Expert Systems with Applications 161 (2020) 113703

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Inference on historical factions based on multi-layered network of historical figures

Myungjun Kim [a], Dong-gi Lee [a], Sangkuk Lee [b], Geun-ho Lee [c], Hyunjung Shin [a,*]

[a] Department of Industrial Engineering, Ajou University, Suwon 16499, South Korea
[b] Department of History, Ajou University, Suwon 16499, South Korea
[c] Department of Korean History, Chungnam University, Daejeon, 34134, South Korea

ARTICLE INFO

Article history:
Received 22 July 2019
Revised 24 May 2020
Accepted 26 June 2020
Available online 7 July 2020

Keywords:
Historical big data analysis
Historical faction identification
Multi-layered network of historical figures
Semi-supervised learning

ABSTRACT

With immense influx of historical data, quantitative inferences on history based on machine learning is becoming more prevalent, attracting many researchers. In particular, understanding the dynamics of historical factions is important as they shared academic beliefs, political views and interests, in which the interactions between the factions portray general political, social, and economic structure of a certain era. In recent years, studying such dynamics through network-based methods on human networks, constructed from genealogy data, have shown promising results. In this paper, we enhance the identification of historical factions by exploiting multi-layered network of historical figures. To understand the mechanisms of historical factions, it is pivotal to comprehend the change in relation between important historical events. The proposed method consists of constructing a multi-layered network of historical figures and applying semi-supervised learning framework to identify historical factions. The proposed method was applied to the classification of factions in the political turmoil occurred during the 15th to 16th century Korea.

1. Introduction

Many years of huge efforts from historians to create database for historical figures and events have led to easier access to a tremendous pile of records of the old days (Manabe, 1999, 2010; Lee, 2010, 2016; Lee, Lee, Kim, & Shin, 2018). The databases for historical data now provide convenient access to genealogy records, texts from literature, images of artifacts and so on. Following from the rapid growth in influx of historical data, many studies attempt to stretch from human-labored restricted inference to machine-aided comprehensive approaches. Machine learning plays a key role in the latter approach, providing historians with data-driven inference. To give few examples, Malmi, Gionis, and Solin (2018) develops automated methods, using Naïve Bayes approaches, for inferring large-scale genealogical networks. In the domain of art history, deep convolution neural networks are often employed to classify style of fine-art collections (Bar, Levy, & Wolf, 2014; Cetinic, Lipic, & Grgic, 2018; Saleh & Elgammal, 2015). For textual data, Lansdall-Welfare et al. (2017) employs text mining techniques on 150 years of British newspapers from 1800–1950 to extract macroscopic trends in history and culture. In Rochat (2015), network analysis on character network from historical writing is employed. Furthermore, many visualization efforts (Liu, Dai, Wang, Zhou, & Qu, 2017; Lee, Campbell, & Chen, 2010) for historical data are carried out to aid in understanding the structure and contents of history.

From the vast domain of using machine learning for historical big data, studies on past faction politics may convey important inferences on political, social, and economic structure of a certain era. Throughout history, people who shared academic beliefs, political views and interests joined to form factions to pursue a particular aim or purpose (Belloni & Beller, 1976). Analysis on the prevalence of factions and their rivalries is crucial for understanding political decision patterns and power mechanisms of times long past. Furthermore, the political competition of factions may be viewed as competition for goods and services (Persico, Pueblita, & Silverman, 2011), which may yield important inferences in economic perspectives.

In our earlier work of Lee et al. (2018), we developed a systematic framework for classifying historical figures into rival power groups by utilizing graph-based semi-supervised learning (SSL) (KKim & Shinim & Shin, 2013; Zhu, 2006; Zhou, Bousquet, Lal, Weston, & Schölkopf, 2004). The classification was based on a

* Corresponding author.
E-mail addresses: junkim@ajou.ac.kr (M. Kim), ldg1226@ajou.ac.kr (D.-g. Lee), okllsskh@ajou.ac.kr (S. Lee), ekh1117@cnu.ac.kr (G.-h. Lee), shin@ajou.ac.kr (H. Shin).

https://doi.org/10.1016/j.eswa.2020.113703
0957-4174/

network constructed from genealogy data with power groups defined from historical records on political decisions. In general, SSL attempts to incorporate both labeled and unlabeled data aimed at improving performance in situations of limited amount of labeled data compared to that of the unlabeled. For instance, in speech recognition, spam filtering, video categorization, webpage classification, and protein function prediction, obtaining data points is cheap but labeling requires abundant time and labor (Chapelle, Schölkopf, & Zien, 2006). Over the years, many SSL algorithms have been proposed and few of the examples include aforementioned graph-based learning, Transductive Support Vector Machines (TSVMs) (Bennett & Demiriz, 1999; Joachims, 1999; Vapnik, 2013), Laplacian SVM (Belkin, Niyogi, & Sindhwani, 2006), generative models (Kingma, Mohamed, Rezende, & Welling, 2014), etc. For dealing with historical data, SSL is appropriate with large portion of data remaining in the yet unknown region. In Lee et al., 2018, graph-based approach was suitable with limited number of features for historical figures that can be extracted from genealogy data.

In this paper, we extend classification on a single network to multi-layered networks of historical figures by exploiting historical events in chronological order. To grasp the dynamics of past factions it is not only crucial to scrutinize the relation of people at the time of a certain event but also to understand the change of the relation between two successive events in history. In a word, multiple networks of historical figures are arranged in chronological order of events, and are layered one after the other. To the multi-layered structure of network, it is limited to apply existing machine learning algorithms which are designated for a single network. The layered structure of networks accompanies computational difficulties incurred by the huge size of matrix representing contemporary or chronological relations of people. In Kim, Lee, and Shin (2019), the authors exploited how to treat the difficulties using traits of sparseness of the matrix with graph-based SSL as a base algorithm. Using the results in Kim et al. (2019), we enhance the identification of historical factions on human networks. SSL is well fit to our problem since political propensity or faction is only known for very few people, although we know a considerable number of people are involved in a certain historical event (Lee et al., 2018). Also, the algorithm enable us to observe how the political propensity of a historical figure had been influenced from his ancestors and conveyed to his off-springs or his followers later. The proposed method was applied to political purges during the 15th to 16th century Korea, in which faction politics played a critical role in shaping political system of Joseon Dynasty.

