ABSTRACT

Company's stock prices may depend on the information from different domains. It may be financial news, related to different topics, discussions on social networking sites, tweets etc. In this research, we are trying to divide the informational background into topic groups, study the difference of their impact on market volatility and reveal the most contributing informational patterns.

Keywords: Financial news, Tweets, Stock market, Sentence-Transformer, BERT, Topic classification, Clustering, PCA, k-means, Sentiment analysis, Correlation, TFT

The object of research is financial news, divided into topic groups, and tweets.

The subject of research is news and tweets impact on public companies' stock prices.

The goal of research is to reveal the most influential thematic groups of information on the stock market.

The main problems of the research are:

- Gathering valuable information (financial news and tweets about public companies with company's tickers)
- Splitting gathered information into topic groups
- Annotating news from each group with sentiment labels.
- Gathering information about companies' historical stock prices.
- Performing experiments for evaluating the interrelation between informational topic groups and historical stock prices.

- Performing experiments for evaluating training metrics of the DL model change for defining news topic groups, which can help improve model performance as external data.

The **contributions** of this paper are as follows:

- Performed news topic classification using pre-trained models
- Investigated the application of clustering method to news topic classification
- Evaluated the time series similarity metrics between historical stock prices and informational background sentiments
- Found out if it makes sense to split news into topic groups to improve model predictions' quality.

Definitions and Abbreviations

Tweet - a message in Twitter's blog.

Stock market - a place where shares of publicly owned companies can be bought and sold.

Market volatility - the degree of a trading price series variation over time.

Machine learning - set of data analysis methods, which give the possibility to train analytical models by solving similar problems.

Supervised learning - a type of machine learning, in which the algorithm is trained on a labeled dataset, in other words, a target variable is known.

Classification - a supervised learning technique, which represents the process of grouping objects into predefined categories.

Clustering - the task of dividing the set of data points into groups, so that the points in the same group are more similar to each other and more different from the data points in other groups.

Sentiment analysis - the process of classifying, which defines whether the part of the text is positive, negative or neutral.

Time series - a series of data points, indexed in a time order.

Time series similarity - (in this work) it is the distance between two time series and their correlation strength.

Deep learning - a subset of machine learning, which works with artificial neural networks.

Transformer - a deep learning model that adopts the mechanism of self-attention, which weights the significance of each part of the input data.

Attention mechanism - a technique used in recurrent neural networks (RNN) and convolutional neural networks (CNN) to search for relationships between different parts of input and output data.

Training set - the set, which is utilized for model training.

Testing set - the set, which is utilized for model performance evaluation.

Encoder - the part of the model, which transforms input data to a set of features.

Decoder - the part of the model, which transforms feature representation of the object to understandable language.

Hidden state - the output of the encoder.

Loss function - the function, which reflects the difference between target variable and predictions of the target variable.

Hyper-parameter - a parameter, which is declared during the model creation and is used for learning process control.

Multiple time series forecasting - forecasting target values for several time steps ahead.

Contents

1 Introduction

The connection between a company's stock market and company's news and events has always been one of the hot-button issues for discussion. This area of investigations, represented in our paper, is applied in a wide range of social and economic spheres. For example, the condition of big pharmaceutical companies' markets depends on the recent announcements of clinical trials, as they are intently monitored by the public [9].

More than that, in recent years, the growth of online trading popularity has accelerated the process of accessing the informational background. Significant number of traders started to utilize such sources of information as news sentiments, which are produced by computer algorithms, which can quickly show whether the news article or post in Twitter is positive or negative.

The main question is what kind of news should be taken into account for obtaining a better vision of companies' future market trends. In the current study we are going to find out how to classify financial news and tweets according to their topic domains and figure out news sentiments from which group mostly correlate with public companies' stock markets. Moreover, in this paper we investigate whether it makes sense to divide news into topic groups, when passing it to the deep learning model for stock prices time series prediction as external data. We evaluate whether some topic groups give the possibility to obtain better model performance or if it is still better to pass to the model all the information without any division. For performing experiments we have chosen 5 big technological public companies: Apple, Amazon, Google, Netflix, Tesla.

The main goal of our research is not to develop the ideal way of predicting stock prices or to work out the universal advice, which will guarantee the success for all investors, but to investigate the market and search for insights, connected with its main processes and behavior. In general, this study would be important and interesting not only for investors, but also for market researchers, data scientists and people, who are interested in better understanding of market trends.

2 Literature review

In the next paragraphs we are going to summarize articles about the recent studies, which included ideas, interesting for our research.

2.1 New drugs and stock market: how to predict pharma market reaction to clinical trial announcements [9]

In this paper authors describe the research, which was connected with forecasting the market value change for pharmaceutical companies, according to announcements of clinical trial results. Taking into account the pharmaceutical industry's peculiarities, researchers statistically proved the reasonableness of announcement impact evaluation, developed a framework for efficient preprocessing of clinical announcements and obtained a high quality prediction for the range of pharmaceutical companies' stock prices, by combining the gradient boosting (GB) classifier with graph convolutional network (GCN).

2.2 Stock market prediction using machine learning classifiers and social media, news [12]

The main idea of the research, described in this paper, is measuring sentiment of news from two different domains: social media news and financial news. The news sentiment series for each day is computed using the Stanford NLP package. The series contains the overall sentiment, calculated for each observation day. The series were formed for both Social media and Financial news and added to 3 different models as features: model with only social media news, model with only financial news and model with both types of news. The main goal was to compare 12 machine learning algorithms for classification on mentioned models. The results showed that for models with social media features the Random Forest performed best on the independent testing dataset. For models with only financial news features the Random forest was also the best. The performance of models with both news features is the worst in comparison to other models for all algorithms. In conclusion, the authors summarized that the best classifier is Random Forest because it showed the highest prediction accuracy in training 2 models and its prediction accuracy improves after feature selection and spam reduction.

2.3 Estimating the impact of domain-specific news sentiment on financial assets [4]

In this survey Stephen Kelly and Khurshid Ahmad investigated how stock prices depend on domain-specific news sentiment. They took into account two significant markets, equity and oil, and evaluated two important financial assets: Dow Jones Industrial Average (DJIA) and West Texas Intermediate (WTI) crude oil. Related news for investigation were collected from websites, blogs and online-databases. Then the text corpus was formed out of gathered information. Each text in the corpus was tokenized and, using psychological and financial dictionaries, containing words, divided into categories, according to their tone and sentiment, words' frequencies were count for each category and different periods of time, so that it could be possible to count the sentiment time series and negative words distribution. Sentiment time series vector was taken to the Vector Autoregression model formula as a feature, For evaluation of financial series change the Rolling Regression was used. More than that, for calculation of sentiment variable statistical power the hypothesis test was performed, in particular, the z-score was computed for sentiment variable in each rolling window,

2.4 Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms [13]

In this report qualities of stock predictions, made by different algorithms, are compared. For described in the paper experiment four ML and DL algorithms were chosen. Among them there are Artificial Neural Networks, Support Vector Regression, Random forest and Deep learning. The following results were obtained: stock prediction is a complicated problem, because its behavior is non-linear in the majority of cases. The best benchmarks were shown by the LSTM model, which means that we can suppose Deep learning works better than Machine learning.

2.5 News-based sentiment and bitcoin volatility [10]

This paper describes the investigation in which researchers used several modifications of the HAR-RV model for bitcoin volatility prediction. Using the keyword "Bitcoin" the author downloaded bitcoin-related news from major English speaking newspapers and formed the news corpus. Furthermore, the information about Google search trends was gathered with the R package called gtrendsR. Then the Sentiment Analysis package in R was used to supply positive, negative, and overall sentiment scores for four different sentiment dictionaries, which next were applied to the news corpus. The HAR-RV model takes such benchmarks as realized volatility as a feature. For model evaluation the author added daily RV, weekly RV and monthly RV to the model. This model was trained, its performance was evaluated and conclusions about features' significance were made. Then the researcher added sentiment variables to the model. Training results showed that model performance was improved after adding sentiment features.

2.6 Event Clustering within News Articles [5]

This article presents the special approach of news clustering, which consists of 3 stages: 1) Grouping sentences into pairs and predicting, if sentences in each particular pair describe the same event. The prediction was realized using the BERT and ALBERT model. 2) After the first step authors got the set of pairs with 0 or 1 score. The main idea of the second step was predicting the score for each element in the pair, imagining if it was grouped in pairs with each of all the other elements. For example, if we have the pair (s_i, s_j), then for all s_k, which are not equal s_i and s_j, we count the score of (s_i, s_k) and (s_j, s_k). If the scores of these two pairs are equal, then we can increase the score of pairs (s_i, s_j), as their probability to be in one cluster increases. If the scores are not equal, then the score of (s_i, s_j) should be decreased. 3) Grouping into clusters: the higher the score of the pair, the higher the probability of relating to the same cluster. Summarizing the research, BERT and ALBERT models were compared, and ALBERT model outperformed. Also proposed approach was compared with the baseline Correlation Clustering and this experiment showed that proposed by the researchers approach performs best results.

2.7 Unsupervised News Topic Modelling with Doc2Vec and Spherical Clustering [17]

This report pays attention to spherical clustering with doc2vec modeling. It is a method, which is very similar to k-means, but as a distance metric it takes Cosine distance instead of Euclidean distance. Authors tried to split the data at 11 previously known clusters. In the result, researchers got the list of key words for each cluster, and the list of dominant clusters for each document.

2.8 Clustering News Articles for Topic Detection[3]

This report presents a way of clustering news so that it could be easy to model their topics. The central idea is using hierarchical agglomerative clustering algorithms. After text preprocessing, tokenization and stemming steps of the algorithm are the following: defining each token as one cluster, then counting distances between all clusters, uniting two nearest clusters, repeating mentioned above actions until there is only one cluster, which contains all tokens. After applying this algorithm researchers got the clustering tree. As news often can be related to several domains, which means that clusters can be dependent on each other, the hierarchical tree of clusters is the most appropriate model for topic modeling. More than that, the tree can help us, if we do not know the exact amount of clusters. We can define it by ourselves, cutting the tree at a particular height.

2.9 A survey and an experimental comparison of methods for text clustering: application to scientific articles [1]

In this paper [6] several clustering methods were investigated in application to text documents and scientific reports. The process of clustering consisted of three stages: text preprocessing, text vectorization and vectors' clustering. For clustering performance evaluation the AMI (Adjusted Mutual Information) metric was calculated for each vectorization method and each group of data. According to the results of experiments the best method of clustering was k-means with Paragraph Vectors vectorization for all documents. Paragraph vectors method uses a neural network model, which can predict the words of the document by its vector. Apart from k-means, agglomerative and spectral clustering were also taken into account. More than that, the authors decided not to pay attention to the DB-SCAN algorithm because preliminary testing showed that a lot of objects were not associated with any clusters by DB-SCAN method. In other words, it was creating too much noise.

2.10 Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting [14]

This paper proposes the Temporal Fusion Transformer architecture, which is based on an attention mechanism and can be applied for multi-horizon forecasting. The main parts of the TFT are: Gating mechanism - provides the adaptive network architecture and gives possibility to accommodate a wide range of datasets, Variable selection networks - for selecting relevant input variables, Static covariate encoders - for static features integration, Temporal processing - consists of a sequence-to-sequence layer and multi-head attention block, Prediction intervals - quantile forecasts to obtain the range of most likely target values. On different simple and complex datasets the researchers show that the TFT model shows state-of-the-art prediction quality.

3 Methodology

3.1 Problem Setup

The main problem, which we are going to work on, is defining which news topic groups, connected with a particular public company, correlate mostly with the company's historical stock prices. The results of the research could help improve our understanding of the market and its trends dependency on different kinds of information. More than that, acquired knowledge could be applied for defining, which information should be taken into account during analyzing and predicting future trends of public companies' markets.

3.2 Data labeling

In this subsection, we discuss the major parts of the pipeline for splitting financial news into topic groups and making sentiment extraction of news and tweets for subsequent evaluation of the correlation between historical stock prices and informational background. The schematic representation of the pipeline is presented in Figure 1. The first stage focuses on financial news topic classification. It is necessary to note that we apply classification only to financial news, as all data from twitter we relate to a separate group, taking into account the fact that among tweets there is not only financial information about the company.

After getting topic labels for all financial news, we split the dataset into groups, so that in the same group there is news with the same topic labels. The news groups that we got at the previous stage and tweets, which we previously gathered, become input to the second stage, which concentrates on news and tweets sentiment labeling. After processing the second stage we get several groups of data about the company (including twitter as a separate group), consisting of news or tweets with sentiment labels: positive, negative, neutral.

3.3 Topic classification

One of the most significant problems, which we aimed to solve in this research, is studying how to divide the big collection of news and tweets into topic groups. We tried two different machine learning approaches to formulated tasks: clustering and topic classification (according to preliminary defined topics). The results of the clustering are not satisfying, it will be described and shown in further subsections. In these subsections we are going to focus on the topic classification task and its solution.

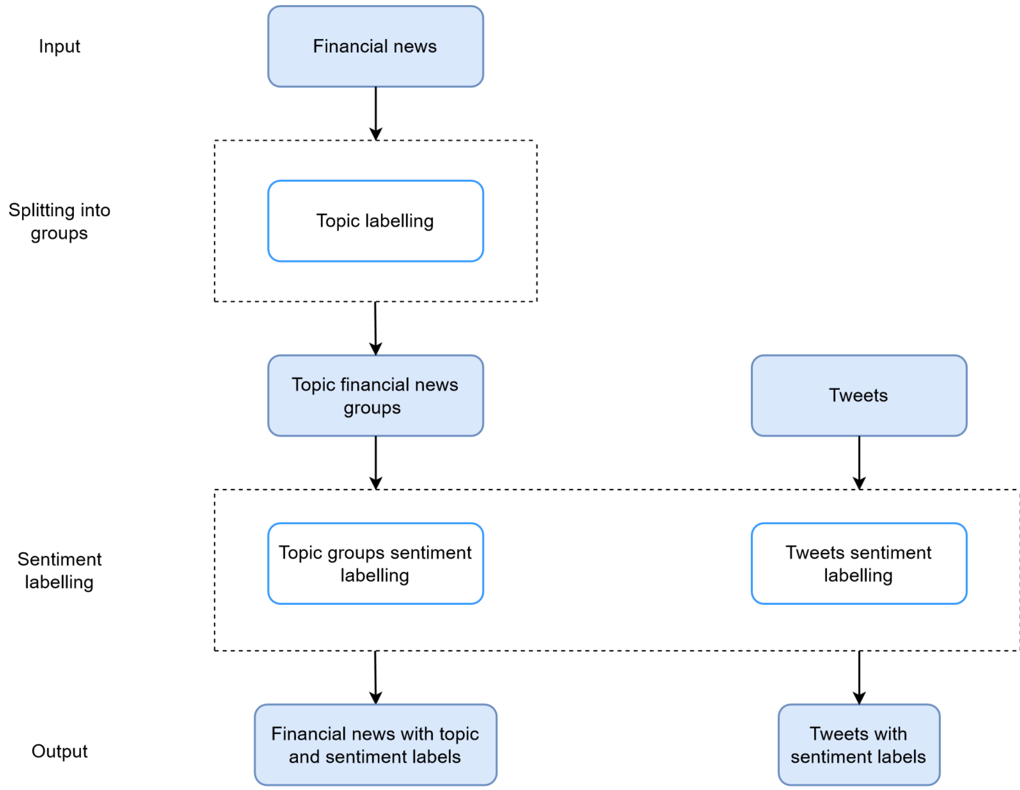

Figure 1: High-level pipeline of data labeling for further experiments. The process of labeling consists of financial news topic labeling according to the list of 20 topics and sentiment labeling of news from each financial topic group and tweets

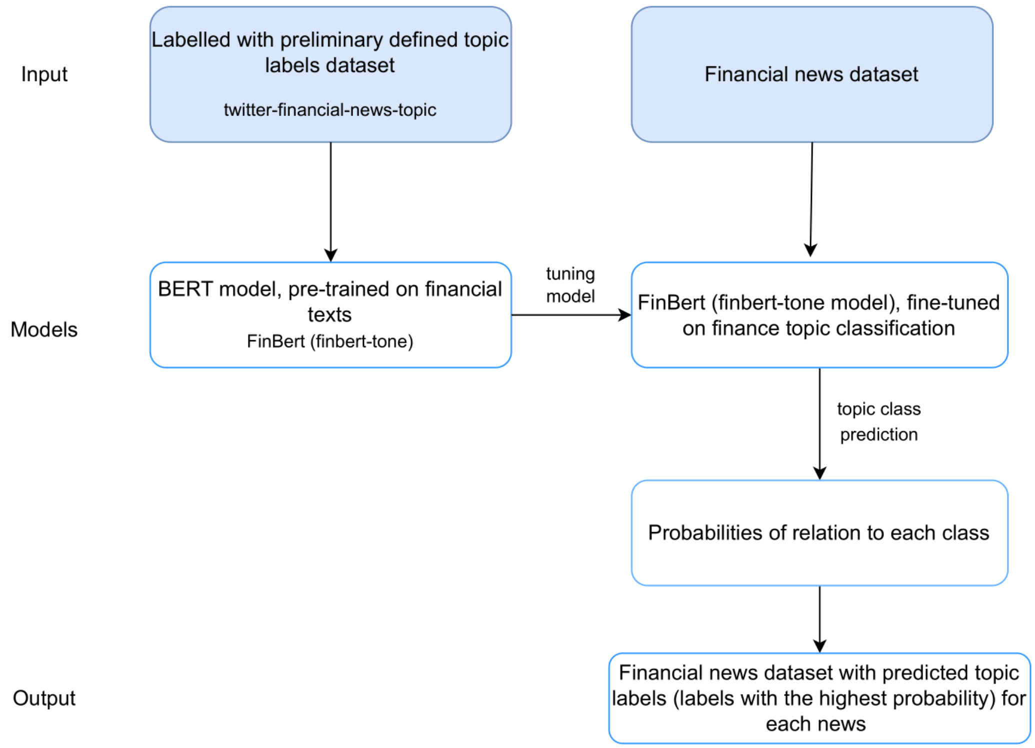

Figure 2: The schema of getting the model for topic labels prediction and making predictions for our dataset with financial news

In Figure 2, the logic of our approach to the classification is represented.

