Irving Fisher Committee on
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

IFC – Bank Indonesia International Workshop and Seminar on *"Big Data for Central Bank Policies / Building Pathways for Policy Making with Big Data"*

Bali, Indonesia, 23-26 July 2018

Measuring stakeholders' expectations for the central bank's policy rate[1]

Alvin Andhika Zulen and Okiriza Wibisono,

Bank Indonesia

[1] This paper was prepared for the meeting. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.

Measuring stakeholders' expectations for the central bank's policy rate

Alvin Andhika Zulen[1] and Okiriza Wibisono[2]

Abstract

In recent decades, the role of market expectation on central bank's policy rate has been increasingly acknowledged in monetary policy formulation. In this research, we develop a machine learning-based technique for identifying the expectation of stakeholders on Bank Indonesia's policy rate. The expectations are extracted from news, starting from 14 days before the monthly Board of Governor's meeting. We achieve an F1-score of 76.8% from out-of-sample evaluation on classification result. The resulting monthly expectation index has 78.6% correlation with the index generated from Bloomberg's monthly survey.

Keywords: policy rate expectation; text mining; machine learning; big data

JEL classification: C02, E52, E58

[1] Statistics Department – Bank Indonesia; E-mail: alvin_az@bi.go.id

[2] Statistics Department – Bank Indonesia; E-mail: okiriza_w@bi.go.id

Contents

1. Background

Expectations on future economic conditions are among the factors that greatly influence the economic actors in making decision. If consumers expect higher inflation in the future, then they increase their consumption expenditures in the present.

One of the indicators that central banks consider in formulating monetary policy is markets' expectation on policy rates. Quoting (Fischer, 2017), "... those times when financial markets and the central bank have different expectations about what a central bank decision will be. Such situations lead to surprises and often to market volatility. " The main objective in measuring expectations on central bank’s policy rate is to avoid market volatility that occurs when market participants have different expectations from the monetary policy taken by central bank. Unexpected movement of Fed Fund Rate is proven in affecting yields of Treasury Bills (Kuttner, 2000) and stock prices (Bernanke & Kuttner, 2004) significantly. If central bank will take a monetary policy that is different from market expectations, a communication strategy is needed so that the volatility in financial markets can be minimized (Fischer, 2017). In addition to avoid volatility, the measurement of policy rate expectations can be an input for projection of macroeconomic indicators, such as inflation and GDP, as implemented by the Monetary Policy Committee (MPC) of the Bank of England (Joyce & Meldrum, 2008).

Because the variable is unobservable, the measurement of expectation on economic indicators, including policy rates, is a nontrivial task. There are two main methods for measuring expectations, i.e. market-based method and survey-based method.

In market-based method, expectations are estimated based on the movement of the price of certain instruments in financial markets. For example, in U.S. financial markets, there is Fed Funds Futures instrument that serves for hedging against changes in The Fed’s monetary policy. The price of this instrument is linked directly with the average of overnight Fed Funds Rate. If the average is decreased then the price of Fed Funds Futures will go up, and vice versa. Thus, expectation on policy rate can be estimated from Fed Funds Futures prices, and changes in expectation are estimated from the instrument’s price movement.

For countries with no interest rate hedging instruments similar to Fed Funds Futures, the measurement of expectation is based on the price of the instrument that moves along with the policy rate, e.g. Treasury Bills, unsecured interbank loan, Forward Rate Agreement (FRA), and Overnight Index Swap (OIS) (Joyce et al., 2008). Nevertheless, measurement with those instruments is more difficult because of additional factors that contribute to pricing, such as credit risk, liquidity risk, and term premium. It is necessary to apply specific calculations and assumptions to exclude these factors in order to obtain an accurate expectation on policy rates.

Survey-based method offers a simpler alternative to measure policy rate expectation. In this method, the survey institution (which can be the central bank itself) asks respondents directly about their expectation on policy rate in the future. This method is also in accordance with recommendation in Manski (2004) that the expectation level can’t be inferred only from the observed choice or action (revealed preference analysis). An expectation measure should be supported by numbers that are explicitly expressed by the respondents.

In Indonesia, Bloomberg conducts a monthly survey of expectations on Bank Indonesia's policy rate (BI 7-day Reverse Repo Rate, formerly BI Rate), i.e. the Economist Estimates Survey. Respondents of the survey are mostly from banking and securities company. Approximately, two weeks before the monthly Board of Governor's Meeting, Bloomberg asked 20-30 respondents about their estimation of Bank Indonesia's policy rate that will be set in the meeting.

This research aims to develop a new measure of stakeholders' expectation on Bank Indonesia's policy rate, as a complement to the Bloomberg survey. From methodological perspective, we show how to utilize textual data to develop the new measure, by employing machine learning-based technique. Based on our observations, a fair amount of expectations on policy rate are quoted in news articles, as seen in Figure 1. Expectations quoted in the news tend to have more varied sources. In addition to market participants, governments, authorities (e.g. Financial Services Authority (OJK), Deposit Insurance Corporation (LPS), Indonesia Stock Exchange (BEI)), and real sector entrepreneurs often express their expectations on Bank Indonesia's policy rate. Hence, it has potential to be used as data source for measuring the expectations.

The paper is organized as follows. In section 2, we provide literature reviews on measuring policy rate expectation and text mining for economic news. In section 3, we discuss the data and methodology. In section 4, we provide a summary of the results and evaluation of the model. In section 5, we conclude the paper and offer some thoughts for future works

Example of Expectation on Bank Indonesia's Policy Rate in News Articles — Figure 1

2. Literature Review

2.1 Survey-Based & Market-Based Method for Measuring Expectation on Policy Rate

Questions on policy rate expectations have been included in various economic and financial surveys. For example, Christensen & Kwan (2014) used the monthly Blue Chip Financial Forecast survey and Survey of Primary Dealers to evaluate whether expectations of market participants are aligned with expectations of the Federal Open Market Committee (FOMC) expectations or not. At Bank Indonesia, the results of Bloomberg survey as described in the previous section are utilized to provide information on policy rate expectation in the Board of Governors Meeting.

Survey-based method has a major advantage over market-based method, i.e. simpler for analysis. Several studies (Christensen & Kwan, 2014; Joyce & Meldrum, 2008; Friedman, 1979) used average or median values to aggregate policy rate expectations of all respondents. For comparison, a research with market-based method (de los Rios & Reid, 2008) used three instrument prices for estimating the probability of Bank of Canada's policy rate changes.

In addition to simpler analysis, we can also calculate the distribution of respondents' expectations with survey-based method. If there are 30 respondents, for example, we can calculate the percentage of respondents who expect a policy rate cut and the percentage of respondents who expect a policy rate hike. In market-based method, the distribution of these expectations can't be provided (Christensen & Kwan, 2014).

However, survey-based method also has several disadvantages compared to market-based method. Market-based method captures the real expectation in the market, i.e. the price of the instrument will move along with market expectation because they are "risking" their money in the instrument (money on the line). Given its subjective nature, in survey-based method, it's possible that the respondents didn't respond according to their actual expectation. Another disadvantage is that the survey-based method is not practical to be done in high frequency (e.g. daily), whereas with market-based method, expectation can be calculated on a daily basis or even from hour-to-hour, if the referred instruments are widely traded.

2.2 Text Mining on Economic News

Text data have been widely used for research in economics and finance. Sahminan (2008) identified keywords that reflect a tight, neutral, or loose monetary policy inclination in the press release statement of Bank Indonesia over the period from January 2004 to December 2007. The econometric analysis shows that monetary policy statements that contain loose or neutral policy inclination tend to lower interbank interest rates, while monetary policy statements with tight policy inclination tend to have no impact on interbank interest rates (asymmetric effect). Rosa & Verga (2007) applied similar method to analyze the impact of European Central Bank (ECB) press releases.

