The Digital Era Stages on Taxation: An Experimental Study of Text Mining and Pattern Recognition for Controlling Tax on Business Online Transaction

ASL Lindawati[1], Rafika Setyawati[2] and Junita[3]

[1]University of Bina Nusantara

ABSTRACT

The main idea of this research is to improve tax control for online businessmen who have not yet reported tax and or not yet have NPWP (Nomor Pokok Wajib Pajak), so it is necessary to clarify and re-reviewed the online business players. The purpose of this research is to design the right system or model for the detection of online business, or business transactions conducted in social media for the benefit of increasing tax revenue. Tracking and imposition tax must be worn by online business in social media, except the business into the marketplace such as, 'Tokopedia', 'Buka Lapak', 'Lazada'. Hitherto this research is conducted there is no single research that focus on this. Based on the Directorate General of Taxation this is an important issue and should be resolved soon, due to the increasingly widespread business through social media in Indonesia last 5 years. Therefore this research needs to be done to anticipate and solve the problem of research through system design model. This detection model is carried out in two stages, the first phase focused on the classification of social media users through the text mining process. The second stage is perfecting the detection of online business through the process model of Pattern Recognition. Therefore, the Indonesia's tax problem increase and complicated particularly in regards with the absence models for control system to increase income tax from online business tax payers, especially SMEs who the major usage social media as the way for a business transactions. This research uses experimental method. Data collection was taken from the social media users through cluster sampling technique, in this study the authors use social media Facebook as experimental data. The analysis associated with text mining is to create a variety of key words that aim to detect business transactions on Facebook. The second phase of analysis implements text mining results with data mining from the bank through design patterns based on time and nominal value. The results of this design model is expected to improve control system and tax revenue.

Keywords: Tax, Business Online, Social Media, Cluster Sampling, Pattern Recognition.

1. INTRODUCTION

The use of internet technology as a means of information and communication exchange in today is already very common. Millions of people in the world today use the internet. By using the internet, all the information we want is readily available. In addition to providing various kinds of information, the internet also allows us to communicate without obstructed distance or time. Seeing the development of the Internet in the field of information and communication, the Internet also plays a role in the economic development. The large number of information media available such as websites and social media, began to be used for business activities. This activity is called e-commerce. E-commerce is a new concept that is usually described as the process of buying and selling goods or services on the World Wide Web Internet (Suyanto, M. 2003: 11).

With this concept, came the term online business. Online business is a promotional activity or offering products in the form of goods or services using internet media ranging from bargaining process to the transaction process without having to meet face to face with the customer. According to DigitalMarketer.id Online business is considered profitable because it does not require much capital. Only capitalize computers or laptops or other gadgets with the internet. Unlike conventional businesses, an online business does not require the seller to have a place of business. In addition to buyers, online business is a practical way of shopping. They do not have to take the time to go shopping.

The birth of online business is utilized by all business size, either from small, medium, and large. Basically, for UMKM, it called the Micro Small Medium Enterprises (SMEs), the online business is very helpful in developing its business to gain profit as much as possible with minimal capital. Therefore, many UMKM in Indonesia are starting to do business online that use social media as the market place for doing business transaction According to Waizly Darwin, Head of Small and Medium Business (SMB) Facebook Indonesia, in 2015 there are 57 million SMEs engaged in doing business in Indonesia and about 15 million SMEs have `go online`. However, along with the number of SMEs online business today is actually causing some problems in the application of taxes. According to Prof. William Fox and Donald Bruce of the University of Tennessee, e-commerce is a replacement and extension of remote sales whose taxes never get collected. Whereas already regulated in PP 46 of 2013 which regulates the taxation for UMKM (SMEs) that is an individual taxpayer or corporate taxpayer does not include a permanent establishment; and receives income from business, excluding income from services in respect of free employment, with gross revenue not exceeding Rp 4,800,000,000.00 (four billion and eight hundred million rupiah) in 1 (one) tax year. With the tariff of the final Income Tax is 1% (one percent). In addition, the provision of online business tax is the same as conventional business. It is stated in the Circular Letter of the Director General of Taxation No. SE/62/PJ/2013 on the Affirmation of the Taxation Provisions on e-Commerce Transaction confirming there is no difference in tax treatment between e-commerce transactions and other trade and/or service transactions.

According to Law no. 7 of 1983 concerning Income Tax subject to tax shall be income, ie any additional economic capability received or obtained by a Taxpayer, whether originating from Indonesia or from outside Indonesia, which may be used for consumption or to increase the property of the Taxpayer concerned, by name and in any form. Therefore, when e-commerce transactions occur and generate profits, the subject tax will be subject to income tax related to the transfer of third parties.

Another problem is that e-commerce transactions happen very quickly and in a short period of time, making anyone transacting hard to trace. Transactions that typically use a transfer system, where the transfer evidence is in the form of electronic evidence is increasingly difficult to detect (Vidyana, Adhysty, 2014: 2).

This is what makes the online business can avoid taxes. Given these limitations and the absence of the right model, the potential tax revenue for e-commerce transactions can be lost. Actually there has been an effort by the government to continue to increase tax revenue for online business, one of them is by issuing Government Regulation in Lieu of Law (Perppu) No 1/2017 on Access to Financial Information for Tax Interest. The government considers that tax is the main source of state revenue. Thus, extensive access is required for tax authorities to receive and obtain financial information for tax purposes. However this has not been effective in preventing obedient taxpayers. Because the business may use other accounts to conduct transactions so that tax officials cannot supervise all transactions.

