2017 International Conference on Information Technology Systems and Innovation (ICITSI)
Bandung, October 23 – 24, 2017
ISBN: 978-1-5386-3100-3

Online Price Prediction System of Consumption Commodities

Neti Sriwulan Sutiawan
School of Electrical Engineering and Informatics
Institut Teknologi Bandung
Bandung, Indonesia
neti@students.itb.ac.id

I.G.B. Baskara Nugraha
School of Electrical Engineering and Informatics
Institut Teknologi Bandung
Bandung, Indonesia
baskara@stei.itb.ac.id

***Abstract*—Growth of online shopping brings possibilities for research and study. We can use online price data for example, for forming consumer price index. This research proposes online price prediction system that use online price commodities from online retailers as data source. To collect many products information as consumption commodities from several online retailer, we use web scraping technique. Collecting consumption commodities from online stores is possible and well proven. We also find the acceptance of proposed system using Technology Acceptance Model (TAM) from user's perspective. The result show that this system can predict price of certain commodity close to the price data survey and 83,34% of respondents believe this system can be implemented as supported system or substituted of Online CPS System in the future.**

Keywords—Online Price Prediction System ; Online Retailer;Consumer Price Index; Web Scraping; TAM

I. INTRODUCTION

Price stability is a goal to control inflation [1]. Inflation is the tendency of the prices of goods and services to rise in general and continuously [2][3]. High inflation may cause a negative impact on the socio-economic conditions [1]. A common indicators that used to measure inflation is Consumer Price Index (CPI). Unstable prices have an effect on unstable inflation so that will create uncertainty for economic actors in making decisions.

The government always tries to monitor inflation and price so that price stability is maintained. As undertaking are to monitor existing commodity price information and make inflation prediction for anticipation [4].

There are some existing information systems that can facilitate government to monitor price commodities. For example, staple food market monitoring system (SP2KP in Indonesian), system of trade ministries that monitor the prices of staple foods each day. This system is used for early warning system against increase or decrease in price significantly. But food prices are monitored only a few commodities (12 commodities). Moreover, it is not equipped with prediction of price data but only show the data from today and earlier. Whereas as the system aims for early warning, prediction can be used to know the price of commodities in future so we can know the prediction price data for anticipating. Data collection of the system obtained through SMS every day. The downside, not all surveyors can be consistent in sending sms every day except conducted intensively monitoring. Another weakness is still prone to data input errors.

BPS-Statistics Indonesia as an official statistics that legally measure CPI and inflation data sourced from the price, also have a price information system (Online Consumer Price Survey (CPS) System) which is used to record price commodities data. Commodities which are collected more than commodities which are collected from SP2KP. The amount of that are from 225 to 462 commodities. This commodities samples are different in every city depend on consumption commodities by its society. In addition, each month government usually asks for the prediction of inflation or CPI from BPS before the inflation or CPI rate release. But prediction results are not published. To predict CPI or Inflation also carried out separately from the price information system.

In this study, we will propose an online price prediction system of consumption commodities. The consumption commodities are CPI commodities. This system aims to record data of commodity prices and also make predictions in one system. Source of data used comes from online retail e-commerce. In this research, we selected Hypermart and Alfacart online retailers as sample. As researches [5][6] that have analyzed the quality of information that fit for data collection in formatting of online daily CPI rates is e-commerce online retailers.

There are some techniques for retrieving data from online source such as online retailers in prediction system. Research on prediction system using online data sources mostly on prediction of stocks price. Web crawling and web scraping technique are commonly use for retrieving data from website. Some research in stock price prediction system, usually use yahoo finance historical price as data source [7][8][9][10][11].

Yahoo finance has web API for serving numerical data source which can be accessed by other system in XML format. Web crawling mostly use for retrieving data because in XML format crawler can get the whole data from web pages without much effort in parsing data later. However for online retailers, their website mostly do not serve an API data and formatted website are HTML. So web scraping technique can solve this kind of problem because it can retrieve exact to the data from particular web pages which parsing it directly from HTML structure [12].