In those studies, the identification of keywords in the press release text is done manually. Researchers read the press releases one by one and record the keywords that appear in the press releases. Nowadays, text mining algorithms are growing rapidly along with the adoption of big data and machine learning. These algorithms can automatically "read" and “extract” relevant information from the text, such as the person's name, the organization’s name, and the keywords. Compared to the manual way, text mining allows us to make use of much larger text data than press releases, including news and social media.

Bollen et al. (2011) proved that the mood expressed by Twitter users can be analyzed to improve the stock market prediction. Moods are identified using keywords, e.g. "I feel ..." and "I’m ...", and then categorized into different types of mood by using OpinionFinder and Google-Profile of Mood States (GPOMS). Similar to (Bollen et al., 2011), O'Connor et al. (2010) created a public sentiment index from positive and negative word occurrences in economic related tweets. This index correlated with the Gallup’s Economic Confidence Index at 73.1% and with the Index of Consumer Sentiment (ICS) from the Reuters/University at 63.5%.

In addition to social media data, news data are also widely used to analyze economic conditions. Baker et al. (2016) developed an Economic Policy Uncertainty (EPU) index by using news articles from 10 leading U.S. newspapers. The EPU index reflects the frequency of articles that contain the following trio of terms: economic (‘‘economic’’ or ‘‘economy’’); policy (‘‘Congress,’’ ‘‘deficit,’’ ‘‘Federal Reserve,’’ ‘‘legislation,’’ ‘‘regulation,’’ or ‘‘White House), and uncertainty (‘‘uncertain’’ or ‘‘uncertainty’’). The EPU indexes have also been constructed for 11 other countries with list of keywords that are tailored to the language and economy.

In terms of monetary policy, Nardelli et al. (2017) developed the Hawkish-Dovish (HD) index that measures media’s perception of ECB communications. The HD index is computed by using two methods: semantic orientation (SO) and support vector machine (SVM). The HD index based on SO method is computed by counting the co-occurrences of strings with a fixed set or pre-determined words/expressions that are normally associated with “hawkish” and “dovish” concepts to determine the tone of the document. For the SVM method, instead of using predefined set of keywords, the algorithm automatically looks for patterns in text documents to select the words with the highest discriminative power and determines the tone of a document based on them. Similar Hawkish-Dovish research has also been done earlier by Lucca & Trebbi (2009) for the FOMC statements.

Those two studies measured media’s perception after each press conference following monetary policy meetings. As far as our observation, there is no research utilizing news data to measure policy rate expectation before the monetary policy meetings.

3. Methodology

3.1 Data

3.1.1 News Articles

The news data used in this research obtained from Bank Indonesia’s Cyber Library. Cyber Library is an internal repository of news articles related to economic and financial topics. The news articles data are available on a daily basis since 1999, thus covering the whole period since Bank Indonesia set the policy rate (BI Rate) in July 2005. The data used in this research are from January 2006 to February 2018.

3.1.2 Policy Rate Expectation Survey

In order to measure the accuracy of policy rate expectation obtained from the news, a benchmark indicator is required for comparison. In this research, we use Economist Estimates Survey from Bloomberg, as described in the first chapter. Survey results are available starting from two weeks before the monthly Board of Governor's meeting, although data from several respondents are often only available close to the date of the meeting. Each respondent gives their estimation on Bank Indonesia's policy rate which they think will be set in the meeting. An example of the survey result is shown in Figure 2.

Example of Bloomberg's Economist Estimates Survey Figure 2

3.2 Machine Learning Model

In order to extract the policy rate expectation from news articles automatically, we build a text mining model by using machine learning-based technique. This section will describe the steps taken in developing the model.

3.2.1 Data Filtering

News articles collected from Bank Indonesia's Cyber Library are not entirely relevant for measuring policy rate expectations. First of all, the news articles are filtered in following steps:

1. Publication Date Filtering

From all the news articles available in Cyber Library, we only used news articles that are published within 14 to 1 days prior to each monthly Board of Governor's meeting.

2. Sentences Spliiting

News articles are splitted into sentences to simplify the extraction of policy rate expectation. Text splitting is done automatically by using Natural Language Toolkit (NLTK) in Python.

3. Keywords Filtering

Sentences from the previous step are filtered again, leaving only sentences that contain keywords related to Bank Indonesia's policy rate, e.g. "BI Rate", "BI 7-days reverse repo rate", and "Bank Indonesia's policy rate".

Thus, the result from these stages is a collection of sentences containing keywords related to Bank Indonesia's policy rate and published on D-14 to D-1 prior to each monthly Board of Governor's meeting. In total, there are 5,700 news articles (2% of overall news in Cyber Library) and 16,000 sentences (0.2% of overall sentences in Cyber Library) that meet the specified criteria.

3.2.2 Annotation

Text mining that make uses of machine learning techniques require annotated datasets for training the algorithms. Annotation is the process of attaching additional information into a collection of texts. Annotation is needed to "teach" the text mining algorithm how to extract the information from the texts, so that the process can be done automatically in the future.

In this research, annotation is done on sentence-level, as the smallest data unit. We added a categorical information about policy rate expectation to each sentence, with 4 (four) possible values as follows:

1. 0: sentence with no expectation information;
2. 1: expecting no change in policy rate;
3. 2: expecting policy rate hike;
4. 3: expecting policy rate cut.

This categorical information will be used as target class in machine learning algorithms.

Each sentence is annotated by two annotators to minimize human error and subjectivity. If a sentence is annotated differently by both annotators, the sentence will be annotated by the third annotator. We also provide an annotation guidance so that the annotations can be given consistently by each annotator.

In total, we collected 4,445 sentences that have been annotated, out of 16,000 sentences generated in previous steps. Table 1 shows the proportion of sentences for each policy rate expectation category.

Annotated Sentences

Table 1

Policy Rate Expectation Category	Number of Annotated Sentences	Percentage (%)
Policy Rate Hike	355	8%
Policy Rate Cut	660	15%
Policy Rate Unchanged	490	11%
No-Expectation	2,940	66%

3.2.3 Pre-processing

After annotating the sentences, one more step is required in order to start training the classification model using machine learning algorithms. Each sentence must be transformed into numerical vector, because machine learning algorithms can only process numerical data.

Each sentence is transformed into numerical vector that contains following information:

1. bag-of-keywords: number of keywords' occurences in the sentence;
2. number of words in the sentence;
3. number of characters in the sentences;
4. numbers and percetages quoted in the sentence;
5. word embedding vector.

All transformations are done by using Pandas and Scikit-learn libraries in Python.

3.2.4 Model Construction

Sentences that have been annotated and transformed into numerical matrix (1 line = 1 sentence) are used as input for machine learning algorithms. Machine learning algorithms will learn the patterns in input data to construct classification model with target function to classify the category of policy rate expectation.

$$\hat{f}(sentence\_vector) \in \{rate\ hike, rate\ cut, rate\ unchanged, no\ expectation\}$$

The data are splitted into 2 datasets: training dataset and test dataset. Training dataset is used to build the classification model. Test dataset is used in model evaluation to provide unbiased evaluation on the model. We split the data using approximately 80:20 ratio (training dataset: 3,645 sentences; test dataset: 800 sentences).

We use 5 (five) machine learning algorithms in this research to find the best classification model for solving the task, i.e.:

1. Logistic regression: modeled the linear relationship between independent variables and the expectation category as dependent variable;
2. Naïve bayes: modeled the probability of expectation category based on Bayes' theorem with the independence assumptions between predictors;
3. Decision tree: modeled the decision tree that predict expectation category (represented in the leaves) based on a set of decsion rules (represented in the branches);
4. Random forest: combined the predictions of multiple decision trees with bootstrapping aggregation; and
5. Xgboost: an implementation of gradient boosted tree by DMLC (http://dmlc.ml/).