The purpose of this research is to provide new innovations in the form of a useful model to facilitate in knowing who the online business of SMEs in Indonesia who are not obedient in paying taxes. With this model provides many benefits such as the implementation of the balance of the tax burden between conventional entrepreneurs and e-commerce, facilitate the Directorate General (DGT) in conducting online tax tracking business, and maximize the state revenue from taxes from e-commerce.

This research will greatly assist the government in overseeing taxation to be more equitable for all MSMEs in Indonesia. Unfortunately, the only limitation of the current government is the absence of a system that can monitor the obligation to pay taxes for SMEs online business. Designing a model from identification to producing a prototype model for taxation for SMEs online business is the focus of this research. The structure of this writing is divided into several parts. The first part is an explanation of the background of the problem that explains the weakness of the taxation system for online business. The second part describe the briefly of literatures that relating to theories of e-commerce taxation, the concept of data mining systems in business processes, and intelligence systems as the basis for the manufacture of prototype models. Followed by the third section will describes the research methodology with a research design using an experimental approach. How the process of data obtained up to its analysis and continued with explanation of research design to implement prototype model. The fourth part is making the prototype of e-commerce transaction tracking which includes two parts: SME online business grouping and online business tracking through SME accounts. Further conclusions and suggestions are put forward in the closing section.

Study References

Electronic commerce (electronic commerce or e-commerce) is a process that consumers make in buying and selling various products electronically from company to company using a computer as an intermediary of business transactions (Loudon 1998). As technological developments become more sophisticated, many people seek to increase their revenues by trading electronically. Currently, the development of information technology has entered into the digital era. E-commerce or electronic commerce has expanded. Today, trading is no longer limited to the ability to use electronic media but has been able to use digital media to conduct business transactions, trade through the use of social media as a form of information technology development in the digital era has been increasingly massive and rampant. In Indonesia, this condition has triggered many business actors conducting business transactions online as well as MSMEs. Harmonizing with the development of online business, so many other things or related factors that need to be developed, including increasing tax revenue and developing control in the field of tax business transactions. Therefore, it is necessary to develop an innovative new model to anticipate the existence of online business model.

Models of E-commerce Transactions by Taxation

Based on the Circular Letter of the Director General of Taxation Number SE-62/PJ/2013 concerning Affirmation of Taxation Provisions on E-Commerce Transactions, the Directorate General of Taxation has set 4 (four) models of e-commerce transactions as follows:

Online Marketplace. Place of business activity in the form of an internet store as an Online Marketplace Merchant to sell goods and/or services. Marketplace generally provides an additional layer of security for every transaction that occurs, such as an escrow payment system or more commonly known as a shared account. So every time a transaction takes place within the marketplace system, the marketplace will be the third party receiving the payment and keep it until the product has been shipped by the seller and received by the buyer. Example: Tokopedia, Lazada.com

Classified Ads. Activities provide the space and time to advertise the goods and/or services performed by the advertisers through the sites provided by Classified Ads Classifieds. Example: Kaskus, OLX.co.id.

Daily Deals. Business activities are almost the same as the Online Marketplace but the payment instrument used in the form of vouchers. Example: Groupon.com.

Online Retail. The activity of selling goods and/or services performed directly by the Online Retail provider to the buyer on the Online Retail website. Example: A seller doing his business in social media.

Tax provisions applicable in online business is not different from conventional stores in general, only the media used in this case is the internet. Based on Income Tax Article 4 paragraph (2) or Final PPM UKM/entrepreneur/online business, online business is one of the tax subject classification. The tax rate charged under Government Regulation of the Republic of Indonesia Number 46 of 2013 is 1 percent deducted from total sales turnover (gross turnover) per month and is paid on the 10th of every month. The object of income tax-deductible SMEs/entrepreneurs/online business is business with total sales turnover (gross circulation) not more than Rp 4.8 Billion in a year. If in certain turnover there is a transaction subject to final tax, then the transaction is not included in earnings turnover with final tax. In accordance with Regulation SE-06/PJ/2015 it is stated that the deduction and collection of income on e-commerce transactions that provide specific affirmation regarding liabilities and collection of income tax on e-commerce transactions.

The growth of online business SMEs in Indonesia is increasingly increasing based on statistical results from the Central Bureau of Statistics (BPS). The perpetrators usually come from among students, college students, young scholars, housewives and others. Therefore, the author seeks to assist the government in the supervision of tax collection for online business by using information systems. According to Lani Sidharta (1995: 11) Information Systems are "Man-made systems that contain an integrated set of manual components and computerized components that aim to collect data, process data, and produce information for users." The information system we developed is Data Mining and Intelligence System. These two systems will help the authors in tracking the online business of UMKM. Explanation of these two systems using Business Process Modeling.

Use of Business Process Model and Notation (BPMN)

Globally Business Process Models and Notations (English: Business Process Models and Notations) are used for business process modeling, or like grammatical modeling options for analyzing and designing information conscious process systems.

The main purpose of BPMN's effort is to provide a notation that is easy for all business users to understand, from business analysts who create the initial concept of the process to the technical developer responsible for implementing the technology that will carry out the process, to the businessman who will manage and monitor the process . (Mendling, J and Matthias, W. 2012)

Business Process Diagram

Business Process Diagram is a process in modeling a process containing sub-processes, which can be graphically represented by another Business Process Diagram connected via a hyperlink to the process symbol. According to Lily (2012) A BPD consists of a set of graphic elements, and the basic categories of elements of BPD are:

(a) Flow Objects are divided into 3 such as; event, activity, and gateway. An event is an activity during a business process. Activity Gateway is used to control branching and merging Sequence Flow.

