ATLANTIS PRESS

Advances in Social Science, Education and Humanities Research, volume 305

International Conference on Ethics of Business, Economics, and Social Science (ICEBESS 2018)

Tourist Destination Mapping Visualization in Yogyakarta using Affinity Propagation and Social Network Analysis: A Case Study on @explorejogja Instagram in 2017

Rahmadi Yotenka
Department of Statistics
Islamic University of Indonesia
rahmadi.yotenka@uii.ac.id

Muhammad Muhajir
Department of Statistics
Islamic University of Indonesia
mmuhajir@uii.ac.id

Nanda Iswandhani
Research and Development MNC Channels
MNC Studios International
niswandhani@gmail.com

Hafizah Ima
Departement of Statistics
Islamic University of Indonesia
15611161@students.uii.ac.id

***Abstract*—This research was conducted in Department of Statistics Islamic University of Indonesia. The data used are primary data obtained by post @explorejogja Instagram account from January until December 2017. In the @explorejogja Instagram account found many tourist destinations that can be visited by tourists both in the country and abroad. Therefore, it is necessary to form a cluster of existing tourist destinations based on the number of likes from user Instagram assumed as the most popular. The purpose of this research is to know the most popular distribution of tourist spot, the cluster formation of tourist destinations, and central popularity of tourist destinations based on @explorejogja Instagram account in 2017. Statistical analysis used is descriptive statistics, Affinity propagation, and social network analysis.**

Keywords—destination, cluster, map, social network analysis

I. INTRODUCTION

A. *Instagram account*

Instagram contains about a variety of account content, ranging from online shop accounts that sell goods and services, campus organization accounts and also Instagram accounts that promote tourism in a country or city. Tourism is a journey from one place to another, temporary, done by individuals or groups, in an attempt to find balance or harmony and happiness with the environment in the social, cultural, natural and scientific dimension [1] (see Table I).

TABLE I. LIST OF INSTAGRAM ACCOUNT RELATED TO TOURIST DESTINATION

Number	Instagram account	Post	Followers	Following
1	Explorejogja	1729	591K	0
2	Wonderfuljogja	3250	452K	7440
3	Dolanjogja	2147	158K	216

From the Table I above there are 5 Instagram accounts that promote tourism in Yogyakarta, of the five accounts Instagram explorejogja account occupy top 1 with the number of followers most as much as 591K

B. *Explorejogja*

Explorejogja is an Instagram account which contains photos about some interesting places in Yogyakarta Special Region (DIY) as depicted in Figure 1. This account is also intended as a witness of the development of tourist attractions in DIY for generations to come [2].

Fig.1. Explorejogja Instagram account

Existence means unity with the outside world. The world belongs in the structure of existence. [3]. It can be interpreted that external impulses can affect existence.

Data mining is one method of statistics that has an important role in extracting large amounts of data. Post photo data in post account Instagram explorejogja require data mining in its analysis. Data posting account Instagram explorejogja is taken at the end of 2017. In this study using 313 post as a sample population of 1537 post overall from account Instagram explorejogja

C. Visualization of tourist destination mapping

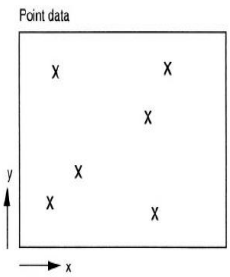
Based on the background that has been described, the writer took the title "Visualization mapping of tourist destination in special region of Yogyakarta using Affinity Propagation method and Social Network Analysis ". in this case aims to see the visualization of tourist destination mapping Yogyakarta province, then grouping based on the level like as the impulse from the outside (the existence) of the user Instagram and apply the graph into the map location of most tourist attractions based on the district. From these results will be obtained information and recommendations of travel destinations to facilitate tourists to know the location of existing destinations according to post account Instagram explorejogja in 2017.

The rest of this paper is organized as follow: Section II describes the proposed method. Section III presents the obtained results and following by discussion. Finally, Section IV concludes this work.

II. PROPOSED METHOD

A. Spatial Analysis

Spatial analysis is a visual inference to a map that is a combination of spatial data and attribute data as depicted in Figure 2. Spatial data refers to a location or position on the surface of the earth, which is the coordinates, raster or administrative boundaries of the region.



Attribute data

3577695	5766285	sL
3575805	5762175	tL
3592365	5774295	lS
3592395	5774085	sL
3595345	5774455	sT
3592497	5774585	sT

Fig. 2. Differences of spatial data and attribute data [4]

B. Data Mining

Data mining is the mining or discovery of new information by searching for a particular pattern or rule from a large amount of data [5]. Data mining is also called a series of processes to explore the added value of knowledge that so far cannot be known manually from a data collection [6]. Data mining has stages like in Figure 3.