3.3 Index Calculation

3.3.1 Expectation Index from News

The best classification model that has been constructed in previous step is then applied to classify the policy rate expectation category on all 16,000 sentences in the dataset. From the classification results, we calculate the monthly policy rate expectation index in following steps:

1. Each sentence with policy rate expectation is given a score: +1 for expecting policy rate hike; -1 for expecting policy rate cut; 0 for expecting no change in policy rate. Sentences with no information on policy rate expectation ware excluded from index calculation.
2. Each news article is given a score: the mean score of sentences (as calculated in 1st step) in the article.
3. The expectation index from news for month t is defined as the mean score of articles (as calculated in 2nd step) that are published in that month.

$$Expectation\ Index\ News_t = \frac{1}{|C_a|}\sum_a score(a) = \frac{1}{|C_a|}\sum_{s_a}\left(\frac{1}{|C_{s_a}|}score(s_a)\right)$$

$|C_a|$ = number of articles in month t

$score(a)$ = score of article a

$|C_{s_a}|$ = number of sentences in article a with policy rate expectation

$score(s_a)$ = score of sentence s in article a

The monthly expectation index has following characteristics:

- Range of index: [-1,+1].
- The index will be close to +1 if there are more news with expectation of policy rate hike.

  The index will be close to 0 if there are more news with expectation of unchanged policy rate.

  The index will be close to -1 if there are more news with expectation of policy rate cut.
- Positive index means more news with expectations of policy rate hike compared to policy rate cut.

  Negative index means more news with expectations of policy rate cut compared to policy rate hike.
- If $index_{t1} > index_{t2}$ then the proportion of news with expectation of policy rate hike is greater in t_1 than in t_2.

  If $index_{t1} < index_{t2}$ then the proportion of news with expectation of policy rate cut is greater in t_1 than in t_2.

3.3.2 Expectation Index from Bloomberg Survey

As described earlier in section 1, in the Economist Estimates Survey, Bloomberg asked respondents about their estimation on Bank Indonesia's policy rate that will be set in the next Board of Governors' meeting. These estimation numbers need to be converted so that they are comparable with the expectation index. The conversion is done as follows:

$$score(x)_t = \begin{cases} +1 : if\ prediction(x)_t > BI\ Rate_{t-1} \\ 0 : if\ prediction(x)_t = BI\ Rate_{t-1} \\ -1 : if\ prediction(x)_t < BI\ Rate_{t-1} \end{cases}$$

$score(x)_t$ = score of respondent x in month t

$prediction(x)_t$ = policy rate prediction respondent x in month t

$BI\ Rate_{t-1}$ = Bank Indonesia's policy rate in month $t-1$

The expectation index from Bloomberg survey for month t is defined as the mean score of all respondents in the month.

$$Expectation\ Index\ Bloomberg_t = \frac{1}{|C_x|}\sum_x score(x)$$

$|C_x|$ = number of respondents in month t

4. Result & Analysis

4.1 Classification Model Evaluation

Classification models that have been trained in the previous steps need to be evaluated in order to measure their accuracy in predicting the target class (i.e. policy rate expectation). We use F1-score as metric for evaluation, in order to get a balanced classification model with the optimal balance of recall and precision.

The result of out-of-sample evaluation for each machine learning model are given in Table 2. We can see that the logistic regression model has the best F1-score (76.8%), compared to other machine learning models. The model also has the best accuracy and precision score. Hence, the logistic regression model becomes our choice for measuring policy rate expectations in the following sections.

Classification Model Evaluation

Table 2

Classification Model	Accuracy	Recall	Precision	F1
Logistic regression	83.4%	**83.2%**	71.2%	**76.8%**
Naïve bayes	80.6%	83.2%	64.5%	72.7%
Decision tree	73.0%	65.7%	53.4%	58.9%
Random forest	78.0%	72.6%	63.3%	67.6%
XGBoost	**84.1%**	75.9%	**75.6%**	75.7%

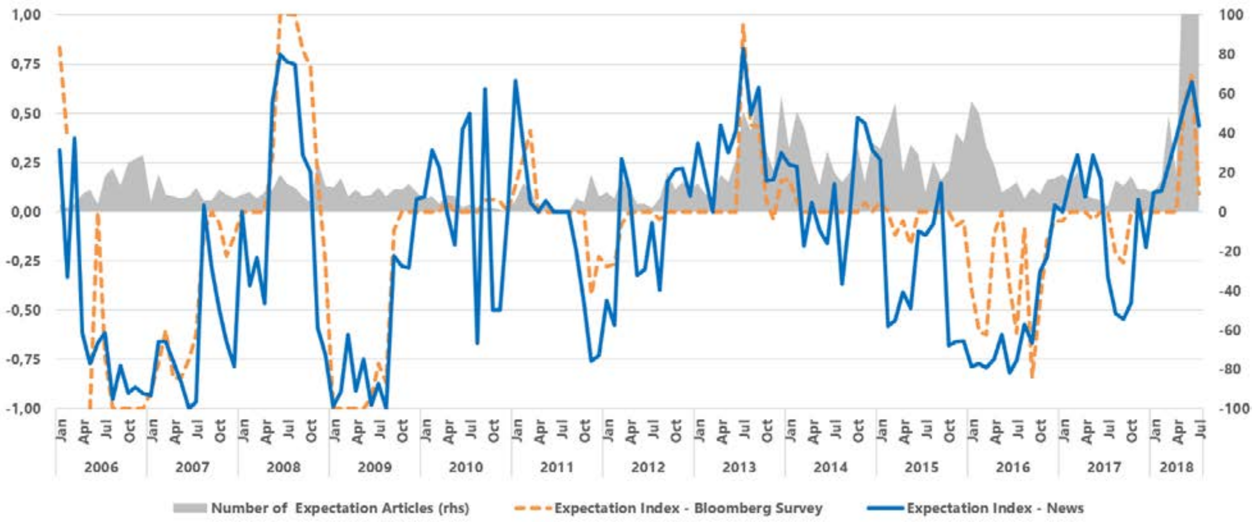
Note: Blue-shaded cells denote the best result for each evaluation metric

4.2 Result Evaluation

For result evaluation, we calculate the correlation between policy rate expectation index generated from news and from Bloomberg survey. Graphs of both indices from January 2012 to July 2018 are presented in Figure 3. We can see that both indices are moving in the same direction generally, with a correlation of 73% (correlation for full data period, i.e. from January 2006, is 78.6%). The correlation value indicates that the policy rate expectation index from news is potential to be used as a new measure of policy rate expectation.

The policy rate expectation index from news tends to be more volatile, e.g. in the second half of 2010. This is likely due to there are some periods (months) where number of sentences containing policy rate expectation in Cyber Library is very low. From 142 months of data, there are 49 months where the number of sentences containing policy rate expectation are less than 10.

Plot of Policy Rate Expectation Index	Figure 3



For some periods, the expectation index from news can "predict" the direction of policy rate more precisely than the expectation index from Bloomberg survey, as presented in Table 3. Overall, compared to the actual change in policy rate, the expectation index from news has a correlation of 76.6%, while the expectation index from Bloomberg survey has a correlation of 84.5%.

Comparison between Expectation Index from News and from Bloomberg Survey — Table 3

Period	Event	Expectation Index from News	Expectation Index from Bloomberg Survey
December 2007	Policy rate cut	-0.73	-0.13
February 2011	Policy rate hike	0.61	0.27
November 2011	Policy rate cut	-0.80	-0.42
February 2012	Policy rate cut	-0.63	-0.27
September 2013	Policy rate hike	0.84	0.44
November 2013	Policy rate hike	0.62	-0.04
February 2015	Policy rate cut	-0.56	0.00
June 2016	Policy rate cut	-0.78	-0.38
September 2017	Policy rate cut	-0.53	-0.26
May 17 2018	Policy rate hike	0.53	0.55
May 30 2018 (additional)	Policy rate hike	0.67	1.00
June 2018	Policy rate hike	0.66	0.69

5. Conclusion & Future Work

5.1 Conclusion

In this research, we develop a new measure of stakeholders' expectation on Bank Indonesia's policy rate. From methodological perspective, we show how to utilize news articles data to develop the new measure, by employing machine learning-based technique. The expectations are extracted from news, starting from 14 days before the monthly Board of Governor's meeting. The machine learning model is trained by using sentences that have been annotated manually.