Figure 1: Flow Objects

(b) Connecting Object elements that connect the flow object. Connecting Object also has 3 types of elements: First Sequence flow is used to indicate the sequence of activities will be done. The Both Messages Flow is used to denote message flow between two entities. Association is used for data associations, information and artifacts with the flow of objects.
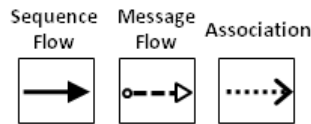
Figure 2: Connecting Object

(c). Swimlanes are represented by lines that separate and classify actors (actors interacting with the system).
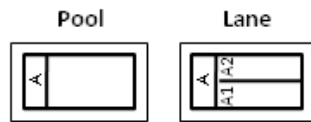
Figure 3: Swimlanes

(d). Artifacts are the elements used to provide additional information of a process. The current BPMN version has 3 types of Artifacts; the first type of data object is a mechanism to show how data is needed or produced by activity. The second type is a group can be used for documentation or analysis purposes. The third type of annotation mechanism for modelers provides additional text information to readers from the BPMN diagram.
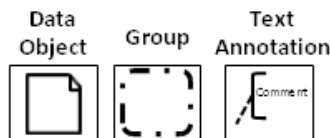
Figure 4: Artifacts

Tracking Using Data Mining Systems

Simply data mining is the mining or discovery of new information by searching for patterns or certain rules of large amounts of data (Davies 2004). Data mining is often referred to as knowledge discovery in database (KDD). KDD is an activity that encompasses collecting, using data, historically to find regularities, patterns, or relationships in large data sets (Santoso 2007).

Stages of Knowledge Discovery in Database process by Han (2006):

1. *Data Cleaning:* The process of removing multiple data, checking inconsistent data, correcting information errors. This will reduce the amount of data handled.

2. *Data Integration:* Added relevant data and merged data from multiple databases into one new database.

3. *Data Selection:* Selection of relevant data and analysis of operational data. The selected data is stored in a separate database.

4. *Data Transformation:* Transforms into specific formats to simplify the process of data mining.

5. *Data Mining:* Look for patterns to filter information on interesting shapes using methods, techniques and algorithms.

6. *Pattern Evaluation:* In this stage the results of data mining techniques are typical patterns as well as predictive models in the evaluation to assess whether existing hypotheses are indeed achieved.

7. *Knowledge Presentation:* Featuring patterns generated from data mining processes, these visualizations help communicate the results of data mining in a form that is easy to understand.

Moreover, Han (2006) stated that the function applied data mining as follows:

1. Association is the process for determining the associative rules of a combination of items at a time.

2. Sequence, almost like an association but the difference is applied in one period only.

3. Clustering, grouping data in groups of data and each cluster filled with similar data.

4. Classification, the process of discovery of models and functions to distinguish concepts or classes on each data.

5. Regression, the mapping process in a prediction value.

6. Forecasting, a process by which estimates prediction values are based on patterns in a data set.

7. Solution, the basic discovery process of the problem of business problems faced.

In this data mining the author designed the pattern to filter the online business of UMKM by using text mining. Text mining is one of the special areas of data mining. Text mining can be defined as a process of digging information in which a user interacts with a set of documents using tools analysis which are components in data mining, one of which is categorization (Feldman 2007).

The initial process of filtering, text mining begins with keywords that are general. The second phase of text mining works by filtering the amount of turnover of UMKM sales per day so that the detection into account bank account easier. Based on Law Number 20 Year 2008 regarding Micro Small and Medium Enterprises grouping of MSMEs can be done through the amount of assets and turnover as follows:

Table 1
The Types of SMEs Based on asset and turnover

| No | Types | Basic Criterias | |
		Asset	*Turnover*
1	Usaha Mikro (Micro)	Maksimal Rp 50 Juta	Maksimal Rp 300 Juta
2	Usaha Kecil (Small)	> Rp 50 Juta - Rp 500 Juta	> Rp 300 Juta - Rp 2,5 Miliar
3	Usaha Menengah (Medium)	> Rp 500 Juta - Rp 10 Miliar	> Rp 2,5 Miliar - Rp 50 Miliar

The third or final stage of text mining can be done with keywords related to the Government Regulation of the Republic of Indonesia Number 46 Year 2013. The stages of text mining are expected to reduce the number of online business search gradually.

Intelligence System Application

Artificial Intelligence (Artificial Intelligence) is a research area, application and instruction related to computer programming to do something that in human view is intelligent (Simon 1987). Making a machine works like a human way of thinking. This system can solve a problem that is unpredictable and wide coverage.

According Suparman (1991) types of Intelligence Systems is divided into 4 as follows:

1. *Natural Language Processing and Voice Technology:* It is a creation of systems to accept natural language that used by humans. NLP in its development changed the computer's natural language into the computer's natural language.

2. *Expert System:* The creation of a system that has a work like an expert has information from a scientist. The system is run can provide useful knowledge for those who run the system.

3. *Pattern Recognition:* Form a system that can recognize a particular pattern. The pattern is usually in the form of text mining. The useful for approaching a search result. Applications of Pattern Recognition include: voice recognition, fingerprint, identification, handwriting, biological slide, and others.

