12-2017

Cryptoeconomics: Data Application for Token Sales Analysis

Seungmin Jin
National Research University Higher School of Economics, sdzhin@edu.hse.ru

Rashid Ali
National Research University Higher School of Economics, rali@edu.hse.ru

A. V. Vlasov
National Research University Higher School of Economics, study@demontroyal.com

Cryptoeconomics: Data Application for Token Sales Analysis

Seungmin Jin [1], Rashid Ali [1], A.V. Vlasov [1]
[1] MSc in Big Data Systems, Faculty of Business and Management
National Research University Higher School of Economics, Moscow, Russian Federation
sdzhin@edu.hse.ru, rali@edu.hse.ru, study@demontroyal.com

Abstract

Initial coin offering, which is called ICO, is rapidly raised its volume as the bypassed crowdfunding. It reached $3.5B in 2017. This is mainly due to five of successful ICOs, Filecoin $257M, Tezos $232M, EOS $185M, Paragon $183M and Bancor $153M. Since these great successes, many startups are jumping into ICO. However, there is no standard criteria to analyze this investment market. Currently, most of token sales rely on their white papers and SNS channels to attract investors, and also we can find its popularity on Google Trends. The goal of this paper is thoughtfully exploring these data sources using big data analysis tools, then building the visualized analytics on this market. Our main approach is based on domain-knowledge to find related criteria. There are several criteria which are widely accepted in this investment markets. We tested the linear correlation between these features and ICO result to validate these criteria. We collected 500 ICO startups to build our database from the ICO-trackers. It consists of founder, white papers, likes and followers in Twitter, ICO results so on. We also collected USD exchange rates of cryptocurrencies to see the price effect of ICO investments. In this data, we considered top 10 ICO startups in 2017 as the successful cases. Using selected criteria, we modeled the prediction of success by the logistic regression method, which may solve our classification problem, successful or not. Our final result is presented as a visual marketing analytics application which is designed on Elasticsearch and Kibana. Kibana is web-based visualization system which connected to the full-search NoSQL database, Elasticsearch. This system analyzes current ICOs status in real-time then visualize whole analysis on the Kibana dashboard. This dashboard provides summarized criteria analysis and the prediction result to monitor ICOs.

Keywords: Data driven analysis, Big Data Application, Visual Analytics, Elasticsearch, Logistic regression, Kibana, Token sales, ICO success, ICO, Cryptoeconomics

Introduction

Nowadays, blockchain technology is being used to launch new cryptocurrency startups. These cryptocurrency startups are using an innovative way of raising the fund which is called Initial Coin Offering (ICO) [1]. Lawrence and Jamie in [2] defined ICO as "it is an event in a project that has cryptographic tokens where part of its token pool is exchanged for money to a community of developers and early adopters". Initial coin offering is rapidly raised its volume by $3.5B in 2017 leaving behind crowdfunding. It is mainly due to five of successful ICOs, Filecoin $257M, Tezos $232M, EOS $185M, Paragon $183M and Bancor $153M [9]. At the same time, initial coin offering is an extremely risky way of investing and raising funds because the process is not regulated by countries, which increases the chances of scam e.g. the Mycelium ICO team members disappeared after raising fund and used that money for their personal needs and hackers stole $7M during the CoinDashâĂŹs ICO by hacking the online wallet of CoinDash [5]. Even so, tons of startups are jumping into ICO, but there are no standard criteria to analyze this token sales market. As the result, risk of investing ICO rapidly growing, since investors even cannot oversee its risk.

The goal of this paper is to explore different ICO data sources, find data features which affect the success of ICO and develop a model to predict and monitor the success of ICO and then develop the visual analytics application on the top of Kibana. We explore data features which can influence the success of ICO. Currently, most of the ICOs rely on white papers, SNS channels, and GitHub to attract investors. Our research work is based on following research questions: (1) Analysis on the format, keywords, and contents of white papers and ICO results. (2) Correlation between SNS channel and ICO results. (3) Correlation between number of tweets, number of followers in SNS channel (Twitter) and ICO results. (4) Correlation between amount of ICO goal (issued tokens) and ICO results. (5) Correlation between the average exchange rate of the coin (in the time frame) and ICO results. (6) Correlation between Google trends (skewness, kurtosis) and ICO results.