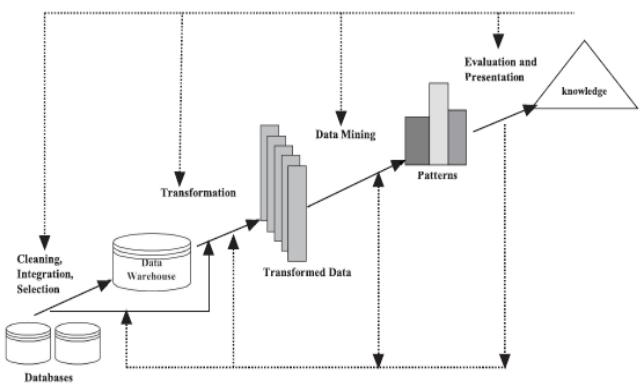


Fig. 3. Data mining process

C. Affinity Progagation

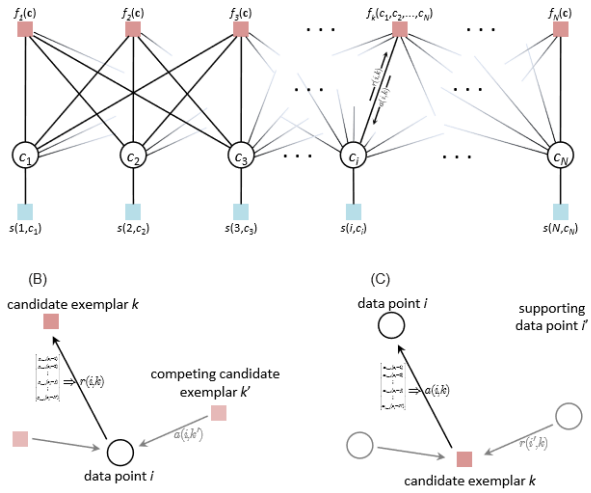


Fig. 4. Affinity propagation algorithm

Affinity propagation is an exemplar-based clustering algorithm that performs belief propagation on the factor graph shown in (A) as depicted in Figure 4. Two kinds of message are passed in the graph; responsibilities (B) are passed from variable nodes to function nodes (i.e., data points to candidate exemplars). Availabilities are passed from function nodes to variable nodes (C), interpreted as candidate exemplars to data points [7] (see Figure 5).

$$F(\mathbf{c};\mathbf{s}) = e^{\overbrace{\sum_{i=1}^{N} s(i,c_i)}^{S \text{ from } (2.9)} + \overbrace{\sum_{k=1}^{N} \log f_k(\mathbf{c})}^{\text{coherence constraint}}} = \prod_{i=1}^{N} e^{s(i,c_i)} \cdot \prod_{k=1}^{N} f_k(\overbrace{c_1, c_2, \ldots, c_N}^{\mathbf{c}})$$

Fig. 5. AP Function

Note that the first term in the exponent involves the net similarity, *S*, from the *k*-median problem, except that similarities are exponentiated to ensure *F*(*c*;*s*) always evaluates to a positive function. The second term contains a coherence constraint defined as follows:

which causes the function to evaluate to zero for the incoherent configuration of a cluster without an exemplar, i.e., a data point *i* has chosen *k* as its exemplar (*ci* = *k*) with *k* having been incorrectly labelled as a non-exemplar (*ck*6=*k*). Each component of *F* (*c*;*s*) is represented by a function node and each label ci is represented by a variable node. Each *fk* (*c*) term appearing in equation (1) has a corresponding function node that is connected to all variables *c*1,*c*2,...,*cN*. In addition, each *s* (*i*,*ci*) term has a corresponding function node that is connected to the single variable *ci*. The log of the global function *F*(*c*;*s*)—in this case *S* (*c*) (previously referred to as net similarity, *S*)—is given by the sum of all the log-functions represented by function nodes

D. Social Network Analysis

Social Network Analysis (SNA) is one of the analysis in data mining that connects several interrelated objects through graph. Objects in SNA called actor terms are the main focus in this analysis.

There are two types of relationships that can be explained in SNA, namely [8]:

- *Directional Relations*: the type of relationship "self-choices" that the relationship that occurs between actors is the choice of each actor and does not apply to each other in opposite, e.g. friendship relationship between *A* and *B*. If *A* recognizes *B* as a friend then not *B* will recognize *A* as his friend. *X*: friendship sociometric *A*, *B* and *C*. If it is known that A is friends with *B* (*A B*), *B* be friends with *C* (*B* → *C*), *C* be friends with *A* (*C* → *A*), dan C be friends with *B* (*C* → *B*), If a relationship is dichotomous then the element of the matrix *X* (X_{ij}), with *i*= *A*, *B*, *C* and *j*= *A*, *B*, *C* are:

$$X = \begin{matrix} - & 1 & 0 \\ 0 & - & 1 \\ 1 & 1 & - \end{matrix} \quad (1)$$