The sequence of steps is the following:

1. The first problem was to find the dataset with labeled financial texts. The dataset, which mostly matched our requirements and task features is twitter-financial-news-topic [15]. It is an english-language dataset, which consists of tweets, related to financial aspects and news.

   The documents in the dataset are labeled with 20 topics:

   - Analyst Update
   - Fed | Central Bank
   - Company | Product News",
   - Treasuries | Corporate Debt
   - Dividend
   - Earnings
   - Energy | Oil
   - Financials
   - Currencies
   - General News | Opinion
   - Gold | Metals | Materials
   - IPO
   - Legal | Regulation
   - M&A | Investments
   - Macro
   - Markets
   - Politics
   - Personnel Change
   - Stock Commentary
   - Stock Movement

2. Then there was a task to find the model for training on chosen dataset or already pre-trained on annotated data model. We managed to find the model, which was fine-tuned on the described labeled dataset. The model is "finbert-tone-fine tuned-finance-topic-classification" model [6]. Figure 3 illustrates how this model was created. First the Bidirectional encoder representations from transformers (BERT) were taken. Then this model was trained on corporate reports, earnings call transcripts and analyst reports, in other words, on financial communication texts. After that the FinBERT model was released. The next step was FinBERT fine-tuning on 10000 sentences from analyst reports, the sentences were preliminary annotated with sentiment labels: positive, negative, neutral. After this stage the finbert-tone model was released. The final stage was training the finbert-tone model on a previously described dataset. The accuracy of the model is 0.910615, which made me conclude that this model is appropriate for topic classification tasks.

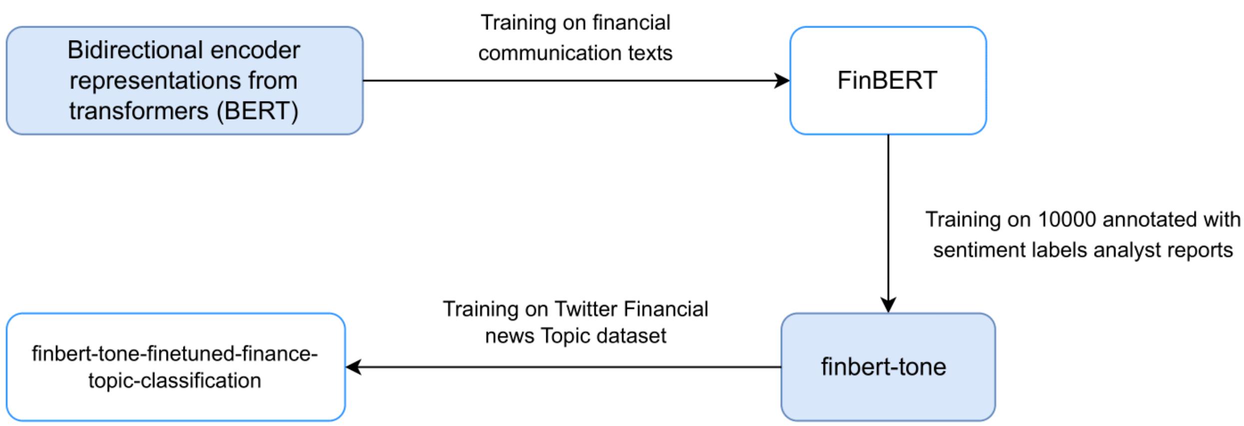

Figure 3: The schema of topic classifier creation

3. The final stage was applying a described model to gather data. As the model was pre-trained only on financial texts, and in twitter dataset there were not only financial news about companies, but also customers' opinions, discussions etc. There was a lot of noise in model predictions, so we decided to apply the model only to financial news and consider tweets as a separate group.

3.4 Sentiment labeling

At the previous stage, which was connected with financial news topic classification, we got 20 topic groups of financial news and the collection of tweets about Big Tech companies. In this subsection we will describe the way how derived topic groups and tweets were annotated with sentiment labels. In Figure 4 there is a description of datasets and models, which were applied to these datasets. For financial news labeling I utilized the finbert-tone model [7], which was described in the previous subsection. We applied this model to each group of news and as the output we obtained 20 topic groups of financial news with sentiment labels (positive, negative, neutral), assigned to each of the news.

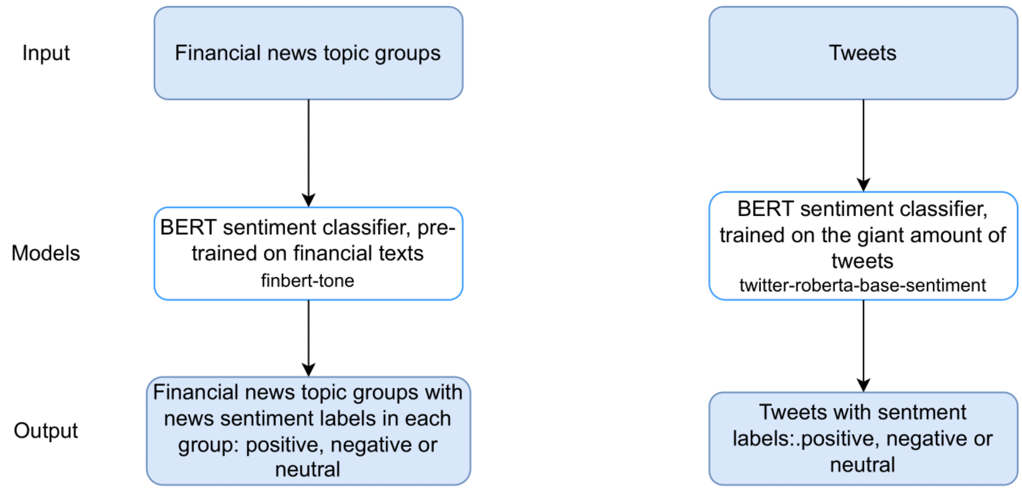

Figure 4: The schema of financial news and tweets sentiment labeling

For tweets sentiment analysis we took the Twitter-roBERTa-base model [16], which was pre-trained on 58 million tweets and then was fine-tuned with TweetEval benchmark for sentiment analysis. After applying this model to twitter collection, we obtained tweets, annotated with sentiment (positive, negative, neutral).

3.5 Clustering

Another approach, which we tried to utilize for splitting financial news into topic groups, is clustering. In this approach the topic groups are not preliminary defined, and splitting is conducted according to the distance between news headlines, which are transformed to vector representations. In Figure 5 there is a pipeline, which reflects the clustering process. The preparation for the clustering was in two steps: transforming to vector representation and dimension reduction. For forming vector representation of news headlines we have chosen the sentence-transformer model all-MiniLM-L6-v2 [11]. This model maps sentences to 384-dimensional vector space. The model was trained on several concatenated datasets, which contained over 1 billion sentence pairs. The main idea of the training method was evaluation of cosine distance between each possible sentence pairs in the batch and applying cross entropy loss, which compared this metric value with true pairs' metrics. This model was downloaded by about 3 million HuggingFace visitors, which proves its reliability.

To reduce the noise and make data representation easier, dimension reduction methods are usually applied. As we are working with a big volume of data, It was important to find the method, which could reduce the dimension fast and efficiently. The method we tried was Principal component analysis. The PCA approach is quite geometric. Its main idea is projecting objects on an axis with the biggest data variance along them. So, the PCA method is more preferable, when it is necessary to simplify an enormous dataset with high complexity.

Finally, as a clustering algorithm, we utilized the k-means algorithm, because in previous research [1] It was claimed as the best algorithm for text clustering.

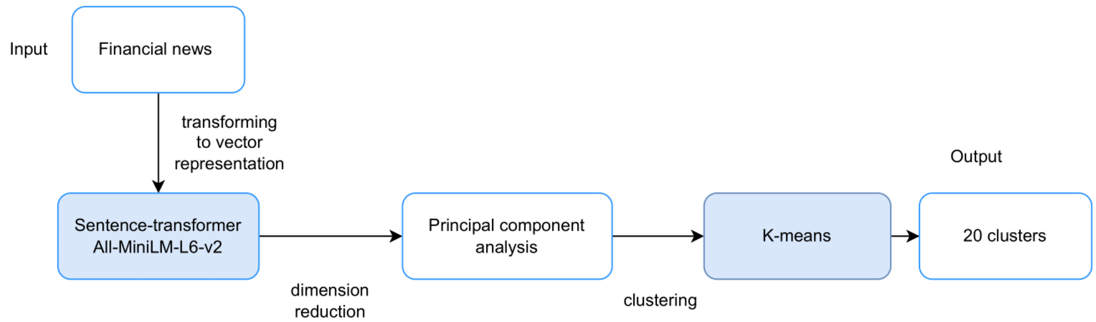

Figure 5: The pipeline of data preparation and clustering process

3.6 Time series similarity evaluation

At the current stage of research, we studied the correlation metrics between topic news sentiments and historical stock prices of 5 public companies. The algorithm of the experiment is the following for each topic group:

- Grouping news and tweets according to particular time intervals (days, months).
- Evaluating the positive news ratio for news or tweets in each group at each period of time.
- Downloading company's stock prices from yahoo finance.
- Grouping stock prices by the time interval, by which the sentiments were grouped in previous points.
- Evaluating the mean stock price for prices, which were observed at each time period.
- Smoothing and scaling the data.
- Evaluating chosen metrics between positive news ratio and mean stock prices or stock prices volatility.
- Analyzing the results of each topic group and comparing them between each other.

As in financial news dataset, the time period is 12 years and there is a very small amount of news published in one day and even in a week, the noticeable results in correlation experiments for financial news were obtained when grouping news by months and evaluating the relation of positive news amount to the overall amount of positive and negative news amounts, comparing it with the monthly situation on company's stock market, expressed in the mean value of all prices per month.

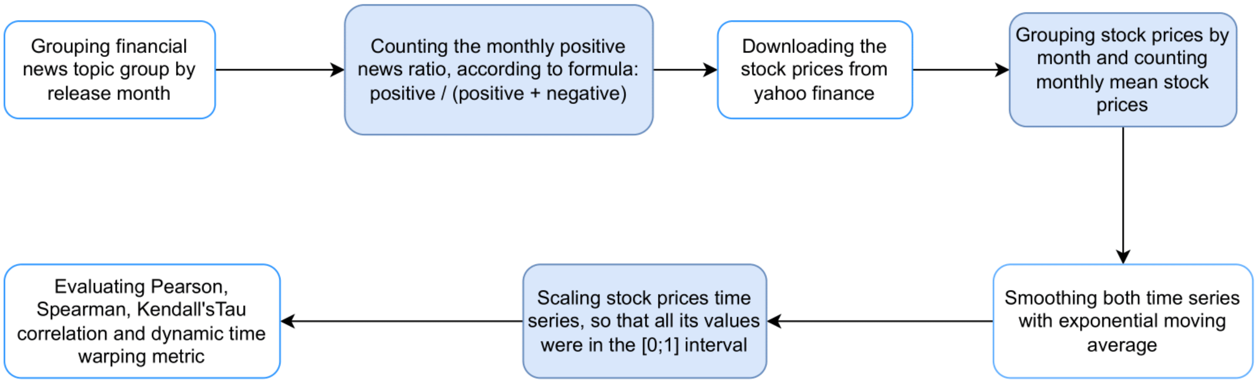

Figure 6 The pipeline of time series similarity evaluation experiment for financial news

The correlation experiment for the Twitter group was performed in another way, as the overall time period in Twitter dataset is about 2 months, and there are a lot of tweets, released in one day. So, for this experiment we grouped the tweets by the release date and tried to match daily tweets' sentiments trends with changes in daily stock prices and, besides the correlation between daily positive news ratio and daily stock prices, for tweets we measured the correlation of daily stock prices volatility with daily news amount and daily positive news ratio.

We decided to evaluate the daily news amount, besides the positive news ratio, as the daily amount of tweets is quite big, compared with the daily amount of financial news. Moreover, it is necessary to mark, that it is the main reason, why for financial news there is no sense to measure the daily amount, but it is much more important to take into account the monthly positive news ratio to understand the level of positive sentiments in the informational background and notice the main changes on the stock market.

Moreover, as the time interval for tweets is measured in days, we evaluated the daily stock prices volatility. Volatility is the degree of variation of a trading price series over time. It is also more representative, when we speak about small time intervals, such as days, while it is not accurate to measure volatility of monthly mean stock prices, which is why we did not use it for financial news. The formula we utilized for volatility evaluation is the Average True Range (ATR) formula. For calculation of the ATR for each day, first, it is necessary to calculate the True Range (TR). The formula for TR looks as follows:

$$TR = max[(high - low), abs(high - close_{prev}), abs(low - close_{prev})],$$

where *high* - the most recent period's (day in our case) high price, *low* - the most recent low price, $close_{prev}$ - the previous period close price. The ATR for the current day is calculated according to the Simple Moving Average (SMA) formula:

$$(\frac{1}{n})\sum_{i=1}^{n} TR_i,$$

where *n* is the number of periods from the beginning of the time series to the current day, including this particular day.

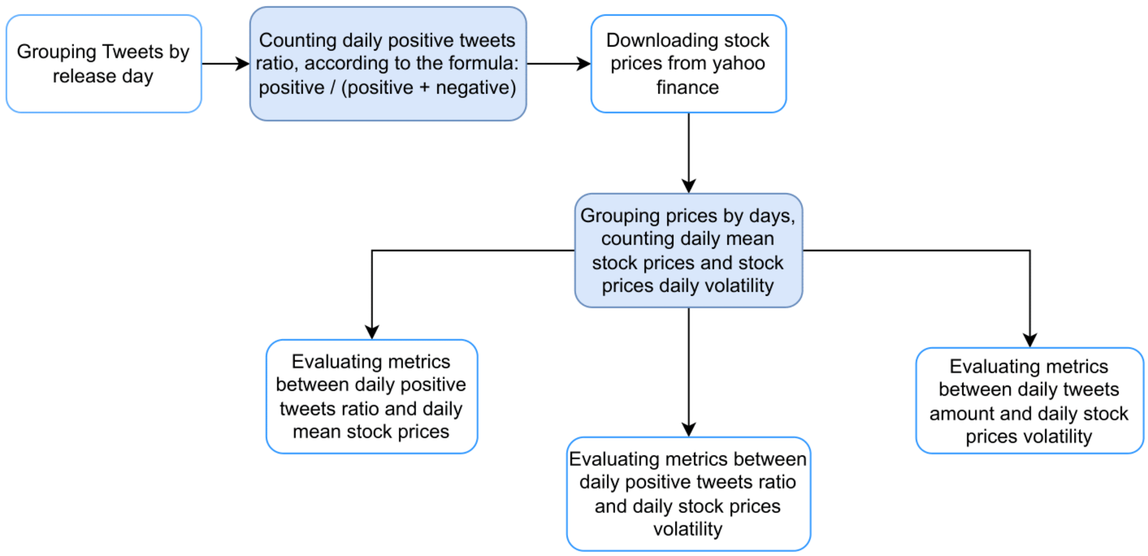

Figure 7 The pipeline of time series similarity evaluation experiment for financial news

To filter noise and produce more smooth data, which contains only main trends of the background sentiments' and stock prices' dynamics, we utilized the Exponential Weighted Moving Average (EWMA). The calculation of the EWMA is described by the following equations:

$$EWMA_1 = value_1$$

$$EWMA_t = const * EWMA_{t-1} + (1 - const) * value_t$$

Summarizing the above recursive equations, we get the resulting formula:

$$EWMA_t = (const^{t-1} * value_1 + const^{t-2} * value_2 + const^{t-3} * value_3 + \ldots + value_t) * (1 - const),$$

where $value_t$ is observation at the moment t.

The Exponential Weighted Moving Average gives bigger weights to the most recent observations and smaller weights to earlier values. This type of moving average is commonly used to smooth out fluctuations and to highlight the main trends.

3.7 Time series predictions

This stage of research was connected with studying a change of time series prediction quality after adding an external data to a deep learning model. The model, which we finally utilized for experiments, was Temporal Fusion Transformer (TFT), as, according to the paper [14], this model outperforms all existing Deep Learning models for time series forecasting. It has an attention-based architecture, which provides high-performance multi-horizon forecasting. The main feature, by which the TFT model stands out among other solutions, is its ability to efficiently build feature representation of each input type and provide qualitative forecasting performance on a big variety of problems. The second advantage of the TFT model is an ability to produce multi-horizon predictions, which give users a possibility to make decisions, thinking multiple steps ahead.

As an implementation of the TFT architecture, described in the paper, we utilized a realization from the Darts library [16], as its interface is simple and understandable.