4. *Robotic:* Merging of all three types of intelligent systems described above. It is a complete system that goals can be accomplished.
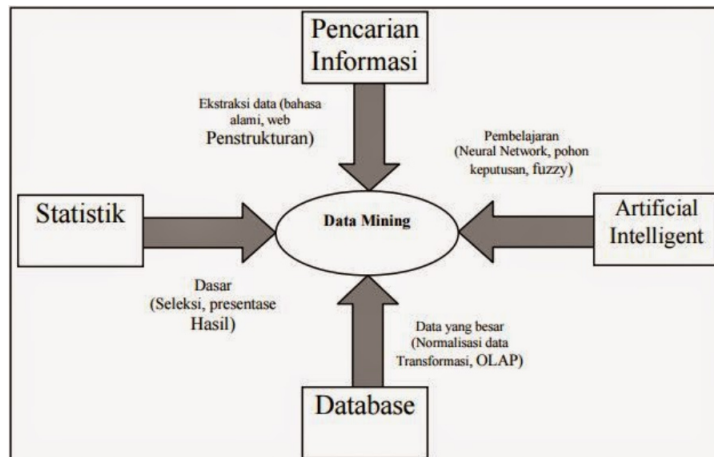
Figure 5: Knowledge Field of Data Mining
Sources: **Pramudiono (2006)**

Data mining inherits many aspects and techniques can be seen from the image above data mining has many long roots from the science of artificial intelligence or Artificial Intelligent, Machine Learning, Database, Statistics, Information (Pramudiono 2006).

An area in artificial intelligence or artificial intelligence relates to the development of techniques that can be programmed and learned from past data. Pattern recognition, data mining and machine learning are often used to refer to the same thing. This field is tangent to the science of probability and statistics as well as optimization. Machine learning becomes an analytical tool in data mining (Santoso 2007).

Theoretical Framework

In online business transactions there is certainly a business process undertaken. The business process consists of several sub processes that have their own clues that will contribute to achieving the goals of the entire process. The use of the business process model aims to define the procedures undertaken to achieve the objectives. This process will be described with using graphic design through the Business Process Diagram (BPD).
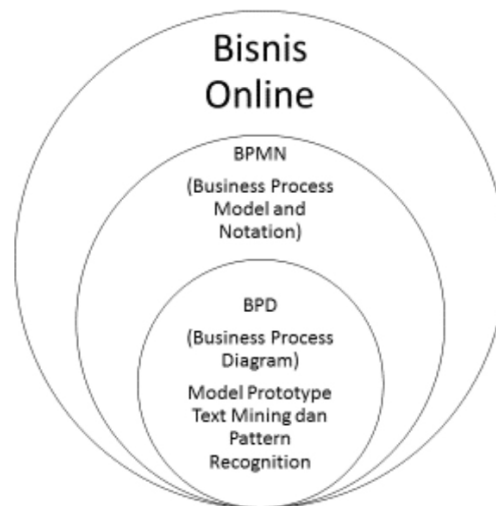
Figure 6: Theoretical Framework

2. RESEARCH METHODS

In defining Research Methods Sugiyono (2008) states that,

> "Research methods are essentially a scientific way of obtaining valid data with the purpose of discoverable, proved, and developed knowledge so that in turn it can be used to understand, solve and anticipate problems".

This research uses experimental research method. Experimental research by Arikunto (2006) is a way to find causal relationships (causal relationship) between two factors that deliberately caused by the researchers by eliminating or reducing or setting aside other disturbing factors. Experimental research is also a deliberate study by researchers by providing treatment or a particular treatment of the subject of research to generate something or event circumstances to be examined how the consequences.

According to Sukardi (2003), there are three types of experimental research designs. Firstly, Pre-Experiment Design. The pre-experiment design attempts to express causal relationships only by involving a group of subjects, so there is no strict control over the extra variables. Secondly, Semu experimental design (Quasi Experimental Research). This quasi-experimental study aims to reveal causal relationships by involving the control group in addition to the experimental group, but the separation of the two groups is not by random technique. Thirdly, the Design of Real Experimental/Pure (True Experimental Research). A pure experimental design is a design used to express causality by involving control groups in addition to the experimental group selected using random techniques.

Research Variables and Sample

Basically, Musdalifah (2014) addressed that the experimental research has 3 variables, namely: Independent variable (Independent) is a variable whose position has an effect on the dependent variable/bound, can be manipulated, changed or replaced; Dependent variable (Dependent) is a variable that is influenced by independent variable; and Variable control is the variable that is not given treatment/experiment but always included in the research process.

Sample repressivity is influenced by several factors (Nafisa, M. 2015), namely: Number of samples. The number of samples is the number of sample groups needed in an experiment. The number of samples is determined by the design of the experiment; and the large sample members. The sample size in the experiment is not determined by the size of the population, but is determined by the strength of the treatment effect of the previous studies. Moreover, Experiments study need sampling technique. Sampling in experimental research can be done two techniques that are random and not random.

Analysis of Experimental Data
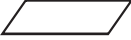
According Samsul Bahri (2014) some data analysis techniques adapted to the experimental design.

1. *Times Series Design:* In this design the group used for the study cannot be selected randomly. Previously given treatment, the group was given pretest up to four times, with a view to knowing the stability and clarity of the group's condition before being treated.

2. *Single Subject Design:* Single subject design generally uses repeatable measurements and implements only single independent variables that are expected to change only one dependent variable.

3. *Control Time Series Design:* This design is a time series design, with a control group.

4. *Separate Sample Pretest-Posttest:* This design is often used in health and family planning studies, the first measurement (pretest) is performed on randomly selected samples from a particular population. A second measurement (posttest) was then performed on another sample group, selected randomly from the same population.

5. *Intact Group Comparison:* The design of this intac group design research is actually derived from the same and related subject groups.

6. *Non-equivalent Control Group Design:* In this design, there are two subject groups where one is treated and one group as a control group. Both get prates and pascates. Differences with non-equivalent groups, that groups are not selected randomly or randomly.

3. RESEARCH DESIGN

The author uses a Quasi Experimental research or Pse experimental design. Because grouping of sample members on control variables and experimental variables is not done randomly. While the data analysis techniques using single subject design, by implementing repeated measurements on a single free variable (Facebook) so as to produce a bound variable (Business online). This research through two stages of Text Mining and Pattern Recognition. The explanation of this research is done with the help of Business Process Model and Notation (BPMN).