The rest of the paper is organized as follows. In Section 2, we briefly explain some background of this study. In Section 3 we outline the proposed methods constructing multi-layered network of historical figures and applying it to identify the historical factions. In Section 4, experiments and results are presented along with enrichment study. We conclude in Section 5.

2. Background

2.1. Political Purges of 1457, 1498, 1519, and 1545 in Joseon Dynasty

During 15th to 16th century, faction politics played a pivotal role in shaping political system that pervaded for the next four centuries in Joseon Dynasty. The prevalence of factions and their political competition allowed longevity of political rivalries. To understand the past political decision patterns or powers mechanisms of Joseon Dynasty up to 1910, it is crucial to analyze the factional strife that existed throughout medieval Joseon (Skillend, 1976). The political struggle led to political purges with persecutions such as sentence to death, exile, or dismissal from official position. The punishments were not limited to those directly involved in the incident, and included people of family relations, and academic or social ties. The descriptions political purges occurred in 1457, 1498, 1519, and 1545 are described as follows:

Sayukshin Incident (1457): After king Sejo forcefully took over the reign in 1453 through a coup, a group of officials conspired to restore back the previous throne of king Danjong. The plot was disclosed to king Sejo, which led to the political purge of 1457.

MuoSahwa (1498): MuoSahwa is conceived as a result of political strife between meritorious elites and neo-Confucians. During the reign of king Sungjong, the neo-Confucians began to increase their political power against meritorious elites. However, in the process of compiling records related to king Sungjong for the Annals of Joseon Dynasty, Kim Il-son, a neo-Confucian scholar, included critics of king Sejo's usurpation, which was disclosed to Yeonsangun by Lee Guk-don, a meritorious elite scholar. As Yeonsangun was the heir of king Sejo, many neo-Confucians were persecuted, resulting in the political purge of 1498.

GimyoSahwa (1519): Despite the suppressions, neo-Confucians took important positions in the national bureau in the early reign of king Joongjong. As a key person for reformation, a neo-Confucian, Jo Gwang-jo acted to remove conspicuous merits earned mostly by meritorious elites. To turn the table, meritorious elites appealed to king Joongjong questioning the authority of Jo Gwang-jo, which led to the political purge of 1519.

EulsaSahwa (1545): EulsaSahwa was largely a result of power struggle between relatives of competing princes. During the reign of king Joonjong, he had two major queens, queen Janggyung and Munjeong. After the reign of king Joongjong, the son of queen Janggyung, king Injong, became the king. As a result, Yoon Im, an elder brother of queen Janggyung, held political power. Shortly after, however, king Injong died which led to the reign of king Myungjong, the son of queen Munjeong. This led to the rise of Yoon Won–hyung, a brother of queen Munjeong. The political struggles between two factions supporting Yoon Im or Yoon Won–hyung eventually led to the political purge of 1545, resulting in persecution of the former faction.

2.2. Data

In this paper, we use two sources of historical literature to construct a multi-layered of historical figures for the four political purges, and to label historical figures to corresponding factions. For network construction, the genealogy, Munhwa Ryu-ssi Gajungbo, is used to extract list of historical figures in corresponding purges and to calculate the degree of closeness between people based on family relations. The use of genealogy is based on the idea that if one's father and grandfather belonged to a particular faction, it is highly likely for the person to belong to the same group (Lee et al., 2018). For label of factions, the Annals of Joseon Dynasty, which records events throughout 472 years of Joseon Dynasty, was used. Fig. 1 shows scanned image and screenshot image of each literature. The description of each data source is described as follows:

Munhwa Ryu-ssi Gajungbo (MRG) (Park & Lee, 2008): MRG is a genealogy record published in 1565 containing names and family relations of 49,635 people throughout 19 generations. As in Lee (2013), the family relations are digitized based on the Henry numbering system (Henry, 1935). MRG also includes, birth year, death year, year of taking position at national bureau, and number of years in incumbency.

The Annals of Joseon Dynasty (AJD,http://sillok.history.go.kr): Registered as a UNESCO World Heritage in 1997, AJD is a historical volume that describes historical facts over 25 reigns, with total of 472 years, from king Taejo (1392) to king Chuljong (1863).

(a) Scanned image of MRG (b) Screenshot of an article in AJD

Fig. 1. Screenshots of MRG and AJD.

3. Proposed method

3.1. Representation of a multi-layered network of historical figures

In general, a network of historical figures, $G(V,E)$ consists of nodes (V) and edges (E), where nodes represent historical figures of a contemporary period with $V = \{x_1, x_2, \cdots, x_N\}$ for N people, and edges represent similarities between the people. The similarities are defined by the weight matrix, W, where elements, W_{ij}, represent the closeness between person x_i and x_j. To exploit historical events in chronological order, the total dataset can be represented by a multi-layered network where each layer is a network of contemporary historical figures and the layers are ordered chronologically with respect to important historical events.

Suppose there are K number of layers composed in multi-layered structure. We denote that network with $G(V,E,S)$, where in addition to nodes and edges S is the set of K layers, $S = \{S_1, S_2, \cdots, S_K\}$. With respect to the K distinct layers, n_k is the size of each layer with $N = n_1 + n_2 + \cdots + n_K$ in total. The resulting weight matrix W is a $N \times N$ block tri-diagonal matrix, where the diagonal blocks represent individual intralayer edges, and banded diagonal blocks represent interlayer edges between adjacent layers. In historical perspective, intralayer edges represent relations between contemporary people and interlayer edges represent relations between people from two different, but adjacent, historical events. The weight matrix contains $K^2 - 3K + 2$ zero blocks resulting in a sparse structure. Fig. 2 represents an exemplary multi-layered network of historical figures and structure of its corresponding weight matrix.

3.2. Network construction

The intuition behind constructing a multi-layered network is that knowing the factions in previous generation or era may aid in prediction process of later events, and vice versa. There are two steps in constructing a network of historical figures: extracting list of people in corresponding events and calculating similarity between them.