From out-of-sample evaluation, we achieve an F1-score of 76.8% on classification accuracy by using logistic regression model. The resulting monthly expectation index has 78.6% correlation with the expectation index generated from Bloomberg's monthly survey.

5.2 Future Work

There are several improvements in the methodology that can be applied for future works.

- Opinion Holder Identification

Currently, the calculation of the expectation index of each month use the average score of the articles. This makes the index is not entirely comparable to expectation measure obtained from Bloomberg survey (news articles vs. survey respondents). We need to identify the opinion holder for each sentence that contains policy rate expectation. Once identified, opinion holders whose expectations are quoted in several articles are counted only once in index calculation.

Another benefit of opinion holder identification is for grouping expectations based on institutional group of the opinion holder, e.g. government, authorities, banking, capital market, industry, academics, and research institutes. Thus, we can further examine which institutional groups expect policy rate hike, cut, or unchanged.

- Data Source Addition

The number of news articles used in this research is not big enough, i.e. 5,700 news articles in 146 months, or about 40 news articles per month. The addition of new data sources can be done with web crawling on online news websites. In addition, we also consider to use news in English language, although additional works are needed to develop a text mining model for English language.

- Classification Model Improvement

Nowadays, artificial neural network (especially deep learning) is state-of-the-art technique for text classification, including opinion extraction task (Irsoy and Cardie, 2014). The currently used classification model, i.e. logistic regression, can be replaced with a neural network model to improve the accuracy. However, it is necessary to annotate more sentences, given the neural network model requires a large amount of training data.

- Expectation vs. Wish vs. Suggestion

Currently, annotated sentences also include phrases of wishes, hopes, and suggestions on the policy rate. We need to separate sentences that contain expectation (or prediction) with sentences that contain wish (or suggestion), so that the index only contains information related to expectations. Rule-based method (using keywords e.g. "expects" vs. "wishes") or machine learning method could be used for the task.

- Expectation Period Identification

Sometimes, sentences that contains policy rate expectations are not referring to the next Board of Governors' meeting, but rather several months or even a year later (e.g. "He predicts BI Rate to be hiked only one more time this year, at the end of 2014."). Such sentences need special handling, i.e. by classifying it as expectation of unchanged policy rate for the next meeting, and as expectation of policy rate hike for meeting at the end of 2014.

References

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics, 131*(4), 1593-1636.

Bernanke, S. B., & Kuttner, K. N. (2004). What Explains the Stock Market's Reaction to Federal Reserve Policy? *The Journal of Finance, 60*(3), 1221-1557.

Bollen, J., Mao, H., & Xiao-Jun, Z. (2011). Twitter Mood Predicts the Stock Market. *Journal of Computational Science, 2*(1), 1-8.

Christensen, J. H., & Kwan, S. (2014). *Assessing Expectations of Monetary Policy*. Retrieved from FRBSF Economic Letter: https://www.frbsf.org/economic-research/publications/economic-letter/2014/september/assessing-expectations-monetary-policy/

de los Rios, A. D., & Reid, C. (2008). Extracting Policy Rate Expectations in Canada. *Capital Markets: Asset Pricing & Valuation eJournal*.

Fischer, S. (2017). *Monetary Policy Expectations and Surprises*. Retrieved from Speeches of Federal Reserve Officials: https://www.federalreserve.gov/newsevents/speech/fischer20170417a.htm

Friedman, B. M. (1979). Interest Rate Expectations Versus Forward Rates: Evidence from an Expectations Survey. *The Journal of Finance, 34*(4), 965-973.

Irsoy, O., & Cardie, C. (2014). Opinion Mining with Deep Recurrent Neural Networks. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (pp. 720-728).

Joyce, M., & Meldrum, A. (2008). Market Expectations of Future Bank Rate. *Bank of England Quarterly Bulletin 2008 Q3*, pp. 274-282.

Joyce, M., Relleen, J., & Sorensen, S. (2008, December). Monetary Policy Expectations from Financial Market Instruments. *ECB Working Paper Series No.978*.

Kuttner, K. N. (2000). Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market. *Journal of Monetary Economics, 47*(3), 523-544.

Lucca, D. O., & Trebbi, F. (2009). Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements. *NBER Working Paper No. 15367*.

Manski, C. F. (2004). Measuring Expectations. *Econometrica, 72*(5), 1329-1376.

Nardelli, S., Tobback, E., & Martens, D. (2017). Between Hawks and Doves: Measuring Central Bank Communication. *ECB Working Paper Series No. 2085*.

O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, (pp. 122-129).

Rosa, C., & Verga, G. (2007). On the Consistency and Effectiveness of Central Bank Communication: Evidence from the ECB. *European Journal of Political Economy, 23*(1), 146-175.

Sahminan, S. (2008). Effectiveness of Monetary Policy Communication in Indonesia and Thailand. *Bank for International Settlements Working Paper No.262*.

Irving Fisher Committee on
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

IFC – Bank Indonesia International Workshop and Seminar on *"Big Data for Central Bank Policies / Building Pathways for Policy Making with Big Data"*

Bali, Indonesia, 23-26 July 2018

Measuring stakeholders' expectations for the central bank's policy rate[1]

Okiriza Wibisono,

Bank Indonesia

[1] This presentation was prepared for the meeting. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.

Measuring Stakeholders' Expectations for the Central Bank's Policy Rate

Okiriza Wibisono, Alvin Andhika Zulen, Anggraini Widjanarti, Hidayah Dhini Ari
Bank Indonesia

Disclaimer

The following opinions are those of the authors and not necessarily those of Bank Indonesia

Expectation on monetary policy

Trend and expectation on inflation

Trend of past policy rate

Trend and expectation on other CB policy rates

(Expectation on) gov't policies, e.g. gas prices

Other info and indicators

BI policy rate will …

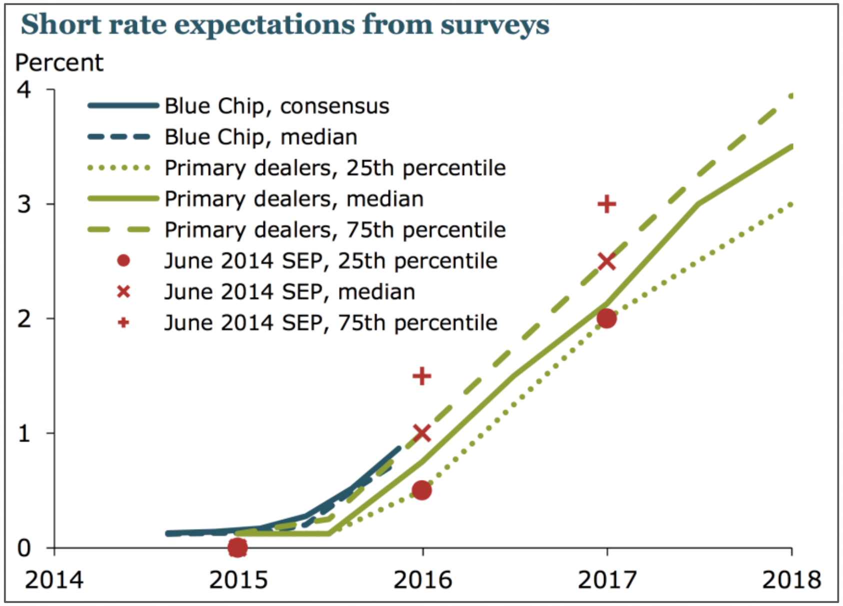
Financial markets, real sector, analysts, central & local gov't, general public

