The situation is well characterised by Želimir Sikonja, a director of INA, who admitted that "... we are disappointed with a lack of reaction from all institutions whose goal is supposed to be the protection of market supply security and ensuring of equal terms for all participants in the gas business in Croatia." The problems with INA and the conflict with the Croatian government also adversely effected MOL. Its shares fell to historical low level in 2013 when it first announced that it is ready to sell its equity in INA, the trade on the Hungarian stock exchange reacted strongly to the announcement which mentioned the debate over management right.

The plunge in share prices is well characterised by the drop of the BUX, the benchmark stock index of the Budapest burse of 1.1%.

As a reaction, the management board started preparations for the sale of MOL's share in INA in November, 2013. This demonstrates the deterioration of the connection and negotiations between the Hungarian oil company, and the Croatian government. The Croatian government's main point is that INA is operated as a subsidiary or daughter company to MOL rather than as an independent company. Experts agreed that the

management's steps were considered as an escalation of the situation. Croatia's macroeconomic situation does not allow for the buyout of the company, however.

As INA is considered to be one of the driving forces and success stories in MOL's recent acquisitions and there is no obvious project on the horizon to invest should proceeds be paid for INA by the Croatian state meaning there is no likely good outcome if selling INA occurred.

An interesting but shocking episode in the INA-MOL story is that the CEO of MOL, Zsolt Hernádi is put under an Interpol arrest warrant by the Croatian state with the accusation of bribery.

As a resolution to the case, where MOL is accused of contractual non-performance, MOL has asked for arbitration. In its statements MOL has referred to its legal obligations to represent shareholder's interests and thus in good faith negotiate. However, discussions have not led to any agreement to date.

### **Introduction to the research**

Quantitative analysis of text (news, tweets, articles, etc) can provide additional information for financial analysis. First of all text contains an additional emotional content (called sentiment) which provide valuable input for further conjectures on a given topic. Another important factor is the opinions and links found in the text to other sources. Third the quantification of some intrinsically qualitative information can be difficult and results in „signal loss”. Fourth, textual information contains some additional value over aggregated and composite quantitative information (Loughran, McDonald, 2014).

Evolution of computer hardware, computing power and storage capacity allowed for the birth and fast evolution of data mining. In addition to the vast amount of data generated every day and every hour it is possible to rely on large databases for analysing information. Dictionaries are used for providing a quick assessment of sentiments found in an article by quickly comparing the contents of the respective text with the built-in words and developing a score. One example is the Harvard Inquireer (<http://www.wjh.harvard.edu/inquirer/>) which allows for deciding on the optimistic or pessimistic nature of an article. Associative dictionaries are also a novelty, they basically function as a thesaurus and allow for establishing the proper context of a text. To visualise context some webpages offer online graphics to provide an eyecatching clue, see e.g. Visuwords or other pages.

Advantage of using such dictionaries is that they provide an unbiased fundament for evaluation, an objective basis which can be referenced and referred in research.

Many techniques exist to reduce the enormous amount of textual input to process thus simplifying analytical work. One interesting and important element is text summarisation. A simple form of summarisation is when we select the sentence(s) with the largest commonality index; that is, a number which represents the similarity between the text and other elements. One of the basic measures used is the Jaccard formula (Jaccard, 1901), which allow for the composition of the Jaccardi matrix. The



(i,j) element of the Jaccardi matrix is given as follows  $J_{ij} = \frac{|S_i \cap S_j|}{|S_i \cup S_j|} = J_{ji}$ . Similarity is

calculated by calculating row sums,  $S_i = \sum_j J_{ij}$ , and a natural ranking according to the significance of a sentence can be starting from the lowest values (e.g. highest information content or novelty).

After having done some analysis and editing tasks and having trimmed the text to our needs the next task is to analyse the text. One important step is to extract sentiments and decide about the message of the text (e.g. optimistic, pessimistic, neutral).

One method for this process is the Bayesian classification, where we use a training set to “teach” the computer to classify documents based on the occurrence of typical terms (so called prior probabilities) and use the definition of the conditional probability to calculate posterior probabilities to classify new documents in the given classes of sentiments.

Another method frequently applied is the support vector model, which separates the datasets using a distance maximisation method (e.g. distance between data groups is maximised by fitting (a) separating hyperplane(s) between).

A simple way can be the word count method where we simple count the number of words with positive and negative sentiment and get a net balance of the text.