For the first step, the year of each purge is set as an orient and list of people of contemporary is selected based on a time-window of the orient. For the time-window, personal information in genealogy is used. To include contemporary people in the network, they must satisfy one of the following conditions:

$$(\text{YP-ALE}) \leqslant \text{Year of Birth} \leqslant (\text{YP-ATP}) \tag{1}$$

$$(\text{YP-INC}) \leqslant \text{Year of Taking Position} \leqslant (\text{YP-1}) \tag{2}$$

where YP is the year of purge, ALE is the average life expectancy, ATP is the mean age of taking position, and INC is the mean incumbency. Here, condition (1) states that the person must be alive and condition (2) points out that the person must be in an official position at the year of purge. Since there are many missing values in the genealogy, condition (1) or (2) must be satisfied for a person to be included in the network.

To calculate similarities between people, the degree of kinship, k_{ij}, derived from genealogy is used. It is defined as the level of the relationship between two persons related by blood, calculated as one degree for each step from a common ancestor. For instance, first-degree of kinship denote parent-offspring relations, siblings as second-degree of kinship, uncle or aunt as third-degree of kinship, and so on. A smaller value of k_{ij} implies closer blood relations between person x_i and x_j, and thus its value can be viewed as a measure of distance between two people.

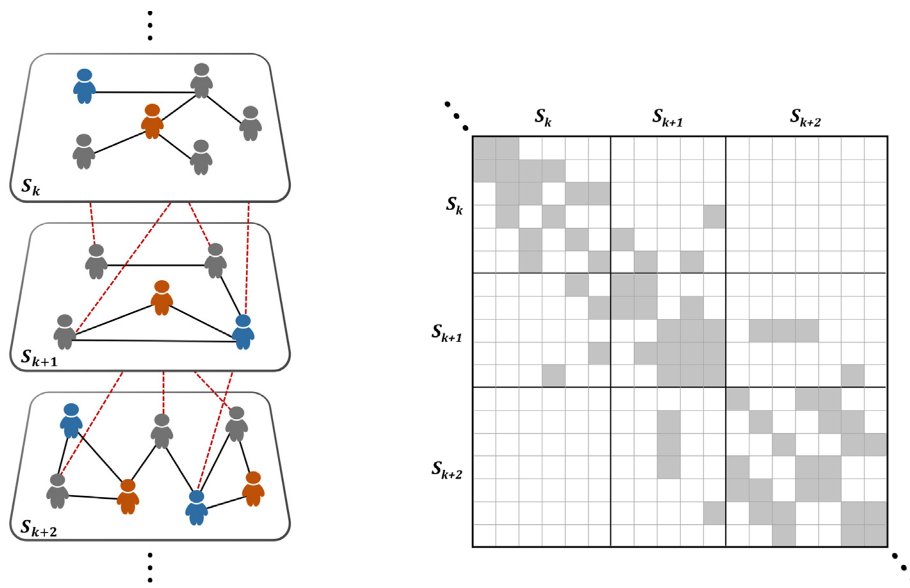
Using the degree of kinship, the weight matrix is given by:

$$W_{ij} = \begin{cases} \frac{2}{1+e^{k_{ij}}} \cdot \delta_{ij} & \text{if } k_{ij} < 9 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where k_{ij} is the degree of kinship between person x_i and x_j, and δ_{ij} is a binary indicator with value 1 if person x_i and x_j belong in same or adjacent event (layer), or otherwise 0. The edge has higher values for close relatives and lower values for far relatives and we ignore relations with $k_{ij} \geqslant 9$.

3.3. Identification of Historical Factions

Before identifying historical factions, it is necessary to define labels which represent specific faction groups for people. In this paper, we make use of the fact that Joseon Dynasty was society ruled by a bipartisan system with the King as the decision maker. For historical purges of Joseon Dynasty, the factions can be represented by victims and assailants. The victims (labeled with '-1') are defined as historical figures who had been sentenced of death,



(a) (b)

Fig. 2. (a) An exemplary multi-layered network of historical figures and the layers are ordered in chronological order. (b) The corresponding structure of the weight matrix of exemplary network in (a). The block tri-diagonal matrix induces a sparse structure that can take benefit of computational advantage.

punished with exile, or dismissal from office during the purges. On the other hand, the assailants (labeled with '+1') are defined as those who incited and supported the persecution of victims. The labels were obtained by searching over the AJD and were verified through results in previous researches regarding Korean history.

Given a network of historical figures and associated labels, we apply graph-based SSL for multi-layered networks (Kim et al., 2019) to identify people with their associated factions. SSL utilize both labeled and unlabeled data where it is known to be effective for large amount of unlabeled data (KKim & Shinim & Shin, 2013). In identifying historical factions, SSL is appropriate with large portion of people are unknown for their factions. Furthermore, we take the benefit of sparseness in the weight matrix arising from huge-sized historical data using multi-layered SSL.

Suppose there are two factions of interest where n_l people belong to a specific faction and n_u people with unknown faction. In such setting, we can assign $Y_l \in \{-1, 1\}$ for n_l people whereas $Y_u = 0$ for unlabeled ones. For graph-based SSL, known labels are propagated to other through the edges. Mathematically, this translates to determining the output vector $f = (f_1, f_2, \cdots, f_n)^T$ by minimizing the following objective function:

$$\min_f (f - Y)^T (f - Y) + \mu f^T L f \tag{4}$$

where L is the graph Laplacian (Chung & Graham, 1997) defined as $D - W$ with $D = \mathrm{diag}(d_i)$ and $d_i = \sum_j w_{ij}$. The first term in (4) is the loss for consistency with actual labels for labeled nodes, and the second term is the smoothness for consistency with geometry of the data. The parameter μ is for trade-off between the loss and smoothness terms. The solution of (4) is given by

$$f = (I + \mu L)^{-1} Y \tag{5}$$

where I is the identity matrix.

To apply SSL for multi-layered networks, we first should account for two different label propagations, interlayer and intralayer label propagation. To control for different propagations, it is desirable to separate the total weight matrix W into intralayer and interlayer relations:

$$W = W^{\{intra\}} + W^{\{inter\}}$$

Now, parallel to (4), the objective function for a multi-layered network is defined as

$$\min_f (f - Y)^T (f - Y) + f^T (\mu_a L^{\{intra\}} + \mu_b L^{\{inter\}}) f \tag{6}$$

where $L^{\{intra\}}$ and $L^{\{inter\}}$ are the graph Laplacians of respective weight matrices. The parameters $\mu_a (\geqslant 0)$ and $\mu_b (\geqslant 0)$ are *smoothness-loss* trade-off parameters for intralayer and interlayer relations, respectively. Similar to (5), the solution is obtained as a closed form,

$$f = (I + \mu_a L^{\{intra\}} + \mu_b L^{\{inter\}})^{-1} Y \tag{7}$$

When $\mu_b = 0$, it reduces to (5).