3.7.1 TFT model parameters

Parameter name	Parameter description	Parameter value
input_chunk_length	number of past time steps that are passed to the forecasting module to predict next time steps	15
output_chunk_length	number of future time steps that should be predicted	3
hidden_size	TFT hidden state size	64
lstm_layers	layers for the Long Short Term Memory Encoder and Decoder number (default is 1)	1
num_attention_heads	attention heads number (default is 4)	4
dropout	fraction of neurons deactivated during dropout.	0.1
batch_size	number of time series used during each training iteration	16
n_epochs	number of training epochs	30
add_relative_index	this parameter gives possibility to use the TFTModel without future_covariates	True
loss_fn	loss function for training	torch.nn.MSELoss()
random_state	parameter, which indicates the randomness of the weight's initialization	42

Such parameters as "input_chunk_length" and "output_chunk_length" were chosen, basing on the argumentation that model performance is better, when the forecast horizon is not big, while the optimal length for input data is the period about two weeks, during which the sufficient amount of events happen. The "n_epochs" parameter was hand-picked for better metrics achievement. The “add_relative_index” parameter should be set to True, as we use past_covariates without future_covariates in our model. The “loss_fn” parameter will be described in the next paragraph. Other parameters were set by default.

3.7.2 Loss function

As a loss function, in the Temporal Fusion transformer we utilized the MSE loss function. It is a criterion, which is based on calculation of the Mean Squared Error (MSE). The function can be described by the following formula:

$$l(x,\ y) = \{l_1, \ldots l_N\}^T, \quad l_n = (x_n - y_n)^2,$$

where x - the input, y - the target, N - the size of the batch.

MSE significantly fines predictions that are considerably different from the actual values by applying a squared operator, but there is also the reverse side of the coin, as it can exaggerate the loss for outliers, which are not so important for us.

By default, the TFT model is probabilistic and uses the likelihood parameter. For getting deterministic forecasts, the documentation recommends using the PyTorch loss functions, and, in particular, the MSE. As in our time series data, there are not a lot of outliers, it is appropriate to use this loss function in our model.

3.7.3 Data processing

For data preparation, we applied the following data transformers from the Darts library:

$MissingValuesFiller$ - the data transformer, which automatically fills missing values using the pd.Dataframe.interpolate() method. The interpolation technique was ‘linear’, which means that all values are treated as equally spaced.

$Scaler$ - transforms the data so that all the values were between 0 and 1.

The overall algorithm of data preparation before passing it to the model:

- Making a time series from a dataset with historical prices, the frequency of the time series is ‘B’ (business days, without weekends).
- Filling missing values in stock prices time series
- Scaling the values in stock prices time series
- Making a time series from a dataset with sentiments, pointing the frequency of the time series as ‘D’ (all days)
- Filling missing values of time series with sentiments
- Scaling sentiments time series
- Making dataframe from obtained sentiments time series
- Transforming dataset into time series with frequency ‘B’
- Cutting the stock prices and sentiments time series, so that they covered the same periods of time.

3.7.4 Training

In the beginning of the process the stock prices time series is divided into two parts: the training set and the test set. The train set is passed to the fit() model. More than that, there is another extremely important parameter, called covariates. In the Darts library, covariates are the external data, which are passed to the model to help improve the predictions. Covariates may be future, past and static. Past covariates are known only into the past, future covariates are known into the future, static covariates are constant over time. As news sentiments (positive news ratio) in real life are known only into the past, the sentiments, as external data, we pass them to past_covariates parameter.

3.7.5 Historical forecasts

To compute the forecasts at multiple time steps, we applied the historical forecasts method.

For multiple forecasting, this method uses an expanding training window. At the first step it takes the window from the beginning of the time series to the time point, which is passed in the "start" parameter (if "train_length" is set to None). Then it trains the model on the training set and produces forecasts for the future period, whose length equals the "forecast_horizon" parameter, then moves the right end of the training set forward by stride time steps, while the left end is fixed at the beginning of the time series.

To make the method return all predicted points, it is necessary to set "last_points_only" parameter to False. Then it will return a sequence of lists of the historical forecasts series.

By default, this method always retrains the models on the entire available data, applying an expanding window strategy. If the "retrain" parameter is set to False, the model will only be trained on the initial training window (up to "start" timestamp), and only if it has not been trained before. For prediction of stock prices with all groups of sentiments apart from "Twitter", we set the "retrain" parameter to False, but as the period of time, covered by tweets, was too short and the prediction quality was poor, we decided to let the model being retrained, while making historical forecasts. As the "start" parameter, we pass the start of the test part of the time series. To the forecast horizon we give the value 3 - the same value, which we passed as the "output_chunk_length" parameter to the model when creating it. The stride parameter is equal to the forecast horizon, so that the method predicts 3 points, then the training set end moves forward by 3 timesteps, and the method predicts next 3 points, this process repeats until the end of the series.

3.7.6 Experiment with TFT predictions

The process of experiment consists of several aspects:

- Training TFT model without sentiments as past covariates (as it is prohibited not to use any covariates, at this stage we pass months and years as past covariates), predicting the values, counting the MAPE between obtained values and values from the test set.

- Creating a new TFT model, adding sentiments to the covariates stack and training this model with sentiments, evaluating the MAPE metric.
- Evaluating the MAPE change in percentage.

4 Experiments

4.1 Datasets

The general dataset, utilized for the trial, was constructed from 3 different datasets, consisting of financial news, connected with 802 publicly traded US companies, and tweets about 11 big technological companies. The general dataset structure and datasets, from which it was formed, will be described in the following subsection.

4.1.1 Dataset description

- **Financial news archive [8]**

This dataset was collected from investing.com, bloomber.com, seekingalpha.com, 247wallst.com, zacks.com and cnbc.com websites. The data was preprocessed by removing images, graphics, ads boxes and punctuation. The data comprises financial news of more than 800 publicly traded companies. The number of samples is 221513. The description of the dataset's field is introduced in Table 1. Moreover, as we are going to work with time series, it is important to visualize amounts of news, released during different periods of time. In other words, to represent the distribution of news release dates. The described visualization is presented in Figure 8.

Data properties		
Attribute name	Description	Type
id	auto-incremented value, which shows the order in which the articles were collected	int64
ticker	an abbreviation, which helps to identify the stock of a particular public company	object
title	news headline	object
category	news/analysis or opinion	object
content	news article text	object
release date	date of news release	object
provider	the author or source of the content	object
url	link to the original source	object
article id	unique identifier of the article	int64

Table 1: Financial news archive dataset properties description

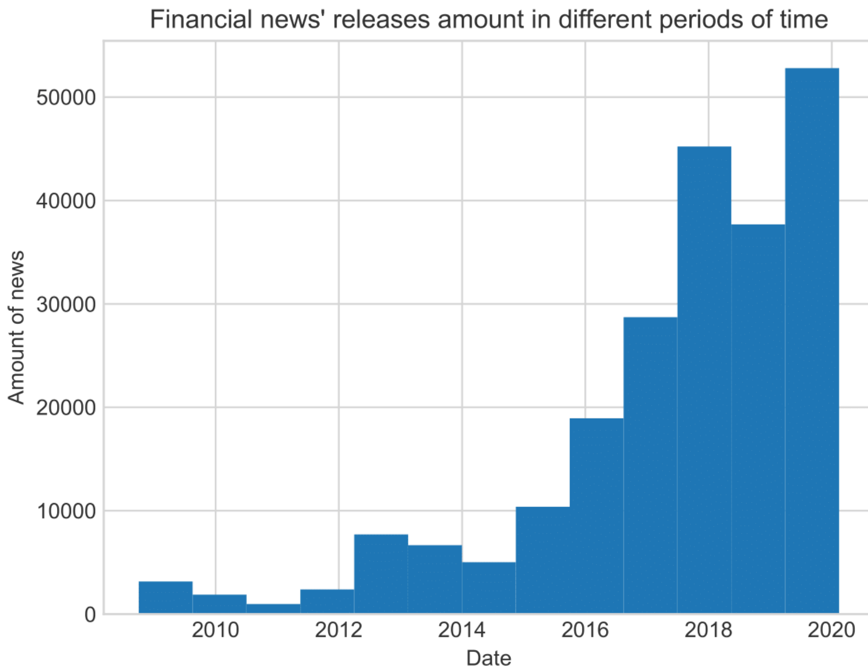

Figure 8: The distribution of financial news' release dates

- **BigTech companies Twitter dataset [2]**

The content of this Twitter dataset represents what people are saying about BigTech companies. It consists of tweets, released during the period from 12.07.2020 till 19.09.2020. The data structure is the following: 866909 samples and 15 columns. The data comprises tweets about 10 companies. The information about several most interesting dataset fields is in Table 2. The amounts of tweets, created during different time periods, are visualized in Figure 9.

Data properties		
Attribute name	Description	Type
created at	the date when the tweet was created	object
file name	the company name	object
followers	the number of tweet's author followers	int64
friends	the number of tweet's author friends	int64
group name	the company name	object
location	the country and the place, where the tweet was published	object
retweet count	the number of tweet's citations	float64
screen name	the author's nickname	object
search query	the query utilized for getting the tweet about particular company	object

text	the full text of the tweet	object
twitter id	the tweet's identificator	float64
username	the author's full name	object
polarity	the tweet's sentiment	float64
partition 0	the company's profile (1 unique value - Technology)	object
partition 1	the company's name	object

Table 2: Big Tech companies twitter dataset properties description

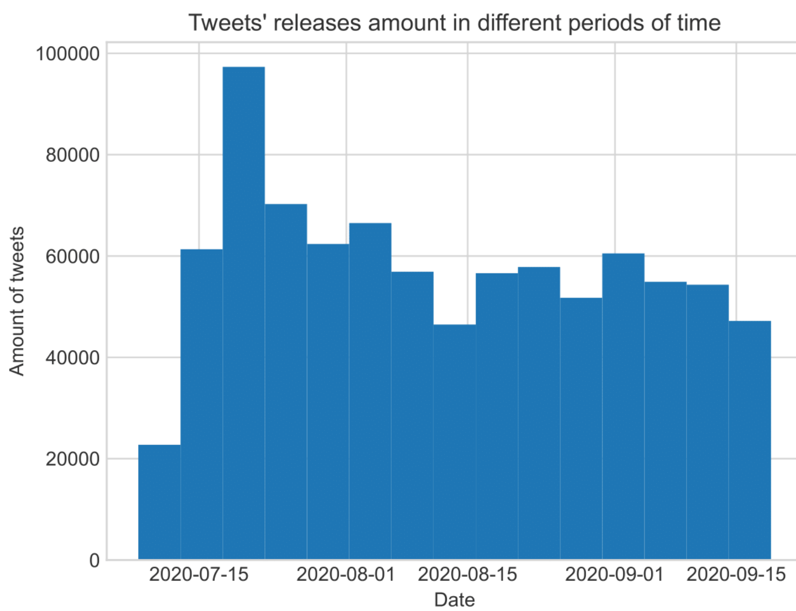

Figure 9: The distribution of tweets' release dates

- **General dataset**

The general dataset, which was gathered from 2 described datasets, contains 1088422 samples. The description of the constructed dataset is in Table 3.

Data properties		
Attribute name	Description	Type

Ticker	an abbreviation, which helps to identify the stock of a particular public company	object
Data source	financial news or Twitter	object
Headline	News' headline	object
Content	News' full text	object
Release date	The date, when particular news or tweet was published	object

Table 3: General dataset properties description

4.2 Metrics

The metrics, which were utilized, are the following:

- **Pearson's correlation coefficient** - measures the strength and the existence of a linear relationship between two variables, if its value is significant, we can conclude that the relationship exists, and the change of one variable is accompanied with the change of another variable. The formula of Pearson correlation is:

$$r_{xy} = \frac{\sum_{i=1}^{n}(d_{xi}d_{yi})}{\sqrt{\sum_{i=1}^{n} d_{xi}^2 \sum_{i=1}^{n} d_{yi}^2}},$$

  where d_{xi} and d_{yi} is the deviation of the i-th observation of x and y variables accordingly from the average value of x and y variables.
- **Spearman's correlation coefficient** - measures the strength of monotonic relationship between two variables, the monotonic means that if the value of one variable increases, the value of the other variable grows too, or vice versa, as the value of one variable increases, the other variable value decreases. The formula of Spearman's correlation:

$$r = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)},$$

  where n - number of data points of the two variables, d_i is the difference between ranks of i-th data points of two variables. Data ranking is achieved by assigning the rank 1 to the biggest number in the column, 2 should be given to the second biggest number and so forth. The smallest value should get the lowest ranking. This should be done for both sets of values.
- **Kendall's tau correlation coefficient** - the ranked correlation coefficient. The formula of the coefficient is:

$$\frac{C-D}{C+D},$$

  where C is the amount of concordant pairs, D - amount of discordant pairs. The pairs (x_i, y_i) and (x_j, y_j) are called concordant if $x_i < x_j$ and $y_i < y_j$ or $x_i > x_j$ and $y_i > y_j$, otherwise the pairs are called discordant.

- **Dynamic time warping** - a distance metric between two input time series, allows two variables to have a good match even if on the x-axes they are not sync. It can indicate the similarity between time series, which are different in speed.
- **MAPE (Mean Absolute Percentage Error)** - the MAPE formula is:

$$\frac{1}{n}\sum\left|\frac{actual - forecast}{actual}\right| * 100,$$

where *actual* - the actual value , *forecast* - the predicted value. The value of MAPE shows the average difference between the forecasted value and the actual value. It is not hard to guess that the lower values of MAPE, the better the prediction quality is. This particular metric was chosen for evaluating the model quality because of its intuitive interpretation.

Clustering metrics:

1) **Distortion** - the average of the squared distances between cluster centers and data points of the respective clusters.
2) **Silhouette score -** the formula is: s(i) = $\frac{b(i) - a(i)}{max\{a(i), b(i)\}}$, where $b(i)$ - distance between i and its nearest cluster centroid, $a(i)$ - average of the distance between i and all of the other points in its own cluster. The value of s(i) is always between -1 and 1. The closer the value to 1, the better the clustering performance.

4.3 Results

4.3.1 Topic classification

The results of the topic classification, presented below, illustrate the amount of news related by the model to each group. We can conclude that the most numerous groups are "Company | Product News", "Markets", "Stock Commentary", "Stock Movement" and "Financials".

- Analyst Update - 7023
- Fed | Central Banks - 3149
- Company | Product News 44401
- Treasuries | Corporate Debt - 2514
- Dividend - 1208
- Earnings - 16595
- Energy | Oil - 6434
- Financials - 17857
- Currencies - 3547
- General News | Opinion 12468
- Gold | Metals | Materials - 2342
- IPO - 833
- Legal | Regulation - 7356
- M&A | Investments - 8724
- Macro - 11210
- Markets - 24264
- Politics - 4619
- Personnel Change - 2690
- Stock Commentary - 24364
- Stock Movement - 19915

To ensure that news flow was splitted reasonably, we extracted defined groups from the general dataset, read the news and tried to analyze if they are all connected with the particular topic and have something in common. we noticed that the groups were defined sensibly. For example, in Figure 10 there is the list of news, related to Analyst Update topic, and each news is connected with some changes (keywords "Upgrades", "gained", "crisis grows", "raised") or the results of analytical reports (keyword "Analyst Blog"). Also, In Figure 11, several news articles about currencies are presented, and, according to the key words ("US dollar", "euro", "Egyptian Pound"), it is obvious that the news are connected with currencies. In "Legal | Regulation" dataset (Figure 12) the key words, which can prove the connection with this topic, are "oppose", "sentenced", "prison", "suspends", "fine". In Figure 13 some stock commentary news are shown, and the majority of news headlines in this group have the key word "Stock". Finally, we can make a conclusion that our approach classified news wisely, and these results can be used in further experiments.