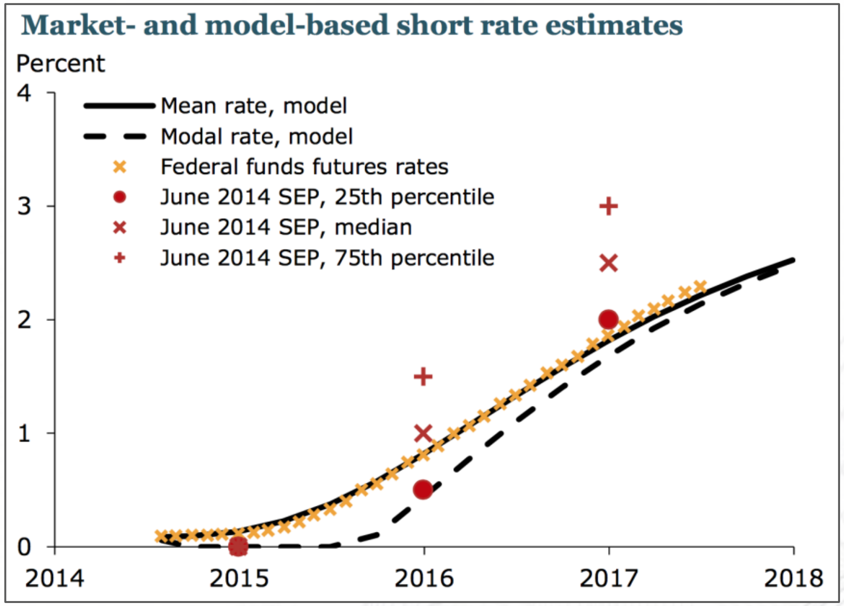
Rise? → Behavior A

Stay? → Behavior B

Decline? → Behavior C

Measuring expectation on policy rate





Survey-based

Market-based

Charts from **Christensen and Swan (2014)**, *Assessing Expectations of Monetary Policy*

Alternative data source for measuring expectation: newspaper articles

Bank Indonesia is expected to hold its reference rate (BI Rate) at 5,75%. The central bank considers the possibility of subsidized gasoline price hike and its impact on inflation.

→ expectation of rate stay

As monetary authority, BI is believed to not be careless in setting the interest rate. **Even if there is gasoline price hike this year, BI will probably only raise reference rate by 50 basis points at most.** "Right now, BI is more pro-growth and real sector," said Chief Economist of Danareksa Research Institute Purbaya Yudhi Sadewa to Investor Daily in Jakarta, Tuesday.

→ expectation of rate hike

A same note is added by Director of Institute for Development of Economics and Finance (Indef) Enny Sri Hartati. Observing the trend of global oil price, Enny is certain that subsidized gasoline prices will not be increased in the near future. **"So, in the next Board of Governors Meeting, Thursday, BI will not change its policy rate.** Apart from its function, BI is also responsible to stabilize business and banking conditions", she remarked.

→ expectation of rate stay

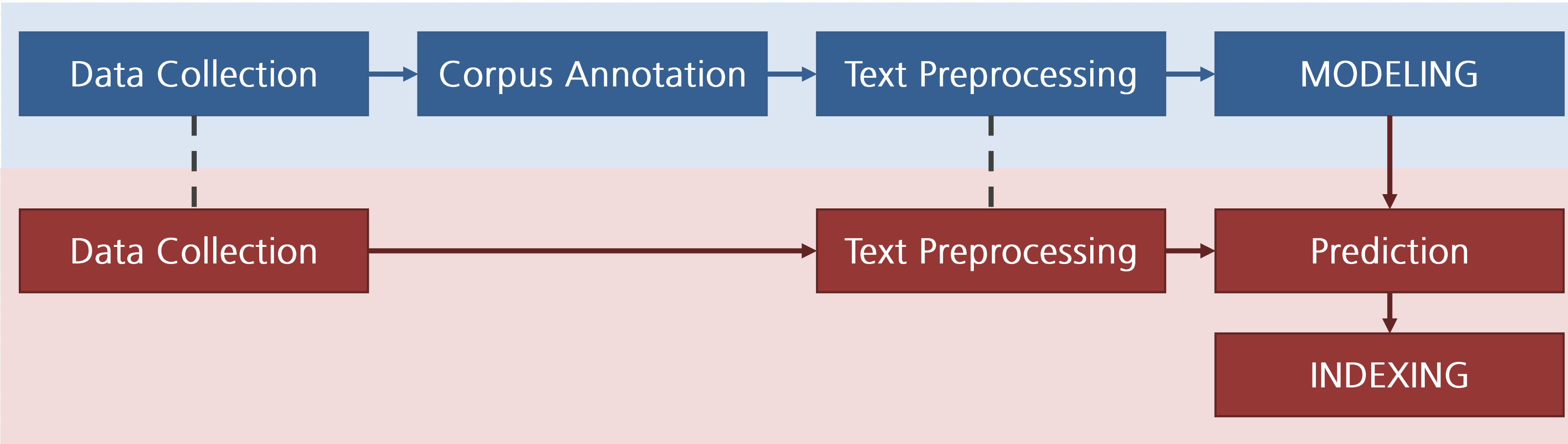
Can we develop a machine learning algorithm for automatically identifying and classifying public expectation on our policy rate from newspaper articles?

Pros-Cons of measuring expectation from news

- ✓ Available for public access
- ✓ Published in real-time
- ✓ Covers wide source of opinion:
  - ❑ Banks, financial institutions
  - ❑ Analysts & economists
  - ❑ Real sector, industries
  - ❑ Academics
  - ❑ Government
- ✓ It's what people read

- o May not reflect "true" expectation, no "money-in-the-game"
- o Likely less accurate prediction than actual survey on professionals
- o No control over "respondents", respondents can change every period

Methodology



MODELING

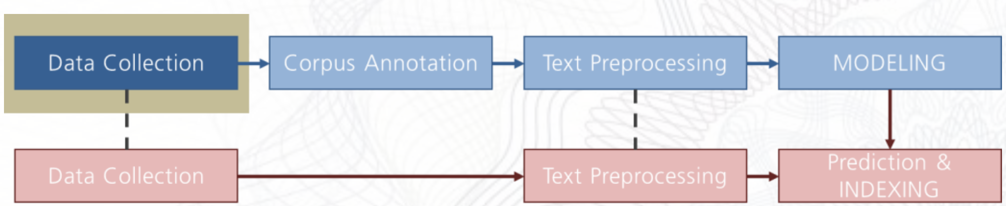
The Modeling step is done <u>once</u>
to train the best machine learning algorithm for the task

INDEXING

The Indexing step is done <u>every time</u>
Bank Indonesia will hold policy rate meeting (regularly once a month)

Modeling: Data collection

- <u>Dataset: CyberLibrary</u>
  - Collection of economic and financial news articles
  - Curated for internal use by Bank Indonesia
- <u>Collection period</u>: Jan 2006 – Feb 2018
  (BI policy rate was established in July 2005)
- <u>Filtering</u>
  - Kept only articles published **within 2 weeks before** scheduled policy rate meeting
  - Kept only <u>sentences</u> that **mention BI policy rate** e.g. "Bank Indonesia reference rate", "BI 7DRR"



Data collection and filtering

Bank Indonesia is expected to keep **BI Rate** at 5,75% in its policy rate meeting this Thursday.