Although solution (7) is sufficient to obtain the predictive outputs, the computation is difficult due to large volume of data involved in matrix inversion. To overcome such problem, it is possible to make use the sparse structure of the weight matrix by using the Woodbury formula (Woodbury, 1950):

$$(A + UBU^T)^{-1} = A^{-1} - A^{-1} U (B^{-1} + U^T A^{-1} U)^{-1} U^T A^{-1}$$

where A is a $n \times n$ invertible matrix, B is a $m \times m$ invertible matrix, and U is a $n \times m$ rectangular matrix. Woodbury formula is known to be effective for dealing with sparse matrices (Hager, 1989). To rephrase the right-side term in (7) using the Woodbury formula, let $A = I + \mu_a L^{\{intra\}}, B$ be the identity matrix I. And let $L^{\{inter\}}$ be decomposed to UU^T where U is the incidence matrix for the Laplacian scaled by μ_b. U is a $n \times p$ matrix where p is the number of edges. By introducing the incident matrix, it becomes possible to make use of the sparseness in the weight matrix. The solution (7) can be rephrased as

$$f = A^{-1} Y - A^{-1} U (I + U^T A^{-1} U) U^T A^{-1} Y \tag{8}$$

4. Results and discussion

4.1. Results of network construction

A multi-layered network was constructed based on four political purges, Sayukshin Incident (1456), MuoSahwa (1498), GimyoSahwa (1519), and EulsaSahwa (1545), during 15th to 16th century Korea. The network was composed of four layers arranged in chronological order. 9,481 people were extracted from MRG with 1,353, 2,240, 2,868, and 3,020 people belonging to layers of Sayukshin Incident, MuoSahwa, GimyoSahwa, and EulsaSahwa, respectively. After connecting the edges, the resulting network density was 0.82%, which is a good indicator for using sparseness in our proposed method. Fig. 3 shows the heatmap of the constructed network. On the heatmap, red diagonal blocks represent intralayer relations and blue banded diagonal blocks represent interlayer relations.

The labels were obtained by searching over the AJD. Total 590 people with 431 victims and 159 assailants were obtained with 22/42, 37/111, 49/151, and 51/127 assailants/victims for Sayukshin Incident, MuoSahwa, GimyoSahwa, and EulsaSahwa, respectively. The labels were validated via examining existing records on Korean history with the aid of Korean historians (The Annals of Joseon Dynasty, 2007; Encyclopedia of Korean Culture, 2012). As a total, this corresponds to 6.22% of historical figures having known labels, which bolsters the necessity of using SSL framework. Table 1 summarizes the overall specifications of the constructed network.

4.1.1. Comparison with single networks

For each of the purges, $p = 5\%, 10\%, 20\%, 40\%, 60\%, 80\%, 100\%$ of labels were assigned to three other layers. For instance, if we were to measure performance on EulsaSahwa layer, labels on other three layers (SayukShin Incident, MuoSahwa, and GimyoSahwa), were varied according to the percentages. Then, to induce SSL setting of $n_l > n_u$, reverse 5-fold cross validation (1-fold as training set and others as test set) AUC (Swets, 2014) was calculated as a performance measurement. The combinations of (μ_a, μ_b) were determined in the range of $\{0.01, 0.1, 1, 10, 100\} \times \{0.01, 0.1, 1, 10, 100\}$. For comparison the performance of ordinary graph-based SSL on single networks were also measured. Experiments were repeated 100 times.

The overall AUC values for each of the purges are shown in Fig. 4. On the figure, ML$(\cdot)$ denote multi-layered networks with $(\cdot)$ percentage of labels on other layers. For single networks, the AUC values are not too much distant from simple random guessing (AUC $= 0.5$). On the other hand, the multi-layered network has an average of 50.20% increase in performance when $p = 100\%$ with maximum increase of 64.09% in the case of GimyoSahwa.

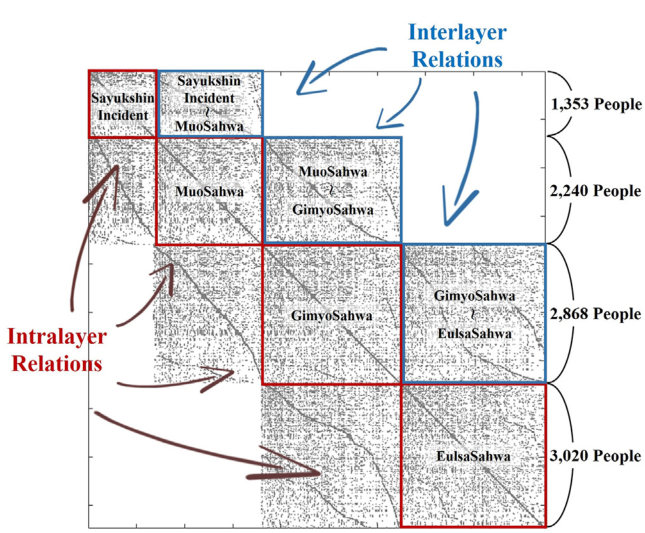


Fig. 3. Heatmap of the constructed network of 9,481 people. Red diagonal blocks represent intralayer relations and blue banded diagonal blocks represent interlayer relations.

Table 1
Summary of the constructed network.

Specification	Value
Historical Figures (Node)	9,481 people
Relations (Edge)	366,539 connections
Factions (Label)	590 ('+1': 159/ '−1': 431) people
Network Density	0.82%
Proportion of Labels	6.22%

Fig. 5 presents the individual AUC comparison when $p = 100\%$. Results of 30 experiments are presented for easier visualization. A point in the scatter plot located above the diagonal line means that the algorithm on the vertical axis performs better. Results on each of the purges are marked with different shapes and colors. From the figure, it can be seen that most of the dots are above the diagonal line, which corresponds to the multi-layered network outperforming single networks in all cases. Additionally, pairwise comparison for each of the purges showed a statistically significant difference (all the p-values were less than 0.0001) between single and multi-layered networks.