Headline
Citigroup Remains Bullish on Facebook
Ally Financial s Ratings Upgraded By Moody s Outlook Stable
Northland Capital upgrades Qualcomm on 3D sensor promise
BofAML ups Nvidia price target to near Street high
Apple to gain unconditional EU approval for Shazam buy sources
Bed Bath   Beyond s Prospects Bleak Sell The Stock Now
Trump calls Boeing a very disappointing company as 737 Max crisis grows
JPMorgan Says Should ve Downgraded Brazil Stocks But Too Late
Terex 1 after upgrade at Jefferies
The Zacks Analyst Blog Highlights Amazon Procter   Gamble Chevron United Parcel Service And Raytheon
Top Analyst Reports For Allergan Ecolab   Exelon
The Zacks Analyst Blog Highlights Wyndham InterContinental Hotels Marriott Snap And Expedia
The Zacks Analyst Blog Highlights Target Tempur Sealy Casey s Chuy s And Insight Enterprises
Zacks Investment Ideas Feature Highlights Chemours Party City Holdco And Tower International
The Zacks Analyst Blog Highlights United Continental Spirit Airlines Delta Air Lines Southwest Airlines And Alaska Air
BMO Overweight on attractive new Communications sector
Barclays upgrades Nvidia to 23 upside
Coach Inc Raised To Strong Buy At Vetr Inc
CMS Energy CMS Upgraded To Buy Here s What You Should Know
4 Contrarian Market Views
Estee Lauder EL Raises View On Solid Q3 Earnings   Sales

Figure 10: Analyst Update news

Headline
USD Extends Losses Against Major Counterparts
Euro Moves Lower Amid ECB Speculations
Euro Soared On LTRO News As JPY CAD AUD Dived
GBP USD Maintaining Momentum And Pushing Towards Critical 1 300
EUR USD Touches Four Week High At 1 31
Daily FX Analysis Euro Surprisingly Well Bid Despite Cyprus
Pound Snapped A Three Day Decline Against The USD
Weak U S dollar worries some but lessens recession
Risk Currencies Strike Positive Tone On Fiscal Cliff Progress
FOREX Dollar slips vs euro in thin market U S data eyed
Euro Rebounds On Broadly Successful Greek Buy Back
Daily Report US Dollar Retreats Against The Yen
Egyptian Pound Hits Record Lows Banks Running Out Of Dollars
USD JPY Holds Gains Despite No BOJ Easing This Time
EUR USD XAU USD Technicals 04 03 2019
The Coronavirus Impact On USD CNH
Dollar Restraint While S P 500 Collapses What Does It Mean
Nikkei battered by yen worry about US financials
FOREX U S dollar rises as investors trim bearish bets
U S Dollar Plummets That Is All you Need To Know
RPT FOREX U S dollar falls to 2009 low vs euro

Figure 11: Currencies news

Headline
Icahn Deason urge Xerox shareholders to oppose Fujifilm deal
Ex Autonomy CFO sentenced in U S to 5 years prison over Hewlett Packard fraud
UPDATE 1 EU probes new China postal law over trade concerns
FINRA fines JPMorgan 1 25 million for failures in employee background checks
Tech giants prep for legal fight against FCC s rule repeal
Fair use matters in dancing toddler copyright case U S court
UPS to pay 8 4 million to resolve U S overcharging probe
Mexico court suspends sale of Roku TV streaming gadgets
U S Congress seeks answers on patient privacy in Google Ascension cloud deal
EU slaps Qualcomm with 1 2B antitrust fine
Google CEO faces first Congressional hearing
America Movil Unit Faces Fine For Network Discrepancy Case

Figure 12: Legal | Regulation news

Headline
Here s Why You Should Hold Alliance Data ADS Stock Now
What Makes Harris HRS A Strong Momentum Stock Buy Now
Are You Looking For A High Growth Dividend Stock Principal Financial PFG Could Be A Great Choice
Here s Where I ll Buy JetBlue
Why Regions Financial RF Is A Great Pick Right Now
Texas Instruments A Growth Stock Paying A 3 2 Dividend Yield
Three Foreign Stocks To Watch STNE QTT NIO
Looking For A Strong Finish To 2014 Top Stock Picks
Top Ranked Growth Stocks To Buy For June 15th

Figure 13: Stock commentary news

4.3.2 Sentiment labeling

To evaluate the quality of model performance on different topic groups, we counted the percentage of the sum of positive and negative labels in the overall amount of labels, as there is not a lot of interest in a big amount of neutral news. The results you can see in Table 4. As a result, I classified all groups, where the percentage is more than 50%. And, according to this consideration, the good classification was performed on such groups, as "Currencies", "Energy | Oil", "Financials", "Gold | Metals | Materials", "Macro", "Markets", "Stock Movement", "Twitter". In further experiments we will check whether these groups will show better results in the improvement of DL model prediction quality.

Sentiment labels distribution				
Topic group name	Positive labels' amount	Negative labels' amount	Neutral labels' amount	Positive + Negative percentage (%)
Analyst Update	1872	774	4377	38%
Fed | Central Banks	384	688	2077	34%
Currencies	742	1133	1672	53%
Dividend	409	55	744	38%

Earnings	4909	2108	9578	42%
Energy | Oil	1605	2158	2671	58%
Financials	11363	5662	832	95%
General News | Opinion	1061	2506	8901	29%
Gold | Metals | Materials	691	615	1036	56%
M&A | Investments	1076	318	7330	16%
IPO	103	87	643	23%
Legal | Regulation	225	1740	5391	27%
Macro	2582	3565	5063	55%
Markets	7984	7472	9540	62%
Personnel Change	109	153	2428	10%
Politics	303	641	3675	20%
Company | Product News	8893	4275	31233	30%
Stock Commentary	9650	1634	13080	46%
Stock Movement	8862	5493	5560	72%
Treasuries | Corporate Debt	335	521	1658	34%
Twitter	389743	87879	392817	55%

Table 4: The table, which shows amounts of sentiment labels of each type inside each topic group

4.3.3 Clustering

Before clustering we applied the PCA method to the pre-processed data. The PCA has a parameter n_components, which indicates how many dimensions will have the space to which we project our data. In Figure 14, the dependency between described parameters and clustering performance is illustrated. As it can be seen from the graph, the best performance of clustering was achieved with only 2 components. In the result, we took n_components, equalled 2, during the final clustering.

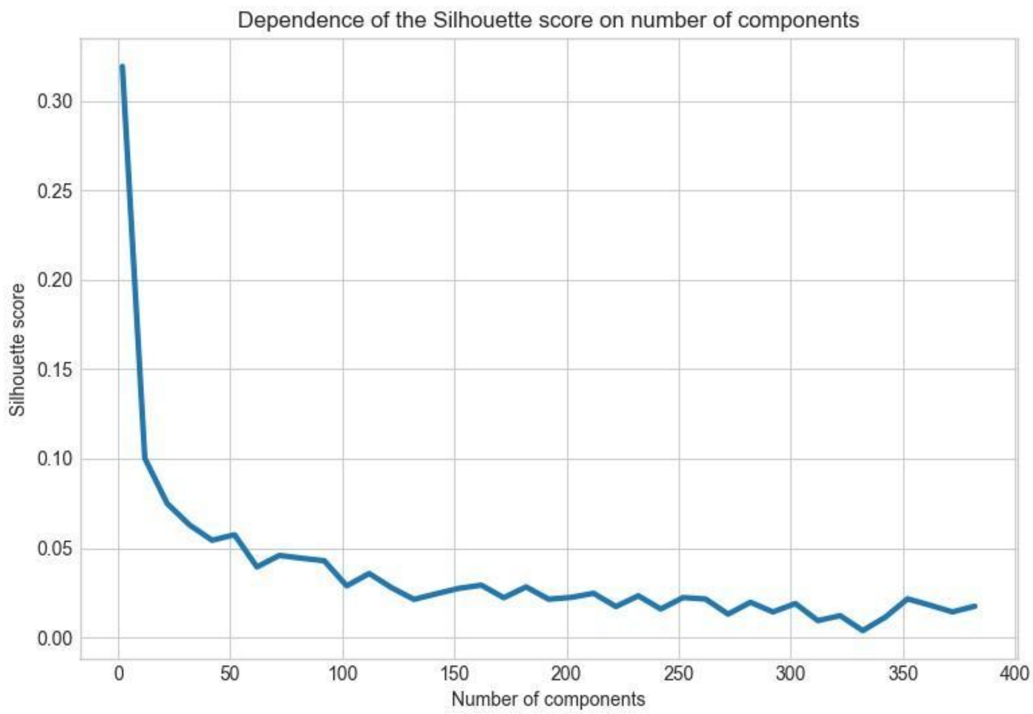

Figure 14: The dependency of clustering performance (Silhouette score) on number of components to which we reduce the dimension

When we started clustering, the question about the amount of clusters was raised. For choosing the amount of clusters we evaluated the metric called "distortion". "Distortion" means squared distance between each object and its cluster center. In Figure 15 The dependency between distortion and clustering performance are presented. The optimal number of classes can be defined by the elbow method, which means that we should find the point, after which the distortion starts to decrease in a linear way. The red point in Figure 15 is elbow point. It means that 10 clusters is the optimal amount.

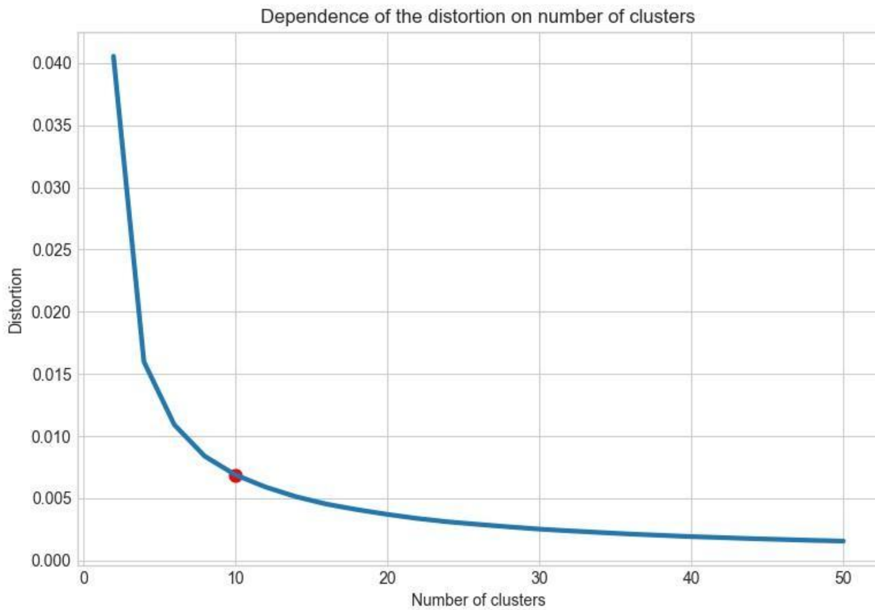

Figure 15: The chart, which shows the distortion dependency on number of clusters, the red point marks the optimal number of clusters according to the elbow method

As the metric for clustering performance evaluation we took Silhouette score, because it does not demand preliminary labeled data. The Silhouette score has a parameter called sample_size. The graph in figure 14 helped to define which size to utilize. Finally, I chose sample_size equal to 50000.

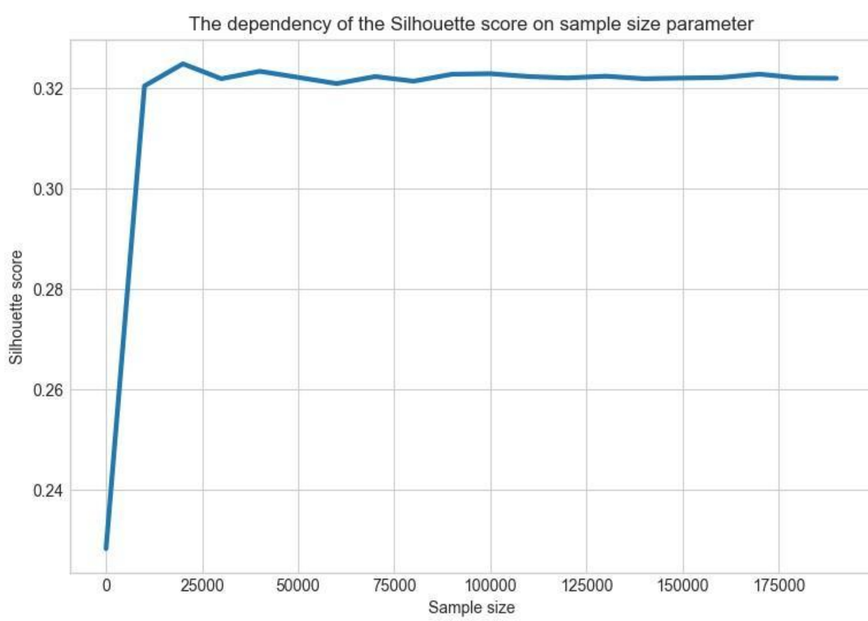

Figure 16: The graph, which shows the dependency of the Silhouette score on its sample size parameter

The results of clustering are illustrated in the Figure17 The Silhouette score = 0.34057847, which means that some clusters are defined, but they intersect each other, so the quality of clustering is not enough for our further experiments.

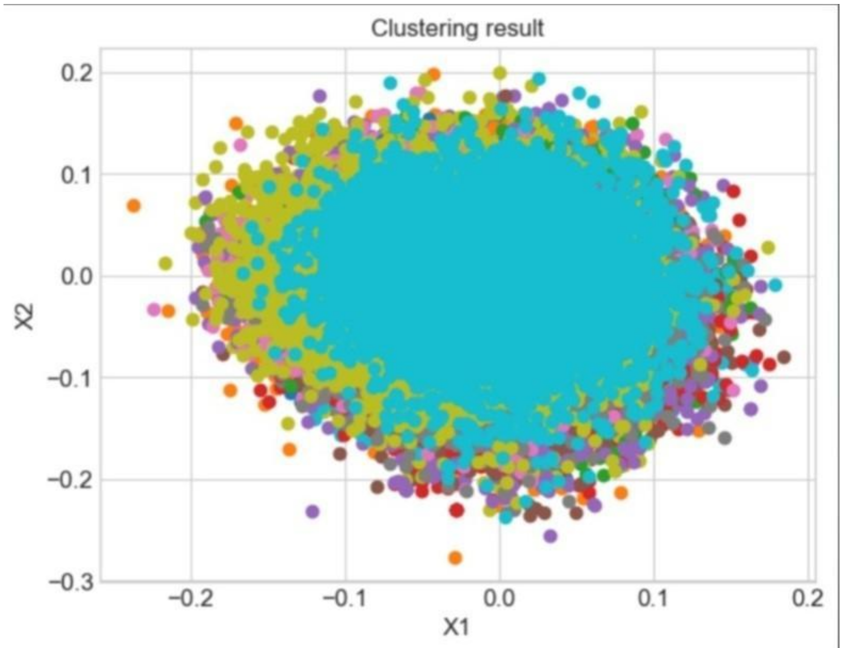

Figure 17: The scatter plot, which visualizes the results of clustering

4.3.4 Time series similarity evaluation

In tables 5, 6, 7, 8, 9 there are the values described in section 4.2 metrics, which evaluate the similarity of 5 big companies' stock prices time series and different news topic groups' positive news ratio series. For each company the line charts with 2 strongly correlated variables on the one graph are presented.

Metrics for Apple				
Topic group name	Pearson's correlation	Spearman's correlation	Kendall's correlation	Dynamic time warping
Analyst Update	0.9	0.96	0.87	1.12
Fed | Central Banks	0.9	0.92	0.75	1.5
Currencies	-0.002	0.27	0.2	3.37
Dividend	0.93	0.92	0.76	2.77
Earnings	0.93	0.94	0.83	0.7
Energy | Oil	0.8	0.78	0.65	1.41

Financials	0.82	0.92	0.78	1.2
General News | Opinion	0.39	0.54	0.44	1.34
Gold | Metals | Materials	0.88	0.95	0.88	1.92
M&A | Investments	0.96	0.95	0.82	1.87
IPO	-0.09	0.33	0.22	4.84
Legal | Regulation	0.06	-0.19	-0.25	2.12
Macro	0.87	0.92	0.76	0.77
Markets	-0.1	-0.52	-0.34	1.65
Personnel Change	0.65	0.65	0.43	2.53
Politics	0.86	0.91	0.73	2.02
Company | Product News	0.56	0.72	0.6	0.7
Stock Commentary	0.75	0.8	0.67	1.04
Stock Movement	0.62	0.78	0.62	1.13
Treasuries | Corporate Debt	0.96	0.96	0.8	2.6

Table 5: The table, which shows the values of evaluated news' and prices' correlation metrics for Apple

According to the table 5, Apple's stock prices mostly correlate with the following topic groups: "Analyst Update", "Fed | Central Banks", "Dividend", "Earnings", "Financials", "Gold | Metals | Materials", "M&A | Investments", "Macro", "Politics", "Treasuries | Corporate Debt". The least Dynamic Time Warping between prices and news is observed for the "Earnings" topic group.

The graph below (Figure 18) presents Apple stock prices' time series and positive news' ratio time series. All news relate to the "M&A | Investments" topic group. It is important to note, that the stock prices time series was scaled to the interval from 0 to 1 inclusive, so that the interval of values was the same for both time series. The same manipulation was performed for all graphs in the current section.

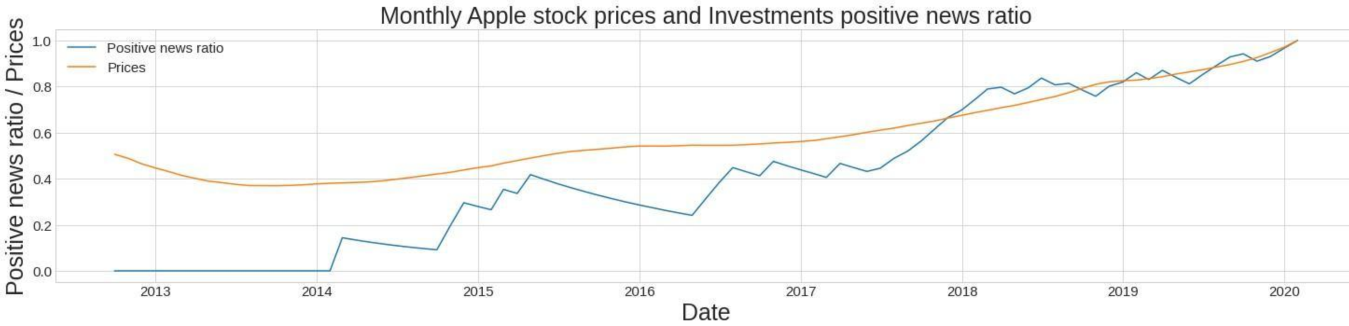

Figure 18: The figure illustrates Apple monthly stock prices and monthly Investments positive news ratio, related to Apple, dynamics

Metrics for Amazon				
Topic group name	Pearson's correlation	Spearman's correlation	Kendall's correlation	Dynamic time warping
Analyst Update	0.6	0.68	0.6	2.32
Fed | Central Banks	-0.04	0.02	-0.06	3.5
Currencies	0.89	0.75	0.58	1.75
Dividend	0.23	0.03	-0.04	2.05
Earnings	0.95	0.92	0.78	0.59
Energy | Oil	0.94	0.8	0.68	1.53
Financials	0.74	0.84	0.78	1.39
General News | Opinion	0.52	0.5	0.41	1.53
Gold | Metals | Materials	0.5	0.45	0.27	1.62
M&A | Investments	0.98	0.97	0.86	1.1
IPO	0.29	0.32	0.12	2.55
Legal | Regulation	0.77	0.72	0.47	2.1
Macro	0.98	0.97	0.88	1.02
Markets	0.61	0.55	0.49	1.78
Personnel Change	0.91	0.84	0.66	1.62
Politics	0.88	0.8	0.6	1.58
Company | Product News	0.82	0.99	0.94	0.87
Stock Commentary	0.96	0.92	0.82	0.58
Stock Movement	0.71	0.7	0.66	1.64

Treasuries | Corporate Debt	0.89	0.76	0.59	1.72

Table 6: The table, which shows the values of evaluated news' and prices' correlation metrics for Amazon

Table 6 indicates that Amazon's stock prices mostly correlate with such topic groups as "Earnings", "Energy | Oil", "M&A | Investments", "Macro", "Personnel Change", "Company | Product news", "Stock Commentary". The least Dynamic Time Warping between prices and news is observed for the "Stock Commentary" and "Earnings" topic groups.