All articles in *CyberLibrary*

Articles published 14 to 1 day before policy rate meeting

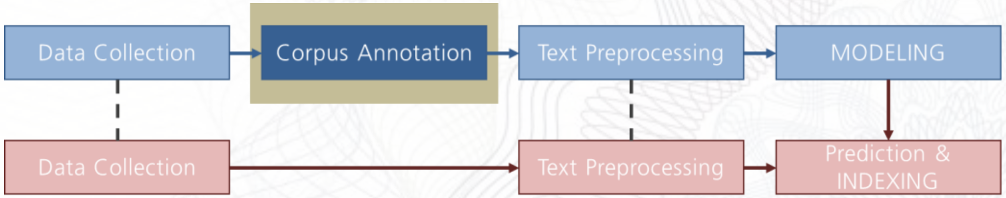
Sentences mentioning BI policy rate

Modeling: Corpus annotation

- Each filtered sentence is annotated as one of **4 classes**:
  - **0**: Non-expectation
  - **1**: Expectation of rate hike
  - **2**: Expectation of rate cut
  - **3**: Expectation of rate stay

- Each sentence is annotated by 2 persons
  - In case of differing class → 3rd annotator

- Total annotated sentences: ~**4,400**
  - 66% non-expectation
  - 8% expectation of rate hike
  - 15% expectation of rate cut
  - 11% expectation of rate stay



Example annotation (in Excel spreadsheet)

date	sent	label	label
2017-08-08	Kalau memang BI ingin melakukan stimulus moneter, hal pertama adalah menurunkan 7-days repo rate dari 4,75% menjadi 4%.	2	turun
2017-08-09	Jika demikian, dia menilai kebijakan pelonggarannya tidak harus melalui pemangkasan suku bunga acuan atau BI 7-Day Repo Rate.	0	nonekspektasi
2017-08-09	Bank Indonesia juga diperkirakan mempertahankan suku bunga 7 days repo rate di level 4.75% sampai akhir 2017.	3	tetap
2017-08-09	BI diperkirakan masih mempertahankan kebijakan suku bunga 7 days repo rate di level 4,75% sampai pertengahan 2018.	3	tetap
2017-08-09	Tingginya minat investor dalam lelang SUN-, menurut Anil, terjadi karena investor mengantisipasi penurunan suku bunga acuan Bank Indonesia (BI).	2	turun
2017-08-11	Di sisi lain, untuk suku bunga acuan Bank Indonesia sejak akhir 2015 sampai saat ini telah menurunkan suku bunga acuan sebesar 150 bps menjadi 4,75%.	0	nonekspektasi
2017-08-11	Wakil Presiden Direktur Bank Central Asia Eugene Keith Galbraith mengatakan, perseroan menilai peluang penurunan bunga kredit masih terjadi karena dari sisi suku bunga acuan Bank Indonesia (BI) masih belum ada perubahan, terutama arah kenaikan.	0	nonekspektasi
2017-08-14	Doddy melanjutkan, perihal Suku Bunga simpanan ini erat kaitannya dengan Suku Bunga acuan bank Indonesia (BI 7 days reverse repo rate).	0	nonekspektasi
2017-08-14	Menurut dia, kebijakan BI yang mempertahankan BI 7 days reverse repo rate di level 4,75% masih sejalan dengan perbaikan kondisi perekonomian global dan proses pemulihan ekonomi domestik yang terus berlanjut.	0	nonekspektasi
2017-08-14	"Sepanjang 2017, BI 7 days reverse repo rate diperkirakan akan tetap flat," ujar dia.	3	tetap
2017-08-14	Chief Economist bank Mandiri Anton Gunawan mengungkapkan, secara industri perbankan, transmisi penurunan Suku Bunga acuan BI pada tahun lalu terhadap Suku Bunga kredit masih belum selesai.	0	nonekspektasi
2017-08-15	Agus di Jakarta, Jumat (4/8), men-gatakan hal tersebut, setelah sembilanbulanberturut-tu-rut BI menahan pelonggaran Suku Bunga acuan "7-Day Reverse Repo Rate" di level 4,75persen.	0	nonekspektasi

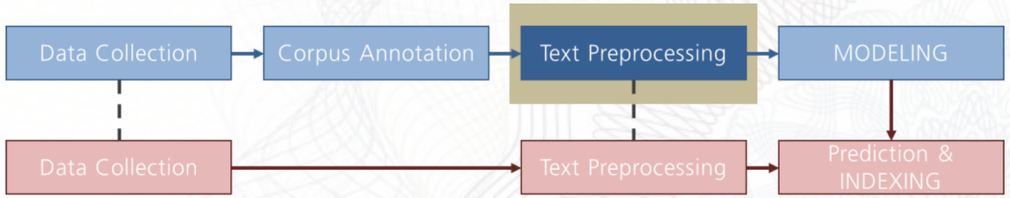
Modeling: Text preprocessing

Each sentence is preprocessed to make learning simpler for the algorithm

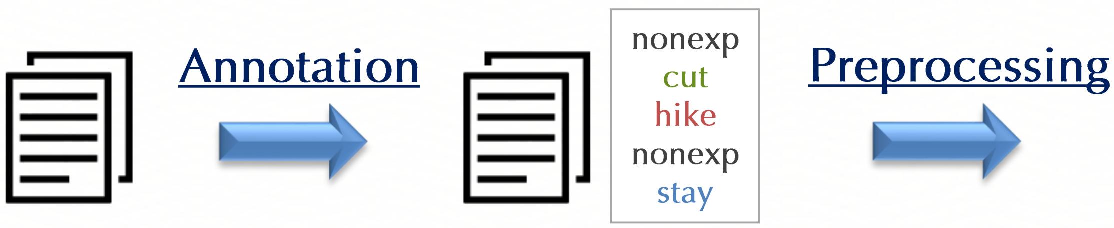
- ❑ Lowercasing
- ❑ Stop words removal
- ❑ Merging synonyms and word forms e.g.
  - ❑ predicting, predicted, expecting, expects → predict
  - ❑ BI Rate, BI policy rate, BI reference rate, BI 7DRR → BIRate

Sentence (in text format) is then transformed into feature vector (numeric)

- ❑ n-gram occurrences (total >5.000 1-4 grams)
- ❑ Sentence length
- ❑ Any numbers and percentages mentioned
- ❑ Sum of word embeddings (from https://fasttext.cc/)



Text preprocessing



	var-1	var-2	…	var-d	label
sent-1	v_{11}	v_{12}	…	v_{1d}	nonexp
sent-2	v_{21}	v_{22}	…	v_{2d}	cut
sent-3	v_{31}	v_{32}	…	v_{3d}	hike
…	…	…	…	…	nonexp
sent-n	v_{n1}	v_{n2}	…	v_{nd}	stay

Filtered sentences

Text + label for each sentence

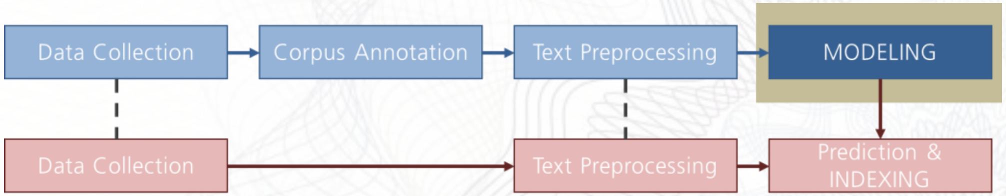
Vector + label for each sentence

Modeling: Training machine learning algorithm

Sentences as feature matrix + annotated labels is ready for machine learning training

Experiment setup:

- ❑ 80% sentences for train/validation set, 20% for test set
- ❑ Models evaluated:
  - ❑ Logistic regression
  - ❑ Naïve bayes
  - ❑ Decision tree
  - ❑ Random forest
  - ❑ XGBoost
- ❑ 25 hyper-parameter settings for each model
  - ❑ 10-fold random split for tuning
  - ❑ Each split has 20% test sentences (taken from the 80% train/validation set)