In this paper, we introduce a novel method to analyze the token sales market. Our method works in six steps: (1) Collecting data of past ICOs of cryptocurrency startups from ICO-tracker. (2) Cleaning the data and preparing data features. (4) Selecting features by analyzing the correlation of each feature with its success criteria. (5) Developing a machine learning model for predicting the success of an initial coin offering event. (6) Developing a visual analytics dashboard using Kibana. Our main approach is based on using domain-knowledge and supervised learning. There are many crowd-sourcing criteria in this market. We analyzed our data with these criteria to validate them. In this research, we found a number of issued tokens, the format of white papers, chatting channels have strong and moderate correlations on ICO results. In the method, we collected 500 ICOs data from the ICO-trackers and social networks to build a database of ICO results, white papers, number of likes and followers, etc. We also collected USD exchange rates on cryptocurrencies to analyze the price effect of ICO investments. We used logistic regression algorithm to predict the result of ICO. Our algorithm performed well on test data with 96% accuracy. Our final result is presented as a visual analytics application on the top of Kibana. Kibana is the web-based visual querying system connected to the NoSQL database, Elasticsearch. This application analyzes current ICOs data then visualizes whole analysis on the Kibana dashboard. Our application is the first visual analytics application that provides deep insight of ICOs status and prediction.

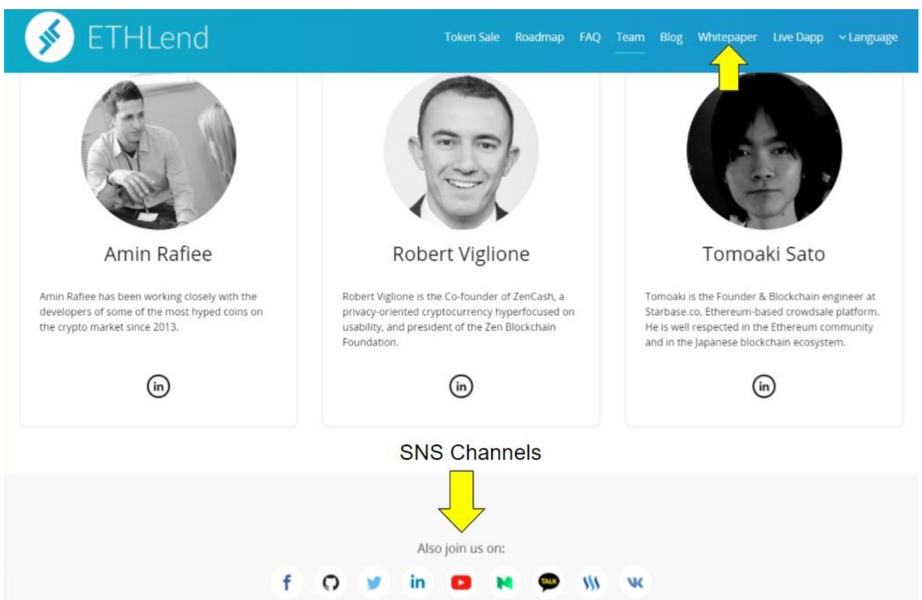

Figure 1: All ICO start-up websites have information of SNS and white paper

Related Work

We focused on analyzing the factors which affect the success of an ICO and predict whether a particular ICO event would be successful or unsuccessful. We want to make discussion on factors which can affect the success of an ICO. In [1], Mohit analyzed interview data of investors related to ICO and his results showed that following factors can affect the success of an ICO: token liquidity, digital community sentiment, quality of information in white papers, local government sentiment towards digital currency, and duration of existence of startups. In [6], Sanam et al. addressed the characteristic of initial coin offering and explored the determinants of ICO. They analyzed data of 253 past ICO and results showed that 81% of ICO was successful to achieve their goal. Their analysis revealed that programming source code, well-organized ICO event, and investorâĂŹs access to project services in future increases the probability of success an ICO event. In our research work, we focused on exploring ICO-trackers, social media communities and quality of information in white papers. We explored data of cryptocurrency startups from social media, Google Trends and ICO-tracker. After collecting data and picking relevant features, we developed machine learning model which gave us outclass results on test data with 96% accuracy.

Method

Here, we explain the method of our study including how we collected and prepared data, which features we found, what kinds of algorithms we used for ICO prediction. The method is designed to answer our minor questions which may lead us to the final answer on ICO prediction. The Followings are our minor questions.