The graph below (Figure 19) presents Amazon stock prices' time series and positive news' ratio time series. All news relates to the "Macro" topic group.

Figure 19: The figure illustrates Amazon monthly stock prices and monthly Macro positive news ratio, related to
Amazon, dynamics

Table 7 shows us that Google stock prices' strongly correlate with "General news | Opinion", M&A | Investments", "Macro", "Markets", "Company | Product news", "Stock Movement", "Treasuries | Corporate Debt". The interesting observation consists in strong negative correlation with news from the "Personnel change" category. The minimal Dynamic Time Warping corresponds with "Macro" news topic group.

Metrics for Google				
Topic group name	Pearson's correlation	Spearman's correlation	Kendall's correlation	Dynamic time warping
Analyst Update	0.12	0.41	0.33	4.85
Fed | Central Banks	0.27	0.47	0.13	2.89
Currencies	0.05	0.02	-0.17	5.53
Dividend	- *	-	-	5.7

Earnings	0.28	0.34	0.2	4.08
Energy | Oil	-0.33	-0.34	-0.3	5.63
Financials	0.73	0.78	0.66	2.53
General News | Opinion	0.88	0.94	0.79	1.25
Gold | Metals | Materials	-0.28	0.07	0.22	4.37
M&A | Investments	0.96	0.9	0.77	1.99
IPO	0.88	0.86	0.64	2.2
Legal | Regulation	0.87	0.87	0.66	5.52
Macro	0.96	0.97	0.82	1.33
Markets	0.97	0.93	0.8	0.73
Personnel Change	-0.47	-0.97	-0.9	5.62
Politics	0.79	0.74	0.54	1.67
Company | Product News	0.95	0.97	0.89	1.25
Stock Commentary	0.83	0.86	0.7	1.007
Stock Movement	0.9	0.94	0.8	1.16
Treasuries | Corporate Debt	0.86	0.83	0.6	5.88

Table 7: The table, which shows the values of evaluated news' and prices' correlation metrics for Google * - the situation, when all the news or tweets in the time series were negative or neutral or there were not any news or tweets.

The graph below (Figure 20) shows Google stock prices' time series and positive news' ratio time series. All news relate to the "Company | Product news" topic group.

Figure 20: The figure illustrates Google monthly stock prices and monthly positive Product news ratio, related to Google, dynamics

The information from the table 8 allows to mark the following topic groups, strongly correlated with Netflix stock prices: "Financials", "General news | Opinion", "Macro", "Company | Product news", "Stock Commentary". The least Dynamic Time Warping in this table relates to the "Financials" news category.

Metrics for Netflix				
Topic group name	Pearson's correlation	Spearman's correlation	Kendall's correlation	Dynamic time warping
Analyst Update	0.14	0.54	0.45	3.39
Fed | Central Banks	-	-	-	4.79
Currencies	0.56	0.43	0.18	1.58
Dividend	-	-	-	4.65
Earnings	0.19	0.39	0.35	2.53
Energy | Oil	0.03	0.36	0.2	2.9
Financials	0.92	0.98	0.89	0.35
General News | Opinion	0.97	0.97	0.83	0.53
Gold | Metals | Materials	-0.14	-0.26	-0.19	2.08
M&A | Investments	0.45	0.44	0.22	1.97
IPO	0.28	0.28	0.23	2.58
Legal | Regulation	0.64	0.73	0.39	2.84
Macro	0.96	0.91	0.72	0.62
Markets	0.43	0.53	0.47	2.15
Personnel Change	-	-	-	4.66
Politics	-0.11	-0.03	-0.19	3.13
Company | Product News	0.94	0.97	0.87	0.46
Stock Commentary	0.95	0.9	0.76	0.53
Stock Movement	0.61	0.59	0.57	1.9

Treasuries | Corporate Debt	0.88	0.79	0.54	2.62

Table 8: The table, which shows the values of evaluated news' and prices' correlation metrics for Netflix

The next graph (Figure 21)illustrates Netflix stock prices' time series and positive news' ratio time series. All news are from the "General news | Opinion" topic group.

Figure 21: The figure illustrates Netflix monthly stock prices and monthly positive General news ratio, related to Netflix, dynamics

According to the next table (Table 9), the strongest correlation Tesla stock prices show with "Politics", "Stock Commentary" and "Stock Movement" news categories. The least Dynamic Time Warping is observed with news from the "Stock Movement" news topic group.

Metrics for Tesla				
Topic group name	Pearson's correlation	Spearman's correlation	Kendall's correlation	Dynamic time warping
Analyst Update	0.87	0.89	0.71	1.51
Fed | Central Banks	-0.39	-0.43	-0.3	6.49
Currencies	0.62	0.58	0.38	2.93
Dividend	0.45	0.61	0.54	2.77
Earnings	0.62	0.72	0.5	1.09
Energy | Oil	0.63	0.5	0.33	2.92
Financials	0.75	0.69	0.46	2.83

General News | Opinion	0.62	0.55	0.4	1.33
Gold | Metals | Materials	0.48	0.35	0.18	3.55
M&A | Investments	0.69	0.63	0.4	3.51
IPO	-	-	-	6.82
Legal | Regulation	0.89	0.87	0.66	4.92
Macro	0.85	0.78	0.6	2.42
Markets	0.23	0.71	0.63	3.44
Personnel Change	-0.38	-0.45	-0.36	5.7
Politics	0.78	0.9	0.73	1.07
Company | Product News	-0.24	0.49	0.46	3.46
Stock Commentary	0.93	0.93	0.78	0.65
Stock Movement	0.94	0.97	0.86	0.48
Treasuries | Corporate Debt	0.72	0.66	0.43	3.14

Table 9: The table, which shows the values of evaluated news' and prices' correlation metrics for Tesla

The following line chart (Figure 22) presents Tesla stock prices' time series and positive news' ratio time series. All news are from the "Stock Movement" topic group.

Figure 22: The figure illustrates Tesla monthly stock prices and monthly Stock Commentary positive news ratio, related to Tesla, dynamics

For tweets I performed a more comprehensive experiment. Besides correlations and Dynamic Time Warping between stock prices and positive tweets' ratio I also evaluated three kinds of correlation coefficient of daily stock prices' volatility with daily positive tweets' ratio and daily stock prices' volatility with daily tweets amount.

Twitter metrics for each company (part 1)				
Company name	Pearson's correlation (sentiment prices)	Spearman's correlation (sentiment prices)	Kendall's correlation (sentiment prices)	Dynamic time warping
Apple	-0.68	-0.86	-0.72	0.67
Amazon	0.45	0.74	0.55	0.17
Google	0.63	0.56	0.32	0.42
Netflix	0.07	0.02	0.02	0.46
Tesla	-0.57	-0.49	-0.36	1.26

Table 10: The table, which shows the values of evaluated for each company metrics, which indicate the connection between tweets about company and its stock prices

Twitter metrics for each company (part 2)						
Company name	Pearson's correlation (amount volatility)	Spearman's correlation (amount volatility)	Kendall's correlation (amount volatility)	Pearson's correlation (sentiment volatility)	Spearman's correlation (sentiment volatility)	Kendall's correlation (sentiment volatility) -
Apple	0.54	0.88	0.71	-0.53	-0.58	-0.4
Amazon	-0.27	0.06	0.05	-0.92	-0.79	-0.67
Google	0.49	-0.47	-0.29	-0.23	-0.16	-0.23
Netflix	-0.85	-0.58	-0.45	0.48	0.42	0.23
Tesla	-0.37	-0.43	-0.33	-0.1	-0.04	-0.0009

Table 11: The table, which shows the values of evaluated for each company metrics, which indicate the connection between tweets about company and its stock prices

Interesting observations from tables 10-11:

- For Apple and Amazon we can observe the pronounced negative correlation between positive tweets' ratio and stock prices. It means that smaller values of positive tweets' ratio corresponds with bigger values of stock price.
- For Amazon and Google there is an explicit positive correlation between positive tweets' ratio and stock prices.
- Netflix prices do not correlate with positive tweets' ratio.

- Apple metrics show that daily tweets amount positively correlates with daily stock prices volatility, but daily positive tweets amount negatively correlates with stock prices' volatility. In other words, the big amount of background information corresponds with the higher value of stock prices variability, however, when there are a lot of positive tweets in one day, the values of stock prices are more stable.
- Amazon daily tweets amount does not correlate with stock prices volatility, however, as for Apple, there is a strong negative correlation of volatility and positive tweets' ratio.

- For Netflix, we can mark the strong negative correlation between the daily amount of tweets and stock prices volatility.

Next 15 (Figure 23 - 37) line charts illustrate the variables, for which previously described Twitter metrics were calculated.

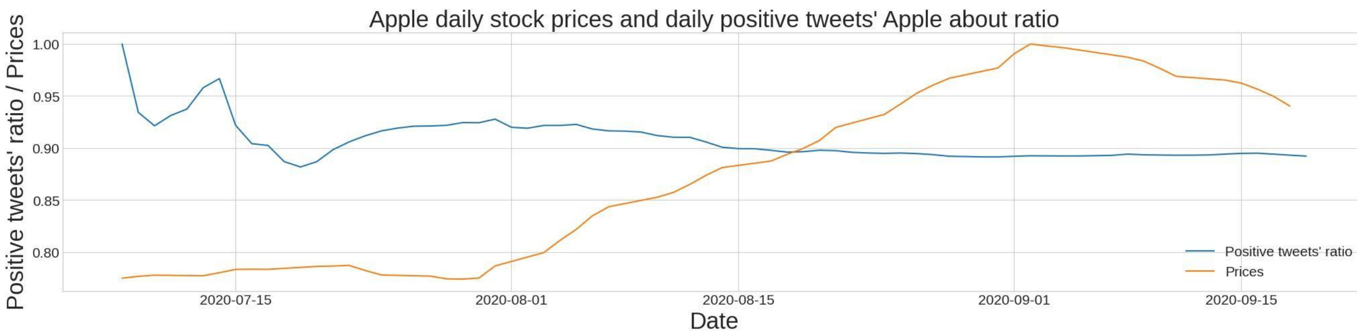

Figure 23: The figure illustrates Apple daily stock prices and daily positive tweets ratio, related to Apple, dynamics

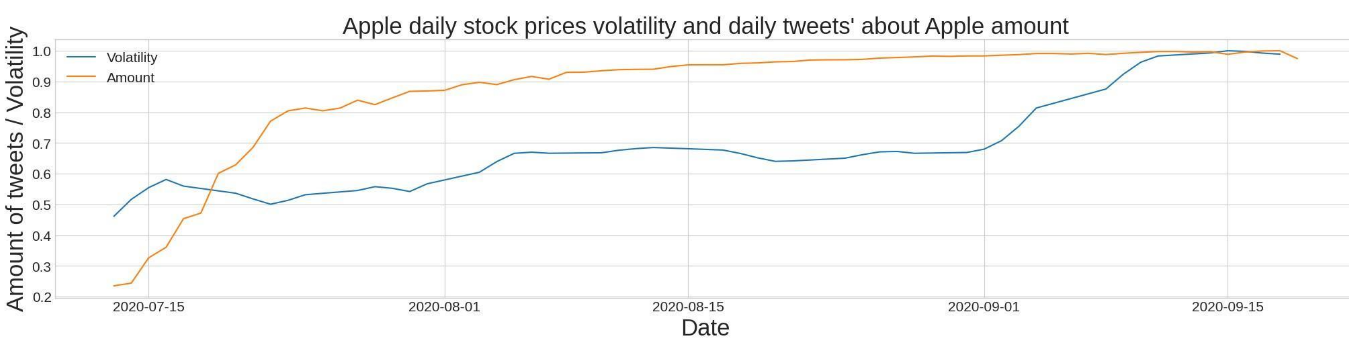

Figure 24: The figure illustrates Apple daily stock prices volatility and daily tweets, related to Apple, amount dynamics

Figure 25: The figure illustrates Apple daily stock prices volatility and daily positive tweets' about Apple ratio dynamics

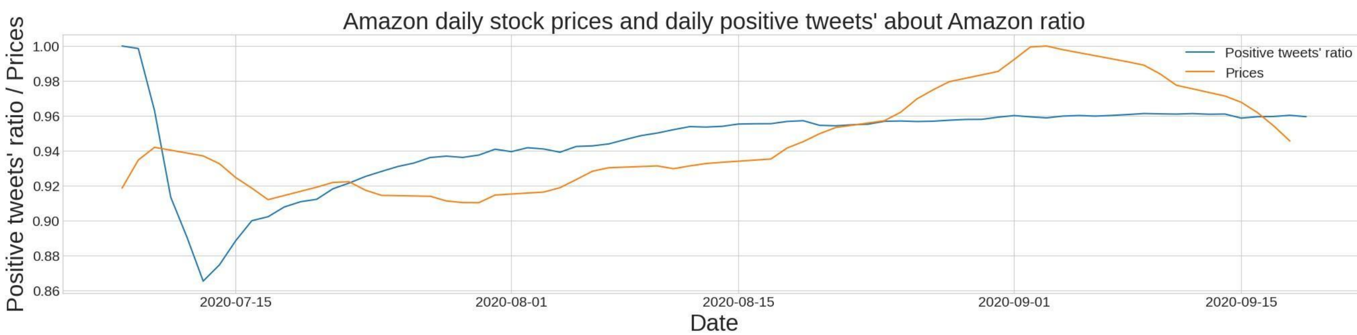

Figure 26: The figure illustrates Amazon daily stock prices and daily positive tweets' about Amazon ratio dynamics

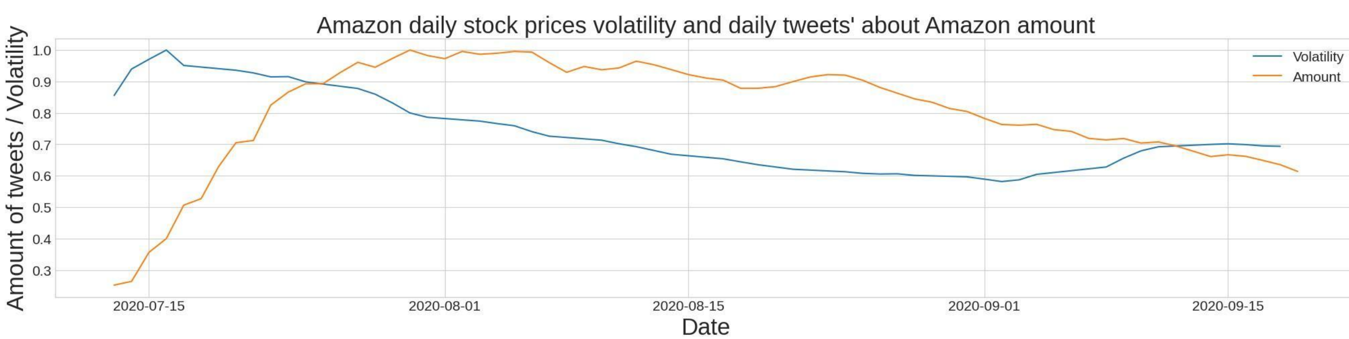

Figure 27: The figure illustrates Amazon daily stock prices volatility and daily tweets, related to Amazon, amount dynamics

