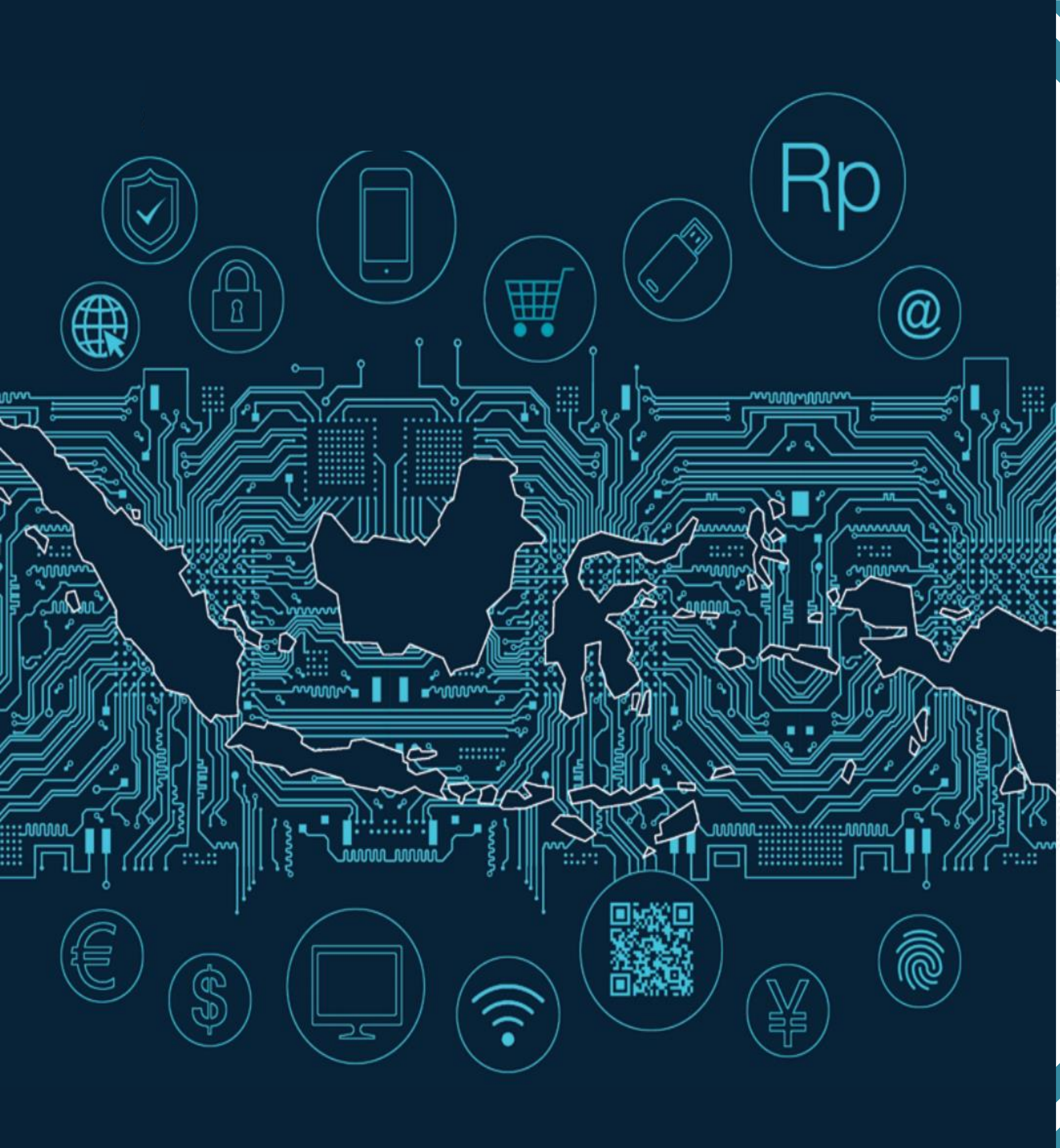


BANK INDONESIA
BANK SENTRAL REPUBLIK INDONESIA

Measuring Payment System Policy Credibility Using Machine Learning

Okiriza Wibisono, Muhammad Abdul Jabbar,
Alvin Andhika Zulen

11th IFC Biennial Conference
25-26 August 2022

Background and Goals

Payment system plays a crucial role *in ensuring the smooth functioning of economic and financial activities. More so in this digital age.*

Public's perception on the payment system *ecosystem may impact their adoption of new developments in payments.*

Previous approach to measure public perception on our payment system policy credibility: semiannual survey to stakeholders (e.g. economists, academics, government, general public).

This research: Utilizing ***Big Data Analytics*** *– text mining to gather public perception regarding payment system ecosystem and its related policies.*

Methodology largely based on our previous use case on measuring monetary policy credibility (Wibisono, 2022).