Thus, in this research we will propose an online price prediction system of consumption commodities. For evaluating the usefulness and acceptance of the proposed system, we will measure the acceptance by user's perspective using Technology Acceptance Model (TAM) after implementing prototype system. We will also find the factors that have positive influence on acceptance of the system from user.

II. Related Work

There are some researches of price prediction using data and information online. R.P Schumaker and H Chen which constructed the stock price prediction system uses linguistic synthesis techniques (prediction of textual documents) on the system Arizona Financial Text System (AZFinText) [13]. This system was extracted data from financial news articles using machine learning techniques. Bakhsi constructed the stock price prediction system using online data with fundamental and technical approach [7]. For fundamental analysis of the data used comes from MarketWatch News (marketwatch.com) and Reuters 'Key Developments' News (reuters.com) is a site that provides information about the stock market, financial news and the latest business while technical analysis uses historical data from Yahoo Finance Historical Price. Chun Yan et al.[14] conducted a research on data extraction of online prices for the decision support system on agriculture price product or commodity. He extracted certain agricultural products from the websites of wholesalers of agricultural and focus on data extraction tables. To meet these goals used an algorithm called MDT-E (Market Data Table Extraction).

Frow previous works, most of the systems use Yahoo finance as source data that has web API for serving numerical data source which can be accessed by other system in XML format. Web crawling mostly use for retrieving data because in XML format crawler can get the whole data from web pages without much effort in parsing data later. However for online retailers, their website mostly don't serve API data and formatted website data are HTML. So web scraping technique can solve this kind of problem because it can retrieve exact to the data from particular web pages which parsing it directly from HTML structure [12].

III. System Design

A. General Overview of System

In this research, we proposed an online price prediction system of consumption commodities that support web-based applications. First, user interacts to the system via website user interface (web browser). System can extract data from data source (online retailers) and store them to database. Prediction process invoked WEKA libraries which prediction models will be built using forcasting algorithm from WEKA libraries in Model Manager Prediction Module. The predicted results show in chart of commodity. The prediction result is monthly prediction so we can see next month prediction of the commodity. Data are also showed in the form of tables, chart of scraping data and prediction commodity. The design of the proposed system can be seen in Fig.1:

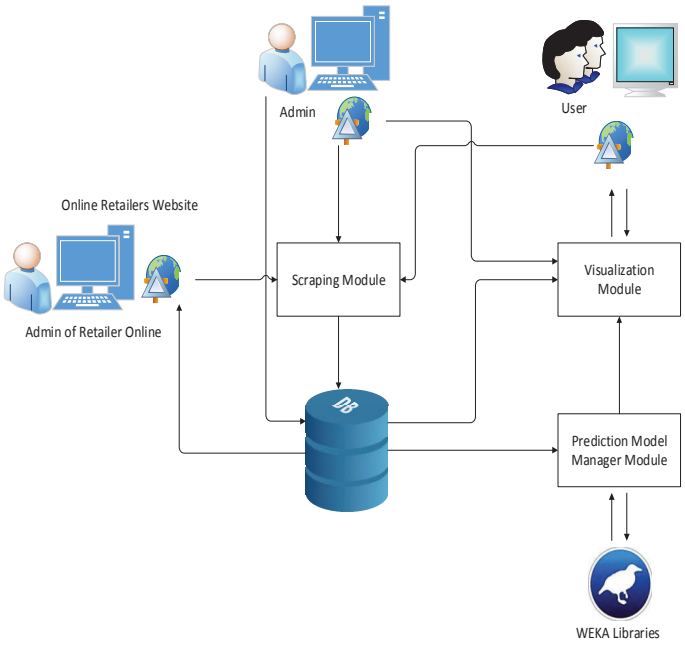

Fig. 1. Online Price Prediction System of Consumption Commodities Architecture

B. Components of Proposed System

The overall design of the proposed system consists of 4 main components as follows (Fig.1):