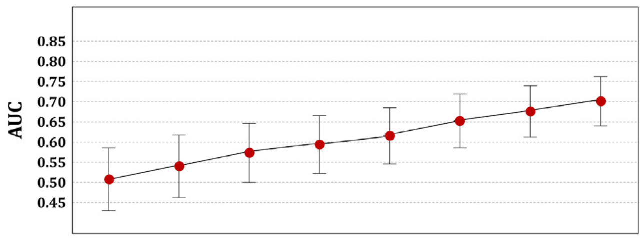
Overall, the results show that the multi-layered network outperforms single networks for all four of the purges. This bolsters that analyzing factions through historical events in chronological order is better than simply examining the single events separately.

To classify the remainder of unlabeled people, we assigned all of 590 labeled people with associated labels and applied the proposed method. As a result, 2,593 people were classified as assailants with 465, 560, 703, and 865 people for Sayukshin Incident, MuoSahwa, GimyoSahwa, and EulsaSahwa, respectively. For victims, 6,888 people were identified with 888, 1,680, 2,165, and 2,155 people in respective layers. Fig. 6 shows a subset network with 100 people in each of the events. Blue and red dots represent people classified as assailants and victims, respectively. The ratio of the groups for each purge is reflected in the figure. The names are shown in Korean.

4.1.2. Comparison with other classification algorithms

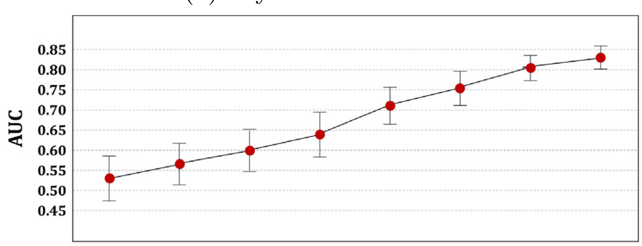
The proposed method ($p = 100\%$) was compared with other classification algorithms: Artificial Neural Network (ANN) (Bishop et al., 1995), two Support Vector Machines (SVM) with linear and Gaussian kernels (Cristianini et al., 2000; Kotsiantis, Zaharakis, & Pintelas, 2006; Schölkopf, Smola, & Bach, 2002; Shin & Cho, 2007), and K-Nearest Neighbor (KNN) (Cover & Hart, 1967; Dudani, 1976). To extract the features, binarization of the Henry numbering system in the genealogy was performed to transform the tree into binary vectors (Lee et al., 2018). As a result, 134 dimensional binary vector were produced for 590 people with known labels. The train/test split was set similar to that of multi-layered networks with $p = 100\%$. For each of the purges, all of the data in other respective layers were assigned in the training set and additional reverse 5-fold cross validation on the layer where the performance was measured. For parameter setting, six hidden layers were set in ANN with 100, 50, 30, 10, 5, and 3 hidden nodes in respective layers. For SVMs, penalizing parameters were set to 10 and 1 for linear and Gaussian SVM, respectively, and the kernel width for Gaussian was set to 25. These values were obtained through cross-validation. For KNN, cosine distance was used as the distance metric with $K = 10$. The experiments were repeated 100 times and the performances were measured with AUC and accuracy values.

Table 2 shows the comparison of average AUC and accuracy for each of the purges. Examining the table, the proposed method achieves the highest performance in all four cases. The difference



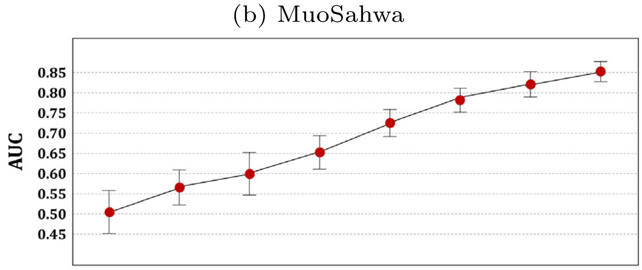
	Single Network	ML (5%)	ML (10%)	ML (20%)	ML (40%)	ML (60%)	ML (80%)	ML (100%)
Avg.	0.5136	0.5465	0.5708	0.5905	0.6251	0.6579	0.6803	0.6980
Std.	0.0744	0.0725	0.0714	0.0670	0.0606	0.0516	0.0449	0.0407

(a) Sayukshin Incident



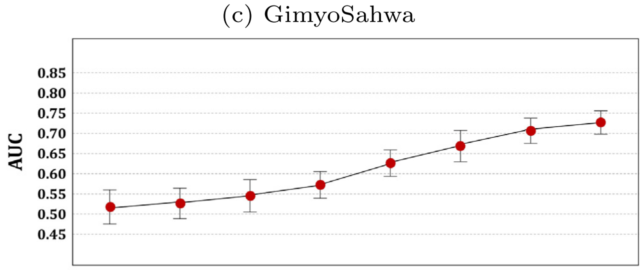
	Single Network	ML (5%)	ML (10%)	ML (20%)	ML (40%)	ML (60%)	ML (80%)	ML (100%)
Avg.	0.5395	0.5705	0.5990	0.6429	0.7125	0.7662	0.8071	0.8357
Std.	0.0478	0.0539	0.0520	0.0521	0.0452	0.0363	0.0295	0.0225

(b) MuoSahwa



	Single Network	ML (5%)	ML (10%)	ML (20%)	ML (40%)	ML (60%)	ML (80%)	ML (100%)
Avg.	0.5118	0.5610	0.5986	0.6552	0.7315	0.7861	0.8268	0.8566
Std.	0.0436	0.0398	0.0395	0.0360	0.0306	0.0274	0.0234	0.0177

(c) GimyoSahwa



	Single Network	ML (5%)	ML (10%)	ML (20%)	ML (40%)	ML (60%)	ML (80%)	ML (100%)
Avg.	0.5220	0.5225	0.5409	0.5791	0.6319	0.6736	0.7068	0.7321
Std.	0.0410	0.0438	0.0411	0.0402	0.0353	0.0320	0.0273	0.0246

(d) EulsaSahwa

Fig. 4. Overall AUC values for four purges, Sayukshin Incident, MuoSahwa, GimyoSahwa, and EulsaSahwa.

may come from the fact that the proposed method takes family relations, namely degree of kinship, into account, unlike other comparing algorithms with simple binarization of Henry numbering system. Thus, the results can be translated that the proposed method incorporates blood-relations from genealogy better than other classification algorithms.