- Correlation between having specific SNS channels and ICO result.
- Correlation between number of likes, reply, followers in Twitter and ICO result.
- Correlation among number of issued tokens, the average price of coin and ICO result in the token sales time frame.
- Correlation between Google Trends and ICO result.
- Correlation between the format of white papers and ICO result.
- Correlation frequents keywords in white papers and ICO result.

Collecting & Preprocessing Data

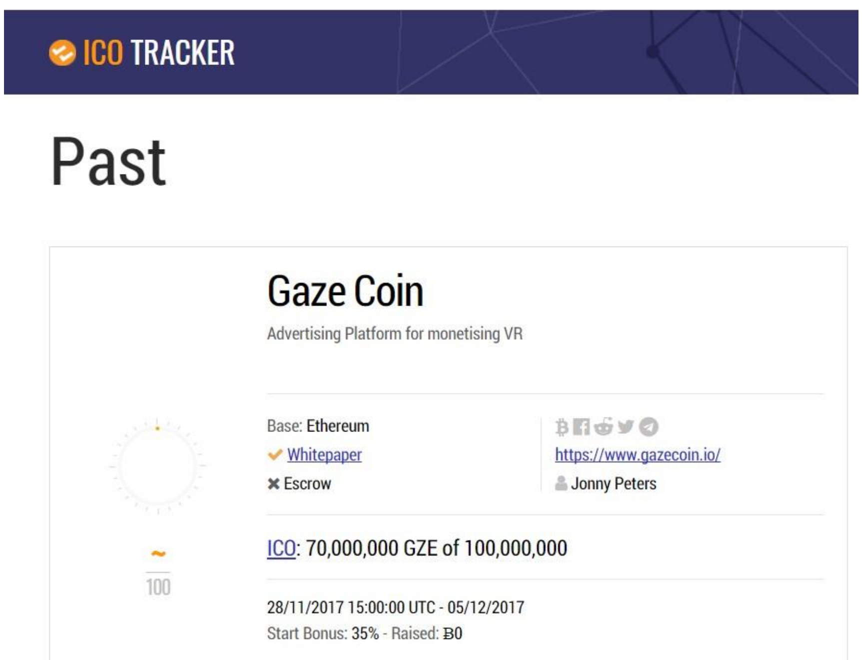

Figure 2: The past ICO section in ICO-tracker.net

Since ICO is still in the early stage, we had to collect and prepare whole data which the research requires, so we made web crawlers to collect list of ICOs from the ICO-tracker websites. ICO-tracker.net [4] provides summary of ICOs with several information. There three main sections in this website, current ICO. upcoming ICO and past ICO. At Nov 24. 2017, we scrapped 517 companies from the past ICO section and created a data table which holds ICO status and hyperlinks of SNS and white paper. This meta data includes name, founder, result, period, white papers hyperlink, SNS hyperlinks. To validate this data, we also built an web crawler to collect ICOs from another ICO-tracker website, Crown+Smith [8]. If there is any difference in ICO data in both website, we dropped the record. After we got all list of ICOs, through specified crawlers, we gathered white papers, SNS data, Google Trends and price exchange rate through the hyperlinks. Although ICO depends on cryptocurrency for their fund raise, not all startups are using Blockchain technology. In our data, 193 start-ups have the keyword, “Blockchain" in their white papers. We also found several startups which have dead hyperlinks. In that case, we considered they don’t have their own white papers.

SNS channels analysis

	Blog	Facebook	Github	LinkedIn	Reddit	Slack	Telegram	Twitter	Youtube	bitcointalk.org
Blog	1.000000	0.126754	0.202748	0.108974	0.250621	0.213902	0.223771	0.222677	0.053683	0.237546
Facebook	0.126754	1.000000	0.006001	0.179677	0.034247	0.147206	0.204070	0.371992	0.153290	0.142436
Github	0.202748	0.006001	1.000000	-0.054138	0.163797	0.149101	0.063204	0.190851	-0.007825	0.118661
LinkedIn	0.108974	0.179677	-0.054138	1.000000	0.047331	0.054757	0.072720	0.107762	0.131028	0.009736
Reddit	0.250621	0.034247	0.163797	0.047331	1.000000	0.188570	0.079539	0.193028	0.026621	0.220859
Slack	0.213902	0.147206	0.149101	0.054757	0.188570	1.000000	0.179554	0.258934	0.079835	0.263367
Telegram	0.223771	0.204070	0.063204	0.072720	0.079539	0.179554	1.000000	0.148782	0.041907	0.253580
Twitter	0.222677	0.371992	0.190851	0.107762	0.193028	0.258934	0.148782	1.000000	0.120336	0.182584
Youtube	0.053683	0.153290	-0.007825	0.131028	0.026621	0.079835	0.041907	0.120336	1.000000	0.073043
bitcointalk.org	0.237546	0.142436	0.118661	0.009736	0.220859	0.263367	0.253580	0.182584	0.073043	1.000000
escrow	0.048096	0.051791	0.111946	0.036596	0.049544	0.198277	-0.025955	0.088418	0.005209	0.270302
ico_goal	0.021782	0.096834	0.012278	0.084228	-0.016569	0.005466	0.034376	0.080019	0.000345	-0.004743
ico_result	0.071002	0.091147	0.017361	0.067267	-0.018981	0.000474	0.023890	0.115626	0.010630	-0.032974