Data & Overview of Methodology

News articles

Source: Cyber Library (internal repository of curated economic and financial news)

~30 domestic news (in Bahasa Indonesia)
~850 articles daily

Whole corpus: since Jan 1999
Training data: Jan 2013 – Sep 2021

Example keywords for filtering the news:

- payment system
- BI-FAST
- BI-RTGS
- SKNBI
- internet banking
- mobile banking
- e-money
- payment service providers
- card payments
- credit cards
- debit cards
- transfer fee
- EDC
- Open API
- QRIS
- consumer protection
- payment service licensing
- supervisory technology
- fraud supervision
- cyber security

1 Filtering
2 Annotation
3 Pre-processing

4 Modelling for Infrastructre Aspect | Modelling for Industry Conduct Aspect | Modelling for Regulation Aspect | Modelling for Entry Policy Aspect | Modelling for Supervision Aspect

5 Model for Infrastructure Aspect | Model for Industry Conduct Aspect | Model for Regulation Aspect | Model for Entry Policy Aspect | Model for Supervision Aspect

Positive / Negative (for each model)

6 Index Calculation

1. Infrastructure
2. Industry Conduct
3. Regulation
4. Entry Policy
5. Supervision

Overall index

5 Credibility Aspects

Payment system infrastructure: policy, developments, and conduct (e.g. reliability, safety, efficiency) of payment system infrastructures, both those that are operated by BI or by the industry

Payment system conduct: conduct of payments services by the industry, how BI's payment system policies are implemented by the industry

Payment system regulation: whether BI's payment system policies are well-formulated and effective in achieving their intended objectives.

Payment system entry policy: effectiveness and efficiency of payment system entry and licensing activities

Payment system supervision: BI's supervision of the payments industry, e.g. related to payment service and payment infrastructure providers, consumer protection

Methodology – Index Construction

1. Annotation

A sample of filtered sentences are annotated as training data for ML classification models.

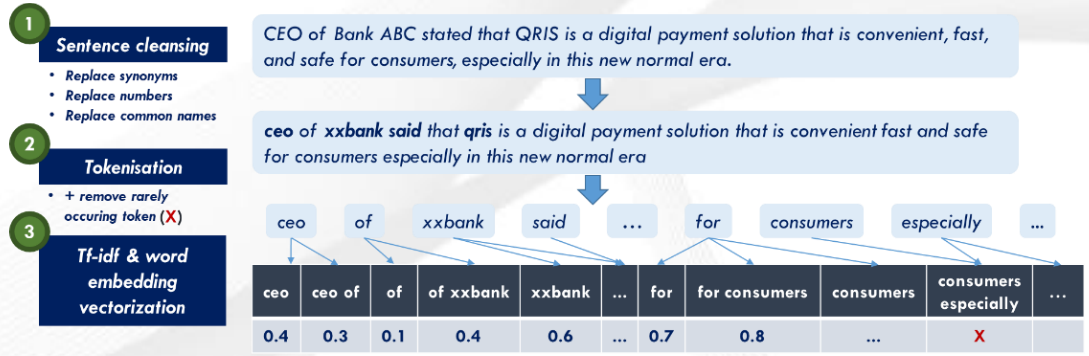
- ➢ Annotated by authors and domain experts within BI.
- ➢ Guidelines incl. examples
- ➢ 2-3 annotators per sentence
- ➢ Most sentence turns out irrelevant for the aspect

Aspect	Positive	Negative	Irrelevant
Infrastructure	561 (6.6%)	24 (0.3%)	7,971 (93.1%)
Industry conduct	1,745 (20.4%)	200 (2.3%)	6,611 (77.3%)
Regulation	1,423 (16.6%)	80 (1.0%)	7,053 (82.4%)
Entry policy	105 (1.2%)	27 (0.3%)	8,424 (98.5%)
Supervision	179 (2.1%)	42 (0.5%)	8,335 (97.4%)
TOTAL	4,013 (9.4%)	373 (0.09%)	41,789 (97.7%)

2. Data preprocessing

Each sentence is transformed from text into tabular-numeric format for training ML models.

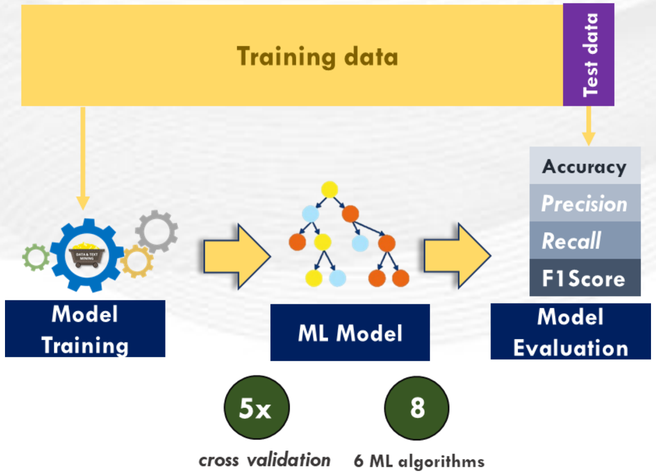
- ➢ Sentence cleansing
- ➢ Tokenization
- ➢ Remove sparse terms
- ➢ N-gram vectorization
- ➢ Word embedding



3. Model training

ML model is trained for classi-fying sentences into pos/neg labels, for each aspect.