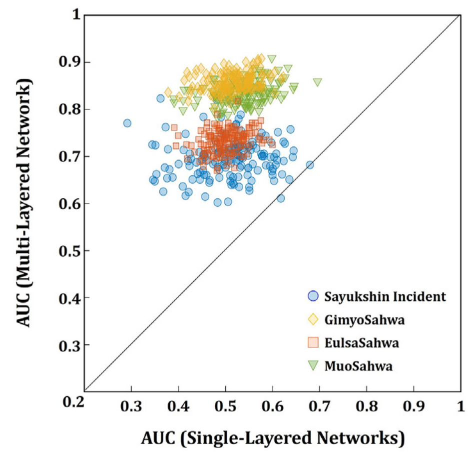


Fig. 5. Individual AUC comparison for 30 experiments when $p = 100\%$. Results on each of the purges are marked with different shapes and colors. Dots above diagonal line indicate higher AUC on the y-axis. The figure shows a better performance of the proposed methods compared to single networks.

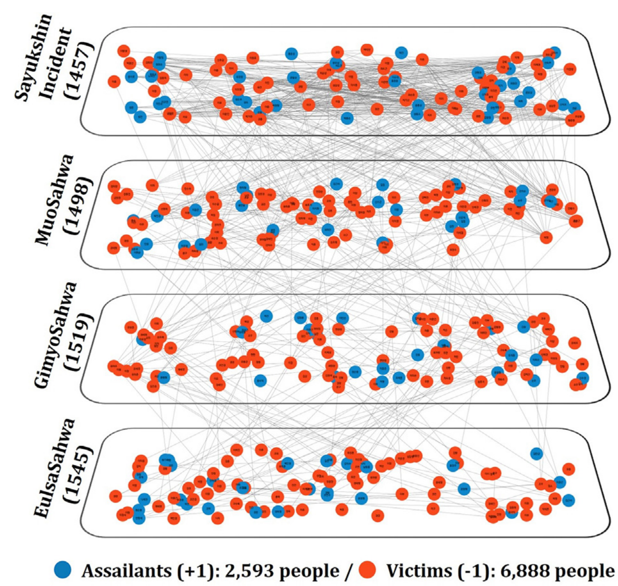


Fig. 6. A subset network with 100 people, for each of the purge, classified to a assailants or victims. The ratio of the groups for each purge is reflected.

4.2. Enrichment study

As an enrichment study, we applied the proposed method to a situation where $p = 100\%$ while erasing all the labels on the layer of interest. In such setting, a single network cannot make any predictions. In contrast, the proposed method makes use of interlayer connections between layers where labels can be propagated from one layer to another, making predictions possible.

The resulting outputs, f_S, f_M, f_G, and f_E, for Sayukshin Incident, MuoSahwa, GimyoSahwa, and EulsaSahwa, are presented in Fig. 7. For each event, the outputs had been sorted in descending order and stretched by a sigmoid function for the ease of visualization and interpretation. The grey line represent unlabeled people and the blue and red dots represents people who were originally

Table 2
Comparison of average AUC values.

Classification Algorithms	AUC	Accuracy
Proposed Method ($p = 100\%$)	**0.6980**	**0.7536**
ANN	0.5451	0.6521
Linear SVM	0.5901	0.6954
Gaussian SVM	0.6281	0.7010
KNN ($K = 10$)	0.6397	0.7276

(a) Sayukshin Incident

Classification Algorithms	AUC	Accuracy
Proposed Method ($p = 100\%$)	**0.8357**	**0.8264**
ANN	0.5467	0.7211
Linear SVM	0.6013	0.7564
Gaussian SVM	0.6494	0.7614
KNN ($K = 10$)	0.5623	0.7194

(b) MuoSahwa

Classification Algorithms	AUC	Accuracy
Proposed Method ($p = 100\%$)	**0.8566**	**0.8400**
ANN	0.5491	0.7237
Linear SVM	0.6836	0.7856
Gaussian SVM	0.7818	0.8202
KNN ($K = 10$)	0.5716	0.7575

(c) GimyoSahwa

Classification Algorithms	AUC	Accuracy
Proposed Method ($p = 100\%$)	**0.7321**	**0.7590**
ANN	0.5421	0.6948
Linear SVM	0.6120	0.7162
Gaussian SVM	0.6362	0.7293
KNN ($K = 10$)	0.5386	0.7129

(d) EulsaSahwa

labeled as assailants ('+1') and victims ('-1'), respectively. The black line represents the threshold for splitting the groups into two where it is given by the ratio of classes for labeled data. The thresholds were used to compute the confusion matrix. The figure shows that most of the blue and red dots are placed on right and left of the threshold, respectively, giving fairly high accuracy for the four purges.

In regards to the utility of the enrichment study, the results can serve as a quantitative baseline for historians. The output values can represent relative likelihood (Nam, Kim, Lee, & Shin, 2016) of belonging to a particular faction: higher of its value implies higher likelihood of belonging to the class of assailants ('+1') and lower its value implies higher likelihood for the opposing group, victims ('-1). Given the tremendous volume of historical literature, if one was to search historical figures for a particular faction, it would be highly efficient to reduce the search space by observing those with very high or low output values through our proposed method.

As an example of Sayukshin Incident, in the top-50 of f_S, Kang Hee-meng was found to assist king Sejo's usurpation and served in important positions during his reign (Kwon, 1996). Also, Lee Jeung, a member of the royal family, supported king Sejo's coup to overthrow his step-brother, king Danjong (Lee, 1996a). Likewise, in the bottom-50, Nam Hyo-on was recorded to be one of the minister who was not persecuted but retired from the official position to keep his incision on king Sejo's wrongdoing (Lee, 1996b). Although these people were not recorded to be assailants in the Sayukshin Incident, such finding suggest their high associations to each of the groups.