### Applying metrics text analysis assessment

When trying to establish the quality of an algorithm in text mining, it is important to apply certain metrics. Originally metrics mean a measurement of distance in mathematics, in the current context they provide a means to test for the goodness of the text mining algorithm. Here we present only a few examples from the literature (see e.g. Das, 2014 and Das and Chen, 2007).

One important element is the confusion matrix, which describes the goodness of classification using a matrix form. Simply put, assuming a  $k$  categories, we have a quadratic  $K \times K$  matrix, where the rows represent actual categories, columns represent assigned categories, and any cell  $(i,j)$  represents a text which is category  $i$  and was assigned to category  $j$ . Obviously only elements in the diagonal of the matrix represent well classified elements, all other elements which are non-zero represent classification error (thus the notation confusion matrix).

The test is based on a  $\chi^2$  critical value, the null hypothesis that in the case of random guessing (a completely useless algorithm) the rows and columns would be independent. Denote with  $O(i,j)$  the actual elements of the confusion matrix and  $E(i,j)$  the expected element under the assumption of no classification (uniformly distributed random values, e.g. the number of observations in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column divided by the total number of observations).

$$\chi^2_{(K-1)^2} = \sum_i \sum_j \frac{(O(i,j) - E(i,j))^2}{E(i,j)}$$

Depending on this test statistics we can decide about accepting the algorithm. Based on the elements of the confusion matrix accuracy can also be tested with the following metrics using the previous notations

$$A = \frac{\sum_i O(i,i)}{\sum_j M(j)} = \frac{\sum_i O(i,i)}{\sum_i M(i)}.$$

This is simply the sum of the diagonal elements divided by the sum of all elements of the matrix.

Incorrect classification can be sometimes more harmful than no classification at all. Incorrect classification can be simply counted as the percentage of elements which are not correctly assigned (this can be weighted). A logical assumption is that the categories are arranged in a manner where neighbouring categories have proximity in their sentimental content, too. Under such arrangement it is expected that a classification which puts a given category to a category with distinctively different meaning causes much more harm than a misclassification to a category in the “vicinity”. In our proposition below we try to give a way to resolve this issue by introducing a vicinity factor in misclassification.

### **Proposed new metrics in text analysis**

One important element in text analysis is classification of text. Besides that, in our proposed method it is possible to identify the main market tendencies according to the followings.

Assume that the information from the market is organised into  $n$  documents (sources) and that at most  $k$  distinctive terms are

#### ***Suggestion for systemic error testing***

If a classification algorithm is completely precise, we would only receive elements in the main diagonal, that is, the rank of the matrix would be full (equalling the number of rows). If on the other hand we have a systemic error in the algorithm, this would mean a tendency of false classification. In this case, a category could be replaced by one or more other categories and the classification would not suffer any loss. For this we suggest a rank probe, that is to calculate the rank of the confusion matrix. If the rank is lower than the order of the matrix ( $k$ ) that means that one category can be reproduced as a linear combination of other (one or more) categories. In that case, the algorithm is generating systemic, inherent errors. If the rank of the confusion matrix is full, then the algorithm contains only random errors.

#### ***Communality matrix and determination of principal vectors in news***

Concerning miscategorisation as a grave error it is logical to identify a measure to deal with this problem. Assume that categories are assigned in a logical order (e.g. decreasing sentiment, etc.) and that the the algorithm is not degenerative, that is the  $K$  confusion matrix is full rank. In that case it is possible to apply linear transformation and gain the Jordan canonical form (Molnár and Szidarovszky, 2002). In that case

there exist at least one real eigenvalue of the matrix, but more importantly the basis of the Jordan-form matrix is composed of the eigenvectors of the original matrix (transformation, or in our case classifying algorithm).

A measure for the degree of miscategorisation can be a simple euclidean distance of the standardised eigenvectors. If the distance is less than a given threshold, then the categorisation can be accepted. If the distance is very large than the algorithm can be considered risky from the aspect of miscategorisation.