1*). User Interface Module*

User Interface module provides web-based applications. In web-based application, users can perform various tasks, for example, users can collect commodities price data via web scraping menu manually depend on needs or every hour automatically. From data that have been collected user can predict the certain price of commodities for next month prediction and the result show on the chart of prediction commodities. Users also can see the table and chart of every commodity per hour, per day or per month that has been stored in database. Beside scraping and prediction, admin also can manage master data which contain of commodities information and also information of data source from a certain online retailers.

2). Scraping Module

- Gathering Data and Scraping Process

Input system is input consumption commodities price data which are taken through the process of scraping from selected URL online retailers. Admin first collecting the URL commodities from online retailers suit with information the commodities of CPI such as quality or brand of commodities. This aims to keep the historical product easily and also to compare with CPI commodities.

Scraping process is retrieving data from a data source and then process it into a data or information to be displayed by system. scraping, the process for extracting price data from web pages online retail as a data source (input) on the system. The process begin with collecting URLs commodities from retailers online. When scraping process is running, the system will read web page and parsing price data from web page then store them into database. To implement process will use cURL library that can parse the data we need exact from web page, such as commodity price. The process can be seen at Fig.2 (left).

3). Prediction Model Manager Module

In this module, we first prepared and preprocessed the data. Latest and historical per quality commodities price data in database were prepared and calculate to the average month of price commodity. We used for example, Multilayer Perceptron Algorithm in WEKA tools as prediction method. This can be changed by other algorithm that can be adjusted by the code while invoked WEKA. Moreover, we also updated and extended database dynamically based on scraping process. Thus, when new data stored into database, a new model can be rebuilt and updated according to aforementioned process. The prediction process can be seen in Fig.2 (right).

Process of WEKA (library) for making prediction models and the result is read by the system through the backend process. Data which is built from WEKA library can be received and sent to web application. Service modules in Prediction Model Manager was implemented by invoking WEKA library. When the result model was done and sent them back to the module in the form of XML File.

4). Visualization Module

This module is used for displaying the commodities information from database. This also show the result of prediction of certain commodity price which is has built from prediction model manager module. So the main output of the system is result of scraping consumption commodity price and prediction result. Data are displayed in the form of tables, chart of scraping data and prediction commodity in a certain periods such as per hour, per day or per month. The predicted results show in chart of commodity. The prediction result is next month prediction of the commodity (monthly prediction).