Model evaluation metric

		Predicted by model	
Actual label		Expectation (label 1, 2, 3)	Non-expectation (label 0)
	Expectation (label 1, 2, 3)	TP	FN (*Type II error*)
	Non-expectation (label 0)	FP (*Type I error*)	TN

Metric	Formula	Interpretation
Accuracy	(TP + TN) / (TP + TN + FP + FN)	% correct predictions
Precision	TP / (TP + FP) = 1 – *type I error rate*	% correct labels from all predicted as expectation
Recall	TP / (TP + FN)	% predicted expectations from all labeled as expectation
F1	2 * Precision * Recall / (Precision + Recall)	Harmonic mean of precision & recall

Evaluation results on test data

Each model with its best hyper-parameter is trained on the whole train/validation set, then evaluated on test set to measure performance

Model	Accuracy	Precision	Recall	F1
Logistic regression	83,4	71,2	83,2	76,7
Naïve bayes	80,6	64,5	83,2	72,7
Decision tree	73,0	53,4	65,7	58,9
Random forest	78,0	63,3	72,6	67,6
XGBoost	84,1	75,6	75,9	75,7

Indexing

At this point we have sufficiently accurate machine learning model for classifying expectation for (the next) policy rate meeting, given unlabeled news sentences.

In Indexing step we aim to aggregate the classifications into a single number for each policy rate meeting.

New, unlabeled sentences	Label
The market is anticipating the result of BI Board of Governors Meeting, which is predicted to keep reference rate at 4.25%.	?
"Rupiah stability in 2017 in the midst of FFR hikes, as well as relatively controlled inflation rate, may serve as the basis for BI to cut its policy rate," he said.	?
There is possibility that BI will raise its interest rate on Thursday's Board meeting.	?
...	...

Trained ML model

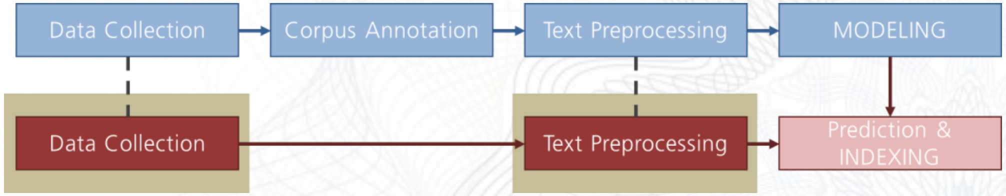
Label, predicted by model
3 (rate stay)
2 (rate cut)
1 (rate hike)
...

Indexing: Data collection and text preprocessing

Data collection and text preprocessing in Indexing is <u>exactly the same</u> as in Modeling step,

<u>except</u> the articles considered are grouped into corresponding policy rate meeting

e.g. articles published between 5 to 19 July 2018 will be grouped for policy rate meeting on 19 July



Indexing: Prediction and index calculation

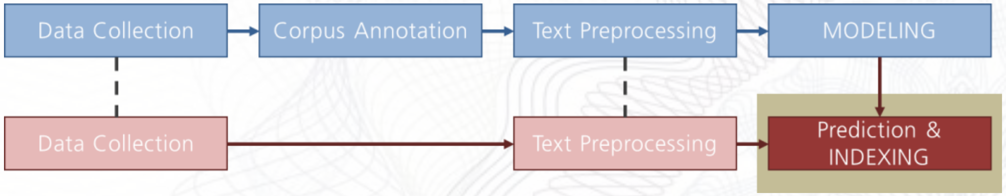
The trained model is used to classify all filtered (and unlabeled) news sentences.

Each sentence is scored based on model's classification:

- 1, if classified as expectation of rate hike
- 0, if classified as expectation of rate stay
- -1, if classified as expectation of rate cut
- Discarded, if classified as **non-expectation**

Each article is scored as the average of its sentences' score.

Expectation index for the next policy rate meeting is the average of article scores.



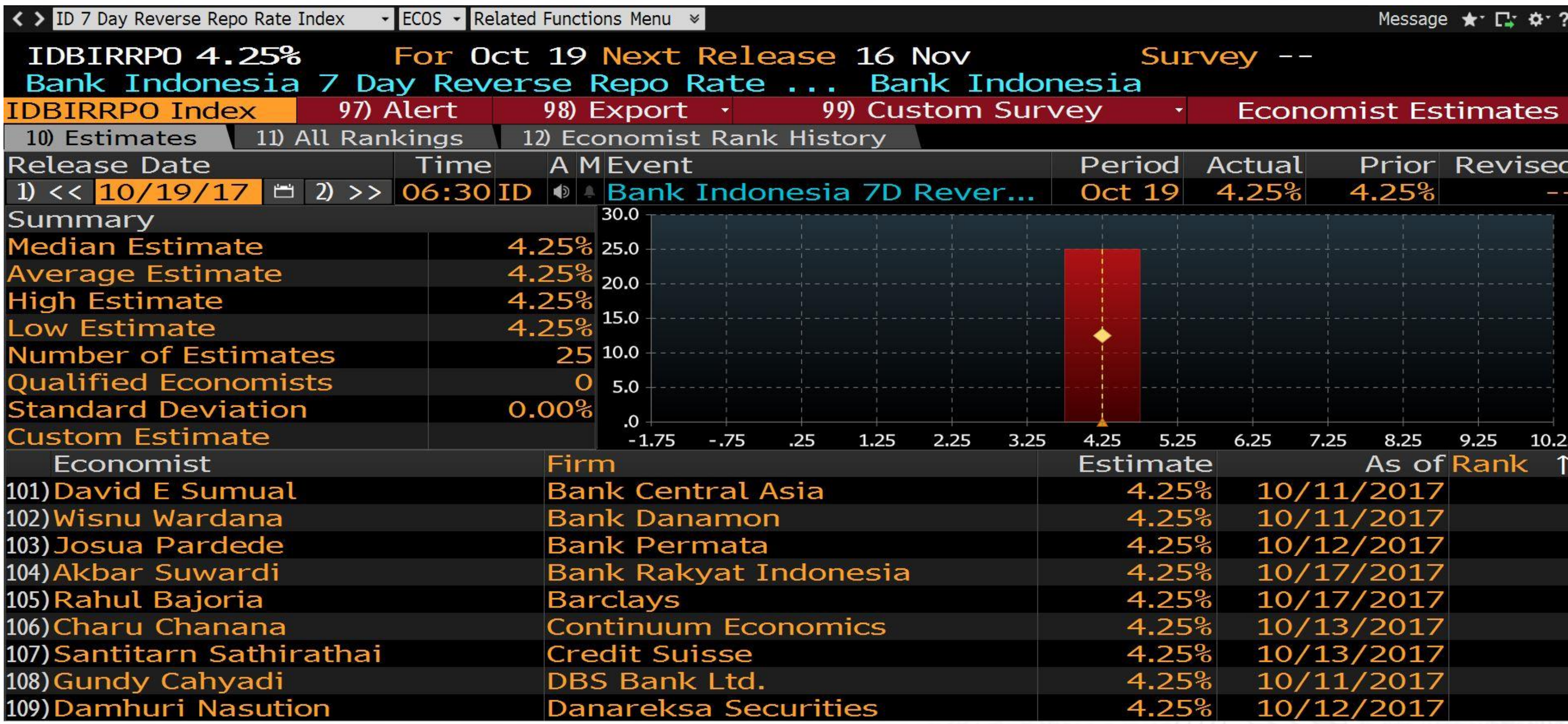
Index properties

The resulting expectation index has following properties:

- Range from -1 to 1
- With more expectation of rate hike, index approaches **1**
  With more expectation of rate stay, index approaches **0**
  With more expectation of rate cut, index approaches **-1**
- **Positive index**: more expectation of rate hike compared to rate cut
  Negative index: more expectation of rate cut compared to rate hike
- If index at t_1 > index at t_2, then there is greater share of rate hike expectation in t_1 compared to t_2