Table 2
Business Process Model and Notation

Simbol	*Nama Simbol*	*Fungsi*
	Terminator	Permulaan/akhir Program
	Proses	Proses pengolaha data
	Input/output data	Proses input/output data, parameter, informasi
	Direct Access Storage	Menunjukkan media penyimpanan data yang dapat dibaca
	Decision	Penyeleksian data yang memberikan pilihan untuk langkas selanjutnya
→	Garis alir (*flow line*)	Arah aliran proses

Text Mining

The process of pattern extraction from a large amount of unstructured data. At this stage, the authors collect data with Cluster Sampling technique. According to Musdalifah (2014) "Cluster Sampling is used when the experimental target population is large enough and the researcher wishes to take only a representative portion of the population (sample)". The goal is to categorize or group text. In the stage will appear experimental variables (unofficial online business) and control variables (official online business).

This excavation is done using some examples of key words. Keywords in this study are:

Table 3
Examples of Some Key Words Collections

Jual Beli	Rekening Bersama	Bonus
Ready Stock	Kontak	Original
Pre Order	Endorsement	Garansi
Bank	Promosi	Belanja Hemat
Nomor Rekening	Stok terbatas	Line
Pengiriman	Paid Promote	Shop
Free Ongkos Kirim	Jasa Titip Beli	Dropship
Grosir	Cash on Delivery	Reseller
High Quality	Fast Respon	Kredit

BNIB	BNOB	JNE
Barang Display	PO	Harga
Mandiri	Tiki	Nego
Gojek	Pembeli	Minta Info
Sista	Agan	Invoice

Pattern Recognition

Pattern Recognition (PR) or Pattern Recognition, is the creation of a system to be able to recognize a particular pattern. This pattern will determine the online business through bank account behavior. The pattern used in research by Transaction Time and Nominal Transaction.

Generally of Working Proses
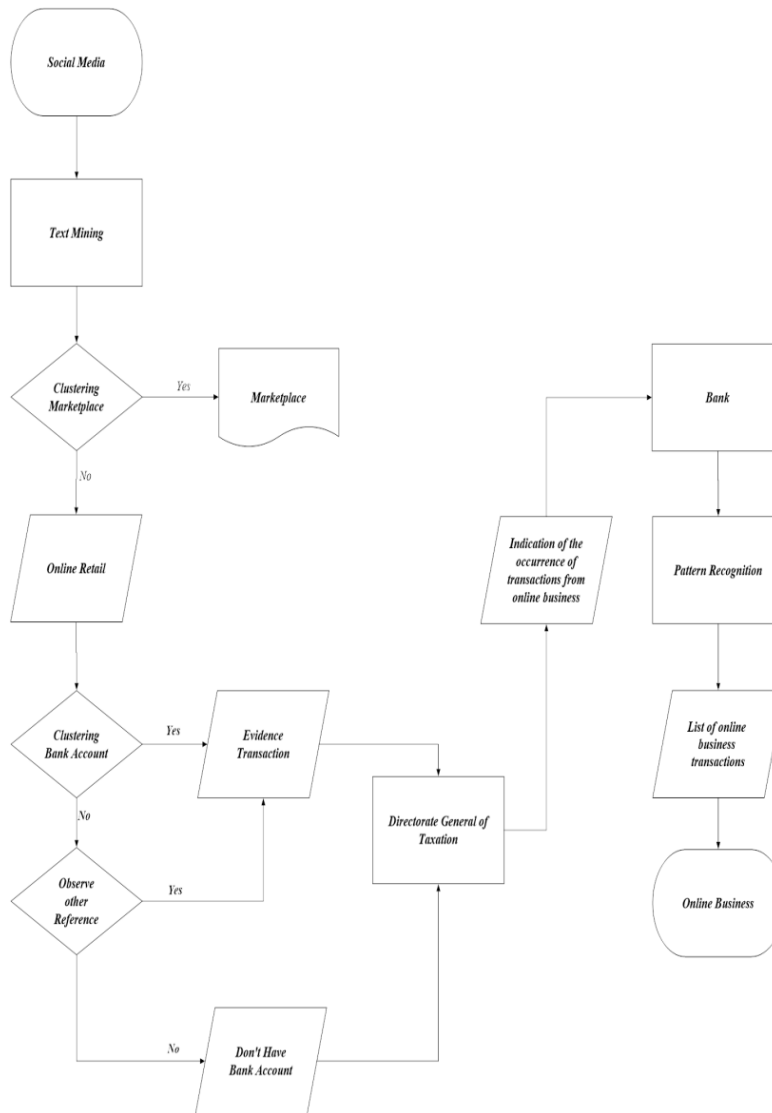
Figure 7: Working Process of Model *Prototype Text Mining* **and** *Pattern Recognition*