Figure 3: Correlation matrix of SNS channels.

What we want to know about SNS channels is its influence power on ICO results. There are many kinds of SNS channels which startups may use for their marketing purpose. To analysis this, we encoded their SNS hyperlinks as the binary category, 0 and 1, which represents existence of the channel, then run correlation matrix on ICO results. Fig. 3. shows the result of correlation matrix on SNS. ICO results mean number of tokens they sold and it is continuous value. These 10 SNS channels comprises 2 chatting, 4 posting, 2 news, 1 open source hub, 1 video. The whole list of SNS channels are follow.

- Slack
- Telegram
- Blog
- Facebook
- LinkedIn
- Twitter
- Reddit

- Bitcointalk.org
- Github
- Youtube

The result shows that Twitter has the first influence position on ICO result, then Facebook stands the second, Blog and LinkedIn are the third and fourth. Although Bitcointalk.org has one of the popular channel about ICO, but it did not show strong correlation with ICO results. Telegram and Slack, Youtube actually showed no linear correlation with ICO results.

Analysis of number of tweets and follower in Twitter

We also wanted to know the influence of Twitter on its ICO event, especially related to cryptocurrency ICOs, because it stands the first position among SNS channels. Do its twitter channels help investors to invest in ICO event of that cryptocurrency-startups? Does the Twitter channel have help raising funds? To answers these questions, we collected number of tweets and number of followers in the Twitter channel of startups. We developed a web crawler to access the Twitter channels of startups based on twitter URLs. The Twitter channels which did not open or did not contain the number of followers and tweets information were dropped. The number of tweets and number of followers was collected from about 400 twitter channels. We calculated the correlation of number of followers and no. of tweets with the ICO-result and ICO-goal of startups using correlation function from Pandas library with Kendall method. We observed a very weak correlation of no. of followers and no. of tweets with ICO-result and ICO-goal which shows that no. of tweets and no. of followers do not show meaningful influence on ICO-result and ICO-goal.

Analysis on number of issued tokens and the average price of cryptocurrency

We analyzed the effect of the average price of cryptocurrency on the number of tokens issued in ICO. The question is whether the average price of cryptocurrency has linear correlation with the token sales or not. To answer this questions, we gathered the average exchange rate of cryptocurrency from start date to end date of ICO. There are minor cryptocurrencies which are created on the basis of a major cryptocurrency like Bitcoin or Ethereum. According to our meta table from ICO-tracker, we downloaded most of cryptocurrency-startups which launched their own cryptocurrency based on the following 10 cryptocurrencies: Bitcoin, Ethereum, Bitshare coin, Litecoin, Waves, ERC20, Counterparty XCP, Aragon, and Ethereum-classic. So, the exchange rate of a new cryptocurrency will depend on the price of its base cryptocurrency during its ICO event. We calculated average exchange rate of every cryptocurrency in USD based on the price of its base cryptocurrency during the period 01-01-2016 to 17-11-2017 [7]. Then, we got the correlation of average exchange rate with ICO-result and ICO-goal, which means the number of initially issued token. We observed a weak positive correlation between ICO-result and average exchange rate which means that average exchange does not have a strong influence on the ICO-results.