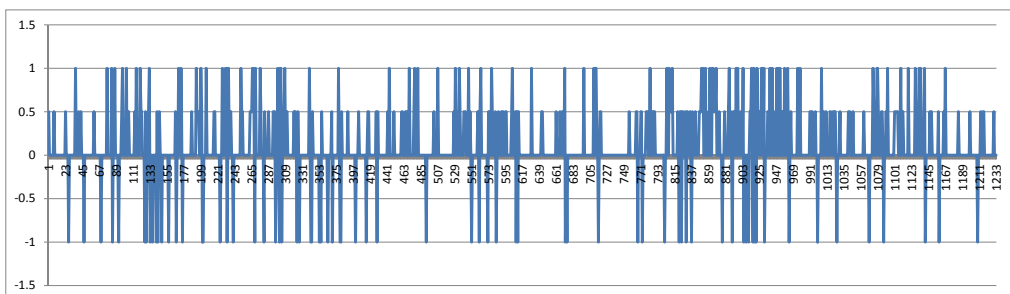
### Empirical results of news analytics

Some elements of the theoretical results were applied to a specific case along the following lines. The Hungarian Oil Companies (MOL) and the Croatian Industrianafta (INA) formed a strategical alliance in 2003 and MOL became the owner of almost 50% of the INA shares. In our short analysis we analysed approximately 850 articles from Hungarian websites (primarily, portfolio.hu). These articles were grouped into three categories based on keyword assessment: bearish (pessimistic), bullish (optimistic) and neutral, and were scored accordingly. In many cases the articles were of political nature and thus had additional layers of information. In cases where multiple messages (perhaps of mixed positive and negative nature) were found the overall aggregate value was considered for that day. This was matched with the daily movement of MOL share prices on the Budapest Stock Exchange (BUX).

These results were combined in simple difference values as follows. If the information gained from news analytics (three discrete values were possible) were matching the daily movement of prices then we assigned a +0.5 value to the forecast.

If the information derived from the analysed news were different from the share price movement we generated a +1 or -1 value depending on the direction of share price change compared to the forecast. The information is summarised in Figure 1.

*Figure 1. Differences between information of analysed news and share price movements, no lag*

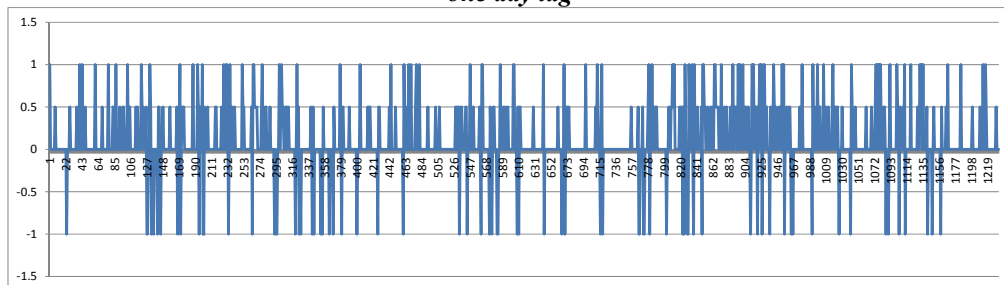


*(1=share price increase w. negative forecast, 0.5 identical movement, correct forecast, -1=share price decrease with positive forecast)*

Source: own calculations

It is well visible that the news analytics performed only partially well in forecasting price movements. As this was the contemporaneous (daily change) it is worth to check for the lag phase behaviour of the forecast. This is shown in Figure 2.

**Figure 2: Differences between information of analysed news and share price movements, one day lag**



(1=share price increase w. negative forecast, 0.5 identical movement, correct forecast, -1=share price decrease with positive forecast)

Source: own calculations

From this chart it can also be observed that a given day prediction from news analysis typically resulted in the next day share price movement following the sentiment of the news text.

## Remarks

Although the above results are of limited scope they show that news analytics require increased attention both from the theoretical view and from the view of technical analysis. There is evidence that the market is not fully informed, at least that that full information principle only holds in a weaker form as news analysis is able to provide additional predictive abilities.

## References

1. Das S. R. Text and Context (2014): Language Analytics in Finance. Foundations and Trends R in Finance, vol. 8, no. 3, pp. 145–260.
2. Das S. ; Chen M. (2007): Yahoo for amazon! sentiment extraction from small talk on the web. Management Science, 53:1375–1388.
3. Jaccard P (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bull. Soc. Vaudoise Sci. Nat. 37: 547-579.
4. Loughran T.; McDonald W. (2014): Measuring readability in financial disclosures. Journal of Finance, 69:1643–1671.
5. Molnár S., Szidarovszky F. (2002): Introduction to Matrix Theory with Applications to Business and Economics, World Scientific, London.
6. <http://www.bloomberg.com/news/articles/2013-11-26/mol-seeks-arbitration-over-ina-as-dispute-escalates>
7. <http://molincroatia.com/inas-investments>
8. <http://molincroatia.com/mols-investment-in-ina>
9. <http://www.bloomberg.com/news/articles/2013-11-11/mol-dives-as-sale-plan-escalates-ina-dispute-budapest-mover>