The results of the proposed method allow us to observe the shift in political propensity of historical figures with respect to the influence from ancestors and how it is conveyed to off-springs or followers in later generations. In such manner, it provides a systematic understanding of the structure of past faction politics over long periods of time. This can essentially aid domain historians to grasp the holistic understanding of historical landscape.

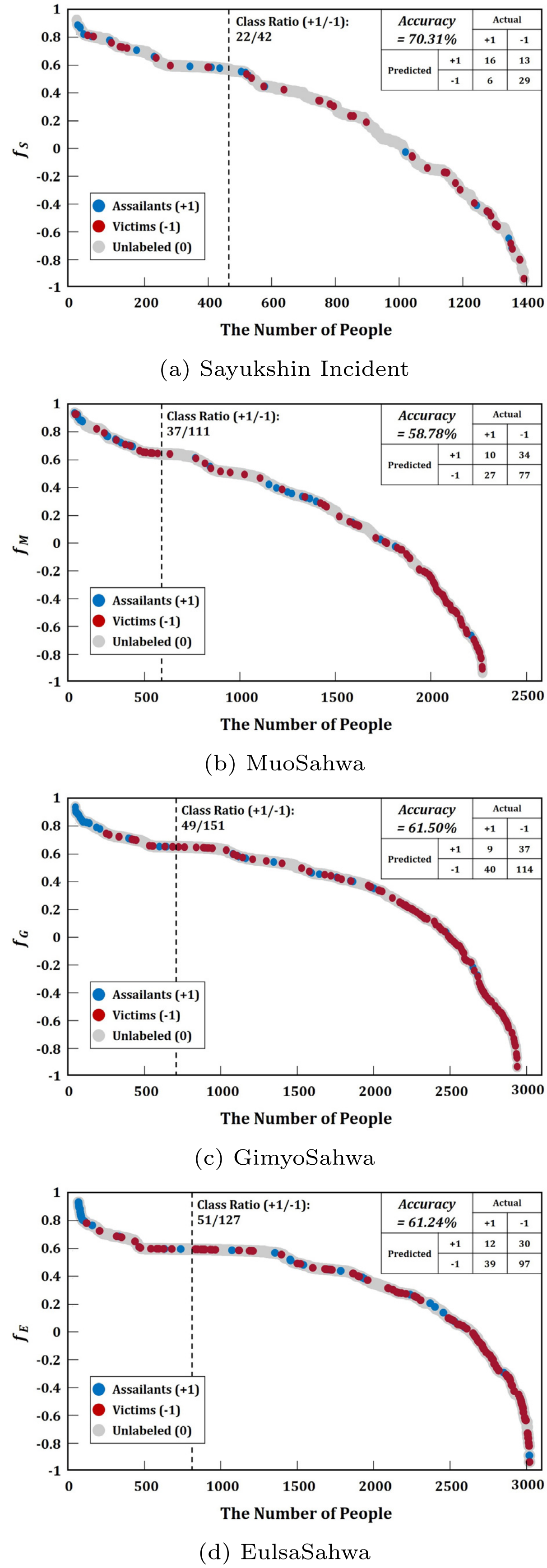


Fig. 7. Predictive values sorted in descending order for each of the purges after erasing all the labels on the layer of interest.

5. Conclusion

In this paper, we employ multi-layered network of historical figures by conveying important historical events in chronological order. To identify the factions, a semi-supervised learning algorithm for multi-layered networks was developed. The proposed method was applied to four purges, Sayukshin Incident (1456), MuoSahwa (1498), GimyoSahwa (1519), and EulsaSahwa (1545), in Joseon Dynasty, where a multi-layered network was constructed based on personal information and degree of kinship found in genealogy. The results showed 50.20% increase in performance compared to single networks. Furthermore, the proposed method outperformed other representative machine learning algorithms, conveying that the former incorporates information in genealogy better than the latter.

Contribution of this research lies in that the proposed method carries the notion that factions in certain era may had been affected by the past events. By taking consideration of interlayer connections beyond intralayer connections, the proposed method analyzes various historical events in a comprehensive manner, where it accounts relations of people not only at the time of a certain event but also between two successive events in history. In such setting, label information in one layer is propagated to others, thereby providing better performances. Macroscopically, this study is considered as a novel work of historical study. Currently, machine learning has enlarged its application area stretching from literature, art, and digital humanities and yet there are much to be explored in the domain of history. The proposed method provides new schemes of analyzing historical methods far from the conventional applications of machine learning, thereby attracting historians.

For our future work, we plan to explore various problem settings where the proposed method can be applied. The proposed algorithm can easily be expanded to other settings. For instance, networks may be constructed based on regional relations in residency information or political relations based on opinions from various historical literature. Such extensions are not necessarily confined to Korean history and can also be planted to any era in any countries. Furthermore, the proposed method can be utilized to current politics, in which political propensities on numerous agenda can be identified. We expect the proposed method to extend human-labored restricted inference to machine-aided comprehensive approach and serve as a sophisticated tool to understand overall structure of politics, society, and the economy.

CRediT authorship contribution statement

Myungjun Kim: Conceptualization, Methodology, Data curation, Formal analysis, Validation, Visualization, Writing - original draft, Writing - review & editing. **Dong-gi Lee:** Conceptualization, Methodology, Data curation, Writing - review & editing. **Sangkuk Lee:** Conceptualization, Data curation, Resources, Validation. **Geun-ho Lee:** Resources, Validation. **Hyunjung Shin:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2018S1A5B6075104/2018R1D1A1B07043524), and the Ajou University research fund.