Google Trends analysis on ICO

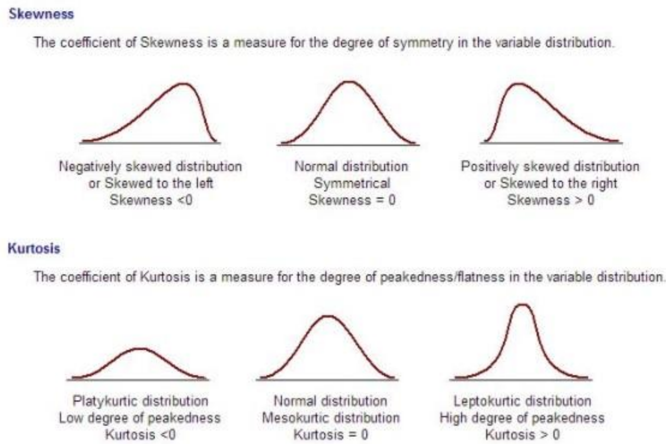

Figure 4: Various chart shapes of skewness and kurtosis

Google Trends is a public web service of Google. This service provides time series data on specific user queries on Google Search. This service shows how often a particular search-term is entered relative to the total search-volume across various regions of the world, and in various languages The horizontal axis of the main graph represents time (starting from 2004), and the vertical is how often a term is searched for relative to the total number of searches, globally. Below the main graph, popularity is broken down by countries, regions, cities and language. [11] In our research, we used 'English' and 'Global' region for query. The predictability on search trends has researched from 2009 and shown good results to represent the trends [12][17][18].

We calculated skewness and kurtosis from 1-year Google Trends to extract popularity of startups. Skewness is asymmetry in a statistical distribution, in which the curve appears distorted or skewed either to the left or to the right. Skewness can be quantified to define the extent to which a distribution differs from a normal distribution. Kurtosis is a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution. [13] That is, data with high kurtosis tend to have heavy tails, or outliers. Data with low kurtosis tend to have light tails, or lack of outliers [4]. figure 4. shows the shape of chart with different skewness and kurtosis. Basically, rapidly increased popularity of startups will shape positive skewness and kurtosis, which means the left-slipped and pointy chart. The reason we chose 1-year time period is all ICO we collected is done in 2017. To achieve this goal, we accessed Google trends API and queried "startup name + founder name" for 1-year trend. The name of startups is often duplicated by other companies or could be very simple words, so we put founder name together to get accurate results. Figure 5. shows the correlation matrix of skewness and kurtosis on whole ICO results. Both have weak positive correlation.

	ico_result	kurtosis	mean	skew	std
ico_result	1.000000	0.124177	-0.088139	0.112844	-0.102308
kurtosis	0.124177	1.000000	-0.787471	0.978336	-0.495041
mean	-0.088139	-0.787471	1.000000	-0.845190	-0.020261
skew	0.112844	0.978336	-0.845190	1.000000	-0.383303
std	-0.102308	-0.495041	-0.020261	-0.383303	1.000000

Figure 5: Google trends analysis result shows rapid popularity has weak correlation on ICO result

format analysis in ICO white papers

	escrow	governance	ico_goal	ico_result
escrow	1.000000	0.018080	-0.043934	-0.060026
governance	0.018080	1.000000	0.124752	0.100859
ico_goal	-0.043934	0.124752	1.000000	0.862450
ico_result	-0.060026	0.100859	0.862450	1.000000
incentives	0.086001	0.290477	0.074339	0.065394
pdf_whitepaper	-0.015336	0.160962	0.148152	0.155216
references	0.110112	0.193537	0.020426	0.006901
roadmap	-0.049441	0.253415	0.074436	0.057449
sci_whitepaper	0.027297	0.106297	-0.074525	-0.100353
solidity	-0.015722	0.478110	0.113604	0.063037
team	-0.055157	0.342435	0.100261	0.037958
token distribution	-0.030223	0.051760	-0.025267	-0.006214
words_num	-0.071891	0.356992	0.124884	0.069236

Figure 6: Correlation matrix of whitepaper format and ICO result

In the concept of value investment, it is important to calculate intrinsic value of the target[14]. The intrinsic value is the actual value of a company or an asset based on an underlying perception of its true value including all aspects of the business, in terms of both tangible and intangible factors. Since startups may not have a real product yet, so this speculation on intrinsic value relies on their indirect product, or information, such like developing teams, technologies they have, vision of the project so on. White papers provide the key information on these ideas to investors. Clearly, the content of white papers is one of the most important things here. However, we focused on the format of white papers instead, because we assumed the content, or popularity, can be observed in SNS and Google Trends. The format of white papers may elevate accessibility and readability, so we analyzed whether these format features have influence on ICO results. Currently, most startups of ICO publish their white papers using one of three file formats and two structure formats.