- *Nondirectional Relations*: the type of relationship between actors symmetrical each other. Example of this relationship is the neighbouring neighbourhood. If *A*-neighbour is next door with *B* then it is definitely *B* after the house again with *A*. This type of relationship will be denoted by a line (without arrows) on the sociogram. In the form of matrix notation, this relationship can be described as follows:

- *X*: sociometric next door neighbour *A*, *B* and *C*. If it is known that *A* next door neighbour with *B* (*A*-*B* and *BA*), adjacent neighbour *B* with *C* (*B*-*C* and *C*-*B*). If a relationship is dichotomous then the element of the X matrix (X_{ij}), with *i*=*A*, *B*, *C* and *j*= *A*, *B*, *C* are:

$$X = \begin{matrix} - & 1 & 0 \\ 1 & - & 1 \\ 0 & 1 & - \end{matrix} \quad (2)$$

In SNA, the relationships between actors can be valuable dichotomies and have value. The relationship is dichotomous if the relation exists then it is worth 1 and if there is no relationship will be worth 0. The relation can also be valuable so that each relationship between actors has different values, it can be worth the strength of the relationship between the actor, the intensity of the relationship or the frequency of the relationship.

E. Research Methods

Data analysis method used in this research is descriptive analysis, tourism destination distribution in the form of web based map, non-hierarchy cluster analysis that is SOM method and application of graph theory on spatial. The tools used in data analysis in this study are Ms. Excel, QGIS 2.10, and Gephi.

III. Results and Discussion

A. Distribution of tourist destination

Instagram accounts explorejogja have a fairly large existence as an account that promotes tourist destinations contained in the province of Yogyakarta. these accounts also get a fairly positive response from Instagram users, especially followers (followers) of the explorejogja account.

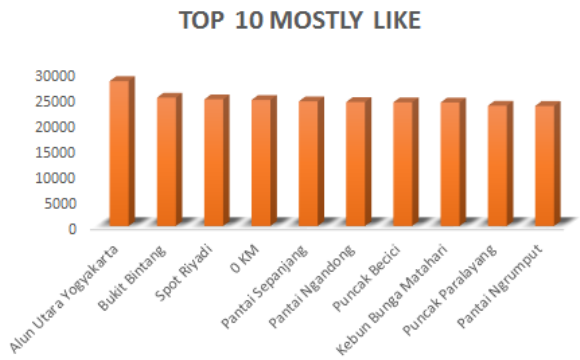


Fig. 6. Top 10 of explorejogja post

From the Figure 6 above can be seen top 10 most popular tourist destination from January until December 2017. Namely is Alun Utara Yogyakarta, Bukit Bintang, Spot Riyadi, 0 KM, Pantai Sepanjang, Pantai ngandong, Puncak becici, kebun bunga Matahari, Puncak paralayang, Pantai Ngrumput.

Based on 113 tourist destinations obtained,then in the form of map web-based tourist destination with 113 point coordinates (see Figure 7).

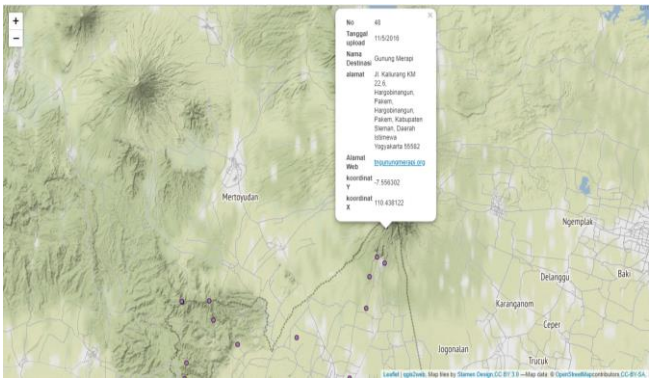


Fig. 7. The map of Tourist Destination Destinations in special region of Yogyakarta

As for the explanation of tourist destinations in accordance with the coordinate point as follows in Table II:

TABLE II. LIST OF TOURIST DESTINATIONS ALONG WITH POINT COORDINATES

No	Destination Names	Coordinate point	
		Latitude	Longitude
1	Ayunan Langit Watu Jaran	-7.735783	110.132603
2	0 KM Yogyakarta	-7.80136	110.364771
3	Air Terjun Curug Gedhe	-7.824035	110.536747
4	Wisata Seribu Batu Songgo Langit	-7.931236	-7.931236

B. Affinity Propagation Cluster

The clusters obtained are depicted in Figure 8.