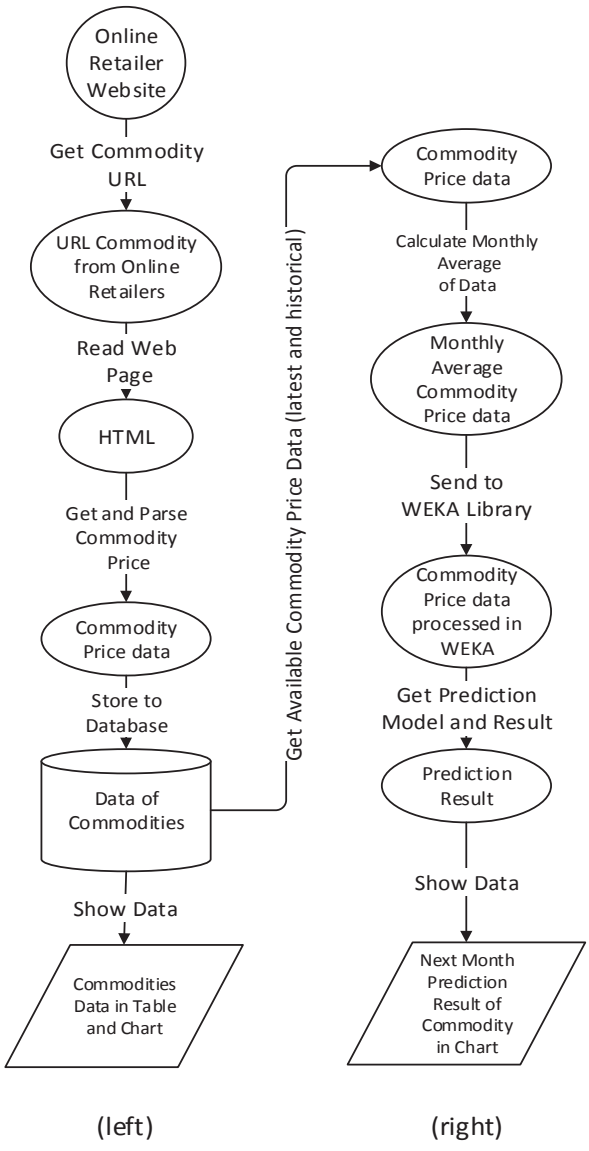

Fig. 2. Scraping Process (left) and Prediction Process (right) in Proposed System

IV. RESULT

A. Impementation

Implementation of the system is built using PHP programming language and Yii Framework 2.0. Yii can be used to build different types of web-based applications, such as: web portal, forum, CMS, e-commerce, web services, and others. In addition, Yii also has features that are common to the framework of other modern, including: MVC (model - view - controller), supports many databases, application templates, easily integrated with the library the other, caching, RESTful, code generators, and so on [15]. Using Apache as a server and MySQL for database.

B. Functionality Testing

Functionality testing was conducted to observe and functional check of the software or also called black box testing and User Acceptance Test (UAT). Evaluation is only its external appearance (interface) and the functionality or just know the input and output without knowing what actually happens in the process in detail. The result showed that the functionality features of the system are executed successfully.

Furthermore, after user finishing UAT test, then System Usability Scale (SUS) method from John Brooke [16] was conducted. This method explains that test steps should include: effectiveness (the ability of users to accomplish tasks using the system, and output quality of the resulting task), efficiency (the level of resources consumed in performing tasks), and satisfaction (user subjective reactions in using the system). This coverage is outlined in 10 system questions on the SUS questionnaires distributed to users. The questionnaire uses a likert scale covering various aspects of system usability, such as the need for support, training, and complexity that have a high degree of validity in measuring usefulness of the system. User filled questionnaire based on that questions.

Based on 30 users who answer the 10 questions from the SUS questionnaire, the results of score found that the average score of the system of all respondents is 74.58. This value is greater than 68 (system category above average or good) indicates that the system has a good usability aspect value and can be accepted by the user.

C. Evaluation Proposed System Acceptance

We want to know how acceptance of proposed system by using an instrument to measure user's perspective. We choose Technology Acceptance Model (TAM) [17][18][19] to find how the acceptance of this proposed system by user's perspective and describe what factors that have positive influenced of the acceptance system are.

Model can be seen in Fig.3. In this model showed the relationship between factors or independent variables and dependent variable which is the acceptance of proposed system.