- File format
  - PDF
  - HTML
  - Google Docs or link
- Structure format
  - Science paper style, IMRAD
  - Solid style, Covering Tech, Team, Road map so on

IMRAD structure represents Introduction, Methods, Results, and Discussion, which is common style in science papers. Solid style explains the details in the structure of ICO contracts which are, Covering Tech, Team, Road map, Token Distribution, Miner incentives, governance. Degree of detail, comprehensiveness, references, depth of thoughts, so on. To analyze these white papers, we extract their file format and text then encode into Boolean value. Figure shows the correlating matrix among style of white papers. Here, solidity means the degree of details in white papers which are counted by number of various sections. In this results, we found PDF file format has weak positive correlation with ICO results and Science paper style has weak negative correlation with ICO results.

Frequent keywords analysis in ICO white papers

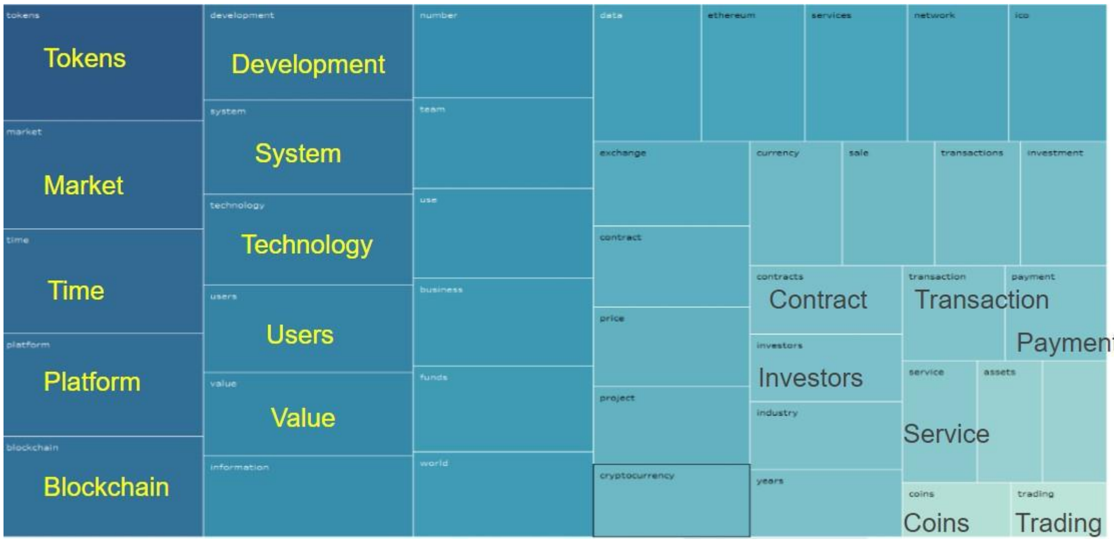

Figure 7: Keywords Treemap based on frequency and ICO result

Keywords in white papers may represent the trend of technologies or there could be some trendy keywords which investors like so that they consider to put their money on. To find these keywords we calculated frequency of keywords from white papers, then draw Treemap visualization using color gradient which will show the ratio of ICO results on total amount of issued tokens. Figure 7. shows the frequent keywords in ICO results in 517 startups. Yellow keywords are discovered in the relatively high ICO results, but green keywords are found in low ICO results. We found "Tokens, Market, Time, Platform, Blockchain, Development, System, Technology, User, Value" in the yellow keywords. "Contract, Investors, Transaction, Service, Coins, Trading" in the green keywords. After we found these keywords, we created the correlation matrix on ICO results for each keywords. Figure 8. shows its the result. Although these keywords form meaningful shape in the Treemap, we discovered there is no strong correlation between these keywords and ICO results.

	blockchain	decentralized	escrow	ico_goal	ico_result
blockchain	1.000000	0.204944	-0.043744	0.095171	0.050595
decentralized	0.204944	1.000000	0.075433	-0.080696	-0.072266
escrow	-0.043744	0.075433	1.000000	-0.043934	-0.060026
ico_goal	0.095171	-0.080696	-0.043934	1.000000	0.862450
ico_result	0.050595	-0.072266	-0.060026	0.862450	1.000000
market	0.854747	0.198859	-0.051864	0.081894	0.050738
pdf_whitepaper	0.419554	0.104204	-0.015336	0.148152	0.155216
platform	0.866327	0.208062	-0.011374	0.109717	0.075502
sci_whitepaper	0.227637	0.262408	0.027297	-0.074525	-0.100353
time	0.884926	0.219452	-0.039187	0.085858	0.030630
tokens	0.832693	0.180221	-0.047810	0.075187	0.052124
words_num	0.821267	0.182754	-0.071891	0.124884	0.069236