Benchmark: Bloomberg Economist Estimates

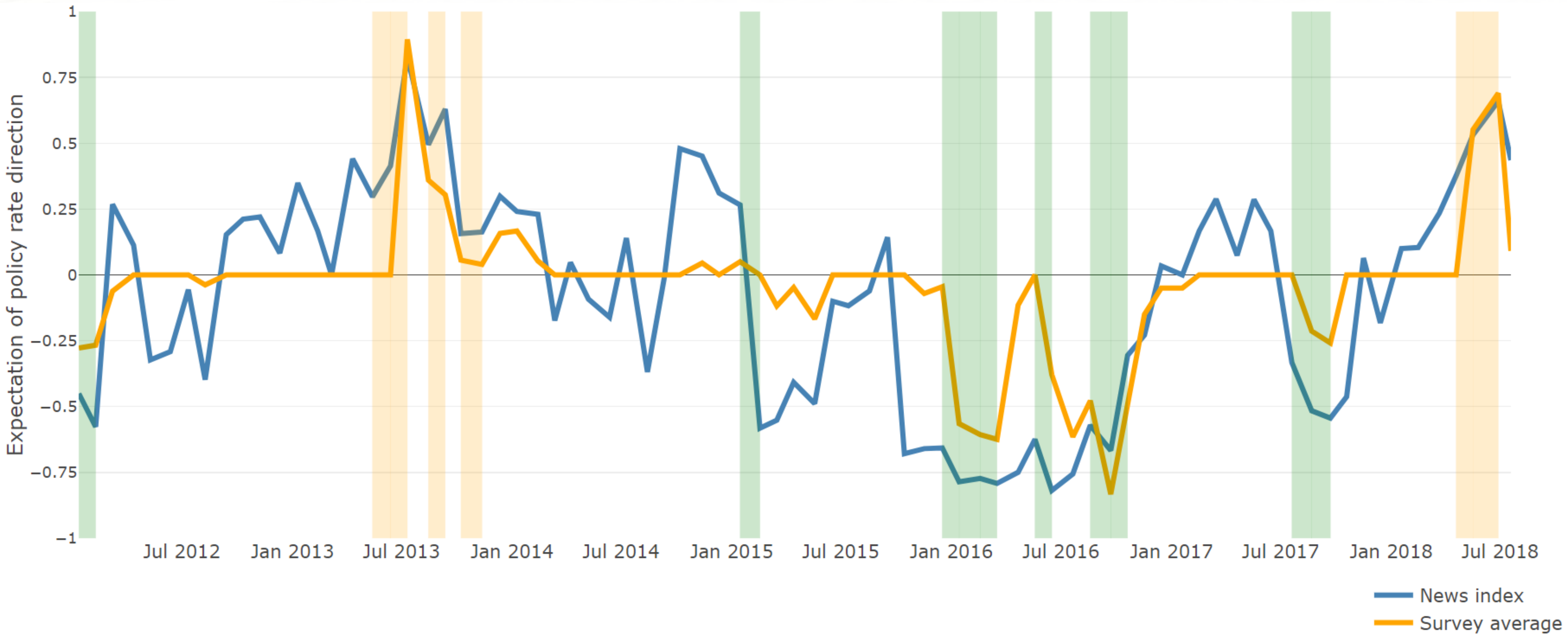
Bloomberg surveys a number of economists for their prediction of BI policy rate that will be set in the next meeting.



Summary	
Median Estimate	4.25%
Average Estimate	4.25%
High Estimate	4.25%
Low Estimate	4.25%
Number of Estimates	25
Qualified Economists	0
Standard Deviation	0.00%
Custom Estimate	

	Economist	Firm	Estimate	As of	Rank
101)	David E Sumual	Bank Central Asia	4.25%	10/11/2017	
102)	Wisnu Wardana	Bank Danamon	4.25%	10/11/2017	
103)	Josua Pardede	Bank Permata	4.25%	10/12/2017	
104)	Akbar Suwardi	Bank Rakyat Indonesia	4.25%	10/17/2017	
105)	Rahul Bajoria	Barclays	4.25%	10/17/2017	
106)	Charu Chanana	Continuum Economics	4.25%	10/13/2017	
107)	Santitarn Sathirathai	Credit Suisse	4.25%	10/13/2017	
108)	Gundy Cahyadi	DBS Bank Ltd.	4.25%	10/11/2017	
109)	Damhuri Nasution	Danareksa Securities	4.25%	10/12/2017	

Results: 2012 – July 2018

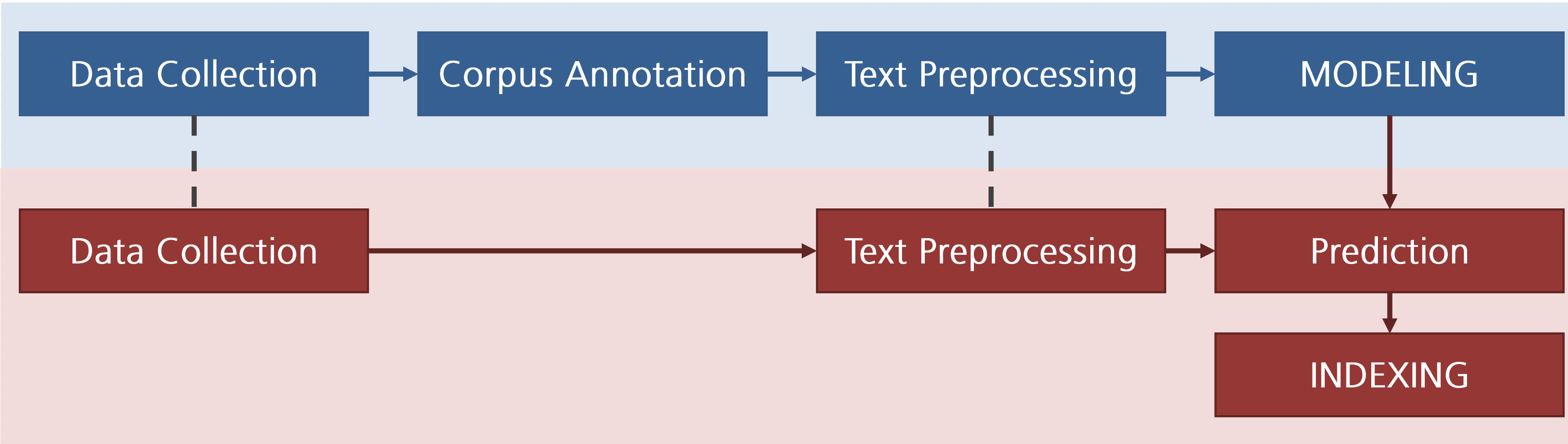


Correlation with Bloomberg survey: 73%

Actual rate hike

Actual rate cut

Methodology (review)



MODELING

The Modeling step is done <u>once</u>
to train the best machine learning algorithm for the task

INDEXING

The Indexing step is done <u>every time</u>
Bank Indonesia will hold policy rate meeting (regularly once a month)

Conclusion and future improvements

- ✓ Have shown machine learning use case for measuring public expectation on our policy rate, exploiting news articles as data source
- ✓ Overall the machine learning model has quite good accuracy, and also the resulting index tracks professional estimates quite well

Future improvements

- ❑ Identifying **opinion holders**: who has what expectation
- ❑ Separating **predictions vs desire**
  - ❑ "what people *think* the rate will be" vs "what they *want* the rate to be"
- ❑ More news **source**, including English ones
- ❑ More **advanced algorithm** for classification
  - ❑ Deep learning? But probably requires (far) more annotation

THANK YOU
Q&A