4. RESEARCH RESULT

4.1. Working Process Data Mining and Intelligence System: First Stage

The author adopted Google, as well as searches done in some social media and the author chose one of the social media that is Facebook. Facebook is filtered by using existing keywords and mentioned in the Writing Method. An example of the results of using Text Mining can be seen in Appendix 1.
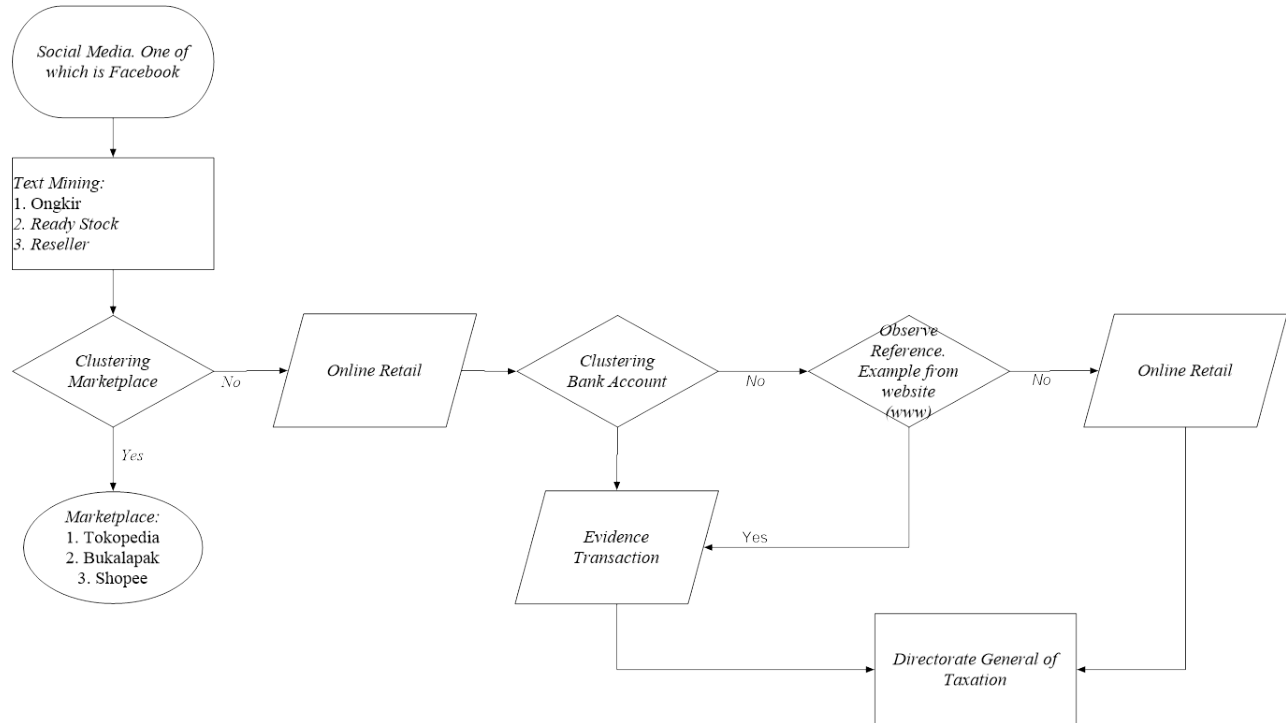
Figure 8: The First Stage

The next process of doing Clustering first, grouping between Official Marketplace entrepreneurs as an example Tokopedia, Bukalapak and Shopee with an unofficial Online Retail. If the online business is from the official Marketplace and Online Retail, the process does not proceed. While the unofficial, the results will be formed into a database that will be done next process is the second Clustering with Bank Account.

Clustering process with Bank Account is divided into two, the first is to do the detection in the "About" or "Home" account of an online business player. Secondly if no Bank Account is detected, the search can be done using other references such as the official website. Examples can be seen in Appendix 2. If the Bank Account is still undetectable, the result will be directly submitted to the Directorate General of Taxation along with the Evidence Transaction result. Examples of Bank Accounts that cannot be detected can be viewed in Appendix 3

The Second Stage

This statement of the Directorate General of Taxation will provide the Indication of the occurrence of transactions from online business or a collection of indications of business actors conducting online
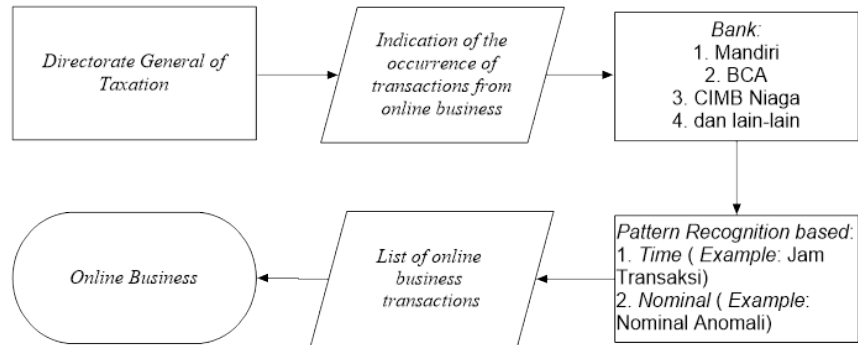
Figure 9: Second Stage

transactions. After that the bank will follow up indication of business actor doing the transaction online by using Pattern Recognition (PR). Pattern Recognition (PR) is an activity that makes the system in order to recognize a particular pattern. The process will result in a List of online business transactions, where the results are unofficial Online Business. Pattern Recognition (PR) can be seen in terms of time and transaction nominal.

Transaction Time, including (a) Frequency of Transaction is how many transactions are performed by the account. Online business can be seen from accounts that actively make transactions. (b) The Transaction Period is that there are certain times for the account in conducting the transaction. For example, at certain hours it appears that the account actively conducts transactions and can also see the account conducting regular transactions to certain accounts. (c) Hours of Transaction ie There are hours of transactions that occur outside of normal working hours.

Nominal Transactions, consisting of (a) Repeat Nominal, describes the nominal amount of incoming or outgoing accounts of the account. (b) Nominal Fluktuatif is the Number of Different Transactions on a daily basis. (c) Fixed Nominal, namely the existence of transactions with a fixed amount to certain accounts. For example deposits. (d) Nominal Anomalies are Transactions that rarely occur but in case of value will be at once large.