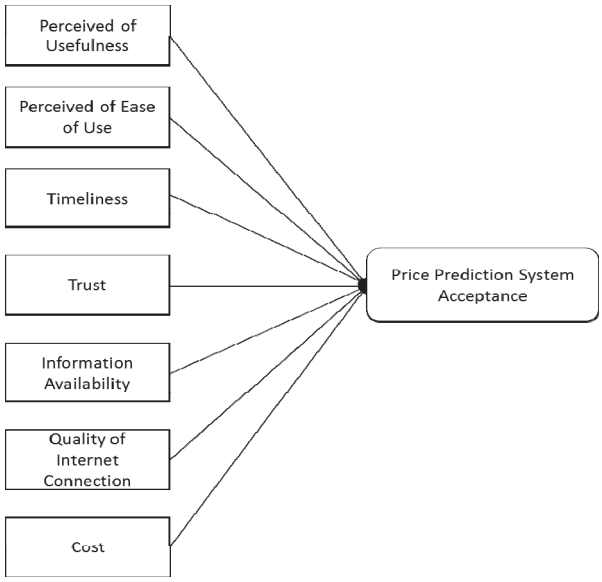

Fig. 3. TAM of Online Price Prediction System Acceptance

Respondents used in this research are Indonesia-Statistics Office (BPS) employees. A total of 33 respondents fill out the questionnaires. Respondents were first asked to see a comparison between Online CPS System and compared to the Proposed System, especially in the system process flow. Comparation process flow between Online CPS System and proposed system can be seen in Fig.4. After that the respondents give the view through questionnaires as acceptance user's perspective on proposed system if implemented in the future.

After data collection from users was done, we use descriptive and regression analysis as statistical testing. We find that 83.34% of respondents agreed (tend to agree - strongly agree) if Online Price Prediction System of Consumption Commodities can be implemented as a supported system and can replace Online CPS System in Official for forming CPI or Inflation in the future (assuming all or most of the commodities are available on the website). Another result are factors that have positive influence on accepted of proposed system. That factors or independent variables are Perceived of Usefulness, Trust and Information Availability.

Based on Perceived of Usefulness factor that has positive influence, we can see that users consider about online price prediction system can reduce time in the process of collecting pricing data, especially the price from supermarket or department store that have online retail sites beside physical store. Thus it can improve performance on commodity price

data collection on CPI formation and make predictions on commodity prices on the same system.

Based on Trust factor that has positive influence, we can see that users consider about trustworthiness or reliability especially about source data that can influence the output that will be resulted from proposed system. Another reason that the source data from online retailers can be used as additional or complementary data source on consumer price surveys.

Based on Information Availability factor that has positive influence, we can see that users consider about availability commodities data from online source and also availability data when accessed from the system. Therefore this proposed system acceptance are influenced by user's perspective of the usefulness of the system, trust for the output and input system and also the availability of the data weather availability of source data and output of the system.

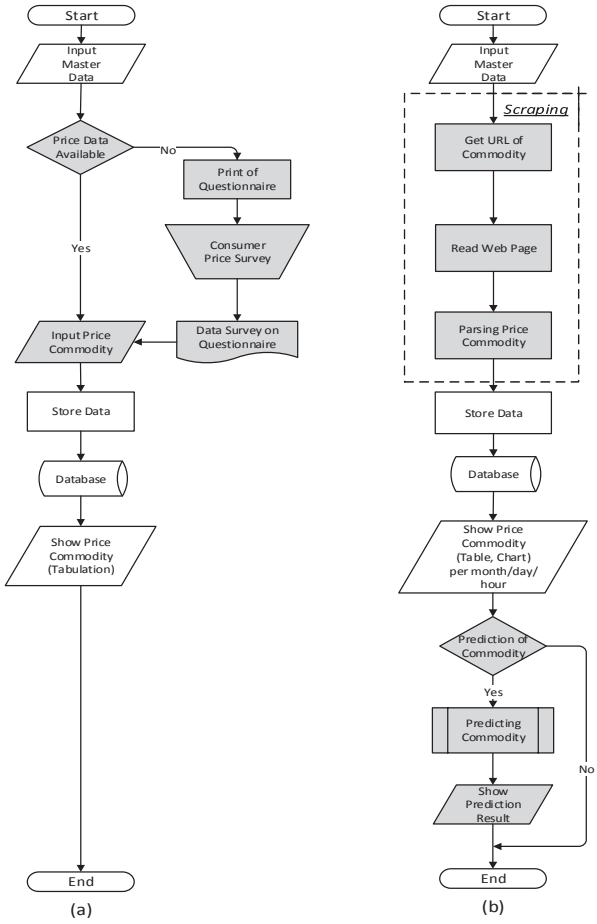

Fig. 4. Comparison Process Between Online CPS System in BPS-Statistics Indonesia (a) and Proposed System (b)

V. CONCLUSION

In this research, we have been successfully developed Online Price Prediction System of Consumption Commodities using online retail data sources and systems can make predictions using the selected prediction method. In this research also measure acceptance of the proposed system when compared with the existing information system (Online CPS System) and result show a good user's perspective. The result show that 83,34% of respondents believe this system can be implemented as supported system. Factors or independent variables that have positive influence on acceptance of implementing proposed system are Perceived of Usefulness, Trust and Information Availability.