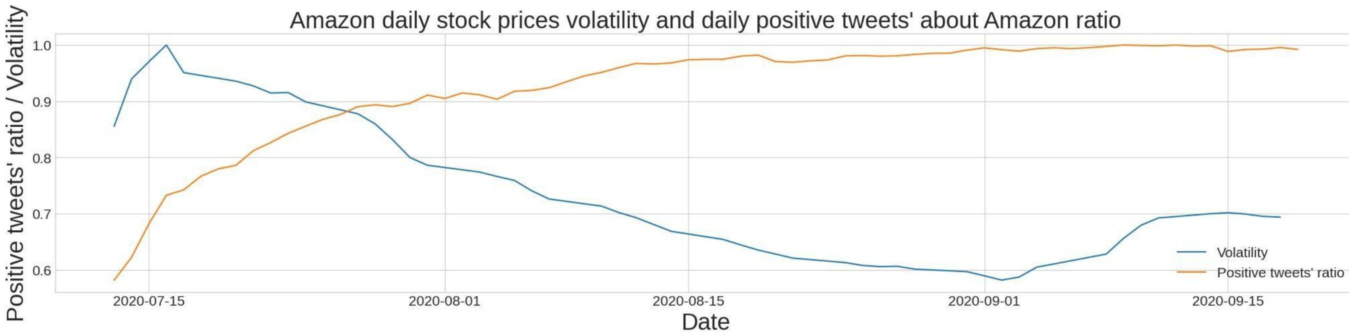

Figure 28: The figure illustrates Amazon daily stock prices volatility and daily positive tweets' about Amazon ratio dynamics

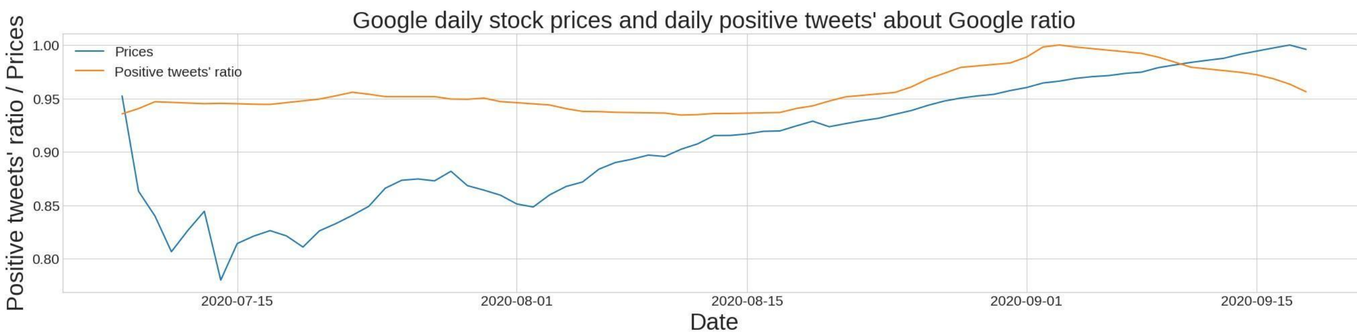

Figure 29: The figure illustrates Google daily stock prices and daily positive tweets' about Google ratio dynamics

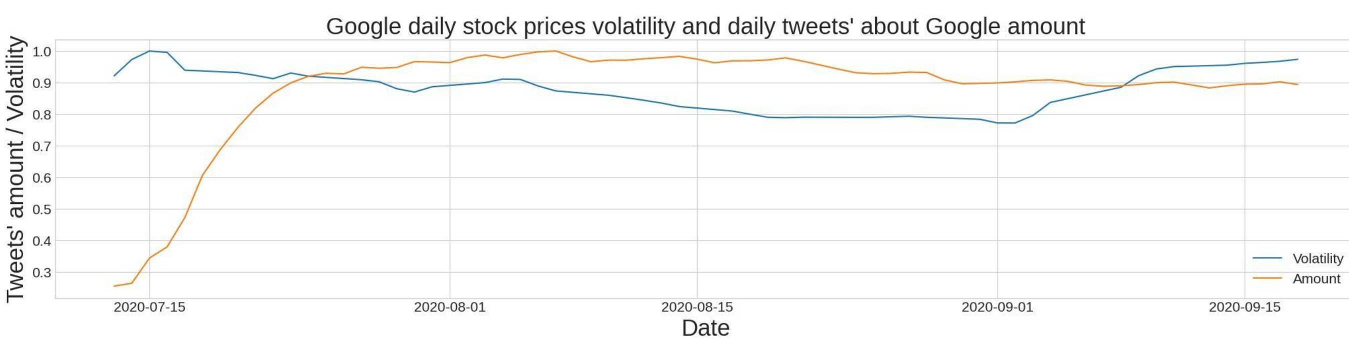

Figure 30: The figure illustrates Google daily stock prices volatility and daily tweets, related to Google, amount dynamics

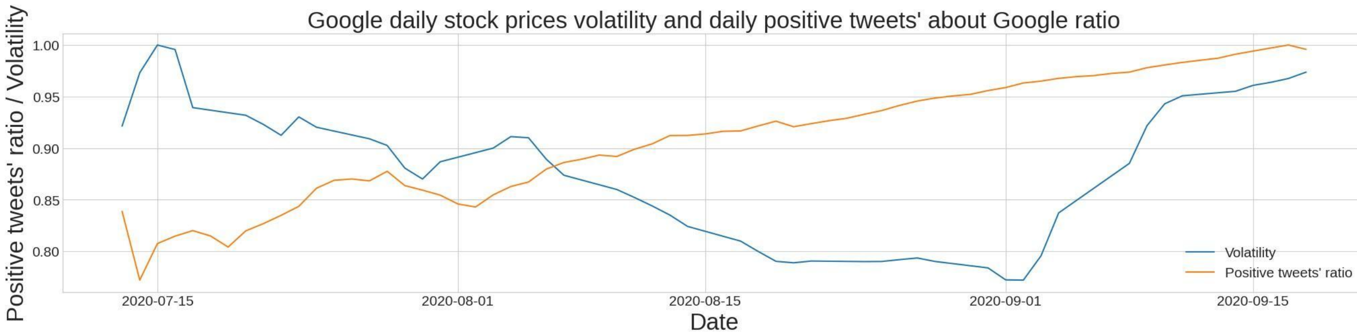

Figure 31: The figure illustrates Google daily stock prices volatility and daily positive tweets' about Google ratio dynamics

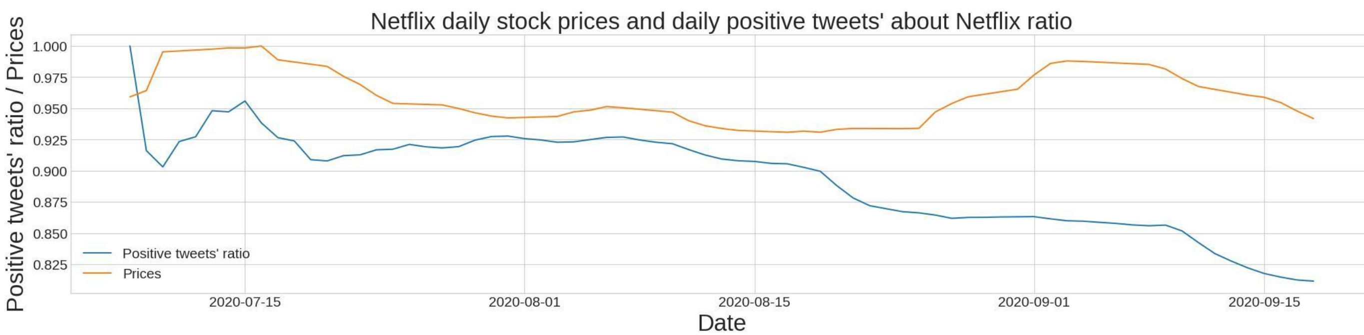

Figure 32: The figure illustrates Netflix daily stock prices and daily positive tweets' about Netflix ratio dynamics

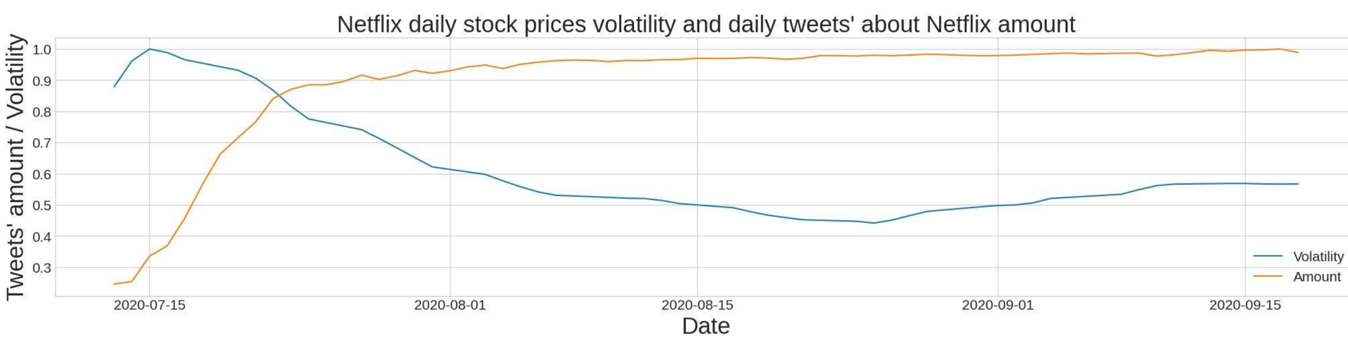

Figure 33: The figure illustrates Netflix daily stock prices volatility and daily tweets, related to Netflix, amount dynamics

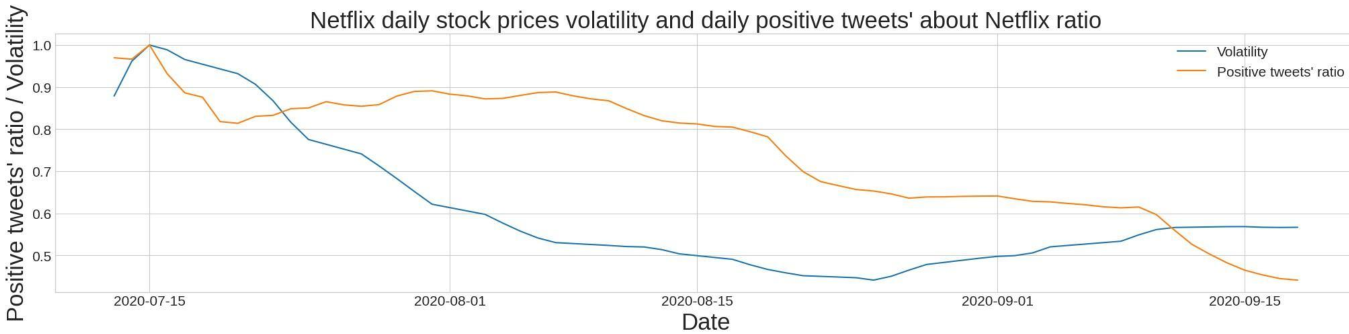

Figure 34: The figure illustrates Netflix daily stock prices volatility and daily positive tweets' about Netflix ratio dynamics

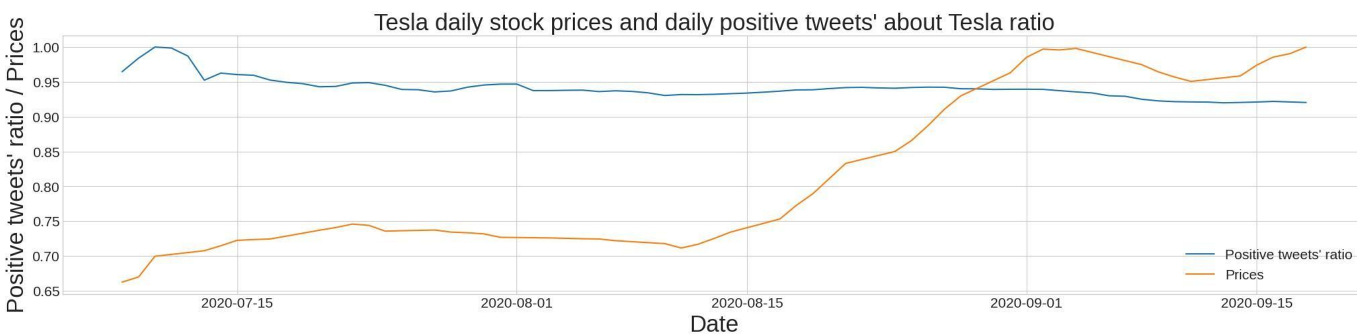

Figure 35: The figure illustrates Tesla daily stock prices and daily positive tweets' about Tesla ratio dynamics

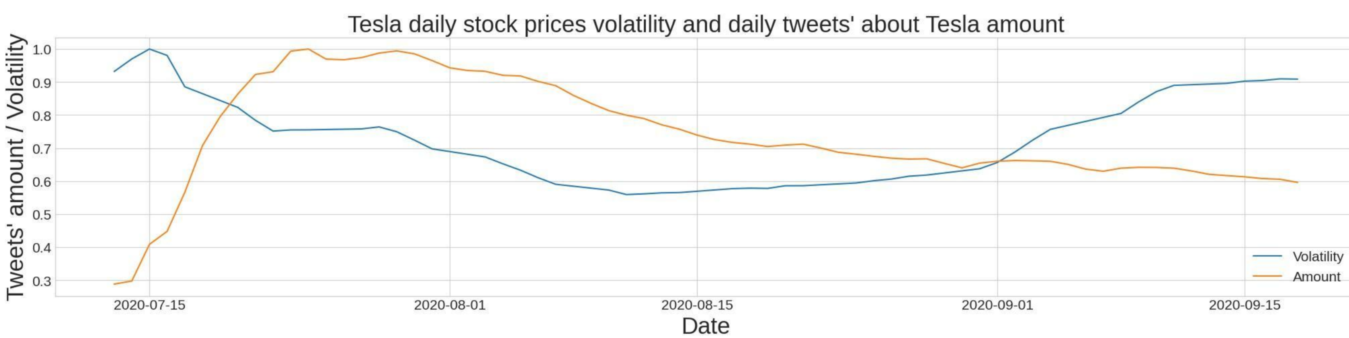

Figure 36: The figure illustrates Tesla daily stock prices volatility and daily tweets, related to Tesla, amount dynamics

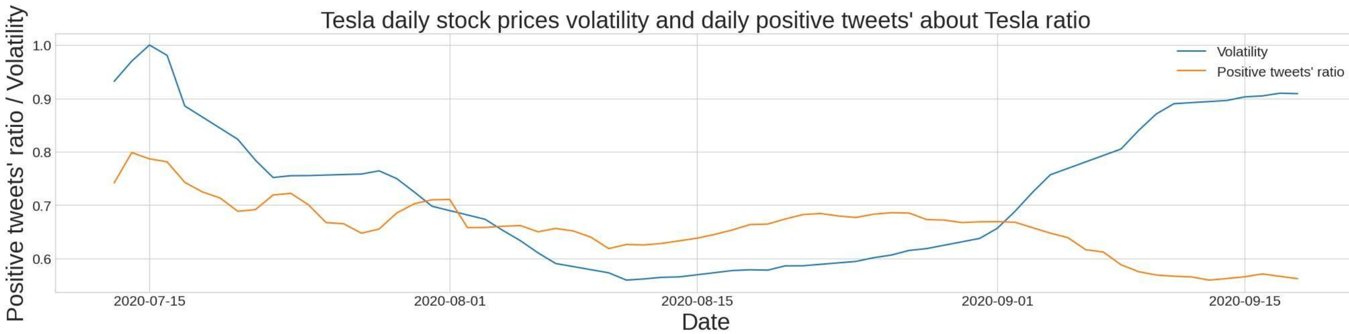

Figure 37: The figure illustrates Tesla daily stock prices volatility and daily positive tweets' about Tesla ratio dynamics

Described above experiments can give us the overall vision of the market and possible connections between its variables. According to evaluated metrics, we can make a conclusion that for different companies the situation primarily varies. However, some common tendencies can be highlighted. For example, the "Macro" news topic group strongly correlates with all companies' stock prices. In further experiments, described in the next section, we tried to investigate the change of training metrics when adding the information about news from different categories sentiments to the Deep Learning model and find out if there is any connection with evaluated correlation and DTW metrics.

4.3.5 Time series predictions

In tables 12-16 some training metrics for the model, which predicts stock prices of 5 companies with financial news as extra data, are presented. In the column "MAPE without sentiments" there are the values of MAPE for the model without news sentiments as external data. As news from different categories are distributed in different periods of time, the metrics for different news topic groups in this case, even if the model is without sentiments, are different, as training performs on different amounts of data. This is the reason why we can not only once evaluate the MAPE of a clean model, trained on all data, and then compare its metrics with the model with sentiments added, but trained on the smaller amount of data. More than that, in the "MAPE change" column there is the MAPE change in percentages from MAPE of the model without sentiments, according to the formula:

$$\frac{MAPE\ (with\ sentiments) - MAPE\ (without\ sentiments)}{MAPE\ (without\ sentiments)} \times\ 100\%$$

Training metrics for Apple			
Topic group name	MAPE (without sentiments)	MAPE (with sentiments)	MAPE change (%)
Analyst Update	6.99%	6.76%	-3.24%
Fed | Central Banks	7.15%	7.97%	11.43%
Currencies	6.38%	6.98%	9.45%

Dividend	6.77%	7.77%	14.73%
Earnings	7.33%	6.81%	-7.13%
Energy | Oil	6.53%	6.81%	4.34%
Financials	7.34%	7.41%	0.89%
General News | Opinion	8.44%	7.52%	-10.86%
Gold | Metals | Materials	6.56%	7.64%	16.45%
M&A | Investments	6.87%	6.99%	1.69%
IPO	4.26%	4.38%	2.91%
Legal | Regulation	6.7%	7.18%	7.14%
Macro	7.75%	7.4%	-4.5%
Markets	7.11%	11.15%	56.9%
Personnel Change	8.72%	8.88%	1.84%
Politics	7.3%	7.33%	0.4%
Company | Product News	6.97%	7.08%	1.64%
Stock Commentary	6.81%	7.99%	17.32%
Stock Movement	7.25%	7.62%	5.13%
Treasuries | Corporate Debt	4.22%	4.39%	3.91%

Table 12: The table, which shows the values of evaluated training metrics for Apple

It is obvious that when MAPE change is negative, there is an improvement in metrics, as low MAPE means high quality of predictions. In tables the improvements are marked with the bold font in the "MAPE change" column.