5. CONCLUSIONS AND RECOMENDATIONS

Conclusions

From this study, the authors can conclude that, (1) Taxation for businesses that move both in conventional and online is not evenly distributed. (2) After the government regulation on access to financial information for tax purposes, tax revenue is expected to increase. But there is no system that can help the increase. (3) Given this limitation, the author wants to develop a model of supervision on taxation, especially for SMEs online business. (4) In developing the model, the author adopted the workings of Google. But this model is only devoted to detecting online businesses that are not yet registered with the Directorate General of Taxes.

Suggestions

Suggestions that writers can provide are: (1) Tax authorities should create a media discussion for online business actors to better understand tax payment mechanisms and regulations (2) The Directorate General

(DGT) of Taxes may make stricter regulation in the supervision of business on line. (3) There is support from the government to continue this research. Because this research is very useful to reduce the potential of tax revenue lost on online business. (4) The existence of cooperation related social media to implement the model that the author made. (5) The existence of cooperation between the banks and the Directorate General (DGT) of Taxes to conduct illegal online business collectors in the sense of having not registered business and not yet have NPWP for the business.
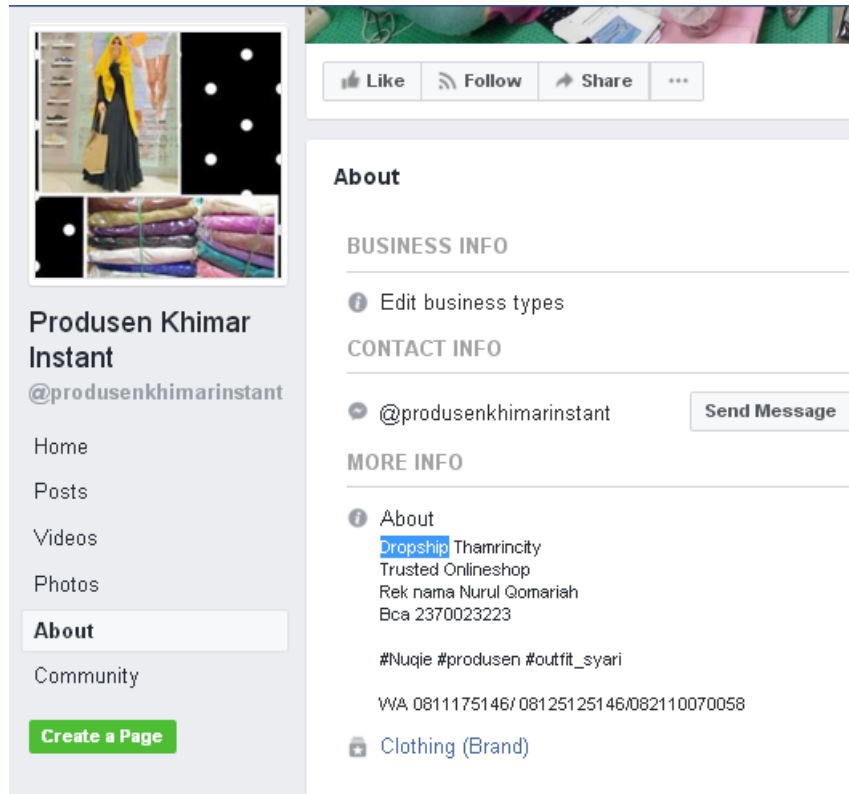
References

Arikunto dan Suharsimi. (2006), *Prosedur Penelitian, Suatu Pendekatan Praktik.* Rineka Cipta. Jakarta.

Bahri, (2014), *Makalah Eksperimen Semu atau Quasi Semu.* Lombok.

Bruce, D dan William F. Fox, (2004), Jurnal State and Local Sales Tax Revenues Losses From E-commerce.

Davies, dan Paul B. (2004), *Database Systems 3rd ed.* Palgrave McMillan. New York.

DigitalMarketer.id. (2015), *Ini Dia Alasan Mengapa Bisnis Online Itu Menguntungkan!.* https://digitalmarketer.id/mindset/ belajar-memahami-bisnis-online-yang-menguntungkan-anda/. Diakses tanggal 11 Agustus 2017.

Feldman, R. dan James, S. (2007), *The Text Mining Handbook.* Cambridge. New York.

Laudon. Kenneth C dan Jane P, Laudon. (1998), *Computer and Information System.* Edisi ke 5, Times Mirror Higher Education Group. United States of America.

Han, J dan Michelin, K. (2006), *Data Mining Concepts and Techniques.* 2nd ed. Morgan Kaufmann Publisher. San Fransisco.

Kompas.com. (2015), *Ini Alasan Facebook Pages Fokus ke UKM di Indonesia.* http://tekno.kompas.com/ read/2015/09/22/15294207/Ini.Alasan.Facebook.Pages.Fokus.ke.UKM.di.Indonesia. Diakses 4 Agustus 2017.

Mendling, J *and* Matthias W. (2012), *Business Process Model and Notation 4th International Workshop.* Vienna. Austria.

Nafisa, M, M. H. Al Ahsad., Dyah A.Y., Marlina A.A.P. (2015), *Penelitian Eksperimen.* Universitas Muhammadiyah Ponorogo.

Pramudiono, I. (2006), *Apa itu Data Mining?.* http://globallavebookx.blogspot.co.id/2015/04/pengertian-data-mining-menurut-ahli.html. Diakses tanggal 14 Agustus 2017.

Santoso, B. (2007), *Data Mining : Teknik Pemanfaatan Data untuk Keperluan Bisnis.* Graha Ilmu. Yogyakarta.

Sidharta, L. (1995), *Pengantar Sistem Informasi Bisnis.* PT. Elex Media Komputindo. Jakarta.