Figure 8: Correlation matrix between white paper format and ICO result

Predicting total failure cases

In previous results, we've understood weight of SNS channel and format of white papers, Google Trends. We also found number of issued tokens and escrow has strong and weak positive correlation on ICO results. Based on these features, we made a classification model to predict whether ICO will totally fails or not. Totally fail cases mean that start-up could not sale their issued tokens at all. We used logistic regression.

Logistic regression

$$h_\theta(x) = g(\theta^T x)$$

$$g(z) = \frac{1}{1+e^{-z}}$$

Logistic regression is a statistical method for analyzing data in which there are one or more independent variables that determine the outcome. [15] The outcome is measured with a dichotomous variable which we used 0 or 1 here. Since we have sparse features and want to classify the ICO results as 0 or 1 which represent failed or not, logistic regression shows good performance this type of classification. Currently, this model can predict ICO fails with 97% accuracy.

Building the data application on Kibana

ICO_META(Table)		
NAME	**TYPE**	**DESC.**
uniq_id	INT	
trackers	STRING	List of Trackers
name	STRING	NAME of ico
desc.	STRING	
founder_profile	STRING	
start_date	DATE	
end_date	DATE	
ico_min	FLOAT	(if it is not set, value is 0
ico_max	FLOAT	(if it is not set, value is 0
whitepaper	STRING	HTML LINK
base	STRING	(BITCOION,ETHERUM ...)
escrow	BOOLEAN	
start_bonus	FLOAT	
rasied	FLOAT	
website	STRING	HTML LINK
bitcoin_talks	STRING	HTML LINK
twitter	STRING	HTML LINK
facebook	STRING	HTML LINK
telegram	STRING	HTML LINK
blog	STRING	HTML LINK
slack	STRING	HTML LINK
github	STRING	HTML LINK
linkedin	STRING	HTML LINK
youtube	STRING	HTML LINK

Figure 9: Schema of visualization app data.

Kibana is an open source data visualization plugin for Elasticsearch [10]. It provides visualization capabilities on top of the content indexed on an Elasticsearch cluster. Users can create bar, line and scatter plots, or pie charts and maps on top of large volumes of data. The combination of Elasticsearch, Logstash, and Kibana (also known as ELK stack or Elastic stack) is available as products or service. Logstash provides an input stream to Elastic for storage and search, and Kibana accesses the data for visualizations such as dashboards. [3]

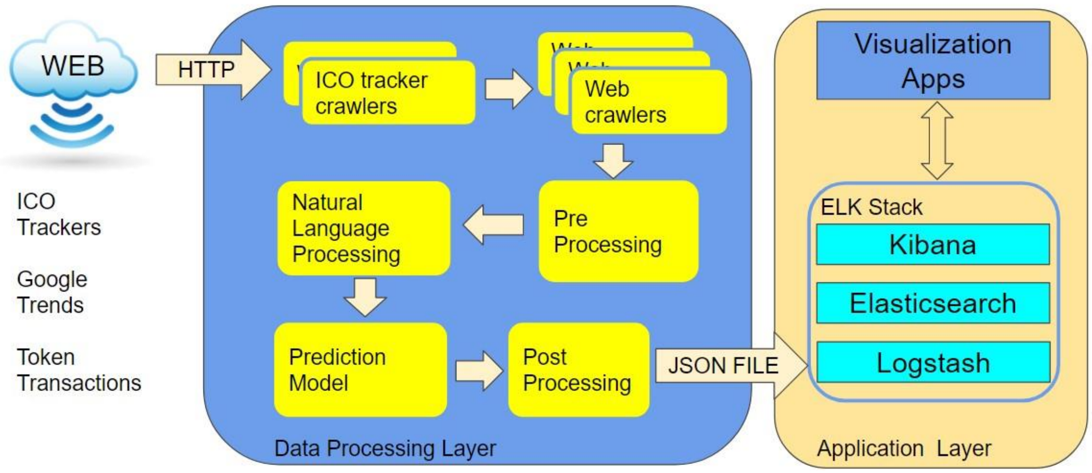

Figure 10: Data application architecture

In the previous steps, we produced data to analyze ICO start-ups cases. Our final goal in this research is creating visual application on the top of Kibana which can show the analysis. Our application will help users to understand ICO status based on our research. Fig. 10 shows how our application works. It consists of two layer, data processing layer and application layer. Processing layer gathers data from web and produces analysis in real time as we studied. After finishing data processing, we get output in JSON format which can store in Elasticsearch. Using this data, we create our visualization and summary. The visual application has three level, monitoring level, prediction level and summary level. Monitoring level shows the analysis of Google Trends, white paper keywords and SNS analysis. Prediction level shows detail prediction and level of risk on the ICO. Summary level shows the list of ICO and summarized feature information. The application handles real time data by hourly based.