- ➢ **5-fold CV**, 80-20 train-validation split
- ➢ **SMOTE** to handle imbalanced labels
- ➢ **2-step classification**: relevant vs irrelevant, positive vs negative
- ➢ Best average **macro F1**: 60.2%
- ➢ Best algorithm: **logistic regression**



4. Index calculation

- The ML models are applied to all sentences, to construct monthly indexes.
- The overall index is a weighted average of the 5 component indexes based on number of sentences.
- Any component with 3 or less sentences in a quarter is excluded.

$$index_{aspect\ k,t} = \frac{\#positive_{k,t} - \#negative_{k,t}}{\#positive_{k,t} + \#negative_{k,t}}$$

$$index_t = \sum_{aspect} \mathbb{1}_{aspect,t} \times (\#positive_{aspect,t} + \#negative_{aspect,t}) \times index_{aspect,t}$$

Example annotated sentences (only relevant sentences)

No	Sentence	Credibility Aspect	Label
1	*Bank Indonesia strengthens payment system infrastructures and promotes noncash payments.*	Infrastructure	*Positive*
2	*Bank Indonesia detected that on Tuesday there have been network disturbance in their BI-RTGS, BI National Clearing System (SKNBI) and BI Scripless Securities Settlement System (BI-SSSS).*	Infrastructure	*Negative*
3	*Since the standardization of electronic payments using QRIS, about 4 million MSMEs have adopted QRIS [for accepting payments].*	Industry conduct	*Positive*
4	*QRIS hampers payment system efficiency since there is requirement to become interoperable and interconnected.*	Industry conduct	*Negative*
5	*Bank Indonesia's payment system policy is aimed at accelerating payment system digitalization to further integrate the digital economy and support national economic recovery.*	Regulation	*Positive*
6	*Regulation on fintech companies by Bank Indonesia and by Financial Services Authority have not been able to provide legal certainty for consumers.*	Regulation	*Negative*
7	*We [Bank Indonesia] are reviewing [payment system] licensing system so that it can be more efficient.*	Entry policy	*Positive*
8	*There are complaints from the industry that the fintech licensing system [by Bank Indonesia] is complex and costly in terms of time required.*	Entry policy	*Negative*
9	*Bank Indonesia closely supervises payment service providers in order to ensure consumer protection in the payment system and digital economy.*	Supervision	*Positive*
10	*Bank Indonesia must recover consumer's loss in this incident of payment system failure.*	Supervision	*Negative*

Result – ML model

Credibility aspect	Best model relevance	Best model sentiment	A: Relevance classification F1	B: Sentiment classification F1	End-to-end F1 (A*B)
Infrastructure	logistic regression	logistic regression	75%	86%	64.5%
Industry conduct	logistic regression	XGBoost	72%	78%	56.2%
Regulation	logistic regression	logistic regression	73%	74%	54.0%
Entry policy	logistic regression	random forest	75%	88%	66.0%
Supervision	XGBoost	decision tree	68%	89%	60.5%
Overall (average)	-	-	**72.6%**	**83.0%**	**60.2%**

Some observation about the results:

- **Logistic regression** algorithm is the most accurate in majority of cases.
- Classifying **sentiment** (positive vs. negative) is relatively **easier** than classifying whether the sentence contains sentiment in the first place (83.0% vs. 72.6% averaged macro-F1).
- Payment system **industry conduct and regulation aspects** have lowest end-to-end F1 (due to lowest sentiment F1). This is somewhat unexpected, since these aspects have the largest share of negative sentences (less imbalance).
- **End-to-end F1** (60.2% average) is **acceptable**, but may warrant further improvement to ensure more robustness of the resulting indexes.

Result – Indexes

Some observation about the indexes:

- **Most** of the indexes are always **positive**, which means that there are more positive sentences about payment system than negative sentences in the news.
- The overall index is also **increasing** over time (86.1% in 2017 to 96.7% in 2021; annual numbers not shown in figure).
- **Only supervision** index ever reaches **negative** value. However, we note that supervision index has **high volatility**, and to less extent, entry policy index as well.
- The indexes have average pairwise **correlation** of 13.4% (30.7% if excluding the volatile supervision index), although industry conduct and regulation has 75% correlation.
- **Example** recent trends captured by the models: positive news about our newly implemented payment system (BI-FAST), positive news about wider use of QRIS.

Conclusion

1. Developed a machine learning methodology for measuring public's perception of payment system policy credibility by utilizing news data.
2. The resulting index shows positive trend: according to the models, news about payment systems in Indonesia in recent years is more positive. But supervision index is highly volatile due to small number of relevant sentences.
3. The models seem to be able to capture relevant developments, such as BI-FAST implementation and wider adoption of QRIS.

Future Works

1. Reidentification/redefinition of the credibility aspects.
2. Model accuracy improvement e.g. by annotating more data, or using specific keywords for each aspect.
3. Econometric analysis (econometric effect of the indexes on macro indicators).
4. Developing method for automatic updating of keywords.