References

Bar, Y., Levy, N., & Wolf, L. (2014). Classification of artistic styles using binarized features derived from a deep neural network. In *Workshop at the European Conference on Computer Vision* (pp. 71–84). Springer.
Belkin, M., Niyogi, P., & Sindhwani, V. (2006). Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of Machine Learning Research, 7*, 2399–2434.
Belloni, F. P., & Beller, D. C. (1976). The study of party factions as competitive political organizations. *Western Political Quarterly, 29*, 531–549.
Bennett, K.P., & Demiriz, A. (1999). Semi-supervised support vector machines. In Advances in Neural Information Processing Systems (pp. 368–374).
Bishop, C. M. et al. (1995). *Neural networks for pattern recognition*. Oxford University Press.
Cetinic, E., Lipic, T., & Grgic, S. (2018). Fine-tuning convolutional neural networks for fine art classification. *Expert Systems with Applications, 114*, 107–118.
Chapelle, O., Schölkopf, B., & Zien, A. (2006). Semi-supervised learning. Cambridge, Massachusettes: The MIT Press.
Chung, F. R., & Graham, F. C. (1997). *Spectral graph theory. 92*. American Mathematical Soc.
Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, 13*, 21–27.
Cristianini, N., Shawe-Taylor, J., et al. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press.
Dudani, S.A. (1976). The distance-weighted k-nearest-neighbor rule. IEEE Transactions on Systems, Man, and Cybernetics, (pp. 325–327).
(2012). Encyclopedia of korean culture (in korean). URL:https://encykorea.aks.ac.kr.
Hager, W. W. (1989). Updating the inverse of a matrix. *SIAM Review, 31*, 221–239.
Henry, R. B. (1935). *Genealogies of the families of the presidents*. Tuttle Company.
Joachims, T. (1999). Transductive inference for text classification using support vector machines. In ICML (pp. 200–209). Vol. 99.
Kim, J., & Shin, H. (2013). Breast cancer survivability prediction using labeled, unlabeled, and pseudo-labeled patient data. *Journal of the American Medical Informatics Association, 20*, 613–618.
Kim, M., Lee, D.-G., & Shin, H. (2019). Semi-supervised learning for hierarchically structured networks. *Pattern Recognition, 95*. 191–20.
Kingma, D.P., Mohamed, S., Rezende, D.J., & Welling, M. (2014). Semi-supervised learning with deep generative models. In Advances in neural information processing systems (pp. 3581–3589).
Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2006). Machine learning: A review of classification and combining techniques. *Artificial Intelligence Review, 26*, 159–190.
Kwon, Y. (1996). Kang hee-meng (in korean). URL:http://encykorea.aks.ac.kr/Contents/Item/E0001550.
Lansdall-Welfare, T., Sudhahar, S., Thompson, J., Lewis, J., Team, F. N., Cristianini, N., Gregor, A., Low, B., Atkin-Wright, T., Dobson, M., et al. (2017). Content analysis of 150 years of british periodicals. *Proceedings of the National Academy of Sciences, 114*, E457–E465.
Lee, D.-G., Lee, S., Kim, M., & Shin, H. (2018). Historical inference based on semi-supervised learning. *Expert Systems with Applications, 106*, 121–131. https://doi.org/10.1016/j.eswa.2018.03.059.
Lee, J. (1996). Gyeyang-gun (in korean). URL:http://encykorea.aks.ac.kr/Contents/Index?contents_id=E0003197.
Lee, J. (1996). Nam hyo-on (in korean). URL:http://encykorea.aks.ac.kr/Contents/Index?contents_id=E0012316.
Lee, J., Campbell, C.D., & Chen, S. (2010). China multi-generational panel dataset liaoning (cmgpd-ln). In User guide (pp. 1749–1909). Inter-university Consortium for Political and Social Research.
Lee, S. (2010). The impacts of birth order and social status on the genealogy register in thirteenth-to fifteenth-century Korea. *Journal of Family History, 35*, 115–127.
Lee, S. (2013). The impact of family background on bureaucratic reproduction in the thirteenth-to-fifteenth century korea: A case study on the kwon-ssi sunghwabo (in korean). *Daedong Munhwa Yeon'gu, 81*, 41–67.
Lee, S. (2016). Conditions and potentials of korean history research based on 'big data' analysis: The beginning of 'digital history'. *Korean Journal of Applied Statistics, 29*, 1007–1023.
Liu, Y., Dai, S., Wang, C., Zhou, Z., & Qu, H. (2017). Genealogyvis: A system for visual analysis of multidimensional genealogical data. *IEEE Transactions on Human-Machine Systems, 47*, 873–885.
Malmi, E., Gionis, A., & Solin, A. (2018). Computationally inferred genealogical networks uncover long-term trends in assortative mating. arXiv preprint arXiv:1802.06055.
Manabe, T. (1999). The digitized kobe collection, phase i: Historical surface marine meteorological observations in the archive of the japan meteorological agency. *Bulletin of the American Meteorological Society, 80*, 2703–2716.
Manoff, M. (2010). Archive and database as metaphor: Theorizing the historical record. portal: Libraries and the. *Academy, 10*, 385–398.
Nam, Y., Kim, M., Lee, K., & Shin, H. (2016). Clash: Complementary linkage with anchoring and scoring for heterogeneous biomolecular and clinical data. *BMC Medical Informatics and Decision Making, 16*, 72.
Park, H., & Lee, S. (2008). A survey of data sources for studies of family and population in korean history. *The History of the Family, 13*, 258–267.

Persico, N., Pueblita, J. C., & Silverman, D. (2011). Factions and political competition. *Journal of Political Economy, 119*, 242–288.
Rochat, Y. (2015). Character network analysis of emile zola's les rougon-macquart. In Digital Humanities 2015 EPFL-CONF-210573.
Saleh, B., & Elgammal, A. (2015). Large-scale classification of fine-art paintings: Learning the right metric on the right feature. arXiv preprint arXiv:1505.00855.
Schölkopf, B., Smola, A. J., Bach, F., et al. (2002). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. MIT Press.
Shin, H., & Cho, S. (2007). Neighborhood property–based pattern selection for support vector machines. *Neural Computation, 19*, 816–855.
Skillend, W. (1976). Edward willett wagner: The literati purges: political conflict in early yi korea. *Bulletin of the School of Oriental and African Studies, 39*. 471–471.
Swets, J. A. (2014). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. Psychology Press.
(2007). The annals of joseon dynasty (in korean). URL:http://sillok.history.go.kr/.
Vapnik, V. (2013). *The nature of statistical learning theory*. Springer Science & Business Media.
Woodbury, M. A. (1950). Inverting modified matrices. *Memorandum Report, 42*, 336.
Zhou, D., Bousquet, O., Lal, T.N., Weston, J., & Schölkopf, B. (2004). Learning with local and global consistency. In Advances in neural information processing systems (pp. 321–328).
Zhu, X. (2006). Semi-supervised learning literature survey. *Computer Science, University of Wisconsin-Madison, 2*, 4.