Conclusion and Discussion

Figure 11: Visual data application on Kibana

In this research, we mainly explored white papers, SNS and Google Trends data to understand ICO result. Using NLP and machine learning, we found Twitter has the strongest influence on ICO results among SNS and PDF and nonscientific style has positive correlation with ICO results. However, Twitter does not show strong correlation between ICO results and its internal information such like number of likes and tweet, followers. It implies there should be small number of users who greatly affect ICO results in Twitter. Theses strong users would work like information hubs. [16] We also put the sentiment analysis for SNS as future works which may show trend of positive and negative view of potential investors. We think these trends will have strong similarity with Google trends. Our prediction model does not predict the level of success now, but only predict the certain ICO will totally fail or not. However, this prediction may help investors to avoid great risk. In future work, we plan to study deeper on semantic analysis of white paper, code quality analysis in Github to understand startups technologies, and sentiment analysis on SNS. We hope our brief research help our colleagues figure out current status in ICO market.

Acknowledgment

We specially thank to our supervisor A.V.Vlasov who helped us understand the current ICO market, status and its risk. He advised and reviewed this research from the scratch.

References

[1] Exploring signals for investing in an Initial Coin Offering (ICO), Mohit Yadav (0387660), Technische UniversitÃďt Berlin, mohit.yadav@campus.tu-berlin.com

[2] All You Need to Know About Initial Coin Offerings. Outlier Ventures Research, a division of Outlier Ventures. Lawrence Lundy, Jamie Burke, Aron van Ammers

[3] Kibana Official (2017, Oct. 20). Retrieved from https://www.elastic.co/products/kibana

[4] ICO Tracker (2017, Oct. 20). Retrieved from https://icotracker.net/past

[5] Crypto Potato (2017, Jul. 22). Retrieved from https://cryptopotato.com/10-keys-evaluating-initial-coinoffering-ICO-investments

[6] Why do businesses go crypto? An empirical analysis of Initial Coin Offerings, Saman Adhami, UniversitÃă Bocconi, Milan-ITALY, Giancarlo Giudici, Politecnico di Milano, Stefano Martinazzi, School of Management, Milan-ITALY

[7] Coin Market Cap (2017, Oct. 20). Retrieved from https://coinmarketcap.com/

[8] Smith and Crown (2017, Oct. 20). Retrieved from https://www.smithandcrown.com/

[9] Coin Schedule Stats (2017, Oct. 20). Retrieved from https://www.coinschedule.com/stats.html?year=2017

[10] Elasticsearch Officail (2017, Oct. 20) Retrieved from https://www.elastic.co/products/elasticsearch

[11] Google Trends, (2017, Oct. 20) Retrieved from https://trends.google.com/trends/

[12] Shimshoni, Yair, Niv Efron, and Yossi Matias. "On the predictability of search trends." Google Research Blog (2009).

[13] Brown, James Dean. "Skewness and kurtosis." Shiken: JALT testing and evaluation SIG Newsletter 1.1 (1997).

[14] Moore, George Edward. "The conception of intrinsic value." (1998).

[15] Hosmer Jr, David W., Stanley Lemeshow, and Rodney X. Sturdivant. Applied logistic regression. Vol. 398. John Wiley & Sons, 2013.

[16] Smith, Marc A., et al. "Mapping Twitter topic networks: From polarized crowds to community clusters." Pew Research Center 20 (2014).

[17] "Predictability of search trends." Google Research Blog (2009). Choi, Hyunyoung, and Hal Varian. "Predicting the present with Google Trends." Economic Record 88.s1 (2012): 2-9.

[18] Preis, Tobias, Helen Susannah Moat, and H. Eugene Stanley. "Quantifying trading behavior in financial markets using Google Trends." Scientific reports 3 (2013): srep01684.