REFERENCES

[1] -------, "Introduction of Inflation - Bank Indonesia," [Online].Available: http://www.bi.go.id/id/moneter/inflasi/pengenalan/Contents/Pentingnya.aspx. [Accessed: 19-Dec-2016].

[2] -------, "Introduction of Inflation - Bank Indonesia," 2013.[Online].Available: www.bi.go.id/id/moneter/inflasi/pengenalan/Contents/Pentingnya.aspx. [Accessed: 19-Dec-2016].

[3] -------, *Manual Book of Consumer Price Survey*. BPS-Statistics Indonesia, 2013.

[4] A. Kadir M, P. R. Widodo, and G. Suryani R, *Implementation of Monetary Policy of Framework of Inflation Targeting in Indonesia*, no. 21. Center for Education and Studies Central Bank BI, 2008.

[5] A. Cavallo, "Online and Official Price Indexes: Measuring Argentina's inflation," *J. Monet. Econ.*, vol. 60, no. 2, pp. 152–165, 2013.

[6] D. P. Manik and Albarda, "Consumer Price Index by using big data in Statistics Indonesia," *Int. Conf. Inf. Technol. Syst. Innov.*, pp. 1–5, 2015.

[7] N. Bakhsi, "Stock Market Prediction Using Online Data : Fundamental and Technical Approaches Nikhil Bakshi Master ' s Thesis in Computer Science," 2008.

[8] D. Mundada, G. Chhaparwal, S. Chaudhari, and T. Bhamare, "Stock Value Prediction System," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 3, no. 4, pp. 2217–2219, 2015.

[9] S. Madge, "Predicting Stock Price Direction using Support Vector Machines," 2015.

[10] R. Aghababaeyan, "Forecasting The Tehran Stock Market by Artificial Neural Network," *Int. J. Adv. Comput. Sci. Appl.*

[11] Y. Wang, "Stock Price Direction Prediction by Directly Using Prices Data : An Empirical Study on The KOSPI and HSI," *Int. J. Bus. Intell. Data Min.*, vol. 9, no. 2, pp. 145–160, 2014.

[12] M. Turland, *php | architect ' s Guide to Web Scraping with PHP*, First. Toronto: Marco Tabini & Associates, Inc., 2010.

[13] R. P. Schumaker and H. Chen, "A Quantitative Stock Prediction System based on Financial News," *Inf. Process. Manag.*, vol. 45, no. 5, pp. 571–583, 2009.

[14] Y. U. Chun-yan, M. A. Jun, and Z. Yu-yan, "Online Price Extraction and Decision Support for Agricultural Products," *Int. Conf. Inf. Manag. Innov. Manag. Ind. Eng.*, pp. 337–340, 2009.

[15] M. Safronov, *Web Application Development with Yii 2 and PHP*, Third Edit. Packt Publishing Ltd., 2014.

[16] J. Brooke, “SUS - A quick and dirty usability scale.”

[17] P. D. D. Dominic and H. Khan, “Evaluation of Online System Acceptance Through Airlines Websites In Malaysia.”

[18] C. A. Colvin and A. Goh, “Validation of The Technology Acceptance Model for Police,” vol. 33, pp. 89–95, 2005.

[19] B. Narteh, S. Lecturer, and C. Management, “Journal of Internet Banking and Commerce,” vol. 17, no. 2, 2012.