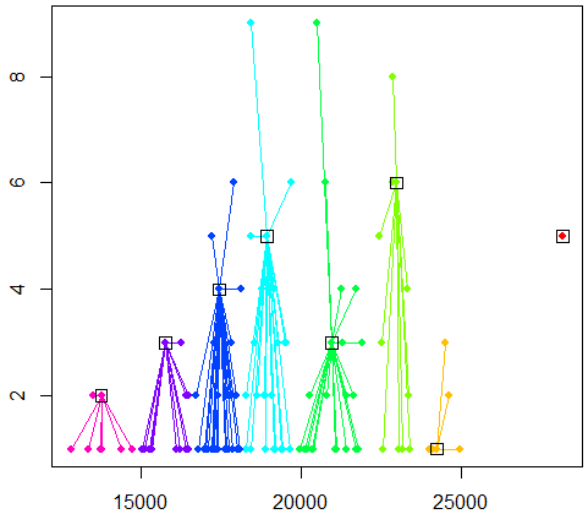


Fig. 8. Cluster result

The AP method processes all objects simultaneously where k is the number of groups. So in the method also determined the number of groups formed, in this study in the form of Five groups as described in Table III.

TABLE III. A LIST OF TOURIST DESTINATIONS ALONG WITH POINT COORDINATES

	Clusters				
	1	2	3	4	5
Number of tourist destination	1	9	29	66	106

C. Social Network Analysis

Graf as depicted in Figure 9 is formed based on the type of tourism consisting of 8 types. namely nature tourism, marine tourism, education, history, culture, religion, tirta, and night.

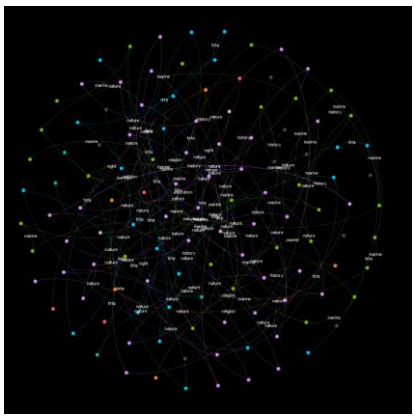


Fig. 9. tourist destinations based on the type of tourism consisting

Based on the picture above can be seen that there are 8 different colors of the plot with purple, green, blue, brown, orange, pink, green tosca, and gray. The explanation related Graf is described in the following Table IV:

TABLE IV. EXPLANATION GRAF LOCATION OF THE SPREAD OF TOURIST DESTINATIONS

No	Type of tourist	Coloring	Percentage
1	Natural tourism	Purple	38.02%
2	Marine tourism	Green	23.14%
3	Tirta tourism	Blue	18.18%
4	Historical tourism	Brown	8.26%
5	Religious tourism	Orange	4.13%
6	Night tourism	Pink	3.31%
7	Culture tourism	Green tosca	3.31%
8	Educational tourism	Gray	1.65%

IV. CONCLUSION AND RECOMMENDATION

The Distribution of tourist destinations in Yogyakarta province is displayed in the form of a web-based dynamic map in which there are 113 points coordinates that symbolize each tourist destination. Of the 113 destinations are then obtained top 10 most popular destinations based on the number of like. Namely, Alun Utara Yogyakarta, Bukit Bintang, Spot Riyadi, 0 KM, Pantai Sepanjang, Pantai ngandong, Puncak becici, kebun bunga Matahari, Puncak paralayang, Pantai Ngrumput. Graf is formed based on the type of tourism consisting of 8 types. namely nature tourism, marine tourism, education, history, culture, religion, tirta, and night. Based on the conclusions obtained by the researcher wants the results of this study can be used by the Yogyakarta government, especially in the field of tourism in order to be used as a reference to improve quality and publications related tourism that has been considered popular. This is certainly in order to bring in more tourists-both from within the country and from abroad.

REFERENCES

[1] Kodyat H., *History of tourism and its development of Indonesia* (Jakarta: Grasindo), 1998.

[2] Huraerah A., *Communication behavior of explorejogja Instagram account in Yogyakarta, tourism journey* (Yogyakarta: Respositary UMY), 2016.

[3] Drikarya, N. and Sudiarja, A. *Complete Works Drikarya* (Gramedia Pustaka Utama), 2017.

[4] Netler, M, and Mitasova, H . 2004. *Open Source GIS: A Grass GIS Approach*. Boston: Kluwer Academic Publisher, 2004.

[5] Davies and Beynon P., *Database Systems Third Edition* (Palgrave Macmillan: New York), 2004.

[6] Pramudiono, *Introduction to Data Mining: Mining Gems of Knowledge in Mount Data*, 2007.

[7] Dueck, D., *Affinity Propagation: Clustering data by passing messages*. University of Toronto, 2009.

[8] Wasserman S, and Faust K., *Social Network Analysis: Methods and Application* (New York: Cambridge University Press), 1994.