Financial news category, which mostly improved the predictions of Apple stock prices: "General news | Opinion".

Figure 38: Apple actual stock prices and predictions, made by the TFT model without external data

Figure 39: Apple actual and predicted stock prices, made by TFT model with positive "General news" ratio

Training metrics for Amazon			
Topic group name	MAPE (without sentiments)	MAPE (with sentiments)	MAPE change (%)
Analyst Update	2.92%	3.69%	26.28%
Fed | Central Banks	2.15%	5.39%	150.33%
Currencies	1.79%	3%	67.64%
Dividend	3.83%	36%	839.13%
Earnings	1.79%	2.5%	39.63%
Energy | Oil	2.55%	2.18%	-14.79%
Financials	3.53%	3.27%	-7.37%
General News | Opinion	4.5%	1.94%	-56.88%
Gold | Metals | Materials	5.07%	3.09%	-39.09%
M&A | Investments	2.28%	2.08%	-9.05%
IPO	2%	2.47%	23.87%
Legal | Regulation	2.76%	2.44%	-11.68%
Macro	3.95%	3.34%	-15.49%
Markets	1.96%	5.48%	179.74%
Personnel Change	1.97%	1.94%	-1.45%
Politics	4.19%	2.32%	-44.69%
Company | Product News	3.17%	2.4%	-24.23%
Stock Commentary	2.5%	2.53%	0.99%

Stock Movement	5.03%	6.25%	24.17%
Treasuries | Corporate Debt	2.05%	3.97%	93.53%

Table 13: The table, which shows the values of evaluated training metrics for Amazon

Financial news category, which mostly improved the predictions of Amazon stock prices: "General news | Opinion".

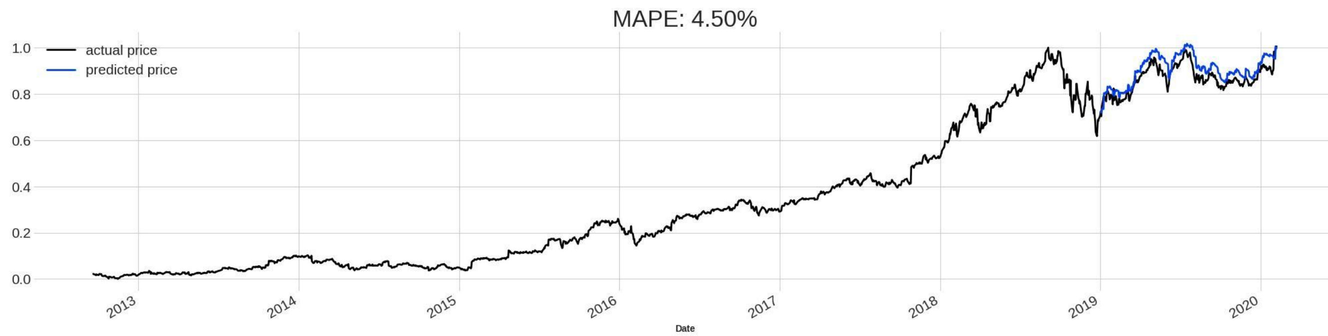

Figure 40: Amazon actual stock prices and predictions, made by the TFT model without external data

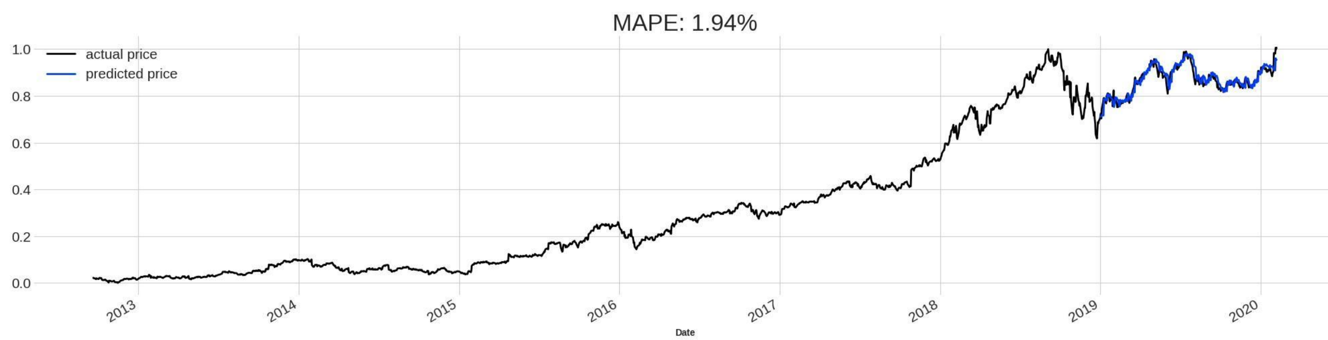

Figure 41: Amazon actual and predicted stock prices, made by TFT model with positive "General news" ratio

Training metrics for Google			
Topic group name	MAPE (without sentiments)	MAPE (with sentiments)	MAPE change (%)
Analyst Update	2.92%	3.44%	17.73%
Fed | Central Banks	2.86%	5.44%	90.35%
Currencies	2.84%	3.15%	11%
Dividend	12.44%	11.65%	-6.39%
Earnings	3.32%	3.53%	6.36%

Energy | Oil	3.4%	2.54%	-25.28%
Financials	4.66%	4.19%	-10.12%
General News | Opinion	3.25%	4.07%	25.35%
Gold | Metals | Materials	4.28%	5.3%	23.77%
M&A | Investments	3.14%	3.82%	21.44%
IPO	2.43%	3.05%	25.62%
Legal | Regulation	5.91%	6.16%	4.24%
Macro	2.74%	4.59%	67.78%
Markets	3.15%	3.67%	16.58%
Personnel Change	4.49%	4.18%	-6.74%
Politics	3.32%	5.74%	72.92%
Company | Product News	5.73%	3.38%	-41.05%
Stock Commentary	4.14%	3.73%	-10.08%
Stock Movement	3.89%	3.58%	-7.97%
Treasuries | Corporate Debt	3.02%	3.06%	1.27%

Table 14: The table, which shows the values of evaluated training metrics for Google

Financial news category, which mostly improved the predictions of Google stock prices: "Company | Product news".

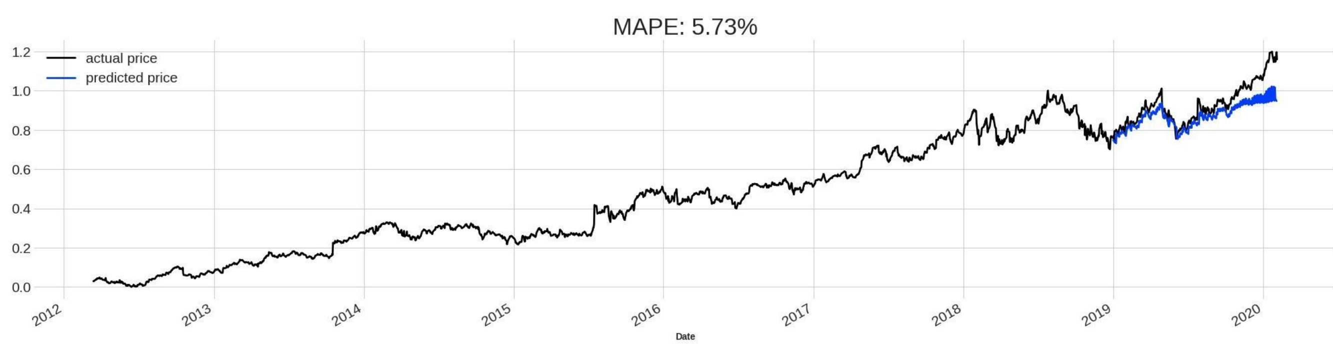

Figure 42: Google actual stock prices and predictions, made by the TFT model without external data

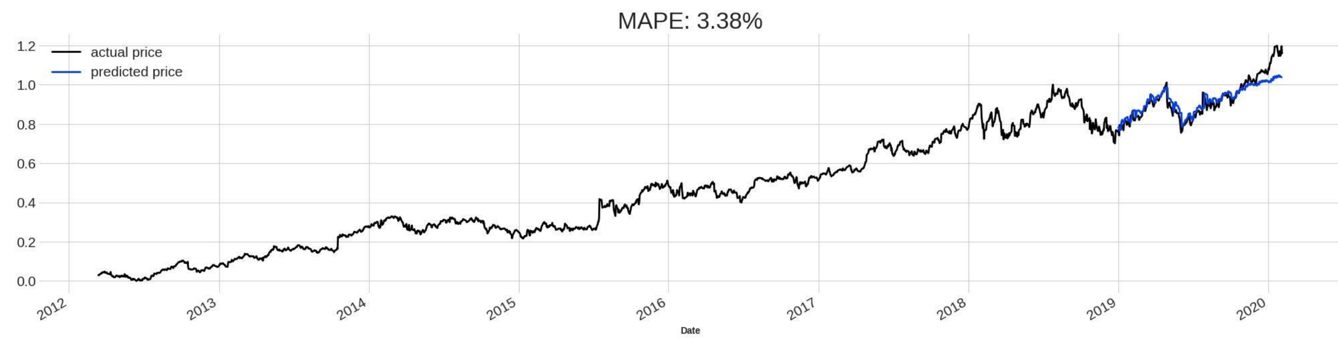

Figure 43: Google actual and predicted stock prices, made by TFT model with positive "Product news" ratio

Training metrics for Netflix			
Topic group name	MAPE (without sentiments)	MAPE (with sentiments)	MAPE change (%)
Analyst Update	3.05%	4.86%	59.42%
Fed | Central Banks	4.33%	3.73%	-13.87%
Currencies	7.03%	10.49%	49.31%
Dividend	4.97%	35.68%	617.68%
Earnings	2.71%	3.24%	19.79%
Energy | Oil	3.86%	6.75%	74.84%
Financials	3.23%	2.48%	-23.04%
General News | Opinion	2.81%	4.22%	50.28%
Gold | Metals | Materials	4.65%	4.8%	3.15%
M&A | Investments	2.68%	3.02%	12.9%
IPO	4.12%	5.95%	44.44%
Legal | Regulation	3.24%	3.28%	0.95%
Macro	5.21%	2.6%	-49.99%
Markets	3.41%	2.98%	-12.57%
Personnel Change	4.35%	5.12%	17.76%
Politics	2.84%	6.63%	133.28%
Company | Product News	3.25%	2.43%	-25.05%
Stock Commentary	3.39%	2.68%	-20.82%
Stock Movement	2.8%	2.77%	-1.35%

Treasuries | Corporate Debt	3.16%	4.76%	50.38%

Table 15: The table, which shows the values of evaluated training metrics for Netflix

Financial news category, which mostly improved the predictions of Netflix stock prices: "Macro".

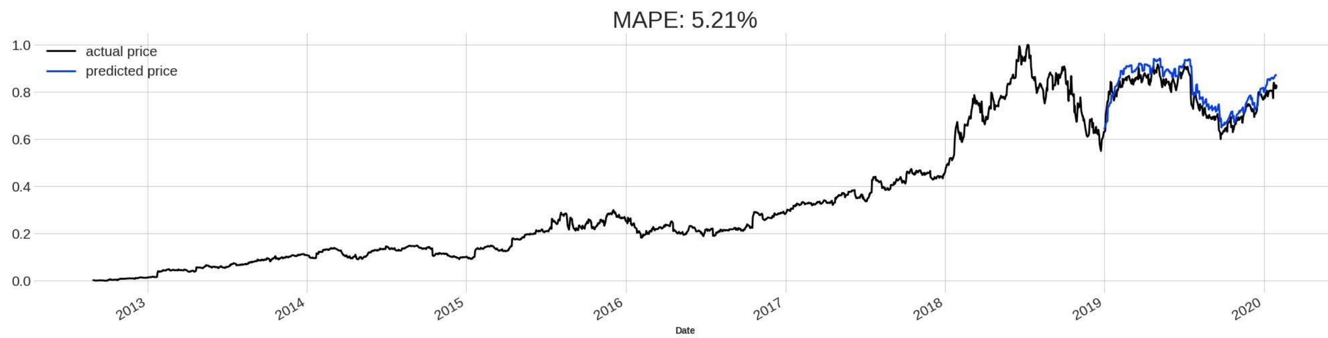

Figure 44: Netflix actual stock prices and predictions, made by the TFT model without external data

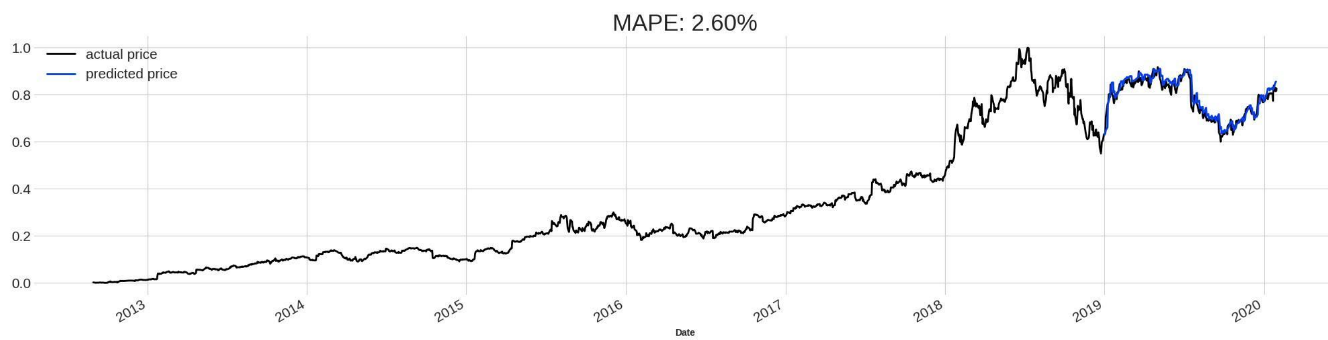

Figure 45: Netflix actual and predicted stock prices, made by TFT model with positive "Macro" news ratio

Training metrics for Tesla			
Topic group name	MAPE (without sentiments)	MAPE (with sentiments)	MAPE change (%)
Analyst Update	7.46%	9.6%	28.72%
Fed | Central Banks	8.15%	12.88%	58.14%
Currencies	12.43%	13.56%	9.05%
Dividend	23.03%	23.67%	2.76%
Earnings	5.77%	5.57%	-3.53%

Energy | Oil	7.3%	7.71%	5.58%
Financials	10.64%	9.82%	-7.65%
General News | Opinion	5.39%	5.74%	6.58%
Gold | Metals | Materials	8.19%	9.75%	18.98%
M&A | Investments	13.16%	9.58%	-27.2%
IPO	9.31%	5.56%	-40.35%
Legal | Regulation	9.11%	9.77%	7.17%
Macro	9.54%	10.69%	12.09%
Markets	5.53%	6.84%	23.77%
Personnel Change	8.48%	12.8%	51%
Politics	4.07%	3.48%	-14.4%
Company | Product News	6.26%	6.51%	3.93%
Stock Commentary	5.07%	5.82%	14.65%
Stock Movement	5.68%	6.48%	14.12%
Treasuries | Corporate Debt	20.25%	83.19%	310.89%

Table 16: The table, which shows the values of evaluated training metrics for Tesla

Financial news category, which mostly improved the predictions of Tesla stock prices: "IPO".

Figure 46: Tesla actual stock prices and predictions, made by the TFT model without external data

Figure 47: Tesla actual and predicted stock prices, made by TFT model with positive "IPO" news ratio

To find out, if the investigation of splitting news into groups makes sense, we performed the following experiment. First we tried to generate a random time series with values from 0 to 1 inclusive (because the positive news ratio is always from 0 to 1) and to pass it to the model. The results (MAPE change) are presented in the first column of table 17. As all values in this column are positive, we can make a conclusion that random time series spoiled the predictions for all companies.

Then we gathered all news topic groups and tried to pass to the model daily positive news ratio among news from all categories together. MAPE change after this stage of experiment is in the second column of Table 17. To the third column we put the best MAPE change, which was received during experiments with topic groups.

MAPE change (in comparison with the model with only time series), %			
Company name	Time series with random sentiments	Time series with all sentiments	Maximum MAPE improvement with topic sentiments
Apple	14.34%	-13.54%	-10.86%
Amazon	76.64%	-58.40%	-56.88%
Google	3.37%	-42.26%	-41.05%
Netflix	86.42%	-43.93%	-49.99%
Tesla	0.5%	-39.88%	-40.35%

Table 17

Conclusions, made based on results from Table 17:

- Random time series spoil the predictions, so it is better to add to the model particular values of news sentiments, not random values.

- For Apple, Amazon, Google all news together give better results, than separate groups.
- For Netflix and Tesla, splitting news into groups can give better results, than utilizing all news together. Especially for Netflix, a separate "Macro" category of news provides significantly bigger improvement than all news together (-49.99% compared to -43.93%).