Simon, H.A. (1987), *Artificial intelligence. In R.J. Corsini (Ed.), Concise encyclopedia of Psychology.* 2nd ed. Wiley. New York.

Sugiyono, D. (2008), Metode Penelitian Bisnis. *Bandung: Pusat Bahasa Depdiknas.*

Sukardi. (2003), *Metodologi Penelitian Pendidikan.* Bumi Aksar. Jakarta.

Suparman. (1991), *Mengenal Artificial Intelligence.* Andi Offset. Yogyakarta.

Suyanto, M. (2003), *E – commerce Perusahaan Top Dunia.* Andi, Yogyakarta.

Undang-Undang Republik Indonesia Nomor 7 Tahun 1983 jo Undang-Undang No. 36 Tahun 2008 Tentang Pajak Penghasilan Pasal 4 (1).

Undang-Undang Nomor 20 Tahun 2008 tentang Usaha Mikro Kecil dan Menengah

Vidyana, A. (2014), *Studi Terhadap Kelemahan Pendeteksian Transaksi E-commerce di Indonesia.* Depok. Indonesia.

Yusuf, M. (2014), *Penelitian Eksperimen.* Yayasan Pendidikan Ujung Pandang.

Appendix 1

1. Keyword   : Dropship
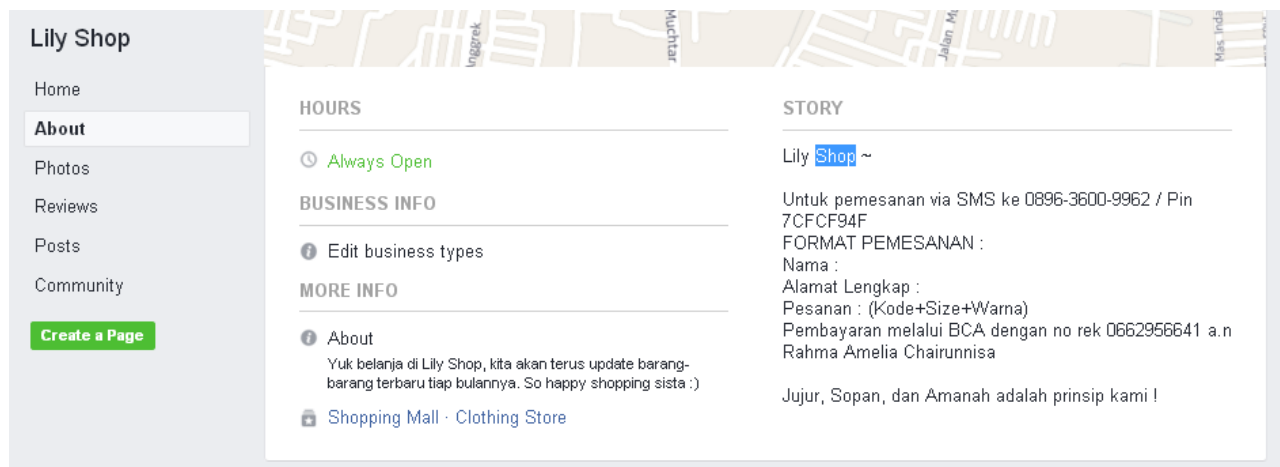
   Result (Hasil) : Nama Bank/Bank Name (BCA), Nama Pemilik Akun/Account Name (Nurul Qomariah), No. Rekening Akun/No. Account (2370023223)

2. Keyword   : Shop

   Result (Hasil) : Nama Bank/Bank Name (BCA), Nama Pemilik Akun/Account Name (Rahma Amelia C), No. Rekening Akun/Account No. (0662956641)
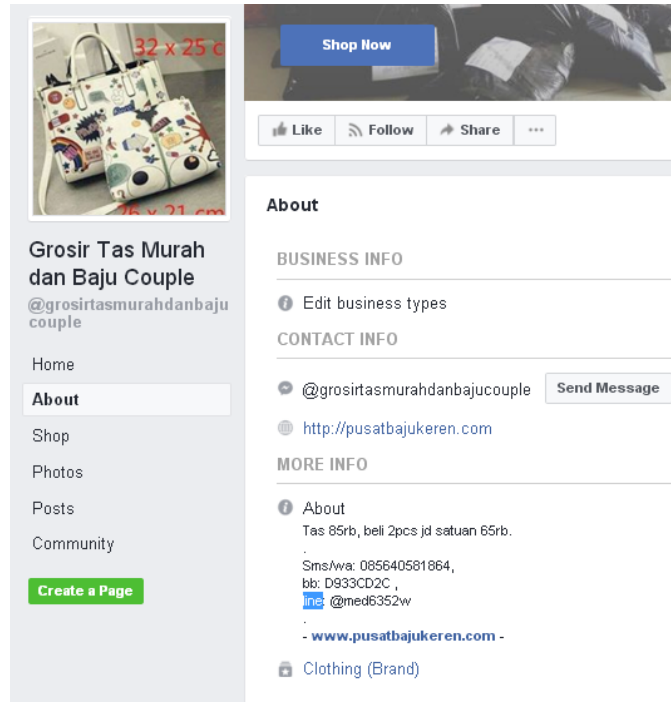
Appendix 2

Keyword   : Line

Hasil   : Referensi (www.pusatbajukeren.com), Nama Bank (BCA, MANDIRI, BRI), Nama Pemilik Akun (Hilda Novita), No. Rekening Akun (BCA : 1691955660, MANDIRI: 1370010342117, BRI: 209901002366502)

Appendix 3

Keyword   : Kredit

Result (Hasil)   : Tidak ditemukan (No found) Bank Account.