For figuring out how the change of the model prediction quality is connected with correlation, evaluated in previous experiments, we calculated the correlations between MAPE change and Pearson, Spearman, Kendall's tau correlations' values. Also we evaluated the correlations between MAPE change and Dynamic Time Warping metric value. All results you can see in Table 18. According to received metrics, in most cases MAPE change has an explicit negative correlation with different kinds of correlation. It can be interpreted as the lower MAPE change values correspond with bigger prediction improvements, which in turn correspond with bigger correlation between news sentiments and prices. More than that, some kind of connection between MAPE and Dynamic Time Warping can be revealed, as for some companies we observe the significant positive correlation between mentioned variables, which gives us the natural assumption that lower Dynamic time warping values match better model performance metrics.

Correlation metrics between MAPE and time series similarity metrics					
	Apple	Amazon	Google	Netflix	Tesla
Pearson correlation (MAPE - Pearson correlation)	-0.39	-0.48	0.1	-0.37	-0.06
Spearman correlation (MAPE - Pearson correlation)	-0.2	-0.23	0.13	-0.3	-0.13
Kendall's tau correlation (MAPE - Pearson correlation)	-0.19	-0.19	0.08	-0.24	-0.12
Pearson correlation (MAPE - Spearman correlation)	-0.6	-0.62	0.17	-0.4	-0.14
Spearman correlation (MAPE - Spearman correlation)	-0.24	-0.37	0.05	-0.19	-0.15

Kendall's tau correlation (MAPE - Spear-man correlation)	-0.18	-0.28	0.06	-0.14	-0.1
Pearson correlation (MAPE - Kendall's tau correlation)	-0.53	-0.56	0.08	-0.34	-0.17
Spearman correlation (MAPE - Kendall's tau correlation)	-0.27	-0.33	-0.05	-0.2	-0.15
Kendall's tau correlation (MAPE - Kendall's tau correlation)	-0.2	-0.29	-0.03	-0.15	-0.09
Pearson correlation (MAPE - Dynamic time warping)	0.11	0.28	-0.24	0.4	0.05
Spearman correlation (MAPE - Dynamic time warping)	0.39	0.5	-0.15	0.22	0.1
Kendall's tau correlation (MAPE - Dynamic time warping)	0.27	0.36	-0.14	0.17	0.08

Table 18

In tables 19-20 there is a common statistics, connected with topic groups' contribution in TFT performance.

Companies, which stock prices' predictions were improved		
Topic group name	Amount of compa-nies	Names of companies
Analyst Update	1	Apple
Fed | Central Banks	1	Netflix
Currencies	0	
Dividend	1	Google
Earnings	2	Apple, Tesla
Energy | Oil	2	Amazon, Google

Financials	4	Amazon, Google, Netflix, Tesla
General News | Opinion	2	Apple, Amazon
Gold | Metals | Materials	1	Amazon
M&A | Investments	2	Amazon, Tesla
IPO	1	Tesla
Legal | Regulation	1	Amazon
Macro	3	Apple, Amazon, Netflix
Markets	1	Netflix
Personnel Change	2	Amazon, Google
Politics	2	Amazon, Tesla
Company | Product News	3	Amazon, Google, Netflix
Stock Commentary	2	Google, Netflix
Stock Movement	2	Google, Netflix
Treasuries |Corporate Debt	0	

Table 19: The table, which for each topic group shows amount and names of companies, which stock prices' predictions were improved by adding the external data about the corresponding news topic group sentiments

According to Table 19, the following topic groups were the most effective and improved predictions for the biggest amount of companies: "Financials", "Macro", "Company | Product news ". If we come back to Table 4, we can conclude that the "Financials" group, which improved predictions for 4 out of 5 companies, was one of the most positively colored topic groups, as the ratio of positive news in the group is 95%. "Macro" group has also relatively big positive news ratio (55%), "Company | Product news" category has only 30% of positive news, but the absolute amount of positive news in this group is the biggest, the small ratio was achieved because of the category's big size. All in all, we can suppose that the ability to improve the model predictions by adding daily positive news ratio as external data, to some extent, is determined by the overall amount and ratio of positive news in a particular topic category.

Amount of topic groups, which improved predictions				
Apple	Amazon	Google	Netflix	Tesla
4	10	7	7	5

Table 20

Company name	MAPE without Twitter	MAPE with Twitter	MAPE change (in comparison with the model without Twitter), %
Apple	20.28 %	22.41%	10.53%
Amazon	50.13 %	53.58%	6.87%
Google	124.6 %	60.68%	-51.3%
Netflix	82.58 %	72.54%	-12.17%
Tesla	43.3%	42.14%	-2.66%

Table 21

The analogical experiment with adding to the TFT model as external data was performed on daily positive tweets' ratio. However, due to the small period of time, in which the data, which I gathered, is distributed, It was impossible to train the model qualitatively. In table 21 you can find the results of MAPE metrics, and they are not impressive. However, here we also can mark some relative improvements in prediction quality for 3 companies out of 5. Graphs 48 - 53 illustrate actual and predicted stock prices of Google, Netflix and Tesla with and without daily positive tweets' ratio.

Figure 48: Google actual stock prices and predictions, made by the TFT model without external data

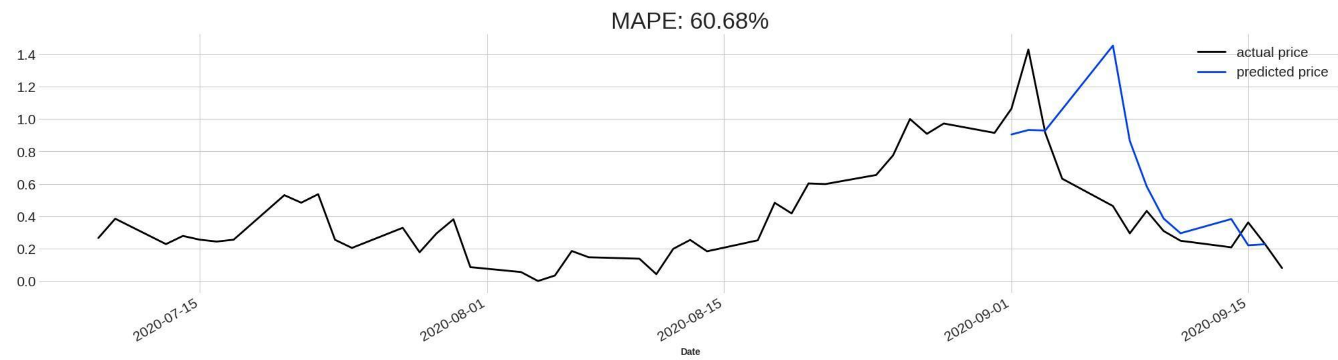

Figure 49: Google actual and predicted stock prices, made by TFT model with positive tweets' ratio

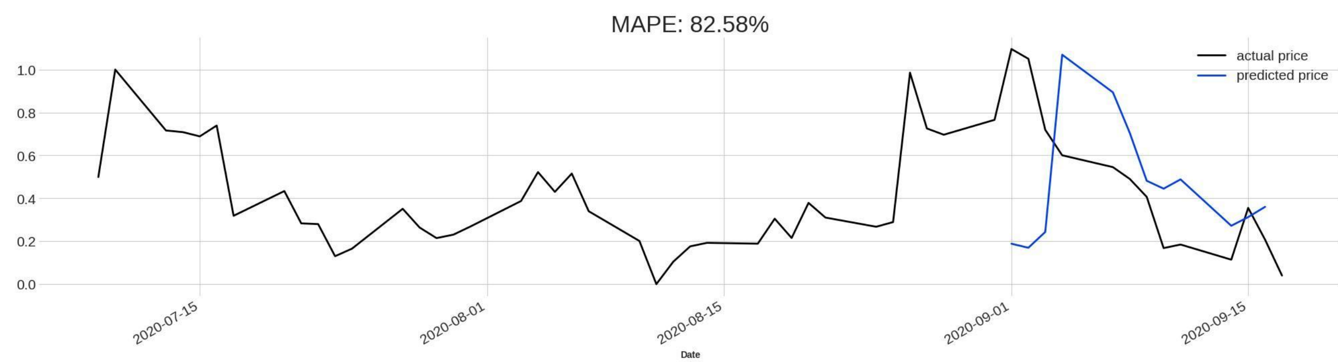

Figure 50: Netflix actual stock prices and predictions, made by the TFT model without external data

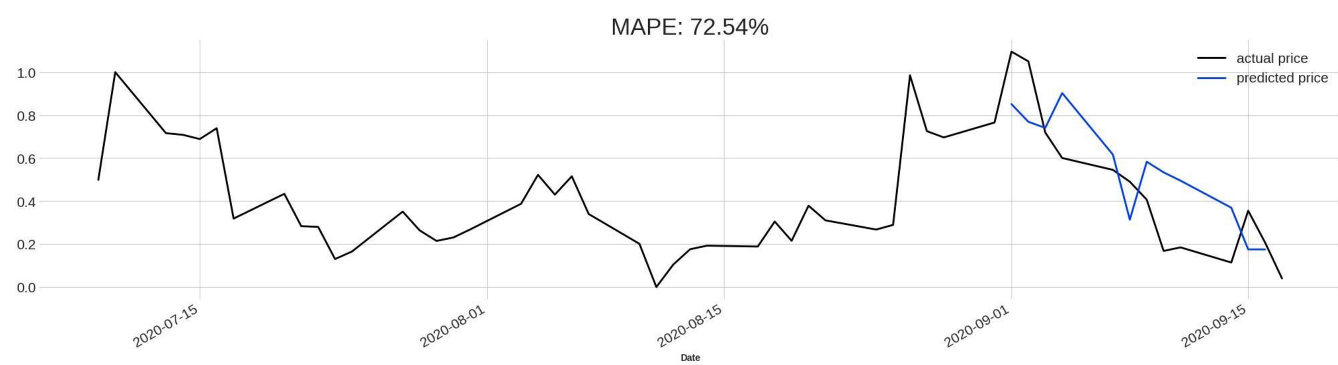

Figure 51: Netflix actual and predicted stock prices, made by TFT model with positive tweets' ratio

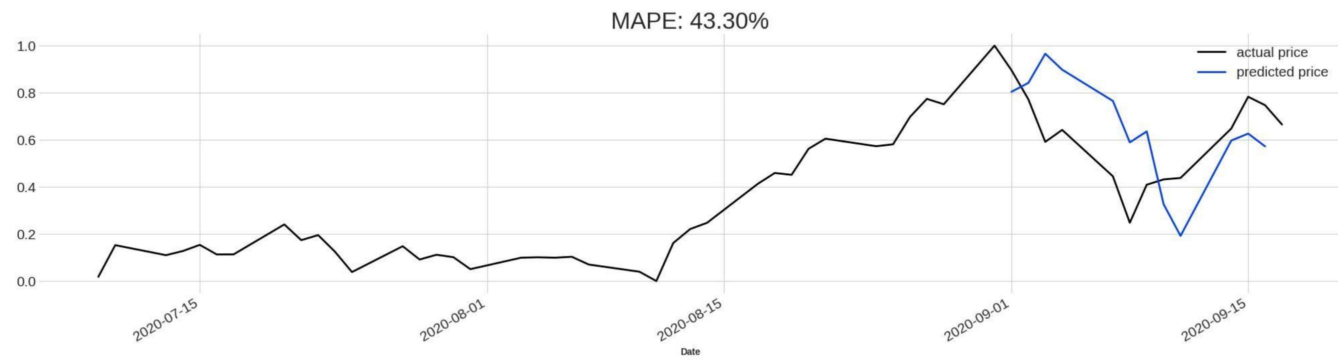

Figure 52: Tesla actual stock prices and predictions, made by the TFT model without external data

Figure 53: Tesla actual and predicted stock prices, made by TFT model with positive tweets' ratio

5 Conclusion

In this research we have achieved the following results:

- Worked out the approach of splitting large amount of news in topic groups
- Annotated the news from each topic group with sentiment labels.
- Evaluated time series similarity metrics between monthly positive news' ratio and monthly mean stock prices of 5 big companies. The metrics, which were evaluated, are: Pearson correlation coefficient, Spearman correlation coefficient, Kendall's tau correlation coefficient, Dynamic Time Warping.
- Evaluated time series similarity metrics for daily positive tweets' ratio with daily stock prices, daily tweets amount with daily stock prices volatility, daily positive tweets' ratio with daily stock prices volatility.
- Run the Temporal Fusion Transformer model 210 times and evaluate MAPE metric, then calculate MAPE change to compare "clean" model and model with external data performances.

The conclusions, which we have made, are the following:

- The situation with the connection between stock prices' behavior and informational background sentiment dynamics varies for different companies, for each company we can find its personal informational patterns, which can help in better understanding of a company's market.
- Some news topic groups strongly correlate with several companies' stock prices and improve TFT model performance for more than 2 companies, which allows us to suppose that there are some common patterns of information, which can work for a large amount of big companies.
- It makes sense to investigate the news topic groups, as sometimes it can help to gain bigger improvement, than by taking into account the full stream of news, investigations in this direction should be continued.

The possible direction for future research is:

- Combining different topic groups with each other. It is reasonable, as according to our results, in some cases it is better to pass to the model not all the news, but its particular groups. It is not always, but the existence of the described situation gives us motivation to continue our research in this sphere.
- Evaluating the MAPE metrics when adding sentiments of news from several categories to the TFT model.
- Comparing the calculated at the previous step metrics with MAPE of the model with news from all categories.
- Taking into account more companies, this will help indicate more common patterns.

References

1. Parhomenko P. A survey and an experimental comparison of methods for text clustering: application to scientific articles. / P. Parhomenko., A. Grigorev, N. Astrakhantsev // *Proceedings of the Institute for System Programming of the RAS* - 29:161–200, 01 2017.
2. Big Tech Companies - Tweet Sentiment [Электронный ресурс] / Kaggle - Режим доступа: https://www.kaggle.com/datasets/wjia26/big-tech-companies-tweet-sentiment, свободный. (дата обращения: 10.04.23)
3. PATEL V., PATEL A. Clustering news articles for topic detection. 2018.
4. Kelly S. Estimating the impact of domain-specific news sentiment on financial assets. / S. Kelly, K. Ahmad // *KnowledgeBased Systems* - 150:116–126, 2018.
5. Örs F. Event clustering within news articles / F. K. Örs, S. R. Yeniterzi // *Proceedings of the Workshop on Automated Extraction of Socio-political Events from News 2020* - pages 63–68, 2020.
6. nickmuchi/finbert-tone-finetuned-finance-topic-classification [Электронный ресурс] / Hugging Face - Режим доступа: https://huggingface.co/nickmuchi/finbert-tone-finetuned-finance-topic-classification, свободный. (дата обращения: 10.04.23)
7. yiyanghkust/finbert-tone [Электронный ресурс] / Hugging Face - Режим доступа: https://huggingface.co/yiyanghkust/finbert-tone, свободный. (дата обращения: 10.04.23)
8. Historical financial news archive [Электронный ресурс] / Kaggle - Режим доступа: https://www.kaggle.com/gennadiyr/us-equities-news-data/tasks, ограниченный. (дата обращения: 10.04.23)
9. Budenny, S. New drugs and stock market: how to predict pharma market reaction to clinical trial announcements / S. Budenny, A. Kazakov, E. Kovtun // [Электронный ресурс]: arXiv:2208.07248 - Режим доступа: https://arxiv.org/abs/2208.07248, свободный (дата обращения: 10.04.23)
10. Sapkota N. News-based sentiment and bitcoin volatility / N. Sapkota // *International Review of Financial Analysis* - 82:102183, 2022.
11. sentence-transformers/all-MiniLM-L6-v2 [Электронный ресурс] / Hugging Face - Режим доступа: https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2, свободный. (дата обращения: 10.04.23)
12. Khan W. Stock market prediction using machine learning classifiers and social media news / W. Khan, M. A. Ghazanfar, M. A. Azam, A. Karami, K. H. Al Youbi, A. S. Alfakeeh // *Journal of Ambient Intelligence and Humanized Computing* - pages 1–24, 2020.
13. Nikou M. Stock price prediction using deep learning algorithms and its comparison with machine learning algorithms. / M. Nikou, G. Mansourfar, J. Bagherzadeh. // *Intelligent Systems in Accounting, Finance and Management* - 26(4):164–174, 2019.

14. Lima B. Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting / Sercan O. Arık, N. Loeff, T. Pfister // [Электронный ресурс]: arXiv:1912.09363 - Режим доступа: https://arxiv.org/pdf/1912.09363.pdf, свободный (дата обращения: 10.04.23)
15. zeroshot/twitter-financial-news-topic [Электронный ресурс] / Hugging Face - Режим доступа: https://huggingface.co/datasets/zeroshot/twitter-financial-news-topic, свободный. (дата обращения: 10.04.23)
16. cardiffnlp/twitter-roberta-base-sentiment [Электронный ресурс] / Hugging Face - Режим доступа: https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment, свободный. (дата обращения: 10.04.23)
17. Budiarto A. Unsupervised news topic modeling with doc2vec and spherical clustering /A. Budiarto, R. Rahutomo, H. N. Putra, T. W. Cenggoro, M. F. Kacamarga, and B. Pardamean // *Procedia Computer Science* - 179:40–46, 2021.
18. Temporal Fusion Transformer [Электронный ресурс] / Darts - Режим доступа: https://unit8co.github.io/darts/generated_api/darts.models.forecasting.tft_model.html, свободный. (дата обращения: 10.04.23)